



THE UNIVERSITY OF
CHICAGO

Computational and Applied Mathematics
&
Statistics Student Seminar

Suzanna Parkinson

Computational and Applied Mathematics
University of Chicago

The Role of Depth in Neural Networks: a Representation Cost Perspective

Tuesday, October 24, 2023

12:30 PM

Jones Laboratory,
Jones 303

ABSTRACT

Even shallow neural networks can approximate any continuous function, so why do deeper networks often outperform shallow ones? One approach to this question is via representation costs, or the norms of parameters needed to represent functions using different neural network architectures. The inductive bias associated with weight decay is towards functions with smaller cost, and so low-cost functions will be easier to learn. I will discuss characterizations of which kinds of functions have low representation cost under various network architecture assumptions, including adding linear layers to a ReLU network. I will also discuss how understanding representation costs can lead to insight into which kinds of functions can be efficiently learned. In particular, there exist functions that can be learned with polynomially many samples using a 3-layer network, but require exponentially many samples to be learned with a 2-layer network.