Recent advances on nonlinear Reduced Order Modelling for stability and bifurcations problems in incompressible fluid dynamics

Giuseppe Pitton, Gianluigi Rozza*

SISSA, International School for Advanced Studies, Mathematics Area, mathLab, via Bonomea 265, 34136, Trieste, Italy

Abstract

In this paper we propose a Reduced Basis framework for the computation of bifurcation and stability problems arising from nonlinear partial differential equations. The proposed method aims at reducing the complexity and the computational time required for the construction of bifurcation and stability diagrams. The method is quite general since it can in principle be specialized to a wide class of nonlinear problems, but in this paper we focus on an application in incompressible Fluid Dynamics at low Reynolds numbers. The validation with a benchmark cavity flow problem is satisfactory.

Keywords: Reduced Basis Method, Proper Orthogonal Decomposition, Steady Bifurcation, Hopf Bifurcation, Navier Stokes, Flow Stability, Spectral Element Method

1. Introduction

The study of stability and bifurcation of nonlinear systems is an established research field, of great importance both from the applied mathematics and engineering perspective.

The aim of Bifurcation Theory is the computation of branches of non-unique solutions for nonlinear problems [1], and the detection of the intersection points between the solution branches, called *bifurcation points*, as a function of some physical or geometrical control parameters.

Some typical control parameters are the geometrical aspect factors, or relevant adimensional quantities (such as the Reynolds number, or the Grashof number), or the boundary conditions, or forcing terms.

Usually the results of a stability analysis are represented in a diagram where the stability region of each branch of solutions is depicted as a function of the control parameters. Bifurcation diagrams show a qualitative or quantitative (pointwise or integral) aspect of a solution as a function of the control parameters, allowing to understand the structure and the physical features of the solution set. See for example [2] for a classical introduction to stability problems in Fluid Mechanics.

For instance, Bifurcation Theory is of evident practical importance in the analysis of stability problems in Elasticity, where it is well known that for many structures there exists a *critical load* which if exceeded will produce a catastrophic collapse (a classic reference on this subject is [3], a more mathematically oriented one is [4]).

Common applications of Bifurcation Theory in Fluid Mechanics include a wide range of industrial problems such as tribology, microfluid dynamics, biomedical industry, biomedicine (blood flows), and many more. In particular, the field of hydrodynamic stability focuses on the classification of the nature of flows as some control parameters are varied [2]. Specifically, the purpose of a stability investigation is, given a well-defined initial or boundary value problem, to classify the asymptotic solutions as stable, or time periodic, or even chaotic in time.

*Corresponding author
Email addresses: gpitton@sissa.it (Giuseppe Pitton), grozza@sissa.it (Gianluigi Rozza)

The construction of stability maps and bifurcation diagrams is a delicate and very expensive task, requiring an important computational effort and prohibitive if the number and range of parameters is large. This is particularly true for three and higher dimensional simulations.

Recent developments of Reduced Order Modelling (ROM) techniques have focused on the reduction of computational time for a wide range of differential problems, while mantaining a prescribed tolerance on error bounds [5]. It is therefore of great interest to further investigate how such methods can be applied to stability problems in Fluid Dynamics and Elasticity in order to reduce the computational power required.

The literature has already shown the effectiveness of Proper Orthogonal Decomposition (POD) both for analysis of principal modes [6] and for the reduction of computational power required by transient simulations. For instance, Terragni and Vega [7] showed how a POD approach could save a considerable amount of computing time for the analysis of bifurcations in some nonlinear dissipative systems. In a recent paper Herrero, Maday and Pla [8] have shown that both POD and the Reduced Basis Method (RBM) can reconstruct the behaviour of velocity and temperature field for a two-dimensional natural convection (Boussinesq) problem with large reduction of the computational power with respect to classical techniques. In particular, stable and unstable solutions are correctly identified, and a surrogate error estimate is always mantained below a prescribed tolerance. A recent remarkable work by Yano et al. [9] introduced a RB Method for the stability of flows under perturbations in the forcing term or in the boundary conditions, based on a space-time framework that allows for particularly sharp error estimates.

Summarizing, given the relatively fast decay of energy spectrum for flows at sufficiently low Reynolds numbers, a Reduced Order Modelling technique could be expected to be an efficient tool for flow stability analysis. Reduced Basis (RB) techniques [10] historically have been proven effective for the study of elastic stability of plates [11], but their application in more complex parametrized stability problems such as in Fluid Mechanics is an open and quite relevant research field.

In this paper, we propose the application of Reduced Order Modeling (ROM) techniques to reduce the quite demanding computing costs for bifurcation and stability analysis of flows. We will consider a benchmark problem well known in the literature, namely a buoyancy-driven flow in a rectangular cavity [12].

The focus of this work is devoted on several improvements with respect to the state of the art for these problems, approached with ROM: in particular we mention approximation stability, sampling, and reduced eigenproblems for stability analysis. Our exposition presents the topic from the different viewpoints of Nonlinear Mathematical Analysis, Applied Mathematics and Numerical Analysis, trying to underline the different aspects of the problem under consideration.

The structure of the paper is the following. In section 2 we present the class of abstract problems that will be considered and we briefly recall some important results on Nonlinear Analysis and Bifurcation Theory. Section 3 is devoted to the presentation of a general ROM technique and particular attention is devoted to the approximation of bifurcation problems. In section 4 the mathematical setting for the approximation of incompressible Fluid Dynamics equation is recalled, and the high-order method used is presented. In section 5 the ROM technique previously developed is specialized to the Navier-Stokes case and finally some numerical results are show and discussed in section 6.

2. Abstract setting

We start our discussion of reduction strategies for bifurcation and stability problems recalling some basic elements of Nonlinear Analysis that will prove to be useful in the applications. Despite being a relatively recent field, there are at least two main frameworks of Nonlinear Analysis of striking effectiveness, a topological one and a variational one. Roughly, the motivation behind these two approaches are respectively:

- consider at the problem as a functional equation between Banach spaces, and study it in a purely abstract setting [1] (e.g. using Implicit Function Theorem, Degree Theory, etc.);
- when possible, exploit the variational structure of the problem, and find solutions as stationary points of some "energy" functional [13].

We will focus on the first approach, since the latter is limited to equations derived from a variational principle, most notably semilinear elliptic equations. Furthermore, most of the existing results of Numerical Analysis for nonlinear problems are cast in such setting. This setting is quite abstract, in fact it encompasses a large class of maps between functional spaces, such as differential and integral equations. As a result the results obtained are very general although we will mainly be concerned on differential problems.

We consider nonlinear problems depending on a parameter $\mu \in \mathcal{D} \subset \mathbb{R}^p$ in the form: find $u \in X$ such that:

$$\langle F(\mu, u), v\rangle = 0 \qquad \forall v \in Y, \tag{1}$$

where $F : \mathcal{D} \times X \to Y'$ is a map, X and Y are Banach spaces and the angled parenthesis denote the duality pairing between Y and its dual space Y'. The family of parameter-dependent solutions $\{u(\mu)\}_{\mu\in\mathcal{D}}$ forms a subset of X, and for some parameter values there may be some qualitative changes in the structure of the solutions. For instance, there may be a loss of uniqueness of the solution, usually followed by a change in the stability properties under infinitesimal or finite perturbations, or the transition from steady state to time dependent solutions, just to mention a few possibilities.

Rigorously, we say that (μ^*, u^*) is a *bifurcation point* for (1) if there exist at least two sequences $(\mu^1_m, u^1_m) \subset \mathcal{D} \times X$, $(\mu^2_n, u^2_n) \subset \mathcal{D} \times X$ such that:

i) $(\mu^1_m, u^1_m) \neq (\mu^2_n, u^2_n) \qquad \forall m, n \in \mathbb{N}$;

ii) $\langle F(\mu^i_m, u^i_m), v\rangle = 0 \qquad \forall v \in Y, \text{ for } i = 1, 2$;

iii) $(\mu^1_m, u^1_m) \to (\mu^*, u^*) \qquad \text{and} \qquad (\mu^2_n, u^2_n) \to (\mu^*, u^*) \qquad \forall m, n \in \mathbb{N}$.

We define the (Fréchet) differential of F with respect to the variable u, defined if exists a linear operator $D_uF(\mu, u) \in \mathscr{L}(X, Y')$ such that:

$$\lim_{\|w\|_X\to 0} \frac{\|F(u + w) - F(u) - D_uF(\mu, u)[w]\|_{Y'}}{\|w\|_X} \qquad \forall w \in X. \tag{2}$$

With $\mathscr{L}(X, Y)$ we denote the set of all linear continuous operators from X to Y, and we denote with $D_uF(\mu, u)[v] \in Y'$ the action on $v \in X$ of the differential of F with respect to the variable u evaluated at the point (μ, u). Many important results in Nonlinear Analysis depend on the existence of partial derivatives of the operator F.

The simplest case is when the partial derivative is an isomorphism from X to Y', and we write $D_uF(\mu, u) \in \mathrm{Iso}(X, Y')$. In this case we say that (μ^*, u^*) is a *regular solution* (or a *nonsingular solution*) of (1), and its existence and uniqueness are ensured locally by the Implicit Function Theorem (IFT) [1]. The fact that the map F is invertible with invertible differential means practically that it is possible to express locally u as a function of μ. This is possible only if the differential $D_uF(\mu^*, u^*)$ is invertible, since in this case $D_uF(\mu^*, u^*)^{-1} : Y' \to X$ maps neighbourhoods into neighbourhoods.

In contrast, if (μ^*, u^*) is a bifurcation point of (1), then $D_uF(\mu^*, u^*)$ is not invertible, and it is impossible to express u as a function of μ directly. In particular, when $D_uF(\mu^*, u^*) \notin \mathrm{Iso}(X, Y')$ two possibilities arise:

$$\begin{aligned} &D_\mu F(\mu^*, u^*) \notin \mathrm{R}(D_uF(\mu^*, u^*)) && \text{limit point;} \\ &D_\mu F(\mu^*, u^*) \in \mathrm{R}(D_uF(\mu^*, u^*)) && \text{bifurcation point;} \end{aligned} \tag{3}$$

where $D_\mu F(\mu^*, u^*)$ is the partial derivative of F with respect to the parameter μ. For both cases we suppose also that $D_uF(\mu^*, u^*)$ has a closed range and satisfies:

$$\dim\ker(D_uF(\mu^*, u^*)) = \dim(\mathrm{R}(D_uF(\mu^*, u^*))^\perp) \tag{4}$$

where $\mathrm{R}(D_uF(\mu^*, u^*))^\perp$ is the orthogonal complement to $\mathrm{R}(D_uF(\mu^*, u^*))$, such that Y' can be written as a direct sum:

$$Y' = \mathrm{R}(D_uF(\mu^*, u^*)) \oplus \mathrm{R}(D_uF(\mu^*, u^*))^\perp, \tag{5}$$

and $\ker F = \{v \in X \text{ s.t. } F(v) = 0\}$ is the kernel of the operator F. We mentioned that the Implicit Function Theorem plays a fundamental role when studying the properties of the solution set of equation (1). The form of the IFT that we consider here is the following: let $F \in C^k(\Lambda \times U, Y')$ with $k \geq 1$, $\Lambda \subseteq \mathcal{D}, U \subseteq X$, and suppose that $(\mu^*, u^*) \in \Lambda \times U$ are such that

$$\langle F(\mu^*, u^*), v\rangle = 0 \quad \forall v \in Y \qquad D_u F(\mu^*, u^*) \in \mathrm{Iso}(X, Y'), \tag{6}$$

that is, u^* solves the problem (1) with the parameter μ^* and the u partial derivative of F is invertible in a neighbourhood of (μ^*, u^*). In particular, there exist neighbourhoods $\Xi \subseteq \mathcal{D}$ of μ^* and $V \subseteq X$ of u^* and a map $\gamma \in C^k(\Xi, X)$ such that:

$$\langle F(\mu, \gamma(\mu)), v\rangle = 0 \qquad \forall v \in V, \forall \mu \in \Xi. \tag{7}$$

In the next paragraphs, the Implicit Function Theorem will be frequently applied to parametrize solution sets in a neighbourhood of a singular point and allow the following of solution branches.

2.1. Limit points

We consider limit points (μ^*, u^*) such that the differential $D_u F(\mu^*, u^*)$ is compact and has a one-dimensional kernel, $\dim\ker D_u F(\mu^*, u^*) = 1$. Let $\varphi_0 \in X$ be a basis for the kernel of $D_u F(\mu^*, u^*)$. In this case the differential map is not an isomorphism, and the Implicit Function Theorem cannot be applied directly to parametrize u with respect to μ. However, introducing a new parameter $s \in [-\varepsilon, \varepsilon]$, the solution set can be parametrized in a neighbourhood of the limit point:

$$\begin{aligned} \mu(s) &= \mu^* + \xi(s) \\ u(s) &= u^* + s\varphi_0 + \gamma(\xi(s), s) \end{aligned} \tag{8}$$

for a map $\xi : [-\varepsilon, \varepsilon] \to \mathbb{R}$. The existence of the map ξ is important both for building an approximation strategy for the fold points and in the following of the solution branch in a neighbourhood of a fold point.

2.2. Bifurcation points

In the case of simple bifurcation points multiple branches of solution issue from a principal branch, hence the study of the solutions set is a particularly delicate task. A classical tool in this case is provided by the Lyapunov-Schmidt reduction, that allows to split the problem in an appropriate way such that the IFT can be applied on some subsets of X and Y'.

At a simple bifurcation point, we require that the differential $D_u F(\mu^*, u^*)$ has a closed range and a one-dimensional kernel, $\dim\ker D_u F(\mu^*, u^*) = 1$ (although these hypotheses can be relaxed, see [14]). Let now $\varphi_0 \in X$ be a basis for the kernel of $D_u F(\mu^*, u^*)$ and $\varphi_0^* \in Y'$ a basis for the topological complement of the range of $D_u F(\mu^*, u^*)$. Next, let $\mathrm{R}(A)$ be the image of X under the operator A. We introduce a projection on the range of the differential $P : Y' \to \mathrm{R}(D_u F(\mu^*, u^*))$ and its complementary projection $Q = I - P : Y' \to (\mathrm{R}(D_u F(\mu^*, u^*)))^\perp$:

$$\begin{cases} Pw = \langle w, \mathcal{R}\varphi_0^*\rangle \varphi_0^* & \forall w \in Y' \\ Qw = w - \langle w, \mathcal{R}\varphi_0^*\rangle \varphi_0^* & \forall w \in Y', \end{cases} \tag{9}$$

where $\mathcal{R} : Y' \to Y$ is the Riesz representation operator. The Lyapunov-Schmidt reduction consists in the projection of equation (1) both on the range of $D_u F(\mu^*, u^*)$ and on its complementary set. With the latter projection we have the *auxiliary equation*

$$\langle QF(\mu, u), v\rangle = 0 \qquad \forall v \in Y \tag{10}$$

for which the Implicit Function Theorem can be applied to introduce two new parameters ξ, s and a map $\gamma : [-\varepsilon_\xi, \varepsilon_\xi] \times [-\varepsilon_s, \varepsilon_s] \to \mathrm{R}(D_u F(\mu^*, u^*))$ such that we can express

$$\begin{aligned} \mu &= \mu^* + \xi \\ u &= u^* + \xi\varphi_0 + \gamma(\xi, s\varphi_0). \end{aligned} \tag{11}$$

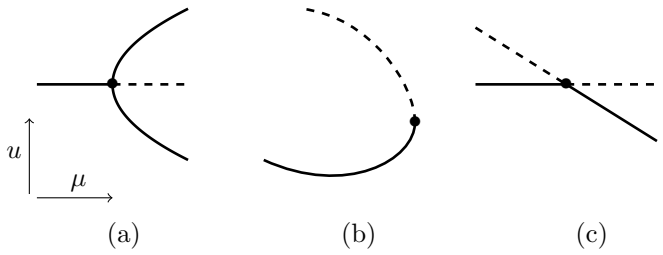

Figure 1: Schematic solutions' behaviour at singular points. Dashed branches are unstable. (a): pitchfork bifurcation, (b): fold, (c): transcritical bifurcation.

Then, applying the dual projection we obtain the *bifurcation equation*:

$$\langle F(\mu^* + \xi, u^* + s\varphi_0 + \gamma(\xi, s\varphi_0)), \mathcal{R}\varphi_0^* \rangle = 0 \tag{12}$$

that can be solved locally for ξ, s and φ_0. Notice how the careful use of the two projections allows to recover the parameter-solution mapping, for sufficiently small values of the parameter s.

To give an intuitive understanding of the solutions set behaviour when a singular point is approached, we draw in figure 1 an illustration of some common bifurcation and fold points.

We remark that for both fold and bifurcation points, the Implicit Function Theorem allows to write a parametrization such that in a neighbourhood of the singular point (μ^*, u^*) there exists a function ξ such that ξ changes sign if (μ^*, u^*) is a bifurcation point, and has a minimum or maximum in zero without changing sign if (μ^*, u^*) is a fold point. This fact will be useful in the applications. The existence of ξ is important also for the practical computation of solution branches near a singular point, as we shall see in section 3.1.

2.3. Branching of periodic solutions

The last ingredient of Bifurcation Theory that we consider is the bifurcation from periodic solutions. The abstract problem (1) can be regarded as a particular case of the following time-dependent problem:

$$\omega \partial_t u - F(\mu, u) = 0 \tag{13}$$

for which the solution u does not depend on time. In equation (13) we suppose that the periodic solutions branching from the steady state solutions have period ω. The abstract theory for the time-independent problem (1) could be recovered with some modifications, but from the practical viewpoint it is usually easier to rely on the *Hopf bifurcation theorem* [1]. According to Hopf result, it is still sufficient to study the linearized operator $D_u F(\mu, u)$, in particular the form:

$$\langle \omega \partial_t u - D_u F(\mu, u)[w], v \rangle = 0 \qquad \forall v \in Y \tag{14}$$

and we assume that

i) $D_u F(\mu^*, u^*)$ is nonsingular and has a pair of simple, purely imaginary eigenvalues: $\pm \mathrm{i}\omega^*$;

ii) there are no other eigenvalues for $D_u F(\mu^*, u^*)$ in the form $\pm \mathrm{i}k\omega^*$, with $k \in \mathbb{N}$;

then, (μ^*, ω^*, u^*) is a bifurcation point for periodic solutions, with period ω^*.

Finally, we remark that the families of singular points briefly described in this section are not the only possibile cases in the framework of Nonlinear PDEs, but we focused on these classes because of their relevance in applications (e.g. Nonlinear Fluid Mechanics).

3. A discrete setting

Many of the results of Nonlinear Analysis exposed in section 2 apply in the finite-dimensional case, providing a rigorous foundation for the numerical approximation and strategies for the actual computation. A milestone in the numerical approximation of nonlinear problems is the theory developed by Brezzi, Rappaz and Raviart [15, 16, 17] (BRR), to which we frequently refer in the rest of the section. A good introduction to this theory is provided by [14]. In this section we rely on BRR theory for the case of Reduced Order Modelling that has many analogies but also some specialities over the traditional Galerkin approximation methods, the most relevant being that the conventional general purpose bases with local support are replaced with problem dependent bases endowed with global support. Referring for instance to [5] for a detailed overview on the Reduced Order Methods of interest, we proceed recalling some basic facts about ROMs and POD that will be useful in the following.

3.1. Truth approximation

The first step of a ROM (RB, POD, ...) is to extract information from a series of known solutions characterized by different values of the parameters. Usually these solutions need to be approximated through a computational method commonly called *truth approximation* in the community, which can be any classical discretization method for PDEs (such as the Finite Element Method, Spectral Element Method, Finite Volume Method, etc). We focus here on (Petrov-)Galerkin projection methods, that recover much of the setting in the original problem (1). The approximated solution $u^{\mathcal{N}}$ is obtained as projection of the full solution u on a finite dimensional subset $X^{\mathcal{N}}$ of X (the superscript $\mathcal{N}$ denotes the dimension of $X^{\mathcal{N}}$), and the test function space is replaced with a finite dimensional subset $Y^{\mathcal{N}}$ of Y. As a result, the truth approximation of problem (1) reads: find $u^{\mathcal{N}} \in X^{\mathcal{N}}$ such that

$$\langle F(\mu, u^{\mathcal{N}}), v\rangle = 0 \qquad \forall v \in Y^{\mathcal{N}}. \tag{15}$$

Throughout this work we assume that $u \mapsto F(\lambda, u)$ is a nonlinear map, and to obtain a linear algebra problem from equation (15) we need to introduce a suitable approximation scheme. For instance, a possible approach consists in constructing a sequence of successive approximations $\{u_k^{\mathcal{N}}\}$ such that the k-th element of the sequence solves a linearized problem. Let us write a decomposition of F into its linear and nonlinear parts, L and N:

$$F(\mu, u) = \mu L(\mu)u + N(\mu, u), \tag{16}$$

then two popular fixed point linearization strategies are as follows.

Picard iterations in this case we rewrite problem (15) as an approximation for a fixed point of F. At iteration $k+1$, solve for $u_{k+1}^{\mathcal{N}}$:

$$\langle L(\mu)u_{k+1}^{\mathcal{N}} + N(\mu, u_k^{\mathcal{N}}), v\rangle = 0 \qquad \forall v \in Y^{\mathcal{N}} \tag{17}$$

Banach-Caccioppoli theorem ensures that a fixed point exists and is unique provided that F is a contraction map on a sufficiently large ball $B(u_0) \subset X$, i.e. there exists a constant $l \in (0, 1)$ such that

$$\|F(\mu, u) - F(\mu, w)\|_{Y'} \leq l\|u - w\|_X \qquad \forall u, w \in B(u_0) \tag{18}$$

and that $u_k^{\mathcal{N}} \to u^{\mathcal{N}}$ in the norm of X.

Newton-Kantorovich iterations in this case also the first differential of F is exploited:

$$\begin{cases} \langle F(\mu, u_k^{\mathcal{N}}) + D_u F(\mu, u_k^{\mathcal{N}})[w_{k+1}^{\mathcal{N}}], v\rangle = 0 \qquad \forall v \in Y^{\mathcal{N}} \\ u_{k+1}^{\mathcal{N}} = u_k^{\mathcal{N}} + w_{k+1}^{\mathcal{N}}. \end{cases} \tag{19}$$

Newton-Kantorovich iterations are convergent if a Lipshitz condition is satisfied on the differential of F:

$$\|D_u F(\mu, u)[v] - D_u F(\mu, u)[w]\|_{Y'} \leq l\|v - w\|_X \qquad \forall u, v, w \in S \tag{20}$$

where S is an appropriate open convex subset of X.

We remark that convergence results are available [18] in both cases also when the operator F is being approximated by a sequence $\{F_h\} \subset \mathrm{Iso}(X, Y')$ such that consistency is preserved, $F_h(\mu, u_k^{\mathcal{N}}) \to F(\mu, u^{\mathcal{N}})$. This fact is of particular importance since often in Numerical Analysis the continuous operators are replaced with discrete operators e.g. exact integrals are approximated by quadrature formulas.

After the linearization, it is possible to derive a linear algebra problem simply by expanding the approximate solution at k-th iteration: $u_k^{\mathcal{N}} = \sum_{i=1}^{\mathcal{N}} u_{k,i}^{\mathcal{N}} \varphi_i$ where $\{\varphi_i\}$ and $\{\psi_j\}$ are basis sets respectively for $X^{\mathcal{N}}$ and $Y^{\mathcal{N}}$. Replacing the series expansion for $u^{\mathcal{N}}$, and choosing $v = \psi_j$, we have a linear algebra problem:

$$\mathrm{A}_k \boldsymbol{u}_k = \boldsymbol{b}_k, \tag{21}$$

for each step k, whose terms have the form:

$$\mathrm{A}_{k,ij} = \langle L(\mu)\varphi_i + D_u F(\mu, u_{k-1}^{\mathcal{N}})[\varphi_i], \psi_j \rangle \qquad \boldsymbol{u}_{k,i} = u_{k,i}^{\mathcal{N}}, \quad \boldsymbol{b}_{k,i} = \langle N(\mu, u_k^{\mathcal{N}}, \psi_i \rangle. \tag{22}$$

System (21) can then be solved by the usual techniques (see for instance [19]). However, our discussion of the algebraic problem is only illustrative, and in the real practice much more complex discretization methods are used. In particular, for medium and large sized problems it is mandatory to distribute the computation between many CPUs, for instance via a Domain Decomposition Method (see [20] or [21]).

The presence of multiple solutions for a given parameter value requires some care to ensure that the computed solutions belong to the same branch, as in some cases the computation could oscillate between two solutions without converging. A popular method to deal with this problem is the *continuation method* [22]. In its simplest version, it consists of two steps:

Step 1. predictor: starting from a known solution (μ^k, u^k), compute a prediction $(\tilde{\mu}^k, \tilde{u}^k) = (\mu^k + \Delta\mu^k, u^k + \Delta u^k)$ on the tangent space to (μ^k, u^k) by solving for Δu^k the problem:

$$D_\mu F(\mu^k, u^k)\Delta\mu^k + D_u F(\mu^k, u^k)\Delta u^k = 0, \tag{23}$$

where $\Delta\mu^k$ is given (and sufficiently small);

Step 2. corrector: starting from the prediction $(\tilde{\mu}^k, \tilde{u}^k)$, solve iteratively the original problem (1) imposing the additional constraint that the solutions are orthogonal to the tangent space at (μ^k, u^k):

$$\begin{cases} D_\mu F(\mu_i^k, u_i^k)\Delta\mu_i^k + D_u F(\mu_i^k, u_i^k)\Delta u_i^k = -F(\mu_{i-1}^k, u_{i-1}^k) \\ (\tilde{\mu}^k, \Delta\mu_i^k)_{\mathcal{D}} + (\tilde{u}^k, \Delta u_i^k)_X = 0 \\ \mu_i^k = \mu_{i-1}^k + \Delta\mu_i^k \qquad u_i^k = u_{i-1}^k + \Delta u_i^k \end{cases} \tag{24}$$

with the starting condition that the sequence of approximant solutions issues from the predictor solution:

$$\mu_0^k = \tilde{\mu}^k \qquad u_0^k = \tilde{u}^k. \tag{25}$$

3.2. Sampling

The information needed for building a reduced order approximation is obtained by constructing a set of properly selected truth solutions $\{u^{\mathcal{N}}(\mu^i)\}_{i=1}^N \subset X^{\mathcal{N}}$ computed for a suitable sequence in the parameter space $\{\mu^i\}_{i=1}^N \subset \mathcal{D}$. There exist different methods[1] for identifying a sequence $\{\mu^i\} \subset \mathcal{D}$, often based on a "worst case" criterion, which given an initial sequence $\{\mu^i\}_{i=1}^k$, aims at searching for the parameter μ^{k+1} whose solution is the worst approximated one within the space spanned by the snapshots $\mathcal{S}_k = \{u^{\mathcal{N}}(\mu^i)\}_{i=1}^k$. Then, the new solution $u^{\mathcal{N}}(\mu^{k+1})$ is added to the snapshots space $\mathcal{S}_{k+1} = \mathcal{S}_k \cup \{u^{\mathcal{N}}(\mu^{k+1})\}$ and the algorithm is restarted until a suitable stopping criterion is satisfied.

Two of the most popular sampling methods based on a worst case strategy are the following.

[1] We refer the interested reader to [5] for a brief overview on the history of the Reduced Basis Method and a review of many sampling techniques.

Greedy Algorithm Very widely spread in the ROM community, we refer to [23] for a comprehensive review and to [24, 25] for some relevant convergence estimates. Suppose that are given $k-1$ linearly independent snapshots $\mathcal{S}_{k-1} = \{u_i^{\mathcal{N}}\} \subset X^{\mathcal{N}}$. Then the k-th snapshot is computed as solution characterized by the parameter value $\mu^k \in \mathcal{D}$ such that

$$\mu^k = \arg\min_{\mu\in\mathcal{D}} \left\| u^{\mathcal{N}}(\mu^k) - \Pi_{\mathcal{S}_k} u^{\mathcal{N}}(\mu^k) \right\|_X , \tag{26}$$

where $\Pi_{\mathcal{S}_k} : X^{\mathcal{N}} \to \mathcal{S}_k$ is the projector on the space generated by $\mathcal{S}_k$. Then $u^{\mathcal{N}}(\mu^k)$ is used to enrich the snapshots space: $\mathcal{S}_k = \mathcal{S}_{k-1} \cup \{u^{\mathcal{N}}(\mu^k)\}$. To simplify the computation of the minimizer μ^k of equation (26), usually two additional constraints are imposed:

- the parameter space $\mathcal{D}$ is replaced by a finite subspace Ξ;
- the projection error $\|u^{\mathcal{N}}(\mu^k) - \Pi_{\mathcal{S}_k} u^{\mathcal{N}}(\mu^k)\|_X$ is replaced by some parameter-dependent estimate $\Delta(\mu)$ such that
$$c_\Delta \Delta(\mu^k) \le \left\| u^{\mathcal{N}}(\mu^k) - \Pi_{\mathcal{S}_k} u^{\mathcal{N}}(\mu^k) \right\|_X \le C_\Delta \Delta(\mu^k) \tag{27}$$
for $c_\Delta, C_\Delta > 0$. This estimate is introduced to avoid computing the explicit solution $u^{\mathcal{N}}(\mu^k)$ and its projection for all $\mu^k \in \Xi$. We discuss on section 3.9 a possible strategy for obtaing an expression for $\Delta(\mu)$.

These two approximations together give origin to the family of *weak greedy* algorithms.

It is easy to verify that the snapshot spaces produced by Greedy Algorithms are *hierarchical spaces*, that is $\mathcal{S}_1 \subset \mathcal{S}_2 \subset \cdots \subset \mathcal{S}_k$. This property has a great practical importance since it allows to enrich the set of snapshots without throwing away the previously computed solutions, since the new one is simply added to the previous set.

Centroidal Voronoi Tessellation Centroidal Voronoi Tessellation (CVT) has been introduced by Du and Gunzburger [26] as sampling strategy for Reduced Order Modelling. We need to define three preliminary concepts:

- a collection of subsets $\{\mathcal{V}_i\}_{i=1}^k$ of $\mathcal{D}$ is called a *tessellation* of $\mathcal{D}$ if they are disjoint $\mathcal{V}_i \cap \mathcal{V}_j = \emptyset$ if $i \neq j$, and covering $\bigcup_{i=0}^k \mathcal{V}_i = \mathcal{D}$;
- we endow the parameter space $\mathcal{D}$ with a distance function $\mathfrak{d}$ induced by the norm of X between the corresponding solutions:
$$\mathfrak{d}(\mu^1, \mu^2) = \|u^{\mathcal{N}}(\mu^1) - u^{\mathcal{N}}(\mu^2)\|_X. \tag{28}$$
Then, given a set of points $\{\mu^i\} \subset \mathcal{D}$, we define the i-th *Voronoi region* $\mathcal{V}_i \subset \mathcal{D}$ as:
$$\mathcal{V}_i = \{\nu \in \mathcal{D} \text{ such that } \mathfrak{d}(\nu, \mu^i) \le \mathfrak{d}(\nu, \mu^j)\, \forall j \neq i\}, \tag{29}$$
and $\{\mu^i\} \subset \mathcal{D}$ are a set of given *generating points*.
- suppose that a density function $\varrho : \mathcal{D} \to \mathbb{R}^+$ is given on the parameter space. Then we define the mass centroid $\xi \in \mathcal{D}$ of a subset $\mathcal{U}$ of $\mathcal{D}$ as the barycenter of $\mathcal{U}$, namely:
$$\xi = \arg\min_{\nu\in\mathcal{U}} \int_{\mathcal{U}} \mathfrak{d}(\upsilon, \nu)\varrho(\upsilon)\, \mathrm{d}\upsilon. \tag{30}$$

Note that in general the generating points defining the Voronoi regions in (29) and the mass centroids defined in equation (30) do not coincide, but if this happens, the tessellation defined by the generating points $\{\mu^i\}$ is named *Centroidal Voronoi Tessellation* (CVT).

A possible choice for the density function is the projection error as defined for the Greedy algorithm:

$$\varrho(\mu) = \left\| u^{\mathcal{N}}(\mu) - \Pi_{X^N} u^{\mathcal{N}}(\mu) \right\|_X , \tag{31}$$

where $\Pi_{X^N} : X \to X^N$ is the projection operator from X to its subspace X^N (although in general X^N need not be a subspace of X; in this case supposing that both X and X^N are imbedded in an "environment" set E, the projection Π_{X^N} can be defined as the element in X^N that minimizes the distance in E from a given element in X). Then, if we define the "weight" of each Voronoi region $\mathcal{V}_i$ as

$$\mathscr{W}_i = \int_{\mathcal{V}_i} \mathfrak{d}(\nu, \xi_i)\varrho(\nu)\,\mathrm{d}\nu \tag{32}$$

it can be shown that the CVT of dimension k of the parameter set is defined by the generating points $\{\mu^i\}_{i=1}^k$ that minimize the "total weight" functional:

$$\mathscr{F} = \sum_{i=1}^{k} \mathscr{W}_i. \tag{33}$$

In this sense the CVT is a best approximation sampling method, and is often combined with the POD defined in section 3.3 to form the CVOD method [27], popular in the ROM community. To help the intuition on the meaning of the symbols entering in the integral (32), we refer to the sketch on figure 3.2. Illustrative picture of some basic quantities entering in the definition of the CVT.

For computational purposes, some additional approximations are introduced, namely:

- as in the weak Greedy algorithm, $\mathcal{D}$ is replaced by Ξ and the projection error (31) by a suitable estimator $\Delta(\mu)$;
- instead of computing all the generating points $\{\mu^k\}$ at each iteration, the first k points are kept fixed and only the $k+1$-th generating point is computed as minimizer of the total weight functional (33);
- another possibility is to consider the Delaunay triangulation (dual of the Voronoi diagram) and take as next point the barycenter of the triangle with the largest weight $\mathscr{W}_i$. This technique has been introduced by Iollo et al. [28] for Delaunay samplings.

We remark that the last two strategies allow to obtain a hierarchical sampling.We sketch on figure 2 an example step of CVT sampling procedure. The enhanced approximation properties of the RB space are visualized by the reduction in density of the isolines of the weight function ϱ.

In the description of Greedy and CVT sampling methods, we implied more or less explicitly that the snapshot space contains only solutions of the original problem (1) for different parameter values, as done in the classical Lagrangean interpolation theory. For this reason, ROMs based on such spaces are often called *Lagrange ROM*.

However, sampling methods are by no means restricted to Lagrange spaces, and can be applied to sample not only the solutions manifold, but also the tangent bundle to the solutions manifold (originating *Hermite ROMs*, [29]) or higher order derivative spaces (originating *Taylor ROMs*, [11]). Note that the solutions manifold and its tangent spaces are all embedded in the original space X, so there is no need to change the discretization method for sampling Hermite or Taylor spaces.

Operatively, the tangent bundle is sampled by linearizing the original problem (1) and chosing among all the possible small linear increments from the solution (μ^k, u^k) the one closer to the solutions manifold. This is equivalent in finding $Tu^k \in X$ such that:

$$\langle D_u F(\mu^k, u^k)[Tu^k], v\rangle = -\langle D_\mu F(\mu^k, u^k), v\rangle \qquad \forall v \in Y. \tag{34}$$

For higher order sampling methods it is sufficient to impose that higher order Taylor expansions of the original problem (1) are canceled out. For instance, a second order expansion would be in the form: find $T^2u^k \in X$ such that:

$$\langle D_{uu} F(\mu^k, u^k)[Tu^k, T^2u^k], v\rangle = -\langle 2D_{u\mu} F(\mu^k, u^k)[Tu^k] + D_{\mu\mu} F(\mu^k, u^k), v\rangle \qquad \forall v \in Y, \tag{35}$$

where $D_{uu}F(\mu, u) \in \mathscr{L}(X, X, Y')$ is a bilinear operator taking two elements of X to Y', or equivalently it can be seen as a linear operator from X to the space of linear operators from X to Y', $D_{uu}F(\mu, u)[w, \cdot] \in \mathscr{L}(X, Y')$. Similar considerations hold for the other second order differentials, $D_{u\mu}F(\mu, u)$ and $D_{\mu\mu}F(\mu, u)$.

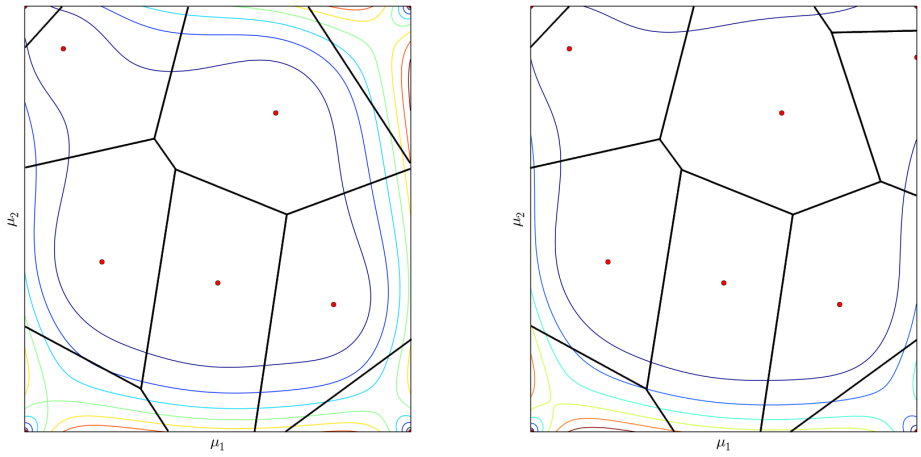

Figure 2: One step of CVT sampling on a two parameters test case. The red dots show the sampled points, the isolines refer to the weight function $\varrho(\nu)$. The black contours are the boundaries of each Voronoi region. On the left there is a representation of a sampling set in the parameters space consisting of 9 points. On the right the same set has been enriched by applying one step of the CVT algorithm. Note the new region on the upper right part of the figure, and how the new sampled point has considerably stretched the region confined inside the isoline of minimum value.

We remark that Hermite spaces may be a particularly good choice when attempting a ROM in the case of parametrized problems with multiple solutions. In fact, as exposed in section 3.1, a popular method to force the computed solutions to lay on the same branch is the continuation method, that requires at each iteration a solution of the tangent problem (34). Our claim is that Hermite sampling is an interesting choice in this case since it would simplify the computation of solutions branches during the online phase:

Step 1. in the predictor phase, impose the solution increment Δu^k of equation (23) to lay on the space spanned by the tangent reduced basis set;

Step 2. during the correction phase, orthogonalize the reduced basis set with respect to the tangent basis set, in order to fulfill by construction the constraint on the orthogonality of the increments (24).

3.3. Proper Orthogonal Decomposition

When dealing with time-dependent problems, it is preferable not to add all the snapshots of a time-dependent run to the snapshots space $\mathcal{S}_k$, otherwise the Reduced Basis Space will likely be too large and many of the basis will be almost parallel, making the online projection phase an ill conditioned problem. A common technique is to extract the "most significant" modes of a time sequence using a RB approximation in combination with a Proper Orthogonal Decomposition [30]. There are many alternative ways to compute the POD modes of a sequence of snapshots $\{u_i^{\mathcal{N}}\} \subseteq X^{\mathcal{N}}$. Here we focus on the method based on the eigenvalues of the correlation matrix [30]. The entries of the correlation matrix $\mathbb{C} \in \mathbb{R}^{N \times N}$ are computed as

$$\mathbb{C}_{ij} = (u_i^{\mathcal{N}}, u_j^{\mathcal{N}})_V, \tag{36}$$

then, if (λ_i, ψ_i) is a couple eigenvalue-eigenvector of $\mathbb{C}$, each basis vector is computed as

$$\zeta_i = \sum_{k=1}^{N} \psi_{i,k} u_k^{\mathcal{N}} \tag{37}$$

where $\psi_{i,k}$ denotes the k-th component of the i-th eigenvalue. The POD modes obtained are automatically orthogonal, but not normal in general. The eigenvalue λ_i associated to each POD mode is related to the fraction of energy stored in the corresponding mode.

A remarkable property of the space generated by the POD bases is that it minimizes the projection error in the norm of the space V of the snapshots used for the construction of the correlation matrix in equation (36):

$$X_{\text{POD}}^N = \text{span}\{\zeta_i\}_{i=1}^N = \arg\min_{X^N \subset X, \dim X = N} \sum_{i=1}^N \left\| u_i^{\mathcal{N}} - \Pi_{X^N} u_i^{\mathcal{N}} \right\|_V^2 . \tag{38}$$

In this sense, the POD modes exhibit a best approximation property. It can be shown (see for instance [30]) that the spaces spanned by the POD are hierarchical once the snapshots are fixed.

The sampling algorithms presented in section 3.2 are frequently combined with a POD for parametrized time-dependent problems. If a Greedy sampling algorithm is used to select the parameters, and the POD is used to recover the most relevant time snapshots, the sampling procedure is called POD-Greedy and we refer to [31] and [32] for details. To clarify the sampling process of time-dependent parametrized problems, we report in Algorithm 1 a possible implementation strategy of the POD-Greedy algorithm.

Algorithm 1 A POD-Greedy strategy for the sampling of parameter dependent evolution problems.

1: **repeat**
2: find μ^{i+1} such that $\mu^{i+1} = \arg\max_{\mu \in \Sigma} \int_0^T \|u^{\mathcal{N}}(\mu^{i+1}) - \Pi_{\mathcal{S}_i} u^{\mathcal{N}}(\mu^{i+1})\|_X \, \mathrm{d}t$
3: compute a sequence of snapshots $\{u_{i+1,k}\}_{k=0}^r$ for the time dependent problem related to the parameter μ^i
4: compute the l POD modes of the sequence $\{\varphi_{i+1,k}\}_{k=1}^l$ such that the retained energy is above a prescribed ratio
5: orthogonalize the time modes with respect to the previous basis sets: $\zeta_{i+1,k} = \varphi_{i+1,k} - \Pi_{\mathcal{S}_i \cup_{j=1}^k \varphi_{i+1,j}} \varphi_{i+1,k}$
6: add the new basis functions $\{\zeta_{i+1,k}\}_{k=1}^l$ to the basis set: $\mathcal{S}_{i+1} = \mathcal{S}_i \cup \{\zeta_{i+1,k}\}$
7: **until** $\max_{\mu \in \Sigma} \Delta(\mu) < \texttt{tol}$

Lastly, we remark that when sampling time dependent problems in view of a POD application, it is important to make sure that the sampling rate is sufficiently high so that the desired time harmonics are well resolved. Usually the sampling rate required for a POD is higher than the minimum defined by the Nyquist criterion, since spurious correlation of noisy data may affect the POD modes [33].

3.4. Construction of the Reduced Basis

The snapshots obtained through the sampling technique exposed in section 3.2 will not in general form a basis for the Reduced Basis spaces. Although the POD could be performed on the entire snapshots space $\mathcal{S}_N$ to compute a series of orthogonal modes with decreasing energy, this approach is avoided for medium and large problems because of the high number of operations involved (due to eigenpairs computation of big and full correlation matrices).

A much cheaper method to build an orthonormal basis starting from a set of linearly independent vectors is the Gram-Schmidt process, as discussed in [5].

Given a set of N_{sn} snapshots $\{u^{\mathcal{N}}(\mu^i)\}_{i=1}^{N_{\text{sn}}}$, we compute a basis function ζ_i for each[2] $i = 1, \dots, N$ by means of the following orthonormalization procedure:

$$\zeta_i = u^{\mathcal{N}}(\mu^i) - \sum_{j=1}^{i} \zeta_j (u^{\mathcal{N}}(\mu^i), \zeta_j)_X \qquad \zeta_i \leftarrow \frac{\zeta_i}{\|\zeta_i\|_X}, \tag{39}$$

[2]Note that in general $N \neq N_{\text{sn}}$.

where $(\cdot,\cdot)_X$ is the inner product in the space X and the last normalization step may or may not follow, as we will discuss in paragraph 3.7.

We now introduce the reduced basis approximation space X^N simply as $X^N = \text{span}\{\zeta_i\}_{i=1}^N$ that will be used for the online approximation.

3.5. Online approximation

After a suitable Reduced Basis is constructed for the problem of interest, one can achieve a substantial computational reduction by constructing a new approximation scheme based on a (Petrov-)Galerkin projection on spaces X^N, Y^N whose dimension is much lower than in the truth approximation case:

$$N = \dim X^N \ll \mathcal{N} = \dim X^{\mathcal{N}}. \tag{40}$$

For the sake of simplicity, we limit ourselves to the case where the test Y^N and trial X^N spaces have the same dimension, but this constraint can be relaxed.

As done in section 3.1, we may simply apply the Galerkin Projection on the Reduced Basis spaces so that the online problem reads: find $u^N \in X^N$ such that:

$$\langle F(\mu, u^N), v\rangle = 0 \qquad \forall v \in Y^N. \tag{41}$$

Again, a suitable linearization of the operator is required in order to obtain a Linear Algebra problem that can be solved with the usual methods, except that this time the linear system has much smaller dimension, allowing for very fast solution.

Also, since the Reduced Basis matrices are computed offline and assembled online, there is the necessity of introducing a rapid way to assemble parameter-dependent matrices. The simplest case is when the map $F(\cdot, u) : \mathcal{D} \to F(\mu, u)$ is affine with respect to the parameters. Then, the operator can be written as

$$F(\mu, u) = \mu L(\mu)u + \sum_{i=1}^{q} \Theta^i(\mu) N^i(u) \tag{42}$$

for appropriate functions $\Theta^i : \mathcal{D} \to \mathbb{R}$, and the matrix arising from the linearization of F and a consequent evaluation of the linearized map on the Reduced Basis spaces can be written as:

$$A^N = \sum_{i=1}^{q} \Theta^i(\mu) A^i. \tag{43}$$

For problems lacking affine parameter dependence, a possible way to recover the decomposition (43) during the online phase is applying the Empirical Interpolation Method [34], still a very active research area.

3.6. Geometric parametrization

The parameter μ has a particular importance when it[3] describes the geometry of the domain on which the problem is posed. The main idea behind this procedure is to refer each of the infinitely many possible geometric configurations to a common "original" domain. This way, the necessity of constructing a new mesh for each computation vanishes, and the problem can be set only on the original domain. The solution obtained on the original domain can then be mapped to the "true" domain.

In the following, we will denote the reference domain simply by $\widehat{\Omega}$, whereas the parametrized domain, depending on a parameter that for simplicity we denote μ will be referred to as $\Omega(\mu)$. Following [5], we now express the affine transformation between the reference domain and the parametrized geometry as:

$$\boldsymbol{x} = \mathcal{T}^{\text{aff}}(\widehat{\boldsymbol{x}}, \mu) = C(\mu) + G(\mu)\widehat{\boldsymbol{x}} \quad \forall \boldsymbol{x} \in \Omega(\mu), \forall \widehat{\boldsymbol{x}} \in \widehat{\Omega}, \tag{44}$$

[3] or at least some of its components

where $C(\mu)$ is a displacement vector and $G(\mu)$ is a deformation tensor. The affine transformation (44) can be used to map the abstract problem (41) from a reference domain $\widehat{\Omega}$ to a domain of interest $\Omega(\mu)$. In pratice, this is done by applying the chain rule and the change of variables theorem:

$$\frac{\partial}{\partial x_i} = \sum_j \frac{\partial \widehat{x}_j}{\partial x_i} \frac{\partial}{\partial \widehat{x}_j} = \sum_j G_{ji}(\mu) \frac{\partial}{\partial \widehat{x}_j} \qquad \mathrm{d}\boldsymbol{x} = |J^{\mathrm{aff}}(\mu)|^{-1} \, \mathrm{d}\widehat{\boldsymbol{x}} \tag{45}$$

for all $\widehat{\boldsymbol{x}} \in \widehat{\Omega}$ and $\boldsymbol{x} \in \Omega(\mu)$, where the Jacobian $J^{\mathrm{aff}}(\mu) = |\det(G(\mu))|$ has been introduced.

3.7. A comment on the condition number of the problem

As we have seen, a good ROM tries to operate at an optimal point between two contrasting phenomena: on one hand we seek high performances during the online phase by decreasing as much as possible the number of reduced bases, on the other hand we try to mantain a good accuracy on the results. We remark that for a large class of problems there is no point in trying to achieve a very high accuracy by adding too many bases during the online phase, since the addition of bases, if not carefully chosen, could severely increase the condition number of the linear system.

For instance, in many elliptic problems the modes obtained with POD or with a GS orthogonalization have an energy decreasing exponentially with the mode number. In this case, we can expect the condition number of the online problem to be no less that the ratio between the energy of the most and least energetic modes, that could easily be of the order of 10^{10}. Even if the modes are normalized after their orthogonalization, the low condition number of the online matrix will not be representative of a well conditioned problem, since cathastrophic cancellation errors during the normalization phase would artificially reduce the condition number without recovering the expected accuracy [35].

3.8. Detection of singular points

We now turn our attention to the approximation properties of the Reduced Basis spaces presented in the previous sections with respect to the detection of the classes of singular points discussed in section 2.3. In this way we aim at building a detecting tool for singular points and branches of non-unique solutions.

From the practical viewpoint, we want to study the eigenvalues of the differential $D_u F(\mu, u(\mu)) \in \mathscr{L}(X, Y')$ as μ is varied. To do so, we project $D_u F$ evaluated on a solution obtained with the Reduced Basis method as explained in section (3.5) on a basis for the spaces X^N and Y^N in order to obtain the differential operator's matrix $\mathrm{T}(\mu)$:

$$\mathrm{T}_{ij}(\mu) = \langle D_u F(\mu, u^N(\mu))[\zeta_i], \varsigma_j \rangle \tag{46}$$

where $\{\varsigma_i\}_{i=1}^N$ is a basis for Y^N and the dependence of $\mathrm{T}(\mu)$ on the parameter-solution couple $(\mu, u^N(\mu))$ is emphasized. Then we compute the eigenvalues of $\mathrm{T}(\mu)$ in order to detect one of the following cases:

- fold point: if μ is a fold point, we expect that in a neighbourhood of μ the spectrum of $\mathrm{T}(\mu)$ will show an eigenvalue approach to and then depart from the imaginary axis;
- bifurcation point: in this case, in a neighbourhood of μ we can find an eigenvalue of $\mathrm{T}(\mu)$ changing sign;
- Hopf bifurcation point: in this case, in a neighbourhood of μ there is a couple of complex conjugate eigenvectors of $\mathrm{T}(\mu)$ crossing the imaginary axis.

The situation described above should however be taken as indicative when the approximation spaces are built with a satisfactory accuracy. Otherwise, the discretization step will introduce perturbations of significative magnitude, that could not only affect the position of the detected singular points, but change the topology of the solutions set, as described in [17]. We refer to section 3.9 for a comment on error certification.

We remark that since $\mathrm{T}(\mu)$ has a small dimension, all the eigenvalues can be computed at a reasonable expense, for instance using QR iterations [19].

Since the stability of both steady state and time-dependent problems is related to the spectral properties of the operator's differential, we might expect that good results of the ROM in the detection of bifurcation points are related to the effictiveness in the approximation of the spectral values.

In this context, we might regard the Reduced Basis method as a Krylov method for approximating the eigenvalues of a parametrized operator, with the peculiarity that the Lanczos or Arnoldi methods for generating a basis set for the Krylov space are replaced by a "smarter", ad-hoc set of basis vectors (we refer to [36] for Krylov methods, and to [37] for Arnoldi methods). This is of particular importance when the dimension of the parameter space is relatively high, since it is the case where the efficiency of the RBM can provide the most striking savings in computational time with respect to standard methods.

Recently, a ROM-based eigenvalue approach has been introduced for the detection of mechanical vibration in the automotive industry [38], showing reliable results.

3.9. Certification

We conclude our brief focus on Reduced Order Methods ideas with a note on error estimation. Basically we are interested in computing error bounds for the approximated solutions both in the steady state and in the time-dependent case, and also error bounds for the location of singular points.

Regarding the error bounds for the Reduced Basis solutions, the framework of BRR theory can by all means be specialized to return error bounds in the form:

$$\Delta(\mu) = \left\| u^{\mathcal{N}}(\mu) - u^N(\mu) \right\|_X . \tag{47}$$

the error bound (47) can be expressed in terms of quantities related to F and its differential (for details see [39, 40], and [10] for a review). For this reason we recall the definition of some constants:

- the *continuity constant* γ of a is, if it exists:

$$\gamma = \sup_{w \in X} \sup_{v \in Y} \frac{\langle D_u F(\mu, u)[w], v \rangle}{\|w\|_X \|v\|_Y} = \|D_u F(\mu, u)\|_{\mathscr{L}(X, Y')} < +\infty; \tag{48}$$

- the *coercivity constant* α of a is, if it exists:

$$\alpha = \inf_{u \in X} \inf_{v \in Y} \frac{\langle F(\mu, u), v \rangle}{\|u\|_X \|v\|_Y} > 0; \tag{49}$$

- the *inf-sup constant* β of a is, if it exists:

$$\beta = \inf_{w \in X} \sup_{v \in Y} \frac{\langle D_u F(\mu, u(\mu))[w], v \rangle}{\|w\|_X \|v\|_Y} > 0; \tag{50}$$

- the *residual* r is obtained computing the operator F on the approximate solution:

$$r^N(\mu) = F(\mu, u^N(\mu)), \tag{51}$$

  and its dual norm is given by:

$$\|r^N(\mu)\|_{Y'} = \sup_{v \in Y} \frac{\langle F(\mu, u^N(\mu)), v \rangle}{\|v\|_Y} \tag{52}$$

We remark that the parameter-solution correspondence implied by the preceding definitions is one-to-one only on each solution branch, and sufficiently far away from bifurcation points. Anyway, in general there is no need of computing explicitly the constants (48)-(50), but only a reasonably sharp upper or lower bound that can be estimated for example via a Successive Constraint Method (SCM, see e.g. [41]).

For regular solution branches, BRR theory in [17] returns error estimates supposing that F_h is a consistent approximation of F and that its differential is Lipshitz continuous:

$$\|D_u F(\mu, u) - D_u F(\mu, v)\|_{\mathscr{L}(X,Y')} \leq L\|u - v\|_X \qquad \forall u, v \in S \tag{53}$$

for some subset S of X, and has a bounded inverse:

$$\left\|D_u F(\mu, u)^{-1}\right\|_{\mathscr{L}(Y',X)} < +\infty. \tag{54}$$

Then, the following bound holds:

$$\begin{aligned} \left\|u(\mu) - u^N(\mu)\right\|_X &\leq 2\left\|D_u F(\mu, u^N(\mu))^{-1}\right\|_{\mathscr{L}(Y',X)} \|F(\mu, u^N(\mu))\|_{Y'} \\ &\leq \frac{C\gamma}{\beta} f(\left\|r^N(\mu)\right\|_{Y'}) \end{aligned} \tag{55}$$

for an appropriate function f, at least when X^N is sufficiently large.

For the approximation of simple limit points, BRR theory considers the case in which the decomposition in linear and nonlinear parts, L and N respectively, of F is such that both L and $D_u N(\mu, u)$ are compact operators from X to Y'. In this case the Implicit Function Theorem can be applied to a regularized version of F as exposed in section 2 to obtain estimates in the form:

$$|\mu(s) - \mu^N(s)| + \|u(s) - u^N(s)\|_X \leq C \inf_{v^N \in X^N} \|u(s) - v\|_X, \tag{56}$$

where $s \in [-\varepsilon, \varepsilon]$ is a real number used to parametrize the solution set in a neighbourhood of the fold point (μ^*, u^*). Similar results are available [42] also for the case of simple quadratic fold points. We remark that in equation (56) the right hand side expresses the approximation properties of the space X^N, and could be specialized for instance applying the theoretical bounds available for POD or Greedy sampling.

Lastly, we comment on the error estimation for time-dependent nonlinear problems in the form:

$$\frac{\partial u}{\partial t} = F(\mu, u). \tag{57}$$

If F is coercive, or if it is possible to find an energy-like bound on the solution, usually there are estimates based respectively on Gronwall lemma or energy conservation that have the form:

$$\|u^{\mathcal{N}}(\mu) - u^N(\mu)\|_X^2 \leq C \exp(-\alpha t)\|r^N(\mu)\|_{Y'}. \tag{58}$$

In general this bound is effective only if some smallness criterion is satisfied, e.g. for sufficiently small times or for small data.

Another possibility is found by noting that in principle the estimates provided by BRR theory can be adapted to the case where X and Y' are Banach spaces defined on a space-time domain, as shown in [43, 9] for the Navier–Stokes equations[4].

In this case the norms of X and Y' should be set up appropriately depending on the problem structure. For instance, for a linear heat equation the classical choice [44] would be $X = L^2([0,T]; H^1(\Omega))$, with the norm:

$$\|u(x,t)\|_X^2 = \int_0^t u(x,s)^2 \,\mathrm{d}s + \int_\Omega |\nabla u(y,t)|^2 \,\mathrm{d}y. \tag{59}$$

Then, the bound (55) holds in the space-time norm of X, that although is still non-decreasing in time, it could be much better than the exponential bound (58).

One disadvantage of this approach is that the algebraic linear system to be solved online, and the computation of an estimate for the residual's dual norm are more demanding. Furthermore, even in this approach some form of control on the operator is necessary, specifically by requiring the diffusion constant to be sufficiently large.

[4] In truth the cited papers develop the theory even further, deriving also bounds for output functionals of the solution and stability criteria for flows under a large class of perturbations.

4. Application to Incompressible Fluid Dynamics

In this section the results of the previous sections are specialized for the case of Navier-Stokes equations. For the sake of generality, we focus in particular on the Rayleigh-Bénard equations for buoyancy driven flows.

4.1. Mathematical model

The strong formulation of the Rayleigh-Bénard cavity problem, as stated in [12], is given in adimensional variables as follows

$$\begin{cases} \frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \nabla \boldsymbol{u} - \Delta \boldsymbol{u} + \nabla p = \mathrm{Gr}\, \vartheta \boldsymbol{\jmath} & \text{on } \Omega \\ \operatorname{div} \boldsymbol{u} = 0 & \text{on } \Omega \\ \boldsymbol{u} = 0 & \text{on } \Gamma_D \\ -\Delta \vartheta = 0 & \text{on } \Omega \\ \vartheta = x & \text{on } \Gamma_D. \end{cases} \tag{60}$$

Where $\boldsymbol{\jmath}$ is the vertical versor, x the horizontal coordinate, Gr the *Grashof number* wich expresses roughly the ratio of buoyancy to viscous forces, and is a parameter. The domain Ω considered in the benchmark is a rectangular bidimensional cavity with unit height and width A. We consider the problem with fully Dirichlet boundary conditions, in symbols we write $\Gamma_D \equiv \partial\Omega$. Solving equations (60) allows to obtain the triple $(\boldsymbol{u}, p, \vartheta)$, representing adimensional fluid velocity, pressure and temperature, respectively.

System (60) consists of three equations, from the top to the bottom expressing momentum, mass and energy balance for an arbitrarily small control volume, treating the fluid as a continuum (see for instance [45] for a detailed exposition of the physical theory). In particular, energy equation as stated on system (60) represents the limit of the more general energy balance equation as the Prandtl number tends to zero. This approximation, along with the simple boundary conditions allows to solve analytically the energy equation, leading to the linear solution in temperature $\vartheta = x$. Navier-Stokes equations and energy equation are in this case uncoupled.

It is convenient to write the variational form of system (60). For this purpose, we introduce the velocity and pressure spaces $\boldsymbol{V} \equiv [H^1_0(\Omega)]^d$ and $Q \equiv L^2_0(\Omega)$, respectively[5], where d is the spatial dimension. Then, multiplying the momentum and mass balance equations in (60) respectively by the test functions $\boldsymbol{v} \in \boldsymbol{V}$ and $q \in Q$, and integrating formally by parts, we get the variational formulation that reads as follows: find $(\boldsymbol{u}, p) \in \boldsymbol{V} \times Q$ such that

$$\begin{cases} m(\boldsymbol{u}, \boldsymbol{v}) + c(\boldsymbol{u}, \boldsymbol{u}, \boldsymbol{v}) + a(\boldsymbol{u}, \boldsymbol{v}) + b(\boldsymbol{v}, p) = f(\boldsymbol{v}) & \forall \boldsymbol{v} \in \boldsymbol{V} \\ b(\boldsymbol{u}, q) = 0 & \forall q \in Q. \end{cases} \tag{61}$$

Where the following bilinear forms have been introduced

$$\begin{aligned} a(\boldsymbol{u}, \boldsymbol{v}) &= \int_\Omega \nabla \boldsymbol{v} : \nabla \boldsymbol{u} \, \mathrm{d}\boldsymbol{x} \\ b(\boldsymbol{v}, q) &= \int_\Omega q \operatorname{div} \boldsymbol{v} \, \mathrm{d}\boldsymbol{x} \\ m(\boldsymbol{u}, \boldsymbol{v}) &= \int_\Omega \boldsymbol{v} \cdot \frac{\partial \boldsymbol{u}}{\partial t} \, \mathrm{d}\boldsymbol{x}, \end{aligned} \tag{62}$$

along with the variational form

$$c(\boldsymbol{u}, \boldsymbol{w}, \boldsymbol{v}) = \int_\Omega \boldsymbol{v} \cdot (\boldsymbol{u} \cdot \nabla \boldsymbol{w}) \, \mathrm{d}\boldsymbol{x} \tag{63}$$

and the linear form

$$f(\boldsymbol{v}) = \int_\Omega \mathrm{Gr}\, \vartheta \boldsymbol{\jmath} \cdot \boldsymbol{v} \, \mathrm{d}\boldsymbol{x}. \tag{64}$$

[5]The notation chosen for the function spaces may need clarification. We refer with H^1_0 to the Sobolev space with zero trace at the boundary (velocity), and with L^2_0 to the Lebesgue L^2 functions with zero average (pressure).

4.2. Parametrized formulation

We are interested in approximating the solutions set of equations (61) for a wide range of aspect ratios for the rectangular geometry, and for an interval of the Grashof number over which the bifurcations take place. Specifically, the independent parameters are the cavity length $\mu \equiv A$ and the Grashof number Gr. To simplify the notation, let us introduce the parameters vector $\boldsymbol{\mu} = (\mu, \mathrm{Gr})$.

As discussed in section 3.6, we introduce a domain parametrization. The main idea behind this procedure is to refer each of the infinitely many geometric configurations to a common "original" domain. This way, there is no need of constructing a new mesh for each computation, and all the problems can be cast on the original domain only. The solution obtained on the original domain can then be mapped to the "true" domain.

In the following, the dependence of the variational forms on the parameters vector $\boldsymbol{\mu}$ will be denoted explicitly, and problem (61) will be cast as: find $(\boldsymbol{u}(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \in \boldsymbol{V} \times Q$ such that

$$\begin{cases} m(\boldsymbol{\mu}; \boldsymbol{u}(\boldsymbol{\mu}), \boldsymbol{v}) + c(\boldsymbol{\mu}; \boldsymbol{u}(\boldsymbol{\mu}), \boldsymbol{u}(\boldsymbol{\mu}), \boldsymbol{v}) + a(\boldsymbol{\mu}; \boldsymbol{u}(\boldsymbol{\mu}), \boldsymbol{v}) + b(\boldsymbol{\mu}; \boldsymbol{v}, p(\boldsymbol{\mu})) = f(\boldsymbol{\mu}; \boldsymbol{v}) & \forall \boldsymbol{v} \in \boldsymbol{V} \\ b(\boldsymbol{\mu}; \boldsymbol{u}(\boldsymbol{\mu}), q) = 0 & \forall q \in Q. \end{cases} \tag{65}$$

The dependence on the parameters $\boldsymbol{\mu}$ of the solution has been explicitly pointed out. We remark that each of the operators can be affinely split into a part which depends on $\boldsymbol{\mu}$ and a part depending on $\boldsymbol{u}$ and $\boldsymbol{x}$. This allows an offline-online decomposition in the computational steps, a feature of great importance for the efficiency of the reduced order method, as discussed in section 3.5.

4.3. Spectral element approximation

Having set the variational form of the Parametrized Partial Differential Equation (P^2DE or μPDE), we proceed with the numerical approximation for problem (65) by means of the Galerkin projection method. The Galerkin method consists in the projection of the continuous problem into a finite dimensional subspace $(\boldsymbol{V}^{\mathcal{N}}, Q^{\mathcal{N}})$ such that $\boldsymbol{V}^{\mathcal{N}} \subset \boldsymbol{V}$ and $Q^{\mathcal{N}} \subset Q$. The finite-dimensional approximation of problem (65) is: find $(\boldsymbol{u}^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}(\boldsymbol{\mu})) \in \boldsymbol{V}^{\mathcal{N}} \times Q^{\mathcal{N}}$ such that

$$\begin{cases} m_h(\boldsymbol{\mu}; \boldsymbol{u}^{\mathcal{N}}(\boldsymbol{\mu}), \boldsymbol{v}^{\mathcal{N}}) + c_h(\boldsymbol{\mu}; \boldsymbol{u}^{\mathcal{N}}(\boldsymbol{\mu}), \boldsymbol{u}^{\mathcal{N}}(\boldsymbol{\mu}), \boldsymbol{v}^{\mathcal{N}}) + a_h(\boldsymbol{\mu}; \boldsymbol{u}^{\mathcal{N}}(\boldsymbol{\mu}), \boldsymbol{v}^{\mathcal{N}}) & \\ \qquad + b_h(\boldsymbol{\mu}; \boldsymbol{v}^{\mathcal{N}}, p^{\mathcal{N}}(\boldsymbol{\mu})) = f_h(\boldsymbol{\mu}; \boldsymbol{v}^{\mathcal{N}}) & \forall \boldsymbol{v}^{\mathcal{N}} \in \boldsymbol{V}^{\mathcal{N}} \\ b_h(\boldsymbol{\mu}; \boldsymbol{u}^{\mathcal{N}}(\boldsymbol{\mu}), q^{\mathcal{N}}) = 0 & \forall q^{\mathcal{N}} \in Q^{\mathcal{N}}. \end{cases} \tag{66}$$

In equation (66), the differential forms show the subscript $_h$ to remind the possible presence of "variational crimes" such as Gaussian quadrature.

Then, given a set of basis functions $\{\boldsymbol{\varphi}_i^{\mathcal{N}}\}_{i=1}^{\mathcal{N}_{\boldsymbol{u}}}$ and $\{\psi_k^{\mathcal{N}}\}_{k=1}^{\mathcal{N}_p}$ for $\boldsymbol{V}^{\mathcal{N}}$ and $Q^{\mathcal{N}}$ respectively, we can expand the approximate solutions as

$$\boldsymbol{u}^{\mathcal{N}}(\boldsymbol{\mu}) = \sum_{i=1}^{\mathcal{N}_{\boldsymbol{u}}} u_i^{\mathcal{N}}(\boldsymbol{\mu}) \boldsymbol{\varphi}_i^{\boldsymbol{N}} \qquad p^{\mathcal{N}}(\boldsymbol{\mu}) = \sum_{k=1}^{\mathcal{N}_p} p_k^{\mathcal{N}}(\boldsymbol{\mu}) \psi_k^{\mathcal{N}}. \tag{67}$$

An aspect of fundamental importance is the choice of the finite-dimensional spaces $\boldsymbol{V}^{\mathcal{N}}$ and $Q^{\mathcal{N}}$. Among the many possibilities developed over the years, we choose the Legendre Spectral Element Method (for details we refer for instance to [46] or [47, 48]) as implemented in the open source software `Nek5000` [49]. In this work, we used the $\mathbb{P}_N - \mathbb{P}_N$ couple for velocity and pressure with polynomials of order 20.

Choosing the basis functions as test functions, we obtain the algebraic form of problem (66). A delicate point when implementing the numerical solver is the choice of the linearization method for the nonlinear convective term. In this work, we adopted the operator splitting-explicit in time third order backward difference/extrapolation formulas. For details on the $\mathbb{P}_N - \mathbb{P}_N$ splitting methods, we refer to [50].

After the linearization, and expanding $\boldsymbol{u}^{\mathcal{N}}$ and $p^{\mathcal{N}}$ as a linear combination of the basis functions, a Linear Algebra problem has been obtained, in the form:

$$\begin{bmatrix} H(\boldsymbol{\mu}) & B^T(\boldsymbol{\mu}) \\ B(\boldsymbol{\mu}) & 0 \end{bmatrix} \begin{pmatrix} \boldsymbol{U}(\boldsymbol{\mu}) \\ \boldsymbol{P}(\boldsymbol{\mu}) \end{pmatrix} = \begin{pmatrix} \boldsymbol{F}(\boldsymbol{\mu}) \\ \boldsymbol{0} \end{pmatrix} \tag{68}$$

where $\boldsymbol{U}_i(\boldsymbol{\mu}) = u_i^{\mathcal{N}}(\boldsymbol{\mu})$ and $\boldsymbol{P}_k(\boldsymbol{\mu}) = p_k^{\mathcal{N}}(\boldsymbol{\mu})$ are the unknowns vectors, and $\boldsymbol{F}(\boldsymbol{\mu})$ is the vector arising from the discretization of the explicit terms. In system (68), $B^T(\boldsymbol{\mu})$ and $B(\boldsymbol{\mu})$ are the discrete gradient and divergence matrices respectively, and $H(\boldsymbol{\mu})$ is the discrete Helmholtz operator, obtained as linear combination of the velocity mass and stiffness matrices.

5. A reduced order modelling technique for Navier–Stokes bifurcation problems

In this section the results of section 2 are specialized to construct a possible framework for the fast solution of bifurcation problems based on the Proper Orthogonal Decomposition and Reduced Basis (POD-RB) methods.

We remark that as for the truth spaces, the reduced basis spaces should be chosen in order to fulfill three fundamental properties. First, it is important that good stability properties are verified, that is, the trial and test spaces must lead to a well-posed problem (*approximation stability*). Secondly, the reduced order spaces should guarantee good approximation properties for all the parameters on a given interval. Third, the reduced order spaces should have the lowest possible dimension while mantaining the required tolerance (*algebraic stability*, i.e. no ill-conditioning in the reduced order matrices). In the following we will discuss how these properties can be satisfied by a careful choice of the spaces $\boldsymbol{V}^N$ and Q^N.

5.1. Reduced order formulation of the Navier-Stokes equations

A basis for the spaces $\boldsymbol{V}^N$ and Q^N can be computed through the techniques discussed in section 3.4, but some additional care is required to ensure the approximation stability of the reduced basis spaces. In general, the basis obtained as described in section 3.4 will not fulfill the Brezzi inf-sup condition:

$$\beta^N \equiv \inf_{q \in Q^N} \sup_{\boldsymbol{v} \in \boldsymbol{V}^N} \frac{b(\boldsymbol{\mu}; q, \boldsymbol{w})}{\|q\|_{Q^N} \|\boldsymbol{v}\|_{\boldsymbol{V}^N}} > 0. \tag{69}$$

As a consequence, the standard stability estimate for mixed problems (see for instance [51]):

$$\inf_{\boldsymbol{w}^N \in \boldsymbol{V}^N_{\text{div}}} \|\boldsymbol{u} - \boldsymbol{w}^N\|_{\boldsymbol{V}} \leq \left(1 + \frac{\delta}{\beta^N}\right) \inf_{\boldsymbol{v}^N \in \boldsymbol{V}^N} \|\boldsymbol{u} - \boldsymbol{v}^N\|_{\boldsymbol{V}} \tag{70}$$

does not guarantee the desired approximation properties of $\boldsymbol{V}^N$ and Q^N as $\beta^N \to 0$. Some possibilities to recover the inf-sup control are the following.

Supremizer enrichment this method consists in the enrichment of the velocity space with an ad-hoc set of basis functions, called *supremizers* and denoted with the symbol $T^{\boldsymbol{\mu}}$. The name supremizers is justified by the way they are computed: each supremizer is obtained imposing that:

$$T^{\boldsymbol{\mu}}\sigma = \arg \sup_{\boldsymbol{w} \in \boldsymbol{v}^{\mathcal{N}}} \frac{b(\boldsymbol{\mu}; \sigma, \boldsymbol{w})}{\|\boldsymbol{w}\|_{\boldsymbol{V}^{\mathcal{N}}}} \tag{71}$$

where σ is a basis for the pressure space Q^N. The supremizers ensure by definition that the inf-sup constant is positive, but require additional offline computations for their construction and produce a velocity space of larger dimension for the online phase with respect to the other two methods discussed here. The final enriched space may depend also on the online parameter value and can be assembled online.

We refer to [35] and [52] for an analysis of the supremizer stabilization, to [53] for a proposed surrogate option, and to [54] for an application to nonlinear parametrized problems, solved along with some rigorous and heuristic insights on the stabilized spaces. In particular, according to [54], usually it is not required to compute a supremizer basis for each element of the pressure basis, but only for some selected pressure basis functions.

Petrov-Galerkin stabilization a less popular option in the ROM community, studied by Rovas [55] for general nonsymmetric problems, Amsallem and Farhat [56] for supersonic fluid-structure interaction problems, and recently by Dahmen [57] and Abdulle and Budác [58] for Stokes problems. This method consists in building different reduced basis spaces for trial and test functions, in order to generate an oblique (i.e. non-orthogonal) online projection that can have better stability properties for non-coercive and non-symmetric problems. A disadvantage of this method is a more complex offline phase, since two different sampling procedures should be set up for the trial and test reduced basis spaces.

Piola transformation introduced in the ROM community by [59], consists in an online preprocessing of the velocity basis set $\{\boldsymbol{\zeta}_i\}$ that allows to obtain a set of (weakly[6]) divergence-free basis functions $\{\boldsymbol{\zeta}_i^{\text{div}}\}$ for each value of the geometric parameter μ. Having divergence-free basis allows to cancel out the pressure term from momentum equation, thus removing any stability issue for mixed problems. Pressure can then be recovered by means of a post-processing step, using the velocity coefficients to obtain the pressure field:

$$p^N(\boldsymbol{\mu}^i) = \sum_{k=1}^{N} u_k^N \sigma_k, \tag{72}$$

or alternatively a Poisson problem for the pressure can be solved online:

$$\Delta p^N(\boldsymbol{\mu}^i) = -\operatorname{div}\left(\boldsymbol{u}^N(\boldsymbol{\mu}^i)\cdot\nabla \boldsymbol{u}^N(\boldsymbol{\mu}^i)\right). \tag{73}$$

We refer for example to [60] for an analysis of velocity-pressure reduced order models.

Each snapshot $\boldsymbol{u}^{\mathcal{N}}(\boldsymbol{\mu}^i)$ is weakly divergence-free in the original domain $\Omega(\mu^i)$, but this is not true when the snapshot is pulled back to the reference domain $\widehat{\Omega}$. In fact, the parametrized formulation imposes:

$$\int_{\Omega(\mu^i)} q \operatorname{div} \boldsymbol{u}(\boldsymbol{\mu}^i)\,\mathrm{d}\boldsymbol{x} = 0 \qquad \forall q \in Q^N, \tag{74}$$

that in coordinates reads:

$$\int_{\Omega(\mu^i)} q\Big(\sum_{j=1}^{d} \frac{\partial \boldsymbol{u}_j(\boldsymbol{\mu}^i)}{\partial x_j}\Big)\,\mathrm{d}\boldsymbol{x} = \int_{\widehat{\Omega}} q\Big(\sum_{j=1}^{d}\sum_{k=1}^{d} G_{jk}(\mu^i)\frac{\partial \boldsymbol{u}_j(\boldsymbol{\mu}^i)}{\partial \widehat{x}_k}\Big) J^{\text{aff}}(\mu^i)\,\mathrm{d}\widehat{\boldsymbol{x}}. \tag{75}$$

It can be concluded that the snapshots do not cancel out the standard divergence on the reference domain:

$$\int_{\widehat{\Omega}} q \sum_{j=1}^{d} \frac{\partial \boldsymbol{u}_j(\boldsymbol{\mu}^i)}{\partial \widehat{x}_j}\,\mathrm{d}\widehat{\boldsymbol{x}} \qquad \forall q \in Q^N \tag{76}$$

but instead the pushed forward, or "stretched" divergence (75).

An advantage of this method is that the online system is smaller because all the pressure-related computations can be avoided, and there is no need of additional offline pre-processing as required by the two previous techniques. A disadvantage is that during the online phase an additional preprocessing step is required to "map" the matrices on the parametrized divergence-free space.

In the following section, we explore the possibilities offered by the Piola transformation to build a divergence-free basis on the reference domain.

The form of the Piola transformation can be inferred by comparing equations (75) and (76):

$$\begin{cases} \boldsymbol{u}_1^{\mathcal{N},\text{div}} = \mathcal{P}_1 \boldsymbol{u}^{\mathcal{N}} = G_{11}(\mu)\boldsymbol{u}_1^{\mathcal{N}} + G_{12}(\mu)\boldsymbol{u}_2^{\mathcal{N}} \\ \boldsymbol{u}_2^{\mathcal{N},\text{div}} = \mathcal{P}_2 \boldsymbol{u}^{\mathcal{N}} = G_{21}(\mu)\boldsymbol{u}_1^{\mathcal{N}} + G_{22}(\mu)\boldsymbol{u}_2^{\mathcal{N}} \end{cases} \tag{77}$$

[6] Due to the variational formulation, the discrete divergence vanishes but in general this is not true for the pointwise divergence.

for the two-dimensional case. In equation (77), $\boldsymbol{u}^{\mathcal{N},\mathrm{div}}$ is the vector $\boldsymbol{u}^{\mathcal{N}}$ after the Piola transformation has been applied with the (vectorial) map $\mathcal{P}$. The superscript $^{\mathrm{div}}$ is used to stress the fact that after being Piola-transformed the vectors are divergence free.

After the reference basis $\{\widehat{\boldsymbol{\zeta}}^{\mathrm{div}}\}$ is built by means of the Gram–Schmidt procedure (39), the mass, stiffness and convection matrices are assembled on the reference domain:

$$\begin{aligned}\widehat{M}_{ij} &= \int_{\widehat{\Omega}} \widehat{\boldsymbol{\zeta}}_i^{\mathrm{div}} \cdot \widehat{\boldsymbol{\zeta}}_j^{\mathrm{div}} \,\mathrm{d}\widehat{\boldsymbol{x}} \\ \widehat{K}_{ij} &= \int_{\widehat{\Omega}} \widehat{\nabla}\widehat{\boldsymbol{\zeta}}_i^{\mathrm{div}} \cdot \widehat{\nabla}\widehat{\boldsymbol{\zeta}}_j^{\mathrm{div}} \,\mathrm{d}\widehat{\boldsymbol{x}} \\ \widehat{C}_{ijk} &= \int_{\widehat{\Omega}} (\widehat{\boldsymbol{\zeta}}_j^{\mathrm{div}} \cdot \widehat{\nabla}\widehat{\boldsymbol{\zeta}}_k^{\mathrm{div}}) \cdot \widehat{\boldsymbol{\zeta}}_i^{\mathrm{div}} \,\mathrm{d}\widehat{\boldsymbol{x}}.\end{aligned} \tag{78}$$

Note that we do not have to assemble any pressure-related matrix.

During the online phase, the divergence-free basis $\{\widehat{\boldsymbol{\zeta}}_i^{\mathrm{div}}\}$ on the reference domain will not satisfy the divergence-free constraint on the original domain. Consequently, the differential forms (62), (63), (64) should not be evaluated on the basis $\{\widehat{\boldsymbol{\zeta}}_i^{\mathrm{div}}\}$ on the reference domain, but instead on the parametrized divergence-free basis $\{\boldsymbol{\zeta}_i^{\mathrm{div}}\}$. Such basis are obtained through a parametrized Piola transformation of the original basis:

$$\begin{cases}\boldsymbol{\zeta}_{i,1}^{\mathrm{div}} = \mathcal{P}_1(\mu)\widehat{\boldsymbol{\zeta}}^{\mathrm{div}} = G_{11}(\mu)\widehat{\boldsymbol{\zeta}}_{i,1}^{\mathrm{div}} + G_{12}(\mu)\widehat{\boldsymbol{\zeta}}_{i,2}^{\mathrm{div}} \\ \boldsymbol{\zeta}_{i,2}^{\mathrm{div}} = \mathcal{P}_2(\mu)\widehat{\boldsymbol{\zeta}}^{\mathrm{div}} = G_{21}(\mu)\widehat{\boldsymbol{\zeta}}_{i,1}^{\mathrm{div}} + G_{22}(\mu)\widehat{\boldsymbol{\zeta}}_{i,2}^{\mathrm{div}}.\end{cases} \tag{79}$$

During the parametrized simulation, it is not necessary to build explicitly the new basis $\{\boldsymbol{\zeta}_i^{\mathrm{div}}\}$, but only to evaluate its effect on the matrices (78) (i. e. on the coeffcients).

Finally, we obtain the reduced order formulation of problem (61) by applying first the Galerkin projection on $\boldsymbol{V}^N \times Q^N$ followed by the Piola transformation. The reduced order problem reads: find $(\boldsymbol{u}^N(\boldsymbol{\mu}), p^N(\boldsymbol{\mu})) \in \boldsymbol{V}^N \times Q^N$ such that

$$m_h(\boldsymbol{\mu}; \boldsymbol{u}^N(\boldsymbol{\mu}), \boldsymbol{v}^N) + c_h(\boldsymbol{\mu}; \boldsymbol{u}^N(\boldsymbol{\mu}), \boldsymbol{u}^N(\boldsymbol{\mu}), \boldsymbol{v}^N) + a_h(\boldsymbol{\mu}; \boldsymbol{u}^N(\boldsymbol{\mu}), \boldsymbol{v}^N) = f_h(\boldsymbol{\mu}; \boldsymbol{v}^N) \tag{80}$$

for all $\boldsymbol{v}^N \in \boldsymbol{V}^N$. The algebraic problem becomes:

$$H_N(\boldsymbol{\mu})\boldsymbol{U}_N(\boldsymbol{\mu}) = \boldsymbol{F}_N(\boldsymbol{\mu}), \tag{81}$$

where H_N is a linear combination of the mass, stiffness and convection matrices. The performance improvements obtained through the ROM are clear, since the (sparse) algebraic system (68) of order $\mathcal{N}_{\boldsymbol{u}} + \mathcal{N}_p$ has been reduced to the (dense) system (81) of order $N_{\boldsymbol{u}}$, with $N_{\boldsymbol{u}} \ll \mathcal{N}_{\boldsymbol{u}} < \mathcal{N}_{\boldsymbol{u}} + \mathcal{N}_p$.

5.2. Branching detection and tracing

5.2.1. Steady bifurcation

To identify a steady state bifurcation point, we follow a technique similiar to Lemma 4 in [61] that we briefly discuss. First, we rewrite the variational problem (61) in the abstract form of section 2:

$$F(\boldsymbol{\mu}; (\boldsymbol{u}(\boldsymbol{\mu}), p(\boldsymbol{\mu}))) = 0, \tag{82}$$

then we introduce the linearized advection operator $\boldsymbol{\mathcal{T}}(\boldsymbol{u}^*) : \boldsymbol{V} \mapsto \boldsymbol{V}$, obtained by taking the Fréchet derivative of the convection term $\boldsymbol{u} \cdot \nabla \boldsymbol{u}$ about a base solution $\boldsymbol{u}^*$:

$$\boldsymbol{\mathcal{T}}(\boldsymbol{u}^*)[\boldsymbol{v}] \equiv \boldsymbol{u}^* \cdot \nabla \boldsymbol{v} + \boldsymbol{v} \cdot \nabla \boldsymbol{u}^*. \tag{83}$$

According to [61], if $(\mu^*, \boldsymbol{u}^*)$ is a bifurcation point, the equation

$$\mu^* \boldsymbol{v} + \boldsymbol{\mathcal{T}}(\boldsymbol{v}^*(\mu^*))[\boldsymbol{v}] = 0 \tag{84}$$

has at least one nonzero solution $\boldsymbol{v}$.

From a reduced order modelling perspective, following [62] we search for a change of sign of the eigenvalues of the matrix $T(\boldsymbol{u}^*)$, defined as

$$\begin{aligned} T_{ij}(\boldsymbol{v}^*) &= \boldsymbol{\mathcal{T}}((\boldsymbol{u}^*, \boldsymbol{\zeta}_j^{\text{div}})_0 \boldsymbol{\zeta}_j^{\text{div}})[\boldsymbol{\zeta}_i^{\text{div}}] \\ &= \sum_{k=1}^{N_{\boldsymbol{u}}} (\boldsymbol{\zeta}_i^{\text{div}}, \boldsymbol{\zeta}_k^{\text{div}} \cdot \nabla \boldsymbol{\zeta}_j^{\text{div}})_0 \boldsymbol{U}_N^k + \sum_{k=1}^{N_{\boldsymbol{u}}} (\boldsymbol{\zeta}_i^{\text{div}}, \boldsymbol{\zeta}_j^{\text{div}} \cdot \nabla \boldsymbol{\zeta}_k^{\text{div}})_0 \boldsymbol{U}_N^k \end{aligned} \tag{85}$$

We portray in figure 3 the eigenvalues of the operator $\boldsymbol{\mathcal{T}}(\boldsymbol{u}^*)$ evaluated in a neighbourhood of a bifurcation point for a reduced order simulation of the cavity toy model problem. In particular, we fix an interval for the Grashof number where there exist both a steady state and a Hopf bifurcation point for the single roll flow, and we evaluate the spectrum of the tangent operator on a basis set obtained from flows with one, two and three rolls, and we compare it with the spectrum computed on a basis set coming from snapshots with only two and three rolls. In the first case we can see both a steady state bifurcation (the eigenvalue crossing the zero) and, for higher values of the Grashof number, a Hopf bifurcation (the pair of complex conjugate eigenvalues crossing the imaginary axis). In the second case no eigenvalue passes through zero or crosses the imaginary axis, hence the second reduced basis set is not able to detect the two bifurcations.

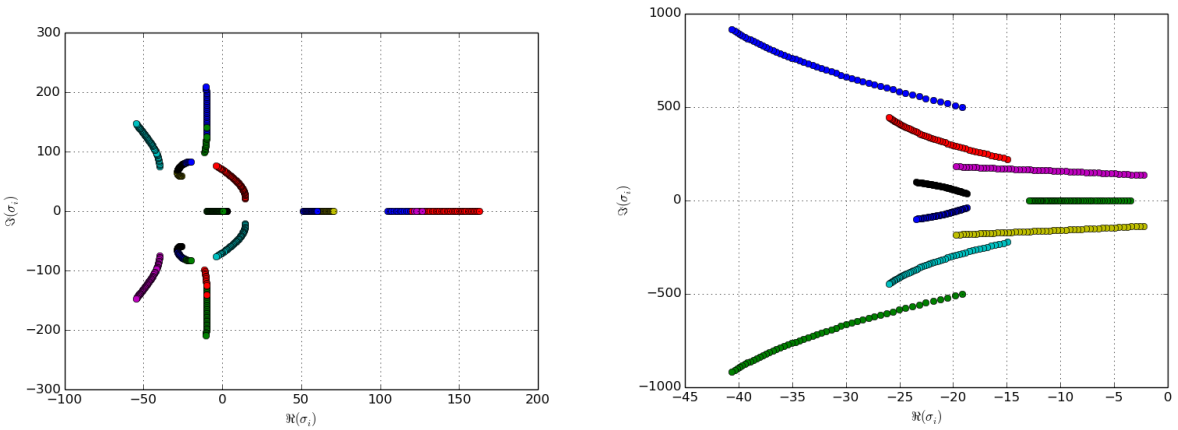

Figure 3: Reduced Model Eigenvalues of the tangent advection operator $\boldsymbol{\mathcal{T}}$ at the bifurcation point for $A = 4$, for $\text{Gr} \in [80 \cdot 10^3, 110 \cdot 10^3]$. On the left the tangent operator is evaluated on a basis set with one, two and three rolls, on the right a set with only two and three roll flows.

To give an intuitive understanding of this behaviour, we draw in figure 4 the projection of a bifurcating line (in black) on two planes. On the blue plane the bifurcation is correctly visualized, but in the green plane the projected curve (in red) no longer bifurcates. This analogy suggests that a similar projection behaviour may affect the reconstruction of bifurcating solutions sets if the reduced basis sets are not chosen properly.

As a further example, we report in figure 5 the path of the eigenvalues in a neighborhood of a steady bifurcation point. Highlighted in red is the path of the eigenvalue changing its sign as the Grashof number is increased. For clarity, we draw in figure 5 the evolution of the minimum eigenvalue as a function of the Grashof number for the same test case of figure 3. It can be seen that the zero is reached approximately at $\text{Gr} = 30.9 \cdot 10^3$, very close to the reference result of [63]. Furthermore, the critical eigenvalue crosses the origin with nonzero velocity, making eigenvalue analysis adequate for bifurcation detection purposes.

5.2.2. Hopf bifurcation

The Hopf bifurcation differs from the steady state bifurcations because it does not regard the branching of steady state solutions. Instead, Hopf bifurcation identifies the point at which time-dependent solutions

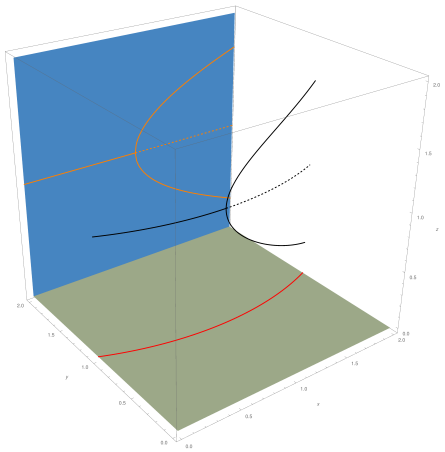

Figure 4: Intuitive visualization of the interpolation behaviour shown in figure 3: the original bifurcating curve (in black) does still bifurcate if projected in the blue plane, that we imagine spanned by the correct basis set, but it does not if projected in the lower plane, spanned by an incomplete basis set.

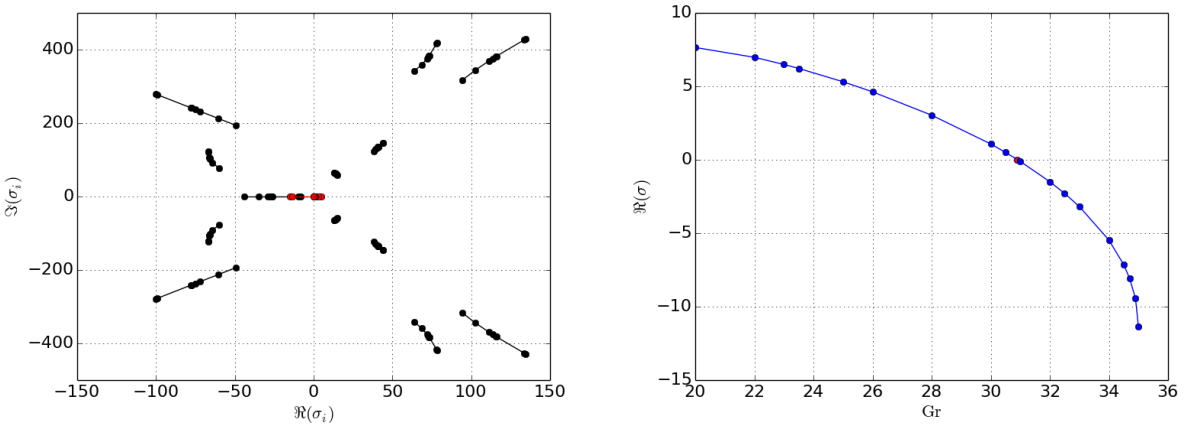

Figure 5: Left: evolution in the complex plane of the Reduced Model eigenvalues of $\mathcal{T}$ with the Grashof number in a neighborhood of a steady bifurcation point. Highlighted in red is the path of the critical eigenvalue, responsible for the bifurcation. Right: Reduced Model minimum eigenvalue of $\mathcal{T}$ as a function of the Grashof number. For clarity, the Grashof number is divided by 10^3.

become stable. This implies that a random perturbation from a steady state does not damp off in time, but after a transient it leads to a new, time-dependent solution.

For the detection of Hopf bifurcation points, the results of section 2.3 apply directly. We write the time-dependent Navier-Stokes equations seeking for a solution in the form of a superposition of a steady state solution $\boldsymbol{u}^*(\boldsymbol{x})$ and a small perturbation $\boldsymbol{u'}(\boldsymbol{x})\mathrm{e}^{\sigma t}$. Neglecting the second order terms, we obtain a linearized equation that rules the time-evolution of the perturbations:

$$\boldsymbol{u}^* \cdot \nabla \boldsymbol{u'} + \boldsymbol{u'} \cdot \nabla \boldsymbol{u}^* - \Delta \boldsymbol{u'} = -\sigma \boldsymbol{u'} \tag{86}$$

that can be seen as an eigenvalue problem for the linearized Navier-Stokes operator $\boldsymbol{\mathcal{L}} : \boldsymbol{V} \mapsto \boldsymbol{V}$:

$$\boldsymbol{\mathcal{L}}(\boldsymbol{u}^*)[\boldsymbol{u'}] = -\sigma \boldsymbol{u'}. \tag{87}$$

If equation (87) admits an eigenvalue σ^* such that $\Re\sigma^* > 0$, the corresponding perturbation will grow in time, and if $\Im\sigma^* > 0$, an oscillatory solution has to be expected, at least until the nonlinear term $\boldsymbol{u'} \cdot \nabla \boldsymbol{u'}$ remains sufficiently small.

Applying this technique to the reduced order model, the Hopf bifurcation is detected when the matrix L associated to the linearized Navier-Stokes operator $\boldsymbol{\mathcal{L}}(\boldsymbol{u}^*)$ admits an eigenvalue satisfying the conditions above. Explicitly, the matrix L has the form

$$L_{ij} = \sum_{k=1}^{N_{\boldsymbol{u}}} (\boldsymbol{\zeta}_i^{\mathrm{div}}, \boldsymbol{\zeta}_k^{\mathrm{div}} \cdot \nabla \boldsymbol{\zeta}_j^{\mathrm{div}})_0 \boldsymbol{U}_N^k + \sum_{k=1}^{N_{\boldsymbol{u}}} (\boldsymbol{\zeta}_i^{\mathrm{div}}, \boldsymbol{\zeta}_j^{\mathrm{div}} \cdot \nabla \boldsymbol{\zeta}_k^{\mathrm{div}})_0 \boldsymbol{U}_N^k + (\nabla \boldsymbol{\zeta}_i^{\mathrm{div}}, \nabla \boldsymbol{\zeta}_j^{\mathrm{div}})_0. \tag{88}$$

Approximating numerically the Reduced Order Model eigenvalues of matrices (85) or (88) is not particularly difficult due to their low dimension, hence direct eigenvalue algorithms can be used, such as the QR algorithm. Usually performing eigenvalue approximations of large sparse matrices arising from the stability analysis of Navier-Stokes equations is a delicate task, as Arnoldi methods can compute only a few eigenvalues. We refer to [64] for a discussion on the approximation of any number of eigenvalues for large systems.

6. Numerical results

We now test the efficiency of the reduced basis method for flow bifurcation problems on a well-known test case, presented at the GAMM benchmark on oscillatory convection of low Prandtl number flows, [12]. Many experimental and numerical references exist on this benchmark, among which we will refer to the results in [63].

In this benchmark the parameters are the domain length $A \in [2, 10]$ and the Grashof number $\mathrm{Gr} \in [50 \cdot 10^3, 1 \cdot 10^6]$, defined in [63] as the conventional Grashof number multiplied by the domain length A. In the considered parameter domain there exist regions admitting unique steady solutions, multiple steady solutions, and unsteady solutions, consequently this represents a good problem to test the proposed numerical method.

Some representative solutions for this benchmark are shown in figure 6. As it can be seen in the visualizations, the solutions are very heterogeneous as the Grashof number and Aspect Ratio are varied within $\mathcal{D}$.

We choose as reference domain the rectangle with $A = 2$ and the affine transformation is given by:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathcal{T}^{\mathrm{aff}}(\widehat{\boldsymbol{x}}, \mu) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{bmatrix} \frac{\mu}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} \widehat{x} \\ \widehat{y} \end{pmatrix}. \tag{89}$$

For this simple geometry, the Piola transformation $\mathcal{P}(\mu, \boldsymbol{x})$, is simply the inverse of $G^{\mathrm{aff}}(\mu)$.

Operatively, it is sufficient to compute once the reduced order matrices on the reference domainand the right hand side vector for all the Reduced Basis functions $\{\widehat{\boldsymbol{\zeta}}^{\mathrm{div}}\}$. Then, during the online phase, each entry of the matrices is multiplied by the corresponding $G_{ij}(\mu)$ in order to evaluate the differential forms on the basis set $\boldsymbol{\zeta}_i^{\mathrm{div}}$ instead of $\widehat{\boldsymbol{\zeta}}_i^{\mathrm{div}}$.

For the time evolution of the online system we chose third order semi-implicit Backward Difference Formula (BDF3).

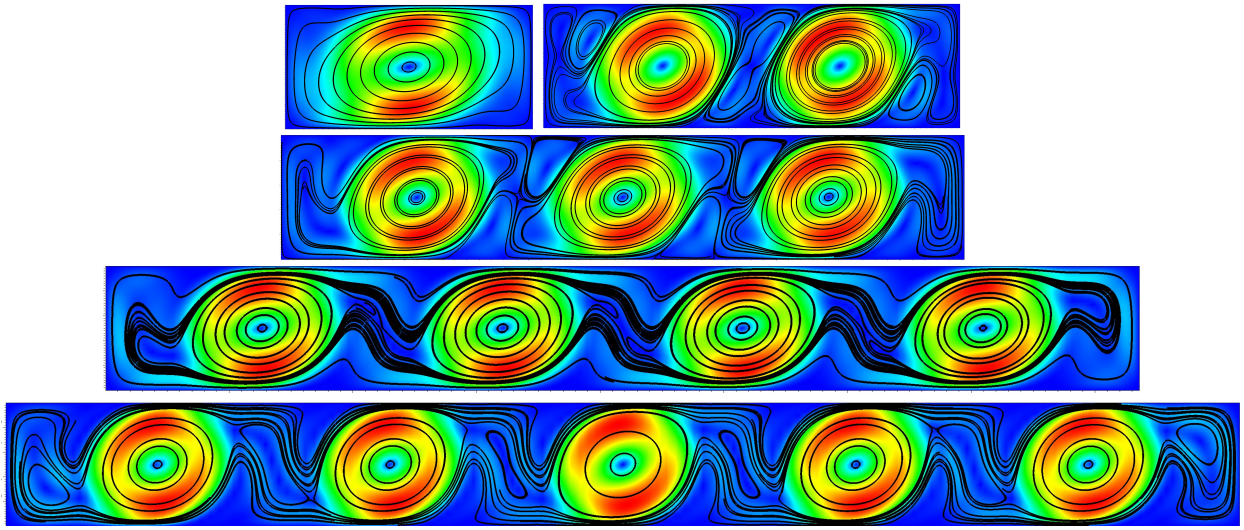

Figure 6: Some snapshots for the GAMM benchmark. From top to bottom, and from left to right: ($\mu = 2, \mathrm{Gr} = 50$); ($\mu = 3.37, \mathrm{Gr} = 264.9$); ($\mu = 5.52, \mathrm{Gr} = 132.1$); ($\mu = 8.36, \mathrm{Gr} = 50.77$); ($\mu = 10, \mathrm{Gr} = 100$). All snapshots are steady state, except the one with three rolls.

6.1. Bifurcation diagram for $A = 4$

In the first test case the geometric aspect ratio is fixed to $A = 4$, and the snapshots are collected until the tolerance computed as in (27) falls below 10^{-4}. The sampling procedure provided 13 snapshots for $\mathrm{Gr} \in [40 \cdot 10^3, 1 \cdot 10^6]$, of which 7 are steady state solutions, and the remaining 6 are computed by a POD of 2 time periodic solutions, with an L^2 energy threshold fixed to 99.9%. To test the proposed methodology (CVT-POD, SEM, Piola transformation) in the usual RB framework we try to rebuild a known bifurcation diagram with the reduced model, with the aim to reconstruct the different solution branches.

According to [63], for this configuration there are three steady solutions up to $\mathrm{Gr} = 120 \cdot 10^3$, after which a Hopf bifurcation occurs and the flow becomes unsteady. The three branches of steady solutions are characterized by a single roll flow up to $\mathrm{Gr} = 25 \cdot 10^3$, then flows with two rolls are stable for $\mathrm{Gr} \leq 100 \cdot 10^3$, and finally three roll flows exist from this last point up to the Hopf bifurcation. The steady state retained snapshots are almost equally distributed between one, two and three roll flows. The time-dependent snapshots are taken at the extreme ends of the Grashof parameter domain, very close to the Hopf bifurcation and at the $\mathrm{Gr} = 1 \cdot 10^6$ extreme. Both the time-dependent simulations required 3 POD modes to store the 99.9% of the energy.

To build the bifurcation diagram, we performed first a reduced time-dependent simulation with an homogeneous initial condition, and for the lowest value of the Grashof number, set at $\mathrm{Gr} = 40 \cdot 10^3$. We remark again that in [63] was used a different definition for the Grashof number, hence the value to be used for the numerical simulations is $\mathrm{Gr} = 10 \cdot 10^3$. This run evolved to a steady state solution (within an L^2 tolerance of 10^{-8}), and successfully rebuilt the first snapshot, consisting in the one roll flow typical of the low Grashof solutions.

Then, some time-dependent simulations are carried out with increasing values of the Grashof number, using as initial condition the result of the previous simulation. The continuation method has been used until the Hopf bifurcation was reached. Finally, the continuation method has been run backwards, with decreasing values of the Grashof number. The results obtained are synthesized in figure 7, where the evolution of the horizontal velocity at a fixed point is plot as a function of the Grashof number. The three branches can clearly be identified, and some hysteresis is present, especially at the one to two rolls transition. For clarity reasons, we also plot the streamlines of some representative solutions.

Note that to build the bifurcation diagram we had to compute 24 solutions for different values of the Grashof number with, $N = 13$ basis functions. The online procedure required slightly less than 600 seconds,

while the full order run would have taken about 24 cpu-hours for each solution, thus requiring approximately 576 cpu-hours to complete, on a IBM iDataPlex DX360M3 cluster. We estimate that in this way we reduced the computation time approximately by a factor of two:

$$\frac{\text{offline computation time + online computation time}}{\text{equivalent fully offline computation time}} \simeq \frac{13 \cdot 24 + \frac{1}{6}}{24 \cdot 24} = 54\%,$$

but this savings estimate can be much better when the number of offline runs becomes large, as we shall see in the next numerical test.

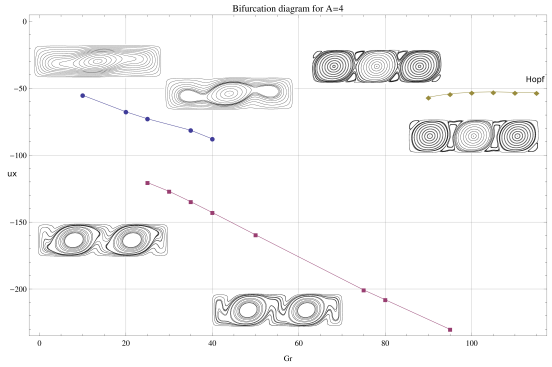

Figure 7: Bifurcation diagram for an aspect ratio of 4. The graph shows the horizontal velocity at the point $(0.7, 0.7)$ as a function of the Grashof number. The three lines are associated to the solutions with 1, 2, and 3 rolls, and the streamlines of some representative solutions are plotted for clarity.

6.2. Stability regions

After the satisfactory results of the fixed-geometry test, we consider the more general case with parametrized geometry. The parameter space $\mathcal{D}$ is now two dimensional, in particular we choose $\mathrm{Gr} \in [50 \cdot 10^3, 1 \cdot 10^6]$ and $A \in [2, 6]$.

At first, we collect a sufficient set of snapshots to extract the information needed for the online phase. We start from a set of 5 snapshots, corresponding to the four vertices and the center of the parameter rectangle $\mathcal{D}$. Such snapshots are used as a starting set for the CVT sampling procedure exposed in section 3.2. For the sake of clarity, we report a scheme of the actual algorithm (Algorithm 2).

We remark that the sampling procedure exposed in Algorithm 2 is divided between two basis sets: one for the steady snapshots and another for the time-dependent snapshots. This separation is necessary because, due to the time averaging errors, the tolerance of the steady state snapshots as computed in equation (91) is an order of magnitude lower than the tolerance for the time-dependent snapshots. If the CVT were performed without this precaution, the steady state regions of $\mathcal{D}$ would be heavily underrepresented, or conversely the number of time-dependent snapshots would be too large. The selection of suitable tolerance criteria for the two sampling procedures is a difficult task, and we chose to set the two tolerance values treating the steady-state and time-dependent cases as separate, imposing tolerances based on experience and on the realizable error bounds.

In this test, the sampling provided a total of 108 basis functions, arising from 30 steady and 21 unsteady snapshots. Almost all the unsteady snapshots required 3 POD modes to store the 99.9% of the energy, except for a few cases where 2 modes were enough. We report in figure 8 the position of the snapshots on the parameter plane (A, Gr), as computed by the Algorithm 2 (CVT).

As a basic validation for the online solver, we run several online simulations with all the 108 basis functions, for the parameters $\boldsymbol{\mu}^i$ used for the snapshot computation where the solution is known to be unique. The steady state snapshots have been correctly rebuilt by the online simulation, with L^2 differences between the snapshots and their online reconstruction alway of the order of the sampling tolerance.

Algorithm 2 The Centroidal Voronoi Tessellation sampling strategy used in section 6.2.

1: **repeat**
2: build the Delaunay triangulation for the points in the parameter space $\boldsymbol{\mu}^i \in \mathcal{D}$
3: find the triangle $K_i \subset \mathcal{D}$ with the largest residual

$$\texttt{res}_{K_i} = \sum_{\boldsymbol{\mu}^j \in K_i} \left\| \boldsymbol{u}(\boldsymbol{\mu}^j) - \boldsymbol{u}^N(\boldsymbol{\mu}^j) \right\|_{L^2} \tag{90}$$

4: compute the barycenter (Gr^i, μ^i) of K_i
5: compute the new snapshot $\boldsymbol{u}(\boldsymbol{\mu}^i)$
6: compute the corresponding basis function $\boldsymbol{\zeta}_i^{\mathrm{div}}$
7: update the tolerance:

$$\texttt{tol} = \max_i \left(\boldsymbol{u}(\boldsymbol{\mu}^i) - \sum_{j=1}^{i} (\boldsymbol{u}(\boldsymbol{\mu}^i), \boldsymbol{\zeta}_j^{\mathrm{div}})_{L^2} \boldsymbol{\zeta}_j^{\mathrm{div}} \right) \tag{91}$$

8: **until** $\max_{\mu \in \Sigma} \Delta(\mu) < \texttt{tol}$

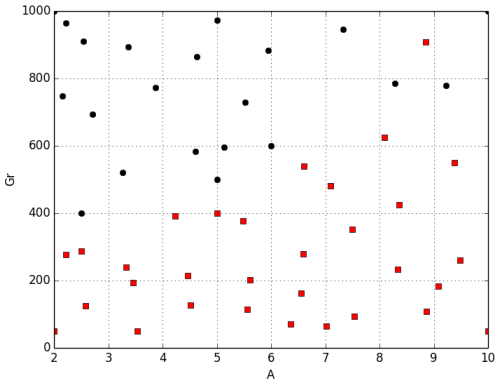

Figure 8: Parameters selected by the sampling Algorithm 2. The red square mark denotes the steady state snapshots, the black circle marks the time-dependent snapshots. The Grashof number is expressed in thousands.

On the other hand, the unsteady simulations show some discrepancies with the snapshots in terms of the oscillation's frequency, but the L^2 difference between the average values of the offline and online simulation are of the same order of magnitude of the sampling tolerance. We suspect that the frequency discrepancy of the online phase should be ascribed to the absence of the rejected POD modes that could be responsible for the frequency drift.

Despite the frequency drift during the online phase, the behaviour of the Hopf bifurcation frequency shows the expected decay as the geometry is varied. In figure 9, we plot the frequency at the onset of Hopf bifurcation for the one, two and three roll cases, as a function of the cavity length A. The asymptotic decay is confirmed by the benchmark [63], and is physically justified since for longer cavities the stability area of one roll flows shrinks to very low Grashof numbers, and consequently the flow must happen at a low velocity in order to mantain the stability of this very long vortex. Also, we note that sudden changes in the stability area happening about $A = 4.8$ in figure 10 are reflected by the abrupt change of sign of the derivative of the red curve in figure 9.

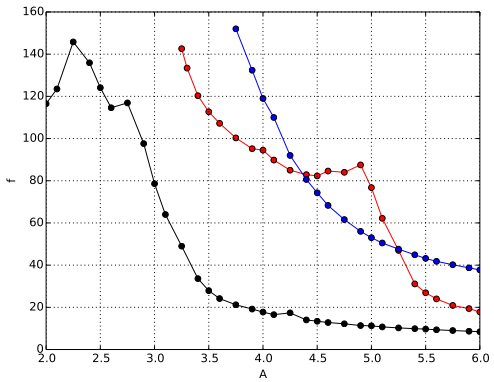

Figure 9: Frequency $f = 1/\omega$ at the onset of the periodic solution as a function of the geometric aspect factor A. The one roll solutions are marked in black, the two rolls solutions in red and the three rolls solutions in blue.

In particular, a remarkable feature of the online phase is that the basis arising from snapshots that are physically impossible for a certain flow configuration are cancelled out by the Galerkin projection. For instance, if we fix $A = 4$, the only admissible solutions for this geometry display one, two, or three rolls, and the basis arising from four or five rolls snapshots are automatically excluded by the online projection (the corresponding coefficients begin very close to machine zero).

Next, we investigate the stability of the reduced basis solutions using the techniques exposed in section 5.2, with the goal of identifying the stability regions of each flow pattern. In particular, we are interested in finding the curves in the parameters' plane that limit the stability of steady flows with one, two, three or four rolls. These curves are formed either by an envelope of steady bifurcation points or Hopf bifurcation points.

First, we compute the Hopf bifurcation point for a series of values of $A \in [2, 6]$. We adopt a subset of the 108 basis consisting of the 25 basis with a single roll. At each run a continuation method on Gr is used, and the reduced system is evolved until a new steady state is reached, with the modulus of the time derivative below a tolerance of 10^{-8}. After the steady solution is reached, the eigenvalues of the reduced linearized Navier-Stokes operator (88) are computed, and the Hopf bifurcation is marked for the Gr whose maximum eigenvalues have a real part sufficiently close to zero.

This operation is repeated for the solution branch with two roll basis (of dimension 14), and for the three and four roll branches (dimension 22 and 9 respectively). For each subset, the geometry parameter A range is bounded below by the snapshot with the lowest value of A.

Then, we compute the steady state bifurcations by running reduced order simulations using basis sets

formed by all the possible couples of snapshots, to be sure that all the bifurcation points are identified. For instance, running a first reduced order simulation with the basis sets arising from one and two rolls snapshots guarantees that we identify the regions where the one roll flows lose stability in favour of two roll flows (this is the most frequent case for low values of the geometric parameter. Then, running a reduced order simulation with basis functions arising from snapshots with one and three rolls, we detect the bifurcation points at which one roll flows become unstable in favour of three roll flows (typical for larger values of the geometric parameter), and so on with all the possible permutations of basis functions.

For a fixed geometrical parameter A, the simulation is started for a Gr not too far from the Hopf bifurcation, then Gr is gradually reduced until there is a sign change in an eigenvalue of the matrix (85). The value of Gr for which there is a steady bifurcation is used to draw the stability region, although the existence region of a flow pattern might be larger due to hysteresis, as shown in the bifurcation diagram 7 and by the reference [63]. In this paper however, we focus on the determination of the bifurcation points and less on the precise computation of stability regions.

We report in figure 10 the results of such analysis, performed for $A \in [2, 6]$ and Gr $\in [50 \cdot 10^3, 1 \cdot 10^6]$. Notice that the one roll flows are stable for all the geometrical aspect ratios if the Grashof number is sufficiently low. The two, three and four rolls flows on the other hand are stable only for some aspect ratios, and for sufficiently large Grashof numbers. Overall, the stability results obtained with the reduced order run are consistent with the high-order results, and the stability isles are visually similiar to the very accurate results of [63].

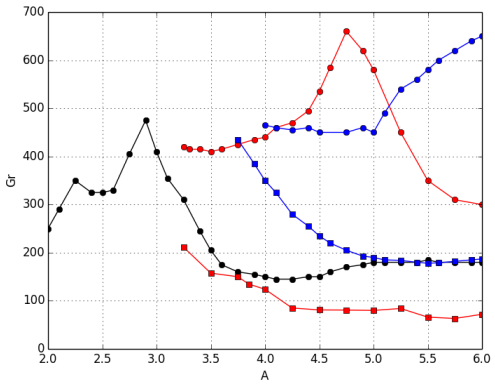

Figure 10: Stability regions in the parameters plane for the one roll flows (in black), the two rolls flows (red) and the three rolls flows (blue). The Hopf bifurcation points are marked with the circles, the steady bifurcation points with the square marks. The Grashof number is expressed in thousands.

Finally, we remark that the construction of the diagram in figure 10 with the reduced order method required the computation of 103 bifurcation points, each taking from 2 to 15 solutions for different values of the Grashof number. This required roughly 43 cpu-hours on a desktop computer, while the same computation whithout using the reduced order method would have taken roughly 5000 cpu-hours. Taking into account the offline phase (the computation of 108 basis functions and the POD-related computations), we get the following estimate for the computational time reduction:

$$\frac{\text{offline computation time} + \text{online computation time}}{\text{equivalent fully offline computation time}} \simeq \frac{103 \cdot 24 + 20 + 103 \cdot 10 \cdot \frac{1}{20}}{103 \cdot 10 \cdot 24} = 10\%,$$

where the numbers refer to 20 hours for the POD computations, and we consider an average of 10 runs per bifurcation diagram, 103 bifurcations computed, and 5 minutes per run during the online phase and 24 hours per run during the offline phase.

6.3. Time-dependent results

Finally, the accuracy of the ROM is tested with a time-dependent benchmark. For this purpose, we compare the results of the high order simulation with the results of the ROM simulation for some representative values of the parameters. In particular, we focus on time-periodic simulations with all the possible flow structure in the parameter range considered in this work. Specifically, we simulate parameter values for which a flow structure spanning from one to four vortices is expected.

For each simulation a basis is built starting from the snapshots with the same number of vortices. In all cases, the reduced order simulations converge to the expected periodic solution. In figure 11 is draw a comparison of the horizontal velocity at a fixed point as a function of time, as computed with a high order simulation, and with the reduced order approximation. We remark that the reduced order dynamics is in good agreement with the high-order dynamics in terms of average value, oscillation amplitude and frequency. In particular, for the considered test case the relative error in oscillation frequency of the reduced model is quantified in 0.51%.

The computational reduction is also relevant, since the reduced order simulation takes less than 5 minutes on a personal computer, whereas the high-resolution simulation requires 24 cpu-hours on the CINECA PLX cluster (an IBM iDataPlex DX360M3).

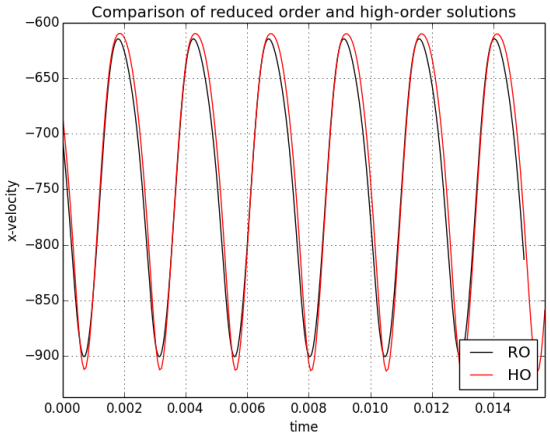

Figure 11: Comparison of the horizontal velocity at a the point $(0.7, 0.7)$ vs time for the high order (in red, HO) and reduced order (in black, RO) simulations. The parameters are Gr $= 963791$, $A = 2.22$, and the resulting flow has a single roll.

7. Perspectives

This work highlights several improvements with respect to the state of the art for parametrized viscous flow stability and bifurcation problems approached with ROM: in particular the focus has been put on approximation stability, proper sampling techniques and reduced eigenproblems for stability analysis and bifurcations detection. The exposition presents the topic from the different viewpoints of Nonlinear Mathematical Analysis, Applied Mathematics and Numerical Analysis, trying to underline the different aspects of the problem under consideration and by bridging these studies with offline high performance computing, on one side, and online advanced reduced order modelling, on the other one. Tests were performed on a well-known cavity benchmark problem.

We plan to extend the proposed ROM framework to more advanced 3D problems, in particular with applications to the study of the *Coanda effect* in hemodynamics [65], and the influence of geometry on symmetry breaking. More complex studies may be devoted to bifurcations and stability of flows with an elastic wall-structure interaction.

Acknowledgements

The authors acknowledge Dr. E. Merzari for his help with the `Nek5000` software and for the useful discussions, and the `Nek5000` community in general, Dr. F. Ballarin for the insights on approximation stability. G. Pitton has been supported by the pre-doc program at SISSA. G. Rozza acknowledges the support of NOFYSAS excellence grant program at SISSA and INDAM-GNCS activity group. The motivation for developing this work came from Prof. A.T. Patera (MIT) and from Prof. J. Rappaz (EPFL). We acknowledge Prof. F. Brezzi (IUSS, Pavia) for insights and some references.

We gratefully thank Prof. A. Quaini for ongoing collaboration on this topic with the Mathematics Department at University of Houston, USA.

The computing resources have been provided by the Sis14_COGESTRA cpu time grant allocation at CINECA, Bologna, Italy.

References

[1] A. Ambrosetti, G. Prodi, A Primer of Nonlinear Analysis, Cambridge University Press, Cambridge, 1993.
[2] S. Chandrasekhar, Hydrodynamic and Hydromagnetic Stability, Dover Publications, 1982.
[3] S. Timoshenko, J. Gere, Theory of Elastic Stability, Dover Civil and Mechanical Engineering, Dover Publications, 2009.
[4] J. Marsden, T. Hughes, Mathematical Foundations of Elasticity, Dover Civil and Mechanical Engineering, Dover Publications, 1994.
[5] G. Rozza, D. Huynh, A. Patera, Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations, Archives of Computational Methods in Engineering 15 (3) (2008) 229–275.
[6] P. Holmes, J. Lumley, G. Berkooz, C. Rowley, Turbulence, Coherent Structures, Dynamical Systems and Symmetry, Cambridge University Press, Cambridge, 2012.
[7] F. Terragni, J. Vega, On the use of POD-based ROMs to analyze bifurcations in some dissipative systems, Physica D: Nonlinear Phenomena (241.17) (2012) 1393–1405.
[8] H. Herrero, Y. Maday, F. Pla, RB (Reduced Basis) for RB (Rayleigh-Bénard), Computer Methods in Applied Mechanics and Engineering (261-262) (2013) 132–141.
[9] M. Yano, A. Patera, A space-time variational approach to hydrodynamic stability theory, Proceedings of the Royal Society A (496.2155).
[10] T. Lassila, A. Manzoni, A. Quarteroni, G. Rozza, Model order reduction in fluid dynamics: challenges and perspectives, in: A. Quarteroni, G. Rozza (Eds.), Reduced Order Methods for modeling and computational reduction, Vol. 9 of Modeling, Simulation and Applications, Springer, Milano, 2014, Ch. 9, pp. 235–273.
[11] A. Noor, J. Peters, Multiple-parameter reduced basis technique for bifurcation and post-buckling analyses of composite materiale, International Journal for Numerical Methods in Engineering (19) (1983) 1783–1803.
[12] B. Roux (Ed.), Numerical Simulation of Oscillatory Convection in Low-Pr Fluids, Vol. 27 of Notes on Numerical Fluid Mechanics and Multidisciplinary Design, Springer, 1990.
[13] A. Ambrosetti, A. Malchiodi, Nonlinear Analysis and Semilinear Elliptic Problems, Cambridge University Press, Cambridge, 2007.
[14] G. Caloz, J. Rappaz, Numerical analysis for nonlinear and bifurcation problems (1997).
[15] F. Brezzi, J. Rappaz, P. Raviart, Finite Dimensional Approximation of Nonlinear Problems. Part I: Branches of Nonsingular Solutions, Numerische Mathematik 36 (1) (1980) 1–25.
[16] F. Brezzi, J. Rappaz, P. Raviart, Finite Dimensional Approximation of Nonlinear Problems. Part II: Limit Points, Numerische Mathematik 37 (1) (1981) 1–28.
[17] F. Brezzi, J. Rappaz, P. Raviart, Finite Dimensional Approximation of Nonlinear Problems. Part III: Simple Bifurcation Points, Numerische Mathematik 38 (1) (1982) 1–30.
[18] S. P. Singh, M. R. Singh, Best approximation in nonlinear functional analysis, in: Q. Ansari (Ed.), Nonlinear Analysis, Trends in Mathematics, Springer, 2014, pp. 165–198.
[19] G. Golub, C. V. Loan, Matrix Computations, Johns Hopkins University Press, 2012.
[20] A. Quarteroni, A. Valli, Domain Decomposition Methods for Partial Differential Equations, Numerical Mathematics and Scientific Computation, Oxford University Press, 1999.
[21] A. Toselli, O. Widlund, Domain Decomposition Methods — Algorithms and Theory, Vol. 34 of Springer Series in Computational Mathematics, Springer Berlin Heidelberg, 2005.
[22] H. Dijkstra, F. Wubs, A. Cliffe, E. Doedel, I. Dragomirescu, B. Eckhardt, A. Gelfgat, A. Hazel, V. Lucarini, A. Salinger, E. Phipps, J. Sanchez-Umbria, H. Schuttelaars, L. Tuckerman, U. Thiele, Numerical bifurcation methods and their application to fluid dynamics: Analysis beyond simulation, Communications in Computational Physics 15 (1) (2014) 1–45.
[23] V. Temlyakov, Greedy approximation, Acta Numerica 17 (2008) 235–409.
[24] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, P. Wojtaszczyk, Convergence rates for greedy algorithms in reduced basis methods, SIAM Journal on Mathematical Analysis 43 (3) (2011) 1457–1472.
[25] A. Buffa, Y. Maday, A. Patera, G. Turinici, A priori convergence of the greedy algorithm for the parameterized reduced basis, ESAIM Mathematical Modelling and Numerical Analysis 46 (3) (2011) 595–603.

[26] Q. Du, M. Gunzburger, Model reduction by proper orthogonal decomposition coupled with centroidal voronoi tessellation, in: Proceedings of Fluids Engineering Division Summer Meeting, Vol. 1, ASME, 2002.
[27] Q. Du, M. Gunzburger, Centroidal voronoi tessellation based proper orthogonal decomposition analysis, Birkhauser Basel, 2003.
[28] M. Bergmann, T. Colin, A. Iollo, D. Lombardi, O. Saut, H. Telib, Reduced Order Models at Work in Aeronautics and Medicine, in: A. Quarteroni, G. Rozza (Eds.), Reduced Order Methods for Modeling and Computational Reduction, Vol. 9 of Modeling, Simulation and Applications, Springer, Milano, 2014.
[29] K. Ito, S. Ravindran, A reduced order method for simulation and control of fluid flows, Journal of Computational Physics 143 (2) (1998) 403–425.
[30] S. Volkwein, Proper orthogonal decomposition: Theory and reduced-order modelling (2013).
[31] B. Haasdonk, M. Ohlberger, Reduced basis method for finite volume approximations of parametrized linear evolution equations, ESAIM Mathematical Modelling and Numerical Analysis 42 (2) (2008) 277–302.
[32] N. Nguyen, G. Rozza, A. Patera, Reduced basis approximation and a posteriori error estimation for the time-dependent viscous burgers' equation, Calcolo 46 (2009) 157–185.
[33] P. Schmid, Dynamic mode decomposition of numerical and experimental data, Journal of Fluid Mechanics 656 (2010) 5–28.
[34] M. Barrault, Y. Maday, N. Nguyen, A. Patera, An "empirical interpolation method": application to efficient reduced-basis discretization of partial differential equations, C. R. Acad. Sci. Paris, Ser. I 339 (2004) 667–672.
[35] G. Rozza, K. Veroy, On the stability of the reduced basis method for Stokes equations on parametrized domains, Computer methods in applied mechanics and engineering (196) (2007) 1244–1260.
[36] Y. Saad, Iterative Methods for Sparse Linear Systems, SIAM, 2003.
[37] Y. Saad, Numerical Methods for Large Eigenvalue Problems, Revised Edition, Vol. 66 of Classics in Applied Mathematics, SIAM, 2011.
[38] N. Gräbner, S. Quraishi, C. Schröder, V. Mehrmann, U. von Wagner, New numerical methods for the complex eigenvalue analysis of disk brake squeal, in: EuroBrake 2014 Conference Proceedings, 2014.
[39] K. Veroy, A. Patera, Certified real-time solution of the parametrized steady incompressible navier-stokes equations: rigorous reduced-basis a posteriori error bounds, International Journal for Numerical Methods in Fluids 47 (8-9) (2005) 773–788.
[40] S. Deparis, G. Rozza, Reduced basis method for multi-parameter dependent steady navier-stokes equations: applications to natural convection in a cavity, Journal of Computational Physics 228 (12) (2009) 4359–4378.
[41] D. Huynh, G. Rozza, S. Sen, A. Patera, A successive constraint linear optimization method for lower bounds of parametric coercivity and inf–sup stability constants, Comptes Rendus Mathematique 345 (8) (2007) 473–478.
[42] F. Brezzi, M. Cornalba, A. D. Carlo, How to get around a simple quadratic fold, Numerische Mathematik 48 (1986) 417–427.
[43] K. Urban, A. Patera, A new error bound for reduced basis approximation of parabolic partial differential equations, Comptes Rendus Mathematique 350 (3-4) (2012) 203–207.
[44] L. Evans, Partial Differential Equations, Vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, 2010.
[45] L. Landau, E. Lifshitz, Fluid Mechanics, Vol. 6 of Course of Theoretical Physics, Butterworth Heinemann, Oxford, 1987.
[46] M. Deville, P. Fischer, E. Mund, High-Order Methods for Incompressible Fluid Flow, Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, Cambridge, 2002.
[47] C. Canuto, M. Hussaini, A. Quarteroni, T. Zhang, Spectral Methods Fundamentals in Single Domains, Scientific Computation, Springer, 2006.
[48] C. Canuto, M. Hussaini, A. Quarteroni, T. Zhang, Spectral Methods Evolution to Complex Geometries and Applications to Fluid Dynamics, Scientific Computation, Springer, 2007.
[49] P. Fischer, J. Lottes, S. Kerkemeier, Nek5000 Web page, http://nek5000.mcs.anl.gov (2008).
[50] A. Tomboulides, J. Lee, S. Orszag, Numerical Simulation of Low Mach Number Reactive Flows, Journal of Scientific Computing 12 (2) (1997) 139–167.
[51] D. Boffi, F. Brezzi, M. Fortin, Mixed Finite Element Methods and Applications, Vol. 44 of Springer Series in Computational Mathematics, Springer, Heidelberg, 2013.
[52] G. Rozza, D. Huynh, A. Manzoni, Reduced basis approximation and a posteriori error estimation for Stokes flows in parametrized geometries: roles of the inf—sup stability constants, Numerische Mathematik 125 (1) (2013) 115–152.
[53] A.-L. Gerner, K. Veroy, Certified reduced basis methods for parametrized saddle point problems, SIAM Journal on Scientific Computing 34 (5) (2012) A2812–A2836.
[54] F. Ballarin, A. Manzoni, A. Quarteroni, G. Rozza, Supremizer stabilization of POD-Galerkin approximation of parametrized Navier-Stokes equations, International Journal for Numerical Methods in Engineering (13).
[55] D. Rovas, Reduced-basis output bound methods for parametrized partial differential equations, Ph.D. thesis, Massachusetts Institute of Technology (2003).
[56] D. Amsallem, C. Farhat, Stabilization of projection-based reduced-order models, International Journal for Numerical Methods in Engineering 91 (2012) 358–377.
[57] W. Dahmen, How to best sample a solution manifold?, in: G. Pfander (Ed.), Sampling Theory - a Renaissance, Applied and Numerical Harmonic Analysis, Springer-Birkhäuser, 2015.
[58] A. Abdulle, O. Budác, A petrov-galerkin reduced basis approximation of the Stokes equation in parametrized geometries, Tech. Rep. MATHICSE Technical Report 35.2014, École polytechnique fédérale de Lausanne (aug 2014).
[59] A. Løvgren, Y. Maday, E. Rønquist, A reduced basis element method for the steady stokes problem, ESAIM: Mathematical

Modelling and Numerical Analysis 40 (3) (2006) 529–552.
[60] A. Caiazzo, T. Iliescu, J. Volker, S. Schyschlowa, A numerical investigation of velocity–pressure reduced order models for incompressible flows, Journal of Computational Physics 259 (2014) 598–616.
[61] G. Galdi, Navier-Stokes Equations: a Mathematical Analysis, in: R. Meyers (Ed.), Mathematics of Complexity and Dynamical Systems, Springer, 2011, pp. 1009–1042.
[62] K. Cliffe, E. Hall, P. Houston, E. Phipps, A. Salinger, Adaptivity and a Posteriori Error Control for Bifurcation Problems III: Incompressible Fluid Flow in Open Systems with O(2) Symmetry, Journal of Scientific Computing 52 (1) (2012) 153–179.
[63] A. Gelfgat, P. Bar-Yoseph, A. Yarin, Stability of multiple steady states of convection in laterally heated cavities, Journal of Fluid Mechanics (388) (1999) 315–334.
[64] H. Elman, K. Meerbergen, A. Spence, M. Wu, Lyapunov Inverse Iteration for Identifying Hopf Bifurcations in Models of Incompressible Flow, SIAM Journal on Scientific Computing 34 (3) (2012) A1584–A1606.
[65] A. Quaini, R. Glowinski, S. Čanić, Symmetry breaking and Hopf bifurcation for incompressible viscous flow in a contraction-expansion channel, submitted, University of Houston, Dept. of Math. Technical Report, 2014.