Xiaochen Du

Cambridge, MA — 626-314-9992 — dux@mit.edu — xiaochendu.github.io

SUMMARY

Fourth-year PhD student focused on leveraging deep learning to accelerate materials simulation and discovery. Track record of publishing in high-profile scientific journals and machine learning workshops. Extensive experience coding in Python, scaling data-driven models, and running large-scale simulations. Skilled in analyzing and manipulating high-dimensional data across diverse scientific disciplines and open-sourcing code repositories. Strong collaborator, experienced in working with interdisciplinary teams.

EDUCATION

Massachusetts Institute of Technology (MIT) Cambridge, MA
PhD Chemical Engineering, Computational Science and Engineering Sep 2021 - June 2026 (expected)
Advisors: Rafael Gómez-Bombarelli, Martin Z. Bazant
Thesis: Accelerating thermodynamic sampling of surface reconstructions at the atomic scale

Duke University Durham, NC
BS Computer Science, AB Chemistry, *summa cum laude* Aug 2017 - Dec 2020

RESEARCH SKILLS

Deep Learning, Computational Chemistry, Generative Modeling, Atomistic Simulations, Density-Functional Theory (DFT), High-Throughput Simulations, Active Learning

EXPERIENCE

MIT, Department of Materials Science and Engineering Cambridge, MA
Graduate Researcher, Rafael Gómez-Bombarelli Group Jan 2022 - present

- Developing deep learning-based methods in PyTorch to accelerate atomistic simulations and materials discovery in fields such as catalysis, electrocatalysis, and renewable energy applications
- Scaling and optimizing graph neural network-based force fields and Transformer-based generative models to study surface reconstructions, enhancing model efficiency and accuracy
- Curating surface reconstruction datasets through high-throughput thermodynamics sampling and large-scale DFT calculations for community use
- Open-sourced new Python packages and expanded existing repositories by more than 10k lines (https://github.com/learningmatter-mit)

Duke University, Department of Computer Science Durham, NC
Researcher and Software Engineer, Alberto Bartesaghi Group Jan 2020 - Apr 2021

- Optimized protein cryo-EM image reconstruction pipeline by integrating denoising algorithms, deep learning models, and statistical inference methods
- Scaled analysis workflows to handle terabytes of noisy image data across 100+ compute nodes
- Led a team of 5 to refactor 30k lines of code and open-source Python codebase (https://github.com/nextpyp)

Duke University, Department of Mechanical Engineering and Materials Science Durham, NC
Undergraduate Researcher and Software Engineer, Volker Blum Group Sep 2017 - Dec 2019

- Developed and launched an open-source materials database (https://github.com/HybriD3-database/MatD3) using Python and Django, providing a user-friendly web interface for materials researchers (https://materials.hybrid3.duke.edu)
- Processed and analyzed high-dimensional data from both experimental and computational sources, including adsorption spectra, X-ray diffraction data, and band structures
- Directed a cross-functional team of 15+ researchers across three institutions and five departments

Emerald Cloud Lab South San Francisco, CA
Scientific Computing Intern May - Aug 2019

- Spearheaded improvements in experimental data and image analysis methods using the Wolfram language and Mathematica, enhancing the automation of scientific wet-lab experiments

SELECTED PUBLICATIONS AND WORKSHOPS

Du, X., Peng, J., Chun, H., Bazant, M.Z. & Gómez-Bombarelli, R. (2024). Accelerating atomic-scale thermodynamic simulations of electrochemical interfaces. *in preparation.*

Du, X., Liu, S. & Gómez-Bombarelli, R. (2024). Scaling autoregressive models for lattice thermodynamics. *under review.*

Du, X., Damewood, J.K., Lunger, J.R., Millan, R., Yildiz, B., Li, L. & Gómez-Bombarelli, R. (2023). Machine-learning-accelerated simulations to enable automatic surface reconstruction. *Nature Computational Science*, 3, 1044. DOI: 10.1038/s43588-023-00571-7
MIT News Feature: https://news.mit.edu/2023/mit-engineers-how-surfaces-materials-behave-1207

Liu, HF.*, Zhou, Y.*, Huang, Q., Piland, J., Jin, W., Mandel, J., **Du, X.**, Martin, J. & Bartesaghi, A. (2023). nextPYP: a comprehensive and scalable platform for characterizing protein variability in situ using single-particle cryo-electron tomography. *Nature Methods*, 20, 1909. DOI: 10.1038/s41592-023-02045-0 * equal contribution

Bouvette, J.*, Liu, HF.*, **Du, X.**, Zhou, Y., Sikkema, A.P., Mello, J.F.R., Klemm, B.P., Huang, R., Schaaper, R.M., Borgnia, M.J. & Bartesaghi, A. (2021). Beam image-shift accelerated data acquisition for near-atomic resolution single-particle cryo-electron tomography. *Nature Communications*, 12, 1957. DOI: 10.1038/s41467-021-22251-8 * equal contribution

Laasner, R., **Du, X.**, Tanikanti, A., Clayton, C., Govoni, M., Galli, G., Ropo, M. & Blum, V. (2020). MatD^3: A Database and Online Presentation Package for Research Data Supporting Materials Discovery, Design, and Dissemination. *Journal of Open Source Software*, 5(45), 1945. DOI: 10.21105/joss.01945

SELECTED PRESENTATIONS

Accelerating thermodynamic simulations of electrochemical interfaces at the atomic scale

- Oral: Materials Research Society Fall Meeting, Boston, MA, Dec 2024

Machine-learning-accelerated simulations for heuristic-free surface reconstruction

- Oral: IPAM New Mathematics for the Exascale: Applications to Materials Science, Los Angeles, CA, May 2023 (delivered virtually)
- Poster: JUAMI 2023: Materials for a Sustainable Future, Nairobi, Kenya, June 2023
- Oral: Materials Research Society Fall Meeting, Boston, MA, Nov-Dec 2022

HONORS & FELLOWSHIPS

NSF Graduate Research Fellowship 2021

Alex Vasilos Memorial Award (for excellence in computer science research) 2021

Phi Beta Kappa 2021

Duke Faculty Scholar (highest undergraduate honor awarded by faculty) 2020