Advait Gadhikar

gadhikaradvait@gmail.com | Phone: +91 83088 25295 | https://advaitgadhikar.github.io/

EDUCATION

CISPA Helmholtz Center for Information Security — Saarbrücken, Germany
PhD in Computer Science (Universität des Saarlandes) — *April 2022 - Nov 2026*

Carnegie Mellon University — Pittsburgh, PA
Master of Science in Electrical and Computer Engineering (4.0/4.0) — *Dec 2020*

Birla Institute of Technology and Science, Pilani — Goa, India
Bachelor of Engineering (Honors) in Electrical and Electronics Engineering (9.17/10) — *Aug 2019*

EXPERIENCE

Qualcomm AI Research — Amsterdam, NL
Senior Researcher, Manager: *Markus Nagel* — *October 2026 - present*

- Building novel quantization methods for edge efficiency.

Microsoft Research — Bengaluru, India
Researcher, Manager: *Manik Varma* — *May - September 2026*

- Working with the retrieval team (ReFoRM) to build efficient dense retrieval models.
- Leveraging novel distillation based methods to push the accuracy-latency Pareto frontier in embedding models.

Relational ML Lab, CISPA — Saarbrücken, Germany
Doctoral Student, Advisor: *Rebekka Burkholz* — *Apr 2022 - Mar 2026*

- Developed methods to build trainable sparse networks from scratch with implicit overparameterization..
- Improved mask identification, initialization and optimization for better sparse training.
- Resulted in five A* publications during the course of my PhD.

Microsoft Research — Cambridge, UK
Research Intern, Advisor: *James Hensman* — *May 2025 - July 2025*

- Designed a data-free optimization method to minimize outliers via rotation for Post Training Quantization.
- Developed theoretical insights from incoherence processing to find optimal rotations in constrained settings.
- Our resulting method reduces parameter incoherence and improves quantization SNR as well as downstream performance on LM eval tasks by up to 2%.

Bosch Center for Artificial Intelligence (BCAI) — Renningen, Germany
Research Intern, Advisor: *Lukas Schott* — *June 2024 - Nov 2024*

- Leveraged dynamic computation for faster inference in large scale Vision Transformers (ViTs).
- Developing methods for learnt routing in the absence of a router for improved inference and stable training.
- Proposed Attention based routing for Mixture-of-Depths (*A-MoD*) models and achieved improved performance with faster convergence.

Optimization, Probability and Learning Lab, CMU — Pittsburgh, PA
Research Assistant, Advisor: *Gauri Joshi* — *May 2020 - June 2021*

- Developed AdaQuantFL, an adaptive quantization method for gradient compression, to minimize the number of bits used for communication during training.
- Created the codebase and conducted experiments to test AdaQuantFL in PyTorch to validate the theoretical propositions about improved communication efficiency for model convergence.
- Benchmarked popular Federated Learning algorithms like FedAvg, FedAdam and FedAdagrad in the Tensorflow Federated framework and verifying their theoretical guarantees for a field guide on Federated Learning in collaboration with Google's Federated Learning team.

Qualcomm Technologies Inc — San Diego, CA
Video Algorithms Intern — *May - Aug 2020*

- Learned the non-linear encoding of a video codec (H-264) and estimated the quality of the encoded video.
- Designed a video denoising model to improve the visual quality of the encoded video, transmitted over a low bitrate (1.2 Mbps) in an unsupervised setting.

Language Technologies Institute, CMU Pittsburgh, PA
Teaching Assistant for Introduction to Deep Learning (11785), Advisor: *Prof Bhiksha Raj* *Dec 2019 - May 2020*

- Collaborated with other TAs to design the Homework such that students can make their own Deep Learning toolkit written in Numpy, termed 'MyTorch' and conducted tutorials on fundamental deep learning concepts.

Indian Institute of Science Bengaluru, India
Research Intern (Bachelor Thesis), Advisor: *Prof Rajiv Soundararajan* *July 2018- Dec 2018*

- Trained a Blind Image Quality Network (VGG-19) without human opinion scores to perform on par with Full Reference quality metrics like the Natural Image Quality Evaluator (NIQE). Integrated the Blind Image Quality Network in a Boundary Equilibrium Generative Adversarial Network to improve the image quality by 16%.

PUBLICATIONS

- Chinmay Joshi, **Advait Gadhikar**, Celia Rubio-Madrigal, Aneet Kumar Dutta, Mridula Singh, and Rebekka Burkholz. "Prune to Protect: Faster Training and Enhanced Privacy by Dynamic Data Pruning." *in the MemFM Workshop at ICML 2026*.
- Tom Jacobs, **Advait Gadhikar**, Celia Rubio-Madrigal, and Rebekka Burkholz. "Hyperbolic Aware Minimization for Sparse Training." *in the International Conference on Learning Representations, 2026.*
- **Advait Gadhikar**, Riccardo Grazzi and James Hensman. "OptRot: Mitigating Weight Outliers via Data-Free Rotations for Post-Training Quantization." *in the ML for Systems Workshop at NeurIPS 2025.*
- Chao Zhou, Tom Jacobs, **Advait Gadhikar**, and Rebekka Burkholz. "Pay Attention to Small Weights." *NeurIPS*, *Dec 2025*.
- **Advait Gadhikar***, Tom Jacobs*, Chao Zhou and Rebekka Burkholz. "Sign-In to the Lottery: Reparameterizing Sparse Training From Scratch." *NeurIPS*, *Dec 2025*.
- **Advait Gadhikar**, Souptik Majumdar, Niclas Popp, Piyapat Saranrittichai, Martin Rapp, and Lukas Schott. "Attention Is All You Need For Mixture-of-Depths Routing." *in the ICLR 2025 Workshop on Scalable Optimization for Efficient and Adaptive Foundation Models (SCOPE)*.
- **Advait Gadhikar**, and Rebekka Burkholz. "Masks, Signs, and Learning Rate Rewinding" *in the International Conference on Learning Representations, 2024.* (Spotlight).
- **Advait Gadhikar**, Sohom Mukherjee, and Rebekka Burkholz. "Why Random Pruning Is All We Need to Start Sparse" *in the International Conference on Machine Learning, 2023.*
- Divyansh Jhunjhunwala, Ankur Mallick, **Advait Gadhikar**, Swanand Kadhe, Gauri Joshi. "Leveraging Spatial and Temporal Correlations in Sparsified Mean Estimation", *NeurIPS*, *Dec 2021.*
- Wang, Jianyu, et al. "A field guide to federated optimization." *arXiv preprint arXiv:2107.06917* (2021).
- Divyansh Jhunjhunwala, **Advait Gadhikar**, Gauri Joshi, Yonina Eldar. "Adaptive Quantization of Model Updates for Communication-Efficient Federated Learning", *International Conference on Acoustics, Speech, and Signal Processing (ICASSP), June 2021.*

CONFERENCE & TALKS

- CPAL 2026 (Recent Spotlight Track), 'Sign-In to the Lottery: Reparameterized Sparse Training.', 'Hyperbolic Aware Minimization for Sparse Training.' and 'Pay Attention to Small Weights.' (March, 2026)
- Microsoft Research, India and Cambridge (UK) 'Signs and Rotations'. (Dec, 2025)
- Qualcomm AI Research, Amsterdam, 'Signs and Rotations'. (Jan, 2026)
- Conference on Parsimony and Learning 2025 (Recent Spotlight Track), 'Masks, Signs and Learning Rate Rewinding'. (March, 2025)
- Helmholtz AI Conference 2024, oral presentation on 'Masks, Signs, and Learning Rate Rewinding'. (June, 2024)
- Cohere AI, invited talk on sparse training for neural networks. (Sept, 2023)

CO-SUPERVISION

- Chinmay Joshi (Msc, 2026), Sree Harsha Nelaturu (MSc, 2025 – now at ZIB, Berlin), Gowtham Krishna (MSc, 2024 – now at NEC labs, Heidelberg)

SERVICE

- Reviewer for ICML, NeurIPS and ICLR from 2023. Top Reviewer (NeurIPS 2024, Gold Reviewer ICML 2026)
- Reviewer for ACL 2026.