# ACMS Statistics Seminar

**Yue Xing**
**Michigan State University**

**Tuesday, March 24, 2026**
**154 Hurley Hall**
**3:30 PM – 4:30 PM**

## Divergence-Based Reinforcement Learning Algorithms for General LLM Alignment

Alignment is the final stage in the training pipeline of large language models (LLMs), aimed at instilling capabilities beyond language comprehension and instruction-following that are learned during pretraining and instruction fine-tuning. While various studies have developed and enhanced alignment algorithms for different tasks and data, there lacks a unified understanding on how and why these algorithms work. In this work, we provide a unified analysis framework from a divergence-based perspective and consider general alignment settings, such as reinforcement learning with verifiable rewards (RLVR) and Preference Alignment (PA). Within this unified framework, we identify some potential issues in existing alignment methods, such as the widely used Group Relative Policy Optimization (GRPO). We further propose $f$-GRPO, a class of on-policy reinforcement learning, and $f$-Hybrid Alignment Loss ($f$-HAL), a hybrid on/off policy objectives, for general LLM alignment based on variational representation of $f$-divergences.

The Department of Applied and Computational Mathematics and Statistics

Please visit acms.nd.edu to view the full list of speakers.