A Topological Approach to Simplification of Three-dimensional Scalar Functions

Attila Gyulassy, Vijay Natarajan, Valerio Pascucci, *Member, IEEE,* Peer-Timo Bremer, *Member, IEEE,* and Bernd Hamann *Member, IEEE*

***Abstract*— This paper describes an efficient combinatorial method for simplification of topological features in a 3D scalar function. The Morse-Smale complex, which provides a succinct representation of a function's associated gradient flow field, is used to identify topological features and their significance. The simplification process, guided by the Morse-Smale complex, proceeds by repeatedly applying two atomic operations that each remove a pair of critical points from the complex. Efficient storage of the complex results in execution of these atomic operations at interactive rates. Visualization of the simplified complex shows that the simplification preserves significant topological features while removing small features and noise.**

***Index Terms*— Morse theory, Morse-Smale complexes, computational topology, computational geometry, simplification, multiresolution, feature detection, volumetric data.**

I. Introduction

UNDERSTANDING and effective exploration of increasingly complex scientific data necessitates the development of sophisticated schemes that represent data sets at multiple resolutions. Such schemes ideally identify and preserve important features in all resolutions while removing insignificant features at lower resolutions. Beginning with the highest resolution, coarser resolutions are obtained by repeated simplification of an input data set. Identification of a feature and ordering features based on their significance is crucial to the construction of good multiresolution representations. Existing methods typically use a geometric approach, where the numerical error associated with the simplified model is used as the measure of approximation quality. Any removal of features caused by these methods is incidental and not always controlled. We adopt a more direct approach by explicitly identifying features, ordering them, and selectively removing the non-significant ones.

Attila Gyulassy, Vijay Natarajan, and Bernd Hamann are with the Institute for Data Analysis and Visualization (IDAV) and the Department of Computer Science, University of California, Davis.

Valerio Pascucci is with the Center for Applied Scientific Computing (CASC), Lawrence Livermore National Laboratory.

Peer-Timo Bremer is with the Department of Computer Science, University of Illinois, Urbana Champaign.

Related work. Scientific data is usually represented as a set of discrete samples of a function defined on a two- or three-dimensional domain. A continuous function is obtained by interpolating the samples within an underlying mesh of the domain. Several mesh simplification methods have been proposed to support efficient visualization of increasingly large and complex data. These methods typically decimate a mesh by repeated application of a fundamental operation, the most common one being *edge contraction* [1]. The error introduced by edge contraction can be computed as the sum of distances to planes that are associated with end points of a contracted edge [2], and hyperplanes are used to measure the error introduced in the function [3]. Originally developed for surface meshes, edge contraction has been successfully extended to tetrahedral meshes [4]–[6] and other higher-dimensional meshes [7]. Also, related schemes have been developed to select data points and insert them iteratively to define a hierarchical approximation [8].

We are interested in preserving *topological features* of scalar functions, which are defined by critical points of the function. Such features correspond to changes in the behavior of isocontour components. For example, for bivariate functions, upon increasing the function value, minima create new isocontour components, maxima destroy components, and saddle points merge or split isocontour components. A purely geometric approach to simplification is able to remove small topological features but does not provide the desired level of control. Considerable work has been done on topological simplification of scalar functions. Initial work focused on simplifying topological features

or preserving them while simplifying mesh geometry [9], [10]. Two data structures are commonly used for explicitly storing topological features: *Reeb graphs* and *Morse-Smale (MS) complexes*.

The Reeb graph [11] traces components of isocontours/isosurfaces as they sweep the domain. In the case of simply connected domains, the Reeb graph has no cycles and is called a *contour tree*. Reeb graphs, contour trees, and their variants have been used successfully to guide the removal of topological features [12]–[16]. The MS complex decomposes the domain of a function into regions having uniform gradient flow behavior [17]. It has been used recently to perform controlled simplification of topological features in functions defined on two-dimensional domains [18], [19] and for purposes of shape analysis [20]. The MS complex allows the simplification to utilize a global view and spatial distribution of a function for detecting, ordering, and removing features. The MS complex also provides the ability to restrict simplification to a local neighborhood of the non-significant feature. Reeb graph-based simplification methods do not enjoy these benefits. Furthermore, when applied to trivariate functions, they are limited to detecting and simplifying features that are associated with the creation and destruction of isosurface components. These features are represented by pairs of critical points consisting of one saddle and one extremum. The MS complex is also able to detect genus changes within the isosurface, which are represented by saddle-saddle pairs. We use this more comprehensive approach for simplifying scalar functions in three variables.

Simplification of topological features has also been studied in the context of vector fields [21], [22]. These methods, however, use numerical approaches to identify, order, and simplify the topology and are therefore prone to instability. In contrast, Reeb graph and MS complex based methods for simplification of scalar fields are combinatorial in nature and hence stable.

Contributions. We describe a new method for explicit removal of topological features of a given trivariate scalar function with the goal of constructing a hierarchical representation. Specifically, we

- introduce two atomic operations that destroy target topological features without affecting the function globally,
- describe a combinatorial algorithm that selectively removes non-significant topological features by repeated application of the two atomic operations, and
- describe a data structure that stores the MS complex and allows optimal execution of the atomic operations.

We extend the results of Bremer et al. [18] and Edelsbrunner et al. [19] to functions defined on three-dimensional domains by following their basic approach and introducing new ideas to address issues that arise due to the added complexity of features in volumetric domains. The extension is made non-trivial by the presence of two different types of saddles of trivariate functions. For bivariate functions, simplification of the function using the MS complex is accomplished by repeated cancellation of critical point pairs, one of which is a saddle and the other an extremum (minimum/maximum). The presence of a new type of saddle of trivariate functions creates the need for an additional type of cancellation. We introduce two atomic operations that simplify a function: a saddle-saddle cancellation and a saddle-extremum cancellation. Repeated application of these two operations on ordered pairs of critical points results in a hierarchy. Critical point pairs are ordered based on the notion of *persistence* [23], which measures the importance of the associated topological feature. We construct an artificial MS complex from a barycentric subdivision of the input mesh such that each vertex becomes critical. Several of the newly inserted critical points are identified as having an infinitesimally small persistence value and cancelled in a pre-processing stage. In an earlier paper [24], we reported the limitation of this pre-processing to small data sizes. Here, we present two new ideas that remove this limitation, namely

- a streaming approach for the removal of several critical point pairs in the pre-processing stage, resulting in a low memory footprint for the particular case when the scalar function is given on a regular grid, and
- a new data structure that supports efficient implementation of the atomic operations.

We demonstrate our approach by simplifying data sets from various application areas.

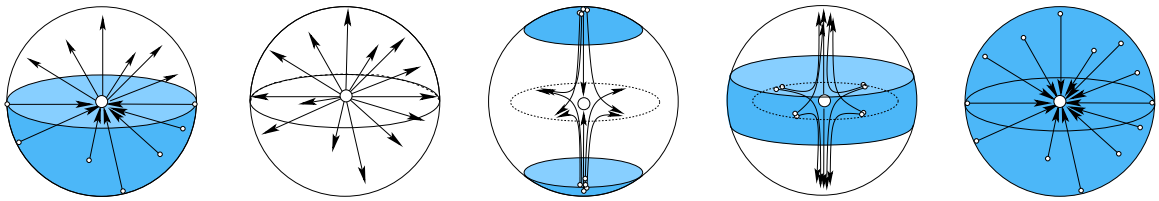


Fig. 1. Local configurations of a regular point and the four types of critical points (minimum, 1-saddle, 2-saddle, and maximum) with shaded oceans, white continents and integral lines. The distinct structure of the oceans allows a combinatorial characterization of the critical points.

II. Morse Theory

Our algorithms are based on Morse theory, which was originally developed for smooth functions. In this section, we provide a brief description of ideas from Morse theory essential to this paper and their extensions to piecewise-linear (PL) functions. For a comprehensive description of the extension, we refer to Edelsbrunner et al. [25]. Morse theory is discussed in detail by Milnor [26] and Matsumoto [27].

Critical points. Let $\mathbb{M}$ be a compact 3-manifold and $f : \mathbb{M} \to \mathbb{R}$ be a real-valued smooth function defined on $\mathbb{M}$. A function f is a *Morse function* if none of its critical points are degenerate (*i.e.*, the Hessian of f is non-singular for all critical points) and no two critical points have the same function value. The *Morse Lemma* states that a Morse function has quadratic behavior within a local neighborhood of every non-degenerate critical point p. This lemma immediately characterizes critical points. The *index* of a critical point is equal to the number of negative eigenvalues of the Hessian. Therefore, minima, 1-saddles, 2-saddles, and maxima have indices equal to $0, 1, 2$, and 3, respectively. Figure 1 shows local neighborhoods of a regular point and the four types of non-degenerate critical points. These local configurations indicate that the criticality of p is characterized by the structure of *oceans* consisting of points x on an infinitesimally small sphere around p, where $f(x) < f(p)$, and *continents* consisting of points x on the sphere, where $f(x) > f(p)$.

MS complex. An integral line of f is a maximal path in $\mathbb{M}$ whose tangent vectors agree with the gradient of f at every point of the path. Each integral line has a natural origin and destination coinciding with critical points of f where the gradient becomes zero. *Ascending manifolds* are obtained as clusters of integral lines having common origin, and *descending manifolds* are obtained as clusters of integral lines having common destination. Ascending and descending manifolds are dual to each other: negating the function transforms one to the other. Descending manifolds have a dimension equal to the index of the destination, and ascending manifolds have a dimension equal to three minus the index of their origin. The descending manifolds of dimension one and two are called *descending arcs* and *descending disks*; the ascending manifolds of dimension one and two are called *ascending arcs* and *ascending disks*. We consider Morse-Smale functions f whose ascending and descending manifolds intersect only transversally. For example, if an ascending disk with origin p intersects a descending disk with destination q, then the points of intersection form a simple path connecting the two saddles p and q. If an ascending disk intersects a descending arc then the point of intersection is a single point, namely the 1-saddle lying on the boundary of the descending arc. The MS complex partitions $\mathbb{M}$ by clustering integral lines that share both common origin and destination. For example, the three-dimensional cells of the MS complex cluster integral lines that originate at a given mini-

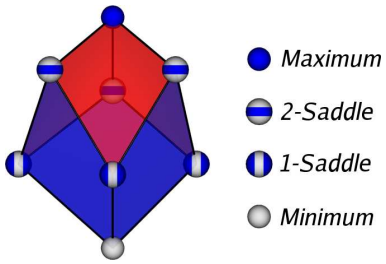


Fig. 2. The boundary of a crystal in the Morse-Smale complex consists of lower-dimensional cells: quads, arcs, and nodes. Every crystal has a unique origin and destination node, the minimum and maximum, respectively, which are end points of integral lines lying within. Note that the glyphs used for critical points indicate their local neighborhood, with shaded oceans and white continents. Quads contained in ascending disks are colored in red; those contained in descending disks are colored blue.

mum and terminate at an associated maximum (see Figure 2). The cells of decreasing dimensions are called *crystals, quads, arcs,* and *nodes,* respectively. The MS complex is equivalently obtained as an overlay of ascending and descending manifolds, which individually partition $\mathbb{M}$. Cells of the MS complex satisfy several combinatorial properties: end points of arcs are critical points whose indices differ exactly by one; quads contain exactly four arcs on their boundary; and the boundary of a crystal contains a collection of quads, arcs, saddles and exactly one minimum and one maximum.

Cancellation. A minimal Morse function is generated from f by repeated cancellation of pairs of critical points. This operation is legal (*i.e.*, it can be realized by a local perturbation of the gradient vector field) if the indices of the two critical points differ by one and they are connected by a unique common arc in the MS complex. For Morse functions defined on three-dimensional domains, we have three types of legal cancellations: minimum and 1-saddle; 1-saddle and 2-saddle; and 2-saddle and maximum. Cancellations play a crucial role in Morse theory for proving important results, including the generalized Poincaré conjecture for higher dimensions [28]. We use cancellations to reduce the number of critical points and hence remove topological features. The local change in the MS complex indicates a smoothing of the gradient vector field and therefore a smoothing of the function f. The ordering of critical point pairs is specified by their persistence value, which quantifies the importance of the topological feature associated with a pair. The persistence of a critical point pair is defined as the absolute difference in value of f between the two points.

PL functions. Scientific data is typically given as a discrete sample over a smooth manifold domain. The smooth manifold is often discretized and represented by a triangulation. Given values of a function f at vertices of a triangulation, we linearly interpolate within edges, triangles, and tetrahedra of the triangulation to obtain a PL function. The spherical neighborhood of a vertex is represented by a two-dimensional triangulation, which contains the oceans and continents. The number of connected components and holes in the ocean uniquely identify the index of the critical point. Gradients, and hence integral lines, are not well defined for PL functions. However, monotonic curves and surfaces corresponding to arcs and quads of the MS complex can be constructed by simulating a separation between integral lines that merge [25]. The function f has its critical points at the nodes of this complex and is monotonic within all arcs, quads, and crystals of the MS complex. We work with this decomposition of the triangulation to determine feature-identifying pairs of critical points as the boundary nodes of a common arc.

III. Data Structure

We introduce a new data structure for storing the MS complex. The design of the data structure is governed by two major objectives: efficient execution of all simplification operations and minimal memory overhead. The data structure stores two types of information: connectivity of the complex and geometry of each cell within the complex. The combinatorial structure of the complex is determined by the connectivity of nodes via arcs, and the geometric structure of the complex is given by the location of nodes, arcs, and all ascending/descending 2- and 3-manifolds. We store the connectivity of the complex as a graph, and augment the graph with geometric components.

Connectivity. We create a list of nodes and a list of arcs to store the connectivity of the MS complex. Each node contains its index of criticality, a list of arcs that originate or terminate at this node, the function value at this point, and relevant geometric information. A design goal is to have fixed-size elements in the lists. Instead of storing a list of arcs incident at a node, we store a reference to exactly one such arc. All arcs have a reference each to the next arc that shares its end points. The list of arcs incident at a specific node is obtained by traversing these references. Besides these two references, each arc also has references to its two end point nodes and geometry. Figure 3 shows our data structure. All nodes of the complex are stored in an array, each element containing multiple fields. G0 refers to the geometric location of a node. The TAG field stores the index of criticality as well as internal flags to represent boundary conditions. If the node is a saddle, then the field G2 stores a reference to the geometry of the 2-manifold originating from that saddle; otherwise, the critical point is an extremum,

in which case the field G3 stores volumetric information of the 3-manifold originating from the extremum. The field A stores a reference to the first arc in the list of arcs incident to this node. Arcs are stored in an array as well, each entry again containing multiple fields. G1 is a reference to the geometry of the arcs (i.e., the integral line), CP1 and CP2 are the two end points of the arc, and A1 and A2 store the next element in the list of arcs incident at CP1 and CP2, respectively.

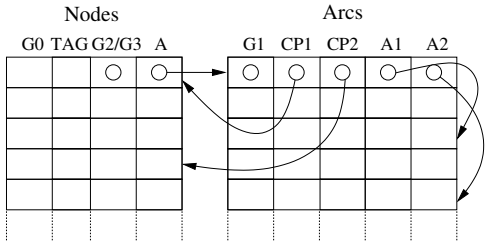


Fig. 3. Data structure for connectivity of nodes and arcs in the MS complex. Nodes and arcs are stored in two arrays. Each entry in these two arrays contains multiple fields that together store the connectivity of the MS complex.

Geometry. We design the geometry components to minimize complexity of cancellations. These components are:

- G0 is a reference to the location of a node, identified by an index.
- G1 is a reference to the path of the arcs, a set of line segments.
- G2 is a reference to the geometry of the ascending/descending 2-manifolds.
- G3 is a reference to the geometry of the ascending/descending 3-manifolds.

G0 is an index referring to an array of input vertices that determines the geometry of a critical point. Storing the geometry of arcs and ascending/descending 2-manifolds is more involved since we want to minimize storage. The key observation is that, upon simplification, arcs and 2-manifolds often merge, and the same line segment or face becomes a part of multiple arcs. This is a common behavior in the pre-processing stage of our simplification algorithm. After a cancellation, several arcs and surfaces are re-routed to pass through the same geometry. Our data structure takes advantage of this redundancy. G1 and G2 are stored as directed acyclic graphs (DAGs), as shown in Figure 4. All leaves of the DAG reachable from a given element in the DAG together constitute the geometry associated with that element. An arc references exactly one element in the DAG to recover the complete geometry of the constituent line segments. The leaves of this DAG are the geometric primitives of the line segments. Similarly, the leaves of the DAG storing the geometry of 2-manifolds contain the faces that compose the 2-manifolds. Each element in the DAG also stores the number of elements that directly reference it. The element is deleted when this counter is zero.

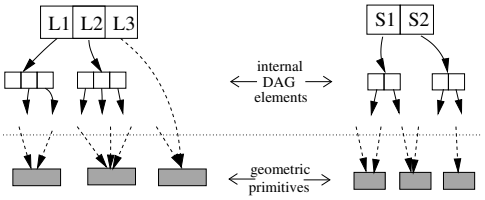
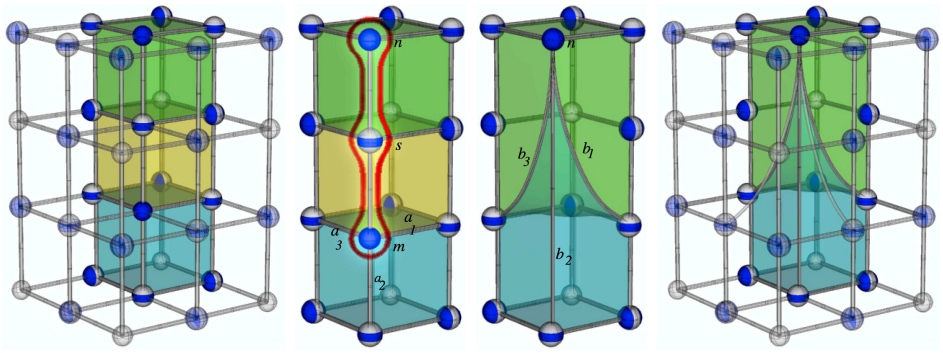


Fig. 4. Geometry of arcs and ascending/descending disks are stored as DAGs to minimize storage. *Left*: our simplification algorithm creates new arcs by merging exactly three other arcs. So, each internal element stores references to three children. *Right*: after a cancellation, two disks merge into one. So, each internal element stores references to two children.

IV. Simplification

We simplify the MS complex of a Morse function f by performing a series of critical point pair cancellations. A cancellation simulates a smoothing operation applied to f by modifying gradient flow in the neighborhood of two critical points. Arcs connecting critical points are lines of steepest descent or ascent, and changing them affects the gradient flow behavior of f. Rules that apply to gradient flow must be adhered to in the simplification process. For example, integral lines must remain disjoint.

Critical point pairs that we consider are end points of an arc in the MS complex and therefore have consecutive indices. We group the pairs into two types: saddle-extremum (indices 1 and 0 or indices 2 and 3) and saddle-saddle (indices 1 and 2). The two types of cancellations are distinct in the way they modify gradient flow behavior. The cancellation procedure is similar to vertex removal used in mesh simplification methods, with a pair of critical points being removed instead of a single vertex, and the reconnection of the complex governed by rules of Morse theory rather than mesh geometry. For reasons of clarity, we illustrate the two types of cancellations using prototypical figures



(a) (b) (c) (d)

Fig. 5. Snapshot of a Morse-Smale complex before and after a saddle-maximum cancellation. (a) Three of the twelve crystals affected by the cancellation are colored. (b) Close-up view of the three crystals and the maximum-saddle-maximum triple that is merged into a single maximum. (c) After cancellation, all ascending arcs and disks that originally flow into the maximum lying below the saddle, flow into the maximum lying above the saddle. One of the three crystals is deleted in this process. (d) Ascending arcs and disks in the other nine crystals are similarly re-routed, resulting in a removal of three more crystals.

of MS complexes. The description, however, holds for all possible configurations.

Saddle-extremum cancellation. The saddle-extremum cancellation removes either a 2-saddle-maximum pair or a 1-saddle-minimum pair. The two pairs are dual to each other as can be seen by negating the function: maxima become minima, 2-saddles become 1-saddles and vice-versa. We restrict our discussion to 2-saddles and maxima. A 2-saddle, by definition, is connected by ascending arcs to exactly two maxima. When one of these maxima is removed in a saddle-maximum cancellation, integral lines ending at this maximum flow toward the second maximum. One can think of a saddle-maximum cancellation as merging three critical points into one maximum. Applying the saddle-maximum cancellation simplifies the function by removing a "bump." Figure 5 shows how the integral lines terminating at the two maxima flow into the remaining maximum after cancellation.

The saddle-maximum cancellation is similar to its two-dimensional analog, which can also be interpreted as merging three critical points. We merge neighboring cells in the ring around the saddle-maximum arc. Therefore, besides removing two critical points, this cancellation also removes several crystals, quads, and arcs from the complex. Let s and m be the saddle and extremum to be cancelled and a be the arc connecting them. We implement the cancellation by performing the following sequence of connectivity-modifying operations:

1) Let n be the end point of the arc originating from s and not equal to m. Let b be this arc that connects s to n.
2) Replace all arcs a_i that terminate at m with new arcs b_i that share their origin with a_i and terminate at n. Add b_i to the arc list of its origin and the arc list of n.
3) Delete nodes s and m and all arcs incident to them. Delete all references of these arcs from the arc lists of their endpoints.

The geometry of arc b_i is obtained as the concatenation of arcs a_i, a, and b. References to the geometry of these arcs are added from a new element added to the DAG. All arcs b_i share the geometry of a and b, which makes the DAG an efficient representation of the geometry. The single 2-manifold corresponding to s is deleted from the complex along with the saddle. Re-direction of arcs as described above changes boundaries of the 2-manifolds that it touches. However, it does not change their surface geometry, and the interior of the ascending/descending 2-manifolds remain simply connected after the cancellation. Therefore, we make no changes to the representation of the surface geometry of these 2-manifolds.

A saddle-maximum cancellation is legal only if the 2-saddle is connected to two distinct maxima. If this condition is not met, then we recognize that the cancellation causes a *strangulation* of the descend-

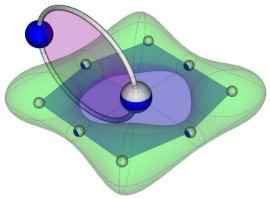
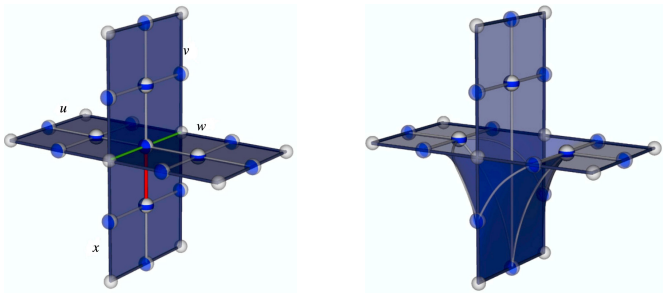


Fig. 6. A saddle-maximum pair that cannot be cancelled. Two integral lines beginning at a 2-saddle flow to the same maximum. Cancelling the saddle-maximum pair causes a strangulation of the blue descending disk since integral lines terminating at the 2-saddle are left without a destination.



(a) (b)

Fig. 7. A saddle-saddle cancellation. (a) Descending disks affected by the cancellation. The red arc connects the pair to be cancelled. All four disks (u, v, w, and x) have two common descending arcs (shown in green) on their boundary, both originating from the 1-saddle to be removed. (b) Descending disks that remain after cancellation. The green descending arcs are deleted from the boundary of the three surviving disks, which now extend to inherit the boundary of x. Ascending manifolds are dual to the descending manifolds and hence modified in a similar way.

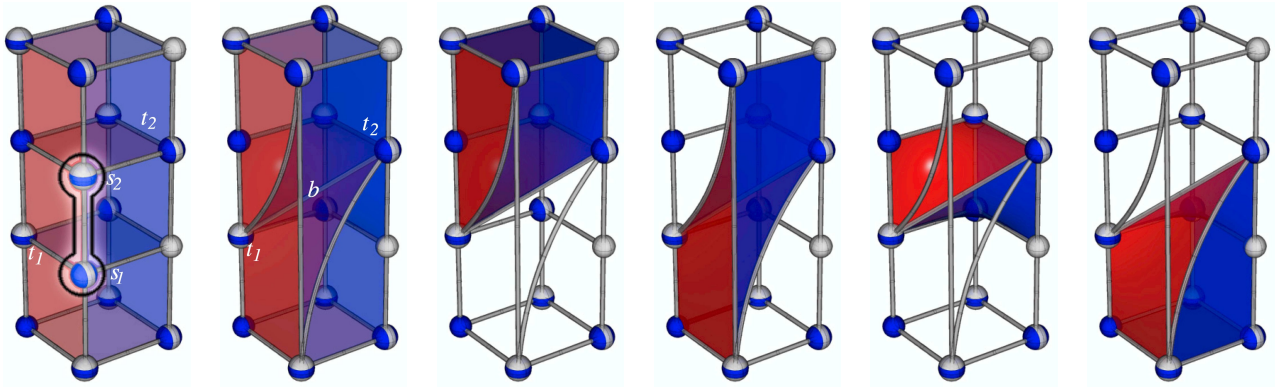
ing disk that originates at the 2-saddle. In fact, it is not possible to route the integral lines terminating at the 2-saddle if we cancel such a saddle-maximum pair. Figure 6 shows this configuration.

Saddle-saddle cancellation. The saddle-saddle cancellation removes a 1-saddle-2-saddle pair, see Figures 7 and 8. This cancellation does not have an analog in lower dimensions. A 1-saddle descends to exactly two minima, and a 2-saddle ascends to exactly two maxima. After cancelling this saddle pair, we need to ensure that the two pairs of extrema originally separated by these saddles remain that way. This constraint necessitates the introduction of new cells to fill in space between the two pairs of extrema. One can think of this cancellation by considering what happens to the descending disk originating from the 2-saddle and the ascending disk originating from the 1-saddle. After cancellation, these two disks disappear, and neighboring disks stretch out and share their boundary. We can no longer consider the cancellation as merging three critical points as we can for the saddle-maximum cancellation. Consider the descending disk that is removed by the cancellation: The boundary of this disk consists of alternating 1-saddles and minima. Arcs lying within the disk connect the 1-saddles on the boundary to the 2-saddle where the disk originates. One of these 1-saddles is involved in the cancellation. This 1-saddle and its two descending arcs are deleted as a result of the cancellation. Descending disks that contain these descending arcs in their boundary expand to share the boundary of the removed disk. Similarly, one ascending disk is removed and its boundary is shared by the neighboring ascending disks. Figure 7 illustrates the operation by showing the descending disks before and after cancellation.

A good way to think about reconnecting the complex after a saddle-saddle cancellation is in terms of ascending and descending disks: All surviving descending disks expand to share the boundary of the deleted disk, thereby creating connections between surviving 2-saddles and 1-saddles on the newly inserted boundary. Similarly, surviving 1-saddles connect to 2-saddles on the newly inserted boundary of their ascending disks, which guarantees full reconnectivity of the complex after a cancellation. Let s_1 and s_2 be a 1-saddle and 2-saddle pair connected by an arc a. We implement the cancellation as a sequence of operations on our data structure:

1) For each pair of 1- and 2-saddles (t_1, t_2) different from (s_1, s_2), where the pair (t_1, s_2) is connected by arc a_1, and the pair (t_2, s_1) is connected by arc a_2, do the following:
   1.1) Create a new arc b from t_1 to t_2.
   1.2) Insert b into the arc lists of its endpoints.
2) Delete all arcs in the arc lists of s_1 and s_2 and delete the two nodes.

New arcs b are created by concatenating the geometry of arcs a_1, a, and a_2. The 2-manifolds of all 1-saddles t_1 merge with the 2-manifold of s_1, and the 2-manifolds of all 2-saddles t_2 merge with the 2-manifold of s_2. Deleting s_1 and s_2 removes a reference to their 2-manifolds. However no geometry needs to be removed since this surface geometry



(a) (b) (c) (d) (e) (f)

Fig. 8. Cells destroyed and created by a saddle-saddle cancellation. The neighborhood of the saddle-saddle pair is divided into four regions of three crystals each. Only one region is shown for reasons of clarity. Cells within the other regions are modified in a similar manner. (a) Three crystals before cancellation of the highlighted pair of saddles. (b) Four crystals after cancellation. (c) Top crystal stretching down to the lower 2-saddle. (d) A crystal defined by a new minimum-maximum pair stretching from lower-left to upper-right corner of the region. (e) Middle crystal shrinking after losing the saddle-saddle pair. (f) Bottom crystal stretching up to the 1-saddle in the upper right.

has become part of many ascending/descending 2-manifolds. When two surfaces merge, we create a new element in the DAG, add references to the merged surfaces, and include a reference to the new element from the saddle.

The MS complex actually gains cells after a saddle-saddle cancellation since re-routing ascending and descending disks creates new intersections between them, thus adding new cells to the complex. Figure 8 shows the cells destroyed and created by this operation. For simplicity, these figures show only three of the twelve crystals destroyed by the cancellation and the cells that reconnect this complex within this region. Introducing new cells is counter-intuitive. Although the MS complex grows in size, the function is smoothed by the removal of saddle pairs. Also, the cells created by the saddle-saddle cancellation are introduced into rings around saddle-extremum pairs. A future saddle-extremum cancellation will remove all these cells, leading to a smoother Morse function and, ultimately, a smaller MS complex. A simple proof that the algorithm terminates, despite the increase in number of cells after a saddle-saddle cancellation, follows from the fact that every cancellation results in the removal of a pair of nodes from the complex.

In theory, the maximum number of cancellations is bounded by half the number of critical points. In practice, however, we stop earlier to preserve more persistent features. The saddle-saddle cancellation introduces significant complexity that does not occur in lower dimensions. Furthermore, the complexity of this cancellation indicates that extensions to higher dimensions are likely to be rather complicated.

Fig. 9. A saddle-saddle pair that cannot be cancelled. Cancelling the pair would create a crystal, called a pouch, that contains exactly one minimum and one maximum as boundary nodes. The boundary of such a pouch cannot be represented as a collection of quads.

An arc connecting two saddles does not guarantee a valid saddle-saddle cancellation. Performing a cancellation could lead to the formation of a *pouch*, see Figure 9. This situation occurs when a crystal incident on the arc has exactly two quads, one connecting the saddle pair to a minimum and the other connecting them to a maximum. Removing the 1-saddle and 2-saddle creates a crystal with zero saddles and zero quads, corresponding to a possibly valid Morse function but resulting in an invalid combinatorial structure for the MS complex. Pouches and strangulations (in the case of saddle-extremum cancellation) do occur in practice, which implies that there exists a minimal MS complex that cannot

be further simplified using legal cancellations.

V. MS Complex Construction

We now describe issues related to the construction of MS complexes from a given data set.

Artificial MS complex. Given a PL function defined on a volumetric mesh, we first construct an artificial MS complex by inserting dummy critical points at the barycenters of all cells having an index of criticality equal to the dimension of the cell. All input data points become local minima. We add new arcs connecting the barycenter of a cell with the barycenters of its faces. The dummy vertices and arcs naturally subdivide each simplex into arcs, quads, and crystals. Minima inherit the function value from the corresponding data point. The function values at the dummy nodes are chosen to be greater, by an infinitesimally small value ε, than the largest value of the vertices of the cell in which it is located. Initial cancellations of ε-persistence critical pairs in this artificial complex remove dummy nodes. A barycentric subdivision when applied to a cube leads to an artificial complex with minima at the originally given data locations. We begin with this artificial complex when input data is given on a cube lattice. Figure 10 illustrates how we subdivide a tetrahedron and a cube.

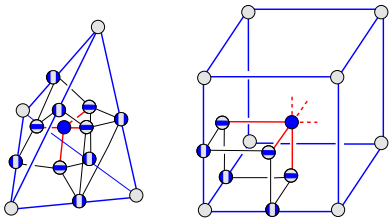


Fig. 10. Creation of an artificial MS complex by subdivision: Dummy critical points are introduced at barycenters of all cells thereby turning original data points into local minima. The function value at a dummy critical point is ε (an infinitesimally small positive value) greater than the largest value of the vertices of the cell in which it is located. Therefore, the dummy critical point has infinitesimally small persistence.

We choose to construct the complex in this way because of its simplicity as opposed to the intricate algorithm described by Edelsbrunner et al. [25]. A disadvantage of starting with this artificial complex is that the number of nodes is equal to the total number of cells in the input mesh, thereby limiting the size of the data set that can be efficiently processed. However, dummy nodes are removed in a pre-processing step, and therefore exploration of the data can still be done interactively. Besides its simplicity, another advantage of the artificial complex is that it automatically resolves the issue of degenerate critical points. *Multiple saddles*, which are typically unfolded into simple 1- and 2-saddles, are not present in the artificial complex. This follows from the fact that each 1-saddle, at the barycenter of an edge, now has exactly two descending arcs connecting it to minima at the end points of the edge. Similarly, all 2-saddles are simple since there are exactly two ascending arcs that connect each one of them to local maxima. Both types of cancellation do not change the number of extrema that connect to a saddle and hence do not introduce any multiple saddles. Multiple saddles can exist in input data sets, but our construction splits these into simple saddles.

Streaming. Pre-processing is costly. It creates seven additional nodes (for cube lattices) in the complex for each data point. Thus, our earlier experiments [24] were restricted to data sizes up to $64 \times 64 \times 64$. We have overcome this restriction by utilizing the regularity of the artificial complex to pre-process the input in a streaming fashion. After removal of ε-persistence pairs, the size of the complex is proportional to the number of critical points in the original data. We create the complex incrementally: We start by adding a couple of slices of the artificial complex, and then cancel all ε-persistence pairs before adding another slice. We do not permit cancellations that affect the slice in order to maintain the regularity of the artificial complex. However, critical point pairs whose neighborhoods are disjoint from an incoming slice are valid pairs for cancellation. Figure 11 shows a snapshot of a complex constructed in this incremental fashion. The streaming approach has several advantages.

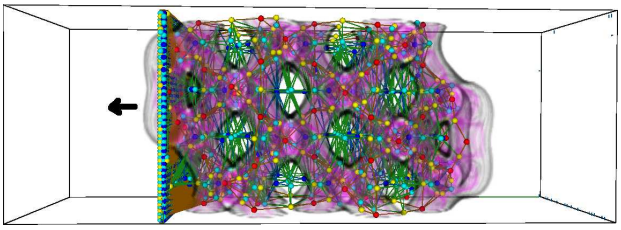


Fig. 11. A snapshot of the artificial complex being constructed one-slice-at-a-time. The unprocessed region is to the left of the slice.

Most importantly, the size of the memory footprint

is $O(n^2+k)$ instead of $O(n^3)$, where n is the leading dimension of the data set and k is the size of the MS complex. We only need to store two full slices in addition to the complex at any time, as opposed to storing the entire artificial complex. Computation of the complex for larger data sets becomes possible via such a streaming method. However, we still pay the $O(n^3)$ time penalty, since all dummy critical points must be removed. The other factor influencing pre-processing time is the maximum number m of arcs incident on any node. Since this is stored as a linked list, search and deletion of arcs in this list cost $O(m)$ time, and can pose serious overhead. Even though the streaming algorithm performs the same number of cancellations as before [24], our use of a more efficient data structure leads to an 8x speedup.

Boundary. Our input data is defined over a volumetric domain that has a non-empty boundary. We use a standard technique from point set topology, called *one-point compactification*, to convert the domain into a 3-manifold without boundary. The compactification involves addition of a vertex at infinity that connects to all boundary vertices. This extension of the triangulation is a simple and efficient way to handle the domain boundary [5], [15], [19]. Instead of explicitly adding the vertex at infinity and simplices connecting it to the boundary, we create a layer around the domain consisting of dummy critical points for each simplex that contains the vertex at infinity. These dummy critical points become nodes of the artificial complex and are removed when we cancel ε-persistence pairs. We restrict all cancellations to pairs that lie completely in the interior of the domain or within the boundary to ensure that we do not change the topology of the domain.

VI. Results

We pre-process the input data by first creating an artificial complex and then removing several dummy critical points by cancelling ε-persistence pairs. This process loads the input one slice at a time to construct the MS complex incrementally. We perform further simplification of the MS complex in an interactive process to identify features. Once the dummy critical points are removed, the complex provides an efficient representation of features in the original data. We identify important features as regions associated with persistent critical points. Similar to Takahashi et al [15], we automatically design a transfer function to enhance critical values that correspond to these features. Our simplification allows us to limit the number of critical values affecting the transfer function to exactly those representing important features.

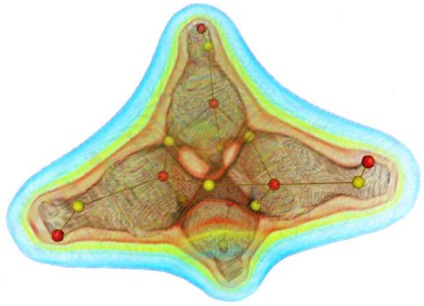


Fig. 12. Atoms and bonds in the C_4H_4 molecule are identified by high-persistence critical points and ascending arcs in the simplified Morse-Smale complex.

Feature identification. We show that our simplification technique extracts known features and removes noise in well-studied data sets. Table I lists the data sets used, their sizes, and the time needed for pre-processing. All experiments were performed on a desktop PC (3GHz Pentium 4, 1GB RAM). The reduction in memory required for the pre-processing stage allows us to handle large data sets when compared to prior results [24]. After pre-processing, the remaining critical pairs can be removed interactively. The first data set represents an electron density distribution in a C_4H_4 molecule. Once ε-persistence critical point pairs are removed, the complex correctly outlines the bond structure of the C_4H_4 molecule. High-persistent maxima correspond to locations of atoms in the molecule, and ascending arcs that connect 2-saddles with these maxima correspond to bonds between atoms, see Figure 12. This correspondence is a visual depiction of the topological approach to identifying atoms in molecules as proposed by the AIM theory [29].

The other data sets listed in Table I were obtained either from simulations or from x-ray scans. Visualization of our results for these data sets are shown in Figures 13 and 14. The hydrogen atom data set describes the spatial electron density in a hydrogen atom subjected to a large magnetic field. The data set exhibits high density around the nucleus, two regions of high density on either side, and a torus

TABLE I
DATA SETS USED IN EXPERIMENTS, THEIR SIZE, AND TIMING RESULTS FOR PRE-PROCESSING (t_{pp}).

Data set	Size	t_{pp}
C_4H_4	$33 \times 33 \times 33$	5s
Silicium	$34 \times 34 \times 98$	15s
Neghip	$64 \times 64 \times 64$	2m 35s
Aneurysm	$128 \times 128 \times 128$	30m
Hydrogen	$128 \times 128 \times 128$	45m

of high density around the nucleus. We correctly identify these features. After the ε-persistence critical points are removed, there are still two disks of saddles separating the maxima in the data set, which correspond to noise. Initial cancellations remove the spurious 1-saddles and 2-saddles, leaving behind four maxima, which represent the four regions of high electron density.

Distinctive features in the silicium, neghip, and aneurysm data sets are revealed using a low threshold. We observed that a threshold value of 10% of the maximum persistence suffices to detect and remove all insignificant features. This can clearly be seen in the neghip data set, where we are able to isolate the different clusters of atoms automatically. We do not, however, expect that the same threshold should be used for all types of data sets. Rather, we view the simplification as an interactive filter that will allow scientists to explore the data at different scales.

Noise removal. We use synthetic data to illustrate how the cancellation of critical point pairs removes noise in a natural manner and hence leads to a robust identification of features. Our input is a simple sum of Gaussian distributions that decrease radially from seed points. Each radial function contributes to a local maximum at its seed point. The local maximum at the center of the domain is the largest and those near the domain boundary are the smallest. Figure 15 shows how the function is successively simplified by removing critical points based on persistence. The smaller maxima represent small perturbations in the data and hence represent noise. The corresponding maxima have low persistence and are removed early during the simplification process. Negligible local maxima are removed first followed by the next tier of maxima with small but considerable persistence, leaving behind the primary feature/maximum at the center.

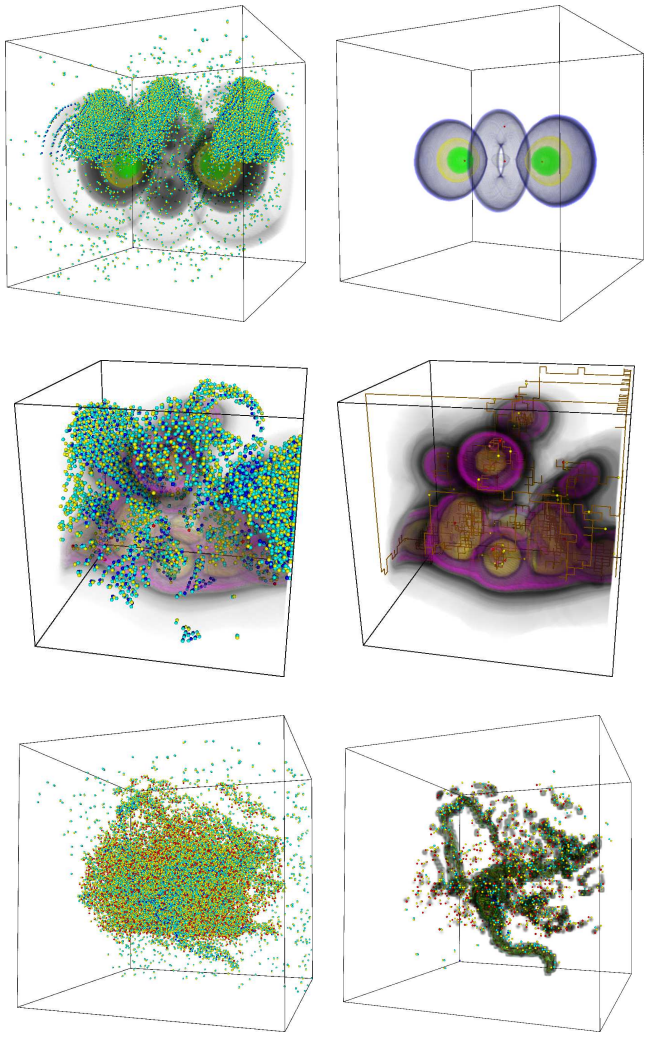


Fig. 13. Topological simplification applied to various data sets (top to bottom: hydrogen, neghip, and aneurysm). The input (left) has a large number of critical points, several of which are identified as being insignificant and removed by repeated application of two atomic operations. Features are identified by the surviving critical points and enhanced in a volume-rendered image, using an automatically designed transfer function (right).

VII. CONCLUSIONS

We have described an algorithm to simplify a trivariate Morse function by cancelling pairs of critical points in its MS complex and demonstrated its application to the identification of features in volumetric scalar fields. This topological approach supports a direct manipulation of features including their detection, ordering, and removal during interactive data exploration. The combinatorial nature of our algorithm leads to a robust and efficient implementation and hence allows us to perform topological analysis on complex and noisy data sets. We use the notion of persistence to define

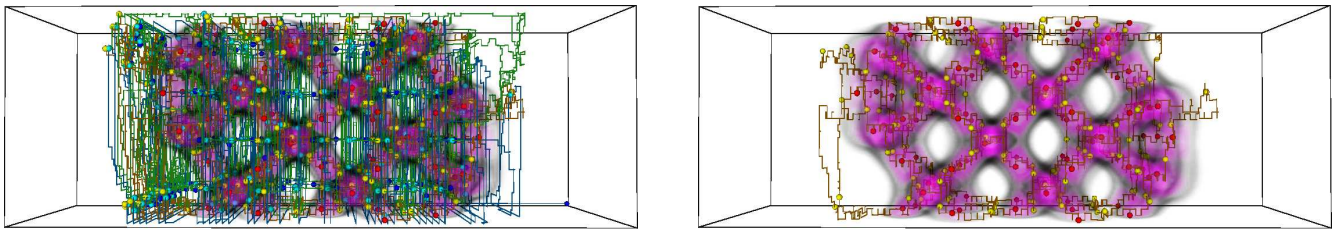


Fig. 14. Topological simplification of the silicium data set. Cancellation of low-persistence critical pairs reveals the shape of the silicium lattice structure.

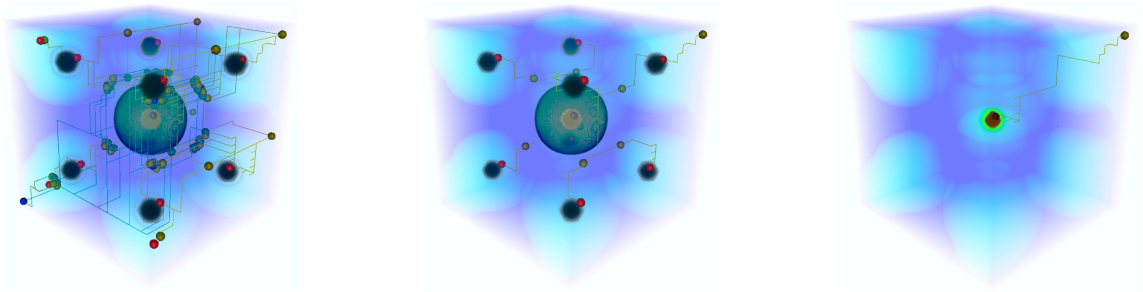


Fig. 15. Noise in a synthetic function introduces features with negligible persistence. Left: The function consists of various spikes with the central one being the largest. Each spike is visualized as a sphere in the volume-rendered image. Middle: All nine spikes are clearly visible after removing noise that created the thin shells surrounding the spheres. Right: Further simplification destroys all maxima except the one representing the crucial features.

the importance of critical points and the features that they represent. This measure has been proven to be robust under the presence of noise [30]. Other spatial measures have also been successfully used to define and detect features [12]. We plan to incorporate such measures into our framework and perform a comparative study. Compared to earlier work [24], we can handle larger data sets since memory overhead is no longer a concern during the pre-processing phase. However, computing the MS complex remains a bottleneck. An algorithm that schedules ε-persistence pairs for cancellation while penalizing the creation of nodes with high valence, would decrease the pre-processing time significantly. Future work will also be directed at designing a multiresolution representation for continuous trivariate functions by performing independent cancellations in the MS complex. A numerical realization of the cancellations discussed in this paper is necessary for the design of a multiresolution representation. As a first step toward this goal, we plan to investigate different methods to create and represent efficiently the geometry of the ascending/descending manifolds consistent with the integral lines. We are also exploring different ideas for effective visualization of the quads and crystals. We are currently not able to utilize them efficiently in the exploration process due to visual clutter.

ACKNOWLEDGMENTS

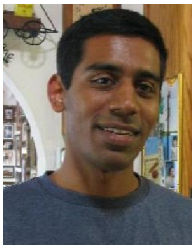
We thank David Cohen-Steiner for suggesting the construction of an artificial MS complex and members of the Visualization and Computer Graphics Research Group at the Institute for Data Analysis and Visualization (IDAV), UC Davis, for various discussions. The C_4H_4 data set is available at http://www.chemistry.mcmaster.ca/aimpac. Other data sets used in this paper are available at http://www.volvis.org. This work was supported by the National Science Foundation under contract ACI 9624034 (CAREER Award), through the Large Scientific and Software Data Set Visualization (LSSDSV) program under contract ACI 9982251, and a large Information Technology Research (ITR) grant; and the National Institutes of Health under contract P20 MH60975-06A2, funded by the National Institute of Mental Health and the National Science Foundation. This work was performed under the auspices of the U.S. Department of Energy by University of California Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

REFERENCES

[1] H. Hoppe, “Progressive meshes,” in *Proc. SIGGRAPH*, 1996, pp. 99–108.
[2] M. Garland and P. S. Heckbert, “Surface simplification using quadric error metrics,” in *Proc. SIGGRAPH*, 1997, pp. 209–216.
[3] ——, “Simplifying surfaces with color and texture using quadric error metrics,” in *Proc. IEEE Conf. Visualization*, 1998, pp. 263–269.
[4] P. Cignoni, D. Constanza, C. Montani, C. Rocchini, and R. Scopigno, “Simplification of tetrahedral meshes with accurate error evaluation,” in *Proc. IEEE Conf. Visualization*, 2000, pp. 85–92.
[5] V. Natarajan and H. Edelsbrunner, “Simplification of three-dimensional density maps,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 10, no. 5, pp. 587–597, 2004.
[6] O. G. Staadt and M. H. Gross, “Progressive tetrahedralizations,” in *Proc. IEEE Conf. Visualization*, 1998, pp. 397–402.
[7] M. Garland and Y. Zhou, “Quadric-based simplification in any dimension,” *ACM Transactions on Graphics*, vol. 24, no. 2, pp. 209–239, 2005.

[8] B. Hamann and J.-L. Chen, "Data point selection for piecewise trilinear approximation," *Computer Aided Geometric Design*, vol. 11, no. 5, pp. 477–489, 1994.
[9] Y.-J. Chiang and X. Lu, "Progressive simplification of tetrahedral meshes preserving all isosurface topologies," *Computer Graphics Forum*, vol. 22, pp. 493–504, 2003.
[10] T. Gerstner and R. Pajarola, "Topology preserving and controlled topology simplifying multiresolution isosurface extraction," in *Proc. IEEE Conf. Visualization*, 2000, pp. 259–266.
[11] G. Reeb, "Sur les points singuliers d'une forme de Pfaff complètement intégrable ou d'une fonction numérique," *Comptes Rendus de L'Académie ses Séances, Paris*, vol. 222, pp. 847–849, 1946.
[12] H. Carr, J. Snoeyink, and M. van de Panne, "Simplifying flexible isosurfaces using local geometric measures," in *Proc. IEEE Conf. Visualization*, 2004, pp. 497–504.
[13] I. Guskov and Z. Wood, "Topological noise removal," in *Proc. Graphics Interface*, 2001, pp. 19–26.
[14] S. Takahashi, G. M. Nielson, Y. Takeshima, and I. Fujishiro, "Topological volume skeletonization using adaptive tetrahedralization," in *Proc. Geometric Modeling and Processing*, 2004, pp. 227–236.
[15] S. Takahashi, Y. Takeshima, and I. Fujishiro, "Topological volume skeletonization and its application to transfer function design," *Graphical Models*, vol. 66, no. 1, pp. 24–49, 2004.
[16] Z. Wood, H. Hoppe, M. Desbrun, and P. Schröder, "Removing excess topology from isosurfaces," *ACM Transactions on Graphics*, vol. 23, no. 2, pp. 190–208, 2004.
[17] S. Smale, "On gradient dynamical systems," *Ann. of Math.*, vol. 74, pp. 199–206, 1961.
[18] P.-T. Bremer, H. Edelsbrunner, B. Hamann, and V. Pascucci, "A topological hierarchy for functions on triangulated surfaces," *IEEE Transactions on Visualization and Computer Graphics*, vol. 10, no. 4, pp. 385–396, 2004.
[19] H. Edelsbrunner, J. Harer, and A. Zomorodian., "Hierarchical Morse-Smale complexes for piecewise linear 2-manifolds," *Discrete and Computational Geometry*, vol. 30, no. 1, pp. 87–107, 2003.
[20] F. Cazals, F. Chazal, and T. Lewiner, "Molecular shape analysis based upon the Morse-Smale complex and the Connolly function," in *Proc. 19th Ann. ACM Sympos. Comput. Geom.*, 2003, pp. 351–360.
[21] W. de Leeuw and R. van Liere, "Collapsing flow topology using area metrics," in *Proc. IEEE Conf. Visualization*, 1999, pp. 149–154.
[22] X. Tricoche, G. Scheuermann, and H. Hagen, "Continuous topology simplification of planar vector fields," in *Proc. IEEE Conf. Visualization*, 2001, pp. 159–166.
[23] H. Edelsbrunner, D. Letscher, and A. Zomorodian., "Topological persistence and simplification," *Discrete and Computational Geometry*, vol. 28, no. 4, pp. 511–533, 2002.
[24] A. Gyulassy, V. Natarajan, V. Pascucci, P.-T. Bremer, and B. Hamann, "Topology-based simplification for feature extraction from 3d scalar fields," in *Proc. IEEE Conf. Visualization*, 2005, pp. 535–542.
[25] H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci., "Morse-Smale complexes for piecewise linear 3-manifolds," in *Proc. 19th Ann. Sympos. Comput. Geom.*, 2003, pp. 361–370.
[26] J. Milnor., *Morse Theory*. New Jersey: Princeton Univ. Press, 1963.
[27] Y. Matsumoto, *An Introduction to Morse Theory*. Amer. Math. Soc., 2002, translated from Japanese by K. Hudson and M. Saito.
[28] S. Smale, "Generalized Poincaré's conjecture in dimensions greater than four," *Ann. of Math.*, vol. 74, pp. 391–406, 1961.
[29] R. F. W. Bader, *Atoms in Molecules*. Oxford: Clarendon Press, 1995.
[30] D. Cohen-Steiner, H. Edelsbrunner, and J. Harer, "Stability of persistence diagrams," in *Proc. 21st Sympos. Comput. Geom.*, 2005, pp. 263–271.

Attila Gyulassy received the B.A. in computer science and the B.A. in applied mathematics at the University of California at Berkeley. He is currently pursuing the Ph.D. degree in computer science at the University of California at Davis. He holds a student employee graduate research fellowship at the Lawrence Livermore National Laboratory.



Vijay Natarajan is a postdoctoral researcher in the Institute for Data Analysis and Visualization (IDAV) at University of California, Davis. His research interests include scientific visualization, computational geometry, computational topology, geometric modeling and meshing. He received the Ph.D. degree in computer science from Duke University in 2004. He holds the B.E. degree in computer science and M.Sc. degree in mathematics from Birla Institute of Technology and Science, Pilani, India.

Valerio Pascucci received the Ph.D. degree in computer science from Purdue University in May 2000 and the EE Laurea (Master's), from the University "La Sapienza" in Rome in December 1993, as a member of the Geometric Computing Group. He has been a computer scientist and project leader at the Lawrence Livermore National Laboratory, Center for Applied Scientific Computing (CASC) since May 2000. Prior to his CASC tenure, he was a senior research associate at the University of Texas at Austin, Center for Computational Visualization, CS and TICAM Departments. He is a member of the IEEE.

Peer-Timo Bremer is currently of Postdoctoral Research Associate at the University of Illinois, Urbana-Champaign. Peer-Timo obtained his Ph.D. in computer science from the University of California at Davis in 2004 while working as a Student Employee Graduate Research Fellow at the Lawrence Livermore National Laboratory. He also holds a Diploma (M.S.) in mathematics from the University of Hanover, Germany. Peer-Timo is a member of the Association for Computing Machinery (ACM) and the Institute of Electrical and Electronics Engineers (IEEE) Computer Society.

Bernd Hamann serves as an Associate Vice Chancellor for Research at the University of California, Davis, where he is a professor of computer science. From 1997 until 2004, he was a Co-Director of the UC Davis Institute for Data Analysis and Visualization (IDAV). His main interests are visualization, geometric modeling,computer graphics, and virtual reality. Bernd Hamann received a B.S. degree in computer science, a B.S. degree in mathematics, and an M.S. degree in computer science from the Technical University of Braunschweig, Germany. He received a Ph.D. degree in computer science from Arizona State University in 1991. He was awarded a 1992 National Science Foundation Research Initiation Award and a 1996 National Science Foundation CAREER Award.