Multiscale Morse theory for science discovery

Valerio Pascucci and Ajith Mascarenhas

1 Scientific Computing and Imaging Institute, School of Computing, University of Utah, Salt Lake City, UT 84112, USA
2 Center for Applied Scientific Computing, Lawrence Livermore National Laboratory, Livermore, CA 94551, USA
E-mail: pascucci@acm.org

Abstract. Computational scientists employ increasingly powerful parallel supercomputers to model and simulate fundamental physical phenomena. These simulations typically produce massive amounts of data easily running into terabytes and petabytes in the near future. The future ability of scientists to analyze such data, validate their models, and understand the physics depends on the development of new mathematical frameworks and software tools that can tackle this unprecedented complexity in feature characterization and extraction problems. We present recent advances in Morse theory and its use in the development of robust data analysis tools. We demonstrate its practical use in the analysis of two large scale scientific simulations: (i) a direct numerical simulation and a large eddy simulation of the mixing layer in a hydrodynamic instability and (ii) an atomistic simulation of a porous medium under impact. Our ability to perform these two fundamentally different analyses using the same mathematical framework of Morse theory demonstrates the flexibility of our approach and its robustness in managing massive models.

1. Introduction

Accurate analysis of massive space-time scientific simulation data introduces unique algorithmic challenges due to the impossibility of fitting all data into main memory and the multiscale nature and complexity of the features in such data. In this papers we focus on the latter challenge. In particular, we study the development of a sound mathematical framework for feature definition and design robust efficient algorithms for discovering and quantifying trends in the data. We also aim at algorithms that provide the user with the flexibility to control the scale of interest and to produce visualizations and collect statistics on the data.

In this paper, we describe ideas from Morse theory that enable effective feature characterization using the powerful language of topology. In addition, our approach is based on a set of combinatorial constructs that are implemented exactly (no numerical approximations) yielding a highly robust data analysis framework that can segment and quantify features at several scales for massive datasets.

We illustrate the effectiveness of our analysis techniques with two applications. In the first application, we enable the discovery and quantification of four major mixing trends in a fundamental hydrodynamic instability simulation. Our techniques also allowed for the first time a feature-based comparison and validation of simulations developed with two different techniques. The hydrodynamic instability (a Rayleigh-Taylor simulation) was developed at Lawrence Livermore National Laboratory [8, 9]. This analysis is described in greater detail in
Figure 1. MS-complex construction, simplification and topologically valid approximation: (a) Morse function with critical points shown; (b) stable manifolds; (c) unstable manifolds; (d) MS-complex; (e) MS-complex and manifold after simplification. Maxima are solid blue, minima are solid white, and saddles are mixed.

[15, 17]. In the second application, we enable the detection of core structures in the atomistic simulation of a porous medium and their tracking in time as a projectile impacts the medium causing structural failures. This analysis is described in detail in [12]. For a more detailed description of Morse-theory based analysis see [7, 10, 11, 3, 4, 20].

2. Morse theory for robust multi-scale feature analysis
To begin, we present some background from Morse theory [16, 18] and from combinatorial and algebraic topology [1, 19].

Smooth maps on manifolds. Let \( f: \mathbb{M} \rightarrow \mathbb{R} \) be a smooth map. A point \( x \in \mathbb{M} \) is a critical point of \( f \) if the gradient of \( f \) vanishes at \( x \), and the value \( f(x) \) is a critical value. Noncritical points and noncritical values are called regular points and regular values, respectively. A critical point \( x \) is nondegenerate if the Hessian (matrix of second-order partial derivatives) at \( x \) is nonsingular. The index of a critical point \( x \), denoted by \( \text{index} x \), is the number of negative eigenvalues of the Hessian. For \( d = 3 \) there are four types of nondegenerate critical points: the minima (index 0), the 1-saddles (index 1), the 2-saddles (index 2), and the maxima (index 3). A function \( f \) is Morse if all critical points are nondegenerate with distinct values.

Morse-Smale complex. An integral line is a maximal path on \( \mathbb{M} \) whose tangent vectors agree with the gradient of \( f \). The stable manifold of a critical point \( x \) is the union of \( x \) and all integral lines that end at \( x \). The unstable manifold of \( x \) is defined symmetrically as the union of a critical point \( x \) and all integral lines that start at \( x \). One can superimpose the stable and unstable manifolds of all critical points to create the Morse-Smale complex (or MS complex) of \( f \) [10, 4], see figure 1(a)–(d). The nodes of this complex are the critical points of \( f \), its arcs are integral lines starting or ending at saddles, and its regions are the nonempty intersections of stable and unstable 2-manifolds. More details on the definition of the MS-complex on 2-manifold triangle meshes and algorithms to compute it are given by Bremer et al. [4].

We briefly describe recent work on 3D Morse-Smale complex computation; See [13] for details. We use a single sweep to construct both the combinatorial structure of the complex and the geometric structure of the ascending manifolds. A second sweep uses this segmentation to guide construction of the geometry of the descending manifolds in a manner consistent with the topology identified in the first step. We view the construction of the ascending and descending manifolds as an iteration through dimensions, where the interior of each \( n \)-manifold is computed via region growing, and its \( n - 1 \)-manifold boundaries are recursively computed. Figure 2 shows how we iterate through dimensions while constructing the ascending manifolds.
Figure 2. (a) The algorithm at each iteration identifies interior and boundary vertices of an ascending $n$-manifold. (b) Using minima as seed points, we grow ascending 3-manifolds, finding vertices that are interior (blue) to a single 3-manifold. The grey region depicts the boundary between ascending 3-manifolds. (c) We next compute ascending 2-manifolds by first identifying interior points. An ascending 2-manifold (green) separates exactly two ascending 3-manifolds. (d) Ascending 1-manifolds (yellow) separate a unique set of ascending 2-manifolds. (e) Finally, the maxima (red) are found as the separating sets between ascending 1-manifolds.

Simplification. It is often useful to simplify an MS-complex to remove noise and to perform multi-scale function analysis. Following [4] we perform cancellations of arc-connected maximum-saddle and minimum-saddle critical point pairs to simplify an MS-complex. Cancellations are ranked by their persistence — the absolute difference in function value between the canceled critical point pair. Figure 1(e) and (f) shows an example of a topological simplification and a corresponding approximation of $f$.

Computation. In practice, one usually deals with piecewise linear (PL)-functions given at the vertices of a triangulation. See [4, 5] for a detailed discussion on how to translate concepts from the generic smooth functions discussed above to PL-functions.

Starting from saddles, we compute the arcs of the MS-complex as steepest non-crossing monotone lines. Because we avoid mesh refinement to handle degeneracies, and instead directly handle merged lines and multi-saddles we can use efficient static data structures to store the triangulation and can compute MS-complexes of large data sets common in simulation.

3. Analysis of turbulent mixing in hydrodynamic instabilities

Understanding the turbulent mixing of fluids is one of the fundamental research problems in the area of fluid dynamics. Turbulent mixing occurs in a broad spectrum of phenomena ranging from boiling water to astrophysics and nuclear fusion. Rayleigh-Taylor instability occurs when two fluids of different density are accelerated opposite the mean density gradient. That is, a heavier fluid is accelerated against a lighter fluid by a force such as gravity. Figure 3 shows the mixing layers and the progression of the mixing process. The heavy fluid accelerates downward, forming “spikes,” while the light fluid moves upward forming “bubbles.” The bubbles and spikes are thought to be one way to characterize the large-scale behavior of the mixing process. Scientists analyzing these simulations are particularly interested in the number of bubbles (and spikes) and their respective evolution. Large-scale models have been proposed based on bubble dynamics in which bubble growth, movement, and interaction are modeled [2]. Our analysis is performed on the envelope surfaces describing the boundary between undisturbed and “mixed” fluids.
**Figure 3.** An overview of the $1152^3$ simulation (periodic in $x$ and $y$) of the Rayleigh-Taylor instability at start ($t=0$), early ($t=200$), middle ($t=400$), and late ($t=700$) time. The light fluid has a density of 1.0, the heavy fluid has a density of 3.0. Two envelope surfaces (at densities 1.02 and 2.98) capture the mixing region. The boundaries of the box show the density field in pseudocolor. heavy fluid is red and the light fluid is blue. Totally mixed fluid and the midplanes to study mixing trends.

**Segmentation of bubbles.** One of the challenges in analyzing the mixing behavior is that there exists no prevalent mathematical definition for what constitutes a bubble/spike. In general, a bubble can be understood as a three dimensional feature composed of lighter density fluid moving upwards (in the $Z$-direction) into a heavier density fluid. We use the topological concepts introduced in Section 2 to define bubbles, spikes and other features of interest. Consider the images of the segmented mixing envelope surface at different times shown to the left, bottom, and right of the plot in figure 4. During early time steps (figure 4 upper left and middle left) it is natural to consider the mixing envelope as a time-varying functional surface defined over the $XY$-plane and associate local maxima with bubbles. This analogy fails at later time steps because the surface becomes non-functional. However, we can generalize this approach by treating the envelope surface as the domain of a function whose value at a point $x$ is the $Z$-coordinate of $x$. It is natural to connect the maxima of this function to bubbles and compute the stable manifold of each maximum as a segmentation of the surface into bubbles. As can be seen in figure 4, this segmentation corresponds very well to the human notion of a bubble. Symmetrically, we use the unstable manifolds of minima to define spikes. Potentially, other functions could be defined on the envelope surface that would result in a robust segmentation as well. For example, the $Z$-velocity at all of the points on the envelope surface could be incorporated to capture the fact that bubbles should be moving upwards into the heavy fluid. Given the complexity of the problem, it has been determined to use the simplest most intuitive segmentation.

In general, topological-based segmentations are often linked to important features: maximal and minimal $Z$-velocities on the midplanes correspond to cores of rising and falling sections of fluids; density extrema correspond to pockets of unmixed fluids. Topological methods are flexible and enable analysis of a variety of phenomena using the same methodology. Furthermore, the MS-complex can be computed combinatorially [10, 4] which translates into provably correct and stable algorithms which are crucial when dealing with large and complex data.

**Multiscale analysis and persistence selection.** The MS-complex, just as any other segmentation, captures noise as well as features. A simplification scheme can be used to remove noise and construct a series of approximations at decreasing resolution. Unlike many other techniques, topological segmentations allow a simplification scheme that is optimal in the $L_\infty$-norm. One can formulate the problem of coarsening a segmentation in the following manner: given a function $f$ and a segmentation $S$ of the domain of $f$, what is the minimal change on $f$ such that $S$ is coarsened? If the segmentation one considers is the MS-complex of $f$, then it can be shown [4] that canceling a critical point pair with persistence $p$ in $f$ requires an approximation $\hat{f}$ with $||f - \hat{f}||_\infty \geq p/2$. Therefore, canceling critical points in order of increasing persistence corresponds to an $L_\infty$-optimal simplification.
For each MS-complex, we compute a sequence of cancellations that optimally simplify the complex down to its minimal configuration. We can thus define a family of segmentations of the envelope surface ranging from persistence $p = 0.0$, where both signal and noise features are segmented, to persistence $p = 1.0$ (full function range), where the entire surface is collapsed into a single component. We can then create statistics showing the number of bubbles over time using a range of persistence thresholds. As shown in Section 3.1, the mixing behavior can differ significantly across scales. Using the simplification sequences, we can capture the behavior on all scales without recomputing the MS-complex. Domain scientists interact with a visualization of the segmented surface and select an appropriate persistence value based on their physical intuition of a correct segmentation of bubbles.

### 3.1. Results: Bubble counts and quantification of mixing phases

The input data consists of 758 time steps, each time step containing about 5.8 GB of density data defined on a $1152^3$ grid. Our analysis was performed on 68 dual-processor nodes of a cluster running Linux. For each time step, we extract an isosurface at density value 2.98 and compute the Morse-Smale complex with function $f$ set to the $z$-coordinate of each point. Isosurface extraction takes about 15 seconds at early time steps and about 30 seconds at the late time steps where the surface is more convoluted. Morse-Smale complex computation takes about 2 seconds at early time-steps and about 25 seconds at late time-steps. We compute a digest of all required data from the Morse-Smale complex that allows us to calculate bubble counts for any choice of persistence efficiently.

Figure 4 describes the count for three choices of persistence values. The inset figures show the segmentation of the isosurface into bubbles. It was not clear what value of persistence was a good choice or if the persistence value would have to change over time. Because our analysis pipeline is flexible, and because we can quickly compute the Morse-Smale complex and bubble count at
any choice of persistence from the digest, we can provide analysts with a variety of visualizations and statistics that they can study and choose an appropriate persistence that produces results that make sense with respect to the application domain. Based on feedback from simulation scientists, we chose a persistence value of 2.39% for the late stages of the mixing. Notice, though, that each of the bubble-count curves show three linear (power-law) regimes. Although the slopes of the three curves are different, the times at which their slopes change are the same, showing that a temporal analysis based on variation of derivatives is scale independent (robust to the choice of input parameters).

Figure 5 shows the results of a more detailed analysis of the bubble count made by the application scientists. Here we have a linear variation of persistence from an early value of nearly zero to the late value of 2.39%. The plot reveals four phases of the mixing process, and provides for the first time quantification of the mixing rates and transition times between the different stages.

3.2. Feature-based comparison of large eddy simulations and direct numerical simulations

Our analysis techniques provide a method to compare two different data sets of the same phenomenon simulated using different classes of simulations. Large eddy simulations (LES) require smaller resolution meshes and less computational resources than direct numerical simulations (DNS). Grid resolutions for LES capture large-scale effects but must model the sub-grid scale effects explicitly. On the other hand, DNS grid resolutions must be fine enough to capture all scales.

Validating the results of these two types of simulations applied to the same phenomenon is of great importance to scientists. We have conducted a feature-based comparison that is shown in figure 6. The LES grid dimension was $1152^3$ and the DNS grid dimension was $3072^3$. We perform the analysis as described previously and compute bubble counts for both types of simulations. After normalizing the counts, we can plot them together and compare them. Figure 6 shows that the simulations are in close agreement and produce similar trends in the mixing layer.
4. Core structure of a porous solid

Porous materials, such as the aerogel used in the Stardust mission [14], are employed in many scientific and industrial applications. Atomistic simulations [6] have been developed to understand their structural properties under a variety of conditions including under impact by a projectile. Detecting the core structure of the porous solid and tracking this structure under impact aids analysis of the simulation and provides important insight into the behaviour of the solid.

We compute the core structure of a porous solid using two methods. First, we compute the Morse-Smale complex using the incremental algorithm presented in [13], with the finger removing and arc smoothing modifications. Then, we analyze the critical point pairs, and filter the arcs to extract the core structure. The full complex shown in Figure 7, is used to plot the distribution of 2-saddle-maximum pairs, shown in figure 8. The features we are interested in are the arcs that connect a low 2-saddle to a high maximum; these critical point pairs are in the top left of the distribution. Flat regions in the integral along each axis reveal that the stable threshold for maxima is 1.5, and the stable threshold for 2-saddles is −0.8. By cancelling all arcs that do not entirely cross the range [−0.8, 1.5], we remove the artifacts and noise. The core structure is extracted as the arcs connecting 2-saddles to maxima that remain after simplification and are entirely contained in the isosurface for isovalue zero. For this particular application, we are interested in the connectivity of the porous solid, therefore we omit arcs that are connected to the structure at only one endpoint from the final core structure.

The results were generated by using an off-the-shelf personal computer, a 3.4 GHz Pentium 4, with 2 GB of memory. The porous solid was represented as real-valued samples on a 230 × 230 × 375 regular grid. The total time required for computation of the initial MS complex was 6 hours 32 minutes and 45 seconds. Additional processing to attain the graph structure took 32 seconds. The total time required to create the clean distance field using the propagation

Figure 6. Feature-based comparison and validation of large eddy simulation (LES) and direct numerical simulation (DNS) of Rayleigh-Taylor instabilities.

Figure 7. The initial computation of the MS complex for the full dataset (left) is simplified revealing the graph structure (right) of the porous solid.
method starting with the same input dataset was 91 seconds. After the initial computation, exploration and further simplification can be done interactively. For large datasets, computation of the full MS complex is not possible; however, we can still perform analysis on a subset of the data to attain the distribution of critical point pairs, assuming features are distributed nearly uniformly in the dataset. The clean distance field method is an efficient implementation that can be applied to large data and benefits from such an analysis.

5. Time-dependent impact data
We have used our methods to explore a simulated dataset of a particle impact on the porous material at several time steps. By computing the clean distance field, we can obtain the density of the porous solid as the ratio of the number of sample points interior to the zero isosurface to the number of sample points outside. The clean distance field ensures that all sample points identified in this way contribute to the structure of the porous material. The structural analysis and comparison between time steps allow us to obtain an important result: there is significant densification of the foam below the crater wall. Such analysis provides a simple, quantitative answer to the amount of densification, as shown in figure 10. In this figure we show the density profile at different times as a function of depth for slices of (a) the whole sample and (b) the center of the sample, including the crater. It can be seen in (b) that the density increases by a factor of two below the crater.

We also have computed the core structure using the MS complex for each time step. Using the distance measures described in [12], we can compare qualitatively how the structure of the material changes between time steps. Figure 9 is a visual depiction of the displacement of filament segments at different times after impact. Note the large displacements near the crater, and the nearly zero displacement well below the crater.

In both figures we can see that density barely changes in the bottom one-third of our sample. This is partly due to the fact that the foam is extremely efficient at absorbing the impact shock wave. Key statistics of the core structure for each time steps are summarized in Table 1.

The ratio of cycle counts before and after the impact supports this observation, as approximately two-thirds of the cycles are destroyed. The ratio of the total length of the filaments before and after the particle impact implies that volume of material displaced by the crater is approximately one-half the volume of the rest of the material. Since this ratio is fairly close to the ratio of the cycle counts, we can say that the majority of the filaments that were

| Metric    | t=500 | t=12750 | t=25500 | t=51000 |
|-----------|-------|---------|---------|---------|
| # Cycles  | 762   | 340     | 372     | 256     |
| Total Length | 34756 | 24316   | 23798   | 18912   |

Table 1. Statistics for each time step.
Figure 9. Top: Volume rendering of the impact of the ball entering the porous solid from the left at time step 500 (a), time step 12750 (b), time step 25500 (c), and time step 51000 (d). Bottom: We compare the core structures of consecutive time steps. The yellow dots represent the core structure of the initial time step, and the teal dots represent the core structure of the next time. The closest arcs between in the core structures at the different time steps are connected via blue and red line segments. The length of these segments corresponds to the displacement of the arc.

broken happened to be in the interior of the crater. The sum of the Hausdorff distances between the time steps is 98.6, giving the maximum distance that any element of the material travelled during the impact. This number is surprisingly high, corresponding to the entire depth of the crater; it indicates that the material of the filaments first hit by the particle was displaced along the trajectory of the particle. The average distance between closest pairs in the graphs of the consecutive time steps was less than 5.0, indicating that the displacement did not propagate into the material, outside the direct path of the particle.

Densification of the foam will vary as a function of impact velocity, and a quantitative characterization of such a function might help to narrow down such velocities when they are not known. In addition, the fact that the foam is getting denser would change, for large particle fluxes, the foam's mechanical and thermal properties.

6. Conclusion
Morse theory provides a sound mathematical framework to study functions on manifolds. This framework can be extended to analyze piecewise linear functions created from real-world scientific data. Using this foundation we develop robust and flexible software that can be fine-tuned to the analysis of a variety of datasets. We present the analysis of two datasets: Rayleigh-Taylor instability simulations and porous solid atomistic simulations.

As current trends continue, computing power will increase rapidly and scientists will develop more sophisticated and larger scale simulations. We believe techniques, such as those described in this paper, will become crucial to address the challenge of analyzing these scientific datasets, and accelerate scientific discoveries.

References
[1] P. S. Alexandrov. Combinatorial Topology. Dover, Mineola, NY, USA, 1998.
[2] U. Alon and D. Shvarts. Two-phase flow model for rayleigh-taylor and richtmeyer-meshkov mixing. In Proc. Fifth International Workshop on Compressible Turbulent Mixing. World Scientific, 1996.
Figure 10. Computation, at each time step, of the filament density of the material (ratio of samples inside the interface surface over those outside the interface). We compute this density profile for the entire data (left) and a cylinder through the impact crater (right).

[3] P.-T. Bremer, H. Edelsbrunner, B. Hamann, and V. Pascucci. A multi-resolution data structure for two-dimensional Morse-Smale functions. In G. Turk, J. J. van Wijk, and R. Moorhead, editors, *Proc. IEEE Visualization ’93*, pages 139–146, Los Alamitos California, 2003. IEEE, IEEE Computer Society Press.

[4] P.-T. Bremer, H. Edelsbrunner, B. Hamann, and V. Pascucci. A topological hierarchy for functions on triangulated surfaces. *IEEE Trans. on Visualization and Computer Graphics*, 10(4):385–396, 2004.

[5] P.-T. Bremer and V. Pascucci. A practical approach to two-dimensional scalar topology. page to appear. 2006.

[6] E. M. Bringa, J. U. Cazamias, P. Erhart, J. Stolken, N. Tanashev, B. D. Wirth, R. E. Rudd, and M. J. Caturla. Atomistic shock hugoniots simulation of single-crystal copper. *Journal of Applied Physics*, 96(7):3793–3799, 2004.

[7] K. Cole-McLaughlin, H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci. Loops in reeb graphs of 2-manifolds. In *Proceedings of the 19th Annual Symposium on Computational Geometry*, pages 344–350. ACM Press, 2003.

[8] A.W. Cook, W. Cabot, and P.L. Miller. The mixing transition in rayleigh-taylor instability. *J. Fluid Mech.*, to appear, 2004.

[9] A.W. Cook and P.E. Dimotakis. Transition stages of rayleigh-taylor instability between miscible fluids. *J. Fluid Mech.*, 443:66–99, 2001.

[10] H. Edelsbrunner, J. Harer, and A. Zomorodian. Hierarchical Morse-Smale complexes for piecewise linear 2-manifolds. *Discrete Comput. Geom.*, 30:87–107, 2003.

[11] H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. In *FOCS ’00: Proceedings of the 41st Annual Symposium on Foundations of Computer Science*, page 454, Washington, DC, USA, 2000. IEEE Computer Society.

[12] A. Gyulassy, M. Duchaineau, V. Natarajan, V. Pascucci, E. Bringa, A. Higginbotham, and B. Hamann. Topologically clean distance fields. *IEEE Transactions on Visualization and Computer Graphics* (to appear), pages xxx—xxx, 2007.

[13] A. Gyulassy, V. Natarajan, V. Pascucci, and B. Hamann. Efficient computation of morse-smale complexes for three-dimensional scalar functions. *IEEE Transactions on Visualization and Computer Graphics*, pages xxx—xxx, 2007.

[14] Friedrich Horz, Ron Bastien, Janet Borg, John P. Bradley, John C. Bridges, Donald E. Brownlee, Mark J. Burchell, Miaolang Chi, Mark J. Cintala, Zu Rong Dai, Zahia Djourdi, Gerardo Dominguez, Thanasis E. Economou, Sam A. J. Fairey, Christine Floss, Ian A. Franchi, Giles A. Graham, Simon F. Green, Philipp Heck, Peter Hoppe, Joachim Huth, Hope Ishii, Anton T. Kearsley, Jochen Kissel, Yan Leitner, Hugues Leroux, Kuljeet Marhas, Keiko Messenger, Craig S. Schwandt, Thomas H. See, Christopher Snead, I. Stadermann, Frank J., Thomas Stephan, Rhonda Stroud, Nick Teslich, Josep M. Trigo-Rodriguez, A. J. Tuzzolino, David Troadec, Peter Tsou, Jack Warren, Andrew Westphal, Penelope Wozniakiewicz, Ian Wright, and Ernst Zimmer. Impact Features on Stardust: Implications for Comet 81P/Wild 2 Dust. *Science*, 314(5806):1716–1719, 2006.

[15] D. Laney, P. T. Bremer, A. Mascarenhas, P. Miller, and V. Pascucci. Understanding the structure of the turbulent mixing layer in hydrodynamic instabilities. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1053–1060, 2006.

[16] Y. Matsumoto. *An Introduction to Morse Theory. Translated from Japanese by K. Hudson and M. Saito*. American Mathematical Society, 2002.

[17] P. Miller, P.-T. Bremer, W. Cabot, A. Cook, D. Laney, A. Mascarenhas, and V. Pascucci. Application of morse theory to analysis of rayleigh-taylor topology. In *Proceedings of the 10th International Workshop on the Physics of Compressible Turbulent Mixing*, 2006.

[18] J. W. Milnor. *Morse Theory*. Princeton Univ. Press, New Jersey, USA, 1963.

[19] J. R. Munkres. *Elements of Algebraic Topology*. Addison-Wesley, Redwood City, CA, USA, 1984.

[20] V. Natarajan, Y. Wang, P.-T. Bremer, V. Pascucci, and B. Hamann. Segmenting molecular surfaces. *Comput. Aided Geom. Des.*, 23(6):495–509, 2006.