

Julian Shun

✉ jshun@eecs.berkeley.edu
www.eecs.berkeley.edu/~jshun

Current Employment

Aug 2015 - **Miller Postdoctoral Research Fellow**, *University of California, Berkeley*, Berkeley, CA.
Current

Research Interests

Large-scale graph analytics and frameworks; parallel programming techniques and tools; parallel algorithms and data structures

Education

Aug 2009 - **Ph.D. in Computer Science**, *Carnegie Mellon University*, Pittsburgh, PA.

May 2015 Thesis: Shared-Memory Parallelism Can Be Simple, Fast, and Scalable
Advisor: Guy Blelloch

Received the ACM Doctoral Dissertation Award (one recipient in 2015)

Received the CMU SCS Doctoral Dissertation Award (one recipient in 2015)

Aug 2004 - **B.A. in Computer Science**, *University of California, Berkeley*, Berkeley, CA.

May 2008 GPA: 3.98/4.0

Ranked 1st in the 2008 graduating class of Computer Science with over 100 students

Awards

2015 Recipient of the ACM Doctoral Dissertation Award (one recipient in 2015)

2015 Recipient of the CMU SCS Doctoral Dissertation Award (one recipient in 2015)

2015–
Current Recipient of a Miller Research Fellowship for postdoctoral research at UC Berkeley

April 2015 Recipient of the Capocelli Prize for the best student-authored paper at the *IEEE Data Compression Conference, 2015*

2013–2014 Recipient of a Facebook Graduate Fellowship

May 2008 Recipient of the Highest Achievement Award in Computer Science for graduating 1st in the 2008 graduating class of over 100 students (UC Berkeley)

May 2007 Inducted into Phi Beta Kappa as a junior

Dec 2006 Inducted into Upsilon Pi Epsilon, an international honor's society for the computing sciences

Recipient of student travel awards for PPOPP 2013, SPAA 2013, SPAA 2014, SODA 2015, ICDE 2015, SPAA 2015, and KDD 2015

Refereed Publications

For certain papers, the authors are listed alphabetically, following the convention in mathematics and theoretical computer science, and others are listed by contribution.

- [1] Laxman Dhulipala, Guy Blelloch and **Julian Shun**. Julienne: A Framework for Parallel Graph Algorithms using Work-efficient Bucketing. *To appear in Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, 2017.
- [2] Julian Labeit, **Julian Shun** and Guy Blelloch. Parallel lightweight wavelet tree, suffix array and FM-index construction. *Journal of Discrete Algorithms*, 2017. **Special issue of DCC 2016**.
- [3] **Julian Shun**. Improved Parallel Construction of Wavelet Trees and Rank/Select Structures. *Proceedings of the IEEE Data Compression Conference (DCC)*, pp. 92–101, 2017.
- [4] **Julian Shun**, Farbod Roosta-Khorasani, Kimon Fountoulakis and Michael Mahoney. Parallel Local Graph Clustering. *Proceedings of the VLDB Endowment*, 9(12), pp. 1041–1052, 2016.
- [5] (*alphabetical order*) Guy Blelloch, Jeremy Fineman, Phillip Gibbons, Yan Gu and **Julian Shun**. Efficient Algorithms with Asymmetric Read and Write Costs. *Proceedings of the European Symposium on Algorithms (ESA)*, pp. 14:1–14:18, 2016.
- [6] (*alphabetical order*) Guy Blelloch, Yan Gu, **Julian Shun** and Yihan Sun. Parallelism in Randomized Incremental Algorithms. *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pp. 467–478, 2016.
- [7] (*alphabetical order*) Naama Ben David, Guy Blelloch, Jeremy Fineman, Phillip Gibbons, Yan Gu, Charles McGuffey and **Julian Shun**. Parallel Algorithms for Asymmetric Read-Write Costs. *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pp. 145–156, 2016.
- [8] Julian Labeit, **Julian Shun** and Guy Blelloch. Parallel Lightweight Wavelet Tree, Suffix Array and FM-Index Construction. *Proceedings of the IEEE Data Compression Conference (DCC)*, pp. 33–42, 2016.
- [9] (*alphabetical order*) Niklas Baumstark, Guy Blelloch and **Julian Shun**. Efficient Implementation of a Synchronous Parallel Push-Relabel Algorithm. *Proceedings of the European Symposium on Algorithms (ESA)*, pp. 106–117, 2015.
- [10] **Julian Shun**. An Evaluation of Parallel Eccentricity Estimation Algorithms on Undirected Real-World Graphs. *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD)*, pp. 1095–1104, 2015.
- [11] (*alphabetical order*) Guy Blelloch, Jeremy Fineman, Phillip Gibbons, Yan Gu and **Julian Shun**. Sorting with Asymmetric Read and Write Costs. *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pp. 1–12, 2015.
- [12] Yan Gu, **Julian Shun**, Yihan Sun and Guy Blelloch. A Top-Down Parallel Semisort. *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pp. 24–34, 2015.
- [13] **Julian Shun** and Kanat Tangwongsan. Multicore Triangle Computations Without Tuning. *Proceedings of the IEEE International Conference on Data Engineering (ICDE)*, pp. 149–160, 2015.
- [14] **Julian Shun**, Laxman Dhulipala and Guy Blelloch. Smaller and Faster: Parallel Processing of Compressed Graphs with Ligma+. *Proceedings of the IEEE Data Compression Conference (DCC)*, pp. 403–412, 2015.

- [15] **Julian Shun**. Parallel Wavelet Tree Construction. *Proceedings of the IEEE Data Compression Conference (DCC)*, pp. 63–72, 2015. **Awarded the Capocelli Prize for Best Student-Authored Paper**
- [16] **Julian Shun**, Yan Gu, Guy Blelloch, Jeremy Fineman and Phillip Gibbons. Sequential Random Permutation, List Contraction and Tree Contraction are Highly Parallel. *Proceedings of the ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pp. 431–448, 2015.
- [17] **Julian Shun**. Fast Parallel Computation of Longest Common Prefixes. *Proceedings of the ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*, pp. 387–398, 2014.
- [18] **Julian Shun** and Guy Blelloch. Phase-concurrent Hash Tables for Determinism. *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pp. 96–107, 2014.
- [19] **Julian Shun**, Laxman Dhulipala and Guy Blelloch. A Simple and Practical Linear-Work Parallel Algorithm for Connectivity. *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pp. 143–153, 2014.
- [20] Aapo Kyrola, **Julian Shun** and Guy Blelloch. Beyond Synchronous: New Techniques for External Memory Graph Algorithms. *Proceedings of the Symposium on Experimental Algorithms (SEA)*, pp. 123–137, 2014.
- [21] **Julian Shun** and Guy Blelloch. A Simple Parallel Cartesian Tree Algorithm and its Application to Parallel Suffix Tree Construction, *ACM Transactions on Parallel Computing (TOPC)*, Vol. 1 Issue 1, Article No. 8, 2014. (Earlier version appears in ALENEX 2011.)
- [22] **Julian Shun**, Guy Blelloch, Jeremy Fineman and Phillip Gibbons. Reducing Contention Through Priority Updates. *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pp. 152–163, 2013.
- [23] **Julian Shun** and Fuyao Zhao (joint first author). Practical Parallel Lempel-Ziv Factorization. *Proceedings of the IEEE Data Compression Conference (DCC)*, pp. 123–132, 2013.
- [24] **Julian Shun** and Guy Blelloch. Ligra: A Lightweight Graph Processing Framework for Shared Memory. *Proceedings of the ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP)*, pp. 135–146, 2013.
- [25] (*alphabetical order*) Guy Blelloch, Jeremy Fineman and **Julian Shun**. Greedy Sequential Maximal Independent Set and Matching are Parallel on Average, *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pp. 308–317, 2012.
- [26] **Julian Shun**, Guy Blelloch, Jeremy Fineman, Phillip Gibbons, Aapo Kyrola, Harsha Vardhan Simhadri and Kanat Tangwongsan. Brief Announcement: The Problem Based Benchmark Suite, *Proceedings of the ACM Symposium on Parallelism in Algorithms and Architectures (SPAA)*, pp. 68–70, 2012.
- [27] (*alphabetical order*) Guy Blelloch, Jeremy Fineman, Phillip Gibbons and **Julian Shun**. Internally Deterministic Parallel Algorithms Can Be Fast, *Proceedings of the ACM Symposium on Principles and Practice of Parallel Programming (PPoPP)*, pp. 181–192, 2012.

- [28] (*alphabetical order*) Guy Blelloch and **Julian Shun**. A Simple Parallel Cartesian Tree Algorithm and its Application to Suffix Tree Construction, *Proceedings of the SIAM Meeting on Algorithm Engineering and Experiments (ALENEX)*, pp. 48–58, 2011.
- [29] (*alphabetical order*) David Aldous and **Julian Shun**. Connected Spatial Networks over Random Points and a Route-Length Statistic, *Statistical Science*, Vol. 25, No. 3, pp. 275–288, 2010.

Talks

Shared-Memory Parallelism Can Be Simple, Fast, and Scalable

- MIT EECS Special Seminar, April 2017
- Yale CS Seminar, March 2017
- UC Davis CS/ECE Seminar, March 2017
- University of Chicago CS Seminar, March 2017
- University of Illinois Urbana-Champaign CS Seminar, February 2017
- University of Maryland College Park CS Seminar, February 2017
- Caltech Frontiers in CMS Symposium, January 2017
- Georgia Tech CSE Seminar, September 2016
- MIT Theory of Computation Seminar, November 2015
- SUNY Stony Brook Seminar, May 2015
- Carnegie Mellon University Ph.D. Thesis Defense, April 2015
- Northwestern University Seminar, April 2015
- Indiana University Bloomington Seminar, March 2015

Improved Parallel Construction of Wavelet Trees and Rank/Select Structures

- Data Compression Conference (DCC), April 2017

Large-Scale Graph Processing in Shared Memory

- Guest lecture in the Advanced Performance Engineering for Multicore Applications (6.S898) course at MIT, February 2017
- Tutorial at the Symposium on Principles and Practice of Parallel Programming (PPoPP), March 2016

Parallel Local Graph Clustering

- MIT Seminar, November 2016
- CMU Systems Design and Implementation (SDI) Seminar, September 2016
- International Conference on Very Large Data Bases (VLDB), September 2016
- UC Berkeley Database Seminar, June 2016
- International Computer Science Institute Lunch Seminar, June 2016
- UC Berkeley AMPLab Retreat, June 2016

Graph Optimization

- Guest lecture in the Performance Engineering of Software Systems course (6.172) at MIT, November 2016

Ligra: A Lightweight Graph Processing Framework for Shared Memory

- ACM San Francisco Bay Area Chapter, July 2016
- Workshop on Algorithms for Modern Massive Data Sets (MMDS), June 2016
- Stanford Seminar, June 2016
- Stanford Software Seminar, January 2016
- UC Berkeley AMPLab Seminar, December 2015
- **Keynote talk** at the High Performance Graph Mining (HPGM) Workshop, August 2015
- UC San Diego Seminar, January 2015
- Georgia Tech CSE Seminar, October 2014
- UCLA Seminar, October 2014
- University of Washington Seminar, October 2014
- CMU Systems Design and Implementation (SDI) Seminar, October 2014
- Intel Labs, Hillsboro, September 2014
- Intel Labs, Santa Clara, January 2014
- UC Berkeley ASPIRE Seminar, October 2013
- Facebook, October 2013
- Symposium on Principles and Practice of Parallel Programming (PPoPP), February 2013

Parallelism in Randomized Incremental Algorithms

- Symposium on Parallelism in Algorithms and Architectures (SPAA), July 2016

Models and Algorithms with Asymmetric Read and Write Costs

- UC Berkeley Benchmarking and Optimization (BeBOP) Seminar, October 2015

A Simple Parallel Cartesian Tree Algorithm and its Application to Parallel Suffix Tree Construction

- UC Berkeley Cloud Computing and Networking Seminar, September 2015
- Microsoft Research, Beijing, July 2011
- CMU Theory Lunch, February 2011
- Meeting on Algorithm Engineering and Experiments (ALENEX), January 2011

An Evaluation of Parallel Eccentricity Estimation Algorithms on Undirected Real-World Graphs

- Conference on Knowledge Discovery and Data Mining (KDD), August 2015
- UC Berkeley AMPLab Seminar, August 2015

Multicore Triangle Computations Without Tuning

- International Conference on Data Engineering (ICDE), April 2015

Smaller and Faster: Parallel Processing of Compressed Graphs with Ligra+

- Data Compression Conference (DCC), April 2015

Parallel Wavelet Tree Construction

- Data Compression Conference (DCC), April 2015

Sequential Random Permutation, List Contraction and Tree Contraction are Highly Parallel

- CMU Theory Lunch, January 2015
- Symposium on Discrete Algorithms (SODA), January 2015

Large-Scale Parallel Graph Algorithms

- University of Maryland (College Park) Seminar, December 2014

Fast Parallel Computation of Longest Common Prefixes

- International Conference for High Performance Computing, Networking, Storage and Analysis (SC), November 2014

Beyond Synchronous Computation: New Techniques for External Memory Graph Algorithms

- Symposium on Experimental Algorithms (SEA), June 2014

Phase-concurrent Hash Tables for Determinism

- Symposium on Parallelism in Algorithms and Architectures (SPAA), June 2014

Sequential and Parallel Hash Tables

- Guest lecture in the Parallel and Sequential Data Structures and Algorithms course (15-210) at CMU, November 2013

Greedy Sequential Maximal Independent Set and Matching are Parallel on Average

- UC Berkeley Theory Lunch, March 2016
- Nanjing University Seminar, January 2014
- MIT EECS Seminar, July 2013
- Symposium on Parallelism in Algorithms and Architectures (SPAA), June 2012
- CMU Theory Lunch, February 2012

Reducing Contention Through Priority Updates

- Symposium on Parallelism in Algorithms and Architectures (SPAA), July 2013

Practical Parallel Lempel-Ziv Factorization

- Data Compression Conference (DCC), March 2013

Brief Announcement: The Problem Based Benchmark Suite

- Symposium on Parallelism in Algorithms and Architectures (SPAA), June 2012

Internally Deterministic Parallel Algorithms Can Be Fast

- CMU SCS Student Seminar Series, March 2012
- Symposium on Principles and Practice of Parallel Programming (PPoPP), February 2012

Grants

2013 **Parallelism Without Concurrency, NSF Grant.**

- Helped write a joint grant among Carnegie Mellon University, Massachusetts Institute of Technology and Georgetown University about linguistics, systems and algorithms for deterministic parallel programming
- Grant has been funded for 4 years (expected \$2.4 million)

Professional Service

Program Committee Member.

- ACM SIGMETRICS, 2018
- ACM SIGMOD, 2018
- Conference on Neural Information Processing Systems (NIPS), 2017 (Review committee)
- IEEE International Conference on High Performance Computing (HiPC), 2017
- High Performance Graph Data Mining and Machine Learning Workshop (HPGDML), 2017
- ACM Symposium on Parallelism in Algorithms and Architectures (SPAA), 2016
- ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI), 2016 (External Review Committee)
- High Performance Graph Processing Workshop (HPGP), 2016
- IEEE International Conference on Cloud and Big Data Computing (CBDCOM), 2016
- IEEE International Conference on High Performance Computing (HiPC), 2016
- High Performance Graph Data Management and Processing Workshop (HPGDMP), 2016 (Program committee co-chair)
- Conference on Neural Information Processing Systems (NIPS), 2016 (Review committee)

Journal Reviewer.

- ACM Transactions on Parallel Computing (TOPC), 2017
- ACM Journal on Experimental Algorithmics (JEA), 2016
- ACM Transactions on Parallel Computing (TOPC), 2016
- ACM Transactions on Parallel Computing (TOPC), 2015
- IEEE Transactions on Parallel and Distributed Systems (TPDS), 2015
- ACM Transactions on Parallel Computing (TOPC), 2014
- Journal of Parallel and Distributed Computing (JPDC), 2014
- ACM Transactions on Algorithms (TALG), 2014
- ACM Transactions on Architecture and Code Optimization (TACO), 2014

Conference Reviewer.

- IEEE Data Compression Conference (DCC), 2017
- IEEE International Parallel & Distributed Processing Symposium (IPDPS), 2017
- ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming (PPoPP), 2017
- USENIX Symposium on Operating Systems Design and Implementation (OSDI), 2016
- International Conference on Parallel Processing (ICPP), 2016
- IEEE International Conference on Data Engineering (ICDE), 2016
- ACM-SIAM Symposium on Discrete Algorithms (SODA), 2016
- ACM Symposium on Parallelism in Algorithms and Architectures (SPAA), 2015
- SIAM Meeting on Algorithm Engineering and Experiments (ALENEX), 2015
- EuroPar, 2013
- IEEE Data Compression Conference (DCC), 2013

Student Volunteer.

- ACM Symposium on Parallelism in Algorithms and Architectures (SPAA), 2012

Teaching Experience

Feb 2017 **Guest Lecturer**, MIT, Cambridge, MA.

Presented a guest lecture on large-scale shared memory graph processing using Ligra in the Advanced Performance Engineering for Multicore Applications (6.S898) course at MIT

Nov 2016 **Guest Lecturer**, MIT, Cambridge, MA.

Presented a guest lecture on optimizing graph algorithms in the Performance Engineering for Software Systems (6.172) course at MIT

March 2016 **Tutorial Presenter**, PPoPP 2016, Barcelona, Spain.

Presented a 3-hour tutorial on Large-Scale Graph Processing in Shared Memory at the Symposium on Principles and Practice of Parallel Programming (PPoPP), 2016

- Jan 2013 - **Teaching Assistant**, *Carnegie Mellon University*, Pittsburgh, PA.
 Dec 2013 Parallel and Sequential Data Structures and Algorithms (15-210; undergraduate-level course)
 - Led weekly recitations and office hours
 - Designed homework assignments, exam problems and recitation notes
 - Gave a guest lecture on sequential and parallel hash tables
 Sept 2012 **Teaching Assistant**, *Carnegie Mellon University*, Pittsburgh, PA.
 CMU SCS: Graph Analytics Workshop
 - Designed written and programming exercises for the workshop
 - Led an exercise session at the workshop, helping participants complete the exercises
 Jan 2012 - **Teaching Assistant**, *Carnegie Mellon University*, Pittsburgh, PA.
 May 2012 Introduction to Computer Systems (15-213; undergraduate-level course)
 - Led weekly recitations and office hours
 - Helped students with assignments during office hours and on the course discussion forum
 - Coordinated with the instructors to write, administer and grade exams
 Jan 2007 - **Teaching Assistant**, *University of California, Berkeley*, Berkeley, CA.
 May 2007 Structure and Interpretation of Computer Programs (CS 61A; undergraduate-level course)
 - Led weekly lab and discussion sessions for an introductory programming course
 - Helped students with homework assignments during office hours and via e-mail
 - Coordinated with the instructor to write, administer and grade exams

Work/Research Experience

- Sept 2009 - **Graduate Research Assistant (advised by Guy Blelloch)**, *Carnegie Mellon University*.
 July 2015
 - Conducted research in large-scale parallel algorithms and frameworks for shared-memory, and techniques for simplifying parallel programming
 - Research topics include: large-scale graph processing frameworks and analytics, deterministic parallelism and programming techniques, parallel string/text processing algorithms, theory of parallel algorithms, benchmarking parallel programs, memory contention in parallel programs, concurrent data structures, data compression algorithms, and external-memory graph algorithms
 July 2008 - **Software Developer**, *Oracle Corporation*, Redwood Shores, CA.
 Jan 2009
 - Member of the Research & Development team in Oracle's Service Engineering division
 - Responsible for researching methods and writing programs for automated and probabilistic diagnosis and prediction of problems in Oracle's On Demand platform
 - Participated in the creation of a Bayesian network for diagnosing database performance problems in On Demand
 - Wrote code for back-end functionality of the diagnosis tool
 - Wrote code for the tool's user interface
 Jan 2008 - **Undergraduate Research Assistant**, *University of California, Berkeley*.
 May 2008
 - Worked with David Aldous in UC Berkeley's Statistics department on a project relating to finding optimal spatial networks
 - Wrote about 2000 lines of C++ code to generate a variety of networks, and compute relevant statistics on the networks
 - Worked on heuristics algorithms to generate more efficient networks
 - Work resulted in a publication in *Statistical Science*
 June 2004 - **Research Assistant**, *Windermere Associates*, Orinda, CA.
 Aug 2004
 - Responsible for compiling information from business articles