Andrew McGregor

CONTACT INFORMATION	College of Information & Computer Sciences University of Massachusetts 140 Governors Drive, Amherst, MA, 01003	*Phone:* (415) 205-4076 *E-mail:* mcgregor@cs.umass.edu *Web:* www.cs.umass.edu/˜mcgregor

RESEARCH INTERESTS

Algorithms for processing massive data sets and data streams; computing with noisy or incomplete data; clustering; communication complexity; coding and information theory. My work appears in theoretical computer science, database, data mining, coding theory, and machine learning publications.

EMPLOYMENT

University of Massachusetts, Amherst, USA   **January 2009 – Present**
Associate Professor in the College of Information and Computer Sciences (since 2014).

University of California, Berkeley, USA   **August 2013 – December 2013**
Long-term invited participant at the Simons Institute for the Theory of Computing.

Google Inc., Mountain View, USA   **August 2013 – December 2013**
Visiting Scientist (one day a week).

Microsoft Research, Mountain View, USA   **June 2008 – January 2009**
Postdoctoral researcher in MSR-SVC.

University of California, San Diego, USA   **October 2006 – May 2008**
Postdoctoral researcher in the Information Theory and Applications Center.

Summer Internships: Bell Labs (2002, 2004, 2005); DIMACS, Rutgers University (2003); University of Glasgow, UK (1997, 1999); and Memorial University, Canada (1998).

EDUCATION

University of Pennsylvania, Philadelphia, USA   **2001 – 2006**
  Ph.D., "Processing Data Streams", August 2007
  M.S.E., Computer Science, 2002

University of Cambridge, Cambridge, UK   **1997 – 2001**
  Certificate of Advanced Study in Mathematics (Distinction), 2001
  B.A., Mathematics (1st Class Honours), 2000

HONORS AND RECOGNITION

Teaching Awards:
  College of Information and Computer Sciences Outstanding Teacher Award (2016)
  Nominated for University 2016-2017 Distinguished Teaching Award (2016)
  University of Massachusetts Lilly Teaching Fellowship (2012)

Research Awards:
  Communication of the ACM Research Highlights (2016) and SIGPLAN Research Highlights Nominated Papers (2013)
  Google Faculty Research Award (2013)
  NEC Laboratories Data Management University Award (2010)
  National Science Foundation CAREER Award (2010)
  SIGACT Significant Papers on New Areas (2007)

  Special Journal Issues for Selected Papers:
    Theory of Computer Systems for Best Papers from ICDT 2017

Trans. Database Syst. for Best Papers from SIGMOD 2011
Journal of Machine Learning for Best Papers from COLT 2007
Trans. Database Syst. for Best Papers from PODS 2007
Theoretical Computer Science for Best Papers from ICALP 2004

According to Google Scholar, my papers have been cited over 3700 times; my h-index is 32; and 14 of my papers have been cited over 100 times.

Advisee Awards:
CRA Outstanding Undergraduate Research Awards, Winner (Daniel Stubbs, 2014)
CRA Outstanding Undergraduate Research Awards, Honorable Mention (Vinay Shah, 2012 and Daniel Stubbs, 2013)
University of Massachusetts "Rising Researcher Award" (Daniel Stubbs, 2013)
NSF Graduate Fellowship (Daniel Stubbs, 2014)
NSF Graduate Fellowship, Honorable Mention (Sofya Vorotnikova, 2015)
Simons Research Fellowship (Sofya Vorotnikova, 2018)

Graduate and Undergraduate: St. Andrew's Society of the State of New York Scholarship (2001), Brown Prize in Pure Mathematics (2001), Gonville and Caius Junior and Senior Scholarship for Mathematics (1998, 2000), Bodey Prize in Applied Mathematics (1999), Daily Telegraph & BAAS Young Science Writer of the Year (1998).

GRANTS AND FUNDING

CAREER: New Directions for Sketching and Stream Computation. NSF CCF-0953754, $515,634, 04/10-03/17. REU Supplements $24,000. (Single PI)

Massive Graph Analysis via Linear Measurements: Towards a Theory of Homomorphic Compression. NSF CCF-1320719, $456,640, 09/13-08/17. (Single PI)

From Data To Users: Providing Interpretable and Verifiable Explanations in Data Mining. NSF IIS-1251110, $250,000 (University of Massachusetts portion), 09/13-08/17. (Multi-institution grant; PI at University of Massachusetts)

AitF: Efficient Memory Management via Randomized, Streaming, and Online Algorithms. NSF CCF-1637536, $500,000, 09/16-08/20. Andrew McGregor (PI) and Emery Berger (co-PI).

Summarizing and Sketching Massive Graphs. Google Faculty Research Award, $43,790, 2013. (Single PI)

Data Analytics in the Cloud: Exact Answers Fast, Approximate Answers Faster. NEC Research Gift, $40,000, 2010. Yanlei Diao (PI) and Andrew McGregor (co-PI).

BOOKS AND THESES

[B1] A. McGregor and S. Muthukrishnan. Data Stream Algorithms. Textbook in Preparation. http://people.cs.umass.edu/~mcgregor/book/book.html

[B2] A. McGregor. Processing Data Streams. Ph.D. Thesis, University of Pennsylvania, 2007. Advisor: Sampath Kannan.

JOURNAL PUBLICATIONS

[J1] A. McGregor and H. Vu. Better Approximation of The Streaming Maximum Coverage Problem. *Theory of Computing Systems*, 2018.

[J2] M. Bury, E. Grigorescu, A. McGregor, M. Monemizadeh, C. Schwiegelshohn, and S. Vorotnikova. Structural Results on Matching Estimation with Applications to

Streaming. *Algorithmica*, 2018.

[J3]   A. Chakrabarti, G. Cormode, and A. McGregor. Robust lower bounds for communication and stream computation. *Theory of Computing*, 12(10):1–35, 2016.

[J4]   D. Barowy, C. Curtsinger, E. Berger, and A. McGregor. AUTOMAN: A Platform for Integrating Human-Based and Digital Computation. *Commun. ACM*, 59(6):102–109, 2016.

[J5]   C. Li, G. Miklau, M. Hay, A. McGregor, and V. Rastogi. The matrix mechanism: optimizing linear counting queries under differential privacy. *The VLDB Journal*, 1–25, 2015.

[J6]   A. McGregor, A. Pavan, S. Tirthapura, and D. Woodruff. Space-Efficient Estimation of Statistics over Sub-Sampled Streams. *Algorithmica*, 1–25, 2015.

[J7]   A. Chakrabarti, G. Cormode, A. McGregor, and J. Thaler. Annotations in Data Streams. *ACM Transactions in Algorithms.* 11(1):1–30, 2014.

[J8]   A. Chakrabarti, G Cormode, R. Kondapally, and A. McGregor. Information Cost Tradeoffs for Augmented Index and Streaming Language Recognition. *SIAM Journal of Computing*, 42(1):61–83, 2013.

[J9]   B. Li, E. Mazur, Y. Diao, A. McGregor, and P. Shenoy. SCALLA: A Platform for Scalable One-Pass Analytics Using MapReduce. *ACM Trans. Database Syst.*, 37(4): 27, 2012.

[J10]   T. Tran, L. Peng, Y. Diao, A. McGregor, and A. Liu. CLARO: Modeling and Processing Uncertain Data Streams. *International Journal on Very Large Databases*, 21(5): 651–676, 2012.

[J11]   A. Chakrabarti, G. Cormode, and A. McGregor. A near-optimal algorithm for computing the entropy of a stream. *ACM Transactions in Algorithms* 6(3):1–21, 2010.

[J12]   A. McGregor and O. Milenkovic. On the hardness of approximating stopping and trapping sets in LDPC codes. *IEEE Trans. Inform. Theory*, 56(4):1640–1650, 2010.

[J13]   S. Guha, A. McGregor, and S. Venkatasubramanian. Sub-linear estimation of entropy and information distances. *ACM Transactions in Algorithms*, 5(4):1–16, 2009.

[J14]   S. Guha and A. McGregor. Stream order and order statistics: Quantile estimation in random-order streams. *SIAM Journal of Computing*, 38(5):2044–2059, 2009.

[J15]   J. Feigenbaum, S. Kannan, A. McGregor, S. Suri, and J. Zhang. Graph distances in the data-stream model. *SIAM Journal on Computing*, 38(5):1709–1727, 2009.

[J16]   T. S. Jayram, A. McGregor, S. Muthukrishnan, and E. Vee. Estimating statistical aggregates on probabilistic data streams. *ACM Trans. Database Syst.*, 33(4):1–30, 2008.

[J17]   S. Guha, P. Indyk, and A. McGregor. Sketching information divergences. *Journal of Machine Learning*, 72(1-2):5–19, 2008.

[J18]   J. Feigenbaum, S. Kannan, A. McGregor, S. Suri, and J. Zhang.   On graph problems in a semi-streaming model. *Theoretical Computer Science*, 348(2-3):207–216, 2005.

[J19]   A. Barg and A. McGregor. Distance distribution of binary codes and the error probability of decoding. *IEEE Trans. Inform. Theory*, 51(12):4237– 4246, 2005.

CONFERENCE
PUBLICATIONS

[C1] M. McCartin-Lim, B. Woolf, and A. McGregor. Connect the Dots to Prove It: A Novel Way to Learn Proof Construction. In *SIGCSE*, 2018.

[C2] A. McGregor, and S. Vorotnikova. A Simple, Space-Efficient, Streaming Algorithm for Matchings in Low Arboricity Graphs. In *SOSA*, 2018.

[C3] A. Mazumdar, A. McGregor, and S. Vorotnikova. Storage Capacity as an Information-Theoretic Analogue of Vertex Cover. In *ISIT*, 2017.

[C4] A. McGregor and H. Vu. Better Approximation of The Streaming Maximum Coverage Problem. In *ICDT*, 2017. **Invited to Special Issue of Theory of Computing Systems**

[C5] A. McGregor and S. Vorotnikova. Planar Matchings in Streams Revisited. In *APPROX-RANDOM*, 2016.

[C6] M. Crouch, A. McGregor, Gregory Valiant, and D. Woodruff. Stochastic Streams: Sample Complexity versus Space Complexity. In *European Symposium on Algorithms*, 2016.

[C7] A. McGregor, S. Vorotnikova, and H. T. Vu. Better Algorithms for Counting Triangles in Graph Streams. In *ACM Symposium on Principles of Database Systems*, 2016.

[C8] A. Abdullah, R. Kumar, A. McGregor, S. Vassilvitskii, and S. Venkatasubramanian. Sketching and Dimensionality Reduction for Information Spaces. *AISTATS*, 2016.

[C9] R. Chitnis, G. Cormode, H. Esfandiari, M. Hajiaghayi, A. McGregor, M. Monemizadeh, and S. Vorotnikova. Kernelization via Sampling with Applications to Dynamic Graph Streams. *ACM-SIAM Symposium on Discrete Algorithms*, 2016.

[C10] O. Simpson, C. Seshadhri, and A. McGregor. Catching the head, tail, and everything in between: a streaming algorithm for the degree distribution. *IEEE International Conference on Data Mining*, 2015.

[C11] M. Bender, S. McCauley, A. McGregor, S. Singh, and H. T. Vu. Run Generation Revisited: What Goes Up May or May Not Come Down. *26th International Symposium on Algorithms and Computation*, 2015.

[C12] A. McGregor, D. Tench, S. Vorotnikova, and H. T. Vu. Densest Subgraph in Dynamic Graph Streams. *Mathematical Foundations of Computer Science*, 2015.

[C13] A. McGregor and H. T. Vu. Evaluating Bayesian Networks via Data Streams. *21st International Computing and Combinatorics Conference*, 2015.

[C14] K. Ahn, G. Cormode, S. Guha, A. McGregor, and A. Wirth. Correlation Clustering in Data Streams. *International Conference on Machine Learning*, 2015.

[C15] S. Guha, A. McGregor, and D. Tench. Vertex and Hyperedge Connectivity in Dynamic Graph Streams. In *ACM Symposium on Principles of Database Systems*, 2015.

[C16] A. Chakrabarti, G. Cormode, A. McGregor, J. Thaler, and S. Venkatasubramanian. On Interactivity in Arthur-Merlin Communication and Stream Computation. *Computational Complexity Conference*, 2015.

[C17] A. McGregor, E. Price, and S. Vorotnikova. Trace Reconstruction Revisited. In *European Symposium on Algorithms*, pages 689–700, 2014.

[C18] M. Crouch, A. McGregor, and D. Stubbs. Dynamic Graphs in the Sliding-Window Model. In *European Symposium on Algorithms*, pages 337–348, 2013.

[C19]   A. McGregor and D. Stubbs. Sketching Earth-Mover Distance on Graph Metrics. In *APPROX-RANDOM*, pages 274–286, 2013.

[C20]   K. Ahn, S. Guha, and A. McGregor. Spectral Sparsification of Dynamic Graph Streams. In *APPROX-RANDOM*, pages 1–10, 2013.

[C21]   K. Krstovski, D. Smith, H. Wallach, and A. McGregor. Efficient Nearest-Neighbor Search in the Probability Simplex. In *International Conference on the Theory of Information Retrieval*, 2013.

[C22]   A. Andoni, A. Goldberger, A. McGregor, and E. Porat. Homomorphic Fingerprints under Misalignments. In *ACM Symposium on Theory of Computing*, pages 931–940, 2013.

[C23]   D. Barowy, C. Curtsinger, E. Berger, and A. McGregor. AUTOMAN: A Platform for Integrating Human-Based and Digital Computation. In *Conference on Object-Oriented Programming Systems, Languages, and Applications*, pages 639–654, 2012. **Invited to CACM Research Highlights. Media coverage including New Scientist and Times of India.**

[C24]   M. McCartin-Lim, A. McGregor, and R. Wang. Approximate Principle Direction Trees. In *International Conference on Machine Learning*, 2012.

[C25]   A. McGregor, A. Pavan, S. Tirthapura, and D. Woodruff. Space-Efficient Estimation of Statistics over Sub-Sampled Streams. In *ACM Symposium on Principles of Database Systems*, pages 273–282, 2012.

[C26]   K. Ahn, S. Guha, and A. McGregor. Graph Sketching: Sparsification, Spanners, and Subgraphs. In *ACM Symposium on Principles of Database Systems*, pages 5–14, 2012.

[C27]   K. Ahn, S. Guha, and A. McGregor. Analyzing Graph Structure via Linear Measurements. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 459–467, 2012.

[C28]   A. McGregor and P. Valiant. The Shifting Sands Algorithm. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 453–458, 2012.

[C29]   M. Crouch and A. McGregor. Periodicity and Cyclic Shifts via Linear Sketches. In *APPROX-RANDOM*, pages 158–170, 2011.

[C30]   B. Li, E. Mazur, Y. Diao, A. McGregor, and P. Shenoy. A platform for scalable one-pass analytics using MapReduce. In *ACM SIGMOD International Conference on Management of Data*, pages 985–996, 2011. **Invited to Special Issue of ACM Trans. Database Syst.**

[C31]   A. McGregor, A. Rudra, and S. Uurtamo. Polynomial Fitting of Data Streams with Applications to Codeword Testing. In *Symposium on Theoretical Aspects of Computer Science*, pages 428–439, 2011.

[C32]   M. Cartright, J. Allan, V. Lavrenko, and A. McGregor. Fast Query Expansion Using Approximations of Relevance Models. In *ACM Conference on Information and Knowledge Management*, pages 1573–1576, 2010.

[C33]   A. Chakrabarti, G Cormode, R. Kondapally, and A. McGregor. Information Cost Tradeoffs for Augmented Index and Streaming Language Recognition. In *IEEE Symposium on Foundations of Computer Science*, pages 387–396, 2010.

[C34]   A. McGregor, I. Mironov, T. Pitassi, O. Reingold, K. Talwar, and S. Vadhan. The Limits of Two-Party Differential Privacy. In *IEEE Symposium on Foundations of Computer Science*, pages 81–90, 2010.

[C35]   T. Tran, A. McGregor, Y. Diao, L. Peng, and A. Liu. Conditioning and Aggregating Uncertain Data Streams: Going Beyond Expectations. In *Proceedings of the VLDB Endowment*, pages 1302-1313, 2010.

[C36]   C. Li, M. Hay, V. Rastogi, G. Miklau, and A. McGregor. Optimizing Linear Counting Queries Under Differential Privacy. In *ACM Symposium on Principles of Database Systems*, pages 123–134, 2010.

[C37]   S. Chien, K. Ligett, and A. McGregor. Space-Efficient Estimation of Robust Statistics and Distribution Testing. In *Innovations in Computer Science*, pages 251–265, 2010.

[C38]   A. Chakrabarti, G. Cormode, and A. McGregor. Annotations in Data Streams. In *International Colloquium on Automata, Languages and Programming*, pages 222–234, 2009.

[C39]   A. McGregor, K. Onak, and R. Panigrahy. The Oil Searching Problem. In *European Symposium on Algorithms*, pages 504–515, 2009.

[C40]   G. Cormode, A. Deligiannakis, M. Garofalakis, and A. McGregor. Probabilistic Histograms for Probabilistic Data. In *Proceedings of the VLDB Endowment*, pages 526–537, 2009.

[C41]   G. Cormode, L. Golab, F. Korn, A. McGregor, D. Srivastava, and X. Zhang. Sampling to Estimate Conditional Functional Dependencies. In *ACM SIGMOD International Conference on Management of Data*, pages 469–482, 2009.

[C42]   K. Chaudhuri and A. McGregor. Finding metric structure in information theoretic clustering. In *Conference on Learning Theory*, pages 391–402, 2008.

[C43]   S. Guha and A. McGregor. Tight multi-pass stream lower bounds via pass elimination. In *International Colloquium on Automata, Languages and Programming*, pages 760–772, 2008.

[C44]   G. Cormode and A. McGregor. Approximation algorithms for clustering uncertain data. In *ACM Symposium on Principles of Database Systems*, pages 191–200, 2008.

[C45]   A. Chakrabarti, G. Cormode, and A. McGregor. Robust lower bounds for communication and stream computation. In *ACM Symposium on Theory of Computing*, pages 641–650, 2008.

[C46]   S. Angelov, K. Kunal, and A. McGregor. Sorting and selection with random costs. In *Latin American Theoretical Informatics Symposium*, pages 48–59, 2008.

[C47]   P. Indyk and A. McGregor. Declaring independence via the sketching of sketches. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 737–745, 2008.

[C48]   A. McGregor and O. Milenkovic. On the hardness of approximating stopping and trapping sets in LDPC codes. In *IEEE Information Theory Workshop*, pages 248–253, 2007.

[C49]   S. Guha and A. McGregor. Lower bounds for quantile estimation in random-order and multi-pass streaming. In *International Colloquium on Automata, Languages and Programming*, pages 704–715, 2007.

[C50]   M. Chu, S. Kannan, and A. McGregor. Checking and spot-checking of heaps. In *International Colloquium on Automata, Languages and Programming*, pages 728–739, 2007.

[C51] S. Guha, P. Indyk, and A. McGregor. Sketching information divergences. In *Conference on Learning Theory*, pages 424–438, 2007. **Invited to Special Issue of the Journal of Machine Learning.**

[C52] T. S. Jayram, A. McGregor, S. Muthukrishnan, and E. Vee. Estimating statistical aggregates on probabilistic data streams. In *ACM Symposium on Principles of Database Systems*, pages 243–252, 2007. **Invited to Special Issue of ACM Trans. Database Syst.**

[C53] S. Guha and A. McGregor. Space-efficient sampling. In *AISTATS*, pages 169–176, 2007.

[C54] A. Chakrabarti, G. Cormode, and A. McGregor. A near-optimal algorithm for computing the entropy of a stream. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 328–335, 2007.

[C55] A. McGregor and B. Shepherd. Island hopping and path coloring with applications to WDM network design. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 864–873, 2007. **SIGACT "Significant papers on new areas published in proceedings 2007" (1 of 13).**

[C56] S. Guha, A. McGregor, and S. Venkatasubramanian. Streaming and sublinear approximation of entropy and information distances. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 733–742, 2006.

[C57] D. Agarwal, A. McGregor, J. M. Phillips, S. Venkatasubramanian, and Z. Zhu. Spatial scan statistics: approximations and performance study. In *ACM International Conference on Knowledge Discovery and Data Mining*, pages 24–33, 2006.

[C58] S. Guha and A. McGregor. Approximate quantiles and the order of the stream. In *ACM Symposium on Principles of Database Systems*, pages 273–279, 2006.

[C59] J. Feigenbaum, S. Kannan, A. McGregor, S. Suri, and J. Zhang. Graph distances in the streaming model: the value of space. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 745–754, 2005.

[C60] S. Kannan and A. McGregor. More on reconstructing strings from random traces: Insertions and deletions. In *IEEE International Symposium on Information Theory*, pages 297–301, 2005.

[C61] B. Harb, S. Kannan, and A. McGregor. Approximating the best-fit tree under l_p norms. In *APPROX-RANDOM*, pages 123–133, 2005.

[C62] A. McGregor. Finding graph matchings in data streams. In *APPROX-RANDOM*, pages 170–181, 2005.

[C63] J. Feigenbaum, S. Kannan, A. McGregor, S. Suri, and J. Zhang. On graph problems in a semi-streaming model. In *International Colloquium on Automata, Languages and Programming*, pages 531–543, 2004. **Invited to Special Issue of Theoretical Computer Science.**

[C64] A. McGregor. A problem in scheduling: Your time starts now.... In *FUN with Algorithms*, pages 34–40, 2004.

[C65] A. Barg and A. McGregor. List decoding of concatenated codes: improved performance estimates. In *IEEE International Symposium on Information Theory*, page 419, 2004.

[C66] T. Batu, S. Kannan, S. Khanna, and A. McGregor. Reconstructing strings from random traces. In *ACM-SIAM Symposium on Discrete Algorithms*, pages 910–918, 2004.

[C67] A. Barg and A. McGregor.   More on the reliability function of the binary symmetric channel. In *IEEE International Symposium on Information Theory*, page 115, 2003.

[C68] A. Barg and A. McGregor. Distance distribution of binary codes and the error probability of decoding. In *Workshop in Coding and Cryptography*, pages 51–61, 2003.

[C69] A. McGregor, E. Miranda, and P. Gawthrop. Physical modeling of musical instruments using bond graphs. In *Brazilian Symposium on Computer Music*, 1999.

TEACHING EXPERIENCE

University of Massachusetts, Amherst, USA   **2009 – Present**

CMPSCI 240 Reasoning About Uncertainty (S10, F11, S13, S14, S15, S16, S17)
CMPSCI 311 Advanced Algorithms (F16)
CMPSCI 596 Independent Study (F12)
CMPSCI 611 Advanced Algorithms (F09, F10, F12, F15, F17)
CMPSCI 696 Independent Study (S10, S11, S14)
CMPSCI 711 Randomized Algorithms (S09)
CMPSCI 711 Data Stream Algorithms (S12)
CMPSCI 891M Theory Seminar (S10, S11, S13)

DEPARTMENT SERVICE

University of Massachusetts, Amherst, USA   **2009 – Present**

Curriculum Committee (2015-17 Chair, 2017-18 Co-chair)
Center for Data Science Executive Committee (2017-18)
Executive Committee (2012-13, 2014-15, 2016-17)
Graduate Admissions (2009-10, 2013-14)
Annual Faculty Review committee (2009-10, 2011-12, 2015-18)
Hiring Committee (2010-11, 2012-13, 2013-14, 2014-15, 2015-16, 2017-18 Chair)
Graduate Committee (2010-11, 2014-15)
Professionalism Seminar (Job Hunting) Panelist (2009, 2010)
Organizer of the Distinguished Lecturer Series (2011-13)

Previous Departments   **2000 – 2008**

Organizer and Co-Creator of the UCSD ITA Seminar (2007-08)
Organizer of the Penn Algorithms and Complexity Seminar (2003-06)
Graduate Student Representative at Penn C.I.S. Faculty Meetings (2003-04)
Organizer and Co-Creator of the Bi-weekly Penn C.I.S. Departmental Social (2002-04)
Organizer of the Penn Theory Lunch (2001-02)
President of the Archimedeans, University of Cambridge Maths Society (2000-01)

INVITED TALKS

Keynote Conference Talks:

12th International Computer Science Symposium in Russia (2017)

Invited Lecture Courses and Tutorials:

KDD Tutorial "Graph Sketching, Streaming, and Space-Efficient Optimization" (2018)
ICML Tutorial "Graph Sketching, Streaming, and Space-Efficient Optimization" (2016)
VLDB Tutorial "Graph Synopses, Sketches and Streams: A Survey" (2012)
L'Ecole de Printemps d'Informatique Théorique (2012)
Stream Processing Workshop, Johns Hopkins Applied Physics Laboratory (2011)
Coding, Complexity, and Sparsity Workshop, University of Michigan (2011)

Stream Processing Workshop, Johns Hopkins Applied Physics Laboratory (2010)
Barbados Workshop on Computational Complexity (2009)

Invited Sessions and Workshops:

Dagstuhl Seminar 17181 on Theory and Applications of Hashing (2017)
ICERM Sublinear Algorithms and Big Data Day, Brown University (2014)
Workshop on Theoretical Aspects of Big Data, Hong Kong (2013)
Dortmund Workshop on Algorithms for Data Streams (2012)
Coding, Complexity, and Sparsity Workshop, University of Michigan (2012)
Third Google Market Algorithms and Optimization Workshop (2012)
NII Shonan Meeting "Large-scale Distributed Computation" (2012)
Bertinoro Workshop on Sublinear Algorithms (2011)
5th Bertinoro Workshop on Randomized Algorithms and Graphs (2010)
Barbados Workshop on Computational Complexity (2009)
Dagstuhl Seminar 08341 on Sublinear Algorithms (2008)
NIPS Workshop "Representations & Inference on Distributions" (2007)
Allerton Conference on Communication, Control, and Computing (2007)
UCSD Information Theory and Applications Workshop (2007)
AMS & MAA Joint Mathematics Meeting (2007)
IIT Kanpur Workshop on Algorithms for Data Streams (2006)
Bertinoro Workshop on Space-Conscious Algorithms (2006)
Dagstuhl Seminar 05291 on Sublinear Algorithms (2005)
DIMACS Workshop on Algebraic Coding Theory and Information Theory (2003)

Seminars (Universities):

Aarhus University, Chinese University of Hong Kong, Dartmouth College, Dortmund University of Technology, Hong Kong University of Science and Technology, Indiana University Bloomington, King's College London, Massachusetts Institute of Technology, McGill University, Penn State University, Purdue University, Tsinghua University, University of Athens, University of California (Merced & San Diego), University of Cambridge, University of Edinburgh, University of Illinois (Urbana Champaign), University of Maryland, University of Massachusetts (Amherst), University of Michigan, Université Paris Diderot, Université Paris Sud, University of Pennsylvania, University of Warwick, University of Utah, Washington University of St. Louis, and Williams College.

Seminars (Industry):

Alcatel-Lucent Bell Labs, AT&T Labs, Google NYC, IBM Almaden, IBM TJ Watson, Microsoft SVC, and Yahoo! Research.

STUDENT SUPERVISION

Ph.D. Thesis Committees (University of Massachusetts):

Cibele Freire, 2015 (Computer Science, Advisor: Neil Immerman)
Dan Barowy, 2015 (Computer Science, Advisor: Emery Berger)
Tom Shelly, 2016 (Mathematics, Advisor: Alexei Oblomkov)
Stephen Oloo, 2014 (Mathematics, Advisor: Tom Braden)
Tobias Wilson, 2014 (Mathematics, Advisor: Alexei Oblomkov)
Md. Ashraful Alam, 2014 (Computer Science, Advisor: Ileana Streinu)
John Bowers, 2014 (Computer Science, Advisor: Ileana Streinu)
Bodou Li, 2014 (Computer Science, Advisor: Yanlei Diao)
Yoonheui Kim, 2012 (Computer Science, Advisor: Victor Lesser)
Jennifer Koonz, 2012 (Mathematics, Advisor: Eric Sommers)
Sam Huston, 2012 (Computer Science, Advisor: Bruce Croft)
Guven Ince, 2012 (Management, Advisor: Iqbal Agha)
Milad Ebtehaj, 2011 (Management, Advisor: Iqbal Agha)

Chao Li, 2011 (Computer Science, Advisor: Gerome Miklau, **2014 Department Outstanding Dissertation Award**)
Thanh Tran, 2010 (Computer Science, Advisor: Yanlei Diao)

Ph.D. Thesis Committees or External Examiner (External):

Sagar Kale, 2017 (Dartmouth College, Advisor Amit Chakrabarti)
Tirodkar Sumedh Vinod, 2016 (IIT Bombay, Advisor: Vinodkumar Prasad)
Christian Konrad, 2013 (Université Paris Diderot, Advisor: Frédéric Magniez)
Valerio Grossi, 2009 (University of Pisa, Advisor: Franco Turini)

Ph.D. Student Supervision:

Michael Crouch (first employment: Bell Labs, Ireland)
Mark McCartin-Lim (co-advised with Beverly Woolf)
David Tench
Sofya Vorotnikova
Hoa Vu (first employment: Boston College)

Undergraduate Supervision (REU and Honors Thesis):

John Brattin
Marco Leandro Carmosino (Ph.D. Student at UC San Diego)
Nicolas Scarrci (**Department Outstanding Achievement in Theory**)
Vinay Shah (**CRA Outstanding Undergraduate Research**, Honorable Mention)
Daniel Stubbs (**CRA Outstanding Undergraduate Research**, Honorable Mention (2013) and Winner (2014) and University of Massachusetts "**Rising Researcher Award**")
Sofya Vorotnikova (2014 **Department Outstanding Achievement in Theory**)

Synthesis Project Supervision:

Md. Ashraful Alam, Daniel Barowy, Marc Cartright, Van Dang, Niall Emmart (2014 Outstanding Synthesis Award), Sam Huston, Bodou Li, and Kriste Krstovski.

COMMUNITY INVOLVEMENT

Editorial:

SICOMP Special Issue for Best Papers from STOC 2012
Springer Encyclopedia of Algorithms Area Editor (2013 – Present)

Organizer:

FOCS Workshop on Linear Sketching as a Tool for Everything (2017)
Communication Complexity & Applications II, Banff Int. Research Station (2017)
Inference Problems Program, Nexus of Information and Computation Theories, Henri Poincare Institute (2016)
Towards a Unified Treatment of Dynamic Graphs, Banff Int. Research Station (2015)
Large-Scale Graph Algorithms, NII Shonan (2014)
Communication Complexity & Applications, Banff Int. Research Station (2014)
STOC Workshop on Distributed and Streaming Computation (2012)
IIT Kanpur Workshop on Algorithms for Data Stream (2009)
DIMACS/DyDAn Workshop on Streaming, Coding, and Compressive Sensing: Unifying Theory and Common Applications to Sparse Signal/Data Analysis and Processing (2009)

Grant Panels:

NSF (2010, 2011, 2012, 2014, 2015, 2016 twice, 2017)

Program Committees:

2nd Symposium on Simplicity in Algorithms (2019)
29th ACM-SIAM Symposium on Discrete Algorithms (2018)

35th Symposium on Theoretical Aspects of Computer Science (2018)
36th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (2017)
23rd Annual European Symposium on Algorithms (2015)
21st Symposium on String Processing and Information Retrieval (2014)
25th ACM-SIAM Symposium on Discrete Algorithms (2014)
20th Symposium on String Processing and Information Retrieval (2013)
32th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (2013)
44th ACM Symposium on Theory of Computing: Judging Panel for Best Student Presentation (2012)
44th ACM Symposium on Theory of Computing (2012)
5th International Frontiers of Algorithmics Workshop (2011)
7th IEEE/ACM International Conference on Distributed Computing in Sensor Systems (2011)
19th ACM Conference on Information and Knowledge Management (2010)
6th IEEE/ACM International Conference on Distributed Computing in Sensor Systems (2010)
21st ACM-SIAM Symposium on Discrete Algorithms (2010)
17th ACM Conference on Information and Knowledge Management (2008)
4th IEEE/ACM International Conference on Distributed Computing in Sensor Systems (2008)

Conference Reviewing:

AAIM, APPROX, COLT, ESA, FOCS, FSTTCS, ICALP, ISAAC, ISIT, ITCS, ITW, MFCS, PODC, PODS, RANDOM, SODA, STACS, STOC, and SWAT.

Journal Reviewing:

SIAM Journal on Computing, Journal of Algorithms, IEEE Transactions on Information Theory, IEEE Transactions on Knowledge and Data Engineering, IEEE Transactions on Communications, Theory of Computing, Theory of Computing Systems, Discrete Applied Mathematics, VLDB Journal, and ACM Transactions on Algorithms.

Collation and Editing of "Open Problems in Data Streams and Related Topics":

www.cse.iitk.ac.in/users/sganguly/data-stream-probs.pdf
See also: sublinear.info

Collation and Editing of "Open Problems In Data Streams, Property Testing, and Related Topics " with P. Indyk, I. Newman, and K. Onak:

www.cs.umass.edu/~mcgregor/papers/11-openproblems.pdf
See also: sublinear.info

Research Blog: polylogblog.wordpress.com