Theor Chem Acc (2002) 107:90–102
DOI 10.1007/s00214-001-0303-0

Theoretical
Chemistry Accounts

Regular article

Implementation of solvent reaction fields for electronic structure

Daniel M. Chipman[1], **Michel Dupuis**[2]

[1] Radiation Laboratory, University of Notre Dame, Notre Dame, IN 46556, USA
[2] Pacific Northwest National Laboratory, EMSL/K1-83, Battelle Blvd, Richland, WA 99352, USA

Received: 2 July 2001 / Accepted: 11 September 2001 / Published online: 19 December 2001

Abstract. Methods are described to incorporate solvent reaction field effects into solute electronic structure calculations. Included are several old and new approaches based on approximate solutions of Poisson's equation through boundary element methods, wherein the solutions are represented in terms of certain apparent surface charge or apparent surface dipole distributions. Practical algorithms to set up and solve the requisite equations are described and implemented in a new general reaction field computer program. Illustrative computational results are presented to show the performance of the program.

Key words: Solvent effects – Reaction field – Boundary element methods – Surface charge distributions – Surface dipole distributions

1 Introduction

Inclusion of a reaction field is becoming very popular in electronic structure calculations as a convenient means to represent important electrostatic effects of a solvent on solute properties [1, 2]. The present work describes and documents practical algorithms implemented in a new computer program for determining the reaction field. A concise treatment of the most immediately relevant literature is given in Ref. [3], to which we also refer for most details of the theory. Through use of boundary element methods, the solutions are expressed in terms of either apparent surface charge or apparent surface dipole distributions. Details are given here on the construction and solution of the relevant equations to determine these distributions by finite matrix manipulations. The computer program described in this work to implement these procedures is a separate module interfaced with the standard electronic structure HONDO package [4]. Computational results are presented to illustrate the performance and precision of the program.

The very simple model of solvation considered here represents the solvent as a linear isotropic homogeneous continuum characterized by its static bulk dielectric constant. A cavity that strictly excludes solvent is carved out of the continuum to hold the solute. The solute polarizes the solvent to produce an equilibrium reaction field that is specified by Poisson's equation together with appropriate boundary conditions. Solution of this equation provides an electrostatic reaction potential, from which an effective solute–solvent potential-energy expression can be immediately obtained and included in the Hamiltonian operator used for electronic structure calculation on the solute. It is emphasized that this simple model only treats the long-range electrostatic part of the solute–solvent interaction. While this is often the dominant effect in polar solution, it should be recognized that many other important solute–solvent interactions also exist [1, 2] that lie outside the scope of the present work.

Unconstrained quantum mechanical calculation of solute electronic structure inevitably leads to a tail of the wave function penetrating outside the cavity that nominally encloses the solute. Exact solution of Poisson's equation to include the effects of this charge penetration requires use of an apparent volume charge density [5, 6, 7] lying outside the cavity in addition to the commonly used apparent surface charge density lying on the cavity surface. This representation of the complete reaction field solution has previously been implemented [8, 9, 10, 11] through the method here denoted SVPE, but exact treatment of the volume polarization is difficult and demanding.

The present work focuses on simpler methods that provide approximate solutions of Poisson's equation by representing the reaction field using only certain apparent surface distributions. Considered here are the previously developed SS(V)PE [7, 8, 11, 12, 13, 14, 15], IEF [16, 17, 18, 19], and COSMO [20, 21, 22, 23, 24, 25, 26, 27] methods that each provide an approximate

Correspondence to: D. M. Chipman

representation of volume polarization effects in terms of apparent surface charge distributions. For comparison, we further consider the large group of most other extant apparent surface charge methods that simply neglect volume polarization, which we collectively classify [3] under the name SPE. Also considered here are analogous new methods [3] called SS(V)PE–μ, IEF–μ, COSMO–μ, and SPE–μ that instead each use an apparent surface dipole distribution to represent the reaction field.

It is shown in Sect. 2 how the formal equations that are fully described in Ref. [3] can be discretized over a grid of points on the cavity surface to allow solution by finite matrix methods. Practical details of how the cavity surface is actually constructed and represented in our computer program are provided in Sect. 3. Details of how the reaction field program is interfaced to a standard electronic structure package [4] are given in Sect. 4. The numerical performance of the program in representative applications is considered in Sect. 5. A brief summary of the work and the major findings are reviewed in Sect. 6. Detailed derivation of a shape factor useful to improve the matrix representations of several of the methods considered is given in the Appendix. A comparison of the accuracy of these various approximate methods is provided in Ref. [3].

2 Solution of reaction field equations

In this section we discuss how the formal equations described in Ref. [3] are discretized to allow their approximate solution using matrix methods. This naturally requires certain information about the cavity surface. In particular, it is necessary to know the locations of a set of N grid points, $\{\mathbf{t}_i\}$, that implicitly represent the actual continuous surface $\mathbf{\Gamma}$, the amount of surrounding surface area, a_i, that each grid point represents, and the unit vector, $\mathbf{n}_i$, along the outward-directed normal to the surface at each grid point. In cases where the surface is not spherical about the expansion center (or a strict union of spheres about a set of expansion centers) it is also useful to know the acute angle, γ_i, between $\mathbf{n}_i$ and the vector to the grid point $\mathbf{t}_i$ from its associated expansion center.

There are many ways in practice to define the cavity surface, each leading to different sets of the quantities $\mathbf{t}_i$, a_i, $\mathbf{n}_i$, and γ_i. The way these items are used during solution of the reaction field equations is more or less independent of how they are obtained, so for purposes of the present section we simply take them as given quantities. Full details of the particular methods available in our program to construct the cavity are given in the following section.

In the SS(V)PE, IEF, COSMO, and SPE methods a system equation of the generic form $\mathscr{K}\sigma(\mathbf{t}) = \mathscr{Y}(\mathbf{t})$ must be solved to obtain the surface charge density, $\sigma(\mathbf{t})$, where $\mathscr{K}$ is the system operator that depends on the cavity construction and $\mathscr{Y}(\mathbf{t})$ is a linear function of the electrostatic potential, $\Phi^\rho(\mathbf{t})$, and/or normal electric field, $F_n^\rho(\mathbf{t})$, generated by the solute at the cavity surface. The different approximate surface charge methods are effectively specified [3] by different operators $\mathscr{K}$ and functions $\mathscr{Y}(\mathbf{t})$.

The basic assumption behind the discretization procedure is that $\sigma(\mathbf{t})$, $\Phi^\rho(\mathbf{t})$, and $F_n^\rho(\mathbf{t})$ vary slowly over the cavity surface so that across any small segment of surface area a_i each can be taken as essentially constant at its value for the representative point $\mathbf{t}_i$ in this area. The total charge on this segment of area can therefore be approximated by $q_i = a_i\sigma(\mathbf{t}_i)$. The generic system equation noted earlier is then approximated on the surface grid by a finite matrix equation of the form

$$\mathbf{Kq} = \mathbf{y} \ ,$$

where the matrix elements K_{ij} of the system matrix $\mathbf{K}$ represent in some sense the integrand of the integral operator $\mathscr{K}$, and elements of the right-hand side (RHS) column vector $\mathbf{y}$ are given by $y_i = \mathscr{Y}(\mathbf{t}_i)$. Expressions for $\mathbf{K}$ and $\mathbf{y}$ are given in Table 1 for each of the approximate surface charge methods considered in this work. These expressions are given in terms of more elementary constitutive matrices and vectors that are discussed in full detail later.

Similarly, in the SS(V)PE–μ, IEF–μ, COSMO–μ, and SPE–μ methods a system equation of the generic form $\mathscr{L}\mu(\mathbf{t}) = \mathscr{Z}(\mathbf{t})$ must be solved for the surface dipole density $\mu(\mathbf{t})$. Assuming that $\mu(\mathbf{t})$ varies only slowly over the surface we can approximate the total dipole on any small segment of surface area by $p_i = a_i\mu(\mathbf{t}_i)$. The different approximate surface dipole methods are effectively specified [3] by different operators $\mathscr{L}$ and functions $\mathscr{Z}(\mathbf{t})$. In this case the generic system equation is approximated on the surface grid by a finite matrix equation of the form

$$\mathbf{Lp} = \mathbf{z} \ ,$$

where the matrix elements L_{ij} of the system matrix $\mathbf{L}$ represent in some sense the integrand of the integral

Table 1. Matrices and column vectors involved in approximate surface charge methods

Method	Matrix $\mathbf{K}$	Vector $\mathbf{y}$
SS(V)PE[a]	$\mathbf{S} - \left(\frac{\epsilon-1}{\epsilon+1}\right)\left(\frac{1}{4\pi}\right)(\mathbf{DAS} + \mathbf{SAD}^*)$	$-\left(\frac{\epsilon-1}{\epsilon+1}\right)\left(\mathbf{I} - \frac{1}{2\pi}\mathbf{DA}\right)\mathbf{v}^\rho$
IEF[a]	$\mathbf{S} - \left(\frac{\epsilon}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{DAS} + \left(\frac{1}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{SAD}^*$	$-\left(\frac{\epsilon}{\epsilon+1}\right)\left(\mathbf{I} - \frac{1}{2\pi}\mathbf{DA}\right)\mathbf{v}^\rho + \left(\frac{1}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{SAf}^\rho$
COSMO	$\mathbf{S}$	$-\left(\frac{\epsilon-1}{\epsilon}\right)\mathbf{v}^\rho$
SPE	$\mathbf{A}^{-1} - \left(\frac{\epsilon-1}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{D}^*$	$-\left(\frac{\epsilon-1}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{f}^\rho$

[a] The $\mathbf{K}$ matrices for the SS(V)PE and IEF surface charge methods represent operators that are formally equivalent to one another. The matrix representation of these operators can also be written in factorized form as $\mathbf{S}\left[\mathbf{I} - \left(\frac{\epsilon-1}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{AD}^*\right]$ or as $\left[\mathbf{I} - \left(\frac{\epsilon-1}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{DA}\right]\mathbf{S}$

operator $\mathscr{L}$, and elements of the RHS column vector $\mathbf{z}$ are given by $z_i = \mathscr{L}(\mathbf{t}_i)$. Expressions for $\mathbf{L}$ and $\mathbf{z}$ are given in Table 2 for each of the approximate surface dipole methods considered in this work.

It remains only to specifically define the matrix elements and vectors corresponding to each of the elementary quantities involved in Tables 1 and 2. For example, the electrostatic potential generated by the solute charge density, $\rho(\mathbf{r})$, at the various grid points is denoted by the column vector $\mathbf{v}^\rho$ collecting the elements $v_i^\rho = \Phi^\rho(\mathbf{t}_i)$, and the corresponding normal electric field is denoted by the column vector $\mathbf{f}^\rho$ collecting the elements $f_i^\rho = F_n^\rho(\mathbf{t}_i)$.

The integrand of the unit operator $\mathscr{I}$ is simply represented by the diagonal unit matrix $\mathbf{I}$, and the diagonal matrix $\mathbf{A}$ has the surface area elements a_i on the diagonal. Some care must be taken with diagonal elements of the matrices $\mathbf{S}, \mathbf{D}$, and $\mathbf{D}^*$ that respectively represent the integral operators $\mathscr{S}$, $\mathscr{D}$, and $\mathscr{D}^*$, which are fully defined in Ref. [3], because of the singularities in their integrands. These singularities are weak and integrable. Thus, the diagonal matrix element shown here for these matrices does not strictly represent the integrand itself, but rather its value averaged over the segment of surface concerned, as obtained by integrating it over the surface segment and dividing the result by the area of this surface segment.

The integrand of the integral operator $\mathscr{D}$ is represented by the matrix $\mathbf{D}$ having elements

$$D_{ij} = \begin{cases} -(2\pi + \sum_{k\neq i} D_{ik} a_k)/a_i & \text{for } i = j \\ -\mathbf{n}_j \cdot (\mathbf{t}_j - \mathbf{t}_i)/|\mathbf{t}_j - \mathbf{t}_i|^3 & \text{for } i \neq j \end{cases} ,$$

and $\mathscr{D}^*$ is represented by the transpose matrix $\mathbf{D}^* = \mathbf{D}^T$. The expression used here for the diagonal elements was derived previously [28] and has also been used in other work [8, 29].

The integrand of the integral operator $\mathscr{S}$ is represented by the matrix $\mathbf{S}$ having elements

$$S_{ij} = \begin{cases} C_s\sqrt{4\pi a_i} f(\gamma_i) & \text{for } i = j \\ 1/|\mathbf{t}_i - \mathbf{t}_j| & \text{for } i \neq j \end{cases} ,$$

in which the diagonal elements require some comment. The factor $\sqrt{4\pi a_i}$ has been derived before [12, 20] and shown to give the exact result for a spherical surface [12, 20] and for a planar surface segment tangent to the surface normal [12], i.e., a segment for which the angle $\gamma_i = 0$. The empirical constant C_s was determined to have the value 1.07 for surfaces of the kind described in the original COSMO work [20] and the value 1.104 for surfaces described through Lebedev integration schemes [12]. Further investigation verifying the latter optimum value of C_s for Lebedev schemes is presented later in this work. The additional shape factor $f(\gamma_i)$ (not to be confused with the solute electric field that has a similar symbol in this work) is new. As shown in detail in the Appendix, it is derived to make the result exact for any small segment of surface area for which the angle γ_i is nonzero. Specifically it is found that

$$f(\gamma_i) = \left(\frac{2}{\pi}\right) K(\sin^2\gamma_i)\sqrt{\cos\gamma_i} \ ,$$

with K being the complete elliptic integral of the first kind (not to be confused with the system matrix that has a similar symbol in this work). The shape factor $f(\gamma_i)$ becomes unity when γ_i is zero and remains numerically very close to unity for small γ_i, so is practically irrelevant on a spherical or approximately spherical surface. However, it provides a significant improvement for severely nonspherical surfaces that contain some large values of the angle γ_i.

Two options are available to solve systems of simultaneous linear equations of the generic form $\mathbf{Kq} = \mathbf{y}$. Completely analogous comments apply to the generic $\mathbf{Lp} = \mathbf{z}$ equation, so henceforth we will explicitly refer only to the former equation. These options are based on slightly modified versions of standard routines available in the literature [30]. The first option calls the routine LUDCMP [30], which utilizes LU decomposition. This approach is relatively slow, but seems to be quite sure in providing a solution to near full machine precision in $\mathcal{O}(N^3)$ steps. This option is available only if the entire computation can be carried out in fast memory.

The generally faster second option to solve the simultaneous linear equations is based on the routine SPARSE [30], which utilizes conjugate gradient iterations. This option can be used even with matrices so large that external disk storage is required to hold the matrix $\mathbf{K}$. The convergence parameter EPS in this routine is set to a default value of 1×10^{-7}, which we have found generally leads to a final energy that is good to nearly full machine double precision (15S). If desired, the user can reduce the value of this convergence parameter to gain speed, at the cost of some sacrifice in the precision of the result. With dense matrices, such as

Table 2. Matrices and column vectors involved in approximate surface dipole methods

Method	Matrix $\mathbf{L}$	Vector $\mathbf{z}$
SS(V)PE–μ	$\mathbf{A}^{-1} - \left(\frac{\epsilon-1}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{D}$	$\left(\frac{\epsilon-1}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{v}^\rho$
IEF–μ[a]	$\mathbf{A}^{-1} - \left(\frac{2\epsilon}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{D} + \left(\frac{\epsilon-1}{\epsilon+1}\right)\left(\frac{1}{2\pi}\right)^2\mathbf{DAD}$	$\left(\frac{\epsilon}{\epsilon+1}\right)\left(\mathbf{I} - \frac{1}{2\pi}\mathbf{DA}\right)\frac{1}{2\pi}\mathbf{v}^\rho - \left(\frac{1}{\epsilon+1}\right)\left(\frac{1}{2\pi}\right)^2\mathbf{SAf}^\rho$
COSMO–μ	$\mathbf{A}^{-1} - \frac{1}{2\pi}\mathbf{D}$	$\left(\frac{\epsilon-1}{\epsilon}\right)\frac{1}{2\pi}\mathbf{v}^\rho$
SPE–μ[a]	$\mathbf{A}^{-1} - \left(\frac{2\epsilon}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{D} + \left(\frac{\epsilon-1}{\epsilon+1}\right)\left(\frac{1}{2\pi}\right)^2\mathbf{DAD}$	$\left(\frac{\epsilon-1}{\epsilon+1}\right)\left(\frac{1}{2\pi}\right)^2\mathbf{SAf}^\rho$

[a] The $\mathbf{L}$ matrices for the IEF–μ and SPE–μ surface dipole methods are identical to one another. They can also be equivalently written in factorized form as $\left[\mathbf{I} - \left(\frac{\epsilon-1}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{DA}\right]\left(\mathbf{A}^{-1} - \frac{1}{2\pi}\mathbf{D}\right)$ or as $\left(\mathbf{A}^{-1} - \frac{1}{2\pi}\mathbf{D}\right)\left[\mathbf{I} - \left(\frac{\epsilon-1}{\epsilon+1}\right)\frac{1}{2\pi}\mathbf{AD}\right]$

typically encountered in the present context, the time required for conjugate gradient iterative solution also seems to grow roughly as $\mathcal{O}(N^3)$, but with a more favorable coefficient than for the LU decomposition strategy.

The original SPARSE [30] routine calls the subroutine ASUB, which computes the matrix vector product $\mathbf{Kx}$, where $\mathbf{x}$ is the current estimate of the solution vector $\mathbf{q}$, and also calls the subroutine ATSUB, which similarly computes $\mathbf{K}^{\mathrm{T}}\mathbf{x}$. For methods that involve a symmetric system matrix so that $\mathbf{K} = \mathbf{K}^{\mathrm{T}}$, we appropriately modified the routine to remove all calls to ATSUB with calls instead to ASUB. For methods that involve a nonsymmetric system matrix we explicitly form the transpose matrix $\mathbf{K}^{\mathrm{T}}$, provided that it can also be held in fast memory, in order to minimize paging while forming the matrix vector product.

With the SS(V)PE and IEF methods, the default used to construct the system matrix corresponds to the symmetric operator listed for the SS(V)PE method in the main body of Table 1. Any of the other nonsymmetric forms noted in Table 1 can be optionally selected by the user. For large nonsymmetric matrix systems solved by conjugate gradient iterations it may happen that the two matrices $\mathbf{K}$ and $\mathbf{K}^{\mathrm{T}}$ do not both fit into fast memory, thus incurring a considerable performance penalty, whereas choosing instead the symmetric form of the system matrix would allow the entire computation to be carried out in fast memory.

When formulated as just described the SS(V)PE and IEF surface charge methods, as well as the IEF–μ and SPE–μ surface dipole methods, require a matrix multiplication of $\mathcal{O}(N^3)$ for construction of the system matrix. This allows subsequent solution of the system equation in one stage. For very large number of surface points such multiplication may be very inefficient, or even not possible to carry out in fast memory. For this situation an alternative two-stage procedure is available that eliminates the matrix multiplication, at the expense of requiring solution of two system equations instead of just one. The default is to allow the program itself to decide whether to use the one-stage or two-stage approach.

The two-stage approach takes advantage of the factorized forms of the relevant operators that are noted in the footnotes to Tables 1 and 2. For example, one alternative expression of the SS(V)PE matrix equation takes the form

$$\left[\mathbf{I} - \left(\frac{\epsilon - 1}{\epsilon + 1}\right)\frac{1}{2\pi}\mathbf{DA}\right]\mathbf{Sq}^{\mathrm{SS(V)PE}} = -\left(\frac{\epsilon - 1}{\epsilon + 1}\right)\left(\mathbf{I} - \frac{1}{2\pi}\mathbf{DA}\right)\mathbf{v}^{\rho} \quad .$$

Solution of this equation can be separated into two stages, the first stage solving

$$\left[\mathbf{I} - \left(\frac{\epsilon - 1}{\epsilon + 1}\right)\frac{1}{2\pi}\mathbf{DA}\right]\mathbf{v}^{\mathrm{SS(V)PE}} = -\left(\frac{\epsilon - 1}{\epsilon + 1}\right)\left(\mathbf{I} - \frac{1}{2\pi}\mathbf{DA}\right)\mathbf{v}^{\rho}$$

to obtain the intermediate quantity $\mathbf{v}^{\mathrm{SS(V)PE}}$, whose elements represent the SS(V)PE reaction potential at the surface (integrated over the surface segment areas), and the second stage then solving the COSMO-like equation

$$\mathbf{Sq}^{\mathrm{SS(V)PE}} = \mathbf{v}^{\mathrm{SS(V)PE}}$$

to finally obtain $\mathbf{q}^{\mathrm{SS(V)PE}}$. Optionally, a two-stage solution based on the other alternative factorized form of the relevant operator, as noted in the footnote to Table 1, can be selected. In the latter case, solution of the first stage produces a temporarily expedient but nonphysical intermediate quantity. With either the SS(V)PE or IEF method, slightly different discretization errors may result from these different solution strategies in practice because of their use of different forms of the system matrix. With either the IEF–μ or the SPE–μ method, which also involve matrix products, these various solution strategies will all lead to the same results in practice because there the matrices involved commute with one another.

3 Cavity definition

In principle, a wide latitude is allowed for defining the cavity surface $\mathbf{\Gamma}$. This is kept as a separate independent section of code as much as possible to provide maximum flexibility in including future alternatives and refinements. Two cavity constructions are currently available in the program. The first and simplest is a sphere having the center and the radius defined by the user. The second is an isodensity surface, i.e., a contour of constant solute electronic density whose value, ρ_0, is defined by the user. The isodensity prescription [29, 31, 32, 33] is known to closely mimic the actual shape of a molecular solute. In view of previous studies [9, 10] to determine the optimum value in neutral solutes, the default is chosen to be $\rho_0 = 0.001e/a_0^3$ in the program.

3.1 Surface grids

At present, only a single-center approach is available in the program for the construction of an isodensity surface. The strategy of using a single-center expansion has been described and used previously [8, 29]. If the cavity is substantially nonspherical this can limit the accuracy of the necessary surface integrations, and in extreme cases can fail to yield a satisfactory surface. The user first selects an origin for the cavity. The default is the center of nuclear change. Other options include the center of nuclear mass, the midpoint of the outermost atoms, the midpoint of the outermost non-hydrogen atoms, and finally arbitrary user-specified (x, y, z) coordinates.

The user is also allowed to specify an overall rotation of the Cartesian axes of the integration grid. The default is rotation to coincide with the principal moments of nuclear charge. Other options allow rotation to coincide with the principal moments of

nuclear mass, and rotation through arbitrary user-specified Euler angles.

Two methods are available for selecting grid points on the surface. Both start by defining points on an initial unit sphere centered at the cavity origin. In the first method, the total number of points is $N_{\theta\phi} = N_\theta N_\phi$, where the user selects the two integers N_θ and N_ϕ that define the number of points in the two angular coordinates of a spherical polar coordinate system. Thus, the dimension in latitude θ is divided into $\Delta\theta = \pi/N_\theta$ equal increments, and the dimension in longitude ϕ is divided into $\Delta\phi = 2\pi/N_\phi$ equal increments. The initial unit sphere weight of a point i is then the area $w_i = \sin\theta_i \Delta\theta\Delta\phi$ of the corresponding surface element. The default values are $N_\theta = 8$ and $N_\phi = 16$. If the solute has a point group symmetry of D_{2h} or any of its subgroups, the user should choose N_θ to be a multiple of 2 and N_ϕ to be a multiple of 4 to ensure that the solvation code maintains the desired symmetry.

The second method available for selecting surface points corresponds to the Lebedev [34] schemes, which are designed to numerically integrate exactly as many spherical harmonics as possible on a unit sphere. We incorporated the code developed by Dmitri N. Laikov and Christoph van Wuellen, as distributed on the internet through the Computational Chemistry List [35]. Users of this code are asked to include Ref. [34] in their publications. The user simply specifies a number, N_{Leb}, of Lebedev points and the code returns the corresponding positions and weights of that number of points on a unit sphere. Possible choices for N_{Leb} are 6, 14, 26, 38, 50, 86, 110, 146, 170, 194, 302, 350, 434, 590, 770, 974, 1202, 1454, 1730, 2030, 2354, 2702, 3074, 3470, 3890, 4334, 4802, 5294, or 5810. These all have octahedral symmetry and so are compatible with solutes having a point group symmetry of D_{2h} or any of its subgroups. The original code also has entries for N_{Leb} of 74, 230, and 266 but we disabled those three sets in our code because they each contain some points having negative weights, which precludes the physical association of weights with surface areas. Users are also warned to be wary of the sets having N_{Leb} of 146, 170, 350, 4802, and 5294, especially the 350 set. In some studies of convergence behavior (see later) we found that these five sets give results that deviate significantly from otherwise smooth plots of the dependence of the solvation energy on N_{Leb} in those methods that involve $\mathbf{S}$ in the system matrix. The default in our program is $N_{\mathrm{Leb}} = 1202$, which for most purposes is more than adequate for small solutes such as those studied in this work, but may be inadequate for very large and/or irregularly shaped solutes.

3.2 Single center cavity construction

To actually locate each grid point, $\mathbf{t}_i$, the ray is followed that emanates from the cavity origin in the direction of point i on the initial unit sphere, extending until the target surface $\mathbf{\Gamma}$ is reached. For a spherical cavity this just means extending the ray to have the length of the given radius. For an isodensity cavity a search is made along each ray until a point is found having a calculated solute electronic density essentially equal to the target value, ρ_0, specified by the user.

The search for the isodensity point along each ray is an iterative procedure, working inward after starting from a point in the direction of the ray on a large sphere whose radius is the distance from the cavity origin to the furthermost nucleus, plus an additional $2a_0$ for good measure. Both the solute density and radial density derivative are determined at the current search point. By default the radial density derivative is determined by projecting the desired radial component of the analytic Cartesian derivatives of the density, but the user can optionally determine it instead from a finite-difference approximation involving the current and the most recent previous point on the ray. It is known [36] that in the vicinity of the target value the radial dependence of the solute density in any particular direction can be closely approximated by a single exponential, $\beta\exp(-\alpha r)$, even when using relatively small Gaussian basis sets, and furthermore that the exponent α is nearly always well within a factor of 2 either way of the value 5 $\mathrm{\AA}^{-1}$. At the first (outermost) search point α is simply taken to have the generic value 5 $\mathrm{\AA}^{-1}$, while afterward the density and its radial derivative at each search point are fit to an exponential to obtain the associated effective exponent α. Assuming the actual density to be an exponential with the exponent α then allows straightforward prediction of the point along the ray that will have the target density of ρ_0. Far from the target this prediction may not be very good, so large jumps are damped by never allowing the distance from the current point to the next predicted point to exceed 20% of the distance from the cavity origin. Also, if the effective α is negative and the current density is less than ρ_0, the prediction is ignored entirely and the next point is instead taken to be 20% closer to the cavity origin. These steps are iteratively continued until the calculated solute electronic density matches the target contour value, ρ_0, within a given tolerance. This tolerance has a default value of 10^{-10} au, or can be otherwise specified by the user. The number of such steps, or iterations, taken to reach the target is capped by a maximum whose default value is 99. If the number of steps along any ray exceeds this maximum, the program is terminated with informational messages about failure to locate the surface.

A single-center treatment of the isodensity cavity surface is predicated on the assumption that each ray extending from the cavity origin outward will intersect the cavity surface only once. If the solute molecule is highly nonspherical, the surface may have one or more pockets that lead to violation of this assumption. To check for such behavior, a second determination of each surface point is made, this time working outward from near the cavity origin. If these two determinations reach essentially the same point on the ray, the surface point is accepted as valid. If they do not, this is taken as evidence that the surface has a pocket that precludes its characterization about this center. In that event, the program is terminated with appropriate informational messages about the failure. In some cases of failure, selecting a different center for the cavity origin may lead to a valid

construction. If the solute is severely nonspherical, a single-center approach to the isodensity cavity construction may not be possible at all. The user is given an option to bypass the cavity validity check, but this is not recommended.

The outward-directed searches carried out for the single-center validity check can be economized, since they need to provide only rough estimates of the surface point locations. Thus, solute densities are here evaluated only on a series of concentric equally spaced shells centered at the cavity origin. By default the innermost shell has radius 0.5 Å and other shells have radii increased by 0.5 Å increments. These values can be overridden by the user, in which case they should be chosen to be small enough to make it unlikely to miss any pockets in the surface. On a given shell, the solute electronic density is calculated at all points where rays intersect the shell. Additional shells are constructed working outward until all the solute densities on the outermost shell have fallen below the target value ρ_0. Along each ray, the two innermost adjacent shells which have solute densities bracketing the target value are noted. This then provides a rough estimate of each ray length good to within half the shell spacing, which is 0.25 Å with the default parameters.

If these searches all succeed, then along each ray we now have a highly refined value of the ray length obtained from working inward, together with a rough estimate of the ray length obtained from working outward. If these two separate determinations of the ray length agree to within the expected bracketing error, which is 0.25 Å for default outward-going shell spacing, the refined cavity surface point obtained by working inward is accepted as valid. If they do not agree within the expected bracketing error, the cavity surface determination is deemed to have failed. Furthermore, if the effective radial exponent α at any final refined point turns out to be negative, failure is declared.

Once the surface grid point locations, $\mathbf{t}_i$ are all satisfactorily determined, a separate evaluation is made at each grid point of the solute density, its analytic first Cartesian derivatives, and optionally also its analytic second Cartesian derivatives. For a spherical cavity, the outward-directed unit normal, $\mathbf{n}_i$, to the surface is parallel to the ray $\mathbf{t}_i$ and the surface area associated with point i is then simply $a_i = w_i t_i^2$, where w_i is the weight on the initial unit sphere (see earlier) and $t_i = |\mathbf{t}_i|$ is the ray length. For an isodensity surface, each unit normal vector to the surface is obtained from the analytic first Cartesian derivatives as the renormalized negative density gradient

$$\mathbf{n}_i = -\frac{\nabla\rho(\mathbf{t}_i)}{|\nabla\rho(\mathbf{t}_i)|} \quad ,$$

and the surface area associated with point i is obtained as

$$a_i = w_i t_i^2 \sec\gamma_i \quad ,$$

where γ_i is the acute angle between the ray $\mathbf{t}_i$ and the normal $\mathbf{n}_i$.

3.3 Evaluation of charge penetration

Enough information is available at this stage to estimate through direct calculation the amount of solute charge penetrating outside the cavity, $\bar{\rho}^{\text{ext}}$. The simplest approximation involves fitting the radially decaying density at each grid point to a single exponential. By "radial" we mean following along the direction of any given ray $\mathbf{t}_i$ from the cavity surface out to infinity. If ρ_i is the solute density at grid point i and ρ_i' is its radial first derivative, one can take $\rho(\mathbf{r}) \approx \beta_i \exp(-\alpha_i r)$ for all $\mathbf{r}$ outside the cavity and lying in the angular cone associated with $\mathbf{t}_i$, with the two fitting parameters having the values $\alpha_i = -\rho_i'/\rho_i$ and $\beta_i = \rho_i/\exp(-\alpha_i t_i)$. Then assuming this same single-exponential behavior remains valid everywhere in this cone outside the cavity, the contribution from this cone to the total penetrating charge is obtained by analytically evaluating the integral of this exponential from the cavity surface out to infinity and multiplying by the angular cone weight factor of $w_i t_i^2$. In typical molecular calculations involving Gaussian basis sets we found that this two-parameter fit typically leads to about 5–15% accuracy in the final result for $\bar{\rho}^{\text{ext}}$.

This estimate can usually be improved if analytic second derivatives of the density are also evaluated at the surface. We found that an effective three-parameter fit procedure can be obtained by taking $\rho(\mathbf{r}) \approx \beta_i[1 + \delta_i(r - t_i)^2]\exp(-\alpha_i r)$ for $\mathbf{r}$ outside the cavity and lying in the angular cone associated with $\mathbf{t}_i$, where β_i and α_i are the same as described earlier and δ_i is obtained by fitting to the additional information provided by the radial second derivative of the solute density. This expression is also easily integrated analytically to obtain its contribution to the total penetrating charge. In practice, the three-parameter fit procedure typically leads to about 1–10% accuracy in the final result for $\bar{\rho}^{\text{ext}}$.

In the particular case of the SPE method a generally more accurate indirect calculation of the penetrating solute charge can be made from knowledge of the apparent surface charges. The discrete analog of the appropriate expression [3] is simply

$$\bar{\rho}^{\text{ext}} = \bar{\rho} + \left(\frac{\epsilon}{\epsilon - 1}\right) \sum_i^{\text{grid}} q_i^{\text{SPE}} \quad ,$$

where $\bar{\rho}$ is the known total charge of the solute, for example, zero for a neutral species.

4 Implementation in HONDO

A program based on the previous description has been written in Fortran for use in the general electronic structure package HONDO [4], keeping it as much as possible in a separate self-contained code module called SVP. This module has four main sections.

An input section, SVPINP, is called only once, at the initialization stage of the entire molecular calculation. It reads all input data as keywords in a namelist format, applies default values for options not specifically declared by the user, and checks for mutual consistency of input parameters.

The remaining three main sections are each called in the self-consistent-field (SCF) determination of the electronic structure. These calls are essentially identical in the four SCF parts of HONDO, which respectively carry out closed-shell Hartree–Fock, open-shell Hartree–Fock, closed-shell density functional, and open-shell density functional calculations. As part of the SCF initialization, a cavity section, SVPSET, is called to determine the locations of all grid points and their associated surface areas on an initial spherical cavity.

The reaction field section, SVPCHG, is called once in each SCF iteration. If the actual cavity is to be an isodensity surface, this section first uses the current solute density to determine the isodensity cavity, starting from either the given initial sphere or the isodensity cavity obtained in the previous SCF iteration. In any case, it then sets up and solves the system matrix equation to obtain the surface point charges or point dipoles from the current solute density.

This step of solving the reaction field equation requires evaluation at the cavity surface of the solute electrostatic potential, $\mathbf{v}^{\rho}$, or the normal electric field, $\mathbf{f}^{\rho}$, or perhaps both depending on the particular reaction field method selected (Tables 1, 2). The solute potential and/or field at each surface point is evaluated essentially exactly by the program. In particular, we emphasize that no kind of multipole expansion is invoked. We frown on multipole expansions for this purpose, whether single-center or multicenter, because such expansions are convergent only outside the charge density they purport to represent. One of the major concerns of our reaction field work is to obtain proper behavior on a cavity surface that lies within a portion of the solute charge density, so any use of multipole expansions would be counterproductive. In any event, it is not computationally demanding to carry out essentially exact calculation of the solute potential and field on the cavity surface. In HONDO all necessary integrals for calculation of the electrostatic potential and field at the cavity points, as well as for calculation of the potential energy resulting from the apparent surface charges or surface dipoles, are based on a general and systematic extension of Rys quadrature for calculation of one- and two-electron integrals over Gaussian basis functions [37, 38, 39, 40].

The Hamiltonian section, SVPHAM, that is called in each SCF iteration handles the calculation of one-electron integrals for all pairs of basis functions over the accumulated surface point charges or point dipoles. These integrals are simply added to the one-electron integrals in the Fock matrix that has already been constructed for the internal energy of the solute. The program then goes on as usual to finish the SCF iteration by diagonalizing the (now modified) Fock matrix to obtain improved orbitals and an improved apparent total SCF energy that then include solvation effects.

The apparent total SCF energy must be corrected to obtain the proper total free energy. For this purpose we make use of several expressions given in Ref. [3] for the total solute–solvent reaction field energy. The most fundamental expression given there has, for a reaction field generated by a surface charge density, the discrete analog

$$\mathscr{E}^{\mathrm{rxn}} = \int_{\mathbf{V}} \mathrm{d}^3\mathbf{r}'\rho(\mathbf{r}') \sum_{i}^{\mathrm{grid}} \left(\frac{q_i}{|\mathbf{r}' - \mathbf{t}_i|} \right) \tag{1}$$

and for a reaction field generated by a surface dipole density has the discrete analog

$$\mathscr{E}^{\mathrm{rxn}} = \int_{\mathbf{V}} \mathrm{d}^3\mathbf{r}'\rho(\mathbf{r}') \sum_{i}^{\mathrm{grid}} p_i \left(\frac{\partial}{\partial \mathbf{n}_{\mathbf{t}_i}} \frac{1}{|\mathbf{r}' - \mathbf{t}_i|} \right) \ . \tag{2}$$

These are easily evaluated in HONDO [4] by making use of routines that evaluate basis set integrals over the potential from a given set of point charges, $\mathbf{q}$, or over the electric field from a given set of point dipoles, $\mathbf{p}$, and then accumulating the contributions of these integrals to the total solute density. For a reaction field generated by a set of surface charges, an alternative secondary expression [3] for the total reaction field energy has the discrete analog

$$\mathscr{E}^{\mathrm{q}} = \mathbf{q}^{\mathrm{T}}\mathbf{v}^{\rho} \ , \tag{3}$$

and for a reaction field generated by a set of surface dipoles an alternative secondary expression [3] for the total reaction field energy has the discrete analog

$$\mathscr{E}^{\mathrm{p}} = -\mathbf{p}^{\mathrm{T}}\mathbf{f}^{\rho} \ . \tag{4}$$

The contribution made to each of these expressions by the nuclear part, $\rho^{\mathrm{nuc}}(\mathbf{r})$, of the total solute charge density, $\rho(\mathbf{r})$, is easily explicitly evaluated manually, whence the contribution made by the electronic part, $\rho^{\mathrm{elec}}(\mathbf{r})$, can then be obtained by simply subtracting the nuclear part from the total.[1]

The total free energy is the internal solute energy plus half the reaction field energy, since the other half of the reaction field energy is expended in polarizing the solvent. By incorporating effects of the point charges or point dipoles in the Fock operator as described earlier, the apparent total SCF energy at this stage includes the full electronic contribution to the reaction field energy in a manner corresponding to fundamental expression of Eq. (1) or Eq. (2), but includes none of the nuclear contribution. Thus, to arrive at the proper total free energy from the apparent total SCF energy, it is necessary to make corrections by subtracting half the reaction field energy due to $\rho^{\mathrm{elec}}(\mathbf{r})$ and by adding half the reaction field energy due to $\rho^{\mathrm{nuc}}(\mathbf{r})$, these corrections being most conveniently evaluated through use of the $\rho^{\mathrm{elec}}(\mathbf{r})$ and $\rho^{\mathrm{nuc}}(\mathbf{r})$ contributions to the secondary expressions of Eq. (3) or Eq. (4). The free energy of solvation is just the total free energy so determined minus the gas-phase energy of the solute.

The normal and recommended route through the HONDO program first carries out a gas-phase calculation and then uses the gas-phase wave function as an initial guess to carry out a self-consistent reaction field calculation. With an isodensity cavity, the cavity surface

[1] By electronic and nuclear contributions to the reaction field energy, we mean here only those contributions arising from the explicit dependence of Eqs. (1), (2), (3) and (4) on $\rho(\mathbf{r})$. There is, of course, also a further implicit dependence on $\rho(\mathbf{r})$ through its effect on determining the reaction field point charges $\mathbf{q}$ or dipoles $\mathbf{p}$.

is updated each SCF iteration to be fully consistent with the polarized solute density. Convergence of the reaction field point charges or point dipoles is monitored by noting the largest absolute change in any point charge or point dipole from its value in the previous iteration. These changes are declared converged when the maximum change falls below a user-specified value, whose default is 1×10^{-7} au. Final convergence of the SCF calculation is not declared until both the usual criterion on the SCF density change and the criterion on the maximum point charge or point dipole change are each separately satisfied.

5 Precision of approximate reaction field calculations

It is generally necessary to have a means to know when the desired level of precision has been reached in a reaction field calculation. This might be just a modest precision if only the total energy is desired, or perhaps a very high precision if energy derivatives are to be evaluated by finite-difference methods for geometry optimization or other purposes. In this section, several performance issues are examined that determine the precision of reaction field calculations. These matters include the convergence with the number of grid points, N, the optimum value of the parameter C_s that scales diagonal elements of the **S** matrix, and the use of alternative forms of the system matrix. Evaluation of the rather different matter of the accuracy of various approximate reaction field methods is covered in Ref. [3].

5.1 Representative solutes

Results are given from calculations on several representative small solutes. Included are the polar neutral solutes H_2O and CH_3CONH_2. The former is roughly spherical, while the latter has a more complicated shape and so is more demanding of the surface integration procedures. The effect of solute charge is considered with the two isoelectronic solutes NO^+ and CN^-.

All the calculations on the neutral and cationic solutes were carried out with the restricted Hartree–Fock (RHF) method with the 6-31G** basis set [41] at the optimum gas-phase RHF/6-31G** geometries. For the anionic solute, diffuse functions were also included through use of the 6-31 + G** basis set [42], at the optimum gas-phase RHF/6-31 + G** geometry. The cavity is always defined by an isodensity surface having $\rho_0 = 0.001e/a_0^3$, and the dielectric constant is chosen as 78.304 to be representative of water solvent.

With these choices for the computational method, the amount of solute charge penetrating outside the cavity, $\bar{\rho}^{\text{ext}}$, is found to be $0.06e$, $0.16e$, $0.07e$, and $0.17e$ for H_2O, CH_3CONH_2, NO^+, and CN^-, respectively. Perhaps a better way to compare these is to normalize to the total number of solute electrons, whence the values of $\bar{\rho}^{\text{ext}}/N_{\text{elec}}$ for the cation and the two neutrals are all within about 20% of one another, while the value for the anion is almost twice as large. Another way to compare is to normalize instead to the cavity surface areas, which are 42, 99, 45, and 68 Å^2 for H_2O, CH_3CONH_2, NO^+, and CN^-, respectively. On this basis the values of $\bar{\rho}^{\text{ext}}$/area are within 10% of one another for the cation and two neutrals, and again the value for the anion is almost twice as large.

The values of $\bar{\rho}^{\text{ext}}$ given here are strictly valid only for the SPE method, where quite precise calculation is possible using the indirect formula noted earlier involving the total surface charge; however, the values obtained from the direct calculation methods discussed earlier, whereby $\bar{\rho}^{\text{ext}}$ is estimated from fitting the solute charge density and its derivatives at the cavity surface, indicate that the value of $\bar{\rho}^{\text{ext}}$ for a given solute varies only slightly among the various other reaction field methods considered. As might be expected, the variation is smallest for the cation and largest for the anion. Among the approximate surface charge methods the maximum variation is only $0.000001e$ for the cation, only $0.0002e$ for the two neutrals, and only $0.002e$ for the anion. The maximum variation is larger among the approximate surface dipole methods, but is still only $0.001e$ for the cation and only $0.003e$ for the two neutrals, while it becomes a more significant $0.04e$ for the anion.

5.2 Convergence with number of grid points

A major factor limiting the precision of a reaction field calculation arises from the discretization error associated with approximating a continuous surface charge (or dipole) density distribution with a finite set of surface charges (or dipoles). This error can be minimized by increasing the number of surface grid points, and it is therefore of interest to study convergence with respect to increasing N.

The SS(V)PE solvation energies of the four solutes considered are shown in Fig. 1. The horizontal axes are chosen to be linear in log N simply in order to spread the points out to assist visualization, and results are included for all $N \geq 50$. The CH_3CONH_2 results show the largest variation with N, the H_2O results vary somewhat less, and the NO^+ and CN^- results, which are very similar to one another in this respect, show the smallest variation. Comparing the two neutrals to one another, part of the larger variation with N of the CH_3CONH_2 compared to the H_2O results can be rationalized on the basis of the former having a little over twice as much surface area as the latter. Even after taking the size factor into account the variation with N is still greater with CH_3CONH_2 than with H_2O, indicating that the more complicated shape of the former also plays a significant role in determining the precision of the result. The two ionic solutes show considerably smaller variation of the solvation energy with N than do either of the neutral solutes. The variation is very similar for the two ionic solutes compared to one another, even though CN^- has nearly twice as much surface area as NO^+.

Exceptional behavior of the SS(V)PE results in Fig. 1 can be seen for N values of 146, 170, and especially 350, in that these results always lie significantly off the

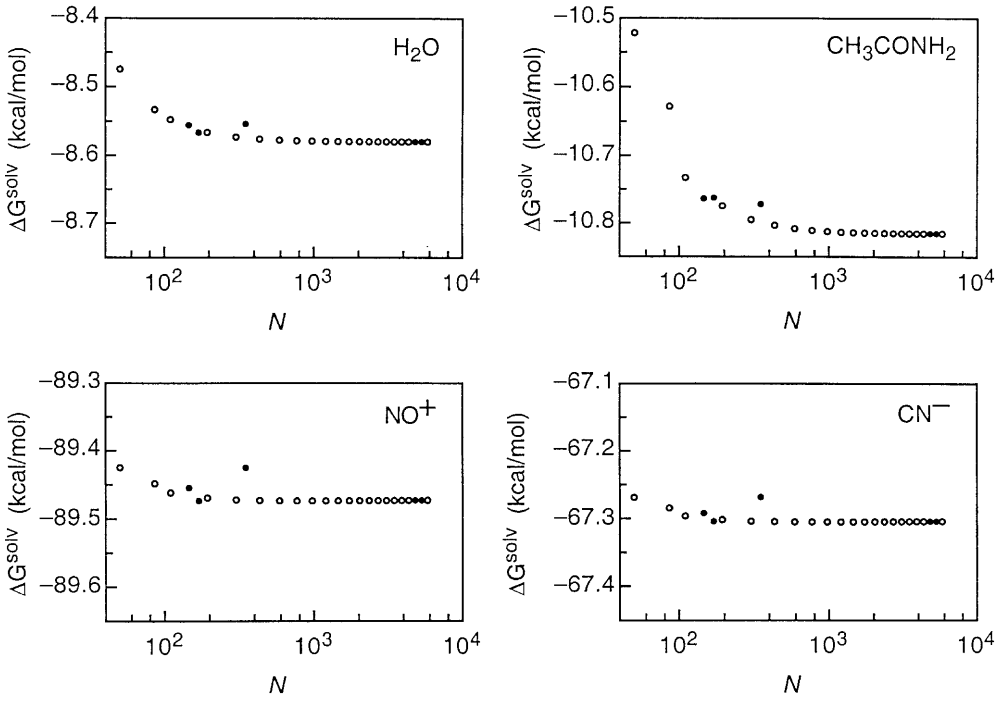

Fig. 1. SS(V)PE solvation free energy for representative solutes as a function of the number of surface grid points, N, on axes linear in $\log N$

otherwise smooth curve joining all other nearby results. Although not obvious on the scale of Fig. 1, examination on an expanded energy scale shows that the results for N values of 4802 and 5294 are also slightly, yet distinctly and consistently, off the otherwise smooth curve joining all other nearby results. Completely analogous behavior is also found with the COSMO and IEF methods, but not with SPE. This suggests that the exceptional behavior of these five integration sets is somehow related to their treatment of the **S** matrix. In any case, it is recommended that the 146, 170, 350, 4802, and 5294 Lebedev sets be excluded from any convergence or extrapolation studies when employing the SS(V)PE, COSMO, or IEF methods, and these exceptional sets are omitted from fitting schemes and discussions of trends given later. To emphasize the matter, these exceptional points are represented by filled-in circles in all the figures, while all other points use open circles.

To have a means for extrapolation to obtain the limiting result corresponding to an infinite number of grid points, N, we searched for the functional behavior of the computed solvation energy at large N. In each case examined we succeeded in finding an apparent asymptotic power-law dependence of the form $\Delta G^{\mathrm{solv}}(N) = G_\infty + g_n/N^n$ by guessing various values of n in increments of 0.5 until a plot of $\Delta G^{\mathrm{solv}}(N)$ versus $1/N^n$ gave a straight line for a broad range of large N values. The intercept of this plot then gives the extrapolated value of $\Delta G^{\mathrm{solv}}(\infty) = G_\infty$ that corresponds to an essentially exact solution of the given reaction field equation.

It is found that asymptotically the SS(V)PE solvation energy converges linearly as $1/N^2$. This is demonstrated in Fig. 2, where use of horizontal axes that are linear in $1/N^2$ allows good straight-line fits to the solvation energies.

From knowledge of the asymptotic dependence on N it is possible to extrapolate to obtain estimates of the limiting results for an infinite number of grid points. The top two panels in Fig. 2 show results with the neutral solutes for $N \geq 974$, using an energy scale having increments of 0.001 kcal/mol. Excellent straight-line fits against $1/N^2$ are obtained there for points having $N \geq 1202$, allowing for extrapolation to obtain the highly precise limiting values of -8.5795 kcal/mol for H_2O and -10.8157 kcal/mol for CH_3CONH_2. The bottom two panels in Fig. 2 show results with the ionic solutes for $N \geq 146$, using an energy scale having increments of 0.01 kcal/mol. Good straight-line fits against $1/N^2$ are obtained there for points in the range $194 \leq N \leq 974$, allowing for extrapolation to obtain the precise limiting values of -89.474 kcal/mol for NO^+ and -67.306 kcal/mol for CN^-.

The IEF and COSMO methods also show apparent asymptotic $1/N^2$ dependence, and with nearly the same coefficient g_2 as for SS(V)PE; thus, increasing the number of points by about a factor of 3 generally leads to about one more digit of precision in the SS(V)PE, COSMO, and IEF results.

It should be mentioned here that for very large N values the SS(V)PE, IEF, and COSMO results begin to deviate from the straight-line fits described earlier, indicating some departure of the convergence rate of the solvation energies from the apparent asymptotic $1/N^2$ behavior. The onset of this departure only occurs when the free energy has converged to about 0.0001 kcal/mol in the neutral solutes and to about 0.001 kcal/mol in the ionic solutes. Thus, with H_2O this departure is seen but slightly and only at the largest available value of N, and with CH_3CONH_2 it is not seen at all; however, for the ionic solutes NO^+ and CN^- this departure is seen for all $N \gtrsim 1000$. This curious behavior is discussed more in the

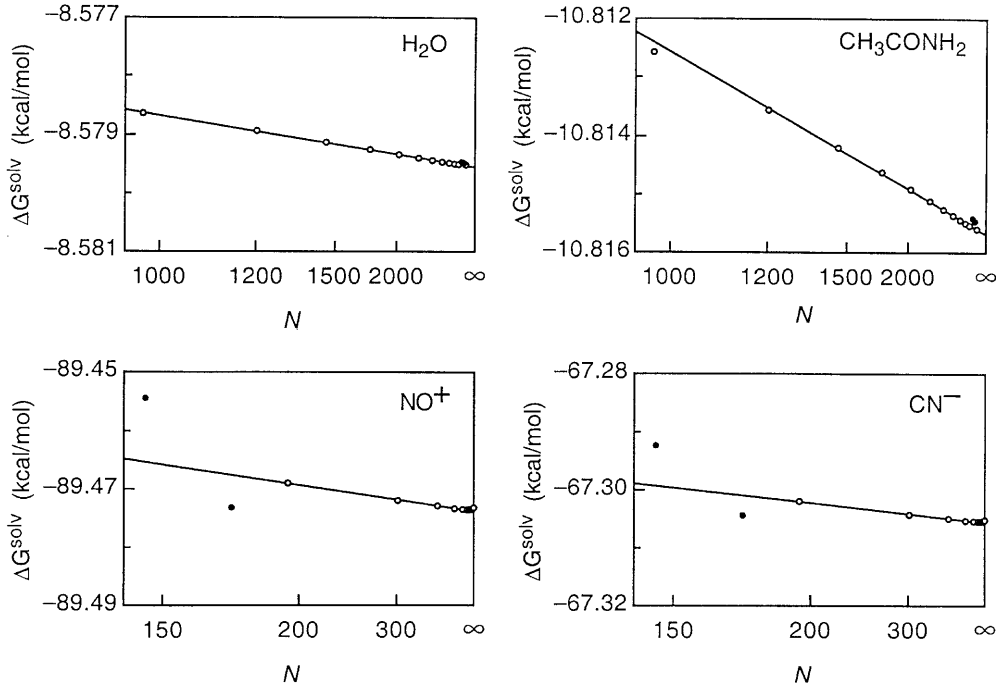

Fig. 2. SS(V)PE solvation free energy for representative solutes as a function of N on axes linear in $1/N^2$

next subsection, where it is suggested that an insufficiently precise value of the C_s parameter is responsible.

The SPE method and the SS(V)PE–μ, IEF–μ, and COSMO–μ methods each show a slightly weaker asymptotic $1/N^{1.5}$ dependence, but in compensation have similar coefficients $g_{1.5}$ that, at least for the range of available N values, are much more favorable than for the SS(V)PE, IEF, and COSMO surface charge methods. Thus, to reach a modest energy precision of, say, 0.1 or 0.01 kcal/mol for the small solutes examined here requires fewer points in SPE, SS(V)PE–μ, IEF–μ, and COSMO–μ than in SS(V)PE, COSMO, and IEF.

The SPE-μ method shows a slightly stronger asymptotic $1/N^{2.5}$ dependence at very large N, which appears to fall back to nearer $1/N^2$ for intermediate values of N. The coefficient is very similar to that of the surface charge SS(V)PE, IEF, and COSMO methods.

5.3 Optimum value of C_s parameter

The value of the empirical scaling factor, C_s used to scale diagonal elements of the **S** matrix was previously optimized [12] as 1.104 for use in Lebedev surface integration, on the basis of numerical studies of model systems with both a spherical cavity and a nonspherical cavity. Since the correction factor $f(\gamma_i)$ that is also included here to improve the precision of diagonal elements of the **S** matrix was not known at that time, it seems worthwhile to reexamine this matter when the $f(\gamma)$ correction factor is included.

The effects of varying C_s are shown for SS(V)PE results in Fig. 3. The middle line in each panel corresponds to the optimum C_s value of 1.104, while the outermost lines correspond to values of 1.070 and 1.140 that are approximately equally spaced below and above the putative optimum value. Examining vertical slices through these plots shows that for any given value of N the solvation free energy is essentially a linear function of C_s, at least over the narrow range of C_s values being considered.

Interestingly, we find that for C_s being either 1.070 or 1.140 the asymptotic convergence rate of the SS(V)PE energy deteriorates to a much weaker $1/\sqrt{N}$ dependence, compared to the strong $1/N^2$ dependence established previously when C_s is 1.104. For that reason, the horizontal axes in Fig. 3 are chosen to be linear in $1/\sqrt{N}$. This allows straight-line fitting and extrapolation of the outermost curves to obtain estimates for an infinite number of grid points, although noting the much cruder energy scales in Fig. 3 compared to Fig. 2 we should not expect such high precision as obtained earlier for C_s being the optimum 1.104. For H_2O, NO^+, and CN^- the range of apparent linearity in $1/\sqrt{N}$ for C_s being either 1.070 or 1.140 allows use of $N \geq 434$ in the extrapolation, while for CH_3CONH_2 only $N \geq 1202$ is used.

Extrapolations of the $C_s = 1.070$ and 1.140 results against $1/\sqrt{N}$ predict limiting values of −8.581 kcal/mol for H_2O, −10.817 kcal/mol for CH_3CONH_2, −89.473 kcal/mol for NO^+, and −67.305 kcal/mol for CN^-, these results being essentially the same for both values of C_s. These are in good agreement, all within 0.001 kcal/mol, of the presumably more precise values found previously from extrapolating the optimum $C_s = 1.104$ results against $1/N^2$. Thus, it is numerically verified that the extrapolated SS(V)PE solvation free energy for an infinite number of grid points is independent of the value of C_s, as it should be.

Very similar results are also obtained with the IEF and COSMO methods. It is concluded that the value of C_s affects the rate of convergence, but not the converged value, of SS(V)PE, IEF, and COSMO calculations. It is

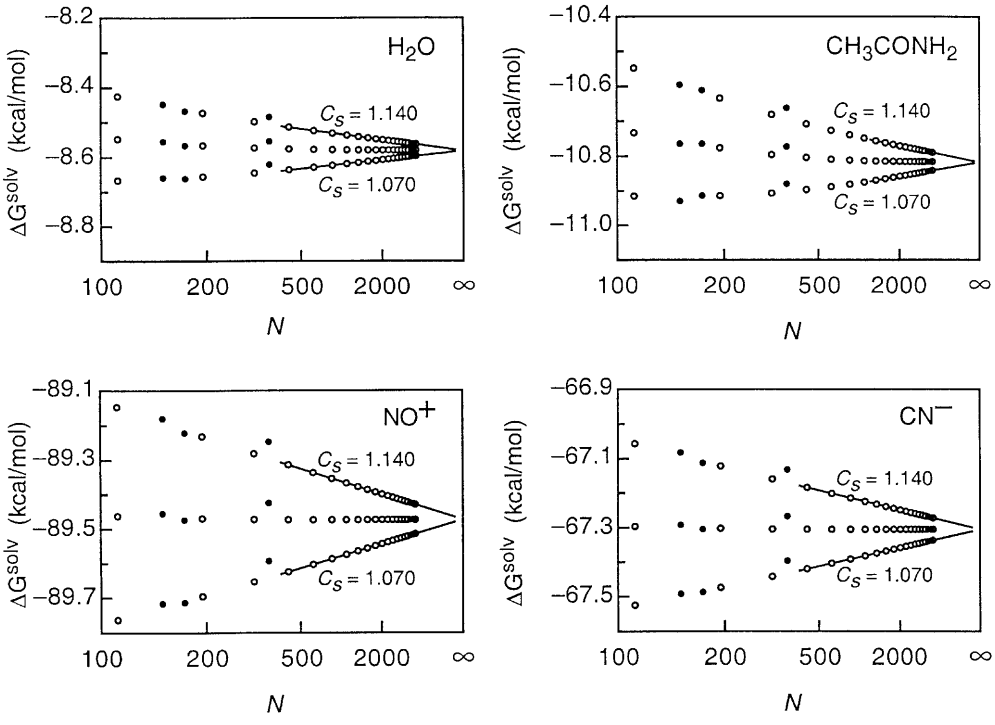

Fig. 3. SS(V)PE solvation free energy for representative solutes with different values of C_s as a function of N on axes linear in $1/\sqrt{N}$

also concluded that the optimum value of C_s for best convergence is close to 1.104.

The performance found here when C_s differs from its optimum value may explain the curious behavior noted in the previous subsection. Recall from there that when using the near-optimum value of 1.104 for C_s with the two ionic solutes, an asymptotic convergence rate of $1/N^2$ is found over a broad range of large N values up to about 1000. For still larger N, where the energy varies less than about 0.001 kcal/mol, the linearity in $1/N^2$ fails, indicating a falloff of the apparent convergence rate. In light of the behavior found in this subsection, where the convergence rate is seen to drop from $1/N^2$ to $1/\sqrt{N}$ when C_s does not have its optimum value, this curious behavior may be explained by the optimum value of about 1.104 for C_s not being precise enough to fully support the favorable $1/N^2$ rate when N gets very large.

At this point C_s is strictly an empirical correction factor. We surmise that if a theoretical approach could be found to reveal the true nature of this parameter, this might provide a more precise value for C_s that would support the $1/N^2$ convergence rate of SS(V)PE, COSMO, and IEF results to very high N values. Furthermore, it might well also suggest an improved formulation for diagonal elements of the **S** matrix that would further accelerate the convergence.

5.4 Alternative forms of the system matrix

Some of the reaction field methods are shown to have several alternative expressions for the system matrix in Tables 1 and 2. For example, the SS(V)PE system matrix given in the main body of Table 1 is symmetric, while the three alternative forms (corresponding to the form given for IEF and to the two factorized forms given in the footnote) are each nonsymmetric. Owing to discretization errors, these four alternative expressions may lead to slightly different results in practice.

SS(V)PE calculations with these alternative expressions show that the symmetric form of the system matrix invariably leads to results about midway between those from the two nonsymmetric factorizable forms. The nonsymmetric IEF form also produces results between the limits of the two nonsymmetric factorizable forms, lying closest to the last listed version of the latter. However, the numerical energy differences between all four alternative forms of the system matrix are generally small. It is most appropriate to compare these differences to the inherent discretization error, which is the difference between the energy at a given value of N and the extrapolated energy corresponding to infinite N. The differences obtained from alternative forms of the system matrix are greatest at small values of N, but even there are at most only about 10% of the discretization error, and at larger N the differences become a still smaller fraction of the discretization error. An independent study [11] has also found only small differences in calculations using alternative forms of the system matrix. It is therefore concluded that distinction of the results from these alternative forms of the system matrix is not very significant in practice.

6 Summary and conclusion

Details were given for practical boundary element solutions of the formal equations defining various approximate reaction field methods by discretization

over the surface to allow use of finite matrix methods. Choices of surface grids were discussed, and considerable detail was given on practical construction of an isodensity cavity surface based on a single-center treatment. A computer program was described to carry out such reaction field calculations, as implemented in the standard HONDO [4] electronic structure code.

For large number of surface points, N, the asymptotic convergence to reach the limiting value was found in each approximate reaction field method considered to vary as g_n/N^n. The SS(V)PE, IEF, and COSMO methods each converge with $n=2$ and with very similar values of the coefficient g_2. The SPE, SS(V)PE–μ, COSMO–μ, and IEF–μ methods each converge with a slightly weaker $n = 1.5$ dependence and with very similar values of the coefficient $g_{1.5}$. The SPE–μ method appears to converge with $n = 2.5$. Despite the less favorable $n = 1.5$ dependence in the SPE, SS(V)PE–μ, IEF–μ, and COSMO–μ methods, the coefficient $g_{1.5}$ is so small that convergence in an absolute sense to say 0.1 or 0.01 kcal/mol is reached in small solutes with a smaller number of points than with the SS(V)PE, IEF, and COSMO methods.

The optimum value of the empirical factor, C_s, scaling diagonal elements of the **S** matrix involved in several of the methods was investigated numerically. It was found that for Lebedev integration schemes the optimum value of C_s is equal to or very near the previously determined value of 1.104, even when a new shape factor, $f(\gamma_i)$, derived here to improve precision is included.

The SS(V)PE and IEF methods each have a system matrix that can be written in several alternative forms that lead to slightly different results owing to discretization error on a finite grid. Numerical studies indicate that the differences between these forms are small in practice, generally being less than about 10% of the inherent discretization error for small N values and a still smaller fraction of the discretization error for larger N values.

Future work on the reaction field program described here is planned in several directions. Concerning the cavity itself, additional definitions of the surface will be incorporated, and the surface integration schemes will be improved to better treat solutes having highly irregular shapes. Concerning the reaction field, inertial and optical responses will be separated to allow determination of vertical excitation energies, and a capability to treat ionic strength effects will be developed.

Acknowledgements. The research described here was supported in part (D.M.C.) by the Office of Basic Energy Sciences of the Department of Energy and in part (M.D.) by the Pacific Northwest National Laboratory. The Pacific Northwest National Laboratory is a multiprogramme national laboratory operated for the U.S. Department of Energy by Battelle Memorial Institute under contract DE-AC06-76RLO 1830. This is contribution no. NDRL-4314 from the Notre Dame Radiation Laboratory.

Appendix

Here we outline the derivation of our formula for diagonal elements, S_{ii}, of the matrix representation **S** of the integral operator $\mathscr{S}$. The surface grid point i lies at a distance t_i from the expansion center, and γ_i is the acute angle between the ray $\mathbf{t}_i$ and the surface normal $\mathbf{n}_i$. This point is representative of some small neighborhood of surface points $\{\mathbf{t}\}$.

With Lebedev surface integration the actual shape of this neighborhood is not specified, so we begin by making a reasonable construction. The notation is considerably simplified, and no generality is lost, if the grid point i under consideration is taken to lie at the north pole of a spherical polar coordinate axis system, in which θ measures the angle between the ray $\mathbf{t}$ and the z-axis. We assume a simple construction whereby the shape of this neighborhood is dictated by where a cone of solid angle having $0 \le \theta \le \theta_i$ and $0 \le \phi \le 2\pi$ intersects with the actual cavity surface. The value of the upper limit, θ_i, will be determined through consideration of the surface area of the segment, which is defined through

$$a_i = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\theta_i} \mathrm{d}\theta\, t(\theta,\phi)^2 \sin\theta \sec\gamma_{\mathbf{t}}(\theta,\phi) \quad .$$

Any point $\mathbf{t}$ in this neighborhood will have a distance from the expansion center that is close to that of the representative point, i.e., $t(\theta,\phi) \approx t_i$. Similarly, the normal angle at any point in this neighborhood will be close to the normal angle at the representative point, i.e., $\gamma_{\mathbf{t}}(\theta,\phi) \approx \gamma_i$. Since θ_i is expected to be small, we can make the expansions

$$t(\theta,\phi) = t_i[1 + u(\phi)\theta + \mathscr{O}(\theta^2)] \quad ,$$

$$\sec\gamma_{\mathbf{t}}(\theta,\phi) = \sec\gamma_i[1 + v(\phi)\theta + \mathscr{O}(\theta^2)] \quad .$$

A short derivation shows that

$$u(\phi) = \tan\gamma_i \sin(\phi + \phi_0) \quad .$$

Our final result will turn out to be independent of the phase shift angle, ϕ_0, so it need not be explicitly evaluated; in fact it can be made to vanish by proper orientation of the x-axis of the coordinate system. The quantity $v(\phi)$ need not be explicitly evaluated either; it is sufficient for the present purposes to note that, like $u(\phi)$, it is a sinusoidal function of period 2π, so its integral vanishes over the full period $0 \le \phi \le 2\pi$.

Using these expansions, we easily find that the area of the segment under consideration is $a_i = \pi t_i^2 \sec\gamma_i \theta_i^2 + \mathscr{O}(\theta_i^4)$. Equating this with the expression $a_i = w_i t_i^2 \sec\gamma_i$ that was previously given in the main body of this work for the same area, where w_i is the corresponding weight on the initial unit sphere, shows that $\sqrt{w_i/\pi} = \theta_i + \mathscr{O}(\theta_i^3)$, which can be turned around to provide a good estimate for the value of θ_i that determines the size of the segment.

Now we can consider the corresponding **S** matrix diagonal element, which is defined by

$$S_{ii} = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\theta_i} \mathrm{d}\theta \frac{t(\theta,\phi)^2 \sin\theta \sec\gamma_{\mathbf{t}}(\theta,\phi)}{|\mathbf{t} - \mathbf{t}_i|} \quad .$$

By introducing the expansions used earlier and noting that

$$\int_0^{2\pi} \mathrm{d}\phi \frac{1}{\sqrt{1+u^2(\phi)}} = 4K(\sin^2\gamma_i)\sqrt{\cos\gamma_i} \ ,$$

we obtain

$$S_{ii} = 4K(\sin^2\gamma_i)t_i\theta_i + \mathcal{O}(\theta_i^3) \ ,$$

which leads to the final result

$$S_{ii} = \left(\frac{2}{\pi}\right)K(\sin^2\gamma_i)\sqrt{4\pi a_i\cos\gamma_i} + \mathcal{O}(\theta_i^3)$$

that is used in the main body of this work. Note that the explicitly evaluated first term on the RHS of the last equation is linear in θ_i and that the first correction term in θ_i^2 vanishes, making the leading nonzero correction of $\mathcal{O}(\theta_i^3)$. Approximation using only the explicitly evaluated term on the RHS should therefore be robust for small θ_i.

References

1. Tomasi J, Persico M (1994) Chem Rev 94: 2027
2. Cramer CJ, Truhlar DG (1999) Chem Rev 99: 2161
3. Chipman DM (2001) Theor Chem Acc
4. Dupuis M, Marquez A, Davidson ER (1999) HONDO 99, a computer program based on HONDO 95.3. Quantum Chemistry Program Exchange, Indiana University, Bloomington
5. Chipman DM (1996) J Chem Phys 104: 3276
6. Mennucci B, Tomasi J (1997) J Chem Phys 106: 5151
7. Chipman DM (1997) J Chem Phys 106: 10194
8. Zhan C-G, Bentley J, Chipman DM (1998) J Chem Phys 108: 177
9. Zhan C-G, Chipman DM (1998) J Chem Phys 109: 10543
10. Zhan C-G, Chipman DM (1999) J Chem Phys 110: 1611
11. Cossi M, Rega N, Scalmani G, Barone V (2001) J Chem Phys 114: 5691
12. Chipman DM (1999) J Chem Phys 110: 8012
13. Chipman DM (2000) J Chem Phys 112: 5558
14. Mennucci B, Cammi R, Tomasi J (1998) J Chem Phys 109: 2798
15. Cancès E, Mennucci B (2001) J Chem Phys 114: 4744
16. Cancès E, Mennucci B, Tomasi J (1997) J Chem Phys 107: 3032
17. Mennucci B, Cancès E, Tomasi J (1997) J Phys Chem B 101: 10506
18. Cancès E, Mennucci B (1998) J Math Chem 23: 309
19. Tomasi J, Mennucci B, Cancès E (1999) J Mol Struct (THEOCHEM) 464: 211
20. Klamt A, Schüürmann G (1993) J Chem Soc Perkin Trans 2: 799
21. Andzelm J, Kölmel C, Klamt A (1995) J Chem Phys 103: 9312
22. Klamt A (1996) J Phys Chem 100: 3349
23. Klamt A, Jonas V (1996) J Chem Phys 105: 9972
24. Baldridge K, Klamt A (1997) J Chem Phys 106: 6622
25. Truong TN, Stefanovich EV (1995) Chem Phys Lett 240: 253
26. Truong TN, Stefanovich EV (1995) J Chem Phys 103: 3709
27. Truong TN, Nguyen UN, Stefanovich EV (1996) Int J Quant Chem Symp 30: 403
28. Purisima EO, Nilar SH (1995) J Comput Chem 16: 681
29. Foresman JB, Keith TA, Wiberg KB, Snoonian J, Frisch MJ (1996) J Chem Phys 100: 16098
30. Press WH, Flannery BP, Teukolsky SA, Vetterling WT (1986) Numerical recipies. The art of scientific computing. Cambridge University Press, Cambridge
31. Bader RFW, Henneker WH, Cade PE (1967) J Chem Phys 46: 3341
32. Bader RFW, Preston HJT (1970) Theor Chim Acta 17: 384
33. Bentley J (2000) J Phys Chem A 104: 9630
34. Lebedev VI, Laikov DN (1999) Dokl Math 59: 477
35. http://www.ccl.net/cca/software/SOURCES/FORTRAN/Lebedev-Laikov-Grids/index.shtml
36. Bentley J (1998) J Phys Chem A 102: 6043
37. Dupuis M, Rys J, King HF (1976) J Chem Phys 65: 111
38. Rys J, Dupuis M, King HF (1983) J Comput Chem 4: 154
39. Dupuis M (2001) Comput Phys Commun 134: 150
40. Dupuis M, Marquez A (2001) J Chem Phys 114: 2067
41. Hariharan PC, Pople JA (1973) Theor Chim Acta 28: 213
42. Clark T, Chandrasekhar J, Spitznagel G, Schleyer PvR (1983) J Comput Chem 4: 294