

Final Report for *University of Utah Group*
Contribution to DE-FG02-09ER25949/DE-SC0002505:
Topology for Statistical Modeling of Petascale Data
(An ASCR-funded collaboration between Sandia
National Labs, Texas A&M U, and U Utah)

Valerio Pascucci Joshua Levine Attila Gyulassy P.-T. Bremer

October , 2013

1 Introduction

The primary goal of our local team is to develop new tools and software implementations for characterizing the topology of the models arising from the analysis of petascale data. Much of the theory has been detailed in the original project proposal, and the activities described in the report represent the work during the lifetime of the project.

2 Introduction

Many commonly used algorithms for mathematical analysis do not scale well enough to accommodate the size or complexity of petascale data produced by computational simulations. The primary goal of this project is to develop new mathematical tools that address both the petascale size and uncertain nature of current data.

At a high level, the approach of the entire team involving all three institutions is based on the complementary techniques of combinatorial topology and statistical modelling. In particular, we use combinatorial topology to filter out spurious data that would otherwise skew statistical modelling techniques, and we employ advanced algorithms from algebraic statistics to efficiently find globally optimal fits to statistical models. The overall technical contributions can be divided loosely into three categories: (1) advances in the field of combinatorial topology, (2) advances in statistical modelling, and (3) new integrated topological and statistical methods.

Roughly speaking, the division of labor between our 3 groups (Sandia Labs in Livermore, Texas A&M in College Station, and U Utah in Salt Lake City) is as follows: the Sandia group focuses on statistical methods and their formulation in algebraic terms, and finds the application problems (and data sets) most relevant to this project, the Texas A&M Group develops new algebraic geometry algorithms, in particular with fewnomial theory, and the Utah group develops new algorithms in

computational topology via Discrete Morse Theory. However, we hasten to point out that our three groups stay in tight contact via videoconference every 2 weeks, so there is much synergy of ideas between the groups.

The following of this document is focused on the contributions that had grater direct involvement from the team at the University of Utah in Salt Lake City.

3 Contributions in Combinatorial Topology

3.1 Visualizing Morse Smale Complexes

Recent advances in practical algorithms for computing Morse-Smale (MS) complexes have made possible multi-resolution analysis of volumetric scalar valued data. Although these approaches are gaining popularity in analysis of scientific data, visualization techniques have not yet explored the full potential of this technology. In [GKK⁺, GBP, WBG], we present novel visualizations using features extracted from MS complexes. We characterize possible visualizations enabled by the robust computation of all dimensional manifolds of the MS complex, and present several examples of these. Furthermore, we developed a framework for selecting features, assigning attributes, and building complex and compelling visualizations. In Figure 1 we show examples of topology-based techniques used to extract features that are hard to detect with more conventional methods not making use of topologic information. For practical use in petascale data we have also worked on improving efficiency and scalable parallel computation [GPPR12].

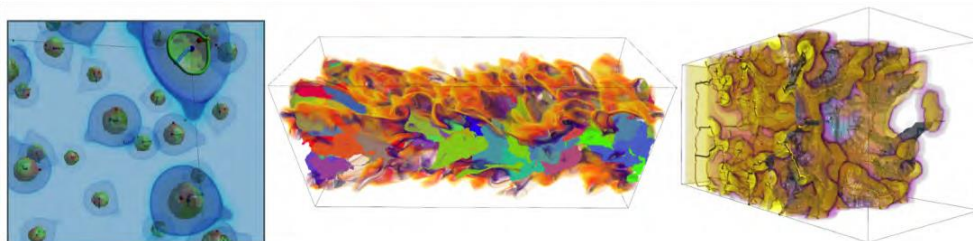


Figure 1: Left: Combinatorial computation of topological invariants results in robust identification of features, even in degenerate cases, such as topological pouches. Center: Using the machinery of topological persistence and simplification, we can visualize the 3-manifolds of the MS complex forming flow basins in a manner oblivious to noise. Right: Derived structures, for example, separating surfaces, can be used to represent non-physical phenomena, such as the “outer surface” of a sponge-like material.

3.2 Unstructured Morse-Smale Complexes

Given the success of topological analysis tools in many domains, there is a need to compute MS complexes for data that is defined on fully unstructured domains. Many of the first software milestones were designed for gridded domains. We have developed software tools which implement Gyulassy’s discrete gradient construction [GBHP08] on regular cell complexes. This resulted in both a video tutorial as well as multimedia submission [GLP11]. The algorithm is illustrated in Figure 2. This generic software allows the construction of MS complexes for multiple

modalities of data. As a result, it will enable topological analysis in for many data sources, allowing for the same demonstrated benefits we have already seen in gridded and PL domains to be leveraged for multiple types of data.

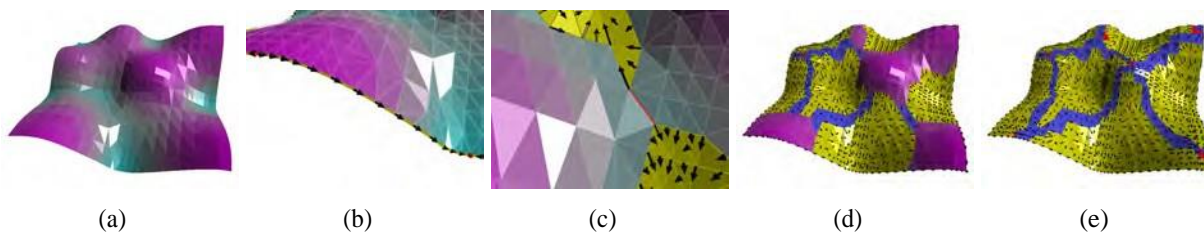


Figure 2: Assigning gradient arrows on a terrain (a). Scalar values (height) are encoded from cyan (low values) to magenta (high values). (b) Boundary cells are paired first. (c) Pairing interior cells finds a saddle (red edge). (d) As pairing continues, a maxima is identified (red triangle). (e) Gradient construction is complete. Ascending 1-manifolds shown as blue cells.

3.3 Edge Maps

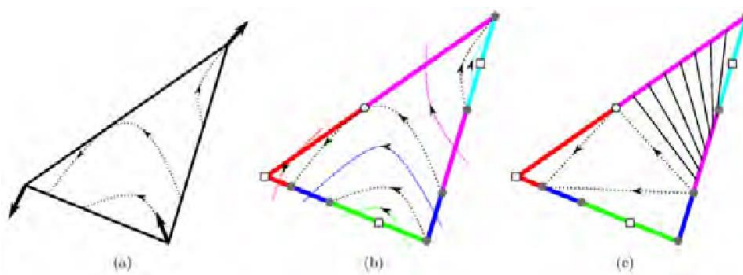


Figure 3: The structure of an edge map of a regular triangle. (a) The original triangle is represented as three vectors, which implies a flow throughout the interior. (b) Our representation subdivides the boundary into a set of intervals, which are broken at internal transition points (white circles), external transition points (white squares), and image points (grey circles). (c) Pairs of intervals are grouped into maps which represent sources and destinations of flow through the triangle intervals.

Robust analysis of vector fields has been established as an important tool for deriving insights from the complex systems these fields model. Traditional analysis and visualization techniques rely primarily on computing streamlines through numerical integration. The inherent numerical errors of such approaches are usually ignored, leading to inconsistencies that cause unreliable visualizations and can ultimately prevent in-depth analysis. In [JBB⁺10, BJB⁺11a, JBB⁺11a, BJB⁺11b] we propose an alternate representation for vector fields on surfaces that explicitly represents the flow behavior of the field through each triangle. This representation, called *edge maps*, complements the traditional approach of storing sample vectors on the vertices of the triangulation. Figure 3 shows the structure of an edge map of a regular triangle.

One piece of this work [JBB⁺10, JBB⁺11a] focuses on the mathematical properties of edge maps. Edge maps allow for a multi-resolution approximation of flow by merging adjacent stream-

lines into an interval based mapping. Consistency is enforced at any resolution if the merged sets maintain an order-preserving property. At the coarsest resolution, we define a notion of equivalency between edge maps, and show that there exist 23 equivalence classes (Figure 4) describing all possible behaviors of piecewise linear flow within a triangle.

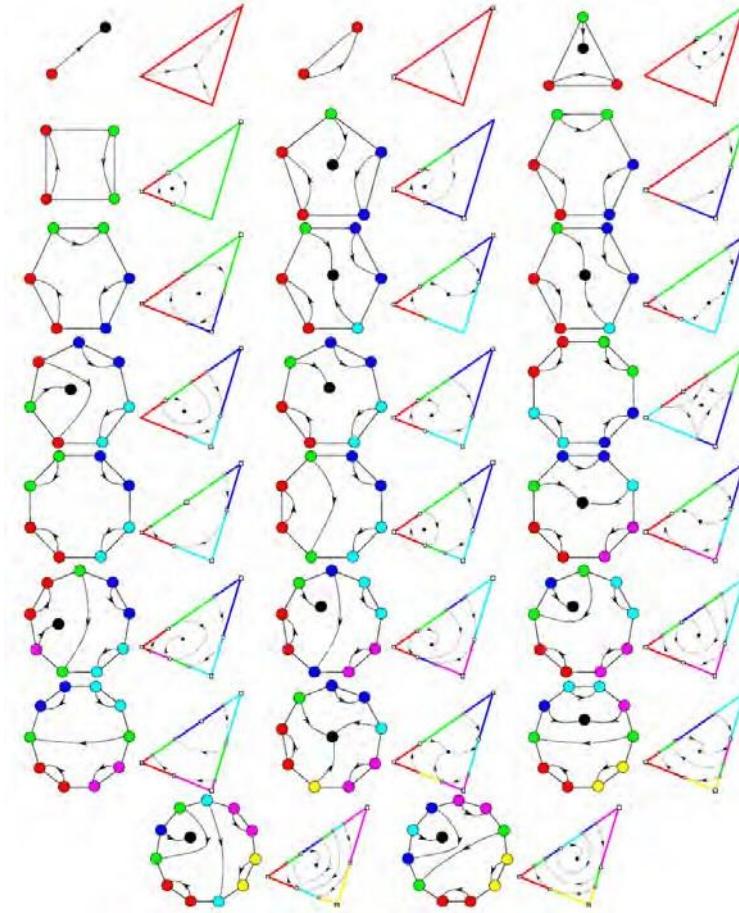


Figure 4: The 23 equivalent classes of mixed graphs that represent piecewise linear flow through a triangle, along with one possible rendition of the edge map. The ordering is of increasing number of links in the map.

A second branch of this work [BJB⁺11a, BJB⁺11b] focuses on encoding the spatial and temporal errors which we use to produce more informative visualizations. This work describes the construction of edge maps, the error quantification, and a refinement procedure to adhere to a user defined error bound (Figure 5). Independent of this error all streamlines computed using edge maps are guaranteed to be consistent, enabling the stable extraction of features such as the topological skeleton. We introduce new visualizations using the additional information provided by the edge maps to indicate the uncertainty involved in computing streamlines and topological structures (Figure 6). [BJB⁺11a] received a best paper award at the 4th IEEE Pacific Visualization Symposium in Hong Kong, China 2011.

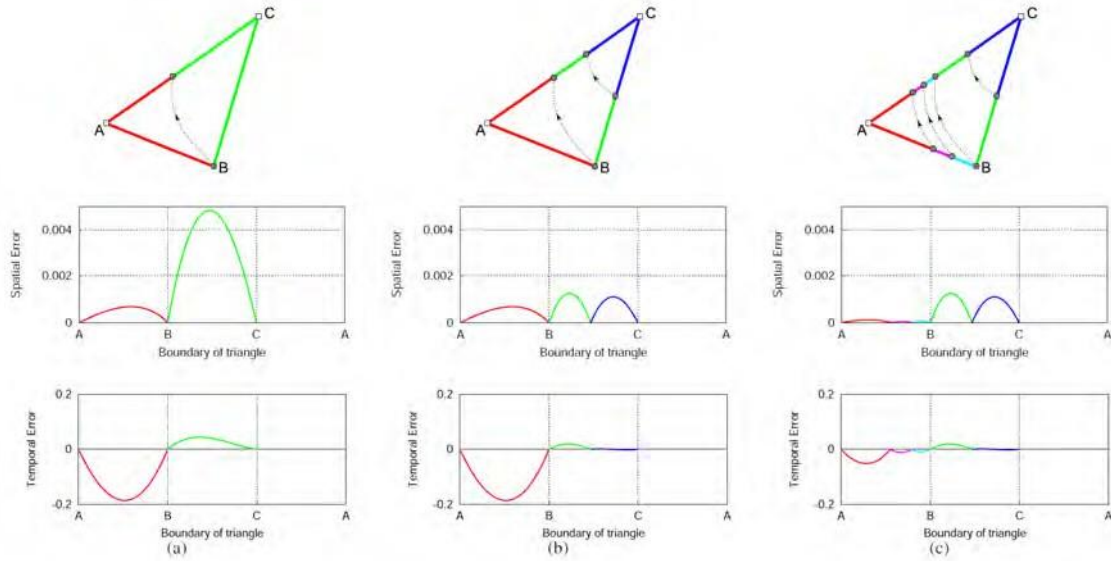


Figure 5: Reducing the mapping error (middle row: spatial error, bottom row: temporal error) by refinement of edge maps (top row). Level of refinement increases from left to right. The length of the edge AC is 0.0354, and the average time taken by a particle to travel across the triangle is 1.7. (a) No refinement. (b) Spatial refinement with an error bound of 0.003 splits the green link into two, creating two new links (green and blue) with smaller spatial and temporal errors. (c) Temporal refinement with an error bound of 0.06 splits the red link twice, creating three new links (red, cyan and magenta) with even smaller spatial and temporal errors.

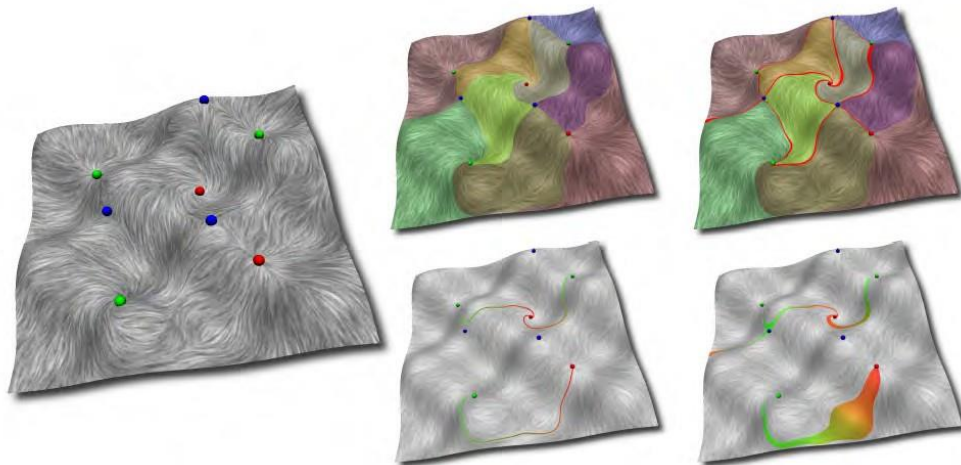


Figure 6: Edge maps enable new views of vector field stability, illustrated with a vector field on this wavy surface. Top row (middle right): A visualization of some colored regions where flow shares the same source (green spheres) and sink (red spheres) is augmented to show how these regions overlap when error is introduced. Bottom row (middle right): Streamwaves (colored green to red as they grow) show the advection of a single particle. In the presence of error, waves can widen and narrow, and bifurcate or merge.

3.4 Quantized 2D Vector Fields

Visualization and analysis of vector fields often hinges on the robust identification of structures such as critical points, separatrices, or closed orbits. Traditional techniques for computing these features fall broadly into two categories: (1) those that use numerical integration and (2) those that rely on purely combinatorial structures. However, the first set of tools often generate inconsistent results due to compounded approximation errors, while the second set often severely reduces the accuracy of the results. Instead, we propose [JBB⁺11b] a new discrete representation of vector fields that approximates the flow up to an arbitrary, user-defined error. By quantizing streamlines along edges of a triangulation we create a graph-based representation of the flow with up to 2^{32} nodes per edge. The graph is implicitly represented by a rasterization scheme and replaces streamline integration with a directed graph traversal. As a result, inherently unstable structures such as separatrices and cycles can be computed exactly and correctly up to the given approximation error. By varying the amount of discretization, we can provide a multi-resolution representation of vector fields that allows for a balance between storage space, computational effort, and fidelity to piecewise linear interpolation. Figure 7 shows our technique on ocean current data, illustrating a more complete topological representation that includes separatrices, stable manifolds, and cycles detected with our approach.

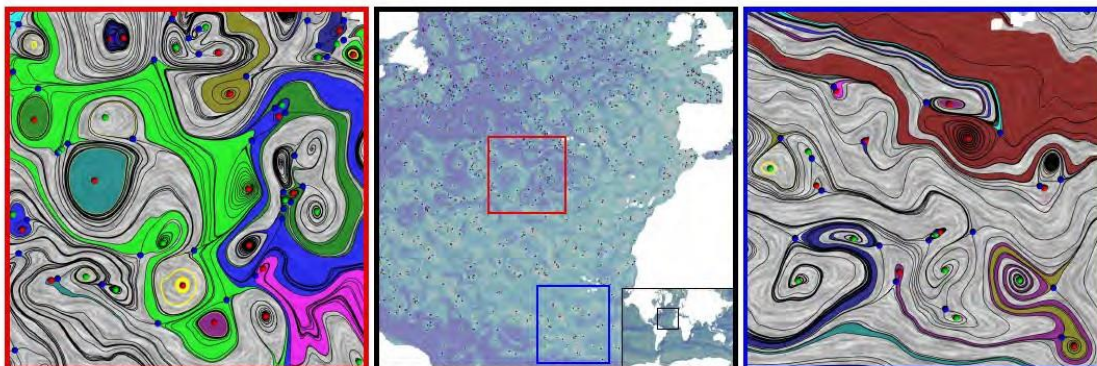


Figure 7: The oceanic currents of the North Atlantic ocean. In the center, we show a 600x600 vertex tile from a larger simulation of oceanic currents. Each image on the side is a zoomed in view visualizing the topology for the tile. Yellow lines are closed streamlines, the colored regions are stable manifolds grown from all the sinks (red balls), and the black lines are separatrices grown from all saddles (blue balls).

3.5 Fiedler Trees for Multiscale Surface Analysis

In [BNPS10] we introduce a new hierarchical decomposition method for multi-scale analysis of surface meshes. In contrast to other multi-resolution methods, our approach relies on spectral properties of the surface to build a binary hierarchical decomposition. Namely, we utilize the first nontrivial eigenfunction (the Fiedler vector) of the Laplace-Beltrami operator to recursively decompose the surface. For this reason we coin our surface decomposition the Fiedler tree. Using the Fiedler tree ensures a number of attractive properties, including: mesh-independent decomposition, well-formed and nearly equi-areal surface patches, and noise robustness. We show how the evenly distributed patches can be exploited for generating multi-resolution high quality

uniform meshes. Additionally, our decomposition permits a natural means for carrying out wavelet methods, resulting in an intuitive method for producing feature sensitive meshes at multiple scales. The method and its generalization to volumetric meshes will be a critical component in building the hierarchical models needed in analysis of petascale data. Figure 8 shows the intrinsic nature of the mesh decomposition.

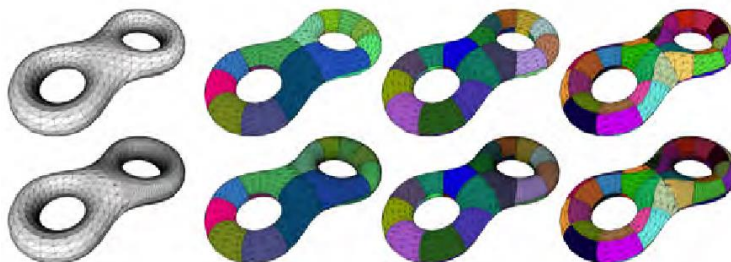


Figure 8: Two input meshes with same shape and different triangulation (left) yield the same sequence of decompositions (in color).

4 New Integrated Topological and Statistical Methods

4.1 Exploring High dimensional Spaces for Uncertainty Quantification

An important goal of scientific data analysis is to understand the behavior of a system or process based on a sample of the system. In many instances it is possible to observe both, input parameters and outputs, and characterize the system as a high-dimensional function. Such data sets arises, for instance, in understanding the uncertainty of large numerical simulations, energy landscapes in optimization problems, or the statistical analysis of image data relating to biological or medical parameters. In [GBPW10,LLBP12] we propose an approach that analyzes and visualizes such data sets. To do so it combines topological and statistical geometric techniques to provide interactive visualizations of discretely sampled high-dimensional scalar fields. The method relies on a segmentation of the parameter space using an approximate Morse-Smale complex on a cloud of point samples. For each crystal of the Morse-Smale complex, a regression of the system parameters with respect to the output yields a curve in the parameter space. The result is a simplified geometric representation of the Morse-Smale complex in the high dimensional input domain. Finally, the geometric representation is embedded in 2D, using dimensionality reduction, to provide a visualization platform. The geometric properties of the regression curves enable the visualization of additional information about each crystal such as local and global shape, width, length, and sampling densities. The method is demonstrated on several synthetic examples and several real scientific problems including: the analysis of manufacturing parameters an their effect on the strength of concrete, the parameters of climate simulations and their relationship to predicted global energy flux, the local concentrations of chemical species in a combustion simulation and their integrations with temperature, and the relationships between MRI brain images and measured clinical variables. Figure 9

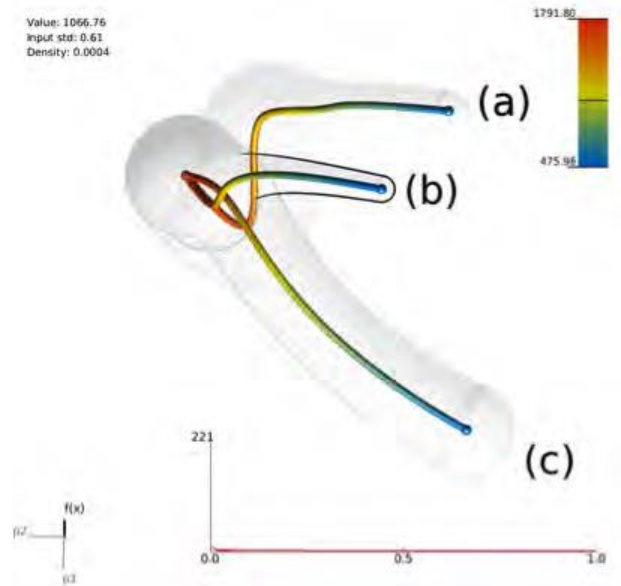


Figure 9: The three distinct minima (blue spheres) correspond to pure fuel, pure oxidizer and extinction/reignition. Graphs of chemical composition plotted against temperature for the crystals indicate that these three minima correspond to extinction (a), pure oxidizer (b) and pure fuel (c). The global maximum (red sphere) corresponds to efficient burning conditions.

shows our visualization technique applied to a high dimensional dataset of chemical composition in relation to heat released during a jet flame combustion simulation. This use of high dimensional analysis and visualization can also be used for visualization of interesting patterns in software [CWRP12].

4.2 Analysis of Large-Scale Scalar Data Using Hixels

One of the greatest challenges for today's visualization and analysis communities is the massive amounts of data generated from state-of-the-art simulations. Traditionally, the increase in spatial resolution has driven most of the data explosion, but more recently ensembles of simulations with multiple results per data point and stochastic simulations storing individual probability distributions are increasingly common. In [TLB⁺11] we introduce a new data representation for scalar data called hixels that store a histogram of values for each sample point of a domain. The histograms may be created by spatial down-sampling, binning ensemble values, or polling values from a given distribution. In this manner, hixels form a compact yet information-rich approximation of large-scale data. In essence, hixels trade off data size and complexity for scalar-value "uncertainty". Based on this new representation we propose new feature detection algorithms using a combination of topological and statistical methods. In particular, we show how to approximate topological structures from hixel data, extract structures from multi-modal distributions, and render uncertain isosurfaces. In all three cases we demonstrate how using hixels compares to traditional techniques and provide new capabilities to recover prominent features that would otherwise be either infeasible to compute or ambiguous to infer.

Fuzzy Isosurfacing When down-sampling larger datasets, hixels enable the preserving the

presence of an isosurface within the data. In particular, when hixels store the counts of all function values present within a block, we can use that to compute the likelihood of the presence of an isosurface within that block. Figure 10 demonstrates the results of this technique for a large combustion jet data set with half a billion grid points.

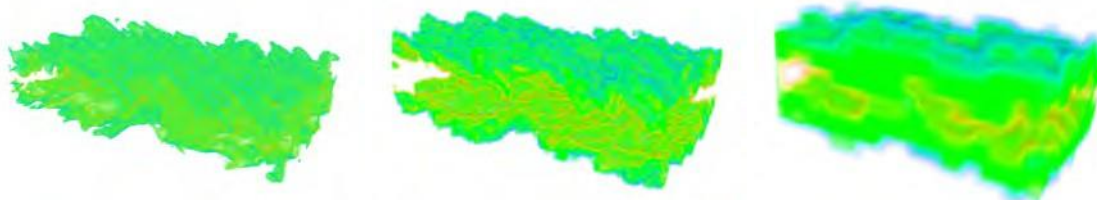


Figure 10: Volume rendering of the jet dataset down-sampled using hixels. We visualize the scalar field g that indicates the likelihood of isovalue $\kappa = 0.506$ lying at that position. From left-to-right, we show hixels that block 2^3 , 8^3 , and 32^3 data values. Opacity is a triangle function centered at $g = 0$ and color is a rainbow map, red for high values, green for middle, and blue for low.

Sampling Topology Hixels encode the potential values along with their distributions at sample locations, a fact that can be exploited in visualizing the uncertainty in topological segmentations of down-sampled data. We use a sampling of the hixels to generate individual instances of the coarser representation, compute the Morse complex on the instance, and aggregate multiple instances of the segmentation to visualize its variability. We generate an instance V_i of the down-sampled data by picking values at each sample from the co-located hixel. The value is picked at random from the distribution encoded by the hixel. By picking values independently from neighboring values, we can simulate any possible down-sampling of the data, assuming all are independent.

We perform convergence tests for a two-dimensional slice through a jet combustion simulation. In this experiment, we computed a hixel representation for the slice with blocks of size 8×8 and 16×16 . We visualize in Figure 11 each aggregate slice for the 8×8 block size, as number of iterations and topological persistence are varied. The convergence of these sequences indicates that the distribution represented by the hixels produces implies stable modes of segmentation.

Extracting Structures from Multi-Modal Distributions As HPC resources increase, ensembles of runs are being computed more frequently to explore the state space of phenomena of interest. The resulting ensemble data comprises a collection of simulation results, each of which represents a state in the system defined by different input parameters and/or models. While ensemble data sets are hailed as a useful mechanism for characterizing the uncertainty in a system, their large size and variability pose significant challenges for existing analysis and visualization techniques. We developed a novel statistical technique for recovering prominent topological features from ensemble data stored in hixel format. This computation is aided by the fact that ensemble data has a statistical dependence between runs that allows us to build a structure representing a predictive link between neighboring hixels. Our algorithm identifies subregions of space and scalar values that are consistent with positive association and we perform topological segmentation on only those regions.

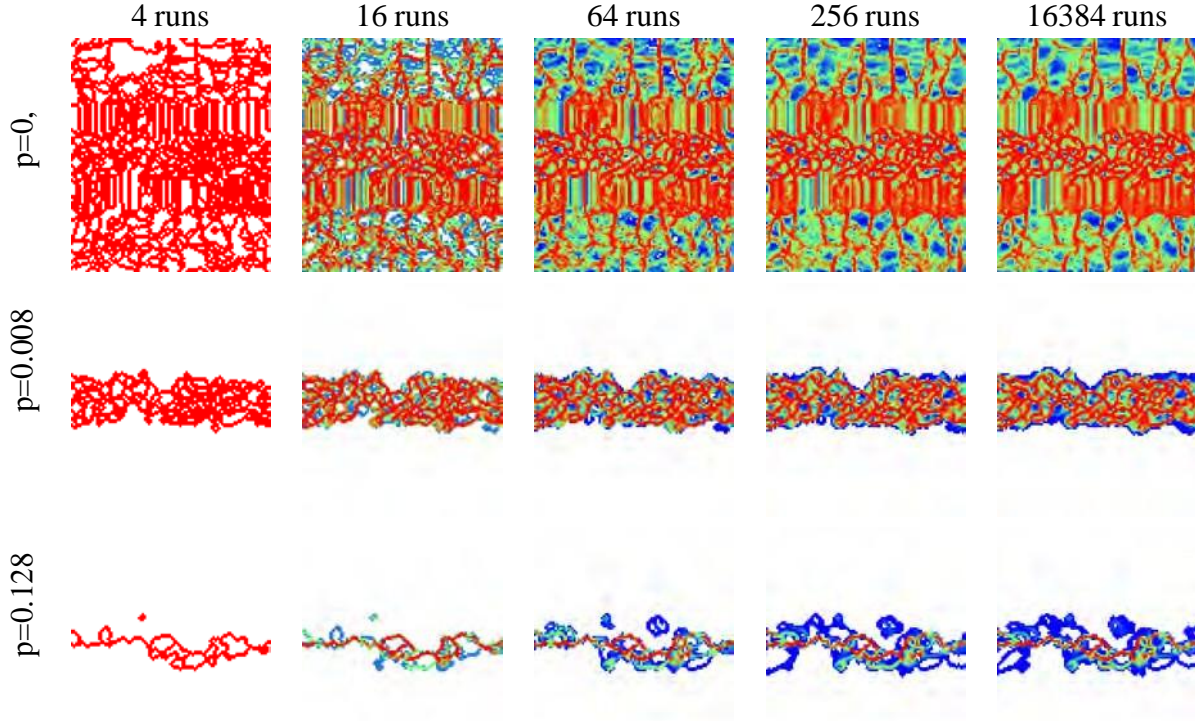


Figure 11: We sample the hixel data for an 8×8 blocking of combustion data, and compute the aggregate segmentation for a number of iterations, also varying the level of persistence simplification. Adjacent white pixels are identified in the interior of the same basin in every single run. The images converge as the number of iterations increases left to right.

We demonstrate results on a mixture of 2 stochastic processes shown in Figure 12. This data highlights the fact that individual hixels can be multi-modal and can behave as both a minimum and maximum. A naive analysis that computes the mean or median of the hixels, followed by standard topological segmentation would fail to incorporate the multi-modal nature of the data. Our method addresses this issue by performing topological analysis directly on sheets of the domain that have likely simultaneously observable sets of behavior. Our approach clearly extracts separate sheets belonging to the two processes and identifies their prominent features.

To compare against down-sampling a large-scale dataset, we also demonstrate results of this method on a hixelated data set generated from the log of the field of a lifted ethylene jet combustion data set with 1.3 billion grid points. The contingency tables between each pair of hixels are computed using observations between neighboring vertices along shared hixel faces. Figure 13 shows the number of buckets per hixel with block sizes of 16 (top-left), 32 (top-middle), and 64 (top-right). The color map ranges from blue at 1 bucket per hixel to red with 27 buckets per hixel and, as is to be expected, the number of buckets per hixel increases significantly as the block size increases. On the right the basins of maxima are shown for corresponding block sizes.



Figure 12: On the left, a volume rendering of a hixel data set generated by sampling images 3200 samples of a Poisson distribution and 9600 samples from a normal distribution. There are 512×512 hixels in this data set, each with 128 bins. The shortest axis in the images corresponds to histogram bins, thus a spatially higher location along that axis indicates a higher function value. Color and opacity are used to illustrate the density of samples. Thus the lower, right corner of shows a hixel with 2 distinct probable function values; the smaller function value is less probable than the larger. The center image shows basins of minima and the right image shows basins of maxima for this data set. By computing basins on sheets we are able to identify prominent features associated with each process in the mixture model.

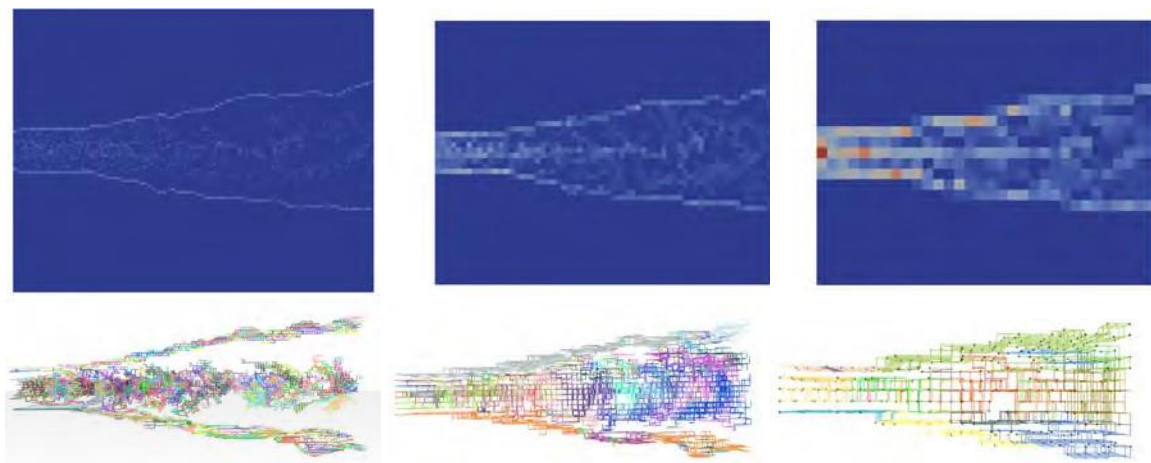


Figure 13: On the top the number of buckets per hixel is displayed for block sizes of 16 (left), 32 (middle) and 64 (right). Blue regions have 1 bucket per hixel while the maximum number of buckets per hixel is 27 and is shown in red. On the bottom the basins of maxima are shown for corresponding block sizes.

4.3 Feature-Based Statistical Analysis of Large Data

In [BKL⁺11a,WSPVJ11] we present a new framework for feature-based statistical analysis of large-scale scientific data and demonstrate its effectiveness by analyzing features from Direct Numerical Simulations (DNS). Combustion scientists use DNS to study fundamental turbulence-chemistry interactions such as extinction and auto-ignition in turbulent jet flames. Of particular interest is the

scalar dissipation rate, χ , which indicates the local rate of molecular mixing, which is enhanced by turbulent flow. Turbulent strains create thin pancake-like features of locally high dissipation

rate whose thickness provides a direct measure of the local mixing length-scale. Understanding the relationship between the thickness and the mean temperature within features is of principal interest to study the relationship between mechanical strains and chemical processes. This analysis is challenging due to the wide range of feature parameters that must be explored and the massive sizes of the simulation.

In our approach we pre-compute merge trees of the χ field which encode the set of features for all possible χ thresholds. Furthermore, we augment the merge trees with attributes, such as statistical moments of various scalar fields, e.g. χ , temperature, etc., as well as length scales computed via spectral analysis. The computation is performed in an efficient streaming manner in a pre-processing step and results in a collection of meta-data that is orders of magnitude smaller than the original simulation data. This meta-data is sufficient to support a fully flexible and interactive analysis of the features, allowing for arbitrary χ thresholds, providing per-feature statistics, and creating various global diagnostics such as Cumulative Density Functions (CDFs), histograms, or time-series. We combine the analysis with a rendering of the features in a linked-view browser that allows scientists to interactively explore, visualize, and analyze the equivalent of one terabyte of simulation data on a commodity desktop. While we have successfully deployed our framework to analyze statistical properties of turbulent combustion, its design and implementation are general and applicable to a wide range of scientific domains.

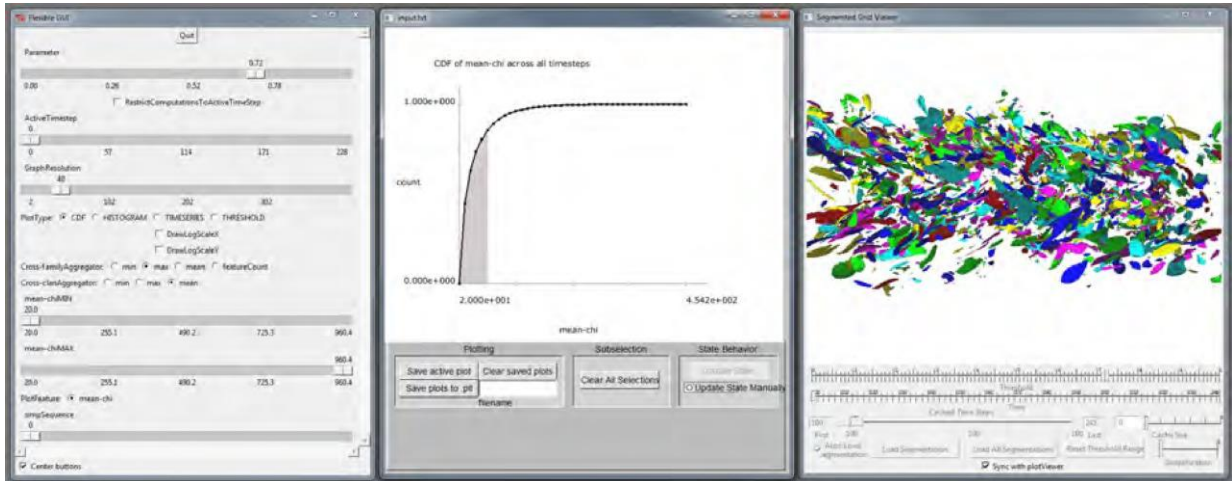


Figure 14: Our framework provides a natural and intuitive work-flow for the exploration of global trends in feature-based statistics. By efficiently encoding hierarchical meta-data in a pre-processing step, interactive data exploration of the equivalent of one terabyte of simulation data is performed on a commodity desktop.

5 Dissemination of results

Refereed Publications

- **11 refereed conference proceedings:** [BNPS10], [WGBP], [BJB⁺11a] (winner of best paper award), [GBP],[GLP11],[JBB⁺11a],[TLB⁺11], [CWRP12], [LLBP12], [WSPVJ11], [GPPR12].
- **7 refereed journal articles:** [GBPW10, BKL⁺11b, WSPVJ11, WPB⁺11, BJB⁺11b, GKK⁺,

JBB⁺11b].

- **1 multimedia submission:** [GLP11]

Presentations

- **24 invited/plenary presentations:**

- Pascucci: Keynote at Jacobi 2013: "High-Performance Computing: Mathematical Models and Algorithms, Kaliningrad, Russia, October 4, 2013.
- Pascucci: Eidgenössische Technische Hochschule Zrich (ETH), Zurich, CH, May 3, 2013.
- Pascucci: KASUT Visualization Summit, Saudi Arabia, April 16, 2013.
- Pascucci: Stanford Linear Accelerator, Menlo Park CA , March 7, 2013.
- Pascucci: ADOBE, Lehi, UT, December 5, 2012.
- Pascucci: Workshop at SIGGRAPH ASIA 2012, Singapore, November 26, 2012.
- Pascucci: Exascale Research Conference, Arlington VA, October 2, 2012.
- Pascucci: DOE CScADS, Snowbird UT, August 1, 2012.
- Pascucci: High Performance Computing, Grids, and Clouds, Cetraro Italy, June 28, 2012.
- Pascucci: Algorithms in the Field at the ACM Symposium on Computational Geometry, UNC, June 19 , 2012.
- Pascucci: BATTELLE headquarters, Columbus OH, May 7, 2012.
- Pascucci: Idaho National Laboratory, Idaho Falls, Apr 12, 2012.
- Pascucci: IAMCS-KAUST Workshop on Computational Biomedicine and Geophysics, Utah, April 5, 2012.
- Pascucci: Institute for Science and Technology (IST), Austria, January 27, 2011.
- Pascucci: Commissariat l'nergie Atomique (CEA) TERA100 High Performance Computing Center, Arpajon, France , February 3, 2011.
- Pascucci: Visualization in Computational Bioscience, February 24, 2011, Texas A&M University, TX.
- Pascucci: CScADS meeting in Lake Tahoe, CA, 2011
- Pascucci: 2nd National Conference in Advancing Tools and Solutions for Nuclear Material Detection, May 3, 2011, Salt Lake City, UT.
- Pascucci: Dagstuhl seminar on Scientific Visualization, Germany, June, 2011.
- Pascucci: Keynote at IX Congress of the Peruvian Society of Computing, October 14, 2010.
- Pascucci: Plenary talk at 23rd SIBGRAPI Conference on Graphics, Patterns and Images, August 31, 2010.
- Pascucci: at the SIBGRAPI 2010 conference
- Pascucci: SCIDAC 2010 conference
- Levine: IRTG meeting in Utah 2010

- **11 refereed conference presentations:** [BNPS10], [WGBP], [BJB⁺11a] (winner of best paper award), [GBP],[GLP11],[JBB⁺11a],[TLB⁺11], [CWRP12], [LLBP12], [WSPVJ11], [GPPR12].

Software Packages (MATLAB, C, C++)

- Demonstration of uncertainty in vector fields.

- Visualization of the structure of 3D Morse-Smale complexes.
- Visualization of fuzzy isosurface in large and/or uncertain data sets.
- Identification of prominent topological features in uncertain data.
- Interactive visualization of discretely sampled high-dimensional scalar fields.
- Aggregation and visualization of feature-based statistics (length scales, descriptive statistics).

Acknowledgements

As always, we humbly thank DOE for their support.

References

- [BHPR10] Osert Bastani, Chris Hillar, Dimitar Popov, and Maurice J. Rojas. Sums of squares, discriminants, and randomization in sparse real root counting. *submitted to AMS Contemporary Mathematics*, 2010.
- [BJB⁺11a] Harsh Bhatia, Shreeraj Jadhav, Peer-Timo Bremer, Guoning Chen, Joshua A. Levine, Luis Gustavo Nonato, and Valerio Pascucci. Edge maps: Representing flow with bounded error. In *4th Pacific Visualization Symposium*, pages 75–82, Hong Kong, China, March 2011. IEEE. **Winner of best paper award.**
- [BJB⁺11b] Harsh Bhatia and Shreeraj Jadhav and Peer-Timo Bremer and Guoning Chen and Joshua A. Levine and Luis Gustavo Nonato and Valerio Pascucci, Flow Visualization with Quantified Spatial and Temporal Errors Using Edge Maps *IEEE Trans. Vis. Comput. Graph.*, 18(9), pp. 1383-1396, 2012.
- [BKL⁺11a] J. C. Bennett, V. Krishnamoorthy, S. Liu, R. W. Grout, E. R. Hawkes, J. H. Chen, J. Shepherd, V. Pascucci, and P.-T. Bremer. Feature-based statistical analysis of combustion simulation data. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1822-1831, 2011.
- [BKL⁺11b] J. C. Bennett, V. Krishnamoorthy, S. Liu, R. W. Grout, E. R. Hawkes, J. H. Chen, J. Shepherd, V. Pascucci, and P.-T. Bremer. Feature-based statistical analysis of combustion simulation data. *IEEE Transactions on Visualization and Computer Graphics*, 17(12):1822-1831, 2011
- [BNPS10] M. Berger, L. G. Nonato, V. Pascucci, and C. T. Silva. Fiedler trees for multiscale surface analysis. *Computers & Graphics*, 34(3):272-281, 2010.
- [CWRP12] A. N. M. Imroz Choudhury and Bei Wang and Paul Rosen and Valerio Pascucci Topological analysis and visualization of cyclical behavior in memory reference traces *PacificVis*, pp. 9-16, IEEE Computer Society, 2012.
- [GBHP08] Attila Gyulassy, Peer-Timo Bremer, Bernd Hamann, and Valerio Pascucci. A practical approach to morse-smale complex computation: Scalability and generality. *IEEE Trans. Vis. Comput. Graph.*, 14(6):1619–1626, 2008.

- [GBP] Attila Gyulassy, Peer-Timo Bremer, and Valerio Pascucci. Towards topology-rich analysis and visualization. *Journal of Physics: Conference Series (scidac 2010 proceedings)*.
- [GBPW10] Samuel Gerber, Peer-Timo Bremer, Valerio Pascucci, and Ross Whiteker. Visual exploration of high dimensional scalar functions. In *IEEE Transaction on Visualization and Computer Graphics*, volume 16, pages 1271–1280, 2010.
- [GKK⁺] A. Gyulassy, N. Kotava, M. Kim, C. Hansen, H. Hagen, and V. Pascucci. Direct feature visualization using Morse-Smale complexes. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1549-1562, September 2012.
- [GLP11] Attila Gyulassy, Joshua A. Levine, and Valerio Pascucci. Visualization of discrete gradient construction (multimedia submission). In *27th Symposium on Computational Geometry*, Paris, France, June 2011. ACM.
- [GPPR12] Attila Gyulassy and Valerio Pascucci and Tom Peterka and Robert B. Ross, The Parallel Computation of Morse-Smale Complexes Proc. IEEE International Parallel and Distributed Processing Symposium (26th IPDPS'12), pp. 484-495, IEEE Computer Society, May 2012.
- [JBB⁺10] Shreeraj Jadhav, Harsh Bhatia, Peer-Timo Bremer, Joshua A. Levine, Luis Gustavo Nonato, and Valerio Pascucci. Consistent approximation of local flow behavior for 2D vector fields using edge maps. Technical Report UUSCI-2010-004, SCI Institute, University of Utah, December 2010.
- [JBB⁺11a] Shreeraj Jadhav, Harsh Bhatia, Peer-Timo Bremer, Joshua A. Levine, Luis Gustavo Nonato, and Valerio Pascucci. Consistent approximation of local flow behavior for 2D vector fields using edge maps. In *4th Workshop on Topology-Based Methods in Data Analysis and Visualization*, Zurich, Switzerland, April 2011.
- [JBB⁺11b] J. A. Levine, S. Jadhav, H. Bhatia, V. Pascucci, and P.-T. Bremer. A quantized boundary representation of 2D flows. *Comput. Graph. Forum*, 31(3):945-954, 2012.
- [LLBP12] Shusen Liu and Joshua A. Levine and Peer-Timo Bremer and Valerio Pascucci. Gaussian mixture model based volume visualization LDAH, pp. 73-77, IEEE, 2012.
- [nIRR10] Martín Avenda, Ashraf Ibrahim, Maurice J. Rojas, and Korben Rusek. Randomized n -completeness for p -adic rational roots of sparse polynomials in one variable. In *ISSAC 2010*. ACM Press, 2010.
- [PRS10] Mikael Passare, Maurice J. Rojas, and Boris Shapiro. New multiplier sequences via discriminant amoebae. *Journal on Foundations of Computational Mathematics*, 2010.
- [PRT10] Philippe Pébay, Maurice J. Rojas, and David C Thompson. Optimizing n -variate $(n+k)$ -nomials for small k . *Theoretical Computer Science*, 2010.
- [PTB10] P. Pébay, D. Thompson, and J. Bennett. Computing contingency statistics in parallel: Design trade-offs and limiting cases. In *2010 IEEE International Conference on*

Cluster Computing, September 2010.

- [PTBM11] Philippe Pébay, David Thompson, Janine Bennett, and Ajith Mascarenhas. Design and performance of a scalable, parallel statistics toolkit. In *12th IEEE International Workshop on Parallel and Distributed Scientific and Engineering Computing (PDSEC-11)*, 2011.
- [RS10] Maurice J. Rojas and Swaminathan Sethuraman. Refined asymptotics for multigraded sums of squares. *submitted to special issue of Theoretical Computer Science*, 2010.
- [TLB⁺11] David C. Thompson, Joshua Levine, Janine C. Bennett, Peer-Timo Bremer, Attila Gyulassy, Valerio Pascucci, and Philippe Pébay. Analysis of large-scale scalar data using hixels. In *IEEE Symposium on Large-Scale Data Analysis and Visualization, LDAV*, pp. 23-30, IEEE, 2011.
- [TLB⁺12] D. Thompson, J. A. Levine, J. Bennett, P.-T. Bremer, A. Gyulassy, V. Pascucci, and P. Pebay. Analysis of uncertain scalar data with hixels. In *Dagstuhl Report*, 2012.
- [WBGPF] Gunther H. Weber, Peer-Timo Bremer, Attila Gyulassy, and Valerio Pascucci. Topology-based feature definition and analysis. In *Numerical Modeling of Space Plasma Flows: ASTRONUM-2010 ASP Conference Series*.
- [WPB⁺11] S Williams, M Petersen, PT Bremer, M Hecht, V Pascucci, J Ahrens, M Hlawitschka, and B Hamann. Adaptive Extraction and Quantification of Geophysical Vortices. *IEEE Trans. Vis. Comput. Graph*, 17(12), pp. 2088-2095, 2011.
- [WSPVJ11] Bei Wang, Brian Summa, Valerio Pascucci, and Mikael Vejdemo-Johansson. Branching and circular features in high dimensional data. *IEEE Trans. Vis. Comput. Graph*, 17(12), pp. 1902-1911, 2011.
- [WSPVJ11] Wathsala Widanagamaachchi and Cameron Christensen and Peer-Timo Bremer and Valerio Pascucci, Interactive exploration of large-scale time-varying data using dynamic tracking graphs, LDAV, pp. 9-17, IEEE, 2012.