Learning Physics-Based Models from Data: Perspectives from Inverse Problems and Model Reduction (Ghattas, O. & Willcox, K. (2021))

OUARDI Ahlam

The UM6P Vanguard Center

May 29, 2025

University Mohammed VI Polytechnic

THE UM6P VANGUARD CENTER

#Empowering Minds.

Workflow to Learn Physics-Based Models from Data

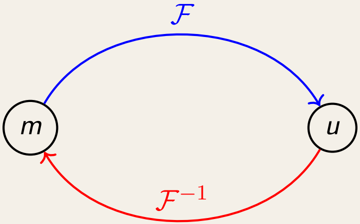
1. **Decide on the Problem**
   Clearly define the physical system and the questions to be answered.
2. **Curate Data**
   Gather, clean, and preprocess relevant experimental or simulated data.
3. **Design an Architecture**
   Choose a model structure: purely data-driven, physics-informed, or hybrid.
4. **Craft a Loss Function**
   Define a metric to quantify mismatch between predictions and data, while incorporating physical constraints.
5. **Employ Optimization**
   Minimize the loss function to learn model parameters using suitable algorithms.

Outline

Introduction

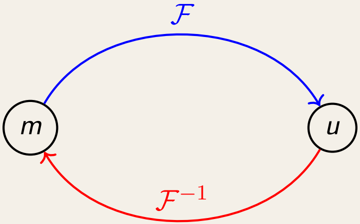
The inverse problem consists of finding m for a given u:

$$\mathcal{F}(m) = u. \tag{1}$$



The inverse problem consists of finding m for a given u:

$$\mathcal{F}(m) = u. \qquad (1)$$



Hadamard (1923) postulated three conditions for problem (1) to be well-posed.

(i) *Existence*. For all $u \in Y$, there exists at least one $m \in X$ such that $F(m) = u$.

(ii) *Uniqueness*. For all $u \in Y$, there is at most one $m \in X$ such that $F(m) = u$.

(iii) *Stability*. The parameter m depends on u continuously.

Poisson's Equation in Electrostatics

The Poisson equation describes the electric potential $u(\mathbf{x})$ in the presence of a charge density $\rho(\mathbf{x})$:

$$-\nabla^2 u = \frac{\rho}{\varepsilon_0}, \quad \mathbf{x} \in \Omega. \tag{2}$$

Example: A thin charged plate ($x \in [0, L]$) with a charge density $\rho(x)$ obeys:

$$-\frac{d^2 u}{dx^2} = \frac{\rho(x)}{\varepsilon_0}, \quad 0 < x < L. \tag{3}$$

where:

- $u(\mathbf{x})$ is the electric potential at position $\mathbf{x}$.
- $\rho(\mathbf{x})$ is the charge density.
- ε_0 is the permittivity of free space.

The solution $u(\mathbf{x})$ describes how the electric potential spreads due to a given charge distribution $\rho(\mathbf{x})$.

Poisson's Equation and Inverse Problem

Let's consider the following Poisson equation:

$$\begin{aligned} -k\frac{\partial^2 u}{\partial x^2} &= m(x), \quad 0 < x < L, \\ u(0) &= u(L) = 0. \end{aligned} \tag{4}$$

where $k > 0$ is a constant. We define the operator $\mathcal{A}$ as follows:

$$\mathcal{A} : u \mapsto -k\frac{\partial^2 u}{\partial x^2}$$

u is the state variable.

Remark

The equation (4) admits a unique solution $u \in H_0^1$ for all $m \in H^{-1}$, i.e., the operator $\mathcal{A}$ is bijective.

Spectral Properties of the Operator

The operator $\mathcal{F} := \mathcal{A}^{-1}$ is well-defined and self-adjoint, thanks to the self-adjointness of $\mathcal{A}$. Its eigenvalues are real, and its eigenfunctions $v_j(x)$, $j = 1, 2, \ldots, \infty$, are given by:

$$v_j(x) = \sin\left(\frac{j\pi x}{L}\right).$$

with associated eigenvalues:

$$\lambda_j(\mathcal{F}) \approx \frac{1}{k}\left(\frac{L}{j\pi}\right)^2.$$

As $j \to \infty$, we have $\lambda_j \to 0$, thus $\mathcal{F}$ is a compact operator.

Spectral Properties of the Operator

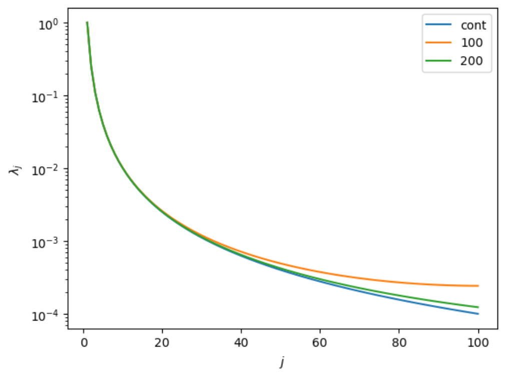


Figure: Spectrum of the continuous operator $\mathcal{F}$ (blue) versus that of the discretized operator $\mathcal{F}$

Solving for the Source $m(x)$

Given $u(x)$, we attempt to solve for the source $m(x)$:

$$\mathcal{F}(m) = u \Longrightarrow m = \mathcal{F}^{-1}u.$$

Using the spectral decomposition of $\mathcal{F}$, we obtain:

$$m = \mathcal{F}^{-1}u = \sum_{j=1}^{\infty} \frac{\langle v_j, u\rangle}{\lambda_j} v_j.$$

where the inner product is:

$$\langle v_j, u\rangle = \int_0^L v_j u dx.$$

For a solution m to exist, the Fourier coefficients of the data, $\langle v_j, u\rangle$, must decay to zero faster than the eigenvalues λ_j.

Solving for the source $m(x)$

We assume, in the following, that there exists additive noise η that represents the discrepancy between the data and the model output for the 'true' parameter m_{true} :

$$\mathcal{F}\left(m_{\text{true}}\right) + \eta = u$$

In this case, we can write

$$m = m_{\text{true}} + \sum_{j=1}^{\infty} \frac{\langle v_j, \eta \rangle}{\lambda_j} v_j$$

Then the error in inferring the source is given by

$$\left\| m - m_{\text{true}} \right\|^2 = \sum_{j=1}^{\infty} \frac{\eta_j^2}{\lambda_j^2}$$

where $\eta_j = \langle \nu_j, \eta \rangle$ are the Fourier components of the noise.

Ill-posed Problems and Regularization

In inverse problems, solving for $m(x)$ from noisy data is an ill-posed problem. **Regularization of Tikhonov** is a method to stabilize the solution:

$$\min_{m} \left\{ \|F(m) - d\|^2 + \beta \|m\|^2 \right\} \tag{5}$$

where:

- $F(m)$ is the forward operator (e.g., solving Poisson's equation).
- d is noisy data.
- β is the regularization parameter.

Regularization of Tikhonov

This regularization parameter is chosen based on **the Morozov discrepancy principle**: select largest β^* such that

$$\|F(m_{\beta^*}) - u\| \leq \delta,$$

where δ is the noise level and m_{β^*} is the solution to the regularized problem.

⚠ In practice, we don't know the noise expression explicitly i.e. δ is unknowable.

Choosing the Regularization Parameter

We consider the Tikhonov-regularized inverse problem:

$$m_\beta = \arg\min_{m\in X}\left(\frac{1}{2}\|F(m)-u\|^2+\frac{\beta}{2}\|m\|^2\right).$$

The Morozov discrepancy principle corresponds to finding:

$$\min_{m\in X}\frac{1}{2}\|m\|^2 \quad \text{subject to } \|F(m)-u\|^2=\delta^2.$$

To solve this, we introduce the Lagrangian:

$$\mathcal{L}(m,\mu)=\frac{1}{2}\|m\|^2+\mu\left(\|F(m)-u\|^2-\delta^2\right).$$

the Lagrange multiplier μ^* at the optimum is related to the regularization parameter β^* by

$$\beta^*=\frac{1}{2\mu^*}.$$

Let's apply this to the example (4), using conjugate gradient to solve the optimization problem (5):

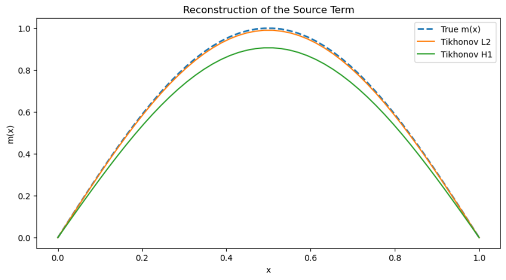
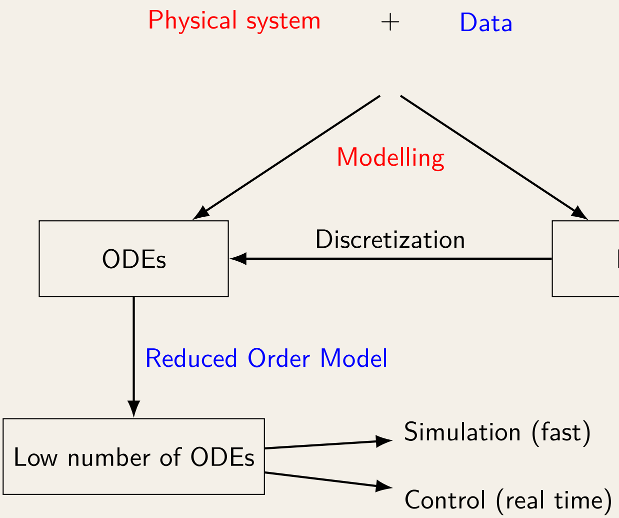


Figure: Recovered source term in the Poisson equation from a noisy data using Tikhonov regularization method.

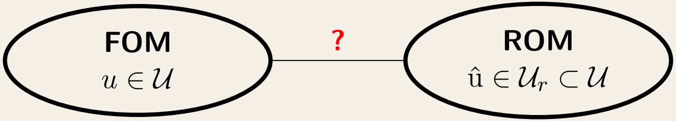
Reduced Order Modelling



Projection-Based Model Reduction

Consider the linear PDE with appropriate boundary and initial conditions:

$$\frac{\partial u}{\partial t} = \mathcal{A}(u) \quad \text{in } \Omega \times (0, t_f)\,, \tag{6}$$



Define r basis vectors $v_1(x), \ldots, v_r(x)$ that span $\mathcal{U}_r$ and that form an orthonormal set, i.e. $\langle v_i, v_j \rangle = \delta_{ij}$. Then

$$u(x,t) \approx \sum_{j=1}^{r} v_j(x)\hat{u}_j(t) = V^T \hat{\mathrm{u}} \tag{7}$$

where $\hat{u}_j, j = 1, \ldots, r$ are the reduced model's coefficients of expansion in the basis v_j.

Galerkin projection

Substituting the approximation (7) into the governing equation (6) yields the residual

$$r(x,t) = \sum_{j=1}^{r} v_j \frac{\mathrm{d}\hat{u}_j}{\mathrm{d}t} - \sum_{j=1}^{r} \mathcal{A}\left(v_j\right)\hat{u}_j.$$

- **Galerkin projection:** $\langle r, v_i \rangle = 0, i = 1, \ldots, r.$

This yields the reduced model

$$\langle r, v_i \rangle = 0 \Rightarrow \frac{\mathrm{d}\hat{u}_i}{\mathrm{d}t} = \sum_{j=1}^{r} \hat{\mathcal{A}}_{ij}\hat{u}_j, \quad i = 1, \ldots, r,$$

where $\hat{\mathcal{A}}_{ij} = \langle v_i, \mathcal{A}\left(v_j\right)\rangle$ is the reduced linear operator, which can be precomputed once the basis is defined.

Proper orthogonal decomposition

Steps:

- Given a set of snapshots $\{u_1, \ldots, u_m\}$ in $\mathbb{R}^n$, define the snapshot matrix $U = [u_1 \ldots u_m]$.
- Compute the correlation matrix: $C = UU^T \in \mathbb{R}^{n \times n}$.
- Solve the eigenproblem: $C\phi_i = \lambda_i \phi_i$.
- Select the first r modes $V = [\phi_1 \ldots \phi_r]$ associated with the largest λ_i.

Energy Criterion: Choose r such that:

$$\frac{\sum_{i=1}^{r} \lambda_i}{\sum_{i=1}^{n} \lambda_i} \geq \kappa \quad (\text{e.g., } \kappa = 0.99)$$

POD minimizes projection error in Frobenius norm.

Model Reduction using POD: Setup

We illustrate the use of Proper Orthogonal Decomposition (POD) to reduce the order of a linear dynamical system.

We start from the semi-discretized heat equation:

$$\frac{dx}{dt} = Ax + Bu(t), \quad y = Cx$$

- $A \in \mathbb{R}^{n \times n}$: 1D Laplacian with Dirichlet BCs
- $B \in \mathbb{R}^{n \times 1}$, $C \in \mathbb{R}^{1 \times n}$: input/output coupling
- Domain: $n = 100$, $L = 1$, $\alpha = 0.01$, $\Delta x = \frac{L}{n+1}$
- Input: $u(t) = \sin(2\pi t)$

Discretization and POD Method

Matrix A has the structure:

$$A = \frac{\alpha}{\Delta x^2} \begin{bmatrix} -2 & 1 & 0 & \cdots & 0 \\ 1 & -2 & 1 & \cdots & 0 \\ 0 & 1 & -2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -2 \end{bmatrix}$$

Simulation over $[0, 5]$ with 300 time steps.

- Collect snapshots $x(t) \in \mathbb{R}^n$ into matrix $X \in \mathbb{R}^{n \times m}$
- Apply SVD: $X = U \Sigma V^T$
- Reduced basis: U_r (first $r = 3$ columns)

Reduced-order model:

$$\frac{dz}{dt} = A_r z + B_r u(t), \quad y_r = C_r z$$

with:

$$A_r = U_r^T A U_r, \quad B_r = U_r^T B, \quad C_r = C U_r$$

FOM vs ROM Output

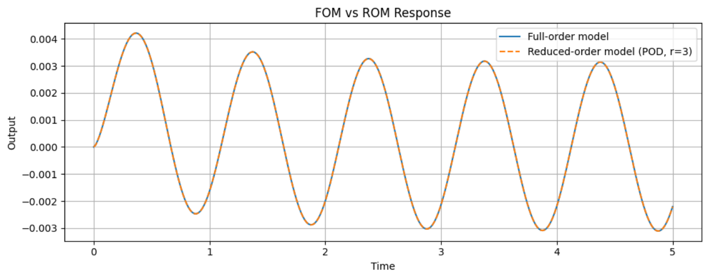


Figure: Output response of FOM and ROM (POD, $r = 3$)

- The reduced model closely matches the full model
- Stability is preserved under POD reduction

Error and Stability Analysis

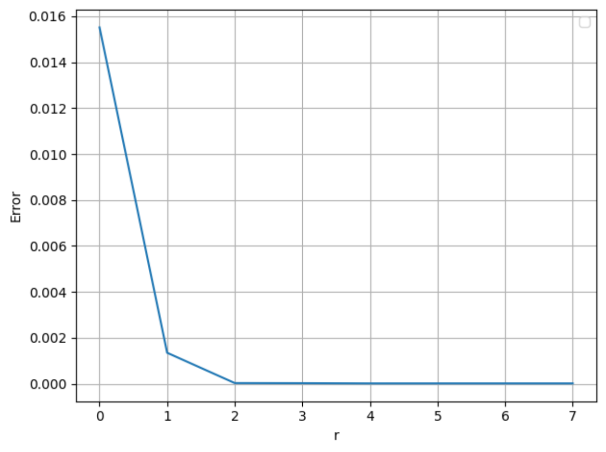


Figure: $\|y_{\text{fom}} - y_{\text{rom}}\|_2$ as a function of r

Error and Stability Analysis

Model	Max Real Part of Eigenvalues	Execution time
Full-order (FOM)	$-0.9869403481356569 < 0$	29.401416301727295
Reduced-order (ROM)	$-0.9920437792538935 < 0$	0.034462213516235

Table: Comparison of stability (all poles in left half-plane), and time of the programs' execution.

- All poles remain in left half-plane $\Rightarrow$ stability preserved
- Significant reduction in computation time

Remark

The results depends on the number m of snapshots.

Model Reduction of Port-Hamiltonian Systems

We consider the port-Hamiltonian (pH) system:

$$\begin{cases} \dot{x}(t) = (J - R)\nabla_x H(x(t)) + Bu(t) \\ y(t) = B^\top \nabla_x H(x(t)) \end{cases}$$

where $J = -J^\top$, $R = R^\top \geq 0$, $H(x)$: Hamiltonian (energy)

- Dissipation inequality:

$$\frac{dH}{dt} = \nabla_{\mathbf{x}} H(\mathbf{x}(t))^\top . \frac{dx}{dt} \leq y(t)^\top u(t)$$

- Galerkin projection using basis $V \in \mathbb{R}^{n \times r}$ yields:

$$\dot{\hat{x}} = (J_r - R_r)\nabla_{\hat{x}} \hat{H}(\hat{x}) + B_r u(t)$$

where $\hat{H}(\hat{x}) = H(V\hat{x})$

$\Rightarrow$ Structure-preserving reduced system satisfies same dissipation inequality.

Stability Preservation under Galerkin Projection

- Original system: $\dot{x} = Ax$, where all eigenvalues of A have strictly negative real parts (i.e. A Hurwitz)

$$\Rightarrow \exists P > 0 \quad \text{s.t. } A^\top P + PA = -Q, \quad \text{for any } Q = Q^\top > 0$$

- Projection with $V \in \mathbb{R}^{n \times r}$, $V^\top V = I$:

$$\dot{\hat{x}} = \hat{A}\hat{x}, \quad \hat{A} = V^\top A V$$

- Reduced Lyapunov equation:

$$\hat{A}^\top \hat{P} + \hat{P}\hat{A} = -\hat{Q}, \quad \hat{P} = V^\top P V > 0$$

- $\Rightarrow \hat{A}$ is Hurwitz $\Rightarrow$ Reduced system is asymptotically stable.

References

- Ghattas, O. & Willcox, K. (2021). *Learning physics-based models from data: perspectives from inverse problems and model reduction*.
- Benning, M., & Burger, M. (2018). *Modern regularization methods for inverse problems. Acta numerica, 27, 1-111.*
- Benner, P., Gugercin, S., & Willcox, K. (2015). *A Survey of Projection-Based Model Reduction Methods*.
- Bui-Thanh, T., Willcox, K., & Ghattas, O. (2008). *Model Reduction for Large-Scale Systems with Parametric Input*.
- Yuezheng Gonga, Qi Wangb, , Zhu Wangb, (2016). *Structure-preserving Galerkin POD reduced-order modeling of Hamiltonian systems*.
- Mark S. Gockenbach, (2011). *Partial Differential Equations : Analytical and Numerical Methodes*.

Questions & Answers

Thank you :)