

AN OVERVIEW OF NONINTRUSIVE CHARACTERIZATION, PROPAGATION, AND SENSITIVITY ANALYSIS OF UNCERTAINTIES IN COMPUTATIONAL MECHANICS

Maarten Arnst & Jean-Philippe Ponthot*

Université de Liège, Département d'Aérospatiale et Mécanique, B-4000 Liège, Belgium

Original Manuscript Submitted: 08/02/2013; Final Draft Received: 08/02/2013

In this paper, we offer a short overview of a number of methods that have been reported in the computational-mechanics literature for quantifying uncertainties in engineering applications. Within a probabilistic framework, we describe the characterization of uncertainties using mathematical statistics methods, the propagation of uncertainties through computational models using either Monte Carlo sampling or stochastic expansion methods, and the sensitivity analysis of uncertainties using variance- and differentiation-based methods. We restrict our attention to nonintrusive methods that can be implemented as wrappers around existing computer programs, thus requiring no modification of the source code. We include some recent advances in the propagation and sensitivity analysis of uncertainties that are characterized by arbitrary probability distributions that may exhibit statistical dependence. Finally, we demonstrate the methods integrated in the proposed overview through a nonlinear engineering application relevant to metal forming.

KEY WORDS: *uncertainty quantification, stochastic modeling, representation of uncertainty, Monte Carlo, polynomial chaos, stochastic response surface method, stochastic sensitivity analysis*

1. INTRODUCTION

Advances in sensing technologies, physical modeling, and high-performance computing are profoundly changing the synergistic integration of experiments, physical understanding, and computation into predictive simulations that support scientific discovery and engineering. Chief among these changes is the increasingly central role attributed to the acknowledgement and examination of *experimental, modeling, and computational limitations* that are inevitably present in attempts to simulate complex natural and engineered systems. The field of uncertainty quantification seeks to establish theory, methods, and computer programs for the characterization, propagation, and management of the

*Correspond to: Maarten Arnst, E-mail: maarten.arnst@ulg.ac.be

1 ensuing *parametric uncertainties, modeling errors, and computational errors* in predictive simulations.

2 A number of frameworks are available for uncertainty quantification, such as those based on fuzzy-set theory,
3 interval theory, evidence theory, and probability theory. Here, we adopt a framework based on probability theory,
4 which facilitates a unified treatment of parametric uncertainties and modeling errors.

5 The first step in a probabilistic uncertainty quantification most often involves the use of mathematical statistics
6 methods to *characterize* the uncertain features associated with the model under study as one or more random vari-
7 ables, random fields, random matrices, or random operators. The second step is to use the model to *propagate* this
8 characterization of inputs into a characterization of predictions. This can be achieved in several ways, for example,
9 using either Monte Carlo sampling or stochastic expansion methods. The latter methods most often involve character-
10 izing the predictions as a polynomial chaos expansion. Several approaches are available to calculate the coefficients
11 in this expansion, for example, embedded projection, nonintrusive projection, and interpolatory collocation. Lastly,
12 the third step involves making the probabilistic model useful in the analysis and design of the natural or engineered
13 system under study, for example, by carrying out *sensitivity analyses* to enable the reduction of uncertainties, by using
14 decision-theoretic methods to validate the analysis [1], or by using optimization methods to improve the design [2].

15 In this paper, we draw together a number of methods reported in the computational-mechanics literature for the
16 characterization, propagation, and sensitivity analysis of uncertainties. We restrict our attention to methods that afford
17 a nonintrusive implementation, that is, they can be applied as wrappers around existing computer programs without
18 requiring modification of the source code. We include recent advances in the propagation and sensitivity analysis
19 of uncertainties characterized by arbitrary probability distributions that may exhibit statistical dependence. We show
20 that these recent advances afford significant simplification of the construction of polynomial chaos expansions by
21 bypassing the need to revert to underlying statistically independent Gaussian or other “labeled” random variables.

22 This paper is self-contained in that we define most quantities and concepts when they first appear and we provide
23 enough details relevant to the implementation of the framework.

24 We do not intend this paper to be an exhaustive account of all the methods available for quantifying uncertainty.
25 Rather, we describe in detail a limited set of representative methods to illuminate key features and discern links
26 between the characterization, propagation, and sensitivity analysis steps in a probabilistic uncertainty quantification.

27 The remainder of this paper is organized as follows. In Sec. 2, we briefly discuss our system of notation. Next,
28 in Sec. 3, we outline a model problem, and in Secs. 4, 5, and 6, we present the proposed overview of methods for
29 the characterization, propagation, and sensitivity analysis of uncertainties, respectively. Subsequently, in Sec. 7, we
30 provide implementation details, and finally, in Secs. 8 and 9, we provide an illustration with numerical results.

2. NOTATION

In this paper, we use the following system of notation:

- A lowercase letter, for example, x , is a real deterministic variable.
- A boldface lowercase letter, for example, $\mathbf{x} = (x_1, \dots, x_m)$, is a real deterministic vector.
- An uppercase letter, for example, X , is a real random variable.
- A boldface uppercase letter, for example, $\mathbf{X} = (X_1, \dots, X_m)$, is a real random vector.
- An uppercase letter enclosed between square brackets, for example, $[A]$, is a real deterministic matrix.
- A boldface uppercase letter enclosed between square brackets, for example, $[\mathbf{A}]$, is a real random matrix.

3. MODEL PROBLEM

In applications in computational mechanics, models are built to understand and predict the behavior and evolution of complex natural and engineered systems. These models are implemented in more and more sophisticated computer programs. Often, an implementation of a model in a computer program—which we term, hereafter, a *computational model*—exhibits certain features that may be considered uncertain. The objective of uncertainty quantification lies in the characterization, propagation, and sensitivity analysis of these uncertainties, ultimately allowing quantitative statements about, and some insight into, the impact that these uncertainties have on predictions. This uncertainty quantification may serve to guide the allocation of resources aimed at reducing uncertainties or constitute an essential prerequisite to model validation or design optimization in the presence of uncertainties, among other purposes.

In this paper, we think of a computational model as a (possibly nonlinear) mapping of a set of input variables into a quantity of interest. Correspondingly, we consider the nonlinear mapping

$$y = g(x_1, \dots, x_m), \quad g : \mathbb{R}^m \rightarrow \mathbb{R}, \quad (1)$$

where g is the computational model, $\mathbf{x} = (x_1, \dots, x_m)$ the input variables, and y the *quantity of interest*.

In this paper, without loss of generality, we assume the quantity of interest to be scalar. We note that although easily obtained, we do not consider the extension to a vector-valued quantity of interest for the sake of simplicity of notation and interpretation. Please note that when a vector-valued quantity of interest is considered, an analysis

1 that treats all its components simultaneously can often provide a more informative uncertainty quantification than an
2 analysis that treats these components separately and thus misses information about their mutual dependence [3].

3 For example, if the computational model were a finite element model for the mechanical deformation of a structure,
4 the input variables could be parameters involved in the description of the geometry, the boundary conditions, and the
5 material properties, and the quantity of interest could concern a displacement component at a prescribed location.

6 We assume that the uncertainties that affect the computational model can be associated with uncertainties in the
7 input variables or a subset of these input variables, which we term, hereafter, the *uncertain input variables*. We
8 will comment on the generality of this assumption in Sec. 4.1. Then, the objective lies in the characterization of the
9 uncertainties in the uncertain input variables, the propagation of these uncertainties through the computational model,
10 and the sensitivity analysis of these uncertainties to allow some insight into their impact on the quantity of interest.

11 4. CHARACTERIZATION OF UNCERTAINTIES

12 4.1 Types of probabilistic approach

13 Within the probabilistic framework, several approaches are available for introducing uncertainties into computational
14 models. Common approaches include the following ones:

15 (i) *Parametric approaches* (refer, for example, to [4–8]) are adequate when the uncertain features of the compu-
16 tational model can be associated with uncertainties in some or all of its parameters. Parametric approaches
17 consider parameters to be geometrical characteristics, boundary conditions, loadings, physical or mechanical
18 properties, or any combination of these. Parametric approaches involve the characterization of some or all
19 parameters as random variables, stochastic processes, or both.

20 (ii) By contrast, *nonparametric approaches* are adequate when the uncertain features of the computational model
21 cannot be associated with uncertainties in some or all of the parameters. Such can be the case, for example,
22 when the uncertainties consist of modeling errors, that is, when the uncertainties stem from various modeling
23 assumptions and simplifications whose impact on the quantity of interest is incompletely known. A nonpara-
24 metric approach is an approach that involves the direct characterization of the computational model as a random
25 model without recourse to a characterization of its parameters as random variables, stochastic processes, or both.
26 For example, in structural dynamics, a class of nonparametric models was obtained in [9] by characterizing the
27 reduced matrices of (a sequence of) reduced-order models as random matrices.

28 (iii) *Output-prediction-error approaches* involve adding a random noise term to the quantity of interest [10–12].

(iv) *Generalized approaches* are hybrid approaches that couple parametric and nonparametric approaches. For example, in structural dynamics, a class of generalized models was obtained in [13, 14] by taking into account parametric uncertainties using the parametric approach and modeling errors using the nonparametric approach.

Throughout this paper, the reader may find it easiest to interpret the model problem in the context of a parametric approach, that is, to think of the uncertain input variables $\boldsymbol{x} = (x_1, \dots, x_m)$ as geometrical characteristics, boundary conditions, loadings, physical or mechanical properties, or any combination of these and of the computational model g as providing a mapping from these uncertain input variables to the quantity of interest y . Nevertheless, we note that the model problem affords a level of abstraction that also allows these uncertain input variables to be the entries of reduced matrices and values taken by noise terms. Owing to this level of abstraction, the discussion of the characterization, propagation, and sensitivity analysis steps to follow is very general, and it indeed applies to parametric, nonparametric, output-prediction-error, and generalized approaches alike.

4.2 Types of probabilistic characterization

The characterization of the uncertain input variables as a random variable $\boldsymbol{X} = (X_1, \dots, X_m)$ with values in \mathbb{R}^m requires the characterization of the probability distribution of this random variable:

$$P_{\boldsymbol{X}} = P_{(X_1, \dots, X_m)}. \quad (2)$$

By the *probability distribution* $P_{\boldsymbol{X}}$ of a random variable \boldsymbol{X} with values in \mathbb{R}^m , probability theory understands a function that assigns, to any meaningful subset \mathcal{B} of \mathbb{R}^m , the *probability* $P_{\boldsymbol{X}}(\mathcal{B})$ that the value taken by \boldsymbol{X} is in \mathcal{B} .

We recall that the characterization of the uncertain input variables as a random variable with a certain probability distribution is very general and allows this random variable to be discrete, continuous, or a combination of these. By a *discrete random variable* \boldsymbol{X} with values in \mathbb{R}^m , probability theory understands a random variable \boldsymbol{X} that can take only specific values from a finite or listable subset of \mathbb{R}^m . By a *continuous random variable* \boldsymbol{X} with values in \mathbb{R}^m , probability theory understands a random variable \boldsymbol{X} that may take any value in a subset (or set of subsets) of \mathbb{R}^m and whose probability distribution $P_{\boldsymbol{X}}$ admits a *probability density function* $\rho_{\boldsymbol{X}}$, that is, a function $\rho_{\boldsymbol{X}}$ from \mathbb{R}^m into $\mathbb{R}^+ = [0, +\infty[$ such that $P_{\boldsymbol{X}}(\mathcal{B}) = \int_{\mathcal{B}} \rho_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x}$ for any meaningful subset \mathcal{B} of \mathbb{R}^m .

Further, we recall that probability theory provides various *statistical descriptors*—such as moments, cumulants, and quantiles—that can be deduced from a given probability distribution. The best-known statistical descriptors are the *mean vector* $\boldsymbol{m}_{\boldsymbol{X}}$ and the *covariance matrix* $[C_{\boldsymbol{X}}]$ given by $\boldsymbol{m}_{\boldsymbol{X}} = \int_{\mathbb{R}^m} \boldsymbol{x} dP_{\boldsymbol{X}}$ and $[C_{\boldsymbol{X}}] = \int_{\mathbb{R}^m} (\boldsymbol{x} -$

1 $\mathbf{m}_X)(\mathbf{x} - \mathbf{m}_X)^T dP_X$ assuming these integrals are bounded; otherwise, the mean vector and the covariance matrix
 2 do not exist. When the random variable \mathbf{X} is continuous, the expressions for \mathbf{m}_X and $[C_X]$ (if they exist) read
 3 as $\mathbf{m}_X = \int_{\mathbb{R}^m} \mathbf{x} \rho_X(\mathbf{x}) d\mathbf{x}$ and $[C_X] = \int_{\mathbb{R}^m} (\mathbf{x} - \mathbf{m}_X)(\mathbf{x} - \mathbf{m}_X)^T \rho_X(\mathbf{x}) d\mathbf{x}$.

4 Two characterizations of the required probability distribution P_X are most often encountered:

5 (i) The first type involves the direct characterization of P_X . Specifically, a characterization of this type is most
 6 often obtained by selecting a probability distribution that depends on a finite number of parameters and then
 7 assigning adequate values to these parameters. For example, if a Gaussian probability distribution was selected
 8 and adequate values assigned to its mean and variance, a characterization of this first type would be obtained.

9 On the one hand, the parameter-dependent probability distribution can be selected as a “labeled” probability
 10 distribution that depends on only a small number of parameters, for example, by selecting it as a Gaussian or
 11 uniform probability distribution. On the other hand, it can be obtained by expressing the required probability
 12 distribution in a versatile manner as a function of a very large number of parameters, for example, by selecting it
 13 as a mixture model [15]. We note that catalogs of available “labeled” probability distributions can be found, for
 14 example, in [16]; if an adequate “labeled” probability distribution is not available, the possibility of constructing
 15 a new, adequate probability distribution can be considered using, for example, the maximum entropy principle,
 16 limit theorems, or coarse-graining or other approaches.

17 (ii) The second type involves an indirect characterization of the required probability distribution P_X . This type
 18 relies on the fact that the probability distribution of a random variable changes when this random variable is
 19 transformed under a (possibly nonlinear) mapping. A characterization of the second type is most often obtained
 20 by fixing the probability distribution P_{Ξ} of a random variable $\Xi = (\Xi_1, \dots, \Xi_d)$ with values in \mathbb{R}^d and then
 21 characterizing the uncertain input variables as a transformation $\mathbf{X} = \mathbf{f}(\Xi)$ of Ξ under a (possibly nonlinear)
 22 mapping \mathbf{f} from \mathbb{R}^d to \mathbb{R}^m , thus implying the probability distribution P_X as the image of P_{Ξ} under \mathbf{f} . By the
 23 image P_X of P_{Ξ} under \mathbf{f} , probability theory understands the probability distribution P_X that assigns, to any
 24 meaningful subset \mathcal{B} of \mathbb{R}^m , the probability $P_X(\mathcal{B}) = P_{\Xi}(\{\xi \in \mathbb{R}^d : \mathbf{f}(\xi) \in \mathcal{B}\})$.

25 The probability distribution P_{Ξ} of the random variable Ξ with values in \mathbb{R}^d is most often fixed as a uniform,
 26 Gaussian, or other “labeled” probability distribution. The mapping \mathbf{f} is most often obtained by selecting a
 27 mapping that depends on a finite number of parameters and then assigning adequate values to these parameters.
 28 As in the first approach, a parameter-dependent mapping can be obtained either by selecting a “labeled” mapping
 29 that depends on only a small number of parameters [6, 17] or by expressing the required mapping in a versatile

manner as a function of a very large number of parameters, for example, by expressing the required mapping as a high-order polynomial [18–22]. We note that Ξ need not have as many components as \mathbf{X} , that is, d may be smaller than m , thus indicating the usefulness of this second type of characterization in obtaining a reduced-dimensional characterization of the uncertain input variables [6, 17].

These two types of characterization are strongly interrelated. From a theoretical point of view, under certain conditions, either one can be converted into the other using, for example, the Rosenblatt transformation [23]. From a computational point of view, the literature on random number generation provides generators for many (possibly multivariate) probability distributions (refer, for example, to [24]). A *generator* is a method for computing an ensemble of samples from a given probability distribution starting from an ensemble of samples from a (possibly multivariate) uniform, Gaussian, or other “labeled” probability distribution. A generator most often computes each sample for the given probability distribution by transforming each corresponding sample for the uniform, Gaussian, or other “labeled” probability distribution through an appropriate mapping. Thus, by identifying the given probability distribution with $P_{\mathbf{X}}$, the uniform, Gaussian, or other “labeled” probability distribution with P_{Ξ} , and the mapping with f , a generator provides, for a characterization of the first type, an equivalent characterization of the second type.

4.3 Types of interpretation of probability

The following two types of interpretation of probability are most often encountered:

- (i) The first type involves the well-known interpretation of probability as describing *variability* or, equivalently, as referring to a frequency of occurrence.
- (ii) The second type involves the interpretation of probability as describing a *state of (possibly incomplete) knowledge*, construing high probabilities as descriptors of possibilities that are most strongly indicated by this state of knowledge and lower probabilities as descriptors of less plausible alternatives. The interpretation of probability as describing a state of knowledge allows probability theory to be developed as a logical framework for inference and decision making in the presence of incomplete knowledge and missing information.

4.4 Types of available information

In applications in computational mechanics, the *available information*—from which the characterization of the uncertain input variables must be inferred—most often consists of the following sources of information:

1 (i) First, *data* are most often available either in the form of results obtained from newly conducted real experiments
2 or in the form of higher fidelity computational models.

3 (ii) Next, in most applications in computational mechanics, there are mechanical and physical laws that apply to
4 the natural or engineered system under consideration. Most often, the applicable laws impose *mechanical*
5 *and physical constraints* that restrict the values that the uncertain input variables may take. Examples include
6 positiveness and symmetry constraints imposed on mechanical properties involved in the description of the con-
7 stitutive behavior of materials [17, 25] and positiveness, symmetry, stability, and causality constraints imposed
8 on reduced matrices and other features of reduced-order models for the dynamical behavior of structures and
9 other systems [8, 9, 26–31]. These constraints act as sources of information when a characterization of the
10 uncertain input variables of a computational model is being inferred because in order to be consistent with the
11 applicable mechanical and physical laws, this characterization of these uncertain input variables must assign a
12 vanishing probability to those values of these uncertain input variables that do not satisfy these constraints.

13 A frequently encountered example is that mechanics and physics require Young’s modulus—a mechanical prop-
14 erty involved in the description of linearly elastic isotropic constitutive behavior of materials—to be positive;
15 thus, in order for a characterization of an uncertain Young modulus to be consistent with mechanics and physics,
16 it must assign a vanishing probability to negative values of this Young modulus.

17 (iii) Various *other sources of information* can also contribute to the available information, for example, previous
18 studies of the computational-mechanics application under consideration. The combined information provided by
19 the mechanical and physical constraints and these other sources of information is often referred to as the *prior*
20 *information* [19, 20, 22, 32–38].

21 Even though they are of a different nature than the sources of information mentioned previously, the following
22 considerations may also play a role in obtaining the characterization of the uncertain input variables:

23 (iv) It is desirable that the characterization of the uncertain input variables provides a context that allows the subse-
24 quent propagation and sensitivity analysis steps to be implemented efficiently and in a manner wherein compu-
25 tational errors are amenable to analysis and can be decreased by expending a higher computational effort.

26 (v) It is desirable that the characterization of the uncertain input variables is parameterized so as to allow those of
27 its features to be varied with respect to which informative sensitivity analyses can be conducted.

4.5 Characterization using mathematical statistics methods

The previous sections indicate that probability theory provides significant freedom as to the type of input variables in which uncertainties can be introduced (Sec. 4.1), the type of characterization that can be adopted for the uncertain input variables (Sec. 4.2), and the type of interpretation that can be conferred on probabilities (Sec. 4.3). Correspondingly, research in computational mechanics has yielded many methods for the characterization of uncertainties, each one conforming to some interpretation and allowing some characterization of uncertain input variables to be deduced from available information. Providing an exhaustive account of all available methods and listing yet-unfinished ones is beyond the scope of this paper; instead, we confine ourselves to a succinct presentation of some of the fundamental methods involving an interplay between modeling considerations and mathematical statistics methods:

(i) In an application wherein probabilities are held to refer to variability, so called “frequentist” mathematical statistics methods (refer, for example, to [39, 40]) can be applied as follows:

(i.1) When a very large amount of data is available, in addition to physical and mechanical constraints, one of the previously mentioned versatile characterizations involving a very large number of parameters can be selected. Adequate values for this very large number of parameters can then be inferred from the very large amount of data using, for example, the kernel density estimation method [39], which can use coordinate transformations or local bandwidth adaptations to account for mechanical and physical constraints.

(i.2) When only a small amount of data is available, in addition to physical and mechanical constraints, the information contained in the data may be too vague to allow a very large number of parameters to be accurately inferred. Then, a “labeled” characterization involving only a small number of parameters can be selected, for example, from one of the available catalogs [16], and adequate values for these parameters can be inferred from the data using, for example, the method of maximum likelihood [40], which is an often used parameter-estimation method from “frequentist” mathematical statistics.

Care should be taken to select a “labeled” characterization that is consistent with the mechanical and physical constraints; for example, the Gaussian probability distribution should not be selected to characterize an uncertain Young modulus because its support is the whole real line and its selection would thus lead to the assignment of a nonvanishing probability to negative values of this Young modulus. If no adequate “labeled” characterization is available, the possibility of constructing a new, adequate one can be considered using, for example, the maximum entropy principle, which allows mechanical and physical constraints to be explicitly taken into account, or limit theorems, coarse graining, or other approaches.

(ii) By contrast, in an application wherein probabilities are held to describe a state of knowledge, Bayesian mathematical statistics methods (refer, for example, to [19, 20, 22, 32–38]) can be applied. These involve an initial encoding of all the available information except for newly available data—that is, the mechanical and physical constraints in addition to the other sources of information mentioned previously—into a *prior probability distribution*. This prior probability distribution is then updated by accounting for newly available data in accordance with Bayes’s rule to obtain a *posterior probability distribution*.

4.6 Characterization of uncertainties in high-dimensional problems

The characterization of uncertainties is especially challenging in high-dimensional problems, that is, when the number of uncertain input variables is large. In fact, when a probabilistic characterization must be inferred for a large number of uncertain input variables, one of the most significant challenges is in constructing this probabilistic characterization in a manner that is consistent with the applicable mechanical and physical constraints. Thus, in recent years, the focus of much research has been on investigating how this characterization can be effected—using mathematical statistics methods such as those that rely on the maximum entropy principle—in a manner that allows mechanical and physical constraints to be explicitly taken into account. Please refer to [17, 25] for examples of positiveness and symmetry constraints imposed on mechanical properties involved in the description of the constitutive behavior of materials and to [8, 9, 26–31] for examples of positiveness, symmetry, stability, and causality constraints imposed on reduced matrices and other features of reduced-order models for the dynamical behavior of structures and other systems.

In addition, recent research has also investigated dimension reduction methods [6, 41] and scalable algorithms [21, 22, 36, 42] as viable strategies for addressing challenges in characterizing uncertainties in high-dimensional problems.

5. PROPAGATION OF UNCERTAINTIES

The next step is to propagate the uncertainties introduced in the input variables through the computational model to the quantity of interest. Probability theory effects this propagation as follows. Once the uncertain input variables are characterized as a random variable \mathbf{X} with values in \mathbb{R}^m , the transformation of \mathbf{X} through the computational model g provides the characterization of the quantity of interest as the random variable Y with values in \mathbb{R} such that

$$Y = g(X_1, \dots, X_m); \quad (3)$$

1 this definition of Y as the transformation of \mathbf{X} through g implies that the probability distribution P_Y of Y is the image
 2 of the probability distribution $P_{\mathbf{X}}$ of \mathbf{X} under g , that is, formally, for any meaningful subset \mathcal{B} of \mathbb{R} ,

$$P_Y(\mathcal{B}) = P_{\mathbf{X}}(\{\mathbf{x} \in \mathbb{R}^m : g(\mathbf{x}) \in \mathcal{B}\}). \quad (4)$$

3 In other words, to obtain the characterization of the quantity of interest, probability theory equates the probabil-
 4 ity $P_Y(\mathcal{B})$ that its value is contained in any meaningful subset \mathcal{B} of \mathbb{R} with the probability $P_{\mathbf{X}}(\{\mathbf{x} \in \mathbb{R}^m : g(\mathbf{x}) \in \mathcal{B}\})$
 5 that the value taken by the uncertain input variables is contained in the corresponding subset $\{\mathbf{x} \in \mathbb{R}^m : g(\mathbf{x}) \in \mathcal{B}\}$
 6 of \mathbb{R}^m , which collects those values in \mathbb{R}^m that the computational model maps into values in \mathcal{B} .

7 By deducing various *statistical descriptors* from it, we can study the probability distribution P_Y . The best known
 8 statistical descriptors are the *mean* m_Y and the *variance* σ_Y^2 , which are defined as follows:

$$m_Y = \int_{\mathbb{R}} y dP_Y, \quad \sigma_Y^2 = \int_{\mathbb{R}} (y - m_Y)^2 dP_Y, \quad (5)$$

9 assuming these integrals are bounded; otherwise, they do not exist. Because integrals with respect to a probability
 10 distribution and an image of it are related by the ‘‘change of variables’’ theorem [43, 44], the mean and variance (if
 11 they exist) are also obtained as $m_Y = \int_{\mathbb{R}^m} g(\mathbf{x}) dP_{\mathbf{X}}$ and $\sigma_Y^2 = \int_{\mathbb{R}^m} (g(\mathbf{x}) - m_Y)^2 dP_{\mathbf{X}}$.

12 The propagation of uncertainties is most often implemented using either the Monte Carlo sampling method or
 13 stochastic expansion methods, as described next.

14 5.1 Monte Carlo sampling method

15 The Monte Carlo sampling method begins by generating an ensemble of ν independent and identically distributed
 16 (i.i.d.) samples from the probability distribution $P_{\mathbf{X}}$, written as follows:

$$\{\mathbf{x}_\ell, 1 \leq \ell \leq \nu\}. \quad (6)$$

17 The computational model is then used to map each sample from $P_{\mathbf{X}}$ into the corresponding sample from P_Y , that is,

$$y_\ell = g(\mathbf{x}_\ell), \quad 1 \leq \ell \leq \nu, \quad (7)$$

1 to obtain the corresponding ensemble of i.i.d. samples from P_Y , written as follows:

$$\{y_\ell, 1 \leq \ell \leq \nu\}. \quad (8)$$

2 Once these i.i.d. samples from P_Y are available, various statistical descriptors can be approximated using methods
3 from mathematical statistics. For example, the mean m_Y and the variance σ_Y^2 (if they exist) can be approximated as

$$m_Y \approx m_Y^\nu = \frac{1}{\nu} \sum_{\ell=1}^{\nu} y_\ell, \quad \sigma_Y^2 \approx (\sigma_Y^\nu)^2 = \frac{1}{\nu} \sum_{\ell=1}^{\nu} (y_\ell - m_Y^\nu)^2. \quad (9)$$

4 This implementation is *nonintrusive* because it requires only the repeated solution—sequentially or in parallel—of
5 the computational model for different values assigned to its uncertain input variables; the computational model itself
6 need not be modified. Because the computational model must be solved for each of the i.i.d. samples from P_X , the
7 computational cost of the Monte Carlo sampling method scales with the number of samples in the ensemble.

8 For many probability distributions, the literature on random number generation (refer, for example, to [24]) pro-
9 vides generators that can be used to obtain the required ensemble of i.i.d. samples from P_X . In addition, the literature
10 provides principles of construction, such as those based on the Rosenblatt transformation mentioned previously, for
11 use to obtain a generator if one should not be already available.

12 From a theoretical point of view, the law of large numbers and the central limit theorem (refer, for example,
13 to [24, 43, 44]) can be used to analyze the convergence of approximations of statistical descriptors of the quantity of
14 interest such as those in (9) with respect to the number of samples. Under certain conditions, the central limit theorem
15 ensures that accuracy improves with the square root of the number of samples. For example, if m_Y exists, the law
16 of large numbers ensures that the approximation m_Y^ν of the mean converges to the exact value m_Y as the number of
17 samples ν increases, and if σ_Y^2 exists, the central limit theorem ensures that the accuracy of m_Y^ν improves with the
18 square root of ν . From a computational point of view, numerical convergence studies can be conducted.

19 Regarding desideratum (iv) in Sec. 4.4, we note that the availability of a generator and the fulfillment of conditions
20 that imply convergence properties can play a role in obtaining the characterization of the uncertain input variables.

21 5.2 Advanced Monte Carlo sampling methods

22 Much recent research has investigated how the computational cost of the Monte Carlo sampling method (Sec. 5.1) can
23 be reduced. Advanced Monte Carlo sampling methods have been proposed, which can reduce the number of samples
24 that must be used to achieve a target accuracy; please refer, for example, to [24] and [45] for details about methods

1 involving antithetic variables, control variates, importance sampling, stratified sampling, Latin hypercube sampling,
 2 and quasi-Monte Carlo sampling. In addition, recent research has investigated the use of these advanced Monte Carlo
 3 sampling methods in concert with reduced-order models; please refer, for example, to [46] and [47] for details about
 4 the multilevel Monte Carlo method and the reduced-basis control-variate Monte Carlo method.

5 5.3 Stochastic expansion methods

6 Stochastic expansion methods most often involve two steps. First, a surrogate model is fitted to the computational
 7 model; then, the characterization of the uncertain input variables is mapped through this surrogate model—instead
 8 of through the computational model—into the characterization of the quantity of interest. A *surrogate model* is any
 9 model that mimics the relationship that the computational model establishes between the uncertain input variables and
 10 the quantity of interest and yet is computationally less expensive. Thus, the attraction of using a surrogate model in
 11 the propagation of uncertainties is most often in gaining a computational speedup. Although many types of surrogate
 12 model have been proposed, polynomial surrogate models are most often encountered.

13 5.3.1 Surrogate model

14 A succinct definition of a polynomial surrogate model is as follows. Let elements $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_m)$ of \mathbb{N}^m be
 15 referred to as *multi-indices* and let a (multivariate) *monomial* \boldsymbol{x}^α associated with a multi-index $\boldsymbol{\alpha}$ be a function
 16 from \mathbb{R}^m into \mathbb{R} defined by $\boldsymbol{x}^\alpha = x_1^{\alpha_1} \times \dots \times x_m^{\alpha_m}$. Let the number $|\boldsymbol{\alpha}| = \alpha_1 + \dots + \alpha_m$ be referred to as the *modulus*
 17 of $\boldsymbol{\alpha}$ and also as the *order* of \boldsymbol{x}^α . Let a (multivariate) *polynomial* be a function from \mathbb{R}^m into \mathbb{R} that maps any \boldsymbol{x} to a
 18 finite sum $\sum_{\boldsymbol{\alpha}} d_{\boldsymbol{\alpha}} \boldsymbol{x}^\alpha$ with real coefficients $d_{\boldsymbol{\alpha}}$. Then, a polynomial surrogate model g^p of order p —hereafter, termed
 19 a *surrogate model*, unless the nature and the order of the polynomial require emphasis—is an m -variate polynomial
 20 that approximates the computational model as precisely as possible in the $P_{\boldsymbol{X}}$ -weighted least-squares sense,

$$g \approx g^p = \sum_{|\boldsymbol{\alpha}|=0}^p c_{\boldsymbol{\alpha}} \boldsymbol{x}^\alpha, \quad \text{where } \boldsymbol{c} = \text{solution of } \min_{\boldsymbol{d} \in \mathbb{R}^\mu} \frac{1}{2} \int_{\mathbb{R}^m} \left| g(\boldsymbol{x}) - \sum_{|\boldsymbol{\alpha}|=0}^p d_{\boldsymbol{\alpha}} \boldsymbol{x}^\alpha \right|^2 dP_{\boldsymbol{X}}, \quad (10)$$

21 where $\boldsymbol{c} = \{c_{\boldsymbol{\alpha}}, 0 \leq |\boldsymbol{\alpha}| \leq p\}$ collects the coefficients and μ is the number of monomials in $\{\boldsymbol{x}^\alpha, 0 \leq |\boldsymbol{\alpha}| \leq p\}$.
 22 Because the objective function of the optimization problem in (10) gauges the precision of the approximation in
 23 the $P_{\boldsymbol{X}}$ -weighted least-squares sense, the precision of the approximation can be expected to be higher over subsets of
 24 values of the uncertain input variables to which $P_{\boldsymbol{X}}$ assigns a higher probability.

25 It follows from standard optimization theory results—specifically, from the results that underpin the projection

1 theorem [48]—that the optimization problem in (10) has at least one solution if the computational model, as well as
 2 all polynomials of order at most p , are $P_{\mathbf{X}}$ -square-integrable, that is,

$$\int_{\mathbb{R}^m} g(\mathbf{x})^2 dP_{\mathbf{X}} < +\infty, \quad (11)$$

3

$$\int_{\mathbb{R}^m} (\mathbf{x}^{\boldsymbol{\alpha}})^2 dP_{\mathbf{X}} < +\infty, \quad 0 \leq |\boldsymbol{\alpha}| \leq p, \quad (12)$$

4 and that the solution is unique if the following additional condition is also fulfilled:

$$\int_{\mathbb{R}^m} \left(\sum_{|\boldsymbol{\alpha}|=0}^p d_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \right)^2 dP_{\mathbf{X}} > 0, \quad \forall \mathbf{d} \in \mathbb{R}^{\mu}, \quad \mathbf{d} \neq \mathbf{0}. \quad (13)$$

5 In fact, the conditions in (11) and (12) ensure that the objective function in (10) is continuous and that given any coeffi-
 6 cients $\tilde{\mathbf{d}}$, the optimization problem in (10) can be formulated equivalently as the optimization of this objective function
 7 over the set of all coefficients \mathbf{d} that satisfy $\int_{\mathbb{R}^m} |g(\mathbf{x}) - \sum_{|\boldsymbol{\alpha}|=0}^p d_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}|^2 dP_{\mathbf{X}} \leq \int_{\mathbb{R}^m} |g(\mathbf{x}) - \sum_{|\boldsymbol{\alpha}|=0}^p \tilde{d}_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}|^2 dP_{\mathbf{X}}$.
 8 This equivalent formulation converts the optimization problem in (10) into the optimization of a continuous objective
 9 function over a closed and bounded set and therefore guarantees the existence of a solution. The condition in (13)
 10 ensures that the objective function in (10) is strictly convex, thus guaranteeing uniqueness.

11 We note that the condition in (12) amounts to requiring $P_{\mathbf{X}}$ to have bounded moments of any order up to $2p$.
 12 Further, it can be shown that the condition in (13) is fulfilled for any order p if $P_{\mathbf{X}}$ is a continuous probability
 13 distribution; however, it is not necessarily fulfilled when $P_{\mathbf{X}}$ is a discrete probability distribution [49].

14 From a theoretical point of view, results from approximation theory (refer, for example, to [49, 50]) can be used
 15 to analyze the convergence of the surrogate model with respect to its order. If the conditions in (11)–(13) are fulfilled
 16 for any order, it is desirable that the precision with which the surrogate model approximates the computational model
 17 can be improved systematically and made arbitrarily high by increasing the order p , that is,

$$\lim_{p \rightarrow +\infty} \int_{\mathbb{R}^m} |g(\mathbf{x}) - g^p(\mathbf{x})|^2 dP_{\mathbf{X}} = 0. \quad (14)$$

18 Unfortunately, this convergence property does not always hold, its presence being dependent on the probability distri-
 19 bution $P_{\mathbf{X}}$. In this regard, the following two results are available. First, it follows immediately from the Weierstrass
 20 theorem that this convergence property holds if $P_{\mathbf{X}}$ has a closed and bounded support. Second, and more generally,
 21 the holomorphicity properties of integral transformations can be used to show that this convergence property holds if

1 there exists a constant $\beta > 0$ such that $\int_{\mathbb{R}^m} \exp(\beta \|\mathbf{x}\|) dP_{\mathbf{X}} < +\infty$ [49, 50]; here, $\|\cdot\|$ is the Euclidean norm. For
 2 example, if $P_{\mathbf{X}}$ were a (possibly multivariate) uniform probability distribution, this convergence property would hold
 3 because of the first result, and if $P_{\mathbf{X}}$ were a (possibly multivariate) Gaussian probability distribution, it would hold
 4 because of the second result. The surrogate model usually converges to the computational model rapidly if the latter
 5 is sufficiently smooth. From a computational point of view, numerical convergence studies can be conducted.

6 Regarding desideratum (iv) in Sec. 4.4, we note that the fulfillment of conditions that imply convergence with
 7 respect to the order can play a role in obtaining the characterization of the uncertain input variables.

8 5.3.2 Propagation of uncertainties

9 Once a surrogate model is available, stochastic expansion methods most often map the characterization of the un-
 10 certain input variables through this surrogate model—instead of through the computational model—to obtain the
 11 characterization of the quantity of interest. Unfortunately, unless the surrogate model perfectly mimics the computa-
 12 tional model, the use of the surrogate model introduces an approximation error in the characterization of the quantity
 13 of interest: the transformation of \mathbf{X} through the surrogate model g^p provides, as an approximation of the random
 14 variable Y , the characterization of the quantity of interest as the random variable Y^p with values in \mathbb{R} such that

$$Y \approx Y^p = g^p(\mathbf{X}), \quad (15)$$

15 thus implying, as an approximation of the probability distribution P_Y of Y , the probability distribution P_Y^p of Y^p as
 16 the image of the probability distribution $P_{\mathbf{X}}$ of \mathbf{X} by g^p , that is, for any meaningful subset \mathcal{B} of \mathbb{R} ,

$$P_Y(\mathcal{B}) \approx P_Y^p(\mathcal{B}) = P_{\mathbf{X}}(\{\mathbf{x} \in \mathbb{R}^m : g^p(\mathbf{x}) \in \mathcal{B}\}). \quad (16)$$

17 The desirability of the convergence property in (14) is further emphasized by the fact that owing to the relationships
 18 that probability theory establishes between probabilistic modes of convergence [43, 44], the convergence of the sur-
 19rogate model g^p to the computational model g in the $P_{\mathbf{X}}$ -weighted least-squares sense implies the convergence of the
 20 probability distribution P_Y^p of $Y^p = g^p(\mathbf{X})$ to the probability distribution P_Y of $Y = g(\mathbf{X})$.

1 5.3.3 Polynomial chaos expansion

2 Let $\{\psi_{\alpha}, 0 \leq |\alpha| \leq p\}$ be a set of polynomials ψ_{α} that span the set of all polynomials of order at most p and
 3 are $P_{\mathbf{X}}$ -orthonormal, that is,

$$\int_{\mathbb{R}^m} \psi_{\alpha}(\mathbf{x})\psi_{\beta}(\mathbf{x})dP_{\mathbf{X}} = \delta_{\alpha\beta}, \quad 0 \leq |\alpha|, |\beta| \leq p, \quad (17)$$

4 where $\delta_{\alpha\beta} = 1$ if $\alpha = \beta$ and $\delta_{\alpha\beta} = 0$ otherwise. Such a set is guaranteed to exist if the conditions in (12) and (13)
 5 are fulfilled, but if it exists, it is not unique. Because $\{\psi_{\alpha}, 0 \leq |\alpha| \leq p\}$ spans the set of all polynomials of order at
 6 most p , the optimization problem in (10) is equivalent to

$$g \approx g^p = \sum_{|\alpha|=0}^p g_{\alpha}\psi_{\alpha}, \quad \text{where } \mathbf{g} = \text{solution of } \min_{\mathbf{h} \in \mathbb{R}^{\mu}} \frac{1}{2} \int_{\mathbb{R}^m} \left| g(\mathbf{x}) - \sum_{|\alpha|=0}^p h_{\alpha}\psi_{\alpha}(\mathbf{x}) \right|^2 dP_{\mathbf{X}}, \quad (18)$$

7 where $\mathbf{g} = \{g_{\alpha}, 0 \leq |\alpha| \leq p\}$ and μ is the number of $P_{\mathbf{X}}$ -orthonormal polynomials in $\{\psi_{\alpha}, 0 \leq |\alpha| \leq p\}$. The
 8 attraction of this reformulation is that upon expressing the stationarity of the objective function in (18), the following
 9 constructive representation of the surrogate model is obtained:

$$g \approx g^p = \sum_{|\alpha|=0}^p g_{\alpha}\psi_{\alpha}, \quad \text{where } g_{\alpha} = \int_{\mathbb{R}^m} g(\mathbf{x})\psi_{\alpha}(\mathbf{x})dP_{\mathbf{X}}, \quad 0 \leq |\alpha| \leq p. \quad (19)$$

10 By constructive, we mean that this representation can facilitate the implementation of the construction of the surrogate
 11 model. Further, this representation allows the random variable Y^p defined previously to be written as follows:

$$Y \approx Y^p = g^p(\mathbf{X}) = \sum_{|\alpha|=0}^p g_{\alpha}\psi_{\alpha}(\mathbf{X}). \quad (20)$$

12 Stochastic expansion methods refer to this characterization of the quantity of interest, $Y^p = \sum_{|\alpha|=0}^p g_{\alpha}\psi_{\alpha}(\mathbf{X})$,
 13 as a *polynomial chaos expansion*. Further, the coefficients g_{α} are referred to as *polynomial chaos coefficients* and
 14 the $P_{\mathbf{X}}$ -orthonormal polynomials ψ_{α} as *polynomial chaos*.

15 5.3.4 Implementation using nonintrusive projection methods

16 Several types of implementation are available for obtaining the surrogate model. Embedded projection methods [6,
 17 51–55], nonintrusive projection methods [51, 53–56], and interpolatory collocation methods [53–55, 57–62] are most

1 often encountered. *Embedded projection methods* numerically determine the surrogate model through a Galerkin
 2 projection of the computational model onto a prescribed basis of orthonormal polynomials; this Galerkin projection
 3 results in a so-called spectral problem whose assembly and solution most often requires modification of the source
 4 code of the computational model. *Nonintrusive projection methods* rely on (multivariate) integration theory: they
 5 numerically determine the surrogate model through an orthogonal projection that involves the use of a quadrature rule
 6 to approximate integrals with respect to the probability distribution of the uncertain input variables. *Interpolatory*
 7 *collocation methods* rely on (multivariate) interpolation theory: they numerically determine the surrogate model by
 8 interpolating between a set of solutions to the computational model. Whereas embedded projection methods require
 9 modification of the computational model, nonintrusive projection and interpolatory collocation methods can be applied
 10 as wrappers around an existing computational model without requiring modification of the source code.

11 We note that the literature sometimes refers to nonintrusive projection methods also as *pseudospectral collocation*
 12 *methods* [54, 56]; further, the literature sometimes refers to nonintrusive projection and interpolatory collocation
 13 methods collectively as *collocation methods* [54, 56].

14 In this paper, we describe only nonintrusive projection methods; we refer the reader to the references mentioned
 15 previously for details about embedded projection, interpolatory collocation, and other methods.

16 Nonintrusive projection methods approximate integrals with respect to the probability distribution $P_{\mathbf{X}}$ using a
 17 quadrature rule. A *quadrature rule* for integration with respect to $P_{\mathbf{X}}$ is a set $\{(\mathbf{x}_\ell^\lambda, w_\ell^\lambda), 1 \leq \ell \leq \lambda\}$ of nodes \mathbf{x}_ℓ^λ and
 18 associated weights w_ℓ^λ that allow the integral of any continuous, integrable function f from \mathbb{R}^m into \mathbb{R} with respect
 19 to $P_{\mathbf{X}}$ to be approximated by a weighted sum of integrand evaluations as $\int_{\mathbb{R}^m} f(\mathbf{x}) dP_{\mathbf{X}} \approx \sum_{\ell=1}^{\lambda} w_\ell^\lambda f(\mathbf{x}_\ell^\lambda)$.

20 A first nonintrusive projection method exploits the fact that for certain “labeled” probability distributions $P_{\mathbf{X}}$,
 21 recurrence relations that can be used to produce sets $\{\psi_\alpha, 0 \leq |\alpha| \leq p\}$ of $P_{\mathbf{X}}$ -orthonormal polynomials ψ_α
 22 are explicitly known and can be read from tables in the literature [6, 51, 53, 54]. For example, the tabulated (ten-
 23 sorized) normalized Hermite and Legendre polynomials constitute sequences of (multivariate) polynomials that are
 24 of increasing order and orthonormal with respect to the (multivariate product of) univariate standard Gaussian and
 25 uniform probability distributions, respectively. When such a set of $P_{\mathbf{X}}$ -orthonormal polynomials is explicitly known
 26 in advance, this first nonintrusive projection method provides the surrogate model as follows:

$$g \approx g^{p,\lambda} = \sum_{|\alpha|=0}^p g_\alpha^\lambda \psi_\alpha, \quad \text{where} \quad g_\alpha^\lambda = \sum_{\ell=1}^{\lambda} w_\ell^\lambda g(\mathbf{x}_\ell^\lambda) \psi_\alpha(\mathbf{x}_\ell^\lambda), \quad 0 \leq |\alpha| \leq p, \quad (21)$$

27 that is, it provides the surrogate model by approximating the integral with respect to $P_{\mathbf{X}}$ involved in the definition of

1 the coefficients in (20) using the quadrature rule for integration with respect to $P_{\mathbf{X}}$.

2 A second nonintrusive projection method, which does not require a set of $P_{\mathbf{X}}$ -orthonormal polynomials to be
 3 explicitly known in advance, is obtained by approximating the integral with respect to $P_{\mathbf{X}}$ involved in the definition
 4 of the objective function in (10) using the quadrature rule for integration with respect to $P_{\mathbf{X}}$. This second nonintrusive
 5 projection method provides the surrogate model as a solution of the following weighted least-squares problem:

$$g \approx \tilde{g}^{p,\lambda} = \sum_{|\alpha|=0}^p c_{\alpha}^{\lambda} \mathbf{x}^{\alpha}, \quad \text{where } \mathbf{c}^{\lambda} = \text{solution of } \min_{\mathbf{d} \in \mathbb{R}^{\mu}} \frac{1}{2} (\mathbf{y} - [M]\mathbf{d})^T [W] (\mathbf{y} - [M]\mathbf{d}), \quad (22)$$

6 where \mathbf{y} is the λ -dimensional vector with $y_{\ell} = g(\mathbf{x}_{\ell}^{\lambda})$, $[M]$ the $(\lambda \times \mu)$ -dimensional matrix with $M_{\ell\alpha} = (\mathbf{x}_{\ell}^{\lambda})^{\alpha}$,
 7 and $[W]$ the λ -dimensional diagonal matrix with $W_{\ell\ell} = w_{\ell}^{\lambda}$. We recall that μ is the number of monomials in $\{\mathbf{x}^{\alpha}, 0 \leq$
 8 $|\alpha| \leq p\}$ and λ the number of nodes in the quadrature for integration with respect to $P_{\mathbf{X}}$. The tilde serves in (22) to
 9 distinguish between surrogate models provided by the first and second nonintrusive projection methods.

10 Many methods can be used to solve the optimization problem in (22). Some initially form the normal equations
 11 and then solve them using a Cholesky factorization. Others avoid forming the normal equations and instead rely on a
 12 QR factorization, a singular value decomposition, or other linear algebra method. There are also methods that cast the
 13 optimization problem in (22) as a generalized least-squares problem, which is then solved using either linear algebra
 14 or quadratic programming. These methods differ in their ability to handle problems that may have multiple solutions
 15 and in their numerical stability: if $[M]$ or $[W]$ is rank-deficient or ill-conditioned, the use of an inadequate method
 16 may cause a disastrous loss of numerical accuracy. Thus far, it appears that only little theoretical and computational
 17 work in the area of uncertainty quantification has examined this issue. Nevertheless, theoretical and computational
 18 studies are available in the field of linear algebra, and we refer the reader to [63–65] for guidance.

19 We note that if the weighted least-squares problem in (22) is ill-conditioned, it can be helpful to normalize the
 20 uncertain input variables prior to the construction of the surrogate model. One way of doing this involves sub-
 21 tracting the mean vector $\mathbf{m}_{\mathbf{X}}$ and multiplying by the inverse $[R_{\mathbf{X}}]^{-T}$ of the Cholesky factor of the covariance
 22 matrix $[C_{\mathbf{X}}] = [R_{\mathbf{X}}^T R_{\mathbf{X}}]$; in fact, the random variable $[R_{\mathbf{X}}]^{-T}(\mathbf{X} - \mathbf{m}_{\mathbf{X}})$ is normalized in the sense that its
 23 mean vector vanishes and its covariance matrix is equal to the identity matrix. The surrogate model then takes the
 24 form $\tilde{g}^{p,\lambda} = \sum_{|\alpha|=0}^p c_{\alpha}^{\lambda} ([R_{\mathbf{X}}]^{-T}(\mathbf{x} - \mathbf{m}_{\mathbf{X}}))^{\alpha}$, wherein the coefficients, although still denoted by c_{α}^{λ} , are obtained,
 25 this time, by solving the weighted least-squares problem $\mathbf{c}^{\lambda} = \text{solution of } \min_{\mathbf{d} \in \mathbb{R}^{\mu}} \frac{1}{2} (\mathbf{y} - [M]\mathbf{d})^T [W] (\mathbf{y} - [M]\mathbf{d})$,
 26 where \mathbf{y} and $[W]$ are as in (22) but $[M]$ is the $(\lambda \times \mu)$ -dimensional matrix with $M_{\ell\alpha} = ([R_{\mathbf{X}}]^{-T}(\mathbf{x}_{\ell}^{\lambda} - \mathbf{m}_{\mathbf{X}}))^{\alpha}$.

27 Many methods can be used to obtain the required quadrature rule for integration with respect to $P_{\mathbf{X}}$. There

1 are probabilistic and number-theoretic integration methods as well as nonprobabilistic integration methods, such as
2 Gaussian, polynomial-based, and other integration methods, sparse-grid and other tensorization methods, and adaptive
3 integration methods (refer, for details, to [24, 45, 49, 66–70]). *Probabilistic integration methods* produce quadrature
4 rules whose nodes are distributed randomly in the domain of integration. By contrast, *nonprobabilistic integration*
5 *methods* produce quadrature rules whose nodes are organized in a systematic way in the domain of integration. When
6 the dimension of the domain of integration is small or moderate and when the integrand is sufficiently smooth, this
7 systematic organization of the nodes allows nonprobabilistic methods to produce very efficient quadrature rules that
8 achieve a high level of accuracy with only a small number of nodes. However, as the dimension of the domain of
9 integration increases, nonprobabilistic methods lose their ability to form accurate quadrature rules with only a small
10 number of nodes and ultimately become less efficient than probabilistic methods.

11 Both of the nonintrusive projection methods mentioned previously are *nonintrusive* because they require only
12 the repeated solution—sequentially or in parallel—of the computational model for different values assigned to its
13 uncertain input variables; the computational model itself need not be modified. The manner in which the coefficients
14 are computed in (21) and the residual is computed in (22) indicates that the computational model must be solved for
15 each node in the quadrature rule for integration with respect to $P_{\mathbf{X}}$. Thus, the computational cost of constructing the
16 surrogate model essentially scales with the number of nodes within this quadrature rule.

17 From a theoretical point of view, results from approximation theory (refer, for example, to [24, 45, 49, 66, 68–70])
18 can be used to analyze the convergence of quadrature-based approximations of integrals with respect to the number
19 of nodes. From a computational point of view, numerical convergence studies can be conducted.

20 Regarding desideratum (iv) in Sec. 4.4, we note that owing to these considerations, the availability of an efficient
21 quadrature rule can play a role in constructing the characterization of the uncertain input variables.

22 5.3.5 Curse-of-dimensionality

23 Especially when the computational model is expensive to solve, it is essentially the quadrature rule that determines
24 the computational cost of a nonintrusive-projection-based implementation. Indeed, as we have already mentioned, a
25 nonintrusive-projection-based implementation requires that the computational model be solved for each node in the
26 quadrature; thus, the smaller the number of nodes, the lower this computational cost.

27 It follows that from among the aforementioned methods available for obtaining the quadrature rule, nonproba-
28 bilistic integration methods are a natural choice because as we have already mentioned, they can be expected to be
29 able to form an accurate quadrature rule with only a small number of nodes, especially when the number of un-

1 certain input variables is small or moderate and when the computational model is sufficiently smooth. However,
 2 nonintrusive-projection-based implementations that use nonprobabilistic integration methods suffer from a so-called
 3 “curse-of-dimensionality” because as the number of uncertain input variables increases, nonprobabilistic integration
 4 methods lose their ability to form accurate quadrature rules with only a small number of nodes [24, 45, 49, 66–70].

5 5.3.6 Approximation of statistical descriptors

6 Once a surrogate model is available, it can be used as a substitute for the computational model in the approximation
 7 of statistical descriptors of the quantity of interest. First, an ensemble of i.i.d. samples $\{\mathbf{x}_\ell, 1 \leq \ell \leq \nu\}$ is
 8 generated from $P_{\mathbf{X}}$. The surrogate model is then used to map each sample \mathbf{x}_ℓ into the corresponding sample of
 9 the quantity of interest, that is, either $y_\ell^{p,\lambda} = g^{p,\lambda}(\mathbf{x}_\ell)$ or $y_\ell^{p,\lambda} = \tilde{g}^{p,\lambda}(\mathbf{x}_\ell)$, depending on whether the first or second
 10 nonintrusive projection method is used. Finally, mathematical statistics methods are applied to the ensemble of i.i.d.
 11 samples $\{y_\ell^{p,\lambda}, 1 \leq \ell \leq \nu\}$. For example, the mean m_Y and the variance σ_Y^2 (if they exist) can be approximated as

$$m_Y \approx m_Y^{p,\lambda,\nu} = \frac{1}{\nu} \sum_{\ell=1}^{\nu} y_\ell^{p,\lambda}, \quad \sigma_Y^2 \approx (\sigma_Y^{p,\lambda,\nu})^2 = \frac{1}{\nu} \sum_{\ell=1}^{\nu} (y_\ell^{p,\lambda} - m_Y^{p,\lambda,\nu})^2. \quad (23)$$

12 Under certain conditions, statistical descriptors of the quantity of interest can be directly approximated starting
 13 from either the coefficients g_α^λ in the representation of the surrogate model in (21) or the coefficients c_α^λ in the
 14 representation of the surrogate model in (22). For example, if the set $\{\psi_\alpha, 0 \leq |\alpha| \leq p\}$ of $P_{\mathbf{X}}$ -orthonormal
 15 polynomials is such that $\psi_0 = 1$, it follows from (17) that the mean m_Y and the variance σ_Y^2 can be approximated as

$$\text{If } \psi_0 = 1, \quad \text{then } m_Y \approx m_Y^\lambda = g_0^\lambda \quad \text{and} \quad \sigma_Y^2 \approx (\sigma_Y^{p,\lambda})^2 = \sum_{|\alpha|=1}^p (g_\alpha^\lambda)^2. \quad (24)$$

16 As in the Monte Carlo sampling method, the law of large numbers and the central limit theorem can be used to
 17 analyze the convergence of approximations such as those in (23) with respect to the number of samples. If convergence
 18 is guaranteed, it suffices to employ a sufficiently large number of samples to ensure that the error from using only a
 19 finite number of samples is much smaller than the error introduced by the use of the surrogate model.

20 Thus, the implementation of a stochastic expansion method most often consists of approximating the computa-
 21 tional model by a sufficiently accurate surrogate model, followed by the use of this surrogate model as a substitute for
 22 the computational model in the approximation of statistical descriptors of the quantity of interest. This approximation
 23 entails virtually no overhead because the computational cost of evaluating polynomials is very low.

1 5.3.7 Reduced-dimensional surrogate model

2 We have seen that the uncertain input variables can be characterized as a transformation $\mathbf{X} = \mathbf{f}(\Xi)$ of a given random
3 variable Ξ with values in \mathbb{R}^d and probability distribution P_Ξ under a mapping \mathbf{f} from \mathbb{R}^d into \mathbb{R}^m . This presents the
4 possibility that the composition of the mapping \mathbf{f} with the computational model g can be approximated by a surrogate
5 model, thus facilitating, if d is smaller than m , the construction of a reduced-dimensional surrogate model [6, 41].

6 5.4 Advanced stochastic expansion methods and related methods

7 Much recent research has investigated how the computational cost of stochastic expansion methods (Sec. 5.3) can be
8 reduced, as well as how their range of applicability can be extended to problems of higher and higher dimension and
9 to computational models that lack smoothness. In particular, alternative (not necessarily polynomial) basis functions
10 and alternative formulations of the construction of the surrogate model are being investigated; please refer, for ex-
11 ample, to [71–77] for details about the use of alternative basis functions including Haar bases, multiwavelet bases,
12 Padé approximants, multi-element polynomial chaos, enrichment functions, preconditioned bases, and separated rep-
13 resentations; and please refer, for example, to [78–83] for details about alternative formulations including Bayesian
14 formulations and Gaussian processes, sensing and other machine-learning formulations, adaptive formulations, as
15 well as formulations that seek to exploit dimension reduction or sparsity, and various combinations of these.

16 5.5 Effectiveness of Monte Carlo sampling and stochastic expansion methods

17 In summary, when the Monte Carlo sampling method is used, accuracy usually improves with the square root of the
18 number of solutions of the computational model; although this convergence rate is rather slow, it is independent of the
19 number of uncertain input variables. When a stochastic expansion method is used, an accurate surrogate model can
20 usually be obtained in a computationally efficient manner if the computational model has only a small or moderate
21 number of uncertain input variables and is sufficiently smooth; once available, the propagation of uncertainties through
22 this surrogate model usually entails virtually no overhead beyond the computational cost of its construction.

23 Thus, within the current state of the art, the Monte Carlo sampling method is most computationally efficient for
24 problems of “very high dimension,” that is, for computational models with a very large number of uncertain input
25 variables. Stochastic expansion methods are indicated most for problems “of low or moderate dimension,” that is, for
26 computational models with a small or moderate number of uncertain input variables, as well as for problems that admit
27 reduction relative to dimensionality. We emphasize that the relative merits and limitations of these methods depend

1 on the problem and implementation specificities; further, we emphasize that much ongoing research is involved with
2 extending the computational efficiency of stochastic expansion methods to problems of higher and higher dimension.

3 **5.6 Reduced-order models**

4 As an alternative to the use of a surrogate model, the computational cost of the uncertainty propagation can also be
5 reduced through the use of a reduced-order model as a substitute for the computational model. In addition to multiscale
6 and coarse-graining approaches, there exist various projection-based methods for obtaining reduced-order models,
7 such as those based on eigendecomposition, proper orthogonal decomposition, Krylov subspaces, and reduced bases.
8 Two significant challenges are in dealing with nonlinearities and in constructing a (family of) reduced-order model(s)
9 that maintains accuracy over a range of values of the uncertain input variables of the computational model [84, 85].

10 **6. SENSITIVITY ANALYSIS OF UNCERTAINTIES**

11 Once the characterization and propagation steps are complete, the objective of the sensitivity analysis of uncertainties
12 is to gain some insight into the manner in which uncertainties introduced in the input variables induce uncertainties
13 in the quantity of interest. Such insight can be very useful for identifying where to direct efforts aimed at reducing
14 uncertainties, and it can constitute a crucial prerequisite to the optimization of designs in the presence of uncertainties
15 and the validation of models, among other purposes.

16 Several types of sensitivity analysis of uncertainties can be used and have been proposed in the literature, such
17 as methods involving scatter plots and regression, correlation, and elementary effects [86–88], methods involving
18 variance analysis [87, 89–91], and methods involving differentiation [92–95]. Variance- and differentiation-based
19 methods are described below; please refer to the references given above for details on the other methods.

20 **6.1 Variance-based methods**

21 Variance-based methods begin by partitioning the uncertain input variables into subsets. These subsets of uncertain
22 input variables are then arranged in an order that reflects their significance in inducing uncertainties in the quantity of
23 interest, thus permitting dominant subsets of uncertain input variables to be distinguished from insignificant ones.

24 In their traditional realm [87, 89, 90], variance-based methods require the uncertain input variables to be parti-
25 tioned into subsets that are statistically independent. *Statistically independent* subsets of uncertain input variables
26 are those between which no physical, causal relationship exists (if uncertainties are held to refer to variability) or
27 no logical relationship is indicated by the available information (if uncertainties are held to refer to a state of possi-

1 bly incomplete knowledge). More recent studies [90, 91] have extended the range of applicability of variance-based
 2 methods to statistically dependent subsets of uncertain input variables. However, for brevity, this section is confined
 3 to variance-based methods appropriate for statistically independent subsets; nevertheless, we will apply one of the
 4 available extensions to statistically dependent subsets in the illustrative problem of Secs. 8 and 9.

5 For example, if the computational model were a finite element model for the mechanical deformation of a structure,
 6 one subset of uncertain input variables could be involved in the characterization of uncertain material properties
 7 and another could be involved in the characterization of an uncertain applied external loading. The variance-based
 8 sensitivity analysis may then involve determining whether either the uncertainties in the material properties or those
 9 in the applied loading are most significant in inducing uncertainties in the deformed shape of the structure.

10 Without loss of generality, let the uncertain input variables be partitioned into n subsets in such a way that the first
 11 subset contains the first m_1 uncertain input variables, the second subset contains the next m_2 uncertain input variables,
 12 and so forth, until the final subset contains the last m_n uncertain input variables. Correspondingly, let the random
 13 variable $\mathbf{X} = (X_1, \dots, X_m)$ be partitioned into n random variables in such a way that $\mathbf{X}^1 = (X_1, \dots, X_{m_1})$,
 14 $\mathbf{X}^2 = (X_{m_1+1}, \dots, X_{m_1+m_2})$, and so forth, until $\mathbf{X}^n = (X_{m_1+\dots+m_{n-1}+1}, \dots, X_{m_1+\dots+m_n})$, that is,

$$\mathbf{X} = \underbrace{(X_1, \dots, X_{m_1})}_{\mathbf{X}^1}, \underbrace{(X_{m_1+1}, \dots, X_{m_1+m_2})}_{\mathbf{X}^2}, \dots, \underbrace{(X_{m_1+\dots+m_{n-1}+1}, \dots, X_{m_1+\dots+m_n})}_{\mathbf{X}^n}. \quad (25)$$

15 Let the random variables $\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^n$ be statistically independent; then, by the rules of probability theory, the
 16 probability distribution of $\mathbf{X} = (\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^n)$ is the product of the probability distributions of $\mathbf{X}^1, \mathbf{X}^2, \dots, \mathbf{X}^n$:

$$P_{\mathbf{X}} = P_{\mathbf{X}^1} \times P_{\mathbf{X}^2} \times \dots \times P_{\mathbf{X}^n}. \quad (26)$$

17 As the main tool for gauging the significance of a subset of uncertain input variables in inducing uncertainties in the
 18 quantity of interest, say, of the j -th subset, variance-based methods provide the following *significance descriptor*:

$$s_{\mathbf{X}^j} = \int_{\mathbb{R}^{m \sim j}} \left(\int_{\mathbb{R}^{m_j}} (g(\mathbf{x}^j, \mathbf{x}^{\sim j}) - m_Y(\mathbf{x}^{\sim j}))^2 dP_{\mathbf{X}^j} \right) dP_{\mathbf{X}^{\sim j}}, \text{ where } m_Y(\mathbf{x}^{\sim j}) = \int_{\mathbb{R}^{m_j}} g(\mathbf{x}^j, \mathbf{x}^{\sim j}) dP_{\mathbf{X}^j}, \quad (27)$$

19 where $\mathbf{X}^{\sim j}$ denotes the random variable that collects all those components of the random variable \mathbf{X} that are not
 20 components of the random variable \mathbf{X}^j , for example, $\mathbf{X}^{\sim 1} = (\mathbf{X}^2, \dots, \mathbf{X}^n)$. Although $s_{\mathbf{X}^j}$ can also be insightfully
 21 interpreted in the context of regression analysis [87, 89, 90], an intuitive interpretation is obtained by recognizing
 22 $\int_{\mathbb{R}^{m_j}} (g(\mathbf{x}^j, \mathbf{x}^{\sim j}) - m_Y(\mathbf{x}^{\sim j}))^2 dP_{\mathbf{X}^j}$ as the variance of $g(\mathbf{X}^j, \mathbf{x}^{\sim j})$, that is, as the variance that the quantity of

1 interest would exhibit owing to the uncertainties in the uncertain input variables contained in the j -th subset if the
 2 value taken by the uncertain input variables contained in all the other subsets was equal to the specific value $\mathbf{x}^{\sim j}$.
 3 Thus, variance-based methods provide the $P_{\mathbf{X}^{\sim j}}$ -weighted average of the variance of $g(\mathbf{X}^j, \mathbf{x}^{\sim j})$ as a significance
 4 descriptor that reflects the significance of the j -th subset of uncertain input variables in inducing uncertainties in the
 5 quantity of interest. Once the n significance descriptors are available,

$$s_{\mathbf{X}^1}, s_{\mathbf{X}^2}, \dots, s_{\mathbf{X}^n}, \quad (28)$$

6 the corresponding subsets of uncertain input variables can be arranged in order of significance—the higher $s_{\mathbf{X}^j}$, the
 7 more significant the j -th subset of uncertain input variables in inducing uncertainties in the quantity of interest—thus
 8 permitting dominant subsets of uncertain input variables to be identified.

9 We note that variance-based methods provide several other significance descriptors that allow a fuller understand-
 10 ing to be gained of how subsets of uncertain input variables influence the quantity of interest. Nevertheless, it appears
 11 that the significance descriptor $s_{\mathbf{X}^j}$ is the backbone of variance-based methods because many of these other sig-
 12 nificance descriptors can be deduced therefrom; for example, the total-effect index given in [87, 89, 90] is obtained
 13 as $s_{\mathbf{X}^j}/\sigma_Y^2$, the main-effect index as $(\sigma_Y^2 - s_{\mathbf{X}^{\sim j}})/\sigma_Y^2$, and the interaction-effect index can be deduced from $s_{(\mathbf{X}^i, \mathbf{X}^j)}$.

14 The computation of $s_{\mathbf{X}^j}$ is a problem of numerical integration. As in the previous section, either probabilis-
 15 tic and number-theoretic integration methods or nonprobabilistic integration methods can be used. Especially when
 16 a probabilistic integration method is used, the computational cost of a straightforward approximation can be pro-
 17 hibitive because of the nested structure of the integrals in (27). Variance-based methods circumvent this issue by
 18 using the law of total variance [43] to express $s_{\mathbf{X}^j}$ equivalently as $s_{\mathbf{X}^j} = \int_{\mathbb{R}^{m \sim j}} \int_{\mathbb{R}^{m_j}} g(\mathbf{x}^j, \mathbf{x}^{\sim j})^2 dP_{\mathbf{X}^j} dP_{\mathbf{X}^{\sim j}}$
 19 $- \int_{\mathbb{R}^{m \sim j}} (\int_{\mathbb{R}^{m_j}} g(\mathbf{x}^j, \mathbf{x}^{\sim j}) dP_{\mathbf{X}^j}) (\int_{\mathbb{R}^{m_j}} g(\mathbf{x}^j, \mathbf{x}^{\sim j}) dP_{\mathbf{X}^j}) dP_{\mathbf{X}^{\sim j}}$, thus allowing $s_{\mathbf{X}^j}$ to be approximated by using
 20 two independent ensembles $\{\mathbf{x}_\ell^j, 1 \leq \ell \leq \nu\}$ and $\{\tilde{\mathbf{x}}_\ell^j, 1 \leq \ell \leq \nu\}$ of i.i.d. samples from $P_{\mathbf{X}^j}$ and one en-
 21 semble $\{\mathbf{x}_\ell^{\sim j}, 1 \leq \ell \leq \nu\}$ of i.i.d. samples from $P_{\mathbf{X}^{\sim j}}$ as follows:

$$s_{\mathbf{X}^j} \approx s_{\mathbf{X}^j}^\nu = \frac{1}{2\nu} \sum_{\ell=1}^{\nu} \left(g(\mathbf{x}_\ell^j, \mathbf{x}_\ell^{\sim j}) - g(\tilde{\mathbf{x}}_\ell^j, \mathbf{x}_\ell^{\sim j}) \right)^2. \quad (29)$$

22 Alternatively, if efficient quadrature rules $\{(\mathbf{x}_k^{j,\kappa}, w_k^{j,\kappa}), 1 \leq k \leq \kappa\}$ and $\{(\mathbf{x}_\ell^{\sim j,\lambda}, w_\ell^{\sim j,\lambda}), 1 \leq \ell \leq \lambda\}$ for

1 integration with respect to $P_{\mathbf{X}^j}$ and $P_{\mathbf{X}^{\sim j}}$, respectively, are available, $s_{\mathbf{X}^j}$ can be straightforwardly approximated as

$$s_{\mathbf{X}^j} \approx s_{\mathbf{X}^j}^{\kappa, \lambda} = \sum_{k=1}^{\kappa} \sum_{\ell=1}^{\lambda} w_k^{j, \kappa} w_{\ell}^{\sim j, \lambda} \left(g(\mathbf{x}_k^{j, \kappa}, \mathbf{x}_{\ell}^{\sim j, \lambda}) - m_{Y, \ell}^{\kappa} \right)^2, \quad \text{where} \quad m_{Y, \ell}^{\kappa} = \sum_{k=1}^{\kappa} w_k^{j, \kappa} g(\mathbf{x}_k^{j, \kappa}, \mathbf{x}_{\ell}^{\sim j, \lambda}). \quad (30)$$

2 We note that the computational cost of computing the significance descriptor $s_{\mathbf{X}^j}$ can be lowered by using a
3 surrogate model (if available) as a substitute for the computational model in either (29) or (30). Finally, we note
4 that under certain conditions, certain surrogate models allow $s_{\mathbf{X}^j}$ to be directly approximated from the coefficients
5 involved in the representation of the surrogate model; please refer to [96, 97] for details.

6 6.2 Differentiation-based methods

7 Insight into how uncertainties induced in the quantity of interest depend on those introduced in the input variables
8 can often also be gained by exploring the sensitivity of statistical descriptors of the quantity of interest with respect to
9 changes in parameters involved in the characterization of the uncertain input variables. Differentiation-based meth-
10 ods [92, 93, 95] lead to such insight by differentiating statistical descriptors of the quantity of interest with respect to
11 parameters involved in the characterization of the uncertain input variables.

12 For example, if the computational model were a finite element model for the mechanical deformation of a structure,
13 an uncertain material property could be characterized as a random field parameterized by a nominal value, a dispersion
14 level, and a spatial correlation length. The differentiation-based sensitivity analysis could then involve determining
15 the sensitivity of the variance of a displacement component at a prescribed location with respect to changes in this
16 nominal value, this dispersion level, and this spatial correlation length.

17 Let the characterization of the uncertain input variables have been obtained by selecting a characterization $P_{\mathbf{X}}(\mathbf{p})$
18 that depends on a finite number of parameters collected in a vector $\mathbf{p} = (p_1, \dots, p_q)$ and then assigning an adequate
19 value $\hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_q)$ to these parameters (refer to Secs. 4.2 and 4.5), that is,

$$P_{\mathbf{X}} \equiv P_{\mathbf{X}}(\hat{\mathbf{p}}), \quad \text{where} \quad \hat{\mathbf{p}} = (\hat{p}_1, \dots, \hat{p}_q). \quad (31)$$

20 Let us consider a statistical descriptor of the quantity of interest that admits a representation as $\int_{\mathbb{R}^m} h(g(\mathbf{x})) dP_{\mathbf{X}}(\hat{\mathbf{p}})$,
21 where h is a function from \mathbb{R} into \mathbb{R} . This representation is quite general and encompasses cases wherein interest might
22 be in the mean $m_Y(\hat{\mathbf{p}})$ and variance $\sigma_Y^2(\hat{\mathbf{p}})$ of the quantity of interest by setting $h(y) = y$ and $h(y) = (y - m_Y(\hat{\mathbf{p}}))^2$,
23 respectively. Differentiation-based methods determine the sensitivity of this statistical descriptor with respect to
24 changes in a parameter involved in the characterization of the uncertain input variables, say, in the j -th parameter,

1 using the partial derivative of $\int_{\mathbb{R}^m} h(g(\mathbf{x})) dP_{\mathbf{X}}(\mathbf{p})$ with respect to p_j evaluated at $\hat{\mathbf{p}}$:

$$s_{p_j} = \left. \frac{\partial \int_{\mathbb{R}^m} h(g(\mathbf{x})) dP_{\mathbf{X}}(\mathbf{p})}{\partial p_j} \right|_{\mathbf{p}=\hat{\mathbf{p}}}; \quad (32)$$

2 the higher the magnitude of the *sensitivity descriptor* s_{p_j} , the more sensitive the statistical descriptor with respect to
3 changes in the j -th parameter. The simplest computation of s_{p_j} involves finite-difference approximation.

4 The problem of numerically approximating the sensitivity descriptor s_{p_j} can also be reformulated as a problem
5 of numerical integration. In fact, if the probability distribution $P_{\mathbf{X}}(\mathbf{p})$ admits a PDF $\rho_{\mathbf{X}}(\mathbf{p})$, if the differentiation
6 and integration in (32) are interchangeable, and if $\rho_{\mathbf{X}}(\mathbf{p})$ is sufficiently differentiable, then s_{p_j} can be expressed
7 equivalently as $s_{p_j} = \int_{\mathbb{R}^m} h(g(\mathbf{x})) \frac{\partial \rho_{\mathbf{X}}(\mathbf{x}, \mathbf{p})}{\partial p_j} \Big|_{\mathbf{p}=\hat{\mathbf{p}}} d\mathbf{x} = \int_{\mathbb{R}^m} h(g(\mathbf{x})) \frac{\partial \ln \rho_{\mathbf{X}}(\mathbf{x}; \mathbf{p})}{\partial p_j} \Big|_{\mathbf{p}=\hat{\mathbf{p}}} \rho_{\mathbf{X}}(\mathbf{x}; \hat{\mathbf{p}}) d\mathbf{x}$, thus allowing s_{p_j} to
8 be approximated using an ensemble $\{\mathbf{x}_\ell, 1 \leq \ell \leq \nu\}$ of i.i.d. samples from $P_{\mathbf{X}}(\hat{\mathbf{p}})$ as

$$s_{p_j} \approx s_{p_j}^\nu = \frac{1}{\nu} \sum_{\ell=1}^{\nu} h(g(\mathbf{x}_\ell)) \frac{\partial \ln \rho_{\mathbf{X}}(\mathbf{x}_\ell; \mathbf{p})}{\partial p_j} \Big|_{\mathbf{p}=\hat{\mathbf{p}}} \quad (33)$$

9 and using a nonprobabilistic quadrature rule $\{(\mathbf{x}_\ell^\lambda, w_\ell^\lambda), 1 \leq \ell \leq \lambda\}$ for integration with respect to $P_{\mathbf{X}}(\hat{\mathbf{p}})$ as

$$s_{p_j} \approx s_{p_j}^\lambda = \sum_{\ell=1}^{\lambda} w_\ell^\lambda h(g(\mathbf{x}_\ell^\lambda)) \frac{\partial \ln \rho_{\mathbf{X}}(\mathbf{x}_\ell^\lambda; \mathbf{p})}{\partial p_j} \Big|_{\mathbf{p}=\hat{\mathbf{p}}}. \quad (34)$$

10 If the characterization of the uncertain input variables as a random variable \mathbf{X} with probability distribution $P_{\mathbf{X}}(\mathbf{p})$
11 corresponds to a characterization as a transformation $\mathbf{X} = \mathbf{f}(\Xi; \mathbf{p})$ (refer to Sec. 4.2), then, assuming again that
12 the differentiation and integration in (32) are interchangeable and that $\mathbf{f}(\cdot; \mathbf{p})$, g , and h are sufficiently differentiable,
13 the sensitivity descriptor s_{p_j} can be expressed as $s_{p_j} = \int_{\mathbb{R}^d} \frac{\partial h(y)}{\partial y} \Big|_{y=g(\mathbf{f}(\xi, \hat{\mathbf{p}}))} (\nabla_{\mathbf{x}} g(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{f}(\xi, \hat{\mathbf{p}})})^T \frac{\partial \mathbf{f}(\xi, \mathbf{p})}{\partial p_j} \Big|_{\mathbf{p}=\hat{\mathbf{p}}} dP_{\Xi}$,
14 thus allowing s_{p_j} to be approximated using an ensemble $\{\xi_\ell, 1 \leq \ell \leq \nu\}$ of i.i.d. samples from P_{Ξ} as

$$s_{p_j} \approx \tilde{s}_{p_j}^\nu = \frac{1}{\nu} \sum_{\ell=1}^{\nu} \frac{\partial h(y)}{\partial y} \Big|_{y=g(\mathbf{f}(\xi_\ell, \hat{\mathbf{p}}))} \left(\nabla_{\mathbf{x}} g(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{f}(\xi_\ell, \hat{\mathbf{p}})} \right)^T \frac{\partial \mathbf{f}(\xi_\ell, \mathbf{p})}{\partial p_j} \Big|_{\mathbf{p}=\hat{\mathbf{p}}} \quad (35)$$

15 and using a nonprobabilistic quadrature rule $\{(\xi_\ell^\lambda, w_\ell^\lambda), 1 \leq \ell \leq \lambda\}$ for integration with respect to P_{Ξ} as

$$s_{p_j} \approx \tilde{s}_{p_j}^\lambda = \sum_{\ell=1}^{\lambda} w_\ell^\lambda \frac{\partial h(y)}{\partial y} \Big|_{y=g(\mathbf{f}(\xi_\ell^\lambda, \hat{\mathbf{p}}))} \left(\nabla_{\mathbf{x}} g(\mathbf{x}) \Big|_{\mathbf{x}=\mathbf{f}(\xi_\ell^\lambda, \hat{\mathbf{p}})} \right)^T \frac{\partial \mathbf{f}(\xi_\ell^\lambda, \mathbf{p})}{\partial p_j} \Big|_{\mathbf{p}=\hat{\mathbf{p}}}. \quad (36)$$

16 We note that fulfillment of the interchangeability assumption mentioned previously can be verified by invoking,

1 for example, the monotone or dominated convergence theorems; for details, please refer to [43, 44, 92]. Further, we
2 note that the computational cost of computing s_{p_j} can be lowered through the use of a surrogate model (if available).

3 Finally, regarding desideratum (v) in Sec. 4.4, we note that the variety of differentiation-based sensitivity analyses
4 that are enabled by the parameterization of the characterization of the uncertain input variables and their relevance to
5 the scientific or engineering questions being asked can play a role in obtaining this characterization.

6 7. IMPLEMENTATION DETAILS

7 The methods described previously lend themselves well to implementations in computational linear algebra packages.
8 In addition, implementations of some of the methods described previously are available in open-source software
9 packages, such as Dakota [98], GPMSA [3], Queso [99], Trilinos::Stokhos [100], and UQTK [101]:

- 10 – Dakota provides nonintrusive algorithms for uncertainty propagation (Monte Carlo sampling, stochastic expansion),
11 sensitivity analysis, design optimization, model calibration, verification, and parameter studies.
- 12 – GPMSA provides nonintrusive algorithms for uncertainty propagation (Gaussian processes), sensitivity analysis,
13 model calibration, and parameter estimation.
- 14 – Queso provides nonintrusive algorithms for uncertainty characterization (Bayesian), decision making under
15 uncertainty, model calibration, and validation.
- 16 – Trilinos::Stokhos provides intrusive algorithms for uncertainty propagation (stochastic expansion), including
17 algorithms for forming and efficiently solving spectral problems arising in embedded projection methods.
- 18 – UQTK provides intrusive and nonintrusive algorithms for uncertainty propagation (stochastic expansion), including
19 algorithms suitable for rapid application prototyping, algorithm research, and community outreach.

20 8. REALIZATION FOR AN ENGINEERING PROBLEM

21 8.1 Context

22 Much of the steel that is produced by steel plants is shaped by forming processes in preparation for its use in as-
23 sembling cars, ships, appliances, and various other products. A *forming process* typically forms pieces of steel, for
24 example, steel sheets, also called steel blanks, by deforming them using classical forming processes such as bending,
25 deep drawing, cold roll forming, or combinations thereof. After leaving the tooling, which corresponds to being re-
26 leased of the forming forces, the material has a tendency to partially return to its original shape because of its elastic

1 recovery. This phenomenon, referred to as the *springback*, is quite complex (refer, for details, to [102] and the ref-
2 erences therein) and depends not only on material properties such as Young's modulus and yield stress but also on
3 many process parameters such as sheet thickness and bending angles. The springback is difficult to predict and is a
4 major quality concern in forming processes because when the springback is smaller or larger than expected, it can
5 cause serious problems to subsequent assembly processes because of geometry mismatches.

6 Often, the same forming process must be applied repeatedly, each time to a nominally identical piece of steel and
7 under nominally identical process conditions. However, even though they are nominally identical, the pieces of steel
8 and the process conditions may exhibit (possibly small) variability; hence, the forming process may persist in yielding
9 a different springback and therefore a permanently deformed piece of steel of a different shape each time that it is
10 repeated. Thus, especially when it must be applied repeatedly, the design of a *robust* forming process can require an
11 assessment of the impact that variability in the pieces of steel and the process conditions has on the springback.

12 8.2 Problem setting

13 We consider a forming process wherein a steel sheet is bent along a straight line: a portion of the steel sheet is clamped
14 (Fig. 1(1)) and the complementary portion is bent downwards by a punch that descends until a rectangular angle is
15 imposed (Fig. 1(2)), after which the punch ascends until it is ultimately removed (Fig. 1(3)). In such a forming
16 process, the steel sheet does not keep its deformed shape after bending; instead, it springs back slightly upwards when
17 the punch is removed. We refer to the angle with which the steel sheet bends upwards again as the *springback angle*.

18 We consider a finite element model implemented in our in-house software METAFOR [102]. This finite element
19 model is based on the geometry depicted in Fig. 1(1); further, it is based on a model of the mechanical behavior
20 of the steel according to an elastoplastic constitutive model with linear isotropic hardening and on the modeling of
21 the contact surfaces between the die, the sheet, and the punch as frictionless contact surfaces. Once the geometrical
22 characteristics in Fig. 1(1) are fixed and values assigned to the parameters involved in the elastoplastic constitutive
23 model with linear isotropic hardening, this finite element model allows the springback angle to be predicted.

24 To examine the impact that variability in the mechanical behavior of the steel has on the springback angle, we
25 consider two of the parameters involved in the elastoplastic constitutive model with linear isotropic hardening, namely,
26 the *hardening modulus* and the *yield stress*, to be uncertain. We assume that the available information—from which we
27 must infer the characterization of the hardening modulus and yield stress—consists of data collected from experiments,
28 in addition to mechanical and physical constraints. Specifically, we assume that a small or moderate number of

1 samples, say n , of the hardening modulus and yield stress are available, written as

$$\{(h_\ell^{\text{obs}}, s_\ell^{\text{obs}}), 1 \leq \ell \leq n\}. \quad (37)$$

2 These samples may have been obtained by simple tensile testing of n steel sheets and characterizing the mechanical
3 behavior of each one in terms of a hardening modulus h_ℓ^{obs} and yield stress s_ℓ^{obs} . We assume that the mechanical and
4 physical constraints require that the hardening modulus and yield stress be positive, for example, as a reflection of the
5 fact that it may be known that the steel hardens—and does not soften—as it deforms.

6 Next, we will demonstrate some of the methods described previously by applying them to the characterization
7 of the uncertainties in the hardening modulus and yield stress (the uncertain input variables), the propagation of
8 these uncertainties through the finite element model (the computational model), and the sensitivity analysis of these
9 uncertainties to allow some insight into their impact on the springback angle (the quantity of interest).

10 8.3 Characterization of uncertainties

11 Because only a small or moderate amount of data is assumed to be available, we proceed by selecting a “labeled”
12 probability distribution that depends on only a small or moderate number of parameters and then inferring adequate
13 values for these parameters from the data using the method of maximum likelihood; refer to Sec. 4.5.

14 For reasons that we will discuss throughout and summarize at the end of this section, we select the bivariate gamma
15 probability distribution, which was introduced in [103], as the “labeled” probability distribution. This distribution
16 admits a probability density function that, when characterizing the hardening modulus and yield stress as random
17 variables H and S , reads as follows:

$$\rho_{(H,S)}(h, s; m_H, \sigma_H^2, m_S, \sigma_S^2, \rho) = \rho_\Gamma(h; m_H, \sigma_H^2) \rho_\Gamma(s; m_S, \sigma_S^2) \sigma(c_\Gamma(h; m_H, \sigma_H^2), c_\Gamma(s; m_S, \sigma_S^2); \rho). \quad (38)$$

18 Here, $\rho_\Gamma(\cdot; m_\Gamma, \sigma_\Gamma^2)$ is the univariate gamma probability density function with mean $m_\Gamma > 0$ and variance $\sigma_\Gamma^2 > 0$:

$$\rho_\Gamma(\gamma; m_\Gamma, \sigma_\Gamma^2) = c \gamma^{\frac{m_\Gamma^2}{\sigma_\Gamma^2} - 1} \exp\left(-\frac{m_\Gamma}{\sigma_\Gamma^2} \gamma\right) \quad \text{if } \gamma \geq 0 \quad \text{and} \quad \rho_\Gamma(\gamma; m_\Gamma, \sigma_\Gamma^2) = 0 \quad \text{otherwise}, \quad (39)$$

19 where c is the normalization constant, $\sigma(\cdot, \cdot; \rho)$ with $-1 \leq \rho \leq 1$ is the function such that

$$\sigma(u, v; \rho) = \frac{1}{\sqrt{1 - \rho^2}} \exp\left(-\frac{(\rho c_\Xi^{-1}(u))^2 - 2\rho c_\Xi^{-1}(u) c_\Xi^{-1}(v) + (\rho c_\Xi^{-1}(v))^2}{2(1 - \rho^2)}\right), \quad (40)$$

1 and $c_\Gamma(\cdot; m_\Gamma, \sigma_\Gamma^2)$ and c_Ξ are the cumulative distribution functions associated with the univariate gamma probability
 2 distribution function with mean m_Γ and variance σ_Γ^2 and the Gaussian probability distribution with zero mean and
 3 unit variance: $c_\Gamma(\zeta; m_\Gamma, \sigma_\Gamma^2) = \int_0^\zeta c\gamma^{m_\Gamma^2/\sigma_\Gamma^2-1} \exp(-(m_\Gamma/\sigma_\Gamma^2)\gamma) d\gamma$ and $c_\Xi(\eta) = \int_{-\infty}^\eta \exp(-\xi^2/2)/\sqrt{2\pi} d\xi$.

4 As we mentioned in Sec. 4.2, for many random variables with ‘‘labeled’’ probability distributions, the litera-
 5 ture provides an equivalent characterization as a transformation of a random variable with a uniform, Gaussian, or
 6 other ‘‘labeled’’ probability distribution. Here, reference [103] provides an equivalent characterization for the ran-
 7 dom variables H and S with probability density function $\rho_{(H,S)}(\cdot, \cdot; m_H, \sigma_H^2, m_S, \sigma_S^2, \rho)$ as a transformation of two
 8 statistically independent Gaussian random variables Ξ_1 and Ξ_2 with zero mean and unit variance,

$$\rho_{(\Xi_1, \Xi_2)}(\xi_1, \xi_2) = \rho_\Xi(\xi_1) \times \rho_\Xi(\xi_2), \quad \text{where} \quad \rho_\Xi(\xi) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\xi^2}{2}\right), \quad (41)$$

9 through a mapping $\mathbf{f}(\cdot, \cdot; m_H, \sigma_H^2, m_S, \sigma_S^2, \rho)$ from $\mathbb{R} \times \mathbb{R}$ into $\mathbb{R}^+ \times \mathbb{R}^+$ such that

$$(H, S) = \mathbf{f}(\Xi_1, \Xi_2; m_H, \sigma_H^2, m_S, \sigma_S^2, \rho) = \left(c_\Gamma^{-1}(c_\Xi(\Xi_1); m_H, \sigma_H^2), c_\Gamma^{-1}(c_\Xi(\rho\Xi_1 + \sqrt{1-\rho^2}\Xi_2); m_S, \sigma_S^2) \right); \quad (42)$$

10 in fact, using the rule that probability theory provides for determining the image of a probability density function under
 11 a differentiable, bijective mapping [43], it can be shown that the image of $\rho_{(\Xi_1, \Xi_2)}$ under $\mathbf{f}(\cdot, \cdot; m_H, \sigma_H^2, m_S, \sigma_S^2, \rho)$
 12 is indeed $\rho_{(H,S)}(\cdot, \cdot; m_H, \sigma_H^2, m_S, \sigma_S^2, \rho)$.

13 The parameters m_H and m_S (σ_H^2 and σ_S^2) involved in the expression of $\rho_{(H,S)}(\cdot; m_H, \sigma_H^2, m_S, \sigma_S^2, \rho)$ are the
 14 mean values (respectively, the variances) of H and S ; further, the parameter ρ controls the statistical dependence
 15 of H and S . We infer adequate values $\hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2$, and $\hat{\rho}$ using the method of maximum likelihood, that is, by
 16 maximizing the product of the values taken at the sample values in (37) by the probability density function in (38):

$$(\hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho}) = \text{solution of} \quad \max_{\substack{m_H, \sigma_H^2, m_S, \sigma_S^2 > 0 \\ -1 \leq \rho \leq 1}} \prod_{\ell=1}^n \rho_{(H,S)}(h_\ell^{\text{obs}}, s_\ell^{\text{obs}}; m_H, \sigma_H^2, m_S, \sigma_S^2, \rho). \quad (43)$$

17 The optimization problem in (43) is a general nonlinear programming problem. It need not have a solution and
 18 if a solution exists, it need not be unique. We use a gradient-based nonlinear programming method to solve the
 19 optimization problem in (43); for details about this method, refer to [103].

20 We emphasize that because the support of the univariate gamma probability density function in (39) is the positive
 21 real line, the support of $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ is the first quadrant and thus consistent with the mechanical
 22 and physical constraints that require that the hardening modulus and yield stress be positive.

1 Hereafter, we will denote the probability density function $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ and the mapping $\mathbf{f}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ by $\rho_{(H,S)}$ and \mathbf{f} unless their dependence on the parameters requires emphasis.

3 We note that uncertainties that may result from the availability of only a small or moderate number of samples can
4 be analyzed using sampling distributions and Bayesian approaches [19, 20, 104], but we do not carry out this analysis.

5 8.4 Propagation of uncertainties

6 Because the hardening modulus and yield stress are the only uncertain input variables of the finite element model and
7 the problem is thus “of low dimension,” we use a stochastic expansion method. We proceed by approximating the
8 finite element model by a surrogate model and then using this surrogate model as a substitute for the finite element
9 model in the approximation of statistical descriptors of the springback angle; please refer to Sec. 5.3.

10 8.4.1 Surrogate model

11 We think of the finite element model as a mapping g of any value of the hardening modulus and yield stress into a value
12 of the springback angle. Assuming that g is $\rho_{(H,S)}$ -square-integrable, that is, $\int_{\mathbb{R}^+} \int_{\mathbb{R}^+} g(h, s)^2 \rho_{(H,S)}(h, s) dh ds < +\infty$,
13 we then obtain the polynomial surrogate model of order p —which we will call, hereafter, the surrogate model—
14 as a bivariate polynomial g^p that approximates g as precisely as possible in the $\rho_{(H,S)}$ -weighted least-squares sense,

$$g \approx g^p = \sum_{\alpha+\beta=0}^p c_{(\alpha,\beta)} h^\alpha s^\beta, \quad \text{where } \mathbf{c} = \text{solution of } \min_{\mathbf{d} \in \mathbb{R}^\mu} \frac{1}{2} \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} \left| g(h, s) - \sum_{\alpha+\beta=0}^p d_{(\alpha,\beta)} h^\alpha s^\beta \right|^2 \rho_{(H,S)}(h, s) dh ds, \quad (44)$$

15 where $\mathbf{c} = \{c_{(\alpha,\beta)}, 0 \leq \alpha + \beta \leq p\}$ and μ is the number of monomials in $\{h^\alpha s^\beta, 1 \leq \alpha + \beta \leq p\}$.

16 It can be shown that the characterization of the hardening modulus and yield stress is such that the precision of the
17 surrogate model can be improved systematically and made arbitrarily high by increasing p :

$$\lim_{p \rightarrow +\infty} \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} |g(h, s) - g^p(h, s)|^2 \rho_{(H,S)}(h, s) dh ds = 0. \quad (45)$$

18 **Proof.** As we mentioned in Sec. 5.3, to show the fulfillment of this convergence property, it suffices to show
19 that there exists a constant $\beta > 0$ such that $\int_{\mathbb{R}^+} \int_{\mathbb{R}^+} \exp(\beta \| (h, s) \|) \rho_{(H,S)}(h, s) dh ds < +\infty$, which we show
20 as follows. We carry out the change of variables $(h, s) = (c_\Gamma^{-1}(c_\Xi(w); \hat{m}_H, \hat{\sigma}_H^2), c_\Gamma^{-1}(c_\Xi(z); \hat{m}_S, \hat{\sigma}_S^2))$ to obtain
21 $\int_{\mathbb{R}} \int_{\mathbb{R}} \exp(c \| (c_\Gamma^{-1}(c_\Xi(w); \hat{m}_H, \hat{\sigma}_H^2), c_\Gamma^{-1}(c_\Xi(z); \hat{m}_S, \hat{\sigma}_S^2)) \|) \exp(-w^2/2)/\sqrt{2\pi} \exp(-z^2/2)/\sqrt{2\pi} \exp(-((\hat{\rho}w)^2 -$
22 $2\hat{\rho}wz + (\hat{\rho}z)^2)/2/(1 - \hat{\rho}^2)) dw dz < +\infty$. Because $\exp(-((\hat{\rho}w)^2 - 2\hat{\rho}wz + (\hat{\rho}z)^2)/2/(1 - \hat{\rho}^2)) \exp(-(1 -$
23 $\delta)w^2/2) \exp(-(1 - \delta)z^2/2) \leq 1$ with $\delta = 1 - (|\hat{\rho}| - \hat{\rho}^2)/(1 - \hat{\rho}^2)$, we obtain $\int_{\mathbb{R}} \int_{\mathbb{R}} \exp(c c_\Gamma^{-1}(c_\Xi(w); \hat{m}_H, \hat{\sigma}_H^2))$

1 $\exp(-\delta w^2/2)/\sqrt{2\pi}dw < +\infty$ and $\int_{\mathbb{R}} \exp(c c_{\Gamma}^{-1}(c_{\Xi}(z); \hat{m}_S, \hat{\sigma}_S^2)) \exp(-\delta z^2/2)/\sqrt{2\pi}dz < +\infty$. We then cast the
2 first of the previous inequalities equivalently as $\int_{\mathbb{R}} \exp(c c_{\Gamma}^{-1}(c_{\Xi}(w); \hat{m}_H, \hat{\sigma}_H^2) - \delta w^2/4) \exp(-\delta w^2/4)/\sqrt{2\pi}dw <$
3 $+\infty$ and subsequently apply Hölder's inequality to obtain $(\int_{\mathbb{R}} \exp((2c/\delta) c_{\Gamma}^{-1}(c_{\Xi}(w); \hat{m}_H, \hat{\sigma}_H^2) - w^2/2)/\sqrt{2\pi}dw)^{\delta/2}$
4 $(\int_{\mathbb{R}} \exp(-\delta w^2/4)^{2/(2-\delta)}/\sqrt{2\pi}dw)^{(2-\delta)/2} < +\infty$. Finally, we use the change of variables to obtain $\int_{\mathbb{R}} \exp((2c/\delta)$
5 $c_{\Gamma}^{-1}(c_{\Xi}(w); \hat{m}_H, \hat{\sigma}_H^2) - w^2/2)/\sqrt{2\pi}dw = \int_{\mathbb{R}^+} \exp((2c/\delta)h) \rho_{\Gamma}(h; \hat{m}_H, \hat{\sigma}_H^2)dh$, thus indicating that the previous in-
6 equality holds if $0 \leq c < \delta/(2\hat{\sigma}_H^2)$. Treating the remaining integral similarly, we conclude that $\int_{\mathbb{R}^+} \int_{\mathbb{R}^+} \exp(c\|(h, s)\|)$
7 $\rho_{(H,S)}(h, s)dhds < +\infty$ if $0 \leq c < \delta/(2 \max(\hat{\sigma}_H^2, \hat{\sigma}_S^2))$ and thus that the convergence property in (45) is fulfilled.

8 8.4.2 Implementation by using a nonintrusive projection method

9 We use a nonintrusive projection method. Because recurrence relations for producing orthonormal polynomials are
10 not explicitly known or tabulated in the literature, we cannot use the first nonintrusive projection method mentioned
11 in Sec. 5.3. Instead, we use the one based on the approximation of the integral with respect to $\rho_{(H,S)}$ involved in the
12 objective function in (44) using a quadrature rule for integration with respect to $\rho_{(H,S)}$.

Because the probability density function $\rho_{(H,S)}$ is the image of the probability density function $\rho_{\Xi} \times \rho_{\Xi}$ under the mapping \mathbf{f} (refer to (41) and (42)) we can use the "change of variables" theorem [43, 44] to obtain

$$\begin{aligned} & \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} \left| g(h, s) - \sum_{\alpha+\beta=0}^p d_{(\alpha,\beta)} h^{\alpha} s^{\beta} \right|^2 \rho_{(H,S)}(h, s) dh ds \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} \left| g(f_1(\xi_1, \xi_2), f_2(\xi_1, \xi_2)) - \sum_{\alpha+\beta=0}^p d_{(\alpha,\beta)} (f_1(\xi_1, \xi_2))^{\alpha} (f_2(\xi_1, \xi_2))^{\beta} \right|^2 \rho_{\Xi}(\xi_1) \rho_{\Xi}(\xi_2) d\xi_1 d\xi_2, \quad (46) \end{aligned}$$

thus allowing us to approximate the integral with respect to the probability density function $\rho_{(H,S)}$ involved in the objective function in (44) using a quadrature rule $\{(\xi_{\ell}^{\lambda}, w_{\ell}^{\lambda}), 1 \leq \ell \leq \lambda\}$ for integration with respect to ρ_{Ξ} as

$$\begin{aligned} & \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} \left| g(h, s) - \sum_{\alpha+\beta=0}^p d_{(\alpha,\beta)} h^{\alpha} s^{\beta} \right|^2 \rho_{(H,S)}(h, s) dh ds \\ & \approx \sum_{k=0}^{\lambda} \sum_{\ell=0}^{\lambda} w_k^{\lambda} w_{\ell}^{\lambda} \left| g(f_1(\xi_k^{\lambda}, \xi_{\ell}^{\lambda}), f_2(\xi_k^{\lambda}, \xi_{\ell}^{\lambda})) - \sum_{\alpha+\beta=0}^p d_{(\alpha,\beta)} (f_1(\xi_k^{\lambda}, \xi_{\ell}^{\lambda}))^{\alpha} (f_2(\xi_k^{\lambda}, \xi_{\ell}^{\lambda}))^{\beta} \right|^2. \quad (47) \end{aligned}$$

13 We use the *Gauss-Hermite quadrature rule* with λ nodes for integration with respect to the zero-mean, unit-variance
14 Gaussian probability density function ρ_{Ξ} , that is, the quadrature rule with λ nodes that has only positive weights and
15 is exact for all polynomials up to order $2\lambda - 1$. The nodes and weights of the Gauss-Hermite quadrature rule can be
16 read from tables in the literature for a range of values of λ or computed using standard methods [66, 69, 70].

1 This approximation amounts to the use of the quadrature rule $\{(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda), w_{(k,\ell)}^\lambda\}, 1 \leq k \leq \lambda, 1 \leq \ell \leq \lambda\}$
 2 with $(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda) = \mathbf{f}(\xi_k^\lambda, \xi_\ell^\lambda)$ and $w_{(k,\ell)}^\lambda = w_k^\lambda w_\ell^\lambda$ as the quadrature rule for integration with respect to $\rho_{(H,S)}$.
 3 Thus, we obtain the surrogate model as a solution of a weighted least-squares problem of the form (22).

4 To improve numerical accuracy in the solution of this weighted least-squares problem, we normalize the harden-
 5 ing modulus and yield stress in preparation for the construction of the surrogate model. Specifically, we normalize
 6 the hardening modulus and yield stress by subtracting the mean values \hat{m}_H and \hat{m}_S and dividing by the standard
 7 deviations $\hat{\sigma}_H$ and $\hat{\sigma}_S$. We then obtain the surrogate model as a solution of the weighted least-squares problem

$$g \approx g^{p,\lambda} = \sum_{\alpha+\beta=0}^p c_{(\alpha,\beta)}^\lambda \left(\frac{h - \hat{m}_H}{\hat{\sigma}_H} \right)^\alpha \left(\frac{s - \hat{m}_S}{\hat{\sigma}_S} \right)^\beta, \quad \text{where } \mathbf{c}^\lambda = \text{solution of } \min_{\mathbf{d} \in \mathbb{R}^\mu} \frac{1}{2} (\mathbf{y} - [M]\mathbf{d})^T [W] (\mathbf{y} - [M]\mathbf{d}), \quad (48)$$

8 where \mathbf{y} is the λ^2 -dimensional vector with $y_{(k,\ell)} = g(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda)$, $[M]$ is the $(\lambda^2 \times \mu)$ -dimensional matrix with
 9 $M_{(k,\ell)(\alpha,\beta)} = ((h_{(k,\ell)}^\lambda - \hat{m}_H)/\hat{\sigma}_H)^\alpha ((s_{(k,\ell)}^\lambda - \hat{m}_S)/\hat{\sigma}_S)^\beta$, and $[W]$ the λ^2 -dimensional diagonal matrix with $W_{(k,\ell)(k,\ell)} = w_{(k,\ell)}^\lambda$. To solve the weighted least-squares problem in (48), we select from among the methods in Sec. 5.3 the one
 10 that involves forming the normal equations and solving by Cholesky factorization:
 11

$$\begin{aligned} & \text{form the normal equations } [M^T W M] \mathbf{c}^\lambda = [M^T W] \mathbf{y}; \\ & \text{compute the Cholesky factorization } [M^T W M] = [R^T R]; \\ & \text{first solve } [R]^T \mathbf{g}^\lambda = [M^T W] \mathbf{y} \text{ and then solve } [R] \mathbf{c}^\lambda = \mathbf{g}^\lambda. \end{aligned} \quad (49)$$

12 We note that it suffices to assign a sufficiently high value to the number of nodes λ to ensure that $[M^T W M]$ is
 13 positive definite and therefore that the Cholesky factorization $[M^T W M] = [R^T R]$ exists and is unique.

14 Further, we note that it is worthwhile inspecting this solution method more closely. By setting $\psi_{(\alpha,\beta)}^\lambda =$
 15 $\sum_{(\bar{\alpha}, \bar{\beta}) \leq (\alpha, \beta)} A_{(\bar{\alpha}, \bar{\beta})(\alpha, \beta)} ((h - \hat{m}_H)/\hat{\sigma}_H)^{\bar{\alpha}} ((s - \hat{m}_S)/\hat{\sigma}_S)^{\bar{\beta}}$, where $[A] = [R]^{-1}$, we obtain polynomials $\psi_{(\alpha,\beta)}^\lambda$ that
 16 have the orthonormality property that $\sum_{k=0}^\lambda \sum_{\ell=0}^\lambda w_{(k,\ell)}^\lambda \psi_{(\alpha,\beta)}^\lambda(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda) \psi_{(\bar{\alpha}, \bar{\beta})}^\lambda(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda) = \delta_{(\alpha,\beta)(\bar{\alpha}, \bar{\beta})}$;
 17 in fact, by defining the $(\lambda^2 \times \mu)$ -dimensional matrix $[\Psi]$ such that $\Psi_{(k,\ell)(\alpha,\beta)} = \psi_{(\alpha,\beta)}^\lambda(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda)$, that is,
 18 $[\Psi] = [MA]$, we obtain $[\Psi^T W \Psi] = [A^T M^T W M A] = [I]$, where $[I]$ is the identity matrix. Then, as in (18)
 19 and (19), we obtain the representation $g^{p,\lambda} = \sum_{\alpha+\beta=0}^p g_{(\alpha,\beta)}^\lambda \psi_{(\alpha,\beta)}^\lambda$, where $g_{(\alpha,\beta)}^\lambda = \sum_{k=1}^\lambda w_{(k,\ell)}^\lambda g(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda)$
 20 $\psi_{(\alpha,\beta)}^\lambda(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda)$. Thus, the computation of $[M^T W M] = [R^T R]$, the solution of $[R]^T \mathbf{g}^\lambda = [M^T W] \mathbf{y}$, and the
 21 solution of $[R] \mathbf{c}^\lambda = \mathbf{g}^\lambda$ can be interpreted as the computational construction of orthonormal polynomials, the com-
 22 putation of the coefficients in the representation of the surrogate model in terms of these orthonormal polynomials,
 23 and the rearrangement of this representation to obtain $g^{p,\lambda} = \sum_{|\alpha|=0}^p c_{(\alpha,\beta)}^\lambda ((h - \hat{m}_H)/\hat{\sigma}_H)^\alpha ((s - \hat{m}_S)/\hat{\sigma}_S)^\beta$,

1 respectively; for details about the computational construction of orthonormal polynomials using computational linear
2 algebra in the area of uncertainty quantification, please refer to [105–109].

3 Finally, we note that this construction of the surrogate model is nonintrusive in that it requires only the repeated
4 solution of the finite element model for different values assigned to the hardening modulus and yield stress. The finite
5 element model must be solved for each node within the quadrature for integration with respect to $\rho_{(H,S)}$.

6 8.4.3 Approximation of statistical descriptors

7 Let $\{\xi_\ell, 1 \leq \ell \leq \nu\}$ and $\{\tilde{\xi}_\ell, 1 \leq \ell \leq \nu\}$ be two independent ensembles of i.i.d. samples from the zero-
8 mean, unit-variance Gaussian probability density function ρ_Ξ ; these ensembles can be easily generated using standard
9 methods [24]. Then, because the probability density function $\rho_{(H,S)}$ is the image of $\rho_\Xi \times \rho_\Xi$ under the mapping \mathbf{f} ,
10 we can use $(h_\ell, s_\ell) = \mathbf{f}(\xi_\ell, \tilde{\xi}_\ell)$ to obtain an ensemble $\{(h_\ell, s_\ell), 1 \leq \ell \leq \nu\}$ of i.i.d. samples from $\rho_{(H,S)}$.

11 Once the ensemble $\{(h_\ell, s_\ell), 1 \leq \ell \leq \nu\}$ is available, we use the surrogate model as a substitute for the finite
12 element model in the approximation of statistical descriptors of the springback angle. Specifically, we first use the
13 surrogate model to map each sample into the corresponding sample $y_\ell^{p,\lambda} = g^{p,\lambda}(h_\ell, s_\ell)$ and then apply mathematical
14 statistics methods to the ensemble $\{y_\ell^{p,\lambda}, 1 \leq \ell \leq \nu\}$ of i.i.d. samples obtained for the springback angle; for example,
15 we can approximate the mean and the variance as follows:

$$m_Y \approx m_Y^{p,\lambda,\nu} = \frac{1}{\nu} \sum_{\ell=0}^{\nu} y_\ell^{p,\lambda} \quad \text{and} \quad \sigma_Y^2 \approx (\sigma_Y^{p,\lambda,\nu})^2 = \frac{1}{\nu} \sum_{\ell=0}^{\nu} \left(y_\ell^{p,\lambda} - m_Y^{p,\lambda,\nu} \right)^2. \quad (50)$$

16 We note that if $\psi_{(0,0)}^\lambda = 1$, we also obtain $m_Y \approx m_Y^\lambda = g_{(0,0)}^\lambda$ and $\sigma_Y^2 \approx (\sigma_Y^{p,\lambda})^2 = \sum_{\alpha+\beta=1}^p (g_{(\alpha,\beta)}^\lambda)^2$.

17 8.5 Sensitivity analysis of uncertainties

18 8.5.1 Variance-based method

19 We consider a variance-based sensitivity analysis that aims to gauge the significance of the uncertainties in the hard-
20 ening modulus and those in the yield stress in inducing uncertainties in the springback angle. For this purpose, we
21 cannot immediately use the variance-based method given in Sec. 6.1 because whereas we characterized the hardening
22 modulus and the yield stress as two statistically dependent random variables, the variance-based method given in
23 Sec. 6.1 can only be used to gauge the significance of (subsets of) uncertain input variables that are characterized as
24 statistically independent (subsets of) random variables. Instead, we use an extension of the variance-based method
25 given in Sec. 6.1, an extension which was described in [90, 91] and which relies on conditional probabilities to allow

1 the significance of (subsets of) uncertain input variables to be gauged, even when they are characterized as statistically
 2 dependent (subsets of) random variables. Based on this extension, we define the following significance descriptors:

$$s_H = \int_{\mathbb{R}^+} \left(\int_{\mathbb{R}^+} (g(h, s) - m_Y(s))^2 \rho_{H|S}(h|s) dh \right) \rho_S(s) ds, \quad m_Y(s) = \int_{\mathbb{R}^+} g(h, s) \rho_{H|S}(h|s) dh, \quad (51)$$

$$s_S = \int_{\mathbb{R}^+} \left(\int_{\mathbb{R}^+} (g(h, s) - m_Y(h))^2 \rho_{S|H}(s|h) ds \right) \rho_H(h) dh, \quad m_Y(h) = \int_{\mathbb{R}^+} g(h, s) \rho_{S|H}(s|h) ds,$$

3 where ρ_H and ρ_S are the marginal probability density functions such that $\rho_H(h) = \int_{\mathbb{R}^+} \rho_{(H,S)}(h, s) ds$ and $\rho_S(s) =$
 4 $\int_{\mathbb{R}^+} \rho_{(H,S)}(h, s) dh$, and $\rho_{H|S}(\cdot|s)$ and $\rho_{S|H}(\cdot|h)$ are the conditional probability density functions such that $\rho_{H|S}(h|s)$
 5 $= \rho_{(H,S)}(h, s) / \rho_S(s)$ and $\rho_{S|H}(s|h) = \rho_{(H,S)}(h, s) / \rho_H(h)$. As in Sec. 6.1, we can interpret s_H as the ρ_S -weighted
 6 average of the variance $\int_{\mathbb{R}^+} (g(h, s) - m_Y(s))^2 \rho_{H|S}(h|s) dh$ that the springback angle would exhibit owing to the
 7 uncertainties in the hardening modulus if the yield stress were set equal to the specific value s , and we can interpret s_S
 8 similarly. Thus, the higher s_H and s_S , the more significant the uncertainties in the hardening modulus and the yield
 9 stress, respectively, are in inducing uncertainties in the springback angle.

10 We note that if the hardening modulus and yield stress were characterized as statistically independent random
 11 variables, that is, if $\rho_{(H,S)} = \rho_H \times \rho_S$ and thus $\rho_{H|S}(\cdot|s) = \rho_H$ and $\rho_{S|H}(\cdot|h) = \rho_S$, the significance descriptors s_H
 12 and s_S would coincide with those of the variance-based method given in Sec. 6.1.

13 The computation of the significance descriptors s_H and s_S is a problem of numerical integration. We carry out
 14 this numerical integration using quadrature rules that are similar to those that we used in the previous section for the
 15 propagation of uncertainties. We provide details about the computation of s_S ; the significance descriptor s_H can be
 16 computed analogously. We use the ‘‘change of variables’’ theorem to obtain

$$s_S = \int_{\mathbb{R}} \int_{\mathbb{R}} \left(g(f_1(\xi_1, \xi_2), f_2(\xi_1, \xi_2)) - m_Y(\xi_1) \right)^2 \rho_{\Xi}(\xi_1) \rho_{\Xi}(\xi_2) d\xi_1 d\xi_2, \quad (52)$$

17 where

$$m_Y(\xi_1) = \int_{\mathbb{R}} g(f_1(\xi_1, \xi_2), f_2(\xi_1, \xi_2)) \rho_{\Xi}(\xi_2) d\xi_2 \quad (53)$$

18 and \mathbf{f} is still defined as in (42), thus allowing us to approximate s_S as follows:

$$s_S \approx s_S^{p,\lambda} = \sum_{k=0}^{\lambda} \sum_{\ell=0}^{\lambda} w_k^{\lambda} w_{\ell}^{\lambda} \left(g^{p,\lambda}(f_1(\xi_k^{\lambda}, \xi_{\ell}^{\lambda}), f_2(\xi_k^{\lambda}, \xi_{\ell}^{\lambda})) - m_Y^{p,\lambda}(\xi_k^{\lambda}) \right)^2, \quad (54)$$

1 where

$$m_Y^{p,\lambda}(\xi_k^\lambda) = \sum_{\ell=0}^{\lambda} w_\ell^\lambda g^{p,\lambda}(f_1(\xi_k^\lambda, \xi_\ell^\lambda), f_2(\xi_k^\lambda, \xi_\ell^\lambda)), \quad (55)$$

2 in which we substituted the surrogate model for the finite element model to lower the computational cost.

3 8.5.2 Differentiation-based method

4 We consider a differentiation-based sensitivity analysis that aims to determine the sensitivity of the variance of the
5 springback angle with respect to changes in parameters involved in the characterization of the hardening modulus and
6 yield stress. Collecting the parameters of the characterization of the hardening modulus and yield stress in a vector

$$\mathbf{p} = (m_H, \sigma_H^2, m_S, \sigma_S^2, \rho) \quad (56)$$

7 and denoting the mean and variance of the springback angle by

$$\begin{aligned} m_Y(\mathbf{p}) &= \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} g(h, s) \rho_{(H,S)}(h, s; \mathbf{p}) dh ds, \\ \sigma_Y^2(\mathbf{p}) &= \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} (g(h, s) - m_Y(\mathbf{p}))^2 \rho_{(H,S)}(h, s; \mathbf{p}) dh ds, \end{aligned} \quad (57)$$

8 we give special interest to the partial derivatives of the variance σ_Y^2 of the springback angle with respect to the
9 variances σ_H^2 , and σ_S^2 of the hardening modulus and yield stress evaluated at the values obtained in (43), that is,

$$s_{\sigma_H^2} = \left. \frac{\partial \sigma_Y^2(\mathbf{p})}{\partial \sigma_H^2} \right|_{\mathbf{p}=\hat{\mathbf{p}}} \quad \text{and} \quad s_{\sigma_S^2} = \left. \frac{\partial \sigma_Y^2(\mathbf{p})}{\partial \sigma_S^2} \right|_{\mathbf{p}=\hat{\mathbf{p}}}. \quad (58)$$

10 Our special interest in sensitivity descriptors such as those in (58) follows from the fact that in a probabilistic investi-
11 gation of manufacturing variability, changes in variance can sometimes be usefully interpreted in terms of changes in
12 manufacturing tolerances; then, determining sensitivity descriptors such as those in (58) can provide valuable insight
13 towards designing, controlling, and optimizing tolerances associated with the manufacturing process.

14 To facilitate the computation of $s_{\sigma_H^2}$ and $s_{\sigma_S^2}$, we interchange the derivatives and the integrals involved in their
15 expressions in (58). In this way, their computation becomes a problem of numerical integration, which we carry out
16 using quadrature rules similar to those that we used previously. We provide details about the computation of $s_{\sigma_H^2}$; the

1 sensitivity descriptor $s_{\sigma_S^2}$ can be computed analogously. We use the ‘‘change of variables’’ theorem to obtain

$$s_{\sigma_H^2} = \frac{\partial \sigma_Y^2(\mathbf{p})}{\partial p_j} \Big|_{\mathbf{p}=\hat{\mathbf{p}}} = \frac{\partial \int_{\mathbb{R}} \int_{\mathbb{R}} \left(g(\mathbf{f}(\xi_1, \xi_2; \mathbf{p})) - m_Y(\mathbf{p}) \right)^2 \rho_{\Xi}(\xi_1) \rho_{\Xi}(\xi_2) d\xi_1 d\xi_2}{\partial \sigma_H^2} \Big|_{\mathbf{p}=\hat{\mathbf{p}}}, \quad (59)$$

2 thus, after interchanging the differentiation and integration operations, we obtain the following approximation:

$$s_{\sigma_H^2} \approx s_{\sigma_H^2}^{p,\lambda} = \sum_{k=0}^{\lambda} \sum_{\ell=0}^{\lambda} w_k^{\lambda} w_{\ell}^{\lambda} \left(2g^{p,\lambda}(\mathbf{f}(\xi_k^{\lambda}, \xi_{\ell}^{\lambda}; \hat{\mathbf{p}})) - 2m_Y^{\lambda}(\hat{\mathbf{p}}) \right) \left(\nabla_{(h,s)} g^{p,\lambda}(h, s) \Big|_{(h,s)=\mathbf{f}(\xi_k^{\lambda}, \xi_{\ell}^{\lambda}; \hat{\mathbf{p}})} \right)^T \frac{\partial \mathbf{f}(\xi_k^{\lambda}, \xi_{\ell}^{\lambda}; \mathbf{p})}{\partial \sigma_H^2} \Big|_{\mathbf{p}=\hat{\mathbf{p}}}, \quad (60)$$

3 where we again substituted the surrogate model for the finite element model to lower the computational cost. We
 4 obtain the needed partial derivatives of $\mathbf{f}(\xi_k^{\lambda}, \xi_{\ell}^{\lambda}; m_H, \sigma_H^2, m_S, \sigma_S^2, \rho)$ using the chain rule and the algorithm in [110].

5 It can be shown that the characterization of the hardening modulus and yield stress is such that the interchange of
 6 the differentiation and integration operations is permitted.

7 **Proof.** We will show that $\partial \int_{\mathbb{R}} (\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, b))^{\alpha} \rho_{\Xi}(\xi) d\xi / \partial b|_{b=\hat{b}} = \int_{\mathbb{R}} \partial (\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, b))^{\alpha} / \partial b|_{b=\hat{b}} \rho_{\Xi}(\xi) d\xi$,
 8 where $\tilde{c}_{\Gamma}^{-1}(\zeta; a, b) = \int_0^{\zeta} \tilde{\rho}_{\Gamma}(\gamma; a, b) d\gamma$, in which $\tilde{\rho}_{\Gamma}(\gamma; a, b) = c\gamma^{a-1} \exp(-\gamma/b)$ if $\gamma \geq 0$ and $\tilde{\rho}_{\Gamma}(\gamma; a, b) = 0$ other-
 9 wise. First, we use the chain rule to obtain $\partial (\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, \hat{b}))^{\alpha} / \partial b|_{b=\hat{b}} = (\alpha/\hat{b})(\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, \hat{b}))^{\alpha}$. Then, because
 10 the function that maps b onto $\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, b)$ is continuous and monotonically increasing, we use the mean-value the-
 11 orem to obtain $|(\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, \hat{b} + \delta b))^{\alpha} - (\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, \hat{b}))^{\alpha}| \leq (\alpha/(\hat{b} - |\delta b|))(\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, \hat{b} + |\delta b|))^{\alpha} |\delta b|$ if $|\delta b|$
 12 is sufficiently small. Because the univariate gamma probability distribution has finite moments of any order, the func-
 13 tion that maps ξ onto $(\alpha/(\hat{b} - |\delta b|))(\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, \hat{b} + |\delta b|))^{\alpha}$ is ρ_{Ξ} -integrable. Thus, because the magnitude of the
 14 finite difference can be bounded from above by a ρ_{Ξ} -integrable function, the dominated convergence theorem implies
 15 that $\lim_{\delta b \rightarrow 0} \int_{\mathbb{R}} (((\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, \hat{b} + \delta b))^{\alpha} - (\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, \hat{b}))^{\alpha}) / \delta b) \rho_{\Xi}(\xi) d\xi = \int_{\mathbb{R}} \lim_{\delta b \rightarrow 0} (((\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, \hat{b} +$
 16 $\delta b))^{\alpha} - (\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); \hat{a}, \hat{b}))^{\alpha}) / \delta b) \rho_{\Xi}(\xi) d\xi$, which concludes the proof. In a similar manner, it can be shown that
 17 $\partial \int_{\mathbb{R}} (\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); a, \hat{b}))^{\alpha} \rho_{\Xi}(\xi) d\xi / \partial a|_{a=\hat{a}} = \int_{\mathbb{R}} \partial (\tilde{c}_{\Gamma}^{-1}(c_{\Xi}(\xi); a, \hat{b}))^{\alpha} / \partial a|_{a=\hat{a}} \rho_{\Xi}(\xi) d\xi$, that the validity of these in-
 18 terchanges extends to the bivariate gamma probability distribution, that the chain rule applies, and therefore, at least
 19 after substituting the surrogate model for the finite element model, that the differentiation and integration operations
 20 may be interchanged in (59) to obtain (60), but we omit the details for the sake of brevity.

21 8.6 Key role played by the characterization of uncertainties

22 To further emphasize the key role played by a judicious characterization of uncertainties, we reiterate that our char-
 23 acterization of the hardening modulus and yield stress is consistent with the applicable mechanical and physical

1 constraints (Sec. 8.3), facilitates the propagation of uncertainties by allowing convergence properties to be demon-
 2 strated and various algorithmic ingredients, such as quadrature rules, to be efficiently obtained (Sec. 8.4), and permits
 3 the sensitivity to uncertainties to be explored by allowing insightful analyses to be carried out (Sec. 8.5).

4 9. NUMERICAL RESULTS

5 We obtained numerical results by assigning the values $l = 20$ mm, $r = 3$ mm, $s = 1$ mm, $u = 6$ mm, $v = 1$ mm,
 6 and $w = 5$ mm to the geometrical characteristics depicted in Fig. 1(1) and values of 210 GPa and 0.3 to the Young's
 7 modulus and Poisson coefficient involved in the elastoplastic constitutive equation with isotropic hardening.

8 We assumed that the data—from which we had to infer adequate values for the parameters involved in the charac-
 9 terization of the hardening modulus and yield stress—consist of the $n = 25$ samples listed in Table 1.

10 We emphasize that the samples in Table 1 have been numerically generated and are not representative of the
 11 variability that may be present in actual mechanical behavior of real steel sheets. Thus, the results to follow cannot be
 12 used to draw conclusions on forming processes but serve only to illustrate some of the methods that we described.

13 9.1 Characterization of uncertainties

First, we solved the optimization problem in (43) to obtain adequate values for the parameters involved in the charac-
 terization of the hardening modulus and yield stress:

$$\hat{m}_H = 1495 \text{ MPa}, \quad \hat{\sigma}_H^2 = 1390 \text{ MPa}^2, \quad \hat{m}_S = 396 \text{ MPa}, \quad \hat{\sigma}_S^2 = 660 \text{ MPa}^2, \quad \hat{\rho} = -0.223.$$

The mean values \hat{m}_H and \hat{m}_S and the variances $\hat{\sigma}_H^2$ and $\hat{\sigma}_S^2$ correspond to the coefficients of variation

$$\frac{\hat{\sigma}_H}{\hat{m}_H} = 2.49\% \quad \text{and} \quad \frac{\hat{\sigma}_S}{\hat{m}_S} = 6.49\%,$$

14 thus indicating that the characterization of the hardening modulus and yield stress as random variables H and S with
 15 probability density function $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ introduces relatively less uncertainties in the hardening
 16 modulus than in the yield stress. Further, because $\hat{\rho} = -0.223$ is negative, the statistical dependence between H
 17 and S is such that if H takes a higher value, the probability that S takes a lower value increases. Figure 2 shows a few
 18 contours of the probability density function $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$.

1 9.2 Propagation of uncertainties

2 9.2.1 Surrogate model and implementation using a nonintrusive projection method

3 Next, we approximated the finite element model by a surrogate model. We obtained results using a range of values for
 4 the order p of the surrogate model and the parameter λ that controls the number of nodes in the quadrature for inte-
 5 gration with respect to the probability density function $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$. We discuss the convergence
 6 with respect to p and λ later; for now, we present detailed results that we obtained with $p = 4$ and $\lambda = 5$.

7 We constructed the Gauss-Hermite quadrature rule, which provided the basis for the quadrature rule for integration
 8 with respect to $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ (Fig. 3(1)), whereupon we solved the finite element model for each
 9 node of this quadrature for integration with respect to $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ (Fig. 3(2)) and we solved the
 10 optimization problem in (48) to obtain the surrogate model (Fig. 3(3)).

11 The computational cost of constructing the surrogate model was dominated by the repeated solution of the finite
 12 element model for each node in the quadrature for integration with respect to $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$; for $p =$
 13 4 and $\lambda = 5$, the construction of the surrogate model required $\lambda^2 = 25$ finite element model solutions.

14 9.2.2 Approximation of statistical descriptors

15 Subsequently, we used the surrogate model as a substitute for the finite element model in approximating statistical
 16 descriptors for the springback angle. We generated two independent ensembles of (a sufficiently large number $v =$
 17 1, 000, 000 of) i.i.d. samples from the zero-mean, unit-variance Gaussian probability density function. On the basis of
 18 these, we generated an ensemble of i.i.d. samples from $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ (Fig. 4(1)), and we used the
 19 surrogate model to map each of these samples into the corresponding sample of the springback angle and then applied
 20 mathematical statistics methods to the ensemble of i.i.d. samples of the springback angle thus obtained (Fig. 4(2)).

Figure 4(2) shows the approximate probability density function of the springback angle obtained for $p = 4$
 and $\lambda = 5$ by applying the kernel density estimation method [39] to the ensemble of i.i.d. samples of the springback
 angle. Further, for $p = 4$ and $\lambda = 5$, we obtained the following approximations of the mean and variance:

$$m_Y^{p,\lambda} = 0.0528 \text{ rad}, \quad (\sigma_Y^{p,\lambda})^2 = 5.34 \times 10^{-6} \text{ rad}^2, \quad \frac{\sigma_Y^{p,\lambda}}{m_Y^{p,\lambda}} = 4.38\%.$$

9.3 Sensitivity analysis of uncertainties

Finally, we carried out a sensitivity analysis of uncertainties to gain insight into the manner in which uncertainties in the hardening modulus and yield stress induce uncertainties in the springback angle. Using the variance-based method, we obtained the following approximations of the significance descriptors for $p = 4$ and $\lambda = 5$:

$$s_H^{p,\lambda} = 0.102 \times 10^{-6} \text{ rad}^2 \quad \text{and} \quad s_S^{p,\lambda} = 5.30 \times 10^{-6} \text{ rad}^2,$$

that is, $s_H^{p,\lambda}/(\sigma_Y^{p,\lambda})^2 = 0.0192$ and $s_S^{p,\lambda}/(\sigma_Y^{p,\lambda})^2 = 0.99$; thus, the uncertainties in the yield stress are more significant than those in the hardening modulus in inducing uncertainties in the springback angle. Using the differentiation-based method, we obtained the following approximations of the sensitivity descriptors for $p = 4$ and $\lambda = 5$:

$$s_{\sigma_H^2}^{p,\lambda} = -0.0395 \times 10^{-9} \text{ rad}^2/\text{MPa}^2 \quad \text{and} \quad s_{\sigma_S^2}^{p,\lambda} = 8.31 \times 10^{-9} \text{ rad}^2/\text{MPa}^2,$$

that is, $(\hat{\sigma}_H^2/(\sigma_Y^{p,\lambda})^2)s_{\sigma_H^2}^{p,\lambda} = -0.0103$ and $(\hat{\sigma}_S^2/(\sigma_Y^{p,\lambda})^2)s_{\sigma_S^2}^{p,\lambda} = 1.02$; thus, the variance of the springback angle is more sensitive to changes in the variance of the yield stress than to changes in the variance of the hardening modulus.

9.4 Numerical convergence study

We conducted a numerical convergence study to examine the impact that the values assigned to p and λ have on the results. Specifically, we repeated the construction of the surrogate model and the approximation of statistical, significance, and sensitivity descriptors of the springback angle for several values of p , setting $\lambda = p + 1$. Figure 5 illustrates the convergence of these approximations. We can observe that the results presented previously for $p = 4$ and $\lambda = 5$ have converged reasonably with respect to the order p of the surrogate model and the parameter λ that controls the number of nodes in the quadrature for integration with respect to $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$.

10. CONCLUSION

We offered a short overview of a number of methods reported in the computational-mechanics literature for quantifying uncertainties in engineering applications. We covered the characterization, propagation, and sensitivity analysis of uncertainties as they apply to parametric, nonparametric, output-prediction-error, and generalized approaches and to problems of low, moderate, and high dimension. We included recent advances in the propagation and sensitivity analysis of uncertainties characterized by arbitrary probability distributions that may exhibit statistical dependence.

1 ACKNOWLEDGMENTS

2 This work was supported by ArcelorMittal and by the Université de Liège through a starting grant allocated to the
3 first author. The authors would like to thank the two anonymous reviewers for their thoughtful suggestions.

4 REFERENCES

- 5 1. Hills, R., Pilch, M., Dowding, K., Red-Horse, J., Paez, T., Babuška, I., and Tempone, R., Validation
6 challenge workshop, *Computer Methods in Applied Mechanics and Engineering*, 197:29–32, 2008, DOI:
7 10.1016/j.cma.2007.10.016.
- 8 2. Capiez-Lernout, E. and Soize, C., Robust design optimization in computational mechanics, *Journal of Applied*
9 *Mechanics, Transactions ASME*, 75:0210011, 2008, DOI: 10.1115/1.2775493.
- 10 3. Higdon, D., Gattiker, J., Williams, B., and Rightley, M., Computer model calibration using
11 high-dimensional output, *Journal of the American Statistical Association*, 103:570–583, 2008, DOI:
12 10.1198/016214507000000888.
- 13 4. Kree, P. and Soize, C., *Mathematics of Random Phenomena*, D. Reidel Publishing Company, Dordrecht, Nether-
14 lands, 1983.
- 15 5. Zhang, D., *Stochastic Methods for Flow in Porous Media: Coping with Uncertainties*, Academic Press, San
16 Diego, California, 2001.
- 17 6. Ghanem, R. and Spanos, P., *Stochastic Finite Elements: A Spectral Approach*, Dover Publications, Mineola,
18 New York, 2003.
- 19 7. Roberts, J. and Spanos, P., *Random Vibration and Statistical Linearization*, Dover Publications, Mineola, New
20 York, 2003.
- 21 8. Batou, A. and Soize, C., Rigid multibody system dynamics with uncertain rigid bodies, *Multibody System*
22 *Dynamics*, 27:285–319, 2012, DOI: 10.1007/s11044-011-9279-2.
- 23 9. Soize, C., Nonparametric model of random uncertainties for reduced matrix models in structural dynamics,
24 *Probabilistic Engineering Mechanics*, 15:277–294, 2000, DOI: 10.1016/S0266-8920(99)00028-4.
- 25 10. Beck, J. and Katafygiotis, L., Updating models and their uncertainties. I: Bayesian statistical framework, *Jour-*
26 *nal of Engineering Mechanics*, 124:455–461, 1998, DOI: 10.1061/(ASCE)0733-9399(1998)124:4(455).
- 27 11. Ljung, L., *System Identification: Theory for the User*, Prentice Hall, Upper Saddle River, New Jersey, 1998.

- 1 12. Kennedy, M. and O'Hagan, A., Bayesian calibration of computer models, *Journal of the Royal Statistical*
2 *Society. Series B: Statistical Methodology*, 63:425–450, 2001, DOI: 10.1111/1467-9868.00294.
- 3 13. Soize, C., Generalized probabilistic approach of uncertainties in computational dynamics using random ma-
4 trices and polynomial chaos decompositions, *International Journal for Numerical Methods in Engineering*,
5 81:939–970, 2010, DOI: 10.1002/nme.2712.
- 6 14. Soize, C., Bayesian posteriors of uncertainty quantification in computational structural dynamics for low- and
7 medium-frequency ranges, *Computers and Structures*, 2013, DOI: 10.1016/j.compstruc.2013.03.020, to appear.
- 8 15. McLachlan, G. and Peel, D., *Finite Mixture Models*, Wiley-Interscience, New York, 2000.
- 9 16. Kotz, S., Balakrishnan, N., and Johnson, N., *Continuous Multivariate Distributions*, Wiley-Interscience, New
10 York, 2000.
- 11 17. Soize, C., Non-Gaussian positive-definite matrix-valued random fields for elliptic stochastic partial dif-
12 ferential operators, *Computer Methods in Applied Mechanics and Engineering*, 195:26–64, 2006, DOI:
13 10.1016/j.cma.2004.12.014.
- 14 18. Desceliers, C., Ghanem, R., and Soize, C., Maximum likelihood estimation of stochastic chaos representations
15 from experimental data, *International Journal for Numerical Methods in Engineering*, 66:978–1001, 2006,
16 DOI: 10.1002/nme.1576.
- 17 19. Ghanem, R. and Doostan, A., On the construction and analysis of stochastic models: Characterization and
18 propagation of the errors associated with limited data, *Journal of Computational Physics*, 217:63–81, 2006,
19 DOI: 10.1016/j.jcp.2006.01.037.
- 20 20. Arnst, M., Ghanem, R., and Soize, C., Identification of Bayesian posteriors for coefficients of chaos expansions,
21 *Journal of Computational Physics*, 229:3134–3154, 2010, DOI: 10.1016/j.jcp.2009.12.033.
- 22 21. Soize, C., Identification of high-dimension polynomial chaos expansions with random coefficients for non-
23 gaussian tensor-valued random fields using partial and limited experimental data, *Computer Methods in Applied*
24 *Mechanics and Engineering*, 199:2150–2164, 2010, DOI: 10.1016/j.cma.2010.03.013.
- 25 22. Soize, C., A computational inverse method for identification of non-Gaussian random fields using the Bayesian
26 approach in very high dimension, *Computer Methods in Applied Mechanics and Engineering*, 200:3083–3099,
27 2011, DOI: 10.1016/j.cma.2011.07.005.
- 28 23. Rosenblatt, M., Remarks on a multivariate transformation, *The Annals of Mathematical Statistics*, 23:470–472,

- 1 1952, DOI: 10.1214/aoms/117772939.
- 2 24. Robert, C. and Casella, G., *Monte Carlo Statistical Methods*, Springer, New York, 2010.
- 3 25. Guillemintot, J. and Soize, C., On the statistical dependence for the components of random elasticity tensors
4 exhibiting material symmetry properties, *Journal of Elasticity*, 2012, DOI: 10.1007/s10659-012-9396-z.
- 5 26. Chebli, H. and Soize, C., Experimental validation of a nonparametric probabilistic model of nonhomogeneous
6 uncertainties for dynamical systems, *Journal of the Acoustical Society of America*, 115:697–705, 2004, DOI:
7 10.1121/1.1639335.
- 8 27. Mignolet, M. and Soize, C., Nonparametric stochastic modeling of linear systems with prescribed
9 variance of several natural frequencies, *Probabilistic Engineering Mechanics*, 23:267–278, 2008, DOI:
10 10.1016/j.probenmech.2007.12.027.
- 11 28. Cottreau, R., Clouteau, D., and Soize, C., Construction of a probabilistic model for impedance
12 matrices, *Computer Methods in Applied Mechanics and Engineering*, 196:2252–2268, 2007, DOI:
13 10.1016/j.cma.2006.12.001.
- 14 29. Ritto, T., Soize, C., and Sampaio, R., Non-linear dynamics of a drill-string with uncertain model
15 of the bit-rock interaction, *International Journal of Non-Linear Mechanics*, 44:865–876, 2009, DOI:
16 10.1016/j.ijnonlinmec.2009.06.003.
- 17 30. Soize, C. and Poloskov, I., Time-domain formulation in computational mechanics for linear viscoelastic media
18 with model uncertainties and stochastic excitation, *Computers and Mathematics with Applications*, 64:3594–
19 3612, 2012, DOI: 10.1016/j.camwa.2012.09.010.
- 20 31. Mignolet, M., Soize, C., and Avalos, J., Nonparametric stochastic modeling of structures with uncertain bound-
21 ary conditions/coupling between substructures, *AIAA Journal*, 51:1296–1308, 2013, DOI: 10.2514/1.J051555.
- 22 32. Jaynes, E., *Probability Theory The Logic of Science*, Cambridge University Press, New York, 1987.
- 23 33. Wang, J. and Zabarab, N., A Bayesian inference approach to the inverse heat conduction problem, *International*
24 *Journal of Heat and Mass Transfer*, 47:3927–3941, 2004, DOI: 10.1016/j.ijheatmasstransfer.2004.02.028.
- 25 34. Marzouk, Y., Najm, H., and Rahn, L., Stochastic spectral methods for efficient Bayesian solution of inverse
26 problems, *Journal of Computational Physics*, 224:560–586, 2007, DOI: 10.1016/j.jcp.2006.10.010.
- 27 35. Liu, W., Siad, L., Tian, R., Lee, S., Lee, D., Yin, X., Chen, W., Chan, S., Olson, G., Lindgen, L., Horstemeyer,
28 M., Chang, Y., Choi, J., and Kim, Y., Complexity science of multiscale materials via stochastic computations,

- 1 *International Journal for Numerical Methods in Engineering*, 80:932–978, 2009, DOI: 10.1002/nme.2578.
- 2 36. Flath, H., Wilcox, L., Akçelik, V., Hill, J., Van Bloemen-Waanders, B., and Ghattas, O., Fast algorithms for
3 Bayesian uncertainty quantification in large-scale linear inverse problems based on low-rank partial Hessian
4 approximations, *SIAM Journal on Scientific Computing*, 33:407–432, 2011, DOI: 10.1137/090780717.
- 5 37. Berry, R., Najm, H., Debusschere, B., Marzouk, Y., and Adalsteinsson, H., Data-free inference of the joint
6 distribution of uncertain model parameters, *Journal of Computational Physics*, 231:2180–2198, 2012, DOI:
7 10.1016/j.jcp.2011.10.031.
- 8 38. Schwab, C. and Stuart, A., Sparse deterministic approximation of Bayesian inverse problems, *Inverse Problems*,
9 28:045003, 2012, DOI: 10.1088/0266-5611/28/4/045003.
- 10 39. Scott, D., *Multivariate Density Estimation: Theory, Practice, and Visualization*, Wiley, New York, 1992.
- 11 40. Cramér, H., *Mathematical Methods of Statistics*, Princeton University Press, Princeton, New Jersey, 1997.
- 12 41. Ganapathysubramanian, B. and Zabaras, N., A non-linear dimension reduction methodology for generat-
13 ing data-driven stochastic input models, *Journal of Computational Physics*, 227:6612–6637, 2008, DOI:
14 10.1016/j.jcp.2008.03.023.
- 15 42. Soize, C., Construction of probability distributions in high dimension using the maximum entropy principle:
16 Applications to stochastic processes, random fields and random matrices, *International Journal for Numerical*
17 *Methods in Engineering*, 76:1583–1611, 2008, DOI: 10.1002/nme.2385.
- 18 43. Casella, G. and Berger, R., *Statistical Inference*, Duxbury, Pacific Grove, California, 2002.
- 19 44. Dudley, R., *Real Analysis and Probability*, Cambridge University Press, Cambridge, United Kingdom, 2002.
- 20 45. Calfish, R., Monte Carlo and quasi-Monte Carlo methods, *Acta Numerica*, 7:1–49, 1998, DOI:
21 10.1017/S0962492900002804.
- 22 46. Cliffe, K., Giles, M., Scheichl, R., and Teckentrup, A., Multilevel Monte Carlo methods and applications
23 to elliptic PDEs with random coefficients, *Computing and Visualization in Science*, 14:3–15, 2011, DOI:
24 10.1007/s00791-011-0160-x.
- 25 47. Boyaval, S. and Lelièvre, T., A variance reduction method for parametrized stochastic differential equations
26 using the reduced basis paradigm, *Communications in Mathematical Sciences*, 8:735–762, 2010.
- 27 48. Bertsekas, D., *Convex Analysis and Optimization*, Athena Scientific, Belmont, Massachusetts, 2003.
- 28 49. Dunkl, C. and Xu, Y., *Orthogonal Polynomials of Several Variables*, Cambridge University Press, Cambridge,

1 United Kingdom, 2001.

- 2 50. Ernst, O., Mugler, A., Starkloff, H., and Ullman, E., On the convergence of generalized polynomial chaos ex-
3 pansion, *Mathematical Modelling and Numerical Analysis*, 46:317–339, 2012, DOI: 10.1051/m2an/2011045.
- 4 51. Soize, C. and Ghanem, R., Physical systems with random uncertainties: Chaos representations
5 with arbitrary probability measure, *SIAM Journal on Scientific Computing*, 26:395–410, 2004, DOI:
6 10.1137/S1064827503424505.
- 7 52. Matthies, H. and Keese, A., Galerkin methods for linear and nonlinear elliptic stochastic partial differ-
8 ential equations, *Computer Methods in Applied Mechanics and Engineering*, 194:1295–1331, 2005, DOI:
9 10.1016/j.cma.2004.05.027.
- 10 53. Le Maître, O. and Knio, O., *Spectral Methods for Uncertainty Quantification: With Applications to Computa-*
11 *tional Fluid Dynamics*, Springer, New York, 2010.
- 12 54. Xiu, D., *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton University
13 Press, Princeton, New Jersey, 2010.
- 14 55. Grigoriu, M., *Stochastic Systems: Uncertainty Quantification and Propagation*, Springer-Verlag, London,
15 2012.
- 16 56. Xiu, D., Efficient collocational approach for parametric uncertainty analysis, *Communications in Computa-*
17 *tional Physics*, 2:293–309, 2007.
- 18 57. Ghanem, R., Hybrid stochastic finite elements and generalized Monte Carlo simulation, *Journal of Applied*
19 *Mechanics*, 65:1004–1009, 1998, DOI: 10.1115/1.2791894.
- 20 58. Ghanem, R. and Ghiocel, D., A new implementation of the spectral stochastic finite element method for stochas-
21 tic constitutive relations, In *ASCE 12th Engineering Mechanics Conference*, La Jolla, California, 1998.
- 22 59. Barthelmann, V., Novak, E., and Ritter, K., High dimensional polynomial interpolation on sparse grids, *Ad-*
23 *vances in Computational Mathematics*, 12:273–288, 2000, DOI: 10.1023/A:1018977404843.
- 24 60. Xiu, D. and Hesthaven, J., High-order collocation methods for differential equations with random inputs, *SIAM*
25 *Journal on Scientific Computing*, 27:1118–1139, 2006, DOI: 10.1137/040615201.
- 26 61. Babuška, I., Nobile, F., and Tempone, R., A stochastic collocation method for elliptic partial differen-
27 tial equations with random input data, *SIAM Journal on Numerical Analysis*, 45:1005–1034, 2007, DOI:
28 10.1137/050645142.

- 1 62. Nobile, F., Tempone, R., and Webster, C., A sparse grid stochastic collocation method for partial differ-
2 ential equations with random input data, *SIAM Journal on Numerical Analysis*, 46:2309–2345, 2008, DOI:
3 10.1137/060663660.
- 4 63. Björck, A., *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia Pennsylvania, 1996.
- 5 64. Golub, G. and Van Loan, C., *Matrix computations*, The Johns Hopkins University Press, Baltimore, Maryland,
6 1996.
- 7 65. Nocedal, J. and Wright, S., *Numerical Optimization*, Springer, New York, New York, 2006.
- 8 66. Abramowitz, M. and Stegun, I., *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathe-*
9 *matical Tables*, Dover Publications, Mineola, New York, 1972.
- 10 67. Novak, E. and Ritter, K., Simple cubature formulas with high polynomial exactness, *Constructive Approxima-*
11 *tion*, 15:499–522, 1999, DOI: 10.1007/s003659900119.
- 12 68. Davis, P. and Rabinowitz, P., *Methods of Numerical Integration*, Dover Publications, Mineola, New York, 2007.
- 13 69. Golub, G. and Meurant, G., *Matrices, Moments and Quadrature with Applications*, Princeton University Press,
14 Princeton, New Jersey, 2009.
- 15 70. Holtz, M., *Sparse Grid Quadrature in High Dimensions with Applications in Finance and Insurance*, Springer,
16 Berlin, Heidelberg, Germany, 2010.
- 17 71. Le Maître, O., Knio, O., Najm, H., and Ghanem, R., Uncertainty propagation using Wiener-Haar expansions,
18 *Journal of Computational Physics*, 197:28–57, 2004, DOI: 10.1016/j.jcp.2003.11.033.
- 19 72. Le Maître, O., Najm, H., Ghanem, R., and Knio, O., Multi-resolution analysis of Wiener-type uncertainty
20 propagation schemes, *Journal of Computational Physics*, 197:502–531, 2004, DOI: 10.1016/j.jcp.2003.12.020.
- 21 73. Ghosh, D., Ghanem, R., and Horse, J.R., Analysis of eigenvalues and modal interaction of stochastic systems,
22 *AIAA Journal*, 43:2196–2201, 2005.
- 23 74. Wan, X. and Karniadakis, G., An adaptive multi-element generalized polynomial chaos method for stochastic
24 differential equations, *Journal of Computational Physics*, 209:617–642, 2005, DOI: 10.1016/j.jcp.2005.03.023.
- 25 75. Chantrasmi, T., Doostan, A., and Iaccarino, G., Padé-Legendre approximants for uncertainty analysis
26 with discontinuous response surfaces, *Journal of Computational Physics*, 228:7159–7180, 2009, DOI:
27 10.1016/j.jcp.2009.06.024.
- 28 76. Alexanderian, A., Le Maître, O., Najm, H., Iskandarani, M., and Knio, O., Multiscale stochastic preconditioners

- 1 in non-intrusive spectral projection, *Journal of Scientific Computing*, 50:306–340, 2012, DOI: 10.1007/s10915-
2 011-9486-2.
- 3 77. Espig, M., Hackbusch, W., Litvinenko, A., Matthies, H., and Wähnert, P., Efficient low-rank approximation of
4 the stochastic Galerkin matrix in tensor formats, *Computers and Mathematics with Applications*, 2012, DOI:
5 10.1016/j.camwa.2012.10.008.
- 6 78. Sacks, J., Schiller, S., and Welch, W., Designs for computer experiments, *Technometrics*, 31:41–47, 1989.
- 7 79. Nouy, A., A generalized spectral decomposition technique to solve a class of linear stochastic partial differ-
8 ential equations, *Computer Methods in Applied Mechanics and Engineering*, 196:4521–4537, 2007, DOI:
9 10.1016/j.cma.2007.05.016.
- 10 80. Bieri, M. and Schwab, C., Sparse high order FEM for elliptic sPDEs, *Computer Methods in Applied Mechanics*
11 *and Engineering*, 198:1146–1170, 2009, DOI: 10.1016/j.cma.2008.08.019.
- 12 81. Doostan, A. and Owhadi, H., A non-adapted sparse approximation of PDEs with stochastic inputs, *Journal of*
13 *Computational Physics*, 230:3015–3034, 2011, DOI: 10.1016/j.jcp.2011.01.002.
- 14 82. Jakeman, J., Archibald, R., and Xiu, D., Characterization of discontinuities in high-dimensional stochas-
15 tic problems on adaptive sparse grids, *Journal of Computational Physics*, 230:3977–3997, 2011, DOI:
16 10.1016/j.jcp.2011.02.022.
- 17 83. Bilonis, I. and Zabaras, N., Multidimensional adaptive relevance vector machines for uncertainty quantifica-
18 tion, *SIAM Journal on Scientific Computing*, 34:B881–B908, 2012, DOI: 10.1137/120861345.
- 19 84. Amsallem, D. and Farhat, C., Interpolation method for adapting reduced-order models and application to aeroe-
20 lasticity, *AIAA Journal*, 46:1803–1813, 2008, DOI: 10.2514/1.35374.
- 21 85. Degroote, J., Vierendeels, J., and Willcox, K., Interpolation among reduced-order matrices to obtain parameter-
22 ized models for design, optimization and probabilistic analysis, *International Journal for Numerical Methods*
23 *in Fluids*, 63:207–230, 2010, DOI: 10.1002/fld.2089.
- 24 86. Helton, J. and Davis, F. Sampling-based methods. In: Saltelli, A., Chan, K., and Scott, E. (Eds.), *Sensitivity*
25 *Analysis*, pp. 101–153. Wiley, West Sussex, United Kingdom, 2000.
- 26 87. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S.,
27 *Global Sensitivity Analysis: The Primer*, Wiley, West Sussex, United Kingdom, 2008.
- 28 88. Lin, G. and Karniadakis, G., Sensitivity analysis and stochastic simulations of non-equilibrium plasma flow,

- 1 *International Journal for Numerical Methods in Engineering*, 80:738–766, 2009, DOI: 10.1002/nme.2582.
- 2 89. Sobol, I., Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, *Math-*
3 *ematics and Computers in Simulation*, 55:271–280, 2001, DOI: 10.1016/S0378-4754(00)00270-6.
- 4 90. Oakley, J. and O’Hagan, A., Probabilistic sensitivity analysis of complex models: a Bayesian approach, *Journal*
5 *of the Royal Statistical Society. Series B*, 66:751–769, 2004, DOI: 10.1111/j.1467-9868.2004.05304.x.
- 6 91. Kucherenko, S., Tarantola, S., and Annoni, P., Estimation of global sensitivity indices for models with depen-
7 dent variables, *Computer Physics Communications*, 183:937–946, 2012, DOI: 10.1016/j.cpc.2011.12.020.
- 8 92. Rubinstein, R., Sensitivity analysis and performance extrapolation for computer models, *Operations Research*,
9 37:72–81, 1989, DOI: 10.1287/opre.37.1.72.
- 10 93. Pellissetti, M. and Ghanem, R., A method for the validation of predictive computations using a stochas-
11 tic approach, *ASME Journal of Offshore Mechanics and Arctic Engineering*, 126:227–234, 2004, DOI:
12 10.1115/OMAE2002-28071.
- 13 94. Reagan, M., Najm, H., Pébay, P., Knio, O., and Ghanem, R., Quantifying uncertainty in chemical systems
14 modeling, *International Journal of Chemical Kinetics*, 37:368–382, 2005, DOI: 10.1002/kin.20081.
- 15 95. Millwater, H., Universal properties of kernel functions for probabilistic sensitivity analysis, *Probabilistic Engi-*
16 *neering Mechanics*, 24:89–99, 2009, DOI: 10.1016/j.probengmech.2008.01.005.
- 17 96. Sudret, B., Global sensitivity analysis using polynomial chaos expansions, *Reliability Engineering and System*
18 *Safety*, 93:964–979, 2008, DOI: 10.1016/j.ress.2007.04.002.
- 19 97. Crestaux, T., Le Maître, O., and Martinez, J.-M., Polynomial chaos expansion for sensitivity analysis, *Reliabil-*
20 *ity Engineering and System Safety*, 94:1161–1172, 2009, DOI: 10.1016/j.ress.2008.10.008.
- 21 98. Eldred, M., Agarwal, H., Perez, V., Wojtkiewicz, S., and Renaud, J., Investigation of reliability
22 method formulations in DAKOTA/UQ, *Structure and Infrastructure Engineering*, 3:199–213, 2007, DOI:
23 10.1080/15732470500254618.
- 24 99. Cheung, S., Oliver, T., Prudencio, E., Prudhomme, S., and Moser, R., Bayesian uncertainty analysis with
25 applications to turbulence modeling, *Reliability Engineering and System Safety*, 96:1137–1149, 2011, DOI:
26 10.1016/j.ress.2010.09.013.
- 27 100. Heroux, M., Bartlett, R., Howle, V., Hoekstra, R., Hu, J., Kolda, T., Lehoucq, R., Long, K., Pawlowski,
28 R., Phipps, E., Salinger, A., Thornquist, H., Tuminaro, R., Willenbring, J., Williams, A., and Stanley, K.,

- 1 An overview of the Trilinos project, *ACM Transactions on Mathematical Software*, 31:397–423, 2005, DOI:
2 10.1145/1089014.1089021.
- 3 101. Debusschere, B., Najm, H., Pébayt, P., Knio, O., Ghanem, R., and Le Maître, O., Numerical challenges in
4 the use of polynomial chaos representations for stochastic processes, *SIAM Journal on Scientific Computing*,
5 26:698–719, 2005, DOI: 10.1137/S1064827503427741.
- 6 102. Papeleux, L. and Ponthot, J., Finite element simulation of springback in sheet metal forming, *Journal of Mate-*
7 *rials Processing Technology*, 125–126:785–791, 2002, DOI: 10.1016/S0924-0136(02)00393-X.
- 8 103. Moran, P., Statistical inference with bivariate gamma distributions, *Biometrika Trust*, 56:627–634, 1969, URL:
9 <http://www.jstor.org/stable/2334670>.
- 10 104. Das, S., Ghanem, R., and Spall, J., Asymptotic sampling distribution for polynomial chaos representation from
11 data: A maximum entropy and Fisher information approach, *SIAM Journal on Scientific Computing*, 30:2207–
12 2234, 2007, DOI: 10.1137/060652105.
- 13 105. Wan, X. and Karniadakis, G., Beyond Wiener-Askey expansions: handling arbitrary PDFs, *Journal of Scientific*
14 *Computing*, 27:455–464, 2006, DOI: 10.1007/s10915-005-9038-8.
- 15 106. Witteveen, J. and Bijl, H., Modeling arbitrary uncertainties using Gram-Schmidt polynomial chaos, In *44th*
16 *AIAA Aerospace Sciences Meeting and Exhibit*, Reno, Nevada, 2006.
- 17 107. Soize, C. and Desceliers, C., Computational aspects for constructing realizations of polynomial chaos in high
18 dimension, *SIAM Journal on Scientific Computing*, 32:2820–2831, 2010, DOI: 10.1137/100787830.
- 19 108. Arnst, M., Ghanem, R., Phipps, E., and Red-Horse, J., Measure transformation and efficient quadrature in
20 reduced-dimensional stochastic modeling of coupled problems, *International Journal for Numerical Methods*
21 *in Engineering*, 2012, In Press, DOI: 10.1002/nme.4368.
- 22 109. Perrin, G., Soize, C., and Funfschilling, C., Identification of polynomial chaos representations in high dimen-
23 sion from a set of realizations, *SIAM Journal on Scientific Computing*, 2012, accepted for Publication.
- 24 110. Moore, R., Algorithm AS 187: Derivatives of the incomplete gamma integral, *Journal of the Royal Statistical*
25 *Society. Series C (Applied Statistics)*, 31:330–335, 1982, URL: <http://www.jstor.org/stable/2348014>.

1 Tables

ℓ [-]	$(h_{\ell}^{\text{obs}}$ [MPa], s_{ℓ}^{obs} [MPa])	ℓ [-]	$(h_{\ell}^{\text{obs}}$ [MPa], s_{ℓ}^{obs} [MPa])	ℓ [-]	$(h_{\ell}^{\text{obs}}$ [MPa], s_{ℓ}^{obs} [MPa])
1	(1488, 375)	10	(1541, 340)	19	(1523, 402)
2	(1485, 403)	11	(1501, 366)	20	(1459, 402)
3	(1514, 407)	12	(1531, 403)	21	(1498, 388)
4	(1500, 377)	13	(1572, 393)	22	(1498, 435)
5	(1569, 348)	14	(1518, 388)	23	(1448, 418)
6	(1452, 384)	15	(1497, 416)	24	(1506, 394)
7	(1439, 393)	16	(1506, 421)	25	(1455, 393)
8	(1475, 399)	17	(1427, 381)		
9	(1530, 452)	18	(1456, 444)		

TABLE 1: Characterization of uncertainties: data.

1 **Figures**

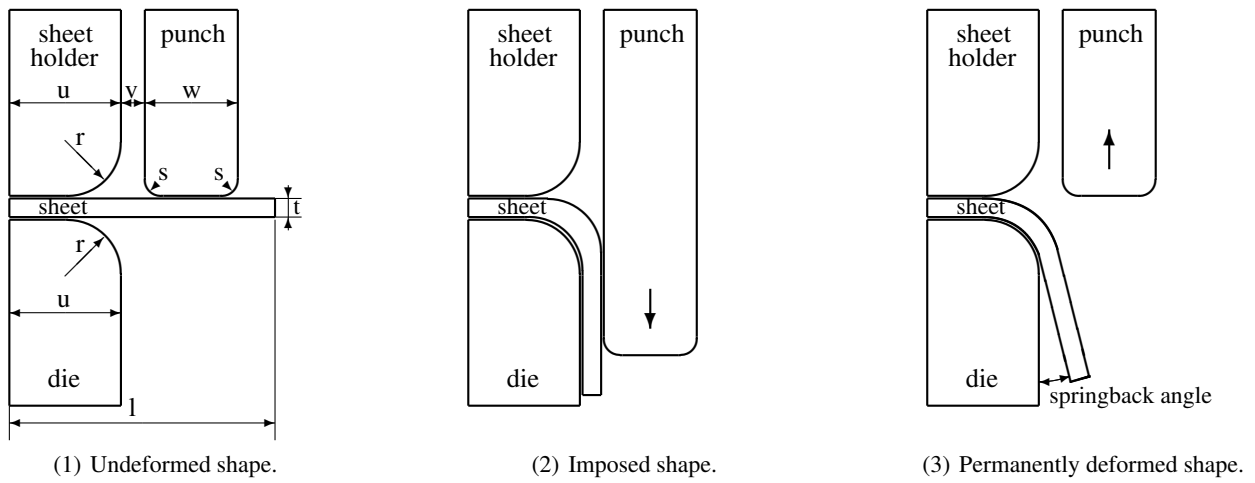
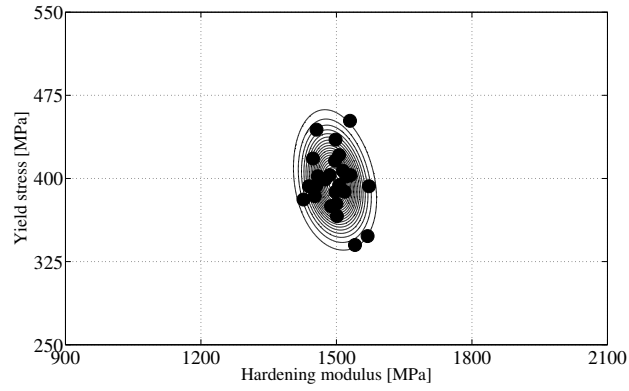
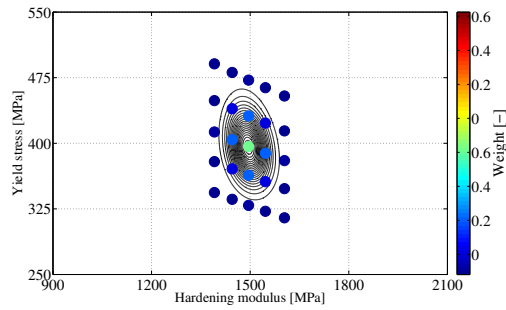


FIG. 1: Schematic representation of the problem. The left portion of the steel sheet is clamped.



$\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ (solid line) and $\{(h_\ell^{\text{obs}}, s_\ell^{\text{obs}}), 1 \leq \ell \leq n\}$ (dots).

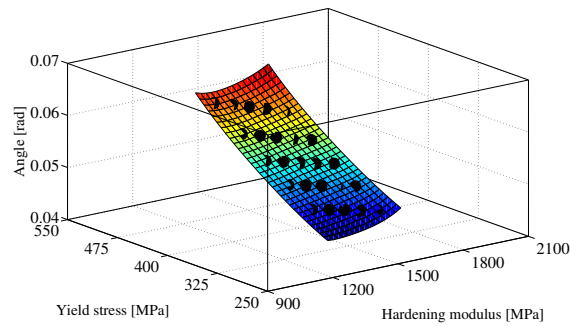
FIG. 2: Characterization of uncertainties: probability density function $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$.



(1) $\{(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda), w_{(k,\ell)}^\lambda\}, 1 \leq k \leq \lambda, 1 \leq \ell \leq \lambda\}$ (dots) and $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ (solid line).

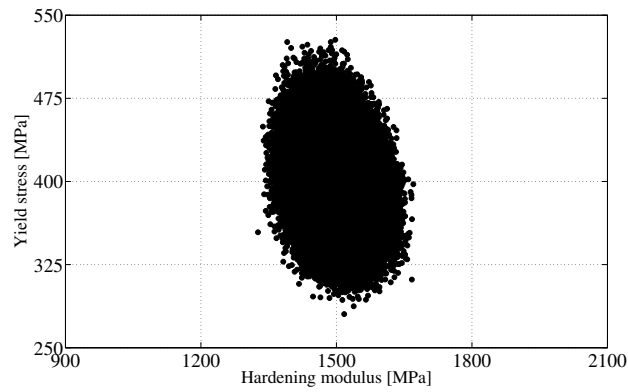


(2) Permanently deformed shapes obtained by solving the finite element model for each $(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda)$.

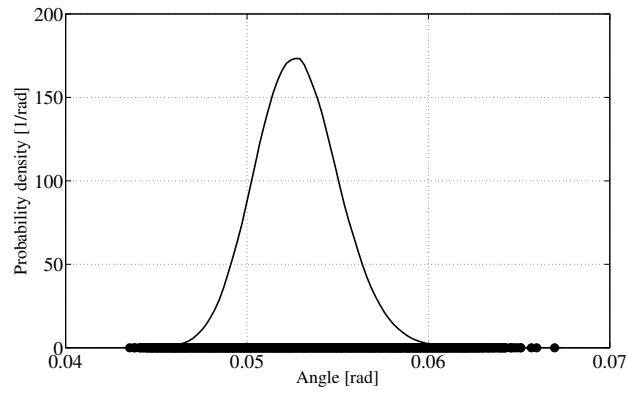


(3) $\{g(h_{(k,\ell)}^\lambda, s_{(k,\ell)}^\lambda), 1 \leq k \leq \lambda, 1 \leq \ell \leq \lambda\}$ (dots) and $g^{p,\lambda}$ (surface).

FIG. 3: Propagation of uncertainties: (1) quadrature rule for integration with respect to $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$, (2) permanently deformed shapes obtained by solving the finite element model for each node that this quadrature rule has, and (3) surrogate model for $p = 4$ and $\lambda = 5$.



(1) $\{(h_\ell, s_\ell), 1 \leq \ell \leq \nu\}$ (dots).



(2) $\{y_\ell^{p,\lambda} = g^{p,\lambda}(h_\ell, s_\ell), 1 \leq \ell \leq \nu\}$ (dots) and $\rho_Y^{p,\lambda}$ (solid line).

FIG. 4: Propagation of uncertainties: (1) ensemble of i.i.d. samples from $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ and (2) corresponding ensemble of i.i.d. samples and probability density function of the springback angle for $p = 4$ and $\lambda = 5$.

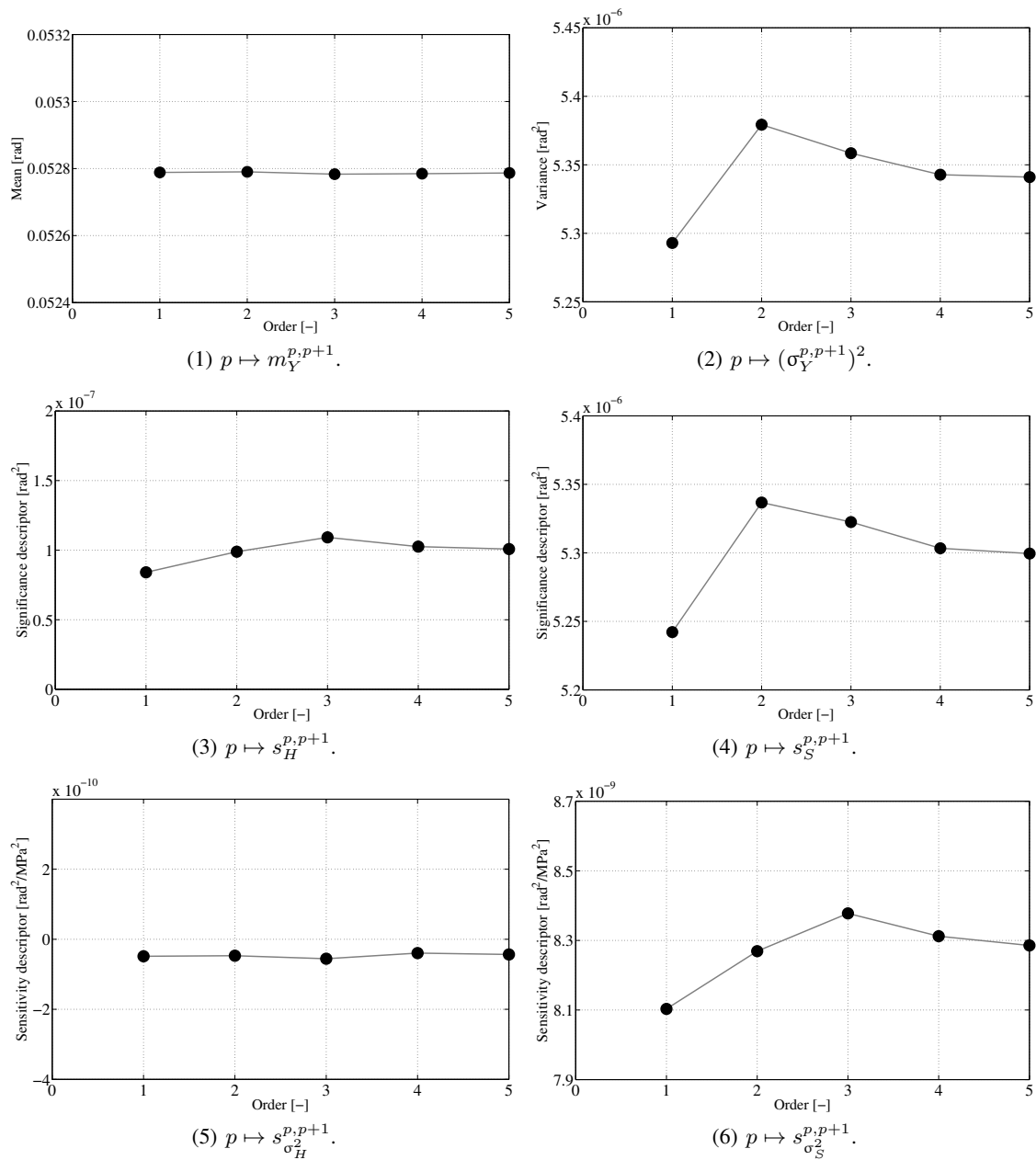


FIG. 5: Convergence analysis: convergence of statistical, significance, and sensitivity descriptors of the springback angle with respect to the order of the surrogate model and the number of nodes that the quadrature rule for integration with respect to $\rho_{(H,S)}(\cdot, \cdot; \hat{m}_H, \hat{\sigma}_H^2, \hat{m}_S, \hat{\sigma}_S^2, \hat{\rho})$ has.