

Friday, February 16 • 2:30 PM

PWEB 175

**Data Sampling and Distillation for Neural Network Potentials**

Dr. Gang Seob (GS) Jung

Research Staff

Computational Sciences and Engineering Division

Oak Ridge National Laboratory, Oak Ridge, TN



**Abstract:** Multiscale modeling methods are typically envisioned as precise and predictive simulation tools to solve complex science and engineering problems. However, even conventional atomistic models often lack computational efficiency and accuracy, making them inadequate for providing reliable information for large-scale continuum models. In this seminar, I will discuss the developments aimed at overcoming these critical limitations.

At the beginning of the talk, I will introduce how atomistic models can enhance our understanding of the experimental observations of crystal growth in 2D materials using empirical reactive forcefield (FF). Although the models offer valuable insights at the atomic scale, the development of reliable FFs is severely limited due to the fixed potential expressions. Recently, neural network potentials (NNPs) have emerged to surpass the longstanding limitations of empirical potentials.

While NNPs can provide higher accuracy than other empirical FFs and lower computational costs compared to quantum calculations, efficient sampling or data generation for training becomes increasingly critical. I will present recent advancements in an automated active learning (AL) framework for NNPs, focusing on accurately simulating bond-breaking in hexane chains through steered molecular dynamics sampling with trained NNPs and assessing model transferability to other alkane chains.

In the end, I will introduce one of the sampling approaches, the multiorder-multithermal (MOMT) ensemble, to capture a broad range of liquid- and solid-phase configurations in a metallic system, nickel. Data distillation through active learning can significantly reduce sampled data without losing much accuracy. The NNP, trained from the distilled data, can predict different energy-minimized closed-pack crystal structures even though those structures were not explicitly part of the initial data. Moreover, this data can be applied to other metallic systems (aluminum and niobium) without repeating the sampling and distillation processes.

The capabilities developed through the research will provide valuable tools for a fundamental understanding of the chemical process and mechanistic insights into the predictive design and interpretive simulations of materials processes and properties.

**Biographical Sketch:** GS Jung is a Research Staff at Oak Ridge National Laboratory. His research interests are in the multiscale modeling of materials to understand their fundamental properties from synthesis and growth to performance and failures. Before joining ORNL, he earned his Ph.D. in multiscale modeling for 2D materials from the Massachusetts Institute of Technology.

For additional information, please contact Prof. Georges Pavlidis at [georges.pavlidis@uconn.edu](mailto:georges.pavlidis@uconn.edu) or Victoria Cerwinski at [victoria.cerwinski@uconn.edu](mailto:victoria.cerwinski@uconn.edu)