Optimal Uncertainty Quantification for Hypervelocity Impact

Tim Sullivan

Mathematics Institute, University of Warwick
and Caltech PSAAP Center

Uncertainty Quantification Short Course

Stanford University, Palo Alto, CA, U.S.A.

1–3 June 2014

THE UNIVERSITY OF WARWICK

http://www.tjsullivan.org.uk/pdf/2014-06-02_stanford.pdf

Credits

Joint work with everyone (!) at the Caltech PSAAP Center, but in particular the core UQ team of Paul-Hervé Kamga, Michael McKerns, Lan Huong Nguyen, Michael Ortiz, Houman Owhadi and Clint Scovel.

Portions of this work were supported by the U. S. Department of Energy NNSA under Award No. 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.

CALTECH PSAAP

National Nuclear Security Administration

Prototypical UQ Problem: Reliability Certification

- $g^\dagger \colon \mathbb{X} \to \mathbb{Y}$ is a system of interest, with random inputs X distributed according to a probability measure $\mu^\dagger$ on $\mathbb{X}$.
- For some subset $\mathcal{F} \subseteq \mathbb{Y}$, the event $[g^\dagger(X) \in \mathcal{F}]$ constitutes failure; we want to know the probability of failure

$$\mathbb{P}_{\mu^\dagger}\left[g^\dagger(X) \in \mathcal{F}\right] \equiv \underbrace{\mathbb{E}_{\mu^\dagger}\left[\mathbb{1}\left[g^\dagger(X) \in \mathcal{F}\right]\right]}_{\substack{\text{“just” an integral} \\ \text{to be evaluated} \\ \text{— directly?} \\ \text{— by MC?} \\ \text{— by quadrature?}}},$$

  or at least to know that it is acceptably small (or unacceptably large!).
- **Problem:** In practical applications, one does not know the Universe's $g^\dagger$ and $\mu^\dagger$ exactly!

Other Quantities of Interest

- For some quantity of interest (measurable function) $q \colon \mathbb{X} \times \mathbb{Y} \to \mathbb{R}$, we want to know

$$\underbrace{\mathbb{E}_{\mu^\dagger}\left[q(X, g^\dagger(X))\right]}_{\substack{\text{“just” an integral} \\ \text{to be evaluated} \\ \text{— directly?} \\ \text{— by MC?} \\ \text{— by quadrature?}}},$$

  or at least to know that it is acceptably small (or unacceptably large!).
- For example:
  - failure probability: $q(x, y) = \mathbb{1}[y \in \mathcal{F}]$,
  - mean performance: $q(x, y) = y$,
  - variance about a nominal output value: $q(x, y) = |y - y_0|^2$.
- Our interest lies in understanding $\mathbb{E}_{\mu^\dagger}\left[q(X, g^\dagger(X))\right]$ when $g^\dagger$ and $\mu^\dagger$ are only imperfectly known (i.e. epistemic uncertainty), and to obtain bounds that are optimal with respect to the known information.

Optimal UQ

- The initial step in the Optimal Uncertainty Quantification approach is specifying a feasible set of admissible scenarios (g, μ) that could be $(g^\dagger, \mu^\dagger)$ according to the available information:

$$\mathcal{A} = \left\{ (g, \mu) \;\middle|\; \begin{array}{c} (g, \mu) \text{ is consistent with the current} \\ \text{information about } (g^\dagger, \mu^\dagger) \\ \text{(e.g. legacy data, models, theory, expert judgement)} \end{array} \right\}.$$

- A priori, all we know about reality is that $(g^\dagger, \mu^\dagger) \in \mathcal{A}$; we have no idea exactly which (g, μ) in $\mathcal{A}$ is actually $(g^\dagger, \mu^\dagger)$.
  - No $(g, \mu) \in \mathcal{A}$ is "more likely" or "less likely" to be $(g^\dagger, \mu^\dagger)$.
  - Particularly in high-consequence settings, it makes sense to adopt a posture of healthy conservatism and determine the best and worst outcomes consistent with the information encoded in $\mathcal{A}$.
- Dialogue between UQ practitioners and the domain experts is essential in formulating — and revising — $\mathcal{A}$.

Optimal UQ

$$\mathcal{A} = \left\{ (g, \mu) \,\middle|\, \begin{array}{c} (g, \mu) \text{ is consistent with the current} \\ \text{information about (i.e. could be) } (g^\dagger, \mu^\dagger) \end{array} \right\}$$

- Optimal bounds (w.r.t. the information encoded in $\mathcal{A}$) on the quantity of interest $\mathbb{E}_{\mu^\dagger}[q(X, g^\dagger(X))]$ are found by minimizing/maximizing $\mathbb{E}_\mu[q(X, g(X))]$ over all admissible scenarios $(g, \mu) \in \mathcal{A}$:

$$\underline{Q}(\mathcal{A}) \leq \mathbb{E}_{\mu^\dagger}[q(X, g^\dagger(X))] \leq \overline{Q}(\mathcal{A}),$$

where $\underline{Q}(\mathcal{A})$ and $\overline{Q}(\mathcal{A})$ are defined by the optimization problems

$$\underline{Q}(\mathcal{A}) := \inf_{(g,\mu)\in\mathcal{A}} \mathbb{E}_\mu[q(X, g(X))],$$

$$\overline{Q}(\mathcal{A}) := \sup_{(g,\mu)\in\mathcal{A}} \mathbb{E}_\mu[q(X, g(X))].$$

- Cf. generalized Chebyshev inequalities in decision analysis (**Smith** (1995)), imprecise probability (**Boole** (1854)), distributionally robust optimization, robust Bayesian inference (surv. **Berger** (1984)).

Reduction of OUQ Problems — LP Analogy

Dimensional Reduction

- A priori, OUQ problems are infinite-dimensional, non-convex*, highly-constrained, global optimization problems.
- However, they can be reduced to equivalent finite-dimensional problems in which the optimization is over the extremal scenarios of $\mathcal{A}$.
- The dimension of the reduced problem is proportional to the number of probabilistic inequalities that describe $\mathcal{A}$.

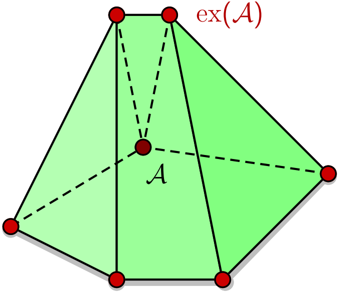


Figure : A linear functional on a convex domain in $\mathbb{R}^n$ finds its extreme value at the extremal points of the domain; similarly, OUQ problems reduce to searches over finite-dimensional families of extremal scenarios.

*But see e.g. **Bertsimas & Popescu** (2005) and **Smith** (1995) for convex special cases.

Reduction of OUQ Problems — Heuristic

Heuristic

If you have N_k pieces of information relevant to the random variable X_k, then just pretend that X_k takes at most $N_k + 1$ values in $\mathbb{X}_k$.

Reduction of OUQ Problems — Heuristic

Heuristic

If you have N_k pieces of information relevant to the random variable X_k, then just pretend that X_k takes at most $N_k + 1$ values in $\mathbb{X}_k$.

- To make this heuristic rigorous, we restrict attention to Radon spaces, "nice" spaces on which every Borel probability measure is inner regular.
- Our theorem builds on now-classical results by **von Weizsäcker & Winkler** (1980) and **Winkler** (1988) characterizing the extremal measures in moment classes, and "nice" linear/affine functionals on such classes.
- Important point: the extremal measures of a moment class

$$\left\{ \mu \in \mathcal{P}(\mathbb{X}) \,\middle|\, \mathbb{E}_\mu[\varphi_1] \leq 0, \dots, \mathbb{E}_\mu[\varphi_n] \leq n \right\}$$

  are the discrete measures that have support on at most $n + 1$ distinct points of $\mathbb{X}$, which we denote by $\Delta_n(\mathbb{X})$.

Reduction of OUQ Problems — Theorem

Heuristic

If you have N_k pieces of information relevant to the random variable X_k, then just pretend that X_k takes at most $N_k + 1$ values in $\mathbb{X}_k$.

Theorem (Generalized moment and indep. constraints)

Suppose that $\mathbb{X} := \mathbb{X}_1 \times \dots \times \mathbb{X}_K$ is a product of Radon spaces. Let

$$\mathcal{A} := \left\{ (g, \mu) \,\middle|\, \begin{array}{c} g \colon \mathbb{X} \to \mathbb{R} \text{ is measurable, } \mu = \mu_1 \otimes \dots \otimes \mu_K \in \bigotimes_{k=1}^{K} \mathcal{P}(\mathbb{X}_k); \\ \langle \text{conditions on } g \text{ alone} \rangle \text{; and, for each } g, \\ \text{for some measurable functions } \varphi_i \colon \mathbb{X} \to \mathbb{R} \text{ and } \varphi_i^{(k)} \colon \mathbb{X}_k \to \mathbb{R}, \\ \mathbb{E}_{\mu}\big[\varphi_i\big] \leq 0 \text{ for } i = 1, \dots, n_0, \\ \mathbb{E}_{\mu_k}\big[\varphi_i^{(k)}\big] \leq 0 \text{ for } i = 1, \dots, n_k \text{ and } k = 1, \dots, K \end{array} \right\}$$

$$\mathcal{A}_\Delta := \left\{ (g, \mu) \in \mathcal{A} \,\middle|\, \begin{array}{c} \mu_k \in \Delta_{N_k}(\mathbb{X}_k) \\ \text{where } N_k := n_0 + n_k \end{array} \right\} \subseteq \mathcal{A}.$$

Then $\underline{Q}(\mathcal{A}) = \underline{Q}(\mathcal{A}_\Delta)$ *and* $\overline{Q}(\mathcal{A}) = \overline{Q}(\mathcal{A}_\Delta)$.

Reduction of OUQ Problems — Consequence

Heuristic

If you have N_k pieces of information relevant to the random variable X_k, then just pretend that X_k takes at most $N_k + 1$ values in $\mathbb{X}_k$.

- Computation of the OUQ bounds $\underline{Q}(\mathcal{A})$ and $\overline{Q}(\mathcal{A})$ is equivalent to finite-dimensional problems in which the optimization variables are
  - the positions of the support points $x_i \in \mathbb{X}$ of the discrete measure μ;
  - the weights $w_i \in [0, 1]$ of the points x_i; and
  - the response values $y_i \in \mathbb{Y}$ corresponding to $g(x_i)$.

  with objective function

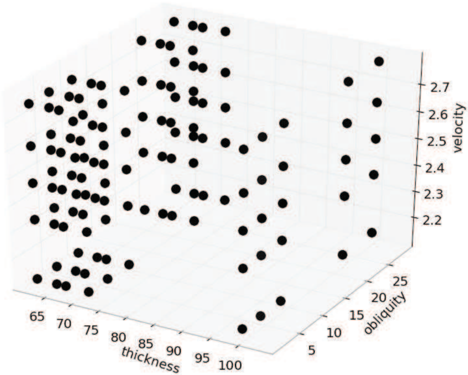
$$\sum_{i=(0,\dots,0)}^{(N_1,\dots,N_K)} w_i \, q(x_i, y_i)$$

  and similar finite sums for the constraints.
- $\implies$ Implementation in the general-purpose open-source Mystic optimization framework, written in Python.

Reduction of OUQ Problems — Consequence

Heuristic

If you have N_k pieces of information relevant to the random variable X_k, then just pretend that X_k takes at most $N_k + 1$ values in $\mathbb{X}_k$.



Optimal Concentration Inequalities

Classical inequalities of probability theory can be seen as OUQ statements:

Example: Markov's Inequality in OUQ Form

$$\mathcal{A}_{\mathrm{M}} := \{\mu \in \mathcal{P}([0,\infty)) \,|\, \mathbb{E}_\mu[X] \leq m\}$$

Suppose 'failure' is $X \geq t$, for $t \geq m$. Then

$$\begin{aligned}\overline{P}(\mathcal{A}_{\mathrm{M}}) &= \sup_{\mu \in \mathcal{A}_{\mathrm{M}}} \mathbb{P}_\mu[X \geq t] \\ &= \sup \left\{ \sum_{i=0}^{1} w_i \mathbb{1}[x_i \geq t] \,\middle|\, w_i, x_i \geq 0, \sum_{i=0}^{1} w_i = 1, \sum_{i=0}^{1} w_i x_i \leq m \right\} \\ &= 1 - \frac{m}{t}.\end{aligned}$$

Optimal Concentration Inequalities

Classical inequalities of probability theory can be seen as OUQ statements:

Example: Markov's Inequality in OUQ Form

$$\mathcal{A}_{\mathrm{M}} := \{\mu \in \mathcal{P}([0,\infty)) \mid \mathbb{E}_\mu[X] \leq m\}$$

Suppose 'failure' is $X \geq t$, for $t \geq m$. Then

$$\overline{P}(\mathcal{A}_{\mathrm{M}}) = \sup_{\mu \in \mathcal{A}_{\mathrm{M}}} \mathbb{P}_\mu[X \geq t] = 1 - \frac{m}{t}.$$

How about other deviation/concentration-of-measure inequalities?

- McDiarmid's inequality: deviations from the mean of bounded-differences functions of independent random variables.
- Hoeffding's inequality: deviations from the mean of sums of independent random variables.

McDiarmid's (a.k.a. Bounded Differences) Inequality

$$\mathcal{A}_{\mathrm{McD}} := \left\{ (g,\mu) \;\middle|\; \begin{array}{c} g \colon \mathbb{X} := \mathbb{X}_1 \times \dots \times \mathbb{X}_K \to \mathbb{R}, \\ \mu = \bigotimes_{k=1}^{K} \mu_k, \text{ (i.e. } X_1, \dots, X_K \text{ independent)} \\ \mathbb{E}_\mu[g(X)] \geq m \geq 0, \\ \operatorname{osc}_k(g) \leq D_k \text{ for each } k \in \{1, \dots, K\} \end{array} \right\},$$

with componentwise oscillations/global sensitivities defined by

$$\operatorname{osc}_k(g) := \sup \left\{ |g(x) - g(x')| \;\middle|\; \begin{array}{c} x, x' \in \mathbb{X}_1 \times \dots \times \mathbb{X}_K, \\ x_i = x'_i \text{ for } i \neq k \end{array} \right\}.$$

Theorem (McDiarmid's Inequality, 1988)

$$\overline{P}(\mathcal{A}_{McD}) := \sup_{(g,\mu) \in \mathcal{A}_{McD}} \mathbb{P}_\mu[g(X) \leq 0] \overset{!!!}{\leq} \exp\left(- \frac{2m^2}{\sum_{k=1}^{K} D_k^2} \right)$$

Optimal McDiarmid and Screening Effects

Theorem (Optimal McDiarmid for $K = 1, 2$)

For $K = 1$,

$$\overline{P}(\mathcal{A}_{McD}) = \begin{cases} 0, & \text{if } D_1 \leq m, \\ 1 - \dfrac{m}{D_1}, & \text{if } 0 \leq m \leq D_1. \end{cases}$$

For $K = 2$,

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

In the highlighted case, $\min\{D_1, D_2\}$ carries no information — not in the sense of 0 bits, but the sense of being a non-binding constraint.

Optimal Hoeffding and the Effects of Nonlinearity

- Similarly, one can consider $\mathcal{A}_{\mathrm{Hfd}}$ "$\subseteq$" $\mathcal{A}_{\mathrm{McD}}$ corresponding to the assumptions of Hoeffding's inequality, which bounds deviation probabilities of sums of independent bounded random variables:

$$\mathcal{A}_{\mathrm{Hfd}} := \left\{ (g, \mu) \,\middle|\, \begin{array}{c} g \colon \mathbb{R}^K \to \mathbb{R} \text{ given by} \\ g(x_1, \dots, x_K) := x_1 + \dots + x_K, \\ \mu = \mu_1 \otimes \dots \otimes \mu_K \text{ supported on a cuboid of} \\ \text{side lengths } D_1, \dots, D_K, \text{ and } \mathbb{E}_\mu[g(X)] \geq m \geq 0 \end{array} \right\}.$$

- Hoeffding's inequality is the bound

$$\overline{P}(\mathcal{A}_{\mathrm{Hfd}}) := \sup_{(g,\mu) \in \mathcal{A}_{\mathrm{Hfd}}} \mathbb{P}_\mu[g(X) \leq 0] \leq \exp\left(- \frac{2m^2}{\sum_{k=1}^K D_k^2} \right).$$

- Interestingly, $\overline{P}(\mathcal{A}_{\mathrm{Hfd}}) = \overline{P}(\mathcal{A}_{\mathrm{McD}})$ for $K = 1$ and $K = 2$, but $\overline{P}(\mathcal{A}_{\mathrm{Hfd}}) \leq \overline{P}(\mathcal{A}_{\mathrm{McD}})$ for $K = 3$, and the inequality can be strict. Thus, sometimes linearity is binding information, sometimes not.

OUQ with Legacy Data

- An interesting class of admissible function-measure pairs arises in the case of partially observed smooth enough functions, e.g.

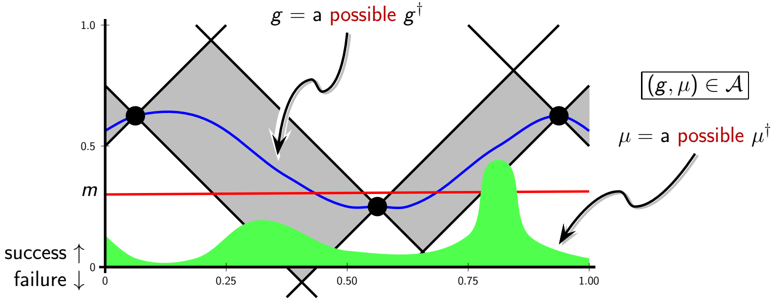
$$\mathcal{A} = \left\{ (g, \mu) \,\middle|\, \begin{array}{c} g \colon \mathbb{X} \to \mathbb{Y} \text{ has prescribed smoothness,} \\ g = g^{\dagger} \text{ on } \mathcal{O} \subseteq \mathbb{X} \text{ (i.e. some legacy data),} \\ \mu \in \mathcal{P}(\mathbb{X}),\ \mathbb{E}_{\mu}[\varphi_i] \leq 0 \text{ for } i = 1, \dots, n \end{array} \right\}$$

- Note that $\mathcal{O}$ need not be statistically representative.
- Simple examples of "smooth enough": Lipschitz constants or Hölder conditions.
- Mathematically interesting interactions between the measure-theoretic constraints and the metric geometry of the space $\mathbb{X}$, e.g. the fact that any parially-defined Lipschitz function can be extended to the whole space without changing the Lipschitz constant (**McShane** (1934)).

Example Reduction: 1 Random Variable, 1 Constraint

The original problem entails optimizing over an infinite-dimensional collection of (g, μ) that could be $(g^\dagger, \mu^\dagger)$. In the reduced problem, we only have to move around and re-weight two Dirac measures (point masses) and the values of g over those two points.

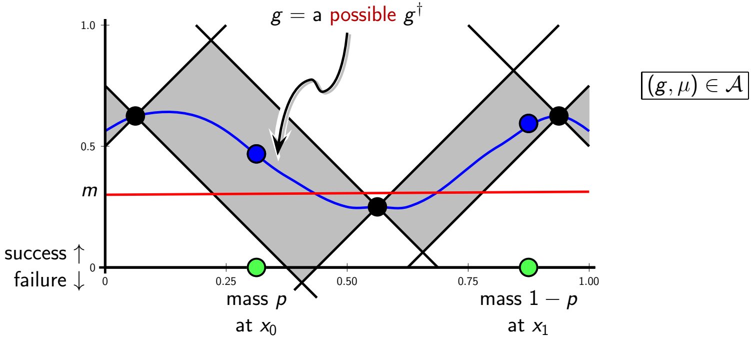
infinite-dimensional problem $\rightsquigarrow$ equivalent 5-dimensional problem!



Example Reduction: 1 Random Variable, 1 Constraint

The original problem entails optimizing over an infinite-dimensional collection of (g, μ) that could be $(g^\dagger, \mu^\dagger)$. In the reduced problem, we only have to move around and re-weight two Dirac measures (point masses) and the values of g over those two points.

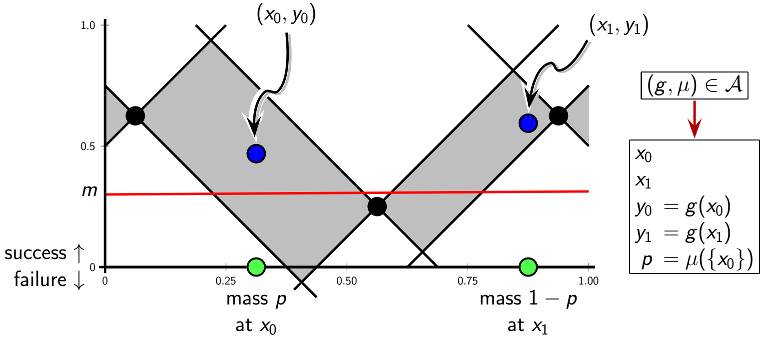
infinite-dimensional problem $\rightsquigarrow$ equivalent 5-dimensional problem!



Example Reduction: 1 Random Variable, 1 Constraint

The original problem entails optimizing over an infinite-dimensional collection of (g, μ) that could be $(g^\dagger, \mu^\dagger)$. In the reduced problem, we only have to move around and re-weight two Dirac measures (point masses) and the values of g over those two points.

infinite-dimensional problem $\rightsquigarrow$ equivalent 5-dimensional problem!



Explicit Solution: 1 Random Variable, 1 Data Point

- The case of a single observation in 1d can be solved explicitly.
- Suppose that you have one observation $(z, g^\dagger(z)) \in [0, \frac{1}{2}] \times \mathbb{R}$ of a function $g^\dagger \colon [0, 1] \to \mathbb{R}$ with Lipschitz constant $L \geq 0$.
- Explicit piecewise and discontinuous least upper bound on $\mathbb{P}_{\mu^\dagger}[g^\dagger(X) \leq 0]$ given L, $(z, g^\dagger(z))$, and that $\mathbb{E}_{\mu^\dagger}[g^\dagger(X)] \geq m$:

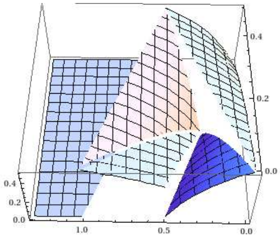


Figure : Surface plot of the least upper bound $\overline{P}$ on $\mathbb{P}_{\mu^\dagger}[g^\dagger(X) \leq 0]$, as a function of the observed data point $(z, g^\dagger(z))$.

Caltech's Hypervelocity Impact Setup

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 $\text{km} \cdot \text{s}^{-1}$.

Caltech's Hypervelocity Impact Setup

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 $\text{km} \cdot \text{s}^{-1}$.

3-Variable Hypervelocity Impact Example

- Legacy data = 32 data points (steel-on-aluminium shots A48–A81, less two mis-fires) from summer 2010 at Caltech's SPHIR facility:

$$X = (h, \alpha, v) \in \mathbb{X} := [0.062, 0.125]\,\text{in} \times [0, 30]\,\text{deg} \times [2300, 3200]\,\text{m/s}.$$

  Output $g^\dagger(h, \alpha, v)$ = the induced perforation area in mm^2; the data set contains results between $6.31\,\text{mm}^2$ and $15.36\,\text{mm}^2$.
- Failure event is $[g^\dagger(h, \alpha, v) \leq \theta]$, for various values of θ.
- Constrain the mean perf. area: $\mathbb{E}_{\mu^\dagger}[g^\dagger(h, \alpha, v)] \geq m := 11.0\,\text{mm}^2$.
- Modified Lipschitz constraint (multi-valued data):

$$L = \left(\frac{175.0}{\text{in}}, \frac{0.075}{\text{deg}}, \frac{0.1}{\text{m/s}} \right) \text{mm}^2$$

$$|y - y'| \leq \sum_{k=1}^{3} L_k |x_k - x'_k| + 1.0\,\text{mm}^2.$$

3-Parameter Hypervelocity Impact Example: Results

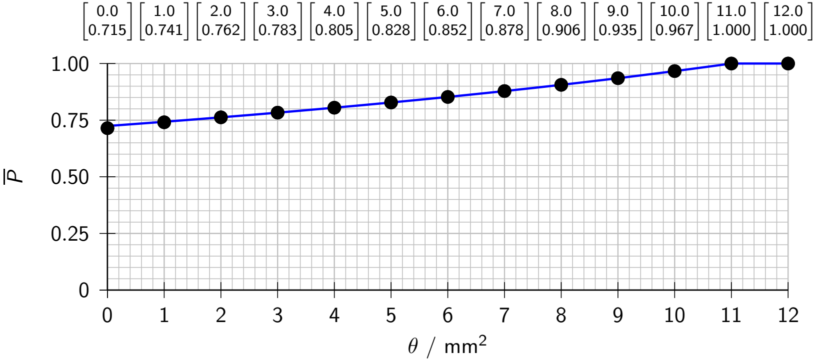


Figure : Maximum probability that perforation area is $\leq \theta$, for various θ, with the data and assumptions of the previous slide, including mean perforation area $\mathbb{E}[g^\dagger(h, \alpha, v)] \geq 11.0\,\mathrm{mm}^2$. For $\theta \geq 2\,\mathrm{mm}^2$, the results are within 10^{-6} of Markov's bound, which indicates that 2 binding data points are those that constrain the maximum of the response function; the other 30 are non-binding.

Models and Neighbourhoods

- One can consider feasible sets in which the constraints on g are of the form $d(g, F) \leq C$ for some model function F.
- There are good and bad choices for the distance function d:

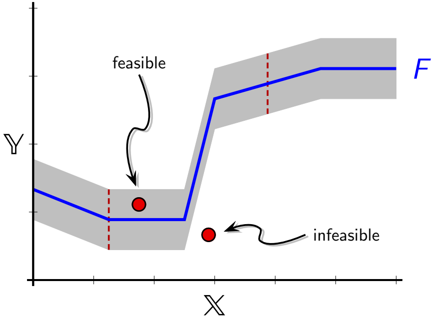


Figure : Assuming that reality $g^\dagger$ is uniformly close to the model F means assuming that the model has approximately the right cliffs in exactly the right places; Hausdorff (graphical) closeness is a much looser assumption.

Models and Neighbourhoods

- One can consider feasible sets in which the constraints on g are of the form $d(g, F) \leq C$ for some model function F.
- There are good and bad choices for the distance function d:

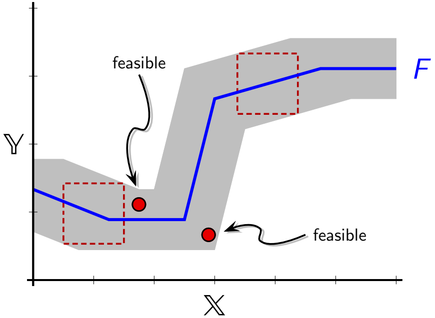


Figure : Assuming that reality $g^{\dagger}$ is uniformly close to the model F means assuming that the model has approximately the right cliffs in exactly the right places; Hausdorff (graphical) closeness is a much looser assumption.

Hypervelocity Impact Application

- System is characterized by three input parameters: target plate thickness h, obliquity α, and impact velocity v, with assumed ranges $h \in \mathbb{X}_h := \{0.5, 1.5, 3.0\}\,\mathrm{mm}$, $\alpha \in \mathbb{X}_\alpha := [0, 60]^\circ$, and $v \in \mathbb{X}_v := [4.5, 7.0]\,\mathrm{km} \cdot \mathrm{s}^{-1}$. Input space is $\mathbb{X} := \mathbb{X}_h \times \mathbb{X}_\alpha \times \mathbb{X}_v$.
- Perforation area is the main performance measure of the system, which is expected to lie in the output space $\mathbb{Y} := [0, 39.73]\,\mathrm{mm}^2$. We want to bound $\mathbb{P}[g^\dagger(h, \alpha, v) \leq \theta]$ for threshold area values θ.
- The model function F is the Optimal Transportation Meshfree method. (A **lot** swept under the carpet here!)
- Model-reality mismatch quantified as $d(g^\dagger, F) \leq \delta$. In practice, we fix a confidence level $0 < \eta < 1$, and use legacy data points to find $\delta(\eta)$ such that $d(g^\dagger, F) \leq \delta(\eta)$ with probability $\geq \eta$.

Mean Constraints on Outputs

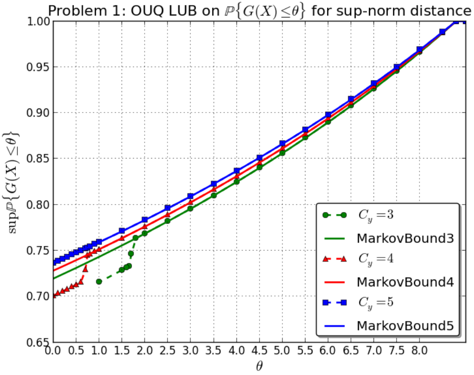


Figure : OUQ least upper bounds on perforation area probabilities given $d(g^\dagger, F) \leq \delta$ and bounds on $\mathbb{E}_\mu[g^\dagger]$. Note the relative insensitivity (for $\theta \geq 2\,\mathrm{mm}^2$) to both δ and the choice of d as the uniform or Hausdorff distance. Note also that the closeness to the Markov bounds (solid curves), indicating that the binding information is the implied maximum perforation area.

Mean Constraints on Outputs

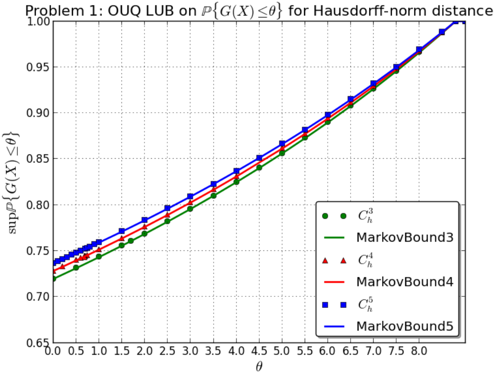


Figure : OUQ least upper bounds on perforation area probabilities given $d(g^\dagger, F) \leq \delta$ and bounds on $\mathbb{E}_\mu[g^\dagger]$. Note the relative insensitivity (for $\theta \geq 2\,\mathrm{mm}^2$) to both δ and the choice of d as the uniform or Hausdorff distance. Note also that the closeness to the Markov bounds (solid curves), indicating that the binding information is the implied maximum perforation area.

Mean Constraints on Inputs

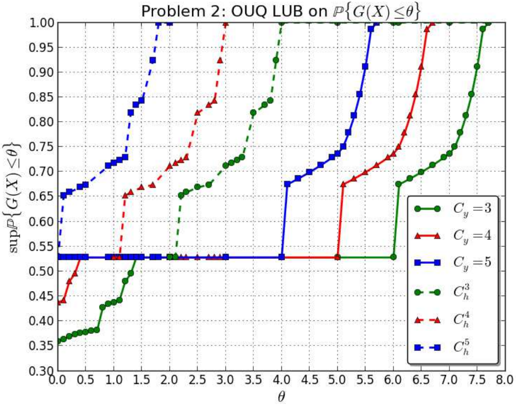


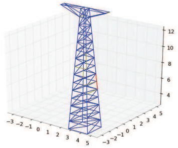
Figure : OUQ least upper bounds on perforation area probabilities given $d(g^\dagger, F) \leq \delta$ and bounds on $\mathbb{E}_\mu[h]$, $\mathbb{E}_\mu[\alpha]$, $\mathbb{E}_\mu[v]$. Note the strong sensitivity to both δ and the choice of d as the uniform (solid curves) or Hausdorff (dashed curves) distance.

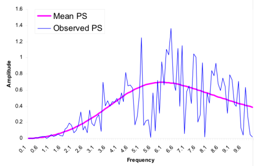
Seismic Safety Certification

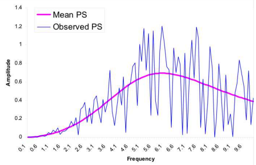
- Consider the survivability of a truss structure under an random earthquake of known intensity drawn from an incompletely specified probability distribution.
- Consider a random ground motion u, with the constraint that the mean power spectrum is the Matsuda–Asano shape function s_{MA}:

$$\mathbb{E}_{u\sim\mu}\left[|\hat{u}(\omega)|^2\right] = s_{\mathrm{MA}}(\omega) \propto \frac{\omega_{\mathrm{g}}^2\omega^2 e^{M_{\mathrm{L}}}}{(\omega_{\mathrm{g}}^2 - \omega^2)^2 + 4\xi_{\mathrm{g}}^2\omega_{\mathrm{g}}^2\omega^2}.$$

- Such shape functions are a common tool in the seismological community, but usually u is generated by filtering white noise through s.
- We used 200 3d Fourier modes, leading to a 1200-dimensional OUQ problem.







Reduction of the Random Power Spectrum

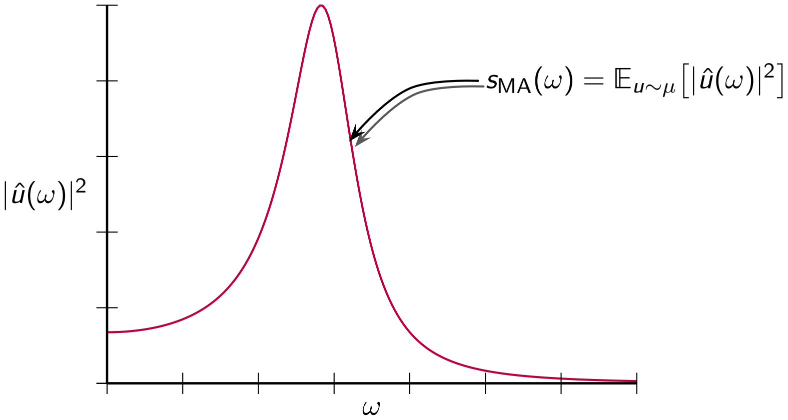


Figure : One mean constraint on each independent random Fourier mode $\hat{u}(\omega)$ (i.e. that $\mathbb{E}_{u\sim\mu}\left[|\hat{u}(\omega)|^2\right] = s_{\mathrm{MA}}(\omega)$) $\implies$ we get to pretend that $u(\omega)$ takes at most two distinct values which together satisfy this mean constraint.

Reduction of the Random Power Spectrum

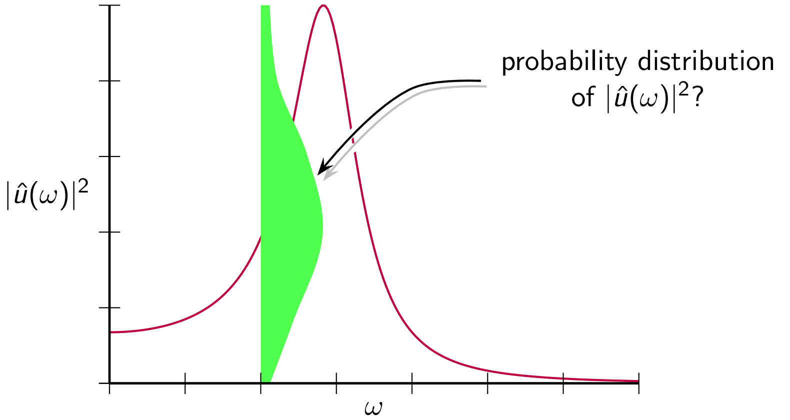


Figure : One mean constraint on each independent random Fourier mode $\hat{u}(\omega)$ (i.e. that $\mathbb{E}_{u\sim\mu}\left[|\hat{u}(\omega)|^2\right] = s_{\mathrm{MA}}(\omega)$) $\implies$ we get to pretend that $u(\omega)$ takes at most two distinct values which together satisfy this mean constraint.

Reduction of the Random Power Spectrum

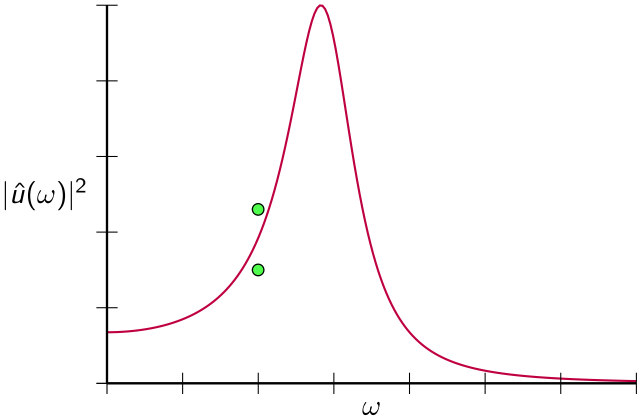


Figure : One mean constraint on each independent random Fourier mode $\hat{u}(\omega)$ (i.e. that $\mathbb{E}_{u\sim\mu}\left[|\hat{u}(\omega)|^2\right] = s_{\mathrm{MA}}(\omega)$) $\implies$ we get to pretend that $u(\omega)$ takes at most two distinct values which together satisfy this mean constraint.

Reduction of the Random Power Spectrum

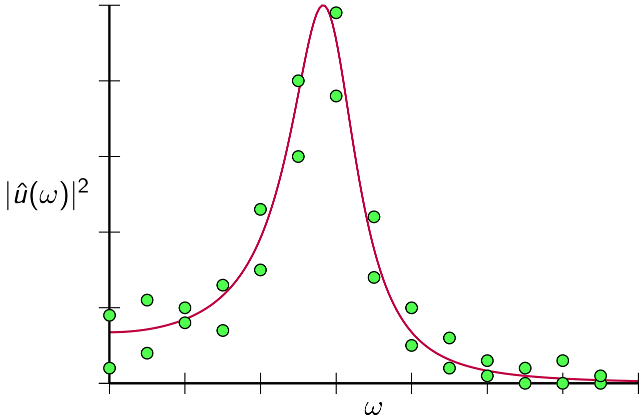


Figure : One mean constraint on each independent random Fourier mode $\hat{u}(\omega)$ (i.e. that $\mathbb{E}_{u\sim\mu}\left[|\hat{u}(\omega)|^2\right] = s_{\mathrm{MA}}(\omega)$) $\implies$ we get to pretend that $u(\omega)$ takes at most two distinct values which together satisfy this mean constraint.

Numerical Vulnerability Curves (CDF Envelopes)

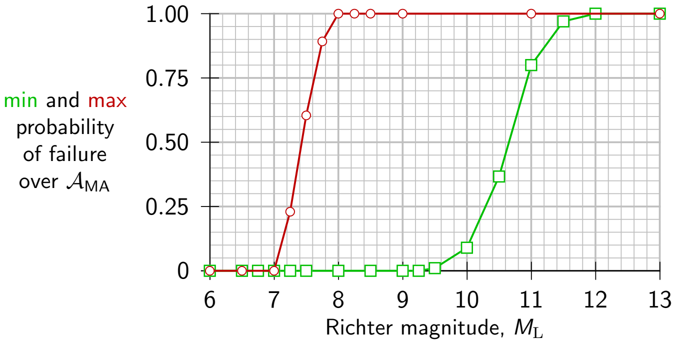


Figure : The minimum and maximum probability of failure as a function of Richter magnitude, M_{L}, where the ground motion u is constrained to have $\mathbb{E}_{\mu}[|\hat{u}|^2] =$ the Matsuda–Asano shape function s_{MA} with natural frequency ω_{g} and natural damping ξ_{g} taken from the 24 Jan. 1980 Livermore earthquake. Each data point required $\approx$ 1 day on 44+44 AMD Opterons (*shc* and *foxtrot* at Caltech). The forward model used 200 Fourier modes for a 3-dimensional ground motion u.

Numerical Vulnerability Curves (CDF Envelopes)

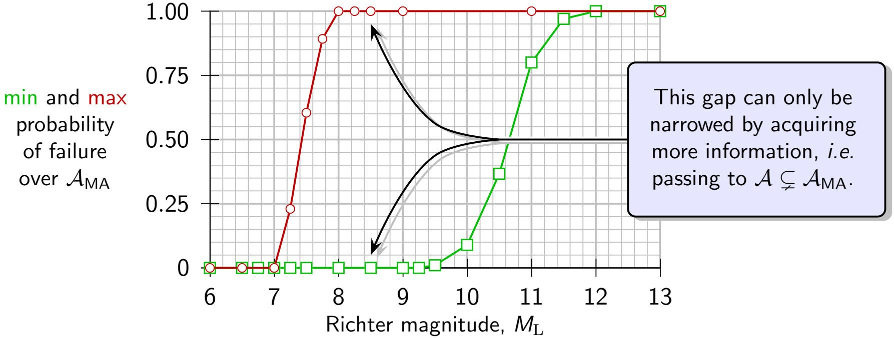


Figure : The minimum and maximum probability of failure as a function of Richter magnitude, M_{L}, where the ground motion u is constrained to have $\mathbb{E}_\mu[|\hat{u}|^2] =$ the Matsuda–Asano shape function s_{MA} with natural frequency ω_{g} and natural damping ξ_{g} taken from the 24 Jan. 1980 Livermore earthquake. Each data point required ≈ 1 day on 44+44 AMD Opterons (*shc* and *foxtrot* at Caltech). The forward model used 200 Fourier modes for a 3-dimensional ground motion u.

Numerical Vulnerability Curves (CDF Envelopes)

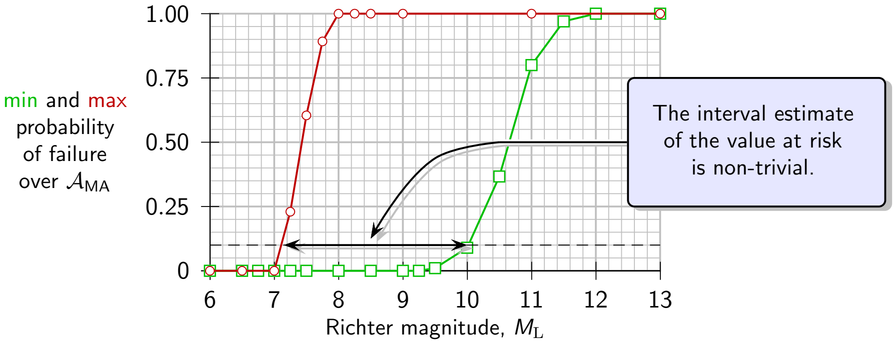


Figure : The minimum and maximum probability of failure as a function of Richter magnitude, M_{L}, where the ground motion u is constrained to have $\mathbb{E}_{\mu}[|\hat{u}|^2] =$ the Matsuda–Asano shape function s_{MA} with natural frequency ω_{g} and natural damping ξ_{g} taken from the 24 Jan. 1980 Livermore earthquake. Each data point required $\approx$ 1 day on 44+44 AMD Opterons (*shc* and *foxtrot* at Caltech). The forward model used 200 Fourier modes for a 3-dimensional ground motion u.

When and How to Use OUQ

- Use OUQ if you are strongly risk-averse, have unavoidable epistemic uncertainties, and have enough time to compute your way through the problem.
- Conversely, for real-time applications with simple and well-understood uncertainties, OUQ is impractical and overkill.
- Good features to include in your optimizer:
    - keep the functional parts of your optimizer as swappable modules, and pay attention to enforcing constraints;
    - cache past function evaluations;
    - look out for convex sub-problems in the non-convex OUQ problem;
    - look out for numerical 'collapse' of the discrete measure (dimension reduction $\implies$ huge cost savings).
- Personal rules of thumb: Differential Evolution works well with pop. size ≈ 40, 200 to 400 generations convergence criterion, run problems with 10s of support points and a fast model on a laptop overnight.

Conclusions

- By posing **UQ as an optimization problem** we
  - place the available information ($\cong$ constraints) about the input uncertainties at the centre of the problem;
  - obtain optimal bounds on output uncertainties w.r.t. that information;
  - get natural notions of information content in optimization-theoretic terms about constraints: active/inactive, binding/non-binding, . . .
- We have theoretical (closed-form pen-and-paper) and real (high-dimensional engineering systems) examples in hand showing these phenomena at work.
- Growing computational resources make large OUQ-type problems increasingly tractable, cf. Bayesian methods in 20th Century.
- Many research questions, especially concerning the inclusion of random sample data, algorithmic properties of OUQ, &c.

References

- Caltech (O)UQ Publications
  - P.-H. T. Kamga & al. "Optimal uncertainty quantification with model uncertainty and legacy data." *J. Mech. Phys. Solids.* Under review.
  - H. Owhadi & al. "Optimal Uncertainty Quantification." *SIAM Rev.* **55**(2):271–345, 2013. doi:10.1137/10080782X
  - T. J. Sullivan & al. "Optimal uncertainty quantification for legacy data observations of Lipschitz functions." *Math. Mod. Num. Anal.* **47**(6):1657–1689, 2013. doi:10.1051/m2an/2013083
  - Kidane & al. "Rigorous model-based uncertainty quantification with application to terminal ballistics. Part I" *J. Mech. Phys. Solids* **60**(5):983–1001, 2011. doi:10.1016/j.jmps.2011.12.001
  - Adams & al. "Rigorous model-based uncertainty quantification with application to terminal ballistics. Part II" *J. Mech. Phys. Solids* **60**(5):1002–1019, 2011. doi:10.1016/j.jmps.2011.12.002
- Software
  - Mystic (optimization): http://trac.mystic.cacr.caltech.edu/project/mystic
  - Pathos (distribution): http://trac.mystic.cacr.caltech.edu/project/pathos