Optimal Uncertainty Quantification

Tim Sullivan
tjs@caltech.edu

California Institute of Technology

AMSTAT Seminar
Mathematics Institute, University of Warwick, U.K.
18 March 2011

Collaborators and Support

- Optimal UQ Theory:
  - Mike McKerns, Michael Ortiz, Houman Owhadi (Caltech)
  - Clint Scovel (Los Alamos National Laboratory)
- Optimal UQ with Legacy Data:
  - Dominik Meyer (Technische Universität München)
  - Florian Theil (University of Warwick)
- Multiphase Steel Application
  - Daniel Balzani (Universität Duisberg-Essen)

CALTECH PSAAP

National Nuclear Security Administration

Portions of this work were supported by the United States Department of Energy National Nuclear Security Administration under Award Number DE-FC52-08NA28613 through the California Institute of Technology's ASC/PSAAP Center for the Predictive Modeling and Simulation of High Energy Density Dynamic Response of Materials.

Outline

1. Introduction
2. Optimal Uncertainty Quantification
3. Consequences of Optimal UQ
4. Computational Examples
5. Further Work and Conclusions

Introduction

What is Uncertainty Quantification?

?

What is Uncertainty Quantification?

The Many Faces of Uncertainty Quantification

- *The Elephant in the Room:* there is a growing consensus that UQ is an essential component of objective science and decision-making.
- *The Blind Men and the Elephant:* unfortunately, as it stands at the moment, UQ has all the hallmarks of an ill-posed problem.

Problems

- Certification
- Extrapolation/Prediction
- Reliability Estimation
- Sensitivity Analysis
- Verification
- Validation
- ...

Methods

- Analysis of Variance
- Bayesian Methods
- Error Bars
- Latin Hypercube Sampling
- (Quasi) Monte Carlo
- Stochastic Collocation
- ...

Types of Uncertainty

- Uncertainties are often divided into two types: epistemic and aleatoric uncertainties.[1]
- An epistemic uncertainty is one that stems from a fundamental lack of knowledge — we don't know the rules that govern the problem.
- An aleatoric uncertainty is one that stems from intrinsic randomness in the system — a "roll of the dice".
- The conventional wisdom is that aleatoric uncertainties are "nicer" than epistemic uncertainties, because the powerful tools of probability theory can be brought to bear.

[1] W. L. Oberkampf, T. G. Trucano & C. Hirsch (2004) "Verification, validation, and predictive capability in computational engineering and physics" *ASME Appl. Mech. Rev.* 57(5):345–384.

Types of Uncertainty

- However, on close inspection, many apparently aleatoric uncertainties are epistemic: are you *really* sure that important parameter X is uniformly distributed, or has sub-Gaussian tails?
- Therefore, theoreticians and practitioners alike tend to be somewhat skeptical of probabilistic methods — "probabilistic reliability studies involve assumptions on the probability densities, whose knowledge regarding relevant input quantities is central."⟨2⟩
- On the other hand, UQ methods based on deterministic worst-case scenarios are oftentimes "too pessimistic to be practical."⟨3⟩

⟨2⟩ I. Elishakoff & M. Ohsaki (2010) *Optimization and Anti-Optimization of Structures Under Uncertainty.* World Scientific, London.

⟨3⟩ R. F. Drenick (1973) "Aseismic design by way of critical excitation" *J. Eng. Mech. Div., Am. Soc. Civ. Eng.* 99:649–667.

Optimal Uncertainty Quantification

- We propose a mathematical framework for UQ as an optimization problem, which we call Optimal Uncertainty Quantification (OUQ), in which knowledge ($\epsilon\pi\iota\sigma\tau\eta\mu\eta$) lies at the heart of the problem formulation.
- The development and application of OUQ to real, complex problems is a collaborative interdisciplinary effort that requires expertise (and stimulates developments) in
  - pure and applied mathematics, especially probability theory,
  - numerical optimization,
  - (massively) parallel computing,
  - the application area (*e.g.* biology, chemistry, economics, engineering, geoscience, meteorology, physics, . . .).

Optimal Uncertainty Quantification

In a nutshell, the OUQ viewpoint is the following:

> *OUQ is the business of computing optimal bounds on quantities of interest that are themselves functions of unknown functions and unknown probability measures, where "optimality" means that those bounds are the sharpest ones possible given the available information on those unknowns.*

This paradigm is easiest to explain in the prototypical context of the certification problem (bounding probabilities of failure). There will be some remarks about other contexts at the end of the talk.

The Certification Problem

- Suppose that you are interested in a system of interest, $G\colon \mathcal{X} \to \mathbb{R}$, which is a real-valued function of some random inputs $X \in \mathcal{X}$ with probability distribution $\mathbb{P}$ on $\mathcal{X}$.
- Some value $\theta \in \mathbb{R}$ is a *performance threshold*: if $G(X) \leq \theta$, then the system fails; if $G(X) > \theta$, then the system succeeds.
- You want to know the probability of failure

$$\mathbb{P}[G(X) \leq \theta],$$

  or at least to know if it exceeds some prescribed maximum acceptable probability of failure ϵ — but you do not know G and $\mathbb{P}$!
- If you have some information about G and $\mathbb{P}$, what are the best rigorous lower and upper bounds that you can give on the probability of failure using that information?

The Certification Problem

- The challenge, then, is to bound $\mathbb{P}[G(X) \leq \theta]$ given some information on or assumptions about G and $\mathbb{P}$.
- Optimality of the bounds is important — the following bounds are true, but useless:
$$0 \leq \mathbb{P}[G(X) \leq \theta] \leq 1.$$
- Robustness of the bounds is also important — *i.e.* to know that the bounds and any safe/unsafe decisions are stable with respect to perturbations of the information/assumptions.

The Importance of Optimality and Robustness. Being overly conservative may lead to huge economic losses, but being overly optimistic may lead to loss of life, environmental damage *& c.*

Space Shuttle *Columbia*, 1 February 2003

Eyjafjallajökull, Iceland, 27 March 2010

Deepwater Horizon, 21 April 2010

Standard UQ Methods

So, how can one show that $\mathbb{P}[G(X) \leq \theta] \leq \epsilon$ when G and $\mathbb{P}$ are only imperfectly known?

- **Monte Carlo methods?**
  We need many independent $\mathbb{P}$-distributed samples of $G(X)$: naïve MC needs $O(\epsilon^{-2} \log \epsilon^{-1})$ samples; QMC needs $O(\epsilon^{-1} (\log \epsilon^{-1})^{\dim \mathcal{X}})$ samples and G to be "well-behaved".
- **Stochastic collocation methods?**
  We need a good representation for $\mathbb{P}$ and rapid decay of the spectrum, and easy exercise of G.
- **Bayesian inference?**
  We need prior distributions in which we genuinely believe; also, if the priors and "reality" disagree greatly, then it may take a very large data set to "correct" the priors into posteriors that are close to "reality", which is vital if the important events are prior-rare.

Standard UQ Methods

- Each of these methods relies, implicitly or explicitly, on the validity of certain assumptions in order to be applicable — or at least efficient. This leads to three main difficulties:
  - the assumptions may not match the information about G and $\mathbb{P}$;
  - the assumptions may vary from method to method, which makes fair comparisons of different methods difficult;
  - the assumptions often cannot be easily perturbed.
- Therefore, in formulating OUQ, we choose to place information on and assumptions about G and $\mathbb{P}$ at the centre of the problem.
- This goes one step beyond Babuška's commandment "thou shalt confess thy sins": in OUQ, your sins precisely describe your problem!

Optimal Uncertainty Quantification

What Problem Should You Solve?

- You want to know about the probability of failure

$$\mathbb{P}[G(X) \leq \theta],$$

  or at least if it's greater than or less than ϵ.
- You want to do this without ignoring or distorting your existing information set, nor making additional assumptions.
- If you had access to The Ultimate Computer, what problem would you try to solve?
- Worry about computational feasibility later!

"Forty-Two?!"

Information / Assumptions

- Write down all the information that you have about the system. For example, this information might come from
  - physical laws;
  - expert opinion;
  - experimental data.
- Let $\mathcal{A}$ denote the set of all pairs (f, μ) that are consistent with your information about $(G, \mathbb{P})$:

$$\mathcal{A} \subseteq \left\{ (f, \mu) \,\middle|\, \begin{array}{c} f \colon \mathcal{X} \to \mathbb{R} \text{ is measurable, and} \\ \mu \text{ is a probability measure on } \mathcal{X} \end{array} \right\}.$$

- All you know about reality is that $(G, \mathbb{P}) \in \mathcal{A}$; any $(f, \mu) \in \mathcal{A}$ is an admissible scenario for the unknown reality $(G, \mathbb{P})$.
- $\mathcal{A}$ is a huge space (probably infinite-dimensional and non-separable in any example of interest); the need to explore it efficiently motivates the reduction theorems that will come later.

The Optimal UQ Problem

With this notation, the Optimal UQ Problem is simply to find the greatest lower bound and least upper bound on the probability of failure among all admissible scenarios $(f, \mu) \in \mathcal{A}$. That is, we want to calculate

$$\mathcal{L}(\mathcal{A}) := \inf_{(f,\mu)\in\mathcal{A}} \mu[f \leq \theta]$$

and

$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu)\in\mathcal{A}} \mu[f \leq \theta].$$

We then have the bounds

$$\mathcal{L}(\mathcal{A}) \leq \mathbb{P}[G(X) \leq \theta] \leq \mathcal{U}(\mathcal{A}),$$

and any bounds other than these would be either (a) not sharp or (b) not conservative.

Rigorous and Optimal Certification Criteria

Given a maximum acceptable probability of failure $\epsilon \in [0, 1]$, calculation of $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$ yields unambiguous, rigorous and optimal criteria for certification of the system:

- if $\mathcal{U}(\mathcal{A}) \leq \epsilon$, then the system is safe even in the worst possible case;
- if $\mathcal{L}(\mathcal{A}) > \epsilon$, then the system is unsafe even in the best possible case;
- if $\mathcal{L}(\mathcal{A}) \leq \epsilon < \mathcal{U}(\mathcal{A})$, then there are some admissible scenarios under which the system is safe and others under which it is unsafe: the information encoded in $\mathcal{A}$ is insufficient to rigorously certify the system, and more information must be sought.
  - The system is (temporarily) deemed unsafe due to lack of information.
  - More information yields a smaller admissible set $\mathcal{A}' \subseteq \mathcal{A}$:

$$\mathcal{L}(\mathcal{A}) \leq \mathcal{L}(\mathcal{A}') \leq \mathbb{P}[G(X) \leq \theta] \leq \mathcal{U}(\mathcal{A}') \leq \mathcal{U}(\mathcal{A}).$$

Simple Mean and Range Constraints

A simple example of an admissible set $\mathcal{A}$ is the following one: our information consists of a lower bound on the mean performance and an upper bound on the diameter of the set of values that the system can take.

$$\mathcal{A} := \left\{ (f, \mu) \,\middle|\, \begin{array}{c} f \colon \mathcal{X} \to \mathbb{R} \text{ is measurable,} \\ \mu \text{ is a probability measure on } \mathcal{X}, \\ \mathbb{E}_\mu[f] \geq m \\ \sup f - \inf f \leq D \end{array} \right\}.$$

This example can be solved exactly:

$$\mathcal{U}(\mathcal{A}) := \sup_{(f,\mu)\in\mathcal{A}} \mu[f \leq \theta] = \left(1 - \frac{(m-\theta)_+}{D}\right)_+,$$

where, for $t \in \mathbb{R}$, $t_+ := \max\{0, t\}$.

More Complicated Mean and Range Constraints

On a product space $\mathcal{X} := \mathcal{X}_1 \times \cdots \times \mathcal{X}_K$, consider

$$\mathcal{A}_{\mathsf{McD}} := \left\{ (f, \mu) \,\middle|\, \begin{array}{c} f \colon \mathcal{X} \to \mathbb{R}, \\ \mathcal{D}_k[f] := \sup |f(x^1, \dots, x^k, \dots, x^K) - \\ \qquad - f(x^1, \dots, \tilde{x}^k, \dots, x^K)| \leq D_k, \\ \mu = \mu_1 \otimes \cdots \otimes \mu_K \text{ on } \mathcal{X}, \\ \mathbb{E}_\mu[f] \geq m \end{array} \right\}.$$

These are the assumptions of McDiarmid's inequality[4] (a.k.a. the bounded differences inequality), which gives the upper bound

$$\mathcal{U}(\mathcal{A}_{\mathsf{McD}}) \leq \exp\left(- \frac{2(m - \theta)_+^2}{\sum_{k=1}^K D_k^2} \right).$$

[4] C. McDiarmid (1989) "On the method of bounded differences" *Surveys in Combinatorics, 1989*, Camb. Univ. Press, 148–188.

Simple Legacy Data Constraints

- You observe a function $G \colon [0, 1] \to \mathbb{R}$ on a fixed finite set $\mathcal{O} = \{z_1, \dots, z_N\} \subseteq [0, 1]$. You want to bound $\mathbb{P}[G(X) \leq \theta]$, and know neither G nor the distribution $\mathbb{P}$ of X exactly.
- Suppose that you do know
  - $G|_{\mathcal{O}}$ (bearing in mind that $\mathcal{O}$ may not be $\mathbb{P}$-distributed),
  - G is differentiable with $|G'| \leq 1$ everywhere,
  - $\mathbb{E}_{\mathbb{P}}[G(X)] \geq m$.
- The corresponding set of admissible scenarios is

$$\mathcal{A} := \left\{ (f, \mu) \,\middle|\, \begin{array}{c} f \colon [0, 1] \to \mathbb{R} \text{ has Lipschitz constant } 1, \\ f = G \text{ on } \mathcal{O}, \\ \mu \text{ is a probability measure on } [0, 1], \\ \mathbb{E}_{\mu}[f] \geq m \end{array} \right\}.$$

Simple Legacy Data Constraints

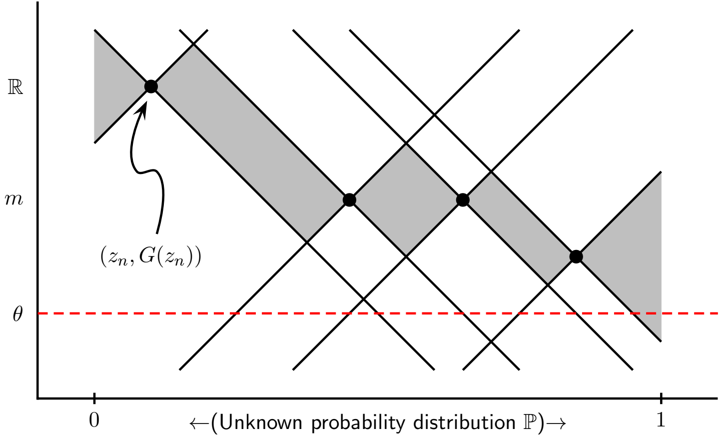


Figure: Shaded, the feasible region for the graph of the unknown function G. How much probability mass can be put on points with G-values below the red line while still satisfying $\mathbb{E}[G] \geq m$?

More Complicated Legacy Data Constraints

More generally, the unknown function G may be a function of many independent inputs of unknown distribution. Whatever information we have about the smoothness of G becomes a constraint on the smoothness of the admissible scenarios f:

$$\mathcal{A} := \left\{ (f, \mu) \,\middle|\, \begin{array}{c} f \colon \mathcal{X} \to \mathbb{R} \text{ is measurable,} \\ \mu = \mu_1 \otimes \cdots \otimes \mu_K \text{ on } \mathcal{X}, \\ f = G \text{ on } \mathcal{O}, \\ \langle \text{some smoothness conditions on } f \rangle, \\ \mathbb{E}_\mu[f] \geq m \end{array} \right\}.$$

Reduction of OUQ

- OUQ problems are global, infinite-dimensional, non-convex, highly-constrained (*i.e.* nasty!) optimization problems.
- The non-convexity is a fact of life, but there are powerful reduction theorems that allow a reduction to a search space of low dimension.
- Instead of searching over all admissible probability measures μ, we need only to search over those with a very simple "extremal" structure: in the simplest case, these are just finite sums of point masses (Dirac measures) on the input parameter space $\mathcal{X}$.
- That is, we can "pretend" that all the random inputs are discrete random variables and just optimize over the possible values and probabilities that those discrete variables might take.

Reduction of OUQ — Linear Inequalities on Moments

Suppose that the admissible set $\mathcal{A}$ has the following form: all the constraints on the measure μ are linear inequalities on generalized moments. That is, for some given functions $g'_1, \dots, g'_{n'} : \mathcal{X} \to \mathbb{R}$,

$$\mathcal{A} = \left\{ (f, \mu) \,\middle|\, \begin{array}{c} f \colon \mathcal{X} \to \mathbb{R} \text{ such that} \\ \langle \text{some conditions on } f \text{ alone} \rangle, \\ \mathbb{E}_\mu[g'_1] \leq 0, \dots, \mathbb{E}_\mu[g'_{n'}] \leq 0 \end{array} \right\}.$$

Theorem (General reduction theorem)

If $\mathcal{X}$ is a Suslin space, then $\mathcal{L}(\mathcal{A}) = \mathcal{L}(\mathcal{A}_\Delta)$ and $\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta)$, where

$$\mathcal{A}_\Delta = \left\{ (f, \mu) \in \mathcal{A} \,\middle|\, \begin{array}{c} \mu \text{ is a sum of at most } n' + 1 \\ \text{weighted Dirac measures on } \mathcal{X} \end{array} \right\}.$$

Note: No constraints $\implies$ conventional, deterministic worst-case analysis.

Linear Inequalities on Moments

- What constraints does the theorem permit?
  - Inequalities on probabilities of certain events, *e.g.*

$$\mu[X \in E] \lesseqqgtr c \text{ or } \mu[f(X) \in E'] \lesseqqgtr c.$$

  - Inequalities on means and higher moments of X, $f(X)$, or any other measurable functions of X, *e.g.*

$$\mathbb{E}_\mu[\langle \ell, X \rangle] \lesseqqgtr c,\ \mathbb{E}_\mu[|f(X)|^p] \lesseqqgtr c,\ \mathbb{E}_\mu[g(X)] \lesseqqgtr c \text{ or } \mathbb{V}_\mu[g(X)] \lesseqqgtr c.$$

- What constraints does the theorem not permit?
  - Relative entropy constraints, *e.g.* for probability measures μ on $\mathbb{R}$ that are absolutely continuous with respect to Lebesgue measure λ, a constraint on

$$D_{\mathsf{KL}}(\mu \| \lambda) := \int_{\mathbb{R}} \frac{\mathrm{d}\mu}{\mathrm{d}\lambda} \log \frac{\mathrm{d}\mu}{\mathrm{d}\lambda} \, \mathrm{d}\lambda.$$

Reduction of OUQ — Independent Inputs

Suppose that we have K independent inputs, *i.e.* $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_K$ and

$$\mathcal{A} = \left\{ (f, \mu) \;\middle|\; \begin{matrix} f \colon \mathcal{X} \to \mathbb{R} \text{ such that} \\ \langle \text{some conditions on } f \text{ alone} \rangle, \\ \mu = \mu_1 \otimes \cdots \otimes \mu_K, \\ \mathbb{E}_\mu[g'_1] \leq 0, \ldots, \mathbb{E}_\mu[g'_{n'}] \leq 0, \\ \mathbb{E}_{\mu_k}[g_i^k] \leq 0 \text{ for } i = 1, \ldots, n_k \text{ and } k = 1, \ldots, K \end{matrix} \right\}.$$

Theorem (Reduction for independent input parameters)

If $\mathcal{X}_1, \ldots, \mathcal{X}_K$ are Suslin spaces, then $\mathcal{L}(\mathcal{A}) = \mathcal{L}(\mathcal{A}_\Delta)$ and $\mathcal{U}(\mathcal{A}) = \mathcal{U}(\mathcal{A}_\Delta)$, where

$$\mathcal{A}_\Delta = \left\{ (f, \mu) \in \mathcal{A} \;\middle|\; \begin{matrix} \mu_k \text{ is a sum of at most } n' + n_k + 1 \\ \text{weighted Dirac measures on } \mathcal{X}_k \end{matrix} \right\}.$$

Reduction of OUQ — Legacy Problem

- Note that the reduction to finite convex combinations of Dirac measures not only simplifies the search over admissible probability measures μ, but also the search over admissible response functions f, because we now only care about the values of f on the (finite!) support of the discrete measure μ.
- So, for example, consider what this means for the legacy OUQ problem with given data on $\mathcal{O} = \{z_1, \dots, z_N\}$:

$$\mathcal{A} := \left\{ (f, \mu) \;\middle|\; \begin{array}{c} f \colon \mathcal{X} \to \mathbb{R} \text{ is measurable,} \\ \mu = \mu_1 \otimes \cdots \otimes \mu_K \text{ on } \mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_K, \\ f = G \text{ on } \mathcal{O} \text{ and } \mathbb{E}_\mu[f] \geq m, \\ |f(x) - f(x')| \leq d_L(x, x') \end{array} \right\}.$$

We now only care about the values of f on $\mathcal{O}$ and on a Hamming cube of 2^K points $\{x_\varepsilon \mid \varepsilon \in \{0,1\}^K\}$ on which the measure μ lives.

Reduction of OUQ — Legacy Problem

McShane's extension theorem[5] ensures that any partially-defined Lipschitz function can be extended to a fully-defined one, and so

$$\mathcal{U}(\mathcal{A}) = \begin{cases} \text{maximize:} & \displaystyle\sum_{\varepsilon \in \{0,1\}^K} \left(\prod_{k=1}^{K} (p_k)^{1-\varepsilon_k} (1-p_k)^{\varepsilon_k} \right) \mathbb{1}_{(-\infty,\theta]}(y_\varepsilon); \\ \text{among:} & x_0, x_1 \in \mathcal{X}; \quad \text{(Track the support of } \mu\text{, with } x_\varepsilon^k := x_{\varepsilon_k}^k) \\ & y \colon \{0,1\}^K \to \mathbb{R}; \quad \text{(Tracks the values of } f \text{ on the support of } \mu) \\ & p \in [0,1]^K; \quad \text{(Tracks the weights of } \mu) \\ \text{subject to:} & \text{for all } \varepsilon, \varepsilon' \in \{0,1\}^K, \\ & \quad |y_\varepsilon - y_{\varepsilon'}| \leq d_L(x_\varepsilon, x_{\varepsilon'}); \\ & \text{for all } \varepsilon \in \{0,1\}^K, n \in \{1, \dots, N\}, \\ & \quad |y_\varepsilon - f(z_n)| \leq d_L(x_\varepsilon, z_n); \\ & \displaystyle\sum_{\varepsilon \in \{0,1\}^K} \left(\prod_{k=1}^{K} (p_k)^{1-\varepsilon_k} (1-p_k)^{\varepsilon_k} \right) y_\varepsilon \geq m. \end{cases}$$

[5] E. J. McShane (1934) "Extension of range of functions" *Bull. Amer. Math. Soc.* 40(12):837–842.

Consequences of Optimal UQ

McDiarmid's Inequality

Consider the following admissible set of scenarios for $(G, \mathbb{P})$, where G has K independent inputs, mean performance $\mathbb{E}[G(X)] \geq m$, and the maximum oscillation of G with respect to changes of its k^{th} argument is at most D_k:

$$\mathcal{A}_{\mathsf{McD}} := \left\{ (f, \mu) \,\middle|\, \begin{array}{c} f \colon \mathcal{X} \to \mathbb{R}, \\ \mathcal{D}_k[f] := \sup |f(x^1, \dots, x^k, \dots, x^K) - \\ \qquad - f(x^1, \dots, \widetilde{x}^k, \dots, x^K)| \leq D_k, \\ \mu = \mu_1 \otimes \dots \otimes \mu_K \text{ on } \mathcal{X} = \mathcal{X}_1 \times \dots \times \mathcal{X}_K, \\ \mathbb{E}_\mu[f] \geq m \end{array} \right\}.$$

In terms of OUQ, McDiarmid's inequality is simply the upper bound

$$\mathcal{U}(\mathcal{A}_{\mathsf{McD}}) \leq \exp\left(-2 \frac{(m - \theta)_+^2}{\sum_{k=1}^K D_k^2} \right).$$

Optimal McDiarmid Inequality

- The reduction theorems mentioned before, along with other reduction theorems that eliminate dependency upon the coordinate positions in the parameter space $\mathcal{X}$, yield finite-dimensional problems that can be solved exactly to give optimal concentration inequalities with the same assumptions as McDiarmid's inequality.
- By a combinatorial induction procedure, $\mathcal{U}(\mathcal{A}_{\mathsf{McD}})$ can be calculated for any $K \in \mathbb{N}$.
- Write $a := (m - \theta)_+$ for the mean performance margin.

Optimal McDiarmid, $K = 1$

$$\mathcal{U}(\mathcal{A}_{\mathsf{McD}}) = \left(1 - \frac{a}{D_1}\right)_+$$

Optimal McDiarmid Inequality

Optimal McDiarmid, $K = 2$

$$\mathcal{U}(\mathcal{A}_{\mathsf{McD}}) = \begin{cases} 0, & \text{if } D_1 + D_2 \leq a, \\ \dfrac{(D_1 + D_2 - a)^2}{4 D_1 D_2}, & \text{if } |D_1 - D_2| \leq a \leq D_1 + D_2, \\ \left(1 - \dfrac{a}{\max\{D_1, D_2\}}\right)_+, & \text{if } 0 \leq a \leq |D_1 - D_2|. \end{cases}$$

- Note that, when there is uncertainty about the response function G, not all parameter sensitivities are created equal!
- If the "sensitivity gap" between the largest parameter sensitivity D_1 and the second-largest one D_2 is big enough, then all the uncertainty in the probability of failure is controlled by D_1 and the performance margin $a := (m - \theta)_+$.

Other Optimal Concentration Inequalities

- Similarly, one can consider the admissible set $\mathcal{A}_{\mathsf{Hfd}}$ that corresponds to the assumptions of Hoeffding's inequality,[6] in which the functions f must be linear:

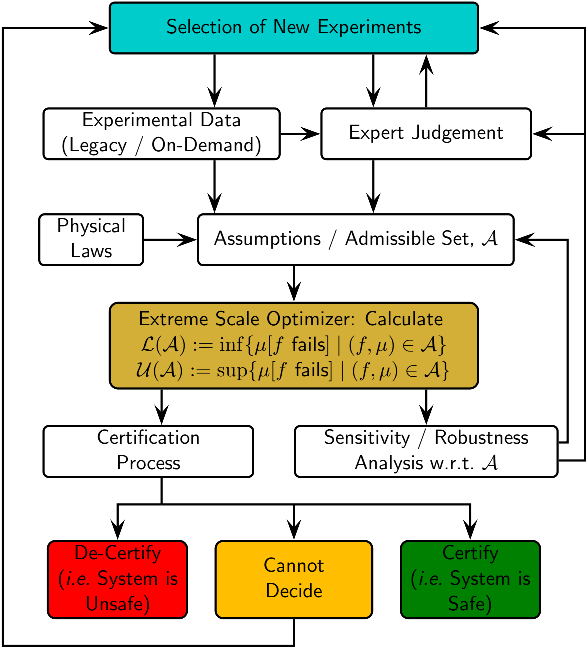
$$\mathcal{A}_{\mathsf{Hfd}} := \left\{ (f, \mu) \;\middle|\; \begin{array}{c} f \colon \mathbb{R}^K \to \mathbb{R} \text{ is given by} \\ f(x_1, \dots, x_K) = x_1 + \dots + x_K, \\ \mu \text{ supported on a cube with sides } D_1, \dots, D_K, \\ \mathbb{E}_\mu[f] \geq m \end{array} \right\}.$$

- Interestingly, $\mathcal{U}(\mathcal{A}_{\mathsf{McD}}) = \mathcal{U}(\mathcal{A}_{\mathsf{Hfd}})$ for $K = 1$ and $K = 2$, but $\mathcal{U}(\mathcal{A}_{\mathsf{McD}}) \geq \mathcal{U}(\mathcal{A}_{\mathsf{Hfd}})$ for $K = 3$ (and the inequality can be strict).

[6] W. Hoeffding (1963) "Probability inequalities for sums of bounded random variables" *J. Amer. Statist. Assoc.* 58(301):13–30.

The OUQ Loop

- The calculation of extremal probabilities of failure over a fixed admissible set $\mathcal{A}$ is not the end of the game.
- A strength of the OUQ viewpoint is that the assumptions/constraints that determine $\mathcal{A}$ can be perturbed to see if the conclusions are robust with respect to those changes.
- Inconclusive results and sensitive assumptions can be selected for further experimentation — and the OUQ process itself can be used to identify the most informative experiments.


Selection of New Experiments
Experimental Data (Legacy / On-Demand)
Expert Judgement
Physical Laws
Assumptions / Admissible Set, $\mathcal{A}$
Extreme Scale Optimizer: Calculate
$\mathcal{L}(\mathcal{A}) := \inf\{\mu[f \text{ fails}] \mid (f,\mu) \in \mathcal{A}\}$
$\mathcal{U}(\mathcal{A}) := \sup\{\mu[f \text{ fails}] \mid (f,\mu) \in \mathcal{A}\}$
Certification Process
Sensitivity / Robustness Analysis w.r.t. $\mathcal{A}$
De-Certify (*i.e.* System is Unsafe)
Cannot Decide
Certify (*i.e.* System is Safe)

Selection of the Best Next Experiment

- Suppose that you are offered a choice of running just one very expensive experiment from a collection $E_1, E_2, \dots$: each experiment E_i will measure some functional $\Phi_i(G, \mathbb{P})$ to very high accuracy. *E.g.*

$$
\begin{aligned}
\Phi_1(f, \mu) &:= \mathbb{E}_\mu[f], \\
\Phi_2(f, \mu) &:= \mu[X \in A] \text{ for some set } A \subseteq \mathcal{X}, \\
\Phi_3(f, \mu) &:= \mathcal{D}_1[f], \\
\Phi_4(f, \mu) &:= f(x_0) \text{ for some point } x_0 \in \mathcal{X}.
\end{aligned}
$$

- Which experiment should you run? How can one objectively say that one experiment is “better” or “worse” than another?
- In the Optimal UQ framework, we can assess how predictive or decisive a potential experiment may be *in advance of performing it*.

Most Predictive Experiments

- If your objective is to have an "accurate" prediction of $\mathbb{P}[G(X) \leq \theta]$ in the sense that $\mathcal{U}(\mathcal{A}) - \mathcal{L}(\mathcal{A})$ is small, then proceed as follows:
- Let $\mathcal{A}_{E,c}$ denote those scenarios in $\mathcal{A}$ that are compatible with obtaining outcome c from experiment E.
- The experiment that is most predictive even in the worst case is defined by a minimax criterion: we seek

$$E^* \in \operatorname*{arg\,min}_{\text{experiments } E} \left(\sup_{\text{outcomes } c} \left(\mathcal{U}(\mathcal{A}_{E,c}) - \mathcal{L}(\mathcal{A}_{E,c}) \right) \right).$$

- Again, the reduction theorems make this kind of OUQ problem computationally tractable. It is a bigger problem than just calculating $\mathcal{L}(\mathcal{A})$ and $\mathcal{U}(\mathcal{A})$, but the presumption is that computer time is cheaper than experimental effort.
- Alternatively, you may want an experiment that is likely to give a decisive safe/unsafe verdict...

Most Decisive Experiments

- Let $J_{\text{safe},\epsilon}(\Phi_i)$ be the closed interval in $\mathbb{R}$ spanned by the possible values of $\Phi_i(f,\mu)$ among all safe scenarios $(f,\mu) \in \mathcal{A}$, *i.e.* those with $\mu[f \leq \theta] \leq \epsilon$.
- Let $J_{\text{unsafe},\epsilon}(\Phi_i)$ be the closed interval in $\mathbb{R}$ spanned by the possible values of $\Phi_i(f,\mu)$ among all unsafe scenarios $(f,\mu) \in \mathcal{A}$, *i.e.* those with $\mu[f \leq \theta] > \epsilon$.
- Determination of these two intervals means solving four OUQ problems.
- What could you conclude if you were told $\Phi_i(G,\mathbb{P})$?

$$\begin{aligned}
\Phi_i(G,\mathbb{P}) &\in J_{\text{safe},\epsilon}(\Phi_i) \setminus J_{\text{unsafe},\epsilon}(\Phi_i) &&\implies \text{system is safe,} \\
\Phi_i(G,\mathbb{P}) &\in J_{\text{unsafe},\epsilon}(\Phi_i) \setminus J_{\text{safe},\epsilon}(\Phi_i) &&\implies \text{system is unsafe,} \\
\Phi_i(G,\mathbb{P}) &\in J_{\text{safe},\epsilon}(\Phi_i) \cap J_{\text{unsafe},\epsilon}(\Phi_i) &&\implies \text{cannot decide,} \\
\Phi_i(G,\mathbb{P}) &\notin J_{\text{safe},\epsilon}(\Phi_i) \cup J_{\text{unsafe},\epsilon}(\Phi_i) &&\implies \text{faulty assumptions!}
\end{aligned}$$

Most Decisive Experiments

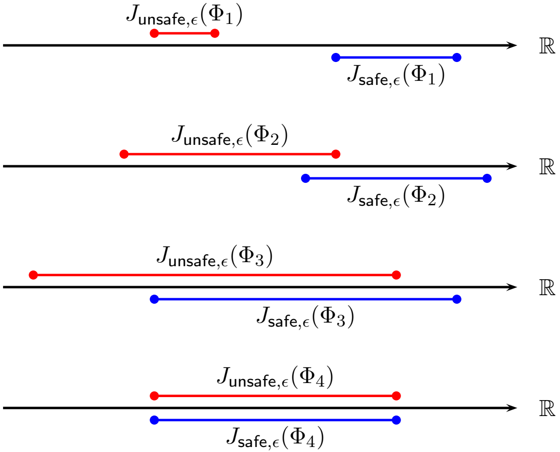


Figure: Outcome intervals for four possible experiments E_1, E_2, E_3 and E_4. E_1 is perfectly decisive; E_4 is completely indecisive; E_2 and E_3 are intermediate cases.

Selection of Experimental Campaigns

- This idea of experimental selection can be extended to plan several experiments in advance, *i.e.* to plan campaigns of experiments.
- This is a kind of infinite-dimensional *Cluedo*, played on spaces of admissible scenarios, against our lack of perfect information about reality, and made tractable by the reduction theorems.

Relevant and Redundant Legacy Data

In many UQ applications, it is important to know which legacy data are those that have any — or the most — relevance to the UQ problem at hand. In this framework, the relevance of legacy data enters naturally in terms of the constraints: relevant data points correspond to non-trivial constraints.

Definition

Given a set of observations of G on $\mathcal{O} \subseteq \mathcal{X}$, say that an observation of G at $z_* \in \mathcal{X}$ is redundant on $S \subseteq \mathcal{X}$ with respect to $\mathcal{O}$ if, whenever the constraints from $G|_{\mathcal{O}}$ are satisfied on S, so is the constraint from $G(z_*)$, *i.e.*

$$\left. \begin{array}{l} \text{for all } n \in \{1, \dots, N\}, \\ \text{for all } x \in S \text{ and } y \in \mathbb{R}, \\ \quad |y - G(z_n)| \leq d_L(x, z_n) \end{array} \right\} \implies |y - G(z_*)| \leq d_L(x, z_*).$$

Relevant and Redundant Legacy Data

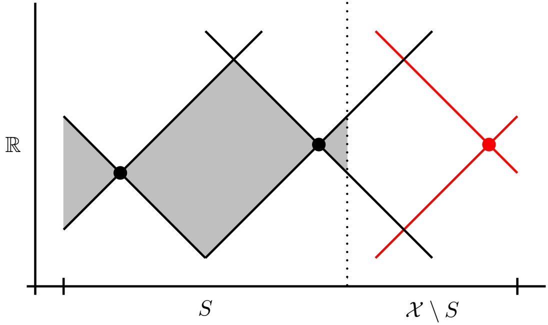


Figure: The red data point on the right is redundant with respect to S and the other data points in S.

Challenge: To develop algorithms that can *efficiently* find maximal redundancy-free subsets of given data sets.

Computational Examples

Example 1: Hypervelocity Impact

Figure: Caltech's Small Particle Hypervelocity Impact Range (SPHIR): a two-stage light gas gun that launches 1–50 mg projectiles at speeds of 2–$10\,\mathrm{km\cdot s^{-1}}$.

Hypervelocity Impact: Surrogate Model

Experimentally-derived deterministic surrogate model for the perforation area (in mm^2), with three independent inputs:

- plate thickness $h \in \mathcal{X}_1 := [1.52, 2.67]\,\mathrm{mm} = [60, 105]\,\mathrm{mil}$;
- impact obliquity $\alpha \in \mathcal{X}_2 := [0, \frac{\pi}{6}]$;
- impact speed $v \in \mathcal{X}_3 := [2.1, 2.8]\,\mathrm{km} \cdot \mathrm{s}^{-1}$.

$$H(h, \alpha, v) := 10.396 \left(\left(\frac{h}{1.778} \right)^{0.476} (\cos \alpha)^{1.028} \tanh \left(\frac{v}{v_{\mathrm{bl}}} - 1 \right) \right)_{+}^{0.468}$$

The quantity $v_{\mathrm{bl}}(h, \alpha)$ given by

$$v_{\mathrm{bl}}(h, \alpha) := 0.579 \left(\frac{h}{(\cos \alpha)^{0.448}} \right)^{1.400}$$

is called the ballistic limit, the impact speed below which no perforation occurs. The failure event is non-perforation, *i.e.* $[H = 0] \equiv [H \leq 0]$.

Hypervelocity Impact: Surrogate Model

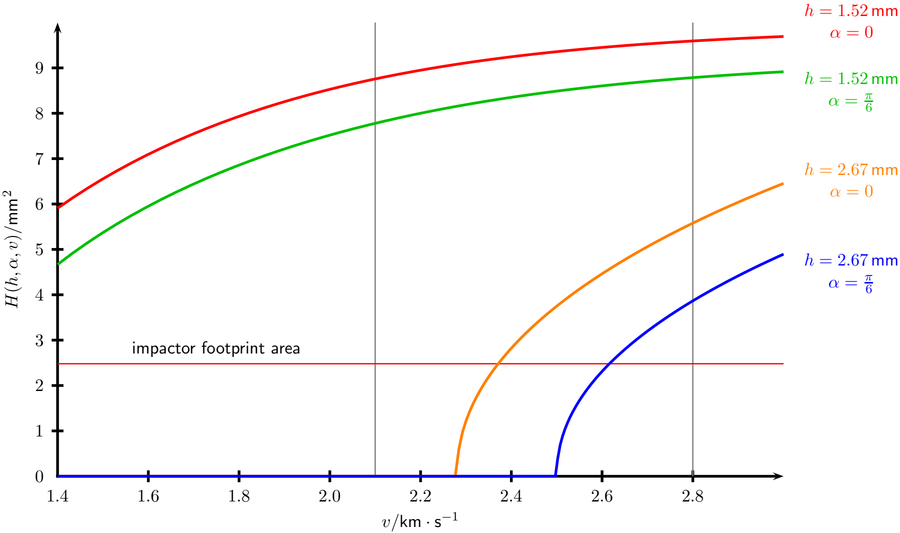


Figure: The surrogate perforation area model of the previous slide.

Admissible scenarios, $\mathcal{A}$	$\mathcal{U}(\mathcal{A})$	Method
$\mathcal{A}_{\mathrm{McD}}$: independence, oscillation and mean constraints (exact response H not given)	$\leq 66.4\%$ $= 43.7\%$	McD. ineq. Opt. McD.
$\mathcal{A} := \{(f, \mu) \mid f = H \text{ and } \mathbb{E}_\mu[H] \in [5.5, 7.5]\}$	$\overset{\text{num}}{=} 37.9\%$	OUQ
$\mathcal{A} \cap \left\{ (f, \mu) \,\middle\vert\, \begin{matrix} \mu\text{-median velocity} \\ = 2.45\,\mathrm{km} \cdot \mathrm{s}^{-1} \end{matrix} \right\}$	$\overset{\text{num}}{=} 30.0\%$	OUQ
$\mathcal{A} \cap \left\{ (f, \mu) \,\middle\vert\, \mu\text{-median obliquity} = \frac{\pi}{12} \right\}$	$\overset{\text{num}}{=} 36.5\%$	OUQ
$\mathcal{A} \cap \left\{ (f, \mu) \,\middle\vert\, \text{obliquity} = \frac{\pi}{6}\ \mu\text{-a.s.} \right\}$	$\overset{\text{num}}{=} 28.0\%$	OUQ

Warning!

It is tempting to say that some of these bounds are "sharper" than others. Except for the first line, every one of these bounds is sharp given the available information, modulo issues of numerical accuracy. In the case of asymmetric information, think before describing a bound as "not sharp".

Numerical Convergence

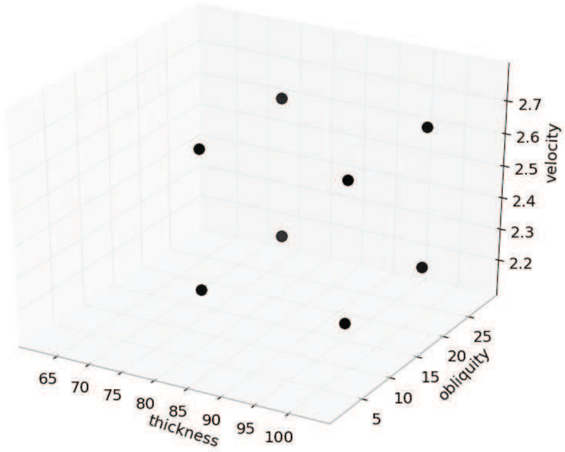


Figure: Support of the $2 \times 2 \times 2$-point measure μ at iteration 0.

Numerical Convergence

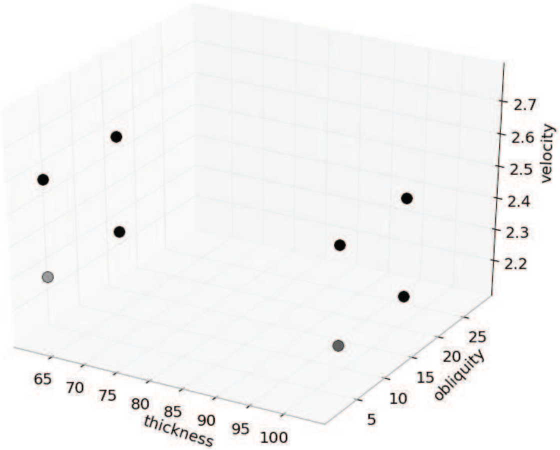


Figure: Support of the $2 \times 2 \times 2$-point measure μ at iteration 150.

Numerical Convergence

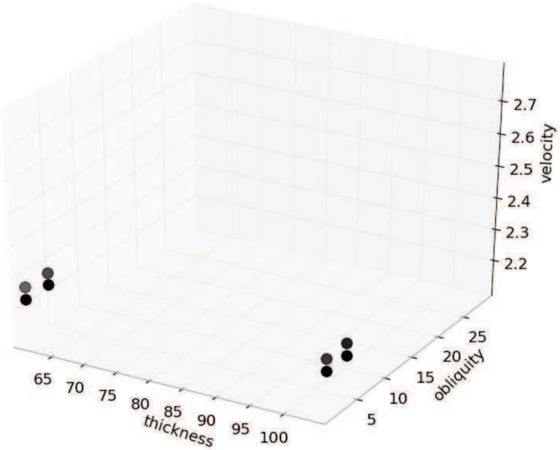


Figure: Support of the $2 \times 2 \times 2$-point measure μ at iteration 200.

Numerical Convergence

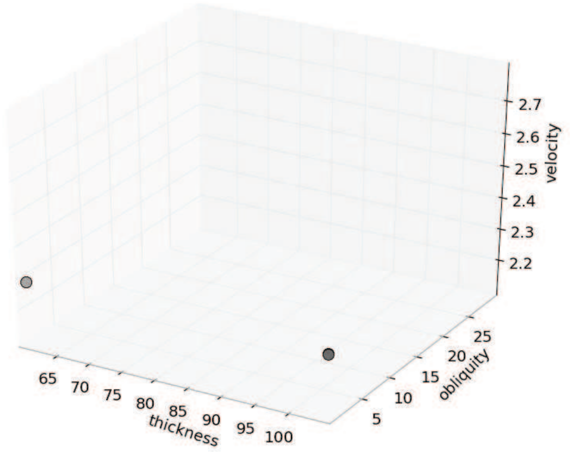


Figure: Support of the $2 \times 2 \times 2$-point measure μ at iteration 1000.

Numerical Convergence

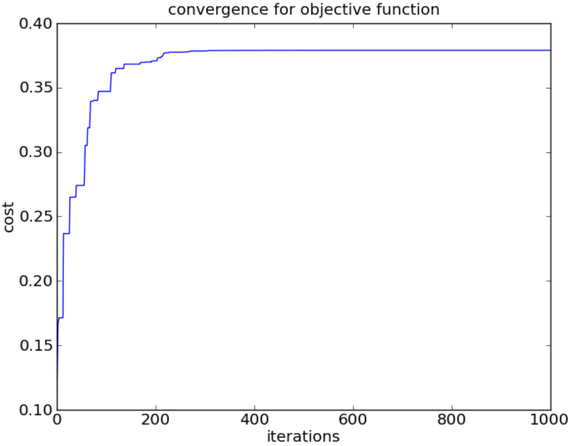
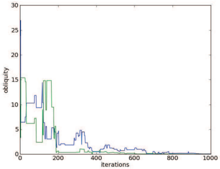
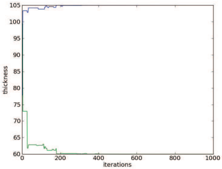


Figure: Numerical convergence of the maximal probability of non-perforation.

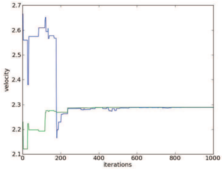
Numerical Convergence



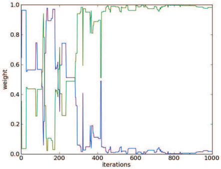
(a) oblquity positions



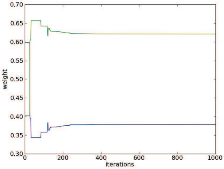
(b) thickness positions



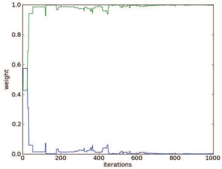
(c) velocity positions



(d) obliquity masses



(e) thickness masses



(f) velocity masses

Figure: Numerical convergence of the positions and probability masses of the components of the $2 \times 2 \times 2$ measure μ.

Comments

- Over these parameter ranges, the oscillations of H are:

$$\mathcal{D}_h[H] = 8.86\,\mathrm{mm}^2, \quad \mathcal{D}_\alpha[H] = 4.17\,\mathrm{mm}^2, \quad \mathcal{D}_v[H] = 7.20\,\mathrm{mm}^2,$$

  so the "screening effects" apply in the optimal McDiarmid inequality.
- The measures that (approximately) maximize the probability of failure yield important information about the "key players" in the system.
- For given mean perforation area, the worst-case probability of failure is not controlled by the impact velocity or the oblquity, but by the thickness of the plate.
- The measure μ's support collapses to
  - the two extremes of the thickness (h) range;
  - the lower extreme of the obliquity (α) range;
  - a single non-trivial value in the velocity (v) range.

Example 2: Seismic Safety Certification

- Now consider a more involved example: the safety of a truss structure under an earthquake.
- For simplicity, we consider a purely elastic response in a truss structure composed of
  - N joints, with positions $u(t) \in \mathbb{R}^{3N}$;
  - J members, with axial strains $y(t) \in \mathbb{R}^{J}$;
  - member j has yield strength S_j, so the structure is safe if

$$\min_{j=1,\dots,J} (S_j - \|y_j\|_\infty) > 0,$$

and fails otherwise.

Equations of Motion

- The time evolution of the structure is governed by the second-order ODE
$$M\ddot{u} + C\dot{u} + Ku = f,$$
where M, C and K are the symmetric positive-definite $3N \times 3N$ mass, damping and stiffness matrices, and f collects any externally-applied loads (*e.g.* wind loads).
- We work instead with $v := u - T\gamma$, *i.e.* u less the motion obtained by rigid translation according to the ground motion:
$$M\ddot{v} + C\dot{v} + Kv = f - MT\ddot{\gamma}.$$
- Let $L \in \mathbb{R}^{J \times 3N}$ map v to the array of member strains, so that $y_j = (Lv)_j$.

Response Function

- Given a history of ground motion acceleration $\ddot{\gamma}$, it is a straightforward calculation in a nodal basis for the structure and using hereditary integrals to solve for v and hence to check the safety of the structure according to the safety criterion

$$\min_{j=1,\dots,J} \left(S_j - \|y_j\|_\infty\right) > 0.$$

- This relationship is deterministic. Randomness enters the problem via the ground motion acceleration, which we treat as a time convolution:[7]

$$\ddot{\gamma}_0 = \psi \star s,$$

where s is the earthquake source and ψ the transfer function.

[7] S. Stein & M. Wysession (2002) *An Introduction to Seismology, Earthquakes, and Earth Structure.* Wiley–Blackwell.

Ground Motion Acceleration

- A reasonable representation for s is as a sequence of I boxcar time impulses of random but independent duration and amplitude:

$$s = \sum_{i=1}^{I} X_i \mathbb{1}_{\left[\sum_{k=1}^{i-1} \tau_k, \sum_{k=1}^{i} \tau_k\right]}$$

  with independent random X_i in $[-a_{\max}, a_{\max}]^3$ of mean 0 and τ_k in $[0, \bar{\tau}_{\max}]$ having mean $\mathbb{E}[\tau_k] \in [\bar{\tau}_1, \bar{\tau}_2]$.

- Similarly, we express the (unknown) transfer function with respect to some basis of a discretization of the time interval of interest — say, of dimension Q.
- Reasonable values for all these parameters ($a_{\max}$, I, *& c.*) can be found in the literature.⟨8⟩

⟨8⟩ L. Esteva (1970) "Seismic risk and seismic design" in *Seismic Design for Nuclear Power Plants* The MIT Press.

Application of OUQ Reduction Theorems

- Crucially, the OUQ reduction theorems apply to give a reduced OUQ problem of dimension $18I + Q$ since
  - there are no constraints on the transfer function (and so the optimizer will find the deterministically "worst" one — one Dirac mass), and there are Q such deterministic components;
  - the mean constraints on the three components of each X_i generate a need for four Dirac masses, and there are I of these;
  - the mean constraints on each τ_k generate need for two Dirac masses, and there are I of these.

Critical Excitation

- Without constraints, worst-case scenarios correspond to focusing the energy of the earthquake in modes of resonances of the structure.
- Without correlations in the ground motion these scenarios correspond to rare events in which independent random variables must conspire to strongly excite a specific resonance mode.
- The lack of information on the transfer function ψ and the mean values $\mathbb{E}[\tau_k]$ permits scenarios characterized by strong correlations in ground motion where the energy of the earthquake can be focused in the above-mentioned modes of resonance.

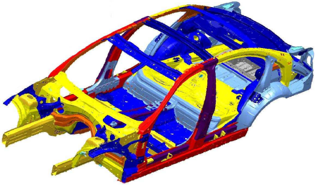
Example 3: Multiphase Steels

Advanced High-Strength Steels (AHSS) offer many advantages in *e.g.* automobile construction:

- light-weight construction;
- enhanced crash safety.

AHSS have complex microstructure, involving two or more phases, leading to a complex macroscopic response (anisotropy, kinematic hardening, *& c.*). Significant sources of uncertainty include:

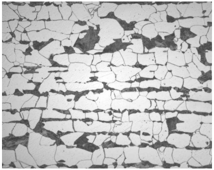
- microstructure morphology;
- material properties of individual phases.



From www.bmw.de

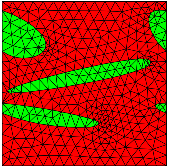
Description of the Material Problem

- We use a micro-macro approach: a microscopic BVP is solved at each macroscopic Gauss point; the microscopic BVP is posed on a statistically similar representative volume element (SSRVE).
- For simplicity, we will assume that there is no error in this model except for the statistical accuracy of the SSRVE. More generally, we would have to incorporate model error, perhaps by a "V&V distance" between computational code and physical reality.



(a) Micrograph of DP steel

(b) SSRVE with spline-shaped inclusions



(c) SSRVE meshed for FE calculation

Description of the Material Problem

- Ferrite matrix phase with material parameters $\boldsymbol{y}_{\mathsf{mat}}$.
- Perlite or austenite inclusion phase with material parameters $\boldsymbol{y}_{\mathsf{inc}}$.
- Inclusion morphology described by a parameter γ.
  - Matrix-inclusion interface described by a collection of splines.
  - The number of splines, and their control points, are part of the parameter γ.
- By "statistical similarity", we mean that the simulated microstructure matches certain statistics of the real microstructure as seen in a scanned micrograph:
  - known ranges/mean values for the parameters $\boldsymbol{y}$ and γ;
  - known range/mean value for volume fraction $\mathcal{P}_V$ occupied by the inclusion phase;
  - higher-order probability functions.

$\mathcal{A}$ for the Material Problem

- Bounds constraints for basic variables:

$$\begin{aligned} \boldsymbol{\gamma} &\in [\boldsymbol{\gamma}^-, \boldsymbol{\gamma}^+], \\ \boldsymbol{y}_{\mathsf{inc}} &\in [\boldsymbol{y}^-_{\mathsf{inc}}, \boldsymbol{y}^+_{\mathsf{inc}}], \\ \boldsymbol{y}_{\mathsf{mat}} &\in [\boldsymbol{y}^-_{\mathsf{mat}}, \boldsymbol{y}^+_{\mathsf{mat}}]. \end{aligned}$$

- Bounds constraints for microstructure statistics:

volume fraction of inclusion phase: $\mathcal{P}_V(\boldsymbol{\gamma}, \boldsymbol{y}) \in [\mathcal{P}_V^-, \mathcal{P}_V^+]$.

- Mean constraint for microstructure statistics:

volume fraction of inclusion phase: $\mathbb{E}_\mu[\mathcal{P}_V(\boldsymbol{\gamma}, \boldsymbol{y})] = \overline{\mathcal{P}_V}$.

- Additional technical constraints:

the individual inclusion phases are not allowed to intersect each other.

A Typical Reliability/Certification Problem

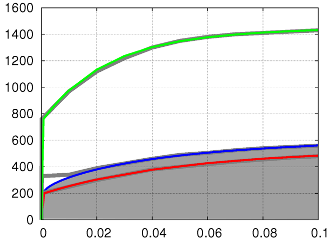
For simplicity, we choose a single scalar performance measure: the area under the (macroscopic) stress-strain curve in a uniaxial tension test:

$$G(\boldsymbol{\gamma}, \boldsymbol{y}) := \int_0^{1/10} \sigma(\epsilon, \boldsymbol{\gamma}, \boldsymbol{y}) \, \mathrm{d}\epsilon.$$

Given a performance threshold θ, we want to certify that the probability of failure

$$\mathbb{P}[G(\boldsymbol{\gamma}, \boldsymbol{y}) \leq \theta]$$

is acceptably small, with respect to the random variable $(\boldsymbol{\gamma}, \boldsymbol{y})$.



Uniaxial test stress-strain curves for several different realizations of the microstructure parameters

Further Work and Conclusions

Other UQ Problems

- Certification — as in the bounding of probabilities of failure — is not the only UQ problem.
- However, certification appears to be a central and prototypical UQ problem; many other UQ problems can be posed as certification problems.
- For example, verification and validation problems are very easily transformed into certification problems.
- "Verification deals with mathematics; validation deals with physics",[9] but otherwise the two problems are very similar.

[9] P. J. Roache (1998) *Verification and Validation in Computational Science and Engineering.* Hermosa Publ., Albuquerque.

Verification

- Verification of a computational model means showing that it is "acceptably close" to the mathematical description of the processes that it is supposed to model.
- Let U denote the exact mathematical solution to the problem, and F the computational model.
- The verification problem is now to show that

$$\mathbb{P}[\|U(X) - F(X)\| \leq \theta] \geq 1 - \epsilon$$

for suitable threshold values θ and ϵ, a suitable norm/distance $\| \cdot \|$, and with $\mathbb{P}$-distributed input parameters X taking values in some verification domain.

Validation

- Validation of a computational model means showing that it is "acceptably close" to the processes that it is supposed to model. (In verification, the problem is mathematical; in validation, the problem is physical.)
- Let F denote the computational model and P the physical process.
- The validation problem is now to show that

$$\mathbb{P}[\|F(X) - P(X)\| \leq \theta] \geq 1 - \epsilon$$

for suitable threshold values θ and ϵ, a suitable norm/distance $\| \cdot \|$, and with $\mathbb{P}$-distributed input parameters X taking values in some validation domain.

Prediction

- What is a "prediction" about some real-valued quantity of interest $G(X)$? In a loose sense, it is an interval $[a, b] \subseteq \mathbb{R}$ (hopefully "as small as possible") such that

$$\mathbb{P}[a \leq G(X) \leq b] \approx 1.$$

- In the OUQ paradigm, this can be made more precise: given a collection of admissible scenarios $\mathcal{A}$ for $(G, \mathbb{P})$, and $\epsilon > 0$, we seek

$$\sup \left\{ a \in \mathbb{R} \,\middle|\, \inf_{(f,\mu) \in \mathcal{A}} \mu[f \geq a] \geq 1 - \frac{\epsilon}{2} \right\},$$

$$\inf \left\{ b \in \mathbb{R} \,\middle|\, \inf_{(f,\mu) \in \mathcal{A}} \mu[f \leq b] \geq 1 - \frac{\epsilon}{2} \right\}.$$

Optimal UQ with Random Sample Data

- Another area of research is what it means to give optimal bounds on the probability of failure when the information is not just deterministic, but also comes from random sample data.
- If we are told that we will observe a realization D of sample data from a space $\mathcal{D}$ (*e.g.* ten IID samples), we seek an “upper-bounder” $\Psi \colon \mathcal{D} \to [0, 1]$ such that

$$\inf_{(f,\mu) \in \mathcal{A}} \mu^{\mathcal{D}} \big[\Psi(D) \geq \mu[f \leq \theta] \big] \geq 1 - \epsilon,$$

  and we want Ψ to be “minimal” in some sense.
- This is a very delicate issue, with great scope for supplier–client conflict and the classical paradoxes of voting theory (*e.g.* Arrow's theorem) to intrude.
- However, it does appear that there is a connection between OUQ with sample data and the notion of a Uniformly Most Powerful hypothesis test.

Conclusions

- UQ is an essential component of modern science, with many high-consequence applications. However, there is no established consensus on how to formally pose "the UQ problem", nor a common language in which to communicate and quantitatively compare UQ methods and results.
- OUQ is an opening gambit. OUQ is not just an effort to provide answers, but an effort to well-pose the question: OUQ is the challenge of optimally bounding functions of unknown responses and unknown probabilities, given some information about them.
- A key feature is that the OUQ viewpoint explicitly requires the user to explicitly state all the assumptions in operation — once listed, they can be perturbed to see if the answers are robust.
- Although the optimization problems involved are large, in many cases of interest, their dimension can be substantially reduced.

References

- H. Owhadi, C. Scovel, T. J. Sullivan, M. McKerns & M. Ortiz. **"Optimal Uncertainty Quantification"** (2010) Submitted to *SIAM Review*. Preprint at

  http://arxiv.org/pdf/1009.0679v1

- Optimization calculations performed using *mystic*:

  http://dev.danse.us/trac/mystic