

Quentin Bertrand

Education & Experience

- 2024–now **Researcher**, *Inria Lyon*.
- 2021–2024 **Postdoctoral Researcher**, *Université de Montréal and Mila*, Montréal.
- 2018–2021 **PhD in Computer Science**, *Inria Paris-Saclay*, 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 recently joined Inria as a permanent researcher. I work on generative models and multi-agent learning. Before I was a postdoctoral researcher at [Mila](#) under the supervision of [Gauthier Gidel](#) and [Simon Lacoste-Julien](#). I work at the intersection of optimization and generative models:

- Representation Learning and Deep Generative Models:
 - I showed that the generalization of deep generative flow matching models [16] comes from the inductive bias of neural networks
 - I wrote a [friendly blog post](#) on normalizing flows and conditional flow matching techniques [13].
 - I showed that [retraining generative models on their own data can yield a stable process](#)[12], additionally I drew connection with reinforcement learning with human feedback [11].
 - I showed that representation learning based on [sparse bilevel optimization yields disentangled representations](#) [6].
- Multi-Agent Learning
 - I proved that Q -learning-guided agents can learn to cooperate in the iterated prisoner's dilemma [15].
 - I extended the notion of [Elo score for cyclic games](#) [10], which has applications to [non-transitive reward models in RLHF](#).
- Optimization:
 - I showed that optimal algorithms for [function value minimization are not optimal for unrolled-estimated Jacobians](#) [9].
 - I developed a [generic algorithm to efficiently solve sparse linear models](#) [7, 14], [code](#).

Previous research. Prior to this position, I completed my PhD [3] 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 [2]:

- Large-scale optimization:
 - Proposed [Anderson extrapolation to accelerate coordinate descent](#) algorithms [4].
 - Showed [support identification and local linear convergence of coordinate descent](#) [5].
- Automatic differentiation:
 - Developed algorithms for [fast hyperparameter optimization of Lasso-type models](#) [1].
 - Provided a [high quality python package](#) for model selection: `sparse-ho` [8].

Internships

- 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**, Q. Klopfenstein, M. Blondel, S. Vaiter, A. Gramfort, and J. Salmon. Implicit differentiation of lasso-type models for hyperparameter optimization. *ICML*, 2020.
- [2] 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.
- [3] Q. **Bertrand**. *Hyperparameter selection for high dimensional sparse learning: application to neuroimaging*. PhD thesis, Université Paris-Saclay, 2021.
- [4] Q. **Bertrand** and M. Massias. Anderson acceleration of coordinate descent. *AISTATS*, 2021.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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. *J. Mach. Learn. Res.*, 2022.
- [9] 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.
- [10] Q. **Bertrand**, W. M. Czarnecki, and G. Gidel. On the limitations of Elo: Real-world games, are transitive, not additive. *AISTATS*, 2023.
- [11] D. Ferbach, Q. **Bertrand**, A. J. Bose, and G. Gidel. Self-consuming generative models with curated data provably optimize human preferences. *NeurIPS*, 2024.
- [12] Q. **Bertrand**, A. J. Bose, A. Duplessis, M. Jiralerspong, and G. Gidel. On the stability of iterative retraining of generative models on their own data. *ICLR*, 2024.
- [13] A. Gagneux, S. Martin, R. Emonet, Q. **Bertrand**, and M. Massias. A visual dive into conditional flow matching. *ICLR Blogpost Track*, 2025.

- [14] Q. **Bertrand** M. Massias, B. MOUFAD. skglm: Improving scikit-learn for regularized generalized linear models. *ICML CODEML Workshop*, 2025.
- [15] Q. **Bertrand**, J. Duque, E. Calvano, and G. Gidel. Self-play q -learners can provably collude in the iterated prisoner's dilemma. *ICML*, 2025.
- [16] Q. **Bertrand**, A. Gagneux, M. Massias, and R. Emonet. On the closed-form of flow matching: Generalization does not arise from target stochasticity. *arXiv preprint arXiv:2506.03719*, 2025.

Teaching

- 2025-2026 [Deep Generative Models](#), ENS Lyon, **Lecturer**
- 2024-2026 [Optimal Transport for Machine and Deep Learning](#), ENS Lyon, **Lecturer**
- 2022-2024 [Adversarial Machine Learning](#), Université de Montréal, **Teaching Assistant**, **Guest Lecturer**, Prof.: [G. Gidel](#).
- 2022 I was responsible for the [deep learning part of the first all-in-French artificial intelligence MOOC](#). 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](#).

Awards

- 2021-2022 [Outstanding 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 (4h a week) and to play chess ([2200 elo](#)).