

# Levent Sagun

---

**Affiliation:** EPFL **DOB:** March 28, 1987 **Contact:** leventsagun@gmail.com

## Research Interests

Probability theory, statistical physics, and their applications on deep learning. Statistical and geometric properties of loss functions in machine learning. The effects and implications of theoretical results to problems that arise in large-scale models and complex datasets.

## Education

- 2011 - 2017** Ph.D. in Mathematics, Courant Institute of Mathematical Sciences, New York University  
*Thesis title:* Explorations on High Dimensional Landscapes: Spin Glasses and Deep Learning  
*Academic advisors:* Gérard Ben Arous and Yann LeCun
- 2006 - 2011** B.S. in Mathematics, Boğaziçi University, Istanbul, Turkey  
*Double major:* B.S. in Physics  
*Exchange program:* Columbia University, New York, Fall 2009  
*Graduated first rank in the Faculty of Arts and Sciences, Class of 2011*
- 2001 - 2006** High School, İstanbul Lisesi, Istanbul, Turkey

## Academic Employment

- 2018 - 2019** Postdoctoral Fellow, École Polytechnique Fédérale de Lausanne, Switzerland  
*Simons Collaboration on Cracking the Glass Problem - Part II*
- 2017 - 2018** Postdoctoral Fellow, École Normale Supérieure Paris & CEA Saclay, France  
*Simons Collaboration on Cracking the Glass Problem - Part I*
- 2015 - 2016** Teaching Assistant, Center for Data Science, New York University
- 2011 - 2017** Research & Teaching Assistant, Department of Mathematics, Courant Institute, NYU

## Industry Employment

- 2016 - 2017** Research Intern and Contractor, Facebook AI Research, New York  
*Project:* Local geometry of solutions and effects of noisy algorithms in deep learning.  
*Manager:* Léon Bottou

## Papers

1. Stefano Spigler, Mario Geiger, Stéphane d’Ascoli, Levent Sagun, Giulio Biroli, Matthieu Wyart. A jamming transition from under- to over-parametrization affects loss landscape and generalization. Integration of Deep Learning Theories Workshop at NeurIPS, 2018.
2. David Lopez-Paz, Levent Sagun. Easing non-convex optimization with neural networks. Workshop Track at ICLR, 2018.

3. Marco Baity-Jesi, Levent Sagun, Mario Geiger, Stefano Spigler, Gérard Ben Arous, Chiara Cammarota, Yann LeCun, Matthieu Wyart, Giulio Biroli. Comparing Dynamics: Deep Neural Networks versus Glassy Systems. Conference Paper at ICML, 2018.
4. Levent Sagun, Utku Evci, V. Uğur Güney, Yann Dauphin, Léon Bottou. Empirical Analysis of the Hessian of Over-Parametrized Neural Networks. Workshop Track at ICLR, 2018.
5. Levent Sagun, Thomas Trogdon, Yann LeCun. Universal halting times in optimization and machine learning. Optimization Workshop at ICML, 2016; and *Quart. Appl. Math.*, **76**, 289-301, 2018.
6. Andrew J. Ballard, Ritankar Das, Stefano Martiniani, Dhagash Mehta, Levent Sagun, Jacob D. Stevenson, David J. Wales. Perspective: Energy Landscapes for Machine Learning. *Phys. Chem. Chem. Phys.*, **19**, 12585-12603, 2017.
7. Matthew Dunn, Levent Sagun, Hale Şirin, Daniel Chen. Early Predictability of Asylum Court Decisions. ML and the Law, ICAIL Pages 233-236, 2017.
8. Pratik Chaudhari, Anna Choromanska, Stefano Soatto, Yann LeCun, Carlo Baldassi, Christian Borgs, Jennifer Chayes, Levent Sagun, Riccardo Zecchina. Entropy-SGD: Biasing Gradient Descent Into Wide Valleys. Conference Track at ICLR, 2017.
9. Levent Sagun, Uğur Güney, Gérard Ben Arous, Yann LeCun. Explorations on high dimensional landscapes. Workshop Track at ICLR, 2015.

## Preprints

1. Mario Geiger, Arthur Jacot, Stefano Spigler, Franck Gabriel, Levent Sagun, Stéphane d’Ascoli, Giulio Biroli, Clément Hongler, Matthieu Wyart. Scaling description of generalization with number of parameters in deep learning. Preprint, available at arXiv:1901.01608, 2019.
2. Mario Geiger, Stefano Spigler, Stéphane d’Ascoli, Levent Sagun, Marco Baity-Jesi, Giulio Biroli, Matthieu Wyart. The jamming transition as a paradigm to understand the loss landscape of deep neural networks. Preprint, available at arXiv:1809.09349, 2018.
3. Matthew Dunn, Levent Sagun, Mike Higgins, Uğur Güney, Volkan Cirik, Kyunghyun Cho. SearchQA: A New Q&A Dataset Augmented with Context from a Search Engine. Preprint, available at arXiv:1704.05179, 2017.
4. Levent Sagun, Léon Bottou, Yann LeCun. Eigenvalues of the Hessian in Deep Learning: Singularity and Beyond. Preprint, available at arXiv:1611.07476, 2017.

## Reviewer duties

ICLR, JMLR, COLT, NeurIPS, TPAMI, SIAM Journal, Cambridge University Press

## Talks and Presentations

A jamming transition from under- to over-parametrization affects loss landscape and generalization, NeurIPS Workshop on Integration of Deep Learning Theories, Montreal, December 2018  
 Over-parametrization in NNs: observations and a definition, Google Brain, Montreal, November 2018  
 Over-parametrization in NNs: observations and a definition, FAIR, Montreal, November 2018  
 Over-parametrization in NNs: observations and a definition, MILA, Montreal, November 2018

An empirical look at the loss landscape: shape, dynamics, and loose ends, Statistical Physics and Machine Learning back together, Institut d'Études Scientifiques de Cargèse, Corsica, August 2018  
 Over-parametrization in Deep Learning, Institut für Theoretische Physik, Heidelberg, July 2018  
 Easing non-convex optimization with neural networks, ICLR 2018 Poster, May 2018  
 Empirical Analysis of the Hessian of Over-Parametrized Neural Networks, ICLR 2018 Poster, May 2018  
 Shape and Dynamics of the Loss Landscape, CILVR Lab Meetings, NYU, New York, April 2018  
 Shape and Dynamics of the Loss Landscape, High-Dimensional Dynamics Workshop, New York, April 2018  
 Comparing Dynamics: Deep Neural Networks versus Glassy Systems, Mathematics, Information and Computation Seminar, NYU, New York, March 2018  
 Over-parametrization in Deep Learning, Invited Visit at Google Brain, Zurich, January 2018  
 Over-parametrization in Deep Learning, Télécom ParisTech, Paris, November 2017  
 Over-parametrization in Deep Learning, Université Paris-Sud, Paris, November 2017  
 Over-parametrization in Deep Learning, Facebook AI Research, Paris, October 2017  
 Over-parametrization in Deep Learning, ENS, Paris, October 2017  
 The Role of Over-Parametrization in Optimization, Nonlinear and Stochastic Optimization at SIAM Conference on Optimization, Vancouver, May 2017  
 Singularity of the Hessian in deep learning, CILVR Lab Meetings, NYU, New York, November 2016  
 Universality in Halting Time, Spotlight talk at Optimization Workshop, ICML, New York, June 2016  
 Universality in Halting Time and Optimization, CILVR Lab Meetings, NYU, New York, April 2016  
 Explorations on High Dimensional Landscapes, Spotlight presentation, MLSS Kyoto, September 2015  
 Explorations on High Dimensional Landscapes, Poster presentation, DLSS, U. of Montreal, August 2015  
 Explorations on High Dimensional Landscapes, Poster presentation, ICLR, San Diego, May 2015  
 Optimization in High Dimensional Landscapes, ML Seminar, Boğaziçi University, January 2015  
 High Dimensional Landscapes: Spin Glasses and Deep Learning, ML Seminar, NYU, September 2014  
 Critical Points in High Dimensional Landscapes, NCAP Presentations by Participants, U of T, August 2014  
 An Introduction to Longest Increasing Subsequences, Boğaziçi University, June 2011  
 Insight into Theory of Large Deviations, Department of Mathematics, Boğaziçi University, May 2010  
 Conjugate Gradient Method, Student Seminars in Mathematics and Physics, Boğaziçi University, April 2010

## Educational Experience

Statistical Physics and Machine Learning back together, Institut d'Études Scientifiques de Cargèse, Corsica, August 2018  
 Deep Learning and Statistical Physics, Beg Rohu Summer School, June 2018  
 Machine Learning Summer School, University of Kyoto, September 2015  
 Deep Learning Summer School, University of Montreal, August 2015  
 CIFAR Neural Computation & Adaptive Perception Summer School, University of Toronto, August 2014  
 Workshop on Stochastic Gradient Methods, IPAM, February 2014  
 St. Petersburg School in Probability and Statistical Physics, Chebyshev Laboratory, June 2012  
 Minicourse on Compressed Sensing by Emmanuel Candès, University of Cambridge, March 2011  
 Summer School on Complex Systems, Institute for Theoretical and Applied Physics, Summer 2010  
 Summer School on Computational Solution of Inverse Problems, Helsinki University, Summer 2010  
 Summer School on Dynamical Systems and Their Applications, Utrecht University, Summer 2009  
 Independent Study on Dynamical Systems, Boğaziçi University, Summer 2009

## Teaching Experience

Recitation leader, Statistical and Mathematical Methods, NYU - Center for Data Science, Fall 2016  
 Recitation leader, Theory of Probability, Courant Institute, Fall 2016  
 Recitation leader, Machine Learning, NYU - Center for Data Science, Spring 2016

Recitation leader, Statistical and Mathematical Methods, NYU - Center for Data Science, Fall 2015  
Recitation leader, Probability and Statistics, Courant Institute, Spring 2015  
Recitation leader, Theory of Probability, Courant Institute, Fall 2014  
Teaching assistant, Introduction to Mathematical Analysis II, Courant Institute, Spring 2014  
Instructor for the Written Exam Workshop, Courant Institute, Fall 2013  
Teaching Assistant for Statistics and Probability, Boğaziçi University, 2009-2010

## Honors, Awards & Fellowships

MacCracken Fellowship, New York University, 2011-2016  
Dora Aksoy Award, Department of Mathematics, Boğaziçi University, 2011  
Hilmi Tolun Award, Faculty of Arts and Sciences, Boğaziçi University, 2011  
Honorable Mention, 9<sup>th</sup> National Mathematical Olympiads, Akdeniz University, Antalya, Turkey, 2004  
Silver Medal, 5<sup>th</sup> National Mathematical Olympiads for Junior High School Students, TÜBİTAK, 2001

## References

### *Research*

**Gérard Ben Arous** - [benarous@cims.nyu.edu](mailto:benarous@cims.nyu.edu)

Professor of Mathematics, Courant Institute of Mathematical Sciences, NYU

Associate Provost for the Quantitative Disciplines, Global Network Professor of Mathematics, NYU Shanghai

*Relationship:* PhD thesis advisor

**Léon Bottou** - [leonb@fb.com](mailto:leonb@fb.com)

Research Lead, Facebook AI Research, New York

Visiting Research Professor, Computer Science Department, NYU

*Relationship:* Manager at FAIR and PhD thesis co-advisor

**Giulio Biroli** - [giulio.biroli@ens.fr](mailto:giulio.biroli@ens.fr)

Professor of Theoretical Physics

Laboratoire de Physique Statistique, École Normale Supérieure, PSL Research University, Paris, France

*Relationship:* Postdoc advisor

### *Teaching*

**Carlos Fernandez-Granda** - [cfgranda@cims.nyu.edu](mailto:cfgranda@cims.nyu.edu)

Assistant Professor of Mathematics and Data Science

Courant Institute of Mathematical Science and Center for Data Science NYU

*Relationship:* I have been his teaching assistant at the Center for Data Science

Last updated: January 10, 2019