

## RESEARCH ARTICLE

WILEY

# Space and chaos-expansion Galerkin proper orthogonal decomposition low-order discretization of partial differential equations for uncertainty quantification

Peter Benner<sup>1,2</sup>  | Jan Heiland<sup>1,2</sup> 

<sup>1</sup>Computational Methods in Systems and Control Theory, Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg, Germany

<sup>2</sup>Faculty of Mathematics, Otto-von-Guericke-University, Magdeburg, Germany

## Correspondence

Jan Heiland, Computational Methods in Systems and Control Theory, Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany.  
Email: [heiland@mpi-magdeburg.mpg.de](mailto:heiland@mpi-magdeburg.mpg.de)

## Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: 314838170

## Abstract

The quantification of multivariate uncertainties in partial differential equations can easily exceed any computing capacity unless proper measures are taken to reduce the complexity of the model. In this work, we propose a multidimensional Galerkin proper orthogonal decomposition that optimally reduces each dimension of a tensorized product space. We provide the analytical framework and results that define and quantify the low-dimensional approximation. We illustrate its application for uncertainty modeling with polynomial chaos expansions and show its efficiency in a numerical example.

## KEYWORDS

model reduction, proper orthogonal decomposition, tensor spaces, uncertainty quantification

## 1 | INTRODUCTION

The statistically sound treatment of modeled uncertainties in simulations comes with significant additional computational costs. Since a deterministic model can already be arbitrarily complex, the computation of statistics for general problems may soon become infeasible unless some kind of model reduction is involved.

In this work, we propose a multidimensional Galerkin proper orthogonal decomposition (POD) that can simultaneously and optimally reduce the physical dimensions of the model and the dimensions related to the uncertainties.

For the quantification of uncertainties in partial differential equation (PDE) models and their numerical discretization, one may distinguish two categories of solvers<sup>1</sup> — sampling based methods, notably the *Monte-Carlo method*, and Galerkin-type projection methods. In this work, we focus on the latter. For a basic explanation and relevant references on the *Monte-Carlo method* and its extensions see Reference 1, for an application in elliptic PDEs see Reference 2, and for a combination with *stochastic collocation* and tensor techniques see Reference 3.

Galerkin-type methods for solving PDEs with uncertainties are also referred to as *spectral stochastic methods* and base on a *polynomial chaos expansion* (PCE) of the random variables describing, for example, uncertain parameters or the candidate solution; see, for example, Reference 4 for a thorough introduction and, for example, References 5 and 6, for example applications.

If the involved randomness is multivariate, multivariate ansatz functions are used to resolve the uncertainty; see Reference 7 for a general description of modeling uncertainties via product spaces as it has been used in, for example, in References 5, 6, and 8.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *International Journal for Numerical Methods in Engineering* published by John Wiley & Sons Ltd.

Since every dimension of the multivariation adds a dimension to the problem, a numerical discretization quickly becomes infeasible in terms of memory requirements, even if the dimensions are treated independent of each other. Several approaches to overcome this complexity have been proposed like *sparse grids*,<sup>9</sup> construction of reduced chaos expansions via, say, *Proper Generalized Decomposition*<sup>10,11</sup> or *Principal Component Analysis* or *Karhunen-Loève* expansions,<sup>12-14</sup> or the use of particular tensor formats to reduce or to handle the data more efficiently.<sup>15-18</sup>

In this work, we extend a Galerkin POD approach from space-time discretizations (as we have considered it in Reference 19 for optimization with PDE constraints) to general tensor product spaces including, in particular, those Hilbert spaces that are used for standard PCE formulations. This approach then defines low-dimensional bases both for the spatial and the uncertainty dimensions that optimally represent the data for a given PCE. As a result we drastically reduce the overall dimension of the problem which enables efficient uncertainty quantification and, prospectively, optimal control of systems with uncertain parameters.

Finding optimal representations for the dimensions is comparable to identifying low-rank tensor structures for the data, as it has been treated in References 15-17, 20, and 21. In contrast to these works, where a predefined structure is adaptively filled to approximate the solution, the proposed POD approach takes a given, possibly high-dimensional, data set and reduces it. The justification of this top-down approach is that the obtained reduction is optimally fitted to the given problem so that it can be used for further efficient explorations – mainly because this approach admits a direct interpretation of the bases for Galerkin discretizations. This relation to Galerkin projections defines the common ground with the PGD approaches,<sup>10</sup> where as in the other approaches mentioned in this paragraph, the optimal bases are constructed in an adaptive bottom-up fashion.

Most similar to our approach is the work<sup>12</sup> on *reduced chaos expansions* of coupled systems, where, basically, a Galerkin POD approach is used for two uncertainty dimensions. There, the authors start with a PCE of bivariate random coefficient and obtain optimal bases via the left and right eigenvectors of a generalized eigenvalue problem involving a covariance matrix and a mass matrix. This approach via the eigenvectors of a covariance matrix is one way to define a POD basis (see, e.g., Reference 22) while the inclusion of the mass matrix provides optimality in the relevant discrete function spaces; see Reference 23. Our approach extends the scope of this work by introducing the tensorized formulation that allows for reduction of multivariate uncertainties together with the spatial dimension in one framework.

The paper is organized as follows. In Section 2, we review the space-time Galerkin POD approach and how it extends to problems with an uncertainty dimension. Then we formulate the Galerkin POD for a product space of arbitrary dimensions and provide the POD compression algorithms and results. Next, in Section 3, we show that a PCE discretization exactly fits into this multidimensional Galerkin POD framework. In Section 4, we illustrate the use of the PCE and its POD reduction for a generic linear convection diffusion PDE. Finally, in Section 5, we provide a numerical example that shows the applicability and efficiency of this approach and show that a naive POD reduction based on random snapshots is not useful for PDEs with uncertain parameters.

## 2 | MULTIDIMENSIONAL GALERKIN POD

We briefly recall the basic notions from our previous work,<sup>19</sup> where we considered a Galerkin POD in the two dimensions space and time. The idea of locating space and time dependent functions

$$x \in L^2((0, T); L^2(\Omega)) : (0, T) \times \Omega \mapsto \mathbb{R},$$

that, for example, solve a partial differential equation, in the space-time outer product space

$$L^2((0, T)) \odot L^2(\Omega) := \{s \cdot y \mid s \in L^2((0, T)), y \in L^2(\Omega)\},$$

naturally extends to functions that depend on space, time and a random parameter  $\alpha$ , that is,

$$x(\alpha) \in L^2((0, T); L^2(\Omega)) : (0, T) \times \Omega \mapsto \mathbb{R},$$

in the space-time-uncertainty product space

$$L^2((0, T)) \odot L^2(\Omega) \odot L^2(\Gamma, \mathbb{P}_\alpha),$$

where  $\Gamma$  is the domain of the random parameter and  $\mathbb{P}_\alpha$  is the associated probability measure; see, for example, Reference 20 where stationary problems are treated in this setup.

Also, the approach of considering the approximation in the product of the finite dimensional spatial  $\mathcal{Y} \subset L^2(\Omega)$  and time  $\mathcal{S} \subset L^2((0, T))$  spaces extends to approximating  $x_\alpha$  in

$$\mathcal{S} \odot \mathcal{Y} \odot \mathcal{W},$$

where  $\mathcal{W}$  is the finite dimensional space that models a polynomial chaos expansion of the functions in  $L^2(\Gamma, \mathbb{P}_\alpha)$ .

And, finally one may approximate a function  $\mathbf{x}$  via its orthogonal projection onto  $\hat{\mathcal{S}} \odot \hat{\mathcal{Y}} \odot \hat{\mathcal{W}}$ , where

$$\hat{\mathcal{Y}} \subset \mathcal{Y}, \quad \hat{\mathcal{S}} \subset \mathcal{S} \quad \text{and} \quad \hat{\mathcal{W}} \subset \mathcal{W},$$

are chosen optimally with respect to  $\mathbf{x}$  for given dimensions of the subspaces.

We provide a general formulation of the product spaces, their discretization, and their optimal low-dimensional approximation. For  $i = 1, 2, \dots, N$ , let

$$\mathcal{V}_i := \text{span}\{\psi_i^1, \psi_i^2, \dots, \psi_i^{d_i}\},$$

be  $d_i$  dimensional Hilbert spaces with inner product  $(\cdot, \cdot)_{\mathcal{V}_i}$  and mass matrix

$$\mathbf{M}_{\mathcal{V}_i} = \left[ (\psi_i^k, \psi_i^\ell)_{\mathcal{V}_i} \right]_{k=1, \dots, d_i, \ell=1, \dots, d_i} \in \mathbb{R}^{d_i \times d_i}.$$

We will use the formal vector of the basis functions

$$\Psi_i = \begin{bmatrix} \psi_i^1 \\ \psi_i^2 \\ \vdots \\ \psi_i^{d_i} \end{bmatrix}, \quad (1)$$

to write, for example,

$$\mathbf{M}_{\mathcal{V}_i} = (\Psi_i, \Psi_i^\top)_{\mathcal{V}_i},$$

via applying the functional  $(\cdot, \cdot) : \psi_i^\ell \psi_i^k \mapsto (\psi_i^\ell, \psi_i^k)$  pointwise to the entries of the formal matrix  $\Psi_i \Psi_i^\top$ . Finally, let  $\mathbf{L}_{\mathcal{V}_i} \in \mathbb{R}^{d_i \times d_i}$  be a factor such that

$$\mathbf{M}_{\mathcal{V}_i} = \mathbf{L}_{\mathcal{V}_i} \mathbf{L}_{\mathcal{V}_i}^\top.$$

We consider the product space

$$\mathcal{V} = \mathcal{V}_1 \odot \mathcal{V}_2 \odot \dots \odot \mathcal{V}_N,$$

of spaces of square integrable functions with the inner product

$$(y, z)_{\mathcal{V}} = \int \int \dots \int y_i z_i \, d_1 \, d_2 \dots d_N,$$

where  $d_i$  denotes the measure associated with  $\mathcal{V}_i$ .

We represent a function  $x \in \mathcal{V}$  via

$$x = \sum_{k_1=1}^{d_1} \sum_{k_2=1}^{d_2} \dots \sum_{k_N=1}^{d_N} \mathbf{x}^{k_1 k_2 \dots k_N} \psi_1^{k_1} \psi_2^{k_2} \dots \psi_N^{k_N},$$

or, equivalently, via the  $N$ -dimensional tensor of the coefficients

$$\mathbf{X} = [\mathbf{x}^{k_1 k_2 \dots k_N}].$$

Note that

$$x = \text{vec}(\mathbf{X})^\top [\Psi_N \otimes \dots \otimes \Psi_2 \otimes \Psi_1], \quad (2)$$

where  $\otimes$  denotes the *Kronecker product* and where  $\text{vec}(\mathbf{X}) \in \mathbb{R}^{d_1 \cdot d_2 \cdots d_N}$  denotes the *vectorization* of a tensor  $\mathbf{X} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N}$  defined as the vector of all entries of  $\mathbf{X}$  in reverse lexicographic order. The *mode-1 matricization*  $\mathbf{X}^{(1)} \in \mathbb{R}^{d_1 \times (d_2 \cdot d_3 \cdots d_N)}$  that is used in the following relation is obtained, if, roughly speaking, the first dimension of  $\mathbf{X}$  is skipped in the vectorization and, instead, the resulting  $d_2 \cdot d_3 \cdots d_N$  vectors of dimension  $d_1$  are used as columns of the matrix  $\mathbf{X}^{(1)} \in \mathbb{R}^{d_1 \times (d_2 \cdot d_3 \cdots d_N)}$ . In the same way, the mode- $\mu$  matricization  $\mathbf{X}^{(\mu)} \in \mathbb{R}^{d_\mu \times (d_1 \cdots d_{\mu-1} d_{\mu+1} \cdots d_N)}$  is defined as the matrix with the vectors obtained from unrolling all dimensions except of the  $\mu$ -th dimension as columns. Moreover, for a matrix  $\mathbf{L} \in \mathbb{R}^{d_* \times d_\mu}$ , with  $d_*$  arbitrary, the  $\mu$ -mode tensor product  $\mathbf{L} \circ_\mu \mathbf{X} \in \mathbb{R}^{d_1 \times \cdots \times d_{\mu-1} \times d_* \times d_{\mu+1} \times \cdots \times d_N}$  is defined as the tensor that is obtained by applying the inverse of the  $\mu$ -mode matricization operation to  $\mathbf{L}\mathbf{X}^{(\mu)}$ .

**Lemma 1.** For a function  $x \in \mathcal{V}$  with its representation  $\mathbf{X}$  as in (2), one has

$$\begin{aligned} \|x\|_{\mathcal{V}}^2 &= \int \int \cdots \int x^2 d_1 d_2 \cdots d_N \\ &= \|\mathbf{L}_{\mathcal{V}_1}^T \mathbf{X}^{(1)} [\mathbf{L}_{\mathcal{V}_N} \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_2}]\|_F^2, \end{aligned}$$

where  $\mathbf{X}^{(1)}$  is the mode-1 matricization of the coefficient tensor  $\mathbf{X}$  and  $\|\cdot\|_F^2 := \|\text{vec}(\cdot)\|_2^2$  denotes the squared Frobenius norm of a tensor (or a vector).

*Proof.* We use the properties of the *Kronecker-product*  $\otimes$ , the  $\mu$ -mode tensor product  $\circ_\mu$ , the vectorization operator  $\text{vec}$ , and the  $\mu$ -mode matricization operator  $\cdot^{(\mu)}$  as introduced right before this Lemma to directly compute

$$\begin{aligned} \|x\|_{\mathcal{V}}^2 &= \int \int \cdots \int x^2 d_1 d_2 \cdots d_N \\ &= \text{vec}(\mathbf{X})^T \int \int \cdots \int [\Psi_N \Psi_N^T \otimes \cdots \otimes \Psi_2 \Psi_2^T \otimes \Psi_1 \Psi_1^T] d_1 d_2 \cdots d_N \text{vec}(\mathbf{X}) \\ &= \text{vec}(\mathbf{X})^T [\mathbf{M}_{\mathcal{V}_N} \otimes \cdots \otimes \mathbf{M}_{\mathcal{V}_2} \otimes \mathbf{M}_{\mathcal{V}_1}] \text{vec}(\mathbf{X}) \\ &= \|\left[\mathbf{L}_{\mathcal{V}_N}^T \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_2}^T \otimes \mathbf{L}_{\mathcal{V}_1}^T\right] \text{vec}(\mathbf{X})\|_2^2 \\ &= \|\left[\mathbf{L}_{\mathcal{V}_N}^T \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_2}^T \otimes I\right] \left[I \otimes \cdots \otimes I \otimes \mathbf{L}_{\mathcal{V}_1}^T\right] \text{vec}(\mathbf{X})\|_2^2 \\ &= \|\left[\mathbf{L}_{\mathcal{V}_N}^T \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_2}^T \otimes I\right] \text{vec}(\mathbf{L}_{\mathcal{V}_1}^T \circ_1 \mathbf{X})\|_2^2 \\ &= \|\text{vec}\left(\left[\mathbf{L}_{\mathcal{V}_N}^T \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_2}^T\right] \circ_2 (\mathbf{L}_{\mathcal{V}_1}^T \circ_1 \mathbf{X})\right)\|_2^2 \\ &= \|\left[\mathbf{L}_{\mathcal{V}_N}^T \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_2}^T\right] \circ_2 (\mathbf{L}_{\mathcal{V}_1}^T \circ_1 \mathbf{X})\|_F^2 \\ &= \|(\mathbf{L}_{\mathcal{V}_1}^T \circ_1 \mathbf{X})^{(1)} [\mathbf{L}_{\mathcal{V}_N} \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_2}]\|_F^2 \\ &= \|\mathbf{L}_{\mathcal{V}_1}^T \mathbf{X}^{(1)} [\mathbf{L}_{\mathcal{V}_N} \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_2}]\|_F^2. \end{aligned}$$

■

By permutations of the tensor  $\mathbf{X}$ , the dimension associated with any  $\mathcal{V}_i$  can take the role of the first dimension with  $\mathcal{V}_1$  in the formula of Lemma 1. To avoid technicalities, we will consider permutations that simply cycle through the dimensions. Therefore, we introduce the operator that permutes a tensor

$$\Pi : \mathbf{X} \in \mathbb{R}^{d_1 \times d_2 \times \cdots \times d_N} \mapsto \Pi \mathbf{X} \in \mathbb{R}^{d_2 \times \cdots \times d_N \times d_1}$$

via

$$[(\Pi \mathbf{X})^{k_1 k_2 \cdots k_N}] = [\mathbf{X}^{k_2 \cdots k_N k_1}].$$

Note that  $\Pi^N \mathbf{X} = \mathbf{X}$  and that, for matrices  $\mathbf{M}$  (where  $N = 2$ ), it holds that  $\Pi \mathbf{M} = \mathbf{M}^T$ .

**Corollary 1** (of Lemma 1). For any  $i \in \{1, \dots, N\}$ , the norm of  $x \in \mathcal{V}$  can be expressed as

$$\|x\|_{\mathcal{V}}^2 = \|\mathbf{L}_{\mathcal{V}_i}^T (\Pi^{i-1} \mathbf{X})^{(1)} [\mathbf{L}_{\mathcal{V}_{i-1}} \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_1} \otimes \mathbf{L}_{\mathcal{V}_N} \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_{i+1}}]\|_F^2,$$

with the convention that  $\mathbf{L}_{\mathcal{V}_{i-1}} \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_1}$  is void for  $i = 1$  as is  $\mathbf{L}_{\mathcal{V}_N} \otimes \cdots \otimes \mathbf{L}_{\mathcal{V}_{i+1}}$  for  $i = N$ .

With these expressions for the norm of the function  $x \in \mathcal{V}$  related to a tensor  $\mathbf{X}$  via (2), we can provide an interpretation of the *higher-order singular value decomposition*<sup>24</sup> in terms of low-dimensional space discretizations as it is the backbone of the POD.

**Theorem 1.** *Given  $x \in \mathcal{V}$ . For any  $i \in \{1, \dots, N\}$  and for a corresponding  $\hat{d}_i \leq d_i$ , the space spanned by*

$$\hat{\Psi}_i = \begin{bmatrix} \hat{\psi}_i^1 \\ \hat{\psi}_i^2 \\ \vdots \\ \hat{\psi}_i^{\hat{d}_i} \end{bmatrix} := V_{i,\hat{d}_i}^\top \mathbf{L}_{\mathcal{V}_i}^{-1} \begin{bmatrix} \psi_i^1 \\ \psi_i^2 \\ \vdots \\ \psi_i^{d_i} \end{bmatrix} = V_{i,\hat{d}_i}^\top \mathbf{L}_{\mathcal{V}_i}^{-1} \Psi_i,$$

where  $V_{i,\hat{d}_i}$  is the matrix of the  $\hat{d}_i$  leading left singular vectors of

$$\mathbf{L}_{\mathcal{V}_i}^\top (\Pi^{i-1} \mathbf{X})^{(1)} [\mathbf{L}_{\mathcal{V}_{i-1}} \otimes \dots \otimes \mathbf{L}_{\mathcal{V}_1} \otimes \mathbf{L}_{\mathcal{V}_N} \otimes \dots \otimes \mathbf{L}_{\mathcal{V}_{i+1}}],$$

optimally approximates  $\mathcal{V}_i$  in the sense that  $x$  is best approximated in

$$\mathcal{V}_1 \cdot \mathcal{V}_2 \cdot \dots \cdot \mathcal{V}_{i-1} \cdot \hat{\mathcal{V}}_i \cdot \mathcal{V}_{i+1} \cdot \dots \cdot \mathcal{V}_N,$$

in the  $\mathcal{V}$ -norm over all subspaces of  $\mathcal{V}_i$  of dimension  $\hat{d}_i$ .

*Proof.* For  $i = 1$ , the claim follows directly from [Lem. 2.5]<sup>19</sup> with considering  $\mathcal{V}_1 \cdot \mathcal{W}$ , and  $\mathcal{W} := \mathcal{V}_2 \cdot \mathcal{V}_3 \cdot \dots \cdot \mathcal{V}_N$ . For any other  $i$ , one can apply Corollary 1 first. ■

For the overall projection error between  $x$  and its projection  $\hat{x}$  onto

$$\hat{\mathcal{V}}_1 \cdot \hat{\mathcal{V}}_2 \cdot \dots \cdot \hat{\mathcal{V}}_N,$$

with  $\hat{\mathcal{V}}_i$  of dimension  $\hat{d}_i$  as defined in Theorem 1, one has that

$$\|x - \hat{x}\|_{\mathcal{V}}^2 \leq \sum_{k_1=\hat{d}_1+1}^{d_1} \sigma_{k_1}^{(1)2} + \sum_{k_2=\hat{d}_2+1}^{d_2} \sigma_{k_2}^{(2)2} + \dots + \sum_{k_N=\hat{d}_N+1}^{d_N} \sigma_{k_N}^{(N)2}, \quad (3)$$

where  $\sigma_k^{(i)}$  is the  $k$ th singular value of  $\mathbf{X}^{(i)}$  as they appear in the SVD for the definition of  $\hat{\mathcal{V}}_i$ . The estimate (3) follows directly from [Equation (24) in Reference 24] if one takes into account the scaling by the factors of the mass matrices. Note that while a single  $\mathcal{V}_i$  is optimally approximated by  $\hat{\mathcal{V}}_i$  by virtue of Theorem 1, the approximation of  $\mathcal{V}$  by  $\prod_{i=1}^N \hat{\mathcal{V}}_i$  might not be optimal in the same sense; see the discussion in [p. 1267 in Reference 24].

### 3 | POLYNOMIAL CHAOS EXPANSION AS PRODUCT SPACE

Let

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N),$$

be a tuple of random variables  $\alpha_i$  that take values in a domain  $\Gamma_i \subset \mathbb{R}$  and that are distributed according to a probability measure  $\mathbb{d}\mathbb{P}_i$ . If  $\tilde{y}$  is a function that depends on  $\alpha$ , that for any value of  $\alpha$  takes on values in a Hilbert space, say,  $L^2(\Omega)$  for a domain  $\Omega$  in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ , and that has a bounded variance with respect to  $\alpha$ , one may approximate  $\tilde{y}$  by a suitable

$$y \in L^2(\Omega) \otimes L^2(\Gamma_1; \mathbb{d}\mathbb{P}_1) \otimes L^2(\Gamma_2; \mathbb{d}\mathbb{P}_2) \otimes \dots \otimes L^2(\Gamma_N; \mathbb{d}\mathbb{P}_N). \quad (4)$$

Note that  $y$  is a random variable and that the expected value  $\mathbb{E}y \in L^2(\Omega)$  of  $y$  is defined as

$$\mathbb{E}y = \int_{\Gamma_N} \cdots \int_{\Gamma_2} \int_{\Gamma_1} y \, d\mathbb{P}_1 \, d\mathbb{P}_2 \cdots d\mathbb{P}_N.$$

A finite dimensional approximation  $y$  to  $\tilde{y}$  can be sought in

$$\mathcal{V} = \mathcal{V}_0 \odot \mathcal{V}_1 \odot \mathcal{V}_2 \odot \cdots \odot \mathcal{V}_N, \quad (5)$$

where  $\mathcal{V}_0 \subset L^2(\Omega)$  is a Finite Element space and where, for  $i = 1, \dots, N$ ,  $\mathcal{V}_i$  is a finite dimensional subspace of  $L^2(\Gamma_i; d\mathbb{P}_i)$  derived from a *Polynomial Chaos Expansion*. Here we will consider  $d_i$ -dimensional spaces

$$\mathcal{V}_i = \text{span}\{\psi_i^1, \psi_i^2, \dots, \psi_i^{d_i}\},$$

with  $\psi_i^k$  being the Lagrange polynomials of degree  $d_i - 1$  defined through the distinct Gaussian quadrature nodes

$$\{\alpha_i^1, \alpha_i^2, \dots, \alpha_i^{d_i}\} \subset \Gamma_i, \quad (6)$$

with respect to the measure  $d\mathbb{P}_i$ ; see Reference 25 for formulas and algorithms. With the corresponding quadrature weights

$$\{w_i^1, w_i^2, \dots, w_i^{d_i}\} \subset \mathbb{R}^{d_i},$$

the quadrature formula

$$\int_{\Gamma_i} z(\alpha_i) \, d\mathbb{P}_i \approx \sum_{k=1}^{d_i} w_i^k z(\alpha_i^k), \quad (7)$$

is exact for polynomials up to degree  $2d_i - 1$ . By virtue of this exactness, and since the correspondingly defined Lagrange polynomials are orthogonal and fulfill  $\psi_i^k(\alpha_i^j) = 1$  if  $k = j$  and  $\psi_i^k(\alpha_i^j) = 0$  if  $k \neq j$ , for the mass matrix  $\mathbf{M}_{\mathcal{V}_i}$ , one has that

$$\mathbf{M}_{\mathcal{V}_i} = \int_{\Gamma_i} \Psi_i \Psi_i^\top \, d\mathbb{P}_i = \begin{bmatrix} w_i^1 & & & \\ & w_i^2 & & \\ & & \ddots & \\ & & & w_i^{d_i} \end{bmatrix}.$$

*Remark 1.* By the fundamental theorem of Gaussian quadrature the nodes  $\{\alpha_i^1, \alpha_i^2, \dots, \alpha_i^{d_i}\}$  in (6) are the roots of that polynomial  $\phi_i^{d_i}$  of degree  $d_i$  that is orthogonal to all polynomials of degree  $d_i - 1$  with respect to the inner product with the measure  $d\mathbb{P}_i$ . For many probability measures  $d\mathbb{P}$ , there exists a commonly preferred choice of the corresponding orthogonal polynomials  $\{\phi^\ell\}_{\ell=1, \dots, d}$ , where the index  $\ell$  also denotes the degree of the polynomial, which can be constructed by explicit recursion formulas; see, for example, [Tab. 4 in Reference 7]. We further note the direct relation of these recursions to the Jacobi matrix (i.e., [Eqns. (4) and (5) and Tab. 1 in Reference 25]) that is used to define the interpolation points (6); see, as well, [Ch. 2.4 in Reference 26] for the relevant fundamental mathematical results.

## 4 | APPLICATION EXAMPLE

For a domain  $\Omega \subset \mathbb{R}^d$ , with  $d = 2$  or  $d = 3$ , for a given right hand side  $f \in L^2(\Omega)$  and a given vector field  $b \in [L^2(\Omega)]^d$ , we consider the generic *convection-diffusion* problem

$$b \cdot \nabla y - \nabla \cdot (\kappa_\alpha \nabla y) = f, \quad (8)$$

where we assume that the diffusivity coefficient depends on a random vector  $\alpha = (\alpha_1, \dots, \alpha_N)$ .

For the following derivation of the model equations, we assume homogeneous Dirichlet conditions or homogeneous Neumann conditions for the boundary. Nonzero boundary conditions can be included in standard ways. For well-posedness one may assume that there is at least one boundary patch with Dirichlet conditions, that the domain is bounded with a boundary smooth enough so that a *Poincaré*-type inequality holds, that  $b$  is sufficiently regular and bounded, and, most importantly, that  $\kappa_\alpha$  is positive and bounded away from zero. If this is the case, then unique solvability of a weak formulation in the corresponding Sobolev spaces follows by standard arguments. In particular, for given  $f$  and  $b$  and  $\kappa_\alpha > 0$ , we may well assume that system (8) has a unique solution  $y$  for any realization of  $\alpha$  and, thus,  $y$  itself can be seen as a random variable depending on  $\alpha$ .

As in standard finite element approaches, for every realization  $\alpha$ , we locate the corresponding solution  $y_\alpha$  in  $H_0^1(\Omega)$  and require (8) to hold in the weak sense, namely

$$\int_{\Omega} v(x)b(x) \cdot \nabla y_\alpha(x) + \kappa_\alpha \nabla v(x) \cdot \nabla y_\alpha(x) \, dx = \int_{\Omega} v(x)f(x) \, dx, \quad (9)$$

for all  $v \in H_0^1(\Omega)$ .

To account for the uncertainty, we assume the solution in the product space of the space variable and the uncertainty dimensions as in (4) and require (9) to hold in a weak sense in the  $\Gamma_i$  inner products (see Sec. 3 in Reference 7), that is,

$$\begin{aligned} & \int_{\Gamma_N} \cdots \int_{\Gamma_2} \int_{\Gamma_1} \int_{\Omega} vb \cdot \nabla y + \kappa_\alpha \nabla v \cdot \nabla y \, dx \, d\mathbb{P}_1 \, d\mathbb{P}_2 \cdots d\mathbb{P}_N = \\ & = \int_{\Gamma_N} \cdots \int_{\Gamma_2} \int_{\Gamma_1} \int_{\Omega} vf \, dx \, d\mathbb{P}_1 \, d\mathbb{P}_2 \cdots d\mathbb{P}_N, \end{aligned} \quad (10)$$

where now  $v$  is a trial function from the ansatz space

$$H_0^1(\Omega) \cdot L^2(\Gamma_1; d\mathbb{P}_1) \cdot L^2(\Gamma_2; d\mathbb{P}_2) \cdots L^2(\Gamma_N; d\mathbb{P}_N).$$

We may cluster the uncertainty dimensions  $\Gamma_i$  into  $\Gamma$  and write

$$\int_{\Gamma} v(\alpha) \, d\mathbb{P} \quad \text{instead of} \quad \int_{\Gamma_N} \cdots \int_{\Gamma_2} \int_{\Gamma_1} v(\alpha_1, \dots, \alpha_N) \, d\mathbb{P}_1 \, d\mathbb{P}_2 \cdots d\mathbb{P}_N.$$

This weak formulation both in the spatial and in the uncertainty domain, is readily discretized by finite elements in space and chaos expansion with quadrature in the uncertainty as follows.

For a finite dimensional approximation, let the FEM space  $\mathcal{V}_0$  be spanned by  $\Psi_0$  (compare (1)) and let  $\mathbf{A}_\alpha \in \mathbb{R}^{d_0 \times d_0}$  be the discrete convection/diffusion operator:

$$\mathbf{A}_\alpha = \int_{\Omega} \Psi_0(b \cdot \nabla \Psi_0^\top - \nabla \cdot \kappa_\alpha \nabla \Psi_0^\top) \, dx = \int_{\Omega} \Psi_0 b \cdot \nabla \Psi_0^\top + \kappa_\alpha \nabla \Psi_0 \cdot \nabla \Psi_0^\top \, dx,$$

where the products and the application of the differential operators are understood component-wise.

To save space in the formal derivation of the equations that include the Polynomial Chaos Expansions, we will *formally* use the *strong* differential operator

$$a_\alpha : y \mapsto b \cdot \nabla y - \nabla \cdot (\kappa_\alpha \nabla y).$$

With that, with discrete ansatz spaces as in (5), and with the ansatz for the solution

$$y = \text{vec}(\mathbf{Y})^\top [\Psi_N \otimes \cdots \otimes \Psi_1 \otimes \Psi_0] = [\Psi_N^\top \otimes \cdots \otimes \Psi_1^\top \otimes \Psi_0^\top] \text{vec}(\mathbf{Y}), \quad (11)$$

where  $\mathbf{Y}$  is the tensor of coefficients (see (2)), Equation (10) is discretized as

$$\int_{\Gamma} \int_{\Omega} [\Psi_N \otimes \cdots \otimes \Psi_1 \otimes \Psi_0] a_\alpha y \, dx \, d\mathbb{P} = \int_{\Gamma} \int_{\Omega} [\Psi_N \otimes \cdots \otimes \Psi_1 \otimes \Psi_0] f \, dx \, d\mathbb{P},$$

where the left-hand side, together with (11), becomes



$$\begin{aligned}
& \int_{\Gamma} \int_{\Omega} [\Psi_N \otimes \cdots \otimes \Psi_1 \otimes \Psi_0] a_{\alpha} [\Psi_N^T \otimes \cdots \otimes \Psi_1^T \otimes \Psi_0^T] dx \, d\mathbb{P} \text{vec}(\mathbf{Y}) = \\
& = \int_{\Gamma} \int_{\Omega} [\Psi_N \otimes \cdots \otimes \Psi_1 \otimes \Psi_0] [\Psi_N^T \otimes \cdots \otimes \Psi_1^T \otimes a_{\alpha} \Psi_0^T] dx \, d\mathbb{P} \text{vec}(\mathbf{Y}) \\
& = \int_{\Gamma} \int_{\Omega} [\Psi_N \Psi_N^T \otimes \cdots \otimes \Psi_1 \Psi_1^T \otimes \Psi_0 a_{\alpha} \Psi_0^T] dx \, d\mathbb{P} \text{vec}(\mathbf{Y}) \\
& = \int_{\Gamma} \left[ \Psi_N \Psi_N^T \otimes \cdots \otimes \Psi_1 \Psi_1^T \otimes \int_{\Omega} \Psi_0 a_{\alpha} \Psi_0^T dx \right] d\mathbb{P} \text{vec}(\mathbf{Y}) \\
& = \int_{\Gamma_N} \cdots \int_{\Gamma_2} \int_{\Gamma_1} [\Psi_N \Psi_N^T \otimes \cdots \otimes \Psi_1 \Psi_1^T \otimes \mathbf{A}_{\alpha}] d\mathbb{P}_1 \, d\mathbb{P}_2 \cdots d\mathbb{P}_N \text{vec}(\mathbf{Y}), \tag{12}
\end{aligned}$$

thanks to the linearity of the involved differential operators and the Kronecker products.

Next we successively approximate the integrals with respect to the probability measures by the corresponding quadrature rules (see (7)) to obtain

$$\begin{aligned}
& \int_{\Gamma_N} \cdots \int_{\Gamma_2} \int_{\Gamma_1} [\Psi_N \Psi_N^T \otimes \cdots \otimes \Psi_1 \Psi_1^T \otimes \mathbf{A}_{\alpha}] d\mathbb{P}_1 \, d\mathbb{P}_2 \cdots d\mathbb{P}_N \text{vec}(\mathbf{Y}) \\
& \approx \int_{\Gamma_N} \cdots \int_{\Gamma_2} \sum_{k_1=1}^{d_1} w_1^{k_1} [\Psi_N \Psi_N^T \otimes \cdots \otimes \Psi_1(\alpha_1^{k_1}) \Psi_1(\alpha_1^{k_1})^T \otimes A_{\alpha_1^{k_1}, \dots, \alpha_N^{k_N}}] d\mathbb{P}_2 \cdots d\mathbb{P}_N \text{vec}(\mathbf{Y}).
\end{aligned}$$

Since the Lagrange polynomials are a nodal basis, it holds for all  $i = 1, \dots, N$ , that  $\Psi_i(\alpha_i^{k_i}) = e_{k_i}$ , where  $e_{k_i} \in \mathbb{R}^{d_i}$  is the  $k_i$ th canonical basis vector. Accordingly, the coefficient matrix for  $\text{vec}(\mathbf{Y})$  becomes

$$\sum_{k_N=1}^{d_N} \cdots \sum_{k_2=1}^{d_2} \sum_{k_1=1}^{d_1} w_N^{k_N} \cdots w_2^{k_2} w_1^{k_1} [e_{k_N} e_{k_N}^T \otimes \cdots \otimes e_{k_2} e_{k_2}^T \otimes e_{k_1} e_{k_1}^T \otimes A_{\alpha_1^{k_1}, \dots, \alpha_N^{k_N}}], \tag{13}$$

which is a completely decoupled system for every combination  $(\alpha_1^{k_1}, \dots, \alpha_N^{k_N})$ .

To derive the Galerkin POD reduced system, we replace  $\Psi_i$  by  $\hat{\Psi}_i$ , for  $i = 0, 1, \dots, N$  in (12). We assume that the reduced bases were obtained as proposed by Theorem 1. The derivation, however, works for any (reduced) basis.

For illustration, we consider the case  $N = 1$ , that is, the spatial dimension and a univariate uncertainty. Then, the reduced system coefficient matrix reads

$$\begin{aligned}
& \int_{\Gamma} \int_{\Omega} [\hat{\Psi}_1 \hat{\Psi}_1^T \otimes \hat{\Psi}_0 a_{\alpha} \hat{\Psi}_0^T] dx \, d\mathbb{P} \text{vec}(\hat{\mathbf{Y}}) = \\
& = \int_{\Gamma} \left[ \hat{\Psi}_1 \hat{\Psi}_1^T \otimes \int_{\Omega} \hat{\Psi}_0 a_{\alpha} \hat{\Psi}_0^T dx \right] d\mathbb{P} \text{vec}(\hat{\mathbf{Y}}) \\
& = \int_{\Gamma} [\hat{\Psi}_1 \hat{\Psi}_1^T \otimes \hat{\mathbf{A}}_{\alpha}] d\mathbb{P} \text{vec}(\hat{\mathbf{Y}}) \\
& = \sum_{k_1=1}^{d_1} w_1^{k_1} [\hat{\Psi}_1(\alpha_1^{k_1}) \hat{\Psi}_1(\alpha_1^{k_1})^T \otimes \hat{\mathbf{A}}_{\alpha_1^{k_1}}] \text{vec}(\hat{\mathbf{Y}}). \tag{14}
\end{aligned}$$

Here, the operator  $\hat{\mathbf{A}}_{\alpha}$  is the POD projection of  $\mathbf{A}_{\alpha}$ :

$$\hat{\mathbf{A}}_{\alpha} = V_{0, \hat{d}_0}^T \mathbf{L}_{V_0}^{-1} \mathbf{A}_{\alpha} \mathbf{L}_{V_0}^{-T} V_{0, \hat{d}_0}, \tag{15}$$

whereas, for the given choice of  $\Psi_1$  and the weights  $w_1^{k_1}$ ,  $k_1 = 1, \dots, d_1$ , one has

$$\begin{aligned}
w_1^{k_1} [\hat{\Psi}_1(\alpha_1^{k_1}) \hat{\Psi}_1(\alpha_1^{k_1})^T] &= w_1^{k_1} V_{1, \hat{d}_1}^T \mathbf{L}_{V_1}^{-1} [\hat{\Psi}_1(\alpha_1^{k_1}) \hat{\Psi}_1(\alpha_1^{k_1})^T] \mathbf{L}_{V_1}^{-T} V_{1, \hat{d}_1} \\
&= w_1^{k_1} V_{1, \hat{d}_1}^T \left( w_1^{k_1} \right)^{-1/2} e_{k_1} e_{k_1}^T \left( w_1^{k_1} \right)^{-1/2} V_{1, \hat{d}_1} \\
&= V_{1, \hat{d}_1}^T e_{k_1} e_{k_1}^T V_{1, \hat{d}_1};
\end{aligned}$$



see Theorem 1. By orthonormality of the POD basis, one obtains that

$$\sum_{k_1=1}^{d_1} w_1^{k_1} \left[ \hat{\Psi}_1(\alpha_1^{k_1}) \hat{\Psi}_1(\alpha_1^{k_1})^T \right] = I;$$

see [Rem. 2.6] Reference 19, which, however, does not help when the terms are multiplied with the (nonconstant)  $\mathbf{A}_\alpha$  in (14). Accordingly, the reduced system does not decouple and, in this univariate case, requires the solution of a  $\hat{d}_0 \cdot \hat{d}_1$ -dimensional system.

The derivation of the general reduced multivariate systems goes along the same lines and results in a possibly fully coupled system of dimension  $\prod_{j=0}^N \hat{d}_j$  which still can be prohibitively large. For these cases, one may consider leaving a certain dimension, say  $\Gamma_N$ , unreduced and rather solve  $d_N$  systems of size  $\prod_{j=0}^{N-1} \hat{d}_j$ . Using this idea recursively one can balance the number of systems and their size.

## 5 | NUMERICAL EXAMPLE

In this section, we illustrate how the general ideas of Section 2, as they were made specific for PDEs with multivariate uncertainties in the parameter in Sections 3 and 4, can be applied for uncertainty quantification in numerical simulations.

Motivated by [example 3.1],<sup>20</sup> we consider a stationary convection diffusion problem as in (8) with uncertainty in the conductivity coefficient.

As the geometrical setup, let  $\Omega \subset \mathbb{R}^3$  be a cylindrical domain of diameter  $D_0 = 2$  without its core of diameter  $D_1 = 0.8$  that is subdivided into four subdomains  $\Omega_i$ ,  $i = 1, 2, 3, 4$ , as illustrated in Figure 1. The height of computational domain is  $H = 0.5$ .

To model the uncertainty in the conductivity coefficient  $\kappa$ , independently on each subdomain  $\Omega_i$ , we assume  $\kappa$  to be a random variable of a random parameter  $\alpha_i$  via

$$\kappa(\alpha)|_{\Omega_i} = \alpha_i,$$

and write  $\kappa(\alpha)$  to express the dependence on the random vector  $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ .

In the presented examples, we let  $\alpha_i$  be distributed on

$$\Gamma_i = [5 \cdot 10^{-4}, 1 \cdot 10^{-3}] \quad \text{for } i = 1, 2, 3, 4,$$

according to two different continuous probability distributions, namely the *uniform* and a *beta* distribution with density functions

$$\rho_i(\alpha_i) = c_u \quad \text{and} \quad \rho_i(z) = c_\beta s(\alpha_i) \cdot (1 - s(\alpha_i))^4,$$

respectively, where  $c_u$  and  $c_\beta$  are normalization parameters and where  $s : \Gamma_i \rightarrow [0, 1]$  is the affine linear mapping that shifts and scales the parameter range to align with the standard range. We will refer to the distributions as *uniform* and *beta-2-5\**.

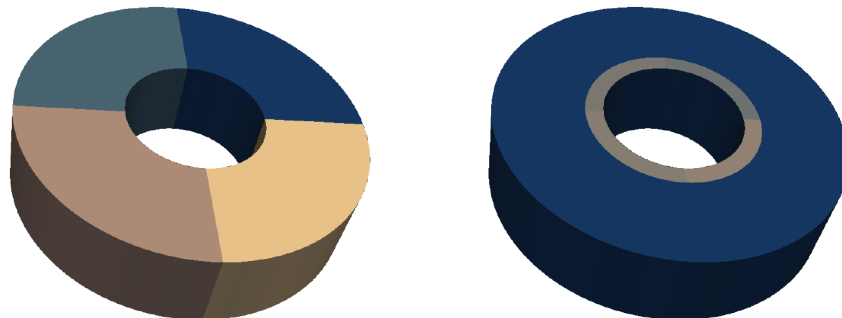


FIGURE 1 The subdomains and the boundary patch used for computing the quantity of interest  $\Upsilon$

As for boundary conditions, we apply zero Dirichlet conditions at the bottom of the domain and zero Neumann conditions elsewhere.

Without particular intentions, the convection  $b$  is chosen as

$$b(s_1, s_2, s_3) = \begin{bmatrix} (s_1^2 + s_2^2 - 1)s_2 \\ -(s_1^2 + s_2^2 - 1)s_1 \\ s_1^2 \sin(2s_3) \end{bmatrix},$$

and the inhomogeneity as

$$f(s_1, s_2, s_3) = \begin{cases} -\sin(2\pi s_1) \sin(4\pi s_2) s_3 (0.5 - s_3), & \text{for } (s_1, s_2, s_3) \in \Omega_1 \cup \Omega_3, \\ 0, & \text{for } (s_1, s_2, s_3) \in \Omega_2 \cup \Omega_4; \end{cases}$$

see Figure 2 for a snapshot of the solution for  $\kappa = 7.5 \cdot 10^{-4}$  being constant across the domain.

Moreover, we introduce  $Y$  as the quantity of interest defined as the spatially averaged value of the solution  $y$  over the surface of a concentric annular ring of diameter 0.1 that is aligned with the inner boundary at the top surface of the domain; see Figure 1 for the arrangement.

For the spatial discretization, we use continuous and piecewise linear finite elements on a discretization of the domain by tetrahedra. Although the mesh is refined at the critical parts, namely the edges of the domain and the surfaces where  $Y$  is computed or where the Dirichlet condition is applied, we need about 240,000 degrees of freedom for the spatial dimension to be within four digits of accuracy for some selected parameters; see Table 1. The values of interest of this numerical study are the expected value  $\mathbb{E}Y$  and the variance  $\mathbb{V}Y$  of the quantity of interest  $Y$  that we approximate by a PCE with various levels of refinement.

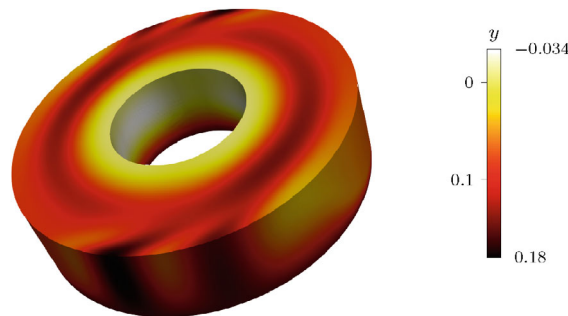


FIGURE 2 The solution  $y$  for  $\kappa = 7.5 \cdot 10^{-4}$  on all subdomains

TABLE 1 Computed  $Y$  for various realizations of the randomness  $\alpha$  on different meshes with increasing degrees of freedom (DOFs)

DOFs	$Y _{\alpha \equiv \bar{\alpha}}$	$Y _{\alpha \equiv \alpha_{\max}}$	$Y_{\alpha \equiv \alpha_{\min}}$	$Y_{\alpha = \alpha_{\delta}}$
90,458	0.60806	0.47561	0.87145	0.61057
127,771	0.60871	0.47652	0.87161	0.61185
154,545	0.60840	0.47663	0.86979	0.61168
192,786	0.60889	0.47712	0.87010	0.61223
<b>237,941</b>	<b>0.60927</b>	<b>0.47764</b>	<b>0.86979</b>	<b>0.61282</b>
262,702	0.60929	0.47780	0.86921	0.61320
282,310	0.60913	0.47758	0.86949	0.61279

Note: Here,  $\bar{\alpha} = 7.5 \cdot 10^{-4}$ ,  $\alpha_{\min} = 5 \cdot 10^{-4}$ ,  $\alpha_{\max} = 10 \cdot 10^{-4}$ , and  $\alpha_{\delta} = (\alpha_{\max}, \alpha_{\min}, \alpha_{\max}, \alpha_{\min})$  and the mesh used for the numerical study later with the seemingly correct digits are printed in bold letters.

In each uncertainty dimension, we apply a PCE with the same number  $d$  of Lagrange polynomials of degree  $d - 1$  that, through the particular choice of the interpolation points, are orthogonal with respect to the inner product induced by the considered density distribution; see Remark 1. We write  $\text{pce}[d]$  to refer to the PCE discretization of dimension  $d$  as well as the expected value/variance of  $Y$  based on this discretization.

Thus, the considered full order model will compute the solution via the system (13) formulated in the outer product of the spaces spanned by the FEM basis  $\Psi_0$  and the four PCE bases of dimension  $d$  represented by the corresponding orthogonal polynomials:

$$\Psi_0 = [\psi_{0,k}]_{k=1}^{237 \cdot 941} \quad \text{and} \quad \Psi_n = [\psi_j^{(d)}]_{j=1}^d, \quad n = 1, 2, 3, 4. \quad (16)$$

As can be seen in Tables 2 and 3, for computing the expected values and variances of  $Y$ , convergence of the PCE discretization is achieved already for low dimensions. However, even though the computations are well parallelized, the computation times for the moderate PCE discretizations are already in the order of days; see Table 4.

This gives motivation for the use of reduced-order models based on a small number of realizations to approximate  $\mathbb{E}Y$  and  $\mathbb{V}Y$ . As we will see, such an approach is capable to improve the estimate of a low-dimensional PCE discretization by one order of magnitude with little computational overhead.

We will investigate three approaches to compute projection bases to approximate the full order FEM simulation by a low-order approximation. Basically, every method computes a reduced approximation space  $\hat{\Psi}_0$  to replace  $\Psi_0$  in the ansatz space defined in (16). Practically, the computation of the reduced spaces comes with definitions of projection matrices, so that the reduced order FEM operator  $\hat{\mathbf{A}}_\alpha$  can be obtained from the full order operator  $\mathbf{A}_\alpha$  as in (15).

- **pcePOD** – Based on the 16 realizations defined by  $\alpha \in \otimes_{i=1}^4 \{\alpha_i^1, \alpha_i^2\}$ , where  $\{\alpha_i^1, \alpha_i^2\}$  are the two nodes that define the  $\text{pce}[2]$  polynomials,  $i = 1, \dots, 4$  and with the reduced-order basis  $\hat{\Psi}_0$  computed as suggested by Theorem 1.
- **wRB (weighted reduced bases)** – As suggested in Reference 27, the basis  $\hat{\Psi}_0$  is computed by a *greedy procedure* that sequentially selects those vector from a set of solutions that most improves the model approximation taking into account the underlying probability distribution through weighting. Following Reference 27, we consider the approximation error weighted by the corresponding probability distribution function. However, the error is not estimated (which is one of the core features of *reduced bases* approaches) but computed directly. This means a possibly overhead in computational effort but allows for unbiased evaluations of the approximation performance as there is no additional estimation error. As the training set, we use 16 random realizations of  $\alpha$ . To account for this uncertainty in the numerical setup, we repeat the experiment 10 times and report the median values.

**TABLE 2** The computed values for  $\mathbb{E}Y/\mathbb{V}Y$  in the `uniform` distribution case based on a PCE discretization, and the difference to the value of the finest computed discretization versus the dimension of the PCE

$d$	$\text{pce}[\cdot]$	Difference to $\text{pce}[5]$
2	0.6141314593/0.0026984301	$-1.5 \cdot 10^{-5} / -1.8 \cdot 10^{-5}$
3	0.6141464092/0.0027161019	$-1.1 \cdot 10^{-7} / -1.0 \cdot 10^{-7}$
4	0.6141465231/0.0027162057	$-1.0 \cdot 10^{-9} / -9.2 \cdot 10^{-10}$
5	0.6141465241/0.0027162066	—

**TABLE 3** The computed values for  $\mathbb{E}Y/\mathbb{V}Y$  in the `beta-2-5` distribution case based on a PCE discretization, and the difference to the value of the finest computed discretization versus the dimension of the PCE

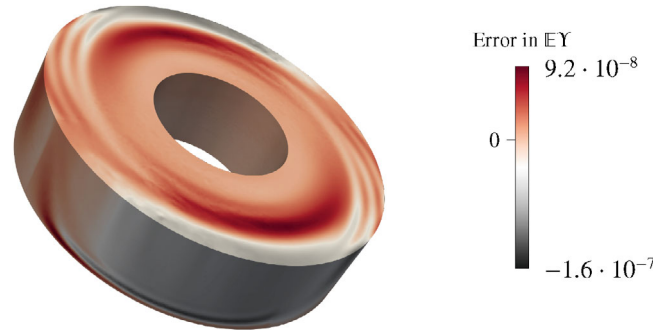
$d$	$\text{pce}[\cdot]$	Difference to $\text{pce}[5]$
2	0.6986880561/0.0014644986	$-5.5 \cdot 10^{-6} / -7.3 \cdot 10^{-6}$
3	0.6986935556/0.0014717976	$-3.4 \cdot 10^{-8} / -3.9 \cdot 10^{-8}$
4	0.6986935895/0.0014718366	$-4.0 \cdot 10^{-10} / -3.3 \cdot 10^{-10}$
5	0.6986935899/0.0014718369	—

**TABLE 4** Overview of the computational tasks and the timings relative to a single evaluation (1h) and a batch evaluation (1H) of 16 runs in parallel on 16 cores<sup>a</sup>

Computational task	Normalized time	Realization
Single solve	1h	407 s for this example on our computing cluster
Computing pce [2]	16h := 1H	642 s with parallelization on 16 cores
Computing pce [5]	645h $\approx$ 39H	About 7 h (on 16 cores, upscaled from 1H $\approx$ 642 s)
SVD of a matrix $\mathbf{X} \in \mathbb{R}^{237941 \times 16}$ of 16 solution vectors	eps	About 7 s (for six singular vectors)
Setup of the pcePOD ROM	1H+ eps	Computing pce [2] plus SVD
Setup of the wRB ROM	1H	Based on 16 full solves, possibly less with estimators, plus a negligible amount of time for comparisons
Computing the pce [5] with the ROM	eps	0.35 s for a 6 dimensional ROM

Note: The timings of the realization were are *wall clock times* (best out of 5 runs) taken for the presented numerical setup on a computing cluster.

<sup>a</sup>With Intel Xeon Silver 4110 (Skylake) CPUs with 8 Cores per CPU (64 Kb L1 Cache, 1024 kb L2 Cache), a Clockrate of 2.1 GHz (3.0 GHz max), and 12 MB Shared L3 Cache per CPU.



**FIGURE 3** The difference in  $\mathbb{E}Y$  in the beta-2-5 distribution case computed via the pce [5] and the pcePOD approximation of dimension  $k' = 6$  on the base of pce [2]

- **rndPOD** (random POD) – We use random realizations of  $\alpha$  for computing a POD basis  $\hat{\Psi}_0$  to be used instead of  $\Psi_0$  in the full order model (16). Again, to account for the involved randomness in the basis computation, we resort to the median value of 10 runs.

We set up the numerical experiments as follows: For varying dimensions (which we denote by  $k'$ ) of the reduced space  $\hat{\Psi}_0$ , we compute the pce [5] approximation of  $Y$ ,  $\mathbb{E}Y$ , and  $\mathbb{V}Y$  by means of the reduced-order model based on

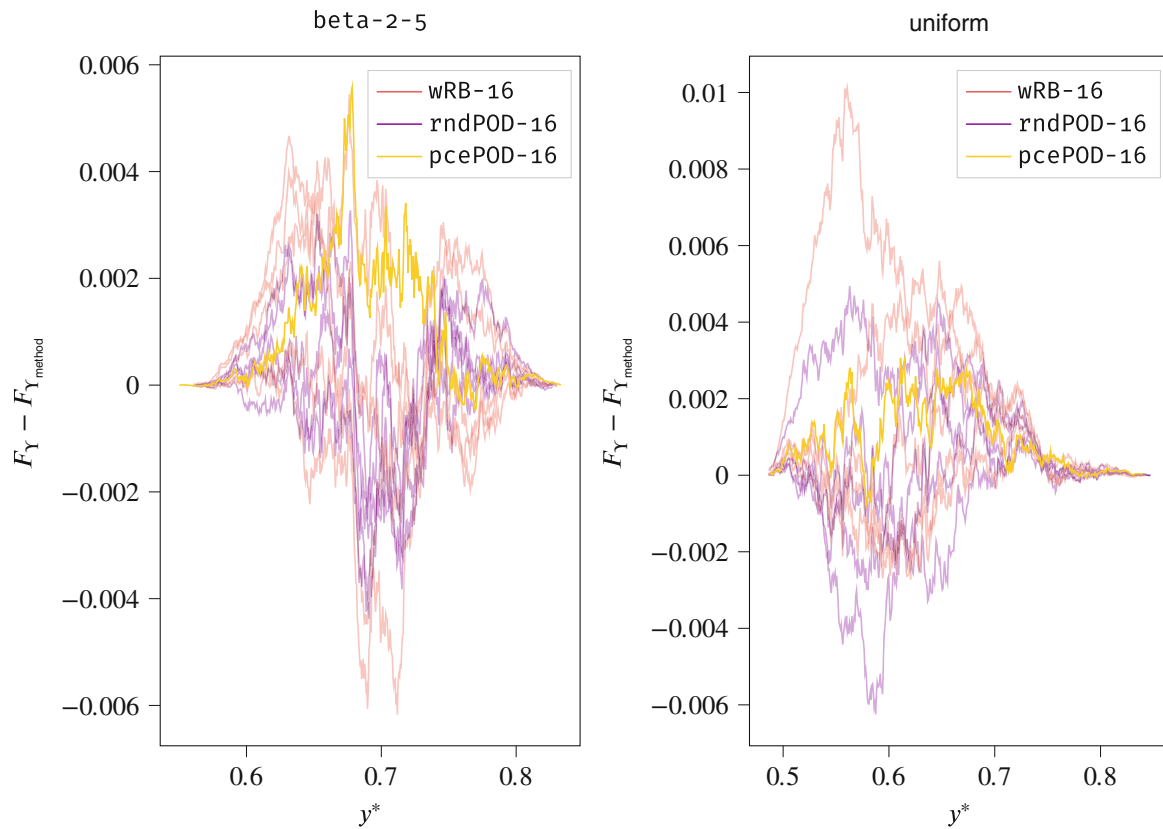
$$\hat{\Psi}_0 \quad \text{and} \quad \Psi_n^{(5)} = \left[ \psi_j^{(5)} \right]_{j=1}^5, \quad n = 1, 2, 3, 4,$$

and compare to the computed expected value and variance of  $Y$  computed by a pce [5] discretization of the full order model based on (16). A view of the difference in the expected value over the full domain is given in Figure 3.

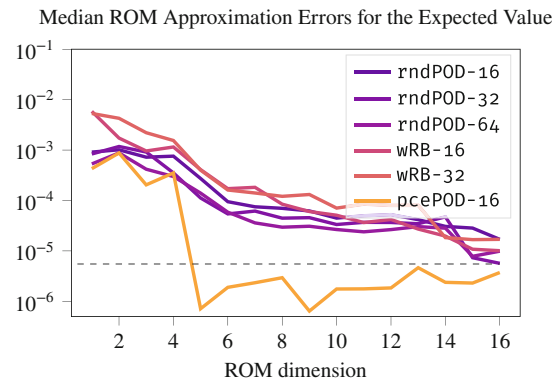
The computed results are laid out in Figures 4 and 5 for plots of the errors and in Figures 6 and 7 for their distribution over 10 realizations (of the wRB and rndPOD approximation) for the two considered distributions of the random variables. There, we also compare to the error level of a full order pce [2] approximation so that every reduced-order model (ROM) result that underbids this level is easily identified as a real improvement in terms of accuracy.

For a measure beyond the comparison of the first and second moment, we employed the *Kolmogorov metric* that measures the maximum difference between the cumulative distribution function of two random variables  $y_1$  and  $y_2$ , that is,

$$m(y_1, y_2) := \max_{y^* \in \mathbb{R}} |F_{y_1}(y^*) - F_{y_2}(y^*)|, \quad (17)$$

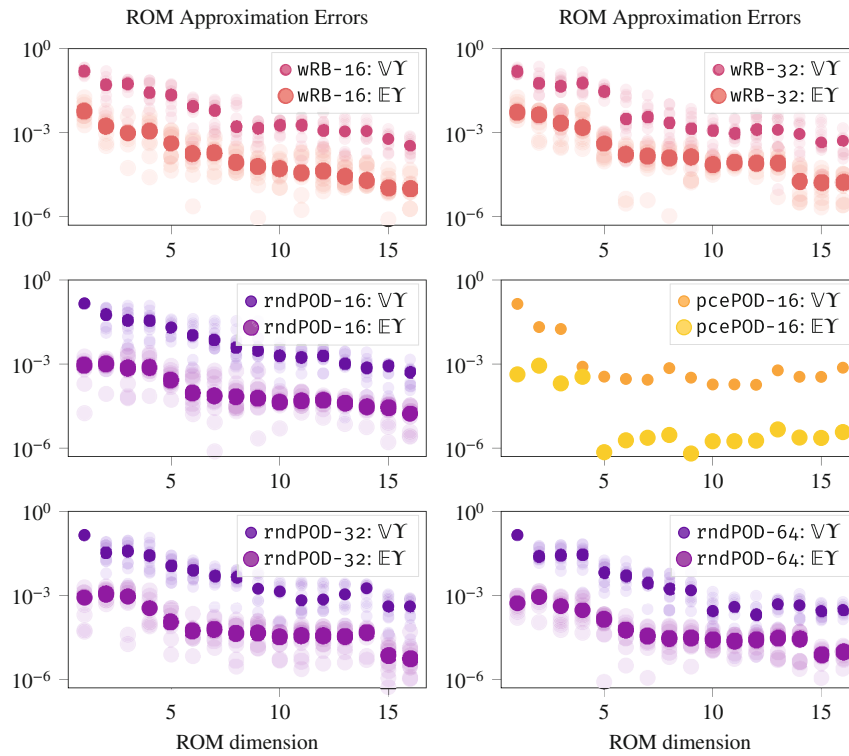


**FIGURE 4** Median errors in the approximation of  $\mathbb{E}Y$  in the beta-2-5 distribution case versus the dimension of the reduced-order model for pcePOD (based 16 full order solves) and for wRB and rndPOD (based on 16 and up to 64 full order solves). The dashed line marks the error level of the pce [2] approximation

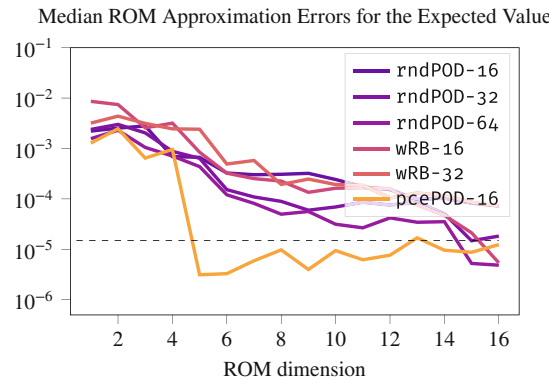


**FIGURE 5** Median errors in the approximation of  $\mathbb{E}Y$  in the uniform distribution case versus the dimension of the reduced-order model for pcePOD (based on 16 full order solves) and for wRB and rndPOD (based on 16 and up to 64 full order solves). The dashed line marks the error level of the pce [2] approximation

where for  $y^* \in \mathbb{R}$  the cumulative distribution function  $F_{y_i} : \mathbb{R} \rightarrow [0, 1]$  describes the probability that the value of the random variable  $y_i$  is less or equal  $y^*$ ,  $i = 1, 2$ . For our experiments, we let  $Y_{\text{pcePOD}}$ ,  $Y_{\text{rndPOD}}$ ,  $Y_{\text{wRB}}$  denote the approximation to the random variable  $Y$  that was computed by means of pcePOD, rndPOD, and wRB, respectively and compute the Kolmogorov metric to measure the distance to full order solution  $Y$ . The numerical evaluation of the metric was conducted for  $k' = 8$  for both considered distributions of  $\alpha$  using the empirical cumulative distribution function based on  $10^6$  samples and a subsequent linear interpolation to a common grid of 2000 points. The values of the measurements are displayed in Table 5. Again, the reported values for rndPOD and wRB are the median value of 10 realizations. Overall,



**FIGURE 6** Detailed plots of the median errors and their distribution over the 10 realizations in the approximation of  $\mathbb{E}Y$  and  $VY$  in the  $\text{beta-2-5}$  distribution case. For a comparison of the median errors in one plot see Figure 5



**FIGURE 7** Detailed plots of the median errors and their distribution over the 10 realizations in the approximation of  $\mathbb{E}Y$  and  $VY$  in the  $\text{uniform}$  distribution case. For a comparison of the median errors in one plot see Figure 5

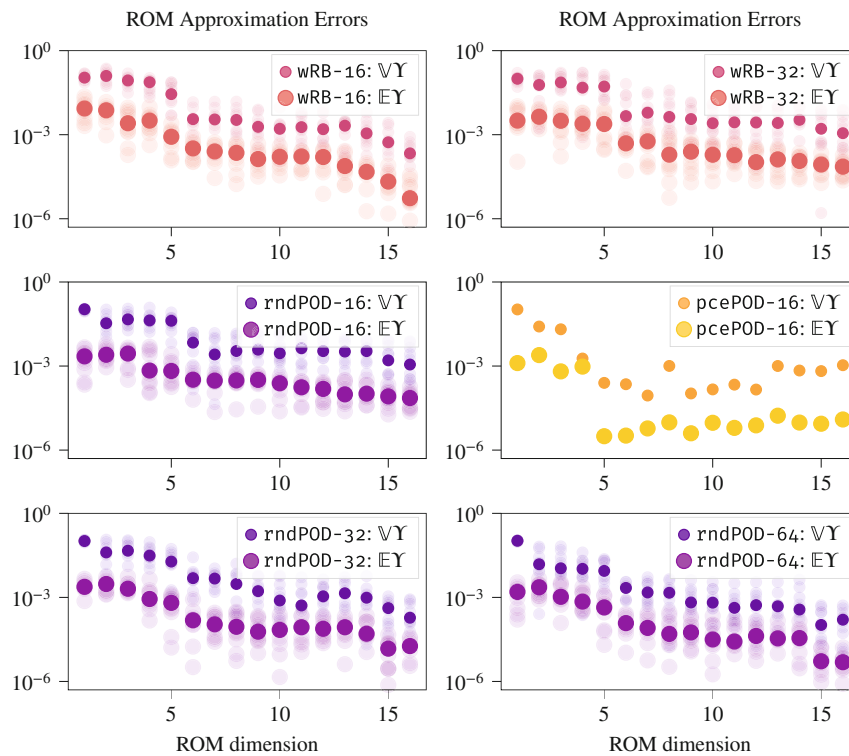
since the sampling introduces an extra layer of randomness to the evaluations, we have repeated the evaluations 5 times and reported the median value. In addition, we have plotted the differences  $F_{Y_{\text{pcePOD}}} - F_Y$ ,  $F_{Y_{\text{rndPOD}}} - F_Y$ , and  $F_{Y_{\text{wRB}}} - F_Y$  (the latter two for five realizations each) in Figure 8.

The improvement in terms of computational effort can be inferred from the measured timings laid out in Table 4. We note that for these small POD dimensions, the effort for computing the POD modes (around 7 s) and computing the pce [5] approximation with the reduced-order models (around 0.35 s) is negligible if compared to the time to compute the data or even the evaluation of pce [5] with the full model; see Table 4.

The presented results can be summarized and interpreted as follows:

- Concerning our proposed approach pcePOD, for both choices of the distribution of  $\alpha$ , we observe that for  $k' = 5$  the error in the reduced-order model drops below the error of a direct pce [2] approximation. Further enrichment of the basis does not show a reliable improvement, which lead us to the conclusion that little accuracy is added by the basis vectors beyond  $k' = 5$ .





**FIGURE 8** The estimated differences in the cumulative distribution function for  $Y$  estimated by a full order FEM and pce [5] approximation and by the reduced order models pcePOD, wRB, and rndPOD for the beta-2-5 (left) and the uniform distribution (right) of  $\alpha$ . Note that for display purposes not every data point is included in the plots so that the peak values are not as high as reported in the numbers in Table 5

**TABLE 5** The Kolmogorov metric for the approximation of the random variable  $Y$  (evaluated by a full order pce [5]) by the reduced order models namely the different reduced FEM basis of size  $k' = 8$  by pcePOD, wRB, and rndPOD and an approximation by a pce [2] discretization with the full order FEM basis

Method	pce [2]	pcePOD-16	wRB-16	wRB-32	rndPOD-16	rndPOD-32	
$m(Y, Y_{\text{method}})$	0.00741	0.00102	0.00212	0.00179	0.00208	0.00201	(beta-2-5)
$m(Y, Y_{\text{method}})$	0.00747	0.00114	0.00226	0.00229	0.00197	0.00200	(uniform)

Note: The upper line are the results for the beta-2-5 distribution and the lower line the results for the uniform distribution of  $\alpha$ .

- In contrast, the wRB and rndPOD approaches show a steady drop in the approximation error with increasing ROM dimensions, though at a higher median error level. Only for  $k' = 15, 16$ , these approaches show better results than the pce [2] approximation with the full order model.
- Since both wRB and rndPOD base on a training set with random realizations, we checked if the approximation can be improved by a larger training set. While for rndPOD a positive effect of a larger training set seems to be indicated by the data, for the wRB approximation some data points even suggest a reverse effect.
- Another observation with wRB and rndPOD is that the realization of the training set can influence the performance by about two orders of magnitude, as can be seen from the scatter plots of 10 realizations in Figures 6 and 7.
- We note that the definition of the bases is optimized towards the expected value of the solution, at least the bases in pcePOD and rndPOD are optimal in the corresponding inner product. Nonetheless, from the computation of  $\mathbb{V}Y$  by a pce [5] approximation with the ROM, we find that the approximation quality of  $\mathbb{V}Y$  follows that of  $\mathbb{E}Y$ .
- The differences in the cumulative distribution function (displayed in Table 5), as estimated via the Kolmogorov metric defined in (17), indicate that a lower order PCE approximation may well approximate the moments but seems to be missing the actual distribution. On the other hand, with roughly the same effort of a lower order PCE, one may set up the reduced order models to estimate the random variable with an improved accuracy of a factor between 3 and 7. Here,



**Code and Data Availability**

The source code of the implementations used to compute the presented results is available from:

doi:10.5281/zenodo.6874085  
github.com/mpimd-csc/multidim-genpod-uq

under the MIT license and is authored by Jan Heiland.

**FIGURE 9** Link to code and data

again the pcePOD approach clearly and reliably outperforms the wRB and rndPOD reduced order models. Moreover, as for the computation of the moments, using a larger set of training data for wRB and rndPOD did not show a visible improvement in accuracy.

The codes that set up, perform, and post process the numerical examples as well as the raw data of the presented cases are available as laid out in Figure 9.

## 6 | CONCLUSION

The theory of multidimensional Galerkin POD naturally applies to problems with multivariate uncertainties and can be made tractable for numerical experiments by exploiting the underlying tensor structures.

We have confirmed that multidimensional POD that includes Polynomial Chaos Expansions of the candidate solutions leads to a significant efficiency gain in the uncertainty quantification. For illustration, we considered a high-dimensional linear convection diffusion example with a multivariate uncertainty in the parameters. In comparison, the direct POD approach based on random realizations and a distribution-informed weighted greedy method to select a *reduced basis* from a training set proved to be similarly capable of improving the computation, though at significantly higher dimensions of the reduced-order model.

In this study, we showed that the very low-dimensional reduced-order models could even improve on the approximation of relevant statistical quantities by the training set. As a future direction we want to state that accurate uncertainty quantification with a very low-dimensional model provides an efficient approach to the more challenging task of optimal control of uncertain PDE systems.

## ACKNOWLEDGMENTS

This research was supported by the DFG through the Research Training Group *Mathematical Complexity Reduction* Project number 314838170. Open Access funding enabled and organized by Projekt DEAL.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Zenodo at <https://doi.org/10.5281/zenodo.6874085>, reference number zenodo.6874085.

## ENDNOTE

\*The numbers in the name for the beta distribution reflect the general way of defining the  $\beta(p, q)$ -distribution on  $(0, 1)$  via the two parameters  $p$  and  $q$  and the density function  $\rho(x) = \frac{x^{p-1}(1-x)^{q-1}}{\int_0^1 s^{p-1}(1-s)^{q-1} ds}$ .

## ORCID

Peter Benner  <https://orcid.org/0000-0003-3362-4103>

Jan Heiland  <https://orcid.org/0000-0003-0228-8522>

## REFERENCES

1. Soize C. *Brief Overview of Stochastic Solvers for the Propagation of Uncertainties*. Springer; 2017:133-139.
2. Cliffe KA, Giles MB, Scheichl R, Teckentrup AL. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Comput Vis Sci*. 2011;14(1):3-15. doi:10.1007/s00791-011-0160-x
3. Haji-Ali AL, Nobile F, Tamellini L, Tempone R. Multi-index stochastic collocation for random PDEs. *Comput. Methods Appl. Mech. Eng*. 2016;306:95-122. doi:10.1016/j.cma.2016.03.029

4. Ghanem RG, Spanos PD. *Stochastic Finite Elements: A Spectral Approach*. Springer; 1991.
5. Babuska I, Tempone R, Zouraris GE. Galerkin finite element approximations of stochastic elliptic partial differential equations. *SIAM J. Numer. Anal.* 2004;42(2):800-825. doi:[10.1137/S0036142902418680](https://doi.org/10.1137/S0036142902418680)
6. Matthies HG, Keese A. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Comput Methods Appl Mech Eng.* 2005;194(12):1295-1331. doi:[10.1016/j.cma.2004.05.027](https://doi.org/10.1016/j.cma.2004.05.027)
7. Soize C, Ghanem RG. Physical systems with random uncertainties: chaos representations with arbitrary probability measure. *SIAM J. Sci. Comput.* 2004;26(2):395-410. doi:[10.1137/S1064827503424505](https://doi.org/10.1137/S1064827503424505)
8. Ghanem R, Red-Horse J. Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element approach. *Phys D: Nonlinear Phenom.* 1999;133(1):137-144. doi:[10.1016/S0167-2789\(99\)00102-5](https://doi.org/10.1016/S0167-2789(99)00102-5)
9. Garcke J. Sparse grids in a nutshell. In: Garcke J, Griebel M, eds. *Sparse Grids and Applications*. Springer; 2013:57-80.
10. Nouy A. Proper generalized decompositions and separated representations for the numerical solution of high dimensional stochastic problems. *Arch. Comput. Methods Eng.* 2010;17(4):403-434. doi:[10.1007/s11831-010-9054-1](https://doi.org/10.1007/s11831-010-9054-1)
11. Tamellini L, Maître OL, Nouy A. Model reduction based on proper generalized decomposition for the stochastic steady incompressible Navier–Stokes equations. *SIAM J. Sci. Comput.* 2014;36(3):A1089-A1117. doi:[10.1137/120878999](https://doi.org/10.1137/120878999)
12. Arnst M, Ghanem R, Phipps E, Red-Horse J. Reduced chaos expansions with random coefficients in reduced-dimensional stochastic modeling of coupled problems. *Int. J. Numer. Methods Eng.* 2014;97(5):352-376. doi:[10.1002/nme.4595](https://doi.org/10.1002/nme.4595)
13. Audouze C, Nair PB. Galerkin reduced-order modeling scheme for time-dependent randomly parametrized linear partial differential equations. *Int. J. Numer. Methods Eng.* 2012;92(4):370-398. doi:[10.1002/nme.4341](https://doi.org/10.1002/nme.4341)
14. Bianchini I, Argiento R, Auricchio F, Lanzarone E. Efficient uncertainty quantification in stochastic finite element analysis based on functional principal components. *Comput. Mech.* 2015;56(3):533-549. doi:[10.1007/s00466-015-1185-7](https://doi.org/10.1007/s00466-015-1185-7)
15. Ballani J, Grasedyck L. Hierarchical tensor approximation of output quantities of parameter-dependent PDEs. *SIAM/ASA J. Uncertain. Quantif.* 2015;3(1):852-872. doi:[10.1137/140960980](https://doi.org/10.1137/140960980)
16. Benner P, Onwunta A, Stoll M. Low-rank solution of unsteady diffusion equations with stochastic coefficients. *SIAM/ASA J. Uncertain. Quantif.* 2015;3(1):622-649. doi:[10.1137/130937251](https://doi.org/10.1137/130937251)
17. Khoromskij BN. Tensor numerical methods for multidimensional PDEs: theoretical analysis and initial applications. *ESAIM: Proc.* 2015;48:1-28. doi:[10.1051/proc/201448001](https://doi.org/10.1051/proc/201448001)
18. Ullmann E. A Kronecker product preconditioner for stochastic Galerkin finite element discretizations. *SIAM J. Sci. Comput.* 2010;32(2):923-946. doi:[10.1137/080742853](https://doi.org/10.1137/080742853)
19. Baumann M, Benner P, Heiland J. Space-time Galerkin POD with application in optimal control of semi-linear parabolic partial differential equations. *SIAM J. Sci. Comput.* 2018;40(3):A1611-A1641. doi:[10.1137/17M1135281](https://doi.org/10.1137/17M1135281)
20. Garreis S, Ulbrich M. Constrained optimization with low-rank tensors and applications to parametric problems with PDEs. *SIAM J. Sci. Comput.* 2017;39(1):A25-A54. doi:[10.1137/16M1057607](https://doi.org/10.1137/16M1057607)
21. Khoromskij BN, Schwab C. Tensor-structured Galerkin approximation of parametric and stochastic elliptic PDEs. *SIAM J. Sci. Comput.* 2011;33(1):364-385. doi:[10.1137/100785715](https://doi.org/10.1137/100785715)
22. Rowley CW. Model reduction for fluids, using balanced proper orthogonal decomposition. *Internat. J. Bifur. Chaos Appl. Sci Eng.* 2005;15(3):997-1013. doi:[10.1142/s0218127405012429](https://doi.org/10.1142/s0218127405012429)
23. Baumann M, Heiland J, Schmidt M. Discrete input/output maps and their relation to proper orthogonal decomposition. In: Benner P, Bollhöfer M, Kressner D, Mehl C, Stykel T, eds. *Numerical Algebra, Matrix Theory, Differential-Algebraic Equations and Control Theory*. Springer International Publishing; 2015:585-608.
24. De Lathauwer L, De Moor B, Vandewalle J. A multilinear singular value decomposition. *SIAM J. Matrix Anal. Appl.* 2000;21(4):1253-1278. doi:[10.1137/S0895479896305696](https://doi.org/10.1137/S0895479896305696)
25. Fernandes AD, Atchley WR. Gaussian quadrature formulae for arbitrary positive measures. *Evol Bioinform Online.* 2007;2:251-259.
26. Wilf HS. *Mathematics for the Physical Sciences*. Dover Books Adv. Math. Dover Publications; 1978.
27. Chen P, Quarteroni A, Rozza G. A weighted reduced basis method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.* 2013;51(6):3163-3185. doi:[10.1137/130905253](https://doi.org/10.1137/130905253)

**How to cite this article:** Benner P, Heiland J. Space and chaos-expansion Galerkin proper orthogonal decomposition low-order discretization of partial differential equations for uncertainty quantification. *Int J Numer Methods Eng.* 2023;124(12):2801-2817. doi: [10.1002/nme.7229](https://doi.org/10.1002/nme.7229)