# Data and Models - Introductory Comments

Wolfgang Dahmen

DASIV Spring-School Feb 22 - 25, 2018

# Contents



UNIVERSITY OF SOUTH CAROLINA

# Contents

# Contents

# Contents

# Outline



UNIVERSITY OF SOUTH CAROLINA

# Contents



UNIVERSITY OF SOUTH CAROLINA

# Guiding Theme

- Mathematical modeling and simulation affect virtually all branches of science

# Guiding Theme

- Mathematical modeling and simulation affect virtually all branches of science
- Models are getting increasingly complex

UNIVERSITY OF SOUTH CAROLINA

# Guiding Theme

- Mathematical modeling and simulation affect virtually all branches of science
- Models are getting increasingly complex
- Joint efforts: natural sciences, engineering, computer science,..., mathematics

# Guiding Theme

- Mathematical modeling and simulation affect virtually all branches of science
- Models are getting increasingly complex
- Joint efforts: natural sciences, engineering, computer science,..., mathematics ... the art of "shortcuts"

UNIVERSITY OF SOUTH CAROLINA

# Guiding Theme

- Mathematical modeling and simulation affect virtually all branches of science
- Models are getting increasingly complex
- Joint efforts: natural sciences, engineering, computer science,..., mathematics ... the art of ”shortcuts”
- ”Optimal Computation”

UNIVERSITY OF SOUTH CAROLINA

# Guiding Theme

- Mathematical modeling and simulation affect virtually all branches of science
- Models are getting increasingly complex
- Joint efforts: natural sciences, engineering, computer science,..., mathematics ... the art of ”shortcuts”
- ”Optimal Computation”
- What does ”optimality” mean?

# Guiding Theme

- Mathematical modeling and simulation affect virtually all branches of science
- Models are getting increasingly complex
- Joint efforts: natural sciences, engineering, computer science,..., mathematics ... the art of ”shortcuts”
- ”Optimal Computation”
- What does ”optimality” mean?
  - identify relevant model - benchmark scenarios

UNIVERSITY OF SOUTH CAROLINA

# Guiding Theme

- Mathematical modeling and simulation affect virtually all branches of science
- Models are getting increasingly complex
- Joint efforts: natural sciences, engineering, computer science,..., mathematics ... the art of ”shortcuts”
- ”Optimal Computation”
- What does ”optimality” mean?
  - identify relevant model - benchmark scenarios
  - what can be achieved at best?

UNIVERSITY OF SOUTH CAROLINA

# Guiding Theme

- Mathematical modeling and simulation affect virtually all branches of science
- Models are getting increasingly complex
- Joint efforts: natural sciences, engineering, computer science,..., mathematics ... the art of ”shortcuts”
- ”Optimal Computation”
- What does ”optimality” mean?
  - identify relevant model - benchmark scenarios
  - what can be achieved at best?
  - try to contrive algorithms meeting these lower bounds

UNIVERSITY OF SOUTH CAROLINA

# Guiding Theme

- Mathematical modeling and simulation affect virtually all branches of science
- Models are getting increasingly complex
- Joint efforts: natural sciences, engineering, computer science,..., mathematics ... the art of ”shortcuts”
- ”Optimal Computation”
- What does ”optimality” mean?
  - identify relevant model - benchmark scenarios
  - what can be achieved at best?
  - try to contrive algorithms meeting these lower bounds

- Discrete $\leftrightarrow$ continuous ($\infty$-dimensional) settings;
- nonlinear and adaptive concepts;
- optimization

UNIVERSITY OF SOUTH CAROLINA

# Playgrounds

Data (-Science): extract information from observations/measurements ...

# Playgrounds

Data (-Science): extract information from observations/measurements ...

- social networks
- autonomous driving
- structural imaging (microscopy, spectroscopy,...), etc.

# Playgrounds

Data (-Science): extract information from observations/measurements ...

$\rightsquigarrow$ "Big-Data Problem"

- social networks
- autonomous driving
- structural imaging (microscopy, spectroscopy,...), etc.

UNIVERSITY OF SOUTH CAROLINA

# Playgrounds

Data (-Science): extract information from observations/measurements ...

$\rightsquigarrow$ "Big-Data Problem"

- social networks
- autonomous driving
- structural imaging (microscopy, spectroscopy,...), etc.

Models: mathematically "encoding" real world processes approximately

# Playgrounds

Data (-Science): extract information from observations/measurements ...

$\rightsquigarrow$ "Big-Data Problem"

- social networks
- autonomous driving
- structural imaging (microscopy, spectroscopy,...), etc.

Models: mathematically "encoding" real world processes approximately

- atomistic models (micro-scale)
- kinetic models (meso-scale)
- continuum models, balance laws (macro-scale)
- multiscale models

UNIVERSITY OF SOUTH CAROLINA

# Playgrounds

Data (-Science): extract information from observations/measurements ...

$\rightsquigarrow$ "Big-Data Problem"

- social networks
- autonomous driving
- structural imaging (microscopy, spectroscopy,...), etc.

Models: mathematically "encoding" real world processes approximately

$\rightsquigarrow$ prediction capability

- atomistic models (micro-scale)
- kinetic models (meso-scale)
- continuum models, balance laws (macro-scale)
- multiscale models

UNIVERSITY OF SOUTH CAROLINA

# Playgrounds

Data (-Science): extract information from observations/measurements ...

$\rightsquigarrow$ "Big-Data Problem"

- social networks
- autonomous driving
- structural imaging (microscopy, spectroscopy,...), etc.

Models: mathematically "encoding" real world processes approximately

$\rightsquigarrow$ prediction capability

- atomistic models (micro-scale)
- kinetic models (meso-scale)
- continuum models, balance laws (macro-scale)
- multiscale models

Quantify information/prediction accuracy

UNIVERSITY OF SOUTH CAROLINA

# ... Combine Models and Data ... why?

- models are often incomplete or deficient

# ... Combine Models and Data ... why?

- models are often incomplete or deficient
- models may depend on (many) parameters that need to be determined

# ... Combine Models and Data ... why?

- models are often incomplete or deficient
- models may depend on (many) parameters that need to be determined
- there may be too few data because the searched for functional interrelation depends on too many parameters
  ... a model may then serve as a prior

# ... Combine Models and Data ... why?

- models are often incomplete or deficient
- models may depend on (many) parameters that need to be determined
- there may be too few data because the searched for functional interrelation depends on too many parameters
  ... a model may then serve as a prior

Examples:

# ... Combine Models and Data ... why?

- models are often incomplete or deficient
- models may depend on (many) parameters that need to be determined
- there may be too few data because the searched for functional interrelation depends on too many parameters
  ... a model may then serve as a prior

Examples:

- meteorology: climate modeling, Hurricane prediction, weather forecast,....
- underground exploration
- structural imaging, material science

# ... Combine Models and Data ... why?

- models are often incomplete or deficient
- models may depend on (many) parameters that need to be determined
- there may be too few data because the searched for functional interrelation depends on too many parameters
  ... a model may then serve as a prior

Examples:

- meteorology: climate modeling, Hurricane prediction, weather forecast,....
- underground exploration
- structural imaging, material science

Model identification, data assimilation

# ... Combine Models and Data ... why?

- models are often incomplete or deficient
- models may depend on (many) parameters that need to be determined
- there may be too few data because the searched for functional interrelation depends on too many parameters
  ... a model may then serve as a prior

Examples:

- meteorology: climate modeling, Hurricane prediction, weather forecast,....
- underground exploration
- structural imaging, material science

Model identification, data assimilation

**Issues:**

- High spatial dimensionality

# ... Combine Models and Data ... why?

- models are often incomplete or deficient
- models may depend on (many) parameters that need to be determined
- there may be too few data because the searched for functional interrelation depends on too many parameters
  ... a model may then serve as a prior

Examples:

- meteorology: climate modeling, Hurricane prediction, weather forecast,....
- underground exploration
- structural imaging, material science

Model identification, data assimilation

**Issues:**

- High spatial dimensionality
- What are proper metrics for data and model outputs?

# ... Combine Models and Data ... why?

- models are often incomplete or deficient
- models may depend on (many) parameters that need to be determined
- there may be too few data because the searched for functional interrelation depends on too many parameters
  ... a model may then serve as a prior

Examples: $\rightsquigarrow$ ”Small-Data Problem”

- meteorology: climate modeling, Hurricane prediction, weather forecast,....
- underground exploration
- structural imaging, material science

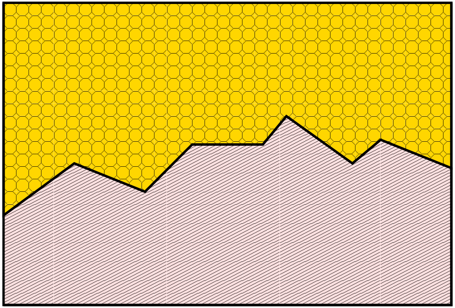
Model identification, data assimilation

**Issues:**

- High spatial dimensionality
- What are proper metrics for data and model outputs?

# Contents

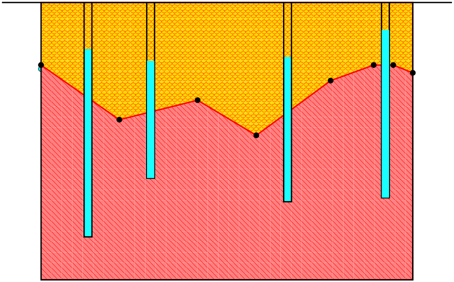
# Porous Media Flow



Model:

$$-\mathrm{div}\,(a(y)\nabla u) = f \quad (+b.c.'s)$$

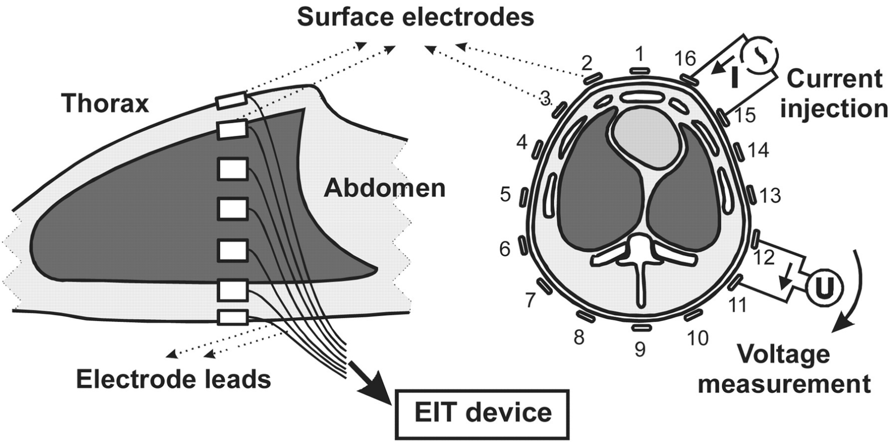
# Porous Media Flow



Model:

$$-\mathrm{div}\,(a(y)\nabla u) = f \quad (+b.c.'s)$$

# EIT




UNIVERSITY OF SOUTH CAROLINA

# Outline

# Data Driven Approaches

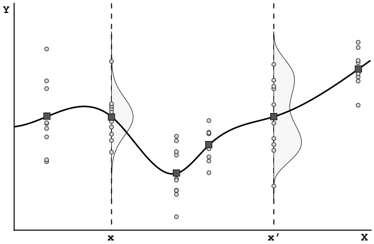
Big Data Problem

- Given data/observations/measurements
  $\mathcal{Z}_n = \{(z_i = (x_i, y_i) : i = 1, \ldots, n\} \subset \mathbb{X} \times \mathbb{Y}$

  learn ”functional law”: $f : \mathbb{X} \to \mathbb{Y}$, $f(x_i) \approx y_i$, explaining the data
- Stochastic model: $z_i$ i.i.d. with respect to (often) unknown probability density $\rho$ on $\mathbb{X} \times \mathbb{Y}$
- Typical goal: find an estimator $\hat{f}_{\mathcal{Z}_n} : \mathbb{X} \to \mathbb{Y}$ that approximates $f$ in a certain sense
- $\mathbb{Y} =$ continuum $\rightsquigarrow$ regression
- $\#\mathbb{Y} < \infty \rightsquigarrow$ classification

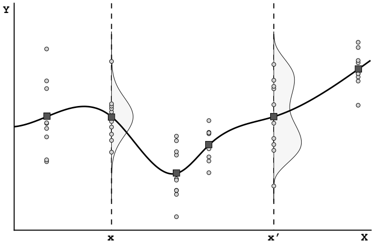
UNIVERSITY OF SOUTH CAROLINA

# Regression

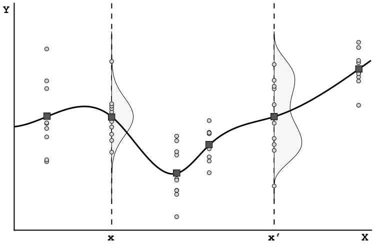


# Regression

$$d\rho(x,y) = d\rho(y|x)d\rho_X(x)$$



# Regression



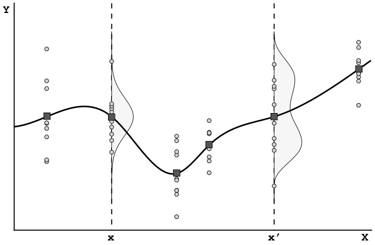

$$d\rho(x,y) = d\rho(y|x)d\rho_X(x)$$

regression function

$$f_\rho(x) := \int_Y y d\rho(y|x) = \mathbb{E}(y|x)$$

UNIVERSITY OF SOUTH CAROLINA

# Regression




$$d\rho(x,y) = d\rho(y|x)d\rho_X(x)$$

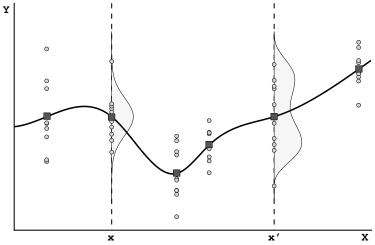
regression function

$$f_\rho(x) := \int_Y y d\rho(y|x) = \mathbb{E}(y|x)$$

Risk functional:

$$\mathcal{E}(f) := \int_{\mathbb{X}\times\mathbb{Y}} (y - f(x))^2 d\rho$$

# Regression




$$d\rho(x,y) = d\rho(y|x)d\rho_X(x)$$

regression function

$$f_\rho(x) := \int_Y y d\rho(y|x) = \mathbb{E}(y|x)$$

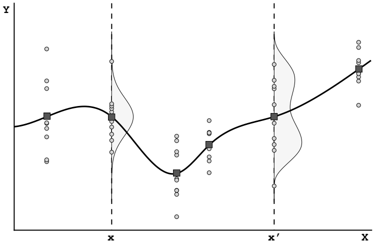
Risk functional:

$$\mathcal{E}(f) := \int_{\mathbb{X}\times\mathbb{Y}} (y - f(x))^2 d\rho$$

$\rightsquigarrow$

$$\mathcal{E}(f) = \mathcal{E}(f_\rho) + \|f - f_\rho\|^2_{L_2(\mathbb{X},\rho_{\mathbb{X}})}, \quad \|\cdot\| := \|\cdot\|_{L_2(\mathbb{X},\rho_{\mathbb{X}})}$$

# Regression




$$d\rho(x,y) = d\rho(y|x)d\rho_X(x)$$

regression function

$$f_\rho(x) := \int_Y y d\rho(y|x) = \mathbb{E}(y|x)$$

Risk functional:

$$\mathcal{E}(f) := \int_{\mathbb{X}\times\mathbb{Y}} (y - f(x))^2 d\rho$$

$\rightsquigarrow$

$$\mathcal{E}(f) = \mathcal{E}(f_\rho) + \|f - f_\rho\|^2_{L_2(\mathbb{X},\rho_{\mathbb{X}})}, \quad \|\cdot\| := \|\cdot\|_{L_2(\mathbb{X},\rho_{\mathbb{X}})}$$

Ideal desirable performance bound:

$$\mathbb{P}_{\rho^n}\left\{\mathcal{Z} : \|f_\rho - \hat{f}_{\mathcal{Z}}\| \geq c(d)\left(\frac{\log n}{n}\right)^{\frac{s}{2s+1}}\right\} \leq Cn^{-\beta}$$

UNIVERSITY OF SOUTH CAROLINA

# Issues

- meaning of ”best estimators” - suitable continuous model class

# Issues

- meaning of ”best estimators” - suitable continuous model class
- universality, what does the estimator have to ”know” to perform best

# Issues

- meaning of ”best estimators” - suitable continuous model class
- universality, what does the estimator have to ”know” to perform best
- role of nonlinearity, adaptivity

# Issues

- meaning of ”best estimators” - suitable continuous model class
- universality, what does the estimator have to ”know” to perform best
- role of nonlinearity, adaptivity
- high spatial dimensionality

# Issues

- meaning of ”best estimators” - suitable continuous model class
- universality, what does the estimator have to ”know” to perform best
- role of nonlinearity, adaptivity
- high spatial dimensionality

Applications: imaging, spectral imaging, learning noise models

# Model Based Approaches

Modeling: Physical, technological process described by $u$: find "balance law" describing $u$

$$F(u, f) = 0$$

where $F(u, f) = G(u, \partial_t u, \partial_x^\alpha u) - f$ is a PDE, integral equation, etc.

# Model Based Approaches

Modeling: Physical, technological process described by $u$: find "balance law" describing $u$

$$F(u, y, f) = 0, \quad y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$$

where $F(u, f) = G(u, \partial_t u, \partial_x^\alpha u) - f$ is a PDE, integral equation, etc.

UNIVERSITY OF SOUTH CAROLINA

# Model Based Approaches

Modeling: Physical, technological process described by $u$: find "balance law" describing $u$

$$F(u, y, f) = 0, \quad y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$$

where $F(u, f) = G(u, \partial_t u, \partial_x^\alpha u) - f$ is a PDE, integral equation, etc.

Forward problem: given $f$ (and $y$), find $u$ s.t.

$$F(u, f) = 0 \qquad \text{e.g.} \quad -\Delta u = f \text{ in } \Omega, \; u|_{\partial\Omega} = 0$$

# Model Based Approaches

Modeling: Physical, technological process described by $u$: find "balance law" describing $u$

$$F(u, y, f) = 0, \quad y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$$

where $F(u, f) = G(u, \partial_t u, \partial_x^\alpha u) - f$ is a PDE, integral equation, etc.

Forward problem: given $f$ (and $y$), find $u$ s.t.

$$F(u, f) = 0 \qquad \text{e.g.} \quad -\Delta u = f \text{ in } \Omega, \; u|_{\partial\Omega} = 0$$

Discretization: Contrive a discrete model

$$F_h(u_h, , f) = 0$$

# Model Based Approaches

Modeling: Physical, technological process described by $u$: find "balance law" describing $u$

$$F(u, y, f) = 0, \quad y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$$

where $F(u, f) = G(u, \partial_t u, \partial_x^\alpha u) - f$ is a PDE, integral equation, etc.

Forward problem: given $f$ (and $y$), find $u$ s.t.

$$F(u, f) = 0 \qquad \text{e.g.} \quad -\Delta u = f \text{ in } \Omega, \; u|_{\partial\Omega} = 0$$

Discretization: Contrive a discrete model

$$F_h(u_h, , f) = 0$$

Issues:

- fixed discretization: efficient solvers, preconditioning

# Model Based Approaches

Modeling: Physical, technological process described by $u$: find ”balance law” describing $u$

$$F(u, y, f) = 0, \quad y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$$

where $F(u, f) = G(u, \partial_t u, \partial_x^\alpha u) - f$ is a PDE, integral equation, etc.

Forward problem: given $f$ (and $y$), find $u$ s.t.

$$F(u, f) = 0 \qquad \text{e.g.} \quad -\Delta u = f \text{ in } \Omega,\ u|_{\partial\Omega} = 0$$

Discretization: Contrive a discrete model

$$F_h(u_h, , f) = 0$$

Issues:

- optimize discretization, adaptive solvers

UNIVERSITY OF SOUTH CAROLINA

# Model Based Approaches

Modeling: Physical, technological process described by $u$: find "balance law" describing $u$

$$F(u, y, f) = 0, \quad y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$$

where $F(u, f) = G(u, \partial_t u, \partial_x^\alpha u) - f$ is a PDE, integral equation, etc.

Forward problem: given $f$ (and $y$), find $u$ s.t.

$$F(u, f) = 0 \qquad \text{e.g.} \quad -\Delta u = f \text{ in } \Omega, \; u|_{\partial\Omega} = 0$$

Discretization: Contrive a discrete model

$$F_h(u_h, y, f) = 0 \quad y \in \mathcal{Y}$$

Issues:

- high-dimensionality

UNIVERSITY OF SOUTH CAROLINA

# Model Based Approaches

Modeling: Physical, technological process described by $u$: find "balance law" describing $u$

$$F(u, y, f) = 0, \quad y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$$

where $F(u, f) = G(u, \partial_t u, \partial_x^\alpha u) - f$ is a PDE, integral equation, etc.

Forward problem: given $f \in \mathbb{F} = ???$ (and $y$), find $u \in \mathbb{U} = ???$ s.t.

$$F(u, f) = 0 \qquad \text{is well conditioned}$$

Discretization: Contrive a discrete model

$$F_h(u_h, y, f) = 0 \quad y \in \mathcal{Y}$$

Issues:

- choice of $\mathbb{U}, \mathbb{F}$ finite $\leftrightarrow$ infinite-dimensional problem

# Ideal Stability

$F(u, f) = 0$ : $\bar{u}$ approximation to $u$ - would like to have:

$$\|u - \bar{u}\|_{\mathbb{U}} \eqsim \|F(\bar{u}, f)\|_{???}$$

# Ideal Stability

$F(u,f)=0$ : $\bar{u}$ approximation to $u$ - would like to have:

$$\|u-\bar{u}\|_{\mathbb{U}} \eqsim \|F(\bar{u},f)\|_{\mathbb{V}'}$$

How to get this?

$$\langle F(u,f), v\rangle = 0 \ \forall\, v \in \mathbb{V}$$

# Ideal Stability

$F(u,f)=0$ : $\bar{u}$ approximation to $u$ - would like to have:

$$\|u-\bar{u}\|_{\mathbb{U}} \eqsim \|F(\bar{u},f)\|_{\mathbb{V}'}$$

How to get this?

$$\langle F(u,f),v\rangle = 0\ \forall\, v\in\mathbb{V} \quad \text{is "good" if } F(\cdot,f): w\to F(w,f)\in\mathbb{V}'$$

- $F$ nonlinear $\rightsquigarrow$ linearize

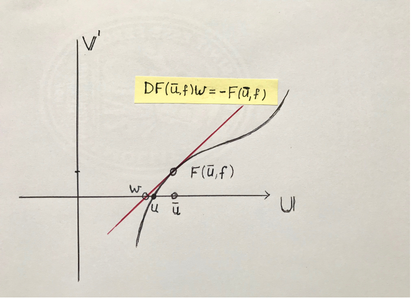
# Ideal Stability

$F(u,f) = 0$ : $\bar{u}$ approximation to $u$ - would like to have:

$$\|u - \bar{u}\|_{\mathbb{U}} \eqsim \|F(\bar{u},f)\|_{\mathbb{V}'}$$

How to get this?

$$\langle F(u,f), v\rangle = 0 \ \forall\, v \in \mathbb{V} \quad \text{is "good" if } F(\cdot,f) : w \to F(w,f) \in \mathbb{V}'$$



# Ideal Stability - well conditioned variational formulations

$F(u,f)=0$ : $\bar{u}$ approximation to $u$ - would like to have:

$$\|u-\bar{u}\|_{\mathbb{U}} \eqsim \|F(\bar{u},f)\|_{\mathbb{V}'} \eqsim \Big(\sum_{K\in\mathcal{T}} \eta_K(\bar{u},f)^2\Big)^{1/2}$$

How to get this?

$$\langle F(u,f),v\rangle = 0 \ \forall\, v\in\mathbb{V} \quad \text{is "good" if } F(\cdot,f): w\to F(w,f)\in\mathbb{V}'$$

- $F$ nonlinear $\rightsquigarrow$ linearize
- $F(u,f)=\mathcal{B}u-f$ linear $\rightsquigarrow$ Babuska-Banach-Necas Theorem
  inf-sup condition $\rightsquigarrow$ DPG framework

# Ideal Stability - well conditioned variational formulations

$F(u,f)=0$ : $\bar{u}$ approximation to $u$ - would like to have:

$$\|u-\bar{u}\|_{\mathbb{U}} \eqsim \|F(\bar{u},f)\|_{\mathbb{V}'} \eqsim \Big(\sum_{K\in\mathcal{T}} \eta_K(\bar{u},f)^2\Big)^{1/2}$$

How to get this?

$$\langle F(u,f),v\rangle = 0 \;\forall\, v\in\mathbb{V} \quad \text{is "good" if } \; F(\cdot,f): w\to F(w,f)\in\mathbb{V}'$$

- $F$ nonlinear $\rightsquigarrow$ linearize
- $F(u,f)=\mathcal{B}u-f$ linear $\rightsquigarrow$ Babuska-Banach-Necas Theorem
  inf-sup condition $\rightsquigarrow$ DPG framework
- A posteriori error indicators that drive adaptive refinements

UNIVERSITY OF SOUTH CAROLINA

# Ideal Stability - well conditioned variational formulations

$F(u,f)=0$ : $\bar{u}$ approximation to $u$ - would like to have:

$$\|u-\bar{u}\|_{\mathbb{U}} \eqsim \|F(\bar{u},f)\|_{\mathbb{V}'} \eqsim \Big(\sum_{K\in\mathcal{T}} \eta_K(\bar{u},f)^2\Big)^{1/2}$$

How to get this?

$$\langle F(u,f),v\rangle = 0 \;\forall\, v\in\mathbb{V} \quad \text{is "good" if } \; F(\cdot,f) : w \to F(w,f) \in \mathbb{V}'$$

- $F$ nonlinear $\rightsquigarrow$ linearize
- $F(u,f)=\mathcal{B}u-f$ linear $\rightsquigarrow$ Babuska-Banach-Necas Theorem
  inf-sup condition $\rightsquigarrow$ DPG framework
- A posteriori error indicators that drive adaptive refinements
- Residuals $\rightsquigarrow$ surrogates for greedy constructions of
  reduced bases

# Example: Flooded City

Courtesy of Nils Gerhard and Siegfried Müller, RWTH Aachen, Germany

| | |
|---|---|
| Reference: | Experiment in a Channel |
| Computational Domain: | $36 \times 3.6\ m^2$ |
| Size of the Buildings: | $0.3 \times 0.3\ m^2$ |
| Width of Streets between Buildings: | $0.1\ m$ |
| $\#$ Buildings: | 25 |

# Example: Flooded City

Courtesy of Nils Gerhard and Siegfried Müller, RWTH Aachen, Germany

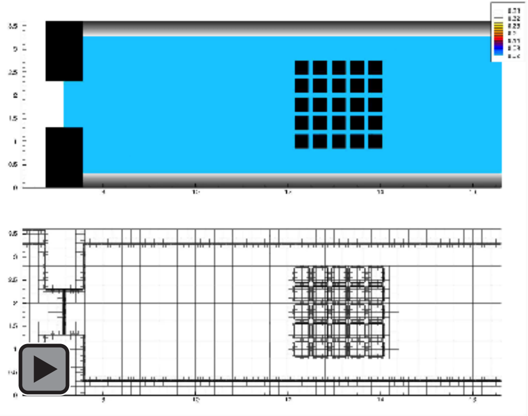
| | |
|---|---|
| # mesh cells: | $13,271,040$ |
| # of D.o.Fs in Uniform Discretization: | $79,626,240$ |
| maximal # of cells in Adaptive Mesh: | $623,025$ |
| | ($4,7\%$ of Reference Method) |
| # of Time Steps: | ca $1,000,000$ |
| CPU Time: | 48 Hours on 320 CPUs |
| | (Intel Xeon Cluster) |

UNIVERSITY OF SOUTH CAROLINA

# Example: Flooded City

Courtesy of Nils Gerhard and Siegfried Müller, RWTH Aachen, Germany

(click video to play)



# Outline

1. What is this about?
   - Prelude
   - Examples
2. Two Worlds
3. How Should the Two Worlds Talk to Each Other?

UNIVERSITY OF SOUTH CAROLINA

# The Role of Reduced Models

- Parametric family of PDEs: $F(u, y, f) = 0, \; y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$

# The Role of Reduced Models

- Parametric family of PDEs: $F(u, y, f) = 0$, $y \in \mathcal{Y}$, $\mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$
- Frequent parameter queries (design tasks, optimal control, calibration, parameter estimation, etc.)

# The Role of Reduced Models

- Parametric family of PDEs: $F(u, y, f) = 0, \; y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$
- Frequent parameter queries (design tasks, optimal control, calibration, parameter estimation, etc.)
- Given $\epsilon > 0$, find possibly small space $\mathbb{U}_\epsilon \subset \mathbb{U}$ such that

$$\sup_{y \in \mathcal{Y}} \inf_{w \in \mathbb{U}_\epsilon} \|u(y) - w\|_{\mathbb{U}} =: \underset{\mathbb{U}}{\mathrm{dist}}(\mathcal{M}, \mathbb{U}_\epsilon) \leq \epsilon$$

UNIVERSITY OF SOUTH CAROLINA

# The Role of Reduced Models

- Parametric family of PDEs: $F(u, y, f) = 0, \; y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$
- Frequent parameter queries (design tasks, optimal control, calibration, parameter estimation, etc.)
- Given $\epsilon > 0$, find possibly small space $\mathbb{U}_\epsilon \subset \mathbb{U}$ such that

$$\sup_{y \in \mathcal{Y}} \inf_{w \in \mathbb{U}_\epsilon} \|u(y) - w\|_{\mathbb{U}} =: \operatorname{dist}_{\mathbb{U}}(\mathcal{M}, \mathbb{U}_\epsilon) \leq \epsilon \quad \text{or} \quad \int_{\mathcal{Y}} \|u(y) - P_{\mathbb{U}_\epsilon} u(y)\|_{\mathbb{U}}^2 dy \leq \epsilon^2$$

UNIVERSITY OF SOUTH CAROLINA

# The Role of Reduced Models

- Parametric family of PDEs: $F(u, y, f) = 0, \ y \in \mathcal{Y}, \quad \mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$
- Frequent parameter queries (design tasks, optimal control, calibration, parameter estimation, etc.)
- Given $\epsilon > 0$, find possibly small space $\mathbb{U}_\epsilon \subset \mathbb{U}$ such that

$$\sup_{y \in \mathcal{Y}} \inf_{w \in \mathbb{U}_\epsilon} \|u(y) - w\|_{\mathbb{U}} =: \operatorname{dist}_{\mathbb{U}}(\mathcal{M}, \mathbb{U}_\epsilon) \leq \epsilon \quad \text{or} \quad \int_{\mathcal{Y}} \|u(y) - P_{\mathbb{U}_\epsilon} u(y)\|_{\mathbb{U}}^2 dy \leq \epsilon^2$$

- $\rightsquigarrow$ speeds up forward simulations needed in inverse tasks:

# The Role of Reduced Models

- Parametric family of PDEs: $F(u, y, f) = 0$, $y \in \mathcal{Y}$, $\mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$
- Frequent parameter queries (design tasks, optimal control, calibration, parameter estimation, etc.)
- Given $\epsilon > 0$, find possibly small space $\mathbb{U}_\epsilon \subset \mathbb{U}$ such that

$$\sup_{y \in \mathcal{Y}} \inf_{w \in \mathbb{U}_\epsilon} \|u(y) - w\|_{\mathbb{U}} =: \operatorname{dist}_{\mathbb{U}}(\mathcal{M}, \mathbb{U}_\epsilon) \leq \epsilon \quad \text{or} \quad \int_{\mathcal{Y}} \|u(y) - P_{\mathbb{U}_\epsilon} u(y)\|_{\mathbb{U}}^2 dy \leq \epsilon^2$$

- $\rightsquigarrow$ speeds up forward simulations needed in inverse tasks:

  - state estimation

  - parameter estimation

# The Role of Reduced Models

- Parametric family of PDEs: $F(u, y, f) = 0$, $y \in \mathcal{Y}$, $\mathcal{M} = \{u(y) : y \in \mathcal{Y}\}$
- Frequent parameter queries (design tasks, optimal control, calibration, parameter estimation, etc.)
- Given $\epsilon > 0$, find possibly small space $\mathbb{U}_\epsilon \subset \mathbb{U}$ such that

$$\sup_{y \in \mathcal{Y}} \inf_{w \in \mathbb{U}_\epsilon} \|u(y) - w\|_{\mathbb{U}} =: \operatorname{dist}_{\mathbb{U}}(\mathcal{M}, \mathbb{U}_\epsilon) \leq \epsilon \quad \text{or} \quad \int_{\mathcal{Y}} \|u(y) - P_{\mathbb{U}_\epsilon} u(y)\|_{\mathbb{U}}^2 dy \leq \epsilon^2$$

- $\rightsquigarrow$ speeds up forward simulations needed in inverse tasks:
  - state estimation
  - parameter estimation

  using data/observations

# The Role of Reduced Models

## Issues:

- How to construct $\mathbb{U}_\epsilon$?

# The Role of Reduced Models

## Issues:

- How to construct $\mathbb{U}_\epsilon$?
- How to represent data? ($\infty$-dimensional setting essential)

# The Role of Reduced Models

## Issues:

- How to construct $\mathbb{U}_\epsilon$?
- How to represent data? ($\infty$-dimensional setting essential)
- The use of a reduced model is a regularization - understand the bias

# The Role of Reduced Models

## Issues:

- How to construct $\mathbb{U}_\epsilon$?
- How to represent data? ($\infty$-dimensional setting essential)
- The use of a reduced model is a regularization - understand the bias

Overarching question:

First discretize then process/analyze/optimize
or

# The Role of Reduced Models

## Issues:

- How to construct $\mathbb{U}_\epsilon$?
- How to represent data? ($\infty$-dimensional setting essential)
- The use of a reduced model is a regularization - understand the bias

Overarching question:

First discretize then process/analyze/optimize
or
First process/analyze/optimize then discretize (according to current needs)
???

UNIVERSITY OF SOUTH CAROLINA

# The Role of Reduced Models

## Issues:

- How to construct $\mathbb{U}_\epsilon$?
- How to represent data? ($\infty$-dimensional setting essential)
- The use of a reduced model is a regularization - understand the bias

Overarching question:

First discretize then process/analyze/optimize
or
First process/analyze/optimize then discretize (according to current needs)
???

This workshop:
- foundational aspects of reduced models, data assimilation, state estimation
- structural imaging, optimization
- DPG - a framework for well conditioned variational formulations