Tuesday 14:00-15:30

Part 4
Examples of imprecise probability in engineering

by Edoardo Patelli

Examples of imprecise probability in engineering

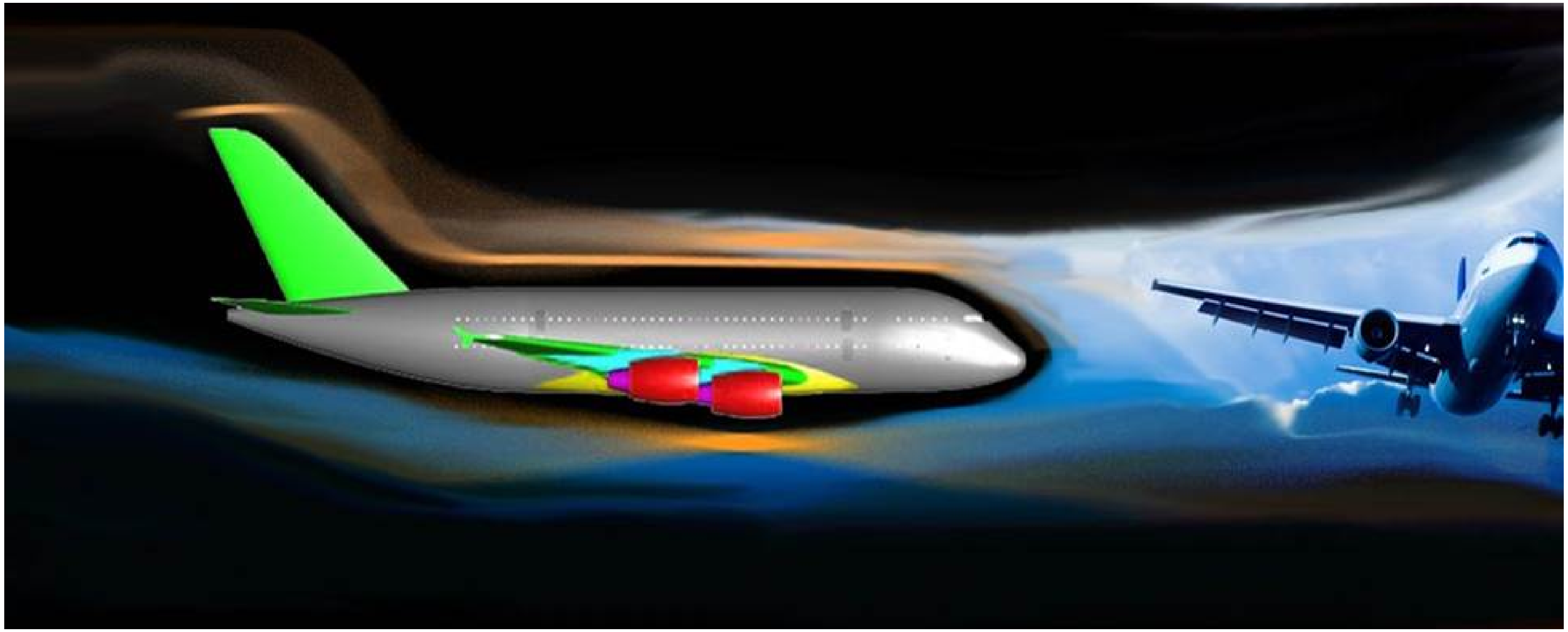
Outline

Outline

Problem Statement

Mathematical and numerical models

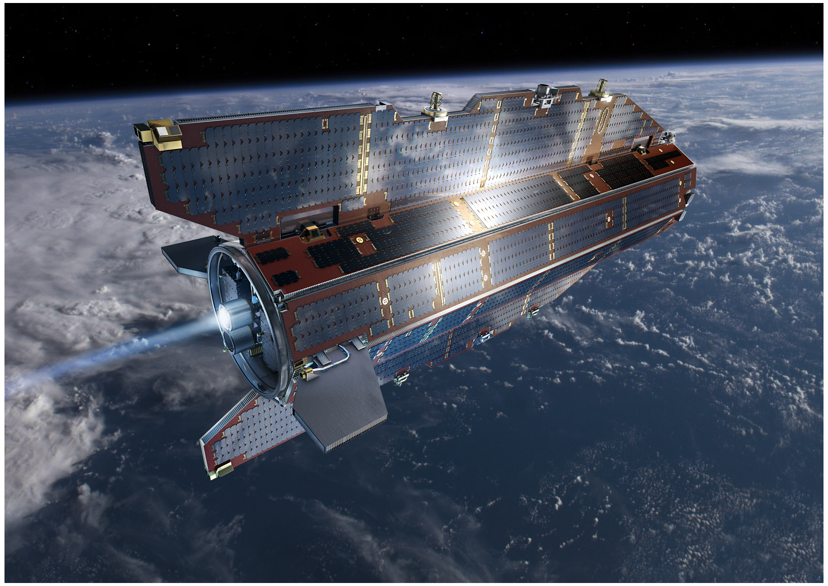
- Adopted to simulate, study and forecast future outcomes
- Models must capture the essentials of the real response.
- Models contain a fairly large set of parameters whose "true" values are not precisely known, i.e. they are uncertain.

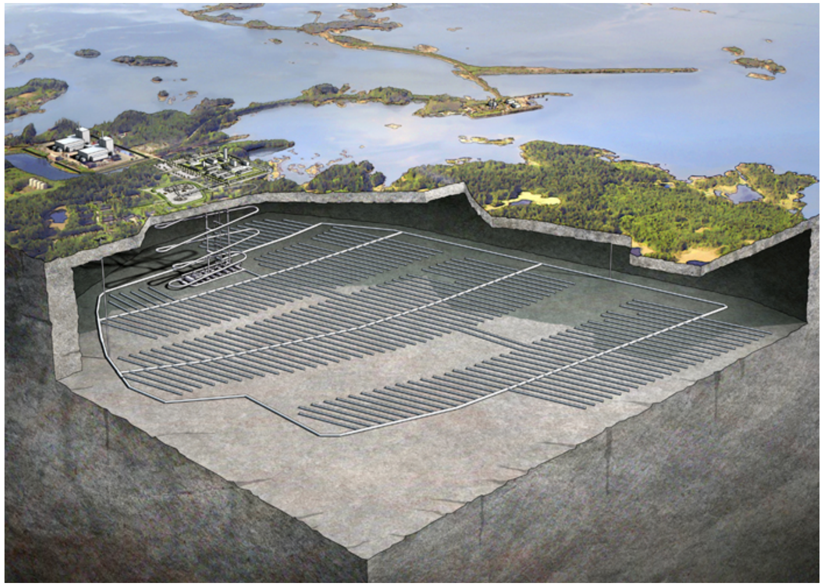


Current challenges in Engineering

Uncertainty quantification and management

- Complex systems often must be designed to operate in harsh domains with a wide array of operating conditions
- Quantitative data is either very sparse or prohibitively expensive to collect

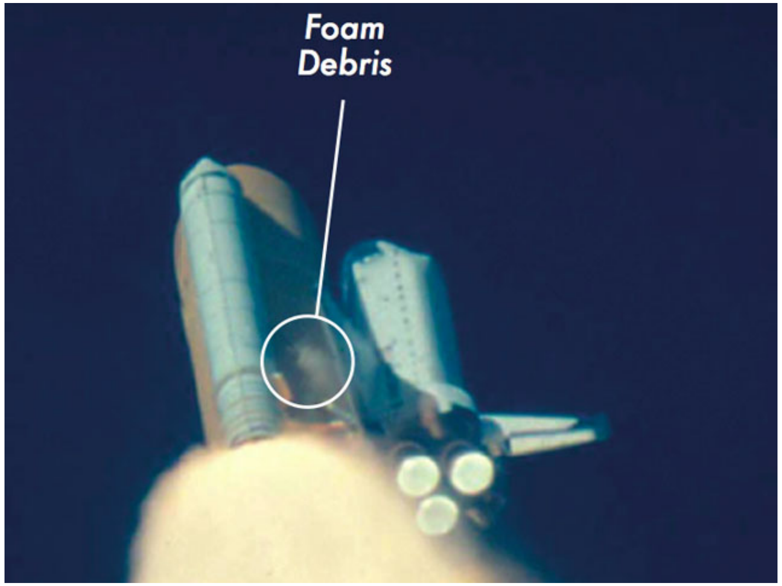




Current challenges in Engineering

High-consequence and safety-critical systems

- Risk is often misestimated
- Models are deterministic without incorporating any measure of uncertainty (Columbia accident report)
- Inadequate assessment of uncertainties, unjustified assumptions (NASA-STD-7009)
- Looking for the "black swan" (e.g. Fukushima)



System failures

Accident triggered by natural events I

- 02 October 2002 - Golf Mexico
- Hurricane Lili -BP Eugene Island 322 Platform A

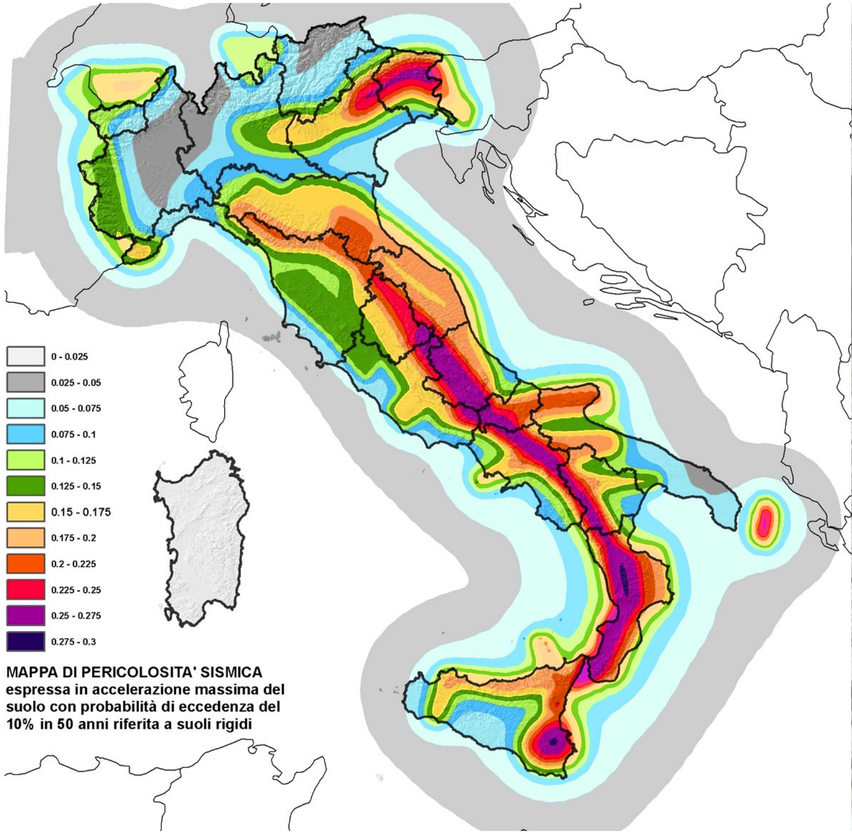
System failures

Accident triggered by natural events II

- Fukushima Daiichi nuclear disaster (Japan) on 11 March 2011
  - 9.0 MW Tohoku earthquake
  - maximum ground accelerations of 0.56 g (design tolerances of 0.45 g)
  - 15 m tsunami arriving 41 minutes later (design seawall of 5.7 m)

System failures

Accident triggered by natural events III



Source: repubblica.it

- Amatrice (ITALY), 24 September 2016, (6.0 magnitude)

System failures

Accident triggered by natural events IV

Source: tt.com

- Landslide (Matrei, Austria, 14^{th} May, 2013)

System failures

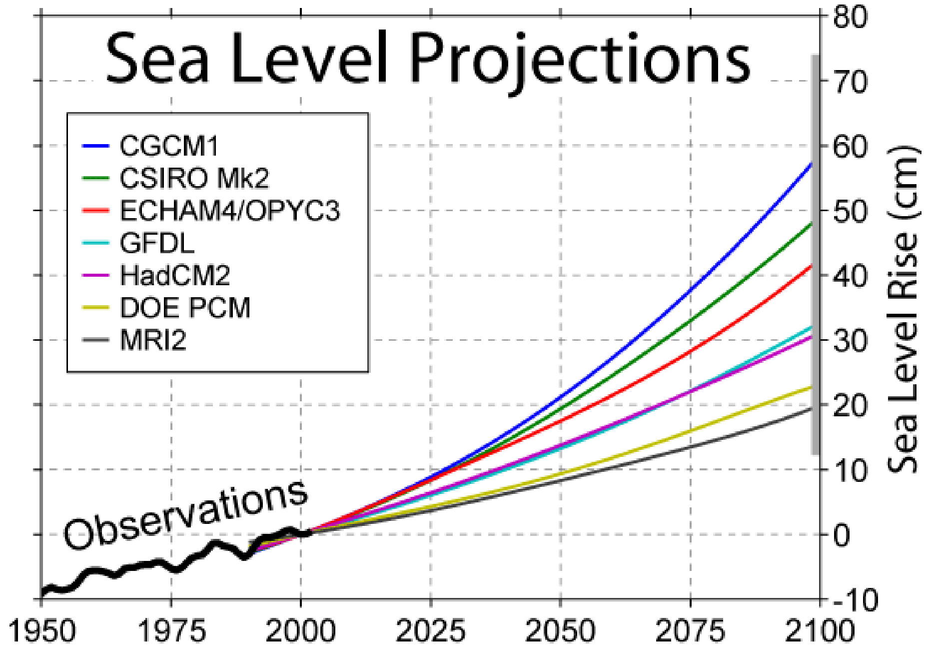
Accident caused by poor maintenance

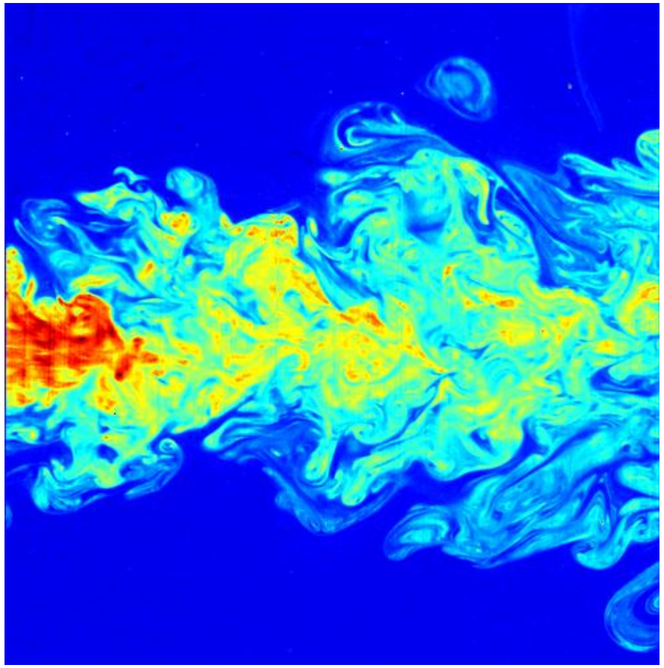
- Southwest B733 near Yuma on Apr 1st 2011
  - Hole in fuselage, sudden decompression
- Viareggio (Italy) on July 1st 2009
  - The failure of an axle on the wagon, derailment and subsequent fire

Unavoidable uncertainties I

Irreducible (aleatory) uncertainties

- Parameters intrinsically uncertain
- Value varies at each experiment
- Future environmental conditions, chaotic and stochastic process

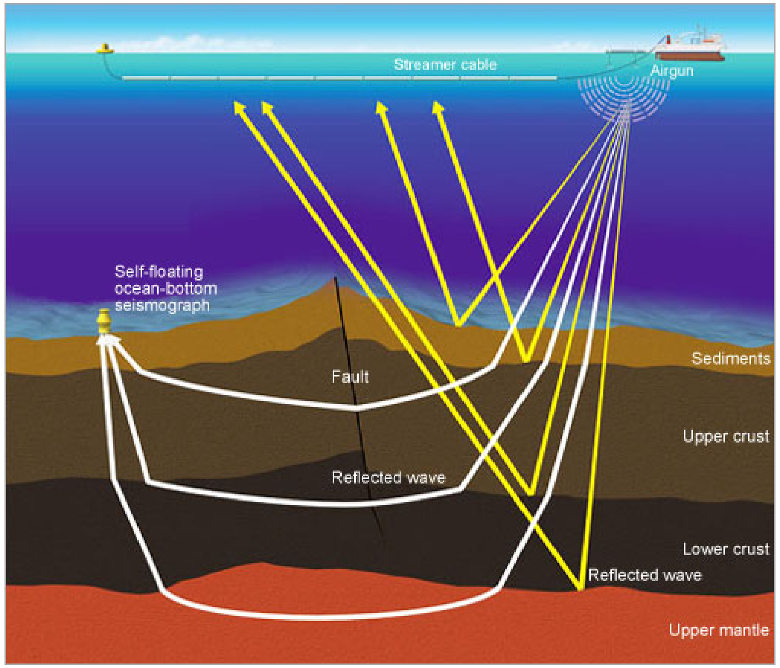


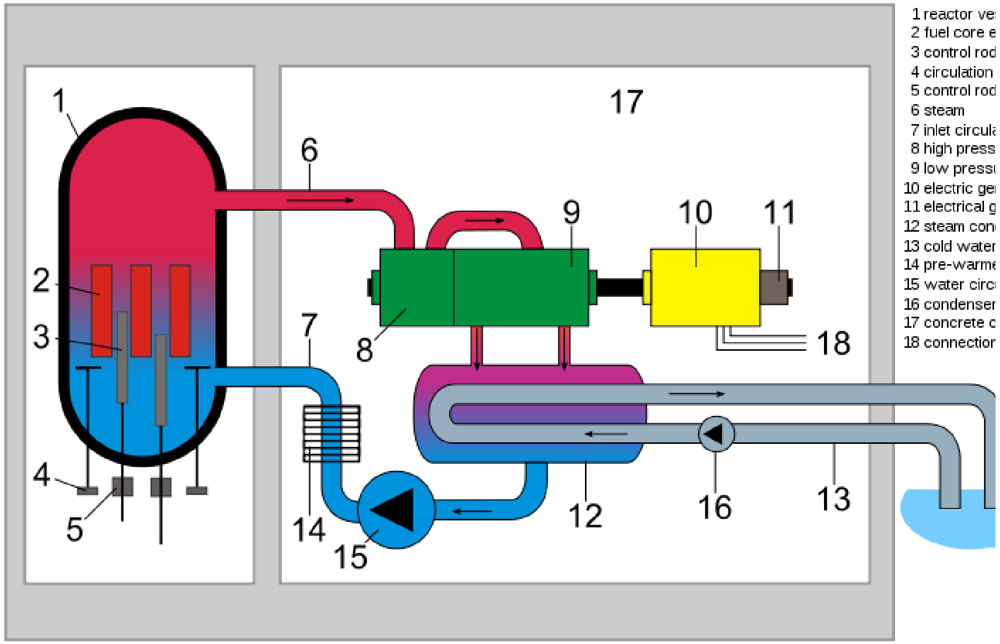


Unavoidable uncertainties II

Reducible (epistemic) uncertainties

- Quantities that could be determined in theory
- Practically they are not measured
- Field properties, simplified model

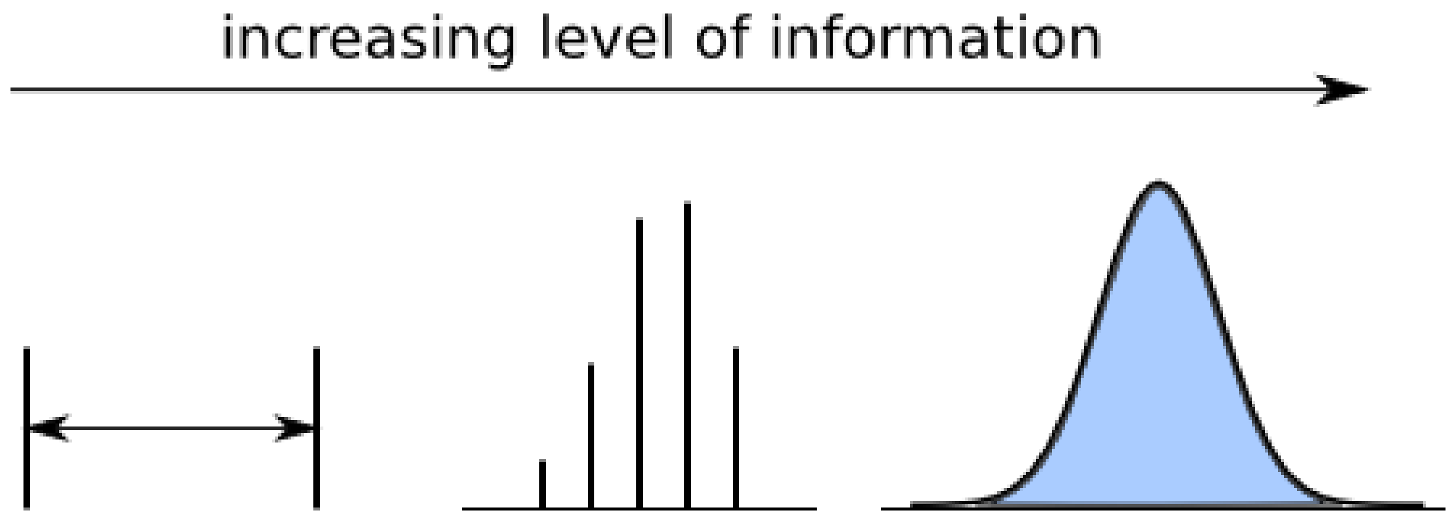




Unavoidable uncertainties III

"Lack of knowledge" and mixed information

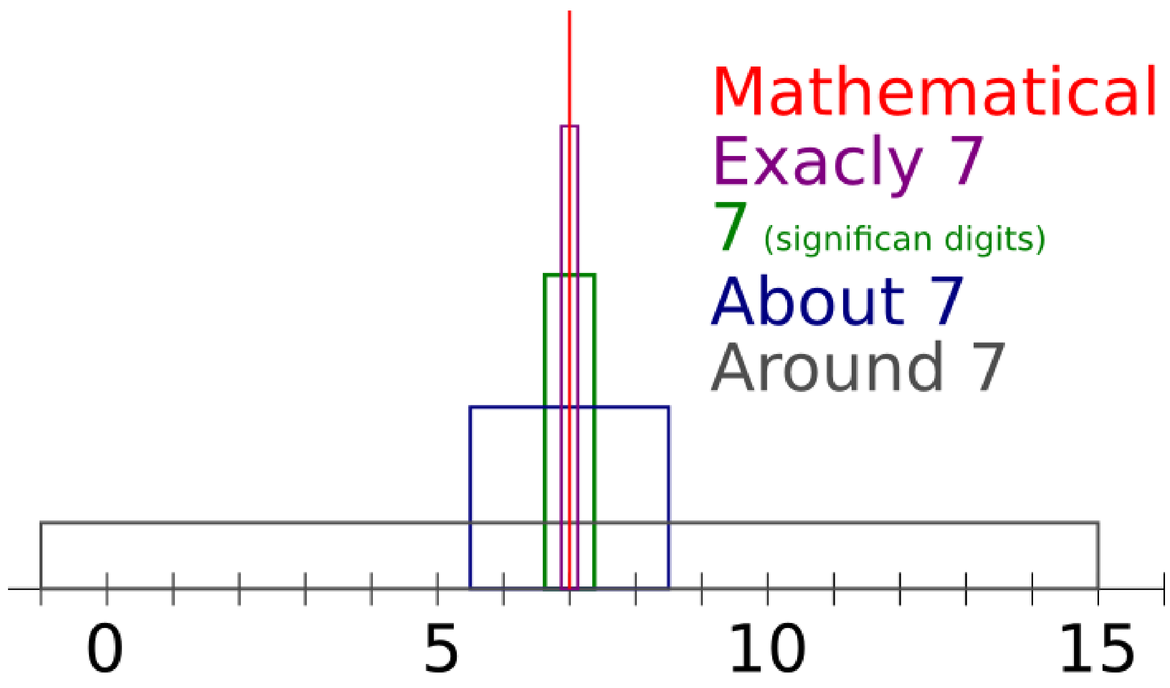
- Statistical information often not available (unique structure)
- Few or missing data (expensive to collect)



Unavoidable uncertainties III

"Lack of knowledge" and mixed information

- Statistical information often not available (unique structure)
- Few or missing data (expensive to collect)
- Qualitative information (expert judgements)
- Heterogeneous information from different sources and in different format

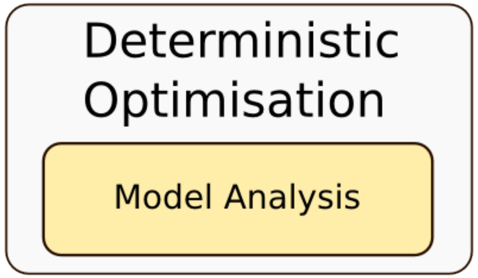


Examples of imprecise probability in engineering

Outline

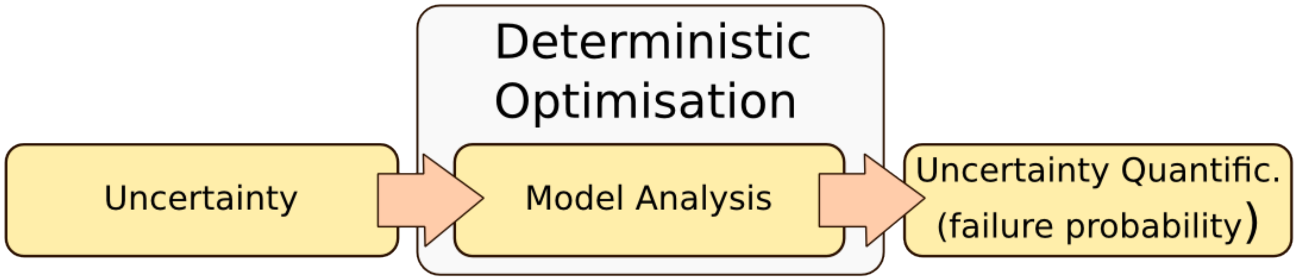
How to deal with uncertainties?

- Effects of uncertainty need to be included at the design stage
- From “Intuitive analysis” to “Quantitative analysis”



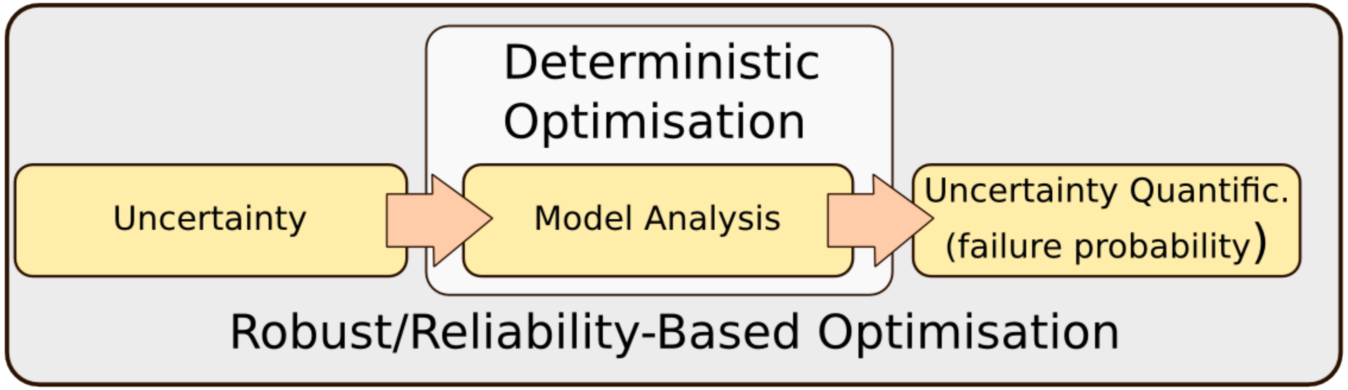
How to deal with uncertainties?

- Effects of uncertainty need to be included at the design stage
- From "Intuitive analysis" to "Quantitative analysis"



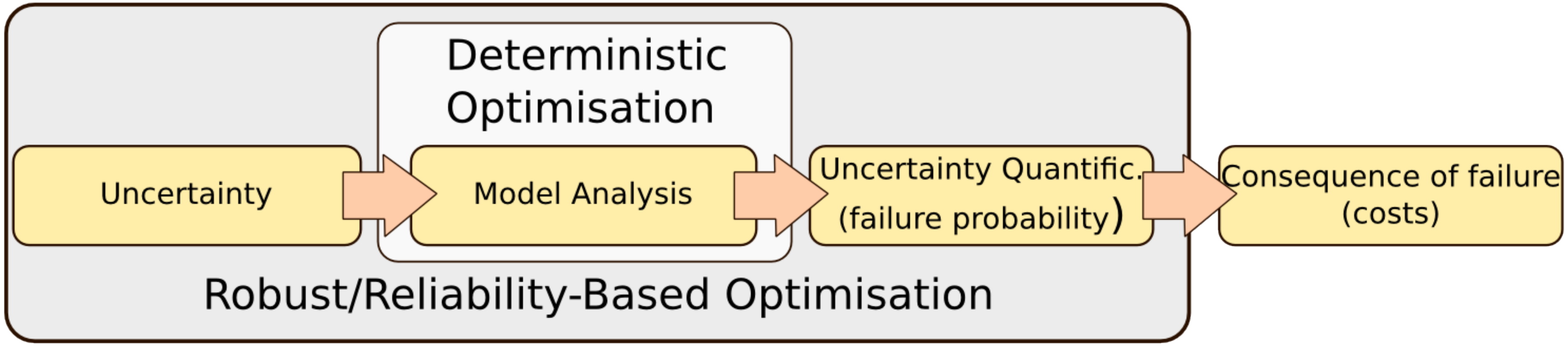
How to deal with uncertainties?

- Effects of uncertainty need to be included at the design stage
- From "Intuitive analysis" to "Quantitative analysis"



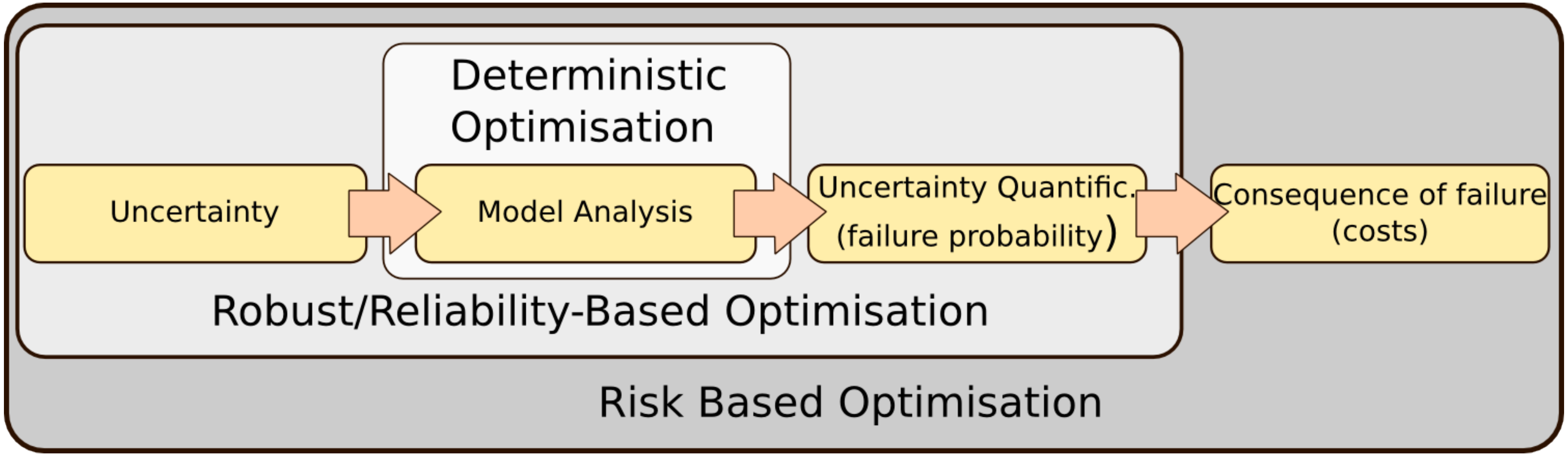
How to deal with uncertainties?

- Effects of uncertainty need to be included at the design stage
- From "Intuitive analysis" to "Quantitative analysis"



How to deal with uncertainties?

- Effects of uncertainty need to be included at the design stage
- From “Intuitive analysis” to “Quantitative analysis”



Dealing with uncertainty: Requirements

Quantifying

- **Modelling and refinement** of uncertainty based on experimental data, simulations and/or expert opinion
- **Propagation** of **uncertainties** through system models
- Parameter ranking and **sensitivity analysis** in the presence of uncertainty

Dealing with uncertainty: Requirements

Quantifying

- **Modelling and refinement** of uncertainty based on experimental data, simulations and/or expert opinion
- **Propagation** of **uncertainties** through system models
- Parameter ranking and **sensitivity analysis** in the presence of uncertainty

Managing

- Identification of the parameters whose uncertainty is the most/least consequential
- Worst-case system **performance assessment**
- **Design** in the presence of uncertainty

Examples of imprecise probability in engineering

Outline

Traditional modelling of uncertainty

Probability: A real number $\mathbb{P}(A)$ assigned to every event $A \subset \Omega$

Traditional modelling of uncertainty

Probability: A real number $\mathbb{P}(A)$ assigned to every event $A \subset \Omega$

$\mathbb{P}$ must satisfy some natural properties, called axioms of probability:

1. $\mathbb{P}(A) \geq 0$ (Nonnegativity)

Traditional modelling of uncertainty

Probability: A real number $\mathbb{P}(A)$ assigned to every event $A \subset \Omega$

$\mathbb{P}$ must satisfy some natural properties, called axioms of probability:

1. $\mathbb{P}(A) \geq 0$ (Nonnegativity)
2. $\mathbb{P}(\Omega) = 1$ (Normalisation)

Traditional modelling of uncertainty

Probability: A real number $\mathbb{P}(A)$ assigned to every event $A \subset \Omega$

$\mathbb{P}$ must satisfy some natural properties, called axioms of probability:

1. $\mathbb{P}(A) \geq 0$ (Nonnegativity)
2. $\mathbb{P}(\Omega) = 1$ (Normalisation)
3. If $A \cap B = \varnothing$ then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ (Additivity)

Subjective probability

Bayesian approach

Combination of rare data and prior expert knowledge

$$\mathbb{P}(\boldsymbol{\theta}|\boldsymbol{x}) = \frac{\mathbb{P}(\boldsymbol{x}|\boldsymbol{\theta})\mathbb{P}(\theta)}{\mathbb{P}(x)} = \frac{\mathbb{P}(\boldsymbol{x}|\boldsymbol{\theta})\mathbb{P}(\boldsymbol{\theta})}{\sum_{i=1}^{n} \mathbb{P}(\boldsymbol{x}|\theta_i)\mathbb{P}(\theta_i)}$$

- Update of expert knowledge by means of data $\mathbb{P}(\boldsymbol{\theta}|\boldsymbol{x})$
- Influence of subjective assessment (prior distribution,$\mathbb{P}(\boldsymbol{\theta})$) decays quickly with increase of sample size

Statistical analysis of imprecise and rare data

Open questions?

- Evaluation of remaining subjective uncertainty
- Consideration of imprecision (data and expert knowledge)
- Very small sample size or no data

Statistical analysis of imprecise and rare data

Open questions?

- Evaluation of remaining subjective uncertainty
- Consideration of imprecision (data and expert knowledge)
- Very small sample size or no data

Statistical analysis of imprecise and rare data

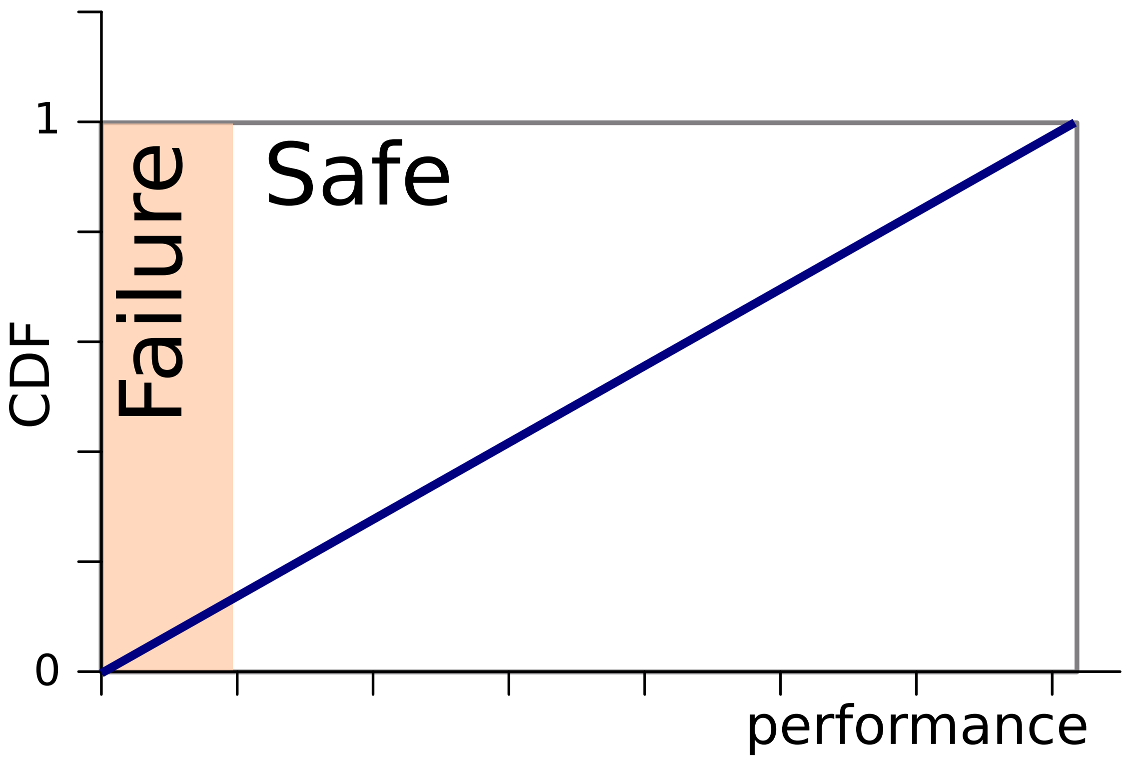
Open questions?

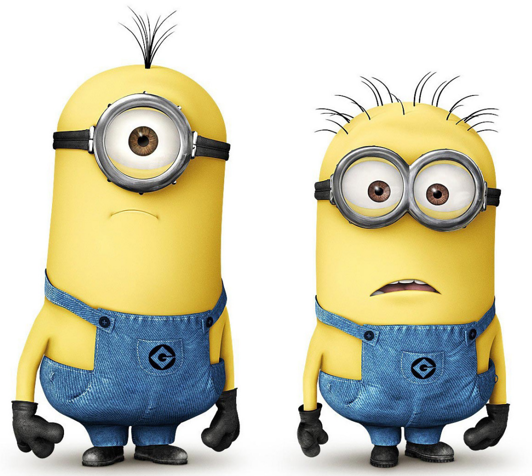
- Evaluation of remaining subjective uncertainty
- Consideration of imprecision (data and expert knowledge)
- Very small sample size or no data

Naive approach

Model epistemic uncertainty as aleatory uncertainty

- Information available as interval only (i.e. bounds) $\rightarrow$ **uniform distribution**

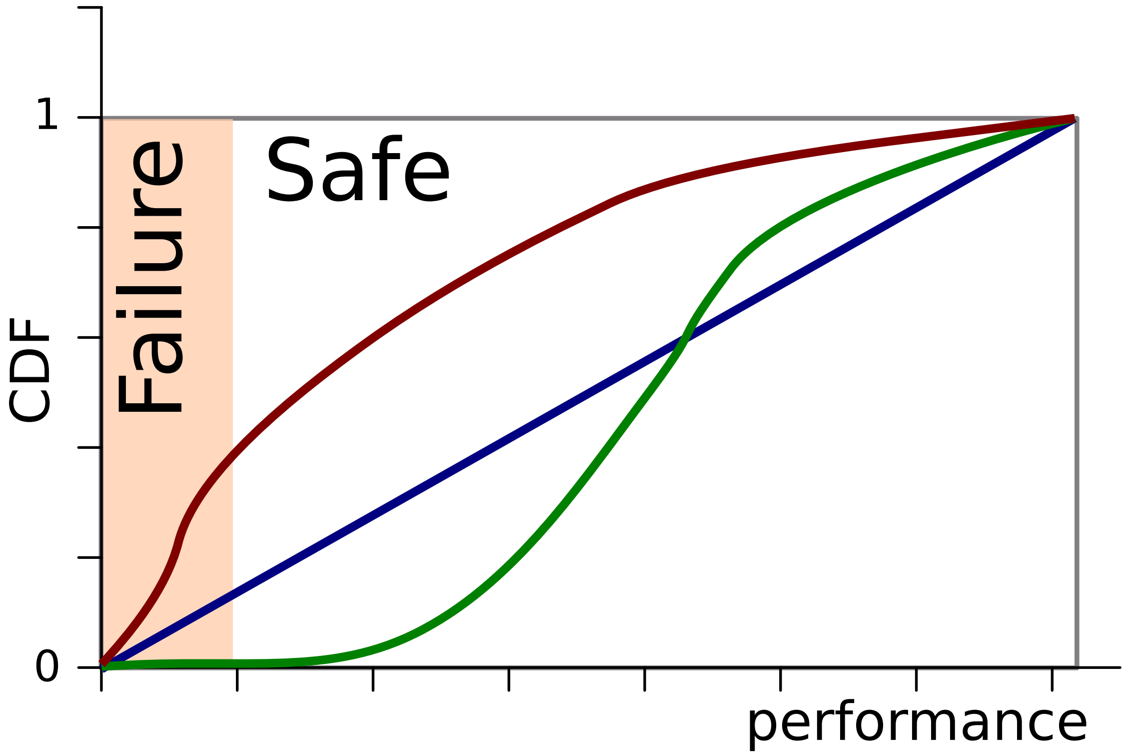


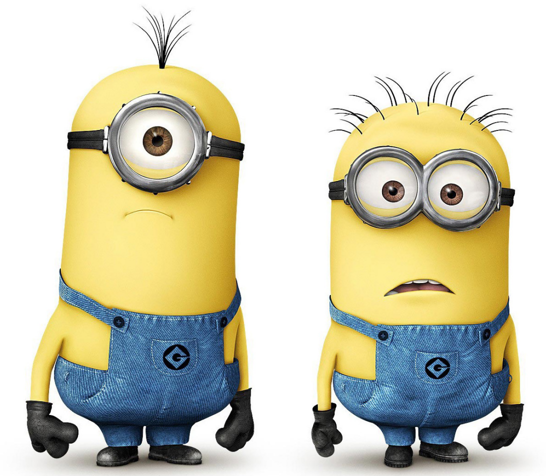


Naive approach

Model epistemic uncertainty as aleatory uncertainty

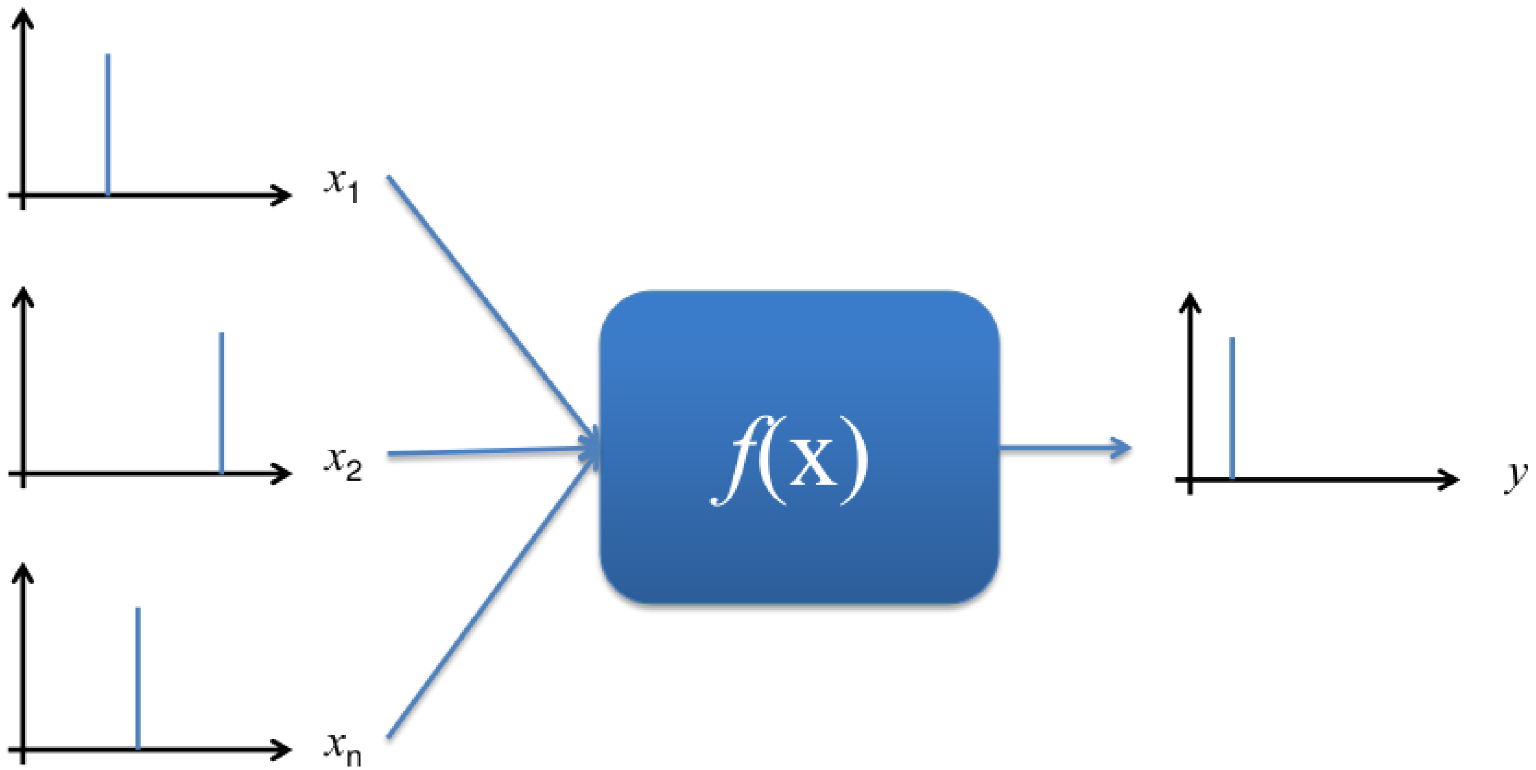
- Information available as interval only (i.e. bounds) $\rightarrow$ **uniform distribution**
- It is a BIG and UNJUSTIFIED (wrong) assumption!
- Give a false sense of confidence





Interval analysis
Epistemic Uncertainty

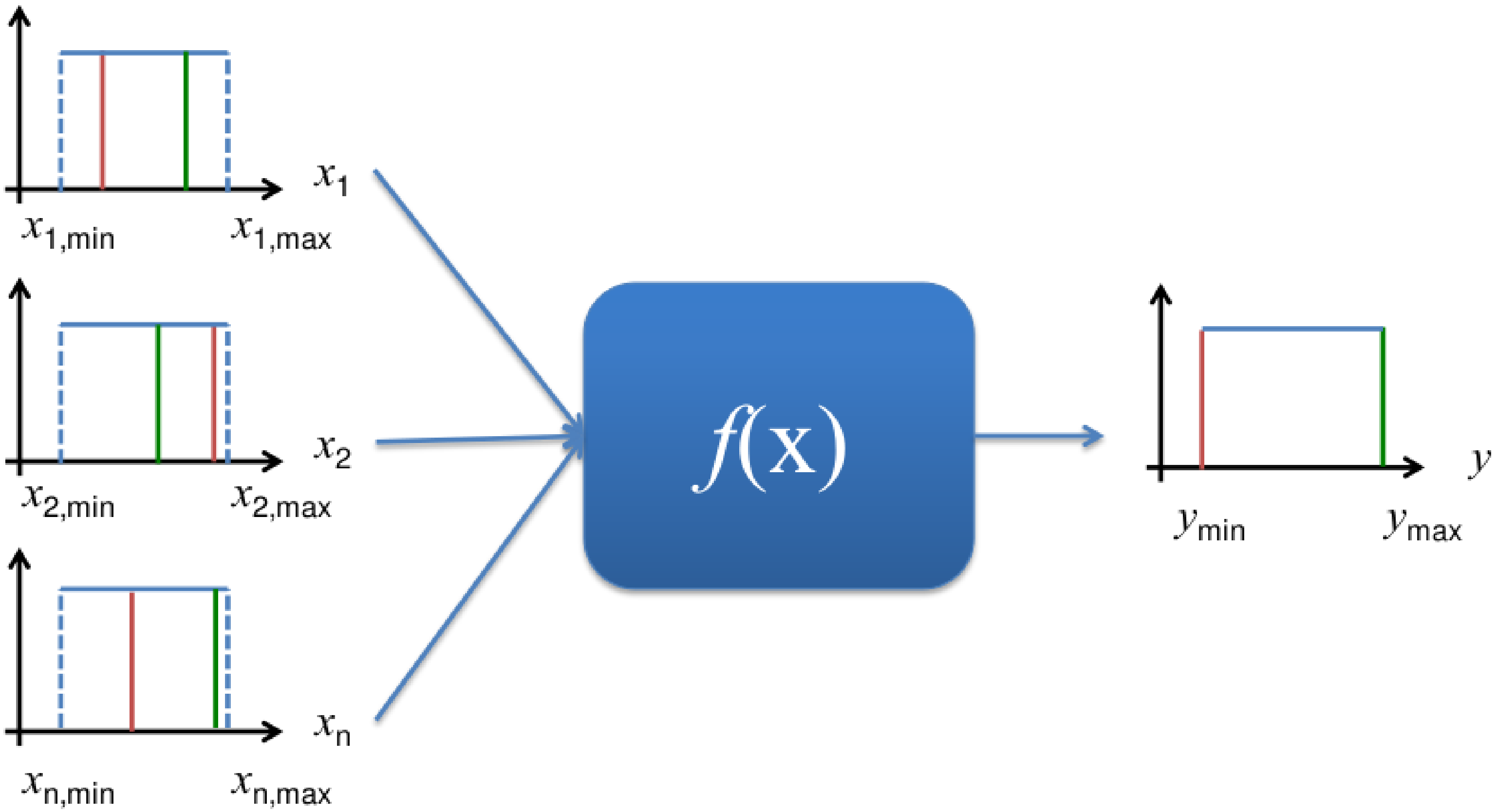
- Intervals as epistemic uncertainty of parametric values



Interval analysis

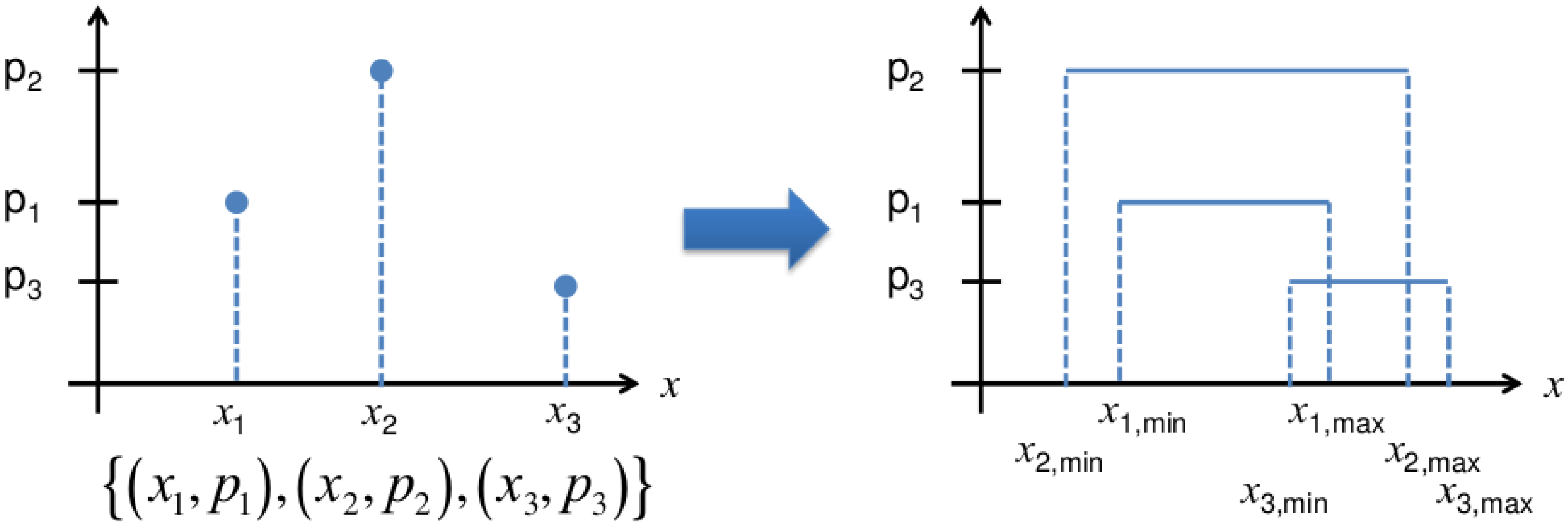
Epistemic Uncertainty

- Intervals as epistemic uncertainty of parametric values
- Uncertainty propagation is an optimization problem



Dempster-Shafer structure
Collection of intervals

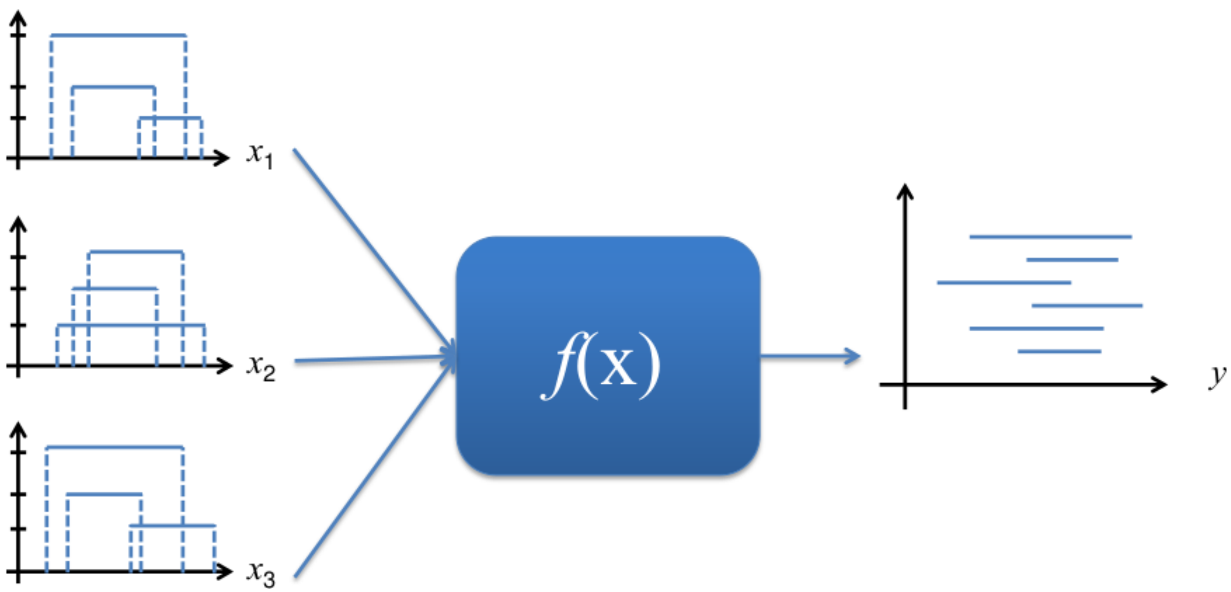
- Different probability masses associated to distinct intervals (e.g. weighting expert options or assumptions)



Dempster-Shafer structure

Collection of intervals

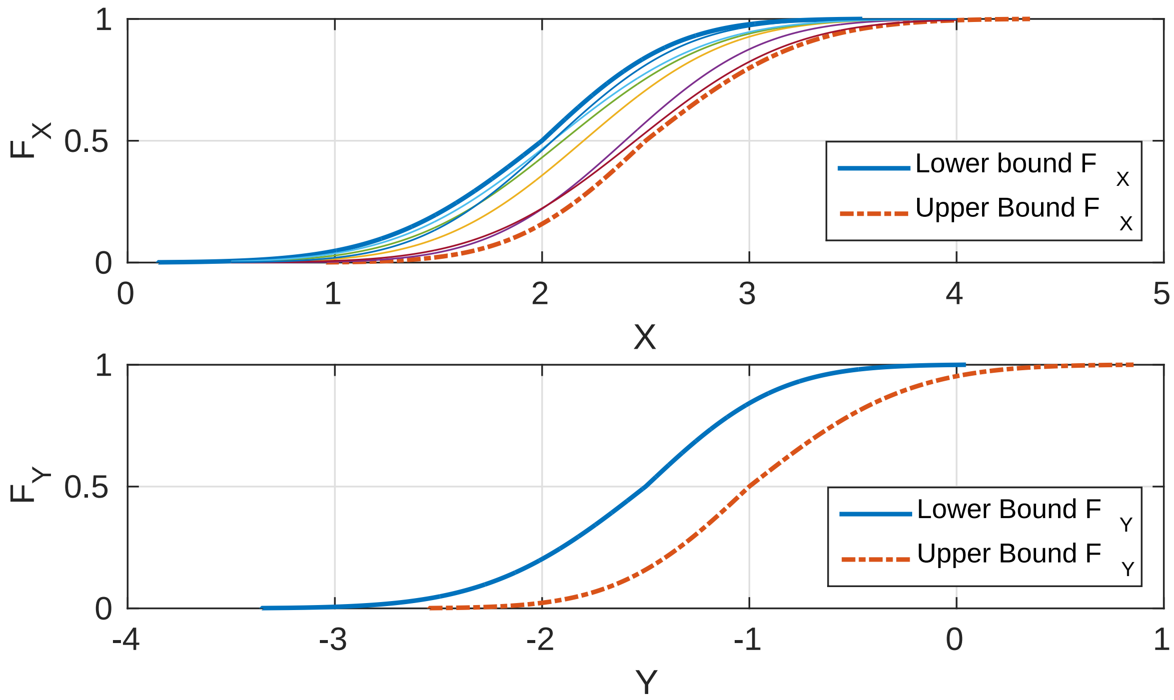
- Different probability masses associated to distinct intervals (e.g. weighting expert options or assumptions)
- UQ by sampling intervals (solving an optimization problem for each sample)



Probability box

Epistemic + Aleatory uncertainty

- Distributional probability box
- Distributional-free probability box



References

- S. Ferson, V.Kreinovich, L.Ginzburg, D.S.Myers, & K.Sentz, Constructing probability boxes and Dempster-Shafer structures Sandia National Laboratories, 2003

Definition of a random set

Succinct introduction

- Random set Γ is like a random variable whose realizations γ are **sets** in $\mathscr{F}$, not numbers
- $\mathscr{F}$ is a focal set
- $\gamma := \tilde{\Gamma}(\alpha) \in \mathscr{F}$ is a focal element

When all focal elements of $\mathscr{F}$ are singletons, Γ becomes a random variable X

References

- Alvarez, D. A., On the calculation of the bounds of probability of events using infinite random sets, *International Journal of Approximate Reasoning*, Vol. 43, 2006, pp. 241–267.
- Alvarez, D. A., *Infinite random sets and applications in uncertainty analysis*, Ph.D. thesis, Unit of applied mathematics, University of Innsbruck, 2007
- Alvarez, D. A., A Monte Carlo-based method for the estimation of lower and upper probabilities of events using infinite random sets of indexable type, *Fuzzy Sets and Systems*, Vol. 160, 2009, pp. 384–401.

Using intervals and d-dimensional boxes as elements of $\mathscr{F}$, is enough to model (without making any supposition at all):

- Possibility distributions (also known as normalized fuzzy sets)
- Cumulative distribution functions (CDFs)
- Intervals and families of intervals (Dempster-Shafer structures)
- Distribution-free and distributional probability boxes (p-boxes)

and their joint combinations (using Copulas).

Definition

A copula is a multivariate CDF $C : [0,1]^d \to [0,1]$ such that each of its marginal CDFs is uniform on the interval $[0,1]$.

Relationship between Random Sets, CDFs and finite families of intervals

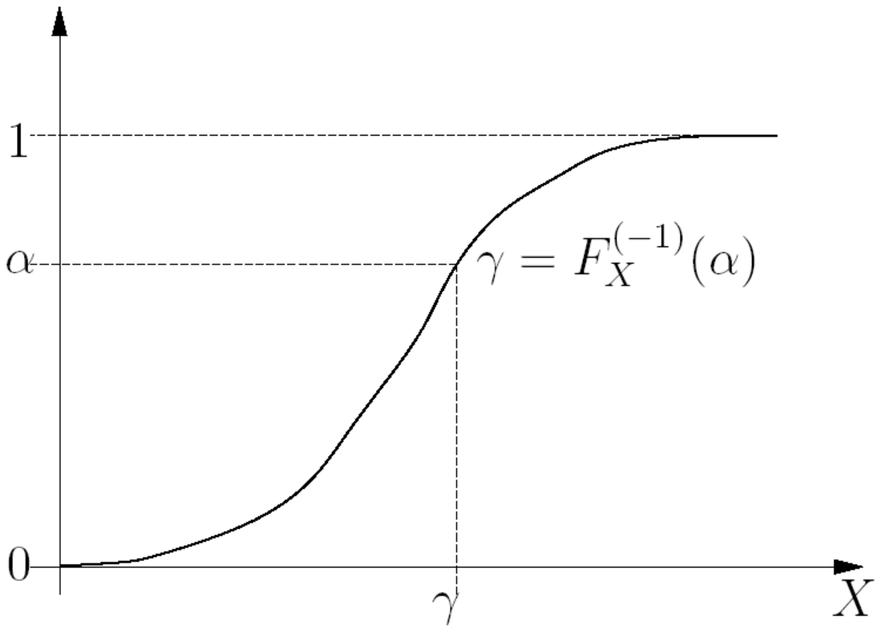


Figure 1: Focal element of a CDF

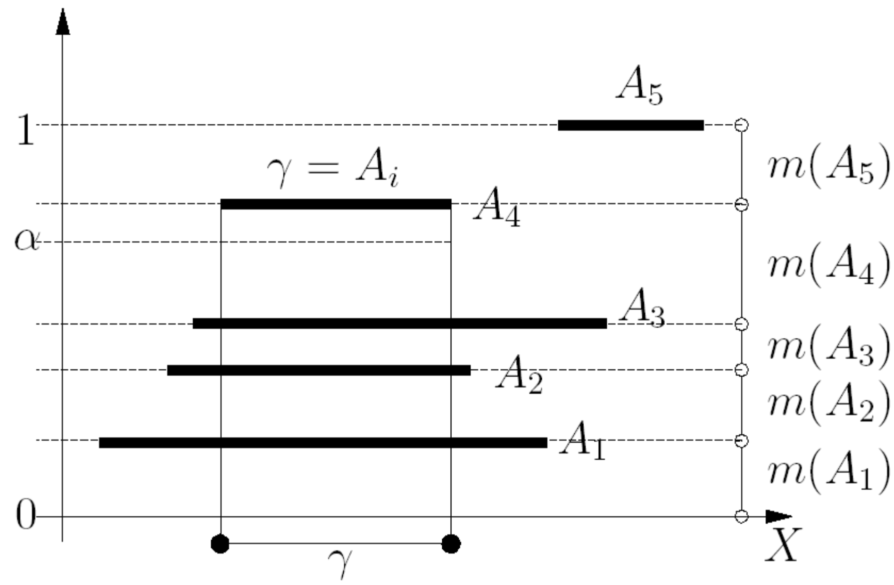


Figure 2: Focal element of a family of intervals

Relationship between Random Sets, p-boxes and fuzzy set

Obtained by generating an α from a uniform distribution on $(0, 1]$ and then, obtaining the corresponding focal element $\gamma = \tilde{\Gamma}(\alpha)$.

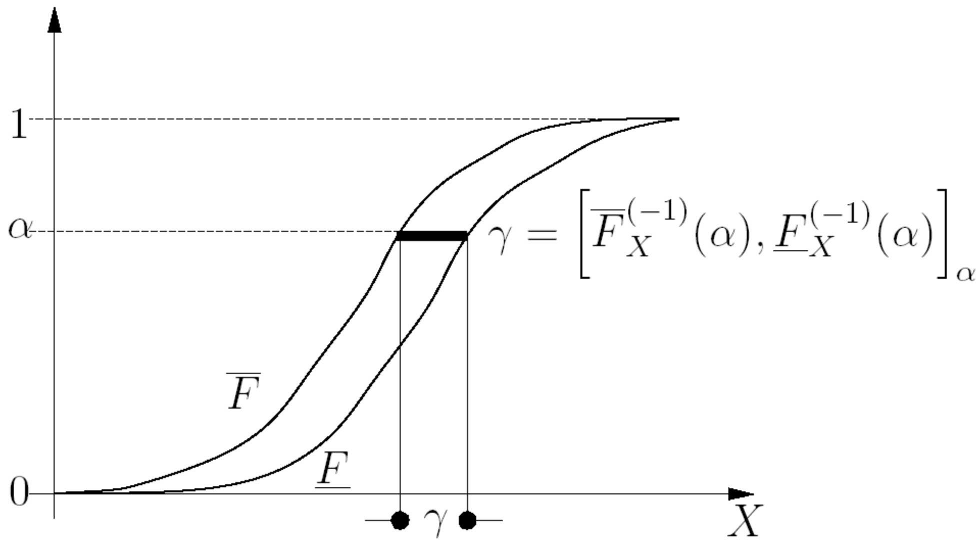


Figure 3: Distribution-free probability box

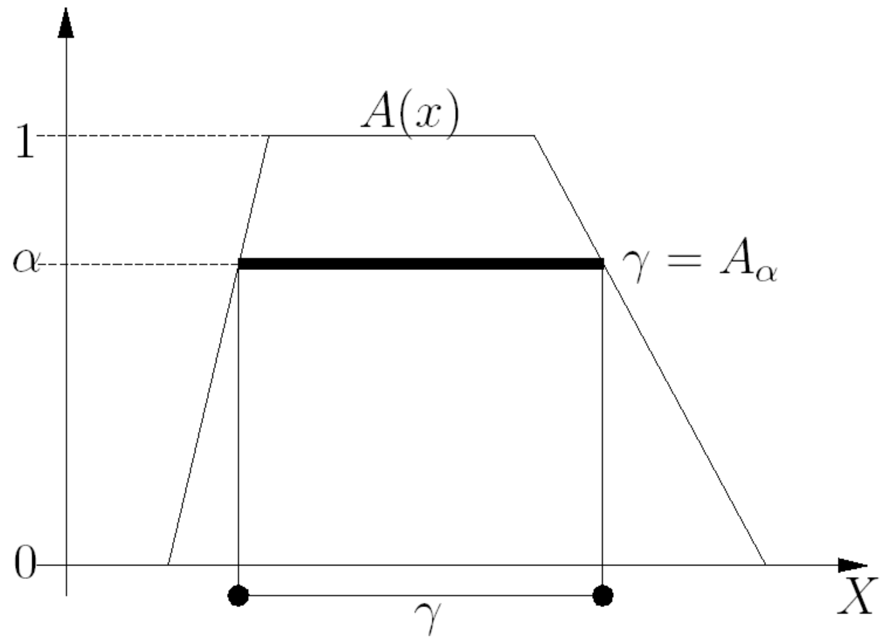


Figure 4: Possibility distribution

Relationship between Random Sets and distributional probability boxes

Distributional p-boxes do not have a graphical representation. Consider a CDF with m uncertain and independent parameters, (e.g. intervals I_i for $i = 1, 2, \ldots, m$) using the random set representation, a focal element of the probability box can be represented as the input intervals $\{I_i : i = 1, 2, \ldots, m\}$ together with the sample of α which is a uniform random variable on $(0, 1] \equiv \Omega$.
Represented as the random set $\tilde{\Gamma} : \Omega \to \mathscr{F}, \alpha \mapsto \Gamma(\alpha)$ (i.e. $(\mathscr{F}, P_\Gamma)$) defined on $\mathscr{R}$ where $\mathscr{F}$ is the system of focal elements $\{\alpha \times I_1 \times \cdots \times I_m : \alpha \in \Omega\}$.

Uncertainty propagation of a random set

Obtained by generating an α from a uniform distribution on $(0, 1]$ and then, obtaining the corresponding focal element $\gamma = \tilde{\Gamma}(\alpha)$.

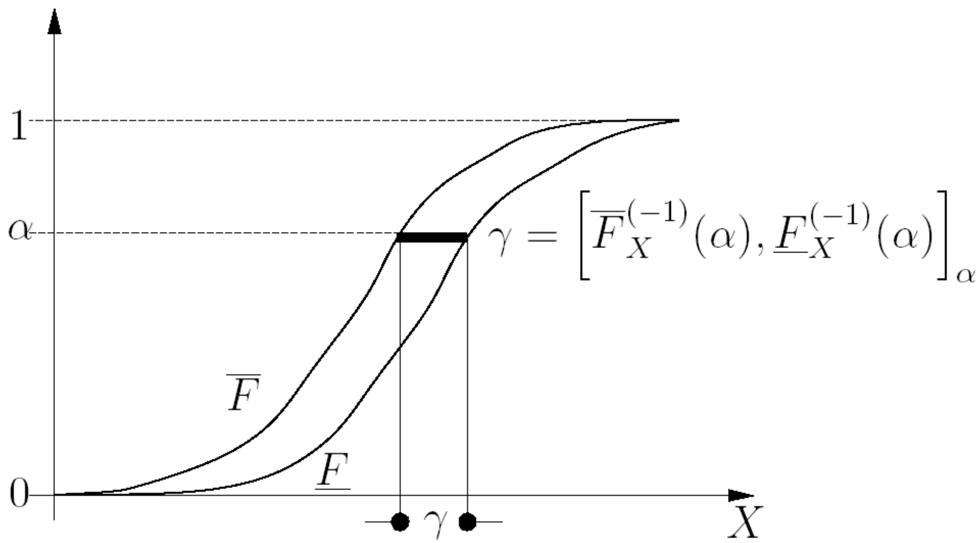


Figure 5: Distribution-free p-box

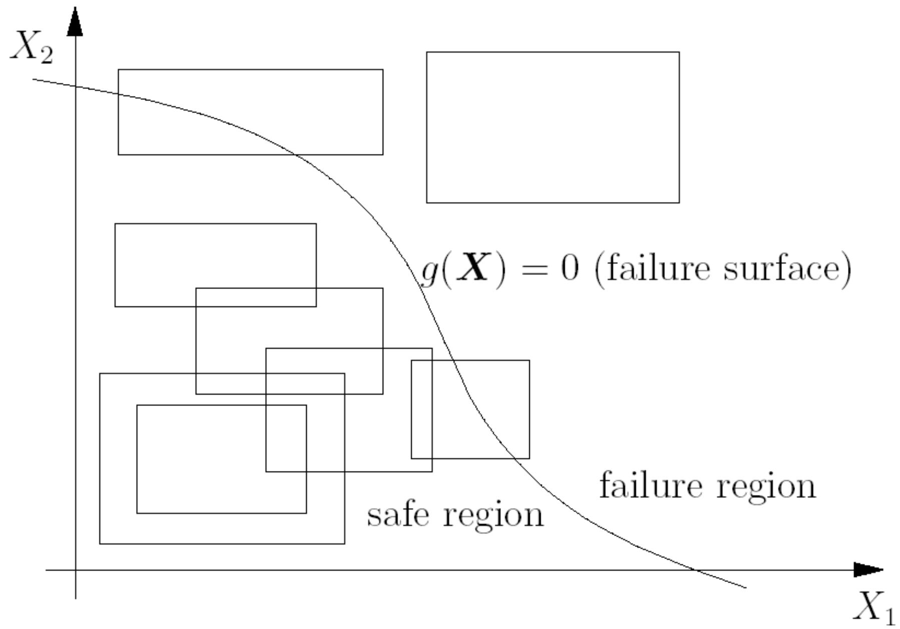


Figure 6: A focal set as a d-dimensional box $\boldsymbol{\gamma} = \times_{i=1}^{d} \gamma_i$ in X

Focal set

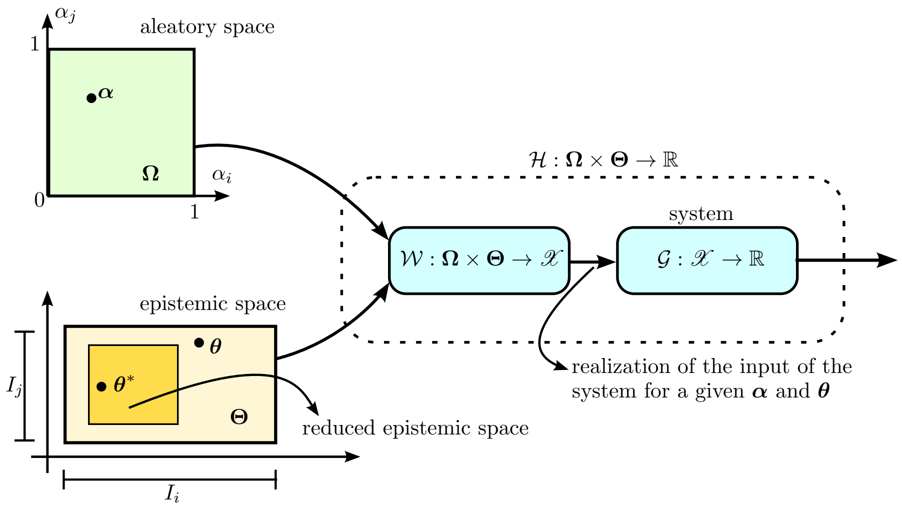
Lower and upper probability measures

The probability measure of a set $\mathscr{F} \in \mathscr{P}(\mathscr{X})$ is bounded

$$\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) \leq P(F) \leq \mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F)$$

where,

$$\mathrm{UP}_{(\mathscr{F},P_\Gamma)}(F) = P_\Gamma\left\{\gamma : \gamma \cap F \neq \emptyset\right\}$$
$$\mathrm{LP}_{(\mathscr{F},P_\Gamma)}(F) = P_\Gamma\left\{\gamma : \gamma \subseteq F\right\}$$



Combination of focal set

Representation either as a d-dimensional box in X or as a point in $(0,1]^d$

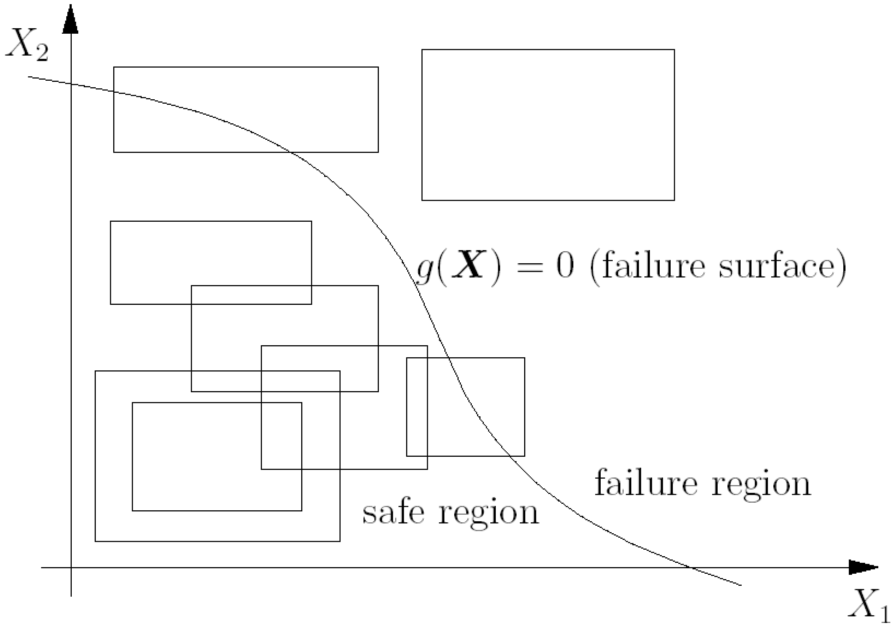


Figure 7: A focal set as a d-dimensional box $\boldsymbol{\gamma} = \times_{i=1}^{d} \gamma_i$ in X

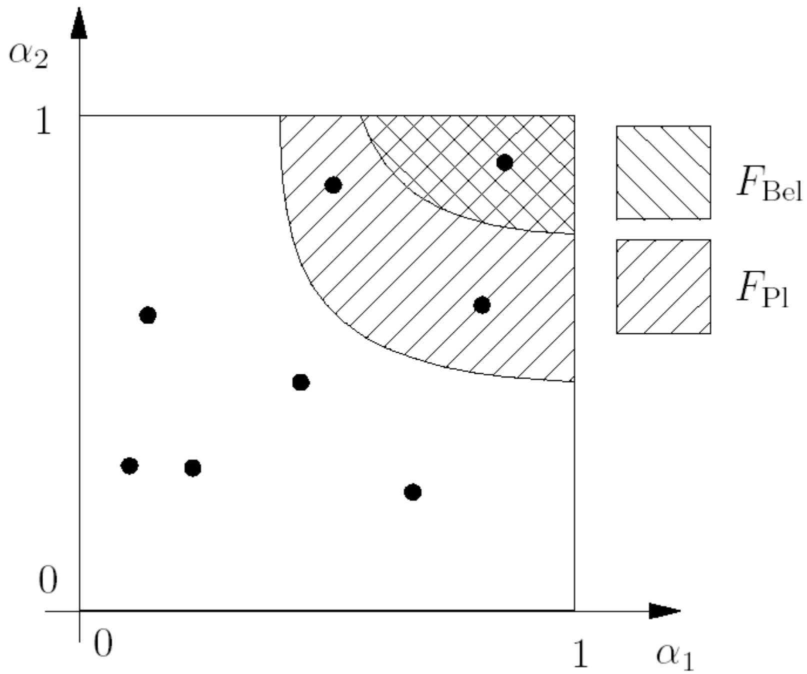
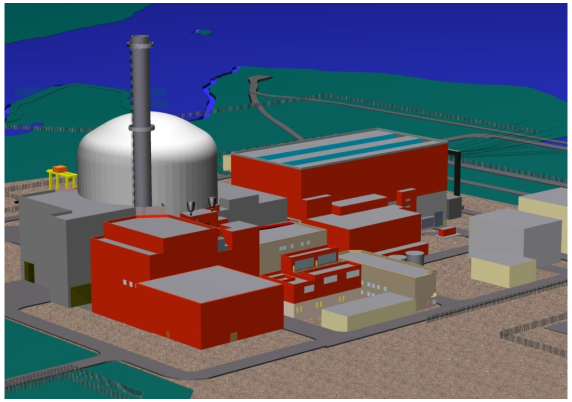
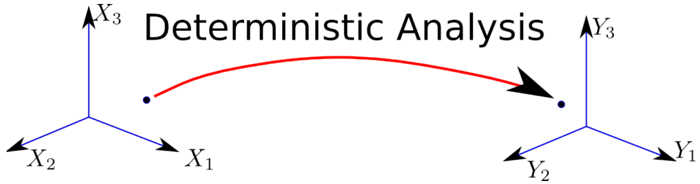


Figure 8: A focal set as a point $\boldsymbol{\alpha} := [\alpha_1, \alpha_2, \ldots, \alpha_n]$ in $(0,1]^d$

Challenges

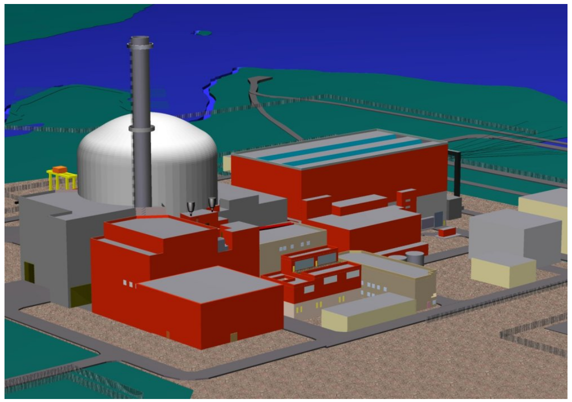
Computational Aspects

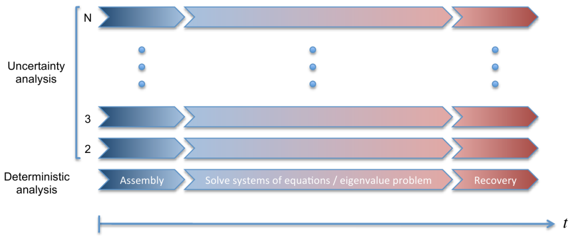


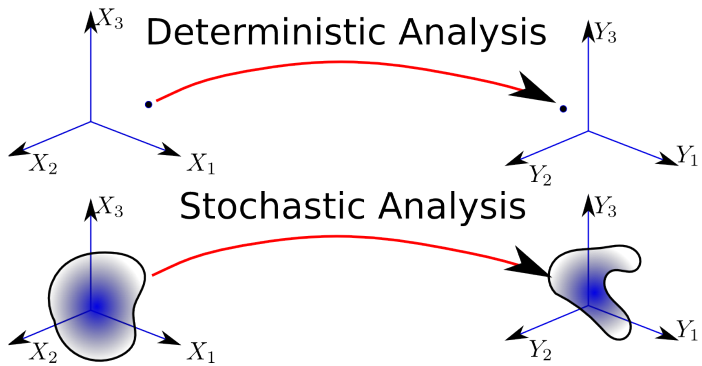


Challenges

Computational Aspects

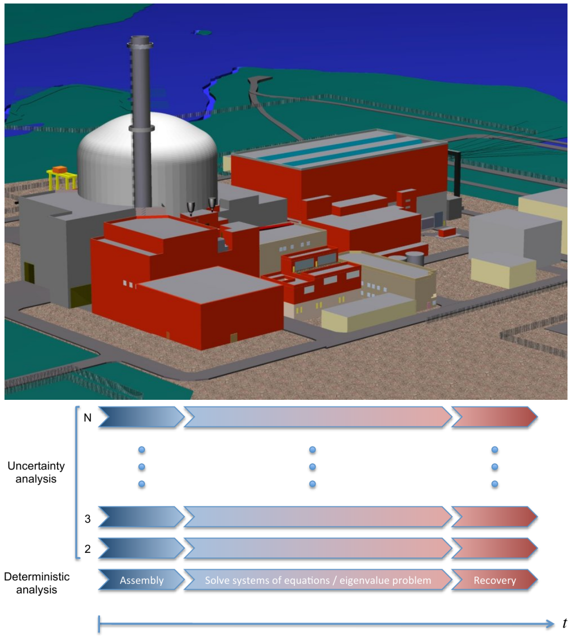


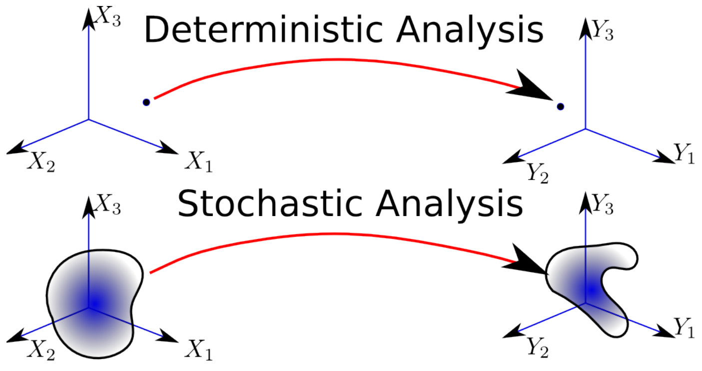


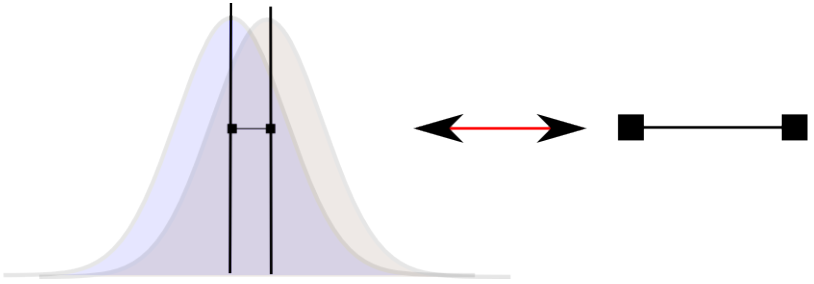


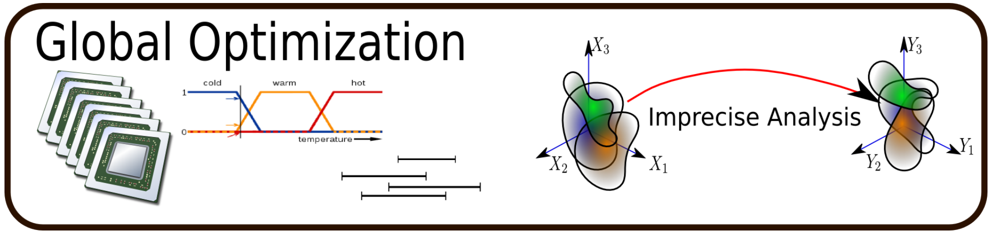
Challenges

Computational Aspects









Needs for improved methods for quantifying and managing uncertainty

Outline

NASA Langley UQ challenge problem

Motivations and Timeline

Aim

- Determine limitations and ranges of applicability of existing UQ methodologies
- Develop **new discipline-independent** UQ methods
- Advance the state of the practice in UQ problems of direct interest to NASA

Timeline

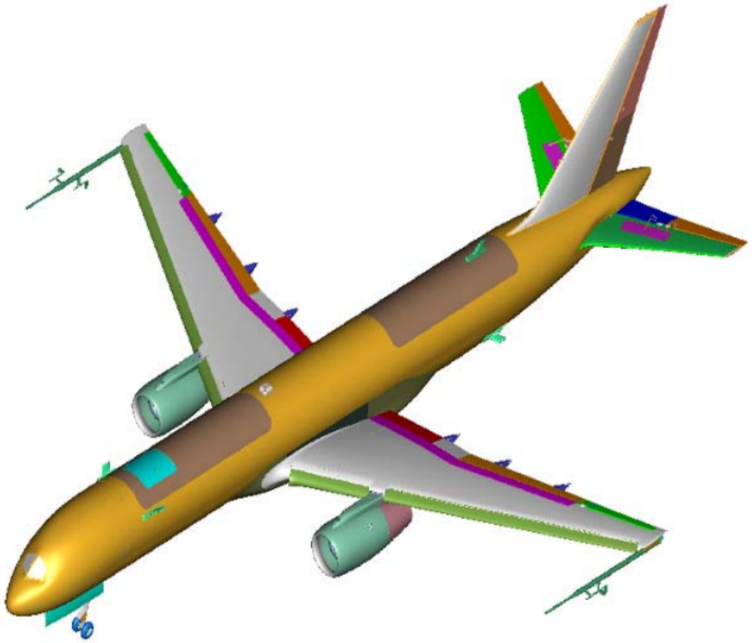
- January 2013: 100+ UQ experts were invited to participate
- January 2014: 11 groups presented at Scitech 2014
- January 2015: Special edition AIAA Journal of Aerospace Information Systems

Outline

Physical System

NASA Langley Generic Transport Model (GTM)

- 5.5% dynamically scaled, remotely piloted, twin-turbine, research aircraft
- Aircraft motion outside the normal flight envelope
- Dynamics are driven by **nonlinearities** and **coupling**, having **oscillatory** and **divergent** behaviour



Generic Transport Model (GTM)

Dynamically scaled, highly instrumented, flight test article

Generic Transport Model (GTM)

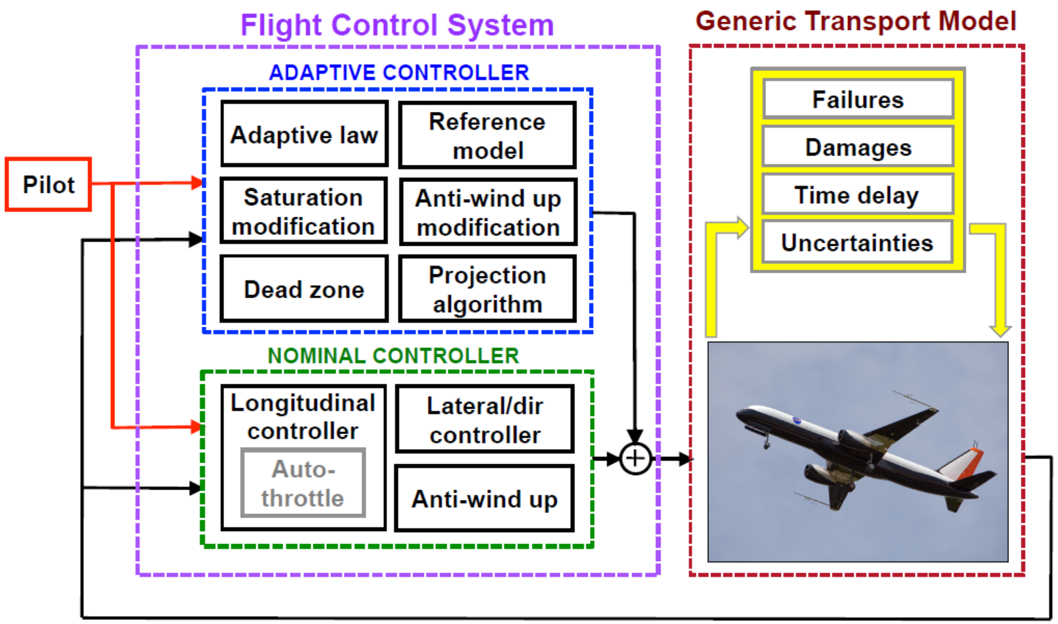
Dynamically scaled, highly instrumented, flight test article

Generic Transport Model (GTM)
AIRSTAR, pilot commands

The mathematical Model

Simulation model

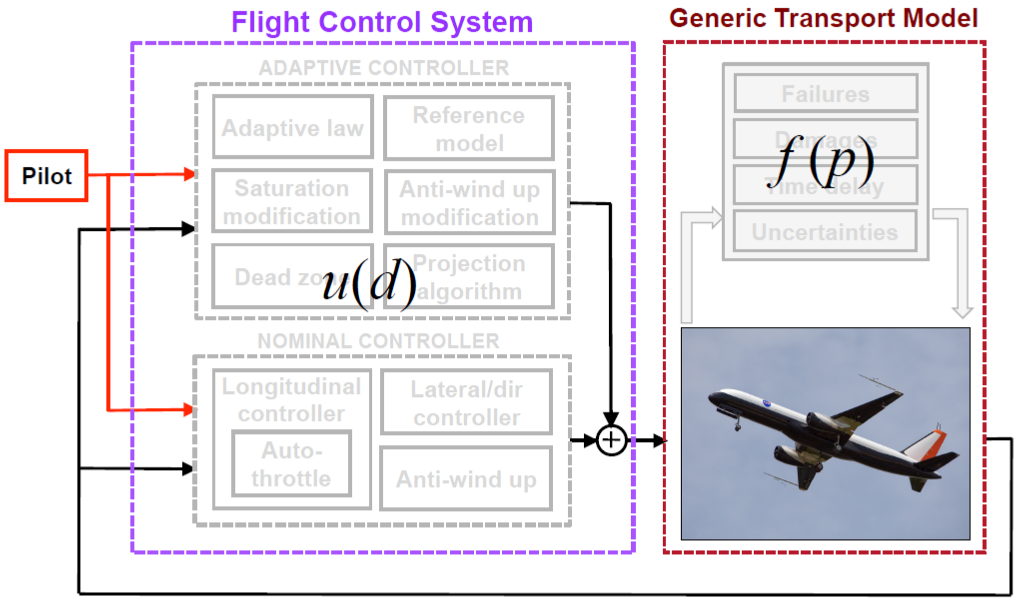
- A high-fidelity aerodynamic mathematical model
- Match closely the dynamics of the test (based on system identification experiments and wind tunnel data



The mathematical Model

Black-box model

- **Specialized aircraft knowledge is not required**
- 5 tasks: Uncertainty Characterization, Sensitivity analysis, Uncertainty Quantification, Extreme Case analysis and Robust Design



The mathematical Model

Black-box model

- 21 Uncertain Parameters $\boldsymbol{p}$: loss in control effectiveness, actuator failure, icing, deadzone, and desired range in operating conditions

The mathematical Model

Black-box model

- 21 Uncertain Parameters $\boldsymbol{p}$: loss in control effectiveness, actuator failure, icing, deadzone, and desired range in operating conditions
- 5 Intermediate variables $\boldsymbol{x}$ (e.g. control effectiveness of elevator, time delay due telemetry and communications)

The mathematical Model

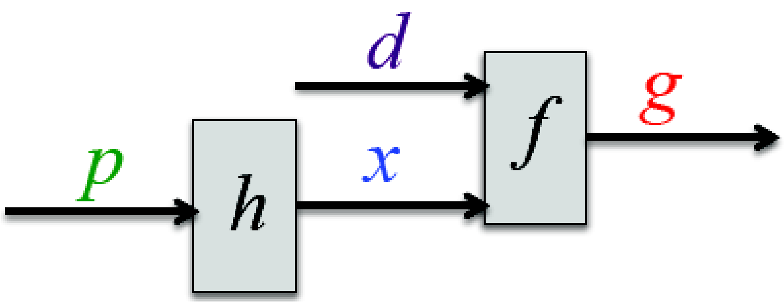
Black-box model

- 21 Uncertain Parameters $\boldsymbol{p}$: loss in control effectiveness, actuator failure, icing, deadzone, and desired range in operating conditions
- 5 Intermediate variables $\boldsymbol{x}$ (e.g. control effectiveness of elevator, time delay due telemetry and communications)
- 14 Design variables $\boldsymbol{d}$: controller gains

The mathematical Model

Black-box model

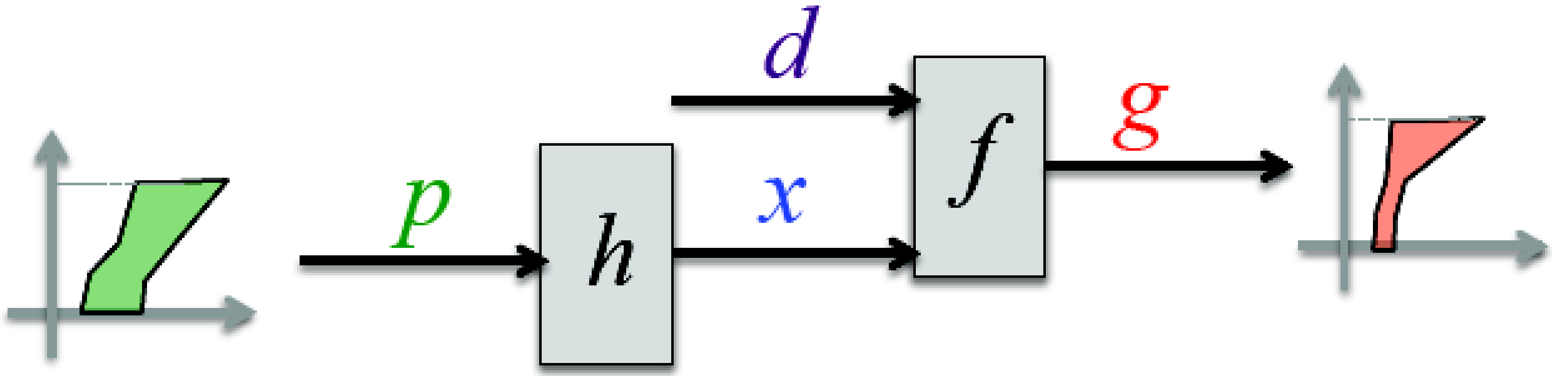
- 21 Uncertain Parameters $\boldsymbol{p}$: loss in control effectiveness, actuator failure, icing, deadzone, and desired range in operating conditions
- 5 Intermediate variables $\boldsymbol{x}$ (e.g. control effectiveness of elevator, time delay due telemetry and communications)
- 14 Design variables $\boldsymbol{d}$: controller gains
- 8 Performance metrics $\boldsymbol{g}$ (e.g. Lon stability, lat/dir stability, elevator actuation)



$$\boldsymbol{x} = h(\boldsymbol{p}) \qquad g = f(\boldsymbol{p}, \boldsymbol{x})$$

Uncertain Parameters (p)

Challenges



Uncertainty models for $\boldsymbol{p}$ are given

Cat. I Random Variables (aleatory uncertainty)

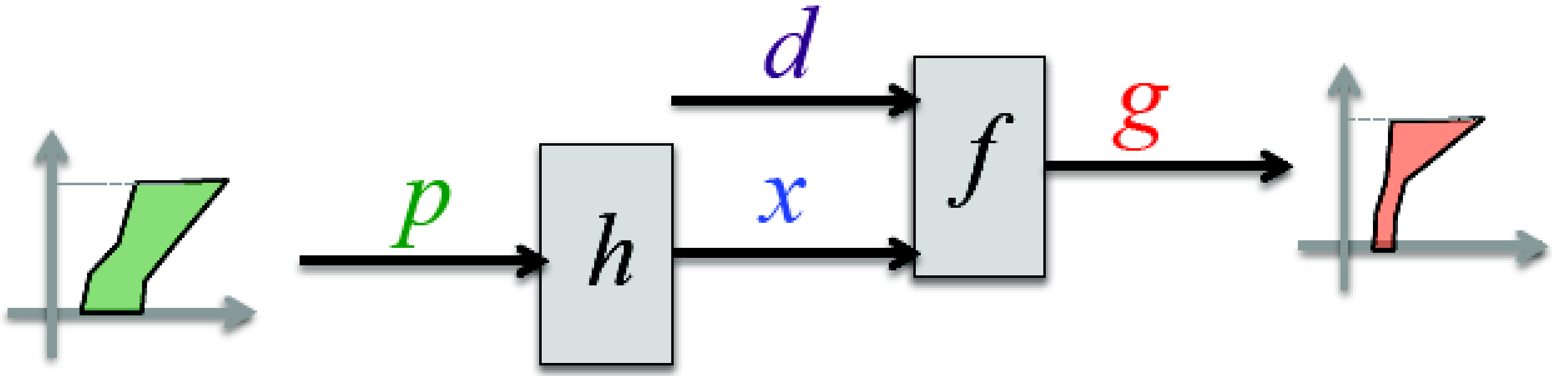
Cat. II Intervals (epistemic uncertainty)

Cat. III Probability boxes (aleatory + epistemic uncertainty)

Some of these Uncertainty Models can be reduced/improved

Uncertain Parameters (p)

Challenges



Uncertainty models for $\boldsymbol{p}$ are given

Cat. I Random Variables (aleatory uncertainty)

Cat. II Intervals (epistemic uncertainty)

Cat. III Probability boxes (aleatory + epistemic uncertainty)

Some of these Uncertainty Models can be reduced/improved

The propagation of $\boldsymbol{p}$ for a fixed $\boldsymbol{d}$ makes $\boldsymbol{g}$ a p-box

Objectives

Tasks

- **Uncertainty Characterization** Advance methods for the refinement of uncertainty models using limited experimental data
- **Sensitivity Analysis** Develop methods for the identification of critical parameters from within a multidimensional parameter space
- **Uncertainty Propagation** Deploy approaches for propagating uncertainties in multidisciplinary systems subject to both aleatory and epistemic uncertainty
- **Extreme Case Analysis** Identify the combination of uncertainties that lead to best- and worst-case outcomes according to two probabilistic measures of interest
- **Robust Design** Determine design points that provide optimal worst-case probabilistic performance. This objective is added as an optional element

NASA Langley UQ challenge problem

Proposed strategy

Solving each problem with at least **two** different approaches

- Cross validate results
- Increase confidence
- Test different hypotheses
- Different numerical implementations

- Theoretical framework: Generalized probabilistic approach (Random Set Theory)
- Computational framework: OpenCossan

Reference

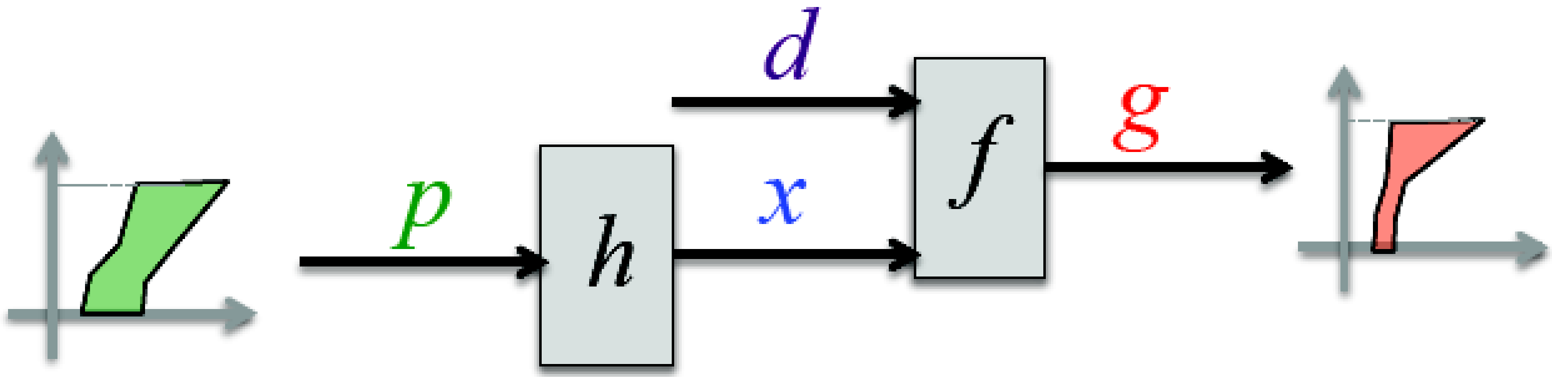
Patelli et al. Uncertainty management in multidisciplinary design of critical safety systems Journal of Aerospace Information Systems, 2014 (DOI:10.2514/1.I010273)

Examples of imprecise probability in engineering

Outline

Uncertainty Characterization

Reduce uncertainty



Sub-Model: $x_1 = h_1(p_1, p_2, p_3, p_4, p_5)$

- Refine the uncertainty model of $\boldsymbol{p}$ given n observations of x_1
- 2 sets of 25 realizations of x_1
- What is the effect of the number of observations n on the quality of the resulting Uncertainty Model?

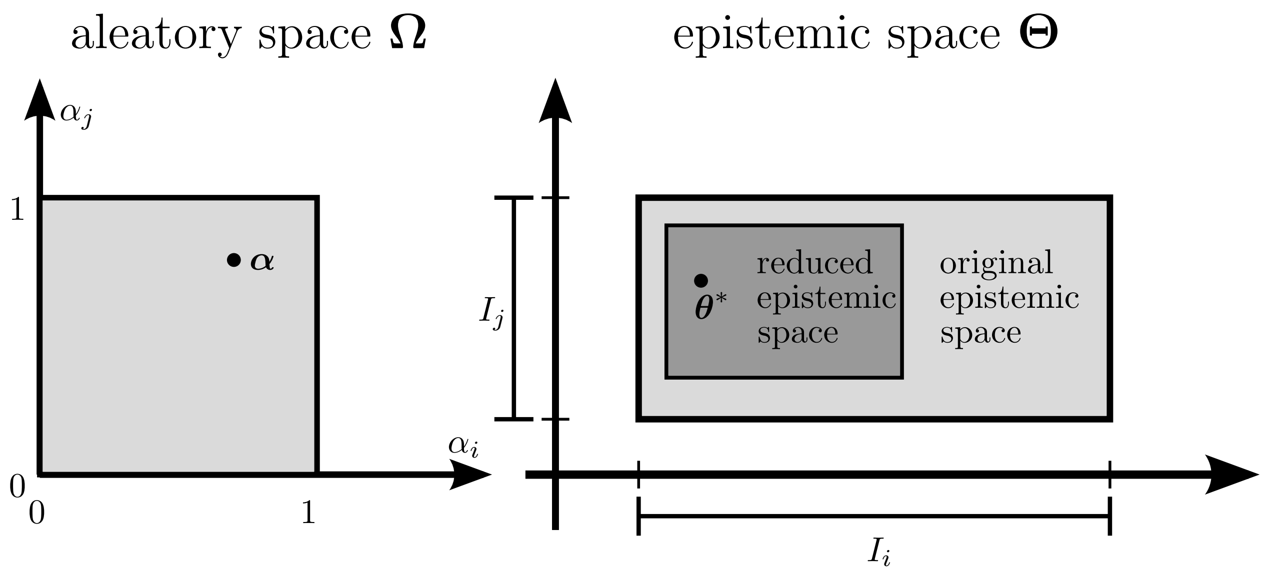
Uncertainty Characterization

Reduce uncertainty

Variable	Category	Aleatory component	Epistemic component	Description
p_1	III	Unimodal Beta (α_1)	$I_1 = [3/5,\ 4/5]$	Interval of $E[p_1]$
			$I_2 = [1/50,\ 1/25]$	Interval of $V[p_1]$
p_2	II		$I_3 = [0,\ 1]$	Interval
p_3	I	Uniform(0,1] (α_2)		Random variable
p_4, p_5	III	Multivariate Gaussian (α_3,α_4)	$I_4 = [-5,\ 5]$	Interval of $E[p_4]$
			$I_5 = [1/400,\ 4]$	Interval of $V[p_4]$
			$I_6 = [-5,\ 5]$	Interval of $E[p_5]$
			$I_7 = [1/400,\ 4]$	Interval of $V[p_4]$
			$I_8 = [-1,\ 1]$	Interval of $\rho(p_4, p_5)$

4 aleatory terms and 8 epistemic components.

Aim and strategy

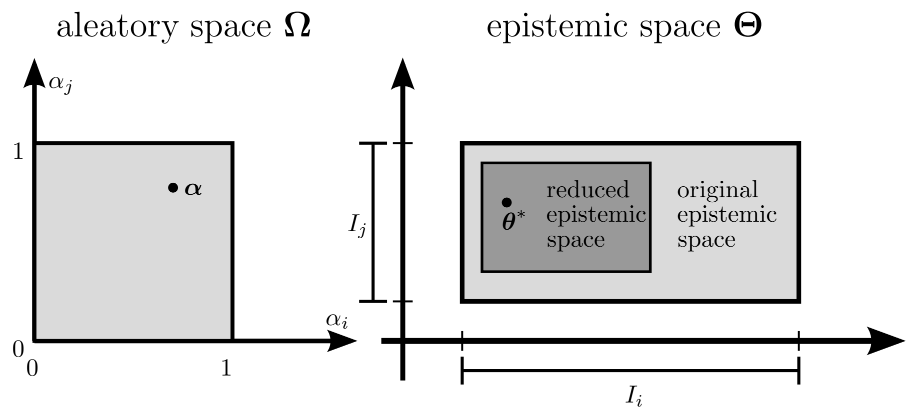


Aleatory space, defined as $\boldsymbol{\Omega} = (0, 1]^4$ serves as the support of the copula which models the dependence between variables p_1, p_3, p_4 and p_5.

Uncertainty Characterization

Aim and strategy

- Refine the uncertainty model of $\boldsymbol{p}$ given n observations
- What is the effect of the number of observations n?



defined as $\boldsymbol{\Theta} = \times_{i=1}^{8} I_i$; "true" parameter vector $\boldsymbol{\theta}^*$; when additional information is available, the epistemic uncertainty is reduced.

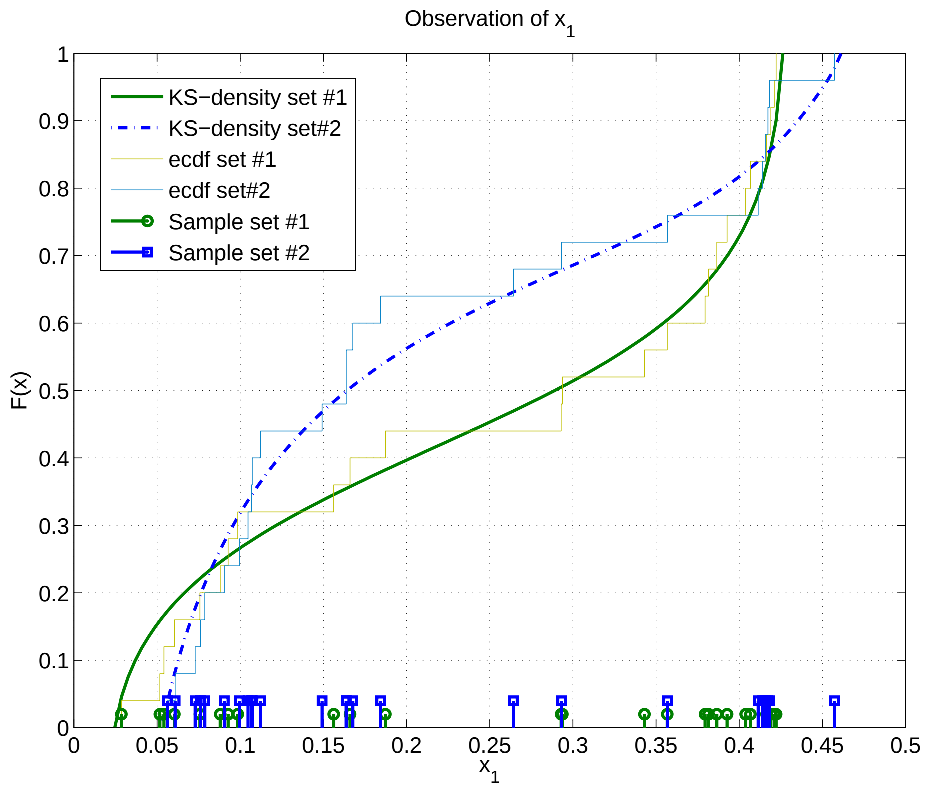
Strategies

- Bayesian updating on the epistemic space
- Non-parametric statistical methods

Available Data

Observations

- 2 sets of 25 realizations of x_1
- Observations come from stochastic model (aleatory uncertainty)
- eCDF of two data sets are quite different
- Gaussian Smoother approach



Each realization associate to a Gaussian distribution:

$$\hat{F}_h(x) = \frac{1}{n\sigma\sqrt{2\pi}} \sum_{k=1}^{n} \exp\left(\frac{-(x-x_k)^2}{2\sigma^2}\right)$$

Pradlwarter, H. and Schuëller, G., The Use of Kernel Densities and Confidence Intervals to cope with Insufficient Data in Validation Experiments, *Computer Methods in Applied Mechanics and Engineering*, 197(29-32), **2008**, 2550–2560.

Bayesian updating on the epistemic space

The updated belief about $\boldsymbol{\theta}^*$ after observing the given dataset $\mathcal{D}_n$, is modelled by the posterior PDF $p(\boldsymbol{\theta}|\mathcal{D}_n)$:

$$p(\boldsymbol{\theta}|\mathcal{D}_n) = \frac{p(\mathcal{D}_n|\boldsymbol{\theta})p(\boldsymbol{\theta})}{P(\mathcal{D}_n)}.$$

Likelihood function:

$$p(\mathcal{D}_n|\boldsymbol{\theta}) = \prod_{i=1}^{n} p(\boldsymbol{x_i};\boldsymbol{\theta})$$

is related to the probability of observing the samples $\mathcal{D}_n = \{\boldsymbol{x}_i, i = 1, 2, \ldots, n\}$ assuming that the true parameters underlying the model PDF $p(\boldsymbol{x};\boldsymbol{\theta})$ is $\boldsymbol{\theta}$ when a set of independent and identically distributed observations $\mathcal{D}_n$ is available.

Bayesian updating on the epistemic space

Likelihood estimation through kernel density

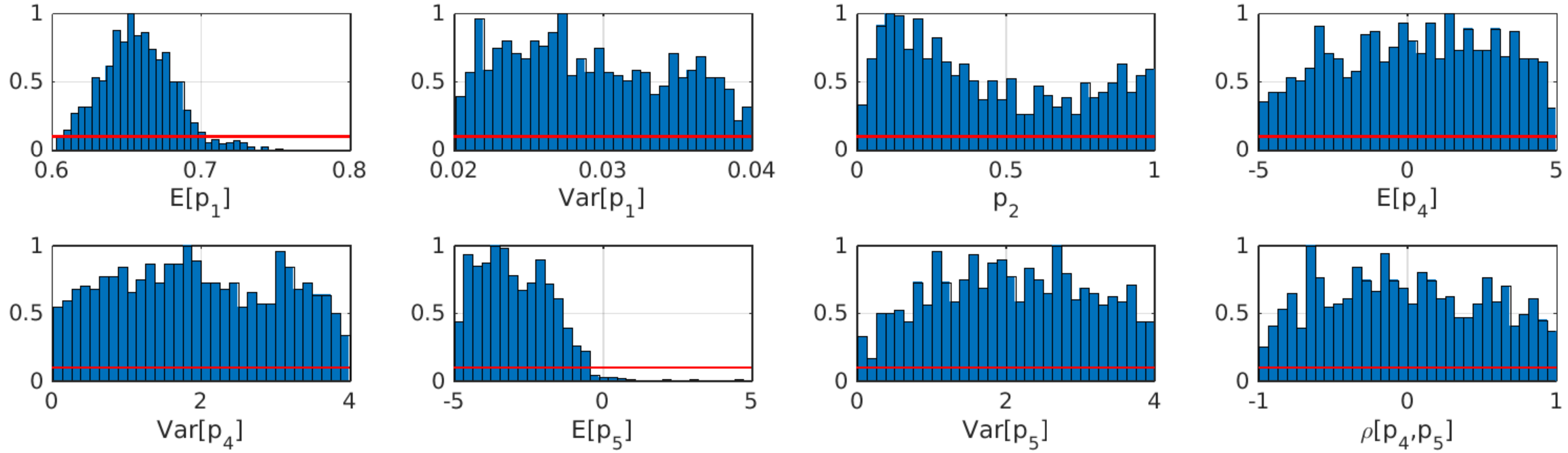
Assume that the samples $\mathcal{D}_n$ were drawn from $p(\boldsymbol{x}; \boldsymbol{\theta}^*)$, where $\boldsymbol{\theta}^* \in \Omega$; the likelihood $p(\mathcal{D}_n|\boldsymbol{\theta}_i)$ will be defined in the following way:

- draw $M = 100$ points from the aleatory space $\boldsymbol{\Omega}$, using copula C; we will call these samples $\{\boldsymbol{\omega}_j : j = 1, \ldots, M\}$;
- calculate $x_1^{(j)} := h_1(\boldsymbol{\omega}_j, \boldsymbol{\theta}_l)$ for $j = 1, \ldots, M$;
- using **kernel density** estimation and the samples $\{x_1^{(j)} : j = 1, \ldots, M\}$, estimate the PDF $p(\boldsymbol{x}|\boldsymbol{\theta}_i) \equiv p(\boldsymbol{x}; \boldsymbol{\theta}_i)$
- calculate the likelihood function as

$$p(\mathcal{D}_n|\boldsymbol{\theta}_i) = \prod_{k=1}^{n} p(\boldsymbol{x}_k|\boldsymbol{\theta}_i)$$

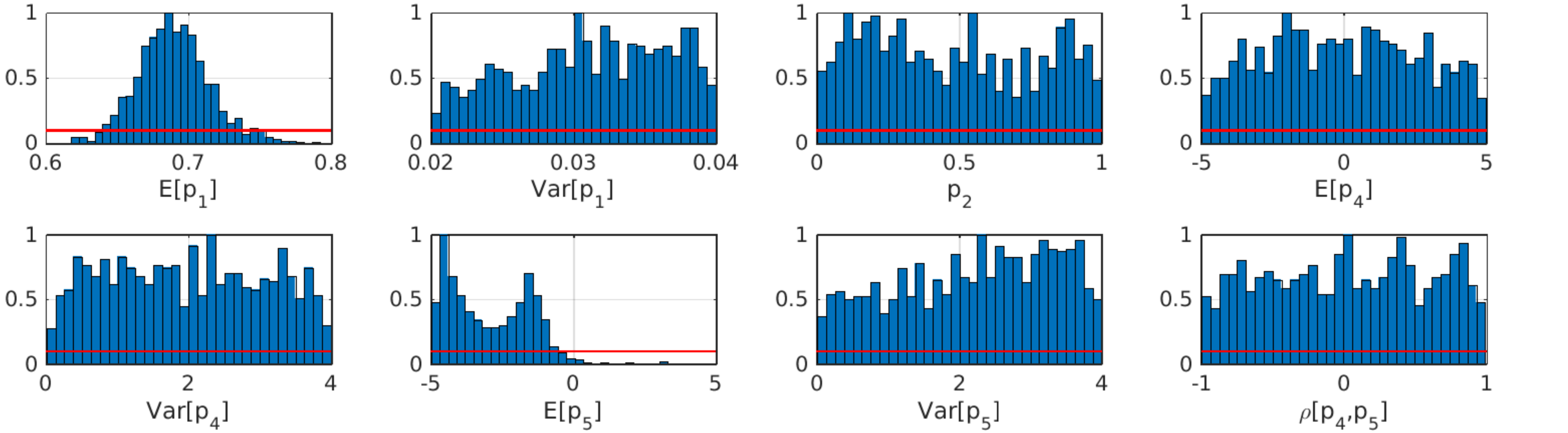
Reduced epistemic space

Bayesian updating (25 observations)



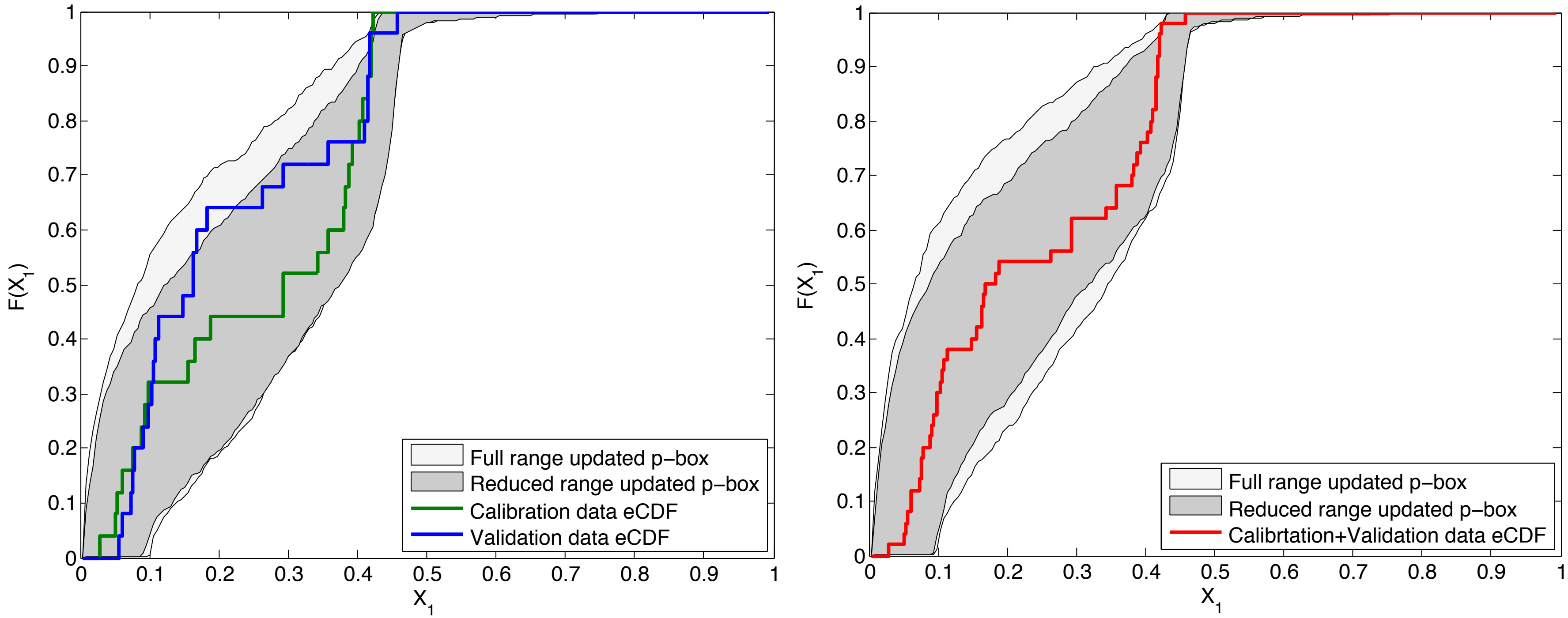
Reduced epistemic space

Bayesian updating (50 observations)



Reduced epistemic space

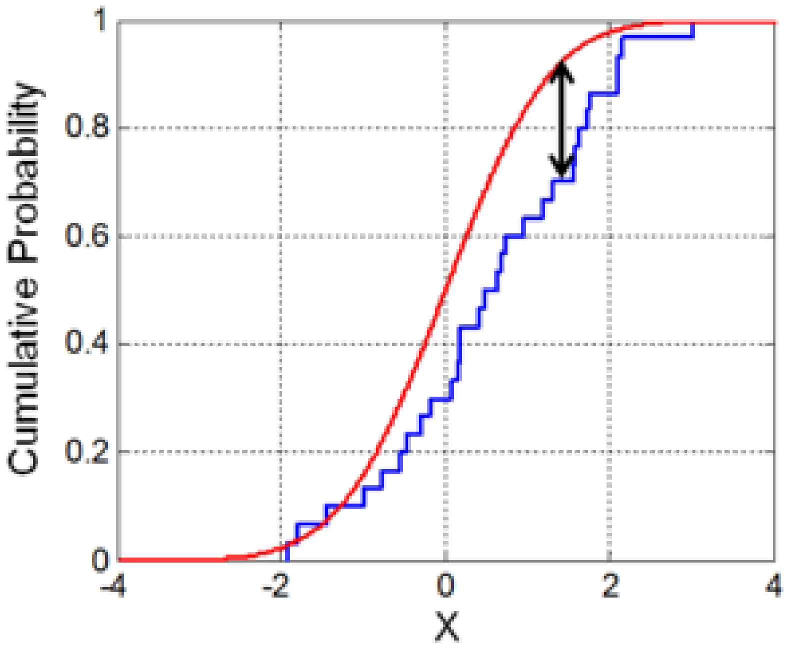
Bayesian updating



Uncertainty Characterization

Non parametric approach

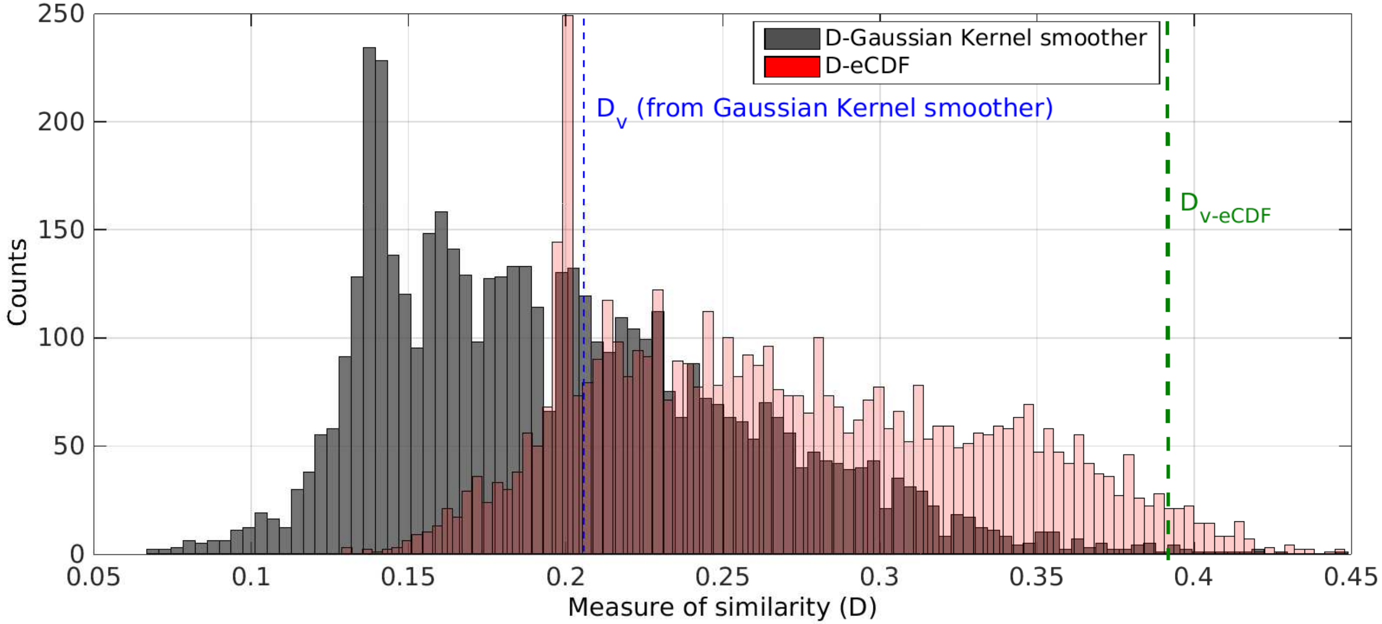
1. Generate n_i realizations on the epistemic space $\boldsymbol{\theta}_i$
2. Evaluate the model $x_1^i := h_1(\boldsymbol{\alpha}_j; \boldsymbol{\theta}_i)$ for $j = 1, \ldots, n_j$;
3. Estimate the empirical CDF $\hat{F}(\cdot|\boldsymbol{\theta}_i)$
4. Compute Kolmogorov-Smirnov test (measure of similarity D_i)



Uncertainty Characterization

Non parametric approach

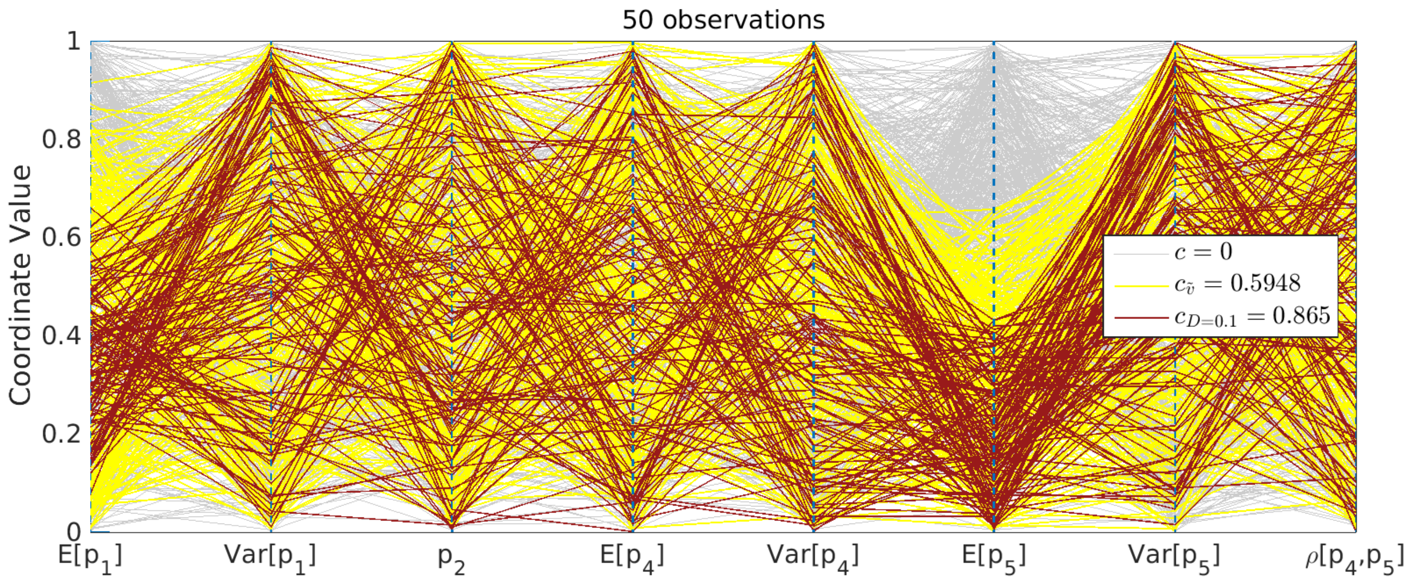
1. Generate n_i realizations on the epistemic space $\boldsymbol{\theta}_i$
2. Evaluate the model $x_1^i := h_1(\boldsymbol{\alpha}_j; \boldsymbol{\theta}_i)$ for $j = 1, \ldots, n_j$;
3. Estimate the empirical CDF $\hat{F}(\cdot|\boldsymbol{\theta}_i)$
4. Compute Kolmogorov-Smirnov test (measure of similarity D_i)
5. If $D_i < D_{\tilde{v}}$ collect $\boldsymbol{\theta}_i$.



Uncertainty Characterization

Non parametric approach

1. Generate n_i realizations on the epistemic space $\boldsymbol{\theta}_i$
2. Evaluate the model $x_1^i := h_1(\boldsymbol{\alpha}_j; \boldsymbol{\theta}_i)$ for $j = 1, \ldots, n_j$;
3. Estimate the empirical CDF $\hat{F}(\cdot|\boldsymbol{\theta}_i)$
4. Compute Kolmogorov-Smirnov test (measure of similarity D_i)



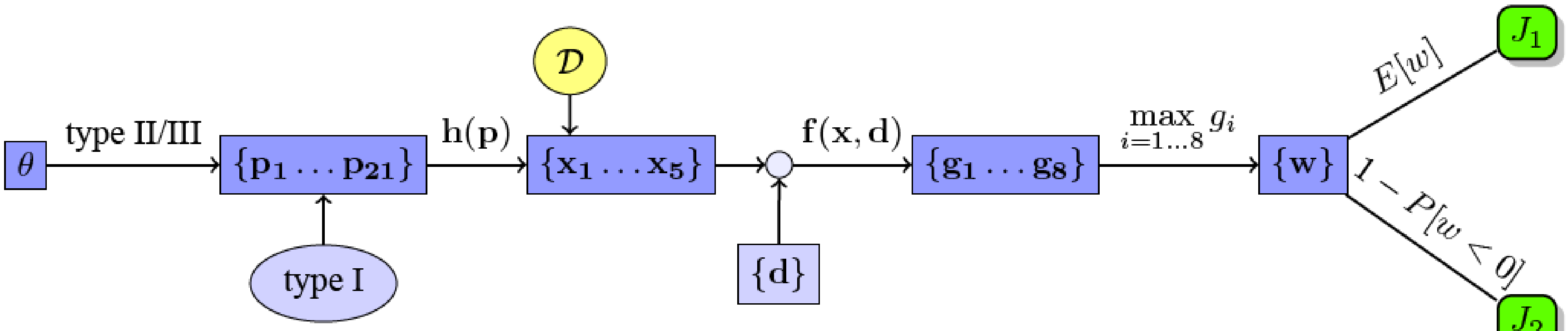
Examples of imprecise probability in engineering

Outline

General remarks

Aim

Rank 4 category II & III parameters $(x_1, \ldots, x_5)$
Rank 17 category II & III parameters $(J_1\ J_2)$



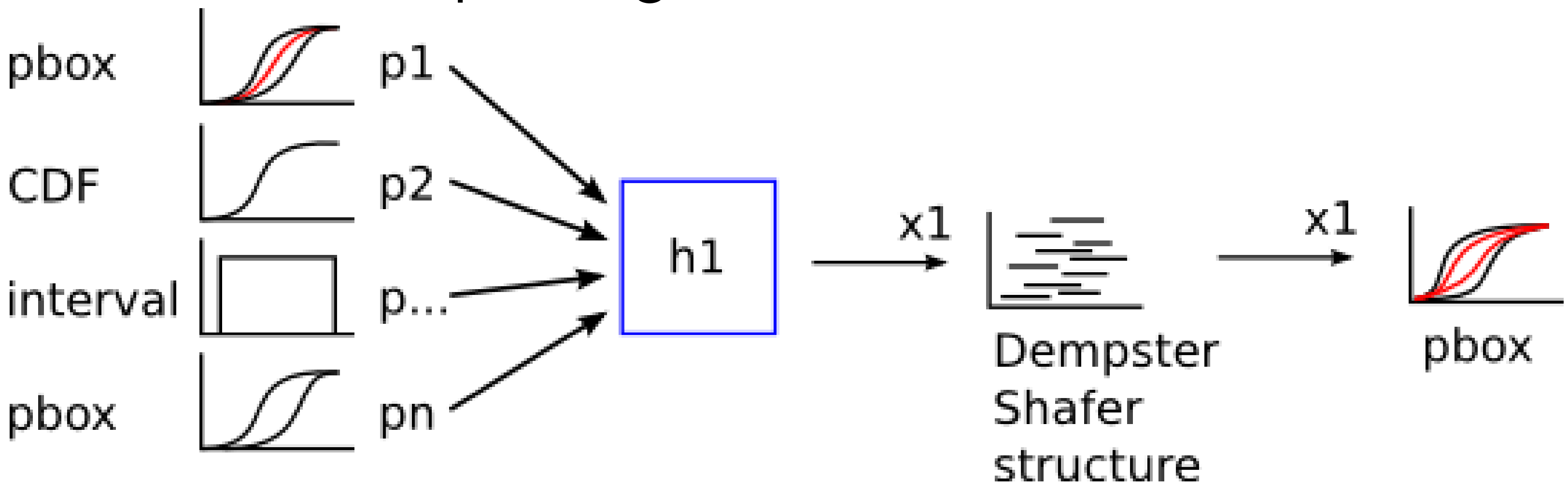
Two strategies were employed:

- Nonspecificity techniques
- Global Sensitivity Analysis

Nonspecificity techniques

Procedure and definitions

Pinching: $\underline{I_i} + p \cdot (\overline{I_i} - \underline{I_i})$ where $p = \{0.1,\ 0.3,\ 0.5,\ 0.7,\ 0.9\}$
Nonspecificity: measure of epistemic uncertainty, of the output random set after pinching was estimated



- Klir, G. J. and Wierman, M. J., *Uncertainty-Based Information: Elements of Generalized Information Theory*, Vol. 15 of *Studies in Fuzziness and Soft Computing)*, Physica-Verlag, Heidelberg, Germany, 1998.
- Klir, G. J., *Uncertainty and Information : Foundations of Generalized Information Theory*, John Wiley and Sons, New Jersey, 2006.
- Alvarez, D. A., *Reduction of uncertainty using sensitivity analysis methods for infinite random sets of indexable type, International Journal of Approximate Reasoning*, Vol. 50, No. 5, 2009, pp. 750 – 762.

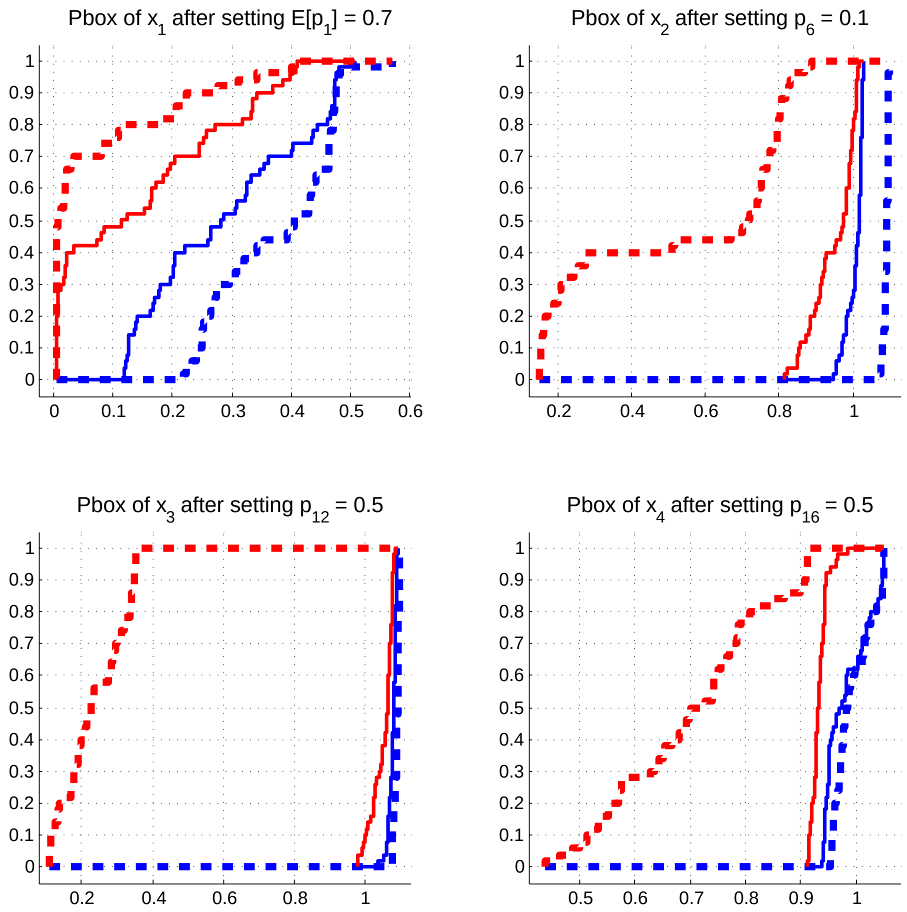
Nonspecificity approach

Some examples

Probability boxes corresponding to the original (dashed lines) and pinched (continuous lines) output probability boxes.

Tools:

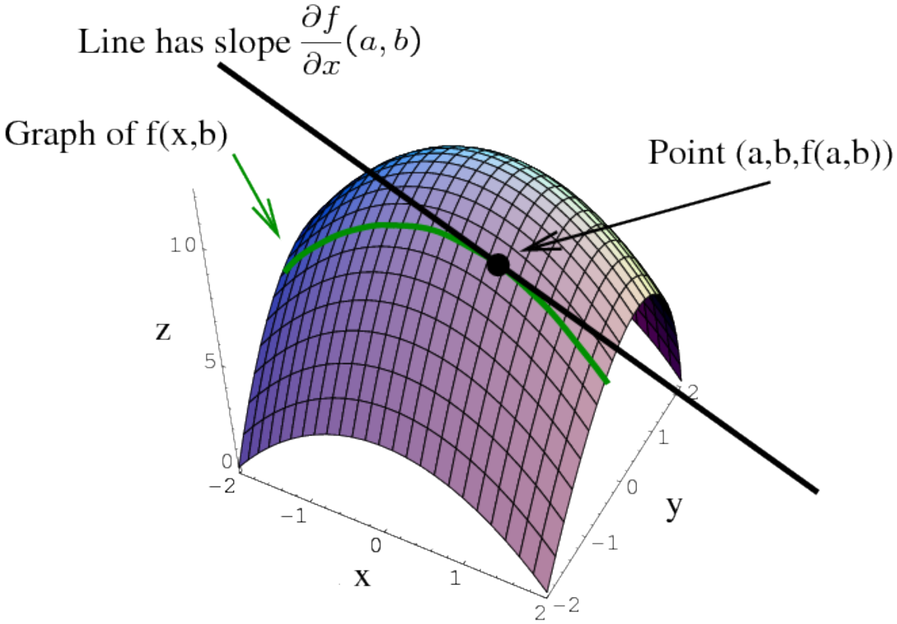
- 50 samples for each pinching
- GA for each focal element (30000 individuals and 10 generations)



Sensitivity Analysis
Background

Local Sensitivity Analysis

- Obtained varying input factors one at time and holding others fixed to a nominal value Z_i^*
- Involve partial derivatives



$$s_{Z_i} = V(Y|Z_i = z_i^*) = \frac{\partial f(\boldsymbol{Z})}{\partial z_i}|_{Z_i=z_i^*} * \sigma_{z_i}$$

Sensitivity Analysis

Background

Global Sensitivity Analysis

- Obtained averaging the local sensitivity analysis over the distribution of uncertain factors, $E(V(Y|Z_i = z_i^*))$
- Variation of the output is taken globally (the $supp(Z)$ is explored)
- Identify "active" input factors at a low computational cost
- Sensitivity measures (Sobol' indexes) can by estimated via Monte Carlo Method

Global sensitivity analysis

General remarks

Compute Sobol' indices and Total indices

$$S_i = \frac{Var_{X_i}[E_{X\sim i}(Y \mid x_i)]}{V[Y]} \quad T_i = 1 - \frac{\mathrm{Var}_{\mathbf{X}_{\sim i}}\left(E_{X_i}(Y \mid \boldsymbol{X}_{\sim i})\right)}{\mathrm{Var}(Y)}$$

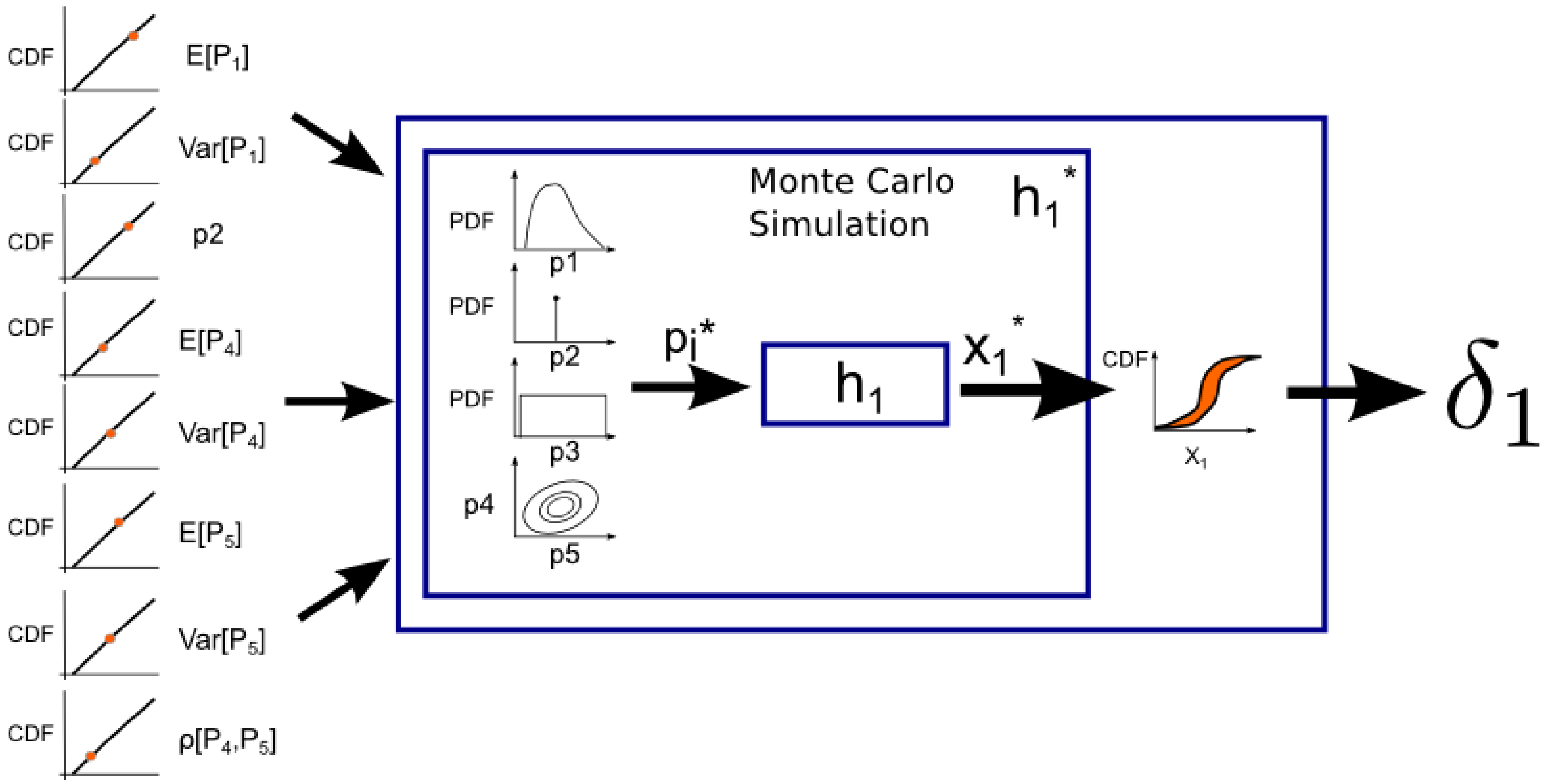
Requirements:

- Exact knowledge of PDF
- Variance of a measurable output

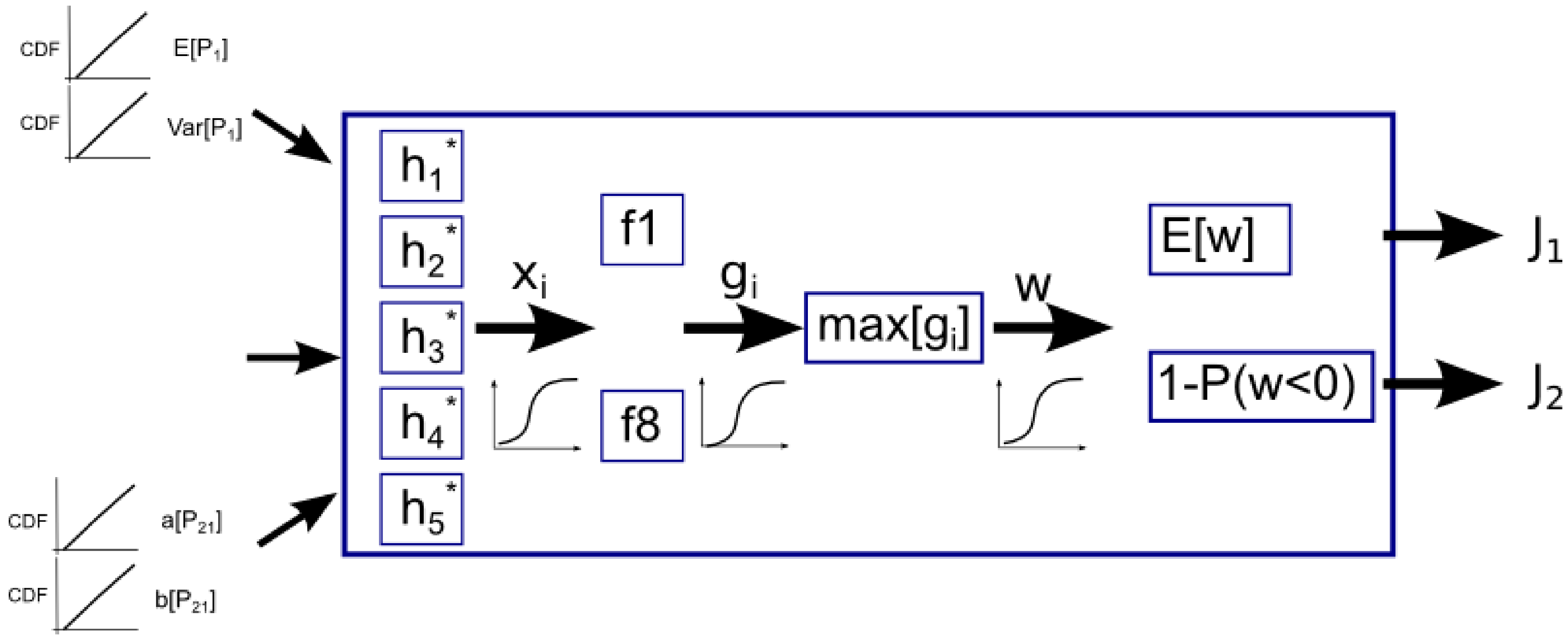
Strategy:

- Redefined model h^*: no p-boxes and a scalar output
- $\delta_i = \int_{-\infty}^{+\infty} |F_i(x) - F_{ref}(x)| dx$ (Problem B1)
- extended-FAST method and Saltelli's method

Redefined model for sensitivity analysis

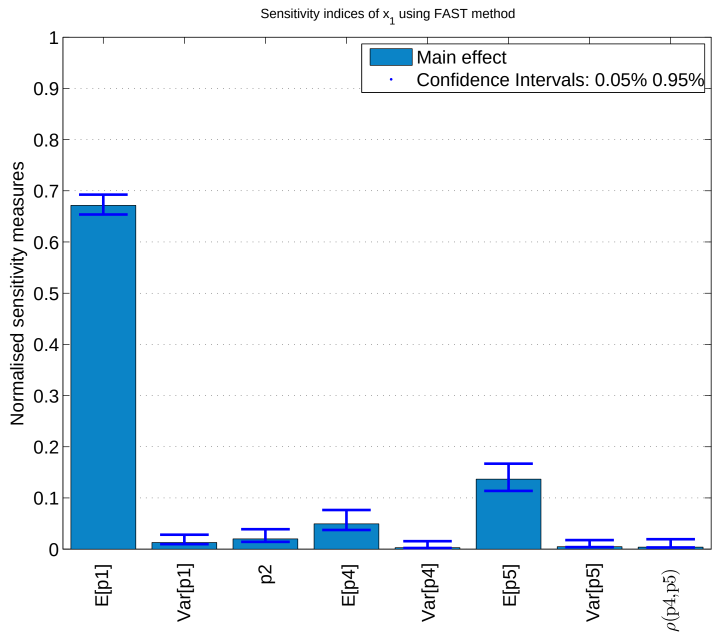


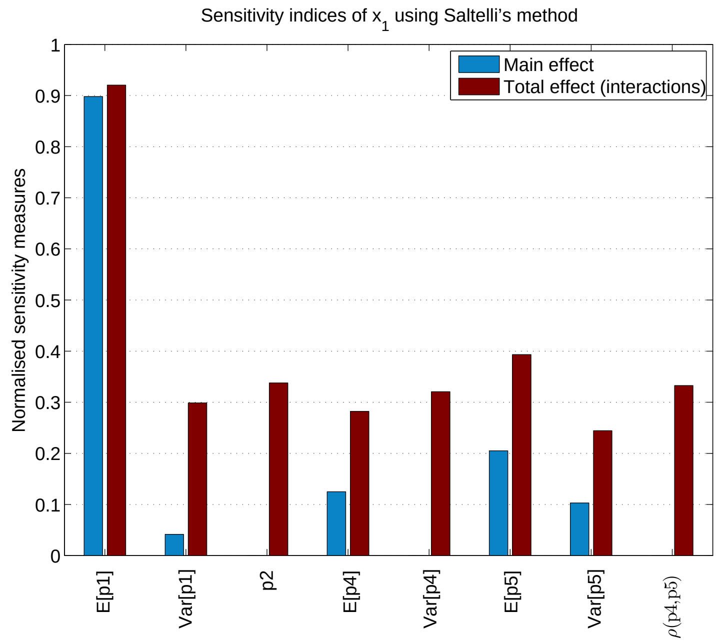
Redefined model for sensitivity analysis



Global Sensitivity Analysis

Results tasks B1





extended-FAST approach and Saltelli's method for x_1

Uncertainty Propagation

General remarks

Aim

Find the range of the metrics with *reduced* and *improved* models

$$J_1 = E[w(\boldsymbol{p}, \boldsymbol{d}_{\text{baseline}})] \quad \text{and} \quad J_2 = 1 - P[w(\boldsymbol{p}, \boldsymbol{d}_{\text{baseline}}) < 0]$$

Two strategies were employed:

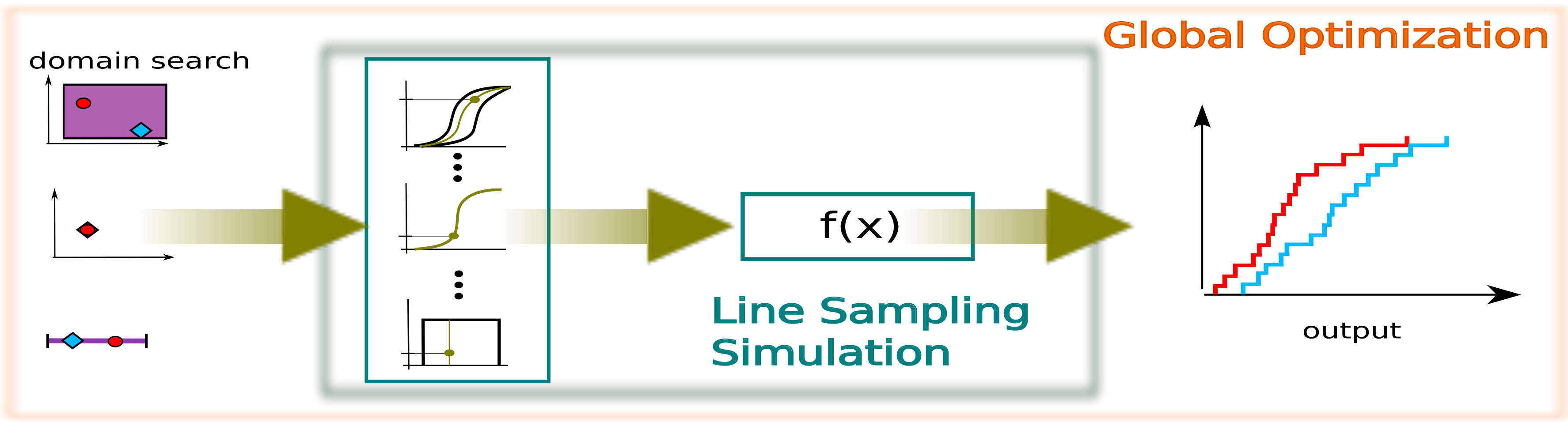
- Approach **A**: propagate intervals obtained from given distribution-free p-boxes and construct Dempster-Shafer structure
- Approach **B**: global optimization in the epistemic space (search domain). Monte Carlo Simulation to estimate J_1 and J_2

Uncertainty Quantification

Epistemic + Aleatory uncertainty

Double loop approach

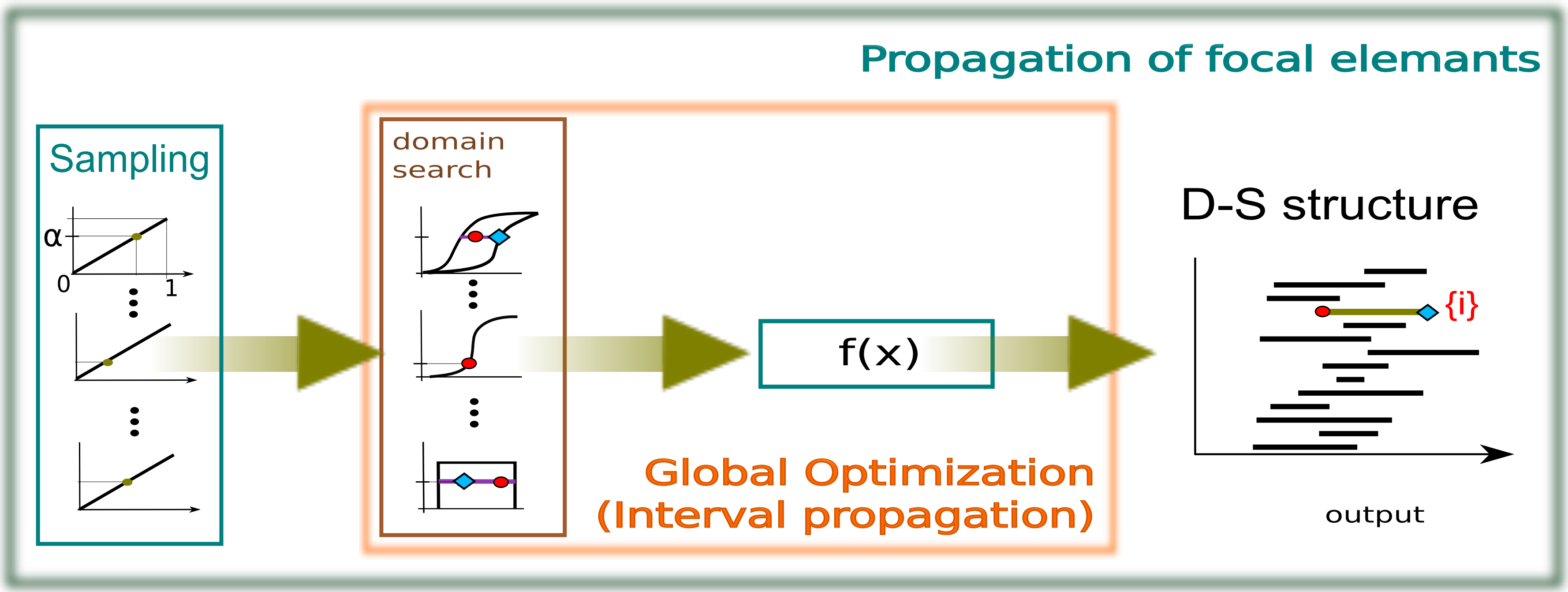
- Sampling intervals and reliability analysis for each sample
- Difficult to treat *distribution-free* p-boxes
- Huge computational efforts
- Allows to identify "extreme realizations"



Uncertainty Quantification

Propagation of Focal Elements

- Sample focal elements in the α-space
- Propagate intervals and construct Dempster-Shafer structure
- Limitation: treats p-boxes as *distribution-free* p-boxes



Examples of imprecise probability in engineering

Outline

Extreme Case Analysis

Aim and strategy

Aim

- Identify epistemic realizations that yields extreme values of

$$J_1 = E\left[w(\mathbf{p},\ \mathbf{d}_{baseline})\right]; \quad J_2 = 1 - P\left[w(\mathbf{p},\ \mathbf{d}_{baseline}) < 0\right]$$

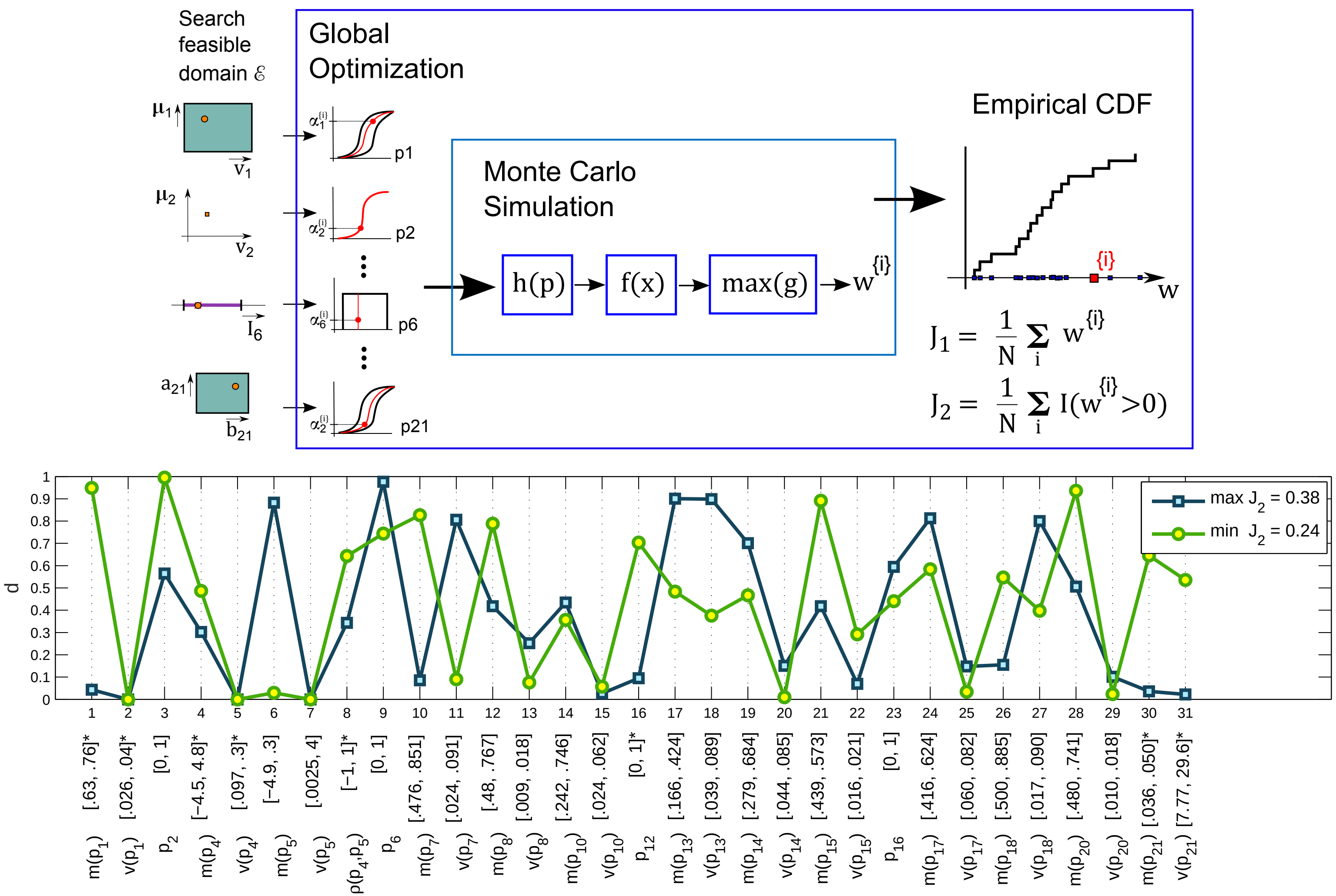
- Qualitatively describe the relation between sub-disciplines $\mathbf{x}$ and failure modes $\mathbf{g}$.

Strategies:

- Identify realizations from Subproblem C - Approach B
- Analyses range variability of the performance functions $\mathbf{g} = f(\mathbf{x})$

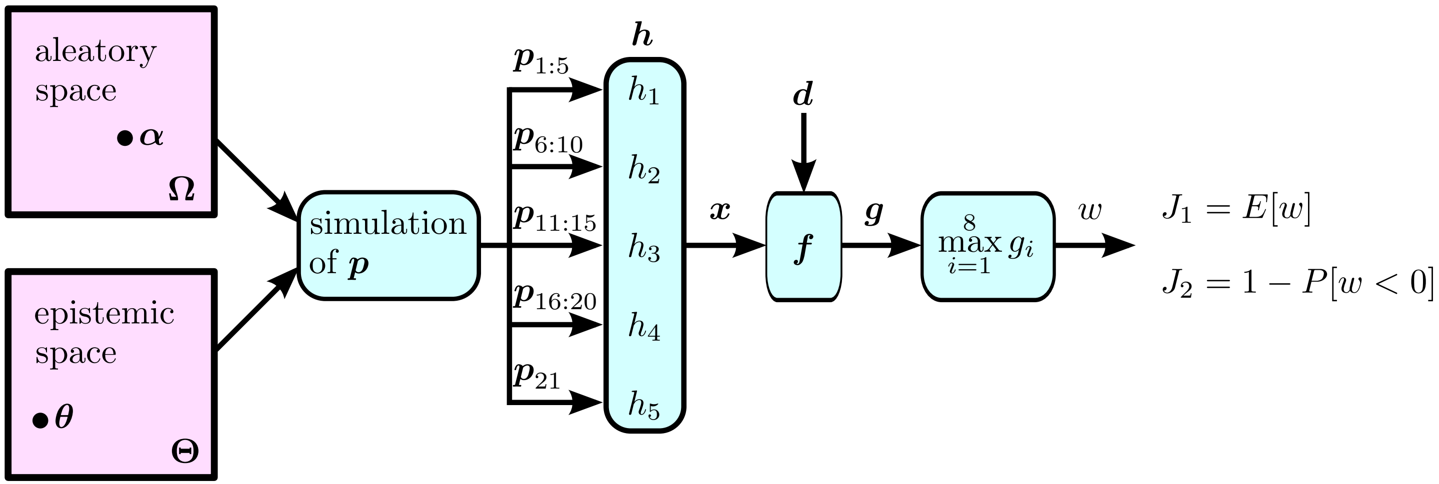
Extreme Case analysis

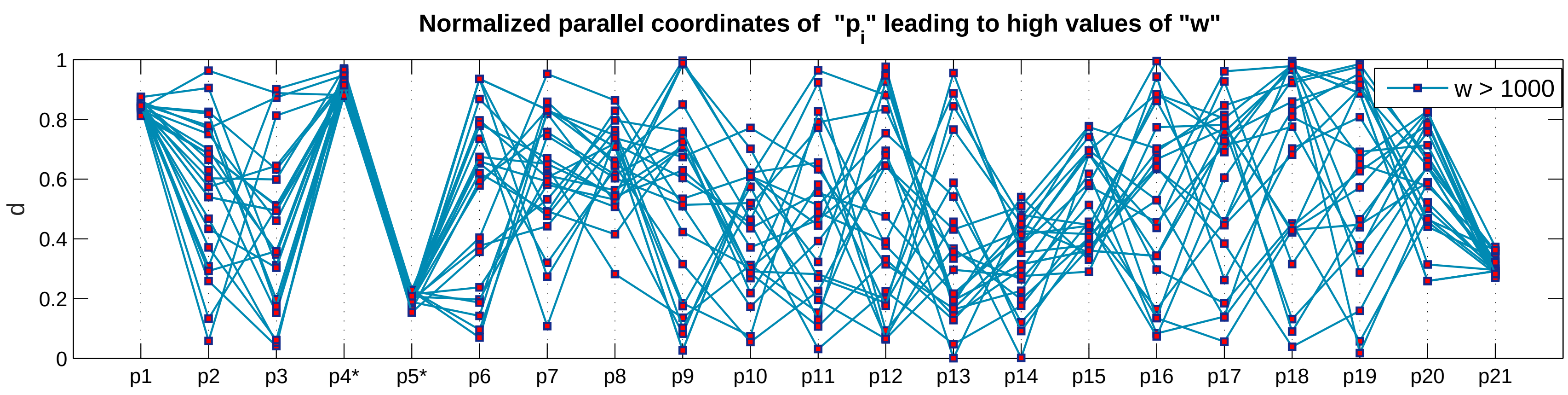
Identify realizations



Extreme Case analysis

Analysis of the failure modes





Examples of imprecise probability in engineering

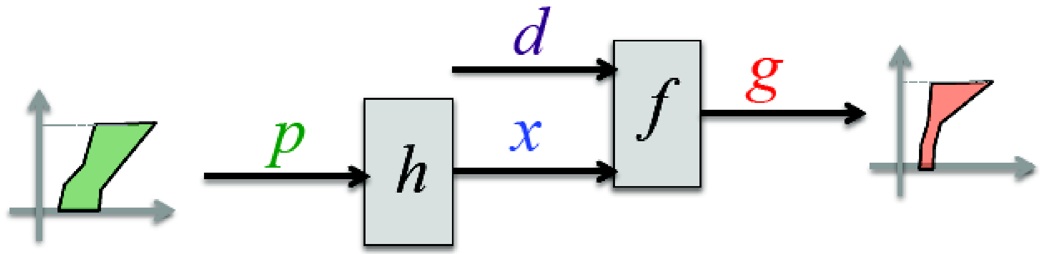
Outline

- Uncertainty Management in Engineering
  - Motivation
  - Method of analysis
  - Limitation of traditional approaches
  - Random Set
- The NASA UQ Challenge Problem
  - The model
  - Uncertainty Characterization (Subproblem A)
  - Sensitivity Analysis (Subproblem B)
  - Uncertainty Propagation (Subproblem C)
  - Extreme Case Analysis (Subproblem D)
  - Robust Design (Subproblem E)
- Conclusions

Robust Design

Aim and strategy

Robust design in presen



- **Design Variable:** 14 control parameters ($\boldsymbol{d}$)
- **Objective Function:** minimize $max(J_1)$ (expected value)
- **Objective Function:** minimize $max(J_2)$ (failure probability)
- Sensitivity analysis of the obtained designs

Computational challenges

- Each candidate solution: $\approx$ 3 days (i.e. $max(J_1)$, $max(J_2)$)
- Running time: $\approx$ 3 days (for each solution)

Strategies:

- Surrogate model (Artificial Neural Networks)
- Optimizer: Genetic Algorithms (and BOBYQA)

Robust Design

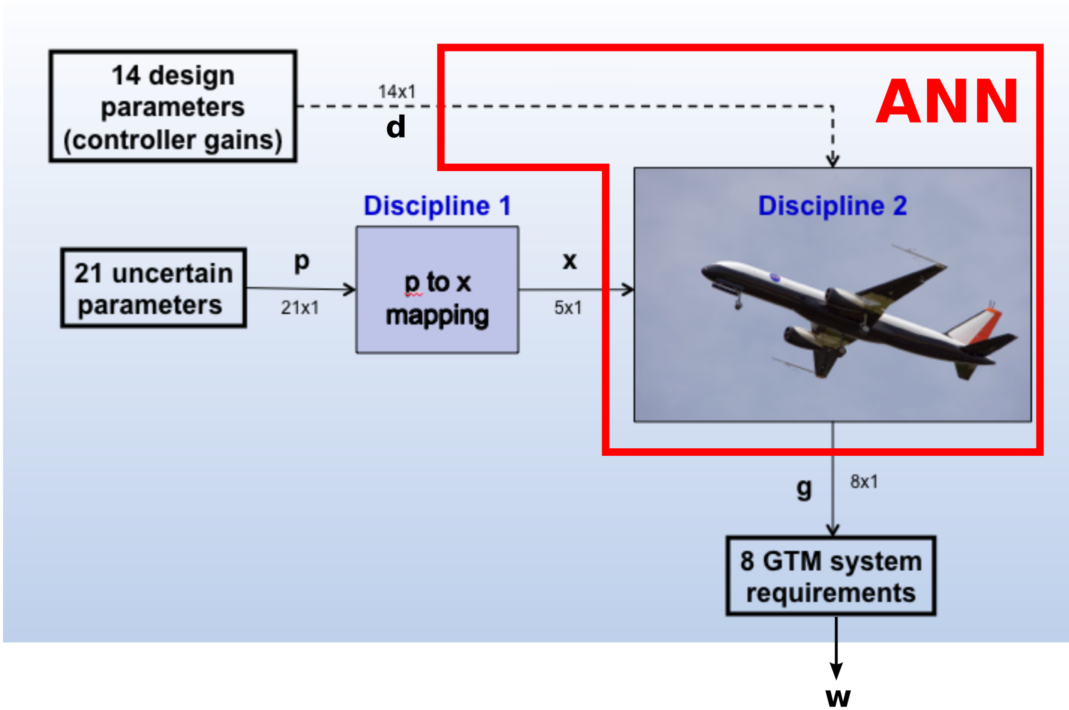
Surrogate model (Artificial Neural Networks)

Approximation of the most computationally expensive part

- Inputs: $\mathbf{x}$ and $\mathbf{d}$
- Outputs: $\mathbf{g}$

Results:

- $max(J_1) = 0.0044$ (baseline 3.05)
- $max(J_2) = 0.34$ (baseline 0.41)



Interval/Random Predictor Model

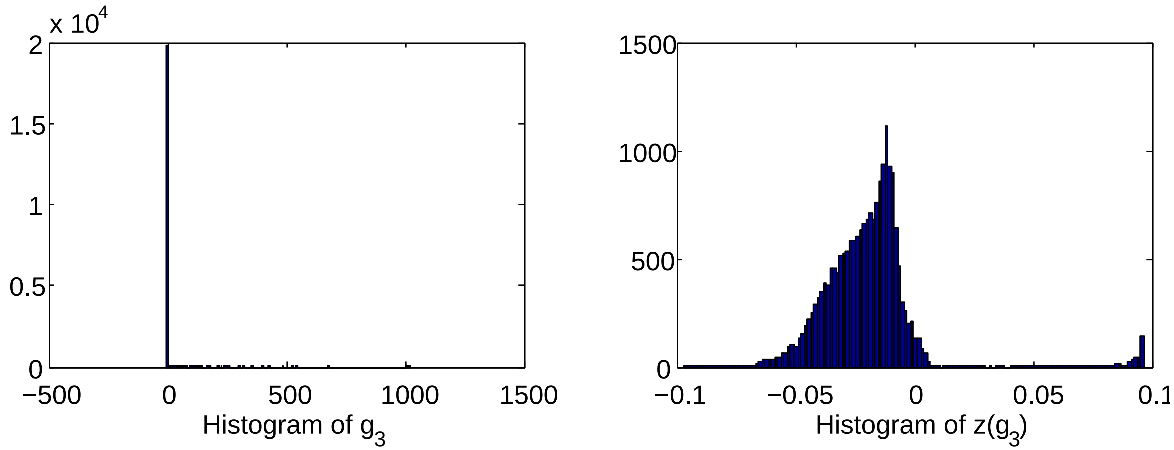
Friday 09:00 - 11:00 Part 12

Training data issues

- Positive and negative values of g concentrated around zero
- Few extremely high values

Non linear transformation: $z(g_j) := \frac{1}{200|\min(g_j)|} - \frac{1}{100(g_j + 2|\min(g_j)|)}$

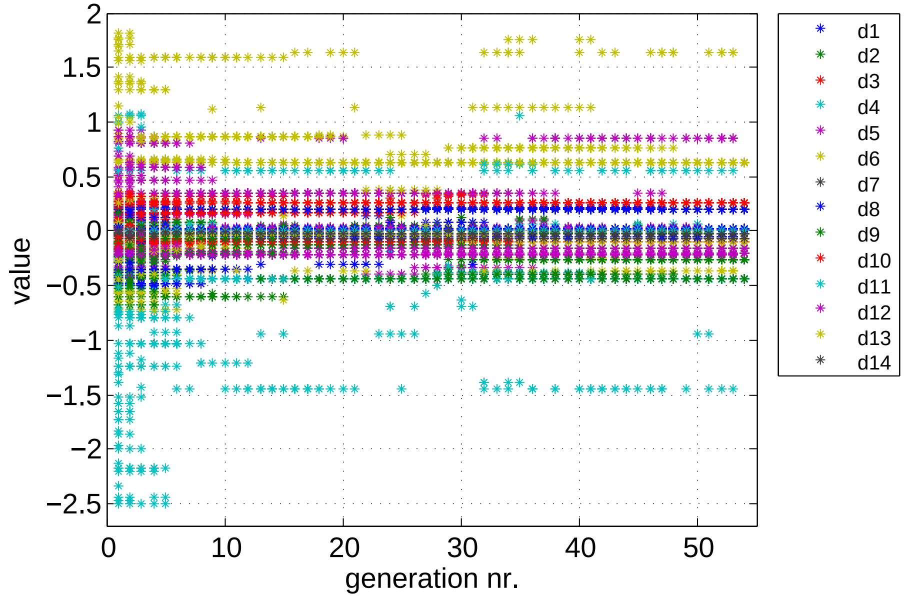
j	$\min(g_j)$	$\max(g_j)$
1	-1.4353	5.57
2	-5.3903	15.90
3	-0.0533	1036.2
4	-0.0015	1159.2
5	-0.4375	1018.8
6	-0.0182	1080.1
7	-0.0002	1137.2
8	-0.0129	1072.5

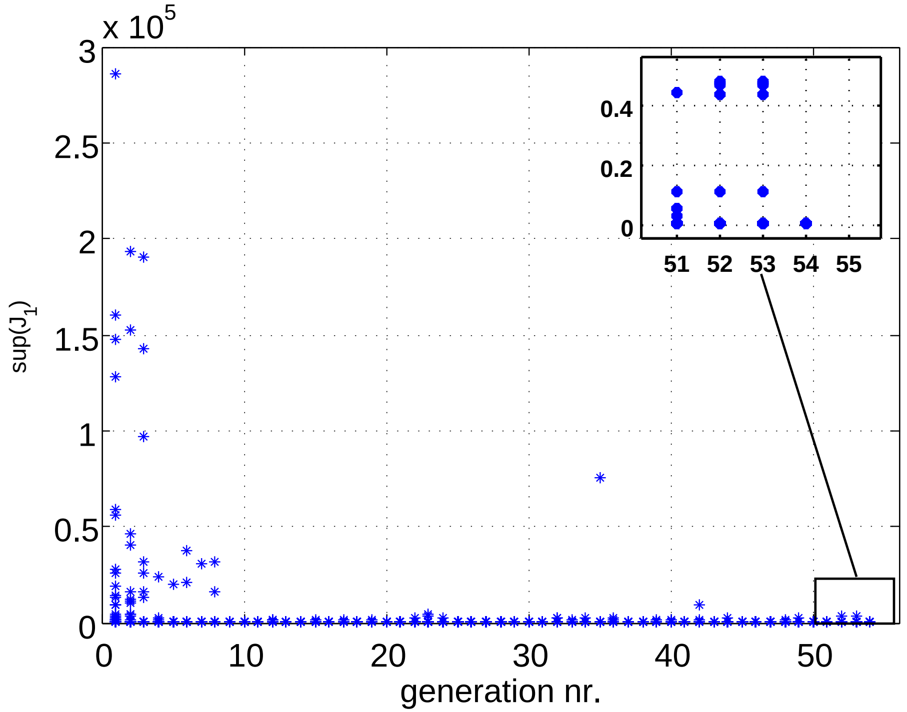


Original and transformed values of g_3

Optimization: Genetic Algorithms

Task E1: minimization of $\max(J_1)$





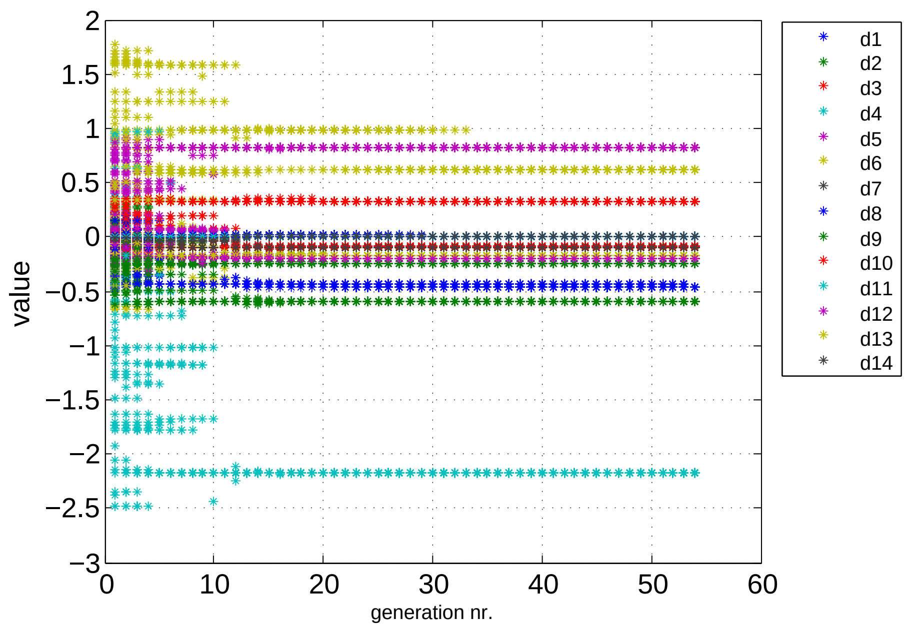
$\mathbf{d}_{E1} = [$ 0.0140, -0.2568, -0.0944, -0.4405, -0.1508, -0.1029, -0.0713, 0.2002, -0.4431, 0.2579, 0.0044, -0.2086, 0.6330, -0.0166 $]$

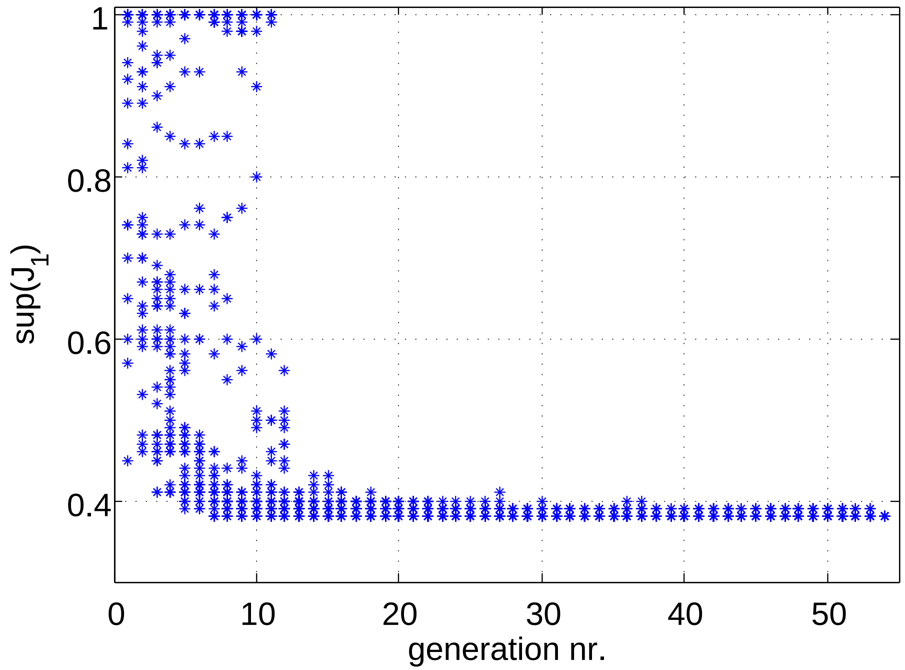
$max(J_1) = 0.0044$ (baseline 3.05)

Optimization: Genetic Algorithms

Task E2: minimization of $\max(J_2)$

Objective: minimization of $\max(J_2)$





$\mathbf{d}_{E2} =$ [0.0102, -0.2409, -0.0844, -2.1800, 0.8192, -0.1642, -0.0981, -0.4674, -0.5958, 0.3230, 0.0015, -0.1975, 0.6249, 0.0052]

$max(J_2) = 0.34$ (baseline 0.41)

Outline

Uncertainty quantification in Engineering

Generalised Probabilistic method

- Rigorous framework to deal with scarce data, imprecision and vagueness without making any assumption
- Dealing with different representations of uncertainties
- Utilisation of traditional stochastic methods and techniques
- Modelling epistemic uncertainties as aleatory uncertainty might lead to **severe over/under estimations**
- Provides bounds of the estimations (traditional probabilistic results always included)

Challenges

Still computational intensive analysis

Summary of the approaches

Drawbacks and limitations

Uncertainty Characterization

- Bayesian: $> 10^6$ model evaluations and difficult to implement
- Non-Parametric: very fast and easy to implement

Summary of the approaches

Drawbacks and limitations

Uncertainty Characterization

- Bayesian: $> 10^6$ model evaluations and difficult to implement
- Non-Parametric: very fast and easy to implement

Sensitivity method

- Global sensitivity method: 5×10^5. (using traditional tools)
- Nonspecificity method: 5×10^7 (required dedicated tools).

Summary of the approaches

Drawbacks and limitations

Uncertainty Characterization

- Bayesian: $> 10^6$ model evaluations and difficult to implement
- Non-Parametric: very fast and easy to implement

Sensitivity method

- Global sensitivity method: 5×10^5. (using traditional tools)
- Nonspecificity method: 5×10^7 (required dedicated tools).

Uncertainty propagation

- 250×10^6 model evaluations ($N\alpha = 1000$ focal elements + GA with 10000 individuals and 50/55 iterations)

Summary of the approaches

Drawbacks and limitations

Uncertainty Characterization

- Bayesian: $> 10^6$ model evaluations and difficult to implement
- Non-Parametric: very fast and easy to implement

Sensitivity method

- Global sensitivity method: 5×10^5. (using traditional tools)
- Nonspecificity method: 5×10^7 (required dedicated tools).

Uncertainty propagation

- 250×10^6 model evaluations ($N\alpha = 1000$ focal elements + GA with 10000 individuals and 50/55 iterations)

Robust Design

- Requires surrogate model (Friday for further details)

Computationally demanding approach (global optimization)