# Lixue Cheng, Ph.D.

lixuecheng@ust.hk @SherryLixueC
Google Scholar sherrylixuecheng
https://sherrylixuecheng.github.io/ https://ai4qc.github.io/

## Education

August 2017 – June 2022 — **Ph.D. in Chemistry,** California Institute of Technology, Pasadena, CA.
Advisor: Thomas F. Miller III & William A Goddard III
Thesis: *Accurate and transferable molecular-orbital-based machine learning for molecular modeling* [Link]

September 2012 – May 2016 — **Bachelor of Sciences in Honors,** University of Wisconsin-Madison, Madison, WI.
Majors: Chemistry (Honor), Mathematics (Honor), Biochemistry, Molecular Biology. Minor (Certificate): Computer Sciences.

## Work & Research Experiences

August 2025 – Present — **Assistant Professor in Chemistry,** The Hong Kong University of Science and Technology (HKUST)
AI for Physical Science lab: A lab focuses on physics-driven multi-scale modeling strategies via highly accurate simulators, emulators, predictors and evaluators/agents to achieve a collaborative effort of AI, computational chemistry and experimentalists.

June 2023 – May 2025 — **Researcher,** Microsoft Research AI4Science Lab (Berlin).
AI for quantum chemistry: DeepQMC and MLFF; LLMs for Science: ability of GPT-4 in chemistry.

July 2022 – April 2023 — **Research Scientist,** Tencent Quantum Lab, Tencent.
Machine learning assisted quantum computing algorithm design for quantum chemistry.

August 2017 – March 2022 — **PhD Student,** Miller Group, California Institute of Technology.
MOB-ML: Molecular Orbital-Based Machine Learning; INSPIRE: Computational Parameterization of Nucleic Acid Secondary Structure Models.

January 2013 – July 2017 — **Undergraduate Researcher, Experimental Biophysics**, supervised by Prof. Thomas M. Record, Jr., University of Wisconsin-Madison.
Interpreting and predicting Hofmeister salt ion, polyol, PEG and sugar effects on biopolymer processes using solubility assay and solute-partitioning model.

August 2015 – May 2016 — **Undergraduate Researcher, Computational Math**, supervised by Prof. Qin Li, University of Wisconsin-Madison.
Randomized Singular Value Decomposition algorithm for low rank matrices & its numerical applications.

### Teaching Experiences

2026 Spring — **Instructor** CHEM 3030 Introductory to AI for Science, The Hong Kong University of Science and Technology.

2019 Fall — **TA for Ch. 125b. Elements of Quantum Chemistry,** California Institute of Technology.

2017 Spring — **TA for Chem 562 Physical Chemistry II,** University of Wisconsin-Madison.

2016 Spring — **Grader/TA for Math 341 Linear Algebra (Honor) and Math 561 Differential Geometry,** University of Wisconsin-Madison.

## Work & Research Experiences (continued)

2015 Fall & 2016 Fall

- **Grader/TA for Chem/Biochem 565/665 Biophysical Chemistry,** University of Wisconsin-Madison.

### Community Services

- Leader, AI for Science community
- Organizer, ICML 2024 and NeurIPS 2025 AI4Science Workshop
- Member, Materials Genome Division, Chinese Materials Research Society
- Member, IAS Center for Complex Quantum Systems, HKUST
- Member, IAS Center for AI for Scientific Discoveries, HKUST

## Publications

Google Scholar Citations: 1047; h-index: 13; i10-index: 17

25 Li, W., Ren, J., **Cheng**, **L.** & Gong, C. Autonomous Quantum Simulation through Large Language Model Agents. Jan. 2026. arXiv: 2601.10194 [quant-ph].

24 Wang, C., Huang, L., Wei, X., Qin, T., **Cheng**, **L.** & Zhang, J. Scalable Machine Learning Force Fields for Macromolecular Systems Through Long-Range Aware Message Passing. 2026. arXiv: 2601.03774.

23 Song, Z., Lu, J., Du, Y., Yu, B., Pruyn, T., Huang, Y., Guo, K., Luo, X., Qu, Y., Qu, Y., Wang, Y., Wang, H., Guo, J., Gan, J., Shojaee, P., Luo, D., Bran, A., Li, G., Zhao, Q., Luo, S.-X. L., Zhang, Y., Zou, X., Zhao, W., Zhang, Y., Zhang, W., Zheng, S., Zhang, S., Khan, S., Rajabi-Kochi, M., Paradi-Maropakis, S., Baltoiu, T., Xie, F., Chen, T., Huang, K., Luo, W., Fang, M., Yang, X., **Cheng**, **L.**, He, J., Hassoun, S., Zhang, X., Wang, W., Reddy, C., Zhang, C., Zheng, Z., Wang, M., Cong, L., Gomes, C., Hsieh, C.-Y., Nandy, A., Schwaller, P., Kulik, H., Jia, H., Sun, H., Moosavi, S. & Duan, C. Evaluating Large Language Models in Scientific Discovery. 2025. arXiv: 2512.15567.

22 Chen, Y., **Cheng**, **L.**, Jing, Y. & Zhong, P. Benchmarking a foundation potential against quantum chemistry methods for predicting molecular redox potentials. *arXiv preprint arXiv:2510.24063* (2025).

21 Jacobson, G. M., Webb, A. W., **Cheng**, **L.** & McCoy, A. B. Diffusion Monte Carlo Study of the Structure and Spectroscopy of H3O–. *TJ. Phys. Chem. A* **129,** 10928–10939 (2025).

20 Simm, G. N. C., Hélie, J., Schulz, H., Chen, Y., Simeon, G., Kuzina, A., Martinez-Baez, E., Gasparotto, P., Tocci, G., Chen, C., Li, Y., **Cheng**, **L.**, Wang, Z., Nguyen, B. H., Smith, J. A. & Sun, L. SimPoly: Simulation of Polymers with Machine Learning Force Fields Derived from First Principles. 2025. arXiv: 2510.13696 [physics.chem-ph].

19 Foster, A., Schätzle, Z., Szabó, P. B., **Cheng**, **L.**, Köhler, J., Cassella, G., Gao, N., Li, J., Noé, F. & Hermann, J. An ab initio foundation model of wavefunctions that accurately describes chemical bond breaking. *arXiv preprint arXiv:2506.19960* (2025).

18 Sun, J., **Cheng**, **L.** & Zhang, S.-X. Stabilizer ground states for simulating quantum many-body physics: theory, algorithms, and applications. *Quantum* **9,** 1782 (2025).

17 Jacobson, G. M., **Cheng**, **L.**, Bhethanabotla, V. C., Sun, J. & McCoy, A. B. Machine learning approaches for developing potential surfaces: Applications to OH–(H2O)n (n= 1–3) complexes. *J. Phys. Chem. A* **129,** 2958–2972 (2025).

16 **Cheng**, **L.**, Szabó, P. B., Schätzle, Z., Kooi, D. P., Köhler, J., Giesbertz, K. J., Noé, F., Hermann, J., Gori-Giorgi, P. & Foster, A. Highly accurate real-space electron densities with neural networks. *J. Chem. Phys.* **162** (2025).

15 **Cheng**, **L.**, Chen, Y.-Q., Zhang, S.-X. & Zhang, S. Quantum approximate optimization via learning-based adaptive optimization. *Comm. Phys.* **7,** 83 (2024).

14 Sun, J., **Cheng**, **L.** & Li, W. Toward Chemical Accuracy with Shallow Quantum Circuits: A Clifford-Based Hamiltonian Engineering Approach. *J. Chem. Theory Comput.* **20,** 695–707 (2024).

13 AI4Science, M. & Quantum, M. The Impact of Large Language Models on Scientific Discovery: a Preliminary Study using GPT-4. 2023. arXiv: 2311.07361.

12 Li, W., Allcock, J., **Cheng**, **L.**, Zhang, S.-X., Chen, Y.-Q., Mailoa, J. P., Shuai, Z. & Zhang, S. TenCirChem: An Efficient Quantum Computational Chemistry Package for the NISQ Era. *J. Chem. Theory Comput.* **19,** 3966–3981 (2023).

11 **Cheng**, **L.**, Sun, J., Deustua, J. E., Bhethanabotla, V. C. & Miller III, T. F. Molecular-orbital-based machine learning for open-shell and multi-reference systems with kernel addition Gaussian process regression. *J. Chem. Phys.* **157,** 154105 (2022).

10 Sun, J., **Cheng**, **L.** & Miller, T. F. Molecular dipole moment learning via rotationally equivariant derivative kernels in molecular-orbital-based machine learning. *J. Chem. Phys.* **157,** 104109 (2022).

9 **Cheng**, **L.**, Yang, Z., Liao, B., Hsieh, C. & Zhang, S. ODBO: Bayesian optimization with search space prescreening for directed protein evolution. *Preprint at https://arXiv.org/abs/2205.09548* (2022).

8 **Cheng**, **L.**, Sun, J. & Miller III, T. F. Accurate molecular-orbital-based machine learning energies via unsupervised clustering of chemical space. *J. Chem. Theory Comput.* **18,** 4826–4835 (2022).

7 Lu, F., **Cheng**, **L.**, DiRisio, R. J., Finney, J. M., Boyer, M. A., Moonkaen, P., Sun, J., Lee, S. J., Deustua, J. E., Miller III, T. F., *et al.* Fast Near Ab Initio Potential Energy Surfaces Using Machine Learning. *J. Phys. Chem. A* **126,** 4013–4024 (2022).

6 Sun, J., **Cheng**, **L.** & Miller III, T. F. Molecular Energy Learning Using Alternative Blackbox Matrix-Matrix Multiplication Algorithm for Exact Gaussian Process. in *NeurIPS 2021 AI for Science Workshop* (2021).

5 Husch, T., Sun, J., **Cheng**, **L.**, Lee, S. J. & Miller III, T. F. Improved Accuracy and Transferability of Molecular-orbital-based Machine Learning: Organics, Transition-metal Complexes, Non-covalent Interactions, and Transition States. *J. Chem. Phys.* **154,** 064108 (2021).

4 **Cheng**, **L.**, Kovachki, N. B., Welborn, M. & Miller III, T. F. Regression Clustering for Improved Accuracy and Training Costs with Molecular-orbital-based Machine Learning. *J. Chem. Theory Comput.* **15,** 6668–6677 (2019).

3 **Cheng**, **L.**, Welborn, M., Christensen, A. S. & Miller III, T. F. A Universal Density Matrix Functional from Molecular Orbital-based Machine Learning: Transferability across Organic Molecules. *J. Chem. Phys.* **150,** 131103. ISSN: 0021-9606 (Apr. 2019).

2 Welborn, M., **Cheng**, **L.** & Miller III, T. F. Transferability in Machine Learning for Electronic Structure via the Molecular Orbital Basis. *J. Chem. Theory Comput.* **14,** 4772–4779 (Sept. 2018).

1 Knowles, D., Shkel, I. A., Phan, N. M., Sternke, M., Lingeman, E., Cheng, X., **Cheng**, **L.**, O'Connor, K. & Record, M. T. Chemical interactions of polyethylene glycols (PEGs) and glycerol with protein functional groups: applications to effects of PEG and glycerol on protein processes. *Biochemistry* **54,** 3528–3542 (2015).

## Research Interests

- **AI for quantum chemistry:** Orbital-based machine learning for highly accurate ground and excited state wavefunction theory emulation; Ab initio wavefunction learning via Deep Quantum Monte Carlo.
- **AI x quantum computing x quantum chemistry:** Quantum computing for quantum chemistry; Machine Learning for variational quantum algorithms.
- **AI for material design:** Generative modeling and property predictions for advanced materials
- **AI for experimental design:** Bayesian optimization for experimental design.
- **LLM for science:** Agentic AI4S and Scientific agents. Benchmark LLMs in scientific tasks.

## Honors & Awards

2017 **Caltech General Fellowship**, Division of Chemistry and Chemical Engineering, Caltech.

## Honors & Awards (continued)

2016 **Mary Ellen Rudin Foundation Scholarship Award**, Department of Mathematics, UW-Madison.

2015 **David H. Durra Scholarship**, College of Letters & Sciences, UW-Madison.

**Eugene and Patricia Kreger Herscher Undergraduate Scholarship**, Department of Chemistry, UW-Madison.

**Walter W. & Young-Ja C. Toy Summer Research Fellowship**, Department of Chemistry, UW-Madison.

2014 **Biochemistry Undergraduate Summer Research Scholarship**, Department of Biochemistry, UW-Madison.

**Gerald W. and Tui G. Hedstrom Scholarship**, College of Letters & Science, UW-Madison.

**Martha Gunhild Weeks Undergraduate Scholarship**, Department of Chemistry, UW-Madison.

**The Honor Society of Phi Kappa Phi Member**, University of Wisconsin-Madison Chapter.

**Trewartha Senior Thesis Grant**, Letters & Science Honors Program, UW-Madison.

**Wisconsin Hilldale Undergraduate/Faculty Research Fellowship**, UW-Madison.

2013 **John & Elizabeth Moore Awards for Excellence**, Department of Chemistry, UW-Madison.

## Patents

2024 **Object determining method and apparatus, computer device, and storage medium.** US Patent 18745916. [Link]

**Molecular energy prediction method and apparatus, device, and storage medium.** US Patent 18619052. [Link]

2020 **Systems and methods for determining molecular structures with molecular-orbital-based features.** US Patent 16817489. [Link]