JINGZHE SHI

Phone: (+86) **182-2151-1369**
Homepage: **https://jingzheshi.github.io/**
Email: **sjzworking@outlook.com** | shi-jz21@tsinghua.org.cn

EDUCATION BACKGROUND

Yao Class, IIIS, Tsinghua University Sep. 2021-Jun.2025

- Recommended to Yao Class due to extraordinary performance in physics competition
- Cumulative GPA : 3.86/4.0, Specialized GPA: 3.90/4.0
- Selected AI-related Courses (4.0/4.0): Machine Learning, Natural Language Processing, Advanced Computer Graphics, Design Thinking in the Age of Artificial Intelligence, Computer Architecture, Operating Systems and Distributed Systems, Mathematics for Computer Science, Theory of Computation, The Physics of Information, Research Immersion Training.

HONORS AND AWARDS

Technological Innovation Scholarship in 2023 & 2024, Tsinghua Univ. Nov. 2023 & Nov. 2024

- Scholarship awarded to students with innovative research or technological outcome at Tsinghua University. (<10/89)

Excellent Voluntary & Public Service Scholarship in 2024, Tsinghua Univ. Nov. 2024

- Scholarship awarded to students with excellent voluntary or public service experience at Tsinghua University. (<10/89)

First-Class Freshmen Scholarship in 2021, Tsinghua Univ. Nov. 2021

- Top scholarship for freshmen at Tsinghua University. (<10/89)

Gold Medal winner in International Physics Olympiad Jul. 2021

- Gold Medal winner in **IPhO 2021** (the 51st **International Physics Olympiad**), **ranking 10th globally**.

National Team member of China for International Physics Olympiad Nov. 2020

- One of the **top 5 students nation wide selected as National Team member** in a series of domestic Physics Olympiads in China mainland to attend IPhO 2021. Won a series of prizes, including **the gold medal in CPhO 2020** (Chinese Physics Olympiads).

SELECTED PUBLICATIONS AND PREPRINTS

(* for equal contribution, # for equal correspondence)

PRISM-Physics: Causal DAG-Based Process Evaluation for Physics Reasoning Publicly available since Oct. 2025
Wanjia Zhao, Qinwei Ma*, **Jingzhe Shi***, Shirley Wu, Jiaqi Han, Yijia Xiao, Si-Yuan Chen, Xiao Luo, Ludwig Schmidt, James Zou*

- **In submission to a top conference**, workshop version accepted by **NeurIPS 2025 Math AI workshop**.
- ArXiv: `https://arxiv.org/abs/2510.03185`.

Intrinsic Entropy of Context Length Scaling in LLMs May. 2025
Jingzhe Shi, Qinwei Ma*, Hongyi Liu*, Hang Zhao#, Jeng-Neng Hwang, Lei Li#(Post-doc at UW,UCPH)*

- **In submission to a top conference**.
- ArXiv: `https://www.arxiv.org/abs/2502.01481`.

Physics Supernova: AI Agent Matches Elite Gold Medalists at IPhO 2025 Sep. 2025
Jiahao Qiu, **Jingzhe Shi***, Xinzhe Juan, Zelin Zhao, Jiayi Geng, Shilong Liu, Hongru Wang, Sanfeng Wu, Mengdi Wang*

- Accepted by **NeurIPS 2025 LLM Eval workshop (oral) and NeurIPS 2025 LAW workshop (spotlight)**.
- ArXiv: `https://arxiv.org/abs/2509.01659`.

Scaling Law for Time Series Forecasting May. 2024
Jingzhe Shi, Qinwei Ma*, Huan Ma, Lei Li(Post-doc at UW,UCPH)*

- Accepted by **NeurIPS 2024**.
- Openreview: `https://openreview.net/forum?id=Cr2jEHJB9q`.

CHOPS: CHat with custOmer Profile Systems for Customer Service with LLMs Apr. 2024
Jingzhe Shi, Jialuo Li, Qinwei Ma, Zaiwen Yang, Huan Ma, Lei Li(Post-doc at UW,UCPH)

- Accepted by **COLM 2024**.
- Openreview: `https://openreview.net/forum?id=9Wmdk94oKF`.

Large Trajectory Models are Scalable Motion Predictors and Planners Oct. 2023
*Qiao Sun, Shiduo Zhang, Danjiao Ma, **Jingzhe Shi**, Derun Li, Simian Luo, Yu Wang, Ningyi Xu, Guangzhi Cao, Hang Zhao*

- ArXiv: `https://arxiv.org/abs/2310.19620`.

RESEARCH EXPERIENCE

Physics Supernova: agent workflows to enhance LLMs' physics reasoning ability targeting IPhO Jul.2025 - Sep. 2025

- Co-first authored work accepted to NeurIPS workshop, advised by **Professor Mengdi Wang** and **Professor Sanfeng Wu** at **Princeton University**.
- We use agent and tool paradigm to enhance LLMs' physics reasoning ability, reaching elite gold medalist level at International Physics Olympiad 2025 theory problems.

Prism Physics: process-score and rule-based formula evaluation of LLM Physics reasoning abilities Jun. 2025-Oct. 2025

- Co-first authored work in submission, advised by **Professor James Zou** at **Stanford University**.
- We propose a novel DAG-based process-score evaluation of LLMs, and a rule-based framework for physics formula comparison, together with a carefully curated dataset, to assess the physics reasoning abilities of LLMs.

Context Length Scaling and Bounds for Language Models from Intrinsic Entropy Perspective Sep. 2024 - May. 2025

- Co-first authored work in submission, advised by **Professor Hang Zhao** at **Tsinghua University** and **Doctor Lei Li**, a **Post-doc** at **UW and UCPH**.
- Co-first authored work in submission.
- We explain context length scaling from Intrinsic Space perspective, with theoretical assumptions and deductions validated by experiments on Natural Language and Synthetic Dataset; and analyze the trade-off leading to an optimal context length.

Scaling Law for Time Series Forecasting Jan. 2024 - May. 2024

- In collaboration with Doctor Lei Li, who is currently a Post-Doc at UW and UCPH.
- Co-first authored work accepted by **NeurIPS 2024** (the 38th Annual Conference on Neural Information Processing Systems (main track)).
- Proposing Scaling Law for Time Series Forecasting from both theoretical and experimental perspective. Taking into account the impact of look-back context length.
- Showing both theoretically and empirically that long context length may hurt performance in TSF.
- Personal Contribution: Proposing the idea and be responsible for a rough theoretical framework and all the code and experiments.

Applications of LLMs for Customer Service Sep. 2023 - Feb. 2024

- In collaboration with Doctor Lei Li, who is currently a Post-Doc at UW and UCPH.
- First authored work accepted by **COLM 2024** (the 1st Conference on Language Modeling, acceptance rate: 28.8%).
- Proposing an agent-based architecture for leveraging large and small LLMs in Customer Service, providing effective performance/cost trade-off.
- Proposing a Dataset with Database, APIs, guiding files and QA pairs for Customer Service collected from CPHOS, a real-scenario of online Physics Olympiad.
- Personal Contribution: Responsible for the idea and main workload for collecting dataset, conducting experiments and designing agent architecture.

Scalable Model for Motion Prediction and Planning in Autonomous Driving Aug. 2023 - Oct. 2023

- Advised by Professor Hang Zhao at Institute for Interdisciplinary Information Sciences, Tsinghua University.
- ArXiv link: `https://arxiv.org/abs/2310.19620`.
- Leveraging successful backbones in NLP for trajectory prediction, demonstrating scalability on diverse datasets and achieving state-of-the-art performance on Nuplan dataset
- Personal Contribution: Responsible for the decoder part. Utilize DDPM to generate trajectory in Key Point Space to capture multi-modal distribution of future trajectories.

SOCIAL WORK EXPERIENCE

CPHOS Dec. 2020 - Present
Co-founder, former tech group leader, council member

- CPHOS (`https://cphos.cn`) is an academical nonprofit organization dedicated to providing Physics Olympiad simulations 4-5 times per year for high school contestants for free through an online platform.
- CPHOS was founded in late 2020 by a group of 10 (including myself), now with 100+ members, mainly from top universities in China. 1000+ students from 200+ high schools participate in most Olympiads held by CPHOS.
- I led the tech group to develope tools supporting online Olympiads, including an LLM-based Replyer for Customer Service (**which developed into the CHOPS research project accepted by COLM later**), etc.

Optiver Jul. 2024 - Aug. 2024
Trading Intern, at Optiver Shanghai Office

- Optiver (`https://optiver.com`) is a leading global market making firm.
- I was involved in sim-trading and a research project.

Shanghai Qi Zhi Institute Aug. 2023 - Sep. 2023
Research Intern, Advised by Hang Zhao

- Shanghai Qi Zhi Institute is a Public Institution affiliated to Science and Technology Commission of Shanghai Municipality, led by Andrew Chi-Chih Yao.
- I worked on state transformers for trajectory prediction supported by computational resources at the institute.

SERVICE

Reviewer

- ICLR 2025&2026, COLM 2025, NeurIPS 2025 reviewer.

IPhO 2022 Marker
Invited online marker for the 52nd International Physics Olympiad (IPhO 2022)

- In 2022 IPhO was held in Switzerland and due to pandemic IPhO had to invite extra markers.
- I was invited and fulfilled my job as an online marker to mark, discuss with my marker partner and to do rebuttals with team leaders from countries and regions all around the world through online meetings.

SKILLS AND OTHERS

Languages: Chinese (native); English (TOEFL: 113; R30,L28,S26,W29; test taken in Nov. 2024); Japanese (daily dialogue).

Programming languages: Python, C/C++, etc.

Tools: Git, LaTeX, SQL, etc.