Session 8

PDE Solvers

S8-A6 Miranda J S Horne — Hard Constraint Projection in a Fourier Neural Operator

S8-A7 Leonardo Boledi — Accelerating Reactive Transport Simulations with KAN-based Surrogate Models

S8-A10 Xiaowei Jin — Operator learning method with adaptive activation function for high-speed shock tube problems and airfoil flows

S8-A11 Mohammad Sheikholeslami — From Discretization to Neural Solvers: A Spectrum of Methods for the Lid-Driven Cavity Problem

S8-A13 Patrick Vaudrevange — Tina Vartziotis, Michael Keckeisen

S8-A14 Kie Okabayashi — Time-stepping Model of Cavitation Flow around a Clark-Y11.7% Hydrofoil Trained by CFD Data

S8-P1 Aleksandar Jemcov — Conditional Givens Networks for Scalable Unitary Liouville Transport

S8-P2 Senwei Zheng — Physics-encoded resolution-generalizable neural network for truncation error correction of solving nonlinear PDEs

S8-P3 Paul Horvath — Fourier Neural Operator for Journal Bearing Modelling

S8-P4 Kosov Vladimir — Influence of temperature on concentration convection caused by instability of the equilibrium of ternary gas mixtures based on data using CFD modeling improved with the …

S8-P5 Xukun Wang — Accelerating high-order discontinous Galerkin solver using auto-differentiation and neural networks

S8-P8 Francis Padilla — Accelerating Kinetic Fokker-Planck simulations using a GPU-optimized deep neural network surrogate: Application to rarefied internal and hypersonic external flows

S8-P9 Ruihan Zhang — Super-Resolution Predictive Initialization Using Neural Operators for Accelerating CFD Simulations

Hard Constraint Projection in a Fourier Neural Operator

[1]Miranda J S Horne*; [1]Peter K Jimack; [2]Amirul Khan; [3]He Wang;

[1]School of Computer Science,[2]School of Civil Engineering, [12]University of Leeds, Woodhouse Lane, Leeds, LS2 9JT
[3]UCL Centre for Artificial Intelligence, University College London, Gower Street, London, WC1E 6BT
**Corresponding email: scmho@leeds.ac.uk*

Abstract

In recent years, physics informed machine learning for fluid dynamics has shown great potential in a variety of applications. In theory, by embedding the known governing equations in a machine learning setting, the models can generate results that better observe the physics of the system along with reducing reliance on the quality and quantity of training data. One of the primary methods for incorporating a known partial differential equation (PDE) in a machine learning setting, the residual penalty method (proposed by Raissi et al. [doi.org/10.1016/j.jcp.2018.10.045]), uses the residual of the PDE as an additional term in the loss function. This results in a multi-objective optimisation and introduces multiple new hyperparameters that require problem specific tuning. Furthermore, constructing the residual of the PDE via automatic differentiation has been found to be a computational bottleneck during the training process.

In this work, we strictly impose the discretised form of the governing PDE within a Fourier neural operator (FNO) [https://doi.org/10.48550/arXiv.2010.08895]. During training and evaluation, the FNO generates an initial prediction of the solution field on a uniform mesh. Each solution prediction is then projected locally to the hyperplane of solutions to a finite difference approximation of the governing PDE, employing the hard constraint projection (HCP) method used by Chen et al. [doi.org/10.1016/j.jcp.2021.110624]. With physical knowledge embedded in the unlearnable final projection layer, the loss function depends only on the data error, avoiding the multi-objective challenge found in models which append the loss function a physics informed penalty term. We present results from an HCP-FNO trained on the nonlinear 1D Burgers' PDE, along with providing insight into best practice when implementing the HCP method.

Keywords: Physics Informed, Hard Constraints, Fourier Neural Operator, Burgers' PDE

Accelerating Reactive Transport Simulations with KAN-based Surrogate Models

[1]L. Boledi*; [1]R. Santoso; [1]J. Poonoosamy;

[1]Institute of Fusion Energy and Nuclear Waste Management (IFN-2): Nuclear Waste Management,4 Forschungszentrum Jülich GmbH, 52425 Jülich, Germany

Abstract

The mobility of chemical species in subsurface environments is governed by the tight coupling of fluid flow, solute transport, and geochemical reactions across space and time. Reactive transport modelling (RTM) provides the quantitative framework to describe these coupled processes and to predict the evolution of complex geochemical systems. These simulations require the repeated solution of chemical reactions, often billions of times per run, leading to prohibitive computational costs. This challenge becomes critical in uncertainty quantification and risk assessment tasks where large numbers of simulations are required, motivating the development of fast and reliable surrogate models.

Multilayer perceptrons (MLPs) are commonly used as surrogates for chemical reactions [1], but recent alternatives may improve performance. In particular, Kolmogorov-Arnold Networks (KANs) have shown promise for approximating PDE solutions, achieving improved accuracy with fewer parameters than MLPs [2]. This efficiency makes them attractive for chemical reaction modeling, an application that, to the best of our knowledge, has not yet been explored. An additional key problem is introduced when integrating these surrogates in reactive transport simulations: prediction errors can accumulate over space and time, potentially violating physical constraints. To address this issue, we incorporate governing equations, namely the Gibbs energy minimization, directly into the training process, similarly to physics-informed neural networks.

In this work, we propose both data-driven and physics-based KANs for chemical reactions and evaluate them on a widely used calcite-dolomite reactive transport benchmark [3], which describes the dissolution of calcite and precipitation of dolomite in a one-dimensional column flushed with a magnesium-rich fluid. We investigate and compare their accuracy and computational speed-up relative to classical solvers. Finally, we assess their potential for integration into reactive transport frameworks.

Keywords: Reactive Transport Modelling, Chemical Calculations, Surrogate Model, Kolmogorov-Arnold Network

References

[1] N. I. Prasianakis et al., Geochemistry and machine learning: methods and benchmarking, Environmental Earth Sciences 84 (5) (2025) 121. doi:10.1007/s12665-024-12066-3.

[2] Z. Liu et al., KAN: Kolmogorov-Arnold Networks, arXiv:2404.19756 [cs] (Feb. 2025). doi:10.48550/arXiv.2404.19756.

[3] P. Engesgaard and K. L. Kipp, A geochemical transport model for redox-controlled movement of mineral fronts in groundwater flow systems: A case of nitrate removal by oxidation of pyrite, Water Resour. Res., 28, 2829–2843 (1992). doi: 10.1029/92WR01264.

Operator learning method with adaptive activation function for high-speed shock tube problems and airfoil flows

[1]Xiaowei Jin*; [2]Shengkai Xu; [1,2]Hui Li

[1]*Key Lab of Smart Prevention and Mitigation of Civil Engineering Disasters of the Ministry of Industry and Information Technology, Harbin Institute of Technology, Harbin 150090, China*

[2] *School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150090, China*

** E-mail:xiaowei.jin@hit.edu.cn*

Abstract

To accurately capture sharp features like shock and expansion waves in high-speed compressible flows we introduce the OL-AAF (Operator learning method with adaptive activation function), a novel framework designed to learn solution operators for parametric governing equations for high-speed shock tube problems and airfoil flows. The primary novelty of the neural network lies in a layer-wise gated Bridge branch that establishes interaction information channels between the latent representations of neural networks. By utilizing learnable gating parameters to adaptively modulate coupled hidden features at every layer, the architecture's representational capacity is markedly increased. The approximation ability of the network is also improved by a adaptive Z-shape-ReLU activation function in the basis function network's terminal layers. Through the simultaneous optimization of its positive-axis slope and truncation threshold, Z-shape-ReLU counteracts the numerical 'blurring' typical of SiLU or other activation functions, ensuring high-fidelity approximations of strong discontinuities. Validated against high-speed shock tube and airfoil flow datasets, the OL-AAF framework outperforms DeepONet in shock-position accuracy and displays strong generalization. This study gains a robust architectural template for high-speed, data-driven aerodynamic prediction.

Keywords: Operator learning, adaptive activation function, high-speed flows, gated Bridge branch

From Discretization to Neural Solvers: A Spectrum of Methods for the Lid-Driven Cavity Problem

[1]Mohammad Sheikholeslami*; [2]Saeed Salehi; [1]Wengang Mao; [3]Håkan Nilsson; [1]Arash Eslamdoost;

[1]Division of Marine Technology, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg SE-412 96, Sweden

[2]Division of Applied Thermodynamics and Fluid Mechanics, Department of Management and Engineering, Linköping University, SE-581 83 Linköping, Sweden

[3]Division of Fluid Dynamics, Department of Mechanics and Maritime Sciences, Chalmers University of Technology, Gothenburg SE-412 96, Sweden

Abstract

This study introduces a spectrum of numerical approaches for solving the lid-driven cavity (LDC) problem, spanning from traditional discretization-based solvers to fully neural formulations. At one end lies the pure finite volume method (FVM), where the Navier–Stokes equations are discretized and solved using conventional solvers, while at the other end is the physics-informed neural network (PINN), in which a neural network approximates the solution by minimizing physics-based residuals with automatic differentiation. Between these extremes, some hybrid approaches are investigated. For instance, one uses a finite volume discretization while solving the resulting system of equations through gradient-based optimization, and another employs a PINN in which the residuals are computed using a finite-volume-style discretization. The LDC problem serves as a representative benchmark because it contains strong velocity gradients near the top corners and exhibits a wide range of spatial frequencies in the solution. These characteristics make it suitable for evaluating how different numerical paradigms capture multiscale flow behavior. By positioning all approaches along a continuous methodological spectrum, this study highlights their conceptual relationships and evaluates their relative strengths in terms of stability, accuracy, and their ability to resolve high-gradient regions. Since the LDC problem reflects challenges common to many incompressible flow simulations, the conclusions of this work can be extended to a broader class of computational fluid dynamics problems.

Keywords: PINN, FVM, optimization, lid-driven cavity

AI-Accelerated Multi-Fidelity Fire Simulation and Hybrid Mesh Generation for Industrial CFD

Patrick Vaudrevange[1]*; Tobias Roessler[1]; Tina Vartziotis[1,2,3]; Michael Keckeisen[1]

[1]*TWT GmbH Science & Innovation*

[2]*Harvard University*

[3] *National Technical University of Athens*

Abstract

Fire simulations are essential in safety-critical engineering processes. However, high-fidelity simulations of complex geometries are computationally expensive. They must be repeated multiple times to iteratively identify the optimal safety configuration and repeated if major design modifications are introduced. This iterative process leads to significant computational efforts in large-scale computational fluid dynamics (CFD) analyses. In this work, we present an automated AI workflow for fast and accurate multi-fidelity AI surrogate models that predict key KPIs for fire safety. High-fidelity CFD simulations are combined with lower-fidelity data to train data-driven surrogate models capable of approximating relevant quantities such as temperature distribution and smoke propagation characteristics. We compare different approaches to the AI model and discuss their respective advantages and disadvantages with respect to prediction accuracy, computational efficiency, and robustness under geometric variations. The objective is to significantly reduce simulation turnaround time while preserving engineering-relevant accuracy for safety assessment. In addition, we address automatic mesh generation as a prerequisite for running such simulations. Appropriate mesh generation is essential for many simulation domains, including finite elements and computational fluid dynamics. The problem is difficult, as contradicting requirements must be satisfied simultaneously: computational time of the simulation should be minimal, while maximal accuracy of the results must still be ensured. We compare two approaches for repetitive meshing of similar but not identical three-dimensional objects: an AI-based approach and a physics-based method. For our specific task, the AI approach shows advantages in generalizability, while the physics-based method enables accurate meshing of difficult regions. Finally, we discuss a possible combination of both approaches towards faster mesh generation for efficient simulations. Finaly, we highlight possible industrial applications and outline further directions of applied research.

Keywords: Fire simulation, Computational Fluid Dynamics (CFD), Multi-fidelity surrogate modeling, Machine learning for CFD, Automatic mesh generation, Industrial simulation workflows

Time-stepping Model of Cavitation Flow around a Clark-Y11.7% Hydrofoil Trained by CFD Data

[1]Kie Okabayashi*; [1]Yuina Motozono; [1]Kazuyuki Tanaka;

[1]*The University of Osaka*

Abstract

To establish a novel prediction framework for cavitation flow that replaces computational fluid dynamics (CFD) analysis, we developed a data-driven time-stepping surrogate model using convolutional long short-term memory (LSTM). This model is an autoregressive model that predicts the flow field at the next time step from the current and past flow fields discretized in space and time. It can be considered as extracting the solution to the flow's temporal evolution equation itself from the field data.

For two-dimensional cavitation flow around an airfoil, we trained the neural network to learn an autoregressive time-stepping function using CFD data from a small number of flow conditions as training data. As a result, the model not only accurately predicted the velocity, pressure, and void fraction distributions of the training data but also demonstrated generalization capability by successfully predicting flow field not used in training. Notably, despite using only a few unsteady cases and single-phase conditions as training data, the model successfully predicted the time evolution of steady flow not included in the training data. Under periodically fluctuating unsteady conditions, the model maintained stable time stepping over six or more cycles without increasing error, demonstrating its robustness. However, under any conditions, challenges remain in the reproducibility of pressure field, and it is particularly difficult to predict the sudden pressure increase that occurs when the cavity disappears.

This surrogate model can temporally evolve the flow over time at speeds 40 times faster than CFD, enabling efficient parametric analysis and aiding design. Furthermore, unlike CFD, it does not require a physical model. To date, no physical model exists that can uniformly represent diverse cavitation phenomena, making this a groundbreaking fluid analysis method for cavitation.

Keywords: Cavitation, Hydrofoil, PDE Solver, Surrogate Model, Generalization

Conditional Givens Networks for Scalable Unitary Liouville Transport

Aleksandar Jemcov* **and Scott C. Morris**

Department of Aerospace and Mechanical Engineering,
University of Notre Dame, Notre Dame, IN 46556, USA

*Corresponding author: ajemcov@nd.edu

Abstract

The Koopman–von Neumann formulation recasts the classical Liouville equation as a unitary quantum evolution suitable for near-term quantum hardware. Two obstacles have limited practical use: circuit depth grows exponentially with system size, and small deviations from exact orthogonality cause probability errors that compound over many time steps.

Conditional Givens networks address both obstacles. The propagator is built as a product of Givens rotations whose angles are predicted by a small conditioned neural network. Orthogonality is guaranteed by construction, independent of training quality. Training on low-dimensional two-mode subsystems and combining them with Lie–Trotter splitting yields a full-system circuit whose depth scales only with the subsystem size.

On a three-mode Navier–Stokes triad, unconstrained networks produce probabilities of order 10^{292} within 50 steps, while the proposed networks conserve probability to machine precision by construction. A matrix-exponential parameterization achieves higher classical fidelity but requires a dense-unitary synthesis step whose gate count grows quadratically with the subsystem dimension; the Givens parameterization outputs angles that directly specify photonic beam-splitter or single-qubit rotation settings, and extends to an unseen viscosity in the tests reported here without retraining.

Keywords: Conditional Givens networks, Koopman–von Neumann, unitary propagation, quantum simulation, Navier–Stokes triad

1 Introduction

The Liouville equation governs the evolution of probability density functions for deterministic dynamical systems. In the Koopman–von Neumann (KvN) formulation [Koopman, 1931, von Neumann, 1932, Joseph, 2020], it becomes a Schrödinger-type equation $\partial_t \psi = L\psi$ for a wavefunction ψ and an anti-symmetric generator L. The propagator $U = e^{L\Delta t}$ over a time step Δt is therefore orthogonal, and the total probability $P = \int |\psi|^2 \, d\boldsymbol{a}$, integrated over the vector of mode amplitudes $\boldsymbol{a}$, is conserved exactly.

For an N-mode system discretized on M grid points per mode, this formulation maps $K = M^N$ classical degrees of freedom onto a Hilbert space encodable in $N\lceil \log_2 M \rceil$ qubits or K photonic modes. Two obstacles prevent practical quantum algorithms. First, generic decompositions of a K-dimensional unitary require $O(K^2)$ gates [Shende et al., 2006]. Because $K = M^N$, both state preparation and propagation scale exponentially in N. Second, learned operators V that are not exactly orthogonal produce probability errors that compound as $(1+\epsilon)^n$ over n time steps, where ϵ denotes the per-step orthogonality deviation. After 50 steps with $\epsilon = 10^{-2}$, the norm inflates by a factor of 1.64; with $\epsilon = 10^{-1}$ the factor reaches 117. This divergence is structural: the parameterization admits non-orthogonal matrices, and the optimizer exploits larger norms to reduce training loss.

Addressing these two problems is a prerequisite for scalable quantum simulation of high-dimensional nonlinear systems such as fluid turbulence, which remains intractable with generic circuit synthesis. The architecture presented below combines a Givens parameterization that is orthogonal by construction with Trotter composition from tractable subsystems.

2 Koopman–von Neumann Generator and Dimensional Transfer

For a Galerkin system $\dot{a}_k = f_k(\boldsymbol{a})$, the joint probability density $\rho(\boldsymbol{a}, t)$ satisfies the Liouville equation $\partial_t \rho = -\sum_k \partial_{a_k}(f_k \rho)$. The wavefunction $\psi = \sqrt{\rho}\, e^{i\phi}$, with phase ϕ, yields the Schrödinger-type equation $\partial_t \psi = L\psi$, where the Weyl-ordered (symmetrically-ordered) generator is

$$L = -\frac{1}{2} \sum_{k=1}^{N} \left(F_k D_k + D_k F_k \right), \tag{1}$$

with D_k a skew-symmetric summation-by-parts operator and $F_k = \mathrm{diag}(f_k(\boldsymbol{a}))$. The symmetrized product ensures $L = -L^T$, so $U = e^{L\Delta t}$ is orthogonal.

For the three-mode Navier–Stokes triad with viscosity ν,

$$\dot{a}_1 = -\nu a_1 + a_2 a_3, \quad \dot{a}_2 = -4\nu a_2 + a_1 a_3, \quad \dot{a}_3 = -9\nu a_3 - 2 a_1 a_2, \tag{2}$$

setting $M = 8$ yields $K = 512$. The full propagator requires depth $O(M^3)$. Lie–Trotter splitting reduces this to depth $O(M^2)$:

$$e^{L\Delta t} \approx e^{L_{(12|3)}\Delta t}\, e^{L_{(3|12)}\Delta t}, \tag{3}$$

where $L_{(12|3)}$ evolves modes (a_1, a_2) at fixed a_3 and $L_{(3|12)}$ evolves a_3 at fixed (a_1, a_2). Each factor is block-diagonal, preserving orthogonality. Training data are generated on $K_2 = M^2 = 64$ subsystems; the full propagator is assembled by composition. The splitting error is $O(\Delta t^2)$ in state error; at $\Delta t = 0.02$ the single-step infidelity is 7.9×10^{-8}. The approach extends to $N > 3$ modes because Navier–Stokes nonlinearities consist of $O(N^2)$ triadic interactions, each handled by the same M^2 Givens architecture.

3 Givens Parameterization

A Givens rotation $G_{ij}(\theta)$ is orthogonal for any θ:

$$[G_{ij}(\theta)]_{ii} = \cos\theta, \quad [G_{ij}(\theta)]_{ij} = -\sin\theta, \quad [G_{ij}(\theta)]_{ji} = \sin\theta, \quad [G_{ij}(\theta)]_{jj} = \cos\theta. \tag{4}$$

Let $U(\nu, a_3)$ denote the exact conditional propagator and $V(\nu, a_3)$ the learned Givens product that approximates it. The triadic structure of the Navier–Stokes nonlinearity makes conditional training tractable: fixing a_3 reduces each factor to a two-dimensional interaction, so the MLP maps $\mathbb{R}^2$ to angles for a $K_2 = M^2$ orthogonal matrix. Trotter splitting then composes these conditional factors; each block is unitary, a direct sum of unitary blocks is unitary, and a product of unitaries is unitary, so the full-system propagator is orthogonal regardless of the accuracy of the learned V.

$$V = \prod_{\ell=1}^{K_2} \prod_{p=1}^{K_2/2} G_{i(\ell,p),\, j(\ell,p)}\big(\theta_{\ell p}\big), \tag{5}$$

with $K_2 = 64$ layers of 32 rotations each (2,048 angles). These angles are produced by a multilayer perceptron (MLP) with two hidden layers of width 48 (tanh activations) mapping $\mathbb{R}^2$ to $\mathbb{R}^{2048}$ (102,848 parameters). The product is orthogonal at every stage of training. Combined with Trotter splitting, the architecture removes three exponential costs: data generation ($O(M^{2N/3})$), gate decomposition (angles are the circuit), and on-hardware variational optimization.

4 Training and Numerical Results

Exact conditional propagators are computed for six viscosities $\nu \in \{0.05, \ldots, 0.50\}$ and all $M = 8$ values of a_3, yielding 48 training samples. The loss, evaluated on a Gaussian initial wavefunction ψ_0 and the MLP parameters $\boldsymbol{w}$, is two-step infidelity:

$$\mathcal{L}(\boldsymbol{w}) = \frac{1}{2} \sum_{s=1}^{2} \Big(1 - \big| \langle V^s \psi_0 \,|\, U^s \psi_0 \rangle \big|^2 \Big). \tag{6}$$

Training uses Adam with cosine annealing for 800 epochs.

When the network outputs matrix entries without constraint, the optimizer drives $\|V\psi_0\|$ above unity because the loss in Eq. (6) is not bounded below in this regime; within 50 steps the probability reaches values of order 10^{292}. A single-strength orthogonality penalty ($\lambda = 1$) does not arrest the divergence: the fidelity gradient along norm-increasing directions exceeds the penalty gradient in the regime visited during optimization, and training gradient norms of $O(10^{10})$ for the unconstrained methods versus $O(10^{-2})$ for architecturally orthogonal ones document the qualitative separation. Increasing λ to dominate the fidelity signal suppresses the training gradient altogether, so the two objectives cannot be jointly satisfied by a penalty of any fixed strength.

The two architecturally orthogonal parameterizations (Givens and matrix exponential, $V = e^A$ with $A = -A^T$) both remain normalized to machine precision but differ in quantum deployment cost. The matrix exponential attains the higher 50-step fidelity ($\mathcal{F}_{50} = 0.955$ versus 0.879 for Givens at the same 400-epoch checkpoint), but extracting a quantum circuit from e^A requires the $O(K_2^2)$-gate unitary synthesis of Shende et al. [2006], which the present framework avoids. Each Givens angle, in contrast, directly specifies a photonic beam splitter or controlled-$\mathrm{R_y}$ rotation, so the trained output is itself a compiled circuit; the measured fidelity gap is the price of this direct-deployment property.

Table 1: Architecture comparison (400 epochs, $\nu = 0.15$). †: orthogonal by construction. $\mathcal{F}_{50}$: fidelity of $V^{50}\psi_0$ against $U^{50}\psi_0$.

Method	Params	Loss	$\mathcal{F}_{50}$	$\lvert P-1\rvert_{50}$	$\lVert V^TV-I\rVert_{\max}$
Givens†	102,848	2.2×10^{-4}	0.879	6×10^{-15}	2×10^{-15}
Expm†	101,280	1.6×10^{-6}	0.955	5×10^{-15}	9×10^{-16}
Unconstr.	203,200	diverges	—	∞	2×10^{5}
Unc.+pen.	203,200	diverges	—	∞	2×10^{5}

Figure 2 shows rollout fidelity over 50 steps. At the training viscosity $\nu = 0.15$ the 50-step fidelity is 0.879 at the 400-epoch checkpoint used in Table 1 and 0.924 at the best checkpoint (epoch 200); at the unseen viscosity $\nu = 0.07$ it is 0.934. The unseen value exceeds the training value because lower viscosity places $U(\nu)$ closer to the identity and is therefore easier to approximate. At $M = 12$ ($K_2 = 144$), 100 epochs yield single-step fidelity 0.9999.

Probability is conserved exactly by construction for any initial condition (IC), but the learned fidelity is directional: for Gaussians centered far from the training mean, 50-step fidelity can fall to $\sim 10^{-3}$ while $|P-1|$ remains below 10^{-14}. This directional sensitivity arises because the per-wavefunction loss in Eq. (6) supplies gradient signal only along basis directions excited by ψ_0; the behavior is a limitation of the training procedure (addressable through multi-IC training or a Frobenius-norm matrix loss) rather than of the architecture.

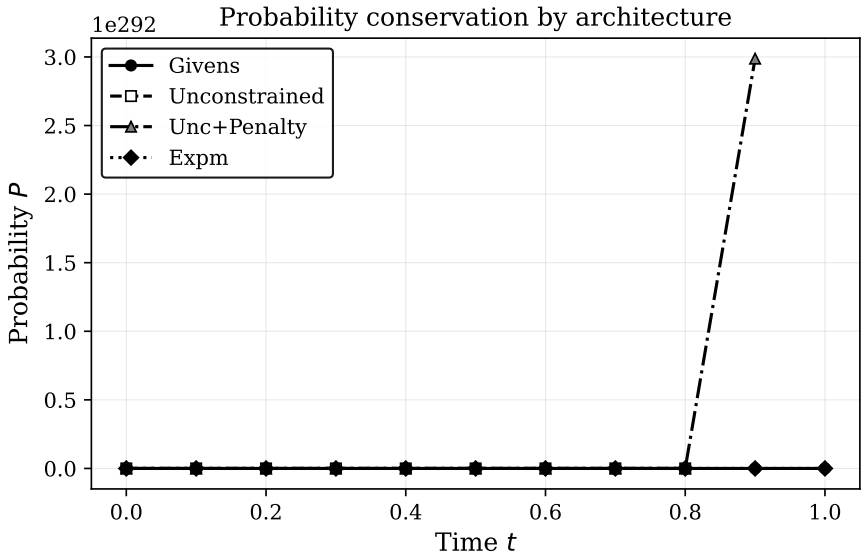


Figure 1: Total probability P over 50 Trotter steps ($\nu = 0.15$). Givens and Expm conserve $P = 1$ exactly. The penalized unconstrained case diverges to $P \sim 10^{292}$.

5 Quantum Circuit Interpretation

Each Givens rotation $G_{ij}(\theta)$ maps to hardware through standard constructions. On photonic processors it is a beam splitter with transmissivity $\cos^2\theta$ in a Reck or Clements mesh [Reck et al., 1994, Clements et al., 2016]. On qubit processors the K_2-dimensional space uses 6 qubits, and each rotation becomes a controlled-$\mathrm{R_y}$ gate (a controlled rotation about the Bloch-sphere y-axis) [Möttönen et al., 2004]. The total gate count remains polynomial in K_2; by contrast, decomposing a dense unitary into elementary gates requires $O(K_2^2)$ gates [Shende et al., 2006]. Once trained, the propagator applies to any IC without recompilation.

6 Conclusion

Conditional Givens networks address the two main obstacles to practical unitary Liouville propagation on near-term quantum hardware. Dimensional transfer via Trotter splitting combined with a Givens parameterization that is orthogonal by construction yields circuits whose depth scales only with subsystem size; probability conservation holds to machine precision at every time step, independent of training quality. On the Navier–Stokes triad, the method conserves probability exactly, achieves best 50-step fidelity of 0.924 at the training viscosity and 0.934 at an unseen viscosity, and produces circuits compatible with photonic and qubit gate sets without an intermediate synthesis step. The measured fidelity gap relative to the matrix-exponential parameterization is the cost of direct circuit interpretability.

The principal limitations are the directional dependence of fidelity on the training initial condition and the lack of tests beyond three modes. Both are testable within the present framework: multi-IC or Frobenius-norm training addresses the first, and five-mode systems are accessible classically while systems with seven or more modes are candidates for near-term quantum hardware. Whether the Trotter ordering over the $O(N^2)$ triadic interactions in a higher-mode system preserves the coherent inter-scale coupling underlying the Kolmogorov energy spectrum remains an open question that the present architecture makes empirically tractable.

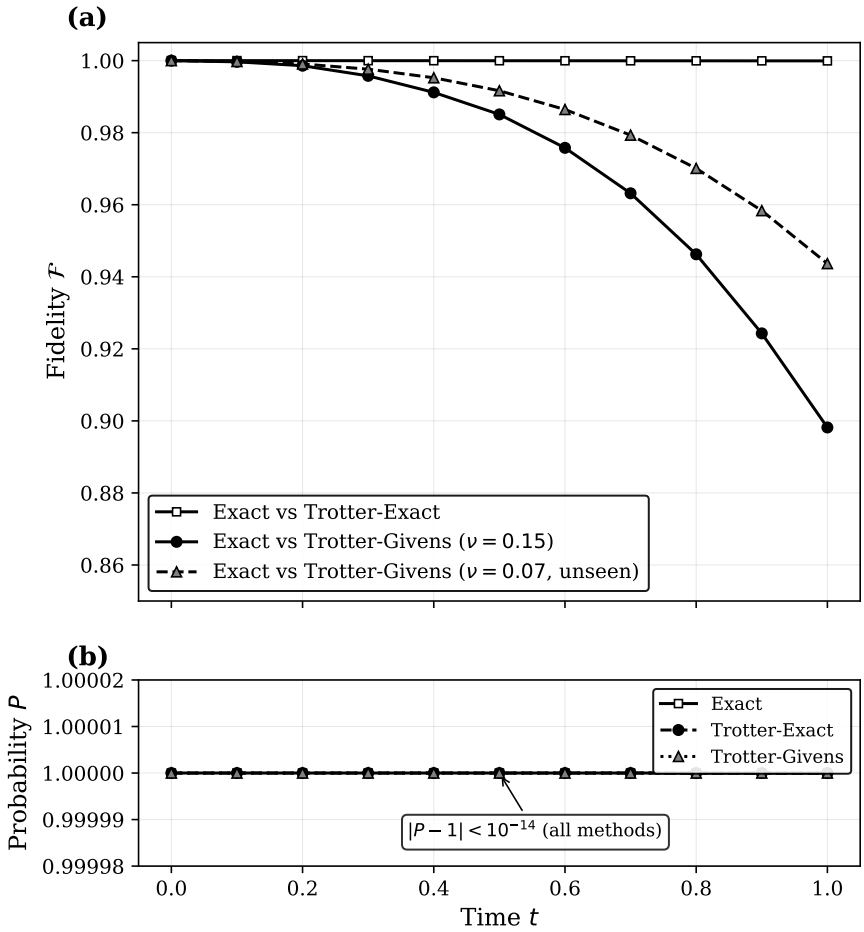


Figure 2: Fidelity over 50 time steps ($\Delta t = 0.02$). Givens-Trotter fidelity decays smoothly at the training viscosity ($\nu = 0.15$) and at the unseen value ($\nu = 0.07$). Trotter-exact fidelity exceeds 0.9999 over the full horizon.

References

B. O. Koopman. Hamiltonian systems and transformations in Hilbert space. *Proc. Natl. Acad. Sci. USA*, 17:315–318, 1931.

J. von Neumann. Zur Operatorenmethode in der klassischen Mechanik. *Ann. Math.*, 33:587–642, 1932.

I. Joseph. Koopman–von Neumann approach to quantum simulation of nonlinear classical dynamics. *Phys. Rev. Research*, 2:043102, 2020.

M. Arjovsky, A. Shah, and Y. Bengio. Unitary evolution recurrent neural networks. In *Proc. ICML*, pp. 1120–1128, 2016.

L. Jing et al. Tunable efficient unitary neural networks (EUNN) and their application to RNNs. In *Proc. ICML*, pp. 1733–1741, 2017.

K. Helfrich, D. Willmott, and Q. Ye. Orthogonal recurrent neural networks with scaled Cayley transform. In *Proc. ICML*, pp. 1969–1978, 2018.

M. Lezcano-Casado and D. Martínez-Rubio. Cheap orthogonal constraints in neural networks. In *Proc. ICML*, pp. 3794–3803, 2019.

M. Reck, A. Zeilinger, H. J. Bernstein, and P. Bertani. Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.*, 73:58–61, 1994.

W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley. Optimal design for universal multiport interferometers. *Optica*, 3:1460–1465, 2016.

M. Möttönen, J. J. Vartiainen, V. Bergholm, and M. M. Salomaa. Quantum circuits for general multiqubit gates. *Phys. Rev. Lett.*, 93:130502, 2004.

V. V. Shende, S. S. Bullock, and I. L. Markov. Synthesis of quantum-logic circuits. *IEEE Trans. CAD*, 25:1000–1010, 2006.

J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.*, 9:4812, 2018.

A. Arrasmith, M. Cerezo, P. Czarnik, L. Cincio, and P. J. Coles. Effect of barren plateaus on gradient-free optimization. *Quantum*, 5:558, 2021.

D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. In *Proc. ICLR*, 2015.

I. Loshchilov and F. Hutter. SGDR: Stochastic gradient descent with warm restarts. In *Proc. ICLR*, 2017.

A Truncation-Error-Correction Method Based on a Resolution-Generalizable Neural Network for Nonlinear PDE Solvers

[1,2,3]Senwei Zheng; [1,2,3]Jiaqing Kou*; [1,2,3]Weiwei Zhang;

[1]*School of Aeronautics, Northwestern Polytechnical University, Xi'an 710072, China;*

[2]*International Joint Institute of Artificial Intelligence on Fluid Mechanics, Northwestern Polytechnical University, Xi'an 710072, China;*

[3]*National Key Laboratory of Aircraft Configuration Design, Xi'an 710072, China;*

**Corresponding author: jqkou@nwpu.edu.cn*

Keywords: truncation error correction; resolution generalization; Euler equations; incompressible Navier-Stokes equations

Abstract. We propose a deep learning approach to compensate the truncation error of traditional numerical methods for nonlinear PDEs, named as Resolution-Generalizable Neural Network (RGNN). The method augments a coarse-grid PDE solver with a machine-learned physics-based corrective forcing to achieve accuracy comparable to that of costly fine-grid PDE solvers. RGNN represents truncation error by separately modelling scale-independent differential operators and scale-dependent coefficients, which enables robust generalization across different grid resolutions. This short manuscript focuses on the two-dimensional compressible Euler equations. In the Euler isentropic-vortex test, RGNN is trained on 30 x 30 and 40 x 40 grids over t in [0,1] and evaluated up to t = 10 on grids as fine as 400 x 400, reducing the minimum-density relative error from 87.7% to 1.6% on 200 x 200 and from 77.4% to 1.0% on 400 x 400. In contrast to standard convolutional neural network for error compensation, the proposed correction significantly reduces non-physical oscillations and long-term instabilities.

1. Introduction

Numerical simulation of conservation laws governed by partial differential equations has become an essential tool for reproducing complex physical phenomena and reducing reliance on expensive experiments. High-fidelity simulations, however, usually require fine spatial discretization, causing sharp growth in degrees of freedom, memory footprint and time-integration cost. Improving the accuracy of coarse-grid simulations is therefore important, and it is closely linked to the estimation and control of discretization errors.

Classical approaches such as Richardson extrapolation, the Grid Convergence Index, truncation-error estimation and error-transport equations provide useful tools for discretization-error estimation, grid-convergence verification and adaptation. Nevertheless, these tools are still used mainly as assessment procedures rather than as direct correction mechanisms for time-dependent coarse-grid solutions.

Deep learning has recently been introduced into CFD through surrogate models and hybrid approaches. Hybrid strategies are attractive because they embed learning within classical numerical solvers, but most learned corrections entangle resolution-dependent discretization errors with resolution-independent physical structures. Since both the magnitude and the structure of discretization errors vary with grid resolution, this coupling limits cross-resolution generalization.

RGNN addresses this limitation by representing truncation error through the multiplication of scale-independent derivative terms and scale-dependent coefficients. The Resolution-Invariant Network extracts transferable derivative features through a Derivative Operator Neural Network, while the Resolution-Dependent Network represents grid-dependent coefficients through Taylor-expansion terms. This design targets cross-resolution generalization, stability through operator pretraining and filtering, and a correction coupled to the original numerical scheme.

2. Methodology

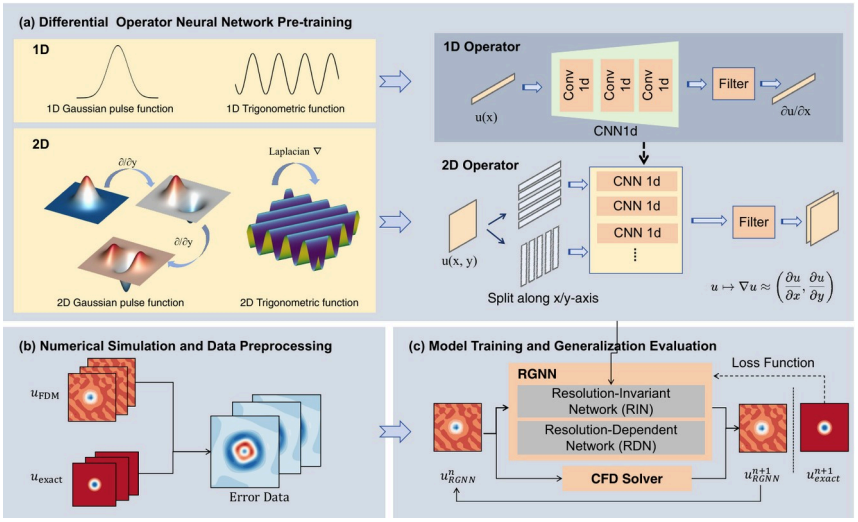


Figure 1. Training and inference workflow of RGNN. DONN is pre-trained on analytical functions, one-step error data are generated from coarse-grid and reference solutions, and the RIN-RDN correction is coupled with the CFD solver during time integration.

2.1 Corrected coarse-grid formulation

Consider a system of time-dependent PDEs on a spatial domain with continuous operator R. After spatial discretization, the numerical solution is advanced by a discrete approximation of this operator. RGNN introduces a learnable corrective forcing into the semi-discrete equation so that the added term approximates the truncation error induced by the discrete operator.

$$\frac{\partial \mathbf{u}}{\partial t} + \mathcal{R}(\mathbf{u}) = \mathbf{f}, \qquad \frac{\partial \hat{\mathbf{u}}_h}{\partial t} + \mathcal{R}_h(\hat{\mathbf{u}}_h) = \mathbf{f}$$

$$\frac{\partial \hat{\mathbf{u}}_h}{\partial t} + \mathcal{R}_h(\hat{\mathbf{u}}_h) + \mathcal{F}_\theta(\hat{\mathbf{u}}_h, h) = \mathbf{f}$$

2.2 Truncation-error structure and RGNN

The truncation error is defined as the difference between the continuous operator projected onto the discrete solution space and the discrete operator evaluated using the projected exact solution. When an exact solution is unavailable, a high-fidelity fine-grid solution is projected onto the coarse grid to construct the reference truncation error.

$$\boldsymbol{\tau}_h \equiv I_h \mathcal{R}(\mathbf{u}) - \mathcal{R}_h(I_h \mathbf{u})$$

For a local stencil, the discrete operator can be written as a weighted sum of neighboring values. Taylor expansion then reformulates the discrete operator as a series of differential terms whose coefficients depend on grid spacing. The resulting truncation error decomposes into scale-dependent coefficients and scale-independent derivative features evaluated on the grid.

$$[\mathcal{R}_h(I_h \mathbf{u})]_i = \sum_{m=0}^{\infty} A_{m,i}(h)\, \partial_x^m \mathbf{u}(x_i, t)$$

$$[\boldsymbol{\tau}_h]_i = \sum_{m=p}^{\infty} B_{m,i}(h)\, \partial_x^m \mathbf{u}(x_i, t), \qquad B_{m,i}(h) = c_m - A_{m,i}(h)$$

Based on this structure, RGNN decomposes the correction into a Resolution-Invariant Network and a Resolution-Dependent Network. The RIN extracts derivative features from the discrete flow field, and its DONN component is pre-trained using analytical functions with closed-form derivatives. The RDN uses grid spacing to construct a Taylor-expanded polynomial basis and learn coefficient weights, embedding the power-law truncation-error prior into the network.

$$\mathcal{F}_\theta(\hat{\mathbf{u}}_h, h) = \sum_{m=p}^{K} \widehat{B}_m(h)\, \widehat{\mathcal{D}}_m(\hat{\mathbf{u}}_h)$$

$$\widehat{\mathcal{D}}_{m,i} = \sum_{s=-K_f}^{K_f} G_s\, \widetilde{\mathcal{D}}_{m,i-s}, \qquad G_s = \frac{\exp[-(s\Delta x)^2/(2\sigma_f^2)]}{\sum_{n=-K_f}^{K_f} \exp[-(n\Delta x)^2/(2\sigma_f^2)]}$$

3. Numerical Experiments

All numerical simulations, network training and inference procedures were conducted in PyTorch on a platform equipped with an Intel Core i9-14900K processor and a single NVIDIA RTX 4090 GPU. This conference version reports only the two-dimensional compressible Euler isentropic vortex case requested here.

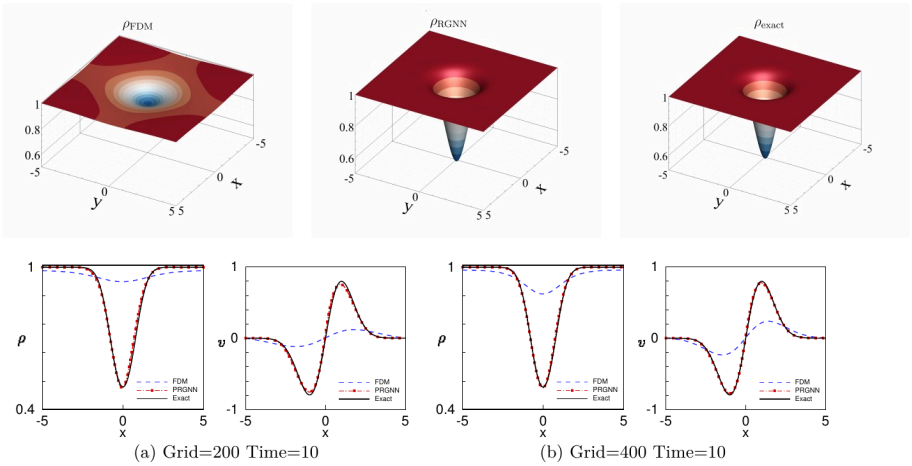


Figure 2. flow-field comparison and profiles for the two-dimensional Euler isentropic vortex at t = 10. The RGNN-corrected solution preserves the vortex amplitude and profile shape much more accurately than the baseline FDM solution on both 200 x 200 and 400 x 400 test grids.

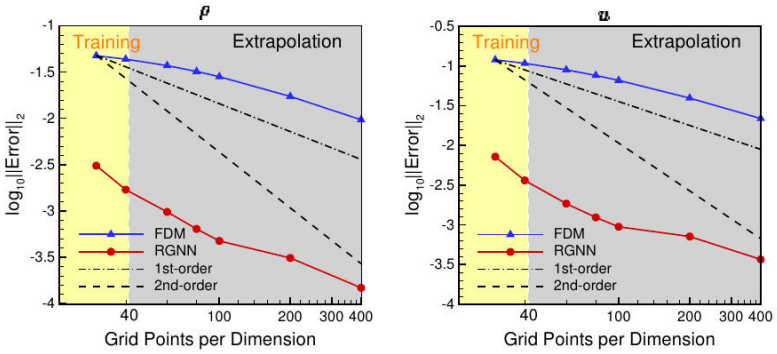


Figure 3. Grid convergence of the L2 error for the Euler isentropic vortex. The two panels correspond to density, velocities. Compared with the baseline and the standard CNN, RGNN reduces the error level while preserving the grid-convergence trend across all variables. Reference first- and second-order slopes are shown for comparison.

3.1 Compressible Euler isentropic vortex

The two-dimensional compressible Euler equations are used to assess applicability to multi-dimensional nonlinear systems with strong coupling among density, momentum and energy.The initial condition is an isentropic vortex superimposed on a uniform mean flow . With vortex strength and center .The computational domain is [0,10] x [0,10] with periodic boundary conditions. The model is trained on 30 x 30 and 40 x 40 grids using snapshots over t in [0,1.0], and it is evaluated on refined grids from 40 x 40 to 400 x 400 over time intervals up to t = 10.0. This imposes a 16-fold increase in total grid points and a tenfold temporal extrapolation relative to the training horizon.

Fig. 2 illustrates the distribution results of density and velocity at T = 10 on test grid. The results demonstrate that RGNN substantially improves both test grids. On the 200 × 200 grid, the minimum density changes from ρ_min = 0.9386 for the baseline to ρ_min = 0.492 for RGNN, while the exact value is 0.5; the relative error therefore drops from 87.7% to 1.6%. On the 400 × 400 grid, the corresponding relative errors decrease from 77.4% to 1.0%.

Fig. 3 illustrates the grid convergence of the Euler vortex L2 error, comparing with the baseline and the standard CNN, RGNN (using errors measured at T=10s as reference). The results further demonstrate that the RGNN maintains a lower error level while preserving a convergence rate essentially consistent with the baseline. Fig.4(Left) presents long-term simulations over ten convective periods, demonstrating that the output filter width governs the accuracy-stability trade-off. An appropriate filter range ensures sustained RGNN stability, whereas the standard CNN diverges after approximately one period. Fig.4(Right) illustrates the accuracy-cost trade-off, showing that the RGNN achieves the same accuracy with a 25×reduction in computational cost.

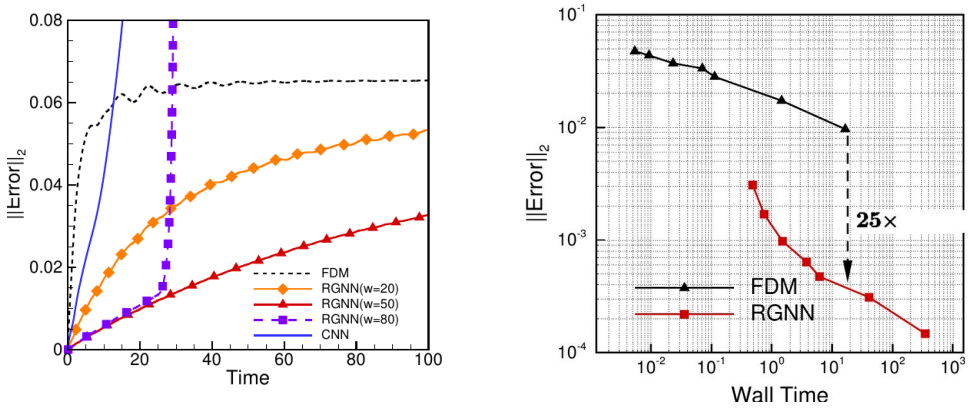


Figure 4. (Left) Temporal evolution of the global L2 error for the Euler isentropic vortex, illustrating the effect of output filtering in the RGNN. Stronger filtering improves long-time stability, weaker filtering yields lower error levels. (Right) Accuracy-cost trade-off for the Euler isentropic vortex, comparing the baseline FDM and the RGNN-corrected scheme on a logarithmic wall-time axis.

Fig5 shows modified-wavenumber comparison for a one-dimensional linear advection test. To further characterize the spectral properties of the correction operator, an approximate dispersion relation (ADR) analysis is performed based on the modified wavenumber response. The ADR results show that RGNN achieves higher dispersion-relation accuracy than the baseline and the CNN correction in the low-to-moderate wavenumber range, while the standard CNN exhibits anti-dissipative behavior at moderate wavenumbers. RGNN maintains a strictly dissipative response across all wavenumbers and reduces the excessive numerical dissipation of the baseline scheme while preserving a physically reasonable dissipation distribution. The comparison of filter parameters confirms that the spectral behavior can be directly controlled: weaker filtering yields lower dissipation and higher spectral accuracy, while stronger filtering enhances robustness by regularizing the high-wavenumber response.

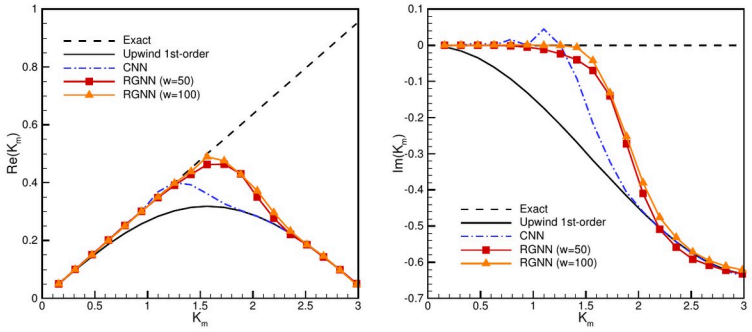


Figure 5. Modified-wavenumber comparison for a one-dimensional linear advection test. Compared with the standard CNN correction, the RGNN correction remains closer to the exact dispersion relation and avoids the anti-dissipative behavior that appears in the CNN dissipation curve. Two representative filter settings are shown to illustrate how filtering changes the spectral response of the corrected scheme.

4. Conclusions

RGNN was developed to correct coarse-grid PDE solutions by encoding the truncation-error structure into the correction. The model separates resolution-dependent coefficients from resolution-invariant differential terms, so the learned correction remains tied to both the continuous operator and the underlying discretization. This structure allows the correction learned from limited coarse-grid data to transfer to finer grids and longer integration times.

Acknowledgments

This work was supported by the National Key Research and Development Program of China (2024YFB3310401). S. Zheng and J. Kou also acknowledge support from the National Natural Science Foundation of China Excellent Young Scientists Fund (Overseas).

4. References

[1] K. Duraisamy, G. Iaccarino, H. Xiao, Turbulence modeling in the age of data, Annual Review of Fluid Mechanics 51 (2019) 357-377.

[2] S. L. Brunton, B. R. Noack, P. Koumoutsakos, Machine learning for fluid mechanics, Annual Review of Fluid Mechanics 52 (2020) 477-508.

[3] 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.

[4] I. B. Celik, U. Ghia, P. J. Roache, C. J. Freitas, Procedure for estimation and reporting of uncertainty due to discretization in CFD applications, J. Fluids Eng. 130 (2008) 078001.

[5] F. Fraysse, J. de Vicente, E. Valero, The estimation of truncation error by tau-estimation revisited, Journal of Computational Physics 231 (2012) 3457-3482.

[6] Y. Bar-Sinai, S. Hoyer, J. Hickey, M. P. Brenner, Learning data-driven discretizations for partial differential equations, Proc. Natl. Acad. Sci. U.S.A. 116 (2019) 15344-15349.

[7] D. Kochkov, J. A. Smith, A. Alieva, Q. Wang, M. P. Brenner, S. Hoyer, Machine learning-accelerated computational fluid dynamics, Proc. Natl. Acad. Sci. U.S.A. 118 (2021) e2101784118.

[8] F. Manrique de Lara, E. Ferrer, Accelerating high order discontinuous Galerkin solvers using neural networks, Journal of Computational Physics 489 (2023) 112253.

[9] H. C. Yee, N. D. Sandham, M. J. Djomehri, Low-dissipative high-order shock-capturing methods using characteristic-based filters, Journal of Computational Physics 150 (1999) 199-238.

[10] A. J. Chorin, A numerical method for solving incompressible viscous flow problems, Journal of Computational Physics 2 (1967) 12-26.

Proceedings of the 1st International Symposium on AI and Fluid Mechanics
Paper No []

Fourier Neural Operator for Journal Bearing Modelling

Paul Horvath[1,2], **Marian Staggl**[1,2], **Stefan Posch**[1,2]

[1] *CD Laboratory for Physics-driven Machine Learning in lndustrial Applications, Graz, Austria*

[2] *Institute of Thermodynamics and Sustainable Propulsion Systems, Graz University of Technology, Austria*

Corresponding author: Paul Horvath (paul.horvath@tugraz.at)

Keywords: journal bearing, Fourier neural operator, Reynolds equation

Abstract. Modeling hydrodynamic journal bearings typically relies on time-dependent numerical solvers that resolve pressure buildup, cavitation, and elastic deformation effects. While these approaches offer high accuracy, their iterative and strongly coupled nature makes large parameter studies, optimization, and real-time applications computationally expensive. In this work, we propose a purely data-driven journal bearing operator based on a Fourier Neural Operator (FNO), trained on reference numerical solutions. The training data is generated using a finite volume solver and consists of pairs of non-dimensional film thickness fields (including temporal gradients) and their corresponding pressure fields. By formulating the problem in a non-dimensional setting, the operator learns the nonlinear spatial mapping $\mathcal{G} : h(\mathbf{x}; \boldsymbol{\mu}) \mapsto p(\mathbf{x}; \boldsymbol{\mu})$ from the film thickness fields to the hydrodynamic pressure distributions across a wide range of eccentricities ($\epsilon \in [0.1, 0.95]$). The learned operator demonstrates high predictive accuracy on standard, unseen nominal configurations. Moreover, the model successfully predicts pressure distributions arising from deformation-induced film thickness variations excluded from training, highlighting its robustness and generalizability in realistic operating scenarios.

1. Introduction

In modern engineering, surrogate models represent a state-of-the-art approach to avoid the necessity of repeatedly evaluating computationally costly numerical simulations in applications such as uncertainty quantification, optimization, and real-time evaluation [1]. The modeling of hydrodynamic journal bearings relies on such numerical solvers [2] to resolve the Reynolds equation [3] and obtain the pressure distribution. To overcome this computational bottleneck [4], modern machine learning architectures offer a promising alternative. Specifically, the Fourier Neural Operator (FNO) [5] has emerged as a prominent architecture due to its elegant spectral formulation, which captures global correlations through spectral transformations while retaining resolution invariance. FNOs have been successfully applied to complex fluid modeling problems, with applications in weather prediction [6], geoscience [7], and aerodynamic design [8]. Because journal bearings operate within a closed cylindrical geometry, the resulting physical fields are inherently periodic. The FNO is naturally suited for such periodic signals, combining the predictive accuracy of detailed numerical bearing models with the computational efficiency of reduced representations such as spring-damper assumptions [9, 10]. Recently, deep learning architectures have demonstrated the potential to significantly accelerate these bearing simulations. One approach is the use of Physics-Informed Neural Networks (PINNs), as proposed by Rom [11], which embed the governing Reynolds equation directly into the network's loss function to model effects such as mass-conserving cavitation. While PINNs reduce the reliance on pre-computed training datasets, extending these networks to accommodate wider parametric design spaces significantly increases the computational burden and duration of the training phase [11], alongside the challenges of balancing complex, competing loss formulations. Alternatively, FNOs offer fast inference speeds by learning the mathematical operator mapping geometries to pressure fields. Notably, Yang et al. [12] recently deployed an FNO for supercritical CO_2 bearings, achieving inference speedups of four orders of magnitude. However, due to complex thermodynamic coupling, their model required the explicit parameterization of physical operating conditions alongside geometry, leading to a large computational effort in data generation. To extend the FNO's

applicability to hydrodynamic lubrication problems, we present an efficient, non-dimensionalized formulation tailored for oil-lubricated journal bearings. By mathematically non-dimensionalizing the Reynolds equation and applying the half-Sommerfeld cavitation condition [13], we decouple the linear operational parameters (speed and viscosity) from the underlying geometric mapping. This reduction enables the model to learn the physics using a dataset of only 200 unique eccentricity parameter configurations. Furthermore, we explicitly evaluate the model's out-of-distribution generalization on deformed shaft geometries (longitudinal bowing), assessing its robustness. The remainder of this paper is structured as follows. Section 2 outlines the theoretical framework, detailing the mathematical non-dimensionalization of the Reynolds equation and the generation of the training dataset. Subsequently, the architecture and optimization of the proposed FNO are presented, followed by the evaluation of its predictive performance on both rigid and deformed journal geometries in Section 3 and a discussion in Section 4.

2. Methodology

To develop a robust surrogate model, the Reynolds equation is first non-dimensionalized to decouple the geometric mapping from linear operational parameters. Based on this formulation, a numerical solver is implemented to generate the training data. The data generation process and FNO architecture are detailed below.

2.1. Numerical Solver and Data Generation

The dataset was generated by solving the non-dimensionalized Reynolds equation using a Finite Volume Method (FVM) solver implemented in Python. The lubricant is modeled as an incompressible, isoviscous, and Newtonian fluid. To account for cavitation, which naturally occurs in the diverging regions between the journal and the bush (see Figure 1), we applied the Gümbel (half-Sommerfeld) condition [13]. This condition truncates any negative relative pressures, strictly enforcing $p = \max(p, 0)$. While more sophisticated mass-conserving cavitation models exist [14, 15], the half-Sommerfeld approach allows for a straightforward non-dimensionalization of the governing equation and provides a baseline to validate the predictive capabilities of the FNO architecture in journal bearing modeling. To assess the trained FNO, it was evaluated across two distinct test scenarios. The first scenario uses a baseline test set containing unseen parameter configurations for a rigid shaft (see Figure 2, top). The second scenario evaluated the model's generalization capabilities by superimposing a quadratic profile over the rigid height field (see Figure 2, bottom). This quadratic profile was chosen to simulate a longitudinal shaft deformation. To maintain physical validity and prevent non-physical negative film thicknesses at maximum eccentricity ($\epsilon = 0.95$), the induced deformation levels were strictly capped at 2.0% of the maximum fluid film height.

The governing Reynolds equation for an incompressible fluid is given by:

$$\underbrace{\frac{\partial}{\partial x}\left(\frac{h^3}{12\eta}\frac{\partial p}{\partial x}\right)+\frac{\partial}{\partial y}\left(\frac{h^3}{12\eta}\frac{\partial p}{\partial y}\right)}_{\text{Poiseuille}}=\underbrace{u_m\frac{\partial h}{\partial x}}_{\text{Couette}}+\underbrace{\frac{\partial h}{\partial t}}_{\text{Displacement}}, \tag{1}$$

where $u_m = \frac{u_\mathrm{I}+u_\mathrm{II}}{2}$ is the mean surface velocity. We non-dimensionalize the spatial coordinates, time, pressure, and film height as follows: $x^* = \frac{x}{r}$, $y^* = \frac{y}{r}$, $t^* = \frac{tu_m}{r}$, $p^* = \frac{p}{p_0}$, $h^* = \frac{h}{c}$. Substituting these variables and defining the reference pressure as $p_0 = \frac{12\eta r u_m}{c^2}$ (where c is the radial clearance) yields the simplified non-dimensional Reynolds equation:

$$\frac{\partial}{\partial x^*}\left(h^{*3}\frac{\partial p^*}{\partial x^*}\right)+\frac{\partial}{\partial y^*}\left(h^{*3}\frac{\partial p^*}{\partial y^*}\right)=\frac{\partial h^*}{\partial x^*}+\frac{\partial h^*}{\partial t^*}. \tag{2}$$

To obtain the pressure field p for a given non-dimensional height field h^*, we discretize the equation using

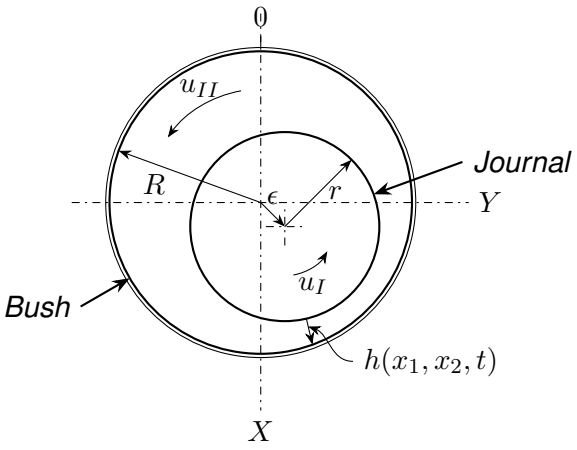


Figure 1: Cross-section of a hydrodynamic journal bearing.

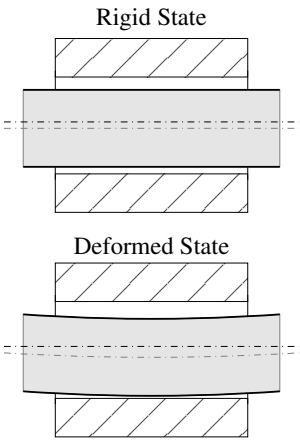


Figure 2: Rigid journal bearing (top), imposed parabolic form error (bottom)

FVM, solve the resulting system, and rescale the result ($p = p^* p_0$). We applied periodic boundary conditions (BCs) in the circumferential x-direction and Dirichlet BCs at the axial boundaries ($p^*(x, \Omega_y) = 0$). Shaft eccentricity ϵ was sampled uniformly within predefined bounds ($\epsilon \in [0.1, 0.95]$) using Latin Hypercube Sampling (LHS). The journal radius R was fixed to 20.0×10^{-3} m with an L/D ratio of 0.5. Radial clearance was fixed to $c = R \times 10^{-3}$, according to [16], and the rotational velocity of the bush was set to $u_{II} = 0$ for simplicity. For each of the 200 sampled eccentricity configurations, the pressure field was solved across 100 uniformly distributed angular orientations of the shaft, yielding a total dataset of 20,000 input-output field pairs. The dataset was partitioned into an 80/10/10 train/validation/test split based on the initial parameter configurations. Through non-dimensionalization, the linear scaling effects of rotational speed and dynamic viscosity were decoupled from the dataset generation. This allowed the model to focus entirely on learning the non-linear spatial mappings, with the physical pressure field subsequently recovered via the scaler p_0. Finally, the non-dimensional temporal height derivative was computed using a backward finite difference scheme with a dimensionless time step of $\Delta t^* \approx 2.6 \times 10^{-5}$.

2.2. Model setup

The model's input consists of three fields: (i) a non-dimensionalized film height field, (ii) the corresponding temporal height gradient field, and (iii) a parameter encoding composed of sine and cosine functions. Figure 3 illustrates the learning setup, where the FNO approximates the non-linear mapping $\mathcal{G} : (h^*, \nabla h^*, \theta) \mapsto p^*$.

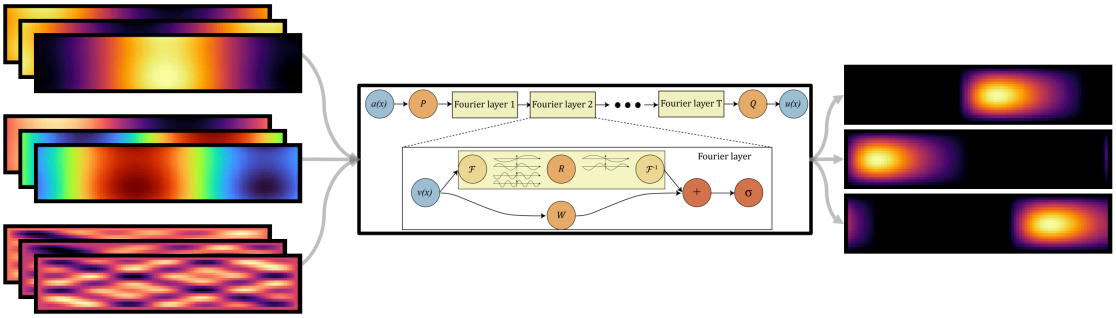


Figure 3: FNO training pipeline. Left (top to bottom): input fields comprising height, height gradient, and parameter-encoded geometry. Center: FNO architecture. Right: predicted pressure distribution.

Hyperparameter optimization was executed using Optuna (Tree-structured Parzen Estimator, 50 trials) to minimize the validation Mean Squared Error (MSE). The optimal configuration yielded a validation MSE of 7.04×10^{-5} using the following parameters:

- **Architecture:** 4 layers, 24 hidden channels; Fourier modes: 10 (x-dimension), 6 (y-dimension).
- **Training setup:** Adam optimizer, batch size 64, maximum 500 epochs.
- **Learning rate:** Initial 1.4×10^{-3}, exponential decay rate 0.98 per epoch.
- **Regularization:** Weight decay 9.2×10^{-4}, dropout rate 0.25.
- **Convergence criteria:** Early stopping with a patience of 25 epochs.

For comparative analysis, the FNO was benchmarked against Multi-Layer Perceptron (MLP) networks utilizing flattened spatial inputs. To ensure a comprehensive evaluation, two separate MLP baselines were constructed: a parameter network using the same amount of parameters as the FNO ($\approx$ 100000 parameters) for an equivalent comparison, and a high-capacity network (scaled to roughly ten times the parameter count) to test performance under over-parameterized conditions. The hyperparameters for both MLP architectures were independently optimized via Optuna to maximize their respective predictive accuracies. The optimal MLP configurations and training parameters are detailed in Appendix A.

3. Results

The trained models were evaluated on the held-out test set. To achieve physical interpretation and visualization, the non-dimensional model predictions (p^*) were mapped back to the dimensional pressure field ($p = p^* p_0$) via the scalar $p_0 = \frac{12\eta r u_m}{c^2}$. For these visualizations, the dynamic viscosity and rotational speed were fixed at $0.008\,\mathrm{Pa\,s}$ and $2000\,\mathrm{rpm}$, respectively, reflecting typical passenger vehicle engine operating conditions [16, 17]. Crucially, because the neural operator learns a purely non-dimensional mapping, its predictive validity is not bound to these specific operational parameters. Evaluating new kinematic conditions requires only a recalculation of the scalar p_0. When evaluated across the undeformed non-dimensional test set, the FNO achieved an overall MSE of 2.76 $\times 10^{-5}$. In contrast, the iso-parameter MLP baseline stalled at an MSE of 3.5×10^{-2}, with the high-capacity MLP achieving 9.4 $\times 10^{-4}$. This performance disparity is attributed to the MLP's lack of spatial awareness. By flattening the input grid, the network loses the spatial correlations required to capture global dynamics. The corresponding parity plot comparing the predicted and true pressures for both models is presented in Figure 4, while the full spatial pressure field predicted by the FNO is illustrated in Figure 5.

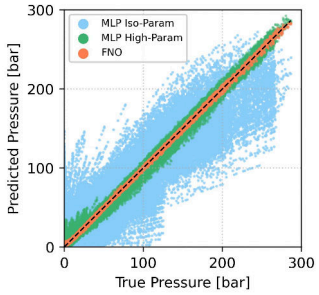


Figure 4: Parity plot comparing the predictive performance of the FNO, the iso-parameter MLP (≈100k parameters), and the scaled MLP (≈1M parameters) on the test set.

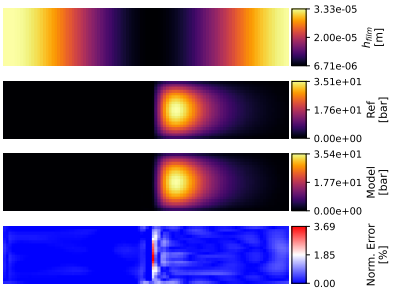
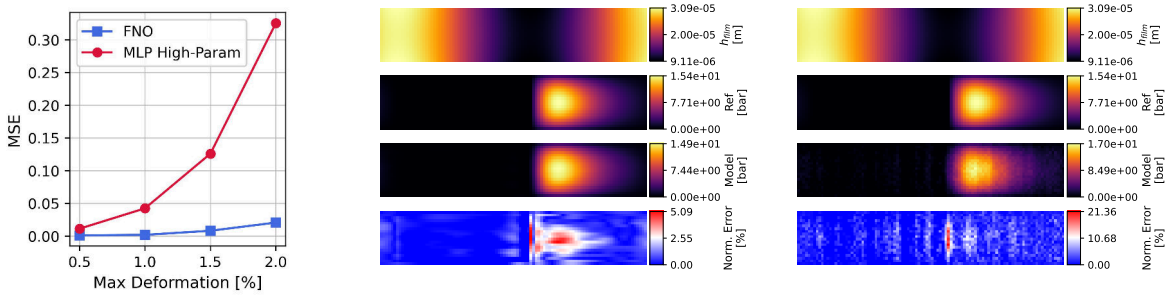


Figure 5: FNO prediction fields showing (top to bottom): height field, ground truth, model prediction, and normalized absolute error (percentage of peak ref. pressure)

To evaluate the model's ability to generalize to deformed shafts, we checked MSE loss for increasing peak deformations, as explained in Section 2.1. Figure 6 below illustrates generalizability of the FNO compared to the high-parameter MLP. The FNO shows superior generalization capabilities with a tenth of the total parameter count on the deformed test set compared to the MLP implementation, reaching significantly lower MSE values across increasing deformation values.



(a) Error Trend (b) FNO Prediction (c) MLP Prediction

Figure 6: Generalization performance on the deformed test set. (a) Mean Squared Error (MSE) of the FNO and the high-parameter MLP as a function of maximum shaft deformation. (b) Predictive fields generated by the FNO for a deformed sample with 2% max. deformation. (c) Corresponding prediction fields from the high-parameter MLP.

4. Discussion and conclusion

This study investigates the applicability of FNO as surrogate models for hydrodynamic journal bearing analysis. By non-dimensionalizing the governing Reynolds equation, the operational parameters of speed and viscosity were fully decoupled from the geometric input space, allowing the problem to be formulated in a universal way. In addition to evaluating the model on the standard test set, we tested its robustness in non-ideal settings by applying the learned operator $\mathcal{G} : h(\mathbf{x}; \boldsymbol{\mu}) \mapsto p(\mathbf{x}; \boldsymbol{\mu})$ to a dataset containing perturbed shaft geometries simulating longitudinal bowing. Crucially, the FNO predicted the pressure fields of the deformed dataset with high accuracy despite never being trained on such topologies, highlighting the operator's extrapolative capabilities. Unlike the MLP baseline, the FNO consistently maintained its predictive precision across increasing peak deformations. Overall, these investigations demonstrate that for oil-lubricated journal bearing modeling, FNOs provide a promising alternative to traditional solvers that successfully adapts to geometric deviations. Future work will investigate the incorporation of physics-informed loss formulations directly into the training phase to further enforce physical consistency and reduce data demand.

Acknowledgments

The financial support by the Austrian Federal Ministry of Economy, Energy and Tourism, and the Christian Doppler Research Association is gratefully acknowledged.

References

[1] G. I. Schuëller, "On the treatment of uncertainties in structural mechanics and analysis," *Computers & Structures*, vol. 85, no. 5-6, pp. 235–243, 2007.

[2] B. J. Hamrock, S. R. Schmid, and B. O. Jacobson, *Fundamentals of Fluid Film Lubrication*. CRC Press, 2 ed., 2004.

[3] O. Reynolds, "On the theory of lubrication and its application to mr. beauchamp tower's experiments, including an experimental determination of the viscosity of olive oil," *Philosophical Transactions of the Royal Society of London*, vol. 177, pp. 157–234, 1886.

[4] J. S. Larsen and I. F. Santos, "Efficient solution of the non-linear reynolds equation for compressible fluid using the finite element method," *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, vol. 37, no. 3, pp. 945–957, 2015.

[5] Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar, "Fourier neural operator for parametric partial differential equations," 2021.

[6] J. Pathak, S. Subramanian, P. Harrington, S. Raja, A. Chattopadhyay, M. Mardani, T. Kurth, D. Hall, Z. Zhai, A. Anandkumar, *et al.*, "Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators," *arXiv preprint arXiv:2202.11214*, 2022.

[7] G. Wen, Z. Li, K. Azizzadenesheli, A. Anandkumar, and S. M. Benson, "U-fno: An enhanced fourier neural operator-based deep-learning model for multiphase flow," *Advances in Water Resources*, vol. 163, p. 104180, 2022.

[8] Z. Li, D. Z. Huang, B. Liu, and A. Anandkumar, "Fourier neural operator with learned deformations for pdes on general geometries," *Journal of Machine Learning Research*, vol. 24, no. 388, pp. 1–26, 2023.

[9] D. W. Childs, "Turbomachinery rotordynamics: Phenomena, modeling, and analysis," 1993.

[10] M. Lucassen, G. Jacobs, and B. Lehmann, "Efficient dynamic simulations of planetary journal bearings in wind turbine gearboxes," *Forschung im Ingenieurwesen*, vol. 89, 2 2025.

[11] M. Rom, "Physics-informed neural networks for the reynolds equation with cavitation modeling," *Tribology International*, vol. 179, p. 108141, 2023.

[12] J. Yang, D. Shin, and A. Palazzolo, "Fourier neural operator for flow-induced rotordynamics force prediction and application to a sco2 magnetic bearing-rotor system," *Mechanical Systems and Signal Processing*, vol. 232, p. 112750, 2025.

[13] L. Gümbel, "Vergleich der ergebnisse der rechnerischen behandlung des lagerschmierungsproblems mit neueren versuchsergebnissen," *Monatsblätter des Berliner Bezirksvereins deutscher Ingenieure (VDI)*, pp. 125–128, 1921. In German.

[14] B. Jakobsson and L. Floberg, "The finite journal bearing, considering vaporization (das gleitlager von endlicher breite mit verdampfung)," in *Transactions of Chalmers University of Technology*, 1957.

[15] K.-O. Olsson, "Cavitation in dynamically loaded bearings," in *Transactions of Chalmers University of Technology*, 1965.

[16] M. M. Khonsari and E. R. Booser, *Journal Bearings*, ch. 8, pp. 255–328. John Wiley & Sons, Ltd, 2017.

[17] Fuels and Lubricants TC 1 Engine Lubrication, *Engine Oil Viscosity Classification*. SAE International, 5 2024.

A. Hyperparameters

Table 1: Optimized hyperparameters for the baseline MLP architectures, determined via hyperparameter optimization.

Hyperparameter	**Iso-parameter MLP**	**High-parameter MLP**
Hidden Neurons (N_{hidden})	12	118
Hidden Layers (N_{layers})	4	5
Activation Function	GELU	GELU
Initial Learning Rate	1.23×10^{-3}	6.13×10^{-4}
Scheduler Decay Rate	0.995	0.982
Weight Decay	2.54×10^{-8}	9.11×10^{-7}
Dropout Rate	0.0	0.0
Total parameters	98,480	1,002,286

Influence of temperature on concentration convection caused by instability of the equilibrium of ternary gas mixtures based on data using CFD modeling improved with the help of artificial intelligence

V N Kossov[1], V Mukamedenkyzy[2] and A G Tolepbergen[2]

[1]*Abai Kazakh National Pedagogical University, Almaty, Kazakhstan*
[2]*Al Farabi Kazakh National University, Almaty, Kazakhstan*

Keywords: multicomponent diffusion, concentration convection, temperature effects, CFD modeling, AI-assisted calibration, mutual diffusion coefficients, instability, pressure.

Abstract. The influence of temperature on combined mixing processes in multicomponent gas systems at high pressures is a key factor determining the intensity of partial mass transfer and the development of instability caused by the violation of mechanical equilibrium of mixtures. Despite the existence of classical computational and theoretical models, the quantitative description of temperature-concentration effects in real multicomponent mixtures remains difficult and requires the comparison of experimental results with numerical data. This paper presents experimental and numerical studies on diffusion and concentration-driven gravitational convection in a H_2–N_2 –CH_4 gas system in vertical cylindrical channels with fixed geometric characteristics and specified compositions in the temperature range from 276 to 353 K at high pressures. The experimental data were used to validate the CFD model of component transport implemented within the framework of a multicomponent diffusion approach. To reconcile the calculated and experimental results, an AI-assisted approach to calibrating mutual diffusion coefficients was applied, based on iterative data-driven adjustment of transport parameters without the use of explicit machine learning models. This approach allows for effective consideration of the temperature dependence of diffusion characteristics while maintaining the physical interpretability of the model. It is shown that in the temperature range 276 K – 353 K, satisfactory agreement between experimental and numerical results is achieved, confirming the applicability of the proposed methodology for analyzing concentration-temperature effects in problems of multicomponent gas diffusion and convective instability at high pressure.

Accelerating high-order discontinuous Galerkin simulation by coupling differentiable solver with the corrective forcing and the reconstruction

[1]Xukun Wang*; [1] Oscar A. Marino; [1,2] Esteban Ferrer;

[1] *ETSIAE-UPM-School of Aeronautics, Universidad Politécnica de Madrid, Plaza Cardenal Cisneros 3, E-28040 Madrid, Spain*

[2] *Center for Computational Simulation, Universidad Politécnica de Madrid, Campus de Montegancedo, Boadilla del Monte, 28660 Madrid, Spain*

Abstract

High-order Discontinuous Galerkin Spectral Element Methods (DGSEM) offer exceptional accuracy for complex flow simulations; however, their computational overhead scales aggressively with polynomial order. This work proposes a differentiable DGSEM solver designed to achieve high-order fidelity using low-order computational costs through two distinct neural network (NN) strategies.

Approach 1: Solver-in-the-Loop Correction

The first approach introduces a learned corrective forcing term applied to low-order simulations. Leveraging the solver's full differentiability, we employ gradient-based optimization and interactive (solver-in-the-loop) training. This method effectively mitigates the "data-shift" problem common in static, offline learning. Validations on the 1D viscous Burgers' equation and 2D decaying homogeneous isotropic turbulence (DHIT) demonstrate that interactive training with extended unrolling horizons significantly enhances both numerical precision and long-term stability.

Approach 2: NN-Based PnPm Reconstruction

The second approach draws inspiration from PnPm and reconstructed DG (rDG) methods. We derive the exact formulation of the corrective term and utilize solution reconstruction to recover high-order accuracy. While classical least-squares reconstructors often introduce excessive dissipation—smearing high-order components—we propose a pretrained NN-based reconstructor. This "p-super-resolution" model learns the mapping from low-order Legendre coefficients within a stencil to their high-order counterparts. Our results indicate that the NN-reconstructor holds significant potential, particularly in under-resolved regimes, providing a robust framework for accelerating high-fidelity fluid dynamics simulations.

(1747)

Keywords: High order discontinuous Galerkin; Differentiable solver; Neural Networks (NN); Corrective forcing; Reconstruction.

Two sesstions: 1. Machine Learning; 8. PDE Solvers;

Accelerating Kinetic Fokker-Planck simulations using a GPU-optimized deep neural network surrogate: Application to rarefied internal and hypersonic external flows

Ehsan Roohi, Francis Padilla

Mechanical and Industrial Engineering, University of Massachusetts Amherst, 160 Governors Dr., Amherst, MA 01003, USA

Particle-based Fokker-Planck (FP) models provide precise simulations of rarefied gas dynamics but face a major computational challenge: the "closure problem," which involves solving complex linear systems for high-order moments at each timestep. This research introduces a novel approach that replaces the traditional physics-based closure solver with a GPU-optimized Deep Neural Network (DNN) surrogate. To avoid Input/Output (I/O) delays, the trained model parameters are extracted and run as a lightweight, batched matrix multiplication directly within the GPU simulation loop. The method is thoroughly tested across various flow regimes. In 1-D Couette flow, it shows excellent interpolation and extrapolation over a wide range of Knudsen numbers (Kn ≈ 0.0015-0.7). In 2-D lid-driven cavity tests, the surrogate accurately predicts complex recirculating flows, achieving speedups of 1.63×- 1.73×, close to the maximum predicted by Amdahl's Law. For external hypersonic flow over a cylinder, the framework captures extreme nonequilibrium features, such as the detached bow shock and surface coefficients (Cf, Cp, Ch). The GPU-native surrogate achieves a 3.85× overall speedup over the baseline solver, effectively removing the closure bottleneck and accurately modeling the detached bow shock. These results demonstrate that the GPU-native surrogate incurs virtually no additional cost relative to particle movement, shifts the computational bottleneck away from the closure solver, and preserves physical accuracy in both internal and high-speed external flows.

Super-Resolution Predictive Initialization Using Neural Operators for Accelerating CFD Simulations

[1]Ruihan Zhang, [1]Jingru Xing, [2]Shoubao Dai*, [1]Yuanyi Wu, [2]Shangwen Gao, [1]Junfu Lyu, [1]Yuxin Wu*
[1]*Department of Energy and Power Engineering, Key Laboratory of Thermal Science and Power Engineering, Ministry of Education, Tsinghua University, Beijing, China*
[2]*Harbin Ship Boiler Turbine Research Institute, State Key Laboratory of Marine Thermal Energy and Power, Harbin, China.*

Keywords: CFD convergence acceleration, Predictive initialization, Neural operator

Abstract. This paper proposes a convergence acceleration method based on predictive initialization. A neural operator was employed to predict the approximately converged fields of all primary variables in the governing equations. The predicted results are then used as initial fields for CFD simulations, which subsequently proceed to full convergence through numerical solving to guarantee numerical accuracy. Such hybrid approach suppresses the drastic change in fields at the beginning of computation. Leveraging the mesh-invariant property of neural operators, our method realizes zero-shot super-resolution initialization on unstructured meshes. The acceleration performance is demonstrated through a condenser simulation problem with varying design parameters, where the multi-physics CFD simulations with nonlinear source terms are solved using finite-volume method. Experiments showed that the maximum acceleration was achieved when all primary variables as well as key source terms were predictive initialized. Compared with constant initialization, our approach achieves a median 3.0× speedup on 9.3× higher resolutions, significantly reducing the time required for high resolution simulations.

1. Introduction

Computational Fluid Dynamics (CFD) is indispensable for modern industrial design, yet its application in PDE-constrained optimization and flow control is often hindered by prohibitive computational costs. To mitigate this, supervised machine learning (ML) surrogate models—ranging from radial basis functions to deep neural networks—have been developed to provide rapid flow field approximations. While these models offer near-instantaneous inference, their efficacy is frequently constrained by data scarcity and poor generalization in complex engineering systems. Consequently, a promising direction is to combine ML's efficiency with the physical rigor of numerical solvers. Existing approaches include coarse-grid correction[1-3], reducing iteration[4-6], and the replacement of specific solvers[7]. However, these methods typically require high-precision predictions to maintain stability and ensure effective speedup.

In this study, we focus on reducing numerical simulations required to achieve convergence in a steady-state problem. The key idea is to provide predicted fields for CFD initialization. Such strategy relaxes accuracy requirements for surrogate models since the numerical solver performs the final refinement and is relatively robust to the initial conditions. Previous studies on predictive initialization mostly rely on convolutional neural networks[6, 8], which restricts models to fixed-resolution structured meshes and precludes generalization across varying mesh points. Unstructured meshes are the mainstream of industrial applications. To achieve generalization of any unstructured meshes, the surrogate model needs to be more flexible in describing the field variables.

In this study, we leverage point-based neural operator to achieve a mesh-independent predictive initialization strategy. The flow field is described by a point cloud consist of cell centres and boundary face centres. A Transolver[9] neural operator trained with data predicts initial fields for all transport variables from design parameters and spatial coordinates. The predicted fields are used to initialize CFD solver and are refined to numerical accuracy. We test our strategy on a condenser simulation problem, in which a strongly coupled nonlinear source term has great influence on convergence behaviour. We found the source term also needs a proper initial field to guarantee effective acceleration. Validation on condenser cases demonstrates a 2–3 acceleration across diverse resolutions without compromising numerical accuracy. On super-resolution meshes, the method achieves a consistent acceleration rate.

2. Methodology

Ruihan Zhang: zhang-rh23@mails.tsinghua.edu.cn
Yuxin Wu: wuyx09@mail.tsinghua.edu.cn*

The convergence behaviour of steady-state is sensitive to initial fields. Initializing the flow field with values proximate to the final solution can reduce the simulation difficulty of complex flow problems. The proposed predictive initialization framework is illustrated in Figure 1. Variables in a CFD solver can be categorized into primary variables, source terms and auxiliary variables based on their mathematical roles: Primary variables are directly solved from governing equations through iterative updates. In flow problems, the transport variables are primary variables. Source terms in the governing equations are strongly coupled with primary variables and are critical for stability. In this study, the source terms are given in the form of implicit algebraic equations, which are solved through under-relaxed updates in the numerical solver. Auxiliary variables refer to those do not explicitly appear in the governing equations and have simple expressions.

To achieve an accelerated convergence, all the primary variables and source terms should be initialized to a state closer to the converged solution. For primary variables, we use surrogate models to predict their solutions and initialize their value at the cell centres. The network takes Cartesian coordinates $\boldsymbol{x}$ and design parameters $\boldsymbol{\beta}$ as input and outputs variable value at corresponding points (equation (1)). Source terms are not predicted by surrogate models. Instead, they are computed by their expressions based on the predicted values of the primary variables. To eliminate the error of initial guess, source terms having implicit expressions are updated multiple times during initialization (equation (2)).

$$\phi_{\text{init}} \leftarrow \hat{\phi} = \text{NN}_{\phi}(\boldsymbol{x}, \boldsymbol{\beta}; \boldsymbol{\theta}_{\phi}) \tag{1}$$

$$S_{\text{init}} \leftarrow \hat{S} = f_S(\hat{\phi}, S_0) \tag{2}$$

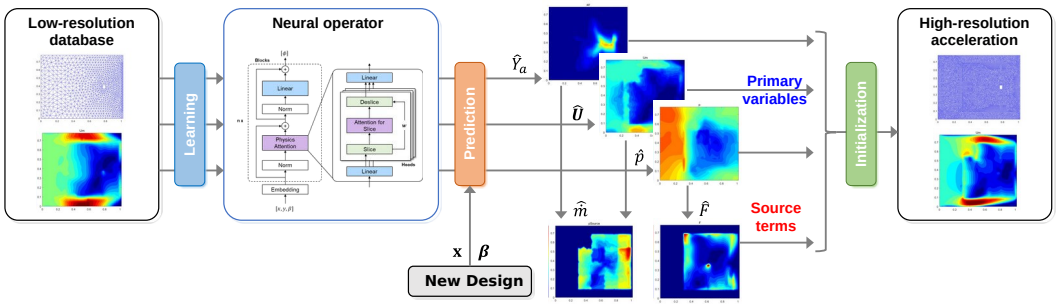


Figure 1 Framework of predictive initialization method. The predicted variables include all primary variables, while source terms are computed based on primary variables.

We adopt Transolver[9], a Transformer-based neural operator architecture, to perform point-based flow field prediction. The core innovation of Transolver lies in its Physics-Attention mechanism, which addresses the computational inefficiency of standard self-attention when applied to dense mesh points. In this study, the predictive network consists of 8 stacked Transolver blocks with 16 attention heads each. The hidden dimension is set to 256, and GELU activation function is employed throughout the architecture.

3. Data acquisition and preprocessing

In design optimization, while boundary conditions or design parameters vary, the governing equations remain consistent. This consistency results in similarity of solutions across design iteration, allowing the surrogate model to learn from previous simulation data and estimate initial fields for new configurations. In this study, the predictive initialization is applied to accelerate flow simulations in shell-and-tube condensers with different design configurations. The governing equations of this problem includes NS equations and a species transport equation.

$$\nabla \cdot (\beta \rho \mathbf{U}) = -\dot{m} \tag{3}$$

$$\nabla \cdot (\beta \rho \mathbf{U}\mathbf{U}) = \nabla \cdot \bar{\bar{\tau}} - \beta \nabla p - \dot{m}\mathbf{U} - \beta \mathbf{F} \tag{4}$$

$$\nabla \cdot (\beta \rho Y_a \mathbf{U}) = \nabla \cdot (\beta \rho D_e \nabla Y_a) \tag{5}$$

The design variation of the condensers is modelled as adjustment of local porosity values in a partitioned region. Training data were generated from five partition patterns, each with 480 sets of porosity values generated through porosity optimization processes (If porosity values are generated randomly, it will more likely to result in non-convergent cases). After filtering out non-convergent cases, the dataset was split into training, validation, and testing sets with a ratio of 8:1:1.

A pronounced skewness in data scale distribution was observed in pressure fields. To avoid scaling most data into a narrow range, we adopted probability integral transform on pressure dataset before feeding into network. This method maps the original variable distribution to a uniform distribution on (0, 1) using the cumulative distribution function (CDF) of the variable estimated from the training dataset. This nonlinear transformation effectively eliminates skewness and enhances sample distinguishability, thereby facilitating neural network training. Other variables are scaled through minmax or zscore normalization.

Data are generated on a 1902-cells unstructured mesh. To prevent network from overfitting the fixed sampling coordinates, the cell points in each case are randomly divided into several subsets to introduce coordinate variability. Three independent Transolver networks were trained to predict pressure, velocity, and air mass fraction using MSE loss and the Adam optimizer. Training was conducted for 2000 epochs with a learning rate of 0.001 and a batch size of 32.

4. Results

The condenser flow field surrogate model was embedded into the initialization process of our homemade CFD solver. The solver monitors convergence via scaled maximum residuals, with a convergence criterion of . Initialization using constant primary variables and source terms serves as the baseline. To quantify efficiency gains, the Acceleration Rate (AR) is defined as ratio of iterations needed for convergence between constant and predictive initialization.

The acceleration effect was first evaluated on the test dataset with different β and training mesh resolution (1902-cells). Figure 2 shows the AR and variables' prediction error distribution, the median AR reached 2.3 on test cases. However, there is no significant correlation between the calculated acceleration ratio and the predicted error. On this steady-state problem, initial fields do not affect the final convergent solution, but we found some cases that diverge with constant initialization can achieve convergence under predictive initialization.

Predictions were directly made on denser meshes without any fine-tuning to test the super-resolution acceleration effect. As shown in Figure 3, a median AR of 1.9 is achieved on 4580-cell mesh, and 3.0 on 17654-cell mesh. Figure 4 shows the evolution of the flow fields during the convergence process. The predicted initial field, even if not perfectly accurate, provides variable fields already close to the converged state, thereby effectively reduce the oscillation of flow fields during the warmup stage. As the demand of computation resource increases with the number of mesh cell, high-resolution acceleration saves much more time than low-resolution acceleration under the same AR.

[4] R. Djeddi, A. Kaminsky, and K. Ekici, Convergence Acceleration of Fluid Dynamics Solvers Using a Reduced-Order-Model. 2017.
[5] C. Ning, J. Kou, and W. Zhang, "An effective convergence accelerator of fluid simulations via generative diffusion probabilistic model," Aerospace Science and Technology, vol. 158, 2025, doi: 10.1016/j.ast.2024.109917.
[6] O. Obiols-Sales, A. Vishnu, N. Malaya, and A. Chandramowlishwaran, "CFDNet: a deep learning-based accelerator for fluid simulations," arXiv e-prints, p. arXiv:2005.04485, 2020, doi: 10.48550/arXiv.2005.04485.
[7] J. Tompson, K. Schlachter, P. Sprechmann, and K. Perlin, "Accelerating eulerian fluid simulation with convolutional networks," in International conference on machine learning, 2017: PMLR, pp. 3424-3433.
[8] X. Chen et al., "Developing an advanced neural network and physics solver coupled framework for accelerating flow field simulations," Engineering with Computers, vol. 40, no. 2, pp. 1111-1126, 2024/04/01 2024, doi: 10.1007/s00366-023-01861-4.
[9] H. Wu, H. Luo, H. Wang, J. Wang, and M. Long, "Transolver: A fast transformer solver for pdes on general geometries," arXiv preprint arXiv:2402.02366, 2024.