Modeling and Quantification of Physical Systems Uncertainties in a Probabilistic Framework

Americo Cunha Jr.

Abstract Uncertainty quantification (UQ) is a multidisciplinary area, that deals with quantitative characterization and reduction of uncertainties in applications. It is essential to certify the quality of numerical and experimental analyses of physical systems. The present manuscript aims to provide the reader with an introductory view about modeling and quantification of uncertainties in physical systems. In this sense, the text presents some fundamental concepts in UQ, a brief review of probability basics notions, discusses, through a simplistic example, the fundamental aspects of probabilistic modeling of uncertainties in a physical system, and explains what is the uncertainty propagation problem.

Keywords Uncertainty quantification · Stochastic modeling of uncertainties · Probabilistic approach

1 An Introductory Overview on UQ

Typically, highly complex engineering projects use both numerical simulations and experimental tests on prototypes to specify a certain system or component with desired characteristics. These two tools are used in a similar way by scientists to investigate physical phenomena of interest. However, none of these approaches provides a response that is an exact reproduction of the physical system behaviour, because computational model and test rig are subject to uncertainties, which are intrinsic to modeling process (lack of knowledge on the physics) and model parameters (measurement inaccuracies, manufacturing variabilities, etc.).

In order to improve the reliability level of numerical results and experimental data, it is necessary to quantify the underlying uncertainties. The cautious experimentalists have been doing this for many decades, leading to a high level competence in what concerns the specification of the level of uncertainty in an experiment. It is worth

A. Cunha Jr. (✉)
NUMERICO - Nucleus of Modeling and Experimentation with Computers,
Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
e-mail: americo@ime.uerj.br

remembering that an experiment that does not specify the level of uncertainty is not well seen by the technical/scientific community. On the other hand, just recently the numerical community has begun to pay attention to the need of specifing the level of confidence for computer simulations.

Uncertainty quantification (UQ) is a multidisciplinary area that deals with quantitative characterization and the reduction of uncertainties in applications. One reason that UQ has gained such popularity over the last years, in numerical world, is due to several books on the subject have recently emerged [1–12]. To motivate its study, we present three important scenarios where UQ is an essential tool:

Decision making: Some risk decisions, which negative result can cause catastrophic failure or huge financial costs, need to be well analysed before a final opinion by the responsible party. The possible variabilities that generate uncertain scenarios need to be taken into account in the analysis. The evaluation of these uncertain scenarios has the task of assisting the responsible party to minimize the chances of a wrong decision. Briefly, and in this context, UQ is essential to provide the necessary certification for a risk decision.

Model validation: Experimental data are widely used to check the accuracy of a computational model which is used to emulate a real system. Although this procedure is already being used by scientists and engineers for many decades, there is still no universally accepted criteria to ensure the model quality. However, it is known that any robust criteria of model validation must take into account the simulation and experiment uncertainties.

Robust design: An increasingly frequent requirement in several projects is the robust design of a component which consists make a specific device low sensitive to variation on its properties. This requires the quantification of model and parameters uncertainties.

In a very simplistic way, we can summarize UQ objectives as (i) *add error bars to experiments and simulations*, and (ii) *define a precise notion of the validated model*.

The first objective is illustrated in Fig. 1a, which shows the comparison between a simulation result with experimental data, and in Fig. 1b, that presents the previous graph with the inclusion of an envelope of reliability around the simulation. As careful experimentalists, which use error bars for a long time, UQ mainly focuses on "error bars for simulations".

Moreover, a possible notion of validated model is illustrated in Fig. 2, where experiment and simulation are compared, and the computational model is considered acceptable if the admissible range for the experimental value (defined by the point and its error bar) is contained within the reliability envelope around the simulation.

This chapter is organised into six sections. Besides this introduction, there is a presentation of some fundamental concepts of UQ in Sect. 2; a brief review on probability theory basics in Sect. 3; an exposure of the fundamental aspects of probabilistic modeling of uncertainties, through a simplistic example, in Sect. 4; the presentation of the uncertainty propagation problem in Sect. 5; and the final remarks in Sect. 6.
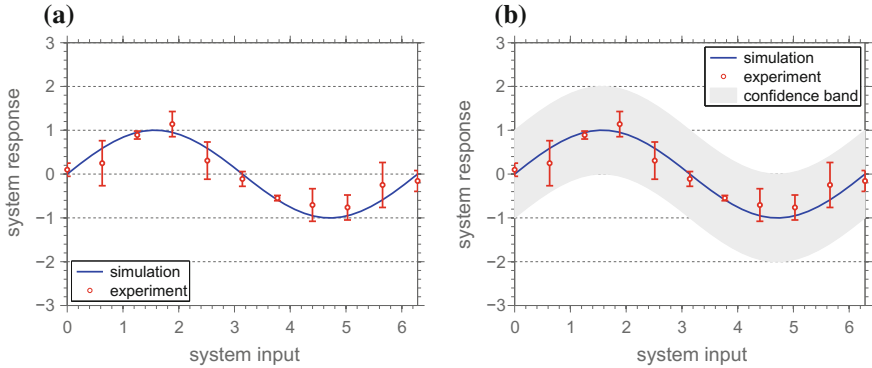
Fig. 1 **a** Comparison between simulation and experimental data, without an envelope of reliability for the simulation, and **b** including this envelope
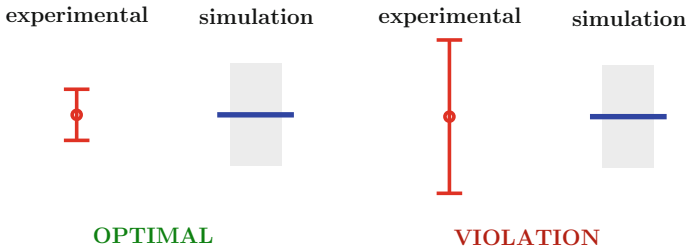
**Fig. 2** Illustration of a possible notion of validated model

It is noteworthy that many of the ideas that are presented in this manuscript are very influenced by courses taught by the author's doctoral supervisor, Prof. Christian Soize [13–15]. Lectures of Prof. Gianluca Iaccarino, Prof. Alireza Doostan, and collaborators were also very inspiring [16–18].

2 Some Fundamental Concepts on UQ

This section introduce some fundamental notions in the context of UQ.

2.1 Errors and Uncertainties

Unfortunately, until the present date, there is still no consensus in UQ literature about the notions of errors and uncertainties. This manuscript presents the definitions we think make more sense, introduced by [19].
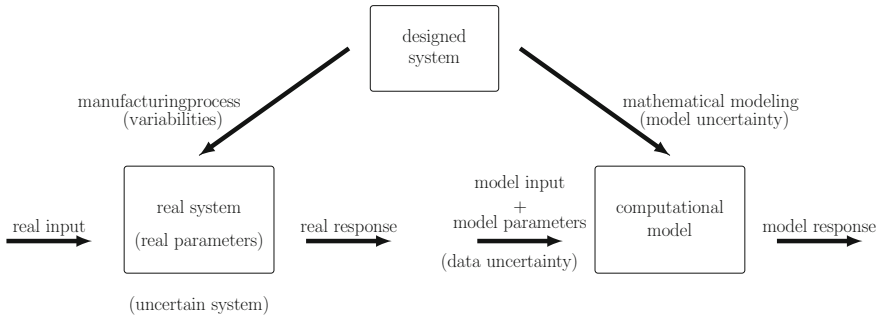
Fig. 3 Schematic representation of the relationship between the designed system, the real system and the computational model [19]

Let's start with three conceptual ideas that will be relevant to the stochastic modeling of physical systems: *designed system*, *real system* and *computational model*. A schematic illustration of these concepts is shown in Fig. 3.

Designed system: The designed system consists of an idealized project for a physical system. It is defined by the shape and geometric dimensions, material properties, connection types between components (boundary conditions), and many other parameters. This ideal system can be as simple as a beam or as complex as an aircraft [19].

Real system: The real system is constructed through a manufacturing process taking the designed system as reference. In contrast to the designed system, the real system is never known exactly, as the manufacturing process introduces some variabilities in the system geometric dimensions, on its materials properties, etc. No matter how controlled the construction process is, these deviations from the conceptual project are impossible to eliminate, since any manufacturing process is subjected to finite accuracy. Thus, the real system is uncertain with respect to the designed system [19].

Computational model: In order to analyze the real system behaviour, a computational model should be used as predictive tool. The construction of this computational model initially performs a physical analysis of the designed system, identifies the associated physical phenomena and makes hypotheses and simplifications about its behaviour. The identified physical phenomena are then translated into equations in a mathematical formulation stage. Using the appropriate numerical methods, the model equations are then discretized and the resulting discrete system of equations is solved, providing an approximation to the computational model response. This approximate response is then used to predict the real system behaviour [19].

Numerical errors: The response obtained with the computational model is, in fact, an approximation to the model equation's true solution. Inaccuracies, intrinsic to the discretization process, are introduced in this step giving rise to *numerical errors* [19]. Other source of errors are: (i) the finite precision arithmetic that is used to perform the calculations, and (ii) possible bugs in the computer code implementation of the computational model.

Uncertainties on the data: The computational model is supplied with model input and parameters, which are (not exact) emulations of the real system input and parameters, respectively. Thus, it is uncertain with respect to the real system. The discrepancy between the real system and computational model supplied information is called *data uncertainties* [4, 19].

Uncertainties on the model: In the conception of the computational model, considerations made may or may not be in agreement with reality, which should introduce additional inaccuracies known as model uncertainties. This source of uncertainty is essentially due to lack of knowledge about the phenomenon of interest and, usually, is the largest source of inaccuracy in computational model response [4, 19].

Naturally, uncertainties affect the response of a computational model, but they should not be considered errors because they are physical in nature. Errors are purely mathematical in nature and can be controlled and reduced to a negligible level if the numerical methods and algorithms used are well known by the analyst [4, 19]. This differentiation is summarized in Fig. 4.
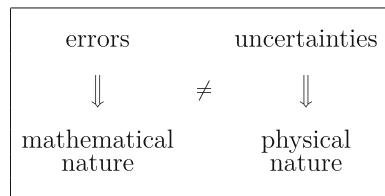
2.2 Verification and Validation

Today *verification* and *validation*, also called V&V, are two concepts of fundamental importance for any carefully done work in UQ. Early works advocating in favor of these ideas, and showing their importance, date back to the late 1990s and early 2000s [20–23]. The impact on the numerical simulation community was not immediate, but has been continuously growing over the years, conquering a prominent space in the last ten years, especially after the publication of Oberkampf and Roy's book [24].

These notions are well characterized in terms of two questions:

Verification:

Are we solving the equation right?

Fig. 4 The difference between errors and uncertainties

Validation:

<p style="text-align:center">Are we solving the right equation?</p>

Although extremely simplistic, the above "definitions" communicate, directly and objectively, the key ideas behind the two concepts. Verification is a task whose goal is to make sure that the model equation's solution is being calculated correctly. In other words, it is to check if the computational implementation has no critical bug and the numerical method works well. It is an exercise in mathematics. Meanwhile, validation is a task which aims to check if the model equations provide an adequate representation of the physical phenomenon/system of interest. The proper way to do this "validation check up" is through a direct comparison of the model responses with experimental data carefully obtained from the real system. It is an exercise in physics. In Fig. 5 the reader can see a schematic representation of the difference between the two notions.

An example in V&V: A skydiver jumps vertically in free fall, from a helicopter that is stopped in flight, from a height of $y_0 = 2000$ m with velocity $v_0 = 0$ m/s. Such situation is illustrated in Fig. 6. Imagine we want to know the skydiver height in every moment of the fall. To do this we develop a (toy) model where the falling man is idealized as point mass $m = 70$ kg, under the action of gravity $g = 9.81$ m/s^2. The height at time t is denoted by $y(t)$.

The skydiver's height at time t can be determined through the following initial value problem (IVP)

$$m\,\ddot{y}(t) + m\,g = 0, \tag{1}$$
$$\dot{y}(0) = v_0,$$
$$y(0) = y_0,$$

where the upper dot is an abbreviation for a time derivative, i.e., $\dot{\square} := d\square/dt$. This toy model is obtained from Newton's 2nd law of motion and considers the weigh as the only force acting on the skydiver body.

Imagine that we have developed a computer code to integrate this IVP using a standard 4th order Runge-Kutta method [25]. The model response obtained with this computer code is shown in Fig. 7.
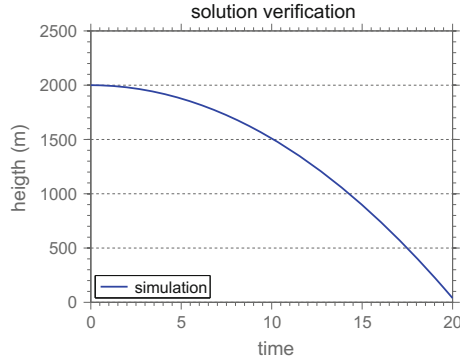
Fig. 5 The difference between verification and validation

solution verification		model validation
⇓	≠	⇓
mathematics		physics
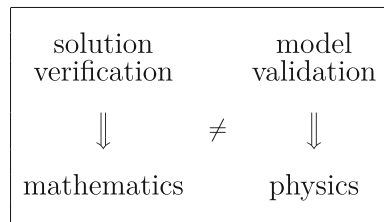
Fig. 6 V&V example: a skydiver in free fall from an initial height y_0
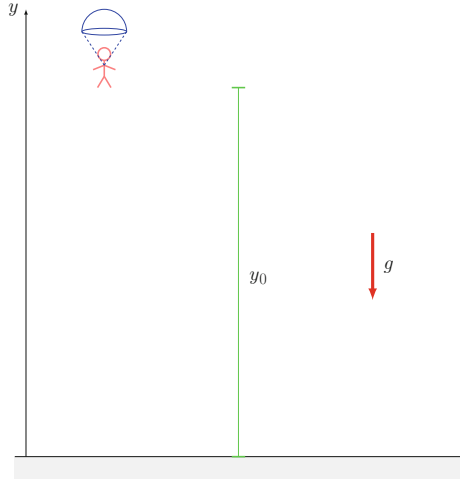

Fig. 7 Response obtained with the toy model

To check accuracy of the numerical method and its implementation we have at our disposal the analytical (reference) solution of the IVP, given by

$$y(t) = -\frac{1}{2}g\,t^2 + v_0\,t + y_0. \tag{2}$$

In Fig. 8a we can see the comparison between toy model response (solid blue curve —) and the reference solution (dashed red curve - - -). We note that both curves are in excellent agreement, but if we look at Fig. 8b, which shows the difference between numerical and analytical solutions, it is evident the effectiveness of the numerical method and the robustness of its implementation become ever clearer.
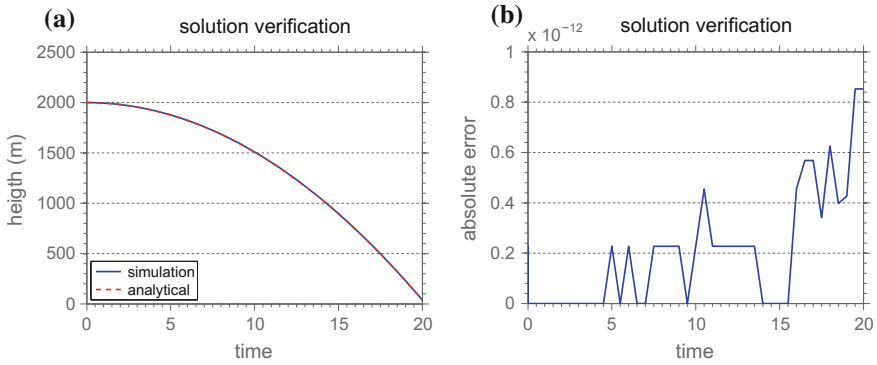
Fig. 8 **a** Solution verification: comparison between toy model response and reference solution;
b absolute error of Runge-Kutta method approximation

Here the verification was made taking as reference the real solution of the model
equation. In the most frequent case, the model equations solution is not known. In
such a situation, the verification task can be performed, for instance, using the *method
of manufactured solutions* [24, 26–28].

Now let's turn our attention to model validation, and compare simulation results
with experimental data, such as shown in Fig. 9a. We note that the simulation is
completely in disagreement with the experimental observations. In other words, the
model does not provide an adequate representation of the real system behaviour.

The toy model above take into account the gravitational force which attracts the
skydiver toward the ground, but neglects air resistance effects. This is the major
reason for the observed discrepancy, the model deficiency (model uncertainty). If
the air drag force effects are included, the improved model below is obtained
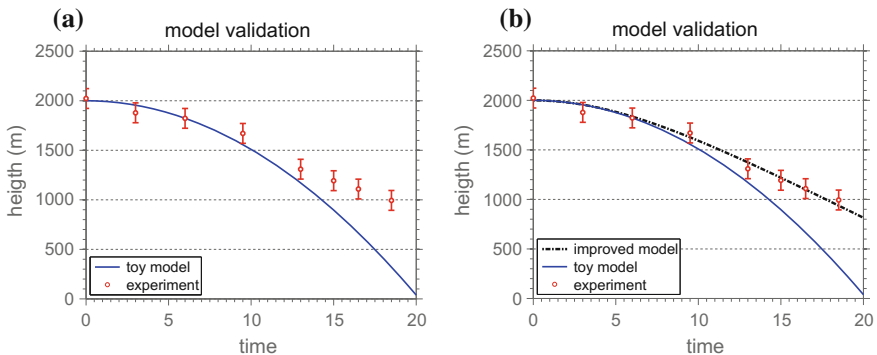
Fig. 9 **a** Model validation: comparison between experimental data and the toy model, **b** compari-
son between experimental data, the toy model, and the improved model

$$m\,\ddot{y}(t) + m\,g - \frac{1}{2}\rho\,A\,C_D\,(\dot{y}(t))^2 = 0, \tag{3}$$

$$\dot{y}(0) = v_0,$$

$$y(0) = y_0,$$

where ρ is the air mass density, A is the cross-sectional area of the falling body, and C_D is the (dimensionless) drag coefficient.

With this new model, a better agreement between simulation and experiment is expected. In Fig. 9b the reader can see the comparison between experimental data and the responses of both models, where we note that the improved model provides more plausible results.

An important message, implicit in this example, is that epistemic uncertainties can be reduced by increasing the actual knowledge about the phenomenon/system of interest [22, 24].

2.3 Two Approaches to Model Uncertainties

Being uncertainties in physical system the focus of stochastic modeling, two approaches are found in the scientific literature to deal with then: *probabilistic*, and *non-probabilistic*.

Probabilistic approach: This approach uses probability theory to model the physical system uncertainties as random mathematical objects. This approach is well-developed and very consistent from the mathematical foundations point of view for this reason, there is a consensus among the experts that it is preferable whenever possible to use it [4].

Non-probabilistic approach: This approach uses techniques such as interval analysis, fuzzy finite element, imprecise probabilities, evidence theory, probability bounds analysis, fuzzy probabilities, etc. In general these techniques are less suitable for problems in high stochastic dimension. Usually they are applied only when the probabilistic approach can not be used [4].

Because of their aleatory nature, data uncertainties are, quite naturally, well represented in a probabilistic environment. Thus, the *parametric probabilistic approach* is an appropriate method to describe this class of uncertainties. This procedure consists in describe the computational model random parameters as random objects (random variables, random vectors, random processes and/or random fields) and then consistently construct their joint probability distribution. Consequently, the model response becomes aleatory, and starts to be modeled by another random object, depending on the nature of the model equations. The model response is calculated using a stochastic solver. For further details, we recommend [4, 19, 29–31].

When model uncertainties are the focus of analysis, the non-probabilistic techniques receive more attention. Since the origin of this type of uncertainty is epistemic (lack of knowledge), it is not naturally described in a probabilistic setting. More details on non-probabilistic techniques can be seen in [32–34]. However, the use of probability theory for model uncertainties is still possible through a methodology called *nonparametric probabilistic approach*. This method, which also take into account the data uncertainty, was proposed in [35], and describes the mathematical operators in the computational model (not the parameters) as random objects. The probability distribution of these objects must be constructed in a consistent way, using the Principle of Maximum Entropy. The methodology lumps the model level of uncertainty into a single parameter, which can be identified by solving a parameter identification problem when (enough) experimental data is available. An overview of this technique can be seen in [19, 31].

A *generalized probabilistic approach* describing model and data uncertainties on different probability spaces, with some advantages, is presented in [36, 37].

3 A Brief on Probability Theory

This section presents a brief review of probability basic concepts. Such exposition is elementary, being insufficient for a solid understanding of the theory. Our objective is only to equip the reader with basic probabilistic vocabulary necessary to understand UQ scientific literature. For deeper studies on probability theory, we recommend the references [38–41].

3.1 Probability Space

The mathematical framework in which a random experiment is described consists of a triplet $(\Omega, \Sigma, \mathbb{P})$, where Ω is called *sample space*, Σ is a *σ-algebra* over Ω, and $\mathbb{P}$ is a *probability measure*. The trio $(\Omega, \Sigma, \mathbb{P})$ is called *probability space*.

Sample space: The set which contains all possible outcomes (events) for a certain random experiment is called sample space, being represented by Ω. An elementary event in Ω is denoted by ω. Sample spaces may contain a number of events that is finite, denumerable (countable infinite) or non-denumerable (non-countable infinity). The following three examples, respectively, illustrate the three situations:

Example 3.1 (finite sample space) Rolling a given cube-shaped fare die, where the faces are numbered from 1 through 6, we have $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Example 3.2 (denumerable sample space) Choosing randomly an integer even number, we have $\Omega = \{\dots, -8, -6, -4, -2, 0, 2, 4, 6, 8, \dots\}$.

Example 3.3 (non-denumerable sample space) Measuring the temperature (in Kelvin) at Rio de Janeiro city during the summer, we have $\Omega = [a, b] \subset [0, +\infty)$.

σ-algebra: In general, not all of the outcomes in Ω are of interest so that, in a probabilistic context, we need to pay attention only to the relevant events. Intuitively, the σ-algebra Σ is the set of relevant outcomes for a random experiment. Formally, Σ is σ-algebra if:

- $\phi \in \Sigma$ (contains the empty set);
- for any $\mathcal{A} \in \Sigma$ we also have $\mathcal{A}^c \in \Sigma$ (closed under complementation);
- for any countable collections of $\mathcal{A}_i \in \Sigma$, it is true that $\bigcup_{i=1}^{\infty} \mathcal{A}_i \in \Sigma$ (closed under denumerable unions).

Example 3.4 Consider the experiment of rolling a die with sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$ where we are interested in knowing if the result is odd or even. In this case, a suitable σ-algebra is $\Sigma = \{\Omega, \{1, 3, 5\}, \{2, 4, 6\}, \phi\}$. On the other hand, if we are interested in knowing the upper face value after rolling, an adequate 2^{ω} (set of all subsets of ω). Different σ-algebras generate distinct probability spaces.

Probability measure: The probability measure is a function $\mathbb{P} : \Sigma \to [0, 1] \subset \mathbb{R}$ which indicates the level of expectation that a certain event in Σ occurs. In technical language, $\mathbb{P}$ has the following properties:

- $\mathbb{P}\{\mathcal{A}\} \geq 0$ for any $\mathcal{A} \in \Sigma$ (probability is nonnegative);
- $\mathbb{P}\{\Omega\} = 1$ (entire space has probability one);
- for any denumerable collection of mutually disjoint events $\mathcal{A}_i$, it is true that $\mathbb{P}\left\{\bigcup_{i=1}^{\infty} \mathcal{A}_i\right\} = \sum_{i=1}^{\infty} \mathbb{P}\{\mathcal{A}_i\}$.

Note that $\mathbb{P}\{\phi\} = 0$ (empty set has probability zero).

3.2 Random Variables

A mapping $\mathbb{X} : \Omega \to \mathbb{R}$ is called a *random variable* if the preimage of every real number under $\mathbb{X}$ is a relevant event, i.e.,

$$\mathbb{X}^{-1}(x) = \{\omega \in \Omega : \mathbb{X}(\omega) \leq x\} \in \Sigma, \qquad \text{for every } x \in \mathbb{R}. \tag{4}$$

We denote a realization of $\mathbb{X}$ by $\mathbb{X}(\omega)$.

Random variables provide numerical characteristics of interesting events, in such a way that we can forget the sample space. In practice, when working with a probabilistic model, we are concerned only with the possible values of $\mathbb{X}$.

Example 3.5 The random experiment is now toss a two fare dice, then $\Omega = \{(d_1, d_2) : 1 \leq d_1 \leq 6 \text{ and } 1 \leq d_2 \leq 6\}$. Define the random variables $\mathbb{X}_1$ and $\mathbb{X}_2$ in such way that $\mathbb{X}_1(\omega) = d_1 + d_2$ and $\mathbb{X}_2(\omega) = d_1 d_2$. The former is a numerical indicator of the sum of dice upper faces values, while the latter characterizes the product of these numbers.

3.3   Probability Distribution

The *probability distribution* of $\mathbb{X}$, denoted by $P_{\mathbb{X}}$, is defined as the probability of the elementary event $\{\mathbb{X} \leq x\}$, i.e.,

$$P_{\mathbb{X}}(x) = \mathbb{P}\left\{\mathbb{X} \leq x\right\}. \tag{5}$$

$P_{\mathbb{X}}$ has the following properties:

- $0 \leq P_{\mathbb{X}}(x) \leq 1$ (it is a probability);
- $P_{\mathbb{X}}$ is non-decreasing, and right-continuous;
- $\lim_{x \to -\infty} P_{\mathbb{X}}(x) = 0$, and $\lim_{x \to +\infty} P_{\mathbb{X}}(x) = 1$;

so that

$$P_{\mathbb{X}}(x) = \int_{\xi=-\infty}^{x} dP_{\mathbb{X}}(\xi), \tag{6}$$

and

$$\int_{\mathbb{R}} dP_{\mathbb{X}}(x) = 1. \tag{7}$$

$P_{\mathbb{X}}$ is also known as *cumulative distribution function* (CDF).

3.4   Probability Density Function

If the function $P_{\mathbb{X}}$ is differentiable, then we call its derivative the *probability density function* (PDF) of $\mathbb{X}$, using the notation $p_{\mathbb{X}}$.

Given that $p_{\mathbb{X}} = dP_{\mathbb{X}}/dx$, we have $dP_{\mathbb{X}}(x) = p_{\mathbb{X}}(x)\, dx$, and then

$$P_{\mathbb{X}}(x) = \int_{\xi=-\infty}^{x} p_{\mathbb{X}}(\xi)\, d\xi. \tag{8}$$

The PDF is a function $p_{\mathbb{X}} : \mathbb{R} \to [0, +\infty)$ such that

$$\int_{\mathbb{R}} p_{\mathbb{X}}(x)\, dx = 1. \tag{9}$$

3.5 Mathematical Expectation Operator

Given a function $g : \mathbb{R} \to \mathbb{R}$, the composition of g with the random variable $\mathbb{X}$ is also a random variable $g(\mathbb{X})$.

The *mathematical expectation* of $g(\mathbb{X})$ is defined by

$$\mathbb{E}\{g(\mathbb{X})\} = \int_{\mathbb{R}} g(x)\, p_{\mathbb{X}}(x)\, dx. \tag{10}$$

With the aid of this operator, we define

$$
\begin{aligned}
m_{\mathbb{X}} &= \mathbb{E}\{\mathbb{X}\} \\
&= \int_{\mathbb{R}} x\, p_{\mathbb{X}}(x)\, dx,
\end{aligned} \tag{11}
$$

$$
\begin{aligned}
\sigma_{\mathbb{X}}^2 &= \mathbb{E}\left\{ (\mathbb{X} - m_{\mathbb{X}})^2 \right\} \\
&= \int_{\mathbb{R}} (x - m_{\mathbb{X}})^2\, p_{\mathbb{X}}(x)\, dx,
\end{aligned} \tag{12}
$$

and

$$\sigma_{\mathbb{X}} = \sqrt{\sigma_{\mathbb{X}}^2}, \tag{13}$$

which are the *mean value*, *variance*, and *standard deviation* of $\mathbb{X}$, respectively. Note further that

$$\sigma_{\mathbb{X}}^2 = \mathbb{E}\{\mathbb{X}^2\} - m_{\mathbb{X}}^2. \tag{14}$$

The ratio between standard deviation and mean value is called *coefficient of variation* of $\mathbb{X}$

$$\delta_{\mathbb{X}} = \frac{\sigma_{\mathbb{X}}}{m_{\mathbb{X}}}, \quad m_{\mathbb{X}} \neq 0. \tag{15}$$

These scalar values are indicators of the random variable behaviour. Specifically, the mean value $m_{\mathbb{X}}$ is a central tendency indicator, while variance $\sigma_{\mathbb{X}}^2$ and standard deviation $\sigma_{\mathbb{X}}$ are measures of dispersion around the mean. The difference in these dispersion measures is that $\sigma_{\mathbb{X}}$ has the same unit as $m_{\mathbb{X}}$ while $\sigma_{\mathbb{X}}^2$ is measured in $m_{\mathbb{X}}$

unit squared. Once it is dimensionless, the coefficient of variation is a standardized measure of dispersion.

For our purposes, it is also convenient to define the entropy of $p_{\mathbb{X}}$

$$S\left(p_{\mathbb{X}}\right) = -\mathbb{E}\left\{\ln\left(p_{\mathbb{X}}(\mathbb{X})\right)\right\}, \tag{16}$$

which (see Eq. 10) is equivalent to

$$S\left(p_{\mathbb{X}}\right) = -\int_{\mathbb{R}} p_{\mathbb{X}}(x)\,\ln\left(p_{\mathbb{X}}(x)\right)dx. \tag{17}$$

Entropy provides a measure for the level of uncertainty of $p_{\mathbb{X}}$ [42].

3.6 Second-Order Random Variables

The mapping $\mathbb{X}$ is a *second-order random variable* if the expectation of its square (second-order moment) is finite, i.e.,

$$\mathbb{E}\left\{\mathbb{X}^2\right\} < +\infty. \tag{18}$$

The inequality expressed in (18) implies that $\mathbb{E}\left\{\mathbb{X}\right\} < +\infty$ ($m_{\mathbb{X}}$ is also finite). Consequently, with the aid of Eq. (14), we see that a second-order random variable $\mathbb{X}$ has finite variance, i.e., $\sigma_{\mathbb{X}}^2 < +\infty$.

This class of random variables is very relevant for stochastic modeling, once, for physical considerations, typical random parameters in physical systems have finite variance.

3.7 Joint Probability Distribution

Given the random variables $\mathbb{X}$ and $\mathbb{Y}$, the *joint probability distribution* of $\mathbb{X}$ and $\mathbb{Y}$, denoted by $P_{\mathbb{X}\mathbb{Y}}$, is defined as

$$P_{\mathbb{X}\mathbb{Y}}(x, y) = \mathbb{P}\left\{\{\mathbb{X} \le x\} \cap \{\mathbb{Y} \le y\}\right\}. \tag{19}$$

The function $P_{\mathbb{X}\mathbb{Y}}$ has the following properties:

- $0 \le P_{\mathbb{X}\mathbb{Y}}(x, y) \le 1$ (it is a probability);
- $P_{\mathbb{X}}(x) = \lim_{y \to +\infty} P_{\mathbb{X}\mathbb{Y}}(x, y)$, and $P_{\mathbb{Y}}(y) = \lim_{x \to +\infty} P_{\mathbb{X}\mathbb{Y}}(x, y)$
  (marginal distributions are limits);

such that

$$P_{\mathbb{X}\mathbb{Y}}(x, y) = \int_{\xi=-\infty}^{x} \int_{\eta=-\infty}^{y} dP_{\mathbb{X}\mathbb{Y}}(\xi, \eta), \tag{20}$$

and

$$\int\int_{\mathbb{R}^2} dP_{\mathbb{X}\mathbb{Y}}(x, y) = 1. \tag{21}$$

$P_{\mathbb{X}\mathbb{Y}}$ is also known as *joint cumulative distribution function.*

3.8 Joint Probability Density Function

If the partial derivative $\partial^2 P_{\mathbb{X}\mathbb{Y}}/\partial x\,\partial y$ exists, for any x and y, then it is called *joint probability density function* of $\mathbb{X}$ and $\mathbb{Y}$, being denoted by

$$p_{\mathbb{X}\mathbb{Y}}(x, y) = \frac{\partial^2 P_{\mathbb{X}\mathbb{Y}}}{\partial x\,\partial y}(x, y). \tag{22}$$

Hence, we can write $dP_{\mathbb{X}\mathbb{Y}}(x, y) = p_{\mathbb{X}\mathbb{Y}}(x, y)\,dy\,dx$, so that

$$P_{\mathbb{X}\mathbb{Y}}(x, y) = \int_{\xi=-\infty}^{x} \int_{\eta=-\infty}^{y} p_{\mathbb{X}\mathbb{Y}}(\xi, \eta)\,d\eta\,d\xi. \tag{23}$$

The joint PDF is a function $p_{\mathbb{X}\mathbb{Y}} : \mathbb{R} \to [0, +\infty)$ which satisfies

$$\int\int_{\mathbb{R}^2} p_{\mathbb{X}\mathbb{Y}}(x, y)\,dy\,dx = 1. \tag{24}$$

3.9 Conditional Probability

Consider the pair of random events $\{\mathbb{X} \leq x\}$ and $\{\mathbb{Y} \leq y\}$, where the probability of occurrence of the second one is non-zero, i.e., $\mathbb{P}\{\{\mathbb{Y} \leq y\}\} > 0$. The *conditional probability* of event $\{\mathbb{X} \leq x\}$, given the occurrence of event $\{\mathbb{Y} \leq y\}$, is defined as

$$\mathbb{P}\{\{\mathbb{X} \leq x\} \,|\, \{\mathbb{Y} \leq y\}\} = \frac{\mathbb{P}\{\{\mathbb{X} \leq x\} \cap \{\mathbb{Y} \leq y\}\}}{\mathbb{P}\{\{\mathbb{Y} \leq y\}\}}. \tag{25}$$

3.10   Independence of Random Variables

The event $\{\mathbb{X} \leq x\}$ is said to be *independent* of event $\{\mathbb{Y} \leq y\}$ if the occurrence of the former does not affect the occurrence of the later, i.e.,

$$\mathbb{P}\left\{\{\mathbb{X} \leq x\} \,\middle|\, \{\mathbb{Y} \leq y\}\right\} = \mathbb{P}\left\{\{\mathbb{X} \leq x\}\right\}. \tag{26}$$

Consequently, if the random variables $\mathbb{X}$ and $\mathbb{Y}$ are independent, from Eq. (25) we see that

$$\mathbb{P}\left\{\{\mathbb{X} \leq x\} \cap \{\mathbb{Y} \leq y\}\right\} = \mathbb{P}\left\{\mathbb{X} \leq x\right\} \mathbb{P}\left\{\mathbb{Y} \leq y\right\}. \tag{27}$$

This also implies that

$$P_{\mathbb{X}\mathbb{Y}}(x, y) = P_{\mathbb{X}}(x)\, P_{\mathbb{Y}}(y), \tag{28}$$

and

$$p_{\mathbb{X}\mathbb{Y}}(x, y) = p_{\mathbb{X}}(x)\, p_{\mathbb{Y}}(y). \tag{29}$$

3.11   Random Process

A *random process* $\mathbb{U}$, indexed by $t \in \mathcal{T}$, is a mapping

$$\mathbb{U} : (t, \omega) \in \mathcal{T} \times \Omega \to \mathbb{U}(t, \omega) \in \mathbb{R}, \tag{30}$$

such that, for fixed t, the output is a random variable $\mathbb{U}(t, \cdot)$, while for fixed ω, $\mathbb{U}(\cdot, \omega)$ is a function of t. In other words, it is a collection of random variables indexed by a parameter. Roughly speaking, a random process, also called *stochastic process*, can be thought of as a time-dependent random variable.

4   Parametric Probabilistic Modeling of Uncertainties

This section discusses the use of the parametric probabilistic approach to describe uncertainties in physical systems. Our goal is to provide the reader with some key ideas behind this approach and call attention to the fundamental issues that must be taken into account. The exhibition is based on [13, 15] and use a simplistic example to discuss the theory.

Fig. 10 Mechanical system composed by a fixed spring and a constant force
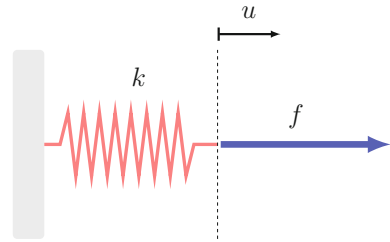
4.1 A Simplistic Stochastic Mechanical System

Consider the mechanical system which consists of a spring fixed on the left side of a wall and being pulled by a constant force on the right side (Fig. 10). The spring stiffness is k, the force is represented by f, and the spring displacement is denoted by u. A mechanical-mathematical model to describe this system behaviour is given by

$$k\,u = f, \tag{31}$$

from where we get the system response

$$u = k^{-1}f. \tag{32}$$

4.2 Stochastic Model for Uncertainties Description

We are interested in studying the case where the above mechanical system is subject to uncertainties on the stiffness parameter k. To describe the random behaviour of the mechanical system, we employ the parametric probabilistic approach.

Let us use the probability space $(\Omega, \Sigma, \mathbb{P})$, where the stiffness k is modeled as the random variable $\mathbb{K} : \Omega \to \mathbb{R}$. Therefore, due to result of the relationship imposed by Eq. (32), the displacement u is also uncertain, being modeled as a random variable $\mathbb{U} : \Omega \to \mathbb{R}$, which respects the equilibrium condition given by the following stochastic equation

$$\mathbb{K}\,\mathbb{U} = f. \tag{33}$$

It is reasonable to assume that the deterministic model is minimally representative, and corresponds to the mean of $\mathbb{K}$, i.e., $m_{\mathbb{K}} = k$. Additionally, for physical reasons, $\mathbb{K}$ must have a finite variance. Thus, $\mathbb{K}$ is assumed to be a second-order random variable, i.e., $\mathbb{E}\left\{\mathbb{K}^2\right\} < +\infty$.

4.3 The Importance of Knowing the PDF

Now that we have the random parameter described in a probabilistic context, and a stochastic model for the system, we can ask ourselves some questions about the system response. For instance, to characterize the system response central tendency, it is of interest to know the mean of $\mathbb{U}$, denoted by $m_{\mathbb{U}}$.

Since $m_{\mathbb{K}}$ is a known information about $\mathbb{K}$ (but $p_{\mathbb{K}}$ is unknown), we can ask ourselves: *Is it possible to compute $m_{\mathbb{U}}$ with this information only?* The answer for this question is negative. The reason is that $\mathbb{U} = \mathbb{K}^{-1} f$, so that

$$
\begin{aligned}
m_{\mathbb{U}} &= \mathbb{E}\left\{\mathbb{K}^{-1} f\right\} \\
&= \int_{\mathbb{R}} k^{-1} f\, p_{\mathbb{K}}(k)\, dk,
\end{aligned}
$$

and the last integral can only be calculated if $p_{\mathbb{K}}$ is known. Once the map $g(k) = k^{-1} f$ is nonlinear, $\mathbb{E}\left\{g\left(\mathbb{K}\right)\right\} \neq g\left(\mathbb{E}\left\{\mathbb{K}\right\}\right)$.

Conclusion: *In order to obtain any statistical information about model response, it is absolutely necessary to know the probability distribution of model parameters.*

4.4 Why Can't We Arbitrate Distributions?

As the knowledge of the probability distribution of $\mathbb{K}$ is necessary, let's assume that it is Gaussian distributed. In this way,

$$
p_{\mathbb{K}}(k) = \frac{1}{\sqrt{2\pi\, \sigma_{\mathbb{K}}^2}} \exp\left\{-\frac{(k - m_{\mathbb{K}})^2}{2\, \sigma_{\mathbb{K}}^2}\right\}, \tag{34}
$$

whose support is the entire real line, i.e., $\mathrm{Supp}\, p_{\mathbb{K}} = (-\infty, +\infty)$.

The attentive reader may question, at this point, that from the physical point of view, make no sense use a Gaussian distribution to model a stiffness parameter, since $\mathbb{K}$ is always positive. This is true and makes the arbitrary choice of a Gaussian distribution inappropriate. However, this is not the only reason against this choice.

For physical considerations, it is necessary that the model response $\mathbb{U}$ be a second-order (finite variance) random variable, i.e., . $\mathbb{E}\left\{\mathbb{U}^2\right\} < +\infty$. *Is this possible when we arbitrate the probability distribution as Gaussian?* No way! Just do a simple calculation

$$\mathbb{E}\left\{\mathbb{U}^2\right\} = \mathbb{E}\left\{\mathbb{K}^{-2} f^2\right\}$$
$$= \int_{\mathbb{R}} k^{-2} f^2\, p_{\mathbb{K}}(k)\, dk$$
$$= \int_{k=-\infty}^{+\infty} k^{-2} f^2 \left(\frac{1}{\sqrt{2\pi\,\sigma_{\mathbb{K}}^2}} \exp\left\{-\frac{(k-m_{\mathbb{K}})^2}{2\,\sigma_{\mathbb{K}}^2}\right\}\right) dk$$
$$= +\infty. \tag{35}$$

In fact, we also have $\mathbb{E}\left\{\mathbb{U}\right\} = m_{\mathbb{U}} = +\infty$.

The Gaussian distribution is a bad choice since $\mathbb{K}$ must be a positive-valued random variable (almost sure). Thus, we know the following information about $\mathbb{K}$:

- $\operatorname{Supp} p_{\mathbb{K}} \subseteq (0, +\infty) \Longleftrightarrow \mathbb{K} > 0$ *a.s.*
- $m_{\mathbb{K}} = k > 0$ is known
- $\mathbb{E}\left\{\mathbb{K}^2\right\} < +\infty$

All these requirements are verified by the exponential distribution, in which the PDF is given by the function

$$p_{\mathbb{K}}(k) = \mathbb{1}_{(0,+\infty)}(k)\frac{1}{m_{\mathbb{K}}} \exp\left\{-\frac{k}{m_{\mathbb{K}}}\right\}, \tag{36}$$

where $\mathbb{1}_{(0,+\infty)}$ the indicator function of the interval $(0, +\infty)$.

However, we still have

$$\mathbb{E}\left\{\mathbb{U}^2\right\} = \mathbb{E}\left\{\mathbb{K}^{-2} f^2\right\}$$
$$= \int_{\mathbb{R}} k^{-2} f^2\, p_{\mathbb{K}}(k)\, dk$$
$$= \int_{k=0}^{+\infty} k^{-2} f^2 \left(\frac{1}{m_{\mathbb{K}}} \exp\left\{-\frac{k}{m_{\mathbb{K}}}\right\}\right) dk$$
$$= +\infty, \tag{37}$$

once the function $k \mapsto k^{-2}$ diverges in $k = 0$. Thus, in order to $\mathbb{E}\left\{\mathbb{U}^2\right\} < +\infty$, we must have $\mathbb{E}\left\{\mathbb{K}^{-2}\right\} < +\infty$.

Conclusion: *Arbitrate probability distributions for parameters can generate a stochastic model that is inconsistent from the physical/mathematical point of view.*

4.5 An Acceptable Distribution

In short, an adequate distribution must satisfy the conditions below

- $\text{Supp}\, p_{\mathbb{K}} \subseteq (0, +\infty) \implies \mathbb{K} > 0$ *a.s.*
- $m_{\mathbb{K}} = k > 0$ is known
- $\mathbb{E}\left\{\mathbb{K}^2\right\} < +\infty$
- $\mathbb{E}\left\{\mathbb{K}^{-2}\right\} < +\infty$.

The gamma distribution satisfies all the conditions above so that it is an acceptable choice. Its PDF is written as

$$p_{\mathbb{K}}(k) = \mathbb{1}_{(0,+\infty)}(k) \frac{1}{m_{\mathbb{K}}} \frac{\delta_{\mathbb{K}}^{-2\delta_{\mathbb{K}}^{-2}}}{\Gamma\left(\delta_{\mathbb{K}}^{-2}\right)} \left(k/m_{\mathbb{K}}\right)^{\delta_{\mathbb{K}}^{-2}-1} \exp\left\{-\frac{k/m_{\mathbb{K}}}{\delta_{\mathbb{K}}^2}\right\}, \tag{38}$$

where $0 \le \delta_{\mathbb{K}} = \sigma_{\mathbb{K}}/m_{\mathbb{K}} < 1/\sqrt{2}$ is a dispersion parameter, and Γ denotes the gamma function

$$\Gamma(\alpha) = \int_{t=0}^{+\infty} t^{\alpha-1}\, e^{-t}\, dt. \tag{39}$$

Conclusion: *Probability distributions for model parameters must be objectively constructed (never arbitrated), and take into account all available information about the parameters.*

4.6 *How to Safely Specify a Distribution?*

In the previous example, we have chosen a suitable probability distribution by verifying if the candidate distributions satisfy the constraints imposed by physical and mathematical properties of the model parameter/response. However, this procedure is not practical and does not provide a unique distribution as a possible choice. For instance, in the spring example, uniform, lognormal and an infinitude of other distributions are also acceptable (compatible with the restrictions).

Thus, it is natural to ask ourselves if it is possible to construct a consistent stochastic model in a systematic way. The answer for this question is affirmative, and the objective procedure to be used depends on the scenario.

Scenario 1: large amount of experimental data is available

The usual procedure in this case employs nonparametric statistical estimation to construct the random parameter distribution from the available data [13, 15, 43].

Suppose we want to estimate the probability distribution of a random variable $\mathbb{X}$, and for that we have N independent samples of $\mathbb{X}$, respectively denoted by X^1, X^2, …, X^N.

Assuming, without loss of generality, that $X^1 < X^2 < \cdots < X^N$, we consider an estimator for $P_{\mathbb{X}}(x)$ given by

$$\widehat{P}_N(x) = \frac{1}{N} \sum_{n=1}^{N} \mathcal{H}(x - X^n), \tag{40}$$

where $\mathcal{H}$ is defined as

$$\mathcal{H}(x - X^n) = \begin{cases} 1 & \text{if } x \geq X^n \\ 0 & \text{if } x < X^n. \end{cases} \tag{41}$$

This estimator, which is mean-square consistent

$$\mathbb{E}\left\{\widehat{P}_N(x)\right\} = P_{\mathbb{X}}(x), \tag{42}$$

and unbiased

$$\lim_{N \to +\infty} \mathbb{E}\left\{\left(\widehat{P}_N(x) - P_{\mathbb{X}}(x)\right)^2\right\} = 0, \tag{43}$$

is known as the *empirical distribution function* or the *empirical CDF* [13, 15, 43, 44].

If the random variable admits a PDF, it is more common to estimate its probability distribution using a *histogram*, that is an estimator for $p_{\mathbb{X}}(x)$. To construct such a histogram, the first step is to divide the random variable support into a denumerable number of bins $\mathcal{B}_m$, where

$$\mathcal{B}_m = [(m-1)h, mh], \qquad m \in \mathbb{Z}, \tag{44}$$

being h the *bin width*. Then we count the number of samples in each of the bins $\mathcal{B}_m$, denoting this number by ν_m. After that, we normalize the counter (dividing by Nh) to obtain the *normalized relative frequency* $\nu_m/(Nh)$. Finally, for each bin $\mathcal{B}_m$, we plot a vertical bar with height $\nu_m/(Nh)$ [43, 44].

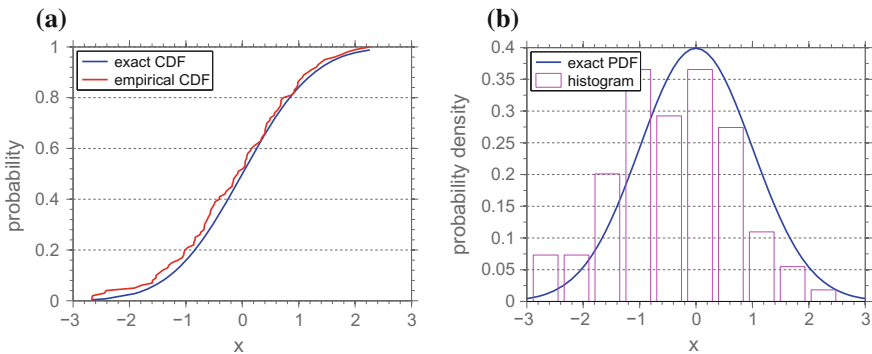In analytical terms (see [43, 44]) we can write this as PDF estimator as

$$\widehat{p}_N(x) = \frac{1}{Nh} \sum_{m=-\infty}^{+\infty} \nu_m \mathbb{1}_{B_m}(x), \tag{45}$$

where $\mathbb{1}_{B_m}(x)$ is the *indicator function* of $\mathcal{B}_m$, defined as

$$\mathbb{1}_{B_m}(x) = \begin{cases} 1 & \text{if } x \in \mathcal{B}_m \\ 0 & \text{if } x \notin \mathcal{B}_m. \end{cases} \tag{46}$$
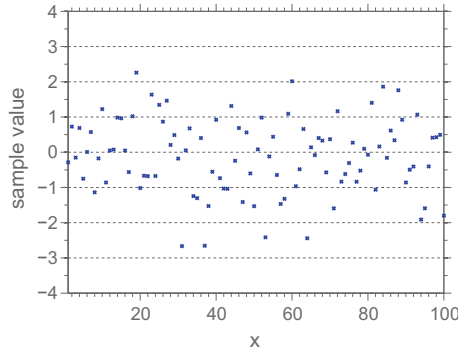
Fig. 11 These samples are realizations of a standard Gaussian random variable

Fig. 12 **a** Estimators for the probability distribution of $\mathbb{X}$: the empirical CDF, and **b** a histogram

Both estimators above are easily constructed, but they require a large number of samples in order to obtain a reasonable approximation [43, 44].

In practice, these estimators are used when we do not know the random variable distribution. However, to illustrate the use of these tools, let us consider a dataset with $N = 100$ samples obtained from the (standard) Gaussian random variable $\mathbb{X}$, with zero mean and unity standard deviation. Such samples are illustrated in Fig. 11. Considering these samples, we can construct the two estimators shown in Fig. 12, with the empirical CDF on the left and a histogram on the right.

Scenario 2: little or even none experimental data is available

When very little or no experimental data is available, to the best of the author's knowledge, the most conservative approach uses the *Maximum Entropy Principle* (MEP) [15, 45, 46, 48], with parametric statistical estimation, to construct the random parameter distribution. If no experimental data is available, this approach takes into account only theoretical information which can be inferred from the model physics and its mathematical structure to specify the desired distribution.

The MEP can be stated as follows: *Among all the (infinite) probability distributions, consistent with the known information about a random parameter, the most unbiased is the one which corresponds to the maximum of entropy PDF.*

Using it to specify the distribution of a random variable $\mathbb{X}$ presupposes finding the unique PDF which maximizes the entropy (objective function)

$$S\left(p_{\mathbb{X}}\right) = -\int_{\mathbb{R}} p_{\mathbb{X}}(x) \ln\left(p_{\mathbb{X}}(x)\right) dx, \tag{47}$$

respecting $N + 1$ constraints (known information) given by

$$\int_{\mathbb{R}} g_k(\mathbb{X}) p_{\mathbb{X}}(x) dx = \mu_k, \qquad k = 0, \dots, N, \tag{48}$$

where g_k are known real functions, with $g_0(x) = 1$, and μ_k are known real values, being $\mu_0 = 1$. The restriction associated with $k = 0$ corresponds to the normalization condition of $p_{\mathbb{X}}$, while the other constraints, typically, but not exclusively, represent statistical moments of $\mathbb{X}$.

To solve this problem, the method of Lagrange multipliers is employed, and introduces other $(N + 1)$ unknown real parameters λ_k (Lagrange multipliers). We can show that if this optimization problem has a solution, it actually corresponds to a maximum and is unique, being written as

$$p_{\mathbb{X}}(x) = \mathbb{1}_{\mathcal{K}}(x) \exp\left(-\lambda_0\right) \exp\left(-\sum_{k=1}^{N} \lambda_k g_k(x)\right), \tag{49}$$

where $\mathcal{K} = \text{Supp}\, p_{\mathbb{X}}$ here denotes the support of $p_{\mathbb{X}}$, and $\mathbb{1}_{\mathcal{K}}(x)$ is the indicator function of $\mathcal{K}$.

The Lagrange multipliers, which depend on μ_k and $\mathcal{K}$, are identified with the aid of the restriction defined in Eq. (48) using techniques of parametric statistics.

4.7 Using the Maximum Entropy Principle

In this section we exemplify the use of the MEP to consistently specify the probability distribution of a random variable $\mathbb{X}$.

Suppose that $\text{Supp}\, p_{\mathbb{X}} = [a, b]$ is the only information we know about $\mathbb{X}$. In this case, a consistent (unbiased) probability distribution for $\mathbb{X}$ is obtained solving the following optimization problem:

Maximize

$$S\left(p_{\mathbb{X}}\right) = -\int_{\mathbb{R}} p_{\mathbb{X}}(x) \ln\left(p_{\mathbb{X}}(x)\right) dx$$

$$= -\int_{x=a}^{b} p_{\mathbb{X}}(x) \ln\left(p_{\mathbb{X}}(x)\right) dx,$$

subjected to the constraint

$$1 = \int_{\mathbb{R}} p_{\mathbb{X}}(x) \, dx$$

$$= \int_{x=a}^{b} p_{\mathbb{X}}(x) \, dx.$$

To solve this optimization problem, first we define the Lagrangian

$$\mathcal{L}\left(p_{\mathbb{X}}, \lambda_0\right) = -\int_{x=a}^{b} p_{\mathbb{X}}(x) \ln\left(p_{\mathbb{X}}(x)\right) dx \; - (\lambda_0 - 1)\left(\int_{x=a}^{b} p_{\mathbb{X}}(x)\, dx - 1\right), \quad (50)$$

where $\lambda_0 - 1$ is the associated Lagrange multiplier. It is worth mentioning that λ_0 depends on the known information about $\mathbb{X}$, i.e. $\lambda_0 = \lambda_0(a, b)$.

Then we impose the necessary conditions for an extreme

$$\frac{\partial \mathcal{L}}{\partial p_{\mathbb{X}}}\left(p_{\mathbb{X}}, \lambda_0\right) = 0, \;\; \text{and} \;\; \frac{\partial \mathcal{L}}{\partial \lambda_0}\left(p_{\mathbb{X}}, \lambda_0\right) = 0, \quad (51)$$

whence we conclude that

$$p_{\mathbb{X}}(x) = \mathbb{1}_{[a,b]}(x)\, e^{-\lambda_0}, \;\; \text{and} \;\; \int_{\mathbb{R}} p_{\mathbb{X}}(x)\, dx = 1. \quad (52)$$

The first equation in Eq. (52) provides the PDF of $\mathbb{X}$ in terms of the Lagrange multiplier λ_0, while the second equation corresponds to the known information about this random variable (the normalization condition).

In order to represent $p_{\mathbb{X}}$ in terms of the known information (a and b), we need to find the dependence of λ_0 with respect to these parameters. To this end, let's go to replace the expression of $p_{\mathbb{X}}$ into the second equation of Eq. (52), so that

$$\int_{\mathbb{R}} \mathbb{1}_{[a,b]}(x)\, e^{-\lambda_0}\, dx = 1 \;\; \implies \;\; e^{-\lambda_0}\,(b-a) = 1, \;\; \implies \;\; e^{-\lambda_0} = \frac{1}{b-a}, \quad (53)$$

from where we get

$$p_{\mathbb{X}}(x) = \mathbb{1}_{[a,b]}(x)\, \frac{1}{b-a}, \quad (54)$$

Table 1 Maximum entropy distributions for given known information

Support	Known information	Maximum entropy PDF
$[a,b]$	–	$p_{\mathbb{X}}(x) = \mathbb{1}_{[a,b]}(x) \frac{1}{b-a}$ (**uniform in** $[a,b]$)
$[a,b]$	$\mathbb{E}\{\mathbb{X}\} = m_{\mathbb{X}} \in [a,b]$	$p_{\mathbb{X}}(x) = \mathbb{1}_{[a,b]}(x) \exp\left(\lambda_0 - x\,\lambda_1\right)$ $\lambda_0 = \lambda_0(a,b,m_{\mathbb{X}})$ $\lambda_1 = \lambda_1(a,b,m_{\mathbb{X}})$
$[a,b]$	$\mathbb{E}\{\mathbb{X}\} = m_{\mathbb{X}} \in [a,b]$ $\mathbb{E}\{\mathbb{X}^2\} = m_{\mathbb{X}}^2 + \sigma_{\mathbb{X}}^2$	$p_{\mathbb{X}}(x) = \mathbb{1}_{[a,b]}(x) \exp\left(\lambda_0 - x\,\lambda_1 - x^2\,\lambda_2\right)$ $\lambda_0 = \lambda_0(a,b,m_{\mathbb{X}},\sigma_{\mathbb{X}})$ $\lambda_1 = \lambda_1(a,b,m_{\mathbb{X}},\sigma_{\mathbb{X}})$ $\lambda_2 = \lambda_2(a,b,m_{\mathbb{X}},\sigma_{\mathbb{X}})$
$[0,1]$	$\mathbb{E}\{\ln(\mathbb{X})\} = p,\ |p| < +\infty$ $\mathbb{E}\{\ln(1-\mathbb{X})\} = q,\ |q| < +\infty$	$p_{\mathbb{X}}(x) = \mathbb{1}_{[0,1]}(x) \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}$ $a = \left(m_{\mathbb{X}}/\delta_{\mathbb{X}}^2\right)\left(1/m_{\mathbb{X}} - \delta_{\mathbb{X}}^2 - 1\right)$ $b = a\left(1/m_{\mathbb{X}} - 1\right)$ (**beta with shape parameters** a **and** b)
$(0,+\infty)$	$\mathbb{E}\{\mathbb{X}\} = m_{\mathbb{X}} > 0$	$p_{\mathbb{X}}(x) = \mathbb{1}_{(0,+\infty)}(x) \frac{1}{m_{\mathbb{X}}} \exp\left(-\frac{x}{m_{\mathbb{X}}}\right)$ (**exponential with mean** $m_{\mathbb{X}}$)
$(0,+\infty)$	$\mathbb{E}\{\mathbb{X}\} = m_{\mathbb{X}} > 0$ $\mathbb{E}\{\ln(\mathbb{X})\} = q,\ |q| < +\infty$	$p_{\mathbb{X}}(x) = \mathbb{1}_{(0,+\infty)}(x) \frac{1}{m_{\mathbb{X}}} \frac{\delta_{\mathbb{X}}^{-2\delta_{\mathbb{X}}^{-2}}}{\Gamma(\delta_{\mathbb{X}}^{-2})} (x/m_{\mathbb{X}})^{\delta_{\mathbb{X}}^{-2}-1} \exp\left\{-\frac{x/m_{\mathbb{X}}}{\delta_{\mathbb{X}}^2}\right\}$ (**gamma with mean** $m_{\mathbb{X}}$ **and variation coefficient** $\delta_{\mathbb{X}}$)
$(0,+\infty)$	$\mathbb{E}\{\ln(\mathbb{X})\} = \mu \in \mathbb{R}$ $\mathbb{E}\{(\ln(\mathbb{X}) - \mu)^2\} = \sigma^2,\ \sigma > 0$	$p_{\mathbb{X}}(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\}$ $\mu = \ln\left(m_{\mathbb{X}}/\sqrt{1+\delta_{\mathbb{X}}^2}\right)$ $\sigma = \sqrt{\ln\left(1+\delta_{\mathbb{X}}^2\right)}$ (**lognormal with location** μ **and scale** σ)
$(-\infty,+\infty)$	$\mathbb{E}\{\mathbb{X}\} = m_{\mathbb{X}} \in \mathbb{R}$ $\mathbb{E}\{\mathbb{X}^2\} = m_{\mathbb{X}}^2 + \sigma_{\mathbb{X}}^2$	$p_{\mathbb{X}}(x) = \frac{1}{\sqrt{2\pi\sigma_{\mathbb{X}}^2}} \exp\left\{-\frac{(x-m_{\mathbb{X}})^2}{2\sigma_{\mathbb{X}}^2}\right\}$ (**normal with mean** $m_{\mathbb{X}}$ **and variance** $\sigma_{\mathbb{X}}^2$)

which corresponds to the PDF of a uniform distributed random variable over the interval $[a,b]$.

Other cases of interest, where the optimization problem solution is a known distribution, are shown in Table 1. In the fourth line of this table the maximum entropy PDF corresponds to a gamma distribution. Once any gamma random variable has finite variance, and $\mathbb{E}\{\ln(\mathbb{X})\} = q,\ |q| < +\infty$, which implies $\mathbb{E}\{\mathbb{K}^{-2}\} < +\infty$, the known information in this case is equivalent to those listed in Sect. 4.5, required to be satisfied by the distribution of $\mathbb{K}$. For this reason, we presented the gamma distribution as the acceptable choice in Sect. 4.5. It corresponds to the most unbiased choice for that set of information.

For other possible applications of the maximum entropy principle and to go deeper into the underlying mathematics, we recommend the reader to see the references [15, 47–54].

5   Calculation of Uncertainty Propagation

Once one or more of the model parameters are described as random objects, the system response itself becomes random. To understand how the variabilities are transformed by the model, and influence in the response distribution, is a key issue in UQ, known as *uncertainty propagation problem*. This problem can only be attacked after the construction of a consistent stochastic model.

Very succinctly, we understand the uncertainty propagation problem as to determine the probability distribution of model response once we know the distribution of model input/parameters. A schematic representation of this problem is can be seen in Fig. 13.

The methods for calculation of uncertainty of propagation are classified into two types: *non-intrusive* and *intrusive*.

Non-intrusive methods: These methods of stochastic calculation obtain the random problem response by running an associated deterministic problem multiple times (they are also known as sampling methods). In order to use a non-intrusive method, it is not necessary to implement the stochastic model in a new computer code. If a deterministic code to simulate the deterministic model is available, the stochastic simulation can then be performed by running the deterministic program several times, changing only the parameters that are randomly generated [55].

Intrusive methods: In this class of stochastic solvers, the random problem response is obtained by running a customized computer code only once. This code is not based on the associated deterministic model, but on a stochastic version of the computational model [2].

5.1   Monte Carlo Method: A Non-intrusive Approach

The most frequently used technique to compute the propagation of uncertainties of random parameters through a model is the Monte Carlo (MC) method, originally proposed by [56], or one of its variants [57].

An overview of the MC algorithm can be seen in the Fig. 14. First, the MC method generates N realizations (samples) of the random parameters according to their joint

Fig. 13 Schematic representation of uncertainty propagation problem

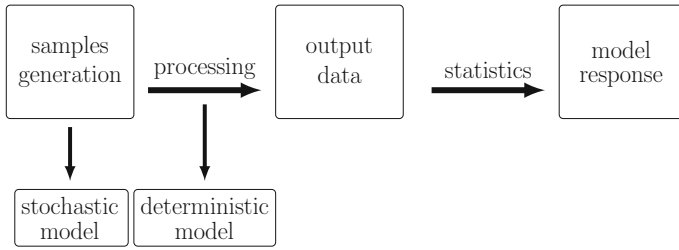Uncertainty Propagation
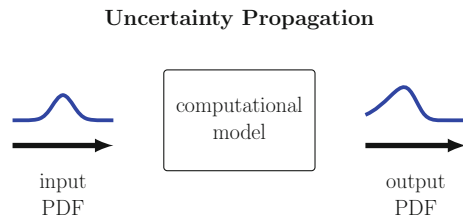
Monte Carlo method

Fig. 14 An overview of monte carlo algorithm

distributions (stochastic model). Each of these realizations defines a deterministic problem which is then solved (processing) using a deterministic technique, generating a certain amount of data. Then, these data are combined through statistics, to access the response of the random system [55, 58]. By the nature of the algorithm, we note that MC is a non-intrusive method.

It can be shown that if N is large enough, the MC method describes very well the statistical behaviour of the random system. However, the rate of convergence of this non-intrusive method is very slow—proportional to the inverse of number of samples square root, i.e., $\sim 1/\sqrt{N}$. Therefore, if the processing time of a single sample is very large, this slow rate of convergence makes MC a very time-consuming method—unfeasible to perform simulation of complex models. Meanwhile, the MC algorithm can easily be parallelized, once each realization can be processed separately and then the results aggregated to compute the statistics [55].

Because of its simplicity and accuracy, MC is the best method to compute the propagation of uncertainties, whenever its use is feasible. Thus, it is recommended that anyone interested in UQ master this technique. Many good references about MC method are available in the literature. For further details, we recommend [58–64].

5.2 Stochastic Galerkin Method: An Intrusive Approach

When the use of MC method is unfeasible, the state of art strategy is based on the so-called stochastic Galerkin method. This spectral approach was originally proposed by [65, 66], and became very popular in the last 15 years, especially after work of [67]. It uses a *Polynomial Chaos Expansion* (PCE) to represent the stochastic model response combined with a Galerkin projection to transform the original stochastic equations into a system of deterministic equations. The resulting unknowns are the coefficients of the linear combination underlying to the PCE.

Once PCE theory is quite rich and extensive, we do not have space in this manuscript to cover it in enough detail, but to the reader interested in digging deeper on this subject is encouraged to see the references [2, 3, 8, 68–70].

6 Concluding Remarks

In this manuscript, we have argued about the importance of modeling and quantification of uncertainties in engineering projects, advocating in favor of the probabilistic approach as a tool to take into account the uncertainties. It is our thought that specifying an envelope of reliability for curves obtained from numerical simulations is an irreversible tendency. We also introduced the basic probabilistic vocabulary to prepare the reader for deeper literature on this subject, and discussed the key points of the stochastic modeling of physical systems, using a simplistic mechanical system as a more in-depth example.

References

1. L. Biegler, G. Biros, O. Ghattas, M. Heinkenschloss, D. Keye, B. Mallick, Y. Marzouk, L. Tenorio, B.B. Waanders, K. Willcox, *Large-Scale Inverse Problems and Quantification of Uncertainty* (Wiley, 2010)
2. O.P. Le Maître, O.M. Knio, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics* (Springer, 2010)
3. D. Xiu, *Numerical Methods for Stochastic Computations: A Spectral Method Approach* (Princeton University Press, 2010)
4. C. Soize, *Stochastic Models of Uncertainties in Computational Mechanics* (American Society of Civil Engineers, 2012)
5. M. Grigoriu, *Stochastic Systems: Uncertainty Quantification and Propagation* (Springer, 2012)
6. R.C. Smith, *Uncertainty Quantification: Theory, Implementation, and Applications* (SIAM, 2013)
7. H. Bijl, D. Lucor, S. Mishra, C. Schwab, *Uncertainty Quantification in Computational Fluid Dynamics* (Springer, 2013)
8. M.P. Pettersson, G. Iaccarino, J. Nordström, *Polynomial Chaos Methods for Hyperbolic Partial Differential Equations: Numerical Techniques for Fluid Dynamics Problems in the Presence of Uncertainties* (Springer, 2015)
9. R. Ohayon, C. Soize, *Advanced Computational Vibroacoustics: Reduced-Order Models and Uncertainty Quantification* (Cambridge University Press, 2015)
10. T.J. Sullivan, *Introduction to Uncertainty Quantification* (Springer, 2015)
11. S. Sarkar, J.A.S. Witteveen, *Uncertainty Quantification in Computational Science* (World Scientific Publishing Company, 2016)
12. R. Ghanem, D. Higdon, H. Owhadi, *Handbook of Uncertainty Quantification* (Springer, 2017)
13. C. Soize, *Uncertainties and Stochastic Modeling* (Short Course at PUC-Rio, Aug 2008)

14. C. Soize, *Stochastic Models in Computational Mechanics* (Short Course at PUC-Rio, Aug 2010)
15. C. Soize, *Probabilité et Modélisation des Incertitudes: Eléments de base et concepts fondamentaux* (Course Notes, Université Paris-Est Marne-la-Vallée, Paris, 2013)
16. G. Iaccarino, A. Doostan, M.S. Eldred, O. Ghattas, Introduction to uncertainty quantification techniques, in *Minitutorial at SIAM CSE Conference*, 2009
17. G. Iaccarino, *Introduction to Uncertainty Quantification* (Lecture at KAUST, 2012)
18. A. Doostan, P. Constantine, *Numerical Methods for Uncertainty Propagation* (Short Course at USNCCM13, 2015)
19. C. Soize, A comprehensive overview of a non-parametric probabilistic approach of model uncertainties for predictive models in structural dynamics. J. Sound Vib. **288**, 623–652 (2005)
20. Guide for the verification and validation of computational fluid dynamics simulations. Technical Report AIAA G-077-1998 (American Institute of Aeronautics and Astronautics, Reston, 1998)
21. W.L. Oberkampf, T.G. Trucano, Verification and validation in computational fluid dynamics. Technical Report SAND 2002-0529 (Sandia National Laboratories, Livermore, 2002)
22. W. Oberkampf, T. Trucano, C. Hirsch, Verification, validation, and predictive capability in computational engineering and physics. Appl. Mech. Rev. **57**, 345–384 (2004)
23. ASME Guide for Verification and Validation in Computational Solid Mechanics. Technical Report ASME Standard V&V 10-2006 (American Society of Mechanical Engineers, New York, 2006)
24. W.L. Oberkampf, C.J. Roy, *Verification and Validation in Scientific Computing* (Cambridge University Press, 2010)
25. U.M. Ascher, C. Greif, *A First Course in Numerical Methods* (SIAM, 2011)
26. P.J. Roache, Code verification by the method of manufactured solutions. J. Fluids Eng. **124**, 4–10 (2001)
27. C.J. Roy, Review of code and solution verification procedures for computational simulation. J. Comput. Phys. **205**, 131–156 (2005)
28. L.A. Petri, P. Sartori, J.K. Rogenski, L.F. de Souza, Verification and validation of a direct numerical simulation code. Comput. Methods Appl. Mech. Eng. **291**, 266–279 (2015)
29. G.I. Schuëller, A state-of-the-art report on computational stochastic mechanics. Probabilistic Eng. Mech. **12**, 197–321 (1997)
30. G.I. Schuëller, Computational stochastic mechanics recent advances. Comput. Struct. **79**, 2225–2234 (2001)
31. C. Soize, Stochastic modeling of uncertainties in computational structural dynamics—recent theoretical advances. J. Sound Vib. **332**, 2379–2395 (2013)
32. D. Moens, D. Vandepitte, A survey of non-probabilistic uncertainty treatment in finite element analysis. Comput. Methods Appl. Mech. Eng. **194**, 1527–1555 (2005)
33. D. Moens, M. Hanss, Non-probabilistic finite element analysis for parametric uncertainty treatment in applied mechanics: recent advances. Finite Elem. Anal. Des. **47**, 4–16 (2011)
34. M. Beer, S. Ferson, V. Kreinovich, Imprecise probabilities in engineering analyses. Mech. Syst. Signal Process. **37**, 4–29 (2013)
35. C. Soize, A nonparametric model of random uncertainties for reduced matrix models in structural dynamics. Probabilistic Eng. Mech. **15**, 277–294 (2000)
36. C. Soize, Generalized probabilistic approach of uncertainties in computational dynamics using random matrices and polynomial chaos decompositions. Int. J. Numer. Methods Eng. **81**, 939–970 (2010)
37. A. Batou, C. Soize, M. Corus, Experimental identification of an uncertain computational dynamical model representing a family of structures. Comput. Struct. **89**, 1440–1448 (2011)
38. G. Grimmett, D. Welsh, *Probability: An Introduction*, 2nd edn. (Oxford University Press, 2014)
39. J. Jacod, P. Protter, *Probability Essentials*, 2nd edn. (Springer, 2004)
40. A. Klenke, *Probability Theory: A Comprehensive Course*, 2nd edn. (Springer, 2014)
41. A. Papoulis, S.U. Pillai, *Probability, Random Variables and Stochastic Processes*, 4th edn. (McGraw-Hill, 2002)

42. C.E. Shannon, A mathematical theory of communication. Bell Syst. Tech. J. **27**, 379–423 (1948)
43. L. Wasserman, *All of Nonparametric Statistics* (Springer, 2007)
44. L. Wasserman, *All of Statistics: A Concise Course in Statistical Inference* (Springer, 2004)
45. E.T. Jaynes, Information theory and statistical mechanics. Phys. Rev. Ser. **II**(106), 620–630 (1957)
46. E.T. Jaynes, Information theory and statistical mechanics II. Phys. Rev. Ser. **II**(108), 171–190 (1957)
47. J.N. Kapur, H.K. Kesavan, *Entropy Optimization Principles with Applications* (Academic Press, 1992)
48. J.N. Kapur, *Maximum Entropy Models in Science and Engineering* (New Age, 2009)
49. F.E. Udwadia, Response of uncertain dynamic systems. I. Appl. Math. Comput. **22**, 115–150 (1987)
50. F.E. Udwadia, Response of uncertain dynamic systems. II. Appl. Math. Comput. **22**, 151–187 (1987)
51. F.E. Udwadia, Some results on maximum entropy distributions for parameters known to lie in finite intervals. SIAM Rev. **31**, 103–109 (1989)
52. K. Sobczyk, J. Trębicki, Maximum entropy principle in stochastic dynamics. Probabilistic Eng. Mech. **5**, 102–110 (1990)
53. K. Sobczyk, J. Trębicki, Maximum entropy principle and nonlinear stochastic oscillators. Phys. A: Stat. Mech. Appl. **193**, 448–468 (1993)
54. J. Trębicki, K. Sobczyk, Maximum entropy principle and non-stationary distributions of stochastic systems. Probabilistic Eng. Mech. **11**, 169–178 (1996)
55. A. Cunha Jr., R. Nasser, R. Sampaio, H. Lopes, K. Breitman, Uncertainty quantification through Monte Carlo method in a cloud computing setting. Comput. Phys. Commun. **185**, 1355–1363 (2014)
56. N. Metropolis, S. Ulam, The Monte Carlo method. J. Am. Stat. Assoc. **44**, 335–341 (1949)
57. C. Lemieux, *Monte Carlo and Quasi-Monte Carlo Sampling* (Springer, 2009)
58. D.P. Kroese, T. Taimre, Z.I. Botev, *Handbook of Monte Carlo Methods* (Wiley, 2011)
59. J.S. Liu, *Monte Carlo Strategies in Scientific Computing* (Springer, 2001)
60. G. Fishman, *Monte Carlo: Concepts, Algorithms, and Applications*, corrected edn. (Springer, 2003)
61. R.Y. Rubinstein, D.P. Kroese, *Simulation and the Monte Carlo Method*, 2nd edn. (Wiley, 2007)
62. S. Asmussen, P.W. Glynn, *Stochastic Simulation: Algorithms and Analysis* (Springer, 2007)
63. R.W. Shonkwiler, F. Mendivil, *Explorations in Monte Carlo Methods* (Springer, 2009)
64. C.P. Robert, G. Casella, *Monte Carlo Statistical Methods* (Springer, 2010)
65. R. Ghanem, P.D. Spanos, Polynomial chaos in stochastic finite elements. J. Appl. Mech. **57**, 197–202 (1990)
66. R. Ghanem, P.D. Spanos, *Stochastic Finite Elements: A Spectral Approach*, 2nd edn. (Dover Publications, 2003)
67. D. Xiu, G.E. Karniadakis, The Wiener-Askey Polynomial Chaos for stochastic differential equations. SIAM J. Sci. Comput. **24**, 619–644 (2002)
68. P. Vos, Time-dependent polynomial chaos. Master Thesis, Delft University of Technology, Delft, 2006
69. P. Constantine, in *A Primer on Stochastic Galerkin Methods*. Lecture Notes, 2007
70. A. O'Hagan, in *Polynomial Chaos: A Tutorial and Critique from a Statistician's Perspective*, (submitted to publication, 2013)