Learning semilinear **N**eural **O**perators: A unified recursive framework for prediction and **D**ata **A**ssimilation.

Ashutosh Singh, Ricardo Augusto Borsoi*, Deniz Erdogmus, Tales Imbiriba*

-ICLR 2024

Northeastern University
College *of* Engineering

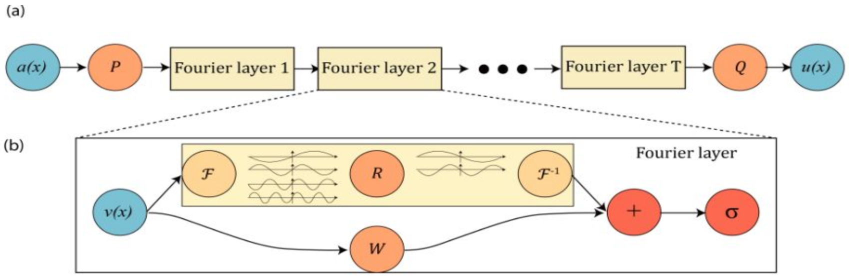
cnrs

UNIVERSITÉ
DE LORRAINE

This work was supported by NIH OT2OD030524, NSF 2316420, ANR23-CE23-0024, ANR-23-CE94-0001

Parametric PDEs and Neural Operator

- Many problems in science and engineering involve solving complex partial differential equation (PDE) systems.
- Examples: molecular dynamics, micro-mechanics, turbulent flows, weather fluctuations etc.
- Traditional solvers : slow and do not scale to large systems well.
- Neural Operators : fast, scalable and resolution invariant.



Parametric PDE :

$$(L_a z)(x,t) = f(x,t)\,,$$
$$z(x,t) = 0\,,$$

Parametric mapping to approximate the solution of the PDE:

$$\mathcal{G}_\theta : a \mapsto z$$
$$\mathcal{G}(a) = (\mathcal{Q} \circ \mathcal{W}_L \circ \cdots \circ \mathcal{W}_1 \circ \mathcal{P})(a)$$

Optimisation to solve :

$$\min_\theta \frac{1}{N} \sum_{i=1}^{N} \|z_i - \mathcal{G}_\theta(a_i)\|_{\mathcal{Z}}^2$$

Challenges:

International Conference On Learning Representations

- Real world system - irregular measurement data, noisy measurement etc.
- Vanilla NO doesn't maintain history or have a well motivated recurrent structure.
- NO drawbacks : error accumulation for long rollouts and sensitivity to noise in the data.
- NO is incapable of making use of future measurements.
- NO literature lacks a systematic framework for data assimilation.
- Infinite dimensional Kalman filters : estimate the state of a system governed by some underlying infinite-dimensional PDEs based on measurements.
- Kalman filters based observers requires tremendous computational resources, limiting their application to large-scale systems.
- In this paper we propose a general framework for data assimilation for NO by exploiting the structure of semilinear PDEs .

Main Contribution:

- We extend the NO theory by leveraging the observer design of semilinear PDEs.
- We break the observer solution into prediction and update steps.
- The resulting recurrent framework can estimate solutions using arbitrary amounts of time measurements.

Semilinear PDE system and observer design:

- System equation based on semilinear evolution:

$$\begin{aligned} \frac{\partial z(t)}{\partial t} &= Az(t) + G(z(t), t) + \omega(t)\,, \\ z(0) &= z_0\,, \\ y(t) &= Cz(t) + \eta(t)\,. \end{aligned}$$

- Observer design:

$$\frac{\partial \hat{z}(t)}{\partial t} = A\hat{z}(t) + G(\hat{z}(t), t) + K(t)\big[y(t) - C\hat{z}(t)\big]$$

$$\hat{z}(t_k) = \underbrace{T(t_k - t_{k-1})\hat{z}(t_{k-1}) + \int_{t_{k-1}}^{t_k} T(t_k - s)G(\hat{z}(s), s)ds}_{\text{Prediction}} + \underbrace{\int_{t_{k-1}}^{t_k} T(t_k - s)K(s)\big[y(s) - C\hat{z}(s)\big]ds}_{\text{Correction}}$$

NODA framework:

- NODA prediction and update step,

$$\hat{z}_{t_k}^{\text{pred}} = \hat{z}_{t_{k-1}} + \mathcal{W}(\hat{z}_{t_{k-1}})$$

Prediction step

$$\hat{z}_{t_k} = \hat{z}_{t_k}^{\text{pred}} + K(\hat{z}_{t_k}^{\text{pred}})\left[y(t_k) - E(\hat{z}_{t_k}^{\text{pred}})\right]$$

Correction step

- Kalman gain,

$$K(\hat{z}_{t_k}^{\text{pred}})[u] = \tanh\left(W_z E(\hat{z}_{t_k}^{\text{pred}}) + W_y y(t_k) + b\right) \odot \left(\hat{C}^* u\right)$$

- Training loss :

$$\mathcal{J}(\phi) = \frac{1}{SN}\sum_{i=1}^{S}\sum_{k=1}^{N}\left\|\hat{z}_{t_k}^{(i)} - z_D^{(i)}(t_k)\right\|_2 + \frac{\lambda}{SH}\sum_{i=1}^{S}\sum_{k=1}^{H}\left\|y^{(i)}(t_k) - E(\hat{z}_{t_k}^{(i)})\right\|_2$$

Results:

Table 2: Averaged RelMSE ($\times 10^3$) for prediction ($\alpha = 0\%$) on the Navier-Stokes equation as a function of the sequence length t_f and of the SNR.

SNR	20 dB			30 dB			∞		
t_f	500	750	1000	500	750	1000	500	750	1000
MNO	48±7	72±6	96±12	39±4	51±7	87±6	18±3	22±3	26±5
FNO	192±10	294±12	351±14	189±19	265±21	317±16	136±11	177±14	265±16
MWNO	172±10	226±9	278±11	147±9	192±9	240±10	78±5	104±6	188±11
C-LSTM	487±15	522±10	604±18	397±20	426±17	481±19	366±11	417±13	529±11
NODA	**10**±2	**18**±4	**32**±6	**8**±1	**13**±2	**26**±4	**7**±1	**8**±1	**13**±2

Table 3: Averaged RelMSE ($\times 10^3$) of NODA for the Navier-Stokes equation as a function of α.

α	0%	10%	20%	30%
Averaged RelMSE	26±4	18±4	13±3	9±3

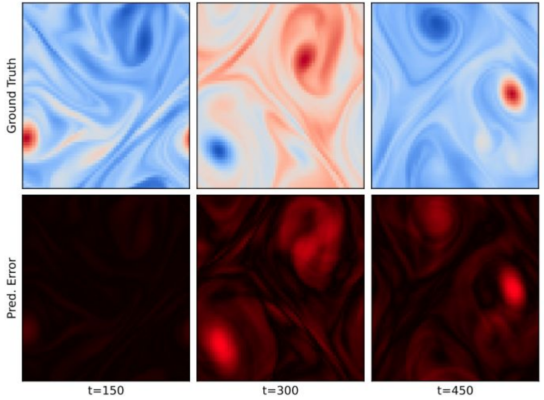


Figure 1: Samples of a realization of a true trajectory of the Navier-Stokes equation for $t \in \{150, 300, 450\}$ (**Top Row**), and elementwise error plots for the corresponding predictions by NODA (**Bottom Row**).

Thank you for listening!