# Jiajun Zhu

3200106048@zju.edu.cn · (+86) 173-009-89120 · lanczoschu.github.io

## Education

**Zhejiang University** *Sept. 2020 - Now*

B.S. in Mathematics *Rank: 5%*

- Core Courses: Mathematical Analysis I/II/III (95/93/93), Advanced Algebra II (95), Ordinary Differential Equation (95), Mathematical Software (99)
- Rank: 1/64 in sophomore year

## Publication

1. **Jiajun Zhu**, Peihao Wang, Ruisi Cai, Jason D. Lee, Pan Li, Zhangyang Wang. **Rethinking Addressing in Language Models via Contextualized Equivariant Positional Encoding**, *The Thirteenth International Conference on Learning Representations (ICLR)*, 2025. Submitted.
2. **Jiajun Zhu**, Siqi Miao, Rex Ying, Pan Li. **Towards Understanding Sensitive and Decisive Patterns in Explainable AI: A Case Study of Model Interpretation in Geometric Deep Learning**, *Nature Machine Intelligence*, 2024. Under review. (Manuscript is deposited on arXiv)
3. Peihao Wang, Ruisi Cai, Yuehao Wang, **Jiajun Zhu**, Pragya Srivastava, Zhangyang Wang, Pan Li, **Understanding Bottlenecks of State Space Models through the Lens of Recency and Over-smoothing**, *The Thirteenth International Conference on Learning Representations (ICLR)*, 2025. Submitted.

## Research

**Equivariant Positional Encoding for Long-Context Transformers** *May. 2024 - Sept. 2024*

Research Intern Supervised by Prof. Zhangyang Wang *University of Texas at Austin*

- Introduced TAPE, a framework that enhances positional embeddings in transformers by incorporating sequence content across layers and enforcing permutation and orthogonal equivariance for generalization.
- Extended conventional positional embeddings into multi-dimensional representations, enabling seamless integration into pre-trained models and facilitating parameter-efficient fine-tuning.
- Showed empirically TAPE excels in language modeling and downstream tasks like arithmetic reasoning and long-context retrieval, achieving state-of-the-art performance and reducing perplexity for long sequences.

**Interpretability of Geometric Deep Learning for Scientific Tasks** *Dec. 2022 - Apr. 2024*

Research Intern Supervised by Prof. Pan Li *Georgia Institute of Technology*

- Adapted 12 interpretability techniques from graph neural networks to geometric deep learning models, which are widely employed in scientific tasks, and benchmarked their performance.
- Defined two critical concepts in the domain of interpretability: *sensitive patterns* and *decisive patterns*, highlighting their misalignment, an aspect previously overlooked by researchers.
- Established fundamental observations that clarify the different applications of the two categories of interpretability techniques.

**Structural Divergence in Graph Neural Networks Fine-tuning** *Aug. 2022 - Oct. 2022*

Research Intern Supervised by Prof. Yang Yang *Zhejiang University*

- Implemented five baseline methods and analyze their performances across eight molecular datasets.
- Conducted a theoretical analysis of our method from the perspective of Taylor decomposition and
- Finalized theoretical analysis and experimental sections of the paper, and co-authored the accepted submission to AAAI 2024

## Miscellaneous

- Programming languages: Python, C/C++, CUDA.
- Software & Frameworks: LaTeX, Git, PyTorch, PyTorch Geometric, Transformers.