

Pengfei Cai

pengfeic@mit.edu | [LinkedIn](#) | [Github](#) | [Google Scholar](#) | caipengfei.me

Differentiable simulations & PDEs | ML for materials discovery | Degradable polymers | Generative models

EDUCATION

Massachusetts Institute of Technology (MIT)

Cambridge, MA

PhD in Computational Materials Science and Engineering

08/2022 – 05/2027

- Computational Science and Engineering coursework: *Deep Learning (A)*, *Fast Methods for Partial Differential and Integral Equations (A)*, *Deep Generative Models (ongoing)*.

National University of Singapore (NUS)

Singapore

B. Eng (Honors) in Materials Science and Engineering

08/2018 – 05/2022

- GPA: **4.83**/5.00 (Highest Distinction)

EXPERIENCE

Department of Materials Science and Engineering, MIT

Graduate Research Assistant, PI: Rafael Gomez-Bombarelli

01/2023 – Present

- Proposed and developed a **generative model**, inspired by functional **flow matching**, as a corrector to refine neural operator predictions, thereby extending rollout accuracies of **long-term spatiotemporal dynamics predictions by 50%**. I will present this work at NeurIPS 2024 workshop on ML for Physical Sciences.
- Developed a **local attention-enhanced Fourier neural operator** to improve the predictions of local instabilities in reaction-diffusion problems with **85% improvement** in errors at the 100th step. Compared method with architectures such as **U-Nets**, **transformers**, and other spectral convolution layers.
- Developed a JAX-based **differentiable finite element solver for end-to-end learning** of unknown physics in reaction-diffusion partial differential equations (PDEs) for frontal polymerization (FP)-based manufacturing.
- Applied **gradient-based PDE-constrained optimizations** and the adjoint method to **learn process conditions and cure kinetics from videos**. Presented preliminary work at ICML workshop on AI for Science.
- Currently pioneering a multiscale method to learn continuum **PDEs from molecular dynamics simulations**, with a focus on Li⁺ diffusion in polymer electrolytes.
- Developing **differentiable pseudospectral solvers** (in JAX / Julia) and **neural ODEs** to learn copolymer kinetics and closure terms from multimodal experimental data (i.e. IR videos, DSC curves); paper in progress.
- Developed a **reverse Monte Carlo method** to model the complex graph network of **degradable crosslinked copolymers**. By matching experimental fragment mass spectra, we can now retrospectively determine network structure and reaction parameters (reactivity ratios, crosslink probabilities) for the first time; paper in progress.
- Applying copolymer property prediction ensemble models for the low-data regime and large language models for molecular embeddings, and currently developing diffusion models for polymer graph generation.

Department of Chemical and Biomolecular Engineering, NUS

Undergraduate Research Assistant, PIs: Bin Liu, Xiaonan Wang

08/2019 – 08/2022

- Developed a **Bayesian optimization-based active learning** framework to accelerate the discovery of photosensitizer (PS) molecules using **graph neural networks** for **molecular property prediction**.
- From a design space of >7 million molecules, suggested >5,300 potential high-performance PSs, with **4 predicted PSs synthesized** by my collaborator with superior or comparable performances to commercial PSs.
- Completed molecular space generation, active learning loops, and **density functional theory (DFT)** calculations for a total of **>14,000 molecules**; workflow code is open-sourced.
- Spearheaded **virtual screening of NIR-II fluorophores** with machine learning, suggested compounds for experimental validation, performed multistep chemical synthesis, and co-led the project towards a first-authored paper.
- Developed and compared neural network architectures (CNN, LSTM, GRU) to predict the compositions of gold nanoclusters from their UV-Vis absorption spectra.

Nanolumi

Product Development Intern, CEO: Jax Lee

07/2020 – 12/2020

- Spearheaded and programmed a graphical user interface for colour gamut and colour spectral analysis (software is used by the company internally).
- Supported the materials startup's anti-counterfeit vertical across multiple projects, from proof of concept to mass production. Analyzed production site data to drive data-centric decision-making.

- **Towards Long Rollout of Neural Operators with Local Attention and Flow Matching-inspired Correction: An Example in Frontal Polymerization PDEs.**
Pengfei Cai, Sulin Liu, Qibang Liu, Philippe Geubelle, Rafael Gomez-Bombarelli. (2024). *Presenting at NeurIPS 2024 Workshop on ML for Physical Sciences*. [Preprint](#).
- **Learning Cure Kinetics of Frontal Polymerization PDEs using Differentiable Simulations.**
Pengfei Cai, Qibang Liu, Philippe Geubelle, Rafael Gomez-Bombarelli. (2024). *ICML 2024 Workshop on AI for Science; NeurIPS Workshop on Data-driven and Differentiable Simulations, Surrogates, and Solvers*. [Preprint](#).
- **Univariate Conditional Variational Autoencoder for Morphogenic Patterns Design in Frontal Polymerization-based Manufacturing.**
Qibang Liu, Pengfei Cai, Diab Abueidda, Seid Koric, Rafael Gomez-Bombarelli, Philippe Geubelle. (2024). *Submitted: Computer Methods in Applied Mechanics and Engineering*. [Preprint](#).
- **A Database of Computed Raman Spectra of Inorganic Compounds with Accurate Hybrid Functionals.**
Yuheng Li, Damien K. J. Lee, Pengfei Cai, Ziyi Zhang, Prashun Gorai, Pieremanuele Canepa. (2024). *Scientific Data*. [Paper link](#).
- **Identification of Chemical Compositions from “Featureless” Optical Absorption Spectra: Machine Learning Predictions and Experimental Validations.**
Tiankai Chen*, Jiali Li*, Pengfei Cai, Qiaofeng Yao, Zekun Ren, Yixin Zhu, Saif Khan, Jianping Xie, Xiaonan Wang. (2023). *Nano Research*. [Paper link](#).
- **Accelerated Design of Near-Infrared-II Molecular Fluorophores via First-Principle Understanding and Machine Learning.**
Shidang Xu*, Pengfei Cai*, Jiali Li, Xianhe Zhang, Xianglong Liu, Xiaonan Wang, Bin Liu. (2022). *ChemRxiv preprint (experimental validation underway)*. [Preprint](#).
- **Machine Learning-Assisted Accurate Prediction of Molecular Optical Properties upon Aggregation.**
Shidang Xu*, Xiaoli Liu*, Pengfei Cai, Jiali Li, Xiaonan Wang, Bin Liu. (2022). *Advanced Science*. [Paper link](#).
- **Self-Improving Photosensitizer Discovery System via Bayesian Search with First-Principle Simulations.**
Shidang Xu*, Jiali Li*, Pengfei Cai, Xiaoli Liu, Bin Liu, Xiaonan Wang. (2021). *Journal of the American Chemical Society*. [Paper link](#).

SKILLS

Programming and Frameworks: Python (PyTorch, JAX, Tensorflow), Julia

Machine Learning: Scientific machine learning, differentiable simulations, generative models (flow matching, DDPM), neural operators, transformers, active learning, cheminformatics, graph neural networks.

AWARDS AND SCHOLARSHIPS

Young NUS Fellow	2023 – 2024
MIT Presidential Fellowship (three students selected in the cohort)	2022 – 2023
Dean’s List (top 5% of cohort)	2021, 2022
A*STAR Undergraduate Scholarship	2018 – 2022

NON-ACADEMIC PROJECTS

- TinyJobs** | *Co-founder* 01/2020 – 08/2022
- Started a platform for employers to post jobs and shortlist candidates. Developed a chatbot to allow >71,000 users across channels to apply for freelance gigs, full-time jobs, and internships in minutes.
 - The bot helped >8,400 job posts and >73,000 applications, and is featured by MyCareersFuture during covid-19.
- sgTutions** | *Co-founder* 07/2018 – 08/2020
- Developed a chatbot to enable students to connect with tutors directly for academic help.
 - Grew the platform to >12,000 users and added features such as shortlisting of tutors and community reports.
- OrcaTech (College Tech Community)** | *President* 01/2019 – 06/2020
- Proposed, launched and furnished the first makerspace in NUS University Town, enabling students in my college to transform their ideas into projects (i.e. drones, vertical farming, chatbots).
 - Conducted free courses, such as 3D printing, Python, and Arduino programming, to residents in my college.
- thenotesrefinery** | *Founder* 06/2017 – 05/2018
- Built a free notes-sharing site for students to contribute educational resources to build an open repository.
 - Ranked the site on the 1st page of Google SG for free notes with search engine optimization.