Structure-Preserving Learning of High-Dimensional Lagrangian and Hamiltonian Systems

Boris Krämer

ICERM Workshop on Computational Learning for Model Reduction
January 6-10, 2025

UC San Diego
JACOBS SCHOOL OF ENGINEERING
Mechanical and Aerospace Engineering

Collaborators

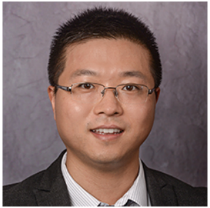
Harsh Sharma
(UC San Diego)

Iman Adibnazari
(UC San Diego)

David Najera
(UC San Diego)

Michael Todd
(UC San Diego)

Michael Tolley
(UC San Diego)



Zhu Wang
(U. South Carolina)

Thanks to the following sponsors of the presented research

- DoD Newton Award: *Geometric Structure-Preserving Model Reduction for Large-Scale Interconnected Systems*
- NSF: *CAREER: Goal-oriented Variable Transformations for Efficient Reduced-order and Data-driven Modeling*
- Korean Institute for the Advancement of Technology: *Prediction, damage analysis and risk assessment for a gas power plant via fast and accurate reduced models*
- Office of Naval Research: *Nonlinear Data-driven and Structure-Preserving Hamiltonian Model Reduction*

Motivating thoughts

- With the ever-increasing data volumes we can access, there is a growing demand to **learn computationally efficient surrogate models** of **high-dimensional dynamical systems** for optimization, uncertainty quantification, and long-term prediction

- Unconstrained learning for nonlinear finite-dimensional systems (reduced or full-order)

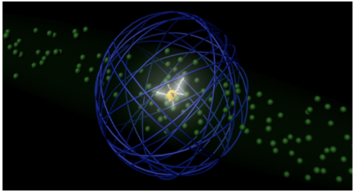
$$\dot{\mathbf{x}}(t,\mu) = \mathbf{f}(\mathbf{x}(t),\mu) + \mathbf{g}(\mathbf{x})\mathbf{u}(t) \quad \in \mathbb{R}^n$$

  a la machine learning may be very expressive, but often fails to extrapolate in time.

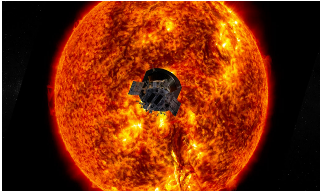
- In general, if n is very large (e.g., $n \geq 1,000$) then almost all methods require dimension reduction (linear or nonlinear)

- **Incorporating model knowledge** into learning framework is imperative for **predictions**: geometric/mechanical structure, nonlinear terms, inputs, etc.

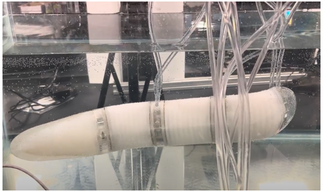
Structured physical systems are everywhere!



(a) Bose-Einstein condensate



(b) Solar plasma



(c) Soft-robotic fish

- Physical systems have interesting properties like conservation laws, symplecticity, reversibility or configuration space structure
- Structure-preserving methods preserve underlying geometric structure
  - Conserve discrete quantities which are close to continuous quantity
  - Reproduce long-time behavior
- Long-time numerical simulation of large-scale systems using structure-preserving methods is computationally prohibitive

Need for physics-preserving reduced-order models

Part I: Learning Hamiltonian reduced models via structure-preserving optimization

- Hamiltonian operator inference: physics-preserving learning of reduced-order models for Hamiltonian systems. *Sharma/Wang/K.*, Physica D: Nonlinear Phenomena, Vol. 431, 133122, 2022.
- Gradient preserving operator inference: data-driven reduced-order models for equations with gradient structure. *Geng/Singh/Ju/K./Wang*, Computer Methods in Applied Mechanics and Engineering, Vol. 427, 117033, 2024.

Gradient systems

- General infinite-dimensional Hamiltonian system

$$\frac{\partial y(x,t)}{\partial t} = \mathcal{L}\frac{\delta \mathcal{H}}{\delta y}$$

where $\mathcal{L}$ is a linear differential operator, $\frac{\delta \mathcal{H}}{\delta y}$ is the variational derivative of

$$\mathcal{H}[y] = \int \Big(\underbrace{H_{\mathsf{quad}}(y, y_x, \dots)}_{\text{quadratic terms}} + \underbrace{H_{\mathsf{nl}}(y)}_{\text{spatially local terms}} \Big) \, \mathrm{d}x$$

1. If $\mathcal{L}$ **is skew adjoint**, $\mathcal{H}$ is referred to as the Hamiltonian, is constant, and the PDE is **conservative**.
2. If $\mathcal{L}$ **is negative semi-definite** (resp. definite), $\mathcal{H}$, referred to as a Lyapunov/energy function, is nonincreasing (resp. monotonically decreasing) and the PDE is **dissipative**.

- Structure-preserving space discretization leads to finite-dimensional Hamiltonian models

$$\dot{\mathbf{y}} = \mathbf{D}\nabla_{\mathbf{y}} H_{\mathrm{d}}(\mathbf{y})$$

where H_{d} is the space-discretized Hamiltonian function.

Goal: Learn low-dimensional Hamiltonian systems from trajectory data $\mathbf{y}(t_1), \dots, \mathbf{y}(t_f)$.

Special forms of gradient systems

(i) When $\mathbf{D}$ is skew-symmetric (a.k.a. skew-adjoint), $\mathbf{D} = -\mathbf{D}^\intercal$, the system is Hamiltonian. A special case is when $\mathbf{D} = \left[\begin{smallmatrix} 0 & \mathbf{I} \\ -\mathbf{I} & 0 \end{smallmatrix}\right]$ the system is a canonical Hamiltonian system and the solution flow is symplectic. Thus, the internal energy of the system, $H(\mathbf{y})$, is conserved, e..g, for any $t_1, t_2 \in I$ with $t_1 < t_2$, we have

$$H(\mathbf{y}(t_2)) - H(\mathbf{y}(t_1)) = \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t} H(\mathbf{y}(t))\,\mathrm{d}t = \int_{t_1}^{t_2} (\nabla_{\mathbf{y}} H(\mathbf{y}))^\intercal\, \mathbf{D}\, \nabla_{\mathbf{y}} H(\mathbf{y})\,\mathrm{d}t = 0.$$

(ii) When $\mathbf{D}$ is negative semi-definite, the system represents a gradient flow and is dissipative:

$$H(\mathbf{y}(t_2)) - H(\mathbf{y}(t_1)) = \int_{t_1}^{t_2} \frac{\mathrm{d}}{\mathrm{d}t} H(\mathbf{y}(t))\,\mathrm{d}t = \int_{t_1}^{t_2} (\nabla_{\mathbf{y}} H(\mathbf{y}))^\intercal\, \mathbf{D}\, \nabla_{\mathbf{y}} H(\mathbf{y})\,\mathrm{d}t \le 0.$$

Note that if $\mathbf{D}$ is negative definite, $H(\mathbf{y})$ is strictly decreasing.

Numerical schemes to solve these equations recognize the special gradient structure in time discretization, e.g., geometric integrators and average vector field methods.

This talk: Canonical Hamiltonian Systems

- We focus on canonical Hamiltonian systems (for other gradient systems, see[1])

$$\dot{\mathbf{y}} = \begin{bmatrix} \dot{\mathbf{q}} \\ \dot{\mathbf{p}} \end{bmatrix} = \mathbf{J}_{2n} \nabla_{\mathbf{y}} H_{\mathrm{d}}(\mathbf{q}, \mathbf{p}) = \begin{bmatrix} \nabla_{\mathbf{p}} H_{\mathrm{d}}(\mathbf{q}, \mathbf{p}) \\ -\nabla_{\mathbf{q}} H_{\mathrm{d}}(\mathbf{q}, \mathbf{p}) \end{bmatrix}$$

- Key features
  1. **Canonical Hamiltonian structure**, i.e. $\mathbf{J}_{2n} = \begin{bmatrix} \mathbf{0} & \mathbf{I}_n \\ -\mathbf{I}_n & \mathbf{0} \end{bmatrix}$
  2. State vector $\mathbf{y} \in \mathbb{R}^{2n}$ can be partitioned as $\mathbf{y} = [\mathbf{q}^\top, \mathbf{p}^\top]^\top$ where $\mathbf{q}, \mathbf{p} \in \mathbb{R}^n$ both have distinct **physical interpretation**
  3. **Symmetry** in linear FOM operators due to structure-preserving space discretization

Assumption: Functional form of $H_{\mathsf{nl}}(q, p)$ is known, whereas $H_{\mathsf{quad}}(q, p)$ and details about spatial discretization are unavailable.

[1] Gradient preserving operator inference: data-driven reduced-order models for equations with gradient structure. *Geng/Singh/Ju/K./Wang*, CMAME 427, 117033, 2024.

Problem formulation: FOM data

Given: Snapshot data matrices from Hamiltonian FOM simulations

$$\mathbf{Q} = \left[\mathbf{q}_1 \cdots \mathbf{q}_K\right] \in \mathbb{R}^{n \times K}, \qquad \mathbf{P} = \left[\mathbf{p}_1 \cdots \mathbf{p}_K\right] \in \mathbb{R}^{n \times K}$$

- Use knowledge about Hamiltonian functional to define nonlinear forcing snapshot matrices

$$\mathbf{F}_{\mathbf{q}} = \left[\mathbf{f}_{\mathbf{q}}(\mathbf{y}_1) \cdots \mathbf{f}_{\mathbf{q}}(\mathbf{y}_K)\right] \in \mathbb{R}^{n \times K}, \qquad \mathbf{F}_{\mathbf{p}} = \left[\mathbf{f}_{\mathbf{p}}(\mathbf{y}_1) \cdots \mathbf{f}_{\mathbf{p}}(\mathbf{y}_K)\right] \in \mathbb{R}^{n \times K}$$

- Build snapshot matrices of time-derivative data via finite difference

$$\dot{\mathbf{Q}} = \left[\dot{\mathbf{q}}_1 \cdots \dot{\mathbf{q}}_K\right] \in \mathbb{R}^{n \times K}, \qquad \dot{\mathbf{P}} = \left[\dot{\mathbf{p}}_1 \cdots \dot{\mathbf{p}}_K\right] \in \mathbb{R}^{n \times K}$$

Next step: Project FOM data onto low-dimensional symplectic subspaces

Problem formulation: symplectic projection

- Symplectic projection via proper symplectic decomposition (PSD)

$$\begin{bmatrix} \mathbf{q} \\ \mathbf{p} \end{bmatrix} \approx \begin{bmatrix} \mathbf{\Phi} & \mathbf{0} \\ \mathbf{0} & \mathbf{\Phi} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{q}} \\ \hat{\mathbf{p}} \end{bmatrix}$$

  where $\mathbf{V}_{\mathbf{q}} = \mathbf{V}_{\mathbf{p}} = \mathbf{\Phi} \in \mathbb{R}^{n \times r}$ are obtained via cotangent the lift algorithm.[2]
- Projecting FOM data to obtain
  1. Reduced snapshot data
$$\hat{\mathbf{Q}} = \mathbf{V}_{\mathbf{q}}^{\top} \mathbf{Q} \in \mathbb{R}^{r \times K}, \qquad \hat{\mathbf{P}} = \mathbf{V}_{\mathbf{p}}^{\top} \mathbf{P} \in \mathbb{R}^{r \times K}$$
  2. Reduced time-derivative data
$$\dot{\hat{\mathbf{Q}}} = \mathbf{V}_{\mathbf{q}}^{\top} \dot{\mathbf{Q}} \in \mathbb{R}^{r \times K}, \qquad \dot{\hat{\mathbf{P}}} = \mathbf{V}_{\mathbf{p}}^{\top} \dot{\mathbf{P}} \in \mathbb{R}^{r \times K}$$
  3. Reduced nonlinear forcing data
$$\hat{\mathbf{F}}_{\mathbf{q}} = \mathbf{V}_{\mathbf{p}}^{\top} \mathbf{F}_{\mathbf{q}} \in \mathbb{R}^{r \times K}, \qquad \hat{\mathbf{F}}_{\mathbf{p}} = \mathbf{V}_{\mathbf{q}}^{\top} \mathbf{F}_{\mathbf{p}} \in \mathbb{R}^{r \times K}$$

Next step: Fit reduced operators to the projected trajectories in a structure-preserving way

[2]Peng L, Mohseni K. Symplectic model reduction of Hamiltonian systems. SIAM Journal on Scientific Computing. 2016;38(1):A1–A27

Problem formulation: model form for learning

- Reduced Hamiltonian in terms of the inferred reduced operators $\hat{\mathbf{D}}_{\mathbf{q}} \in \mathbb{R}^{r \times r}$ and $\hat{\mathbf{D}}_{\mathbf{p}} \in \mathbb{R}^{r \times r}$

$$\hat{H}(\hat{\mathbf{q}}, \hat{\mathbf{p}}) = \frac{1}{2}\hat{\mathbf{q}}^\top \hat{\mathbf{D}}_{\mathbf{q}} \hat{\mathbf{q}} + \frac{1}{2}\hat{\mathbf{p}}^\top \hat{\mathbf{D}}_{\mathbf{p}} \hat{\mathbf{p}} + \hat{H}_{\mathsf{nl}}(\hat{\mathbf{q}}, \hat{\mathbf{p}})$$

- Model form for learning Hamiltonian ROMs based on $\hat{H}(\hat{\mathbf{q}}, \hat{\mathbf{p}})$

$$\dot{\hat{\mathbf{q}}} = \frac{\partial \hat{H}}{\partial \hat{\mathbf{p}}} = \hat{\mathbf{D}}_{\mathbf{p}} \hat{\mathbf{p}} + \mathbf{V}_{\mathbf{p}}^\top \mathbf{f}_{\mathbf{q}}(\mathbf{V}_{\mathbf{q}} \hat{\mathbf{q}}, \mathbf{V}_{\mathbf{p}} \hat{\mathbf{p}})$$

$$\dot{\hat{\mathbf{p}}} = -\frac{\partial \hat{H}}{\partial \hat{\mathbf{q}}} = -\hat{\mathbf{D}}_{\mathbf{q}} \hat{\mathbf{q}} - \mathbf{V}_{\mathbf{q}}^\top \mathbf{f}_{\mathbf{p}}(\mathbf{V}_{\mathbf{q}} \hat{\mathbf{q}}, \mathbf{V}_{\mathbf{p}} \hat{\mathbf{p}})$$

Hamiltonian Operator Inference

- Constrained optimization problem[3] to compute $\hat{\mathbf{D}}_{\mathbf{q}}$ and $\hat{\mathbf{D}}_{\mathbf{p}}$

$$\min_{\substack{\hat{\mathbf{D}}_{\mathbf{q}}=\hat{\mathbf{D}}_{\mathbf{q}}^{\top},\\ \hat{\mathbf{D}}_{\mathbf{p}}=\hat{\mathbf{D}}_{\mathbf{p}}^{\top}}} \left\| \begin{bmatrix} \dot{\hat{\mathbf{Q}}} - \hat{\mathbf{F}}_{\mathbf{q}}(\hat{\mathbf{Q}}, \hat{\mathbf{P}}) \\ \dot{\hat{\mathbf{P}}} + \hat{\mathbf{F}}_{\mathbf{p}}(\hat{\mathbf{Q}}, \hat{\mathbf{P}}) \end{bmatrix} - \begin{bmatrix} \mathbf{0} & \hat{\mathbf{D}}_{\mathbf{p}} \\ -\hat{\mathbf{D}}_{\mathbf{q}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{Q}} \\ \hat{\mathbf{P}} \end{bmatrix} \right\|_F$$

  where symmetric constraints on $\hat{\mathbf{D}}_{\mathbf{q}}$ and $\hat{\mathbf{D}}_{\mathbf{p}}$ ensure that the learned reduced operators retain symmetric property of full-model operators
- Separate, symmetric linear least-squares problems of the form

$$\min_{\hat{\mathbf{D}}_{\mathbf{p}}=\hat{\mathbf{D}}_{\mathbf{p}}^{\top}} \left\| \underbrace{\dot{\hat{\mathbf{Q}}} - \hat{\mathbf{F}}_{\mathbf{q}}(\hat{\mathbf{Q}}, \hat{\mathbf{P}})}_{\mathbf{R}_{\mathbf{p}}} - \hat{\mathbf{D}}_{\mathbf{p}}\hat{\mathbf{P}} \right\|_F \rightarrow (\hat{\mathbf{P}}\hat{\mathbf{P}}^{\top})\hat{\mathbf{D}}_{\mathbf{p}} + \hat{\mathbf{D}}_{\mathbf{p}}(\hat{\mathbf{P}}\hat{\mathbf{P}}^{\top}) = \hat{\mathbf{P}}\hat{\mathbf{R}}_{\mathbf{p}}^{\top} + \hat{\mathbf{R}}_{\mathbf{p}}\hat{\mathbf{P}}^{\top}$$

$$\min_{\hat{\mathbf{D}}_{\mathbf{q}}=\hat{\mathbf{D}}_{\mathbf{q}}^{\top}} \left\| \underbrace{\dot{\hat{\mathbf{P}}} + \hat{\mathbf{F}}_{\mathbf{p}}(\hat{\mathbf{Q}}, \hat{\mathbf{P}})}_{-\mathbf{R}_{\mathbf{q}}} + \hat{\mathbf{D}}_{\mathbf{q}}\hat{\mathbf{Q}} \right\|_F \rightarrow (\hat{\mathbf{Q}}\hat{\mathbf{Q}}^{\top})\hat{\mathbf{D}}_{\mathbf{q}} + \hat{\mathbf{D}}_{\mathbf{q}}(\hat{\mathbf{Q}}\hat{\mathbf{Q}}^{\top}) = \hat{\mathbf{Q}}\hat{\mathbf{R}}_{\mathbf{q}}^{\top} + \hat{\mathbf{R}}_{\mathbf{q}}\hat{\mathbf{Q}}^{\top}$$

  can be solved via Lyapunov equations.

[3] Based on Operator Inference: [Data-driven operator inference for nonintrusive projection-based model reduction, Peherstorfer & Willcox, CMAME, 306, 196-215 351 2016]

Error bound on the learned operators

1. Time stepping scheme for the FOM is convergent

$$\max_{i\in\{1,\cdots,T/\Delta t\}} \left|\left|\mathbf{y}_i - \mathbf{y}(t_i)\right|\right|_2 \to 0 \quad as \quad \Delta t \to 0$$

2. Derivatives approximated from projected states converge to $\frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{y}}(t_k)$

$$\max_{i\in\{1,\cdots,T/\Delta t\}} \left|\left|\dot{\hat{\mathbf{y}}}_i - \frac{\mathrm{d}}{\mathrm{d}t}\hat{\mathbf{y}}(t_i)\right|\right|_2 \to 0 \quad as \quad \Delta t \to 0$$

Theorem 1 ([Sharma/Wang/K, 2022][4])

Let $\tilde{\mathbf{D}}_{\mathbf{q}}$ and $\tilde{\mathbf{D}}_{\mathbf{p}}$ be the intrusively projected ROM operators. If the snapshot data matrix has full column rank, then for every $\epsilon > 0$, there exists $2r \leq 2n$ and $\Delta t > 0$ such that for the difference between the learned operators $\hat{\mathbf{D}}_{\mathbf{q}}, \hat{\mathbf{D}}_{\mathbf{p}}$ and the projection-based $\tilde{\mathbf{D}}_{\mathbf{q}}, \tilde{\mathbf{D}}_{\mathbf{p}}$, we have

$$||\hat{\mathbf{D}}_{\mathbf{q}} - \tilde{\mathbf{D}}_{\mathbf{q}}||_F \leq \epsilon, \qquad ||\hat{\mathbf{D}}_{\mathbf{p}} - \tilde{\mathbf{D}}_{\mathbf{p}}||_F \leq \epsilon.$$

[4]Sharma H, Wang Z, Kramer B. Hamiltonian operator inference: Physics-preserving learning of reduced-order models for canonical Hamiltonian systems. Physica D: Nonlinear Phenomena, 431, 133122, 2022.

Error bound on the solutions

Theorem 1 ([Geng/Singh/Ju/K./Wang, 2024][5])

Let $\mathbf{y}(t)$ be the solution of the FOM on $[0,T]$ and $\mathbf{y}_r(t)$ be the solution of the structure-preserving ROM on the same interval. Suppose $\nabla_{\mathbf{y}} H(\mathbf{y})$ is Lipschitz continuous, then the ROM approximation error satisfies

$$\int_0^T \|\mathbf{y} - \mathbf{\Phi}\mathbf{y}_r\|^2 \,\mathrm{d}t \leq$$

$$C(T)\left(\underbrace{\int_0^T \|\mathbf{y} - \mathbf{\Phi}\mathbf{\Phi}^\intercal\mathbf{y}\|^2 \mathrm{d}t}_{\textit{projection error}} + \underbrace{\int_0^T \|\dot{\mathbf{y}} - \mathcal{D}_t[\mathbf{y}]\|^2 \,\mathrm{d}t}_{\textit{data error}} + \underbrace{\int_0^T \|\mathbf{\Phi}^\intercal\mathcal{D}_t[\mathbf{y}] - \mathbf{D}_r\mathbf{\Phi}^\intercal\nabla_{\mathbf{y}} H(\mathbf{y})\|^2 \,\mathrm{d}t}_{\textit{optimization error}}\right),$$

where $C(T) = \max\{1 + C_2^2, 2\}T\alpha(T)$, $\alpha(T) = 2\int_0^T e^{2C_1(T-\tau)}\,\mathrm{d}\tau$, and the constants $C_1 = \mathcal{C}_{\mathrm{log-Lip}}[\mathbf{\Phi}\mathbf{D}_r\mathbf{\Phi}^\intercal\nabla_{\mathbf{y}} H]$ and $C_2 = \|\mathbf{\Phi}\mathbf{D}_r\mathbf{\Phi}^\intercal\|\,\mathcal{C}_{\mathrm{Lip}}[\nabla_{\mathbf{y}} H]$.

[5] Gradient preserving operator inference: data-driven reduced-order models for equations with gradient structure. *Geng/Singh/Ju/K./Wang*, CMAME 427, 117033, 2024.

Nonlinear Schrödinger equation

- Cubic Schrödinger equation

$$i\psi_t + \psi_{xx} + \gamma|\psi|^2\psi = 0$$

- Writing $\psi = p + iq$ with space-time continuous Hamiltonian $\mathcal{H}$

$$\mathcal{H}(q,p) = \int \frac{1}{2}\left[p_x^2 + q_x^2 - \frac{\gamma}{2}[q^2 + p^2]^2\right]\,dx$$

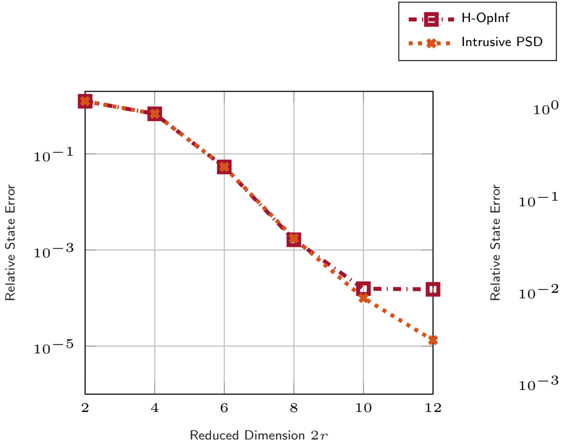
leads to canonical Hamiltonian PDE form $y_t = J\frac{\partial \mathcal{H}}{\partial y}$ for $y = [p,q]^\top$

$$y_t = \begin{bmatrix} p_t \\ q_t \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\delta \mathcal{H}}{\delta p} \\ \frac{\delta \mathcal{H}}{\delta q} \end{bmatrix} = \begin{bmatrix} -q_{xx} - \gamma(q^2 + p^2)q \\ p_{xx} + \gamma(q^2 + p^2)p \end{bmatrix}$$

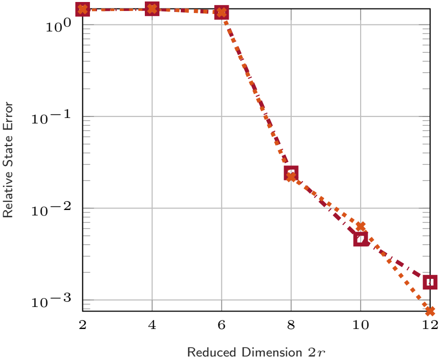
- Mass and momentum invariants of motion

$$\mathcal{M}_1(q,p) = \int \left[p^2 + q^2\right]\,dx \qquad \mathcal{M}_2(q,p) = \int \left[p_x q - q_x p\right]\,dx$$

Nonlinear Schrödinger equation ($2n = 128$): state error

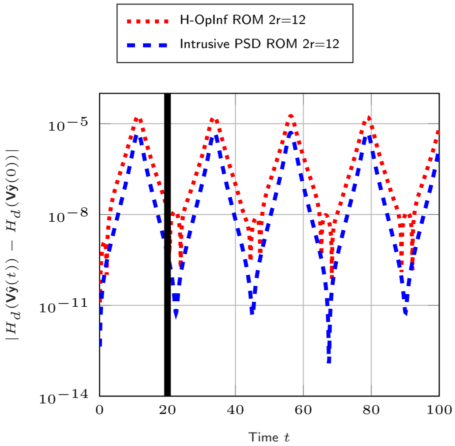


(a) Training data

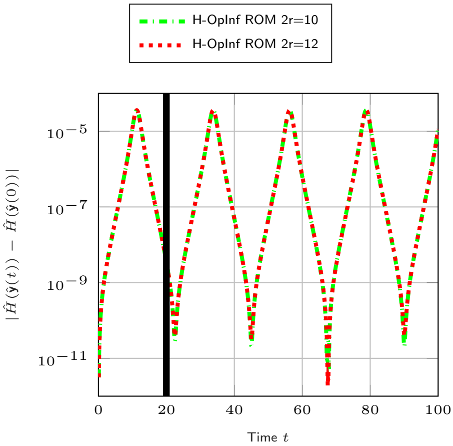


(b) Test data (400% outside)

Nonlinear Schrödinger equation: energy error

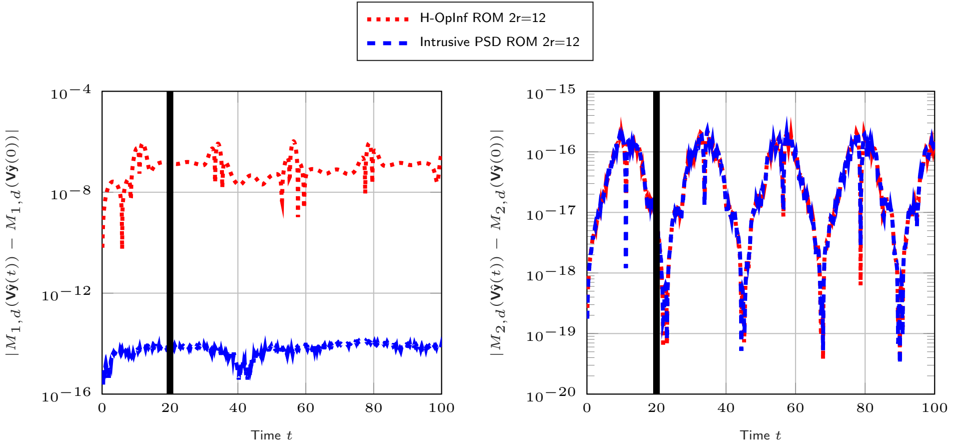


(a) FOM energy error



(b) ROM energy error

Nonlinear Schrödinger equation: invariants of motion



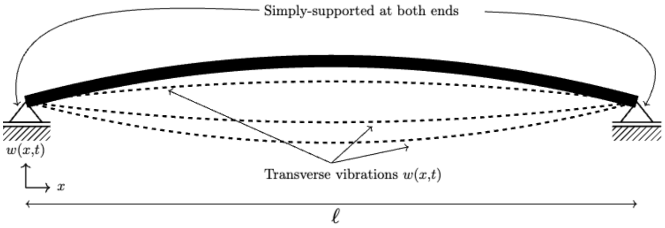
(a) Mass conservation **(b)** Momentum conservation

H-OpInf conserves mass and momentum invariants of motion

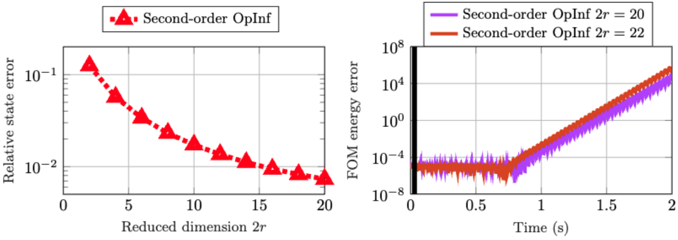
Part II: Learning Lagrangian reduced models from high-dimensional data

- *Preserving Lagrangian structure in data-driven reduced-order modeling of large-scale mechanical systems*, Sharma, H. & Kramer, B., Physica D: Nonlinear Phenomena, Vol 462, 134128, 2024.

Why do we need to preserve the Lagrangian structure?



Euler-Bernoulli beam with transverse vibrations in response to a nonzero initial condition.



(a) State error (training data)

(b) FOM energy error

- Unconstrained second-order model learning provides monotonically decaying state error in training
- BUT: predictive capabilties limited due to energy growth and finite-time blowup.

Lagrangian mechanics: simple mechanical systems

Consider a Lagrangian system with a finite-dimensional configuration manifold Q, state space TQ and a Lagrangian $L : \mathsf{TQ} \to \mathbb{R}$. The **forced Euler-Lagrange equations** define the dynamics:

$$\frac{\partial L(\mathbf{q}, \dot{\mathbf{q}})}{\partial \mathbf{q}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L(\mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{\mathbf{q}}}\right) + \mathbf{f}(\mathbf{q}, \dot{\mathbf{q}}, t) = \mathbf{0}.$$

For simple mechanical systems with configuration manifold $\mathsf{Q} = \mathbb{R}^n$:

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \underbrace{T(\dot{\mathbf{q}})}_{\text{kinetic energy}} - \underbrace{U(\mathbf{q})}_{\text{potential energy}} = \frac{1}{2}\dot{\mathbf{q}}^\top \mathbf{M}\dot{\mathbf{q}} - \frac{1}{2}\mathbf{q}^\top \mathbf{K}\mathbf{q}, \qquad \mathbf{M} = \mathbf{M}^\top \succ 0, \quad \mathbf{K} = \mathbf{K}^\top.$$

The force is often modeled via a dissipative force and an external time-dependent input as

$$\mathbf{f}(\mathbf{q}, \dot{\mathbf{q}}, t) = -\mathbf{C}\dot{\mathbf{q}} + \mathbf{B}\mathbf{u}(t),$$

with $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times n}$ the damping and input matrix, $\mathbf{u}(t) \in \mathbb{R}^m$ the time-dependent inputs. The resulting linear equations of motion are

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}\dot{\mathbf{q}}(t) + \mathbf{K}\mathbf{q}(t) = \mathbf{B}\mathbf{u}(t).$$

Lagrangian mechanics: nonlinear wave equations

Consider for illustration the class of 1d nonlinear wave equations

$$\frac{\partial^2 q}{\partial t^2} - \frac{\partial^2 q}{\partial x^2} + \frac{\mathsf{d}U_{\mathsf{nl}}(q)}{\mathsf{d}q} = 0, \tag{1}$$

where $U_{\mathsf{nl}}(q)$ is the nonlinear component of the potential energy. Discretization of the space-time continuous Lagrangian (w. symmetric FDs or pseudo-spectral methods) at n equally spaced points:

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\sum_{i=1}^{n}\left(\left(\frac{\partial q_i}{\partial t}\right)^2 - \left(\sum_{k=1}^{n} D_{ik} q_k\right)^2\right) - \sum_{i=1}^{n} U_{\mathsf{nl}}(q_i), \qquad \mathbf{q} = [q_1, q_2, \cdots, q_n]^\top$$

where $q_i := q(t, x_i)$, and $\frac{\partial q}{\partial x}(x_i) \approx \sum_{k=1}^{n} D_{ik} q_k$. The Euler-Lagrange equations are

$$\ddot{\mathbf{q}} = \mathbf{K}\mathbf{q} + \frac{\mathsf{d}U_{\mathsf{nl}}(\mathbf{q})}{\mathsf{d}\mathbf{q}}, \quad \mathbf{K} = \mathbf{K}^\top. \tag{2}$$

The nonlinear FOM described by (2) conserves the total energy

$$E(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2}\dot{\mathbf{q}}^\top\dot{\mathbf{q}} + \frac{1}{2}\mathbf{q}^\top\mathbf{K}\mathbf{q} + \sum_{i=1}^{n} U_{\mathsf{nl}}(q_i). \tag{3}$$

High-dimensional Lagrangian systems and (data-driven) model reduction

1. Intrusive structure-preserving model reduction for Lagrangian systems ([Lall et al., 2004], [Carlberg et al., 2015])
   Drawback: Requires access to FOM operators
2. Structure-preserving neural networks ([Cranmer et al., 2019], [Lutter et al., 2019], [Gupta et al., 2020])
   Drawback: Ill-suited for high-dimensional systems
3. Nonintrusive model reduction via operator inference (OpInf)
   - Operator inference for nonlinear systems ([Peherstorfer and Willcox, 2016], [Benner et al., 2020])
   - Lift & Learn ([Qian et al., 2020], [Swischuk et al., 2020])

   Drawback: Does not preserve the Lagrangian structure
4. Operator Inference for linear mechanical systems [Filanova et al., MSSP 200 (2023): 110620]. $\Rightarrow$ Similar to our approach, independently derived.

Our contribution: Embed Lagrangian structure into a Operator Inference learning framework

The data we need for learning

Given: Solutions from Lagrangian FOM simulation with inputs and outputs stored in matrices

$$\mathbf{Q} = [\mathbf{q}_1, \cdots, \mathbf{q}_K] \in \mathbb{R}^{n \times K}, \quad \mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_K] \in \mathbb{R}^{p \times K}, \quad \mathbf{U} = [\mathbf{u}(t_1), \cdots, \mathbf{u}(t_K)] \in \mathbb{R}^{m \times K}$$

- Compute proper orthogonal decomposition basis via SVD

$$\mathbf{Q} = \mathbf{V}\Xi\mathbf{W}^\top, \qquad \mathbf{V} \in \mathbb{R}^{n \times n}, \Xi \in \mathbb{R}^{n \times n}, \mathbf{W} \in \mathbb{R}^{K \times n}$$

- Project FOM data to obtain reduced snapshot data

$$\hat{\mathbf{Q}} = \mathbf{V}_r^\top \mathbf{Q} = [\hat{\mathbf{q}}_1, \cdots, \hat{\mathbf{q}}_K] \in \mathbb{R}^{r \times K}$$

- Generate or collect reduced time-derivative data

$$\dot{\hat{\mathbf{Q}}} = [\dot{\hat{\mathbf{q}}}_1, \cdots, \dot{\hat{\mathbf{q}}}_K] \in \mathbb{R}^{r \times K}, \qquad \ddot{\hat{\mathbf{Q}}} = [\ddot{\hat{\mathbf{q}}}_1, \cdots, \ddot{\hat{\mathbf{q}}}_K] \in \mathbb{R}^{r \times K}$$

Next step: Fit reduced operators to the projected trajectories in a structure-preserving way

Problem formulation: model form for learning ROM

- Reduced Lagrangian with reduced mass matrix $\hat{\mathbf{M}} = \mathbb{I}_r$

$$\hat{L}_r(\hat{\mathbf{q}}, \dot{\hat{\mathbf{q}}}) = \frac{1}{2}\dot{\hat{\mathbf{q}}}^\top \dot{\hat{\mathbf{q}}} - \frac{1}{2}\hat{\mathbf{q}}^\top \hat{\mathbf{K}}\hat{\mathbf{q}},$$

- Reduced forcing

$$\hat{\mathbf{f}}(\hat{\mathbf{q}}, \dot{\hat{\mathbf{q}}}, t) = \hat{\mathbf{C}}\dot{\hat{\mathbf{q}}} - \hat{\mathbf{B}}\mathbf{u}(t)$$

- Model form for learning Lagrangian ROMs based on $\hat{L}_r(\hat{\mathbf{q}}, \dot{\hat{\mathbf{q}}})$

$$\ddot{\hat{\mathbf{q}}}(t) + \hat{\mathbf{C}}\dot{\hat{\mathbf{q}}}(t) + \hat{\mathbf{K}}\hat{\mathbf{q}}(t) = \hat{\mathbf{B}}\mathbf{u}(t)$$

along with the reduced output equation

$$\mathbf{y}(t) = \hat{\mathbf{E}}\hat{\mathbf{q}}(t)$$

Model form ensures that the reduced models are Lagrangian

Lagrangian Operator Inference for simple mechanical systems

- Constrained optimization problem to compute $\hat{\mathbf{C}} \in \mathbb{R}^{r\times r}, \hat{\mathbf{K}} \in \mathbb{R}^{r\times r}$, and $\hat{\mathbf{B}} \in \mathbb{R}^{r\times m}$

$$\min_{\substack{\hat{\mathbf{K}}=\hat{\mathbf{K}}^\top\succ 0, \hat{\mathbf{C}}=\hat{\mathbf{C}}^\top\succ 0,\\ \hat{\mathbf{B}}}} ||\ddot{\hat{\mathbf{Q}}} + \hat{\mathbf{C}}\dot{\hat{\mathbf{Q}}} + \hat{\mathbf{K}}\hat{\mathbf{Q}} - \hat{\mathbf{B}}\mathbf{U}||_F$$

  where the specific choice of $\hat{\mathbf{M}}$ simplifies the constrained inference problem ([Gosea, Gugercin, and Werner, 2023])
- Separate linear least-squares problem to compute $\hat{\mathbf{E}} \in \mathbb{R}^{p\times r}$

$$\min_{\hat{\mathbf{E}}} ||\mathbf{Y} - \hat{\mathbf{E}}\hat{\mathbf{Q}}||_F$$

- Constrained optimization problem solved using the semidefinite programming mode in CVX[6]

[6] M. Grant and S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, 2014

Lagrangian Operator Inference for nonlinear wave equations

$$\mathcal{L}(x, q, q_x, q_t) = \frac{1}{2}\left(\left(\frac{\partial q}{\partial t}\right)^2 - \left(\frac{\partial q}{\partial x}\right)^2\right) - U_{\mathsf{nl}}(q)$$

- Use knowledge about the nonlinear potential energy U_{nl} at the PDE level to build the nonlinear forcing snapshot data matrix

$$\mathbf{F}_{\mathsf{nl}} = [\mathbf{f}_{\mathsf{nl}}(\mathbf{q}_1), \cdots, \mathbf{f}_{\mathsf{nl}}(\mathbf{q}_K)] \in \mathbb{R}^{n \times K}$$

- Projecting FOM snapshot data $\mathbf{Q}$ and forcing snapshot data $\mathbf{F}_{\mathsf{nl}}$

$$\hat{\mathbf{Q}} = \mathbf{V}_r^\top \mathbf{Q} \in \mathbb{R}^{r \times K}, \qquad \hat{\mathbf{F}}_{\mathsf{nl}} = \mathbf{V}_r^\top \mathbf{F}_{\mathsf{nl}} \in \mathbb{R}^{r \times K}$$

- Constrained optimization problem to compute $\hat{\mathbf{K}} \in \mathbb{R}^{r \times r}$

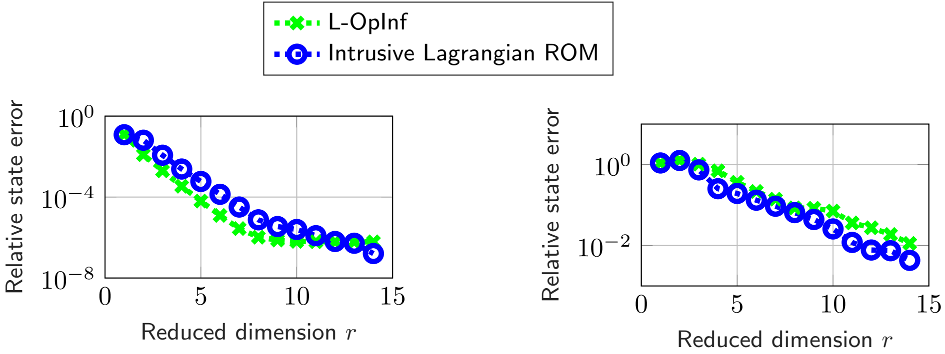
$$\min_{\hat{\mathbf{K}} = \hat{\mathbf{K}}^\top} \|\ddot{\hat{\mathbf{Q}}} - \hat{\mathbf{F}}_{\mathsf{nl}} - \hat{\mathbf{K}}\hat{\mathbf{Q}}\|_F$$

Learned ROM operator $\hat{\mathrm{K}}$ respects the symmetric property introduced during the structure-preserving spatial discretization

Sine-Gordon equation (n=2000): state error

Nonlinear hyperbolic PDE with a nonpolynomial nonlinearity

$$\frac{\partial^2 q}{\partial t^2} = \frac{\partial^2 q}{\partial x^2} - \sin(q)$$

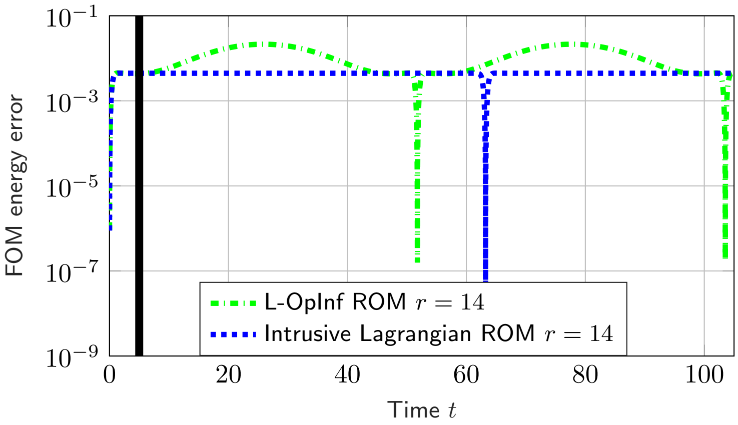


(a) Training regime $[0, 5]$s

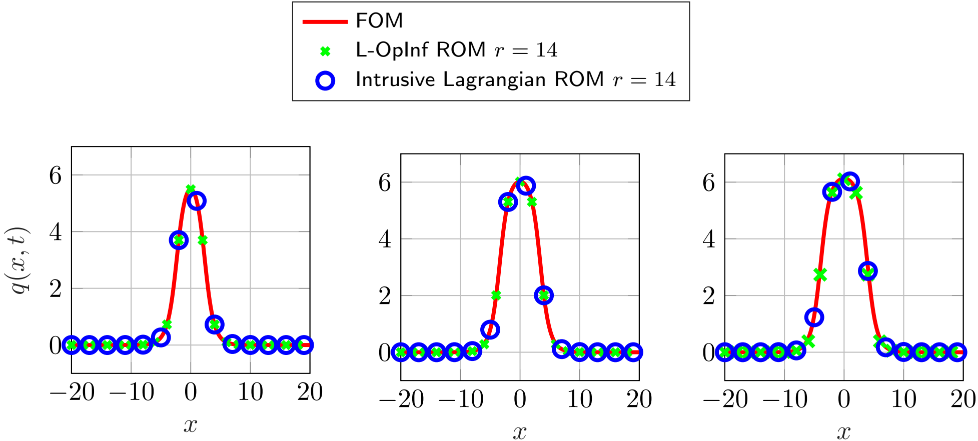
(b) Testing regime $[5, 25]$s

Sine-Gordon equation ($n = 2000$): bounded energy error

Preserving Lagrangian structure yields stable ROMs with bounded energy error far outside the training data regime (here: over 20x past training interval)



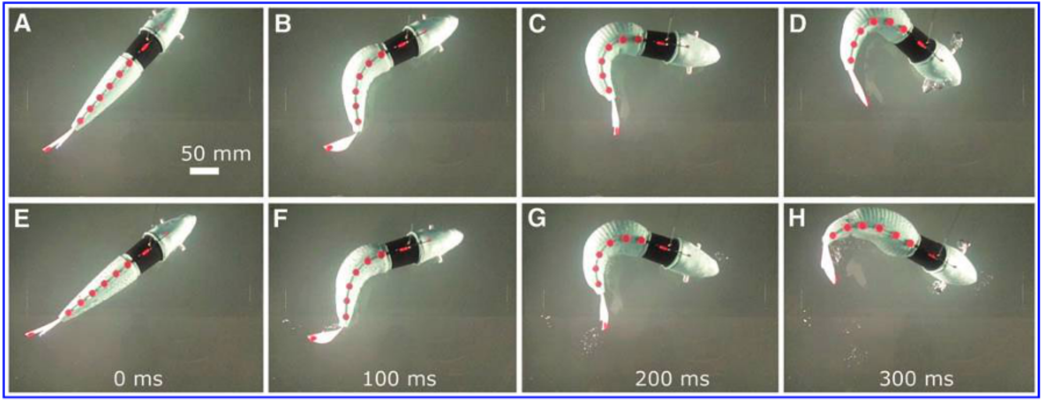
Sine-Gordon equation ($n = 2000$): extrapolation in time



(a) $t = 5$ **(b)** $t = 15$ **(c)** $t = 25$

Accurate predictions 400% outside training time interval

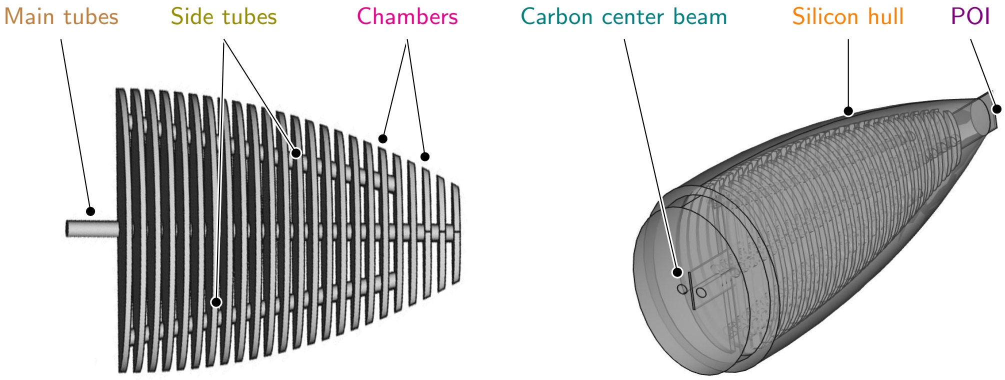
Benchmark soft-robotic fishtail



- Soft robotic fish[7] designed to emulate escape responses in addition to forward swimming because such maneuvers require rapid body accelerations and continuum-body motion
- Fish's soft body is an array of fluidic elastomer actuators

[7] A. D. Marchese, C. D. Onal, and D. Rus, Autonomous soft robotic fish capable of escape maneuvers using fluidic elastomer actuators, Soft Robotics, 1 (2014), pp. 75–87.

Benchmark fishtail CAD model[8]

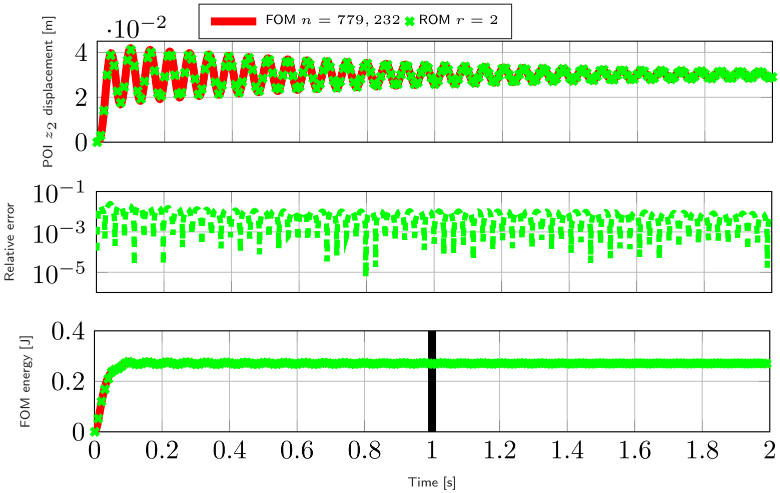


(a) Fluid chamber system **(b)** Complete fishtail model

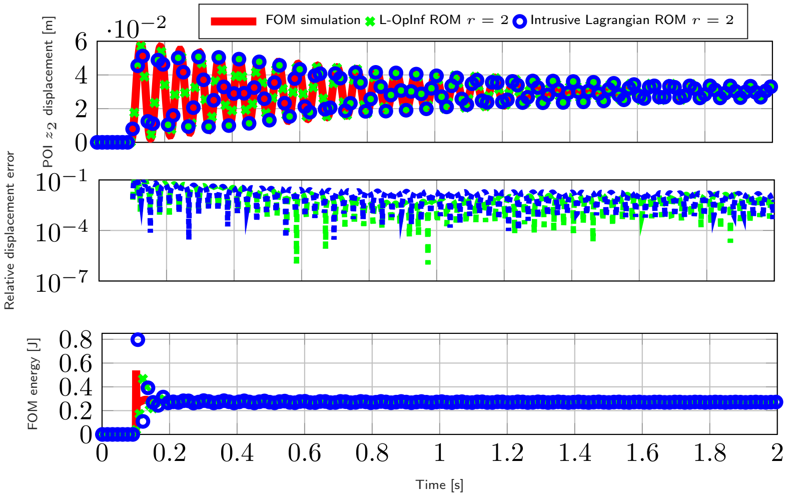
$$\rho \frac{\partial^2 \underline{q}(t, \mathbf{z})}{\partial t^2} = \nabla_{\mathbf{z}} \cdot \underline{\sigma}(t, \mathbf{z})$$

[8] D. Siebelts, A. Kater, and T. Meurer, Modeling and motion planning for an artificial fishtail, IFAC-PapersOnLine, 51 (2018), pp. 319–324. Available at `https://morwiki.mpi-magdeburg.mpg.de/morwiki/index.php/Artificial_Fishtail`

Soft-robotic fishtail ($n = 779,232$): sigmoid input



Soft-robotic fishtail ($n = 779,232$): step input



L-OpInf works well even for unknown control inputs

Full-Body Optimal Control of a Swimming Soft Robot Enabled by Data-Driven Model Reduction

Iman Adibnazari, Harsh Sharma, Jacobo Cervera Torralba, Boris Krämer , Michael T. Tolley
UC San Diego

Collaboration with the Bioinspired Robotics and Design Lab, UCSD (Prof. Mike Tolley)

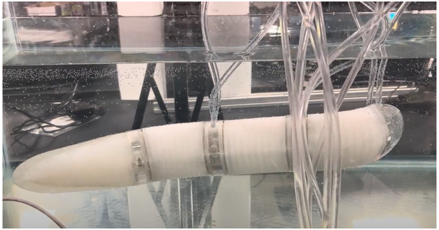
A swimming trout....

Taken from [Beal, D. N. et al (2006)]
What's the catch in the left video?

Video from the Tolley Lab

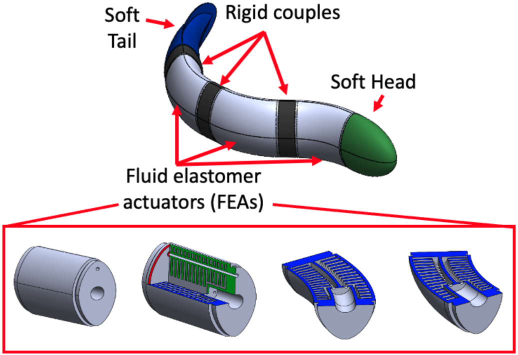
The SERPENT V1 aquatic soft robot

- Autonomous underwater vehicles (AUVs) are mechanically safe and provide silent operation
- Problem: Soft robots are (infinite)-dimensional, so hard to control and simulate.



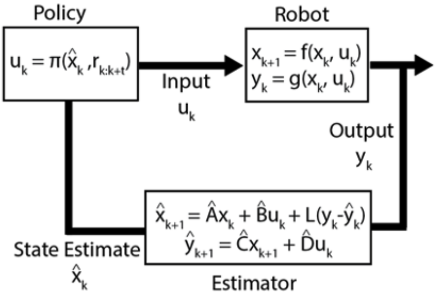
Built "SERPENT V1" from Bioinspired Robotics and Design Lab, UCSD (Prof. Mike Tolley)

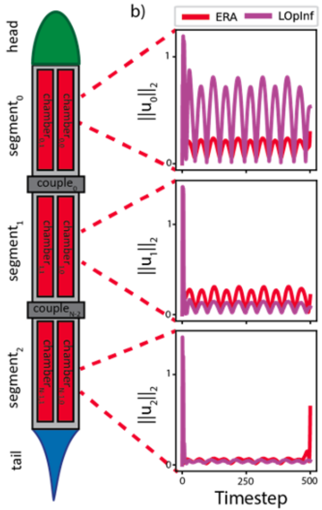
- Segmentation and computational discretization in SOFA ($n = 251,000$)
- $m = 6$ controls: fluid-elastomer actuators
- $p = 40$ outputs: centerline trajectories



Model predictive control w LOPINF for anguilliform swimming

- Control loop requires fast online state estimation
- Discretized model not available ($\rightarrow$ SOFA), need data-driven ROMs
- Lagrangian OPINF due to the second-order nature. Linear for faster online estimation.





Actuation effort in each segment

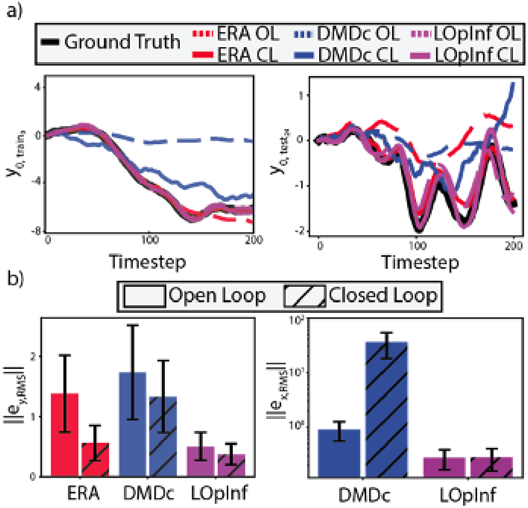
Preliminary results

Compared prevalent linear system identification methods

1. Dynamic Mode Decomposition w/ control (DMDc)
2. Eigensystem Realization Algorithm/Observer/Kalman Identification Algorithm (ERA/OKID)
3. Structure-preserving LOPINF

Current takeaways:

- LOPINF ROM most accurate in RMS error over many trajectories (from random ICs)
- Works well in open and closed-loop
- Certain regimes are too nonlinear ⇒ need to embed nonlinear terms in LOPINF



Nonlinear Lagrangian ROMs

H. Sharma, D. Najera, M. Todd, B. Kramer
UC San Diego

Lagrangian operator inference enhanced with structure-preserving machine learning for nonintrusive model reduction of mechanical systems, *Sharma/Najera/Todd/K.*, CMAME, Vol. 423, 116865, 2024.

Nonlinear Lagrangian mechanical systems

Consider Lagrangian system with a finite-dimensional configuration manifold Q, state space TQ and a Lagrangian $L : \mathsf{TQ} \to \mathbb{R}$. The forced Euler-Lagrange equations define the dynamics:

$$\frac{\partial L(\mathbf{q}, \dot{\mathbf{q}})}{\partial \mathbf{q}} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L(\mathbf{q}, \dot{\mathbf{q}})}{\partial \dot{\mathbf{q}}}\right) + \mathbf{f}(\mathbf{q}, \dot{\mathbf{q}}, t) = \mathbf{0}.$$

For simple mechanical systems with configuration manifold $\mathsf{Q} = \mathbb{R}^n$:

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \underbrace{T(\dot{\mathbf{q}})}_{\text{kinetic energy}} - \underbrace{U(\mathbf{q})}_{\text{potential energy}} = \frac{1}{2}\dot{\mathbf{q}}^\top \mathbf{M}\dot{\mathbf{q}} - \frac{1}{2}\mathbf{q}^\top \mathbf{K}\mathbf{q} - U_{\mathsf{nl}}(\mathbf{q}), \qquad \mathbf{M} = \mathbf{M}^\top \succ 0$$

We model viscous damping

$$\mathbf{f}(\dot{\mathbf{q}}) = -\mathbf{C}\dot{\mathbf{q}} - \frac{\partial \mathcal{F}_{\mathsf{nl}}(\dot{\mathbf{q}})}{\partial \dot{\mathbf{q}}},$$

with $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{n \times n}$ the damping and input matrix, $\mathbf{u}(t) \in \mathbb{R}^m$ the time-dependent inputs. The resulting equations of motion are

$$\mathbf{M}\ddot{\mathbf{q}}(t) + \mathbf{C}\dot{\mathbf{q}}(t) + \frac{\partial \mathcal{F}_{\mathsf{nl}}(\dot{\mathbf{q}})}{\partial \dot{\mathbf{q}}} + \mathbf{K}\mathbf{q}(t) + U_{\mathsf{nl}}(\mathbf{q}) = \mathbf{B}\mathbf{u}(t).$$

LOPINF enhanced with structure-preserving machine learning ('LOPINF-SpML')

We now have a nonlinear constrained optimization problem

$$\min_{\hat{\mathbf{M}}=\hat{\mathbf{M}}^\top\succ\mathbf{0},\hat{\mathbf{C}}=\hat{\mathbf{C}}^\top\succ\mathbf{0},\hat{\mathbf{K}}=\hat{\mathbf{K}}^\top\succ\mathbf{0},\widehat{\mathcal{F}}_{\mathsf{nl}},\widehat{U}_{\mathsf{nl}}} \left\| (\hat{\mathbf{I}}+\hat{\mathbf{M}})\ddot{\hat{\mathbf{Q}}} + \hat{\mathbf{C}}\dot{\hat{\mathbf{Q}}} + \frac{\partial \widehat{\mathcal{F}}_{\mathsf{nl}}(\dot{\hat{\mathbf{Q}}})}{\partial \dot{\hat{\mathbf{Q}}}} + \hat{\mathbf{K}}\hat{\mathbf{Q}} + \frac{\partial \widehat{U}_{\mathsf{nl}}(\hat{\mathbf{Q}})}{\partial \hat{\mathbf{Q}}} = \mathbf{0}, \right\|_F$$

Our approach: a two-step approach

- Step 1: LOPINF to learn the linear reduced stiffness matrix $\hat{\mathbf{K}}$ and the linear reduced damping matrix $\hat{\mathbf{C}}$ ✓

$$\min_{\hat{\mathbf{K}}=\hat{\mathbf{K}}^\top\succ\mathbf{0},\hat{\mathbf{C}}=\hat{\mathbf{C}}^\top\succ\mathbf{0}} \left\| \ddot{\hat{\mathbf{Q}}} + \hat{\mathbf{C}}\dot{\hat{\mathbf{Q}}} + \hat{\mathbf{K}}\hat{\mathbf{Q}} \right\|_F .$$

- Step 2: Structure-preserving machine learning to learn the reduced mass matrix $\hat{\mathbf{M}}$, the nonlinear components of the reduced potential energy function $\widehat{U}_{\mathsf{nl}}(\hat{\mathbf{q}})$, and the nonlinear components of the reduced dissipation function $\widehat{\mathcal{F}}_{\mathsf{nl}}(\dot{\hat{\mathbf{q}}})$.

Structure-preserving neural networks

Parametrization of the nonlinear terms via polynomial-augmented multilayer perceptrons (MLPs); we choose $(P_1 = P_2 = 4)$

$$\widehat{U}_{\mathsf{NN}}(\hat{\mathbf{q}};\boldsymbol{\alpha},\boldsymbol{\lambda},\boldsymbol{\theta}_{\widehat{U}_{\mathsf{NN}}}) = \sum_{i_1,i_2,\cdots,i_r}^{i_1+i_2+\cdots+i_r=P_1} \alpha_{i_1,i_2,\cdots,i_r}\hat{q}_1^{i_1}\hat{q}_2^{i_2}\cdots\hat{q}_r^{i_r} + \sum_{i}^{N}\lambda_i\mathcal{U}^{(i)}(\hat{\mathbf{q}}),$$

$$\widehat{\mathcal{F}}_{\mathsf{NN}}(\dot{\hat{\mathbf{q}}};\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\theta}_{\widehat{\mathcal{F}}_{\mathsf{NN}}}) = \sum_{i_1,i_2,\cdots,i_r}^{i_1+i_2+\cdots+i_r=P_2} \beta_{i_1,i_2,\cdots,i_r}\dot{\hat{q}}_1^{i_1}\dot{\hat{q}}_2^{i_2}\cdots\dot{\hat{q}}_r^{i_r} + \sum_{i}^{N}\gamma_i\mathcal{F}^{(i)}(\dot{\hat{\mathbf{q}}}),$$

We parametrize the reduced kinetic energy term as

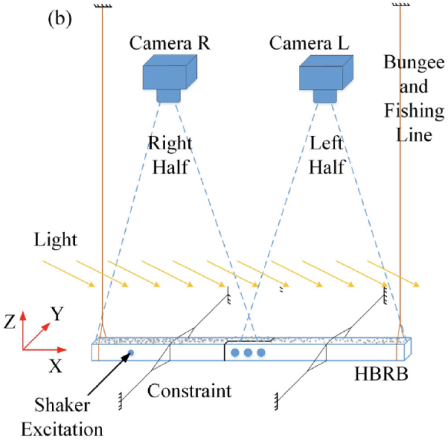
$$\widehat{T}_{\mathsf{NN}}(\dot{\hat{\mathbf{q}}};\boldsymbol{\zeta}) = \sum_{i_1,i_2,\cdots,i_r}^{i_1+i_2+\cdots+i_r=2} \zeta_{i_1,i_2,\cdots,i_r}\dot{\hat{q}}_1^{i_1}\dot{\hat{q}}_2^{i_2}\cdots\dot{\hat{q}}_r^{i_r} \quad\Rightarrow\quad [\hat{\mathbf{M}}_{\mathsf{NN}}]_{ij} = [\hat{\mathbf{M}}_{\mathsf{NN}}]_{ji} = \frac{\partial^2\widehat{T}_{\mathsf{NN}}\left(\dot{\hat{\mathbf{q}}}\right)}{\partial\dot{\hat{q}}_i\partial\dot{\hat{q}}_j}$$

We then minimize the squared loss function under structure-preserving constraints

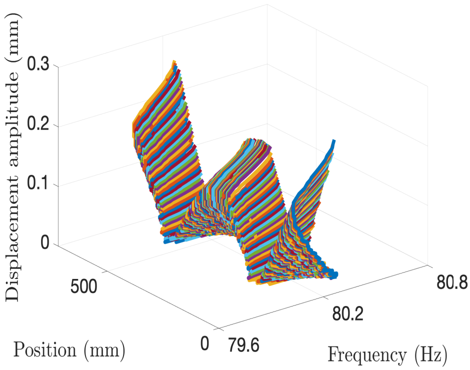
$$\min_{\boldsymbol{\zeta},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\lambda},\boldsymbol{\gamma},\boldsymbol{\theta}_{\widehat{U}_{\mathsf{NN}}},\boldsymbol{\theta}_{\widehat{\mathcal{F}}_{\mathsf{NN}}}} \mathcal{J}(\boldsymbol{\zeta},\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\lambda},\boldsymbol{\gamma},\boldsymbol{\theta}_{\widehat{U}_{\mathsf{NN}}},\boldsymbol{\theta}_{\widehat{\mathcal{F}}_{\mathsf{NN}}}) \quad \text{such that} \quad \frac{1}{2}\dot{\hat{\mathbf{q}}}^\top(\mathbf{I}_r + \hat{\mathbf{M}}_{\mathsf{NN}})\dot{\hat{q}} > 0,\ \widehat{\mathcal{F}}(\dot{\hat{\mathbf{q}}}) \geq 0.$$

Experimental data from the half Brake-Reuss beam

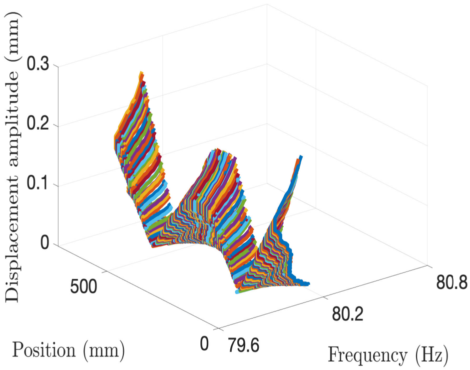
- Modeling of jointed structures remains a challenging problem due to the strong nonlinearities at the frictional interfaces found in joints
- Experimental dataset from [Chen et al., Measurement and identification of the nonlinear dynamics of a jointed structure using full-field data, Part I: Measurement of nonlinear dynamics. MSSP, 2022;166:108401]; high-speed cameras combined with digital image correlation provide the full-field response of the structure
- Repository: `https://github.com/mattiacenedese/BRBtesting`.



Backbone curves: comparison of FOM and LOPINF-SpML



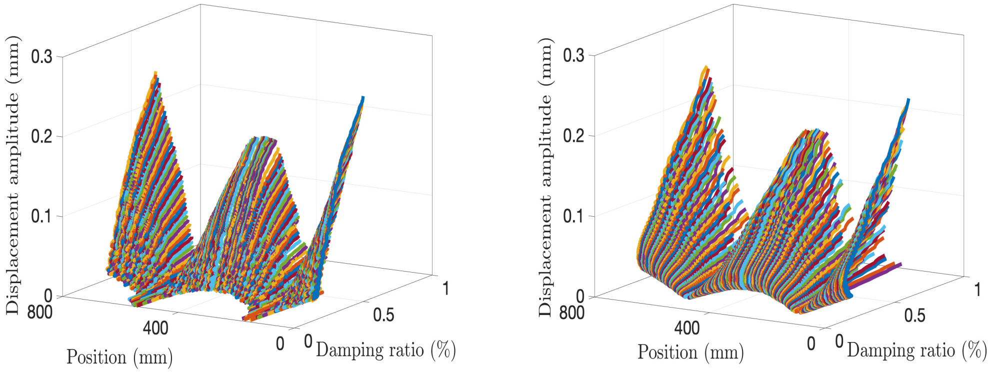
(a) Experimental data



(b) LOpInf-SpML ROM $r = 3$

Figure: Half Brake-Reuß beam. The LOpInf-SpML ROM of size $r = 3$ accurately predicts the amplitude-dependent frequency characteristics and yields backbone curves that appear to agree with the backbone curves obtained directly from the experimental data.

Damping plots: comparison of FOM and LOPINF-SpML



(a) Experimental data

(b) LOpInf-SpML ROM $r = 3$

Figure: Half Brake-Reuß beam. The amplitude-dependent damping plots based on the LOpInf-SpML ROM of dimension $r = 3$ are reasonably similar to the plots obtained from the experimental data.

Today's talk: review and outlook

Review: Embedding geometric/mechanical structure (Hamiltonian, Lagrangian) in model learning creates

- Physically interpretable and analyzable models that engineers are familiar with
- Stable ROMs with bounded energy error
- Accurate long-time predictions far outside the training data regime
- Reduces the need for large training data

Looking ahead:

- Neural networks are expressive, but not as nicely interpretable as symbolic expressions. Good methods for incorporating nonlinearities (polynomial or non-polynomial) in an interpretable manner required (e.g, SINDy, Higher-order OPINF, etc).
- Better methods to extrapolate in parameter space

Papers on Hamiltonian/Gradient systems

1. Hamiltonian operator inference: physics-preserving learning of reduced-order models for Hamiltonian systems. *Sharma/Wang/K.*, Physica D: Nonlinear Phenomena, Vol. 431, 133122, 2022.
2. Gradient preserving operator inference: data-driven reduced-order models for equations with gradient structure. *Geng/Singh/Ju/K./Wang*, CMAME 427, 117033, 2024.
3. Symplectic model reduction of Hamiltonian systems using data-driven quadratic manifolds. *Sharma/Mu/Buchfink/Geelen/Glas/K*. CMAME, 417, 116402, 2023.
4. Bayesian identification of nonseparable Hamiltonians with multiplicative noise using deep learning and reduced-order modeling, *Galioto/Sharma/K./Gorodetsky*, CMAME, 430, 117194, 2024.
5. Bayesian Identification of nonseparable Hamiltonian systems using stochastic dynamic models, *Sharma/Galioto/Gorodetsky/K.* 2022 IEEE 61st Conference on Decision and Control (CDC), 2022, pp. 6742-6749.

Papers on Lagrangian systems

1. Full-Body Optimal Control of a Swimming Soft Robot Enabled by Data-Driven Model Reduction. *Adibnazari/Sharma/Torralba/Kramer/Tolley*, 2023 Southern California Robotics (SCR) Symposium, September 14-15, 2023.
2. Lagrangian operator inference enhanced with structure-preserving machine learning for nonintrusive model reduction of mechanical systems, *Sharma/Najera/Todd/K.*, CMAME, Vol. 423, 116865, 2024.
3. Preserving Lagrangian structure in data-driven reduced-order modeling of large-scale mechanical systems, *Sharma/K.*, Physica D: Nonlinear Phenomena, Vol 462, 134128, 2024.