Model Reduction for Large-Scale Applications in Probabilistic Analysis and Inverse Problems

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

POD Workshop
Bordeaux, France
April 1, 2008

Acknowledgements

- Omar Bashir, Tan Bui-Thanh, Krzysztof Fidkowski, David Galbally, Chad Lieberman (MIT)
- George Biros (University of Pennsylvania)
- Omar Ghattas (University of Texas at Austin)
- Matthias Heinkenschloss, Dan Sorensen (Rice University)
- Bart van Bloemen Waanders, Judy Hill (Sandia National Labs)

- Funding: AFRL (Dr. Cross), AFOSR (Dr. Fahroo), NSF (Dr. Darema), Singapore-MIT Alliance, Sandia National Laboratories (Computer Science Research Institute)

Outline

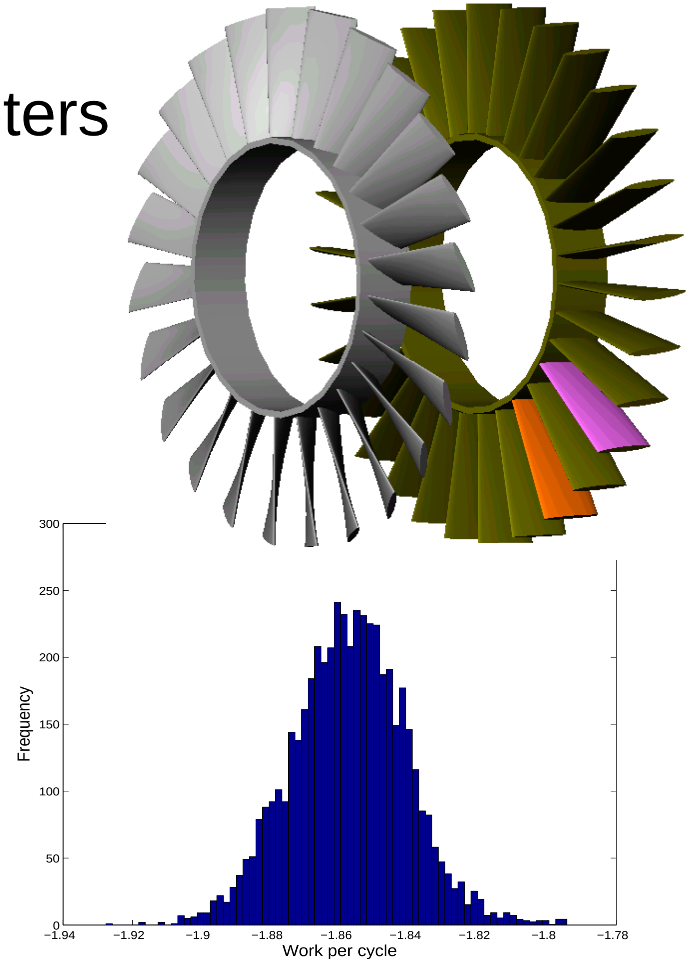
- Model reduction for optimization and probabilistic applications
  - A model-constrained optimization framework
- Efficient nonlinear model reduction

- Examples
  - Probabilistic analysis of compressor unsteady aerodynamics
  - Real-time inverse problems for contaminant transport
  - Bayesian inference of combustion parameters

Why Model Reduction?

Key challenge: probabilistic analyses of large-scale (e.g. CFD) models

- Compressor blade mistuning: small variations in blade structural parameters and blade shape can have a large impact on blade row performance
- <u>2D</u> CFD model unsteady analysis for <u>two</u> blade passages:
~3 minutes per geometry
- Monte Carlo simulation:
5000 samples ≈10 days
50,000 samples ≈ 3.5 months



Parameterized Dynamical Systems

$$\boxed{\begin{array}{rcl} \dot{\mathbf{x}} & = & A(\mathbf{z})\mathbf{x} + B(\mathbf{z})\mathbf{u} \\ \mathbf{y} & = & C(\mathbf{z})\mathbf{x} \end{array}}$$

$$\boxed{\begin{array}{rcl} \dot{\mathbf{x}} & = & f(\mathbf{x}, \mathbf{z}, \mathbf{u}) \\ \mathbf{y} & = & g(\mathbf{x}, \mathbf{z}, \mathbf{u}) \end{array}}$$

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

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

$\mathbf{z} \in \mathbf{R}^r$: parameter vector

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

Example: CFD Systems

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

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

- $\mathbf{x}(t)$: vector of n flow unknowns
  e.g. 2D Euler, N grid points, $n = 4N$
  $\mathbf{x} = [\rho_1 \ (\rho u)_1 \ (\rho v)_1 \ e_1 \ \rho_2 \cdots\cdots \rho_N \ (\rho u)_N \ (\rho v)_N \ e_N]^T$
- $\mathbf{z}$: input parameters
  e.g. shape parameters, PDE coefficients
- $\mathbf{u}(t)$: forcing inputs
  e.g. flow disturbances, wing motion
- $\mathbf{y}(t)$: outputs
  e.g. flow characteristic, lift force

Reduced-Order Dynamical Systems

$$\underset{n \times 1}{\begin{bmatrix} \\ \\ \mathbf{x} \\ \\ \\ \end{bmatrix}} = \begin{bmatrix} \\ \\ & V & \\ \\ \\ \end{bmatrix} \underset{m \times 1}{\begin{bmatrix} \\ \mathbf{x}_r \\ \\ \end{bmatrix}} \qquad W^T V = I$$

$$\boxed{\begin{aligned} \dot{\mathbf{x}} &= A\mathbf{x} + B\mathbf{u} \\ \mathbf{y} &= C\mathbf{x} \end{aligned}} \xrightarrow{\mathbf{x} \approx V\mathbf{x}_r} \boxed{\begin{aligned} \mathbf{r} &= V\dot{\mathbf{x}}_r - AV\mathbf{x}_r - B\mathbf{u} \\ \mathbf{y}_r &= CV\mathbf{x}_r \end{aligned}}$$

$$\Big\downarrow \; W^T\mathbf{r} = 0$$

$$\begin{aligned} A_r &= W^T A V \\ B_r &= W^T B \\ C_r &= C V \end{aligned} \qquad \boxed{\begin{aligned} \dot{\mathbf{x}}_r &= A_r\mathbf{x}_r + B_r\mathbf{u} \\ \mathbf{y}_r &= C_r\mathbf{x}_r \end{aligned}}$$

Large-Scale Reduction Methods

- Determine the projection $\mathbf{x} = \mathbf{V}\mathbf{x}_r$ where V contains *m* basis vectors

  $$\mathbf{V} = [\mathbf{V}_1 \quad \mathbf{V}_2 \quad \cdots \quad \mathbf{V}_m]$$

  so that *m* << *n* and system dynamics are captured accurately: $\mathbf{y}_r \approx \mathbf{y}$

- Many model reduction methods for large-scale systems
  - Krylov-based methods, proper orthogonal decomposition, balanced truncation, reduced basis, modal analysis, Fourier model reduction, optimization approaches, et al.

- Methodology “mature” for linear time-invariant systems with few inputs/few outputs
  - Key challenges: nonlinear systems, parametric inputs, sampling, rigorous error guarantees
  - Model reduction for optimization versus just simulation

Model Reduction for Optimization and Probabilistic Applications

- Sampling of input space (parameters and time) to determine the basis is a critical question for model reduction application to design, optimization, inverse problems, probabilistic analysis

- Krylov-based methods (LTI systems): Iterative Rational Krylov Algorithm *(Gugercin, Antoulas & Beattie, 2006)*

- POD basis for optimal control: Adaptive methods *(Afanasiev & Hinze, 1999)*, TRPOD *(Arian, Fahl & Sachs, 2000)*, OS-POD *(Kunisch & Volkwein, 2006)*

- General reduced basis construction: Greedy algorithm *(Veroy, Prud'homme, Rovas & Patera, 2003; Grepl & Patera, 2005)*
  - Adaptive heuristic to choose “good” sample points
  - Sample the location in parameter space of maximum error between full and reduced-order model outputs

Sampling: Model-Constrained Optimization

- Formulate the task of finding the POD snapshot sample points as a model-constrained optimization problem
  - Adaptive sampling (sequence of optimization problems)
  - Update POD basis after each greedy cycle
- Linear problem with initial-condition parameters: explicit solution via eigenvalue problem *(Bashir, Willcox, Ghattas, van Bloemen Waanders, Hill, 2007)*

$$
\begin{array}{rcl}
\max\limits_{z} & & \dfrac{1}{2}\displaystyle\int_0^T \|y - y_r\|_2^2 dt \\
\text{subject to} & & \\
\dot{x} & = & A(z)x + B(z)u \\
x(0) & = & x^0 \\
y & = & C(z)x \\
\dot{x}_r & = & V^T A(z)Vx + V^T B(z)u \\
x_r(0) & = & x_r^0 \\
y_r & = & C(z)Vx_r \\
z_l \leq & z & \leq z_u
\end{array}
$$

- Nonlinear parametric dependence: solve with tailored PDE-constrained optimization algorithm *(Bui-Thanh, Willcox, Ghattas, 2008)*

Model-Constrained Optimization

Nonlinear Parametric Dependence

$$
\begin{array}{rcl}
\max\limits_{z} & & \dfrac{1}{2}\displaystyle\int_0^T \|y - y_r\|_2^2 dt \\
\text{subject to} & & \\
\dot{x} & = & A(z)x + B(z)u \\
x(0) & = & x^0 \\
y & = & C(z)x \\
\dot{x}_r & = & V^T A(z)Vx + V^T B(z)u \\
x_r(0) & = & x_r^0 \\
y_r & = & C(z)Vx_r \\
z_l \leq & z & \leq z_u
\end{array}
$$

Model-Constrained Optimization

Nonlinear Parametric Dependence

$$\max_{z} \quad \frac{1}{2}\int_0^T \|y - y_r\|_2^2 dt$$

subject to

$$\begin{aligned} \dot{x} &= A(z)x + B(z)u \\ x(0) &= x^0 \end{aligned}$$

Objective function targets error in current reduced-order model outputs over the parameter space

x

$(z)Vx + V^T B(z)u$

$$\begin{aligned} y_r &= C(z)Vx_r \\ z_l \leq & \; z \leq z_u \end{aligned}$$

Model-Constrained Optimization

Nonlinear Parametric Dependence

$$
\begin{array}{rcl}
\max\limits_{z} & & \frac{1}{2}\int_0^T \|y - y_r\|_2^2 dt \\
\text{subject to} & & \\
\dot{x} & = & A(z)x + B(z)u \\
x(0) & = & x^0 \\
y & = & C(z)x \\
\dot{x}_r & = & V^T A(z)Vx + V^T B(z)u \\
x_r(0) & = & x_r^0 \\
y_r & = & C(z) \\
z_l \leq & z & \leq z_u
\end{array}
$$

Large-scale governing equations as constraints to define *y*(*t*)

Model-Constrained Optimization

Nonlinear Parametric Dependence

$$\max_{z} \quad \frac{1}{2}\int_0^T \|y - y_r\|_2^2 dt$$

subject to

$$\begin{aligned}
\dot{x} &= A(z) \\
x(0) &= x^0 \\
y &= C(z)x
\end{aligned}$$

$$\boxed{\begin{aligned}
\dot{x}_r &= V^T A(z) V x + V^T B(z) u \\
x_r(0) &= x_r^0 \\
y_r &= C(z) V x_r
\end{aligned}}$$

$$z_l \leq \quad z \quad \leq z_u$$

Reduced-order model as constraints to define $y_r(t)$

Model-Constrained Optimization

Nonlinear Parametric Dependence

$$
\begin{aligned}
\max_{z} \quad & \frac{1}{2}\int_0^T \|y - y_r\|_2^2 dt \\
\text{subject to} \quad & \\
\dot{x} &= A(z)x + B(z)u \\
x(0) &= x^0 \\
y &= C(z)x \\
\dot{x}_r &= V^T A(z \\
x_r(0) &= x_r^0 \\
y_r &= C(z)Vx_r \\
z_l \leq \quad z \quad & \leq z_u
\end{aligned}
$$

Bounds on parameters

Model-Constrained Optimization

Nonlinear Parametric Dependence

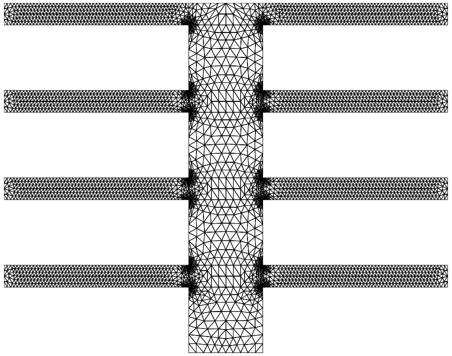
$$1 \int^{T}$$

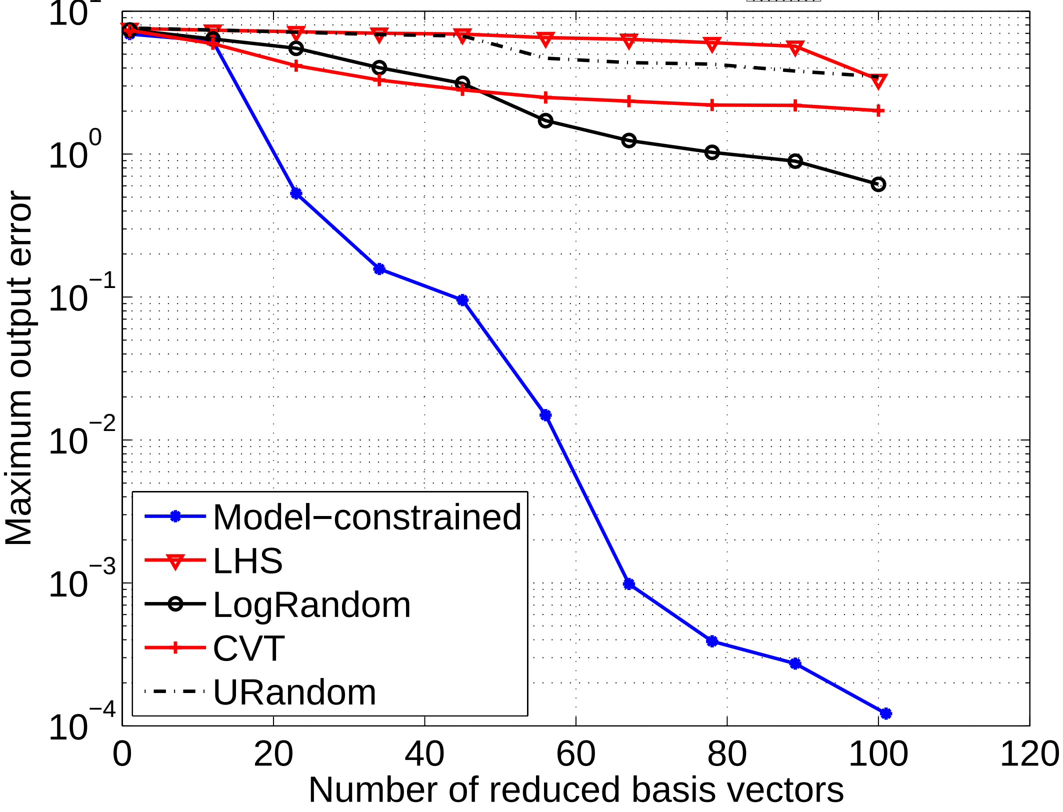
- Solve efficiently using state-of-the-art PDE-constrained optimization methodology (interior-reflective trust-region inexact-Newton conjugate-gradient method)
- For reduced model of fixed size, computational cost scales ~linearly with the dimension of the parameter space
- Can replace actual error with error estimator or error indicator (e.g. residual) to eliminate full-scale system constraints

$$z_l \leq z \leq z_u$$

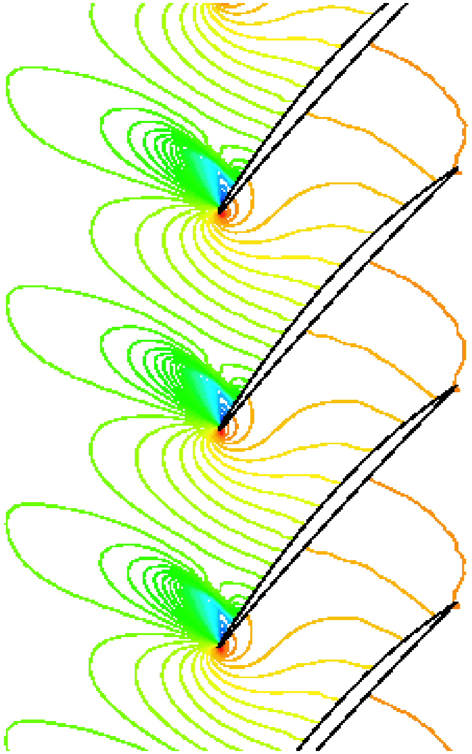
Steady Parametric Problem: Thermal Fin Design

- Model-constrained optimization sampling approach for parametric input spaces
  - Application to thermal fin design problem with 21 parameter input space
  - Finite element model: 17,899 states
  - Reduced model has 3-4 orders magnitude lower error compared to Latin hypercube, log-random, and other sampling methods





Unsteady Parametric Problem: Blade Shape Variations



$$\frac{\partial w}{\partial t} + \nabla \cdot \mathcal{F}(w) = 0$$

$$w = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ \rho E \end{pmatrix} \quad \mathcal{F}^x = \begin{pmatrix} \rho u \\ \rho u^2 + P \\ \rho u v \\ \rho u H \end{pmatrix}, \quad \mathcal{F}^y = \begin{pmatrix} \rho v \\ \rho u v \\ \rho v^2 + P \\ \rho v H \end{pmatrix}$$

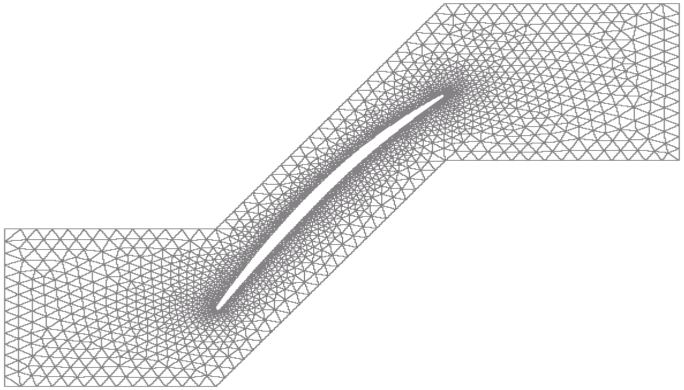
- Subsonic rotor blade, Mach 0.113
- 2D linearized Euler equations, DG CFD model
- 51,504 states per blade passage
- Small variations in blade structural parameters and blade shape can have large impact on blade row performance
- Inputs: blade plunging motion, blade shape parameters
- Output: blade lift forces

Goal: create a reduced-order model that captures input/output mapping between plunging motion input and lift force output *over a range of blade geometries*.

Unsteady Parametric Problem: Blade Shape Variations

- Blade geometric variations parameterized by a set of geometric modes, w_i

$$g = g_0 + \sum_{i=1}^{ns} \sigma_i z_i w_i$$



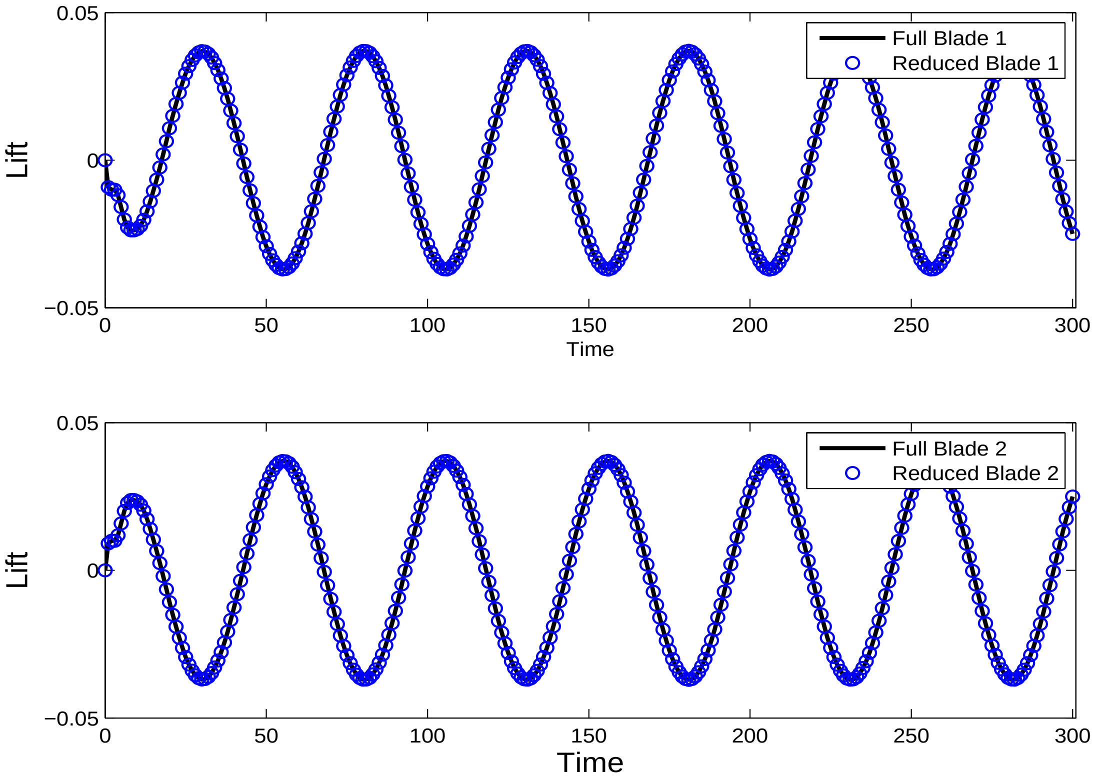
- Described by ns geometry parameters z_i

- With some assumptions, CFD model dependence on geometry can be written

$$\begin{aligned} \dot{x} &= \left(A_0 + \sum_{i=1}^{ns} A_i z_i\right) x + \left(B_0 + \sum_{i=1}^{ns} B_i z_i\right) u \\ y &= \left(C_0 + \sum_{i=1}^{ns} C_i z_i\right) x \end{aligned}$$

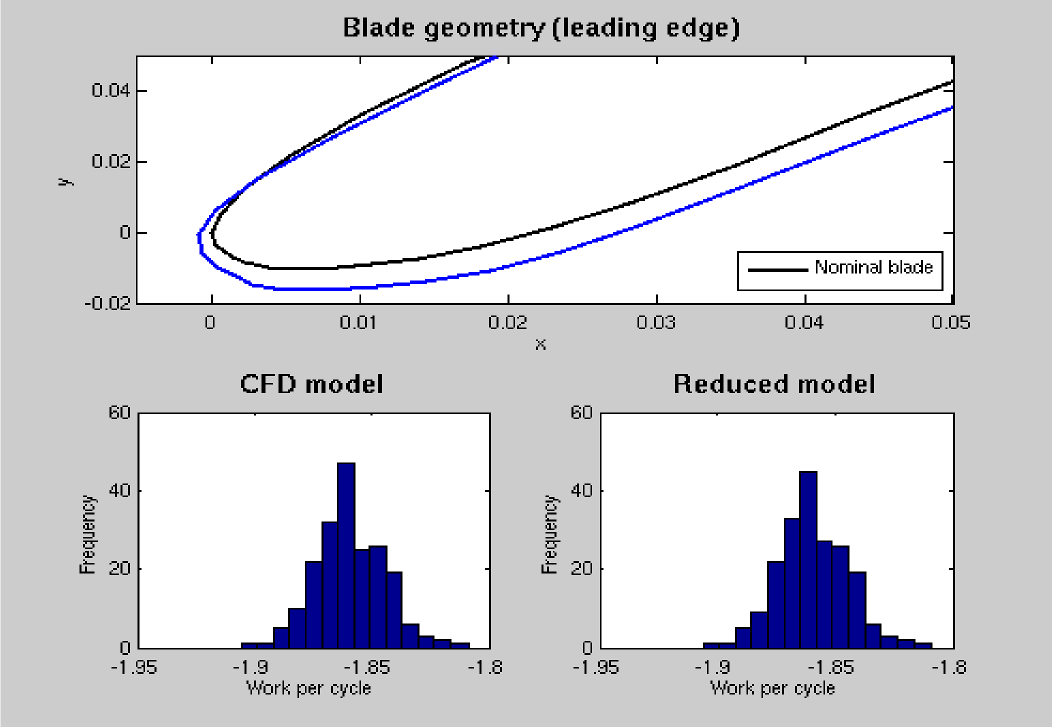
Unsteady Plunging Motion

- Forced response of two blade passages to sinusoidal plunging motion (180° interblade phase angle)
- Full model: *n*=103,008
- Reduced model: *m*=201, but offline cost ~3 hours



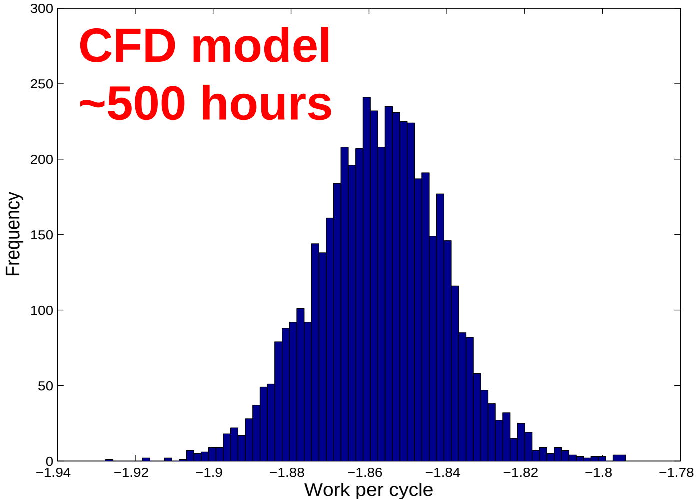
Monte Carlo Simulations

- Parameterized reduced model used to evaluate unsteady response over a range of geometry variations
- Full model: n=103,008; ~3 mins per geometry
- Reduced model: m=201, <0.1 secs per geometry



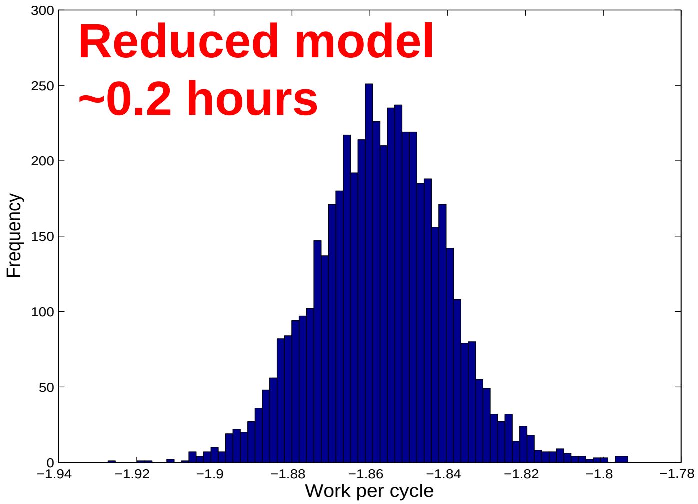
Monte Carlo Simulation Results

- Work per cycle for plunging motion as blade geometry varies
- Same 10,000 random samples in each case



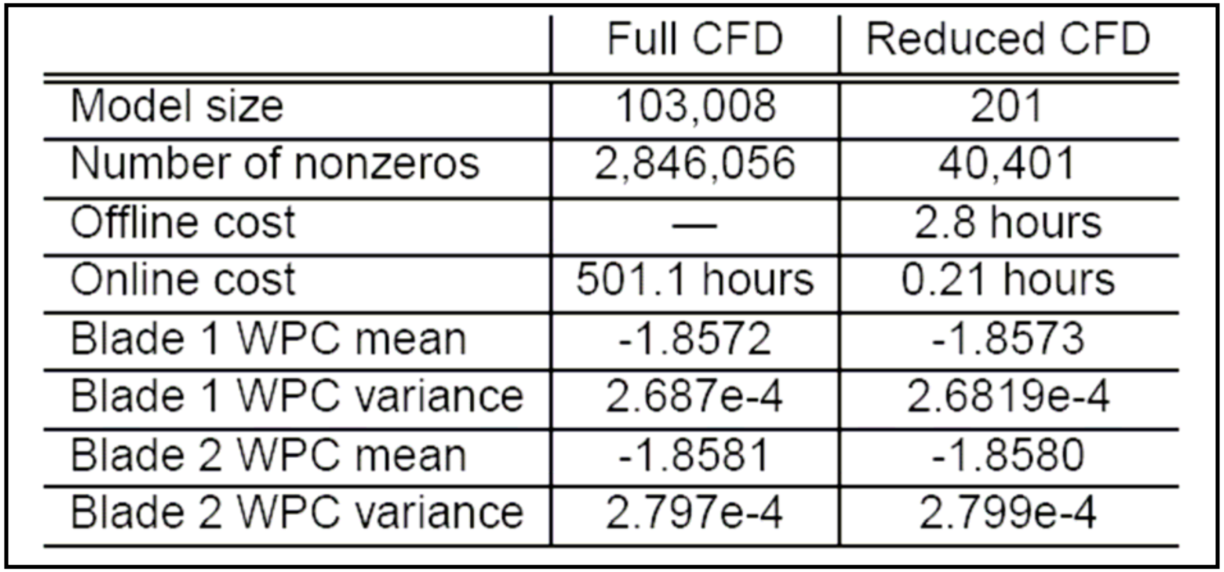
CFD model

~500 hours



Reduced model

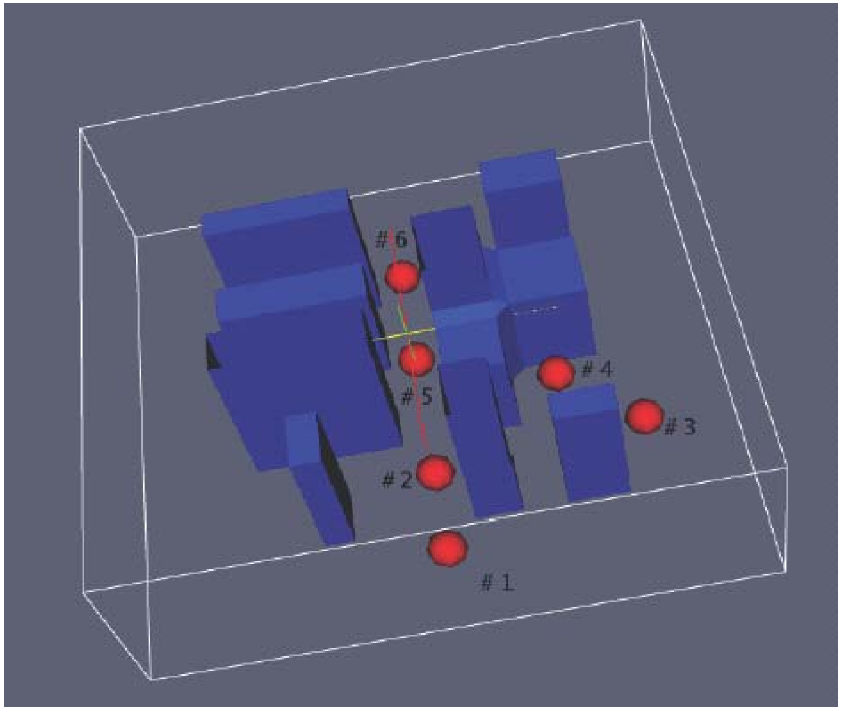
~0.2 hours



	Full CFD	Reduced CFD
Model size	103,008	201
Number of nonzeros	2,846,056	40,401
Offline cost	—	2.8 hours
Online cost	501.1 hours	0.21 hours
Blade 1 WPC mean	-1.8572	-1.8573
Blade 1 WPC variance	2.687e-4	2.6819e-4
Blade 2 WPC mean	-1.8581	-1.8580
Blade 2 WPC variance	2.797e-4	2.799e-4

Real-Time Inverse Problem Applications

- Real-time inversion problems can be formulated as PDE-constrained optimization problems
  - Emergency response, hazard assessment, structural health monitoring, etc.
- Typical scenario: contaminant transport through 3D urban canyon
  - Wind from mesoscale models (MM5)
  - Sparse sensor readings of concentration
  - Solve inversion problem to determine initial condition
  - Solve forward problem to determine response



Goal: create a reduced-order model that predicts outputs of interest *over “all important” initial conditions*

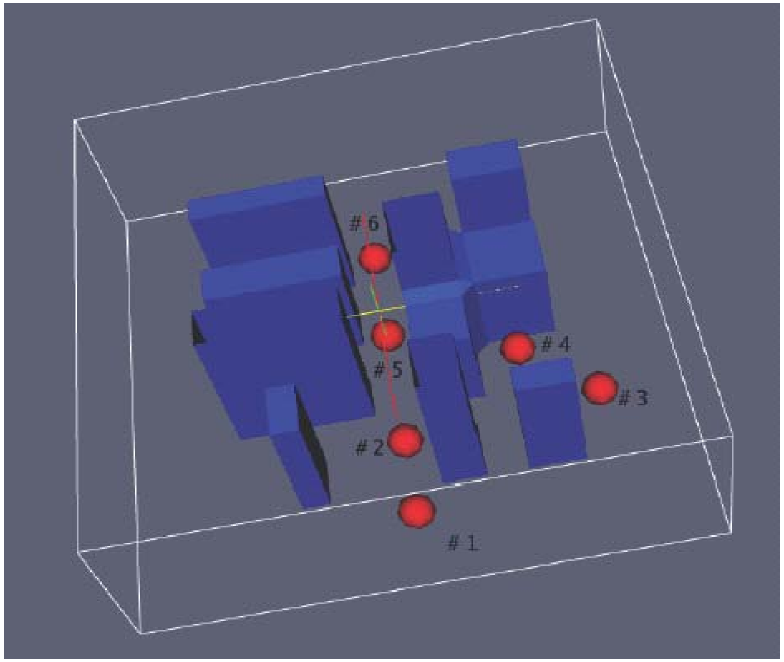
Model-Constrained Optimization: Linear Case

$$\begin{array}{rcll} \dfrac{\partial w}{\partial t} + \vec{v} \cdot \nabla w - \kappa \nabla^2 w & = & 0 & \text{in } \Omega \times (0, t_f) \\ w & = & 0 & \text{on } \Gamma_D \times (0, t_f) \\ \dfrac{\partial w}{\partial n} & = & 0 & \text{on } \Gamma_N \times (0, t_f) \\ w & = & w_0 & \text{in } \Omega \text{ for } t = 0 \end{array}$$

- Linear problem: greedy sampling via eigenvectors of the Hessian

$$w_0^* = \arg\max_{w_0} \frac{1}{2} w_0^T \mathbf{H}_e w_0$$

- Guaranteed error bound for reduced model
- *p*=31 eigenvectors leads to reduced model with *m*=137

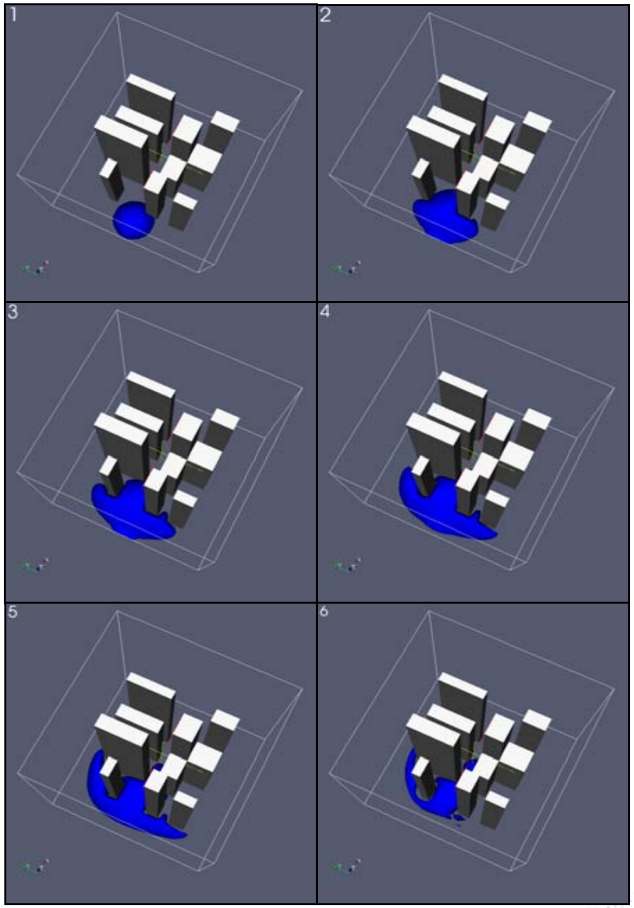


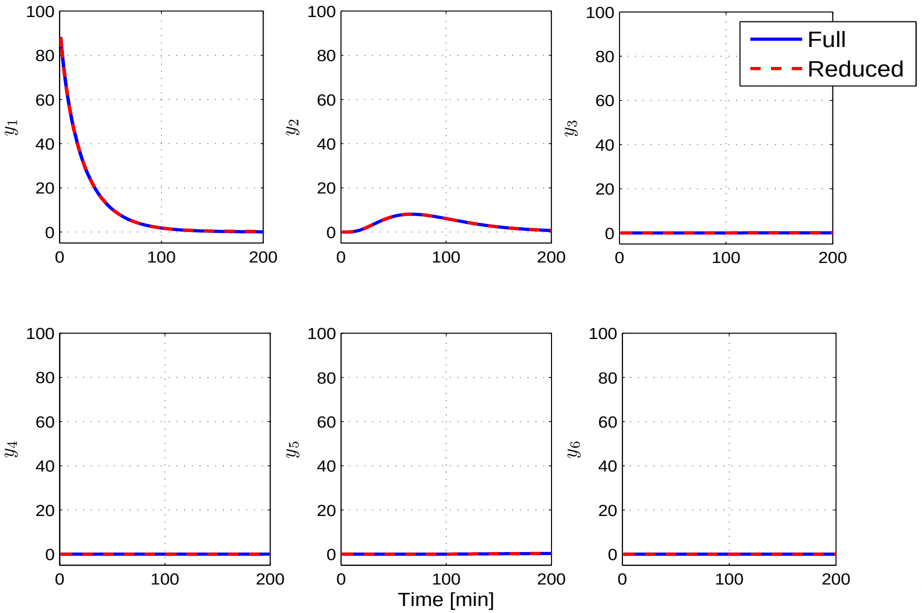
Contaminant Transport Inverse Problem

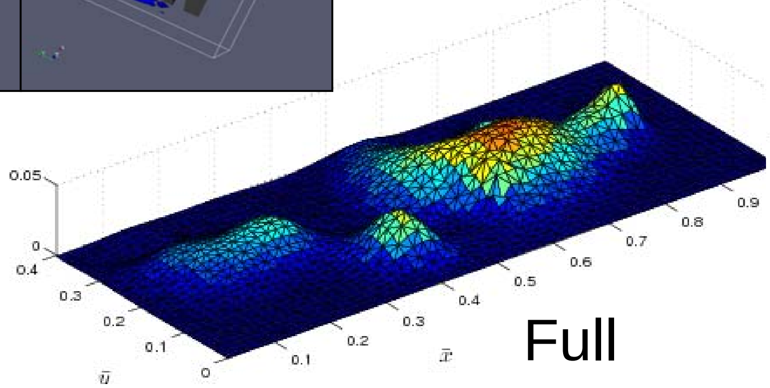
Full model: 68,921 states

Reduced model: 137 states

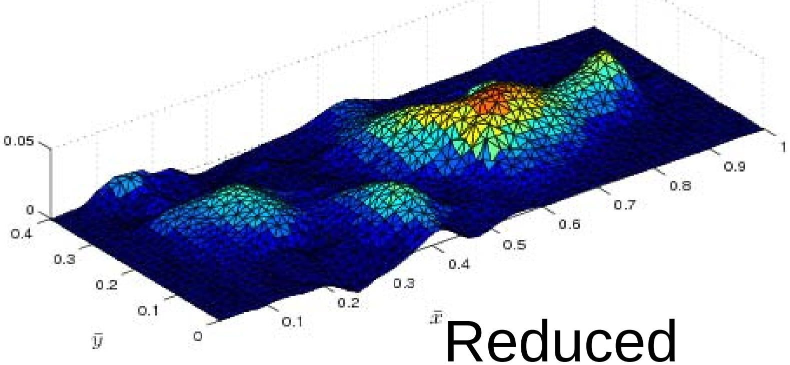
Six sensor locations, *Pe*=900







Full



Reduced

Reduction of Nonlinear Systems

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

- Nonlinear systems: standard projection approach leads to a model that is low order but still expensive to solve
- Proposed approach: interpolation *(Barrault, Maday, Nguyen, Patera, 2004)*
  - Approximate the nonlinear term using interpolation
  - Leads to an efficient offline/online scheme

Efficient Reduction of Nonlinear Systems

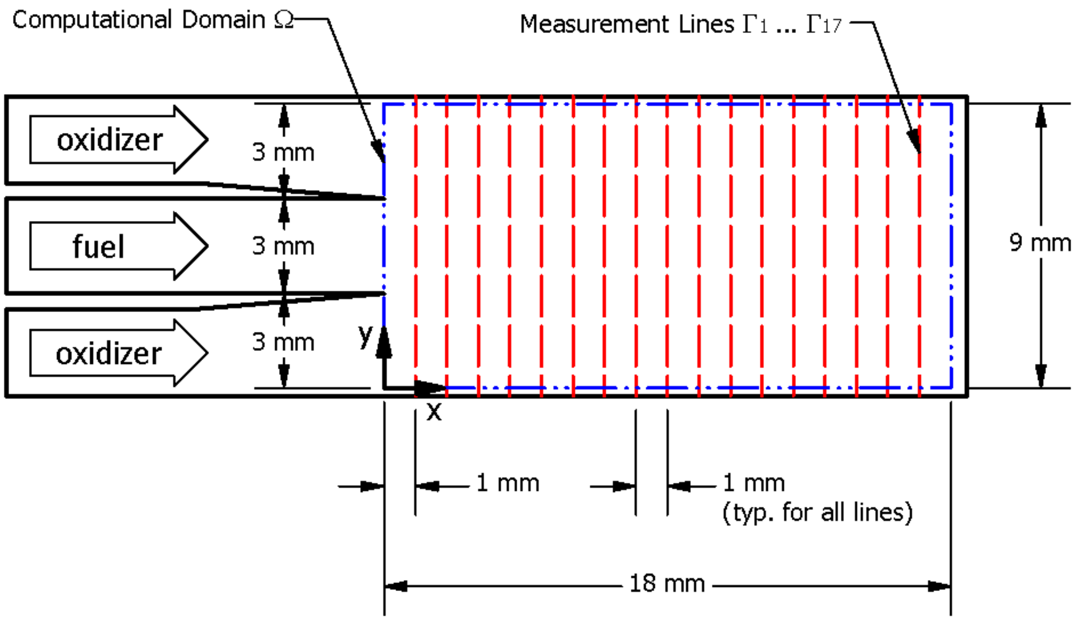
- Compute the snapshot set
    - For each snapshot, compute the state and the nonlinear term
- Compute the POD basis that spans the state (*N* basis vectors)
- Compute the POD interpolation basis that spans the nonlinear terms (*M* basis vectors)
- Compute a set of spatial interpolation points (various heuristic methods)
- Offline: form and store the parameter-independent matrices
- Online solution of reduced model cost:

$O(MN^2)$ per Newton iteration

M=number of interpolation points

N=number of state basis functions

Nonlinear Combustion Chamber Model

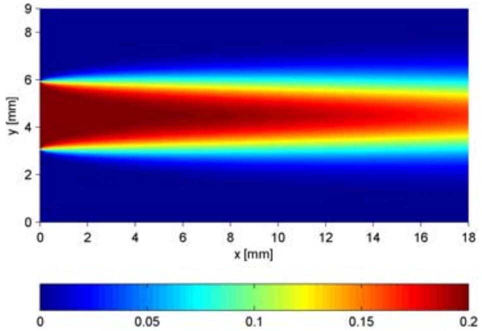


Convection-diffusion-reaction problem:

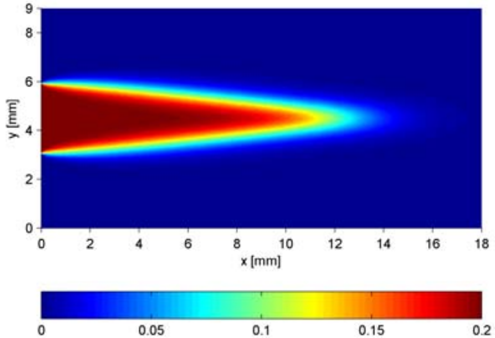
$$\mathbf{U} \cdot \nabla w - \nabla(\kappa \nabla w) + s(w; \mathbf{z}) = f$$

$$s(w; \mathbf{z}) = Aw(c - w)e^{\frac{-E}{d-w}}$$

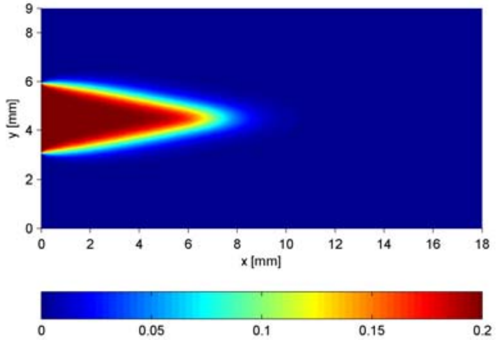
$$\begin{aligned} w &= w_D \quad \text{on} \quad \Gamma_{\text{in}} \\ \nabla w \cdot \hat{\mathbf{n}} &= 0 \quad \text{on} \quad \partial\Omega \setminus \Gamma_{\text{in}} \end{aligned}$$



Molar fraction of fuel for (ln(A), E) = (5.00, 0.15)



Molar fraction of fuel for (ln(A), E) = (7.25, 0.15)



Molar fraction of fuel for (ln(A), E) = (7.25, 0.05)

Model Reduction for CDR Problem

- Finite element model leads to nonlinear system of equations:

$$\boxed{\underbrace{A\mathbf{x}(\mathbf{z})}_{\textit{Linear term}} + \underbrace{\mathbf{S}\left(\mathbf{x}(\mathbf{z});\mathbf{z}\right)}_{\textit{Nonlinear term}} = \mathbf{F}}$$

- Reduced model has the form:

Usual projection matrices

Nonlinear term operates on M-dimensional vector

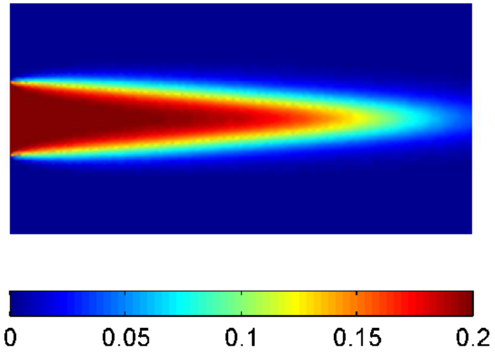
Usual projection vector

$$\boxed{A_0 + A_N\,\mathbf{x}_{N,M} + E_{N,M}\,\mathbf{s}\left(\bar{\mathbf{x}}_M + D_{M,N}\,\mathbf{x}_{N,M};\ \mathbf{z}\right) = \mathbf{F}_N}$$

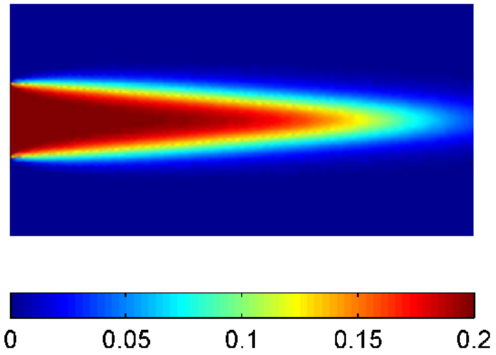
Products of state and interpolation basis vectors integrated over domain

State basis vectors evaluated at interpolation points

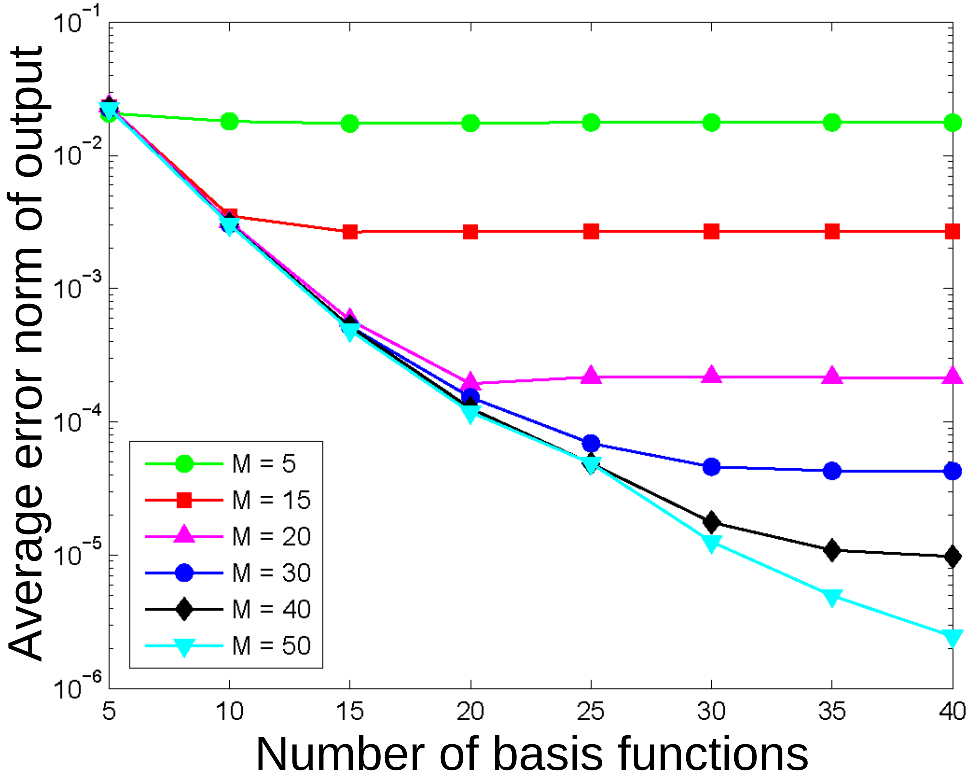
Model Reduction for CDR Problem



FEM solution



ROM: 40 basis functions, 50 interp. points



		Reduced-order model			FEM
N	M	Avg. rel. error	Max. rel. error	Online time	Comp. time
5	50	2.25 E − 02	9.73 E − 02	1.59 E − 05	1
10	50	3.03 E − 03	2.78 E − 02	1.61 E − 05	1
20	50	1.18 E − 04	2.00 E − 03	1.63 E − 05	1
30	50	1.26 E − 05	4.48 E − 04	1.71 E − 05	1
40	50	2.47 E − 06	1.34 E − 04	2.00 E − 05	1

CDR Inverse Problem

- Inverse problem: given sparse measurements of concentration at certain locations in the combustor, can we infer the values of the Arrhenius parameters *A* and *E*?

- Deterministic approach: solve a PDE-constrained optimization problem to find the solution that best matches the measurements
  - Combine data measurements with knowledge of governing equations
  - No way to account for measurement error
  - Yields no information regarding uncertainty in the "best estimate"

- Bayesian formulation: solution is a *distribution* of parameter estimates
  - Incorporates measurement error
  - Incorporates *a priori* knowledge of parameter values
  - Requires Markov Chain Monte Carlo simulation, i.e. thousands of forward solves. Expensive!

CDR Inverse Problem

- Inverse problem: given sparse measurements, **y***, infer the values of the Arrhenius parameters, **z**:

$$\begin{aligned} \min_{\mathbf{x}_{N,M},\mathbf{z}} \quad \mathcal{J}(\mathbf{x}_{N,M},\mathbf{z}) &= \frac{1}{2}\|\mathbf{y}_{N,M} - \mathbf{y}*\|_2^2 \\ \text{s.t.} \quad \left. \begin{aligned} \mathbf{R}_N(\mathbf{x}_{N,M};\mathbf{z}) &= 0 \\ \mathbf{y}_{N,M} &= \mathbf{C}_N\mathbf{x}_{N,M} + \mathbf{C}_0 \end{aligned} \right\} & \textit{Reduced model} \\ \mathbf{z}_{\min} \leq \mathbf{z} &\leq \mathbf{z}_{\max} \end{aligned}$$

- Bayes theorem relates forward and inverse probabilities:

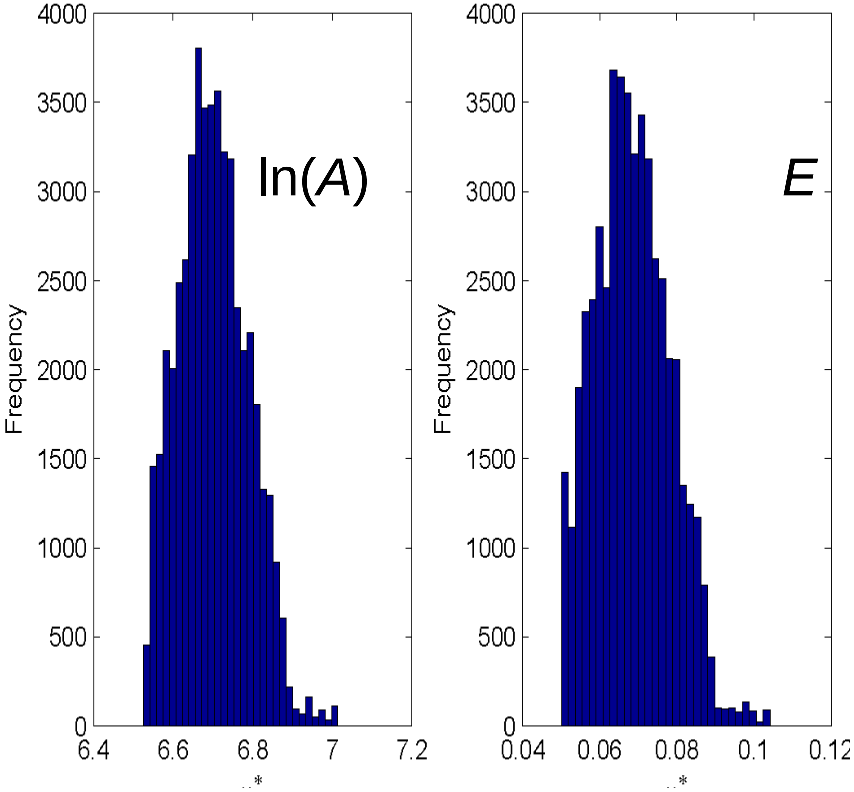
$$p(\mathbf{z}|\mathbf{y}^*) = \frac{1}{p(\mathbf{y}^*)} p(\mathbf{y}^*|\mathbf{z}) p(\mathbf{z})$$

Knowledge of parameters given a set of measurements (posterior probability)

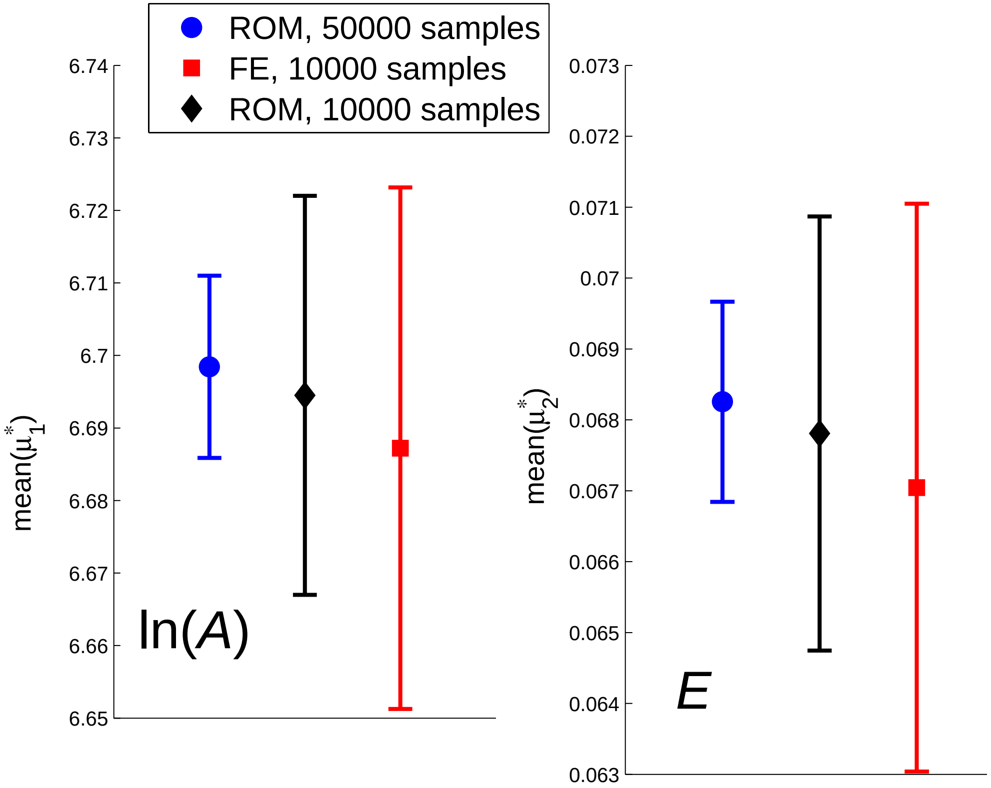
Conditional probability of outputs given input parameters (forward probability)

Knowledge of parameters before collecting measurements (prior probability)

Bayesian Uncertainty Quantification of CDR Inverse Problem



Marginal posterior histograms for Arrhenius parameters. Note that parameters can vary within the ranges [5.00, 7.25] and [0.05, 0.15].

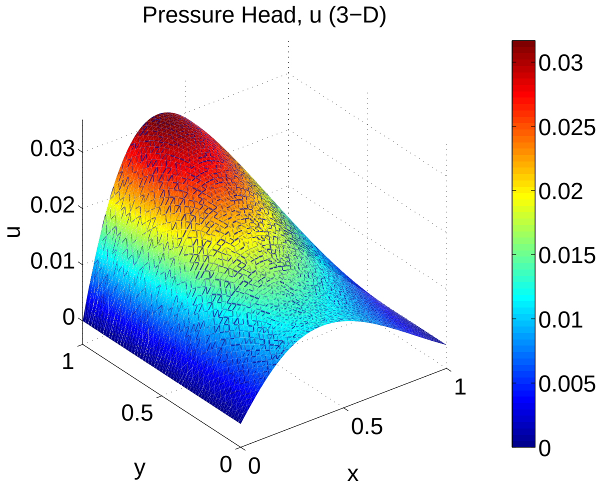


95% confidence intervals for the mean estimates of the Arrhenius parameters.

FEM: 10,000 samples, ~110 hours
ROM: 10,000 samples, ~100 seconds

Summary

- Successful application of model reduction in a broad range of fields
  - Optimal control, fluid dynamics, structural dynamics, circuit design, geophysics, atmospheric modeling
- Research challenges being addressed to transition from reduced-order models for simulation to reduced-order models for optimization
  - Model-constrained optimization approach to sampling and building the reduced basis
  - Opens a new class of problems: design, inverse problem applications, probabilistic analyses
- Ongoing research
  - Bayesian approach to inverse problems
  - The role of reduced models in the probabilistic setting
  - Sampling in infinite-dimensional parameter space


Pressure Head, u (3−D)