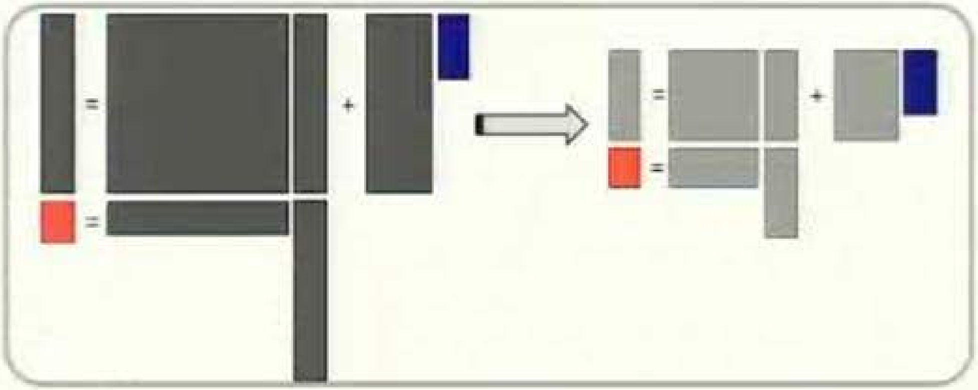
Data-Driven Model Reduction to Support Decision Making in Complex Systems

Karen Willcox
Joint work with Tiangang Cui, Laura Mainini, Youssef Marzouk, Benjamin Peherstorfer

SIAM CSE15

Salt Lake City, UT
March 14, 2015



PROJECTION-BASED REDUCED MODELS

Data-driven reduced models

Reduced models are typically built and used in a static way:

- *offline phase*: sample a high-fidelity model, build a low-dimensional basis, project to build the reduced model
- *online phase*: use the reduced model

A single and/or static model is insufficient in many situations

- Complex physical system
- Other information sources, besides models
- Decision goal changes dynamically
- Physical system changes dynamically

→ A need for (online) adaptation of models

Data-driven reduced models

Adaptation is data-driven, where data could be:

- Sensor data collected online (e.g., structural sensors on board an aircraft)
- Simulation data collected online (e.g., over the path to an optimal solution)

Achieve adaptation in a variety of ways:

- Adapt the basis
- Adapt the way in which nonlinear terms are approximated
- Adapt the reduced model itself
- Construct localized reduced models; adapt through model choice

Start with a physics-based model

Arising, for example, from systems of ODEs or spatial discretization of PDEs describing the system of interest

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

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

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

$\mathbf{x} \in \mathbf{R}^N$: 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-based model reduction

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

$\mathbf{x} \approx \mathbf{V}\mathbf{x}_r$

$$\begin{aligned} \mathbf{r} &= \mathbf{V}\dot{\mathbf{x}}_r - \mathbf{A}\mathbf{V}\mathbf{x}_r - \mathbf{B}\mathbf{u} \\ \mathbf{y}_r &= \mathbf{C}\mathbf{V}\mathbf{x}_r \end{aligned}$$

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

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

$$\begin{aligned} \mathbf{A}_r(\mathbf{p}) &= \mathbf{W}^T\mathbf{A}(\mathbf{p})\mathbf{V} \\ \mathbf{B}_r(\mathbf{p}) &= \mathbf{W}^T\mathbf{B}(\mathbf{p}) \\ \mathbf{C}_r(\mathbf{p}) &= \mathbf{C}(\mathbf{p})\mathbf{V} \end{aligned}$$

$\mathbf{x} \in \mathbf{R}^N$: 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

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

$\mathbf{V} \in \mathbf{R}^{N \times n}$: reduced basis

Projection-based model reduction of nonlinear systems

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

- Nonlinear systems: standard projection approach leads to a model that is low order but still expensive to solve

- The cost of evaluating the nonlinear term

$$\mathbf{f}_r(\mathbf{x}_r,\mathbf{u}) = \mathbf{V}^T\mathbf{f}(\mathbf{V}\mathbf{x}_r,\mathbf{u})$$

still depends on *N*, the dimension of the large-scale system

- Can achieve efficient nonlinear reduced models via interpolation, e.g., Missing Point Estimation [Astrid et al., 2008], (Discrete) Empirical Interpolation Method [Barrault et al., 2004; Chaturantabut & Sorensen, 2010]

$$\dot{\mathbf{x}}_r = \mathbf{A}_r\mathbf{x}_r + \mathbf{E}_r\,\mathbf{f}_r(\mathbf{D}_r\mathbf{x}_r,\mathbf{u})$$

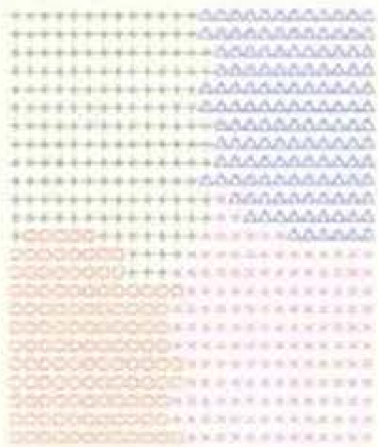
Computing the Basis: Proper Orthogonal Decomposition (POD)

(aka Karhunen-Loève expansions, Principal Components Analysis, Empirical Orthogonal Eigenfunctions, ...)

- Consider *K* **snapshots** $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K \in \mathcal{R}^N$ [Sirovich, 1991] (solutions at selected times or parameter values)
- Form the snapshot matrix $X = [\mathbf{x}_1 \ \mathbf{x}_2 \ \ldots \ \mathbf{x}_K]$
- Choose the *n* basis vectors $\mathbf{V} = [\mathbf{V}_1 \ \mathbf{V}_2 \cdots \mathbf{V}_n]$ to be left singular vectors of the snapshot matrix, with singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq \sigma_{n+1} \geq \cdots \geq \sigma_K$
- This is the optimal projection in a least squares sense:

$$\boxed{\min_{\mathbf{V}} \sum_{i=1}^{K} ||\mathbf{x}_i - \mathbf{V}\mathbf{V}^T\mathbf{x}_i||_2^2 = \sum_{i=n+1}^{K} \sigma_i^2}$$

- Similar idea to construct a basis for the nonlinear term using snapshots of $\mathbf{f}(\mathbf{x}, \mathbf{u})$ in DEIM approximation

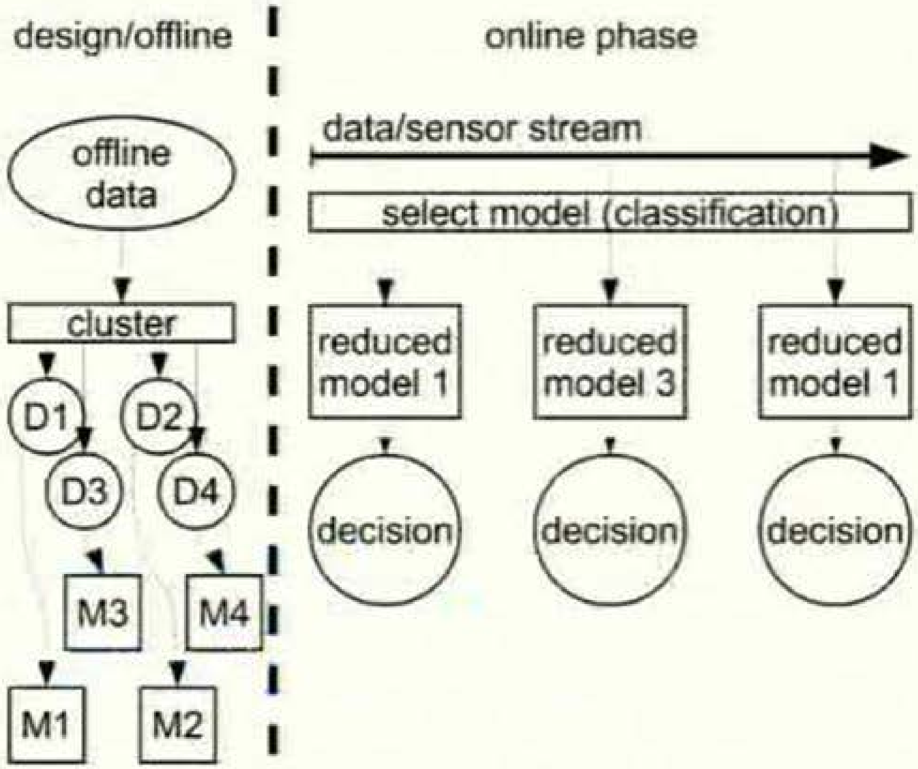


LOCALIZED REDUCED MODELS

Localized and adaptive reduced models

[Peherstorfer, Butnaru; W., Bungartz; SISC 2014]

- Automatic model management based on machine learning
  - Cluster set of snapshots $\mathcal{S} = \{x_1, \dots, x_M\} \subset \mathbb{R}^N$ into $\mathcal{S} = \mathcal{S}_1 \uplus \cdots \uplus \mathcal{S}_k$ (using e.g. k-means)
  - Create a separate local (reduced) model for each cluster
  - Derive a basis $Q \in \mathbb{R}^{N \times m}, m \ll N$ to obtain low-dimensional indicator $z_i = Q^T x_i$ that describes state x_i
  - Learn a classifier $g: Z \to \{1, \dots, k\}$ to map from low-dimensional indicator z to model index (using e.g. nearest neighbors)
  - Classify current state/indicator online and select model



→ Reduced models are tailored to local system behavior

Localized and adaptive reduced model

[Peherstorfer, Butnaru, W., Bungartz; SISC 2014]

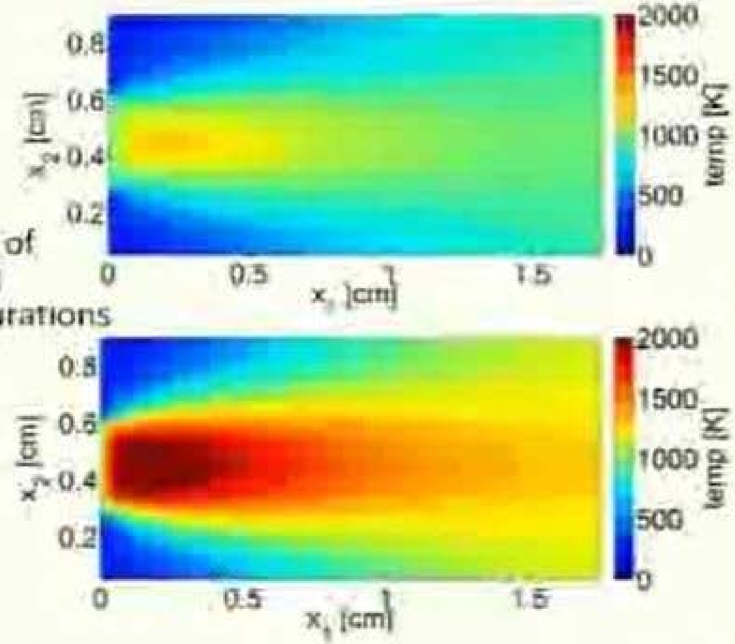
- Example: Reacting flow with one-step reaction

$$2H_2 + O_2 \rightarrow 2H_2O$$

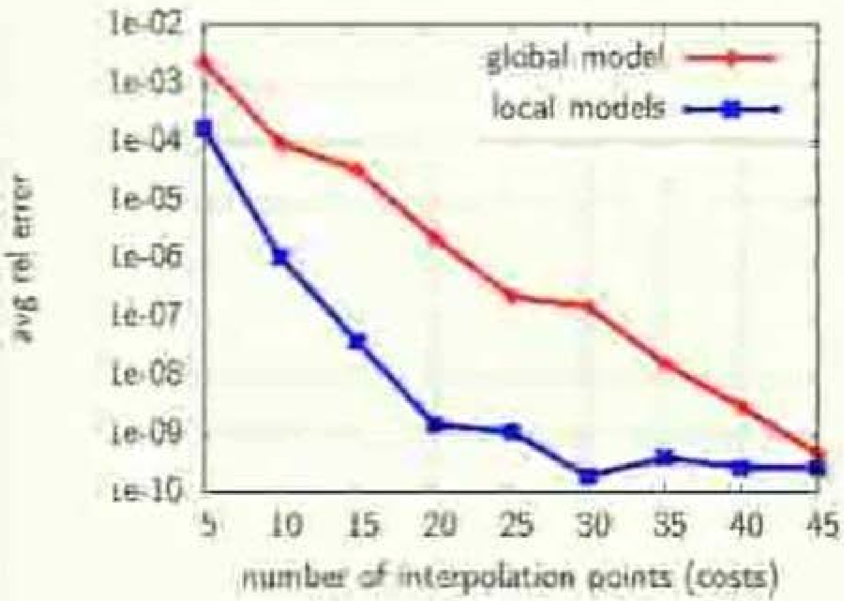
- Governed by convection-diffusion-reaction equation

$$\kappa \Delta \boldsymbol{y} - \nu \nabla \boldsymbol{y} + \boldsymbol{F}(\boldsymbol{y}, \boldsymbol{\mu}) = 0 \quad \text{in } \Omega$$

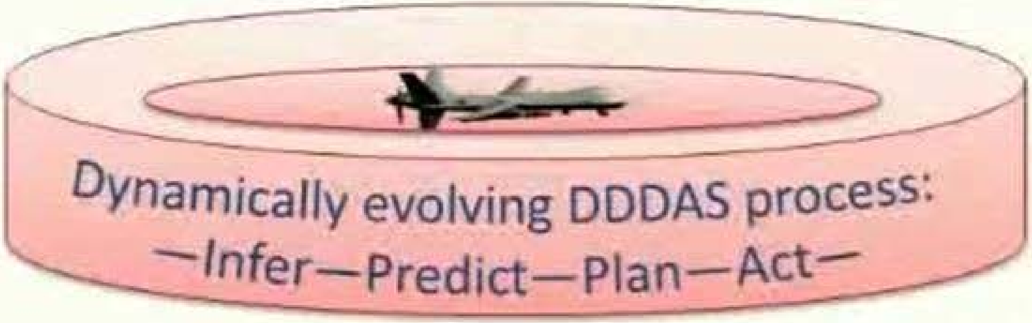
- Exponential nonlinearity (Arrhenius-type source term)
- POD-DEIM reduced model



Temperature field of flame for different parameter configurations



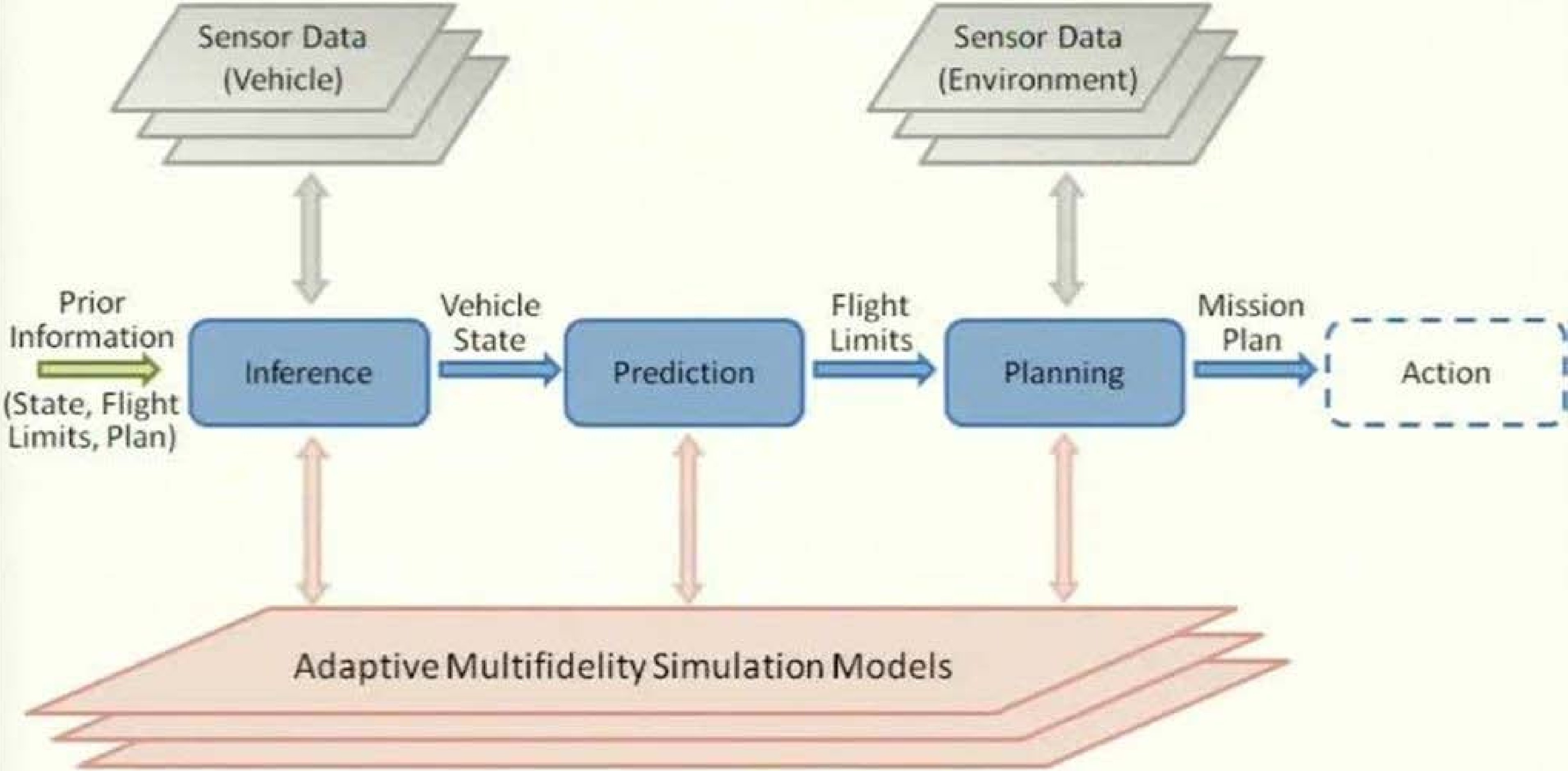
Combining **4 local models** with machine-learning-based model management achieves **accuracy improvement by up to two orders of magnitude** compared to a single, global model



DYNAMIC DATA-DRIVEN DECISIONS VIA ADAPTIVE REDUCED MODELS

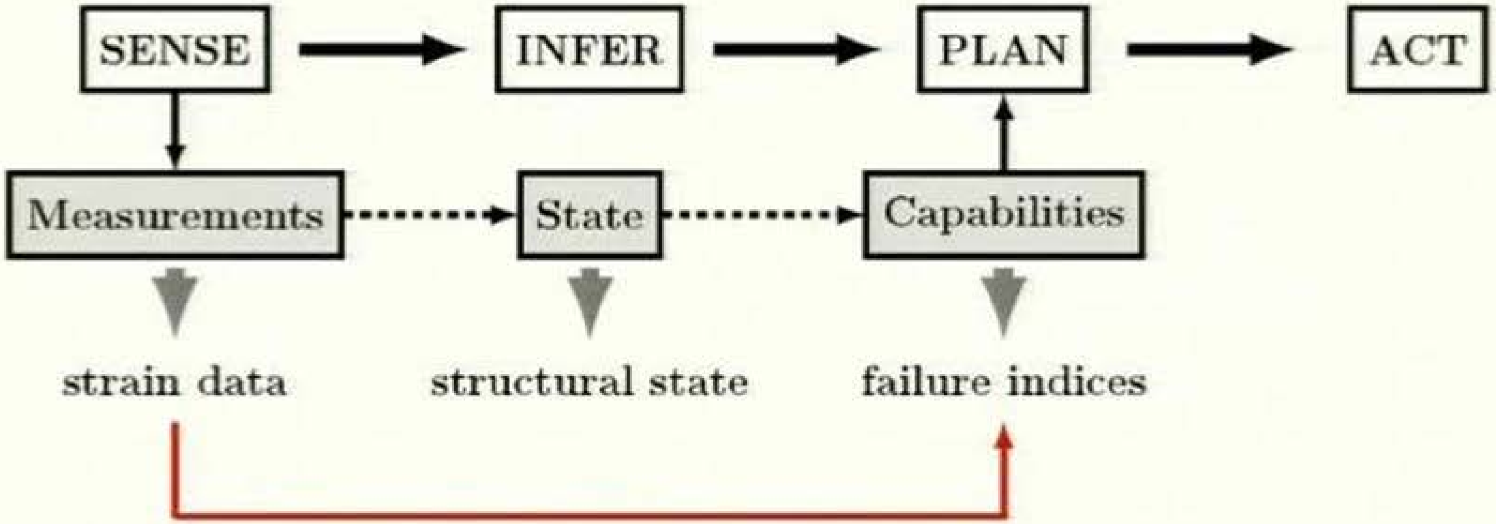
Motivating application: Dynamic data-driven decisions

- A **self-aware aerospace vehicle** can dynamically adapt the way it performs missions by gathering information about itself and its surroundings and responding intelligently



Dynamic data-driven structural assessment

[Mainini & W., AIAA 2015]

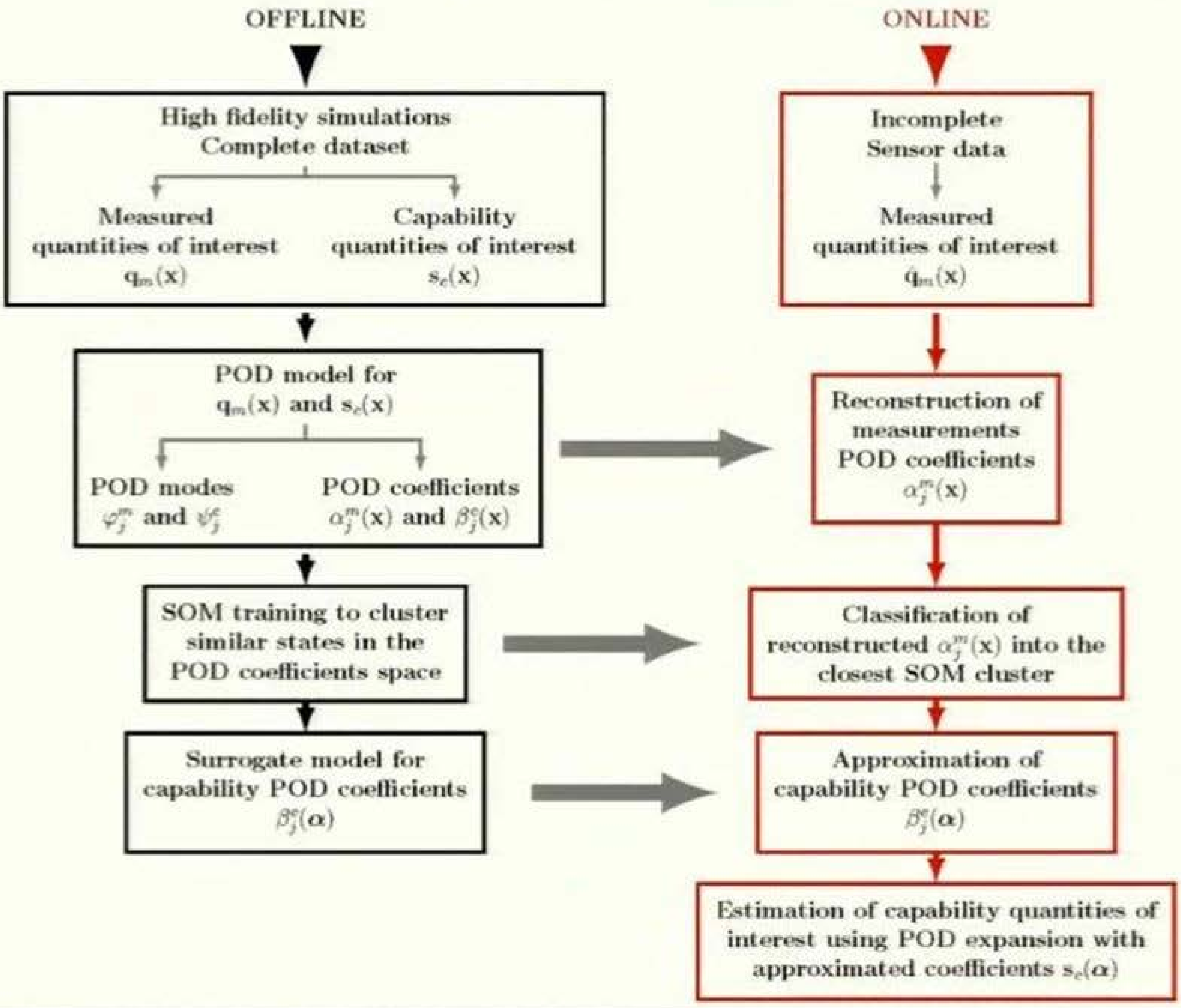


We want to:

use sensed structural data
└→ estimate the structural state
└→ update the flight capabilities
└→ dynamically re-plan the mission

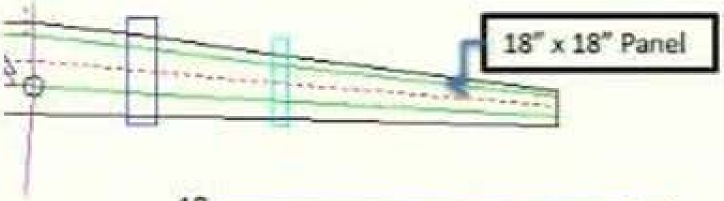
adaptive surrogate modeling to map from sensed data to capabilities

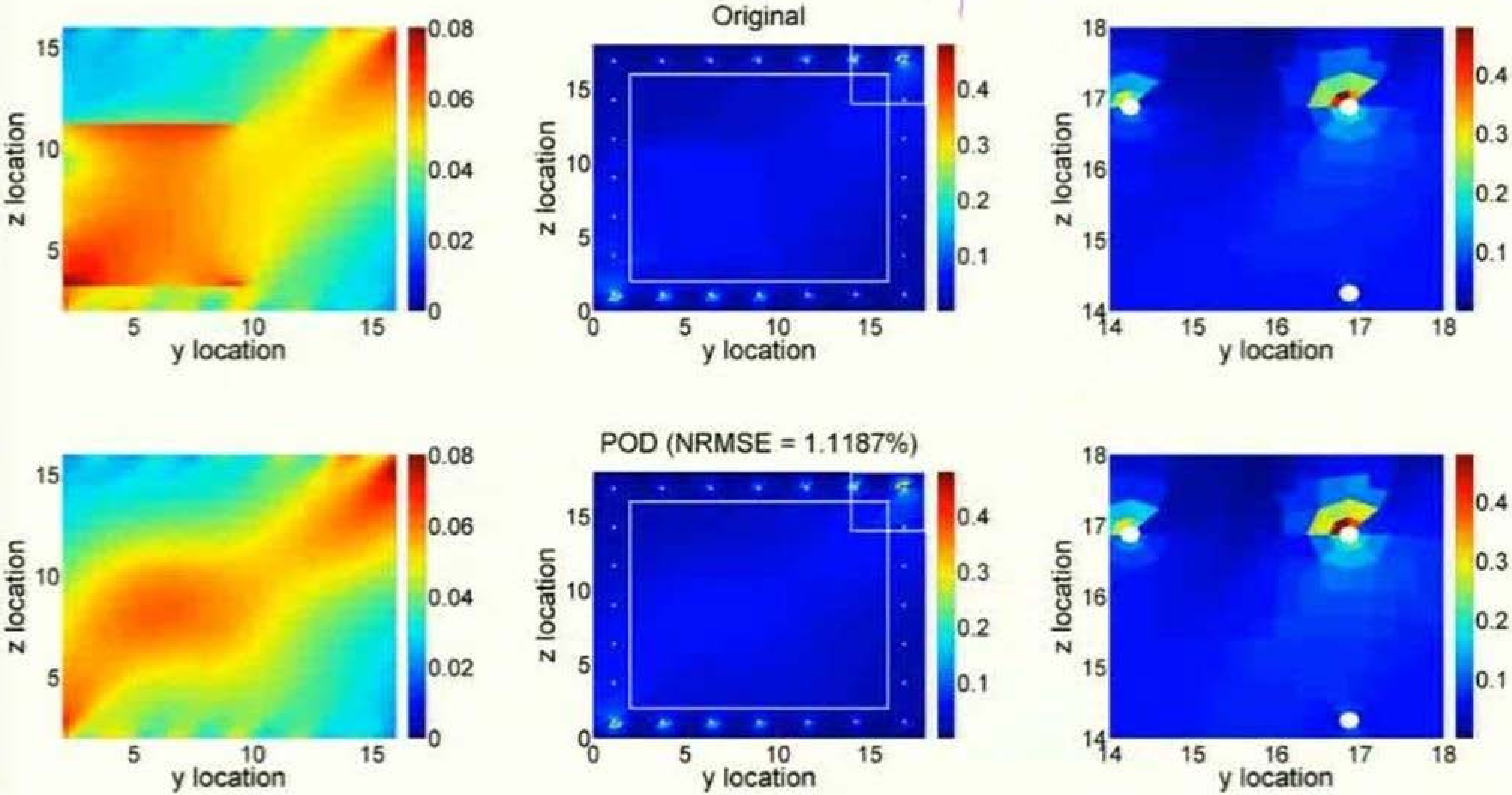
Localization and adaptation using POD and self-organizing maps



Dynamic data-driven structural assessment

Damage: 7.85 x 9.91 in^2, at (5.47, 7.15), ply 4 to 6





measure strain data → POD + SOM models → estimate Failure Index (FI)

$$\pi(p|d) \sim L(d|p)\, \pi_0(p)$$

DATA-DRIVEN REDUCED MODELS FOR INVERSE PROBLEMS

SUMMARY

Data-driven reduced models: Summary

Data-driven reduced models via adaptation, where data could be:

- Sensor data collected online (e.g., structural sensors on board an aircraft)
- Simulation data collected online (e.g., over the path to an optimal solution)

Achieve adaptation in a variety of ways:

- Adapt the basis
  - Data-driven reduced models for inverse problems *(Cui et al., 2014)*
- Adapt the way in which nonlinear terms are approximated
  - Adaptive DEIM *(Peherstorfer, MS260 Wed 1210)*
- Adapt the reduced model itself
  - Low-rank updates *(Peherstorfer & W., in review)*
- Construct localized reduced models; adapt through model choice *(Peherstorfer et al., 2014; Mainini & W., 2015; Mainini MS55)*

Conclusions

- The next generation of complex engineered systems will be endowed with sensors and computing capabilities that enable new modes of decision-making
- New opportunities for model reduction and surrogate modeling, in particular through data-driven reduced models that exploit the synergies of physics-based computational modeling and physical data
  - different sources of information (models, sensors, etc.) tell us different things about the decision problem, with the collective information they provide being greater than the individual parts
  - need methods that leverage complementary perspectives from physics-based modeling and data-driven approaches