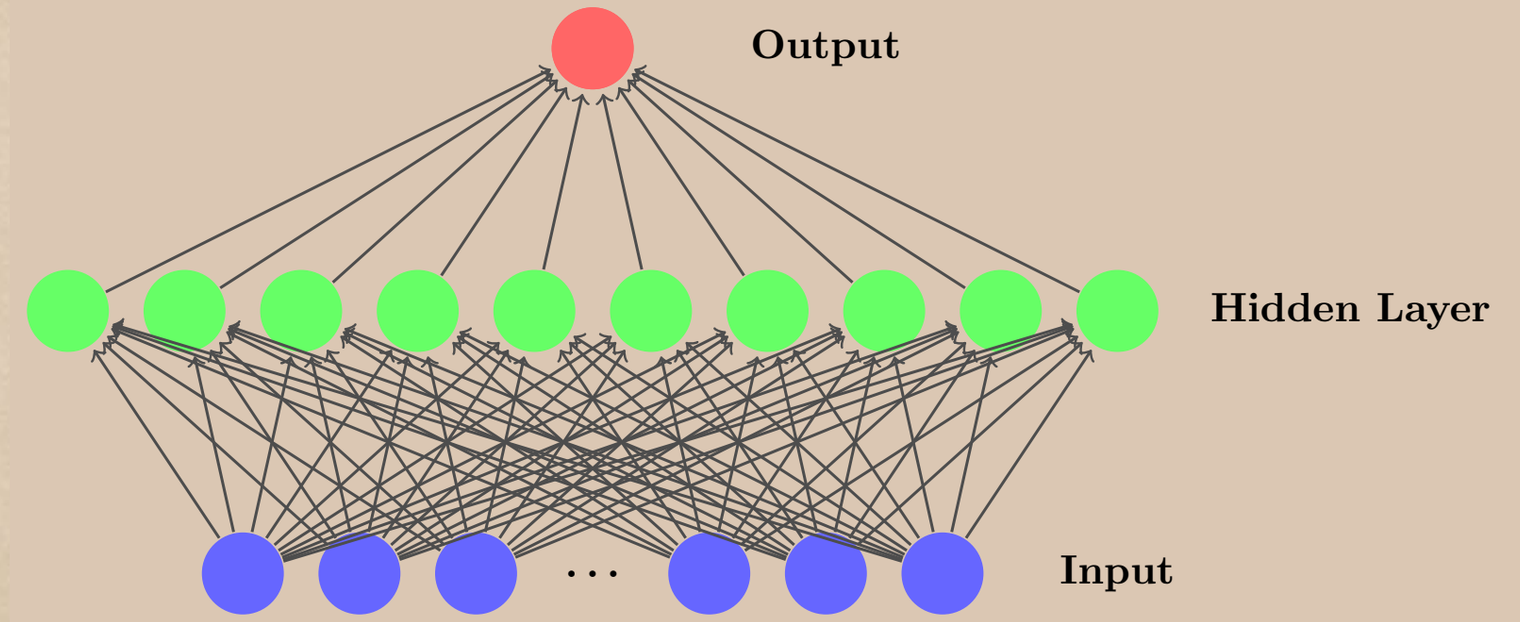
ResNet with one-neuron hidden layers is universal approximator

Hongzhou Lin, Stefanie Jegelka

Poster #28

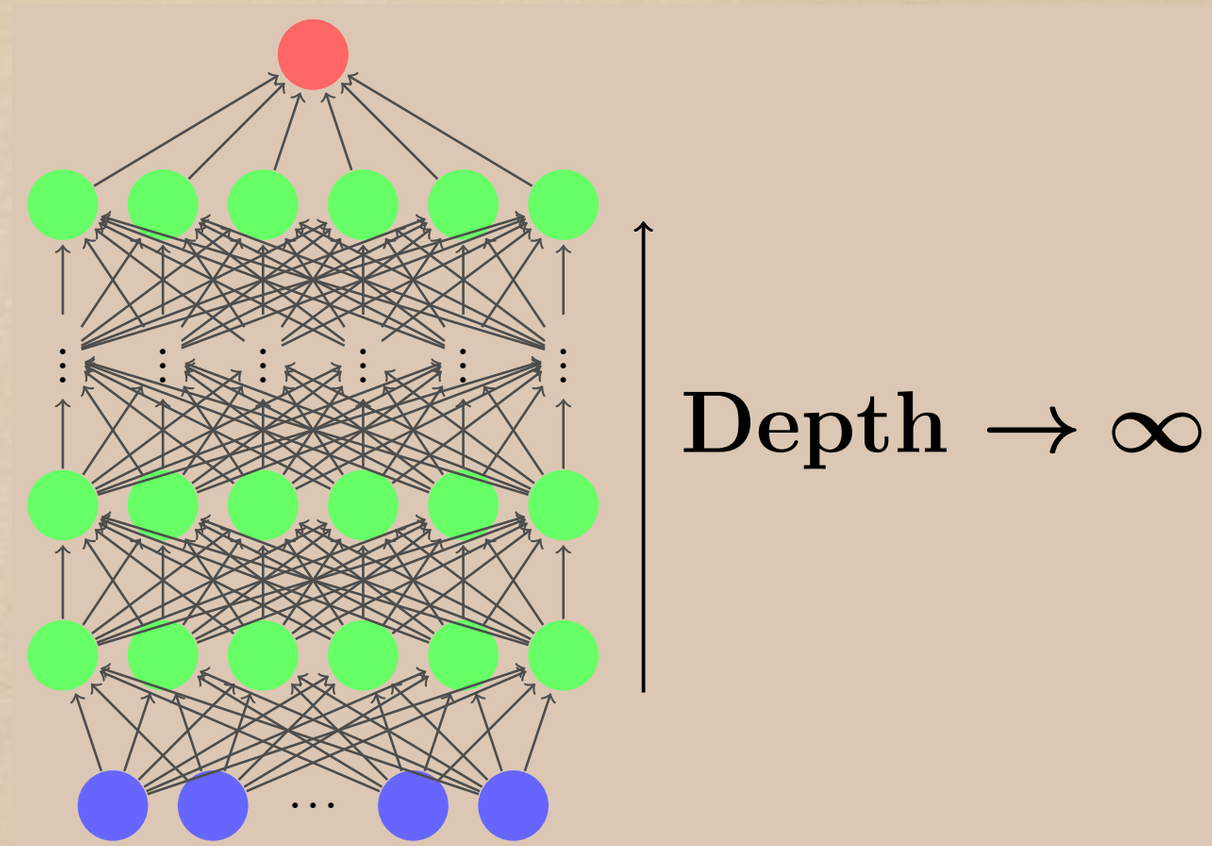
In the 90's: *Universal approximation theorem*



1 hidden layer, width go to infinity →**universal approximation**

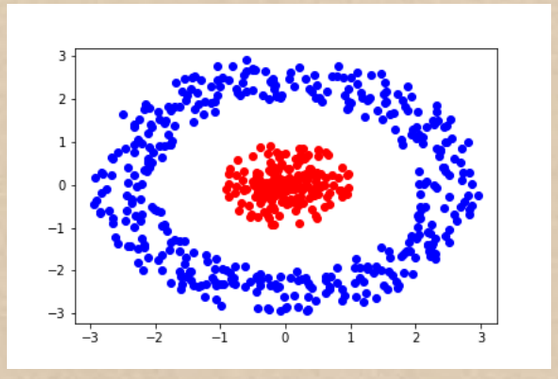
[Cybenko 1989, Funahashi 1989, Hornik et al 1989, Kurková 1992]

Deep Learning



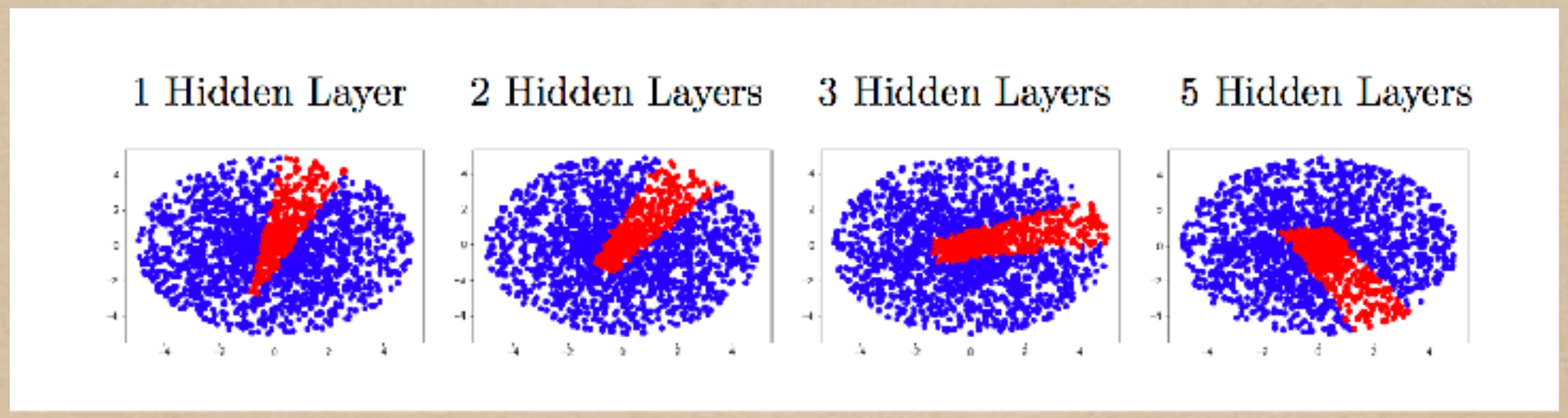
As the depth go to infinity, how many neurons per layer do we need in order to guarantee the theorem?

Classifying the unit ball distribution



Narrow fully connected networks fail!

Narrow: # of neurons per layer ⩽ input dimension **d**



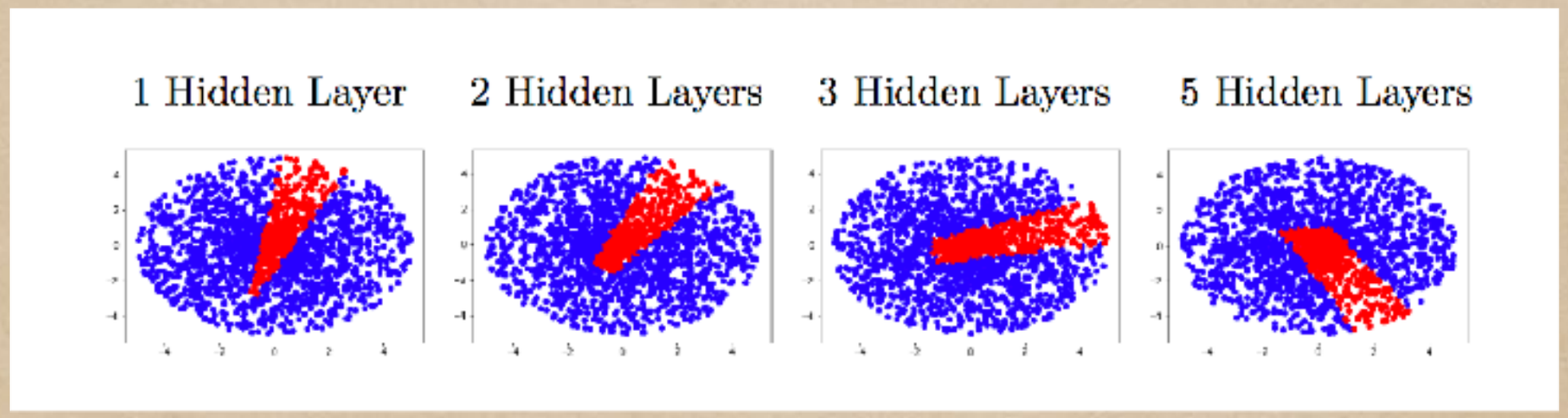
Depth increases

Classifying the unit ball distribution

Theorem [Lu et al 2017, Hanin and Sellke 2017]:
The decision boundary of a narrow FNN is always unbounded.

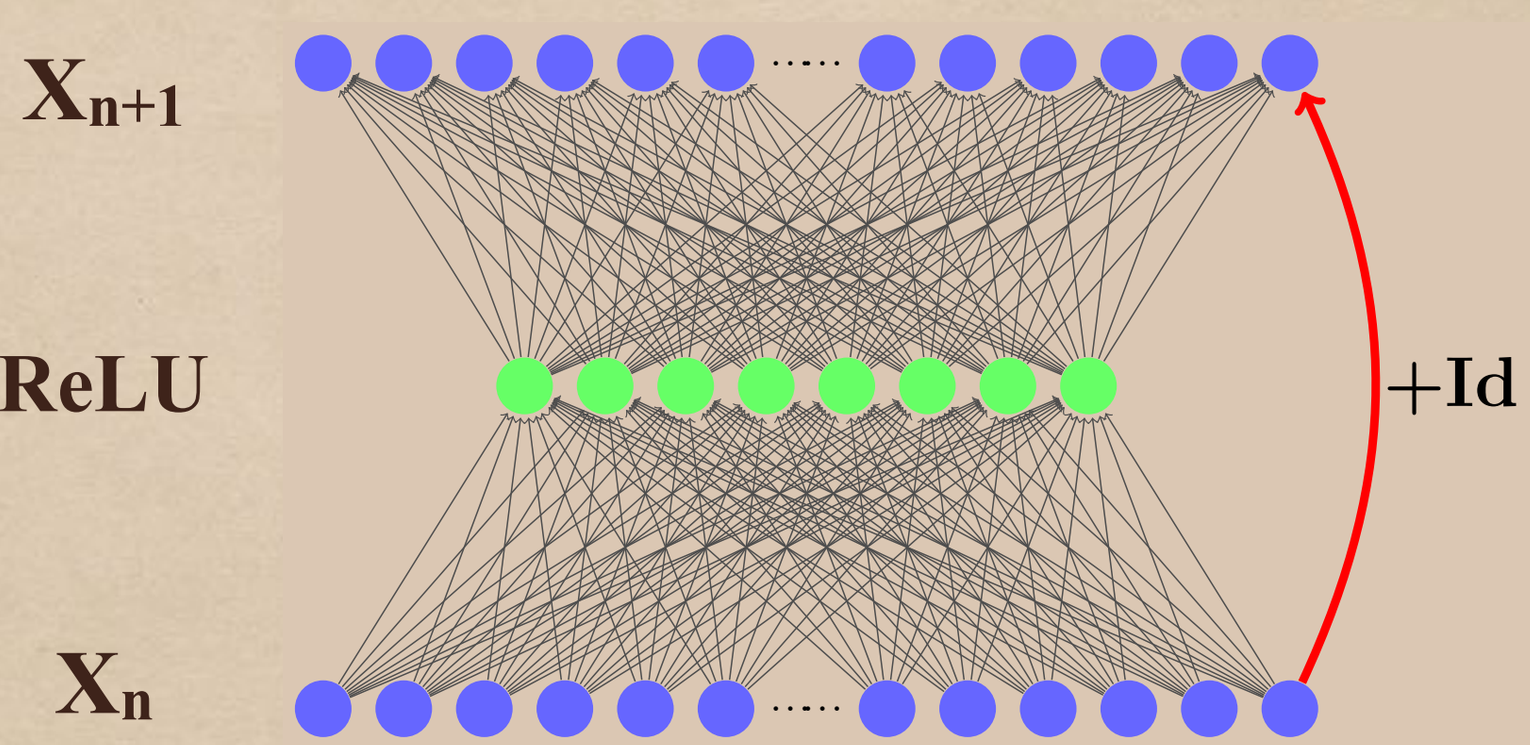
Narrow fully connected networks fail!

Narrow: # of neurons per layer ⩽ input dimension **d**



Depth increases

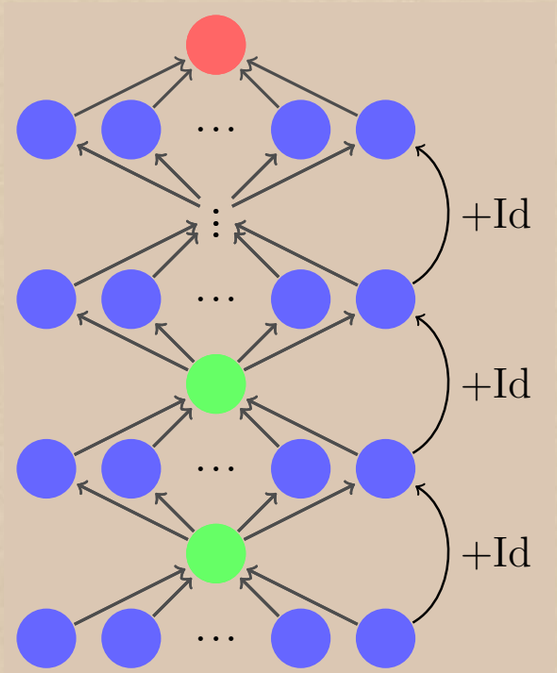
ResNet: residual network

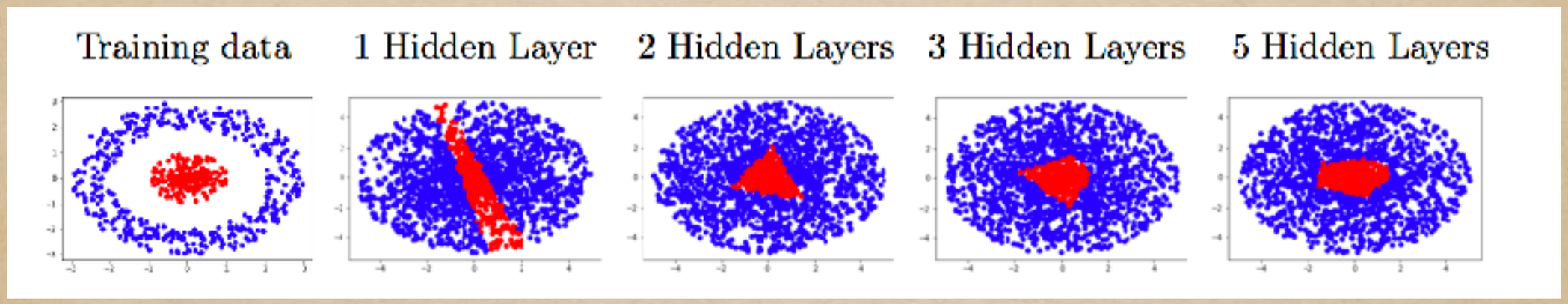


$$X_{n+1} = X_n + V_n \, \mathrm{ReLU}(W_n X_n + b_n)$$

[He et al 2016a, 2016b, Hardt and Ma 2017]

ResNet with one-neuron hidden layers

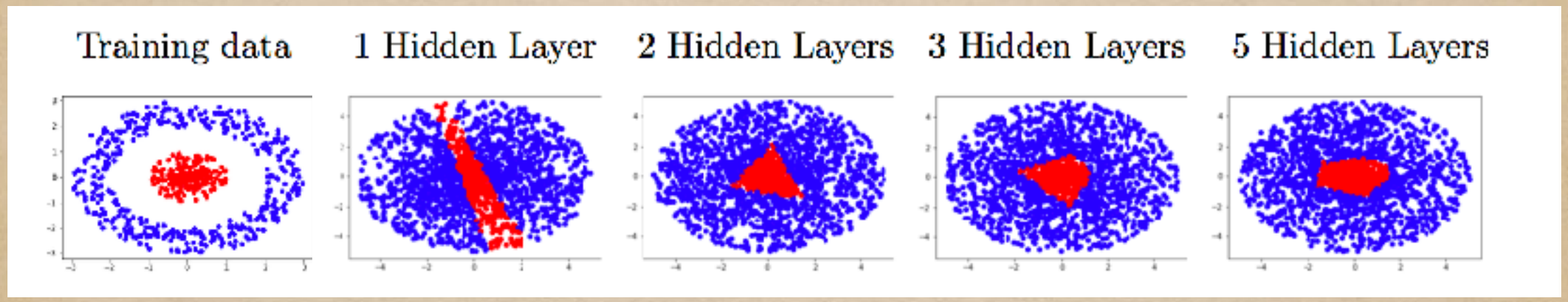




Depth increases

ResNet with one-neuron hidden layers

Theorem: ResNet with **one-neuron** hidden layers is a universal approximator when the depth go to infinity.



Depth increases

Thank you!

Poster #28

05:00 -- 07:00 PM

@ Room 210 & 230 AB