Received: 7 November 2020 | Revised: 26 June 2021 | Accepted: 29 June 2021

DOI: 10.1002/nme.6778

RESEARCH ARTICLE

WILEY

Uncertainty quantification by optimal spline dimensional decomposition

Steven Dixler | **Ramin Jahanbin** | **Sharif Rahman**

College of Engineering, The University of Iowa, Iowa City, Iowa, USA

Correspondence
Sharif Rahman, College of Engineering, The University of Iowa, Iowa City, IA 52242, USA.
Email: sharif-rahman@uiowa.edu

Abstract

An optimal version of spline dimensional decomposition (SDD) is unveiled for general high-dimensional uncertainty quantification analysis of complex systems subject to independent but otherwise arbitrary probability measures of input random variables. The resulting method involves optimally derived knot vectors of basis splines (B-splines) in some or all coordinate directions, whitening transformation producing measure-consistent orthonormalized B-splines equipped with optimal knots, and Fourier-spline expansion of a general high-dimensional output function of interest. In contrast to standard SDD, there is no need to select the knot vectors uniformly or intuitively. The generation of optimal knot vectors can be viewed as an inexpensive preprocessing step toward creating the optimal SDD. Analytical formulas are proposed to calculate the second-moment properties by the optimal SDD method for a general output random variable in terms of the expansion coefficients involved. It has been shown that the computational complexity of the optimal SDD method is polynomial, as opposed to exponential, thus mitigating the curse of dimensionality by a discernible magnitude. Numerical results affirm that the optimal SDD method developed is more precise than polynomial chaos expansion, sparse-grid quadrature, and the standard SDD method in calculating not only the second-moment statistics, but also the cumulative distribution function of an output random variable. More importantly, the optimal SDD outperforms standard SDD by sustaining nearly identical computational cost.

KEYWORDS

B-splines, polynomial chaos expansion, sparse grids, spline chaos expansion, spline dimensional decomposition, stochastic analysis

1 | INTRODUCTION

Uncertainty quantification, commonly referred to as UQ, is a contemporary scientific discipline, cutting across traditional research areas of engineering and applied sciences, numerical analysis, and probability and statistics. A long-standing staple for UQ analysis involves measure-consistent orthogonal polynomials in input random variables, such as those readily exploited by the polynomial chaos expansion (PCE)[1-3] and polynomial dimensional decomposition (PDD)[4-6] methods. These methods are known to offer significant computational advantages over crude Monte Carlo simulation (MCS), especially when the output variable is globally smooth over the entire domain of input random

variables. While the polynomial-based expansion methods have become mainstream and found numerous applications, they are not well suited to deal with locally significant changes, including discontinuity and nonsmoothness, in stochastic responses of interest. In this regard, two relatively new expansion methods introducing orthonormal version of basis splines (B-splines), namely, the spline chaos expansion (SCE)[7] and spline dimensional decomposition (SDD),[8,9] have been created to tackle locally pronounced, highly nonlinear, or nonsmooth responses. In consequence, a low-degree SCE or SDD approximation with an adequate mesh size is capable of producing a markedly more accurate estimate of the output variance than PCE or PDD approximations with exceedingly large expansion orders.[7-9]

While the SCE and SDD methods both entail orthonormalized B-splines as their basis functions, the SDD method, akin to its predecessor, the PDD method, is better adapted to handling high-dimensional UQ problems than SCE or PCE. However, the basis functions of the existing SDD are often predicated on knot vectors with uniform spacing in all coordinate directions. As a result, they are not necessarily maximally empowered to capture locally abrupt changes in stochastic responses, if they exist, in an efficient manner. The authors argue that the knot vectors should be selected optimally, leading to more powerful basis functions, thereby imparting superior approximation quality to the resulting SDD expansion. The development of such SDD featuring optimally derived knot vectors is the principal motivation for this work.

Two other UQ methods, frequently used by engineers and applied mathematicians, are the stochastic collocation method[10,11] and sparse-grid quadrature.[12,13] Depending on the problem at hand, both of them are generally more efficient than crude MCS. However, for truly high-dimensional problems, a collocation method like PCE or SCE also requires an astronomically large number of basis functions or coefficients, succumbing to the curse of dimensionality. While the B-splines have also been employed to construct the sparse-grid quadrature, they are neither orthogonal nor measure-consistent, meaning that the basis functions are not adapted to the probability measure of input random variables. More importantly, the foundational idea of sparse grids can be traced to the referential dimensional decomposition (RDD),[14] also known as anchored decomposition[15] or cut-high-dimensional model representation,[16] of a high-dimensional function. By contrast, SDD is rooted in the analysis-of-variance (ANOVA) dimensional decomposition (ADD), which is generally superior to RDD. Indeed, an error analysis reveals the suboptimality of RDD approximations, meaning that an RDD approximation, regardless of how the reference point is chosen, cannot be better than an ADD approximation for identical degrees of interaction.[14]

This article presents a novel optimal SDD method for general uncertainty quantification analysis of complex systems subject to independent but otherwise arbitrary probability measures of input random variables. The method involves (1) optimally derived knot vectors of B-splines in some or all coordinate directions; (2) whitening transformation producing measure-consistent orthonormalized B-splines equipped with optimal knots; and (3) Fourier-spline expansion of a general high-dimensional output function of interest. The article is structured as follows. Section 2 formally defines a general UQ problem, including a list of requisite assumptions for input and output random variables. Section 3 describes the construction of measure-consistent orthonormalized B-splines. A general SDD method is summarized in Section 4, including derivation of the output statistics from the truncated expansion. Two approximate methods for estimating the SDD coefficients are presented in Section 5. Section 6 explains the derivation of the optimal knots from two distinct approaches, thereby yielding the new optimal SDD method. The results from two numerical examples and an industrial-scale engineering application are reported in Sections 7 and 8, respectively. Finally, conclusions are drawn in Section 9.

2 | A GENERAL UQ PROBLEM

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

2.1 | Input random variables

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, including external loads, displacement boundary conditions, material properties, and geometry, of a complex mechanical system. Denote by $F_{\mathbf{X}}(\mathbf{x}) := \mathbb{P}(\cap_{k=1}^{N}\{X_k \le x_k\})$ the joint cumulative distribution function (CDF) of $\mathbf{X}$. The kth component of $\mathbf{X}$ is a random variable X_k, which has the marginal CDF $F_{X_k}(x_k) := \mathbb{P}(X_k \le x_k)$. The positive integer N, which represents the total number of input random variables, is often referred to as the dimension of the stochastic or UQ problem. As an example, consider a cantilevered beam with length L and Young's modulus E, which is subjected to a vertically applied concentrated force F and moment M at its free end. If all of these input parameters are modeled as random variables, then $\mathbf{X} := (L, E, F, M)^\intercal$ with stochastic dimension $N = 4$.

Depending on the UQ problem, some of the input parameters may possess spatial variability, suggesting a need for their random field description. Common examples are elastic properties of engineering materials, size and shape characteristics of mechanical components, and wind and snow loads in structural systems, to name a few. For computational purposes, the random fields must be discretized into a finite number of constituent random variables. An often-used approach entails the Karhunen–Loève (K–L) expansion,[17] leading to an infinite series expansion of the random field consisting of deterministic functions of space and uncorrelated random variables. Recent works on the K–L approximation are premised on Galerkin isogeometric[18] and isogeometric collocation methods,[19,20] where the latter, by eliminating one dimension-order of domain integration, offers a hefty computational advantage over the former. Readers interested in further detail are directed to the aforementioned works.

In summary, the randomness in the UQ problem may stem from spatially invariant random variables alone, from the discretization of spatially variant random fields, or both. If the problem contains spatially invariant random variables and random fields, then N represents the total number of random variables.

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

Assumption 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 $l \in \mathbb{N}$,

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

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 CDF $F_{X_k}(x_k)$ and continuous marginal probability density function (PDF) $f_{X_k}(x_k) := \partial F_{X_k}(x_k)/\partial x_k$ with a bounded support $[a_k, b_k] \subset \mathbb{R}$. Consequently, with Items (1) and (2) in mind, the joint CDF $F_{\mathbf{X}}(\mathbf{x})$ and joint PDF $f_{\mathbf{X}}(\mathbf{x}) := \partial^N F_{\mathbf{X}}(\mathbf{x})/\partial x_1 \ldots \partial x_N$ of $\mathbf{X}$ are obtained from

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

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

Assumption 1 ensures 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 work.

2.2 | An output random variable of interest

Given an input random vector $\mathbf{X} := (X_1, \ldots, X_N)^\intercal : (\Omega, \mathcal{F}) \to (\mathbb{A}^N, \mathcal{B}^N)$ with known PDF $f_{\mathbf{X}}(\mathbf{x})$ on $\mathbb{A}^N \subset \mathbb{R}^N$, denote by $y(\mathbf{X}) := y(X_1, \ldots, X_N)$ a real-valued, square-integrable, measurable transformation on $(\Omega, \mathcal{F})$. Here, $y : \mathbb{A}^N \to \mathbb{R}$ represents a relevant function from a mathematical model, describing an output response of interest for a UQ problem. A major objective of UQ analysis is to estimate the probabilistic characteristics of an output random variable $Y = y(\mathbf{X})$, including its statistical moments and CDF, when the probability law of the input random vector $\mathbf{X}$ is prescribed. More often than not, Y is assumed to belong to a reasonably large class of random variables, such as the weighted L^2 space

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

which is a Hilbert space with the inner product

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

and norm

$$\|y(\mathbf{X})\|_{L^2(\Omega,\mathcal{F},\mathbb{P})} := \sqrt{(y(\mathbf{X}), y(\mathbf{X}))_{L^2(\Omega,\mathcal{F},\mathbb{P})}} = \sqrt{\int_\Omega y^2(\mathbf{X}(\omega)) \,\mathrm{d}\mathbb{P}(\omega)} = \sqrt{\int_{\mathbb{A}^N} y^2(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) \,\mathrm{d}\mathbf{x}}.$$

3 | MEASURE-CONSISTENT ORTHONORMAL SPLINES

For the coordinate direction $k = 1, \dots, N$, define a nonnegative integer $p_k \in \mathbb{N}_0$ and a positive integer $n_k \geq p_k + 1$, representing the degree and number of basis functions, respectively. Then, a knot vector

$$\boldsymbol{\xi}_k := \{\xi_{k,i_k}\}_{i_k=1}^{n_k+p_k+1} = \{a_k = \xi_{k,1}, \xi_{k,2}, \dots, \xi_{k,n_k+p_k+1} = b_k\}$$

is defined on the interval $[a_k, b_k]$ by a nondecreasing sequence of real numbers, where ξ_{k,i_k} is the i_kth knot with $i_k = 1, 2, \dots, n_k + p_k + 1$. Any knot may appear up to $p_k + 1$ times in the sequence. Hence, the knot vector can be rewritten as

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

$$a_k = \zeta_{k,1} < \zeta_{k,2} < \cdots < \zeta_{k,r_k-1} < \zeta_{k,r_k} = b_k,$$

where ζ_{k,j_k}, $j_k = 1, 2, \dots, r_k$ are the r_k unique knots, each of which has multiplicity $1 \leq m_{k,j_k} \leq p_k + 1$. For more details, readers are referred to Appendix A of this article and Chapter 2 of the book by Cottrell et al.[21] A knot vector is called $(p_k + 1)$-open if the end knots have multiplicities $p_k + 1$. In this work, only $(p_k + 1)$-open knot vectors are considered.

Denote by $B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k)$ the i_kth univariate B-spline with degree p_k. Given the knot vector $\boldsymbol{\xi}_k$ and *zero*-degree basis functions, all higher-order* B-spline functions on $[a_k, b_k]$ are defined recursively, where $1 \leq k \leq N$, $1 \leq i_k \leq n_k$, and $1 \leq p_k < \infty$. See Appendix A for an explicit definition of $B^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k)$.

The B-splines are endowed with a number of desirable properties, which can generally deliver tremendous approximating power to numerical methods. More specifically, they are[21,22]: (1) nonnegative; (2) locally supported on the interval $[\xi_{k,i_k}, \xi_{k,i_k+p_k+1})$ for all i_k; (3) linearly independent; (4) committed to partition of unity; and (5) pointwise C^∞-continuous everywhere except at the knots ζ_{k,j_k} of multiplicity m_{k,j_k} for all j_k, where they are $C^{p_k - m_{k,j_k}}$-continuous, provided that $1 \leq m_{k,j_k} < p_k + 1$.

3.1 | Univariate orthonormalized B-splines

The aforementioned B-splines, although they form a basis of the spline space of degree p_k and knot vector $\boldsymbol{\xi}_k$, are not orthogonal with respect to the probability measure $f_{X_k}(x_k)\,\mathrm{d}x_k$ of X_k. A linear transformation, originally proposed in the prequel,[8] is summarized here in three steps to generate their orthonormal version.

(1) Given a set of B-splines of degree p_k, create an auxiliary set by replacing any element, arbitrarily chosen to be the first, with *one*. Arrange the elements of the set into an n_k-dimensional vector

$$\mathbf{P}_k(x_k) := \left(1, B^k_{2,p_k,\boldsymbol{\xi}_k}(x_k), \dots, B^k_{n_k,p_k,\boldsymbol{\xi}_k}(x_k)\right)^\intercal$$

comprising the auxiliary B-splines. The linear independence of the auxiliary B-splines is preserved.[7]

*Degree and order are used interchangeably in this article.

(2) Construct an $n_k \times n_k$ spline moment matrix

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

The matrix $\mathbf{G}_k$ exists because X_k has finite moments up to order $2p_k$, as stated in Assumption 1. Furthermore, it is symmetric and positive-definite,[7] ensuring the existence of a nonsingular $n_k \times n_k$ whitening matrix $\mathbf{W}_k$ such that

$$\mathbf{W}_k^\intercal \mathbf{W}_k = \mathbf{G}_k^{-1}.$$

(3) Apply a whitening transformation to create a vector of orthonormalized B-splines

$$\boldsymbol{\psi}_k(x_k) = \mathbf{W}_k \mathbf{P}_k(x_k),$$

consisting of uncorrelated components

$$\psi^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k), \quad i_k = 1, \ldots, n_k, \quad k = 1, \ldots, N.$$

Note that the invertibility of $\mathbf{G}_k$ does not uniquely determine $\mathbf{W}_k$. Indeed, there are several ways to choose $\mathbf{W}_k$ such that the condition described in Step 2 is satisfied.[7] One prominent, relatively stable option is to invoke the Cholesky factorization $\mathbf{G}_k = \mathbf{Q}_k \mathbf{Q}_k^\intercal$, leading to

$$\mathbf{W}_k = \mathbf{Q}_k^{-1}.$$

Figure 1(A–D) depicts a set of second-order ($p = 2$) B-spline functions on $[-1, 1]$ with an arbitrarily chosen, nonuniformly spaced knot vector $\boldsymbol{\xi} = \{-1, -1, -1, -0.75, 0, 0.75, 1, 1, 1\}$ before and after orthonormalization. The six B-splines in Figure 1(A) were derived using the Cox–de Boor formula in Appendix A. Once again, they are nonnegative and locally supported, but not orthonormal with respect to the probability measure of the random variable X with the bounded domain $[-1, 1]$. By contrast, Figure 1(B–D) illustrates the associated orthonormalized B-splines, obtained when X is a uniform, truncated Gaussian, and beta random variable, respectively, with their PDFs $f_X : [-1, 1] \to \mathbb{R} \setminus (-\infty, 0)$ given as

$$\frac{1}{2}, \quad \frac{2\phi(2x+1)}{\Phi(3) - \Phi(-1)}, \quad \text{and} \quad \frac{\Gamma(5)(x+1)^2(1-x)}{16\Gamma(3)\Gamma(2)},$$

where Φ and ϕ respectively denote the probability distribution and density functions of a standard Gaussian random variable. For all three cases, the spline moment matrix was calculated analytically and the whitening matrix was obtained by inverting the Cholesky factorization, as suggested previously. The shapes of the spline functions depend on both the spacing of the knots and the probability measure of X. Unlike in Figure 1(A), all nonconstant functions in Figure 1(B–D) have a mean of *zero*. Furthermore, the orthonormalized B-splines are neither nonnegative nor locally supported. However, an orthonormal basis is an essential ingredient in constructing an SDD expansion.

As illustrated by the examples in Figure 1(B–D), the spline moment matrix can be evaluated analytically for some PDFs. If this is not the case, it can be approximated using a Gauss-type quadrature rule. Since no potentially expensive output function evaluations are involved, the numerical integration can be performed with an arbitrary level of precision. While it is possible to do this using a single set of measure-consistent quadrature points and weights of sufficiently high order corresponding to the entire domain $[a_k, b_k]$, it is often more accurate to generate measure-consistent quadrature points and weights for each subinterval defined between successive, unique knots. In this case, the integral is split into several integrals, each operating on a polynomial of order $2p_k$ multiplied by a part of the PDF. Whichever method used to compute the spline moment matrix can also be used to compute its derivatives, given in Appendix B, with respect to the internal knots for purposes that will be explained later.

3.2 | Multivariate orthonormalized B-splines

Due to the product-type probability measure of independent random input variables, measure-consistent multivariate orthonormalized B-splines in N variables are easily built from the N-dimensional tensor product of measure-consistent

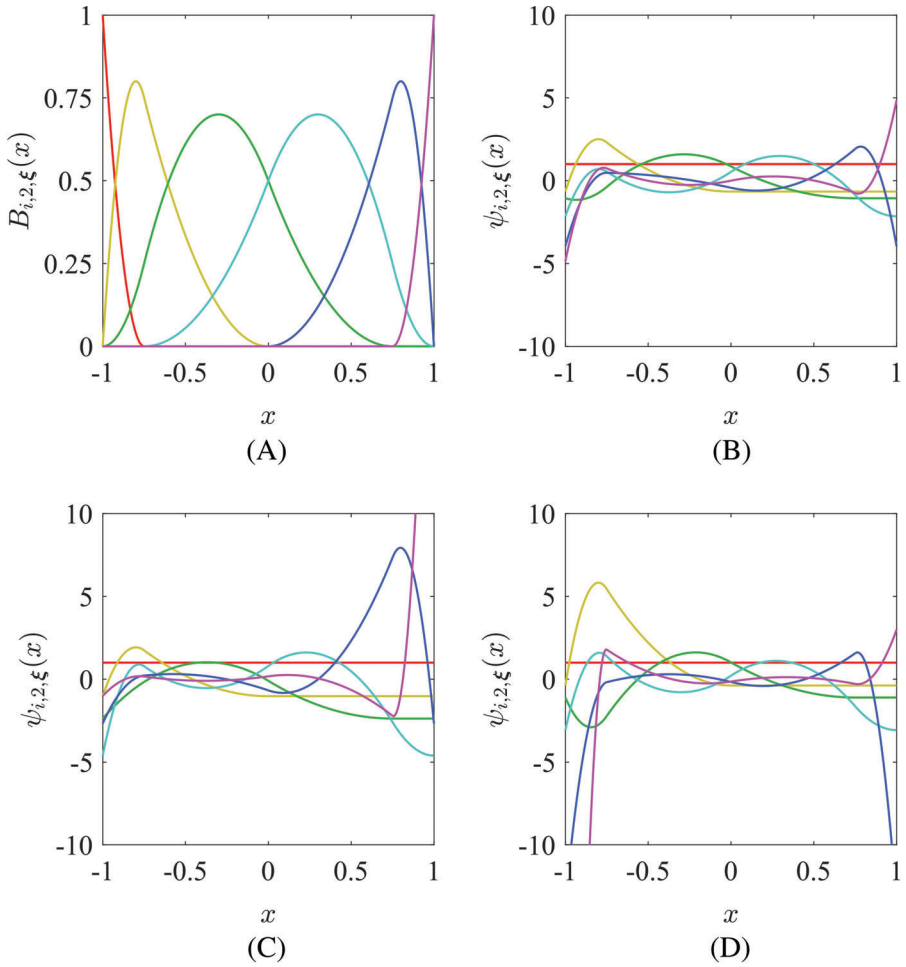


FIGURE 1 A set of B-splines associated with the knot vector $\boldsymbol{\xi} = \{-1, -1, -1, -0.75, 0, 0.75, 1, 1, 1\}$ and order $p = 2$: (A) nonorthonormal basis; (B) orthonormal basis for uniform measure; (C) orthonormal basis for truncated Gaussian measure; and (D) orthonormal basis for beta measure

univariate orthonormalized B-splines. However, forming such a tensor product in a high-dimensional setting is not recommended. Instead, the authors advocate constructing a series of tensor products in a dimensionwise manner.

Denote by $\emptyset \neq u = \{k_1, \ldots, k_{|u|}\} \subseteq \{1, \ldots, N\}$ a nonempty subset of the index set $\{1, \ldots, N\}$ with cardinality $1 \leq |u| \leq N$. For such a subset, let $\mathbf{X}_u := (X_{k_1}, \ldots, X_{k_{|u|}})^\intercal$ be a subvector of $\mathbf{X}$ defined on the abstract probability space $(\Omega^u, \mathcal{F}^u, \mathbb{P}^u)$, where Ω^u is the sample space of $\mathbf{X}_u$, $\mathcal{F}^u$ is a σ-algebra on Ω^u, and $\mathbb{P}^u$ is a probability measure. As $\mathbf{X}$ comprises independent random variables, the PDF of $\mathbf{X}_u$ is

$$f_{\mathbf{X}_u}(\mathbf{x}_u) = \prod_{k \in u} f_{X_k}(x_k) = \prod_{l=1}^{|u|} f_{X_{k_l}}(x_{k_l}), \ \mathbf{x}_u := (x_{k_1}, \ldots, x_{k_{|u|}})^\intercal.$$

Define three multiindices $\mathbf{i}_u := (i_{k_1}, \ldots, i_{k_{|u|}}) \in \mathbb{N}^{|u|}$, $\mathbf{n}_u := (n_{k_1}, \ldots, n_{k_{|u|}}) \in \mathbb{N}^{|u|}$, and $\mathbf{p}_u := (p_{k_1}, \ldots, p_{k_{|u|}}) \in \mathbb{N}_0^{|u|}$, representing the knot indices, numbers of basis functions, and degrees of splines, respectively, in all $|u|$ coordinate directions. Denote by $\boldsymbol{\Xi}_u := \{\boldsymbol{\xi}_{k_1}, \ldots, \boldsymbol{\xi}_{k_{|u|}}\}$ a family of all $|u|$ knot sequences. Associated with $\mathbf{i}_u$, define an index set

$$\mathcal{I}_{u,\mathbf{n}_u} := \left\{\mathbf{i}_u = (i_{k_1}, \ldots, i_{k_{|u|}}) : 1 \leq i_{k_l} \leq n_{k_l}, \ l = 1, \ldots, |u|\right\} \subset \mathbb{N}^{|u|}$$

with cardinality

$$|\mathcal{I}_{u,\mathbf{n}_u}| = \prod_{k \in u} n_k.$$

For the coordinate direction k_l, define by

$$I_{k_l} = r_{k_l} - 1,$$

the number of subintervals corresponding to the knot vector $\boldsymbol{\xi}_{k_l}$ with r_{k_l} distinct knots. Then the partition, defined by the knot vectors $\boldsymbol{\xi}_{k_1}, \ldots, \boldsymbol{\xi}_{k_{|u|}}$, decomposes the $|u|$-dimensional rectangle $\mathbb{A}^u := \times_{k\in u}[a_k, b_k]$ into smaller rectangles

$$\left\{\mathbf{x}_u = (x_{k_1}, \ldots, x_{k_{|u|}}) : \zeta_{k_l j_{k_l}} \le x_{k_l} < \zeta_{k_l, j_{k_l+1}},\ l = 1, \ldots, |u|\right\},\ j_{k_l} = 1, \ldots, I_{k_l},$$

where $\zeta_{k_l j_{k_l}}$ is the j_{k_l}th distinct knot in the coordinate direction k_l. A mesh is defined by the partition of $\mathbb{A}^u$ into such rectangular elements. Define the largest element size in each coordinate direction $k \in u$ by

$$h_{u,k_l} := \max_{j_{k_l}=1,\ldots,I_{k_l}} \left(\zeta_{k_l j_{k_l}+1} - \zeta_{k_l j_{k_l}}\right),\ l = 1, \ldots, |u|.$$

Then, given the knot vectors $\boldsymbol{\Xi}_u = \{\boldsymbol{\xi}_{k_1}, \ldots, \boldsymbol{\xi}_{k_{|u|}}\}$,

$$\mathbf{h}_u := (h_{u,k_1}, \ldots, h_{u,k_{|u|}}) \text{ and } h_u := \max_{l=1,\ldots,|u|} h_{u,k_l}$$

define a vector of the largest element sizes in all $|u|$ coordinates and the global mesh size, respectively, for the domain $\mathbb{A}^u$. Consequently, for $\emptyset \neq u = \{k_1, \ldots, k_{|u|}\} \subseteq \{1, \ldots, N\}$, with $\mathbf{p}_u = (p_{k_1}, \ldots, p_{k_{|u|}}) \in \mathbb{N}_0^{|u|}$ and $\boldsymbol{\Xi}_u = \{\boldsymbol{\xi}_{k_1}, \ldots, \boldsymbol{\xi}_{k_{|u|}}\}$ in mind, the multivariate orthonormalized B-splines in $\mathbf{x}_u = (x_{k_1}, \ldots, x_{k_{|u|}})$ consistent with the probability measure $f_{\mathbf{X}_u}(\mathbf{x}_u)\, \mathrm{d}\mathbf{x}_u$ are

$$\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{x}_u) = \prod_{k\in u} \psi^k_{i_k,p_k,\boldsymbol{\xi}_k}(x_k) = \prod_{l=1}^{|u|} \psi^{k_l}_{i_{k_l},p_{k_l},\boldsymbol{\xi}_{k_l}}(x_{k_l}),\quad \mathbf{i}_u = (i_{k_1}, \ldots, i_{k_{|u|}}) \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}, \tag{1}$$

where

$$\overline{\mathcal{I}}_{u,\mathbf{n}_u} := \left\{\mathbf{i}_u = (i_{k_1}, \ldots, i_{k_{|u|}}) : 2 \le i_{k_l} \le n_{k_l},\ l = 1, \ldots, |u|\right\} \subset (\mathbb{N} \setminus \{1\})^{|u|}$$

is a reduced index set, which has cardinality

$$|\overline{\mathcal{I}}_{u,\mathbf{n}_u}| := \prod_{k\in u}(n_k - 1). \tag{2}$$

The key difference between the index sets $\overline{\mathcal{I}}_{u,\mathbf{n}_u}$ and $\mathcal{I}_{u,\mathbf{n}_u}$ is that the former limits the range of index i_{k_l}, $l = 1, \ldots, |u|$, associated with the k_lth variable x_{k_l}, to $2, \ldots, n_{k_l}$. The exclusion of $i_{k_l} = 1$ removes the first constant element of $\boldsymbol{\psi}_{k_l}(x_{k_l})$ in order to prevent reduction of the degree of interaction of the corresponding multivariate spline basis below $|u|$.

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 $\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{X}_u)$, $\emptyset \neq u \subseteq \{1, \ldots, N\}$, $\mathbf{i}_u \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}$, become functions of random input variables. Then, for $\emptyset \neq u, v \subseteq \{1, \ldots, N\}$, $\mathbf{i}_u \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}$, and $\mathbf{j}_v \in \overline{\mathcal{I}}_{v,\mathbf{n}_v}$, the first- and second-order moments of multivariate orthonormalized B-splines are[9]

$$\mathbb{E}\left[\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{X}_u)\right] = 0 \tag{3}$$

and

$$\mathbb{E}\left[\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{X}_u)\Psi^v_{\mathbf{j}_v,\mathbf{p}_v,\boldsymbol{\Xi}_v}(\mathbf{X}_v)\right] = \begin{cases} 1, & u = v \text{ and } \mathbf{i}_u = \mathbf{j}_v, \\ 0, & \text{otherwise}, \end{cases} \tag{4}$$

respectively. The *zero*-mean and orthonormal properties of spline functions, defined in Equations (3) and (4), are crucial ingredients of SDD.

It is important to underscore that the multivariate and univariate orthonormalized B-splines derived in this section are both consistent with an arbitrary probability measure of input random variables with bounded domains. However, there are many UQ problems where there exist random variables with unbounded domains. In such a case, an appropriate probability preserving transformation, mapping a random variable with unbounded domain to another with bounded domain solves the problem. The transformation will be revisited when numerical examples are discussed.

4 | A GENERAL SPLINE DIMENSIONAL DECOMPOSITION

A principal objective of solving the UQ problem is to effectively estimate the relevant probabilistic characteristics of $y(\mathbf{X}) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$. The dimension N of a real-life stochastic problem often exceeds 10 and may even be in the realm of hundreds, where the output response function $y(\mathbf{X})$ is highly nonlinear with locally significant changes, including discontinuity and nonsmoothness, with respect to the random input $\mathbf{X}$. Although somewhat subjective, here, the stochastic dimension exceeding 10 is viewed as a high-dimensional UQ problem. Nonetheless, stochastic computation for complex mechanical systems in a high-dimensional stochastic domain $\mathbb{A}^N$ is an expensive initiative.

4.1 | ANOVA dimensional decomposition

For any nonempty subset u and subvector $\mathbf{X}_u := (X_{k_1}, \ldots, X_{k_{|u|}})^\intercal$, denote by $-u := \{1, \ldots, N\} \setminus u$ and $\mathbf{X}_{-u} := \mathbf{X}_{\{1,\ldots,N\}\setminus u}$ their complementary subset and subvector, respectively. Then, for a given $\emptyset \neq u \subseteq \{1, \ldots, N\}$, the marginal density function of $\mathbf{X}_u$, defined on $\mathbb{A}^u := \times_{k\in u}\mathbb{A}^{\{k\}} \subset \mathbb{R}^{|u|}$, is

$$f_{\mathbf{X}_u}(\mathbf{x}_u) := \int_{\mathbb{A}^{-u}} f_{\mathbf{X}}(\mathbf{x})\, \mathrm{d}\mathbf{x}_{-u} = \prod_{k\in u} f_{X_k}(x_k),$$

where the second equality forms due to statistical independence of the input random variables as per Assumption 1. Hence, it can be shown that, for any function $y \in L^2(\Omega, \mathcal{F}, \mathbb{P})$, there exists a unique, finite, hierarchical expansion[14,23,24]

$$y(\mathbf{X}) = y_\emptyset + \sum_{\emptyset\neq u\subseteq\{1,\ldots,N\}} y_u(\mathbf{X}_u), \tag{5a}$$

$$y_\emptyset := \int_{\mathbb{A}^N} y(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x})\, \mathrm{d}\mathbf{x}, \tag{5b}$$

$$y_u(\mathbf{X}_u) := \int_{\mathbb{A}^{-u}} y(\mathbf{X}_u, \mathbf{x}_{-u}) f_{\mathbf{X}_{-u}}(\mathbf{x}_{-u})\, \mathrm{d}\mathbf{x}_{-u} - \sum_{v\subset u} y_v(\mathbf{X}_v), \tag{5c}$$

referred to as ADD, where y_u is a $|u|$-variate component function describing a constant or a $|u|$-variate interaction of $\mathbf{x}_u = (x_{k_1}, \ldots, x_{k_{|u|}})$ on y when $|u| = 0$ or $|u| > 0$. Here, $(\mathbf{X}_u, \mathbf{x}_{-u})$ denotes an N-dimensional vector whose kth component is X_k if $k \in u$ and x_k if $k \notin u$. The summation in Equation (5a) comprises $2^N - 1$ terms with each term depending on a group of variables indexed by a particular subset of $\{1, \ldots, N\}$.

The decomposition presented in Equations (5a)–(5c) has two notable properties:[14]

(1) Any nonconstant component function $y_u(\mathbf{X}_u)$ has a *zero* mean, that is,

$$\mathbb{E}[y_u(\mathbf{X}_u)] = 0,\ \emptyset \neq u \in \{1, \ldots, N\}. \tag{6}$$

(2) Any two distinct component functions $y_u(\mathbf{X}_u)$ and $y_v(\mathbf{X}_v)$ are mutually orthogonal, that is,

$$\mathbb{E}[y_u(\mathbf{X}_u) y_v(\mathbf{X}_v)] = 0,\ u, v \in \{1, \ldots, N\},\ u \neq v. \tag{7}$$

Readers interested in further details of ADD are directed to prior works.[14,23,24]

It is elementary to show that all ADD component functions of $y(\mathbf{X})$ are members of respective subspaces of $L^2(\Omega, \mathcal{F}, \mathbb{P})$. Unfortunately, the subspaces are infinite-dimensional. Therefore, a further discretization or refinement is

necessary. In contrast to the past works on polynomial refinements, a new spline adaptation of the subspaces, spanning measure-consistent orthonormalized B-splines, is proposed.

4.2 | Fourier spline expansion

Consider any nonconstant component function of ADD defined in Equation (5c). Given the second-moment properties in Equations (3) and (4), $y_u(\mathbf{X}_u)$ belongs to the Hilbert space

$$S_u := \left\{ y_u(\mathbf{X}_u) \in L^2(\Omega^u, \mathcal{F}^u, \mathbb{P}^u) : \mathbb{E}\left[y_u(\mathbf{X}_u) y_v(\mathbf{X}_v)\right] = 0 \text{ if } u \neq v,\ u, v \subseteq \{1, \ldots, N\} \right\}.$$

Denote by

$$\{\Psi^u_{\mathbf{i}_u, \mathbf{p}_u, \boldsymbol{\Xi}_u}(\mathbf{X}_u) : \mathbf{i}_u \in \overline{\mathcal{I}}_{u, \mathbf{n}_u}\}$$

a set of measure-consistent multivariate orthonormalized B-splines in $\mathbf{X}_u$. The cardinality of the set is $|\overline{\mathcal{I}}_{u,\mathbf{n}_u}|$, defined in Equation (2). It is controlled by the numbers of basis functions n_{k_l}, which in turn is determined by the lengths of the knot vectors $\boldsymbol{\xi}_{k_l}$ and the corresponding orders p_{k_l}. Clearly, any increase in the lengths of the knot vectors increases the number of bases, thereby expanding $\overline{\mathcal{I}}_{u,\mathbf{n}_u}$. With this in mind, consider a refinement process with fixed $\mathbf{p}_u$ in which n_{k_l} or, equivalently, the length of $\boldsymbol{\xi}_{k_l}$ is increased in all $|u|$ coordinate directions, in such a way that the largest element size h_{u,k_l} is monotonically reduced. The result is an increasing family of the sets of such basis functions. In the limit, when $n_{k_l} \to \infty$, $k_l = 1, \ldots, |u|$, denote by $\boldsymbol{\xi}_{k_l,\infty}$ and $\boldsymbol{\Xi}_{u,\infty} = (\boldsymbol{\xi}_{1,\infty}, \ldots, \boldsymbol{\xi}_{|u|,\infty})$ the associated knot vector in the k_lth coordinate direction and the family of such $|u|$ knot vectors, respectively. Then, there exists an infinite number of basis functions with the associated index set

$$\overline{\mathcal{I}}_{u,\infty} := \left\{ \mathbf{i}_u = (i_{k_1}, \ldots, i_{k_{|u|}}) : 2 \leq i_{k_l} < \infty,\ l = 1, \ldots, |u| \right\},$$

representing the infinite counterpart of $\overline{\mathcal{I}}_{u,\mathbf{n}_u}$. Consequently, the infinite set of multivariate orthonormalized B-splines $\{\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_{u,\infty}}(\mathbf{X}_u) : \mathbf{i}_u \in \overline{\mathcal{I}}_{u,\infty}\}$ forms an orthonormal basis of $\mathcal{S}_u$, yielding

$$\mathcal{S}_u = \overline{\text{span}\{\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_{u,\infty}}(\mathbf{X}_u) : \mathbf{i}_u \in \overline{\mathcal{I}}_{u,\infty}\}},$$

where the overline stands for set closure.

According to the standard Hilbert space theory, every $y_u(\mathbf{X}_u) \in \mathcal{S}_u$ can be expanded in terms of the aforementioned spanning set, resulting in

$$y_u(\mathbf{X}_u) \sim \sum_{\mathbf{i}_u \in \overline{\mathcal{I}}_{u,\infty}} C^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_\infty} \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_{u,\infty}}(\mathbf{X}_u), \tag{8}$$

where

$$C^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_{u,\infty}} := \int_{\mathbb{A}^u} y_u(\mathbf{x}_u) \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_{u,\infty}}(\mathbf{x}_u) f_{\mathbf{X}_u}(\mathbf{x})\, \mathrm{d}\mathbf{x}_u = \int_{\mathbb{A}^N} y(\mathbf{x}) \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_{u,\infty}}(\mathbf{x}_u) f_{\mathbf{X}}(\mathbf{x})\, \mathrm{d}\mathbf{x},\ \mathbf{i}_u \in \overline{\mathcal{I}}_{u,\infty},$$

are the expansion coefficients. Here, the integral in the second equality forms when Equation (5c) is applied to the first integral. Finally, Equations (5c) and (8) can be combined to obtain the Fourier spline expansion

$$y(\mathbf{X}) \sim y_\emptyset + \sum_{\emptyset \neq u \subseteq \{1,\ldots,N\}} \sum_{\mathbf{i}_u \in \overline{\mathcal{I}}_{u,\infty}} C^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_{u,\infty}} \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_{u,\infty}}(\mathbf{X}_u), \tag{9}$$

where the constant component function $y_\emptyset$ is already defined in Equation (5b). The expansion in Equation (9) behaves like a Fourier series and is referred to as SDD in this article. According to Equation (9), the SDD of any random

variable $y(\mathbf{X}) \in L^2(\Omega, \mathcal{F}, \mathbb{P})$ is a dimensionwise orthogonal projection onto the spline space spanning the set of associated measure-consistent multivariate orthonormalized B-splines.

The connection between SDD in Equation (9) and ADD in Equations (5a)–(5c) is obvious, as the former is derived by exploiting the spline adaptation of the latter. In addition, because of the *zero*-mean value and orthonormality of basis functions, as described in Equations (3) and (4), SDD inherits all the desirable properties of ADD, including the second-moment properties of $y_u(\mathbf{X}_u)$ described in Equations (6) and (7).

4.3 | Truncation of SDD

In a practical setting, the lengths of all knot vectors, and therefore the total number of basis functions, are finite. In this case, a truncated set $\{\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{X}_u) : \mathbf{i}_u \in \overline{I}_{u,\mathbf{n}_u}\}$ is used to approximate $y(\mathbf{X})$, resulting in the SDD approximation

$$y_{\mathbf{p},\Xi}(\mathbf{X}) := y_\emptyset + \sum_{\emptyset \neq u \subseteq \{1,\dots,N\}} \sum_{\mathbf{i}_u \in \overline{I}_{u,\mathbf{n}_u}} C^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u} \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{X}_u) \tag{10}$$

with the expansion coefficients

$$C^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u} := \int_{\mathbb{A}^N} y(\mathbf{x}) \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{x}_u) f_{\mathbf{X}}(\mathbf{x})\, \mathrm{d}\mathbf{x}, \quad \mathbf{i}_u \in \overline{I}_{u,\mathbf{n}_u}. \tag{11}$$

From Equations (10) and (11), there are

$$L_{\mathbf{p},\Xi} = 1 + \sum_{\emptyset \neq u \subseteq \{1,\dots,N\}} \prod_{k \in u} (n_k - 1) = \prod_{k=1}^{N} n_k \tag{12}$$

basis functions. By inspection of Equation (12), it is clear that the SDD approximation suffers from the curse of dimensionality if all terms in Equation (10) are retained. However, many higher-variate interaction terms of SDD contribute only a negligible amount to the function value and can therefore be safely ignored. Due to the dimensional hierarchical structure of SDD, this can be done by simply keeping all basis functions in at most $1 \leq S \leq N$ variables, thereby retaining the degrees of interaction among input variables less than or equal to S. The result is an S-variate SDD approximation

$$y_{S,\mathbf{p},\Xi}(\mathbf{X}) := y_\emptyset + \sum_{\substack{\emptyset \neq u \subseteq \{1,\dots,N\} \\ 1 \leq |u| \leq S}} \sum_{\mathbf{i}_u \in \overline{I}_{u,\mathbf{n}_u}} C^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u} \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{X}_u) \tag{13}$$

of $y(\mathbf{X})$, comprising

$$L_{S,\mathbf{p},\Xi} = 1 + \sum_{\substack{\emptyset \neq u \subseteq \{1,\dots,N\} \\ 1 \leq |u| \leq S}} \prod_{k \in u} (n_k - 1) \leq \prod_{k=1}^{N} n_k = L_{\mathbf{p},\Xi} \tag{14}$$

basis functions. When $S = 1$ or $S = 2$, the resulting SDD approximations are referred to as univariate and bivariate SDD approximations, respectively. In such cases, the functions $y_{1,\mathbf{p},\Xi}(\mathbf{X})$ or $y_{2,\mathbf{p},\Xi}(\mathbf{X})$ should not be interpreted as first- and second-order approximations, as S does not limit the accuracy of SDD in capturing the potential nonlinearity of $y(\mathbf{X})$. On the contrary, depending on how the orders and knot vectors are chosen, arbitrarily high-order univariate and bivariate terms of $y(\mathbf{X})$, including discontinuity and nonsmoothness, could be lurking inside $y_{1,\mathbf{p},\Xi}(\mathbf{X})$ or $y_{2,\mathbf{p},\Xi}(\mathbf{X})$.

The SDD approximations in Equations (10) and (13) are both convergent to the correct limit in mean-square, in probability, and in distribution. Readers interested in formal proofs are directed to a prior work.[8]

4.4 | Computational expense

Due to matching hierarchical structures of function decompositions, the SDD method is endowed with the same order of computational complexity as the existing PDD method. In general, if $S \ll N$, as is often the case in practical applications,

the number of basis functions in the S-variate SDD approximation decreases rapidly, resulting in significant computational savings. To explain the scalability of SDD with respect to the stochastic dimension N, consider the univariate ($S = 1$) approximation $y_{1,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ and bivariate ($S = 2$) approximation $y_{2,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ of $y(\mathbf{X})$. In either approximation, the required computational effort can be assessed by the associated number of basis functions involved. For instance, there are

$$L_{1,\mathbf{p},\boldsymbol{\Xi}} = 1 + \sum_{k=1}^{N} (n_k - 1)$$

and

$$L_{2,\mathbf{p},\boldsymbol{\Xi}} = 1 + \sum_{k=1}^{N} (n_k - 1) + \sum_{k_1=1}^{N-1} \sum_{k_2=k_1+1}^{N} (n_{k_1} - 1)(n_{k_2} - 1)$$

basis functions involved in forming $y_{1,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ and $y_{2,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$, respectively. Hence, given the values of $n_k, k = 1, \ldots, N$, which are determined by $\mathbf{p}$ and $\boldsymbol{\Xi}$, the computational effort with respect to N increases linearly for univariate approximation and quadratically for bivariate approximation. For example, when $N = 12$ and $n_1 = \cdots = n_{12} = 6$, the univariate and bivariate SDD approximations involve 61 and 1711 basis functions, respectively. By contrast, the numbers of basis functions in the SCE and tensor-product-truncated PCE approximations with six bases in each direction both jump to 6^{12}, which is substantially larger than that required by either of the two SDD approximations. In general, from Equation (14), the computational effort by an S-variate SDD approximation scales S-degree-polynomially with respect to N. Therefore, the computational complexity of a truncated SDD is polynomial, as opposed to exponential, thereby alleviating the curse of dimensionality by a substantial magnitude.

4.5 | Output statistics and other properties

The S-variate SDD approximation $y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ can be treated as an inexpensive surrogate of a function $y(\mathbf{X})$, which may require expensive computation to evaluate. Accordingly, the relevant statistical properties of $y(\mathbf{X})$, such as its first two moments, can be estimated by those of $y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$.

Evaluating the expectation of $y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ in Equation (13) and recognizing Equation (3), its mean

$$\mathbb{E}\left[y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] = y_{\emptyset} = \mathbb{E}\left[y(\mathbf{X})\right] \tag{15}$$

is independent of S, $\mathbf{p}$, and $\boldsymbol{\Xi}$. More importantly, the SDD approximation always yields the exact mean, provided that the expansion coefficient $y_{\emptyset}$ is determined exactly.

Applying the expectation operator on $[y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X}) - y_{\emptyset}]^2$ and using Equation (4) results in the variance

$$\operatorname{var}\left[y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})\right] = \sum_{\substack{\emptyset \neq u \subseteq \{1,\ldots,N\} \\ 1 \leq |u| \leq S}} \sum_{\mathbf{i}_u \in \overline{I}_{u,\mathbf{n}_u}} \left(C^{u}_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}\right)^2 \leq \operatorname{var}\left[y(\mathbf{X})\right] \tag{16}$$

of $y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$. Therefore, the second-moment statistics of an SDD approximation in Equations (15) and (16) are determined based on a reduced set of expansion coefficients corresponding to at most S-variate interactions between input variables. The formulas for the mean and variance of the SDD approximation are identical to those established for the PDD approximation, although the respective expansion coefficients and corresponding basis functions are not. The fundamental reason for this similarity is rooted in the use of hierarchically ordered orthonormal basis functions in both decompositions.

Being convergent in probability and in distribution, the CDF and PDF of $y(\mathbf{X})$, if they exist, can also be estimated economically by resampling $y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$. This will be illustrated in a numerical example.

4.6 | A single-index version

While the compact notations enable a concise description of SDD, a version using a single index notation should impart an alternative interpretation of the corresponding approximation. Besides, the single-index version allows a simpler explanation of the calculation of expansion coefficients, to be discussed in the next section.

Recall that the set

$$\{\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{X}_u) : 1 \le |u| \le S,\ \mathbf{i}_u \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}\},$$

consisting of $L_{S,\mathbf{p},\Xi}$ basis functions, is required to construct the S-variate SDD approximation $y_{S,\mathbf{p},\Xi}(\mathbf{X})$ of $y(\mathbf{X})$. Given $\Psi_1(\mathbf{X};\Xi) = 1$ but otherwise with an arbitrary order of choice, arrange the elements of the set by

$$\left\{\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{X}_u) : 1 \le |u| \le S,\ \mathbf{i}_u \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}\right\} = \left\{\Psi_2(\mathbf{X};\Xi),\ \dots,\Psi_{L_{S,\mathbf{p},\Xi}}(\mathbf{X};\Xi)\right\},\quad \Psi_1(\mathbf{X};\Xi) = 1,$$

such that $\Psi_i(\mathbf{X};\Xi)$, $i = 1,\ \dots,L_{S,\mathbf{p},\Xi}$, represents the ith basis function of the aforementioned set. In this way, the same basis functions in the set can be indexed with a single integer i. Obviously, $\Psi_i(\mathbf{X};\Xi)$ also depends on $\mathbf{p}$, but the latter symbol is suppressed for brevity.

Associated with each $i = 1,\ \dots,L_{S,\mathbf{p},\Xi}$, denote by $C_i(\Xi) \in \mathbb{R}$ the ith SDD coefficient. As a result, the S-variate SDD approximation can also be written as

$$y_{S,\mathbf{p},\Xi}(\mathbf{X}) := \sum_{i=1}^{L_{S,\mathbf{p},\Xi}} C_i(\Xi)\Psi_i(\mathbf{X};\Xi)$$

with the expansion coefficients

$$C_i(\Xi) := \int_{\mathbb{A}^N} y(\mathbf{x})\Psi_i(\mathbf{X};\Xi)f_{\mathbf{X}}(\mathbf{x})\,\mathrm{d}\mathbf{x},\quad i = 1,\ \dots,L_{S,\mathbf{p},\Xi}. \tag{17}$$

Thenceforth, the mean and variances of $y_{S,\mathbf{p},\Xi}(\mathbf{X})$ are calculated from the expansion coefficients as

$$\mathbb{E}\left[y_{S,\mathbf{p},\Xi}(\mathbf{X})\right] = C_1(\Xi) = \mathbb{E}\left[y(\mathbf{X})\right]$$

and

$$\mathrm{var}\left[y_{S,\mathbf{p},\Xi}(\mathbf{X})\right] = \sum_{i=2}^{L_{S,\mathbf{p},\Xi}} C_i^2(\Xi) \le \mathrm{var}\left[y(\mathbf{X})\right], \tag{18}$$

respectively.

5 | CALCULATION OF EXPANSION COEFFICIENTS

The determination of the SDD coefficients $C_i(\Xi)$, $i = 1,\ \dots,L_{S,\mathbf{p},\Xi}$, involves various high-dimensional integrations. For an arbitrary function y and an arbitrary probability distribution of random input $\mathbf{X}$, an exact evaluation of these coefficients from definition alone is impossible. Two approximate methods are proposed to estimate the coefficients as follows.

5.1 | Numerical integration

A natural instinct is to approximate the coefficients by numerical integration, for instance, by a general anisotropic $(Q_{k_1},\ \dots,Q_{k_N})$-point, multivariate, tensor-product Gauss-type quadrature rule with $Q_{k_1},\ \dots,Q_{k_N} \in \mathbb{N}$, yielding the estimate

$$\overline{C}_i(\Xi) = \underbrace{\sum_{j_1=1}^{Q_1}\cdots\sum_{j_N=1}^{Q_N}}_{N\ \text{sums}} y(x_1^{(j_1)},\ \dots,x_N^{(j_N)})\Psi_i(x_1^{(j_1)},\ \dots,x_N^{(j_N)};\Xi)\prod_{k=1}^{N} w_k^{(j_k)} \tag{19}$$

of C_i, where, for each $k = 1, \ldots, N$, $x_k^{(j_k)}$ and $w_k^{(j_k)}$ are the integration points and matching weights, consistent with the probability measure $f_{X_k}(x_k)\,\mathrm{d}x_k$. As is the case for computing the spline moment matrix by numerical integration, it is generally best practice to split the integrals in each direction k into I_k integrals between the corresponding distinct knots. In real-life applications, the output function y is often determined algorithmically by performing time-consuming finite-element analysis (FEA), isogeometric analysis (IGA), or other numerical calculations. Clearly, for a high-dimensional problem, say, with N exceeding 10, evaluating the N-dimensional sums in Equation (19) is computationally formidable and likely prohibitive. Furthermore, if any alterations are made to any of the knot vectors, the output function must be recomputed at an entirely different set of integration points, further increasing computational cost. While methods, such as the dimension-reduction techniques,[25] have been developed to tackle the former issue, they do not address the latter. Therefore, a practical alternative to numerical integration, such as regression analysis, is often necessary to estimate these coefficients.

5.2 | Least-squares regression

The standard least-squares (SLS) regression is predicated on the optimal approximation quality of the SDD approximation. In other words, the S-variate SDD approximation $y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ is the best approximation of $y(\mathbf{X})$ in the sense that

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

Thereafter, the approximate expansion coefficients of $y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$ are determined from the minimization of

$$\mathbb{E}\left[y(\mathbf{X}) - \sum_{i=1}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} C_i(\boldsymbol{\Xi})\Psi_i(\mathbf{X};\boldsymbol{\Xi})\right]^2. \tag{20}$$

Given a UQ problem with known distribution of random input $\mathbf{X}$ and an output function $y : \mathbb{A}^N \to \mathbb{R}$, consider an input–output data set $\{\mathbf{x}^{(l)}, y(\mathbf{x}^{(l)})\}_{l=1}^{L}$ of size $L \in \mathbb{N}$. The mapping y can be as simple as an explicitly defined mathematical function or as intricate as an implicitly described function obtained via computational simulation, such as FEA or IGA of complex mechanical systems. In either case, the data set, often referred to as the experimental design, can be generated by calculating the function $y(\mathbf{x}^{(l)})$ at each input data point $\mathbf{x}^{(l)}$. Various sampling methods, namely, standard MCS, quasi MCS, and Latin hypercube sampling, can be used to build the experimental design.

According to SLS, the expansion coefficients of the S-variate SDD approximation are estimated by minimizing

$$\hat{e}_{S,\mathbf{p},\boldsymbol{\Xi}} := \frac{1}{L}\sum_{l=1}^{L}\left[y(\mathbf{x}^{(l)}) - \sum_{i=1}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} C_i(\boldsymbol{\Xi})\Psi_i(\mathbf{x}^{(l)};\boldsymbol{\Xi})\right]^2,$$

which is an empirical analogue of Equation (20). Denote by

$$\mathbf{A} := \begin{bmatrix} \Psi_1(\mathbf{x}^{(1)};\boldsymbol{\Xi}) & \ldots & \Psi_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\mathbf{x}^{(1)};\boldsymbol{\Xi}) \\ \vdots & \ddots & \vdots \\ \Psi_1(\mathbf{x}^{(L)};\boldsymbol{\Xi}) & \ldots & \Psi_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\mathbf{x}^{(L)};\boldsymbol{\Xi}) \end{bmatrix} \text{ and } \mathbf{b} := \left(y(\mathbf{x}^{(1)}), \ldots, y(\mathbf{x}^{(L)})\right)^{\intercal} \tag{21}$$

an $L \times L_{S,\mathbf{p},\boldsymbol{\Xi}}$ matrix and an L-dimensional column vector comprising evaluations of the orthonormal spline basis functions and output function at the data points, respectively. Then, the estimated coefficients $\hat{C}_i(\boldsymbol{\Xi})$, $i = 1, \ldots, L_{S,\mathbf{p},\boldsymbol{\Xi}}$, are obtained as

$$\hat{\mathbf{c}}(\boldsymbol{\Xi}) := \left(\hat{C}_1(\boldsymbol{\Xi}), \ldots, \hat{C}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\boldsymbol{\Xi})\right)^{\intercal} = (\mathbf{A}^{\intercal}\mathbf{A})^{-1}\mathbf{A}^{\intercal}\mathbf{b}. \tag{22}$$

Here, $\mathbf{A}^{\intercal}\mathbf{A}$ is an $L_{S,\mathbf{p},\boldsymbol{\Xi}} \times L_{S,\mathbf{p},\boldsymbol{\Xi}}$ matrix, often referred to as the information or data matrix. A necessary condition for the SLS solution is $L > L_{S,\mathbf{p},\boldsymbol{\Xi}}$, that is, the data size must be larger than the number of coefficients. Even when this condition is satisfied, the experimental design must be judiciously selected in such a way that the information matrix is well conditioned.

6 | OPTIMAL SPLINE DIMENSIONAL DECOMPOSITION

The SDD method presented in Section 4 is general in the sense that the knot vectors are yet to be specified, meaning that they can be chosen freely. While such a freedom is normally welcome, an indiscriminate selection of knot vectors may also produce inadequate or weak basis functions, vitiating the approximation quality of SDD. A standard approach in selecting the basis functions entails uniformly spaced knots. Such an approach makes sense when there is no prior knowledge of the behavior of the function being approximated. By contrast, a potentially better alternative is possible if the information from the function is propagated to the knot vectors, which are subsequently derived in an optimal fashion. Indeed, this can be achieved by minimizing the error of the SDD approximation in the mean-square sense, as is the primary objective in calibrating the SDD coefficients. However, there are several key differences between the optimization of the coefficients and that of the knot locations.

6.1 | An overview of the optimization problem

The optimization of the knot locations is considerably more complex than that of the SDD coefficients. Firstly, the dependence of the mean-square error of the SDD approximation on the knot locations is, in general, highly nonlinear, in contrast to its dependence on the coefficients. Thus, an iterative procedure is required to derive an optimal solution. Furthermore, given that the knot sequences in each coordinate direction are nondecreasing by definition, the optimal solution space must be restricted by a corresponding set of linear inequality constraints. In this work, it will be assumed for simplicity that the internal knots in all coordinate directions are simple, that is, they each have a multiplicity of *one*. Thus, to avoid coincident knots, the aforementioned constraints are implemented by enforcing a user-specified minimum separation distance $\delta > 0$ between any two adjacent knots. These constraints will be explicitly defined later in each of the alternative optimization approaches described in the next two sections.

Although the derivations of the expansion coefficients and optimal knot locations both entail a minimization of the mean-square error between the series expansion and the output function, the two corresponding optimization subproblems are uncoupled, thereby mitigating the complexity introduced by the knot sequence optimization, to the extent possible. This may seem counterintuitive, given that the dependence of the splines on the knot locations is imparted onto the expansion coefficients. Nevertheless, for a given UQ problem with a fixed selection of spline degrees $\mathbf{p}$, the solution of the coefficients can be uniquely determined by a set of knot vectors $\boldsymbol{\Xi}$, whether or not they are optimally selected. Consequently, the mean-square error can effectively be reduced to a function of $\boldsymbol{\Xi}$ alone when the expansion coefficients are optimally derived. Hence, the minimization of the mean-square error entails a constrained nonlinear optimization with respect to the knot locations, wherein the expansion coefficients are computed during each iteration such that the mean-square error between the output function and its surrogate is minimized for each successive guess of the family of optimal knot vectors, denoted by $\boldsymbol{\Xi}^* = (\boldsymbol{\xi}_1^*, \ \dots \ , \boldsymbol{\xi}_N^*)$.

In this work, the SDD methods entailing uniformly spaced knots and optimally selected knots are referred to as the standard SDD and optimal SDD methods, respectively. There are at least two approaches to generate optimal knots for the S-variate optimal SDD approximation, depending on whether or not the basis functions are orthonormalized. While SDD is fundamentally rooted in an orthonormal basis of the spline space, as explained in Section 3, a nonorthonormalized basis is also possible for a similar expansion. Therefore, both approaches, described in the next two sections, will be studied and tested during numerical implementation. The optimization problems formulated in either of these approaches are traditionally solved by employing a suitable gradient-based method, such as the well-known sequential linear programming (SLP) and sequential quadratic programming (SQP) methods. In this regard, the gradients of the corresponding objective functions with respect to the knot locations can be computed using the relevant equations provided in Appendix B.

6.2 | Optimal knots with orthonormal basis

In the first approach, the optimal knots are generated using the orthonormal basis set

$$\left\{ \Psi_1(\mathbf{X}; \boldsymbol{\Xi}), \ \dots \ , \Psi_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\mathbf{X}; \boldsymbol{\Xi}) \right\}, \ \Psi_1(\mathbf{X}) = 1,$$

as described in Section 4, for the function expansion. Define

$$e_{L_{S,\mathbf{p},\boldsymbol{\Xi}}} := \mathbb{E}\left[y(\mathbf{X}) - \sum_{i=1}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} C_i(\boldsymbol{\Xi})\Psi_i(\mathbf{X};\boldsymbol{\Xi})\right]^2$$

to be the mean-square error committed by $y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$. Here, $C_i(\boldsymbol{\Xi})$ and $\Psi_i(\mathbf{X};\boldsymbol{\Xi})$ are the ith coefficient and basis function of the SDD approximation, respectively, each depending on the family $\boldsymbol{\Xi} = (\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_N)$ of knot vectors in all coordinate directions. The definition and construction of the coefficients and basis functions are already described in Section 4 for an arbitrary $\boldsymbol{\Xi}$. Thereafter, the determination of the optimal knot vectors requires one to

$$\begin{aligned}&\min_{\boldsymbol{\Xi}}\ e_{L_{S,\mathbf{p},\boldsymbol{\Xi}}} := \mathbb{E}\left[y(\mathbf{X}) - \sum_{i=1}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} C_i(\boldsymbol{\Xi})\Psi_i(\mathbf{X};\boldsymbol{\Xi})\right]^2,\\ &\text{subject to } \xi_{k,i_k} - \xi_{k,i_k-1} \geq \delta,\ i_k = p_k+2, \ldots, n_k+1,\ k = 1, \ldots, N.\end{aligned} \tag{23}$$

The objective function on the right-hand side of Equation (23) can be further reduced by expanding the squared term inside the expectation operator and invoking the orthonormal properties of the basis functions. In so doing,

$$e_{L_{S,\mathbf{p},\boldsymbol{\Xi}}} = \mathbb{E}\left[y^2(\mathbf{X})\right] - \sum_{i=1}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} C_i^2(\boldsymbol{\Xi}),$$

where the second term represents the second raw moment of $y_{S,\mathbf{p},\boldsymbol{\Xi}}(\mathbf{X})$. Since $C_1(\boldsymbol{\Xi}) = \mathbb{E}[y(\mathbf{X})]$ and $\mathbb{E}[y^2(\mathbf{X})]$ do not depend on $\boldsymbol{\Xi}$, the optimization problem is reduced to

$$\begin{aligned}&\min_{\boldsymbol{\Xi}}\ \tilde{e}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}} := -\sum_{i=2}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} C_i^2(\boldsymbol{\Xi}),\\ &\text{subject to } \xi_{k,i_k} - \xi_{k,i_k-1} \geq \delta,\ i_k = p_k+2, \ldots, n_k+1,\ k = 1, \ldots, N.\end{aligned} \tag{24}$$

When the output function is simple and low-dimensional, the expectations in Equation (23) or (24) can be determined analytically or numerically, say, using Gauss quadrature, as explained in Section 5. In this case, Equation (23) or (24) is readily solved, as will be demonstrated in one numerical example.

For practical applications involving high-dimensional functions, a discrete version of Equation (23) or (24) is necessary. Furthermore, the SDD expansion coefficients are estimated by an approximate method, say, by the SLS regression. In this case, the optimization problem in Equation (23) is simplified to

$$\begin{aligned}&\min_{\boldsymbol{\Xi}}\ \hat{e}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}} := \frac{1}{L}\sum_{l=1}^{L}\left[y(\mathbf{x}^{(l)}) - \sum_{i=1}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} \hat{C}_i(\boldsymbol{\Xi})\Psi_i(\mathbf{x}^{(l)};\boldsymbol{\Xi})\right]^2,\\ &\text{subject to } \xi_{k,i_k} - \xi_{k,i_k-1} \geq \delta,\ i_k = p_k+2, \ldots, n_k+1,\ k = 1, \ldots, N,\end{aligned} \tag{25}$$

where the estimated SDD coefficients $\hat{C}_i(\boldsymbol{\Xi})$, $i = 1, \ldots, L_{S,\mathbf{p},\boldsymbol{\Xi}}$, are obtained from Equation (22).

The primary advantage of the first approach in formulating the optimization problem in Equation (25) lies in a well-conditioned information matrix $\mathbf{A}^\intercal\mathbf{A}$. The result is an efficient and stable estimation of the SDD coefficients. This is obviously due to the orthonormal basis employed in forming the information matrix, which is symmetric, diagonally dominant, and sparse. It is important to note that, due to the nature of the SLS approximation, the SDD estimate of the variance of $y(\mathbf{X})$ may exceed the exact variance in some cases, which is not the case when the coefficients are calculated by integration, as stated in Equation (18). Consequently, the knot vectors should be optimized according to Equation (25), rather than Equation (24), when obtaining the coefficients by SLS regression, to avoid convergence issues. Furthermore, as $\boldsymbol{\Xi}$ evolves during the iterations of the knot vector optimization, there is a tedious need to recalculate the whitening matrix during the orthonormalization process. Therefore, the first approach is appropriate when the whitening matrix is produced accurately and efficiently.

6.3 | Optimal knots with nonorthonormal basis

In the second approach, the optimal knots are generated employing the nonorthonormal basis set for the function expansion. They are obtained from a similar dimensionwise tensorization of univariate B-splines, as done in Equation (1), yielding

$$\Phi^{u}_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{x}_u) = \prod_{k\in u} B^{k}_{i_k,p_k,\xi_k}(x_k) = \prod_{l=1}^{|u|} B^{k_l}_{i_{k_l},p_{k_l},\xi_{k_l}}(x_{k_l}), \quad \mathbf{i}_u = (i_{k_1},\, \dots\, , i_{k_{|u|}}) \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}.$$

Then, arrange the elements of the set by

$$\left\{ \Phi^{u}_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{X}_u) \,:\, 1 \le |u| \le S,\ \mathbf{i}_u \in \overline{\mathcal{I}}_{u,\mathbf{n}_u} \right\} = \left\{ \Phi_2(\mathbf{X};\boldsymbol{\Xi}),\, \dots\, , \Phi_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\mathbf{X};\boldsymbol{\Xi}) \right\}, \quad \Phi_1(\mathbf{X};\boldsymbol{\Xi}) = 1,$$

such that $\Phi_i(\mathbf{X};\boldsymbol{\Xi})$, $i = 1,\, \dots\, , L_{S,\mathbf{p},\boldsymbol{\Xi}}$, represents the ith nonorthonormal basis function of the aforementioned set. In this manner, the same basis functions in the set can also be indexed with a single integer i. Again, $\Phi_i(\mathbf{X};\boldsymbol{\Xi})$ also depends on $\mathbf{p}$, but the latter symbol is dropped for conciseness.

Henceforth, the optimization problem from Equation (23), but formulated with respect to the nonorthonormal basis, is rephrased to

$$\begin{aligned} \min_{\boldsymbol{\Xi}}\ & \bar{e}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}} := \mathbb{E}\left[y(\mathbf{X}) - \sum_{i=1}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} \overline{C}_i(\boldsymbol{\Xi})\Phi_i(\mathbf{X};\boldsymbol{\Xi}) \right]^2, \\ \text{subject to}\ & \xi_{k,i_k} - \xi_{k,i_k-1} \ge \delta,\ i_k = p_k + 2,\, \dots\, , n_k + 1,\ k = 1,\, \dots\, , N, \end{aligned} \tag{26}$$

where $\overline{C}_j$, $j = 1,\, \dots\, , L_{S,\mathbf{p},\boldsymbol{\Xi}}$, are new expansion coefficients relative to nonorthonormal basis functions. However, the expansion in Equation (26) does not lead to SDD, as the basis is no longer orthonormal with respect to the probability measure of input random variables. Consequently, there does not exist a simplified variant of the optimization problem, as recognized by Equation (24).

Finally, the discrete version of Equation (26) requires one to

$$\begin{aligned} \min_{\boldsymbol{\Xi}}\ & \hat{\bar{e}}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}} := \frac{1}{L}\sum_{l=1}^{L} \left[y(\mathbf{x}^{(l)}) - \sum_{i=1}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} \hat{\overline{C}}_i(\boldsymbol{\Xi})\Phi_i(\mathbf{x}^{(l)};\boldsymbol{\Xi}) \right]^2, \\ \text{subject to}\ & \xi_{k,i_k} - \xi_{k,i_k-1} \ge \delta,\ i_k = p_k + 2,\, \dots\, , n_k + 1,\ k = 1,\, \dots\, , N. \end{aligned} \tag{27}$$

Then, the coefficients $\hat{\overline{C}}_i(\boldsymbol{\Xi})$, $i = 1,\, \dots\, , L_{S,\mathbf{p},\boldsymbol{\Xi}}$, are estimated from the SLS regression as

$$\hat{\overline{\mathbf{c}}}(\boldsymbol{\Xi}) := \left(\hat{\overline{C}}_1(\boldsymbol{\Xi}),\, \dots\, , \hat{\overline{C}}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\boldsymbol{\Xi}) \right)^{\intercal} = \left(\overline{\mathbf{A}}^{\intercal}\overline{\mathbf{A}} \right)^{-1} \overline{\mathbf{A}}^{\intercal}\mathbf{b},$$

where

$$\overline{\mathbf{A}} := \begin{bmatrix} \Phi_1(\mathbf{x}^{(1)};\boldsymbol{\Xi}) & \dots & \Phi_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\mathbf{x}^{(1)};\boldsymbol{\Xi}) \\ \vdots & \ddots & \vdots \\ \Phi_1(\mathbf{x}^{(L)};\boldsymbol{\Xi}) & \dots & \Phi_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\mathbf{x}^{(L)};\boldsymbol{\Xi}) \end{bmatrix},$$

and $\mathbf{b}$ was defined in Equation (21). When formulating the optimization problem in Equation (27), the information matrix $\overline{\mathbf{A}}^{\intercal}\overline{\mathbf{A}}$ is constructed using nonorthonormal basis functions. Therefore, it is generally dense and may encounter a high condition number. In consequence, the estimation of the associated expansion coefficients, which involves inversion of the information matrix, is relatively more difficult to obtain in general. Hence, the second approach is relevant primarily for low-order SDD approximations. Nevertheless, the second approach eliminates the need to perform the whitening transformation for each iteration of the knot optimization, rendering the process more computationally expedient than that in the first approach. Once the optimal knot vectors are derived, it is generally useful to orthonormalize the corresponding

basis functions, especially if the first or second moments of the underlying SDD approximation are of interest. This will be discussed in the following section.

It is worth noting that an optimization problem involving nonorthonormal, standard B-spline expansion in a single random variable was applied for data-driven UQ analysis.[26] By contrast, the B-spline expansion in Equation (27) stems from dimensionwise decomposition of a high-dimensional function. In other words, the expansion exploited here is not merely due to a straightforward tensor-product construction of univariate B-splines, which is impractical for addressing high-dimensional UQ problems. This will be more evident when numerical examples are presented.

6.4 | Optimal SDD and output statistics

Once the optimal knot vectors are determined, in conjunction with either the orthonormal (first approach) or nonorthonormal basis (second approach), the optimal S-variate SDD approximation is expressed by

$$y_{S,\mathbf{p},\boldsymbol{\Xi}^*}(\mathbf{X}) = \sum_{i=1}^{L_{S,\mathbf{p},\boldsymbol{\Xi}^*}} C_i(\boldsymbol{\Xi}^*)\Psi_i(\mathbf{X};\boldsymbol{\Xi}^*)$$

with the expansion coefficients

$$C_i(\boldsymbol{\Xi}^*) := \int_{\mathbb{A}^N} y(\mathbf{x})\Psi_i(\mathbf{X};\boldsymbol{\Xi}^*) f_{\mathbf{X}}(\mathbf{x})\, \mathrm{d}\mathbf{x}, \quad i = 1, \ldots, L_{S,\mathbf{p},\boldsymbol{\Xi}^*}.$$

Here, $\Psi_i(\mathbf{X};\boldsymbol{\Xi}^*)$ and $C_i(\boldsymbol{\Xi}^*)$ are the ith orthonormal basis function and ith expansion coefficient of an S-variate SDD approximation generated from a family of optimal knot vectors.

Henceforth, the mean and variances of $y_{S,\mathbf{p},\boldsymbol{\Xi}^*}(\mathbf{X})$ are calculated from the expansion coefficients as

$$\mathbb{E}\left[y_{S,\mathbf{p},\boldsymbol{\Xi}^*}(\mathbf{X})\right] = C_1(\boldsymbol{\Xi}^*) \tag{28}$$

and

$$\mathrm{var}\left[y_{S,\mathbf{p},\boldsymbol{\Xi}^*}(\mathbf{X})\right] = \sum_{i=2}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} C_i^2(\boldsymbol{\Xi}^*), \tag{29}$$

respectively. In practice, the coefficients are estimated, say, by SLS regression, as discussed in Section 6. In this case, the approximate expansion coefficients must be used in Equations (28) and (29).

7 | NUMERICAL EXAMPLES

Two numerical problems are presented to highlight the proposed optimal SDD method for obtaining the second-moment statistics and probability distribution of various responses. Each example follows a particular objective. Example 1 features a two-dimensional UQ problem entailing three mathematical functions and provides a rigorous comparison among the optimal SDD, standard SDD, and a few existing UQ methods from the current literature. Example 2 represents a 16-dimensional UQ problem from linear elasticity. The latter example encompasses stochastic stress analysis of a complex horseshoe geometry in three physical dimensions. The stresses were calculated using the standard method of IGA.[21]

All knot vector optimizations were performed using the well-known method of SQP. The minimum separation distance δ used in the constraints of the knot vector optimization was taken to be 10% of the uniform knot spacing in both examples. Until this point, it was assumed for the sake of generality that the knot sequences in all coordinate directions are chosen independently. However, for a high-dimensional UQ problem, simultaneous optimization of all internal knots in all coordinate directions may be computationally infeasible. For example, if SDD with four subintervals were employed in 50 different coordinate directions, a total of 150 knot locations would need to be optimized. Furthermore, it is often the case in realistic UQ problems that several coordinate directions pertain to input variables that have similar physical

meaning, are identically or similarly distributed, and are expected to have similar influence on the output of interest. In such instances as this, it is recommended that the knot sequences in each coordinate direction be arranged into groups accordingly, and a set of knot sequences associated with each group be optimized, thereby reducing the number of optimization parameters and thus the required computational effort. This was done in the following numerical examples, as will be explained in further detail.

7.1 | Example 1: Three mathematical functions

In the first example, consider three bivariate output functions $y_i : [-1,1]^2 \to \mathbb{R}$, $i = 1,2,3$, of two independent and identically distributed input random variables X_1 and X_2, each with a uniform distribution over $[-1,1]$, as follows:

1. Exponential drop function[8]:

$$y_1(X_1,X_2) = g(X_1) + g(X_2) + \frac{1}{5}g(X_1)g(X_2),$$

where, for $i = 1,2$,

$$g(x_i) = \begin{cases} 1, & -1 \le x_i \le 0, \\ \exp(-10x_i), & 0 < x_i \le 1. \end{cases}$$

2. Continuous function[27]:

$$y_2(X_1,X_2) = \exp(-2|X_1| - 2|X_2|).$$

3. Product peak function[27]:

$$y_3(X_1,X_2) = \frac{1}{(1+X_1^2)(1+X_2^2)}.$$

From the graphs of these functions, depicted in Figure 2, the exponential drop and continuous functions are both nonsmooth, whereas the product peak function is smooth. The objective of this example is to evaluate the accuracy of the proposed optimal SDD method for estimating the second-moment statistics of all three functions in comparison with those obtained from the standard SDD method, PCE, and sparse-grid quadrature. It is important to recognize that some of these functions are well-known Genz functions introduced to examine various numerical integration methods for calculating their means.[27] Here, instead, the variances of these functions are evaluated, which is relatively more challenging to compute. Nonetheless, these functions, although purely mathematical, provide a wonderful prospect to rigorously scrutinize the proposed SDD method before delving into complex engineering problems.

For the SDD approximations, the degrees, numbers of subintervals, and knot vectors in both coordinate directions were chosen to be identical, that is, $p_1 = p_2 = p$ (say), $I_1 = I_2 = I$ (say), and $\boldsymbol{\xi}_1 = \boldsymbol{\xi}_2 = \boldsymbol{\xi}$ (say). Therefore, only one knot vector was optimized, but it was used in both coordinate directions, resulting in a total of $I - 1$ parameters in the optimization. Given the low dimensionality of these functions, the SDD coefficients and exact variances were computed analytically by integration. Hence, the knot vector was optimized according to Equation (24).

7.1.1 | Optimal knots

The optimal knot locations corresponding to the linear ($p = 1$) and quadratic ($p = 2$) SDD approximations, each with five distinct numbers of subintervals $I = 4, 6, 8, 10, 12$ in each coordinate direction, were calculated for all three functions and are displayed in Figure 3(A–C). They were determined by solving the optimization problem in Equation (24) with the SDD coefficients exactly calculated from their definition in Equation (17). As all three functions satisfy the symmetry property, that is, $y_i(x_1,x_2) = y_i(x_2,x_1)$, $i = 1,2,3$, the same respective knot locations were used in both coordinate directions.

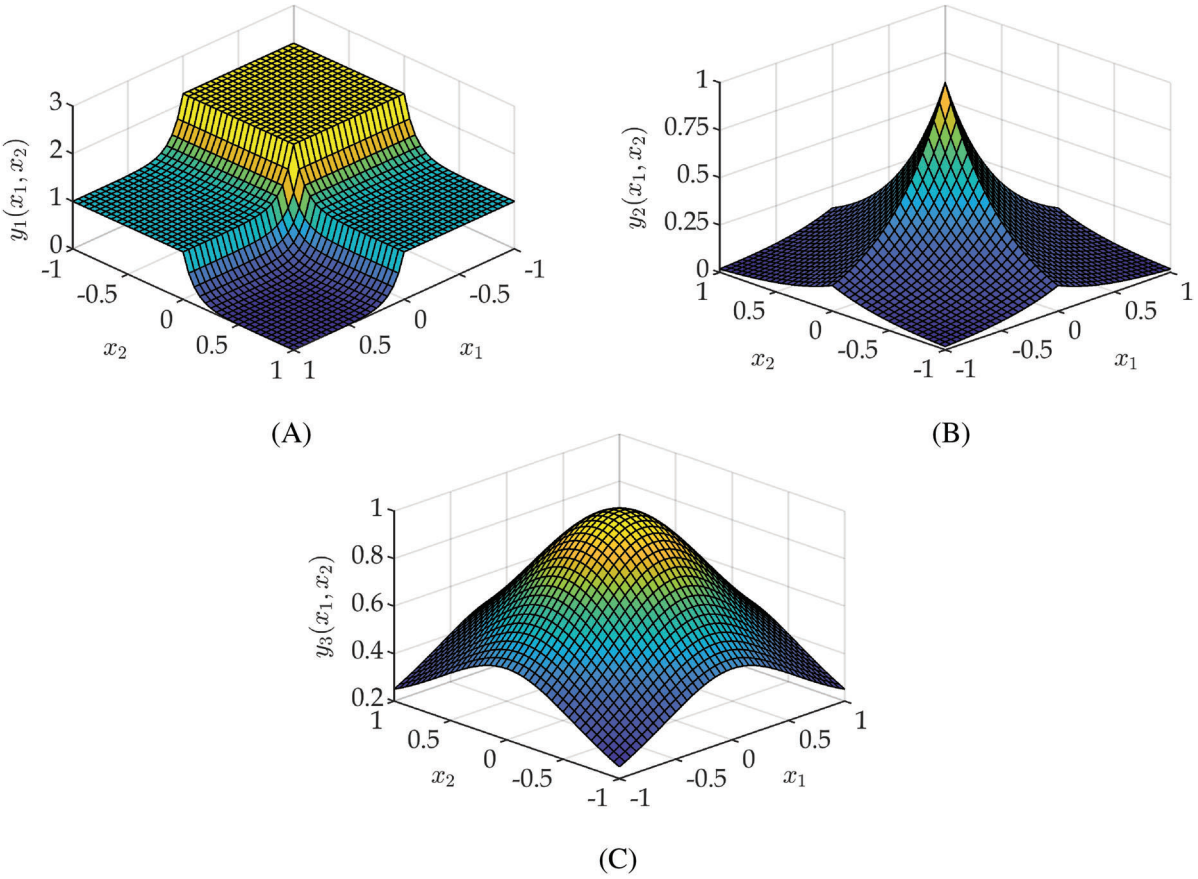


FIGURE 2 Three mathematical functions in Example 1: (A) exponential drop; (B) continuous; and (C) product peak

From Figure 2(A), the exponential drop function is flat for $-1 \leq x_i \leq 0$. Consequently, the objective function is insensitive to the locations of all knots belonging to a subinterval to the left of *zero*, placing the corresponding optimal knots arbitrarily, as depicted in Figure 3(A). For example, for $p = 1,\ I = 4$, the location of the first internal knot (around -0.45) in Figure 3(A) is arbitrary, but the location of the second internal knot (around -0.02) is not, because the subinterval between the second and third internal knots is not entirely to the left of *zero*, since the third internal knot is located around 0.17. The density of the remaining knots, that is the proximity of adjacent knots, is high near *zero* and gradually decreases as x_i increases. However, a phenomenon similar to that occurring with the knots to the left of *zero* can be seen with the last internal knot for $p = 1,\ I = 12$ and $p = 2,\ I = 8, 10$, as well as the last two knots for $p = 2,\ I = 12$; these locations do not follow the pattern seen in the other nonarbitrary knot locations because the exponential drop function decays as x_i increases, therefore the objective function loses sensitivity to the knot locations closer to *one*. The crowding of the knots near *zero* occurs in part because the function changes more abruptly in that region, so smaller subintervals are required there. For $p = 2$, this crowding is particularly pronounced for the two knots located closest to *zero*. This is because there is a discontinuity in the slope of the response at $x_i = 0$, which the quadratic basis requires a repeated knot to capture. However, a limit was placed on how close two adjacent knots could be placed in the optimization process, as shown in the constraints of Equation (24), so these knots were not permitted to coincide. Nevertheless, the near-coincidence of adjacent knots at a location of nonsmoothness in the response is expected in general, which motivates a further modification to the optimization process to be pursued in a future work.

According to Figure 2(B,C), the continuous and product peak functions are also even functions. Therefore, the optimal knot sequences exhibit the same symmetry, as can be seen in Figure 3(B,C). Similar to what occurred with the exponential drop function, the density of the knots were largest near *zero* and decreased in either direction, due to the higher variation of the response near *zero*. The continuous function has a slope discontinuity at $x_i = 0$, where the optimal knots for $p = 2$ are nearly coincident, once again. It is worth noting, however, that although there are *three* consecutive knots that converge in that region, a multiplicity of only *two* would be sufficient to capture the nonsmoothness of the response.

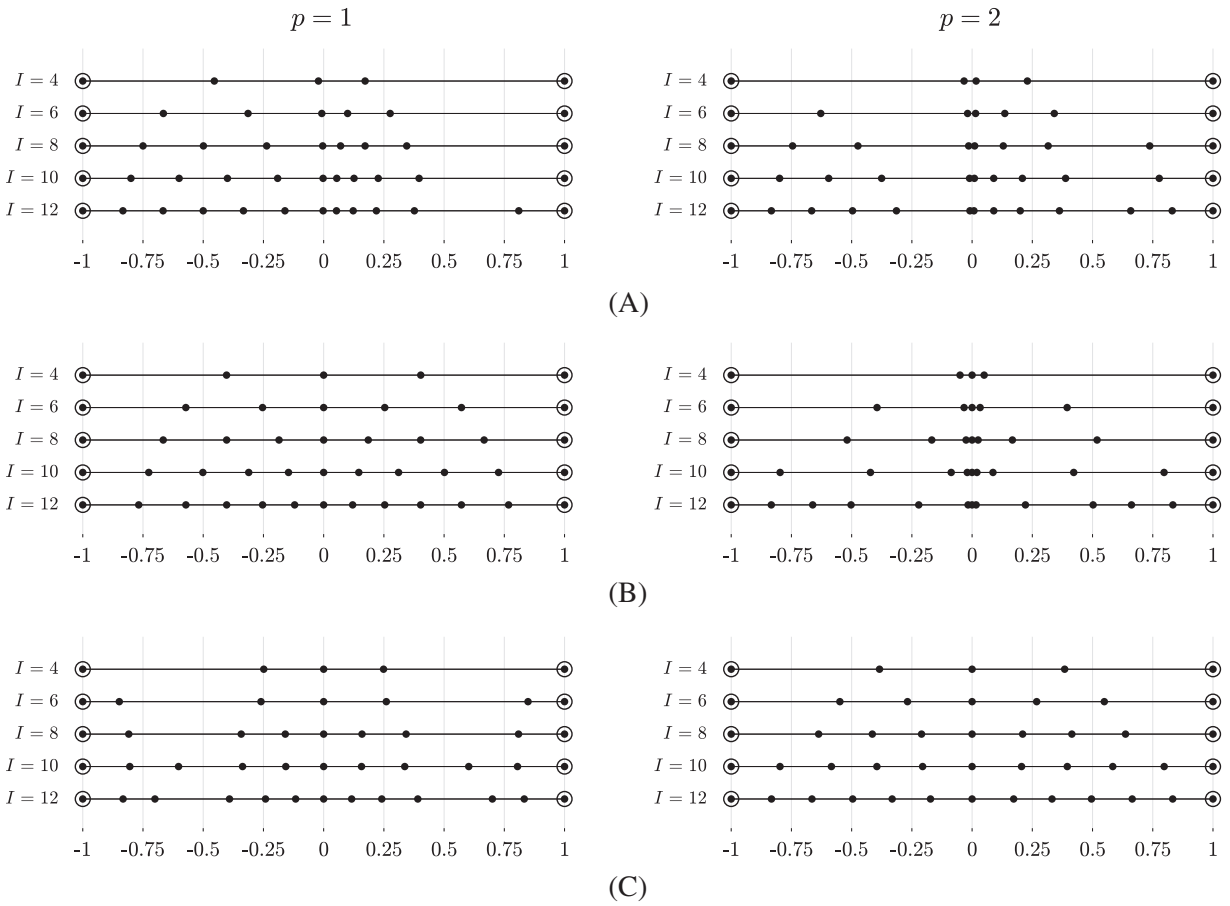


FIGURE 3 Optimal knot spacing for all three functions in Example 1: (A) exponential drop; (B) continuous; and (C) product peak

Due to the smoothness of the product peak function, none of the corresponding optimal knots came close to coinciding. However, there was considerable relocation of the knots to improve the accuracy of the surrogate model. For all three functions, the greatest change in the optimal knot locations relative to the those of the uniformly spaced knots is seen for smaller values of I. This makes sense intuitively, as the knots have more freedom when there are fewer of them, and the corresponding spline expansion is generally of relatively lower accuracy, offering the greatest potential for improvement. This suggests that optimization of the knot locations is most relevant when using relatively few subintervals.

7.1.2 | Error analysis

Given the probability distribution of input variables, the output variances for all three functions were obtained by four different approximate methods: (1) optimal SDD with the knots generated by solving the optimization problem in Equation (24); (2) standard SDD with uniformly spaced knots;[8,9] (3) tensor-product PCE,[1,3] and (4) sparse grids with the Clenshaw–Curtis quadrature rule.[12,13] For the SDD methods, whether optimal or standard, both linear ($p = 1$) and quadratic ($p = 2$) splines were employed. The coefficients of both SDD methods were determined exactly. The variances of optimal SDD and standard SDD were calculated using Equations (18) and (29), respectively. To grade the approximation quality of each method, the exact variances of the exponential drop, continuous, and product peak functions were calculated to be 0.550270499, $2.52956894 \times 10^{-2}$, and $3.25578478 \times 10^{-2}$, respectively.

Figure 4(A–C) describes how the absolute errors in the variance, calculated by the aforementioned methods, decay with respect to the number of basis functions for the exponential drop, continuous, and product peak functions, respectively. According to Figure 4(A,B), the standard SDD method appears to perform about as well as PCE, which consistently outperforms sparse grids. The sharp slope discontinuity of the continuous function causes the standard linear SDD to perform better than both the standard quadratic SDD and PCE methods, as is apparent in Figure 4(B). More importantly,

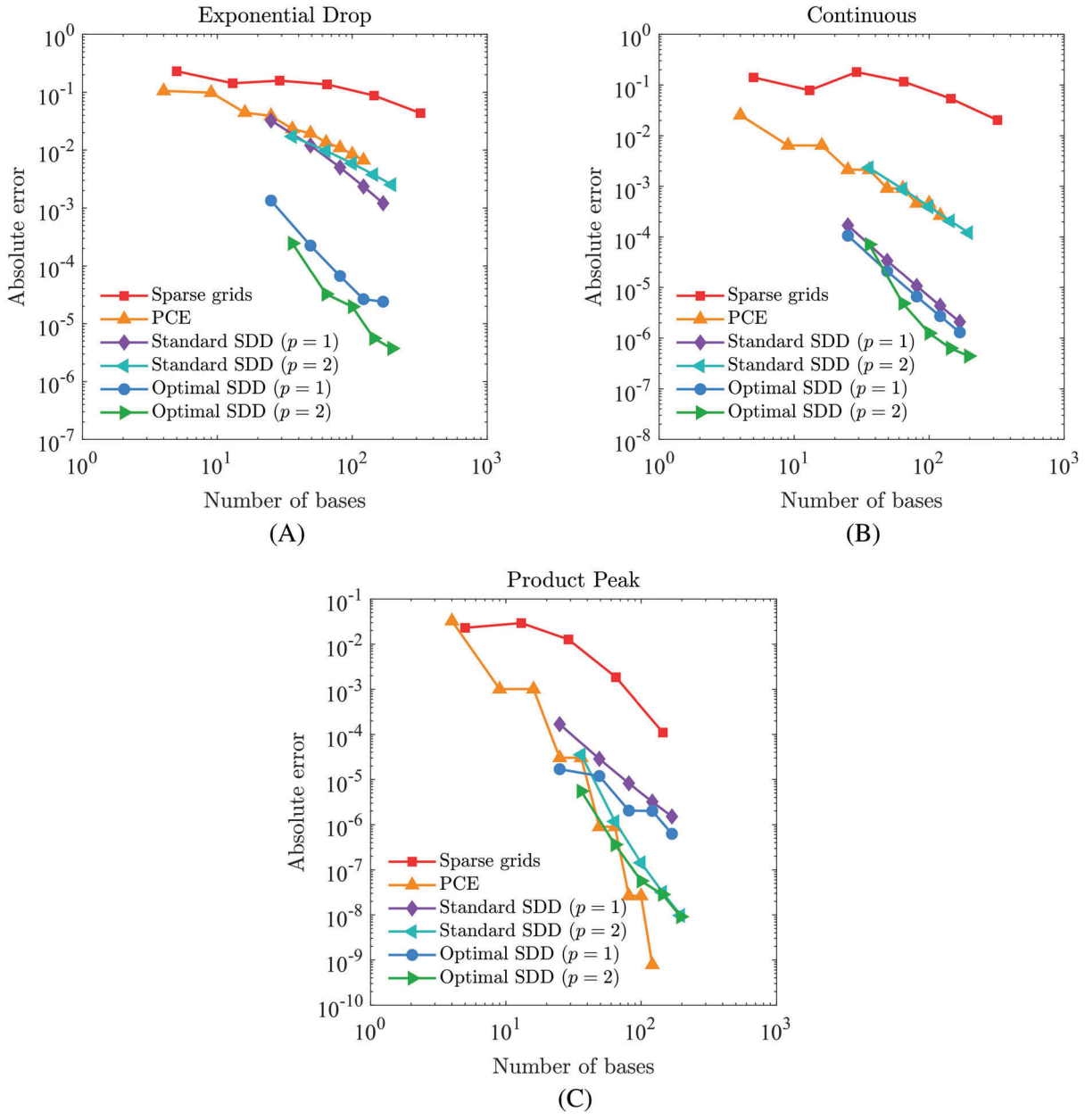


FIGURE 4 Errors in variances for all three functions in Example 1: (A) exponential drop; (B) continuous; and (C) product peak

optimization of the knot locations causes substantial decrease in the variance error for the exponential drop and continuous functions in comparison with sparse-grid quadrature, PCE, and standard SDD. What is particularly noteworthy is that the optimization led to nearly repeated knots for the quadratic optimal SDD approximation of the continuous function, allowing it to capture the slope discontinuity more closely and offering a marked improvement over the standard quadratic SDD approximation. This indicates that the proposed optimization scheme is a promising method to improve the accuracy of the existing methods, including the standard SDD approximation, when the response is not smooth. On the other hand, as can be seen in Figure 4(C), use of optimal knots offers only marginal improvement over the corresponding SDD approximations with uniform knots when the response is smooth, although most methods perform quite well. Hence, an SDD, whether optimal or standard, may not be needed in this case. It is interesting to note that the convergence rates of the linear and quadratic SDD approximations appear to be roughly equal for both the standard and optimal knot sequences. Furthermore, the basis functions of odd orders do not contribute to the PCE approximations for continuous and product peak functions, as they are even functions. Overall, numerical evidence reveals significant improvement in the accuracy of the SDD method by optimization of the knot locations.

7.2 | Example 2: A horseshoe under asymmetric loading

The second example tackles a linear elasticity problem with a geometrically complex yet single-patch three-dimensional horseshoe shape and a 16-dimensional random input. The solid horseshoe is constructed by executing a U-shaped sweep of a cross-section consisting of a square with side length $L = 20$ units subtracted by a quarter disk of radius $r_2 = 4$ units, as shown in Figure 5(A). Furthermore, the center of the quarter disk is located at a distance $r_1 = 3$ units from the origin, and the straight section of the horseshoe has height $h = 20$ units.

For material properties, the Poisson's ratio $\nu = 1/3$. Moreover, the Young's modulus $E(\mathbf{z}; \cdot)$ is a homogeneous lognormal random field with mean $\mu_E = 210 \times 10^9$ units and coefficient of variation $v_E = 0.1$. As a translation random field, $E(\mathbf{z}; \cdot)$ is written as

$$E(\mathbf{z}; \cdot) = C_\alpha \exp[\alpha(\mathbf{z}; \cdot)],$$

where $\mathbf{z}$ is a spatial point in the physical domain of the horseshoe,

$$C_\alpha = \frac{\mu_E}{\sqrt{1 + v_E^2}},$$

and $\alpha(\mathbf{z}; \cdot)$ is a homogeneous Gaussian random field with mean *zero* and covariance function

$$\Gamma(\mathbf{z}, \mathbf{z}') = \sigma^2 \exp\left(-\frac{||\mathbf{z} - \mathbf{z}'||}{bL}\right), \ \mathbf{z}, \mathbf{z}' \in \mathcal{D} \subset \mathbb{R}^3$$

with $\sigma^2 = 0.01$ and $b = 1$.

For the stress analysis, the top surfaces of the horseshoe are subjected to a pair of random displacements of magnitude u_s, which is distributed uniformly over $[-0.005, 0.01]$ units, in the directions shown in Figure 5(A). This causes asymmetric deformation and resultant twisting of the horseshoe. Furthermore, allowing the applied displacements to change direction introduces a slope discontinuity in the von Mises stress, which is the output quantity of interest. This necessitates the application of optimal SDD to be demonstrated in the context of a realistic linear elasticity problem. The stochastic problem requires calculating the various probabilistic characteristics of stresses due to uncertain boundary conditions and material properties.

Employing IGA with quadratic nonuniform rational B-splines (NURBS), the geometry was constructed precisely. First, a base eight-element mesh was generated with quadratic NURBS; the corresponding knot vectors are reported

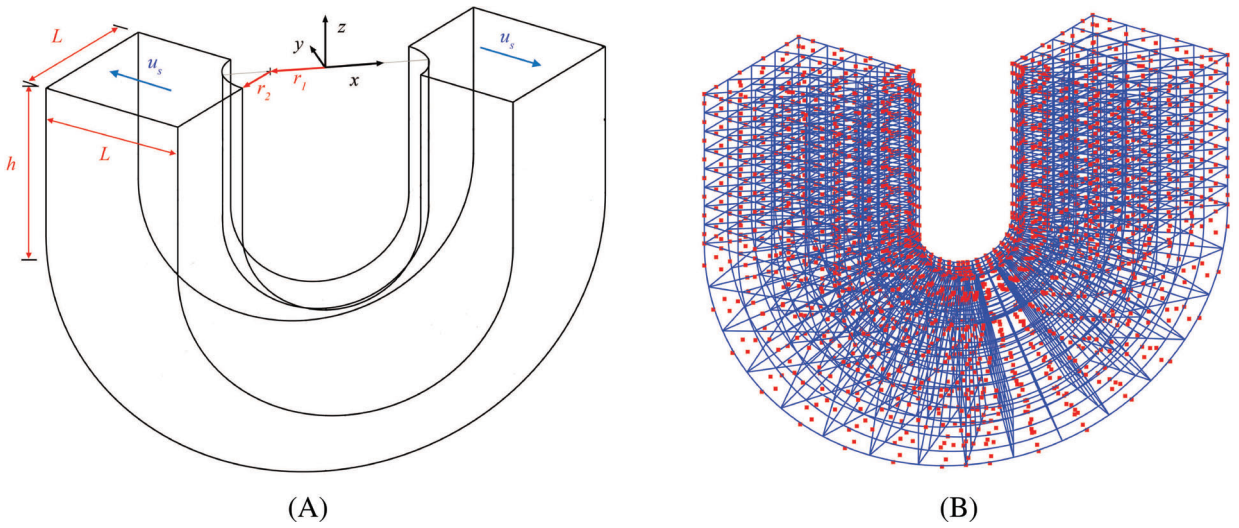


FIGURE 5 A horseshoe under asymmetric loading in Example 2: (A) problem schematic; (B) refined mesh employed in the analyses. The control points are illustrated by red closed squares

in a prior work.[9] Second, the fine mesh illustrated in Figure 5(B) was constructed by simple knot insertion, which consists of 1024 quadratic elements and 2442 control points. The control points are depicted by red closed squares. Readers interested in modeling techniques of three-dimensional objects via NURBS are directed to the book by Cottrell et al.[21]

The random field describing Young's modulus was discretized by the K–L expansion, exploiting the isogeometric collocation method[19] and considering 15 independent standard Gaussian random variables. As the 15th eigenvalue was only 2.47% of the first eigenvalue, this number was deemed adequate to satisfactorily approximate the random field. The domain discretization and NURBS objects, such as orders and knot vectors, for the K–L expansion and deterministic IGA were identical. The eigenpairs $\{\lambda_i, \phi_i(\mathbf{z})\}$, $i = 1, \ldots, 15$, were then numerically calculated. By dint of the collocation method, the random field is efficiently discretized with adequate accuracy. In aggregate, there are 16 random variables in this problem.

In the horseshoe problem, a crude MCS with a sample size of 10,000 provided a benchmark solution for univariate SDD with uniform and optimal knot vectors. From prior experience,[9] no tangible improvement was noticed when implementing bivariate SDD over univariate SDD, indicating that the original function is dominantly univariate with weak interaction terms, hence the focus on univariate SDD here. Furthermore, the same degrees and numbers of subintervals were adopted in all 16 coordinate directions, that is, $p_k = p$ (say) and $I_k = I$ (say) for $k = 1, 2, \ldots 16$, while the knot vectors in all but the coordinate direction associated with the applied displacement were identical. More specifically, only the knot locations associated with the applied displacement were optimized, whereas the remaining 15 directions were assigned uniformly spaced knots. Given the high dimension of the problem and relatively expensive computation involved in evaluating the output function, the SDD coefficients were estimated by SLS regression (Section 5.2), rather than numerical integration, using a subset of the MCS-generated input-output data set. A nonorthonormal spline basis was used during the knot vector optimization process, according to Equation (27), after which the coefficients with respect to an orthonormal basis were computed for the resulting optimal knot sequence according to Equation (22).

For the SDD methods, each standard Gaussian random variable X_i was transformed to a truncated Gaussian variable, where the PDF was truncated at $x_i = \pm 3$. Such a transformation was necessary because splines require bounded support by definition. It is worth noting that transforming to other desired probability measures may affect the results. However, it is best practice to choose a transformation yielding as little difference between the original and mapped distributions as is possible, since in general, heavily nonlinear transformations are to be avoided to obtain better approximations. Hence, the truncated Gaussian is an appropriate choice of mapped distribution in this context. Univariate SDD with a linear ($p = 1$) basis and two or four subintervals ($I = 2, 4$) was employed. The expansion coefficients were estimated by SLS regression employing an MCS-generated input–output data set of size 350 for $I = 2$ and 650 for $I = 4$. To avoid the need to compute the whitening matrix during each iteration, the optimization problem in Equation (27) was used in this example. Both second moment and probabilistic distribution analyses were performed for this example, the results of which are presented next.

7.2.1 | Second-moment analysis

The first set of analyses entails the evaluation of the second-moment properties of the von Mises stress. The crude MCS was chosen to provide the benchmark solution along with four univariate, linear SDD methods: standard SDD method with $p = 1,\ I = 2$ and $p = 1,\ I = 4$; and optimal SDD method with $p = 1,\ I = 2$ and $p = 1,\ I = 4$. Figures 6 and 7 display the contour plots of the mean and SD, respectively, of the von Mises stress σ_v, obtained by MCS and the aforementioned SDD methods. All four SDD methods satisfactorily estimate the mean of the von Mises stress in comparison with those obtained from the MCS to the extent that any differences in the contours are barely distinguishable. Almost the same can be said about the SD of the von Mises stress, although there is a slight degradation in the accuracy of the estimated results by the standard SDD method with $p = 1,\ I = 2$. It is evident that the optimal SDD method with $p = 1,\ I = 2$ is capable of delivering better results than the standard SDD method without increasing the number of subintervals. The error in the maximum values of the SD of the von Mises stress estimated by the optimal SDD is below 0.3%. More importantly, the numbers of function evaluations (IGA) required for all four SDD methods are only 350–650. This reveals that the SDD approximations with four subintervals are more than adequate for computing the second-moment properties of the stochastic responses, requiring only a fraction of the computational effort by MCS.

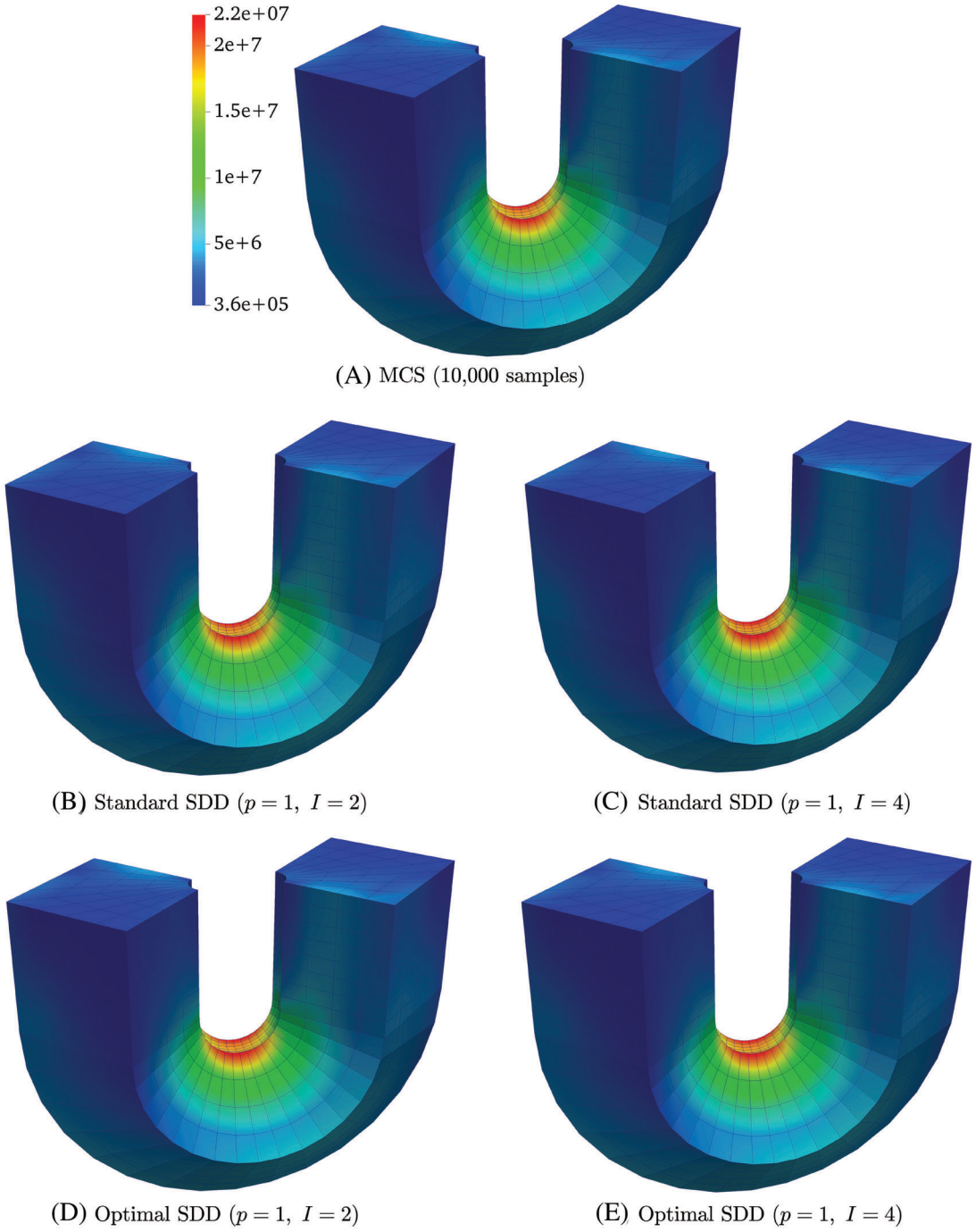


FIGURE 6 Contour plots for the mean of the von Mises stress obtained by various methods in consistent units in Example 2

7.2.2 | Probability distribution analysis

In this section, the CDF of a relevant stress at a critical point of the horseshoe is discussed. Let the point at which the mean of the von Mises stress σ_v is maximum be the critical point and denote by $\sigma_{v,c}$ the random von Mises stress at that point. For CDF analysis of this specific von Mises stress, the crude MCS was selected along with the same four SDD methods defined in the preceding section. With each SDD surrogate constructed, one is able to resample the SDD approximation with a relatively large sample size—10^5 in this example—to estimate the CDF of $\sigma_{v,c}$.

Figure 8(A,B) illustrates the CDF of $\sigma_{v,c}$ calculated by the optimal and standard SDD methods for $I = 2$ and $I = 4$, respectively. For comparison, the MCS estimate is also shown. Evidently, the optimal SDD approximations not only

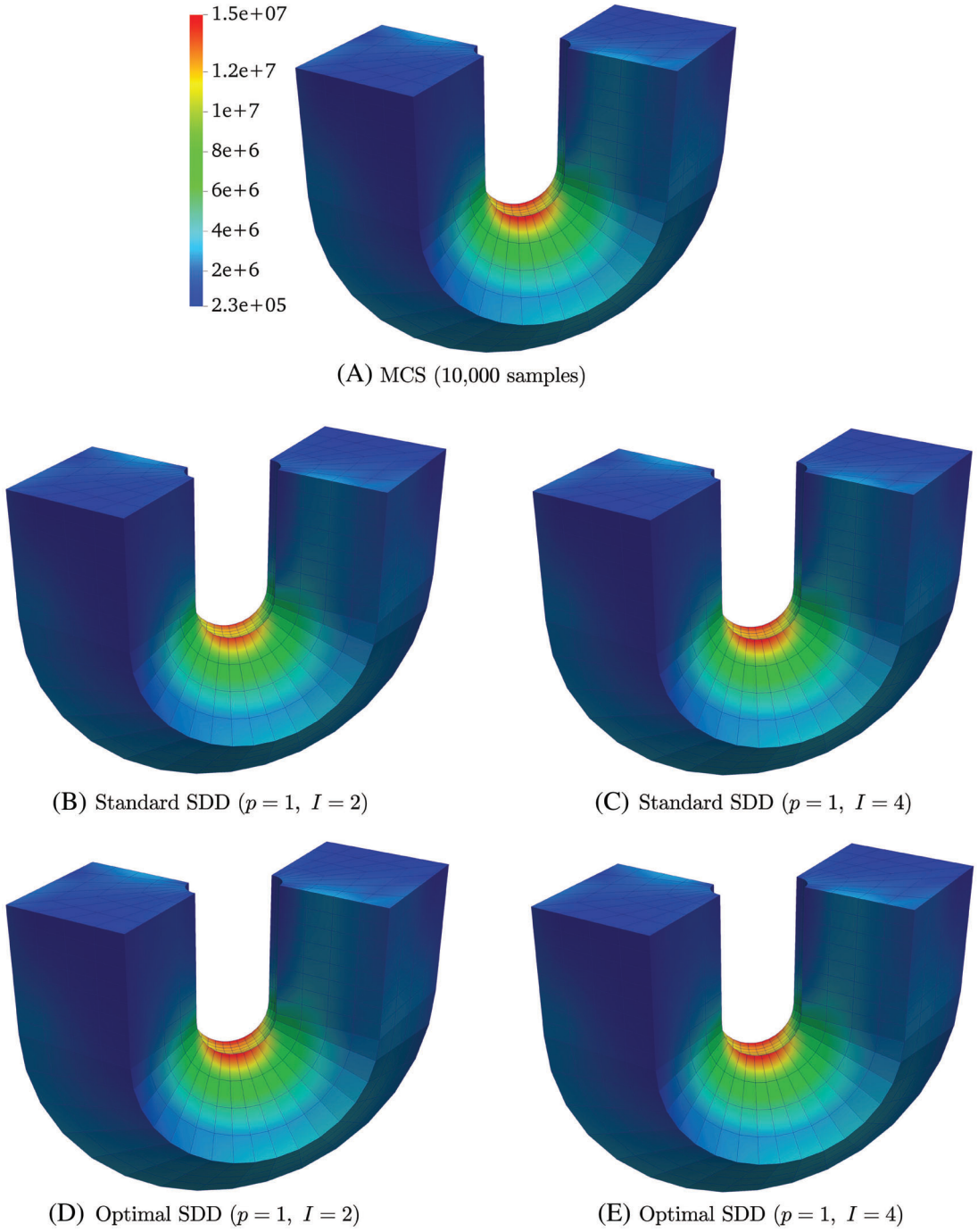
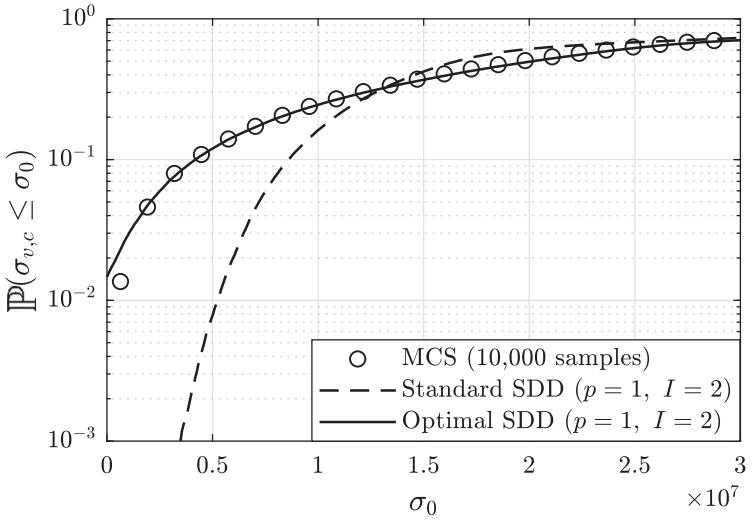


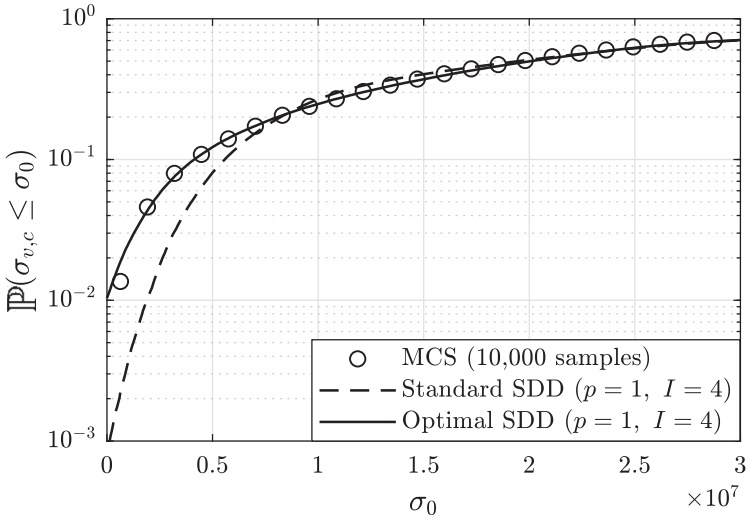
FIGURE 7 Contour plots for the SD of the von Mises stress obtained by various methods in consistent units in Example 2

assiduously capture the second-moment statistical properties of the response, as observed previously, but also provide an impressive approximation of the CDF with only 350 to 650 response (IGA) evaluations, compared with 10,000 in the case of MCS. Furthermore, the knot vector optimization offers a marked improvement in accuracy of the CDF over the standard SDD, especially in the case of SDD with only two subintervals, as can be seen in Figure 8(A). It should be noted that, because von Mises stress is positive by definition, there is a sharp decrease in the CDF near *zero* according to crude MCS. In this regard, the SDD surrogate models may yield invalid results, as they can become negative. This is a known problem faced by surrogate methods entailing orthogonal expansions, including SDD.

In summary, the optimal SDD method can reliably estimate the statistical moments and probability distribution of a random response for bounded or unbounded distributions.



(A)



(B)

FIGURE 8 Probability distribution of the von Mises stress at a critical point in Example 2 obtained by crude MCS and univariate SDD methods: (A) $p = 1,\ I = 2$; (B) $p = 1,\ I = 4$

8 | APPLICATION

This section illustrates the power of the optimal SDD method in solving a large-scale, 43-dimensional practical engineering problem. The application involves a free-free modal analysis of a model aircraft, which is comprised of several materials with uncertain properties, and subsequent estimation of the statistical moments of the natural frequencies and mode shapes.

8.1 | A blended wing-body model aircraft

The subject of the modal analysis is an airplane prototype designed by Airbus, known as the Model Aircraft for Validation and Experimentation of Robust Innovative Controls (MAVERIC), which is shown in Figure 9(A–C). A finite-element mesh consisting of 115,774 linear tetrahedral elements, 44 linear brick elements, and 27,175 nodes was employed to solve the underlying problem. While the specific material composition of this model aircraft is not publicly available, five

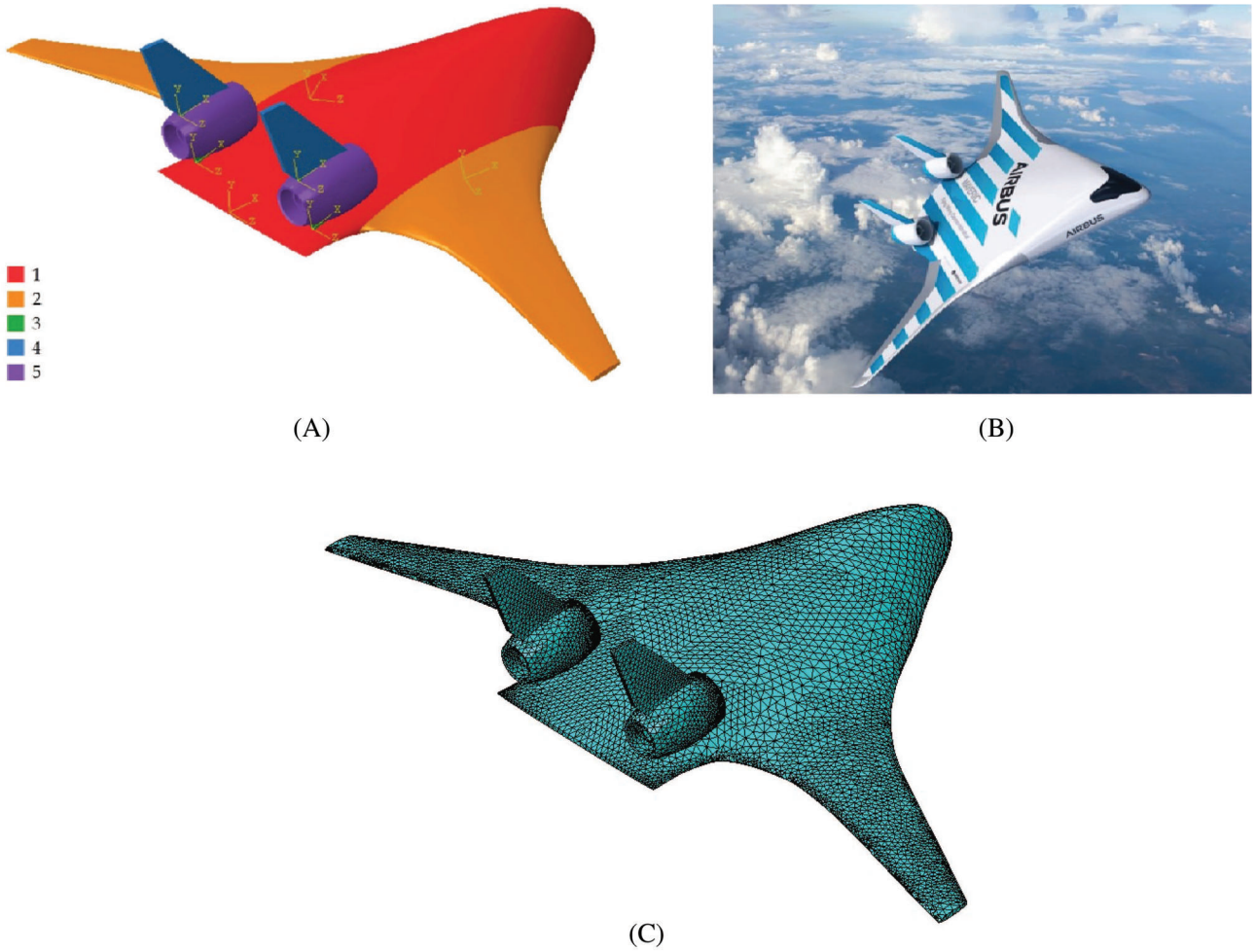


FIGURE 9 MAVERIC model aircraft: (A) computer-aided design model and material designation; (B) actual prototype;[28] (C) finite element mesh

materials—four composites and one metal alloy—commonly used in aerospace applications were assigned to different components for the purpose of this analysis. These materials are numbered and color-coded according to Figure 9(A), which also shows the material orientations of the four composites that were modeled as orthotropic. A total of 43 material properties among the five materials constitute the stochastic input of the problem: (1) the mass densities of all five materials; (2) the nine elastic constants, consisting of three Young's moduli (E_x, E_y, E_z), three shear moduli (G_{xy}, G_{xz}, G_{yz}), and three Poisson's ratios (ν_{xy}, ν_{xz}, ν_{yz}), of the four orthotropic materials; and (3) the two elastic constants, consisting of the Young's modulus and Poisson's ratio of the isotropic material. In this way, a realistic high-dimensional problem is proposed without the need to select a large number of materials. The mean values of these properties, provided in Table 1, are representative of several variants of carbon fiber for materials 1–3, fiberglass for material 4, and aluminum for material 5. It should be noted that, although the mean values of the moduli of the orthotropic materials are equal in two directions, they are treated as independent random variables in the context of this problem. No damping was included.

The 43 material properties were divided into three groups, each with a different probability distribution. The five mass densities comprised the first group, which followed an upper-truncated Weibull distribution with a coefficient of variation of 0.15 and a domain of two SDs below and three above the mean. The second group consisted of 25 properties, including the Young's moduli and shear moduli of the four orthotropic materials as well as the Young's modulus of the isotropic material. The samples corresponding to this group were drawn from a similar distribution as those in the first group, but with a coefficient of variation of 0.1. The final group included the three Poisson's ratios of each of the four orthotropic materials, as well as that of the isotropic material, for a total of 13 properties. The samples in this group were drawn from a beta distribution with a coefficient of variation of 0.05 and a domain of three SDs below and two above

TABLE 1 Mean values of random material properties for the model aircraft

Orthotropic materials										
Material	**Density (kg/m^3)**	E_x **(GPa)**	E_y **(GPa)**	E_z **(GPa)**	G_{xy} **(GPa)**	G_{xz} **(GPa)**	G_{yz} **(GPa)**	ν_{xy}	ν_{xz}	ν_{yz}
1	2100	96	36	96	17	21	17	0.22	0.06	0.22
2	2060	36	36	141	14	20	20	0.18	0.25	0.18
3	2140	98	37	98	18	21	18	0.21	0.06	0.21
4	2375	8.417	4.257	5.511	1.764	2.356	1.797	0.217	0.16	0.225
Isotropic materials										
Material	**Density (kg/m^3)**				**Young's modulus (GPa)**			**Poisson's ratio**		
5	2810				71.7			0.33		

the mean. It is worth noting that because the mean values of all 43 properties are different, with the exception of some orthotropic moduli, the variables in each group are not necessarily identically distributed.

Hence, a total of three knot vectors were used—one for each of the aforementioned groups of variables. The values of the variables within each group could then be obtained by linear transformations of independent draws from the probability distribution corresponding to the group. For simplicity, the degrees, p, and numbers of subintervals, I, for all three knot sequences were chosen to be identical. Furthermore, the minimum separation distance δ used in the constraints of the knot vector optimization was taken to be 40% of the uniform knot spacing. This tightening of the constraints, relative to those in Examples 1 and 2 of Section 7, was done to mitigate issues encountered for lower values of δ with the conditioning of the information matrix used to compute the expansion coefficients. Despite this restriction, it is believed that a sufficient amount of freedom was allowed in the optimal solution, as there were not expected to be any slope discontinuities or rapid changes in the response that would necessitate more closely spaced knots.

8.2 | Results

A mode-based steady-state dynamic analysis was performed to obtain the eigensolutions of the model aircraft, which represent its natural frequencies and mode shapes. This was done using the Lanczos method,[29] available in ABAQUS (version 2019).[30] Since the model was unconstrained, the first six eigensolutions corresponding to the rigid-body modes were excluded from the results presented here.

Due to the uncertainty in the material properties, the eigensolutions of the model aircraft are stochastic. Three univariate optimal SDD methods, two of linear order ($p = 1, I = 2, 4$) and one of quadratic order ($p = 2, I = 4$), were employed to estimate the first and second moment properties, while a crude MCS entailing 10,000 samples provided the benchmark solution. The SDD expansion coefficients were again estimated by SLS regression of an MCS-generated input–output data set consisting of only 750 samples (FEA).

Table 2 lists the means and SDs of the natural frequencies of the model aircraft associated with the first 10 non-rigid-body modes, obtained using the aforementioned optimal SDD methods and MCS. A comparison of respective statistics of the eigenfrequencies by these four methods indicate a high level of accuracy of all three optimal SDD methods. For example, the relative errors in the SDs of the 10th natural frequency calculated by each of the three optimal SDD methods with respect to MCS are 2.7%, 1.6%, and 1.1%, respectively. The quadratic optimal SDD approximation with four subintervals, despite having the largest number of bases, is able to produce especially accurate results while only requiring 750 samples (FEA) for this 43-dimensional stochastic problem.

In addition to the frequency analysis, the nodal displacement magnitudes were calculated to assess the second moment properties of the mode shapes. Contour plots of the SDs of the first and third non-rigid-body mode shapes, calculated using crude MCS and the three optimal SDD methods, are shown in Figures 10(A–D) and 11(A–D), respectively. Once more, the optimal SDD methods yield such a degree of accuracy that the corresponding contour plots are nearly indistinguishable from those generated by MCS for both mode shapes. Hence, for both the eigenfrequency and eigenmode analyses, the proposed optimal SDD methods deliver highly accurate second moment statistics, while requiring less than 10% of the computational effort mandated by crude MCS. The success of the UQ analysis for the

TABLE 2 Second-moment statistics of first 10 non-rigid-body eigenfrequencies of the model aircraft by various methods

	Univariate optimal SDD methods (750 samples)						Crude MCS (10,000 samples)	
	$p = 1, I = 2$		$p = 1, I = 4$		$p = 2, I = 4$			
Mode	Mean (Hz)	SD (Hz)	Mean (Hz)	SD (Hz)	Mean (Hz)	SD (Hz)	Mean (Hz)	SD (Hz)
1	59.4406	4.4204	59.4427	4.4181	59.4430	4.4221	59.4224	4.4218
2	60.7739	4.4636	60.7757	4.4659	60.7757	4.4658	60.7523	4.4653
3	121.0259	9.4724	121.1010	9.6434	121.0426	9.7088	120.7733	9.7645
4	128.7155	8.1245	128.6913	8.2728	128.6898	8.3703	128.4790	8.4839
5	136.7336	7.2396	136.6542	7.4320	136.7404	7.5433	136.8959	7.7858
6	143.9950	7.6669	143.9679	7.9409	144.0157	8.0612	144.1646	8.4261
7	259.2875	15.8346	259.3240	16.1396	259.3251	16.1679	259.1993	16.4193
8	265.1022	15.9591	265.1741	16.2490	265.1863	16.3037	265.0101	16.5504
9	295.8141	20.2159	295.6427	20.4128	295.6370	20.5291	295.5515	20.6350
10	299.8269	20.3640	299.7538	20.5772	299.6423	20.6819	299.5501	20.9212

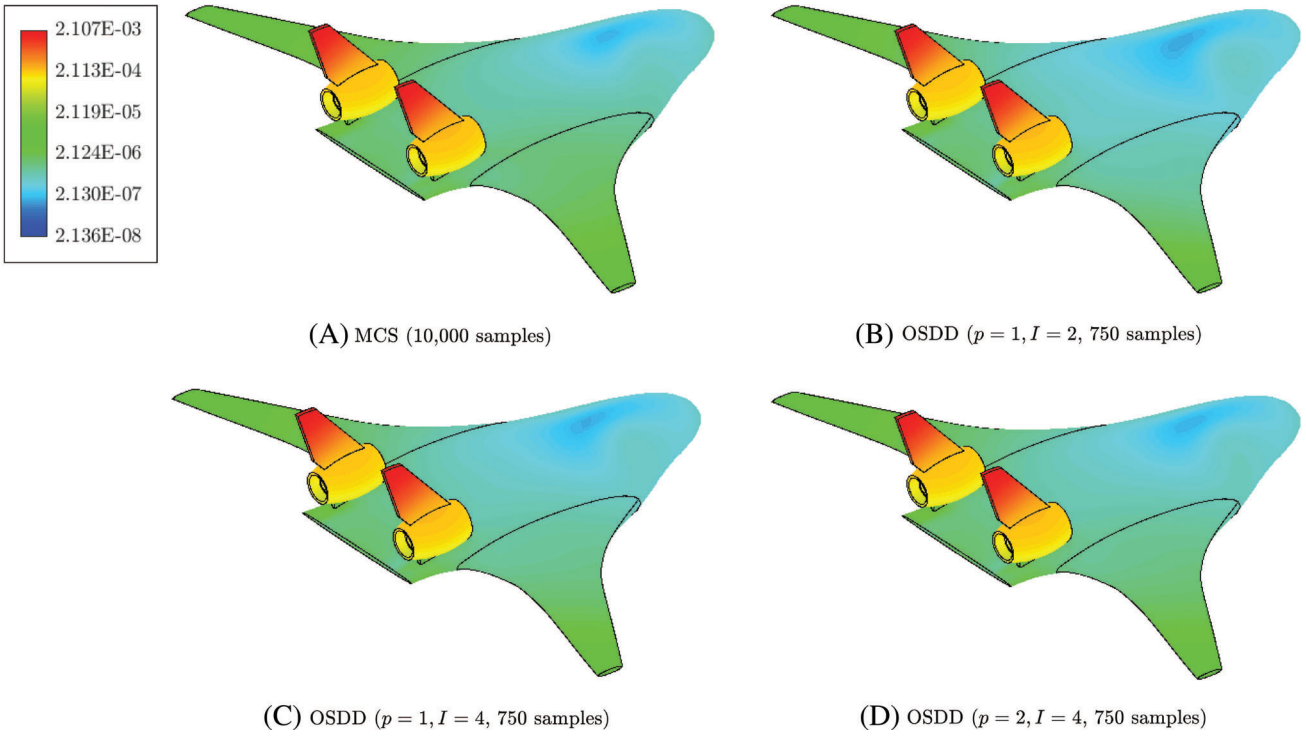


FIGURE 10 Contour plots for the SD of the first non-rigid-body mode shape of the model aircraft obtained by crude MCS and various optimal SDD (OSDD) methods

model aircraft demonstrates the viability of optimal SDD in solving high-dimensional, industrial-scale engineering problems.

9 | CONCLUSION

A novel dimensional decomposition, designated as optimal SDD, is introduced for general high-dimensional uncertainty quantification analysis of complex systems. The method is premised on optimally derived knot vectors of B-splines in

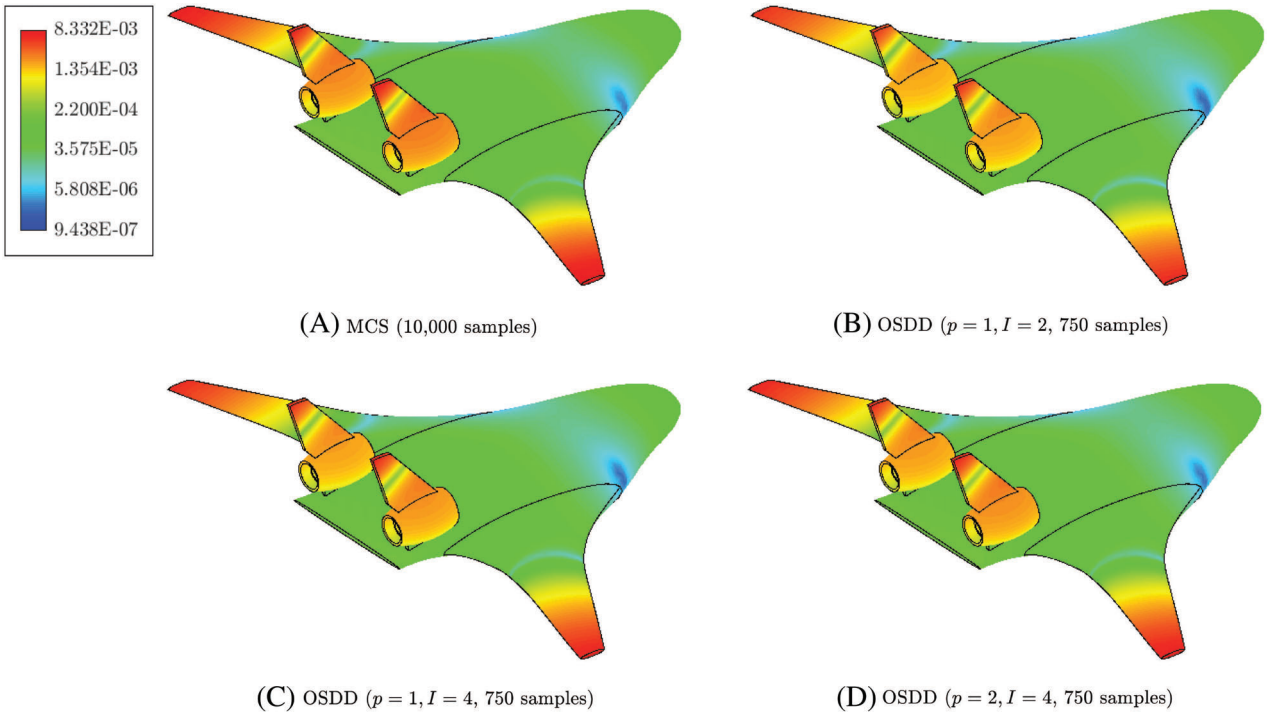


FIGURE 11 Contour plots for the SD of the third non-rigid-body mode shape of the model aircraft obtained by crude MCS and various optimal SDD (OSDD) methods

some or all coordinate directions, whitening transformation producing measure-consistent orthonormalized B-splines equipped with optimal knots, and Fourier-spline expansion of a general high-dimensional output function of interest. Compared with the standard SDD, developed in a prior work, the knot vectors are not mandated to be uniformly spaced or selected instinctively. Two distinct approaches, one exploiting an orthonormal basis and the other utilizing a nonorthonormal basis, are advocated for determining the optimal locations of the knots by minimizing the mean-squared approximation error. In doing so, no expensive function evaluations, in excess of those required to estimate the expansion coefficients, are incurred, meaning that the optimization can be performed with little additional effort. Indeed, the generation of optimal knot vectors can be viewed as an inexpensive preprocessing step toward creating the optimal SDD. Thereafter, the optimal SDD method is formed by retaining relevant tensor products of univariate splines associated with a chosen degree of interaction from the ANOVA decomposition. Analytical formulas have been proposed to calculate the second-moment properties by the optimal SDD method for a general output random variable in terms of the expansion coefficients involved. Akin to that of the standard SDD method, the computational complexity of the optimal SDD method is polynomial, as opposed to exponential, thus alleviating the curse of dimensionality to an appreciable magnitude.

Numerical results indicate that the optimal SDD method is more precise than the standard SDD method in predicting not only the variance, but also the CDF of an output random variable, with both methods demanding practically the same computational resources. Moreover, a low-order optimal SDD approximation with an adequate mesh size generates a substantially more accurate estimate of the output variance than a high-order approximation from existing polynomial chaos expansion or sparse-grid quadrature. The optimal SDD method proposed is most relevant under locally nonlinear or nonsmooth behavior often recognized in engineering applications.

ACKNOWLEDGMENT

This work was supported by the U.S. National Science Foundation under grant CMMI-1933114.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Sharif Rahman https://orcid.org/0000-0003-0837-9871

REFERENCES

1. Wiener N. The homogeneous chaos. *Am J Math*. 1938;60(4):897-936.
2. Cameron RH, Martin WT. The orthogonal development of non-linear functionals in series of Fourier-Hermite functionals. *Ann Math*. 1947;48:385-392.
3. Xiu D, Karniadakis GE. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J Sci Comput*. 2002;24:619-644.
4. Rahman S. Mathematical properties of polynomial dimensional decomposition. *SIAM/ASA J Uncertain Quantif*. 2018;6:816-844.
5. Rahman S. A polynomial dimensional decomposition for stochastic computing. *Int J Numer Methods Eng*. 2008;76:2091-2116.
6. Yadav V, Rahman S. Adaptive-sparse polynomial dimensional decomposition for high-dimensional stochastic computing. *Comput Methods Appl Mech Eng*. 2014;274:56-83.
7. Rahman S. A spline chaos expansion. *SIAM/ASA J Uncertain Quantif*. 2020;8:27-57.
8. Rahman S, Jahanbin R. A spline dimensional decomposition for uncertainty quantification. *SIAM/ASA J Uncertain Quantif*. 2020.
9. Jahanbin R, Rahman S. Stochastic isogeometric analysis in linear elasticity. *Comput Methods Appl Mech Eng*. 2020;364(112928):1-38.
10. Babuska I, Nobile F, Tempone R. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J Numer Anal*. 2007;45(3):1005-1034.
11. Xiu D, Hesthaven JS. High-order collocation methods for the differential equation with random inputs. *SIAM J Sci Comput*. 2005;27:1118-1139.
12. Gerstner T, Griebel M. Numerical integration using sparse grids. *Numer Algorithms*. 1998;18:209-232.
13. Smolyak S. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl Akad Nauk*. 1963;4:240-243.
14. Rahman S. Approximation errors in truncated dimensional decompositions. *Math Comput*. 2014;83(290):2799-2819.
15. Griebel M, Holtz M. Dimension-wise integration of high-dimensional functions with applications to finance. *J Complex*. 2010;26(5):455-489.
16. Rabitz H, Alis O. General foundations of high dimensional model representations. *J Math Chem*. 1999;25:197-233. https://doi.org/10.1023/A:1019188517934
17. Loève M. *Probability Theory*. Vol II. Berlin/Heidelberg, Germany; New York, NY: Springer; 1977.
18. Rahman S. A Galerkin isogeometric method for Karhunen-Loève approximation of random fields. *Comput Methods Appl Mech Eng*. 2018;338:533-561.
19. Jahanbin R, Rahman S. An isogeometric collocation method for efficient random field discretization. *Int J Numer Methods Eng*. 2019;117(3):344-369.
20. Jahanbin R, Rahman S. Isogeometric methods for Karhunen-Loève representation of random fields on arbitrary multi-patch domains. *Int J Uncertain Quantif*. 2021;11:27-57.
21. Cottrell JA, Hughes TJR, Bazilevs Y. *Isogeometric Analysis: Toward Integration of CAD and FEA*. Hoboken, NJ: John Wiley & Sons; 2009.
22. De Boor C. On calculation with B-splines. *J Approx Theory*. 1972;6:50-62.
23. Efron B, Stein C. The jackknife estimate of variance. *Ann Stat*. 1981;9(3):586-596.
24. Hoeffding W. A class of statistics with asymptotically normal distribution. *Ann Math Stat*. 1948;19(3):293-325.
25. Xu H, Rahman S. A generalized dimension-reduction method for multi-dimensional integration in stochastic mechanics. *Int J Numer Methods Eng*. 2004;61:1992-2019.
26. Dertimanis V, Spiridonakos M, Chatzi E. Data-driven uncertainty quantification of structural systems via B-spline expansion. *Comput Struct*. 2018;207:245-257.
27. Genz A, Malik A. An adaptive algorithm for numerical integration over an *n*-dimensional rectangular region. *J Comput Appl Math*. 1980;6:295-302.
28. Airbus www.airbus.com. Accessed October 12, 2020.
29. Cullum JK, Willoughby RA. *Lanczos Algorithms for Large Symmetric Eigenvalue Computations: Theory*. Classics in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics; 2002.
30. ABAQUS Analysis user's manual standard, version 6.12; 2012; Dassault Systems Simulia Corporation.
31. Cox MG. The numerical evaluation of B-splines. *J Inst Math Appl*. 1972;10:134-149.

How to cite this article: Dixler S, Jahanbin R, Rahman S. Uncertainty quantification by optimal spline dimensional decomposition. *Int J Numer Methods Eng*. 2021;122:5898–5934. https://doi.org/10.1002/nme.6778

APPENDIX A. 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 appendix briefly describes the paraphernalia of univariate B-splines.

A.1 Knot Vector

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

Definition 1. A knot vector $\boldsymbol{\xi}_k$ for the interval $[a_k, b_k] \subset \mathbb{R}$, given $n_k > p_k \geq 0$, is a vector comprising nondecreasing sequence of real numbers

$$\begin{aligned} \boldsymbol{\xi}_k := \{\xi_{k,i_k}\}_{i_k=1}^{n_k+p_k+1} &= \{a_k = \xi_{k,1}, \xi_{k,2}, \ \dots \ , \xi_{k,n_k+p_k+1} = b_k\}, \\ \xi_{k,1} &\leq \xi_{k,2} \leq \cdots \leq \xi_{k,n_k+p_k+1}, \end{aligned} \tag{A1}$$

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

According to Equation (A1), there are a total of $n_k + p_k + 1$ knots, which may be equally or unequally spaced. To monitor knots without repetitions, denote by $\zeta_{k,1}, \ \dots \ , \zeta_{k,r_k}$ the r_k distinct knots in $\boldsymbol{\xi}_k$ with respective multiplicities $m_{k,1}, \ \dots \ , m_{k,r_k}$. Then the knot vector in Equation (A1) can be expressed more compactly by

$$\begin{aligned} \boldsymbol{\xi}_k = \{a_k = \overbrace{\zeta_{k,1}, \ \dots \ , \zeta_{k,1}}^{m_{k,1} \text{ times}}, \overbrace{\zeta_{k,2}, \ \dots \ , \zeta_{k,2}}^{m_{k,2} \text{ times}}, \ \dots \ , \overbrace{\zeta_{k,r_k-1}, \ \dots \ , \zeta_{k,r_k-1}}^{m_{k,r_k-1} \text{ times}}, \overbrace{\zeta_{k,r_k}, \ \dots \ , \zeta_{k,r_k}}^{m_{k,r_k} \text{ times}} = b_k\}, \\ a_k = \zeta_{k,1} < \zeta_{k,2} < \cdots < \zeta_{k,r_k-1} < \zeta_{k,r_k} = b_k, \end{aligned} \tag{A2}$$

which consists of a total number of

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

knots. As shown in Equation (A2), each knot, whether interior or exterior, may appear $1 \leq m_{k,j_k} \leq p_k + 1$ times, where m_{k,j_k} is referred to as its multiplicity. The multiplicity has important implications on the regularity properties of B-spline functions. A knot vector is called open if the end knots have multiplicities $p_k + 1$. In this case, definitions of more specific knot vectors are in order.

Definition 2. A knot vector is said to be $(p_k + 1)$-open if the first and last knots appear $p_k + 1$ times, that is, if

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

Definition 3. A knot vector 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

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

A $(p_k + 1)$-open knot vector with or without simple knots is commonly found in applications.[21]

A.2 B-Splines

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

Definition 4. Let $\boldsymbol{\xi}_k$ be a general knot vector of length at least $p_k + 2$ for the interval $[a_k, b_k]$, as defined by Equation (A1). 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,

$$B^k_{i_k,0,\boldsymbol{\xi}_k}(x_k) := \begin{cases} 1, & \xi_{k,i_k} \leq x_k < \xi_{k,i_k+1}, \\ 0, & \text{otherwise}, \end{cases} \tag{A3}$$

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

$$B^k_{i_k,p_k,\xi_k}(x_k) := q_{i_k,p_k,\xi_k}(x_k) B^k_{i_k,p_k-1,\xi_k}(x_k) + \left(1 - q_{i_k+1,p_k,\xi_k}(x_k)\right) B^k_{i_k+1,p_k-1,\xi_k}(x_k), \tag{A4}$$

where

$$q_{i_k,p_k,\xi_k}(x_k) := \begin{cases} \frac{x_k - \xi_{k,i_k}}{\xi_{k,i_k+p_k} - \xi_{k,i_k}}, & \xi_{k,i_k} < \xi_{k,i_k+p_k}, \\ 0, & \text{otherwise}, \end{cases} \tag{A5}$$

and $1 \le k \le N$, $1 \le i_k \le n_k$, $1 \le p_k < \infty$.

The recursive formula in Definition 4 was derived by Cox[31] and de Boor.[22]

APPENDIX B. OBJECTIVE FUNCTION GRADIENTS

This appendix serves the purpose of providing explicit expressions involved in evaluating the sensitivity, with respect to the knot locations, of the objective functions given in Sections 6.2 and 6.3, which is needed when using a gradient-based method, such as SLP or SQP, to solve the underlying constrained nonlinear optimization problems. As mentioned previously, it is assumed for the sake of simplicity that every internal knot has a multiplicity of *one*. Hence, the number of internal knots in each coordinate direction is equal to $n_k - p_k - 1$. The procedure described in this appendix is repeated to obtain derivatives with respect to all internal knots in all coordinate directions, that is for all ξ_{k,j_k}, $p_k + 2 \le j_k \le n_k$, $1 \le k \le N$.

B.1 Sensitivity of B-Splines

From Equation (A3), it is clear that $\partial B^k_{i_k,0,\xi_k}(x_k)/\partial \xi_{k,j_k} = 0$ for all admissible values of k, i_k, j_k, $\boldsymbol{\xi}_k$, and x_k. Hence, the propagation of the influence of the internal knot locations in $\boldsymbol{\xi}_k$ begins with that on the quotients $q_{i_k,p_k,\xi_k}(x_k)$, given in Equation (A5). Accordingly, the partial derivatives are

$$\frac{\partial q_{i_k,p_k,\xi_k}(x_k)}{\partial \xi_{k,j_k}} = \begin{cases} \frac{x_k - \xi_{k,i_k+p_k}}{(\xi_{k,i_k+p_k} - \xi_{k,i_k})^2}, & j_k = i_k \quad \text{and} \quad \xi_{k,i_k} < \xi_{k,i_k+p_k}, \\ \frac{\xi_{k,i_k} - x_k}{(\xi_{k,i_k+p_k} - \xi_{k,i_k})^2}, & j_k = i_k + p_k \quad \text{and} \quad \xi_{k,i_k} < \xi_{k,i_k+p_k}, \\ 0, & \text{otherwise}. \end{cases}$$

The sensitivity of the higher-order B-splines can then be derived recursively by differentiation of Equation (A4), yielding

$$\begin{aligned} \frac{\partial B^k_{i_k,p_k,\xi_k}(x_k)}{\partial \xi_{k,j_k}} =& \frac{\partial q_{i_k,p_k,\xi_k}(x_k)}{\partial \xi_{k,j_k}} B^k_{i_k,p_k-1,\xi_k}(x_k) + q_{i_k,p_k,\xi_k}(x_k) \frac{\partial B^k_{i_k,p_k-1,\xi_k}(x_k)}{\partial \xi_{k,j_k}} - \frac{\partial q_{i_k+1,p_k,\xi_k}(x_k)}{\partial \xi_{k,j_k}} B^k_{i_k+1,p_k-1,\xi_k}(x_k) \\ &+ \left(1 - q_{i_k+1,p_k,\xi_k}(x_k)\right) \frac{\partial B^k_{i_k+1,p_k-1,\xi_k}(x_k)}{\partial \xi_{k,j_k}}. \end{aligned} \tag{B1}$$

It is important to note the distinction between the derivatives of the B-spline functions with respect to the knot locations $\boldsymbol{\xi}_k$ and those with respect to the coordinate x_k. At this point, the sensitivity of the B-splines propagates in different ways, depending on which approach, described in Sections 6.2 and 6.3, is followed to optimize the internal knot locations.

B.2 Sensitivity with Orthonormal Basis

When an orthonormal basis is employed during the optimization procedure, the dependence of the whitening transformation on the internal knots must be considered. Recall that this transformation is given by

$$\boldsymbol{\psi}_k(x_k) = \mathbf{W}_k \mathbf{P}_k(x_k), \tag{B2}$$

where $\boldsymbol{\psi}_k(x_k)$ consists of the univariate orthonormal spline functions $\psi^k_{i_k,p_k,\xi_k}(x_k)$, $\mathbf{W}_k$ is the $n_k \times n_k$ whitening matrix, and $\mathbf{P}_k(x_k) := \left(1, B^k_{2,p_k,\xi_k}(x_k), \ldots, B^k_{n_k,p_k,\xi_k}(x_k)\right)^\intercal$. Differentiating Equation (B2) yields

$$\frac{\partial \boldsymbol{\psi}_k(x_k)}{\partial \xi_{k,j_k}} = \frac{\partial \mathbf{W}_k}{\partial \xi_{k,j_k}} \mathbf{P}_k(x_k) + \mathbf{W}_k \frac{\partial \mathbf{P}_k(x_k)}{\partial \xi_{k,j_k}}. \tag{B3}$$

Whereas $\partial \mathbf{P}_k(x_k)/\partial \xi_{k,j_k}$ simply involves the derivatives given in Equation (B1), with the first element replaced by *zero*, the sensitivity of $\mathbf{W}_k$ is yet to be determined. Recall that

$$\mathbf{W}_k^\intercal \mathbf{W}_k = \mathbf{G}_k^{-1}, \tag{B4}$$

where $\mathbf{G}_k$ is the spline moment matrix. Equation (B4) can be implicitly differentiated to obtain

$$\frac{\partial \mathbf{W}_k^\intercal}{\partial \xi_{k,j_k}} \mathbf{W}_k + \mathbf{W}_k^\intercal \frac{\partial \mathbf{W}_k}{\partial \xi_{k,j_k}} = -\mathbf{G}_k^{-1} \frac{\partial \mathbf{G}_k}{\partial \xi_{k,j_k}} \mathbf{G}_k^{-1}. \tag{B5}$$

Equation (B5) cannot be solved for $\partial \mathbf{W}_k/\partial \xi_{k,j_k}$ by simple matrix inversion or rearrangement of terms. However, when obtained by inverting the Cholesky factorization of the spline moment matrix, $\mathbf{W}_k$, and its derivatives by extension, are $n_k \times n_k$ lower triangular matrices. Furthermore, since $\mathbf{G}_k$ is an $n_k \times n_k$ symmetric matrix, so too is the term on the right-hand side of Equation (B5), which is consistent with the obvious symmetry of the sum on the left-hand side. Consequently, there are $n_k(n_k+1)/2$ equations in the $n_k(n_k+1)/2$ unknown elements of $\partial \mathbf{W}_k/\partial \xi_{k,j_k}$.

The solution to this system of equations can be found by performing a half-vectorization on $\partial \mathbf{W}_k/\partial \xi_{k,j_k}$ and the matrix resulting from evaluation of the right-hand side of Equation (B5), and constructing a $n_k(n_k+1)/2 \times n_k(n_k+1)/2$ coefficient matrix containing a pattern of the elements of $\mathbf{W}_k$, some of which are doubled due to the sum, that multiply the respective unknown elements of $\partial \mathbf{W}_k/\partial \xi_{k,j_k}$. This matrix can be derived by expanding the left-hand side of Equation (B5), collecting like terms of $\partial \mathbf{W}_k/\partial \xi_{k,j_k}$, and expressing the result as a matrix–vector product. More practically, a very simple computer algorithm can be implemented to construct this matrix for any arbitrary number of basis functions n_k. This exercise is left up to the reader, if desired.

Before Equation (B5) can be solved, the gradients of the spline moment matrix must be computed. Recall that

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

This equation can be differentiated with respect to the knot locations to obtain

$$\frac{\partial \mathbf{G}_k}{\partial \xi_{k,j_k}} = \mathbb{E}\left[\frac{\partial \mathbf{P}_k(X_k)}{\partial \xi_{k,j_k}} \mathbf{P}_k^\intercal(X_k) + \mathbf{P}_k(X_k) \frac{\partial \mathbf{P}_k^\intercal(X_k)}{\partial \xi_{k,j_k}}\right]. \tag{B6}$$

Once again, the derivatives of $\mathbf{P}_k$ have already been determined. Therefore, the derivatives of $\mathbf{G}_k$ are determined by Equation (B6), which in turn allow the solution of Equation (B5) for the derivatives of $\mathbf{W}_k$, which can finally be substituted into Equation (B3) to obtain the derivatives of $\psi^k_{i_k,p_k,\xi_k}(x_k)$. The next step is to compute the gradients of the multivariate orthonormalized B-spline functions, related to $\psi^k_{i_k,p_k,\xi_k}(x_k)$ by

$$\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{x}_u) = \prod_{k \in u} \psi^k_{i_k,p_k,\xi_k}(x_k), \quad \mathbf{i}_u = (i_{k_1}, \ldots, i_{k_{|u|}}) \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}, \tag{B7}$$

where $\emptyset \neq u = \{k_1, \ldots, k_{|u|}\} \subseteq \{1, \ldots, N\}$ and $\overline{\mathcal{I}}_{u,\mathbf{n}_u} := \left\{\mathbf{i}_u = (i_{k_1}, \ldots, i_{k_{|u|}}) : 2 \leq i_{k_l} \leq n_{k_l},\ l = 1, \ldots, |u|\right\} \subset (\mathbb{N} \setminus \{1\})^{|u|}$. Differentiating Equation (B7) yields

$$\frac{\partial \Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{x}_u)}{\partial \xi_{k,j_k}} = \sum_{s \in u} \frac{\partial \psi^s_{i_s,p_s,\xi_s}(x_s)}{\partial \xi_{k,j_k}} \prod_{\substack{t \in u \\ t \neq s}} \psi^t_{i_t,p_t,\xi_t}(x_t), \quad \mathbf{i}_u = (i_{k_1}, \ldots, i_{k_{|u|}}) \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}. \tag{B8}$$

Note that the sum in Equation (B8) need only be evaluated over the values of s that pertain to coordinate directions that share a knot sequence with coordinate direction k. Hence, if the knot sequences in all coordinate directions are to be optimized independently, then the sum can be omitted and s can be replaced with k.

Once the multivariate orthonormalized B-spline functions are converted to their single-index analogues according to

$$\left\{ \Psi^{u}_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{X}_u) \,:\, 1 \le |u| \le S,\ \mathbf{i}_u \in \overline{I}_{u,\mathbf{n}_u} \right\} = \left\{ \Psi_2(\mathbf{X};\boldsymbol{\Xi}),\ \dots\ , \Psi_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\mathbf{X};\boldsymbol{\Xi}) \right\},\quad \Psi_1(\mathbf{X};\boldsymbol{\Xi}) = 1,$$

the SDD expansion coefficients are calculated in one of two ways.

One option is to compute the coefficients according to

$$C_i(\boldsymbol{\Xi}) := \int_{\mathbb{A}^N} y(\mathbf{x})\Psi_i(\mathbf{X};\boldsymbol{\Xi}) f_{\mathbf{X}}(\mathbf{x})\ \mathrm{d}\mathbf{x},\quad i = 1,\ \dots\ , L_{S,\mathbf{p},\boldsymbol{\Xi}}, \tag{B9}$$

where $L_{S,\mathbf{p},\boldsymbol{\Xi}}$ is the total number of basis functions. Equation (B9) can be differentiated to obtain

$$\frac{\partial C_i(\boldsymbol{\Xi})}{\partial \xi_{k,j_k}} = \int_{\mathbb{A}^N} y(\mathbf{x}) \frac{\partial \Psi_i(\mathbf{X};\boldsymbol{\Xi})}{\partial \xi_{k,j_k}} f_{\mathbf{X}}(\mathbf{x})\ \mathrm{d}\mathbf{x},\quad i = 1,\ \dots\ , L_{S,\mathbf{p},\boldsymbol{\Xi}}. \tag{B10}$$

Finally, differentiation of the corresponding objective function, given in the problem statement of Equation (24)

$$\tilde{e}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}} := -\sum_{i=2}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} C_i^2(\boldsymbol{\Xi}),$$

yields

$$\frac{\partial \tilde{e}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}}{\partial \xi_{k,j_k}} = -2 \sum_{i=2}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} C_i(\boldsymbol{\Xi}) \frac{\partial C_i(\boldsymbol{\Xi})}{\partial \xi_{k,j_k}}. \tag{B11}$$

So the derivatives of $\psi^{k}_{i_k,p_k,\xi_k}(x_k)$ obtained by Equation (B3) are substituted into Equation (B8) to obtain the derivatives of $\Psi^{u}_{\mathbf{i}_u,\mathbf{p}_u,\boldsymbol{\Xi}_u}(\mathbf{x}_u)$, which are substituted into Equation (B10) to obtain the derivatives of $C_i(\boldsymbol{\Xi})$, which are finally substituted into Equation (B11) to obtain the derivatives of $\tilde{e}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}$.

Alternatively, if the SDD expansion coefficients are computed by the SLS regression

$$\hat{\mathbf{c}}(\boldsymbol{\Xi}) := \left(\hat{C}_1(\boldsymbol{\Xi}),\ \dots\ , \hat{C}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\boldsymbol{\Xi}) \right)^{\intercal} = (\mathbf{A}^{\intercal}\mathbf{A})^{-1}\mathbf{A}^{\intercal}\mathbf{b}, \tag{B12}$$

where

$$\mathbf{A} := \begin{bmatrix} \Psi_1(\mathbf{x}^{(1)};\boldsymbol{\Xi}) & \dots & \Psi_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\mathbf{x}^{(1)};\boldsymbol{\Xi}) \\ \vdots & \ddots & \vdots \\ \Psi_1(\mathbf{x}^{(L)};\boldsymbol{\Xi}) & \dots & \Psi_{L_{S,\mathbf{p},\boldsymbol{\Xi}}}(\mathbf{x}^{(L)};\boldsymbol{\Xi}) \end{bmatrix} \text{ and } \mathbf{b} := \left(y(\mathbf{x}^{(1)}),\ \dots\ , y(\mathbf{x}^{(L)}) \right)^{\intercal}$$

and $\{\mathbf{x}^{(l)}, y(\mathbf{x}^{(l)})\}_{l=1}^{L}$ is the input-output data set of size L, the sensitivity of the coefficients can be evaluated by differentiation of Equation (B12), resulting in

$$\frac{\partial \hat{\mathbf{c}}(\boldsymbol{\Xi})}{\partial \xi_{k,j_k}} = (\mathbf{A}^{\intercal}\mathbf{A})^{-1} \left[\frac{\partial \mathbf{A}^{\intercal}}{\partial \xi_{k,j_k}}\mathbf{b} - \left(\frac{\partial \mathbf{A}^{\intercal}}{\partial \xi_{k,j_k}}\mathbf{A} + \mathbf{A}^{\intercal}\frac{\partial \mathbf{A}}{\partial \xi_{k,j_k}} \right) \hat{\mathbf{c}}(\boldsymbol{\Xi}) \right], \tag{B13}$$

where $\partial\mathbf{A}/\partial\xi_{k,j_k}$ simply contains evaluations of the corresponding derivatives of the multivariate orthonormalized B-spline functions, given in Equation (B8), at the input data points $\mathbf{x}^{(l)}$. Finally, differentiation of the corresponding objective function, given in the problem statement of Equation (25),

$$\hat{e}_{L_{S,\mathbf{p},\boldsymbol{\Xi}}} := \frac{1}{L}\sum_{l=1}^{L} \left[y(\mathbf{x}^{(l)}) - \sum_{i=1}^{L_{S,\mathbf{p},\boldsymbol{\Xi}}} \hat{C}_i(\boldsymbol{\Xi})\Psi_i(\mathbf{x}^{(l)};\boldsymbol{\Xi}) \right]^2,$$

yields

$$\frac{\partial \hat{e}_{L_{S,\mathbf{p},\Xi}}}{\partial \xi_{k,j_k}} = -\frac{2}{L}\sum_{l=1}^{L}\sum_{i=1}^{L_{S,\mathbf{p},\Xi}}\left(\frac{\partial \hat{C}_i(\Xi)}{\partial \xi_{k,j_k}}\Psi_i(\mathbf{x}^{(l)};\Xi) + \hat{C}_i(\Xi)\frac{\partial \Psi_i(\mathbf{x}^{(l)};\Xi)}{\partial \xi_{k,j_k}}\right)\left(y(\mathbf{x}^{(l)}) - \sum_{j=1}^{L_{S,\mathbf{p},\Xi}}\hat{C}_j(\Xi)\Psi_j(\mathbf{x}^{(l)};\Xi)\right). \tag{B14}$$

So the derivatives of $\psi^k_{i_k,p_k,\xi_k}(x_k)$ obtained by Equation (B3) are substituted into Equation (B8) to obtain the derivatives of $\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{x}_u)$, which are evaluated at the input data points and substituted into Equation (B13) to obtain the derivatives of $\hat{C}_i(\Xi)$, which are finally substituted along with the evaluations of the derivatives of $\Psi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{x}_u)$ into Equation (B14) to obtain the derivatives of $\hat{e}_{L_{S,\mathbf{p},\Xi}}$.

B.3 Sensitivity with nonorthonormal Basis

When a nonorthonormal basis set is employed for the function expansion, as described in Section 6.3, the gradients of the corresponding objective function with respect to the internal knot locations are simpler to compute. In this case, the first step is to compute the gradients of the multivariate B-spline functions, related to $B^k_{i_k,p_k,\xi_k}(x_k)$ by

$$\Phi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{x}_u) = \prod_{k\in u} B^k_{i_k,p_k,\xi_k}(x_k), \quad \mathbf{i}_u = (i_{k_1},\ \dots\ , i_{k_{|u|}}) \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}. \tag{B15}$$

Differentiating Equation (B15) yields

$$\frac{\partial \Phi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{x}_u)}{\partial \xi_{k,j_k}} = \sum_{s\in u}\frac{\partial B^s_{i_s,p_s,\xi_s}(x_s)}{\partial \xi_{k,j_k}}\prod_{tu\in t\neq s} B^t_{i_t,p_t,\xi_t}(x_t), \quad \mathbf{i}_u = (i_{k_1},\ \dots\ , i_{k_{|u|}}) \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}, \tag{B16}$$

where $\partial B^k_{i_k,p_k,\xi_k}(x_s)/\partial \xi_{k,j_k}$ are evaluated according to Equation (B1). Once again, this sum only needs to be evaluated over the values of s that pertain to coordinate directions that share a knot sequence with coordinate direction k. The multivariate B-spline functions must then be converted to their single-index analogues according to

$$\left\{\Phi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{X}_u) : 1 \le |u| \le S,\ \mathbf{i}_u \in \overline{\mathcal{I}}_{u,\mathbf{n}_u}\right\} = \left\{\Phi_2(\mathbf{X};\Xi),\ \dots\ ,\Phi_{L_{S,\mathbf{p},\Xi}}(\mathbf{X};\Xi)\right\}, \quad \Phi_1(\mathbf{X};\Xi) = 1.$$

The expansion coefficients are then calculated according to

$$\hat{\overline{\mathbf{c}}}(\Xi) := \left(\hat{\overline{C}}_1(\Xi),\ \dots\ ,\hat{\overline{C}}_{L_{S,\mathbf{p},\Xi}}(\Xi)\right)^{\mathsf{T}} = \left(\overline{\mathbf{A}}^{\mathsf{T}}\overline{\mathbf{A}}\right)^{-1}\overline{\mathbf{A}}^{\mathsf{T}}\mathbf{b}, \tag{B17}$$

where

$$\overline{\mathbf{A}} := \begin{bmatrix} \Phi_1(\mathbf{x}^{(1)};\Xi) & \dots & \Phi_{L_{S,\mathbf{p},\Xi}}(\mathbf{x}^{(1)};\Xi) \\ \vdots & \ddots & \vdots \\ \Phi_1(\mathbf{x}^{(L)};\Xi) & \dots & \Phi_{L_{S,\mathbf{p},\Xi}}(\mathbf{x}^{(L)};\Xi) \end{bmatrix}.$$

The sensitivity of the coefficients can then be evaluated by differentiation of Equation (B17), resulting in

$$\frac{\partial \hat{\overline{\mathbf{c}}}(\Xi)}{\partial \xi_{k,j_k}} = \left(\overline{\mathbf{A}}^{\mathsf{T}}\overline{\mathbf{A}}\right)^{-1}\left[\frac{\partial \overline{\mathbf{A}}^{\mathsf{T}}}{\partial \xi_{k,j_k}}\mathbf{b} - \left(\frac{\partial \overline{\mathbf{A}}^{\mathsf{T}}}{\partial \xi_{k,j_k}}\overline{\mathbf{A}} + \overline{\mathbf{A}}^{\mathsf{T}}\frac{\partial \overline{\mathbf{A}}}{\partial \xi_{k,j_k}}\right)\hat{\overline{\mathbf{c}}}(\Xi)\right], \tag{B18}$$

where $\partial \overline{\mathbf{A}}/\partial \xi_{k,j_k}$ simply contains evaluations of the corresponding derivatives of the multivariate nonorthonormal, standard B-spline functions, given in Equation (B16), at the input data points $\mathbf{x}^{(l)}$. Finally, differentiation of the corresponding objective function, given in the problem statement of Equation (27),

$$\hat{e}_{L_{S,\mathbf{p},\Xi}} := \frac{1}{L}\sum_{l=1}^{L}\left[y(\mathbf{x}^{(l)}) - \sum_{i=1}^{L_{S,\mathbf{p},\Xi}}\hat{\overline{C}}_i(\Xi)\Phi_i(\mathbf{x}^{(l)};\Xi)\right]^2,$$

yields

$$\frac{\partial \hat{\tilde{e}}_{L_{S,\mathbf{p},\Xi}}}{\partial \xi_{k,j_k}} = -\frac{2}{L}\sum_{l=1}^{L}\sum_{i=1}^{L_{S,\mathbf{p},\Xi}}\left(\frac{\partial \hat{\tilde{C}}_i(\Xi)}{\partial \xi_{k,j_k}}\Phi_i(\mathbf{x}^{(l)};\Xi) + \hat{\tilde{C}}_i(\Xi)\frac{\partial \Phi_i(\mathbf{x}^{(l)};\Xi)}{\partial \xi_{k,j_k}}\right)\left(y(\mathbf{x}^{(l)}) - \sum_{j=1}^{L_{S,\mathbf{p},\Xi}}\hat{\tilde{C}}_j(\Xi)\Phi_j(\mathbf{x}^{(l)};\Xi)\right). \tag{B19}$$

So the derivatives of $B^k_{i_k,p_k,\xi_k}(x_k)$ obtained by Equation (B1) are substituted into Equation (B16) to obtain the derivatives of $\Phi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{x}_u)$, which are evaluated at the input data points and substituted into Equation (B18) to obtain the derivatives of $\hat{\tilde{C}}_i(\Xi)$, which are finally substituted along with the evaluations of the derivatives of $\Phi^u_{\mathbf{i}_u,\mathbf{p}_u,\Xi_u}(\mathbf{x}_u)$ into Equation (B19) to obtain the derivatives of $\hat{\tilde{e}}_{L_{S,\mathbf{p},\Xi}}$.