

Mathieu Carrière

Topological Data Analysis and Machine Learning

+1 917-941-5182

✉ mathieu.carriere3@gmail.com

Address: 390 Ft Washington Ave,
10033 NYC, USA

<https://mathieucarriere.github.io/website>

Skype: mathieu.carriere

French and American citizenship

Education

Now **Postdoctoral Research Scientist**, *Rabadan Lab, Columbia University, New York, USA.*

11/2017 **Ph.D. in Applied Mathematics and Informatics**, *DataShape, Inria Saclay, Palaiseau, France.*
Title: On metric and statistical properties of topological descriptors for geometric data.

02/2015 **Engineering Degree**, *Ecole Centrale Paris, Châtenay-Malabry, France.*

12/2014 **M.Sc. in Mathematics, Vision and Learning**, *ENS Cachan, Cachan, France.*

Research contributions and impact

My research focuses on topological data analysis (TDA) and statistical machine learning (ML). I contributed to the analysis of topological descriptors and on the use of ML kernel methods for them with appropriate kernels, signatures and metrics. This work resulted in several scientific articles in top conference proceedings and scientific journals, and has been used in different fields, like bioinformatics and computer graphics. I also implemented part of this work in the C++ GUDHI library.

Research Articles

Machine Learning

- Stable Topological Signatures for Points on 3D Shapes. M. Carrière, S. Oudot, M. Ovsjanikov, *Proceedings of the Eurographics Symposium on Geometry Processing, 2015.*
Used topological descriptors to improve accuracy in 3D point classification on the dataset of Chen et al.
- Sliced Wasserstein Kernel for Persistence Diagrams. M. Carrière, M. Cuturi, S. Oudot, *Proceedings of the International Conference on Machine Learning, 2017.*
Derived a new kernel for Persistence Diagrams improving accuracy on various datasets, such as texture images.
- Are Cycles Important Topological Features? A Case Study in Brain Image Classification. M. Carrière, C. Chen, F. Liu, X. Ni, T. Wu, J. Fan, R. Rabadan, *Preprint, 2019.*
Used Persistence Diagrams and Cycles to improve autism classification from MRIs.
- A General Neural Network Architecture for Persistence Diagrams and Graph Classification. M. Carrière, F. Chazal, Y. Ike, T. Lacombe, M. Royer, Y. Umeda, *Preprint, 2019.*
Defined general Neural Net architecture for Persistence Diagrams with application in graph classification.

Statistical Analysis

- Structure and Stability of the 1-Dimensional Mapper. M. Carrière, S. Oudot, *Journal on Foundations of Computational Mathematics, 2017.*
Defined appropriate metrics to prove the stability of the Mapper clustering algorithm.
- Statistical Analysis and Parameter Selection for Mapper. M. Carrière, B. Michel, S. Oudot, *Journal of Machine Learning Research, 2017.*
Defined confidence intervals and convergence rates for the Mapper clustering algorithm.
- Two-Tier Mapper: a user-friendly clustering method for global gene expression based on topology. R. Jeitziner, M. Carrière, J. Rougemont, S. Oudot, K. Hess, C. Brinken, *Bioinformatics, 2019.*
Used a modified Mapper clustering algorithm for genomics.
- Topological Data Analysis of single-cell Hi-C contact maps. M. Carrière, R. Rabadan, *Proceedings of the Abel Symposium, 2019.*
Used the Mapper clustering algorithm to establish confidence regions for single-cell Hi-C contact maps.

Metric Properties

- Local Equivalence and Intrinsic Metrics between Reeb graphs. M. Carrière, S. Oudot, *Proceedings of the Symposium on Computational Geometry, 2017*.
Established a local equivalence between the bottleneck and the Gromov Hausdorff distances for Reeb graphs.
- On the Metric Distortion of Embedding Persistence Diagrams into separable Hilbert spaces. M. Carrière, U. Bauer, *Proceedings of the Symposium on Computational Geometry, 2019*.
Established negative results about the equivalence of Persistence Diagram distances and Hilbert space metrics.

Skills

Languages French (native), English (professional TOEFL 627/677), Spanish (B1 level).

Code C++, Python (proficient), R, Matlab (prior experience).

Coding projects

- Cover complex module of the C++ GUDHI library, http://gudhi.gforge.inria.fr/doc/latest/group__cover__complex.html.
- sklearn-tda: a scikit-learn compatible Python library implementing various vectorization and kernel methods used in TDA, https://github.com/MathieuCarriere/sklearn_tda.

Grants

- Mobility Grant (1000 euros) from the DAAD exchange program.
- Mobility Grant (1000 euros) from the STIC doctoral school.
- Best Scientific Contribution 2017 (2nd Prize – 600 euros) from the STIC doctoral school.
- Funding Support (1800 dollars) from ICML 2017.
- Thiessé de Rosemont / Schneider Prize (10,000 euros) from Chancellerie des Universités de Paris.

Teaching Activities

I was a teaching assistant for the following courses.

2015–2017 *Topological Data Analysis*, Ecole Polytechnique, Palaiseau, France.

2016–2017 *Basics of Algorithmic and Programming*, Ecole Polytechnique, Palaiseau, France.

Presentations for Workshops and Conferences

I gave presentations at the following international conferences.

06/2015 *Symposium on Geometry Processing*, TU Graz, Graz, Austria.

11/2015 *Journées de Géométrie Algorithmique*, IESC, Cargèse, Corsica.

12/2015 *Computational and Methodological Statistics*, London University, London, UK.

04/2016 *Stochastic Geometry and its Applications*, Université de Nantes, Nantes, France.

06/2016 *Symposium on Computational Geometry*, Tufts University, Boston, USA.

07/2016 *Applied Topology: Methods, Computation and Science*, Politecnica di Torino, Torino, Italy.

08/2017 *International Conference on Machine Learning*, Convention Centre, Sydney, Australia.

09/2017 *France Japan Machine Learning Workshop*, ENS Ulm, Paris, France.

10/2017 *Amazon Graduate Research Symposium*, Amazon Meeting Center, Seattle, USA.

12/2017 *Journées de Géométrie Algorithmique*, Centre Paul Langevin, Aussois, France.

08/2018 *TRIPODS Bootcamp on Topology and Machine Learning*, ICERM, Providence, USA.

01/2019 *AMS Special Session on Topological Data Analysis*, Convention Center, Baltimore, USA.

02/2019 *Applied Algebraic Topology Research Network*, via Bluejeans.

04/2019 *Topology, Geometry and Data Seminar*, OSU, Columbus, USA.

References

Steve Oudot

DataShape team

Inria Saclay

91120 Palaiseau, France

`steve.oudot@inria.fr`

+33 174 854 216

Marco Cuturi

CREST - ENSAE

Université Paris-Saclay

91120 Palaiseau, France

`marco.cuturi@ensae.fr`

+33 170 266 857

Raul Rabadan

Columbia University

New-York, USA

`rr2579@columbia.edu`