The Theory of Parameter Condensation in Neural Networks

Invited Speaker
Tao Luo
Shanghai Jiao Tong University
Date: **August 4, 2025 (Monday)**
Time: **10:30am- 12:00am (Hong Kong Time)**
Zoom Meeting: **984 9526 8280**

Biography

Tao Luo is an Associate Professor at the School of Mathematical Sciences/Institute of Natural Sciences at Shanghai Jiao Tong University (SJTU). His research focuses on the mathematical theory of machine learning and materials science. He received his Bachelor's degree in 2012 from Zhiyuan College at SJTU and his Ph.D. from the Hong Kong University of Science and Technology (HKUST) in 2017, where he was awarded the Hong Kong Mathematical Society's Best PhD Thesis Award. From 2017 to 2020, he served as a Golomb Visiting Assistant Professor in the Department of Mathematics at Purdue University. He has made contributions to research in the frequency principle and condensation phenomena in deep learning, as well as the Peierls-Nabarro model, epitaxial growth, and the Cauchy-Born rule in materials science. His work has been published in top international conferences and journals, including the SIAM journal series, Arch. Ration. Mech. Anal., J. Mach. Learn. Res., NeurIPS, ICLR.

Abstract

In this talk, we will first introduce the phenomenon of parameter condensation in neural networks, which refers to the tendency of certain parameters to converge towards the same values during training. Then, for certain types of networks, we prove that condensation occurs in the early stages of training. We further analyze which hyperparameters and training strategies influence parameter condensation. In some cases, we even provide a phase diagram that delineates whether parameter condensation occurs. We will also briefly discuss the relationship between parameter condensation and generalization ability. Finally, towards the end of the training, we study the set of global minima and present a detailed analysis of its geometric structure and convergence properties.