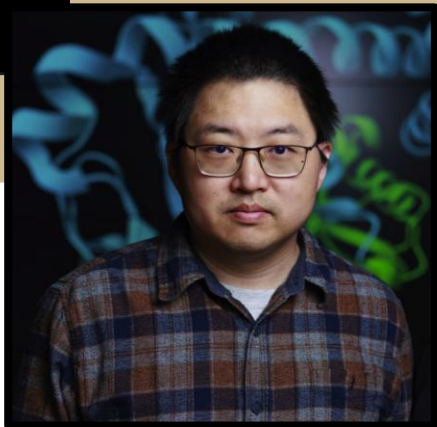


DEPARTMENT COLLOQUIUM

Professor Ming Chen

Department of Chemistry
Purdue University

"Artificial Intelligence Meets Physics"



In recent years, artificial-intelligence (AI) methodologies have been developed for a broad range of scientific applications, including molecular and materials property prediction, protein structure determination, drug discovery, and materials design. These AI-for-science approaches primarily leverage the strong expressivity of deep neural networks together with the massive volumes of experimental and computational data accumulated over decades. Despite the impressive preliminary successes of these models, major challenges remain. In particular, achieving data efficiency, ensuring physical consistency, and enabling reliable extrapolation to regimes not represented in the training data remain open questions. Incorporating physics into AI models represents a promising strategy to address these challenges. In this lecture, I will focus on three complementary directions through which physics can be integrated into AI to enhance accuracy, interpretability, and transferability.

(1) Physics can guide AI at the inference stage. I will present two recently developed methods from my group that demonstrate the strategy of using physics-based constraints to bias inference in both large language models and diffusion models. These approaches enable generation of protein conformations consistent with specific structural constraints or external environments, including protein–environment interactions and experimental restraints. (2) Physics can provide principled frameworks for AI method design. I will discuss reaction-coordinate discovery, formulated as a non-linear manifold-learning problem. My group has recently introduced a physics-guided coarse-graining approach that establishes a general framework for applying non-linear manifold-learning methods to enhanced-sampling trajectories. (3) Physics can directly impose structural constraints, such as symmetry and topology, on AI models. I will highlight our recent work on training an AI-based reweighting method in collective-variable space. By constructing a diffusion model that respects the topology of the collective variables, the model can accurately predict biased data distributions, which is essential for reweighting enhanced sampling simulations. Together, these examples illustrate multiple avenues through which physics can enhance AI models in the physical sciences. The methods discussed demonstrate how integrating physical principles at inference time, during model design, and within model architectures can substantially improve the accuracy, data efficiency, and transferability of AI-for-science frameworks.