



HyperSPNs: Compact and Expressive Probabilistic Circuits

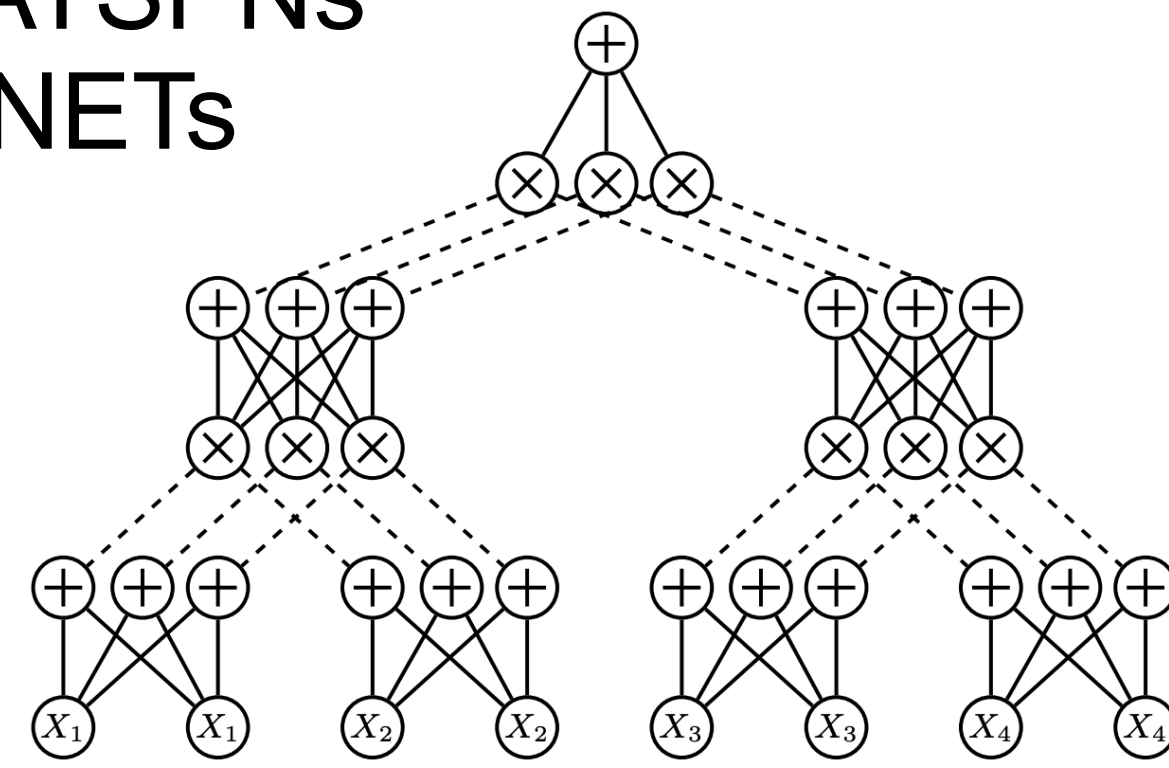
Andy Shih, Dorsa Sadigh, Stefano Ermon {andyshih,dorsa,ermon}@cs.stanford.edu

Paper / Code:



Large SPNs proposed [Peharz et al. '19, '20]

- RATSPNs
- EiNETs



Regularization Choices

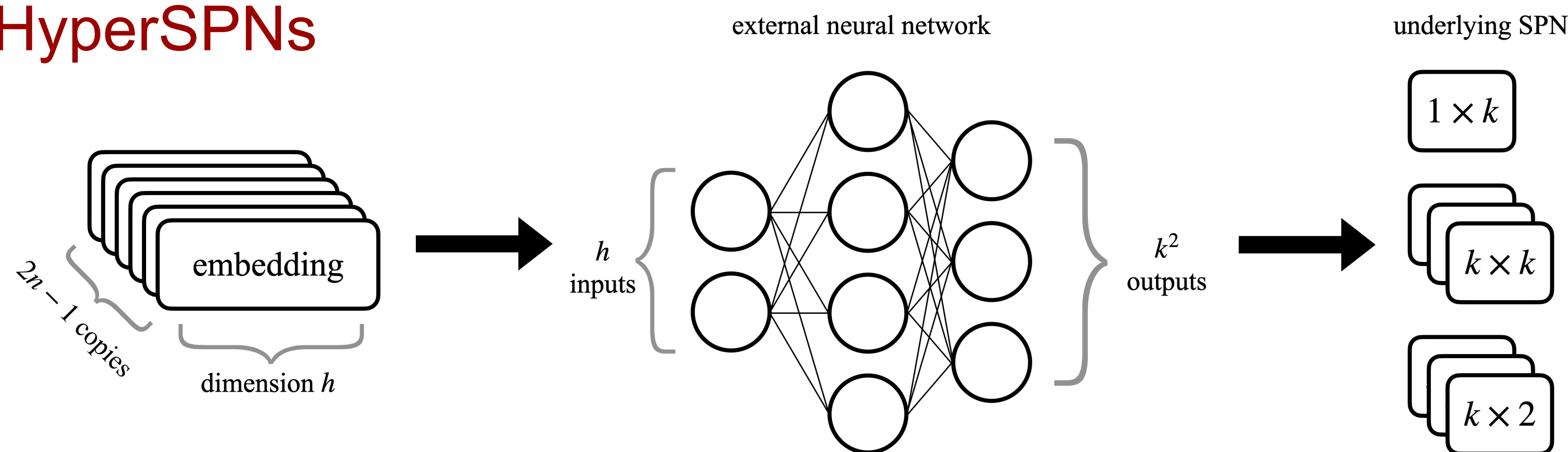
Dropout	Discriminative only
Weight Decay	Many parameters
HyperSPN	Few parameters Memory efficient Better generalization

our proposal

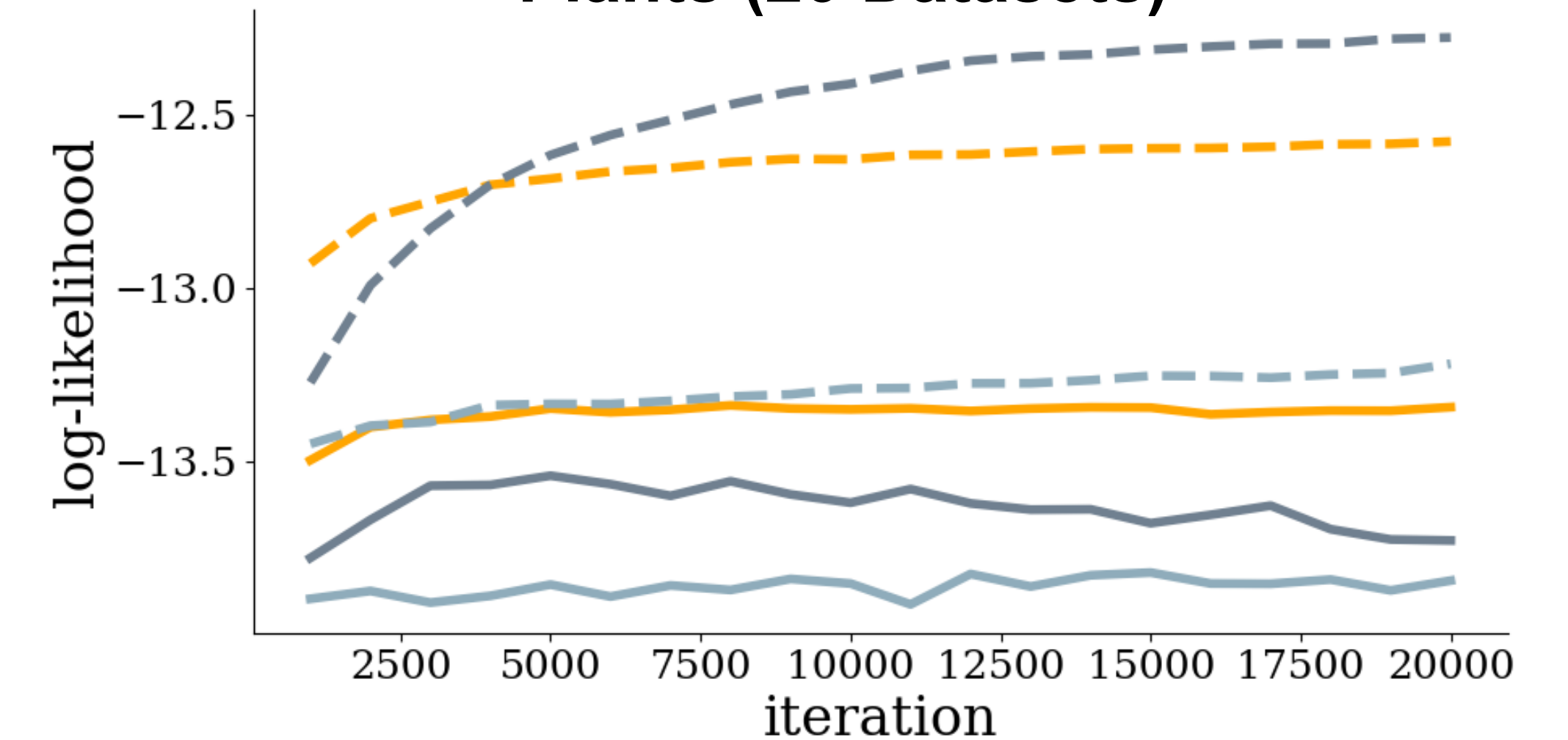
Experiments

Synthetic (Test Log-LL)		
	Log-Likelihood	# Params
SPN-Large	-166.90 ± 0.03	640050
SPN-Small	-167.00 ± 0.01	102450
HyperSPN	-166.32 ± 0.04	129115

HyperSPNs

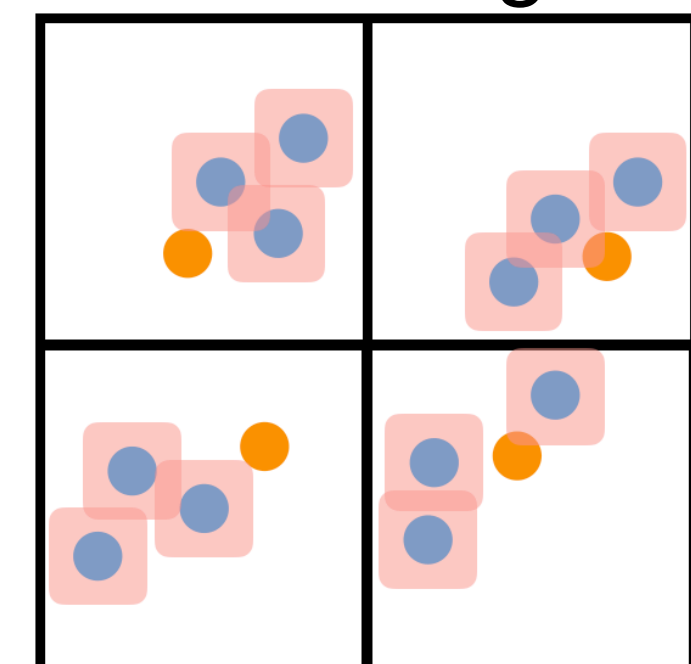


Plants (20 Datasets)



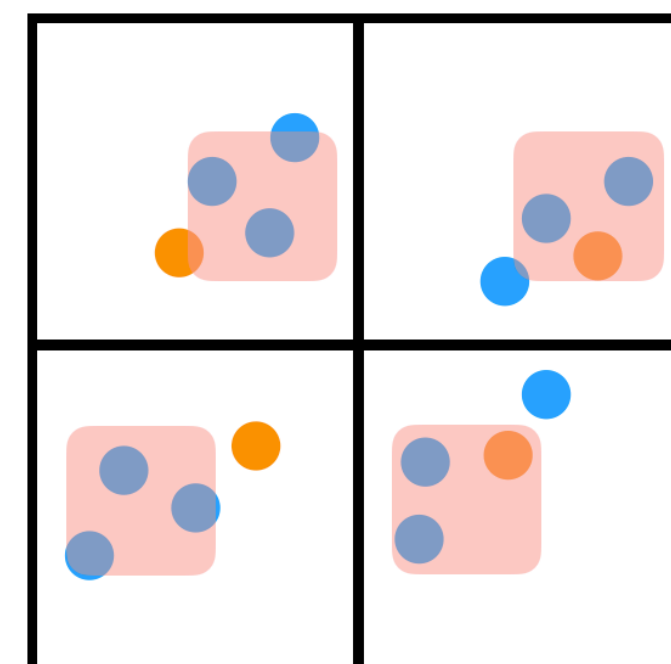
- - spn-large (train) - - spn-small (train)
 — spn-large (valid) — spn-small (valid)
 - - hyperspn (train)
 — hyperspn (valid)

SPN-Large



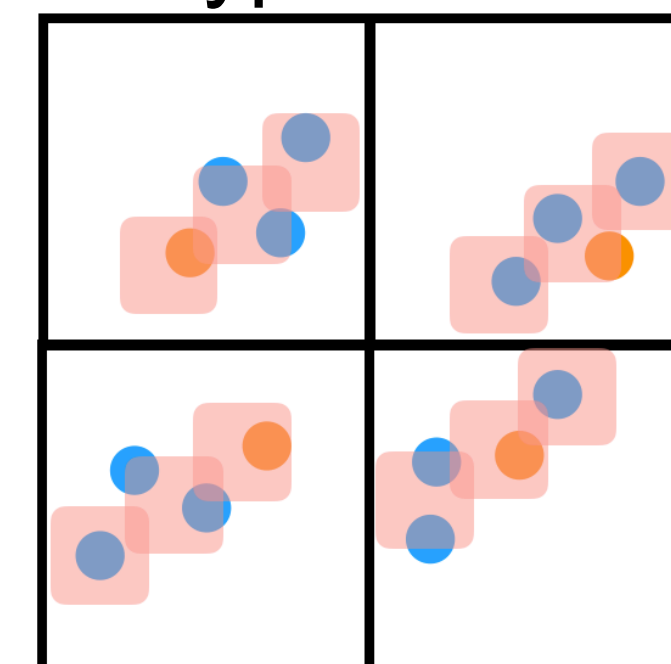
Many clusters
Not constrained ❌

SPN-Small



Few clusters ❌

HyperSPN



Many clusters
Constrained ✅

● train ● test ■ cluster

Better than weight decay on

- 16 / 20 of Twenty Datasets
- 14 / 15 of Amazon Baby Reg.