Valentin De Bortoli

Senior Research Scientist · Google DeepMind

London, UK valentin.debortoli@gmail.com +33 6 66 36 23 17 vdeborto.github.io Google Scholar

EDUCATION

École Normale Supérieure Saclay, PhD 2017 – 2020
- **Title**: Non-local statistics in images: modelisation, estimation and sampling
- **Advisors**: Agnès Desolneux, Bruno Galerne, Arthur Leclaire.

École Normale Supérieure Saclay, Msc MVA (Master Vision Apprentissage) 2016 – 2017
- The most competitive machine learning and computer vision masters in France
- **Coursework**: Optimization, Graphical Models, Computer Vision, Wavelet Theory, PDE Analysis.

École Normale Supérieure Saclay, Agrégation externe 2015 – 2016
- Ranked 19th out of 300.
- Agrégation externe is the most competitive exam for the civil service sector in the French education system.
- **Coursework**: Advanced Linear Algebra, Group Theory, Fourier Analysis, Probability Theory, Differential Calculus.

École Normale Supérieure Saclay, Bsc – Msc, pure mathematics 2013 – 2015
- **Coursework**: Group theory, Fourier Analysis, Functional Calculus, PDE analysis, Differential Geometry, Advanced Linear Algebra, Probability

Classe préparatoire aux Lazaristes, preparation for competitive exams, *Lyon* 2011 – 2013

EXPERIENCE

Senior Research Scientist, Google DeepMind, *London, UK* 2025 – Present
- Collaborating with GenMedia Unit within Google DeepMind for next frontier generative models (James Thornton).
- Collaboration with Protein Design teams in the Science Unit (Kevin Li, Alex Chu, Jiaxin Shi).
- Technical lead on *Hackable Diffusion* (github.com/google/hackable_diffusion), an OSS diffusion framework with 150+ stars, used for finetuning of Gemma Diffusion (github.com/google-deepmind/gemma).

Research Scientist, Google DeepMind, *London, UK* 2023 – 2025
- Working on next generation generative modeling in the Foundational Unit, collaborating with Arnaud Doucet.
- Collaboration with robotics teams for dexterous manipulation (Stannis Zhou, Kevin Murphy).
- Collaboration with Isomorphic Labs on sampling problems for protein modeling (Peter Wirnsberger, Michael Hutchinson).

Associate Professor, École Normale Supérieure, *Paris, France* 2022 – Present
- CNRS (Centre National de la Recherche Scientifique) chargé de recherche.
- Member of the DATA team (leads: Giulio Biroli, Gabriel Peyré, Stéphane Mallat).
- On leave until 2026.

Postdoctoral researcher, Oxford University, *Oxford, UK* 2020 – 2022
- Member of the COSINE project under the supervision of Arnaud Doucet.

Teaching Assistant, École Normale Supérieure, *Paris, France* 2017 – 2020
- Differential calculus (undergraduate level), reference: Frédéric Pascal.
- Optimization (graduate level), reference: Alain Trouvé.
- Hilbertian analysis and Fourier analysis (agrégation), reference: Frédéric Pascal.

Research Internship, San Diego State University, *California, US* 2015

- Review of wavelet-based unsupervised texture segmentation, advantage of adaptive wavelets under the supervision of Jérôme Gilles.

SELECTED PUBLICATIONS

My main interests revolve around generative modeling and its value in scientific applications. I have been promoting the use of diffusion models since 2020. I pioneered many of the methodological advances of diffusion models such as Riemannian diffusion models, discrete diffusion models and the links between those models and Optimal Transport. A detailed list of my publications can be found on my Google Scholar. I have published in most major machine learning conferences, for example UAI, ICML, ICLR, NeurIPS and machine learning journal venues such as JMLR and TMLR. Below are three publications that I've selected that best reflect my work.

Riemannian Score-Based Generative Modeling 2022
Valentin De Bortoli, Émile Mathieu, Michael Hutchinson, James Thornton, Yee Whye Teh, Arnaud Doucet
NeurIPS · BEST PAPER AWARD

Convergence Of Denoising Diffusion Models Under The Manifold Hypothesis 2022
Valentin De Bortoli
TMLR

Diffusion Schrödinger Bridge With Applications To Score-Based Generative Modeling 2021
Valentin De Bortoli, James Thornton, Jeremy Heng, Arnaud Doucet
NeurIPS · SPOTLIGHT PAPER

I have also been involved in the development of generative modeling techniques for protein modeling. Notably, through a collaboration with MIT and Baker Lab · UW Institute for Protein Design, I conceived some of the methods used in RF Diffusion. Below are two publications reflecting this work.

De novo design of protein structure and function with RFdiffusion 2023
Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eisenach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, Basile IM Wicky, Nikita Hanikel, Samuel J Pellock, Alexis Courbet, William Sheffler, Jue Wang, Preetham Venkatesh, Isaac Sappington, Susana Vázquez Torres, Anna Lauko, **Valentin De Bortoli**, Emile Mathieu, Sergey Ovchinnikov, Regina Barzilay, Tommi S Jaakkola, Frank DiMaio, Minkyung Baek, David Baker
Nature

SE(3) Diffusion Model with Application to Protein Backbone Generation 2023
Jason Yim, Brian L Trippe, **Valentin De Bortoli**, Emile Mathieu, Arnaud Doucet, Regina Barzilay, Tommi Jaakkola
ICML

Further to my work on protein modeling, I have special interest in the theoretical foundations of generative models and their links with Markov processes and Optimal Transport. In particular, I am interested in new numerical solvers for Optimal Transport tasks leveraging the effectiveness of diffusion models. Below are three publications reflecting this work.

Dynamical regimes of diffusion models 2024
Giulio Biroli, Tony Bonnaire, **Valentin De Bortoli**, Marc Mézard
Nature Communications

From denoising diffusions to denoising markov models 2022
Joe Benton, Yuyang Shi, **Valentin De Bortoli**, George Deligiannidis, Arnaud Doucet
Journal of the Royal Statistical Society Series B: Statistical Methodology · READ PAPER

Quantitative uniform stability of the iterative proportional fitting procedure 2021
George Deligiannidis, **Valentin De Bortoli**, Arnaud Doucet
Annals of Applied Probability

ORGANIZED SEMINARS AND WORKSHOPS

I have organized a number of seminars and workshops exploring the themes of generative modeling, sampling and stochastic control.

2022 **SIAM Imaging** minisymposium on New Directions in Stochastic Bayesian Imaging Methodology
2022 **NeurIPS** workshop on score-based models
2022 **MIA** Plug and Play workshop
2022 **CNRS** – Machine learning assisted scientific computing and sampling
2023 **NeurIPS** workshop on diffusion models
2023 **ICML** workshop on New Frontiers in Learning, Control, and Dynamical Systems
2024 **NeurIPS** workshop on Scientific Methods for Understanding Deep Learning
2025 **NeurIPS** workshop on Principles of Generative Modeling (PriGM)
2026 **ICLR** workshop on Generative models in Genomics
2026 **ICLR** workshop on Deep Generative Model in Machine Learning: Theory, Principle and Efficacy
2026 **ICML** workshop on Foundations of Deep Generative Models: Understanding Memorization, Generalization, and Reasoning
2026 **NeurIPS** workshop on Principles of Generative Modeling (PriGM)
2026 **NeurIPS** – AI for Stochastic Dynamics: From Theoretical Foundations to Scientific Applications

MENTORING AND RESPONSIBILITY

I have supervised a number of students both in industry and academia.

2022 – 2026 **Maxence Noble-Bourillot**, PhD (École Polytechnique) – co-supervised with Alain Durmus
2022 **Roland Andrews**, internship (ENS Paris) – co-supervised with James Thornton
2024 **Andrew Campbell**, internship (Google DeepMind) – co-supervised with Arnaud Doucet
2024 **Hippolyte Pilchen**, internship (Google DeepMind) – co-supervised with Arthur Gretton
2025 **Tommaso Martorella**, internship
2025 **Christopher Marouani**, internship
2026 **Justin Deschenaux**, internship

Conference committees: NeurIPS Area Chair · ICLR Senior Area Chair · ICML Area Chair

TECHNOLOGIES

Languages: Python (proficient), Matlab (proficient), C (basics)

Technologies: deep learning and autodifferentiation frameworks such as Flax, PyTorch and TensorFlow. Familiarity with large scale distributed training. Basics on kernel development (Pallas kernel).