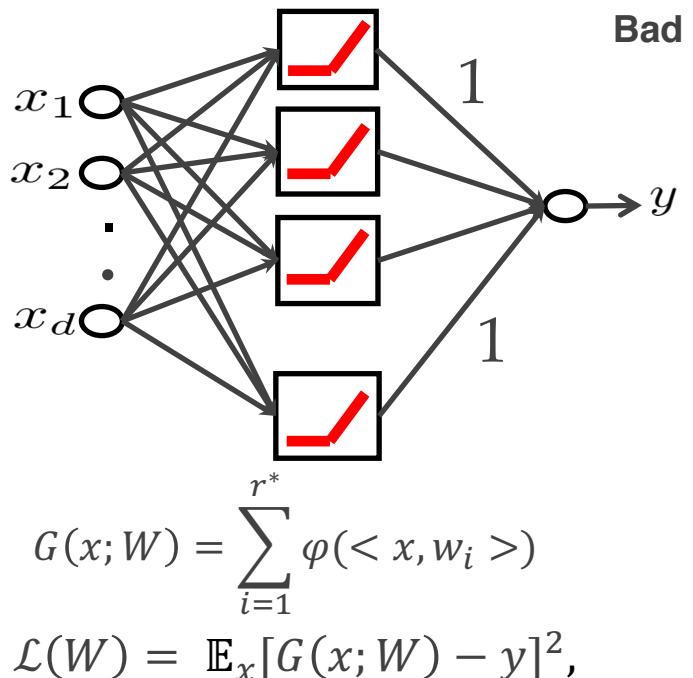


Porcupine Neural Networks: Approximating Neural Network Landscapes

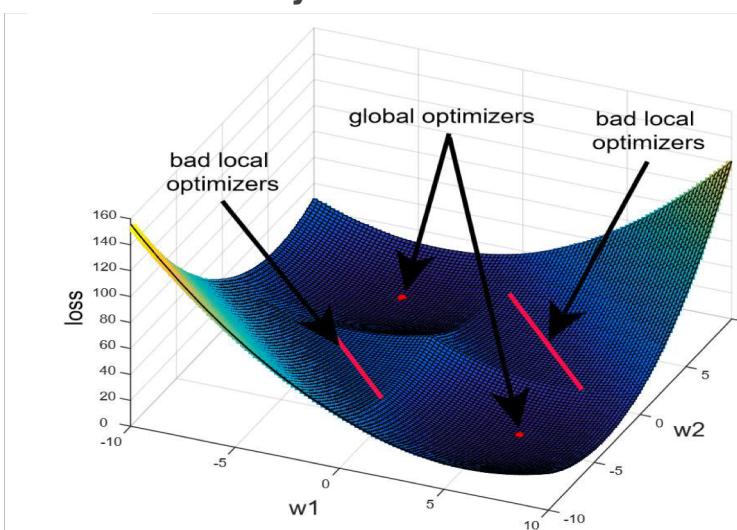
Soheil Feizi¹, Hamid Javadi², Jesse Zhang³, David Tse³

1: University of Maryland, College Park 2: Rice University 3: Stanford University

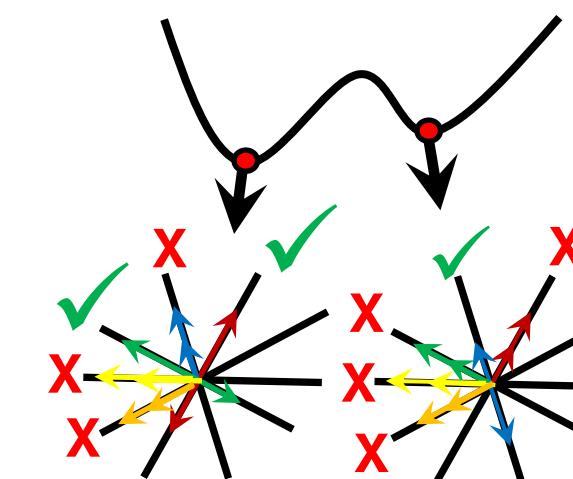
Optimizing Two-layer ReLU Network



Bad local minima may exist!

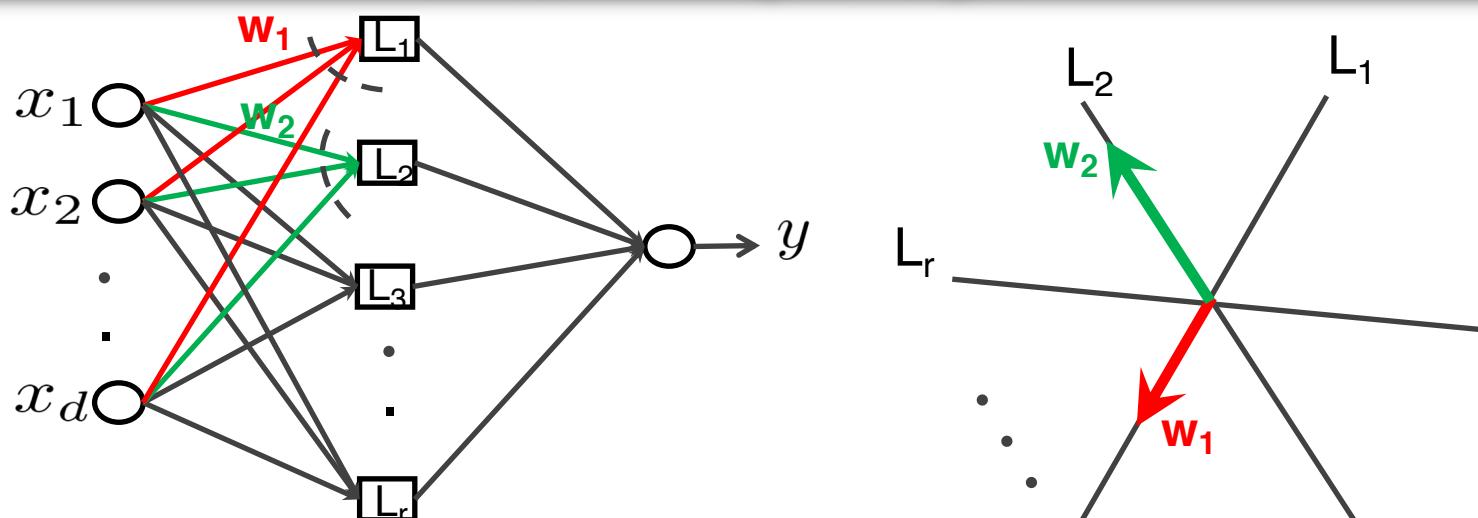


Optimization Landscapes of PNNs



Theorem: Local optima with at least d good lines are global optima.

Idea: Constraining the Weights



- ❖ PNNs have good optimization landscapes: (most) Local optima = Global optima
- ❖ PNNs have good approximation power

Approximation Power of PNNs

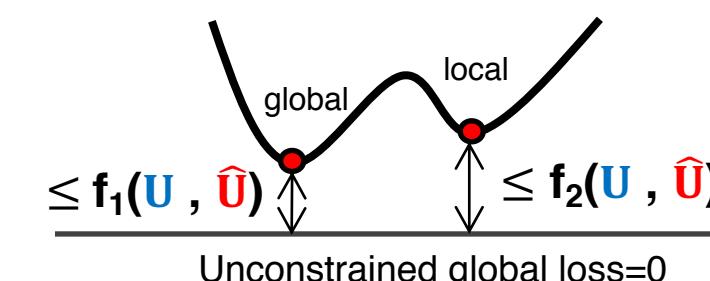
True unit norm vectors: $\mathbf{U}=(\mathbf{u}_1, \dots, \mathbf{u}_{r^*})$

PNN unit norm vectors: $\widehat{\mathbf{U}}=(\widehat{\mathbf{u}}_1, \dots, \widehat{\mathbf{u}}_r)$

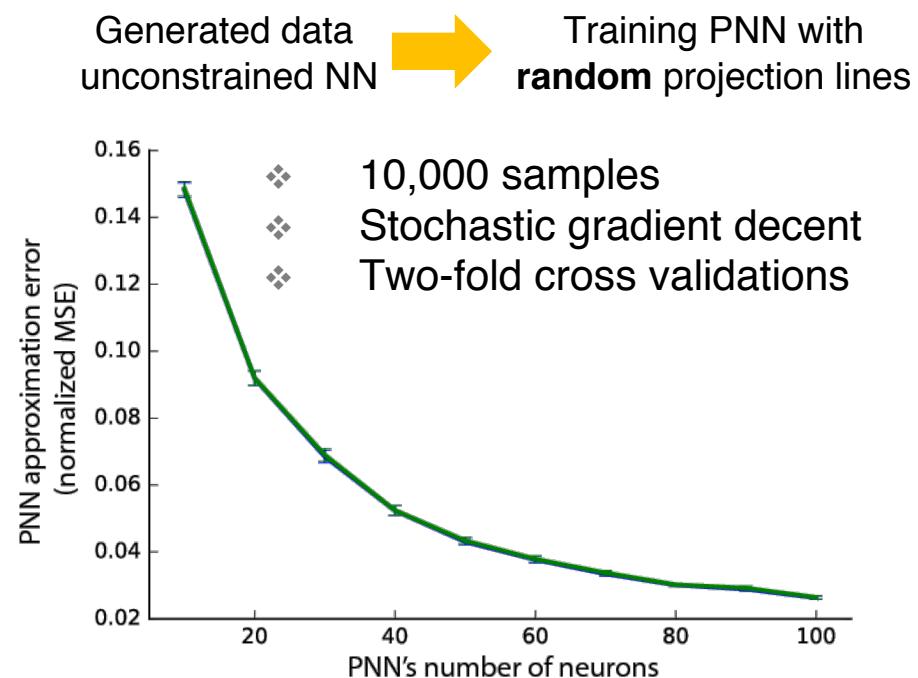
$$\Psi(K) = \begin{pmatrix} \Psi(\widehat{\mathbf{U}}^t \widehat{\mathbf{U}}) & \Psi(\widehat{\mathbf{U}}^t \mathbf{U}) \\ \Psi(\mathbf{U}^t \widehat{\mathbf{U}}) & \Psi(\mathbf{U}^t \mathbf{U}) \end{pmatrix}$$

Theorem: $f_1(\mathbf{U}, \widehat{\mathbf{U}}) = \|\Psi(K)/\Psi(K_{11})\|$

Similar (slightly worst) bound for f_2



PNNs in Practice



Asymptotic Analysis

High dimensional regime

\mathbf{U} and $\widehat{\mathbf{U}}$ uniformly random

$d, r \rightarrow \infty, \gamma = r/d = \text{fixed}$

$$f_1(\mathbf{U}, \widehat{\mathbf{U}}) \rightarrow \left(1 + \frac{r^*}{r}\right) \left(1 - \frac{2}{\pi}\right)$$

