Model-operator fusion for scientific machine learning

Tushar*
Department of Applied Mechanics
Indian Institute of Technology Delhi, Hauz Khas, New Delhi - 110016, India.
amz218314.iitd@gmail.com

Souvik Chakraborty
Department of Applied Mechanics
Yardi School of Artificial Intelligence (ScAI)
Indian Institute of Technology Delhi, Hauz Khas - 110016, India.
souvik@am.iitd.ac.in

Abstract

The governing physics in science and engineering is often based on assumptions and approximations, leading to analyses and designs that are also approximate. While data-driven machine learning models have emerged to address this issue, they often suffer from (a) lack of interpretability, (b) reliance on large datasets, and (c) poor generalization beyond their training domain. Although operator learning has shown promise, it still faces challenges due to its purely data-driven nature. A suitable alternative lies in data-physics fusion, where data-driven models are used to correct the low-fidelity physics-based models. In this context, we introduce a novel framework, "Model operator fusion for scientific machine learning", designed for solving parametric partial differential equations (PDEs). The proposed framework integrates differentiable physics with the recently developed Wavelet Neural Operator (WNO), combining WNO's data-driven learning capabilities with the interpretability and generalization of physics-based solvers. We demonstrate the effectiveness of this approach by solving parametric PDEs through few benchmark examples from diverse scientific and engineering fields. The results clearly highlight the efficacy of the proposed method.

1 INTRODUCTION

Physical systems are governed by the laws of physics, which are often expressed as Partial Differential Equations (PDEs). The study of PDEs is well-established, with techniques such as Finite Element Methods (FEM), Finite Difference Methods (FDM), and Finite Volume Methods (FVM) widely accessible. However, these established physical laws are frequently derived from specific assumptions and approximations. An alternative approach involves utilizing data-driven methods where recent advancements in neural operators [1] have demonstrated their effectiveness in learning complex nonlinear partial differential equations (PDEs). However, purely data-driven models have certain limitations: (a) they often lack interpretability, (b) they require large amounts of data, and (c) they may struggle to generalize beyond the training domain. To overcome these challenges, a potential solution resides in data-physics fusion, where data-driven models are employed to learn the missing physics and thus act as model correctors.

In this research work, our objective is to develop a novel hybrid framework that combines neural operators with conventional physics-based models. The proposed framework combines the principles of differentiable physics [2] with the recently introduced Wavelet Neural Operator (WNO) [1]. It

*Corresponding author:amz218314.iitd@gmail.com

1st Conference on Applied AI and Scientific Machine Learning (CASML 2024).

utilises WNO's capacity to learn from data while maintaining the interpretability and generalisation of physics-based solvers. The salient features of the proposed framework are as follows:

- **Physics-WNO Integration:** The proposed framework introduces an innovative approach that integrates the Wavelet Neural Operator (WNO) with a low-fidelity physics model directly within the governing equation.
- **In-Equation Augmentation:** This framework emphasize on augmentation within the governing equation itself. This in-equation augmentation enhances generalization compared to extrusive augmentation methods.
- **End-to-End Training:** The approach supports end-to-end training, which ensures computational efficiency during the training phase.
- **Differentiable Physics:** The framework incorporates a differentiable physics solver [2] into the WNO. This integration allows the model to be trained using the backpropagation algorithm, eliminating the need for numerical approximations during the training process.

2 Proposed Approach

Neural operators can learn the mapping from functional parametric dependencies to the solutions of partial differential equations (PDEs), enabling them to capture the solution operator for an entire family of PDEs from data. However, in a scientific setting, we generally have access to both scientific models (low fidelity) and data (sparse). Therefore, we propose a framework based on the fusion of physics based model and data driven Neural Operators. In this work we utilize recently developed Wavelet Neural operator (WNO) as our operator learning framework. The fundamental idea is to integrate the WNO to the low fidelity physics model, which can be shown as:

$$\frac{\partial \mathbf{u}}{\partial t} = \mathcal{G}_k(\mathbf{u}, \mathbf{u_x}, \mathbf{u_{xx}} \ldots, \mathbf{p}_k) + \mathcal{W}(\mathbf{u}; \boldsymbol{\theta}); \tag{1}$$

where, $\mathcal{G}_k$ represents some general operator representing the known low fidelity PDE and $\boldsymbol{p}$ represents the parameters of the PDE. $\mathcal{W}(\mathbf{u}; \boldsymbol{\theta})$ represents the WNO. The hypothesis here is that WNO will act as the physics corrector. Since we fuse the WNO with the low fidelity physics model we refer to the proposed approach as the model operator fusion (MOF). For using MOF in practice, one needs to train the WNO in Eq. 1. However, we don't have access to the samples/measurements of missing physics. Therefore we employ an end-to-end approach, which involves integration of differentiable physics solver with WNO. A schematic representation of the framework is shown in Fig. 1.

3 Numerical Problems

In this research, we assess the performance of the proposed framework using a series of benchmark problems with multiple cases to evaluate the framework's effectiveness in handling various scenarios and complexities, one of which is shown here. For illustrative purposes, a synthetic low-fidelity physics model was created by omitting certain portions of the PDE. The solutions produced by the proposed framework are compared with the ground truth to assess accuracy as shown in Fig 2. Particularly, we demonstrate the extrapolation and generalization capabilities of the proposed framework. For quantitative assessment, mean squared error (MSE) values were calculated for all benchmark examples, comparing the performance of different models relative to the ground truth, as shown in Table 1.

3.1 One dimensional PDE

In this example, we consider the Nagumo equation, which has various applications including modeling of wave propagation in neurons. Specifically, it is used to study the dynamics of voltage across nerve cells and the impulses in nerve fibers. The Nagumo equation incorporating periodic boundary

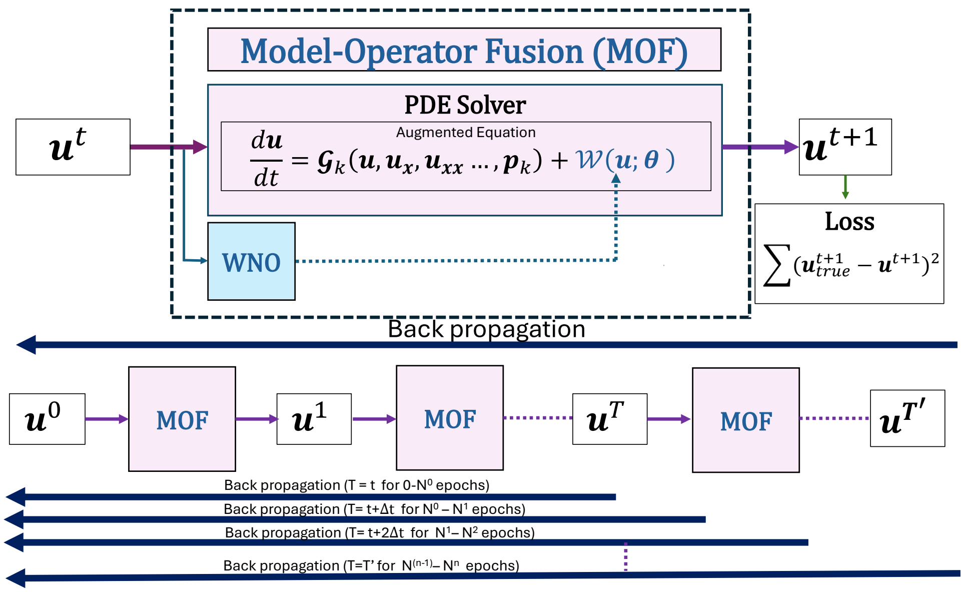

Figure 1: Schematic representation of the framewok: The two major components are the the WNO block and the PDE solver. The WNO when additively augmented with the known PDE within the equation form, is trained to learn the missing physics and thus acts as the physics corrector. The augmented PDE is then solved using the numerical solver to get the corrected output. The framework is trained by progressively unrolling in time as shown.

Method	Cases	Proposed (Er)	WNO only (Er)	Known physics (Er)
Nagumo	Missing Cubic	0.1873	35.685	1.1312
	Missing Diffusion	0.0042	34.2061	0.2310
Allen Cahn	Only Diffusion	0.1583	72.033	2.6089
	Missing Diffusion	0.0011	28.6927	0.1046
Smoke-plume	Inexact Buoyancy	0.00399	0.3605	0.3881

Table 1: The prediction errors (Er: Mean square error (MSE)) for the different examples obtained using the proposed approach, WNO only (data-driven WNO), and known physics.

condition is expressed as follows:

$$\begin{aligned}
&\frac{\partial u}{\partial t} - \epsilon\frac{\partial^2 u}{\partial x^2} = u(1-u)(u-\alpha), \quad x \in (0,1), t \in (0,T)\\
&u(x=0,t) = u(x=1,t), \quad x \in (0,1), t \in (0,T)\\
&u_x(x=0,t) = u_x(x=1,t), \quad x \in (0,1), t \in (0,T)\\
&u(x,t=0) = u_0(x), \quad x \in (0,1)
\end{aligned} \tag{2}$$

where, the parameter $\alpha \in \mathbb{R}$ determines the wave's speed along the axon, while $\epsilon > 0$ controls the diffusion rate.

3.1.1 Missing cubic term in Nagumo equation

First, we examine the scenario where only the diffusion term is present in the known physics, while the cubic term is absent. The equation of known physics takes the following form:

$$\frac{\partial u}{\partial t} - \epsilon\frac{\partial^2 u}{\partial x^2} = 0, \quad x \in (0,1), t \in (0,T) \tag{3}$$

The initial and boundary conditions remain the same as in Eq. 2.

In this case, we have considered $\epsilon = 0.0001$ and $\alpha = -10$ to emphasize the disparity between partial physics and complete physics. The objective is to employ the proposed approach as a surrogate to learn the temporal evolution from the stochastic initial condition to the output.

During the testing phase, a total of 100 initial conditions were randomly selected for evaluation. The initial conditions are of the form $\boldsymbol{u}_0(x) = \alpha \sin(\eta\pi x)$. The parameter α is sampled from the uniform distributions $\mathcal{U}(-10, 10)$ (different from the training data) and the parameter η, can take values of $\{2, 3, 4, 5\}$ (same as training)

The accurate predictions in Fig. 2 demonstrate the extrapolative ability of the proposed framework, extending to twice the training window. Additionally, since the three distinct initial conditions were not included in the training data, this highlights the model's strong generalization capability.

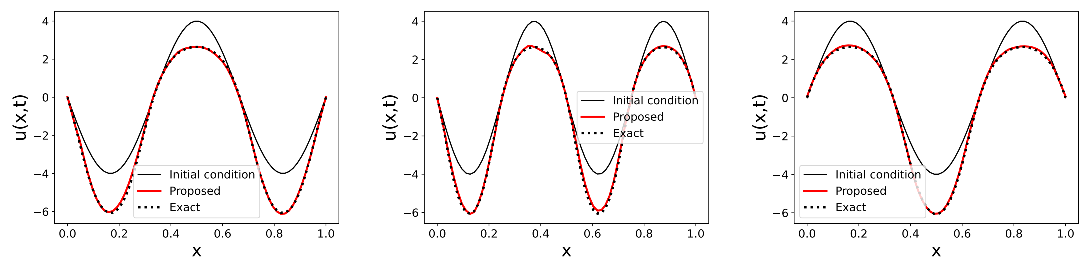

Figure 2: Comparison of predictions from the proposed approach with the ground truth for the Nagumo equation for the case with missing cubic term in the known equation, tested across three different initial conditions (three columns), with $t > t_{train}$ to demonstrate the extrapolative capabilities

3.1.2 Out-of -distribution generalization

In this section, we assess the out-of-distribution generalization of the proposed algorithm. To that end, we considered initial conditions generated from Gaussian process with three different kernels: Radial Basis Function (RBF), Exponential Sine Squared, and Matern. The results obtained are shown in Fig. 3. We observe that the proposed approach easily generalizes to out-of-distribution input without any further training. This illustrates that model operator fusion allows the model to generalize to out-of-distribution cases. To further strengthen our claim, the mean-squared error for the three cases are reported in Table 2, offering insights into the accuracy of the results obtained.

Exponential Sine Squared	**RBF**	**Matérn**
0.0220	0.0444	0.0418

Table 2: The prediction error (Mean Square Error (MSE)) using different kernels obtained using the proposed approach. MSE is computed by taking all 100 test samples up to 150 time steps for all the kernels.

3.2 Two dimensional PDE

In this case we consider a smoke plume simulation by solving the Navier Stokes and transport equation along with the Boussinesq approximation

$$
\begin{aligned}
&\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\frac{1}{\rho}\nabla p + \nu \nabla \cdot \nabla \mathbf{u} + \boldsymbol{\eta} d \quad \text{s.t.} \quad \nabla \cdot \mathbf{u} = 0, \\
&\frac{\partial d}{\partial t} + \mathbf{u} \cdot \nabla d = 0
\end{aligned} \tag{4}
$$

where $\mathbf{u}$ represents the velocity, ρ, p, ν denote density, pressure and viscosity respectively, the term $\boldsymbol{\eta} d$ represents the Boussinesq buoyancy, where d denotes the smoke marker density. The constraint $\nabla \cdot \mathbf{u} = 0$ enforces the condition of incompressibility (i.e volume preserving motion). As a test setup we have considered an enclosed space of 150×150 units, assuming circular smoke source with

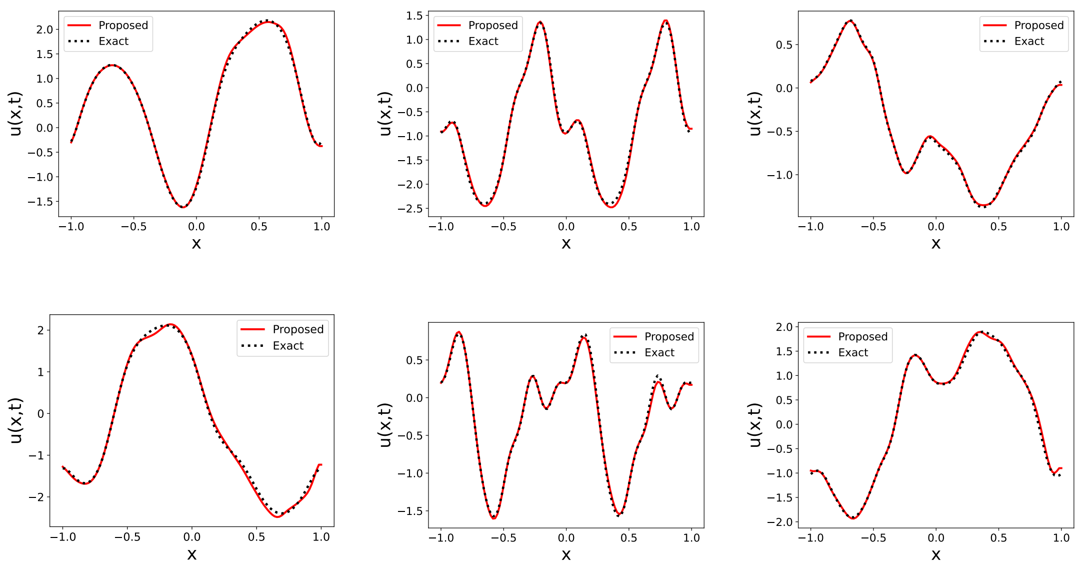

Figure 3: Allen-Cahn equation with missing cubic term. The comparison is shown for different Gaussian inputs taking three different kernel functions: Radial Basis Function (RBF), Exponential Sine Squared, and Matern (left to right). Top and bottom corresponds to two different initial conditions corresponding to same kernel. All the predictions are shown at $t = 98\Delta t$, way beyond the training time window of $t = 50\Delta t$

diameter $\phi = 10$ units. By varying the location of smoke source horizontally from 30 to 120 units and keeping the vertical position to be constant ($= 15$ units), we have generated 50 training sample simulations with 64×64 spacial resolution upto 50 time steps with $\Delta t = 1$ sec. The framework is trained upto 20 time steps. Here, the known governing equation contains a mis-specified buoyancy factor, and the integrated WNO is designed to correct for this by learning the discrepancy between the known equation and the actual data.

Figure 4 compare the ground truth with results from the proposed approach and the known physics model. In both cases, the proposed framework yields results that closely align with the ground truth. This superior performance can be attributed to the framework's utilization of both sparse data and approximate physics.

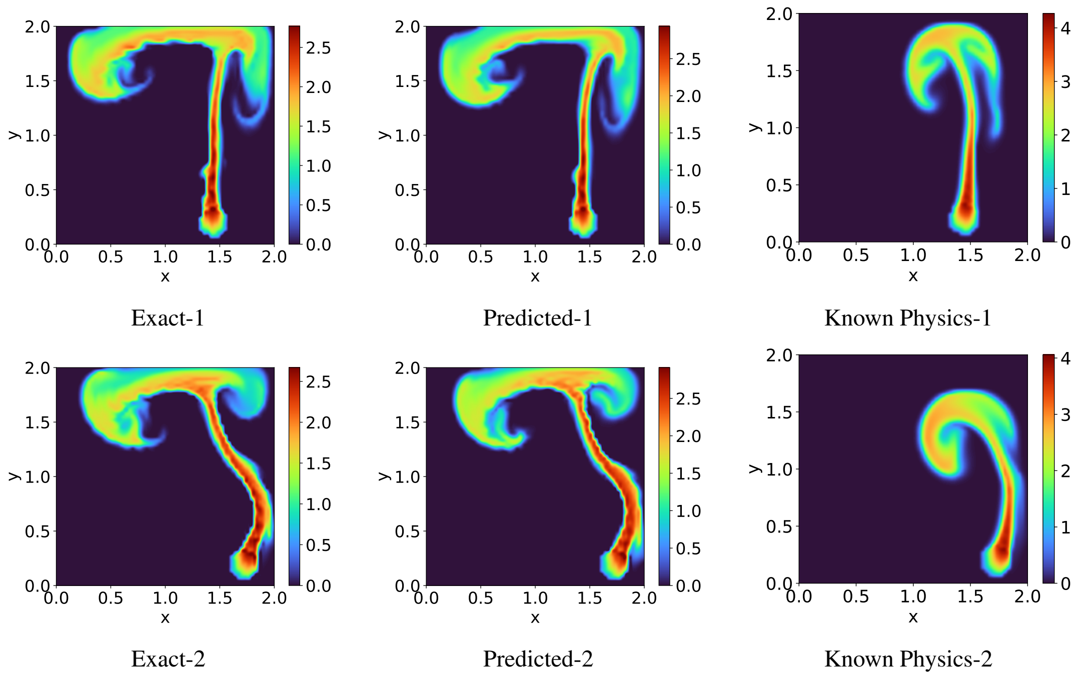

Figure 4: Comparison of predicted smoke density $s(x, y, t = 44\Delta t)$ obtained using the proposed MFO, with the ground truth and known physics for Smoke plume equation. The comparison is shown for three different initial conditions from the test samples.

References

[1] T. Tripura and S. Chakraborty, “Wavelet neural operator for solving parametric partial differ- ential equations in computational mechanics problems,” Computer Methods in Applied Me- chanics and Engineering, vol. 404, p. 115783, 2023.

[2] K. Um, R. Brand, Y. R. Fei, P. Holl, and N. Thuerey, “Solver-in-the-loop: Learning from differentiable physics to interact with iterative pde-solvers,” Advances in Neural Information Processing Systems, vol. 33, pp. 6111–6122, 2020

[3] S. Chakraborty, T. Chatterjee, R. Chowdhury, S. Adhikari, A surrogate based multi-fidelity approach for robust design optimization, Applied Mathematical Modelling 47 (2017) 726– 744.

[4] P. E. Gill, W. Murray, M. H. Wright, Practical optimization, SIAM, 2019.

[5] S. S. Afshari, F. Enayatollahi, X. Xu, X. Liang, Machine learning-based methods in structural reliability analysis: A review, Reliability Engineering and System Safety 219 (2022) 108223.

[6] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, L. Yang, Physics- informed machine learning, Nature Reviews Physics 3 (2021) 422–440.

[7] S. Cai, Z. Mao, Z. Wang, M. Yin, G. E. Karniadakis, Physics-informed neural networks (pinns) for fluid mechanics: A review, Acta Mechanica Sinica 37 (2021) 1727–1738.

[8] S. Chakraborty, Transfer learning based multi-fidelity physics informed deep neural network, Journal of Computational Physics 426 (2021) 109942.

[9] M. Raissi, P. Perdikaris, G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational physics 378 (2019) 686–707.

[10] S. Rezaei, A. Harandi, A. Moeineddin, B.-X. Xu, S. Reese, A mixed formulation for physics- informed neural networks as a potential solver for engineering problems in heterogeneous domains: comparison with finite element method, Computer Methods in Applied Mechanics and Engineering 401 (2022) 115616.

[11] S. Goswami, C. Anitescu, S. Chakraborty, T. Rabczuk, Transfer learning enhanced physics informed neural network for phase-field modeling of fracture, Theoretical and Applied Fracture Mechanics 106 (2020) 102447.

[12] N. Navaneeth, T. Tripura, S. Chakraborty, Physics informed wno, Computer Methods in Applied Mechanics and Engineering 418 (2024) 116546.

[13] S. Garg, S. Chakraborty, B. Hazra, Physics-integrated hybrid framework for model form error identification in nonlinear dynamical systems, Mechanical Systems and Signal Pro- cessing 173 (2022) 109039.

[14] V. K. Dertimanis, E. Chatzi, S. E. Azam, C. Papadimitriou, Input-state-parameter estima- tion of structural systems from limited output information, Mechanical Systems and Signal Processing 126 (2019) 711–746.

[15] D. J. MacKay, et al., Introduction to gaussian processes, NATO ASI series F computer and systems sciences 168 (1998) 133–166.

[16] N. Takeishi, A. Kalousis, Physics-integrated variational autoencoders for robust and in- terpretable generative modeling, Advances in Neural Information Processing Systems 34 (2021) 14809–14821.

[17] U. Tripathi, S. Garg, R. Nayek, S. Chakraborty, Physics-integrated deep learning for uncer- tainty quantification and reliability estimation of nonlinear dynamical systems, Probabilistic Engineering Mechanics 72 (2023) 103419.

[18] S. Li, Y. Yang, Data-driven identification of nonlinear normal modes via physics-integrated deep learning, Nonlinear Dynamics 106 (2021) 3231–3246.

[19] S. Chakraborty, et al., Deep physics corrector: A physics enhanced deep learning archi- tecture for solving stochastic differential equations, Journal of Computational Physics 479 (2023) 112004.