



ELSEVIER

Contents lists available at SciVerse ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion

S. Oladyshkin*, W. Nowak

SRC Simulation Technology, Institute for Modelling Hydraulic and Environmental Systems (LH2), University of Stuttgart, Pfaffenwaldring 61, 70569 Stuttgart, Germany

ARTICLE INFO

Article history:

Received 21 January 2011

Received in revised form

15 May 2012

Accepted 19 May 2012

Available online 29 May 2012

Keywords:

Data-driven

Polynomial chaos

Arbitrary distribution

Orthonormal basis

Uncertainty

Random

Stochastic

ABSTRACT

We discuss the arbitrary polynomial chaos (aPC), which has been subject of research in a few recent theoretical papers. Like all polynomial chaos expansion techniques, aPC approximates the dependence of simulation model output on model parameters by expansion in an orthogonal polynomial basis. The aPC generalizes chaos expansion techniques towards arbitrary distributions with arbitrary probability measures, which can be either discrete, continuous, or discretized continuous and can be specified either analytically (as probability density/cumulative distribution functions), numerically as histogram or as raw data sets. We show that the aPC at finite expansion order only demands the existence of a finite number of moments and does not require the complete knowledge or even existence of a probability density function. This avoids the necessity to assign parametric probability distributions that are not sufficiently supported by limited available data. Alternatively, it allows modellers to choose freely of technical constraints the shapes of their statistical assumptions. Our key idea is to align the complexity level and order of analysis with the reliability and detail level of statistical information on the input parameters. We provide conditions for existence and clarify the relation of the aPC to statistical moments of model parameters. We test the performance of the aPC with diverse statistical distributions and with raw data. In these exemplary test cases, we illustrate the convergence with increasing expansion order and, for the first time, with increasing reliability level of statistical input information. Our results indicate that the aPC shows an exponential convergence rate and converges faster than classical polynomial chaos expansion techniques.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

The lack of information about the properties of physical systems, such as material parameters or boundary values, can lead to model uncertainties up to a level where the quantification of prediction uncertainties may become the dominant question in application tasks. Most physical processes appearing in nature are non-linear and, as a consequence, the required mathematical models are non-linear. Traditional and very well-known approaches for stochastic simulation are brute-force Monte Carlo simulation (e.g. [29]) and related approaches (e.g. latin hypercube sampling [15]). Unfortunately, for large and complex models, Monte Carlo techniques are inadequate. Even single deterministic simulations may require parallel high-performance computing. As a reasonably fast and attractive alternative, stochastic model reduction techniques based on the polynomial chaos expansion can be applied.

Polynomial chaos expansion. A large number of studies for diverse applications is based on the polynomial chaos expansion

(PCE) introduced by Wiener [52] in 1938. The chaos expansion offers an efficient high-order accurate way of including non-linear effects in stochastic analysis. PCE can be seen, intuitively, as a mathematically optimal way to construct and obtain a model response surface in the form of a high-dimensional polynomial in uncertain model parameters. The chances and limitations of polynomial chaos and related expansion techniques were discussed in [3]. The paper [43] showed how to use PCE for robust design under uncertainty with controlled failure probability. Recently, the sensitivity analysis based on PCE decomposition [5,9,34] has received increased attention. The papers [42,45] demonstrate correspondingly how classical PCE and its new aPC version can deliver the information required for global sensitivity analysis at low computational costs. Also, FORM and SORM methods (e.g. [17]) could be extended to higher-order accuracy via PCE, however, this has not yet been achieved.

The PCE technique can mainly be sub-divided into intrusive and non-intrusive approaches for the involved projection integral. The intrusive approach requires manipulation of the governing equations and can sometimes provide semi-analytical solutions for stochastic analysis. The best-known method from this group is the stochastic Galerkin technique, which originated from structural

* Corresponding author.

E-mail address: sergey.oladyshkin@iws.uni-stuttgart.de (S. Oladyshkin).

mechanics [20] and has been applied in studies for modeling uncertainties in flow problems [19,30,58]. However, because of the necessary symbolic manipulations, the procedure may become very complex and analytically cumbersome. For that reason, non-intrusive approaches like sparse quadrature [23] and the probabilistic collocation method [16,26] have been receiving increasing attention.

Polynomial chaos expansion for non-Gaussian distributions. The original PCE is based on Hermite polynomials, which are optimal for normally distributed random variables. Unfortunately, natural phenomena and uncertainty in engineering are often not that simple, and the distribution of physical or model parameters often cannot be considered Gaussian. However, it is possible to put into conformity a physical variable with a normal variable by an adequate transformation called Gaussian anamorphosis or normal score transformation (e.g. [48]) or approximate parametric transformations [11]. Using transformed variables for expansion cannot be considered an optimal choice because it leads to slow convergence of the expansion (e.g. [57,58]). In recent years, the PCE technique has been extended to the generalized polynomial chaos (gPC), based on the Askey scheme [2] of orthogonal polynomials by [57,58]. The gPC extends PCE towards a counted number of parametric statistical distributions (Gamma, Beta, Uniform, etc.). However, application tasks demand further adaptation of the chaos expansion technique to a larger spectrum of distributions. In [50,32], the authors presented a multi-element generalized polynomial chaos (ME-gPC) method. It is based on a decomposition of the random space into local elements, and subsequently implements gPC locally within the individual elements. An error control theory for the ME-gPC method was developed in [51] for elliptic problems. The ME-gPC is the first adaptive piecewise approach helping to deal with discontinuity of distributions or of model responses, and provides the desired adaptation to a wide spectrum of distributions. The ME-gPC conception offered in [50] provides a flexible tool for stochastic modeling, but interprets these data as an exactly known probability distribution, and considerably increases the computational effort for multidimensional stochastic problems.

Limited availability of data. The methods discussed above assume an exact knowledge of the involved probability density functions. Unfortunately, information about the distribution of data is very limited in realistic engineering applications, especially when environmental influences or natural phenomena are involved, or when predicting or engineering the environment (see also [46]). Applied research on (partially) natural or complex realistic systems often faces the problem of limited information about the model parameters and even about their probability distributions. For example, material properties of underground reservoirs are insufficiently available to provide a full picture of their distribution. Moreover, the statistical distribution of model parameters can be nontrivial, e.g., bounded, skewed, multi-modal, discontinuous, etc. Also, the dependence between several uncertain input parameters might be unknown, compare [25]. Depending on the modeling task and circumstances, statistical information on model parameters may be available either discrete, continuous, or discretized continuous, they could exist analytically as PDF/CDF or numerically as histogram. The key shortcoming of current PCE approaches in this context is twofold. First, they are heavily restricted in handling most of these conditions, and second they assume that this information is complete and perfect.

Small samples or data sets do not contain perfect or complete information on the probability distribution of model input parameters. For example, the study [46] demonstrated that limited information on input statistics introduces its own type of uncertainty in quantifying statistical model output distribution. Also, any attempt to construct probability density functions of any particular shape from samples of limited size or from sparse

information introduces additional subjectivity into the analysis, which bears the severe risk of leading to biased results. In a related application study [44], we illustrate that errors or additional (and mostly subjective) assumptions in data interpretation can severely bias uncertainty quantification and risk assessment, and hence could lead to failing designs. Methods of maximum entropy [18], closely related to known as the exponential polynomial method [13] in reliability engineering, and minimum relative entropy [56] are often used in the engineering sciences to construct a probability distribution from sparse information (mostly in the form of a few statistical moments and bounds) that may be available from different instances of the same object or from different objects with supposedly similar properties or conditions. Although these two methods are designed to minimize subjectivity and even though they can preserve the sample moments up to arbitrary order, they are heavily debated within the statistical community (e.g. [33]). In fact, they still introduce new assumptions and impose a specific assumption on distribution shape. The same is true for other typical methods to construct PDFs from moments in the field of reliability engineering, such as the Hermite polynomial transformation [53]. Such methods, however, are more subjective than entropy-based methods, since they cannot keep the original sample moments up to higher orders unchanged. If one still desires to fit a PDF as a pragmatic tool to filter raw data against noise, one should have full freedom in the chosen distribution shapes, not restricted by the technical constraints of PCE or gPC.

Approach and novelties. To overcome the first part of the problem, we claim that it is not even necessary to cast the available statistical information into probability density functions. Instead, the available information can directly and most purely be used in stochastic analysis, when using our data-driven formulation of PCE, see Section 2. We argue that applied tasks demand direct handling of arbitrary data distributions without additional assumptions for stochastic analysis. To overcome the second part of the problem, we suggest to perform a robustness analysis around the PCE (here: aPC) to assess the impact of incomplete statistical input information. Overall, we propose to align the complexity level and order of analysis with the reliability and detail level of statistical information on the input parameters.

The concept we propose in the current paper is to approach the problem in a highly parsimonious and yet fully data-driven description of randomness. We draw attention to the arbitrary polynomial chaos (aPC) that has recently been touched upon in a few theoretical papers. Two studies focusing on proofs of existence were published in the mathematical stochastic community [14,37]. Constructing the aPC polynomials by Gram–Schmidt orthogonalization was presented in the field of aerospace engineering [55,54]. These studies did not discuss the aPC in the light of data availability, limited reliability of data and assumptions in data interpretation.

The aPC extends chaos expansion techniques by employing a global polynomial basis for arbitrary distributions of data. In a certain sense, it allows to return back to a global basis with the new freedom of arbitrary polynomial chaos that the ME-gPC [50,51] uses only within piecewise local elements.

The most important property of the aPC that we will install and exploit in the current paper is that the aPC can work with probability measures that may be (if necessary) implicitly and incompletely defined via their moments only, and that it requires no additional information. In fact, our equations will show explicitly (in closed form) that statistical moments are the only source of information that is propagated in all polynomial expansion-based stochastic approaches. Thus, exact probability density functions do not have to be known and do not even have

to exist. For finite-order expansion, only a finite number of moments has to be known. This opens the path to data-driven applications, where data samples with limited size merely allow inference of a few statistical moments, but are not sufficient to support, without a substantial degree of subjectivity, the assumption of an exactly known probability measure (see discussion in Section 4.1). Fully in line with the demands of application tasks, the statistical data of modeling parameters can be specified either analytically (as probability density/cumulative distribution functions), numerically as histogram as a raw data sets.

In Section 2, we deliver the necessary mathematical material, and provide the necessary properties and proofs in Section 3. The convergence rate of the aPC will be illustrated in the context of an example problem in Section 4. How to address issues arising from the incomplete and inaccurate character of raw data sets used as statistical input information is discussed and illustrated in Section 5.

2. The arbitrary polynomial chaos expansion

2.1. One-dimensional aPC

We will consider a stochastic process in the probability space (Ω, A, Γ) with space of events Ω , σ -algebra A and probability measure Γ , see e.g. [22]. Let us consider a stochastic model $Y = f(\xi)$ with model input $\xi \in \Omega$ and model output Y . For a stochastic analysis of Y , the model $f(\xi)$ may be expanded as follows:

$$Y(\xi) \approx \sum_{i=1}^d c_i P^{(i)}(\xi), \tag{1}$$

where d is the order of expansion, c_i are the expansion coefficients that are determined by Galerkin projection, numerical integration or collocation, and $P^{(i)}(\xi)$ are the polynomials forming the basis $\{P^{(0)}, \dots, P^{(d)}\}$ that is orthogonal (or even orthonormal) with respect to the measure Γ (see Eq. (6)). The only difference between aPC and previous PCE methods is that the measure Γ can have an arbitrary form, and thus the basis $\{P^{(0)}, \dots, P^{(d)}\}$ has to be found specifically for the probability measure Γ appearing in the respective application.

This opens the path to data-driven applications of aPC. If a function $Y(\xi)$ is expanded in the orthonormal polynomial basis $\{P^{(0)}, \dots, P^{(d)}\}$, then characteristic statistical quantities of $Y(\xi)$ can be evaluated directly from the expansion coefficients c_i . For example, the mean and variance of $Y(\xi)$ are given by the following simple analytical relations:

$$\mu_Y = c_1, \quad \sigma_Y^2 = \sum_{i=2}^N c_i^2. \tag{2}$$

Notice that, in the current paper, we will focus on mono-dimensional stochastic input (i.e., only one uncertain parameter) for simplicity, but without loss of generality (see Section 2.2).

2.2. Multi-dimensional aPC

Most realistic applications feature multi-dimensional model input ξ , i.e. $\xi = \{\xi_1, \xi_2, \dots, \xi_N\}$. Here, the total number of input parameters is equal to N . The model parameters can be design or control parameters that can be chosen by the operator of a system, and uncertain parameters that describe our (incomplete) knowledge of the system properties. Hence, to investigate the influence of all input parameters ξ_1, ξ_2, \dots on the model output Y , the model output Y can be represented by a multivariate polynomial expansion as follows

$$Y(\xi_1, \xi_2, \dots, \xi_N) \approx \sum_{i=1}^M c_i \Phi_i(\xi_1, \xi_2, \dots, \xi_N). \tag{3}$$

Here, the coefficients c_i quantify the dependence of the model output Y on the input parameters $\xi_1, \xi_2, \dots, \xi_N$. The number M of terms in the expansion (3) depends on the total number of input parameters N and on the order d of the expansion, according to the combinatory formula $M = (N + d)! / (N! d!)$. The function Φ_i is a simplified notation of the multi-variate orthogonal polynomial basis for $\xi_1, \xi_2, \dots, \xi_N$. Assuming that the input parameters within $\xi_1, \xi_2, \dots, \xi_N$, are independent (e.g. [20]), the multi-dimensional basis can be constructed as a simple product of the corresponding univariate polynomials

$$\Phi_i(\xi_1, \xi_2, \dots, \xi_N) = \prod_{j=1}^N P_j^{(\alpha_j^i)}(\xi_1, \xi_2, \dots, \xi_N), \tag{4}$$

$$\sum_{j=1}^N \alpha_j^i \leq M, \quad i = 1, \dots, N,$$

where α_j^i is a multivariate index that contains the combinatoric information how to enumerate all possible products of individual univariate basis functions. In other words, the index α can be seen as $M \times N$ matrix, which contains the corresponding degree (e.g. 0, 1, 2, etc.) for parameter number j in expansion term k .

Let us mention that, in the current state of science for polynomial chaos expansions, the random variables have to be statistically independent or may be correlated in a linear fashion only. Linear correlation can be removed by adequate linear transformation, such as the KL-expansion [26], also called proper orthogonal decomposition [28] or principal component analysis [31] in other disciplines. Construction of a joint polynomial basis for statistically dependent random variables beyond linear dependence is a very important issue for future research.

2.3. Stochastic analysis based on PCE

Eqs. (1) and (3) can be interpreted as a model response surface for $Y = f(\xi_1, \xi_2, \dots, \xi_N)$, and represent the basic key element for: (I) uncertainty quantification; (II) robust design and (III) global sensitivity analysis.

(I) The simplest way to quantify uncertainty is via the analytical relations, see Eq. (2). However, in order to evaluate more complex statistical quantities, Monte Carlo simulation can be performed directly and immensely fast on the obtained polynomial given by Eq. (3) (see e.g. [44,49]). For Monte Carlo simulation in the absence of precise statistical information, Oladyshkin et al. [44] discuss and suggest the maximum entropy method for PDF estimation.

(II) Including design and control parameters together with uncertain parameters in expansion (3) provide an effective basis for robust design. The paper [43] showed how to use PCE for robust design under uncertainty with controlled failure probability.

(III) Expansion (3) also delivers the information required for global sensitivity analysis including simultaneous influences of different modeling parameters at low computational costs. For example, Sobol indices [35,36] or Weighted indices [45] can be computed directly from the coefficients α_i .

3. Moment-based analysis

Let us define the polynomial $P^{(k)}(\xi)$ of degree k in the random variable $\xi \in \Omega$:

$$P^{(k)}(\xi) = \sum_{i=0}^k p_i^{(k)} \xi^i, \quad k = \overline{0, d}, \tag{5}$$

where $p_i^{(k)}$ are coefficients in $P^{(k)}(\xi)$.

Our goal is to construct the polynomials in Eq. (5) such that they form an orthonormal basis for arbitrary distributions. The arbitrary distributions for the framework presented in our paper can be either discrete, continuous, discretized continuous, specified analytically, as histograms, raw data sets or by their moments. In this paper, we exploit this freedom and show how to treat any given probability distribution solely defined by the statistical moments of ξ . For limited-order expansion, this allows to work with arbitrary probability measures that are implicitly and incompletely defined by a limited number of moments only.

The goals of this section are to: (1) Derive a constructive rule to obtain the orthonormal basis which clarifies that only moments of ξ are important, (2) show that finite-order expansion only requires a finite number of moments, (3) provide conditions for the existence of an aPC even for sampled data.

3.1. Constructing the aPC from moments

Orthonormality for polynomials $P^{(k)}$ of degree k and $P^{(l)}$ of degree l is defined as

$$\int_{\xi \in \Omega} P^{(k)}(\xi)P^{(l)}(\xi) d\Gamma(\xi) = \delta_{kl} \quad \forall k, l = \overline{0, d}, \tag{6}$$

where δ_{kl} is the Kronecker delta. For the further development, we will make use of only the orthogonality condition

$$\int_{\xi \in \Omega} P^{(k)}(\xi)P^{(l)}(\xi) d\Gamma(\xi) = 0 \quad \forall k \neq l. \tag{7}$$

Instead of the normality condition, we will at first introduce an intermediate auxiliary condition by demanding that the leading coefficients of all polynomials be equal to 1

$$p_k^{(k)} = 1 \quad \forall k. \tag{8}$$

The following conditions apply for the orthogonal polynomial basis $\{P^{(k)}(\xi)\}$ ($k = \overline{0, d}$). First, for the zero-degree polynomial $P^{(0)}$, we obtain directly from (7) and (8) that $p_0^{(0)} = 1$, which also satisfies the normality condition (6). The orthogonality conditions for $P^{(1)}$ are as follows:

$$\int_{\xi \in \Omega} p_0^{(0)} \left[\sum_{i=0}^1 p_i^{(1)} \xi^i \right] d\Gamma(\xi) = 0, \tag{9}$$

$$p_1^{(1)} = 1.$$

This procedure can be continued for the construction of all following polynomials to obtain an orthogonal basis. The generalized conditions of orthogonality for any polynomial $P^{(k)}$ of degree k with all lower-order polynomials can be written in the following form:

$$\int_{\xi \in \Omega} p_0^{(0)} \left[\sum_{i=0}^k p_i^{(k)} \xi^i \right] d\Gamma(\xi) = 0,$$

$$\int_{\xi \in \Omega} \left[\sum_{i=0}^1 p_i^{(1)} \xi^i \right] \left[\sum_{i=0}^k p_i^{(k)} \xi^i \right] d\Gamma(\xi) = 0,$$

$$\vdots$$

$$\int_{\xi \in \Omega} \left[\sum_{i=0}^{k-1} p_i^{(k-1)} \xi^i \right] \left[\sum_{i=0}^k p_i^{(k)} \xi^i \right] d\Gamma(\xi) = 0,$$

$$p_k^{(k)} = 1. \tag{10}$$

The system of equations given by (10) is closed and defines the unknown polynomial coefficients $p_i^{(k)}$ ($i = \overline{0, k}$) of the required basis. Obviously, the above definition of the orthogonal polynomial of degree k uses the definition of all polynomials of lower degrees $0, \dots, k-1$. We will use that particular property to simplify the system in Eq. (10) by substituting the first equation into the second, the first and the second into the third, and so on. In addition, we will apply condition (8). Hence, without loss of generality, the system in Eq. (10) can be reduced to

$$\int_{\xi \in \Omega} \sum_{i=0}^k p_i^{(k)} \xi^i d\Gamma(\xi) = 0,$$

$$\int_{\xi \in \Omega} \sum_{i=0}^k p_i^{(k)} \xi^{i+1} d\Gamma(\xi) = 0,$$

$$\vdots$$

$$\int_{\xi \in \Omega} \sum_{i=0}^k p_i^{(k)} \xi^{i+k-1} d\Gamma(\xi) = 0,$$

$$p_k^{(k)} = 1. \tag{11}$$

Note that this rearrangement defines the k th orthogonal polynomial independent of all other polynomials from the orthogonal basis. The k th raw moment of the random variable ξ is defined as

$$\mu_k = \int_{\xi \in \Omega} \xi^k d\Gamma(\xi). \tag{12}$$

This allows to re-write Eq. (11) based on only the raw moments of ξ :

$$\sum_{i=0}^k p_i^{(k)} \mu_i = 0,$$

$$\sum_{i=0}^k p_i^{(k)} \mu_{i+1} = 0,$$

$$\vdots$$

$$\sum_{i=0}^k p_i^{(k)} \mu_{i+k-1} = 0,$$

$$p_k^{(k)} = 1. \tag{13}$$

Alternatively, the system of linear equations (13) can be written in the more convenient matrix form:

$$\begin{bmatrix} \mu_0 & \mu_1 & \dots & \mu_k \\ \mu_1 & \mu_2 & \dots & \mu_{k+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{k-1} & \mu_k & \dots & \mu_{2k-1} \\ 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} p_0^{(k)} \\ p_1^{(k)} \\ \vdots \\ p_{k-1}^{(k)} \\ p_k^{(k)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{14}$$

As a direct consequence, an orthogonal polynomial basis up to order d can be constructively defined for any arbitrary probability measure Γ under the following conditions: the coefficients $p_i^{(k)}$ can be constructed if and only if the square matrix of moments in the left-hand side of Eq. (14) is not singular. In the Appendix, we provide a proof for this under the condition that the number of support points in the distribution of ξ is greater than k and that all moments up to order $2k-1$ are finite. This holds for all continuous random variables, under the condition that $2k-1$ moments exist. If the moments of ξ are evaluated directly from a data set of limited size or from a discrete probability distribution featuring a finite number of possible outcomes, there need to be k or more

distinct values in the data set or distribution. All moments are always finite if no element of the data set is infinite.

From the Eqs. (7) to (14), it becomes evident that moments are the only required form of information on input distributions for constructing the basis and thus to operate the aPC. For finite-order expansion, a finite number of moments is sufficient. Hence, if a raw data set is the only form of available input information, computing its moments is sufficient, and estimating a full PDF from the data is not necessary. The same is true if only a limited number of moments are provided as input characterization. Also, arbitrary parametric distribution can be addressed, simply by working with their moments. This implies that any difference between distributions that becomes visible only in moments of order higher than $2d-1$ will be invisible to any order d polynomial expansion technique.

3.2. Explicit form of data-driven polynomial chaos

In this section, we present an analytical explicit form of the coefficients for moderate degrees of polynomials, which can be easily used for diverse data-driven application tasks, such as uncertainty quantification, global sensitivity analysis and probabilistic risk assessment. The coefficients for the higher degrees of polynomials can be obtained using the implicit scheme presented above (14), via recursive relations (see [1, Chapter 22]), via Gram-Schmidt orthogonalization (see [55,54]) or via the Stieltjes procedure [39].

To simplify the explicit form of coefficients, we will assume a normalized distribution of data with zero mean and unit variance after linear transformation

$$\xi' = \frac{(\xi - \mu)}{\sigma}, \tag{15}$$

which leads to a centralization and standardization of all moments.

Thus, the orthogonal polynomial basis $\{P^{(k)}(\xi)\}$ ($k = \overline{0,d}$) can be presented as

$$P^{(d)}(\xi) = \sum_{i=0}^d p_i^{(d)} \left(\frac{\xi - \mu}{\sigma} \right)^i, \tag{16}$$

where $p_i^{(d)}$ are the coefficients of polynomial $P^{(d)}$ defined explicitly through the raw moments of ξ' from the relations below. Due to Eq. (15), the raw moments of ξ' are related to the central moments $\hat{\mu}_k(\xi)$ of ξ via:

$$\hat{\mu}_k(\xi') = \hat{\mu}_k(\xi) \hat{\mu}_2(\xi)^{-k/2}.$$

Coefficients for polynomial of 0 degree

$$p_0^{(0)} = 1. \tag{17}$$

Coefficients for polynomial of 1st degree

$$p_0^{(1)} = 0, \quad p_1^{(1)} = 1. \tag{18}$$

Coefficients for polynomial of 2nd degree

$$p_0^{(2)} = -1, \quad p_1^{(2)} = -\mu_3, \quad p_2^{(2)} = 1. \tag{19}$$

Coefficients for polynomial of 3rd degree

$$p_0^{(3)} = \mu_3^2 - \mu_3^3 + \mu_3\mu_4 - \mu_5, \quad p_1^{(3)} = -\mu_3\mu_5 + \mu_3^2 - \mu_4 + \mu_3\mu_4,$$

$$p_2^{(3)} = -\mu_3\mu_4 + \mu_5 - \mu_3,$$

$$p_3^{(3)} = 1 - \mu_3 + \mu_3^2, \tag{20}$$

Coefficients for polynomial of 4th degree

$$p_0^{(4)} = \mu_3^2\mu_5\mu_4 + \mu_3^3\mu_7 - \mu_3^2\mu_5^2 - 2\mu_3^2\mu_4\mu_6 + 2\mu_4^2\mu_5 - \mu_4^3,$$

$$p_1^{(4)} = -\mu_3^3 + \mu_4^2\mu_3^2 - \mu_3^2\mu_4\mu_5 + \mu_4\mu_5\mu_6 + \mu_3\mu_5\mu_6,$$

$$+ \mu_3\mu_5\mu_7 - \mu_3\mu_4\mu_7 - \mu_3^3\mu_5 + \mu_3^3\mu_6 - \mu_3\mu_6^2,$$

$$p_2^{(4)} = -\mu_4^2\mu_3^2 + \mu_3\mu_4\mu_7 + \mu_4\mu_5^2 - \mu_4^2\mu_6 + \mu_3^3\mu_5 - \mu_3\mu_5\mu_6,$$

$$p_3^{(4)} = \mu_3^3\mu_4 - \mu_5\mu_4 - \mu_3^2\mu_7 + \mu_5^2 + \mu_4\mu_6,$$

$$p_4^{(4)} = -\mu_3^4 + \mu_3^2\mu_6 - 2\mu_5\mu_4 + \mu_4^2. \tag{21}$$

3.3. Normalization

The above orthogonal polynomial basis can be used directly for analysis. However, an orthonormal basis has more useful properties (see Eq. (2) and Section 3.4). Thus, the next step is to normalize the orthogonal basis. We will use the norm for the polynomial P^k introduced in Eq. (6):

$$\|P^{(k)}\|^2 = \int_{\xi \in \Omega} [P^{(k)}(\xi)]^2 d\Gamma(\xi). \tag{22}$$

Hence, a valid orthonormal polynomial basis $\{\Psi^{(k)}(\xi)\}$ ($k = \overline{0,d}$) is:

$$\Psi^{(k)}(\xi) = \frac{1}{\|P^{(k)}\|} \sum_{i=0}^k p_i^{(k)} \xi^i. \tag{23}$$

For normalization, the evaluation of $\|P^{(k)}\|$ for $k=d$ additionally requires finiteness and availability of the $2d$ th moment.

3.4. Summarized properties of the orthonormal basis

As consequence of the derivations in Sections 3.1 and 3.3, the polynomial basis in Eq. (23) has the following properties:

Property I. The orthonormal basis can be constructed without any hierarchical conditions or recurrence relations that are used in [1, Chapter 22] and in [54,55].

Property II. Existence of the moments μ_0, \dots, μ_{2d} is the necessary and sufficient condition for constructing an orthonormal basis $\{\Psi^{(0)}, \dots, \Psi^{(d)}\}$ up to degree d , together with the condition that the number of supports points of ξ is greater than d if ξ is a discrete variable or is represented by a data set.

Property III. The orthonormal polynomial basis for arbitrary probability measures is based on the corresponding moments only, and does not require the knowledge (or even existence) of a probability density function.

Property IV. All the zeros of the orthogonal polynomials are real, simple and located in the interior of the interval of orthogonality [1]. This property is useful for numerical integration, especially for bounded distributions.

Property V. As particular cases, the Hermite, Laguerre, Jacoby polynomials, etc. from the Askey scheme and the polynomials for log-normal variables by Ernst et al. [14] can be reconstructed within a multiplicative constant.

Property VI. All distributions that share the same moments up to order $2d$ will also share the same basis, and thus will lead to identical results in an expansion up to order d .

4. Data-driven modeling

The arbitrary polynomial chaos expansion presented in Sections 2 and 3 provides a simple and efficient tool for analysing stochastic systems. We will consider a very simple model in order to focus all attention on our data-driven concept, which is based directly on the moments of sampled data without intermediate

steps of data reinterpretation. This avoids the subjectivity usually introduced when choosing among a small limited number of theoretical distributions to represent a natural phenomenon, and so avoids the problems of subjectivity under limited data availability is discussed in Section 1. These problems will be illustrated in Section 4.1. An application to a problem with a realistic level of complexity and a detailed discussion of expert's subjectivity in uncertainty analysis is presented in [44]. That paper demonstrates how subjectivity of interpreting limited data sets can easily lead to substantial prediction bias, and that the subjective choice of distribution shapes has a similar relevance as uncertainties due to physical conceptualization, numerical codes and parameter uncertainty.

Here, for simplicity, we consider the exponential decay differential equation which was already used in [58] to illustrate the Askey scheme:

$$\frac{dY(t)}{dt} = -\xi Y, \quad Y(0) = 1. \quad (24)$$

Let Y_{PC} be the solution obtained using the polynomial chaos expansion (1) for the problem defined in Eq. (24). We use a Monte Carlo simulation as reference solution and define the time dependent relative error $\epsilon(t)$ between the polynomial chaos expansion solution $Y_{PC}(t)$ and the Monte Carlo solution $Y_{MC}(t)$ as

$$\epsilon(t) = \frac{|Y_{PC}(t) - Y_{MC}(t)|}{|Y_{MC}(t)|}. \quad (25)$$

4.1. Fidelity of data-driven interpretation

To illustrate the fidelity of the data-driven chaos expansion, we will consider a synthetic example for an empirical data distribution, see Fig. 1, and apply both aPC and the classical PCE for comparison. The illustrative data set (with sample size $N=500$) presented in the left plot of Fig. 1 was generated as superposition of normal and log-normal distributions and contains statistical noise introduced due to the small size of the data sample.

The classical approach would be to introduce a parametric probability distribution, e.g., with fitted mean and variance or with maximum likelihood parameters. Here, for illustration, we select the Normal, Lognormal and Gamma distribution, see the right plot of Fig. 1. Evidently, the list of possible candidate distributions for fitting to the considered data can be very long. Introducing a full probability density function (PDF) resembles a strong assumption on all higher moments up to infinite order, and

claims to know the exact shape, e.g., also of the extreme value tails. Such assumptions on the alleged shape of the underlying probability density function, unfortunately, can lead to substantial errors in data interpretation. The data-driven approach strongly alleviates this situation, because it can directly handle a set of moments (e.g., the mean, variance, skewness, kurtosis, and so forth), without any further assumptions on higher-order moments and without having to introduce a PDF at all. This also provides the freedom to work with only a small number of moments obtained via expert elicitation, without asking for a full PDF.

We will apply different orders of the polynomial chaos expansion (1) to the test problem (24) using two sources of input information about the data distribution: (1) the three introduced assumptions on PDFs (right plot of Fig. 1) and (2) the pure raw data sample (left plot of Fig. 1). To observe the pure impact of data interpretation regardless of numerical techniques, we treat all four cases with the aPC, i.e., we construct an optimal orthonormal polynomial base (see Section 2) for each case. This avoids the non-linear transformations that are usually necessary to map the assumed PDFs onto the normal PDF. Such transformations would introduce additional errors, as discussed in Sections 4.2 and 4.3.

Technically, the coefficients c_i in the chaos expansion can be obtained, e.g., by Galerkin projection (e.g. [19,30,58]) or by the Collocation method (e.g. [16,26,60]). Both methods lead to the same result when using the optimal polynomial basis in the case of univariate analysis. Fig. 2 illustrates the convergence of the mean and variance (at time $t=1$) for the assumed PDFs (left plot of Fig. 1) and for the pure raw data (right plot of Fig. 1).

All considered cases reproduce an acceptable approximation with the linear expansion. Increasing the expansion order shows strong convergence for the data-driven polynomial chaos expansion. However, increasing the order does not assure convergence for the expansions based on interpreted data. The problem does not lie in poor numerical properties when treating these distributions, but in accurate convergence to a wrong value. This error is introduced only by fitting parametric PDFs to the data instead of letting the data fully manifest themselves. In fact, the PDF-based analysis are good only up to first order, because only moments up to the second are represented accurately by the matched PDFs. This means that all effort spent for higher-order expansion (especially for complex problems) is invested to negligible improvement, if not matched with an adequate effort for accurate data interpretation. The key advantage of aPC in this respect is that (1) it allows full freedom in the used type of input information, and (2) our analysis in Section 3 makes explicitly clear what amount of information (i.e., the moments up to a certain number)

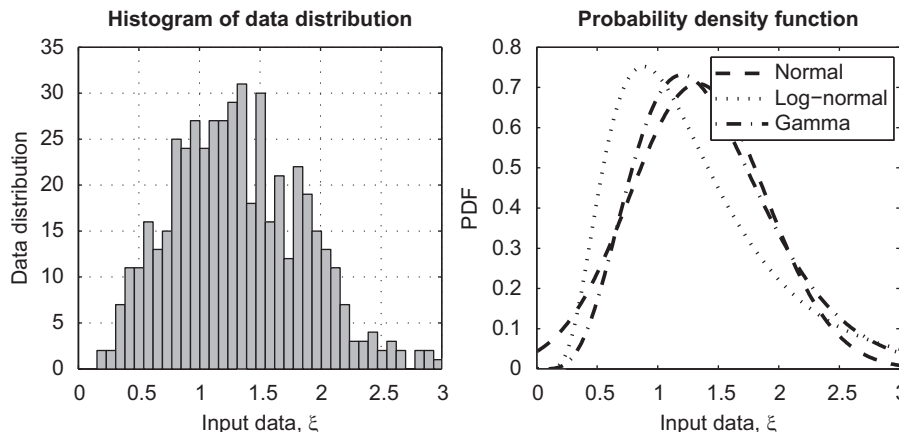


Fig. 1. Data distribution (left plot) and assumed stochastic distribution: Normal, Lognormal and Gamma (right plot).

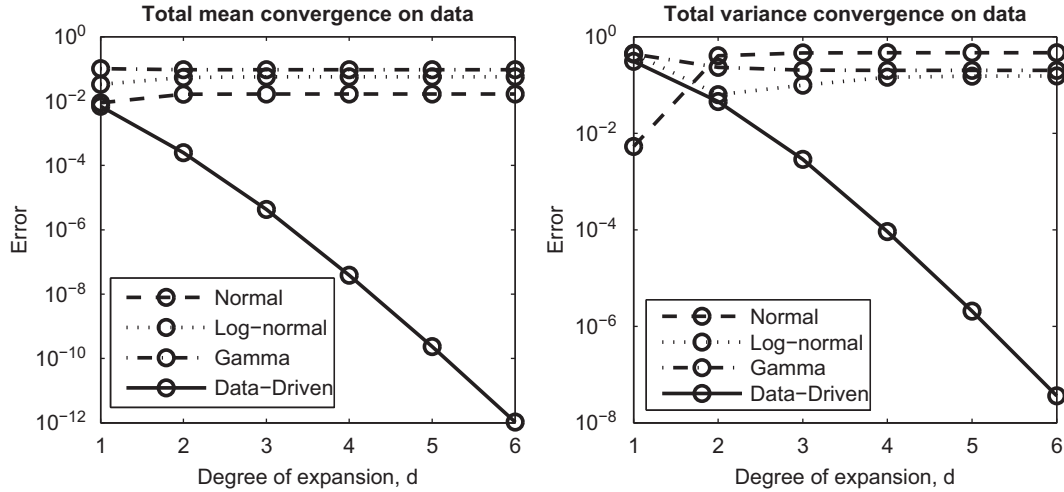


Fig. 2. Convergence on data: estimation of mean (left plot) and variance (right plot).

enters the analysis at what expansion order. This allows to align the complexity and order of analysis with the reliability and detail level of statistical information on the input parameters.

4.2. Evidence of improved convergence

In this section we will illustrate the efficiency of analysis within an optimal (data-driven) polynomial basis. We wish to show the improved convergence rate of the arbitrary polynomial chaos compared to the classical PCE technique. The classical PC requires non-linear transformations to map non-normal input data distributions onto the normal PDF. In this technical aspect, the classical PCE does not differ from the gPC, which requires transformation onto one of possible PDF from Askey scheme.

Eq. (24) can be expanded (see Eq. (1)) in the orthogonal polynomial basis $\Psi_i(\xi)$. The projection coefficients are defined as

$$c_i(t) = \int_{\xi \in \Omega} Y(t)\Psi_i(\xi) d\Gamma(\xi), \quad i = \overline{0, d}. \tag{26}$$

We will apply both the aPC and the classical PCE (including the mapping onto the normal distribution) to the example (24). For these two expansions techniques, we will apply both Galerkin projection (intrusive) and Gauss quadrature (non-intrusive) to evaluate the integral in Eq. (26). In both intrusive and non-intrusive approaches, the resulting values $c_i(t)$ from Eq. (26) will change with the distribution of the random variable ξ . For the optimal basis (used in aPC), however, the results from Galerkin projection and numerical integration coincide. This yields 3 distinguishable techniques. For all three techniques, we analyse the performance for diverse exemplary distributions of the univariate input ξ (Rayleigh, Weibull, Log-normal). Detailed descriptions of these distributions can be found in [38]. For the classical PCE, the random variable ξ is not distributed in the same space as the polynomial basis $\Psi_i(\xi)$, and an additional conversion is required. Thus we map the model variable ξ onto a corresponding normal variable ξ_N by Gaussian anamorphosis or normal score transform [48]. Fig. 3 illustrates the convergence of the mean and variance of Y (at time $t=1$) for our three exemplary distributions of the model input ξ . As previously demonstrated for the gPC [57], expansion in the optimal polynomial basis without transformation shows at least an exponential convergence. Convergence with a non-optimal basis (here: Hermite) after transformation strongly depends on the nonlinearity of the required transformation from ξ to ξ_N .

4.3. Clarification of error types

In Section 4.1, we demonstrated the possible errors introduced by subjective data interpretation when fitting parametric PDFs to raw data. In that analysis, we deliberately excluded the error by the different numerical accuracies of aPC and classical PCE. Now, we will demonstrate the faster numerical convergence of aPC compared to classical PCE with transformation, but this time excluding the error of data interpretation. We classify the cause of error into two types: I—transformation expansion error and II—numerical integration error. Modeling within the non-optimal basis using Galerkin projection leads to errors by expanding and truncating the transformation, which we denote here as transformation expansion error (type I). Modeling within the non-optimal basis using Gauss quadrature entails numerical integration error (type II). Using the optimal basis provides identical results for both intrusive and non-intrusive methods, because numerical integration is exact when using the roots of the $d+1$ order polynomial from the optimal basis, and because no transformation from ξ to ξ_N is necessary.

Transformation expansion error (type I). For intrusive manipulation, the anamorphosis transformation from ξ to ξ_N has to be expanded in ξ_N . The finite number d of terms in this expansion causes the first type of error. The difference between expansion in an optimal basis and expansion of $Y(\xi)$ after transformation to ξ_N is the so-called “aliasing error” (see [59]). Fig. 4 illustrates the nature of the this type of error. Expansion of the transformation (here $\xi = \exp(\xi_N)$) at different orders is shown in the left plot of Fig. 4. The right plot of Fig. 4 demonstrates the corresponding mapping of a normal probability density function (PDF) back to physical space using the expanded and truncated expansion. In these examples, the normal PDF should transform to a log-normal PDF. The strong nonlinearity of the logarithmic transformation leads to a poor approximation with a finite number of terms. Thus, the choice of a non-optimal polynomial basis for the model input ξ leads to a wrong representation of the probability measure Γ_ξ . This leads to the erroneous analysis of model output $Y(\xi)$ visible for Galerkin-based computations in Fig. 3.

Numerical integration error (type II). The accuracy of numerical integration (especially sparse) strongly depends on the choice of integration points. For example, in Gauss–Hermite integration, the polynomial basis defines the positions ξ_i of integration points in the space of the input variable by the roots of the polynomial of degree $d+1$. Thus, using a non-optimal polynomial basis provides a non-optimal choice of the integration points, which causes the

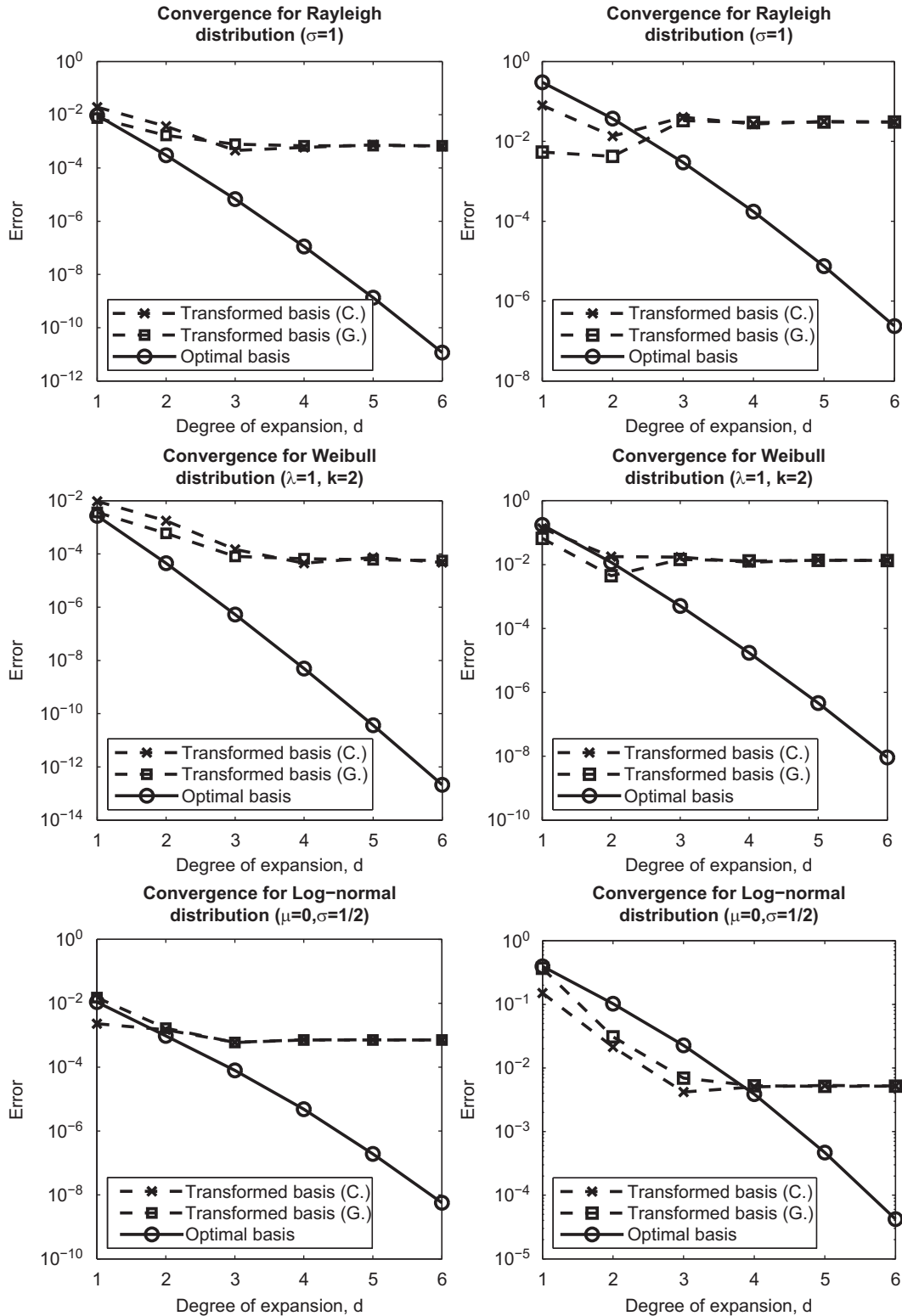


Fig. 3. Convergence of mean (left) and variance (right) estimation based on optimal basis and transformed basis using Galerkin projection (G.) and numerical integration (C.). For the optimal basis (used in aPC), the results from Galerkin projection and numerical integration coincide.

second type of error. To illustrate this type of error, let us consider a stochastic model with a random variable ζ that follows a non-Gaussian distribution. The selected model is an extremely simple non-linear one

$$Y(\zeta_{ph}) = \zeta^6. \tag{27}$$

In our example, the input parameter ζ is distributed according to the Chi-square distribution. We will construct two expansions: one based on Hermite polynomials with adequate Gaussian anamorphosis, and one based on optimal polynomials for the Chi-square distribution of the input data. In both cases, we employ Gauss quadrature and compare the results to a reference

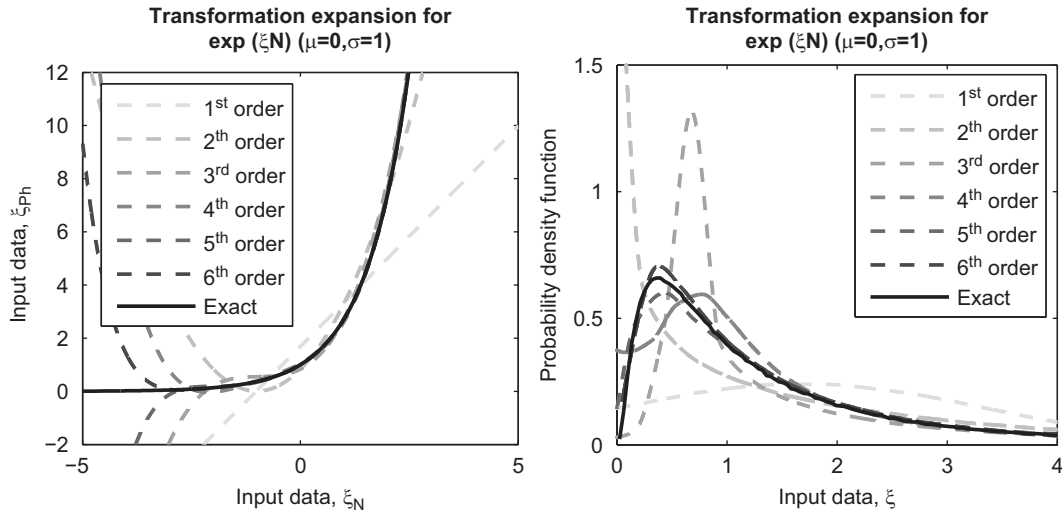


Fig. 4. Transformation expansion error: expansion series (left plot) and corresponding distribution (right plot).

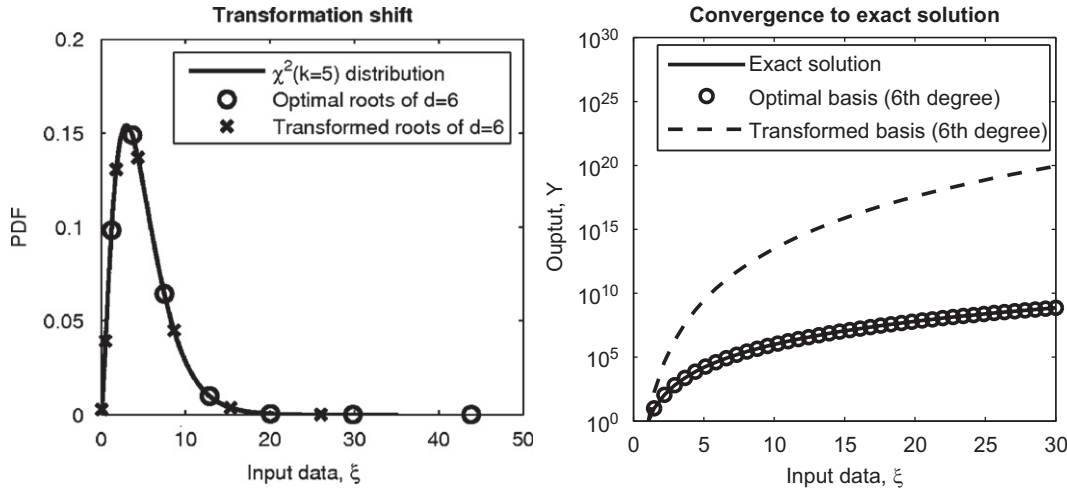


Fig. 5. Numerical integration error: transformation shift (left plot) and convergence to exact analytical solution (right plot).

solution. The supposedly optimal location of the integration points for Hermite polynomials correspond to the roots of the Hermite polynomial of order $d + 1$, back-transformed from ξ_N to ξ by anamorphosis. However, the truly optimal distribution of the integration points are the roots of the optimal polynomials that are orthogonal for ξ without further transformation. The obvious difference is shown in the left plot of Fig. 5. The transformed points are shifted against the optimal ones and thus cannot be considered as an optimal choice for numerical integration. Therefore, strong nonlinearity in the transformation leads to significant errors in PCE techniques that derive their numerical integration rules from the involved basis. Evidently, the example in Eq. (27) has an analytical solution, which can be reproduced by the expansion of 6th order within the optimal basis (see the right plot in Fig. 5). However, the transformed Hermite chaos combined with non-optimal Gauss quadrature does not converge to the known analytical solution even for the expansion degree $d=6$ that should be, in theory, fully accurate construct $Y = \xi^6$.

5. Robustness analysis for inaccurate input data

The presented approach can handle different forms of input information. In particular, it can directly handle raw data, which

can be useful for practical applications. However, when the input data set is small, the sample moments are only uncertain estimates of real moments. Hence, a direct application of the method presented becomes less robust. In that case, it would be useful to apply some standard methods to assess the robustness in the estimation of moments, such as Jackknife or Bootstrapping (e.g. [12]). In the field of reliability engineering, Bootstrap methods have been applied to construct upper confidence limits for unreliability in [10]. Bootstrap-based confidence intervals caused by the uncertainty representing computationally demanding models by meta-models has been investigated in the paper [41] via regression based sensitivity analysis. The recent paper [4] explores sparse and partially random integration techniques for PCE, applied to sensitivity analysis, and provides Monte Carlo based estimates for the error introduced by the random character of the used integration rules. In this section, we focus on the robustness of data-driven expansions with respect to the limited size of a raw data set, that represents the underlying probability distributions of model input only inaccurately.

For that, we repeatedly ($N=1000$) generated raw data according to the assumed underlying theoretical distribution. Each time, we constructed a new data-driven basis and performed a projection of the model output Y (Eq. (24)) to the corresponding data-driven polynomial basis and computed the mean μ_Y and variance

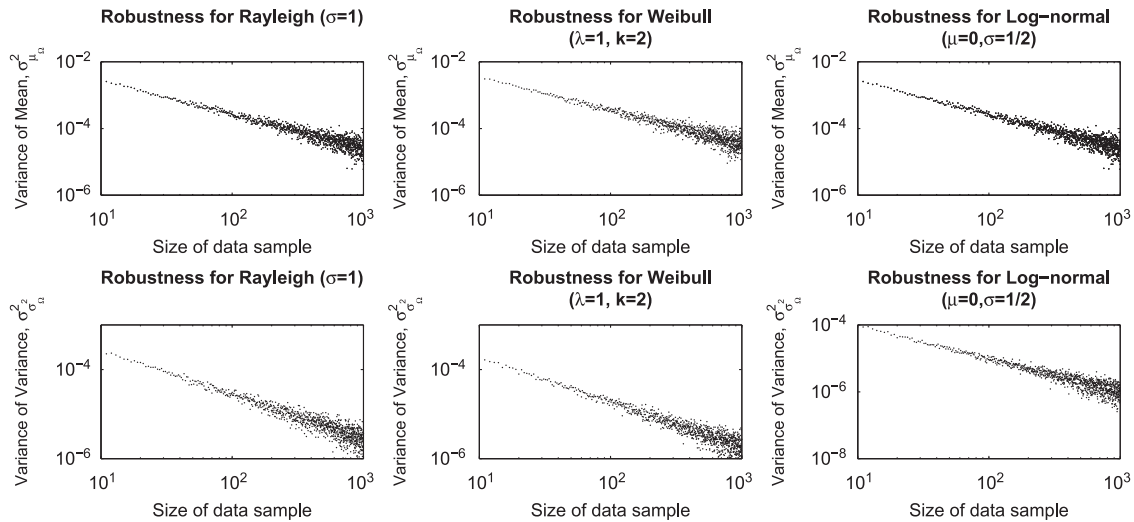


Fig. 6. Robustness of data-driven expansion with respect to the size of a raw data sample: variance of the mean (top) and variance of the variance (bottom).

σ_Y^2 of the model output $Y(t=1)$ in each repetition. From this, we computed the variance of the mean σ_{μ}^2 and the variance of the variance σ_{σ}^2 . This entire nested Monte Carlo analysis was repeated for sizes of the raw data set ranging from $N=20$ to 1000. Fig. 6 shows the results for the distributions considered in Section 4.2. For this illustration, we used a 3rd degree of expansion. Other degrees of expansion (1–6) show similar results, all of them having error variances inversely proportional to the size of the sampled data set ($1/N$), i.e. having error standard deviations proportional to $1/\sqrt{N}$. This rate is visible as the slope of the scatter plots in Fig. 6.

The scatter is caused by the finite number of Monte Carlo repetitions used in the error estimation. It corresponds to the Monte Carlo error of the error estimate. Only visually, the scatter increases with increasing size N of the data sets due to the logarithmic scale of the ordinate. The important aspect of the plots in the Fig. 6, however, is not the degree of scatter (i.e. the uncertainty of the error estimation), but the average slope (i.e. the error estimate itself).

Apparently, the data-driven chaos expansion has a convergence rate proportional to $1/\sqrt{N}$ for the standard deviation and confidence intervals of computed model output statistics. This convergence rate is well known for the variance of sample statics and from Monte Carlo techniques in general [6]. This means that the aPC does not modify the robustness and convergence properties with respect to insufficient sample size in comparison to moments from classical sample statics or Monte Carlo simulation.

The analysis we performed here cannot be done with computationally expensive models, or if a single real data set is all that is available. In that case, one can perform a Jackknife or Bootstrapping method to estimate the sampling distribution of the used moments from that data set, and propagate the resulting randomized moments through the response surface obtained with a constant set of integration points. This will estimate the variances of the PCE solution due to the limited size of the data set, and corresponded to the value for a single given N in Fig. 6.

Especially in such cases with very small data sets, expert opinion can be very useful to filter the data set, remove alleged outliers, fit a simple or complex PDF, and so forth. In our proposed approach, an expert will have total freedom of data interpretation (not restricted to the selection among standard PDFs) and can provide much more sophisticated information (e.g. lower and higher moments, complex and even non-parametric distributions, etc.). According to our approach, expert opinion (in a most general

sense) will be incorporated directly without any additional transformation or additional subjectivity when translating it to the stochastic numerical framework. The presented methods allow experts to choose freely of technical constraints the shapes of their statistical assumptions.

6. Remaining issues for future research

The polynomial basis for continuous or discrete random variables can be constructed if and only if the number of support points (distinct values) in the distribution (within the available data set) is greater than the desired degree of the basis (see Property II in Section 3.4). However, for discrete cases we cannot guarantee that the integration points (see Section 4) will be distributed only within the space of a random variable. Still, it would be possible for each integration point to find a neighbouring point that belongs to the discrete space, but the convergence is not guaranteed, and this remains an open question for future research.

Formal knowledge about the convergence of the polynomial basis for diverse random spaces can be very useful for practical needs, such as the convergence for the Hermite basis in normal space [7]. Karl Weierstrass established his approximation theorem in 1885, which states that every continuous function defined on an interval can be uniformly approximated as closely as desired by a polynomial function. A generalization of the Weierstrass theorem was proposed in the Stone-Weierstrass theorem where, instead of the interval, an arbitrary compact Hausdorff space is considered and, instead of the algebra of polynomial functions, approximation with elements from more general sub-algebras were investigated [40]. Thus, if the random space is an arbitrary compact Hausdorff space, uniform convergence is guaranteed and the polynomial space is dense in the arbitrary compact Hausdorff space. However, it is not assured that any polynomial expansion (including the optimal orthonormal basis) will uniformly converge in any random space and the definition of the random space will define convergence in that space.

The classical theorem of [47] characterizes the problem of polynomial density by the unique solvability of a moment problem, which means that the distribution function is required to be uniquely defined by the sequence of its moments. Ernst et al. [14] discussed such aspects for the generalized chaos expansion (gPC), and showed that the moment problem is not

uniquely solvable for the lognormal distribution. However, the mentioned works demand existence and precise knowledge about the probability density function, which is neither required nor desired for the arbitrary polynomial chaos (aPC). We observed convergence of the aPC in Section 4 for a counted number of useful cases, however research on a formal proof of convergence is a remaining question for future research.

As an outlook for future development, we point out the construction of a joint basis for parameters that have a complex statistical dependence beyond correlation. Following that direction, the aPC could be the first PCE family member that will allow to handle non-linear statistical dependence between input variables.

7. Summary and conclusions

In the current paper, we presented the arbitrary polynomial chaos expansion (aPC). The aPC conception provides a constructive and simple tool for uncertainty quantification, global sensitivity analysis and robust design. It offers a new data-driven approach for stochastic analysis that avoids the subjectivity of assigning parametric probability distributions that are not sufficiently supported by available data. We show that a global orthonormal polynomial basis for finite-order expansion demands the existence of a finite number of moments only, and does not require exact knowledge or even existence of a probability density function. Thus, the aPC can be constructed for arbitrary parametric and non-parametric distributions of data, even if the statistical model output characterization of input data is incomplete.

Also, the orthonormal basis can be constructed without using any hierarchical conditions or recurrence relations with polynomials of lower-order. For discrete random variables, the aPC can be constructed if and only if the number of discrete values of the random variable is greater than the largest considered degree of the basis. In case of continuous random variables, the aPC can be constructed from a number of moments which equals to two times the degree of the basis. If desired, the method can work directly with raw sampled data sets to represent the uncertainty and possible variational ranges of input data. The presented methods allow experts to choose freely of technical constraints the shapes of their statistical assumptions and makes explicitly clear what amount of information (i.e., the moments up to a certain order) enters the analysis at what expansion order. Overall, this allows to align the complexity and order of analysis with the reliability and detail level of statistical information on the input parameters.

We also provided numerical studies for diverse exemplary distributions, where we illustrated convergence rates for optimal and non-optimal polynomial bases using intrusive and non-intrusive methods. This analysis strictly illustrated that using a non-optimal polynomial basis provides slow convergence of the chaos expansion and therefore causes additional errors in subsequent analysis. Modeling results within our new data-driven aPC show, at least, an exponential convergence for all examined cases. We defined, discussed and illustrated the difference between numerical integration error and transformation expansion error. Both errors lead to wrong estimates of statistical characteristics when using a non-optimal basis. Thus, the aPC not only provides freedom for modeling physical systems with unknown probability distribution function, when only data sets of very limited size are available, but it also provides better convergence rates than conventional polynomial chaos techniques.

When the statistical information used as input is inaccurate or uncertain, i.e. when using small sets of raw data, a new form of uncertainty enters into the analysis and has to be considered in convergence assessments. We propose to apply Jackknife or

Bootstrapping methods to assess the uncertainty of moments from small data sets, and then propagate that uncertainty through the aPC onto the output statistics to assess robustness. In an illustrative test case, we observe the classical Monte Carlo convergence rate for the results of the PCE analysis with respect to the size of the raw data set.

Acknowledgements

The authors would like to thank the German Research Foundation (DFG) for its financial support of the project within the Cluster of Excellence in Simulation Technology (EXC 310/1) at the University of Stuttgart.

The authors would also like to express their sincere thanks to Prof. Dr. Helmut Harbrecht of the Institute of Applied Analysis and Numeric Simulation, University of Stuttgart, and Prof. Dr. Guido Schneider of the Institute of Analysis, Dynamic and Modeling, University of Stuttgart, for helpful discussions. Also, we would like to thank the anonymous reviewers for their constructive comments.

Appendix: Non-singularity of the moments matrix

Let us write the square matrix of Eq. (14) in the following decomposed form:

$$\mathbf{M} = \begin{bmatrix} \mathbf{H} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \quad (28)$$

where

$$\mathbf{H} = \begin{bmatrix} \mu_0 & \cdots & \mu_{k-1} \\ \vdots & \vdots & \vdots \\ \mu_{k-1} & \cdots & \mu_{2k-2} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mu_k \\ \vdots \\ \mu_{2k-1} \end{bmatrix},$$

$$\mathbf{C} = [0 \ \dots \ 0], \quad \mathbf{D} = [1].$$

Evidently, \mathbf{D} is always invertible, and hence the determinant of \mathbf{M} is given by:

$$\det(\mathbf{M}) = \det(\mathbf{D})\det(\mathbf{H} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}) \quad (29)$$

Because $\det(\mathbf{D}) = 1$ and $\mathbf{C} = [0, \dots, 0]$, we obtain:

$$\det(\mathbf{M}) = \det(\mathbf{H}). \quad (30)$$

The matrix \mathbf{H} is also known as the Hankel matrix of moments. The properties of its determinant were studied in the paper [27]. Moreover, Karlin [24] showed that $\det(\mathbf{H})$ for $\text{rank}(\mathbf{H}) = k$ is zero if and only if the distribution of ξ has only k or fewer points of support. Thus, \mathbf{M} is non-singular if and only if the number of support points in the distribution of ξ is greater than k and if all moments up to order $2k-2$ are finite.

References

- [1] Abramowitz M, Stegun IA, editors. Handbook of mathematical functions with formulas, graphs, and mathematical tables. New York: Dover; 1965, ISBN 0-486-61272-4.
- [2] Askey R, Wilson J. Some basic hypergeometric polynomials that generalize Jacobi polynomials, Memoirs of the American Mathematical Society. Providence, RI: AMS; 1985. p. 319.
- [3] Augustin F, Gilg A, Paffrath M, Rentrop P, Wever U. Polynomial chaos for the approximation of uncertainties: Chances and limits. European Journal of Applied Mathematics 2008;19:149–90.
- [4] Blatman G, Sudret B. Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. Reliability Engineering and System Safety 2010;95:1216–29.
- [5] Buzzard GT. Global sensitivity analysis using sparse grid interpolation and polynomial chaos. Reliability Engineering and System Safety, 2011, <http://dx.doi.org/10.1016/j.res.2011.07.011>.

- [6] Caflisch RE. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica* 1998;7:1–49.
- [7] Cameron R, Martin W. The orthogonal development of non-linear functionals in series of Fourier–Hermite functionals. *Annals of Mathematics* 1947;48.
- [9] Crestaux T, LeMaitre O, Martinez JM. Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering and System Safety* 2009;94:1161–72.
- [10] Dargahi-Noubary GR, Razzaghi M. Bootstrap construction of the upper confidence limit for unreliability. *Reliability Engineering and System Safety* 1996;37(1):1–6.
- [11] Ditlevsen O, Madsen HO. *Structural reliability method*. New York: John Wiley & Sons; 1992 p. 384.
- [12] Efron B. The Jackknife, the bootstrap, and other resampling plans. *Society for Industrial Mathematics* 1987;100:55.
- [13] Er GK. A method for multi-parameter PDF estimation of random variables. *Structural Safety* 1998;20(1):25–36.
- [14] Ernst OG, Mugler A, Starkoff H-J, Ullmann E. On the convergence of generalized polynomial chaos expansions. *Mathematical Modelling and Numerical Analysis* 2012;46:317–39.
- [15] Helton JC, Davis FJ. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering and System Safety* 2003;81:23–69.
- [16] Isukapalli SS, Roy A, Georgopoulos PG. Stochastic response surface methods (SRSMs) for uncertainty propagation: application to environmental and biological systems. *Risk Analysis* 1998;18(3):351–63.
- [17] Jang YS, Sitar N, Kiureghian AD. Reliability analysis of contaminant transport in saturated porous media. *Water Resources Research* 1994;30(8):2435–48.
- [18] Jaynes ET. On the rationale of maximum-entropy methods. *Proceedings of the IEEE* 1982;106:939–52.
- [19] Ghanem R, Spanos P. A stochastic Galerkin expansion for non-linear random vibration analysis. *Probabilistic Engineering Mechanics* 1993;8:255–64.
- [20] Ghanem RG, Spanos PD. *Stochastic finite elements: a spectral approach*. New York: Springer-Verlag; 1991.
- [22] Grigoriu M. *Stochastic calculus: applications in science and engineering*. Boston: Birkhauser; 2002 755pp.
- [23] Keese A, Matthies HG. Sparse quadrature as an alternative to Monte Carlo for stochastic finite element techniques. *Proceedings in Applied Mathematics and Mechanics* 2003;3:493–4. <http://dx.doi.org/10.1002/pamm.200310516>.
- [24] Karlin S. *Total positivity*, vol. I. Stanford University Press; 1968, 576 pp.
- [25] Der Kiureghian A, Liu PL. Structural reliability under incomplete probability information Sparse quadrature as an alternative to Monte Carlo for stochastic finite element techniques. *Journal of Engineering Mechanics* 1986;112(1):85–104.
- [26] Li H, Zhang D. Probabilistic collocation method for flow in porous media: comparisons with other stochastic methods. *Water Resources Research* 2009;43:44–8.
- [27] Lindsay BG. On the determinants of moment matrices. *Annals of Statistics* 1989;17(2):711–21.
- [28] Lumley JL. The structure of inhomogeneous turbulent flows. In: Yaglom AM, Tatarski I, editors. *Atmospheric turbulence and wave propagation*; 1967. p. 166–78.
- [29] Maltz FH, Hitzl DL. Variance reduction in monte carlo computations using multi-dimensional hermite polynomials. *Journal of Computational Physics* 1979;2:345–76.
- [30] Matthies HG, Keese A. Galerkin methods for linear and non-linear elliptic stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering* 2005;194:1295–331.
- [31] Pearson K. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine* 1901;2(6):559–72.
- [32] Prempraneerach P, Hover FS, Triantafyllou MS, Karniadakis GE. Uncertainty quantification in simulations of power systems: Multi-element polynomial chaos methods. *Reliability Engineering and System Safety* 2010;95:632–46.
- [33] Sambucini V. A reference prior for the analysis of a response surface. *Journal of Statistical Planning and Inference* 2007;137(4):1119–28.
- [34] Sandoval EH, Anstett-Collin F, Basset M. Sensitivity study of dynamic systems using polynomial chaos. *Reliability Engineering and System Safety* 2012;104:15–26.
- [35] Sobol IM. On sensitivity estimation for nonlinear mathematical models. *Matematicheskoe Modelirovanie* 1990;2(1):112–8.
- [36] Sobol IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* 2001;55(1–3):271–80.
- [37] Soize C, Ghanem R. Physical systems with random uncertainties: chaos representations with arbitrary probability measure. *SIAM Journal on Scientific Computing* 2004;26(2):395–410.
- [38] Evans M, Hastings N, Peacock B. *Statistical distributions*. 3rd ed. Wiley-Interscience; 2000 (ISBN-10: 0471371246, 221 pp).
- [39] Stieltjes TJ. *Quelques Recherches sur la Théorie des Quadratures dites Mécaniques. Oeuvres I*; 1884. pp. 377–96.
- [40] Stone MH. Applications of the theory of boolean rings to general topology. *Transactions of the American Mathematical Society* 1937;41(3):375–481. <http://dx.doi.org/10.2307/1989788>.
- [41] Storlie C, Swiler L, Helton J, Sallaberry C. Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering and System Safety* 2009;94:1735–63.
- [42] Sudret B. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering and System Safety* 2008;93:964–79.
- [43] Oladyskhin S, Class H, Helmig R, Nowak W. An integrative approach to robust design and probabilistic risk assessment for CO₂ storage in geological formations. *Computational Geosciences* 2011;15(3):565–77.
- [44] Oladyskhin S, Class H, Helmig R, Nowak W. A concept for data-driven uncertainty quantification and its application to carbon dioxide storage in geological formations. *Advances in Water Resources* 2011;34:1508–18.
- [45] Oladyskhin S, de Barros FPJ, Nowak W. Global sensitivity analysis: a flexible and efficient framework with an example from stochastic hydrogeology. *Advances in Water Resources* 2012;37:1508–18.
- [46] Red-Horse JR, Benjamin AS. A probabilistic approach to uncertainty quantification with limited information. *Reliability Engineering and System Safety* 2004;85:183–90.
- [47] Riesz M. Sur le problème des moments et le théorème de Parseval correspondant. *Acta Litt. Acad. Sci. Szeged*. 1923;1:209–25.
- [48] Wackernagel H. *Multivariate geostatistics, an introduction with applications*. second ed. Berlin: Springer; 1998 ISBN 3-540-64721-X, 291 pp.
- [49] Walter L, Binning P, Oladyskhin S, Flemisch B, Class H. Brine migration resulting from CO₂ injection into saline aquifers—an approach to risk estimation including various levels of uncertainty. *International Journal of Greenhouse Gas Control* 2012;9:495–506. <http://dx.doi.org/10.1016/j.ijggc.2012.05.004>.
- [50] Wan X, Karniadakis GE. Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM Journal on Scientific Computing* 2006;28(3):901–28.
- [51] Wan X, Karniadakis GE. Error control in multi-element generalized polynomial chaos method for elliptic problems with random coefficients. *Communication in Computational Physics* 2009;5:793–820.
- [52] Wiener N. The homogeneous chaos. *American Journal of Mathematics* 1938;60:897–936.
- [53] Winterstein SR. Nonlinear vibration models for extremes and fatigue. *Journal of Engineering Mechanics ASCE* 1998;114:1772–90.
- [54] Witteveen JAS, Sarkar S, Bijl H. Modeling physical uncertainties in dynamic stall induced fluid-structure interaction of turbine blades using arbitrary polynomial chaos. *Computers and Structures* 2007;85:866–87.
- [55] Witteveen JAS, Bijl H. Modeling arbitrary uncertainties using Gram–Schmidt polynomial chaos. 44th AIAA aerospace sciences meeting and exhibit, Reno, Nevada, AIAA-2006-896; 2006.
- [56] Woodbury AD, Ulrych TJ. Minimum relative entropy: forward probabilistic modeling. *Water Resources Research* 1993;29(8):2847–60.
- [57] Xiu D, Karniadakis GE. The Wiener–Askey Polynomial Chaos for stochastic differential equations. *SIAM Journal on Scientific Computing* 2002;24(2):619–44.
- [58] Xiu D, Karniadakis GE. Modeling uncertainty in flow simulations via generalized polynomial chaos. *Journal of Computational Physics* 2003;187:137–67.
- [59] Xiu D. Efficient collocational approach for parametric uncertainty analysis. *Communications in Computational Physics* 2007;2(2):293–309.
- [60] Shi L, Yang JD, Li H. Probabilistic Collocation Method for Unconfined Flow in Heterogeneous Media. *Journal of Hydrology* 2009;365(1–2):4–10.