

EDUCATION

- **Boston University** Boston, MA
Ph.D. Candidate in System Engineering; GPA: 4.0/4.0 Sep. 2018 – Present
- **Stony Brook University (SBU)** Stony Brook, NY
Ph.D. Candidate in Applied Mathematics; GPA: 3.74/4.0 Aug. 2016 – Aug.2018
- **University of Science and Technology of China (USTC),** Hefei, China
B.S. in Applied Mathematics; GPA: 3.64/4.0 Aug. 2012– June 2016

RESEARCH INTEREST

My interests lie in **stochastic optimization** and **machine learning**. I currently focus on understanding and designing algorithms for machine learning, specifically, stochastic gradient descent and its variants, and adaptive gradient descent methods. I am interested in both theory and application.

PUBLICATIONS

Xiaoyu Li*, Zhenxun Zhuang*, Francesco Orabona. **Exponential Step Sizes in Non-Convex Optimization.** *arXiv preprint, 2020* <https://arxiv.org/abs/2002.05273>

Xiaoyu Li, Francesco Orabona. **On the Convergence of Stochastic Gradient Descent with Adaptive Stepsizes.** In: *The 22nd International Conference on Artificial Intelligence and Statistics, AISTATS. 2019*

TECHNICAL SKILLS

- **Programming Languages:** Matlab, Python, C/C++(MPI) **Deep Learning Package:** PyTorch

WORKING EXPERIENCE

- **Research Intern, Nokia Bell Labs, Murray Hill, NJ**
Fundamentals of Neural Networks Dimensioning *Mentor: Carl Nuzman* *June 2019 - Aug. 2019*
 - Give the theoretical bounds of the dimension of Neural Networks when they are used for classification of Gaussian Mixture Models, on the basic of random construction.
 - Verify the theoretical results with Matlab simulation.

RESEARCH EXPERIENCE

- **Research Assistant, Boston University & Stony Brook University**
(Ongoing) Stochastic Optimization and Machine Learning *Advisor : Francesco Orabona* *Oct 2017 - Present*
 - Provide theoretical guarantees to existing popular-used algorithms which have no theoretical understanding, such as Generalized AdaGrad in the non-convex setting.
 - Propose novel stochastic optimization methods for machine learning, with theoretical support.
 - Implement the algorithms to compare their performances with existing popular algorithms.
- **Research Assistant, Stony Brook University**
Optimization and Design of Supercomputer Network Topologies *June 2017 - Aug. 2017*
 - Design parallelized simulated annealing using C and C++ with MPI to optimize regular graphs.
 - Manage to improve computing performance of cluster using optimized graphs as network topology.
- **Undergraduate Exchange Student Research, National Tsing-Hua University**
Comparisons of Probability Structure of Extended Poisson Distributions with Over-dispersion *July 2015 - Aug. 2015*
 - Compare the probability structure of four extended Poisson Models with character of over-dispersion in terms of flexibility and range of application.

TEACHING EXPERIENCE

- **Grad Teaching Assistant, SBU** Fundamental of Computing, *17Fall*; Elements of Statistics, *16Fall and 17Spring*
- **Undergrad Teaching Assistant, USTC** Single Variable Calculus, *15Fall*

HONORS AND AWARDS

- **Honorable Mentioned** Mathematical Contest in Modeling; 2015
- **Outstanding Student Scholarship** USTC; 2014-2015, 2013-2014, 2012-2013