YANGZESHENG (ANDREW) SUN

University of Minnesota, Minneapolis 55455, MN, USA *(+1)-612-666-3155*
Email: **sun00032@umn.edu** | Website: **victoriacity.github.io** | GitHub: **github.com/victoriacity**

EDUCATION

University of Minnesota – Twin Cities
M. S. Computer Science **(GPA: 4.0/4.0)** 2020 – PRESENT
Ph. D. Computational Chemistry **(GPA: 4.0/4.0)** 2017 – PRESENT
Representative courses: **Advanced Algorithms & Data Structures, Parallel Computing, Machine Learning, Real-time Simulation & Planning**

Fall 2016
University of California, Berkeley
Visiting Undergraduate Student, Computer Science & Physics **(Visiting GPA: 4.0/4.0)**

2013 – 2017
Wuhan University
B. S. with Honors, Chemistry

TECHNICAL SKILLS

Programming Languages	Python, C++, C#, C, MATLAB, JavaScript, HTML/CSS
Machine Learning	PyTorch, TensorFlow, Keras, OpenCV, Scikit-learn
Big Data Analytics	NumPy, Pandas, Matplotlib, Seaborn, SQL
High-performance Computing	CUDA, MPI, OpenMP, HDF5
Graphics and Simulation	Unity Engine, OpenGL, Blender

SOFTWARE PROJECTS

Real-time Neural Style Transfer for Particle-based Fluid Simulations github.com/victoriacity/nssim

- Implemented real-time fluid simulation methods using smoothed particle hydrodynamics and material point method
- Developed **a neural style transfer method for fluid simulations** based on differentiable simulation and rendering
- Designed an asynchronous simulation/stylization pipeline **achieving more than 80 times speedup** than SOTA results

Interactive molecular and macroscopic dynamics using the Taichi programming language github.com/victoriacity/taichimd

- Extended capabilities of the Taichi programming language in computer graphics to molecular simulations
- Implemented molecular dynamics algorithms and hybrid Eulerian-Lagrangian **solvers for graphics simulations**
- Implemented agent-based simulations and controls using reinforcement learning
- Developed a real-time **implicit surface renderer** and approximate global illumination algorithms for simulation visualization

High definition road network creation through procedural generation and crowdsourced gameplay github.com/citiesskylines-csur/CSUR

- Designed an offline **procedural generation** system for road networks from a semantic representation of lane configurations and traffic modalities based on **a Blender backend**
- Developed an automated object creation pipeline based on **Unity Engine** and in-house game API infrastructure
- Developed graphics algorithms for user-friendly design of **high-fidelity road geometries**
- Led game communities of **over 30,000 users** producing diverse virtual environments for complex road networks

PROFESSIONAL EXPERIENCE

Graduate Research Assistant, *University of Minnesota – Twin Cities* 2018 – PRESENT

3D computer vision on spatial datasets from molecular simulations

- Developed dataset generation and augmentation methods for **sparse periodic 3D voxel grids** with non-orthogonal axes
- Achieved generation of diverse structural patterns in well agreement with simulation results using a **Pix2Pix** image translation model with **3D-UNet**
- Explored incorporation of **physics-based symmetry and equivariance** in machine learning models for various spatial simulation trajectory representations, including voxel grids and point clouds

Physics-informed machine learning on simulation data for clean energy materials discovery

- Collaborated with academic and industrial researchers to identify key scientific and machine learning challenges
- Developed probabilistic, physics-informed machine learning algorithms for **noisy simulation data**
- Developed **meta-learning methods** for accurate few-shot learning of large numbers of simulation tasks with limited data
- Evaluated **transfer learning** and **multi-task learning** techniques within and among simulation systems
- Achieved significant **improvement by factors of 10–10^3** compared with domain-specific models on complex simulations

Extreme-scale molecular simulations on high-performance computing systems

- Deployed and scaled molecular simulations for **10^4–10^5 CPU cores** on national leadership-class supercomputers
- Developed high-performance simulation and data collection software using **MPI, CUDA, and HDF5** on TB-sized datasets
- Built and administered high-performance computing infrastructure and **Linux toolchains** for on-prem clusters
- Authored technical reports and proposals leading to **over 10 million CPU hours awarded** annually

SELECTED PUBLICATIONS

Peer-reviewed conferences and workshops

Sun, Y.-Z.-S.; Josephson, T. R.; Siepmann, J. I. Interpretable Learning of Complex Multicomponent Adsorption Equilibria from Self-attention, **NeurIPS 2020** Machine Learning for Molecules Workshop, Online, 2020.

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.

Scientific journals

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*ria. **2020**, 112785.

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, *Journal of Chemical and Engineering Data*, **2020**, 65, 1330–1344.

Sun, Y.-Z.-S.; DeJaco, R. F; Siepmann, J. I. Deep neural network learning of complex binary sorption equilibria from molecular simulation data, *Chemical Science* **2019**, 10, 4377–4388. ***(Cover article)***