

HERMITE POLYNOMIALS FOR EFFICIENT GENERATION OF PSEUDOLABELS IN SSL



NON-SMOOTH RELUS TO POLYNOMIAL ACTIVATIONS

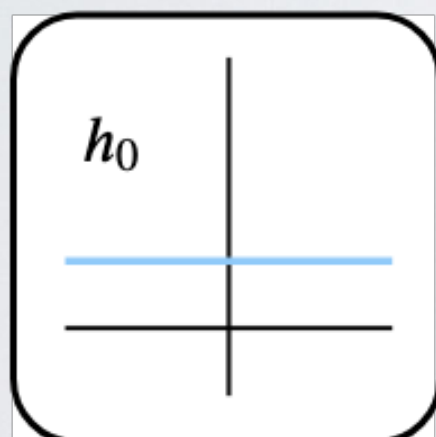


- Ge et al showed that we can avoid spurious local minima using Hermites
- Idea: use the lower order terms in the Hermite polynomial series **Expansion of ReLU** as **Activation Functions**



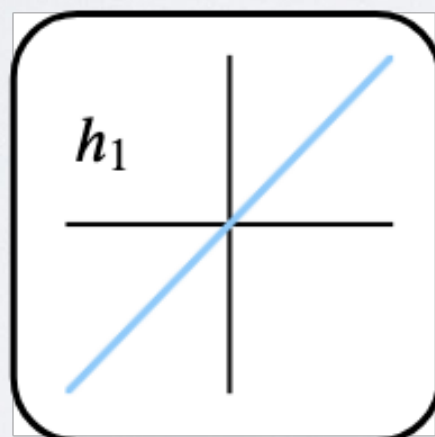
HERMITE POLYNOMIALS AS ACTIVATIONS

h_0



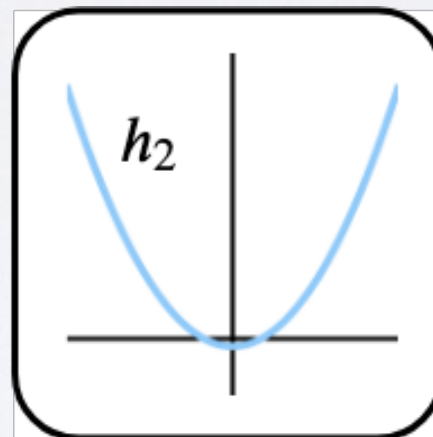
1

h_1



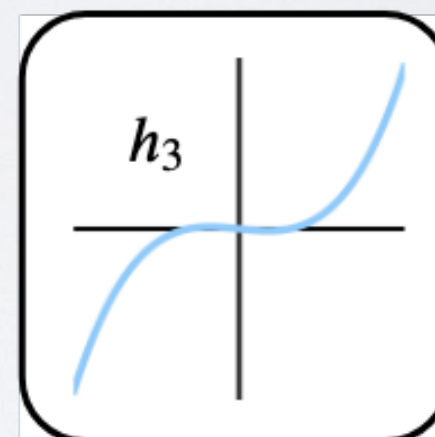
x

h_2



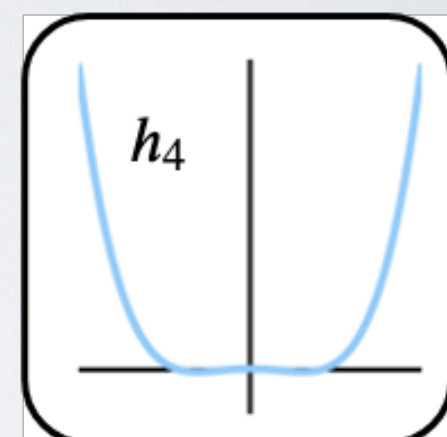
$\frac{x^2 - 1}{2!}$

h_3



$\frac{x^3 - 3x}{3!}$

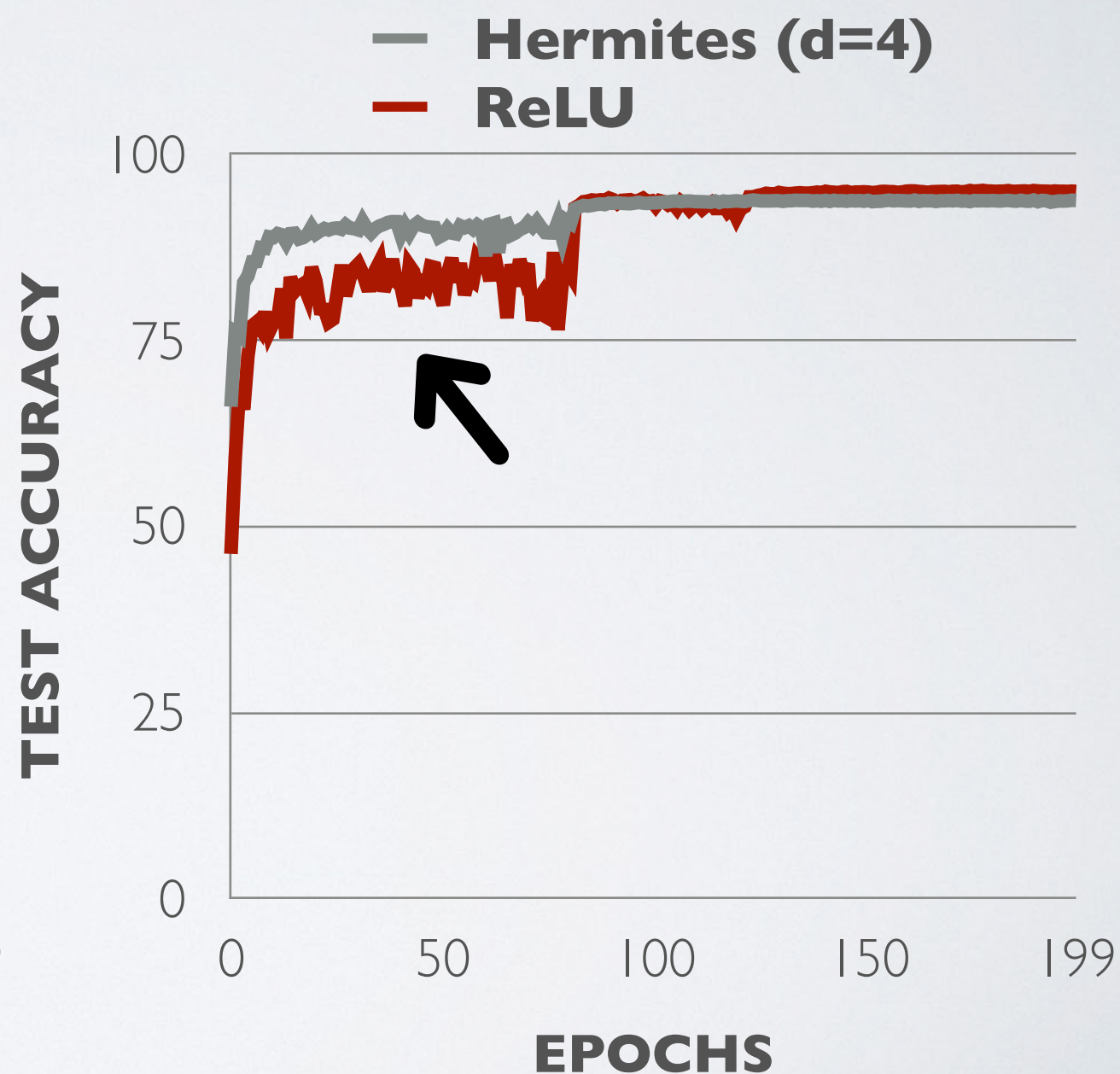
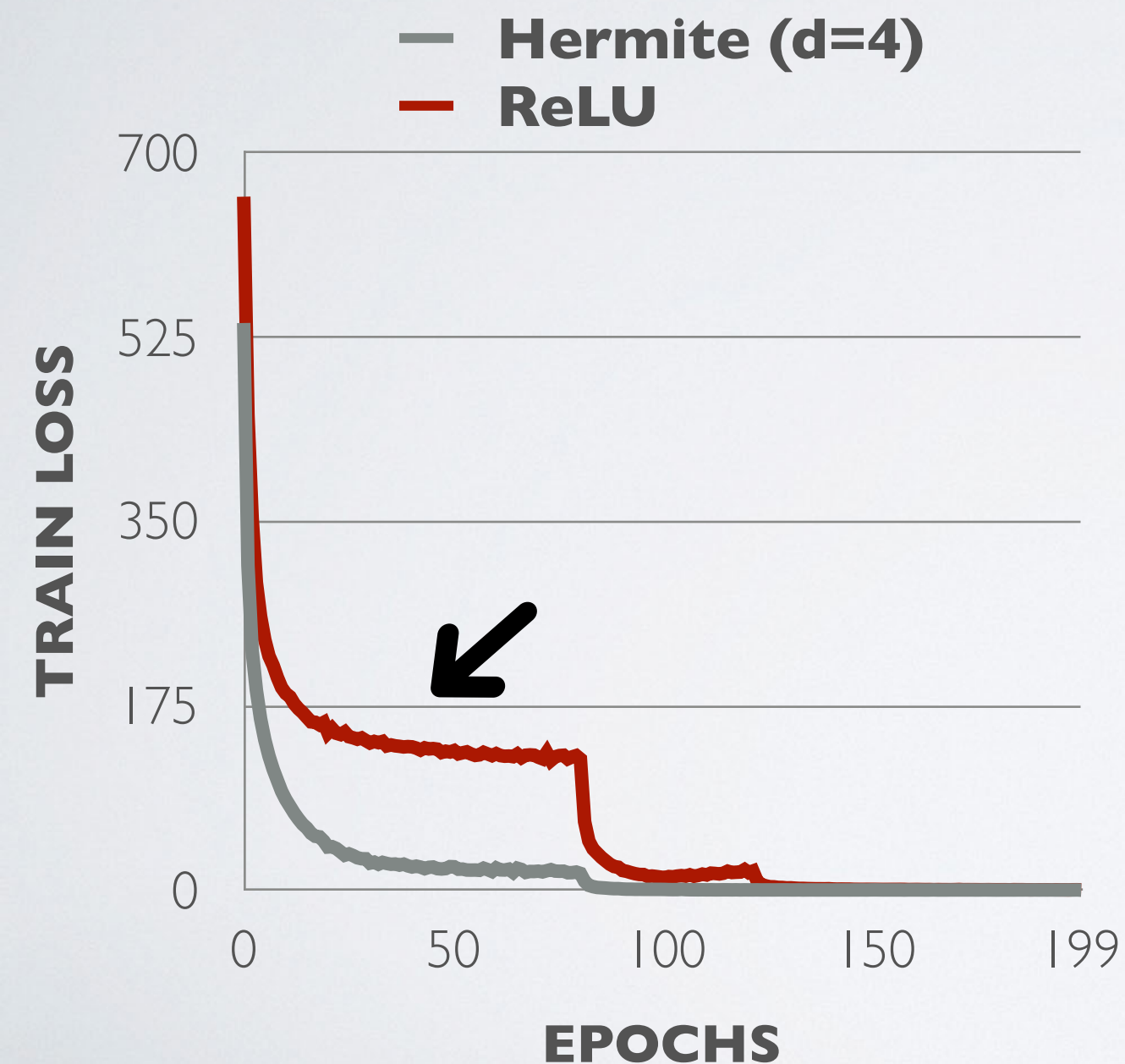
h_4



$\frac{x^4 - 6x^2 + 3}{4!}$



HERMITES POSSESS EARLY RISER PROPERTY



HERMITES MAKE CONSCIOUS CLASSIFICATIONS



	Train Similar to Test	Train Different From Test	
ReLU	High Confidence Predictions	High Confidence Predictions	
Hermite	High Confidence Predictions	Low Confidence Predictions	

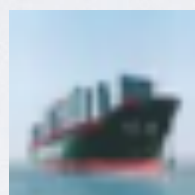
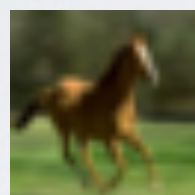
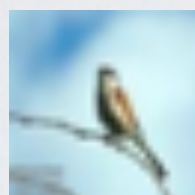
When the test data is different from the training data,
Hermite networks consciously make (**approximately**)
random predictions unlike ReLU networks

THE SAAS CONCEPT

PSEUDO-LABELS GENERATION VIA THE **BANDERSNATCH** PHENOMENON

Good Labels → Least Training Time

CIFAR10

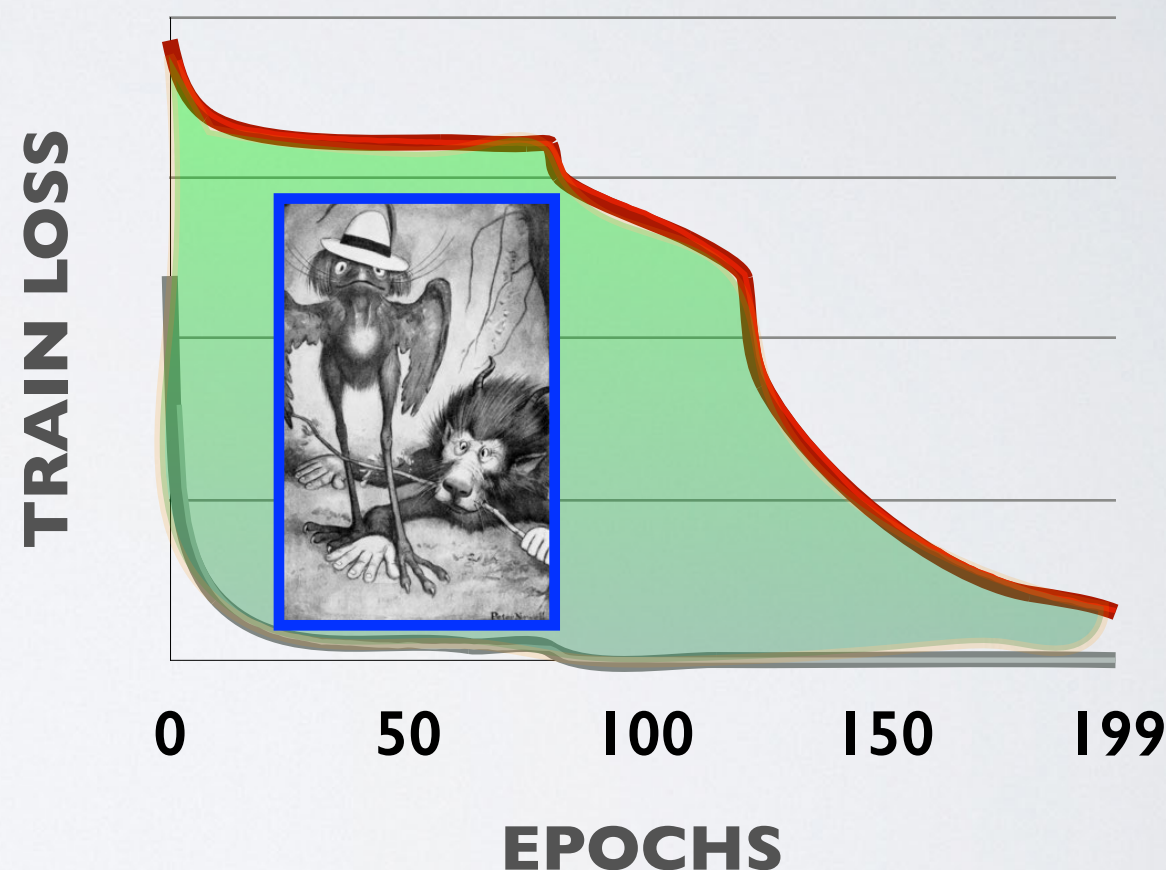


**TRUE
LABELS**

Bird Horse Ship

**RANDOM
LABELS**

Cat Truck Dog



SaaS: Find labels with least training time

COMPUTATIONAL BENEFITS





On AWS
p3.16xlarge

HERMITE-SAAS **TRAINS FASTER** THAN RELU-SAAS

CIFAR10

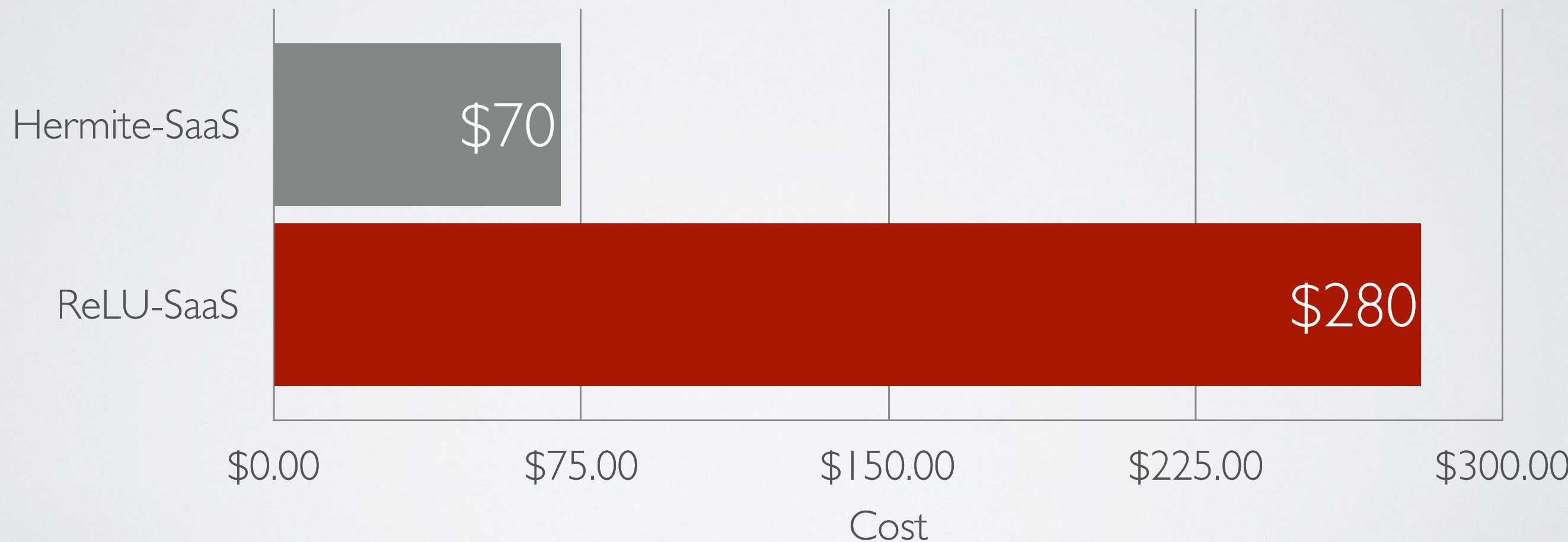




On AWS
p3.1 6xlarge

HERMITE-SAAS **COSTS LESS** THAN RELU-SAAS

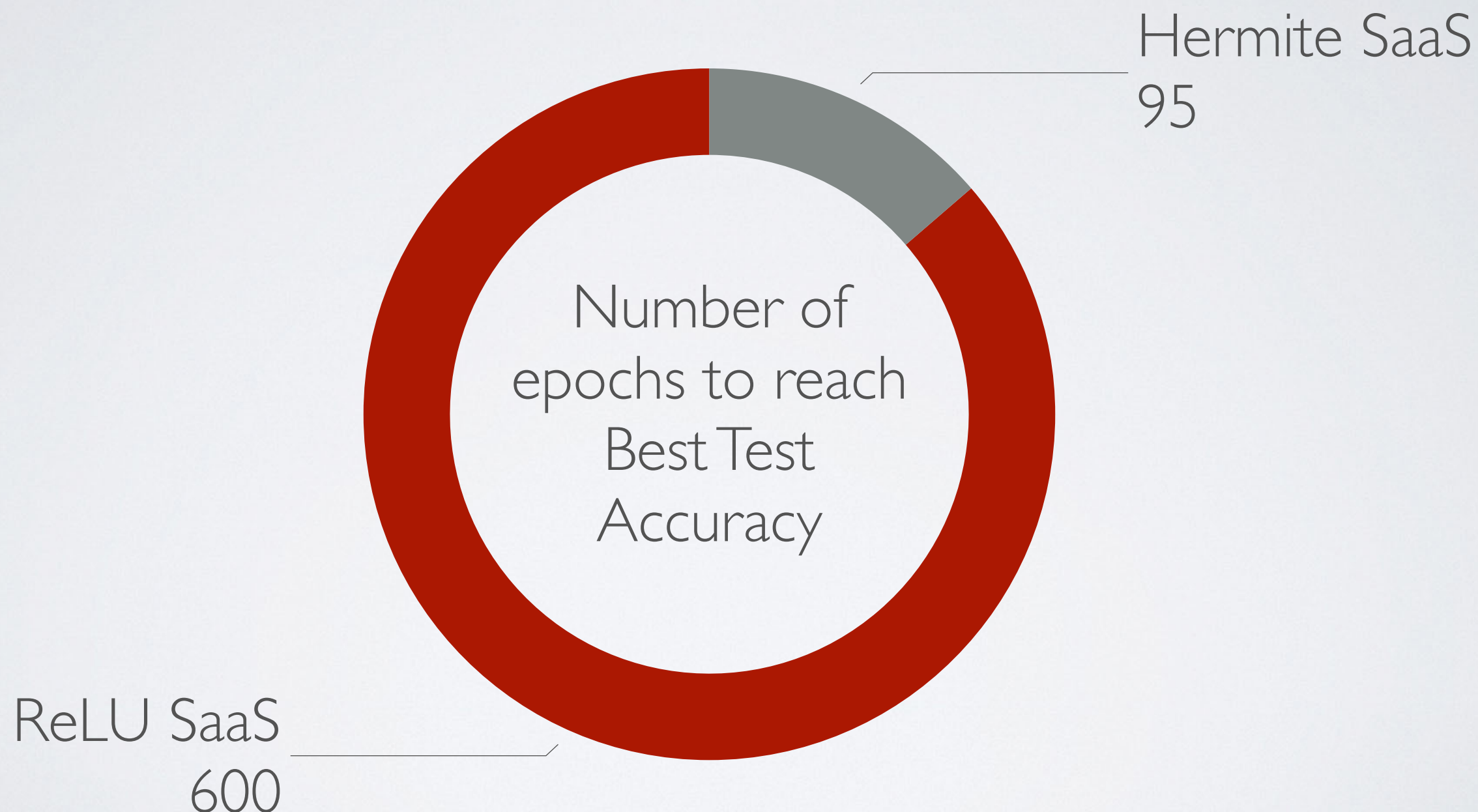
CIFAR10





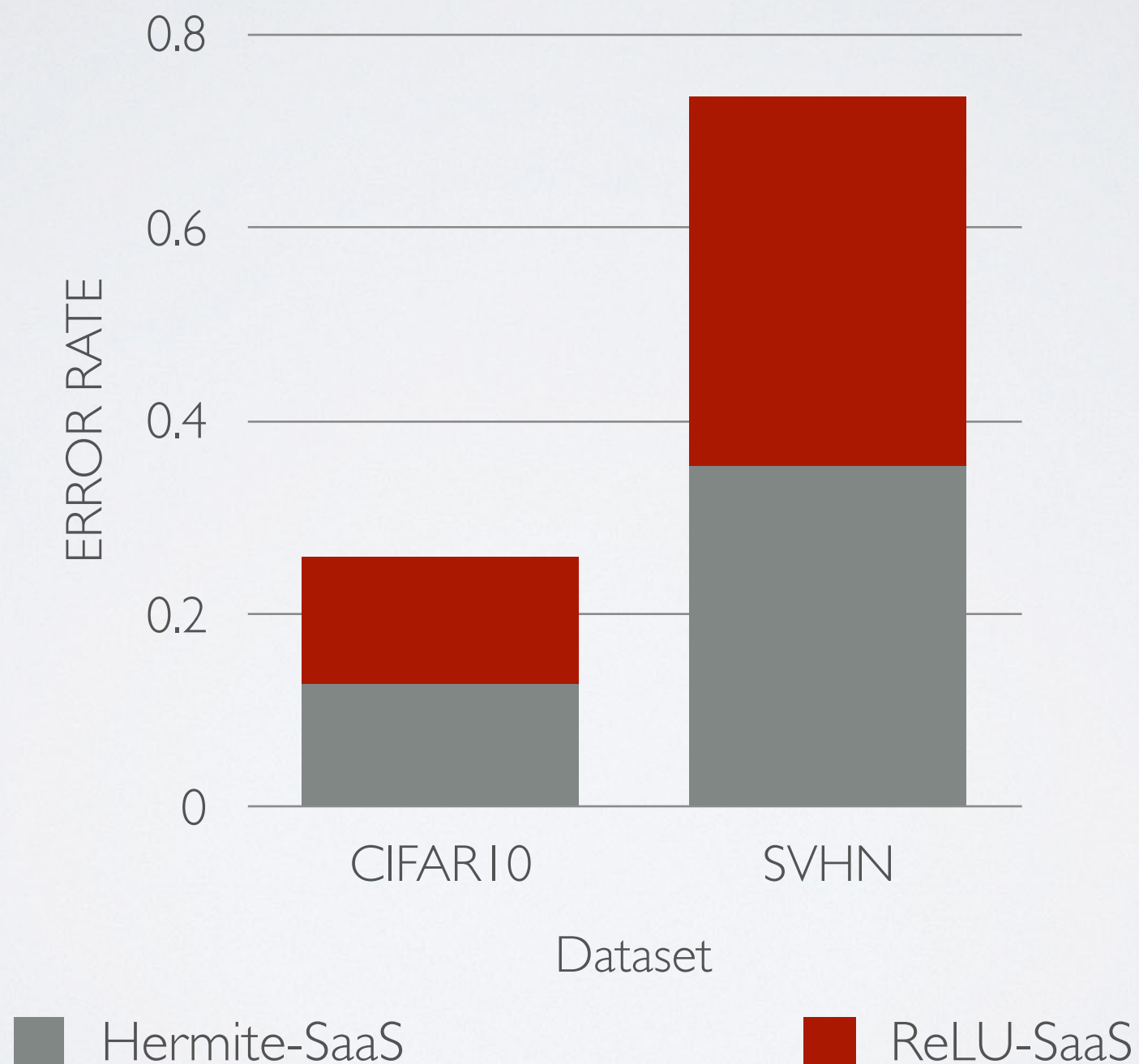
Adding 30% uniform
noise to labelled data

HERMITE-SAAS **MORE RESILIENT TO NOISE** THAN RELU-SAAS





HERMITE-SAAS **GENERALIZES BETTER** THAN RELU-SAAS





Hermite polynomials
are nice