Harry Winston Sullivan

Curriculum Vitae

Department of Chemical Engineering
University of Minnesota
801-6419157
h.sully2015@gmail.com
winstonsullivan.netlify.app/

Professional Summary Statement

- Chemical Engineering PhD student with a focus on computational science and engineering. My diverse background contains professional research in computer science, physics, chemical engineering, machine learning, and finance.
- Highly adept and creative researcher with a proven track record of swiftly conceptualizing, designing, and implementing innovative solutions to intricate engineering challenges. Possesses strong communication skills to effectively convey complex ideas.

Education

2024 – Present **PhD in Chemical Engineering**, *University of Minnesota*, Minneapolis, MN.
Bio-molecular simulation, statistical mechanics, thermodynamics, Bayesian machine learning, non-equilibrium thermodynamics.

2023 – 2024 **Masters in Chemical Engineering**, *University of Utah*, Salt Lake City, UT, GPA: 4.0.
Neutron scattering analysis, molecular simulation, classical and quantum statistical mechanics, thermodynamics, Bayesian statistics and machine learning, continuum mechanics, catalysis, density functional theory. Incomplete.

2019 – 2023 **B.S. in Computational Physics**, *University of Utah*, Salt Lake City, UT, GPA: 3.868.
Classical mechanics, electricity and magnetism, statistical mechanics, quantum mechanics, solid state physics, scientific computing.

2019 – 2023 **Minor in Computer Science**, *University of Utah*, Salt Lake City, UT.
Bayesian machine learning, uncertainty quantification, model validation, data structures and algorithms, neural networks.

Research Experience

2023 – 2024 **Graduate Research assistant in neutron scattering analysis with quantum molecular simulation and machine learning**, *University of Utah*, Salt Lake City, UT, **Dr. Michael Hoepfner**, *Associate Professor* (*link*).
Utilized quantum statistical mechanics in tandem with variational optimization to determine interatomic forces from microstructure of soft matter.

2023 – 2023 **Analysis of cryptocurrency markets using machine learning and scientific computing**, *University of Utah*, Salt Lake City, UT, **Blank Page Studios** (*link*).
3 months of contracted research into the application of machine learning, scientific computing, and data analysis for algorithmically trading cryptocurrency and analyzing NFT sales data. Managed by Kenny Abitbol of Coinbase and Marcus Corbett.

2021 – 2023 **Undergraduate researcher in neutron scattering analysis with molecular simulation and machine learning**, *University of Utah*, Salt Lake City, UT, **Dr. Michael Hoepfner**, *Associate Professor* (*link*).
Application of classical statistical mechanics and Bayesian machine learning to determine interatomic forces from microstructure of soft matter.

2021 – 2021 **Research on laser interferometer gravitational-wave data analysis**, *University of Utah*, Salt Lake City, Utah, **Dr. Yue Zhao**, *Assistant Professor* (*link*).
Summer research opportunity refactoring LIGO Analysis software using Python and Fortran. Created a standard operating procedure for analysis for future use of the tool.

Industry Experience

2023 **Trading Data Analyst**, *Blank Page Studios*.
Utilized machine learning to analyze cryptocurrency market data and NFT sales data.

Publications

In Preparation **H.W. Sullivan** B.L. Shanks, and M.P. Hoepfner, *Exploring Liquid Structure: Uncertainty Quantification and Inference of Radial Distribution Functions with Non-Stationary Gaussian Processes*.
Abstract: The radial distribution function, which characterizes the spatial arrangement of atoms, is a cornerstone in liquid state theory that serves as a vital benchmark for molecular simulations. Our understanding of the liquid state relies heavily on established theoretical relationships that link the radial distribution function to thermodynamic properties and interatomic forces. Despite these profound and intricate connections, atomic structure is often relegated to a mere validation step in molecular modeling, with preference typically given to training force field parameters using macroscopic thermodynamic data or on interatomic potentials computed from quantum mechanical methods. While this approach can yield models that accurately reproduce the thermophysical properties of fluids, it frequently fails to capture what we know about the atomic structure as observed in experiments. We argue that, to truly align molecular models with the principles of statistical mechanics, radial distribution functions must take center stage in force field optimization and design. In this study, we present a probabilistic machine learning framework to estimate total or partial RDFs with UQ/P from neutron scattering data using non-stationary GP regression. We show how non-stationary GPs with a physics-informed mean and kernel conditioned on state-of-the-art experimental data enables the complete reconstruction of the atomic structure as well as thermophysical properties from both simulation and experimentally derived total structure factors.

Dec, 2024 B.L. Shanks, **H.W. Sullivan**, and M.P. Hoepfner, *Bayesian Analysis Reveals the Key to Extracting Pair Potentials from Neutron Scattering Data*, J. Phys. Chem. Lett. 2024, 15, 12608-12618, https://pubs.acs.org/doi/10.1021/acs.jpclett.4c02941.(*link*).
Abstract: The inverse problem of statistical mechanics is an unsolved, century-old challenge to learn classical pair potentials directly from experimental scattering data. This problem was extensively investigated in the 20th century but was eventually eclipsed by standard methods of benchmarking pair potentials to macroscopic thermodynamic data. However, it is becoming increasingly clear that existing force field models fail to reliably reproduce fluid structures even in simple liquids, which can result in reduced transferability and substantial misrepresentations of thermophysical behavior and self-assembly. In this study, we revisited the structure inverse problem for a classical Mie fluid to determine to what extent experimental uncertainty in neutron scattering data influences the ability to recover classical pair potentials. Bayesian uncertainty quantification was used to show that structure factors with random noise smaller than 0.005 to $\sim 30A^{-1}$ are required to accurately recover pair potentials from neutron scattering. Notably, modern neutron instruments can achieve this precision to extract classical force models to within approximately $\pm$ 1.3 for the repulsive exponent, $\pm 0.068A^{-1}$ for atomic size, and 0.024 kcal/mol in the potential well-depth with 95% confidence. Our results suggest the exciting possibility of improving molecular simulation accuracy through the incorporation of neutron scattering data, advancement in structural modeling, and extraction of model independent measurements of local atomic forces in real fluids.

March, 2024 B.L. Shanks, **H.W. Sullivan**, A. R. Shazed and M.P. Hoepfner, *Accelerated Bayesian Inference for Molecular Simulations using Local Gaussian Process Surrogate Models*, J. Chem. Theory Comput. 2024, 20, 3798-3808, https://doi.org/10.1021/acs.jctc.3c01358 (*link*).

Abstract: While Bayesian inference is the gold standard for uncertainty quantification and propagation, its use within physical chemistry encounters formidable computational barriers. These bottlenecks are magnified for modeling data with many independent variables, such as X-ray/neutron scattering patterns and electromagnetic spectra. To address this challenge, we employ local Gaussian process (LGP) surrogate models to accelerate Bayesian optimization over these complex thermophysical properties. The time-complexity of the LGPs scales linearly in the number of independent variables, in stark contrast to the computationally expensive cubic scaling of conventional Gaussian processes. To illustrate the method, we trained a LGP surrogate model on the radial distribution function of liquid neon and observed a 1,760,000-fold speed-up compared to molecular dynamics simulation, beating a conventional GP by three orders-of-magnitude. We conclude that LGPs are robust and efficient surrogate models poised to expand the application of Bayesian inference in molecular simulations to a broad spectrum of experimental data.

Conference Presentations

Contributed Talks

Nov, 2023 **Bayesian Inverse Quantum Theory**, *U of U Theoretical Chemistry Department*, Salt Lake City, UT.

A guided paper review of J.C. Lemm's year 2000 paper on the application of Bayesian field theory to quantum mechanical force field prediction from generic quantum observation.

Sept, 2023 **Learning Interatomic Forces from Fluid Structure with Machine Learning Accelerated Bayesian Optimization**, *ACS*, Laramie, WY.

Abstract: In the realm of liquid state physics, the complex phenomena of self-assembly and structural ordering have fascinated researchers, intimately connecting macroscopic and microscopic behavior. The pursuit of accurate predictions for material behavior from the atomic scale has been a central objective, but the absence of comprehensive theories and computational techniques has posed a challenge for diverse systems. By tackling the computational cost associated with evaluating model force field structures, this study introduces a groundbreaking solution. Through the development of a local Gaussian process surrogate model, we have successfully accelerated the estimation of structural ordering for arbitrary potential energy functions. This advancement enables the application of a Bayesian optimization approach to force field development, unlocking unprecedented computational efficiency. Our results showcase the viability of employing machine learning to expedite interatomic force reconstruction, utilizing experimental diffraction data. This transformative tool not only enables profound insights into structural analysis but also empowers the optimization of force fields.

Poster Presentations

July, 2024 **Foundations of Molecular Modeling and Simulation**, *Department of Energy*, SLC, UT.

Estimation and Uncertainty Quantification of Experimental Radial Distribution Functions Using Non-Stationary Gaussian Processes

Mar, 2022 **Recent Advances in Machine Learning Accelerated Molecular Dynamics**, *Centre Européen de Calcul Atomique et Moléculaire*, Trieste, IT.

Bayesian optimized force fields enabled by a radial distribution function surrogate model. Presented with Brennon L. Shanks PhD Candidate.