# An Introduction to Model Reduction for Large-Scale Applications

Karen Willcox
Aerospace Computational Design Laboratory
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology

September 23, 2005

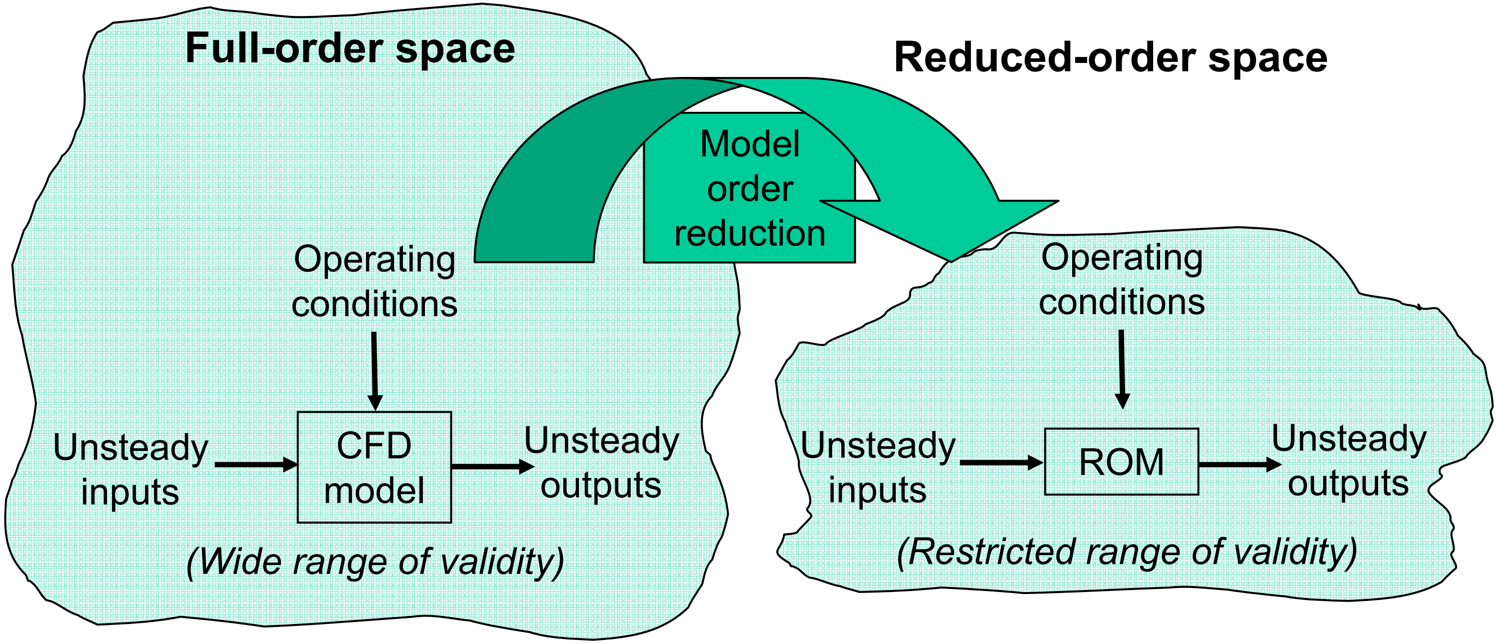
ACDL Seminar
MIT

# Outline

- Model reduction overview
- Projection framework
- CFD Examples
- Proper orthogonal decomposition

# Model Order Reduction

Systematic model order reduction to replicate high-fidelity dynamical system results over a restricted range of inputs.



**Full-order space**

**Reduced-order space**

Model order reduction

Operating conditions

Unsteady inputs → CFD model → Unsteady outputs

*(Wide range of validity)*

Operating conditions

Unsteady inputs → ROM → Unsteady outputs

*(Restricted range of validity)*

→ Reduction level is several orders of magnitude *e.g.* 2-D Euler : from ~$10^4$ states to ~$10^1$.

# *Why Model Reduction?*

When is the high-fidelity dynamical system model too expensive?

- Multidisciplinary applications
  - Aeroelasticity (fluid/structure)
  - Flow control
- Real-time applications
- Design and optimization
- Probabilistic applications

# Dynamical Systems

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u} \\ \mathbf{y} &= C\mathbf{x} \end{aligned}$$

$$\begin{aligned} \dot{\mathbf{x}} &= f(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} &= g(\mathbf{x}) \end{aligned}$$

$\mathbf{x} \in \mathbf{R}^n$: state vector

$\mathbf{u} \in \mathbf{R}^p$: input vector

$\mathbf{y} \in \mathbf{R}^q$: output vector

# CFD Dynamical Systems

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u} \\ \mathbf{y} &= C\mathbf{x} \end{aligned}$$

$$\begin{aligned} \dot{\mathbf{x}} &= f(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} &= g(\mathbf{x}) \end{aligned}$$

- $\mathbf{x}(t)$: vector of $n$ flow unknowns

  e.g. 2D Euler, $N$ grid points, $n = 4N$

  $$\mathbf{x} = [\rho_1 \; (\rho u)_1 \; (\rho v)_1 \; e_1 \; \rho_2 \cdots\cdots \rho_N \; (\rho u)_N \; (\rho v)_N \; e_N]^T$$

- $\mathbf{u}(t)$: inputs

  e.g. flow disturbances, wing motion

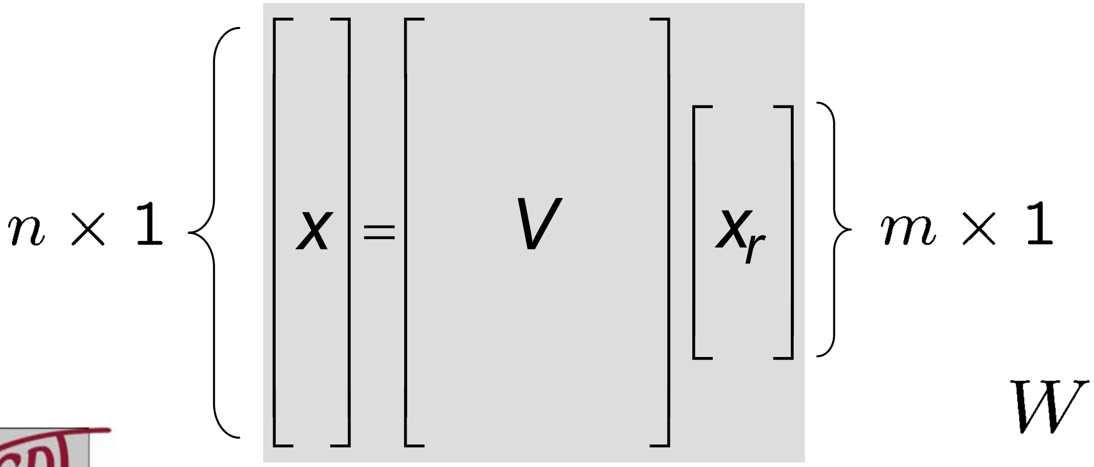
- $\mathbf{y}(t)$: outputs

  e.g. flow characteristic, lift force

# Reduced-Order Projection

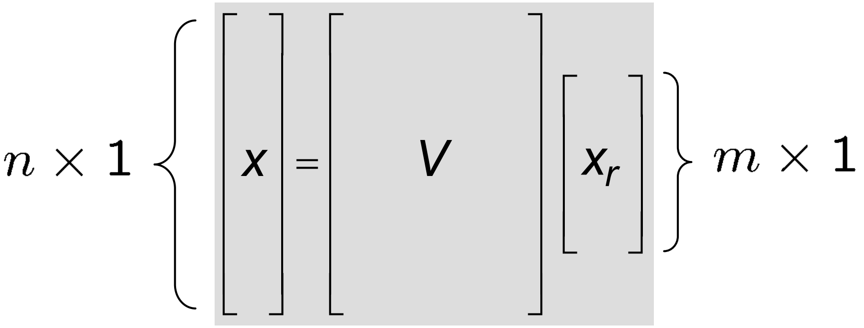
$$\mathbf{x}(t) = \sum_{i=1}^{m} V_i \alpha_i(t)$$

$$\mathbf{x}(t) = V\mathbf{x}_r(t)$$



$$W^T V = I$$

# Reduced-Order Dynamical Systems



$$W^T V = I$$

$$\begin{aligned} \dot{\mathbf{x}} &= f(\mathbf{x}, \mathbf{u}) \\ \mathbf{y} &= g(\mathbf{x}) \end{aligned} \xrightarrow{\mathbf{x} = V\mathbf{x}_r} \begin{aligned} \dot{\mathbf{x}}_r &= W^T f(V\mathbf{x}_r, \mathbf{u}) \\ \mathbf{y}_r &= g(V\mathbf{x}_r) \end{aligned}$$

$$\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u} \\ \mathbf{y} &= C\mathbf{x} \end{aligned} \xrightarrow{\mathbf{x} = V\mathbf{x}_r} \begin{aligned} \dot{\mathbf{x}}_r &= A_r\mathbf{x}_r + B_r\mathbf{u}_r \\ \mathbf{y}_r &= C_r\mathbf{x}_r \end{aligned}$$

$$A_r = W^T A V, \;\; B_r = W^T B, \;\; C_r = CV$$

# Reduced-Order Basis

- Determine the projection $x = Vx_r$ where *V* contains *m* basis vectors

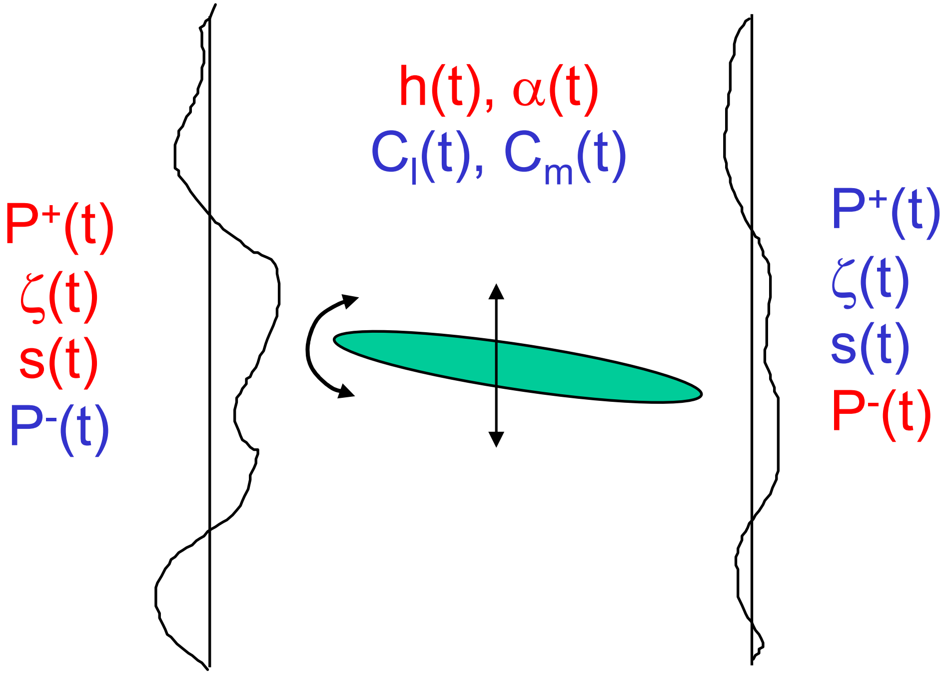
  $$V = \begin{bmatrix} v_1 & v_2 & \cdots & v_m \end{bmatrix}$$

  so that *m* << *n* and system dynamics are captured accurately: $y_r \approx y$
- Most, but not all, reduction techniques use projection framework (although all can be interpreted in a projection framework)

# *Large-Scale Reduction Methods*

- Proper orthogonal decomposition (POD)
  - Use data to generate empirical eigenfunctions
  - Time and frequency domain methods
- Krylov subspace methods
  - Arnoldi, Lanczos methods
- Balanced truncation
  - Recent work to extend to large-scale systems
  - Close connection between POD and balanced truncation
- Fourier model reduction

# Example: Aeroelastic Model

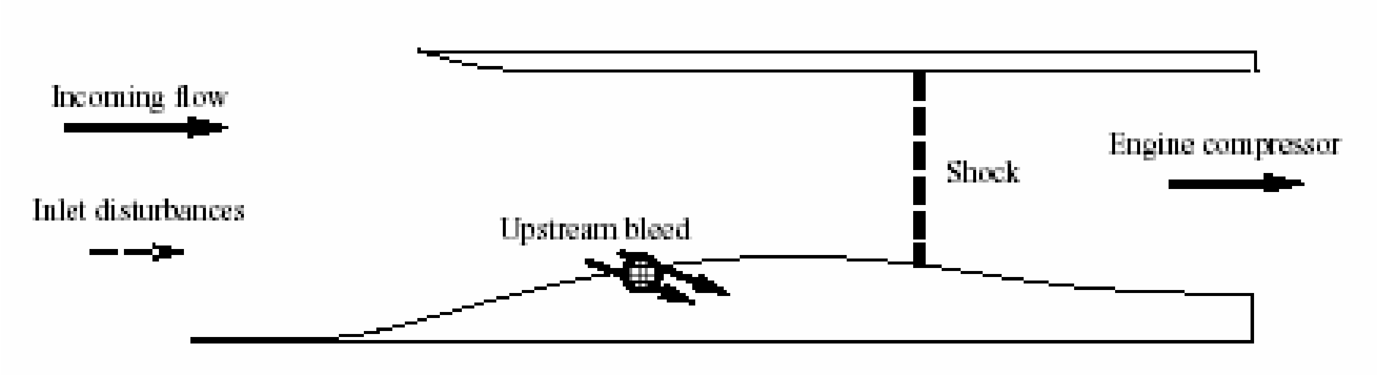


Inputs : $u = [h, \alpha, P^+_{-\infty}, \zeta_{-\infty}, s_{-\infty}, P^-_{+\infty}]$

Outputs : $y = [C_l, C_m, P^+_{+\infty}, \zeta_{+\infty}, s_{+\infty}, P^-_{-\infty}]$

# *Example: Supersonic Diffuser Active Flow Control*

- Supersonic diffuser: freestream Mach = 2.2
- Control the position of the shock in response to incoming disturbances using upstream bleed



- Inputs: flow disturbance, bleed actuation
- Outputs: average Mach number at throat

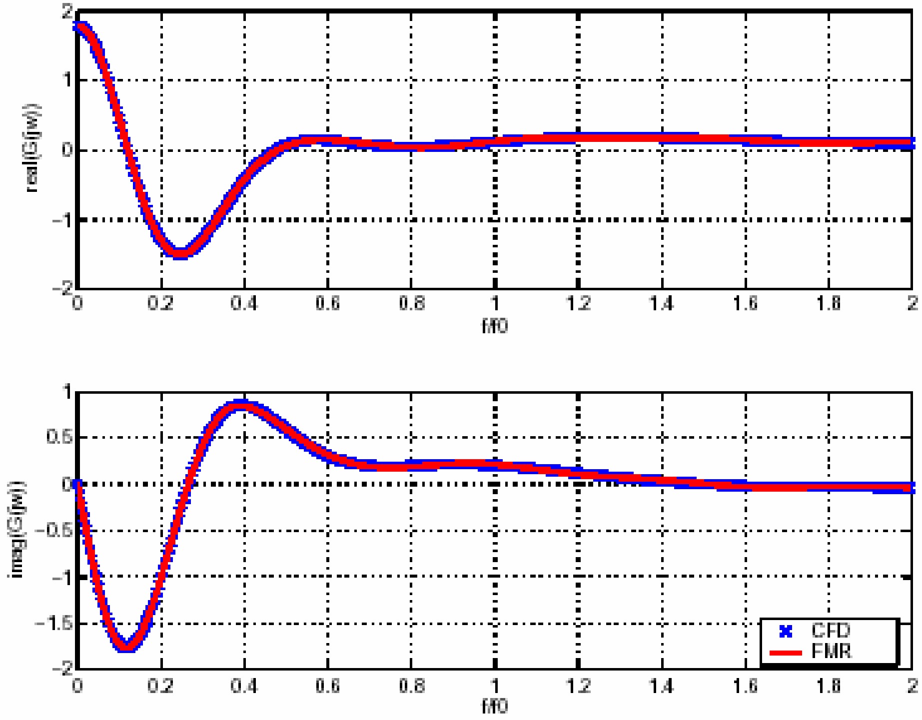
# *Supersonic Diffuser CFD*

Steady flow Mach contours, inflow Mach 2.2

- CFD model has 11,730 states
- Complicated flow dynamics
  - Disturbance transfer function describes a delay
- Reduced-order models yield very accurate results with O(10) states
  - Over frequency range of interest

# *Supersonic Diffuser: Bleed Actuation*

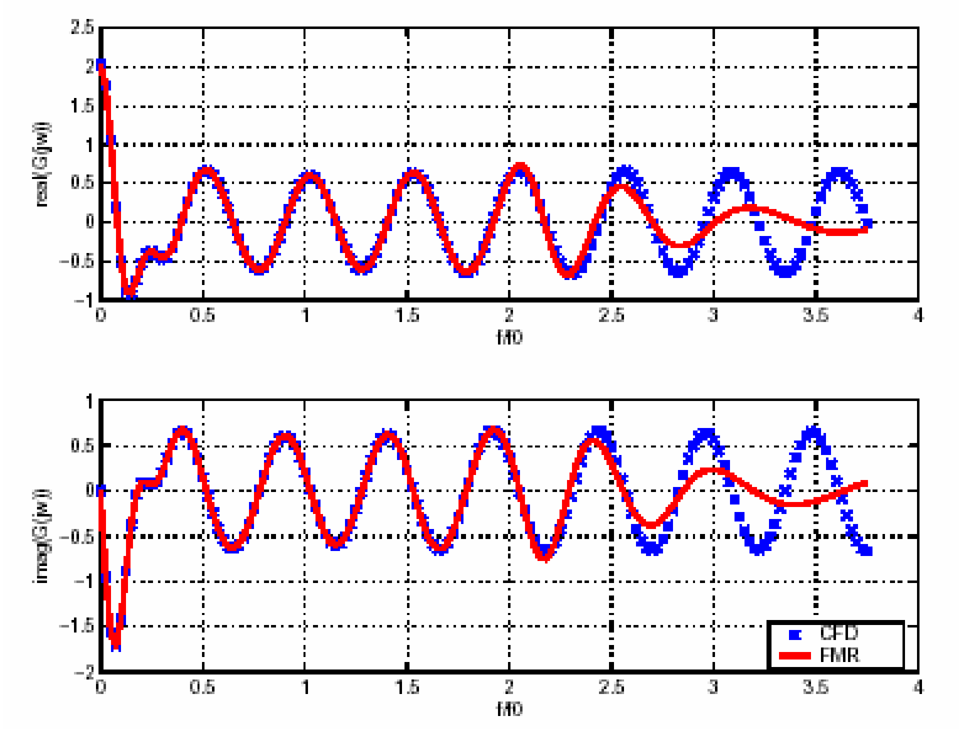
Transfer function: bleed actuation to average throat Mach number.



x CFD *n*=11,730 – FMR *k=20*

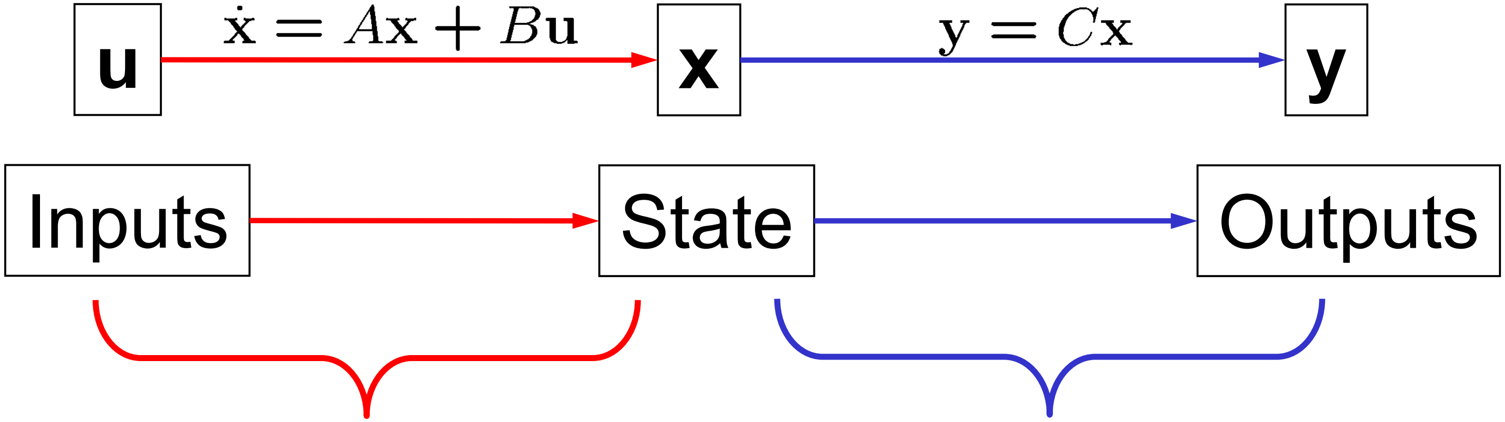
# Supersonic Diffuser: Disturbance Dynamics

Transfer function: incoming density disturbance to average throat Mach number.



**x** CFD *n*=11,730 **–** FMR *k=20*

# *Which States Are Important?*



$$\begin{tikzcd}
\mathbf{u} \arrow[r, "\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}"] & \mathbf{x} \arrow[r, "\mathbf{y} = C\mathbf{x}"] & \mathbf{y} \\
\text{Inputs} \arrow[r] & \text{State} \arrow[r] & \text{Outputs}
\end{tikzcd}$$

“Controllable” modes
- easy to reach, require small control energy

“Observable” modes
- generate large output energy

# *How Many States Are Needed?*

Hankel matrix: relates past inputs to future outputs

$$\mathbf{y} = \Gamma \mathbf{u}$$

$$\begin{bmatrix} \vdots \\ y(-2) \\ y(-1) \\ y(0) \\ y(1) \\ y(2) \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & & & & & & \\ \cdots & g_0 & & & & & \\ \cdots & g_1 & g_0 & & & & \\ \cdots & g_2 & g_1 & g_0 & & & \\ \cdots & g_3 & g_2 & g_1 & g_0 & & \\ \cdots & g_4 & g_3 & g_2 & g_1 & g_0 & \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ u(-2) \\ u(-1) \\ u(0) \\ u(1) \\ u(2) \\ \vdots \end{bmatrix}$$

Singular values of $\Gamma$: Hankel singular values, $\sigma_j$

— describe effectiveness of inputs translating to outputs

— corresponding mode: controllability and observability

# *Proper Orthogonal Decomposition*

- Consider *M* snapshots $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M \in \mathcal{R}^n$ (instantaneous state solutions)
- Construct kernel $K \in \mathcal{R}^{n \times n}$

$$\boxed{K = \sum_{j=1}^{M} \mathbf{x}_j \mathbf{x}_j^T}$$

*n×n eigenvalue problem*

- $V = [\mathbf{v}_1 \ \mathbf{v}_2 \ \mathbf{v}_3 \cdots]$ are eigenvectors of *K* with $\lambda_1 > \lambda_2 > \lambda_3 > \cdots$
- If $V_m$ contains the first $m$ eigenvectors, then $Q_m = V_m V_m^T$ is the optimal projection in a least squares sense:

$$\boxed{\min_{Q_m} \sum_{i=1}^{M} ||\mathbf{x}_i - Q_m \mathbf{x}_i||_2^2 = \sum_{i=m+1}^{M} \lambda_i}$$

  - Note: optimality and error bound applies to the reconstruction of sampled data, not to the ROM.

# Method of Snapshots

Sirovich: the eigenvectors of the kernel are linear combinations of the snapshots:

$$\boxed{\mathbf{v}_i = \sum_{j=1}^{M} \alpha_j^i \mathbf{x}_j}$$

The eigenvalue problem becomes:

$$\boxed{R\alpha^i = \lambda_i \alpha^i} \qquad \alpha^i = \begin{bmatrix} \alpha_1^i \\ \alpha_2^i \\ \vdots \end{bmatrix}$$

$M \times M$ *eigenvalue problem*

where $R$ is the $M \times M$ correlation matrix:

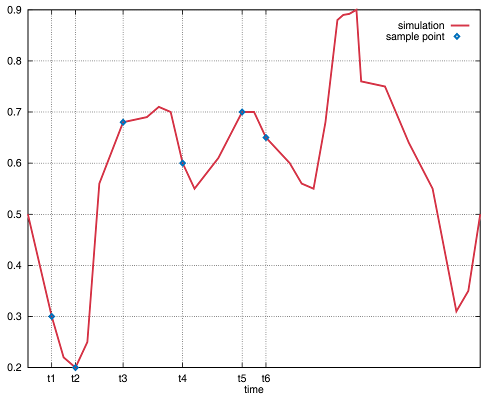
$$\boxed{R_{ij} = \mathbf{x}_i^T \mathbf{x}_j}$$

# Reduction via POD

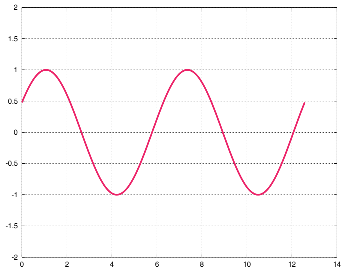
1. Simulate the high-order system to get $M$ snapshots $\mathbf{x}_j$, $j$=1, 2, …, $M$ (could be for different parameters)
2. Construct the correlation matrix $R_{ij} = \mathbf{x}_i^T \mathbf{x}_j$
3. Calculate the eigenvectors $\alpha^i$ and eigenvalues $\lambda_i$ of $R$
4. Construct the basis functions $$\mathbf{v}_i = \sum_{j=1}^{M} \alpha_j^i \mathbf{x}_j$$
5. Select the most energetic basis functions using $\lambda_i$
6. Project the governing equations onto the reduced basis

# *Frequency Domain*

For linear systems, response can be decomposed into temporal harmonics:



$$= \sum$$



For $u = \bar{u}e^{j\omega t}$, $X = \bar{X}e^{i\omega t}$ and $y = \bar{y}e^{i\omega t}$

$$\boxed{\begin{array}{rcl} \bar{X} & = & [j\omega I - A]^{-1} B\bar{u} \\ \bar{y} & = & C\bar{X} \end{array}}$$

# Frequency Domain POD

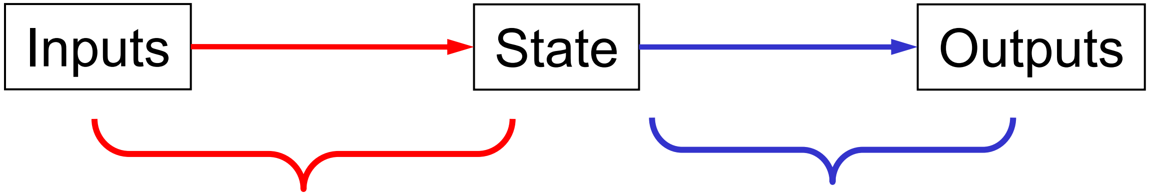
- Difficulty in choosing appropriate time simulation to obtain snapshots
- Instead, pick a set of sample frequencies $\omega_k$
- Solve frequency domain equations at each frequency to obtain complex snapshots

$$\bar{X}(\omega_k) = [j\omega_k I - A]^{-1} B$$

$$K = \frac{1}{M} \sum_{k=1}^{M} \bar{X}(\omega_k) \bar{X}^*(\omega_k)$$

# *A Different View of POD*

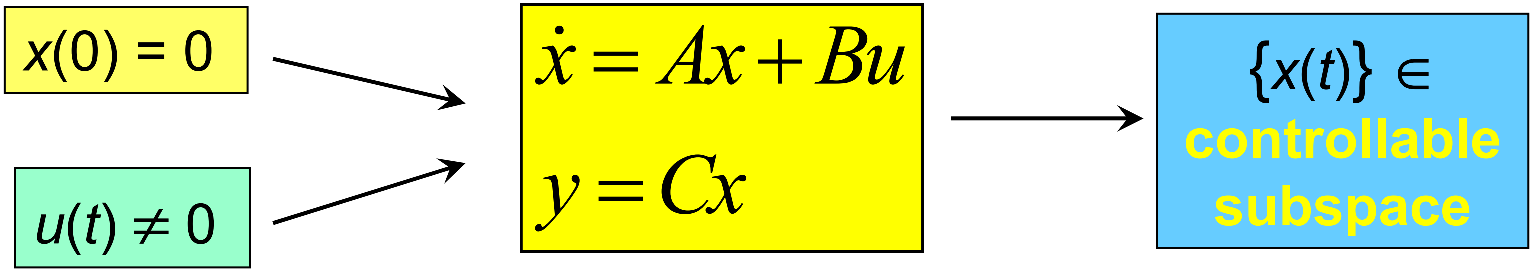
- POD often thought of in least squares context
  - POD basis vector provides least squares fit to snapshot data
- **Balanced truncation** provides different insight
  - Concepts of **controllability** and **observability**



“Controllable” modes
- easy to reach, require small control energy

“Observable” modes
- generate large output energy

# Subspaces



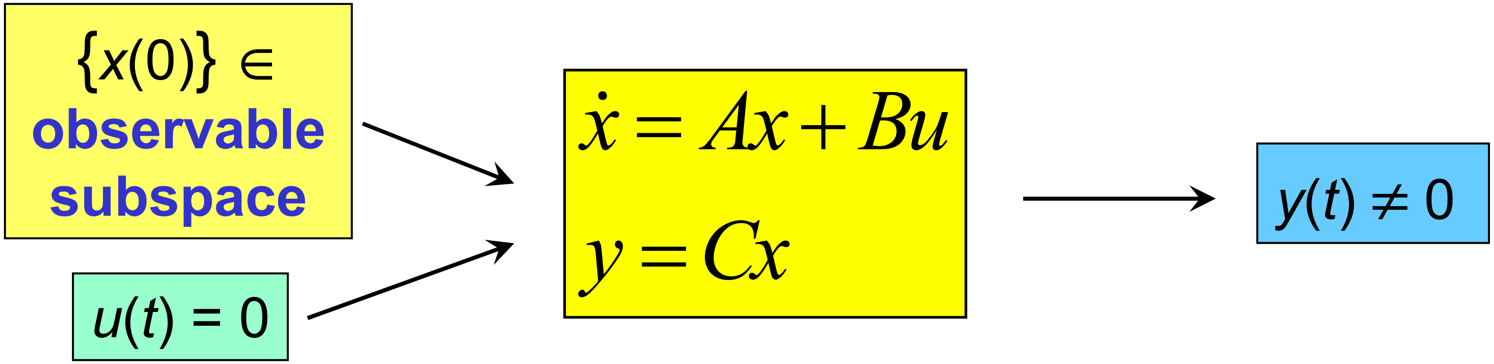
$x(0) = 0$

$u(t) \neq 0$

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

$\{x(t)\} \in$ **controllable subspace**

The ***controllable subspace*** is that set of states that can be obtained with zero initial state and a given input (reachable states).



$\{x(0)\} \in$ **observable subspace**

$u(t) = 0$

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

$y(t) \neq 0$

The ***observable subspace*** is that set of states which as initial conditions produce a non-zero output with no external input.

# Gramians

Controllability Gramian:

$$W_c = \int_0^\infty \underbrace{e^{At}B}_{x_\delta(t)} B^T e^{A^T t}\, dt$$

eigenvectors of $W_c$ span the controllable subspace

$e^{At}B$ is the impulse response

Observability Gramian:

$$W_o = \int_0^\infty \underbrace{e^{A^T t}C^T}_{Z_\delta(t)} C e^{At}\, dt$$

eigenvectors of $W_o$ span the observable subspace

$e^{A^T t}C^T$ is the dual impulse response

# Dual System

Linearized model (state vector *x*)

$$\dot{x} = Ax + Bu$$

$$y = Cx$$

Dual system (dual state vector *Z*)

$$\dot{Z} = A^T Z + C^T u_d$$

$$y_d = B^T Z$$

# *Balanced Realization*

The linear system is said to be **balanced** when the **gramians** are **diagonal and equal**.

The balancing transformation $T$ is given by the eigenvectors of the **gramian product**:

$$W_{co} = W_c W_o = T^{-1} \Lambda T$$

$\lambda_i$ are positive, real

$\sigma_i = \sqrt{\lambda_i}$ are the **Hankel singular values**

$W_{co}$

**Eigenvector $T_i$**

**- state through which input is transmitted to output**

**Eigenvalue $\lambda_i$**

**- relative importance of state**

**- independent of realization**

**Balanced truncation**: keep only states with large HSVs.

# Frequency Domain Gramians

Controllability Gramian:

$$W_c = \frac{1}{2\pi}\int_{-\infty}^{\infty} \underbrace{(j\omega I - A)^{-1}B}_{\bar{X}_\omega} B^T \left(-j\omega I - A^T\right)^{-1} d\omega$$

*(jωI-A)*$^{-1}$*B* is the response to sinusoidal forcing

Observability Gramian:

$$W_o = \frac{1}{2\pi}\int_{-\infty}^{\infty} \underbrace{\left(j\omega I - A^T\right)^{-1}C^T}_{\bar{Z}_\omega} C(-j\omega I - A)^{-1} d\omega$$

*(jωI-A*$^T$*)*$^{-1}$*C*$^T$ is the dual response to sinusoidal forcing

# POD Versus Balanced Truncation

POD Kernel:

$$K = \frac{1}{M}\sum_{j=1}^{M}\left(j\omega_j I - A\right)^{-1} BB^{*}\left(-j\omega_j I - A^{*}\right)^{-1}$$

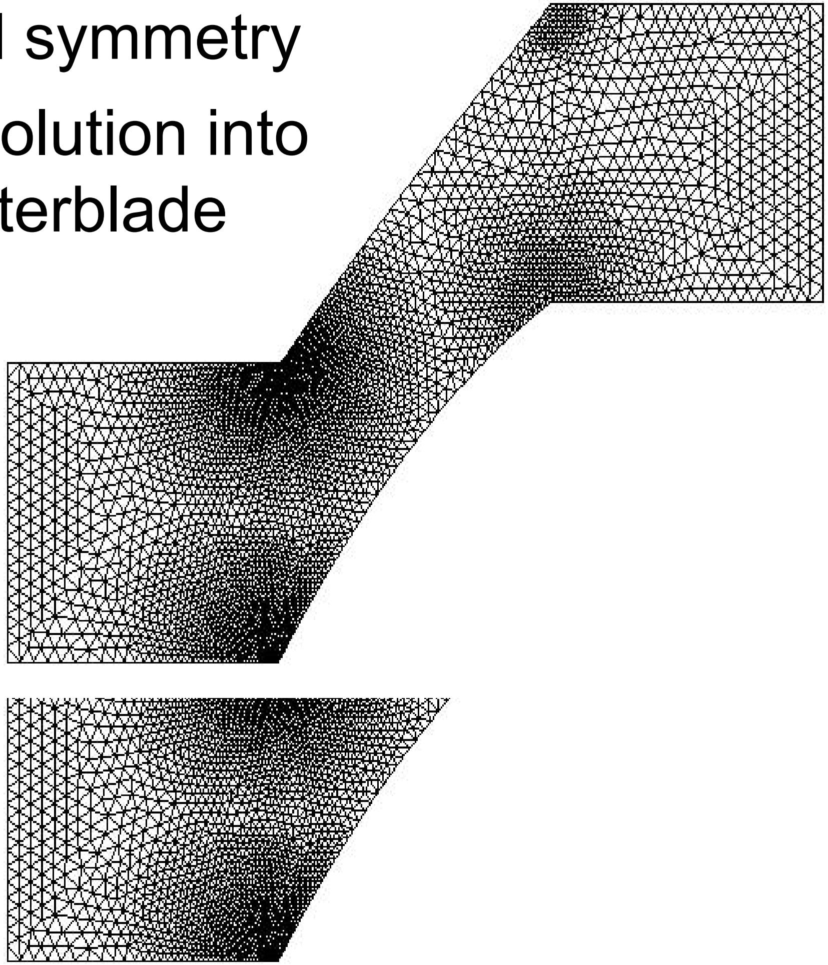
Controllability Gramian:

$$W_c = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(j\omega I - A\right)^{-1} BB^{*}\left(-j\omega I - A^{*}\right)^{-1} d\omega$$

POD is an approximation to the controllability Gramian with rectangular rule.

# Turbomachinery Applications

- Problems have spatial symmetry
- Decompose general solution into spatial harmonics – interblade phase angles
- Solve efficiently in frequency domain on a single passage



# First Standard Configuration

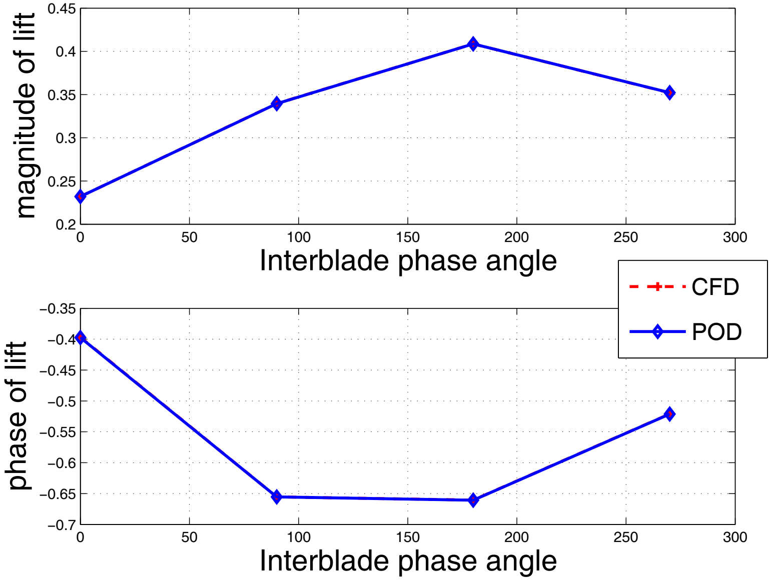
- Test case:
  - First standard configuration
  - Inflow Mach number $M = 0.18$
  - Inflow angle $\beta = 62^\circ$
  - Unsteady rigid plunging motion
    - Inputs: displacements ($h$), velocities ($\dot{h}$) for each blade
    - Output: lift coefficients ($C_l$) for each blade
  - 76,000 states per blade passage
- BPOD
  - 51 equally spaced frequencies in $\omega = [0,4]$
  - Gramians numerically integrated with Simpson’s rule

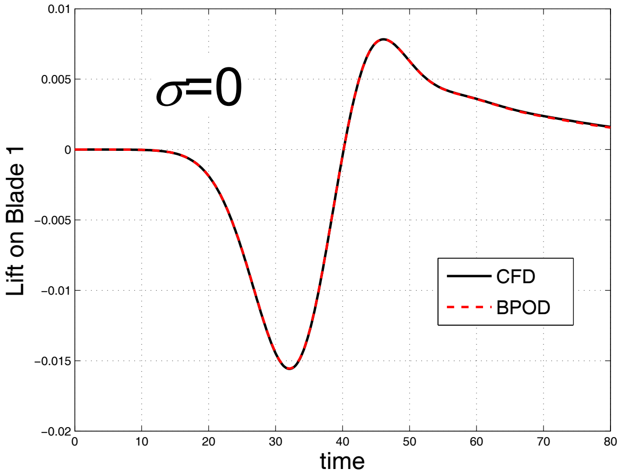
# *BPOD Results*

(T. Bui-Thanh, MIT)

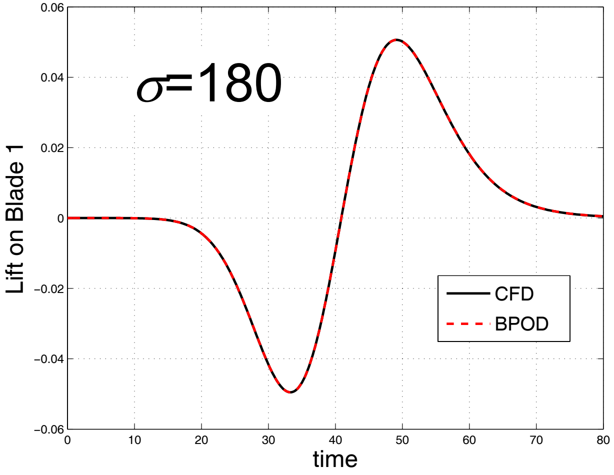
4 blades
CFD: 302,000 states
ROM: 92 states
23 states for each IBP

Unsteady lift for harmonic plunge at $\omega = 0.85$

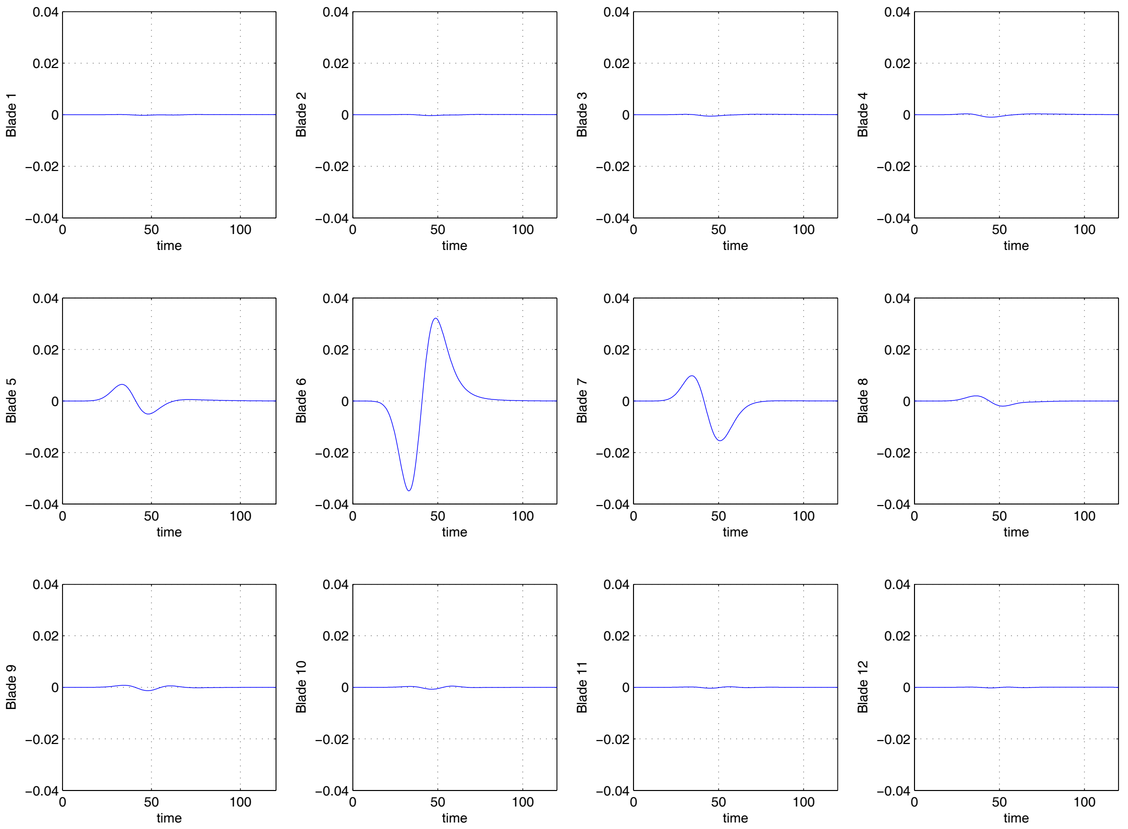




Unsteady lift for pulse input in plunge



# 12 blades: reduce 912,000 states to 276 states: POD



# Summary

- Recent advances in model reduction of large-scale systems
  - Success in many different applications
- Projection framework underlies most large-scale reduction methods
- Several challenges remain:
  - Reduction of nonlinear systems
  - Reduction of parametrically varying systems
  - Rigorous guarantees in the large-scale setting