# ANKUR MALLICK

⚲ankurmallick.github.io ✉ ankurmallick@microsoft.com

## EDUCATION

**Carnegie Mellon University, Pittsburgh, USA** *Sep 2017 - Aug 2022*
*Ph.D in Electrical and Computer Engineering, GPA 4.0/4.0*
*Thesis:* Probabilistic Methods for Mitigating Uncertainties in Large-Scale Computing and Machine Learning

**Indian Institute of Technology Bombay, Mumbai, India** *Jul 2011 - Jun 2016*
*B.Tech & M.Tech in Electrical Engineering, GPA 9.12/10.00*
*Minor in Computer Science*
*Thesis:* Estimation of Spatial Fields from Samples obtained at Unknown Random Locations

## WORK EXPERIENCE

**Microsoft Corporation, Redmond, USA** *Sep 2022 - Present*
*Role:* Senior Researcher
- Efficient machine learning inference and machine learning for large-scale systems

**Sony Corporation, Atsugi, Japan** *Jul 2016 - Aug 2017*
*Role:* R&D Engineer
- Computational imaging and machine learning for a multispectral image sensor

## AWARDS

- **ACM SIGMETRICS Best Paper Award** *2020*
- **Qualcomm Innovation Fellowship** *2019*
- Carnegie Institute of Technology Dean's Fellowship *2017*
- IIT Bombay Undergraduate Research Award *2016*
- IEEE Signal Processing Society Student Travel Grant *2016*
- Govt. of India INSPIRE Scholarship *2011*

## RESEARCH

**Carnegie Mellon University, Pittsburgh, USA (Doctoral Research)** *Sep 2017 - Aug 2022*
- Designed algorithms using rateless erasure codes for fast distributed computing with straggling (slow) nodes
- Derived theoretical guarantees showing near-ideal asymptotic performance of the proposed algorithm
- Demonstrated up to $3 \times -$ speedup over naive distributed computing approaches on Amazon EC2
- *Received a Best Paper Award at ACM SIGMETRICS 2020 for this work*

**Microsoft Research, Redmond, USA (Internship)** *Jun 2020 - Aug 2020*
- Designed algorithms for data drift mitigation in continuous machine learning for large scale systems
- Obtained upto $8 \times -$ speedup, and upto $20\%$ accuracy gain over baselines in production systems
- *Implemented algorithms in internal products, published a paper and filed a US patent following internal review*

**Lawrence Livermore National Laboratory, Livermore, USA (Internship)** *May 2018 - Aug 2018*
- Designed sample efficient and uncertainty aware algorithms for machine learning on scientific datasets
- *Resulted in a 3 year collaboration, 2 publications, and a mention in Andrew Ng's weekly digest on deep learning*

## PUBLICATIONS

1. **Rateless Codes for Near-Perfect Load Balancing in Distributed Matrix-Vector Multiplication** [Link]
   A.Mallick*, M.Chaudhari, U.Sheth, G.Palanikumar, G.Joshi
   *Communications of the ACM 2022* (**Research Highlight**)

2. **Rateless Sum Recovery Codes for Distributed Non-linear Computations**
   A.Mallick*, G.Joshi
   *IEEE Information Theory Workshop (ITW) 2022*

3. **Matchmaker: Data Drift Mitigation in Machine Learning for Large-Scale Systems** [Link]
   A.Mallick*, K.Hsieh, B.Arzani, G.Joshi
   *Conference on Machine Learning and Systems (MLSys) 2022*

4. **Leveraging Spatial and Temporal Correlations in Sparsified Mean Estimation** [Link]
   D.Jhunjhunwala, A.Mallick*, A.Gadhikar, S.Kadhe, G.Joshi
   *Conference on Neural Information Processing Systems (NeurIPS) 2021*

5. **Rateless Codes for Distributed Non-linear Computations** [Link]
   A.Mallick*, S.Smith, G.Joshi
   *IEEE International Symposium on Topics in Coding (ISTC) 2021*

6. **Deep Kernels with Probabilistic Embeddings for Small-Data Learning** [Link]
   A.Mallick*, C.Dwivedi, B.Kailkhura, G.Joshi, T.Yong-Jin Han
   *Conference on Uncertainty in Artificial Intelligence (UAI) 2021* (**Oral Presentation**)

7. **Rateless Codes for Near-Perfect Load Balancing in Distributed Matrix-Vector Multiplication** [Link]
   A.Mallick*, M.Chaudhari, U.Sheth, G.Palanikumar, G.Joshi
   *ACM SIGMETRICS 2020* (**Best Paper Award**)

8. **Probabilistic Neighbourhood Component Analysis: Sample Efficient Uncertainty Estimation in Deep Learning** [Link]
   A.Mallick*, C.Dwivedi, B.Kailkhura, G.Joshi, T.Yong-Jin Han
   *ICML Workshop on Uncertainty and Robustness in Deep Learning (UDL), 2020*
   ***Featured in Andrew Ng's weekly digest on deep learning research highlights*** [Link]

9. **Rateless Codes for Distributed Computations with Sparse Compressed Matrices** [Link]
   A.Mallick*, G.Joshi
   *IEEE International Symposium on Information Theory (ISIT), 2019*

10. **Fast and Efficient Distributed Matrix-Vector Multiplication Using Rateless Fountain Codes** [Link]
    A.Mallick*, M.Chaudhari, G.Joshi
    *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019*

11. **Bandlimited Field Reconstruction from Samples Obtained on a Discrete Grid with Unknown Random Locations** [Link]
    A.Mallick*, A.Kumar
    *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016*

## SERVICE

- **Editorial Board Member:** Journal of Big Data

- **Program Committee Member:** Conference on Machine Learning and Systems (MLSys) 2023

- **Area Chair:** ICML Workshop on Information-Theoretic Methods for Rigorous, Responsible, and Reliable Machine Learning (ITR3@ICML) 2021

- **Conference Reviewer:**
  1. International Conference on Learning Representations (ICLR) 2021, 2022, 2023
  2. Neural Information Processing Systems (NeruIPS) 2021, 2022
  3. International Conference on Machine Learning (ICML) 2020
  4. International Symposium on Information Theory (ISIT) 2019, 2021
  5. International Conference on Accoustics Speech and Signal Processing (ICASSP) 2019

- **Journal Reviewer:**
  1. Transactions on Machine Learning Research

2. IEEE Journal on Selected Areas in Information Theory

3. IEEE Transactions on Information Theory

4. IEEE Transactions on Signal Processing

5. IEEE/ACM Transactions on Networking

## INVITED TALKS

**Algorithms and Machine Learning for Large-Scale Computing Systems**

- Microsoft Research, Redmond, USA                                                              *Mar 2022*

**Fast and Efficient Distributed Matrix-Vector Multiplication Using Rateless Codes**

- Joint Mathematics Meeting (JMM) 2019, Baltimore, USA                            *Jan 2019*

- Indian Institute of Technology Bombay, Mumbai, India                               *Jan 2019*

## POSITIONS OF RESPONSIBILITY

**Teaching Assistantship**

- 18-661: Introduction to Machine Learning for Engineers                           *Spring 2019*

- 18-847F: Foundations of Cloud and Machine Learning Infrastructure          *Fall 2018*

- EE 342: Control and Communcations                                                            *Spring 2016*

- EE 603: Digital Signal Processing and Applications                                       *Fall 2015*

- MA 207: Partial Differential Equations                                                          *Fall 2014*

**Leadership and Mentoring**

- ECE Department Representative, Graduate Student Assembly (GSA) , CMU          *2019-2021*

- Vice-President, Indian Graduate Student Association (IGSA), CMU                     2019-2021

- Treasurer, Indian Graduate Student Association (IGSA), CMU                            2018-2019

- Institute Student Mentor, IIT Bombay                                                              2014-2016

## TECHNICAL SKILLS

**Programming:**  Python, MATLAB, C++ (Basic), Java (Basic)
**Software & Tools:**  Tensorflow, PyTorch, LaTex