



Structure-preserving Lift & Learn: Scientific machine learning for nonlinear conservative partial differential equations

Harsh Sharma¹ · Juan Diego Draxl Giannoni² · Boris Kramer²

Received: 5 July 2025 / Accepted: 13 February 2026

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2026

Abstract

This work presents *structure-preserving Lift & Learn*, a scientific machine learning method that employs lifting variable transformations to learn structure-preserving reduced-order models for nonlinear partial differential equations (PDEs) with conservation laws. We propose a hybrid learning approach based on a recently developed energy-quadratization strategy that uses knowledge of the nonlinearity at the PDE level to derive an equivalent quadratic lifted system with quadratic system energy. The lifted dynamics obtained via energy quadratization are linear in the old variables, making model learning very effective in the lifted setting. Based on the lifted quadratic PDE model form, the proposed method derives quadratic reduced terms analytically and then uses those derived terms to formulate a constrained optimization problem to learn the remaining linear reduced operators in a structure-preserving way. The proposed hybrid learning approach yields computationally efficient quadratic reduced-order models that respect the underlying physics of the high-dimensional problem. We demonstrate the generalizability of quadratic models learned via the proposed structure-preserving Lift & Learn method through three numerical examples: the one-dimensional wave equation with exponential nonlinearity, the two-dimensional sine-Gordon equation, and the two-dimensional Klein-Gordon-Zakharov equations. The numerical results show that the proposed learning approach is competitive with the state-of-the-art structure-preserving data-driven model reduction method in terms of both accuracy and computational efficiency.

Keywords Model reduction · Lift & Learn · Scientific machine learning

Mathematics Subject Classification (2010) 37K06 · 37N30 · 65P10 · 70G45

Communicated by: Benjamin Peherstorfer

✉ Harsh Sharma
hsharma33@wisc.edu

¹ Department of Mechanical Engineering, University of Wisconsin-Madison, Madison, WI, USA

² Department of Mechanical and Aerospace Engineering, University of California San Diego, San Diego, CA, USA

1 Introduction

Model reduction methods seek to derive reduced-order models (ROMs) that provide accurate approximations of the full-order model (FOM) solutions. The low-dimensional nature of ROMs enables many-query tasks such as optimization, design, uncertainty quantification, and control, where models need to be evaluated many times for different settings. For applications where the high-dimensional problem possesses additional qualitative features like conservation laws, it is desirable to construct a ROM that respects the underlying physics of the problem. The field of structure-preserving model reduction has developed such physics-preserving ROMs for a wide class of applications ranging from computational physics [1–3] to structural mechanics [4, 5] to soft robotics [6–8].

In this work, we focus on structure-preserving nonintrusive model reduction for conservative PDEs of the form

$$\frac{\partial^2 \phi(\mathbf{x}, t)}{\partial t^2} = \Delta \phi(\mathbf{x}, t) - f_{\text{non}}(\phi(\mathbf{x}, t)), \quad (1)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_d) \in \Omega$ is the spatial variable, t is time, Δ is the Laplacian operator in \mathbb{R}^d , $\phi(\mathbf{x}, t)$ is the scalar state field, and the nonlinear component of the vector field, $f_{\text{non}}(\phi(\mathbf{x}, t)) := \nabla_{\phi}(g(\phi(\mathbf{x}, t)))$, is derived from a smooth nonlinear function $g(\phi(\mathbf{x}, t))$. A key property of nonlinear wave equations of the form (1) is that the total energy

$$\mathcal{E}[\phi(\mathbf{x}, t)] := \int_{\Omega} \left(\frac{1}{2} \left(\frac{\partial \phi(\mathbf{x}, t)}{\partial t} \right)^2 + \frac{1}{2} (\nabla \phi(\mathbf{x}, t))^2 + g(\phi(\mathbf{x}, t)) \right) \mathrm{d}\mathbf{x}, \quad (2)$$

is a conserved quantity provided the boundary conditions enforce zero normal energy flux through the boundary, i.e., $\partial_n \phi(\mathbf{x}, t) \frac{\partial \phi(\mathbf{x}, t)}{\partial t} = 0$ on $\partial\Omega$. This includes common conservative settings such as periodic boundary conditions, homogeneous Neumann boundary conditions, and unbounded domains where $\phi(\mathbf{x}, t)$ and its gradients decay sufficiently fast at infinity. The integrand in (2) typically has the physical interpretation of energy density with the first term in the integrand representing the kinetic energy and the other two terms representing the potential energy.

Structure-preserving model reduction for conservative PDEs was first explored from the Hamiltonian viewpoint in [9] where the authors derived Hamiltonian ROMs by projecting the Hamiltonian FOM operators onto a symplectic subspace obtained via proper symplectic decomposition (PSD). In a similar direction, the work in [10] developed reduced basis methods for structure-preserving model reduction of parametric Hamiltonian systems. The authors in [11] presented a modified projection technique based on the proper orthogonal decomposition (POD) basis such that the Hamiltonian structure is preserved after the Galerkin projection step. Building on these methods, a variety of structure-preserving model reduction methods have been developed recently [12–15]. All of these methods, however, are *intrusive*, which means that the projection step requires access to the high-dimensional FOM operators. Moreover, all

of the aforementioned structure-preserving approaches suffer from computational efficiency issues as the evaluation of the nonlinear components of the structure-preserving ROM vector field still scales with the FOM dimension. To tackle this computational bottleneck, the authors in [16] developed a structure-preserving variant of the discrete empirical interpolation method (DEIM) for Hamiltonian systems. The authors in [17] improved on this earlier work and presented a gradient-preserving hyper-reduction approach with theoretical guarantees for the preservation of the FOM Hamiltonian. This approach, however, requires a computationally demanding offline phase that constructs the DEIM basis from the Jacobian snapshot data matrix.

In another research direction, projection-based model reduction via lifting [18–22] has emerged as a promising approach for nonlinear dynamical systems. By quadratizing the nonlinear dynamics before projection, these methods eliminate the need for the additional hyper-reduction step. In general, however, quadratic ROMs derived via lifting are not guaranteed to respect the underlying physics of the high-dimensional problem. The authors in [23] presented an energy-quadratization strategy for structure-preserving model reduction for conservative PDEs of the form (1). Even though the energy-quadratization strategy yields a computationally efficient structure-preserving ROM without any additional hyper-reduction, the Galerkin projection step is intrusive in the sense that it requires access to the linear and quadratic FOM operators in the lifted setting. For many practical applications, the lifted FOM operators are not available as this would require implementing a discretization for a new PDE, which motivates the need for a *nonintrusive* approach for learning structure-preserving quadratic ROMs directly from data.

The Operator Inference method for nonintrusive model reduction of FOMs with linear and/or low-order polynomial terms was introduced in [24, 25]. This nonintrusive approach has been extended to nonlinear FOMs with nonpolynomial systems in [26] where knowledge of nonpolynomial nonlinear terms at the PDE level is leveraged to learn the remaining reduced operators via Operator Inference. Building on this gray-box approach, the authors in [27] developed Hamiltonian Operator Inference for learning physics-preserving ROMs of canonical Hamiltonian systems. Since then, a variety of structure-preserving Operator Inference methods have been developed for different classes of conservative FOMs [28–31]. However, all of these structure-preserving Operator Inference approaches, similarly to their intrusive counterparts, suffer from computational efficiency issues for nonlinear systems and need an additional structure-preserving hyper-reduction step to reduce the computational cost in the online stage.

The authors in [32] integrated lifting transformations with the Operator Inference framework to develop a method called Lift & Learn for nonintrusively learning quadratic ROMs of general nonlinear FOMs. This method circumvents the need for hyper-reduction by learning quadratic ROMs from projections of the FOM snapshot data in the lifted setting. Even though this machine learning approach learns computationally efficient quadratic ROMs from data, the learned ROMs are not guaranteed to respect the underlying physics of the problem (see Section 2). The main goal of this work is to develop a structure-preserving Lift & Learn approach to derive structure-

preserving quadratic ROMs of nonlinear conservative PDEs. The main contributions of this work are as follows:

1. We present a nonintrusive model reduction method—a special class of scientific machine learning—to derive structure-preserving quadratic ROMs from data generated by the high-dimensional nonlinear models of conservative PDEs. The proposed approach (see Fig. 1 for a schematic overview) leverages knowledge of the nonpolynomial nonlinearity at the conservative PDE level to ensure that the learned quadratic ROMs respect the underlying physics of the problem.
2. We present a theoretical result that shows that the structure-preserving Lift & Learn ROMs conserve a perturbed lifted FOM energy exactly.
3. We learn structure-preserving quadratic ROMs for three nonlinear conservative PDEs. The numerical results demonstrate the learned ROMs’ ability to provide accurate and stable predictions outside the training data regime while also achieving computational efficiency similar to the state-of-the-art nonintrusive Hamiltonian ROMs with structure-preserving hyper-reduction.

The paper is structured as follows. Section 2 summarizes the standard Lift & Learn approach for learning quadratic ROMs of general nonlinear systems via lifting transformations. Section 3 introduces structure-preserving Lift & Learn, a nonintrusive model reduction approach for learning quadratic ROMs that respect the underlying physics of the problem. Section 4 demonstrates the proposed nonintrusive approach on three nonlinear conservative PDEs with increasing complexity: one-dimensional nonlinear wave equation with exponential nonlinearity, two-dimensional sine-Gordon

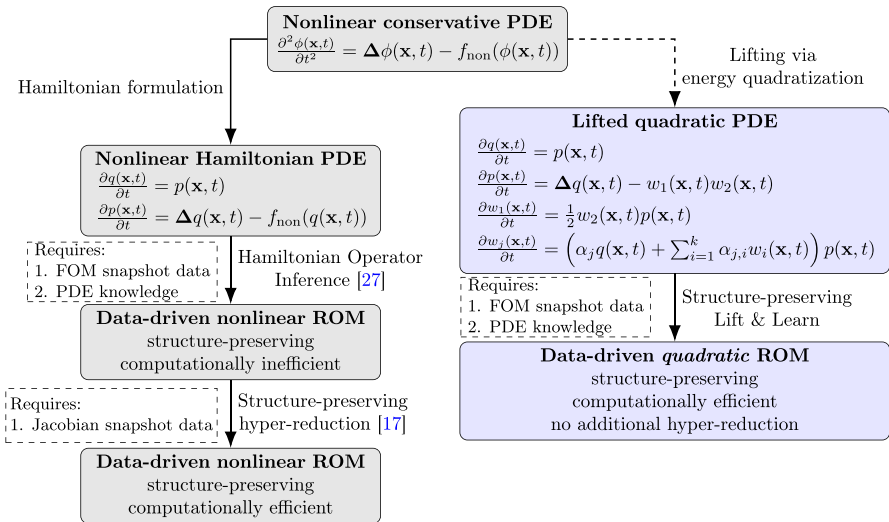


Fig. 1 The proposed structure-preserving Lift & Learn approach (right path) learns computationally efficient quadratic ROMs without any additional hyper-reduction step whereas the Hamiltonian Operator Inference approach (left path) requires additional Jacobian snapshot data for structure-preserving hyper-reduction. The dashed line style for the arrow from the nonlinear conservative PDE to the lifted quadratic FOM indicates that we only derive the symbolic form of the lifted quadratic FOM, and constructing the lifted FOM operators is *not* required

equation, and two-dimensional Klein-Gordon-Zakharov equation. Finally, Section 5 provides concluding remarks and suggests future research directions.

2 Background

In Section 2.1, we review Lift & Learn [32], a model reduction method for deriving quadratic ROMs of general nonlinear FOMs nonintrusively from data. In Section 2.2, we motivate the need for structure-preserving Lift & Learn by demonstrating how the standard Lift & Learn approach violates the conservative nature of the original high-dimensional problem.

2.1 Standard Lift & Learn

Consider a nonlinear PDE

$$\frac{\partial y(\mathbf{x}, t)}{\partial t} = f(y(\mathbf{x}, t)), \quad y(\mathbf{x}, 0) = y_0(\mathbf{x}), \tag{3}$$

where $f(y(\mathbf{x}, t))$ is a differentiable nonlinear function that maps the d_s -dimensional vector consisting of field variables $y(\mathbf{x}, t)$ to its time derivative and $y_0(\mathbf{x})$ is the initial condition. The corresponding space-discretized model is a high-dimensional system of nonlinear ODEs

$$\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t)), \quad \mathbf{y}(0) = \mathbf{y}_0, \tag{4}$$

where $\mathbf{y}(t) \in \mathbb{R}^{n \cdot d_s}$ is the spatially discretized state vector, \mathbf{y}_0 is the space-discretized initial condition, and $\mathbf{f}(\mathbf{y}(t))$ is an $n \cdot d_s$ -dimensional vector of real-valued nonlinear functions. We summarize the three steps of the standard Lift & Learn [32] next and summarize our notation in Table 1.

- Expose quadratic structure via lifting transformations:** Given a nonlinear PDE of the form (3), exploit knowledge of the functional form of the nonlinearity in $f(y(\mathbf{x}, t))$ to identify a lifting transformation which yields an augmented vector

Table 1 Summary of notation and dimensions for the original (non-lifted) and lifted systems introduced in Section 2.1

	Original variables (non-lifted)	Lifted variables
PDE level	$\partial y(\mathbf{x}, t) / \partial t = f(y(\mathbf{x}, t))$	$\partial \bar{y}(\mathbf{x}, t) / \partial t = \bar{a}(\bar{y}(\mathbf{x}, t)) + \bar{b}(\bar{y}(\mathbf{x}, t))$
(space-time continuous)	with d_s -dimensional state vector $y(\mathbf{x}, t)$	with d_a -dimensional state vector $\bar{y}(\mathbf{x}, t)$
FOM level	$\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t))$	$\dot{\bar{\mathbf{y}}}(t) = \bar{\mathbf{A}}\bar{\mathbf{y}}(t) + \bar{\mathbf{B}}(\bar{\mathbf{y}}(t) \otimes \bar{\mathbf{y}}(t))$
(space-discretized)	with $\mathbf{y} \in \mathbb{R}^{n \cdot d_s}$	with $\bar{\mathbf{y}} \in \mathbb{R}^{n \cdot d_a}$
ROM level	$\dot{\mathbf{y}}_r(t) = \mathbf{f}_r(\mathbf{y}_r(t))$	$\dot{\bar{\mathbf{y}}}_r(t) = \bar{\mathbf{A}}_r \bar{\mathbf{y}}_r(t) + \bar{\mathbf{B}}_r(\bar{\mathbf{y}}_r(t) \otimes \bar{\mathbf{y}}_r(t))$
(reduced-order)	with $\mathbf{y}_r \in \mathbb{R}^r$	with $\bar{\mathbf{y}}_r \in \mathbb{R}^{\bar{r}}$

of field variables in which the system dynamics have quadratic structure. The key idea of lifting is to find a quadratization $w \in \mathbb{R}^{d_w}$ of the nonlinear PDE (3) in terms of d_w additional variables $w_i = \tau_i(y)$ for $i = 1, \dots, d_w$ and then employ a transformation $\tau_{\text{lift}} : \mathbb{R}^{d_s} \rightarrow \mathbb{R}^{d_a}$ with $d_a = d_s + d_w$ that transforms the original nonlinear PDE in field variables $y \in \mathbb{R}^{d_s}$ into an equivalent lifted PDE in field variables $\bar{y} = [y^\top, w^\top]^\top \in \mathbb{R}^{d_a}$ with quadratic dynamics, i.e.,

$$\frac{\partial y(\mathbf{x}, t)}{\partial t} = f(y(\mathbf{x}, t)) \quad \xrightarrow{\bar{y}=\tau_{\text{lift}}(y)} \quad \frac{\partial \bar{y}(\mathbf{x}, t)}{\partial t} = \bar{a}(\bar{y}(\mathbf{x}, t)) + \bar{b}(\bar{y}(\mathbf{x}, t)), \quad (5)$$

where \bar{a} and \bar{b} consist of linear functions \bar{a}_j and quadratic functions \bar{b}_j , respectively, for $j = 1, \dots, d_a$.

- Construct lifted training data:** At the space-discretized level, the lifting transformation transforms the nonlinear FOM (4) into an equivalent lifted quadratic FOM, i.e.,

$$\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t)) \quad \xrightarrow{\bar{y}=\tau_{\text{lift}}(\mathbf{y})} \quad \dot{\bar{\mathbf{y}}}(t) = \bar{\mathbf{A}}\bar{\mathbf{y}}(t) + \bar{\mathbf{B}}(\bar{\mathbf{y}}(t) \otimes \bar{\mathbf{y}}(t)), \quad (6)$$

where $\bar{\mathbf{y}}(t) \in \mathbb{R}^{\bar{n}}$ with $\bar{n} = n \cdot d_a$ is the space-discretized state vector in the lifted setting, the notation \otimes denotes the Kronecker product of vectors, and $\bar{\mathbf{A}} \in \mathbb{R}^{\bar{n} \times \bar{n}}$ and $\bar{\mathbf{B}} \in \mathbb{R}^{\bar{n} \times \bar{n}^2}$ are the linear and quadratic FOM operators in the lifted setting, respectively. To obtain the data in the lifted variables, the first step is to simulate the original non-lifted FOM (4) for K time steps to construct the high-dimensional snapshot data $\mathbf{Y} \in \mathbb{R}^{n \cdot d_s \times K}$. Then, the lifting map is applied to this snapshot data matrix to obtain the lifted snapshot data matrix $\bar{\mathbf{Y}} \in \mathbb{R}^{\bar{n} \times K}$. For projection-based model reduction, the lifted state is approximated as $\bar{\mathbf{y}}(t) \approx \bar{\mathbf{V}}\bar{\mathbf{y}}_r(t)$ with a basis matrix $\bar{\mathbf{V}} \in \mathbb{R}^{\bar{n} \times \bar{r}}$ and then the reduced snapshot data matrix $\hat{\bar{\mathbf{Y}}}_r \in \mathbb{R}^{\bar{r} \times K}$ is obtained by projecting $\bar{\mathbf{Y}}$ onto the subspace spanned by $\bar{\mathbf{V}}$. Finally, the reduced time-derivative data matrix $\hat{\dot{\bar{\mathbf{Y}}}}_r \in \mathbb{R}^{\bar{r} \times K}$ is constructed from $\hat{\bar{\mathbf{Y}}}_r$ using a finite difference scheme.

- Solve the least-squares problem via Operator Inference:** The postulated quadratic form for the nonintrusive ROM based on the lifted FOM in (6) is given by

$$\dot{\bar{\mathbf{y}}}_r(t) = \bar{\mathbf{A}}_r\bar{\mathbf{y}}_r(t) + \bar{\mathbf{B}}_r(\bar{\mathbf{y}}_r(t) \otimes \bar{\mathbf{y}}_r(t)), \quad (7)$$

where the reduced operators $\bar{\mathbf{A}}_r \in \mathbb{R}^{\bar{r} \times \bar{r}}$ and $\bar{\mathbf{B}}_r \in \mathbb{R}^{\bar{r} \times \bar{r}^2}$ in the lifted setting are learned from data. Based on (7), Lift & Learn formulates the following minimization problem:

$$\min_{\bar{\mathbf{A}}_r \in \mathbb{R}^{\bar{r} \times \bar{r}}, \bar{\mathbf{B}}_r \in \mathbb{R}^{\bar{r} \times \bar{r}^2}} \|\hat{\dot{\bar{\mathbf{Y}}}}_r - \bar{\mathbf{A}}_r\hat{\bar{\mathbf{Y}}}_r - \bar{\mathbf{B}}_r(\hat{\bar{\mathbf{Y}}}_r \otimes \hat{\bar{\mathbf{Y}}}_r)\|_F. \quad (8)$$

The ability to learn the reduced operators in (7) without assuming access to the lifted FOM operators in (6) is a key advantage of the Lift & Learn approach. However, this learning approach based on exploiting the quadratic structure in the lifted setting

is not guaranteed to be structure-preserving for nonlinear conservative PDEs. This is best understood with an example.

2.2 Motivation: Lift & Learn does not conserve additional properties

Consider the one-dimensional sine-Gordon equation

$$\frac{\partial^2 \phi(x, t)}{\partial t^2} = \frac{\partial^2 \phi(x, t)}{\partial x^2} - \sin(\phi(x, t)). \tag{9}$$

The boundary conditions are periodic and the initial conditions are $(\phi(x, 0), \frac{\partial \phi}{\partial t}(x, 0)) = (0, 4/\cosh(x))$. We define $q(x, t) = \phi(x, t)$ and $p(x, t) = \frac{\partial \phi(x, t)}{\partial t}$ to rewrite the nonlinear conservative PDE (9) in first-order form. A structure-preserving spatial discretization of these first-order PDEs would lead to a $2n$ -dimensional nonlinear conservative FOM of the form

$$\dot{\mathbf{q}}(t) = \mathbf{p}(t), \quad \dot{\mathbf{p}}(t) = \mathbf{D}\mathbf{q}(t) - \sin(\mathbf{q}(t)), \tag{10}$$

which conserves the space-discretized nonlinear FOM energy $E(\mathbf{q}, \mathbf{p}, t) = \frac{1}{2}\mathbf{p}(t)^\top \mathbf{p}(t) - \frac{1}{2}\mathbf{q}(t)^\top \mathbf{D}\mathbf{q}(t) + \sum_{i=1}^n (1 - \cos(q_i(t)))$. Note that the symmetric discretization matrix $\mathbf{D} = \mathbf{D}^\top \in \mathbb{R}^{n \times n}$, also known as the symmetric discrete Laplacian, contains all the information about the spatial discretization scheme. From hereon, we simplify the notation by omitting explicit dependence on time: the space-discretized state vectors $\mathbf{q}(t)$ and $\mathbf{p}(t)$ at time t are therefore denoted as \mathbf{q} and \mathbf{p} , respectively.

Since the lifting map τ_{lift} in (5) is generally non-unique, there are multiple ways of deriving a quadratization for (9). We presented an energy-quadratization strategy in [23] that, in combination with intrusive Galerkin projection, leads to a structure-preserving quadratic ROM. Based on this energy-quadratization strategy, we introduce two auxiliary variables $w_1 = \sin(q/2)$ and $w_2 = \cos(q/2)$ that quadratize the non-linear energy and transform (10) into a $4n$ -dimensional quadratic lifted FOM

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{p}, \\ \dot{\mathbf{p}} &= \mathbf{D}\mathbf{q} - 2\mathbf{w}_1 \odot \mathbf{w}_2, \\ \dot{\mathbf{w}}_1 &= \frac{1}{2}\mathbf{w}_2 \odot \mathbf{p}, \\ \dot{\mathbf{w}}_2 &= -\frac{1}{2}\mathbf{w}_1 \odot \mathbf{p}, \end{aligned} \tag{11}$$

where the notation \odot denotes the (Hadamard) component-wise product of two vectors. This quadratic lifted FOM possesses a quadratic invariant in the lifted variables $E_{\text{lift}}(\mathbf{q}, \mathbf{p}, \mathbf{w}_1, \mathbf{w}_2) = \frac{1}{2}\mathbf{p}^\top \mathbf{p} - \frac{1}{2}\mathbf{q}^\top \mathbf{D}\mathbf{q} + 2\mathbf{w}_1^\top \mathbf{w}_1$. We use a block-diagonal basis matrix $\bar{\mathbf{V}} \in \mathbb{R}^{4n \times 4r}$ that preserves the coupling structure

$$\bar{\mathbf{V}} = \text{blkdiag}(\Phi, \Phi, \mathbf{V}_1, \mathbf{V}_2) \in \mathbb{R}^{4n \times 4r},$$

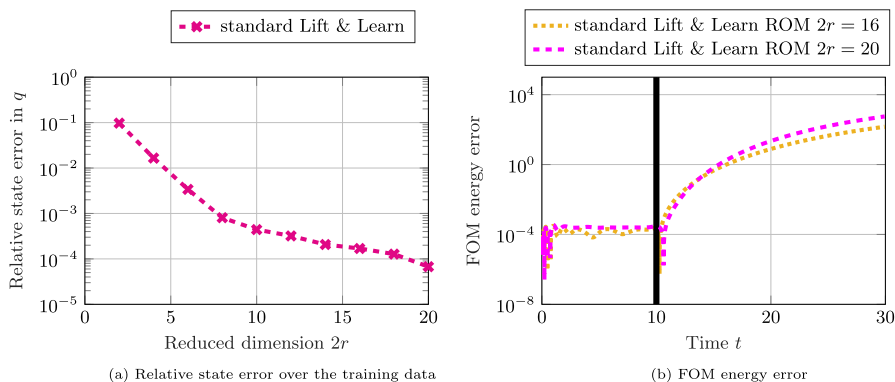


Fig. 2 One-dimensional sine-Gordon PDE. Plot **a** shows that the standard Lift & Learn approach yields quadratic ROMs that achieve relative state error below 10^{-2} for ROMs of size $2r > 6$ in the training data regime. However, the energy error comparison in plot **b** demonstrates that the standard learning approach yields unstable ROMs that exhibit unbounded energy error growth outside the training data regime. The solid black line in plot **b** indicates end of the training time interval

where Φ is the proper symplectic decomposition (PSD) basis matrix computed using the cotangent lift algorithm in [9], and \mathbf{V}_1 and \mathbf{V}_2 are the POD basis matrices for \mathbf{w}_1 and \mathbf{w}_2 , respectively.¹

We build a training dataset by simulating the nonlinear FOM of dimension $2n = 400$ from $t = 0$ to $t = 10$ with fixed time step $\Delta t = 0.005$. We then compute a block-diagonal basis matrix for the lifted FOM state and then learn quadratic ROMs from projected data. We numerically integrate the quadratic ROMs using the implicit midpoint method which is a fully implicit energy-conserving integrator for quadratic vector fields. Figure 2a shows the relative state error in the training data regime for learned quadratic ROMs of different dimensions. We observe that the state approximation error over the training data decreases monotonically from $2r = 2$ to $2r = 20$. However, the energy error plots in Fig. 2b show that standard Lift & Learn ROMs demonstrate unbounded energy error growth despite using an energy-conserving numerical integrator. This illustrates that preserving the quadratic structure alone in Lift & Learn is not enough for nonlinear FOMs with conservation laws.

3 Structure-preserving Lift & Learn

We introduce *structure-preserving Lift & Learn* (sp Lift & Learn) as a nonintrusive model reduction method to learn structure-preserving ROMs for nonlinear conservative PDEs. We leverage energy-quadratising lifting transformations that ensure the ROM is both energy-conserving and its lifted dynamics have only linear and quadratic terms. This has the additional benefit that we explicitly know the quadratic terms in the ROM (they arise from the lifting), and we only need to learn the linear terms that contain the reduced spatial-derivative operators.

¹ We showed in [23] that choosing the PSD basis matrix for \mathbf{q} and \mathbf{p} ensures that the quadratic ROM derived via intrusive projection of the lifted FOM (11) conserves the lifted FOM energy.

In Section 3.1, we derive the nonlinear conservative FOM for nonlinear conservative PDEs of the form (1). In Section 3.2, we describe energy-quadrating lifting transformations at the PDE level. In Section 3.3, we use these lifting transformations to obtain the lifted reduced data from a nonlinear conservative FOM simulation. In Section 3.4, we introduce the proposed hybrid approach to learn structure-preserving quadratic ROMs from the lifted reduced data. In Section 3.5, we compare the proposed method with the Hamiltonian Operator Inference [27] method for nonlinear wave equations. In Section 3.6, we revisit the motivational example from Section 2.2 to demonstrate the advantages of using the proposed structure-preserving Lift & Learn approach over standard Lift & Learn.

3.1 Model formulation

We rewrite (1) in first-order form by defining $q(\mathbf{x}, t) := \phi(\mathbf{x}, t)$ and $p(\mathbf{x}, t) := \frac{\partial}{\partial t}(\phi(\mathbf{x}, t))$ to obtain

$$\frac{\partial q(\mathbf{x}, t)}{\partial t} = p(\mathbf{x}, t), \quad \frac{\partial p(\mathbf{x}, t)}{\partial t} = \Delta q(\mathbf{x}, t) - f_{\text{non}}(q(\mathbf{x}, t)), \quad (12)$$

which conserves the total energy

$$\mathcal{E}[q(\mathbf{x}, t), p(\mathbf{x}, t)] := \int_{\Omega} \left(\frac{1}{2} p(\mathbf{x}, t)^2 + \frac{1}{2} (\nabla q(\mathbf{x}, t))^2 + g(q(\mathbf{x}, t)) \right) \mathrm{d}\mathbf{x}. \quad (13)$$

We assume that the nonlinear potential energy component $g(q(\mathbf{x}, t))$ in (13) is nonnegative which is typically the case for conservative PDEs. A structure-preserving spatial discretization of the first-order PDEs in (12) would have the form

$$\dot{\mathbf{q}} = \mathbf{p}, \quad \dot{\mathbf{p}} = \mathbf{D}\mathbf{q} - \mathbf{f}_{\text{non}}(\mathbf{q}), \quad (14)$$

where $\dot{\mathbf{q}} \in \mathbb{R}^n$ and $\dot{\mathbf{p}} \in \mathbb{R}^n$ denote the time derivatives of the space-discretized state vectors $\mathbf{q} \in \mathbb{R}^n$ and $\mathbf{p} \in \mathbb{R}^n$, respectively, $\mathbf{D} = \mathbf{D}^\top \in \mathbb{R}^{n \times n}$ is the symmetric discrete Laplacian, and $\mathbf{f}_{\text{non}}(\mathbf{q}) \in \mathbb{R}^n$ is the nonlinear component of the space-discretized vector field. The FOM (14) conserves the space-discretized nonlinear FOM energy $E(\mathbf{q}, \mathbf{p}) = \frac{1}{2} \mathbf{p}^\top \mathbf{p} - \frac{1}{2} \mathbf{q}^\top \mathbf{D}\mathbf{q} + \sum_{i=1}^n g(q_i)$. We assume that we know the symbolic form of the PDE (12), but we do not know how it is discretized in space and time and do not have access to the discretized operators in (14).

3.2 Exposing structure-preserving variables that quadratize energy

Given a nonlinear PDE of the form (12) with nonlinear potential energy component $g(q)$ in (13) satisfying the nonnegativity condition, we use the energy-quadrization strategy we presented in [23] to introduce the first auxiliary variable $w_1 = \tau_1(q)$ defined by $w_1^2 = \kappa^2 g(q)$, where $\kappa \in \mathbb{R}$ is a free scalar parameter. We then define the second auxiliary variable as $w_2 = \tau_2(q) = \frac{f_{\text{non}}(q)}{\kappa w_1}$ with another free

scalar parameter $\bar{\kappa} \in \mathbb{R}$. The auxiliary state dynamics for w_1 can be written as $\frac{\partial w_1(\mathbf{x}, t)}{\partial t} = \frac{\kappa^2 \bar{\kappa}}{2} w_2(\mathbf{x}, t) p(\mathbf{x}, t)$. Thus, these first two auxiliary variables ensure that the time evolution equations for $\{q, p, w_1\}$ are quadratic in terms of the lifted variables $\{q, p, w_1, w_2\}$, independent of the form of the nonlinearity in $g(q)$. We then compute auxiliary state dynamics for w_2 , and if the dynamics are not quadratic in terms of the lifted variables $\{q, p, w_1, w_2\}$, then we introduce another auxiliary variable w_3 that yields quadratic dynamics for w_2 in terms of the lifted variables $\{q, p, w_1, w_2, w_3\}$. We continue this process of introducing auxiliary variables until we find a lifted PDE with quadratic dynamics in the lifted variables.

Assuming the energy-quadratization strategy yields a quadratization with k auxiliary variables, the resulting quadratic lifted PDE can be symbolically written as

$$\begin{aligned} \frac{\partial q(\mathbf{x}, t)}{\partial t} &= p(\mathbf{x}, t), \\ \frac{\partial p(\mathbf{x}, t)}{\partial t} &= \Delta q(\mathbf{x}, t) - \bar{\kappa} w_1(\mathbf{x}, t) w_2(\mathbf{x}, t), \\ \frac{\partial w_1(\mathbf{x}, t)}{\partial t} &= \frac{\kappa^2 \bar{\kappa}}{2} w_2(\mathbf{x}, t) p(\mathbf{x}, t), \\ \frac{\partial w_j(\mathbf{x}, t)}{\partial t} &= \left(\alpha_j q(\mathbf{x}, t) + \sum_{i=1}^k \alpha_{j,i} w_i(\mathbf{x}, t) \right) p(\mathbf{x}, t), \quad \text{for } j = 2, \dots, k, \end{aligned} \tag{15}$$

where $\alpha_2, \dots, \alpha_k$ and $\alpha_{i,1}, \dots, \alpha_{i,k}$ for $i = 2, \dots, k$ are real-valued constant coefficients such that the constants in the set $\boldsymbol{\alpha}_i := \{\alpha_i, \alpha_{i,1}, \dots, \alpha_{i,k}\}$ can not be all zero for $i = 2, \dots, k$. This quadratic lifted PDE possesses a quadratic invariant in terms of the lifted variables

$$E_{\text{lift}}[q(\mathbf{x}, t), p(\mathbf{x}, t), w_1(\mathbf{x}, t), \dots, w_k(\mathbf{x}, t)] := \int_{\Omega} \left(\frac{1}{2} p(\mathbf{x}, t)^2 + \frac{1}{2} (\nabla q(\mathbf{x}, t))^2 + \frac{1}{\kappa^2} w_1(\mathbf{x}, t)^2 \right) dx. \tag{16}$$

3.3 Constructing reduced snapshot data in the lifted setting

If we were to discretize (15), the model would have the form

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{p}, \\ \dot{\mathbf{p}} &= \mathbf{D}\mathbf{q} - \bar{\kappa} \mathbf{w}_1 \odot \mathbf{w}_2, \\ \dot{\mathbf{w}}_1 &= \frac{\kappa^2 \bar{\kappa}}{2} \mathbf{w}_2 \odot \mathbf{p}, \\ \dot{\mathbf{w}}_j &= \left(\alpha_j \mathbf{q} + \sum_{i=1}^k \alpha_{j,i} \mathbf{w}_i \right) \odot \mathbf{p}, \quad \text{for } j = 2, \dots, k. \end{aligned} \tag{17}$$

This lifted FOM can be rewritten in a standard quadratic form by expressing each Hadamard product as a Kronecker-product term:

$$\dot{\mathbf{q}} = \mathbf{p},$$

$$\begin{aligned}
 \dot{\mathbf{p}} &= \mathbf{D}\mathbf{q} + \mathbf{H}_p(\mathbf{w}_1 \otimes \mathbf{w}_2), \\
 \dot{\mathbf{w}}_1 &= \mathbf{H}_{w_1}(\mathbf{w}_2 \otimes \mathbf{p}), \\
 \dot{\mathbf{w}}_j &= \mathbf{H}_{w_j}(\mathbf{q} \otimes \mathbf{p}) + \sum_{i=1}^k \mathbf{H}_{w_{j,i}}(\mathbf{w}_i \otimes \mathbf{p}), \quad \text{for } j = 2, \dots, k,
 \end{aligned}
 \tag{18}$$

where the matrices $\mathbf{H}_p \in \mathbb{R}^{n \times n^2}$, $\mathbf{H}_{w_1} \in \mathbb{R}^{n \times n^2}$, and $\mathbf{H}_{w_j} \in \mathbb{R}^{n \times n^2}$, $\mathbf{H}_{w_{j,1}} \in \mathbb{R}^{n \times n^2}, \dots, \mathbf{H}_{w_{j,k}} \in \mathbb{R}^{n \times n^2}$ for $j = 2, \dots, k$ are sparse matrices of the coefficients of the quadratic terms in (17). In this lifted setting, the quadratic FOM (17) conserves the lifted FOM energy

$$E_{\text{lift}}(\mathbf{q}, \mathbf{p}, \mathbf{w}_1, \dots, \mathbf{w}_k) = \frac{1}{2} \mathbf{p}^\top \mathbf{p} - \frac{1}{2} \mathbf{q}^\top \mathbf{D}\mathbf{q} + \frac{1}{\kappa^2} \mathbf{w}_1^\top \mathbf{w}_1.
 \tag{19}$$

We emphasize that since we derived the algebraic relations of the lifted variables to the original variables, i.e., $w_i = \tau_i(q)$ for $i = 1, \dots, k$, we can directly insert those relationships into the discretized system (c.f., the quadratic terms in (17)). In other words, we do not need to know the specific discretization scheme used for deriving (14) to determine those quadratic terms.

To construct the basis matrix for the augmented state vector in the lifted quadratic FOM (17) without ever solving that system, we first build position and momentum snapshot data matrices by simulating the nonlinear conservative FOM (14). Let $(\mathbf{q}_1, \mathbf{p}_1), \dots, (\mathbf{q}_K, \mathbf{p}_K)$ be the solutions of (14) at time t_1, \dots, t_K computed with a structure-preserving time integrator [33]. Then, the position and momentum snapshot data matrices are defined as

$$\mathbf{Q} = [\mathbf{q}_1, \dots, \mathbf{q}_K] \in \mathbb{R}^{n \times K}, \quad \mathbf{P} = [\mathbf{p}_1, \dots, \mathbf{p}_K] \in \mathbb{R}^{n \times K}.
 \tag{20}$$

We compute the lifted states as $\mathbf{w}_{i,j} = \tau_i(\mathbf{q}_j)$ and construct the lifted snapshot data matrix for each lifted variable

$$\mathbf{W}_i = [\mathbf{w}_{i,1}, \dots, \mathbf{w}_{i,K}] \in \mathbb{R}^{n \times K}, \quad i = 1, \dots, k.
 \tag{21}$$

To obtain low-dimensional data via projection, we use a block-diagonal basis matrix $\tilde{\mathbf{V}} \in \mathbb{R}^{\tilde{n} \times \tilde{r}}$ that preserves the coupling structure

$$\tilde{\mathbf{V}} = \text{blkdiag}(\Phi, \Phi, \mathbf{V}_1, \dots, \mathbf{V}_k) \in \mathbb{R}^{\tilde{n} \times \tilde{r}},
 \tag{22}$$

where Φ is the PSD basis matrix for \mathbf{q} and \mathbf{p} computed using the cotangent lift algorithm, and \mathbf{V}_i is the POD basis matrix that contains as columns the POD basis vectors for \mathbf{w}_i for $i = 1, \dots, k$.² We obtain projections of the snapshot data matrices as

$$\hat{\mathbf{Q}} = \Phi^\top \mathbf{Q} \in \mathbb{R}^{r \times K}, \quad \hat{\mathbf{P}} = \Phi^\top \mathbf{P} \in \mathbb{R}^{r \times K}, \quad \hat{\mathbf{W}}_i = \mathbf{V}_i^\top \mathbf{W}_i \in \mathbb{R}^{r \times K}, \quad i = 1, \dots, k.
 \tag{23}$$

² The PSD basis matrix Φ in the cotangent lift algorithm is computed via the singular value decomposition of the extended snapshot data matrix $\mathbf{Y}_e := [\mathbf{Q}, \mathbf{P}] \in \mathbb{R}^{n \times 2K}$ whereas the POD basis matrix \mathbf{V}_i for \mathbf{w}_i is computed via the singular value decomposition of the corresponding snapshot data matrix $\mathbf{W}_i \in \mathbb{R}^{n \times K}$.

Using the reduced momentum trajectories, we obtain reduced time-derivative data $\hat{\mathbf{P}} \in \mathbb{R}^{r \times K}$ via an eighth-order central finite difference scheme. Although lower-order schemes (such as second- or third-order finite difference schemes) may suffice, we use an eighth-order scheme to make negligible derivative approximation errors that could affect the learned ROM operators.

3.4 Learning structure-preserving quadratic ROMs via constrained operator inference

Based on the block-diagonal basis matrix $\tilde{\mathbf{V}}$ in (22), the postulated model form—which would arise if we were to project the lifted FOM (17) with $\tilde{\mathbf{V}}$ —for the structure-preserving ROM is

$$\begin{aligned} \hat{\mathbf{q}} &= \hat{\mathbf{p}}, \\ \hat{\mathbf{p}} &= \hat{\mathbf{D}}\hat{\mathbf{q}} + \Phi^\top \mathbf{H}_p (\mathbf{V}_1 \hat{\mathbf{w}}_1 \otimes \mathbf{V}_2 \hat{\mathbf{w}}_2), \\ \hat{\mathbf{w}}_1 &= \mathbf{V}_1^\top \mathbf{H}_{w_1} (\mathbf{V}_2 \hat{\mathbf{w}}_2 \otimes \Phi \hat{\mathbf{p}}), \\ \hat{\mathbf{w}}_j &= \mathbf{V}_j^\top \mathbf{H}_{w_j} (\Phi \hat{\mathbf{q}} \otimes \Phi \hat{\mathbf{p}}) + \sum_{i=1}^k \mathbf{V}_j^\top \mathbf{H}_{w_{j,i}} (\mathbf{V}_i \hat{\mathbf{w}}_i \otimes \Phi \hat{\mathbf{p}}), \quad \text{for } j = 2, \dots, k, \end{aligned} \tag{24}$$

where $\hat{\mathbf{D}} = \hat{\mathbf{D}}^\top \in \mathbb{R}^{r \times r}$ is the nonintrusive analogue of the symmetric ROM operator that would arise if we were to project intrusively. Since none of the quadratic terms in the lifted FOM (17) requires approximation of spatial derivatives, the corresponding reduced quadratic terms in (24) after projection can be computed analytically. Hence, $\hat{\mathbf{D}} = \hat{\mathbf{D}}^\top \in \mathbb{R}^{r \times r}$ is the only matrix that we need to learn from data. We simplify the notation by rewriting the quadratic ROM form in (24) as

$$\begin{aligned} \hat{\mathbf{q}} &= \hat{\mathbf{p}}, \\ \hat{\mathbf{p}} &= \hat{\mathbf{D}}\hat{\mathbf{q}} + \hat{\mathbf{H}}_p (\hat{\mathbf{w}}_1 \otimes \hat{\mathbf{w}}_2), \\ \hat{\mathbf{w}}_1 &= \hat{\mathbf{H}}_{w_1} (\hat{\mathbf{w}}_2 \otimes \hat{\mathbf{p}}), \\ \hat{\mathbf{w}}_j &= \hat{\mathbf{H}}_{w_j} (\hat{\mathbf{q}} \otimes \hat{\mathbf{p}}) + \sum_{i=1}^k \hat{\mathbf{H}}_{w_{j,i}} (\hat{\mathbf{w}}_i \otimes \hat{\mathbf{p}}), \quad \text{for } j = 2, \dots, k. \end{aligned} \tag{25}$$

where $\hat{\mathbf{H}}_p := \Phi^\top \mathbf{H}_p (\mathbf{V}_1 \otimes \mathbf{V}_2) \in \mathbb{R}^{r \times r^2}$, $\hat{\mathbf{H}}_{w_1} := \mathbf{V}_1^\top \mathbf{H}_{w_1} (\mathbf{V}_2 \otimes \Phi) \in \mathbb{R}^{r \times r^2}$, $\hat{\mathbf{H}}_{w_j} := \mathbf{V}_j^\top \mathbf{H}_{w_j} (\Phi \otimes \Phi) \in \mathbb{R}^{r \times r^2}$, and $\hat{\mathbf{H}}_{w_{j,i}} := \mathbf{V}_j^\top \mathbf{H}_{w_{j,i}} (\mathbf{V}_i \otimes \Phi) \in \mathbb{R}^{r \times r^2}$ for $i = 1, \dots, k$ and $j = 2, \dots, k$. As mentioned above, the matrices $\hat{\mathbf{H}}_p$, $\hat{\mathbf{H}}_{w_1}$, and $\hat{\mathbf{H}}_{w_j}$, $\hat{\mathbf{H}}_{w_{j,1}}, \dots, \hat{\mathbf{H}}_{w_{j,k}}$ for $j = 2, \dots, k$ are merely representations of the quadratic products that we already derived analytically in (24). Thus, there is no need to learn them.

Next, we use the analytically constructed quadratic terms in (25) to infer $\hat{\mathbf{D}}$ from projections of the FOM snapshot data. Based on the postulated model form for the

quadratic ROM in (25), we formulate the following constrained optimization problem for learning $\widehat{\mathbf{D}}$:

$$\min_{\widehat{\mathbf{D}}=\widehat{\mathbf{D}}^\top} \|\widehat{\mathbf{P}} - \widehat{\mathbf{H}}_p(\widehat{\mathbf{W}}_1 \otimes \widehat{\mathbf{W}}_2) - \widehat{\mathbf{D}}\widehat{\mathbf{Q}}\|_F. \tag{26}$$

Although this paper focuses on nonlinear conservative PDEs of the form (1), we note that the proposed approach can be flexibly adapted to learn structure-preserving ROMs for a more general class of coupled conservative PDEs. In Section 4.3, we provide evidence for the wide scope of the proposed approach through the numerical example of two-dimensional Klein-Gordon-Zakharov equations, a system of coupled conservative PDEs that does not have a canonical Hamiltonian formulation.

Proposition 1 *The structure-preserving Lift & Learn ROM (25) conserves the perturbed lifted FOM energy*

$$\widehat{E}_{\text{lift}}(\widehat{\mathbf{q}}, \widehat{\mathbf{p}}, \widehat{\mathbf{w}}_1, \dots, \widehat{\mathbf{w}}_k) := E_{\text{lift}}(\Phi\widehat{\mathbf{q}}, \Phi\widehat{\mathbf{p}}, \mathbf{V}_1\widehat{\mathbf{w}}_1, \dots, \mathbf{V}_k\widehat{\mathbf{w}}_k) + \Delta E_{\text{lift}}(\widehat{\mathbf{q}}) \tag{27}$$

with a perturbation of the form $\Delta E_{\text{lift}}(\widehat{\mathbf{q}}) = \frac{1}{2}\widehat{\mathbf{q}}^\top(\Phi^\top\mathbf{D}\Phi - \widehat{\mathbf{D}})\widehat{\mathbf{q}}$ where $\mathbf{D} = \mathbf{D}^\top \in \mathbb{R}^{n \times n}$ is from (17).

Proof Given a lifted quadratic FOM of the form (17), we evaluate the lifted FOM energy function E_{lift} (19) at the lifted state approximation to obtain

$$\begin{aligned} E_{\text{lift}}(\Phi\widehat{\mathbf{q}}, \Phi\widehat{\mathbf{p}}, \mathbf{V}_1\widehat{\mathbf{w}}_1, \dots, \mathbf{V}_k\widehat{\mathbf{w}}_k) &= \frac{1}{2}(\Phi\widehat{\mathbf{p}})^\top(\Phi\widehat{\mathbf{p}}) - \frac{1}{2}(\Phi\widehat{\mathbf{q}})^\top\mathbf{D}(\Phi\widehat{\mathbf{q}}) + \frac{1}{\kappa^2}(\mathbf{V}_1\widehat{\mathbf{w}}_1)^\top(\mathbf{V}_1\widehat{\mathbf{w}}_1) \\ &= \frac{1}{2}\widehat{\mathbf{p}}^\top\widehat{\mathbf{p}} - \frac{1}{2}\widehat{\mathbf{q}}^\top(\Phi^\top\mathbf{D}\Phi)\widehat{\mathbf{q}} + \frac{1}{\kappa^2}\widehat{\mathbf{w}}_1^\top\widehat{\mathbf{w}}_1. \end{aligned} \tag{28}$$

Defining $\Delta E_{\text{lift}}(\widehat{\mathbf{q}}) := \frac{1}{2}\widehat{\mathbf{q}}^\top(\Phi^\top\mathbf{D}\Phi - \widehat{\mathbf{D}})\widehat{\mathbf{q}}$, we see that the energy associated with the lifted ROM (25) is—analogueous to (19)—given by

$$\begin{aligned} \widehat{E}_{\text{lift}}(\widehat{\mathbf{q}}, \widehat{\mathbf{p}}, \widehat{\mathbf{w}}_1, \dots, \widehat{\mathbf{w}}_k) &= \frac{1}{2}\widehat{\mathbf{p}}^\top\widehat{\mathbf{p}} - \frac{1}{2}\widehat{\mathbf{q}}^\top(\Phi^\top\mathbf{D}\Phi)\widehat{\mathbf{q}} + \frac{1}{\kappa^2}\widehat{\mathbf{w}}_1^\top\widehat{\mathbf{w}}_1 \\ &+ \frac{1}{2}\widehat{\mathbf{q}}^\top(\Phi^\top\mathbf{D}\Phi - \widehat{\mathbf{D}})\widehat{\mathbf{q}} = \frac{1}{2}\widehat{\mathbf{p}}^\top\widehat{\mathbf{p}} - \frac{1}{2}\widehat{\mathbf{q}}^\top\widehat{\mathbf{D}}\widehat{\mathbf{q}} + \frac{1}{\kappa^2}\widehat{\mathbf{w}}_1^\top\widehat{\mathbf{w}}_1. \end{aligned}$$

We now compute the time-derivative of $\widehat{E}_{\text{lift}}(\widehat{\mathbf{q}}, \widehat{\mathbf{p}}, \widehat{\mathbf{w}}_1, \dots, \widehat{\mathbf{w}}_k)$

$$\begin{aligned} \frac{d}{dt}\widehat{E}_{\text{lift}}(\widehat{\mathbf{q}}, \widehat{\mathbf{p}}, \widehat{\mathbf{w}}_1, \dots, \widehat{\mathbf{w}}_k) &= \widehat{\mathbf{p}}^\top\widehat{\dot{\mathbf{p}}} - (\widehat{\mathbf{D}}\widehat{\mathbf{q}})^\top\widehat{\dot{\mathbf{q}}} + \frac{2}{\kappa^2}\widehat{\mathbf{w}}_1^\top\widehat{\dot{\mathbf{w}}}_1 \\ &= -\bar{\kappa}\widehat{\mathbf{p}}^\top\Phi^\top(\mathbf{V}_1\widehat{\mathbf{w}}_1 \odot \mathbf{V}_2\widehat{\mathbf{w}}_2) + \bar{\kappa}\widehat{\mathbf{w}}_1^\top\mathbf{V}_1^\top(\mathbf{V}_2\widehat{\mathbf{w}}_2 \odot \Phi\widehat{\mathbf{p}}) \\ &= 0. \end{aligned}$$

Thus, the quadratic ROM learned via structure-preserving Lift & Learn conserves $\widehat{E}_{\text{lift}}(\widehat{\mathbf{q}}, \widehat{\mathbf{p}}, \widehat{\mathbf{w}}_1, \dots, \widehat{\mathbf{w}}_k)$. □

We showed in [23] that the structure-preserving quadratic ROM obtained via intrusive lifting conserves the lifted FOM energy exactly, which means that it possesses a

quadratic invariant of the form $\tilde{E}_{\text{lift}}(\hat{\mathbf{q}}, \hat{\mathbf{p}}, \hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_k) := E_{\text{lift}}(\Phi\hat{\mathbf{q}}, \Phi\hat{\mathbf{p}}, \mathbf{V}_1\hat{\mathbf{w}}_1, \dots, \mathbf{V}_k\hat{\mathbf{w}}_k)$. Building on this result, the quadratic invariant $\hat{E}_{\text{lift}}(\hat{\mathbf{q}}, \hat{\mathbf{p}}, \hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_k)$ in (27) can be interpreted as a perturbation of $\tilde{E}_{\text{lift}}(\hat{\mathbf{q}}, \hat{\mathbf{p}}, \hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_k)$. Moreover, this perturbation term is bounded as follows:

$$|\Delta E_{\text{lift}}(\hat{\mathbf{q}})| = \frac{1}{2}\hat{\mathbf{q}}^\top(\tilde{\mathbf{D}} - \hat{\mathbf{D}})\hat{\mathbf{q}} \leq \frac{1}{2}\|\tilde{\mathbf{D}} - \hat{\mathbf{D}}\|\|\hat{\mathbf{q}}\|^2, \tag{29}$$

where $\tilde{\mathbf{D}} := \Phi^\top \mathbf{D} \Phi \in \mathbb{R}^{r \times r}$ is the linear ROM operator obtained via intrusive projection. Thus, the structure-preserving Lift & Learn ROM trajectories conserve a perturbed lifted FOM energy where the magnitude of the perturbation depends on the difference between the learned linear ROM operator $\hat{\mathbf{D}}$ and the intrusively projected linear ROM operator $\tilde{\mathbf{D}}$.

Remark 1 The nonlinear FOM for conservative PDEs of the form (1) can also be written in the Lagrangian form where the governing equations are a set of n coupled second-order ordinary differential equations. The Lagrangian Operator Inference method [29] for nonlinear wave equations also uses knowledge of the nonlinear potential energy at the PDE level to learn low-dimensional nonlinear Lagrangian ROMs from high-dimensional FOM data. However, these learned Lagrangian ROMs suffer from computational efficiency issues as the evaluation of the nonlinear ROM vector field scales with the FOM dimension. To tackle this challenge and have efficient ROMs, the Lagrangian Operator Inference approach would need an additional structure-preserving hyper-reduction step, which to the best of the authors’ knowledge remains an open problem. In contrast, the proposed structure-preserving Lift & Learn approach yields computationally efficient quadratic ROMs that do not require a second level of approximation.

Remark 2 While our knowledge of the symbolic form of the quadratic lifted PDE (15) enables us to derive the quadratic ROM form (24), we do not formally define, or numerically solve, the lifted PDE (15).

3.5 Comparison of structure-preserving Lift & Learn and Hamiltonian Operator Inference

In this section and in the numerical results, we compare the proposed structure-preserving Lift & Learn approach with Hamiltonian Operator Inference (HOPI) [27], a structure-preserving learning approach for canonical Hamiltonian PDEs. The HOPI method also uses knowledge of the functional form of the nonlinear potential energy at the PDE level to postulate a Hamiltonian model of the form

$$\hat{\mathbf{q}} = \hat{\mathbf{D}}_{\mathbf{q}}\hat{\mathbf{p}}, \quad \hat{\mathbf{p}} = \hat{\mathbf{D}}_{\mathbf{p}}\hat{\mathbf{q}} + \Phi^\top \mathbf{f}_{\text{non}}(\Phi\hat{\mathbf{q}}), \tag{30}$$

where the symmetric reduced matrices $\hat{\mathbf{D}}_{\mathbf{q}} = \hat{\mathbf{D}}_{\mathbf{q}}^\top \in \mathbb{R}^{r \times r}$ and $\hat{\mathbf{D}}_{\mathbf{p}} = \hat{\mathbf{D}}_{\mathbf{p}}^\top \in \mathbb{R}^{r \times r}$ are inferred from data. The symmetric constraints on $\hat{\mathbf{D}}_{\mathbf{q}}$ and $\hat{\mathbf{D}}_{\mathbf{p}}$ ensure that the linear components of the learned Hamiltonian ROM retain the symmetric property of the linear FOM operator introduced during the structure-preserving discretization.

Given K reduced state snapshots and the reduced time-derivative data, HOPIInf formulates the following separate symmetric linear least-squares problems for learning $\widehat{\mathbf{D}}_q$ and $\widehat{\mathbf{D}}_p$:

$$\min_{\widehat{\mathbf{D}}_q = \widehat{\mathbf{D}}_q^T} \|\widehat{\mathbf{Q}} - \widehat{\mathbf{D}}_q \widehat{\mathbf{P}}\|_F, \quad \min_{\widehat{\mathbf{D}}_p = \widehat{\mathbf{D}}_p^T} \|\widehat{\mathbf{P}} - \widehat{\mathbf{F}}_{\text{non}}(\widehat{\mathbf{Q}}) - \widehat{\mathbf{D}}_p \widehat{\mathbf{Q}}\|_F. \tag{31}$$

The HOPIInf approach, similar to Lagrangian Operator Inference, still requires evaluating the nonlinear vector field (30) that scales with the FOM dimension. To reduce the online computational cost of simulating (30), the HOPIInf approach needs an additional structure-preserving hyper-reduction step. The authors in [17] presented a gradient-preserving DEIM strategy that yields nonlinear Hamiltonian ROMs that conserve the nonlinear FOM energy in an asymptotic sense at a computational cost independent of the FOM dimension. We therefore consider the HOPIInf approach combined with structure-preserving DEIM (spDEIM) as the state of the art and use it to compare our structure-preserving Lift & Learn method with.

3.6 Revisiting the one-dimensional sine-Gordon equation

For a numerical illustration, we revisit the sine-Gordon equation example from Section 2. The proposed structure-preserving Lift & Learn method learns $4r$ -dimensional quadratic ROMs of the form

$$\begin{aligned} \widehat{\mathbf{q}} &= \widehat{\mathbf{p}}, \\ \widehat{\mathbf{p}} &= \widehat{\mathbf{D}}\widehat{\mathbf{q}} + \widehat{\mathbf{H}}_p(\widehat{\mathbf{w}}_1 \otimes \widehat{\mathbf{w}}_2), \\ \widehat{\mathbf{w}}_1 &= \widehat{\mathbf{H}}_{w_1}(\widehat{\mathbf{w}}_2 \otimes \widehat{\mathbf{p}}), \\ \widehat{\mathbf{w}}_2 &= \widehat{\mathbf{H}}_{w_{2,1}}(\widehat{\mathbf{w}}_1 \otimes \widehat{\mathbf{p}}), \end{aligned} \tag{32}$$

where the quadratic ROM operators $\widehat{\mathbf{H}}_p \in \mathbb{R}^{r \times r^2}$, $\widehat{\mathbf{H}}_{w_1} \in \mathbb{R}^{r \times r^2}$, and $\widehat{\mathbf{H}}_{w_{2,1}} \in \mathbb{R}^{r \times r^2}$ are computed analytically, and the symmetric linear ROM operator $\widehat{\mathbf{D}} = \widehat{\mathbf{D}}^T \in \mathbb{R}^{r \times r}$ is learned from data.

In Fig. 3, we compare the accuracy and energy error behavior for structure-preserving Lift & Learn ROMs with $2r$ -dimensional nonlinear Hamiltonian ROMs obtained via HOPIInf with $r_{\text{spDEIM}} = 2r$. The relative state error comparison in Fig. 3a shows that both approaches achieve similar accuracy in the training data regime up to $2r = 10$. For $2r > 10$, we observe that the accuracy in the training data does not decrease as favorably with an increase in the reduced dimension. The energy error comparison in Fig. 3b shows that both approaches achieve bounded FOM energy error that stays below 10^{-1} for all times. Compared to Fig. 2b, the bounded energy error 200% past the training time interval for the structure-preserving Lift & Learn ROM shows that the proposed approach learns ROMs that respect the underlying physics of the problem.

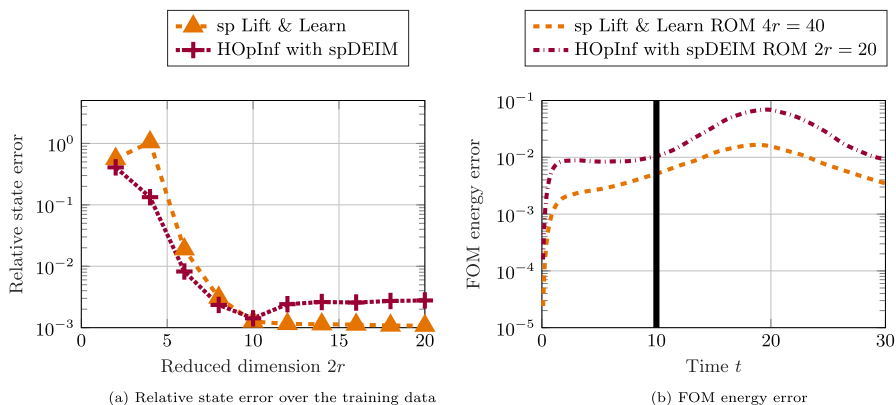


Fig. 3 Comparison of structure-preserving Lift & Learn and HOPInf with spDEIM. Plots **a** and **b** show that both approaches learn accurate and stable ROMs with bounded FOM energy error

4 Numerical results

In this section, we investigate the numerical performance of the proposed structure-preserving Lift & Learn method for three nonlinear conservative PDEs. In Section 4.1, we study the one-dimensional nonlinear wave equation with exponential nonlinearity. In Section 4.2, we consider the two-dimensional sine-Gordon equation. Finally, in Section 4.3, we consider a nonlinear conservative FOM with 960,000° of freedom to derive structure-preserving ROMs for the two-dimensional Klein-Gordon-Zakharov equations, a system of coupled conservative PDEs that does not have a canonical Hamiltonian formulation. Below, we list practical details regarding the numerical implementation and the error measures used in the numerical experiments.

- The numerical experiments in Sections 4.1 and 4.2 are implemented in MATLAB 2022a on a quad-core Intel i7 processor with 2.3 GHz and 32 GB memory. The numerical experiments for the two-dimensional Klein-Gordon-Zakharov equations in Section 4.3 are implemented in MATLAB 2022b using compute nodes of the Triton Shared Computing Cluster [34] equipped with 8 processing cores of Intel Xeon Platinum 64-core CPU at 2.9 GHz and 1 TB memory.
- The constrained optimization problems to learn the linear reduced operators in this work are solved using the CVX 2.2 [35], a MATLAB-based modeling system for convex optimization. The symmetric constraints and the optimization objectives for a given constrained optimization problem are specified using standard MATLAB expression syntax in the CVX package.
- For nonlinear Hamiltonian ROMs obtained via HOPInf [27], we use the implicit midpoint rule for numerical time integration. The implicit midpoint rule is a second-order symplectic integrator that exhibits bounded energy error for nonlinear Hamiltonian systems [36].
- For quadratic ROMs obtained via the proposed structure-preserving Lift & Learn approach, we use Kahan's method for numerical time integration. Kahan's method is a second-order structure-preserving numerical integrator for quadratic vector fields (see [37] for more details about Kahan's method).

- The *relative state error* in q is computed in the entire training or testing intervals as

$$\text{Relative state error in } q = \frac{\|\mathbf{Q} - \Phi\widehat{\mathbf{Q}}\|_F^2}{\|\mathbf{Q}\|_F^2}, \tag{33}$$

where $\widehat{\mathbf{Q}} = [\widehat{\mathbf{q}}_1, \dots, \widehat{\mathbf{q}}_K] \in \mathbb{R}^{r \times K}$ is the ROM snapshot data obtained from the ROM simulations and $\Phi\widehat{\mathbf{Q}} \in \mathbb{R}^{n \times K}$ is the approximation of the FOM snapshot data \mathbf{Q} .

- We measure the common cost/accuracy tradeoff for ROMs using the *efficacy* metric which is computed as

$$\text{Efficacy} = \frac{1}{\text{relative state error in training data regime} \times \text{wall-clock time in seconds}}, \tag{34}$$

where the MATLAB wall-clock time is obtained by calculating the average over 20 runs. In comparisons based on this metric, the model reduction approach with higher efficacy is considered advantageous. The reported models are sufficiently accurate such that reporting efficacy is sensible.

- The *FOM energy error* is computed as follows:

$$\text{FOM energy error} = |E(\Phi\widehat{\mathbf{q}}(t), \Phi\widehat{\mathbf{p}}(t)) - E(\Phi\widehat{\mathbf{q}}(0), \Phi\widehat{\mathbf{p}}(0))|, \tag{35}$$

where $E(\Phi\widehat{\mathbf{q}}(t), \Phi\widehat{\mathbf{p}}(t))$ is the FOM energy approximation obtained by evaluating the space-discretized nonlinear FOM energy E for the FOM state approximation at time t .

4.1 Nonlinear wave equation with exponential nonlinearity

The one-dimensional nonlinear wave equation with exponential nonlinearity considered here arises from the Johnson–Mehl–Avrami–Kolmogorov theory [38–40] of nucleation and growth reactions for modeling the kinetics of phase change.

4.1.1 PDE formulation and the corresponding nonlinear conservative FOM

We consider the FOM setup from [41]. Let $\Omega = (0, \pi) \subset \mathbb{R}$ be the spatial domain and consider the one-dimensional nonlinear wave equation

$$\frac{\partial^2 \phi(x, t)}{\partial t^2} = \frac{\partial^2 \phi(x, t)}{\partial x^2} + \exp(-\phi(x, t)), \tag{36}$$

with scalar field variable $\phi(x, t)$ at spatial location $x \in \Omega$ and time $t \in (0, T]$. We consider homogeneous Dirichlet boundary conditions $\phi(0, t) = \phi(\pi, t) = 0$ for $t \in (0, T]$. The corresponding initial conditions are $(\phi(x, 0), \frac{\partial \phi}{\partial t}(x, 0)) = (0.5x(\pi - x), 0)$. We define $q(x, t) := \phi(x, t)$ and $p(x, t) := \frac{\partial \phi(x, t)}{\partial t}$ to rewrite (36) in first-order form and then discretize the system of first-order PDEs using $n = 200$ equally spaced grid points to derive the nonlinear conservative FOM

$$\dot{\mathbf{q}} = \mathbf{p}, \quad \dot{\mathbf{p}} = \mathbf{D}\mathbf{q} + \exp(-\mathbf{q}),$$

where $\mathbf{D} = \mathbf{D}^\top$ denotes the symmetric discrete Laplacian. The nonlinear conservative FOM conserves the space-discretized energy

$$E(\mathbf{q}, \mathbf{p}) = \frac{1}{2} \mathbf{p}^\top \mathbf{p} - \frac{1}{2} \mathbf{q}^\top \mathbf{D} \mathbf{q} + \sum_{i=1}^n (\exp(-q_i)).$$

4.1.2 Model form for structure-preserving Lift & Learn based on energy quadratization

Based on the exponential term in the nonlinear FOM energy expression, we solve $w^2 = \kappa^2 g(q)$ with $g(q) = \exp(-q)$ and choose $\kappa = 1$ to obtain the first auxiliary variable $w = \exp(-q/2)$, which quadratizes the nonlinear energy in the lifted variables. Since the time evolution equations for $\{\mathbf{q}, \mathbf{p}, \mathbf{w}\}$ are quadratic in terms of $\{\mathbf{q}, \mathbf{p}, \mathbf{w}\}$ in this example, we do not need to introduce additional auxiliary variables. The resulting lifted FOM

$$\begin{aligned} \dot{\mathbf{q}} &= \mathbf{p}, \\ \dot{\mathbf{p}} &= \mathbf{D} \mathbf{q} + \mathbf{w} \odot \mathbf{w}, \\ \dot{\mathbf{w}} &= -\frac{1}{2} \mathbf{w} \odot \mathbf{p}, \end{aligned}$$

conserves the lifted FOM energy $E_{\text{lift}}(\mathbf{q}, \mathbf{p}, \mathbf{w}) = \frac{1}{2} \mathbf{p}^\top \mathbf{p} - \frac{1}{2} \mathbf{q}^\top \mathbf{D} \mathbf{q} + \mathbf{w}^\top \mathbf{w}$. We consider a block-diagonal basis matrix of the form

$$\tilde{\mathbf{V}} = \text{blkdiag}(\Phi, \Phi, \mathbf{V}) \in \mathbb{R}^{3n \times 3r},$$

where the PSD basis matrix $\Phi \in \mathbb{R}^{n \times r}$ for \mathbf{q} and \mathbf{p} is computed using the cotangent lift algorithm and the POD basis matrix $\mathbf{V} \in \mathbb{R}^{n \times r}$ for the auxiliary variable \mathbf{w} is computed via the singular value decomposition of the lifted snapshot data matrix $\mathbf{W} \in \mathbb{R}^{n \times r}$. The model form for the nonintrusive quadratic ROM is

$$\begin{aligned} \hat{\mathbf{q}} &= \hat{\mathbf{p}}, \\ \hat{\mathbf{p}} &= \hat{\mathbf{D}} \hat{\mathbf{q}} + \hat{\mathbf{H}}_{\mathbf{p}} (\hat{\mathbf{w}} \otimes \hat{\mathbf{w}}), \\ \hat{\mathbf{w}} &= \hat{\mathbf{H}}_{\mathbf{w}} (\hat{\mathbf{w}} \otimes \hat{\mathbf{p}}), \end{aligned}$$

where the reduced quadratic operators $\hat{\mathbf{H}}_{\mathbf{p}} \in \mathbb{R}^{r \times r^2}$ and $\hat{\mathbf{H}}_{\mathbf{w}} \in \mathbb{R}^{r \times r^2}$ are computed analytically.

4.1.3 Numerical results

The training dataset is built by integrating the nonlinear FOM until time $t = 10$ using the symplectic midpoint rule with $\Delta t = 0.005$. We consider FOM snapshots from $t = 10$ to $t = 100$ as the test dataset to study the predictive capability of quadratic ROMs learned via structure-preserving Lift & Learn. For comparison, we consider quadratic ROMs obtained via lifting and intrusive projection, and nonlinear ROMs obtained via HOpInf with $r_{\text{sPDEIM}} = r$.

In Fig. 4, we evaluate the numerical performance of the aforementioned approaches by comparing the relative state error (33) for different reduced dimensions. The state error plots in Fig. 4a and b show that all three approaches yield ROMs with similar accuracy in both training and test data regimes.

We compare the computational efficiency of structure-preserving Lift & Learn against the HOpInf with spDEIM approach with $r_{\text{spDEIM}} = r$ through efficacy plots in Fig. 5a. The $3r$ -dimensional quadratic ROMs learned via structure-preserving Lift & Learn achieve higher efficacy than the $2r$ -dimensional nonlinear Hamiltonian ROMs obtained via HOpInf with spDEIM. Thus, compared to HOpInf with spDEIM, the proposed approach learns ROMs that achieve similar state error at a substantially lower computational cost for this example. The FOM energy error plots in Fig. 5b show that all three approaches achieve bounded FOM energy error that stays below 10^{-3} for all times. In this example, the FOM energy is approximately 2.80, while the perturbation term ΔE_{lift} discussed in (27) remains on the order of 10^{-4} – 10^{-5} throughout the simulation. These results demonstrate the ability of the learned quadratic ROM to provide accurate and stable predictions far outside the training data regime.

4.2 Two-dimensional sine-Gordon wave equation

We now consider the two-dimensional analogue of the sine-Gordon equation from Section 2.

4.2.1 PDE formulation and the corresponding nonlinear conservative FOM

We consider the FOM setup from [42]. Let $\Omega = (-7, 7) \times (-7, 7) \subset \mathbb{R}^2$ be the spatial domain and consider the two-dimensional sine-Gordon equation

$$\frac{\partial^2 \phi}{\partial t^2}(x, y, t) = \frac{\partial^2 \phi}{\partial x^2}(x, y, t) + \frac{\partial^2 \phi}{\partial y^2}(x, y, t) - \sin(\phi(x, y, t)), \tag{37}$$

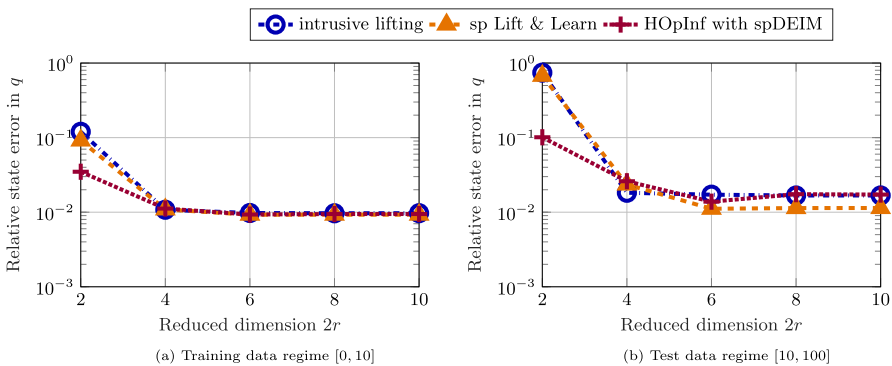


Fig. 4 Nonlinear wave equation with exponential nonlinearity. Quadratic ROMs obtained via structure-preserving Lift & Learn achieve similar state error to nonlinear ROMs obtained via intrusive lifting and HOpInf with spDEIM in both training and test data regimes

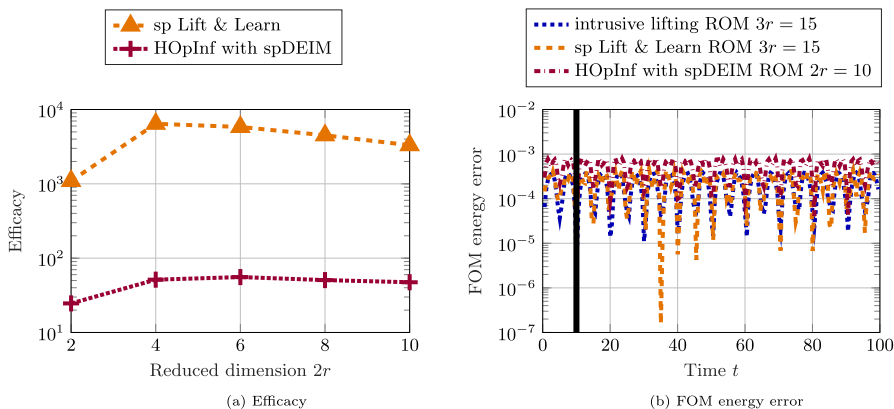


Fig. 5 Nonlinear wave equation with exponential nonlinearity. The efficacy comparison in plot **a** shows that the structure-preserving Lift & Learn approach achieves higher efficacy than the HOPIInf with spDEIM approach. Plot **b** shows that all ROMs demonstrate bounded energy error. The solid black line in plot **b** indicates the end of the training data regime

for $t \in (0, T]$. The boundary conditions are periodic and the corresponding initial conditions are

$$\phi(x, y, 0) = 4 \tan^{-1} \left(\exp \left((3 - \sqrt{x^2 + y^2}) \right) \right), \quad \frac{\partial \phi}{\partial t}(x, y, 0) = 0.$$

We define $q(x, y, t) := \phi(x, y, t)$ and $p(x, y, t) := \partial \phi(x, y, t) / \partial t$ and rewrite the second-order conservative PDE (37) in the first-order form. We then discretize the computational domain Ω with $n_x = n_y = 100$ equally spaced grid points in both spatial directions leading to a nonlinear FOM of dimension $2n = 20,000$

$$\dot{\mathbf{q}} = \mathbf{p}, \quad \dot{\mathbf{p}} = \mathbf{D}\mathbf{q} - \sin(\mathbf{q}),$$

where $\mathbf{D} = \mathbf{D}^\top$ denotes the symmetric discrete Laplacian in the two-dimensional setting. The nonlinear FOM conserves the space-discretized energy

$$E(\mathbf{q}, \mathbf{p}) = \frac{1}{2} \mathbf{p}^\top \mathbf{p} - \frac{1}{2} \mathbf{q}^\top \mathbf{D}\mathbf{q} + \sum_{i=1}^n (1 - \cos(q_i)).$$

4.2.2 Model form for structure-preserving Lift & Learn based on energy quadratization

Since the two-dimensional sine-Gordon equation (37) has the same form of nonlinearity as its one-dimensional counterpart in equation (9), the nonintrusive ROM form has the same structure as the nonintrusive ROM for the one-dimensional sine-Gordon example in equation (32).

4.2.3 Numerical results

We build the training dataset by integrating the nonlinear FOM for the two-dimensional sine-Gordon equation from $t = 0$ to $t = 10$ and we consider FOM snapshots from $t = 10$ to $t = 12.5$ (predicting 25% outside training data) as the test dataset. We compare the relative state error of the structure-preserving Lift & Learn ROMs against intrusive lifting ROMs and HOpInf ROMs with $r_{\text{spDEIM}} = 6r$ in Fig. 6. The relative state error plots in Fig. 6a show that the quadratic ROMs learned via the proposed structure-preserving Lift & Learn approach achieve accuracy similar to the intrusive lifting ROMs in the training data regime. Notably, the relative state error for HOpInf ROMs with $r_{\text{spDEIM}} = 6r$ does not decrease as favorably with an increase for $2r > 20$. This could be remedied by further increasing r , which, however, would make the HOpInf ROM increasingly more costly to simulate. The comparison plots in Fig. 6b show that all three approaches achieve similar accuracy in the test data regime.

In Fig. 7a, we compare the efficacy of quadratic ROMs obtained via structure-preserving Lift & Learn and nonlinear Hamiltonian ROMs obtained via HOpInf with spDEIM. We observe that both approaches exhibit a similar trend where the efficacy increases with an increase with the reduced dimension up to $2r = 20$. For $2r > 20$, the efficacy decreases with an increase in the reduced dimension. The FOM energy error plots in Fig. 7b compare energy error performance of the proposed approach against intrusive lifting and HOpInf with spDEIM. All three approaches demonstrate similar energy error behavior with the intrusive lifting ROM of dimension $4r = 80$ performing marginally better than the other two nonintrusive approaches. For prediction beyond $t \approx 12.5$, the energy error can become relatively large but remains bounded, reflecting the limited predictive capability of a reduced basis constructed from training data on $t \in [0, 10]$. The FOM energy in this example is approximately 151 in the time interval $[0, 10]$, while the perturbation term ΔE_{lift} discussed in (27) remains several orders of magnitude smaller (below 10^{-11}) over the entire simulation.

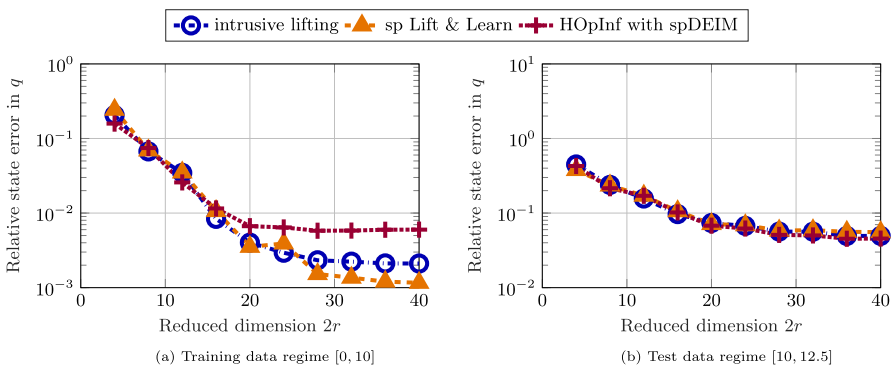


Fig. 6 Two-dimensional sine-Gordon equation. The relative state error comparison in plot a shows that the proposed structure-preserving Lift & Learn approach achieves higher accuracy than the HOpInf with the spDEIM approach in the training regime. Plot b shows that all three approaches achieve similar accuracy in the test data regime

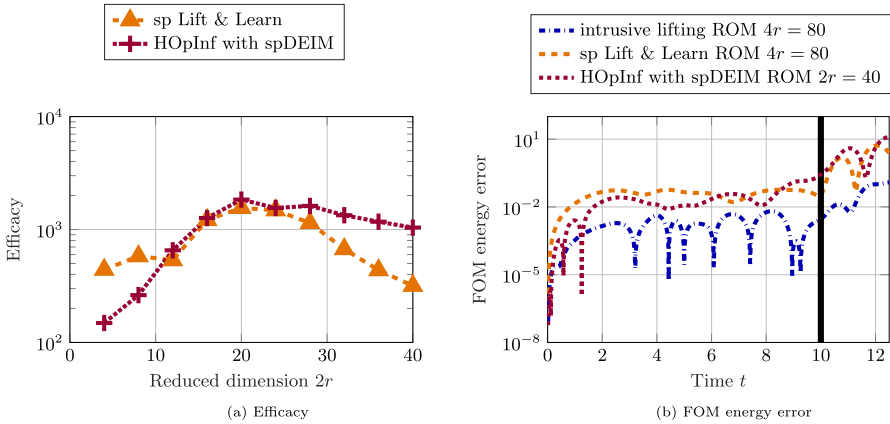


Fig. 7 Two-dimensional sine-Gordon wave equation. Plot **a** shows that the quadratic ROMs obtained via the nonintrusive lifting approach achieve higher efficacy than the HOplnf ROMs with spDEIM. The energy error comparison in plot **b** shows that all approaches yield bounded energy error behavior in the training data regime with the intrusive lifting ROM achieving the lowest FOM energy error. The solid black line in plot **b** indicates the end of the training time interval

4.3 Two-dimensional Klein-Gordon-Zakharov equations

The Klein-Gordon-Zakharov (KGZ) equations [43] are a system of nonlinear dispersive PDEs used to describe the mutual interaction between the Langmuir waves and ion acoustic waves in a plasma. These coupled equations play a crucial role in the study of the dynamics of strong Langmuir turbulence in plasma physics [44]. Unlike the previous two numerical examples, the nonlinear conservative FOM for the KGZ equations is not of the form (14) as it does not have a canonical Hamiltonian formulation.

4.3.1 System of coupled PDEs and the corresponding nonlinear conservative FOM

We consider the FOM setup from [45]. Let $\Omega = (-20, 20) \times (-20, 20) \subset \mathbb{R}^2$ be the spatial domain and consider the coupled nonlinear equations

$$\begin{aligned} \frac{\partial^2 \psi}{\partial t^2}(x, y, t) &= \frac{\partial^2 \psi}{\partial x^2}(x, y, t) + \frac{\partial^2 \psi}{\partial y^2}(x, y, t) - \psi(x, y, t) - \psi(x, y, t)\phi(x, y, t) - |\psi(x, y, t)|^2 \psi(x, y, t), \\ \frac{\partial^2 \phi}{\partial t^2}(x, y, t) &= \frac{\partial^2 \phi}{\partial x^2}(x, y, t) + \frac{\partial^2 \phi}{\partial y^2}(x, y, t) + \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) |\psi(x, y, t)|^2, \end{aligned}$$

where $t \in (0, T]$ is time, $\psi(x, y, t)$ is a complex-valued scalar field that describes the fast time scale component of the electric field raised by electrons, and $\phi(x, y, t)$ is a real-valued scalar field that describes the deviation of the ion density from its equilibrium. The boundary conditions are periodic and the initial conditions are

$$\begin{aligned} \psi(x, y, 0) &= \operatorname{sech}(-(x - 2)^2 - y^2) + \operatorname{sech}(-x^2 - (y - 2)^2), & \frac{\partial \psi}{\partial t}(x, y, 0) &= 0, \\ \phi(x, y, 0) &= \operatorname{sech}(-(x - 2)^2 - y^2) + \operatorname{sech}(-x^2 - (y - 2)^2), & \frac{\partial \phi}{\partial t}(x, y, 0) &= 0. \end{aligned}$$

To rewrite the KGZ equations in first-order form, we first write the complex-valued function ψ in terms of its real and imaginary parts as $\psi = q_1 + iq_2$ and then define $p_1 = \partial q_1 / \partial t$ and $p_2 = \partial q_2 / \partial t$. We then discretize the two-dimensional spatial domain Ω with $n_x = n_y = 400$ equally spaced grid points in both spatial directions to derive the nonlinear FOM of dimension $6n = 6n_x n_y = 960,000$ as

$$\begin{aligned} \dot{\mathbf{q}}_1 &= \mathbf{p}_1, \\ \dot{\mathbf{q}}_2 &= \mathbf{p}_2, \\ \dot{\mathbf{p}}_1 &= \mathbf{D}\mathbf{q}_1 - \mathbf{q}_1 - \boldsymbol{\phi} \odot \mathbf{q}_1 - (\mathbf{q}_1^2 + \mathbf{q}_2^2) \odot \mathbf{q}_1, \\ \dot{\mathbf{p}}_2 &= \mathbf{D}\mathbf{q}_2 - \mathbf{q}_2 - \boldsymbol{\phi} \odot \mathbf{q}_2 - (\mathbf{q}_1^2 + \mathbf{q}_2^2) \odot \mathbf{q}_2, \\ \dot{\boldsymbol{\phi}} &= \boldsymbol{\phi} + (\mathbf{q}_1^2 + \mathbf{q}_2^2), \\ \dot{\boldsymbol{\phi}} &= \mathbf{D}\boldsymbol{\phi}, \end{aligned}$$

with the space-discretized energy

$$E(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2, \boldsymbol{\phi}, \boldsymbol{\phi}) = \frac{1}{2} \mathbf{y}^\top \mathbf{M} \mathbf{y} + \boldsymbol{\phi}^\top (\mathbf{q}_1^2 + \mathbf{q}_2^2) + \sum_{i=1}^n \frac{1}{2} (q_{1,i}^2 + q_{2,i}^2)^2,$$

where $\mathbf{y} = [\mathbf{q}_1^\top, \mathbf{q}_2^\top, \mathbf{p}_1^\top, \mathbf{p}_2^\top, \boldsymbol{\phi}^\top, \boldsymbol{\phi}^\top]^\top \in \mathbb{R}^{6n}$ and $\mathbf{M} = \text{blkdiag}(\mathbf{I}_n - \mathbf{D}, \mathbf{I}_n - \mathbf{D}, \mathbf{I}_n, \mathbf{I}_n, -\mathbf{D}, \mathbf{I}_n) \in \mathbb{R}^{6n \times 6n}$.

4.3.2 Model form for structure-preserving Lift & Learn based on energy quadratization

We follow the energy-quadratization strategy and introduce an auxiliary variable $w = q_1^2 + q_2^2$ to lift the nonlinear FOM to a quadratic lifted FOM

$$\begin{aligned} \dot{\mathbf{q}}_1 &= \mathbf{p}_1, \\ \dot{\mathbf{q}}_2 &= \mathbf{p}_2, \\ \dot{\mathbf{p}}_1 &= \mathbf{D}\mathbf{q}_1 - \mathbf{q}_1 - \boldsymbol{\phi} \odot \mathbf{q}_1 - \mathbf{w} \odot \mathbf{q}_1, \\ \dot{\mathbf{p}}_2 &= \mathbf{D}\mathbf{q}_2 - \mathbf{q}_2 - \boldsymbol{\phi} \odot \mathbf{q}_2 - \mathbf{w} \odot \mathbf{q}_2, \\ \dot{\boldsymbol{\phi}} &= \boldsymbol{\phi} + \mathbf{w}, \\ \dot{\boldsymbol{\phi}} &= \mathbf{D}\boldsymbol{\phi}, \\ \dot{\mathbf{w}} &= 2\mathbf{q}_1 \odot \mathbf{p}_1 + 2\mathbf{q}_2 \odot \mathbf{p}_2, \end{aligned}$$

with a quadratic FOM energy in the lifted state variables

$$E_{\text{lift}}(\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1, \mathbf{p}_2, \boldsymbol{\phi}, \boldsymbol{\phi}, \mathbf{w}) = \frac{1}{2} \mathbf{y}^\top \mathbf{M} \mathbf{y} + \boldsymbol{\phi}^\top \mathbf{w} + \frac{1}{2} \mathbf{w}^\top \mathbf{w}.$$

To approximate the lifted FOM state, we consider a block-diagonal projection matrix $\tilde{\mathbf{V}}$ of the form

$$\tilde{\mathbf{V}} = \text{blkdiag}(\Phi, \Phi, \Phi, \Phi, \mathbf{V}, \mathbf{V}, \mathbf{V}) \in \mathbb{R}^{7n \times 7r},$$

where $\Phi \in \mathbb{R}^{n \times r}$ is computed from the FOM snapshots of $\mathbf{q}_1, \mathbf{q}_2, \mathbf{p}_1$, and \mathbf{p}_2 , $\mathbf{V} \in \mathbb{R}^{n \times r}$ is computed from the FOM snapshots of φ, ϕ , and the lifted \mathbf{w} snapshots. The model form for the nonintrusive quadratic ROM is

$$\begin{aligned} \hat{\mathbf{q}}_1 &= \hat{\mathbf{p}}_1, \\ \hat{\mathbf{q}}_2 &= \hat{\mathbf{p}}_2, \\ \hat{\mathbf{p}}_1 &= \hat{\mathbf{D}}_{\mathbf{q},1} \hat{\mathbf{q}}_1 - \hat{\mathbf{q}}_1 + \hat{\mathbf{H}}_{\mathbf{p}}(\hat{\phi} \otimes \hat{\mathbf{q}}_1) + \hat{\mathbf{H}}_{\mathbf{p}}(\hat{\mathbf{w}} \otimes \hat{\mathbf{q}}_1), \\ \hat{\mathbf{p}}_2 &= \hat{\mathbf{D}}_{\mathbf{q},2} \hat{\mathbf{q}}_2 - \hat{\mathbf{q}}_2 + \hat{\mathbf{H}}_{\mathbf{p}}(\hat{\phi} \otimes \hat{\mathbf{q}}_2) + \hat{\mathbf{H}}_{\mathbf{p}}(\hat{\mathbf{w}} \otimes \hat{\mathbf{q}}_2), \\ \hat{\phi} &= \hat{\phi} + \hat{\mathbf{w}}, \\ \hat{\phi} &= \hat{\mathbf{D}}_{\varphi} \hat{\varphi}, \\ \hat{\mathbf{w}} &= \hat{\mathbf{H}}_{\mathbf{w}}(\hat{\mathbf{q}}_1 \otimes \hat{\mathbf{p}}_1 + \hat{\mathbf{q}}_2 \otimes \hat{\mathbf{p}}_2), \end{aligned}$$

where the quadratic reduced operators $\hat{\mathbf{H}}_{\mathbf{p}} \in \mathbb{R}^{r \times r^2}$ and $\hat{\mathbf{H}}_{\mathbf{w}} \in \mathbb{R}^{r \times r^2}$ are computed analytically, and the linear reduced matrices $\hat{\mathbf{D}}_{\mathbf{q},1} = \hat{\mathbf{D}}_{\mathbf{q},1}^T \in \mathbb{R}^{r \times r}$, $\hat{\mathbf{D}}_{\mathbf{q},2} = \hat{\mathbf{D}}_{\mathbf{q},2}^T \in \mathbb{R}^{r \times r}$, and $\hat{\mathbf{D}}_{\varphi} = \hat{\mathbf{D}}_{\varphi}^T \in \mathbb{R}^{r \times r}$ are learned from projections of the high-dimensional data. We point out once more that all quadratic forms introduced by the lifting procedure do not have to be learned, but are known.

4.3.3 Numerical results

We build a training dataset consisting of FOM snapshots from $t = 0$ to $t = 4$. For the test dataset, we consider FOM snapshots from $t = 4$ to $t = 5$ (predicting 25% outside

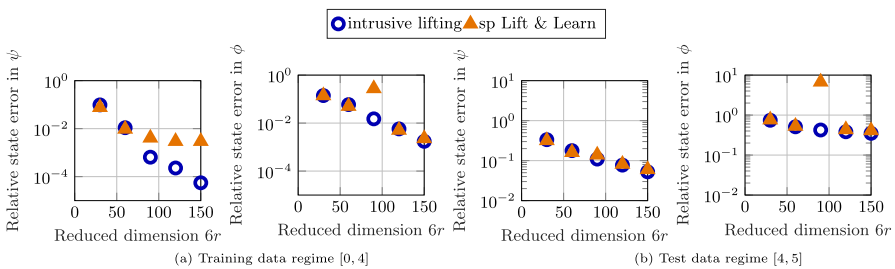


Fig. 8 Klein-Gordon-Zakharov equations. The relative state error comparison for ψ in plot **a** shows that both approaches achieve similar accuracy up to $6r = 60$. For $6r > 60$, the relative state error for structure-preserving Lift & Learn does not decrease as favorably with an increase in the reduced dimension. The relative state error comparison for ψ in plot **b** shows that both approaches achieve similar state error performance in the test data regime. The relative state error comparison for ϕ in plots **a** and **b** show that, except $6r = 90$ for structure-preserving Lift & Learn, intrusive and nonintrusive approaches achieve similar accuracy in both training and test data regimes

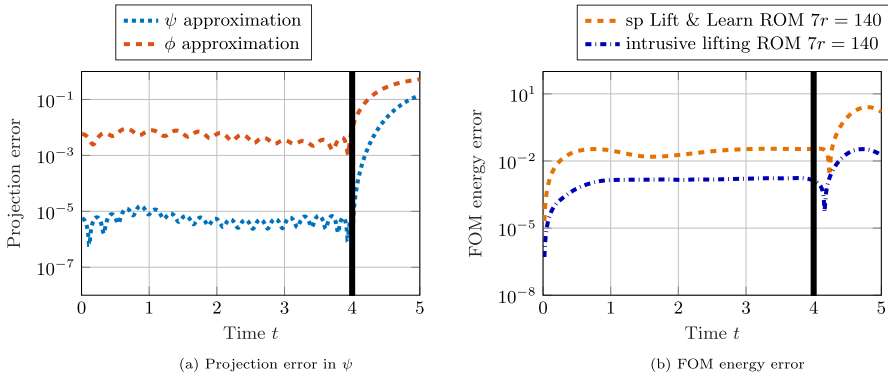


Fig. 9 Klein-Gordon-Zakharov equations. The time evolution of projection error for ψ and ϕ in plot **a** shows that the basis matrix provides accurate approximations in the training data regime. However, the projection error for both field variables grows substantially outside the training data regime. Despite the projection error growth in the test data regime, the energy error comparison in plot **b** shows that both structure-preserving Lift & Learn and intrusive lifting approaches yield ROMs with bounded energy error

training data) to study the generalization capability of the learned ROMs. Due to the non-canonical nature of KGZ equations, the HOPInf with the spDEIM approach is not applicable to this example. Therefore, we only consider structure-preserving Lift & Learn ROMs and structure-preserving intrusive lifting ROMs for this numerical example.

In Fig. 8, we compare the relative state error of the structure-preserving Lift & Learn approach against the intrusive lifting approach for the complex-valued scalar field $\psi(x, y, t)$ and the real-valued scalar field $\phi(x, y, t)$. For training data, the relative state error comparison for ψ in Fig. 8b shows that the quadratic ROMs obtained via structure-preserving Lift & Learn and intrusive lifting achieve similar accuracy up to $6r = 60$. For $6r > 60$, we observe that the state error for ψ levels off in the training data regime for the structure-preserving Lift & Learn ROMs. In Fig. 8d, both nonintrusive and intrusive ROMs exhibit similar accuracy in the test data regime for ψ . The relative state error comparison for ϕ in Fig. 8b shows that structure-preserving Lift & Learn ROMs and intrusive lifting ROMs achieve similar accuracy in the training data regime for all reduced dimensions except $6r = 90$. For $6r = 90$, we observe that the learned quadratic ROM encounters conditioning issues during numerical time integration.³ In Fig. 8d, we observe that both approaches fail to achieve relative state error below 10^{-1} for ϕ in the test data regime.

In Fig. 9a, we show the time evolution of the projection error in ψ and ϕ , which illustrates that the basis matrix provides accurate approximations with relative error below 10^{-2} for both field variables in the training data regime. However, the basis matrix yields substantially higher projection error in the test data regime with relative

³ We investigated adding a Tikhonov-type regularization term to the learning problem for $6r = 90$ as is commonly done in unconstrained Operator Inference [46]. However, this did not lead to improved accuracy or stability in our numerical experiments. Since the proposed structure-preserving Lift & Learn formulation enforces hard constraints derived from the underlying physics, it already provides physics-based regularization, and additional regularization may therefore be ineffective.

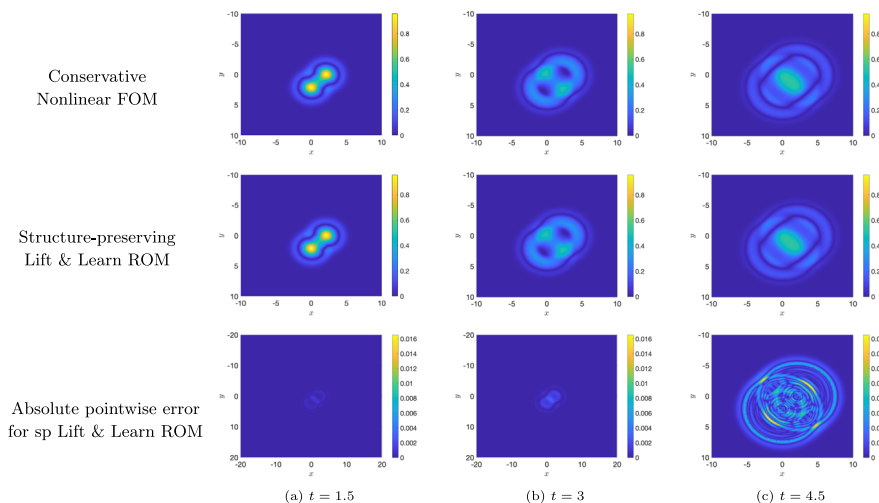


Fig. 10 Klein-Gordon-Zakharov equations. Plots compare the absolute value of the complex-valued field $\psi(x, y, t)$ at selected time instances $t \in \{1.5, 3, 4.5\}$ for the nonlinear conservative FOM (top row) and the structure-preserving quadratic ROM of dimension $7r = 140$ (middle row). The bottom row shows the absolute pointwise error between the structure-preserving Lift & Learn ROM and the conservative FOM. The data-driven quadratic ROM learned via structure-preserving Lift & Learn provides accurate approximate solutions for $|\psi(x, y, t)|$ at all three time instances

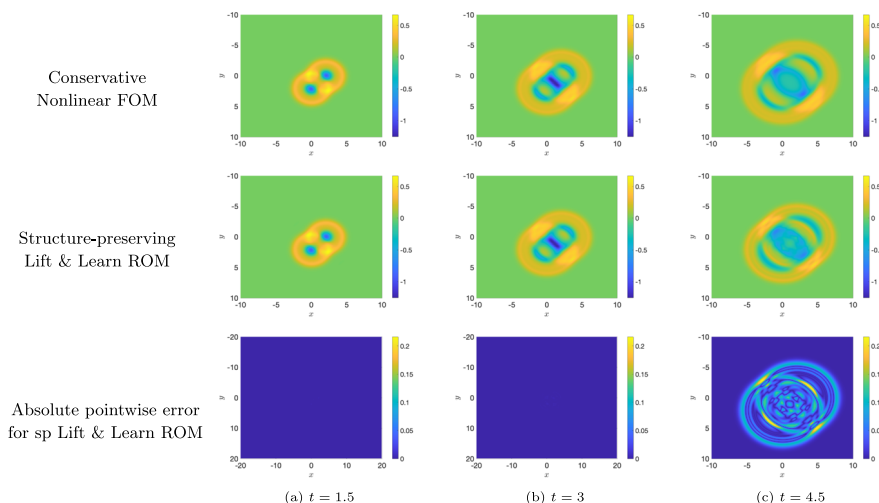


Fig. 11 Klein-Gordon-Zakharov equations. Plots compare the evolution of the scalar field $\phi(x, y, t)$ at selected time instances $t \in \{1.5, 3, 4.5\}$ for the nonlinear conservative FOM (top row) and the structure-preserving quadratic ROM of dimension $7r = 140$ (middle row). The bottom row shows the absolute pointwise error between the structure-preserving Lift & Learn ROM and the conservative FOM. The proposed structure-preserving Lift & Learn method provides accurate approximation of $\phi(x, y, t)$ solution over the two-dimensional computational domain

error higher than 10^{-1} for both ψ and ϕ at $t = 5$. This growth in the projection error is due to the transport-dominated nature of the problem. Notwithstanding the substantial projection error growth outside the training data regime, the energy error comparison in Fig. 9b shows that both structure-preserving Lift & Learn and intrusive lifting yield quadratic ROMs with bounded FOM energy error in the training data regime.

We compare the FOM solution and the structure-preserving Lift & Learn ROM solution, and also include the corresponding absolute pointwise error, for $\psi(x, y, t)$ and $\phi(x, y, t)$ in Figs. 10 and 11, respectively. Figure 10 shows that the structure-preserving Lift & Learn ROM of dimension $7r = 140$ accurately captures the time evolution of the complex-valued scalar field $\psi(x, y, t)$ and provides accurate prediction at $t = 4.5$, which is 12.5% outside the training time interval. Figure 11 shows that the structure-preserving Lift & Learn ROM of dimension $7r = 140$ provides an accurate approximation of the scalar solution field $\phi(x, y, t)$ over the two-dimensional computational domain at $t = 1.5$, $t = 3$, and $t = 4.5$. In terms of computational efficiency, numerical time integration of the structure-preserving quadratic ROM of size $7r = 140$ requires approximately 1.48 s compared to the approximate FOM run time of 89 min, which is a factor of $3608\times$ speedup. Thus, the quadratic ROM learned via the proposed method provides accurate and stable predictions at a substantially lower computational cost.

5 Conclusions

We have presented *structure-preserving Lift & Learn*, a nonintrusive model reduction method which uses lifting transformations to learn low-dimensional quadratic ROMs that respect the conservative nature of the high-dimensional problem. In contrast to the standard Lift & Learn approach [32] that was designed to match the algebraic structure of the lifted model (here: quadratic), the proposed structure-preserving Lift & Learn approach enforces additional conservative properties in the learning process. It does this by leveraging prior knowledge of the conservative PDE through the energy-quadratrization strategy, which enables the derivation of quadratic reduced terms analytically. The remaining linear reduced operators are then learned via a constrained optimization problem. We presented a theoretical result that shows the learned quadratic ROMs conserve a perturbed lifted FOM energy exactly.

Numerical experiments on two canonical Hamiltonian PDEs demonstrate the ability of the proposed approach to learn structure-preserving ROMs that achieve accuracy and computational efficiency similar to nonintrusive Hamiltonian Operator Inference [27] ROMs with spDEIM [17]. The numerical results also show that the proposed structure-preserving Lift & Learn approach yields generalizable quadratic ROMs that provide accurate predictions with bounded energy error outside the training data regime. The final numerical example with the Klein-Gordon-Zakharov equations shows that the proposed approach can be flexibly adapted to learn structure-preserving ROMs for a wider class of coupled conservative PDEs that do not have a canonical Hamiltonian formulation.

Future research directions motivated by this work are as follows: automating the process of deriving quadratic reduced terms in structure-preserving Lift & Learn via

the optimal quadratization framework [47]; studying the long-term boundedness of nonintrusive quadratic ROMs using the theoretical framework from [48]; deriving error bounds for the difference between nonlinear FOM energy approximation obtained via intrusive and nonintrusive quadratic ROMs; and combining the proposed method with the probabilistic learning framework in [49] to use it for learning structure-preserving ROMs from noisy data.

Acknowledgements We thank Iman Adibnazari for his help with running numerical experiments for the two-dimensional Klein-Gordon-Zakharov equations on the Triton Shared Computing Cluster (TSCC) allocation supported by ONR DURIP Award No. N000142412222.

Author contribution Harsh Sharma: writing—original draft, visualization, validation, software, methodology, formal analysis, conceptualization. Juan Diego Draxl Giannoni: writing—review and editing, software, methodology. Boris Kramer: writing—review and editing, resources, methodology, funding acquisition, formal analysis, conceptualization.

Funding H. Sharma and B. Kramer were in part financially supported by the Applied and Computational Analysis Program of the Office of Naval Research under award N000142212624 and the National Science Foundation under award NSF-CMMI 2144023.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Klein, R., Sanderse, B.: Energy-conserving hyper-reduction and temporal localization for reduced order models of the incompressible Navier-Stokes equations. *J. Comput. Phys.* **499**, 112697 (2024)
2. Tyranowski, T.M., Kraus, M.: Symplectic model reduction methods for the Vlasov equation. *Contrib. Plasma Phys.* **63**(5–6), e202200046 (2023)
3. Hesthaven, J., Pagliantini, C., Ripamonti, N.: Adaptive symplectic model order reduction of parametric particle-based Vlasov-Poisson equation. *Math. Comput.* **93**(347), 1153–1202 (2024)
4. Carlberg, K., Tuminaro, R., Boggs, P.: Preserving Lagrangian structure in nonlinear model reduction with application to structural dynamics. *SIAM J. Sci. Comput.* **37**(2), B153–B184 (2015)
5. Sharma, H., Najera-Flores, D.A., Todd, M.D., Kramer, B.: Lagrangian operator inference enhanced with structure-preserving machine learning for nonintrusive model reduction of mechanical systems. *Comput. Methods Appl. Mech. Eng.* **423**, 116865 (2024)
6. Adibnazari, I., Sharma, H., Torralba, J.C., Kramer, B., Tolley, M.T.: Full-body optimal control of a swimming soft robot enabled by data-driven model reduction. In: Southern California Robotics Symposium (2023)
7. Lepri, M., Bacciu, D., Della, S.C.: Neural autoencoder-based structure-preserving model order reduction and control design for high-dimensional physical systems. *IEEE Contr Syst Lett.* **8**, 133–138 (2023)
8. Sharma, H., Adibnazari, I., Cervera-Torralba, J., Tolley, M.T., Kramer, B.: Data-driven model reduction for soft robots via Lagrangian Operator Inference. *IFAC-PapersOnLine.* **58**(17), 91–96 (2024)
9. Peng, L., Mohseni, K.: Symplectic model reduction of Hamiltonian systems. *SIAM J. Sci. Comput.* **38**(1), A1–A27 (2016)
10. Afkham, B.M., Hesthaven, J.S.: Structure preserving model reduction of parametric Hamiltonian systems. *SIAM J. Sci. Comput.* **39**(6), A2616–A2644 (2017)
11. Gong, Y., Wang, Q., Wang, Z.: Structure-preserving Galerkin POD reduced-order modeling of Hamiltonian systems. *Comput. Methods Appl. Mech. Eng.* **315**, 780–798 (2017)
12. Hesthaven, J.S., Pagliantini, C., Ripamonti, N.: Structure-preserving model order reduction of Hamiltonian systems. In: International Congress of Mathematics, pp. 5072–5097 (2022)

13. Buchfink, P., Glas, S., Haasdonk, B.: Symplectic model reduction of Hamiltonian systems on nonlinear manifolds and approximation with weakly symplectic autoencoder. *SIAM J. Sci. Comput.* **45**(2), A289–A311 (2023)
14. Sharma, H., Mu, H., Buchfink, P., Geelen, R., Glas, S., Kramer, B.: Symplectic model reduction of Hamiltonian systems using data-driven quadratic manifolds. *Comput. Methods Appl. Mech. Eng.* **417**, 116402 (2023)
15. Gruber, A., Tezaur, I.: Variationally consistent Hamiltonian model reduction. *SIAM J. Appl. Dyn. Syst.* **24**(1), 376–414 (2025)
16. Wang Z.: Structure-preserving Galerkin POD-DEIM reduced-order modeling of Hamiltonian systems. [arXiv:2103.00388](https://arxiv.org/abs/2103.00388) (2021)
17. Pagliantini, C., Vismara, F.: Gradient-preserving hyper-reduction of nonlinear dynamical systems via discrete empirical interpolation. *SIAM J. Sci. Comput.* **45**(5), A2725–A2754 (2023)
18. Gu, C.: QLMOR: a projection-based nonlinear model order reduction approach using quadratic-linear representation of nonlinear systems. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **30**(9), 1307–1320 (2011)
19. Benner, P., Breiten, T.: Two-sided projection methods for nonlinear model order reduction. *SIAM J. Sci. Comput.* **37**(2), B239–B260 (2015)
20. Kramer, B., Willcox, K.E.: Nonlinear model order reduction via lifting transformations and proper orthogonal decomposition. *AIAA J.* **57**(6), 2297–2307 (2019)
21. Benner, P., Goyal, P., Gugercin, S.: \mathcal{H}_2 -quasi-optimal model order reduction for quadratic-bilinear control systems. *SIAM J. Matrix Anal. Appl.* **39**(2), 983–1032 (2018)
22. Kramer, B., Willcox, K.: Balanced truncation model reduction for lifted nonlinear systems. In: Beattie, C., Benner, P., Embree, M., Gugercin, S., Lefteriu, S. (eds.) *Realization and Model Reduction of Dynamical Systems: A Festschrift in Honor of the 70th Birthday of Thanos Antoulas*, pp. 157–174. Springer International Publishing, Cham (2022)
23. Sharma, H., Draxl Giannoni, J.D., Kramer, B.: Nonlinear energy-preserving model reduction with lifting transformations that quadratize the energy. *Physica D* **483**, 134954 (2025)
24. Peherstorfer, B., Willcox, K.: Data-driven operator inference for nonintrusive projection-based model reduction. *Comput. Methods Appl. Mech. Eng.* **306**, 196–215 (2016)
25. Kramer, B., Peherstorfer, B., Willcox, K.E.: Learning nonlinear reduced models from data with operator inference. *Annu. Rev. Fluid Mech.* **56**(1), 521–548 (2024)
26. Benner, P., Goyal, P., Kramer, B., Peherstorfer, B., Willcox, K.: Operator inference for non-intrusive model reduction of systems with non-polynomial nonlinear terms. *Comput. Methods Appl. Mech. Eng.* **372**, 113433 (2020)
27. Sharma, H., Wang, Z., Kramer, B.: Hamiltonian operator inference: physics-preserving learning of reduced-order models for canonical Hamiltonian systems. *Physica D* **431**, (2022)
28. Gruber, A., Tezaur, I.: Canonical and noncanonical Hamiltonian operator inference. *Comput. Methods Appl. Mech. Eng.* **416**, 116334 (2023)
29. Sharma, H., Kramer, B.: Preserving Lagrangian structure in data-driven reduced-order modeling of large-scale dynamical systems. *Physica D* **462**, 134128 (2024)
30. Filanova, Y., Duff, I.P., Goyal, P., Benner, P.: An operator inference oriented approach for linear mechanical systems. *Mech. Syst. Signal Process.* **200**, 110620 (2023)
31. Vijaywargiya, A., McQuarrie, S.A., Gruber, A.: Tensor parametric Hamiltonian operator inference. [arXiv:2502.10888](https://arxiv.org/abs/2502.10888) (2025)
32. Qian, E., Kramer, B., Peherstorfer, B., Willcox, K.: Lift & learn: Physics-informed machine learning for large-scale nonlinear dynamical systems. *Physica D* **406**, 132401 (2020)
33. Sharma, H., Patil, M., Woolsey, C.: A review of structure-preserving numerical methods for engineering applications. *Comput. Methods Appl. Mech. Eng.* **366**, 113067 (2020)
34. Center, S.D.S.: Triton Shared Computing Cluster. University of California, San Diego (2022)
35. Grant, M., Boyd, S.: CVX: Matlab Software for Disciplined Convex Programming, version 2.2, (2020). <http://cvxr.com/cvx>
36. Hairer, E., Lubich, C., Wanner, G.: *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*, vol. 31. Springer Science & Business Media, (2006)
37. Celledoni, E., McLachlan, R.I., Owren, B., Quispel, G.R.W.: Geometric properties of Kahan’s method. *J. Phys. A: Math. Theor.* **46**(2), 025201 (2012)
38. Avrami, M.: Kinetics of phase change. I General theory. *J. Chem. Phys.* **7**(12), 1103–1112 (1939)

39. Johnson, W.A.: Reaction kinetics in process of nucleation and growth. *Trans. Am. Instit. Mining Metallurgic. Eng.* **135**, 416–458 (1939)
40. Kolmogorov, A.N.: On the statistical theory of the crystallization of metals. *Bull. Acad. Sci. URSS (Cl Sci Math Nat)*. **3**, 335 (1937)
41. Li, D., Sun, W.: Linearly implicit and high-order energy-conserving schemes for nonlinear wave equations. *J. Sci. Comput.* **83**, 1–17 (2020)
42. Bo, Y., Wang, Y., Cai, W.: Arbitrary high-order linearly implicit energy-preserving algorithms for Hamiltonian PDEs. *Num. Algo.* 1–28, (2022)
43. Boling, G., Guangwei, Y.: Global smooth solution for the Klein-Gordon-Zakharov equations. *J. Math. Phys.* **36**(8), 4119–4124 (1995)
44. Bergé, L., Bidégaray, B., Colin, T.: A perturbative analysis of the time-envelope approximation in strong Langmuir turbulence. *Physica D* **95**(3–4), 351–379 (1996)
45. Guo, J., Zhuang, Q.: Energy stable schemes for the Klein-Gordon-Zakharov equations. *Comput. Math. Appl.* **147**, 150–163 (2023)
46. McQuarrie, S.A., Huang, C., Willcox, K.E.: Data-driven reduced-order models via regularised operator inference for a single-injector combustion process. *J. R. Soc. N. Z.* **51**(2), 194–211 (2021)
47. Bychkov, A., Issan, O., Pogudin, G., Kramer, B.: Exact and optimal quadratization of nonlinear finite-dimensional nonautonomous dynamical systems. *SIAM J. Appl. Dyn. Syst.* **23**(1), 982–1016 (2024)
48. Schlegel, M., Noack, B.R.: On long-term boundedness of Galerkin models. *J. Fluid Mech.* **765**, 325–352 (2015)
49. Galioto, N., Sharma, H., Kramer, B., Gorodetsky, A.A.: Bayesian identification of nonseparable Hamiltonians with multiplicative noise using deep learning and reduced-order modeling. *Comput. Methods Appl. Mech. Eng.* **430**, 117194 (2024)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.