

Ruichen Xu

ruichen.xu@stonybrook.edu | [Homepage](#)
[LinkedIn](#) | [GitHub](#) | [Google Scholar](#) | [OpenReview](#)

OBJECTIVE

Ph.D. candidate in Applied Mathematics & Statistics (Stony Brook University) seeking an **internship** focused on applying **large language models (LLMs)** and **diffusion models** to real-world applications. Interests include agentic LLM workflows (tool use, structured generation, retrieval), multimodal/diffusion modeling for generation and inverse problems, and rigorous evaluation. I aim to build reliable, reproducible ML systems—from rapid prototyping to scalable training/inference—and collaborate to turn research into usable products and tools.





EXPERIENCE

- **Stony Brook University — Department of Applied Mathematics & Statistics (AMS)** Aug 2022 – Present
Graduate Researcher, Teaching Assistant & Instructor (Advisor: Prof. Yuefan Deng) Stony Brook, NY, USA
 - Research in **neural operators**, **AI for PDEs**, **simulated annealing**; large-scale runs on **Seawulf/NVwulf** (SLURM, CUDA, job arrays, logging).
 - **Teaching Assistant**: AMS 310 (F'22, S'23), AMS 326 (S'24), AMS 510 (F'24; led recitations), AMS 595 (F'24), AMS 528 (S'25), AMS 502 (S'25); ran discussions, grading, office hours.
 - **Instructor**: AMS 394 (W'23, W'24, Su'24, Su'25); AMS 326 (Su'25)—designed/delivered lectures, created/graded assessments, mentored students.
 - Lightning Speaker, **IACS Annual Meeting** (2025, 2024).
- **Courant Institute of Mathematical Sciences, New York University (NYU)** Jan 2021 – May 2022
Recitation Leader New York, NY, USA
 - Courses: Mathematics for Economics II (Sp'21, Sp'22); Probability, Statistics & Decision Making (Fa'21, Sp'22). Led recitations, graded, held office hours.
- **Courant Institute of Mathematical Sciences, New York University (NYU)** Fall 2020, Fall 2021
Grader New York, NY, USA
 - Courses: Analysis (Fa'20); Special Topics (Fa'21).

EDUCATION

- **Stony Brook University (The State University of New York)** Aug. 2022 – May. 2027 (Expected)
Ph.D., Computational Applied Mathematics (GPA 4.00/4.00) Stony Brook, NY, USA
- **Courant Institute of Mathematical Sciences, New York University (NYU)** Sept. 2020 – May. 2022
M.S., Mathematics (GPA 4.00/4.00) New York, NY, USA
- **University of California, Davis** Sept. 2019 – Jun. 2020
M.S., Statistics (GPA 3.94/4.00) Davis, CA, USA
- **Beijing University of Chemical Technology (BUCT)** Sept. 2015 – Jun. 2019
Bachelor, Financial Mathematics (89.33/100); Minor: Commercial Management Beijing, China

PROJECTS

- **ORACLE: LLM-Guided Edits + Simulated Annealing for SBDD (ongoing)** 2025 – Present
Tools: Python, PyTorch; Topics: structure-based design, multi-objective optimization (docking, QED, SA-Score) 
 - Couples contact-aware, LLM-driven molecular edits with SA under a unified score to improve pocket interaction quality and drug-likeness while preserving pocket fidelity.
 - Builds a reproducible pipeline for scoring/selection and ablation across objectives; emphasizes careful control of docking/QED/SA-Score trade-offs.
- **RL-Guided Simulated Annealing (QESA) (ongoing)** 2024 – Present
Tools: Python, PyTorch; HPC: Seawulf & NVwulf; Topic: adaptive temperature control via quasi-equilibrium 
 - Learns a temperature-adjustment policy that maintains near-equilibrium sampling, stabilizing exploration and reducing schedule hand-tuning.
 - End-to-end sweeps and logging on Seawulf/NVwulf (job arrays, checkpoints, metrics), with comparisons to classical SA schedules and move schemes.
- **Physics-Informed Active Learning for Neural Operators (ongoing)** 2025 – Present
Tools: PyTorch; Topics: operator learning, physics-aware acquisition 
 - Designs a physics-aware query strategy over parameters/initial conditions to cut training data while maintaining operator accuracy.
 - Targets robust generalization under limited sensing; integrates uncertainty and PDE residuals into acquisition.
- **PartialObs-PDEBench: Benchmarking PDE Solvers with Partial Observations (ongoing)** 2025 – Present
Tools: Python; Topics: standardized sparse sensing, MAP recovery interface 
 - Standardizes sensing protocols (random/uniform/subdomain/FFT masks) and a MAP-recovery API for fair, reproducible comparison under partial observations.
 - Provides loaders/format and leaderboard hooks to evaluate solvers consistently across PDE families.

PUBLICATIONS

Peer-Reviewed — Conferences

- [c.1] **Discretization-invariance? On the Discretization Mismatch Errors in Neural Operators.**
Wenhan Gao, **Ruichen Xu**, Yuefan Deng, Yi Liu. *International Conference on Learning Representations (ICLR)*, 2025. ([OpenReview](#)) ([ICLR Page](#))
Contribution: formalizes discretization-mismatch errors and introduces CROP for robust cross-resolution operator learning.
- [c.2] **Kolmogorov–Arnold Representation for Symplectic Learning: Advancing Hamiltonian Neural Networks.**
Z. Wu, **R. Xu**[†], L. Chen, G. Kementzidis, S. Wang, Y. Deng. *International Joint Conference on Neural Networks (IJCNN) 2025, accepted*. ([arXiv](#))
Role: [†]co-first author; project lead.
Contribution: Kolmogorov–Arnold–based parameterization for symplectic HNNs preserving energy and long-horizon phase-space structure.

Peer-Reviewed — Journals

- [j.1] **Dynamic Schwartz–Fourier Neural Operator for Enhanced Expressive Power.**
Wenhan Gao, Jian Luo, **Ruichen Xu**, Yi Liu. *Transactions on Machine Learning Research (TMLR)*, 2025. ([OpenReview](#))
Contribution: augments FNOs with dynamic Schwartz operators for adaptive cross-frequency interactions on physics tasks.
- [j.2] **Coordinate Transform Fourier Neural Operators for Symmetries in Physical Modeling.**
Wenhan Gao, **Ruichen Xu**, Haochun Wang, Yi Liu. *TMLR*, 2024. ([OpenReview](#))
Contribution: enforces coordinate-transform consistency in FNOs to respect domain symmetries and improve cross-frame generalization.

Peer-Reviewed Workshops

- [w.1] **APOD: Adaptive PDE-observation diffusion for physics-constrained sampling.**
Ruichen Xu, Haochun Wang, Georgios Kementzidis, Chenhao Si, Yuefan Deng. *ICML 2025 Workshop on Assessing World Models (Poster)*. ([OpenReview](#))
Contribution: couples diffusion sampling with PDE constraints to generate fields consistent with sparse observations and physics.
- [w.2] **RL-QESA: Reinforcement-Learning Quasi-Equilibrium Simulated Annealing.**
Ruichen Xu, Kai Li, Haochun Wang, Georgios Kementzidis, Wei Zhu, Yuefan Deng. *2nd AI for Math Workshop @ ICML 2025 (Poster)*. ([OpenReview](#))
Contribution: learns temperature adjustments that keep SA near quasi-equilibrium for steadier exploration.
- [w.3] **An Iterative Framework for Generative Backmapping of Coarse-Grained Proteins.**
Georgios Kementzidis, Erin Wong, John Nicholson, **Ruichen Xu**, Yuefan Deng. *ICML 2025 GenBio Workshop (Poster)*. ([OpenReview](#))
Contribution: iterative generative pipeline reconstructing atomistic detail from coarse-grained structures.
- [w.4] **SP: Learning Physics from Sparse Observations — Three Pitfalls of PDE-Constrained Diffusion Models.**
Ruichen Xu, Haochun Wang, Georgios Kementzidis, Chenhao Si, Yuefan Deng. *ICML 2025 Workshop on Assessing World Models (Poster)*.
Contribution: identifies noise-phase futility, grid-locking, and scale mismatch; outlines APOD to mitigate them.

Preprints

- [s.1] **Velocity-Inferred Hamiltonian Neural Networks: Learning Energy-Conserving Dynamics from Position-Only Data.**
R. Xu, Z. Wu, L. Chen, G. Kementzidis, S. Wang, H. Wang, Y. Shi, Y. Deng. *arXiv preprint*, 2025. ([arXiv](#))
Contribution: infers latent velocities from position-only trajectories to train energy-preserving HNNs with reduced sensing needs.
- [s.2] **The Impact of Move Schemes on Simulated Annealing Performance.**
R. Xu, H. Wang, Y. Deng. *arXiv preprint*, 2025. ([arXiv](#))
Contribution: systematic study of SA move strategies, identifying regimes where single-coordinate and other proposals converge fastest.
- [s.3] **An Iterative Framework for Generative Backmapping of Coarse-Grained Proteins.**
G. Kementzidis, E. Wong, J. Nicholson, **R. Xu**, Y. Deng. *arXiv preprint*, 2025. ([arXiv](#))
Contribution: iterative generative pipeline reconstructing atomistic detail from coarse-grained structures with improved fidelity.

TALKS & PRESENTATIONS

• Lightning Talk — Diffusion4PDE

IACS Annual Meeting, Institute for Advanced Computational Science (Stony Brook University)

2025



- Topic: diffusion-based sampling under PDE constraints.
- **Lightning Talk — Active Learning for Neural Operators**
 IACS Annual Meeting, Institute for Advanced Computational Science (Stony Brook University)
 - Topic: physics-aware acquisition strategies for neural operators.
- **Student Seminar Speaker — AI for PDEs**
 IACS Student Seminar (Stony Brook University)
 - Overview: neural operators and optimization; current projects.

2024



2025



SKILLS

- **Programming Languages:** Python, C/C++, MATLAB, R
- **Web Technologies:** Basic HTML/CSS for docs & sites
- **Database Systems:** HDF5
- **Data Science & Machine Learning:** PyTorch, PyTorch Lightning, scikit-learn; diffusion models; neural operators (FNO/PINO variants)
- **High-Performance Computing:** Seawulf & NVwulf clusters (SLURM, CUDA, job arrays, logging); Linux servers
- **DevOps & Version Control:** Git/GitHub, Docker, experiment tracking
- **Specialized Area:** Scientific ML for PDEs, Simulated Annealing, Reinforcement Learning (policy-guided optimization)
- **Mathematical & Statistical Tools:** Numerical PDEs, optimization, spectral/FFT methods, uncertainty evaluation
- **Other Tools & Technologies:** LaTeX, Markdown, shell scripting
- **Research Skills:** Reproducible pipelines, dataset preparation, benchmarking, paper writing

REFERENCES

1. **Prof. Yuefan Deng**
 Professor, Department of Applied Mathematics & Statistics
 Stony Brook University
 Email: yuefan.deng@stonybrook.edu
 Relationship: *Ph.D. Advisor*
2. **Prof. Bertal H. Aktas**
 Assistant Professor of Medicine, Department of Medicine
 Brigham and Women's Hospital; Harvard Medical School
 Email: huseyin_aktas@hms.harvard.edu
 Relationship: *Key Collaborator*