HAL open science

Recursive POD expansion for reaction-diffusion equation

M. Azaïez, F. Ben Belgacem, T. Chacón Rebollo

To cite this version:

M. Azaïez, F. Ben Belgacem, T. Chacón Rebollo. Recursive POD expansion for reaction-diffusion equation. Advanced Modeling and Simulation in Engineering Sciences, 2016, Model order reduction: POD, PGD and reduced bases, 3 (1), pp.1-22. 10.1186/s40323-016-0060-1 . hal-04458367

HAL Id: hal-04458367

https://hal.science/hal-04458367v1

Submitted on 20 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Autorisation HAL

HAL Authorization

Azaïez et al. Adv. Model. and Simul. in Eng. Sci.(2016)3:3
DOI 10.1186/s40323-016-0060-1

Advanced Modeling and Simulation in Engineering Sciences
a SpringerOpen Journal

RESEARCH ARTICLE Open Access

Recursive POD expansion for reaction-diffusion equation

M. Azaïez[1*], F. Ben Belgacem[2] and T. Chacón Rebollo[3,4]

*Correspondence: azaiez@enscbp.fr
[1]Bordeaux INP, I2M, (UMR CNRS 5295), 33607 Pessac, France
Full list of author information is available at the end of the article

Abstract

This paper focuses on the low-dimensional representation of multivariate functions. We study a recursive POD representation, based upon the use of the power iterate algorithm to recursively expand the modes retained in the previous step. We obtain general error estimates for the truncated expansion, and prove that the recursive POD representation provides a quasi-optimal approximation in L^2 norm. We also prove an exponential rate of convergence, when applied to the solution of the reaction-diffusion partial differential equation. Some relevant numerical experiments show that the recursive POD is computationally more accurate than the Proper Generalized Decomposition for multivariate functions. We also recover the theoretical exponential convergence rate for the solution of the reaction-diffusion equation.

Keywords: Recursive POD, High Order SVD, Model Reduction, Multivariate functions, PGD

Background

Model Reduction methods are nowadays basic numerical tools in the treatment of large-scale parametric problems appearing in real-world problems. They are applied with success, for instance, in signal processing, analysis of random data, solution of parametric partial differential equations and control problems, among others. In signal processing, Karhunen-Loève's expansion (KLE) provides a reliable procedure for a low dimensional representation of spatiotemporal signals (see [12,20]). Different research communities use different terminologies for the KLE. It is named the proper orthogonal decomposition (POD) in mechanical computation (see [3]), referred to as the principal components analysis (PCA) in statistics (see [17,18,24]) and data analysis or called singular value decomposition (SVD) in linear algebra (see [13]). These techniques allow large reduction of computational costs, thus making affordable the solution of many parametric problems of practical interest, otherwise out of reach. Let us mention some representative references [5,6,11,16,26,27], although this list, by far, is not exhaustive.

The extension of KLE to the tensor representation of multivariate functions is, however, a challenging problem. Real problems are quite often multivariate. Let us mention, for instance the analysis of multivariate stochastic variables, simulation and control of thermal flows and multi-component mechanics, among many others. Some recent techniques have been introduced to build low-dimensional tensor decompositions of multivariate functions and data. Among them, the High-Order Singular Value Decomposition

© 2016 Azaïez et al. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

(HOSVD) provides low-dimensional approximation of tensor data, in a similar way as the Singular Value Decomposition allows to approximate bi-variate data (see [7,8,21]). Also, the Proper Generalized Decomposition (PGD) appears to be well suited in many cases to approximate multivariate functions by low-dimensional varieties (see [14,15]). However, in general there is not an optimal tensor of rank three or larger to approximate a given high-dimensional tensor (see [9]).

In this paper we study an alternative method to build low-dimensional tensor decompositions of multivariate functions. This is a recursive POD (R-POD), based upon the successive application of the bivariate POD to each of the modes obtained in the previous step. In each step only one of the parameters is active, while the set of the remaining parameters is considered as a passive single parameter. We introduce a feasible version of the R-POD, in which the expansion is truncated whenever the singular values are smaller than a given threshold. This provides a fast algorithm, as only a small number of modes is computed, just those required to achieve a targeted error level.

As an application, we analyze the velocity of convergence of the R-POD applied to approximate the solution of the reaction-diffusion equation. We prove that the expansion converges with exponential rate. We use as main theoretical tool the Courant-Fischer-Weyl Theorem, that allows to reduce the error analysis of the POD expansion to that of the polynomial approximation of the function to be expanded. We also use the analytic dependence of the solution on the diffusivity and reaction rate coefficients, that yields the exponential convergence rate. This analysis is based upon the one introduced in [2]. Further, we use sub-sequent bounds for the singular values to construct a practical truncation error estimator, which is used to recursively compute the expansion by the Power Iterate (PI) method [1]. This avoids to compute the full singular value decomposition of the correlation matrix, just the modes needed to attempt a given error threshold are computed. The PI method provides a fast and reliable tool to build the POD expansion of bivariate functions, of which we take advantage to recursively build the R-POD expansion.

We present a battery of numerical tests, in which we apply the R-POD to the representation of three-variate functions, and in particular to the solution of the reaction-diffusion. We compare the R-POD to the PGD expansion. The PGD expansion can be interpreted as the PI method applied to the effective computation of the POD for bivariate functions (see [23]). We here consider its extension to multivariate functions. We obtain exponential convergence rates for both R-POD and PGD expansions, although the R-POD is in general more accurate than the PGD for the same number of modes. We also recover an exponential rate of convergence for the R-POD approximation of the solution of the reaction-diffusion equation, in quite good agreement with the theoretical expectations.

The guidelines of the paper are as follows. “The Karhunen-Loève decomposition on Hilbert spaces” section recalls the POD or Karhunen-Loève expansion in Hilbert spaces. In “Recursive POD representation” section, we introduce the recursive POD decomposition of multivariate functions, and make a general error analysis. “Analysis of solutions of the reaction-diffusion” section deals with the error analysis for the R-POD expansion of the solution of the reaction-diffusion equation. Finally, in “Numerical tests” section we present some numerical tests where we analyze the practical performances of the recursive POD decomposition of multivariate functions.

Notation—Let $X \subset \mathbb{R}^d$ be a given Lipschitz domain and G a measure space. We denote by $L^2(G, X)$ the Bochner space of measurable and square integrable functions from G on X (cf. [10]).

The Karhunen-Loève decomposition on Hilbert spaces

The Karhunen-Loève decomposition, also known as Proper Orthogonal decomposition (POD in the sequel) provides a technique to obtain low-dimensional approximations of parametric functions. To describe it, let us consider a Hilbert space H endowed with a scalar product $(\cdot,\cdot)_H$, and a parameter measure space G. Let us consider a function $f \in L^2(G, H)$, and introduce the POD operator

$$A : H \mapsto H, \quad \mathrm{A}v = \int_G f(\gamma)\,(f(\gamma), v)_H\, d\gamma \quad \text{for } v \in H.$$

POD operator is linear and bounded. Moreover, it is self-adjoint and non-negative. Indeed, it holds $A = B^*B$, where $B : H \mapsto L^2(G)$ and its adjoint operator $B^* : L^2(G) \mapsto H$ are given by

$$\begin{aligned} &(Bv)(\gamma) = (f(\gamma), v)_H \quad \text{for } v \in H, \\ &B^*\varphi = (\varphi, f)_{L^2(G)} = \int_G f(\gamma)\,\varphi(\gamma)\, d\gamma \quad \text{for } \varphi \in L^2(G). \end{aligned} \tag{1}$$

Furthermore, the operator A is compact. This arises because the operator B is compact by the Kolmogorov compactness criterion in $L^2(G)$ (cf. Muller [22], Chapter 2).

Consequently, there exists a complete orthonormal basis of H formed by eigenvectors $(v_m)_{m\geq 0}$ of A, associated to non-negative eigenvalues $(\lambda_m)_{m\geq 0}$, that we assume to be ordered in decreasing value. Each non-zero eigenvalue has a finite multiplicity, and 0 is the only possible accumulation point of the spectrum. If H is infinite-dimensional, then $\lim_{m\to\infty} \lambda_m = 0$.

Moreover, consider the correlation operator $C = BB^* : L^2(G) \mapsto L^2(G)$,

$$\mathrm{C}\varphi(\gamma) = \int_G (f(\gamma), f(\mu))_H\, \varphi(\mu)\, d\mu \quad \text{for } \varphi \in L^2(G).$$

Then the sequence $(\varphi_m)_{m\geq 0}$, with

$$\varphi_m(\gamma) = \frac{1}{\sigma_m}(Bv_m)(\gamma) = \frac{1}{\sigma_m}\,(f(\gamma), v_m)_H,\ \sigma_m = \sqrt{\lambda_m}\ (\text{it also holds } v_m = \frac{1}{\sigma_m}\, B^*\varphi_m) \tag{2}$$

is an orthogonal basis of $L^2(G)$. This yields the abstract Karhunen-Loève decomposition,

Corollary 0.1 *It holds*

$$f(\gamma) = \sum_{m\geq 0} \sigma_m\, \varphi_m(\gamma)\, v_m, \quad a.\ e.\ in\ G,$$

where the series is convergent in $L^2(G, H)$.

The main interest of the POD is the following best-approximation property (cf. [22], Chapter 2):

Lemma 0.2 *Let $V_l = Span(\varphi_1, \ldots, \varphi_l) \subset H$. Let W_l be any sub-space of H of dimension l. Then*

$$\int_G d_H(f(\gamma), V_l)^2\, d\gamma \leq \int_G d_H(f(\gamma), W_l)^2\, d\gamma,$$

where

$$d_H(v, W_l) = \inf_{w \in W_l} \|v - w\|_H \quad \textit{for } v \in H$$

denotes the distance from the element $\varphi \in H$ to the sub-space W_l.

Let us next consider the case of bivariate functions. Assume that $X \subset \mathbb{R}^d$ and $Y \subset \mathbb{R}^s$ are two bounded domains, d and s are integers ≥ 1. Let T be a given function in $L^2(X \times Y)$, that we want to approximate in a low-dimensional variety. Let us consider the integral operator with kernel T expressed as

$$v \mapsto B\,v, \;\; (B\,v)(y) = \int_X T(x, y)\, v(x)\, dx. \tag{3}$$

The operator B maps $L^2(X)$ into $L^2(Y)$, is bounded and has an adjoint operator B^* defined from $L^2(Y)$ into $L^2(X)$ as

$$\varphi \mapsto B^*\,\varphi, \;\; (B^*\,\varphi)(x) = \int_Y T(x, y)\, \varphi(y) dy. \tag{4}$$

We are thus in a particular case of the previous abstract setting, with $G = Y$, $H = L^2(X)$ and $f(\gamma)(x) = T(x, \gamma)$.

Recursive POD representation

The POD expansion may be recursively adapted to the representation of multi-parametric functions. Let us consider the case of trivariate functions to avoid unnecessary complexities. Consider a bounded domain $Z \subset \mathbb{R}^q$ for some integer number $q \geq 1$, and a trivariate function $T \in L^2(X \times Y \times Z)$. We identify T with a function of $L^2(Y \times Z, L^2(X))$ as both spaces are isometric. From Corollary 0.1 we deduce that there exist two orthonormal sets $(v_m)_{m\geq 0}$ and $(\varphi_m)_{m\geq 0}$ which are respectively complete in $L^2(X)$ and in $L^2(Y \times Z)$ such that T admits the representation

$$T(x, y, z) = \sum_{m\geq 0} \sigma_m \varphi_m(y, z)\, v_m(x), \tag{5}$$

where the sum is convergent in $L^2(Y \times Z, L^2(X))$. Moreover the singular values σ_m (that we assume to be ordered in decreasing value) are non-negative and converge to zero.

We next apply the POD expansion to each mode $\varphi_m(y, z)$. There exists two orthonormal sets $(u_k^{(m)})_{k\geq 1}$ and $(w_k^{(m)})_{k\geq 1}$ which are respectively complete in $L^2(Y)$ and $L^2(Z)$, such that φ_m admits the representation

$$\varphi_m(y, z) = \sum_{k\geq 0} \sigma_k^{(m)}\, u_k^{(m)}(y)\, w_k^{(m)}(z), \tag{6}$$

where the expansion is convergent in $L^2(Z, L^2(Y))$, which is isometric to $L^2(Y \times Z)$. Also, the singular values $(\sigma_k^{(m)})_{k\geq 0}$ are non-negative and decrease to zero. We then haves

Lemma 0.3 *The function $T \in L^2(X \times Y \times Z)$ admits the expansions*

$$T = \sum_{m\geq 0} \sum_{k\geq 0} \sigma_m\, \sigma_k^{(m)}\, v_m \otimes u_k^{(m)} \otimes w_k^{(m)} = \sum_{m\geq 0} \sigma_m\, v_m \otimes \left(\sum_{k\geq 0} \sigma_k^{(m)}\, u_k^{(m)} \otimes w_k^{(m)} \right), \tag{7}$$

where both sums are convergent in $L^2(X \times Y \times Z)$.

Proof It is enough to prove that any of both series is absolutely convergent. Consider the partial absolute sum for the first one

$$S_N = \sum_{0\le m\le N} \sum_{0\le k\le N} \|\sigma_m \sigma_k^{(m)} v_m \otimes u_k^{(m)} \otimes w_k^{(m)}\|^2_{L^2(X\times Y\times Z)}.$$

As the eigenfunctions are orthonormal (in their corresponding spaces),

$$S_N = \sum_{0\le m\le N} \sum_{0\le k\le N} |\sigma_m|^2 |\sigma_k^{(m)}|^2 \le \sum_{0\le m\le N} |\sigma_m|^2 \le \|T\|^2_{L^2(X\times Y\times Z)},$$

where the first inequality holds because $\|\varphi_m\|_{L^2(Y\times Z)} = 1$, and then $\sum_{k\ge 0} |\sigma_k^{(m)}|^2 = 1$. □

Feasible recursive POD representation

To build up a feasible recursive POD (R-POD) representation, consider a partial sum of the POD representation (7),

$$T_{P_M} = \sum_{0\le m\le M} \sigma_m v_m \otimes \left(\sum_{0\le k\le K_m} \sigma_k^{(m)} u_k^{(m)} \otimes w_k^{(m)} \right), \tag{8}$$

for some given integers $K_1 \ge 1, \ldots, K_M \ge 1$. The notation P_M is a short for the multi-index $(M, K_1, \ldots, K_M)$. We have

$$\begin{aligned} \|T - T_{P_M}\|^2_{L^2(X\times Y\times Z)} &= \sum_{m\ge M+1} |\sigma_m|^2 \left(\sum_{k\ge 0} |\sigma_k^{(m)}|^2 \right) + \sum_{0\le m\le M} |\sigma_m|^2 \left(\sum_{k\ge K_m+1} |\sigma_k^{(m)}|^2 \right) \\ &\le \sum_{m\ge M+1} |\sigma_m|^2 + \sum_{0\le m\le M} |\sigma_m|^2 \left(\sum_{k\ge K_m+1} |\sigma_k^{(m)}|^2 \right) \end{aligned} \tag{9}$$

This estimate suggests a practical strategy for the IP method to construct the expansion (8) within a targeted error:

Algorithm FR-POD (Feasible recursive POD representation)
Assume that some estimates of the remainders are computable:

$$\sum_{m\ge M+1} |\sigma_m|^2 \le |\alpha_M|^2, \quad \sum_{k\ge K_m+1} |\sigma_k^{(m)}|^2 \le |\beta_K^{(m)}|^2. \tag{10}$$

Set a tolerance $\varepsilon > 0$. Let

$$A = 1/\sqrt{2}, \quad B = 1/(\sqrt{2}\,\|T\|_{L^2(X\times Y\times Z)}). \tag{11}$$

- **Step 1:** Compute the modes φ_m and v_m and singular values σ_m for $m = 1, \ldots, M_\varepsilon$, until $\alpha_{M_\varepsilon} \le A\,\varepsilon$.
- **Step 2:** For each $m = 1, \ldots, M_\varepsilon$, compute the modes $u_m^{(k)}$ and $w_k^{(m)}$ and the singular values $\sigma_m^{(k)}$ for $k = 1, \ldots, K_m$, until $\beta_{K_m}^{(m)} \le B\,\varepsilon$.

For smooth functions the singular values decrease very fast, so that good estimators of the remainders are $\alpha_M = \sigma_{M+1}$, $\beta_K^{(m)} = \sigma_{K+1}^m$. For less smooth functions, some more summands of the series defining the remainders could be need. In "Analysis of solutions of the reaction-diffusion" section we shall obtain estimators α_M and $\beta_K^{(m)}$ when T is the solution of the reaction-diffusion, considered as a function depending on three parameters: The diffusivity, the reaction speed and the space-time variable. These estimators decrease exponentially with M.

Lemma 0.4 *Let T_ε the representation of T provided by Algorithm FR-POD within an error level ε. It holds*

$$\|T - T_\varepsilon\|_{L^2(X\times Y\times Z)} \le \varepsilon. \tag{12}$$

Proof From estimate (9) and (10),

$$\|T - T_{P_M}\|^2_{L^2(X\times Y\times Z)} \le |\alpha_{M_\varepsilon}|^2 + \sum_{0\le m\le M} |\sigma_m|^2\, |\beta^{(m)}_{K_m}|^2 \le (A^2 + \|T\|^2_{L^2(X\times Y\times Z)} B^2)\, \varepsilon^2 < \varepsilon^2,$$

where we have used that $\sum_{0\le m\le M} |\sigma_m|^2 \le \|T\|^2_{L^2(X\times Y\times Z)}$. □

In practice we recursively compute the expansion by the PI method [1]. This avoids to compute the full singular value decomposition of the correlation matrix, we just compute the modes needed to reach a given error threshold.

Quasi-optimality of recursive POD representation

The POD representation in general provides the most accurate representation in L^2 norm, for a given number of truncation modes. This is due to the best-approximation property stated in Lemma 0.2. Let us consider a trivariate approximation of T with M modes, of the form

$$\hat{T}_M(x, y, z) = \sum_{0\le m\le M} \hat{X}_m(x)\, \hat{Y}_m(y)\, \hat{Z}_m(z), \quad \text{for } (x, y, z) \in X \times Y \times Z. \tag{13}$$

Lemma 0.5 *Let $T \in L^2(X \times Y \times Z)$. It holds*

$$\|T - T_M\|_{L^2(X\times Y\times Z)} \le \|T - \hat{T}_M\|_{L^2(X\times Y\times Z)}, \tag{14}$$

where

$$T_M(x, y, z) = \sum_{0\le m\le M} \sigma_m \varphi_m(y, z)\, v_m(x), \tag{15}$$

and $\hat{T}_M$ is any trivariate approximation of T with M modes, of the form (13).

Proof Let V_M the space spanned by $v_1, \ldots, v_M$ in $L^2(Y \times Z)$. Observe that T_M is the orthogonal projection in $L^2(Y \times Z)$ of T on V_M. Let W_M be any sub-space of dimension M of $L^2(Y \times Z)$. Then, due to Lemma 0.2, it holds

$$\int_X \|(T - T_M)(x)\|^2_{L^2(Y\times Z)}\, dx \le \int_X \|(T - S_M)(x)\|^2_{L^2(Y\times Z)}\, dx$$

for any $S_M \in W_M$, where we denote $T(x)(y, z) = T(x, y, z)$, and similarly $T_M(x)$ and $S_M(x)$. As the spaces $L^2(X, L^2(Y \times Z))$ and $L^2(X \times Y \times Z)$ are isometric, taking $S_M = \hat{T}_M$, the inequality (14) follows. □

Note that in particular this implies that the POD expansion (15) is more accurate than the three-variate PGD one.

The following result states the quasi-optimality of the feasible R-POD with representations.

Lemma 0.6 *It holds*

$$\|T - T_\varepsilon\|_{L^2(X\times Y\times Z)} \le \|T - \hat{T}_M\|_{L^2(X\times Y\times Z)} + \varepsilon/\sqrt{2}, \tag{16}$$

for any trivariate approximation $\hat{T}_M$ of T with M modes, of the form (13).

Proof We have

$$\begin{aligned}\|T - T_\varepsilon\|_{L^2(X\times Y\times Z)} &\le \|T - T_M\|_{L^2(X\times Y\times Z)} + \|T_M - T_\varepsilon\|_{L^2(X\times Y\times Z)}\\ &\le \|T - T_M\|_{L^2(X\times Y\times Z)} + \left(\sum_{0\le m\le M} |\sigma_m|^2 \left(\sum_{k\ge K_m+1} |\sigma_k^{(m)}|^2\right)\right)^{1/2}\\ &\le \|T - T_M\|_{L^2(X\times Y\times Z)} + \varepsilon/\sqrt{2} \le \|T - \hat{T}_M\|_{L^2(X\times Y\times Z)} + \varepsilon/\sqrt{2},\end{aligned}$$

where the second-to-last estimate is obtained similarly to the proof of Lemma 0.4, and the last one follows from Lemma 0.5. □

Then, the feasible R-POD representation is more accurate than $\hat{T}_M$, for ε small enough, if the inequality in (16) is strict. If (16) is an equality, this means that $\hat{T}_M$ is optimal. In this case the accuracy of the feasible R-POD representation can be made arbitrarily close to the optimal one. It should be noted, however, that the R-POD contains more modes than $\hat{T}_M$. Anyhow, we present some numerical experiments in "Numerical tests" section that show than the R-POD representation is more accurate than the PGD one, for the same number of modes.

Analysis of solutions of the reaction-diffusion

Let us now consider the homogeneous Dirichlet boundary value problem of the linear reaction-diffusion equation,

$$\begin{cases}\partial_t T - \gamma\,\Delta T + \alpha\, T = f & \text{in } \mathcal{Q},\\ T = 0 & \text{in } (0,b)\times\partial\Omega,\\ T(x,0) = T_0(x) & \text{in } \Omega,\end{cases} \tag{17}$$

where $\gamma > 0$ and $\alpha \ge 0$ respectively denote the diffusivity and the reaction rate, and $\mathcal{Q} = \Omega \times (0,b)$. This problem fits into the functional framework of constant-coefficient linear parabolic equations, and admits a unique solution $T \in L^2((0,b), H^1(\Omega))$ such that $\partial_t T \in L^2(\mathcal{Q})$ if $f \in L^2(\mathcal{Q})$ and $T_0 \in L^2(\Omega)$. We shall assume that the pair (γ,α) ranges in a set $\mathcal{G} = [\gamma_m,\gamma_M]\times[\alpha_0,\alpha_M]$ with $0 < \gamma_m < \gamma_M$, $0 \le \alpha_0 \le \alpha_M$. Our purpose in this section is to analyze the rate of convergence of the approximation of T by a recursive POD expansion in separated tensor form:

$$T((x,t),(\gamma,\alpha)) \simeq T_P((x,t),(\gamma,\alpha)) = \sum_{m=0}^{M}\sum_{i=0}^{I} \tau_i^{(m)}\,\varphi_i^{(m)}(\gamma)\,w_i^{(m)}(\alpha)\,v_m(x,t), \tag{18}$$

where $P = (M,I)$, the $\tau_i^{(m)}$ are real numbers and $\varphi_i^{(m)} \in L^2(\gamma_m,\gamma_M)$, $w_i^{(m)} \in L^2(0,\alpha_M)$ and $v_m \in L^2(\mathcal{Q})$ are eigenmodes. To obtain this expression, let us start from the POD expansion of T where $\mu = (\gamma,\alpha) \in \mathcal{G}$ and $z = (x,t) \in \mathcal{Q}$,

$$T((x,t),(\gamma,\alpha)) = \sum_{m\ge 0} \sigma_m\,\varphi_m(\gamma,\alpha)\,v_m(x,t), \tag{19}$$

where the expansion converges in $L^2(\mathcal{G}\times\mathcal{Q})$. As $\varphi_m \in L^2(\mathcal{G})$, it also admits a POD expansion

$$\varphi_m(\gamma,\alpha) = \sum_{i\ge 0} \sigma_i^{(m)}\,u_i^{(m)}(\gamma)\,w_i^{(m)}(\alpha), \tag{20}$$

which is convergent in $L^2(\mathcal{G})$, where $\{u_i^{(m)}\}_{i\geq 0}$ is an orthonormal basis of $L^2(\gamma_m, \gamma_M)$ and $\{w_i^{(m)}\}_{i\geq 0}$ is an orthonormal basis of $L^2(0, \alpha_M)$. If we truncate the expansion (19) for T to $M+1$ summands and that (20) for φ_m at $I+1$ summands, then we recover the expression for T_P in (18) where $\tau_i^{(m)} = \sigma_m \sigma_i^{(m)}$.

To analyze the rate of convergence of T_P towards T, we need some technical tools. Let us consider the orthonormal Fourier basis $\{e_k\}$ of $L^2(\Omega)$ formed by the eigenfunctions of the Laplace operator. It holds

$$-\Delta e_k = \lambda_k e_k \quad \text{in } \Omega, \quad e_k = 0 \text{ in } \partial\Omega, \tag{21}$$

where $\lambda_k > 0$ is the eigenvalue associated to e_k. The sequence $\{\lambda_k\}_{k\geq 0}$ is ordered to be non-decreasing with $\lim_{k\to\infty} \lambda_k = 0$. We decompose T_0 and f as

$$T_0(x) = \sum_{k\geq 0} a_k e_k(x), \quad f(x,t) = \sum_{k\geq 0} f_k(t) e_k(x), \text{ with } a_k = (T_0, e_k)_{L^2(\Omega)}, \; f_k(t) = (f(\cdot,t), e_k)_{L^2(\Omega)},$$

where the series are respectively convergent in $L^2(\Omega)$ and $L^2(\mathcal{Q})$, and

$$\|T_0\|^2_{L^2(\Omega)} = \sum_{k\geq 0} |a_k|^2, \quad \|f\|^2_{L^2(\mathcal{Q})} = \sum_{k\geq 0} \|f_k\|^2_{L^2(0,b)}. \tag{22}$$

The solution of the reaction-diffusion equations is then expanded in terms of the eigenfunctions e_k,

$$T((x,t),(\gamma,\alpha)) = \sum_{k\geq 0} \theta_k(t,(\gamma,\alpha))\, e_k(x), \tag{23}$$

where the coefficients θ_k are defined by

$$\theta_k(t,(\gamma,\alpha)) = a_k\, e^{-(\gamma\lambda_k+\alpha)t} + \int_0^t f_k(s)\, e^{-(\gamma\lambda_k+\alpha)(t-s)}\, ds.$$

We shall consider T as a mapping from $\mathcal{G}$ into $L^2(\mathcal{Q})$ that brings a couple $(\gamma,\alpha) \in \mathcal{G}$ into the function $T((\cdot,\cdot),(\gamma,\alpha)) \in L^2(\mathcal{Q})$, that we denote $T_{(\gamma,\alpha)}$.

Our main result is the following.

Theorem 0.7 *The truncated POD series expansion T_P given by (18) satisfies the error estimate*

$$\|T - T_P\|_{L^2(\mathcal{G}\times\mathcal{Q})} \leq C_\rho\, (\rho^{-M} + \sqrt{M}\, \rho^{-I}), \tag{24}$$

for any $1 < \rho < \rho_*$, where $C_\rho > 0$ is a constant depending on ρ, unbounded as $\rho \to 1$, and $\rho_* = (\sqrt{\gamma_M} + \sqrt{\gamma_m})/(\sqrt{\gamma_M} - \sqrt{\gamma_m})$.

Therefore, the recursive POD expansion converges with spectral accuracy in terms of the number of truncation modes in the main and secondary expansions.

The proof of this result is essentially based upon the analyticity of T with respect to diffusivity γ and reaction rate α. It is rather technical, and will come up after several lemmas, the first of which is

Lemma 0.8 *The mapping $(\gamma,\alpha) \in \mathcal{G} \mapsto T_{(\gamma,\alpha)} \in L^2(\mathcal{Q})$ is analytic.*

Proof According to (23), T is the sum of two contributions, coming from the initial condition T_0 and the source f. We prove the analyticity for each of them.

i.—Let us consider the part generated by the initial condition, corresponding to

$$\theta_k(t,(\gamma,\alpha)) = a_k\, e^{-(\gamma\lambda_k+\alpha)t}.$$

Let us bound the residual

$$\sup_{\gamma\geq\epsilon,\alpha\geq 0}\left\|\sum_{k\geq L}\theta_k(t,(\gamma,\alpha))\,e_k(x)\right\|^2_{L^2(\mathcal{Q})} = \sup_{\gamma\geq\epsilon,\alpha\geq 0}\sum_{k\geq L}(a_k)^2\int_0^b e^{-2(\gamma\lambda_k+\alpha)t}\,dt$$
$$= \sup_{\gamma\geq\epsilon,\alpha\geq 0}\sum_{k\geq L}(a_k)^2\frac{1-e^{-2(\gamma\lambda_k+\alpha)b}}{2(\gamma\lambda_k+\alpha)} \leq \frac{1}{2\epsilon\lambda_0}\sum_{k\geq L}(a_k)^2,$$

for any $L > 0$. Then the series uniformly converges on each set $[\epsilon,+\infty[\times[0,+\infty[$, for all $\epsilon > 0$. As each term in the series (23) determines an analytic function from $\mathcal{G}$ into $L^2(\mathcal{Q})$, then the limit is analytic from $(0,+\infty)\times(0,+\infty)$ into $L^2(\mathcal{Q})$.

ii.—Let us now investigate the part arisen from the source f, corresponding to

$$\theta_k(t,(\gamma,\alpha)) = \int_0^t f_k(s)\,e^{-(\gamma\lambda_k+\alpha)(t-s)}\,ds. \tag{25}$$

This needs the preliminary statement.

Let $g\in L^2(0,b)$ and $\lambda > 0$, $\alpha\geq 0$ be given, the function

$$G:(\gamma,\alpha)\mapsto\int_0^t g(s)\,e^{-(\gamma\lambda+\alpha)(t-s)}\,ds,$$

mapping $]0,+\infty[\times]0,+\infty[$ into $L^2(0,b)$, *is analytic.*

To prove it, we show that $(\gamma,\alpha)\mapsto G(\gamma,\alpha)$ is locally expressed as a convergent power series. Let $\gamma_0 > 0$, $\alpha_0 > 0$ be fixed. On account of the analyticity of the exponential we derive that

$$G(\gamma,t) = \sum_{n\geq 0}\frac{[\lambda(\gamma-\gamma_0)+(\alpha-\alpha_0)]^n}{n!}\int_0^t g(s)[-(t-s)]^n e^{-(\gamma_0\lambda+\alpha_0)(t-s)}\,ds$$
$$:= \sum_{n\geq 0}\frac{[\lambda(\gamma-\gamma_0)+(\alpha-\alpha_0)]^n}{n!}G_n(t).$$

This series is absolutely convergent in $L^2(0,b)$. Indeed, the integral term being a convolution, then Young's inequality can be used which implies that

$$\sum_{n\geq 0}\frac{[\lambda(\gamma-\gamma_0)+(\alpha-\alpha_0)]^n}{n!}\|G_n\|_{L^2(0,b)}$$
$$\leq \sum_{n\geq 0}\frac{[\lambda(\gamma-\gamma_0)+(\alpha-\alpha_0)]^n}{n!}\|g\|_{L^2(0,b)}\|(-t)^n e^{-(\gamma_0\lambda+\alpha_0)t}\|_{L^1(0,\infty)}$$
$$= \|g\|_{L^2(0,b)}\sum_{n\geq 0}\frac{[\lambda(\gamma-\gamma_0)+(\alpha-\alpha_0)]^n}{(\lambda\gamma_0+\alpha_0)^n}.$$

The geometrical series is convergent for (λ,α) such that $|(\lambda\gamma+\alpha)-(\lambda\gamma_0+\alpha_0)| < \eta$ provided that $\eta < \lambda\gamma_0+\alpha_0$. Then, the function $G:]0,+\infty[\times]0,+\infty[\mapsto L^2(0,b)$ is analytic.

To finish the proof, let us check out that the series (23) with θ_k given by (25) is uniformly convergent in $[\epsilon,+\infty[\times[0,+\infty[$, for all $\epsilon > 0$. For a given L we have

$$\sup_{\gamma\geq\epsilon,\alpha\geq 0}\left\|\sum_{k\geq L}\theta_k(t,(\gamma,\alpha))\,e_k(x)\right\|^2_{L^2(\mathcal{Q})} = \sup_{\gamma\geq\epsilon,\alpha\geq 0}\sum_{k\geq L}\|\theta_k(t,(\gamma,\alpha))\|^2_{L^2(0,b)}$$
$$\leq \sup_{\gamma\geq\epsilon,\alpha\geq 0}\sum_{k\geq L}\|f_k\|^2_{L^2(0,b)}\|e^{-(\gamma\lambda_k+\alpha)t}\|^2_{L^1(0,\infty)}$$
$$\leq \frac{1}{(\epsilon\lambda_0)^2}\sum_{k\geq L}\|f_k\|^2_{L^2(0,b)}.$$

Then the series (23) of analytic functions is uniformly convergent. As a result, the limit is also analytic. The proof is complete. □

Another preliminary tool required in our study is related to the polynomial approximation of regular vector-valued functions. We shall adapt a result by S. Bernstein (in 1912), stated for complex-valued functions, and improved since then in many works (see for instance [19]). For some $\rho > 1$, let the set E_ρ in the complex plan be defined as

$$E_\rho = \left\{\zeta \in \mathbf{C}; \quad |\zeta - 1| + |\zeta + 1| \leq \rho + \rho^{-1}\right\}.$$

Consider a function $F : E_\rho \to H$ where H is a Hilbert space. For a given integer number $M \geq 0$ let F_M be the truncated Chebyshev polynomial series expansion of F of degree M with coefficients in H. The shape of the polynomial F_M will be fixed later on (see Remark 0.2). Following the proof as exposed in [19], we come up with

Lemma 0.9 *Assume that F is analytic and bounded in E_ρ. There holds that*

$$\max_{\xi \in [-1,1]} \|F(\xi) - F_M(\xi)\|_H \leq C_\rho\, \rho^{-M}.$$

Remark 0.1 The constant in the lemma may be fixed to (see [25, Theorem 8.2])

$$C_\rho = \frac{2}{\rho - 1} \|F\|_{L^\infty(E_\rho, H)},$$

that blows up as ρ goes to unity.

We now need to derive similar approximation estimates for analytic vector valued functions defined from $\mathcal{G}$ into $L^2(\mathcal{Q})$. The following result holds

Lemma 0.10 *For any $\alpha \in [0, \alpha_M]$ there exists a polynomial $S_M^{(\alpha)}$ ranging from $[\gamma_0, \gamma_M]$ into $L^2(\mathcal{Q})$, with degree $\leq M$, such that for all ρ $(1 < \rho < \rho_*)$,*

$$\max_{(\gamma,\alpha) \in \mathcal{G}} \|T(\gamma, \alpha) - S_M^{(\alpha)}(\gamma)\|_{L^2(\mathcal{Q})} \leq \hat{C}_\rho\, \rho^{-M},$$

where $\hat{C}_\rho$ is a non-negative constant, possibly unbounded as $\rho \to 1$.

Proof We only give a sketch of the proof. Following the result by Lemma 0.8, for any given $\alpha \geq 0$, the vector-valued function $\gamma \in \mathbf{C} \mapsto T(\gamma, \alpha)$ is analytic in $\mathcal{R}e\gamma > 0$. This implies that provided that $\rho < \rho_*$, the ellipse

$$\mathcal{E}_\rho = \left\{\zeta \in \mathbf{C}; \quad |\zeta - \gamma_M| + |\zeta - \gamma_m| \leq \frac{\gamma_M - \gamma_m}{2}(\rho + \rho^{-1})\right\}$$

is included in the analyticity set of T. Consider thus the coordinates transformation

$$\zeta = \tau(\hat{\zeta}) := \frac{\gamma_M - \gamma_m}{2}\hat{\zeta} + \frac{\gamma_M + \gamma_m}{2}, \qquad \hat{\zeta} \in E_\rho.$$

It is affine and bijective from E_ρ into $\mathcal{E}_\rho$ and transforms the reference interval $[-1, 1]$ into $G = [\gamma_m, \gamma_M]$. This transformation makes it possible to construct such a polynomial $S_M^{(\alpha)}$. In fact, we start by constructing the truncated Chebyshev series expansion $\hat{S}_M^{(\alpha)}(\hat{\zeta})$ of the (transformed) function $\hat{T}^{(\alpha)}(\hat{\zeta}) = T(\zeta, \alpha)$. Then, back to the interval $[\gamma_m, \gamma_M]$, we set $S_M^{(\alpha)}(\zeta) = \hat{S}_M^{(\alpha)}(\hat{\zeta})$.

To obtain the error estimate, from Lemma 0.9 we obtain

$$\max_{\gamma \in G} \|T(\gamma, \alpha) - S_M^{(\alpha)}(\gamma)\|_{L^2(\mathcal{Q})} \leq \frac{2}{\rho - 1} \|T(\cdot, \alpha)\|_{L^\infty(\mathcal{E}_\rho, L^2(\mathcal{Q}))}\, \rho^{-M} \leq \frac{2K}{\rho - 1}\, \rho^{-M},$$

where

$$K = \sup_{\alpha\in[0,\alpha_M]} \|T(\cdot,\alpha)\|_{L^\infty(\mathcal{E}_\rho,L^2(\mathcal{Q}))},$$

which is finite and independent of $\alpha \in [0, \alpha_M]$ due to the uniform boundedness of $T(\gamma, \alpha)$ in compact sets of $(0, +\infty) \times (0, +\infty)$. The proof is complete. □

Remark 0.2 The polynomial $S_M^{(\alpha)}$ may be put under the form

$$S_M^{(\alpha)}(\gamma) = \sum_{0\le m\le M} w_m^{(\alpha)}\, U_m(\gamma), \qquad \forall\gamma \in G,$$

where U_m stands for the polynomial obtained by transporting the Chebyshev polynomial of degree m defined in $[-1, 1]$ to the interval G, and the coefficients $(w_m^{(\alpha)})_{0\le m\le M}$ belong to $L^2(\mathcal{Q})$.

Proof of Theorem 0.7 Let us consider the truncated primary expansion

$$T_M((x,t),(\gamma,\alpha)) = \sum_{m=0}^{M} \sigma_m\, \varphi_m(\gamma,\alpha)\, v_m(x,t),$$

for some integer $M \ge 0$. Let S_M be the vector-valued polynomial (considered as a function of (γ, α)) constructed in Lemma 0.10. In view of Lemma 0.2 and Remark 0.2, the following identity holds,

$$\|T - T_M\|_{L^2(\mathcal{G}\times\mathcal{Q})} \le \|T - S_M\|_{L^2(\mathcal{G}\times\mathcal{Q})} \le |\mathcal{G}|^{1/2} \max_{(\gamma,\alpha)\in\mathcal{G}} \|T(\gamma,\alpha) - S_M^{(\alpha)}(\gamma)\|_{L^2(\mathcal{Q})}.$$

Applying the result stated in Lemma 0.10 it follows that

$$\|T - T_M\|_{L^2(\mathcal{G}\times\mathcal{Q})} \le \hat{C}_\rho\, \rho^{-M}. \tag{26}$$

Next, observe that as the sequence $(v_m)_{m\ge0}$ is orthonormal in $L^2(\mathcal{Q})$, then

$$\|T_M - T_P\|^2_{L^2(\mathcal{G}\times\mathcal{Q})} \le \sum_{m=0}^{M} \sigma_m^2\, \|\varphi_m - \varphi_m^{(I)}\|^2_{L^2(\mathcal{G})}, \tag{27}$$

where

$$\varphi_m^{(I)}(\gamma,\alpha) = \sum_{i=1}^{I} \sigma_i^{(m)}\, u_i^{(m)}(\gamma)\, w_i^{(m)}(\alpha)$$

is the truncated POD expansion of φ_m to $I + 1$ terms. Also, by (2),

$$\varphi_m(\gamma,\alpha) = \frac{1}{\sigma_m} \int_{\mathcal{Q}} T((x,t),(\gamma,\alpha))\, v_m(x,t)\, dx\, dt. \tag{28}$$

Then φ_m is an analytic function from $(0, +\infty) \times (0, +\infty)$ into $\mathbb{R}$. By an argument similar to that of Lemma 0.10, we prove that for any $\alpha \in [0, \alpha_1]$ there exists a polynomial in γ, $r_{I,\alpha}^{(m)}(\gamma)$, of degree least or equal than I such that

$$\max_{\gamma\in[\gamma_m,\gamma_M]} |\varphi_m(\gamma,\alpha) - r_{I,\alpha}^{(m)}(\gamma)| \le \frac{2}{1-\rho}\, \|\varphi_m(\cdot,\alpha)\|_{L^\infty(\mathcal{E}_\rho)}\, \rho^{-I}. \tag{29}$$

From (28), we deduce $\sigma_m\, |\varphi_m(\gamma,\alpha)| \le \|T(\gamma,\alpha)\|_{L^2(\mathcal{Q})}$ for all $(\gamma,\alpha) \in \mathcal{G}$, and then

$$\sigma_m \max_{(\gamma,\alpha)\in\mathcal{G}} |\varphi_m(\gamma,\alpha) - r_{I,\alpha}^{(m)}(\gamma)| \le \frac{K}{1-\rho}\, \rho^{-I},$$

for some constant $K \geq 0$ independent of m and α. Consequently, in view of Lemma 0.2,

$$\sigma_m^2 \, \|\varphi_m - \varphi_m^{(I)}\|^2_{L^2(\mathcal{G})} \leq \sigma_m^2 \, \|\varphi_m - r_I^{(m)}\|^2_{L^2(\mathcal{G})} \leq |\mathcal{G}| \, \sigma_m^2 \, \|\varphi_m - r_I^{(m)}\|^2_{L^\infty(\mathcal{G})} \leq \overline{C}_\rho^2 \, \rho^{-2I}, \tag{30}$$

where $r_I^{(m)}(\gamma, \alpha) = r_{I,\alpha}^{(m)}(\gamma)$ and $\overline{C}_\rho = |\mathcal{G}|^{1/2} \dfrac{K}{1-\rho}$. From (27) we deduce that

$$\|T_M - T_P\|_{L^2(\mathcal{G}\times\mathcal{Q})} \leq \overline{C}_\rho \sqrt{M} \, \rho^{-I}.$$

Combining this estimate with (26) completes the proof. □

Remark 0.3

- The constant C_ρ in estimate (24) also depends on the parameters domain $\mathcal{G}$. We do not make explicit this dependence to simplify the notation.
- The limit value for the convergence rates ρ_* only depends on the ratio γ_M/γ_m, as

$$\rho_* = \frac{2}{\sqrt{\frac{\gamma_M}{\gamma_m}} - 1} + 1.$$

- In view of estimate (24), in general a quasi-optimal choice for I is $I = M + \dfrac{1}{2}\log M$ (actually, the closest integer to this number). In this case,

$$\|T - T_P\|_{L^2(\mathcal{G}\times\mathcal{Q})} \leq C_\rho \, \rho^{-M}.$$

  We thus obtain the same asymptotic convergence order when $M \to \infty$ as for $\|T - T_M\|_{L^2(\mathcal{G}\times\mathcal{Q})}$.
- For more general parameter-depending parabolic equations, the above technique applies if the elliptic operator is symmetric. This allows to diagonalize the problem and expand the solution as a series in terms of the eigenfunctions of the elliptic operator. The use of Courant–Fischer–Weyl Theorem [19] allows to reduce the estimate of the truncation error of the POD expansion to the estimate of the interpolation error of the solution with respect to one of the parameters, eventually by polynomial functions. Then the convergence rate of the POD expansion will depend on the smoothness of the solution with respect to the parameters of the problem.

Reordering of recursive POD expansion

A practical way to re-order expansion (18) is in decreasing order of the values $\tau_m^{(i)} = \sigma_m \, \sigma_m^{(i)}$. This leads to an expansion of the form

$$T_P((x,t),(\gamma,\alpha)) = \sum_{l=0}^{L} \tilde{\tau}_l \, \tilde{\varphi}_l(\gamma) \, \tilde{w}_l(\alpha) \, \tilde{v}_l(x,t), \quad L = (M+1)(I+1), \tag{31}$$

where the sets $\{\tau_m^{(i)}, \, m = 0, \ldots, M, \, i = 1, \ldots, I\}$ and $\{\tilde{\tau}_l, \, l = 0, \ldots, L\}$ coincide, and $\tilde{\tau}_0 \geq \tilde{\tau}_1 \geq \cdots \geq \tilde{\tau}_L$.

To analyze the rate of convergence of this rearrangement of the RPOD expansion, let us at first remark that Theorem 0.7 allows, as a by-product, to estimate the singular values σ_m and $\sigma_m^{(i)}$. Indeed, denote by $\mathcal{L}(E, F)$ the set of linear bounded mappings from a Banach space E into a Banach space F, the following bound holds,

$$\sigma_{M+1} = \min_{B_M \in \mathcal{L}(L^2(\mathcal{G}), L^2(\mathcal{Q})), \operatorname{rank} B_M \leq M} \|B - B_M\|_{\mathcal{L}(L^2(\mathcal{G}), L^2(\mathcal{Q}))}, \tag{32}$$

where

$$(B\varphi)(z) = \int_G T(\gamma, z)\varphi(\gamma) \, d\gamma. \, \forall z \in \mathcal{Q}.$$

Consider the operator

$$(\tilde{B}_M\varphi)(z) = \int_G T_M(\gamma, z)\varphi(\gamma)\, d\gamma,\ \forall z \in \mathcal{Q}. \tag{33}$$

Then by estimate (26)

$$\sigma_{M+1} \le \|B - \tilde{B}_M\|_{\mathcal{L}} = \|T - T_M\|_{L^2(G\times\mathcal{Q})} \le \hat{C}_\rho\, \rho^{-M}. \tag{34}$$

Similarly,

$$\sigma_{i+1}^{(m)} \le \min_{E_I^{(m)}\in\mathcal{L},\mathrm{rank}E_I^{(m)}\le I} \|E^{(m)} - E_I^{(m)}\|_{\mathcal{L}(L^2(G),L^2([0,\alpha_1])},$$

where $(E^{(m)}u)(\alpha) = \int_G \varphi^{(m)}(\gamma,\alpha)\, u(\gamma)\ d\gamma, \quad \forall\alpha \in [0,\alpha_1]$. Let us assume that the $\varphi^{(m)}$ satisfy the additional (slightly) stronger boundedness property

$$\sup_{\alpha\in[0,\alpha_1],\, m=0,1,\ldots} \|\varphi_m(\cdot,\alpha)\|_{L^\infty(\mathcal{E}_\rho)} < +\infty. \tag{35}$$

Then, in view of estimate (29),

$$\sigma_{I+1}^{(m)} \le \|E^{(m)} - \tilde{E}_I^{(m)}\|_{\mathcal{L}(L^2(G),L^2([0,\alpha_1])} = \|\varphi^{(m)} - r_I^{(m)}\|_{L^2(\mathcal{G})} \le \hat{C}_\rho^{(m)}\, \rho^{-I}, \tag{36}$$

where $(\tilde{E}_I^{(m)}u)(\alpha) = \int_G r_I^{(m)}(\gamma,\alpha)\, u(\gamma)\, d\gamma, \quad \forall\alpha \in [0,\alpha_1]$.

Then the error associated to this reordering, for large M and I, is estimated by

$$\|T - T_P\|_{L^2(\mathcal{G}\times\mathcal{Q})} \le D_\rho\, L^{1/4}\, \rho^{-2\sqrt{L}} \tag{37}$$

for any $1 < \rho < \rho_*$, where D_ρ is a constant, possibly unbounded as $\rho \to 1$. To justify it, let us write T_P as

$$T_P((x,t),(\gamma,\alpha)) = \sum_{k=0}^{K}\sum_{m+i=k} \tau_i^{(m)}\, \varphi_i^{(m)}(\gamma)\, w_i^{(m)}(\alpha)\, v_m(x,t),$$

where for simplicity we assume that $M$and I are such that $L = (K+1)(K+2)/2$ for some integer $K \ge 0$. For other values there will appear a residual corresponding to high order modes that will be asymptotically negligible, as it is of larger order with respect to ρ. If estimates (34) and (36) are sharp, it holds

$$\tau_i^{(m)} \simeq A_\rho\, \rho^{-(m+i)} \tag{38}$$

for some constant A_ρ. Then, $\tau_i^{(m)} < \tau_j^{(n)}$ if $i+m > j+n$ and consequently the set $\{\tilde{\tau}_l,\ (k(k+1)/2) - 1 \le l \le (k+1)(k+2)/2\}$ coincides with the set $\{\tau_i^{(m)},\ m+i = k\}$. Then, due to estimate (38),

$$\|T - T_P\|^2_{L^2(\mathcal{G}\times\mathcal{Q})} \le \sum_{k\ge K+1}\sum_{m+i=k} |\tau_i^{(m)}|^2 \le A_\rho \sum_{k\ge K+1} (k+1)\, \rho^{-2k}. \tag{39}$$

As $\sum_{k\ge K+1} (k+1)\, \rho^{-2k} \simeq (K+1)\, \rho^{-2(K+1)}$, $K \simeq \sqrt{2L}$, then (37) follows.

Practical implementation

Assume again that estimate (36) is sharp. Then $\sum_{i\ge I+1} |\sigma_i^{(m)}|^2 \simeq C'_m\, \rho^{-2I} \simeq |\sigma_{I+1}^{(m)}|^2$. Thus, we may set the estimator $\beta_I^{(m)} = \sigma_{I+1}^{(m)}$, and similarly $\alpha_M = \sigma_{M+1}$, in (10). This suggests to

consider a different number of summands in the secondary expansions of (18), what leads to an expansion as in (8),

$$T_P((x,t),(\gamma,\alpha)) = \sum_{m=0}^{M}\sum_{i=0}^{I_m} \tau_i^{(m)}\,\varphi_i^{(m)}(\gamma)\,w_i^{(m)}(\alpha)\,v_m(x,t), \tag{40}$$

where M and I_m are determined to fit the error tolerance tests $\sigma_M \le A\,\varepsilon$ and $\sigma_{I_m}^{(m)} \le B\,\varepsilon$ where A and B are given in (11). In practice for simplicity these may be replaced by $\sigma_{M+1} \le \varepsilon$ and $\sigma_{I_m+1}^{(m)} \le \varepsilon$.

Also, in view of (38) and (39), we deduce that a good estimator the error $\|T - T_P\|_{L^2(\mathcal{G}\times\mathcal{Q})}$ is $\tau_I^{(M)}$, associated to the last computed mode, such that $I + M = K$.

Numerical tests

This section is devoted to the comparison of the practical performances of the feasible R-POD expansion. In particular, we confirm the exponential rate of convergence of the truncated POD expansion for the diffusion-reaction equation proved in section "Analysis of solutions of the reaction-diffusion". We are also interested in comparing the rate of convergence of R-POD and PGD expansions, as the latter is particularly well suited to approximate multivariate functions. We have considered functions with high and low smoothness, as the smoothness plays a crucial role in the decreasing of the size of the modes in both expansions. In addition we have tested the ability of both representations to approximate functions that already have a separated tensor structure. For completeness we describe in the Appendix the application of the PGD expansion to approximate multivariate functions.

Multi-variate functions

In this test we apply the R-POD and the PGD to approximate multivariate functions. Actually we consider tri-variate functions a generic test to determine the relative performances of both expansions. We have considered the following tests:

Case 1: Function with tensor structure.

$$S_1(x,y,z) = x + y + z. \tag{41}$$

Case 2: Function with non tensor structure.

$$S_2(x,y,z) = \sin(xyz). \tag{42}$$

Case 3: Function with low regularity

$$S_3(x,y,z) = \sqrt{x + 2y + z + 4}. \tag{43}$$

The space domain is fixed to $\Omega = X \times Y \times Z$, with $X = Y = Z =]-1,1[$ and Gauss–Lobatto–Legendre quadrature is used (see [4]) with the polynomial degree equal to $N = 64$. These formulas are used to evaluate the matrix representation of the operators B and A.

We set the tolerance error in $L^2(X \times Y \times Z)$ in the residual of both RPOD and PGD expansions to $\mu = 10^{-7}$. This corresponds to $\varepsilon = 10^{-14}$ in Algorithm FR-RPOD. We have displayed in Figs. 1, 2, 3 the comparison of the convergence history of the feasible R-POD and PGD processes, for all the three-variate functions considered. The x-axis represents the number of eigen-modes while the y-axis represents the $L^2(X \times Y \times Z)$ error, in

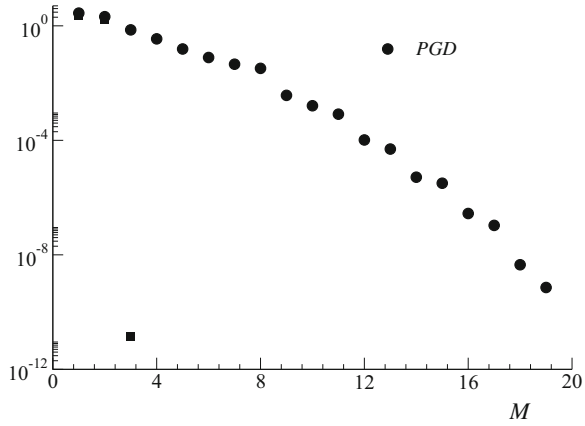

Fig. 1 Comparison of errors for feasible R-POD and PGD. Function $S_1(x, y, z) = x + y + z$

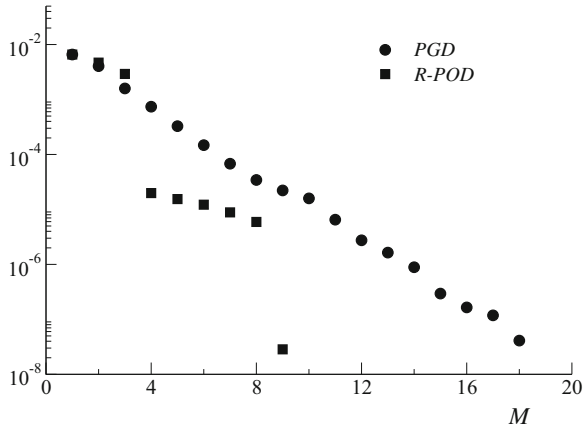

Fig. 2 Comparison of errors for feasible R-POD and PGD. Function $S_2(x, y, z) = \sin(xyz)$

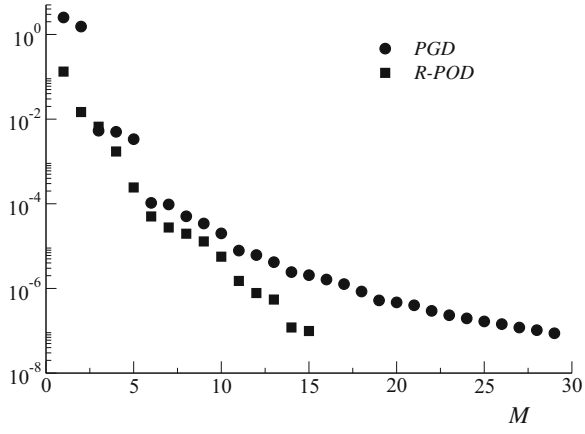

Fig. 3 Comparison of errors for feasible R-POD and PGD. Function $S_3(x, y, z) = \sqrt{x + 2y + z + 4}$

logarithmic coordinates. We observe in Fig. 1 that the R-POD just needs 3 modes to fit a function that already has a separated tensor structure, while the PGD requires 17 modes to reach the error level. Further, that for functions with low smoothness both expansions

require approximately the same number of modes to reach a moderate accuracy, however the R-POD is more efficient to reach high accuracy in all cases. Finally, that the error associated to the R-POD expansions is almost in all cases below the one associated to the PGD one for the same number of modes.

Table 1 displays the number of modes required by each expansion to reach the error below the threshold $\mu = 10^{-7}$. We observe that both expansions appear to converge for all cases considered, although in general both require a larger number of modes to approximate functions with lower smoothness. Also, that in all cases considered the R-POD requires less modes than the PGD.

Reaction-Diffusion equation

This part is devoted to determining the effective convergence rate of the R-POD approximation of some solutions to the transient reaction-diffusion equation when parameterized by the diffusivity and reaction coefficients. We assess the exponential convergence rate and investigate the variation of this rate with respect to the set $\mathcal{G} = [\gamma_m, \gamma_M] \times [\alpha_m, \alpha_M]$.

Test 1: Exponential convergence rate.

We consider the time-dependent reaction-diffusion equation in the domain $\mathcal{Q} = (0, 1) \times (0, 1)$ and we select three possible pairs of source terms and initial conditions, given by

$$\begin{aligned}
&\textbf{Data 1}: && f(t,x) = \sqrt{|x-t-0.3|}, && T_0(x) = 0,\\
&\textbf{Data 2}: && f(t,x) = 0, && T_0(x) = |x-0.4|,\\
&\textbf{Data 3}: && f(t,x) = \sqrt{|x-t-0.3|}, && T_0(x) = |x-0.4|.
\end{aligned}$$

These data have mild singularities, so the temperature solutions of (17), have a reduced regularity with respect to x and t, in particular for $t = 0$ for the two last data. The heat problem is discretized by an Euler scheme/Gauss–Lobatto–Legendre spectral method see [4] (the time step is $\delta t = 10^{-2}$ and the polynomial degree is $N = 64$).

Calculation for the matrix representations of the operators B and A are realized by means of accurate quadrature formulas. Indeed, various integrals (with respect to either γ, α or (t, x)) are computed using Gauss-Lobatto quadrature formulas with high resolution in the corresponding intervals.

Figure 4 shows the convergence history of the R-POD expansion for the reaction-diffusion equation (40) in terms of the total number of modes in the expansion. We have considered the sets of diffusivities $\gamma \in [1, 51]$, and reaction rates $\alpha \in [0, 100]$. The error is measured in $L^2(\mathcal{Q})$ norm. The numbers of secondary modes I_m has been determined to fit the test $\sigma_m^{(I_m+1)} \leq \varepsilon = 10^{-10}$. In practice a small amount of secondary modes (actually, $I_m \simeq 4$) is needed to fit this test. The modes have been re-arranged in decreasing order of the effective singular values $\tau_i^{(m)} = \sigma_m \sigma_i^{(m)}$ (denoted by a ● symbol). We observe that the $\tau_i^{(m)}$ indeed are good error estimators for this re-arranged expansion, as was argued in "Practical implementation" section.

Table 1 Comparison of feasible R-POD and PGD for trivariate functions

	Modes R-POD	Modes PGD
Case 1	3	17
Case 2	9	17
Case 3	21	43

To assess the regularity of the eigenmodes associated to the conductivity parameter γ we choose to plot the three first corresponding to the most important singular values. The computational is made for case of Data 3. Based on Fig. 5 we clearly observe that these functions are regular. Same observation is made for the reaction parameter

Test 2: Dependence of the convergence rate with respect to the parameters range.

The dependence with respect to the ratio of diffusivities $R = \gamma_M/\gamma_m$ of the exponential convergence rate, stated by Theorem 0.7, is illustrated in Fig. 6. We depict the convergence

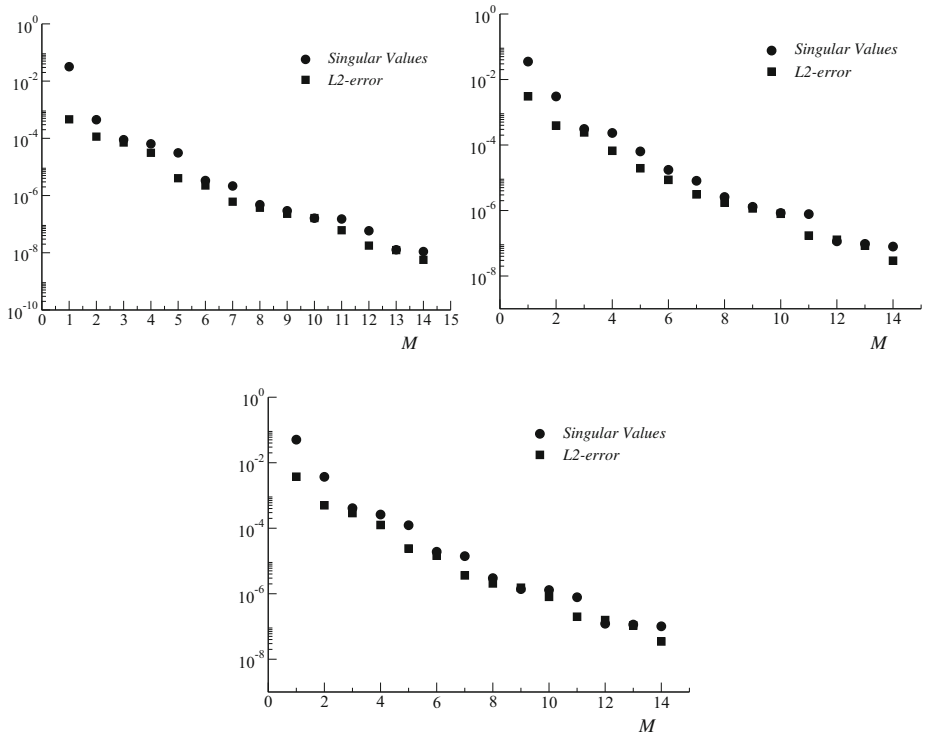

Fig. 4 Convergence history for POD expansion of the solution of the reaction-diffusion equation. Data 1 (*top left*), Data 2 (*top right*) and Data 3 (*bottom*)

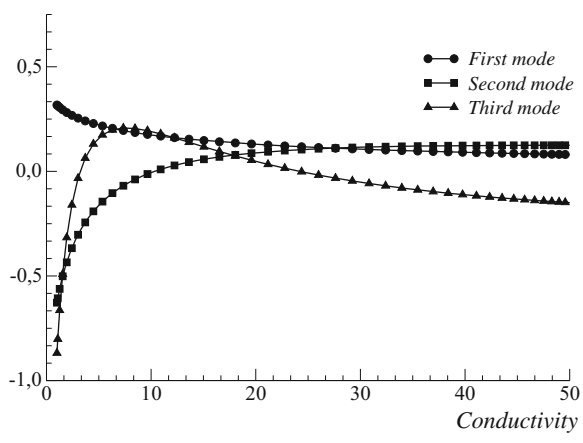

Fig. 5 Three first eigenmodes associated to the conductivity parameter (γ)

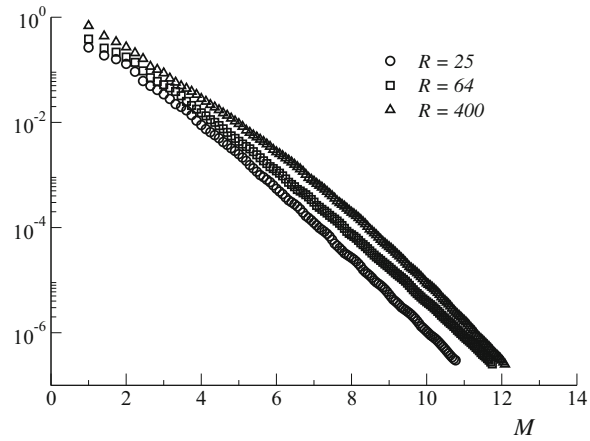

Fig. 6 Variation of the R-POD errors (in logarithmic scale) with respect to the ratio $R = \gamma_M/\gamma_m$, for fixed $\alpha_m = 0, \alpha_M = 100$. The variable M stands for the square root of the number of modes

history for Data 3, computed for $R = 25, 64$ and 400, in all cases with a fixed interval of reaction rates $[\alpha_m, \alpha_M] = [0, 100]$, with respect to the square root of the numbers of modes, $M = \sqrt{L}$. We can point out that the convergence rate degrades as R increases, in accordance with the fact that

$$\rho_* = \frac{2}{\sqrt{\frac{\gamma_M}{\gamma_m} - 1}} + 1.$$

We observe some gap between the purely exponential decay of the error and the computed one, as the error curve in logarithmic coordinates appears to be a slightly concave curve instead of a straight line. This is consistent with the presence of the factor $L^{1/4}$ in estimate (37).

In Table 2, we present the computed exponential convergence rate $\alpha_c = 2 \log \rho_c$, so that the $L^2(\mathcal{G} \times \mathcal{Q})$ error, in terms of the number of modes after rearranging the RPOD series, is assumed to satisfy

$$e(L) = C\, e^{-\alpha_c \sqrt{L}},$$

and the theoretical one given by $\alpha^* = 2 \log \rho_*$. The value α_c is calculated by exponential regression. We indeed recover an exponential rate of convergence with respect to the square root of the number of modes, with an effective convergence rate larger than the theoretical one. We numerically state that the computed rate in all cases is larger than one (see Table 2). We thus observe a kind of super-convergence effect.

We next test the dependence of the convergence rate with respect to the interval of reaction rates $[\alpha_0, \alpha_M]$. We show in Fig. 7 the convergence rates history corresponding

Table 2 Computed and theoretical convergence rates, for different values of $R = \gamma_M/\gamma_m$ and fixed $\alpha_m = 0, \alpha_M = 100$ (for Data 3)

$R = \gamma_M/\gamma_m$	α_c	α_*
25	1.48	0.81
36	1.44	0.67
64	1.38	0.50
100	1.36	0.40
400	1.35	0.20

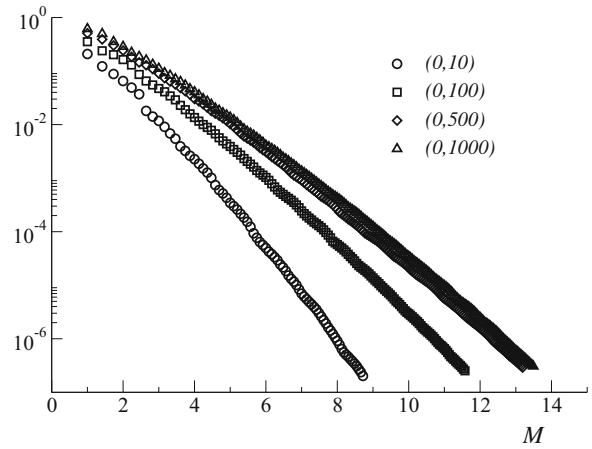

Fig. 7 Variation of the R-POD errors with respect to the reaction rate $\boldsymbol{\alpha}$. The *curves* correspond to $\boldsymbol{\alpha}_0 = 0$, $\boldsymbol{\alpha}_M = 10, 100, 500, 1000$, with $\gamma_m = 1$, $\gamma_M = 51$ in all cases. The variable M stands for the square root of the number of modes

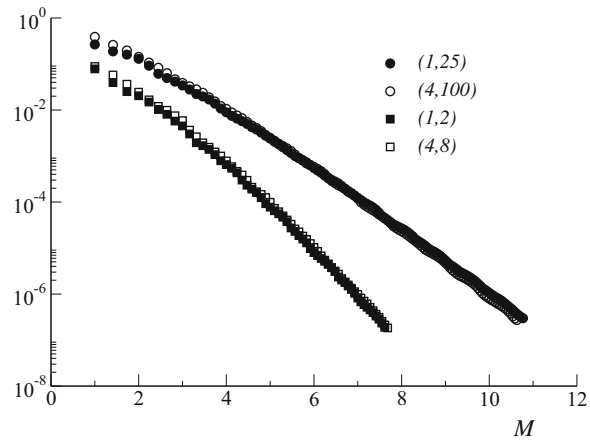

Fig. 8 Analysis of dependence of the R-POD errors with respect to the ratio $R = \gamma_M/\gamma_m$. The *curves* correspond to the indicated pairs (γ_m, γ_M). The variable M stands for the square root of the number of modes

to $\alpha_m = 0$, $\alpha_M = 10, 100, 500, 1000$ for fixed $\gamma_m = 1$, $\gamma_M = 51$. We observe a decrease of the rate as α_M increases, that however appears to be uniformly bounded, in agreement with estimate (37), where the dependence of the error bound with respect to $[\alpha_0, \alpha_M]$ only appears through the coefficient D_ρ.

The last numerical experiment studies wether the dependence of the exponential convergence on the diffusivities range $[\gamma_m, \gamma_M]$ indeed takes place in terms of the ratio $R = \gamma_M/\gamma_m$. This is confirmed by the result plot in Fig. 8, where we consider the couples $(\gamma_m, \gamma_M) = (1, 2)$ and $(4, 8)$, corresponding to $R = 2$, and $(\gamma_m, \gamma_M) = (1, 25)$ and $(4, 100)$, corresponding to $R = 25$, with fixed $\alpha_0 = 0$, $\alpha_M = 100$.

Conclusion

We have introduced in this paper a recursive POD (RPOD) expansion to approximate multivariate functions. The approach consists in building truncated recursive POD expan-

sions of the modes that appear in the expansions at the previous level, to a given tolerance error. We have constructed a practical truncation error estimator by means of bounds for the singular values, which is used to recursively compute the expansion by the Power Iterate (PI) method. This allows to compute just the modes needed to attempt a given error threshold. We have proved the quasi-optimality of this RPOD expansion in L^2, similar to that of the POD expansion.

We have proved the exponential rate of convergence of the RPOD expansion for the solution of the reaction-diffusion equation, based upon the analyticity of its solution with respect to those parameters.

We have finally performed some relevant numerical tests that on one hand show that the RPOD is more accurate than the PGD expansion for three-variate functions, and that on another hand confirm the exponential rate of convergence for the solution of the reaction-diffusion equation, presenting a good agreement with the qualitative and quantitative theoretical expectations.

Further extensive tests for more complex multivariate functions, in particular of practical interest for engineering applications, are in progress and will appear in a forthcoming paper.

Authors' contributions
MA, FBB and TCR participated to the development of the mathematical proves and the numerical investigations. They checked the results and wrote the manuscript. All authors read and approved the final manuscript.

Author details
[1]Bordeaux INP, I2M, (UMR CNRS 5295), 33607 Pessac, France , [2]LMAC, EA 2222, Université de Technologie de Compiègne, BP 20529, 60205 Compiègne Cedex, France, [3]I2M, IPB (UMR CNRS 5295), Université de Bordeaux, 33607 Pessac, France, [4]Departamento EDAN & IMUS, Universidad de Sevilla, C/Tarfia, s/n, 41012 Sevilla, Spain.

Acknowledgements
None.

Competing interests
The authors declare that they have no competing interests.

Appendix: The PGD representation of multivariate functions

We describe in this section the procedure to calculate the PGD representation of a multivariate functions. We focus on trivariate functions for the sake of clarity. Its extension to general multivariate functions is straightforward.

The PGD approximation of a trivariate function T searches for an expansion of the form

$$T(x, y, z) = \sum_{m\geq 0} X_m(x)\, Y_m(y)\, Z_m(z), \text{ for } (x, y, z) \in X \times Y \times Z. \tag{44}$$

The leading term $X_0 \otimes Y_0 \otimes Z_0$ is initially computed by means of an adaptation of the Power Iteration algorithm: Assume known an approximation $X_0^{(n-1)} \otimes Y_0^{(n-1)} \otimes Z_0^{(n-1)}$.

Step 1. Find $Z_0^{(n)} \in L^2(Z)$ such that for all $Z^* \in L^2(Z)$,

$$\left(X_0^{(n-1)} \otimes Y_0^{(n-1)} \otimes Z_0^{(n)} - T,\ X_0^{(n-1)} \otimes Y_0^{(n-1)} \otimes Z^* \right)_{L^2(X\times Y\times Z)} = 0. \tag{45}$$

Step 2. Find $\tilde{X}_0^{(n)} \in L^2(X)$ such that for all $X^* \in L^2(X)$,

$$\left(\tilde{X}_0^{(n)} \otimes Y_0^{(n-1)} \otimes Z_0^{(n)} - T,\ X^* \otimes Y_0^{(n-1)} \otimes Z_0^{(n)} \right)_{L^2(X\times Y\times Z)} = 0. \tag{46}$$

Set

$$X_0^{(n)} = \frac{\tilde{X}_0^{(n)}}{\|\tilde{X}_0^{(n)}\|_{L^2(X)}}.$$

Step 3. Find $\tilde{Y}_0^{(n)} \in L^2(Y)$ such that for all $Y^* \in L^2(Y)$,

$$\left(X_0^{(n)} \otimes \tilde{Y}_0^{(n)} \otimes Z_0^{(n)} - T,\ X_0^{(n)} \otimes Y^* \otimes Z_0^{(n)} \right)_{L^2(X \times Y \times Z)} = 0. \tag{47}$$

Set

$$Y_0^{(n)} = \frac{\tilde{Y}_0^{(n)}}{\|\tilde{Y}_0^{(n)}\|_{L^2(Y)}}.$$

The procedure is to be iterated until the error eventually is below a given tolerance.

The Mth mode $X_M \otimes Y_M \otimes Z_M$ is computed in the same way, by replacing the function T by the residual $T - \hat{T}_{M-1}$, where now

$$\hat{T}_{M-1}(x, y, z) = \sum_{0 \le m \le M-1} X_m(x)\, Y_m(y)\, Z_m(z), \text{ for } (x, y, z) \in X \times Y \times Z. \tag{48}$$

In this way, the residual $T - \hat{T}_M$ is orthogonal to $Span(X_M \otimes Y_M \otimes Z_M)$.

There is no proof, up to the knowledge of the authors, that the PGD expansion (44) exists for functions $T \in L^2(X \times Y \times Z)$ or perhaps with additional regularity, nor that the alternate Power Iteration process (45)–(47) converges. There is a proof, however, that for general functions depending on three or more parameters, there does not exist optimal sub-spaces of finite dimension 3 or larger, that satisfy the optimal approximation property set by Theorem 0.2 (see [9]).

Received: 24 November 2015 Accepted: 11 February 2016
Published online: 09 March 2016

References

1. Azaïez M, Belgacem Ben F. Karhunen-Loève's truncation error for bivariate functions. Comput Methods Appl Mech Eng. 2015;290:57–72.
2. Azaiez M, Ben Belgacem F and Chacón Rebollo T. Error bounds for POD expansions of parameterized transient temperatures. Submitted to Comp. Methods App. Mech. Eng.
3. Berkoz G, Holmes P, Lumley JL. The proper orthogonal decomposition in the analysis of turbulent flows. Annu Rev Fluids Mech. 1993;25:539–75.
4. Bernardi C, Maday Y. Approximations spectrales de problèmes aux limites elliptiques, Mathématiques et applications. Berlin: Springer; 1992.
5. Chinesta F, Keunings R, Leygue A. The Proper Generalized Decomposition for Advanced Numerical Simulations: A Primer. New York: Springer Publishing Company, Incorporated; 2013.
6. Chinesta F, Ladevèse P, Cueto E. A Short Review on Model Order Reduction Based on Proper Generalized Decomposition. Arch Comput Methods Eng. 2011;18:395–404.
7. De Lathauwer L, De Moor B, Vandewalle J. A multilinear singular value decomposition. SIAM J Matr Anal Appl. 2000;21(4):12531278.
8. De Lathauwer L, De Moor B, Vandewalle J. On the Best Rank-One and Rank-R1; R2;.; RN Approximation of Higher Order Tensors. SIAM J Matr Anal Appl. 2000;21(4):13241342.
9. De Silva V, Lim LH. Tensor rank and the ill posedness of the best low-rank approximation problem. SIAM J Matrix Anal Appl. 2008;20(3):1084–127.
10. Diestel J and Uhl. Vector measures. AMS J J; 1977.
11. Epureanu BI, Tang LS, Paidoussis MP. Coherent structures and their influence on the dynamics of aeroelastic panels. Int J Non-Linear Mech. 2004;39:977–91.
12. Ghanem R and Spanos P. Stochastic finite elements: a spectral approach. Springer-Verlag; 1991.
13. Golub GH, Van Loan CF. Matrix Computations. 3rd ed. Baltimore: The Johns Hopkins University Press; 1996.
14. Heyberger C, Boucard PA, Néron D. A rational strategy for the resolution of parametrized problems in the PGD framework. Comp Meth Appl Mech Eng. 2013;259:40–9.
15. Heyberger C, Boucard PA, Néron D. Multiparametric Analysis within the Proper Generalized Decomposition Framework. Comput Mech. 2012;49(3):277–89.
16. Holmes P, Lumley JL, Berkooz G. Coherent Structures, Synamical Systems and Symmetry, Cambridge Monographs on Mechanis. Cambridge: Cambridge University Press; 1996.
17. Hotelling H. Analysis of a complex of statistical variables into principal componentse. J Educ Psychol. 1933;24:417–41.
18. Jolliffe IT. Principal Component Analysis. Springer; 1986.
19. Little G, Reade JB. Eigenvalues of analytic kernels. SIAM J Math Anal. 1984;15:133–6.
20. Loève MM. Probability Theory. Princeton: Van Nostrand; 1988.
21. Lorente LS, Vega JM, Velazquez A. Generation of Aerodynamic Databases Using High-Order Singular Value Decomposition. J Aircraft. 2008;45(5):1779–88.
22. Muller M. On the POD Method. An Abstract Investigation with Applications to Reduced-Order Modeling and Suboptimal Control, Ph D Thesis. Georg-August Universität, Göttingen; 2008.

23. Nouy A. A generalized spectral decomposition technique to solve a class of linear stochastic partial differential equations. Comput Meth Appl Mech Eng. 2007;196:4521–37.
24. Pearson K. On lines and planes of closest fit system of points in space. Philo Mag J Sci. 1901;2:559–72.
25. Trefethen LN. Approximation theory and approximation practice. Software, Environments, and Tools. Philadel-phia: Society for Industrial and Applied Mathematics (SIAM); 2013.
26. Willcox K, Peraire J. Balanced Model Reduction via the Proper Orthogonal Decomposition. AIAA. 2002;40:2323–30.
27. Yano M. A Space-Time Petrov-Galerkin Certified Reduced Basis Method: Application to the Boussinesq Equations. SIAM J Sci Comput. 2013;36:232–66.

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ► Immediate publication on acceptance
- ► Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com