Symmetry breaking in neural network training

Keegan Stoner
Northeastern University

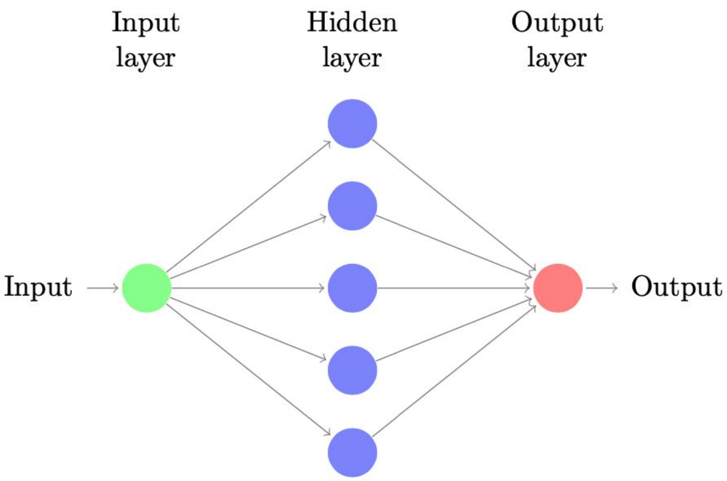
string_data
2020

NN-QFT Correspondence

Based on work with Jim Halverson and Anindita Maiti - arxiv:2008.08601

GP / asymptotic NN	Free QFT
inputs $(x_1, \ldots, x_k)$	external space or spacetime points
kernel $K(x_1, x_2)$	Feynman propagator
asymptotic NN $f(x)$	free field
log-likelihood	free action S_{GP}

$$f_{\theta,N} : \mathbb{R}^{d_{\mathrm{in}}} \to \mathbb{R}^{d_{\mathrm{out}}},$$



$$\sigma(x) = \frac{\exp\left(W\,x + b\right)}{\sqrt{\exp\left[2(\sigma_b^2 + \frac{\sigma_W^2}{d_{\mathrm{in}}} x^2)\right]}},$$

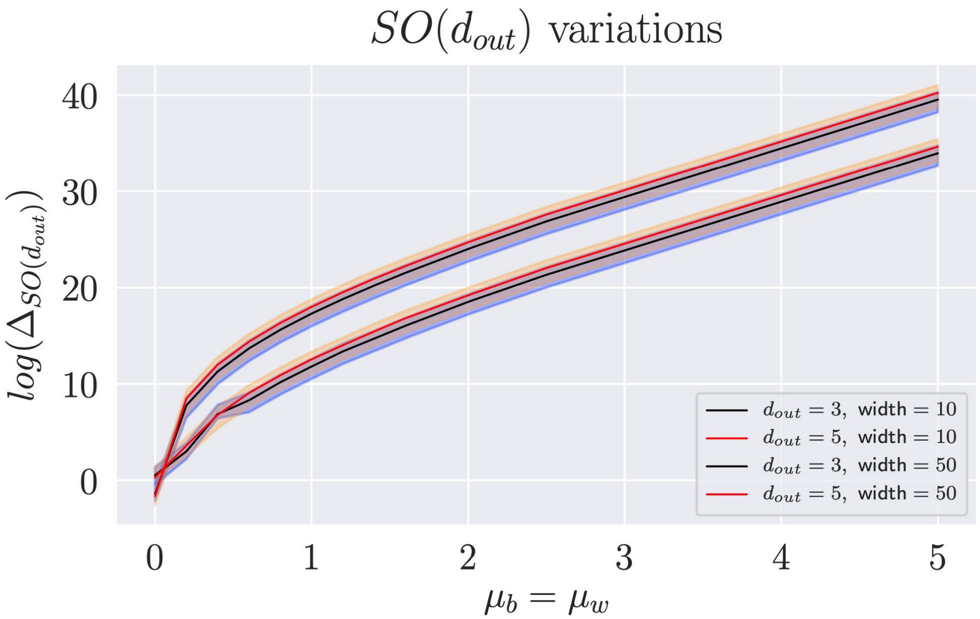
$$K_{\mathrm{Gauss}}(x, x') = \sigma_b^2 + \sigma_W^2\ \exp\left[-\frac{\sigma_W^2 |x - x'|^2}{2 d_{\mathrm{in}}}\right],$$

Symmetries at initialization

When $d_{out} > 1$, output components are uncorrelated.

Initialized network outputs have rotational symmetry when parameters are sampled with **mean zero.**

As the parameter mean increases, the *SO(d*

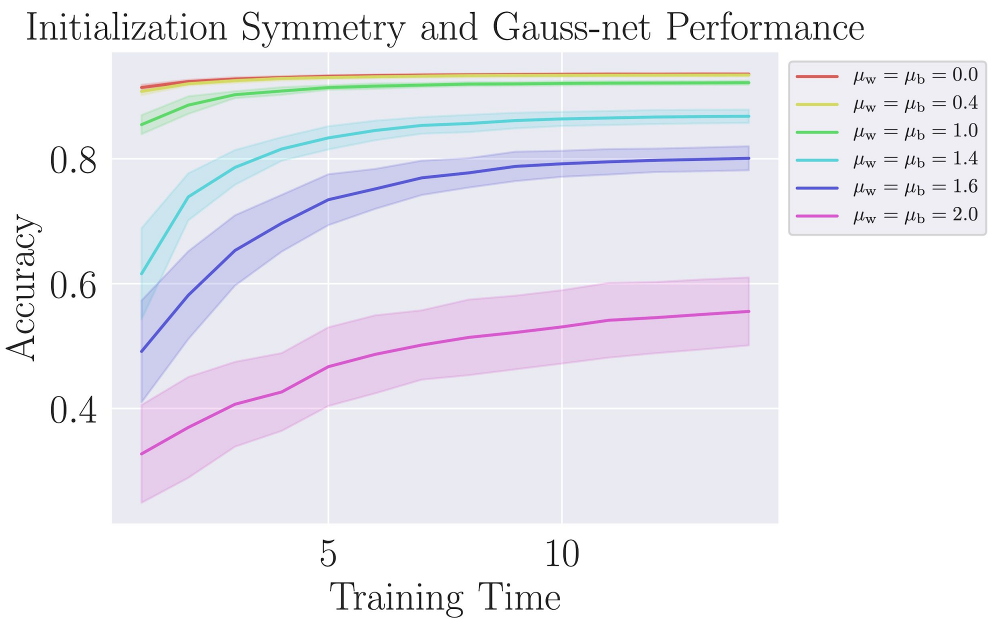


MNIST performance

MNIST training accuracy reflects real-life ability to learn. **Gauss-net can learn MNIST to ~97% accuracy!**

For MNIST, d_{out} = 10. Performed 20 experiments each with $N = 10$.

Idea: Networks with more symmetry at initialization train better.*



*Preliminary results from Gauss-net. How well this statement generalizes is the subject of active research.

Thank you!