

Oliver Sieberling

osieberl@mit.edu | [Homepage](#) | [Google Scholar](#) | [GitHub](#)

EDUCATION

Massachusetts Institute of Technology

Ph.D. Computer Science

- Advisor: Yoon Kim

Cambridge, USA

Sep. 2025 – Present

ETH Zurich

B.Sc. Computer Science

- Thesis: *Dynamic Monotone Functions on the (1+1)-Evolutionary Algorithm: Drift Minimization Does Not Maximize Hardness*. Advised by Johannes Lengler, Institute of Theoretical Computer Science.
- Bronze Medal, ICPC SWERC 2024.

Zurich, Switzerland

Sep. 2021 – Aug. 2025

Princeton University

Departmental Exchange Program

Princeton, USA

Sep. 2023 – Jan. 2024

EXPERIENCE

MIT-IBM Watson AI Lab

Research Intern

- Worked on improving linear attention, with Bharat Runwal and Rameswar Panda.

Cambridge, USA

June 2026 – Aug. 2026

ETH Zurich

Teaching Assistant

- Led weekly tutorials for *Algorithms & Probability*, taught by Angelika Steger and Rasmus Kyng.

Zurich, Switzerland

Feb. 2024 – June 2024

PUBLICATIONS

- **Oliver Sieberling**, Bharat Runwal, Rameswar Panda, Yoon Kim, *Dynamic Short Convolutions Improve Transformers*, Preprint, 2026. [Link](#).
- Zhaofeng Wu, **Oliver Sieberling**, Shawn Tan, Rameswar Panda, Yury Polyanskiy, Yoon Kim, *Variable-Width Transformers*, Preprint, 2026. [Link](#).
- Johannes von Oswald*, Nino Scherrer*, Seijin Kobayashi, Luca Versari, Songlin Yang, Maximilian Schlegel, Kaitlin Maile, Yanick Schimpf, **Oliver Sieberling**, Alexander Meulemans, Guillaume Lajoie, Rif A. Saurous, Charlotte Frenkel, Razvan Pascanu, Blaise Agüera y Arcas, João Sacramento, *MesaNet: Sequence Modeling by Locally Optimal Test-Time Training*, ICLR, 2026. [Link](#).
- Shengkun Tang, **Oliver Sieberling**, Eldar Kurtic, Zhiqiang Shen, Dan Alistarh, *DarwinLM: Evolutionary Structured Pruning of Large Language Models*, COLM, 2026. [Link](#).
- Roberto Castro*, Andrei Panferov*, Soroush Tabesh, **Oliver Sieberling**, Jiale Chen, Mahdi Nikdan, Saleh Ashkboos, Dan Alistarh, *Quartet: Native FP4 Training Can Be Optimal for Large Language Models*, NeurIPS, 2025. [Link](#).
- **Oliver Sieberling**, Denis Kuznedelev, Eldar Kurtic, Dan Alistarh, *EvoPress: Accurate Dynamic Model Compression via Evolutionary Search*, ICML, 2025. [Link](#).
- Johannes Lengler ^{$\alpha\beta$} , Leon Schiller ^{$\alpha\beta$} , **Oliver Sieberling** ^{$\alpha\beta$} , *Plus Strategies are Exponentially Slower for Planted Optima of Random Height*, GECCO, 2024. [Link](#). Extended journal version published in *Algorithmica*, 2026.
- Marc Kaufmann ^{$\alpha\beta$} , Maxime Larcher ^{$\alpha\beta$} , Johannes Lengler ^{$\alpha\beta$} , **Oliver Sieberling** ^{$\alpha\beta$} , *Hardest Monotone Functions for Evolutionary Algorithms*, EvoCOP, 2024. Best Paper Award Nomination. [Link](#). Extended journal version published in *SN Computer Science*, 2025.

*Equal contribution. ^{$\alpha\beta$} Authors listed alphabetically (TCS).

TALKS

- **Cohere Labs Community**: Dynamic Short Convolutions Improve Transformers. August 2026.
- **Pinterest**: Dynamic Short Convolutions Improve Transformers. August 2026.
- **EvoCOP 2024**: Hardest Monotone Functions for Evolutionary Algorithms. April 2024.

ACADEMIC SERVICE

- **Reviewer**: NeurIPS 2026, ICML 2026, TMLR 2026.