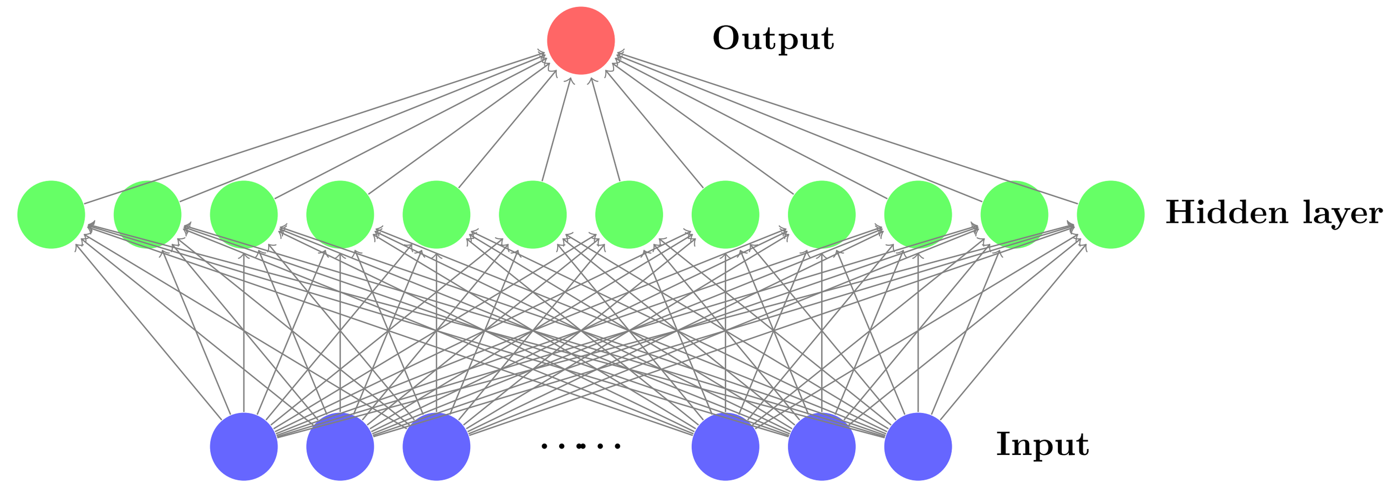
ResNet with one-neuron hidden layer is a Universal Approximator

Hongzhou Lin, Stefanie Jegelka

The representational power of Neural Network

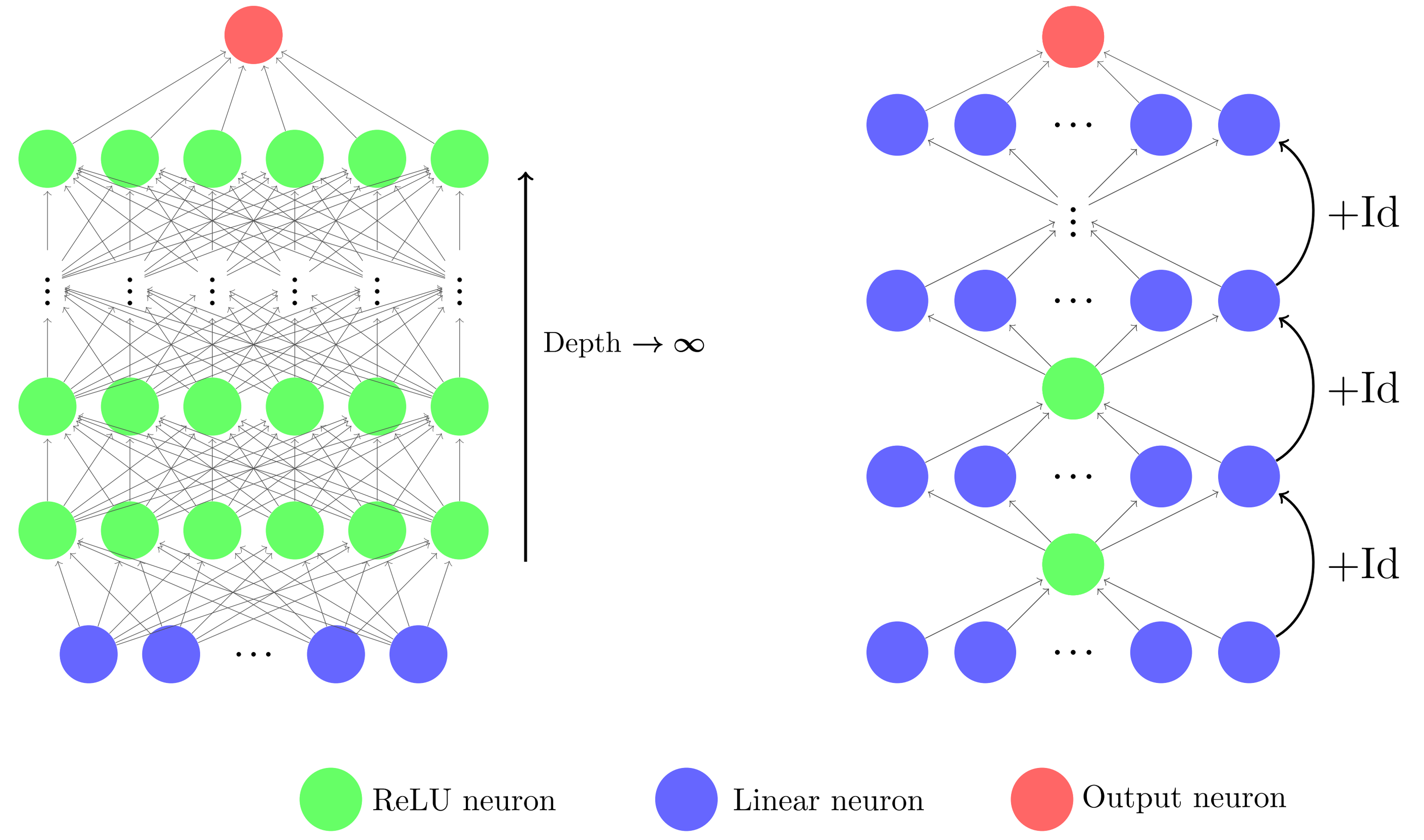
In the 90's: *Universal approximation theorem*



1 Hidden layer, width goes to infinity →**universal approximation**

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

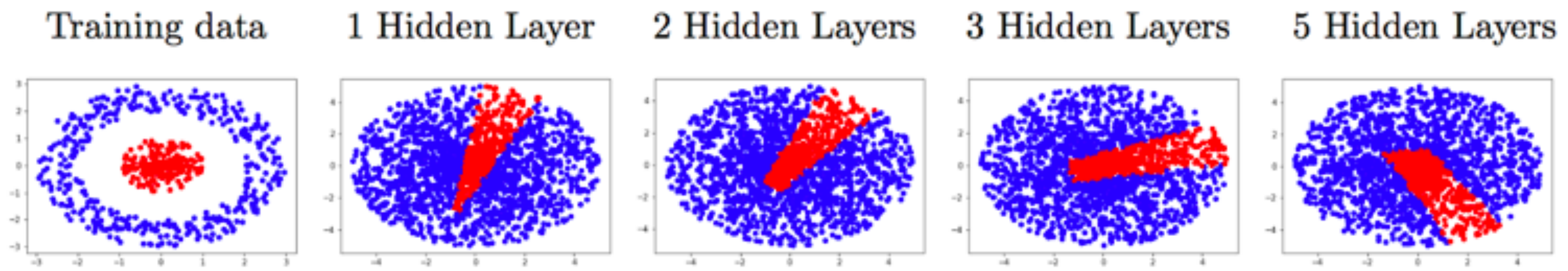
Recently: *Deep Learning with hundreds of layers*



Question: **what is the minimum condition on the width such that universal approximation still hold when the depth goes to infinity?**

A motivating example: Classifying the unit ball distribution

Failure case: Narrow Fully connected Network

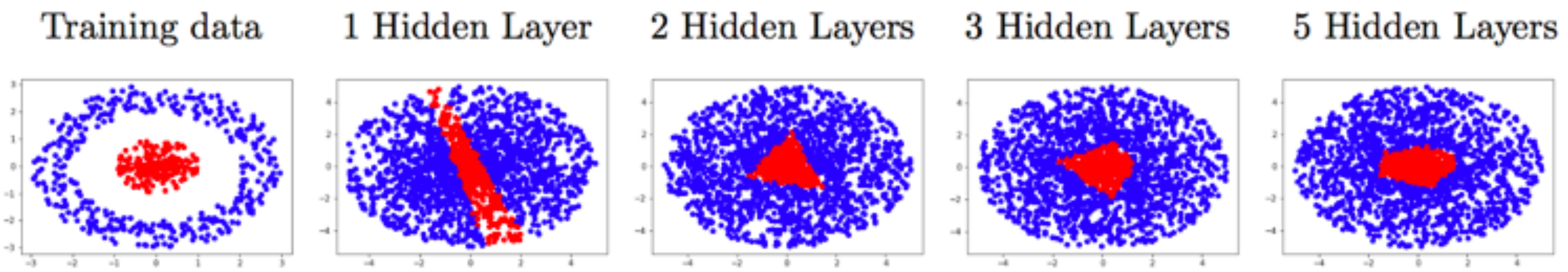


✦ **Narrow**: number of neurons in each hidden layer ⩽ the input dimension.

✦ Here, we apply FNN with 2 neurons per hidden layer using ReLU activation.

Theorem: In the input features are in R^d, a fully connected network with **d** neurons per layer always has unbounded decision boundary. [Lu et al 2017, Hanin et al 2017]

Success case: Residual network with one-neuron hidden layer



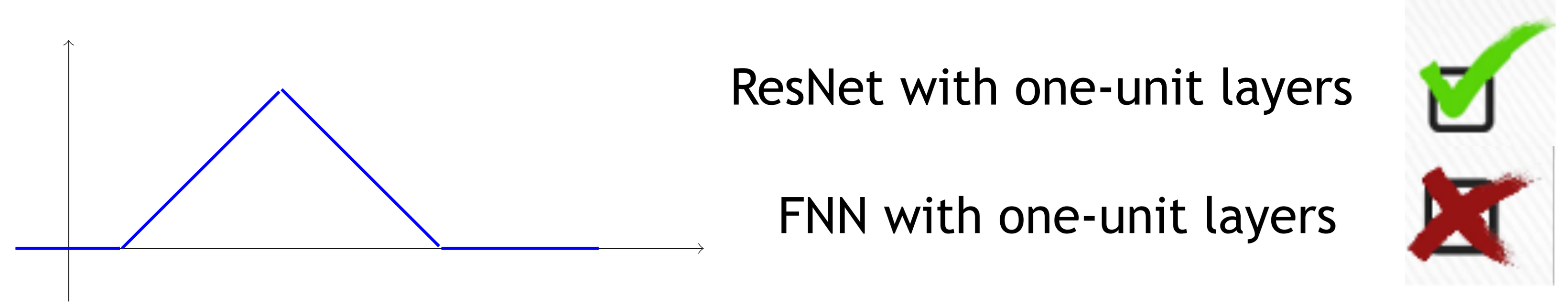
Theorem: ResNet with **one-neuron hidden layer** is a universal approximator in the space of integrable functions $\ell_1(R^d)$. In other words, for any $\epsilon>0$, there is a ResNet R with finitely many layers such that

$$\int_{\mathbb{R}^d} |f(x) - R(x)| dx \leq \epsilon.$$

✦ The result holds for any input dimension **d.**

✦ ResNet: O(d) parameters **vs** Fully connected network: $O(d^2)$ parameters.

A sketch of proof for one dimension case

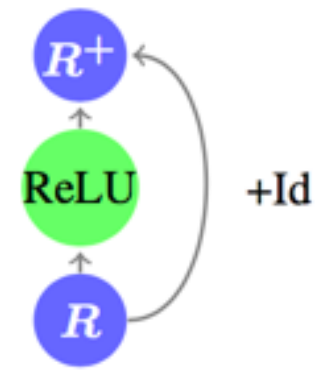


Operations realizable by one-neuron ResNet:

(a) ***Shifting by a constant***: $R^+ = R + c$.

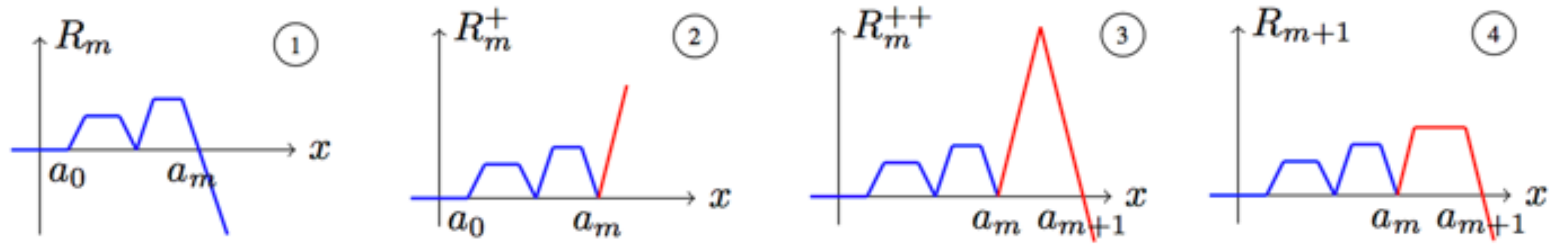
(b) ***Min/Max with a constant:*** $R^+ = \min\{R,c\}$ *or* $R^+ = \max\{R,c\}$.

(c) ***Min/Max with a linear transformation***: $R^+ = \min\{R, \alpha R + \beta\}$.

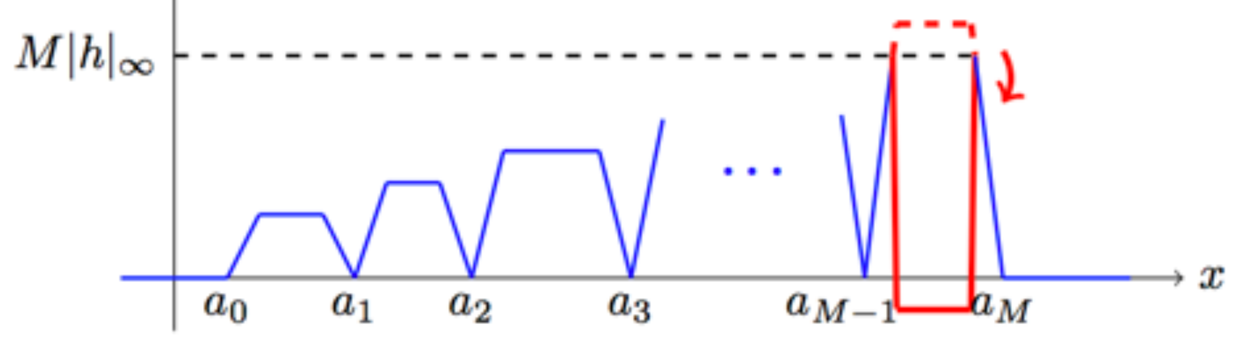


Main strategy: approximate piecewise constant functions.

✦ First, construct an increasing trapezoid function;



✦ Second, adjust the function value on each subdivision.



Take away message:

✦ **ResNet architecture increases the representational power.**

✦ **Stands in sharp contrast to fully connected network.**

NIPS 2018, December 6th, Montreal, Canada