

Blaise Delattre

Science Tokyo (Tokyo Institute of Science)

✉ bldelattre@gmail.com

Post-doctoral Researcher in Trustworthy AI 🌐 <https://www.lamsade.dauphine.fr/~bdelattre>

Research Interests

Certified robustness and trustworthy AI; Lipschitz-constrained networks; randomized smoothing; stability and efficiency in deep learning; robust foundation models.

Education

- 2025–Present **Post-doctoral Researcher**, *Science Tokyo*, Tokyo, Japan
Supervised by Prof. Yang Cao. Research on provable robustness and trustworthiness of foundation models.
- 2022 – 2025 **PhD in Computer Science**, *Université Paris-Dauphine PSL*, Paris, France
Supervised by Prof. Alexandre Allauzen and Dr. Quentin Barthélemy. Thesis on Lipschitz neural networks and adversarial certified robustness.
- 2020 – 2021 **Master 2 MASH (Mathematics, Learning and Data Science)**, *Université Paris-Dauphine & ENS Ulm PSL*, Paris, France
Joint PSL degree in Data Science (Dauphine–ENS Ulm). Coursework in machine learning, optimization, statistics, reinforcement learning, NLP, and computer vision.
- 2016 – 2019 **MSc in Financial Engineering & Quantitative Finance**, *Ensimag Grenoble INP & IAE Grenoble*, Grenoble, France

Publications

(* denotes equal contribution)

1. **B. Delattre**, P. Caillon, E. Fagnou, Q. Barthélemy, A. Allauzen. *Bridging the Theoretical Gap in Randomized Smoothing*. AISTATS, 2025.
2. E. Fagnou, P. Caillon, **B. Delattre**, A. Allauzen. *Accelerated Training through Iterative Gradient Propagation Along the Residual Path*. ICLR, 2025 (**oral**).
3. E. Fagnou, P. Caillon, **B. Delattre**, A. Allauzen. *Chain and Causal Attention for Efficient Entity Tracking*. ACL, 2024.
4. **B. Delattre**, A. Araujo, Q. Barthélemy, A. Allauzen. *The Lipschitz–Variance–Margin Tradeoff for Enhanced Randomized Smoothing*. ICLR, 2023.
5. **B. Delattre**, Q. Barthélemy, A. Araujo, A. Allauzen. *Efficient Bound of Lipschitz Constant for Convolutional Layers by Gram Iteration*. ICML, 2023.
6. A. Araujo, A. Havens, **B. Delattre**, A. Allauzen, B. Hu. *A Unified Algebraic Perspective on Lipschitz Neural Networks*. ICLR, 2023 (**spotlight**).
7. L. Meunier*, **B. Delattre***, A. Araujo, A. Allauzen. *A Dynamical System Perspective for Lipschitz Neural Networks*. ICML, 2022 (**oral**).

Research Experience

- 2025–Present **Post-doctoral Researcher**, *Science Tokyo*, Tokyo, Japan
Work on provable robustness and trustworthy foundation models.
- 2022 – 2025 **Research Scientist (CIFRE)**, *Foxstream*, Lyon, France
Research on robust deep learning.

Talks

- 2025 **Tutorial (accepted)**, *Tutorial on Robust Certificates for Neural Networks*, ADC 2025, Bali, Indonesia
- 2025 **Invited Talk**, *Toward Certified Robustness in Foundation Models*, Kobe International Workshop on Trustworthy Foundation Models (KTFM), Kobe, Japan
Invited talk at KTFM 2025.
- 2025 **Invited Talk**, *Lipschitz-Constrained Networks for Certified Robustness*, JSPS–NSFC Joint Research Workshop (Online)
Half-day workshop on Lipschitz-constrained networks and certified robustness.
- 2024 **Reading Group Presentation**, *Spectral Norm Estimation in Deep Learning*, ENS Lyon, France
Invited presentation at the ENS Lyon Reading Group.
- 2023 **Reading Group Presentation**, *The Lipschitz–Variance–Margin Tradeoff for Enhanced Randomized Smoothing*, University of Illinois Urbana–Champaign (UIUC) (Online)
Invited presentation at the UIUC Adversarial Attacks Reading Group.
- 2022 – 2023 **Conference & Workshop Presentations**, *A Dynamical System Perspective for Lipschitz Neural Networks*, ICML 2022 (Baltimore, USA); SCAI 2023 (Paris, France)
Oral presentation at ICML 2022 and at the Sorbonne Center for AI workshop “Mathematical Foundations of AI” (2023).

Academic Service

Reviewer: NeurIPS (2025), ICLR (2025–2026), ICML (2024–2025), AISTATS (2026), AAAI (2026), IEEE TACON, L4DC (2024), ACML Journal Track (2024)

Poster & Encore Track Chair: ADMA 2025, Kyoto, Japan.

Teaching Experience

- 2023–2025 Trustworthy Machine Learning in Practice (Lecturer), Executive Master, Université Paris-Dauphine PSL.

Skills

- Programming Python, PyTorch
- Languages French (native), English (fluent)