

QIN ZHANG

Computer Science Department
Indiana University Bloomington
Luddy Hall, RM 3044,
700 North Woodlawn Avenue,
Bloomington, IN 47405, USA

Email: qzhangcs@indiana.edu
Work: +1 (812) 855-2567
Cell: +1 (408) 802-5610
homes.sice.indiana.edu/qzhangcs

RESEARCH INTERESTS ◊ *Algorithms for Big Data*: communication-efficient algorithms for distributed data; streaming and sketching algorithms; algorithms for massive parallel computation; lower bounds.
◊ *Theoretical Foundations for Machine Learning and AI*: collaborative learning; distributed learning; learning augmented algorithms.

EMPLOYMENT ◊ Indiana University, Bloomington, IN, USA 2019 – Present
Associate Professor (with tenure)
Core faculty in Data Science
◊ Indiana University, Bloomington, IN, USA 2013 – 2019
Assistant Professor
◊ Theory Group, IBM Research Almaden, San Jose, CA, USA 2012 – 2013
Post-doctoral Researcher
◊ Center for Massive Data Algorithmics, Aarhus University, Denmark 2010 – 2012
Post-doctoral Associate

EDUCATION ◊ Ph.D. in Computer Science and Engineering 2010
Hong Kong University of Science and Technology
Dissertation: *Communication-Efficient Algorithms for Tracking Distributed Data Streams*
◊ B.S. in Computer Science 2006
Fudan University, Shanghai, China

Note: All papers use alphabetic ordering of authors (equal contribution) except mentioned otherwise.
Ph.D. students I have worked with are underlined.

INVITED JOURNAL PAPERS ◊ Q. Zhang. Technical Perspective: Fair Near Neighbor Search via Sampling. **SIGMOD Record**, 50(1): 41, 2021.
◊ S. Guha, Y. Li, and Q. Zhang. Distributed Partial Clustering. *ACM Transactions on Parallel Computing (TOPC)*, 6(3): 11:1-11:20, 2019.
◊ D. P. Woodruff and Q. Zhang. When Distributed Computation is Communication Expensive. **Distributed Computing**, 30(5): 309-323, 2017.
◊ G. Cormode, S. Muthukrishnan, K. Yi, and Q. Zhang. Continuous Sampling from Distributed Streams. *Journal of the ACM (JACM)*, 59(2): 10, 2012.
JOURNAL PAPERS ◊ Z. Huang, B. Radunovic, M. Vojnovic, and Q. Zhang. Communication Complexity of Approximate Matching in the Message-Passing Model. **Distributed Computing**, 33(6): 515-531, 2020.

QIN ZHANG

- ◇ Y. Song, H. Tang, H. Zhang, and Q. Zhang. Overlap Detection on Long, Error-Prone Sequencing Reads via Smooth q-Gram. **Bioinformatics**, published on 20 April 2020.
<https://doi.org/10.1093/bioinformatics/btaa252>.
- ◇ Z. Huang, K. Yi, and Q. Zhang. Randomized Algorithms for Tracking Distributed Count, Frequencies, and Ranks. **Algorithmica**, 81(6): 2222-2243, 2019.
- ◇ Y. Yu, D. Belazzougui, C. Qian, and Q. Zhang (*ordered by contribution*). Memory-efficient and Ultra-fast Network Lookup and Forwarding using Othello Hashing. *ACM Transactions on Networking (TON)*, 26(3): 1151-1164, 2018.
- ◇ J. Chen and Q. Zhang. Improved Algorithms for Distributed Entropy Monitoring. **Algorithmica**, 78(3): 1041-1066, 2017.
- ◇ J. M. Phillips, E. Verbin, and Q. Zhang. Lower Bounds for Number-in-Hand Multiparty Communication Complexity, Made Easy. *SIAM Journal of Computing (SICOMP)*, 45(1): 174-196, 2016
- ◇ R. Pagh, Z. Wei, K. Yi, and Q. Zhang. Cache-Oblivious Hashing. **Algorithmica**, 69(4): 864-883, 2014.
- ◇ E. Verbin and Q. Zhang. The Limits of Buffering: A Tight Lower Bound for Dynamic Membership in the External Memory Model. *SIAM Journal of Computing (SICOMP)*, 42(1): 212-229, 2013.
- ◇ K. Yi and Q. Zhang. Optimal Tracking of Distributed Heavy Hitters and Quantiles. **Algorithmica**, 65(1): 206-223, 2013.
- ◇ K. Yi and Q. Zhang. Multi-Dimensional Online Tracking. *ACM Transactions on Algorithms (TALG)*, 8(2), Article 12, April 2012.

PEER-
REVIEWED
CONFERENCE
PAPERS

- ◇ N. Karpov and Q. Zhang. Batched Coarse Ranking in Multi-Armed Bandits. *In Proceedings of the 34th Annual Conference on Neural Information Processing Systems (NeurIPS '20)*. Virtual conference, December 2020.
- ◇ N. Karpov, Q. Zhang and Y. Zhou. Collaborative Top Distribution Identifications with Limited Interaction. *In Proceedings of the 61st IEEE Symposium on Foundations of Computer Science (FOCS '20)*. Virtual conference, November 2020.
- ◇ K. Dong, Y. Li, Q. Zhang and Y. Zhou. Multinomial Logit Bandit with Low Switching Cost. *In Proceedings of the 37th International Conference on Machine Learning (ICML '20)*. Virtual conference, July 2020.
- ◇ H. Zhang and Q. Zhang. MinSearch: An Efficient Algorithm for Similarity Search under Edit Distance. *In Proceedings of the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '20)*, **oral presentation**. Virtual conference, August 2020.
- ◇ M. Khan, G. Pandurangan, N. D. Pham, A. Vullikanti, and Q. Zhang. A Multi-criteria Approximation Algorithm for Influence Maximization with Probabilistic Guarantees. *In Proceedings of the SIAM Symposium on Algorithm Engineering and Experiments (ALENEX '20)*, pages 81-93. Salt Lake City, UT, U.S.A., January 2020.
- ◇ C. Tao, Q. Zhang and Y. Zhou. Collaborative Learning with Limited Interaction: Tight Bounds for Distributed Exploration in Multi-Armed Bandits. *In Proceedings of the 60th IEEE Symposium on Foundations of Computer Science (FOCS '19)*, pages 126-146. Baltimore, MD, U.S.A., November 2019.
- ◇ H. Zhang and Q. Zhang. MinJoin: Efficient Edit Similarity Joins via Local Hash Minima. *In Proceedings of the 25th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '19)*, **oral presentation**, pages 1093-1103. Anchorage, AK, U.S.A., August 2019.
- ◇ S. Assadi, N. Karpov, and Q. Zhang. Distributed and Streaming Linear Programming in Low Dimensions. *In Proceedings of the 38th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS '19)*, pages 236-253. Amsterdam, The Netherlands, June-July 2019.

Invited to Special Issue of ACM Transactions on Database Systems.

QIN ZHANG

- ◇ J. Chen, Q. Zhang, and Y. Zhou. Tight Bounds for Collaborative PAC Learning via Multiplicative Weights . In *Proceedings of the 32nd Annual Conference on Neural Information Processing Systems (NeurIPS '18)*, pages 3602-3611. Montreal, Canada, December 2018.
 - ◇ J. Chen, E. Sadeqi-Azer, and Q. Zhang. A Practical Algorithm for Distributed Clustering and Outlier Detection. In *Proceedings of the 32nd Annual Conference on Neural Information Processing Systems (NeurIPS '18)*, pages 2253-2262. Montreal, Canada, December 2018.
 - ◇ H. Zhang, Q. Zhang, and H. Tang (*ordered by contribution*). Smooth q -Gram, and Its Applications to Detection of Overlaps among Long, Error-Prone Sequencing Reads. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM '18)*, pages 267-276. Turin, Italy, October 2018.
 - ◇ J. Chen and Q. Zhang. Distinct Sampling on Streaming Data with Near-Duplicates. In *Proceedings of the 37th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS '18)*, pages 369-382. Houston, TX, U.S.A., June 2018.
 - ◇ D. P. Woodruff and Q. Zhang. Distributed Statistical Estimation of Matrix Products with Applications. In *Proceedings of the 37th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS '18)*, pages 383-394. Houston, TX, U.S.A., June 2018.
 - ◇ H. Zhang and Q. Zhang. Communication-Efficient Distributed Skyline Computation. In *Proceedings of the 26th ACM Conference on Information and Knowledge Management (CIKM '17)*, pages 437-446. Singapore, November, 2017.
 - ◇ Y. Yu, D. Belazzougui, C. Qian, and Q. Zhang (*ordered by contribution*). A Concise Forwarding Information Base for Scalable and Fast Name Switching. In *Proceedings of the 25th IEEE International Conference on Network Protocols (ICNP '17)*, pages 1-10. Toronto, Canada, October 2017.
 - ◇ H. Zhang and Q. Zhang. EmbedJoin: Efficient Edit Similarity Joins via Embeddings. In *Proceedings of the 23rd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '17)*, **oral presentation**, pages 585-594. Halifax, Nova Scotia, Canada, August 2017.
 - ◇ J. Chen, X. Chen, Q. Zhang, and Y. Zhou. Adaptive Multiple-Arm Identification. In *Proceedings of the 34th International Conference on Machine Learning (ICML '17)*, pages 722-730. Sydney, Australia, August 2017.
 - ◇ S. Guha, Y. Li, and Q. Zhang. Distributed Partial Clustering. In *Proceedings of the 29th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA '17)*, pages 143-152. Washington D.C., U.S.A., July 2017. **Best paper award**.
- Invited to Special Issue of ACM Transactions on Parallel Computing.**
- ◇ J. Chen and Q. Zhang. Bias-Aware Sketches. In *Proceedings of the 43rd International Conference on Very Large Data Bases (VLDB '17)*, pages 961-972. Munich, Germany, August-September 2017.
 - ◇ J. Chen, H. Sun, D. Woodruff, and Q. Zhang. Communication-Optimal Distributed Clustering. In *Proceedings of the 30th Annual Conference on Neural Information Processing Systems (NIPS '16)*, pages 3720-3728. Barcelona, Spain, December 2016.
 - ◇ D. Belazzougui and Q. Zhang. Edit Distance: Sketching, Streaming and Document Exchange. In *Proceedings of the 57th IEEE Symposium on Foundations of Computer Science (FOCS '16)*, pages 51-60. New Brunswick, NJ, U.S.A., October, 2016.
 - ◇ D. Chen and Q. Zhang. Streaming Algorithms for Robust Distinct Elements. In *Proceedings of the 35th ACM SIGMOD International Conference on Management of Data (SIGMOD '16)*, pages 1433-1447. San Francisco, CA, USA, June 2016.
 - ◇ A. Andoni, J. Chen, R. Krauthgamer, B. Qin, D. P. Woodruff, and Q. Zhang. On Sketching Quadratic Forms. In *Proceedings of the 7th Innovations in Theoretical Computer Science (ITCS '16)*, pages 311-319. Cambridge, MA, USA, January 2016.

QIN ZHANG

- ◇ Q. Zhang. Communication-Efficient Computation on Distributed Noisy Datasets. *In Proceedings of the 27th ACM Symposium on Parallelism in Algorithms and Architectures (SPAA '15)*, pages 313-322. Portland, Oregon, USA, June 2015.
- ◇ D. Van Gucht, R. Williams, D. P. Woodruff, and Q. Zhang. The Communication Complexity of Distributed Set-Joins with Applications to Matrix Multiplication. *In Proceedings of the 34th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS '15)*, pages 199-212. Melbourne, VIC, Australia, May-June 2015.
- ◇ Z. Huang, B. Radunovic, M. Vojnovic, and Q. Zhang. Communication Complexity of Approximate Matching in Distributed Graphs. *In Proceedings of the 32nd Symposium on Theoretical Aspects of Computer Science (STACS '15)*, pages 460-473. Munich, Germany, March 2015.
- ◇ D. Chen, C. Konrad, K. Yi, W. Yu, and Q. Zhang. Robust Set Reconciliation. *In Proceedings of the 33rd ACM SIGMOD International Conference on Management of Data (SIGMOD '14)*, pages 135-146. Snowbird, UT, USA, June 2014.
- ◇ D. P. Woodruff and Q. Zhang. An Optimal Lower Bound for Distinct Elements in the Message Passing Model. *In Proceedings of the 25th ACM-SIAM Symposium on Discrete Algorithms (SODA '14)*, pages 718-733. Portland, OR, USA, January 2014.
- ◇ D. P. Woodruff and Q. Zhang. When Distributed Computation is Communication Expensive. *In Proceedings of the 27th International Symposium on Distributed Computing (DISC '13)*, pages 16-30. Jerusalem, Israel, October 2013.
- Invited to Distributed Computing.**
- ◇ D. P. Woodruff and Q. Zhang. Subspace Embeddings and Lp Regression Using Exponential Random Variables. *In Proceedings of the 26th Annual Conference on Learning Theory (COLT '13)*, pages 546-567. Princeton, NJ, USA, June 2013.
- ◇ L. K. Lee, M. Lewenstein, and Q. Zhang. Parikh Matching in the Streaming Model. *In Proceedings of the 19th International Symposium on String Processing and Information Retrieval (SPIRE '12)*, pages 336-341. Cartagena de Indias, Colombia, October 2012.
- ◇ E. Verbin and Q. Zhang. Rademacher-Sketch: A Dimensionality-Reducing Embedding for Sum-Product Norms, with an Application to Earth-Mover Distance. *In Proceedings of the 39th International Colloquium on Automata, Languages and Programming (ICALP '12)*, pages 834-845. Warwick, UK, July 2012.
- ◇ Z. Huang, K. Yi, and Q. Zhang. Randomized Algorithms for Tracking Distributed Count, Frequencies, and Ranks. *In Proceedings of the 31st ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS '12)*, pages 295-306. Scottsdale, Arizona, USA, May 2012.
- ◇ D. P. Woodruff and Q. Zhang. Tight Bounds for Distributed Functional Monitoring. *In Proceedings of the 44th ACM Symposium on Theory of Computing (STOC '12)*, pages 941-960. New York, NY, USA, May 2012.
- ◇ J. M. Phillips, E. Verbin, and Q. Zhang. Lower Bounds for Number-in-Hand Multiparty Communication Complexity, Made Easy. *In Proceedings of the 23rd ACM-SIAM Symposium on Discrete Algorithms (SODA '12)*, pages 486-501. Kyoto, Japan. January 2012.
- ◇ M. T. Goodrich, N. Sitchinava, and Q. Zhang. Sorting, Searching and Simulation in the MapReduce Framework. *In Proceedings of the 22nd International Symposium on Algorithms and Computation (ISAAC '11)*, pages 374-383. Yokohama, Japan. December 2011.
- ◇ H. L. Chan, T. W. Lam, L. K. Lee, J. Pan, H. F. Ting, and Q. Zhang. Edit Distance to Monotonicity in Sliding Windows. *In Proceedings of the 22nd International Symposium on Algorithms and Computation (ISAAC '11)*, pages 564-573. Yokohama, Japan. December 2011.

QIN ZHANG

- ◇ J. Li, K. Yi, and Q. Zhang. Clustering with Diversity. *In Proceedings of the 37th International Colloquium on Automata, Languages and Programming (ICALP '10)*, pages 188-200. Bordeaux, France, July 2010.
- ◇ R. Pagh, Z. Wei, K. Yi, and Q. Zhang. Cache-Oblivious Hashing. *In Proceedings of the 29th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS '10)*, pages 297-304. Indianapolis, IN, USA, June 2010.
- ◇ G. Cormode, S. Muthukrishnan, K. Yi, and Q. Zhang. Optimal Sampling From Distributed Streams. *In Proceedings of the 29th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS '10)*, pages 77-86. Indianapolis, IN, USA, June 2010.

Invited to Journal of the ACM.

- ◇ E. Verbin and Q. Zhang. The Limits of Buffering: A Tight Lower Bound for Dynamic Membership in the External Memory Model. *In Proceedings of the 42nd ACM Symposium on Theory of Computing (STOC '10)*, pages 447-456. Cambridge, MA, USA, June 2010.
- ◇ K. Yi and Q. Zhang. On the Cell Probe Complexity of Dynamic Membership. *In Proceedings of the 21st ACM-SIAM Symposium on Discrete Algorithms (SODA '10)*, pages 123-133. Austin, TX, USA, January 2010.
- ◇ Z. Wei, K. Yi, and Q. Zhang. Dynamic External Hashing: The Limit of Buffering. *In Proceedings of the 21st ACM Symposium on Parallelism in Algorithms and Architectures (SPAA '09)*, pages 253-259. Calgary, Canada, August 2009.
- ◇ K. Yi and Q. Zhang. Optimal Tracking of Distributed Heavy Hitters and Quantiles. *In Proceedings of the 28th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (PODS '09)*, pages 167-174. Providence, RI, USA, June - July 2009.
- ◇ K. Yi and Q. Zhang. Multi-Dimensional Online Tracking. *In Proceedings of the 20th ACM-SIAM Symposium on Discrete Algorithms (SODA '09)*, pages 1098-1107. New York, NY, USA, January 2009.
- ◇ J. Jia, Q. Zhang, Q. Zhang, and M. Liu (*ordered by contribution*). Revenue Generation for Truthful Spectrum Auction in Dynamic Spectrum Access. *In Proceedings of the 10th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '09)*, pages 3-12. New Orleans, LA, USA, May 2009.
- ◇ Q. Zhang, F. Li, and K. Yi (*ordered by contribution*). Finding Frequent Items in Probabilistic Data. *In Proceedings of the 27th ACM SIGMOD International Conference on Management of Data (SIGMOD '08)*, pages 819-832. Vancouver, Canada, June 2008.

PEER-
REVIEWED
SHORT
PAPERS

- ◇ Y. Yu, D. Belazzougui, C. Qian, and Q. Zhang (*ordered by contribution*). A Fast, Small, and Dynamic Forwarding Information Base. *In Proceedings of the 2017 International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS '17)*, pages 41-42. Urbana-Champaign, IL, USA, June 2017.

PAPERS
UNDER
SUBMISSION /
MANUSCRIPTS

- ◇ N. Karpov and Q. Zhang. Instance-Sensitive Algorithms for Pure Exploration in Multinomial Logit Bandit. *Submitted*.
- ◇ Y. Li, Y. Song, and Q. Zhang. Learning to Cluster via Same-Cluster Queries. *Submitted*.
- ◇ A. Andoni, J. Chen, R. Krauthgamer, B. Qin, D. P. Woodruff, and Q. Zhang. On Sketching Quadratic Forms. *Manuscript; journal version of the ITCS'16 paper; submitted to JACM, under revision; arXiv: <http://arxiv.org/abs/1511.06099>.*
- ◇ J. Chen, H. Nguyen, and Q. Zhang. Submodular Maximization over Sliding Windows. *Manuscript. arXiv: <https://arxiv.org/abs/1611.00129>*

QIN ZHANG

- THESIS ◇ Qin Zhang. Communication-Efficient Algorithms for Tracking Distributed Data Streams. Ph.D. Thesis. Hong Kong University of Science and Technology, 2010.
- PATENTS ◇ D. Woodruff and Q. Zhang. “Robust regression analysis techniques using exponential random variables”. US 10007643 B2 (granted)
- ◇ G. Cormode, K. Yi, and Qin Zhang. “Sampling from distributed streams of data”. US 9130827 B2 (granted)
- FUNDING ◇ **PI for all grants. Total personal share USD 2.12M.**
- ◇ **(PI)** *Collaborative Research: AF: Small: Parallel Reinforcement Learning with Communication and Adaptivity Constraints.*
NSF CCF-2006591, 10/2020-09/2023, **\$242,214.**
Total award \$499,959, with Yuan Zhou (UIUC).
- ◇ **(Sole PI)** *CAREER: Foundation of Communication-Efficient Distributed Computation and Monitoring.*
NSF CCF-1844234, 06/2019-05/2024, **\$499,698.**
- ◇ **(PI)** *AF: Small: Redundancy Exploiting Algorithms for High Throughput Genomics.*
NSF CCF-1619081, 08/2016-07/2020, **\$400,000.**
Former PI: Cenk Sahinalp; former Co-PI: Funda Ergun.
- ◇ **(PI)** *BIGDATA: Collaborative Research: F: Efficient Distributed Computation of Large-Scale Graph Problems in Epidemiology and Contagion Dynamics.*
NSF IIS-1633215, 09/2016-08/2020, **\$530,114.**
Total award \$1.8M, with Gopal Pandurangan (Univ. of Houston) and Anil Vullikanti (Virginia Tech).
- ◇ **(Sole PI)** *AF: Small: Efficient Algorithms for Querying Noisy Distributed/Streaming Datasets.*
NSF CCF-1525024, 06/2015-05/2019, **\$444,320.**
- ◇ **(Sole PI)** *A Foundation of Communication-Efficient Online Distributed Data Processing.*
Indiana University Faculty Research Support Program (FRSP), 06/2015-05/2016, **\$38,000.**
- ◇ **(PI)** *Data Structure Complexity in Data Streams and MapReduce Models.*
KFKT2011A12, 01/2011-01/2013, **CNY 40,000.**
With Yitong Yin (Nanjing University, China).
- HONORS AND AWARDS ◇ NSF CAREER Award, 2019.
- ◇ Best Paper Award, *ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, 2017.
- ◇ Four conference papers (PODS 2020, SPAA 2017, DISC 2013, PODS 2010) are invited to *ACM Transactions on Database Systems*, *ACM Transactions on Parallel Computing*, *Distributed Computing*, and *Journal of the ACM* respectively.
- ◇ Top Reviewer for NeurIPS 2019 and ICML 2020.
- ◇ HKUST Ph.D. Scholarship, 2006 – 2010, Annual.
- ◇ Scholarship of Fudan University, 2002 – 2006, Annual.
- ◇ *Fudi* Scholarship for University Study, 2002 – 2006, Annual.
- ◇ First Prize in China National Olympiad in Mathematics, 2002.
- ◇ Second Prize in China National Olympiad in Mathematics, 2001.

QIN ZHANG

- TEACHING EXPERIENCE
- ◇ Indiana University, Bloomington, USA 2013 – Present
 - A. Course*
 - Spring 2020: Instructor, CSCI-B403 Introduction to Algorithm Design and Analysis (undergraduate)
 - Fall 2019: Instructor, CSCI-B503 Introduction to Algorithm Design and Analysis (graduate)
 - Fall 2018: Instructor, CSCI-B669 Sublinear Algorithms for Big Data (graduate)
 - Spring 2018: Instructor, CSCI-B403 Introduction to Algorithm Design and Analysis (undergraduate)
 - Fall 2017: Instructor, CSCI-B669 Sublinear Algorithms for Big Data (graduate)
 - Spring 2017: Instructor, CSCI-B403 Introduction to Algorithm Design and Analysis (undergraduate)
 - Fall 2015: Instructor, CSCI-B561 Advanced Database Concepts (2 sessions) (graduate)
 - Fall 2014: Instructor, CSCI-B561 Advanced Database Concepts (2 sessions) (graduate)
 - Spring 2014: Instructor, CSCI-B490 Mining the Big Data (undergraduate)
 - Fall 2013: Instructor, CSCI-B669 Sublinear Algorithms for Big Data (graduate)
 - B. Independent Study*
 - Fall 2019: Mentor, CSCI-Y790 Collaborative Selection under Limited Communication (graduate)
 - Fall 2019: Mentor, CSCI-Y790 String Similarity Joins and Search on Edit Distance (graduate)
 - Fall 2019: Mentor, CSCI-Y790 Clustering with Same-cluster Query Oracle (graduate)
 - Fall 2018: Mentor, CSCI-Y790, Overlap detection on DNA reads via Smooth q -gram (graduate)
 - Spring 2018: Mentor, CSCI-Y790 Sublinear Algorithms for Big Data (graduate)
 - Fall 2017: Mentor, CSCI-Y790 Practical Distributed Clustering (graduate)
 - Fall 2017: Mentor, CSCI-Y790 Matching q -Grams with Edits (graduate)
 - Fall 2016: Mentor, CSCI-Y790 Scalable Similarity Search on Sparse Vectors (graduate)
 - Fall 2016: Mentor, CSCI-Y790 Efficient Edit Similarity Joins (graduate)
 - Fall 2016: Mentor, CSCI-Y790 Bias-Aware Sketches (graduate)
 - Spring 2016: Mentor, CSCI-Y790 Communication-Optimal Distributed Clustering (graduate)
 - Spring 2016: Mentor, CSCI-Y790 Distributed Learning (graduate)
 - Spring 2015: Mentor, CSCI-Y790: Algorithms for Numerical Linear Algebra (graduate)
 - Spring 2015: Mentor, CSCI-Y390 Advanced Data Mining (undergraduate)
 - Fall 2014: Mentor, CSCI-Y790 Advanced Data Mining (graduate)
 - Spring 2014: Mentor, CSCI-Y790 Advanced Randomized Algorithms (graduate)
 - ◇ Aarhus University, Denmark 2011
 - Spring 2011: Instructor, Streaming Algorithms
 - ◇ Hong Kong University of Science and Technology 2006 – 2010, 2012
 - Summer 2012: Instructor, Linear Sketch and Its Applications in Data Streams and Compressive Sensing
 - 2006, 2007, 2008, 2009, 2010: Teaching Assistant, for courses Discrete Mathematics, Combinatorial Optimization, Design and Analysis of Algorithms, Computational Geometry
- INVITED WORKSHOP TALKS
- ◇ Collaborative Learning with Communication Constraints. *Workshop on Algorithms for Large Data*, Online. August 2021.
 - ◇ Parallel Learning with Limited Communication. *Second Hawaii Workshop on Parallel Algorithms and Data Structures*, Honolulu, HI, USA. December 2019.
 - ◇ Distributed Statistical Estimation of Matrix Products with Applications. *Workshop on Interactive Complexity*, Simons Institute, Berkeley CA, October 2018.
 - ◇ Edit Distance: Sketching, Streaming and Document Exchange. *NII Shonan Meeting on Processing Big Data Streams*, Shonan Village, Japan. June 2017.

QIN ZHANG

- ◇ Efficient Algorithms for Streaming Datasets with Near-Duplicates. *Theory and Applications of Hashing*, Dagstuhl Seminar, Germany. May 2017.
- ◇ Some New Questions in Communication Complexity. *Communication Complexity and Applications, II*, Banff, Alberta, Canada. March 2017.
- ◇ Communication Complexity for Distributed Graphs. *ADGA: Workshop on Advances in Distributed Graph Algorithms*, Paris, France. September 2016.
- ◇ Lower Bound Techniques for Multiparty Communication Complexity. *Nexus of Information and Computation Theories*, Henri Poincare Institute, Paris, France. February 2016.
- ◇ Streaming Algorithms for Robust Distinct Elements. *Workshop on Multi-dimensional Proximity Problems*, University of Maryland, College Park, MD, USA. January 2016.
- ◇ Algorithms for Querying Noisy Distributed/Streaming Datasets. *Sublinear Algorithms Workshop*, Johns Hopkins University, Baltimore, MD, USA. January 2016.
- ◇ A Sketching Algorithm for Spectral Graph Sparsification. *NII Shonan Meeting on Algorithms for Large Scale Graphs*, Shonan Village, Japan. October 2014.
- ◇ New Directions in Distributed Monitoring. *Bertinoro Workshop on Sublinear Algorithms*, Bertinoro (Forlì-Cesena), Italy. May 2014.
- ◇ Multiparty Communication Complexity in the Message-Passing Model. *Workshop on Theoretical Aspects of Big Data*, Hong Kong. July 2013.
- ◇ Rademacher Embedding, with Application to Earth-Mover Distance. *Workshop on Algorithms for Data Streams*, Dortmund, Germany. July 2012.
- ◇ Tight Bounds for Distributed Functional Monitoring. *NII Shonan Meeting on Large-scale Distributed Computation*, Shonan Village, Japan. January 2012.
- ◇ Lower Bounds for Number-in-Hand Multiparty Communication Complexity. *Workshop: Synergies in Lower Bounds*, Aarhus, Denmark. June 2011.
- ◇ External Memory Data Structures with $o(1)$ -I/O Updates. *China Theory Week*, Beijing, China. September 2010.
- ◇ Other Invited Workshops (cannot make them due to visa, family, etc.)
 - *Workshop on Data Summarization*, University of Warwick, UK. March 2018.
 - *Hawaiian Workshop on Parallel Algorithms and Data Structures*, Honolulu, Hawaii, USA. December 2017.
 - *BIRS Workshop: Towards a Unified Treatment of Dynamic Graphs*, Banff, Alberta, Canada. March 2015.
 - *BIRS Workshop: Communication Complexity and Applications*, Banff, Alberta, Canada. August 2014.
 - *NII Shonan Meeting on Compact Data Structures for Big Data*, Shonan Village, Japan. September 2013.

SEMINARS (EXTERNAL UNIVERSI- TIES)

- ◇ Collaborative Learning with Limited Interaction: Tight Bounds for Distributed Exploration in Multi-Armed Bandits
 - *Shanghai University of Finance and Economics*, Shanghai, China. May 2019.
- ◇ Similarity Joins on Sequence Data
 - *Biostatistics Colloquium*, IUPUI, Indianapolis, IN, USA. April 2019.
- ◇ Communication-Efficient Computation on Distributed Data: Theory and Practice.
 - *Computer Science Colloquium*, *University of Utah*, Salt Lake City, UT, USA. January 2019.
 - *Computer Science Colloquium*, *Purdue University*, West Lafayette, IN, USA. February 2019.
 - *Computer Science&Engineering Colloquium*, *University of Minnesota*, Minneapolis, MN, USA. February 2019.
 - *Computer Science Colloquium*, *University of Illinois at Chicago*, Chicago, IL, USA. April 2019.

QIN ZHANG

- ◇ Distributed Partial Clustering.
 - *Hong Kong University of Science and Technology*, Hong Kong. June 2017.
 - *University of Illinois at Urbana-Champaign*, Champaign, IL, USA. October 2017.
- ◇ The Communication Complexity of Distributed Set-Joins.
 - Computer Science Colloquium, *University of Houston*, Houston, TX, USA. November 2016.
- ◇ Edit Distance: Sketching, Streaming and Document Exchange.
 - *Aarhus University*, Aarhus, Denmark. September 2016.
 - *Shanghai University of Finance and Economics*, Shanghai, China. June 2017.
- ◇ Multiparty Communication Complexity in the Message-Passing Model.
 - *Stanford University*, CA, USA. April 2013.
 - *Purdue University*, IN, USA. October 2013.
- ◇ Rademacher Embedding, with application to Earth-Mover Distance.
 - *Tsinghua University*, Beijing, China. January 2012.
- ◇ Tight Bounds for Distributed Functional Monitoring.
 - *Hong Kong University of Science and Technology*, Hong Kong. August 2012.
- ◇ Taming the Data Deluge.
 - *Nanyang Technological University*, Singapore, March 2012.
 - *Fudan University*, Shanghai, China, September 2012.
 - *Shanghai Jiao Tong University*, Shanghai, China. September 2012.
- ◇ Distributed Streaming.
 - *IT University of Copenhagen*, Copenhagen, Denmark. December 2011.
 - *Shanghai Jiao Tong University*, Shanghai, China. January 2012.
- ◇ Lower Bounds for Number-in-Hand Multiparty Communication Complexity.
 - *LIAFA*, Paris, France. October 2011.
- ◇ Clustering with Diversity.
 - *University of Hong Kong*, Hong Kong, China. June 2010.
 - *Shanghai Jiao Tong University*, Shanghai, China. July 2010.
- ◇ On the Cell Probe Complexity of Dynamic Membership.
 - *Fudan University*, Shanghai, China. September 2009.
- ◇ The Art of Metric Embeddings.
 - *University of Leicester*, UK. August 2007.

SEMINARS (INTERNAL)

- ◇ A Brief Introduction to Big Data Algorithms. *IU Math Club*, Bloomington, IN, USA. October 2021.
- ◇ Efficient Computation on Distributed Data: Theory and Practice. *Computer Science Colloquium*, Bloomington, IN, USA. September 2018.
- ◇ Edit Distance: Sketching, Streaming and Document Exchange. *Theory Lunch Seminar*, Bloomington, IN, USA. September 2016.
- ◇ Efficient Algorithms for Querying Noisy Disturbed/Streaming Datasets. *IU Research Horizon*, Bloomington, IN, USA. September 2016.
- ◇ The Communication Complexity of Distributed Set-Joins with Applications to Matrix Multiplication. *Theory Lunch Seminar*, Bloomington, IN, USA. September 2015
- ◇ Computational Models for Big Data. *Department of Statistics Colloquium*, Bloomington, IN, USA. April 2015
- ◇ Efficient Statistical Estimations for Distributed Inconsistent Data. *Theory Lunch Seminar*, Bloomington, IN, USA. September 2014

QIN ZHANG

- ◇ Lower Bound Techniques for Multiparty Communication. *Theory Lunch Seminar*, Bloomington, IN, USA. February 2014
- ◇ A Crash Introduction on Data Stream Algorithms. *Theory Lunch Seminar*, Bloomington, IN, USA. September 2013

SEMINARS (INDUSTRY)

- ◇ Multiparty Communication Complexity in the Message-Passing Model.
– *IBM Research Almaden*, San Jose, CA. USA. April 2012.
- ◇ Taming the Data Deluge.
– *IBM Research Almaden*, San Jose, CA. USA. November 2012.
– *Microsoft Research Asia*, Beijing, China. March 2012.
- ◇ Sorting, Searching and Simulation in the MapReduce Framework.
– *Microsoft Research Asia*, Beijing, China. January 2012.
- ◇ Optimal Sampling from Distributed Streams.
– *Microsoft Research Asia*, Beijing, China. September 2010.
- ◇ Dynamic Dictionary and Membership.
– *Microsoft Research Asia*, Beijing, China. July 2009.

MEDIA

- ◇ Interview “Lower Bounds for Distinct Elements in the Message Passing Model” by *Abstract Talk* (Collaborative Interviews on Science and Technology).

CONFERENCE TALKS

- ◇ Collaborative Learning with Limited Interaction: Tight Bounds for Distributed Exploration in Multi-Armed Bandits. *FOCS '19*, Baltimore, MD, USA, November 2019.
- ◇ Distributed Statistical Estimation of Matrix Products with Applications. *PODS '18*, Houston, TX, USA. June 2018.
- ◇ Distributed Patial Clustering. *SPAA '17*, Washington D.C., USA. July 2017.
- ◇ Edit Distance: Sketching, Streaming and Document Exchange. *FOCS '16*, New Brunswick, NJ, USA. October 2016.
- ◇ Communication-Efficient Computation on Distributed Noisy Datasets. *SPAA '15*, Portland, OR, USA. June 2015.
- ◇ An Optimal Lower Bound for Distinct Elements in the Message Passing Model. *SODA '14*, Portland, OR, USA. January 2014.
- ◇ Subspace Embeddings and Lp Regression Using Exponential Random Variables. *COLT '13*, Princeton, NJ, USA. June 2013.
- ◇ Tight Bounds for Distributed Functional Monitoring. *STOC '12*, New York, NY, USA. May, 2012.
- ◇ Lower Bounds for Number-in-Hand Multiparty Communication Complexity. *SODA '12*, Kyoto, Japan. January 2012.
- ◇ On the Cell Probe Complexity of Dynamic Membership. *SODA '10*, Austin, TX, USA. January 2010.
- ◇ Dynamic External Hashing: The Limit of Buffering. *SPAA '09*, Calgary, Canada. August 2009.
- ◇ Optimal Tracking of Distributed Heavy Hitters and Quantiles. *PODS '09*, Providence, RI, USA. June 2009.
- ◇ Multi-Dimensional Online Tracking. *SODA '09*, New York, USA. January 2009.
- ◇ Dynamic External Hashing: The Limit of Buffering. *AAAC '09*, Hang Zhou, China. April 2009.

QIN ZHANG

- ◇ Finding Frequent Items in Probabilistic Data. *SIGMOD '08*, Vancouver, Canada. June 2008.
- ◇ Shannon Coding for the Discrete Noiseless Channel and Related Problems. *AAAC '08*, Hong Kong, China. April 2008.

AWARDS EARNED BY STUDENT ADVISEES

- ◇ Luddy Outstanding Research Award, Nikolai Karpov (Ph.D. student), May, 2021.
- ◇ NeurIPS 2019 Travel Award, Boli Fang (Ph.D. student), December 2019.
- ◇ NeurIPS 2019 Travel Award, Chao Tao (Ph.D. student), December 2019.
- ◇ ICML 2019 Travel Award, Boli Fang (Ph.D. student), June 2019.
- ◇ Purdom Fellowship, Nikolai Karpov (Ph.D. student), September 2019.
- ◇ NSF Travel Award for CIKM 2018, Haoyu Zhang (Ph.D. student), October 2018.
- ◇ NeurIPS 2018 Travel Award, Jiecao Chen (Ph.D. student), December 2018.
- ◇ ACM SIGIR Travel Award, and NSF and ACM SIGWEB Travel Award for CIKM 2017, Haoyu Zhang (Ph.D. student), August 2017.
- ◇ Purdom Fellowship, Boli Fang (Ph.D. student), August 2017.
- ◇ ICML Travel Award, Jiecao Chen (Ph.D. student), August 2017.
- ◇ IU Graduate Research Award, Jiecao Chen (Ph.D. student), April 2017.
- ◇ STOC Travel Award, Jiecao Chen (Ph.D. student), June 2015

STUDENT SUPERVISION

- ◇ Current Ph.D. Students
 - Nikolai Karpov (since April 2018, co-advising with Prof. Cenk Sahinalp)
 - Yan Song (since Fall 2018)
 - Boli Fang (since Spring 2020, co-advising with Prof. David Crandall)
- ◇ Graduated Ph.D. Students (I serve as the *chairperson* in the thesis committee for all)
 - Jiecao Chen (Ph.D. 2019; first employment: Software Engineer & Researcher at Google AI)
Dissertation: *Communication-Efficient Algorithms for Distributed Computation & Machine Learning*
 - Haoyu Zhang (Ph.D. 2020; first employment: Research Scientist at Facebook)
Dissertation: *String Similarity Joins and Search under Edit Distance*
 - Chao Tao (Ph.D. 2021; co-advising with Prof. Yuan Zhou; first employment: Google)
Dissertation: *New Efficient Decision-Making Strategies in Selected Multi-Armed Bandit Problems*
- ◇ Ph.D. Thesis Committees
 - Yifan Pan, 2015 (Advisor: Yuqing Wu)
 - Lei Le, 2019 (Advisor: Martha White)
 - Erfan Sadeqi Azer, 2019 (Advisor: Funda Ergun)
 - Lei Wang, 2019 (Advisor: Haixu Tang)
 - Wontack Han, 2021 (Advisor: Yuzhen Ye)
 - Md. Taufique Hussain, 2021 (Advisor: Ariful Azad at ISE)
- ◇ External Ph.D. Thesis Committees/Reviewers
 - Chenggang Wu, 2014 (at Tsinghua)
 - Yang Li, 2017 (at Upenn)
 - Linxiao Huang, 2017 (at Tsinghua)
 - Yong Gu, 2021 (at Tsinghua)
- ◇ Ph.D. Advisory Committees (all at CS/IU)
 - Erfan Sadeqi Azer (Advisor: Funda Ergun)
 - Inhak Hwang

QIN ZHANG

- ◇ Summer Interns
 - Charilaos Pipis (B.S. at National Technical Univ. of Athens, Greece; through SICE GTAP program), Summer 2020
 - Yanbo Chen (B.S. at Fudan Univ., China; through SICE GTAP program), Summer 2020
 - Yiming Li (B.S. at IU), Summer 2017
 - Grace Li (M.S. at IU), Summer 2016
 - Bo Qin (Ph.D. at HKUST), Summer 2014/2015
 - Di Chen (Ph.D. at HKUST), Summer 2014
- ◇ Undergraduate/Masters Independent Study Supervision (all at IU)
 - Min Chen (M.S. at CS), Spring 2018
 - Rachel Lowden (B.S. at Neuroscience), Spring 2015
 - Zhihua Chen (M.S. at CS), Fall 2014
 - Soheil Jazayeri (M.S. at CS), Spring 2014
 - Wen Chen (M.S. at CS), Spring 2014

UNIVERSITY SERVICE ◇ General Education Committee (2019 - 2020)

SCHOOL SERVICE ◇ SICE Curriculum Committee (2019 - 2020)
◇ Data Science Admissions Committee (2017 - 2018, 2018 - 2019)
◇ Data Science Curriculum Committee (2014 - 2015)
◇ Science Fest Organization Committee (2016)

DEPARTMENT SERVICE ◇ Chair Search Committee (2019 - 2020)
◇ Hiring Committee (2015 - 2016)
◇ Graduate Education Committee (2015 - 2016, 2017 - 2018)
◇ Graduate Admissions and Award Committee (2013 - 2014, 2014 - 2015, 2016 - 2017, 2019 - 2020, 2020 - 2021)
◇ Undergraduate Admissions and Award Committee (2017 - 2018, 2018 - 2019)
◇ Theory Seminar Organizer (2013 - Now)

PROFESSIONAL SERVICES ◇ Journal Editor: Guest Editor, ACM Transactions on Algorithms, SODA 2016 Special Issue

- ◇ Conference/Workshop Program Committees
 - AAAI, AAAI Conference on Artificial Intelligence (2018, 2020)
 - BeyondMR, Algorithms and Systems for MapReduce and Beyond (2018)
 - CIKM, Conference on Information and Knowledge Management (2014, 2015, 2016)
 - DB3, International Workshop on Big Dynamic Distributed Data (2013)
 - ICDE, IEEE International Conference on Data Engineering (2021)
 - ICLR, International Conference on Learning Representations (2018)
 - ICML, International Conference on Machine Learning (2018, 2019, 2020, 2021)
 - IJCAI, International Joint Conference on Artificial Intelligence (2019)
 - IPDPS, International Parallel & Distributed Processing Symposium (2019)
 - ISAAC, International Symposium on Algorithms and Computation (2019)
 - MASSIVE, Workshop on Massive Data Algorithmics (2014, 2015, 2016)
 - NIPS/NeurIPS, Annual Conference on Neural Information Processing Systems (2016, 2017, 2018, 2019, 2020, 2021 (Area Chair))

QIN ZHANG

- PODS, ACM Symposium on Principles of Database System (2017, 2020)
- SIGMOD, ACM International Conference on Management of Data (2021)
- SODA, ACM-SIAM Symposium on Discrete Algorithms (2016)
- SOSA, SIAM Symposium on Simplicity in Algorithms (2021)
- TAMC, Conference on Theory and Applications of Models of Computation (2013)
- VLDB, International Conference on Very Large Data Bases (2021)
- WADS, Algorithms and Data Structures Symposium (2017)
- WAOA, Workshop on Approximation and Online Algorithms (2017)

◇ Journal Review

- Algorithmica
- Distributed and Parallel Databases
- EJOR, European Journal of Operational Research
- IPL, Information Processing Letters
- JACM, Journal of the ACM
- JOCO, Journal of Combinatorial Optimization
- JCSS, Journal of Computer and System Sciences
- SICOMP, SIAM Journal of Computing
- TALG, ACM Transactions on Algorithms
- TDS, ACM Transactions on Data Science
- TKDE, IEEE Transactions on Knowledge and Data Engineering
- TIT, IEEE Transactions on Information Theory
- TOCT, ACM Transactions on Computation Theory
- TODS, ACM Transactions on Database Systems
- VLDB, Very Large Data Bases Journal
- WWWJ, World Wide Web Journal

◇ Conference Review

- ACSC, Australasian Computer Science Conference (2011)
- ALENEX, Algorithm Engineering and Experiments (2011)
- APPROX, International Workshop on Approximation Algorithms for Combinatorial Optimization Problems (2012, 2013, 2015)
- COCOON, International Computing and Combinatorics Conference (2011, 2014)
- COLT, Conference on Learning Theory (2019)
- CPM, Combinatorial Pattern Matching (2017)
- ESA, European Symposium on Algorithms (2013, 2014, 2015, 2018)
- FOCS, IEEE Symposium on Foundations of Computer Science (2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, 2021)
- ICALP, International Colloquium on Automata, Languages and Programming (2009, 2012, 2013, 2015, 2016, 2017, 2018)
- ICDT, The International Conference on Database Theory (2017)
- IPDPS, International Parallel & Distributed Processing Symposium (2015)
- ISAAC, International Symposium on Algorithms and Computation (2012, 2013, 2014, 2015, 2018)

QIN ZHANG

- ITCS, Innovations in Theoretical Computer Science (2018)
 - PODS, ACM Symposium on Principles of Database System (2010, 2012, 2014, 2015, 2018, 2021)
 - RANDOM, International Workshop on Randomization and Computation (2015)
 - SOCG, ACM Symposium on Computational Geometry (2011, 2013, 2018, 2019, 2020)
 - SODA, ACM-SIAM Symposium on Discrete Algorithms (2008, 2011, 2012, 2013, 2014, 2015, 2017, 2018, 2019, 2020, 2021)
 - SOSA, Symposium on Simplicity in Algorithms (2018)
 - STACS, Symposium on Theoretical Aspects of Computer Science (2007, 2010, 2012, 2013, 2014, 2017)
 - STOC, ACM Symposium on Theory of Computing (2009, 2014, 2016, 2017, 2019, 2020, 2021)
 - WINE, Conference on Web and Internet Economics (2010)
- ◇ Funding Agency Service:
- Proposal review for the Danish Council for Independent Research (2014)
 - Panel for National Science Foundation (2015, 2018)
 - Proposal review for the United States – Israel Binational Science Foundation (2016)
 - Proposal review for the Israel Science Foundation (2016)
 - Proposal review for the Research Commission of the ETH Zurich (2021)
- ◇ Workshop Organization: 67th Midwest Theory Day (April 15-16, 2017)