Geometric Multiscale Reduction for Autonomous and Controlled Nonlinear Systems

Jake Bouvrie[1] and **Mauro Maggioni**[2]

[1]Laboratory for Computational and Statistical Learning, MIT
[2]Mathematics, Computer Science, ECE Departments, Duke University

IEEE CDC, December 2012

Motivation

Many interesting dynamical systems have natural **multiscale structure**, in time and/or in space.

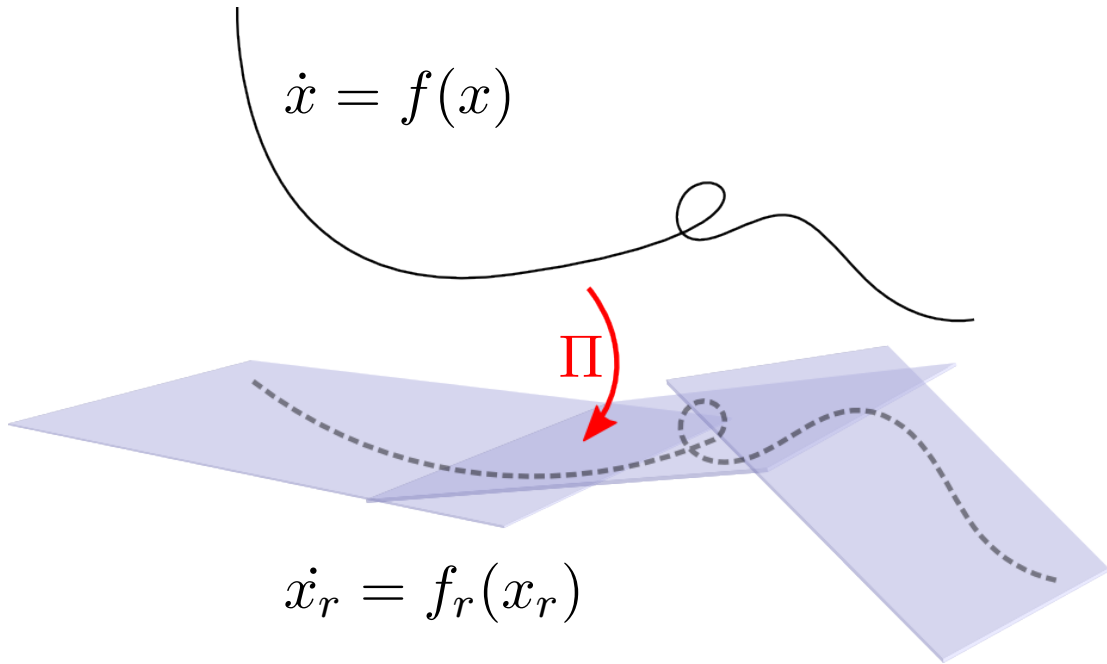
- Goal: *Statistical model reduction* that systematically takes advantage of this structure whenever possible.

Advantage: Higher-accuracy at lower complexity:

- Simpler controllers
- Faster simulations
- Revealing analyses/interpretations

Approach

A system will be approximated by a collection of *local, low-dimensional* systems at different scales:



A globally defined, reduced-order system is obtained by blending or switching between local sub-systems.

Approach

- **Step (1)**. Analyze the geometry of empirical trajectories $\Rightarrow$ build multiscale data tree.
  - Nodes are subsets of samples.
- **Step (2)**. Estimate approximating subspaces at nodes of the tree *(multiscale data-dependent dictionary)*, and project.
  - POD + Galerkin projection
  - Balanced truncation + Petrov-Galerkin projection
- **Step (3)**. Combine subsystems.

Geometric Data Tree Construction

Given a precision $\epsilon > 0$, build a multiscale tree $\mathcal{T} = \{\mathsf{C}_{j,k}\}$ organizing a dataset of N samples $\{x_i\}_{i=1}^N$ into a spatially hierarchical structure (with e.g., recursive spectral partitioning, METIS, iterated k-means,...).

- Each node corresponds to a subset of the samples (snapshots), $\mathsf{C}_{j,k}$ (scale j, cell k).
- Children partition the samples of their parent.
- *Bottom-up pruning:* If parent can be encoded to precision ϵ with $d \leq d_{\mathsf{max}}$ PCA components, then delete children.

Collection of leaf nodes of $\mathcal{T}$ is interpreted as a *partitioning of the statespace into clusters occurring at different scales*.

Local Reduction: POD

Suppose we are given a tree $\mathcal{T}$ of snapshots for the high-dimensional system

$$\begin{cases} \dot{\mathbf{X}} &= f(\mathbf{X}, \mathbf{u}) \\ \mathbf{y} &= h(\mathbf{X}) \end{cases}$$

with state $\mathbf{X} \in \mathbb{R}^D$, input $\mathbf{u} \in \mathbb{R}^p$, and output $\mathbf{y} \in \mathbb{R}^q$.
For each leaf-node cell $\mathsf{C}_{j,k}$, $(j,k) \in \text{Leaf}$, define the local mean

$$c_{j,k} = |\mathsf{C}_{j,k}|^{-1} \sum_{x \in \mathsf{C}_{j,k}} x$$

and covariance

$$\text{cov}_{j,k} = |\mathsf{C}_{j,k}|^{-1} \sum_{x \in \mathsf{C}_{j,k}} (x - c_{j,k})(x - c_{j,k})^*.$$

Local Reduction: POD

Take the SVD

$$\mathrm{cov}_{j,k} = U_{j,k}\Sigma^2_{j,k}U^*_{j,k}.$$

- Define $P_{j,k}$: matrix whose rows are the columns of $U_{j,k}$ corresponding to the $r < D$ largest singular values.

Local, affine Galerkin projection:

- Consider the dynamics of $\mathbf{X}' = \mathbf{X} - c_{j,k}$
- Project down to the subspace spanned by the r orthonormal rows of $P_{j,k}$.

A reduced-order dynamical system valid in the vicinity of $c_{j,k}$ is given by

$$\begin{cases} \dot{\mathbf{x}}_{j,k} = P_{j,k} f\left(P^*_{j,k}\mathbf{x}_{j,k} + c_{j,k}\right) \\ \mathbf{y} = h\left(P^*_{j,k}\mathbf{x}_{j,k} + c_{j,k}\right) \end{cases}$$

for each $(j,k) \in$ Leaf.

Local Reduction: POD - Comments

- Approximation of the statespace with an arrangement of low-dimensional planes.
- Approximating subspaces/dynamics only need to be good locally.
- Local systems need not have the same dimension (chosen adaptively based on ϵ).
- Approximation captures local input-to-state behavior, but ignores *state-to-output* properties.

Local Reduction: Balanced Truncation

Approach is similar to (Lall, Marsden, Glavaski '02), but adapted to the localized setting with some modifications.

Local Empirical Balanced Truncation

- Controllability tree + local POD.
- Observability tree + local observability subspace estimation.
- Local truncation & projection.

Local Reduction: Balanced Truncation

Controllability tree construction & local POD:

- Simulate the system with $\mathbf{u}(t) = \mathbf{e}_i \delta(t)$ for $i = 1, \ldots, p$ and $\mathbf{X}(0) = 0$, collect N_c state trajectory samples.
- Build a controllability tree $\mathcal{T}_c$ to precision ϵ.
- For each leaf node in the tree, compute the projections (partial isometries) $P_{j,k}$ described before.

Local Reduction: Balanced Truncation

Local observability assessment:

Let $\{\phi_{j,k}^{(i)}\}_{i=1}^{r}$ denote the top r orthonormal basis vectors spanning the POD subspace associated to node (j,k) of $\mathcal{T}_c$.

- For each $(j,k) \in \text{Leaf}(\mathcal{T}_c)$, run r simulations of the original system from initial conditions $\mathbf{X}(0) = \phi_{j,k}^{(i)} + c_{j,k}$, $i = 1, \ldots, r$, respectively, with $\mathbf{u}(t) = 0$. Collect N_o snapshots each.
- Assumes that if dim r POD reduction locally well-approximates the dynamics, then output trajectories should approximately lie in a subspace of dimension $\leq r$.
- Total of $r \cdot |\text{Leaf}|$ *short* simulations.

Local Reduction: Balanced Truncation

Collect the snapshot vectors $v(\alpha,\beta) = \left(\mathbf{y}_\alpha^{(i)}(t_\beta)\right)_{i=1}^{r} \in \mathbb{R}^r$ into a $(r \times N_o q)$ matrix $Y_{j,k}$.

A rank-r approximation of the true *local* observability gramian $W_{j,k}^o$ is given by

$$W_{j,k}^o \approx (P_{j,k}^* Y_{j,k})(P_{j,k}^* Y_{j,k})^* = P_{j,k}^* \widetilde{W}_{j,k}^o P_{j,k}$$

$\widetilde{W}_{j,k}^o = Y_{j,k} Y_{j,k}^*$ is the $(r \times r)$ reduced observability gramian local to leaf node (j,k) of $\mathcal{T}_c$.

Note: We do not explicitly form the large $(D \times D)$ matrix $W_{j,k}^o$.

Local Reduction: Balanced Truncation

Compute local balancing transformations $T_{j,k}$:

Define the $(N_o q \times N_c p)$ Hankel matrix of inner-products local to node (j,k),

$$H_{j,k} = Y_{j,k}^* P_{j,k} X_{j,k}$$

where

- $Y_{j,k}$ is the $(r \times N_o q)$ matrix of observability snapshots
- $X_{j,k}$ is the $(d \times N_c p)$ matrix of controllability snapshots

Compute the (sorted, economy) SVD $H_{j,k} = U_{j,k} \Sigma_{j,k} V_{j,k}^*$.
Truncated balancing coordinate transformation:

$$T_{j,k}^* = P_{j,k}^* Y_{j,k} \tilde{U}_{j,k} \tilde{\Sigma}_{j,k}^{-1/2}, \qquad T_{j,k}^{-1} = X_{j,k} \tilde{V}_{j,k} \tilde{\Sigma}_{j,k}^{-1/2}.$$

where $\tilde{U}_{j,k}, \tilde{V}_{j,k}, \tilde{\Sigma}_{j,k}$ denote the d-truncated submatrices.

Local Reduction: Balanced Truncation

Local Petrov-Galerkin Projection at node (j,k):

A balanced, reduced order system valid in a local region of the statespace near $c_{j,k}$ is obtained by performing a Petrov-Galerkin projection onto the d directions corresponding to the largest singular values of $H_{j,k}$:

$$\begin{cases} \dot{\mathbf{x}}_{j,k} = T_{j,k} f\left(T_{j,k}^{-1}\mathbf{x}_{j,k} + c_{j,k}\right) \\ \mathbf{y} = h\left(T_{j,k}^{-1}\mathbf{x}_{j,k} + c_{j,k}\right) \end{cases}$$

for each $(j,k) \in$ Leaf.

Note: *Time* enters via the locality of the transformations.

Combining Subsystems: Switching

Definition

The statespace **regions of responsibility** are defined as the Voronoi regions of the statespace induced by the centers $\{c_{j,k}\}_{(j,k)\in\text{Leaf}}$ under the metric ρ. In other words, a particular local system with index (j', k') is responsible for a state $s \in \mathbb{R}^D$ if s is ρ-closer to $c_{j',k'}$ than any other leaf node center.

Motivates the simple switching rule:

The active system at time t_+ is the system associated to the plane with center $c_{j,k}$ closest to $\mathbf{X}(t_-)$.

Combining Subsystems: Switching

Hand-off from one system to another is via initial conditions:

Parameterize subsystems with tuples $\{(\Phi_i, \Psi_i, c_i, f)\}_{i=1}^{|\text{Leaf}|}$ and define

$$f_i(x) \triangleq \Phi_i f\left(\Psi_i^* x + c_i\right) \qquad h_i(x) \triangleq h\left(\Psi_i^* x + c_i\right)$$

with $f_i : \mathbb{R}^{d_i} \to \mathbb{R}^{d_i}, d_i = \dim(\text{range } \Phi_i)$,

- Suppose we switch from system i_0 to i_1 at time t_{i_0,i_1}.

The global reduced system for $t_0 \le t \le t_{i_1,i_2}$ may be defined as

$$\dot{\mathbf{x}}(t) = \begin{cases} f_{i_0}(\mathbf{x}),\ \mathbf{x}(t_0) = \Phi_{i_0}(\mathbf{X}(t_0) - c_{i_0}), \\ \qquad\qquad\qquad\qquad \text{for } t_0 < t \le t_{i_0,i_1} \\ f_{i_1}(\mathbf{x}),\ \mathbf{x}(t_{i_0,i_1}) \leftarrow \Phi_{i_1}\left(\Psi_{i_0}^* \mathbf{x}(t_{i_0,i_1}) + c_{i_0} - c_{i_1}\right), \\ \qquad\qquad\qquad\qquad \text{for } t_{i_0,i_1} < t \le t_{i_1,i_2}. \end{cases}$$

Combining Subsystems: Blending

Transitions can be smoothed by blending systems in a neighborhood of a boundary. However,

- Local systems are expressed in different coordinate systems.
- Dimension of local systems can be different.

Solution:
Interpolate in a common space just large enough to include the planes participating in the blend. Common space is found with an SVD.

Combining Subsystems: Blending

Example: Blending two systems.

- $S \subseteq \mathbb{R}^D$: high dimensional statespace
- $S_i \supseteq \text{range}\Phi_i$: d_i dimensional subspace associated to local system i.
- Blending neighborhood $\mathcal{N} = \mathcal{N}(t)$ consists of two arbitrary (but distinct) local systems i and j.

A plane in S describing the common subspace $S_\mathcal{N} = S_i \oplus S_j$ may be found via the SVD:

$$U_\mathcal{N}^* \Sigma_\mathcal{N} V_\mathcal{N} = \text{svd}\big([(\Phi_i^* + c_i \mathbf{1}_{d_i}^\top)\ (\Phi_j^* + c_j \mathbf{1}_{d_j}^\top)]\big). \tag{1}$$

The common plane has dimension $d_\mathcal{N} \leq d_i + d_j$, center $c_\mathcal{N} = U_\mathcal{N}^* \mathbf{e}_1$, and orthogonal basis vectors $\{U_\mathcal{N}^* \mathbf{e}_k\}_{k=2}^{d_\mathcal{N}}$.

Combining Subsystems: Blending

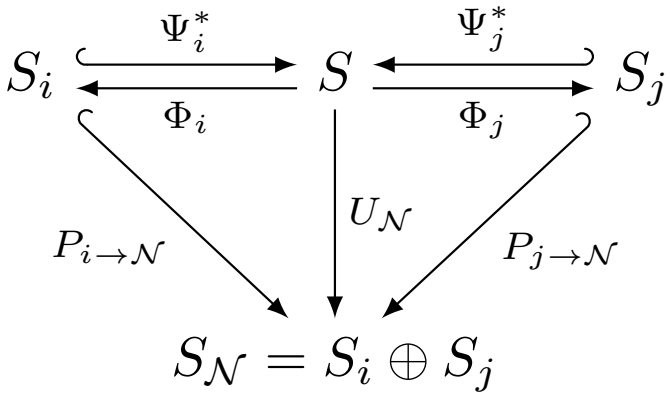


Diagram chasing yields:

$$P_{i\to\mathcal{N}} = U_{\mathcal{N}}\Psi_i^*$$
$$P_{\mathcal{N}\to i} = \Phi_i U_{\mathcal{N}}^*.$$

Global reduced system:

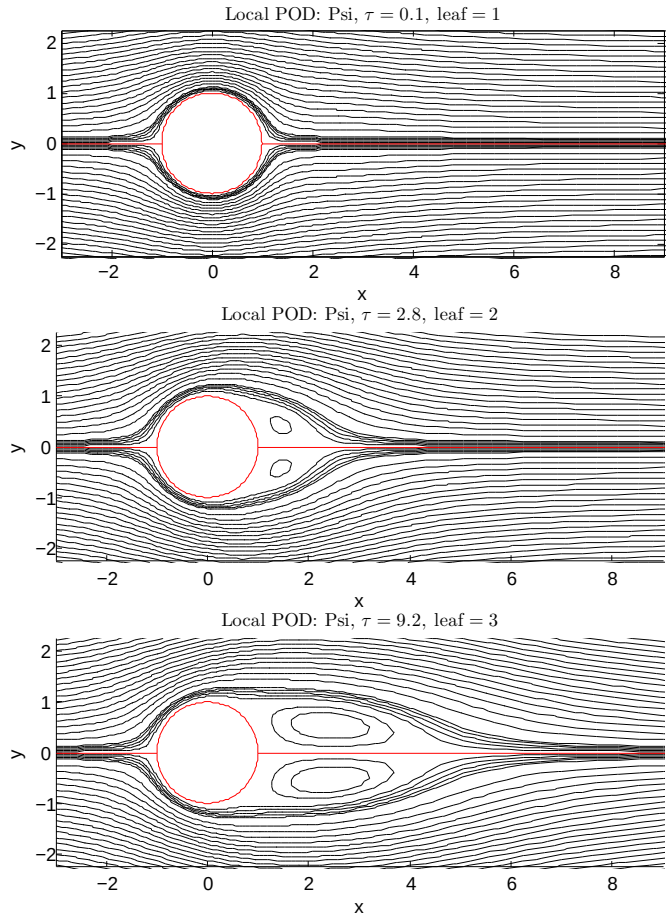
$$\begin{cases} \dot{\mathbf{x}}(t) = \sum_{i\in\mathcal{N}} w_i(\mathbf{x},t) P_{i\to\mathcal{N}} f_i(P_{\mathcal{N}\to i}\mathbf{x}) \\ \mathbf{y}(t) = \sum_{i\in\mathcal{N}} w_i(\mathbf{x},t) h_i(P_{\mathcal{N}\to i}\mathbf{x}) \end{cases} \tag{2}$$

where w_i satisfy $w_i \geq 0, \sum_{i\in\mathcal{N}} w_i = 1$ for all $\mathbf{x}, t$.

Demonstration: Fluid Simulation

- Stable ($R_e = 40$) 2-D flow past a stationary cylinder. 61×61 grid, D=3721.
- Mixture of implicit and explicit finite difference schemes to numerically integrate the unsteady Navier-Stokes equations in vorticity/stream-function form.
- Local POD: snapshots every 30 steps ($\delta t = 10^{-3}$), n=665 samples.
- Tree: $\epsilon = 10^{-3}$, 5 node tree, 3 leaf nodes of dimension $d = 3$.
- Comparison: Global POD, 5 modes.
- Reduced simulation: switching rule.

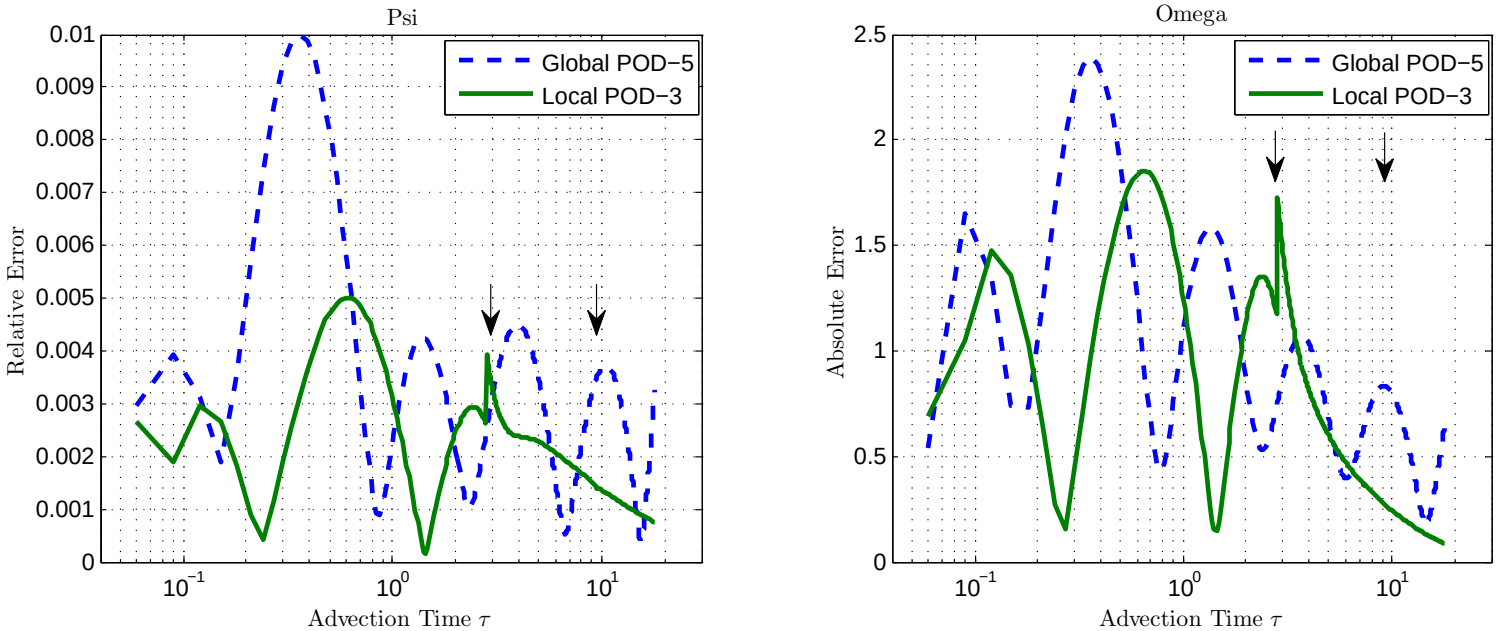
Demonstration: Fluid Simulation



Streamlines illustrating the local POD flow simulation near the beginning (top), after a transition from leaf node model 1 to 2 (middle), and after a transition from leaf node model 2 to 3 (bottom).

Demonstration: Fluid Simulation

Evolution of the simulation error for global POD (5 modes) vs. local POD (3 modes):



Vorticity ω error (right) and stream function ψ error (left). Arrows indicate where switching between models occurred in the local POD simulation.

Summary

- Snapshot-based framework for localized, multiscale reduction
- Statespace decomposed into a collection of planes determined by geometry of empirical trajectories.
- Defined local POD/balanced-POD projections $\Rightarrow$ multiscale data-dependent dictionary designed to efficiently encode behavior of a nonlinear system as a collection of simpler systems.
- Global reduced system described by combining local, low-dimensional systems.

Advantages:

- Can decompose complex phenomena into multi-resolution hierarchy of simple parts.
- Better approximation at lower complexity: approximating subspaces do not need to capture global behavior.
- Control with simple local controllers ; decentralized possibilities.

Open problem: approximate the local dynamics with simple, easy to evaluate maps.