Rémi Bardenet

✉ remi.bardenet@gmail.com
rbardenet.github.io

Research interests

Probabilistic modelling and inference; Monte Carlo methods; interacting particles; applications to machine learning, signal processing, particle physics, and cell biology.

Academic positions

Feb. 2015- **Chargé de recherche**, *CNRS & CRIStAL, Université de Lille*, Lille (France).
CRIStAL [link] is the dept. of computer science, signal processing and automatic control.

2013-2015 **Postdoctoral fellow**, *Department of Statistics, University of Oxford*, Oxford (UK).
2020 Science fellowship of the EPSRC, working with Chris Holmes on large-scale Markov chain Monte Carlo methods, motivated by biological data analysis.

Education

2009-2012 **Doctorat (Ph.D.)**, *Université Paris-Sud XI*, Orsay (France), *très honorable*.
Computer science, *Towards adaptive learning and inference - Applications to hyperparameter tuning and astroparticle physics*, under the supervision of *Balázs Kégl* [link].

2008-2009 **Master (M.Sc.)**, *Ecole Normale Supérieure*, Cachan (France), *highest honours*.
Mathematics, computer vision and machine learning ('MVA').

2008 **Agrégation**, *Université Louis Pasteur*, Strasbourg (France), *national rank 82*.
Mathematics major, probability & statistics minor.

Recent (2016–) management of funding resources

2020-2024 **National individual Artificial intelligence Chair**, *role:* PI, **880 k€**.
Project BACCARAT "Bayesian learning of expensive models, with applications to cell biology".

2019-2024 **ERC Starting Grant**, *role:* PI, **1.5 M€**.
Project BLACKJACK "Fast Monte Carlo integration with repulsive processes".

2017-2021 **LAI**, *role:* PI, **20 k€**.
Project TUNNELING EFFECT between Univ. Lille and Univ. Oxford (UK).
LAIs ("associated international labs") are initiatives from Univ. Lille to consolidate existing strong international scientific collaborations. They are also sesames to PhD funding from the university.

2016-2020 **ANR JCJC grant**, *role:* PI, **172 k€**.
Project BOB "Bayesian inference on a budget".

2016 **CNRS PEPS JCJC**, *role:* PI, **10 k€**.
Project DPPMC "Monte Carlo with determinantal point processes".

Awards

2013 **2nd prize of the *Gilles Kahn* award**, *Société Informatique de France*.
Awarded by the French society of Computer Science for the best French Ph.D. of the academic year.

Recent (2016–) international seminars and visits

- ○ Acoustics research institute, **Vienna, Austria** (1 week, postponed until after covid)
- ○ Dept. of probability and statistics, National University of Singapore, **Singapore** (2 weeks, postponed until after covid)
- ▷ Machine learning group, CWI, **Amsterdam, Netherlands** (5 days, 2019)
- ▷ European space agency technical labs, **Leiden, Netherlands** (2 days, 2019)
- ▷ Dept. of mathematics, **Univ. Bristol, UK** (3 days, 2017)
- ▷ Dept. of mathematics, **Univ. Aalborg, Denmark** (1 week, 2017)
- ▷ Probabilistic numerics group, **Max Planck Institut Tübingen, Germany** (2 days, 2016);
- ▷ Dept. of statistics, **Univ. Kent, UK** (2 days, 2016);
- ▷ Machine learning group, **Microsoft Sunnyvale, USA** (2 days, 2016);
- ▷ Dept. of medical physiology, **Univ. Utrecht, Netherlands** (3 days, 2016);
- ▷ Dept. of statistics, **Harvard Univ., USA** (3 days + 1 week 2016);
- ▷ Depts. of statistics and of computer science, **Univ. Oxford, UK** (a total of 6 weeks since 2016);

Recent (2016–) national seminars and visits

- ▷ Medical graduate school, **Univ. Lille**, Lille (2020, one day)
- ▷ Dept. of mathematics, **Univ. Paris-Descartes**, Paris (2019, one day)
- ▷ Dept. of mathematics, **Univ. Paris-Saclay**, Paris (2018, one day).
- ▷ *Journée algorithmes stochastiques* [link], **Univ. Paris-Dauphine**, Paris (2017, one day).
- ▷ Dept. of mathematics *Jean Leray*, **Univ. Nantes**, Nantes (2017, two days).
- ▷ Dept. of signal proc. and automatic control *GIPSA-lab*, **Univ. Grenoble-Alpes** (2016, 3 days).
- ▷ Dept. of applied mathematics *Jean Kuntzmann*, **Univ. Grenoble-Alpes** (2016, one day).
- ▷ Project-team *Mistis*, **Inria Grenoble**, Grenoble (2016, one day).

Recent (2016–) invited talks at international workshops or special sessions of international conferences

- ○ Plenary speaker at the international conference on *Mathematical Methods in Physics* [link], Marrakech, Morocco, April 2021 (to come).
- ▷ ICML workshop on *Negative dependence in machine learning*, International Conference on Machine Learning (ICML), Vienna, July 2020 (videoconference).
- ○ Plenary speaker at *Mathematical statistics and learning* [link], Univ. Pompeu Fabra, Barcelona, June 2020 (postponed to 2021).
- ○ Session on *Kernel methods for uncertainty quantification*, SIAM conference on uncertainty quantification, Munich, March 2020 (postponed).
- ▷ Workshop on *Stochastic Geometry, Stereology and Image Analysis*, Sandbjerg, Denmark, June 2019.
- ▷ Workshop on *Point processes in Space, Time and Beyond*, Skagen, Denmark, May 2019.
- ▷ *Workshop* [link] on *Bayesian nonparametrics for signal/image processing*, Bordeaux, France, July 2018.
- ▷ MLSS *African machine learning summer school*, Algiers, Algeria, June 2018.
- ▷ SSP conference on *Statistical signal processing*, Freiburg, Germany, June 2018.
- ▷ Workshop on *Bayesian nonparametrics for signal/image processing*, Bordeaux, France, June 2018.
- ▷ Workshop on *Cardiac modelling* of the Royal Statiscal Society, Chicheley, UK, February 2017.
- ▷ MCQMC conference *Monte Carlo and quasi-Monte Carlo methods*, Stanford Univ., Palo Alto, USA, August 2016.
- ▷ Workshop on *High-Dimensional Statistical Models & Big Data*, Alan Turing Institute, London, UK, February 2016.
- ▷ MCMSki conference on *Monte Carlo methods*, Lenzerheide, Switzerland, January 2016.

Selected recent (2016–) invited talks in national workshops/conferences

- ○ Invited opening speaker at the workshop *DPPs in the Alps*, postponed until after covid.
- ▷ Invited speaker at the French *Académie des Sciences* for a mini-workshop on determinantal point processes, June 2018.
- ▷ Invited speaker at the Physics colloquium of ENS Lyon, France, March 2018.
- ▷ Plenary speaker at the *StatLearn* [link] workshop, Univ. Lyon, April 2017.
- ▷ Plenary speaker at the *GRETSI* [link] conference, September 2017.
  GRETSI is the main French-speaking event on signal processing, held every other year since 1967, with 400+ regular participants. Plenaries are prestigious and usually given by more senior academics.
- ▷ Bayes in Paris *[link] national seminar series*, Paris, April 2016.

Recent (2016–) teaching experience

CNRS positions come with no teaching duty. Out of personal inclination, I still maintain a small teaching activity at the master level.

2019– **Lecturer**, *ENS Paris-Saclay*, Cachan (France).
16 hours per year, on Bayesian machine learning ("MVA" master programme; the most demanded machine learning master nationwide.)

2019– **Lecturer**, *Univ. Lille*, Lille (France).
18 to 24 hours per year, on Model-based machine learning (master-level students in data science).

2016–2019 **Lecturer**, *ENSAE ParisTech*, Paris (France).
9 hours per year, on Bayesian nonparametrics (master-level students in statistics and econometry).

2015–2019 **Lecturer**, *École centrale de Lille*, Lille (France).
15 hours per year, on practical machine learning with applications to bankruptcy prediction (master-level engineering students). I organized this course around a machine learning competition, see e.g. the *2019 edition* [link].

Recent (2016–) scientific responsibilities and research management

2020 Serving as jury president for the French PhD prize in *signal & image processing and computer vision*.

2019 co-organised *GRETSI* [link], the main French-speaking event in signal processing. This is my biggest event so far (∼450 participants).

2019 co-organised an international *workshop on DPPs and fermions* [link] across physics, mathematics, and signal processing at Univ. Lille.

2017–2018 I ran a small workgroup on *reproducible research* in Lille, where software engineers transfer selected skills to researchers to achieve high standards of reproducibility. See our *tutorial* [link] for instance, or a *concrete example* [link] of our standards: our software package on DPP sampling for machine learning. More software activity on my *GitHub account* [link]

2016 co-organised a one-day national workshop with academic and industrial guests on *Big data: modeling, estimation and selection* [link] at Univ. Lille.

2016 organised a special session on *statistical applications of determinantal point processes* [link] at the MAS days of the French society for applied and industrial mathematics (SMAI).

2016– Reviewer for journals such as *Annals of Statistics, Annals of Applied Probability, Journal of the Royal Statistical Society B, Biometrika, Bernoulli, Journal of Machine Learning Research (JMLR), Electronic Journal of Statistics (EJS), Journal of Computational and Graphical Statistics (JCGS)*, and conferences such as *International Conference on Machine Learning (ICML), Advances in Neural Information Processing Systems (NIPS), International conference on learning theory (COLT)*.

Supervision

Percentages display my own share of the supervision.

2020– **Postdoc**, *CNRS*, Lille (France), 100%.
I am supervising Barbara Pascal's postdoc on *Repulsive point processes in signal processing*.

2020– **Ph. D. student**, *Ecole Centrale de Lille*, Lille (France), 50%.
I am co-supervising Yoann Jayer's Ph.D. with Mylène Maida (Prof. Univ. Lille & Laboratoire Painlevé), on *Monte Carlo integration with repulsive stochastic processes*.

2020– **Ph. D. student**, *Univ. Lille*, Lille (France), 50%.
I am co-supervising Diala Hawat's Ph.D. with Raphaël Lachièze-Rey (HDR, Univ. Paris & MAP5), on *Repulsive points patterns in statistical signal processing*.

2019– **Postdoc**, *CNRS*, Lille (France), 100%.
I am supervising Arnaud Poinas's postdoc on *Repulsive point processes in statistics*. Arnaud was co-supervised by Adrien Hardy (Univ. Lille & Laboratoire Painlevé) during his first year of postdoc.

2017–2020 **Ph. D. student**, *Ecole Centrale de Lille*, Lille (France), 50%.
I am co-supervising Ayoub Belhadji's Ph.D. with Pierre Chainais (Prof. Centrale Lille & CRIStAL), on *Determinantal point processes for dimension reduction in signal processing*.

2016–2020 **Ph. D. student**, *Ecole Centrale de Lille*, Lille (France), 50%.
I am co-supervising Guillaume Gautier's Ph.D. with Michal Valko (HDR, Inria Lille & CRIStAL), on *Fast sampling of determinantal point processes*. Guillaume will defend in March 2020.

2014–2018 **Ph. D. student**, *Univ. Oxford*, Oxford (UK), 20%.
I have co-supervised Ross Johnstone's Ph.D. with Gary Mirams and David Gavaghan (Oxford Computer Science), with industrial collaborators at *Roche labs* (Basel, Switzerland) on *Uncertainty characterisation in action potential modelling for cardiac drug safety*.

Recent (2016–) track Record

In this section, I present my research interests in a loose chronological order, restricting to themes in which I have been active in the period 2016–2020.

Scalable Bayesian inference. Loosely speaking, Bayesian inference for big data faces two main challenges:

big data can be *tall* – it contains an overwhelming number of individual data points – and/or it can be *large* – each data point contains a large number of components. Since 2013, I have had a keen interest into scaling up Bayesian inference to tall datasets while maintaining the theoretical properties that make Bayesian inference desirable in the first place. Restricting to the period 2016–2020, I have proposed the first MCMC algorithm that can provably use less than the original number of data points while maintaining the same accuracy as if using the entire dataset [8]. Focusing on large datasets, this time, we obtained efficient reduction of dimensionality in linear regression using randomized column subset selection [1]. These days, I am thinking of going back to tall data with an entirely different approach, rephrasing inference for tall data as a tractable statistical decision problem that is constrained by data access, instead of looking for scalable versions of algorithms that are designed to solve an unconstrained decision problem.

Applications to biology. I have been part of the 2020 science network – a UK-based network of young computational scientists. In the 2016–2020 time frame, I have transferred statistical tools to computational biology, in particular cardiac modelling [43, 41, 40] and cell tissue tracking [42]. Currently, I'm collaborating with US FDA and its British equivalent, to apply our Bayesian methodology in [41] to assess the dangerous side-effects of anti-arrhythmic drugs.

Determinantal point processes. This has been my most active theme in 2016–2020. Determinantal point processes (DPPs) are point process models that naturally encode diversity or repulsiveness between the points of a given realization, using algebraic arguments. They arise in many fields, e.g. random matrix theory, combinatorics, or quantum physics. Our landmark result was to show that DPPs lead to fast Monte Carlo integration [3], leading to a natural stochastic version of Gaussian quadrature, and we have proven faster-than-Monte-Carlo central limit theorems for our estimators. See also [15] for a different description of the train of thought that took us to the results in [3]; [16, 13] for computational improvements to sampling DPPs; and [4] for a Python toolbox on DPP sampling. Another line of attack of the same problem led us to generic bounds for numerical integration in reproducing kernel Hilbert spaces using DPPs [12]. These works open the way towards more efficient Monte Carlo integration for expensive-to-evaluate integrands, such as those arising in inference for complex biological systems. This is the main topic of my ERC starting grant BLACKJACK, which just started in February 2020.

Applications to signal processing. After a stimulating *talk* [link] by Patrick Flandrin, I was intrigued by the seemingly repulsive point process formed by the zeros of the short-time Fourier transform of white noise. We proved [6] that this process is an instance of zeros of a Gaussian analytic function, an object of recent booming interest in the random matrix community. This opened a probabilistic toolbox full of theorems that apply directly to the design of filters for signal processing. This first result sparked a lot of interest and raised new deep questions. In our follow-up work [5], we gave new links to an analytic wavelet transform as well as discrete windowed Fourier transforms. This gives yet another viewpoint and a stimulating new field of application for repulsive point processes such as DPPs. These questions are part of the new ANR project ASCETES (PI Sylvain Meignen), to which I'm participating.

Publications

My research is multidisciplinary, and publishing habits differ among the relevant disciplines. See my *full CV* [link] for a list of notable differences. I group computer science, computational statistics, applied probability, and signal processing under the acronym 'CS' below.

CS journals

[1] A. Belhadji, **R. Bardenet**, and P. Chainais. A determinantal point process for column subset selection. *Journal of Machine Learning Research (JMLR)*, 2020.

[2] G. Gautier, **R. Bardenet**, and M. Valko. Fast sampling of β-ensembles. *Statistics and Computing (in press)*, 2020.

[3] **R. Bardenet** and A. Hardy. Monte Carlo with determinantal point processes. *Annals of Applied Probability*, 2020.

[4] G. Gautier, **R. Bardenet**, G. Polito, and M. Valko. DPPy: Sampling determinantal point processes with Python. *Journal of Machine Learning Research; Open Source Software (JMLR MLOSS)*, 2019.

[5] **R. Bardenet** and A. Hardy. Time-frequency transforms of white noises and Gaussian analytic functions. *Applied and Computational Harmonic Analysis*, 2019.

[6] **R. Bardenet**, J. Flamant, and P. Chainais. On the zeros of the spectrogram of white noise. *Applied and Computational Harmonic Analysis*, 2018.

[7] **R. Bardenet**, F. Lavancier, X. Mary, and A. Vasseur. On a few statistical applications of determinantal point processes. *ESAIM: Proceedings and Surveys*, 60:180–202, 2017.
[8] **R. Bardenet**, A. Doucet, and C. Holmes. On Markov chain Monte Carlo methods for tall data. *Journal of Machine Learning Research (JMLR)*, 2017.
[9] **R. Bardenet** and O.-A. Maillard. Concentration inequalities for sampling without replacement. *Bernoulli*, 2015.
[10] **R. Bardenet**, O. Cappé, G. Fort, and B. Kégl. Adaptive MCMC with online relabeling. *Bernoulli*, 21(3):1304–1340, 2015.

CS refereed proceedings

[11] A. Belhadji, **R. Bardenet**, and P. Chainais. Kernel interpolation with continuous volume sampling. In *International Conference on Machine Learning (ICML)*, 2020.
[12] A. Belhadji, **R. Bardenet**, and P. Chainais. Kernel quadrature with determinantal point processes. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.
[13] G. Gautier, **R. Bardenet**, and M. Valko. On two ways to use determinantal point processes for Monte Carlo integration. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.
[14] **R. Bardenet**, P. Chainais, J. Flamant, and A. Hardy. A correspondence between zeros of time-frequency transforms and Gaussian analytic functions. In *Proceedings of the International conference on Sampling Theory and its Applications (SampTA)*, 2019.
[15] **R. Bardenet** and A. Hardy. From random matrices to Monte Carlo integration via Gaussian quadrature. In *Proceedings of the IEEE Statistical Signal Processing workshop (SSP)*, 2018.
[16] G. Gautier, **R. Bardenet**, and M. Valko. Zonotope hit-and-run for efficient sampling of projection dpps. In *International Conference on Machine Learning (ICML)*, 2017.
[17] **R. Bardenet** and M. K. Titsias. Inference for determinantal point processes without spectral knowledge. In *Advances in Neural Information Processing Systems (NIPS)*, pages 3375–3383, 2015.
[18] **R. Bardenet**, A. Doucet, and C. Holmes. Towards scaling up MCMC: an adaptive subsampling approach. In *Proceedings of the International Conference on Machine Learning (ICML)*, 2014.
[19] **R. Bardenet**, M. Brendel, B. Kégl, and M. Sebag. Collaborative hyperparameter tuning. In *International Conference on Machine Learning (ICML)*, 2013.
[20] **R. Bardenet**, O. Cappé, G. Fort, and B. Kégl. An adaptive Metropolis algorithm with online relabeling. In *Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 22, pages 91–99, April 2012.
[21] J. Bergstra, **R. Bardenet**, B. Kégl, and Y. Bengio. Algorithms for hyperparameter optimization. In *Advances in Neural Information Processing Systems (NIPS)*, volume 24. The MIT Press, 2011.
[22] **R. Bardenet** and B. Kégl. Surrogating the surrogate: accelerating Gaussian-process-based global optimization with a mixture cross-entropy algorithm. In *Proceedings of the 27th International Conference on Machine Learning (ICML)*, 2010.

CS refereed workshops

[23] G. Gautier, **R. Bardenet**, and M. Valko. On two ways to use determinantal point processes for Monte Carlo integration. Technical report, ICML workshop on Negative dependence in machine learning, 2019.
[24] **R. Bardenet**, M. Brendel, B. Kégl, and M. Sebag. SCoT: surrogate-based collaborative tuning for hyperparameter learning that remembers the past. NIPS workshop on Bayesian Optimization and Decision Making, 2012.
[25] **R. Bardenet**, B. Kégl, and G. Fort. Relabeling MCMC algorithms in Bayesian mixture learning. Snowbird Learning workshop, 2011.
[26] **R. Bardenet** and B. Kégl. Sampling-based optimization with mixtures. NIPS workshop on Optimization for Machine Learning, 2009.

CS book chapters

[27] **R. Bardenet**. *Proceedings of the 2012 IN2P3 School of Statistics*, chapter Monte Carlo methods. EDP Sciences, 2013.

Physics Journals (selected papers)

[28] P. Abreu et al. The lateral trigger probability function for UHE cosmic rays showers detected by the Pierre Auger Observatory. *Astroparticle Physics*, 35:266–276, 2011. Auger Collaboration paper.
[29] P. Abreu et al. Anisotropy and chemical composition of ultra-high energy cosmic rays using arrival directions measured by the Pierre Auger Observatory. *Journal of Cosmology and Astroparticle Physics*, page 022, 2011. Auger Collaboration paper.
[30] P. Abreu et al. Advanced functionality for radio analysis in the Offline software framework of the Pierre Auger Observatory. *Nuclear Instruments and Methods in Physics Research A*, 635:92–102, 2011. Auger Collaboration paper.
[31] P. Abreu et al. The Pierre Auger Observatory scaler mode for the study of solar activity modulation of galactic cosmic rays. *JINST*, 6:01003, 2011. Auger Collaboration paper.
[32] P. Abreu et al. Search for first harmonic modulation in the right ascension distribution of cosmic rays detected at the Pierre Auger Observatory. *Astroparticle Physics*, 34:627–639, 2011. Auger Collaboration paper.
[33] P. Abreu et al. The exposure of the hybrid detector of the Pierre Auger Observatory. *Astroparticle Physics*, 34:368–381, 2011. Auger Collaboration paper.
[34] P. Abreu et al. Update on the correlation of the highest energy cosmic rays with nearby extragalactic matter. *Astroparticle Physics*, 34:314–326, 2010. Auger Collaboration paper.

Physics refereed proceedings

[35] **R. Bardenet** and B. Kégl. An adaptive Monte Carlo Markov chain algorithm for inference from mixture signals. In *Proceedings of ACAT'11, Journal of Physics: Conference series*, 2012.

Physics technical reports

[36] B. Kégl, R. Busa-Fekete, K. Louedec, **R. Bardenet**, X. Garrido, I.C. Mariş, D. Monnier-Ragaigne, S. Dagoret-Campagne, and M. Urban. Reconstructing $N_{\mu 19}(1000)$. Technical reports 2011-054, Auger Project Technical Note, 2011.
[37] **R. Bardenet**, B. Kégl, and D. Veberic. Single muon response: The signal model. Technical Report 2010-110, Auger Project Technical Note, 2010.
[38] R. Dallier, **R. Bardenet**, S. Gambetta, H. Lyberis, and C. Macolino. Radio spectrum measurements at Auger, part 2. Technical Report 2010-032, Auger Project Technical Note, 2010.

Biology Journals

[39] B. J. Ridder, D. J. Leishman, M. Bridgland-Taylor, M. Samieegohar, X. Han, W. W. Wu, A. Randolph, P. Tran, J. Sheng, T. Danker, A. Lindqvist, D. Konrad, S. Hebeisen, L. Polonchuk, E. Gissinger, M. Renganathan, B. Koci, H. Wei, J. Fan, P. Levesque, J. Kwagh, J. Imredy, J. Zhai, M. Rogers, E. Humphries, R. Kirby, S. Stoelzle-Feix, N. Brinkwirth, M. Giustin, N. Becker, S. Friis, M. Rapedius, T. A. Goetze, T. Strassmaier, G. Okeyo, J. Kramer, Y. Kuryshev, C. Wu, H. Himmel, G. R. Mirams, D. G. Strauss, **R. Bardenet**, and Z. Li. A systematic strategy for estimating herg block potency and its implications in a new cardiac safety paradigm block potency and its implications in a new cardiac safety paradigm. *Toxicology and Applied Pharmacology*, 2020.
[40] K. A. Beattie, A. P. Hill, **R. Bardenet**, Y. Cui, J. I. Vandenberg, D. J. Gavaghan, T. P. de Boer, and G. R. Mirams. Sinusoidal voltage protocols for rapid characterization of ion channel kinetics. *Journal of Physiology*, 2018.
[41] R. H. Johnstone, **R. Bardenet**, D. J. Gavaghan, and G. R. Mirams. Hierarchical Bayesian inference for ion channel screening dose-response data. *Wellcome Open Research*, 2016.
[42] J. Kursawe, P. Brodskiy, C. Narcisso, J. J. Zartmann, R. E. Baker, **R. Bardenet**, and A. G. Fletcher. Robust cell tracking in epithelial tissues through identification of maximum common subgraphs. *Journal of the Royal Statistical Society Interface*, 2016.
[43] R. H. Johnstone, E. T. Y. Chang, **R. Bardenet**, T. P. De Boer, D. J. Gavaghan, P. Pathmanathan, R. H. Clayton, and G. R. Mirams. Uncertainty and variability in models of the cardiac action potential: Can we build trustworthy models? *Journal of molecular and cellular cardiology*, 96:49–62, 2016.