Materials for the Working Group on Topological Data Analysis
July 2014, EPFL

V. Karpova

June 16, 2014

References

Books

[1] H. EDELSBRUNNER AND J. HARER *Computational topology: An introduction*, AMS **(2009)**. (pdf)

[2] A. ZOMORODIAN *Topology for computing* Cambridge Monographs on Applied and Computational Mathematics (No. 16), **(2009)**. (pdf)

[3] T. KACZYNSKI ET AL. *Computational homology* Springer **(2004)**.

[4] V . PASCUCCI ET AL. *Topological Methods in Data Analysis and Visualization: Theory, Algorithms, and Applications*, Springer **(2010)**.

- This book addresses many other applications of topology in addition to the ones on which we mainly focus here (i.e, persistence homology, barcodes, Mapper). Thus, it is good for broadening the horizon, but might be off-topic, depending on our purposes.

[5] A. ZOMORODIAN *Short course in computational topology*, Advances in applied and computational topology, AMS **(2011)**.

EPFL projects

[6] C. SPAHN AND R. JEITZINER *Topology and data*, Semester project, EPFL **(2013)**.

[7] C. SPAHN *Proximity and Stability in Persistent Homology*, Semester project, EPFL **(2014)**.

[8] R. JEITZINER *Progression analysis of disease applied to breast cancer research*, Master Project, EPFL **(2014)**.

[9] R. BESSE *Application of Computational Homology to the Analysis of Breast Cancer Data*, Master Project, EPFL and Institute for Mathematics and its Applications, University of Minnesota **(2014)**.

Articles: foundations & surveys

[10] G. CARLSSON *Topology And Data*, Bulletin (New Series) of the American Mathematical Society Volume 46, Number 2, **(2009)**. (pdf)

[11] A. ZOMORODIAN *Topological Data Analysis* Proceedings of Symposia in Applied Mathematics, **(2011)**. (pdf)

[12] V. NANDA AND R. SAZDANOVIC *Simplicial models and topological inference in biological systems*, Discrete and Topological Models in Molecular Biology, 109 Natural Computing Series, Springer-Verlag **(2014)**. (pdf)

DSGA, PAD, Mapper

Publications

[13] M. NICOLAU ET AL. *Disease-specific genomic analysis: identifying the signature of pathologic biology* Bioinformatics Vol. 23 no. 8 **(2007)**. (pdf)

   - Introduces the DSGA method.

[14] M. NICOLAU ET AL. *Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival* PNAS Systems Biology vol. 108 no. 17 **(2011)**. (pdf)

   - Introduces the PAD method.

[15] G. SINGH ET AL. *Topological Methods for the Analysis of High Dimensional Data Sets and 3D Object Recognition*, Eurographics Symposium on Point-Based Graphics **(2007)**. (pdf)

   - Introduces Mapper.

Software

DSGA (Disease-Specific Genomic Analysis) and **PAD** (Progression Analysis of Disease) http://www.stanford.edu/~nicolau/SOFTWARE.pdf

A tutorial for PAD is available at http://comptop.stanford.edu/pad/PADtutorial.pdf

Mapper http://comptop.stanford.edu/programs/

Iris (commercial version of Mapper) http://www.ayasdi.com

Persistent homology

Theory

[16] D. COHEN-STEINER ET AL. *Stability of Persistence Diagrams*, Discrete and Computational Geometry, 37(1) **(2007)**. (pdf)

[17] G. CARLSSON AND A. ZOMORODIAN *The theory of multidimensional persistence*, SCG'07, Gyeongju, South Korea, **(2007)**. (pdf)

[18] H. EDELSBRUNNER AND J. HARER *Persistent Homology - a survey* Surveys on Discrete and Computational Geometry. Twenty Years Later, 257-282, Contemporary Mathematics 453, Amer. Math. Soc., Providence, Rhode Island **(2008)**. (pdf)

[19] R. GHRIST *Barcodes: the persistent topology of data*, Bull (New Series) of the AMS Volume 45, Number 1, **(2008)**. (pdf)

[20] P. BUBENIK AND J. SCOTT *Categorification of Persistent Homology*, available on the arXiv http://arxiv.org/pdf/1205.3669v3.pdf **(2014)**.

Computational approach

[21] A. ZOMORODIAN AND G. CARLSSON *Computing persistent homology*, Discrete Comput. Geom, 33 **(2005)**. (pdf)

[22] G. CARLSSON ET AL. *Computing multidimensional persistence*, Algorithms and Computation, Springer Lecture Notes in Computer Science Volume 5878, **(2009)**. (pdf)

Software

For computing persistent homology: [26], [28].

For computing (ordinary) homology of simplicial complexes: [28], [23], [24].

[23] THE CAPD GROUP *CAPD::RedHom*, http://redhom.ii.uj.edu.pl.

[24] THE COMPUTATIONAL HOMOLOGY PROJECT *CHOM*, http://chomp.rutgers.edu.

Building simplicial complexes from data: algorithms and software

Čech complex

See [10] or [12] for definition.

[25] S. DANTCHEV AND I. IVRISSIMTZIS *Efficient construction of the Čech complex*, Comput. Graph., vol36, no.6, **(2002)**. (pdf)

[26] D.MOROZOV *Dionysus software library*, http:www.mrzv.org/software/dionysus

Vietoris-Rips complex

See [10] or [12] for definition.

[27] A. ZOMORODIAN *Fast construction of the Vietoris-Rips complex*, Computer and Graphics, **(2010)**. (pdf)

[28] V. NANDA *Perseus: the persistent homology software*, http://www.math.rutgers.edu/?vidit.

Witness complex

See [10] for definition.

[29] V. DE SILVA AND G. CARLSSON *Topological estimation using witness complexes*, SPBG'04 Proceedings of the First Eurographics conference on Point-Based Graphics, **(2004)**. (pdf)

Alpha complex

[30] H. EDELSBRUNNER *The union of balls and its dual shape*, Discrete Comput. Geom., vol.13, **(1995)**. (pdf)

Articles: recent theoretical developments

Here's a selction from the lists generated by the following searches on the arXiv.

- *"persistent homology"*

[31] A. ROMERO ET AL. *Defining and computing persistent $\mathbb{Z}$-homology in the general case* http://arxiv.org/pdf/1403.7086.pdf (**2014**)

[32] F. BELCHI AND A. MURILLO A_∞-*Persistence* http://arxiv.org/pdf/1403.2395.pdf (**2014**)

[33] A. BLUMBERG ET AL. *Robust statistics, hypothesis testing, and confidence intervals for persistent homology on metric measure spaces* http://xxx.lanl.gov/pdf/1206.4581v2.pdf (**2013**)

[34] P. BUBENIK ET AL. *Metrics for generalized persistence modules* http://arxiv.org/pdf/1312.3829.pdf (**2013**)

[35] P. FROSINI *G-invariant persistent homology* http://arxiv.org/pdf/1212.0655.pdf (**2013**)

[36] P. SAVELIEV *Robustness of topology of digital images and point clouds* http://arxiv.org/pdf/1101.2008v2.pdf (**2011**)

[37] A. TAUSZ AND G. CARLSSON *Applications of zigzag persistence to topological data analysis* http://arxiv.org/abs/1108.3545 (**2011**)

[38] V. DE SILVA ET AL. *Dualities in persistent (co)homology* http://arxiv.org/pdf/1107.5665.pdf (**2011**)

[39] G. CARLSSON, V. DE SILVA AND D. MOZOROV *Zigzag persistent homology and real-valued functions* http://comptop.stanford.edu/u/preprints/zigzags-socg-final.pdf (**2009**)

[40] G. CARLSSON AND V. DE SILVA *Zigzag persistence* http://comptop.stanford.edu/u/preprints/zigzagCdS.pdf (**2009**)

- *"multidimensional persistence"*

[41] M.P. LESNICK *Multidimensional interleavings and applications to topological interference* (PhD Thesis) http://arxiv.org/pdf/1206.1365.pdf (**2012**)

[42] W. CHACHÓLSKI, M. SCOLAMIERO, AND F. VACCARINO *Combinatorial resolutions of multigraded modules and multipersistent homology* http://xxx.lanl.gov/pdf/1206.1819v1.pdf (**2012**)

- *"barcodes"*

[31], [32]

[43] A. ADCOCK ET AL. *The Ring of Algebraic Functions on Persistence Bar Codes*, http://arxiv.org/pdf/1304.0530.pdf (**2013**)

[44] P. BUBENIK *Statistical topological data analysis using persistence landscapes* http://arxiv.org/pdf/1207.6437.pdf (**2013**)

- *"topological data analysis"*

[45] A.J. BLUMBERG AND M.A. MANDELL *Quantitative Homotopy Theory in Topological Data Analysis* http://arxiv.org/pdf/1309.6628.pdf (**2013**)

[46] U. BAUER ET AL. *Measuring Distance between Reeb Graphs* http://arxiv.org/pdf/1307.2839.pdf (**2013**)

Applications of TDA

A recent survey-type paper (written for non-mathematicians):

[47] P.Y. Lum et al. *Extracting insights from the shape of complex data using topology*, Sci Rep 3:1236, **(2013)**. (pdf)

Cancer

[14]

[48] M. Nicolau, A. Levine, and G. Carlsson *Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival,* PNAS 108, 7265-727, **(2011)**. (pdf)

[49] D. De Woskin et al. *Applications of computational homology to the analysis of treatment response in breast cancer patients*, Topology and its Applications 157 **(2010)**. (pdf)

[50] J. Arsuaga et al. *Topological analysis of gene expression arrays identifies high risk molecular subtypes in breast cancer,* in *Applicable Algebra in Engineering, Communication and Computing*, Special issue on Computer Algebra in Algebraic Topology and Its Applications, **(2012)**.

[51] L. Seemann et al. *A Robust Topology-Based Algorithm for Gene Expression Profiling*, ISRN Bioinformatics **(2012)**. (pdf)

Neuroscience

[52] G. Singh et al. *Topological analysis of population activity in visual cortex*, Journal of Vision, vol. 8, no. 8 **(2008)**. (pdf)

[53] Y. Dabaghian et al. *A topological paradigm for hippocampal spatial map formation using persistent homology*, PLoS Comput. Biol., vol. 8, no. 8 **(2012)**. (pdf)

[54] J. Brown, T. Gedeon *Structure of the afferent terminals in terminal ganglion of a cricket and persistent homology*, PLoS ONE, **(2012)**. (pdf)

[55] D. Romano et al. *Topological Methods Reveal High and Low Functioning Neuro-Phenotypes Within Fragile X Syndrome* Human Brain Mapping **(2014)**. (pdf)

Image treatment

[56] G. Carlsson et al. *On the local behavior of spaces of natural images*, International Journal of Computer Vision, (76), 1, **(2008)**. (pdf)

Miscellaneous

[57] A. Colins et al. *A barcode shape descriptor for curve point cloud data*, Computers and Graphics 28 **(2004)**. (pdf)

[58] V. de Silva and R. Ghrist *Coverage in sensor networks via persistent homology*, Algebr. Geom. Topol. 7 **(2007)**. (pdf)

[59] G. R. Bowman et al. *Structural insight into RNA hairpin folding intermediates*, Journal of the American Chemical Society Communications, **(2008)**. (pdf)

[60] D. Horak et al. *Persistent homology of complex networks*, Journal of Statistical Mechanics: Theory and Experiment, **(2009)**. (pdf)

[61] J.M. Chan et al. *Topology of viral evolution* PNAS, vol. 110 no. 46, **(2013)**. (pdf)

[62] J. Gamble and G. Heo *Exploring uses of persistent homology for statistical analysis of landmark-based shape data* Journal of Multivariate Analysis, Elsevier, vol. 101, issue 9, **(2010)**. (pdf)

Articles: recent applications

See the "Recent preprints" section of the Stanford CompTop project (http://comptop.stanford.edu).

[63] M. Bampasidou and T. Gentimis *Modeling Collaborations with Persistent Homology*, available on the arXiv http://arxiv.org/pdf/1403.5346v1.pdf **(2014)**

[64] K. Emmett and R. Rabadan *Characterizing Scales of Genetic Recombination and Antibiotic Resistance in Pathogenic Bacteria Using Topological Data Analysis* http://arxiv.org/pdf/1406.1219.pdf **(2014)**

People

- **Gunnar Carlsson**: Stanford; founder of TDA, co-founder of Ayasdi.
  http://math.stanford.edu/~gunnar/

- **Ayasdi**
  http://www.ayasdi.com

- **Monica Nicolau**: Stanford; developed the DSGA and PAD methods.
  http://www.stanford.edu/~nicolau/

- **Afra Zomorodian**: currently works at D.E. Shaw Group (an investment and technology development firm). Has been working on computational applications of algebraic topology.

- **Pawel Dlotko**: University of Pennsylvania; works on applied computational topology. Also, knows a lot of people in the field and can give good advice on whom to contact for which purpose.
  http://www.math.upenn.edu/~dlotko/index.html

Websites

- Institute of mathematics and its applications (IMA) video resources page. The 2013-2014 program is dedicated to Scientific and Engineering Applications of Algebraic Topology. See http://www.ima.umn.edu/videos/#ap for workshops and lectures.

- Stanford CompTop project dedicated to applied and computational algebraic topology. http://comptop.stanford.edu . Note the "Events" and "Recent preprints" sections on the right of the screen.
  See also the Software list http://comptop.stanford.edu/programs/ for programs, libraries and packages.

Courses

- Course on Computational Topology and Data Analysis, given in 2013 by Tamal K.Dey (Ohio State University, http://www.cse.ohio-state.edu/%7Etamaldey)

  http://www.cse.ohio-state.edu/~tamaldey/course/CTDA/CTDA.html

- The Ayasdi resources page contains a publications list and links to various webinars, slides, presentations.

  http://www.ayasdi.com/resources/