

The University of Texas at Austin Oden Institute for Computational **Engineering and Sciences** 

# **Projection-based Model Reduction Formulations for Physics-based Machine Learning**

Professor Karen E. Willcox Oden Institute for Computational Engineering and Sciences

Minisymposium on Scientific Machine Learning SIAM CSE19 February 25, 2019



# Contributors

- Dr. Boris Kramer (MIT)
- Dr. Laura Mainini (UTRC)
- Prof. Benjamin Peherstorfer (Courant)
- Elizabeth Qian (MIT)
- Renee Swischuk (MIT)













Physics-based models have structure

But they are large-scale and expensive to solve, and often embodied in black-box solvers Consider physics-based models represented as systems of ODEs or spatial discretization of PDEs describing the system of interest

which in turn arise from governing physical principles (conservation laws, etc.)

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{p})\mathbf{x} + \mathbf{B}(\mathbf{p})\mathbf{u}$$
 
$$\dot{\mathbf{x}} = f(\mathbf{x}, \mathbf{p}, \mathbf{u})$$
 
$$\mathbf{y} = \mathbf{C}(\mathbf{p})\mathbf{x}$$
 
$$\mathbf{y} = g(\mathbf{x}, \mathbf{p}, \mathbf{u})$$

 $\mathbf{x} \in \mathbf{R}^{N}$ : state vector  $\mathbf{u} \in \mathbf{R}^{N_{i}}$ : input vector  $\mathbf{p} \in \mathbf{R}^{N_{p}}$ : parameter vector  $\mathbf{y} \in \mathbf{R}^{N_{o}}$ : output vector

# Projection preserves structure

Reduced models: Low-cost but accurate approximations of high-fidelity models via projection onto a lowdimensional subspace

Interpretable & analyzable

$$\dot{\mathbf{x}} = \mathbf{A}(\mathbf{p})\mathbf{x} + \mathbf{B}(\mathbf{p})\mathbf{u}$$

$$\mathbf{y} = \mathbf{C}(\mathbf{p})\mathbf{x}$$
FOM
FOM
$$\mathbf{FOM}$$

$$\mathbf{FOM}$$

$$\mathbf{FOM}$$

$$\mathbf{FOM}$$

$$\mathbf{FOM}$$

$$\mathbf{FOM}$$

$$\mathbf{W}^{T}\mathbf{r} = \mathbf{0}$$

 $\mathbf{x} \in \mathbf{R}^N$ : state vector  $\mathbf{p} \in \mathbf{R}^{N_p}$ : parameter vector  $\mathbf{u} \in \mathbf{R}^{N_i}$ : input vector  $\mathbf{y} \in \mathbf{R}^{N_o}$ : output vector

 $\dot{\mathbf{x}}$ 

У

 $\mathbf{x}_r \in \mathbf{R}^n$ : reduced state vector  $\mathbf{V} \in \mathbf{R}^{N \times n}$ : reduced basis

### **Machine learning**

"Machine learning is a field of computer science that uses statistical techniques to give computer systems the ability to "learn" with data, without being explicitly programmed." [Wikipedia]

### **Reduced order modeling**

"Model order reduction (MOR) is a technique for reducing the computational complexity of mathematical models in numerical simulations." [Wikipedia]

The difference in fields is perhaps largely one of history and perspective: model reduction methods have grown from the scientific computing community, with a focus on *reducing* high-dimensional models that arise from physics-based modeling, whereas machine learning has grown from the computer science community, with a focus on *creating* low-dimensional models from black-box data streams. Yet recent years have seen an increased blending of the two perspectives and a recognition of the associated opportunities. [Swischuk et al., *Computers & Fluids*, 2018]

## Outline

domain-aware interpretable analyzable

- 1. Reduced models can be learned from data
- 2. Basis expansions can be constructed so as to respect physical constraints
- 3. Structure can be exposed through variable transformations

This work was supported in part by the U.S. Air Force **Center of Excellence on Multi-Fidelity Modeling of Rocket Combustor Dynamics**, the AFOSR **MURI on Managing Multiple Information Sources of Multi-physics Systems**, the U.S. Department of Energy Applied Mathematics program as part of the **DiaMonD Multifaceted Mathematics Integrated Capability Center**, the AFOSR **Dynamic Data Driven Application System** Program, and the **MIT-SUTD International Design Center**.





# 1. Reduced models can be learned from data

Given state snapshot data (simulation or experimental), learn the dynamical system that (may have) generated it



ROM: 
$$\widehat{\mathbf{E}}\dot{\widehat{\mathbf{x}}}=\widehat{\mathbf{A}}\widehat{\mathbf{x}}+\widehat{\mathbf{B}}\mathbf{u}$$

ROM preserves linear structure:

 $\widehat{\mathbf{A}} = \mathbf{V}^{\top} \mathbf{A} \mathbf{V}, \ \widehat{\mathbf{B}} = \mathbf{V}^{\top} \mathbf{B}, \widehat{\mathbf{E}} = \mathbf{V}^{\top} \mathbf{E} \mathbf{V}$ 

## **Quadratic Model**

FOM: 
$$\mathbf{E}\dot{\mathbf{x}} = \underbrace{\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}}_{\text{linear}} + \underbrace{\mathbf{H}(\mathbf{x} \otimes \mathbf{x})}_{\text{quadratic}}$$

ROM:  $\widehat{\mathbf{E}}\dot{\widehat{\mathbf{x}}} = \widehat{\mathbf{A}}\widehat{\mathbf{x}} + \widehat{\mathbf{B}}\mathbf{u} + \widehat{\mathbf{H}}(\widehat{\mathbf{x}}\otimes\widehat{\mathbf{x}})$ 

ROM preserves quadratic structure:

$$\widehat{\mathbf{H}} = \mathbf{V}^{\top} \mathbf{H} (\mathbf{V} \otimes \mathbf{V})$$

Projection-based model reduction gives us the mathematical lens through which to learn physics-based low-dimensional models from data Given state data, learn the system

## **Operator Inference**

Peherstorfer & W. Data-driven operator inference for nonintrusive projection-based model reduction, *Computer Methods in Applied Mechanics and Engineering*, 2016

$$\mathbf{E}\dot{\mathbf{x}} = \underbrace{\mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}}_{\text{linear}} + \underbrace{\mathbf{H}(\mathbf{x} \otimes \mathbf{x})}_{\text{quadratic}}$$

Given state data (X) and velocity data ( $\dot{X}$ ):

$$\mathbf{X} = \begin{bmatrix} | & | \\ \mathbf{x}(t_1) & \dots & \mathbf{x}(t_K) \\ | & | \end{bmatrix} \qquad \dot{\mathbf{X}} = \begin{bmatrix} | & | \\ \dot{\mathbf{x}}(t_1) & \dots & \dot{\mathbf{x}}(t_K) \\ | & | \end{bmatrix}$$

Find the operators **A**, **B**, **E**, **H** by solving the least squares problem:

 $\min_{A,B,E,H} \left\| \mathbf{X}^{\top} \mathbf{A}^{\top} + (\mathbf{X} \otimes \mathbf{X})^{\top} \mathbf{H}^{\top} + \mathbf{U}^{\top} \mathbf{B}^{\top} - \dot{\mathbf{X}}^{\top} \mathbf{E} \right\|$ 

In a global basis, here via the proper orthogonal decomposition (POD)

### **Operator Inference** [Peherstorfer & Willcox, 2016]

1. Generate full state trajectories (from high-fidelity simulation)

$$\mathbf{X} = \begin{bmatrix} | & | \\ \mathbf{x}(t_1) & \dots & \mathbf{x}(t_K) \\ | & | \end{bmatrix} \qquad \dot{\mathbf{X}} = \begin{bmatrix} | & | \\ \dot{\mathbf{x}}(t_1) & \dots & \dot{\mathbf{x}}(t_K) \\ | & | \end{bmatrix}$$

In a global basis, here via the proper orthogonal decomposition (POD)

### **Operator Inference** [Peherstorfer & Willcox, 2016]

- 1. Generate full state trajectories (from high-fidelity simulation)
- 2. Compute POD basis from these trajectories

 $X = V \: \Sigma \: W^\top$ 

In a global basis, here via the proper orthogonal decomposition (POD)

### **Operator Inference** [Peherstorfer & Willcox, 2016]

- 1. Generate full state trajectories (from high-fidelity simulation)
- 2. Compute POD basis from these trajectories
- Project trajectories onto POD basis, to obtain trajectories in low-dimensional POD coordinate space

$$\widehat{\mathbf{X}} = \mathbf{V}^{\top}\mathbf{X}$$

In a global basis, here via the proper orthogonal decomposition (POD)

### **Operator Inference** [Peherstorfer & Willcox, 2016]

- 1. Generate full state trajectories (from high-fidelity simulation)
- 2. Compute POD basis from these trajectories
- Project trajectories onto POD basis, to obtain trajectories in low-dimensional POD coordinate space
- 4. Solve least squares minimization problem to infer the low-dimensional model

 $\min_{\widehat{A},\widehat{B},\widehat{E},\widehat{H}} \left\| \widehat{X}^{\top} \widehat{A}^{\top} + \left( \widehat{X} \otimes \widehat{X} \right)^{\top} \widehat{H}^{\top} + U^{\top} \widehat{B}^{\top} - \dot{\widehat{X}}^{\top} \widehat{E} \right\|$ 

In a global basis, here via the proper orthogonal decomposition (POD)

### **Operator Inference** [Peherstorfer & Willcox, 2016]

- 1. Generate full state trajectories (from high-fidelity simulation)
- 2. Compute POD basis from these trajectories
- 3. Project trajectories onto POD basis, to obtain trajectories in low-dimensional POD coordinate space
- 4. Solve least squares minimization problem to infer the low-dimensional model

Under certain conditions, recovers the intrusive POD reduced model

→ convenience of black-box learning + rigor of projection-based reduction + structure imposed by physics



2. Basis expansions can be constructed so as to respect physical constraints Using particular solutions to enforce constraints by construction

## Representing a high-dimensional state in a low-dimensional basis

 $\mathbf{x}(t) \approx \mathbf{V} \, \hat{\mathbf{x}}(t) = \sum_{i=1}^{r} \mathbf{V}_i \, \hat{x}_i(t)$ 

Proper orthogonal decomposition (POD):

- Basis vectors  $\mathbf{V}_i$  are linear combinations of snapshots
- If snapshots satisfy homogeneous conditions (e.g., BCs, divergence, etc.)
   → basis vectors satisfy those conditions
   → reconstructed solution satisfies conditions

We can enforce other (non-homogeneous) conditions using particular solutions



Computing particular solutions

Also known in model reduction literature as "static corrections"

Romanowski & Dowell, 1994; Hall, Thomas & Dowell, 2000; Willcox, 2000 Simple example: representing the temperature profile T(z, t) in a 1D heated rod

$$z = 0 \xrightarrow{Z} z = L$$

$$T(0,t) = \gamma_0 f(t) \qquad T(L,t) = \gamma_L$$

Auxiliary problem 1: solution  $\overline{T}^{L}(z)$ BCs: T(0,t) = 0 T(L,t) = 1

Auxiliary problem 2: solution  $\overline{T}^{0}(z)$ 

BCs: T(0, t) = 1

BCs:

$$T(L,t)=0$$

# Particular solutions

used to enforce nonhomogeneous boundary conditions

Swischuk, Mainini, Peherstorfer & W., *Computers & Fluids*, 2018 Modify snapshots to satisfy homogenous BCs:

$$\tilde{T}(z,t_j) = T(z,t_j) - \underbrace{\gamma_0 f(t_j) \,\overline{T}^0(z) - \gamma_L \overline{T}^L(z)}_{\swarrow}$$

particular solutions scaled by the BC for that snapshot

- POD basis vectors  $V_i$ , i = 1, ..., r computed from modified snapshots satisfy homogeneous conditions
- Temperature solution expanded in POD basis with modal coefficients  $\hat{T}_i$ , i = 1, ..., r
- Reconstructed solution:

$$T(x,t) = \gamma_0 f(t) \,\overline{T}^0(z) - \gamma_L \overline{T}^L(z) + \underbrace{\sum_{i=1}^{\prime} V_i(z) \,\widehat{T}_i(t)}_{\text{enforces a specified set of BCs}} + \underbrace{\sum_{i=1}^{\prime} V_i(z) \,\widehat{T}_i(t)}_{\text{satisfies homogeneous BCs}}$$



## 3. Lifting

Structure can be exposed through variable transformations

# Very simple example

Lifting a 4<sup>th</sup>-order ODE to quadratic-bilinear form.

Can either lift to a system of ODEs or to a system of DAEs Consider the fourth order system

Introduce auxiliary variables:

Chain rule:

$$\dot{x} = x^4 + u$$

s: 
$$w_1 = x^2 \quad w_2 = w_1^2$$

$$\dot{w}_1 = 2x[w_1^2 + u] = 2x[w_2 + u]$$
$$\dot{w}_2 = 2w_1\dot{w}_1 = 4xw_1[w_2 + u]$$

Need additional variable to make auxiliary dynamics quadratic:  $w_3 = xw_1$   $\dot{w}_3 = \dot{x}w_1 + x\dot{w}_1$   $= w_1w_2 + w_1u + 2w_1w_2 + 2w_1u$ QB-ODE  $\dot{x} = w_2 + u$  QB-DAE

$$\dot{w}_1 = 2xw_2 + 2xu \\ \dot{w}_2 = 4w_2w_3 + 4w_3u$$

 $\dot{w}_3 = 3w_1w_2 + 3w_1u_1$ 



## Lifting example: Euler equations

Transformation to specific volume form (use  $1/\rho$  in place of  $\rho$ ) yields a quadratic system of ODEs

$$\frac{\partial}{\partial t} \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix} + \frac{\partial}{\partial x} \begin{pmatrix} \rho u \\ \rho u^{2} + p \\ (E + p)u \end{pmatrix} = 0$$
$$E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho u^{2}$$



primitive variables

- Define auxiliary variable:  $q = 1/\rho$
- Take derivative:  $\frac{\partial q}{\partial t} = \frac{-1}{\rho^2} \frac{\partial \rho}{\partial t} = \frac{-1}{\rho^2} \left( -\rho \frac{\partial u}{\partial x} u \frac{\partial \rho}{\partial x} \right) = q \frac{\partial u}{\partial x} u \frac{\partial q}{\partial x}$

lifted variables

## Lifting example: Tubular reactor

 $\psi$ : species concentration  $\theta$ : temperature  $\mathcal{D}$  : Damköhler number Pe: Pèclet number  $\beta, B, \theta_0, \gamma$ : known constants

#### **Governing equations:**

$$\dot{\psi} = \frac{1}{Pe} \psi_{ss} - \psi_s - \mathcal{D}\psi e^{\gamma - \frac{\gamma}{\theta}}, \qquad s \in (0, 1), \ t > 0$$
$$\dot{\theta} = \frac{1}{Pe} \theta_{ss} - \theta_s - \beta(\theta - \theta_{ref}) + \mathcal{B}\mathcal{D}\psi e^{\gamma - \frac{\gamma}{\theta}}$$

Lifting transformations:  $w_1 = e^{\gamma - \frac{\gamma}{\theta}}, \quad w_2 = \theta^{-2}, \quad w_3 = \theta^{-1}$ 

Auxiliary dynamics for lifted variables are quartic:

$$\dot{w}_1 = w_1(\gamma \ \theta^{-2})\dot{\theta} = \gamma \underbrace{w_1 w_2 \dot{\theta}}_{\text{quartic}}$$
$$\dot{w}_2 = -2\theta^{-3}\dot{\theta} = -2 \underbrace{w_2 w_3 \dot{\theta}}_{\text{quartic}}$$
$$\dot{w}_3 = -\theta^{-2}\dot{\theta} = -\underbrace{w_2 \dot{\theta}}_{\text{cubic}},$$

...while the original equations become quadratic

$$\dot{\psi} = \underbrace{\frac{1}{Pe}\psi_{ss} - \psi_s}_{\text{linear}} - \mathcal{D}\underbrace{\psi w_1}_{\text{quadratic}}$$
$$\dot{\theta} = \underbrace{\frac{1}{Pe}\theta_{ss} - \theta_s}_{\text{linear}} - \beta(\theta - \theta_{\text{ref}}) + \mathcal{B}\mathcal{D}\underbrace{\psi w_1}_{\text{quadratic}}$$

22

## Lifting example: Tubular reactor

Quadratic-bilinear form achievable with differential-algebraic equations (DAEs) To get to QB form, need additional auxiliary variables:

$$w_4 = \psi w_1, \quad w_5 = w_2 w_3, \quad w_6 = w_1 w_2$$

The lifted system then becomes a quadratic-bilinear DAE:

$$\dot{\psi} = \underbrace{\frac{1}{Pe}\psi_{ss} - \psi_s - \mathcal{D}w_4}_{\text{linear}}$$
$$\dot{\theta} = \underbrace{\frac{1}{Pe}\theta_{ss} - \theta_s - \beta(\theta - \theta_{\text{ref}}) + \mathcal{B}\mathcal{D}w_4}_{\text{linear}}$$
$$\dot{w}_1 = \gamma \ w_6 \left[\frac{1}{Pe}\psi_{ss} - \psi_s\right] + \gamma \mathcal{B}\mathcal{D} \ w_4 w_6$$
$$\dot{w}_2 = -2 \ w_5 \left[\frac{1}{Pe}\psi_{ss} - \psi_s\right] - 2\mathcal{B}\mathcal{D} \ w_4 w_5$$

$$\dot{w}_3 = -w_2 \left[ \frac{1}{Pe} \psi_{ss} - \psi_s \right] - \mathcal{BD} \ w_2 w_4$$
$$0 = w_4 - w_1 \psi$$
$$0 = w_5 - w_2 w_3$$

 $0 = w_6 - w_1 w_2$ 

## Lifting summary: **Tubular reactor**

- Introduce six auxiliary variables; state increase from 2n to 8n
- Lift to a QB-DAE

$$\dot{\psi} = \frac{1}{Pe}\psi_{ss} - \psi_s - \mathcal{D}\psi e^{\gamma - \frac{\gamma}{\theta}}$$

$$\dot{\theta} = \frac{1}{Pe}\theta_{ss} - \theta_s - \beta(\theta - \theta_{ref}) + \mathcal{B}\mathcal{D}\psi e^{\gamma - \frac{\gamma}{\theta}}$$
original equations
$$\dot{\psi} = \underbrace{\frac{1}{Pe}\psi_{ss} - \psi_s - \mathcal{D}w_4}_{\text{linear}}$$

$$\dot{\theta} = \underbrace{\frac{1}{Pe}\theta_{ss} - \theta_s - \beta(\theta - \theta_{ref}) + \mathcal{B}\mathcal{D}w_4}_{\text{linear}}$$

$$\dot{w}_1 = \gamma \ w_6 \left[\frac{1}{Pe}\psi_{ss} - \psi_s\right] + \gamma \mathcal{B}\mathcal{D} \ w_4 w_6$$

$$\dot{w}_2 = -2 \ w_5 \odot \left[\frac{1}{Pe}\psi_{ss} - \psi_s\right] - 2\mathcal{B}\mathcal{D} \ w_4 w_5$$

$$\dot{w}_3 = -w_2 \odot \left[\frac{1}{Pe}\psi_{ss} - \psi_s\right] - \mathcal{B}\mathcal{D} \ w_2 w_4$$

$$0 = w_4 - w_1 \psi$$

0 =

$$\underbrace{Pe^{\circ ss} \quad \forall s \quad \forall e^{\circ} \quad \forall e^$$

## Tubular reactor: QB-POD ROM

- Finite difference discretization with n points per unknown
- Recorded snapshots for POD every  $\Delta t = 0.01s$

#### **Original system**

 $\dot{\mathbf{x}} = f(\mathbf{x}) + \mathbf{B}\mathbf{u}$ 

$$\begin{aligned} \dot{\boldsymbol{\psi}} &= \mathbf{A}_{\boldsymbol{\psi}} \boldsymbol{\psi} + \mathbf{b}_{\boldsymbol{\psi}} u(t) - \mathcal{D} \ \boldsymbol{\psi} \odot e^{\gamma - \frac{\gamma}{\boldsymbol{\theta}}} \\ \dot{\boldsymbol{\theta}} &= \mathbf{A}_{\boldsymbol{\theta}} \boldsymbol{\theta} + \mathbf{b}_{\boldsymbol{\theta}} u(t) + \mathcal{B} \mathcal{D} \ \boldsymbol{\psi} \odot e^{\gamma - \frac{\gamma}{\boldsymbol{\theta}}} \end{aligned}$$

#### Lifted system

 $\mathbf{E}\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} + \mathbf{H}(\mathbf{x}\otimes\mathbf{x}) + \sum_{k=1}^{m}\mathbf{N}_{k}\mathbf{x}u_{k}$ 

$$\dot{\boldsymbol{\psi}} = \mathbf{A}_{\boldsymbol{\psi}} \boldsymbol{\psi} + \mathbf{b}_{\boldsymbol{\psi}} u(t) - \mathcal{D} \mathbf{w}_{4}$$
$$\dot{\boldsymbol{\theta}} = \mathbf{A}_{\boldsymbol{\theta}} \boldsymbol{\theta} + \mathbf{b}_{\boldsymbol{\theta}} u(t) + \mathcal{B} \mathcal{D} \mathbf{w}_{4}$$
$$\dot{\mathbf{w}}_{1} = \gamma \mathbf{w}_{6} \odot [A_{2}\boldsymbol{\theta} + \mathbf{b}_{\boldsymbol{\theta}} u(t)] + \gamma \mathcal{B} \mathcal{D} \mathbf{w}_{4} \odot \mathbf{w}_{6}$$
$$\dot{\mathbf{w}}_{2} = -2 \mathbf{w}_{5} \odot [A_{2}\boldsymbol{\theta} + \mathbf{b}_{\boldsymbol{\theta}} u(t)] - 2\mathcal{B} \mathcal{D} \mathbf{w}_{4} \odot \mathbf{w}_{5}$$
$$\dot{\mathbf{w}}_{3} = -\mathbf{w}_{2} \odot [A_{2}\boldsymbol{\theta} + \mathbf{b}_{\boldsymbol{\theta}} u(t)] - \mathcal{B} \mathcal{D} \mathbf{w}_{2} \odot \mathbf{w}_{4}$$
$$0 = \mathbf{w}_{4} - \mathbf{w}_{1} \odot \boldsymbol{\psi}$$
$$0 = \mathbf{w}_{5} - \mathbf{w}_{2} \odot \mathbf{w}_{3}$$
$$0 = \mathbf{w}_{6} - \mathbf{w}_{1} \odot \mathbf{w}_{2}$$



## Summary

- Physics-based models have structure; learned models should exploit/respect that structure
- **Projection**: a structure-preserving lens

## Learn

Infer a low-dimensional model directly from data of the original system, but through the lens of projection

## Lift

Introduce transformations and auxiliary variables to express the physics in a structured form, then learn a reduced model

→ Elizabeth Qian talk, Tuesday 1035h (401A) & Thursday 1505h (401C)

## **Particularize**

Use particular solutions to enforce boundary conditions and other physical constraints  $\rightarrow$  <u>Renee Swischuk poster</u>, Tuesday PM

# **Data-driven** decisions

building the mathematical foundations and computational methods to enable design of the next generation of engineered systems

## 



Aerospace Engineering and Engineering Mechanics Cockrell School of Engineering





The University of Texas at Austin Institute for Computational Engineering and Sciences