Fang Sun

Ph.D. Candidate in Computer Science, University of California, Los Angeles

fts@cs.ucla.edu | web.cs.ucla.edu/~fts | Google Scholar | GitHub | LinkedIn

Research Interests

My research centers on **AI for Quantum Computing and Quantum Chemistry (QCQC)**. I develop **geometric deep learning** methods that tighten the *simulation–design loop* in molecular and materials science across two pillars: (I) accelerating molecular and quantum simulation, including amortized neural initialization for quantum-chemistry circuits and neural surrogates for molecular dynamics, and (II) property-conditioned generative design of molecules and crystalline materials. Keywords: quantum-classical hybrid algorithms for electronic structure, graph neural networks, neural ODEs, normalizing flows and diffusion models, molecular dynamics, and time-series foundation models.

Education

University of California, Los Angeles — Ph.D. in Computer Science 2023 – Present

- Advisor: Prof. Yizhou Sun. Prospectus: *Charting the Chemical Landscape: Geometric Deep Learning for Simulation-Driven Molecular Design.*
- Committee: Yizhou Sun (Chair), Wei Wang, Jason Cong (CS), Alexander A. Balandin (MSE).

University of California, Los Angeles — M.S. in Computer Science 2023 – 2026
Peking University — B.Sc. in Computer Science & Technology, *Summa cum laude* 2019 – 2023
Peking University — B.Sc. in Chemistry (Double-Degree Program) 2019 – 2023

Selected Research

Pillar I — AI for Molecular & Quantum Simulation

ML-LUCJ: Amortized Neural Initialization for Quantum-Chemistry Circuits 2025 – Present

Ongoing — in collaboration with IBM Quantum and UC Berkeley.

- Replace expensive per-instance optimization of variational quantum-chemistry circuits with a single neural-network forward pass, mapping classically computed CCSD amplitudes directly to Local Unitary Cluster Jastrow (L-UCJ) ansatz parameters (orbital rotations **U** and diagonal Coulomb matrices **J**).
- Designed a permutation-equivariant *(2,1)-graph transformer* over orbital-pair tokens (expressivity beyond 1-WL) that ingests the 4-index t_2 tensor as an attention bias and jointly decodes **U** and **J** with symmetry enforced by construction.
- Two-stage training: supervised pre-training on compressed double-factorization teachers (Transition1x, ∼30k reactive geometries) followed by PPO reinforcement-learning fine-tuning with a quantum Monte Carlo energy reward.
- Outputs are hardware-agnostic L-UCJ parameters consumed by sample-based quantum diagonalization (QSCI/SQD); targets near-term IBM heavy-hex processors. Cost per molecule: minutes–hours → milliseconds.

GF-NODE: Graph Fourier Neural ODEs for Molecular Dynamics 2024 – 2025

Published at Transactions on Machine Learning Research (TMLR).

- Unified spectral graph decomposition (graph Fourier transform) with continuous-time neural ODEs to model spatial–temporal multi-scales in MD; state-of-the-art on MD17/rMD17/MD22 and alanine dipeptide, with 3–6 orders-of-magnitude speedup over *ab initio* MD and stable long-horizon rollouts.

Pillar II — Generative Design of Molecules & Materials

TopoConnect: Generative Inverse Design of Topological Semimetal Interconnects 2025 – Present

Ongoing — in collaboration with TSMC's TCAD-US team.

- Built the first large-scale DFT benchmark of 7,263 topological semimetals with transport figures of merit ($\rho_0\lambda$, cohesive energy, and a Fermi-arc surface-conduction metric) for beyond-copper interconnect conductors.
- Adapted a property-conditioned crystal diffusion model (MatterGen) via TSM-focused pre-training and *composable classifier-free guidance*; a Pareto sweep over guidance weights yields novel, DFT-validated crystals with $\rho_0\lambda$ below copper and ruthenium.

GraphVF: Multi-modal Variational Flow for Structure-Based Drug Design 2023 – 2025

Published at the Learning on Graphs Conference (LoG 2025).

- Combined a 2D junction-tree topology encoder with a 3D autoregressive normalizing flow for protein-conditioned molecule generation; 31.1% high-affinity rate on CrossDocked with zero-shot controllability over substructures and physicochemical properties (no retraining).

Publications

(* denotes equal contribution. Author name in **bold**.)

Manuscripts in Preparation

- ML-LUCJ: Amortized Neural Initialization for Quantum-Chemistry Circuits.
  Fang Sun, et al.
  In preparation, 2026 (with IBM Quantum / UC Berkeley).
- TopoConnect: Generative Inverse Design of Topological Semimetal Interconnects.
  Fang Sun, et al.
  In preparation, 2026 (with TSMC).

Pillar I — Molecular & Physical Simulation

- Graph Fourier Neural ODEs: Modeling Spatial-temporal Multi-scales in Molecular Dynamics.
  Fang Sun, Zijie Huang, Haixin Wang, Huacong Tang, Xiao Luo, Wei Wang, Yizhou Sun.
  Transactions on Machine Learning Research (TMLR), 2025.
- DoMiNO: Decomposing Molecular Dynamics with Multi-Scale Neural Graph Ordinary Differential Equations.
  Fang Sun, Zijie Huang, Yadi Cao, Xiao Luo, Wei Wang, Yizhou Sun.
  ACM Transactions on Knowledge Discovery from Data (TKDD), 2026. (Earlier version: Oral, ICLR 2025 MLMP.)
- Symmetry-Preserving Conformer Ensemble Networks for Molecular Representation Learning.
  Yijia Zhu, Yushi Shi, Yuning Chen, **Fang Sun**, Yizhou Sun, Wei Wang.
  NeurIPS, 2025.
- CoRGI: GNNs with Convolutional Residual Global Interactions for Lagrangian Simulation.
  Ethan Ji, Yuanzhou Chen, Arush Ramteke, **Fang Sun**, Tianrun Yu, Jai Parera, Wei Wang, Yizhou Sun.
  KDD, 2026.
- FD-Bench: A Modular and Fair Benchmark for Data-driven Fluid Simulation.
  Haixin Wang, Ruoyan Li, Fred Xu, **Fang Sun**, Kaiqiao Han, Zijie Huang, Guancheng Wan, Ching Chang, Xiao Luo, Wei Wang, Yizhou Sun.
  KDD (DB Track), 2026.

- SpectralFlowNet: Resolution-Invariant Continuous Neural Dynamics for Mesh-Based PDE Modeling.
  Tianrun Yu*, **Fang Sun***, Haixin Wang, Xiao Luo, Yizhou Sun.
  ICLR 2025 MLMP Workshop.
- Analysis of Neural ODE Performance in Long-term PDE Sequence Modeling.
  Fang Sun, Maxwell Dalton, Yizhou Sun.
  ICLR 2025 MLMP Workshop.

Pillar II — Molecular & Materials Generation

- Controllable Generation of Drug-like Molecular Materials with Multi-modal Variational Flow (GraphVF).
  Fang Sun, Zhihao Zhan, Hongyu Guo, Ming Zhang, Jian Tang, Yizhou Sun.
  Learning on Graphs Conference (LoG), 2025. Preprint: arXiv:2304.12825.
- RetroAgent: Agentic Retrosynthesis Planning with Search Over Structured Memory.
  Yanqiao Zhu, Jingru Gan, Xiaoqi Sun, **Fang Sun**, Yidan Shi, Md Mofijul Islam, Chao Shang, Wenhao Gao, Connor W. Coley, Yizhou Sun, Wei Wang.
  Conference on Language Modeling (COLM), 2026.
- Automated Molecular Concept Generation and Labeling with Large Language Models.
  Shichang Zhang, Botao Xia, Zimin Zhang, Qianli Wu, **Fang Sun**, Ziniu Hu, Yizhou Sun.
  COLING, 2025.

Foundations — Graph Neural Networks

- A Comprehensive Survey on Deep Graph Representation Learning.
  Wei Ju, Zheng Fang, . . . , **Fang Sun**, . . . , Xiao Luo, Ming Zhang.
  Neural Networks, 2024. (Authored the "Molecular Generation" section.)
- DisenCTR: Dynamic Graph-based Disentangled Representation for Click-Through Rate Prediction.
  Yifan Wang, Yifang Qin, **Fang Sun**, Bo Zhang, Xuyang Hou, Ke Hu, Jia Cheng, Jun Lei, Ming Zhang.
  SIGIR, 2022.
- DisenPOI: Disentangling Sequential and Geographical Influence for Point-of-Interest Recommendation.
  Yifang Qin*, Yifan Wang*, **Fang Sun***, Wei Ju, Xuyang Hou, Zhe Wang, Jia Cheng, Jun Lei, Ming Zhang.
  WSDM, 2023.
- Over-Smoothing Effect of Graph Convolutional Networks: Spectral and Topological Analysis with Practical Remedies.
  Fang Sun.
  IEEE SoutheastCon, 2026.

Foundations — Time Series

- Position: Beyond Prediction — Toward Verifiable Physiological Waveform Reasoning with Foundation Models and Agentic LLMs.
  Xiaoda Wang, Ching Chang, Defu Cao, Kaiqiao Han, **Fang Sun**, . . . , Yizhou Sun, Wei Wang, Carl Yang.
  ICML, 2026.
- Accelerating Time Series Foundation Models with Speculative Decoding.
  Pranav Subbaraman*, **Fang Sun***, Yuxuan Yao, Huacong Tang, Xiao Luo, Yizhou Sun.
  Preprint, 2025.
- Beyond the Time Domain: Recent Advances on Frequency Transforms in Time Series Analysis.
  Qianru Zhang, Peng Yang, . . . , **Fang Sun**, . . . , Yizhou Sun, Hongzhi Yin.
  Preprint, 2025.

Other Publications

- Deep Knowledge Tracing for Explainable Problem Recommendations on Codeforces.
  Junyang Zhao, **Fang Sun**, Yizhou Sun.
  IJCAI, 2025.

- Why Are We Moral? An LLM-based Agent Simulation Approach to the Study of Moral Evolution.
  Ziheng Zhou, Huacong Tang, Mingjie Bi, Wanying He, **Fang Sun**, . . . , Yipeng Kang, Fangwei Zhong.
  ACL, 2026.
- RevMax: Revenue-Maximizing Recommendation System Competition.
  Fang Sun, Paul Zhang, Pranav Subbaraman, Yizhou Sun.
  EAAI-26, 2025.
- Synthesis and Applications of Borylenes.
  Junhao Li, **Fang Sun**, Honglin Du, Huilong Hong, Kunhao Wang, Jiang Bian.
  University Chemistry, 2019.

Honors & Awards

- **Gold Reviewer Award** (top 25% reviewer), ICML 2026.
- **John Hopcroft Scholarship**, Peking University, 2022.
- **Canon Scholarship** (4/25), Peking University, 2021.
- **Second Award**, PKU-CPC Programming Contest, 2021.
- **National Scholarship** (2/155), Peking University, 2019.

Research & Industry Experience

Academic Mentor, Research in Industrial Projects for Students (RIPS) 2024

- Sponsored by IPAM & Relay Therapeutics. Mentored a team on generative molecular design, structural diversity, and protein–ligand interaction analysis.

Research Intern, Meituan Inc. — Euler-Ads Team 2021 – 2022

- Developed *Euler-Ads*, a distributed GNN framework integrating large-scale graph data into online CTR prediction; deployed via A/B testing on Meituan's advertising platform.

Professional Service

- **Tutorials Chair**, Learning on Graphs Conference (LoG), 2025.
- **Session Chair**, AI and Predictive Modeling, IEEE SoutheastCon, 2026.
- **Reviewer**: ICML, ICLR, NeurIPS, CIKM, LoG.

Technical Skills

- **Deep Learning**: PyTorch, PyTorch Geometric, Hugging Face Transformers, JAX.
- **Quantum Computing & Chemistry**: Qiskit, ffsim, PySCF, Q-Chem, quantum Monte Carlo, sample-based quantum diagonalization (QSCI/SQD).
- **Cheminformatics & Materials**: RDKit, ASE, pymatgen, Synopsys QuantumATK (DFT/transport).
- **Scientific Computing**: Python, C++, NumPy, SciPy, CUDA; HPC/SLURM clusters.
- **Infrastructure**: AWS, GCP, Docker, Git.