

1154-68-1109      **Y Cooper\*** ([yaim@math.ias.edu](mailto:yaim@math.ias.edu)). *The loss function of overparameterized neural networks.*

Training modern neural networks relies on the use of gradient descent and related methods to minimize a non-convex loss function. Though this loss function is not convex, gradient descent comes close to finding global minima remarkably often in many real world settings. This is well established empirically, but in the early stages of being understood theoretically. In this talk, we will discuss recent progress in understanding some geometric properties of the loss function. This can help us understand the ways in which it is not convex, as well as ways in which it is better behaved than an arbitrary nonconvex function. (Received September 13, 2019)