Konstantin Mishchenko

http://konstmish.com/ konsta.mish@gmail.com Last update: October 24, 2024 Google Scholar

Positions

- 2024-now AI Research Scientist at Meta, Paris, France
- 2023–2024 Research Scientist at Samsung AI Center, Cambridge, UK
- 2021–2022 Postdoctoral Researcher at Inria Sierra, Paris, France
 - 2020 Research Intern at Google Brain, remote
 - 2018 Applied Scientist Intern at Amazon, Seattle, US
 - 2016 C++ Software Engineering Intern, AIM Tech (High Frequency Trading), Moscow, Russia

Education

- 2017-2021 KAUST, PhD in Computer Science, Adviser: Peter Richtárik
- 2016–2017 ENS Cachan and Paris-Dauphine, MSc in Machine Learning
- 2012–2016 Moscow Institute of Physics and Technology, BSc in Computer Science and Physics

Achievements and awards

- 2023 ICML Outstanding Paper Award
- 2022 ICML Outstanding Reviewer (Top 10%)
- 2022 ICLR Highlighted Reviewer
- 2021 A Rising Star in Data Science by the University of Chicago (link)
- 2021 NeurIPS 2021 Outstanding Reviewer (Top 8%)
- 2021 ICML 2021 Best Reviewer (Top 10%)
- 2021 ICLR 2021 Outstanding Reviewer
- 2020 ICML 2020 Top Reviewer
- 2020 One of 12 **Outstanding Program Committee members** for AAAI 2020 selected from >6000 reviewers
- 2019 NeurIPS 2019 Best Reviewer Award
- 2018 71st place worldwide in IEEEXtreme team programming competition
- 2017-2020 Dean's Award (\$5000 annually for 3 years), given to a few top students accepted to KAUST
 - 2017 1st place in the Plume Labs machine learning competition on air pollution prediction
 - 2017 123rd place worldwide in IEEEXtreme team programming competition
- 2016-2017 Paris Graduate School of Mathematics fellowship (awarded to 24 people from 12 countries)
 - 2015 1st prize in Higher School of Economics Olympiad on Applied Math and Informatics
 - 2012 **Top-1** (max score) at the National Exam in math (only 54 participants out of >800k scored max)
 - 2012 1st prize in Moscow Mathematical Olympiad

Papers

Conference papers

- A. Defazio, X. Yang, H. Mehta, K. Mishchenko, A. Khaled, A. Cutkosky The Road Less Scheduled NeurIPS, 2024
- 17. Y. Malitsky, K. Mishchenko

Adaptive proximal gradient method for convex optimization NeurIPS, 2024 16. K. Mishchenko, A. Defazio Prodigy: An Expeditiously Adaptive Parameter-Free Learner **ICML**, 2024 15. A. Khaled, K. Mishchenko, C. Jin DoWG Unleashed: An Efficient Universal Parameter-Free Gradient Descent Method NeurIPS, 2023 14. A. Defazio, K. Mishchenko Learning-Rate-Free Learning by D-Adaptation **ICML**, 2023 **Outstanding Paper Award** 13. B. Woodoworth, K. Mishchenko, F. Bach Two Losses Are Better Than One: Faster Optimization Using a Cheaper Proxy **ICML**, 2023 12. K. Mishchenko, F. Bach, M. Even, B. Woodworth Asynchronous SGD Beats Minibatch SGD Under Arbitrary Delays NeurIPS, 2022 11. K. Mishchenko, G. Malinovsky, S. Stich, P. Richtárik ProxSkip: Yes! Local Gradient Steps Provably Lead to Communication Acceleration! Finally! **ICML**, 2022 10. K. Mishchenko, A. Khaled, P. Richtárik Proximal and Federated Random Reshuffling **ICML**, 2022 9. K. Mishchenko, B. Wang, D. Kovalev, P. Richtárik IntSGD: Adaptive Floatless Compression of Stochastic Gradients ICLR, Spotlight, 2022 8. K. Mishchenko, A. Khaled, P. Richtárik Random Reshuffling: Simple Analysis with Vast Improvements NeurIPS, 2020 7. Y. Malitsky, K. Mishchenko Adaptive Gradient Descent Without Descent **ICML**, 2020 6. K. Mishchenko, F. Hanzely, P. Richtárik 99% of Worker-Master Communication in Distributed Optimization Is Not Needed **UAI**. 2020 5. K. Mishchenko, D. Kovalev, E. Shulgin, Y. Malitsky, P. Richtárik Revisiting Stochastic Extragradient **AISTATS**, 2020 4. A. Khaled, K. Mishchenko, P. Richtárik Tighter Theory for Local SGD on Identical and Heterogeneous Data **AISTATS**, 2020 3. S. Soori, K. Mishchenko, A. Mokhtari, M. Dehnavi, M. Gürbüzbalaban DAve-QN: A Distributed Averaged Quasi-Newton Method with Local Superlinear Convergence Rate **AISTATS**, 2020 2. F. Hanzely, K. Mishchenko, P. Richtárik SEGA: Variance Reduction via Gradient Sketching **NeurIPS**, 2018 1. K. Mishchenko, F. lutzeler, J. Malick, M.-R. Amini A Delay-Tolerant Proximal-Gradient Algorithm for Distributed Learning **ICML**, 2018

Journal papers

- 6. N. Doikov, K. Mishchenko, Yu. Nesterov Super-Universal Regularized Newton Method SIAM Journal on Optimization (**SIOPT**)
- 5. K. Mishchenko Regularized Newton Method with Global $O(1/k^2)$ Convergence SIAM Journal on Optimization (**SIOPT**)
- 4. S. Horváth, D. Kovalev, K. Mishchenko, S. Stich, P. Richtárik Stochastic Distributed Learning with Gradient Quantization and Double Variance Reduction Optimization Methods and Software
- A. Salim, L. Condat, K. Mishchenko, P. Richtárik Dualize, Split, Randomize: Fast Nonsmooth Optimization Algorithms Journal of Optimization Theory and Applications (JOTA)
- 2. K. Mishchenko, E. Gorbunov, M. Takáč, P. Richtárik Distributed Learning with Compressed Gradient Differences Optimization Methods and Software
- K. Mishchenko, F. lutzeler, J. Malick A Distributed Flexible Delay-tolerant Proximal Gradient Algorithm SIAM Journal on Optimization (SIOPT)

Workshop papers

 G. Malinovsky, K. Mishchenko, P. Richtárik Server-Side Stepsizes and Sampling Without Replacement Provably Help in Federated Optimization

DistributedML, 2023

- A. Khaled, K. Mishchenko, P. Richtárik Better Communication Complexity for Local SGD NeurIPS, Oral at Federated Learning for Data Privacy and Confidentiality workshop, 2019
- D. Kovalev, K. Mishchenko, P. Richtárik Stochastic Newton and Cubic Newton Methods with Simple Local Linear-Quadratic Rates NeurIPS, Spotlight at Beyond First-Order Methods in ML workshop, 2019

K. Mishchenko Sinkhorn Algorithm as a Special Case of Stochastic Mirror Descent NeurIPS, Optimal Transport & Machine learning workshop, 2019

- A. Khaled, K. Mishchenko, P. Richtárik First Analysis of Local GD on Heterogeneous Data NeurIPS, Federated Learning for Data Privacy and Confidentiality workshop, 2019
- K. Mishchenko, M. Montgomery, F. Vaggi A Self-supervised Approach to Hierarchical Forecasting with Applications to Groupwise Synthetic Controls ICML, Time Series workshop, 2019

Preprints

- H. Chen, W. Luk, K. F. C. Yiu, R. Li, K. Mishchenko, S. I. Venieris, H. Fan Hardware-Aware Parallel Prompt Decoding for Memory-Efficient Acceleration of LLM Inference arXiv:2405.18628
- A. Defazio, A. Cutkosky, H. Mehta, K. Mishchenko When, Why and How Much? Adaptive Learning Rate Scheduling by Refinement arXiv:2310.07831
- K. Mishchenko, R. Islamov, E. Gorbunov, S. Horváth Partially Personalized Federated Learning: Breaking the Curse of Data Heterogeneity arXiv:2305.18285
- S. Horváth, K. Mishchenko, P. Richtárik Adaptive Learning Rates for Faster Stochastic Gradient Methods arXiv:2208.05287

- X. Qian, A. Sailanbayev, K. Mishchenko, P. Richtárik MISO is Making a Comeback With Better Proofs and Rates arXiv:1906.01474
- K. Mishchenko, P. Richtárik A Stochastic Decoupling Method for Minimizing the Sum of Smooth and Non-Smooth Functions arXiv:1905.11535
- K. Mishchenko, P. Richtárik A Stochastic Penalty Model for Convex and Nonconvex Optimization with Big Constraints arXiv:1810.13387

Reviewing and serving as Program Committee Member

2023-Now Transactions on Machine Learning Research (TMLR), Action Editor

- 2023 Mathematical Programming, Reviewer
- 2023 International Conference on Machine Learning (ICML), Reviewer
- 2023 International Conference on Learning Representations (ICLR), Reviewer
- 2022 Conference on Neural Information Processing Systems (NeurIPS), Program Committee Member
- 2022 Asian Conference on Machine Learning (**ACML**), **Senior** Program Committee member (Area Chair)
- 2022 Transactions on Machine Learning Research (TMLR), Reviewer
- 2022 International Conference on Machine Learning (ICML), Reviewer
- 2022 Journal of Machine Learning Research (JMLR), Reviewer
- 2022 International Conference on Learning Representations (ICLR), Reviewer
- 2021 Journal of Optimization Theory and Applications (JOTA), Reviewer
- 2021 NeurIPS New Frontiers in Federated Learning: Privacy, Fairness, Robustness, Personalization and Data Ownership Workshop, Program Committee Member
- 2021 NeurIPS Optimization for Machine Learning Workshop, Program Committee Member
- 2021 NeurIPS "I (Still) Can't Believe It's Not Better!" Workshop, Reviewer
- 2021 IEEE Transactions on Pattern Analysis and Machine Intelligence, Reviewer
- 2021 Conference on Neural Information Processing Systems (NeurIPS), Program Committee Member
- 2021 International Conference on Machine Learning (ICML), Expert Reviewer
- 2021 International Conference on Artificial Intelligence and Statistics (AISTATS), Reviewer
- 2021 International Conference on Learning Representations (ICLR), Reviewer
- 2020 Conference on Neural Information Processing Systems (NeurIPS), Program Committee Member
- 2021 Journal of Machine Learning Research (JMLR), Reviewer
- 2020 Conference on Uncertainty in Artificial Intelligence (UAI), Program Committee Member
- 2020 **IJCAI-PRICAI**, **Workshop** on Federated Learning for User Privacy and Data Confidentiality, Program Committee member
- 2020 (\times 2) Journal of Machine Learning Research (JMLR), Reviewer
 - 2020 International Conference on Machine Learning (ICML), Program Committee Member (top 33% ranking reviewer)
 - 2020 International Joint Conference on Artificial Intelligence (IJCAI-PRICAI), Program Committee Member
 - 2020 NeurIPS 2019 Reproducibility Challenge, Reviewer
 - 2019 Journal of Optimization Theory and Applications (JOTA), Reviewer
 - 2019 Bridging Game Theory and Deep Learning (NeurIPS Workshop), Reviewer
 - 2019 AAAI Conference on Artificial Intelligence (AAAI), Program Committee Member, One of 12 outstanding PC members

- 2019 Conference on Neural Information Processing Systems (**NeurIPS**), Program Committee Member, Best Reviewer Award
- 2019 Mathematical Programming, Reviewer
- 2019 Conference on Uncertainty in Artificial Intelligence (UAI), Program Committee Member
- 2019 International Conference on Machine Learning (ICML), Program Committee Member

Conference organization

- 2020 SIAM Conference on Optimization, organizer of 2 minisymposia (event cancelled due to COVID)
- 2020 SIAM Conference on Mathematics of Data Science, organizer of session on Optimization for Deep Learning
- 2019 International Conference on Continuous Optimization, organizer of 3 sessions
- 2018 Informs Optimization Society Meeting, organizer of 1 session

Talks

- 10/2024 Imperial College London Optimization and Control Seminar, UK
- 03/2024 Information Sciences and Systems, Princeton, US
- 02/2024 Invited lecture at Cambridge University, UK
- 06/2023 MIT Operations Research Center, Online
- 05/2023 Google Brain Seminar, Online
- 03/2023 AIDRC Seminar, AIDRC Seminar, UAE (remote)
- 12/2022 NeurIPS Workshop on Federated Learning: Recent Advances and New Challenges, US (remote)
- 07/2022 EUROPT Workshop on Advances in Continuous Optimization, Portugal
- 05/2022 Google Deepmind London, UK
- 12/2021 One-World Optimization Seminar, Online
- 08/2021 Modeling and Optimization: Theory and Applications (MOPTA), Online
- 07/2021 Beyond First-Order Methods in ML Systems Workshop at ICML, Online
- 03/2021 Federated Learning One World Seminar (FLOW), Online
- 11/2020 INFORMS Annual Meeting, Online
- 10/2020 Google Brain Montreal, Online
- 10/2020 JetBrains Research Machine Learning Seminar, Online
- 06/2020 All-Russian Optimization Seminar, Online
- 03/2020 Statistics Department of London School of Economics, UK
- 02/2020 Imperial College of London Reading Group, UK
- 02/2020 Oxford Data Science seminar, UK
- 02/2020 Gatsby Unit, University College of London, UK
- 02/2020 Google Deepmind London, UK
- 02/2020 Facebook Artificial Intelligence Research New York, USA
- 01/2020 Sierra team at Inria, France
- 12/2019 LIONS group at EPFL, Switzerland
- 10/2019 Boris Polyak's seminar on theory of automatic control, Institute for Control Sciences, Russia
- 10/2019 Seminar on applied mathematics, Moscow Institute of Physics and Technology, Russia
- 10/2019 Modern optimization methods seminar, Moscow Institute of Physics and Technology, Russia
- 08/2019 International Conference on Continuous Optimization, TU Berlin, Germany
- 06/2019 Numerical Analysis seminar, Bath University, UK
- 03/2019 Machine Learning and Optimization Laboratory seminar, EPFL, Switzerland

- 11/2018 Microsoft Research Seattle, USA
- 07/2018 International Symposium on Mathematical Programming, Bordeaux, France
- 03/2018 Informs Optimization Society Meeting, University of Colorado, USA
- 10/2017 Optimization at Work, Moscow Institute of Physics and Technology, Russia