YANGZESHENG (ANDREW) SUN

University of Minnesota, Minneapolis 55455, MN, USA *(+1)-612-666-3155*

sun00032@umn.edu | victoriacity.github.io | github.com/victoriacity

EDUCATION

University of Minnesota – Twin Cities

Ph. D., Computational Chemistry (GPA: 4.0/4.0) 2017 – PRESENT

M. S., Computer Science (GPA: 4.0/4.0) 2020 – PRESENT

Related courses: Computer Graphics, Intelligent Agents, Advanced Algorithms & Data Structures

University of California, Berkeley

Visiting Undergraduate Student, Computer Science & Physics (Visiting GPA: 4.0/4.0) Fall 2016

Wuhan University

B. S. with Honors, Chemistry 2013 – 2017

TECHNICAL SKILLS

Languages Python, C++, C#, MATLAB, HTML/CSS

Frameworks PyTorch, Tensorflow, NumPy, CUDA, OpenGL, MPI, OpenMP

Tools Unity Engine, Blender, Linux

RESEARCH EXPERIENCE

University of Minnesota – Twin Cities

Graduate Research Assistant, Department of Chemistry 2018 – Present

Generative machine learning for the spatial distribution of adsorption in molecular simulations

- Developed a parallel HDF5 program collecting trajectory data from 10^5 simulations
- Developed dataset generation and augmentation methods for periodic 3D grids with non-orthogonal axes
- Implemented a Pix2Pix model with 3D-UNet on prediction of trajectories and achieved generation of diverse crystal structure patterns

Predicting hydrogen storage in nanoporous materials using meta-learning

- Developed a meta-learning architecture to extract representations of multiple datasets describing the hydrogen adsorption properties in nanoporous materials as functions of thermodynamic conditions
- Achieved extrapolation outside the range of training data and attained 10x better few-shot performance than domain-specific models
- Applied the meta-learning model to guide the design of hydrogen storage materials for fuel-cell vehicles with prediction results verified by experiments

Machine learning of adsorption equilibria in complex chemical systems

- Developed a learning algorithm which minimizes the KL divergence between the statistical mechanics distribution from molecular simulations and the approximating distribution parametrized by a neural network.
- Achieved improvement by a factor of 10^3 in predicting simulation results compared with domain-specific models
- Developed an interpretable neural network formulation of multicomponent adsorption equilibria based on the Transformer model for natural language processing

Extreme-scale molecular simulations on high-performance computing systems

- Designed and developed high-throughput molecular simulation workflows on 10^4–10^5 CPU cores on leadership-class supercomputers
- Contributed to research proposals awarding more than 10 million CPU hours each year

SOFTWARE PROJECTS

Cities: Skylines Urban Road (CSUR) github.com/citiesskylines-csur/CSUR

Offline procedural generation of realistic road environments in Cities: Skylines

- Designed an offline procedual generation system for road networks based on a custom high-level API over Blender Python backend to produce textured 3D meshes and Unity prefab data containing game mechanics
- Developed a 2D graphics library based on Cairo for the visualization of functionalities of procedually-generated assets
- Developed an automated asset creation pipeline for Cities: Skylines based on the Unity Engine and in-house APIs provided by the game
- Organized public beta programs for extensive play-testing and user feedback
- Developed vector algorithms for the geometric design of roads with player-friendly specifications
- The completed asset/mod pack released on Steam Workshop enjoyed exceptional reception by the Cities: Skylines community and gained more than **18,000** users

TaichiMD github.com/victoriacity/taichimd

Interactive, GPU-accelerated Molecular Dynamics using the Taichi programming language

- Extends capabilities of the Taichi programming language in computer graphics to molecular simulation education and research
- Achieves interactive, real-time molecular dynamics simulations accelerated by GPUs
- Provides a platform for rapid implementation of novel simulation algorithms and machine-learned simulations

SELECTED PUBLICATIONS

- Li, Z.; Bucior, B. J. ; Chen, H.; **Sun, Y.-Z.-S.**; Haranczyk, M.; Siepmann, J. I.; Snurr, R. Q.; Predicting adsorption of n-alkanes and mixtures in metal-organic frameworks using host-guest energy histograms, *J. Phys. Chem. B,* in preparation (invited article for special issue on machine learning)
- **Sun, Y.-Z.-S.**; DeJaco, R. F.; Li, Z.; Tang, D.; Sholl, D. S.; Snurr, R. Q.; Thommes, M.; Hartmann, M.; Siepmann, J. I. Fingerprinting nanoporous materials for hydrogen storage using meta-learning, in preparation.
- Rahbari, A.; Josephson, T. R.; **Sun, Y.-Z.-S.**; Moultos, O.A.; Dubbeldam, D.; Siepmann, J. I.; Vlugt, T. J. H. Multiple linear regression and thermodynamic fluctuations are equivalent for computing thermodynamic derivatives from molecular simulation, *Fluid Phase Equilib.* **2020**, in press.
- Eggimann, B. L.; **Sun, Y.-Z.-S.**; DeJaco, R. F.; Singh, R.; Ahsan, M.; Josephson, T. R.; Siepmann, J. I. Assessing the quality of molecular simulations for vapor–liquid equilibria: an analysis of the TraPPE database, *J. Chem. Eng. Data* **2020**, 65, 1330–1344. DOI: 10.1021/acs.jced.9b00756
- **Sun, Y.-Z.-S.**; DeJaco, R. F.; Siepmann, J. I. Predicting hydrogen storage in nanoporous materials using meta-learning, NeurIPS 2019 Machine Learning and the Physical Sciences Workshop, Vancouver, Canada, **2019**. https://ml4physicalsciences.github.io/files/NeurIPS_ML4PS_2019_47.pdf
- **Sun, Y.-Z.-S.**; DeJaco, R. F; Siepmann, J. I. Deep neural network learning of complex binary sorption equilibria from molecular simulation data, *Chem. Sci.* **2019**, 10, 4377–4388. ***(Cover article)*** DOI: 10.1039/C8SC05340E

PEER REVIEW CONTRIBUTIONS

NeurIPS 2019 Machine Learning and the Physical Sciences Workshop

Journal of Chemical Theory and Computation