ESAIM: M2AN 46 (2012) 731–757
DOI: 10.1051/m2an/2011053

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

GALERKIN APPROXIMATION WITH PROPER ORTHOGONAL DECOMPOSITION: NEW ERROR ESTIMATES AND ILLUSTRATIVE EXAMPLES

Dominique Chapelle[1], Asven Gariah[1] and Jacques Sainte-Marie[1,2]

Abstract. We propose a numerical analysis of proper orthogonal decomposition (POD) model reductions in which a priori error estimates are expressed in terms of the projection errors that are controlled in the construction of POD bases. These error estimates are derived for generic parabolic evolution PDEs, including with non-linear Lipschitz right-hand sides, and for wave-like equations. A specific projection continuity norm appears in the estimates and – whereas a general uniform continuity bound seems out of reach – we prove that such a bound holds in a variety of Galerkin bases choices. Furthermore, we directly numerically assess this bound – and the effectiveness of the POD approach altogether – for test problems of the type considered in the numerical analysis, and also for more complex equations. Namely, the numerical assessment includes a parabolic equation with super-linear reaction terms, inspired from the FitzHugh-Nagumo electrophysiology model, and a 3D biomechanical heart model. This shows that the effectiveness established for the simpler models is also achieved in the reduced-order simulation of these highly complex systems.

Mathematics Subject Classification. 65M60, 35A35, 35B45.

Received March 3, 2011. Revised June 30, 2011.
Published online February 3, 2012.

1. Introduction

In general, the simulation of partial differential equations resorts to discretization techniques such as finite differences, finite elements, or finite volumes. This typically results in discrete systems of large dimensions, hence the solution process can be rather costly, especially in situations when many computational iterations are required, as often occurs in design, control applications and inverse modeling.

In order to obtain reduced-order models, two main approaches are generally used. The first one consists in analyzing the dynamics operator of the system considered and retaining only the "most significant parts". Modal analysis (linear or non-linear normal modes), but also the moment matching method [2,7] and balanced truncation [11,26] belong to this first family. Unfortunately, for complex and large systems, these tools can be difficult to use in practice, since *e.g.* the eigenmodes are costly to obtain.

Keywords and phrases. POD, Galerkin approximation error estimates, non-linear parabolic problems, cardiac models.

[1] Inria Rocquencourt Laboratoire Saint-Venant Domaine de Voluceau Rocquencourt, B.P. 105, 78153 Le Chesnay, France. jacques.sainte-marie@inria.fr
[2] CETMEF, 2 boulevard Gambetta, 60200 Compiègne, France.

Article published by EDP Sciences

© EDP Sciences, SMAI 2012

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

The second strategy is more data-oriented in the sense that it mainly uses snapshots of the system to perform its reduction. The reduced basis [20, 21, 23, 24, 27] and the proper orthogonal decomposition (POD) [12, 15, 17, 19, 28] are two techniques belonging to this second family. This second approach consists in projecting the system onto subspaces of reduced sizes, albeit containing the major part of the expected dynamical solution. The aim is to obtain low-dimensional systems capturing the essence of the phenomena of interest.

Proper orthogonal decomposition, also known as Karhunen-Loève decomposition or principal component analysis, is a method initially introduced for analyzing multidimensional data. This method essentially provides an orthonormal basis for representing the given data in an optimal manner with respect to a quadratic criterion. The work in [16] has been pioneering in the development of the POD technique. In fluid mechanics, POD has been successfully used to access the coherent structures in turbulent flows [14], and it is now widely used in engineering in general.

Despite its relative simplicity of development and use, the POD technique has some limitations, since in particular it does not guarantee stability *e.g.* when parametric variations are considered [1]. Moreover, existing error estimates are expressed with respect to quantities which are not controlled in the construction of the POD basis [13]. This latter important issue is our primary concern here.

In this article, we propose new error estimates for the POD-based Galerkin approximation of the solutions of some classical and widely used PDE systems. First, we briefly recall the foundations of the POD decomposition. Then we derive the estimates for linear and non-linear parabolic equations, and also for linear hyperbolic systems. Finally, the theoretical results are confronted with numerical tests in various situations including a complex 3D biomechanical heart model.

2. Classical principles of POD reduction

In this section we briefly summarize the general principles and construction rules for POD reductions.

2.1. Construction of the POD basis

Let V be a separable Hilbert space with scalar product $\langle \cdot, \cdot \rangle$ and norm $\| \cdot \|$. Let $z(t)$, $t \in [0, T]$, be a function with regularity

$$z \in L^2(0, T; V).$$

We introduce $\mathrm{Cov} : V \to V$ the *covariance operator* defined by

$$\mathrm{Cov}\, \varphi = \int_0^T \langle z(t), \varphi \rangle z(t)\, \mathrm{d}t.$$

Performing the POD (time-continuous here) of rank l of z over $[0, T]$ consists in finding the orthogonal projector π^l of rank l solution of

$$\min_{\tilde{\pi}^l} \| z(t) - \tilde{\pi}^l z(t) \|_{L^2(0,T;V)}. \tag{2.1}$$

The integer l is called the *POD rank*. The time-discrete POD consists in solving the problem

$$\min_{\tilde{\pi}^l} \sum_i^N \| z(t_i) - \tilde{\pi}^l z(t_i) \|_V^2, \tag{2.2}$$

and for a detailed presentation of the solution of (2.2), see *e.g.* [18, 19] and references therein. In the continuous case, *i.e.* for the problem (2.1), the results and their proofs are, to some extent, similar but some aspects need to be specified. The four following propositions are the cornerstones of the solution of (2.1). The complete proofs of these propositions rely on straightforward adaptations of results contained in [18].

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Proposition 2.1. *There exists a unique sequence* $(\lambda_i)_{i\in I}$ *with* I *finite or countable, such that*

$$\begin{gathered}\lambda_i > 0, \quad \forall i \in I,\\ \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N, \quad \textit{if } I \textit{ finite} \quad (I = \{1, 2, \ldots, N\}),\\ \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_i \geq \ldots, \quad \lambda_i \underset{i\to\infty}{\longrightarrow} 0, \quad \textit{otherwise} \quad (I = \mathbb{N}^*),\end{gathered}$$

and an orthonormal sequence $(\varphi_i)_{i\in I}$ *of* V *satisfying*

$$\operatorname{Cov}\varphi_i = \lambda_i\varphi_i, \quad \forall i \in I,$$

such that $(\varphi_i)_{i\in I}$ *is total in the orthogonal complement of the kernel of* Cov, *i.e.*

$$V = \operatorname{Ker}\operatorname{Cov} \overset{\perp}{\oplus} \overline{\operatorname{Span}\{\varphi_i,\ i \in I\}}. \tag{2.3}$$

To understand (2.3) it is helpful to characterize the kernel of Cov with respect to z.

Proposition 2.2. *The kernel of* Cov *is made of the vectors that are orthogonal to* $z(t)$ *for almost every* $t \in [0, T]$, *i.e.*

$$\operatorname{Ker}\operatorname{Cov} = \{\varphi\ ;\ \langle z(t), \varphi\rangle = 0 \text{ a.e. } t \in [0, T]\}.$$

Then we have the classical result.

Proposition 2.3. *For all* $1 \leq l \leq \operatorname{Card} I$, *a solution* π^l *of problem* (2.1) *is determined by*

$$V^l = \operatorname{Im}\pi^l = \operatorname{Span}(\varphi_1, \ldots, \varphi_l).$$

Moreover,

$$\|z - \pi^l z\|_{L^2(0,T;V)} = \left\{\sum_{i=l+1}^{\operatorname{Card} I} \lambda_i\right\}^{1/2}. \tag{2.4}$$

The POD reduction is interesting when the sequence of eigenvalues λ_i tends rapidly to zero. Indeed, for small values of l we then have the approximation

$$z(t) \approx \sum_{i=1}^{l} \langle z(t), \varphi_i\rangle\varphi_i.$$

In practice, the dimension of V can be very large and it is costly to use the operator Cov. So it is convenient to introduce $\widetilde{\operatorname{Cov}}$ the $L^2(0, T) \to L^2(0, T)$ operator defined by

$$\widetilde{\operatorname{Cov}}\, v\,(s) = \int_0^T \langle z(t), z(s)\rangle v(t)\,\mathrm{d}t = \left\langle \int_0^T v(t)z(t)\,\mathrm{d}t,\ z(s)\right\rangle,$$

for a.e. $s \in [0, T]$. We have the following proposition.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Proposition 2.4. Cov *and* $\widetilde{\text{Cov}}$ *share the same non-zero eigenvalues, with identical multiplicities. Moreover* $\widetilde{\text{Cov}}$ *is compact.*

Thus there exists an orthonormal sequence $(v_i)_{i\in I}$ of $L^2(0,T)$ eigenvectors of $\widetilde{\text{Cov}}$, in finite number for each non-zero eigenvalue,

$$\widetilde{\text{Cov}}\, v_i = \lambda_i v_i, \quad \forall i \in I,$$

such that $(v_i)_{i\in I}$ is total in the orthogonal complement of the kernel of $\widetilde{\text{Cov}}$, *i.e.*

$$L^2(0,T) = \text{Ker}\,\widetilde{\text{Cov}} \overset{\perp}{\oplus} \overline{\text{Span}\{v_i,\ i \in I\}}.$$

We define $\forall i \in I$ the V element

$$\varphi_i = \frac{1}{\sqrt{\lambda_i}} \int_0^T v_i(t) z(t)\, \mathrm{d}t.$$

Then, $(\varphi_i)_{i\in I}$ is an orthonormal sequence of eigenvectors of Cov, with the same sequence of corresponding eigenvalues

$$\text{Cov}\,\varphi_i = \lambda_i \varphi_i, \quad \forall i \in I.$$

2.2. Reduced-order modeling

Considering $z(x,t)$ the solution of a PDE problem, the POD-based reduced order modeling, or more simply POD reduction, consists in building a spatial Galerkin approximation $z^l(x,t)$ of $z(x,t)$ in the POD space $V^l = \text{Span}(\varphi_1, \ldots, \varphi_l)$. Then the key point is to be able to control the *reduction error*, namely

$$\|z - z^l\|_{L^2(0,T;V)}.$$

We tackle this problem in the following section.

3. New estimates for the POD reduction error

In this section, our objective is to derive POD-reduction error estimates bounded by approximation terms which can be conveniently controlled in the construction of the POD basis.

For the sake of generality and homogeneity with the existing literature, we introduce the classical abstract mathematical framework. Nevertheless, to fix the ideas the reader can keep in mind that in the examples considered, the abstract spaces H and V will typically correspond to $L^2(\Omega)$ and $H_0^1(\Omega)$, respectively.

Let $(V,\ ((\cdot,\cdot)),\ \|\cdot\|)$ and $(H,\ (\cdot,\cdot),\ |\cdot|)$ be two separable Hilbert spaces with continuous and dense embedding $V \hookrightarrow H$, *i.e.*

$$|v| \le C_\Omega \|v\|, \quad \forall v \in V.$$

We choose H as the pivot space – namely, we perform the identification of H with its dual space H' – and then

$$V \hookrightarrow H \hookrightarrow V'.$$

Let a be a symmetric bilinear form on V, continuous, and coercive, namely,

$$a(v,w) \le C_a \|v\|\|w\|, \quad \forall v, w \in V,$$
$$a(v,v) \ge c_a \|v\|^2, \quad \forall v \in V.$$

Then a also defines a scalar product on V and we denote by $\|\cdot\|_a$ the associated norm.

We point out that in our estimations we use C to denote a generic positive constant, independent of all discretization parameters, and that may take different values at various occurrences, including in the same equation.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

3.1. Galerkin estimates for linear parabolic problems with H-orthogonal projectors

We formally introduce the *abstract parabolic equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}(u(t),v)+a(u(t),v)=(f(t),v),\quad \forall v\in V, \tag{3.1}$$

$$u(0)=u_0. \tag{3.2}$$

Equation (3.1) is to be understood in the sense of distributions in time. We then have the following existence and uniqueness result, see [8], XVIII, Section 3.2, Theorem 1 and Section 3.3, Theorem 2.

Proposition 3.1. *Assume $f\in L^2(0,T;H)$ and $u_0\in H$. Then there exists a unique solution u of equations (3.1)–(3.2) such that*

$$u\in L^2(0,T;V)\cap C([0,T];H),\quad \frac{\mathrm{d}u}{\mathrm{d}t}\in L^2(0,T;V').$$

Considering now a finite-dimensional subspace V^l, we formally introduce the spatial Galerkin approximation u^l of u

$$u^l(t)\in V^l, \tag{3.3}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(u^l(t),v^l)+a(u^l(t),v^l)=(f(t),v^l),\quad \forall v^l\in V^l, \tag{3.4}$$

$$u^l(0)=u_0^l. \tag{3.5}$$

Since V^l is a finite-dimensional space, it is easy to prove the following result [8], XVIII, Section 3.3.1, Lemma 1.

Proposition 3.2. *Assume $f\in L^2(0,T;H)$ and $u_0^l\in V^l$. Then there exists a unique solution u^l of equations (3.4)–(3.5) such that*

$$u^l\in C([0,T];V^l),\quad \frac{\mathrm{d}u^l}{\mathrm{d}t}\in L^2(0,T;V^l).$$

Note that we have more regularity here than for the continuous solution only because we are considering a finite dimensional problem, and of course the corresponding estimates are not uniform with respect to the discretization.

Finally, let π_H^l and π_V^l respectively denote the H-orthogonal and V-orthogonal projectors of V onto the reduction space V^l. For all $v\in V$

$$|v-\pi_H^l v|=\inf_{v^l\in V^l}|v-v^l|,$$
$$\|v-\pi_V^l v\|=\inf_{v^l\in V^l}\|v-v^l\|.$$

With a view to estimating the *reduction error* $\|u-u^l\|_{L^2(0,T;V)}$, classical error estimates are of the form [8]

$$\|u-u^l\|_{L^2(0,T;V)}\le C\left(\|u-\pi_V^l u\|_{L^2(0,T;V)}+\left\|\frac{\partial}{\partial t}(u-\pi_V^l u)\right\|_{L^2(0,T;V)}+|u_0^l-\pi_V^l u_0|\right). \tag{3.6}$$

However, the POD criterion (2.1) does not provide a direct control on the time-derivative term in the right-hand side, and our objective is to circumvent this difficulty. To that end we use the *H-projection error*, still in the same $L^2(0,T;V)$-norm, *i.e.*

$$\|u-\pi_H^l u\|_{L^2(0,T;V)}.$$

and the following result holds.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Proposition 3.3. *For all $T > 0$,*

$$\|u - u^l\|_{L^2(0,T;V)} \leq C(|\pi_H^l u_0 - u_0^l| + \|u - \pi_H^l u\|_{L^2(0,T;V)}).$$

Proof. We split $u - u^l$ into two parts

$$u - u^l = p^l + q^l,$$

where $p^l = u - \pi_H^l u$ and $q^l = \pi_H^l u - u^l$. Since $q^l \in V^l$, and using the definition of u^l, q^l satisfies the variational equation

$$\frac{\mathrm{d}}{\mathrm{d}t}(q^l(t), v^l) + a(q^l(t), v^l) = -(f(t), v^l) + \frac{\mathrm{d}}{\mathrm{d}t}(\pi_H^l u(t), v^l) + a(\pi_H^l u(t), v^l), \quad \forall v^l \in V^l.$$

The projection π_H^l satisfies $(\pi_H^l u(t), v^l) = (u(t), v^l)$, so that using the definition of u we get

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} q^l(t), v^l\right) + a(q^l(t), v^l) = -a(p^l(t), v^l).$$

Taking $v^l = q^l(t)$, we then obtain the energy estimate

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\{|q^l|^2\}(t) + \|q^l(t)\|_a^2 = -a(p^l(t), q^l(t)),$$

which we now integrate on $[0, T]$ to obtain,

$$\int_0^T \|q^l(t)\|_a^2 \,\mathrm{d}t \leq \frac{1}{2}|q^l(0)|^2 + \left|\int_0^T a(p^l(t), q^l(t)) \,\mathrm{d}t\right|,$$

where we dropped the term in $|q^l(T)|^2$ in the left-hand side. Hence, by Young's inequality,

$$\int_0^T \|q^l(t)\|_a^2 \,\mathrm{d}t \leq |q^l(0)|^2 + \int_0^T \|p^l(t)\|_a^2 \,\mathrm{d}t.$$

This directly entails

$$\|q^l\|_{L^2(0,T;V)} \leq C\left(|q^l(0)|^2 + \|p^l\|_{L^2(0,T;V)}\right),$$

using the properties of the scalar product a, and the triangle inequality

$$\|u(t) - u^l(t)\| \leq \|p^l(t)\| + \|q^l(t)\|$$

concludes the proof. □

Therefore, *via* the introduction of π_H^l we avoid the time derivative appearing in the right-hand side of the more standard estimate (3.6). However, we now need to deal with the approximation term $\|u - \pi_H^l u\|_{L^2(0,T;V)}$, which is the topic of the next section.

3.2. Galerkin estimates for linear parabolic problems with V-orthogonal projectors

Note first that, since V^l is finite-dimensional, we have an inverse inequality of the form

$$\exists \alpha^l > 0, \quad \forall v^l \in V^l, \quad \|v^l\| \leq \alpha^l |v^l|.$$

Hence, π_H^l is continuous as an endomorphism of V, as can be seen by directly writing

$$\|\pi_H^l v\| \leq \alpha^l |\pi_H^l v| \leq \alpha^l |v| \leq C\alpha^l \|v\|.$$

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

However, as inverse inequality constants blow up when the dimension of the V^l subspace increases, we will obtain some better insight by using the V-projection as follows

$$\begin{aligned}\|\pi_H^l v\| &\leq \|\pi_H^l v - \pi_V^l v\| + \|\pi_V^l v\| \\ &\leq \alpha^l |\pi_H^l v - \pi_V^l v| + \|v\| = \alpha^l |\pi_H^l (v - \pi_V^l v)| + \|v\| \\ &\leq \alpha^l |v - \pi_V^l v| + \|v\|.\end{aligned}$$

Denoting by $\mathcal{L}(V)$ the space of V-endomorphisms and by $\mathcal{L}(V,H)$ the space of linear operators from V into H, this entails

$$\|\pi_H^l\|_{\mathcal{L}(V)} \leq 1 + \alpha^l \|\operatorname{Id} - \pi_V^l\|_{\mathcal{L}(V,H)},$$

where Id is the identity operator and the inverse inequality constant is now multiplied by a projection error term which can be conjectured to vanish in various cases when increasing l, since V is more regular than H. Hence, we can transform any estimate with $\|v - \pi_H^l v\|$ into an estimate with $\|v - \pi_V^l v\|$. Indeed, as π_H^l and π_V^l project onto the same subspace we have

$$\begin{aligned}v - \pi_H^l v &= (\operatorname{Id} - \pi_H^l)(v - \pi_V^l v) \\ &= (\operatorname{Id} - (\pi_H^l - \pi_V^l))(v - \pi_V^l v).\end{aligned}$$

Since $(\pi_H^l - \pi_V^l)v = \pi_H^l(v - \pi_V^l v)$ for all $v \in V$, we remark that

$$(\pi_H^l - \pi_V^l)v = \begin{cases} \pi_H^l v & \text{if } v \in (V^l)^\perp, \\ 0 & \text{if } v \in V^l, \end{cases}$$

denoting by $(V^l)^\perp$ the V-orthogonal complement of V^l. Hence,

$$\|\pi_H^l - \pi_V^l\|_{\mathcal{L}(V)} \leq \|\pi_H^l\|_{\mathcal{L}(V)}.$$

Defining

$$\rho_l = \|\pi_H^l\|_{\mathcal{L}(V)}, \quad \sigma_l = \|\pi_H^l - \pi_V^l\|_{\mathcal{L}(V)},$$

we can then convert the projector by writing

$$\|v - \pi_H^l v\| \leq (1 + \sigma_l)\, \|v - \pi_V^l v\| \tag{3.7}$$

$$\leq (1 + \rho_l)\, \|v - \pi_V^l v\|, \tag{3.8}$$

from which we directly infer the following estimate.

Corollary 3.4. *For all $T > 0$,*

$$\|u - u^l\|_{L^2(0,T;V)} \leq C \left(|\pi_H^l u_0 - u_0^l| + (1 + \sigma_l) \|u - \pi_V^l u\|_{L^2(0,T;V)} \right).$$

However, we do not formally assert that, for a general reduction space, neither ρ_l nor σ_l have a bounded behavior with respect to l. This behavior is likely to be dependent on the specific types of variational problem and Galerkin reduction considered, and can be numerical assessed when no analytical treatment is at hand.

Note that the last term in the right-hand side of (3.4) is the quantity that is in fact minimized in the construction of POD subspaces. Furthermore, for POD reduction subspaces the sequences (σ_l) and (ρ_l) remain bounded in a large class of situations, see Section 3.5 for some theoretical insight.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

3.3. Extension to a non-linear parabolic equation

We formally introduce the *abstract non-linear parabolic equation*

$$\frac{\mathrm{d}}{\mathrm{d}t}(u(t),v)+a(u(t),v)=(f(t,u(t)),v),\quad \forall v\in V, \tag{3.9}$$

$$u(0)=u_0. \tag{3.10}$$

Unlike in equation (3.1), f is some $[0,T]\times V\to V$ function. The general theory is very delicate. Especially the solution may explode in finite time. We provide the following proposition, where we assume that f is Lipschitz-continuous in the second variable. As proven in the appendix, this guarantees, for any $T>0$, the well-posedness of equations (3.9)–(3.10) in the same spaces as in the linear case.

Proposition 3.5. *Assume $u_0\in H$, $f\in C([0,T]\times H;H)$, and that f is L-Lipschitz continuous in its second variable, i.e. that there exists a constant L such that*

$$\forall t\in[0,T],\quad \forall h_1,h_2\in H,\quad |f(t,h_1)-f(t,h_2)|\le L|h_1-h_2|.$$

Assume also that the embedding $V\hookrightarrow H$ is compact. Then there exists a unique solution u of equations (3.9)–(3.10) such that

$$u\in L^2(0,T;V)\cap C([0,T];H),\quad \frac{\mathrm{d}u}{\mathrm{d}t}\in L^2(0,T;V').$$

Let now u^l be the spatial Galerkin approximation of u in V^l

$$u^l(t)\in V^l, \tag{3.11}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(u^l(t),v^l)+a(u^l(t),v^l)=(f(t,u^l(t)),v^l),\quad \forall v^l\in V^l, \tag{3.12}$$

$$u^l(0)=u_0^l. \tag{3.13}$$

More simply, with the Peano existence theorem, we obtain the following result.

Proposition 3.6. *Assume $u_0^l\in V^l$, $f\in C([0,T]\times H;H)$ and f is L-Lipschitz continuous in its second variable. Then there exists a unique solution u^l of (3.11)–(3.13) such that*

$$u^l\in C^1([0,T];V^l).$$

The proof of this result can also be seen as contained in that of Proposition 3.5, proven in the appendix.

We now show the following result for the reduction error.

Proposition 3.7. *For all $T>0$,*

$$\|u-u^l\|_{L^2(0,T;V)}\le C_1(L,T)\big(|\pi_H^l u_0-u_0^l|+C_2(L)\|u-\pi_H^l u\|_{L^2(0,T;V)}\big), \tag{3.14}$$

where, for all $L>0$,

$$C_1(L,T)=C\mathrm{e}^{LT},\quad C_2(L)=C(L+1).$$

In addition, we have

$$\|u-u^l\|_{L^2(0,T;V)}\le C_1(L,T)\big(|\pi_H^l u_0-u_0^l|+(1+\sigma_l)C_2(L)\|u-\pi_V^l u\|_{L^2(0,T;V)}\big). \tag{3.15}$$

Moreover, under the condition $L<\frac{c_a}{C_\Omega^2}$, we have the improved constants

$$C_1(L,T)=C_1(L)=\frac{C}{\sqrt{\frac{c_a}{C_\Omega^2}-L}},\quad C_2(L)=\frac{C}{\sqrt{\frac{c_a}{C_\Omega^2}-L}},$$

where C_1 is now independent of T.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Proof. We split $u - u^l$ into two parts

$$u - u^l = p^l + q^l,$$

where $p^l = u - \pi_H^l u$ and $q^l = \pi_H^l u - u^l$. Using the same property of π_H^l as in the proof of Proposition 3.3, we obtain

$$\left(\frac{\mathrm{d}q^l}{\mathrm{d}t}(t), v^l\right) + a(q^l(t), v^l) = -a(p^l(t), v^l) + (f(t, u(t)) - f(t, u^l(t)), v^l).$$

Taking $v^l = q^l(t)$, and integrating on $[0, t]$, $0 \le t \le T$, we obtain

$$\begin{aligned}
\frac{1}{2}|q^l(t)|^2 + c_a\|q^l\|^2_{L^2(0,t;V)} &\le \frac{1}{2}|q^l(0)|^2 + C_a\|p^l\|_{L^2(0,t;V)}\|q^l\|_{L^2(0,t;V)} \\
&\quad + L\int_0^t |u(s) - u^l(s)| \cdot |q^l(s)|\,\mathrm{d}s \\
&\le \frac{1}{2}|q^l(0)|^2 + C_a\|p^l\|_{L^2(0,t;V)}\|q^l\|_{L^2(0,t;V)} \\
&\quad + L\int_0^t (|p^l(s)| + |q^l(s)|) \cdot |q^l(s)|\,\mathrm{d}s \\
&\le \frac{1}{2}|q^l(0)|^2 + (C_a + LC_\Omega^2)\|p^l\|_{L^2(0,t;V)}\|q^l\|_{L^2(0,t;V)} \\
&\quad + L\|q^l\|^2_{L^2(0,t;H)}.
\end{aligned} \tag{3.16}$$

Let us first assume $L < \frac{c_a}{C_\Omega^2}$. Hence, for $t = T$ and using the continuous embedding $V \hookrightarrow H$, Equation (3.16) entails

$$(c_a - LC_\Omega^2)\|q^l\|^2_{L^2(0,T;V)} \le \frac{1}{2}|q^l(0)|^2 + (C_a + c_a)\|p^l\|_{L^2(0,T;V)}\|q^l\|_{L^2(0,T;V)}.$$

Using Young's inequality, we can then conclude as in Proposition 3.3.

Let us now consider the general case for L. By Young's inequality on (3.16),

$$|q^l(t)|^2 + c_a\|q^l\|^2_{L^2(0,t;V)} \le |q^l(0)|^2 + \frac{(C_a + LC_\Omega^2)^2}{c_a}\|p^l\|^2_{L^2(0,T;V)} + 2L\|q^l\|^2_{L^2(0,t;H)}. \tag{3.17}$$

Then, we use Gronwall's inequality for $t \mapsto |q^l(t)|^2$, which leads to

$$|q^l(t)|^2 \le \mathrm{e}^{2Lt}\left(|q^l(0)|^2 + \frac{(C_a + LC_\Omega^2)^2}{c_a}\|p^l\|^2_{L^2(0,T;V)}\right).$$

Finally, re-incorporating this estimate in (3.17) gives, for $t = T$,

$$\|q^l\|^2_{L^2(0,T;V)} \le \frac{\mathrm{e}^{2LT}}{c_a}\left(|q^l(0)|^2 + \frac{(C_a + LC_\Omega^2)^2}{c_a}\|p^l\|^2_{L^2(0,T;V)}\right),$$

and we conclude for (3.14) as in Proposition 3.3.

Of course (3.15) directly follows like in Corollary 3.4. □

3.4. Galerkin estimates for the wave-like equation

Let us now consider the wave-like equation

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(y(t), v) + a(y(t), v) = (f(t), v), \quad \forall v \in V, \tag{3.18}$$

$$y(0) = y_0, \quad \frac{\mathrm{d}y}{\mathrm{d}t}(0) = \dot{y}_0, \tag{3.19}$$

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

for which we have the classical existence and uniqueness results, see [8], XVIII, Section 5.5.2, Theorem 1 and Section 5.5.3, Theorem 2.

Proposition 3.8. *We assume $f \in L^2(0,T;H)$, $y_0 \in V$ and $\dot{y}_0 \in H$. Then there exists a unique solution y of equations (3.18)–(3.19) such that*

$$y \in C([0,T];V), \quad \frac{\mathrm{d}y}{\mathrm{d}t} \in C([0,T];H), \quad \frac{\mathrm{d}^2 y}{\mathrm{d}t^2} \in L^2(0,T;V').$$

As in Section 3.1, we formally introduce the spatial Galerkin approximation y^l of y

$$y^l(t) \in V^l, \tag{3.20}$$

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(y^l(t), v^l) + a(y^l(t), v^l) = (f(t), v^l), \quad \forall v^l \in V^l, \tag{3.21}$$

$$y^l(0) = y_0^l, \quad \frac{\mathrm{d}y^l}{\mathrm{d}t}(0) = \dot{y}_0^l. \tag{3.22}$$

And the following holds, see [8], XVIII, Section 5.3.1, Lemma 2.

Proposition 3.9. *We assume $f \in L^2(0,T;H)$, $y_0 \in V^l$ and $\dot{y}_0 \in V^l$. Then there exists a unique solution y^l of equations (3.20)–(3.22) such that*

$$y^l \in C^1([0,T];V^l), \quad \frac{\mathrm{d}^2 y^l}{\mathrm{d}t^2} \in L^2(0,T;V^l).$$

The error estimate between the solutions of problems (3.18)–(3.19) and (3.20)–(3.22) is given by the following proposition.

Proposition 3.10. *For all $T > 0$,*

$$\begin{aligned} \|y - y^l\|_{L^2(0,T;V)} + \left\| \frac{\mathrm{d}}{\mathrm{d}t}(y - y^l) \right\|_{L^2(0,T;H)} \leq C \Bigg\{ & \sqrt{T} \left(\|y_0 - \pi_H^l y_0\| + \|\pi_H^l y_0 - y_0^l\| + |\pi_H^l \dot{y}_0 - \dot{y}_0^l| \right) \\ & + \|y - \pi_H^l y\|_{L^2(0,T;V)} + (1+T) \left\| \frac{\mathrm{d}}{\mathrm{d}t} \left(y - \pi_H^l y \right) \right\|_{L^2(0,T;V)} \Bigg\}, \end{aligned} \tag{3.23}$$

and

$$\begin{aligned} \|y - y^l\|_{L^2(0,T;V)} + \left\| \frac{\mathrm{d}}{\mathrm{d}t}(y - y^l) \right\|_{L^2(0,T;H)} \leq C \Bigg\{ & \sqrt{T} (\|y_0 - \pi_H^l y_0\| + \|\pi_H^l y_0 - y_0^l\| + |\pi_H^l \dot{y}_0 - \dot{y}_0^l|) \\ & + (1+\sigma_l) \left(\|y - \pi_V^l y\|_{L^2(0,T;V)} + (1+T) \left\| \frac{\mathrm{d}}{\mathrm{d}t}(y - \pi_V^l y) \right\|_{L^2(0,T;V)} \right) \Bigg\}. \end{aligned} \tag{3.24}$$

Proof. We split $y - y^l$ into two parts

$$y - y^l = p^l + q^l,$$

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

where $p^l = y - \pi_H^l y$ and $q^l = \pi_H^l y - y^l$. Since $q^l \in V^l$, and using the definition of y^l, q^l verifies the variational equation

$$\frac{\mathrm{d}^2}{\mathrm{d}t^2}(q^l(t), v^l) + a(q^l(t), v^l) = -(f(t), v^l) + \frac{\mathrm{d}^2}{\mathrm{d}t^2}(\pi_H^l y(t), v^l) + a(\pi_H^l y(t), v^l), \quad \forall v^l \in V^l.$$

The projector π_H^l verifies $(\pi_H^l y(t), v^l) = (y(t), v^l)$, so that using the definition of y we get

$$\left(\frac{\mathrm{d}^2 q^l}{\mathrm{d}t^2}(t), v^l\right) + a(q^l(t), v^l) = -a(p^l(t), v^l).$$

We infer the energy balance by taking $v^l = \frac{\mathrm{d}q^l}{\mathrm{d}t}(t)$, *viz.*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\left\{\left|\frac{\mathrm{d}q^l}{\mathrm{d}t}\right|^2 + \|q^l\|_a^2\right\}(t) = -a\left(p^l(t), \frac{\mathrm{d}q^l}{\mathrm{d}t}(t)\right).$$

Performing an integration by parts over time and using Young's inequality in the right-hand side, we have

$$\left|\frac{\mathrm{d}q^l}{\mathrm{d}t}(t)\right|^2 + (1-\eta)\|q^l(t)\|_a^2 \le \left|\frac{\mathrm{d}q^l}{\mathrm{d}t}(0)\right|^2 + \|q^l(0)\|_a^2 + 2a(p^l(0), q^l(0)) + \frac{1}{\eta}\|p^l(t)\|_a^2 + \theta\|q^l\|_{L^2(0,T;a)}^2 + \frac{1}{\theta}\left\|\frac{\mathrm{d}p^l}{\mathrm{d}t}\right\|_{L^2(0,T;a)}^2.$$

By integration on $[0,T]$ again and taking $\eta = \frac{1}{4}$, $\theta = \frac{1}{4T}$, we get

$$\left\|\frac{\mathrm{d}q^l}{\mathrm{d}t}\right\|_{L^2(0,T;H)}^2 + \|q^l\|_{L^2(0,T;a)}^2 \le C\left\{T\left(\|p^l(0)\|_a^2 + \|q^l(0)\|_a^2 + \left|\frac{\mathrm{d}q^l}{\mathrm{d}t}(0)\right|^2\right) + \|p^l\|_{L^2(0,T;a)}^2 + T^2\left\|\frac{\mathrm{d}p^l}{\mathrm{d}t}\right\|_{L^2(0,T;a)}^2\right\}.$$

Using the properties of the scalar product a, we get back to the $\|\cdot\|$ norm, and the triangular inequality

$$\|y(t) - y^l(t)\| \le \|p^l(t)\| + \|q^l(t)\|$$

ends the proof for (3.23), whence (3.24) directly follows. □

3.5. Boundedness of (σ^l)

As already mentioned, we need some characterization of the behavior of the sequences $(\rho_l)_{l\ge1}$ and $(\sigma_l)_{l\ge1}$ in order for the above estimations to be meaningful. To provide some insight into this issue we give some examples of reduction subspaces for which these sequences can be proven to be bounded.

Let us start by showing this boundedness when the Galerkin subspace is given by finite element discretization procedures. To fix the ideas we consider a standard $\mathbf{P}_1$ discretization, but this result can be extended with ease to most other finite element procedures.

Proposition 3.11. *Let Ω be an open convex polyhedral subset of $\mathbb{R}^2$ or $\mathbb{R}^3$. Let $H = L^2(\Omega)$ and $V = H_0^1(\Omega)$. Let $(\mathcal{T}_h)_{h>0}$ be a quasi-uniform family of triangulations of Ω, and V_h the $\mathbf{P}_1$-Lagrange finite element subspace of V built on $\mathcal{T}_h$, with $\pi_{h,H}$ the H-orthogonal projector onto V_h. Then $(\pi_{h,H})_{h>0}$ is bounded in $\mathcal{L}(V)$, i.e.*

$$\forall h > 0, \quad \forall v \in V, \quad \|\pi_{h,H} v\| \le C\|v\|. \tag{3.25}$$

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Proof. Let us introduce a family of Clément interpolation operators $(\mathcal{C}_h)_{h>0}$ associated with $(\mathcal{T}_h)_{h>0}$ and uniformly bounded from V to V_h [6]. Since $(\mathcal{T}_h)_{h>0}$ is quasi-uniform, an inverse inequality holds, see [5], Theorem 3.2.6, so that

$$\begin{aligned}\|\pi_{h,H}v\| &\leq \|\pi_{h,H}v - \mathcal{C}_h v\| + \|\mathcal{C}_h v\| \\ &\leq Ch^{-1}|\pi_{h,H}v - \mathcal{C}_h v| + \|\mathcal{C}_h v\|.\end{aligned}$$

Remark that, by the characterization of an orthogonal projector,

$$|\pi_{h,H}v - \mathcal{C}_h v| \leq |v - \pi_{h,H}v| + |v - \mathcal{C}_h v| \leq 2|v - \mathcal{C}_h v|. \tag{3.26}$$

Now we use the property

$$\forall h > 0, \quad \forall v \in V, |v - \mathcal{C}_h v| \leq Ch\|v\|, \tag{3.27}$$

and the boundedness of $(\mathcal{C}_h)_{h>0}$ in $\mathcal{L}(V)$ [6], Theorem 2. This shows our result. □

We now consider spectral analysis, namely, taking Galerkin subspaces provided by the eigenmodes of the bilinear form a. We thus assume the embedding $V \hookrightarrow H$ to be compact, which is satisfied when Ω is bounded, $H = L^2(\Omega)$ and $V = H_0^1(\Omega)$. Then there exists a Hilbertian basis of H, (w_i), characterized by

$$\begin{aligned}a(w_i, v) &= \omega_i^2 (w_i, v), \\ 0 < \omega_1 \leq \omega_2 \leq \ldots, \quad \omega_i &\underset{i\to\infty}{\longrightarrow} +\infty.\end{aligned}$$

Introducing $\tilde{w}_i = \frac{1}{\omega_i} w_i$, $(\tilde{w}_i)_{i\geq 1}$ is a Hilbertian basis of V for the scalar product associated with a.

Proposition 3.12. *Assuming $V^l = \mathrm{Span}(w_1, \ldots, w_l)$, the sequences (ρ_l) and (σ_l) are bounded.*

Proof. We remark

$$a(v, \tilde{w}_i)\tilde{w}_i = (v, w_i)w_i.$$

Summing this identity from 1 to l directly entails that $\pi_H^l = \pi_a^l$, where π_a^l is the a-orthogonal projector of V onto V^l. Moreover

$$c_a^{1/2}\|\pi_a^l v\| \leq \|\pi_a^l v\|_a \leq \|v\|_a \leq C_a^{1/2}\|v\|,$$

which leads to

$$\|\pi_a^l\|_{\mathcal{L}(V)} \leq \left(\frac{C_a}{c_a}\right)^{1/2}$$

and the property $\pi_H^l = \pi_a^l$ concludes the proof. □

As a third example, we will consider the case of the POD subspaces arising from the analysis of the homogeneous wave-like equation. The following result is very straightforward to establish by decomposing the solution on the eigenmodes. We also refer to [9] for related discussions.

Proposition 3.13. *Let y be the solution of the homogeneous wave equation, namely, (3.18)–(3.19) with $f = 0$. Denoting by $(\varphi_i(T))_{i=1}^l$ the POD basis constructed with y over $[0, T]$, for $T \in [0, \infty)$,*

$$\varphi_i(T) \underset{T\to\infty}{\longrightarrow} \widetilde{w}_{\sigma(i)},$$

where σ describes a certain reordering determined by the initial conditions. Therefore

$$\rho(l, T) = \|\pi_H^l(T)\|_{\mathcal{L}(V)} \underset{T\to\infty}{\longrightarrow} C.$$

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

4. Numerical validations

In this section, we provide some numerical validations of the above error estimates for some examples of one-dimensional problems. As in the rest of the paper, we only consider the case of *self-reduction, i.e.* when the reduction space we use is the POD space generated from the trajectory of the reference solution u itself. In particular, we aim at assessing whether or not the sequences (σ_l) and (ρ_l) are bounded in several examples. Of course, since the reference solution is needed to compute the POD space, it is mostly a theoretical study on synthetic data. However, this is an important first step before tackling the practical situation of *parametric variations*, when a unique POD space is used to reduce a family of solutions. This issue will be addressed in forthcoming papers.

4.1. Discretization and corresponding reduction for parabolic problem

Here, we consider the reduction of (3.9)–(3.10) with the one-dimensional non-linear equation

$$\begin{gathered}\partial_t u - \partial^2_{xx} u = f(t,u) \quad \text{in } (0,T)\times(0,1),\\ u(t,0) = u(t,1) = 0,\\ u(0,x) = u_0(x) \quad \text{in } (0,1),\end{gathered}$$

where now f is simply a $[0,T]\times\mathbb{R}\to\mathbb{R}$ function. In the sequel, $H = L^2(0,1)$ and $V = H^1_0(0,1)$. We keep the notations $(\cdot,\cdot)$ and $a(\cdot,\cdot)$ for their respective scalar products.

4.1.1. Semi-discrete solution and reduced form

Let u_h be the $\mathbf{P}_1$ approximation of u on the regular mesh $(x_i)_{i=1}^{N_h}$

$$x_i = ih, \quad 1\le i\le N_h, \quad h = \frac{1}{N_h+1},$$

associated with the basis of shape functions $(e_i)_{i=1}^{N_h}$. The discrete solution u_h is defined by

$$\frac{\mathrm{d}}{\mathrm{d}t}(u_h(t),e_i) + a(u_h(t),e_i) = (f(t,u_h(t)),e_i), \quad 1\le i\le N_h, \tag{4.1}$$

$$u_h(0,x) = u_{h,0}(x). \tag{4.2}$$

This discrete solution is the reference solution with which the reduced solutions will be compared. However, the POD basis $(\varphi_1,\dots,\varphi_l)$ itself will be constructed based on the fully-discrete solution u_h^n described below. Nevertheless, we emphasize that we do not consider time discretization issues in this paper, hence in our numerical trials we choose the time step "sufficiently small" for the discrete solution to be converged in time.

The corresponding reduced form u_h^l of u_h satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}(u_h^l(t),\varphi_k) + a(u_h^l(t),\varphi_k) = (f(t,u_h^l(t)),\varphi_k), \quad 1\le k\le l, \tag{4.3}$$

$$u_h^l(0,x) = u_{h,0}^l(x). \tag{4.4}$$

As before, we have local existence and uniqueness of the solutions u_h and u_h^l in the classical sense.

We also similarly introduce the $L^2(0,1)$-orthogonal and $H^1_0(0,1)$-orthogonal projectors $\pi^l_{L^2}$ and $\pi^l_{H^1_0}$ from V_h onto V^l, and the corresponding sequences

$$\begin{aligned}\rho_l &= \|\pi^l_{L^2}\|_{\mathcal{L}(H^1_0)},\\ \sigma_l &= \|\pi^l_{L^2} - \pi^l_{H^1_0}\|_{\mathcal{L}(H^1_0)},\end{aligned}$$

that still verify

$$\sigma_l \le \rho_l.$$

We can then directly adapt Proposition 3.7.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Proposition 4.1. *Assume $u_{h,0} \in V_h$, $u^l_{h,0} \in V^l$, $f \in C([0,T]\times\mathbb{R};\mathbb{R})$, and f is Lipschitz continuous in its second variable. Then there exists unique classical and global solutions u_h and u^l_h of equations (4.1)–(4.2) and (4.3)–(4.4), respectively. Moreover, for all $T>0$,*

$$\begin{aligned}\|u_h - u^l_h\|_{L^2(0,T;H^1_0)} &\le C\left(\|\pi^l_{L^2}u_{h,0} - u^l_{h,0}\|_{L^2(\Omega)} + \|u_h - \pi^l_{L^2}u_h\|_{L^2(0,T;H^1_0)}\right)\\ &\le C\big(\|\pi^l_{L^2}u_{h,0} - u^l_{h,0}\|_{L^2(\Omega)}\\ &\quad +(1+\sigma_l)\|u_h - \pi^l_{H^1_0}u_h\|_{L^2(0,T;H^1_0)}\big).\end{aligned} \tag{4.5}$$

Note that – if we assume that the POD basis is constructed in a continuous-time discrete-space setting – the last term in this error estimate directly corresponds to the POD remainder, recall (2.4), hence it is perfectly controlled in the POD construction itself.

4.1.2. Full discretization

We use the classical θ-method as a time discretization scheme. In order to compute the reference solution u_h, we need the non-reduced mass matrix M and stiffness matrix K

$$M = [(e_j,e_i)]_{1\le i,j\le N_h}, \quad K = [a(e_j,e_i)]_{1\le i,j\le N_h},$$

and the reaction term application $F:\mathbb{R}^{N_h}\to\mathbb{R}^{N_h}$ of coefficient

$$(F(t,\beta))_i = \int_0^1 f\Big(t,\sum_{k=1}^{N_h}\beta_k e_k(x)\Big)e_i(x)\,\mathrm{d}x.$$

Then the vector $U_h(t)\in\mathbb{R}^{N_h}$ concatenating the coordinates of $u_h(x,t)$ in $(e_i(x))_{i=1}^{N_h}$ satisfies

$$\begin{aligned}M\dot{U}_h(t) + KU_h(t) &= F(t,U_h(t)),\\ U_h(0) &= U_{h,0}.\end{aligned}$$

Next we apply a semi-implicit time scheme by the θ-method

$$\begin{aligned}\frac{1}{\Delta t}M\left(U_h^{n+1}-U_h^n\right) + K\left(\theta U_h^{n+1} + (1-\theta)U_h^n\right) &= \theta F(t^{n+1},U_h^{n+1})\\ &\quad + (1-\theta)F(t^n,U_h^n),\end{aligned}$$

leading to a non-linear problem in U_h^{n+1} once U_h^n is known, which we can solve for using a Newton algorithm.

For the reduced solutions u^l_h, we follow exactly the same path (spatial discretization $U^l_h(t)$, full discretization $(U^{l,n}_h)$), except that we substitute the POD basis $(\varphi_i)_{i=1}^l$ for the finite element basis $(e_i)_{i=1}^{N_h}$. This gives the reduced mass and stiffness matrices M^l and K^l, and the reduced reaction term F^l. We emphasize that although these reduced matrices are of limited size, they are full. We call Φ^l the matrix

$$\Phi^l = [\varphi_1,\ldots,\varphi_l] \quad \in\mathbb{R}^{N_h\times l},$$

where vectors φ_i are expressed as column vectors of coordinates in $(e_i)_{i=1}^{N_h}$. Then we obtain the following relations between reduced and non-reduced operators

$$\begin{aligned}M^l &= (\Phi^l)^{\mathrm{T}}M\Phi^l,\\ K^l &= (\Phi^l)^{\mathrm{T}}K\Phi^l,\\ F^l(t,\beta^l) &= (\Phi^l)^{\mathrm{T}}F(t,\Phi^l\beta^l),\\ \mathrm{d}_{\beta^l}F^l(t,\beta^l) &= (\Phi^l)^{\mathrm{T}}\mathrm{d}_\beta F(t,\Phi^l\beta^l)\Phi^l,\end{aligned}$$

where $\mathrm{d}_{\beta^l}F^l$ and $\mathrm{d}_\beta F$ denote the differential quantities needed in the Newton algorithm computations.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

4.2. Sharpness indicators for the new estimates

Here, we define the quantities that we need to check to ensure the *sharpness* of the new estimates. Note first that the POD eigenvalues λ_i typically decrease exponentially, hence the maximum POD rank to be considered is set as

$$\lambda_{l_{\max}+1} \leq 10^{-12}\lambda_1 \leq \lambda_{l_{\max}},$$

in order to preserve sufficient K-orthogonality of the POD basis $(\varphi_i)_{i=1}^{l}$ when we perform the diagonalization of the covariance matrix. Indeed, since the covariance matrix is ill-conditioned, this orthogonality tends to rapidly deteriorate with l and we should preserve

$$\left\|(\Phi^l)^{\mathrm{T}} K \Phi^l - \mathrm{Id}_l\right\| \leq \varepsilon_{\mathrm{tol}}.$$

4.2.1. Summary of the estimation chain

In Proposition 4.1, we mainly handle three error terms:

- the *reduction error* $R(l)$
$$R(l) = \|u_h - u_h^l\|_{L^2(0,T;H_0^1)};$$
- the L^2*-projection error* $Q(l)$
$$Q(l) = \|u_h - \pi_{L^2}^l u_h\|_{L^2(0,T;H_0^1)};$$
- and the H_0^1*-projection error* $P(l)$
$$P(l) = \|u_h - \pi_{H_0^1}^l u_h\|_{L^2(0,T;H_0^1)}, \tag{4.6}$$

  that coincides, in this situation of self-reduction, with the *POD remainder* $\varepsilon(l)$

$$\varepsilon(l) = \left\{\sum_{i>l} \lambda_i\right\}^{1/2}.$$

Note that we do not need the reduced solution u_h^l to compute $Q(l)$ nor $P(l)$. Moreover, we point out that these quantities – except for $P(l)$ which can be obtained as a by-product of the covariance computation – are auxiliary quantities only computed to evaluate the reduction performance and accuracy.

If we prescribe the initial condition

$$u_{h,0}^l = \pi_{L^2}^l u_{h,0},$$

then the first term in the right-hand side of equation (4.5) vanishes. Thus we summarize the estimation chain by

$$R(l) \leq CQ(l) \leq C(1+\sigma_l)P(l).$$

Let us introduce the following *sharpness indicator*

$$\mathcal{S}_{\mathrm{Gal}}(l) = \frac{R(l)}{Q(l)},$$

which is clearly bounded under the assumptions of Proposition 4.1, but can be considered in a more general framework. By contrast, note that for the second inequality that only relies on (3.7), the bound

$$\frac{Q(l)}{(1+\sigma_l)P(l)} \leq 1$$

always holds.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Finally, we aim at numerically verifying, in various cases, that:

- the maximum POD rank $l_{\max}$ is reasonably limited compared to the number of degrees of freedom of the system
$$l \ll N_h,$$
for the POD subspace to accurately approximate the solution, recall (4.6);
- the quantity $\max_{1\le l\le l_{\max}} \rho_l$, that is an upper bound of $\max_{1\le l\le l_{\max}} \sigma_l$, remains small;
- the indicator $\mathcal{S}_{\mathrm{Gal}}(l)$, $1 \le l \le l_{\max}$, remains numerically bounded, especially in cases of strong non-linearities.

4.2.2. Computation of the (ρ_l) and (σ_l) sequences

In order to compute ρ_l, it is useful to manipulate the L^2-orthonormal basis $(\psi_k)_k$ that results from a Gram–Schmidt L^2-orthonormalization on the H^1_0-orthonormal POD basis (φ_k). Thus

$$V^l = \mathrm{Span}(\psi_1, \dots, \psi_l),$$

with ψ_i independent of the POD rank l. Denoting by Ψ^l the matrix

$$\Psi^l = [\psi_1, \dots, \psi_l],$$

where the ψ_i elements are expressed as column vectors of coordinates in $(e_i)_{i=1}^{N_h}$, we notice that $\pi^l_{L^2}$ has the following matrix in $(e_i)_{i=1}^{N_h}$

$$\Pi^l_{L^2} = \Psi^l (\Psi^l)^{\mathrm{T}} M \quad \in \mathbb{R}^{N_h \times N_h}.$$

We use the definition of ρ_l

$$\begin{aligned}\rho_l^2 &= \sup_{v\in V_h\setminus\{0\}} \frac{\int_0^1 \left([\pi^l_{L^2} v]'(x)\right)^2 \mathrm{d}x}{\int_0^1 v'(x)^2 \,\mathrm{d}x} \\ &= \sup_{\beta\in\mathbb{R}^{N_h}\setminus\{0\}} \frac{\beta^{\mathrm{T}} (\Pi^l_{L^2})^{\mathrm{T}} K \Pi^l_{L^2} \beta}{\beta^{\mathrm{T}} K \beta}.\end{aligned}$$

Then ρ_l is the solution of the "largest K-eigenvalue" problem

$$\rho_l = \sup\left\{\omega \ge 0 \mid \exists \beta \in \mathbb{R}^{N_h} \setminus \{0\},\ (\Pi^l_{L^2})^{\mathrm{T}} K \Pi^l_{L^2} \beta = \omega^2 K \beta\right\}.$$

Similarly, let Φ^l be the matrix

$$\Phi^l = [\varphi_1, \dots, \varphi_l],$$

and $\widetilde{\Pi}^l_{L^2}$ be the matrix of the truncated projector $(\pi^l_{L^2} - \pi^l_{H^1_0})$

$$\widetilde{\Pi}^l_{L^2} = \Psi^l (\Psi^l)^{\mathrm{T}} M - \Phi^l (\Phi^l)^{\mathrm{T}} K.$$

Then σ_l is the solution of the problem

$$\sigma_l = \sup\left\{\omega' \ge 0 \mid \exists \beta \in \mathbb{R}^{N_h} \setminus \{0\},\ (\widetilde{\Pi}^l_{L^2})^{\mathrm{T}} K \widetilde{\Pi}^l_{L^2} \beta = \omega'^2 K \beta\right\}.$$

These properties will allow the numerical evaluation of the sequences.

4.3. Numerical experiments and validation for the parabolic problems

We present three numerical cases of POD reduction on the generic discrete parabolic equations (4.3)–(4.4). The corresponding parameters are gathered in Table 1. In all these cases, we take $\theta = \frac{2}{3}$ in the θ-method for time discretization, and $N_h = 100$ for the spatial discretization.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

TABLE 1. Cases of study for the reduction of parabolic equations.

Case	Case A	Case B	Case C
Nb. timesteps	10^3	10^3	800
Δt	10^{-4}	10^{-4}	10^{-5}
$u_0(x)$	$\mathbb{1}_{[\frac{1}{3},\frac{2}{3}]}(x)$	$\frac{27}{4}x^2(1-x)$	$\frac{27}{4}x^2(1-x)$
$f(t,u)$	$9u$	$10u^2$	$100u^2$

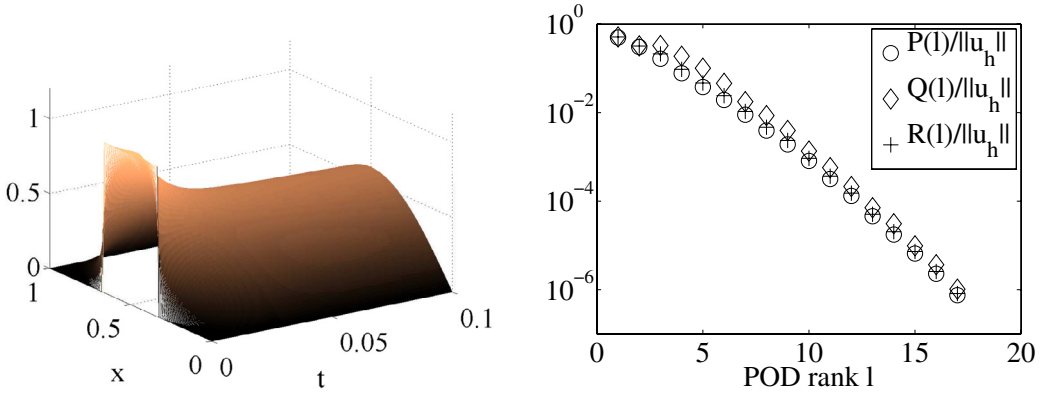


FIGURE 1. Case A. Left: full solution. Right: relative errors of H_0^1-projection, L^2-projection and reduction.

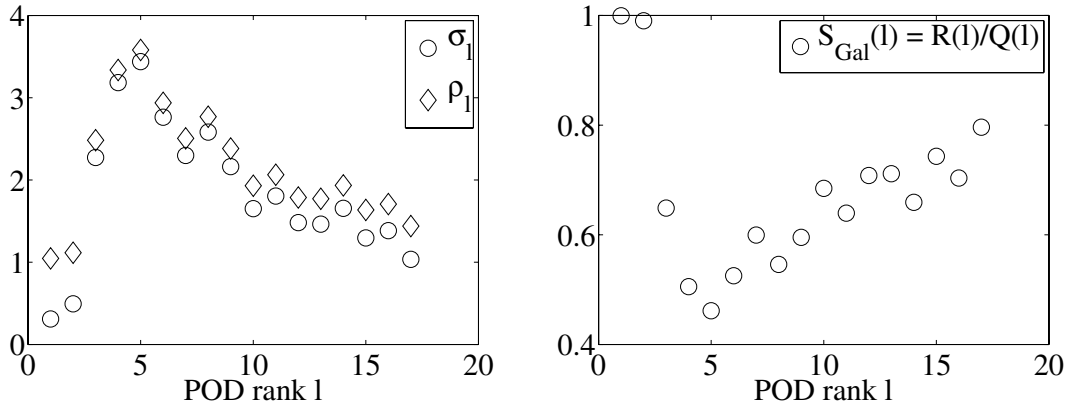


FIGURE 2. Case A. Left: POD sequences (ρ_l) and (σ_l). Right: sharpness indicator $(\mathcal{S}_{\mathrm{Gal}}(l))$.

4.3.1. Case A: Lipschitz continuous reaction term

Case A satisfies the assumptions of Proposition 4.1 since f is linear with respect to u. We display the corresponding results in Figures 1 and 2. In these figures the POD rank l varies from 1 to $l_{\max}$.

Figure 1, as an indication, shows the shape in space and time of the non-reduced solution u_h, and a numerical comparison between the indicators $P(l)$, $Q(l)$ and $R(l)$.

Figure 2 displays the sequence of POD constants ρ_l and their truncated versions σ_l, together with the sharpness indicator $\mathcal{S}_{\mathrm{Gal}}(l)$.

In this simple case, all our verifications are successful, namely,

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

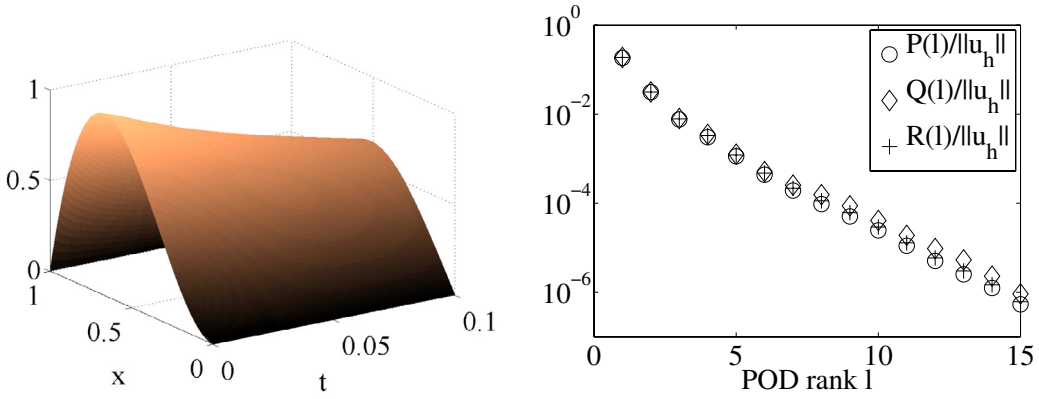


FIGURE 3. Case B. Left: full solution. Right: relative errors of H_0^1-projection, L^2-projection and reduction.

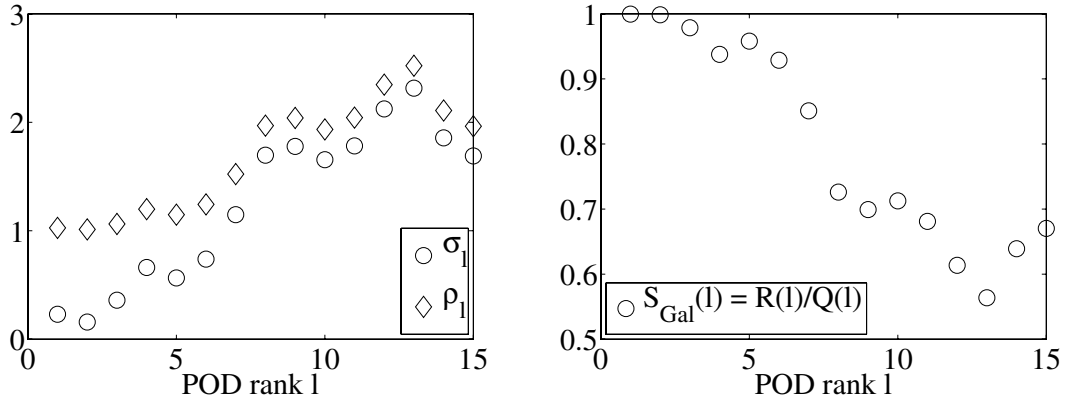


FIGURE 4. Case B. Left: POD sequences (ρ_l) and (σ_l). Right: sharpness indicator $(\mathcal{S}_{\mathrm{Gal}}(l))$.

- the POD remainder decreases at an exponential rate, and $l_{\max} = 10$;
- the POD constants ρ_l are of magnitude $O(1)$ and remain bounded with l. Also, the improvement provided by σ_l is limited;
- as expected with Proposition 4.1, the sharpness indicator is bounded and of small value, *viz.*

$$\mathcal{S}_{\mathrm{Gal}} \in [0.3,\, 0.8].$$

4.3.2. Cases B and C: super-linear reaction term

By contrast, the other two cases B and C are beyond the assumptions of Proposition 4.1 because we consider super-linear reaction terms. Moreover, while case B remains bounded, case C appears to explode in finite time, which is why we reduced the time range in this case while keeping a similar number of time steps, see Table 1. Even though our above error estimates do not hold in these cases, we can still compute the same numerical error quantities for illustrative purposes. Case B is reported on in Figures 3 and 4, and case C in Figures 5 and 6.

In fact, the maximum POD rank as well as the behavior and magnitude of the indicators ρ_l, σ_l, and $\mathcal{S}_{\mathrm{Gal}}$ reveal no significant difference compared to case A. Furthermore, the reduction error still decreases as fast as the POD remainder.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

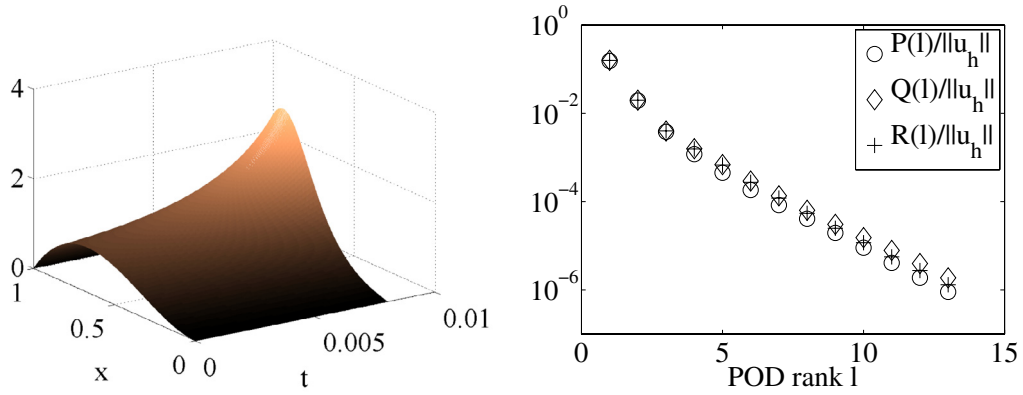


FIGURE 5. Case C. Left: full solution. Right: relative errors of H_0^1-projection, L^2-projection and reduction.

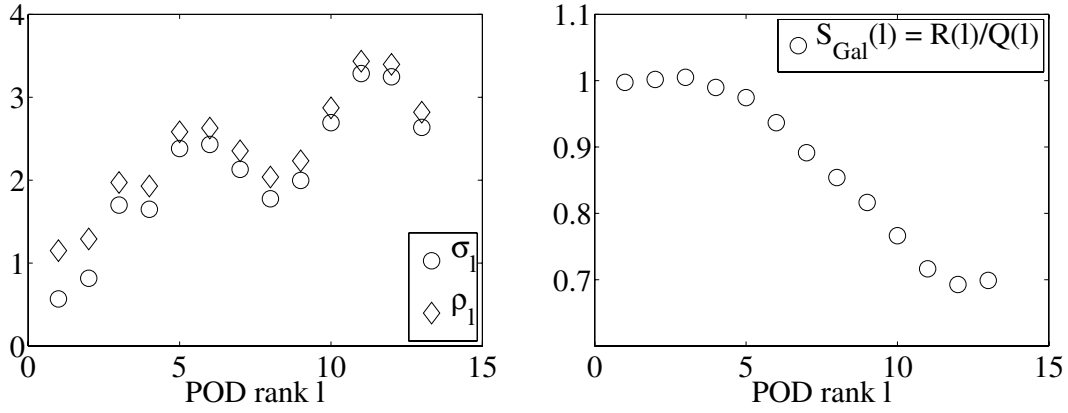


FIGURE 6. Case C. Left: POD sequences (ρ_l) and (σ_l). Right: sharpness indicator $(\mathcal{S}_{\mathrm{Gal}}(l))$.

4.4. Numerical assessment of the wave equation reduction

Considering now the 1D homogeneous wave equation,

$$\begin{aligned} \partial_{tt}^2 y - c^2 \partial_{xx}^2 y = 0 \quad &\text{in } (0,T)\times(0,1), \\ y(t,0) = y(t,1) &= 0, \\ y(0,x) = y_0(x) \quad &\text{in } (0,1), \\ \partial_t y(0,x) = \dot{y}_0(x) \quad &\text{in } (0,1), \end{aligned}$$

we report on the numerical values obtained for the various error terms. We discretize in space with finite elements on a regular mesh, and in time with a Newmark scheme according to the classical parameters $\beta = \frac{1}{4}$ and $\gamma = \frac{1}{2}$, see *e.g.* [22]. We take a regular cutoff function for $y_0(x)$, and $\dot{y}_0(x) = 0$. The corresponding results are shown in Figures 7 and 8.

We verify that the POD basis is very close to a set of $H_0^1(0,1)$-eigenmodes $(\tilde{w}_i)$ of the Dirichlet Laplacian as substantiated in Section 3.5. This also explains why σ_l is much lower than ρ_l, since the L^2 and H_0^1 projectors onto eigenspaces coincide.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

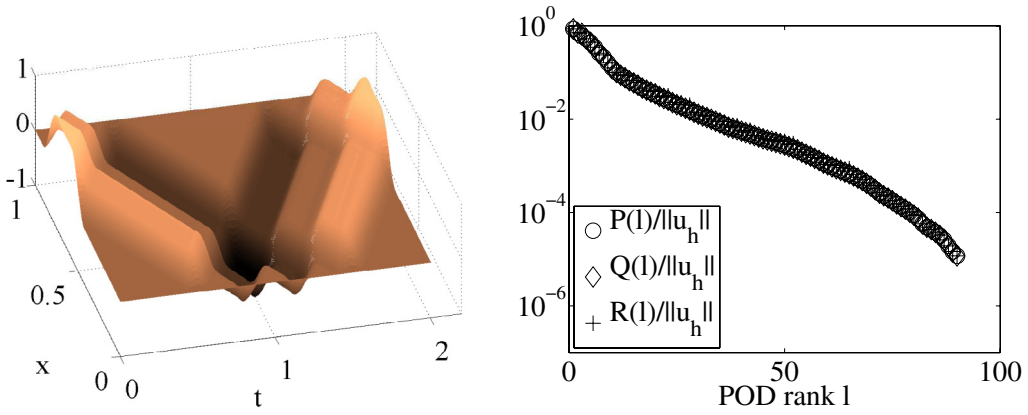


FIGURE 7. Homogeneous wave equation. Left: full solution. Right: relative errors of H_0^1-projection, L^2-projection and reduction.

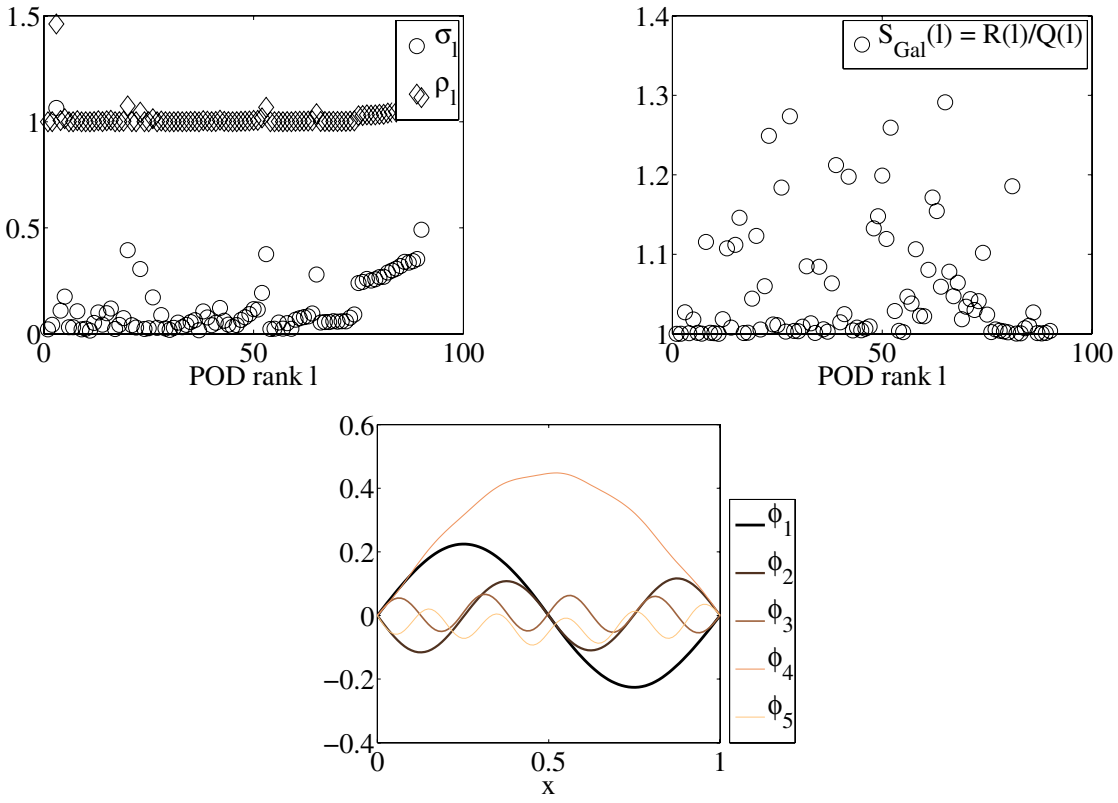


FIGURE 8. Homogeneous wave equation. Top-left: POD sequences (ρ_l) and (σ_l). Top-right: sharpness indicator $(\mathcal{S}_{\mathrm{Gal}}(l))$. Bottom: the first five POD modes.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Note that the estimate of Proposition 3.13 contains the first-order time derivative $\frac{\partial}{\partial t}(u - \pi^l_{H^1_0} u)$ which is not controlled by the POD construction. Nevertheless, we observe from Figure 7 that the POD reduction is very effective, and indeed converges nearly-exponentially with the POD-rank.

5. Reduction of a complex system: a biomechanical heart model

In this section, we test our reduction technique and estimates with the 3D continuum mechanics model of a beating heart. The electromechanical heart model and its discretization are described in [25], and a validation of the model by confrontation with clinical data is given in [4].

A model describing the three-dimensional electromechanical behavior of the heart requires several important ingredients, namely,

- a constitutive law accounting for both the active and passive aspects in the behavior of the muscle fibres;
- a representation of the electrical activation – the input in the constitutive law – that can be obtained from modeling approaches of various types and complexities;
- a geometrical (or "anatomical") description of the myocardium incorporating the fibre directions;
- a model of the blood circulation inside and outside of the heart cavities, and also a model describing the opening and closure of the valves that separate the cavities from each other and from the external circulation.

Let $\underline{y}$ be the displacement field, the deformation gradient is defined by $\underline{\underline{F}} = \underline{\underline{I}} + \underline{\underline{\nabla}}\,\underline{y}$ and the Green–Lagrange strain tensor is given by $\underline{\underline{e}} = \frac{1}{2}(\underline{\underline{F}}^T \cdot \underline{\underline{F}} - \underline{\underline{I}})$ and $J = \det \underline{\underline{F}}$. Based on the above modeling ingredients, the second Piola-Kirchhoff stress tensor $\underline{\underline{\Sigma}}$ contains the active cardiac fibre law, a viscous stress component and a hyperelastic potential accounting for passive effects, these components being combined by means of a rheological model of Hill-Maxwell type.

Using a total Lagrangian formulation and denoting by Ω_H the reference domain corresponding to cardiac tissue, while the part of the boundary corresponding to ventricular endocardium is denoted by Γ, the principle of virtual work then gives

$$\int_{\Omega_H} \rho\, \underline{\ddot{y}} \cdot \underline{v}\, \mathrm{d}\Omega + \int_{\Omega_H} \underline{\underline{\Sigma}} : \mathrm{d}_{\underline{y}}\, \underline{\underline{e}} \cdot \underline{v}\, \mathrm{d}\Omega + \int_{\Gamma} P_0 \underline{\nu} \cdot \underline{\underline{F}}^{-1} \cdot \underline{v} J\, \mathrm{d}\Gamma = 0 \quad \forall \underline{v} \in V,$$

where V denotes a suitable space of displacement test functions, ρ the mass per unit volume, $\mathrm{d}_{\underline{y}}\, \underline{\underline{e}}$ the differential of the Green-Lagrange strain tensor with respect to the displacement, while P_0 is a prescribed intraventricular pressure.

For the simulations presented hereafter an idealized left ventricle embedded with active fibers has been considered. The discretization is performed with $\mathbf{P}_1$-Lagrange finite elements in space (with about 1000 degrees of freedom), and a Newmark scheme in time [3]. We show some snapshots of the solution for the full finite element model in Figure 9.

Although we do not have a theoretical estimate for the reduction error in this complex non-linear case, the three error terms appearing in the linear estimation chain still feature excellent decreasing rate and correlation, see Figures 10 and 11. Analyzing our indicators reveals that their magnitude may slightly differ from the linear one-dimensional case, but again shows the effectiveness of the POD reduction, namely,

- $l_{\max} = 36$;
- ρ_l and σ_l are almost identical, and numerically bounded;
- the sharpness indicator established for the linear case is still bounded, and more precisely

$$\mathcal{S}_{\mathrm{Gal}} \in [0.6,\ 1.0].$$

In Figure 10, we also display the evolution in time of the relative residual e^l defined as

$$\mathrm{e}^l(t) = \frac{\|\underline{y}(t) - \underline{y}^l(t)\|_{L^\infty(\Omega)}}{\|y\|_{C([0,T];L^\infty(\Omega))}}. \tag{5.1}$$

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

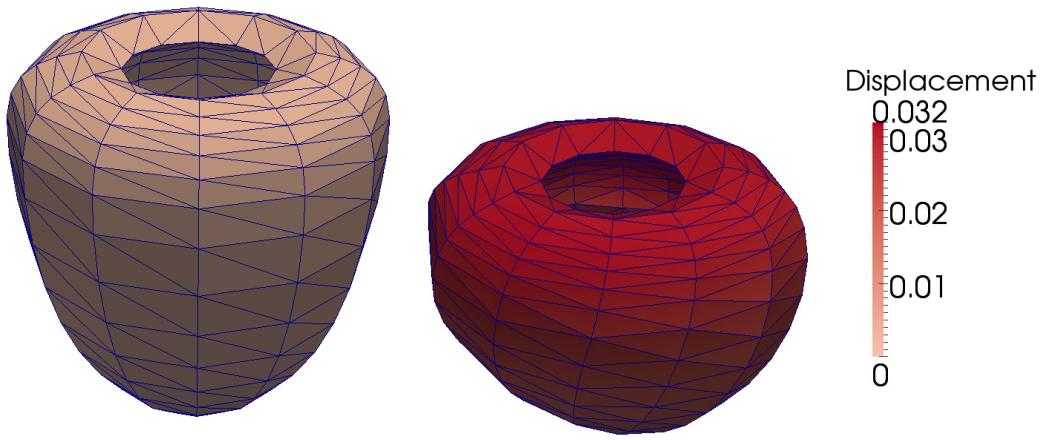


FIGURE 9. Model of a ventricle: snapshots of the displacement field at the beginning (left) and 40% (right) of the first cardiac cycle for the full model.

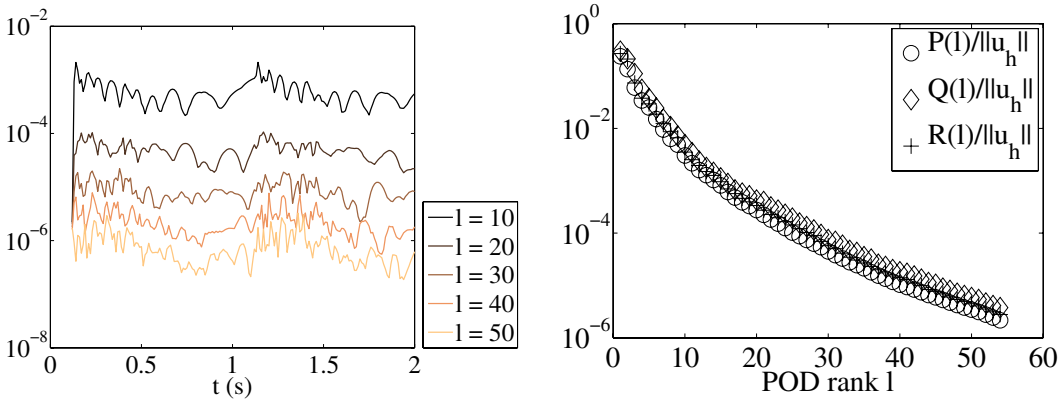


FIGURE 10. Model of a ventricle. Left: evolution in time of the residual e^l (see (5.1)), for several POD ranks. Right: relative errors of H_0^1-projection, L^2-projection and reduction.

We observe an excellent behavior of e^l, which roughly decreases by an order of magnitude for each addition of 10 modes in the POD basis.

6. CONCLUSION

We have proposed Galerkin estimates for the Proper Orthogonal Decomposition reduction of some classical PDEs. The numerical implementation of the reduction and some verifications were presented in this article. We have also demonstrated reduced simulations of a complex three-dimensional electromechanical model of the heart, where the validity of similar Galerkin estimates is numerically verified, although no formal proof can be given in this case.

A special emphasis was placed on the derivation of POD-reduction error estimates in convenient norms and which can be controlled in the construction of the POD basis.

The present study can be extended in many directions. Firstly, as far as POD reduction of PDEs is concerned, one of the major difficulties lies in achieving stability of the POD basis with respect to *e.g.* parameter variations,

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

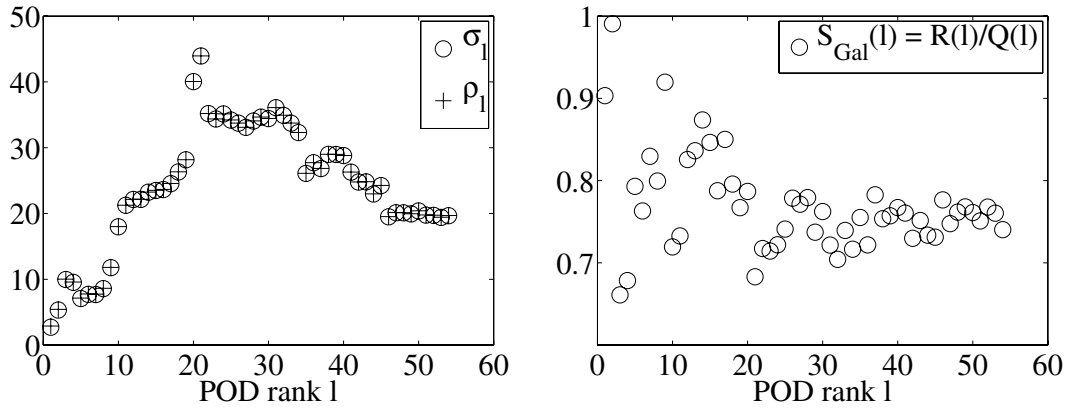


FIGURE 11. Model of a ventricle. Left: POD sequences (ρ_l) and (σ_l). Right: sharpness indicator $(\mathcal{S}_{\mathrm{Gal}}(l))$.

initial and boundary conditions, and so on. This subject needs to be further investigated. Secondly, filtering and estimation techniques for inverse modeling are extremely costly from a computational standpoint. This justifies – or even often requires – the use of POD-based reduced models and/or reduced filters and hence, the derivation of error estimates for such problems is crucial.

APPENDIX A. EXISTENCE AND UNIQUENESS OF SOLUTIONS OF VARIATIONAL EQUATIONS WITH A LIPSCHITZ CONTINUOUS REACTION TERM

Although some results pertaining to this type of problem exist in the literature, for the sake of completeness we provide the sketch of a self-contained proof for the specific result that we need in our case, namely, Proposition 3.5.

Since we assume the embedding $V \hookrightarrow H$ to be compact, we can use the Hilbertian bases (w_i) and $(\tilde{w}_i)$ of H and V – respectively – made up by the eigenvectors, as already introduced in Section 3.5. Let W_k, $k \geq 1$, denote the subspace

$$W_k = \mathrm{Span}(w_1, \ldots, w_k) = \mathrm{Span}(\tilde{w}_1, \ldots, \tilde{w}_k),$$

and P_k the orthogonal projector from H onto W_k, *i.e.*

$$P_k h = \sum_{i=1}^{k} (h, w_i) w_i, \quad \forall h \in H.$$

It coincides with the orthogonal projector from (V, a) onto W_k defined by

$$P_k^a v = \sum_{i=1}^{k} a(v, \tilde{w}_i) \tilde{w}_i = P_k v, \quad \forall v \in V.$$

We will show the existence result by a Galerkin approach using the sequence of eigenspaces.

Proposition A.1. *For all $k \geq 1$, there exists a unique global solution $u_k \in C^1(\mathbb{R}^+; W_k)$ such that*

$$\frac{\mathrm{d}}{\mathrm{d}t}(u_k(t), v_k) + a(u_k(t), v_k) = (f(t, u_k(t)), v_k), \quad \forall v_k \in W_k, \tag{A.1}$$

$$u_k(0) = P_k u_0.$$

Moreover, for all $T > 0$, (u_k) is bounded in $C([0, T]; H)$.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Proof. Let us first prove uniqueness. Consider two solutions u_k^1, u_k^2, then

$$\frac{\mathrm{d}}{\mathrm{d}t}|u_k^1 - u_k^2|^2(t) \leq 2L|u_k^1 - u_k^2|^2(t),$$

and since $u_k^1(0) = u_k^2(0)$, we infer $u_k^1 = u_k^2$.

We now tackle the global existence. By the Peano existence theorem (see *e.g.* [10], Sect. 2.4.4), we have local existence. By uniqueness, the maximum time of existence $T_k^\star \in (0, \infty]$ is well-defined. We test the equation with $v_k = u_k(t)$, *i.e.*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|u_k(t)|^2 + \|u_k(t)\|_a^2 = (f(t, u_k(t)) - f(t,0), u_k(t)) + (f(t,0), u_k(t)).$$

By Young's inequality (here $|(f(t,0), u_k(t))| \leq L|u_k(t)|^2 + 1/(4L)|f(t,0)|^2$),

$$\frac{\mathrm{d}}{\mathrm{d}t}|u_k(t)|^2 \leq 4L|u_k(t)|^2 + \frac{1}{2L}|f(t,0)|^2.$$

Then, by Gronwall's lemma, for all $T > 0$,

$$\lim_{t \to T^-} |u_k(t)| \leq C(T) < \infty,$$

with $C(T) = |u_0| + \frac{1}{2L}\int_0^T e^{4L(T-s)}|f(t,0)|^2\,\mathrm{d}t$. Finally, on the one hand we deduce global existence, *i.e.* $T_k^\star = \infty$ (*e.g.* [10], Sects. 2.4.3, 2.4.4), and on the other hand we get the boundedness in $C([0,T];H)$. □

In order to show that (u_k) is a Cauchy sequence in the Banach spaces $C([0,T];H)$ and $L^2(0,T;V)$, let us consider the decomposition

$$u_{k+p} - u_k = P_k(u_{k+p} - u_k) + (\mathrm{Id} - P_k)u_{k+p}. \tag{A.2}$$

For the first term in the right-hand side, we get the following estimates.

Lemma A.2. *For all $0 \leq t_0 \leq t_1$ and all $k, p \geq 1$,*

$$\max\left(\|P_k(u_{k+p} - u_k)\|_{C([t_0,t_1];H)},\ \sqrt{2c_a}\|P_k(u_{k+p} - u_k)\|_{L^2(t_0,t_1;V)}\right) \leq |(u_{k+p} - u_k)(t_0)| + \sqrt{2L(t_1 - t_0)}\|u_{k+p} - u_k\|_{C([t_0,t_1];H)}.$$

Proof. For all $v_k \in W_k$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(u_{k+p}(t) - u_k(t), v_k) + a(u_{k+p}(t) - u_k(t), v_k) = (f(t, u_{k+p}(t)) - f(t, u_k(t)), v_k). \tag{A.3}$$

Testing this equation with $v_k = P_k(u_{k+p} - u_k)$, using orthogonality properties of P_k, and finally integrating on $[t_0, t]$, $t \in [t_0, t_1]$, we have

$$\frac{1}{2}|P_k(u_{k+p} - u_k)(t)|^2 + c_a\|P_k(u_{k+p} - u_k)(t)\|_{L^2(t_0,t;V)}^2 \leq \frac{1}{2}|P_k(u_{k+p} - u_k)(t_0)|^2 + L\|u_{k+p} - u_k\|_{L^2(t_0,t;H)}^2.$$

□

Using the diagonalisation of a, we obtain the following estimate for the second term in (A.2) in the $C([t_0,t_1];H)$-norm.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Lemma A.3. *For all $0 \le t_0 \le t_1$ and all $k, p \ge 1$*

$$\|(\mathrm{Id} - P_k)u_{k+p}\|_{C([t_0,t_1];H)} \le |(\mathrm{Id} - P_k)u_{k+p}(t_0)| + \frac{C}{\omega_{k+1}}.$$

Proof. Applying now (A.1) for u_{k+p} with the test function $v^l = (\mathrm{Id} - P_k)u_{k+p}(t)$ yields

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|(\mathrm{Id} - P_k)u_{k+p}|^2(t) + \|(\mathrm{Id} - P_k)u_{k+p}(t)\|_a^2 = \Big(f(t, u_{k+p}(t)), (\mathrm{Id} - P_k)u_{k+p}(t)\Big). \tag{A.4}$$

Note that

$$\|(\mathrm{Id} - P_k)u_{k+p}(t)\|_a^2 \ge \omega_{k+1}^2 |(\mathrm{Id} - P_k)u_{k+p}(t)|^2,$$

so that by Young's inequality on the right-hand side of (A.4), we infer

$$\Big(f(t, u_{k+p}(t)), (\mathrm{Id} - P_k)u_{k+p}(t)\Big) \le \frac{1}{4\omega_{k+1}^2}|f(t, u_{k+p}(t))|^2 + \|(\mathrm{Id} - P_k)u_{k+p}(t)\|_a^2.$$

We conclude using the Lipschitz character of f and the boundedness of (u_k) in $C([0,T];H)$. □

We are ready to show the first convergence result.

Proposition A.4. *For all $T > 0$, (u_k) converges in $C([0,T];H)$ to some limit u.*

Proof. Let $\tau = \frac{1}{8L}$ and $j \ge 1$. Lemmas A.2 and A.3 lead to

$$\frac{1}{2}\|u_{k+p} - u_k\|_{C([(j-1)\tau, j\tau];H)} \le |(u_{k+p} - u_k)((j-1)\tau)| + |(\mathrm{Id} - P_k)u_{k+p}((j-1)\tau)| + \frac{C}{\omega_{k+1}}. \tag{A.5}$$

We prove by induction that the statement

$$\mathcal{P}(j): \quad (u_k) \text{ is a Cauchy sequence in } C([0, j\tau];H)$$

holds for all $j \ge 1$.

We easily show $\mathcal{P}(1)$. Assume now that $\mathcal{P}(j-1)$ holds for some $j \ge 2$. Let u be the limit of (u_k) in $C([0,(j-1)\tau];H)$. Then we decompose in (A.5)

$$|(\mathrm{Id} - P_k)u_{k+p}((j-1)\tau)| \le |(u_{k+p} - u)((j-1)\tau)| + |(\mathrm{Id} - P_k)u((j-1)\tau)|,$$

which proves that (u_k) is a Cauchy sequence in $C([(j-1)\tau, j\tau];H)$, and hence that $\mathcal{P}(j)$ holds. □

Remark that we directly obtain

$$u(0) = u_0. \tag{A.6}$$

Next, we get an estimate for the second term in (A.2) in the $L^2(0,T;V)$ norm.

Lemma A.5. *For all $T > 0$ and all $k, p \ge 1$,*

$$\|(\mathrm{Id} - P_k)u_{k+p}\|_{L^2(0,T;V)} \le C(|(\mathrm{Id} - P_k)u_0| + g_{k,p}), \tag{A.7}$$

where the sequence $g_{k,p}$ is defined as

$$g_{k,p} = \|(\mathrm{Id} - P_k)f(\cdot, u_{k+p})\|_{L^2(0,T;H)},$$

and verifies

$$\forall \varepsilon > 0,\ \exists k_0,\ \forall k \ge k_0,\ \forall p \ge 0, \quad g_{k,p} \le \varepsilon. \tag{A.8}$$

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

Proof. We consider again (A.4) that we rewrite as

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|(\mathrm{Id}-P_k)u_{k+p}|^2(t) + \|(\mathrm{Id}-P_k)u_{k+p}(t)\|_a^2 = \Big((\mathrm{Id}-P_k)f(t,u_{k+p}(t)),(\mathrm{Id}-P_k)u_{k+p}(t)\Big).$$

Then by integration,

$$c_a\|(\mathrm{Id}-P_k)u_{k+p}\|^2_{L^2(0,T;V)} \le \frac{1}{2}|(\mathrm{Id}-P_k)u_0|^2 + g_{k,p}\|(\mathrm{Id}-P_k)u_{k+p}\|_{L^2(0,T;V)},$$

and by Young's inequality, we obtain the estimate (A.7).

Now, by strong convergence in $C([0,T];H)$ and continuity of f,

$$|(\mathrm{Id}-P_k)f(s,u_n(s))| \underset{n\to\infty}{\longrightarrow} |(\mathrm{Id}-P_k)f(s,u(s))|.$$

Also,

$$|(\mathrm{Id}-P_k)f(s,u_n(s))| \le L\overline{C} + |f(s,0)| \le C,$$

so that by the dominated convergence theorem,

$$\bar{g}_{k,n} = \|(\mathrm{Id}-P_k)f(\cdot,u_n)\|_{L^2(0,T;H)} \underset{n\to\infty}{\longrightarrow} \bar{g}_k = \|(\mathrm{Id}-P_k)f(\cdot,u)\|_{L^2(0,T;H)}.$$

Moreover, by the Parseval theorem, $\bar{g}^2_{k,n}$ and $\bar{g}^2_k$ are the remainders of some positive converging series, so that in particular $\bar{g}_k \underset{k\to\infty}{\longrightarrow} 0$, and $(\bar{g}_{k,n})_k$ is a decreasing sequence for each n. Finally for $\varepsilon > 0$, there exists K such that $\bar{g}_K \le \frac{\varepsilon}{2}$, and n_0 such that for all $n \ge n_0$, $|\bar{g}_{K,n} - \bar{g}_K| \le \frac{\varepsilon}{2}$. We conclude by taking $k_0 = \max(K, n_0)$. □

This entails the second convergence result.

Proposition A.6. *For all $T > 0$, (u_k) converges in $L^2(0,T;V)$ and its limit is u.*

Proof. By the decomposition (A.2) and Lemma A.2,

$$\|u_{k+p} - u_k\|_{L^2(0,T;V)} \le C\Big(|(\mathrm{Id}-P_k)u_0| + \|u_{k+p} - u_k\|_{C([0,T];H)} + g_{k,p}\Big).$$

Using (A.8), (u_k) is also a Cauchy sequence in $L^2(0,T;V)$. Let $\tilde{u}$ be its limit. Since $L^2(0,T;V)$ and $C([0,T];H)$ are both continuously embedded in $L^2(0,T;H)$, then $\tilde{u} = u$. □

We can finally conclude.

Proof of Proposition 3.5. Using the previous convergence results we can reinterpret the limit u as satisfying equation (3.9) in the distribution sense. Given the regularity of u, we have that $-a(u(t),v) + (f(t,u(t)),v)$ is in $L^2(0,T)$ for any $v \in V$, hence it directly follows that $\frac{\mathrm{d}u}{\mathrm{d}t} \in L^2(0,T;V')$. We finally prove the uniqueness as in Proposition A.1. □

References

[1] D. Amsallem and C. Farhat, Interpolation method for adapting reduced-order models and application to aeroelasticity. *AIAA J.* **46** (2008) 1803–1813.

[2] A. Astolfi, Model reduction by moment matching for linear and nonlinear systems. *IEEE Trans. Automat. Cont.* **55** (2010) 2321–2336.

[3] K.J. Bathe, *Finite Element Procedures.* Prentice Hall (1996).

[4] R. Chabiniok, D. Chapelle, P.-F. Lesault, A. Rahmouni and J.-F. Deux, Validation of a biomechanical heart model using animal data with acute myocardial infarction, in *MICCAI Workshop on Cardiovascular Interventional Imaging and Biophysical Modelling (CI2BM09)* (2009).

[5] P.G. Ciarlet, *The Finite Element Method for Elliptic Problems.* North-Holland (1987).

[6] P. Clément, Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.* **8** (1975) 77–84.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press

[7] L. Daniel, C.S. Ong, S.C. Low, H.L. Lee and J. White, A multiparameter moment-matching model-reduction approach for generating geometrically parameterized interconnect performance models. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* **23** (2004) 678–693.

[8] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology* **5** (1992).

[9] B.F. Feeny and R. Kappagantu, On the physical interpretation of proper orthogonal modes in vibrations. *J. Sound Vib.* **211** (1998) 607–616.

[10] T.M. Flett, *Differential Analysis*. Cambridge University Press (1980).

[11] S. Gugercin and A.C. Athanasios, A survey of model reduction by balanced truncation and some new results. *Int. J. Control* **77** (2004) 748–766.

[12] M. Hinze and S. Volkwein, Proper orthogonal decomposition surrogate models for nonlinear dynamical systems: Error estimates and suboptimal control, in Dimension Reduction of Large-Scale Systems, edited by T.J. Barth, M. Griebel, D.E. Keyes, R.M. Nieminen, D. Roose, T. Schlick, P. Benner, D.C. Sorensen and V. Mehrmann. *Lect. Notes Comput. Sci. Eng.* **45** (2005) 261–306.

[13] M. Hinze and S. Volkwein, Error estimates for abstract linear-quadratic optimal control problems using proper orthogonal decomposition. *Comput. Optim. Appl.* **39** (2008) 319–345.

[14] P. Holmes, J. Lumley and G. Berkooz, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press, Cambridge (1996).

[15] M. Kahlbacher and S. Volkwein, Galerkin proper orthogonal decomposition methods for parameter dependent elliptic systems. Discussiones Mathematicae: Differential Inclusions, *Control and Optimization* **27** (2007) 95–117.

[16] D.-D. Kosambi, Statistics in function space, *J. Indian Math. Soc. (N.S.)* **7** (1943) 76–88.

[17] K. Kunisch and S. Volkwein, Galerkin proper orthogonal decomposition methods for parabolic problems. *Numer. Math.* **90** (2001) 117–148.

[18] K. Kunisch and S. Volkwein, Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM J. Numer. Anal.* **40** (2002) 492–515 (electronic).

[19] K. Kunisch and S. Volkwein, Proper orthogonal decomposition for optimality systems. *ESAIM: M2AN* **42** (2008) 1–23.

[20] Y. Maday, A.T. Patera and G. Turinici, *A priori* convergence theory for reduced-basis approximations of single-parameter elliptic partial differential equations. *J. Sci. Comput.* **17** (2002) 437–446.

[21] C. Prud'homme, D.V. Rovas, K. Veroy and A.T. Patera, A mathematical and computational framework for reliable real-time solution of parametrized partial differential equations. *ESAIM: M2AN* **36** (2002) 747–771. Programming.

[22] P.-A. Raviart and J.-M. Thomas, *Introduction à l'Analyse Numérique des Equations aux Dérivées Partielles*. Collection Mathématiques Appliquées pour la Maîtrise (in French), Masson (1983).

[23] D.V. Rovas, L. Machiels and Y. Maday, Reduced-basis output bound methods for parabolic problems. *IMA J. Numer. Anal.* **26** (2006) 423–445.

[24] G. Rozza, D.B.P. Huynh and A.T. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: application to transport and continuum mechanics. *Arch. Comput. Methods Eng.* **15** (2008) 229–275.

[25] J. Sainte-Marie, D. Chapelle, R. Cimrman and M. Sorine, Modeling and estimation of the cardiac electromechanical activity. *Comput. Struct.* **84** (2006) 1743–1759.

[26] T. Stykel, Balanced truncation model reduction for semidiscretized Stokes equation. *Linear Algebra Appl.* **415** (2006) 262–289.

[27] K. Veroy, C. Prud'homme and A.T. Patera, Reduced-basis approximation of the viscous Burgers equation: rigorous a posteriori error bounds. *C. R. Math. Acad. Sci. Paris* **337** (2003) 619–624.

[28] K. Willcox and J. Peraire, Balanced model reduction *via* the proper orthogonal decomposition. *AIAA J.* **40** (2002) 2323–2330.

https://doi.org/10.1051/m2an/2011053 Published online by Cambridge University Press