

Preface

The continued improvements in high performance computing and high resolution sensing capabilities are resulting in data of unprecedented size and complexity. Historically topological and statistical techniques have been deployed as independent alternatives in the analysis of a variety of data types. However, the continued increases in size, dimensionality, and number of variables create new challenges that traditional approaches cannot address. New methods that leverage the mutual strengths of both topological and statistical techniques are needed to support the management, analysis and visualization of such complex data.

In an effort to characterize the current challenges and research trends, and to foster collaborations, we organized the Workshop on the Analysis of Large-scale, High-dimensional, and Multivariate Data using Topology and Statistics, held June 12–14 in Le Barp, France. Around 30 researchers from 20 European and American universities, companies, and national research laboratories were in attendance. The program comprised 18 presentations, including a keynote talk by Herbert Edelsbrunner from the Institute of Science and Technology Austria, titled “Persistent Homology: Theory and Practice.” A number of interesting challenges were addressed during the workshop, with presentations covering a wide range of topics, including topological techniques for large data, high-dimensional data analysis, computational challenges, multivariate visualization and analysis techniques.

In this book, we present 16 peer-reviewed chapters, divided into 6 parts as the outcome of this workshop. Parts I and II focus on large-scale data, Parts III and IV focus on multivariate data, and Parts V and VI focus on high-dimensional data. The chapters in Part I include recent results in the area of in-situ and distributed analysis. We start with a distributed-memory algorithm for labeling connected components in simulation data (Harrison et al.), followed by a discussion of in-situ visualization in fluid mechanics (Ribes et al.). Part I concludes with a survey of recent discoveries in sublinear algorithms for extreme-scale data analysis (Seshadhri et al.). Part II focuses on the efficient representation of large functions and includes

a report on optimal general simplification of scalar fields on surfaces (Tierny et al.), an algorithm for piecewise polynomial monotonic interpolation of 2D gridded data (Allemand-Giorgis et al.), and a technique for shape analysis using real functions (Biasotti et al.). The chapters in Part III focus on structural techniques for multivariate data. This part includes a survey on 3D symmetric tensor fields that highlights what we know and where to go next (Zhang and Zhang), followed by a comparison of Pareto Sets and Jacobi Sets (Huttenberger and Garth), and a report on deformations preserving total curvature (Berres et al.). Part IV focuses on classification and visualization of vector fields, and includes a presentation of Lyapunov time for 2D Lagrangian visualization (Sadlo), followed by a survey of geometric algebra for vector field analysis and visualization (Ausoni and Frey). This part concludes with a report on a technique for computing accurate Morse-Smale complexes from gradient vector fields (Gyulassy et al.). Part V includes chapters focused on the exploration of high-dimensional models, including a presentation of high-dimensional sampling techniques (Ebeida et al.), and a report on the realization of regular maps of large genus (Razafindrazaka and Polthier). Lastly, Part VI presents recent results in the analysis of large, high-dimensional systems, and includes a technique for faster localized solving of systems of equations (Anthony et al.), followed by a system for ensemble analysis of electrical circuit simulations (Crossno et al.).



Fig. 1 Workshop participants, June 2013

In summary, this book brings together recent results from some of the most prominent and recognized leaders in the fields of statistics, topology, and computer science. The book's contents cover both theory and application, providing an overview of important key concepts and the latest research trends. The algorithms detailed in this book are broadly applicable and can be used by application scientists to glean insight from complex data sets.

Acknowledgements

The Workshop on the Analysis of Large-Scale, High-Dimensional, and Multivariate Data Using Topology and Statistics was co-organized by the French Atomic Energy Commission and Alternative Energies (CEA), Sandia National Laboratories, and Kitware, SAS. The organizing committee greatly acknowledges Julien Jomier and Charles Marion for their administrative support in preparation for the workshop, Sophie Vivodtzev for her logistical help, and thanks all of the workshop participants for making it such an enjoyable and informative event. Lastly, the editors would like to thank all of the contributors of this volume for their insightful contributions.

Livermore, CA, USA
Le Barp, France
Salt Lake City, UT, USA
July 2014

Janine Bennett
Fabien Vivodtzev
Valerio Pascucci

Topological and Statistical Methods for Complex Data
Tackling Large-Scale, High-Dimensional, and
Multivariate Data Spaces

Bennett, J.C.; Vivodtzev, F.; Pascucci, V. (Eds.)

2015, XV, 297 p. 120 illus., 101 illus. in color.,

Hardcover

ISBN: 978-3-662-44899-1