

# Haewon Jeong | Curriculum Vitae

📞 +1 412 417 4054 • ✉ haewon@seas.harvard.edu

## Research Interest

---

My primary research interests are in using tools from Information Theory for **reliable** and **responsible** machine learning (ML), with the application in education.

- **Reliable ML:** By marrying Coding Theory and distributed computing, developed many state-of-the-art coded computing algorithms that can mitigate unreliabilities in large-scale ML algorithms.
- **Responsible ML for education:** Understanding the impacts of ML algorithms in classrooms, especially on minority students in STEM classes, and developing theoretically-grounded ML tools that can promote fairness and diversity in early education.

## Current Position

---

**Harvard University**  
Postdoctoral Fellow

**Cambridge, MA**  
June 2020 – Now

- Advised by: Prof. Flavio Calmon (Harvard), Prof. Muriel Medard (MIT), Prof. Nilanjana Dasgupta (UMass Amherst)

## Education

---

**Carnegie Mellon University (CMU)**  
Ph.D. in Electrical & Computer Engineering

**Pittsburgh, PA**  
Feb 2020

- Thesis: Fully-Decentralized Coded Computing for Reliable Large-Scale Computing (Advised by Prof. Pulkit Grover)
- Committee: Alex Dimakis (UT Austin), Christian Engelmann (ORNL), Tze Meng Low (CMU), Pulkit Grover (CMU, Chair), Virginia Smith (CMU)

**Korea Advanced Institute of Science & Technology (KAIST)**  
B.Eng. in Electrical Engineering & B.S. in Computer Science

**Daejeon, Korea**  
Dec 2013

- *Cum Laude*, Major GPA: 3.94/4.0

## Selected Publications

---

- **Haewon Jeong**, Ateet Devulapalli, Viveck Cadambe, and Flavio Calmon, “ $\epsilon$ -Approximate Coded Matrix Multiplication is Nearly Twice as Efficient as Exact Multiplication”, To be presented at ISIT 2021
- **Haewon Jeong**, Yaoqing Yang, Vipul Gupta, Christian Engelmann, Tze Meng Low, Viveck Cadambe, Kannan Ramchandran, and Pulkit Grover, “3D Coded SUMMA: Communication-Efficient and Robust Parallel Matrix Multiplication”, European Conference on Parallel Processing 2020 (Euro-Par 2020)
- **Haewon Jeong**, Mohammad Fahim, Farzin Haddadpour, Sanghamitra Dutta, Viveck Cadambe, Pulkit Grover, “On the optimal recovery threshold of coded matrix multiplication”, IEEE Transactions on Information Theory 2019
- **Haewon Jeong** and Pulkit Grover, “Energy-adaptive Error Correcting For Dynamic and Heterogeneous Networks”, Proceedings of the IEEE, vol. 107, April 2019

## Works in Progress

---

- **Haewon Jeong**, Flavio Calmon, Nilanjana Dasgupta, and Muriel Medard, “Risks of machine learning in STEM classrooms on minority students”, to be submitted for publication in summer 2021.
- **Haewon Jeong**, Hao Wang, Flavio Calmon, and Nilanjana Dasgupta, “Amis Predictions from Imputed Data: A Study of Fairness in Missing Values”, to be submitted for publication in summer/fall 2021.

## Honors & Awards

---

- Harvard Data Science Initiative (HDSI) Postdoctoral Research Fund Award, 2021
- ITA Workshop Graduation Day Award, Gold Prize, 2020
- NSDI Community Award, 2014
- Samsung Human-Tech Thesis Award, Silver Prize, 2014
- KAIST Undergraduate Research Program Award, 3rd Prize, 2012
- National Science and Engineering Undergraduate Scholarship, Full tuition, 2009–2012

## Professional Activities

---

- Organizer, the ICML-21 Workshop on Information-Theoretic Methods for Rigorous, Responsible, and Reliable Machine Learning (ITR3@ICML21), 2021
- Program Committee, the 11th Workshop on Fault Tolerance for HPC at eXtreme Scale (FTXS)

## Teaching / Mentoring Experience

---

### Teaching

- Teaching Assistant, '18-461/661: Intro to Machine Learning for Engineers' (CMU), Fall 2018, Spring 2019
- Teaching Assistant, '18-758: Wireless Communications' (CMU), Fall 2017
- Teaching Assistant, 'CS101: Introduction to Programming' (KAIST), Fall 2011, Spring 2012
- Winter Camp Instructor (volunteer), 'Science Share program' – Designed and instructed a week-long science camp in Jeju island for underprivileged elementary school students, 2011
- Teacher (volunteer), 'EduShare' – taught math and science to underprivileged middle school students at a local community center, 2010

### Mentoring

- Mentoring undergraduate research interns: Yuk Wong and Yuqiu Zhang (2018–2019, published [11]), Quang Minh Nyguen (2019–2020, published [16]), Jessica Edwards (2020), Michael Wu (2021–present)

## Internship Experience

---

### Oak Ridge National Lab

*HPC Resilience Visiting Researcher*

**Oak Ridge, TN**

*June 2019–Aug 2019*

- Hosted by: Dr. Christian Engelmann. Implemented coded matrix multiplication algorithms using MPI and MKL libraries and ran large-scale experiments on HPC systems to evaluate the performance of coded computing strategies.

### NVIDIA

*Resilience Research Intern*

**Santa Clara, CA**

*May 2017–August 2017*

- Conducted research on the design of error detection/correction (EDC) schemes at different communication layers for NVLink, a high-speed interconnect between processors. Devised a mathematical analysis technique to bridge EDC schemes at different layers, and evaluated the theory through simulations using Python and Sage.

### Ium Socius

*Software Developer*

**Seoul, Korea**

*Sep 2012–Dec 2012*

- Comprehensive web development experience at a startup that launched the first online dating service in Korea. Analyzed real-time user activity log of 5 million users and implemented a real-time user behavior detection system to improve the matching algorithm.

## All Publications

---

- [20] **Haewon Jeong**, Viveck Cadambe, Flavio Calmon, and Ateet Devulapalli, " $\epsilon$ -Approximate Coded Matrix Multiplication is Nearly Twice as Efficient as Exact Multiplication", Under minor revision for JSAIT Special Issue on Coded Computing
- [19] **Haewon Jeong**, Viveck Cadambe, Flavio Calmon, and Ateet Devulapalli, " $\epsilon$ -Approximate Coded Matrix Multiplication is Nearly Twice as Efficient as Exact Multiplication", To be presented at ISIT 2021
- [18] **Haewon Jeong**, Yaoqing Yang, Vipul Gupta, Christian Engelmann, Tze Meng Low, Viveck Cadambe, Kannan Ramchandran, and Pulkit Grover, "3D Coded SUMMA: Communication-Efficient and Robust Parallel Matrix Multiplication", European Conference on Parallel Processing 2020 (Euro-Par 2020)
- [17] **Haewon Jeong**, Sanghamitra Dutta, Yaoqing Yang, Viveck Cadambe, Tze Meng Low, and Pulkit Grover, "Addressing Unreliability in Emerging Devices and Non-von Neumann Architectures Using Coded Computing", Proceedings of the IEEE, vol. 108, no. 8
- [16] Quang Minh Nyguen, **Haewon Jeong**, Pulkit Grover, "Coded QR Decomposition", IEEE International Symposium on Information Theory 2020 (ISIT 2020)
- [15] **Haewon Jeong**, Yaoqing Yang, Vipul Gupta, Viveck Cadambe, Kannan Ramchandran, Christian Englemann, Tze Meng Lowe, Pulkit Grover, "Coded 2.5D SUMMA: Coded Matrix Multiplication for High Performance Computing", ICML 2019 CodML Workshop
- [14] **Haewon Jeong**, Mohammad Fahim, Farzin Haddadpour, Sanghamitra Dutta, Viveck Cadambe, Pulkit Grover, "On the optimal recovery threshold of coded matrix multiplication", IEEE Transactions on Information Theory 2019
- [13] **Haewon Jeong**, Yaoqing Yang, and Pulkit Grover, "Systematic Matrix Multiplication Codes", IEEE International Symposium on Information Theory 2019 (ISIT 2019)
- [12] **Haewon Jeong** and Pulkit Grover, "Energy-adaptive Error Correcting For Dynamic and Heterogeneous Networks", Proceedings of the IEEE, vol. 107, April 2019

- [11] Yuk Wong, Yuqiu Zhang, **Haewon Jeong**, and Pulkit Grover, "Robust Molecular Dynamics Simulation on Cloud ", IEEE International Conference on Acoustics, Speech, and Signal Processing 2019 (ICASSP 2019)
- [10] Utsav Sheth, **Haewon Jeong**, Sanghamitra Dutta, Malhar Chaudhari, Yaoqing Yang, Jukka Kohonen, Teemu Roos, and Pulkit grover, "An Application of Storage-Optimal MatDot Codes for Coded Matrix Multiplication: Fast k-Nearest Neighbors Estimation" IEEE International Conference on Big Data 2018
- [9] Sanghamitra Dutta, Ziqian Bai, **Haewon Jeong**, Tze Meng Low, and Pulkit Grover, "Coded Training of Model Parallel Deep Neural Networks under Soft-Errors", IEEE International Symposium on Information Theory 2018 (ISIT 2018)
- [8] **Haewon Jeong**, Fangwei Ye, and Pulkit Grover, "Locally Recoverable Coded Matrix Multiplication", 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton 2018)
- [7] **Haewon Jeong**, Tze Meng Low, and Pulkit Grover, "Coded FFT and Its Communication Overhead", 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton 2018)
- [6] **Haewon Jeong**, Mohammad Fahim, Farzin Haddadpour, Sanghamitra Dutta, Viveck Cadambe, Pulkit Grover, "On the optimal recovery threshold of coded matrix multiplication", 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton 2017)
- [5] **Haewon Jeong**, Chris Blake, and Pulkit Grover, "Energy-Adaptive Polar Codes: Trading off Reliability and Decoding Energy with Adaptive Polar Coding Circuits", IEEE International Symposium on Information Theory 2017 (ISIT 2017)
- [4] **Haewon Jeong** and Pulkit Grover, "Energy-Adaptive Code", 53th Annual Allerton Conference on Communication, Control, and Computing (Allerton 2015)
- [3] Pulkit Grover, Shawn Kelly, Jeff Weldon, Praveen Venkatesh and **Haewon Jeong**, "An information theoretic technique for harnessing attenuation of high spatial frequencies to design ultra-high-density EEG", 53th Annual Allerton Conference on Communication, Control, and Computing (Allerton 2015)
- [2] EunYoung Jeong, Shinae Woo, Muhammad Jamshed, **Haewon Jeong**, Sunghwan Ihm, Dongsu Han, and KyoungSoo Park, "mTCP: a Highly Scalable User-level TCP Stack for Multicore Systems", 11th USENIX conference on Networked systems design and implementation 2014 (NSDI 2014)
- [1] **Haewon Jeong**, Si-Hyeon Lee, Sae-Young Chung, "Flashcast", 18th Asia-Pacific Conference on Communications 2012

## Invited talks / Other talks

---

- SIAM CSE 21, Invited talk, "Coding Theory Meets High-Performance Computing – 3D Coded SUMMA and More", 2021
- ITA Workshop, Invited talk, "Fully-Decentralized Coded Computing For Reliable Large-Scale Computing", 2020
- ICML CodML Workshop, Poster presentation, "Coded 2.5D SUMMA: Coded Matrix Multiplication for High Performance Computing", 2019
- CMU '18-847F: Foundations of Cloud and ML Infrastructure' course, Guest talk, "Coded Fourier Transform", 2018
- Machine Learning in Science & Engineering 2018 (MLSE18), Poster presentation, "Reliable Machine Learning Using Unreliable Components: From Matrix Operations to Neural Networks", 2018
- IEEE North American School of Information Theory (NSAIT16), Poster presentation, "Towards Designing Universal Energy-Efficient Codes", 2016

## Skills

---

- **Computer:** Python (Pandas, Scikit-learn, Jupyter), MATLAB, C, TensorFlow, Java, Sage, Linux, SQL, Gurobi
- **Language:** English, Korean, Spanish (beginner)