PRELIMINARY PREPRINT VERSION – DO NOT CITE
This is not the version of record. The final version will appear in the Proceedings of IJCAI-ECAI 2026.

UNOP: Physics-Constrained Unsupervised Neural Operator for Long-Horizon PDE Learning on Generalized Geometries

Xinrui Cheng[1] , **Tianqi Zhao**[1] , **Zhaodong Zhang**[1] , **Ngai Wong**[2] , **Zhongjie Wang**[1] and **Ruihan Hu**[1,*]
[1]Harbin Institute of Technology
[2]The University of Hong Kong
cxrxin0530@163.com, ztq@stu.hit.edu.cn, 25s003081@stu.hit.edu.cn, nwong@eee.hku.hk, rainy@hit.edu.cn, rh.hu@hit.edu.cn

Abstract

Unsupervised learning of neural operators is constrained by numerical instability, causing predictions to diverge in long-horizon rollouts. To address this, we present a physics-constrained unsupervised neural operator for long-horizon PDE learning on generalized geometries (UNOP). This framework replaces differential constraints with integral consistency for stable, label-free learning. Unlike prior works, UNOP is built upon Latent Integral Physics Embedding (LIPE), which enforces physical consistency through integral constraints. To extend integral formulations to generalized geometries, the Geometry-Agnostic Latent Adapter (GALA) projects them onto a unified latent grid of PDE inputs, providing a regularized domain for spatial integral evaluation. Based on this shared embedding, the Gated Spectral Evolution Operator (GSEO) performs stable temporal integration while retaining spatial regions with sharp gradients and fine-scale structures, with the evolution constrained by the LIPE objective. Experiments on 1D, 2D, and 3D benchmarks show UNOP outperforms state of the art methods, reducing error accumulation by up to 60% in 20-step rollouts. Code is available at https://github.com/chengxinrui/UNOP.

1 Introduction

The mathematical basis for explaining physical laws is provided by partial differential equations (PDEs), which have numerous uses in energy optimization, aircraft design, and climate prediction [Raissi *et al.*, 2019; Karniadakis *et al.*, 2021], as well as emerging digital twin paradigms for 6G and space-ground integrated networks [Wang *et al.*, 2023; Chen *et al.*, 2023]. These applications involve complex spatiotemporal dynamics and generalized geometries, making traditional numerical solvers prohibitive for real-time tasks, especially in resource-constrained scenarios like mobile edge computing [Gao *et al.*, 2019]. Inspired by the success of large-scale pre-trained models [Zeng *et al.*, 2022], the exorbitant cost of labeled data is a major obstacle to scaling neural PDE solvers: each high-fidelity CFD/FEM simulation may need hundreds of GPU hours, rendering supervised techniques unsustainable. Consequently, shifting from label intensive supervised paradigms to physics-constrained unsupervised learning has become a necessity for scalable solvers.

Under the data-driven paradigm, three fundamental challenges emerge: (i) **Label scarcity**: Supervised neural operators require thousands of labeled trajectories, but generating each via CFD/FEM is expensive [Herde *et al.*, 2024; Wu *et al.*, 2024], creating a scalability bottleneck; (ii) **Geometric constraints**: Real-world domains such as vehicle surfaces, airfoils, and engine intakes [Li *et al.*, 2023a; Li *et al.*, 2023b; Lam *et al.*, 2023] exhibit irregular geometries represented by unstructured meshes or point clouds, making them incompatible with the regular grids required by spectral operators [Li *et al.*, 2021]; (iii) **Long-horizon instability**: Autoregressive rollout accumulates errors rapidly, often causing predictions to diverge within 10–20 steps [Brandstetter *et al.*, 2022; Lippe *et al.*, 2023]. These challenges compound: irregular geometries make standard derivative computation unstable, tainting physics residuals and accelerates long-horizon divergence.

Recent research pursues three directions. Spectral neural operators [Li *et al.*, 2021] like FNO achieve efficient global modeling, but depend on regular grids; handling irregular geometries requires coordinate mappings [Li *et al.*, 2023a] that incur significant computational cost. Geometric deep learning [Bronstein *et al.*, 2021] leverages GNNs [Fortunato *et al.*, 2022] and Transformers [Hao *et al.*, 2023] to handle generalized geometries. However, these approaches are impractical when labeled data is scarce or costly. To bypass the dependency on labeled data, Physics-informed learning (PINNs) [Raissi *et al.*, 2019] allows for **unsupervised learning** using PDE residuals; however, it is prone to memory inflation and gradient pathologies when computing higher-order derivatives [Yu *et al.*, 2022]. Alternatives like variational formulations [Kharazmi *et al.*, 2021] and probabilistic models [Zhang *et al.*, 2025] mitigate oscillations but not the memory cost. However, these methods neglecting temporal structure suffer error accumulation. Crucially, current paradigms impose an trade-off: spectral methods sacrifice geometric flexibility for efficiency, geometric learners demand expensive supervision, and PINNs compromise stability for label-free training. As a result, these requirements remain difficult to jointly achieve within a single model.

To overcome these issues, we explore physics restrictions from a holistic standpoint. Traditional differential residuals require pointwise derivatives [Raissi *et al.*, 2019], which are unstable on irregular geometries. Instead, we reformulate PDE evolution as state propagation via integral consistency. This replaces differential residuals with integral consistency constraints, avoiding derivative computation and enabling label-free learning. This integral formulation, developed for advection-diffusion systems [Karniadakis *et al.*, 2021], enhances stability for parabolic PDEs. We bridge the gap between spectral operators and geometric learning by projecting generalized geometries onto a shared regular latent space, combining spectral efficiency and geometric flexibility.

To realize this vision, we present the **Physics-Constrained Unsupervised Neural Operator for Long-Horizon PDE Learning on Generalized Geometries (UNOP)**. This framework delivers three key contributions:

- **Latent Integral Physics Embedding (LIPE):** A derivative-free objective enforcing physical laws via stochastic spatiotemporal integral sampling. Designed for advection-diffusion systems, it replaces unstable differential residuals to enable robust unsupervised learning without simulation labels.
- **Geometry-Agnostic Latent Adapter (GALA):** A unified interface that projects generalized geometries into a regular latent space by aggregating features from grid and their neighboring vertices. This latent representation provides a consistent domain for LIPE, enabling accurate evaluation of spatial integrals while decoupling geometric complexity from operator computation.
- **Gated Spectral Evolution Operator (GSEO):** A spectral evolution module employing curvature-aware gating on spectral features to preserve spatial regions with high curvature. It accurately captures multiscale dynamics and maintains shocks and discontinuities in the solution field during long-horizon rollouts, preventing the numerical dissipation that affects standard spectral methods.

2 Related Work

Current research on neural PDE solvers primarily focuses on three aspects: architecture efficiency, physics constraints, and long-horizon stability. However, achieving all three simultaneously on generalized geometries remains an elusive goal.

2.1 Spectral Methods and Geometric Learning

Spectral methods achieve efficient global modeling. FNO [Li *et al.*, 2021] learns integral kernels with $O(N \log N)$ complexity, while DeepONet [Lu *et al.*, 2021] establishes the operator learning framework [Kovachki *et al.*, 2023]. Extensions such as U-FNO [Wen *et al.*, 2022], U-NO [Rahman *et al.*, 2023], CNO [Raonic *et al.*, 2023], and adaptive FNO [Guibas *et al.*, 2021] enhance multiresolution performance but rely on regular grids. Geo-FNO [Li *et al.*, 2023a] and domain-adaptive methods [Lingsch *et al.*, 2024] attempt mapping irregular geometries to regular grids but remain topologically constrained. Conversely, geometric deep learning [Bronstein *et al.*, 2021] handles generalized geometries using GNNs [Sanchez-Gonzalez *et al.*, 2020; Pfaff *et al.*, 2021; Fortunato *et al.*, 2022] or Transformers [Hao *et al.*, 2023; Hao *et al.*, 2024; Wu *et al.*, 2024]. Specialized models like GAOT [Wen *et al.*, 2025] and AeroGTO [Liu *et al.*, 2025] target 3D aerodynamics. While flexible, these methods depend on expensive CFD/FEM simulations for supervision, motivating the search for data-efficient, physics-constrained alternatives.

2.2 Physics-Informed Learning

To minimize supervision, PINNs [Raissi *et al.*, 2019] embed PDE governing equations into loss functions via automatic differentiation, with variational formulations like Deep Ritz [E and Yu, 2018] as alternatives. Subsequent works address gradient pathology [Yu *et al.*, 2022] and propagation failure [Wu *et al.*, 2025]. PINO [Li *et al.*, 2024] combines data and residual losses. However, computing higher-order derivatives on neural outputs amplifies high-frequency noise and causes prohibitive memory costs on 3D geometries [Yu *et al.*, 2022]. Weak-form methods [Kharazmi *et al.*, 2021] and MCNP [Zhang *et al.*, 2025] offer limited mitigation. MCNP employs Monte Carlo integration but is primarily limited to low-dimensional regular grids due to interpolation constraints. Moreover, these methods often struggle to maintain stability over long time horizons.

2.3 Long-Term Prediction

Error accumulation during autoregressive rollout is a key bottleneck [Lippe *et al.*, 2023]. Existing techniques, including noise injection [Brandstetter *et al.*, 2022], temporal bundling [Brandstetter *et al.*, 2022], and solver-based refinement [Giorgini *et al.*, 2025; Lippe *et al.*, 2023], provide partial mitigation. Pure data-driven methods [Cachay *et al.*, 2023] often drift over long horizons due to the lack of physical constraints. While recent studies incorporate physics-based regularization [Wang and Perdikaris, 2023], most rely on differential residuals, which are sensitive to the spatial and temporal discretization of the solution, leading to error amplification during long-term rollouts.

2.4 Our Contribution

UNOP unifies these strengths by decoupling geometry from physics. We introduce GALA to project irregular geometries into a regular latent space, enabling efficient spectral processing via GSEO. To ensure stability, we propose LIPE, which replaces unstable differential residuals with integral consistency. By enforcing physical laws via stochastic sampling, LIPE eliminates gradient pathologies, stabilizes predictions, and prevents error accumulation in long-horizon rollouts.

3 Methodology

3.1 Problem Formulation

We consider physical systems governed by the advection-diffusion equation [Karniadakis *et al.*, 2021] on a domain $\Omega \subset \mathbb{R}^d$ ($d \in \{1, 2, 3\}$):

$$\frac{\partial u}{\partial t} + \mathbf{v} \cdot \nabla u = \nu \nabla^2 u + f \tag{1}$$

where $u(\mathbf{x}, t) : \Omega \times [0, T] \to \mathbb{R}^C$ is the state field comprising C physical quantities, $\mathbf{v} \in \mathbb{R}^d$ is the advection velocity, $\nu \geq$

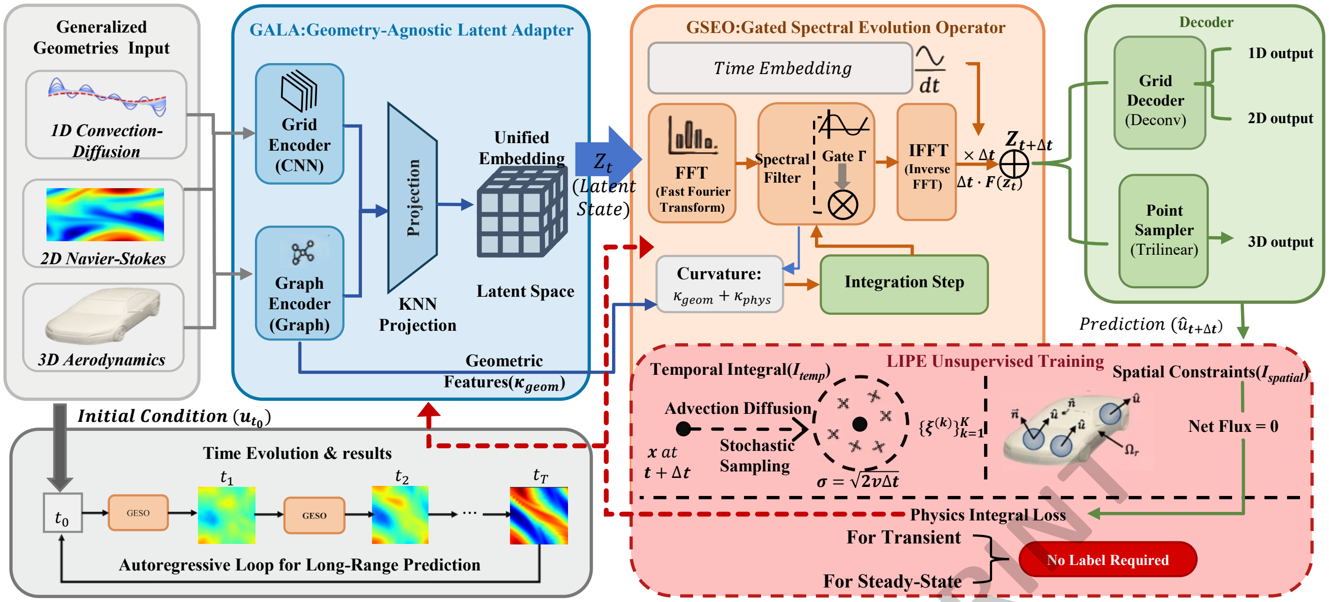

Figure 1: Architecture of UNOP. The framework consists of two main interconnected modules for its forward pass: (1) **GALA** (blue block, left) encodes generalized inputs into a unified latent space; (2) **GSEO** (orange block, middle) evolves the latent state via gated spectral operations. A **Decoder** (green block, right) projects the evolved state back to physical space. The entire framework is then trained unsupervisedly under the **LIPE** (red block, bottom right) paradigm, enforcing derivative-free physics constraints through an integral consistency objective. **This objective uses two integral operators: $\mathcal{I}_{\text{temp}}$ for transient problems (left part of red box) via stochastic advection-diffusion sampling, and $\mathcal{I}_{\text{spatial}}$ for steady-state problems (right part of red box) via control volume integration.**

0 is the diffusion coefficient (possibly spatially varying $\nu(\mathbf{x})$), and f is an external source term.

The domain is discretized as N samples $\{(\mathbf{x}_i, \mathbf{u}_i)\}_{i=1}^{N}$, where $\mathbf{x}_i \in \mathbb{R}^d$ is the spatial coordinate and $\mathbf{u}_i = u(\mathbf{x}_i, t) \in \mathbb{R}^C$ denotes the physical quantities at the i-th point. This discretization may reside on regular grids or irregular geometries, all of which are treated uniformly by our framework.

Our goal is to learn the solution operator $\mathcal{G} : u_t \mapsto u_{t+\Delta t}$ in an unsupervised manner without labeled simulation data, maintaining long-horizon stability.

3.2 Framework Overview

As illustrated in Figure 1, UNOP unifies three modules into a single pipeline. Crucially, LIPE (Sec. 3.4) constitutes the core physical principle, while GALA and GSEO provide the numerical realization.

To clarify the data flow and symbol usage, we distinguish three operational spaces:

- **Physical Space ($\mathbf{x}$):** The original domain containing irregular input points $\mathbf{x}_i$.
- **Latent Space (Z):** A unified representation where generalized inputs are mapped for efficient evolution.
- **Integral Space ($\mathcal{I}$):** The stochastic domain where LIPE constraints are evaluated via interpolation on latent grid.

The forward pass is defined as:

$$\hat{u}_{t+\Delta t} = \mathcal{D}\Big(Z_t + \Delta t \cdot \mathcal{S}(Z_t; \Gamma)\Big) \tag{2}$$

where $Z_t = \mathcal{E}(u_t)$ is the encoded latent state, $\mathcal{S}$ is the GSEO evolution operator, and $\mathcal{D}$ is the decoder. The pipeline is trained end-to-end via the LIPE objective to solve Eq. (1).

3.3 GALA: Geometry-Agnostic Latent Adapter

Standard spectral methods require regular grids, but real-world data often comes as irregular point clouds. GALA projects the discretized PDE data $\{(\mathbf{x}_i, \mathbf{u}_i)\}_{i=1}^{N}$ onto a unified regular latent grid, enabling efficient spectral processing and stable interpolation for LIPE.

Local Feature Extraction

To accommodate the discretized PDE data $\{(\mathbf{x}_i, \mathbf{u}_i)\}_{i=1}^{N}$ shown in Figure 1, we employ an encoding strategy. We extract local features $h_i \in \mathbb{R}^D$ based on the spatial arrangement of the points:

$$h_i = \begin{cases} \mathcal{E}_{\text{CNN}}(\mathbf{u}, \mathbf{x})_i & \text{for regular} \\ \mathcal{E}_{\text{Graph}}\big([\mathbf{u}_i; \mathbf{x}_i], \{[\mathbf{u}_j; \mathbf{x}_j]\}_{j \in \mathcal{N}_i}\big) & \text{for irregular} \end{cases} \tag{3}$$

For regular inputs, we use a lightweight CNN [Krizhevsky *et al.*, 2012] to extract local spatial features, while $\mathcal{E}_{\text{Graph}}$ is implemented as a physics-token based GATv2 [Brody *et al.*, 2022] to handle complex topologies. For the graph branch, $\mathbf{u}_i$ and $\mathbf{x}_i$ denote the physical quantities and coordinates of point i, aggregating information from neighbors $j \in \mathcal{N}_i$. This step ensures that local geometric features are extracted before unified projection.

Latent Grid Projection

To construct a structured representation, features $\{h_i\}$ are projected onto a regular latent grid $\mathcal{G}$ with N_g vertices

$\{\mathbf{g}_k\}_{k=1}^{N_g}$. Each vertex aggregates nearby input points in the neighborhood $\mathcal{N}_k$, weighted by a learnable kernel ϕ based on spatial distance:

$$Z(\mathbf{g}_k) = \frac{\sum_{i\in\mathcal{N}_k} \phi(\mathbf{g}_k - \mathbf{x}_i) \cdot h_i}{\sum_{i\in\mathcal{N}_k} \phi(\mathbf{g}_k - \mathbf{x}_i) + \epsilon} \tag{4}$$

Here, $Z(\mathbf{g}_k)$ is the feature vector at vertex $\mathbf{g}_k$, corresponding to a component of the latent state Z_t. The resulting tensor Z_t serves as the uniform substrate for GSEO evolution. The kernel ϕ is based on a Gaussian approximation with learnable parameters (see Appendix for implementation and weight derivation).

Importantly, the dimensionality of this latent grid does not correspond to the physical dimensionality of the input domain. Regardless of whether the original PDE is defined in 1D, 2D, or 3D, GALA embeds the geometry into a unified latent representation Z_t, enabling a shared evolution mechanism. Furthermore, the number of grid vertices N_g and the latent feature dimension D are hyperparameters independent of the input size, allowing the framework to scale to high-resolution problems without architectural changes.

3.4 LIPE: Latent Integral Physics Embedding

LIPE defines the unsupervised training objective. Instead of minimizing unstable differential residuals, we employ integral consistency derived from the physics of diffusion, enforced directly on the decoded prediction $\hat{u} = \mathcal{D}(Z)$ as shown in Eq. (2).

Integral Targets

Drawing inspiration from the Feynman-Kac formula [Karniadakis *et al.*, 2021] which links PDEs to stochastic processes, we reformulate Eq. 1 into integral forms.

Temporal Target: For transient dynamics, the state at $t + \Delta t$ is theoretically given by the expectation:

$$u(\mathbf{x}, t+\Delta t) = \mathbb{E}_{\boldsymbol{\xi}}\Big[u\big(\mathbf{x} - \mathbf{v}\Delta t + \boldsymbol{\xi}, t\big)\Big] + f(\mathbf{x})\Delta t \tag{5}$$

Here, $\mathbf{x} - \mathbf{v}\Delta t$ represents deterministic advection, $\boldsymbol{\xi} \sim \mathcal{N}(\mathbf{0}, 2\nu\Delta t\, \mathbf{I}_d)$ represents Brownian motion diffusion [Zhang *et al.*, 2025], and $f(\mathbf{x})\Delta t$ accounts for the integration of the external source term f in Eq. (1).

Spatial Target: For steady-state ($\partial_t u = 0$), mass conservation dictates zero net flux across any control volume Ω_r:

$$\oint_{\partial\Omega_r} \mathbf{u} \cdot \mathbf{n}\, dS = 0 \tag{6}$$

where $\mathbf{u}$ is the velocity vector, $\mathbf{n}$ is the outward unit normal vector on the boundary $\partial\Omega_r$, and dS denotes the surface area element.

Stochastic Discretization

To compute these targets during training, we define the **integral operator** $\mathcal{I}$ via stochastic sampling.

For Temporal Dynamics: We approximate the expectation $\mathbb{E}_{\boldsymbol{\xi}}$ in Eq. 5 using K stochastic samples:

$$\mathcal{I}_{\text{temp}}[u](\mathbf{x}) = \sum_{k=1}^{K} w_k \cdot u\big(\mathbf{x} - \mathbf{v}\Delta t + \boldsymbol{\xi}^{(k)}\big) \tag{7}$$

where $\boldsymbol{\xi}^{(k)}$ are sampled noise vectors and w_k are Gaussian weights normalized to sum to unity. Crucially, evaluating u at the off-grid locations $\mathbf{x} - \mathbf{v}\Delta t + \boldsymbol{\xi}^{(k)}$ is enabled by interpolation on the GALA latent grid.

For Spatial Constraints: We approximate the surface integral via discrete summation:

$$\mathcal{I}_{\text{spatial}}[\hat{u}_{t+\Delta t}](\Omega_r) = \sum_{j=1}^{M} \hat{u}_{t+\Delta t}(\mathbf{s}_j) \cdot \mathbf{n}_j \cdot \frac{4\pi r^2}{M} \tag{8}$$

where $\mathbf{s}_j$ are sampled points on the spherical surface S_r, $\mathbf{n}_j$ are the corresponding outward normals, r is the radius of S_r, and M is the number of surface samples. Here, $\hat{u}_{t+\Delta t} = \mathcal{D}(Z_t + \Delta t\, \mathcal{S}(Z_t))$ denotes the decoded prediction for the next time step. In practice, we select K and M to balance computational cost and variance reduction, with ablation studies in Sec. 4 confirming stability.

Training Objective

While the integral targets are derived from physical variables u, the LIPE objective is computed on the decoded network prediction $\hat{u} = \mathcal{D}(Z)$. This formulation ensures that the supervision signal from physical laws backpropagates through the decoder to guide the latent evolution in GALA and GSEO. The temporal loss is:

$$\mathcal{L}_{\text{trans}} = \big\|\hat{u}_{t+\Delta t} - (\mathcal{I}_{\text{temp}}[u_t] + f\Delta t)\big\|_2^2 \tag{9}$$

where the term $f\Delta t$ represents the explicit integration of the external source field f, consistent with Eq. 5.

For spatial problems, we minimize the flux residual:

$$\mathcal{L}_{\text{spatial}} = \mathbb{E}_{r,\mathbf{c}}\Big[\big(\mathcal{I}_{\text{spatial}}[\hat{u}_{t+\Delta t}](\Omega_r)\big)^2\Big] \tag{10}$$

This loss term explicitly penalizes local mass violations. Crucially, these losses provide the only supervision signal, guiding the entire framework to discover physical dynamics from scratch.

3.5 GSEO: Implementing Latent Evolution

GSEO provides a differentiable evolution prior in latent space. Without an explicit evolution operator, LIPE alone would lead to ill-posed optimization, as infinitely many mappings can satisfy integral constraints locally while violating global dynamics.

Spectral Processing

Global mixing is performed in the frequency domain:

$$\tilde{Z} = \mathcal{F}^{-1}\big(\mathbf{R} \odot \mathcal{F}(Z_t)\big) \tag{11}$$

where $\mathcal{F}$ is the Fourier transform and $\mathbf{R}$ are learnable weights. This effectively captures global advection features.

Gated Update with Curvature Awareness

Latent states corresponding to generalized geometries often contain high curvature regions, such as shocks or sharp gradients, which are common in realistic physical systems. To address this, we introduce a curvature-aware gate Γ that modulates the spectral evolution:

$$Z_{t+\Delta t} = Z_t + \Delta t \cdot \tau \cdot \Gamma \odot (\tilde{Z} - Z_t) \tag{12}$$

		$\beta = 0.01$			$\beta = 0.1$		
Model	**Roll.**	E1	E2	E3	E4	E5	E6
PINO	10	0.25	1.40	3.58	3.32	3.94	4.44
	20	0.08	0.48	1.02	1.57	1.11	1.44
PI-DON	10	3.69	5.68	3.42	8.60	5.17	9.63
	20	1.52	2.96	2.68	3.32	6.86	9.23
MCNP	10	0.19	0.23	0.26	0.24	0.19	0.47
	20	0.12	0.15	0.23	0.20	0.19	0.25
Ours	10	**0.07**	**0.12**	**0.08**	**0.09**	**0.05**	**0.11**
	20	0.09	0.24	**0.07**	0.33	**0.09**	**0.10**

Table 1: 1D Convection-Diffusion (%). Details in Sec. 4.

		Zero		Li		Kolmogorov	
Model	**Roll.**	E1	E2	E3	E4	E5	E6
PINO	10	1.22	6.98	7.24	9.54	6.24	8.53
	20	2.75	8.46	5.97	8.12	5.16	13.5
SP-FNO	10	4.78	6.25	5.42	18.7	27.6	18.5
	20	3.82	8.68	3.64	6.32	6.68	13.6
MCNP	10	7.55	4.56	2.56	5.48	4.68	9.54
	20	2.40	4.23	2.66	5.77	4.61	9.95
Ours	10	**1.05**	**1.26**	**2.06**	**3.25**	**4.13**	**5.33**
	20	**1.59**	**1.38**	**1.30**	**1.97**	**1.27**	**3.02**

Table 2: 2D Navier-Stokes (Rel. $\mathcal{L}_2$ %). Details in Sec. 4.

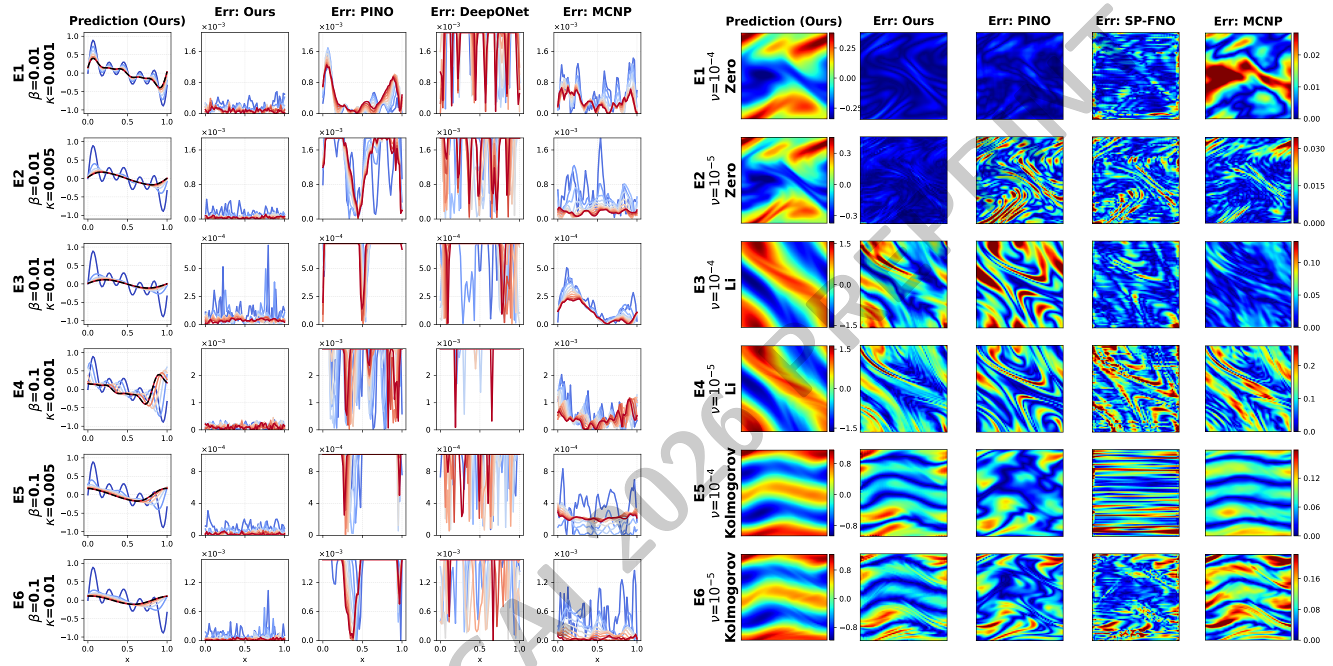

Figure 2: **Qualitative Results on Spatiotemporal Accuracy. Left:** Absolute error evolution for **1D Convection-Diffusion**. UNOP maintains negligible error levels. In contrast, PINO and PI-DeepONet exhibit significant deviations; their error curves are frequently truncated at the visualization limit, appearing as saturated spikes that far exceed the unified scale. **Right:** Error maps for **2D Navier-Stokes**. UNOP exhibits a uniformly deep blue distribution indicating low error. Conversely, **SP-FNO** shows distinct striping artifacts caused by spectral leakage, and **MCNP** displays red high-error regions in turbulent zones where it fails to resolve fine scales.

The gate Γ is defined as a function of both geometric and physical curvature indicators:

$$\Gamma = \sigma\big(\alpha \cdot \bar{\kappa}_{\text{geom}}(Z_t) + \beta \cdot \bar{\kappa}_{\text{phys}}(Z_t) + \gamma\big) \tag{13}$$

Here, $\bar{\kappa}_{\text{geom}}(Z_t)$ characterizes geometric curvature induced by the spatial arrangement of input points, while $\bar{\kappa}_{\text{phys}}(Z_t)$ is the Laplacian magnitude of the latent state, reflecting sharp physical variations. The learnable parameters α, β, and γ balance their contributions, enabling adaptive modulation of the update strength in high-curvature regions and mitigating the smoothing artifacts of standard spectral methods.

3.6 Decoding

The final prediction is obtained by decoding the evolved latent state $Z_{t+\Delta t}$:

$$\hat{u}_{t+\Delta t} = \mathcal{D}(Z_{t+\Delta t}) + \psi(h) \tag{14}$$

Here, the decoder $\mathcal{D}$ reconstructs physical values from the latent grid using interpolation, and may incorporate a residual network that leverages the pre-projection features h (Eq. 3) to recover high-frequency details lost during grid projection.

In summary, each component of UNOP addresses a specific limitation: GALA handles generalized geometries, LIPE enforces stable and consistent evolution over time, and GSEO for preserving fine-scale details in long-horizon predictions. The effectiveness of this design is confirmed in Sec. 4.

4 Experiments

Datasets and Tasks. We evaluate UNOP on three benchmarks, each specifically chosen to validate a core component of our framework: (1) **1D Convection-Diffusion**: This benchmark tests UNOP's long-horizon stability on abrupt

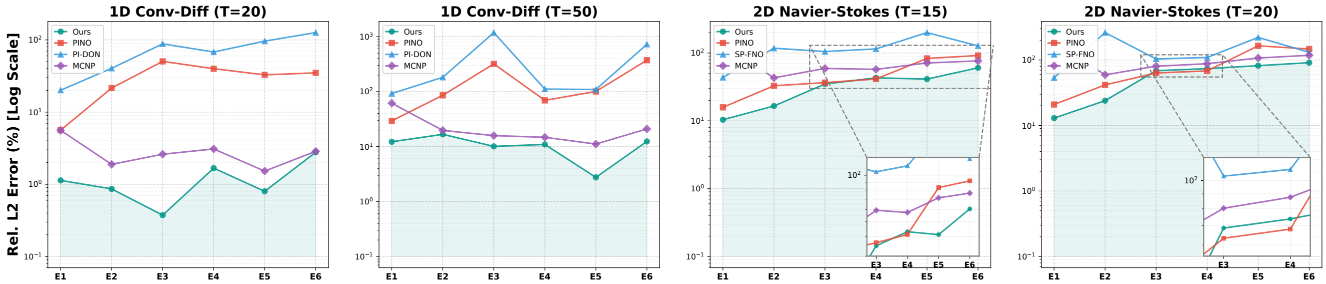

Figure 3: **Quantitative Long-term Stability (Log-Scale).** UNOP (Green) maintains a flat, low-error trajectory across both tasks. Note the **order-of-magnitude gap** versus baselines, confirming that integral constraints mitigate the rapid error growth typical of differential methods.

changes, highlighting LIPE's role in managing discontinuities. We define six regimes based on advection velocity β and diffusion coefficient κ: Regimes E1–E3 use low advection $\beta = 0.01$ with diffusion $\kappa \in \{0.001, 0.005, 0.01\}$ respectively. Regimes E4–E6 use high advection $\beta = 0.1$ with the same diffusion sequence.(2) **2D Navier-Stokes**: This chaotic system evaluates GSEO's ability to capture complex dynamics and maintain stable long-term predictions. Regimes E1 and E2 use Zero forcing, E3 and E4 use Li forcing, and E5 and E6 use Kolmogorov forcing. Within each forcing type, the odd-numbered regime uses high viscosity $\nu = 10^{-4}$, while the even-numbered regime uses low viscosity $\nu = 10^{-5}$.(3) **3D ShapeNet-Car**: A spatial domain task on unstructured point clouds, aimed at demonstrating GALA's adaptation to generalized geometries.

Comparison Methods. UNOP is compared with state-of-the-art unsupervised neural operators: **PI-DeepONet** [Wang and Perdikaris, 2023] serves as a physics-informed baseline for 1D functional mapping. **PINO** [Li *et al.*, 2024] combines data and residual losses within an FNO architecture. **SP-FNO** [Tran *et al.*, 2023] utilizes a factorized FNO variant for efficient 2D modeling. **MCNP** [Zhang *et al.*, 2025] acts as a integral-based competitor using Monte Carlo integration.

Implementation Details. All models are implemented in PyTorch and trained on single NVIDIA 4090 GPU. Detailed hyperparameters are provided in Appendix. All quantitative results are reported as the mean of three independent experiments.

Evaluation Metrics. We assess performance using **Relative $\mathcal{L}_2$ Error**. For 3D Car, we report the **Drag Coefficient Error** C_d and physics violation metrics $\mathcal{E}_{BC}$ and $\mathcal{E}_{Mass}$.

4.1 Main Results

Results on 1D/2D Benchmarks. Tables 1 and 2 present the quantitative comparison. In Table 1, regarding the **1D Convection-Diffusion** benchmark, UNOP achieves the lowest error across all regimes. Notably, in the high-advection regime E4 ($\beta = 0.1, \kappa = 10^{-3}$), UNOP reduces the error to **0.09%**, outperforming PINO at 3.32% and PI-DeepONet at 8.60% by orders of magnitude. This improvement mainly comes from LIPE's integral formulation, which successfully eliminates the high-frequency oscillations inherent in differential baselines. Table 2 details the results for the **2D Navier-Stokes** benchmark, where UNOP consistently outperforms all other baselines. For instance, in the Kolmogorov forcing task E6, UNOP achieves **3.02%** error at rollout 20, significantly lower than SP-FNO at 13.6% and MCNP at 9.95%. This confirms that GSEO's curvature-adaptive gating effectively mitigates spectral leakage, preserving fine-scale turbulent structures that pure spectral methods tend to smooth out.

Spatiotemporal Fidelity. Figure 2 provides a qualitative comparison. In 1D tasks, UNOP achieves near-zero deviation from the ground truth shock, with mean squared error around 10^{-4}, while baseline methods exhibit noticeable spurious oscillations. In 2D tasks, UNOP maintains low error across the entire domain, accurately capturing smooth low-frequency patterns while preserving high-frequency structures such as sharp gradients and vortices. In contrast, SP-FNO produces striping artifacts, and MCNP fails in high-gradient regions.

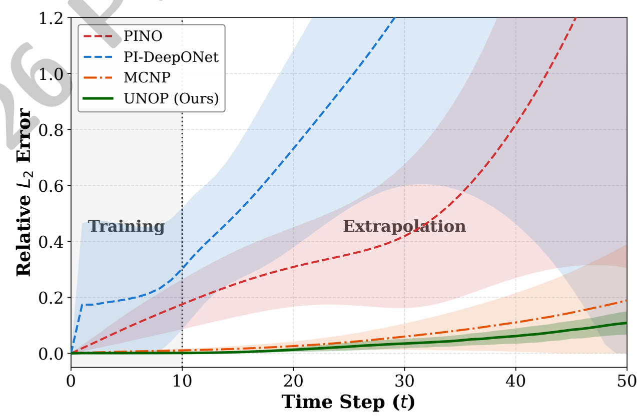

Figure 4: **1D Error Evolution.** The vertical line at $t = 10$ marks the training horizon; baseline errors rise beyond this point, while UNOP remains stable throughout.

Long-term Stability Analysis. Figure 3 and Figure 4 analyze error evolution during extended rollouts. As shown in Figure 4, during 1D extrapolation at $T = 50$, baseline errors surge: PI-DeepONet diverges beyond 120% relative error, while MCNP drifts up to 20%. In contrast, UNOP maintains a stable error trajectory below 15% throughout the rollout. LIPE stabilizes long-term rollouts by enforcing integral consistency, effectively suppressing autoregressive error accumulation. In 2D turbulence, UNOP preserves sharp vortex structures at $T = 20$ in Figure 5, whereas MCNP blurs fine details. This performance arises from a careful balance of global and local dynamics: low-frequency convective patterns are retained to stabilize the rollout, while GSEO's curvature adaptive gating selectively injects high frequency residuals into regions with sharp gradients. By doing so, the method recovers fine scale details lost to numerical dissipation in purely spectral or interpolation based approaches.

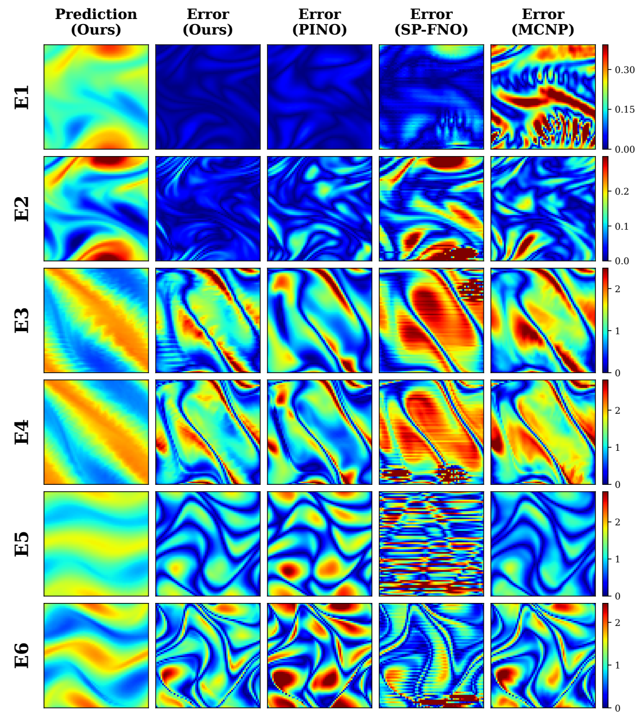

Figure 5: **2D Rollout Stability.** In these 2D Navier–Stokes heatmaps, redder regions correspond to higher prediction errors, highlighting areas where each model struggles.

Method	$\mathcal{L}_2$ (%) ↓	C_d (%) ↓	$\mathcal{E}_{BC}$ ↓	$\mathcal{E}_{Mass}$ ↓
PINO (10%)	47.68	32.25	5.744	0.157
PI-DeepONet (10%)	49.28	35.36	5.711	0.153
MCNP (10%)	17.36	5.83	3.428	0.086
UNOP (0%)	23.91	8.02	3.818	0.091
UNOP (10%)	**18.97**	**4.06**	**2.942**	**0.074**
UNOP (100%)	18.21	8.05	3.126	0.077

Table 3: 3D Aerodynamics on ShapeNet-Car. Comparison under Semi-Supervised (10% Label) settings. $\mathcal{E}_{BC/Mass}$: Physics violations ($\times 10^{-3}$). $\mathcal{L}_2$: Relative error (%).

Results on 3D Generalization. Table 3 reports the performance on the semi-supervised ShapeNet-Car task using 10% labels. UNOP achieves an $\mathcal{L}_2$ error of **18.97%** and a C_d error of **4.06%**, significantly outperforming the best baseline MCNP by **30%**. The methods such as PINO and PI-DeepONet exhibit substantial performance degradation with errors exceeding 47%. This highlights the inherent limitation of standard operator architectures when applied to irregular geometries without dedicated geometric adaptation, whereas UNOP's success validates GALA's ability to decouple geometric complexity from physical learning. UNOP with 10% supervision outperforms the fully supervised UNOP shown in gray, suggesting that sparse supervision combined with integral physics regularizes overfitting.

Configuration	2D (E4)	3D (C_d)
Full UNOP	**3.25**	**4.06**
w/o LIPE Constraint	33.13	-
w/o GALA Adapter	-	102.03
w/o GSEO (All)	44.16	25.37
– *w/o Curvature Gate*	23.09	18.79
– *w/o Spectral Filter*	37.25	20.61

Table 4: Ablation study on Component Effectiveness (Rel. $\mathcal{L}_2$ %). '-' denotes inapplicable configurations.

4.2 Ablation Studies

Table 4 validates the contribution of each component. **Effect of LIPE.** Replacing LIPE with standard differential loss increases the 2D error from 3.25% to 33.13%. This dramatic 10× error increase highlights LIPE's critical role in UNOP's stability. **Effect of GALA.** Removing GALA in 3D tasks leads to training failure with errors exceeding 100%, demonstrating its necessity for handling irregular geometries. **Effect of GSEO.** Disabling the Curvature Gate increases the error to 23.09% on 2D tasks, indicating its importance in preserving sharp dynamics that spectral filters tend to smooth.

4.3 3D Visualization Analysis

3D Boundary Adherence. Figure 6 visualizes surface pressure and velocity. UNOP maintains strict boundary adherence, whereas baselines show significant violations near boundaries, further confirming the efficacy of our integral boundary constraints.

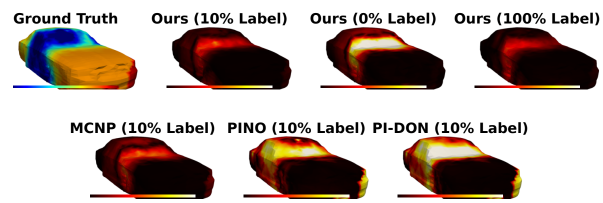

Figure 6: **3D Surface Pressure and Velocity Consistency.** UNOP maintains strict boundary adherence, whereas baselines show significant violations (bright/white regions).

5 Conclusion

We propose UNOP, a framework for unsupervised PDE learning that enforces integral consistency via LIPE rather than differential residuals. By reformulating physical constraints in latent space, UNOP achieves a stable balance between geometric flexibility and temporal evolution, demonstrating superior accuracy and long-term stability across 1D–3D benchmarks while capturing both global patterns and local structures such as shocks and vortices. Via the GALA interface, it handles complex geometries in high dimensions, consistently outperforming interpolation and spectral baselines in preserving physical fidelity over extended rollouts.

Limitations and Future Work. The current LIPE formulation targets advection-diffusion PDEs via the Feynman-Kac framework. Extending it to hyperbolic systems with strong shocks or multi-physics couplings remains an important direction for future work.

Contribution Statement

Xinrui Cheng and Ruihan Hu contributed equally to this work. *Ruihan Hu is the corresponding author.

Ethical Statement

There are no ethical issues.

Acknowledgments

This work was supported in part by the National Key Research and Development Program under Grant 2023YFF0905500, in part by the National Key Research and Development Program of Heilongjiang under Grant 2022ZX01A11, in part by the Natural Science Foundation of Heilongjiang Province under Grant LH2023F001, in part by the ZTE Industry-University-Institute Cooperation Funds under Grant No. 2022H2ZTE02-01-P4, and in part by the Spring Goose Talent Team Program of Heilongjiang.

References

[Brandstetter *et al.*, 2022] Johannes Brandstetter, Daniel E. Worrall, and Max Welling. Message passing neural PDE solvers. In *International Conference on Learning Representations*, 2022.

[Brody *et al.*, 2022] Shaked Brody, Uri Alon, and Eran Yahav. How attentive are graph attention networks? In *International Conference on Learning Representations*, 2022.

[Bronstein *et al.*, 2021] Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021.

[Cachay *et al.*, 2023] Salva Rühling Cachay, Bo Zhao, Hailey Joren, and Rose Yu. DYffusion: A dynamics-informed diffusion model for spatiotemporal forecasting. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

[Chen *et al.*, 2023] Xinyu Chen, Qiang Zhang, and Guanghui Lu. Key technologies of digital twin in space-air-ground integrated networks. *ZTE Communications*, 29(3):51–58, 2023.

[E and Yu, 2018] Weinan E and Bing Yu. The deep ritz method: A deep learning-based numerical algorithm for solving variational problems. *Communications in Mathematics and Statistics*, 6(1):1–12, 2018.

[Fortunato *et al.*, 2022] Meire Fortunato, Tobias Pfaff, Peter Wirnsberger, Alexander Pritzel, and Peter Battaglia. MultiScale MeshGraphNets. In *ICML 2022 2nd AI for Science Workshop*, 2022.

[Gao *et al.*, 2019] Zhipeng Gao, Congcong Yao, and Kaile Xiao. Mobile edge computing: Architecture, applications and challenges. *ZTE Communications*, 25(3):23–30, 2019.

[Giorgini *et al.*, 2025] Ludovico T. Giorgini, Andre N. Souza, Domenico Lippolis, Predrag Cvitanović, and Peter Schmid. Learning dissipation and instability fields from chaotic dynamics. *Physica D: Nonlinear Phenomena*, 481:134865, 2025.

[Guibas *et al.*, 2021] John Guibas, Morteza Mardani, Zongyi Li, Andrew Tao, Anima Anandkumar, and Bryan Catanzaro. Efficient token mixing for transformers via adaptive fourier neural operators. In *International Conference on Learning Representations*, 2021.

[Hao *et al.*, 2023] Zhongkai Hao, Zhengyi Wang, Hang Su, Chengyang Ying, Yinpeng Dong, Songming Liu, Ze Cheng, Jian Song, and Jun Zhu. GNOT: A general neural operator transformer for operator learning. In *Proceedings of the 40th International Conference on Machine Learning*, pages 12556–12569, 2023.

[Hao *et al.*, 2024] Zhongkai Hao, Chang Su, Songming Liu, Julius Berner, Chengyang Ying, Hang Su, Anima Anandkumar, Jian Song, and Jun Zhu. DPOT: Auto-regressive denoising operator transformer for large-scale PDE pretraining. In *Proceedings of the 41st International Conference on Machine Learning*, 2024.

[Herde *et al.*, 2024] Maximilian Herde, Bogdan Raonic, Tobias Rohner, Roger Käppeli, Roberto Molinaro, Emmanuel de Bézenac, and Siddhartha Mishra. Poseidon: Efficient foundation models for PDEs. In *Advances in Neural Information Processing Systems*, volume 37, 2024.

[Karniadakis *et al.*, 2021] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.

[Kharazmi *et al.*, 2021] Ehsan Kharazmi, Zhongqiang Zhang, and George E. M. Karniadakis. hp-VPINNs: Variational physics-informed neural networks with domain decomposition. *Computer Methods in Applied Mechanics and Engineering*, 374:113547, 2021.

[Kovachki *et al.*, 2023] Nikola Kovachki, Zongyi Li, Burigede Liu, Kamyar Azizzadenesheli, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Neural operator: Learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research*, 24(1):1–97, 2023.

[Krizhevsky *et al.*, 2012] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, volume 25, pages 1097–1105, 2012.

[Lam *et al.*, 2023] Remi Lam, Alvaro Sanchez-Gonzalez, Matthew Willson, Peter Wirnsberger, Meire Fortunato, et al. Learning skillful medium-range global weather forecasting. *Science*, 382(6677):1416–1421, 2023.

[Li *et al.*, 2021] Zongyi Li, Nikola Borislavov Kovachki, Kamyar Azizzadenesheli, Burigede Liu, Kaushik Bhattacharya, Andrew Stuart, and Anima Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021.

[Li *et al.*, 2023a] Zongyi Li, Daniel Zhengyu Huang, Burigede Liu, and Anima Anandkumar. Fourier neural operator with learned deformations for PDEs on general geometries. *Journal of Machine Learning Research*, 24(1):1–26, 2023.

[Li *et al.*, 2023b] Zongyi Li, Nikola Kovachki, Chris Choy, Boyi Li, Jean Kossaifi, Shourya Otta, Mohammad Amin Nabian, Maximilian Stadler, Christian Hundt, Kamyar Azizzadenesheli, and Animashree Anandkumar. Geometry-informed neural operator for large-scale 3d PDEs. In *Advances in Neural Information Processing Systems*, volume 36, pages 35836–35854, 2023.

[Li *et al.*, 2024] Zongyi Li, Hongkai Zheng, Nikola Kovachki, David Jin, Haoxuan Chen, Burigede Liu, Kamyar Azizzadenesheli, and Anima Anandkumar. Physics-informed neural operator for learning partial differential equations. *ACM/IMS Journal of Data Science*, 1(3):1–27, 2024.

[Lingsch *et al.*, 2024] Levi E. Lingsch, Mike Yan Michelis, Emmanuel de Bézenac, Sirani M. Perera, Robert K. Katzschmann, and Siddhartha Mishra. Beyond regular grids: Fourier-based neural operators on arbitrary domains. In *Proceedings of the 41st International Conference on Machine Learning*, pages 30610–30629, 2024.

[Lippe *et al.*, 2023] Phillip Lippe, Bastiaan S. Veeling, Paris Perdikaris, Richard E. Turner, and Johannes Brandstetter. PDE-refiner: Achieving accurate long rollouts with neural PDE solvers. In *Advances in Neural Information Processing Systems*, volume 36, 2023.

[Liu *et al.*, 2025] Pengwei Liu, Pengkai Wang, Xingyu Ren, Hangjie Yuan, Zhongkai Hao, Chao Xu, Shengze Cai, and Dong Ni. AeroGTO: An efficient graph-transformer operator for learning large-scale aerodynamics of 3d vehicle geometries. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pages 18924–18932, 2025.

[Lu *et al.*, 2021] Lu Lu, Pengzhan Jin, Guofei Pang, Zhongqiang Zhang, and George Em Karniadakis. Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence*, 3(3):218–229, 2021.

[Pfaff *et al.*, 2021] Tobias Pfaff, Meire Fortunato, Alvaro Sanchez-Gonzalez, and Peter Battaglia. Learning mesh-based simulation with graph networks. In *International Conference on Learning Representations*, 2021.

[Rahman *et al.*, 2023] Md Ashiqur Rahman, Zachary E Ross, and Kamyar Azizzadenesheli. U-NO: U-shaped neural operators. *Transactions on Machine Learning Research*, 2023.

[Raissi *et al.*, 2019] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[Raonic *et al.*, 2023] Bogdan Raonic, Roberto Molinaro, Tim De Ryck, Tobias Rohner, Francesca Bartolucci, Rima Alaifari, Siddhartha Mishra, and Emmanuel de Bézenac. Convolutional neural operators for robust and accurate learning of PDEs. In *Advances in Neural Information Processing Systems*, volume 36, pages 77187–77200, 2023.

[Sanchez-Gonzalez *et al.*, 2020] Alvaro Sanchez-Gonzalez, Jonathan Godwin, Tobias Pfaff, Rex Ying, Jure Leskovec, and Peter Battaglia. Learning to simulate complex physics with graph networks. In *Proceedings of the 37th International Conference on Machine Learning*, pages 8459–8468, 2020.

[Tran *et al.*, 2023] Alasdair Tran, Alexander Mathews, Lexing Xie, and Cheng Soon Ong. Factorized fourier neural operators. In *The Eleventh International Conference on Learning Representations*, 2023.

[Wang and Perdikaris, 2023] Sifan Wang and Paris Perdikaris. Long-time integration of parametric evolution equations with physics-informed DeepONets. *Journal of Computational Physics*, 475:111855, 2023.

[Wang *et al.*, 2023] Weili Wang, Lun Tang, and Qianbin Chen. Digital twin network based intelligent O&M for 6G networks. *ZTE Communications*, 29(3):8–14, 2023.

[Wen *et al.*, 2022] Gege Wen, Zongyi Li, Kamyar Azizzadenesheli, Anima Anandkumar, and Sally M. Benson. U-FNO: An enhanced fourier neural operator-based deep-learning model for multiphase flow. *Advances in Water Resources*, 163:104180, 2022.

[Wen *et al.*, 2025] Shizheng Wen, Arsh Kumbhat, Levi Lingsch, Sepehr Mousavi, Yizhou Zhao, Praveen Chandrashekar, and Siddhartha Mishra. Geometry aware operator transformer as an efficient and accurate neural surrogate for PDEs on arbitrary domains. *arXiv preprint arXiv:2505.18781*, 2025.

[Wu *et al.*, 2024] Haixu Wu, Huakun Luo, Haowen Wang, Jianmin Wang, and Mingsheng Long. Transolver: A fast transformer solver for PDEs on general geometries. In *Proceedings of the 41st International Conference on Machine Learning*, pages 53681–53705, 2024.

[Wu *et al.*, 2025] Haixu Wu, Yuezhou Ma, Hang Zhou, Huikun Weng, Jianmin Wang, and Mingsheng Long. ProPINN: Demystifying propagation failures in physics-informed neural networks. *arXiv preprint arXiv:2502.00803*, 2025.

[Yu *et al.*, 2022] Jeremy Yu, Lu Lu, Xuhui Meng, and George Em Karniadakis. Gradient-enhanced physics-informed neural networks for forward and inverse PDE problems. *Computer Methods in Applied Mechanics and Engineering*, 393:114823, 2022.

[Zeng *et al.*, 2022] Wei Zeng, Teng Su, Hui Wang, et al. Pangu: Large-scale autoregressive chinese pre-trained language model and its applications. *ZTE Communications*, (2):28, 2022.

[Zhang *et al.*, 2025] Rui Zhang, Qi Meng, Rongchan Zhu, Yue Wang, Wenlei Shi, Shihua Zhang, Zhi-Ming Ma, and Tie-Yan Liu. Monte carlo neural PDE solver for learning PDEs via probabilistic representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47(6):5059–5075, 2025.