

# KAIYUE WEN

📞 13670156595    ✉️ [kaiyuewen3@gmail.com](mailto:kaiyuewen3@gmail.com)    📄 [Personal Site](#)    🌐 [github.com/WhenWen](https://github.com/WhenWen)

## Education

**BS, Institute for Interdisciplinary Information, Tsinghua University**

**Expected 07/2024**

*Overall GPA: 3.95/4.00*

## Awards and Honors

### Competitive Mathematics

1 <sup>st</sup> Prize in National High School Mathematics Olympics Competition	11/2019
1 <sup>st</sup> Prize in National High School Mathematics Olympics Competition	11/2018
Silver Medal in S.-T. Yau College Student Mathematics Contest on Probability and Statistics (rank 3)	05/2021
Bronze Medal in S.-T. Yau College Student Mathematics Contest Team Track	05/2021
Silver Medal in S.-T. Yau College Student Mathematics Contest on Probability and Statistics (rank 3)	09/2022
Silver Medal in S.-T. Yau College Student Mathematics Contest Team Track (rank 2)	09/2022

### Honors

Comprehensive Merit Scholarship of Tsinghua	10/2021
Comprehensive Merit Scholarship of Tsinghua	10/2022
Silver Medal in Yao Award (top scholarship in IIS; 3 students institute-wide)	09/2023
National Scholarship (top 0.2% national-wide)	10/2023

## Publications and Manuscripts

(\* stands for equal contribution.)

[6](**Submission to Annals of Statistics, Major Revision**) Kaiyue Wen\*, Tengyao Wang\*, Yuhao Wang.

“Residual Permutation Test for High-Dimensional Regression Coefficient Testing”

[5](**Neurips 2023**) Kaiyue Wen, Yuchen Li, Bingbin Liu, Andrej Risteski. “(Un) interpretability of Transformers: a case study with Dyck grammars”

[4](**Neurips 2023, Oral**) Kaiyue Wen, Zhiyuan Li, Tengyu Ma. “Sharpness Minimization Algorithms Do Not Only Minimize Sharpness To Achieve Better Generalization”

[3](**ICLR 2023**) Kaiyue Wen, Tengyu Ma, Zhiyuan Li. “How Sharpness-Aware Minimization Minimizes Sharpness?”

[2](**ICLR 2023**) Kaiyue Wen\*, Jiaye Teng\*, Jingzhao Zhang. “Benign Overfitting in Classification: Provably Counter Label Noise with Larger Models”

[1](**EMNLP 2022**) Xiaozhi Wang\*, Kaiyue Wen\*, Zhengyan Zhang, Lei Hou, Zhiyuan Liu, Juanzi Li. “Finding Skill Neurons in Pre-trained Transformer-based Language Models”

## Experience

**Research on the interpretability of Transformers when trained on Dyck Grammar**    **02/2023 – 05/2023**

*Core group member, Supervised by Andrej Risteski, CMU*

- Investigate the loss landscape of 2-layer Transformers when trained on GPT task on bounded depth Dyck Grammar.
- Exhibit a variety of “uninterpretable” attention patterns that can perfectly generate Dyck through theoretical calculation and empirical validations.

**Research on the limiting dynamics of Sharpness Aware Minimization**

**06/2022 – 09/2022**

*Core group member, Supervised by Zhiyuan Li and Tengyu Ma, Stanford*

- Investigate the limiting flow of Sharpness Aware Minimization when the learning rate and perturbation radius converges to 0 under the assumption that global minimizers form a Riemannian manifold
- Provably show different implicit bias for Sharpness Aware Minimization in full-batch and stochastic settings
- Show the necessity of the stop gradient step by proving that without which the method will have a different bias

**Research on understanding soft prompt tuning via neuron activation pattern**

**08/2021 – 04/2022**

*Core group member, Supervised by Xiaozhi Wang and Zhiyuan Liu, Tsinghua NLP Lab*

- Discover highly predictive neurons in pretrained language model which consistently emerges in prompt tuning
- Show the importance of these neurons to the performance under various ways of parameter-efficient tuning

## Skills

---

**Languages:** Familiar with Python and has written C++, R, Matlab, Bash

**Maths:** Familiar with mathematics analysis, measure theory, linear algebra, abstract algebra, probability theory, statistics, causal inference, and discrete mathematics

**Leadership:** Class Monitor of Yao Class from 2021 to 2024, vice president of the IIIS Student Union from 2023 to 2024

**Miscellaneous:** Chinese debating (PB: rank 2 school-wide), science fiction novel writing