

Mert Cemri

LinkedIn: [mert-cemri-4a727913a](#)

Google Scholar: [Mert Cemri](#)

Email : mert_cemri@berkeley.edu

Mobile : +1-510-229-2502

EDUCATION

- **University of California, Berkeley** Berkeley, CA, USA
• *Ph.D. in Electrical Engineering and Computer Sciences* *August, 2023 - May 2028*
Co-advisors: Prof. Ion Stoica, Prof. Kannan Ramchandran and Prof. Alex Dimakis
Affiliations: Berkeley AI Research (BAIR), Sky Computing Lab, and BLISS
CGPA: 4.0/4.0
- **Bilkent University** Ankara, Turkey
• *B.S. in Electrical and Electronics Engineering* *September, 2019 - June, 2023*
CGPA: 4.00/4.00
- **École Polytechnique Fédérale de Lausanne (EPFL)** Lausanne, Switzerland
• *Exchange Semester in Computer Science* *February, 2022 - July, 2022*
Continued research afterwards with Prof. Ali H. Sayed's Adaptive Systems Laboratory

RESEARCH INTERESTS

- Large Language Models, Multi-Agent and Agentic Systems, Reliability and Failure Analysis of AI Systems, AI for Scientific and Algorithmic Discovery, Reinforcement Learning and Post-Training, Test-Time Scaling, Distributed Optimization

PUBLICATIONS

- *Google Scholar: 1,456 citations, h-index 9, i10-index 9 (September 2026).*
- **Peer-Reviewed Conference and Journal Papers**
 1. S. Liu, S. Agarwal, M. Maheswaran, **M. Cemri**, Z. Li, Q. Mang, A. Naren, E. Boneh, A. Cheng, M. Z. Pan, A. Du, K. Keutzer, A. Cheung, A. G. Dimakis, K. Sen, M. Zaharia, and I. Stoica, 'EvoX: Meta-Evolution for Automated Discovery,' in *Conference on Language Modeling (COLM)*, 2026. [**Accepted**]
 2. T. Qin, F. Bai, T.-Y. Hu, R. Vemulapalli, H. S. Koppula, Z. Xu, B. Jin, **M. Cemri**, J. Lu, Z. Wang, and M. Cao, 'COMPASS: Benchmarking Constrained Optimization in LLM Agents,' in *Conference on Language Modeling (COLM)*, 2026. [**Accepted**]
 3. S. Liu, **M. Cemri**, S. Agarwal, A. Krentsel, A. Naren, Q. Mang, Z. Li, A. Gupta, M. Maheswaran, A. Cheng, M. Pan, E. Boneh, K. Ramchandran, K. Sen, M. Zaharia, A. G. Dimakis, and I. Stoica, 'SkyDiscover: A Flexible, Adaptive Framework for AI-Driven Scientific and Algorithmic Discovery,' in *Proc. ACM Conference on AI and Agentic Systems (CAIS)*, pp. 1223–1227, 2026. [**Industry Spotlight / Demonstration**]
 4. **M. Cemri**, M. Pan, S. Yang, L. Agrawal, B. Chopra, R. Tiwari, K. Keutzer, A. Parameswaran, D. Klein, K. Ramchandran, M. Zaharia, J. E. Gonzalez, and I. Stoica, 'Why Do Multi-Agent LLM Systems Fail?,' in *Advances in Neural Information Processing Systems (NeurIPS), Datasets and Benchmarks Track*, 2025. [**Spotlight; 880+ citations**]
 5. H. Bai, Y. Zhou, **M. Cemri**, J. Pan, A. Suhr, S. Levine, and A. Kumar, 'DigiRL: Training In-The-Wild Device-Control Agents with Autonomous Reinforcement Learning,' in *Advances in Neural Information Processing Systems (NeurIPS), Main Track*, 2024. [**260+ citations**]
 6. **M. Cemri**, V. Bordignon, M. Kayaalp, V. Shumovskaia, and A. H. Sayed, 'Asynchronous Social Learning,' in *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2023.
 7. V. Shumovskaia, M. Kayaalp, **M. Cemri**, and A. H. Sayed, 'Discovering Influencers in Opinion Formation over Social Graphs,' *IEEE Open Journal of Signal Processing*, vol. 4, pp. 188–207, 2023.
- **Peer-Reviewed Workshop Papers**
 1. **M. Cemri***, A. Cojocaru*, M. Pan, S. Liu, S. Agarwal, A. Krentsel, J. Tang, K. Ramchandran, J. E. Gonzalez, M. Zaharia, A. Dimakis, and I. Stoica, 'Fantastic Adaptive Taxonomies and How to Use Them,' in *Proc. ICML Workshop on Failure Modes in Agentic AI (FAGEN)*, 2026. [**Best Paper Award; *equal contribution**]
 2. Y. Wu, A. J. Li, W. Ma, Z. Zhou, L. Cao, **M. Cemri**, S. Liu, Y. Xiu, C. Yan, H. Zhao, B. Yu, I. Stoica, and D. Song, 'Frontier Task Synthesis via Solution-Centric Evolution,' in *Proc. 3rd AI for Math Workshop: Toward Self-Evolving Scientific Agents*, 2026.
 3. **M. Cemri**, N. Rajaraman, R. Tiwari, X. Liu, K. Keutzer, I. Stoica, K. Ramchandran, A. Beirami, and Z. Sun, 'SPECS: Faster Test-Time Scaling through Speculative Drafts,' in *Proc. ICML 2025 Workshop on Efficient Systems for Foundation Models (ES-FoMo III)*, 2025.

4. **M. Cemri**, M. Pan, S. Yang, L. Agrawal, B. Chopra, R. Tiwari, K. Keutzer, A. Parameswaran, D. Klein, K. Ramchandran, M. Zaharia, J. E. Gonzalez, and I. Stoica, ‘Why Do Multi-Agent LLM Systems Fail?’, in *Proc. ICLR 2025 Workshop on Building Trust in Language Models and Applications*, 2025.
5. **M. Cemri**, A. Jalal, and K. Ramchandran, ‘Discrete Diffusion Posterior Sampling for Protein Design,’ in *Proc. ICML 2024 Workshop on Structured Probabilistic Inference & Generative Modeling*, 2024.
- **Preprints and Papers Under Review**
 1. Y. Wu, A. J. Li, W. Ma, L. Cao, Z. Zhou, **M. Cemri**, S. Liu, Y. Xiu, C. Yan, H. Zhao, B. Yu, I. Stoica, and D. Song, ‘BenchEvolver: Frontier Task Synthesis via Solution-Centric Evolution,’ arXiv preprint arXiv:2606.01286, 2026.
 2. S. Agarwal, A. Krentsel, S. Liu, **M. Cemri**, A. Cheng, R. Meng, T. Pfister, C.-L. Li, S. Ratnasamy, A. Parameswaran, M. Zaharia, I. Stoica, and M. Lesani, ‘Inductive Deductive Synthesis: Enabling AI to Generate Formally Verified Systems,’ arXiv preprint arXiv:2605.23109, 2026.
 3. S. Liu, A. Krentsel, S. Agarwal, **M. Cemri**, Z. Mao, S. Ponnappalli, A. G. Dimakis, S. Ratnasamy, M. Zaharia, A. Parameswaran, and I. Stoica, ‘The Time is Here for Just-in-Time Systems: Challenges and Opportunities,’ arXiv preprint arXiv:2605.24096, 2026.
 4. **M. Cemri**, S. Agrawal, A. Gupta, S. Liu, A. Cheng, Q. Mang, A. Naren, L. E. Erdogan, K. Sen, M. Zaharia, A. Dimakis, and I. Stoica, ‘AdaEvolve: Adaptive LLM Driven Zeroth-Order Optimization,’ arXiv preprint arXiv:2602.20133, 2026.
 5. A. Cheng, S. Liu, M. Pan, Z. Li, S. Agarwal, **M. Cemri**, B. Wang, A. Krentsel, T. Xia, J. Park, S. Yang, J. Chen, L. Agrawal, A. Naren, S. Li, R. Ma, A. Desai, J. Xing, K. Sen, M. Zaharia, and I. Stoica, ‘Let the Barbarians In: How AI Can Accelerate Systems Performance Research,’ arXiv preprint arXiv:2512.14806, 2025.
 6. B. Jin, T. J. Collins, D. Yu, **M. Cemri**, S. Zhang, M. Li, J. Tang, T. Qin, Z. Xu, J. Lu, G. Yin, J. Han, and Z. Wang, ‘Controlling Performance and Budget of a Centralized Multi-agent LLM System with Reinforcement Learning,’ arXiv preprint arXiv:2511.02755, 2025.
 7. A. Cheng, S. Liu, M. Pan, Z. Li, B. Wang, A. Krentsel, T. Xia, **M. Cemri**, J. Park, S. Yang, J. Chen, L. Agrawal, A. Desai, J. Xing, K. Sen, M. Zaharia, and I. Stoica, ‘Barbarians at the Gate: How AI is Upending Systems Research,’ arXiv preprint arXiv:2510.06189, 2025.
 8. **M. Cemri**, T. Çukur, and A. Koç, ‘Unsupervised Simplification of Legal Texts,’ arXiv preprint arXiv:2209.00557, 2022.

HONORS AND AWARDS

- Best Paper Award, ICML Workshop on Failure Modes in Agentic AI (FAGEN) 2026
- Spotlight paper, NeurIPS Datasets and Benchmarks Track 2025
- EECS Departmental Fellowship, UC Berkeley 2023-2024
- Research Excellence Award, Bilkent University 2023
- Academic Excellence Award, Bilkent University 2023
- Merit scholarship due to outstanding academic success, Bilkent University 2021-2023
- Tuition scholarship due to success in University Entrance Exam, Bilkent University 2019-2023
- Bronze Medal in the International Baltic Sea Philosophical Essay Event 2018
- Silver Medal in National Science and Technology Projects Competition for High School Students 2018

ACADEMIC SERVICE

- **Conference and Workshop Organization**
 - **Program Chair, Area Chair and Organizing Committee member**, NeurIPS 2026 Workshop “Who Verifies the Agents?”, Sydney, Australia, December 2026.
 - **Program Chair and Organizing Committee member**, inaugural ACM CAIS 2026 Workshop “AI Agents for Discovery in the Wild”, San Jose, CA, May 26, 2026. The workshop featured invited talks, contributed papers and posters, demonstrations, and panels.
 - **Organizer**, Berkeley Laboratory for Information and System Sciences (BLISS) Seminar, UC Berkeley.
- **Program Committee and Reviewing**
 - **Conference Program Committee / Reviewer**: AAAI 2027, NeurIPS 2026, ICML 2026, COLM 2026, ICLR 2026.
 - **Journal Reviewer**: Transactions on Machine Learning Research (TMLR).
 - **Workshop Reviewer**: NeurIPS 2026 Workshop “Who Verifies the Agents?”; ACM CAIS 2026 Workshop “AI Agents for Discovery in the Wild”.
- **Invited Talks and Panels**
 - **Panel Moderator**, ACM CAIS 2026 Workshop “AI Agents for Discovery in the Wild”, with Ahmad Beirami, Alex Dimakis, and Mohammad Alizadeh, San Jose, CA, May 2026.

RESEARCH EXPERIENCE

- **University of California, Berkeley** Berkeley, CA
Berkeley AI Research (BAIR), Sky Computing Lab and BLISS - Graduate Researcher August 2023 - Present
 - Lead author of the first large-scale empirical study of failure modes in multi-agent LLM systems, which produced the MAST taxonomy and an accompanying benchmark; the resulting paper was selected as a Spotlight at the NeurIPS 2025 Datasets and Benchmarks Track. Extended this line of work into automatically adaptive failure taxonomies (AdaMAST), which received a Best Paper Award at the ICML 2026 FAGEN workshop.
 - Design efficient inference and test-time scaling algorithms for large language models, including speculative-draft cascades (SPECS) that cut latency while preserving or improving answer quality.
 - Build LLM-driven frameworks for scientific and algorithmic discovery and for zeroth-order optimization (SkyDiscover, AdaEvolve, EvoX), where language models propose, evaluate, and evolve candidate algorithms under an adaptive search budget.
 - Developed novel discrete diffusion and posterior sampling guidance algorithms, with an application to designing unseen protein sequences.
- **École Polytechnique Fédérale de Lausanne (EPFL)** Lausanne, Switzerland
Adaptive Systems Lab (ASL) - Undergraduate Researcher February 2022 - June 2023
 - Worked under the supervision of Prof. Ali H. Sayed on distributed optimization and social learning.
 - Developed the theoretical model of social learning for the asynchronous case. Also conducted research on identifying the most influential agent over a social graph, and recovering the graph topology by analyzing interactions among users in social networks.
- **National Magnetic Resonance Research Center (UMRAM)** Ankara, Turkey
ICON Lab - Undergraduate Researcher March 2021 - July 2023
 - Worked on the applications of neural networks to NLP problems and on developing novel tools to analyze graphical data using graph neural networks under the supervision of Prof. Tolga Çukur and Prof. Aykut Koç.

INDUSTRY EXPERIENCE

- **Google** Sunnyvale, CA
Systems Research Group (SRG), Student Researcher April 2026 - Present
 - Research on AI for systems: using LLM-driven agents to accelerate systems performance research and to synthesize formally verified systems components.
- **Apple Inc.** Cupertino, CA
AI/ML Team - Machine Learning Intern May 2025 - August 2025
 - Worked on agentic post-training of the Apple Foundation Model (AFM) for tool-using capabilities, through new synthetic environment creation and by applying and developing new reinforcement learning (RL) algorithms.
- **Titra Technology** Ankara, Turkey
Autonomous Driving Unit - Project Engineer September 2022 - June 2023
 - Worked on reinforcement learning (RL) based autopilot systems for fixed-winged unmanned aerial vehicles.
- **The Scientific and Technological Research Council of Turkey (TÜBİTAK)** Ankara, Turkey
Advanced Technologies Research Institute (ATRI) - Intern June 2021 - August 2021
 - Worked on electronic warfare technologies. Implemented pseudo-random number generation algorithms on an FPGA board (with VHDL), produced Gaussian noise, and harvested the noise on a Jupyter notebook.
 - Built a neural network on a PYNQ FPGA, and studied how deep learning algorithms are accelerated with FPGAs.

SKILLS

- **Languages:** English (Fluent, TOEFL IBT: 110/120), German (Intermediate, Goethe Zertifikat B1), Turkish (Native)
- **Programming Languages:** Python, Matlab, C/C++, VHDL, Assembly Language, L^AT_EX
- **Frameworks:** vLLM, verl, PyTorch, NumPy, Pandas, Keras, Linux

SELECTED PROJECTS

- **Reinforcement Learning based Autopilot** In this work, we develop an RL-based controller for the attitude and landing controls of a fixed wing unmanned air vehicle. We are using PPO to do online learning.
- **Meta Learning with Zeroth Order Oracle** Studied the zeroth-order (ZO) methods in a learning to learn (L2L) framework with respect to their generalizability. In particular, I analyzed the zero-shot performance of these methods on new datasets and applied several optimization tricks to improve the generalization of this framework. We showed that the enhancements we made accelerate the convergence of the algorithm, and help prevent overfitting.
- **Autonomous Vehicle (EPFL Human-Robot Tandem Race '22)** Developed an autonomous vehicle that can track a particular human, and participated in EPFL Human-Robot Tandem Race with this vehicle. For detection of a selected person, we used a combination of Yolov5 and Pifpaf models, and for tracking, we used DeepSort and ReID methods.
- **Transmitter-Receiver for a AWGN Channel that Adds Random Rotation to Messages** Built a transmitter and a receiver that creates 6-bit latent representations of 7-bit ASCII characters, send it through a WGN channel that also adds a random phase (rotation) to the signals, and successfully recover original word of 7-bit ASCII characters.
- **Gaussian Noise using PYNQ-Z1 board** Implemented Taus-88 algorithm for pseudo-random number generation (PRNG) purposes on PYNQ-Z1 FPGA and generated a Gaussian noise by combining such PRNG algorithms using the Central Limit Theorem. Harvested this noise by building a DMA channel and demonstrated it on a Jupyter notebook.