Neural Network Approximations Of PDEs Beyond Linearity: A Representational Perspective

Tanya Marwah[1], Zachary C. Lipton[1], Jianfeng Lu[2], Andrej Risteski[1]

[1]Machine Learning Department, Carnegie Mellon University
[2]Department of Mathematics, Duke University

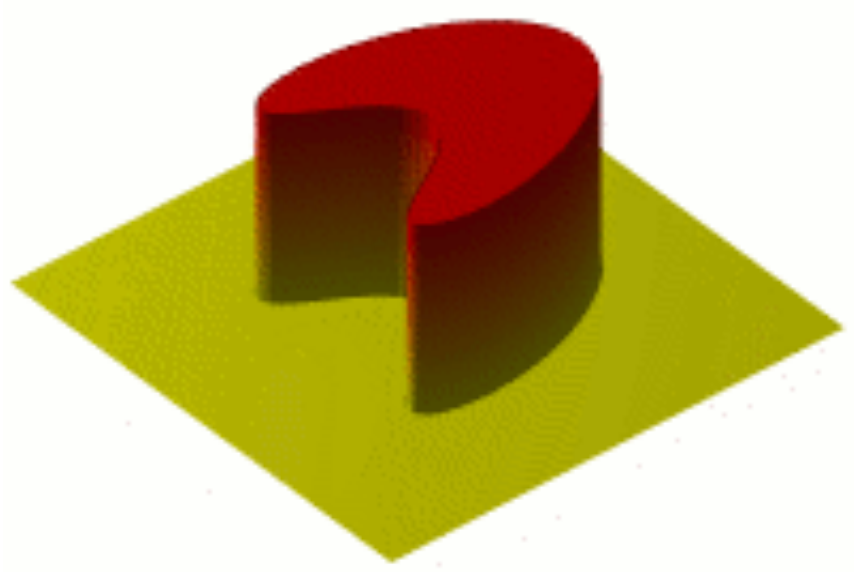
ICML International Conference On Machine Learning

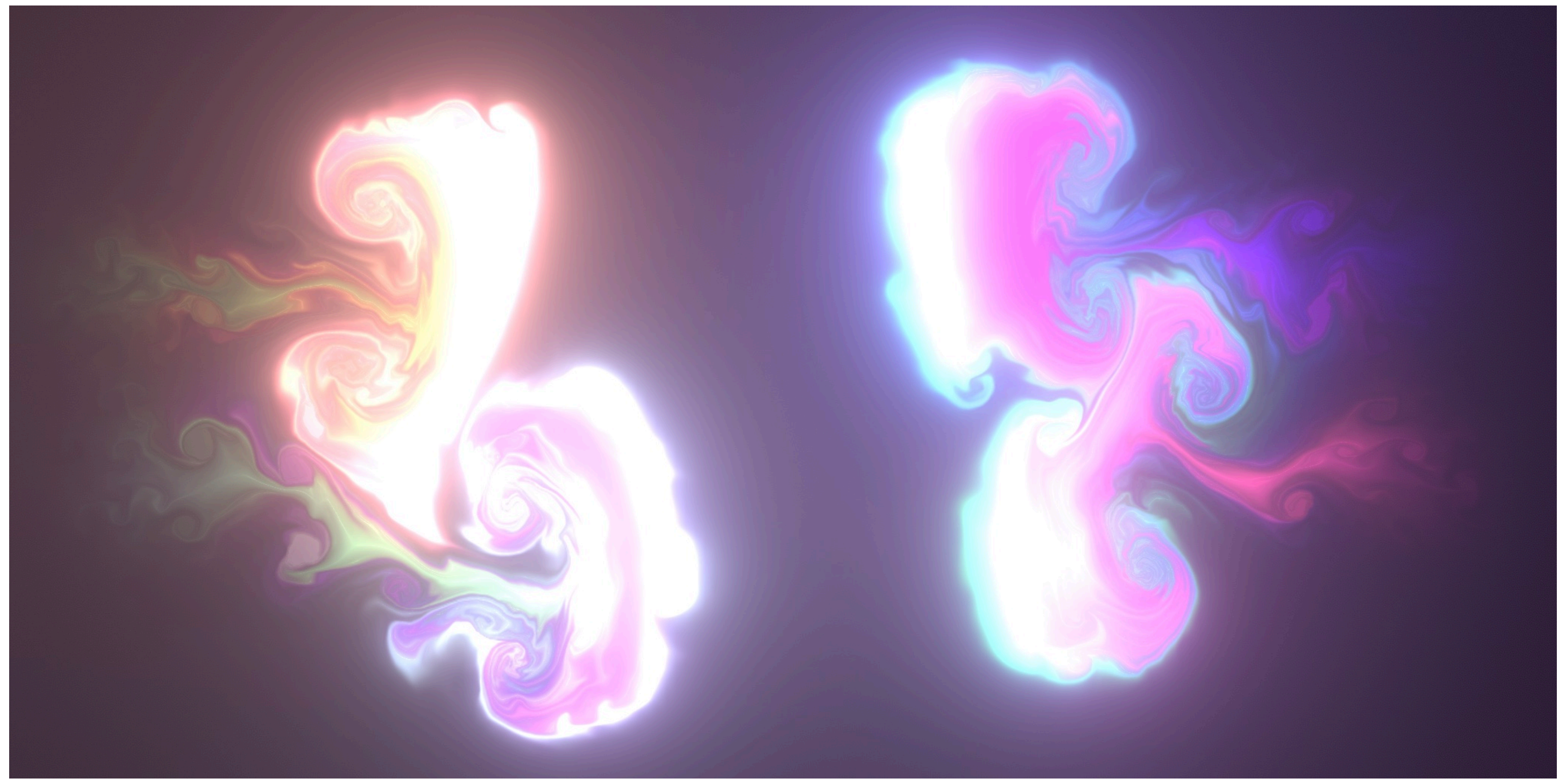
Carnegie Mellon University

Duke University

Partial Differential Equations

A partial differential equation (PDE) relates a multivariate function defined over some domain to its partial derivates.





Solving PDEs using Neural Networks

Why Neural networks?

- Mesh free. Trained using Monte-Carlo approximation of the variational formulation (E and Yu. 2017)
- Can be used to approximate solutions to entire families of PDEs (Li et al. 2020)

Theoretical Analysis:

For what families of PDEs, can the solution be represented by a small neural network?

Previous work (Marwah et al (2021), and Chen at al (2021)) show that for *linear elliptic PDEs* approximating neural networks have a polynomial dependence on input dimension.

We extend these works to a family of nonlinear PDEs

Nonlinear Elliptic PDE

For a strongly convex function $L : \Omega \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$, we first define an Energy functional:

$$\mathcal{E}(u) = \int_\Omega \Big(L(x, u(x), \nabla u(x)) - f(x)u(x) \Big) dx.$$

Nonlinear Elliptic PDE:

$$D\mathcal{E}(u) := -\operatorname{div}(\partial_{\nabla u} L(x, u, \nabla u)) + \partial_u L(x, u, \nabla u) = f$$

for all $x \in \Omega$ and $u(x) = 0$ for all $x \in \partial\Omega$.

Here div denotes the divergence operator: Given a vector field $F : \mathbb{R}^d \to \mathbb{R}^d$, $\operatorname{div} F = \nabla \cdot F = \sum_{i=1}^{d} \frac{\partial F_i}{\partial x_i}$

Barron Norm

For $f : [0,1]^d \to \mathbb{R}$, Fourier transform:

$$\hat{f}(\omega) = \int_{[0,1]^d} f(x) e^{-i2\pi x^T \omega} dx, \quad \omega \in \mathbb{N}^d.$$

Then the **Barron Norm** ($\|\cdot\|_{\mathscr{B}}$) of the function f is defined as:

$$\|f\|_{\mathscr{B}} = \sum_{\omega \in \mathbb{N}^d} (1 + \|\omega\|_2) \, |\hat{f}(\omega)| \, .$$

*Classical result from **Barron (1993)** shows that if $\|f\|_{\mathscr{B}}$ is bounded then the function f can be represented by a $\epsilon-$approximated by a two layer neural network with $O\left(\frac{\|f\|_{\mathscr{B}}^2}{\epsilon}\right)$ parameters.*

Our Result

If composing a function with Barron norm b with partial derivatives of L produces a function of Barron norm at most $B_L b^p$, then solution to the nonlinear elliptic PDE can be $\epsilon-$approximated in the L^2 sense by a function with Barron norm

$$O\left(\left(dB_L\right)^{\max\{p\log(1/\epsilon), p^{\log(1/\epsilon)}\}}\right).$$

<u>Proof Sketch:</u>

Neurally simulating a preconditioned gradient descent (for a strongly-convex loss) in an appropriate Hilbert space and bounding the growth of the Barron norm of each iterate as well as bounding the total iterations.

Thank You!