SIAM/ASA J. UNCERTAINTY QUANTIFICATION
Vol. 8, No. 1, pp. 27–57
© 2020 Sharif Rahman

A Spline Chaos Expansion*

Sharif Rahman†

Abstract. A spline chaos expansion, referred to as SCE, is introduced for uncertainty quantification analysis. The expansion provides a means for representing an output random variable of interest with respect to multivariate orthonormal basis splines (B-splines) in input random variables. The multivariate B-splines are built from a whitening transformation to generate univariate orthonormal B-splines in each coordinate direction, followed by a tensor-product structure to produce the multivariate version. SCE, as it stems from compactly supported B-splines, tackles locally prominent responses more effectively than the polynomial chaos expansion (PCE). The approximation quality of the expansion is demonstrated in terms of the modulus of smoothness of the output function, leading to the mean-square convergence of SCE to the correct limit. Analytical formulae are proposed to calculate the mean and variance of an SCE approximation for a general output variable in terms of the requisite expansion coefficients. Numerical results indicate that a low-order SCE approximation with an adequate mesh is markedly more accurate than a high-order PCE approximation in estimating the output variances and probability distributions of oscillatory, nonsmooth, and nearly discontinuous functions.

Key words. uncertainty quantification, B-splines, polynomial chaos expansion, stochastic analysis

AMS subject classifications. 26B49, 41A61, 49K30, 60H35, 65C60

DOI. 10.1137/19M1239702

1. Introduction.

Uncertainty quantification (UQ) of complex mathematical models is a cross-cutting research topic with broad impacts on engineering and applied sciences [10, 18, 19]. A frequently employed method for UQ analysis entails polynomial chaos expansion (PCE), which describes an infinite series expansion of a square-integrable output random variable in terms of measure-consistent orthogonal polynomials in input random variables [1, 7, 22]. The expansion is largely predicated on the smoothness assumption of the output function, because the polynomial basis of PCE is globally supported. While polynomials have many attractive properties, they possess one undesirable feature: polynomials may oscillate wildly [17]. As soon as the expansion degree or order[1] exceeds four or five, a PCE approximation becomes prone to unstable swings. This is chiefly because polynomials are inflexible if they are too smooth, long heralded as a virtue. They are analytic, which means that the behavior of a polynomial in an arbitrarily small region determines the behavior everywhere. In the physical world, though, the output function is frequently of a disjointed nature, meaning that the behavior in one region may be completely unrelated to the behavior in another region. In

*Received by the editors January 22, 2019; accepted for publication (in revised form) November 5, 2019; published electronically January 8, 2020.
https://doi.org/10.1137/19M1239702
Funding: This work was supported by the U.S. National Science Foundation under grant CMMI-1607398.
†Applied Mathematics and Computational Sciences, The University of Iowa, Iowa City, IA, 52245 (sharif-rahman@uiowa.edu).

[1]The nouns *degree* and *order* of a polynomial or spline expansion are used synonymously in the paper.

this case, the convergence property of PCE or other polynomial-based methods may become markedly deteriorated. In an effort to enhance the performance of global supported PCE, domain decomposition techniques, such as multielement formulation of PCE, have been introduced [21]. However, in the presence of large subdomains of discontinuities, the multielement PCE becomes computationally inefficient, especially when there are many input random variables. Therefore, alternative UQ methods, proficient in tackling locally pronounced highly nonlinear or nonsmooth output functions, are desirable.

This paper presents a new, alternative orthogonal expansion, referred to as spline chaos expansion or SCE, for UQ analysis subject to independent but otherwise arbitrary probability measures of input random variables. The paper is structured as follows. Section 2 starts with mathematical preliminaries and assumptions. A brief exposition of univariate basis splines (B-splines) is given in section 3. This is followed by a presentation of orthonormal B-splines, including their second-moment properties, in section 4. Section 5 describes the construction of multivariate B-splines and explains how they form an orthonormal basis of a spline space of interest. Section 6 formally presents SCE for a square-integrable random variable and then demonstrates the convergence and optimality of SCE. The formulae for the mean and variance of an SCE approximation are derived. The results from three numerical examples are reported in section 7. Section 8 discusses future work. Finally, conclusions are drawn in section 9.

2. Input random variables.

Let $\mathbb{N} := \{1, 2, \ldots\}$, $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, and $\mathbb{R} := (-\infty, +\infty)$ represent the sets of positive integers (natural), nonnegative integers, and real numbers, respectively. Denote by $[a_k, b_k]$ a finite closed interval, where $a_k, b_k \in \mathbb{R}$, $b_k > a_k$. Then, given $N \in \mathbb{N}$, $\mathbb{A}^N = \times_{k=1}^N [a_k, b_k]$ represents a closed bounded domain of $\mathbb{R}^N$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where Ω is a sample space representing an abstract set of elementary events, $\mathcal{F}$ is a σ-algebra on Ω, and $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a probability measure. Defined on this probability space, consider an N-dimensional input random vector $\mathbf{X} := (X_1, \ldots, X_N)^\intercal$, describing the statistical uncertainties in all system parameters of a stochastic or UQ problem. Denote by $F_{\mathbf{X}}(\mathbf{x}) := \mathbb{P}(\cap_{i=1}^N \{X_k \leq x_k\})$ the joint distribution function of $\mathbf{X}$. The kth component of $\mathbf{X}$ is a random variable X_k, which has the marginal probability distribution function $F_{X_k}(x_k) := \mathbb{P}(X_k \leq x_k)$. In the UQ community, the input random variables are also known as basic random variables. The nonzero, finite integer N represents the number of input random variables and is often referred to as the dimension of the stochastic or UQ problem.

A set of assumptions on input random variables used or required by SCE is as follows.

Assumption 2.1. *The input random vector $\mathbf{X} := (X_1, \ldots, X_N)^\intercal$ satisfies all of the following conditions:*

1. *All component random variables X_k, $k = 1, \ldots, N$, are statistically independent but not necessarily identically distributed.*
2. *Each input random variable X_k is defined on a bounded interval $[a_k, b_k] \subset \mathbb{R}$. Therefore, all moments of X_k exist, that is, for all $l \in \mathbb{N}_0$,*

$$\mathbb{E}\left[X_k^l\right] := \int_\Omega X_k^l(\omega) d\mathbb{P}(\omega) < \infty,$$

where $\mathbb{E}$ is the expectation operator with respect to the probability measure $\mathbb{P}$.

3. *Each input random variable X_k has absolutely continuous marginal probability distribution function $F_{X_k}(x_k)$ and continuous marginal probability density function $f_{X_k}(x_k) := \partial F_{X_k}(x_k)/\partial x_k$ with a bounded support $[a_k, b_k] \subset \mathbb{R}$. Consequently, with items 1 and 2 in mind, the joint probability distribution function $F_{\mathbf{X}}(\mathbf{x})$ and joint probability density function $f_{\mathbf{X}}(\mathbf{x}) := \partial^N F_{\mathbf{X}}(\mathbf{x})/\partial x_1 \cdots \partial x_N$ of $\mathbf{X}$ are obtained from*

$$F_{\mathbf{X}}(\mathbf{x}) = \prod_{k=1}^{N} F_{X_k}(x_k) \quad \textit{and} \quad f_{\mathbf{X}}(\mathbf{x}) = \prod_{k=1}^{N} f_{X_k}(x_k),$$

respectively, with a bounded support $\mathbb{A}^N \subset \mathbb{R}^N$ of the density function.

Assumption 2.1 asserts the existence of a relevant sequence of orthogonal polynomials or splines consistent with the input probability measure. The discrete distributions and dependent variables are not dealt with in this paper.

Given the abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$ of $\mathbf{X}$, there exists an image probability space $(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}} d\mathbf{x})$, where $\mathbb{A}^N$ is the image of Ω from the mapping $\mathbf{X} : \Omega \to \mathbb{A}^N$ and $\mathcal{B}^N := \mathcal{B}(\mathbb{A}^N)$ is the Borel σ-algebra on $\mathbb{A}^N \subset \mathbb{R}^N$. Relevant statements and objects in the abstract probability space have obvious counterparts in the associated image probability space. Both probability spaces will be exploited in this paper.

3. Univariate B-splines.

Let $\mathbf{x} = (x_1, \ldots, x_N)$ be an arbitrary point in $\mathbb{A}^N$. For the coordinate direction k, $k = 1, \ldots, N$, define a positive integer $n_k \in \mathbb{N}$ and a nonnegative integer $p_k \in \mathbb{N}_0$, representing the total number of basis functions and polynomial degree, respectively. The rest of this section briefly describes paraphernalia of univariate B-splines.

3.1. Knot sequence.

In order to define B-splines, the concept of knot sequence, also referred to as knot vector by some, for each coordinate direction k is needed.

Definition 3.1. *A knot sequence $\boldsymbol{\xi}_k$ for the interval $[a_k, b_k] \subset \mathbb{R}$, given $n_k > p_k \geq 0$, is a nondecreasing sequence of real numbers*

$$\begin{aligned} \boldsymbol{\xi}_k := \{\xi_{k,i_k}\}_{i_k=1}^{n_k+p_k+1} &= \{a_k = \xi_{k,1}, \xi_{k,2}, \ldots, \xi_{k,n_k+p_k+1} = b_k\}, \\ \xi_{k,1} \leq \xi_{k,2} &\leq \cdots \leq \xi_{k,n_k+p_k+1}, \end{aligned} \tag{3.1}$$

where ξ_{k,i_k} is the i_kth knot with $i_k = 1, 2, \ldots, n_k + p_k + 1$ representing the knot index for the coordinate direction k. The elements of $\boldsymbol{\xi}_k$ are called knots.

According to (3.1), the total number of knots is $n_k + p_k + 1$. The knots may be equally spaced or unequally spaced, resulting in a uniform or nonuniform distribution. More importantly, the knots, whether they are exterior or interior, may be repeated, that is, a knot ξ_{k,i_k} of the knot sequence $\boldsymbol{\xi}_k$ may appear $1 \leq m_{k,i_k} \leq p_k + 1$ times, where m_{k,i_k} is referred to as its multiplicity. The multiplicity has important implications on the regularity properties of B-spline functions. To monitor knots without repetitions, say, there are r_k distinct knots $\zeta_{k,1}, \ldots, \zeta_{k,r_k}$ in $\boldsymbol{\xi}_k$ with respective multiplicities $m_{k,1}, \ldots, m_{k,r_k}$. Then the knot sequence in (3.1) can be expressed more precisely by

$$\boldsymbol{\xi}_k = \{a_k = \overbrace{\zeta_{k,1}, \ldots, \zeta_{k,1}}^{m_{k,1} \text{ times}}, \overbrace{\zeta_{k,2}, \ldots, \zeta_{k,2}}^{m_{k,2} \text{ times}}, \ldots, \overbrace{\zeta_{k,r_k-1}, \ldots, \zeta_{k,r_k-1}}^{m_{k,r_k-1} \text{ times}}, \overbrace{\zeta_{k,r_k}, \ldots, \zeta_{k,r_k}}^{m_{k,r_k} \text{ times}} = b_k\},$$
$$a_k = \zeta_{k,1} < \zeta_{k,2} < \cdots < \zeta_{k,r_k-1} < \zeta_{k,r_k} = b_k,$$

which consists of a total number of

$$\sum_{i_k=1}^{r_k} m_{k,i_k} = n_k + p_k + 1$$

knots. A knot sequence is called open if the end knots have multiplicities $p_k + 1$. In this case, definitions of more specific knot sequences are in order.

Definition 3.2. *A knot sequence is said to be (p_k+1)-open if the first and last knots appear $p_k + 1$ times, that is, if*

$$\boldsymbol{\xi}_k = \{a_k = \overbrace{\zeta_{k,1}, \ldots, \zeta_{k,1}}^{p_k+1 \text{ times}}, \overbrace{\zeta_{k,2}, \ldots, \zeta_{k,2}}^{m_{k,2} \text{ times}}, \ldots, \overbrace{\zeta_{k,r_k-1}, \ldots, \zeta_{k,r_k-1}}^{m_{k,r_k-1} \text{ times}}, \overbrace{\zeta_{k,r_k}, \ldots, \zeta_{k,r_k}}^{p_k+1 \text{ times}} = b_k\}, \tag{3.2}$$
$$a_k = \zeta_{k,1} < \zeta_{k,2} < \cdots < \zeta_{k,r_k-1} < \zeta_{k,r_k} = b_k.$$

Definition 3.3. *A knot sequence is said to be (p_k+1)-open with simple knots if it is (p_k+1)-open and all interior knots appear only once, that is, if*

$$\boldsymbol{\xi}_k = \{a_k = \overbrace{\zeta_{k,1}, \ldots, \zeta_{k,1}}^{p_k+1 \text{ times}}, \zeta_{k,2}, \ldots, \zeta_{k,r_k-1}, \overbrace{\zeta_{k,r_k}, \ldots, \zeta_{k,r_k}}^{p_k+1 \text{ times}} = b_k\},$$
$$a_k = \zeta_{k,1} < \zeta_{k,2} < \cdots < \zeta_{k,r_k-1} < \zeta_{k,r_k} = b_k.$$

A $(p_k + 1)$-open knot sequence with or without simple knots is commonly found in applications [2].

3.2. B-splines. The B-spline functions for a given degree are defined in a recursive manner using the knot sequence as follows.

Definition 3.4. *Let $\boldsymbol{\xi}_k$ be a general knot sequence of length at least $p_k + 2$ for the interval $[a_k, b_k]$, as defined by (3.1). Denote by $B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k)$ the i_kth univariate B-spline function with degree $p_k \in \mathbb{N}_0$ for the coordinate direction k. Given the zero-degree basis functions,*

$$B^k_{i_k,0,\boldsymbol{\xi}_k}(x_k) := \begin{cases} 1, & \xi_{k,i_k} \le x_k < \xi_{k,i_k+1}, \\ 0, & \textit{otherwise}, \end{cases}$$

for $k = 1, \ldots, N$, all higher-order B-spline functions on $\mathbb{R}$ are defined recursively by

$$B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k) := \frac{x_k - \xi_{k,i_k}}{\xi_{k,i_k+p_k} - \xi_{k,i_k}} B^k_{i_k,p_k-1,\boldsymbol{\xi}_k}(x_k) + \frac{\xi_{k,i_k+p_k+1} - x_k}{\xi_{k,i_k+p_k+1} - \xi_{k,i_k+1}} B^k_{i_k+1,p_k-1,\boldsymbol{\xi}_k}(x_k),$$

where $1 \le k \le N$, $1 \le i_k \le n_k$, $1 \le p_k < \infty$, and $0/0$ is considered as zero.

The recursive formula in Definition 3.4 is due to Cox [3] and de Boor [5]. However, a similar formula was reported by Popoviciu and Chakalov in the 1930s [6]. For alternative definitions, such as those involving divided differences, readers should consult the seminal work of Schoenberg [16].

The B-spline functions satisfy the following desirable properties [3, 5, 12].

Property 3.5. *They are nonnegative, that is,* $B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k) \geq 0$ *for all* i_k *and* x_k.

Property 3.6. *They are locally supported on the interval* $[\xi_{k,i_k}, \xi_{k,i_k+p_k+1})$ *for all* i_k.

Property 3.7. *They are linearly independent, that is, if*

$$\sum_{i_k=1}^{n_k} c^k_{i_k} B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k) = 0,$$

then $c^k_{i_k} = 0$ *for all* i_k.

Property 3.8. *They form a partition of unity, that is,*

$$\sum_{i_k=1}^{n_k} B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k) = 1, \;\; x_k \in [\xi_{k,1}, \xi_{k,n_k+p_k+1}].$$

Property 3.9. *They are pointwise* C^∞*-continuous everywhere except at the knots* ξ_{k,i_k} *of multiplicity* m_{k,i_k}*, where it is* $C^{p_k-m_{k,i_k}}$*-continuous, provided that* $1 \leq m_{k,i_k} < p_k + 1$.

For an illustration, consider $k = 1$, $a_1 = 0$, $b_1 = 1$, $p_1 = 2$, and two open knot sequences:

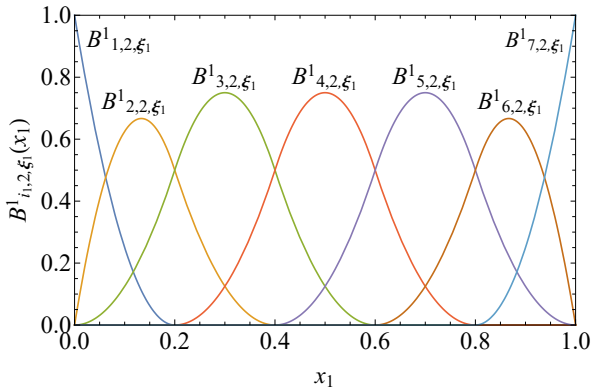
$$\begin{aligned}&(1)\ \boldsymbol{\xi}_1 = \{0, 0, 0, 0.2, 0.4, 0.6, 0.8, 1, 1, 1\},\\ &(2)\ \boldsymbol{\lambda}_1 = \{0, 0, 0, 0.2, 0.4, 0.6, 0.6, 0.8, 1, 1, 1\}.\end{aligned}$$

Here, $\boldsymbol{\xi}_1$ is a three-open knot sequence with simple knots because the multiplicity of each interior knot is one. In contrast, $\boldsymbol{\lambda}_1$ is merely a three-open knot sequence, as the multiplicity of the sixth knot is two. Consequently, there are seven and eight univariate quadratic B-spline basis functions for these two cases: $B^1_{i_1,2,\boldsymbol{\xi}_1}(x_1)$, $i_1 = 1, \ldots, 7$; and $B^1_{i_1,2,\boldsymbol{\lambda}_1}(x_1)$, $i_1 = 1, \ldots, 8$, which are illustrated in Figures 1(a) and 1(b), respectively. The basis functions for the first case are C^1-continuous at all interior knots, whereas the basis functions for the second case are C^0-continuous at $\lambda_{1,6} = \lambda_{1,7} = 0.6$ and C^1-continuous at other interior knots. Clearly, the regularities of B-splines depend on the degree and multiplicities of the knots selected.

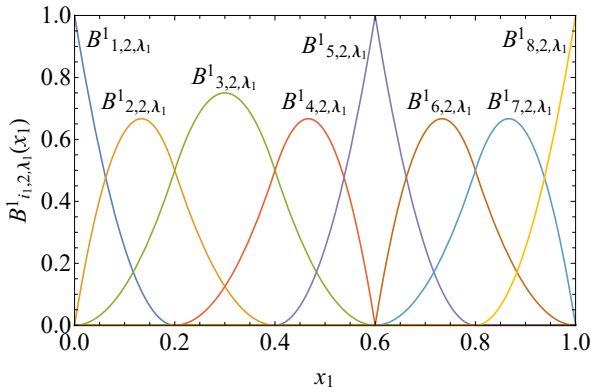
3.3. Spline space. Suppose for $n_k > p_k \geq 0$, a knot sequence $\boldsymbol{\xi}_k$ has been specified on the interval $[a_k, b_k]$. The associated spline space of degree p_k, denoted by $\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k}$, is conveniently defined using an appropriate polynomial space. Define such a polynomial space as a finite-dimensional linear space

$$\Pi_{p_k} := \left\{ g(x_k) = \sum_{l=0}^{p_k} c_{k,l} x_k^l : c_{k,l} \in \mathbb{R} \right\}$$

of real-valued polynomials in x_k of degree at most p_k.



(a)



(b)

Figure 1. *Quadratic B-splines generated on the interval* [0,1]*;* (a) *seven B-splines for* $\boldsymbol{\xi}_1 = \{0, 0, 0, 0.2, 0.4, 0.6, 0.8, 1, 1, 1\}$*;* (b) *eight B-splines for* $\boldsymbol{\lambda}_1 = \{0, 0, 0, 0.2, 0.4, 0.6, 0.6, 0.8, 1, 1, 1\}$.

Definition 3.10 (Schumaker [17]). *For* $n_k > p_k \geq 0$*, let* $\boldsymbol{\xi}_k$ *be a* (p_k+1)*-open knot sequence on the interval* $[a_k, b_k]$*, as defined by* (3.2)*. Then the space*

(3.3)
$$\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k} := \left\{ \begin{array}{l} g_k : [a_k, b_k] \to \mathbb{R} : \ \textit{there exist polynomials } g_{k,1}, g_{k,2}, \ldots, g_{k,r_k-1} \textit{ in } \Pi_{p_k} \\ \textit{such that } g_k(x_k) = g_{k,i_k}(x_k) \textit{ for } x_k \in [\xi_{k,i_k}, \xi_{k,i_k+1}), \ i_k = 1, \ldots, r_k - 1, \\ \textit{and } \dfrac{\partial^{j_k} g_{k,i_k-1}}{\partial x_k}(\xi_{k,i_k}) = \dfrac{\partial^{j_k} g_{k,i_k}}{\partial x_k}(\xi_{k,i_k}) \textit{ for } j_k = 0, 1, \ldots, p_k - m_{k,i_k}, \\ i_k = 2, \ldots, r_k - 1 \end{array} \right\}$$

is defined as the spline space of degree p_k *with distinct knots* $\zeta_{k,1}, \ldots, \zeta_{k,r_k}$ *of multiplicities* $m_{k,1} = p_k + 1$*,* $1 \leq m_{k,2} \leq p_k + 1$*, …,* $1 \leq m_{k,r_k-1} \leq p_k + 1$*,* $m_{k,r_k} = p_k + 1$*.*

The spline space is uniquely determined by distinct interior knots $\zeta_{k,2}, \ldots, \zeta_{k,r_k-1}$ of multiplicities $m_{k,2}, \ldots, m_{k,r_k-1}$. Indeed, the multiplicities decide the nature of $\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k}$ by controlling the smoothness of the splines at interior knots. For instance, if $m_{k,i_k} = p_k+1$, $i_k = 2, \ldots, r_k-1$,

then two polynomial pieces g_{k,i_k-1} and g_{k,i_k} in the subintervals adjoining the knot ξ_{k,i_k} are unrelated, possibly forming a jump discontinuity at ξ_{k,i_k}. In this case, $\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k}$ will be the roughest space of splines. If $m_{k,i_k} < p_k + 1$, $i_k = 2, \ldots, r_k - 1$, then the two aforementioned polynomial pieces are connected smoothly in the sense that the first $p_k - m_{k,i_k}$ derivatives are all continuous across the knot. More specifically, if $m_{k,i_k} = 1$, $i_k = 2, \ldots, r_k - 1$, then there are simple knots with the corresponding spline space becoming the smoothest space of piecewise polynomials of degree at most p_k.

Proposition 3.11 (Schumaker [17]). *The spline space $\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k}$ is a linear space of dimension*

$$\dim \mathcal{S}_{k,p_k,\boldsymbol{\xi}_k} = n_k = \sum_{i_k=2}^{r_k-1} m_{k,i_k} + p_k + 1. \tag{3.4}$$

Proposition 3.12 (Schumaker [17]). *For $n_k > p_k \geq 0$, let $\boldsymbol{\xi}_k$ be a (p_k+1)-open knot sequence on the interval $[a_k, b_k]$. Denote by*

$$\left\{B^k_{1,p_k,\boldsymbol{\xi}_k}(x_k), \ldots, B^k_{n_k,p_k,\boldsymbol{\xi}_k}(x_k)\right\} \tag{3.5}$$

a set of n_k B-splines of degree p_k. Then

$$\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k} = \operatorname{span}\left\{B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k)\right\}_{i_k=1,\ldots,n_k}.$$

4. Orthonormal B-splines. The B-splines presented in the preceding section, although they form a basis of the spline space $\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k}$, are obtained without any explicit consideration of the probability law of X_k. Therefore, they are not orthogonal with respect to the probability measure $f_{X_k}(x_k)dx_k$. A popular choice for constructing orthogonal or orthonormal basis is the Gram–Schmidt procedure [9]. However, it is known to be ill-conditioned. Therefore, more stable methods are needed to compute orthonormal splines consistent with the input probability measure. In this section, a linear transformation is proposed to generate their orthonormal version. The latter splines facilitate an orthogonal series expansion in a Hilbert space, resulting in concise forms of the expansion and second-moment properties of an output random variable of interest.

4.1. Spline moment matrix. In reference to the set of B-splines in (3.5), consider replacing any one of its elements with an arbitrary nonzero constant, thus creating an auxiliary set. Without loss of generality, let

$$\left\{1, B^k_{2,p_k,\boldsymbol{\xi}_k}(x_k), \ldots, B^k_{n_k,p_k,\boldsymbol{\xi}_k}(x_k)\right\} \tag{4.1}$$

be such a set, obtained by replacing the first element of (3.5) with 1. Proposition 4.1 shows that the auxiliary B-splines are also linearly independent.

Proposition 4.1. *The auxiliary set of B-splines in (4.1) is linearly independent.*

Proof. For constants $\bar{c}^k_{i_k} \in \mathbb{R}$, $i_k = 1, \ldots, n_k$, set

$$\bar{c}^k_1 + \sum_{i_k=2}^{n_k} \bar{c}^k_{i_k} B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k) = 0. \tag{4.2}$$

Using Property 3.8, write (4.2) as

$$\bar{c}^k_1 B^k_{1,p_k,\boldsymbol{\xi}_k}(x_k) + \sum_{i_k=2}^{n_k} \left(\bar{c}^k_1 + \bar{c}^k_{i_k}\right) B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k) = 0. \tag{4.3}$$

From Property 3.7, $\{B^k_{1,p_k,\boldsymbol{\xi}_k}(x_k), \ldots, B^k_{n_k,p_k,\boldsymbol{\xi}_k}(x_k)\}$ is linearly independent, meaning that the coefficients of (4.3) must all vanish. Consequently,

$$\bar{c}^k_{i_k} = 0, \;\; i_k = 1, \ldots, n_k,$$

completing the proof. ∎

When the input random variable X_k, instead of the real variable x_k, is inserted in the argument, the elements of the auxiliary set become random B-splines. A formal definition of the spline moment matrix follows.

Definition 4.2. *Let*

$$\mathbf{P}_k(X_k) := (1, B^k_{2,p_k,\boldsymbol{\xi}_k}(X_k), \ldots, B^k_{n_k,p_k,\boldsymbol{\xi}_k}(X_k))^\intercal$$

be an n_k-dimensional vector of constant or random B-splines. Then the $n_k \times n_k$ matrix, defined by

$$\mathbf{G}_k := \mathbb{E}[\mathbf{P}_k(X_k)\mathbf{P}^\intercal_k(X_k)],$$

is called the spline moment matrix of $\mathbf{P}_k(X_k)$. The matrix $\mathbf{G}_k$ exists as X_k has finite moments up to order $2p_k$, as mandated by Assumption 2.1.

Here, any element of $\mathbf{G}_k$ represents the expectation of the product between two random splines. However, $\mathbf{G}_k$ is not the covariance matrix of $\mathbf{P}_k(X_k)$, as the means of B-splines are not *zero*.

Proposition 4.3. *The spline moment matrix $\mathbf{G}_k$ is symmetric and positive-definite.*

Proof. By definition, $\mathbf{G}_k = \mathbf{G}^\intercal_k$. From Proposition 4.1, the elements of $\mathbf{P}_k(x_k)$ are linearly independent. Hence, the spline moment matrix is a Gram matrix and is, therefore, positive-definite. ∎

4.2. Whitening transformation. From Proposition 4.3, $\mathbf{G}_k$ is positive-definite and therefore invertible. Consequently, there is a nonsingular whitening matrix $\mathbf{W}_k \in \mathbb{R}^{n_k \times n_k}$ such that the factorization

$$\mathbf{W}^\intercal_k \mathbf{W}_k = \mathbf{G}^{-1}_k \quad \text{or} \quad \mathbf{W}^{-1}_k \mathbf{W}^{-\intercal}_k = \mathbf{G}_k \tag{4.4}$$

holds. This leads to a set of orthonormal B-splines.

Definition 4.4. *Let* $\mathbf{X} := (X_1, \ldots, X_N)^\intercal$ *be a vector of* $N \in \mathbb{N}$ *input random variables fulfilling Assumption 2.1. Recall, for* $n_k > p_k \geq 0$ *and a specified knot sequence* $\boldsymbol{\xi}_k$*, that* $\mathbf{P}_k(X_k)$ *represents an* n_k*-dimensional vector of B-splines of degree* p_k*. Then the corresponding* n_k*-dimensional vector*

$$\boldsymbol{\psi}_k(X_k) := \left(\psi^k_{1,p_k,\boldsymbol{\xi}_k}(X_k), \ldots, \psi^k_{n_k,p_k,\boldsymbol{\xi}_k}(X_k)\right)^\intercal$$

of orthonormal B-splines, also of degree p_k*, is obtained from the whitening transformation*

$$\boldsymbol{\psi}_k(X_k) = \mathbf{W}_k \mathbf{P}_k(X_k), \tag{4.5}$$

where $\mathbf{W}_k \in \mathbb{R}^{n_k \times n_k}$ *is a nonsingular whitening matrix satisfying* (4.4)*.*

The whitening transformation in Definition 4.4 is a linear transformation that converts $\mathbf{P}_k(X_k)$ into $\boldsymbol{\psi}_k(X_k)$ in such a way that the latter has uncorrelated random B-splines. The transformation is called "whitening" because it changes one random vector to the other, which has statistical properties akin to that of a white noise vector. However, the condition (4.4) does not uniquely determine the whitening matrix $\mathbf{W}_k$. There are infinitely many choices of $\mathbf{W}_k$ satisfying (4.4). All of these choices result in a linear transformation, decorrelating $\mathbf{P}_k(X_k)$ but producing different random vectors $\boldsymbol{\psi}_k(X_k)$ [11, 14].

A prominent choice for $\mathbf{W}_k$, obtained from the Cholesky factorization $\mathbf{G}_k = \mathbf{Q}_k \mathbf{Q}_k^\intercal$, is

$$\mathbf{W}_k = \mathbf{Q}_k^{-1}, \tag{4.6}$$

where $\mathbf{Q}_k$ is an $n_k \times n_k$ lower-triangular matrix. The rest of the paper will use the Cholesky factorization. Nonetheless, other whitening matrices, in conjunction with (4.4), can be used to generate orthonormal B-splines.

Proposition 4.5. *Given the preambles of Propositions 3.12 and 4.1, the set of elements of* $\boldsymbol{\psi}_k(x_k)$ *from Definition 4.4 also spans the spline space* $\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k}$*, that is,*

$$\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k} := \operatorname{span}\left\{\psi^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k)\right\}_{i_k=1,\ldots,n_k}.$$

A proof of Proposition 4.5 can be obtained by recognizing the elements of $\boldsymbol{\psi}_k(x_k)$ to be linearly independent.

4.3. Statistical properties. Similar to $\mathbf{P}_k(X_k)$, $\boldsymbol{\psi}_k(X_k)$ is also a function of random input variable X_k. Proposition 4.6 describes its second-moment properties.

Proposition 4.6. *Let* $\mathbf{X} := (X_1, \ldots, X_N)^\intercal : (\Omega, \mathcal{F}) \to (\mathbb{A}^N, \mathcal{B}^N)$ *be a vector of* $N \in \mathbb{N}$ *input random variables fulfilling Assumption 2.1. If the whitening matrix is selected as* $\mathbf{Q}_k^{-1}$*, then the first- and second-order moments of the vector of orthonormal B-splines* $\boldsymbol{\psi}_k(X_k) = \mathbf{Q}_k^{-1}\mathbf{P}_k(X_k)$*,* $k = 1, \ldots, N$*, are*

$$\mathbb{E}\left[\boldsymbol{\psi}_k(X_k)\right] = (1, 0, \ldots, 0)^\intercal \tag{4.7}$$

and

$$\mathbb{E}\left[\boldsymbol{\psi}_k(X_k)\boldsymbol{\psi}_k^\intercal(X_k)\right] = \mathbf{I}_{n_k}, \tag{4.8}$$

respectively, where $\mathbf{I}_{n_k}$ *is the* $n_k \times n_k$ *identity matrix.*

Proof. Using (4.6) in the whitening transformation (4.5),

$$\begin{aligned}\mathbb{E}[\boldsymbol{\psi}_k(X_k)\boldsymbol{\psi}_k^\intercal(X_k)] &= \mathbf{Q}_k^{-1}\mathbb{E}[\mathbf{P}_k(X_k)\mathbf{P}_k^\intercal(X_k)]\mathbf{Q}_k^{-\intercal} \\ &= \mathbf{Q}_k^{-1}\mathbf{G}_k\mathbf{Q}_k^{-\intercal} \\ &= \mathbf{Q}_k^{-1}\mathbf{Q}_k\mathbf{Q}_k^\intercal\mathbf{Q}_k^{-\intercal} = \mathbf{I}_{n_k},\end{aligned}$$

obtaining (4.8). Recognize that $\psi^k_{1,p_k,\boldsymbol{\xi}_k}(X_k)$, the first element of $\boldsymbol{\psi}_k(X_k)$, is *one*. Then, using (4.8), the expectations of products between the first row of $\boldsymbol{\psi}_k(X_k)$ and all n_k columns of $\boldsymbol{\psi}_k^\intercal(X_k)$ produce (4.7). ■

5. Multivariate B-splines.

As the input vector $\mathbf{X} = (X_1, \ldots, X_N)^\intercal$ comprises independent random variables, its joint probability density function is the product of its marginal density functions. Consequently, measure-consistent multivariate orthonormal B-splines can be easily constructed from the tensor-product of measure-consistent univariate B-splines.

5.1. Tensor-product spline space.

For each $k = 1, \ldots, N$, suppose the knot sequence $\boldsymbol{\xi}_k$ on the interval $\mathbb{A}^{\{k\}} = [a_k, b_k]$, number of basis functions n_k, and degree p_k have been specified. The associated vector of measure-consistent univariate orthonormal splines in x_k is

$$\boldsymbol{\psi}_k(x_k) := \left(\psi^k_{1,p_k,\boldsymbol{\xi}_k}(x_k), \ldots, \psi^k_{n_k,p_k,\boldsymbol{\xi}_k}(x_k)\right)^\intercal.$$

Correspondingly, the spline space is $\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k}$, as expressed by (3.3). To define tensor-product B-splines in N variables and the associated spline space, define a multi-index $\mathbf{p} := (p_1, \ldots, p_N) \in \mathbb{N}_0^N$, representing the degrees of splines in all N coordinate directions. Denote by $\boldsymbol{\Xi} := \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N\}$ a family of all N knot sequences. Because of the tensor nature of the resulting space, many properties of univariate splines carry over, described as follows.

Definition 5.1. *Given $\mathbf{p} := (p_1, \ldots, p_N)$ and $\boldsymbol{\Xi} := \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N\}$, the tensor-product spline space, denoted by $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$, is defined by*

$$\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}} := \bigotimes_{k=1}^{N} \mathcal{S}_{k,p_k,\boldsymbol{\xi}_k},$$

where the symbol $\bigotimes$ stands for tensor product.

It is clear from Definition 5.1 that $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$ is a linear space of dimension $\prod_{k=1}^N n_k$. Here, n_k, the dimension of the spline space $\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k}$, is obtained from (3.4) when each knot sequence is chosen according to (3.2). Each spline $g \in \mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$ is defined on the N-dimensional rectangular domain

$$\mathbb{A}^N := \times_{k=1}^N \mathbb{A}^{\{k\}} = \times_{k=1}^N [a_k, b_k].$$

Define two additional multi-indices $\mathbf{i} := (i_1, \ldots, i_N) \in \mathbb{N}^N$ and $\mathbf{n} := (n_1, \ldots, n_N) \in \mathbb{N}^N$, representing the knot indices and numbers of univariate basis functions, respectively, in all N coordinate directions. Associated with $\mathbf{i}$, define an index set

$$\mathcal{I}_{\mathbf{n}} := \{\mathbf{i} = (i_1, \ldots, i_N) : 1 \le i_k \le n_k,\ k = 1, \ldots, N\} \subset \mathbb{N}^N$$

which has cardinality

$$|\mathcal{I}_{\mathbf{n}}| = \prod_{k=1}^{N} n_k,$$

thus matching the dimension of $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$. Then the partition defined by the knot sequences $\boldsymbol{\xi}_k$, $k = 1, \ldots, N$, splits $\mathbb{A}^N$ into smaller N-dimensional rectangles

$$\mathbb{A}^N_{\mathbf{i}} = \{\mathbf{x} : \zeta_{k,i_k} \leq x_k < \zeta_{k,i_k+1},\ k = 1, \ldots, N\},$$
$$\mathbf{i} \in \{\mathbf{i} = (i_1, \ldots, i_N) : 1 \leq i_k \leq r_k - 1,\ k = 1, \ldots, N\} \subseteq \mathcal{I}_{\mathbf{n}}.$$

A mesh is defined by the partition of $\mathbb{A}^N$ into rectangular elements $\mathbb{A}^N_{\mathbf{i}}$. Define the largest element size in each coordinate direction k by

$$h_k := \max_{1 \leq l \leq r_k - 1} (\zeta_{k,l+1} - \zeta_{k,l}),\ k = 1, \ldots, N.$$

Then, given the family of knot sequences $\boldsymbol{\Xi} = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N\}$,

$$\mathbf{h} := (h_1, \ldots, h_N) \quad \text{and} \quad h := \max_{1 \leq k \leq N} h_k$$

define a vector of the largest element sizes in all N coordinates and the global element size, respectively, for the domain $\mathbb{A}^N$.

5.2. Tensor-product orthonormal B-splines. Given the B-splines for all N coordinate directions, a formal definition of tensor-product B-splines is as follows.

Definition 5.2. *Let $\mathbf{X} := (X_1, \ldots, X_N)^\intercal : (\Omega, \mathcal{F}) \to (\mathbb{A}^N, \mathcal{B}^N)$ be a vector of $N \in \mathbb{N}$ input random variables fulfilling Assumption 2.1. Suppose the univariate orthonormal B-splines consistent with the marginal probability measures in all coordinate directions have been obtained as the sets $\{\psi^k_{1,p_k,\boldsymbol{\xi}_k}(x_k), \ldots, \psi^k_{n_k,p_k,\boldsymbol{\xi}_k}(x_k)\}$, $k = 1, \ldots, N$. Then, for $\mathbf{p} = (p_1, \ldots, p_N) \in \mathbb{N}_0^N$ and $\boldsymbol{\Xi} = \{\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N\}$, the multivariate orthonormal B-splines in $\mathbf{x}$ consistent with the probability measure $f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$ are defined as*

$$\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x}) := \prod_{k=1}^{N} \psi^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k), \quad \mathbf{i} = (i_1, \ldots, i_N) \in \mathcal{I}_{\mathbf{n}}.$$

5.3. Statistical properties. When the input random variables $X_1, \ldots, X_N$, instead of real variables $x_1, \ldots, x_N$, are inserted in the argument, the multivariate splines $\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$, $\mathbf{i} \in \mathcal{I}_{\mathbf{n}}$, become functions of random input variables. Therefore, it is important to establish their second-moment properties, to be exploited in section 6.

Proposition 5.3. *Let $\mathbf{X} := (X_1, \ldots, X_N)^\intercal : (\Omega, \mathcal{F}) \to (\mathbb{A}^N, \mathcal{B}^N)$ be a vector of $N \in \mathbb{N}$ input random variables fulfilling Assumption 2.1. Then the first- and second-order moments of multivariate orthonormal B-splines $\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$, $\mathbf{i}, \mathbf{j} \in \mathcal{I}_{\mathbf{n}}$, are*

$$\mathbb{E}\left[\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] = \begin{cases} 1, & \mathbf{i} = \mathbf{1} := (1, \ldots, 1), \\ 0, & \mathbf{i} \neq \mathbf{1}, \end{cases}$$

and

$$\mathbb{E}\left[\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\Psi_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] = \begin{cases} 1, & \mathbf{i} = \mathbf{j}, \\ 0, & \mathbf{i} \neq \mathbf{j}, \end{cases}$$

respectively.

The statistical properties of univariate orthonormal B-splines in Proposition 4.6, with statistical independence in mind, lead to the result of Proposition 5.3.

5.4. Orthonormal basis. The following proposition shows that the multivariate orthonormal splines from Definition 5.2 span the spline space of interest.

Proposition 5.4. *Let* $\mathbf{X} := (X_1, \ldots, X_N)^\intercal : (\Omega, \mathcal{F}) \to (\mathbb{A}^N, \mathcal{B}^N)$ *be a vector of* $N \in \mathbb{N}$ *input random variables fulfilling Assumption 2.1. Then* $\{\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x}) : \mathbf{i} \in \mathcal{I}_{\mathbf{n}}\}$*, the set of multivariate orthonormal B-splines for a chosen degree* $\mathbf{p}$ *and family of knot sequences* $\boldsymbol{\Xi}$*, consistent with the probability measure* $f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$*, is a basis of* $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$*. That is,*

$$\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}} = \operatorname{span}\left\{\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x})\right\}_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} = \bigotimes_{k=1}^{N} \operatorname{span}\left\{\psi^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k)\right\}_{i_k=1,\ldots,n_k}, \quad |\mathcal{I}_{\mathbf{n}}| = \prod_{k=1}^{N} n_k.$$

The statistical properties in Proposition 5.3 result in linear independence of the elements of $\{\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x})\}_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}}$. The desired result is obtained readily.

6. SCE. Given an input random vector $\mathbf{X} := (X_1, \ldots, X_N)^\intercal : (\Omega, \mathcal{F}) \to (\mathbb{A}^N, \mathcal{B}^N)$ with the probability density function $f_{\mathbf{X}}(\mathbf{x})$ on $\mathbb{A}^N \subset \mathbb{R}^N$, let $y(\mathbf{X}) := y(X_1, \ldots, X_N)$ be a real-valued, square-integrable, measurable transformation on $(\Omega, \mathcal{F})$. Here, $y : \mathbb{A}^N \to \mathbb{R}$ represents an output function from a mathematical model, describing relevant stochastic performance of a complex system. Associated with the image probability space $(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}}d\mathbf{x})$, define

$$L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}}d\mathbf{x}) := \left\{ y : \mathbb{A}^N \to \mathbb{R} : \int_{\mathbb{A}^N} |y(\mathbf{x})|^2 f_{\mathbf{X}}(\mathbf{x})d\mathbf{x} < \infty \right\}$$

to be a weighted L^2-space of interest. Clearly, $L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}}d\mathbf{x})$ is a Hilbert space, which is endowed with the inner product

$$(y(\mathbf{x}), z(\mathbf{x}))_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})} := \int_{\mathbb{A}^N} y(\mathbf{x})z(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$$

and induced norm

$$\|y(\mathbf{x})\|_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})} = \sqrt{(y(\mathbf{x}), y(\mathbf{x}))_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})}}.$$

Similarly, for the abstract probability space $(\Omega, \mathcal{F}, \mathbb{P})$, there is an isomorphic Hilbert space

$$L^2(\Omega, \mathcal{F}, \mathbb{P}) := \left\{ Y : \Omega \to \mathbb{R} : \int_{\Omega} |y(\mathbf{X}(\omega))|^2 \, d\mathbb{P}(\omega) < \infty \right\}$$

of equivalent classes of output random variables $Y = y(\mathbf{X})$ with the corresponding inner product

$$(y(\mathbf{X}), z(\mathbf{X}))_{L^2(\Omega,\mathcal{F},\mathbb{P})} := \int_\Omega y(\mathbf{X}(\omega))z(\mathbf{X}(\omega))d\mathbb{P}(\omega)$$

and norm

$$\|y(\mathbf{X})\|_{L^2(\Omega,\mathcal{F},\mathbb{P})} := \sqrt{(y(\mathbf{X}), y(\mathbf{X}))_{L^2(\Omega,\mathcal{F},\mathbb{P})}}.$$

It is elementary to show that $y(\mathbf{X}(\omega)) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ if and only if $y(\mathbf{x}) \in L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}}d\mathbf{x})$.

6.1. SCE approximation. An SCE approximation of a square-integrable random variable $y(\mathbf{X}) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ is simply its orthogonal projection onto the spline space $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$, formally presented as follows.

Theorem 6.1. *Let $\mathbf{X} := (X_1, \ldots, X_N)^\intercal : (\Omega, \mathcal{F}) \to (\mathbb{A}^N, \mathcal{B}^N)$ be a vector of $N \in \mathbb{N}$ input random variables fulfilling Assumption 2.1. Given a degree $\mathbf{p}$ and a family of knot sequences $\boldsymbol{\Xi}$, recall that $\{\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X}) : \mathbf{i} \in \mathcal{I}_\mathbf{n}\}$ represents the set comprising multivariate orthonormal B-splines that is consistent with the probability measure $f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$. Then, for any random variable $y(\mathbf{X}) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, there exists an orthogonal expansion in multivariate orthonormal splines in $\mathbf{X}$, referred to as an SCE approximation*

$$y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X}) := \sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X}) \tag{6.1}$$

of $y(\mathbf{X})$, where the SCE expansion coefficients $C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}} \in \mathbb{R}$, $\mathbf{i} \in \mathcal{I}_\mathbf{n}$, are defined as

$$C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}} := \mathbb{E}\left[y(\mathbf{X})\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] := \int_{\mathbb{A}^N} y(\mathbf{x})\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}, \quad \mathbf{i} \in \mathcal{I}_\mathbf{n}. \tag{6.2}$$

Furthermore, the SCE approximation is the best approximation of $y(\mathbf{X})$ in the sense that

$$\mathbb{E}\left[y(\mathbf{X}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right]^2 = \inf_{g\in\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}} \mathbb{E}\left[y(\mathbf{X}) - g(\mathbf{X})\right]^2$$

or, equivalently,

$$\|y(\mathbf{x}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x})\|_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})} = \inf_{g\in\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}} \|y(\mathbf{x}) - g(\mathbf{x})\|_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})}.$$

Proof. Consider an arbitrary function $y(\mathbf{x}) \in L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}}d\mathbf{x})$. Then an orthogonal projection operator $P_{\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}} : L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}}d\mathbf{x}) \to \mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$, defined by

$$P_{\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}}y := \sum_{\mathbf{i}\in\mathcal{I}_\mathbf{n}} C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x}), \tag{6.3}$$

can be constructed. By definition of the random vector $\mathbf{X}$, the sequence $\{\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\}_{\mathbf{i}\in\mathcal{I}_\mathbf{n}}$ is a basis of the spline subspace $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$ of $L^2(\Omega, \mathcal{F}, \mathbb{P})$, inheriting the properties of the basis $\{\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x})\}_{\mathbf{i}\in\mathcal{I}_\mathbf{n}}$ of the spline subspace $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$ of $L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}}d\mathbf{x})$.[2] Therefore, (6.3) leads to the expansion in (6.1).

[2]With a certain abuse of notation, $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$ is used here as a set of spline functions of both real variables ($\mathbf{x}$) and random variables ($\mathbf{X}$).

For deriving the expression of the expansion coefficients, define a second moment

$$e_{\mathrm{SCE}} := \mathbb{E}\left[y(\mathbf{X}) - \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right]^2 \tag{6.4}$$

of the difference between $y(\mathbf{X})$ and its SCE approximation. Differentiate both sides of (6.4) with respect to $C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}$, $\mathbf{i}\in\mathcal{I}_{\mathbf{n}}$, to write

$$\begin{aligned}
\frac{\partial e_{\mathrm{SCE}}}{\partial C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}} &= \frac{\partial}{\partial C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}}\mathbb{E}\left[y(\mathbf{X}) - \sum_{\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} C_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}\Psi_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right]^2 \\
&= \mathbb{E}\left[\frac{\partial}{\partial C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}}\left\{y(\mathbf{X}) - \sum_{\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} C_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}\Psi_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right\}^2\right] \\
&= 2\mathbb{E}\left[\left\{\sum_{\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} C_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}\Psi_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X}) - y(\mathbf{X})\right\}\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] \\
&= 2\left\{\sum_{\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} C_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}\mathbb{E}\left[\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\Psi_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] - \mathbb{E}\left[y(\mathbf{X})\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right]\right\} \\
&= 2\left\{C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}} - \mathbb{E}\left[y(\mathbf{X})\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right]\right\}.
\end{aligned} \tag{6.5}$$

Here, the second, third, fourth, and last lines are obtained by interchanging the differential and expectation operators, performing the differentiation, swapping the expectation and summation operators, and applying Proposition 5.3, respectively. Setting $\partial e_{\mathrm{SCE}}/\partial C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}} = 0$ in (6.5) produces the desired result in (6.2).

Any spline function $g \in \mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$ can be expressed by

$$g(\mathbf{X}) = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \bar{C}_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X}) \tag{6.6}$$

with some real-valued coefficients $\bar{C}_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}$, $\mathbf{i}\in\mathcal{I}_{\mathbf{n}}$. To minimize $\mathbb{E}[\{y(\mathbf{X}) - g(\mathbf{X})\}^2]$, its derivatives with respect to the coefficients must be *zero*, that is,

$$\frac{\partial}{\partial \bar{C}_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}}\mathbb{E}\left[\{y(\mathbf{X}) - g(\mathbf{X})\}^2\right] = \frac{\partial}{\partial \bar{C}_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}}\mathbb{E}\left[\left\{y(\mathbf{X}) - \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \bar{C}_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right\}^2\right] = 0, \quad \mathbf{i}\in\mathcal{I}_{\mathbf{n}}.$$

From (6.5) and the following text, the derivatives are *zero* only when the coefficients $\bar{C}_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}$, $\mathbf{i}\in\mathcal{I}_{\mathbf{n}}$, match the expansion coefficients defined in (6.2). Therefore, the SCE approximation is the best one, as claimed. ■

Proposition 6.2. *For any $y(\mathbf{X}) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, let $y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ be the SCE approximation associated with a chosen degree $\mathbf{p}$ and family of knot sequences $\boldsymbol{\Xi}$. Then the truncation error $y(\mathbf{X}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ is orthogonal to the subspace $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}} \subset L^2(\Omega, \mathcal{F}, \mathbb{P})$.*

Proof. Let g described in (6.6), with arbitrary coefficients $\bar{C}_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}$, $\mathbf{i} \in \mathcal{I}_{\mathbf{n}}$, be an arbitrary element of $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$. Then

$$\begin{aligned}
& \mathbb{E}\left[\{y(\mathbf{X}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\}\, g(\mathbf{X})\right] \\
= \; & \mathbb{E}\left[\left\{y(\mathbf{X}) - \sum_{\mathbf{j}\in\mathcal{I}_{\mathbf{n}}} C_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}\Psi_{\mathbf{j},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right\} \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} \bar{C}_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] \\
= \; & \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}\bar{C}_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}} - \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}\bar{C}_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}} \\
= \; & 0,
\end{aligned}$$

where the third line follows from (6.2) and Proposition 5.3. Hence, the proposition is proved. ∎

Proposition 6.3. *The projection operator $P_{\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}} : L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}}d\mathbf{x}) \to \mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$ is a linear, bounded operator.*

Proof. The operator $P_{\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}}$ is obviously linear. To prove its boundedness, use Proposition 6.2 and then invoke the Pythagoras theorem, yielding

$$\mathbb{E}[\{y(\mathbf{X}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\}^2] + \mathbb{E}[y^2_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})] = \mathbb{E}[y^2(\mathbf{X})].$$

Therefore,

$$\mathbb{E}[y^2_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})] \le \mathbb{E}[y^2(\mathbf{X})]$$

for any $y(\mathbf{X}) \in L^2(\Omega, \mathcal{F}, \mathbb{P}$. This is equivalent to the assertion that

$$\left\|P_{\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}} y(\mathbf{x})\right\|_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})} \le \|y(\mathbf{x})\|_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})}$$

for any $y(\mathbf{x}) \in L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}}d\mathbf{x})$. ∎

From the general properties of orthogonal projection, the proofs of Theorem 6.1 and Propositions 6.2 and 6.3 are straightforward and may be deemed unnecessary to the eye of an expert reader. Nonetheless, they are documented here for the paper to be self-contained.

6.2. Approximation quality and convergence. A preferred approach among approximation theorists to measure the quality of approximations by polynomials and splines involves the modulus of smoothness [4, 17, 20]. Formal definitions of the modulus of smoothness in each coordinate direction k, followed by a tensorized version, are presented as follows.

Definition 6.4 (Schumaker [17]). *Given a positive integer $\alpha_k \in \mathbb{N}$ and $0 < h_k \le (b_k - a_k)/\alpha_k$, the α_kth modulus of smoothness of a function $y(x_k) \in L^2[a_k, b_k]$ in the L^2-norm is a function defined by*

$$\omega_{\alpha_k}(y; h_k)_{L^2[a_k,b_k]} := \sup_{0\le u_k\le h_k} \left\|\Delta^{\alpha_k}_{u_k} y(x_k)\right\|_{L^2[a_k,b_k-\alpha_k u_k]}, \quad h_k > 0,$$

where

$$\Delta_{u_k}^{\alpha_k} y(x_k) := \sum_{i=0}^{\alpha_k} (-1)^{\alpha_k - i} \binom{\alpha_k}{i} y(x_k + i u_k)$$

is the α_kth forward difference of y at x_k for any $0 \le u_k \le h_k$.

Moreover, given a multi-index $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N) \in \mathbb{N}^N$ and any vector $\mathbf{u} \ge \mathbf{0}$, let

$$\Delta_{\mathbf{u}}^{\boldsymbol{\alpha}} = \prod_{k=1}^{N} \Delta_{u_k}^{\alpha_k}.$$

Then the $\boldsymbol{\alpha}$-modulus of smoothness of a function $y(\mathbf{x}) \in L^2(\mathbb{A}^N)$ in the L^2-norm is the function defined by

$$\omega_{\boldsymbol{\alpha}}(y; \mathbf{h})_{L^2(\mathbb{A}^N)} := \sup_{\mathbf{0} \le \mathbf{u} \le \mathbf{h}} \|\Delta_{\mathbf{u}}^{\boldsymbol{\alpha}} y(\mathbf{x})\|_{L^2(\mathbb{A}_{\boldsymbol{\alpha},\mathbf{u}}^N)}, \quad \mathbf{h} > \mathbf{0},$$

where

$$\mathbb{A}_{\boldsymbol{\alpha},\mathbf{u}}^N = \left\{ \mathbf{x} \in \mathbb{A}^N : \mathbf{x} + \boldsymbol{\alpha} \otimes \mathbf{u} \in \mathbb{A}^N \right\}, \quad \boldsymbol{\alpha} \otimes \mathbf{u} = (\alpha_1 u_1, \ldots, \alpha_N u_N).$$

The book by Schumaker [17] provides a slightly general definition of the modulus of smoothness for $y \in L^q[a_k, b_k]$ (Chapter 2) or $y \in L^q(\mathbb{A}^N)$ (Chapter 13), $1 \le q < \infty$, including a summary of their elementary properties.

From Definition 6.4, as h_k approaches *zero*, so does $0 \le u_k \le h_k$. Taking the limit $u_k \to 0$ inside the integral of the L^2 norm, which is permissible for a finite interval and uniformly convergent integrand, the forward difference

$$\lim_{u_k \to 0} \Delta_{u_k}^{\alpha_k} y(x_k) = y(x_k) \sum_{i=0}^{\alpha_k} (-1)^{\alpha_k - i} \binom{\alpha_k}{i} = 0$$

as the sum vanishes for any $\alpha_k \in \mathbb{N}$. Consequently, the coordinate modulus of smoothness

$$\omega_{\alpha_k}(y; h_k)_{L^2[a_k, b_k]} \to 0 \quad \text{as } h_k \to 0 \quad \forall \alpha_k \in \mathbb{N}.$$

Following similar considerations, the tensor modulus of smoothness

$$\omega_{\boldsymbol{\alpha}}(y; \mathbf{h})_{L^2(\mathbb{A}^N)} \to 0 \quad \text{as } \mathbf{h} \to \mathbf{0} \quad \forall \boldsymbol{\alpha} \in \mathbb{N}^N.$$

These limits, in conjunction with Lemma 6.5, will be used to prove the L^2-convergence of the SCE approximations.

Lemma 6.5. *Let $L^2(\mathbb{A}^N)$ be an unweighted Hilbert space, defined as*

$$L^2\left(\mathbb{A}^N\right) := \left\{ y : \mathbb{A}^N \to \mathbb{R} : \int_{\mathbb{A}^N} |y(\mathbf{x})|^2 d\mathbf{x} < \infty \right\}$$

with standard norm $\|\cdot\|_{L^2(\mathbb{A}^N)}$. Then, for any function $y(\mathbf{x}) \in L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}} d\mathbf{x})$, it holds that

$$\|y(\mathbf{x})\|_{L^2(\mathbb{A}^N, \mathcal{B}^N, f_{\mathbf{X}} d\mathbf{x})} \le \sqrt{\|f_{\mathbf{X}}(\mathbf{x})\|_{L^\infty(\mathbb{A}^N)}} \, \|y(\mathbf{x})\|_{L^2(\mathbb{A}^N)},$$

where $\|\cdot\|_{L^\infty(\mathbb{A}^N)}$ is the infinity norm. Here, additionally, it is assumed that $f_{\mathbf{X}} \in L^\infty(\mathbb{A}^N)$.

Proof. From definition,

$$\begin{aligned} \|y(\mathbf{x})\|^2_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})} &:= \int_{\mathbb{A}^N} y^2(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) d\mathbf{x} \\ &= \left(y^2(\mathbf{x}), f_{\mathbf{X}}(\mathbf{x})\right)_{L^2(\mathbb{A}^N)} \\ &\leq \left\|y^2(\mathbf{x})\right\|_{L^1(\mathbb{A}^N)} \cdot \|f_{\mathbf{X}}(\mathbf{x})\|_{L^\infty(\mathbb{A}^N)} \\ &= \|y(\mathbf{x})\|^2_{L^2(\mathbb{A}^N)} \cdot \|f_{\mathbf{X}}(\mathbf{x})\|_{L^\infty(\mathbb{A}^N)}, \end{aligned} \tag{6.7}$$

where the third line stems from Hölder's inequality. As $\|f_{\mathbf{X}}(\mathbf{x})\|_{L^\infty(\mathbb{A}^N)}$ is positive, applying the square root on (6.7) yields the desired result. ■

Proposition 6.6. *For any $y(\mathbf{X}) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, a sequence of SCE approximations $\{y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\}_{\mathbf{h}>\mathbf{0}}$, with $\mathbf{h} = (h_1, \ldots, h_N)$ representing the vector of largest element sizes, converges to $y(\mathbf{X})$ in mean-square, that is,*

$$\lim_{\mathbf{h}\to\mathbf{0}} \mathbb{E}\left[|y(\mathbf{X}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})|^2\right] = 0.$$

Furthermore, the sequence of SCE approximations converges in probability, that is, for any $\epsilon > 0$,

$$\lim_{\mathbf{h}\to\mathbf{0}} \mathbb{P}\left(|y(\mathbf{X}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})| > \epsilon\right) = 0,$$

and converges in distribution, that is, for all points $\xi \in \mathbb{R}$ where $F(\xi)$ is continuous,

$$\lim_{\mathbf{h}\to\mathbf{0}} F_{\mathbf{p},\boldsymbol{\Xi}}(\xi) = F(\xi)$$

such that $F_{\mathbf{p},\boldsymbol{\Xi}}(\xi) := \mathbb{P}(y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X}) \leq \xi)$ and $F(\xi) := \mathbb{P}(y(\mathbf{X}) \leq \xi)$ are distribution functions of $y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ and $y(\mathbf{X})$, respectively. If $F(\xi)$ is continuous on $\mathbb{R}$, then the distribution functions converge uniformly.

Proof. From Lemma 6.5,

$$\|y(\mathbf{x}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x})\|_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})} \leq \sqrt{\|f_{\mathbf{X}}(\mathbf{x})\|_{L^\infty}}\, \|y(\mathbf{x}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x})\|_{L^2(\mathbb{A}^N)}. \tag{6.8}$$

Recognize from Proposition 6.3 that $P_{\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}}$ is a linear, bounded operator. Therefore, invoke Theorem 12.8 of Schumaker's book [17], which states that for a bounded linear operator, the unweighted L^2-error from the SCE approximation is bounded by

$$\|y(\mathbf{x}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x})\|_{L^2(\mathbb{A}^N)} \leq C' \omega_{\mathbf{p}+\mathbf{1}}(y; \mathbf{h})_{L^2(\mathbb{A}^N)}, \tag{6.9}$$

where C' is a constant that depends only on $\mathbf{p}$ and N, and $\mathbf{p} + \mathbf{1} = (p_1 + 1, \ldots, p_N + 1)$. Combining (6.8) and (6.9) produces

$$\|y(\mathbf{x}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x})\|_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})} \leq C \omega_{\mathbf{p}+\mathbf{1}}(y; \mathbf{h})_{L^2(\mathbb{A}^N)}, \tag{6.10}$$

where $C = C'\sqrt{\|f_{\mathbf{X}}(\mathbf{x})\|_{L^\infty}}$ is another constant, depending on $\mathbf{p}$, N, and now $f_{\mathbf{X}}(\mathbf{x})$.

Equation (6.10) gives a result on the L^2-distance of a function y to the spline space $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}}$. From the discussion related to Definition 6.4, the modulus of smoothness

$$\omega_{\mathbf{p}+\mathbf{1}}(y; \mathbf{h})_{L^2(\mathbb{A}^N)} \to 0 \quad \text{as } \mathbf{h} \to \mathbf{0} \quad \forall \mathbf{p} \in \mathbb{N}_0^N.$$

Therefore,

$$\lim_{\mathbf{h}\to\mathbf{0}} \|y(\mathbf{x}) - y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x})\|_{L^2(\mathbb{A}^N,\mathcal{B}^N,f_{\mathbf{X}}d\mathbf{x})} = 0,$$

thus proving the mean-square convergence of $y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ to $y(\mathbf{X})$ for any degree $\mathbf{p} \in \mathbb{N}_0^N$. In addition, as the SCE approximation converges in mean-square, it does so in probability. Moreover, as the expansion converges in probability, it also converges in distribution. ■

6.3. A special case of SCE. The well-known PCE approximation, especially its tensor-product version, can be derived from the SCE approximation proposed.

Proposition 6.7. *Given $k = 1, \ldots, N$, $0 \le p_k < \infty$, and an interval $[a_k, b_k] \subset \mathbb{R}$, let*

$$\boldsymbol{\xi}_k^{'} = \{\overbrace{a_k, \ldots, a_k}^{p_k+1 \text{ times}}, \overbrace{b_k, \ldots, b_k}^{p_k+1 \text{ times}}\} \tag{6.11}$$

be a $(p_k + 1)$-open knot sequence with no internal knots and $\boldsymbol{\Xi}^{'} = \{\boldsymbol{\xi}_1^{'}, \ldots, \boldsymbol{\xi}_N^{'}\}$. Then the resulting SCE approximation reduces to a PCE approximation.

For the knot sequence $\boldsymbol{\xi}_k^{'}$ in (6.11), the resulting B-splines are related to the well-known Bernstein polynomials of degree p_k. Since the set of Bernstein polynomials of degree p_k forms a basis of the polynomial space Π_{p_k}, the spline space $\mathcal{S}_{k,p_k,\boldsymbol{\xi}_k^{'}} = \Pi_{p_k}$. Then, going through the standard tensor-product construction, it is trivial to show that, indeed, the multivariate spline space $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}^{'}}$ is spanned by the set $\{\Psi_{\mathbf{i}}(\mathbf{x}) : \mathbf{0} \le \mathbf{i} \le \mathbf{p}\}$ of multivariate orthonormal polynomials in $\mathbf{x}$ that are consistent with the probability measure $f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}$. This results in a $\mathbf{p}$th-degree tensor-product PCE

$$y_{\mathbf{p},\boldsymbol{\Xi}^{'}}(\mathbf{X}) = y_{\mathbf{p}}(\mathbf{X}) := \sum_{\mathbf{0}\le\mathbf{i}\le\mathbf{p}} C_{\mathbf{i}}^{'}\Psi_{\mathbf{i}}(\mathbf{X}) \tag{6.12}$$

of $y(\mathbf{X}) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathbf{p} \in \mathbb{N}_0^N$ and

$$C_{\mathbf{i}}^{'} := \int_{\mathbb{A}^N} y(\mathbf{x})\Psi_{\mathbf{i}}(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})d\mathbf{x}, \quad \mathbf{i} \in \mathbb{N}_0^N,$$

are its expansion coefficients. Hence, a $\mathbf{p}$th-degree SCE approximation with no internal knots becomes identical to a $\mathbf{p}$th-degree PCE approximation.

As described in the preceding paragraph, there is no mesh in the deduction of the PCE approximation from the SCE approximation. Therefore, a refinement by reducing the element sizes is not possible. However, a refinement is still possible by degree elevation, that is, by increasing gradually the degree p_k in all coordinate directions. Indeed, when $p_k \to \infty$ for all $k = 1, \ldots, N$, the right side of (6.12) becomes the full PCE representation of $y(\mathbf{X})$.

6.4. Output statistics and other probabilistic characteristics. The SCE approximation $y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ can be viewed as a surrogate of $y(\mathbf{X})$. Therefore, relevant probabilistic characteristics of $y(\mathbf{X})$, including its first two moments and probability density function, if it exists, can be estimated from the statistical properties of $y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$.

Applying the expectation operator on $y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ in (6.1) and imposing Proposition 5.3, its mean is

$$\mathbb{E}\left[y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] = C_{\mathbf{1},\mathbf{p},\boldsymbol{\Xi}} = \mathbb{E}\left[y(\mathbf{X})\right], \ \mathbf{1} = (1,\ldots,1),$$

which is independent of $\mathbf{p}$ and $\boldsymbol{\Xi}$. More importantly, the SCE approximation always yields the exact mean.

Applying the expectation operator again, this time on $[y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X}) - C_{\mathbf{1},\mathbf{p},\boldsymbol{\Xi}}]^2$, and employing Proposition 5.3 one more time results in the variance

$$\operatorname{var}\left[y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] = \sum_{\mathbf{i}\in(\mathcal{I}_{\mathbf{n}}\setminus\{\mathbf{1}\})} C^2_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}} = \sum_{\mathbf{i}\in\mathcal{I}_{\mathbf{n}}} C^2_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}} - C^2_{\mathbf{1},\mathbf{p},\boldsymbol{\Xi}} \tag{6.13}$$

of $y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$. It is elementary to show that $\operatorname{var}[y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})] \le \operatorname{var}[y(\mathbf{X})]$.

The second-moment properties of an SCE approximation are solely determined by the expansion coefficients. The formulae for the mean and variance of the SCE approximation are same as those reported for the PCE approximation, although the respective expansion coefficients involved are not. The primary reason for this similarity stems from the use of orthonormal basis in both expansions.

Being convergent in probability and in distribution, the probability density function of $y(\mathbf{X})$, if it exists, can also be estimated by that of $y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$. However, deriving analytical formula for the density function is hopeless in general. Nonetheless, the density function can be estimated by Monte Carlo simulation (MCS) of the SCE approximation, that is, by resampling of $y_{\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ involving inexpensive evaluations of simple spline functions.

6.5. SCE as an infinite series. The set of orthonormal B-splines $\{\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}}(\mathbf{x}) : \mathbf{i} \in \mathcal{I}_{\mathbf{n}}\}$ from (6.1) has its size equal to $\prod_{k=1}^{N} n_k$. Therefore, the size is controlled by the number of basis functions n_k, which, in succession, is decided by the length of the knot sequence $\boldsymbol{\xi}_k$ and order p_k in each coordinate direction. Obviously, the longer the sequence $\boldsymbol{\xi}_k$, the larger the value of n_k and, hence, the size of the set. For a refinement process with a fixed p_k, consider increasing the length of $\boldsymbol{\xi}_k$ or n_k in all N coordinate directions in such a way that the largest element size h_k is monotonically reduced. The result is an increasing family of the sets of such basis functions. In the limit, when $n_k \to \infty$ or $h_k \to 0$, $k = 1,\ldots,N$, denote by $\boldsymbol{\xi}_{k,\infty}$ and $\boldsymbol{\Xi}_\infty = \{\boldsymbol{\xi}_{1,\infty},\ldots,\boldsymbol{\xi}_{N,\infty}\}$ the associated knot sequence in the kth coordinate direction and the family of such N knot sequences, respectively. Then there exists a set of infinite number of basis functions $\{\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}_\infty}(\mathbf{x}) : \mathbf{i} \in \mathbb{N}^N\}$ with the index set of knot indices

$$\{\mathbf{i} = (i_1,\ldots,i_N) : 1 \le i_k < \infty, \ k = 1,\ldots,N\} = \mathbb{N}^N.$$

In consequence, $\{\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}_\infty}(\mathbf{x}) : \mathbf{i} \in \mathbb{N}^N\}$ forms an orthogonal basis of $\mathcal{S}_{\mathbf{p},\boldsymbol{\Xi}_\infty}$, yielding

$$L^2(\Omega,\mathcal{F},\mathbb{P}) = \overline{\operatorname{span}\{\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}_\infty}(\mathbf{X})\}_{\mathbf{i}\in\mathbb{N}^N}},$$

where the overline stands for set closure. Hence, every $y(\mathbf{X}) \in L^2(\Omega,\mathcal{F},\mathbb{P})$ can be expanded in terms of the aforementioned spanning set, resulting in an infinite series

$$y(\mathbf{X}) \sim \sum_{\mathbf{i}\in\mathbb{N}^N} C_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}_\infty}\Psi_{\mathbf{i},\mathbf{p},\boldsymbol{\Xi}_\infty}(\mathbf{X}), \tag{6.14}$$

which is referred to as SCE in the paper. Here the symbol $\sim$ represents equality in the mean-square sense. From Proposition 6.6, the partial sums of (6.14) converge to $y(\mathbf{X})$ in L^2. Therefore, the infinite series is mean-square convergent to the correct limit.

Comparing (6.1) and (6.14), the former is a truncated version of the latter. Therefore, the designation "SCE" employed in the paper makes sense even though the SCE approximation in (6.1) represents a finite sum or expansion.

7. Numerical experiments. Three examples describing one-, two-, and four-dimensional UQ problems, where the output function is explicitly defined or obtained from the solution of an ordinary differential equation (ODE), are presented. The random input $\mathbf{X}$ fulfills Assumption 2.1, and the output function $y(\mathbf{X})$ is in $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Therefore, SCE and PCE approximations can be applied to estimate their second-moment statistics and probability distributions. The objectives are to evaluate the approximation power of the SCE approximation in terms of the second-moment statistics or probability distributions of $y(\mathbf{X})$ and contrast the SCE results with those obtained from the existing PCE approximation.

The coordinate degrees for SCE or PCE approximations in the second and third examples are identical, that is, $p_1 = p_2 = p_3 = p_4 = p$ (say). So are the knot sequences for SCE, that is, $\boldsymbol{\xi}_1 = \boldsymbol{\xi}_2 = \boldsymbol{\xi}_3 = \boldsymbol{\xi}_4 = \boldsymbol{\xi}$ (say) with a uniform mesh of element sizes $h_1 = h_2 = h_3 = h_4 = h$. In all three examples, the degree p and/or element size h were varied as desired. The basis for a pth-degree PCE approximation was obtained from an appropriate set of Legendre orthonormal polynomials in input variables, whereas the basis for an SCE approximation, given a degree p and a knot sequence of element size h, was generated from the Cholesky factorization of the spline moment matrix. From the uniform distribution, the spline moment matrix was constructed analytically. All knot sequences are $(p+1)$-open and consist of uniformly spaced distinct knots with even and/or odd numbers of elements, depending on the example. The PCE and SCE coefficients, which are one-, two-, and four-dimensional integrals, were calculated exactly.

Define, for the first two examples, two approximation errors in the variances,

$$e_{p,h} := \frac{|\mathrm{var}[y(\mathbf{X})] - \mathrm{var}[y_{p,h}(\mathbf{X})]|}{\mathrm{var}[y(\mathbf{X})]} \quad \text{and} \quad e_p := \frac{|\mathrm{var}[y(\mathbf{X})] - \mathrm{var}[y_p(\mathbf{X})]|}{\mathrm{var}[y(\mathbf{X})]},$$

committed by the SCE approximation $y_{p,h}(\mathbf{X}) := y_{p,\boldsymbol{\xi}}(X)$ or $y_{(p,p),\{\boldsymbol{\xi},\boldsymbol{\xi}\}}(X_1, X_2)$ and the PCE approximation $y_p(\mathbf{X}) := y_p(X)$ or $y_{(p,p)}(X_1, X_2)$, respectively, of $y(\mathbf{X})$. The exact variance $\mathrm{var}[y(\mathbf{X})]$ was obtained analytically, whereas the SCE variance $\mathrm{var}[y_{p,h}(\mathbf{X})]$ and PCE variance $\mathrm{var}[y_p(\mathbf{X})]$ were also determined analytically from (6.13) and similar formula, respectively. Therefore, all approximation errors were calculated exactly.

7.1. Example 1: Three univariate functions. Consider a family of three functions of a real-valued, uniformly distributed random variable X over $[-1, 1]$:

$$y(X) = \begin{cases} \sin(3\pi X) & \text{(smooth, oscillatory)}, \\ \exp(-3|X|) & \text{(nonsmooth)}, \\ \Phi(20X) & \text{(nearly discontinuous)}. \end{cases} \tag{7.1}$$

Here, $\Phi(u) = (1/\sqrt{2\pi}) \int_{-\infty}^{u} \exp(-\xi^2/2) d\xi$ is the cumulative probability distribution function of a Gaussian random variable with *zero* mean and *unit* variance. From top to bottom, (7.1)

comprises oscillatory yet smooth, nondifferentiable, and nearly discontinuous functions that are progressively more difficult to approximate by polynomials.

The knot sequences for the oscillatory function include simple knots and consist of even numbers of elements with varying element sizes: $h = 1/2, 1/4, 1/8, 1/12, 1/16$. For the nonsmooth and nearly discontinuous functions, however, the knot sequences comprise both even and odd numbers of elements, producing the following element sizes: $h = 2/5, 2/9, 2/17, 2/25, 2/33$ for odd numbers of elements and $h = 1/2, 1/4, 1/8, 1/12, 1/16$ for even numbers of elements. The odd numbers of elements are relevant when the location of the point where the function is nondifferentiable or nearly discontinuous is unknown. However, if the aforementioned point is known, then it is possible to employ even numbers of elements by deploying knot(s) at that point as well. In the latter case, double knots (multiplicity of two for $p = 2$) were placed for the nondifferentiable function, whereas a single knot was assigned for the nearly discontinuous function.

Figures 2, 3, and 4 depict the comparisons of PCE and SCE approximations for the oscillatory, nonsmooth, and nearly discontinuous functions, respectively. For the oscillatory

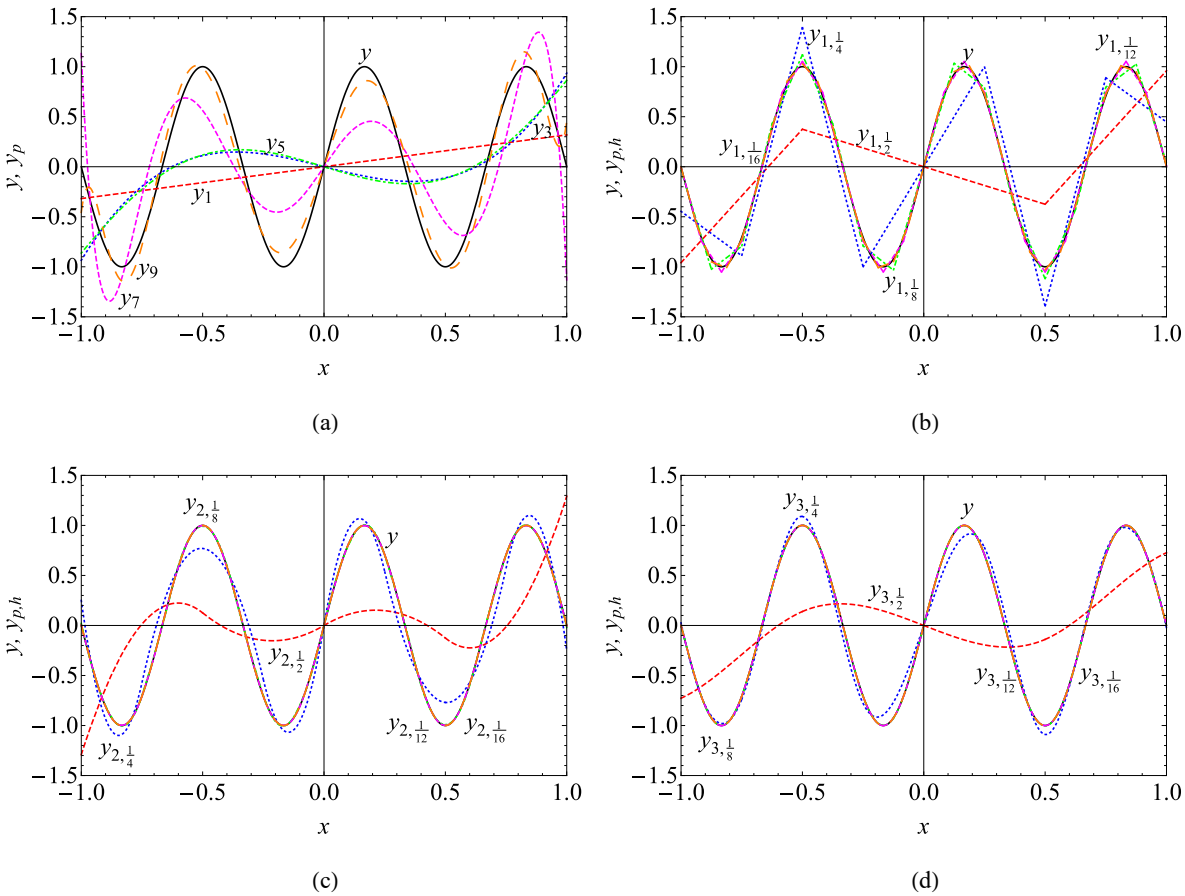


Figure 2. *Oscillatory function:* $y(x) = \sin(3\pi x)$*;* (a) *PCE approximations for* $p = 1, 3, 5, 7, 9$*;* (b) *linear SCE approximations for* $h = 1/2, 1/4, 1/8, 1/12, 1/16$*;* (c) *quadratic SCE approximations for* $h = 1/2, 1/4, 1/8, 1/12, 1/16$*;* (d) *cubic SCE approximations for* $h = 1/2, 1/4, 1/8, 1/12, 1/16$.

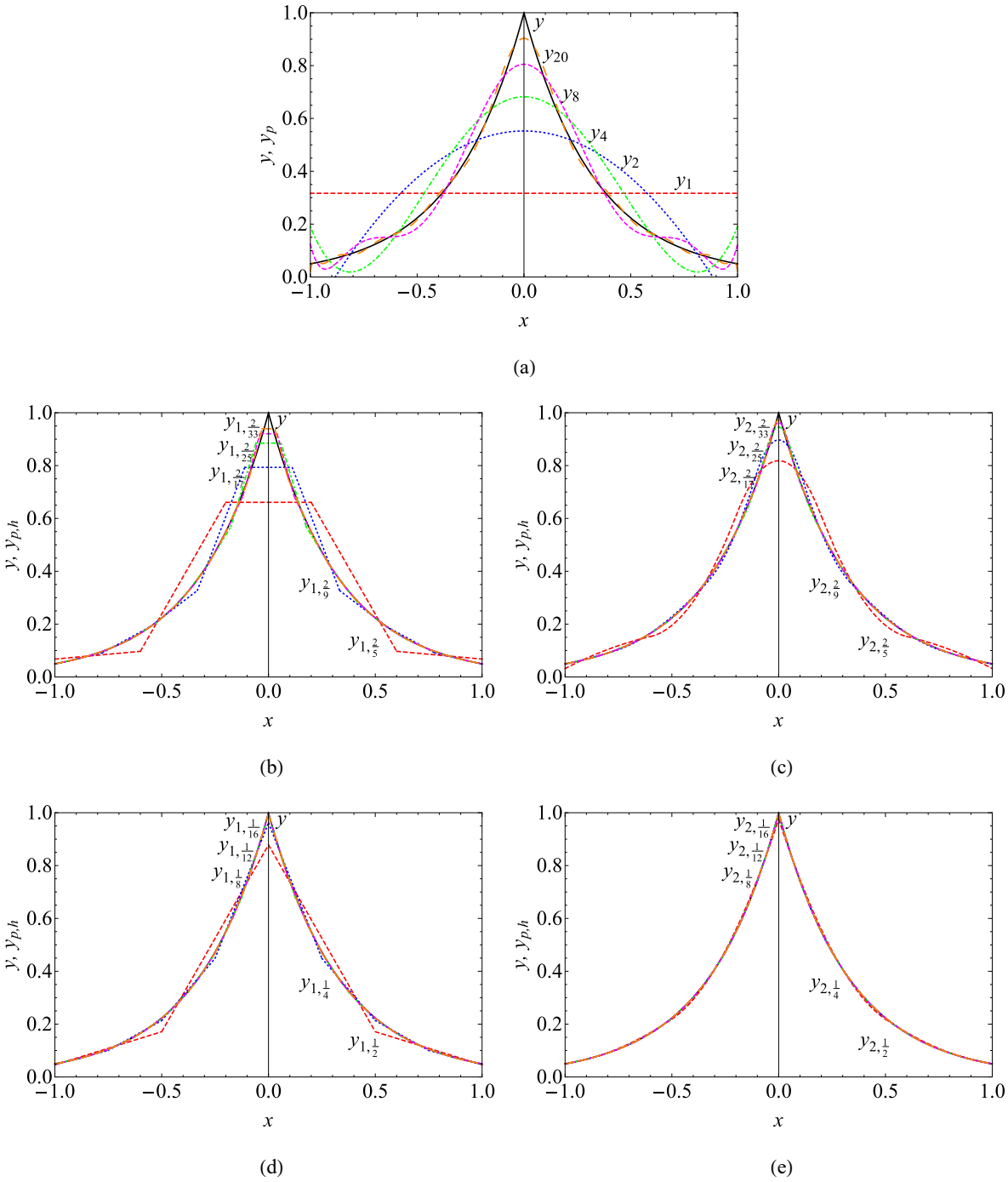


Figure 3. *Nonsmooth function:* $y(x) = \exp(-3|x|)$*;* (a) *PCE approximations for* $p = 1, 2, 4, 8, 20$*;* (b) *linear SCE approximations for* $h = 2/5, 2/9, 2/17, 2/25, 2/33$ *(odd);* (c) *quadratic SCE approximations for* $h = 2/5, 2/9, 2/17, 2/25, 2/33$ *(odd);* (d) *linear SCE approximations for* $h = 1/2, 1/4, 1/8, 1/12, 1/16$ *(even);* (e) *quadratic SCE approximations for* $h = 1/2, 1/4, 1/8, 1/12, 1/16$ *(even).*

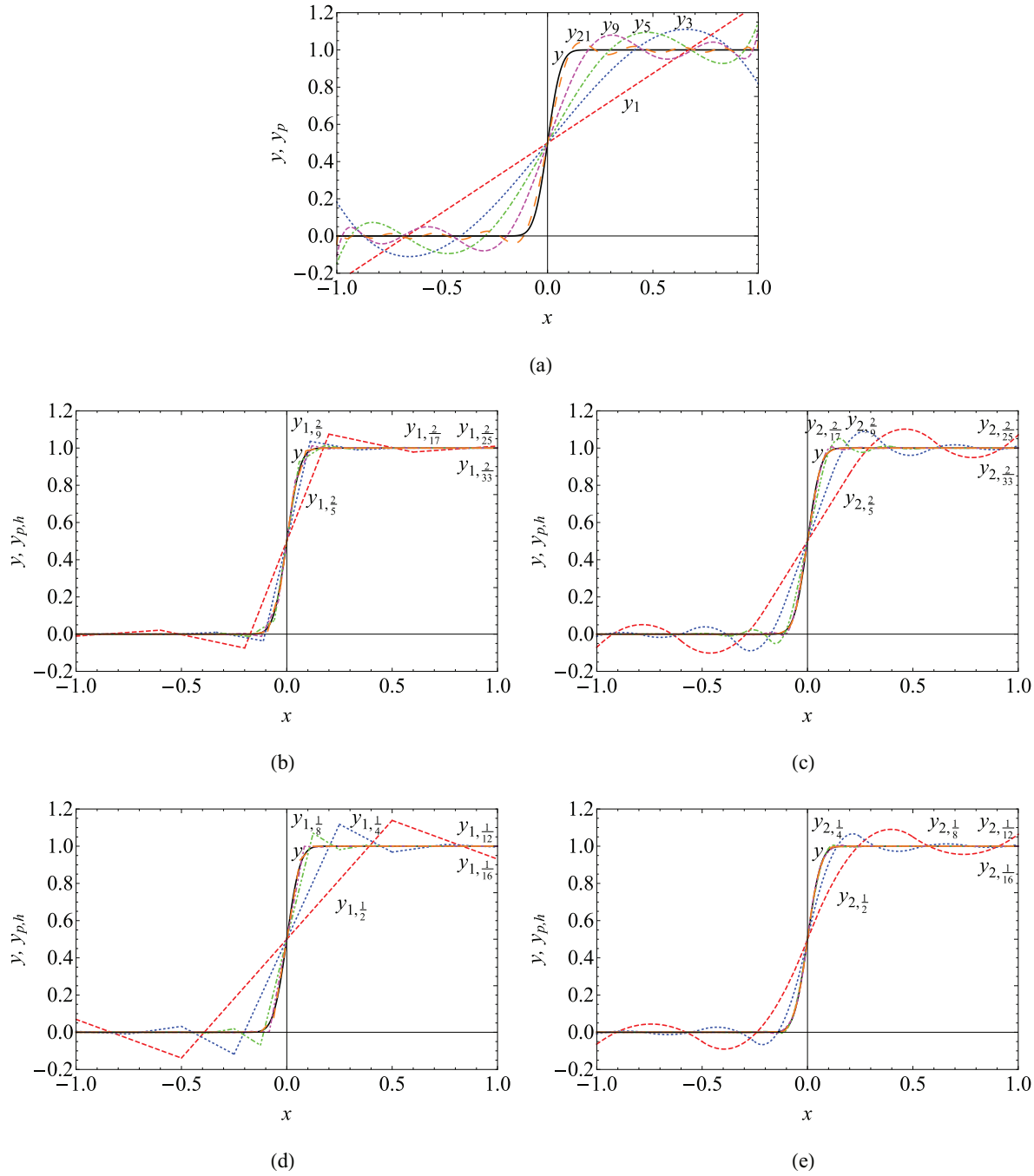


Figure 4. *Nearly discontinuous function:* $y(x) = \Phi(20x)$*;* (a) *PCE approximations for* $p = 1, 3, 5, 9, 21$*;* (b) *linear SCE approximations for* $h = 2/5, 2/9, 2/17, 2/25, 2/33$ *(odd);* (c) *quadratic SCE approximations for* $h = 2/5, 2/9, 2/17, 2/25, 2/33$ *(odd);* (d) *linear SCE approximations for* $h = 1/2, 1/4, 1/8, 1/12, 1/16$ *(even);* (e) *quadratic SCE approximations for* $h = 1/2, 1/4, 1/8, 1/12, 1/16$ *(even).*

function, the PCE approximation improves with p as shown in Figure 2(a), but at the cost of the 9th-degree approximation to be fairly acceptable. Such requirement becomes stringent for the nonsmooth (Figure 3(a)) or nearly discontinuous (Figure 4(a)) functions, where 20th- or 21st-degree PCE approximations are warranted. In contrast, the SCE approximations for the oscillatory function, exhibited in Figure 2(b), look satisfactory, if not great, even for a linear spline ($p = 1$), as long as the mesh is adequately fine ($h \leq 1/8$). For $p = 2$ or 3 and $h \leq 1/8$, any distinction between an SCE approximation and actual function in Figure 2(c) or Figure 2(d) is indiscernible to the naked eye.

For the nonsmooth and nearly discontinuous functions, there are two sets of linear ($p = 1$) and quadratic ($p = 2$) SCE approximations, obtained separately for odd and even numbers of elements; they are displayed in Figures 3 and 4. According to Figures 3(b) and 3(d), the approximation quality of linear SCE approximations for the nonsmooth function is visibly better when there are even numbers of elements, as expected. The same observation holds for quadratic SCE approximations, where even a much coarser mesh produces excellent approximation for even numbers of elements. The SCE results for the nearly discontinuous function are qualitatively the same. However, there are still some oscillations in SCE approximations when the mesh is too coarse, pointing to the Gibbs' type phenomenon commonly observed in polynomial-based approximations. Zhang and Martin [24] reported such behavior for a cubic spline approximation of the Heaviside function and found that the oscillation near discontinuity never goes away for a uniform knot sequence. Clearly, a better, if not optimal, selection of knot sequences is required.

Finally, Table 1 presents the errors $e_{p,h}$ and e_p in the variances of all three functions, obtained using SCE and PCE approximations, respectively, for various chosen degrees and knot sequences. Clearly, the SCE approximation commits much lower errors than does the PCE approximation for the same degree p. To attain an accurate approximation using splines, one is not interested in large values of p. Instead, the motivation is to keep p fixed to a low value, but increase (decrease) the number of knots (element size). Indeed, Table 1 demonstrates that a low-degree SCE approximation with an adequate mesh is capable of producing significantly more accurate estimates of the variance than the PCE approximation even when its degree of expansion is excessively large. All approximations errors reported in Table 1 are consistent with the plots displayed in Figures 2 through 4.

7.2. Example 2: Solution of a stochastic ODE. The second example involves a stochastic boundary-value problem, described by the ODE

$$-\frac{d}{d\xi}\left(\exp(|X_1|)\frac{d}{d\xi}y(\xi;X_1,X_2)\right) = \exp(|X_2|),\ 0 \leq \xi \leq 1,\ y(\xi;X_1,X_2) \in \mathbb{R}, \tag{7.2}$$

with boundary conditions

$$y(0;X_1,X_2) = 0,\ \ \exp(|X_1|)\frac{dy}{d\xi}(1;X_1,X_2) = 1.$$

Here, X_1 and X_2 are two real-valued, independent, and identically distributed random variables, each following a uniform distribution over $[-1, 1]$. Originally studied by the author [14], the ODE is slightly modified here by introducing the absolute-value function, thus producing a nonsmooth solution.

Table 1

Relative errors in the variances of three univariate functions by PCE and SCE approximations.

(a) smooth, oscillatory function: $y(X) = \sin(3\pi X)$

			$e_{p,h}$		
p	e_p	h	$p=1$	$p=2$	$p=3$
1	0.932453	1/2	0.719505	0.755454	0.815119
3	0.823578	1/4	0.0936391	0.0404491	0.0107727
5	0.822617	1/8	3.56392×10^{-3}	1.5349×10^{-4}	8.20356×10^{-6}
7	0.292768	1/12	6.05231×10^{-4}	9.8752×10^{-6}	1.95148×10^{-7}
9	0.0279709	1/16	1.80817×10^{-4}	1.56899×10^{-6}	1.6×10^{-8}

(b) nonsmooth function: $y(X) = \exp(-3|X|)$

			$e_{p,h}$ (odd no. of elements)			$e_{p,h}$ (even no. of elements)	
p	e_p	h	$p=1$	$p=2$	h	$p=1$	$p=2$
1	1	2/5	0.122349	0.0212933	1/2	0.0167023	4.96606×10^{-4}
2	0.325922	2/9	0.026662	3.74539×10^{-3}	1/4	1.13075×10^{-3}	9.37131×10^{-6}
4	0.124885	2/17	4.42555×10^{-3}	5.48633×10^{-4}	1/8	7.00943×10^{-5}	1.76989×10^{-7}
8	0.0301413	2/25	1.43968×10^{-3}	1.71708×10^{-4}	1/12	1.377×10^{-5}	1.68832×10^{-8}
20	0.0037569	2/33	6.36083×10^{-4}	7.45087×10^{-5}	1/16	4.34601×10^{-6}	3.14213×10^{-9}

(c) nearly discontinuous function: $y(X) = \Phi(20X)$

			$e_{p,h}$ (odd no. of elements)			$e_{p,h}$ (even no. of elements)	
p	e_p	h	$p=1$	$p=2$	h	$p=1$	$p=2$
1	0.209125	2/5	0.0198118	0.0574063	1/2	0.0983968	0.0308548
3	0.0966401	2/9	2.59428×10^{-3}	0.0184093	1/4	0.0299556	5.54174×10^{-3}
5	0.0543964	2/17	2.19365×10^{-4}	2.05094×10^{-3}	1/8	3.89483×10^{-3}	5.25409×10^{-5}
9	0.0212929	2/25	1.82967×10^{-4}	1.81128×10^{-4}	1/12	4.98215×10^{-4}	2.10786×10^{-5}
21	0.0017763	2/33	7.0312×10^{-5}	1.50297×10^{-5}	1/16	9.23004×10^{-5}	1.036×10^{-5}

A direct integration of (7.2) yields the exact solution:

$$y(\xi; X_1, X_2) = \frac{1}{\exp(|X_1|)} \left[\xi + \left(\xi - \frac{\xi^2}{2} \right) \exp(|X_2|) \right].$$

Therefore, the first two raw moments of $y(\xi; X_1, X_2)$ can be obtained easily. For instance, at $\xi = 1$, the two moments of $y(1; X_1, X_2)$, denoted briefly as $y(X_1, X_2)$, are

$$\mathbb{E}[y(X_1, X_2)] = \frac{1}{e} \left[1 + \frac{1}{2}(e-1) \right] (e-1) \approx 1.1752,$$

$$\mathbb{E}[y^2(X_1, X_2)] = \frac{1}{16e^2}(e^2 + 8e - 1)(e^2 - 1) \approx 1.52048.$$

The exact solutions were used to benchmark the approximate results from SCE and PCE approximations.

Figures 5 and 6 display three-dimensional (left) and contour (right) plots of the exact function $y(x_1, x_2)$ and several approximations from PCE and SCE. Because of the absolute-value function, the exact solution is saddle-shaped with slope discontinuities at the center, as shown in Figure 5(a). The second-order PCE approximation exhibited in Figure 5(b) commits a variance error of $e_2 = 0.116812$ and is clearly inadequate. The 16th-order PCE approximation in Figure 5(c) shows some improvement by reducing the error to $e_{16} = 2.26714 \times 10^{-3}$, but not to an extent expected from such an impractically high expansion order.

In contrast, the linear ($p = 1$) SCE approximation in Figure 6(a), obtained for an even number of elements with an element size of $h = 1/10$, matches the exact function extremely well, producing a variance error of $e_{1,1/10} = 1.86149 \times 10^{-6}$. The quadratic ($p = 2$) SCE approximation in Figure 6(b), generated using the same mesh, yields an error of $e_{2,1/10} = 3.54972 \times 10^{-4}$ and is better than the 16th-order PCE approximation yet inferior to that in Figure 6(a). This apparent anomaly of a linear SCE approximation producing a better result than a quadratic SCE approximation can be explained by examining the knot sequences used. Due to even numbers of elements, there exists a central knot in each coordinate direction for both cases of $p = 1$ and $p = 2$. However, for $p = 2$, the first-order derivatives are continuous across the central knot in both directions. This is why the quadratic SCE approximation is smoother than the linear SCE approximation or the exact function. However, as $y(x_1, x_2)$ is not differentiable at the central knot, the linear approximation performs better than the quadratic approximation. However, if the central knot is repeated (multiplicity of two) in the knot sequences, the quadratic SCE approximation, shown in Figure 6(c), is even better than the linear SCE approximation, resulting in an error of $e_{2,1/10} = 4.0056 \times 10^{-10}$. Having said so, such manipulations of the knot sequences are not possible in general if the locations of slope discontinuities are not known a priori. In this case, the quadratic SCE approximation in Figure 6(b) is perhaps more realistic, and the result of the linear SCE approximation should be deemed fortuitous for this specific problem.

7.3. Example 3: A nonsmooth function of four variables. In the final example, consider a nonsmooth function

$$y(\mathbf{X}) = \prod_{i=1}^{4} \frac{|4X_i - 2|^{b_i} + a_i}{1 + a_i}, \;\; a_i, b_i \in \mathbb{R}, \;\; i = 1, 4,$$

of four independent random variables X_i, $i = 1, 2, 3, 4$, each of which is uniformly distributed over $[0, 1]$. The function parameters are as follows: $a_1 = 0$, $a_2 = 1$, $a_3 = 2$, $a_4 = 4$; $b_1 = b_2 = b_3 = b_4 = 3/5$. Clearly, y is a nondifferentiable function where the exponent b_i controls its nonlinearity. Compared with $b_i = 1$, the smaller the value of the exponent, the more nonlinear the function becomes in the ith coordinate direction. This type of function, especially with *unit* exponents, has been used for global sensitivity analysis [15].

Figure 7 depicts the probability distribution functions of $y(\mathbf{X})$ calculated by three methods: (1) crude MCS; (2) second-, fourth-, and eighth-order PCE approximations; and (2) quadratic SCE approximations with three element sizes: $h = 1/2$, $h = 1/4$, and $h = 1/8$. In SCE calculations, there are even numbers of elements for the chosen meshes with repeated central knots ($x_k = 0.5$) in each coordinate direction. Although the basis functions and corresponding expansion coefficients of SCE and PCE approximations were calculated exactly, there is no

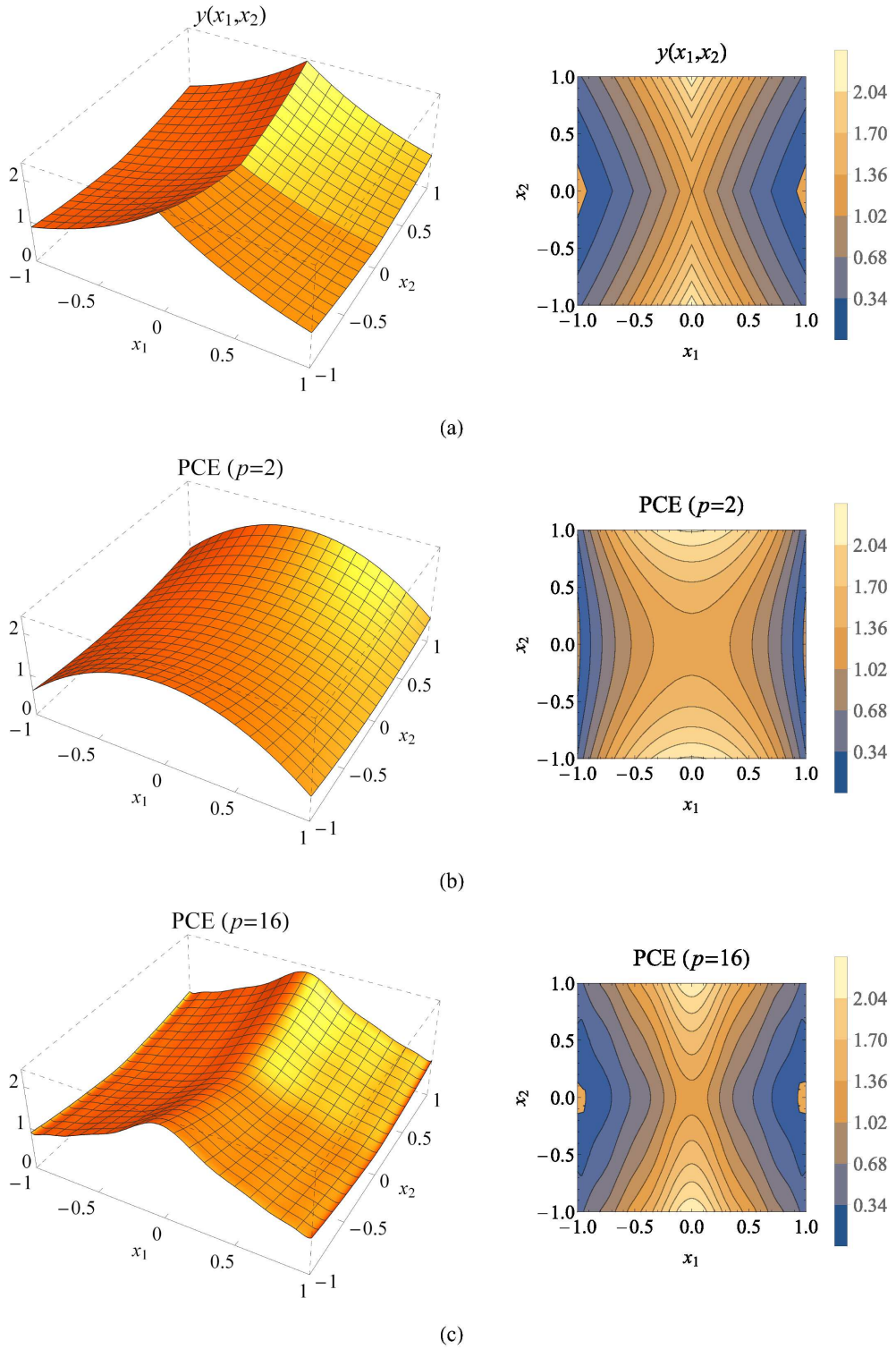


Figure 5. *Three-dimensional and contour plots of the exact and two PCE solutions of ODE;* (a) *exact solution* $y(x_1, x_2)$*;* (b) *second-order PCE approximation;* (c) 16*th-order PCE approximation.*

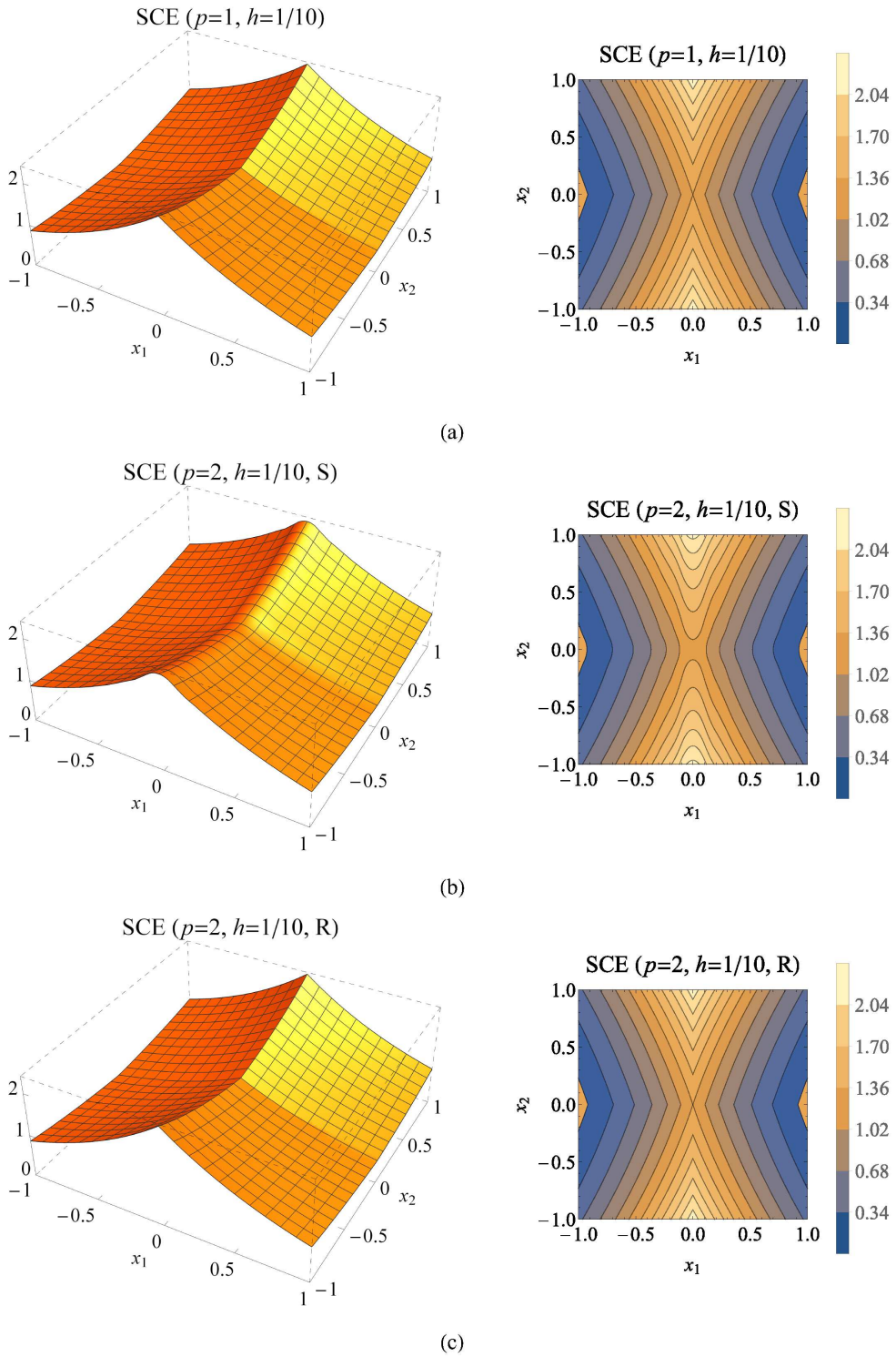


Figure 6. *Three-dimensional and contour plots of three SCE solutions of ODE;* (a) *linear SCE approximation for* $h = 1/10$*;* (b) *quadratic SCE approximation for* $h = 1/10$ *and simple ("S") knots;* (c) *quadratic SCE approximation for* $h = 1/10$ *and a repeated ("R") central knot.*

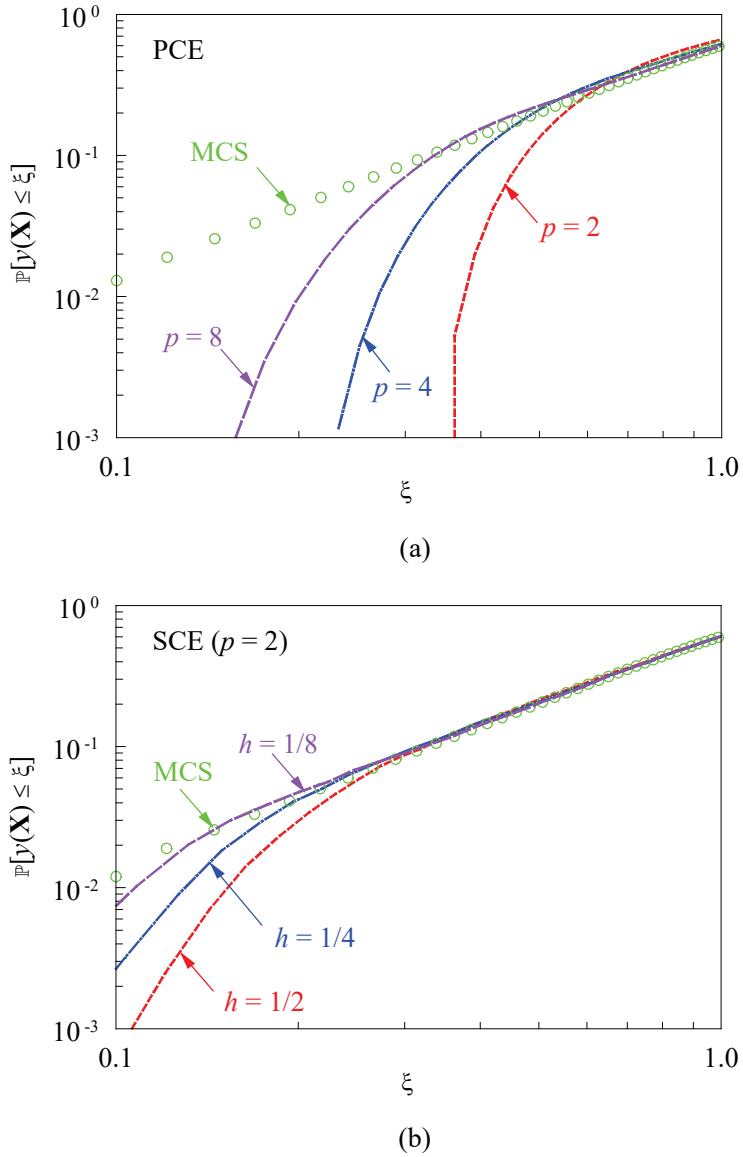


Figure 7. *Probability distributions of $y(\mathbf{X})$ calculated by three distinct methods;* (a) *crude MCS and several PCE approximations;* (b) *crude MCS and several SCE approximations.*

analytical means to determine their probability distributions. Instead, the PCE and SCE approximations once built were resampled to generate their associated distributions. The sample size for both crude MCS and resampling is 10,000, which should be adequate for examining the tail probabilistic characteristics up to a probability of 10^{-3}. Compared with the MCS result, the convergence of probability distributions by the SCE approximations in Figure 7(b) is markedly faster than that by the PCE approximations in Figure 7(a). It appears that low-order SCE approximations also yield more accurate estimates of the probability distributions than a high-order PCE approximation for nonsmooth functions.

8. Discussion. While the paper is aimed at fundamental mathematical development of SCE, a brief deliberation on the practical significance of the work is justified. First, the success of SCE is dependent on its effective implementation for UQ analysis of a general computational model. For more realistic problems not considered here, the expansion coefficients of SCE approximations cannot be calculated exactly. Therefore, computationally efficient methods or techniques for estimating the expansion coefficients are direly needed. In this regard, methods, such as dimension-reduction techniques [23] and sparse-grid quadrature [8], come to mind. Some of these methods, when appropriately adapted, may potentially aid in calculating the SCE coefficients economically.

Second, the SCE approximation proposed is designed to account for locally prominent and highly nonlinear stochastic responses, including discontinuity and nonsmoothness, emanating from multiple failure modes of complex systems. On the contrary, if the response is smooth and moderately nonlinear, then existing PCE equipped with globally supported basis is adequate. In the latter case, there is no significant advantage of SCE over PCE approximations.

Third, and more importantly, the use of tensor-product structure to form multivariate B-splines is not always suitable. Indeed, for high-dimensional UQ problems, tensor-product expansions in the context of SCE or PCE approximations will require an astronomically large number of terms or coefficients, succumbing to the curse of dimensionality. Therefore, developments of alternative computational methods capable of exploiting low effective dimensions of high-dimensional functions, à la dimensional decomposition methods [13], are desirable.

These topics are subjects of current research in the author's group.

9. Conclusion. A new chaos expansion, namely, SCE of a square-integrable random variable, comprising measure-consistent multivariate orthonormal B-splines in independent random variables, is unveiled. Under prescribed assumptions, a whitening transformation is proposed to decorrelate univariate B-splines in each coordinate direction into their orthonormal version. The transformed set of B-splines was proved to form a basis of a general spline space comprising splines of specified degree and knot sequence. Through a tensor-product structure, multivariate orthonormal B-splines were constructed, spanning the space of multivariate splines of specified degrees and knot sequences in all coordinate directions. The result is an expansion of a general L^2-function with respect to measure-consistent multivariate orthonormal B-splines. Compared with the existing PCE, SCE, rooted in compactly supported B-splines, deals with locally prominent stochastic responses in a more proficient manner. The approximation quality of the expansion was demonstrated in terms of the modulus of smoothness of the function, leading to the mean-square convergence of SCE to the correct limit. The weaker modes of convergence, such as those in probability and in distribution, follow readily. The optimality of SCE, including deriving PCE as a special case of SCE, was demonstrated. Analytical formulae akin to those found in the PCE literature are proposed to calculate the mean and variance of an SCE approximation for a general output variable in terms of the expansion coefficients. Numerical results obtained for one-, two-, and four-dimensional UQ problems entailing oscillatory, nonsmooth, and nearly discontinuous functions indicate that a low-order SCE approximation with an adequate mesh is capable of producing a substantially more accurate estimates of the output variance and probability distribution than a PCE with an overly large order of approximation.

Acknowledgments. The author thanks two anonymous reviewers and the associate editor for providing a number of helpful comments on an earlier draft of the paper.

REFERENCES

[1] R. H. Cameron and W. T. Martin, *The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals*, Ann. of Math., 48 (1947), pp. 385–392.
[2] J. A. Cottrell, T. J. R. Hughes, and Y. Bazilevs, *Isogeometric Analysis: Toward Integration of CAD and FEA*, John Wiley & Sons, Hoboken, NJ, 2009.
[3] M. G. Cox, *The numerical evaluation of B-splines*, IMA J. Appl. Math., 10 (1972), pp. 134–149.
[4] W. Dahmen, R. de Vore, and K. Scherer, *Multidimensional spline approximation*, SIAM J. Numer. Anal., 17 (1980), pp. 380–402.
[5] C. De Boor, *On calculation with B-splines*, J. Approx. Theory, 6 (1972), pp. 50–62.
[6] C. De Boor and A. Pinkus, *The B-spline recurrence relations of Chakalov and of Popoviciu*, J. Approx. Theory, 124 (2003), pp. 115–123.
[7] O. G. Ernst, A. Mugler, H. J. Starkloff, and E. Ullmann, *On the convergence of generalized polynomial chaos expansions*, ESAIM Math. Model. Numer. Anal., 46 (2012), pp. 317–339.
[8] T. Gerstner and M. Griebel, *Numerical integration using sparse grids*, Numer. Algorithms, 18 (1998), pp. 209–232.
[9] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed., John Hopkins University Press, Baltimore, MD, 1996.
[10] M. Grigoriu, *Stochastic Calculus: Applications in Science and Engineering*, Birkhaüser, Basel, 2002.
[11] A. Kessy, A. Lewin, and K. Strimmer, *Optimal whitening and decorrelation*, Amer. Statist., 72 (2018), pp. 309–314.
[12] L. A. Piegl and W. Tiller, *The NURBS Book*, 2nd ed., Springer-Verlag, Berlin, 1997.
[13] S. Rahman, *Mathematical properties of polynomial dimensional decomposition*, SIAM/ASA J. Uncertain. Quantif., 6 (2018), pp. 816–844.
[14] S. Rahman, *A polynomial chaos expansion in dependent random variables*, J. Math. Anal. Appl., 464 (2018), pp. 749–775.
[15] A. Saltelli and I. M. Sobol, *About the use of rank transformation in sensitivity analysis of model output*, Reliab. Engrg. System Safety, 50 (1995), pp. 225–239.
[16] I. Schoenberg, *On spline functions*, in Inequalities, O. Shisha, ed., Academic Press: New York, NY, 1967, pp. 255–286.
[17] L. J. Schumaker, *Spline Functions: Basic Theory*, 3rd ed., Cambridge University Press, Cambridge, UK, 2007.
[18] R. Smith, *Uncertainty Quantification: Theory, Implementation, and Applications*, SIAM: Philadelphia, PA, 2013.
[19] T. J. Sullivan, *Introduction to Uncertainty Quantification*, Springer, New York, NY, 2015.
[20] A. F. Timan, *Theory of Approximation of Functions of a Real Variable*, Dover New York, NY, 1994.
[21] X. Wan and G. Karniadakis, *An adaptive multi-element generalized polynomial chaos method for stochastic differential equations*, J. Comput. Phys., 209 (2005), pp. 617–642.
[22] N. Wiener, *The homogeneous chaos*, Amer. J. Math., 60 (1938), pp. 897–936.
[23] H. Xu and S. Rahman, *A generalized dimension-reduction method for multi-dimensional integration in stochastic mechanics*, Internat. J. Numer. Methods Engrg., 61 (2004), pp. 1992–2019.
[24] Z. Zhang and C. F. Martin, *Convergence and Gibbs' phenomenon in cubic spline interpolation of discontinuous functions*, J. Comput. Appl. Math., 87 (1997), pp. 359–371.