

# Mher Safaryan | CV

✉ mher.safaryan@ist.ac.at • 🌐 mher-safaryan.github.io  
IST Austria, Am Campus 1, 3400 Klosterneuburg, Austria

## Current Position

- **Postdoctoral Researcher** **ISTA, Austria**  
*Institute of Science and Technology Austria (ISTA)*  
*Optimization Theory and Algorithms for Machine Learning, advisor: Prof. Dan Alistarh*  
Nov 2022–present
- **Marie Skłodowska-Curie Fellowship** **ISTA, Austria**  
*Marie Skłodowska-Curie Actions (MSCA) COFUND IST-BRIDGE*  
Nov 2022–Apr 2025
- **Industrial Secondment** **Neural Magic Inc., USA**  
*New Optimization Methods for LLM Quantization, advisor: Dr. Alexandre Marques*  
Sep 2024–Feb 2025

## Research Interests

- ◇ optimization (theory and algorithms), machine learning, federated learning
- ◇ large-scale, convex/non-convex, stochastic/deterministic optimization, variance reduction
- ◇ communication/computation/memory efficient and scalable optimization algorithms
- ◇ collaborative learning (asynchronous, adversarial, local training, heterogeneity, etc.)
- ◇ model compression (knowledge distillation, pruning, sparse optimization, quantization)
- ◇ information theory (compression, encoding schemes, vector quantization)

## Education

- **Ph.D. in Mathematics** **Yerevan State University, Armenia**  
*Department of Mathematics, Chair of Theory of Functions*  
*Thesis: On estimates for maximal operators associated with tangential regions*  
Sep 2015–Jun 2018
- **M.Sc. in Mathematics (GPA 20/20)** **Yerevan State University, Armenia**  
*Department of Mathematics, Chair of Theory of Functions*  
*Thesis: Some generalizations of theorems of Fatou and Littlewood*  
Sep 2013–Jun 2015
- **B.Sc. in Mathematics (GPA 19.64/20)** **Yerevan State University, Armenia**  
*Department of Mathematics and Mechanics*  
*Thesis: Some properties of convergent and divergent convolution type operators*  
Sep 2009–Jun 2013

## Experience

- **Postdoctoral Research Fellow** **KAUST, Saudi Arabia**  
*King Abdullah University of Science & Technology (KAUST)*  
*Department of Applied Mathematics and Computational Sciences*  
*Optimization for Machine Learning, advisor: Prof. Peter Richtárik*  
Oct 2019–Oct 2022
- **Teaching Assistance**  
*Special Topics in Federated Learning (Spring 2020): Prof. Peter Richtárik*  
*Stochastic Gradient Descent Methods (Fall 2020): Prof. Peter Richtárik*

- Research Technician**

○ King Abdullah University of Science & Technology (KAUST)  
Computer, Electrical and Mathematical Sciences & Engineering (CEMSE) Division  
KAUST SRI, Center for Uncertainty Quantification in Computational Science and Engineering

▷ Computer Algebra for Differential Equations Nov 2016–Oct 2019  
Automation of symbolic PDE analysis with Wolfram Mathematica, advisor: Prof. Diogo Gomes

    - Finding conservation and dissipation laws for a system of time-dependent evolution equations
    - Symbolic methods for overdetermined systems of linear PDEs with free parameters

▷ [collaboration] Big Data Optimization in Machine Learning Jan 2019–Oct 2019  
Stochastic optimization methods, advisor: Prof. Peter Richtárik

**KAUST, Saudi Arabia**
  - Junior Researcher**

○ Institute of Mathematics of National Academy of Sciences  
Real Analysis Department, advisor: Prof. Grigori Karagulyan  
Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals

**Yerevan, Armenia**  
Aug 2014–June 2019
  - Search Engine Developer**

○ Teamable Software

Working extensively on data quality and all aspects of search engine in the product.  
Building intelligent, advanced and scalable search engine with Python and Apache Solr.

**Yerevan, Armenia**  
Apr 2014–Nov 2016
  - Assistant Teacher of Olympiad Mathematics**

○ Quantum School

**Yerevan, Armenia**  
2011–2012
- Internships and Visits**.....
- Participant**

○ Okinawa Institute of Science and Technology (OIST)  
Machine Learning Summer School (MLSS)  
Poster presentation on “Knowledge Distillation Performs Partial Variance Reduction”

**Okinawa, Japan**  
2024, Mar 4–16
  - Internship Student**

○ King Abdullah University of Science & Technology (KAUST)  
Computer, Electrical and Mathematical Science and Engineering (CEMSE) Division  
Automation of basic operations in analysis of PDEs using Wolfram Mathematica: variational derivative of a functional, integration by parts, generating polynomials with respect to certain symmetry groups and simplifying integral identities.

**KAUST, Saudi Arabia**  
2016, April–June
  - Visiting Student**

○ Hausdorff Research Institute for Mathematics (HIM)  
Winter School on Advances in Mathematics of Signal Processing

**Bonn, Germany**  
2016, Jan 11–15
  - Programming Intern**

○ Instigate Training Center, Instigate Mobile CJSC

**Yerevan, Armenia**  
Oct 2012–Jul 2013

## Publications

---

### Conference papers

- ✉ Thomas Robert, Mher Safaryan, Ionut-Vlad Modoranu, Dan Alistarh  
**LDAdam: Adaptive Optimization from Low-Dimensional Gradient Statistics**,  
International Conference on Learning Representations (ICLR) 2025  
(acceptance rate: 32.08%, total submissions: 11,565)

- ✉ Diyuan Wu, Ionut-Vlad Modoranu, Mher Safaryan, Denis Kuznedelev, Dan Alistarh  
**The Iterative Optimal Brain Surgeon: Faster Sparse Recovery by Leveraging Second-Order Information**  
*Conference on Neural Information Processing Systems (NeurIPS) 2024*  
 (acceptance rate: 25.8%, total submissions: 15,671)
- ✉ Ionut-Vlad Modoranu, Mher Safaryan, Grigory Malinovsky, Eldar Kurtic, Thomas Robert, Peter Richtárik, Dan Alistarh  
**MicroAdam: Accurate Adaptive Optimization with Low Space Overhead and Provable Convergence**  
*Conference on Neural Information Processing Systems (NeurIPS) 2024*  
 (acceptance rate: 25.8%, total submissions: 15,671)
- ✉ Rustem Islamov, Mher Safaryan, Dan Alistarh  
**AsGrad: A Sharp Unified Analysis of Asynchronous-SGD Algorithms**  
*International Conference on Artificial Intelligence and Statistics (AISTATS) 2024*  
 (acceptance rate: 27.5%, total submissions: 1980)
- ✉ Mher Safaryan, Alexandra Peste, Dan Alistarh  
**Knowledge Distillation Performs Partial Variance Reduction**  
*Conference on Neural Information Processing Systems (NeurIPS) 2023*  
 (acceptance rate: 26.1%, total submissions: 13,330)
- ✉ Bokun Wang, Mher Safaryan, Peter Richtárik  
**Theoretically Better and Numerically Faster Distributed Optimization with Smoothness-Aware Quantization Techniques**  
*Conference on Neural Information Processing Systems (NeurIPS) 2022*  
 (acceptance rate: 25.6%, total submissions: 10,411)
- ✉ Mher Safaryan, Rustem Islamov, Xun Qian, Peter Richtárik  
**FedNL: Making Newton-Type Methods Applicable to Federated Learning**  
*International Conference of Machine Learning (ICML) 2022*  
 (acceptance rate: 21.9%, total submissions: 5630)
- ✉ Xun Qian, Rustem Islamov, Mher Safaryan, Peter Richtárik  
**Basis Matters: Better Communication-Efficient Second Order Methods for Federated Learning**  
*International Conference on Artificial Intelligence and Statistics (AISTATS) 2022*  
 (acceptance rate 29%, total submissions: 1685)
- ✉ Mher Safaryan, Filip Hanzely, Peter Richtárik  
**Smoothness Matrices Beat Smoothness Constants: Better Communication Compression Techniques for Distributed Optimization**  
*Conference on Neural Information Processing Systems (NeurIPS) 2021*  
 (acceptance rate: 26%, total submissions: 9122)
- ✉ Mher Safaryan, Peter Richtárik  
**Stochastic Sign Descent Methods: New Algorithms and Better Theory**  
*International Conference of Machine Learning (ICML) 2021*  
 (acceptance rate: 21.5%, total submissions: 5513)

## Journal papers.....

- ✉ Arto Maranjyan, Mher Safaryan, Peter Richtárik  
**GradSkip: Communication-Accelerated Local Gradient Methods with Better Computational Complexity**  
Transactions on Machine Learning Research (TMLR), 2025
- ✉ Rustem Islamov, Xun Qian, Slavomír Hanzely, Mher Safaryan, Peter Richtárik  
**Distributed Newton-Type Methods with Communication Compression and Bernoulli Aggregation**  
Transactions on Machine Learning Research (TMLR), 2023
- ✉ Aleksandr Beznosikov, Samuel Horváth, Peter Richtárik, Mher Safaryan  
**On Biased Compression for Distributed Learning**  
*Journal of Machine Learning Research (JMLR)*, 2023
- ✉ Mher Safaryan, Egor Shulgin, Peter Richtárik  
**Uncertainty Principle for Communication Compression in Distributed and Federated Learning and the Search for an Optimal Compressor**  
*Information and Inference: A Journal of the IMA*, 2021
- ✉ Mher Safaryan  
**On Generalizations of Fatou's Theorem in  $L^p$  for Convolution Integrals with General Kernels**  
*The Journal of Geometric Analysis*, Volume 31, pp. 3280–3299, 2021
- ✉ Mher Safaryan  
**On an equivalency of rare differentiation bases of rectangles**  
*Journal of Contemporary Mathematical Analysis*, Volume 53(1), pp. 57-61, 2018
- ✉ Grigori Karagulyan, Mher Safaryan  
**On a theorem of Littlewood**  
*Hokkaido Mathematical Journal*, Volume 46(1), pp. 87-106, 2017
- ✉ Grigori Karagulyan, Davit Karagulyan, Mher Safaryan  
**On an equivalency of differentiation basis of dyadic rectangles**  
*Colloquium Mathematicum*, Volume 146, pp. 295-307, 2017
- ✉ Grigori Karagulyan, Mher Safaryan  
**On generalizations of Fatou's theorem for the integrals with general kernels**  
*The Journal of Geometric Analysis*, Volume 25(3), pp. 1459-1475, 2014
- ✉ Yuri Movsisyan, Sergey Davidov, Mher Safaryan  
**Construction of free  $g$ -dimonoids**  
*Algebra and Discrete Mathematics*, Volume 18(1), pp. 138–148, 2014

## Preprints.....

- ✉ Andrei Panferov, Alexandra Volkova, Ionut-Vlad Modoranu, Vage Egiazarian, Mher Safaryan, Dan Alistarh  
**Unified Scaling Laws for Compressed Representations**, arXiv:2506.01863, 2025
- ✉ Alex Iacob, Lorenzo Sani, Mher Safaryan, Paris Giampouras, Samuel Horváth, Andrej Jovanovic, Meghdad Kurmanji, Preslav Aleksandrov, William F. Shen, Xinchu Qiu, Nicholas D. Lane  
**DES-LOC: Desynced Low Communication Adaptive Optimizers for Training Foundation Models**, arXiv:2505.22549, 2025

- ✉ Ionut-Vlad Modoranu, Mher Safaryan, Erik Schultheis, Dan Alistarh  
**SVD-Free Low-Rank Adaptive Gradient Optimization for Large Language Models**, arXiv:2505.17967, 2025
- ✉ Alyazeed Albasyoni, Mher Safaryan, Laurent Condat, Peter Richtárik  
**Optimal Gradient Compression for Distributed and Federated Learning**, arXiv:2010.03246, 2020
- ✉ Diogo A. Gomes, Mher Safaryan, Ricardo de Lima Ribeiro, Mohammed Sayyari  
**A Surprisingly Effective Algorithm for the Simplification of Integrals and Sums Arising in the Partial Differential Equations and Numerical Methods**, KAUST Repository, 2020

## Awards

---

- ✉ **Top Reviewer Award at NeurIPS 2022**  
  - Marie Skłodowska-Curie Fellowship **IST Austria**  
MSCA COFUND IST-BRIDGE *Nov 2022 - Apr 2025*
- ✉ **Top Reviewer Award at AISTATS 2022**  
  - Nominal Fellowship Djrbashian** **Yerevan State University**  
*Given to one student for excellence and research* *Spring 2015*
  - Nominal Fellowship Mergelyan** **Yerevan State University**  
*Given to one student for excellence and research* *Spring 2014*
  - YSU bronze medal** **Yerevan State University**  
*YSU best student competition, Department of Mathematics* *2013*
- ✉ **Third Prize (2011, 2013), Honorable mention (2012)** **American University in Bulgaria**  
*International Mathematics Competition (IMC) for University Students* *2011-2013*

## Teaching

---

- Set Theory** **Yerevan State University**  
*Lecturer (informal mini-course)* *2015*
  - Calculus** **Yerevan State University**  
*Teaching Assistant* *2016 Jan-Apr*
  - Stochastic Gradient Descent Methods** **KAUST**  
*Teaching Assistant* *Fall 2020*
  - Special Topics in Federated Learning** **KAUST**  
*Teaching Assistant* *Spring 2020*

## Co-supervision of Master's Theses

---

Artavazd Maranjyan, Yerevan State University, Armenia, (Jan 2022 – Oct 2022).....

- ✉ Arto Maranjyan, Mher Safaryan, Peter Richtárik  
**GradSkip: Communication-Accelerated Local Gradient Methods with Better Computational Complexity**  
Transactions on Machine Learning Research (TMLR), 2025  
*Currently: PhD student at KAUST, Saudi Arabia (August 2023 - present)*

Rustem Islamov, Institut Polytechnique de Paris, France (Apr 2023 – Sep 2023).....

✉ Rustem Islamov, Mher Safaryan, Dan Alistarh

**AsGrad: A Sharp Unified Analysis of Asynchronous-SGD Algorithms**

International Conference on Artificial Intelligence and Statistics (AISTATS), 2024

*Currently: PhD student at The University of Basel, Switzerland (October 2023 - present)*

## Reviewing

---

- ELLIS PhD Programm (Evaluator): 2023, 2024.
- International Conference on Learning Representations (ICLR): 2020, 2021, 2022
- International Conference on Machine Learning (ICML): 2020, 2021, 2023
- Conference on Neural Information Processing Systems (NeurIPS): 2020, 2021, 2022, 2023
- International Conference on Artificial Intelligence and Statistics (AISTATS): 2022, 2024
- ICML 2021 Workshop on Federated Learning for User Privacy and Data Confidentiality (FL-ICML'21)
- NeurIPS OPT Workshop on Optimization for Machine Learning: 2023, 2024
- International Symposium on Distributed Computing (DISC): 2024
- Journal of Machine Learning Research (JMLR)
- IEEE Transactions on Information Theory (TIT)
- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)

## Selected Talks

---

- International Conference on Machine Learning, July 17-23, 2022, Baltimore, Maryland USA (**spotlight**)
- Mohamed bin Zayed University of AI (MBZUAI), guest lecture, April 26, 2022 (virtual) (**invited**)
- Toyota Technological Institute at Chicago (TTIC) reserach seminar, April 6, 2022 (virtual) (**invited**)
- Rising Stars in AI Symposium 2022, March 13-15, KAUST (**invited**)
- Conference on Neural Information Processing Systems, December 6-14, 2021 (virtual)
- Federated Learning One World (FLOW) Seminar, August 4, 2021 (virtual)
- International Conference on Machine Learning, July 18-24, 2021 (virtual)
- ICLR Distributed and Private Machine Learning (DPML) Workshop, 2021 (virtual)
- NeurIPS International Workshop on Scalability, Privacy, and Security in Federated Learning (SpicyFL) 2020 (virtual)
- Federated Learning One World (FLOW) Seminar, November 25, 2020 (virtual)
- YerevaNN Machine Learning Research Seminar, April 25, 2020 (virtual)