

Geometry-Guided Computation of 3D Electrostatics for Large Molecular Datasets

Xuejun Hao*, Amitabh Varshney

Graphics and Visual Informatics Laboratory, Department of Computer Science and UMIACS, A. V. Williams Building, University of Maryland, College Park, MD 20742

Abstract

Electrostatic interactions play a central role in biological processes. Development of fast computational methods to solve the underlying Poisson-Boltzmann equation (PBE) is vital for biomolecular modeling and simulation package. In this paper, we propose new methods for efficiently computing the electrostatic potentials for large molecules by using the geometry of the molecular shapes to guide the computation. The accuracy and stability of the solution to the PBE is quite sensitive to the boundary layer between the solvent and the solute which defines the molecular surface. In this paper, we present a new interface-layer-focused PBE solver. First, we analytically construct the molecular surface of the molecule and compute a distance field from the surface. We then construct nested iso-surface layers outwards and inwards from the surface using the distance field. We have developed a volume simplification algorithm to adaptively adjust the density of the irregular grid based on the importance to the PBE solution. We have generalized the finite difference methods using Taylor series expansion on the irregular grids. Our algorithm achieves about three times speedup in the iterative solution process of PBE, with more accurate results on an analytical solvable testing case, compared with the popular optimized DelPhi program.

Key words: Tetrahedron decomposition, iso-surface generation, level-of-detail of hierarchy, scalar field, Poisson-Boltzmann equation, finite difference methods

* Corresponding author. Address: Unit 74, 1051 Riverside Drive, Psychiatry Department, Columbia University, New York, NY 10032

Email addresses: xh2108@columbia.edu (Xuejun Hao), varshney@cs.umd.edu (Amitabh Varshney).

1 Introduction

Electrostatic interactions are of central importance for many biological processes (Leach, 2001; Tainer et al., 1985). Experiments have shown that electrostatics influence nearly all biochemical reactions, such as macromolecular folding and conformational stability. Electrostatics also determine the structural and functional properties of biological samples, such as their shapes, binding energies, and association rates. Molecular modeling packages (Humphrey et al., 1996) have invested significant effort in correctly and efficiently modeling the electrostatics to simulate the static structure and binding energy, in addition to modeling user-defined conformations (Kreylos et al., 2003) or trajectories (Leach, 2001). The successful modeling of electrostatics has great practical, as well as, theoretical importance, for structure-based drug design and protein folding. There are two ways to model the electrostatic properties of biological samples – quantum mechanical methods and classical electrostatics. Quantum mechanical methods are more accurate, but due to their immense computational demands, can only be applied to small molecules. Thus the application of quantum mechanical methods to large molecules, such as the ones we consider here, is currently not possible for real-time systems.

Classical electrostatic interactions are modeled as the interactions between partial atomic charges (also called net atomic charges). Partial atomic charges arise since electronegative elements, such as Oxygen, attract electrons more readily than elements such as Hydrogen. This give rise to an unequal distribution of charges in a molecule. The electrostatics of molecules depend not only on their 3D structures and charge distributions, but also on their environment. Biological processes occur in aqueous solution, so solvent plays an important role in determining the electrostatics of the solute molecules. Solvent properties are usually described in terms of average values. Thus, instead of treating each solvent atom explicitly, we treat them as a continuum with average properties. The more important solute molecules are treated explicitly (Honig and Nicholls, 1995). The Poisson-Boltzmann equation (PBE) describes the electrostatic interactions in solution. Several methods have been developed to solve the PBE efficiently. They normally solve the equation and compute the electrostatic potentials either on the surface of the solute molecules, or on the 3D grid discretizing the space, or on a slice of the 3D space.

In this paper, we propose a new method for efficiently computing the electrostatic potentials. Our method is based on the observation that the accuracy and stability of the solution to PBE is quite sensitive to the boundary layer between the solvent and the solute. So an accurate construction of this boundary with adaptively controlled grid density should improve the solution. In our algorithm, we first analytically construct the molecular surface, then build a tetrahedral decomposition of the 3D space around the surface and a distance

field from the surface. Next we build iso-surfaces by the marching-tetrahedra method on the distance field with progressively greater distances. This results in nested isosurfaces at varying distances from the molecular surface. After that, we apply an edge collapse algorithm to simplify the tetrahedral grids to adaptively adjust the grid point density in space according to their influence on the solution. We maintain a higher resolution for the solution-sensitive region in the vicinity of the molecular surface. We have found that this improves the accuracy and stability of the solution while speeding up the computation. In our work, we have generalized the traditional finite difference computations on regular grids to irregular grids by a Taylor series expansion. We validate our algorithm on an analytical solvable case and compare the results with the popular DelPhi program.

The main contributions of this paper are:

1. We show that PBE can be solved more efficiently by taking advantage of the exact geometry around the solution-sensitive boundary layer.
2. We achieve better results by using an application-driven hierarchy of detail for 3D space decomposition.

2 Previous Work

Theoretical models for molecular electrostatics fall into two categories. *Microscopic models* explicitly treat each atom in the protein or solvent molecule, and each ion in the surrounding solution. Macroscopic or *Continuum models* describe properties of groups of molecules or ions in terms of average values. The earliest models were all macroscopic. An increase in the computational power has gradually allowed us to incorporate greater atomic-level detail for solute molecules. Currently, the most widely used models still treat the solvent as a continuum (G. Allen (editor), 1999).

Modern electrostatic models are based on non-linear Poisson-Boltzmann equation (PBE). If there are no highly charged molecules or high ionic strengths, the equation can be well-approximated by a linear PBE, and an analytic solution is possible (Tanford and Kirkwood, 1957). In most cases, however, such an analytical solution does not exist, and numeric methods have been developed to solve linear (Warwicker and Watson, 1982) or nonlinear PBE. Among them the finite difference method (FDM) (Warwicker and Watson, 1982) is the most widely used. In finite difference methods the molecule is mapped onto a 3D grid. Ionizable atoms are assigned to grid points and the electrostatic potential at each grid point is calculated using the finite difference approximation

of the PBE. The accuracy of the results is highly dependent on grid spacing, while the computational cost increases steeply with the number of grid points. One approach to reduce the cost is *focusing* (Gilson et al., 1988), in which the mesh of the grid is reduced only in the vicinity of ionizable groups of particular interest with potentials from coarser grids used as initial guesses. A more powerful approach is the *multi-grid method* (Oberoi and Allewell, 1993), in which the solution on a given grid (generally the finest grid), is obtained by iterating over a hierarchy of coarser grids. The key advantage is that the accuracy of the solution is iteratively improved by solving the problem on the coarser grids where the computational cost is low with infrequent visits to the finer grids where the computational cost is high. One drawback of the multi-grid method is that, while it works gracefully for linear PBE, the solution might not converge when applied to non-linear PBE.

The finite difference algorithms using regular grid, though quite successful, have several shortcomings. First, their computational cost is proportional to the cube of the grid size, which makes it very hard to increase the resolution. Second, they do not scale well. For fixed grid size, the resolution will decrease as the dataset becomes bigger. Third, the low resolution approach, even with multi-grid refinements, may introduce visual artifacts at the visualization stage.

Adaptive space-subdivision approach (Baker et al., 2000; Holst et al., 2000) has been used to address the high cost of using a regular grid. This approach increases the accuracy of the solution by explicitly giving a higher spatial resolution to the solvent-solute boundary region. However, since the adaptive space-subdivision approach does not start from an analytical definition of the solvent-solute boundary (the molecular surface), it tends to over-subdivide around the boundary region. This over-subdivision results in two drawbacks. First, it increases the number of grid points and therefore the time for each iteration of the PBE solver. Second, it increases the number of iterations to reach a desired convergence threshold since a greater number of closely-spaced points near the boundary increases the time to propagate the solvent-solute boundary effects.

Our new approach solves the linear or non-linear PBE on an irregular grid. It has the advantage that the PBE-solution-sensitive boundary layer is constructed analytically. We generate a progressive level of grid detail with the highest levels of detail around the solution-sensitive solvent-solute boundary region that gradually change to coarser levels of detail farther from the solvent-solute boundary. Since we start with an analytical definition of the solvent-solute boundary, we do not suffer from the problems of over-subdivision in trying to straddle that boundary. This allows us to speed up the time to compute the PBE solution with lesser number of points than previous work.

3 Fundamentals of Molecular Electrostatics

The current trends in real-time molecular electrostatics follow the principles of classical electrostatics, treating the solvent as a continuum. In this section we outline the physical foundations for molecular electrostatics and explain the application of different forms of the Poisson-Boltzmann Equation.

3.1 *Electrostatics in Uniform Dielectric Medium*

Electrostatics has a simple form when all the charges and field considered are in a uniform dielectric medium, including vacuum. The electrical potential then satisfies the Poisson equation (Jackson, 1975):

$$\nabla^2\phi(\vec{r}) + \frac{4\pi\rho(\vec{r})}{\epsilon} = 0$$

where $\phi(\vec{r})$ is the electrostatic potential, $\rho(\vec{r})$ is the charge density, and both $\phi(\vec{r})$ and $\rho(\vec{r})$ are functions of position. Here the dielectric constant, ϵ , is independent of the position in uniform media.

As an example, the electric potential field generated by a point charge is given by Coulomb's Law (Jackson, 1975): $\phi(r) = \frac{q}{\epsilon r}$, where the point charge q is assumed to be at the origin and r is the distance from the origin. Here the linear superposition rule holds and the electric potential field generated by a set of point charges is the summation of the fields generated by each point charge (Jackson, 1975): $\phi(\vec{r}) = \sum_{i=1}^n \phi_i(\vec{r}) = \sum_{i=1}^n \frac{q_i}{\epsilon|\vec{r}-\vec{r}_i|}$, where n is the number of point charges, q_i is the charge and \vec{r}_i is the position vector of point charge i .

3.2 *Electrostatics in Nonuniform Medium with Environmental Response*

The Poisson equation given in previous subsection assumes uniform medium and without the environmental response. If the dielectric ϵ varies through space, then we arrive at a general form of the Poisson equation:

$$\nabla[\epsilon(\vec{r})\nabla(\phi(\vec{r}))] + 4\pi\rho(\vec{r}) = 0$$

where $\epsilon(\vec{r})$ is a function of position. Normally the solute is treated as a uniform medium with a low relative dielectric of about $2 \sim 4$. The solvent is also treated as a uniform medium with a relative dielectric of about 80 (G. Allen (editor), 1999).

The environmental response consists of three physical processes that screen the effects of charge: (a) electronic polarization; (b) reorientation of permanent dipole in polar materials; and (c) redistribution of charges, such as mobile ions. Combining these factors, we get a general form for the molecular electrostatics – the Poisson-Boltzmann Equation (PBE):

$$\nabla[\epsilon(\vec{r})\nabla(\phi(\vec{r}))] - \kappa'^2(\vec{r})\sinh[\phi(\vec{r})] + 4\pi\rho(\vec{r}) = 0$$

where κ' is the modified Debye–Hückel parameter and defined as: $\kappa'^2 = \frac{8\pi N_a e^2 I}{1000kT}$, where N_a is Avogadro’s number, e is the electron charge, k is Boltzmann constant, T is the absolute temperature, and I is the ionic strength of the bulk solution. The variables ϕ , ϵ , κ' , and ρ are all functions of the position vector \vec{r} . The general PBE above incorporates electronic and dipole polarization through ϵ and ion-screening through κ' .

If there are no highly-charged molecules and ionic strengths are low, we can make an approximation to linearize the sinh term: $\sinh[\phi(\vec{r})] \approx \phi(\vec{r})$, and then the general PBE simplifies to the linear PBE:

$$\nabla[\epsilon(\vec{r})\nabla(\phi(\vec{r}))] - \kappa'^2(\vec{r})\phi(\vec{r}) + 4\pi\rho(\vec{r}) = 0$$

If there are no mobile ions present in the system, the modified Debye–Hückel parameter κ' will be equal to zero, and PBE reduces to the general Poisson equation.

4 Finite Difference Method for PBE

The analytical solution to the PBE exists for very few cases, with highly simplifying assumptions. For almost all cases, we have to rely on numerical methods to arrive at the solution. The most widely-used numerical method for the solution of the PBE is the finite difference method (FDM), introduced by Warwicker and Watson (Warwicker and Watson, 1982) in 1982.

In FDM the molecule and a region of the surrounding solvent are mapped onto a 3D grid. Each grid point represents a small region of either the molecule or the solvent. Values are assigned at each point for the charge density, dielectric constant, and ionic strength parameters in the PBE. With a fine enough grid scale, variation in the dielectric response can be represented at the atomic resolution. The electrostatic potential at each grid point is calculated using the finite difference approximation of the PBE:

$$\phi_0^{new} = \frac{\sum \epsilon_i \phi_i + 4\pi q_0/h}{\sum \epsilon_i + \kappa'^2 h^2 [1 + \phi_0^2/3! + \dots + \phi_0^{2n}/(2n+1)! + \dots]}$$

where the non-linear term is represented as an infinite series, which equals 1 for linear PBE, h is the grid spacing in Å, ϕ_0 is the electrostatic potential at the central grid point, q_0 is the charge at this grid point, and the summations are over the six neighboring grid points ($i = 1-6$) (Warwicker and Watson, 1982).

To assign charge density, dielectric constant, and ionic strength parameters to the grid points, we first define the molecule-solvent boundary, which is the smooth solvent-accessible surface (Lee and Richards, 1971), using a probe radius for the water molecule (typically 1.4 Å).

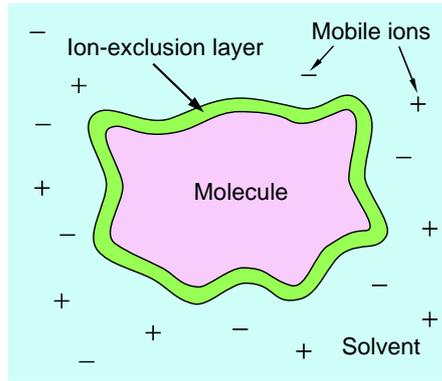


Fig. 1. 2D view of the electrostatic model (based on (Holst, 1993))

Figure 1 shows the two-dimensional view of the electrostatic model. The grid points within the molecule are normally assigned a uniform dielectric of $\epsilon_1 = 2$ as an approximation to the high-frequency dielectric constant of organic liquids. All grid points within the solvent region and the ion-exclusion layer are assigned a dielectric constant $\epsilon_2 = 80$. The modified Debye-Hückel parameter κ' is zero inside the molecule and at the ion-exclusion layer, where there are no mobile ions; κ' is non-zero in the solvent region. With these considerations we get following forms of the PBE:

$$\begin{cases} \epsilon_1 \nabla^2 \phi(\vec{r}) + 4\pi\rho(\vec{r}) = 0 & \text{inside molecule} \\ \epsilon_2 \nabla^2 \phi(\vec{r}) + 4\pi\rho(\vec{r}) = 0 & \text{at ion-exclusion layer} \\ \epsilon_2 \nabla^2 \phi(\vec{r}) - \kappa'^2(\vec{r}) \sinh[\phi(\vec{r})] + 4\pi\rho(\vec{r}) = 0 & \text{within solvent} \end{cases}$$

The numerical solvers for partial differential equations using FDM initialize the grid boundary values using various methods. The boundary values, once estimated, do not change. We use the analytical approximation obtained using Debye-Hückel potentials (Gilson et al., 1988) that are accurate if the solution grid is large enough relative to the size of the molecule.

5 Interface-Layer-focused FDM for PBE

From the discussion above, we find that the molecular surface boundary layer is critical to the accuracy of the FDM solution of the PBE. Not only are all the atomic charges within the molecule, but also there are large differences in the dielectric constants between the two regions separated by the boundary layer. As several biological processes occur at or near the molecular surface, a high accuracy for the solution to PBE close to this boundary layer is critical. Not coincidentally, the stability and accuracy of numerical methods also depend largely on the discretization of the grid in this region. To the best of our knowledge, no previous algorithm for solving PBE for molecules exists that builds the tetrahedral grid based on the molecular surface at the solvent-solute boundary. In this paper we present a new algorithm to solve the PBE more efficiently by building an adaptive tetrahedral space-decomposition about the molecular surface. The main idea is to give higher priority and resolution to the boundary region, and lower priority and resolution to other non-critical regions. We also adjust the grid density to be close to uniform in each region.

5.1 Analytical Molecular Surface

Previous methods to solve PBE approximate the molecular surface after building a 3D grid around the molecule. For each grid point, a binary marker indicates whether it is inside the molecule or inside the solvent. The molecular surface is then defined as passing between those grid points that have dissimilar markers. With such methods the accuracy of the molecular surface is limited to the grid resolution; the actual molecular surface points do not in general coincide with the grid points.

Several analytical molecular surface algorithms have been published (Akkiraju and Edelsbrunner, 1996; Bajaj et al., 2003; Connolly, 1983; Klein et al., 1990; Sanner and Olson, 1997; Varshney et al., 1994). We analytically generate the molecular surface using the approach in (Varshney et al., 1994) and then incorporate it in the 3D grid used for the solution of the PBE. Guaranteeing that the grid points at the solvent-solute boundary are actually on the exact surface improves the accuracy and speed of the algorithm.

5.2 Distance-Field-based Tetrahedralization

The accuracy of the FDM solution to PBE depends on the ionic strength assignment, which is zero in the 2 Å ion-exclusion layer from the molecular surface, and constant outside. Therefore we need an accurate ionic-screening

surface that is 2 Å offset outwards from the molecular surface. Our tetrahedralization algorithm is based on the distance field from the molecular surface and can generate the ionic-screening surface as well as provide an adaptive space decomposition.

We use an odd/even scheme for splitting rectilinear and curvilinear grids into tetrahedra as done in (Max et al., 1990). We use a method similar to the one described in (Gibson, 1998) to build a signed-distance map of the space that measures the distance of each grid point to the molecular surface (points inside the molecule are assigned negative distances).

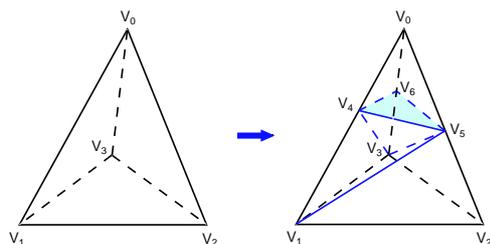


Fig. 2. Marching-tetrahedra-based Iso-surfaces and Tetrahedral grid refinement

Next, we generate a sequence of iso-surfaces from the distance map using a tetrahedral variant of the Marching Cubes algorithm (Lorensen and Cline, 1987; Nielson and Hamann, 1991). We use tetrahedra instead of cubes for simplicity and stability. We insert new grid points into the 3D grid such that they form surfaces at a fixed distance away from the molecular surface. One case of the marching tetrahedra is shown in Figure 2. Here the processing of tetrahedron $V_1V_2V_3V_0$ generates triangle $V_4V_5V_6$, which splits the original tetrahedron into four new tetrahedra: $V_4V_5V_6V_0$, $V_4V_3V_5V_6$, $V_1V_5V_3V_4$, and $V_1V_2V_3V_5$.

5.3 Adaptive Tetrahedralization

Multiresolution hierarchies have been well-researched for triangle meshes (Luebke et al., 2002) and tetrahedral meshes (Floriani, 2002). Such hierarchies have been used for domain decomposition in finite element analysis (Hebert, 1994; Maubach, 1995; Rivara and Levin, 1992), and used extensively in scientific visualization (De Floriani and M.Lee, 2003; Gerstner and Rumpf, 1999; Gregorski et al., 2002; Lee et al., 2001; Pascucci and Bajaj, 2000; Roxborough and Nielson, 2000; Trotts et al., 1999; Zhou et al., 1997). In our case, adaptive tetrahedral space decomposition is driven by the twin goals of accuracy and efficiency of the PBE solution. We desire a finer grid near the solvent-solute boundary for accuracy and a sparser grid elsewhere for efficiency.

We depict a 2D version of our adaptive tetrahedral grid in Figure 3(c). Here the thick red curve represents the molecular surface, while the thin red curves are

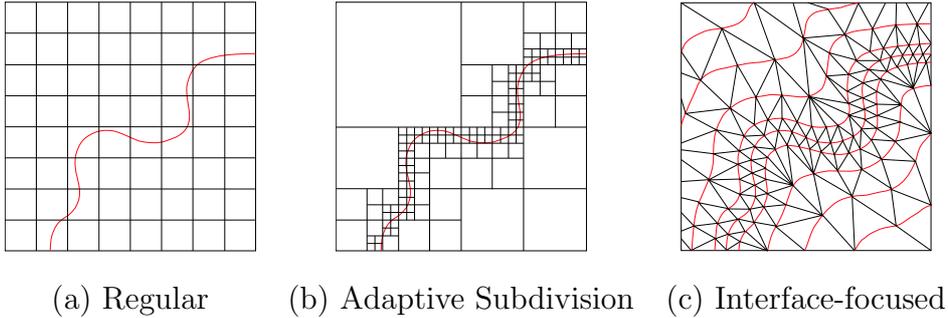


Fig. 3. Different Space Decompositions for PBE solvers

the iso-distance layers from the surface. Figure 3 clarifies the conceptual difference between our approach and the approaches using regular grids (Figure 3(a)) or octree-based space sub-division scheme (Figure 3(b)). The advantage of our approach is that we refine the grid directly on or close to the molecular surface. Regular grids are not adaptive and hence suffer from low accuracy, or high computational costs, or both. The adaptive octree-based subdivision scheme in most cases generates an excessively fine grid around the molecular surface to approximate it well. Our approach can adjust the resolution progressively and seamlessly based on the distance from the molecular surface.

We achieve adaptive tetrahedral decomposition by using edge collapses. Multiresolution tetrahedral grid hierarchies have been built using bottom-up edge-collapses (Cignoni et al., 2004; Trotts et al., 1999) or top-down longest edge bisections (Gregorski et al., 2002). An edge collapse will decrease the triangle count on the iso-surfaces, as well as the tetrahedra count. We use a half-edge collapse scheme so that an edge will collapse to one of its vertices and no new vertices need to be generated. Each edge collapse decreases the vertex count by one, and decreases the triangle and tetrahedron count based on its local connectivity. As an example, Figure 4 shows the collapse of edge E that results in decimation of tetrahedron $V_0V_1V_2V_5$ and vertex V_2 , while the tetrahedron $V_0V_2V_3V_5$ is changed to $V_0V_1V_3V_5$. The finite difference method (FDM) computes the value at a grid point from the values at its neighboring points (Section 4). During volume simplification we have to be careful not to simplify the volume into a state in which some grid points lose some of their neighbors necessary for FDM.

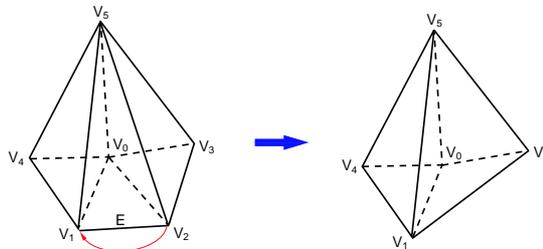


Fig. 4. Edge collapse for tetrahedral decimation

We also carefully avoid generating tetrahedra with negative volume (i.e., tetrahedra with a wrong orientation) and flipped triangles with reversed normals during the simplification process. In Figure 5, we show a 2D projection of one of these cases. Here the collapse of edge E results in grid point V_0 losing one of the six neighbors necessary for FDM. To avoid this we check and invalidate edge collapses that result in a new tetrahedron with three collinear vertices. In Figure 5, we get a new tetrahedron with collinear vertices V_0 , V_3 , and V_1 and therefore the edge E 's collapse should not be allowed.

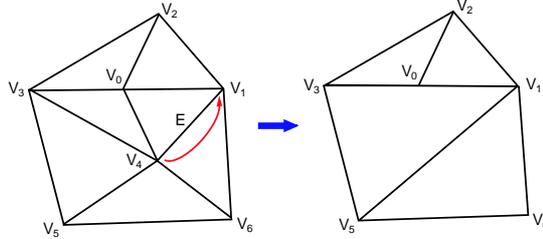


Fig. 5. An invalid edge collapse may lose a neighbor necessary for FDM solver

Another constraint for edge collapses is to preserve the spatial grid's outer (volumetric) boundary. We do this to ensure that the total volume of all the tetrahedra in the grid does not change as a result of the edge-collapse-based simplification. If a candidate edge for collapse has one of its vertices on the grid boundary we collapse the edge to the boundary vertex. If the candidate edge for collapse has both of its vertices on the grid boundary we carry out the collapse only if it will not result in a change in the total volume of all the affected tetrahedra.

5.4 Derivative Computation for Irregular Grids

The FDM solver for the PBE has to compute first and second derivatives of the 3D potential field at each grid point. The derivatives can be computed for regular grids by taking the finite differences between the potential value at each grid point with values at their six axis-aligned neighboring grid points. The regular structure of the regular grids makes this procedure straight forward. For irregular grids the situation is complicated by the fact that not only do the distances between grid points vary, but also the neighboring points are rarely axis-aligned.

We compute the derivatives of the potential at a grid point i by using the values at the vertices of the tetrahedra that are adjacent to i and include the principal axes from point i . As an example, consider the derivatives along x -axis for point V_0 . First, we find the two tetrahedra that share the vertex V_0 extend towards positive and negative x -axis from point V_0 as shown in Figure 6. This can be done by a simple orientation test.

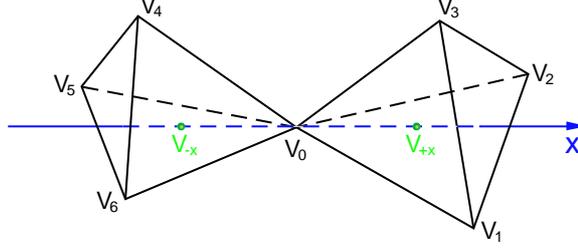


Fig. 6. Neighboring tetrahedra of V_0 along x -axis

After we identify the two tetrahedra we locate points V_{-x} and V_{+x} , that are equidistant from V_0 and along the x -axis. Let the distances between V_{-x} and V_0 and between V_{+x} and V_0 be h , then the second-order derivative of potential at V_0 can be approximated by:

$$\left. \frac{\partial}{\partial x} \left[\epsilon(\vec{r}) \frac{\partial}{\partial x} \phi(\vec{r}) \right] \right|_{V_0} \approx \frac{\epsilon(V_{+x})\phi'_x(V_{+x}) - \epsilon(V_{-x})\phi'_x(V_{-x})}{2h}$$

The first partial derivative of potential along x at V_{-x} and V_{+x} can be estimated by using a Taylor series expansion. We express potential values at vertices of each tetrahedron in terms of the value and derivatives at point V_{-x} and V_{+x} . As an example, $\phi'_x(V_{+x})$ can be estimated by solving the following system of four linear equations in four unknowns ($\phi(V_{+x}), \phi'_x(V_{+x}), \phi'_y(V_{+x}), \phi'_z(V_{+x})$):

$$\begin{cases} \phi(V_0) = \phi(V_{+x}) + \phi'_x(V_{+x})(-h) \\ \phi(V_1) = \phi(V_{+x}) + \phi'_x(V_{+x})\Delta x_1 + \phi'_y(V_{+x})\Delta y_1 + \phi'_z(V_{+x})\Delta z_1 \\ \phi(V_2) = \phi(V_{+x}) + \phi'_x(V_{+x})\Delta x_2 + \phi'_y(V_{+x})\Delta y_2 + \phi'_z(V_{+x})\Delta z_2 \\ \phi(V_3) = \phi(V_{+x}) + \phi'_x(V_{+x})\Delta x_3 + \phi'_y(V_{+x})\Delta y_3 + \phi'_z(V_{+x})\Delta z_3 \end{cases}$$

where $[\Delta x_i, \Delta y_i, \Delta z_i], i = 1, 2, 3$, is the vector difference between 3D positional vector of point V_i and V_{+x} : $[\Delta x_i, \Delta y_i, \Delta z_i] = \vec{r}(V_i) - \vec{r}(V_{+x})$.

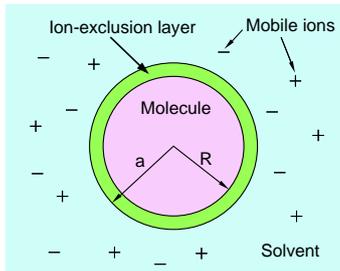
The above equations can be solved analytically. With such a solution we can express the second-derivative of the potential at V_0 along x axis as the linear-weighted sum of the potentials at V_0 to V_6 . We similarly compute the derivatives along y and z axes. This method of computing the derivatives for irregular meshes will have the same degree of accuracy as the method used for regular grids because both use the first-order Taylor series expansion to connect the values at neighboring points.

6 Results and Discussion

In this section we discuss the results obtained using our algorithm. We have used a 2GHz Pentium 4 PC running Windows 2000 with a nVIDIA GeForce3 graphics card. We present results on an analytical solvable case and compare them with the results by the well-known DelPhi system for computing molecular electrostatics. Our results show clear advantages of our algorithm over the standard DelPhi algorithm. We achieve better accuracy with less computation time. We then show our results on some real molecular datasets. We display the results by color coding smooth molecular surfaces as well as by direct volume rendering.

6.1 Results and Comparisons on an Analytical Solvable Case

Normally it is difficult (or impossible) to obtain analytical solutions to the PBE. In some special cases we may have analytical solutions to the linearized PBE. One example is that of a spherical molecule, with total charge q uniformly distributed on the surface, immersed in a solvent containing mobile ions, as shown in Figure 7.



(a)

The analytical solution here is (Holst, 1993):

$$\begin{cases} \phi(r) = \frac{q}{\epsilon_2 R} \left(1 - \frac{R\kappa}{1+\kappa a}\right) & \text{inside molecule} \\ \phi(r) = \frac{q}{\epsilon_2 r} \left(1 - \frac{r\kappa}{1+\kappa a}\right) & \text{ion-exclusion layer} \\ \phi(r) = \frac{qe^{\kappa a}}{\epsilon_2(1+\kappa a)} \cdot \frac{e^{-\kappa r}}{r} & \text{inside solvent} \end{cases}$$

(b)

Fig. 7. An analytically solvable case of PBE

We have tested the analytical solvable case using our algorithm and the DelPhi (V.4) program. The results are summarized in Table 1. Here we have used a spherical surface charge with a diameter of 27 Å and a positive charge of 20e (where e is the charge of an electron). The sphere is immersed in a cubic solvent volume of each side 66 Å.

The average error in Table 1 is the average of the relative error over all grid points. Peak-signal-to-noise-ratio (PSNR) is $20 \log_{10}(\frac{\text{signal energy}}{\text{noise energy}})$. The *signal energy* is defined as the sum of the squares of the potential values over all grid points. The *noise energy* is defined as the sum of the squares of the errors over all grid points. PBE time is the time for solving linear PBE on the grid. One can see from Table 1 the advantages of our method. To get the same accuracy, our method needs only 27K points instead of several million

	DelPhi			Our Method
Grid size	67 ³	133 ³	199 ³	N/A
# of pts	300,763	2,352,637	7,880,599	26,987
PSNR	8.17	19.1	25.1	27.7
Avg. error	30.88%	17.91%	13.27%	15.98%
PBE Time	0.31 sec	4.50 sec	20.09 sec	0.25 sec

Table 1
Comparison of our method with DelPhi

needed by DelPhi, and takes only 0.25 seconds to converge, compared with several seconds by DelPhi. For about the same amount of time, our method is much more accurate than DelPhi, e.g., 15.98% instead of 30.88% error. Our algorithm with 27K points has even higher PSNR than DelPhi with about 8M points.

6.2 Results on Molecules

We now show our results for some molecules. The first dataset is superoxide dismutase (SOD) enzyme, which consists of 2196 atoms. Our second dataset is a channel on the outer membrane of the Escherichia coli (Ecoli) bacterium molecule (Sukharev et al., 2001), which consists of 10585 atoms. The results are shown in Figures 8 and 9. We have used about 2000 directions to precompute the accumulated transparencies for each contributing voxel.

Figures 8(a) and 9(a) display the smooth molecular surfaces of SOD and Ecoli membrane channel using the SURF algorithm (Varshney et al., 1994). Figures 8(b) and Figure 9(b) display the electrostatic potential on the surfaces, with red for negative and blue for positive potential; both use the potential information to modulate lighting color with grey for neutral potential. Figure 8(c) and 9(c) show the volume rendered 3D potential field, from the viewer up to the molecular surface. Figure 8(d), (e), and (f) are closeups of Figure 8(a), (b), and (c) respectively. Electrostatic potential is traditionally displayed on molecular surfaces (Rocchia et al., 2001). We find it is more informative to use direct volume rendering for the potential between the viewer and the molecular surface. This gives more information compared to the on-surface only potential display. Comparing Figure 8(e) with 8(f), one can clearly see that in front of the negative on-surface potential in the central region, there is a sizable positive potential region. It would have been hard to use the traditional electrostatic potential display methods such as color-coded molecular surfaces or electrostatic iso-potential surfaces to convey the same amount of visual information.

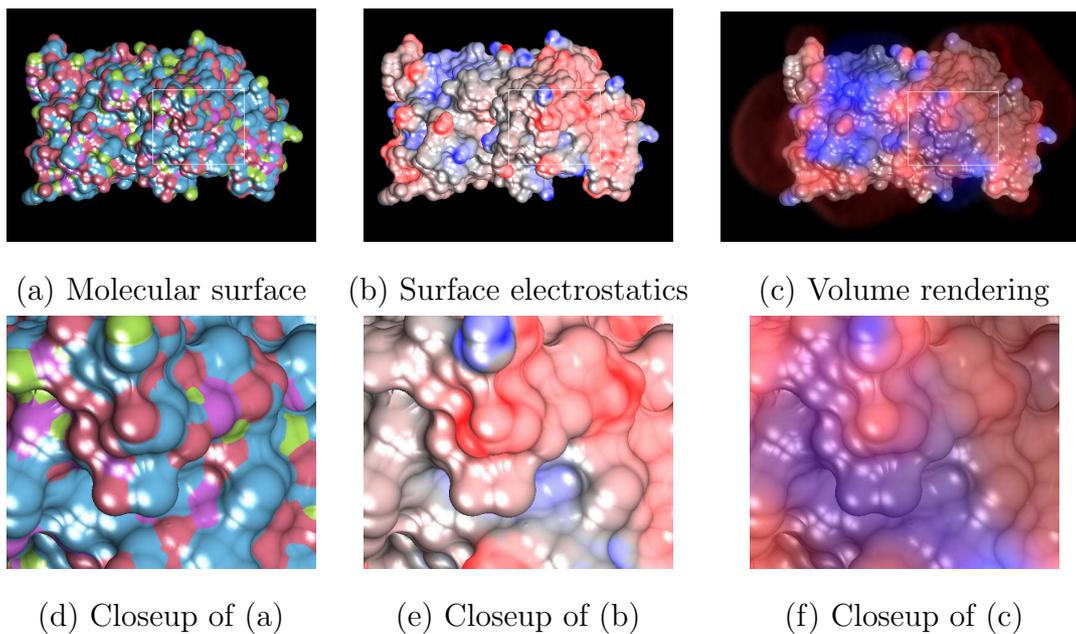


Fig. 8. Electrostatics on SuperOxide Dismutase (SOD) dataset (red is for negative potential, and blue is for positive potential)

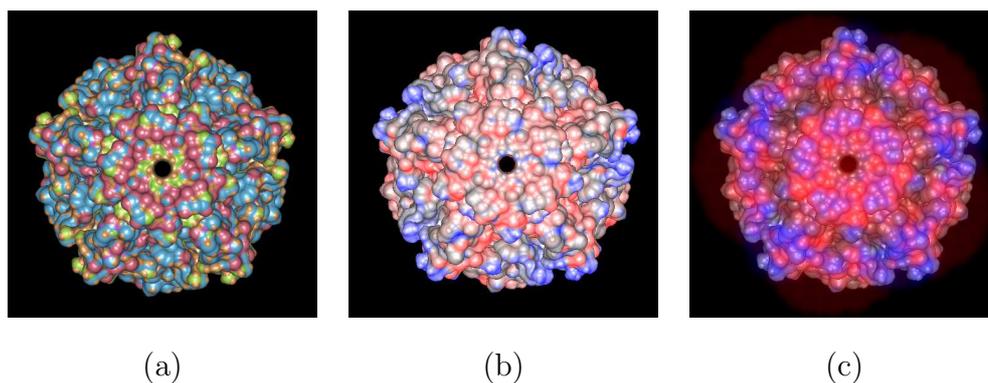


Fig. 9. Ecoli's mechanosensitive channel shown as (a) its molecular surface, (b) its surface color-coded by electrostatics potential (red for negative potential, and blue for positive potential), and (c) volume rendering of the electrostatics field around the channel surface

7 Conclusions

We have presented new algorithms for efficiently computing electrostatic potentials for large molecular datasets. Our methods give higher priority and resolution to the solution-sensitive region to improve the accuracy and accelerate convergence rates. We build a 3D tetrahedral partition of the space directly from an analytically constructed interface layer. We also provide an algorithm to control the density and uniformity of the grid by using an edge collapse scheme. Compared with the state-of-the-art method using analytically-solvable testing case, our method is faster and more accurate. The advantage

of our algorithm in solving partial differential equations directly from the geometrical point of view gives it a broad range of possible applications in other application domains.

With the advances presented in this paper our electrostatic computation methods are now fast enough to be used in interactive molecular docking experiments with interleaved computation and visualization of large molecules. We plan to explore this in the near future.

Acknowledgements

We will like to acknowledge several valuable discussions on this topic with Dianne O’Leary, Leila De Floriani, and Sergei Sukharev. This work has been supported in part by the NSF grants: IIS-00-81847, ACR-98-12572/02-96148, CCF 04-29753, and CNS 04-03313.

References

- Akkiraju, N., Edelsbrunner, H., 1996. Triangulating the surface of a molecule. *Discrete Appl. Math.* 71, 5–22.
- Bajaj, C., Pascucci, V., Shamir, A., Holt, R., Netravali, A., 2003. Dynamic maintenance and visualization of molecular surfaces. *Discrete Appl. Math.* 127, 23–51.
- Baker, N., Holst, M., Wang, F., 2000. Adaptive multilevel finite element solution of the poisson-boltzmann equation ii: refinement at solvent accessible surfaces in biomolecular systems. *J. Comput. Chem.* 21, 1343–1352.
- Cignoni, P., De Floriani, L., Magillo, P., Puppo, E., Scopigno, R., 2004. Selective refinement queries for volume visualization of unstructured tetrahedral meshes. *IEEE Transactions on Visualization and Computer Graphics* 10(1), 29–45.
- Connolly, M. L., 1983. Solvent-accessible surfaces of proteins and nucleic acids. *Science* 221, 709–713.
- De Floriani, L., M.Lee, 2003. Selective refinement on nested tetrahedral meshes. In: G.Brunett, B., H.Mueller (Eds.), *Geometric Modelling for Scientific Visualization*. Springer Verlag.
- Floriani, L., Oct. 2002. Multi-resolution data structures. In: *Multi-resolution modeling, visualization and compression of volumetric data (Tutorial Notes)*, IEEE Visualization 2003. Seattle.
- G. Allen (editor), 1999. *Protein : a comprehensive treatise*, vol 2, 61 – 97. JAI Press.
- Gerstner, T., Rumpf, M., 1999. Multiresolutional parallel isosurface extraction

- based on tetrahedral bisection. In: Proceedings 1999 Symposium on Volume Visualization. IEEE Computer Society, pp. 267 – 278.
- Gibson, S. F., 1998. Using distance maps for accurate surface representation in sampled volumes. In: IEEE Symposium on Volume Visualization. pp. 23–30.
- Gilson, M. K., Sharp, K. A., Honig, B., 1988. Calculating electrostatic interactions in biomolecules: method and error assessment. *J. Comp. Chem.* 9, 327 – 335.
- Gregorski, B., Duchaineau, M., Lindstrom, P., Pascucci, V., Joy, K., Oct. 2002. Interactive view-dependent rendering of large isosurfaces. In: Proceedings IEEE Visualization 2002. IEEE Computer Society, pp. 475 – 484.
- Hebert, D. J., May 1994. Symbolic local refinement of tetrahedral grids. *Journal of Symbolic Computation* 17 (5), 457–472.
- Holst, M., Baker, N., Wang, F., 2000. Adaptive multilevel finite element solution of the poisson-boltzmann equation i: algorithms and examples. *J. Comput. Chem.* 21, 1319–1342.
- Holst, M. J., 1993. Multilevel methods for the poisson-boltzmann equation. Ph.D. thesis, Numerical Computing Group, University of Illinois at Urbana-Champaign.
- Honig, B., Nicholls, A., 1995. Classical electrostatics in biology and chemistry. *Science* 268, 1144 – 1149.
- Humphrey, W., Dalke, A., Schulten, K., 1996. VMD—visual molecular dynamics. *J. Mol. Graphics* 14, 33–38.
- Jackson, J. D., 1975. *Classical Electrodynamics*, 2nd Edition. John Wiley and Sons.
- Klein, T. E., Huang, C. C., Pettersen, E. F., Couch, G. S., Ferrin, T. E., Langridge, R., 1990. A real-time malleable molecular surface. *J. Mol. Graphics* 8(1), 16–24 and 26–27.
- Kreylos, O., Max, N., Hamann, B., Crivelli, S. N., Bethel, E. W., 2003. Interactive protein manipulation. In: IEEE Visualization '03. pp. 581–588.
- Leach, A. R., 2001. *Molecular Modelling: Principles and Applications*, 2nd Edition. Prentice Hall.
- Lee, B., Richards, F. M., 1971. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* 55, 379 – 400.
- Lee, M., De Floriani, L., Samet, H., May 2001. Constant-time neighbor finding in hierarchical tetrahedral meshes. In: Proceedings International Conference on Shape Modeling & Applications. Genova, Italy, pp. 286–295.
- Lorensen, W., Cline, H., Jul. 1987. Marching cubes: a high resolution 3D surface construction algorithm. *Computer Graphics (SIGGRAPH '87 Proceedings)* 21 (4), 163–169.
- Luebke, D., Reddy, M., Cohen, J., Varshney, A., Watson, B., Huebner, R., 2002. *Level of Detail for 3D Graphics*. Morgan-Kaufmann, San Francisco.
- Maubach, J. M., Jan. 1995. Local bisection refinement for n -simplicial grids generated by reflection. *SIAM Journal on Scientific Computing* 16 (1), 210–227.

- Max, N., Hanrahan, P., Crawfis, R., 1990. Area and volume coherence for efficient visualization of 3D scalar function. San Diego Workshop on Volume Visualization, *Computer Graphics* 24 (5), 27–33.
- Nielson, G. M., Hamann, B., 1991. The asymptotic decider: Removing the ambiguity in marching cubes. In: *Visualization '91*. pp. 83–91.
- Oberoi, H., Allewell, N. M., 1993. Multigrid solution of the nonlinear Poisson-Boltzmann equation and calculation of titration curves. *Biophys. J.* 65, 48 – 55.
- Pascucci, V., Bajaj, C. L., Oct. 2000. Time critical isosurface refinement and smoothing. In: *Proceedings IEEE Symposium on Volume Visualization*. IEEE Computer Society, Salt Lake City, UT, pp. 33–42.
- Rivara, M., Levin, C., 1992. A 3D refinement algorithm for adaptive and multigrid techniques. *Communications in Applied Numerical Methods* 8, 281–290.
- Rocchia, W., Alexov, E., Honig, B., 2001. Extending the applicability of the nonlinear poisson-boltzmann equation: Multiple dielectric constants and multivalent ions. *J. Phys. Chem. B* 105, 6507–6514.
- Roxborough, T., Nielson, G., Oct. 2000. Tetrahedron-based, least-squares, progressive volume models with application to freehand ultrasound data. In: *Proceedings IEEE Visualization 2000*. IEEE Computer Society, pp. 93–100.
- Sanner, M. F., Olson, A. J., 1997. Real time surface reconstruction for moving molecular fragments. In: Altman, R. B., Dunker, A. K., Hunter, L., Klein, T. E. (Eds.), *Pacific Symposium on Biocomputing '97*. pp. 385–396.
- Sukharev, S., Durell, S. R., Guy, H. R., 2001. Structural models of the mscl gating mechanism. *J. Biophys* 81(2), 917–936.
- Tainer, J. A., Getzoff, E. D., Sayre, J., Olson, A. J., 1985. Modeling intermolecular interactions: topography, mobility and electrostatic recognition. *J. Mol. Graphics* 3, 103–105.
- Tanford, C., Kirkwood, J. G., 1957. Theory of protein titration curves. I. General equations for impenetrable spheres. *J. Am. Chem. Soc.* 79, 5333 – 5339.
- Trotts, I. J., Hamann, B., Joy, K. I., Jul./Sep. 1999. Simplification of tetrahedral meshes with error bounds. *IEEE Transactions on Visualization and Computer Graphics* 5 (3), 224–237.
- Varshney, A., Brooks, Jr., F. P., Wright, W. V., September 1994. Computing smooth molecular surfaces. *IEEE Computer Graphics & Applications* 15 (5), 19–25.
- Warwicker, J., Watson, H. C., 1982. Calculation of electrostatic potential in the active site cleft due to α -helix dipoles. *J. Mol. Biol.* 155, 53 – 62.
- Zhou, Y., Chen, B., Kaufman, A., Oct. 1997. Multiresolution tetrahedral framework for visualizing regular volume data. In: Yagel, R., Hagen, H. (Eds.), *Proceedings IEEE Visualization '97*. Phoenix, AZ, pp. 135–142.