# KONSTANTINOS KONSTANTINIDIS

Ph.D.
Electrical and Computer Engineering
Iowa State University
Ames, IA

## Personal Information

| | |
|---|---|
| Address: | Redwood City, CA 94063 |
| Phone: | 515-357-2016 |
| Email: | k.m.konstantinidis@gmail.com |
| Website: | kkonstantinidis.github.io |
| GitHub: | github.com/kkonstantinidis |
| Google Scholar: | https://scholar.google.com/citations?user=pLi_5zwAAAAJ&hl=en |

## Education

**January 2017 - December 2022** — **Ph.D. in Electrical and Computer Engineering**
Department of Electrical and Computer Engineering,
Iowa State University, Ames, IA.
Dissertation: "*Leveraging coding techniques and redundancy for speeding up distributed computing and robustifying distributed learning.*"
Advisor: Professor Aditya Ramamoorthy.
GPA: 3.94/4.0.

**January 2017 - August 2022** — **M.Eng. in Electrical and Computer Engineering**
Department of Electrical and Computer Engineering,
Iowa State University, Ames, IA.
GPA: 3.94/4.0.

**September 2011 - December 2016** — **Diploma in Electrical and Computer Engineering (5-year program)**
Department of Electrical and Computer Engineering,
Technical University of Crete, Chania, Greece.
Thesis: "*Fast synchronization of OQPSK signals.*"
Advisor: Professor George N. Karystinos.
GPA: 3.4/4.0.

## Research Interests

- Advanced Computing.
- Machine Learning.
- Network Coding.

## Academic Experience

| | |
|---|---|
| January 2017 - May 2022 | **Graduate Research Assistant**<br>Leveraging coding techniques and redundancy for speeding up distributed computing and robustifying distributed learning.<br>Iowa State University. |
| Fall 2020 | **Graduate Teaching Assistant**<br>EE 571: Introduction to Convex Optimization<br>Enrollment: 20+ students<br>Iowa State University. |
| Spring 2019 | **Graduate Teaching Assistant**<br>EE 322: Probabilistic Methods for Electrical Engineers<br>Enrollment: 70+ students<br>Iowa State University. |
| Fall 2018 | **Graduate Teaching Assistant**<br>EE 322: Probabilistic Methods for Electrical Engineers<br>Enrollment: 50+ students<br>Iowa State University.<br>Received the Teaching Excellence Award upon nomination of the instructor. |

## Graduate Research Experience

January 2020 - October 2023

Department of Electrical and Computer Engineering,
Iowa State University.

**Project**: Securing Distributed Learning Against Failures and Attacks

**Description**: Developed novel filtering and detection mechanisms for distributed machine learning scenarios in which computing devices may return erroneous or malicious gradients, which can derail the training. *Coding-theoretic* ideas are paired with *design theory* to minimize the fraction of corrupted computations. *Graph theory* is combined with computational redundancy to optimally assign tasks to devices, and *clique-finding* allows for the detection and exclusion of misbehaving devices from the training. The methods are robust to the most sophisticated attacks and achieve, on average, a 25% increase in top-1 accuracy on the CIFAR-10 dataset over defenses suggested by prior work. They maintain training convergence even when 30% of the devices behave adversarially, and the corresponding reduction of the fraction of corrupted gradients ranges from 16% to 99%.

**Supervisor**: Prof. Aditya Ramamoorthy.

January 2018 - January 2019

Department of Electrical and Computer Engineering,
Iowa State University.

**Project**: Straggler-Resilient and Low-Cost Distributed Matrix Multiplication

**Description**: Developed a technique to multiply high-dimensional matrices in distributed setups that suffer from straggling machines. This operation is of fundamental importance to big data problems where the sheer size of the matrices precludes the computation on a single node. The *stragglers* in the cluster are treated as *erasures* in coding theory, and the multiplication can be completed as soon as a minimum number of machines return their assigned tasks of multiplying coded submatrices; this number is called the *recovery threshold*. We prove that as long as the elements of the matrices are bounded by a maximum value, the recovery threshold can be significantly reduced compared to prior approaches and demonstrate a tradeoff between the assumed absolute value bounds of the matrix entries and the recovery threshold. Our experiments on Amazon Web Services (AWS) highlight the superiority of our method. In particular, we achieve a 37% reduction in the computation time when up to 60% of the machines are stragglers.

**Supervisor**: Prof. Aditya Ramamoorthy.

January 2017 - May 2020

Department of Electrical and Computer Engineering,
Iowa State University.

**Project**: Coding Techniques for Speeding up Distributed Computing

**Description**: Developed algorithms to reduce *MapReduce* communication time across the servers on the *AWS EC2* platform. The protocol uses *Single Parity Check* codes and design theory constructions to assign tasks to servers and splits files less finely than prior work. Tweaked the baseline *TeraSort* algorithm, popular for sorting large datasets (generated and fetched within the *HDFS* system), and adapted it to our scheme. Our method uses *MPI* to facilitate server communication and achieves significant speedups of up to 4.7×. Extended this work to the case when the desired functions can be aggregated (amenable to deep learning applications). Its speedup is 4.3× over the baseline approach. The latter work on *aggregated* MapReduce achieves state-of-the-art communication load but with an exponentially smaller requirement on the minimum number of jobs.

**Supervisor**: Prof. Aditya Ramamoorthy.

## Industry Experience

December 2023 - Present
**Research Scientist at Meta Platforms, Inc. (Facebook)**
I work on the Marketplace Recommendation & Delivery Infrastructure team building highly-scalable data pipelines to compute, store, and retrieve machine learning features & training data, primarily in C++ and Python. Also, I implement system changes, flexible APIs, and monitoring to improve the latency and reliability of offline and online feature serving.

September 2022 - November 2023
**Software Engineer, Platform at C3.ai, Inc.**
As a member of the Platform-Data team, I worked on machine learning infrastructure problems. I researched, built, and optimized the *feature store*, a system that stores machine learning features and serves them to models for predicting phenomena or events. My methods transform and fetch the data from the database with low latency. I analyzed the complexity of my algorithms mathematically and experimentally and identified bottlenecks to speed them up and reduce cloud computing costs on AWS and GCP. Finally, I organized meetings with data scientists and engineers to discuss my findings, analyze efficiency, and investigate optimizations to reduce computation time across multiple machines.

May 2022 - August 2022
**Software Engineer Intern at Meta Platforms, Inc. (Facebook)**
Developed multiple debugging components for machine learning feature authoring used in the data pipelines of Facebook Marketplace. The main component was a framework that categorizes errors during feature compilation, generates alerts, and assigns tasks to the appropriate team; this framework was integrated with the CI/CD. Another end product of my work was an internal UI tool to fetch and display feature values from low-latency storage after a series of transformations.

June 2021 - August 2021
**Software Engineering Intern, Platform at C3.ai, Inc.**
Implemented an end-to-end framework for cluster failure prediction; the framework has two components. The first is the data pipeline which loads cluster health metrics, handles missing data, and creates a training data set. The second component is the ML pipeline which trains a model and makes predictions regarding the cluster's state as soon as new test data becomes available. Followed the process of continuous integration/continuous deployment (CI/CD).

## Publications

Journal papers

- **K. Konstantinidis**, N. Vaswani, and A. Ramamoorthy, "Detection and Mitigation of Byzantine Attacks in Distributed Training," *IEEE/ACM Transactions on Networking (ToN)*, October 2023.
  Source code
- **K. Konstantinidis** and A. Ramamoorthy, "Resolvable Designs for Speeding up Distributed Computing," *IEEE/ACM Transactions on Networking (ToN)*, August 2020.
  Source code
- L. Tang, **K. Konstantinidis** and A. Ramamoorthy, "Erasure Coding for Distributed Matrix Multiplication for Matrices With Bounded Entries," *IEEE Communications Letters (COMML)*, November 2018.
  Source code

Conference papers

- **K. Konstantinidis** and A. Ramamoorthy, "Aspis: Robust Detection for Distributed Learning," *IEEE International Symposium on Information Theory (ISIT)*, June 2022.
  Source code

Presentation video

- **K. Konstantinidis** and A. Ramamoorthy, "ByzShield: An Efficient and Robust System for Distributed Training," *Machine Learning and Systems (MLSys)*, April 2021.
  Source code
  Presentation video
- **K. Konstantinidis** and A. Ramamoorthy, "CAMR: Coded Aggregated MapReduce," *IEEE International Symposium on Information Theory (ISIT)*, July 2019.
- **K. Konstantinidis** and A. Ramamoorthy, "Leveraging Coding Techniques for Speeding up Distributed Computing," *IEEE Global Communications Conference (GLOBECOM)*, December 2018.
  Source code

## Awards

| | |
|---|---|
| May 2022 | **Research Excellence Award**<br>Iowa State University, Ames, IA. |
| June 2019 | **Best Student Poster Award** (link)<br>Midwest Machine Learning Symposium (MMLS), Madison, WI. |
| May 2019 | **Teaching Excellence Award**<br>Iowa State University, Ames, IA. |
| June 2018 | **Graduate Scholarship**<br>Gerondelis Foundation, Lynn, MA. |
| March 2018 | **John Hatsios and Andromache Tsandes Award**<br>Iowa State University, Ames, IA. |

## Seminar

| | |
|---|---|
| July 2020 | **Speeding Up Distributed Computing via Coding** (video)<br>Dependable Data-Driven Discovery (D4) Institute, Ames, IA. |

## Reviewing Service

- *IEEE* Transactions on Communications (TCOM) (2023, 2022, 2020, 2019).
- *IEEE/ACM* Transactions on Networking (ToN) (2021).
- *IEEE* Transactions on Information Theory (TIT) (2023, 2022).
- International Conference on Artificial Intelligence and Statistics (AISTATS) (2023).
- *IEEE* International Symposium on Information Theory (ISIT) (2021, 2020, 2019).
- *IEEE* Access (2024).

## Undergraduate Internship And Research Experience

| | |
|---|---|
| July 2016 - August 2016 | Research Intern at School of Mineral Resources Engineering<br>Technical University of Crete<br>I interned at Geodesy & Geomatics Lab and SenseLab of the university.<br>Project 1: Development of an Android app that stores geodetic measurements on server.<br>Project 2: Representation of a cylindrical geological core in horizontal plane.<br>Supervisors: Professor Panagiotis Partsinevelos, Dimitrios Galanakis, Achilleas Ilias Tripolitsiotis. |
| January 2015 | Carried out research on probabilistic graphical models.<br>School of Electrical and Computer Engineering<br>Technical University of Crete.<br>Project: Implementation of forward/backward inference (Viterbi) as well as learning (Baum Welch) algorithms on the Dishonest Casino problem.<br>Supervisor: Professor Aggelos Bletsas. |

## Selected Graduate Coursework

The following is a partial list of the graduate coursework I have completed at Iowa State University.

- **COMS573: Machine Learning**
  Grade: A.
- **EE525X: Data Analytics in Electrical and Computer Engineering**
  Grade: A.
- **EE523: Random Processes for Communications and Signal Processing**
  Grade: A.
- **EE526X: Deep Learning: Theory and Practice**
  Grade: A-.
- **COMS525: Numerical Analysis of High Performance Computing**
  Grade: A.

## Selected Undergraduate Coursework

The following is a subset of the coursework I have completed at the Technical University of Crete.

- **TEL416: Information Theory and Coding**
  Grade: 10/10.
- **TEL415: Statistical Signal Processing for Telecommunications**
  Grade: 10/10.
- **TEL606: Probabilistic Graphical Models (graduate course)**
  Grade: 8/10.
- **TEL413: Convex Optimization**
  Grade: 8.5/10.
- **TEL414: Modeling and Performance Evaluation of Communication Networks**
  Grade: 10/10.

## Skills

**Programming Languages**
**Proficient**: Python, Java, MATLAB, **Good**: C++, SQL, Bash.

**Interfaces/Frameworks**
**Proficient**: AWS, PyTorch, MPI, MapReduce, **Good**: Hadoop, NumPy, Git, Jenkins.

**Networking**
FTP, SSH, DDNS, VPN, WOL.

**Miscellaneous**
Windows, Linux, Excel, LaTeX.

## Volunteering

September 2021 - May 2022 **Mathematics Tutor for CyMath Kids**
Volunteered in teaching mathematics to 3rd-grade students on a weekly basis. The lessons involved creative problem solving as well as different methods to approach a problem by means of fun challenges. The program is part of the Iowa State University 4U Promise initiative aiming to inspire and motivate students with an increased interest in STEM fields in their early years of education.