Jan Chleboun
Elements of uncertainty modeling

Persistent URL: http://dml.cz/dmlcz/702739

ELEMENTS OF UNCERTAINTY MODELING[*]

Jan Chleboun

Abstract

The goal of this contribution is to introduce some approaches to uncertainty modeling in a way accessible to non-specialists. Elements of the Monte Carlo method, polynomial chaos method, Dempster-Shafer approach, fuzzy set theory, and the worst (case) scenario method are presented.

1 Introductory comments on modeling and uncertain data

Where can uncertainty analysis be placed in computational modeling? Typically, uncertainty propagation analysis leads to solving "two level" problems. This means that we can distinguish both an inner problem and an outer problem that together constitute an uncertainty analysis problem. Such a structure is not uncommon; a PDE-constrained optimization also falls into this category, for instance. Indeed, if a problem of this kind is solved by successive optimization steps, then the inner PDE is repeatedly solved during the optimization process to deliver necessary data to a constrained optimization algorithm keeping the optimized variables in an admissible set.

In engineering-oriented problems, the inner problem, commonly known as the *state problem*, represents the mathematical model of a physical phenomenon or design (imagine a temperature field in a heated body or a mechanical stress distribution in a loaded body, for example). Often, the inner problem is rather standard and even easily solvable for given unique and crisp input values such as thermal conductivity coefficients, heat capacity, intensity of heat sources, boundary condition parameters, loading forces, Young modulus, etc.

The outer problem originates from the fact that the values of input parameters are usually not known exactly. Then a question arises how uncertainty can be measured in inputs, how it propagates through the state problem, and how it can be measured in state problem outputs.

We can view uncertainty in modeling from yet another perspective. To see this, let us recall the general layout of a modeling process.

(I) A situation (phenomenon) we wish to model to get an insight and to predict the behavior we are interested in.

(II) Available input information that we can use in our models. This step is closely related to (III) next.

(III) A model chosen from a hierarchy of mathematical models. This step is crucial because it determines the rest of the modeling sequence. Choosing an adequate model can be a difficult task in which the model complexity and solvability as well as the model adequacy to our needs have to be taken into account. Modeling a beam-like body can be a good example. Models of different complexity are at one's disposal, take Bernoulli beam, Timoshenko beam, perhaps a 2D shell if one dimension of the body is significantly smaller than the others, and a full 3D model. We can, however, also consider a hierarchy of material models from linear models to nonlinear ones. In choosing a model, a compromise has to be made to end up with a model that is not too complex and computationally demanding, but can still deliver information that we need.

(IV) Validation [16, 17]. It is a process of gaining trust in the mathematical model. Roughly speaking, the mathematical formulation should be adequate to both the phenomenon that we model and the questions we wish to answer through the model. Although a theoretical analysis is always valuable and can significantly contribute to the validation process, validation is unthinkable without computational modeling, see (VI).

(V) Approximation. Once the mathematical model is defined, we find ourselves in a situation similar to (III). We have to choose a numerical method to obtain an approximate solution. Again, different or even contradictory factors should be balanced.

(VI) Verification [16, 17]. It is a process of gaining trust in the numerical method, its implementation, and its accuracy. This trust originates from various sources. At least, it is necessary to solve benchmark problems and to numerically check theoretical convergence rates. Verification is a matter of mathematics and, unlike validation, it is independent of the modeled object. On the other hand, only verified numerical models allow us to put "hands on" the mathematical model through the numerical solution that is believed[1] to be close to the exact solution of the mathematical model.

(VII) Model output and the desired information. The former may not be equal to the latter, and a post-processing may be required. It is important to obtain outputs that contain, though possibly hidden, the desired information. As a consequence, (VII) is closely related to (III).

(VIII) Interpretation of results. This step can be more demanding than it appears to be at first glance. There is the danger of misinterpretation caused by our expectations that might be seemingly confirmed by the obtained results.

The goal of modeling is to step out of the area of well-proved solutions, and make a new prediction. Then the ultimate goal is a guaranteed prediction. That is, the solution accompanied by the evaluation of the inaccuracy caused by the model selection, by the approximation method, and by other effects. Among them, the effect of uncertainties in input data is of great importance.

---

[1]This belief must not be a blind belief, it should be a well-founded belief. Unfortunately, we can never be entirely sure that our (complex) software is correct.

To close this section, let us recall two kinds of uncertainty; see [15], for example.

Epistemic uncertainty is caused by the lack of knowledge. In principle, it can (often) be reduced through improving measuring instruments as well as data collecting and mining.

Aleatory uncertainty is caused by the inherent variation associated with the modeled system. Take, for example, the randomness of material parameters, or the variability of the weather.

Consequently, our mathematical models are burdened with uncertainty in input data.

1.1 Notation, basic setting

In the sequel, we will use $\mathcal{U}_{\mathsf{ad}}$ to denote the set of values of input parameters. For example, $\mathcal{U}_{\mathsf{ad}}$ is an interval if a scalar parameter is uncertain; $\mathcal{U}_{\mathsf{ad}}$ is a subset of $\mathbb{R}^n$ if an n-tuple of real values is uncertain; $\mathcal{U}_{\mathsf{ad}}$ is a set of functions if a function is uncertain.

Next, $D(a)u = f(a)$ will stand for a state problem dependent on $a \in \mathcal{U}_{\mathsf{ad}}$ where the right-hand side f can also depend on a. Consequently, the state solution $u \equiv u(a)$ also depends on $a \in \mathcal{U}_{\mathsf{ad}}$. Examples include a boundary value problem for an ordinary or a partial differential equation dependent on $a \in \mathcal{U}_{\mathsf{ad}}$, an initial value problem dependent on $a \in \mathcal{U}_{\mathsf{ad}}$, or a variational inequality dependent on $a \in \mathcal{U}_{\mathsf{ad}}$ (then the equality symbol is inappropriate).

Finally, the quantity of interest, $\Phi(a, u(a)) \in \mathbb{R}$, will be the third ingredient of uncertainty modeling. The quantity of interest, also known as the criterion function (or criterion functional), evaluates the input data both directly and indirectly (through $u(a)$). Displacement, temperature, local mechanical stress or stress invariants, and concentration of chemicals can serve as examples.

We assume that $D(a)u = f(a)$ is uniquely solvable for each $a \in \mathcal{U}_{\mathsf{ad}}$ and that $\Psi(a) \equiv \Phi(a, u(a))$, where $a \in \mathcal{U}_{\mathsf{ad}}$, is continuous and bounded. We did not specified the set $\mathcal{U}_{\mathsf{ad}}$ but, generally speaking, we assume that $\mathcal{U}_{\mathsf{ad}}$ is a connected and compact subset of a Banach space.

2 Stochastic approaches to uncertainty

2.1 Monte Carlo method

The idea of the Monte Carlo method is quite simple. Random samples of a taken from $\mathcal{U}_{\mathsf{ad}}$ are evaluated through Ψ, the values $\Psi(a)$ are collected, and the collection is then statistically analyzed to infer probabilistic characteristics of the model behavior; see [14, 18, 21]. An advantageous feature is that the method easily allows evaluating multiple samples in parallel to speed up collecting output data.

Figure 1 shows an example. Let us assume that we are to predict the tip displacement of a loaded cantilever beam with a constant but uncertain thickness a. It is further assumed that the thickness is random with a known probability distribution. By using this distribution, we generate (pseudo)random samples (Figure 1,
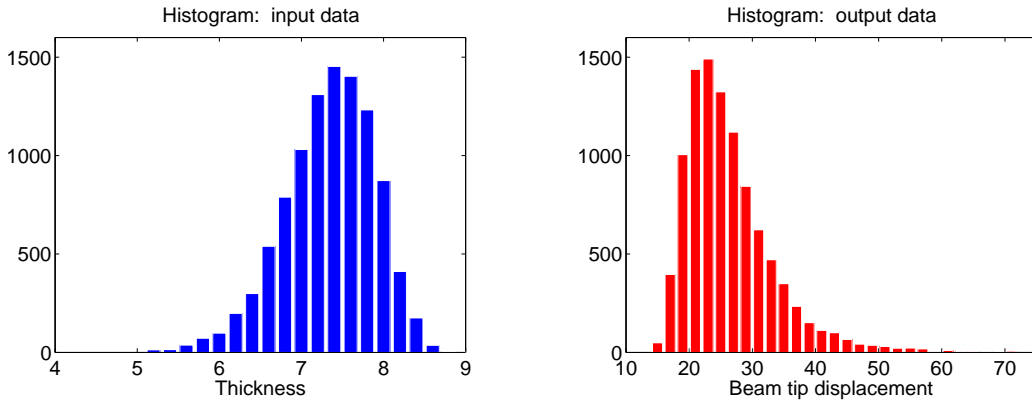
Fig. 1: *Monte Carlo method. Histogram of input data (left). Histogram of output data (right).*

left) of the thickness and calculate the beam tip displacement $\Psi(a)$ for each sample a. Then an approximation of the probabilistic behavior of $\Psi(a)$ can be inferred from the histogram of $\Psi(a)$ (Figure 1, right).

Although Monte Carlo simulation has proved fruitful and is commonly used in the modeling of uncertainty propagation through a model, one should be aware of some possible pitfalls.

The probability distribution of input data can be difficult to identify or its parameters can be uncertain.

Dependencies between input parameters are possible and, moreover, often uncertain, which complicates both the sampling procedure and the credibility of results.

If N is the number of samples, then, in general, the convergence rate of estimated probabilistic parameters is equal to $\mathcal{O}(N^{-1/2})$. Although sophisticated sampling methods can be a partial remedy, the necessary number of state solutions can still be prohibitive if the state problem is computationally demanding.

2.2 Polynomial chaos

Polynomial chaos is also known as the Wiener polynomial chaos or Hermite chaos or, in a generalized form, the Askey chaos. In computational applications, the method is also called stochastic finite elements; see [11, 12, 22].

Analogously to the Monte Carlo method, it is assumed that the state problem input parameters can be represented by a random process; let us denote it by $X(\theta)$. The idea is to express or, in calculations, to approximate the random process through separate spatial (or temporal) deterministic variables and independent random variables.

To give an example [22], let us consider a two-dimensional expansion $\sum_{i=0}^{\infty} c_i \phi_i$ using Hermite polynomials and ξ_1, ξ_2, two independent Gaussian random variables with zero mean and unit variance, that is,

$$X(\theta) = c_0 + c_1\xi_1 + c_2\xi_2 + c_3(\xi_1^2 - 1) + c_4(\xi_1\xi_2) + c_5(\xi_2^2 - 1) + \dots$$

$$= c_0 + c_1\phi_1 + c_2\phi_2 + c_3\phi_3 + c_4\phi_4 + c_5\phi_5 \cdots = \sum_{i=0}^{\infty} c_i\phi_i,$$

where $c_i \in \mathbb{R}$ for $i = 0, 1, \dots$.

Next, a weight w related to the used random variables is introduced to make the system $\{\phi_i\}$ w-orthogonal. For $\xi = (\xi_1, \xi_2)$ where ξ_1 and ξ_2 are Gaussian random variables, we take

$$w(\xi) = \frac{1}{2\pi}e^{-(\xi_1^2 + \xi_2^2)/2},$$

which simplifies the weighted inner product $\langle \cdot, \cdot \rangle_w$ of ϕ_i, ϕ_j, i.e.,

$$\langle \phi_i, \phi_j \rangle_w = \int_{S_w} \phi_i(\xi)\phi_j(\xi)w(\xi)\,\mathrm{d}\xi = \langle \phi_i, \phi_j \rangle_w \delta_{ij} \text{ (Kronecker } \delta),$$

where S_w is the domain of w; it is $S_w = \mathbb{R}^2$ for our choice of ξ.

A finite part of the expansion of random inputs is employed in the state problem to obtain its approximate solution as a random process. Let us illustrate this through an example taken from [22].

Let us consider the following initial value problem

$$y'(t) = -ky(t), \quad y(0) = y_\star, \tag{1}$$

where $k \equiv k(\theta)$ is a random variable with probability density function f.

Remark: For (1) and a given $k \in \mathbb{R}$, the deterministic solution $y(t) = y_\star e^{-kt}$ allows a direct probabilistic characterization of relevant quantities without the use of a polynomial chaos expansion. Take, for instance, the mean of the stochastic solution at t

$$\overline{y}(t) = y_\star \int_S e^{-kt} f(k)\,\mathrm{d}k,$$

where S is the support of f. For educational purposes, however, we will not use the approach based on the deterministic solution but we will apply the polynomial chaos expansion to (1). $\square$

Let us assume that k is a Gaussian random variable. Then it is recommended to expand the random variables into functions originating from Hermite polynomials applied to Gaussian random variables $\xi_1, \dots, \xi_n$ ($n = 2$ in our example). The solution y of (1) is a random process $y(t, \theta)$ at any $t \in \mathbb{R} \setminus \{0\}$. We approximate both $y(t, \theta)$ and $k(\theta)$ by finite parts of their polynomial chaos expansions, i.e.,

$$y(t, \theta) \approx \widehat{y}(t, \xi) = \sum_{j=0}^{P} y_j(t)\phi_j(\xi), \quad k(\theta) \approx \widehat{k}(\xi) = \sum_{i=0}^{P} k_i\phi_i(\xi),$$

where y_j are unknown functions of t, $k_i \in \mathbb{R}$ are unknown constants, and ϕ_i are known w-orthogonal functions. By inserting these finite sums into $y'(t) = -ky(t)$, we arrive at

$$\sum_{i=0}^{P} y_i'(t)\phi_i = -\sum_{i=0}^{P}\sum_{j=0}^{P} \phi_i\phi_j k_i y_j(t).$$

After multiplying by ϕ_ℓ ($\ell = 0, 1, \dots, P$), integrating with the weight w, and exploiting the orthogonality, we obtain

$$\langle \phi_\ell, \phi_\ell \rangle_w y_\ell'(t) = -\sum_{i=0}^{P}\sum_{j=0}^{P} e_{ij\ell} k_i y_j(t), \quad \ell = 0, 1, \dots, P, \tag{2}$$

where $e_{ij\ell} = \int_{S_w} \phi_i(\xi)\phi_j(\xi)\phi_\ell(\xi)w(\xi)\,\mathrm{d}\xi$ and S_w is the domain of w.

We observe that the original stochastic problem (1) is transformed into a system of numerically solvable (deterministic) ordinary differential equations (2).

By having $y(t, \theta) \approx \widehat{y}(t, \xi) = \sum_{j=0}^{P} y_j(t)\phi_j(\xi)$, the orthogonality, and $\phi_0 = 1$, we can approximate the mean solution at t by

$$E[\widehat{y}(t,\xi)] = \overline{\widehat{y}(t,\xi)} = \int_{S_w} \sum_{j=0}^{P} y_j(t)\phi_j(\xi)1w(\xi)\,\mathrm{d}\xi = \langle \phi_0, \phi_0 \rangle_w y_0(t)$$

and the variance of the solution at t by

$$\begin{aligned}
\mathrm{Var}\,(\widehat{y}(t,\xi)) &= E\left[\left(\widehat{y}(t,\xi) - \overline{\widehat{y}(t,\xi)}\right)^2\right] \\
&= \int_{S_w} \left(\sum_{j=0}^{P} y_j(t)\phi_j(\xi) - \langle \phi_0, \phi_0 \rangle_w y_0(t)\right)^2 w(\xi)\,\mathrm{d}\xi \\
&= \int_{S_w} \sum_{i=0}^{P}\sum_{j=0}^{P} y_i(t)y_j(t)\phi_i(\xi)\phi_j(\xi)w(\xi)\,\mathrm{d}\xi \\
&\quad - 2\langle \phi_0, \phi_0 \rangle_w y_0(t) \int_{S_w} \sum_{j=0}^{P} y_j(t)\phi_j(\xi)w(\xi)\,\mathrm{d}\xi + \langle \phi_0, \phi_0 \rangle_w^2 y_0^2(t) \\
&= \sum_{i=0}^{P} \langle \phi_i, \phi_i \rangle_w y_i^2(t) - 2\langle \phi_0, \phi_0 \rangle_w^2 y_0^2(t) + \langle \phi_0, \phi_0 \rangle_w^2 y_0^2(t) \\
&= \sum_{i=1}^{P} \langle \phi_i, \phi_i \rangle_w y_i^2(t).
\end{aligned}$$

Although the complexity of the method increases with the number of random variables and, consequently, with the complexity of the expansion, the method can be

$10^2 \times - 10^4 \times$ faster than the Monte Carlo method delivering the required probabilistic characteristics with the same accuracy; see [22].

Other probability density functions can be considered in input parameters. Relevant weights w and w-orthogonal polynomials are then used in the analysis; see [22].

The approximate stochastic solution $\widehat{y}$ can be further evaluated by the criterion functional whose probability characteristics are to be inferred.

2.3 Transformation to a deterministic problem; Karhunen-Loève expansion (KLE)

The underlying idea is identical to the idea presented in the previous subsection. However, unlike the polynomial chaos expansion, which is rather heuristic, the Karhunen-Loève expansion is mathematically more rigorous. More details on the transformation can be found in [2, 3, 4, 5].

Let $\Omega \subset R^d$, where $d \in \{1, 2, 3\}$, be a domain. Let us consider g, a stochastic function on Ω. Values $g(s)$ and $g(t)$, where $s, t \in \Omega$ and $s \neq t$, are random variables that can be coupled to some extent. In the probability theory, such non-deterministic couplings are characterized by the covariance function $\text{cov}[g] : \Omega \times \Omega \to \mathbb{R}$ that is defined as follows

$$\text{cov}[g](s, t) = E\left[(g(s) - E[g(s)])\,(g(t) - E[g(t)])\right], \quad s, t \in \Omega,$$

where $E[\omega]$ stands for the mean (expected value) of a random quantity ω. We assume that $\text{cov}[g]$ is continuous and bounded on $\Omega \times \Omega$.

We define an operator $T_g : L^2(\Omega) \to L^2(\Omega)$ by

$$T_g v(\cdot) = \int_\Omega \text{cov}[g](x, .)v(x)\,dx \quad \forall v \in L^2(\Omega).$$

It can be shown that the operator is compact, selfadjoint, and that its eigenvalues are non-negative. Let $\{\lambda_i\}_{i=1}^\infty$ be a non-increasing sequence of the eigenvalues of T_g and let $\{b_i\}_{i=1}^\infty$ be a sequence of the corresponding $L^2(\Omega)$-orthonormal eigenfunctions, i.e., $T_g b_i = \lambda_i b_i$.

The stochastic function $g(x, \theta)$, where $x \in \Omega$, can be expressed by the Karhunen-Loève expansion from which, however, we take only the first N terms to approximate g by g_N, that is,

$$g_N(x, \theta) = E[g](x) + \sum_{i=1}^N \sqrt{\lambda_i}\,b_i(x)Y_i(\theta), \tag{3}$$

where the real random variables, $\{Y_i\}_{i=1}^N$, are uncorrelated, have zero mean and unit variance, i.e., $E[Y_i] = 0$ and $E[Y_i Y_j] = \delta_{ij}$.

The amplitude of λ_i can provide guidance for choosing N (if λ_i is "small", we cut off the rest of the infinite expansion) and, analogously, for determining the size of the polynomial chaos expansion introduced in Subsection 2.2.

48

Let us illustrate the above ideas through a boundary value problem with stochastic functions; see [5].

Let $\Omega \subset \mathbb{R}^d$ be a domain. We consider the following elliptic boundary value problem

$$-\operatorname{div}(a(x,\theta)\nabla u(x,\theta)) = f(x,\theta) \text{ in } \Omega, \text{ i.e, } x \in \Omega, \tag{4}$$

$$u(x,\theta) = 0 \text{ on } \partial\Omega, \tag{5}$$

where a, f (and u) are stochastic functions.

We assume that
a) $a(x,\theta) = a(x, Y_1(\theta), \ldots, Y_N(\theta))$, where Y_i are the functions introduced in (3);
b) $f(x,\theta) = f(x, Y_1(\theta), \ldots, Y_N(\theta))$;
c) Γ_i, the range of Y_i, is a bounded interval in $\mathbb{R}$ for $i = 1, 2, \ldots, N$;
d) the random variable Y_i has a known density function $\rho_i : \Gamma_i \to \mathbb{R}^+$ (nonnegative real numbers) with $\rho_i \in L^\infty(\Gamma_i)$, where $i = 1, 2, \ldots, N$.

As a consequence, $u(x,\theta) = u(x, Y_1(\theta), \ldots, Y_N(\theta))$.

Let $\rho = \rho(y) : \Gamma \to \mathbb{R}^+$ be the joint probability density function of the random vector $Y = (Y_1, \ldots, Y_N)$, where $\Gamma = \prod_{i=1}^N \Gamma_i \subset \mathbb{R}^N$.

We also assume that quantities Y_i are not only uncorrelated but mutually independent. Consequently, $\rho(y) = \prod_{i=1}^N \rho_i(y_i)$, where $y = (y_1, \ldots, y_N)$.

Like a deterministic boundary value problem, the stochastic problem (4)-(5) has its stochastic variational counterpart, see [5]. Although it is omitted here, we introduce its deterministic equivalent: Find $u \in H_0^1(\Omega) \otimes L_\rho^2(\Gamma)$ such that

$$\int_\Gamma \rho\,(a\nabla_x u, \nabla_x v)_{[L^2(\Omega)]^d}\,\mathrm{d}y = \int_\Gamma \rho\,(f, v)_{L^2(\Omega)}\,\mathrm{d}y \quad \forall v \in H_0^1(\Omega) \otimes L_\rho^2(\Gamma), \tag{6}$$

where ∇_x indicates that only the partial derivatives with respect to the spatial variables are included into the gradient, $H_0^1(\Omega)$ stands for the Sobolev space of once differentiable functions with traces vanishing on $\partial\Omega$,

$$L_\rho^2(\Gamma) = \left\{ v : \Gamma \to \mathbb{R} \mid \int_\Gamma \rho(y) v^2(y)\,\mathrm{d}y < +\infty \right\},$$

and the tensor space $H_0^1(\Omega) \otimes L_\rho^2(\Gamma)$ is a Hilbert space with the inner product defined as follows

$$(u, v)_{H_0^1(\Omega)\otimes L_\rho^2(\Gamma)} = \int_\Gamma \rho(y)(u(\cdot, y), v(\cdot, y))_{H^1(\Omega)}\,\mathrm{d}y.$$

Problem (6) is purely deterministic because the stochastic features of (4) have been transformed into the weight ρ. However, we pay for it by the increased number of dimensions. In (6), functions a, f, u, and v are functions of $d + N$ variables.

Let us confine ourselves to a few comments on solving multidimensional boundary value problems though this subject would deserve a more detailed treatment.

In Monte Carlo Galerkin finite element method (FEM) [5], samples of the state solution u are obtained via realizations of a and f. For each realization, i.e., $a(\cdot, y_1, \ldots, y_N)$, $f(\cdot, y_1, \ldots, y_N)$, where $(y_1, \ldots, y_N)$ is fixed, the state solution is obtained through a standard Galerkin FEM. The samples are weighted by their probability[2] and the expected value (i.e., mean) of the stochastic state solution can be calculated, for example.

FEM methods for multidimensional BVPs have been designed. They use, for instance, h-FEM basis functions in the spatial variable (d-tuple) x and p-FEM basis functions in the probabilistic variable (n-tuple) y; special polynomials to approximate the probabilistic part of u and compute its mean efficiently; a reduced set of basis functions. It is said that problems exhibiting $\approx 10 - 20$ dimensions, or ≈ 100 dimensions in special cases, are solvable by present means.

Although the KLE-based methods are anchored in a rigorous mathematical analysis and can deliver theoretical results as well as error estimates, one should be aware that the KLE and ρ_i identification are crucial and demanding prerequisites.

Remark: We omit a huge class of differential equations perturbed by (white) noise (Brownian motion, Wiener process). They are known as stochastic differential equations and a special (Itô) calculus has been proposed for their analysis. □

We conclude the section on stochastic methods by the observation that these methods can deliver extremely important and valuable assessments of uncertainty and its propagation to model outputs. To perform well, however, they need input data that is sometimes (if not even often) difficult to obtain in the required quality and/or quantity.

3 Non-stochastic methods

Stochastic methods can also be viewed as methods where input values are weighted (by their probability) and the goal is to infer the weights of model output values. Since the weights have to fulfil rather strong assumptions (recall the requirements placed on probability measures), it can be advantageous to have methods whose assumptions about input data are relaxed.

We will present two such methods and one method where inputs are not weighted.

3.1 The worst (case) scenario method (WSM)

In this approach, input parameter values are considered equally possible and of the same weight. The name of the method reflects a common goal in practice — to find the particular value of a quantity of interest that is most unfavorable from an application point of view; see [13].

Since maxima are often important (maximum temperature, maximum mechanical stress, for example), to determine the "worst" scenario (also known as anti-optimization [6, 10]), we maximize Ψ by searching for

---

[2]In Monte Carlo, it is common to divide the probability domain into subdomains of equal probability and to take a representative sample from each subdomain.

$$a^0 = \arg\max_{a \in \mathcal{U}_{\mathsf{ad}}} \Psi(a). \tag{7}$$

If also the "best" scenario

$$a_0 = \arg\min_{a \in \mathcal{U}_{\mathsf{ad}}} \Psi(a) = \arg\max_{a \in \mathcal{U}_{\mathsf{ad}}} (-\Psi(a)) \tag{8}$$

is found, then the range of $\Psi|_{\mathcal{U}_{\mathsf{ad}}}$ is given by

$$I_\Psi = [\Psi(a_0), \Psi(a^0)]. \tag{9}$$

To obtain (7)-(9), we assume that $\mathcal{U}_{\mathsf{ad}}$ is a compact and convex set and that $\Psi : \mathcal{U}_{\mathsf{ad}} \to \mathbb{R}$ is a continuous map. In practice, non-convex $\mathcal{U}_{\mathsf{ad}}$ makes maximization (minimization) of Ψ difficult.

Remark: The formulation of the worst scenario problem (7) (or its modification (8)) is identical to the formulation of other already established problems; take optimization problems or parameter identification problems. Indeed, for the latter, where a desirable output u_{given} is known on a domain Ω, the goal could be to minimize

$$\Psi(a) = \int_\Omega (u(a) - u_{\mathsf{given}})^2 \, \mathrm{d}x,$$

over $\mathcal{U}_{\mathsf{ad}}$, for instance. □

The method delivers the guaranteed range of $\Psi|_{\mathcal{U}_{\mathsf{ad}}}$, which fits to the concept of the guaranteed prediction, see Section 1. The WSM does not use weighted data, which eliminates the difficulties associated with the determination of weights. On the other hand, the method neglects the fact that the occurrence of extremal values of Ψ is rare in many practical problems.

Remark: The worst scenario approach is also possible in stochastic problems. Take $\mathcal{U}_{\mathsf{ad}}$ representing a set of admissible cumulative distribution functions, for instance. The extremal probabilistic features of Ψ are then in the focus of the analysis; see [7].

3.2 Dempster-Shafer evidence theory

In this approach, the entire sets of input data are weighted. The weights, though resembling probability measures, are defined less strictly, however.

Let us introduce the elements of the evidence theory; a more detailed treatment can be found in [1, 6, 8, 20].

Let X be a universal set and P_X be the power set of X. A map $m : P_X \to [0, 1]$, called the basic probability assignment, is defined. It holds $m(\emptyset) = 0$ and $\sum_{A \in P_X} m(A) = 1$. For simplicity, we assume that $m(A_i) > 0$ only for a finite number of sets $A_i \in P_X$, where $i = 1, 2, \ldots, k$; these sets A_i are called *focal elements*. We can interpret $m(A_i)$ as the weight associated with A_i; see Figure 2 (left).
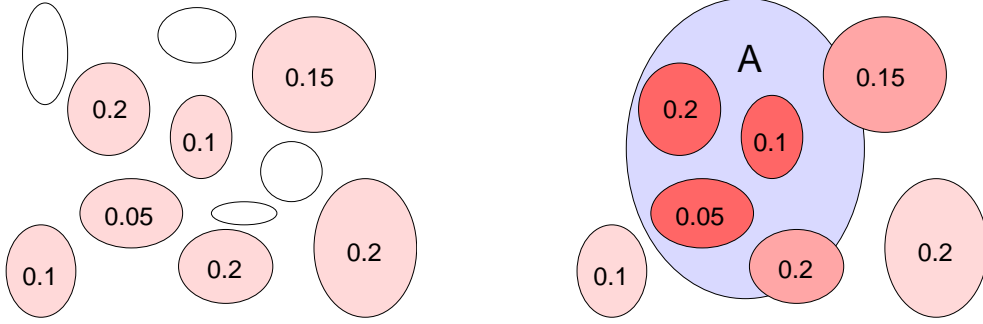
Fig. 2: *Focal elements and their weights (left). Focal elements and a set A (right).*

Next, two functions are defined on P_X, namely *belief* and *plausibility*

$$\text{Bel}(A) = \sum_{A_i \subseteq A} m(A_i), \quad \text{Pl}(A) = \sum_{A_i \cap A \neq \emptyset} m(A_i), \quad A \in P_X. \tag{10}$$

For the set A and the focal elements depicted in Figure 2, we obtain

$$\text{Bel}(A) = 0.2 + 0.1 + 0.05 = 0.35, \quad \text{Pl}(A) = 0.2 + 0.1 + 0.05 + 0.15 + 0.2 = 0.7.$$

Various interpretations of Bel and Pl can be found in [15]: $\text{Pl}(A)$ is the largest probability for A that is consistent with all available evidence, $\text{Bel}(A)$ is the smallest probability for A that is consistent with all available evidence. Or, Pl is an upper limit and Bel is a lower limit on the strength of evidence at hand.

The latter interpretation is close to our weight-oriented perspective. Indeed, we can interpret $\text{Pl}(A)$ ($\text{Bel}(A)$) as an upper (lower) weight A can have, given the weights of focal elements.

Our ultimate goal is to infer weights attributed to (at least) some sets of values produced by the criterion functional Ψ. In other words, we have to establish a set of focal elements in the space of the quantity of interest. To achieve this goal, we employ the worst scenario method.

Let A_i, where $i = 1, \ldots, N$, be the focal elements of a probability assignment m in the space of model inputs. Let us interpret each A_i as an admissible set and calculate the intervals $\Psi|_{A_i}$, where $i = 1, \ldots, N$; see (7)-(9).

These intervals can serve as focal elements, but we have to take into consideration that more than one admissible set can be mapped to one interval. For the quantity of interest, basic probability assignment m_Ψ and its focal elements I_Ψ^k are defined as follows (extension principle)

$$m_\Psi(I_\Psi^k) = \sum_{\{i:\ I_\Psi^k = \Psi|_{A_i}\}} m(A_i), \quad k = 1, \ldots, K, \tag{11}$$

where K is the total number of different intervals $\Psi|_{A_i}$.

Once m_Ψ, the basic probability assignment in the space of the quantity of interest, is established, the relationship between the focal elements I_Ψ^k and various sets can be assessed through Bel and Pl.

Example: Let us imagine that some input parameters of a computational model of a structure are not given as crisp numbers but they are known only through inexact measurements performed by four groups of students. The measurements have resulted in four data sets denoted by $A_1, \ldots, A_4$. Since the groups do not share the same level of experience, the credibility of their results is also different, which is represented by weights m_i that we attribute to A_i, where $i = 1, 2, 3, 4$.

The range of a quantity of interest Ψ is calculated for each A_i; let

$$\Psi|_{A_1} = [5, 7.5], \quad \Psi|_{A_2} = [6, 10], \quad \Psi|_{A_3} = [6, 10], \quad \Psi|_{A_4} = [9.5, 12]. \tag{12}$$

By applying (11) to (12), we obtain three focal elements and their probability assignment, i.e.,

$$I_\Psi^1 = [5, 7.5], \quad I_\Psi^2 = [6, 10], \quad I_\Psi^3 = [9.5, 12]; \tag{13}$$
$$m_\Psi(I_\Psi^1) = m_1, \quad m_\Psi(I_\Psi^2) = m_2 + m_3, \quad m_\Psi(I_\Psi^3) = m_4. \tag{14}$$

To get some insight into the behavior of the quantity of interest, we will use m_Ψ to calculate Bel and Pl for various intervals of length 3.

To this end, let us consider $x \in [1, 13]$ and define two functions

$$f_{\mathrm{Bel}}(x) = \mathrm{Bel}([x, x+3]), \quad f_{\mathrm{Pl}}(x) = \mathrm{Pl}([x, x+3]).$$

It is, for instance, $f_{\mathrm{Pl}}(4.9) = \mathrm{Pl}([4.9, 7.9]) = m_\Psi(I_\Psi^1) + m_\Psi(I_\Psi^2) = m_1 + m_2 + m_3$.

In Figure 3 ($m_1 = 0.2$, $m_2 = 0.1$, $m_3 = 0.25$, and $m_4 = 0.45$), two clusters of intervals attract our attention. We observe that intervals $[x, x+3]$ determined by $x \in [6.5, 7.5]$ have plausibility equal to 1 but zero belief and that intervals $[x, x+3]$
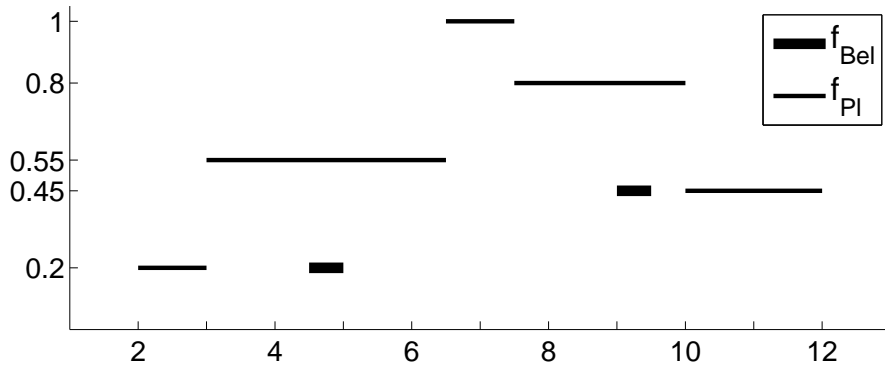

Fig. 3: *The graphs of f_{Bel} and f_{Pl}.*

determined by $x \in [9, 9.5]$ have plausibility 0.8 and their belief value is equal to 0.45. The intervals from the former set can be in agreement with the quantity of interest "hidden" in the intervals I_Ψ^i but no agreement is guaranteed. The intervals from the latter set cannot comply with all I_Ψ^i but at least partial correspondence is guaranteed. $\square$

It is worth noting that the Dempster-Shafer theory includes rules of combination that allow to combine two different probability assignments. In other words, different opinions of different experts can be combined into one assessment; see [1, 19], for instance. Although fruitful in uncertainty analysis, the combination fails if the included opinions are too different [19].

3.3 Fuzzy set theory

A fuzzy set $U \subset Z$, where Z is a basic set, is identified through μ_U, the *membership function*

$$\mu_U : Z \to [0, 1],$$

where the real value in $[0, 1]$ represents the degree to which $z \in Z$ belongs to the set U. The higher the value, the stronger the membership. For the theory and applications, see [1, 6, 9, 23].

For our purposes and in accordance with our weight-oriented standpoint, we assume that an admissible set $\mathcal{U}_{\mathsf{ad}}$ is given together with its membership function $\mu_{\mathcal{U}_{\mathsf{ad}}} : \mathcal{U}_{\mathsf{ad}} \to [0, 1]$ and that if $\mu_{\mathcal{U}_{\mathsf{ad}}}(x) = 0$, then $x \in \partial \mathcal{U}_{\mathsf{ad}}$ where $\partial \mathcal{U}_{\mathsf{ad}}$ is the boundary of $\mathcal{U}_{\mathsf{ad}}$. As in the worst scenario section, it is advantageous to assume that $\mathcal{U}_{\mathsf{ad}}$ is a compact and convex subset of a Banach space. We assume, moreover, that $\mu_{\mathcal{U}_{\mathsf{ad}}}$ is a continuous and concave function on $\mathcal{U}_{\mathsf{ad}}$.

For $\alpha \in [0, 1]$, a subset $^\alpha\mathcal{U}_{\mathsf{ad}}$ comprising all $x \in \mathcal{U}_{\mathsf{ad}}$ such that $\mu_{\mathcal{U}_{\mathsf{ad}}}(x) \geq \alpha$ is called α-cut. Obviously $\mathcal{U}_{\mathsf{ad}} \equiv {}^0\mathcal{U}_{\mathsf{ad}}$.

By knowing the α-cuts of a fuzzy set, we are able to restore its membership function, which is the goal we want to achieve in the set $I_\Psi = \{\Psi(a) | \ a \in \mathcal{U}_{\mathsf{ad}}\}$ that is fuzzy due to the fuzziness of $\mathcal{U}_{\mathsf{ad}}$.

To this end, we temporarily fix $\alpha \in [0, 1]$ and, by employing the best and the worst scenario (see (7)-(9)), we determine

$$^\alpha I_\Psi = [\Psi(a_{0,\alpha}), \Psi(a^{0,\alpha})], \tag{15}$$

where $^\alpha I_\Psi$ is the α-cut of $I_\Psi \equiv {}^0 I_\Psi$.

By computing (15) for all $\alpha \in [0, 1]$, we can construct the membership function μ_{I_Ψ} through

$$\mu_{I_\Psi}(y) = \max\{\alpha | \ y \in {}^\alpha I_\Psi\}, \quad y \in I_\Psi,$$

to asses the fuzziness of Ψ, the quantity of interest. In practice, of course, $^\alpha I_\Psi$ is found only for a finite number of input α-cuts and only an approximation of μ_{I_Ψ} is obtained.

54

4 Concluding remarks

Uncertainty propagation analysis fits into the framework of decision analysis. The analyst has a variety of approaches at his or her disposal. Their applicability depends on what is known about input data, what kind of uncertainty is most relevant to the available data and to the problem under investigation.

In practice, the weights attributed to input data are often uncertain because they originate from a limited number of (inaccurate) measurements, expert opinions, hypotheses, and estimates. As a consequence, the validity of uncertainty analysis results is a delicate matter.

Although ignored in this paper, sensitivity analysis is an important part of uncertainty analysis. The goal of sensitivity analysis is to assess the influence of a (small) change in inputs on the quantity of interest. It helps to identify input parameters that have a weak influence on the value of the quantity of interest; these parameters can be excluded from the uncertainty analysis. Sensitivity analysis is also beneficial in algorithms searching for the minimum or maximum of a quantity of interest.

The worst scenario method can be used as a stand-alone method. It is also an integral part of other methods as we observed in the sections on the Dempster-Shafer theory and fuzzy set theory. Moreover, the WSM shares many features with optimal shape design, PDE constrained optimization, and inverse problems, for instance. As a consequence, various well-tried tools for theoretical as well as computational analysis are at our disposal. However, their tailoring for uncertain input data problems is desirable.

References

[1] Ayyub, B.M. and Klir, G.J.: *Uncertainty modeling and analysis in engineering and the sciences.* Chapman & Hall/CRC, Taylor & Francis Group, Boca Raton, 2006.

[2] Babuška, I. and Chatzipantelidis, P.: On solving elliptic stochastic partial differential equations. Comput. Methods Appl. Mech. Engrg. **191** (2002), 4093–4122.

[3] Babuška, I. and Liu, K.M.: On solving stochastic initial-value differential equations. Math. Models Methods Appl. Sci. **13** (2003), 715–745.

[4] Babuška, I., Liu, K.M., and Tempone, R.: Solving stochastic partial differential equations based on the experimental data. Math. Models Methods Appl. Sci. **13** (2003), 415–444.

[5] Babuška, I., Tempone, R., and Zouraris, G.E.: Solving elliptic boundary value problems with uncertain coefficients by the finite element method: The stochastic formulation. Comput. Methods Appl. Mech. Engrg. **194** (2005), 1251–1294.

[6] Bernardini, A.: What are the random and fuzzy sets and how to use them for uncertainty modelling in engineering systems? In: I. Elishakoff (Ed.), *Whys and hows in uncertainty modelling, probability, fuzziness and anti-optimization.* CISM Courses and Lectures No. 388, Springer–Verlag, Wien, New York, 1999.

[7] Chleboun, J.: An approach to the Sandia workshop static frame challenge problem: A combination of elementary probabilistic, fuzzy set, and worst scenario tools. Comput. Methods Appl. Mech. Eng. **197** (2008), 2500–2516.

[8] Dempster, A.P.: Upper and lower probabilities induced by a multivalued mapping. Ann. Math. Stat. **38** (1967), 325–339.

[9] Dubois, D. and Prade, H. (Eds.): *Fundamentals of fuzzy sets. Foreword by Lotfi Zadeh, The Handbooks of Fuzzy Sets Series*, vol. 7. Kluwer Academic Publishers, Dordrecht, 2000.

[10] Elishakoff, I.: An idea of the uncertainty triangle. Shock Vib. Dig. **22** (1990), 1.

[11] Ghanem, R. and Red-Horse, J.: Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element approach. Physica D **133** (1999), 137–144.

[12] Ghanem, R.G. and Spanos, P.D.: *Stochastic finite elements: A spectral approach.* Springer–Verlag, New York, 1991.

[13] Hlaváček, I., Chleboun, J., and Babuška, I.: *Uncertain input data problems and the worst scenario method, North-Holland Series in Applied Mathematics and Mechanics*, vol. 46. Elsevier, Amsterdam, 2004.

[14] Madras, N.: *Lectures on Monte Carlo methods, Fields Institute Monographs*, vol. 16. American Mathematical Society, Providence, RI, 2002.

[15] Oberkampf, W.L., Helton, J.C., and Sentz, K.: Mathematical representation of uncertainty. Research Article AIAA 2001-1645, American Institute of Aeronautics and Astronautics, Reston, VA, 2001.

[16] Oberkampf, W.L. and Trucano, T.G.: Verification and validation in computational fluid dynamics. Progress in Aerospace Sciences **38** (2002), 209–272.

[17] Roache, P.J.: *Verification and validation in computational science and engineering.* Hermosa, Albuquerque, 1998.

[18] Rubinstein, R.Y. and Kroese, D.P.: *Simulation and the Monte Carlo method.* John Wiley & Sons, Hoboken, 2008, second edition.

[19] Sentz, K. and Ferson, S.: Combination of evidence in Dempster-Shafer theory. Sandia Report SAND2002-0835, Sandia National Laboratories, Albuquerque, New Mexico, 2002.

[20] Shafer, G.: *A mathematical theory of evidence.* Princeton University Press, Princeton, NJ, 1976.

[21] Sobol′, I.M.: *A primer for the Monte Carlo method.* CRC Press, Boca Raton, FL, 1994.

[22] Xiu, D. and Karniadakis, G.E.: The Weiner-Askey polynomial chaos for stochastic differential equations. Tech. Rep., Brown University, Providence, RI, 2003. Available on the Internet.

[23] Zimmermann, H.J.: *Fuzzy set theory — and its applications.* Kluwer Academic Publishers, Boston, 2001, fourth edition.