Gabriel Arpino

CONTACT
INFORMATION

ga442@cam.ac.uk   Cambridge, UK
https://gabrielarpino.github.io

EDUCATION

University of Cambridge, UK
Ph.D. in Engineering, 2021 — Present.
- Supervisor: Prof. Ramji Venkataramanan.
- Advisor: Prof. Ioannis Kontoyiannis.
- Subject: Computational challenges and methods for high-dimensional heterogeneous regression.

ETH Zürich, Switzerland
M.Sc. in Electrical Engineering and Information Technology, 2019 — 2021.
- Depth Area: Applied Probability, Statistics, Optimization.
- Supervisor: Prof. Afonso Bandeira.
- M.Sc. Thesis: *"Computational Hardness of Sparse High-Dimensional Linear Regression"*.

University of Toronto, Canada
B.A.Sc. in Engineering Science (with Honours), 2014 — 2019.
- B.A.Sc. Thesis: *"Tightening PAC-Bayes Bounds using Data-Dependent Priors"*.
- Supervisor: Prof. Daniel Roy.

RESEARCH
INTERESTS

I am interested in problems at the interface of machine learning, statistics, applied probability, and optimization. I seek to understand and to develop efficient methods for tackling inference problems that arise in practice, often from the perspectives of information theory and statistical physics.

PUBLICATIONS

Gabriel Arpino, Xiaoqi Liu, Julia Gontarek, Ramji Venkataramanan. "Inferring Change Points in High-Dimensional Regression via Approximate Message Passing". *Under Review.*

Gabriel Arpino, Xiaoqi Liu, Ramji Venkataramanan. "Inferring Change Points in High-Dimensional Linear Regression via Approximate Message Passing". *International Conference on Machine Learning (ICML), 2024.*

Gabriel Arpino, Daniil Dmitriev, Nicolo Grometto. "Greedy Heuristics and Linear Relaxations for the Random Hitting Set Problem". *International Conference on Approximation Algorithms for Combinatorial Optimization Problems (APPROX), 2024.*

Gabriel Arpino, Ramji Venkataramanan. "Statistical-Computational Tradeoffs in Mixed Sparse Linear Regression". *Conference on Learning Theory (COLT), 2023.*

[1] ***Gabriel Arpino**, *Nicolo Grometto, Afonso Bandeira. "Group Testing in the High Dilution Regime". *International Symposium on Information Theory (ISIT), 2021.*

Gintare Karolina Dziugaite, Kyle Hsu, Waseem Gharbieh, **Gabriel Arpino**, Daniel M. Roy. "On the role of data in PAC-Bayes bounds". *Artificial Intelligence and Statistics Conference (AISTATS), 2021.*

Gintare Karolina Dziugaite, **Gabriel Arpino**, and Daniel Roy. "Towards generalization guarantees for SGD: Data-dependent PAC-Bayes priors". *Neural Information Processing Systems (NeurIPS) Workshop on Bayesian Deep Learning, 2018.*

---

[1] * denotes equal contribution.

Gabriel Arpino, Kyle Morris, Sasanka Nagavalli, Katia Sycara. "Using Information Invariants to Compare Swarm Algorithms and General Multi-Robot Algorithms". *IEEE International Conference on Robotics and Automation (ICRA), 2018.*

Kyle Morris, **Gabriel Arpino**, Sasanka Nagavalli, Katia Sycara. "Full Stack Swarm Architecture". *RISS Working Papers Journal, 2017.*

Johnathon N. Caguiat, **Gabriel Arpino**, Sally G. Krigstin, Donald W. Kirk, Charles Q. Jia. "Dependence of supercapacitor performance on macro-structure of monolithic biochar electrodes". *Biomass and Bioenergy, 2018.*

AWARDS

- G-Research Ph.D. Runner-Up Prize 2024 (£4,000).
- Cambridge Trust Fellowship to study at Gonville & Caius College, University of Cambridge (£45,000/year).
- Scholarship to join the Intelligent Co-ordination and Logistics Lab as part of the Robotics Institute Summer Scholars (RISS) program ($3,000).
- University of Toronto 2017 Scholar Award for academic performance ($2,000).

EXPERIENCE

Google
Student Researcher, November 2024 – February 2025.
- Developing Deep Learning models for vessel path reconstruction as part of the Global Networking (GN) team.

Invenia Labs
Junior Researcher, September 2017 — September 2018.
- Led the development of scalable and distributed Bayesian Deep Learning and Gaussian Process routines to model the US electricity grid, implemented in Tensorflow and as an open-source package in Julia: `https://github.com/invenia/GPForecasting.jl`.
- Developed AWS cloud and GPU infrastructure to run models on over 50 gigabytes of data, improving accuracy by 55% over previous models.
- Composed statistical kernels for gaussian process forecasting in the electricity market, resulting in forecast accuracies beating the state of the art.

Robotics Institue, Carnegie Mellon University
RISS Researcher, May 2017 — September 2017.
- Led a team of researchers through the development of a paper on information invariants in multi-agent robotic systems, accepted as a conference paper to ICRA 2018, supervised by Prof. Katia Sycara
- Developed a novel full-stack swarm robot control architecture implemented and tested on ROS, leading to the publication of a paper at the RISS 2017 Working Papers Journal and achieving over 95% reproduction accuracy on real-world simulations.

CEAR Lab, Technion - Israel Institute of Technology
Researcher, May 2016 — September 2016.
- Implemented 3D SLAM localization algorithms in C++ onto Clearpath Field Robots, resulting in the creation of accurate point cloud representations of $100m^2$ orchards, supervised by Prof. Amir Degani.
- Refined robotic vision algorithms in C++ and Python using ROS and PCL for orchard tree identification, resulting in robust code that identified 90% of orchard tree clusters.

Green Technology Laboratory, University of Toronto
Researcher, March 2015 — September 2015.
- Developed procedures for the performance testing of biochar supercapacitors, leading to a Biomass and Bioenergy Journal publication supervised by Prof. Charles Q. Jia.

INVITED TALKS	• *Statistical-Computational Tradeoffs in Mixed Sparse Linear Regression*, University of Cambridge Information Theory seminar, November 2023.
	• *Statistical-Computational Tradeoffs in Mixed Sparse Linear Regression*, ETH Zürich DACO seminar, October 2023.
	• *Dilution Group Testing: Novel Bounds via Practical Decoders*, Professor Helmut Bölcskei's lab seminar, ETH Zürich. June 2020.
	• *Noisy Group Testing: Achievable Rates*, Professor Afonso Bandeira's lab seminar, ETH Zürich. April 2020.
WORKSHOPS AND POSTERS	• *Participant.* 11th World Congress in Probability and Statistics, 2024.
	• *Participant.* Youth in High-Dimensions: Recent Progress in Machine Learning, High-Dimensional Statistics and Inference, 2023.
	• *Participant.* Les Houches Summer school on Statistical Physics & Machine learning, 2022.
GRADUATE COURSES	Mathematical Statistics, Probability Theory, Information Theory, Functional Analysis, Advanced Machine Learning, Computational Complexity, Measure Theory, Control Theory, Empirical Processes, Mathematical Optimization, Neural Network Theory.
TEACHING	• *Teaching Assistant.* University of Cambridge Engineering 3F7: Information Theory, Michaelmas 2022 & 2023.
	• *Teaching Assistant.* University of Cambridge Engineering 3F4: Data Transmission, Lent 2023.
	• *Teaching Assistant.* University of Cambridge Engineering 3F8: Inference, Lent 2022.
ACADEMIC SERVICE	*Reviewer.* International Conference on Algorithmic Learning Theory (ALT). Conference on Neural Information Processing Systems (NeurIPS). International Conference on Machine Learning (ICML). IEEE Transactions on Information Theory. IEEE Transactions on Signal Processing.
LANGUAGES	• Portuguese, English, Italian, French, Spanish.
	• C++, Python (Numpy, JAX, Tensorflow, Pytorch), Julia, MATLAB.
EXTRA CURRICULAR	Jazz musician and leader, having performed and led bands at the professional level on upright and electric bass in venues such as *The Rex* in Toronto, *Hot Numbers* in Cambridge, UK, *Next Door Records* in London, UK, *Moods* and *Lebewohlfabrik* in Zürich.