Mathurin Massias

Machine Learning researcher

email: mathurin.massias@gmail.com
website: https://mathurinm.github.io

Research experience

Nov. 2021 – present — INRIA (Lyon, France): Researcher (*chargé de recherche*). DANTE Team.
Efficient and frugal machine learning

Jan. 2020 – Oct. 2021 (2 years) — Università di Genova (Genova, Italy): Post-doctoral researcher with L. Rosasco and S. Villa.
Statistical learning and optimisation, designing new implicit regularization methods machine learning and inverse problems
Publications: [2, 3, 9]

Sep. 2016 – Dec. 2019 (3 years) — INRIA (Université Paris-Saclay, France): PhD, supervised by A. Gramfort and J. Salmon.
"*High dimensional sparse regression with heteroscedastic noise: application to neural source localization*", obtained Summa cum laude.
Keywords: optimisation, neuro-imaging, inverse problems, sparsity, high dimension
Publications: [1, 4, 5, 6, 7, 8]

Feb. 2019 – May 2019 (3 months) — U. of Tokyo/RIKEN (Japan), Deep Learning Theory team: intern, supervised by T. Suzuki.
Work on gradient Langevin dynamics for non-convex regression in RKHS
Keywords: stochastic differential equations
Publications: [10]

Education

Sep. 2014 – Apr. 2015 — **ENS Cachan** (Cachan, France): MSc in Machine Learning (MVA)
Summa cum laude (average grade: 16.8/20)

Sep. 2011 – Apr. 2015 — **Ecole Centrale Paris** (Paris, France): Engineering degree
Major in Applied Mathematics and Data Science (average grade: 16.3/20)

Publications

Journal publications

[1] M. Massias, S. Vaiter, A. Gramfort, and J. Salmon. Dual extrapolation for sparse Generalized Linear Models. *Journal of Machine Learning Research*, 21(234):1–33, 2020.

Proceedings of rank A international conferences

[2] C. Molinari, M. Massias, L. Rosasco, and S. Villa. Iterative regularization for convex regularizers. In *AISTATS*, 2021.

[3] Q. Bertrand and M. Massias. Anderson acceleration of coordinate descent. In *AISTATS*, 2021.

[4] M. Massias*, Q. Bertrand*, A. Gramfort, and J. Salmon. Support recovery and sup-norm convergence rates for sparse pivotal estimation. In *AISTATS*, 2020.

[5] P. Ablin, T. Moreau, M. Massias, and A. Gramfort. Learning step sizes for unfolded sparse coding. In *NeurIPS*, 2019.

[6] Q. Bertrand*, M. Massias*, A. Gramfort, and J. Salmon. Concomitant Lasso with repetitions: beyond averaging multiple realizations of heteroscedastic noise. In *NeurIPS*, 2019.

[7] M. Massias, A. Gramfort, and J. Salmon. Celer: a fast solver for the Lasso with dual extrapolation. In *ICML*, 2018.

[8] M. Massias, O. Fercoq, A. Gramfort, and J. Salmon. Heteroscedastic multitask concomitant lasso for sparse multimodal regression. In *AISTATS*, 2018.

Preprints

[9] Q. Bertrand, Q. Klopfenstein, M. Massias, M. Blondel, S. Vaiter, A. Gramfort, and J. Salmon. Implicit differentiation for fast hyperparameter selection in non-smooth convex learning. 2021. URL: https://arxiv.org/abs/2105.01637.

[10] B. Muzellec, K. Sato, M. Massias, and T. Suzuki. Dimension-free convergence rates for gradient Langevin dynamics in RKHS. 2020. URL: https://arxiv.org/abs/2003.00306.

Teaching

Dec. 2021 (30 h)	EMINES Marrakech: Teacher for the one week *Data Science* class.
2020 – 2021 (70 h)	École Polytechnique Executive Education: Teacher for the *Data Science Starter Program*
2019 – 2022 (3 × 42 h)	École Polytechnique/HEC "Data Science for Business" Master: Teacher for the *Python for Data Science* class
2017 – 2019 (2 × 40 h)	Université Paris-Saclay "Data Science" Master: Teaching assistant and partial lecturer for the *Optimization for Data Science* class
2016 – 2017 (56 h)	Télécom Paris: Teaching assistant for: Analysis and Probabilities (MDI 113/114, Bachelor, 10 h), Linear Models (SD 204, Master, 10 h) Machine Learning and Data Mining (MDI 343, Executive Master, 20 h) Practical Machine Learning (SD 207, Master, 10 h), Tools for signals and images (SI 101, Bachelor, 6 h)

Open source software

Summary on my GitHub page: https://github.com/mathurinm

- celer (python implementation of fast algorithms to solve sparse Generalized Linearized Models): lead developer
- sparse-ho (hyperparameter tuning for sparse machine learning models): core developer
- benchopt (automatic benchmarking of optimization packages on standard ML tasks): core developer
- scikit-learn (machine learning in python): contributor
- MNE-python (brain imaging with magneto and electro-encephalographic modalities): contributor

Community service

Reviewer for NeurIPS 2020 (top 10 % reviewer), 2019 (top 400), 2018 (top 800), ICML 2020, 2019, AISTATS 2022, 2021, 2020, SPARS 2019, ACML 2019, 2018, and for JMLR, SIAM Journal on Optimization, IEEE TSP, Signal Processing.

Grants and awards

- 2019: Best PhD prize of Programme Gaspard Monge Optimisation (PGMO)
- 2019: Best PhD prize of Télécom Paris
- 2018: 1500 € from the GdR ISIS to fund a 1 month visit to the University of Washington (Seattle, USA)
- 2018: 1000 € from the STIC doctoral school to fund SPARS 2017 conference and summer school attendance
- 2017: Best presentation award at JDSE conference (Orsay, France)

Selected invited talks

- DANTE seminar (Lyon), 01/2021, *Fast and cheap regularization of inverse problems.*
- GAIA seminar (Grenoble), 12/2020, *Fast and cheap regularization of inverse problems.*
- SIAM Imaging Science (Toronto), 07/2020, symposium *Advances in non-smooth/non-convex optimization for inverse problems.*
- LCSL seminar, Genova, 01/2020: "*Support recovery and sup-norm convergence rates for sparse pivotal estimation*".
- SIERRA team seminar, Inria (Paris), 12/2019: "*The smoothed multivariate square-root Lasso: optimizational and statistical handling of correlated noise*".

- IMAG Probability and Statistics seminar, Université de Montpellier (Montpellier, France), 09/2019: "*Concomitant Lasso with repetitions: smoothing the nuclear norm to handle non homoscedastic noise*".
- Data-Driven Biomedical Science team seminar, Riken AIP (Nagoya), 04/2019: "*Exploiting regularity in sparse Generalized Linear Models solvers*".
- MOKAPLAN team seminar, Inria (Paris), 12/2018: "*Dual extrapolation for sparse Generalized Linear models*".
- MILO team seminar, EPFL (Lausanne), 12/2018: "*Celer: a fast solver for the Lasso with duality improvements*".
- University of Washington (Seattle), 05/2018: "*Solving Lasso-type problems with aggressive Gap safe rules*".
- CMStats (London), 12/2017: "*From safe screening rules to working sets for faster Lasso-type solvers*".

Research summary

In many domains, the size of the collected data is evergrowing. Their analysis requires to solve large scale optimization problems, whose computational cost is a limiting factor. My work makes the resolution of this type of problems easier in principled ways: by accelerating algorithms, by proposing cheaper formulations, and by developing new regularization techniques. It lies at the intersection of statistical learning, optimization and inverse problems.

- In [1, 7], I have proposed a state-of-the-art algorithm to solve sparse problems such as the Lasso and sparse logistic regression, making the resolution faster by at least one order of magnitude. I have relased the `celer` Python library which brings these speed-ups to a wide datascience audience, with scheduled plans to replace `scikit-learn`'s.
- In [3], I have proposed the first parameter and restart-free, practically efficient acceleration of coordinate descent based on Anderson extrapolation. The proposed algorithm has nearly optimal convergence rates, is easy to implement, and applies to many regularized, unregularized or non-convex machine learning problems.
- In [2], I have proposed an implicit/iterative regularization approach for low complexity, non-smooth and non strongly convex penalties. These penalties cover a wealth of popular choices in data science and applications: sparsity, low rank, planeity of images, etc. Instead of the classical penalized approach which requires costly hyperparameter tuning, we solve a single problem, obtaining equivalent solutions up to $\times 100$ speed-ups.
- In [4, 6, 8], I have designed a sparse regression estimator able to handle correlated and repeated measurements. This estimator showed state-of-the-art performance in the neural source estimation problem it was designed for. I have proposed an efficient algorithm to estimate it, and showed that it possessed minimax optimal support recovery rates.