Predictive Data Science for physical systems

From model reduction to scientific machine learning

Professor Karen E. Willcox
SC19 | November 21, 2019

ODEN INSTITUTE
FOR COMPUTATIONAL ENGINEERING & SCIENCES

TEXAS
The University of Texas at Austin

The Team

Michael Kapteyn

MIT

Stefanie Salinger

UT Austin

Dr. David Knezevic

CTO, Akselos

Luwen Huang

MIT Mapping Lab

Cory Kays

Aurora Flight Sciences

Funding sources: US Air Force **Computational Math** Program (F. Fahroo);
US Department of Energy **AEOLUS MMICC** (S. Lee, W. Spotz);
US Air Force **Dynamic Data Driven Application Systems** Program (E. Blasch);
The **Boeing** Company;
SUTD-MIT **International Design Centre**

Outline

1 **The Unreasonable Effectiveness of Computational Science**

From forward simulation to predictive data science

---

2 **Predictive Digital Twin**

Reduced models enable scalable system-level modeling & rapid updating with dynamic data

---

3 **Conclusions & Outlook**

The Unreasonable Effectiveness of Computational Science

From forward simulation to predictive data science

ODEN INSTITUTE
FOR COMPUTATIONAL ENGINEERING & SCIENCES

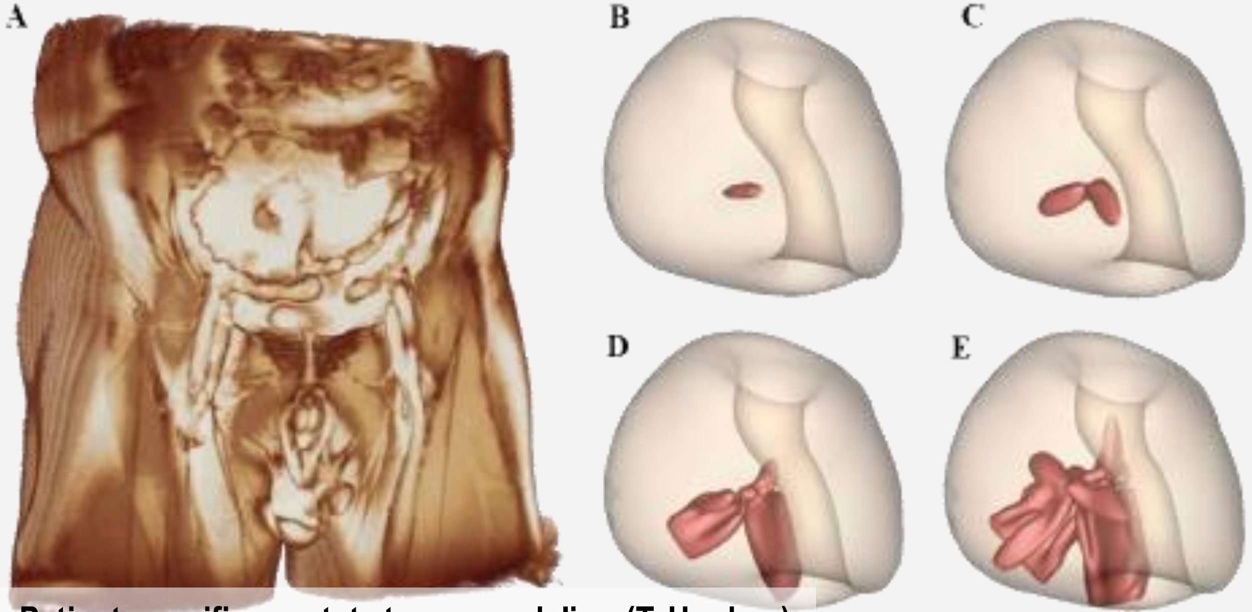
How do we harness the explosion of data to extract knowledge, insight and decisions?

BIG DECISIONS

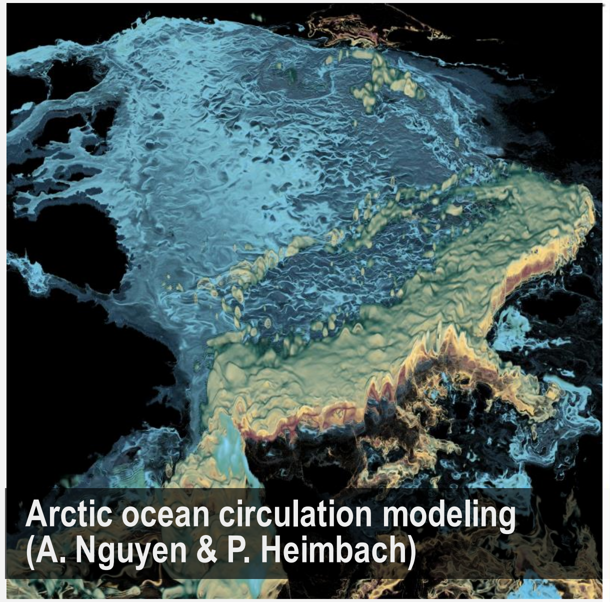
need more than just big data…

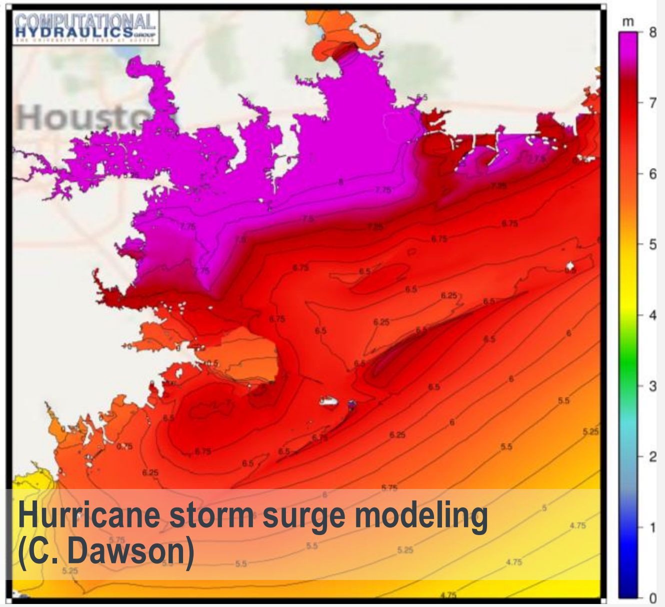
They need **BIG MODELS** too.

Inspired by
Coveney, Dougherty, Highfield "Big data need big theory too"



Patient-specific prostate tumor modeling (T. Hughes)


Arctic ocean circulation modeling (A. Nguyen & P. Heimbach)



Hurricane storm surge modeling (C. Dawson)

BIG DECISIONS

need more than just big data…

1. **Complex multiscale multiphysics phenomena**
   driving the dynamics of high-consequence applications
2. **High dimensional parameters**
   underlying the characterization of scientific and engineering systems
3. **Data are sparse, intrusive and expensive to acquire**
   especially in the most critical regimes
4. **Uncertainty quantification**
   in model inference and certified predictions in regimes beyond training data

BIG DECISIONS

need **BIG MODELS** too.

> "
> **Computational Science**
> **or Computational Science & Engineering (CSE)**
>
> is an interdisciplinary field that uses mathematical modeling and advanced computing to understand and solve complex problems. At its core CSE involves developing models and simulations to understand physical/natural systems.

BIG DECISIONS must incorporate the **predictive power**, **interpretability**, and **domain knowledge** of physics-based models.

What is a physics-based model?

A representation of the **governing laws of nature** that innately embeds the concepts of **time, space, and causality**

In solving the governing equations of the system, we constrain the **predictions** to lie on the **solution manifold** defined by the laws of nature

Example: equations of linear elasticity

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial \sigma}{\partial x} + \frac{\partial \sigma}{\partial y} + F$$

equation of motion (Newton's 2nd law)

$$\varepsilon = \frac{1}{2}[\nabla u + (\nabla u)^{\top}]$$

strain-displacement equations

$$\sigma = C : \varepsilon$$

constitutive equations

+ boundary conditions
+ initial conditions

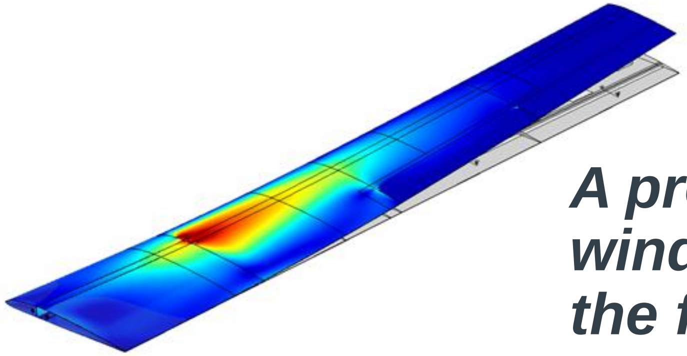
a mathematical model of how solid objects deform, relating stress σ, strain ε, displacement u, and loading F

The unreasonable effectiveness of physics-based models [Wigner, 1960]

Solving a physics-based model:

Given initial conditions, boundary conditions, loading conditions, and system parameters

Compute solution trajectories $\sigma(x, y, t), \varepsilon(x, y, t)$, $\mathrm{u}(x, y, t)$, …



A predictive window on the future

Forward simulations
Advancing scientific discovery & engineering innovation

Optimization & inverse problems
Advancing estimation, design & control

Uncertainty quantification
Towards Predictive Science

Scientific machine learning
Towards Predictive Data Science

6 decades of Computational Science & Engineering

How do we harness the explosion of data to extract knowledge, insight and decisions?

INSTITUTE FOR COMPUTATIONAL ENGINEERING & SCIENCES
ODEN

Learning from data through the lens of models…

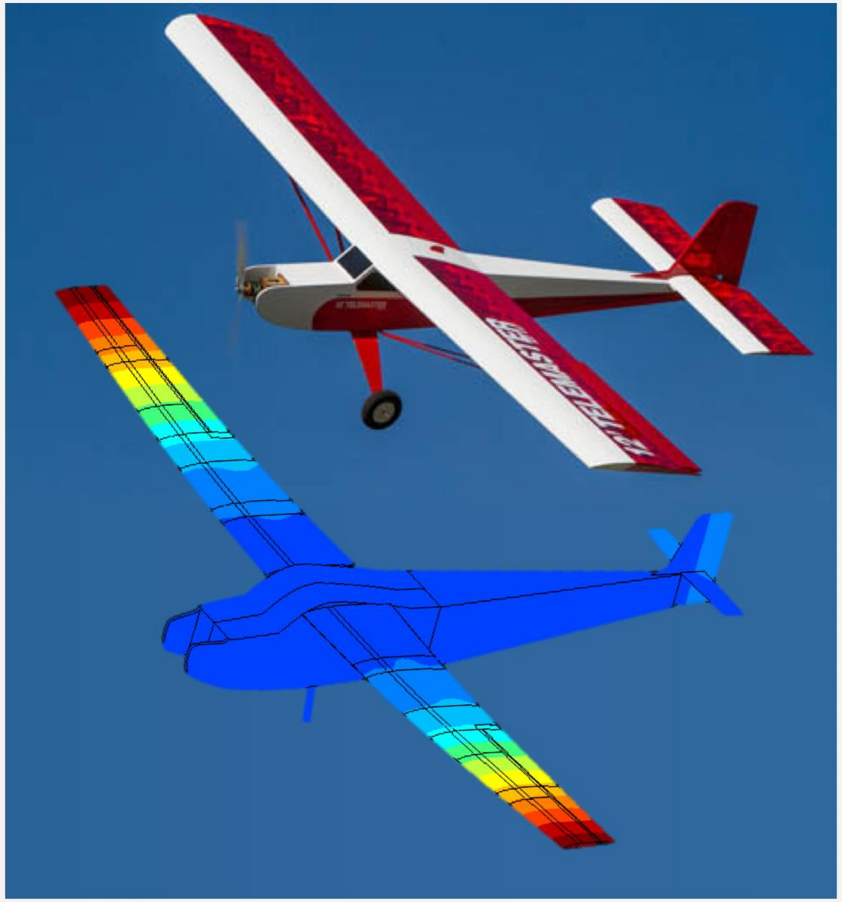
dynamical systems
meshing methods
uncertainty quantification
Bayesian inference
large scale optimization
boundary values
finite volumes
inverse problems
adjoints
convergence analysis
high performance computing
regularization
image processing
finite elements
randomized methods
physics-based models
projection-based model reduction
partial differential equations
data assimilation
approximation theory
big data analysis
low rank approximation
error analysis
importance sampling

Learning from data through the lens of models…

dynamical systems
uncertainty quantification
meshing methods
Bayesian inference
finite volumes
inverse problems
large scale optimization
boundary values
adjoints
convergence analysis
high performance computing
regularization
image processing
randomized methods
finite elements
physics-based models
projection-based model reduction
partial differential equations
data assimilation
approximation theory
big data analysis
low rank approximation
error analysis
importance sampling

Predictive Digital Twin

Component-based reduced models enable
scalable predictive modeling &
rapid model updating with dynamic data



Digital twins enable data-driven decisions

High-consequence decisions require digital twins that are
predictive • reliable • explainable

Vehicle & environmental data

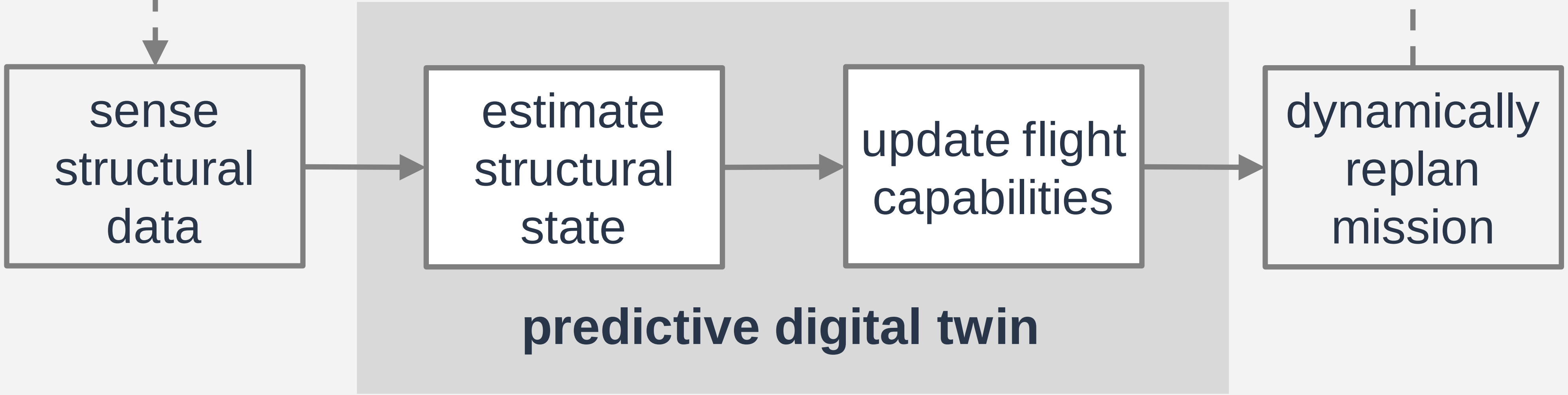
Physics-based predictive models

Predictive Digital Twin

self-aware aircraft

Our **digital twin adapts** to the **evolving UAV structural health…**

providing **near real-time capability estimates** that enable **dynamic decision-making**

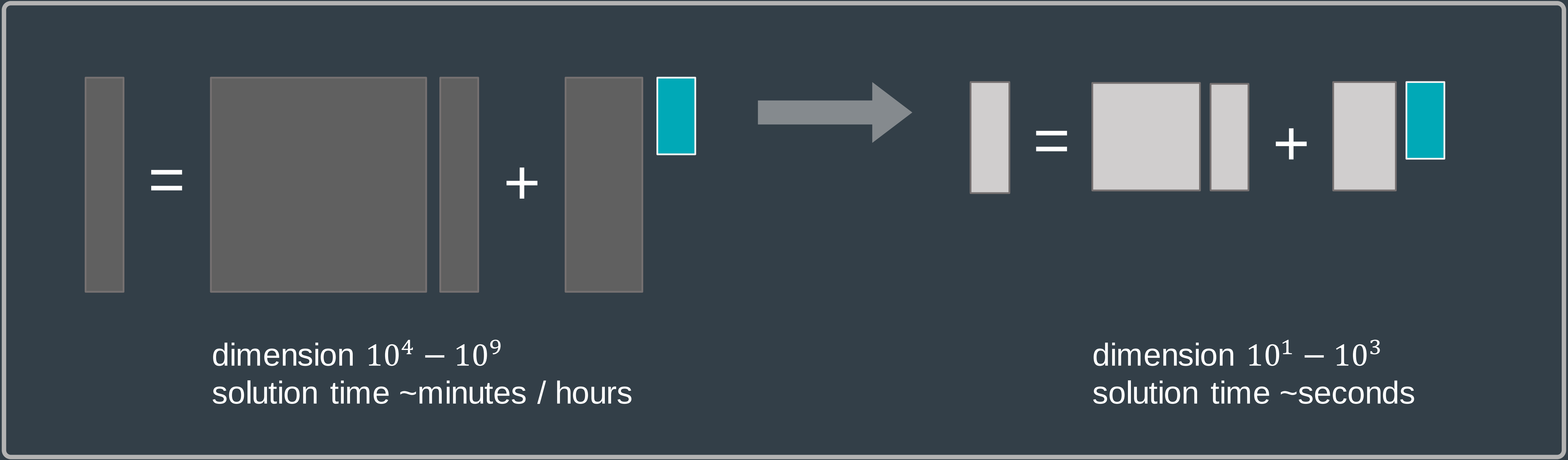


self-aware aircraft

Physics-based models

- simulate new previously unseen scenarios
- obey the laws of physics
- have quantifiable uncertainty
- parameters represent real-world quantities

But physics-based models are too complex and too expensive for use in near real-time onboard decision-making…



Projection-based model reduction

1 Train: Solve PDEs to generate training data

2 Identify structure: Compute a low-dimensional basis

3 Reduce: Project PDE model onto the low-dimensional subspace

Machine learning

"The scientific study of algorithms & statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns & inference instead." [Wikipedia]

Reduced-order modeling

"Model order reduction (MOR) is a technique for reducing the computational complexity of mathematical models in numerical simulations." [Wikipedia]

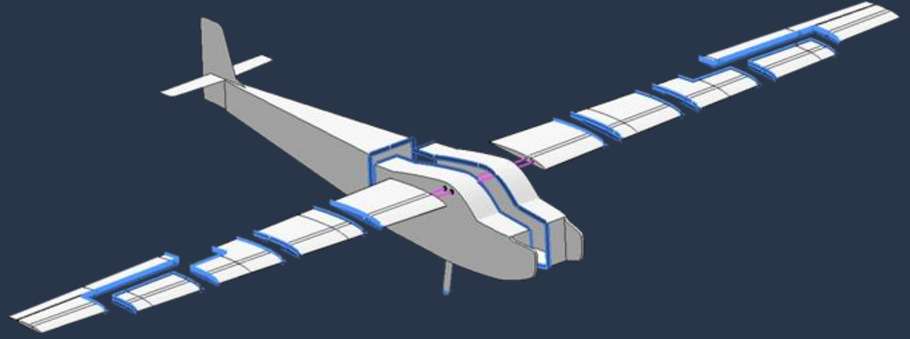
What is the connection between reduced-order modeling and machine learning?

Model reduction methods have grown from CSE, with a focus on ***reducing* high-dimensional models** that arise from physics-based modeling, whereas machine learning has grown from CS, with a focus on ***creating* low-dimensional models** from black-box data streams.

[Swischuk et al., *Computers & Fluids*, 2018]

Can we get the best of both worlds?

Reduced-order modeling leads to low-cost physics-based models that enable predictive digital twins



Challenges & limitations

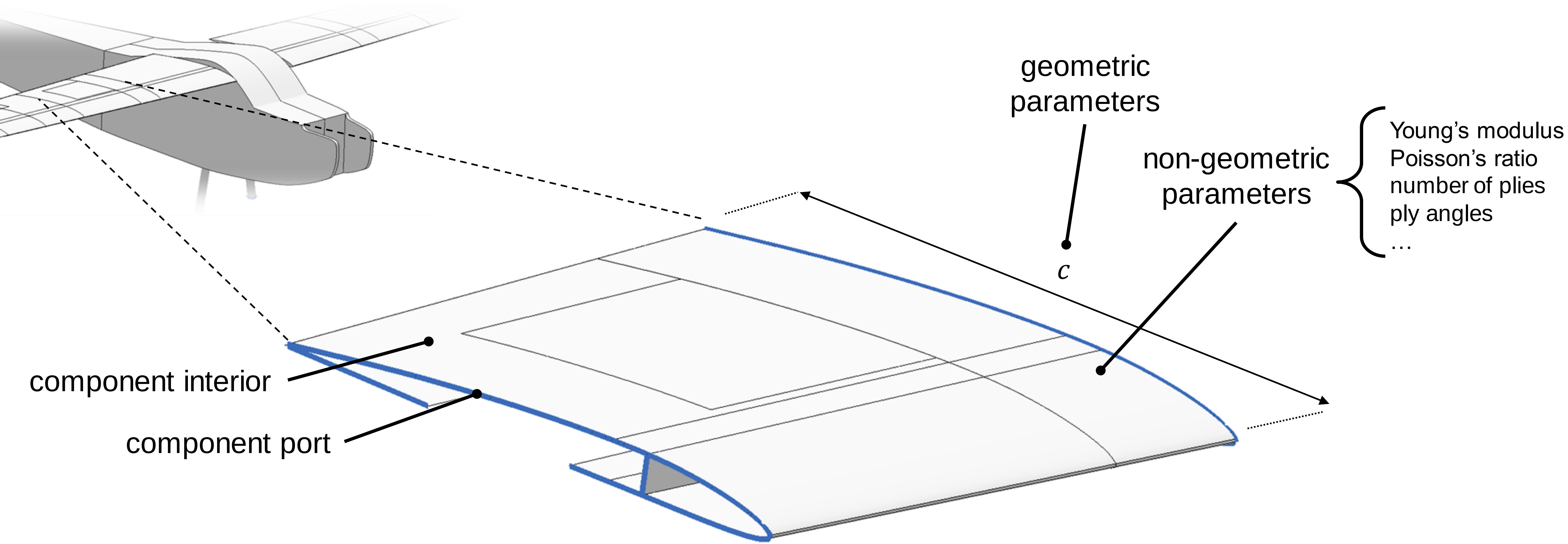
- **training** is expensive
- scaling to **high-dimensional parameters**
- dealing with **discontinuous parameter** dependence

Approach

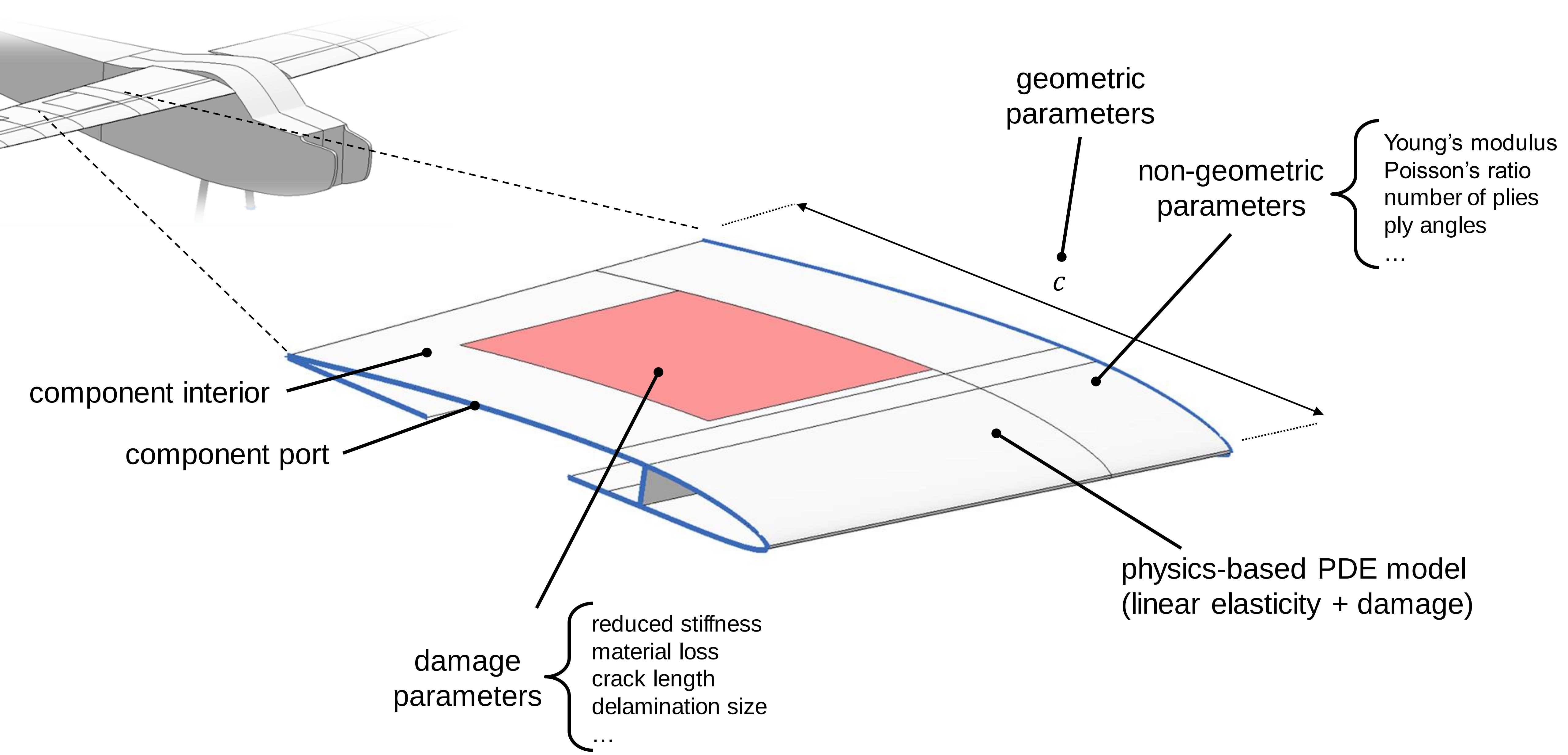
- Static-Condensation Reduced-Basis-Element (SCRBE) method [Huynh 2013]

"Divide and conquer"

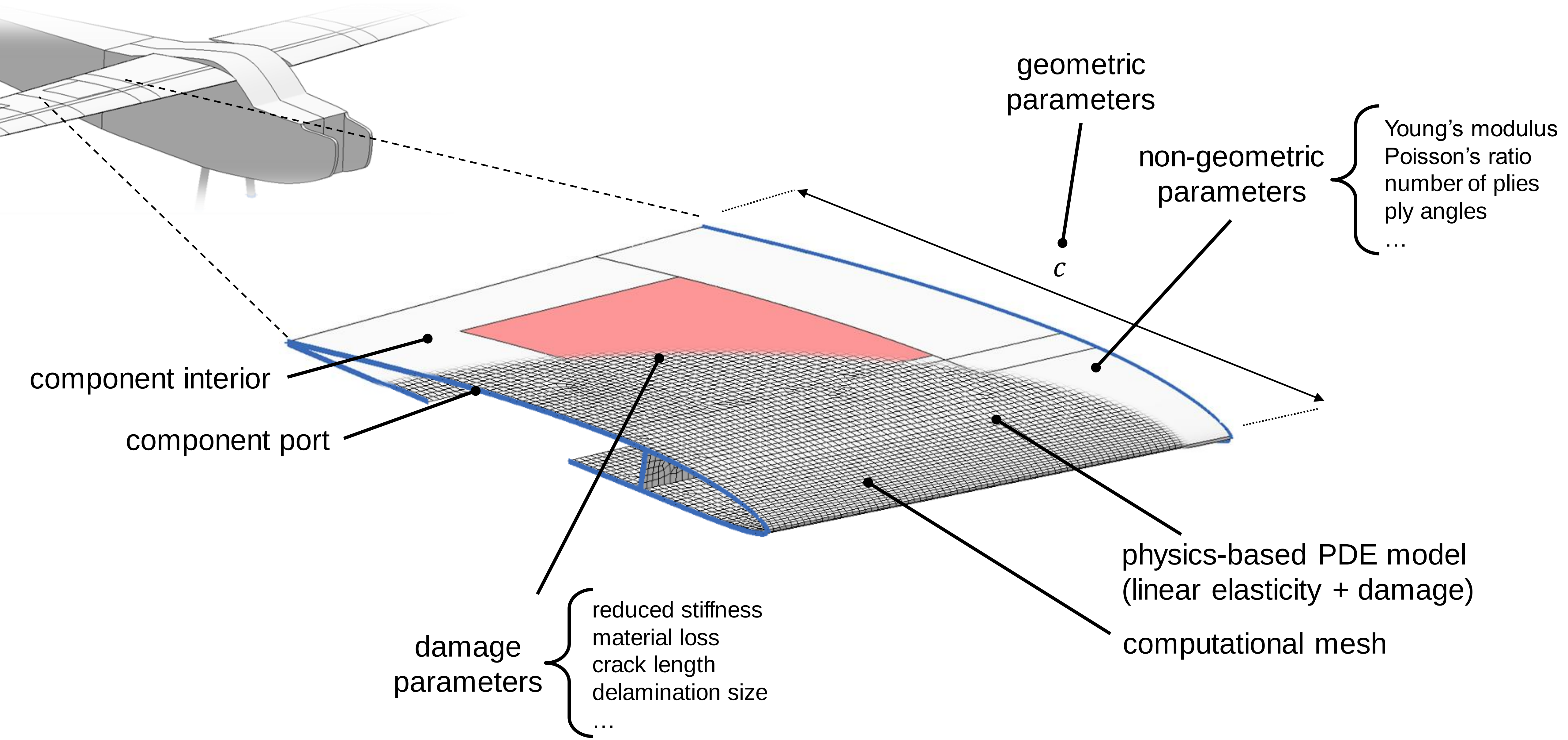
Example component: section of a wing

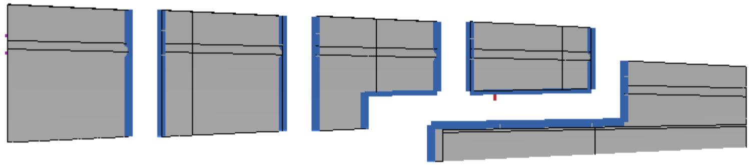


Example component: section of a wing

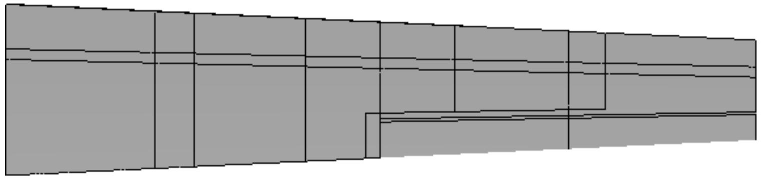


Example component: section of a wing

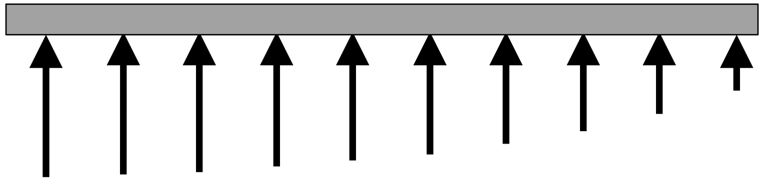




local effects
(component parameters)

+ 

interactions
(assembly parameters)

+ 

context
(loads parameters)

A complex nonlinear system is more than just the sum of its pieces

Flight test vehicle

Customized 12ft Telemaster aircraft

Custom wing sets: pristine & damaged

Accelerometers + vibration + dynamic strain sensors

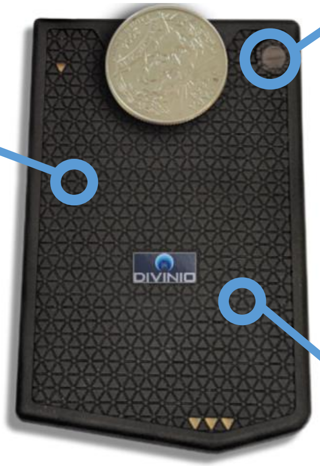
24 strain gauges per wing

INSTITUTE FOR COMPUTATIONAL ENGINEERING & SCIENCES ODEN

MIT AEROASTRO

Aurora FLIGHT SCIENCES A BOEING COMPANY

3 axis accelerometer
3 axis gyro



Temperature, pressure and humidity sensors

Dual high-frequency dynamic strain and vibration sensors

DIVINIO

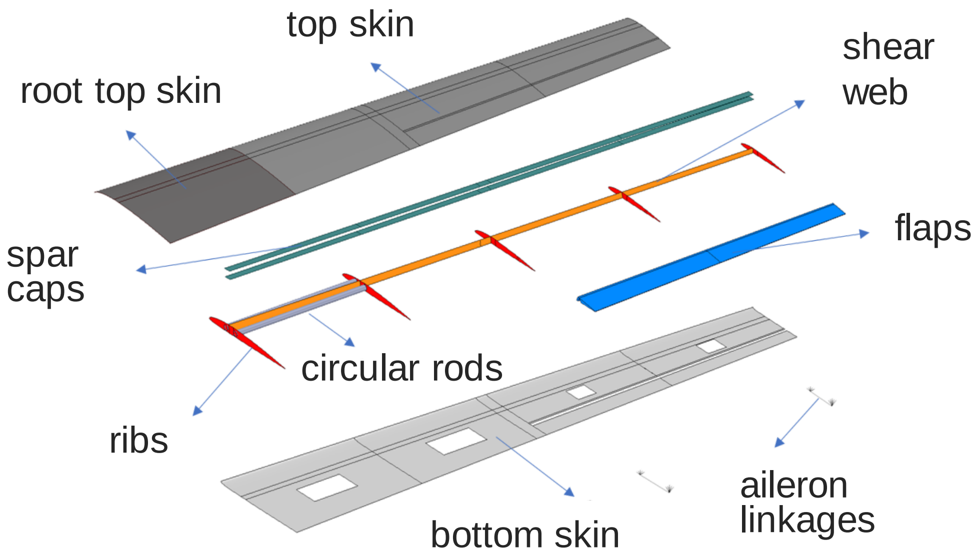
Internal structure

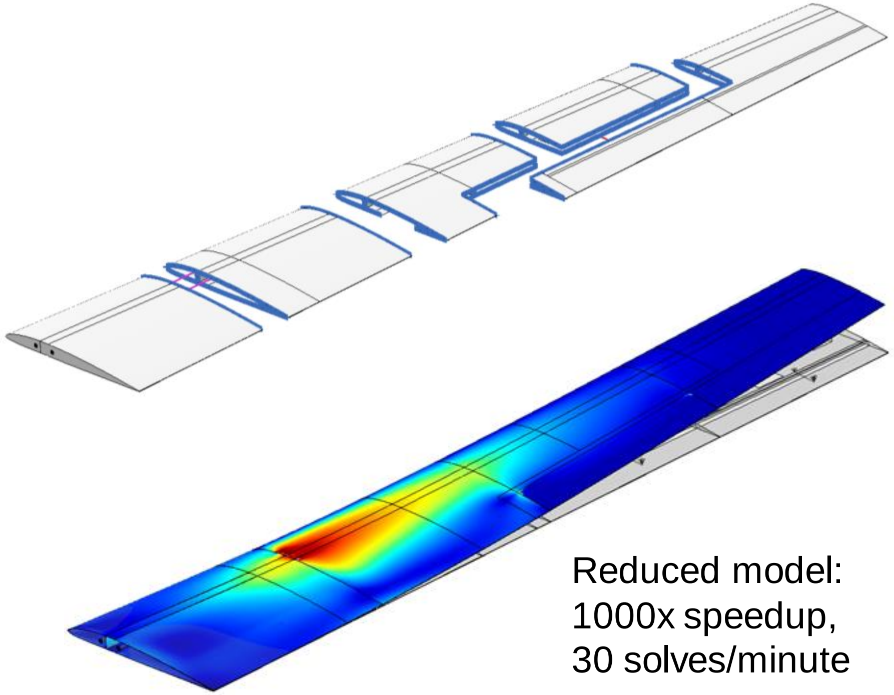
*Willcox has a family member who is co-founder of Divinio. Purchase of the sensors for use in the research was reviewed and approved in compliance with all applicable MIT policies and procedures.

Physics-based Digital Twin

Finite element model: multiple material types (carbon fiber, carbon rod, plywood, foam) & multiple element types (solid, shell, beam)

Reduced model: 0.03 seconds per structural analysis (cf. 55 seconds for the finite element model)





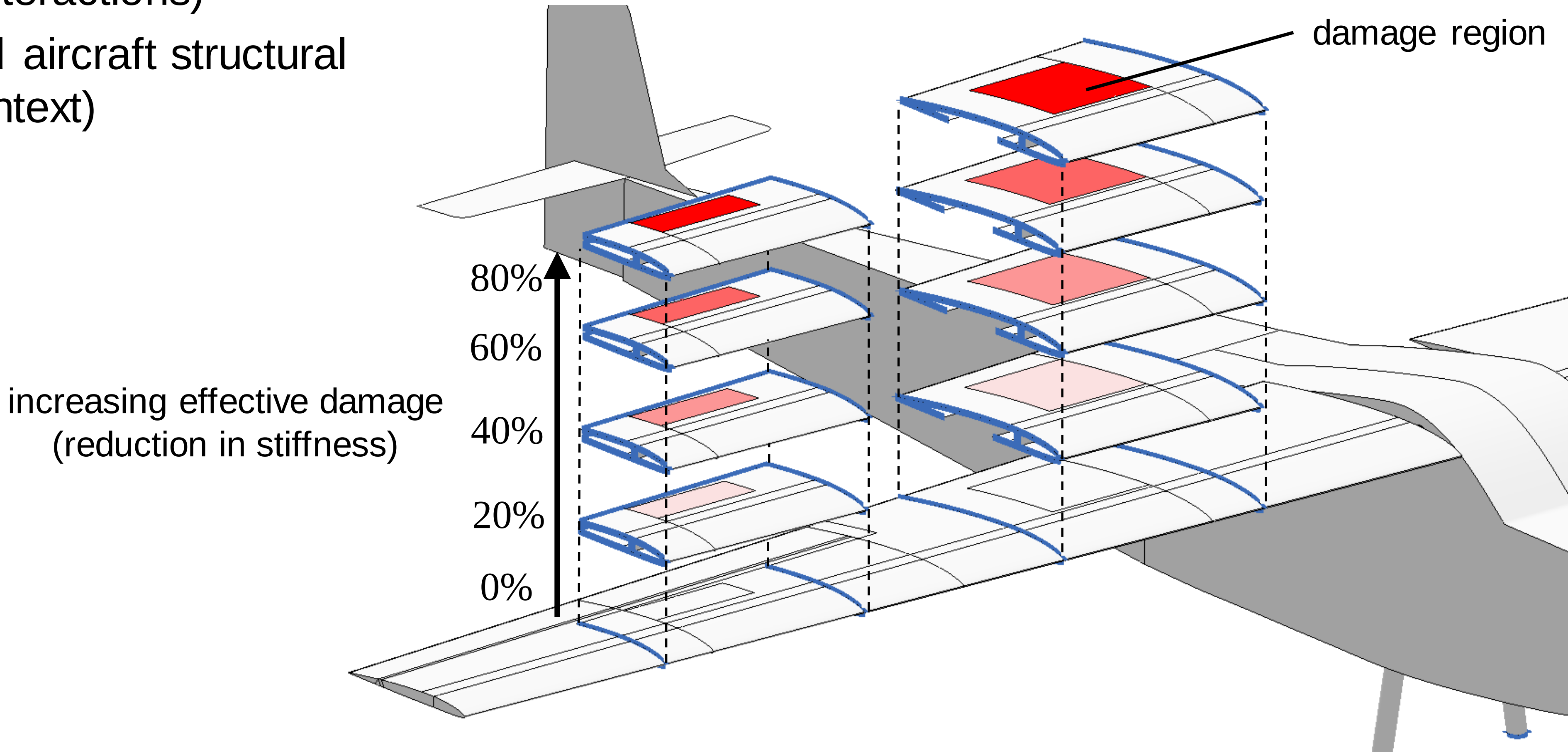
Reduced model:
1000x speedup,
30 solves/minute

Internal structure

From **component-based model** to **digital twin: Physics-based library**

Offline: Construct a library of damage states for each component

- Create multiple copies of each component
- Train components for parameter ranges of interest (local + interactions)
- Compute associated aircraft structural load constraints (context)

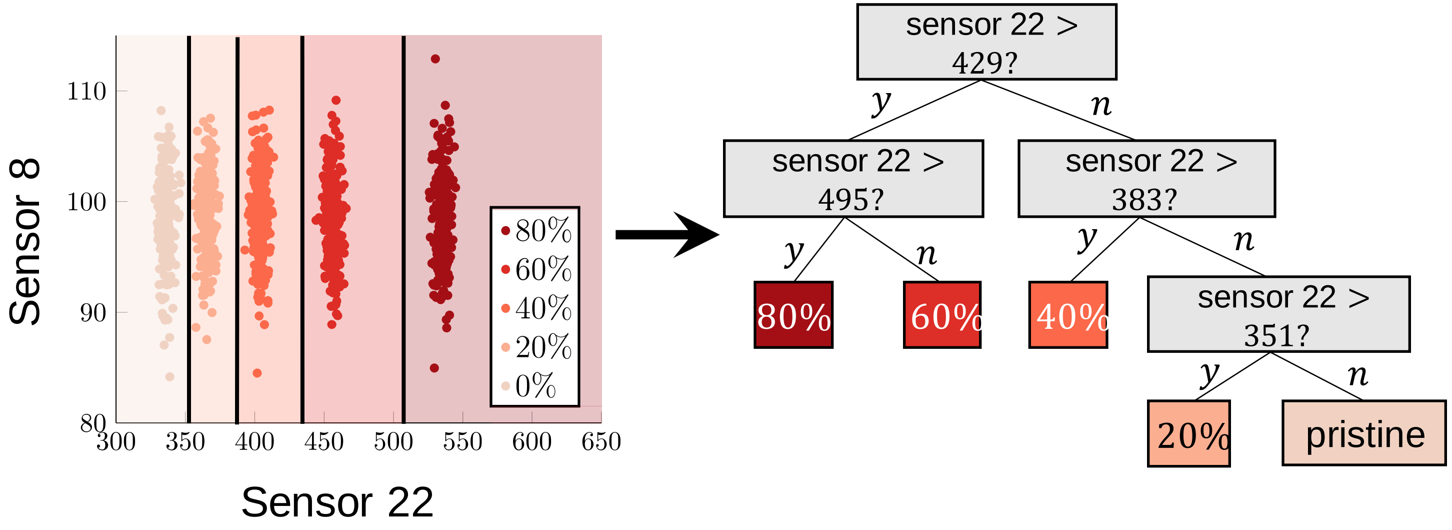


From **component-based model** to **digital twin: Interpretable machine learning**

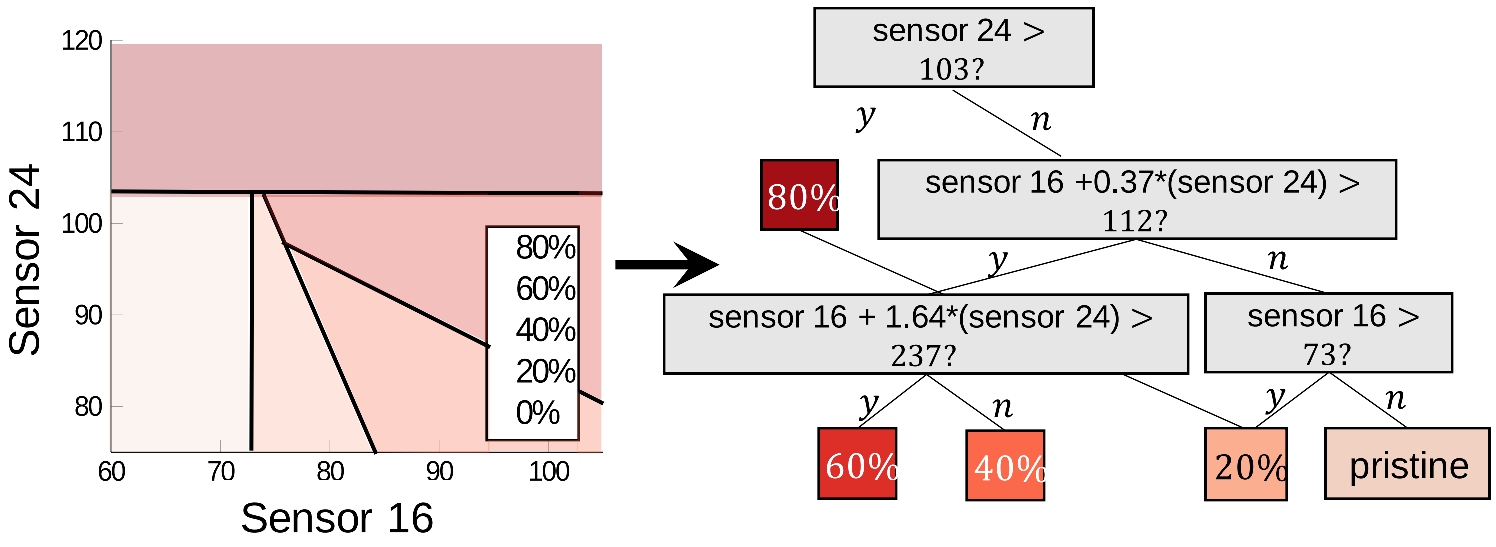
Offline: Train a classifier using simulation data

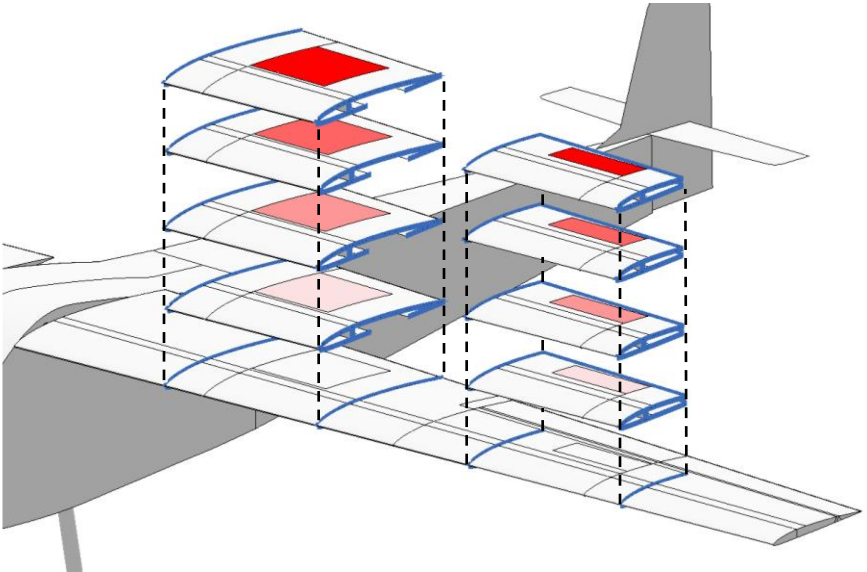
- Optimal Classification Trees
  [Bertsimas & Dunn, 2017]
- Highly interpretable
- Natural framework for sensor selection
- Online classification is rapid

Component 1

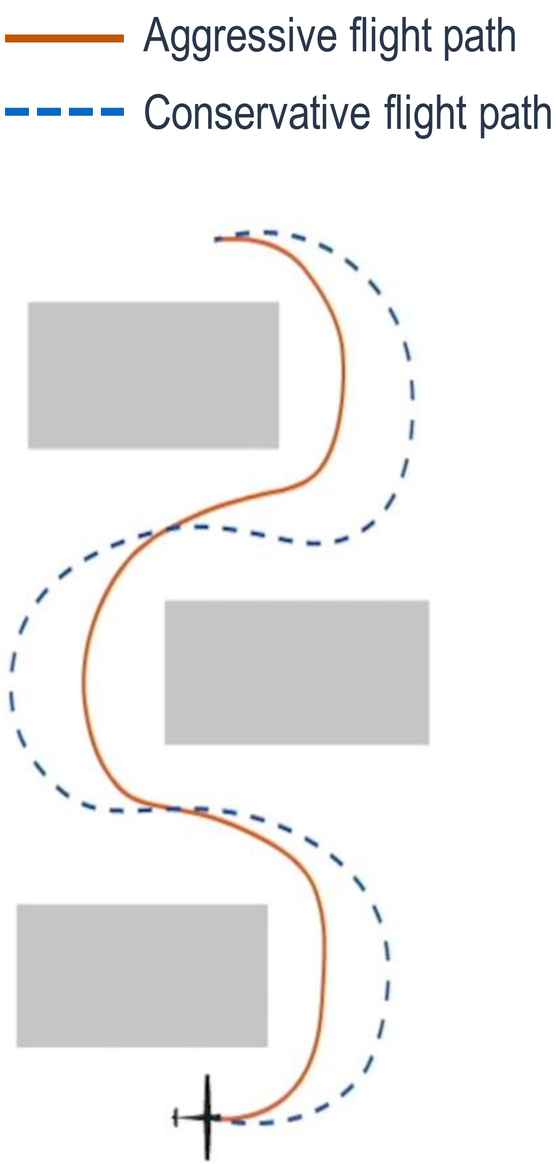


Component 2

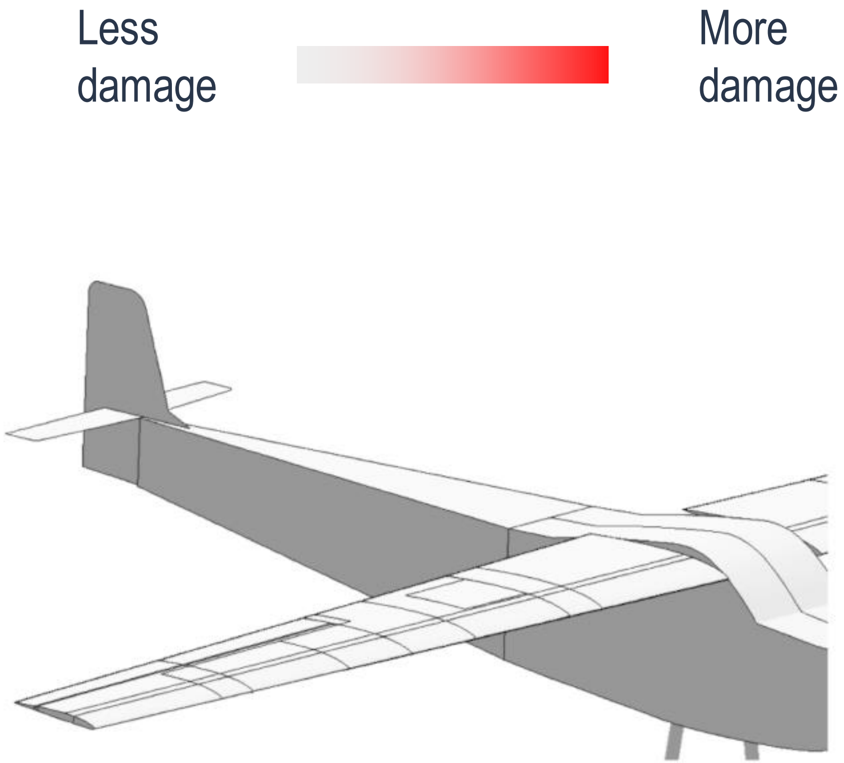




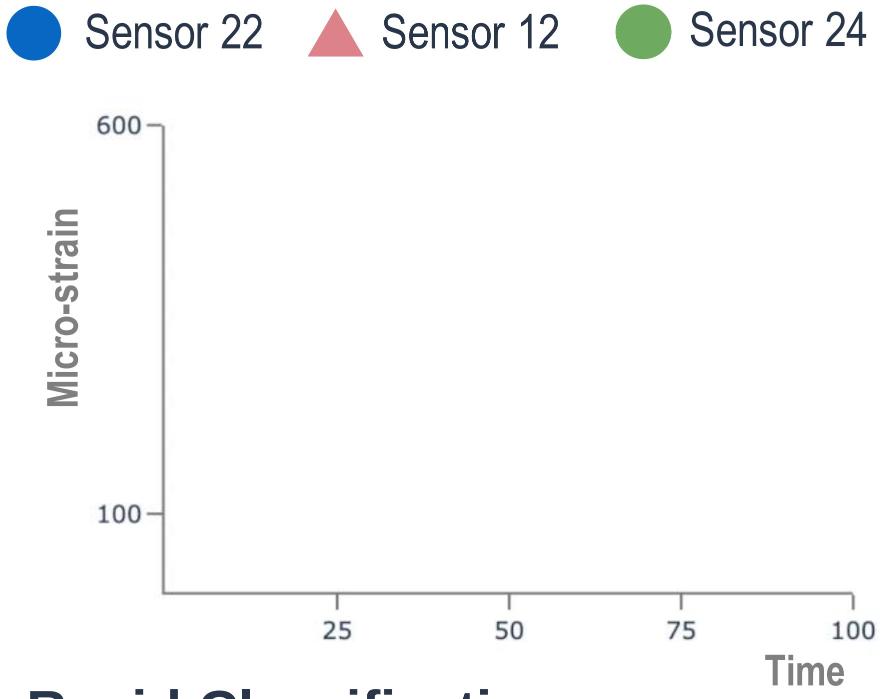
Flight of the UAV



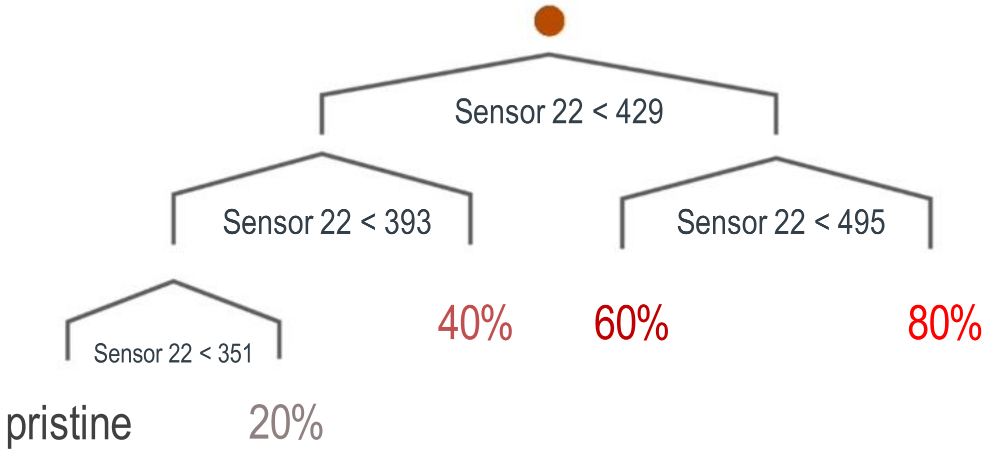
Health estimates



Strain Measurements



Rapid Classification

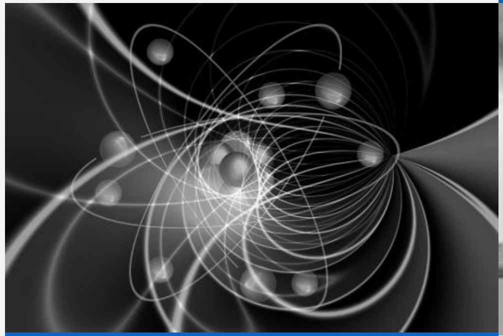


Conclusions & Outlook

From forward simulations to
predictive data science

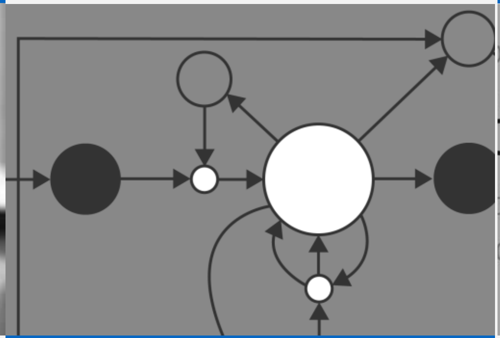
Predictive Data Science

Learning from data through the lens of models is a way to exploit structure in an otherwise intractable problem



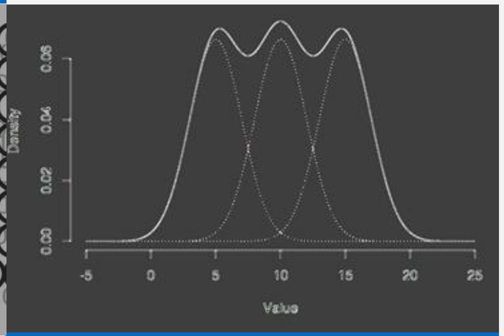
Respect physical constraints

Embed domain knowledge



Bring interpretability to results

Integrate heterogeneous, noisy & incomplete data



Get predictions with quantified uncertainties

Data Science

Computational Science & Engineering

Predictive Data Science

Revolutionizing decision-making for **high-consequence applications** in science, engineering & medicine

Data Science

Computational Science & Engineering

Predictive Data Science

Needs **interdisciplinary research & education** at the interfaces of computer science, mathematics, statistics, high performance computing, and applications across science, engineering and medicine

Interdisciplinary research & education in computational engineering & sciences

developing high-performance computing solutions to society's big problems

ODEN.UTEXAS.EDU

ODEN INSTITUTE
FOR COMPUTATIONAL ENGINEERING & SCIENCES