# Xingjian Bai

St John's College, Oxfords, UK, OX1 3JP [Github] [Scholar] xingjianbai0914@gmail.com

Education

Master of Mathematics and Computer Science, University of Oxford

Oct 2023 - Present

#### Bachelor of Arts, University of Oxford

Oct 2020 - July 2023

Mathematics and Computer Science

• First Year: Distinction with Gibbs Prize (the best performance in CS)

• Second & Third Year: First Class

## Research Experience

# Visual Geometry Group (VGG), Oxford

Oct 2023 - Present

Student Researcher

Supervisor: Prof. Christian Rupprecht, Luke Melas-Kyriazi

Topics: Develop diffusion models parametrized by fixed-point dynamic systems, which enable dynamic allocation of computational resources across denoising timesteps.

# Stanford Vision & Learning Lab (SVL)

Jul 2023 - Sep 2023

Undergraduate Visiting Research Intern

Supervisor: Prof. Jiajun Wu

Topics: Enhance the compositionality of diffusion models with neural-symbolic control; distill the understanding of abstract relations from Large Language Models.

## Algorithms and Complexity Theory Group, Oxford

Mar 2023 - Aug 2023

Student Researcher

Supervisor: Prof. Christian Coester

Topics: Innovated sorting algorithms leveraging erroneous predictions from machine learning models; obtained optimal, sub- $O(n \log n)$  comparison complexity with good predictions.

#### Mathematics Institute, Oxford

Jul 2022 - Apr 2023

Summer Research Intern

Supervisor: Prof. Jan Obłój

Topics: Proposed adversarial attack algorithms grounded in distributional robust optimization (DRO) sensitivity analysis; advanced the understanding of robustness of neural networks.

#### AI Safety Research Lab, Oxford

Nov 2022 - Mar 2023

Student Researcher

Mentor: Joar Skalse

Topics: Explored reward hacking due to over-optimization in Reinforcement Learning settings; developed a geometric explanation and an early-stopping algorithm to prevent it in training.

#### **Publications**

Xingjian Bai, Christian Coester "Sorting with Predictions." NeurIPS 2023. [arXiv]

Xingjian Bai, Guangyi He, Yifan Jiang, Jan Obloj "Wasserstein Distributional Robustness of Neural Networks." *NeurIPS* 2023. [arXiv]

Xingjian Bai, Luke Melas-Kyriazi "Fixed Point Diffusion Models." CVPR under review. [PDF]

Jacek Karwowski, Oliver Hayman, **Xingjian Bai**, Klaus Kiendlhofer, Charlie Griffin, Joar Skalse "Goodhart's Law in reinforcement learning." *ICLR 2024.* [arXiv] [Post]

Xingjian Bai, Ruining Ma, Yulong Lou "Containing Invasive Species via Cellular Automaton and AI." *Journal of Undergraduate Mathematics and Its Applications (UMAP)*, 2021.

Hannah Rose Kirk, Yennie Jun, Paulius Rauba, Gal Wachtel, Ruining Li, Xingjian Bai, Noah Broestl, Martin Doff-Sotta, Aleksandar Shtedritski, Yuki M. Asano "Memes in the Wild: Assessing the Generalizability of the Hateful Memes Challenge Dataset." *Proceedings of the 5th Workshop on Online Abuse and Harms (WOAH)*, 2021. [arXiv]

# Awards & Honors

NeurIPS Scholar Award Conference on Neural Information Processing Systems (NeurIPS)	2023
Regional Gold Medalist, going to ICPC World Final International Collegiate Programming Contest (ICPC)	2023
Outstanding Winner & American Maths Society Best Paper (1 / 10053) 37th Mathematical Contest in Modeling	2021
"Hack the Hackers' Hack" award, best out of 66 teams Oxford Hackathon	2020
Full Score USA Computing Olympiad Open	2019
First place among the national team Canadian Computing Olympiad	2018
Silver Medalist Chinese National Olympiad in Informatics	2018
First place in Beijing, 395 / 400 points Chinese National Olympiad in Informatics Provincial - middle school division	2016
Other Experience	
Oxford Student Ambassador  Mathematics Institute & Computer Science department  Participate in outreach events; teach algorithms to students from underdeveloped areas.	Present
Reviewer Workshop on Self-Supervised Learning - Theory and Practice, NeurIPS	2023
Practicals Demonstrator, Compilers, Oxford Computer Science department	2022
Reviewer Workshop on Online Abuse and Harms, NAACL	2022
Skills & Interests	

**Programming Languages**: Proficient in C++, Python; experienced in Julia, Java, Scala, Haskell. **Hobbies**: Marathon (4h 7min), tennis, table tennis, the game of Go (3 Dan).