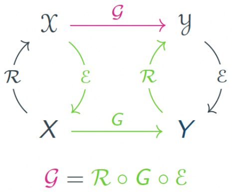
AI in the Sciences and Engineering 2024: Lecture 15

Siddhartha Mishra

Seminar for Applied Mathematics (SAM), D-MATH (and),
ETH AI Center,
ETH Zürich, Switzerland.

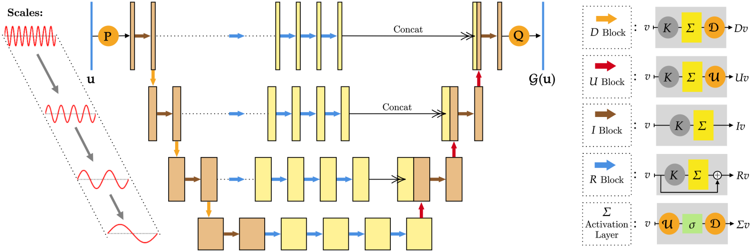
What you learnt so far

- Operator learning: Given Abstract PDE: $\mathcal{D}_a(u) = f$
- Learn Solution Operator: $\mathcal{G} : \mathcal{X} \mapsto \mathcal{Y}$ with $\mathcal{G}(a, f) = u$
- Enforce Continuous-Discrete Equivalence via ReNO:



- Neither CNN nor FNO are ReNOs.
- SNO/DeepONet can be ReNOs but perform poorly !!
- CNO is ReNO that works

CNO



- CNO instantiated as a modified Operator UNet
- Built for multiscale information processing

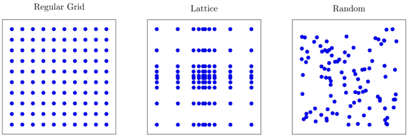
Empirical Results

- Extensive Empirical evaluation on RPB benchmarks.

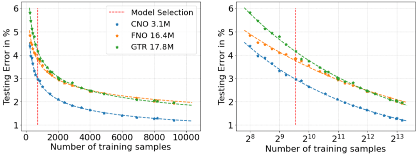
	In/Out	FFNN	GT	UNet	ResNet	DON	FNO	CNO
Poisson Equation	In	5.74%	2.77%	0.71%	0.43%	12.92%	4.98%	**0.21%**
	Out	5.35%	2.84%	1.27%	1.10%	9.15%	7.05%	**0.27%**
Wave Equation	In	2.51%	1,44%	1.51%	0.79%	2.26%	1.02%	**0.63%**
	Out	3.01%	1.79%	2.03%	1.36%	2.83%	1.77%	**1.17%**
Smooth Transport	In	7.09%	0.98%	0.49%	0.39%	1.14%	0.28%	**0.24%**
	Out	650.6%	875.4%	1.28%	0.96%	157.2%	3.90%	**0.46%**
Discontinuous Transport	In	13.0%	1.55%	1.31%	**1.01%**	5.78%	1.15%	**1.01%**
	Out	257.3%	22691.1%	1.35%	1.16%	117.1%	2.89%	**1.09%**
Allen-Cahn Equation	In	18.27%	0.77%	0.82%	1.40%	13.63%	**0.28%**	0.54%
	Out	46.93%	2.90%	2.18%	3.74%	19.86%	**1.10%**	2.23%
Navier-Stokes Equations	In	8.05%	4.14%	3.54%	3.69%	11.64%	3.57%	**2.76%**
	Out	16.12%	11.09%	10.93%	9.68%	15.05%	9.58%	**7.04%**
Darcy Flow	In	2.14%	0.86%	0.54%	0.42%	1.13%	0.80%	**0.38%**
	Out	2.23%	1.17%	0.64%	0.60%	1.61%	1.11%	**0.50%**
Compressible Euler	In	0.78%	2.09%	0.38%	1.70%	1.93%	0.44%	**0.35%**
	Out	1.34%	2.94%	0.76%	2.06%	2.88%	0.69%	**0.59%**

CNO/FNO: Issues

- Data on Non-uniform Grids.



- Time-dependent problems
- Scaling with data



Time-dependent PDEs

- Of the Abstract form:

$$u_t + \mathcal{L}(t, x, u) = 0, \quad u(0) = \bar{u}.$$

- Solution operator: $\mathcal{S} : (0, T) \times \mathcal{X} \mapsto \mathcal{X}$; $\mathcal{S}(t, \bar{u}) = u(t)$
- Fo any time increment: $\mathcal{S}(\Delta t, u(t)) = u(t + \Delta t)$.
- Generated data is the form of Trajectories:

$$(u(0), u(t_1), u(t_2), \dots, u(T)) = (\bar{u}, \mathcal{S}(t_1, \bar{u}), \mathcal{S}(t_2, \bar{u}), \dots, u(T))$$
$$= (\bar{u}, \mathcal{S}(t_1, \bar{u}), \mathcal{S}(t_2 - t_1, u(t_1)), \dots, u(T))$$

- Learning Task:
- Given $\bar{u}$ + BC: generate the solution trajectory $u(t)$, for all $t \in (0, T]$

Neural Operators for Time-dependent PDEs

- Direct Evaluation with FNO/CNO.
- $\mathrm{NO}_k : \bar{u} \mapsto \mathrm{NO}_k(\bar{u}) \approx \mathcal{S}(k\Delta t, \bar{u})$
- Lot of compute as K-different NOs need to be trained.
- Only evaluation at discrete time levels

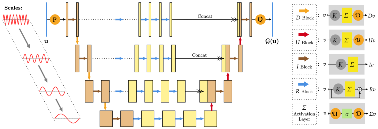
Autoregressive Evaluation

- Assume Trajectory data on uniformly spaced timepoints: $u(t_k) = u(k\Delta t)$.
- Define $\mathrm{NO}_{\Delta t}(u(t_\ell)) \approx u(t_\ell + \Delta t)$
- Then Autoregressive Rollout is

$$u(t_k) \approx \underbrace{\mathrm{NO}_{\Delta t} \circ \ldots \mathrm{NO}_{\Delta t} \circ \mathrm{NO}_{\Delta t}}_{k \text{ times}} \bar{u}.$$

- Issues:
  - Needs uniform spacing.
  - Long rollouts lead to training issues.
  - Error Accumulation
  - Only evaluation at discrete time levels

Time Conditioning



- Lead Time as an Input Channel
- $\mathrm{CNO}\ (\bar{t}, u(t)) \approx \mathcal{S}(\bar{t}, u(t)) = u(t + \bar{t})$.
- Add Conditional Normalizations after each layer !!

$$\mathcal{N}(w) = g_N(t) \odot \frac{w - \mathbb{E}(w)}{\sqrt{\mathrm{Var}\ (w) + \epsilon}} + h_N(t),$$

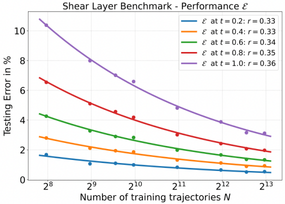
- g_N, h_N are MLPs in general.
- Instance,Batch,Layer Normalizations.

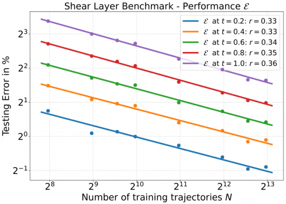
Training Strategies

- One at a Time training based on:
- Input-Target Pairs: $\bar{u}, \mathcal{S}(t_k, \bar{u}) = u(t_k)$
- For $t_K = T$, K training samples per trajectory.
- all2all training based on:
- Input-Target Pairs: $u(t_i), \mathcal{S}(t_j - t_i, u(t_i)) = u(t_j)$, $\forall i < j$
- $\frac{K^2+K}{2}$ training samples per trajectory !!
- Inference is Direct or Autoregressive
- Multiple possibilities for Autoregressive Rollouts
- Evaluation at any time $t > 0$ including Out-of-distribution times.

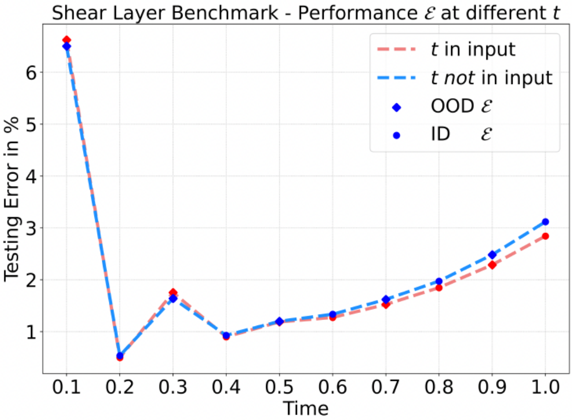
Results for Shear Layer

Error vs. Time

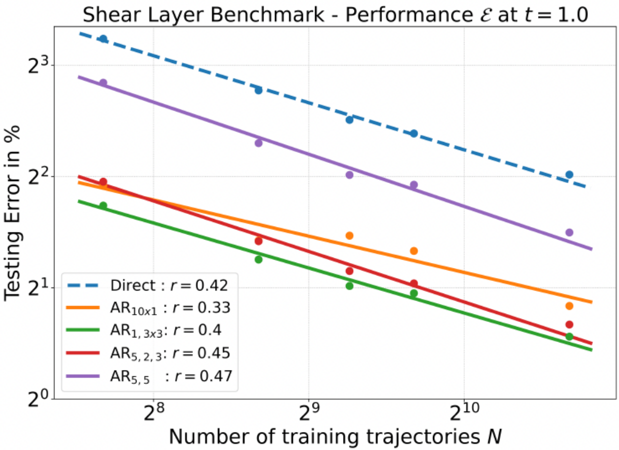




Results at OOD time levels.



Results for Different Strategies I.



Results for Different Strategies II.