Stefano **Sarao Mannelli**

Gatsby Computational Neuroscience Unit & Sainsbury Wellcome Centre, University College London, UK.

stefano.sarao@gmail.com - scholar profile - ORCID

I am a curious and self-motivated researcher who enjoys working on interdisciplinary topics. I believe that scientific exchange and communication are key to progress, thus I actively promote healthy research environments by organising events that serve as communication channels. In my work, I am keen on understanding learning at a fundamental level, in both machines and animals. I am a theoretical physicist by training, specialised in statistical physics of disordered systems, and physics gave me the mindset and some of the tools I use to probe neural networks. While part of my research is devoted to providing new technical tools, the majority of my studies focus on specific aspects of learning by building parsimonious models that capture the relevant features of a phenomenon while retaining analytical tractability. My scientific goal is to reconcile the small and the big picture by analysing how microscopic aspects –such as data structure, learning rule, and loss metric– can affect macroscopic observations –such as test error and accuracy.

Research Experience

- Jan. 2023 – present: **Senior Research Fellow** in the Gatsby Computational Neuroscience Unit and Sainsbury Wellcome Centre, University College London, UK - Supervisor: **Andrew Saxe**;
    - *My research focuses on comparing learning in neural networks and animals by building exactly solvable models to probe complex learning protocols, i.e. curriculum, transfer, continual, and reinforcement learning. From this comparison, I get insights into animal behaviour and neural networks (mis)functioning that lead to new counter-intuitive understandings, in particular to a theory for the ineffectiveness of curricula in ML –in contrast with animal learning– and to a new experimental design.*
- Jun. 2021 – Dec.2022: **Research Fellow** in the Gatsby Computational Neuroscience Unit and Sainsbury Wellcome Centre, University College London, UK - Supervisor: **Andrew Saxe**;
- Oct. 2020 – Jun. 2021: **Research Fellow** in the Department of Experimental Psychology, University of Oxford, UK - Supervisor: **Andrew Saxe**;
- Mar. 2020 - Apr. 2020: **Visiting researcher** with **Eric Vanden-Eijnden** at NYU, USA;
    - *I studied analytically the role of over-parameterisation on the loss landscape of two-layer quadratic neural networks teacher-student models –including the phase retrieval problem– and characterised a trivialisation transition –i.e. when the non-convexity of the landscape disappears.*
- Jan. 2019 - Mar. 2019: **Visiting researcher** at KITP, UCSB, U.S.A.;
- Feb. 2018: **Visiting researcher** at Duke University, U.S.A.;
- Mar. 2017 - Jul. 2017: **Research intern** with **Lenka Zdeborová** at IPhT, CEA, France;
- Jan. 2016 - Apr. 2016: **Research intern** with **Nicolas Vayatis** at CMLA, ENS Cachan, France;

Education

- Oct. 2017 – Oct. 2020: **Ph.D. in Theoretical Physics** at IPhT, CEA, France - Supervisor: **Lenka Zdeborová**;
    - *In my thesis, I developed new methods to study full-batch gradient descent in prototypical non-convex high-dimensional problems. My works reconciled dynamics and geometrical properties debunking common misconceptions about the role of spurious minima in non-convex optimisation.*
- Sept. 2016 - Oct. 2017: **M.Sc. in Electronic Engineering** at Politecnico di Torino, Italy, grade: *110/110 cum laude*;
- Sept. 2014 - Jul. 2016: **M.Sc. in Physics of Complex Systems** at Politecnico di Torino-SISSA, Italy, grade: *110/110 cum laude*;
- Sept. 2014 - Jul. 2016: **M2, Physique Théorique** at Paris Diderot, UPMC, and ENS Cachan, France;
- Sept. 2014 - Jul. 2016: **Alta Scuola Politecnica Diploma** at Politecnico di Torino, Politecnico di Milano, Italy;
- Sept. 2014 - Jun. 2017: **M.Sc. in Engineering Physics** at Politecnico di Milano, Italy, grade: *110/110 cum laude*;
- Sept. 2011 - Jul. 2014: **B.Sc. in Mathematics for Engineering** at Politecnico di Torino, Italy, grade: *110/110 cum laude*.

Organisation of Scientific Events

- (future event) A.Y. 2025-26: KITP program on *Psychophysics and Statistical Physics of Learning in Higher Level Cognition*, Santa Barbara, USA [link][Proposal submitted];

- ⸪ (future event) Aug. 2024: School on *Analytical Connectionism*, New York, USA [link][Confirmed];
- ⸪ (future event) Mar. 2024: Workshop on *Bridging Analytical and Experimental Insights into Representational Change* as part of CoSyNe, Lisbon, Portugal [link][Confirmed];
- ⸪ Oct. 2023: Workshop on *Analytical Approaches for Neural Network Dynamics*,Paris,FRA [link][Secured 10K EUR];
- ⸪ Aug.-Sep. 2023: School and workshop on *Analytical Connectionism*, London, UK [link][Secured 42K GBP];
- ⸪ Jun. 2022: Workshop *Communication Across Communities in Machine Learning Research and Practice* as part of FAccT 2022, Seoul, South Korea [link];
- ⸪ May 2021: Workshop *Science and Engineering of Deep Learning* as part of ICLR 2021, virtual [link].

Grants & Awards

- ⸪ Nov. 2023: UK-IT Trustworthy AI Exchange Programme, 4500 GBP, The Alan Turing Institute;
- ⸪ Mar. 2023: SCGB Conference Awards, 3000 GBP, Simons Foundation;
- ⸪ Sep. 2017: Ph.D. scholarship from CEA;
- ⸪ Feb. 2017 & Feb. 2016: mobility grant *Tesi su proposta*, 2200 EUR/ea., Politecnico di Torino;
- ⸪ Sep. 2013, Sep. 2012 & Sep. 2011: accommodation scholarship *Borsa talenti* from Fondazione CEUR;
- ⸪ Sep. 2012: 3rd prize at individual math competition *Alfa Class*, 500 EUR, from Fondazione CRT.

Teaching and Mentoring

- ⸪ 13-17 Jun. 2022: **Teaching assistant** at Statistical Physics of Deep Learning summer school, Como, Italy [8h];
- ⸪ Mar. 2019-May 2019: **Teaching assistant** in "Statistical Physics" with **Gregory Schehr** at Supelec, France [15h];
- ⸪ Mar. 2018-Jun. 2018: **Teaching assistant** in "Statistical Physics" with **Gregory Schehr** at Supelec, France [20h];
- ⸪ Sept.-Dec. 2023: **Mentoring** Gabriel Ocana Santero, Ph.D. student, University of Oxford;
- ⸪ Mar.-Dec. 2023: **Mentoring** Jin Lee, Ph.D. student, University College London;
- ⸪ July-Sept. 2023: **Mentoring** Anchit Jain, BSc, University of Cambridge;
- ⸪ Sept. 2022-Aug. 2023: **Mentoring** Nishil Patel, Research Assistant, University College London;
- ⸪ Feb.-July 2022: **Mentoring** Gaspard Abel, MSc, École normale supérieure;
- ⸪ Apr.-June 2023: **Mentoring** Léo Touzo, MSc, École normale supérieure.

Positions of Responsibility

- ⸪ A.Y. 2022-2023: *Organiser* of Neuroscience external seminars at Gatsby Computational Neuroscience Unit;
- ⸪ A.Y. 2021-2022: *Organiser* of SaxeLab seminars.
- ⸪ Review for NeurIPS, ICML, ICLR, ScyPost, JSTAT, JPhysA, Physica A.

Conferences and Seminars

- ⸪ Jul. 2023: *Speaker and poster* at Statistical Physics and ML Back Together Again, Cargese, France;
- ⸪ Jun. 2022: *Speaker and poster* at ICML, Baltimore, Maryland (USA);
- ⸪ Jun. 2022: *Posters* at Youth in High-Dimensions, Trieste, Italy;
- ⸪ May 2022: *Speaker* at Tri-Centre meeting at Hebrew University, Jerusalem, Israel;
- ⸪ Apr. 2022: *Seminar* at Bocconi University, Milan, Italy;
- ⸪ Mar. 2022: *Posters* Cosyne, Lisbon & Cascais, Portugal;
- ⸪ Nov. 2021: *Seminar* at King's College, London, UK;
- ⸪ Jun. 2021: *Posters* at Youth in High-Dimensions, online conference;
- ⸪ Dec. 2020: *Posters* at NeurIPS, online conference;
- ⸪ Sep. 2020: *Speaker* at Computational Phase Transitions Simons Institute, Berkeley, California (USA);
- ⸪ Mar. 2020: *Seminar* at CUNY, NY, USA;
- ⸪ Feb. 2020: *Seminar* at EPFL, Switzerland;
- ⸪ Dec. 2019: *Speaker* and poster at NeurIPS, Vancouver (Canada);
- ⸪ Jul. 2019: *Speaker* at Theoretical Advances in Deep Learning, Turkey;
- ⸪ Jun. 2019: *Speaker* and poster at ICML, California (USA);
- ⸪ Jun. 2019: *Speaker* Day on methods for dynamical mean field equation, Paris, France;
- ⸪ Jan. 2019: *Poster* at The Rough High-Dimensional Landscape Problem KITP, UCSB, California (USA);
- ⸪ Feb. 2019: *Poster* at Machine Learning for Quantum Many-Body Physics KITP, UCSB, California (USA);
- ⸪ Sep. 2017: *Speaker* at Junior Conference on Data Science and Engineering, France;

Publication List

Patel, N., Lee, S., <u>Sarao Mannelli</u>, S., Goldt, S., & Saxe, A. (2023). *The RL Perceptron: Dynamics of Policy Learning in High Dimensions.* ICLR 2023 Workshop on Physics4ML. Full paper submitted (arXiv preprint 2306.10404);

Gerace, F., Doimo, D., <u>Sarao Mannelli</u>, S., Saglietti, L., & Laio, A. (2023). *Optimal transfer protocol by incremental layer defrosting.* Submitted (arXiv preprint 2303.01429);

<u>Sarao Mannelli</u>, S., Gerace, F., Rostamzadeh, N., Saglietti, L. (2022). *Bias-inducing geometries: exactly solvable data model with fairness implications.* Submitted (arXiv preprint 2205.15935);

Lee, S., <u>Sarao Mannelli</u>, S., Clopath, C., Goldt, S., Saxe, A. (2022). *Maslow's Hammer in Catastrophic Forgetting: Node Re-Use vs. Node Activation.* ICML 2022;

Saglietti, L., <u>Sarao Mannelli</u>, S., Saxe, A. (2022). *An Analytical Theory of Curriculum Learning in Teacher-Student Networks.* J.Stat.Mech 2022(11), 114014 and NeurIPS 2022;

Gerace, F., Saglietti, L., <u>Sarao Mannelli</u>, S., Saxe, A., Zdeborová, L. (2021). *Probing transfer learning with a model of synthetic correlated datasets.* Machine Learning: Science and Technology 3 (1);

<u>Sarao Mannelli</u>, S., & Urbani, P. (2021). *Just a Momentum: Analytical Study of Momentum-Based Acceleration Methods in Paradigmatic High-Dimensional Non-Convex Problem.* NeurIPS 2021;

Baker, A., Biazzo, I., Braunstein, A., Catania, G., Dall'Asta, L., Ingrosso, A., Krzakala, F., Mazza, F., Mézard, M., Muntoni, A.P., Refinetti, M., <u>Sarao Mannelli</u>., S., Zdeborová, L. (2020). *Epidemic mitigation by statistical inference from contact tracing data.* Proceedings of the National Academy of Sciences PNAS 118 (32);

<u>Sarao Mannelli</u>, S., Vanden-Eijnden, E., & Zdeborová, L. (2020). *Optimization and Generalization of Shallow Neural Networks with Quadratic Activation Functions.* NeurIPS 2020;

<u>Sarao Mannelli</u>, S., Biroli, G., Cammarota, C., Krzakala, F., Urbani, P., & Zdeborová, L. (2020). *Complex Dynamics in Simple Neural Networks: Understanding Gradient Flow in Phase Retrieval.* NeurIPS 2020;

<u>Sarao Mannelli</u>, S., & Zdeborová, L. (2020). *Thresholds of descending algorithms in inference problems.* J. Stat. Mech. (2020) 034004;

<u>Sarao Mannelli</u>, S., Biroli, G., Cammarota, C., Krzakala, F., & Zdeborová, L. (2019). *Who is Afraid of Big Bad Minima? Analysis of Gradient-Flow in a Spiked Matrix-Tensor Model.* NeurIPS 2019 (spotlight);

<u>Sarao Mannelli</u>, S., Krzakala, F., Urbani, P., & Zdeborova, L. (2019). *Passed & Spurious: analysing descent algorithms and local minima in spiked matrix-tensor model.* ICML 2019;

<u>Sarao Mannelli</u>, S., Biroli, G., Cammarota, C., Krzakala, F., Urbani, P., & Zdeborová, L. (2018). *Marvels and pitfalls of the Langevin algorithm in noisy high-dimensional inference.* Phys. Rev. X 10, 011057.