Curriculum Vitae
Justin Thaler

Personal Information

Full Name: Justin Ross Thaler
Email: justin.thaler@georgetown.edu
URL: http://people.cs.georgetown.edu/jthaler/

Research Interests

Verifiable Computation, Algorithms for Massive Data Sets, Computational Learning Theory, Quantum Computing

Education

Ph.D. November 2013 - School of Engineering and Applied Sciences,
Harvard University, Cambridge MA
Adviser: Michael Mitzenmacher

B.S. 2005-2009 - Yale University, New Haven, CT
Summa Cum Laude

Positions

- **Assistant Professor. Department of Computer Science, Georgetown University, Washington D.C. August 2016-present.**
- **Research Scientist. Scalable Machine Learning Group. Yahoo Labs, New York, NY. June 2014-July 2016.**
- **Research Fellow. Simons Institute for the Theory of Computing, Berkeley, CA. August 2013-May 2014**.
- **Ph.D. Student. Harvard University, Cambridge, MA. September 2009-July 2013**.
- **Adjunct at Center for Computing Sciences, Institute for Defense Analyses, Bowie, MD (2009-2014), Research Intern during Summer 2009.**

Academic Honors

- NSF CAREER Award. Project Title: The Polynomial Method in Complexity and Cryptography. (2019).
- Best Newcomer Paper Award, *International Conference on Database Theory (ICDT)* (2016).
- Best Paper Award, *Symposium on Parallel Algorithms and Architectures (SPAA)* (2014).
- Best Paper Award, *International Colloquium on Automata, Languages, and Programming (ICALP)*, Track A (2013).
- NSF Graduate Research Fellowship Recipient (2010).
- National Defense Science and Engineering Graduate Fellow (2009-2012).
- Phi Beta Kappa (2009).
- Yale University Computer Science Prize (2009). Awarded by Yale University's Department of Computer Science to the graduating senior who ranks highest in scholarship.
- Anthony D. Stanley Memorial Prize (2009). Awarded by Yale University's Department of Mathematics for excellence in pure and applied mathematics.

Publications

Journal Papers

(J1) **Dual Polynomials for Collision and Element Distinctness**. Mark Bun and Justin Thaler. *Theory of Computing*, 2016.

(J2) **Parallel Peeling Algorithms**. Jiayang Jiang, Michael Mitzenmacher, and Justin Thaler. *ACM Transactions on Parallel Computing*. (Special issue devoted to *SPAA* 2014, Extended version of C20).

(J3) **Dual Lower Bounds for Approximate Degree and Markov-Bernstein Inequalities**. Mark Bun and Justin Thaler. *Information and Computation*, 2015. (Special issue devoted to *ICALP* 2014, Extended Version of C25).

(J4) **Annotations in Data Streams**. Amit Chakrabarti, Graham Cormode, Andrew McGregor, and Justin Thaler. *ACM Transactions on Algorithms*, 2014.

(J5) **Streaming Graph Computations with a Helpful Advisor**. Graham Cormode, Michael Mitzenmacher, and Justin Thaler. *Algorithmica*, 2013. (Extended Version of C35).

(J6) **External-Memory Multimaps**. Elaine Angelino, Michael T. Goodrich, Michael Mitzenmacher, and Justin Thaler. *Algorithmica*, 2013. (Special issue devoted to *ISAAC* 2011, Extended Version of C34).

Conference Papers

(C1) **Quantum Algorithms and Approximating Polynomials for Composed Functions with Shared Inputs.** Mark Bun, Robin Kothari, and Justin Thaler. In *Symposium on Discrete Algorithms (SODA)*, 2019.

(C2) **Approximate Degree and the Complexity of Depth Three Circuits**. Mark Bun and Justin Thaler. In *International Conference on Randomization and Computation (RANDOM)*, 2018.

(C3) **Doubly-efficient zkSNARKs without trusted setup**. Riad S. Wahby, Ioanna Tzialla, abhi shelat, Justin Thaler and Michael Walfish. In *IEEE Symposium on Security and Privacy (S&P)*, 2018.

(C4) **The Polynomial Method Strikes Back: Tight Quantum Query Bounds via Dual Polynomials**. Mark Bun, Robin Kothari, and Justin Thaler. In *ACM Symposium on the Theory of Computing (STOC)*, 2018. Also presented at the 2018 *Conference on Quantum Information Processing (QIP)* as a **plenary talk**. **Invited to Theory of Computing**.

(C5) **A Nearly Optimal Lower Bound on the Approximate Degree of AC^0**. Mark Bun and Justin Thaler. In *IEEE Symposium on Foundations of Computer Science* (*FOCS*), 2017. **Invited to SICOMP (Special Issue for FOCS 2017)**.

(C6) **On the Power of Statistical Zero Knowledge**. Adam Bouland, Lijie Chen, Dhiraj Holden, Justin Thaler, Prashant Nalini Vasudevan. In *IEEE Symposium on Foundations of Computer Science* (*FOCS*), 2017. **Invited to SICOMP (Special Issue for FOCS 2017)**.

(C7) **Full Accounting for Verifiable Outsourcing**. Riad S. Wahby, Ye Ji, Andrew J. Blumberg, abhi shelat, Justin Thaler, Michael Walfish, and Thomas Wies. In *Conference on Computer and Communications Security (CCS)*, 2017.

(C8) **A High-Performance Algorithm for Identifying Frequent Items in Data Streams**. Daniel Anderson, Pryce, Bevin, Kevin Lang, Edo Liberty, Lee Rhodes, and Justin Thaler. In *Internet Measurement Conference (IMC)*, 2017.

(C9) **Reliably Learning the ReLU in Polynomial Time**. Surbhi Goel, Varun Kanade, Adam Klivans, and Justin Thaler. In *Conference on Learning Theory (COLT)*, 2017.

(C10) **Determining Tournament Payout Structures for Daily Fantasy Sports**. Christopher Musco, Maxim Sviridenko, and Justin Thaler. In *Meeting on Algorithm, Engineering & Experiments (ALENEX)*, 2017. **Invited to ACM Journal of Experimental Algorithmics (Special Issue for ALENEX 2017)**.

(C11) **Improved Bounds on the Sign-Rank of AC^0**. Mark Bun and Justin Thaler. In *International Colloquium on Automata, Languages and Programming (ICALP)*, 2016.

(C12) **Lower Bounds for the Approximate Degree of Block-Composed Functions**. Justin Thaler. In *International Colloquium on Automata, Languages and Programming (ICALP)*, 2016.

(C13) **Semi-Streaming Algorithms for Annotated Graph Streams**. Justin Thaler. In *International Colloquium on Automata, Languages and Programming (ICALP)*, 2016.

(C14) **Space Lower Bounds for Itemset Frequency Sketches**. Edo Liberty, Michael Mitzenmacher, Justin Thaler, and Jonathan Ullman. In *Principles of Database Systems (PODS)*, 2016.

(C15) **A Framework for Estimating Stream Expression Cardinalities**. Anirban Dasgupta, Kevin Lang, Lee Rhodes, and Justin Thaler. In *International Conference on Database Theory (ICDT)*, 2016. **Best Newcomer Paper Award. Invited to ACM Transactions on Database Systems (Special Issue for ICDT 2016).**

(C16) **Streaming Verification in Data Analysis**. Samira Daruki, Justin Thaler, and Suresh Venkatasubramanian. In *International Symposium on Algorithms and Computation (ISAAC)*, 2015.

(C17) **Hardness Amplification and the Approximate Degree of Constant-Depth Circuits**. Mark Bun and Justin Thaler. In *International Colloquium on Automata, Languages and Programming (ICALP)*, 2015.

(C18) **Variable Selection is Hard**. Dean Foster, Howard Karloff, and Justin Thaler. In *Conference on Learning Theory (COLT)*, 2015.

(C19) **Verifiable Stream Computation and Arthur-Merlin Communication**. Amit Chakrabarti, Graham Cormode, Andrew McGregor, Justin Thaler, and Suresh Venkatasubramanian. In *Computational Complexity Conference (CCC)*, 2015.

(C20) **Parallel Peeling Algorithms**. Jiayang Jiang, Michael Mitzenmacher, and Justin Thaler. In *Symposium on Parallelism in Algorithms and Architectures (SPAA)*, 2014. **Best Paper Award. Invited to ACM Transactions on Parallel Computing (special issue for SPAA 2014)**.

(C21) **Distribution-Independent Reliable Learning**. Varun Kanade and Justin Thaler. In *Conference on Learning Theory (COLT)*, 2014.

(C22) **Faster Private Release of Marginals on Small Databases**. Karthekeyan Chandrasekaran, Justin Thaler, Jonathan Ullman, Andrew Wan. In *Innovations in Theoretical Computer Science (ITCS)*, 2014.

(C23) **Annotations for Sparse Data Streams**. Amit Chakrabarti, Graham Cormode, Navin Goyal, and Justin Thaler. In *Symposium on Discrete Algorithms (SODA)*, 2014.

(C24) **Time-Optimal Interactive Proofs for Circuit Evaluation**. Justin Thaler. In *International Cryptology Conference (CRYPTO)*, 2013.

(C25) **Dual Lower Bounds for Approximate Degree and Markov-Bernstein Inequalities**. Mark Bun and Justin Thaler. In *International Colloquium on Automata, Languages and Programming (ICALP)*, 2013. **Best Paper Award for Track A. Invited to Information and Computation (special issue for ICALP 2013).**

(C26) **Cache-Oblivious Dictionaries and Multimaps with Negligible Failure Probability**. Michael Goodrich, Dan Hirschberg, Michael Mitzenmacher, and Justin Thaler. In *Mediterranean Conference on Algorithms (MedAlg)*, 2012.

(C27) **Verifying Computations with Streaming Interactive Proofs**. Graham Cormode, Justin Thaler, and Ke Yi. In *VLDB*, 2011.

(C28) **Faster Algorithms for Privately Releasing Marginals**. Justin Thaler, Jonathan Ullman, and Salil Vadhan. In *International Colloquium on Automata, Languages and Programming (ICALP)*, 2012.

(C29) **Attribute-Efficient Learning and Weight-Degree Tradeoffs for Polynomial Threshold Functions**. Rocco Servedio, Li-Yang Tan, and Justin Thaler. In *Conference on Learning Theory (COLT)*, 2012.

(C30) **Verifiable Computation with Massively Parallel Interactive Proofs**. Justin Thaler, Mike Roberts, Michael Mitzenmacher, and Hanspeter Pfister. In *USENIX Workshop on Hot Topics in Cloud Computing (HotCloud)*, 2012.

(C31) **Continuous Time Channels with Interference**. Ioana Ivan, Michael Mitzenmacher, Justin Thaler, and Henry Yuen. In *International Symposium on Information Theory (ISIT)*, 2012.

(C32) **Hierarchical Heavy Hitters with the Space Saving Algorithm**. Michael Mitzenmacher, Thomas Steinke, and Justin Thaler. In *Meeting on Algorithm, Engineering & Experiments (ALENEX)*, 2012.

(C33) **Practical Verified Computation with Streaming Interactive Proofs**. Graham Cormode, Michael Mitzenmacher, and Justin Thaler. In *Innovations in Theoretical Computer Science (ITCS)*, 2012.

(C34) **External-Memory Multimaps**. Elaine Angelino, Michael T. Goodrich, Michael Mitzenmacher, and Justin Thaler. In *International Symposium on Algorithms and Computation (ISAAC)*, 2011.

(C35) **Streaming Graph Computations with a Helpful Advisor**. Graham Cormode, Michael Mitzenmacher, and Justin Thaler. In *European Symposium on Algorithms (ESA)*, 2010.

(C36) **Graph Covers and Quadratic Minimization**. Nicholas Ruozzi, Justin Thaler, and Sekhar Tatikonda. In *Allerton Conference on Communication, Control, and Computing*, 2009.

Invited Papers and Workshop Papers

(I1) **Catching Lies (and Mistakes) in Offloaded Computation**. Michael Mitzenmacher and Justin Thaler. Communications of the ACM (CACM), February 2016. (Invited Technical Perspective).

(I2) **Data Stream Verification**. Justin Thaler. *Encyclopedia of Algorithms*. Springer Berlin Heidelberg, 2015. (Invited Survey Article).

(I3) **Peeling Arguments and Double Hashing**. Michael Mitzenmacher and Justin Thaler. Appears as an invited paper in *Allerton Conference on Communication, Control, and Computing*, 2012.

(I4) **On the Zero-Error Capacity Threshold for Deletion Channels**. Ian Kash, Michael Mitzenmacher, Justin Thaler, and Jonathan Ullman. In *Information Theory and Applications Workshop (ITA)*, 2011.

Manuscripts

(M1) **The Large-Error Approximate Degree of $\mathbf{AC^0}$.** Mark Bun and Justin Thaler. 2018.

(M2) **A Note on the GKR Protocol.** Justin Thaler. 2015.

(M3) **Verifiable Computation Using Multiple Provers.** Andrew J. Blumberg, Justin Thaler, Victor Vu, and Michael Walfish. 2014.

Selected Invited Talks and Lecture Series

- **Approximate Degree: A Survey**.
  - Invited talk at CMO 2018 workshop on Analytic Techniques in Theoretical Computer Science
- **The Polynomial Method Strikes Back**.
  - Invited talk at the Harvard/MIT/MSR Theory Reading Group (October 2017)
  - Invited talk at NYU Theory Seminar (December 2017)
  - Invited talk at University of Maryland's QuICS Seminar (January 2017)
- **Verifiable Computing: Between Theory and Practice**.
  - Invited survey talk at the *STOC* 2017 workshop on *Probabilistically Checkable and Interactive Proofs (PCP/IP): Between Theory and Practice* (June 2017).
- **A Nearly Optimal Lower Bound on the Approximate Degree of $\mathbf{AC^0}$**.
  - Invited talk at BIRS 2017 workshop on Communication Complexity and its Applications II
  - Invited talk at Columbia University Theory Seminar (April 2017)
- **Chebyshev Polynomials, Approximate Degree, and their Applications**.
  - Invited survey talk at the *FOCS* 2016 workshop on *(Some) Orthogonal Polynomials and their Applications to TCS*.
- **Verifiable Computation.**
  - Invited speaker at the Fourteenth Bellairs' Crypto-Workshop (Invited Lecture Series). Delivered 15 hours of lectures surveying modern techniques for proof-based verifiable computation. Workshop organized by Claude Crépeau.
- **Interactive Proofs and Argument Systems.**
  - Invited speaker at the Summer School on Secure and Oblivious Computation and Outsourcing at the University of Notre Dame. Delivered 3 hours of lectures surveying modern techniques for proof-based verifiable computation. Summer school organized by Marina Blanton.

Grants and Funding

- **NSF CAREER Award #1845125**.
  - Project Title: The Polynomial Method in Complexity and Cryptography.
  - NSF Program: Algorithmic Foundations.
  - Role: PI.
  - Amount: $549,045.
  - Dates: June 2019-May 2024.

- **Research Seed Grant from Georgetown University's Massive Data Institute**.
- Project Title: Enabling Analysis of Sensitive Data via Zero-Knowledge Proofs.
- Role: PI.
- Amount: $40,000.
- Dates: July 2017-June 2018

U.S. Patent Applications

- **Automatic Fantasy Sports Data Analysis Method and Apparatus**. Michael Lazarus, Maxim Sviridenko, and Justin Thaler. Application number US20180001215A1. Application filed June 30, 2016 and published January 4, 2018.
- **Fantasy Sports Data Analysis for Game Structure Development**. Justin Thaler, Maxim Sviridenko, Edo Liberty, Ron Belmarch, Jerry Shen, and Prerit Uppal. Application number US20170095739A1. Application filed October 6, 2015 and published April 6, 2017.

Open Source Software Development

- Co-creator and core contributor to an open source library of highly optimized streaming algorithms, called *DataSketches: A Java software library of stochastic streaming algorithms* (url: `https://datasketches.github.io`). The library is used within several companies, including Oath/Yahoo, Amazon, and Splice Machine. It also has been incorporated into a popular open source graph database library called Gaffer that is maintained by the British intelligence agency GCHQ, and into a low-latency open source data store called Druid. As of October 2018, the library is averaging over 55,000 downloads per month from Maven Central.

External Service and Professional Activities

- **PC Member**: STOC 2019, SOSA 2019, SODA 2018, FSTTCS 2017, ICALP 2016, ALENEX 2016, SDM 2015.
- **Grant Reviews and Panels**: National Science Foundation Panel (2017), External reviewer for Research Grants Council of Hong Kong (2018).
- **Workshop Organization**: Co-organizer of Capital Area Theory Day (2018). Co-organizer and chair of the 5-day 2019 workshop on Probabilistically Checkable and Interactive Proof Systems at Simons Institute for the Theory of Computing.
- **Journal reviewer**: *SICOMP*, *Computational Complexity*, *Journal of Cryptology*, *Theory of Computing*, *SIAM Journal on Discrete Mathematics*, *Information and Computation*, *Algorithmica*, *Communications of the ACM*, *Discrete and Computational Geometry*, *Theoretical Computer Science*, *Frontiers in ICT (Big Data Section)*, *Information Processing Letters*.
- **External conference reviewer**: STOC, FOCS, CCC, SODA, CRYPTO, ICALP, ITCS, RANDOM, PODS, NIPS, ICDT, PODC, DISC, TCC, ESORICS, ESA.

Student Advising and Committee Membership

- Ph.D. advisor for Pryce Bevan, currently a second-year Ph.D. student in Georgetown University's Department of Computer Science (Fall 2017-present).
- Ph.D. advisor for Shuchen Zhu, currently a first-year Ph.D. student in Georgetown University's Department of Computer Science (Spring 2018-present).
- Mentor of postdoctoral scholar Nikhil Mande (Spring 2019-present).
- Co-mentor of postdoctoral scholar Michael Clear, joint with Adam O'Neill (Fall 2017-Fall 2018).
- Hosted 5-week visit from Nikhil Mande, graduate student at Tata Institute of Fundamental Research.
- Ph.D. qualifying examination committee member for Jeffrey Cohn, Georgetown University Department of Physics (December 2017).
- Ph.D. proposal and thesis committee member. Samira Daruki, University of Utah. Proposal defense occurred in September 2015, and thesis defense in May 2017.
- Mentored two Georgetown University undergraduates, Pryce Bevan and Daniel Anderson, on a year-long research project on streaming algorithms that led to publication (C8) above.
- Co-host for Yahoo Labs intern Christopher Musco, Ph.D. student at MIT (Summer 2015).

Teaching

- COSC 548 – Streaming Algorithms (Fall 2016, Fall 2018)
- ANLY 550 – Structures and Algorithms for Analytics (Spring of: 2017, 2018 (2 sections), 2019)
- COSC 544 – Probabilistic Proof Systems (Fall 2017)

Service to Georgetown University

- CS Colloquium Committee (Fall 2016-present).
- Steering Committee Member for Georgetown's M.S. in Analytics (Fall 2017 and Spring 2018).
- Computer Science Faculty Advisory Committee (Fall 2018-present)
- Admissions Committee Member for Georgetown's M.S. in Analytics (Spring 2017). Personally reviewed more than 250 applications over 13 weeks.