Grigory Yaroslavtsev
http://grigory.ai

Engineering 4427
Department of Computer Science
4400 University Drive, Fairfax, VA
grigory@grigory.us
U.S. citizen

Research Interests

Algorithms and theory for large-scale data analysis, artificial intelligence, machine learning and data science.

Academic Appointments

2021–current **Assistant Professor of Computer Science**, *George Mason University*, Fairfax, VA

2020–2021 **Adjunct Assistant Professor of Computer Science**, *Indiana University*, Bloomington, IN

2019 **Visiting Researcher**, *Alan Turing Institute for Data Science and AI*, London, UK

2018–2020 **Adjunct Assistant Professor of Statistics**, *Indiana University*, Bloomington, IN

2017–2020 **Founding Director, Center for Algorithms and Machine Learning (CAML)**, *Indiana University*, Bloomington, IN, http://caml.indiana.edu

2016–2020 **Assistant Professor of Computer Science**, *Indiana University*, Bloomington, IN

2014–2016 **Warren Center Postdoctoral Fellow**, *University of Pennsylvania*, Philadelphia, PA, Departments of Computer and Information Sciences and Statistics at the Wharton Business School. Mentors: Michael Kearns (CIS) and Elchanan Mossel (Stat)

2013–2014 **ICERM Postdoctoral Fellow**, *Brown University*, Providence, RI, Mentor: Philip Klein

Education

2010–2013 **Ph.D.**, *Pennsylvania State University*, State College, PA, Computer Science & Engineering
Thesis: "Efficient Combinatorial Techniques in Sparsification, Summarization and Testing of Large Datasets." Advisor: Sofya Raskhodnikova
Joined by invitation, didn't apply to any other Ph.D. programs

2008–2010 **M.Sc.**, *Academic University of the Russian Academy of Sciences*, St. Petersburg, Russia, Applied Math and Physics, 1st student in the pilot TCS class

2004–2008 **B.Sc.**, *St. Petersburg State Polytechnic University*, St. Petersburg, Russia, Physics and Technology, ranked 1st on the entry test for the department

Research Internships

Summer 2013 **Theory group**, **Microsoft Research**, Redmond, WA, Mentor: Konstantin Makarychev

Fall 2012 **Theory group**, **Microsoft Research**, Mountain View, CA, Mentor: Alex Andoni

Summer 2012 **Theory group**, **IBM Research**, Almaden, CA, Mentor: David Woodruff

Summer 2011 **Database theory group**, **AT&T Labs–Research, Shannon Laboratory**, Florham Park, NJ, Mentors: Graham Cormode, Cecilia M. Procopiuc, Divesh Srivastava and Howard Karloff

Selected Awards and Honors

2019 **NeurIPS'19 Highest-Scoring Reviewer (Top 400)**

2019 **Alan Turing Institute Visiting Researcher**, *London, UK*

2017 **Facebook Faculty Research Award**

2014–2016 **Warren Center Postdoctoral Fellowship**, *University of Pennsylvania*

2013–2014 **Institute Postdoctoral Fellowship**, *Brown University, ICERM*

2012 **Best Graduate Research Assistant Award at CSE Department**, *Penn State*

2010 **TopCoder Open Algorithm Competition Finalist (Top-24 worldwide)**

2010–2013 **College of Engineering Fellowship**, *Penn State*

2010–2011 **University Graduate Fellowship**, *Penn State*

2009–2010 **Yandex Personal Research Grant**, *Academic University of the RAS*

2008–2009 **Coach of the top team in St. Petersburg High-School Olympiad in Informatics**

2004 **Ranked 1st on the entry test for the Department of Physics and Technology**, *St. Petersburg State Polytechnic University*

Media Coverage

For PNAS paper "Private Algorithms for the Protected in Social Network Search":

- PBS Newshour "The secret things you give away through your phone metadata"
- Schneier on Secuity "Research on Balancing Privacy with Surveillance."
- Association of American Universities "Penn Researchers Balance Privacy and Security in Network Analysis."
- ACM Tech News / The Daily Pennsylvanian: "Penn Professor's Computer Algorithm Could Fight Terrorism While Protecting Privacy."
- Quartz: "There may be a way to allow mass surveillance and preserve our privacy at the same time."
- Pacific Standard: "Searching Private Data, and Ensuring It Stays Private."
- Wired (German): "Ein neuer Überwachungs-Algorithmus soll in Social Media nur auf Terroristen zielen."
- Vice Motherboard: "Algorithms Claim to Hunt Terrorists While Protecting the Privacy of Others."
- The Naked Scientists Podcast: "National Security Algorithm."

Selected Coverage in Books and Classes at Other Universities

STOC'14 paper "Parallel Algorithms for Geometric Graph Problems":

- **Harvard** CS 229r, Fall'13, "Algorithms for Big Data". Taught by Jelani Nelson.
- **Columbia** COMS 6998-9, Fall'15, "Algorithmic Techniques for Massive Data". Taught by Alexandr Andoni.
- **ETH Zurich**, Spring'19, "Massively Parallel Algorithms". Taught by Mohsen Ghaffari.

STOC'14 paper "L_p-Testing":

- **MIT** 6.889, Fall'17, "Sublinear Time Algorithms". Taught by Ronitt Rubinfeld.
- **Introduction to Property Testing**, by Oded Goldreich.
- **Encyclopedia of Algorithms**, article by Sofya Raskhodnikova.

SAT'09 paper "Finding Efficient Circuits using SAT-solvers":

- **Stanford** CS354, Fall'11, Spring'14, Spring'16, "Topics in Circuit Complexity". Taught by Ryan Williams.
- **"The Art of Computer Programming", Volume 4**, by Donald E. Knuth.

Funding

2017–2021 **NSF CRII Award**, *"Algorithms for Noise-Tolerant Function Testing with Applications to Deep Learning"*, Sole PI, award amount: $\$175,000$

2018–2020 **Google Cloud Platform Credit**, Award amount: $\$15,000$

2017 **Facebook Faculty Research Award**, Award amount: $\$35,000$

Journal Papers

PNAS 16 Private Algorithms for the Protected in Social Network Search, with M. Kearns, A. Roth and S. Wu, **Proceedings of the National Academy of Sciences**.

Algo 16 Certifying Equality with Limited Interaction, with J. Brody, A. Chakrabarti, R. Kondapally and D. Woodruff, **Algorithmica, special issue on "Information Complexity and Applications**.

TODS 14 Private Analysis of Graph Structure, with V. Karwa, S. Raskhodnikova and A. Smith, ACM Transactions on Database Systems.

Comb 14 Steiner Transitive-Closure Spanners of Low-Dimensional Posets, with P. Berman, A. Bhattacharyya, E. Grigorescu, S. Raskhodnikova and D. Woodruff, Combinatorica.

I&C 12 Approximation Algorithms for Spanner Problems and Directed Steiner Forest, with P. Berman, A. Bhattacharyya, K. Makarychev and S. Raskhodnikova, **Information and Computation, special issue for ICALP'11**.

IPL 10 New upper bounds on the Boolean Circuit Complexity of Symmetric Functions, with E. Demenkov, A. Kojevnikov and A. Kulikov, Information Processing Letters.

Peer-Reviewed Conference and Workshop Papers

Authors listed in alphabetical order unless marked with ★ for ordering by contribution:

IJCAI 23 HOUDINI: Escaping from Moderately Constrained Saddles, with D. Avdiukhin, 32nd International Joint Conference on Artificial Intelligence.

AAAI 23 ★ Tree Learning: Optimal Algorithms and Sample Complexity, D. Avdiukhin, G.Yaroslavtsev, D.Vainstein, O.Fischer, S.Das, F. Mirza, 37th AAAI Conference on Artificial Intelligence.

NeurIPS 21 Escaping Saddle Points with Compressed SGD, with D.Avdiukhin, 35th Conference on Neural Information Processing Systems.

AAAI 21 ★ Objective-Based Hierarchical Clustering of Deep Embedding Vectors, S.Naumov, G. Yaroslavtsev and D.Avdiukhin, 35th AAAI Conference on Artificial Intelligence.

AISTATS 20 Bisect and Conquer: Hierarchical Clustering via Max-Uncut Bisection, with S. Ahmadian, V. Chatziafratis, A. Epasto, E. Lee, K. Makarychev and M. Mahdian, 23rd International Conference on Artificial Intelligence and Statistics.

AISTATS 20 ★ "Bring Your Own Greedy"+Max: Near-Optimal 1/2-Approximations for Submodular Knapsack, G. Yaroslavtsev, S. Zhou and D. Avdiukhin, 23rd International Conference on Artificial Intelligence and Statistics.

SOSA@SODA 20 Fast Fourier Sparsity Testing, with S. Zhou, 3rd SIAM Symposium on Simplicity in Algorithms.

OPT@NeurIPS 19 Escaping Saddle Points with Inequality Constraints via Noisy Sticky Projected Gradient Descent, with D. Avdiukhin and C. Jin, 11th OPT Workshop on Optimization for Machine Learning, **Oral + poster**.

RANDOM 19 Approximate $\mathbb{F}_2$-Sketching of Valuation Functions, with S. Zhou, 23rd International Workshop on Randomization and Computation.

CCC 19 Optimality of Linear Sketching under Modular Updates, with K. Hosseini and S. Lovett, 34th Conference on Computational Complexity.

KDD 19 Adversarially Robust Submodular Maximization under Knapsack Constraints, with D. Avdiukhin, S. Mitrovic and S. Zhou, 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Research Track, **Oral presentation (9.2% acceptance rate)**.

VLDB 19 Multi-Dimensional Balanced Graph Partitioning via Projected Gradient Descent, with D. Avdiukhin and S. Pupyrev, 45th International Conference on Very Large Data Bases, Research Track.

AISTATS 19 Hierarchical Clustering for Euclidean Data, with M. Charikar, V. Chatziafratis and R. Niazadeh, 22nd International Conference on Artificial Intelligence and Statistics.

ICML 18 ★ Massively Parallel Algorithms and Hardness for Single-Linkage Clustering under ℓ_p-Distances, G. Yaroslavtsev and A. Vadapalli., 35th International Conference on Machine Learning, **Long talk (8.6% accceptance rate)**.

CCC 18 Linear Sketching over $\mathbb{F}_2$, with S. Kannan, E. Mossel and S. Sanyal, 33rd Conference on Computational Complexity.

SODA 16 Tight Bounds on Linear Sketches of Approximate Matchings, with S. Assadi, S. Khanna and Y. Li, 27th Annual ACM-SIAM Symposium on Discrete Algorithms.

ICALP 15 Amplification of One-Way Information Complexity via Codes and Noise Sensitivity, with M. Molinaro and D. Woodruff, 42nd International Colloquium on Automata, Languages and Programming.

STOC 15 Near Optimal LP Rounding Algorithm for Correlation Clustering on Complete and Complete k-partite Graphs, with S. Chawla, K. Makarychev and T. Schramm, 47th ACM Symposium on the Theory of Computing.

RANDOM 14 Certifying Equality with Limited Interaction, with J. Brody, A. Chakrabarti, R. Kondapally and D. Woodruff, 18th International Workshop on Randomization and Computation.

PODC 14 Beyond Set Disjointness: The Communication Complexity of Finding the Intersection, with J. Brody, A. Chakrabarti, R. Kondapally and D. Woodruff, 33rd ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing.

STOC 14 Parallel Algorithms for Geometric Graph Problems, with A. Andoni, K. Onak and A. Nikolov, 46th ACM Symposium on the Theory of Computing.

STOC 14 L_p-testing, with P. Berman and S. Raskhodnikova, 46th ACM Symposium on the Theory of Computing.

CCC 14 Lower Bounds for Testing Properties of Functions over Hypergrid Domains, with E. Blais and S. Raskhodnikova, 29th IEEE Conference on Computational Complexity.

ICDE 13 ★ Accurate and Efficient Private Release of Datacubes and Contingency Tables, G. Yaroslavtsev, G. Cormode, C. Procopiuc and D. Srivastava, 29th IEEE International Conference on Data Engineering.

SODA 13 Beating the Direct Sum Theorem in Communication Complexity with Implications for Sketching, with Marco Molinaro and David Woodruff, 24th Annual ACM-SIAM Symposium on Discrete Algorithms, **Invited to a special issue of "Algorithmica" on "Information Complexity and Applications"**.

SODA 13 Learning Pseudo-Boolean k-DNF and Submodular Functions, with S. Raskhodnikova, 24th Annual ACM-SIAM Symposium on Discrete Algorithms.

APPROX 12 Primal-Dual Algorithms for Node-Weighted Network Design in Planar Graphs, with P. Berman, 15th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems.

VLDB 11 Private Analysis of Graph Structure, with V. Karwa, S. Raskhodnikova and A. Smith, 37th International Conference on Very Large Data Bases, Research Track.

ICALP 11 Improved Approximation for the Directed Spanner Problem, with P. Berman, A. Bhattacharyya, K. Makarychev and S. Raskhodnikova, 38th International Colloquium on Automata, Languages and Programming, **Runner-up for the Best Paper Award, invited to a special issue of a journal "Information and Computation"**.

ICALP 11 Steiner Transitive-Closure Spanners of Low-Dimensional Posets, with P. Berman, A. Bhattacharrya, E. Grigorescu, S. Raskhodnikova and D. Woodruff, 38th International Colloquium on Automata, Languages and Programming.

SAT 09 Finding Efficient Circuits using SAT-solvers, with A. Kojevnikov and A. Kulikov, 12th International Conference on Theory and Applications of Satisfiability Testing.

Professional Activities

ML and AI program committees (reviewer level)

- AAAI (AAAI Conference on Artificial Intelligence): 2021, 2020.
- AISTATS (International Conference on Artificial Intelligence and Statistics): 2021, 2020, 2019.
- "New in ML"@NeurIPS(Neural Information Processing Systems): 2021, 2019.

- NeurIPS (Neural Information Processing Systems): 2021, 2020, 2019, 2018, 2017, 2016.
- ICLR (International Conference on Learning Representations): 2022, 2021, 2020, 2019, 2018.
- ICML (International Conference on Machine Learning): 2021, 2020, 2019, 2018.
- UAI (Uncertainty in Artificial Intelligence): 2019.

Theory program committees

- **32nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'21).**
- 21st International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (APPROX'18).
- 5th Workshop on Algorithms and Systems for MapReduce and Beyond (BeyondMR'18) at SIGMOD/PODS'18.
- 23rd International Computing and Combinatorics Conference (COCOON'17).
- **28th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'17).**
- 24th Annual European Symposium on Algorithms (ESA'16), Design and Analysis Track.
- 41st International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM'15), Foundations of Computer Science Track.

Organization

- **Recent Trends in Clustering and Classification**
  3-day workshop at Toyota Technological Institute, Chicago, Sep 2019.
  `http://caml.indiana.edu/rtcc.html`
- **Seminar of the Center for Algorithms and Machine Learning**
  Weekly seminar on research in Algorithms and ML at Indiana University (2018 – 2020).
  `http://caml.indiana.edu/`
- **Linear Sketching as a Tool for Everything**
  1-day workshop at IEEE FOCS, Oct 2017.
  `http://caml.indiana.edu/linear-sketching.focs.html`
- **67th Midwest Theory Day**
  2-day workshop at Indiana University, Bloomington, Apr 2017.
  `http://caml.indiana.edu/mtd.html`
- **Big Data through the Lens of Sublinear Algorithms**
  2-day workshop at Rutgers University, DIMACS, Aug 2015.
  `http://grigory.us/mpc-workshop-dimacs.html`
- **Algorithmic Frontiers of Modern Massively Parallel Computation**
  1-day workshop at ACM FCRC/STOC, Jun 2015.
  `http://grigory.us/mpc-workshop-fcrc.html`
- **Sublinear Algorithms and Big Data Day**
  Brown University, ICERM, Apr 2014.
  `http://grigory.us/big-data-day.html`
- **Theory Seminar**
  Univeristy of Pennsylvania Computer and Information Sciences Department (2014 – 2016).
  `http://theory.cis.upenn.edu/seminar/`
- **Theory Seminar**
  Brown CS Department and ICERM (2013 – 2014).
  `http://grigory.us/theory-seminar-brown-spring14.html`

Other Service

Service to federal funding agencies

- Panelist for grant proposals for NSF IIS core programs: 2020.
- Panelist for grant proposals for NSF CCF core programs: 2019, 2018, 2017.

- Reviewer for Israeli Science Foundation grant proposals: 2017.

Internal service at Indiana University

- Founding head of the departmental Graduate Research Award committee ('16–'20)
- Graduate admission committee member ('16–'17, '18–'19)
- Graduate education committee member ('17–'18)
- Undergraduate education committee member ('18–'20)

Selected Talks

May 2023 **Google Research**, New York, NY, Google Tech Talk. *Learning from Tuples*.

May 2023 **New York University**, New York, NY, Theory Seminar. *Learning from Tuples*.

May 2023 **Columbia University**, New York, NY, Theory Seminar. *Learning from Tuples*.

Apr 2022 **University of Texas**, Austin, TX, Theory Seminar. *Hierarchical Clustering for Everyone*.

Oct 2021 **University of Maryland**, College Park, MD, Capital Area Theory Seminar. *Hierarchical Clustering for Everyone*.

May 2020 **University of California, Davis**, Davis, CA, Computer Science Colloquium (Virtual). *Hierarchical Clustering for Everyone*.

May 2020 **University of Wisconsin-Madison**, Madison, WI, SILO Seminar (Virtual). *Advances in Gradient Descent Methods for Non-Convex Optimization*.

Feb 2020 **Brown University**, Providence, RI, Computer Science Colloquium. *Hierarchical Clustering for Everyone*.

Feb 2020 **George Mason University**, Fairfax, VA, Computer Science Colloquium. *Hierarchical Clustering for Everyone*.

Jan 2020 **University of California, Davis**, Davis, CA, Math Colloquium. *Hierarchical Clustering for Everyone*.

Oct 2019 **University of Illinois**, Urbana-Champaign, IL, CSL Seminar. *Advances in Hierarchical Clustering of Vector Data*.

Oct 2019 **University of California, Riverside**, Riverside, CA, Departmental Colloquium, CSE. *Advances in Hierarchical Clustering of Vector Data*.

Oct 2019 **California Institute of Technology**, Pasadena, CA, CMI Seminar. *Advances in Linear Sketching over Finite Fields*.

Oct 2019 **University of California, San Diego**, San Diego, CA, Theory Seminar. *Advances in Hierarchical Clustering of Vector Data*.

Oct 2019 **University of Southern California**, Los Angeles, CA, Theory Lunch. *Advances in Hierarchical Clustering of Vector Data*.

Aug 2019 **Google Research**, Menlo Park, CA, Tech Talk. *Advances in Hierarchical Clustering of Vector Data*.

Jun 2019 **University of Warwick**, Warwick, UK, Discrete Mathematics and Applications Seminar. *Advances in Hierarchical Clustering of Vector Data*.

May 2019 **University of Oxford**, Oxford, UK, Algorithms and Complexity Seminar. *Advances in Hierarchical Clustering of Vector Data*.

Mar 2019 **Facebook Core Data Science**, Menlo Park, CA, Tech Talk. *Advances in Hierarchical Clustering of Vector Data*.

Mar 2019 **Johns Hopkins University**, Baltimore, MD, Algorithms and Complexity Seminar. *Advances in Hierarchical Clustering of Vector Data*.

Mar 2019 **Northwestern University**, Evanston, IL, Computer Science Seminar. *Advances in Hierarchical Clustering of Vector Data*.

Oct 2018 **Simons Institute for the Theory of Computing, UC Berkeley**, Berkeley, CA, Workshop on Interactive Complexity. *Advances in Linear Sketching over Finite Fields*.

Aug 2018 **IBM Almaden Research Center**, San Jose, CA, Theory Seminar. *Massively Parallel Algorithms and Hardness for Single-Linkage Clustering under ℓ_p-Distances*.

Jun 2018 **Massachusetts Institute of Technology**, Cambridge, MA, 2nd Workshop on Local Algorithms. *Badger Rampage: Multidimensional Balanced Partitioning of Facebook-scale Graphs*.

May 2018 **Stanford University**, Palo Alto, CA, Theory Seminar. *Massively Parallel Algorithms and Hardness for Single-Linkage Clustering under ℓ_p-Distances*.

Mar 2018 **University of Warwick**, Warwick, UK, Workshop on Data Summarization. *Massively Parallel Algorithms and Hardness for Single-Linkage Clustering under ℓ_p-Distances*.

Sep 2018 **University of Michigan**, Ann Arbor, MI, Theory Seminar. *Linear Sketching for Functions over the Boolean Hypercube*.

Apr 2018 **Toyota Technological Institute**, Chicago, IL, 68th Midwest Theory Workshop. *Linear Sketching for Functions over the Boolean Hypercube*.

Oct 2017 **IEEE FOCS**, Berkeley, CA, Workshop "Linear Sketching as a Tool for Everything". *Linear Sketching for Functions over the Boolean Hypercube*.

Oct 2017 **Facebook**, Menlo Park, CA, Tech Talk. *Clustering on Clusters 2049: Massively Parallel Algorithms for Clustering Graphs and Vectors*.

Jun 2017 **ITMO University**, St. Petersburg, Russia, Departmental Colloquium. *Computational and Communication Complexity in Massively Parallel Computing*.

Jun 2017 **Higher School of Economics**, Moscow, Russia, Workshop on Complexity of Computation, Communication, Descriptions and Proofs. *Computational and Communication Complexity in Massively Parallel Computing*.

Feb 2017 **Facebook**, Menlo Park, CA, Tech Talk. *Clustering on Clusters: Massively Parallel Algorithms for Clustering Graphs and Vectors*.

May 2017 **St. Petersburg Department of Steklov Institute of Mathematics**, St. Petersburg, Russia, Theory Seminar. *Linear Sketching over F_2*.

May 2017 **Moscow State University**, Moscow, Russia, Kolmogorov Seminar. *Linear Sketching over F_2*.

Mar 2017 **BIRS Research Center**, Banff, Canada, Banff Workshop on Communication Complexity and Applications II. *Linear Sketching over F_2*.

Nov 2016 **Columbia University**, New York, NY, Theory Seminar. *Linear Sketching over F_2*.

Oct 2016 **University of Pennsylvania**, Philadelphia, PA, Theory Seminar. *Linear Sketching over F_2*.

Sep 2016 **University of Utah**, Salt Lake City, UT, Theory Seminar. *Linear Sketching over F_2*.

Aug 2016 **University of Illinois**, Urbana, IL, Theory Seminar. *Linear Sketching over F_2*.

Jun 2016 **Microsoft Research**, Redmond, WA, Theory Seminar. *Linear Sketching over F_2*.

Mar 2016 **Drexel University**, Philadelphia, PA, Departmental Colloquium. *What's New in "The Big Data Theory"?*.

Feb 2016 **Georgetown University**, Washington, DC, Departmental Colloquium. *What's New in "The Big Data Theory"?*.

Feb 2016 **Indiana University**, Bloomington, IN, Departmental Colloquium. *What's New in "The Big Data Theory"?*.

Feb 2016 **University of Colorado**, Boulder, CO, Departmental Colloquium. *What's New in "The Big Data Theory"?*.

Feb 2016 **Boston University**, Boston, MA, Departmental Colloquium. *What's New in "The Big Data Theory"?*.

Feb 2016 **College of William and Mary**, Williamsburg, VA, Departmental Colloquium. *What's New in "The Big Data Theory"?*.

Aug 2015 **University of Wisconsin**, Madison, WI, Theory Seminar. *Fast Fourier Sparsity Testing over the Boolean Hypercube*.

Jul 2015 **ISMP'15**, Pittsburgh, PA, 22nd International Symposium on Mathematical Programming. *Parallel Algorithms for Geometric Problems*.

May 2015 **Cornell University**, Ithaca, NY, Theory Seminar. *Near Optimal LP Rounding for Correlation Clustering*.

Apr 2015 **Massachusetts Institute of Technology**, Cambridge, MA, Algorithms and Complexity Seminar. *Near Optimal LP Rounding for Correlation Clustering*.

Mar 2015 **Microsoft Research**, Redmond, WA, Theory Seminar. *Near Optimal LP Rounding for Correlation Clustering*.

Feb 2015 **Google Research**, New York, NY, Tech Talk. *Near Optimal LP Rounding for Correlation Clustering*.

Jan 2015 **Rutgers University**, New Brunswick, NJ, Theory Seminar. *Near Optimal LP Rounding for Correlation Clustering*.

Jan 2015 **Carnegie Mellon University**, Pittsburgh, PA, Theory Lunch. *Near Optimal LP Rounding for Correlation Clustering*.

Jan 2015 **Pennsylvania State University**, State College, PA, Departmental Colloquium. *Near Optimal LP Rounding for Correlation Clustering*.

Nov 2014 **Johns Hopkins University**, Baltimore, MD, Theory Seminar. *Parallel Algorithms for Geometric Problems*.

Oct 2014 **University of Maryland**, College Park, MD, Capital Area Theory Seminar. *Parallel Algorithms for Geometric Problems*.

Aug 2014 **University of Pennsylvania**, Philadelphia, PA, Theory Seminar. *Parallel Algorithms for Geometric Problems*.

May 2014 **University of Massachusetts**, Amherst, MA, Theory Seminar. *Parallel Algorithms for Geometric Problems*.

May 2014 **Massachusetts Institute of Technology**, Cambridge, MA, Theory of Distributed Systems Seminar. *Beyond Set Disjointness: The Communication Complexity of Finding the Intersection*.

Mar 2014 **Google Research**, New York, NY, Tech Talk. *Parallel Algorithms for Geometric Problems*.

Mar 2014 **Sandia Labs**, Livermore, CA, Special Seminar. *Parallel Algorithms for Geometric Problems*.

Mar 2014 **Stanford University**, Palo Alto, CA, Theory Seminar. *Parallel Algorithms for Geometric Problems*.

Mar 2014 **Georgia Tech**, Atlanta, GA, Departmental Colloquium. *Approximating Graph Problems: The Old and The New*.

Feb 2014 **Massachusetts Institute of Technology**, Cambridge, MA, Algorithms and Complexity Seminar. *Approximating Graph Problems: The Old and The New*.

Feb 2014 **Yahoo! Research**, New York, NY, Tech Talk. *Approximating Graph Problems: The Old and The New*.

Feb 2014 **Toyota Technological Institute**, Chicago, IL, Departmental Colloquium. *Approximating Graph Problems: The Old and The New*.

Nov 2014 **University of Pennsylvania**, Philadelphia, PA, Statistics Seminar. *L_p-Testing*.

Oct 2014 **Columbia University**, New York, NY, Theory Seminar. *L_p-Testing*.

Jan 2014 **Microsoft Research**, Redmond, WA, Theory Lunch. *L_p-Testing*.

Nov 2014 **Harvard University**, Cambridge MA, Theory Seminar. *L_p-Testing*.

Nov 2013 **Brown University**, Providence RI, Theory Seminar. *L_p-Testing*.

Oct 2013 **IBM Almaden Research Center**, San Jose, CA, Theory Seminar. *L_p-Testing*.

Sep 2013 **Massachusetts Institute of Technology**, Cambridge, MA, Algorithms and Complexity Seminar. *Property Testing and Communication Complexity*.

Jun 2013 **Microsoft Research**, Redmond, WA, Theory Seminar. *Learning and Testing Submodular Functions*.

May 2013 **Aarhus University**, Aarhus, Denmark, Theory Seminar. *Beating the Direct Sum in Communication Complexity with Implications for Sketching*.

Apr 2013 **University of Melbourne**, Melbourne, Australia, Theory Seminar. *Learning and Testing Submodular Functions*.

Apr 2013 **University of Sydney**, Sydney, Australia, Theory Seminar. *Advances in Directed Spanners*.

Feb 2013 **UCLA**, Los Angeles, LA, Theory Seminar. *Learning and Testing Submodular Functions*.

Dec 2012 **Massachusetts Institute of Technology**, Cambridge, MA, Algorithms and Complexity Seminar. *Beating the Direct Sum in Communication Complexity with Implications for Sketching*.

Dec 2012 **Weizmann Institute of Science**, Rehovot, Israel, Theory Seminar. *Learning and Testing Submodular Functions*.

Dec 2012 **Harvard University**, Cambridge, MA, Theory Seminar. *Learning and Testing Submodular Functions*.

Dec 2012 **Carnegie Mellon University**, Pittsburgh, PA, Theory Lunch. *Learning and Testing Submodular Functions*.

Dec 2012 **Carnegie Mellon University**, Pittsburgh, PA, Operations Research Seminar. *Learning and Testing Submodular Functions*.

Nov 2012 **Princeton University**, Princeton, NJ, Theory Lunch. *Beating the Direct Sum in Communication Complexity with Implications for Sketching*.

Nov 2012 **IBM T.J. Watson Research Center**, Yorktown Heights, NY, Integer Programming Seminar. *Learning and Testing Submodular Functions*.

Nov 2012 **Columbia University**, New York, NY, Theory Seminar. *Learning and Testing Submodular Functions*.

Oct 2012 **Microsoft Research Silicon Valley**, Mountain View, CA, Lab Seminar. *Parallel Algorithms for Geometric Problems*.

Oct 2012 **Microsoft Research Silicon Valley**, Mountain View, CA, Theory Seminar. *Learning and Testing Submodular Functions*.

May 2012 **IBM Almaden Research Center**, San Jose, CA, Theory Seminar. *Learning and Testing Submodular Functions*.

Nov 2011 **Carnegie Mellon University**, Pittsburgh, PA, Theory Lunch. *Advances in Directed Spanners*.

Nov 2011 **University of Maryland**, College Park, MD, Capital Area Theory Seminar. *Advances in Directed Spanners*.

Aug 2011 **AT&T Labs – Research, Shannon Laboratory**, Florham Park, NJ, Lab Seminar. *Private Analysis of Graph Structure*.

Jun 2011 **AT&T Labs – Research, Shannon Laboratory**, Florham Park, NJ, Mathematics Research Colloquium and Informal Seminar. *Improved Approximation for the Directed Spanner Problem*.

May 2011 **Moscow State University**, Moscow, Russia, Combinatorial Optimization Seminar. *Improved Approximation for the Directed Spanner Problem*.

Apr 2011 **IBM T.J. Watson Research Center**, Yorktown Heights, NY, Integer Programming Seminar. *Improved Approximation for the Directed Spanner Problem*.

Dec 2010 **ITMO University**, St. Petersburg, Russia, Algorithms Seminar. *Improved Approximation for the Directed Spanner Problem*.

Apr 2010 **Pennsylvania State University**, State College, PA, Theory Seminar. *Linear Bounds on Circuit Complexity and Feebly One-Way Permutations*.

Teaching

- "**Analysis of Algorithms" (M.Sc. level)**
  George Mason University, CS 583, Fall 2022.
- "**Advanced Algorithms" (Ph.D. level)**
  George Mason University, CS 630, Spring 2022.
- "**Analysis of Algorithms" (M.Sc. level)**
  George Mason University, CS 583, Fall 2021.
- "**Data Structures" (Undergraduate level)**
  (Honors) Indiana University, Bloomington, CSCI-H343, Spring 2018 and 2019.
  Indiana University, Bloomington, CSCI-C343, Fall 2018.
- "**Applied Algorithms" (M.Sc. level)**
  Indiana University, Bloomington, CSCI-B505, Fall 2017 and 2019, Spring 2020.
- "**Foundations of Data Science" (Ph.D. level)**
  Indiana University, Bloomington, CSCI-B609, Fall 2016 and 2017.
- "**Algorithms for Big Data" (Ph.D. level)**
  University of Pennsylvania, CIS 700, Fall 2015.
- "**Computational Learning Theory" (All levels)**
  University of Pennsylvania, CIS 625, Spring 2015 (co-teaching with Michael Kearns).
- "**Sublinear Algorithms for Big Data" (All levels)**
  University of Buenos Aires, Argentina. 15-hour crash course. July–August 2014.

Tutorials

- "**Algorithms for MapReduce and Beyond**" (with Sergei Vassilvitskii, Google)
  24th International Conference on Information and Knowledge Management (CIKM 2015), Melbourne, Australia.

Guest lecturer at undergraduate classes

- CIS 399, "Foundations of Data Science", University of Pennsylvania, Spring 2016.
- CMPSC 464, "Introduction to the Theory of Computing", Pennsylvania State University, Fall 2010.

Extracurricular education for high-school students

- Prepared training contests for the **United States Team in the International Olympiad in Informatics**, 2011.
- Co-founder and coordinator of St. Petersburg network of extracurricular education in informatics for high-school students (http://spbtc.ru) (2009-2010).

Visits and Consulting

- **Facebook**, Menlo Park, CA. Consultant, Spring'17–Spring'19. (Host: Sergey Pupyrev)
- **Google Research**, New York, NY. Weekly visitor in Fall'14–Spring'15. (Host: Silvio Lattanzi)
- **Microsoft Research**, Redmond, WA. 03/08/15–03/14/15, 01/08/14–01/12/14. (Host: Konstantin Makarychev)
- **IBM T.J. Watson Research Center**, Yorktown Heights, NY. 04/19/11–04/21/11, 11/13/12–11/15/12. (Hosts: Konstantin Makarychev, Vishwanath Nagarajan)
- **AT&T Labs – Research, Shannon Laboratory**, Florham Park, NJ. 11/18/11–11/25/11. (Host: Howard Karloff)

- **Weizmann Institute of Science**, Rehovot, Israel. 12/27/12–01/04/13. (Host: Robert Krauthgamer)
- **University of Melbourne**, Australia. 04/12/13–04/20/13. (Host: Anthony Wirth)
- **Aarhus University**, Denmark. 05/17/13–05/25/13. (Host: Joshua Brody)

Mentorship and Supervision

Postdocs supervised at Indiana University

- Samson Zhou, 2018–2019 → postdoc at Carnegie Mellon University.

Ph.D. students supervised at Indiana Unviersity

- Dmitrii Avdyukhin, 2017–.
- Adithya Vadapalli, 2016–2018, joint paper in ICML'18.
- Nikolai Karpov, 2017.

Undergraduate interns mentored

- Farid Arthaud (ENS Paris-Ulm), 2019.
- Jakub Boguta (University of Warsaw), 2019.
- Stanislav Naumov (St. Petersburg ITMO University), 2019.

Ph.D. students mentored while a postdoc

- Sepehr Assadi (Univeristy of Pennsylvania), joint paper in SODA'16. → Assistant Professor at Rutgers, Computer Science.
- Yang Li (University of Pennsylvania), joint paper in SODA'16. → Research Scientist at Facebook, NYC
- Steven Wu (University of Pennsylvania), joint paper in PNAS'16. → Assistant Professor at University of Minnesota, Computer Science
- Tselil Schramm (UC Berkeley), joint paper in STOC'15. → Assistant Professor at Stanford University, Statistics
- Eli Fox-Eppstein (Brown University).
- David Meierfrankenfeld (Brown Univeristy).

Patents

- "A Communication and Message-Efficient Protocol for Computing the Intersection Between Different Sets of Data", with David P. Woodruff. U.S. patent #9438704. IBM Almaden Research Center, San Jose, CA.

Reviewing

Theory conferences

- STOC (ACM Symposium on the Theory of Computing): 2021, 2019, 2018, 2017, 2016, 2015.
- FOCS (IEEE Symposium on Foundations of Computer Science): 2023, 2019, 2017, 2015, 2014, 2013, 2012.
- SODA (ACM Symposium on Discrete Algorithms): 2020, 2019, 2018, 2016, 2013, 2012.
- ICALP (International Colloquium on Automata, Languages and Programming): 2018, 2017, 2015, 2014, 2013.
- CCC (Conference on Computational Complexity): 2018, 2016.
- ITCS (Innovations in Theoretical Computer Science): 2018.
- RANDOM (Workshop on Randomization and Computation): 2017, 2015, 2014.
- APPROX (Workshop on Approximation Algorithms for Combinatorial Optimization Problems): 2012.
- ESA (European Symposium on Algorithms): 2015.

- MFCS (Symposium on Mathematical Foundations of Computer Science): 2013, 2010.

Learning theory conferences

- COLT (Conference on Learning Theory): 2016.
- ALT (Conference on Algorithmic Learning Theory): 2014.

Databases and large-scale data processing conferences

- PODS (ACM Symposium on Principles of Database Systems): 2018, 2017, 2016.
- SPAA (ACM Symposium on Parallelism in Algorithms and Architectures): 2017.
- VLDB (Conference on Very Large Databases): 2012.
- CIKM (ACM International Conference on Information and Knowledge Management): 2014.

Journals

- SICOMP (SIAM Journal on Computing)
- CSUR (ACM Computing Surveys)
- I&C (Information and Computation)
- TKDE (IEEE Transactions on Knowledge and Data Engineering)
- ToC (Theory of Computing)
- RSA (Random Structures and Algorithms)
- Algorithmica

References

Sofya Raskhodnikova

Ph.D. advisor
Professor and Assc Chair of the Faculty
Boston University
Dept. of Computer Science

Michael Kearns

Postdoctoral mentor
Professor and National Center Chair
University of Pennsylvania
Dept. of Computer and Information Science
+ Economics (Wharton)
+ Statistics (Wharton)
+ Operations, Information & Decisions (Wharton)

Elchanan Mossel

Postdoctoral mentor
Professor
Massachusetts Institute of Technology
Dept. of Mathematics

Moses Charikar

Collaborator
Donald E. Knuth Professor
Stanford University
Dept. of Computer Science
+ Mathematics

David Woodruff

Internship mentor
Associate Professor (with tenure)
Carnegie Mellon University
Dept. of Computer Science

Sampath Kannan

Postdoctoral mentor
Henry Salvatori Professor
University of Pennsylvania
Dept. of Computer and Information Science