

## Geometric modeling of subcellular structures, organelles, and multiprotein complexes

Xin Feng<sup>1</sup>, Kelin Xia<sup>2</sup>, Yiying Tong<sup>1,\*</sup>,† and Guo-Wei Wei<sup>2,3,4,\*</sup>,†

<sup>1</sup>*Department of Computer Science and Engineering, Michigan State University, MI 48824, USA*

<sup>2</sup>*Department of Mathematics, Michigan State University, MI 48824, USA*

<sup>3</sup>*Department of Electrical and Computer Engineering, Michigan State University, MI 48824, USA*

<sup>4</sup>*Department of Biochemistry and Molecular Biology, Michigan State University, MI 48824, USA*

### SUMMARY

Recently, the structure, function, stability, and dynamics of subcellular structures, organelles, and multiprotein complexes have emerged as a leading interest in structural biology. Geometric modeling not only provides visualizations of shapes for large biomolecular complexes but also fills the gap between structural information and theoretical modeling, and enables the understanding of function, stability, and dynamics. This paper introduces a suite of computational tools for volumetric data processing, information extraction, surface mesh rendering, geometric measurement, and curvature estimation of biomolecular complexes. Particular emphasis is given to the modeling of cryo-electron microscopy data. Lagrangian-triangle meshes are employed for the surface presentation. On the basis of this representation, algorithms are developed for surface area and surface-enclosed volume calculation, and curvature estimation. Methods for volumetric meshing have also been presented. Because the technological development in computer science and mathematics has led to multiple choices at each stage of the geometric modeling, we discuss the rationales in the design and selection of various algorithms. Analytical models are designed to test the computational accuracy and convergence of proposed algorithms. Finally, we select a set of six cryo-electron microscopy data representing typical subcellular complexes to demonstrate the efficacy of the proposed algorithms in handling biomolecular surfaces and explore their capability of geometric characterization of binding targets. This paper offers a comprehensive protocol for the geometric modeling of subcellular structures, organelles, and multiprotein complexes. Copyright © 2012 John Wiley & Sons, Ltd.

Received 16 August 2012; Revised 16 October 2012; Accepted 2 November 2012

**KEY WORDS:** macromolecules; geometric modeling; Laplace–Beltrami operator; high-order geometric PDEs; surface meshing; Gaussian curvature; mean curvature

### 1. INTRODUCTION

One of major features of biological sciences in the 21st century is their transition from an empirical, qualitative, and phenomenological discipline to a comprehensive, quantitative, and predictive one [1]. Indeed, theoretical description, mathematical modeling, and computer simulation of biological systems have made enormous contribution to the present understanding of biological sciences. The material basis and fundamental underpinning of modern biological sciences are biological macromolecules, especially proteins and nucleic acids, which coil into specific three-dimensional (3D) shapes and are able to carry out most of the functions of cells. The goal of theoretical description, mathematical modeling, and computer simulation is to understand the structure, function, dynamics, and transport of biological macromolecules. A prerequisite to theoretical description, mathematical

\*Correspondence to: Yiying Tong, Department of Computer Science and Engineering and Guo-Wei Wei, Department of Mathematics, Michigan State University, MI 48824, USA.

†E-mail: ytong@msu.edu; wei@math.msu.edu

modeling, and computer simulation of the structure, function, dynamics, and transport of biological macromolecules is the geometric modeling based on their 3D shapes. In addition to straightforward geometric visualization, geometric modeling bridges the gap between imaging and mathematical modeling such that the structural information can be integrated into mathematical models [2].

Macromolecular 3D shapes can be indirectly obtained from a number of experimental means, including macromolecular X-ray crystallography, nuclear magnetic resonance (NMR), electron paramagnetic resonance, cryo-electron microscopy (cryo-EM), multiangle light scattering, confocal laser-scanning microscopy, small angle scattering, ultra fast laser spectroscopy, and others. The main workhorses for single macromolecules are crystallography and NMR. The advanced X-ray crystallography technology is able to provide decisive structural information at Armstrong and sub-Armstrong resolutions, whereas NMR experiments often offer structural information under physiological conditions. Both X-ray crystallography and NMR are technologically relatively well developed, except for their applications in special tasks, such as the crystallization of membrane proteins. However, these approaches are not suitable for proteasomes, subcellular structures, organelles, cells, and tissues. Their study has become one of the major new trends in structural biology. Currently, one of the most powerful tools for subcellular structures, organelles, and large multiprotein complexes is cryo-EM [3].

Cryo-EM, or electron cryo-microscopy, is an increasingly popular transmission EM whose sample is bombarded by electron beams at cryogenic temperatures to improve the signal-to-noise ratio (SNR). Its working principle is based on the projection (thin film) specimen scans collected from many different directions around one or two axes and the creation of 3D images by using the Radon transform. Cryo-EM allows the imaging of specimens in their native environment and is capable of providing 3D mapping of entire cellular proteomes together with their detailed interactions at a nanometer resolution [4–7].

Because most biological specimens are extremely radiation sensitive, they can only sustain the illumination of a limited electron dose. As a result, cryo-EM images are inevitably of low SNR and lead to limited resolutions [3]. In practice, cryo-EM maps often do not contain adequate information to offer unambiguous atomic-scale structural reconstruction of biological specimens. Additional information obtained from other techniques, such as crystallography, NMR, and computer simulation, is utilized to interpret the cryo-EM maps. However, the resolution of cryo-EM maps has been improved dramatically in the past few years, thanks to the technical advances in experimental hardware, noise reduction, and image segmentation techniques. By further taking the advantage of symmetric averaging, many cryo-EM-based virus structures have already achieved a resolution that can be interpreted in terms of an atomic model. Therefore, it is time to utilize cryo-EM images for molecular and atomic-scale mathematical modeling and computer simulation of subcellular structures, organelles, and large multiprotein complexes.

Most 3D imaging data obtained from cryo-EM and many other tomographic modalities are currently presented in a digital format, as a volumetric density distribution, where each Cartesian grid point is assigned with a scalar value associated with the local electron scattering power. For the purpose of visualization, one needs to convert them into images in the form of a series of two-dimensional (2D) images, rendering of the 3D shapes generated by isosurface extraction, or direct volumetric rendering. For the purpose of geometric analysis, structural features in the complex settings of cellular landscapes are further characterized in terms of surface areas, surface-enclosed volumes, and Gaussian and mean curvatures. For the purpose of mathematical modeling and computation, the resulting 3D geometric shape is to be further described in either the Lagrangian representation or the Eulerian representation. The Lagrangian representation is a basis for the Lagrangian formulation of the biological evolution, in which surface elements are directly evolved according to a governing equation or a set of rules [2, 8]. Similarly, the Eulerian representation facilitates the Eulerian formulation of the biological dynamics, in which the biological shape is embedded in a hypersurface function, or a level set function, and such a function is then evolved under prescribed physical and/or biological principles [9–11]. Both Lagrangian and Eulerian approaches have their own pros and cons, and are very useful in mathematical modeling and computation. In the present work, we focus our studies on the Lagrangian representation of 3D imaging data.

Currently, the SNR of 3D imaging data for subcellular structures, organelles, and large multi-protein complexes is typically in the neighborhood of 0.01 [3]. To make the situation worse, the image contrast, which depends on the difference between electron scattering cross sections of cellular components, is also very low in most biological systems. Consequently, appropriate noise reduction is an indispensable process in the structure reconstruction from 3D imaging data. To improve the SNR and image contrast of cryo-EM data, researchers have employed a wide variety of denoising schemes, including wavelet transform techniques [12], nonlinear anisotropic diffusions [13, 14] or Beltrami flow [15], bilateral filter [16–18], and iterative median filtering [19]. Despite much effort, noise reduction of cryo-EM data remains a challenging task and is far from adequate, due to the extremely low SNR and other technical complications [3]. Innovative mathematical approaches are necessary to further tackle this problem [20–22].

Geometric flows, in which the flow motion is governed or influenced by geometric properties, such as curvatures, have become an established approach to image analysis and surface generation in the past few years. Particularly, mean curvature flows have been a popular subject in applied mathematics for image analysis, material design [23–25], and surface processing [26, 27]. The first use of PDEs for image analysis dates back to 1983 [28]. Witkin noticed that the evolution of an image under a diffusion operator is formally equivalent to the standard Gaussian low-pass filter for image denoising [28]. Perona and Malik introduced an anisotropic diffusion equation [29] to protect image edges during the diffusion process. The Perona–Malik equation stimulated much interest in applied mathematics [29–33]. Over the past two decades, many related mathematical techniques, such as the level set formalism devised by Osher and Sethian [25, 34], Mumford–Shah variational functional [35], and the total variation (TV) minimization [36], have been widely used for image analysis [37–41].

To improve the efficiency of noise removing, Wei introduced the first family of arbitrarily high-order nonlinear PDEs for image denoising and restoration in 1999 [31]. Many fourth-order evolution equations were introduced in the literature for image analysis [32, 42–45]. These equations were proposed either as a high-order generalization of the Perona–Malik equation [31, 46] or as an extension of the TV formulation [32, 42–44]. The essential assumption in these high-order evolution equations is that high-order diffusion operators are able to remove high-frequency components more efficiently. High-order geometric PDEs have been widely applied to image and surface analysis [31, 32, 42–48]. Because of the stiffness of high-order nonlinear PDEs, computational techniques for solving higher-order geometric PDEs are of great importance. For instance, alternating direction implicit schemes are developed in the literature for integrating high-order nonlinear PDEs [46, 49].

Image-processing PDEs of the Perona–Malik type and the TV type are mostly designed to function as nonlinear low-pass filters. In 2002, Wei and Jia [33] introduced coupled nonlinear PDEs to behave as high-pass filters. These coupled nonlinear PDEs are demonstrated for image edge detection. The essential idea behind these PDE-based high-pass filters is that when two Perona–Malik type of PDEs evolve at dramatically different speeds, the difference of their solutions gives rise to image edges. This follows from the fact that the difference between an all-pass filter (i.e., identity operator) and a low-pass one is a high-pass filter [33]. The speeds of evolution in these equations are controlled by the appropriate selection of the diffusion coefficients. These PDE-based edge detectors have been shown to work extremely well for images with a large amount of textures [33, 50]. Most recently, the PDE transform is introduced for functional mode decomposition, [51–53] based on arbitrarily high-order PDE high-pass filters. Such an approach has significantly extended the utility of PDEs for image, surface, and data analysis. Similar to wavelet transform, the PDE transform has controllable time-frequency location and perfect reconstruction. The PDE transform has found its success in molecular surface generation of proteins [54].

The use of curvature-controlled PDEs for biomolecular surface construction was initiated in 2005 [55]. Atomic coordinate information of a protein is embedded in 3D Eulerian grids to undergo geometric flow evolution before the protein surface is extracted via the marching cubes method from a level-set type of hypersurface function. This approach was combined with a variational procedure to generate the first variational biomolecular surface model, the minimal molecular surface, for proteins [9, 26]. Molecular interactions were further incorporated in this approach to develop potential and curvature-driven geometric flows for the construction of biomolecular surfaces [46].

Recently, many variational multiscale models have been introduced on the basis of the geometric flow separation of solvent and solute domains [1, 10, 11, 56].

After the surface construction, a further issue in geometric modeling is the surface and volumetric (i.e., boundary and interior) meshing [41, 57, 58]. There are a wide variety of methods that can be used for this purpose. Mesh generation is one of the most important aspects of continuum mechanical analysis [59–63]. Numerous elegant methods, such as the probabilistic methods for centroidal Voronoi tessellations [64, 65], the optimal Delaunay triangulation and graph cut-based variational surface reconstruction [66], and other surface remeshing enhancement methods and technologies [41, 67–70], have been developed for surface reconstruction or surface remeshing during the past two decades. In general, high-quality triangle surface meshes must be low noise, low memory cost, near  $60^\circ$  for majority of element angles and aligned with the physical features. Yu *et al.* discussed the use of adaptive feature-preserving methods for biomolecular surface meshing [2, 71]. They used the constrained Delaunay triangulation implemented in TetGen [72] for volumetric meshing [2].

Curvature is a measure of how much a curve deviates from being straight or a surface from being flat [73]. Curvature has been used to analyze the stereospecificity of molecular surfaces [74]. Curvature has also been applied for finite element analysis crank mechanisms [75]. The essential idea is that geometries of binding partners are locally complementary to each other at the binding site(s). Curvatures are also used as a geometric descriptor to characterize the shape of known protein binding sites so as to identify matching site(s) in other proteins and ligands. However, in real cases, the effect of stereospecificity may be offset by hydrogen bond, polarization, electrostatics, solvation, and allosteric modulation.

The objective of the present work is to explore the efficient computational methods for the geometric modeling of subcellular structures, organelles, and large multiprotein complexes. Specifically, we study the reconstruction of biological structures from noisy 3D imaging data, examine the geometric representation of complex biological shapes, provide accurate calculation of surface areas and surface-enclosed volumes, and investigate the computational algorithm and surface mapping of Gaussian and mean curvatures of macromolecules. Most geometric modeling is carried out in the Lagrangian representation with triangle meshes on the surface.

The rest of this paper is organized as follows. Section 2 is devoted to computational methods and numerical algorithms for geometric modeling. We give a brief description of high-order geometric PDEs and nonlinear PDE-based high-pass filters. Different surface extraction schemes are discussed. Numerical algorithms for calculating surface areas and surface-enclosed volumes are given in the Lagrangian representation. We introduce the state of the art techniques for volumetric meshing of subcellular structures, organelles, and large multiprotein complexes. Efficient schemes for computing Gaussian and mean curvatures are provided. In Section 3, we carry out extensive numerical experiments to validate the proposed methods, algorithms, and schemes. We design analytical cases to test accuracy and convergent order of the proposed algorithms for area, volume, and curvature calculations. Second-order convergence is found in these schemes. Finally, we apply the proposed methods to six biomolecular examples. Our results demonstrate the usefulness, robustness, and efficiency of the proposed approaches. This paper ends with a conclusion in Section 4.

## 2. METHODS

This section provides a variety of mathematical and computational methods for geometric modeling. The goal here is to introduce a repertoire of appropriate computational tools for the applications involving volumetric data and the shapes defined in cryo-EM datasets.

### 2.1. High-order geometric flows

Geometric flows, such as the Laplace–Beltrami flow, play a significant role in image analysis. An important aspect in the geometric flow development is the use of high-order geometric PDEs for image processing or surface analysis. Willmore flow, proposed in 1920s, is a fourth-order geometric

PDE that locally minimizes the difference between two principal curvatures (see detailed description on principal curvatures in Section 2.5.1). Therefore, the Willmore flow prefers spherical shapes, which may be undesirable in general applications. Motivated by the hyperdiffusion in the pattern formation in alloys, glasses, polymer, combustion, and biological systems, Wei introduced the first family of arbitrarily high-order geometric PDEs for edge-preserving image restoration in 1999, using Fick's law [31]

$$\frac{\partial u(\mathbf{r}, t)}{\partial t} = - \sum_q \nabla \cdot \mathbf{j}_q + e(u(\mathbf{r}, t), |\nabla u(\mathbf{r}, t)|, t), \quad q = 0, 1, 2, \dots \quad (1)$$

where the nonlinear hyperflux is given by

$$\mathbf{j}_q = -d_q(u(\mathbf{r}, t), |\nabla u(\mathbf{r}, t)|, t) \nabla \nabla^{2q} u(\mathbf{r}, t), \quad q = 0, 1, 2, \dots \quad (2)$$

where  $\mathbf{r} \in \mathbb{R}^n$ ,  $\nabla = \partial/\partial \mathbf{r}$ ,  $u(\mathbf{r}, t)$  is the processed image function,  $d_q(u(\mathbf{r}, t), |\nabla u(\mathbf{r}, t)|, t)$  are edge sensitive diffusion coefficients, and  $e(u(\mathbf{r}, t), |\nabla u(\mathbf{r}, t)|, t)$  is a nonlinear operator. Equation (1) is subject to the initial image data  $u(\mathbf{r}, 0) = X(\mathbf{r})$  and appropriate boundary conditions. The essential idea of Equation (1) is to accelerate the noise removal in the Perona–Malik equation [29] by higher-order derivatives, which is more efficient in noise dissipation. As a generalization of the Perona–Malik equation, the hyperdiffusion coefficients  $d_q(u, |\nabla u|, t)$  in Equation (2) can also be chosen as the Gaussian form

$$d_q(u(\mathbf{r}, t), |\nabla u(\mathbf{r}, t)|, t) = d_{q0} \exp \left[ -\frac{|\nabla u|^2}{2\sigma_q^2} \right], \quad (3)$$

where the values of constant  $d_{q0}$  depend on the noise level, and  $\sigma_0$  and  $\sigma_1$  were chosen as the local statistical variance of  $u$  and  $\nabla u$

$$\sigma_q^2(\mathbf{r}) = \overline{|\nabla^q u - \overline{\nabla^q u}|^2} \quad (q = 0, 1). \quad (4)$$

The notation  $\overline{Y(\mathbf{r})}$  denotes the local average of  $Y(\mathbf{r})$  centered at position  $\mathbf{r}$ . The measure based on the local statistical variance is important for discriminating image features from noise. As a result, one can bypass the image preprocessing, that is, the convolution of the noise image with a smooth mask in the application of the PDE operator to noisy images. High-order geometric PDEs have found many practical applications [31, 44, 76, 77]. Arbitrarily high-order geometric PDEs are modified for molecular surface formation and evolution [46]

$$\frac{\partial S}{\partial t} = (-1)^q \sqrt{g(|\nabla \nabla^{2q} S|)} \nabla \cdot \left( \frac{\nabla(\nabla^{2q} S)}{\sqrt{g(|\nabla \nabla^{2q} S|)}} \right) + P(S, |\nabla S|), \quad (5)$$

where  $S$  is the hypersurface function,  $g(|\nabla \nabla^{2q} S|) = 1 + |\nabla \nabla^{2q} S|^2$  is the generalized Gram determinant, and  $P$  is a generalized potential term, including microscopic interactions in biomolecular surface construction. When  $q = 0$  and  $P = 0$ , Equation (5) recovers the mean curvature flow used in our earlier construction of minimal molecular surfaces [9]. It reproduces the surface diffusion flow [46] when  $q = 1$  and  $P = 0$ . It has been shown that surface generated with the fourth-order geometric PDE demonstrates a morphology distinguished from that obtained with the mean curvature flow or the Laplace–Beltrami flow [46].

## 2.2. Nonlinear PDE-based high-pass filters

Unfortunately, the studies of geometric flows have been essentially limited to the construction of nonlinear PDE-based low-pass filters. From the point of view of image and signal processing, low-pass filtering is just one specific type of operations, and other filters, such as high-pass filters and band-pass filters, are equally important. An exception is the nonlinear PDE-based high-pass filters introduced by Wei and Jia [33] for image edge detection in 2002,

$$u_t(\mathbf{r}, t) = F_1(u, \nabla u, \nabla^2 u, \dots) + \epsilon_u(v - u) \quad (6)$$

$$v_t(\mathbf{r}, t) = F_2(v, \nabla v, \nabla^2 v, \dots) + \epsilon_v(u - v) \quad (7)$$

where  $u(\mathbf{r}, t)$  and  $v(\mathbf{r}, t)$  are scalar fields,  $\epsilon_u$  and  $\epsilon_v$  are coupling strengths. Here,  $F_1$  and  $F_2$  are general nonlinear diffusion operators and can be chosen as the Perona–Malik operator  $F_1 = \nabla \cdot d_1(|\nabla u|)\nabla$  and  $F_2 = \nabla \cdot d_2(|\nabla v|)\nabla$ . The initial values for both nonlinear evolution equations are chosen to be the same image, that is,  $u(\mathbf{r}, 0) = v(\mathbf{r}, 0) = X(\mathbf{r})$ . As a nonlinear dynamic system, the time evolution of Equations (6) and (7) will eventually lead to a synchronization in the solution for positive nonzero coupling coefficients. For the purpose of image processing, Equations (6) and (7) are designed to evolve at dramatically different time scales, for example, the coefficients  $d_1$  and  $d_2$  are chosen as the Gaussian form in Equation (3) with  $d_{20} \gg d_{10} \geq 0$ . After finite time evolution, the image edges are obtained as the difference [33]

$$w(\mathbf{r}, t) = u(\mathbf{r}, t) - v(\mathbf{r}, t). \quad (8)$$

It was shown that Equation (8) behaves like a band-pass filter when  $d_{20} \gg d_{10} \sim 0$ . The essence of this approach is that when two coupled evolution PDEs are evolving at dramatically different speeds, the difference of two low-pass PDE operators gives rise to a band-pass or high-pass filter. The coupling terms play the role of relative fidelity and balance the disparity of two images. It has been shown that nonlinear PDE-based high-pass filters work extremely well for images with a large number of textures and outperform classical Sobel, Prewitt, and Canny operators [33, 50].

### 2.3. Surface extraction and meshing

Both the 3D imaging data from cryo-EM and the result of the aforementioned Eulerian geometric flows or PDE-based nonlinear filtering are given as functions sampled on Eulerian grids. For the subsequent use in FEMs or Lagrangian geometry processing, the Lagrangian representation is employed, which is often much more efficient than the Eulerian representation due to better adaptivity of the irregular sampling and the reduction from 3D into curved 2D representation. In most geometry processing approaches,  $h$ -refinement (more segments) is preferred over  $p$ -refinement (degree of the polynomial in each segment), because the modern computer architectures can handle a large number of simple objects more efficiently than a small number of complex objects representing the same surface [78]. In the following, we use the extremal  $C^0$  case, that is, piecewise flat surface meshes, the de facto standard data structure in current geometry processing.

The conversion from 3D image data to 2D triangle mesh is often implemented by using the widely used marching cubes algorithm [79]. Without loss of generality, we can assume the isosurface to be extracted is the 0 level set. The vertices can be found on edges with opposite field values on both ends. The exact location of the intersection of the isosurface on the edge can be easily computed on the basis of the trilinear interpolation approximation of the continuous underlying field, which reduces to linear interpolation along the edge. The mesh connectivity is then established by examining each cell and constructing triangles with vertices on the edges of that cell by checking in a predetermined lookup table, which contains the connectivity information of the vertices within each cell for the 256 possible sign configurations of the eight grid points of the cell. The actual lookup table is reduced to 15 cases by using symmetry. For some cases, disambiguation based on actual field values is necessary [80].

We recommend the use of marching cubes for applications that do not require well-shaped elements, as this approach is highly parallelizable. On the other hand, for FEM and other applications with stringent requirement on the maximum and minimum angles of each triangle, we propose to use an alternative based on restricted Delaunay triangulation [81]. A sample implementation is available from the Computational Geometry Algorithms Library (CGAL) [82]. It distributes sample points on the surface and then extracts an interpolating triangle mesh through the 3D triangulation of the points. These points are added iteratively following a Delaunay refinement-like step until the sizes and shapes of the mesh triangles meet specified criteria. Other choices include dual contouring for adaptive octree data [83, 84], and the extended marching cubes for models with sharp features [85], which might be rare for cryo-EM datasets though.

## 2.4. Surface areas and surface-enclosed volumes

Surface area and surface-enclosed volume are crucial components in the mathematical and thermodynamical modeling of biomolecular systems [1, 8, 10, 86]. Given a Lagrangian mesh with piecewise flat segments, it is straightforward to compute the surface area and surface-enclosed volume. For surface area, one simply sums up the area of each surface triangle,

$$A(S) = \int_S 1 \, dA = \sum_{t_l \in T} |t_l| = \sum_{t_l \in T} \frac{1}{2} |(\mathbf{v}_k - \mathbf{v}_i) \times (\mathbf{v}_j - \mathbf{v}_i)|, \quad (9)$$

where  $S$  is the surface of the biomolecule,  $A(S)$  is the total surface area, and  $T$  contains all the surface triangles in a tessellation of  $S$ . Here,  $t_l$  is a triangle mesh element, with  $\mathbf{v}_i$ ,  $\mathbf{v}_j$ , and  $\mathbf{v}_k$  as the coordinates of its three vertices.

To calculate the volume on the mesh, one picks an arbitrary point inside or outside the mesh, for example, the origin, and sums up all the signed volume of the tetrahedron formed by the point and each triangle,

$$V(S) = \int_S \mathbf{x} \cdot \mathbf{n} \, dA = \sum_{t_l \in T} \frac{1}{6} ((\mathbf{v}_k - \mathbf{v}_i) \times (\mathbf{v}_j - \mathbf{v}_i)) \cdot \mathbf{v}_i, \quad (10)$$

where  $V(S)$  is the total volume, and  $\mathbf{n}$  is the outward surface normal at a position  $\mathbf{x}$  on the surface  $S$ . Here, the vertices of each triangle are assumed to be listed in counterclockwise order when viewed from the outside of the surface. Even when a volumetric mesh of the inside is available, summing up the volumes of these thin tetrahedra formed by a fixed point and boundary faces is in general much more efficient than summing up the volumes of all the tetrahedra of the volumetric mesh.

## 2.5. Curvatures

**2.5.1. Brief introduction to the continuous theory.** Local geometric shape of the surface can be described by curvatures. A brief description of curvatures in terms of differential geometry theory can be found in Ref. [9]. The curvature for a point on a curve represents how fast the tangent direction turns, or more precisely, the magnitude of the second derivative of the curve in its arc-length parameterization. For a point on the surface, one can create a planar curve through the intersection of the surface and the local plane spanned by the surface normal and a tangent direction. The curvature of this planar curve is called the normal curvature along the chosen tangent direction. We can denote the maximum curvature among these normal curvatures by  $\kappa_1$  and the minimum curvature by  $\kappa_2$ . These two curvatures are called principal curvatures, and the tangent directions associated with them are called principal directions. Note that these two directions are always orthogonal to each other. It can be further shown that the normal curvature  $\kappa$  along an arbitrary direction can be determined by  $\kappa_1$  and  $\kappa_2$  (the principal curvatures) and the angle  $\theta$  that the chosen tangent direction makes with the maximum curvature direction

$$\kappa = \cos^2(\theta)\kappa_1 + \sin^2(\theta)\kappa_2.$$

The second-order approximation (quadric surface patch) around that point will be the same if two principal curvatures are identical. To see this, we describe the neighborhood of a point on the surface by the deviation of the surface from the tangent plane (a height function  $z = f(x, y)$  in a local coordinate system with the origin aligned to the point and the  $xy$ -plane aligned to the tangent plane). The second-order approximation is

$$z = \frac{1}{2} \begin{pmatrix} x & y \end{pmatrix} \text{Hess} \begin{pmatrix} x \\ y \end{pmatrix} + \dots,$$

where Hess is the Hessian (the symmetric second derivative matrix) of  $f$ ,

$$\text{Hess} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}.$$

The actual shape of the second-order approximation depends only on the eigenvalues of Hess, because applying a rotation in the tangent plane can diagonalize the Hessian. Thus, we can align two local approximation shapes through a rotation, as long as the diagonalized Hessians are the same. By the definition of curvature, one can immediately see that these eigenvalues are  $-\kappa_1$  and  $-\kappa_2$ . Here, we follow the convention in which bending towards the normal indicates a negative curvature, and bending away from the normal indicates a positive curvature. In this way, the curvatures for spheres will be positive. Note that some authors use the opposite sign.

Alternatively, it is often advantageous to use the Gaussian curvature and the mean curvature defined by

$$K = \kappa_1\kappa_2, \tag{11}$$

$$H = \frac{1}{2}(\kappa_1 + \kappa_2) \tag{12}$$

where  $K$  is the Gaussian curvature, and  $H$  is the mean curvature. They correspond to, respectively, the determinant and half of the trace of the aforementioned Hessian matrix, which is another way of prescribing the rotation invariants.

On the basis of the signs of the Gaussian curvature and the mean curvature, the neighborhood of a surface point can be roughly classified as one of the eight different shapes, namely, pit, valley, saddle ridge, flat, minimal surface, saddle valley, ridge, and peak. In Table I, we specify the type of shapes for each possible combination of signs. The actual shapes can be found in Figure 1.

Considering the quadratic approximations they represent, we can see that local shapes with opposite signs of mean curvatures and same signs of Gaussian curvatures may fit together.

Table I. Surface types based on signs of Gaussian curvature and mean curvature as illustrated in Figure 1.

	$K > 0$	$K = 0$	$K < 0$
$H > 0$	Peak	Ridge	Saddle ridge
$H = 0$	None	Flat	Minimal surface
$H < 0$	Pit	Valley	Saddle valley

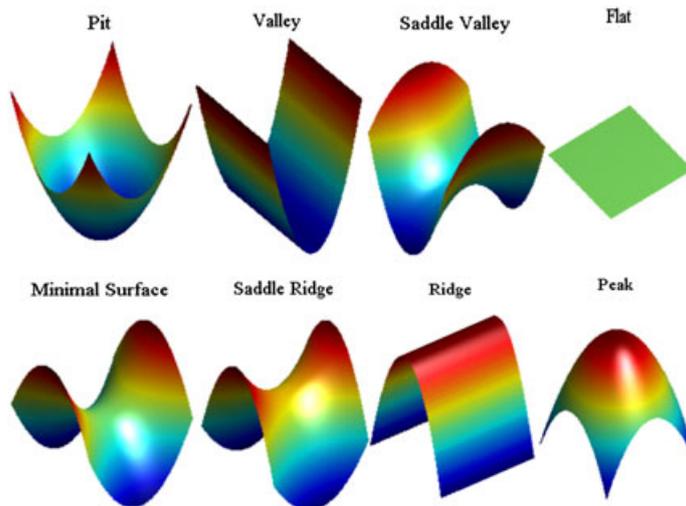


Figure 1. Representative image gallery of surface types based on signs of Gaussian curvature and mean curvature listed in Table I.

To give intuitive descriptions of the shape, another pair of continuous variables, the shape index  $s$  and the curvedness  $c$ , can also be defined [73]

$$s = -\frac{2}{\pi} \arctan \left( \frac{\kappa_1 + \kappa_2}{\kappa_1 - \kappa_2} \right),$$

$$c = \sqrt{\frac{1}{2} (\kappa_1^2 + \kappa_2^2)}.$$

Here,  $s$  describes the relation between the principal curvatures and  $c$  describes how nonflat the shape is.

2.5.2. *Curvature estimates on triangle meshes.* Meshing is a crucial component in geometric and mechanical analysis [58, 87–94]. In this work, we employ triangle meshes for the geometric representation of macromolecules. Therefore, we discuss schemes for curvature calculation on the triangle meshes.

The integral of Gaussian curvature over a small region and the integral of the geodesic curvature (deviation of a surface curve from a geodesic curve or a locally shortest curve) over its boundary sum up to  $2\pi$ , according to the Gauss–Bonnet theorem. For a triangle mesh, the integral of the geodesic curvature around the dual loop around a vertex (e.g., the loop in Figure 3) is the same as the sum of the tip angles of triangles containing that vertex. Thus, the Gaussian curvature integral for a dual cell around the vertex is often estimated by the angle defect (or angle deficit), the difference between  $2\pi$ , and the sum, which is also called the Gauss–Bonnet scheme. To obtain a pointwise estimate, we can divide it by the area of the neighborhood around the vertex, as shown in Figure 2.

The mean curvature normal can be estimated by using the Laplace–Beltrami operator applied to the surface description [95], which is essentially an estimate of the trace of the Hessian of the local description of the surface. The neighborhood area of the vertex could be estimated either by barycentric dual cell area (one third of the sum of the neighbor triangles’ area) or the Voronoi dual cell area (the area of the region containing points closer to the vertex than to any other vertices).

The discrete estimates of the curvatures can be formulated as follows:

$$K_i = \frac{1}{A_i} \left( 2\pi - \sum_{\theta_j \in \Theta_i} \theta_j \right), \tag{13}$$

where  $K_i$  is the estimated curvature at vertex  $i$ ,  $\theta_j$  is the angle of triangle  $j$  at vertex  $i$ ,  $\Theta_i$  is a collection of all angles around vertex  $i$  as shown in Figure 2, and  $A_i$  is one of the dual cell areas (Voronoi dual, barycentric dual, and mixed dual areas) in Figure 3 used to estimate the curvature. As shown in Ref. [95], the Voronoi dual cell area guarantees the least estimated errors for meshes

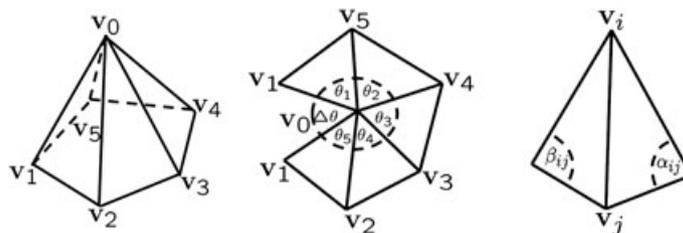


Figure 2. Schematic illustration of curvature algorithms. Left: a typical ‘one-ring’ neighborhood of a vertex ( $v_0$ ); Middle: flattening the one-ring by ‘cutting open’ along the edge  $v_0v_1$ , we can measure the ‘angle deficit’ used in Gaussian curvature estimates, denoted here by  $\Delta\theta = 2\pi - \sum_{i=1}^5 \theta_i$ . Right: angles used in the cotangent formula for the Laplace–Beltrami estimate of mean curvature.

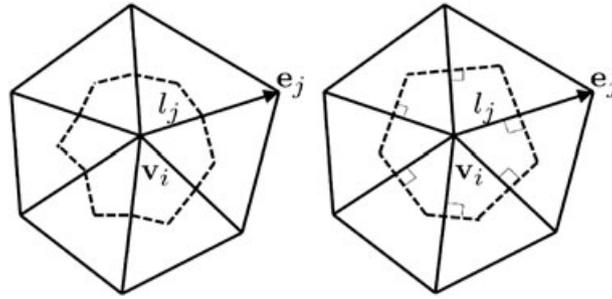


Figure 3. An illustration of dual cells defined around a vertex. Left: the area of the barycentric dual cell around a vertex (the cell formed by connecting consecutive barycenters of the triangles and edges incident to the center vertex  $v_i$ ), here  $l_j$  is the length of the part of edge  $e_j$  inside the neighborhood; Right: the area of the Voronoi dual cell of a vertex (the region containing all points closer to the center vertex  $v_i$  than to any other vertices).

with nonobtuse triangles. The straightforward estimation suggested by the authors for the Voronoi cell area around a vertex on mesh is

$$A_i = \frac{1}{8} \sum_{v_j \in N_i} (\cot \alpha_{ij} + \cot \beta_{ij}) \|\mathbf{v}_i - \mathbf{v}_j\|^2, \tag{14}$$

where  $N_i$  is the collection of all vertices immediately adjacent to vertex  $v_i$ , also known as the one-ring neighborhood as shown in Figure 3. For one-ring neighborhoods containing obtuse angles, a modification called mixed area can be applied [95]. In practice, the formula using Voronoi dual area produces better results even when there are negative cotangents.

The mean curvature ( $H_i$ ) at vertex  $i$  can be computed as

$$\mathbf{H}_i = \frac{1}{4A_i} \sum_{v_j \in N_i} (\cot \alpha_{ij} + \cot \beta_{ij})(\mathbf{v}_i - \mathbf{v}_j), \tag{15}$$

where  $\mathbf{H}_i = H_i \mathbf{n}_i$  is the mean curvature normal, and  $\mathbf{n}_i$  is a unit vector representing an estimated surface normal at vertex  $i$ . The sign of  $H_i$  can be determined by the sign of  $\mathbf{H}_i \cdot \mathbf{n}$ , where  $\mathbf{n}$  is the normal of any of the triangles incident to vertex  $i$ . Here,  $\mathbf{v}_i$  and  $\mathbf{v}_j$  are the coordinates of vertex  $i$ , and  $j$ ,  $\alpha_{ij}$  and  $\beta_{ij}$  are the opposite angles of the same edge in two triangles incident to the edge (see, e.g., Figure 2 right).

If an estimated curvature tensor is required, a commonly used approach is to take the average of the curvature tensor evaluated on the edges inside a certain neighborhood [96]

$$\mathbf{C}(\mathbf{v}_i) = \frac{1}{A'_i} \sum_{e_j \in E(\mathbf{v}_i)} \beta(e_j) l_j \bar{\mathbf{e}}_j \bar{\mathbf{e}}_j^T, \tag{16}$$

where  $\mathbf{C}(\mathbf{v}_i)$  is the estimated curvature tensor at vertex  $i$  expressed as a symmetric  $3 \times 3$  matrix in the global Euclidean coordinates, and  $A'_i$  is the area of a specific neighborhood, for which common choices include, for example, the intersection of a sphere which gives radius around  $i$ , a geodesic disk on the surface around  $i$ , or the one-ring of vertex  $i$ . Here,  $E(\mathbf{v}_i)$  is the set of all edges intersecting the neighborhood around vertex  $i$ ,  $\beta(e_j)$  is the signed dihedral angle between the normals of the faces sharing edge  $e_j$  (negative when the faces bend towards the surface normal and positive otherwise),  $\bar{\mathbf{e}}_j$  the unit direction along  $e_j$  (choosing either orientation of the edge will result in the same tensor),  $(\cdot)^T$  denotes the matrix transpose operation, and  $l_j$  is the length of the part of edge  $e_j$  inside the neighborhood. To find two principal curvatures and two principal directions, one may perform an eigen-decomposition of  $\mathbf{C}(\mathbf{v}_i)$ . The eigenvalue with the smallest absolute value is always nearly 0, and the associated eigenvector is an estimate of the local surface normal. Other two eigenvalues are the principal curvatures, and their associated eigenvectors are two principal directions in the tangent plane. The larger the chosen neighborhood is, the less accurate the result is.

However, choosing an overly small neighborhood results in noisy estimates when the resolution of the mesh is low.

### 2.6. Volumetric meshing

Volumetric meshing refers to the interior meshing. It is possible to generate tetrahedron meshes directly from 3D images with theoretical bounds on dihedral angles using algorithms such as isosurface stuffing [97]. Isosurface stuffing uses regular patterns to tetrahedralize grid cells completely inside the surface, followed by a marching-cubes-like boundary treatment, which shifts some of grid points near the boundary for attaining a better element shape. The algorithm is extremely fast, and the surface can approximate smooth isosurfaces well under reasonable assumptions, but the element shape is not optimal, and its adaptivity is restricted to octree-like structures.

Other available popular algorithms include TetGen [72] and NetGen [98], both providing user control on the size and shape of tetrahedra. The NetGen can take either a constructive solid geometry (shapes composed of primitive shapes combined through Boolean operations, i.e., union, intersection, and subtraction) or a boundary surface representation. However, NetGen can be less robust than other algorithms, such as TetGen [99]. TetGen produces tetrahedron meshes through constrained Delaunay tetrahedralization. If the tetrahedron mesh is required to conform to a boundary triangle mesh, the TetGen can be the method of choice. However, restricting the boundary to the given mesh can make the quality of the volumetric mesh dependent on the surface triangle mesh given by the user.

Another recent algorithm using interleaved Delaunay refinement and mesh optimization [100] can generate quality meshes that satisfy a set of user-defined criteria, which can be useful, for example, when importance of the sampling density is determined by the local chemical structure. In the Delaunay refinement step, sample points are inserted to satisfy the user-specified quality requirements. In the optimization step, a target function called the optimal Delaunay triangulation energy is used, whose minimization leads to a high-quality mesh. A final step perturbing the locations of vertices of slivers (flat tetrahedra) further improves the mesh quality.

## 3. RESULTS AND DISCUSSION

In this section, we examine the accuracy and demonstrate the usefulness of the methods and algorithms presented in the last section. A collection of all cryo-EM data can be found at the EMDataBank (<http://emdatbank.org/index.html>), which was established to create a global deposition and retrieval network for cryo-EM maps and associated metadata. It also works as a portal of software tools for standardized map format conversion, segmentation, model assessment, visualization, and data integration. A list of EM software can be found in the website (<http://emdatbank.org/emsoftware.html>). Cryo-EM data are stored as a 3D grid of voxels (volumetric cell), each with a value corresponding to the density of electrons. The data format was developed by the Medical Research Council Laboratory of Molecular Biology and is supported by almost every molecular graphics software that supports volumetric data, such as visual molecular dynamics, PyMOL (<http://www.pymol.org/>), UCSF Chimera (<http://www.cgl.ucsf.edu/chimera/>), and so on. The Medical Research Council file format is used in cryo-EM data, and its specifications can be found from the website ([http://ami.scripps.edu/software/mrctools/mrc\\_specification.php](http://ami.scripps.edu/software/mrctools/mrc_specification.php)).

In the present work, we consider six representative cryo-EM maps from the EMDataBank. With the help of visualization tool visual molecular dynamics (<http://www.ks.uiuc.edu/Research/vmd/>), we extract their surfaces with the recommended iso-values, and the results are displayed in Figure 4. The details of these data are summarized as follows.

- Figure 4A (EMD1048): The baseplate of bacteriophage T4. It is a multiprotein molecular machine that controls host cell recognition, attachment, tail sheath contraction, and viral DNA ejection.
- Figure 4B (EMD1129): GDP-tubulin. It is a GDP-bound tubulin. The rope-like polymers of tubulin, which are components of the cytoskeleton, can grow as long as 25  $\mu\text{m}$  and are highly dynamic.

- Figure 4C (EMD1265): Bacteriophage  $\phi 29$ . It is a viral DNA-packaging motor, which translocates and compresses genomic DNA with tremendous velocity into a preformed protein shell (the procapsid).
- Figure 4D (EMD1590): *Manduca sexta* vacuolar ATPase complex. It is a V-ATPase, which acidifies a wide array of intracellular organelles and pumps protons across the plasma membranes. V-ATPases couple the energy of ATP hydrolysis to proton transport across intracellular and plasma membranes of eukaryotic cells.
- Figure 4E (EMD1617): *Shigella flexneri* T3SS needle complex. The type three secretion system (T3SS) is a protein appendage found in several Gram-negative bacteria. In pathogenic bacteria, the needle-like structure is used as a sensory probe to detect the presence of eukaryotic organisms and secrete proteins that help the bacteria infect them.
- Figure 4F (EMD5119): Clathrin coats. It is a polyhedral lattice that surrounds the vesicle to safely transport molecules between cells. The endocytosis and exocytosis of vesicles allow cells to transfer nutrients, to import signaling receptors, and to mediate an immune response.

### 3.1. Data denoising and surface extraction

In this work, we explore our integrated tools on EMD datasets to test the strategies proposed used for geometric modeling. Six different EMD objects shown in Figure 4 are employed for the present study. It can be seen from Figure 4 that the electron tomography sometimes produces extremely noisy and low contrast 3D density maps. The poor signal-to-noise ratio (SNR) hinders visualization and interpretation. Therefore, some noise filtering techniques are indispensable. Many important methods and schemes, like wavelet transform techniques, nonlinear anisotropic diffusions, Beltrami flow, bilateral filter, and iterative median filtering have been used for noise reduction [12, 15–19, 101]. In this work, to improve the noise removal, we make use of the high-order geometric flows. Basically, it is a set of high-order geometric PDE based low-pass filters for image processing or surface analysis. An example of noise removal of EMD1617 is demonstrated in Figure 5. The basic structure of the T3SS needle complex is preserved while the noise amplitude is dramatically reduced. During the process of noise reduction, the surface of the protein is smoothed.

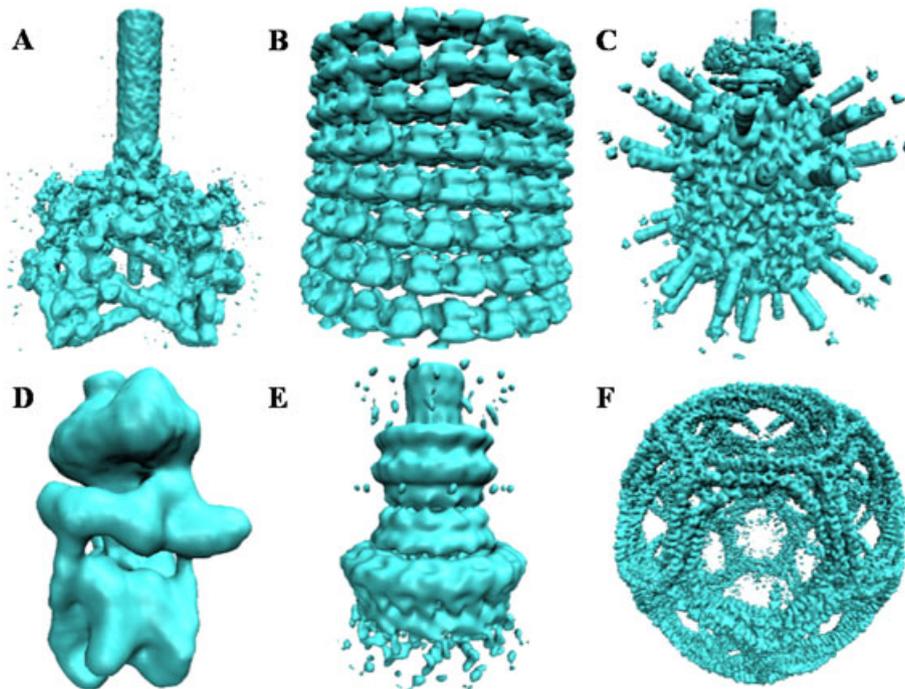


Figure 4. Image gallery of representative cryo-electron microscopy maps used in this study. The visual molecular dynamics is used for visualization.

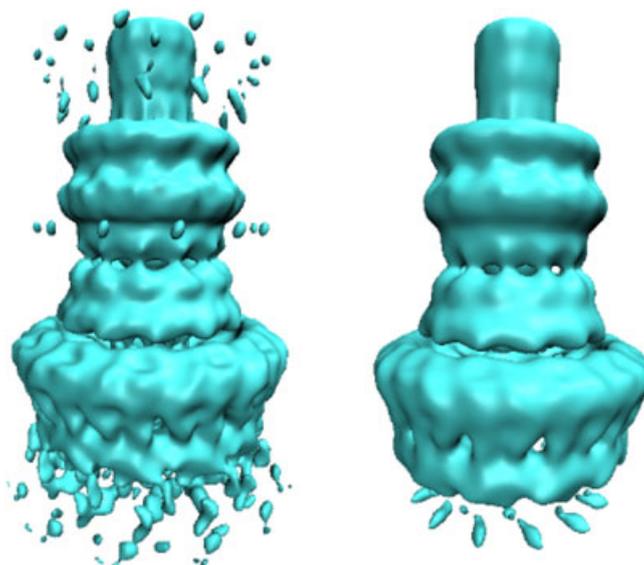


Figure 5. Noise reduction of EMD1617. Left: before filtering; right: after filtering by high-order geometric PDEs.

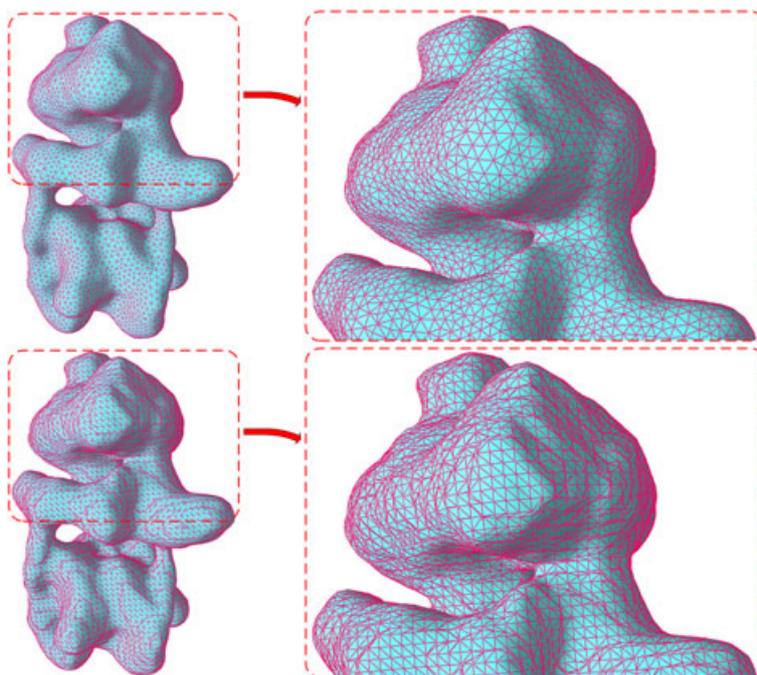


Figure 6. Comparison of surface meshes. Top: marching cubes result of EMD1590; bottom: Computational Geometry Algorithms Library result of EMD1590.

Artificial sharp edge and sharp tips are naturally removed. From the energy minimization point of view, these features are not favorable in the biological surface formation [9]. Therefore, the loss of these features does not lead to degradation in accuracy when dealing with biological data.

Among those surface extraction methods, shown in Section 2.3, we compare the results of the marching cubes method and CGAL's Delaunay-based method in Figure 6. The marching cubes method is highly parallelizable because the lookup table used is precomputed and stored. For all the files that we tested, it costs less than 5 s to process a cryo-EM map with up to  $200 \times 200 \times 200$  Cartesian grids. The direct result from marching cubes for EMD1590 shows that it may have a large

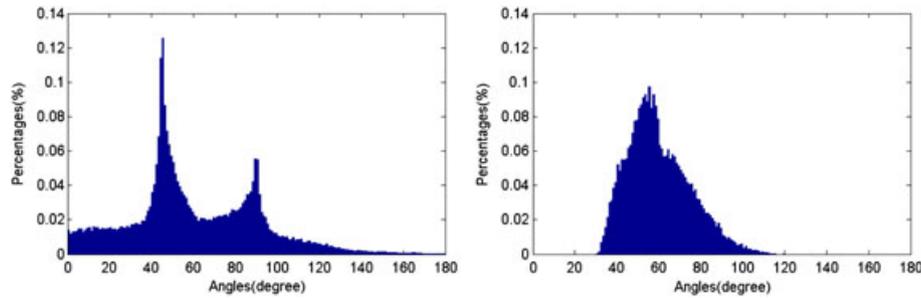


Figure 7. Comparison of surface mesh angle distributions. Left: angle histogram of marching cubes result of EMD1590; Right: angle histogram of Computational Geometry Algorithms Library isosurface extraction result of EMD1590.

number of skinny triangles, and the overall shape may contain terracing artifacts for a large proportion of the triangles. Many triangles have sharp angles less than  $30^\circ$ . The lack of element quality control is the intrinsic weakness of the marching cubes methods. Thus, the result often needs post-processing to improve the mesh quality. It may also require a large number of triangles to store (14,170 vertices and 28,360 triangles in the example shown) at a given accuracy due to the lack of adaptivity.

On the other hand, CGAL's running time and generated surface mesh quality depend heavily on the criteria chosen for the Delaunay triangulation. The criteria are controlled by three parameters: angular bound for mesh triangles' minimal angle, radius bound for the maximum surface Delaunay balls' radius (a surface Delaunay ball circumscribes a mesh triangle and centered on surface), and distance bound for the maximum distance between triangle's circumcenter and surface Delaunay ball center. However, these parameters can usually be easily tuned to achieve proper results with appropriate triangle shapes and sizes given some domain knowledge of the intended application. We use  $30^\circ$  as the angular bound, 0.8 as the radius bound, and 0.8 as the distance bound to directly extract the EMD1590 surface mesh from its cryo-EM map file. It takes about 4 s to run the algorithm to obtain the extracted surface with 10,100 vertices and 20,220 triangles. If we set smaller parameter values, we will obtain more detailed models but may suffer from longer running time and increased mesh size due to the smaller mesh triangles. Compared with the marching cubes result for EMD1590, the CGAL result gives triangle angles always greater than  $30^\circ$ , triangles with almost the same size, and reduced mesh sizes without losing surface shape accuracy, as shown in Figure 6.

The histograms of the triangle angle for both meshes are given in Figure 7. From the figure, one can see that the angles in the marching cubes results have a large distribution in the low range, which may result in accuracy problems in mathematical modeling involving PDEs. In contrast, CGAL's results are guaranteed to meet the angle requirements while reducing the mesh size significantly.

### 3.2. Surface mesh improvement

The surface generated from the marching cubes method or its variants often does not fit the need of applications relying on finite element, finite difference, or finite volume methods, such as geometric reconstruction of the internal structure or numerical simulation of electrostatics [1, 8, 10, 102, 103]. The process of creating a mesh that satisfies the new requirements while remaining close to the original mesh is called remeshing [104]. For instance, the mesh for EMD1590 produced by the marching cubes method can be remeshed into a mesh with rather uniform well-shaped triangles as shown in Figure 8. Even for meshes with well-shaped elements (for triangle meshes, this means nearly equilateral triangles, measured by the ratio of the circumcircle radius to the length of the shortest edge [105]), it is still possible to perform remeshing to reduce vertex count while remaining faithful to the original underlying surface. In this case, the procedure is also called mesh simplification.

Figure 9 presents a collection of meshes of six cryo-EM maps generated by using the CGAL approaches. The bacteriophage  $\phi 29$  and clathrin lattice have small-scale features. In particular, clathrin lattice data are quite noisy. It is seen from Figure 9 that the CGAL library is very robust and reliable for cryo-EM meshing.

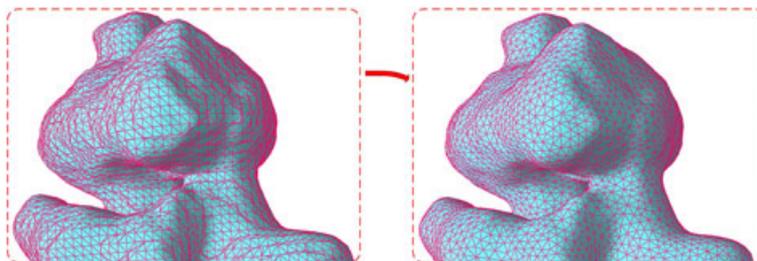


Figure 8. Mesh improvement with Delaunay remeshing. Left: marching cubes result of EMD1590; Right: remeshing result using the result to the left as the input.

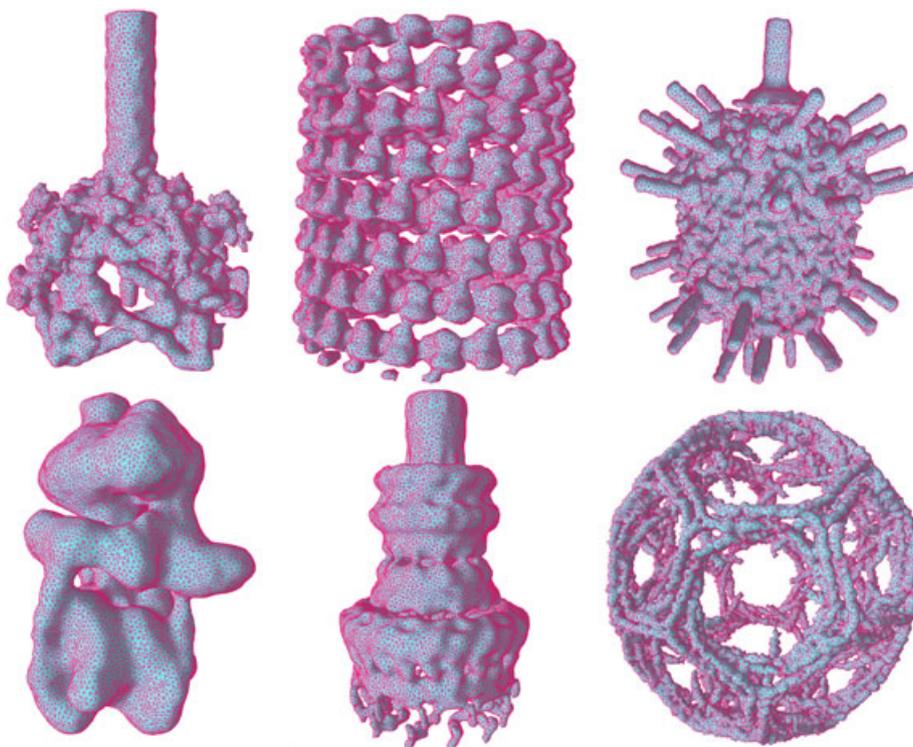


Figure 9. Computational Geometry Algorithms Library results of surface meshes. From upper left to lower right: EMD1048, bacteriophage T4 baseplate [106]; EMD1129, GDP-tubulin [107]; EMD1265, bacteriophage  $\phi$ 29 [108]; EMD1590, vacuolar ATPase motor [109]; EMD1617, *Shigella flexneri* T3SS needle complex [110]; and EMD5119, clathrin lattice [111].

### 3.3. Areas, volumes and curvatures

Surface areas and surface-enclosed volumes are frequently used in mathematical models of biomolecular systems [1, 8, 10, 86]. Accurate estimation of surface areas and surface-enclosed volumes is important in theoretical biology. The validation of the presented numerical methods is described in the succeeding paragraphs.

We compute surface areas and surface-enclosed volumes for spheres with different radii by the proposed methods and give the comparison between the theoretical values and their estimates in Table II. The radii used in our tests are 1,  $\sqrt{2}$ , and  $\sqrt{3}$ , whose theoretical values of area and volume are straightforward to compute. It can be seen that the straightforward methods proposed in this work are accurate for the high-quality meshes generated using the proposed methods.

As shown in Section 2.5, the curvature at a point on the surface describes the local geometric feature. Curvature analysis is useful for the identification of protein–protein and protein–ligand

Table II. Comparison of theoretical values and computer estimate of sphere’s areas and volumes.

Radius	Total area (est.)	Total area (theo.)	Total volume (est.)	Total volume (theo.)
1	12.53	12.57	4.16	4.19
$\sqrt{2}$	25.09	25.13	11.81	11.85
$\sqrt{3}$	37.66	37.70	21.72	21.77

Table III. The curvatures estimated using barycentric dual cell area.

radius	$\mu_K$	$\sigma_K^2$	$K_{\text{theo}}$	$\mu_H$	$\sigma_H^2$	$H_{\text{theo}}$
1	1.003239	0.026352	1	1.000031	0.018632	1
$\sqrt{2}$	0.500804	0.011344	0.5	0.707106	0.011323	0.707107
$\sqrt{3}$	0.333685	0.009292	0.333333	0.577343	0.010881	0.577350

Here,  $\mu_K$  (resp.,  $\mu_H$ ) is the average of Gaussian curvature  $K$  (mean curvature  $H$ ),  $\sigma_K^2$  ( $\sigma_H^2$ ) is the standard deviation of  $K$  ( $H$ ), and  $K_{\text{theo}}$  ( $H_{\text{theo}}$ ) is the theoretical value of  $K$  ( $H$ ).

Table IV. The curvatures estimated using Voronoi cell area.

Radius	$\mu_K$	$\sigma_K^2$	$K_{\text{theo}}$	$\mu_H$	$\sigma_H^2$	$H_{\text{theo}}$
1	1.003238	0.010534	1	1.000030	0.002627	1
$\sqrt{2}$	0.500804	0.007382	0.5	0.707107	0.003701	0.707107
$\sqrt{3}$	0.333684	0.007158	0.333333	0.577343	0.005481	0.577350

Here,  $\mu_K$  (resp.,  $\mu_H$ ) is the average of Gaussian curvature  $K$  (mean curvature  $H$ ),  $\sigma_K^2$  ( $\sigma_H^2$ ) is the standard deviation of  $K$  ( $H$ ), and  $K_{\text{theo}}$  ( $H_{\text{theo}}$ ) is the theoretical value of  $K$  ( $H$ ).

interaction sites. It can also be used to help understand the protein–DNA binding specificity. In this work, we first validate the accuracy and convergence order of the numerical methods proposed in Section 2.5. We then demonstrate the usefulness of these methods for cryo-EM data analysis.

As discussed in Section 2.5, around each vertex, the one-ring area can be chosen in two different ways, the barycentric dual cell area and the Voronoi cell area. The accuracy of these approaches are examined by spheres of radii ( $r$ ) 1,  $\sqrt{2}$ , and  $\sqrt{3}$ . Their Gaussian and mean curvatures are given by  $1/r^2$  and  $1/r$ , respectively. These spheres are tessellated with triangles of similar sizes. The results of estimated curvatures obtained with the barycentric dual areas are shown in Table III, together with theoretical values. Both our results for both Gaussian and mean curvatures are accurate in the tests. The resulting standard deviations show that the difference between the computed value and the theoretical value is small relative to the mesh size. For a comparison, we also listed our results obtained with the Voronoi dual areas in Table IV, under the same mesh. It is seen that the curvatures computed with the Voronoi dual areas are essentially the same as those with the barycentric dual areas. However, the Voronoi dual area approach offers smaller standard deviations in Gaussian and mean curvature estimations than does the barycentric dual area approach. Therefore, the Voronoi dual area approach performs better and is utilized in the rest of this work.

To further explore the accuracy and convergence of our curvature estimate, we design tests on different analytical models with different geometric types, including all cases of the intrinsically nonflat ones, namely, peak, pit, saddle ridge, minimal surface, and saddle valley. The pit type is identical to the peak case because of the symmetry. Specifically, the test results on a patch of a sphere are shown here (Figure 10). The convergence orders of our curvature method are measured by  $L_\infty$  and  $L_2$  error norms. Tables V and VI show the orders for Gaussian curvature and mean curvature, respectively. The average  $L_\infty$  order is about 1.4, whereas the second accuracy is achieved for the  $L_2$  order. These results indicate the robustness and reliability of the proposed methods for curvature evaluation.

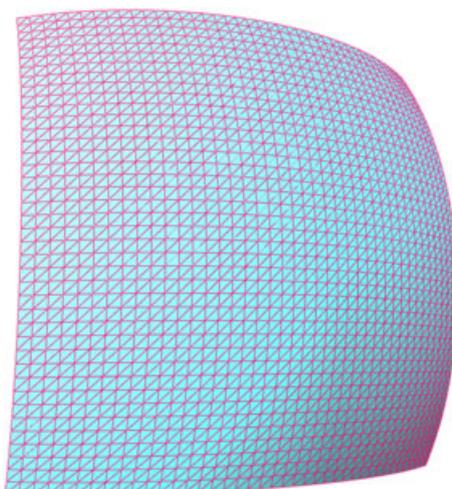


Figure 10. The analytical geometry model: a patch of a sphere.

Table V. The convergence orders for Gaussian curvatures on a patch of a sphere.

Maximal edge length	$L_\infty$	Order	$L_2$	Order
$1.00 \times 10^{-1}$	$8.325 \times 10^{-3}$		$4.815 \times 10^{-3}$	
$5.00 \times 10^{-2}$	$2.676 \times 10^{-3}$	1.64	$1.149 \times 10^{-3}$	2.07
$2.50 \times 10^{-2}$	$1.031 \times 10^{-3}$	1.38	$2.803 \times 10^{-4}$	2.04
$1.25 \times 10^{-2}$	$4.544 \times 10^{-4}$	1.18	$6.920 \times 10^{-5}$	2.02

Table VI. The convergence orders for mean curvatures on a patch of a sphere.

Maximal edge length	$L_\infty$	Order	$L_2$	Order
$1.00 \times 10^{-1}$	$1.275 \times 10^{-3}$		$5.228 \times 10^{-4}$	
$5.00 \times 10^{-2}$	$3.441 \times 10^{-4}$	1.89	$1.409 \times 10^{-4}$	1.89
$2.50 \times 10^{-2}$	$8.820 \times 10^{-5}$	1.96	$3.623 \times 10^{-5}$	1.96
$1.25 \times 10^{-2}$	$2.226 \times 10^{-5}$	1.99	$9.167 \times 10^{-6}$	1.98

3.3.1. *Applications of curvature estimates cryo-electron microscopy to maps.* Having established the accuracy and convergence of proposed numerical methods for curvature estimation, we apply these methods for the curvature calculation of six cryo-EM map entries. Note that these complexes vary in dimensions. The absolute value of curvatures increases as the dimension decreases as shown in the analytically expressions given in the last section.

First, we evaluate Gaussian curvatures and illustrate the results in Figure 11. Because the Gaussian curvature is an intrinsic measure of curvature and does not depend on the surface embedding, it is a convenient tool for identifying peak, pit, saddle ridge, and saddle valley. These features are clearly demonstrated in Figure 11. Taking the *Shigella exneri* T3SS needle complex as an example, Gaussian curvatures are mostly negative along the ring regions, which can be identified as saddle valleys, whereas Gaussian curvatures are positive on peaks and noisy dots.

We next consider the mean curvatures of six cryo-EM map entries. In contrast to the Gaussian curvature, mean curvature is an extrinsic measure of curvature, and it reflects the local characteristic of a surface. Figure 12 plots the mean curvature maps of six biomolecular complexes. Overall, mean curvatures are mostly positive for these complexes, indicating the main geometric features of peaks, ridges, and noisy dots. However, regions with very negative mean curvature can be found for pits and valleys, which are clearly potential binding targets of other smaller compounds.

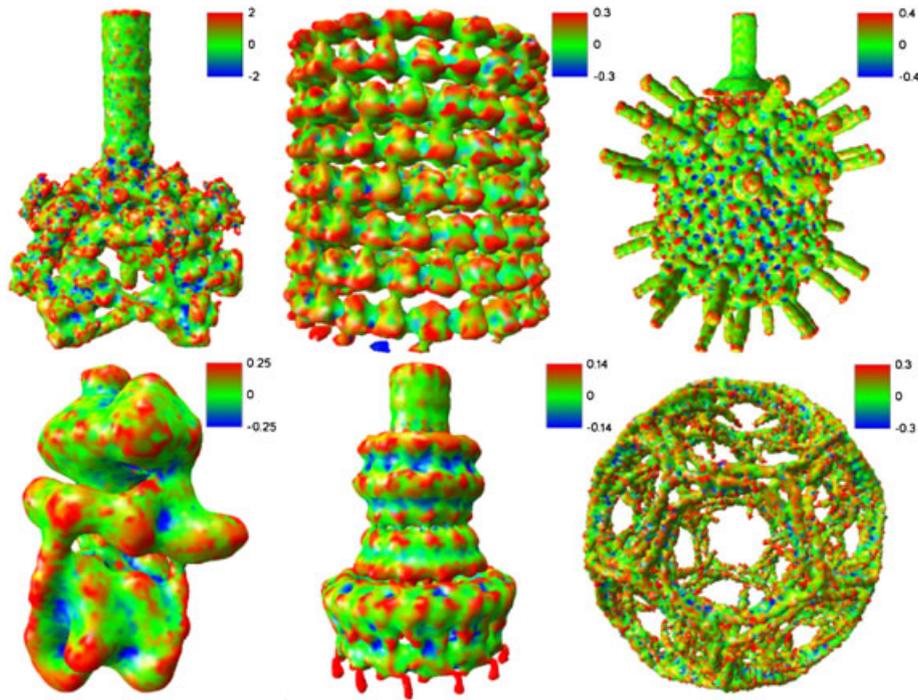


Figure 11. Gaussian curvature estimates for six cryo-electron microscopy map entries.

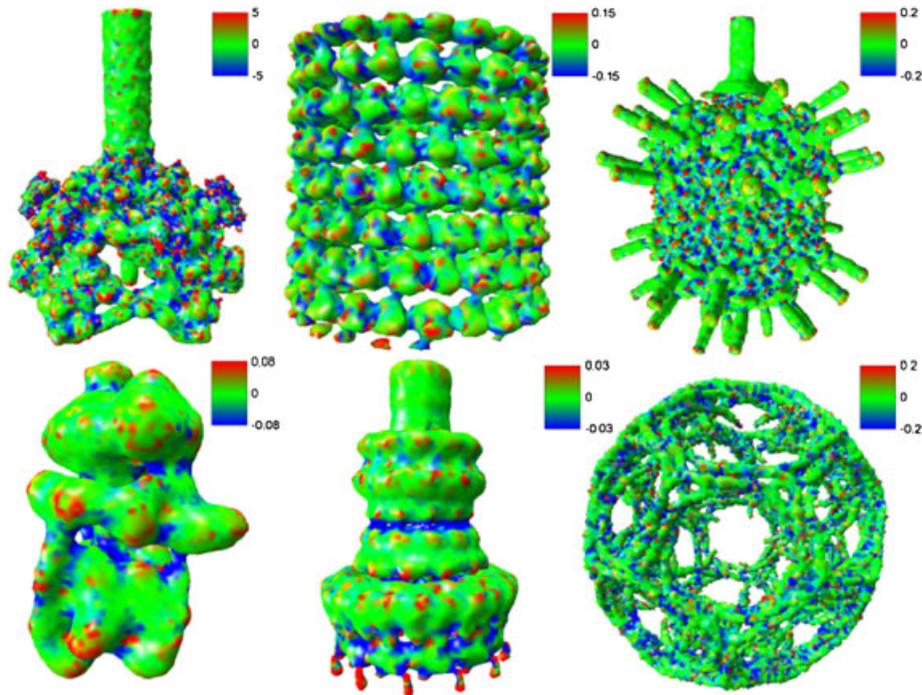


Figure 12. Mean curvature estimates for six cryo-electron microscopy map entries.

To further utilize the power of the present curvature estimates, we investigate the behavior of the first and second principal curvatures. The accuracy and convergence of the present curvature estimates established in the last section enable us to accurately compute principal curvatures as well by Equations (11) and (12). The maximum curvatures,  $\kappa_1$ , are plotted in Figure 13. It is interesting

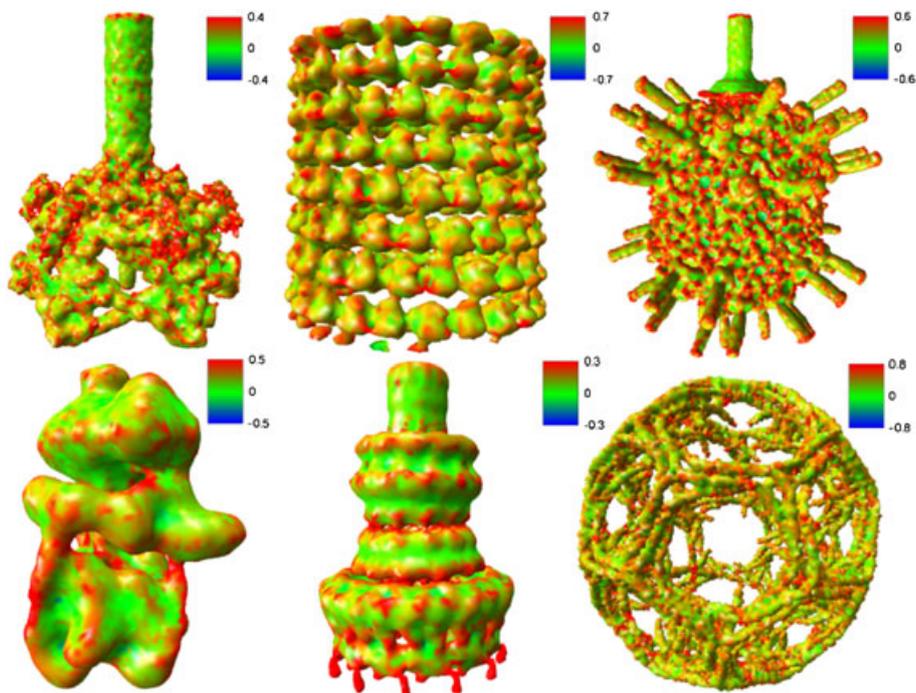


Figure 13. Maximum curvature ( $\kappa_1$ ) estimates for six cryo-electron microscopy map entries.

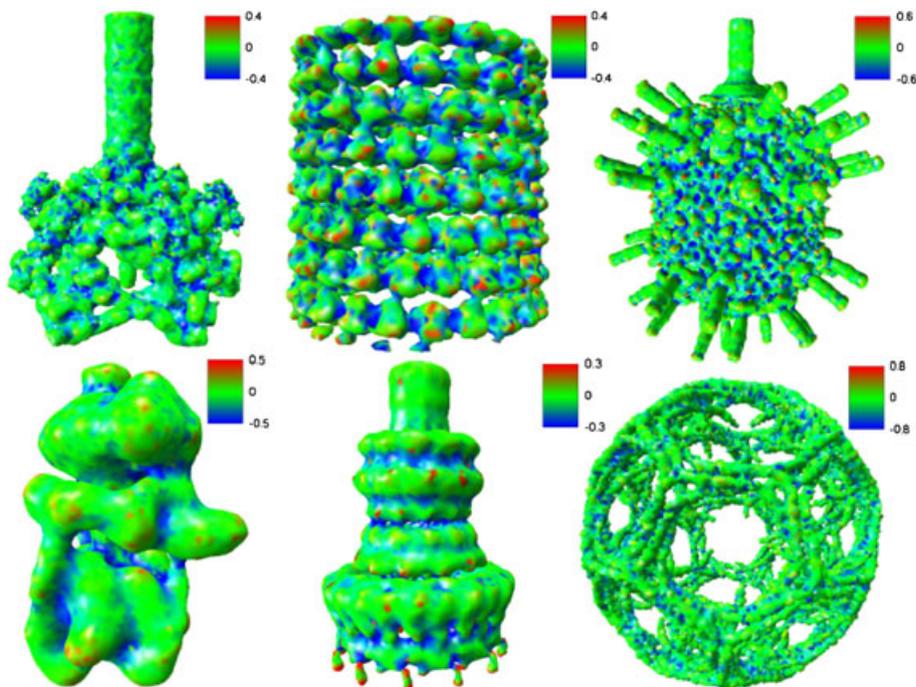


Figure 14. Minimum curvature ( $\kappa_2$ ) estimates for six cryo-electron microscopy map entries.

to note that the maximum curvature is a very good indicator for peaks and ridges of the biomolecular complex and possible noisy dots. Therefore, with a good confidence, one can exclude these regions with very large positive  $\kappa_1$  values from being targets of small binding compounds.

Finally, we investigate the behavior of the minimum curvature,  $\kappa_2$ . Results are depicted in Figure 14 for six cryo-EM entries. As expected, large negative curvatures indicate pits and valleys

(pockets), which are potential binding sites of small compounds. We believe that the second principal curvature can be used as a promising binding indicator for practical docking, drug design, and protein design analysis. This aspect, together with the electrostatic analysis, is further analyzed elsewhere for proteins [112].

### 3.4. Volumetric meshing

If a Lagrangian surface readily exists, its volumetric meshing is a separate task. There are a number of strategies for volumetric meshing. First, we can tetrahedralize the surface mesh files by using CGAL library functions. To reduce the work load of the tetrahedralization process, we use CGAL to extract the surface mesh with nearly same-sized triangles. Then, we use CGAL's tetrahedralization functionality to produce the tetrahedron mesh. The CGAL API function has five parameters to fine tune the tetrahedralization process: angular bound for surface mesh triangles' minimal angle, triangle size bound for the maximum surface Delaunay ball radius, triangle distance bound for the maximum distance between a triangle's circumcenter and the surface Delaunay ball center, cell radius edge ratio bound for the maximum ratio of the circumradius of a cell to its shortest edge, and cell size bound for the maximum cell circumradius. Setting smaller values for the latter three parameters will lead to more sampling points in the tetrahedralization step, which will increase the number of vertices and tetrahedra in meshes. If the user needs smaller cells near the surface mesh and larger cells far from the surface mesh, a large cell size and small surface triangle size bound could be adopted. The CGAL library tetrahedralization process has provable guarantees on the surface mesh quality, through tetrahedralization of the interior regions with constrained Delaunay triangulation.

TetGen is another popular choice for tetrahedralization step with high performance. The TetGen library has a large number of parameters to easily meet various requirements by the users. The most commonly modified parameters for tetrahedralizing a surface mesh is the maximum volume constraint on tetrahedron and the cell radius edge ratio bound. The default value of the cell radius edge ratio bound is 2.0, which can be lowered by the user to remove most cases of low-quality element shapes. A comparison of the results from TetGen and CGAL is given in Figure 15. As the surface triangle meshes are of similar good quality, both produced satisfiable results.

To observe the quality of the tetrahedron meshes generated by the proposed methods, we can show the planar cross-sectional views of the tetrahedron meshes as in Figure 15. Alternatively,



Figure 15. Comparison of volumetric meshing for an EMD1590 cut open in the middle. Left: TetGen result; Right: Computational Geometry Algorithms Library result.

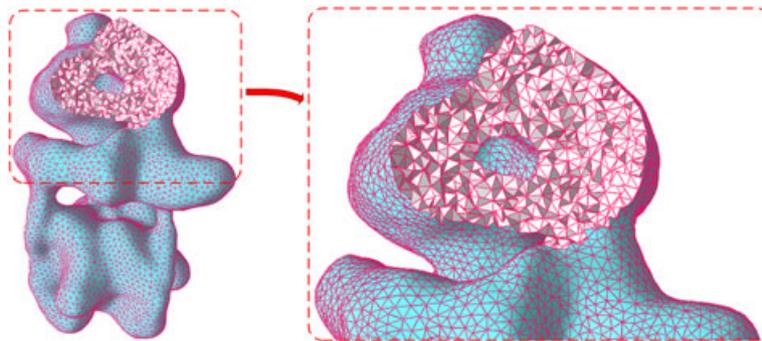


Figure 16. Cross section view of Computational Geometry Algorithms Library result of EMD1590.

we can also observe the internal structure by generating a cross section by removing a connected piece of volume as shown in Figure 16. For this purpose, we first choose a surface face as a seed face, then use breadth-first search algorithm to find a number of tetrahedra connected to the seed face. If we set a constant number of tetrahedra to remove as the stop criterion, we can expose the internal elements in a curved cross section at approximately the same distance from the seed face. This gives us a cutaway view to illustrate the interior meshing quality after tetrahedralization.

Both CGAL and TetGen share a parameter to tune the cell radius edge ratio bound of a generated tetrahedron. This parameter is highly effective in controlling the quality of the tetrahedra. All well-shaped tetrahedra have small values (less than 3) for the ratio, and most of the badly shaped tetrahedra have large values. This does not mean that the limit can be set arbitrarily small, because the value has a lower bound of 0.612 (the value for an equilateral tetrahedron). The one case of a badly shaped tetrahedron with small (cell radius to shortest edge) ratio is called ‘sliver’, which has a flat and near-degenerate shape. Its cell radius edge ratio can go as low as 0.707. One effective way to prevent slivers from being created is to incorporate a minimum volume constraint or to employ a procedure called sliver exudation as is done in CGAL.

#### 4. CONCLUSION

One of the most important new trends in structural biology is the investigation of the structures, functions, and dynamics of subcellular structures, organelles, and large multiprotein complexes. Currently, a main workhorse for this investigation is the cryo-EM. However, cryo-EM maps are well known for their low resolution and low reliability. Geometric modeling provides tools for improving cryo-EM resolution and bridges the gap between molecular structures and mathematical modeling, which is necessary for the understanding of the function and dynamics of subcellular structures and complexes. This work presents a comprehensive protocol for the geometric modeling of cryo-EM data and other volumetric density distributions used in macromolecules. We first introduce high-order geometric PDEs for noise removal of cryo-EM data and volumetric data. High-order geometric PDEs are more efficient for denoising because they have controllable time-frequency localization and can be tuned up for specific noise distribution. Additionally, we discuss two surface extraction schemes, the regular marching cubes, and the CGAL. Moreover, we explore efficient and reliable algorithms for volumetric meshing of macromolecular complexes, bearing in mind the further needs of other geometric tasks. Furthermore, we discuss some simple but robust schemes for the estimation of surface areas and surface-enclosed volumes based on the Lagrangian triangle-based surface representation. Finally, we introduce barycentric dual cell and Voronoi cell approaches for surface curvature estimations of biological complexes. We present explicit formula for evaluating conventional Gaussian curvature and mean curvature. On the basis of these curvatures, one can calculate the first and second principal curvatures as well.

The computational methods and algorithms introduced in the present work are carefully validated for their accuracy and convergence by analytically solvable systems. We show that the simple methods based on triangle meshes work well for estimating surface area and surface-enclosed

volumes. The second-order convergence of Gaussian and mean curvature estimations, particularly in the  $L_2$  norm, has been numerically observed. The  $L_\infty$  norm is also of second-order accuracy for the mean curvature calculations. Therefore, the resulting Gaussian and mean curvatures can be directly used for calculating two principal curvatures. The rigorous validations conducted in the present work provide us with empirical evidence in applying the proposed methods to the geometric analysis of cryo-EM data and other volumetric data for biomolecular systems with high-quality meshes, despite the known fact that these curvature estimates do not converge when there is no control over mesh quality.

Extensive applications are considered to cryo-EM data denoising, surface extraction, volumetric meshing, surface area and volume calculation, and curvature analysis. To this end, we utilize a total of six representative cryo-EM complexes, including bacteriophage T4 baseplate [106], GDP-tubulin [107], bacteriophage  $\phi 29$  [108], vacuolar ATPase motor [109], Shigella flexneri T3SS needle complex [110], and clathrin lattice [111]. We have demonstrated the efficiency of high-order geometric PDEs for noise removal of cryo-EM data. We also investigate the performance of marching cubes and CGAL schemes for surface extraction. Marching cubes method is readily parallelizable and is recommended for applications that do not require well-shaped elements. Whereas, if high-quality meshes are needed, the CGAL is a better choice. The comparison among different methods has been carried out for triangle meshing. Specifically, we analyze the performance of four algorithms, the isosurface stuffing [97], TetGen [72], NetGen [98], and Delaunay refinement and interleaved mesh optimization [100] for the volumetric meshing of biomolecules. TetGen and NetGen both perform well, while the last one is very suitable for quality improvement, capable of remeshing the surface from possibly low-quality surface mesh inputs. Interesting results are found in curvature analysis. It is found that the maximum and minimum curvature maps of cryo-EM complexes can be used for binding site characterization. Specifically, the maximum curvature can also be used to exclude regions from the binding targets of small molecules, whereas the minimum curvature serves a promising indicator of binding targets. These findings will be further investigated by combining curvature and electrostatic characteristics in our future work [112].

#### ACKNOWLEDGEMENTS

This work was supported in part by NSF grants CCF-0936830, IIS-0953096, and DMS-1160352, and NIH grant R01GM-090208.

#### REFERENCES

1. Wei GW. Differential geometry based multiscale models. *Bulletin of Mathematical Biology* 2010; **72**:1562–1622.
2. Yu Z, Holst M, Cheng Y, McCammon JA. Feature-preserving adaptive mesh generation for molecular shape modeling and simulation. *Journal of Molecular Graphics and Modeling* 2008; **26**:1370–1380.
3. Volkman N. Methods for segmentation and interpretation of electron tomographic reconstructions. *Methods in Enzymology* 2010; **483**:31–46.
4. Nickell S, Kofler C, Leis AP, Baumeister W. A visual approach to proteomics. *Nature Reviews Molecular Cell Biology* 2006; **7**:225–230.
5. Robinson CV, Sali A, Baumeister W. The molecular sociology of the cell. *Nature* 2007; **450**:973–982.
6. Leis A, Rockel B, Andrees L, Baumeister W. Visualizing cells at the nanoscale. *Trends in Biochemical Sciences* 2009; **34**:60–70.
7. Tocheva EI, Li Z, Jensen GJ. *Electron Cryotomography*, Chapter in Cold Spring Harbor Perspectives Biology, Vol. 46, Bacterial Cell, 2010. A003442.
8. Chen Z, Baker NA, Wei GW. Differential geometry based solvation models II: Lagrangian formulation. *Journal of Mathematical Biology* 2011; **63**:1139–1200.
9. Bates PW, Wei GW, Zhao S. Minimal molecular surfaces and their applications. *Journal of Computational Chemistry* 2008; **29**(3):380–391.
10. Chen Z, Baker NA, Wei GW. Differential geometry based solvation models I: Eulerian formulation. *Journal of Computational Physics* 2010; **229**:8231–8258.
11. Chen D, Chen Z, Wei GW. Quantum dynamics in continuum for proton transport II: Variational solvent-solute interface. *International Journal for Numerical Methods in Biomedical Engineering* 2012; **28**:25–51.
12. Stoschek A, Hegerl R. Denoising of electron tomographic reconstructions using multiscale transformations. *Journal of Structural Biology* 1997; **120**:257–265.

13. Frangakis AS, Hegerl R. Noise reduction in electron tomographic reconstructions using nonlinear anisotropic diffusion. *Journal of Structural Biology* 2001; **135**:239–250.
14. Fernandez JJ, Li S. An improved algorithm for anisotropic nonlinear diffusion for denoising cryo-tomograms. *Journal of Structural Biology* 2003; **144**:152–161.
15. Fernandez JJ. Tomobflow: feature-preserving noise filtering for electron tomography. *BMC Bioinformatics* 2009; **178**:1–10.
16. Tomasi C, Manduchi R. Bilateral filtering for gray and color images. *Proceedings of the International Conference on Computer Vision (ICCV)* 1998; **98**:839–846.
17. Jiang W, Baker ML, Wu Q, Bajaj C, Chiu W. Applications of a bilateral denoising filter in biological electron microscopy. *Journal of Structural Biology* 2003; **144**:114–122.
18. Pantelic RS, Rothnagel R, Huang CY, Muller D, Woolford D, Landsberg MJ, McDowall A, Pailthorpe B, Young PR, Banks J, Hankamer B, Ericksson G. The discriminative bilateral filter: an enhanced denoising filter for electron microscopy data. *Journal of Structural Biology* 2006; **155**:395–408.
19. van der Heide P, Xu XP, Marsh BJ, Hanein D, Volkmann N. Efficient automatic noise reduction of electron tomographic reconstructions based on iterative median filtering. *Journal of Structural Biology* 2007; **158**:196–204.
20. Tsai K, Ma J, Ye D, Wu J. Curvelet processing of mri for local image enhancement. *International Journal for Numerical Methods in Biomedical Engineering* 2012; **28**:661–677.
21. Pan M-S, Tang J-T, Rong Q-S, Zhang F. Medical image registration using modified iterative closest points. *International Journal for Numerical Methods in Biomedical Engineering* 2011; **27**:1150–1166.
22. Radaelli AG, Peiro J. On the segmentation of vascular geometries from medical images. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**(1):3–34.
23. Sochen N, Kimmel R, Malladi R. A general framework for low level vision. *IEEE Transactions on Image Processing* 1998; **7**(3):310–318.
24. Osher S, Fedkiw RP. Level set methods: an overview and some recent results. *Journal of Computational Physics* 2001; **169**(2):463–502.
25. Sethian JA. Evolution, implementation, and application of level set and fast marching methods for advancing fronts. *Journal of Computational Physics* 2001; **169**(2):503–555.
26. Bates PW, Wei GW, Zhao S. The minimal molecular surface. *arXiv:q-bio/0610038v1*, [q-bio.BM], 2006.
27. Zhang Y, Bajaj C, Xu G. Surface smoothing and quality improvement of quadrilateral/hexahedral meshes with geometric flow. *Communications in Numerical Methods in Engineering* 2009; **25**:1–18.
28. Witkin A. Scale-space filtering: a new approach to multi-scale description. *Proceedings of IEEE International Conference on Acoustic Speech Signal Processing* 1984; **9**:150–153.
29. Perona P, Malik J. Scale-space and edge-detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1990; **12**(7):629–639.
30. Soltanianzadeh H, Windham JP, Yagle AE. A multidimensional nonlinear edge-preserving filter for magnetic-resonance image-restoration. *IEEE Transactions on Image Processing* 1995; **4**(2):147–161.
31. Wei GW. Generalized Perona–Malik equation for image restoration. *IEEE Signal Processing Letters* 1999; **6**(7):165–167.
32. Chan T, Marquina A, Mulet P. High-order total variation-based image restoration. *SIAM Journal on Scientific Computing* 2000; **22**(2):503–516.
33. Wei GW, Jia YQ. Synchronization-based image edge detection. *Europhysics Letters* 2002; **59**(6):814–819.
34. Osher S, Sethian J. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations. *Journal of Computational Physics* 1988; **79**(1):12–49.
35. Mumford D, Shah J. Optimal approximations by piecewise smooth functions and associated variational problems. *Communications on Pure and Applied Mathematics* 1989; **42**(5):577–685.
36. Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. In *Proceedings of the Eleventh Annual International Conference of the Center for Nonlinear Studies on Experimental Mathematics: Computational Issues in Nonlinear Science*. Elsevier North-Holland, Inc.: Amsterdam, The Netherlands, The Netherlands, 1992; 259–268.
37. Blomgren P, Chan TF. Color TV: total variation methods for restoration of vector-valued images. *IEEE Transactions on Image Processing* 1998; **7**(3):304–309.
38. Carstensen V, Kimmel R, Sapiro G. Geodesic active contours. *International Journal of Computer Vision* 1997; **22**:61–79.
39. Osher S, Rudin LI. Feature-oriented image enhancement using shock filters. *SIAM Journal on Numerical Analysis* 1990; **27**(4):919–940.
40. Sapiro G, Ringach DL. Anisotropic diffusion of multivalued images with applications to color filtering. *IEEE Transactions on Image Processing* 1996; **5**(11):1582–1586.
41. Sazonov I, Nithiarasu P. Semi-automatic surface and volume mesh generation for subject-specific biomedical geometries. *International Journal for Numerical Methods in Biomedical Engineering* 2012; **28**:133–157.
42. You Y, Kaveh M. Fourth-order partial differential equations for noise removal. *IEEE Transactions on Image Processing* 2002; **9**(10):1723–1730.
43. Tasdizen T, Whitaker R, Burchard P, Osher S. Geometric surface processing via normal maps. *ACM Transactions on Graphics* 2003; **22**(4):1012–1033.

44. Lysaker M, Lundervold A, Tai XC. Noise removal using fourth-order partial differential equation with application to medical magnetic resonance images in space and time. *IEEE Transactions on Image Processing* 2003; **12**(12):1579–1590.
45. Greer JB, Bertozzi AL. H-1 solutions of a class of fourth order nonlinear equations for image processing. *Discrete and Continuous Dynamical Systems* 2004; **10**(1–2):349–366.
46. Bates PW, Chen Z, Sun YH, Wei GW, Zhao S. Geometric and potential driving formation and evolution of biomolecular surfaces. *Journal of Mathematical Biology* 2009; **59**:193–231.
47. Chambolle A, Lions PL. Image recovery via total variation minimization and related problems. *Numerische Mathematik* 1997; **76**(2):167–188.
48. Greer JB, Bertozzi AL. Traveling wave solutions of fourth order PDEs for image processing. *SIAM Journal on Mathematical Analysis* 2004; **36**(1):38–68.
49. Witelski TP, Bowen M. ADI schemes for higher-order nonlinear diffusion equations. *Applied Numerical Mathematics* 2003; **45**(2–3):331–351.
50. Sun YH, Wu PR, Wei GW, Wang G. Evolution-operator-based single-step method for image processing. *International Journal of Biomedical Imaging* 2006; **83847**:1–27.
51. Wang Y, Wei GW, Yang S-Y. Mode decomposition evolution equations. *Journal of Scientific Computing* 2012; **50**:495–518.
52. Wang Y, Wei GW, Yang S-Y. Partial differential equation transform–variational formulation and Fourier analysis. *International Journal for Numerical Methods in Biomedical Engineering* 2011; **27**:1996–2020.
53. Wang Y, Wei GW, Yang S-Y. Selective extraction of entangled textures via adaptive PDE transform. *International Journal in Biomedical Imaging* 2012; **2012**:Article ID 958142.
54. Zheng Q, Yang SY, Wei GW. Molecular surface generation using PDE transform. *International Journal for Numerical Methods in Biomedical Engineering* 2012; **28**:291–316.
55. Wei GW, Sun YH, Zhou YC, Feig M. Molecular multiresolution surfaces. *arXiv:math-ph/0511001v1*, pages 1–11, 2005.
56. Zhao S. Pseudo-time-coupled nonlinear models for biomolecular surface representation and solvation analysis. *International Journal for Numerical Methods in Biomedical Engineering* 2012; **27**:1964–1981.
57. d’Otreppe V, Boman R, Ponthot J-P. Generating smooth surface meshes from multi-region medical images. *International Journal for Numerical Methods in Biomedical Engineering* 2012; **28**:642–660.
58. Marchandise E, Compere G, Willemet M, Bricteux G, Geuzaine C, Remacle JF. Quality meshing based on stl triangulations for biomedical simulations. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**:876–889.
59. Pan W, Wheel M. A finite-volume method for solids with a rotational degrees of freedom based on the 6-node triangle. *International Journal for Numerical Methods in Biomedical Engineering* 2011; **27**:1411–1426.
60. Qian J, Lu J. Point-cloud method for image-based biomechanical stress analysis. *International Journal for Numerical Methods in Biomedical Engineering* 2011; **27**:1493–1506.
61. Salo Z, Beek M, Whyne CM. Evaluation of mesh morphing and mapping techniques in patient specific modelling of the human pelvis. *International Journal for Numerical Methods in Biomedical Engineering* 2012; **28**:904–913.
62. Xiao YX, Shu S, Zhang P, Tan M. An algebraic multigrid method for isotropic linear elasticity problems on anisotropic meshes. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**:534–553.
63. Younes A, Fahs M, Ackerer P. An efficient geometric approach to solve the slope limiting problem with the discontinuous Galerkin method on unstructured triangles. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**:1824–1835.
64. Du Q, Faber V, Gunzburger M. Centroidal Voronoi tessellations: applications and algorithms. *SIAM Review* 1999; **41**(4):637–676.
65. Ju LL, Du Q, Gunzburger M. Probabilistic methods for centroidal Voronoi tessellations and their parallel implementations. *Parallel Computing* 2002; **28**:1477–1500.
66. Wan M, Wang Y, Wang DS. Variational surface reconstruction based on Delaunay triangulation and graph cut. *International Journal for Numerical Methods in Engineering* 2011; **85**:206–229.
67. Sazonov I, Yeo SY, Bevan LT, Xie XH, Loon RV, Nithiarasu P. Modelling pipeline for subject-specific arterial blood flow – A review. *International Journal for Numerical Methods in Biomedical Engineering* 2011; **27**:1868–1910.
68. Saksono PH, Nithiarasu P, Sazonov I. Numerical prediction of heat transfer patterns in a subject-specific human upper airway. *Journal of Heat Transfer* 2012; **134**(031022):1–9.
69. Weatherill NP, Hassan O. Efficient three-dimensional Delaunay triangulation with automatic point creation and imposed boundary constraints. *International Journal for Numerical Methods in Engineering* 1994; **37**:2005–2039.
70. George PL, Hecht F, Saltel E. Automatic mesh generator with specified boundary. *Computer Methods in Applied Mechanics and Engineering* 1991; **92**:269–288.
71. Yu Z, Holst MJ, Hayashi T, Bajaj CL, Ellisman MH, McCammon JA, Hoshijima M. Three-dimensional geometric modeling of membrane-bound organelles in ventricular myocytes: bridging the gap between microscopic imaging and mathematical simulation. *Journal of Structural Biology* 2008; **164**:304–313.
72. Si H. Constrained Delaunay tetrahedral mesh generation and refinement. *Finite Elements in Analysis and Design* 2010; **46**:33–46.

73. Koenderink JJ, van Doorn AJ. Surface shape and curvature scales. *Image and Vision Computing* 1992; **10**(8):557–564.
74. Cipriano G, Phillips Jr., GN, Gleicher M. Multi-scale surface descriptors. *IEEE Transactions on Visualization and Computer Graphics* 2009; **15**:1201–1208.
75. Kuo WL, Cleghorn YL. Curvature- and displacement-based finite element analyses of flexible slider crank mechanisms. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**:1228–1245.
76. Gilboa G, Sochen N, Zeevi YY. Image sharpening by flows based on triple well potentials. *Journal of Mathematical Imaging and Vision* 2004; **20**(1–2):121–131.
77. Gilboa G, Sochen N, Zeevi YY. Forward-and-backward diffusion processes for adaptive image enhancement and denoising. *IEEE Transactions on Image Processing* 2002; **11**(7):689–703.
78. Botsch M, Kobbelt L, Pauly M, Alliez P, Lévy B. *Polygon Mesh Processing*. A K Peters: Abingdon, Oxford, UK, 2010.
79. Lorensen WE, Cline HE. Marching cubes: a high resolution 3d surface reconstruction algorithm. *Computer Graphics* 1987; **21**:163–169.
80. Montani C, Scateni R, Scopigno R. A modified look-up table for implicit disambiguation of marching cubes. *The Visual Computer* 1994; **10**(6):353–355.
81. Boissonnat J-D, Oudot S. Provably good sampling and meshing of surfaces. *Graphical Models* 2005; **67**(5):405–451.
82. Damiand G. Combinatorial maps. In *CGAL User and Reference Manual*, 4.0 edition. CGAL Editorial Board, 2012. [www.cgal.org/Manual/4.0/doc\\_html/cgal\\_manual/packages.html#Pkg:CombinatorialMaps](http://www.cgal.org/Manual/4.0/doc_html/cgal_manual/packages.html#Pkg:CombinatorialMaps).
83. Ju T, Losasso F, Schaefer S, Warren J. Dual contouring of hermite data. *ACM Transactions on Graphics (SIGGRAPH)* 2002; **21**(3):339–346.
84. Bischoff S, Pavic D, Kobbelt L. Automatic restoration of polygon models. *ACM Transactions on Graphics* 2005; **24**(4):1332–1352.
85. Kobbelt LP, Botsch M, Schwanecke U, Seidel H-P. Feature sensitive surface extraction from volume data. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '01*. ACM: New York, NY, USA, 2001; 57–66.
86. Wei G-W, Zheng Q, Chen Z, Xia K. Variational multiscale models for charge transport. *SIAM Review* 2012; **54**(4):699–754.
87. Botti L, Piccinelli M, Ene-Iordache B, Remuzzi A, Antiga L. An adaptive mesh refinement solver for large-scale simulation of biological flows. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**:86–100.
88. Lv G, Shen L, Shen Z. Numerical methods for energy flux of temperature diffusion equation on unstructured meshes. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**:646–665.
89. Phongthanapanich S, Dechaumphai P. Finite volume method for convection-diffusion-reaction equation on triangular meshes. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**:716–727.
90. Yamakawa S, Shimada K. 88-Element solution to schneiders' pyramid hex-meshing problem. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**:1700–1712.
91. Ko JH, Park SH, Park HC, Byun D. Finite macro-element-based volume grid deformation for large moving boundary problems. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**:1656–1673.
92. Miki T, Imai Y, Ishikawa T, Wada S, Aoki T, Yamaguchi T. A fourth-order Cartesian local mesh refinement method for the computational fluid dynamics of physiological flow in multi-generation branched vessels. *International Journal for Numerical Methods in Biomedical Engineering* 2011; **27**:424–435.
93. Mohapatra J, Natesan S. Uniformly convergent numerical method for singularly perturbed differential-difference equation using grid equidistribution. *International Journal for Numerical Methods in Biomedical Engineering* 2011; **27**:1427–1445.
94. Scott MH, Hamutcuoglu OM. Analytical sensitivity of interpolatory quadrature in force-based frame elements. *International Journal for Numerical Methods in Biomedical Engineering* 2010; **26**:1586–1595.
95. Desbrun M, Meyer M, Schröder P, Barr A. Implicit fairing of irregular meshes using diffusion and curvature flow. *ACM SIGGRAPH* 1999;317–324.
96. Cohen-Steiner D, Morvan J-M. Second fundamental measure of geometric sets and local approximation of curvatures. *Journal of Differential Geometry* 2006; **74**(3):391–404.
97. Labelle F, Shewchuk JR. Isosurface stuffing: fast tetrahedral meshes with good dihedral angles. *ACM Transactions on Graphics (SIGGRAPH)* 2007; **26**(3):57:1–10.
98. Schöberl J. Netgen - an advancing front 2d/3d mesh generator based on abstract rules. *Computing and Visualization in Science* 1997; **1**(1):41–52.
99. Lizier M, Shepherd J, Nonato L, Comba J, Silva C. Comparing techniques for tetrahedral mesh generation. *Inaugural International Conference of the Engineering Mechanics Institute*, Minneapolis, MN, 2008.
100. Tournois J, Wormser C, Alliez P, Desbrun M. Interleaving Delaunay refinement and optimization for practical isotropic tetrahedron mesh generation. *ACM Transactions on Graphics (SIGGRAPH)* 2009; **28**(3):75:1–75:9.
101. Fernandez JJ, Li S, Lucic V. Three-dimensional anisotropic noise reduction with automated parameter tuning: application to electron cryotomography. In *Current Topics in Artificial Intelligence*, Vol. 4788. AIP: Melville, NY, 2007; 60–69.

102. Yu SN, Geng WH, Wei GW. Treatment of geometric singularities in implicit solvent models. *Journal of Chemical Physics* 2007; **126**:244108.
103. Yu SN, Wei GW. Three-dimensional matched interface and boundary (MIB) method for treating geometric singularities. *Journal of Computational Physics* 2007; **227**:602–632.
104. Alliez P, Ucelli G, Gotsman C, Attene M. Recent advances in remeshing of surfaces. In *Shape Analysis and Structuring*, de Floriani L, Spagnuolo M (eds). Springer-Verlag: Berlin, 2007; 53–82.
105. Shewchuk JR. Delaunay refinement mesh generation. *Ph.D. Thesis*, School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, May 1997. Available as Technical Report CMU-CS-97-137.
106. Kostyuchenko VA, Leiman PG, Chipman PR, Kanamaru S, van Raaij MJ, Arisaka F, Mesyanzhinov VV, Rossmann MG. Three-dimensional structure of bacteriophage t4 baseplate. *Natural Structural Biology* 2003; **10**(9):688–693.
107. Wang H, Nogales E. Nucleotide-dependent bending flexibility of tubulin regulates microtubule assembly. *Nature* 2005; **435**(7044):911–915.
108. Xiang Y, Morais M, Battisti A, Grimes S, Jardine P, Anderson D, Rossmann M. Structural changes of bacteriophage phi29 upon DNA packaging and release. *EMBO Journal* 2006; **25**(21):5229–5239.
109. Muench SP, Huss M, Song CF, Phillips C, Wiczorek H, Trinick J, Harrison MA. Cryo-electron microscopy of the vacuolar atpase motor reveals its mechanical and regulatory complexity. *Journal of Molecular Biology* 2009; **386**(4):989–999.
110. Hodgkinson JL, Horsley A, Stabat D, Simon M, Johnson S, da Fonseca PC, Morris EP, Wall JS, Lea SM, Blocker AJ. Three-dimensional reconstruction of the Shigella t3ss transmembrane regions reveals 12-fold symmetry and novel features throughout. *Nature Structural & Molecular Biology* 2009; **16**(5):477–485.
111. Fotin A, Cheng Y, Sliz P, Grigorieff N, Harrison SC, Kirchhausen T, Walz T. Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* 2004; **432**(7017):573–579.
112. Feng X, Xia K, Tong Y, Wei G-W. Geometric and topological modeling of macromolecules I: Surface generation, meshing and classification. *Journal of Molecular Graphics and Modelling* submitted 2012.