A Robust Training Framework for Physics-Informed Neural Operators in Stiff Systems: Application to DNP Transport in Molten Salt Reactors

Sinho Kim[a], Juhyeong Lee[a], Joon Young Bae[a], JinHo Song[a], Joon Eon Yang[a], Sung Joong Kim[a,b*]

[a]*Department of Nuclear Engineering, Hanyang University, 222 Wangsimnri-ro, Seongdong-gu, Seoul 04763, Republic of Korea*
[b]*Institute of Nano Science and Technology, Hanyang University, 222 Wangsimnri-ro, Seongdong-gu, Seoul 04763, Republic of Korea*
[*]Corresponding author: *sungjkim@hanyang.ac.kr*

Keywords : Molten Salt Reactors (MSRs), Physics-Informed Neural Operator (PINO), Training Stability

1. Introduction

The global pursuit of clean and sustainable energy has renewed interest in Molten Salt Reactors (MSRs), a promising Generation IV nuclear technology. Unlike conventional reactors, MSRs utilize a liquid fuel where fissile material is dissolved into a molten salt, which serves as both fuel and coolant. Liquid fuel confers several key benefits, such as high-temperature, low pressure operation, which enhances thermal efficiency and intrinsic safety. However, the unique configuration also introduces novel challenges in reactor physics and safety analysis. The circulation of the fuel salt creates a tightly coupled multiphysics environment where neutronics, thermal-hydraulics, and material transport are inextricably linked.

One representative challenge in modeling MSRs is the transport of delayed neutron precursors (DNPs). In conventional reactors, DNPs remain static within the solid fuel elements. In an MSR, however, they are transported with the fuel salt throughout the entire primary circuit. The fuel circulation has two critical effects on reactor dynamics. First, a significant fraction of DNPs may decay outside the active core, reducing the effective delayed neutron fraction βeff. A lower βeff narrows the margin to prompt criticality, demanding more precise control and analysis. Second, the spatial distribution of the delayed neutron source becomes dependent on the fuel's velocity field. During transients, such as an accidental loss of primary flow, the reduced salt velocity increases the DNP residence time within the core. Consequently, a larger fraction of precursors decay in the neutronically important core region, inserting positive reactivity. This feedback mechanism, which can counteract the negative temperature feedback, must be precisely modeled for safety analysis [1].

Capturing these feedback mechanisms with high fidelity simulations requires resolving detailed Multiphysics phenomena, which is computationally expensive. Such computational cost can also pose significant challenges when analyzing various scenarios, including MSR design optimization and predicting reactor behavior during transients. This computational bottleneck necessitates the development of fast and accurate surrogate models capable of replicating the underlying physics at a fraction of the cost.

Instead of purely data-driven, the physics-informing provides an alternative by embedding physical laws by incorporating the system's governing equations into the learning objective. This approach is intended to enhance physical consistency and generalization, but it faces a unique challenge in our context. The governing equations for DNP transport are mathematically "stiff," characterized by the coexistence of physical processes with vastly different time scales—fast advection and slow radioactive decay. The resulting stiffness can create a highly non-convex and complex loss landscape during model training. While the optimization challenges of applying Physics Informed Neural Networks (PINNs) to stiff systems are documented [2], our study investigates the phenomenon within the architecturally distinct Fourier Neural Operator (FNO) framework.

We demonstrate that for the stiff DNP transport problem, a naive application of the physics-informed loss acts as a maladaptive regularizer, disrupting the learning process and leading to unstable outcomes. To address the problem, we systematically compare three approaches: a purely data-driven model (Vanilla FNO), a PINO with a constant physics-loss weight, and a PINO trained with a loss annealing strategy. Our findings reveal that while the annealing strategy successfully mitigates training instability, the true value of the physics-informed approach emerges in challenging extrapolation scenarios. The annealed PINO demonstrates superior generalization and stability when predicting unseen conditions, highlighting its potential for building trustworthy surrogate models where data is inherently scarce and reliability is paramount.

2. Methology

2.1. Governing Physics and Data Generation

The spatio-temporal evolution of the g-th Delayed Neutron Precursor group concentration, $C_g(x,t)$, is governed by the advection-decay equation:

$$\frac{\partial C_g}{\partial t} + \nabla \cdot (u\, C_g) = -\lambda_g C_g$$

where λg is the radioactive decay constant for the g-th group, and u is the velocity field of the molten salt. In this study, we focus specifically on the DNP transport phenomena under steady-state flow conditions; thus, u is treated as static (time-invariant, only dependent on parameter Re) for each simulation case. The system's stiffness arises from the disparity in time scales between the fast advection term $\nabla \cdot (u\,Cg)$ and the comparatively slow decay term $-\lambda gCg$, particularly for precursors with long half-lives λg. This disparity poses a fundamental challenge for physics-informed learning methods.

To generate the ground truth dataset, we considered an incompressible, fully developed laminar flow (Poiseuille flow) in a 2D rectangular channel (Lx=0.2 m, Ly=0.05 m). The FLiBe salt properties were evaluated at a constant temperature of 700°C. The analysis included six DNP groups from U-235 thermal fission.

The system is driven by a constant source of precursors at the inlet, modeled with a uniform Dirichlet boundary condition: Cg (0, y, t) = 1.0. Ground truth data was generated by numerically solving these equations on a 256×64 grid. Each simulation was run for 500 time steps with a step size of 0.1 seconds, capturing the evolution towards a steady state. Simulations were performed at various inlet velocities corresponding to different Reynolds numbers (Re), representing distinct operational conditions.

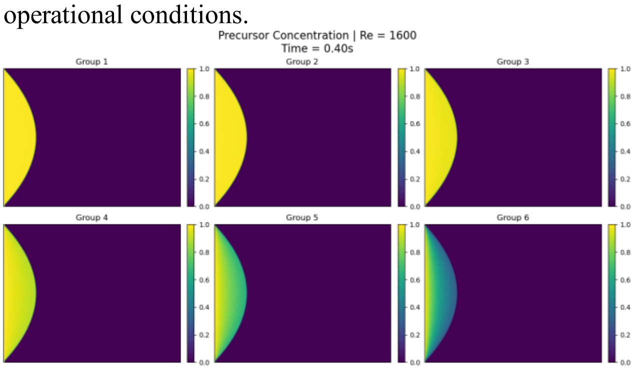

Fig. 1. Ground truth data set visualization

2.2. Physics-Informed Fourier Neural Operator

We employ the Fourier Neural Operator (FNO) architecture to learn the solution operator mapping flow conditions (parameterized by Re) to the DNP concentration fields. The FNO learns a resolution invariant mapping by performing global convolution operations efficiently in the frequency domain via the Fast Fourier Transform (FFT), enabling it to model complex, long-range dependencies [3].

To enhance data efficiency and enforce physical consistency, we augment the standard data-driven loss with a physics-based loss term that penalizes the PDE residual. The total loss function is a weighted sum:

$L(tot) = L_{(data)} + \alpha_{(phys)} L_{(phys)}$

where $L_{(data)}$ is the mean squared error (MSE) between the model's prediction and the simulation data, and $L_{(phys)}$ is the MSE of the PDE residual calculated on collocation points. The hyperparameter α balances the contribution of the two terms.

The conceptual architecture of our Physics-Informed Fourier Neural Operator (PI-FNO) is illustrated in Figure 2.

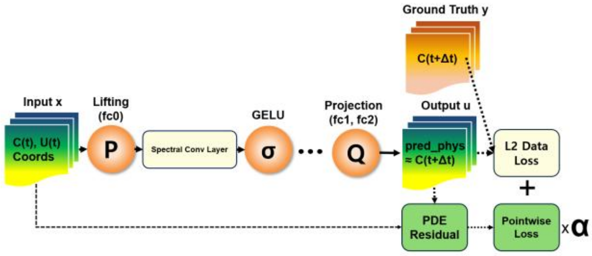

Fig. 2. Schematic of PINO

2.3. Naive Enforcement vs. Loss Annealing

For stiff systems, a naïve Constant PINO approach, where a fixed weight α>0 is applied from the start, presents a critical optimization challenge. The optimizer is immediately forced to navigate a difficult loss landscape where large gradients from the advection term conflict with smaller gradients from the decay term. Such gradient conflict can trap the model in poor local minima or lead to divergent behavior.

To overcome the instability, we introduce a physics loss annealing strategy. Training begins as a purely data driven FNO α. The physics loss is then introduced and its weight α is gradually increased to a target value of 10^{-5}. The rate of increase follows a cosine schedule, which allows the model to first learn the basic solution structure from data before being gently guided by the physical constraints.

3. Results and Analysis

3.1. Benchmark Configuration and Metrics

We conducted a benchmark study comparing the performance and training stability of three models: the data-driven Vanilla FNO (α =0), the Constant PINO, and the Annealing PINO. All models share an identical FNO architecture and were trained on a dataset comprising simulations at two operating points (Re=1200 and Re=2000). To ensure statistical robustness, each model configuration was trained 10 times with different random seeds.

Performance was evaluated using an autoregressive rollout on two unseen test cases: an interpolation case (Re=1600) and an extrapolation case (Re=2100). This closed-loop evaluation can rigorously test the model's temporal stability and generalization capabilities. We report two primary metrics: the Overall root Mean Squared Error (rMSE) across the time series and the Final-State rMSE, which measures convergence to the correct steady-state solution. We also report the number of training runs that failed to converge to a physically plausible solution.

3.2. Comparative Analysis

Interpolation Performance (Re=1600): As shown in Table 1, the Vanilla FNO serves as a strong performance baseline, successfully converging in 9 out of 10 runs and achieving the lowest average errors. In stark contrast, the Constant PINO exhibited severe training instability, failing in 30% of the runs. Its average errors were higher, and the large standard deviations indicate unpredictable and unreliable performance. Such unreliable performance validates our hypothesis that naively enforcing the PDE residual disrupts learning in stiff systems. The Annealing PINO successfully resolves the instability, reducing the failure rate and performance variance to levels comparable to the Vanilla FNO. However, its accuracy is slightly lower than the data driven baseline, suggesting that in data-rich interpolation scenario, the physics-loss term acts as a slight constraint.

Table 1: Comparative analysis of interpolation performance (Re=1600)

Case	Overall rMSE Avg	Overall rMSE StDev	Final-State rMSE Avg	Final-State rMSE StDev	Fail converging
Vanilla FNO	**28.86**	**11.96**	**8.11**	**8.65**	**1/10**
Constant PI-FNO	**36.55**	**24.14**	**23.45**	**36.9**	**3/10**
Annealing PI-FNO	**32.16**	**9.81**	**10.53**	**10.2**	**2/10**

(Unit: e-03, Device: NVIDIA GeForce RTX 3060 Ti Vram 8GB)

Extrapolation Performance (Re=2100): The true value of the physics-informed approach becomes evident in the more challenging extrapolation task, as detailed in Table 2. Here, the Vanilla FNO struggles significantly, failing to converge in half of the runs and exhibiting extremely high errors and variance. This demonstrates the limitations of purely data-driven models when confronted with out-of-distribution data. The Annealing PINO, however, shows remarkable robustness. It maintains a low failure rate and achieves significantly lower average errors and variance compared to the Vanilla FNO. Such superior generalization capability stems from the physical inductive bias provided by the PDE loss, which constrains the model to physically plausible solutions even in data-scarce regions.

Table 2: Comparative analysis of extrapolation performance (Re=2100)

Case	Overall rMSE Avg	Overall rMSE StDev	Final-State rMSE Avg	Final-State rMSE StDev	Fail converging
Vanilla FNO	**84.10**	**92.14**	**82.29**	**129.91**	**5/10**
Annealing PI-FNO	**60.54**	**76.48**	**43.87**	**96.72**	**2/10**

4. Conclusions and Discussions

In this study, a critical guideline for applying Physics Informed Neural Operators to the stiff systems was developed. It was shown that, in MSR DNP transport, the use of a constant loss weight for physics enforcement causes pathological training behavior and undermines model reliability. The proposed loss annealing strategy, however, proved to be an effective and robust solution, guaranteeing stable convergence by first allowing the model to learn from data before gradually imposing physical constraints.

The key contribution of this study is the finding that, although purely data-driven models may perform slightly better in data-rich interpolation tasks, the annealed PINO achieves markedly superior generalization and reliability in data-scarce extrapolation scenarios.

A methodological limitation arises from the use of the standard FNO architecture, which relies on the FFT and is therefore most suitable for uniform grids. Practical MSR simulations with complex geometries will require adopting alternative frameworks capable of handling unstructured meshes, such as Geometric FNO (Geo-FNO) or Physics-Informed DeepONets. Nevertheless, the proposed annealing strategy is architecture-agnostic, and future work will focus on integrating it with these advanced architectures. The long-term vision is to enhance physical fidelity by including fission source terms and loop recirculation, leveraging this robust PINO framework to develop a tool for the rapid and reliable estimation of the effective delayed neutron fraction.

ACKNOWLEDGEMENT

This work was supported by Korea Foundation Of Nuclear Safety (KoFONS) using the financial resource granted by the RS-2025-02312295, and the Human Resources Development of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government Ministry of Knowledge Economy (RS-2024-00439210).

REFERENCES

[1] Chen, S., Li, R., Zuo, X., Cheng, M., & Dai, Z. (2025). Dynamic Effect of the Delayed Neutron Precursor Distribution on System Safety Analysis in Liquid-Fueled Molten Salt Reactor. *Energies*, 18(3), 670.

[2] Ji, W., Qiu, W., Shi, Z., Pan, S., & Deng, S. (2021). Stiff-PINN: Physics-informed neural network for stiff chemical kinetics. *The Journal of Physical Chemistry A*, 125(36), 8098-8106.

[3] Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., & Anandkumar, A. (2021). Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations (ICLR)*.

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