Reduced Order Data and Modelling

M. Azaïez (Bdx INP), T. Chacon (IMS) et al

CIMPA2019, Sciences des données

Table of Contents

Modern day modelisation problem

Increasing complexity of computer modelled problems :

- Multiphysics simulation
- Optimisation and control
- Big data analysis
- Identification and Inverse problems

Example : *Optimizing a flow in a multiparametric framework*

$$\begin{cases} \frac{D\mathbf{v}}{Dt} = \frac{1}{Re}\nabla^2\mathbf{v} - \nabla p \\ \mathbf{v}(x, t=0) = \mathbf{v}_0(x) \quad on\ \Omega \\ \mathbf{v} = \mathbf{f} \quad on\ \partial\Omega \end{cases}$$

Possible parameters : x, t, Re, $\mathbf{v}_0$, $\mathbf{f}$
Example : 5D optimization $\longrightarrow$ reduced order modelling.

Two directions :

- Tensor Reduction
- Projection on reduction basis

Table of Contents

Proper Orthogonal Decompositions

- POD is a powerful and elegant method of data analysis aimed at obtaining lowdimensional approximate descriptions of high-dimensional processes.
- POD was developed by several people among the first was Kosambi, J. (1943)[1]
- It is also known as Principal Component Analysis , the Karhunen–Loéve Decomposition, and the Single Value Decomposition (SVD).

POD has been used

- Approximate, low-dimensional descriptions of turbulent fluid flows
- Structural vibration
- Damage detection
- Control of fluid flows

POD has been extensively appleid to

- Image processing,
- Signal analysis
- Data compression.

Kosambi, J. Indian Math. Soc., 1943, 7, 76–88.

Karhunen-Loève expansion : what is it ?

- T is a given function in the Lebesgue space $L^2(X \times Y)$. $X \subset \mathbb{R}^d$ and $Y \subset \mathbb{R}^s$
- The integral operator B with kernel T and its adjoint operator B^* expressed as

$$\varphi \mapsto B\,\varphi, \qquad (B\,\varphi)(y) = \int_X T(x,y)\varphi(x)\,dx.$$

$$v \mapsto B^*\,v, \qquad (B^*\,v)(x) = \int_Y T(x,y)v(y)dy.$$

- $A = B^*B$ is an integral operator whose kernel K is

$$K(x,\xi) = \int_Y T(x,y)\,T(\xi,y)\,dy.$$

Résultat 1

$\exists$ a Hilbert basis $(\varphi_m)_{m\geq 0}$ in $L^2(X)$ where φ_m is an eigenvector of A related to a non-negative eigenvalue λ_m, such as

$$A\,\varphi_m = \lambda_m\,\varphi_m, \qquad \forall m \geq 0. \tag{1}$$

$$K(x,\xi) = \sum_{m\geq 0} \lambda_m \varphi_m(x)\varphi_m(\xi), \qquad \forall (x,\xi) \in X \times X.$$

Karhunen-Loève expansion

Résultat 2

There exists a system $(\varphi_m, v_m, \sigma_m)_{m\geq 0}$ such that $(\varphi_m)_{m\geq 0}$ is an orthonormal basis in $L^2(X)$, $(v_m)_{m\geq 0}$ an orthonormal system in $L^2(Y)$ and $(\sigma_m)_{m\geq 0}$ a sequence of nonnegative real numbers such that

$$B\,\varphi_m = \sigma_m\, v_m, \qquad B^*\, v_m = \sigma_m\, \varphi_m, \qquad \sigma_m = \sqrt{\lambda_m}.$$

The sequence $(\sigma_m)_{m\geq 0}$ is ordered decreasingly and decays toward zero.

Résultat 3

A direct result is the Karhunen-Loève/POD expansion (POD).

$$T(x,y) = \sum_{m\geq 0} \sigma_m\ \varphi_m(x) v_m(y), \qquad \forall (x,y) \in X \times Y.$$

The *(KL-approximation)* of function T of order M is denoted T_M and is

$$T_M(x,y) = \sum_{m=0}^{M} \sigma_m\ \varphi_m(x) v_m(y), \qquad \forall (x,y) \in X \times Y.$$

Approximation errors

Résultat 4

$$\frac{\|T - T_M\|_{L^2(X\times Y)}}{\|T\|_{L^2(X\times Y)}} = \sqrt{\frac{\sum_{m\geq M+1}\lambda_m}{\sum_{m\geq 0}\lambda_m}}. \tag{2}$$

Résultat 5

Let $(\psi_m)_{m\geq 0}$ be a Hilbertian basis in $L^2(X)$. We set $u_m(y) = \int_X T(x,y)\psi_m(x)\,dx$ and

$$S_M = \sum_{0\leq m\leq M} \psi_m \otimes u_m.$$

$$\|T - T_M\|_{L^2(X\times Y)} \leq \|T - S_M\|_{L^2(X\times Y)}. \tag{3}$$

Résultat 6

$X = I =]-1,1[$. Assume that $T \in H^\tau(I, L^2(Y))$ for $\tau \geq 0$. Then, the following bound holds

$$\|T - T_M\|_{L^2(I\times Y)} \leq C_T\, M^{-\tau}.$$

Time/Space separation : Transient temperature

Continuous problem

We consider the case of having a heat source term The heat model to handle is the following : *find a temperature field T satisfying*

$$\begin{aligned} \partial_t T - \text{div}\,(\gamma \nabla T) + \beta T &= S, && \text{in } \Omega \times J, \\ T &= 0, && \text{in } \partial\Omega \times J, \\ T(\cdot, 0) &= a_0(x), && \text{in } \Omega. \end{aligned}$$

- $t \in J =]0, b[$ and $x \in \Omega$
- γ is the conduction parameter
- β is the heat transfer coefficient
- a_0 is an initial conduction
- The source can be $S(x,t) = a(x) \otimes \theta(t)$ or not.

Azaïez, M. and Ben Belgacem, F., Karhunen-Loève's truncation error for bivariate functions, Computer Methods in Applied Mechanics and Engineering, Vol 290, pp 57-72, 2015

Numerics : Separeted source term

- $I = (0, 1)$ and $J = (0, 1)$.
- We consider

$$S_1(t, x) = (\theta \otimes a)(x, t) = e^t(x - 0.4),$$
$$S_2(t, x) = (\theta \otimes a)(x, t) = e^t|x - 0.4|.$$

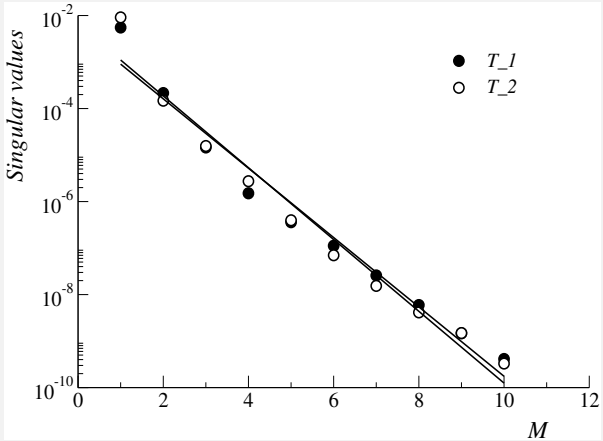
- The corresponding solutions are denoted by T_1 and T_2, respectively.
- The heat problem is discretized by an Euler scheme/Gauss-Lobatto-Legendre spectral method (the time step is $\delta t = 10^{-2}$ and the polynomial degree is $N = 64$.
- Quadrature formulas are used to evaluate the matrix representation of the operators B and A.

Remarks

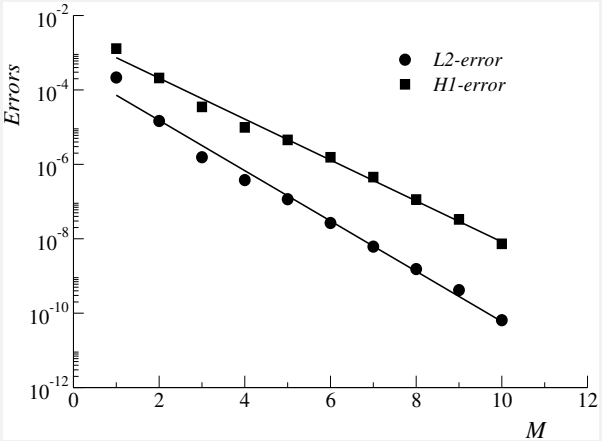
- S_1 indefinitely smooth ans S_2 enjoys moderate spacial regularity ($a \in H^\tau(I)$ with τ no bigger than $3/2$)
- Our aim is to show that the regularity has not much importance in the separation aptitude of both temperature fields T_1 and T_2

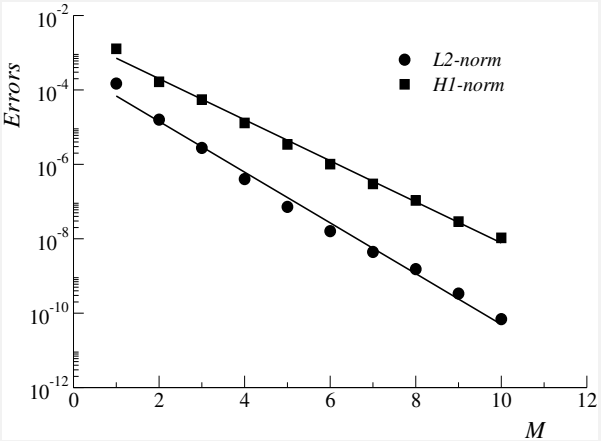
Numerics : Separeted source term

Singular values for T_1 and T_2



Truncation errors versus the cut-off M, for T_1 and T_2





Numerics : Non separeted source term

- $I = (0,1)$ and $J = (0,1)$.
- We consider

$$S_3(t,x) = \sqrt{|x - t - 0.3|}$$

- The corresponding solutions are denoted by T_3
- The solution of the corresponding heat problem is such that

$$T_3 \in H^{(3)-}(I; L^2(J)) \cap L^2(I; H^{(3/2)-}(J)).$$

- T_3 is not accessible. We use a reference computed field. It is computed using $\delta t = 10^{-3}$ and $N = 256$.

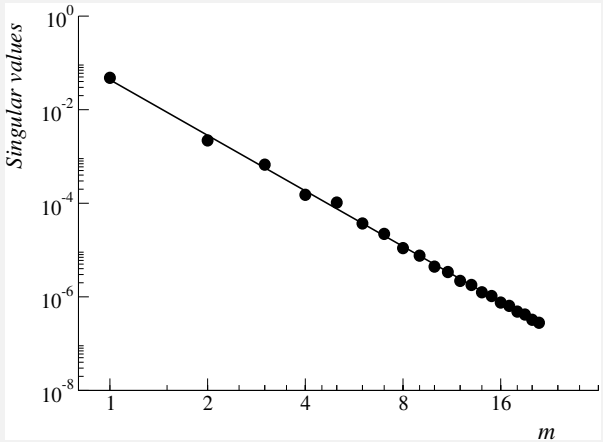
Remark

According to the estimate (2), the L^2-norm of the truncation error is expected to behave —for lower cut-off frequencies— as

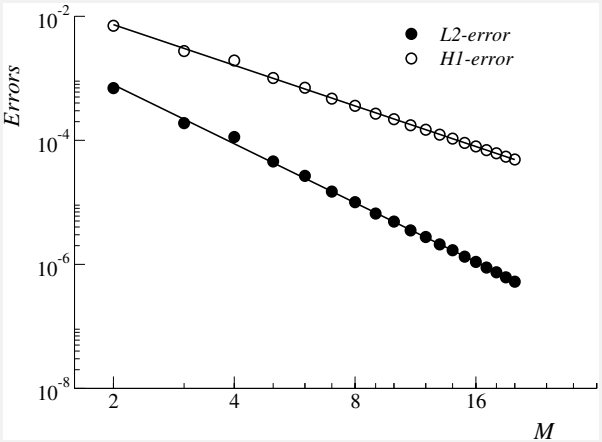
$$\|T - T_M\|_{L^2(I\times J)} \leq C\sqrt{\sum_{m\geq M+1}(\sigma_m)^2} \approx C\sqrt{\sum_{m\geq M+1} m^{-7.88}} \approx CM^{-3.44}.$$

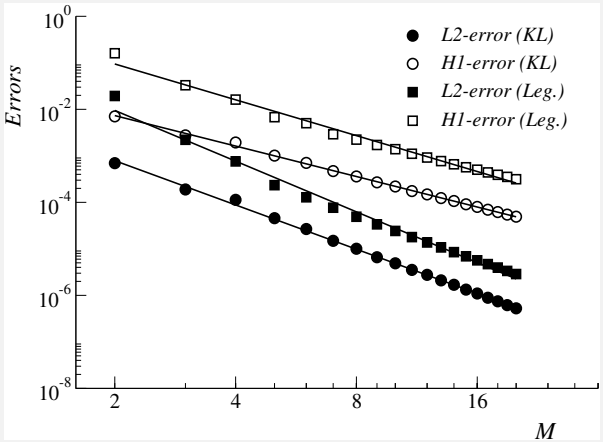
Numerics : Non separeted source term

Singular values for T_3



K-L errors for T_3 (left). Both K.-L. and Legendre polynomial errors (right).





Transient temperature

Explanation : Heat eqution with no source term

The heat model to handle is the following : *find a temperature field* T *satisfying*

$$\begin{aligned} \partial_t T - \text{div}\,(\gamma \nabla T) + \beta T &= 0, && \text{in } \Omega \times J, \\ T &= 0, && \text{in } \partial\Omega \times J, \\ T(\cdot, 0) &= a_0(x), && \text{in } \Omega. \end{aligned}$$

- $t \in J =]0, b[$ and $x \in \Omega$
- γ is the conduction parameter
- β is the heat transfer coefficient

Eigenvalue problem

$$\begin{aligned} -\,\text{div}\,(\gamma \nabla e) + \beta e &= r e, && \text{in } \Omega, \\ e &= 0, && \text{on } \partial\Omega. \end{aligned} \tag{4}$$

- The sequence of eigenvalues $(r_m)_{m \geq 1}$ is positive and grows to infinity like $C m^{2/s}$.
- $(e_m)_{m \geq 1}$ form an orthonormal basis in $L^2(\Omega)$; it is orthogonal and dense in $H_0^1(\Omega)$

Explanation

- Using this basis, the solution of the boundary value heat problem writes

$$T(x,t) = \sum_{m\geq 1} a_m e^{-r_m t} e_m(x), \qquad \forall (x,t) \in \Omega \times J. \tag{5}$$

  Fourier coefficients $(a_m)_{m\geq 1}$ are square summable ($\in \ell^2(\mathcal{R})$).

- It is easily seen that

$$\|\nabla T\|^2_{L^2(\Omega\times J)} = \frac{1}{2}\sum_{m\geq 1}(a_m)^2(1-e^{-2r_m b}) < \infty.$$

- Weak smoothness on T with respect to x and t. Hence,

The following bound on the (KL)-approximation error

$$\|T - T_M\|_{L^2(\Omega\times J)} \leq C_a\, M^{-1}.$$

It is not in accordance with the observations made in many papers $\Longrightarrow$ The geometry of the system $(e^{-r_m t})_{m\geq 1}$ plays an important role in enhancing the estimate.

Explanation

- The kernel K related to A is given by

$$K(x,\xi) = \sum_{k\geq 1}\sum_{m\geq 1} a_m a_k \left(\int_{(0,b)} e^{-(r_m+r_k)t}\, dt \right) e_k(x) e_m(\xi)$$
$$= \sum_{k\geq 1}\sum_{m\geq 1} a_k \frac{1-e^{-(r_m+r_k)b}}{r_m+r_k} a_m\, e_k(x) e_m(\xi).$$

- Now, we expand the function φ as follows

$$\varphi(x) = \sum_{m\geq 1} f_m e_m(x), \qquad \forall x \in \Omega.$$

- After replacing in the eigenvalue equation, we get

$$\sum_{k\geq 1} \left(\sum_{m\geq 1} a_k \frac{1-e^{-(r_m+r_k)b}}{r_m+r_k} a_m f_m \right) e_k(x) = \lambda \sum_{k\geq 1} f_k\, e_k(x).$$

This eigenvalue problem can be written under a matrix form, as follows

$$\mathcal{G}\boldsymbol{\varphi} = \lambda \boldsymbol{\varphi}. \tag{6}$$

Explanation

The **Gram matrix** is defined as

$$\mathcal{G} = \mathcal{A}\mathcal{C}_b\mathcal{A}$$

The entries of different matrices for all $k, m \geq 1$ are provided by

$$g_{km} = a_k c_{km} a_m \qquad c_{km} = \frac{1 - e^{-(r_m + r_k)b}}{r_m + r_k}, \qquad a_{km} = a_k \delta_{km} \tag{7}$$

- The matrix $\mathcal{C}_b$ is symmetric and positive definite.
- The matrix $\mathcal{G}$ is non-negative definite.
- The kernel $N(\mathcal{G})$ coincides with the kernel $N(\mathcal{A})$.

Resultat 7

The eigenvalues decay fast towards zero.
Assume that $a_m = 0, \forall m \geq M$ for a large integer M. The following bound holds

$$\sigma_m \leq C\|a\|_{L^2(\Omega)} \exp\left(-\frac{\mu m}{\log(r_M)}\right), \qquad 1 \leq m \leq M.$$

μ is a positive real number.

A Mechanical illustration : DNS Lid Driven Cavity data

Adimensional Navier-Stokes equation, (ψ, ω) formulation,

$$\begin{cases} \nabla^2 \psi = -\omega \\ \frac{\partial \omega}{\partial t} + (\vec{V} \cdot \nabla)\omega = \frac{1}{Re} \nabla^2 \omega \end{cases}$$

Simulation settings : $dt = 0.002$,$dx = 1/257$, $\Omega = [0,1] \times [0,1]$, $\mathcal{T} = [0, 1000]$

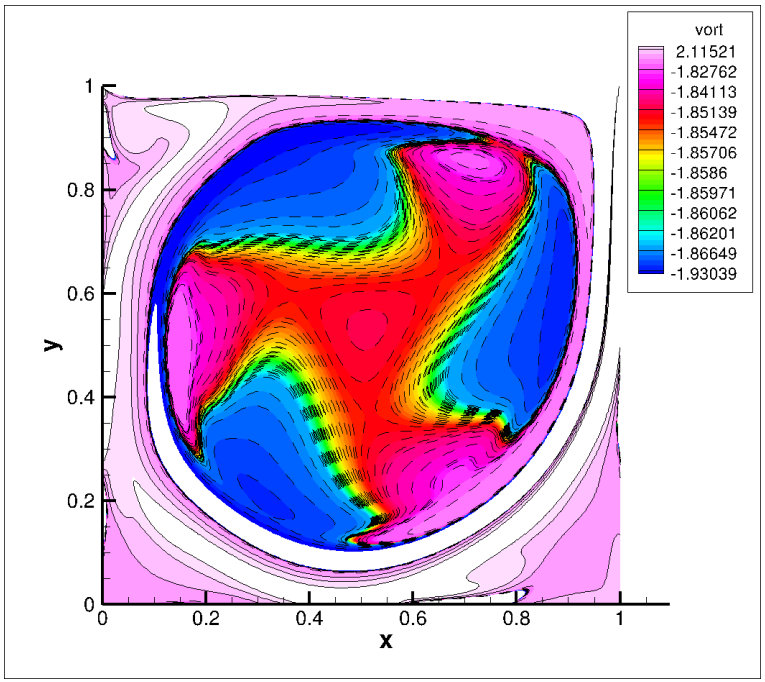


FIGURE 1 – Vorticity contour of the DNS simulation of Lid driven cavity at Re=8900, t=900.

POD on the LDC simulation

POD parameters : $t \in [500, 600]$, 500 snapshots, target error : $\varepsilon = 1\%$

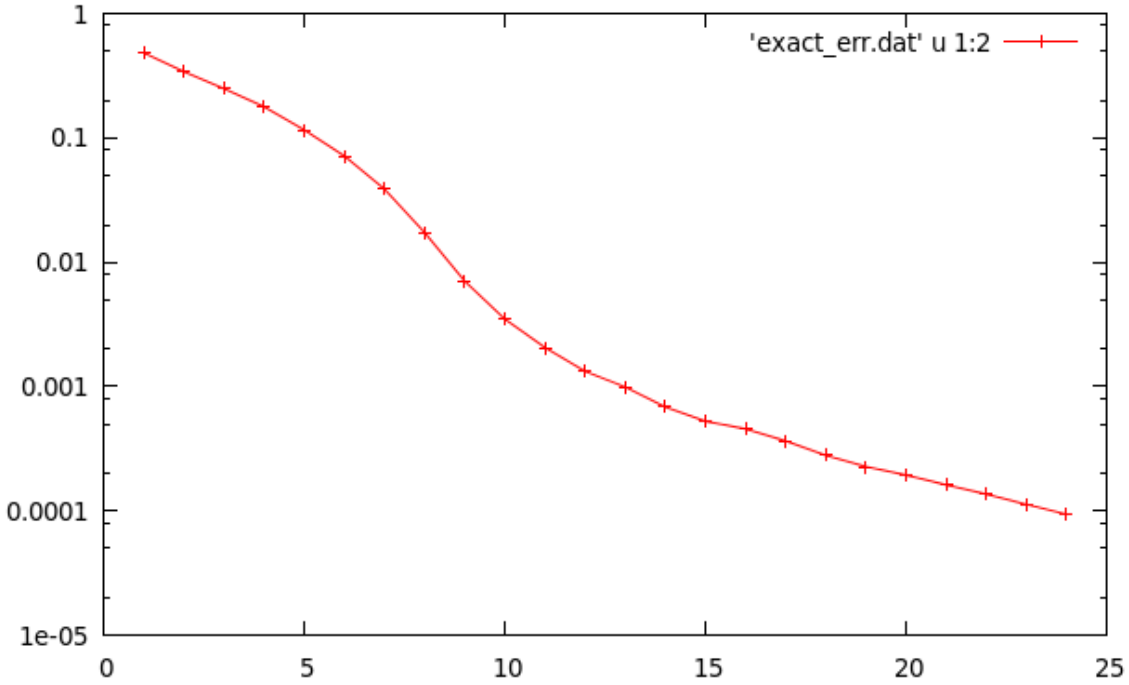


FIGURE 2 – Vorticity representation error decreases quickly toward 1%, K=15 is enough.

Type	Double number	Mem size
Full	$N_t \times N_x$	$\simeq 250Mo$
POD, K=15	$K \times (N_t + N_x)$	$\simeq 7Mo$

TABLE 1 – Storage cost comparison for POD and full data

What does a POD mode looks like ?

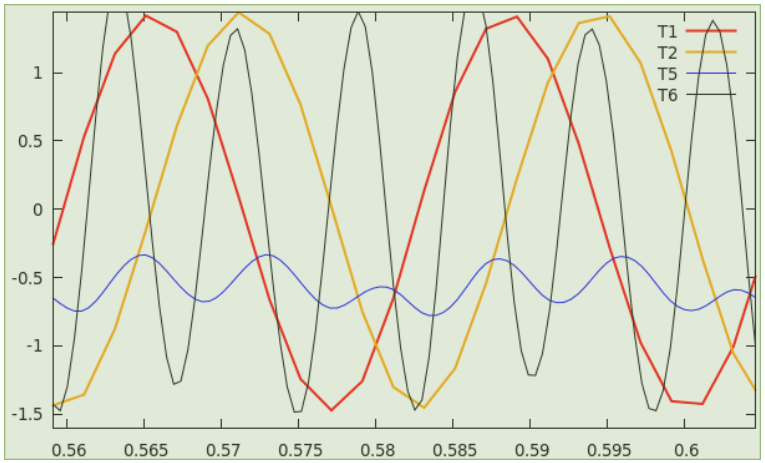


FIGURE 3 – A few amplitudes (time modes) for LDC simulation POD

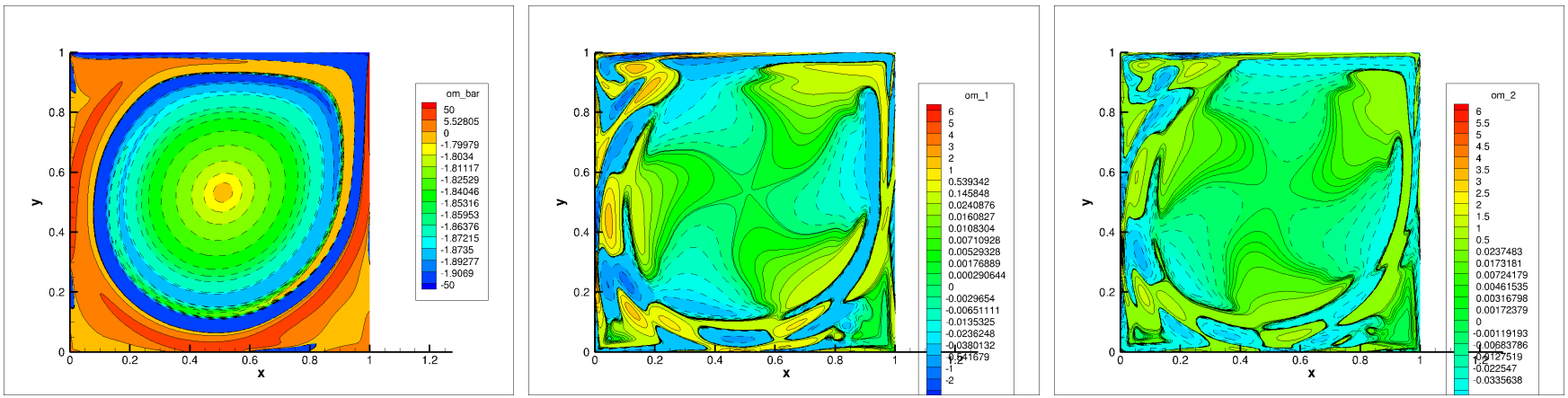


FIGURE 4 – Spatial modes for LDC simulation POD (from left to right : $\bar{\omega}, X_1, X_2$)

Parameter/Time-Space separation

As an example we consider the parameterized transient temperatures

$$\begin{aligned} \partial_t T - \gamma\,\Delta T &= f && \text{in } Q, \\ T &= 0 && \text{in } \partial\Omega \times (0,b), \\ T(x,0) &= a(x) && \text{in } \Omega, \end{aligned} \tag{8}$$

- γ varies $G \subset \mathbb{R}$ and $\mathcal{Q} = \Omega \times]0, T^*]$ is a space-time bounded domains.
- T is function in $L^2(G \times \mathcal{Q})$

K-L approximation

$$T(\gamma, z) \approx T_M(\gamma, z) = \sum_{m}^{M} \psi_m(\gamma) w_m(z), \qquad \forall (\gamma, z = (x,t)) \in G \times Q.$$

M. Azaïez, F. Ben Belgacem and T. Chacón Rebollo., Error Bounds for POD expansions of parameterized transient temperatures Computer Methods in Applied Mechanics and Engineering, 305 (2016) 501-511

K-L for Parameterized transient temperatures

- Define B maps $L^2(G)$ into $L^2(\mathcal{Q})$ with kernel T expressed as

$$\varphi \mapsto B\,\varphi, \qquad (B\,\varphi)(z) = \int_G T(\gamma, z)\varphi(\gamma)\, d\gamma. \tag{9}$$

- B^* its adjoint defined from $L^2(\mathcal{Q})$ into $L^2(G)$ as

$$v \mapsto B^*\, v, \qquad (B^*\, v)(\gamma) = \int_{\mathcal{Q}} T(\gamma, z)v(z)dz. \tag{10}$$

- $A = B^*B$ is an integral operator whose kernel K. It is expressed by

$$K(\gamma, \mu) = \int_{\mathcal{Q}} T(\gamma, z)\, T(\mu, z)\, dz. \tag{11}$$

Resultat 7

- The self-adjoint operator A is linear, bounded and compact.
- There exists a complete orthonormal basis of $L^2(G)$ formed by eigenvectors $(\varphi_m)_{m\geq 0}$ of A, associated to non-negative eigenvalues $(\lambda_m)_{m\geq 0}$
- Using Mercer's theorem yields the following decomposition

$$K(\gamma,\mu)=\sum_{m\geq 0}\lambda_m\,\varphi_m(\gamma)\varphi_m(\mu),\qquad \text{in } G\times G.$$

Resultat 8

- Setting $\sigma_m=(\lambda_m)^{1/2}$ (the singular values of B), the sequence $(v_m)_{m\geq 0}$ given by $v_m=\frac{1}{\sigma_m}B\varphi_m$ is an orthogonal basis of $L^2(\mathcal{Q})$
- A direct result is the Karhunen-Loève/POD expansion (POD).

$$T(\gamma,z)=\sum_{m\geq 0}\sigma_m\,\varphi_m(\gamma)v_m(z),\qquad \forall(\gamma,z)\in G\times Q.$$

The *(KL-approximation)* of function T of order M is denoted T_M and is is

$$T_M(\gamma,z)=\sum_{m=0}^{M}\sigma_m\,\varphi_m(\gamma)v_m(z),\qquad \forall(x,y)\in X\times Y.$$

Theorem

Assume that $f \in L^2(\mathcal{Q})$ and $a \in L^2(\Omega)$. The truncated POD series expansion T_M satisfies the error estimate

$$\|T - T_M\|_{L^2(G\times Q)} \leq C_\rho \ \rho^{-M}, \qquad \forall \rho,\ 1 < \rho < \rho_*, \tag{12}$$

where $C_\rho > 0$ is a constant depending on ρ.

The proof is based on

- The function $\gamma \mapsto T(\gamma)$, mapping $]0, +\infty[$ into $L^2(\mathcal{Q})$, is analytic.
- $\exists$ a polynomial S_M ranging from G into $L^2(\mathcal{Q})$, with degree $\leq M$, s.t.

$$\max_{\gamma\in[\gamma_m,\gamma_M]} \|T(\gamma) - S_M(\gamma)\|_{L^2(\mathcal{Q})} \leq C_\rho\, \rho^{-M},\ \forall \rho\, (1 < \rho < \rho_*),$$

with

$$\rho_* = \frac{\left(\sqrt{\gamma_m} + \sqrt{\gamma_M}\right)^2}{\gamma_M - \gamma_m} > 1.$$

and

$$C_\rho = \frac{2}{\rho - 1}\|T\|_{L^\infty(\mathcal{E}_\rho)}.$$

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

Numerical experiments

- We consider the time-dependent heat equation in the domain $\mathcal{Q} = (0,1) \times (0,1)$.

$$\begin{aligned} \partial_t T - \gamma \, \Delta T &= f && \text{in } Q, \\ T &= 0 && \text{in } \partial\Omega \times (0,b), \\ T(x,0) &= a(x) && \text{in } \Omega, \end{aligned} \tag{13}$$

- We select three possible pairs of source terms and boundary conditions, given by

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

Remark

These data have singularities, so the temperature solutions of (13), have a low regularity with respect to x and t

Numerical experiments

The heat problem is discretized by

- $G = [1, 100]$
- Euler scheme/Gauss-Lobatto-Legendre spectral method
- $\delta t = 10^{-2}$ and the polynomial degree is $N = 64$)
- B and A are computed by means of Gauss-Lobatto quadrature formulas

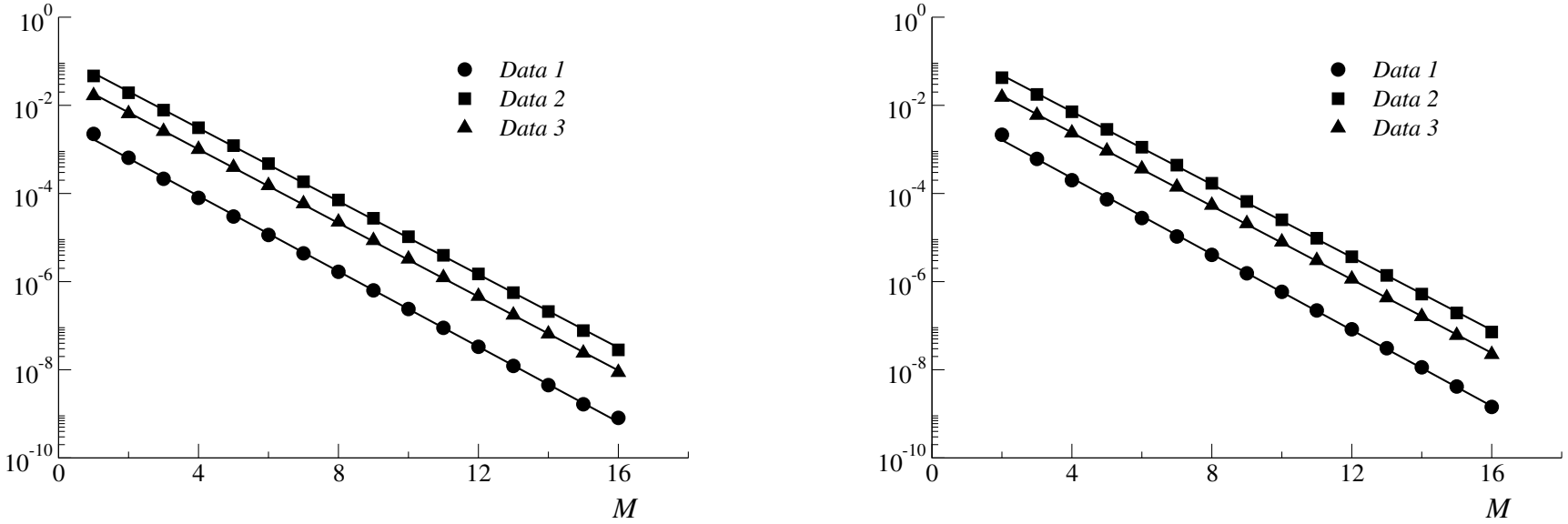


FIGURE 5 – History for the POD-error (left). Largest singular values (σ_M) (right).

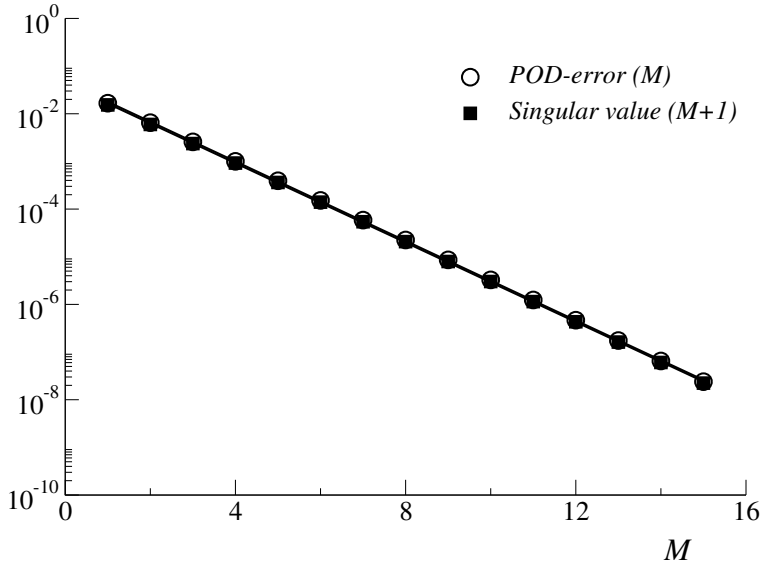


FIGURE 6 – POD-error $\|T - T_M\|_{L^2}$ and the singular value σ_{M+1} (for Data 1).

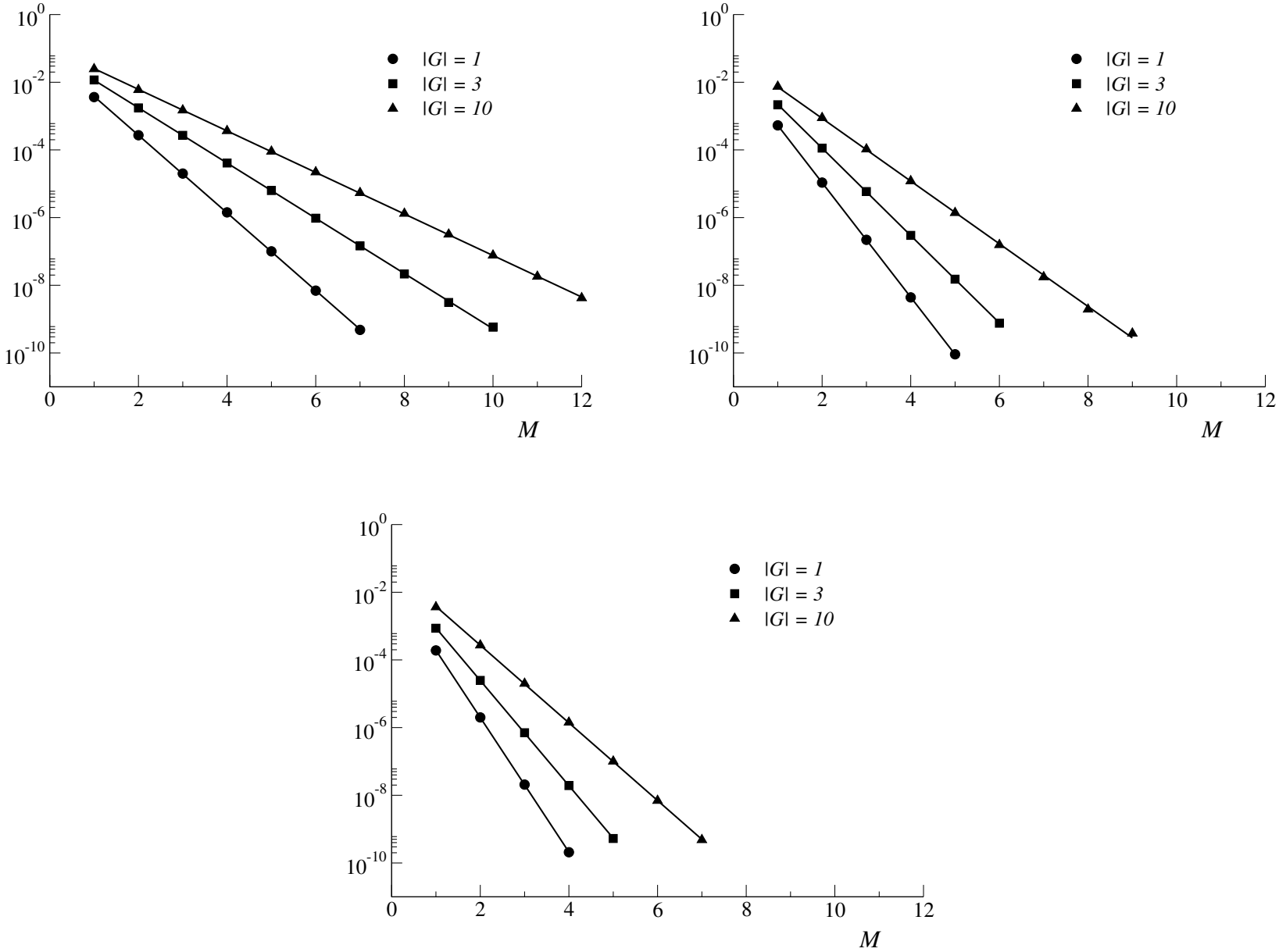


FIGURE 7 – Variation of the POD-errors. Each diagram corresponds to a different location of $\gamma_m = 1, 5$ or 10. In each plot three cases are considered, the conductivities range equals $1, 3$ or 10.

Table of Contents

Multivariate Data Reduction Techniques

Definition : tensor (d-way array)

$A \in \mathbb{R}^{\mathcal{I}}$ where $\mathcal{I} = I_1 \times ... \times I_d$, $d \in \mathbb{N}$ is a (algebraic) tensor.

Tensor Reduction : *find an efficient way to reduce the storage cost of* A.

For K "small" find a good approximation A_K of A with

$$A_K = \sum_{k=1}^{K} \sigma_k \otimes_{i=1}^{d} u_i^k \quad \{u_i^k\}_{i,k} \in \mathbb{R}^{I_i} \tag{14}$$

Equivalent functional formulation

Let $f : \Omega \in \mathbb{R}^d \longrightarrow \mathbb{R}$ find a (finite) separated expression to f.

$$f(x_1, \ldots, x_d) \approx f_K(x_1, \ldots, x_d) = \sum_{k=1}^{K} \varphi_1^k(x_1) \ldots \varphi_d^k(x_d)$$

Discret approach

- Higher-order singular value decomposition (HOSVD) : introduced by Tucker [1]
- Truncated higher-order singular value decomposition (T-HOSVD) : introduced by et Lathauwer et al. [2]
- Sequentially truncated higher-order singular value decomposition (ST-HOSVD) : introduced by Vannieuwenhoven et al. [3]


L.R. Tucker, *Some mathematical notes on three-mode factor analysis*, Psikometrica, **31** (1966), pp. 279–311.


L. De Lathawer, B. De Moor, J. Vandewalle, *On the best rank-*1 *and rank-*$(R_1, R_2, \ldots, R_N)$ *approximation of higher-order tensors*, SIAM J. Matrix. Anal. Appl. **21** (2000), pp. 1324–1342.


N. Vannieuwenhoven, R. Vandebril, K. Meerbergen, A new truncation strategy for the higher-order singular value decomposition, SIAM J. Matrix. Anal. Appl. **21** (2000), pp. 1324–1342.

T-HOSVD

Input : Data $f(x_1, x_2, x_3)$ and Expanding Reduced Order (r_1, r_2, r_3)

- Compute for each $i = 1, 2, 3$

$$\left(u_k^i\right)_{k=1}^{r_i} \Longleftarrow POD[f(\cdot, x_i, \cdot)]$$

- Compute the core

$$S_{ijk} = \int\int\int f(\mathbf{x}) u_i^1(x_1) u_j^2(x_2) u_k^3(x_3)\, d\mathbf{x}$$

Output : Reduced Data

$$f_h(\mathbf{x}) = \sum_{i=1}^{r_1}\sum_{j=1}^{r_2}\sum_{k=1}^{r_3} S_{ijk} u_i^1(x_1) u_j^2(x_2) u_k^3(x_3)$$

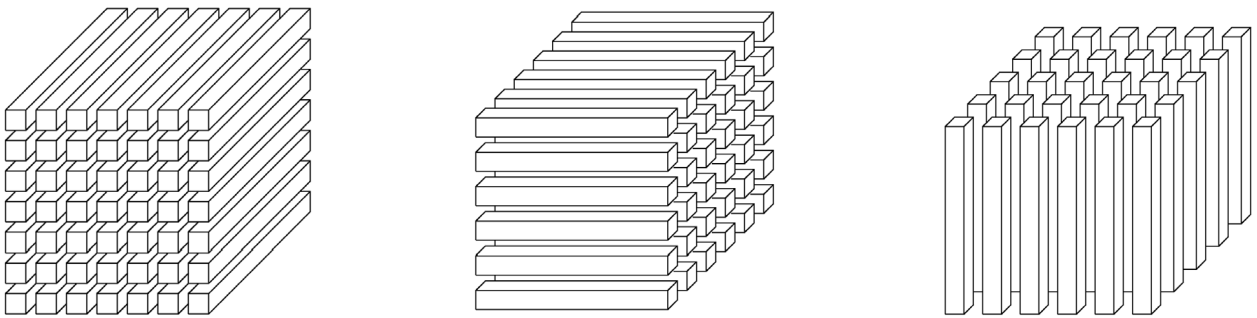


FIGURE 8 – The fibers of a third order tensor.

ST-HOSVD

Input : Data $f(x_1, x_2, x_3)$ and Expanding Reduced Order (r_1, r_2, r_3)

- Compute

$$\left(u_k^1(x_1)\right)_{k=1}^{r_1} \text{ and } \left(\varphi_k^1(x_2, x_3)\right)_{k=1}^{r_1} \Longleftarrow POD[f(x_1, x_2, x_3)]$$

and

$$\hat{f}_1 = \sum_{i=1}^{r_1} u_i^1(x_1)\varphi_k^1(x_2, x_3)$$

- Compute

$$\left(u_k^2(x_2)\right)_{k=1}^{r_2} \text{ and } \left(\varphi_k^2(x_1, x_3)\right)_{k=1}^{r_2} \Longleftarrow POD[\hat{f}_1(x_1, x_2, x_3)]$$

and

$$\hat{f}_2 = \sum_{i=1}^{r_2} u_i^2(x_2)\varphi_k^2(x_1, x_3)$$

- Compute

$$\left(u_k^3(x_2)\right)_{k=1}^{r_3} \text{ and } \left(\varphi_k^3(x_1, x_3)\right)_{k=1}^{r_2} \Longleftarrow POD[\hat{f}_2(x_1, x_2, x_3)]$$

and

$$\hat{f}_3 = \sum_{i=1}^{r_3} u_i^3(x_3)\varphi_k^3(x_1, x_2)$$

ST-HOSVD

Output : Reduced Data

$$\hat{f}_h(\mathbf{x}) = \sum_{i=1}^{r_1} \sum_{j=1}^{r_2} \sum_{k=1}^{r_3} S_{ijk} u_i^1(x_1) u_j^2(x_2) u_k^3(x_3)$$

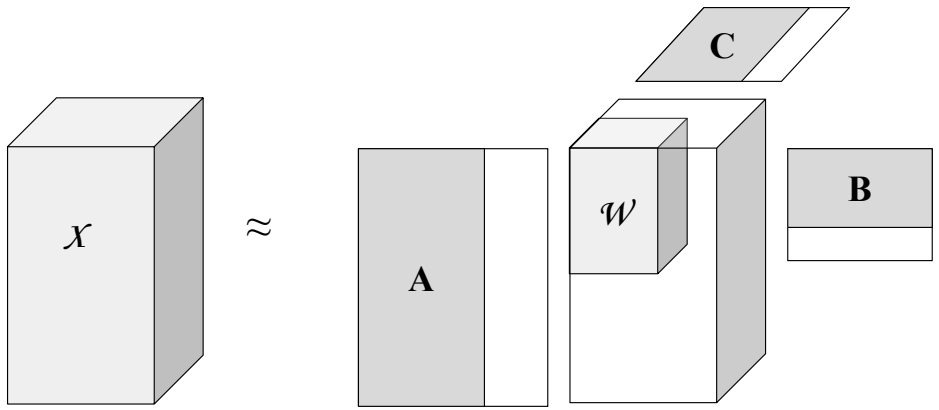


FIGURE 9 – ST-HOSVD / Tucker Decomposition of a third order array $\mathcal{X}$

PGD for data tensorisation

Objective and Concept

$$f(x_1, x_2, \ldots, x_n) \approx f_K(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{K} F_i^1(x_1) F_i^2(x_2) \cdots F_i^n(x_n). \tag{15}$$

Let $\mathcal{S}_1 = \{z(x_1, ..., x_n) = \prod_{i=1}^{n} X_i(x_i)\}$ and $\mathcal{S}_m = \{v = \sum_{j=1}^{m} z_j, z_i \in \mathcal{S}_1\}$. The problem reads :

$$u_m = \arg \min_{u \in S_m} \|f - u\|_{L^2} \tag{16}$$

Case of $n = 3$

u_{M-1} and $\mathcal{R}_{M-1}(\mathbf{x}) = f(\mathbf{x}) - u_{M-1}(\mathbf{x})$ are computed, then $F_M^1(x_1)$, $F_M^2(x_2)$ and $F_M^3(x_3)$ are solutions of : $\forall F_1^*, F_2^*, F_3^*$

$$\int_\Omega F_M^1(x_1) F_M^2 F_M^3 F_1^*(x_1) F_M^2 F_M^3 d\mathbf{x} = \int_\Omega \mathcal{R}_{m-1}(\mathbf{x})\, F_1^*(x_1) F_M^2 F_M^3 d\mathbf{x},$$

$$\int_\Omega F_M^1 F_M^2(x_2) F_M^3 F_M^1 F_2^*(x_2) F_M^1 d\mathbf{x} = \int_\Omega \mathcal{R}_{m-1}(\mathbf{x})\, F_M^1 F_2^*(x_2) F_M^3 d\mathbf{x},$$

$$\int_\Omega F_M^1 F_M^2 F_M^3(x_3) F_M^1 F_M^2 F_3^*(x_3) d\mathbf{x} = \int_\Omega \mathcal{R}_{m-1}(\mathbf{x})\, F_M^1 F_M^2 F_3^*(x_3) d\mathbf{x},$$

Recursive-POD (HOSVD)

Input : Data $f(x_1, x_2, x_3)$ and Tolerance ϵ

- Compute

$$(\sigma_{i_1})_{i_1=1}^{r_1}\ ,\ (u_{i_1}(x_1))_{i_1=1}^{r_1} \text{ and } (\varphi_{i_1}(x_2, x_3))_{i_1=1}^{r_1} \Longleftarrow POD[f(x_1; x_2, x_3)]$$

- Compute for $i_1 = 1, \ldots, r_1$

$$\left(\sigma_{i_2}^{(i_1)}\right)_{i_2=1}^{r_2(i_1)}\ ,\ \left(u_{i_2}^{(i_1)}(x_2)\right)_{i_2=1}^{r_2(i_1)} \text{ and } \left(\varphi_{i_2}^{(i_1)}(x_3)\right)_{i_2=1}^{r_2(i_1)} \Longleftarrow POD[\varphi_{i_1}(x_2, x_3)]$$

The reduced approximated data is given by

$$\hat{f} = \sum_{i=1}^{r_1} \sum_{i=2}^{r_2(i_1)} \sigma_{i_1} \sigma_{i_2}^{(i_1)} u_{i_1}(x_1) u_{i_2}^{(i_1)}(x_2) \varphi_{i_2}^{(i_1)}(x_3)$$

Tools for Multivariate data reduction

- HOSVD, THOSVD, ST-HOSVD,...
- PGD
- R-POD (R-KL, R-SVD)

  Hierarchical,..

R-HOPOD

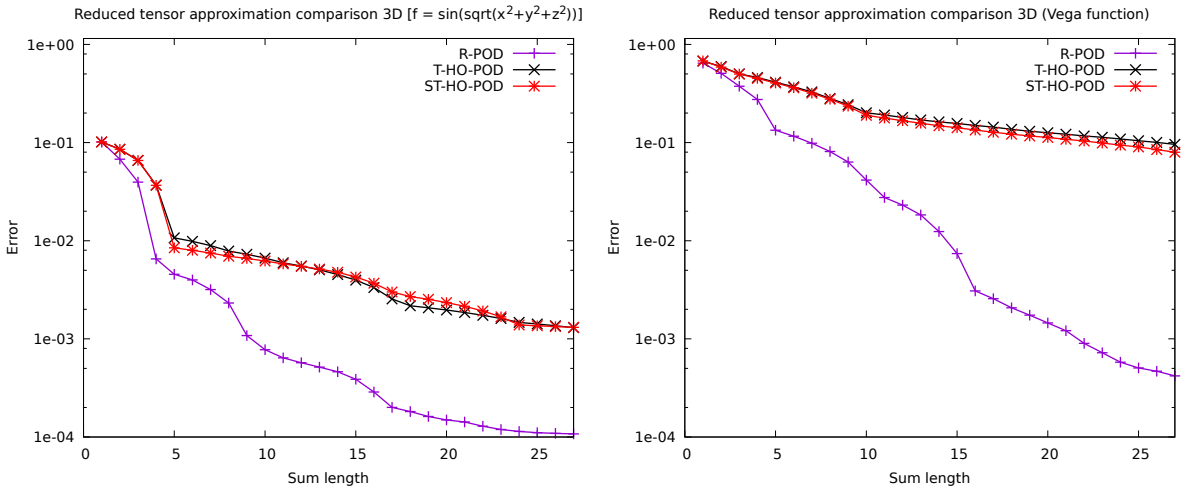


FIGURE 1 – Convergence history of R, T and ST POD, expansion error in logarithmic scale.

- Number of computed and evaluated points : $25^3 = 15625$

Recursive Proper Orthogonal Decomposition (R-POD)

Let $X_i \subset \mathbb{R}^{d_i}$, d_i integers ≥ 1, are n bounded domains. $\mathbb{X}_1^n = X_1 \times X_2 \times \ldots X_n$.

Goal

Approximating, in a low-dimensional variety, a function f in the Lebesgue space $L^2(\mathbb{X}_1^n)$.

• $L^2(\mathbb{X}_1^n)$ and $L^2(\mathbb{X}_2^n, L^2(X_1))$ are isometric, there exist $(\varphi_{i_1})_{i_1 \in \mathbb{N}}$ and $(v_{i_1})_{i_1 \in \mathbb{N}}$, which are respectively complete in $L^2(X_1)$ and $L^2(\mathbb{X}_2^n)$, such that :

$$f(\mathbf{x}_1^n) = \sum_{i_1 \in \mathbb{N}} \sigma_{i_1} \varphi_{i_1}(x_1) v_{i_1}(\mathbf{x}_2^n), \tag{17}$$

where the sum is convergent in $L^2(\mathbb{X}_2^n, L^2(X_1))$ and σ_{i_1} are convergent to zero and non-negative.

• We continue applying the POD expansion to each mode $v_{i_1}(\mathbf{x}_2^n)$.

$L^2(\mathbb{X}_2^n)$ and $L^2(\mathbb{X}_3^n, L^2(X_2))$ are isometric, there exist $(\varphi_{i_2}^{(i_1)})_{i_2 \in \mathbb{N}}$ and $(v_{i_2}^{(i_1)})_{i_2 \in \mathbb{N}}$, which are respectively complete in $L^2(X_2)$ and $L^2(\mathbb{X}_3^n)$, such that :

$$v_{i_1}(\mathbf{x}_2^n) = \sum_{i_2 \in \mathbb{N}} \sigma_{i_2}^{(i_1)} \varphi_{i_2}^{(i_1)}(x_2) v_{i_2}^{(i_1)}(\mathbf{x}_3^n),$$

converges in $L^2(\mathbb{X}_3^n, L^2(X_2))$.

Continuing in this fashion way we justify :

Lemma

The function $f \in L^2(\mathbb{X}_1^n)$ admits the expansion :

$$f = \sum_{i_1 \in \mathbb{N}} \sigma_{i_1} \left(\sum_{i_2 \in \mathbb{N}} \sigma_{i_2}^{(i_1)} \cdots \left(\sum_{i_{n-1} \in \mathbb{N}} \sigma_{i_{n-1}}^{(i_1^{n-2})} \varphi_{i_1} \otimes \varphi_{i_2}^{(i_1)} \right) \otimes \cdots \otimes \varphi_{i_{n-1}}^{(i_1^{n-2})} \right) \otimes v_{i_{n-1}}^{(i_1^{n-2})}, \tag{18}$$

where sums are convergent in $L^2(\mathbb{X}_1^n)$.

Its R-POD approximation is given by

$$f_P = \sum_{i_1=0}^{I_1} \sigma_{i_1} \left(\sum_{i_2=0}^{I_2^{(i_1)}} \sigma_{i_2}^{(i_1)} \cdots \left(\sum_{i_{n-1}=0}^{I_{n-1}^{(\mathbf{i}_1^{n-2})}} \sigma_{i_{n-1}}^{(\mathbf{i}_1^{n-2})} \varphi_{i_1} \otimes \varphi_{i_2}^{(i_1)} \right) \otimes \cdots \otimes \varphi_{i_{n-1}}^{(\mathbf{i}_1^{n-2})} \right) \otimes v_{i_{n-1}}^{(\mathbf{i}_1^{n-2})}. \tag{19}$$

P is a multi-index notation and stands for $P = \left(I_1, I_2^{(i_1)}, \ldots, I_{n-1}^{(\mathbf{i}_1^{n-2})}\right)$.
All elements of the vector P are larger than or equal to one.

Lemma

Assume that the coefficients of the RPOD expansion

$$f_P = \sum_{i_1=0}^{I_1} \sigma_{i_1} \left(\sum_{i_2=0}^{I_2^{(i_1)}} \sigma_{i_2}^{(i_1)} \cdots \left(\sum_{i_{n-1}=0}^{I_{n-1}^{(i_1^{n-2})}} \sigma_{i_{n-1}}^{(i_1^{n-2})} \varphi_{i_1} \otimes \varphi_{i_2}^{(i_1)} \right) \otimes \cdots \otimes \varphi_{i_{n-1}}^{(i_1^{n-2})} \right) \otimes v_{i_{n-1}}^{(i_1^{n-2})}.$$

satisfy the estimates

$$\sum_{I_1 < i_1} \left| \sigma_{i_1} \right|^2 \leq \Sigma_1, \quad \sum_{I_k^{(i_1^{k-1})} < i_k} \left| \sigma_{i_k}^{(i_1^{k-1})} \right|^2 \leq \Sigma_k, \; k = 2, \cdots, n-1, \tag{20}$$

then, the following error estimate in L^2 norm holds :

$$\left|\left| f - f_P \right|\right|^2_{L^2(\mathbb{X}_1^n)} \leq \Sigma_1 + C_P \left|\left| f \right|\right|^2_{L^2(\mathbb{X}_1^n)}, \tag{21}$$

where $C_P = \Sigma_2 + \Sigma_{n-2} + \Sigma_{n-1}$,

Reordering of recursive POD expansion

A natural re-ordering of expansion (18) is in decreasing order of the values $\sigma_\alpha = \sigma_{i_1}\, \sigma_{i_2}^{i_1} \cdots \sigma_{i_{n-1}}^{(i_1,i_2,\cdots,i_{n-2})}$, with $\alpha = (i_1, \cdots, i_{n-1}) \in \mathbb{N}^{n-1}$.

Expansion (18) may be written as

$$f = \sum_{k\geq 1} \sum_{|\alpha|=k} \sigma_\alpha\, \Psi_\alpha, \tag{22}$$

where $|\alpha| = i_1 + i_2 + \cdots + i_{n-1}$ and

$$\Psi_\alpha = \varphi_{i_1} \otimes \varphi_{i_2}^{(i_1)} \otimes \cdots \otimes \varphi_{i_{n-1}}^{(\mathbf{i}_1^{n-2})} \otimes v_{i_{n-1}}^{(\mathbf{i}_1^{n-2})}. \tag{23}$$

Numerical Experiments : Toy data

• In practice frequently it holds $\sigma_\alpha \simeq \rho^{-|\alpha|}$ with $\rho > 1$ for many functions f. In this case we derive a practical error estimate for the truncated expansion

$$f_K = \sum_{k=1}^{K} \sum_{|\alpha|=k} \sigma_\alpha \, \Psi_\alpha. \tag{24}$$

Once we re-order the R-POD expansion :

$$f_K(x_1, x_2, \ldots, x_{n-1}, x_n) = \sum_{l=0}^{L_K} \tilde{\sigma}_l \, \tilde{\varphi}_l^{(1)}(x_1) \tilde{\varphi}_l^{(2)}(x_2) \cdots \tilde{\varphi}_l^{(n-1)}(x_{n-1}) \tilde{\varphi}_l^{(n)}(x_n), \tag{25}$$

where $L_K = \begin{pmatrix} n+K-1 \\ K \end{pmatrix} \simeq \frac{1}{(n-1)!} K^{n-1}$.

This yields to the L^2 error in terms of the number of modes in expansion

Lemma

$$\|f - f_K\|_{L^2(\mathbf{X}^n)} \leq D'_n \, L_K^{\frac{n-2}{2(n-1)}} \, \rho^{-L_K^{\frac{1}{n-1}}}. \tag{26}$$

Numerical Experiments : Toy data

$$f_1(\mathbf{x}) = \sin(x_1 x_4 x_5^2) + \sqrt{9 + x_3 + 4x_2},$$

$$f_2(\mathbf{x}) = \frac{1}{1 + x_1 + x_2 + x_3 + x_4 + x_5},$$

$$\begin{aligned} f_3(\mathbf{x}) =& x_1^2\{\sin[5x_2\pi + 3\log(x_1^3 + x_2^2 + x_4^2 + x_3 + \pi^2)] - 1\}^2 \\ &+ (x_1 + x_3 - 1)(2x_2 - x_3)(4x_5 - x_4)\cos[30(x_1 \\ &\log(6 + x_1^2 x_2^2 + x_3^3) + x_3 + x_4 + x_5)] - 4x_1^2 x_2 x_5^3 (1 - x_3)^{3/2}. \end{aligned}$$

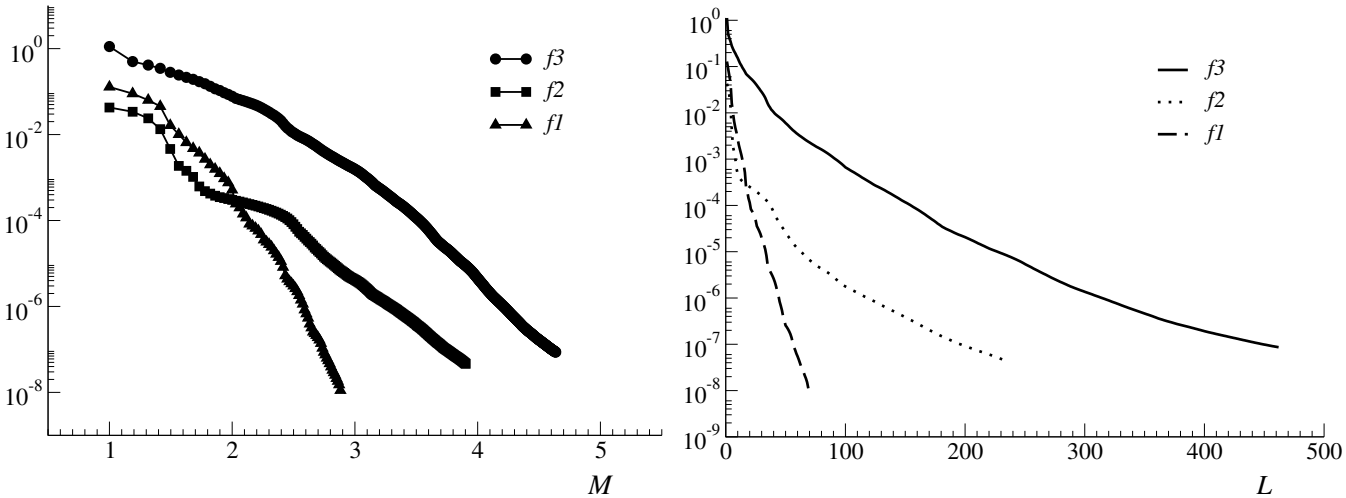


FIGURE 10 – Convergence history of RPOD expansion error in logarithmic scale. The x-axis corresponds to $L^{1/4}$ (Left), where L (Right) is the number of modes in the re-arranged expansion

Case of reaction-diffusion problem

Consider linear reaction-diffusion equation,

$$\left\{\begin{array}{rcl} \partial_t T - \gamma\,\Delta T + \alpha\, T & = & f \text{ in } \mathcal{Q} = \Omega \times (0,b), \\ T & = & 0 \text{ in } (0,b) \times \partial\Omega, \\ T(x,0) & = & T_0(x) \text{ in } \Omega, \end{array}\right. \tag{27}$$

(γ, α) ranges in a set $\mathcal{G} = [\gamma_m, \gamma_M] \times [\alpha_0, \alpha_M]$ with $0 < \gamma_m < \gamma_M$, $0 \leq \alpha_0 \leq \alpha_M$.
T is approximated by a R-POD expansion in separated tensor form :

$$T_P((x,t),(\gamma,\alpha)) = \sum_{m=0}^{M} \sum_{i=0}^{I} \tau_i^{(m)}\, \varphi_i^{(m)}(\gamma)\, w_i^{(m)}(\alpha)\, v_m(x,t), \tag{28}$$

T_P satisfies the error estimate

$$\|T - T_P\|_{L^2(\mathcal{G}\times\mathcal{Q})} \leq C_\rho\, (\rho^{-M} + \sqrt{M}\, \rho^{-I}), \tag{29}$$

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

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

Numerical Experiments : exponential convergence

We aim to assess the exponential convergence rate and investigate the variation of this rate with respect to the set $\mathcal{G} = [\gamma_m, \gamma_M] \times [\alpha_m, \alpha_M]$.
The domain $\mathcal{Q} = (0,1) \times (0,1)$ and we select three possible pairs of source terms and initial conditions, given by

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

These data have mild singularities, so the solutions have a reduced regularity with respect to x and t, The heat problem is discretized by an Euler scheme/Gauss-Lobatto-Legendre spectral method (the time step is $\delta t = 10^{-2}$ and the polynomial degree is $N = 64$).

M. Azaïez, F. Ben Belgacem, T. Chacón Rebollo, Recursive POD expansion for reaction-diffusion equation, Adv.Model. and Simul. in Eng. Sci. (2016) 3 :3. DOI 10.1186/s40323-016-0060-1

Numerical Experiments : exponential convergence

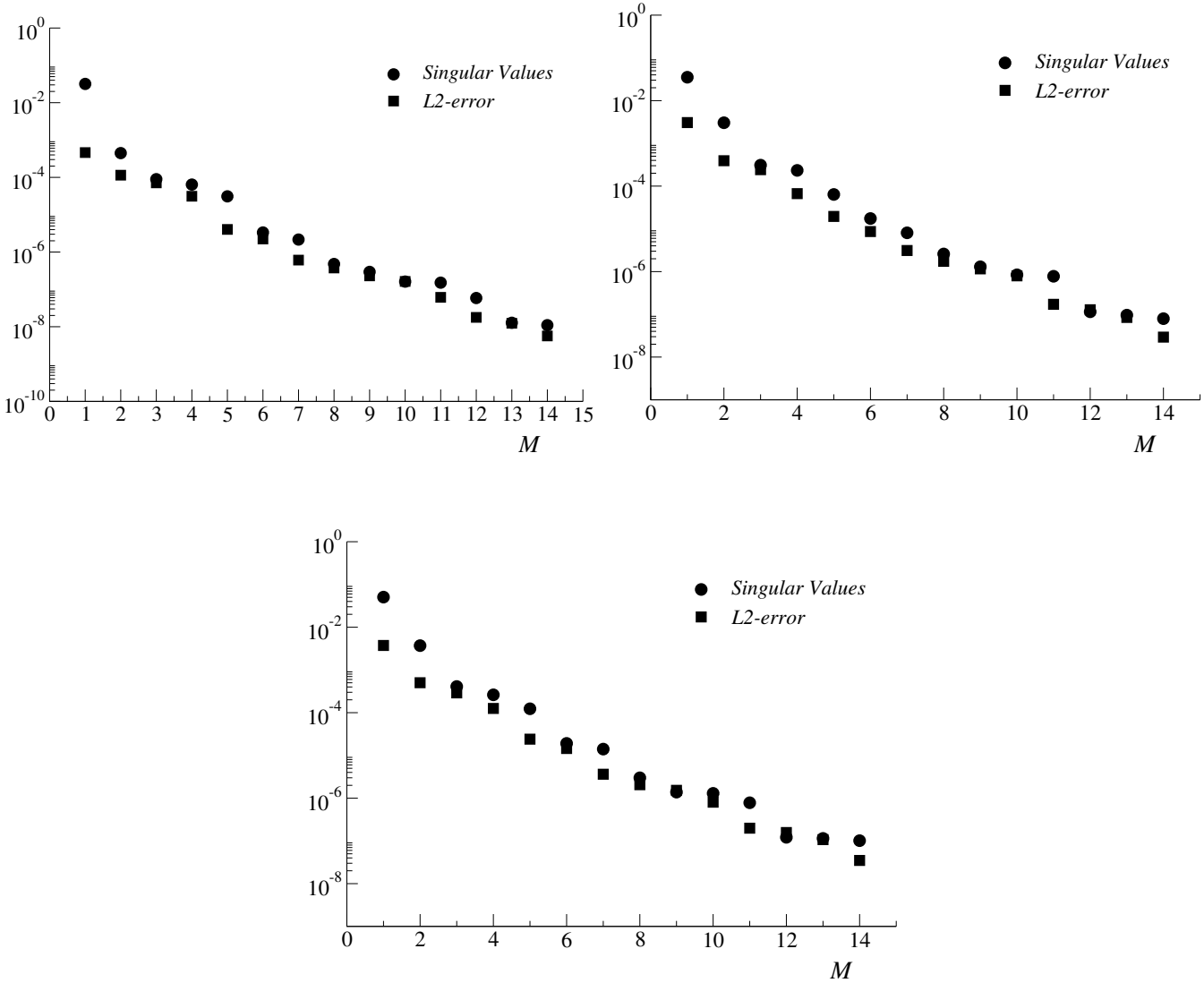


FIGURE 11 – Convergence history for POD expansion. • $\gamma \in [1, 51]$, and $\alpha \in [0, 100]$. Data 1 (top left), Data 2 (top right) and Data 3 (bottom). The error is measured in $L^2(\mathcal{Q})$ norm

Convergence rate with respect to the parameters range.

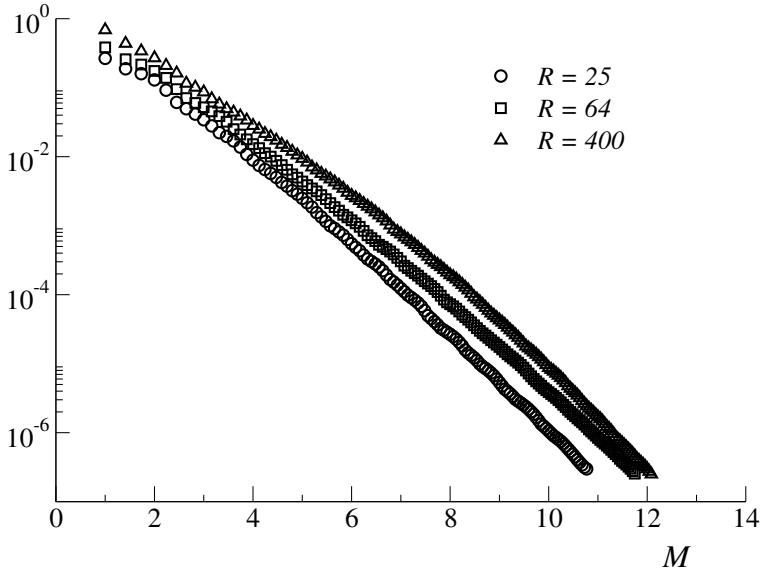


FIGURE 12 – Variation of the R-POD errors (in logarithmic scale) with respect to the ratio $R = \gamma_M/\gamma_m$, for fixed $\alpha_m = 0$, $\alpha_M = 100$. The variable M stands for the square root of the number of modes.

- The convergence rate degrades as R increases, in accordance with the fact that

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

Convergence rate with respect to the interval of reaction rates

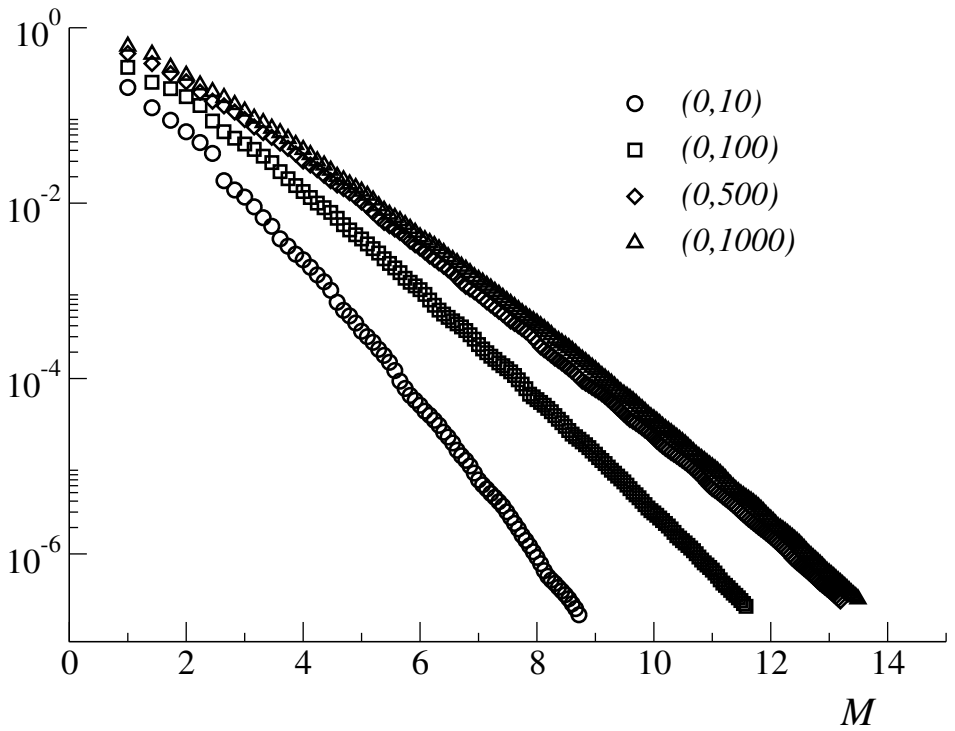


FIGURE 13 – Variation of the R-POD errors with respect to the reaction rate α. The curves correspond to $\alpha_0 = 0$, $\alpha_M = 10,\ 100,\ 500,\ 1000$, with $\gamma_m = 1$, $\gamma_M = 51$ in all cases. The variable M stands for the square root of the number of modes.

Case of advection-diffusion- reaction problem

The linear advection-diffusion-reaction equation in a bounded domain $\Omega \subset \mathbb{R}^d$, $d \geq 1$, and a time interval $(0, b)$, $b > 0$:

$$\left\{ \begin{array}{rcl} \partial_t T + \vec{a} \cdot \nabla T - \gamma \, \Delta T + \alpha \, T & = & f \text{ in } \mathcal{Q} = \Omega \times (0, b), \\ T & = & 0 \text{ in } (0, b) \times \partial\Omega, \\ T(x, 0) & = & T_0(x) \text{ in } \Omega, \end{array} \right. \tag{30}$$

We assume the triad $(\vec{a}, \gamma, \alpha)$ ranges in a set $\mathcal{G} = B_R \times [\gamma_m, \gamma_M] \times [0, \alpha_M]$ where $B_R \subset \mathbb{R}^3$ is the ball of radious $R > 0$, and $0 < \gamma_m < \gamma_M$, $0 \leq \alpha_M$.

- The R-POD approximation of T reads :

$$T_P((x,t),(\vec{a},\gamma,\alpha)) = \sum_{m=0}^{M} \sum_{i=0}^{I} \sum_{j=0}^{J} \tau_j^{(m,i)} \, p_j^{(m,i)}(\vec{a}) \, w_j^{(m,i)}(\alpha) \, u_i^{(m)}(\gamma) \, v_m(x,t), \tag{31}$$

$P = (M, I, J)$ with $M \geq 0$, $I \geq 0$, $J \geq 0$ integer numbers, the $\tau_j^{(m,i)}$ are real numbers and $p_j^{(m,i)} \in L^2(B_R)$, $w_j^{(m,i)} \in L^2(0, \alpha_M)$, $u_i^{(m)} \in L^2(\gamma_m, \gamma_M)$, and $v_m \in L^2(\mathcal{Q})$ are eigenmodes.

M. Azaïez, T. Chacón Rebollo, E. Perracchione and J. M. Vega, On the Recursive-POD approximation of multi-parametric functions : Application to the transient advection-diffusion-reaction equation, To Be or not To Be

R-POD Approximation

Let us define the constants

$$L = \sup_{x \in \Omega} |x| \text{ and } \rho_* = (r+1)/(r-1) \text{ with } r = \sqrt{\gamma_M/\gamma_m}. \tag{32}$$

The rate of convergence of T_P towards T is stated as follows

Theorem

The truncated R-POD series expansion T_P given by (31) satisfies the error estimate

$$\|T - T_P\|_{L^2(\mathcal{G}\times\mathcal{Q})} \leq C_\rho \, e^{RL/\gamma_m} \, (\rho^{-M} + \sqrt{M}\,\rho^{-I} + \sqrt{MI}\,\rho^{-J}), \tag{33}$$

for any $1 < \rho < \rho_*$, where $C_\rho > 0$ is a constant depending on ρ, unbounded as $\rho \to 1$,

- R-POD expansion converges with spectral accuracy in terms of the number of truncation modes in the main, secondary and tertiary expansions.
- The rate of convergence depends on the ratio γ_M/γ_m, while there is an amplification factor that grows exponentially with the velocity
- The error estimates are independent of the reaction rate.

Calculus : Exponential convergence rate

We consider the time-dependent advection-diffusion-reaction equation in the domain $\mathcal{Q} = (0,1) \times (0,1)$ and we the source term and initial condition, given by

$$f(t,x) = \sqrt{|x - t - 0.3|}, \qquad T_0(x) = |x - 0.4|.$$

Strategy :

- The heat problem is discretized by Euler scheme/Gauss-Lobatto-Legendre spectral method
- Calculation for the matrix representations of the operators B and A are realized by means of accurate quadrature formulas
- Various integrals (with respect to either γ, α, $\vec{a}$ or (t,x)) are computed using Gauss-Lobatto quadrature formulas with high resolution in the corresponding intervals.

Convergence rate with respect to the parameters range.

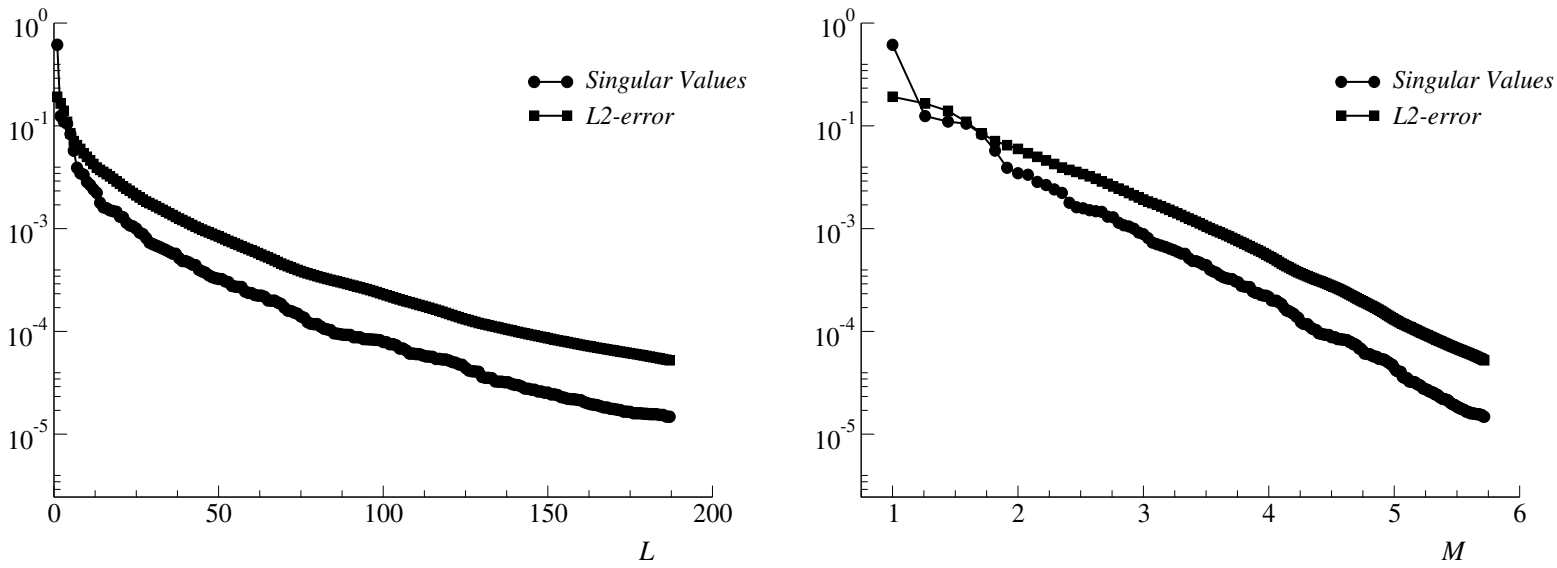


FIGURE 14 – Convergence history for POD expansion of the solution of the advection-diffusion-reaction equation. Left : with respect to L, right with respect to $M = L^{1/4}$

Convergence rate with respect to the parameters range.

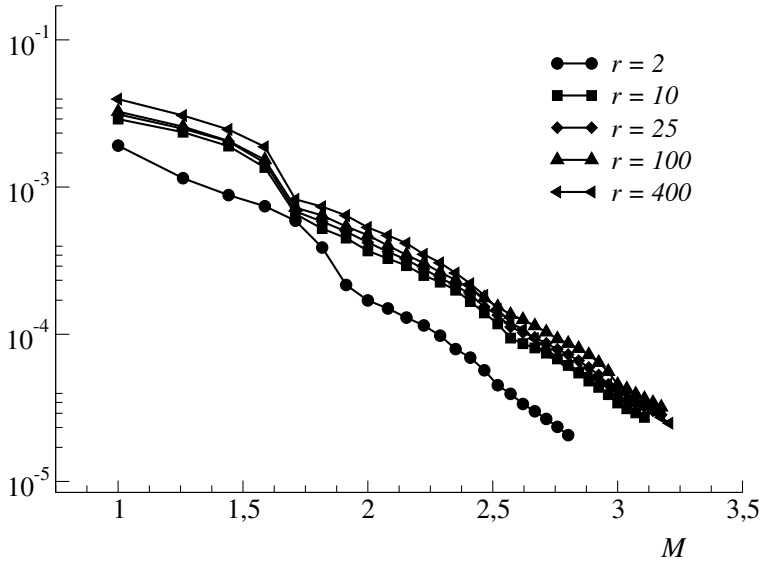


FIGURE 15 – Variation of the R-POD errors (in logarithmic scale) with respect to the ratio $r = \gamma_M/\gamma_m$, for fixed $[a_m, a_M] = [0, 5]$ and $\alpha_m = 0$, $\alpha_M = 10$. The variable M stands for $L^{1/4}$ where L is the number of modes.

Convergence rate with respect to the parameters range.

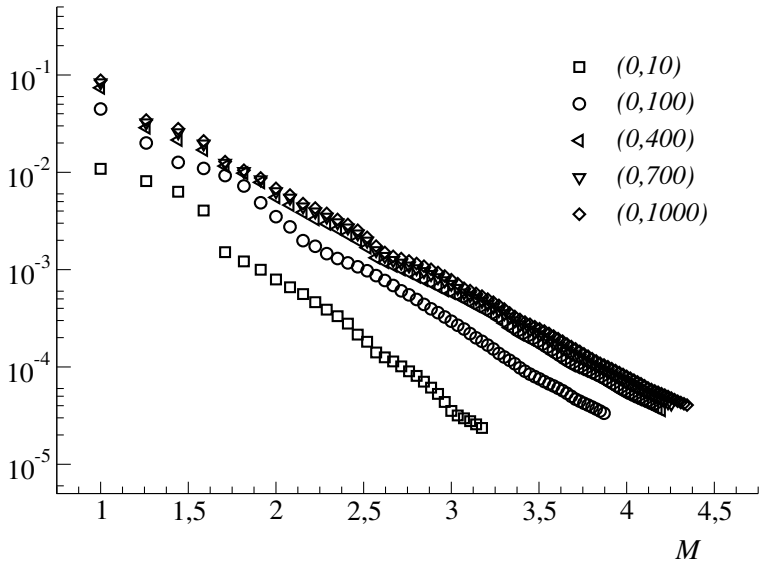


FIGURE 16 – Variation of the R-POD errors with respect to the reaction rate α. The curves correspond to $\alpha_0 = 0$, $\alpha_M = 1$, 25, 50, 75, 100 with $\gamma_m = 1$, $\gamma_M = 51$ and $a_m = 0$, $a_M = 20$ in all cases. The variable M stands for $L^{1/4}$ where L is the number of modes.

Convergence rate with respect to the parameters range.

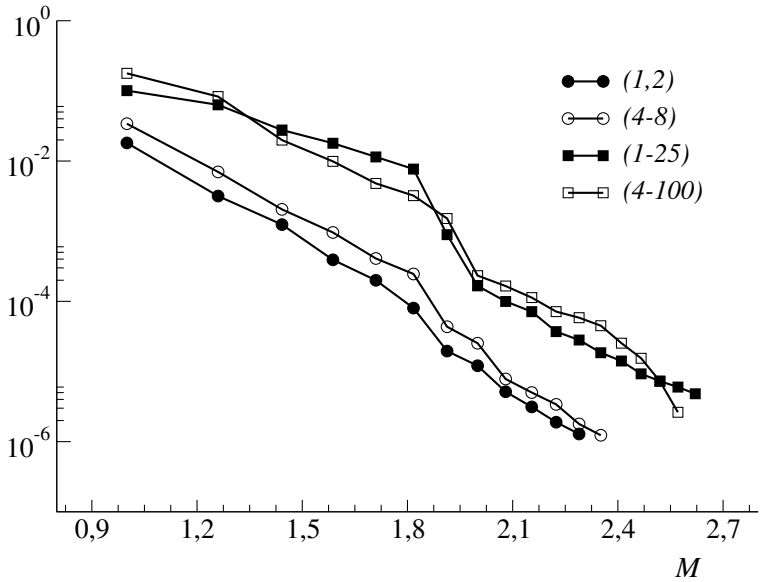


FIGURE 17 – Variation of the R-POD errors with respect to the ratio $r = \gamma_M/\gamma_m$, for fixed $\alpha_m = 0$, $\alpha_M = 10$ and $a_m = 0$ and $a_M = 10$. The variable M stands for $L^{1/4}$ where L is the number of modes.

Convergence rate with respect to the parameters range.

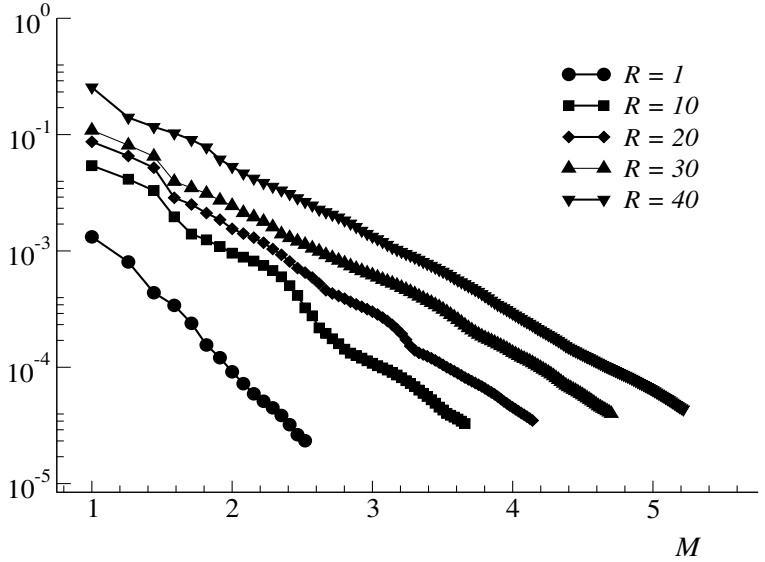
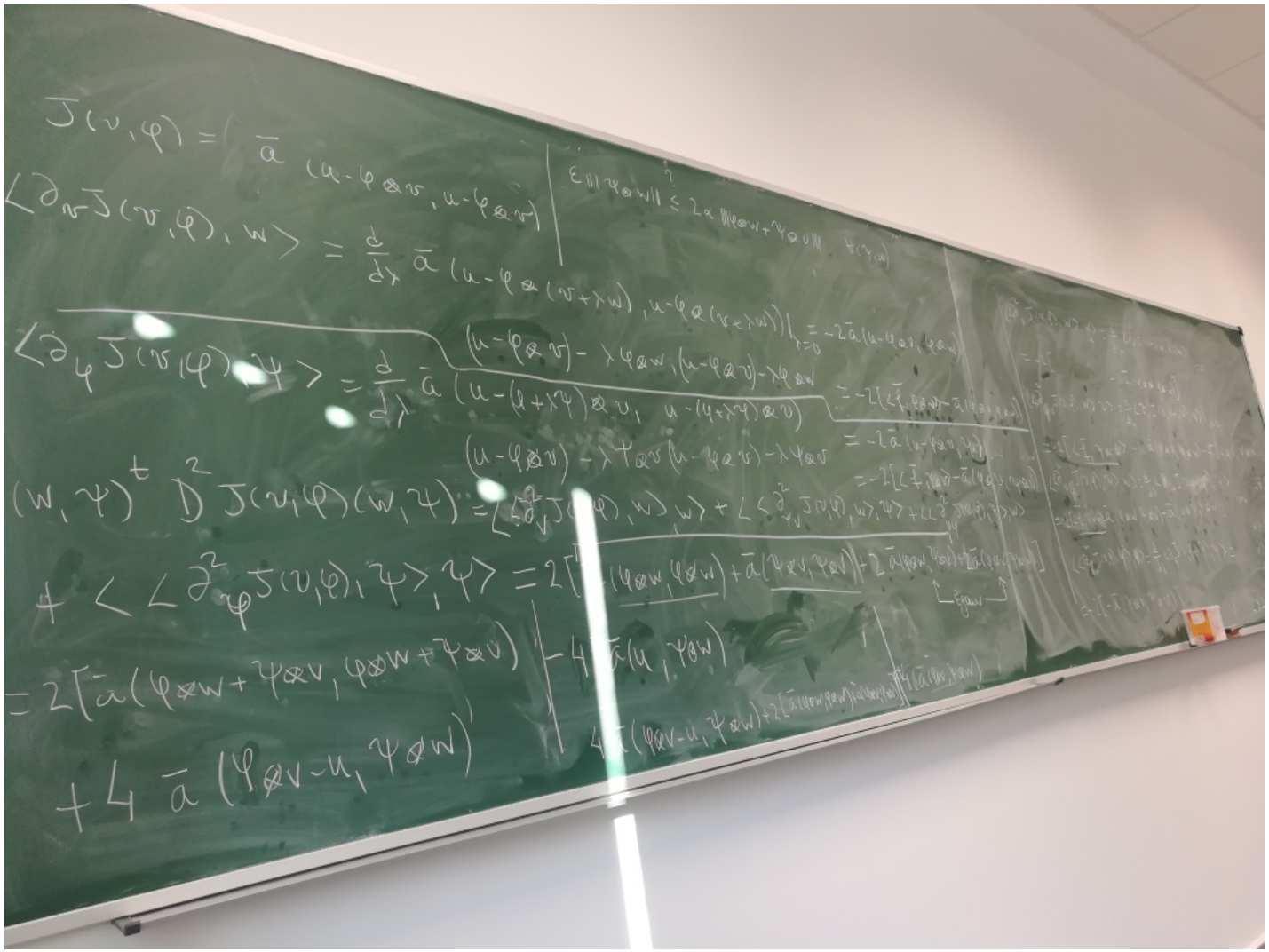


FIGURE 18 – Variation of the R-POD errors with respect to the advection velocity variation $R = a_M$. The curves correspond to $a_0 = 0$, $R = 1$, 10, 20, 30, 40 and 50, with $\gamma_m = 50$, $\gamma_M = 101$ and $\alpha_m = 0$, $\alpha_M = 10$ in all cases. The variable M stands for $L^{1/4}$ where L is the number of modes.

Table of Contents

Una imagen que lo explica todo !



Gracias

por su bienvenida

por su amabilidad conmigo

.... y sus sonrisas.

Y VIVA IMUS