

---

# Embedding conservation structure in neural fields for reduced state dynamics modeling from sparse and noisy measurements

---

**Aviral Prakash**  
Theoretical Division  
Los Alamos National Laboratory  
aviralp@lanl.gov

**Ben S. Southworth**  
Theoretical Division  
Los Alamos National Laboratory  
southworth@lanl.gov

**Marc L. Klasky**  
Theoretical Division  
Los Alamos National Laboratory  
mklasky@lanl.gov

## Abstract

Reduced/latent state dynamics models for parameterized PDEs offer a viable alternative to high-fidelity methods for multi-query applications. These reduced state dynamics approaches rely on high-quality data and struggle with noisy and sparse spatiotemporal measurements typically obtained from experiments and do not satisfy underlying conservation laws. In this article, we propose a reduced state dynamics approach, which we refer to as ECLEIRS, that satisfies conservation laws exactly even for parameters unseen in the model training process. We compare ECLEIRS with other reduced state dynamics approaches, those that do not enforce any physical constraints and those with physics-informed loss functions, for three shock-propagation problems: 1-D advection, 1-D Burgers and 2-D Euler equations. The numerical experiments conducted in this study demonstrate that ECLEIRS provides the most accurate prediction of dynamics for unseen parameters in the presence of highly sparse and noisy measurements.

## 1 Introduction

Despite the availability of numerous accurate high-fidelity methods, these models have a high computational expense making them expensive for multi-query forward and inverse problems such as parameter estimation, design space exploration or optimization and uncertainty quantification. Recently, there is a growing interest in non-intrusive methods [1, 2, 3, 4, 5] to model solution dynamics for these multi-query applications. Reduced/latent state dynamics methods offers a non-intrusive data-driven approach for constructing reduced order models of full spatiotemporal dynamics, thereby accelerating multi-query applications.

Early work on reduced state dynamics models for parameterized problems used autoencoder architecture to reduce dimensionality and parametrized neural ODEs (PNODEs) [6] or Sparse Identification of Nonlinear Dynamics (SINDy) [7, 8] for learning dynamics of low-dimensional reduced states. As these methods rely on autoencoder-decoder architectures, these methods are limited to fixed spatial grid data. To extend these approaches for arbitrary space-time measurements, implicit neural representations (INRs) [9, 10, 11, 12] were used in [13, 14, 15] to reduce dimensionality. As these methods rely solely on data, they struggle when the quality of data deteriorates due to low spatio-temporal resolution and noisy measurements. For such scenarios, adding physical constraints using

a penalty-type PDE-constrained optimization problem [16, 14] during training of a reduced states dynamics model can improve results. However, this approach does not guarantee the satisfaction of conservation laws at parameters or spatio-temporal instances not used in the model training process and can provide nonphysical predictions. This behavior highlights the need for approaches that satisfy conservation laws during the inference stage for arbitrary system parameters. There is recent work on enforcing conservation by projecting solutions through solution of an equality-constrained optimization problem [17, 18, 19]. However, these method must be evaluated on a fixed spatial grid and may not ensure local conservation at an arbitrary spatial location. To our knowledge, there is no prior work on exactly enforcing conservation laws for reduced state dynamics approaches applicable to arbitrary spatio-temporal measurements. Furthermore, there are limited studies that rigorously evaluate the performance of reduced state dynamics approaches for spatiotemporal sparse data with added noise.

In this work, we propose an INR-based reduced state dynamics approach, which we refer to as exact conservation law-embedded identification of reduced states (ECLEIRS), that applies to arbitrarily space-time distributed noisy data and ensures exact satisfaction of conservation laws for any spatial, temporal locations and system parameters, even those that are included in the training dataset. The performance of ECLEIRS is rigorously compared to other INR-based reduced state dynamics methods, those without conservation law information and a physics-informed variant, both of which can by definition also handle sparse data with noise. We demonstrate that ECLEIRS yields overall superior performance when trained on sparse and noisy data compared to the other methods while exactly satisfying local conservation laws.

## 2 Existing Methods

**System description:** We consider a solution vector  $\mathbf{q}(\mathbf{x}, t; \boldsymbol{\mu}) \in \mathbb{R}^{d_e}$  that evolves in time  $t \in \mathbb{R}^+$  within a spatial domain  $\mathbf{x} \in \mathcal{D}_x \subset \mathbb{R}^d$ , where  $d$  is the dimensionality of the spatial domain, subject to the parameters  $\boldsymbol{\mu} \in \mathcal{D}_{\boldsymbol{\mu}} \subset \mathbb{R}^{d_{\boldsymbol{\mu}}}$ , where  $d_{\boldsymbol{\mu}}$  is the dimensionality of system parameters. For many physical phenomena of interest, the evolution of the solution vector is governed by nonlinear conservation laws of the form

$$\frac{\partial \mathbf{q}(\mathbf{x}, t; \boldsymbol{\mu})}{\partial t} + \nabla_{\mathbf{x}} \cdot \mathbf{f}(\mathbf{q}(\mathbf{x}, t; \boldsymbol{\mu})) = 0, \quad (1)$$

where  $\mathbf{f}(\mathbf{q}(\cdot))$  is defined as the flux vector. The divergence operator considered in this article is of the form  $\nabla_{\mathbf{x}} = \frac{\partial}{\partial x_1} \hat{\mathbf{e}}_{x_1} + \frac{\partial}{\partial x_2} \hat{\mathbf{e}}_{x_2} + \frac{\partial}{\partial x_3} \hat{\mathbf{e}}_{x_3}$ , where  $\hat{\mathbf{e}}_{x_1}$ ,  $\hat{\mathbf{e}}_{x_2}$  and  $\hat{\mathbf{e}}_{x_3}$  are unit vectors along the Cartesian coordinate directions for  $d = 3$ .

**Identified reduced states (IRS):** Using autodecoder-based INRs [10] for dimensionality reduction involves approximating the solution as

$$\mathbf{q}^m(\mathbf{x}, t; \boldsymbol{\mu}) = \mathbf{d}_{\boldsymbol{\theta}_d}(\tilde{\mathbf{q}}(\boldsymbol{\mu}, t), \mathbf{x}) \quad \text{and} \quad \tilde{\mathbf{q}}(\boldsymbol{\mu}, t) = \mathbf{h}_{\boldsymbol{\theta}_h}(\boldsymbol{\mu}, t), \quad (2)$$

where  $\mathbf{q}^m(\mathbf{x}, t; \boldsymbol{\mu}) \in \mathbb{R}^{d_e}$  is the modeled solution field,  $\tilde{\mathbf{q}} \in \mathcal{M}$  are reduced states,  $\mathbf{d}_{\boldsymbol{\theta}_d}$  is an multilayer perceptron (MLP) parametrized with weights  $\boldsymbol{\theta}_d$  and  $\mathbf{h}_{\boldsymbol{\theta}_h}$  is an MLP hypernetwork parameterized with weights  $\boldsymbol{\theta}_h$ . Given the data for high-fidelity solution  $\mathbf{q}(\mathbf{x}, t; \boldsymbol{\mu})$ , the unknown parameters  $\boldsymbol{\theta}_d$  and  $\boldsymbol{\theta}_h$  are identified by solving a nonconvex optimization problem

$$\boldsymbol{\theta}_d, \boldsymbol{\theta}_h = \arg \min_{\hat{\boldsymbol{\theta}}_d, \hat{\boldsymbol{\theta}}_h} \left\| \mathbf{d}_{\hat{\boldsymbol{\theta}}_d}(\mathbf{h}_{\hat{\boldsymbol{\theta}}_h}(\boldsymbol{\mu}, t), \mathbf{x}) - \mathbf{q}(\mathbf{x}, t; \boldsymbol{\mu}) \right\|_2^2. \quad (3)$$

As the hypernetwork provides maps time directly to reduced states, it may not yield accurate reduced states for long time periods, which is needed for dynamics forecasting problems. Therefore, we instead use parametrized neural ODEs to determine dynamical equations for reduced states. These neural ODEs are obtained as

$$\frac{d\tilde{\mathbf{q}}(t, \boldsymbol{\mu})}{dt} = \mathbf{g}_{\boldsymbol{\theta}_g}(\tilde{\mathbf{q}}(t), \boldsymbol{\mu}) \quad \text{and} \quad \boldsymbol{\theta}_g = \arg \min_{\hat{\boldsymbol{\theta}}_g} \left\| \frac{d\tilde{\mathbf{q}}(t, \boldsymbol{\mu})}{dt} - \mathbf{g}_{\hat{\boldsymbol{\theta}}_g}(\tilde{\mathbf{q}}(t), \boldsymbol{\mu}) \right\|_2^2, \quad (4)$$

where  $\mathbf{g}_{\boldsymbol{\theta}_g}$  is an MLP with  $\boldsymbol{\theta}_g$  weights. This approach is similar to the methods in [13, 14, 15] with some differences in the neural network architecture.

**Physics-informed INR-based dimensionality reduction (PI-IRS):** Addition of physics conservation information during dimensionality reduction serves as a good regularization in the presence of low-quality measurements. Therefore, this “physics-informed” approach involves the same solution representation used in Eq. (2) while augmenting the optimization problem in Eq. (3) with a conservation law-based constraint. Following this strategy, the nonconvex optimization problem is written as

$$\boldsymbol{\theta}_d, \boldsymbol{\theta}_h = \arg \min_{\boldsymbol{\theta}_d, \boldsymbol{\theta}_h} \underbrace{\left\| \mathbf{d}_{\boldsymbol{\theta}_d}(\mathbf{h}_{\boldsymbol{\theta}_h}(\boldsymbol{\mu}, t), \mathbf{x}) - \mathbf{q}(\mathbf{x}, t; \boldsymbol{\mu}) \right\|_2^2}_{\text{Data loss}} + \lambda \underbrace{\left\| \frac{\partial \mathbf{d}_{\boldsymbol{\theta}_d}(\mathbf{h}_{\boldsymbol{\theta}_h}(\boldsymbol{\mu}, t), \mathbf{x})}{\partial t} + \nabla \cdot \mathbf{f}(\mathbf{d}_{\boldsymbol{\theta}_d}(\mathbf{h}_{\boldsymbol{\theta}_h}(\boldsymbol{\mu}, t), \mathbf{x})) \right\|_2^2}_{\text{Conservation law loss}}, \quad (5)$$

where  $\lambda$  is a user-input penalty parameter that governs the strictness of enforcement of the conservation law constraint. Once the low-dimensional reduced states  $\tilde{\mathbf{q}}$  are identified, the dynamical evolution of reduced states are obtained following the procedure as Eq. (4).

### 3 Proposed method: ECLEIRS

We consider vector fields of dimensionality  $d_e = 1$  to simplify the notation for method formulation. As shown in [20], conservation laws of the form in Eq. (1) can be written as

$$\nabla_w \cdot \mathbf{z}(\mathbf{w}; \boldsymbol{\mu}) = 0, \quad (6)$$

where  $\mathbf{z}(\mathbf{w}; \boldsymbol{\mu}) = [q(\mathbf{w}; \boldsymbol{\mu}) \ \mathbf{f}(q(\mathbf{w}; \boldsymbol{\mu}))]^T \in \mathbb{R}^{d+1}$  is the lifted vector field and  $\nabla_w = \frac{\partial}{\partial t} \hat{t} + \frac{\partial}{\partial x_1} \hat{e}_{x_1} + \frac{\partial}{\partial x_2} \hat{e}_{x_2} + \frac{\partial}{\partial x_3} \hat{e}_{x_3}$  is used to redefine the divergence operator with respect to space-time coordinate vector  $\mathbf{w} = [t \ \mathbf{x}] \in \mathbb{R}^{d+1}$ . For high-dimensional vector fields  $d+1 > 3$ , the form of  $\mathbf{z}$  to satisfy divergence free behavior was studied in [21, 22]. In particular, [22] identified that if there exists a skew-symmetric matrix field  $\mathbf{A}(\mathbf{w}) \in \mathbb{R}^{d+1 \times d+1}$  such that  $\mathbf{z}$  defined as row-wise divergence of  $\mathbf{A}(\mathbf{w})$  exactly satisfies Eq. (6). The above statement implies that  $\mathbf{z}(\mathbf{w}; \boldsymbol{\mu})$  has the form

$$\mathbf{z}^m(\mathbf{w}; \boldsymbol{\mu}) = \begin{bmatrix} q^m(\mathbf{w}; \boldsymbol{\mu}) \\ \mathbf{f}^m(\mathbf{w}; \boldsymbol{\mu}) \end{bmatrix} = \begin{bmatrix} \nabla_w \cdot \bar{\mathbf{A}}_1(\mathbf{w}; \boldsymbol{\mu}) \\ \vdots \\ \nabla_w \cdot \bar{\mathbf{A}}_{d+1}(\mathbf{w}; \boldsymbol{\mu}) \end{bmatrix}, \quad (7)$$

where  $\mathbf{z}^m(\mathbf{w}; \boldsymbol{\mu})$  is the modeled approximation of  $\mathbf{z}(\mathbf{w}; \boldsymbol{\mu})$  and  $\bar{\mathbf{A}}_i(\mathbf{w}; \boldsymbol{\mu})$  is the  $i$ th row of the skew-symmetric matrix field  $\mathbf{A}(\mathbf{w}; \boldsymbol{\mu}) \in \mathbb{R}^{(d+1) \times (d+1)}$ . The  $\frac{d(d+1)}{2}$  components of  $\mathbf{A}(\mathbf{w}; \boldsymbol{\mu})$ , represented as  $\mathbf{a}(\mathbf{w}; \boldsymbol{\mu}) \in \mathbb{R}^{\frac{d(d+1)}{2}}$ , are modeled using implicit neural representation of the form

$$\mathbf{a}(\mathbf{w}; \boldsymbol{\mu}) = \mathbf{d}_{\boldsymbol{\theta}_d}(\tilde{\mathbf{q}}(\boldsymbol{\mu}, t), \mathbf{x}), \quad (8)$$

where  $\mathbf{d}_{\boldsymbol{\theta}_d}$  is the MLP defined in Eq. (2). The detailed derivation of the condition in Eq. (7) can be found in [21, 22, 20]. For a 1-D spatial domain system

$$\mathbf{A}(\mathbf{w}; \boldsymbol{\mu}) = \begin{bmatrix} 0 & a(\mathbf{w}; \boldsymbol{\mu}) \\ -a(\mathbf{w}; \boldsymbol{\mu}) & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{z}^m(\mathbf{w}; \boldsymbol{\mu}) = \begin{bmatrix} q^m \\ f^m \end{bmatrix} = \begin{bmatrix} \frac{\partial a}{\partial x_1} \\ -\frac{\partial a}{\partial t} \end{bmatrix}. \quad (9)$$

Similarly, for a 2-D spatial domain system,

$$\mathbf{A}(\mathbf{w}; \boldsymbol{\mu}) = \begin{bmatrix} 0 & a_1(\mathbf{w}; \boldsymbol{\mu}) & a_2(\mathbf{w}; \boldsymbol{\mu}) \\ -a_1(\mathbf{w}; \boldsymbol{\mu}) & 0 & a_3(\mathbf{w}; \boldsymbol{\mu}) \\ -a_2(\mathbf{w}; \boldsymbol{\mu}) & -a_3(\mathbf{w}; \boldsymbol{\mu}) & 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{z}^m(\mathbf{w}; \boldsymbol{\mu}) = \begin{bmatrix} q^m \\ f_1^m \\ f_2^m \end{bmatrix} = \begin{bmatrix} \frac{\partial a_1}{\partial x_1} + \frac{\partial a_2}{\partial x_2} \\ -\frac{\partial a_1}{\partial t} + \frac{\partial a_3}{\partial x_2} \\ -\frac{\partial a_2}{\partial t} - \frac{\partial a_3}{\partial x_1} \end{bmatrix}. \quad (10)$$

Following the same definition of  $\tilde{\mathbf{q}}(t, \boldsymbol{\mu}) = \mathbf{h}_{\boldsymbol{\theta}_h}(\boldsymbol{\mu}, t)$  as in Eq. (2), the problem reduces to determining the unknown parameters  $\boldsymbol{\theta}_d$  and  $\boldsymbol{\theta}_h$  by solving a nonconvex optimization problem

$$\boldsymbol{\theta}_d, \boldsymbol{\theta}_h = \arg \min_{\boldsymbol{\theta}_d, \boldsymbol{\theta}_h} \left\| \mathbf{z}^m(\mathbf{w}; \boldsymbol{\mu}) - \mathbf{z}(\mathbf{x}, t; \boldsymbol{\mu}) \right\|_2^2. \quad (11)$$

Once the low-dimensional reduced states  $\tilde{\mathbf{q}}$  are identified, the dynamical evolution of reduced states are obtained following the formulation in Eq. (4).

Table 1: Details of the learning and validation datasets for 1-D advection problem. The learning dataset is further randomly divided between training (75%) and testing set (25%) to ensure that the model is not overfitted.

Test case	Dataset Name	Parameter values
1-D advection	Learning	$c \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$ , $q_0 \in \{0.9, 0.95, 1.0, 1.05, 1.1\}$ $x_{in} \in \{0.1, 0.15, 0.2, 0.25, 0.3\}$
	Validation	$c \in \{0.75, 0.85, 0.95, 1.05, 1.15, 1.25\}$ , $q_0 \in \{0.88, 0.93, 0.98, 1.03, 1.08, 1.13\}$ $x_{in} \in \{0.07, 0.12, 0.17, 0.22, 0.27, 0.32\}$
1-D Burgers	Learning	$c \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$ , $q_0 \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$ $\omega \in \{0.8, 0.9, 1.0, 1.1, 1.2\}$
	Validation	$c \in \{0.75, 0.85, 0.95, 1.05, 1.15, 1.25\}$ , $q_0 \in \{0.75, 0.85, 0.95, 1.05, 1.15, 1.25\}$ $\omega \in \{0.75, 0.85, 0.95, 1.05, 1.15, 1.25\}$
2-D Euler	Learning	70 random samples, $x_1^0 \sim \mathcal{U}(0, 1)$ , $x_2^0 \sim \mathcal{U}(0, 1)$
	Validation	30 random samples, $x_1^0 \sim \mathcal{U}(0, 1)$ , $x_2^0 \sim \mathcal{U}(0, 1)$ (different from the learning dataset)

Table 2: Hyperparameters in the neural network architectures for different problems. MLP with the SIREN [11] activation function is used to represent the three sub-networks.

Test case	Hypernetwork ( $h_{\theta_h}$ )	Decoder ( $d_{\theta_d}$ )	Dynamics ( $g_{\theta_g}$ )
1-D advection	4 layers, 64 neurons	4 layers, 40 neurons	4 layers, 32 neurons
1-D Burgers	4 layers, 64 neurons	4 layers, 40 neurons	4 layers, 32 neurons
2-D Euler	4 layers, 100 neurons	4 layers, 100 neurons	4 layers, 100 neurons

## 4 Results and Discussion

In this section, we assess and compare the performance of different reduced state dynamics approaches for: 1) the 1-D advection problem, 2) the 1-D Burgers problem and 3) the 2-D Euler problem. The details for the training and validation data for these test cases are shown in Table 1. The 1-D advection problem is parameterized by initial wave height ( $q_0$ ), initial wave width ( $x_{in}$ ) and wave speed. The 1-D Burgers problem is parameterized by the initial sine wave amplitude ( $q_0$ ), wave frequency ( $\omega$ ) and advection speed ( $c$ ). Lastly, the 2-D Euler problem is parameterized by the size of length ( $x_2^0$ ) and width ( $x_1^0$ ) of the initial density field centered at the origin. The details of the neural network architecture for the three sub-networks: decoder, hypernetwork and reduced dynamics ODE, are mentioned in Table 2. All the reduced state dynamics approaches are trained using ADAM optimizer [23], while the number of neural network parameters are kept similar for each method to enable a fair comparison. The validation dataset comprises of full spatio-temporal resolution solution fields at system parameters unseen in the training dataset. More details on the setup of numerical experiments and comprehensive analysis of results can be found in [24].

A comparison of density fields predicted by different reduced state dynamics models for 2-D Euler problem is shown in Figure 1. For this test case, all the models are trained on significantly sparse data as mentioned in the figure caption. The results indicate that IRS does not capture shock boundaries accurately for this test, whereas both PI-IRS and ECLEIRS perform much better at representing the density field. More quantitative results for all the test cases are tabulated in Table 3. The tabulated results indicate that PI-IRS and ECLEIRS perform better than IRS for all test cases, while ECLEIRS exhibits overall lowest errors. Despite PI-IRS also providing accurate results, these results are sensitive to user-selected  $\lambda$ , which appears to vary with different sparsity, noise levels and test cases. Therefore, obtaining good results using PI-IRS involves a time-consuming human-in-loop tuning step. On the other hand, ECLEIRS provide a tuning-free overall accurate predictions across different data scenarios. Lastly, we observe that both IRS and PI-IRS exhibit a high conservation error which is expected as these methods do not embed any conservation structure in their formulation. Conversely, ECLEIRS provide machine-precision conservation errors for spatio-temporal locations and parameters that are outside the training dataset for all three test cases.

In this work, all the networks were trained to the same number of epochs. As the three approaches use the same number of neural network parameters, the cost of model training is similar between the three approaches, with PI-IRS and ECLEIRS having a slightly higher training cost due to additional backpropagation required for enforcing the nonlinear conservation law. In the model inference stage, as the cost of dynamics evolution depends solely on the cost of evaluation of  $g_{\theta_g}$  and chosen time integration scheme, this cost is the same for the three reduced dynamics approaches. However, additional cost is incurred for ECLEIRS when  $q$  is obtained from the time-evolved latent variables

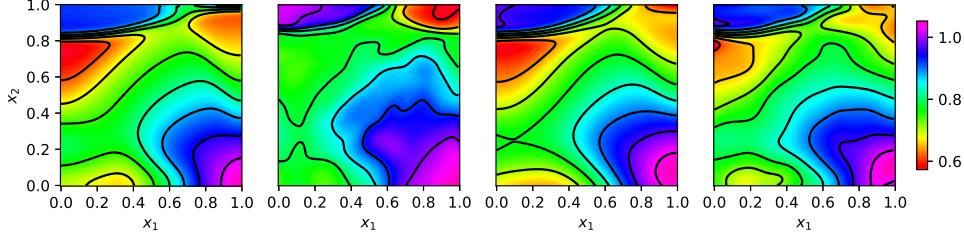


Figure 1: 2-D Euler problem: Predicted density at  $\mu = [0.527, 0.775]$  in the validation dataset for ground truth (left), IRS (center-left), PI-IRS with  $\lambda = 10^{-1}$  (center-right) and ECLEIRS (right). These reduced state dynamics models were trained using data from learning dataset with 1% spatial sparsity, 40% temporal sparsity and added noise with standard deviation of 0.1.

Table 3: Solution and conservation error for different test cases. The 1-D advection and 1-D Burgers test cases are trained using 20% spatial and temporal randomly sampled data, whereas 2-D Euler test cases is trained using 0.5% spatial and 20% temporal randomly sampled sparse data. Different noise levels represent different standard deviation of added Gaussian noise ( $\sigma_1 < \sigma_2 < \sigma_3 < \sigma_4$ ), where  $\sigma = [\sigma_1, \sigma_2, \sigma_3, \sigma_4] = [0, 0.03, 0.1, 0.2]$  for 1-D advection,  $\sigma = [0, 0.05, 0.1, 0.2]$  for 2-D Euler problems and  $\sigma = [0, 0.1, 0.2, 0.4]$  for 1-D Burgers problem. The validation dataset comprises of full resolution clean solution fields at parameter values unseen in the training dataset. Solution and conservation (cons.) errors are quantified over full resolution spatial and temporal locations. The mean and standard deviation (SD) of errors are with respect to parameters in the validation dataset.

Test case	Models	Solution error (mean $\pm$ SD)				Cons. error
		$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_4$	
1-D advection	IRS	0.31 $\pm$ 0.43	0.53 $\pm$ 0.47	0.39 $\pm$ 0.29	0.59 $\pm$ 0.23	$8 \times 10^{-2} \pm 7 \times 10^{-2}$
	PI-IRS, $\lambda = 10^{-3}$	0.19 $\pm$ 0.21	<b>0.09 <math>\pm</math> 0.05</b>	0.31 $\pm$ 0.36	0.29 $\pm$ 0.21	$3 \times 10^{-2} \pm 5 \times 10^{-2}$
	PI-IRS, $\lambda = 10^{-1}$	0.25 $\pm$ 0.43	0.35 $\pm$ 0.58	0.27 $\pm$ 0.49	<b>0.21 <math>\pm</math> 0.16</b>	$9 \times 10^{-3} \pm 10^{-2}$
	ECLEIRS	<b>0.15 <math>\pm</math> 0.09</b>	0.16 $\pm$ 0.19	<b>0.25 <math>\pm</math> 0.25</b>	0.30 $\pm$ 0.24	<b><math>8 \times 10^{-9} \pm 6 \times 10^{-9}</math></b>
1-D Burgers	IRS	0.24 $\pm$ 0.15	0.27 $\pm$ 0.15	0.29 $\pm$ 0.11	0.38 $\pm$ 0.12	$10^{-1} \pm 10^{-1}$
	PI-IRS, $\lambda = 10^{-3}$	<b>0.20 <math>\pm</math> 0.05</b>	0.2 $\pm$ 0.17	0.24 $\pm$ 0.10	<b>0.24 <math>\pm</math> 0.10</b>	$3 \times 10^{-2} \pm 5 \times 10^{-2}$
	PI-IRS, $\lambda = 10^{-1}$	0.29 $\pm$ 0.13	0.25 $\pm$ 0.08	0.27 $\pm$ 0.12	0.29 $\pm$ 0.14	$6 \times 10^{-2} \pm 8 \times 10^{-2}$
	ECLEIRS	0.25 $\pm$ 0.09	<b>0.15 <math>\pm</math> 0.12</b>	<b>0.18 <math>\pm</math> 0.08</b>	<b>0.24 <math>\pm</math> 0.12</b>	<b><math>6 \times 10^{-9} \pm 5 \times 10^{-9}</math></b>
2-D Euler	IRS	0.031 $\pm$ 0.010	0.048 $\pm$ 0.019	0.1 $\pm$ 0.037	0.127 $\pm$ 0.03	N/A
	PI-IRS, $\lambda = 10^{-3}$	0.028 $\pm$ 0.010	<b>0.022 <math>\pm</math> 0.007</b>	0.05 $\pm$ 0.034	0.057 $\pm$ 0.023	$5 \times 10^{-2} \pm 10^{-2}$
	PI-IRS, $\lambda = 10^{-1}$	0.023 $\pm$ 0.013	<b>0.022 <math>\pm</math> 0.007</b>	<b>0.028 <math>\pm</math> 0.023</b>	<b>0.039 <math>\pm</math> 0.019</b>	$9 \times 10^{-3} \pm 6 \times 10^{-3}$
	ECLEIRS	<b>0.013 <math>\pm</math> 0.003</b>	0.028 $\pm$ 0.008	<b>0.033 <math>\pm</math> 0.01</b>	0.052 $\pm$ 0.015	<b><math>2 \times 10^{-9} \pm 10^{-9}</math></b>

due to additional gradient computation. The difference in this additional cost compared to other models will be low if solution  $q$  is only desired at a few spatial and temporal instances.

These results indicate that ECLEIRS gives the overall best performance by consistently providing the most accurate results while not requiring expensive parameter tuning and satisfying conservation laws to machine precision even for parameters unseen during the model learning phase. The ability to ensure exact conservation and accurately predict dynamics makes ECLEIRS a reliable and robust approach for parameterized PDE problems. We plan to extend this formulation for quantify uncertainty especially for out-of-distribution parameters and time instance while also validating these models for more complex physical problems, especially those exhibiting a conservation structure.

## 5 Acknowledgements

We would like to thank Ryosuke Park and Marc Charest as the primary developers of the *flecsim* package we used for 2-D Euler calculations. This work was supported by Laboratory Directed Research and Development program of Los Alamos National Laboratory under project number 20230068DR. BSS was partially supported by the DOE Office of Advanced Scientific Computing Research Applied Mathematics program through Contract No. 89233218CNA000001. Los Alamos National Laboratory Report LA-UR-25-31089.

## References

- [1] Benjamin Peherstorfer and Karen E. Willcox. Data-driven operator inference for nonintrusive projection-based model reduction. *Computer Methods in Applied Mechanics and Engineering*, 306:196–215, 2016.
- [2] Boris Kramer, Benjamin Peherstorfer, and Karen E. Willcox. Learning nonlinear reduced models from data with operator inference. *Annual Review of Fluid Mechanics*, 56:521–548, 2024.
- [3] Vedant Puri, Aviral Prakash, Levent B. Kara, and Yongjie J. Zhang. SNF-ROM: Projection-based nonlinear reduced order modeling with smooth neural fields. *Journal of Computational Physics*, 532:113957, 2025.
- [4] Aviral Prakash and Yongjie J. Zhang. Data-driven identification of stable sparse differential operators using constrained regression. *Computer Methods in Applied Mechanics and Engineering*, 429:117149, 2024.
- [5] Aviral Prakash and Yongjie J. Zhang. Nonintrusive projection-based reduced order modeling using stable learned differential operators. *Computer Methods in Applied Mechanics and Engineering*, 442:117946, 2025.
- [6] Kookjin Lee and Eric J. Parish. Parameterized neural ordinary differential equations: Applications to computational physics problems. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 477, 2021.
- [7] William D. Fries, Xiaolong He, and Youngsoo Choi. LaSDI: Parametric latent space dynamics identification. *Computer Methods in Applied Mechanics and Engineering*, 399:115436, 2022.
- [8] Christophe Bonneville, Xiaolong He, April Tran, Jun S. Park, William Fries, Daniel A. Messenger, Siu W. Cheung, Yeonjong Shin, David M. Bortz, Debojyoti Ghosh, Jiun-Shyan Chen, Jonathan Belof, and Youngsoo Choi. A comprehensive review of latent space dynamics identification algorithms for intrusive and non-intrusive reduced-order-modeling. *arXiv:2403.10748*, 2024.
- [9] Zhiqin Chen and Hao Zhang. Learning implicit fields for generative shape modeling. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5932–5941, 2019.
- [10] Jeong J. Park, Peter Florence, Julian Straub, Richard Newcombe, and Steven Lovegrove. DeepSDF: Learning continuous signed distance functions for shape representation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019.
- [11] Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *Proceedings of the 34th International Conference on Neural Information Processing Systems, NIPS ’20*, 2020.
- [12] Shaowu Pan, Steven L. Brunton, and José N. Kutz. Neural implicit flow: A mesh-agnostic dimensionality reduction paradigm of spatio-temporal data. *Journal of Machine Learning Research*, 24(41):1–60, 2023.
- [13] Yuan Yin, Matthieu Kirchmeyer, Jean-Yves Franceschi, Alain Rakotomamonjy, and Patrick Gallinari. Continuous PDE dynamics forecasting with implicit neural representations. *arXiv:2209.14855*, 2023.
- [14] Tianshu Wen, Kookjin Lee, and Youngsoo Choi. Reduced-order modeling for parameterized PDEs via implicit neural representations. *arXiv:2311.16410*, 11 2023.
- [15] Francesco Regazzoni, Stefano Pagani, Matteo Salvador, Luca Dedé, and Alfio Quarteroni. Learning the intrinsic dynamics of spatio-temporal processes through latent dynamics networks. *Nature Communications*, 15, 12 2024.
- [16] Maziar Raissi, Paris Perdikaris, and George E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

- [17] Kookjin Lee and Kevin T. Carlberg. Deep conservation: A latent-dynamics model for exact satisfaction of physical conservation laws. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(1):277–285, May 2021.
- [18] Geoffrey Négiar, Michael W. Mahoney, and Aditi S. Krishnapriyan. Learning differentiable solvers for systems with hard constraints. *arXiv:2207.08675*, 2023.
- [19] Valentin Duruisseaux, Miguel Liu-Schiaffini, Julius Berner, and Anima Anandkumar. Towards enforcing hard physics constraints in operator learning frameworks. In *ICML 2024 AI for Science Workshop*, 2024.
- [20] Jack Richter-Powell, Yaron Lipman, and Ricky T. Q. Chen. Neural conservation laws: A divergence-free perspective. In *Advances in Neural Information Processing Systems*, 2022.
- [21] Cristian Barboiu. Representation of divergence-free vector fields. *Quarterly of applied mathematics*, 69(2):309–316, 2011.
- [22] James P. Kelliher. Stream functions for divergence-free vector fields. *Quarterly of applied mathematics*, 79(1):163–174, 2021.
- [23] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv:1412.6980*, 2017.
- [24] Aviral Prakash, Ben S. Southworth, and Marc L. Klasky. ECLEIRS: Exact conservation law embedded identification of reduced states for parameterized partial differential equations from sparse and noisy data. *arXiv:2506.18855*, 2025.