Alexandr Andoni

Title: Associate Professor, Dept. of Computer Science
Website: http://www.mit.edu/~andoni
Phone: 212 853 0685, Email: andoni@cs.columbia.edu

Field of Specialization

Theoretical Computer Science, with a particular focus on: algorithmic foundations of massive data, sublinear algorithms, high-dimensional computational geometry, and metric embeddings.

Education

Massachusetts Institute of Technology — Cambridge, MA

2005–'09 PhD in Computer Science. Research adviser: Piotr Indyk.
Thesis title: *Nearest Neighbor Search: the Old, the New, and the Impossible.*
Committee: Piotr Indyk, Robert Krauthgamer, and Ronitt Rubinfeld.

2004–'05 Master of Engineering in Electrical Engineering and Computer Science.
Thesis title: *Approximate Nearest Neighbor Problem in High Dimensions.* Supervised by Piotr Indyk.

2001–'04 Bachelor of Science degree in Computer Science and Engineering, and Bachelor of Science degree in Mathematics. Departmental GPA: 5.0/5.0. Overall GPA: 4.9/5.0.

Politehnica University of Bucharest — Bucharest, Romania

1999–'01 Department of Computer Science and Automated Control. GPA: 9.89/10.0. Transferred to MIT.

Career History

Columbia University, Dept. of Computer Science — New York, NY

2015– Associate Professor. Member of the Data Science Institute.

UCB/Simons Institute for the Theory of Computing — Berkeley, CA

2014–'15 Visiting scientist.
Long-term participant in: Theoretical Foundations of Big Data Analysis (Fall'13), Algorithmic Spectral Graph Theory (Fall'14), Information Theory (Spring'15).

Microsoft Research Silicon Valley — Mountain View, CA

2010–'14 Researcher.

Princeton U/Center for Computational Intractability — Princeton, NJ

2009–'10 Postdoctoral researcher. Hosts: Sanjeev Arora (Princeton) and Assaf Naor (NYU).

Microsoft Research Silicon Valley — Mountain View, CA

Summer'08 Research intern.

IBM Almaden Research Center — San Jose, CA

Summer'07, '06 Research intern.

Google Inc. — Mountain View, CA

Summer'05 Summer Intern (software engineer).

Palo Alto Research Center, Computer Science Lab — Palo Alto, CA

Summer'04 Research intern.

Microsoft Corporation — Redmond, WA

Summer'03 — Software Design Engineer intern in the Security Group.

MIT, CSAIL — Cambridge, MA

Spring'04, June'02 — Undergraduate researcher for Piotr Indyk.

2001–'02 — Undergraduate researcher for Martin Rinard (co-supervised by Darko Marinov and Sarfraz Khurshid).

Teaching

Columbia University, Dept. of Computer Science — New York, NY

Spring'17 — COMS W4995-3 *Advanced Algorithms*. Enrollment: 52 (as of Feb 1, 2017).

Columbia University, Dept. of Computer Science — New York, NY

Fall'16 — CSOR 4231-2 *Analysis of Algorithms*. Enrollment: 104. Evaluation of course overall quality: 4.47 mean. Evaluation of instruction overall quality: 4.50 mean.

Columbia University, Dept. of Computer Science — New York, NY

Fall'15 — COMS 6998-9 *Algorithmic Techniques for Massive Data*. Enrollment: 20. Evaluation of course overall quality: 4.38 mean. Evaluation of instruction overall quality: 4.54 mean.

MADALGO Center for Massive Data Algorithmics — Aarhus, Denmark

August'15 — Taught lectures on sketching and nearest neighbor search during the *MADALGO Summer School on Streaming*. (By invitation.)

International school of Mathematics "Guido Stampacchia" — Erice, Italy

September'14 — Taught lectures on "Sampling in Graphs" at the *Graph Theory, Algorithms and Applications (3rd edition)* summer school. (By invitation.)

University of Copenhagen — Copenhagen, Denmark

July'14 — Taught lectures on high dimensional geometry at the *Summer School on Hashing: Theory and Practice*. (By invitation.)

Moscow State University — Moscow, Russia

August'13 — Taught lectures on "Sketching, Sampling, and other Sublinear Algorithms" at the *School on ALgorithms for MAssive DAta (ALMADA)*. (By invitation.)

MADALGO Center for Massive Data Algorithmics — Aarhus, Denmark

August'11 — Taught lectures on "Embedding and Sketching" during the *MADALGO & CTIC Summer School on High-Dimensional Geometric Computing*. (By invitation.)

MIT, EECS Department — Cambridge, MA

Fall'07 — Teaching Assistant (Head TA) for 6.046, *Introduction to Algorithms*, taught by Ronitt Rubinfeld and Madhu Sudan.

Fall'04 — Teaching Assistant for 6.854, *Advanced Algorithms*, taught by David Karger.

Oct'04–Sep'06 — Coach of MIT's team for the ACM International Collegiate Programming Contest (in a committee of 2–4 coaches).

National Center for Information Technology (CNTI) — Chişinău, Moldova

Oct'98–Feb'99, Aug'99, Aug'00 — Coach of Moldova's team for International Olympiads in Informatics (usually in a committee of 2-3 coaches). Gave lectures and organized training contests.

Students supervised

Current students:

- Peilin Zhong, starting Fall'16 (co-supervised);
- Sandip Sinha, starting Fall'16 (co-supervised);
- Emmanouil Vasileios Vlatakis Gkaragkounis, starting Fall'16 (co-supervised).

Past interns (mentored at MSR Silicon Valley, 2011-2014):

- Huy L Nguyen, 2011 (now faculty at Northeastern University);
- Grigory Yaroslavtsev, 2012 (now faculty at University of Indiana);
- Amirali Abdullah, 2013 (postdoc at U. Michigan, now at Qualtrics);
- Ilya Razenshteyn, 2014 (final-year student, already awarded a 3-year Junior Fellowship of the Simons Society of Fellows).

Student research projects, with academic units (at Columbia University):

- Flora M. Park (undergraduate), Fall'16–Spring'17;
- Nishanth Mohan (MS), Spring'17;
- Negev Shekel-Nosatzki (MS), Spring'16 (results published as C39).

Publications: conference proceedings

C43. **Approximate Near Neighbors for General Symmetric Norms**
by Alexandr Andoni, Aleksandar Nikolov, Ilya Razenshteyn, Erik Waingarten.
In **STOC** *(Symposium on Theory of Computation)*, 2017.

C42. **High Frequency Moments via Max-Stability**
by Alexandr Andoni.
In **ICASSP** *(International Conference on Acoustics, Speech, and Signal Processing)*, special session on *Random Embeddings and Geometry-Preserving Dimensionality Reduction*, 2017.

C41. **Optimal Hashing-based Time-Space Trade-offs for Approximate Near Neighbors**
by Alexandr Andoni, Thijs Laarhoven, Ilya Razenshteyn, Erik Waingarten.
In **SODA** *(Symposium on Discrete Algorithms)*, 2017.
Invited to **T.Alg.** special issue.

C40. **LSH Forest: Practical Algorithms Made Theoretical**
by Alexandr Andoni, Ilya Razenshteyn, Negev Shekel–Nosatzki.
In **SODA** *(Symposium on Discrete Algorithms)*, 2017.

C39. **Impossibility of Sketching of the 3D Transportation Metric with Quadratic Cost**
by Alexandr Andoni, Assaf Naor, Ofer Neiman.
In **ICALP** *(International Colloquium on Automata, Languages and Programming)*, 2016.

C38. **Tight Lower Bounds for Data-Dependent Locality-Sensitive Hashing**
by Alexandr Andoni, Ilya Razenshteyn.
In **SoCG** *(International Symposium on Computational Geometry)*, 2016.

C37. **On Sketching Quadratic Forms**
by Alexandr Andoni, Jiecao Chen, Robert Krauthgamer, Bo Qin, David P. Woodruff, and Qin Zhang.
In **ITCS** *(Innovations in Theoretical Computer Science)*, 2016.

C36. **Interacting with Large Distributed Datasets Using Sketch**
by Mihai Budiu, Rebecca Isaacs, Derek Murray, Gordon Plotkin, Paul Barham, Samer Al-Kiswany, Yazan Boshmaf, Qingzhou Luo, Alexandr Andoni.
In Eurographics Symposium on Parallel Graphics and Visualization, 2016.

C35. **Practical and Optimal LSH for Angular Distance**
by Alexandr Andoni, Piotr Indyk, Thijs Laarhoven, Ilya Razenshteyn, and Ludwig Schmidt.
In **NIPS** *(Conference on Neural Information Processing Systems)*, 2015.

C34. **Optimal Data-Dependent Hashing for Approximate Near Neighbors**
by Alexandr Andoni, Ilya Razenshteyn.
In **STOC** *(Symposium on Theory of Computation)*, 2015.

C33. **Sketching and Embedding are Equivalent for Norms**
by Alexandr Andoni, Robert Krauthgamer and Ilya Razenshteyn.
In **STOC** *(Symposium on Theory of Computation)*, 2015.
Invited to **SICOMP** special issue (in submission).

C32. **Spectral Approaches to Nearest Neighbor Search**
by Amirali Abdullah, Alexandr Andoni, Ravi Kannan, Robert Krauthgamer.
In **FOCS** *(Symposium on Foundations of Computer Science)*, 2014.

C31. **Learning Polynomials with Neural Networks**
by Alexandr Andoni, Rina Panigrahy, Gregory Valiant, Li Zhang.
In **ICML** *(International Conference on Machine Learning)*, 2014.

C30. **Parallel Algorithms for Geometric Graph Problems**
by Alexandr Andoni, Aleksandar Nikolov, Krzysztof Onak, Grigory Yaroslavtsev.
In **STOC** *(Symposium on Theory of Computation)*, 2014.

C29. **Beyond Locality Sensitive Hashing**
by Alexandr Andoni, Piotr Indyk, Huy L. Nguyen, Ilya Razenshteyn.
In **SODA** *(Symposium on Discrete Algorithms)*, 2014.

C28. **Towards $(1+\epsilon)$-Approximate Flow Sparsifiers**
by Alexandr Andoni, Robert Krauthgamer, Anupam Gupta.
In **SODA** *(Symposium on Discrete Algorithms)*, 2014.

C27. **Learning Sparse Polynomial Functions**
by Alexandr Andoni, Rina Panigrahy, Gregory Valiant, Li Zhang.
In **SODA** *(Symposium on Discrete Algorithms)*, 2014.

C26. **Tight Lower Bound for Linear Sketches of Moments**
by Alexandr Andoni, Huy L. Nguyen, Yury Polyanskiy, Yihong Wu.
In **ICALP** *(International Colloquium on Automata, Languages and Programming)*, 2013.

C25. **Homomorphic Fingerprints under Misalignments: Sketching Edit and Shift Distances**
by Alexandr Andoni, Assaf Goldberger, Andew McGregor, Ely Porat.
In **STOC** *(Symposium on Theory of Computation)*, 2013.

C24. **Shift Finding in Sub-linear Time**
by Alexandr Andoni, Haitham Hassanieh, Piotr Indyk, Dina Katabi.
In **SODA** *(Symposium on Discrete Algorithms)*, 2013.

C23. **Eigenvalues of a Matrix in the Streaming Model**
by Alexandr Andoni, Huy L. Nguyen.
In **SODA** *(Symposium on Discrete Algorithms)*, 2013.

C22. **Width of Points in the Streaming Model**
by Alexandr Andoni, Huy L. Nguyen.
In **SODA** *(Symposium on Discrete Algorithms)*, 2012.
Invited to **T.Alg.** special issue (appears as J6).

C21. **Streaming Algorithms via Precision Sampling**
by Alexandr Andoni, Robert Krauthgamer, Krzysztof Onak.
In **FOCS** *(Symposium on Foundations of Computer Science)*, 2011.

C20. **Near Linear Lower Bounds for Dimension Reduction in L1**
by Alexandr Andoni, Moses Charikar, Ofer Neiman, Huy L. Nguyen.
In **FOCS** *(Symposium on Foundations of Computer Science)*, 2011.

C19. **Polylogarithmic Approximation to Edit Distance and Asymmetric Query Complexity**
by Alexandr Andoni, Robert Krauthgamer, Krzysztof Onak.
In **FOCS** *(Symposium on Foundations of Computer Science)*, 2010.
Invited to **SIAM J. Comp.** special issue (declined); full version as `arxiv.org/abs/1005.4033`.

C18. **Global Alignment of Molecular Sequences via Ancestral State Reconstruction**
by Alexandr Andoni, Constantinos Daskalakis, Avinatan Hassidim, Sebastien Roch.
In **ICS** *(Innovations in Computer Science)*, 2010.

C17. **Lower bounds for Edit Distance and Product Metrics via Poincare-Type Inequalities**
by Alexandr Andoni, T.S. Jayram, Mihai Pătraşcu.
In **SODA** *(Symposium on Discrete Algorithms)*, 2010.

C16. **Near-optimal Sublinear Time Algorithms for Ulam Distance**
by Alexandr Andoni, Huy L. Nguyen.
In **SODA** *(Symposium on Discrete Algorithms)*, 2010.

C15. **Efficient sketches for Earth-Mover Distance, with applications**
by Alexandr Andoni, Khanh Do Ba, Piotr Indyk, David Woodruff.
In **FOCS** *(Symposium on Foundations of Computer Science)*, 2009.

C14. **External Sampling**
by Alexandr Andoni, Piotr Indyk, Krzysztof Onak, Ronitt Rubinfeld.
In **ICALP** *(International Colloquium on Automata, Languages and Programming)*, 2009.

C13. **Approximating Edit Distance in Near-Linear Time**
by Alexandr Andoni, Krzysztof Onak.
In **STOC** *(Symposium on Theory of Computation)*, 2009.
Invited to **SIAM J. Comp.** special issue (appears as J5).

C12. **Approximate Line Nearest Neighbor in High Dimensions**
by Alexandr Andoni, Piotr Indyk, Robert Krauthgamer, Huy L. Nguyen.
In **SODA** *(Symposium on Discrete Algorithms)*, 2009.

C11. **Overcoming the L1 Non-Embeddability Barrier: Algorithms for Product Metrics**
by Alexandr Andoni, Piotr Indyk, Robert Krauthgamer.
In **SODA** *(Symposium on Discrete Algorithms)*, 2009.

C10. **Hardness of Nearest Neighbor under ℓ_∞**
by Alexandr Andoni, Dorian Croitoru, Mihai Pătraşcu.
In **FOCS** *(Symposium on Foundations of Computer Science)*, 2008.
Invited to **Discrete & Computational Geometry** (declined).

C9. **The Smoothed Complexity of Edit Distance**
by Alexandr Andoni, Robert Krauthgamer.
In **ICALP** *(International Colloquium on Automata, Languages and Programming)*, 2008.
Journal version appears as J4.

C8. **Earth Mover Distance over High-Dimensional Spaces**
by Alexandr Andoni, Piotr Indyk, Robert Krauthgamer.
In **SODA** *(Symposium on Discrete Algorithms)*, 2008.

C7. **The Computational Hardness of Estimating Edit Distance**
by Alexandr Andoni, Robert Krauthgamer.
In **FOCS** *(Symposium on Foundations of Computer Science)*, 2007.
Invited to **SIAM J. Comp.** special issue (appears as J2).

C6. **Testing k-wise and Almost k-wise Independence**
by Noga Alon, Alexandr Andoni, Tali Kaufman, Kevin Matulef, Ronitt Rubinfeld, Ning Xie.
In **STOC** *(Symposium on Theory of Computation)*, 2007.

C5. **Near-optimal Hashing Algorithms for Approximate Nearest Neighbor in High Dimensions**
by Alexandr Andoni, Piotr Indyk.
In **FOCS** *(Symposium on Foundations of Computer Science)*, 2006.
Invited to **C.ACM** Research Highlights (appears as J1).

C4. **On Optimality of the Dimensionality Reduction Method**
by Alexandr Andoni, Piotr Indyk, Mihai Pătraşcu.
In **FOCS** *(Symposium on Foundations of Computer Science)*, 2006.

C3. **Efficient Algorithms for Substring Near Neighbor Problem**
by Alexandr Andoni, Piotr Indyk.
In **SODA** *(Symposium on Discrete Algorithms)*, 2006.

C2. **Graceful Service Degradation (or, How to Know your Payment is Late)**
by Alexandr Andoni, Jessica Staddon.
In **EC** *(Conference on Electronic Commerce)*, 2005.

C1. **Lower Bounds for Embedding of Edit Distance into Normed Spaces**
by Alexandr Andoni, Michel Deza, Anupam Gupta, Piotr Indyk, Sofya Raskhodnikova.
In **SODA** *(Symposium on Discrete Algorithms)*, 2003.

Publications: Book chapters

B3. **Nearest neighbors in high-dimensional spaces**
by Alexandr Andoni, Piotr Indyk.
Book chapter in *Handbook of Discrete and Computational Geometry (3rd edition)*, Jacob E. Goodman, Joseph O'Rourke, and Csaba D. Tóth (eds), CRC Press LLC, to appear, 2016.

B2. **High-dimensional computational geometry**
by Alexandr Andoni.
Book chapter in *Handbook on Big Data*, Peter Buhlmann, Petros Drineas, Michael Kane, Mark van der Laan (eds.), CRC Press, 2016.

B1. **Locality-sensitive hashing using stable distributions**
Alexandr Andoni, Mayur Datar, Nicole Immorlica, Piotr Indyk, Vahab Mirrokni.
Book chapter in *Nearest Neighbor Methods in Learning and Vision: Theory and Practice*, T. Darrell and P. Indyk and G. Shakhnarovich (eds.), MIT Press, 2006.

Publications: journals

J7. **Universalité des espaces de Wasserstein à floconnage près (Eng. Snowflake Universality of Wasserstein Spaces)**
by Alexandr Andoni, Assaf Naor, Ofer Neiman.
Annales scientifiques de l'ENS (Eng. Scientific Annals of ENS), accepted, 2017.

J6. **Width of Points in the Streaming Model**
by Alexandr Andoni, Huy L. Nguyen.
T.Alg. (ACM Transactions on Algorithms, SODA special issue), 12(1):5, 2016.

J5. **Approximating Edit Distance in Near-Linear Time**
by Alexandr Andoni, Krzysztof Onak.
SICOMP (SIAM J. Comp., STOC special issue), 41(6):1635–1648, 2012.

J4. **The Smoothed Complexity of Edit Distance**
by Alexandr Andoni, Robert Krauthgamer.
T.Alg. (ACM Transactions on Algorithms), 8(4):44, 2012.

J3. **Global Alignment of Molecular Sequences via Ancestral State Reconstruction**
by Alexandr Andoni, Constantinos Daskalakis, Avinatan Hassidim, Sebastien Roch.
Stochastic Processes and their Applications, 122:3852–3874, 2012.

J2. **The Computational Hardness of Estimating Edit Distance**
by Alexandr Andoni, Robert Krauthgamer.
SICOMP (SIAM J. Comp., FOCS special issue), 39(6):2398-2429, 2010.

J1. **Near-optimal Hashing Algorithms for Approximate Nearest Neighbor in High Dimensions**
by Alexandr Andoni, Piotr Indyk.
CACM (Communications of the ACM), 51(1):117–122, 2008.
Over **1500 citations** in Google Scholar, along with the conference version C5.

Non-refereed Publications

M3. **Phylogenetic Reconstruction with Insertions and Deletions**
by Alexandr Andoni, Mark Braverman, Avinatan Hassidim.
Manuscript, available at `http://www.mit.edu/~andoni/papers/phylo.pdf`, 2010.

M2. **Corrigendum to "Efficient similarity search and classification via rank aggregation" by Ronald Fagin, Ravi Kumar and D. Sivakumar (proc. SIGMOD'03)**
by Alexandr Andoni, Ronald Fagin, Ravi Kumar, Mihai Pătraşcu, D. Sivakumar.
In **SIGMOD**, 2008.

M1. **An evaluation of exhaustive testing for data structures**
by Darko Marinov, Alexandr Andoni, Dumitru Daniliuc, Sarfraz Khurshid, Martin Rinard.
Technical Report MIT-LCS-TR-921, MIT CSAIL, Cambridge, MA, 2003.

Grants

Current:

- Simons Foundation: Mathematics and Physical Sciences—Simons Collaboration on Algorithms and Geometry. Role: coPI. Period: 2 years starting Jul'16. Lead PI is Assaf Naor (Princeton), and other coPIs are Sanjeev Arora (Princeton), Tim Austin (NYU), Mark Braverman (Princeton), Jeff Cheeger (NYU), Subhash Khot (NYU), Bruce Kleiner (NYU), Ran Raz (Princeton), Oded Regev (NYU), Mike Saks (Rutgers), Amit Singer (Princeton), David Steurer (Cornell).
- NSF: CCF Algorithmic Foundations. "AF:Small: Nearest Neighbor Search in High Dimensional Spaces". Role: single PI. Period: 3 years starting Sep'16.
- Google: Faculty Research Award. "Data-Dependent Hashing for Near Neighbor Search: Bridging Theory and Practice". Role: single PI. Period: starting 2016 (no expiration).

Organizational Activities

- *(upcoming)* Co-organizer of the "Fast Iterative Methods in Optimization" workshop in the "Bridging Continuous and Discrete Optimization" semester-long program at the Simons Institute for the Theory of Computing, Berkeley, CA. Oct'17.
- *(upcoming)* Organizer of the Simons Collaboration on Algorithms & Geometry Meeting (day-long workshop) on "Graph Algorithms and Continuous Optimization", New York. Mar'17.
- Co-organizer of the New York Area Theory Day (a day-long workshop twice a year), New York. Starting December'15 (two events organized already).
- Co-organizer of the FOCS'16 Workshop/Tutorial Day, New Brunswick, NJ. Oct'16.
- Co-organizer of the STOC'16 Workshop/Tutorial Day, Boston, MA. Jun'16.
- Co-organizer of the DIMACS Workshop on "Big Data through the Lens of Sublinear Algorithms", Rutgers, NJ. Aug'15.
- Co-organizer of the "Bertinoro Workshop on Sublinear Algorithms 2014", Italy. May'14.

- Co-organizer of the "Data Structures (in memory of Mihai Pătraşcu)" workshop at FOCS'12, New Brunswick, NJ. Oct'12.
- Co-organizer of the Workshop on Embeddings as part of the Discrete Analysis programme at the Isaac Newton Institute for Mathematical Sciences, Cambridge, UK. Jan'11.

Service

Service to Columbia University:

- Theory seminar organizer, Spring'15–present;
- PhD admissions committee, Spring'17–present;
- MS adviser for the Machine Learning track, Spring'16–present;
- Distinguished Lecture Series committee, Fall'16–Spring'17;
- Student nominations committee, Fall'15–present;
- Hiring committee for the CS department, Fall'15–Spring'16.

Outside Service:

- Program Committee member of: STOC'16, ESA'16, RANDOM'15, SODA'15, FOCS'13, MASSIVE'12, APPROX'12, ESA'12 (experimental track), CPM'11, STOC'11, RANDOM'10.
- *SIAM Journal of Computing (SICOMP)* guest editor for STOC'16 special issue.
- *Transactions on Algorithms (TALG)* guest editor for SODA'15 special issue.
- *Theory of Computing (Toc)* guest editor for APPROX/RANDOM'12 special issue.
- Panelist for: National Science Foundation (NSF), 2 panels.
- Reviewer for National Science Foundation (NSF), US-Israel Binational Science Foundation (BSF), Swiss National Science Foundation (SNSF).
- Reviewer for: STOC, FOCS, SODA, NIPS, ICML, SoCG, CCC, COLT, ICALP, PODS, DISC, SWAT, RANDOM, ALENEX, ESA, CPM, CIKM, RSA-CT; J. ACM, SICOMP, JoCG, TALG, Information & Computation, IEEE Info. Theory, IPL, Algorithmica, IEEE Trans. on Computers, TPAMI, Random Structures and Algorithms.

Public outreach:

- *upcoming:* Simons Foundation Lecture at the Simons Foundation, New York. May'17.

Selected Invited Talks

- Keynote talk at International Conference on Similarity Search and Applications (SISAP), Tokyo, Japan. October 2016.
- Invited talk at the *Highlights of Algorithms* conference, Paris, France. June 2016.
- "Big Data Boot Camp" lectures on "Algorithmic High-Dimensional Geometry", at the Simons Institute for the Theory of Computing, Berkeley, CA. September 2013.
- Keynote talk at International Symposium on Mathematical Foundations of Computer Science, Warsaw, Poland. August 2011.