

# Avoiding Pathologies in Very Deep Networks

David Duvenaud, Oren Rippel, Ryan Adams, Zoubin Ghahramani

## Nonparametric Priors on Deep Nets

Deep GPs are compositions of functions, each  $f^{(\ell)} \stackrel{\text{ind}}{\sim} \mathcal{GP}(0, k(\mathbf{x}, \mathbf{x}'))$ .

$$\mathbf{f}^{(1:L)}(\mathbf{x}) = \mathbf{f}^{(L)}(\mathbf{f}^{(L-1)}(\dots \mathbf{f}^{(2)}(\mathbf{f}^{(1)}(\mathbf{x})) \dots))$$

Can be derived as either

1. neural nets with nonparametric activation functions
2. neural nets with infinitely-many parametric hidden nodes

