Enric Boix-Adserà

E-mail: eboix@wharton.upenn.edu Website: https://eboix.github.io/

Education and academic positions

Wharton School at the University of Pennsylvania Summer 2025-
Assistant Professor in Statistics and Data Science

Massachusetts Institute of Technology and Harvard Spring 2025
Postdoctoral Associate in MIT Math and Harvard CMSA
Hosted by: Philippe Rigollet and Michael Douglas

UC Berkeley Fall 2024
Research Fellow at the Simons Institute for the Theory of Computing
Program: Large Language Models and Transformers

Massachusetts Institute of Technology 2018-2024
Ph.D. in Electrical Engineering and Computer Science (GPA: 5.0/5.0)
Advisors: Guy Bresler and Philippe Rigollet

Princeton University 2014-2018
A.B. in Mathematics (GPA: 3.96/4.0)
Certificates: Applied and Computational Mathematics; Applications of Computing
Advisor: Emmanuel Abbe

Research summary and selected publications

My research focuses on building a **mathematical science of deep learning**. My aim is to characterize the fundamental mechanisms driving what and how neural networks learn, to enable us to develop more efficient, more interpretable, and safer systems.

Selected publications:

- The power of fine-grained experts: Granularity boosts expressivity in Mixture of Experts
  (alphabetical) **EB**, Philippe Rigollet
  Preprint, 2025
- Towards a theory of model distillation
  EB
  Preprint, 2024
- When can transformers reason with abstract symbols?
  EB*, Omid Saremi, Emmanuel Abbe, Samy Bengio, Etai Littwin, Joshua Susskind
  International Conference on Learning Representations (ICLR), 2024

- SGD learning on neural networks: leap complexity and saddle-to-saddle dynamics

  (alphabetical) Emmanuel Abbe, **EB**, Theodor Misiakiewicz

  Conference on Learning Theory (COLT), 2023

Industry Experience

Part-time at Apple Machine Learning Research — Fall 2022-Summer 2023

Intern at Apple Machine Learning Research — Summer 2022

- Authored three publications on deep learning in collaboration with industry
- Conducted theoretical analysis and ran experiments on Apple's GPU cluster

Software Engineering Intern at Facebook — Summer 2016

- Wrote software to test Facebook's network switches

Teaching and Mentorship Experience

Research Mentor — 2023-present

- Mentored graduate students on research projects in deep learning

Teaching Assistant for *Algorithms for Inference* (6.7810) at MIT — Fall 2022

- Advanced undergraduate / early graduate course with 70 students
- Taught recitations, prepared problem sets, ran office hours, answered Piazza questions, helped write and proctor exams

Mentor for MIT Graduate Application Assistance Program (GAAP) — 2021-2023

- Helped students from underrepresented groups with the grad school application process
- Answered questions and gave iterative feedback on application materials

Undergraduate TA for *Algorithms and Data Structures* (COS 226) at Princeton — Spring 2015

- Early undergraduate course
- Graded and provided written feedback on programming assignments

Awards

Apple AI/ML Scholar	2021
Ernst A. Guillemin AI and Decision Making Master's Thesis Award, 2nd place	2021
Siebel Scholar	2019
NSF Graduate Research Fellowship	2019-present
Phi Beta Kappa	2018
Middleton Miller '29 award for best senior thesis in the math department	2018
Princeton Applied and Computational Math best independent work project	2018
NSF Center for Science of Information Channels Scholar	2017-2018
Princeton University's Shapiro Prize for Academic Excellence	2015
National Merit Scholar	2014
International Olympiad in Informatics: Bronze medal	2014
Spanish Olympiad in Computer Science: 1st place	2014

Additional academic experience

Participant in Simons Institute programs: Fall 2021
Computational Complexity of Statistical Inference, and
Geometric Methods in Optimization and Sampling
Participant in Simons Institute program: Fall 2020
Probability, Geometry, and Computation in High Dimensions (online)
Research in algebraic complexity, advised by Zeev Dvir 2017
Research in algebraic complexity, advised by Rafael Mendes de Oliveira Summer 2015

SERVICE

Reviewer for: Conference on Learning Theory **(COLT)**, Foundations of Computer Science **(FOCS)**, International Conference on Learning Representations **(ICLR)**, International Conference on Machine Learning **(ICML)**, IEEE Transactions on Automatic Control **(IEEE-TAC)**, IEEE Transactions on Control of Network Systems **(IEEE-TCNS)**, Informs Journal on Optimization **(IJOO)**, IEEE International Symposium on Information Theory **(ISIT)**, Innovations in Theoretical Computer Science **(ITCS)**, Conference on Neural Information Processing Systems **(NeurIPS)**, Numerical Algorithms **(NUMA)**, Mathematics of Operations Research **(MOR)**, International Conference on Randomization and Computation **(RANDOM)**, ACM-SIAM Symposium on Discrete Algorithms **(SODA)**, ACM Symposium on Theory of Computer Science **(STOC)**, Theoretical Computer Science **(TCS)**, Transactions on Machine Learning Research **(TMLR)**.

Coffee Hour Chair for MIT EECS Graduate Student Association 2019
- Organized weekly coffee hours for department, attended by about 85 people

Academic Officer for Princeton Mathematics Club 2017-2018
- Organized mathematics club lecture series directed towards undergraduate students

Volunteer at Julia Robinson Mathematics Festival 2017
- Taught problem-solving skills to elementary- and middle-school students

Volunteer at Princeton University Science Olympiad Invitational Tournament 2017
- Proctored exams and helped with logistics for high school student competition

Volunteer at Princeton University Physics Competition 2015
- Graded exams for high school student competition

PAPERS

27. Granularity boosts expressivity in Mixture of Experts
(alphabetical) **EB**, Philippe Rigollet
Preprint, 2025

26. FACT: a first-principles alternative to the Neural Feature Ansatz for how networks learn representations
EB*, Neil Mallinar*, James B. Simon, Mikhail Belkin.
Preprint, 2025

25. Let Me Think! A long chain of thought can be worth exponentially many short ones
Parsa Mirtaheri*, Ezra Edelman*, Samy Jelassi, Eran Malach, **EB**.
Conference on Neural Information Processing Systems (NeurIPS), 2025

24. Aggregate and conquer: detecting and steering LLM concepts by combining nonlinear predictors over multiple layers

    Daniel Beaglehole*, Adityanarayanan Radhakrishnan*, **EB**, Mikhail Belkin

    Preprint, 2025

23. On the inductive bias of infinite-depth ResNets and the bottleneck rank

    EB

    Preprint, 2025

22. Towards a theory of model distillation

    EB

    Preprint, 2024

21. Prompts have evil twins

    Rimon Melamed, Lucas H. McCabe, Tanay Wakhare, Yejin Kim, H. Howie Huang, **EB**

    Empirical Methods in Natural Language Processing, 2024

20. When can transformers reason with abstract symbols?

    EB*, Omid Saremi, Emmanuel Abbe, Samy Bengio, Etai Littwin, Joshua Susskind

    International Conference on Learning Representations (ICLR), 2024

19. Transformers learn through gradual rank increase

    EB*, Etai Littwin*, Emmanuel Abbe, Samy Bengio, Joshua Susskind

    Conference on Neural Information Processing Systems (NeurIPS), 2023

18. Tight conditions for when the NTK approximation is valid

    (alphabetical) **EB**, Etai Littwin

    Transactions on Machine Learning Research (TMLR), 2023

17. SGD learning on neural networks: leap complexity and saddle-to-saddle dynamics

    (alphabetical) Emmanuel Abbe, **EB**, Theodor Misiakiewicz

    Conference on Learning Theory (COLT), 2023

16. GULP: a prediction-based metric between representations

    EB, Hannah Lawrence, George Stepaniants, Philippe Rigollet

    Conference on Neural Information Processing Systems (NeurIPS), 2022

    Oral presentation (top 8%)

15. On the non-universality of deep learning: quantifying the cost of symmetry

    (alphabetical) Emmanuel Abbe, **EB**

    Conference on Neural Information Processing Systems (NeurIPS), 2022

14. The merged-staircase property: a necessary and nearly sufficient condition for SGD learning of sparse functions on two-layer neural networks

    (alphabetical) Emmanuel Abbe, **EB**, Theodor Misiakiewicz

    Conference on Learning Theory (COLT), 2022

13. The staircase property: How hierarchical structure can guide deep learning

    (alphabetical) Emmanuel Abbe, **EB**, Matthew Brennan, Guy Bresler, Dheeraj Nagaraj

    Conference on Neural Information Processing Systems (NeurIPS), 2021

12. Chow-Liu++: Optimal Prediction-Centric Learning of Tree Ising Models

    (alphabetical) **EB**, Guy Bresler, and Frederic Koehler

    Foundations of Computer Science (FOCS), 2021

11. Wasserstein barycenters are NP-hard

    (alphabetical) Jason Altschuler and **EB**

    SIAM Journal on Mathematics of Data Science (SIMODS), 2021

10. Hardness results for Multimarginal Optimal Transport Problems

    (alphabetical) Jason Altschuler and **EB**

    Discrete Optimization (DISOPT), 2020

9. Polynomial-time algorithms for Multimarginal Optimal Transport problems with decomposable structure

    (alphabetical) Jason Altschuler and **EB**

    Mathematical Programming, 2020

8. Wasserstein barycenters can be computed in polynomial time in fixed dimension

    (alphabetical) Jason Altschuler and **EB**

    Journal of Machine Learning Research (JMLR), 2020

7. The Multiplayer Colonel Blotto Game

    (alphabetical) **EB**, Ben Edelman and Siddhartha Jayanti

    Economics and Computation (EC), 2020

6. Sample-Efficient Active Learning of Causal Trees

    Kristjan Greenewald, Dmitriy Katz-Rogozhnikov, Karthikeyan Shanmugam,

    Sara Magliacane, Murat Kocaoglu, **EB** and Guy Bresler

    Conference on Neural Information Processing Systems (NeurIPS), 2019

5. The Average-Case Complexity of Counting Cliques in Erdos-Renyi Hypergraphs

    (alphabetical) **EB**, Matthew Brennan and Guy Bresler

    Foundations of Computer Science (FOCS), 2019

    Invited to SICOMP special edition

4. Subadditivity Beyond Trees and the Chi-Squared Mutual Information

   (alphabetical) Emmanuel Abbe and **EB**

   IEEE International Symposium on Information Theory (ISIT), 2019

3. Randomized Concurrent Set Union and Generalized Wake-Up

   Siddhartha Jayanti*, Robert E. Tarjan*, and **EB**

   Symposium on Principles of Distributed Computing (PODC), 2019

2. Graph powering and spectral robustness

   (alphabetical) Emmanuel Abbe, **EB**, Peter Ralli, and Colin Sandon

   SIAM Journal on Mathematics of Data Science (SIMODS), 2018

1. An Information-Percolation Bound for Spin Synchronization on General Graphs

   (alphabetical) Emmanuel Abbe and **EB**

   Annals of Applied Probability (AAP), 2018

SELECTED POSTERS

5. Mathematics of Deep Learning (MoDL) at Simons Institute in NYC, *When can transformers reason with abstract symbols?*, September 28-29, 2023.

4. Optimization and Statistical Learning Workshop at Les Houches, *GULP: a prediction-based metric between representations*, January 18, 2023.

3. Conference on Neural Information Processing Systems (NeurIPS), *On the non-universality of deep learning: quantifying the cost of symmetry*, December 1, 2022.

2. Mathematics of Deep Learning (MoDL) at Simons Institute in NYC, *On the non-universality of deep learning: quantifying the cost of symmetry*, September 29, 2022.

1. Conference on Neural Information Processing Systems (NeurIPS), *The staircase property: How hierarchical structure can guide deep learning*, December 2021.

SELECTED TALKS

38. UPenn Econometrics Lunch, *Kernels, neural networks, and feature learning*, November 3, 2025.

37. Aspen Conference on Foundation Models, *What is your neural network actually doing?*, October 24, 2025.

36. Lonergan Workshop at Boston College, *Reasoning in AI models: moments of insight?*, June 16, 2025.

35. USC Symposium on the Future of Computing, *Extracting world models with distillation*, March 14, 2025.

34. Les Houches School of Physics, *Distilling world models by using the Linear Representation Hypothesis*, January 17, 2025.

33. Boston University ECE department, *Beyond the black box: Characterizing and improving how neural networks learn*, March 4, 2024.

32. Stanford University Statistics Department, *Beyond the black box: Characterizing and improving how neural networks learn*, February 20, 2024.

31. University of Pennsylvania Computer and Information Science and Electrical and Systems Engineering Departments, *Beyond the black box: Characterizing and improving how neural networks learn*, February 13, 2024.

30. Columbia University Statistics Department, *Beyond the black box: Characterizing and improving how neural networks learn*, February 7, 2024.

29. Wharton Statistics and Data Science Department at the University of Pennsylvania, *Beyond the black box: Characterizing and improving how neural networks learn*, February 5, 2024.

28. University of California San Diego HDSI, *Beyond the black box: Characterizing and improving how neural networks learn*, January 24, 2024.

27. University of Toronto Mathematics Department, *Beyond the black box: Characterizing and improving how neural networks learn*, January 18, 2024.

26. Kempner Research Fellow Talk at Harvard, *Characterizations of what and how neural networks learn*, December 8, 2023.

25. Tomaso Poggio lab meeting at MIT, *The staircase property and low-rank learning*, November 20, 2023.

24. Young Researchers Workshop at Cornell, *When can transformers reason with abstract symbols?*, October 2, 2023.

23. Mathematics of Information Processing Seminar at RWTH Aachen (online), *Transformers learn through gradual rank increase*, July 17, 2023.

22. Conference on Learning Theory (COLT), *SGD learning on neural networks: leap complexity and saddle-to-saddle dynamics*, July 14, 2023.

21. Mathieu Wyart lab meeting at EPFL, *The staircase property and the leap complexity*, June 20, 2023.

20. Youth in High Dimensions Conference at the International Center for Theoretical Physics in Trieste, *The staircase property and the leap complexity*, June 1, 2023.

19. TTIC Young Researcher Series, *The staircase property and the leap complexity*, January 25, 2023.

18. Murat Erdogdu group meeting at University of Toronto (online), *The merged staircase property*, November 14, 2022.

17. Cornell Statistics and Data Science Seminar, *The merged-staircase property*, November 9, 2022.

16. Joan Bruna group meeting at NYU, *The merged-staircase property*, October 3, 2022.

15. Conference on Learning Theory (COLT), *The merged-staircase property: a necessary and nearly sufficient condition for SGD learning of sparse functions on two-layer neural networks*, July 2022.

14. MIT Theory of ML reading group, *Learning staircases with neural networks*, April 8, 2022.

13. Tufts Computational and Applied Math Seminar, *Polynomial-Time Algorithms for Multimarginal Optimal Transport Problems with Structure*, April 4, 2022.

12. Sierra seminar at INRIA Paris, *A necessary and nearly sufficient condition for SGD learning of sparse functions on two-layer neural networks*, March 9, 2022.

11. Simons Institute Deepfoundations Workshop at UC Berkeley, *The staircase property*, December 6, 2021.

10. Tomaso Poggio lab meeting at MIT, *The staircase property*, November 29, 2021.

9. CCSI + GMOS student seminar at Simons Institute at UC Berkeley, *The staircase property: How hierarchical structure can guide deep learning*, September 24, 2021.

8. UToronto Analysis and Applied Math Seminar (online), *The computational complexity of Wasserstein barycenters*, April 23, 2021.

7. MIT Theory Group student lunch (online), *Settling the Computational Complexity of Wasserstein Barycenters*, March 10, 2021.

6. Economics and Computation (EC), *The Multiplayer Colonel Blotto Game*, July 2020.

5. Youth in High Dimensions Conference at the International Center for Theoretical Physics in Trieste (online), *The Average-Case Complexity of Counting Cliques in Erdos-Renyi Hypergraphs*, June 29, 2020.

4. Foundations of Computer Science (FOCS), *The Average-Case Complexity of Counting Cliques in Erdos-Renyi Hypergraphs*, 2019.

3. Constantinos Daskalakis group meeting at MIT, *Directions in average-case complexity*, October 31, 2019.

2. IEEE International Symposium on Information Theory (ISIT), *Subadditivity Beyond Trees and the Chi-Squared Mutual Information*, 2019.

1. LIDS Student Seminar conference, *Clustering Networks with Graph Powering*, January 31, 2019.