

Topological Data Analysis Made Easy with the Topology ToolKit, A Sequel

Martin Falk, Christoph Garth, Charles Gueunet, Joshua A Levine, Jonas Lukasczyk, Julien Tierny, Jules Vidal

To cite this version:

Martin Falk, Christoph Garth, Charles Gueunet, Joshua A Levine, Jonas Lukasczyk, et al.. Topological Data Analysis Made Easy with the Topology ToolKit, A Sequel. IEEE VIS Tutorials, Oct 2019, Vancouver, Canada. hal-02267191

HAL Id: hal-02267191 <https://hal.science/hal-02267191v1>

Submitted on 19 Aug 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Topological Data Analysis Made Easy with the Topology ToolKit, A Sequel

Figure 1: TTK is a software platform for topological data analysis in scientific visualization. It is both easily accessible to end users (ParaView plugins (a), VTK-based generic GUIs (b) or command-line programs (c)) and flexible for developers (Python (d), VTK/C++ (e) or dependence-free C++ (f) bindings). TTK provides an efficient and unified approach to topological data representation and simplification, which enables in this example a discrete Morse-Smale complex (a) to comply to the level of simplification dictated by a piecewise linear persistence diagram (bottom-right linked view, a). Code snippets are provided (d-f) to reproduce this pipeline.

1 LEVEL OF THE TUTORIAL

This tutorial is targeted at a *Beginner* audience.

2 POTENTIAL SCHEDULE CONFLICTS

If possible, we would like to avoid any scheduling overlap with IEEE LDAV 2019 which takes place on October 21. Coauthor Tierny is a co-chair of LDAV this year, and many of the speakers, as well as the potential attendees, are members of both communities.

3 ABSTRACT

This tutorial presents topological methods for the analysis and visualization of scientific data from a user's perspective, with the Topology ToolKit (TTK), a recently released open-source library for topological data analysis. Topological methods have gained considerably in popularity and maturity over the last twenty years and success stories of established methods have been documented in a wide range of applications (combustion, chemistry, astrophysics, material sciences, etc.) with both acquired and simulated data, in both post-hoc and in-situ contexts. Last year, we held the first iteration of this tutorial, that aimed to cover this area at *a software level* and from *a user's point-of-view*. This tutorial aims to continue to fill a gap by providing a beginner's introduction to topological methods for practitioners, researchers, students, and lecturers. In particular, instead of focusing on theoretical aspects and algorithmic details, this tutorial focuses on how topological methods can be useful in practice for concrete data analysis tasks such as segmentation, feature extraction or tracking. The tutorial describes in detail how to achieve these tasks with TTK. First, after an introduction to topological methods and their application in data analysis, a brief overview of TTK's main entry point for end users, namely ParaView, will be presented. Second, an overview of TTK's main features will be given. A running example will be described in detail, showcasing how to access TTK's features via ParaView, Python, VTK/C++, and C++. Third, hands-on sessions will concretely show how to use TTK in ParaView for multiple, representative data analysis tasks. Fourth, the usage of TTK will be presented for developers, in particular

by describing several examples of visualization and data analysis projects that were built on top of TTK. Finally, some feedback regarding the usage of TTK as a teaching platform for topological analysis will be given. Presenters of this tutorial include experts in topological methods, core authors of TTK as well as active users, coming from academia, labs, or industry. A large part of the tutorial will be dedicated to hands-on exercises and a rich material package will be provided to the participants. This tutorial mostly targets students, practitioners and researchers who are not necessarily experts in topological methods but who are interested in using them in their daily tasks. We also target researchers already familiar to topological methods and who are interested in using or contributing to TTK. We kindly ask potential attendees to optionally pre-register at the following address, in order for us to reach out to them ahead of the tutorial with information updates (for instance, last minute updates, instructions for the download of the tutorial material package, etc.): <https://forms.gle/gn7yn3JwzdBN4Mgr7>

Tutorial web page (including all material, TTK pre-installs in virtual machines, code, data, demos, video tutorials, slides, etc): <https://topology-tool-kit.github.io/ieeeVis2019Tutorial.html>

4 TUTORIAL ORGANIZATION

Motivations As scientific datasets become more intricate and larger in size, advanced data analysis algorithms are needed for their efficient visualization and exploration. For scalar field visualization, topological analysis techniques [23, 32, 39] have shown to be practical solutions in various contexts by enabling the concise and complete capture of the structure of the input data into highlevel *topological abstractions* such as contour trees [5, 15, 16], Reeb graphs [30, 31, 42], or Morse-Smale complexes [8, 20, 21]. Such topological abstractions are fundamental data structures that enable advanced data analysis, exploration and visualization techniques, including for instance: small seed set extraction for fast isosurface traversal [6, 34, 45], feature tracking [36], transfer function design for volume rendering [49], data simplification [43] and compression [37], similarity estimation [11, 44], geometry processing [40, 47], or application-driven segmentation and analysis tasks. Successful applications in a variety of fields of science (combustion [4, 19, 26], fluid dynamics [7, 25], material sciences [13, 22, 27], chemistry [3, 14], and astrophysics [35, 38]) have been documented,

which further demonstrates the importance of these techniques.

While reference textbooks have been published [9], up until last year with the first iteration of this tutorial [12] no tutorial has covered this area in recent years at IEEE VIS. The latest tutorial related to topology occurred nearly 10 years ago [48]. Moreover, topological methods have not yet been widely adopted as a standard data analysis tool for end users and developers. We believe one of the reasons for this is the lack of open-source software that implement these algorithms in a generic, user-friendly, and efficient way. Recently, the Topology ToolKit (TTK) [10, 41] has been released (BSD license) to fill this gap. TTK is written in C++ (\sim 100k lines). 11 institutions have contributed to its development so far, including 8 academic institutions and 3 private companies. TTK is supported under Linux, MacOS, and Windows. Since it release in 2017, TTK's website collected 150k page-views, from 15k visitors. These statistics indicate that a user base exists and that further efforts towards the explanation of TTK's usage would be beneficial to the community. Thus, our motivation is to introduce to beginners how topological methods can be useful for analyzing data, and how to do it with TTK.

Target audience This tutorial mostly targets beginners, students, practitioners, and researchers who are not experts in topological methods but who are interested in using them in their daily tasks. It also targets researchers already familiar to topological methods and who are interested in using or contributing to TTK.

Tutorial goals The goals of this tutorial are to present the key tools in topological data analysis (the Persistence diagram, the Reeb graph and its variants, the Morse-Smale complex, etc.) and how they can be used in practice for precise data analysis tasks, including data segmentation and feature extraction. All examples will be illustrated with TTK. This tutorial also aims at presenting TTK's different usage modalities (ParaView, Python, VTK/C++, C++). We expect participants to become capable of using TTK independently, at least with ParaView (possibly with Python), after attending the tutorial.

Hands-on material A large part of the tutorial will be dedicated to hands-on exercises with TTK and ParaView [1]. We will provide a rich material package including TTK pre-installs in virtual machines (to be used by attendees during the tutorial), code, data, demos, video tutorials, etc. Most of this material is already available on TTK's website [12]. Our idea is that participants who bring a laptop will be able to follow along, regardless of their native OS. Attendees who attend just to listen and learn will also benefit from the tutorial and receive sufficient material to try out our examples at home.

Optional pre-registration Last year, we observed that many of our attendees would benefit from having time before the tutorial to install materials and have their systems set up with the hands-on material. If our proposal is accepted, we will use the following on-line form [\(https://forms.gle/gn7yn3JwzdBN4Mgr7\)](https://forms.gle/gn7yn3JwzdBN4Mgr7) to notify potential audiences of where and how to download the material and build a mailing list to help form an informal community for the tutorial.

Proposal strengths In contrast to previous tutorials on topological methods [48], we believe this proposal to have a unique concrete and applicative appeal, by its focus on the *usage* of topological methods rather than on their *foundations*. Thus, we expect it to attract a larger audience than the specific subset of IEEE VIS attendees typically found in traditional topology sessions.

We believe that the list of presenters is also a strength of this proposal. First, it includes topology experts as well as core developers and users of TTK. More importantly, it includes researchers with a variety of experience profiles (Ph.D. students, post-docs, professors) and backgrounds (industry, labs, academia), which will ease interactions with a potentially heterogeneous audience. Moreover, the particularly large number of presenters has two merits. First, it imposes a mini-symposium structure, where speakers will give presentation lasting between 10 and 20 minutes, which will result in a lively rhythm in the overall tutorial. Second, this large number of presenters will be instrumental during the hands-on exercises, as

there will be enough presenters such that one presenter can assist a small group of attendees (typically 3 to 4). Finally, we believe the detailed program of the tutorial (see Sect. 4) achieves a balance between concepts, usage descriptions and application examples.

Detailed content The tutorial is divided into three main parts (each part being subdivided into modules), for a target duration of approximately 3 hours and a half. These three groups of modules can be organized differently to to fit any standard structure for breaks to match the tutorial schedule of IEEE VIS. After the tutorial concludes, we will make all content available from our website.

Followup Survey Since this would be the second iteration of this tutorial, and we hope to continue to run it in future years, we also plan to collect survey data from the participants to help provide feedback to the organizers and shape the content and structure of future tutorials, that might be hosted at additional venues.

A. Preliminaries (60 minutes)

A1. General introduction *(5 minutes, by Julien Tierny)*

A2. TTK docker installation *(5 minutes, by Christoph Garth)*

A3. Introduction to topological methods for data analysis *(30 minutes, by Joshua Levine)* This talk will present the core tools in topological data analysis (the Persistence diagram [9], the Reeb graph and its variants [5, 16–18, 31, 42], the Morse-Smale complex [8, 20, 21]). In particular, it will detail how these tools can be used for data segmentation and feature extraction.

A4. Quick introduction to ParaView's user interface *(25 minutes, by Charles Gueunet)* This talk will provide a brief description of ParaView's main interface [1], in order to support its usage for beginners in the subsequent hands-on session. This will cover the usage of filters, pipeline design, view manipulation, and and Python exports.

B. Hands-on exercises (70 minutes)

B1. General usage of TTK *(10 minutes, by Julien Tierny)* This talk will briefly describe TTK's usage philosophy. It will briefly present how TTK can be used from ParaView, Python, VTK/C++ or C++.

B2. Running TTK with Docker *(5 minutes, by Christoph Garth)*

B3. Segmenting medical data with merge trees *(15 minutes, by Charles Gueunet)* This hands-on TTK/ParaView exercise will be a step-by-step tutorial showing how to extract individual bones in a medical CT scan interactively with merge trees.

B4. Extracting filament structures with the Morse-Smale complex *(15 minutes, by Julien Tierny)* This hands-on TTK/ParaView exercise will show step-by-step how to extract filament structures with the Morse-Smale complex on chemistry data.

B5. Topology-aware data compression *(10 minutes, by Charles Gueunet)* This exercise will show step-by-step how to compress data while guaranteeing feature preservation [37].

B6. Distances, barycenters and clusters *(10 minutes, by Jules Vidal)* This exercise will show step-by-step how to efficiently compare and cluster data sets based on their persistence diagram [46].

C. Advanced usage (80 minutes)

C1. TTK's architecture and core data structures *(10 minutes, by Julien Tierny)* This talk will present TTK's architecture.

C2. Post Hoc Analysis with TTK *(20 minutes, by Jonas Lukasczyk)* Large-scale simulations pose additional challenges to the already complex task of feature tracking and visualization, since the vast number of features and the size of the simulation data make it infeasible to naively identify, track, analyze, render, store, and interact with data. These limitations necessitate *in situ* algorithms that store, at simulation runtime, the least amount of information needed to still support flexible *post hoc* analysis. TTK contains several modules that enable the *in situ* generation of data products (value and depth images [2,33], meta-graphs [50], contour trees [16], etc.) that can be used *post hoc* to interactively generate nested tracking graphs [29], and approximate 3D views of the simulation data [28].

C3. TTK Integration into Inviwo *(20 minutes, by Martin Falk)* Inviwo [24] is a rapid prototyping framework for data visualization. TTK has recently been integrated into Inviwo to extend it with topological analysis. In this talk, we will provide a brief introduction on the concepts utilized in Inviwo like its network editor and the associated data flow paradigm. We then detail our approach of integrating TTK into the data flow within Inviwo, which involves transforming data structures from Inviwo to TTK and back. The seamless integration is demonstrated with a number of examples.

C4. TTK as a teaching platform *(15 minutes, by Joshua Levine)* This talk will provide feedback about our experience in using TTK in our topological data analysis classes.

C5. Panel Discussion *(10 minutes, Multiple speakers)* Open discussion about TTK, its perception from an end user's point of view, its development, its place in our research community.

C6. Concluding remarks *(5 minutes, by Julien Tierny)*

5 BACKGROUND AND CONTACT INFORMATION

The current list is all confirmed speakers. Additional speakers may be added after they confirm their trip to IEEE VIS.

Martin Falk – *martin.falk@liu.se* – is a research fellow in the Scientific Visualization Group at Linköping University. He received his Ph.D. degree (Dr.rer.nat.) from the University of Stuttgart in 2013. His research interests include large-scale volume rendering, visualizations for systems biology, large spatio-temporal data, topological analysis, glyph-based rendering, and GPU-based simulations.

Christoph Garth – *garth@cs.uni-kl.de* – is a professor of computer science at Technische Universität Kaiserslautern, and head of the scientific visualization group there. His research interests encompass the visualization and analysis of large scale data sets using methods from topological analysis, feature extraction, visual analytics, and high-performance computing, among others. In this context, he has employed TTK in teaching, to provide students with an in-depth understanding of topological methods, as well as for his research, as a robust and mature basis to develop novel visualization algorithms. Charles Gueunet – *charles.gueunet@kitware.com* – is currently an R&D engineer at Kitware. He received his PhD from Sorbonne

Université (Paris, France) in 2019. He worked on high performance topological data analysis using level-set based abstractions. He is the author of several contour tree and Reeb graph modules in TTK. Joshua A. Levine – *josh@email.arizona.edu* – is an assistant professor in the Department of Computer Science at University of Arizona. Prior to starting at Arizona, he was an assistant professor at Clemson University, and before that a postdoctoral research as-

sociate at the University of Utah's SCI Institute. He received his PhD from The Ohio State University. His research interests include visualization, geometric modeling, topological analysis, mesh generation, vector fields, performance analysis, and computer graphics.

Jonas Lukasczyk – $il@jluk.de$ – is a Ph.D. student in the Visual Information Analysis Group at the Technical University of Kaiserslautern (TU-KL). He received his Bachelor and Master degree in Computer Science at the TU-KL in 2012 and 2015. His work focuses on Topology-Based Visual Analytics of Large-Scale Simulations. His research interests include topological analysis, HPC, visualization, and web technologies. Several of his approaches use TTK as a backbone for reliable and reproducible data analysis.

Julien Tierny – *julien.tierny@sorbonne-universite.fr* – received the Ph.D. degree in Computer Science from Lille 1 University in 2008 and the Habilitation degree (HDR) from Sorbonne Universite in 2016. He is currently a CNRS permanent research scientist, affiliated with Sorbonne Universite. Prior to his CNRS tenure, he held a Fulbright fellowship (U.S. Department of State) and was a post-doc researcher at the SCI Institute at the University of Utah. His expertise includes topological data analysis for scientific visualization. He is the lead developer of the Topology ToolKit (TTK).

Jules Vidal – *jules.vidal@sorbonne-universite.fr* – is a Ph.D. student at Sorbonne Universite. He received the engineering degree in 2018 from ENSTA ParisTech. His notable contributions to TTK

include the efficient and progressive approximation of distances, barycenters and clusterings of persistence diagrams [46].

ACKNOWLEDGMENTS

This work is partially supported by the European Commission grant H2020- FETHPC "VESTEC" (ref. 800904). This work is also supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, under Award Number(s) DE-SC-0019039.

REFERENCES

- [1] J. Ahrens, B. Geveci, and C. Law. Paraview: An end-user tool for large-data visualization. *The Visualization Handbook*, pp. 717–731, 2005. [2] J. Ahrens, S. Jourdain, P. O'Leary, J. Patchett, D. H. Rogers, and M. Petersen. An Image-Based Approach to
- Extreme Scale In Situ Visualization and Analysis. In *Proceedings of the International Conference for High Perfor-*
- mance Computing, Networking, Storage and Analysis, pp. 424-434. IEEE Press, 2014.

[3] H. Bhatia, A. G. Gyulassy, V. Lordi, J. E. Pask, V. Pascucci, and P.-T. Bremer. Topoms: Comprehensive topological

exploration for mole
- simulations using topology-based data segmentation. IEEE TVCG, 2011.
[5] H. Carr, J. Snoeyink, and U. Axen. Computing contour trees in all dimensions. In Symp. on Dis. Alg., 2000.
[6] H. Carr, J. Snoeyink, and M. van de Pa
- *IEEE VIS*, 2004.
- [7] F. Chen, H. Obermaier, H. Hagen, B. Hamann, J. Tierny, and V. Pascucci. Topology analysis of time-dependent
multi-fluid data using the reeb graph. Co*mputer Aided Geometric Design,* 2013.
[8] L. De Floriani, U. Fugacci
-
- analysis: discrete models and algorithms. Comp. Grap. For., 2015.

[9] H. Edelsbrunner and J. Harer. Computational Topology: An Introduction. American Mathematical Society, 2009.

[10] M. Falk, G. Favelier, C. Gueunet, P. 2019.
- [11] G. Favelier, N. Faraj, B. Summa, and J. Tierny. Persistence Atlas for Critical Point Variability in Ensembles. *IEEE TVCG*, 2018. [12] G. Favelier, C. Gueunet, A. Gyulassy, J. Jomier, J. Levine, J. Lukasczyk, D. Sakurai, M. Soler, J. Tierny, W. Usher,
- and Q. Wu. Topological data analysis made easy with the Topology Toolkit. In IEEE VIS Tutorials, 2018. [https:](https://topology-tool-kit.github.io/ieeeVisTutorial.html)[//topology-tool-kit.github.io/ieeeVisTutorial.html](https://topology-tool-kit.github.io/ieeeVisTutorial.html).
[13] G. Favelier, C. Gueunet, and J. Termy. Visualizing ensem
- actions in chemical systems. *IEEE TVCG*, 2014.
- [15] C. Gueunet, P. Fortin, J. Jomier, and J. Tierny. Contour forests: Fast multi-threaded augmented contour trees. In *IEEE LDAV*, 2016.
- [16] C. Gueunet, P. Fortin, J. Jomier, and J. Tierny. Task-based Augmented Merge Trees with Fibonacci Heaps,. In
- *IEEE LDAV*, 2017. [17] C. Gueunet, P. Fortin, J. Jomier, and J. Tierny. Task-based Augmented Contour Trees with Fibonacci heaps. *IEEE TPDS*, 2019.
- [18] C. Gueunet, P. Fortin, J. Jomier, and J. Tierny. Task-based Augmented Reeb Graphs with Dynamic ST-Trees. In *Eurographics Symposium on Parallel Graphics and Visualization*, 2019. [19] A. Gyulassy, P. Bremer, R. Grout, H. Kolla, J. Chen, and V. Pascucci. Stability of dissipation elements: A case
- study in combustion. *Comp. Graph. For.*, 2014. [20] A. Gyulassy, P. T. Bremer, B. Hamann, and V. Pascucci. A practical approach to morse-smale complex computa-
- tion: Scalability and generality. *IEEE TVCG*, 2008. [21] A. Gyulassy, D. Guenther, J. A. Levine, J. Tierny, and V. Pascucci. Conforming morse-smale complexes. *IEEE*
- *TVCG*, 2014.
- [22] A. Gyulassy, V. Natarajan, M. Duchaineau, V. Pascucci, E. Bringa, A. Higginbotham, and B. Hamann. Topologi-
| cally Clean Distance Fields. *IEEE TVCG, 2007.*
| C. Heine, H. Leitte, M. Hlawitschka, F. Iuricich, L. De F
- survey of topology-based methods in visualization. Comp. Grap. For., 2016.
[24] D. Jönsson, P. Steneteg, E. Sundén, R. Englund, S. Kottravel, M. Falk, A. Ynnerman, I. Hotz, and T. Ropinski.
Inviwo a visualization system
- [org/](https://inviwo.org/). [25] J. Kasten, J. Reininghaus, I. Hotz, and H. Hege. Two-dimensional time-dependent vortex regions based on the
- acceleration magnitude. *IEEE TVCG*, 2011.
[26] D. E. Laney, P. Bremer, A. Mascarenhas, P. Miller, and V. Pascucci. Understanding the structure of the turbulent
mixing layer in hydrodynamic instabilities. *IEEE TVCG*, 2006
-
- Viscous fingering: A topological visual analytic approach. In PMVMSP, 2017.
[28] J. Lukasczyk, E. Kinner, J. Ahrens, H. Leitte, and C. Garth. VOIDGA: A View-Approximation Oriented Image
Database Generation Approach. In IEE
-
-
- [29] J. Lukasczyk, G. Weber, R. Maciejewski, C. Garth, and H. Leitte. Nested Tracking Graphs. *CGF*, 2017.
[30] S. Parsa. A deterministic o(m log m) time algorithm for the reeb graph. In Symp. on Comp. Geom., 2012.
[31] V.
- [32] V. Pascucci, X. Tricoche, H. Hagen, and J. Tierny. *Topological Methods in Data Analysis and Visualization:*
17 Hoory, Algorithms and Applications. Springer, 2010.
133] D. Rogers, J. Woodring, J. Ahrens, J. Patchett,
- v1.2. Technical Report LA-UR-17-25072, Los Alamos National Laboratory, 2018.
- [34] E. Santos, J. Tierny, A. Khan, B. Grimm, L. Lins, J. Freire, V. Pascucci, C. Silva, S. Klasky, R. Barreto, and N. Podhorszki. Enabling advanced visualization tools in a web-based simulation monitoring system. In Proc. *IEEE eScience*, 2009. [35] N. Shivashankar, P. Pranav, V. Natarajan, R. van de Weygaert, E. P. Bos, and S. Rieder. Felix: A topology based
-
- framework for visual exploration of cosmic filaments. *IEEE TVCG*, 2016.
[36] M. Soler, M. Plainchault, B. Conche, and J. Tierry. Lifted wasserstein matcher for fast and robust topology
tracking. In IEEE Symposium on Large
- 2018. [38] T. Sousbie. The persistent cosmic web and its filamentary structure: Theory and implementations. *Royal Astronomical Society*, 2011. <http://www2.iap.fr/users/sousbie/web/html/indexd41d.html>.
- [39] J. Tierny. *Topological Data Analysis for Scientific Visualization*. Springer, 2018.
[40] J. Tierny, J. Daniels, L. G. Nonato, V. Pascucci, and C. Silva. Interactive quadrangulation with Reeb atlases and
connectivity
-
- [41] J. Tierny, G. Favelier, J. A. Levine, C. Gueunet, and M. Michaux. The Topology ToolKit. IEEE TVCG, 2017.
 <https://topology-tool-kit.github.io/>.

[42] J. Tierny, A. Gyulassy, E. Simon, and V. Pascucci. Loop surgery for
- [43] J. Tierny and V. Pascucci. Generalized topological simplification of scalar fields on surfaces. *IEEE TVCG*, 2012. [44] J. Tierny, J.-P. Vandeborre, and M. Daoudi. Partial 3D shape retrieval by reeb pattern unfolding. *Comp. Grap. For.*,
- 28:41–55, 2009. [45] M. van Kreveld, R. van Oostrum, C. Bajaj, V. Pasucci, and D. Schikore. Contour trees and small seed sets for
-
- isosurface traversal. In Symp. on Comp. Geom., 1997.
[46] J. Vidal, J. Budin, and J. Tierny. Progressive Wasserstein Barycenters of Persistence Diagrams. IEEE TVCG, 2019.
[47] A. Vintescu, F. Dupont, G. Lavoué, P. Memari,
- [48] G. Weber, P.-T. Bremer, H. Carr, and A. Gyulassy. Scalar topology in visual data analysis. In *IEEE VIS Tutorials*, 2009. [49] G. Weber, S. E. Dillard, H. Carr, V. Pascucci, and B. Hamann. Topology-controlled volume rendering. *IEEE*
- *TVCG*, 2007.
- [50] W. Widanagamaachchi, C. Christensen, V. Pascucci, and P.-T. Bremer. Interactive Exploration of Large-Scale
Time-Varying Data using Dynamic Tracking Graphs. In Large data analysis and visualization (LDAV), 2012 IEEE
Sy