

Thomas Steinke

Research Staff Member
IBM Research – Almaden

cv@thomas-steinke.net
<http://www.thomas-steinke.net/>

Research Interests

Broadly, my research interests lie in the foundations of computer and data science. More specifically, my current research focuses are data privacy (namely, differential privacy) and its connection to learning theory (specifically, adaptive data analysis). I also work in the area of pseudorandomness, with an emphasis on pseudorandom generators, small-space derandomization, and discrete Fourier analysis.

Positions & Education

IBM Almaden Research Center, San Jose, CA, USA, 2016–present.

Research Staff Member. 2018–present

Postdoctoral Researcher. 2016–2018

Harvard University, Cambridge, MA, USA, 2010–2016.

PhD & SM in Computer Science.

Thesis: Upper and Lower Bounds for Privacy and Adaptivity in Algorithmic Data Analysis.

Advisor: Prof. Salil P. Vadhan.

University of Canterbury, Christchurch, New Zealand, 2007–2010.

MSc in Mathematics.

Thesis: Constructive Notions of Compactness in Apartness Spaces.

Supervisor: Prof. Douglas S. Bridges.

Awarded with distinction.

BSc(Hons) in Mathematics and Computer Science.

First class honours. (*GPA*: 9.00 out of 9.00.)

Manuscripts

Kobbi Nissim, Adam Smith, Thomas Steinke, Uri Stemmer, Jonathan Ullman. *The Limits of Post-Selection Generalization*. Manuscript 2018. (In submission.)

<https://arxiv.org/abs/1806.06100>

Mark Bun, Jarosław Błasiok, Aleksandar Nikolov, Thomas Steinke. *Towards an Optimal Algorithm for Concentrated Differential Privacy*. Manuscript 2017. (Poster at Third Workshop on the Theory and Practice of Differential Privacy (TPDP) 2017.)

Thomas Steinke, Jonathan Ullman. *Subgaussian Tail Bounds via Stability Arguments*. Manuscript 2017. (In submission.)

<https://arxiv.org/abs/1701.03493>

Publications

Vitaly Feldman, Thomas Steinke. *Calibrating Noise to Variance in Adaptive Data Analysis*. 31st Annual Conference on Learning Theory (COLT) 2018.

<https://arxiv.org/abs/1712.07196>

Mark Bun, Cynthia Dwork, Guy N. Rothblum, Thomas Steinke. *Composable and Versatile Privacy via Truncated CDP*. 50th Annual ACM Symposium on the Theory of Computing (STOC) 2018. (Poster at Third Workshop on the Theory and Practice of Differential Privacy (TPDP) 2017.)

Invited to special issue of Journal of Privacy and Confidentiality.

Thomas Steinke, Jonathan Ullman. *Tight Lower Bounds for Differentially Private Selection*. 58th Annual IEEE Symposium on Foundations of Computer Science (FOCS) 2017.

<https://arxiv.org/abs/1704.03024>

Vitaly Feldman, Thomas Steinke. *Generalization for Adaptively-Chosen Estimators via Stable Median*. 30th Annual Conference on Learning Theory (COLT) 2017.

<https://arxiv.org/abs/1706.05069>

Mark Bun, Thomas Steinke, Jonathan Ullman. *Make Up Your Mind: The Price of Online Queries in Differential Privacy*. ACM-SIAM Symposium on Discrete Algorithms (SODA) 2017.

<https://arxiv.org/abs/1604.04618>

Mark Bun, Thomas Steinke. *Concentrated Differential Privacy: Simplifications, Extensions, and Lower Bounds*. Theory of Cryptography Conference (TCC) 2016.

<https://arxiv.org/abs/1605.02065>

Raef Bassily, Kobbi Nissim, Adam Smith, Thomas Steinke, Uri Stemmer, Jonathan Ullman. *Algorithmic Stability for Adaptive Data Analysis*. 48th ACM Annual Symposium on the Theory of Computing (STOC) 2016.

<http://arxiv.org/abs/1511.02513>

Invited to special issue of SIAM Journal on Computing. (In submission.)

Cynthia Dwork, Adam Smith, Thomas Steinke, Jonathan Ullman, Salil Vadhan. *Robust Traceability from Trace Amounts*. 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS) 2015.

<http://ieeexplore.ieee.org/document/7354420>

Thomas Steinke, Jonathan Ullman. *Between Pure and Approximate Differential Privacy*. Journal of Privacy and Confidentiality 2017. (Presented at First Workshop on the Theory and Practice of Differential Privacy (TPDP) 2015.)

<http://arxiv.org/abs/1501.06095>

Mark Bun, Thomas Steinke. *Weighted Polynomial Approximations: Limits for Learning and Pseudorandomness*. 19th International Workshop on Randomization and Computation (RANDOM) 2015.

<http://arxiv.org/abs/1412.2457>

Thomas Steinke, Jonathan Ullman. *Interactive Fingerprinting Codes and the Hardness of Preventing False Discovery*. 28th Annual Conference on Learning Theory (COLT) 2015.

<http://arxiv.org/abs/1410.1228>

Thomas Steinke, Salil Vadhan, Andrew Wan. *Pseudorandomness and Fourier Growth Bounds for Width 3 Branching Programs*. Theory of Computing 2017. (Originally appeared in 18th International Workshop on Randomization and Computation (RANDOM) 2014.)

<http://theoryofcomputing.org/articles/v013a012/>

Invited to special issue.

Omer Reingold, Thomas Steinke, Salil Vadhan. *Pseudorandomness for Regular Branching Programs via Fourier Analysis*. 17th International Workshop on Randomization and Computation (RANDOM) 2013. <http://arxiv.org/abs/1306.3004>

Thomas Steinke. *Pseudorandomness for Branching Programs Without the Group Theory*. Electronic Colloquium on Computational Complexity (ECCC) 2012. <http://eccc.hpi-web.de/report/2012/083/>

Michael Mitzenmacher, Justin Thaler, Thomas Steinke. *Hierarchical Heavy Hitters with the Space Saving Algorithm*. Algorithm Engineering and Experiments (ALENEX) 2012. <http://arxiv.org/abs/1102.5540>

Varun Kanade, Thomas Steinke. *Learning Hurdles for Sleeping Experts*. ACM Transactions on Computation Theory 2014. (Originally appeared in Innovations in Theoretical Computer Science (ITCS) 2012.) <http://eccc.hpi-web.de/report/2011/115/>

Invited to special issue.

Raazesh Sainudiin, Thomas Steinke. *A Rigorous Extension of the Schönhage-Strassen Integer Multiplication Algorithm Using Complex Interval Arithmetic*. Reliable Computing 2013. (Presented at Seventh International Conference Computability and Complexity in Analysis (CCA) 2010.) <http://arxiv.org/abs/1006.0405>

Theses & Surveys

Thomas Steinke. *Upper and Lower bounds for Privacy and Adaptivity in Algorithmic Data Analysis*. PhD Thesis, Harvard University 2016. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:33840662>

Cynthia Dwork, Adam Smith, Thomas Steinke, Jonathan Ullman. *Exposed! A Survey of Attacks on Private Data*. Annual Review of Statistics and Its Application 2017. <http://www.annualreviews.org/doi/abs/10.1146/annurev-statistics-060116-054123>

Kobbi Nissim, Thomas Steinke, Alexandra Wood, Mark Bun, Marco Gaboardi, David R. O'Brien, Salil Vadhan. *Differential Privacy: A Primer for a Non-technical Audience*. Manuscript 2017. (Abstract accepted to Privacy Law Scholars Conference. Berkeley, CA, USA, June 2017.) <http://privacytools.seas.harvard.edu/files/privacytools/files/pedagogical-document-dp.pdf>

Thomas Steinke. *Constructive Notions of Compactness in Apartness Spaces*. Masters Thesis, University of Canterbury 2011. <http://ir.canterbury.ac.nz/handle/10092/5682>

Teaching

2015: Teaching Fellow at Harvard University, Pseudorandomness (CS225).
Awarded Certificate of Distinction for Teaching.

2015: Advised undergraduate summer students Ally Kaminsky and Jimmy Jiang as part of NSF Privacy Tools for Sharing Research Data project.

2014: Advised undergraduate summer student Sitan Chen. Project title: Pseudorandomness for Read-Once, Constant-Depth Formulas.

2013: Teaching Fellow at Harvard University, Algorithms for Big Data (CS229r).

2013–2016: Non-resident tutor in Lowell House at Harvard University.

2013: Tutor and lecturer at the New Zealand Olympiad Informatics (NZOI) camp in Auckland, New Zealand. (This is a training camp for high school students who will represent New Zealand at the International Olympiad in Informatics (IOI).)

2012: Head Teaching Fellow at Harvard University, Introduction to the Theory of Computation (CS121).
Awarded Certificate of Distinction for Teaching.

2010: Mathematics Tutor at the University of Canterbury, second-year real analysis and calculus (MATH243, MATH264).

2010: Mentor for Jamie McCloskey and Logan Glasson, who both won bronze medals for New Zealand at the IOI, in Waterloo, Canada.

2010: Tutor and lecturer at the NZOI camp in Auckland, New Zealand.

2009: Mathematics Tutor at the University of Canterbury, second-year calculus and linear algebra (MATH254, MATH264).

2009: Tutor at the NZOI camp in Auckland, New Zealand.

2008: Mathematics Tutor at the University of Canterbury, first-year calculus and linear algebra (MATH108, MATH109).

2008: Tutor at NZOI camp in Christchurch, New Zealand.

Research Talks

Banff International Research Station. *Concentrated Differential Privacy*. Banff, AB, Canada, April 2018.
<http://www.birs.ca/events/2018/5-day-workshops/18w5189/videos>

University of Waterloo. *Protecting Privacy and Guaranteeing Generalization by Controlling Information*. Waterloo, ON, Canada, April 2018. (Invited talk.)

University of Maryland. *Protecting Privacy and Guaranteeing Generalization by Controlling Information*. College Park, MD, USA, April 2018. (Invited talk.)

Georgetown University. *Protecting Privacy and Guaranteeing Generalization by Controlling Information*. Washington, DC, USA, April 2018. (Invited talk.)

University of Massachusetts. *Protecting Privacy and Guaranteeing Generalization with Algorithmic Stability*. Amherst, MA, USA, February 2018. (Invited talk.)

North Carolina State University. *Protecting Privacy and Guaranteeing Generalization with Algorithmic Stability*. Raleigh, NC, USA, February 2018. (Invited talk.)

Ohio State University. *Protecting Privacy and Guaranteeing Generalization with Algorithmic Stability*. Columbus, OH, USA, January 2018. (Invited talk.)

Boston University. *Protecting Privacy and Guaranteeing Generalization with Algorithmic Stability*. Boston, MA, USA, January 2018. (Invited talk.)

Harvard differential privacy research group meeting. *Relaxing differential privacy for adaptive data analysis*. Cambridge, MA, USA, January 2018. (Invited talk.)

Theory of Computing Affiliated - Silicon Valley (TOCA-SV) workshop. *Less is more: Limiting information to guarantee generalization in adaptive data analysis*. Stanford, CA, USA, January 2018. (Invited talk.)

IPAM workshop on Algorithmic Challenges in Protecting Privacy for Biomedical Data. *How well does privacy compose?* Los Angeles, CA, USA, January 2018. (Invited talk.)
<http://www.ipam.ucla.edu/abstract/?tid=15037&pcode=PBD2018>

Simons Institute for the Theory of Computing. *Preventing Overfitting in Adaptive Data Analysis via Stability*. Berkeley, CA, USA, December 2017. (Invited talk.)
<https://simons.berkeley.edu/talks/thomas-steinke-12-1-17>

Theory and Practice of Differential Privacy Workshop (part of ACM CCS 2017). *Concentrating on the Foundations of Differential Privacy*. Dallas, TX, USA, October 2017. (Invited speaker.)

IEEE Symposium on Foundations of Computer Science (FOCS) 2017. *Tight Lower Bounds for Differentially Private Selection*. Berkeley, CA, USA, October 2017.

Conference on Learning Theory 2017. *Generalization for Adaptively-chosen Estimators via Stable Median*. Amsterdam, Netherlands, July 2017.
<https://webcolleges.uva.nl/Mediasite/Play/6123f099ac8c45c4939a93230a85aaf61d>

Simons Institute for the Theory of Computing. *Pseudorandom Generators for Small Space via Fourier Analysis*. Berkeley, CA, USA, March 2017. (Invited talk.)
<https://simons.berkeley.edu/talks/thomas-steinke-2017-03-09>

University of Oxford. *Generalisation for Adaptive Data Analysis*. Oxford, UK, November 2016. (Invited talk.)

Isaac Newton Institute, University of Cambridge. *Generalisation for Adaptive Data Analysis*. Cambridge, UK, November 2016. (Invited talk.)
<https://www.newton.ac.uk/seminar/20161122153016302>

Theory of Cryptography Conference (TCC) 2016-B. *Concentrated Differential Privacy, Simplifications, Extensions, and Lower Bounds*. Beijing, China, November 2016.

World Congress on Probability and Statistics. *Robust Traceability from Trace Amounts*. Toronto, Canada, July 2016. (Invited talk.)

ACM Symposium on the Theory of Computing (STOC) 2016. *Algorithmic Stability for Adaptive Data Analysis*. Cambridge, MA, USA, June 2016.

Information Theory and Applications Workshop (ITA) 2016. *Interactive Fingerprinting Codes and the Hardness of Preventing False Discovery*. San Diego, CA, USA, February 2016. (Invited talk.)

The Chinese University of Hong Kong. *The Power of Adaptivity in Data Analysis*. Hong Kong, January 2016. (Invited talk.)

IEEE Symposium on Foundations of Computer Science (FOCS) 2015. *Robust Traceability from Trace Amounts*. Berkeley, CA, USA, October 2015.
<http://techtalks.tv/beta/talks/robust-traceability-from-trace-amounts/62042/>

Academia Sinica. *On the Power of Adaptivity in Data Analysis*. Taipei, Taiwan, August 2015. (Invited talk.)

China Theory Week 2015. *On the Power of Adaptivity in Data Analysis*. Shanghai, China, August 2015. (Invited talk.)

Conference on Learning Theory (COLT) 2015. *Interactive Fingerprinting Codes and the Hardness of Preventing False Discovery*. Paris, France, July 2015.

http://videlectures.net/colt2015_steinke_false_discovery/

Theory and Practice of Differential Privacy (TPDP) 2015. *Between Pure and Approximate Differential Privacy*. London, UK, April 2015.

Workshop on Randomization and Computation (RANDOM) 2014. *Pseudorandomness and Fourier Growth Bounds for Width 3 Branching Programs*. Barcelona, Spain, September 2014.

MIT Algorithms & Complexity Seminar. *Pseudorandomness for Regular Branching Programs via Fourier Analysis*. Cambridge, MA, USA December 2013. (Invited talk.)

Workshop on Randomization and Computation (RANDOM) 2013. *Pseudorandomness for Regular Branching Programs via Fourier Analysis*. Berkeley, CA, USA, August 2013.

Microsoft Research Silicon Valley Theory Seminar. *Pseudorandomness for Regular Branching Programs via Fourier Analysis*. Mountain View, CA, USA, August 2013. (Invited talk.)

Algorithm Engineering and Experiments (ALENEX) 2012. *Hierarchical Heavy Hitters with the Space Saving Algorithm*. Kyoto, Japan, January 2012.

Stanford Theory Seminar. *Learning Hurdles for Sleeping Experts*, Stanford, CA, USA, November 2011.

Computability and Complexity in Analysis (CCA) 2011. *Constructive Notions of Compactness in Apartness Spaces*. Cape Town, South Africa, January 2011.

Computability and Complexity in Analysis (CCA) 2010. *A Rigorous Extension of the Schönhage-Strassen Integer Multiplication Algorithm Using Complex Interval Arithmetic*. Zhenjiang, China, June 2010.

Miscellaneous

2018: Co-Organiser of workshop on Mathematical Foundations of Data Privacy at the Banff International Research Station, Banff, Canada. <http://www.birs.ca/events/2018/5-day-workshops/18w5189>

2014–2016: Co-Organizer of Harvard Theory of Computing Seminar.

2016: Nominated by department for Derek C. Bok Award for Excellence in Graduate Student Teaching of Undergraduates.

2010–2013: Lord Rutherford Memorial Research Fellowship.

2010: ACM International Collegiate Programming Contest (ACM-ICPC) in Harbin, China. Qualified for world finals as New Zealand champions and second in the south pacific region (Australia & New Zealand).

2009: ACM-ICPC in Stockholm, Sweden. New Zealand Champions and second in the south pacific region.

2007: Member of New Zealand International Olympiad in Informatics (IOI) team that went to Zagreb, Croatia.

Programming: C++, Python.

Languages: English, German (working knowledge), French (very limited), Mandarin Chinese (started learning).