

# YUE SUN

yuesun9308@gmail.com

## EDUCATION

---

### University of Washington

*Ph.D. in Electrical Engineering*

**Academics:** Overall GPA 3.83/4

**Thesis:** *Nonconvex Optimization and Model Representation with Applications in Control Theory and Machine Learning*

Seattle, WA, US

2016.9 - 2022.3

### Tsinghua University

*Bachelor of Engineering in Electronic Engineering*

**Academics:** Overall GPA 89.2/100

Beijing, China

2012.8 - 2016.7

## WORKING EXPERIENCE

---

### Microsoft, BingAds team, Software Engineer

Project: Advertisement Recommendation

Seattle, WA, US

2022.4 - present

- Preprocessed the data to count embedding coverage and prepare for downstream learning (**SCOPE(SQL)**, **python**).
- Applied BERT for initial embedding of URL/user/Ads. Trained downstream CNN and TwinBERT models for predicting user-Ads and URL-Ads click-through rate (**python/pytorch** on GPU).
- Modeled commercial URL classification using semi-supervised learning, with demos in **python/pytorch**.
- Modeled URL-Ads recommendation (user-blind recommendation) by maximum apriori estimator and online learning, with demos in **python/pytorch**.
- Maintained inference pipeline of embedding generation and user-Ads matching (**SCOPE(SQL)**, **Spark**, **Kafka**).

### Nokia Bell Labs, Math & ALGO Bell Labs Summer Intern

Project: Audio Echo Cancellation with Sparse Coefficients

*Host: Shirin Jalali*

Seattle, WA, US

2021.6 - 2021.8

- Reframed audio echo cancellation problem as online least squares and implemented SGD, APA, IML, Catalyst.
- Combined with sparsity of the parameters, incorporated RIP of circulant matrix into algorithm analysis. Derived and implemented proportionate algorithms and regularized algorithms for sparse coefficient estimation.

### Google, Software Engineering Intern (PhD)

Project: Online Learning for Entropy Coding in Next-generation Video Codec

*First authored paper in DCC 2020, with a patent. Host: Jingning Han*

Mountain view, CA, US

2019.6 - 2019.9

- Implemented online update algorithm for probability estimation for entropy coding.
- Generalized baseline algorithm by maximum apriori estimator with online optimization for adaptive learning.
- Implemented experiments on real dataset and the compressed file is 2% smaller than baseline (**C++ & python**).

## RESEARCH EXPERIENCE

---

Papers published in **NeurIPS**, **CDC**, **ICASSP**, **L4DC**, **ICML workshop**, **OJ-CSYS**, **TSP**

### University of Washington, Research Assistant

Project: Towards Understanding the Role of Representation Dimension in Meta-Learning

*First/second authored paper accepted by NeurIPS 2021, TOPML 2021, ICASSP 2021, Advisor: Maryam Fazel*

Seattle, WA, US

2020.9 - 2022.3

- Demonstrated the procedure and obtained statistical guarantee (sample complexity & error) of meta-learning.
- Obtained guarantee of meta-learning accuracy in the overparametrized regimes, theoretically proved the double descent phenomenon and verified it by numerical experiments.
- Generalized to non-linear models, e.g., multiple class classification, validated on image (CNN/Resnet) and text (LSTM/attention/transformer) classification with MAML as baseline (**python/pytorch**, **AWS**).

Project: Learning Linear Dynamical System via Regularization

*First authored paper (oral) in L4DC 2020, Advisor: Maryam Fazel*

2018.10 - 2022.3

- Obtained optimal error rate of unregularized method for linear system identification.
- Learning with provably fewer data and smaller error by Hankel nuclear norm (low rank) regularization.
- Implemented regularized method on Daisy, Gym and Mujoco dynamical systems and observed the advantage in input sensitivity, simplicity of tuning parameters, training data size and estimation error (**python/pytorch**).

- Proved the global convergence of non-convex policy optimization for a generic family of optimal control problems.
- Proposed the connection between policy optimization in RL and convexification method in control theory.

**Simons Institute for the Theory of Computing, Visiting Graduate Student**

Berkeley, CA, US

Project: Perturbed Gradient Descent Method for Non-convex Optimization

2017.9 - 2017.12

*First authored paper in NeurIPS 2019, ICML workshop 2018*

- Derived convergence rate of noisy gradient descent method for nonsmooth nonconvex optimization problems and extended approximation strategy to Riemannian metric.
- Applied to image processing task where the retraction operator is trained by conditional GAN (**python/tensorflow**).

**Ohio State University, Research Assistant**

Columbus, OH, US

Project: Phase Retrieval and Low Rank Matrix Completion

2015.7 - 2015.9

*First/second authored papers in ICASSP 2016, TSP, Advisor: Yuejie Chi*

- Reproduced the derivation and conducted performance test for various phase retrieval algorithms.
- Derived an extension of Wirtinger Flow algorithms by expanding it from rank 1 case to low rank case. Analyzed the advantages with outliers.

**PUBLICATIONS**

---

1. **Yue Sun**, Samet Oymak and Maryam Fazel, “Finite Sample Identification of Low-order LTI Systems via Nuclear Norm Regularization”, accepted by IEEE Open Journal of Control Systems (OJ-CSYS).
2. Yang Zheng, **Yue Sun**, Maryam Fazel and Na Li, “Escaping High-order Saddles in Policy Optimization for Linear Quadratic Gaussian (LQG) Control”, in 61st Conference on Decision and Control (CDC 2022).
3. **Yue Sun**, Adhyayan Narang, Ibrahim Gulluk, Samet Oymak and Maryam Fazel, “Towards Sample-Efficient Overparameterized Meta-Learning”, in 35th Conference on Neural Information Processing Systems (NeurIPS 2021) & 2021 Workshop on the Theory of Overparameterized Machine Learning (TOPML).
4. **Yue Sun** and Maryam Fazel, “Learning Optimal Controllers by Policy Gradient: Global Optimality via Convex Parameterization”, in 60th Conference on Decision and Control (CDC 2021) & 2021 INFORMS Annual Meeting.
5. Ibrahim Gulluk<sup>†</sup>, **Yue Sun**<sup>†</sup>, Samet Oymak and Maryam Fazel, “Sample Efficient Subspace-Based Representations for Nonlinear Meta-Learning”, in 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).
6. **Yue Sun**, Samet Oymak and Maryam Fazel, “Finite Sample System Identification: Improved Rates and the Role of Regularization”, (**oral**) in 2020 Conference of Learning for Dynamics and Control (L4DC).
7. **Yue Sun**, Jingning Han and Yaowu Xu, “Online Probability Model Estimation for Video Compression”, in 2020 Data Compression Conference (DCC). Patent: Probability Estimation for Entropy Coding (WO 2021/092531 A1).
8. **Yue Sun**, Nicolas Flammarion and Maryam Fazel. “Escaping from Saddle Points on Riemannian Manifolds,” in 33rd Conference on Neural Information Processing Systems (NeurIPS 2019).
9. **Yue Sun** and Maryam Fazel. “Escaping Saddle Points Efficiently in Equality-constrained Optimization Problems”, in ICML Workshop on Modern Trends in Nonconvex Optimization for Machine Learning, 2018.
10. Yuanxin Li, **Yue Sun** and Yuejie Chi. “Low-Rank Positive Semidefinite Matrix Recovery From Corrupted Rank-One Measurements”. IEEE Transactions on Signal Processing 65.2 (2017): 397-408.
11. **Yue Sun**, Yuanxin Li and Yuejie Chi. “Outlier-Robust Recovery of Low-Rank Positive Semidefinite Matrices from Magnitude Measurements”, in 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP).
12. Zhong Hua, Yixin Su, **Yue Sun** and Xin Wang, “Handwritten Numeral Recognition Algorithms and Implementation Based on the Multifunction Touch Screen”, in 2014 Asia-Pacific Congress on Sports Technology and Engineering.

**TECHNICAL REPORTS**

---

Scholar page: <https://scholar.google.com/citations?hl=en&user=IxBPfiYAAAAJ>

1. **Yue Sun**, “Linear System Identification with Redundancy in Output”.
2. **Yue Sun**, “A Short Note on Solving Box Inequality and Linear Equality Constrained Optimization Problem”.
3. **Yue Sun**, “Globally Optimizing the Learning Rate of the Heavy Ball Method”.

**REVIEWING SERVICE**

---

NeurIPS 2019/20, L4DC 2020, ICLR 2021/22, ALT 2021, ISIT 2021, KDD 2021, JMLR, SICON, TIT, TAC