

Eduard GORBUNOV

PERSONAL DATA

PLACE AND DATE OF BIRTH: Rybinsk, Russia | 22 November 1996
EMAIL: eduard.gorbunov@mbzuai.ac.ae
WEBSITE: eduardgorbunov.github.io
GOOGLE SCHOLAR: <https://scholar.google.com/citations?user=QPVriwoAAAAJ&hl=en>
WoS RESEARCHER ID: U-1740-2019
SCOPUS AUTHOR ID: 57207780972
ORCID ID: 0000-0002-3370-4130
DBLP: <https://dblp.org/pid/215/5512.html>

WORK EXPERIENCE

APRIL 2024 – NOW	Research Scientist at the ML department , MBZUAI , Abu Dhabi, UAE (hosted by Samuel Horváth and Martin Takác)
SEPTEMBER 2022 – MARCH 2024	Postdoctoral Fellow at the ML department , MBZUAI , Abu Dhabi, UAE (hosted by Samuel Horváth and Martin Takác)
MAY 2020 – AUGUST 2022	Junior Researcher at Laboratory of Advanced Combinatorics and Network Applications , MIPT , Moscow
FEBRUARY 2022 – MAY 2022	Research Consultant at Mila , Montreal (in the group of Gauthier Gidel)
SEPTEMBER 2020 – JANUARY 2022	Junior Researcher at Yandex.Research-MIPT Laboratory , Moscow
MAY 2020 – DECEMBER 2021	Research Assistant at International Laboratory of Stochastic Algorithms and High-Dimensional Inference , HSE , Moscow
NOVEMBER 2019 – APRIL 2020	Junior Researcher at IITP RAS , Moscow
FEBRUARY 2020 – DECEMBER 2020	Junior Researcher at Joint Research Laboratory of Applied Mathematics , RANEPA-MIPT , Moscow
FEBRUARY 2020 – DECEMBER 2020	Junior Researcher at Laboratory of Numerical Methods of Applied Structural Optimization , MIPT , Moscow
AUGUST 2019 – JULY 2019	Researcher at Huawei-MIPT group, Moscow (research, Python)
MAY 2019 – AUGUST 2019	Intern at Huawei Media Lab , Moscow (research, C++)
AUGUST 2017 – OCTOBER 2019	Researcher at Peter Richtárik's Group , MIPT , Moscow

EDUCATION

SEPTEMBER 2020 – DECEMBER 2021	PhD in Computer Science Moscow Institute of Physics and Technology, Moscow Thesis: " Distributed and Stochastic Optimization Methods with Gradient Compression and Local Steps " Advisors: Alexander GASNIKOV and Peter RICHTÁRIK LINKS: slides and video of the defense
SEPTEMBER 2018 – JULY 2020	Master of Science in Applied Mathematics and Physics Moscow Institute of Physics and Technology, Moscow Thesis: " Derivative-free and stochastic optimization methods, decentralized distributed optimization " Advisor: Alexander GASNIKOV
SEPTEMBER 2014 – JULY 2018	Bachelor of Science in Applied Mathematics and Physics Moscow Institute of Physics and Technology, Moscow Thesis: " Accelerated Directional Searches and Gradient-Free Methods with non-Euclidean prox-structure " Advisor: Alexander GASNIKOV

RESEARCH INTERESTS

Optimization, Machine Learning, Federated Learning, Variational Inequalities, Derivative-Free Optimization, Randomized Algorithms

RESEARCH VISITS AND INTERNSHIPS

- **8 June 2021 – 30 September 2021**, [Mila](#), Université de Montréal. Internship in the group of [Gauthier Gidel](#)
- **1 September 2020 – 28 February 2021**, Visual Computing Center, KAUST, Thuwal, Saudi Arabia. I remotely worked in the group of [P. Richtárik](#)
- **2 February – 31 March 2020**, Visual Computing Center, KAUST, Thuwal, Saudi Arabia (worked with [P. Richtárik](#))
- **6 October – 26 October 2019**, [SIERRA](#), [INRIA](#), Paris, France (worked with [A. Taylor](#))
- **13 January – 24 February 2019**, Visual Computing Center, KAUST, Thuwal, Saudi Arabia (worked with [P. Richtárik](#))
- **14 January – 8 February 2018**, Visual Computing Center, KAUST, Thuwal, Saudi Arabia (worked with [P. Richtárik](#))

SCHOLARSHIPS, HONORS AND AWARDS

Selected Awards

- **10 April, 2019.** [The Ilya Segalovich Award – Yandex scientific scholarship](#), highly selective: 9 winners from Russia, Belarus and Kazakhstan (350,000 Russian rubles, internship offer at Yandex.Research, travel grant to attend one international conference; news about award: <https://nplus1.ru/news/2019/04/10/ya-awards>)
- **Outstanding reviewer awards:** NeurIPS (2022, 2021, 2020), ICML (2022, 2021), ICLR (2021)

Other Awards (in the reverse chronological order)

- **September 2021 - January 2021.** [A. M. Raigorodskii Scholarship for Contribution to the Development of Numerical Optimization Methods](#) (main scholarship)
- **February 2021 - June 2021.** [A. M. Raigorodskii Scholarship for Contribution to the Development of Numerical Optimization Methods](#) (main scholarship)
- **February 2020 - June 2020.** Increased State Academic Scholarship for 4 year bachelor and master students at MIPT for scientific achievements (14,000 Russian rubles per month instead of the regular scholarship)
- **15 January, 2020.** [Huawei scholarship for bachelor and master students at MIPT](#) (125,000 Russian rubles)
- **September 2019 - January 2020.** Increased State Academic Scholarship for 4 year bachelor and master students at MIPT for scientific achievements (10,000 Russian rubles per month in addition to the regular scholarship)
- **February 2019 - June 2019.** Increased State Academic Scholarship for 4 year bachelor and master students at MIPT for scientific achievements (10,000 Russian rubles per month in addition to the regular scholarship)

- **September 2018 - January 2019.** Increased State Academic Scholarship for 4 year bachelor and master students at MIPT for scientific achievements (10,000 Russian rubles per month in addition to the regular scholarship)
- **February 2018 - June 2018.** Increased State Academic Scholarship for 4 year bachelor and master students at MIPT for scientific achievements (10,000 Russian rubles per month in addition to the regular scholarship)
- **September 2017 - January 2018.** Increased State Academic Scholarship for 4 year bachelor and master students at MIPT for scientific achievements (10,000 Russian rubles per month in addition to the regular scholarship)
- **November 2017.** Diploma of winner of the Section of Information Transmission Problems, Data Analysis and Optimization at 60th Scientific Conference of MIPT
- **May 2017.** [Third Prize at MIPT's Student Olympiad in Mathematics](#)
- **March 2017.** First Prize at MIPT's Team Mathematical Tournament
- **September 2016 - June 2017.** Abramov scholarship for 1-3 year bachelor students with the best grades at MIPT (12,000 Russian rubles per month)
- **December 2015.** [Third Prize at MIPT's Student Olympiad in Mathematics](#)
- **February 2015 - June 2015.** Abramov scholarship for 1-3 year bachelor students with the best grades at MIPT (12,000 Russian rubles per month)
- **April 2014.** Participant of Final Round of All-Russian Mathematical Olympiad ([scored points: 28 out of 56, 59th place](#))

TEACHING

- **Co-creator and lecturer of the course** “Optimization Methods for Machine Learning” in [MADE](#), Mail.ru Group (Spring 2020, Spring 2021, Spring 2022) and [MIPT](#) (Fall 2020)
- **Teaching assistant for the courses**
 - Fall 2022, Spring 2023: ML712: Distributed and Federated Learning ([MBZUAI](#)): created and presented lectures on 2 topics (out of 8 topics), replied to the students questions, mentored the projects, checked and created quizzes
 - Spring 2019: [Algorithms and Models of Computation](#)
 - Fall 2018: [Probability Theory](#)
 - Spring 2018: [Algorithms and Models of Computation](#)

SUPERVISION & MENTORING

I was supervising/mentoring the work of the following students/interns and research assistants/postdocs on the listed projects (projects are listed in the reverse chronological order).

- [Yury Demidovich](#) (Postdoc), [Petr Ostroukhov](#) (Research assistant), [Grigory Malinovsky](#) (PhD student). Project on the convergence analysis of methods with clipping, random reshuffling, and local steps under the generalized smoothness assumption. Led to the preprint: Methods with Local Steps and Random Reshuffling for Generally Smooth Non-Convex Federated Optimization, [arXiv:2412.02781](#)
- [Nazarii Tupitsa](#) (Research assistant), [Sayantan Choudhury](#) (PhD student), Alen Aliev (Master student). Project on the convergence of different optimization methods under generalized smoothness assumptions. Led to the preprint: Methods for Convex (L_0, L_1) -Smooth Optimization: Clipping, Acceleration, and Adaptivity, [arXiv:2409.14989](#)

- [Viktor Moskvoretskii](#) (Master student), [Nazarii Tupitsa](#) (Research assistant). Project on a new way of training LLMs for low-resource languages. Led to the paper accepted to [EMNLP 2024 \(Findings\)](#): Low-Resource Machine Translation through the Lens of Personalized Federated Learning, [arXiv:2406.12564](#)
- [Savelii Chezhegov](#) (PhD student), Yaroslav Klyukin (Bachelor student), [Andrei Semenov](#) (Bachelor student). Project on the high-probability convergence of versions of AdaGrad and Adam with gradient clipping. Led to the preprint: Gradient Clipping Improves AdaGrad when the Noise Is Heavy-Tailed, [arXiv:2406.04443](#)
- [Sayantan Choudhury](#) (PhD student), [Nazarii Tupitsa](#) (Research assistant). Project on scale-invariant version of AdaGrad – led to the paper accepted to [NeurIPS 2024](#): Remove that Square Root: A New Efficient Scale-Invariant Version of AdaGrad, [arXiv:2403.02648](#)
- [Nazarii Tupitsa](#) (Research assistant). Project on Federated Learning – led to the preprint: Federated Learning Can Find Friends That Are Advantageous, [arXiv:2402.05050](#)
- [Nickolay Kutuzov](#) (Bachelor student). Supervision of the work on numerical experiments for a project on high-probability convergence under heavy-tailed noise – led to the paper accepted to [AISTATS 2024](#): Breaking the Heavy-Tailed Noise Barrier in Stochastic Optimization Problems, [arXiv:2311.04161](#)
- [Grigory Malinovsky](#) (PhD student). Project on Byzantine-robust learning with partial participation – led to the paper accepted to [NeurIPS 2024](#): Byzantine Robustness and Partial Participation Can Be Achieved At Once: Just Clip Gradient Differences, [arXiv:2311.14127](#)
- Nikita Kornilov (Master student). Project on relative error in gradients – led to the preprint: Intermediate Gradient Methods with Relative Inexactness, [arXiv:2310.00506](#)
- [Nazarii Tupitsa](#) (Research assistant), [Abdulla Jasem Almansoori](#) (PhD student), Yanlin Wu (Master student). Project on Byzantine-robust methods for distributed variational inequalities – led to the paper accepted to [NeurIPS 2023](#): Byzantine-Tolerant Methods for Distributed Variational Inequalities, [arXiv:2311.04611](#)
- Nikita Fedin (Master student). Project on Byzantine-robust learning with variance reduction – led to the paper accepted to [MOTOR 2023](#): Byzantine-Robust Loopless Stochastic Variance-Reduced Gradient, [arXiv:2303.04560](#)
- [Abdurakhmon Sadiev](#) (PhD student)
 1. Project on high-probability convergence without bounded variance assumption – led to the paper accepted to [ICML 2023](#): High-Probability Bounds for Stochastic Optimization and Variational Inequalities: the Case of Unbounded Variance, [arXiv:2302.00999](#)
 2. Project on high-probability convergence for composite and distributed optimization problems – led to the paper accepted to [ICML 2024](#): High-Probability Convergence for Composite and Distributed Stochastic Minimization and Variational Inequalities with Heavy-Tailed Noise, [arXiv:2310.01860](#)
- [Konstantin Burlachenko](#) (PhD student). Supervision of the work on numerical experiments for a project on distributed non-convex optimization with unbiased compression – led to the paper accepted to [ICML 2021](#): MARINA: Faster Non-Convex Distributed Learning with Compression, [arXiv:2102.07845](#)
- [Dmitry Makarenko](#) (PhD student). Supervision of the work on numerical experiments for a project on distributed optimization with biased compression – led to the paper accepted as a [spotlight to NeurIPS 2020](#): Linearly Converging Error Compensated SGD, [arXiv:2010.12292](#)

- [Aleksandr Beznosikov](#) (Bachelor student). Project on zeroth-order composite optimization – led to the paper accepted to [IFAC 2020: Derivative-Free Method For Decentralized Distributed Non-Smooth Optimization](#), [arXiv:1911.10645](#)

PUBLICATIONS

Preprints

14. [A. Lobanov](#), [A. Gasnikov](#), [E. Gorbunov](#), [M. Takác](#). **Linear Convergence Rate in Convex Setup is Possible! Gradient Descent Method Variants under (L_0, L_1) -Smoothness**, [arXiv:2412.17050](#)
13. [Y. Demidovich](#), [P. Ostroukhov](#), [G. Malinovsky](#), [S. Horváth](#), [M. Takác](#), [P. Richtárik](#), [E. Gorbunov](#) (*equal contribution). **Methods with Local Steps and Random Reshuffling for Generally Smooth Non-Convex Federated Optimization**, [arXiv:2412.02781](#)
12. [K. Ponkshe*](#), [R. Singhal*](#), [E. Gorbunov](#), [A. Tumanov](#), [S. Horváth](#), [P. Vepakomma](#) (*equal contribution). **Initialization using Update Approximation is a Silver Bullet for Extremely Efficient Low-Rank Fine-Tuning**, [arXiv:2411.19557](#)
11. [S. Khirirat](#), [A. Sadiev](#), [A. Riabinin](#), [E. Gorbunov](#), [P. Richtárik](#). **Error Feedback under (L_0, L_1) -Smoothness: Normalization and Momentum**, [arXiv:2410.16871](#)
10. [E. Gorbunov*](#), [N. Tupitsa*](#), [S. Choudhury](#), [A. Aliev](#), [P. Richtárik](#), [S. Horváth](#), [M. Takác](#) (*equal contribution). **Methods for Convex (L_0, L_1) -Smooth Optimization: Clipping, Acceleration, and Adaptivity**, [arXiv:2409.14989](#)
9. [S. Chezhegov](#), [Y. Klyukin](#), [A. Semenov](#), [A. Beznosikov](#), [A. Gasnikov](#), [S. Horváth](#), [M. Takác](#), [E. Gorbunov](#). **Gradient Clipping Improves AdaGrad when the Noise Is Heavy-Tailed**, [arXiv:2406.04443](#)
8. [N. Tupitsa](#), [S. Horváth](#), [M. Takác](#), [E. Gorbunov](#). **Federated Learning Can Find Friends That Are Advantageous**, [arXiv:2402.05050](#)
7. [N. Kornilov](#), [Y. Dorn](#), [A. Lobanov](#), [N. Kutuzov](#), [I. Shibaev](#), [E. Gorbunov](#), [A. Gasnikov](#), [A. Nazin](#). **Zeroth-order Median Clipping for Non-Smooth Convex Optimization Problems with Heavy-tailed Symmetric Noise**, [arXiv:2402.02461](#)
6. [N. Kornilov](#), [E. Gorbunov](#), [M. Alkousa](#), [F. Stonyakin](#), [P. Dvurechensky](#), [A. Gasnikov](#). **Intermediate Gradient Methods with Relative Inexactness**, [arXiv:2310.00506](#)
5. [S. Khirirat](#), [E. Gorbunov](#), [S. Horváth](#), [R. Islamov](#), [F. Karray](#), [P. Richtárik](#). **Clip21: Error Feedback for Gradient Clipping**, [2305.18929](#)
4. [K. Mishchenko](#), [R. Islamov](#), [E. Gorbunov](#), [S. Horváth](#). **Partially Personalized Federated Learning: Breaking the Curse of Data Heterogeneity**, [2305.18285](#)
3. [E. Gorbunov](#). **Unified analysis of SGD-type methods**, [2303.16502](#)
2. [I. Fatkhullin](#), [I. Sokolov](#), [E. Gorbunov](#), [Z. Li](#), [P. Richtárik](#). **EF21 with Bells & Whistles: Practical Algorithmic Extensions of Modern Error Feedback**, [arXiv:2110.03294](#)
1. [E. Gorbunov](#), [D. Dvinskikh](#), [A. Gasnikov](#). **Optimal Decentralized Distributed Algorithms for Stochastic Convex Optimization**, [arXiv:1911.07363](#)

Conferences

33. [A. Agafonov](#), [P. Ostroukhov](#), [R. Mozhaev](#), [K. Yakovlev](#), [E. Gorbunov](#), [M. Takác](#), [A. Gasnikov](#), [D. Kamzolov](#). **Exploring Jacobian Inexactness in Second-Order Methods for Variational**

Inequalities: Lower Bounds, Optimal Algorithms and Quasi-Newton Approximations
NeurIPS 2024 (spotlight), [arXiv:2405.15990¹](#)

32. S. Choudhury, N. Tupitsa, N. Loizou, S. Horváth, M. Takác, E. Gorbunov. **Remove that Square Root: A New Efficient Scale-Invariant Version of AdaGrad**
NeurIPS 2024, [arXiv:2403.02648](#)
31. G. Malinovsky, P. Richtárik, S. Horváth, E. Gorbunov. **Byzantine Robustness and Partial Participation Can Be Achieved at Once: Just Clip Gradient Differences**
NeurIPS 2024, [arXiv:2311.14127](#)
30. A. Sadiev, G. Malinovsky, E. Gorbunov, I. Sokolov, A. Khaled, K. Burlachenko, P. Richtárik. **Federated Optimization Algorithms with Random Reshuffling and Gradient Compression**
NeurIPS 2024, [arXiv:2206.07021](#)
29. V. Moskvoretskii, N. Tupitsa, C. Biemann, S. Horváth, E. Gorbunov, I. Nikishina. **Low-Resource Machine Translation through the Lens of Personalized Federated Learning**
EMNLP 2024 (Findings), [arXiv:2406.12564](#)
28. E. Gorbunov, A. Sadiev, M. Danilova, S. Horváth, G. Gidel, P. Dvurechensky, A. Gasnikov, P. Richtárik. **High-Probability Convergence for Composite and Distributed Stochastic Minimization and Variational Inequalities with Heavy-Tailed Noise**
ICML 2024 (oral), [arXiv:2310.01860](#)
27. N. Puchkin*, E. Gorbunov*, N. Kutuzov, A. Gasnikov (*equal contribution). **Breaking the Heavy-Tailed Noise Barrier in Stochastic Optimization Problems**
AISTATS 2024, [arXiv:2311.04161](#)
26. A. Rammal, K. Gruntkowska, N. Fedin, E. Gorbunov, P. Richtárik. **Communication Compression for Byzantine Robust Learning: New Efficient Algorithms and Improved Rates**
AISTATS 2024, [arXiv:2310.09804](#)
25. N. Tupitsa, A. J. Almansoori, Y. Wu, M. Takác, K. Nandakumar, S. Horváth, E. Gorbunov. **Byzantine-Tolerant Methods for Distributed Variational Inequalities**
NeurIPS 2023, [arXiv:2311.04611](#)
24. N. Kornilov, O. Shamir, A. Lobanov, D. Dvinskikh, A. Gasnikov, I. Shibaev, E. Gorbunov, S. Horváth. **Accelerated Zeroth-order Method for Non-Smooth Stochastic Convex Optimization Problem with Infinite Variance**
NeurIPS 2023, [arXiv:2310.18763](#)
23. S. Choudhury, E. Gorbunov, N. Loizou. **Single-Call Stochastic Extragradient Methods for Structured Non-monotone Variational Inequalities: Improved Analysis under Weaker Conditions**
NeurIPS 2023, [arXiv:2302.14043](#)
22. A. Sadiev, M. Danilova, E. Gorbunov, S. Horváth, G. Gidel, P. Dvurechensky, A. Gasnikov, P. Richtárik. **High-Probability Bounds for Stochastic Optimization and Variational Inequalities: the Case of Unbounded Variance**
ICML 2023, [arXiv:2302.00999](#)
21. E. Gorbunov, A. Taylor, S. Horváth, G. Gidel. **Convergence of Proximal Point and Extragradient-Based Methods Beyond Monotonicity: the Case of Negative Comonotonicity**
ICML 2023, [arXiv:2210.13831](#)

¹The link to arXiv versions are provided for convenience.

20. N. Fedin, E. Gorbunov. Byzantine-Robust Loopless Stochastic Variance-Reduced Gradient
MOTOR 2023, 2303.04560
19. E. Gorbunov, S. Horváth, P. Richtárik, G. Gidel. Variance Reduction is an Antidote to Byzantines: Better Rates, Weaker Assumptions and Communication Compression as a Cherry on the Top
ICLR 2023, arXiv:2206.00529
18. A. Beznosikov*, E. Gorbunov*, H. Berard*, N. Loizou (*equal contribution). Stochastic Gradient Descent-Ascent: Unified Theory and New Efficient Methods
AISTATS 2023, arXiv:2202.07262
17. E. Gorbunov*, M. Danilova*, D. Dobre*, P. Dvurechensky, A. Gasnikov, G. Gidel (*equal contribution). Clipped Stochastic Methods for Variational Inequalities with Heavy-Tailed Noise
NeurIPS 2022, arXiv:2206.01095
16. E. Gorbunov, A. Taylor, G. Gidel. Last-Iterate Convergence of Optimistic Gradient Method for Monotone Variational Inequalities
NeurIPS 2022, arXiv:2205.08446
15. P. Richtárik, I. Sokolov, I. Fatkhullin, E. Gasanov, Z. Li, E. Gorbunov. 3PC: Three Point Compressors for Communication-Efficient Distributed Training and a Better Theory for Lazy Aggregation
ICML 2022, 2202.00998
14. E. Gorbunov*, A. Borzunov*, M. Diskin, M. Ryabinin (*equal contribution). Secure Distributed Training at Scale
ICML 2022, arXiv:2106.11257
13. M. Danilova, E. Gorbunov. Distributed Methods with Absolute Compression and Error Compensation
MOTOR 2022, arXiv:2203.02383
12. E. Gorbunov, H. Berard, G. Gidel, N. Loizou. Stochastic Extragradient: General Analysis and Improved Rates
AISTATS 2022, arXiv:2111.08611
11. E. Gorbunov, N. Loizou, G. Gidel. Extragradient Method: $\mathcal{O}(1/k)$ Last-Iterate Convergence for Monotone Variational Inequalities and Connections With Cocoercivity
AISTATS 2022, arXiv:2110.04261
10. M. Ryabinin*, E. Gorbunov*, V. Plokhotnyuk, G. Pekhimenko (*equal contribution). Moshpit SGD: Communication-Efficient Decentralized Training on Heterogeneous Unreliable Devices
NeurIPS 2021, arXiv:2103.03239
9. E. Gorbunov, K. Burlachenko, Z. Li, P. Richtárik. MARINA: Faster Non-Convex Distributed Learning with Compression
ICML 2021, arXiv:2102.07845
8. E. Gorbunov, F. Hanzely, P. Richtárik. Local SGD: Unified Theory and New Efficient Methods
AISTATS 2021, arXiv:2011.02828
7. E. Gorbunov, D. Kovalev, D. Makarenko, P. Richtárik. Linearly Converging Error Compensated SGD
NeurIPS 2020 (spotlight), arXiv:2010.12292

6. E. Gorbunov, M. Danilova, A. Gasnikov. **Stochastic Optimization with Heavy-Tailed Noise via Accelerated Gradient Clipping**
NeurIPS 2020, arXiv:2005.10785
5. E. Gorbunov, F. Hanzely, P. Richtárik. **A unified theory of SGD: variance reduction, sampling, quantization and coordinate descent**
AISTATS 2020, arXiv:1905.11261
4. A. Beznosikov, E. Gorbunov, A. Gasnikov. **Derivative-Free Method For Decentralized Distributed Non-Smooth Optimization**
IFAC 2020, arXiv:1911.10645
3. E. Gorbunov, A. Bibi, O. Sener, E. Bergou, P. Richtárik. **A Stochastic Derivative Free Optimization Method with Momentum**
ICLR 2020, arXiv:1905.13278
2. D. Kovalev, E. Gorbunov, E. Gasanov, P. Richtárik **Stochastic Spectral and Conjugate Descent Methods**
NeurIPS 2019, arXiv:1802.03703
1. D. Dvinskikh, E. Gorbunov, A. Gasnikov, P. Dvurechensky, César A. Uribe. **On Dual Approach for Distributed Stochastic Convex Optimization over Networks**
CDC 2019, arXiv:1903.09844

Journals

15. E. Gorbunov, M. Danilova, I. Shibaev, P. Dvurechensky, A. Gasnikov. **High-Probability Complexity Bounds for Non-smooth Stochastic Convex Optimization with Heavy-Tailed Noise**
Journal of Optimization Theory and Applications (JOTA) 2024, arXiv:2106.05958
14. K. Mishchenko, E. Gorbunov, M. Takáč, P. Richtárik. **Distributed Learning with Compressed Gradient Differences**
Optimization Methods and Software (OMS) 2024, arXiv:1901.09269
13. Y. Dorn, N. Kornilov, N. Kutuzov, A. Nazin, E. Gorbunov, A. Gasnikov. **Implicitly normalized forecaster with clipping for linear and non-linear heavy-tailed multi-armed bandits**
Computational Management Science 2024
12. A. Gasnikov, D. Dvinskikh, P. Dvurechensky, E. Gorbunov, A. Beznosikov, A. Lobanov. **Randomized gradient-free methods in convex optimization**
Encyclopedia of Optimization 2023, arXiv:2211.13566
11. A. Beznosikov, B. Polyak, E. Gorbunov, D. Kovalev, A. Gasnikov. **Smooth Monotone Stochastic Variational Inequalities and Saddle Point Problems – Survey**
European Mathematical Society Magazine 2023, arXiv:2208.13592
10. M. Danilova, P. Dvurechensky, A. Gasnikov, E. Gorbunov, S. Guminov, D. Kamzolov, I. Shibaev. **Recent Theoretical Advances in Non-Convex Optimization**
High-Dimensional Optimization and Probability 2022, arXiv:2012.06188
9. E. Gorbunov, A. Rogozin, A. Beznosikov, D. Dvinskikh, A. Gasnikov. **Recent theoretical advances in decentralized distributed convex optimization**
High-Dimensional Optimization and Probability 2022, arXiv:2011.13259
8. E. Gorbunov, P. Dvurechensky, A. Gasnikov. **An Accelerated Method for Derivative-Free Smooth Stochastic Convex Optimization**
SIAM Journal on Optimization (SIOPT) 2022, arXiv:1802.09022

7. A. Gasnikov, E. Gorbunov, S. Guz, E. Chernousova, M. Shirobokov, E. Shulgin. **Lecture Notes on Stochastic Processes (book, in Russian)**, ISBN 978-5-9710-9658-0, Moscow MIPT 2022, (available at [arXiv:1907.01060](https://arxiv.org/abs/1907.01060))
6. E. Bergou, E. Gorbunov, P. Richtárik. **Stochastic Three Points Method for Unconstrained Smooth Minimization**
SIAM Journal on Optimization (SIOPT) 2020, [arXiv:1902.03591](https://arxiv.org/abs/1902.03591)
5. P. Dvurechensky, E. Gorbunov, A. Gasnikov. **An Accelerated Directional Derivative Method for Smooth Stochastic Convex Optimization**
European Journal of Operational Research (EJOR) 2020, [arXiv:1804.02394](https://arxiv.org/abs/1804.02394)
4. E. Vorontsova, A. Gasnikov, E. Gorbunov, P. Dvurechensky. **Accelerated Gradient-Free Optimization Methods with a Non-Euclidean Proximal Operator**, Automation and Remote Control
Automation and Remote Control 2019
3. E. Vorontsova, A. Gasnikov, E. Gorbunov. **Accelerated Directional Search with non-Euclidean prox-structure**
Automation and Remote Control 2019, [arXiv:1710.00162](https://arxiv.org/abs/1710.00162)
2. E. Gorbunov, E. Vorontsova, A. Gasnikov. **On the upper bound for the mathematical expectation of the norm of a vector uniformly distributed on the sphere and the phenomenon of concentration of uniform measure on the sphere**
Mathematical Notes 2019, [arXiv:1804.03722](https://arxiv.org/abs/1804.03722)
1. A. Gasnikov, E. Gorbunov, D. Kovalev, A. Mohammed, E. Chernousova. **The global rate of convergence for optimal tensor methods in smooth convex optimization**
Computer Research and Modeling 2018, [arXiv:1809.00382](https://arxiv.org/abs/1809.00382)

TALKS AND POSTERS

56. 25 July, 2024. **ICML 2024**, Vienna, Austria. Oral Talk and Poster: "High-Probability Convergence for Composite and Distributed Stochastic Minimization and Variational Inequalities with Heavy-Tailed Noise". Links: [slides](#), [poster](#)
55. 3 July, 2024. **EURO 2024**, Copenhagen, Denmark. Talk: "Byzantine Robustness and Partial Participation Can Be Achieved Simultaneously: Just Clip Gradient Differences". Links: [slides](#)
54. 26 June, 2024. **EUROPT 2024**, Lund, Sweden. Talk: "Last-Iterate Convergence of Extragradient-Based Methods". Links: [slides](#)
53. 24 June, 2024. Invited talk at **INSAIT**: "Byzantine Robustness and Partial Participation Can Be Achieved Simultaneously: Just Clip Gradient Differences". Links: [slides](#)
52. 21 June, 2024. **Principles of Distributed Learning**, Nantes, France. Talk: "Byzantine Robustness and Partial Participation Can Be Achieved Simultaneously: Just Clip Gradient Differences". Links: [slides](#)
51. 29 May, 2024. **NETYS 2024**, online. **Keynote talk** "Byzantine Robustness and Partial Participation Can Be Achieved Simultaneously: Just Clip Gradient Differences". Links: [slides](#)
50. 3 May, 2024. **AISTATS 2024**, Valencia, Spain. Poster: "Communication Compression for Byzantine Robust Learning: New Efficient Algorithms and Improved Rates". Links: [poster](#)
49. 3 May, 2024. **AISTATS 2024**, Valencia, Spain. Poster: "Breaking the Heavy-Tailed Noise Barrier in Stochastic Optimization Problems". Links: [poster](#)

48. 7 February, 2024. [Federated Learning One-World Seminar](#), online. Talk “Variance Reduction is an Antidote to Byzantines: Better Rates, Weaker Assumptions and Communication Compression as a Cherry on the Top”. Links: [slides](#), [video](#)
47. 10 December - 16 December, 2023. [NeurIPS 2023](#), New Orleans, USA. Poster: “Single-Call Stochastic Extragradient Methods for Structured Non-monotone Variational Inequalities: Improved Analysis under Weaker Conditions”. Links: [poster](#)
46. 10 December - 16 December, 2023. [NeurIPS 2023](#), New Orleans, USA. Poster: “Accelerated Zeroth-order Method for Non-Smooth Stochastic Convex Optimization Problem with Infinite Variance”. Links: [poster](#)
45. 10 December - 16 December, 2023. [NeurIPS 2023](#), New Orleans, USA. Poster: “Byzantine-Tolerant Methods for Distributed Variational Inequalities”. Links: [poster](#)
44. 15 September, 2023. TES Conference on Mathematical Optimization for Machine Learning, Berlin, Germany. Talk: “Clipped Methods for Stochastic Optimization with Heavy-Tailed Noise”. Links: [slides](#)
43. 27 July, 2023. [ICML 2023](#), Honolulu, USA. Poster: “High-Probability Bounds for Stochastic Optimization and Variational Inequalities: the Case of Unbounded Variance”. Links: [poster](#)
42. 25 July, 2023. [ICML 2023](#), Honolulu, USA. Poster: “Convergence of Proximal Point and Extragradient-Based Methods Beyond Monotonicity: the Case of Negative Comonotonicity”. Links: [poster](#)
41. 6 June, 2023. [Oberseminar at LT Group, University of Hamburg](#), Hamburg, Germany. Talk: “Algorithms for Stochastic Optimization with Heavy-Tailed Noise and Connections with the Training of Large Language Models”. Links: [slides](#)
40. 2 May, 2023. [ICLR 2023](#), Kigali, Rwanda. Poster: “Variance Reduction is an Antidote to Byzantines: Better Rates, Weaker Assumptions and Communication Compression as a Cherry on the Top”. Links: [poster](#)
39. 27 April, 2023. [AISTATS 2023](#), Valencia, Spain. Poster: “Stochastic Gradient Descent-Ascent: Unified Theory and New Efficient Methods”. Links: [poster](#)
38. 13 February, 2023. [PEP talks](#), UCLouvain, Belgium. Talk “Convergence of Proximal Point and Extragradient-Based Methods Beyond Monotonicity: the Case of Negative Comonotonicity”. Links: [video](#), [slides](#)
37. 28 November - 9 December, 2022. [NeurIPS 2022](#), New Orleans, USA. Poster: “Clipped Stochastic Methods for Variational Inequalities with Heavy-Tailed Noise”. Links: [poster](#)
36. 28 November - 9 December, 2022. [NeurIPS 2022](#), New Orleans, USA. Poster: “Last-Iterate Convergence of Optimistic Gradient Method for Monotone Variational Inequalities”. Links: [poster](#)
35. 8 October, 2022. [MBZUAI Workshop on Collaborative Learning: From Theory to Practice](#), Abu Dhabi, UAE. Talk “Variance Reduction is an Antidote to Byzantines: Better Rates, Weaker Assumptions and Communication Compression as a Cherry on the Top”. Links: [slides](#)
34. 9 September, 2022. [All-Russian Optimization Seminar](#), online. Talk “Methods with Clipping for Stochastic Optimization and Variational Inequalities with Heavy-Tailed Noise” (in Russian). Links: [video](#), [slides](#)
33. 21 July, 2022. [ICML 2022](#), Baltimore, USA. Poster: “Secure Distributed Training at Scale”. Links: [poster](#)

32. 21 July, 2022. [ICML 2022](#), Baltimore, USA. Poster: “3PC: Three Point Compressors for Communication-Efficient Distributed Training and a Better Theory for Lazy Aggregation”. Links: [poster](#)
31. 3 July, 2022. [MOTOR 2022](#), Petrozavodsk, Russia. Talk: “Distributed Methods with Absolute Compression and Error Compensation”. Links: [slides](#)
30. 25 April, 2022. [Lagrange Workshop on Federated Learning](#), online. Talk: “Secure Distributed Training at Scale”. Links: [slides](#)
29. 29 March, 2022. [AISTATS 2022](#), online. Poster “Extragradient Method: $\mathcal{O}(1/K)$ Last-Iterate Convergence for Monotone Variational Inequalities and Connections With Co-coercivity”. Links: [poster](#)
28. 28 March, 2022. [AISTATS 2022](#), online. Poster “Stochastic Extragradient: General Analysis and Improved Rates”. Links: [poster](#)
27. 13 March, 2022. [Rising Stars in AI Symposium 2022](#), KAUST, Saudi Arabia. Talk “Extragradient Method: $\mathcal{O}(1/K)$ Last-Iterate Convergence for Monotone Variational Inequalities and Connections With Cocoercivity”. Links: [slides](#), [video](#)
26. 16 February, 2022. [Vector Institute Endless Summer School session “NeurIPS 2021 Highlights”](#), online. Talk “Moshpit SGD: Communication-Efficient Decentralized Training on Heterogeneous Unreliable Devices” (jointly with [Max Ryabinin](#)). Links: [slides](#)
25. 20 December, 2021. [MLO EPFL](#) internal seminar, online. Talk “Moshpit SGD: Communication-Efficient Decentralized Training on Heterogeneous Unreliable Devices”. Links: [slides](#)
24. 10 December, 2021. [NeurIPS 2021](#), online. Poster “Moshpit SGD: Communication-Efficient Decentralized Training on Heterogeneous Unreliable Devices”. Links: [poster](#)
23. 1 December, 2021. [MTL MLOpt](#) (internal seminar), online. Talk “Extragradient Method: $\mathcal{O}(1/K)$ Last-Iterate Convergence for Monotone Variational Inequalities and Connections With Cocoercivity”. Links: [slides](#)
22. 17 November, 2021. [All-Russian Optimization Seminar](#), online. Talk “Extragradient Method: $\mathcal{O}(1/K)$ Last-Iterate Convergence for Monotone Variational Inequalities and Connections With Cocoercivity” (in Russian). Links: [video](#), [slides](#)
21. 3 November, 2021. [Federated Learning One-World Seminar](#), online. Talk “Secure Distributed Training at Scale”. Links: [video](#), [slides](#)
20. 21 July, 2021. [ICML 2021](#), online. Poster “MARINA: Faster Non-Convex Distributed Learning with Compression”. Links: [poster](#)
19. 14 April, 2021. [AISTATS 2021](#), online. Poster “Local SGD: Unified Theory and New Efficient Methods”. Links: [poster](#)
18. 10 March, 2021, [Federated Learning One-World Seminar](#), online. Talk “MARINA: Faster Non-Convex Distributed Learning with Compression”. Links: [video](#), [slides](#)
17. 19 January, 2021. [NeurIPS New Year AfterParty at Yandex](#). Talk “Linearly Converging Error Compensated SGD”. Links: [video](#)
16. 9 December, 2020. [NeurIPS 2020](#), online. Poster “Stochastic Optimization with Heavy-Tailed Noise via Accelerated Gradient Clipping” (presented by [M. Danilova](#)). Links: [video](#), [poster](#)
15. 9 December, 2020. [NeurIPS 2020](#), online. Poster “Linearly Converging Error Compensated SGD”. Links: [video](#), [poster](#)
14. 7 October, 2020, [Federated Learning One-World Seminar](#) and [Russian Optimization Seminar](#), online. Talk “Linearly Converging Error Compensated SGD”. Links: [video](#), [slides](#)

13. 26 – 28 August, 2020, 23rd International Conference on Artificial Intelligence and Statistics (AISTATS 2020), online. I have presented our joint work with [Filip Hanzely](#) and [Peter Richtárik](#) called “A Unified Theory of SGD: Variance Reduction, Sampling, Quantization and Coordinate Descent”. Links: [video](#)
12. 8 July, 2020, [Russian Optimization Seminar](#), online. Talk “On the convergence of SGD-like methods for convex and non-convex optimization problems” (in Russian). Links: [video](#), [slides](#)
11. 28 June – 10 July, 2020, Machine Learning Summer School, online. I have presented our joint work with [Dmitry Kovalev](#), Dmitry Makarenko and [Peter Richtárik](#) called “Linearly Converging Error Compensated SGD”. Links: [video](#), [slides](#)
10. 27 – 30 April, 2020, 8-th International Conference on Learning Representations (ICLR 2020), online. I have presented our joint work with [Adel Bibi](#), [Ozan Sener](#), [El Houcine Bergou](#) and [Peter Richtárik](#) called “A Stochastic Derivative Free Optimization Method with Momentum”. Links: [video](#)
9. 14 December, 2019, NeurIPS 2019 workshop “[Optimization Foundations for Reinforcement Learning](#)”, Vancouver, Canada. [Poster](#) “A Stochastic Derivative Free Optimization Method with Momentum”
8. 13 December, 2019, NeurIPS 2019 workshop “[Beyond First Order Methods in ML](#)”, Vancouver, Canada. [Poster](#) “An Accelerated Method for Derivative-Free Smooth Stochastic Convex Optimization”
7. 18 October, 2019, [SIERRA](#) research seminar, INRIA. Talk “A Unified Theory of SGD: Variance Reduction, Sampling, Quantization and Coordinate Descent”. Links: [slides](#)
6. 1-6 July 2018, [23rd International Symposium on Mathematical Programming](#), Bordeaux, France. [Talk](#) “An Accelerated Directional Derivative Method for Smooth Stochastic Convex Optimization”
5. 10-15 June 2018, Traditional Youth School “Control, Information and Optimization” organized by [Boris Polyak](#) and [Elena Gryazina](#), Voronovo, Russia. [Poster](#) and [Talk](#) “An Accelerated Directional Derivative Method for Smooth Stochastic Convex Optimization”
4. 14 April 2018, Workshop “Optimization at Work”, MIPT, Dolgoprudny, Russia. [Talk](#) “An Accelerated Method for Derivative-Free Smooth Stochastic Convex Optimization”
3. 5-7 February 2018, [KAUST Research Workshop on Optimization and Big Data](#), KAUST, Thuwal, Saudi Arabia. Joint [Poster](#) “Stochastic Spectral Descent Methods” with D. Kovalev and E. Gasanov
2. 25 November 2017, 60th Scientific Conference of MIPT, Section of Information Transmission Problems, Data Analysis and Optimization, IITP, Moscow, Russia. [Talk](#) “About accelerated Directional Search with non-Euclidean prox-structure”
1. 27 October 2017, Workshop “Optimization at Work”, MIPT, Dolgoprudny, Russia. [Talk](#) “Accelerated Directional Search with non-Euclidean prox-structure”

REVIEWING

- Journals: [Journal of Machine Learning Research \(JMLR\)](#) (2020, 2023, 2024), [Mathematical Programming](#) (2024), [Transactions on Machine Learning Research](#) (2022, 2024), [Journal Numerische Mathematik](#) (2023), [SIAM Journal on Mathematics of Data Science \(SIMODS\)](#) (2023), [Science China Mathematics](#) (2021), [SIAM Journal on Optimization \(SIOPT\)](#) (2020, 2021), [Optimization Methods and Software](#) (2019)

- Conferences: [AISTATS](#) (2022, 2023, 2025), [ICLR](#) (2021, 2022, 2023, 2024, 2025), [ICML](#): (2019, 2021, 2022, 2024), [NETYS](#) (2024, 2025)

ORGANIZATION

- Action Editor: [TMLR](#) (July 2024 – Now)
- Organizer of [Russian Optimization Seminar](#): May 2020 – August 2022
- Organizer of [research seminar on Optimization at MIPT](#): March 2020 – June 2020

SUMMER SCHOOLS

- **28 June - 10 July 2020.** Participant of [Machine Learning Summer School](#). I have presented our joint work with [Dmitry Kovalev](#), Dmitry Makarenko and [Peter Richtárik](#) called “Linearly Converging Error Compensated SGD”. Links: [video](#), [slides](#)
- **June 2018.** Participant of Traditional Youth School “Control, Information and Optimization”
- **June 2017.** Participant of Traditional Youth School “Control, Information and Optimization”
- **July 2015.** [Participant](#) of Summer School “Contemporary Mathematics” in Dubna
- **July 2014.** [Participant](#) of Summer School “Contemporary Mathematics” in Dubna

LANGUAGES

RUSSIAN: Mothertongue
ENGLISH: Advanced

COMPUTER SKILLS

Operating Systems: MICROSOFT WINDOWS, LINUX, MAC OS
Programming Languages: PYTHON, \LaTeX , C, C++

INTERESTS

- Wakesurfing, Fitness, Hiking
- Football: 9 years in football school in Rybinsk, Russia. I was also playing for an [amateur team](#)

Last Updated on January 21, 2025