Quentin Bertrand

Education & Experience

2021–now **Postdoctoral researcher**, *Université de Montréal and Mila*, Montréal.
2018–2021 **PhD in Computer Science**, *Inria*, Saclay.
2017–2018 **MS in Computer Science**, *École Normale Supérieure*, Cachan.
2014–2017 **BS and MS in Engineering**, *École polytechnique*, Palaiseau.

Research Highlights

Current research. I am a postdoctoral researcher at Mila under the supervision of Gauthier Gidel and Simon Lacoste-Julien. I work at the intersection of game theory and optimization for fast hyperparameter optimization. In particular:

- I extended the notion of Elo score for cyclic games [12].
- I developed a generic algorithm to efficiently solve sparse linear models [9], code.
- I showed that optimal algorithms for function value minimization are not optimal for unrolled-estimated Jacobians [11].
- I showed that representation learning based on sparse bilevel optimization yields disentangled representations [8].

Previous research. Prior to this position, I completed my PhD [5] in statistics and optimization under the supervision of Joseph Salmon and Alexandre Gramfort (core `scikit-learn` contributor). I worked on model calibration for high dimensional sparse linear regression applied to brain signals reconstruction [4]:

- Coordinate descent algorithms:
  - Proposed Anderson extrapolation to accelerate coordinate descent algorithms [6].
  - Showed support identification and local linear convergence of coordinate descent [7].
- Model calibration as an hyperparameter optimization problem:
  - Developed algorithms for fast hyperparameter optimization of Lasso-type models [3].
  - Provided a high quality python package for model selection: `sparse-ho` [10].
- Model calibration as a statistical problem:
  - Theoretically studied the statistical influence of smoothing parameters for the *square-root Lasso* and the *multivariate square-root Lasso* [2].
  - Formulated optimization problems to handle sparse linear regression with correlated noise as smoothing-based optimization problems, [1], code.

Work Experience

2017 **Stanford Research Institute**, *Research Intern*, Menlo Park, CA.

- Worked on the DARPA project Probabilistic Programming for Advanced Machine Learning.
- Developed and implemented new algorithms to compute exact bounds in graphical models.

References

[1] Q. **Bertrand**, M. Massias, A. Gramfort, and J. Salmon. Handling correlated and repeated measurements with the smoothed multivariate square-root lasso. *NeurIPS*,

2019.

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

[3] Q. **Bertrand**, Q. Klopfenstein, M. Blondel, S. Vaiter, A. Gramfort, and J. Salmon. Implicit differentiation of lasso-type models for hyperparameter optimization. *ICML*, 2020.

[4] P.-A. Bannier, Q. **Bertrand**, J. Salmon, and A. Gramfort. Electromagnetic neural source imaging under sparsity constraints with sure-based hyperparameter tuning. *Medical Imaging meets NeurIPS*, 2021.

[5] Q. **Bertrand**. *Hyperparameter selection for high dimensional sparse learning: application to neuroimaging*. PhD thesis, Université Paris-Saclay, 2021.

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

[7] Q. Klopfenstein, Q. **Bertrand**, A. Gramfort, J. Salmon, and S. Vaiter. Model identification and local linear convergence of coordinate descent. *(Accepted under minor revisions in Optimization Letters)*, 2022.

[8] S. Lachapelle, T. Deleu, D. Mahajan, I. Mitliagkas, Y. Bengio, S. Lacoste-Julien, and Q. **Bertrand**. Synergies between disentanglement and sparsity: a multi-task learning perspective. *arXiv preprint arXiv:2211.14666*, 2022.

[9] Q. **Bertrand**, Q. Klopfenstein, P.-A. Bannier, G. Gidel, and M. Massias. Beyond L1: Faster and better sparse models with skglm. *Advances in neural information processing systems*, 2022.

[10] Q. **Bertrand**, Q. Klopfenstein, M. Massias, M. Blondel, S. Vaiter, A. Gramfort, and J. Salmon. Implicit differentiation for fast hyperparameter selcetion in non-smooth convex learning. *J. Mach. Learn. Res.*, 2022.

[11] D. Scieur, Q. **Bertrand**, G. Gidel, and F. Pedregosa. The curse of unrolling: Rate of differentiating through optimization. *Advances in neural information processing systems*, 2022.

[12] Q. **Bertrand**, W. M. Czarnecki, and G. Gidel. On the limitations of Elo: Real-world games, are transitive, not additive. *AISTATS*, 2023.

Teaching

2022-2023 Game Theory for Machine Learning, Université de Montréal, **Teaching Assistant**, **Guest Lecturer**, Prof.: G. Gidel.

2022 I was responsible for the deep learning part of the first artificial intelligence MOOC all in French. I shoot around ten videos of twenty minutes covering basic (what is a neural network), to advanced techniques (representation / transfer learning).

2020-2021 Optimization for Machine Learning, Data Science Summer School of École polytechnique, **Lecturer**, 6h.

2020-2021 Python for Data Science, MS X-HEC, **Teaching Assistant**, 40h.

2019-2021 Optimization for Data Science, MS Data Science, **Teaching Assistant**, 2*20h, Prof.: A. Gramfort and R. Gower.

2019-2020 Numerical Methods and Applications, BS ENSAE, **Teaching Assistant**, 30h, Prof.: S. M. Kaber.

Awards

2021-2022 Oustanding reviewer award at NeurIPS (top 8%).

2019 NeurIPS travel award, I was awarded a grant from GDRIA to visit Samuel Vaiter.

Miscellaneous

On my free time I like to swim and to play chess (2200 elo).