Helen Xu

1 Cyclotron Rd, B59-4024J
Berkeley, CA 94704
+1 (914) 462 7260
hjxu@lbl.gov
itshelenxu.github.io/
Google Scholar: `https://bit.ly/3vTLj7o`

Research Interests

My research creates theoretically and practically efficient parallel algorithms by using a *locality-first strategy* of first understanding and exploiting locality before introducing parallelism. One example is my work in parallel cache-friendly data structures for **sparse graph and sparse tensor applications**.

Current Position

Mar 2022 - Present **Grace Hopper Postdoctoral Scholar in Computing Sciences**, Lawrence Berkeley National Laboratory.
Host: Dr. Aydın Buluç

Education

Feb 2022 **Ph.D. in Computer Science**, Massachusetts Institute of Technology.
Thesis: The Locality-First Strategy for Developing Efficient Multicore Algorithms
Advisor: Prof. Charles E. Leiserson

May 2016 **B.S. with Honors in Computer Science**, Stony Brook University.
Thesis: Write-optimized Skip Lists
Advisor: Prof. Michael A. Bender
Awards: Ranked first in the 2016 graduating class of over 100 students and received the Undergraduate Teaching Assistant Award (top 1% of TAs in the CS department)

Publications

Published 14 peer-reviewed conference papers (10 full papers and 4 short papers) in top venues for systems and algorithms including SIGMOD, SPAA, IPDPS, and PODS.

Preprints

A Compressed Packed Memory Array with Fast Parallel Batch Updates. Brian Wheatman and **Helen Xu**. Under review at the International Conference on Very Large Data Bases (VLDB), 2023.

Sycamore: Overcoming the Insert-Query Tradeoff for In-Memory Key-Value Stores. **Helen Xu**, Brian Wheatman, Amanda Li, and Prashant Pandey. Under review at the International Conference on Very Large Data Bases (VLDB), 2023.

Code-word Compression Techniques for Parallel Sparse Tensor Algebra. **Helen Xu**, Tao B. Schardl, Michael Pellauer, and Joel S. Emer. Under review at the Data Compression Conference (DCC), 2023.

Conference Publications

ALENEX '23 *Optimizing Search Layouts in Packed Memory Arrays*. Brian Wheatman, Randal Burns, Aydın Buluç, and **Helen Xu**. To appear at the SIAM Symposium on Algorithm Engineering and Experiments (ALENEX), 2023.

ESA '22 *When Are Cache-Oblivious Algorithms Cache Adaptive? A Case Study of Matrix Multiplication and Sorting*. Arghya Bhattacharya, Abiyaz Chowdhury, **Helen Xu**, Rathish Das, Rezaul A. Chowdhury, Rob Johnson, Rishab Nithyanand, Michael A. Bender. In European Symposium on Algorithms, 2022.

ACDA '21 *Multidimensional Included and Excluded Sums*. **Helen Xu**, Sean Fraser, and Charles E. Leiserson. In SIAM Conference on Applied and Computational Discrete Algorithms, 2021.

SIGMOD '21 *Terrace: A Hierarchical Graph Container for Skewed Dynamic Graphs*. Prashant Pandey, Brian Wheatman, **Helen Xu**, and Aydın Buluç. In ACM SIGMOD International Conference on Management of Data, 2021.

ALENEX '21 *A Parallel Packed Memory Array to Store Dynamic Graphs*. Brian Wheatman and **Helen Xu**. In SIAM Symposium on Algorithm Engineering and Experiments, 2021.

APOCS '21 *Beyond Worst-case Analysis of Multicore Caching Strategies*. (in alphabetical order) Shahin Kamali and **Helen Xu**. In SIAM Symposium on Algorithmic Principles of Computer Systems, 2021.

SPAA '20 *Closing the Gap Between Cache-Oblivious and Cache-Adaptive Analysis*. (in alphabetical order) Michael A. Bender, Rezaul A. Chowdhury, Rathish Das, Rob Johnson, William Kuszmaul, Andrea Lincoln, Quanquan C. Liu, Jayson Lynch, and **Helen Xu**. In ACM Symposium on Parallelism in Algorithms and Architectures, 2020.

SPAA '18 *Cache-Adaptive Exploration: Experimental Results and Scan-Hiding for Adaptivity*. (in alphabetical order) Andrea Lincoln, Quanquan C. Liu, Jayson Lynch, and **Helen Xu**. In ACM Symposium on Parallelism in Algorithms and Architectures, 2018.

IPDPS '18 *A Fill Estimation Algorithm for Sparse Matrices and Tensors in Blocked Formats*. Peter Ahrens, **Helen Xu**, and Nicholas Schiefer. In IEEE International Parallel and Distributed Processing Symposium, 2018.

PODS '17 *Write-Optimized Skip Lists*. (in alphabetical order) Michael A. Bender, Martin Farach-Colton, Rob Johnson, Simon Mauras, Tyler Mayer, Cynthia Phillips, and **Helen Xu**. In ACM Symposium on Principles of Database Systems, 2017.

Peer-Reviewed Short Publications

HPEC '20 *Work-Efficient Parallel Algorithms for Accurate Floating-Point Prefix Sums*. Sean Fraser, **Helen Xu**, and Charles E. Leiserson. In IEEE High Performance Extreme Computing Conference, 2020.

SPAA '20 *Brief Announcement: Multicore Paging Algorithms Cannot Be Competitive*. (in alphabetical order) Shahin Kamali and **Helen Xu**. In ACM Symposium on Parallelism in Algorithms and Architectures, 2020.

ICONS '19 *Dynamic Programming with Spiking Neural Computing*. (in alphabetical order) James Aimone, Ojas Parekh, Cynthia Phillips, Ali Pinar, William Severa, and **Helen Xu**. In International Conference on Neuromorphic Systems, 2019.

HPEC '18 *Packed Compressed Sparse Row: A Dynamic Graph Representation*. Brian Wheatman and **Helen Xu**. In IEEE High Performance Extreme Computing Conference, 2018.

Awards

2022 **Grace Hopper Fellowship in Computing Sciences** *for two years of postdoctoral study at Berkeley Lab.*

2022 **Argonne Training Program for Extreme-Scale Computing (ATPESC),** *a fully-funded two-week course on high-performance computing (awarded to about 80 researchers).*

2021 **Finalist for Best Student Presentation** *at ACDA 2021*.

2020 **Rising Stars Workshop** *held virtually at UC Berkeley*.

2020 **Chateubriand Fellowship** *for Spring 2020 at ENS Lyon*.

2016 **National Physical Science Consortium (NPSC) Graduate Fellowship,** *awarded for six years of graduate study*.

2016 **Undergraduate Teaching Assistant Award,** *awarded to around 1% of TAs in the department of Computer Science at Stony Brook University*.

2015 **NSF Research Experience for Undergraduates Supplement,** *awarded for one year of undergraduate research*.

2013-2016 **Christian Mata Scholarship for Excellence in Computer Science,** *awarded to the top ranked computer science student in every year at Stony Brook University*.

Research Experience

Sep 2016 - Feb 2022 **Supertech Group** *Massachusetts Institute of Technology*, Cambridge, MA.
Designed, analyzed, and implemented parallel and cache-friendly algorithms for shared-memory multicores.
Mentor: Prof. Charles E. Leiserson.

Summer 2020 **Architecture Research Group** *NVIDIA Research*, Westford, MA.
Created and evaluated compressed data structures for sparse tensor computations.
Mentor: Dr. Michael Pellauer.

Spring 2020 **Resource Optimization: Models, Algorithms, and Scheduling Group** *ENS Lyon*, Lyon, France.
Studied algorithms for multi-resource job scheduling with theoretical guarantees.
Mentors: Professors Loris Marchal, Frederic Vivien, and Anne Benoit.

Summer 2019 **Systems and Networking Group** *Microsoft Research*, Cambridge, UK.
Developed and implemented algorithms for novel parallel machine learning accelerators.
Mentor: Dr. Nuno Lopes.

Summer 2016 **Discrete Mathematics and Optimization Group** *Sandia National Laboratories*, Albuquerque, NM.
Simplified and refined randomized I/O-efficient data structures.
Mentor: Dr. Cindy Phillips

Summer 2015 **Systems and Security Research Group** *Sandia National Laboratories*, Livermore, CA.
Designed and implemented a network event tracking database using an existing write-optimized fractal tree index.
Mentor: Dr. Tom Kroeger

Teaching

Qualifications

2021 **Kaufman Teaching Certificate Program (KTCP)** *Teaching and Learning Lab*, Massachusetts Institute of Technology, Cambridge, MA.

- Completed a semester-long course consisting of developing teaching skills via weekly workshops.
- Prepared and delivered a teaching demo on algorithms to a general audience and provided peer feedback.

Experience

2017 - 2018 **Teaching Assistant for Performance Engineering of Software Systems** *Department of Electrical Engineering and Computer Science*, Massachusetts Institute of Technology, Cambridge, MA.

- Received an average overall rating of 6.2 out of 7 points.
- Led weekly recitations (TA-led tutorial sessions) and office hours.
- Developed the course cloud infrastructure, including the course development environment and tools suite.
- Coordinated with instructors to write and grade homeworks and exams.
- Helped students with assignments on the course discussion forum.

2014 - 2015 **Teaching Assistant for Discrete Mathematics** *Department of Computer Science*, Stony Brook University, Stony Brook, NY.

Students Advised

Ph.D.

2018-present **Brian Wheatman**, Johns Hopkins University (hosted at Berkeley Lab in Summer 2022).
Compressed Dynamic Data Structures for Efficient Graph Processing.

2020-present **Arghya Bhattacharya**, Stony Brook University.
An Experimental Evaluation of Cache-Adaptive Algorithms.

Masters theses

Feb 2023 (expected) **Amanda Li**, MIT.
Efficient Dynamic Graph Processing via Data Structure Design.

June 2020 **Sean Fraser**, MIT.
Optimizing Parallel Prefix Sums for Scientific Computing.

June 2019 **Brian Wheatman**, MIT (co-advised with Tim Kaler).
Image Alignment and Dynamic Graph Analytics : Two Case Studies of How Managing Data Movement Can Make (Parallel) Code Run Fast.

June 2019 **Stephanie Ren**, MIT (co-advised with Tao B. Schardl).
Vector-Aware Space Cuts in Stencil Computations.

Technical Talks and Presentations

Invited Seminars

Optimizing Dynamic Graph Processing on Multicores with the Locality-First Strategy

- Lawrence Berkeley National Laboratory, 2022.

Data Structure Design for Skewed Dynamic Graphs

- MIT Fast Code Seminar, 2021.
- Williams CS Colloqium, 2021.

Updatable Data Compression Formats for Hierarchical Fiber Abstraction

- NVIDIA Research, 2020.

A Fill-Estimation Algorithm for Sparse Matrices and Tensors in Blocked Formats

- Tel Aviv University, 2019.

Conference Talks

Multidimensional Included and Excluded Sums

- *Finalist for Best Student Presentation* at ACDA, 2021.

Beyond Worst-case Analysis of Multicore Caching Strategies
- APOCS, 2021.

Work-Efficient Parallel Algorithms for Accurate Floating-Point Prefix Sums
- HPEC, 2020.

Multicore Paging Algorithms Cannot Be Competitive
- SPAA, 2020.

Dynamic Time Warping in Strongly Subquadratic Time: Algorithms for the Low-Distance Regime and Approximate Evaluation
- ICALP, 2019 (filling in for Bill Kuszmaul).

Cache-Adaptive Exploration: Experimental Results and Scan-Hiding for Adaptivity
- SPAA, 2018.

Write-Optimized Skip Lists
- PODS, 2017 (Joint talk with Tyler Mayer).
- Guest lecture in MIT's advanced performance engineering course (MIT 6.S898), 2017.

Posters

A Parallel Packed Memory Array for Dynamic Graphs
- Rising Stars Poster Session, 2020.

A Fill-Estimation Algorithm for Sparse Matrices and Tensors in Blocked Formats
- CSAIL Alliances Poster Session, 2018.

Performance Engineering Discussions

Software Performance Engineering at the End of Moore's Law
- Jane Street Symposium, 2020.

Leiserchess Codewalk
- Guest lecture in Performance Engineering of Software Systems (MIT 6.172), 2018.

Communication Workshops

NSF Graduate Research Fellowship Statements Workshop
- MIT EECS Communication Lab, 2020 and 2021.

Research Qualifying Examination Workshop
- MIT EECS Communication Lab, 2019 and 2021.

Professional Service

Program Committees
ALENEX (2022, 2023), SPAA (2022, 2023), SEA (2023).

Paper Reviewer
BigData (2023), JPDC (2022), ICPP (2021), ACDA (2021), WADS (2021), SPAA (2021, 2020), EuroPar (2020), SOSA (2020), SODA (2019, 2018), ESA (2019, 2022).

Availability Reviewer
SIGMOD (2023).

Outreach Activities

2018 - 2022 **Communication Lab Advisor** *with the MIT Electrical Engineering and Computer Science Communications Lab.*

Communication Lab Page: `http://mitcommlab.mit.edu/eecs/`

- Reviewed over 50 papers, talks, statements, and CVs via peer coaching.
- Created and implemented a partnership with Professor Saman Amarasinghe's research group to jointly prepare conference presentations via weekly meetings.
- Prepared and led workshops for the NSF Graduate Research Fellowship application (2020, 2021) and the PhD research qualifying exam presentation (2019, 2021).
- Wrote articles about core technical communications topics, such as thesis proposals and paper abstracts.

2017 **Co-president** *of MIT Graduate Women in Course 6 (GW6).*

Webpage: `http://gw6.scripts.mit.edu/`

- Organized events several times per month for graduate women in the department, including professional development and social activities.

References

Prof. Charles E. Leiserson
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
cel@mit.edu
+1 (617) 642-6521

Prof. Michael A. Bender
Department of Computer Science
Stony Brook University
bender@cs.stonybrook.edu
+1 (631) 987-8035

Prof. Shahin Kamali
Department of Computer Science
York University
kamalis@yorku.ca
+1 (416) 736-2100

Dr. Aydın Buluç
Computational Research Division
Lawrence Berkeley National Laboratories
abuluc@lbl.gov
+1 (510) 325-9659