

# YUNZHEN FENG

✉ felix.feng681@gmail.com · ☎ (+86) 150-1082-6920 · Homepage: <https://fengyzpku.github.io/>

## EDUCATION

---

**Peking University (PKU)**, Beijing, China

Aug'17 - June'21 (expected)

- B.S. in Applied Mathematics.
  - Majored in Data Science and Big Data Technology.
  - Ranked: 1/18. Major GPA: 3.76/4.0.
  - Supported by Elite Undergraduate Training Program of Applied Mathematics (top 15%).
  - Advisor: Prof. Bin Dong

## RESEARCH INTERESTS

---

I am interested in understanding the theories behind current data-driven machine learning and, in return, using them to guide the designs of reliable models and algorithms. I like to adopt new perspectives from related areas, such as statistical physics, game theory, optimal transport, and differential equation.

## PUBLICATIONS AND MANUSCRIPTS

---

1. **Yunzhen Feng**, Yue M. Lu. A Precise High-dimensional Analysis of Laplacian Semi-Supervised Learning, *in preparation for ICML 2021*.
2. **Yunzhen Feng\***, Chenkai Yu\*, Adam Wierman, Bin Dong. Pure Nash Equilibria in Adversarial Robustness, *in preparation*.
3. Chizhou Liu, **Yunzhen Feng**, Ranran Wang, Bin Dong. Enhancing Certified Robustness of Smoothed Classifiers via Weighted Model Ensembling. *arXiv:2005.09363, submitted to ICLR 2021*.
4. **Yunzhen Feng\***, Runtian Zhai\*, Di He, Liwei Wang, Bin Dong. Transferred Discrepancy: Quantifying the Difference Between Representations. *arXiv:2007.12446*.

## RESEARCH

---

### **A Precise High-dimensional Analysis of Laplacian Semi-Supervised Learning (SSL)**

with Prof. Yue M. Lu, SEAS, Harvard University

Aug'20 - Present

- Consider a high-dimensional mixture of two Gaussians in the noisy regime and the semi-supervised learning setting where only partial labelings are known.
- Provide a rigorous analysis of the generalization error of Laplacian regularized convex classifiers in the high-dimensional limit when the number of labeled samples, the number of unlabeled samples, and their dimension go to infinity with a fixed ratio.
  - Generalize previous works to data-dependent regularization and general data covariances.
  - Discuss how the ratios of labeled data and unlabeled data affect the generalization error.
  - Observe a transition between unlabeled data's harms and benefits to the classification.

### **Pure Nash Equilibria in Adversarial Robustness**

with Prof. Bin Dong, BICMR, Peking University

July'20 - Present

- From a game-theoretic perspective, prove the existence of pure Nash equilibria for binary classification in the one-dimensional case, without assumptions on the attack or defense.
  - Locate the solution to the inf-sup problem in a compact set to address the continuous function space.
  - Construct a pure Nash equilibrium from each solution to the inf-sup problem, inspired by optimal transport: each classifier in the inf-sup problem is an optimal defense for any attacks.
- Construct a data-dependent preprocessing on inputs from synthetic datasets. Standard training on the processed data can have non-trivial robustness. This observation explains the robust feature notion.
- Working on combining game-theoretic algorithms with adversarial training.

## Enhancing Certified Robustness with Weighted Ensembling

with Prof. Bin Dong, BICMR, Peking University

Jan'20 - June'20

- Proposed a weighted ensembling scheme (SWEEN) to improve certified robustness for randomized smoothing with extensive experiments.
  - Ensembling several small diverse models outperformed an individual model with more parameters.
  - Improved the SOTA with 31% less training time and 36% fewer FLOPs.
  - Applied an adaptive prediction algorithm to save computational cost by at least 45%.
- Provided theoretical guarantees on the scheme:
  - Generality: SWEEN can achieve optimal certified robustness,
  - Optimization: SWEEN can be provably trained to achieve near-optimal risk with a surrogate loss.

## Quantifying the Difference between Representations

with Dr. Di He, Microsoft Research Asia

Jan'20 - June'20

- Proposed Transferred Discrepancy (TD), a new metric to measure the difference between two networks' representation based on their downstream performances.
- Analyzed TD's connection with Canonical Correlation Analysis in linear probing setting.
- Used TD to investigate the effect of various deep learning factors, and discovered that the factors improving the performance always lead to more similar representations under different random initializations.

## OTHER EXPERIENCES

---

### Competition on Kaggle: Long-term Traffic Flow Prediction

Dec'19

- Optimized 3D-convolution-based and WaveNet-based networks for temporal sequential data.
- Ranked 11 / 156 at school, competing against graduate students and computer science students.

### Research Experiences for Undergraduates Program

with Prof. Bin Dong, BICMR, Peking University

Mar'19 - Jan'20

- Proposed a generalization bound exploiting the connection between deep ResNets and differential equations.
- Analyzed how training labels and label noises affected adversarial robustness with Neural Tangent Kernel.

## SPECIALIZED SKILLS

---

- **Programming:** Python (PyTorch), MATLAB,  $\LaTeX$ , C.
- **Language:** English, TOEFL(108 with S26), GRE(159 + 170 + 3.5).

## SELECTED HONORS AND AWARDS

---

<i>Bronze Medal</i> , S.-T.Yau College Student Mathematics Contests	2020
<i>Meritorious Award</i> , Mathematical Contest in Modeling (top 9%)	2019
The elite undergraduate training program of Applied Math (top 15%)	2019
<i>Gold Medal</i> , China Mathematics Olympiad (top 100 in China, same level as USAMO)	2016

## SELECTED COURSES

---

Statistical Learning 3.85    Deep Learning: Advanced Topics 3.81    Introduction to Data Science 3.88  
Mathematical Theory of Neural Networks Models (Princeton, Prof. Weinan E) 3.97  
Topology 3.85    Partial Differential Equations 3.73    Ordinary Differential Equations 3.85  
Randomized algorithms 3.97    Data Structure and Algorithm 3.83    Introduction to Computation 4.0

## ACTIVITIES AND INTERESTS

---

- 3.0 (NTRP standard) Tennis Player.
- Rock-climber and outdoor hiker trained by committees from Chinese Mountaineering Association.
- One-year volunteer at Arthur M. Sackler Museum of Art and Archaeology, Peking University.