Grigory Yaroslavtsev, http://grigory.us

CONTACT INFORMATION
361 Levine Hall, 3330 Walnut Street
Philadelphia, PA, 19104-6389

Cell phone: +1 (814) 713-1096
E-mail: grigory@grigory.us

RESEARCH INTERESTS
Approximation and sublinear time algorithms for big data: sparsification, summarization, private data release, property testing.

2014– **University of Pennsylvania**, Philadelphia, PA.

Postdoctoral Fellowship at the Warren Center for Network and Data Sciences, hosted by the departments of Computer and Information Sciences and Statistics at the Wharton School.

2013–2014 **Brown University ICERM**, Providence, RI.

Institute Postdoctoral Fellowship in Mathematics.

2010–2013 **Pennsylvania State University**, State College, PA (Joined by invitation, didn't apply to any other Ph.D. programs)

Ph.D., Thesis: "Efficient Combinatorial Techniques in Sparsification, Summarization and Testing of Large Datasets."

2008–2010 **Academic University of the Russian Academy of Sciences**, St. Petersburg, Russia

M.S. in Applied Mathematics and Physics (1^{st} student in the pilot class, joint with Steklov Institute), GPA: 4.9/5.0.

2004–2008 **St. Petersburg State Polytechnic University**, St. Petersburg, Russia

B.S. in Physics and Technology. (1^{st} result in the admission test for the department).

RESEARCH INTERNSHIPS

Microsoft Research, Redmond, May 2013 – August 2013.
Theory group, mentored by Konstantin Makarychev.

- Improved approximation algorithms for correlation clustering (with S. Chawla, K. Makarychev, T. Schramm, STOC'15).

Microsoft Research, SVC, August 2012 – October 2012.
Theory group, mentored by Alexandr Andoni.

- Parallel algorithms for large-scale geometric problems (with A. Andoni, A. Nikolov and K. Onak, STOC'14).

IBM Research, Almaden, May 2012 – July 2012.
Theory group, mentored by David P. Woodruff.

- Optimal direct-sum theorem for one-way communication complexity, showing that parallel repetition is optimal for solving multiple instances of problems, such as augmented indexing (with M. Molinaro and D. Woodruff, SODA'13).
- Almost optimal round vs. communication protocol for computing the intersection of distributed databases (with D. Woodruff, PODC'14, U.S. patent pending).

AT&T Labs — Research, May 2011 — August 2011.
Database theory group, mentored by Graham Cormode, Cecilia M. Procopiuc, Divesh Srivastava and Howard Karloff.

- Design and implementation of efficient differentially private mechanisms for linear queries (with G. Cormode, M. Procopiuc and D. Srivastava, ICDE'13)

Achievements and awards

- **Warren Center Postdoctoral Fellowship** at University of Pennsylvania, 2014 —.
- **Institute Postdoctoral Fellowship in Mathematics** at Brown ICERM, 2013 — 2014.
- **Best Graduate Research Assistant at Computer Science and Engineering Department**, 2012.
- **TopCoder Open Algorithm Competition onsite finalist (Top 24 worldwide, handle "griffon")**, 2010.
- **College of Engineering Fellowship**, 2010 — 2013.
- **University Graduate Fellowship**, 2010 — 2011.
- **Yandex personal research grant**, 2009 — 2010.
- Travel awards: ICDE'13, SODA'13, FOCS'12, STOC'12, STOC'11, ICALP'11, SAT'09.
- Diploma for coaching the best team in St. Petersburg Olympiad in Informatics and Programming for high-school students, 2008.
- 2^{nd} place in St. Petersburg State Polytechnic University Olympiad in Mathematics, 2005.
- Best result in the admissions test for the Department of Physics and Technology in St. Petersburg State Polytechnic University, 2004.

Journal Papers

- **Certifying Equality with Limited Interaction**, with Joshua Brody, Amit Chakrabarti, Ranganath Kondapally and David Woodruff.

  Algorithmica, special issue on "Information Complexity and Applications", to appear.

- **Private Analysis of Graph Structure**, with Vishesh Karwa, Sofya Raskhodnikova and Adam Smith.

  ACM Transactions on Database Systems, Volume 39, Issue 3, 2014.

- **Steiner Transitive-Closure Spanners of Low-Dimensional Posets**, with Piotr Berman, Arnab Bhattacharyya, Elena Grigorescu, Sofya Raskhodnikova and David Woodruff.

  Combinatorica, June 2014, Volume 34, Issue 3, pages 255-277.

- **Approximation Algorithms for Spanner Problems and Directed Steiner Forest**, with Piotr Berman, Arnab Bhattacharyya, Konstantin Makarychev and Sofya Raskhodnikova.

  Information and Computation, special issue for ICALP'11. Volume 222, 2013, pp. 93-107.

- **New upper bounds on the Boolean Circuit Complexity of Symmetric Functions**, with Eugeny Demenkov, Arist Kojevnikov and Alexander Kulikov.

  Information Processing Letters, 110, pp. 264-267, Elsevier, 2010.

Conference Papers

- **Tight Bounds on Linear Sketches of Approximate Matchings**, with Sepehr Assadi, Sanjeev Khanna and Yang Li.

  SODA 2016 (27th Annual ACM-SIAM Symposium on Discrete Algorithms).

- **Near Optimal LP Rounding Algorithm for Correlation Clustering on Complete and Complete k-partite Graphs**, with Shuchi Chawla, Konstantin Makarychev and Tselil Schramm.

  STOC 2015 (47th ACM Symposium on the Theory of Computing).

- **Certifying Equality with Limited Interaction**, with Joshua Brody, Amit Chakrabarti, Ranganath Kondapally and David Woodruff.

  RANDOM 2014 (18th International Workshop on Randomization and Computation).

  Available upon request.

- **Beyond Set Disjointness: The Communication Complexity of Finding the Intersection**, with Joshua Brody, Amit Chakrabarti, Ranganath Kondapally and David Woodruff.

  PODC 2014 (33rd Annual ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing).

- **Parallel Algorithms for Geometric Graph Problems**, with Alexandr Andoni, Krzysztof Onak and Aleksandar Nikolov.

  STOC 2014 (46th ACM Symposium on the Theory of Computing). Available upon request.Tech report: ArXiv:1401.0042.

- L_p**-testing**, with Piotr Berman and Sofya Raskhodnikova.

  STOC 2014 (46th ACM Symposium on the Theory of Computing). Available upon request.

- **Lower Bounds for Testing Properties of Functions over Hypergrid Domains**, with Eric Blais and Sofya Raskhodnikova.

  CCC 2014 (29th IEEE Conference on Computational Complexity). Available upon request. Tech report: ECCC TR13-036.

- [1]**Accurate and Efficient Private Release of Datacubes and Contingency Tables**. Grigory Yaroslavtsev, Graham Cormode, Cecilia M. Procopiuc and Divesh Srivastava.

  ICDE 2013 (29th IEEE International Conference on Data Engineering). Available as ArXiv:1207.6096.

- **Beating the Direct Sum Theorem in Communication Complexity with Implications for Sketching**, with Marco Molinaro and David Woodruff.

  SODA 2013 (24th Annual ACM-SIAM Symposium on Discrete Algorithms).

- **Learning Pseudo-Boolean k-DNF and Submodular Functions**, with Sofya Raskhodnikova.

  SODA 2013 (24th Annual ACM-SIAM Symposium on Discrete Algorithms).

- **Primal-dual algorithms for Node-Weighted Network Design in Planar Graphs**, with Piotr Berman.

  APPROX 2012 (15th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems).

- **Private Analysis of Graph Structure**, with Vishesh Karwa, Sofya Raskhodnikova and Adam Smith.

  VLDB 2011 (37th International Conference on Very Large Data Bases), Research track.

- **Improved Approximation for the Directed Spanner Problem**, with Piotr Berman, Arnab Bhattacharyya, Konstantin Makarychev and Sofya Raskhodnikova.

  ICALP 2011 (38th International Colloquium on Automata, Languages and Programming), Track A.

[1]This is the only paper with non-alphabetical ordering of authors

Runner-up for the Best Paper Award, invited to a special issue of a journal "Information and Computation".

- **Steiner Transitive-Closure Spanners of Low-Dimensional Posets**, with Piotr Berman, Arnab Bhattacharrya, Elena Grigorescu, Sofya Raskhodnikova and David Woodruff.

  ICALP 2011 (38th International Colloquium on Automata, Languages and Programming), Track A.

- **Finding Efficient Circuits using SAT-solvers**, with Arist Kojevnikov and Alexander Kulikov.

  SAT 2009 (12th International Conference on Theory and Applications of Satisfiability Testing).

Research visits

- 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, 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)

Selected Talks

- Fast Fourier Sparsity Testing over the Boolean Hypercube
  - University of Wisconsin, Madison. August 2015.
- Near Optimal LP Rounding for Correlation Clustering
  - Cornell University, Ithaca, NY. Theory Seminar. May 2015.
  - MIT, Boston, MA. Algorithms and Complexity Seminar. April 2015.
  - Microsoft Research, Redmond, WA. Theory Seminar. March 2015.
  - Google Research, NYC. Google Tech Talk. February 2015.
  - Rutgers University, New Brunswick, NJ. Theory Seminar. January 2015.
  - Carnegie Mellon University, Pittsburgh, PA. Theory Lunch. January 2015.
  - Pennsylvania State University, State College, PA. CSE Departmental Colloquium. January 2015.
- Beyond Set Disjointness: The Communication Complexity of Finding the Intersection
  - MIT, Boston, MA. Theory of Distributed Systems Seminar. May 2014.
- "The Big Data Theory" and Randomized Algorithms
  - Georgia Tech, Atlanta, GA. March 2014.
- Approximating Graph Problems: The Old and The New
  - Yahoo! Research, NYC. February 2014.
  - MIT, Boston, MA. Algorithms and Complexity Seminar. February 2014.
  - Toyota Technological Institute, Chicago IL. February 2014.
- L_p-Testing
  - University of Pennsylvania, Statistics Student Seminar, November 2014.
  - Columbia University, New York, NY. Theory Seminar. October 2014.
  - Microsoft Research, Redmond. Theory Lunch. January 2014.
  - Harvard University, Boston MA. Theory Seminar. November 2013.
  - Brown University, Providence RI. Theory Seminar. November 2013.
  - IBM Almaden Research Center, San Jose, CA. October 2013.
- Property Testing and Communication Complexity
  - MIT, Boston, MA. Algorithms and Complexity Seminar. September 2013.

- Accurate and Efficient Private Release of Data Cubes and Contingency Tables
  - Cornell University, CDI Project Meeting. May 2013.
- Beating the Direct Sum in Communication Complexity with Implications for Sketching.
  - Aarhus University, Denmark. Theory Seminar. May 2013.
  - MIT, Boston, MA. Algorithms and Complexity Seminar. December 2012.
  - Princeton University, Princeton, NJ. Theory lunch. November 2012.
- Parallel Algorithms for Geometric Problems
  - 22nd International Symposium on Mathimatical Programming (ISMP 2015). July 2015.
  - Johns Hopkins University, Baltimore, MD. Theory Seminar. November 2014.
  - University of Maryland, College Park, MD. Capital Area Theory Seminar. October 2014.
  - University of Pennsylvania, Philadelphia, PA. Theory Seminar. August 2014.
  - University of Massachusetts, Amherst. Theory Seminar. May 2014.
  - Google Research, NYC. Google Tech Talk. March 2014.
  - Sandia Labs, Livermore, CA. March 2014.
  - Stanford University, Stanford, CA. Theory Seminar. March 2014.
  - Microsoft Research SVC, Mountain View, CA. Lab Meeting. October 2012.
- Learning and Testing Submodular Functions.
  - Microsoft Research, Redmond. Theory Seminar. June 2013.
  - University Of Melbourne, Theory Seminar, April 2013.
  - UCLA, Los Angeles, LA. Theory Seminar. February 2013.
  - Weizmann Institute of Science, Rehovot, Israel. Theory Seminar. December 2012.
  - Harvard University, Boston, MA. Theory Seminar. December 2012.
  - Carnegie-Mellon University, Pittsburgh, PA. Theory Lunch, December 2012.
  - Carnegie-Mellon University, Pittsburgh, PA. Operations Research Seminar. December 2012.
  - IBM T.J. Watson Research Center, Yorktown Heights, NY. Integer Programming for Lunch. November 2012.
  - Columbia University, New York, NY. Theory Seminar. October 2012.
  - Microsoft Research SVC, Mountain View, CA. Theory Seminar. October 2012.
  - IBM Almaden Research Center, San Jose, CA. Theory Seminar. May 2012.
- Advances in Directed Spanners.
  - University of Sydney, Theory Seminar, April 2013.
  - Carnegie-Mellon University, Theory Lunch, November 2011.
  - University of Maryland, Capital Area Theory Seminar, November 2011.
- Private Analysis of Graph Structure
  - AT&T Labs — Research, Florham Park, NJ. August 2011.
- Improved Approximation for the Directed Spanner Problem
  - AT&T Labs — Research, Florham Park, NJ. Mathematics Research Colloquium and Informal Seminar. June 2011.
  - Moscow State University. Combinatorial Optimization Seminar. May 2011.
  - IBM T.J. Watson Research Center, Yorktown Heights, NY. IP for lunch. April 2011.
  - St. Petersburg Institute of Fine Mechanics and Optics. Theory Seminar. December 2010.
- Linear bounds on circuit complexity and feebly one-way permutations
  - Pennsylvania State University, State College, PA. Theory Seminar. April 2010.

Posters

- Beating the Direct Sum Theory in Communication Complexity with Implications for Sketching
  - FOCS 2012, Rutgers University, NJ. October 2012.

- Parallel Algorithms for Geometric Graph Problems
  - FOCS 2012, Rutgers University, NJ. October 2012.
- Overlapping Clustering with Qualitative Information
  - FOCS 2012, Rutgers University, NJ. October 2012.
- Learning and Testing Submodular Functions
  - New York Computer Science and Economics Day, New York, NY. December 2012.
  - FOCS 2012, Rutgers University, NJ. October 2012.
  - EPFL, Lausanne, Switzerland. Algorithmic Frontiers Workshop. June 2012.
  - STOC 2012, New York, NY. May 2012.
- Private Analysis of Graph Structure
  - EPFL, Lausanne, Switzerland. Algorithmic Frontiers Workshop. June 2012.
- Improved Approximation for the Directed Spanner Problem
  - STOC 2011, San Jose, CA. June 2011.

Teaching

Computational Learning Theory

- **University of Pennsylvania**, CIS 625, Spring 2015 (co-teaching with Michael Kearns)

15-hour crash course "Sublinear Algorithms for Big Data":

- **University of Buenos Aires**, Argentina. July – August 2014.

Organized reading groups (running meetings / selection of material) at University of Pennsylvania and Pennsylvania State University.

- Spring 2015: The Big Data Reading Group (based on the survey Sketching as a Tool for Numerical Linear Algebra by David Woodruff)
- Fall 2014: The Big Data Reading Group (based on recent results on distributed and massively parallel algorithms).
- Spring 2013: "Computer Science for the Information Age", based on an eponimous book by John Hopcroft and Ravi Kannan.
- Fall 2011: Selected lectures from classes Analysis of Boolean Functions by Ryan O'Donnell and The PCP Theorem and Hardness of Approximation by Venkatesan Guruswami and Ryan O'Donnell at CMU.
- Spring 2011: "A Theorist's Toolkit", based on notes for a class taught by Sanjeev Arora at Princeton.

Extracurricular education for high-school students:

- Prepared training contests for the United States team in 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).
- Judge for Baltic Science and Engineering Contest (Intel ISEF semifinals), 2010.

Patents

- "A Communication and Message-Efficient Protocol for Computing the Intersection Between Different Sets of Data", with David P. Woodruff. U.S. patent application #20150172425. IBM Almaden Research Center, San Jose, CA.

Service

Program comittees:

- 41st International Conference on Current Trends in Theory and Practice of Computer Science (SOFSEM'15), Foundations of Computer Science Track. Pec pod Snezkou, Czech Republic.
- "Graph Theory and Applications" CSEDays 2012, Theory track. Ekaterinburg, Russia.

Reviewing:

- Journals: SIAM Journal on Computing (SICOMP), SIAM Journal on Discrete Mathematics (SIDMA), Information and Computation (I&C), IEEE Transactions on Knowledge and Data Engineering (TKDE), Theory of Computing (ToC), Random Structures and Algorithms (RSA), Algorithmica.
- Conferences: SODA'16, ESA'15, RANDOM'15, FOCS'15, ICALP'15, STOC'15, CIKM'14, ALT'14, RANDOM'14, FOCS'14, ICALP'14, FOCS'13, MFCS'13, ICALP'13, SODA'13, APPROX'12, FOCS'12, COCOA'12, SWAT'12, SODA'12, VLDB'12, WADS'11, MFCS'10, SAT'10.

Organizer:

- Big Data through the Lens of Sublinear Algorithms: 2-Day workshop at Rutgers University, DIMACS, Aug 2015.
- Algorithmic Frontiers of Modern Massively Parallel Computation: 1-Day workshop at ACM Federated Computing Research Conference, Jun 2015.
- Theory Seminar at the Univeristy of Pennsylvania, Computer and Information Sciences Department (2014 – 2015).
- Sublinear Algorithms and Big Data Day at Brown University, Institute for Computational and Experimental Research in Mathematics.
- Theory Seminar at Brown CS Department and Brown University, Institute for Computational and Experimental Research in Mathematics (2013 – 2014).

PROFESSIONAL MEMBERSHIP: ACM SIGACT, IEEE

TECHNICAL SKILLS: C/C++, STL, Java, Windows/Linux, LaTeX, CPlex/Gurobi/AMPL.