University of Trieste

Department of Mathematics and Geosciences
Master degree in Mathematics

REDUCED ORDER METHODS FOR PARAMETRIC OPTIMAL FLOW CONTROL PROBLEMS

APPLICATIONS IN ENVIRONMENTAL AND MARINE SCIENCES AND ENGINEERING

Candidate:
Maria Strazzullo

Advisor:
Prof. Gianluigi Rozza

Co-advisors:
Prof. Renzo Mosetti
Dr. Francesco Ballarin

Year 2015–2016

Ad Anna, Aurelio
e Francesco

Abstract

This master thesis aims at using reduced order methods in optimal control applications governed by parametrized partial differential equations. From the results obtained we will deduce how much useful reduced methods could be in several scientific and engineering fields. Optimal control problems are widely exploited in many modeling researches. They are computationally very demanding. A numerical method capable to reduce the dimensionality of the problem is an indispensable tool for optimization problems, most of all, in the case where several physical and geometrical configurations have to be considered. Model order reduction is a way to reduce computational costs of the simulations and, despite that, to reach precise results. The reduced basis methods allow to solve these parametrized optimal control problems in a rapid and in an accurate way. Among this work, we focused our analysis on optimal control problems characterized by quadratic cost functional to minimize constrained to linear parametric partial differential equations. We recast them in a saddle-point formulation in order to exploit the consolidated knowledge of this kind of structure.
The reduced basis method is introduced as a Galerkin projection into reduced spaces, generated by basis functions chosen through a proper orthogonal decomposition algorithm. The resolution procedure is divided in two stages: the offline stage, when the basis and the space are built, and the online where the projection is made and the problem is solved. The theoretical knowledge on the reduced methods will be applied to several test cases: they will assert the potentiality of this numerical approach. At the end, the method will be exploited in the field of environmental marine sciences and engineering. In this specific context the great versatility of the method will be shown. Two explicative applications are proposed: a large scale climatological application and a small scale pollutant control in the Gulf of Trieste. The first one is inserted in forecasting modeling and data assimilation context, the second is about the interest in the safeguard of Gulf of Trieste and surrounding areas.

Sommario

Questa tesi si propone di utilizzare metodi ridotti per problemi di controllo ottimo parametrici vincolati a equazioni alle derivate parziali. I risultati ottenuti ci condurranno a dedurre quanto questo approccio possa essere utile in vari contesti applicativi negli ambiti scientifici ed ingegnerisitici. I problemi di controllo ottimo, sebbene molto sfruttati da varie branche della ricerca scientifica, sono computazionalmente molto complessi, persino quando non dipendono da parametri. A maggior ragione, nel caso in cui varie configurazioni fisiche e/o geometriche fossero presenti, un metodo capace di ridurre la dimensionalità del problema di ottimizzazione, potrebbe risultare un utile strumento per risparmiare i costi computazionalii dovuti alla simulazione. Il metodo delle basi ridotte consente di risolvere in maniera rapida e accurata questa tipologia di problemi parametrizzati. La nostra analisi si è concentrata maggiormente su problemi di controllo quadratici nel funzionale e lineari nell'equazione di stato. Essi sono stati studiati nella loro formulazione punto sella, in modo da poter sfruttare le conoscenze teoriche già consolidate per questo particolare tipo di struttura.
In questo lavoro è introdotto il metodo delle basi ridotte come proiezione di Galerkin in uno spazio di dimensioni ridotte, generato da funzioni di base scelte in maniera opportuna tramite l'algoritmo di Proper Orthogonal Decomposition. La procedura computazionale viene divisa in una fase offline/online: nella prima fase si genera la base, nella seconda si attua la proiezione e avviene la risoluzione del problema ridotto.
Quanto introdotto teoricamente, verrà poi applicato a diversi casi test per verificarne e sottolinearne la potenzialità. Infine, per mostrare la grande versatilità del metodo, esso verrà utilizzato in un campo particolare di applicazione: le scienze e l'ingegneria ambientale in ambito marino. Verranno proposti due esempi: uno di tipo climatologico in grande scala, dove il controllo viene sfruttato nel contesto di un'eventuale assimilazione dati; il secondo incentrato sul controllo di un inquinante marino nel Golfo di Trieste. In entrambi i casi verrà sottolineato quanto il metodo permetta di ottenere soluzioni affidabili rispetto a quelle ottenute tramite le classiche tecniche di discretizzazione.

Contents

List of Figures

List of Tables

Introduction

This master thesis aims at analysing numerical methods for optimal control problems governed by parametric Partial Differential Equations (PDEs) involved in environmental marine applications. Optimal control problems (OCPs) are usually very complex and demanding, computationally speaking. Then, the goal of this work is to propose a rapid and suitable approach based on model order reduction: it allows to solve parametric optimal control problems (OCP($\boldsymbol{\mu}$)s) in a low dimensional framework. The final ambition is to compare the performance of the model reduction application with the results obtained through a full order approach. In this thesis, with a full order solutions we refer to Finite Element discretization (FE), whereas with model order reduction techniques we indicate Reduced Basis (RB) approximation. With reduced basis methods, we mean also basis built by Proper Orthogonal Decomposition (POD).
Among all the work, some theoretical and numerical examples of OCPs and OCP($\boldsymbol{\mu}$)s are presented, governed by different state equations: from Laplace to Stokes problems, from advection diffusion to quasi-geostrophic state equations. This thesis is a full breath composition, and shows how reduced order methods can be useful tools to exploit, in order to solve low dimensional optimal control systems, rather than full order optimal control problems, demanding and costly from the computational point of view.
Even if this thesis wants to analyse RB methods and to apply them in marine environmental control, the importance of model reduction methods can be underlined in various fields of knowledge (i.e. see [46, 47, 38]). If we limit to control problems, we can affirm that they are widespread in several engineering applications (we refer to [15, 59, 64, 54, 16]). For this reasons we decided to focus on OCPs in a total general framework. The same we did for Reduced Basis approach: we analysed it in a generic formulation, in order to show how they remarkably simplify the structure of the problem and how much computational cost is saved.
In general, in engineering fields, a major issue is represented by the prediction of quantities in different physical and geometrical configurations: for this reason control problems have wide spread influence and applications. The several scenarios of optimization models are described by physical or geometrical parameters, that might change at every evaluation. In this sense exploiting RB methods will provide accurate solutions in a more rapid way than high fidelity resolution does. One can analyse different control problems: time dependent, non-linear, distributed, on the boundary, etc. In the applications proposed among the chapters, we have focused on control problems with quadratic functional to minimize constrained to linear state equations.
In this thesis, the applications of properly reduced optimal control problems specifically involve examples in environmental marine science. The interest in this kind of application is twofold:

1. on one side, the interest is based on the study of how human activities are changing our planet and how much we have to regard our surroundings. Environmental sci-

ences and environmental engineering are fields of knowledge that are growing and are attracting many resources from several applications. The respect and the safeguard of the environment is a fundamental principle to live in a sustainable way and not to reach unpleasant life conditions.

2. on the other side the interest is totally and purely mathematical: modeling and numerical mathematics can afford environmental problems, allow to simulate trends and to forecast future configurations and, then, to prevent undesirable effects on the environment and on human beings. Numerical analysis with mathematical modeling and scientific computing are a great instrument to reach a deep comprehension of natural phenomena and is adapt to interpret various results. In general, different configurations and scenarios are given by the use of different parameters. RB methods perfectly fit in this framework: they are a versatile tool capable to handle really complex parametric problems and to transform them in low dimensional systems, reducing the computational time of resolution.

In the final part of the work, the dissertation moves toward a specific analysis of two examples:

1. an Oceanographic state solution tracking problem, governed by quasi-geostrophic equation. The interest in this very peculiar example is linked to global forecasting models and to prevision of future climatological scenarios on the Atlantic Ocean;

2. marine pollutant control related to the Gulf of Trieste and the surrounding areas: more precisely, we have built a pollutant control problem governed by advection-diffusion equations and simulate it in the Gulf of Trieste. The significance of this example lies in the control of the damages that marine pollution can cause to the seaside, the coast, to *flora* and *fauna* population and to inhabitants.

In view of the considerations that we have shown up to this point, we have structured the thesis in the following way.

Chapter 1
In this chapter the general theoretical formulation of a nonlinear OCP problem is briefly presented. Firstly, the Lagrangian approach in Banach spaces is discussed, to move toward the linear quadratic application in Hilbert spaces. Then, the saddle-point formulation for linear quadratic control is discussed, analysing the well-posedness of the problem. Finally, some theoretical examples in linear quadratic control problems are shown (the governing equations considered are Laplace, advection-diffusion and Stokes, respectively).

Chapter 2
The second chapter aims at discussing how an optimal control problem can be numerically approximated. After a brief introduction on the discretization techniques usually used in this context, we moved to Galerkin approximation for linear quadratic control problems into the saddle-point framework. The well-posedness of the discrete problem is discussed: we focused on elliptic coercive state equations and Stokes equations. Finally some numerical examples of OCPs are shown: two distributed control problems with Laplace governing equations and Stokes governing equation, respectively.

Chapter 3
In this chapter reduced basis approximation for PDEs is introduced. The main ideas of

the method are described, focusing on Proper Orthogonal Decomposition (POD) as a way to build the reduced space. The Empirical Interpolation Method (EIM) is recalled, since it is needed in Oceanographic application proposed at the end of this chapter. All we have shown among this chapter, wil be exploited and adapted to OCP($\boldsymbol{\mu}$) in the following chapters.

Chapter 4
Here a reduced basis framework for the efficient solution of parametrized linear quadratic optimal control problems governed by coercive elliptic PDEs and Stokes equations is provided. The well-posedness of the RB approximation is proved. Finally, some numerical examples concerning what we treated in the theoretical analysis of the chapter are shown: some numerical test already faced in chapter 2 are proposed in the reduced version, whereas a first environmental application of OCP($\boldsymbol{\mu}$) governed by advection-diffusion equation is presented.

Chapter 5
The final chapter is completely dedicated to Reduced Basis method applied to environmental marine sciences. It is divided in two sections containing two different linear quadratic problems. The first is an Oceanographic climatological OCP($\boldsymbol{\mu}$) governed by quasi-geostrophic equation. We simulated the results on the Atlantic Ocean. The second section concerns a pollutant control problem governed by advection diffusion equations. We formulate the problem in the Gulf of Trieste and study different effects deriving from different weather conditions.

The simulations reported in this work have been done exploiting different softwares. For the full order solutions, we used FEniCS (see [45] and for further information visit the website *https://fenicsproject.org/*). The reduced systems have been built through RBniCS library. The meshes of the Atlantic Ocean and of the Gulf of Trieste have be obtained trough *Freefem++* (see [32], and visit *http://www.freefem.org/*) and *Gmsh* (see as a reference [27], and visit *http://gmsh.info/*).

This work has been carried out in collaboration with the Scuola Internazionale Superiore di Studi Avanzati (SISSA), Mathematics Area, mathLab, and with the Istituto Nazionale di Oceanografia e Geofisica Sperimentale (OGS).

UNITS, University of Trieste,
SISSA, International School for Advanced Studies,
OGS, National Institute of Oceanography and Experimental Geophysics.

Trieste, March 2017.

Chapter 1

A Theoretical Introduction to Optimal Control

The aim of this chapter is to briefly introduce optimal control problems, what are their main features and how they can be formalized from the mathematical point of view. The following analysis is far to be exhaustive: control is a very wide topic. For those who want to analyse deeper this subject, we refer to [29, 24, 4]. In this context, an exhaustive discussion would be impossible. It would be far from our aims: we can only give the taste of how a powerful instrument control is and how well spread its applications are. These problems are a very challenging task and they had fascinated great scientists and minds. Furthermore, optimal control problems have a great impact on our life and they are reaching a certain degree of maturity with deeper studies in mathematics and engineering. They have several applications in very different fields, from natural science, environmental purposes, biology modeling, to industrial development and research (see [42, 18, 64]). We can formulate an Optimal Flow Control problem wherever a canal, an irrigation structure, a pipeline fluid network, a blood vessel, a dam is. In the following we will describe the general features of an optimal control problem.

Optimization or control aims at managing a physical quantity (e.g., flow rate and direction), or another fluid feature (e.g. temperature, concentration), in order to achieve a desired state.
The rigorous formulation of this problem needs essentially three elements:

1. an **objective**, describing what we want to reach through optimization. Mathematically, it is formalized by an *objective (or cost) functional*. There are several objectives used in different applications: flow tracking, prevention (delay) of turbulence and temperature variations, drag minimization and so on;

2. **control variables**, to be chosen in order to minimize the objective functional. When the control variable operates on the domain boundary we are treating a *boundary control* (it can represent an injection or a suction of fluid or a heat exchange or a cooling process). Otherwise, if the control variable acts on the total domain we are facing a *distributed control*. Last, we can talk about *shape controls* if we face domain design, shape optimization or surface roughness problems (i.e. see [49, 31, 17]);

3. **constraints** are the last ingredients. Their role is to set conditions on the optimizers. Constraints characterize the fluid model and they are represented by a set of *partial differential equations*. The are also referred as *state equations*.

Putting all together: solve an optimal control means to seek controls fulfilling constraints and minimizing an objective functional.

In this chapter the abstract formulation of a control problem is introduced. In section 1.1 the Lagrangian Formalism is introduced: existence results are presented and optimality condition for a generic nonlinear optimal control problem are derived. In section 1.2 we will refer to a different problem formulation based on a saddle-point approach. Section 1.3 focuses on linear quadratic optimal control problems. Some examples of distributed control are shown.

1.1 The Lagrangian Formulation

In this section a generic setting for a general steady PDE constrained optimal control problem is described. The theoretical development of this topics is due to *J.L. Lions*, which proved existence and uniqueness of the solution of optimal control problem governed by elliptic, parabolic and hyperbolic PDEs (see [44, 43]). However, this classical approach do not straightforwardly handle a wide class of problems: for example, nonlinear optimal control problems or boundary control problems (see e.g. [23, 56, 13, 14]).
There is also a complementary way to treat this kind of issues: the *Lagrangian approach*. Thanks to the definition of a Lagrangian functional, the optimal control problem can be seen in a constrained minimization formulation: if an optimal solution exists, it will make the derivative of the Lagrangian functional vanish. For proves and theoretical knowledge we refer to [29, 37, 35]. Banach and Hilbert Space theory is required (see [8, 75]).
Let us specify the notation used in the following. A capital letter X will indicate a Banach space. To refer to its dual, the symbol X^* is used, whereas $\langle \cdot, \cdot \rangle_{XX^*}$ indicates dual pairing of X and X^*.

1.1.1 The General Problem: Existence Results

First of all, let us discuss existence and uniqueness of solution for a general nonlinear optimal control problem. Let Y, U be reflexive Banach spaces and Z a Banach space. A generic optimal control problem (OCP from now on) can be formulated as follows:

$$\min_{(y,u)\in Y\times U} J(y,u) \qquad \text{subject to } \mathcal{E}(y,u)=0, \qquad u\in U_{ad}, y\in Y_{ad}. \tag{1.1.1}$$

where $J : Y \times U \to \mathbb{R}$ and $\mathcal{E} : Y \times U \to Z$ are continuous. To be more specific, in our context $J(y,u)$ represents the so called *cost functional*, whereas $\mathcal{E}(y,u) = 0$ is the governing state equation. The subsets $U_{ad} \subseteq U$ and $Y_{ad} \subseteq Y$ represent the control space and the state space, respectively. When $U_{ad} \subsetneq U$ it indicates some bounds on the control, whereas $Y_{ad} \subsetneq Y$ shows a constrain on the state solution. A problem is said unconstrained when $U_{ad} = U$ and $Y_{ad} = Y$. The following assumptions have to be considered to prove existence of an optimal control result:

1. J is sequentially weakly lower semi-continuous,
2. U_{ad} is convex, bounded and closed,
3. Y_{ad} is convex and closed,
4. state equation $\mathcal{E}(y,u) = 0$ has a bounded solution operator $u \in U_{ad} \mapsto y(u) \in Y$

5. $\mathcal{E} : Y \times U \to Z$ is continuous under weak convergence.

Thanks to these hypotheses, the next theorem holds (for the proof we refer to [35, Section 1.5.2]):

Theorem 1.1. *If (1)-(5) hold, then an OCP has an optimal solution* $(y_\star, u_\star)$.

Now that the solution existence is guaranteed, we will focus on solving the OCP.

1.1.2 Adjoint approach for Reduced Problem and Reduced Functional Derivative

Consider Y, U, Z Banach spaces and the general constrained OCP

$$\min_{(y,u)\in Y\times U} J(y,u) \quad \text{subject to } \mathcal{E}(y,u) = 0 \quad u \in U_{ad}. \tag{1.1.2}$$

where J and $\mathcal{E}$ have been chosen as cost functional and state equation, respectively. Next we describe the hypotheses needed to continue our analysis.

Assumptions 1.1. *Let* U_{ad} *be nonempty, closed, bounded and convex. Suppose that* J *and* $\mathcal{E}$ *are continuously Frechét differentiable and that the state equation verifies the following property: for all* $u \in U_{ad}$ *exists a unique* $y = y(u)$ *in* Y. *Additionally we assume that* $\mathcal{E}_y(y(u), u) \in \mathcal{L}(Y, Z)$ *has a bounded inverse for all* $u \in U_{ad}$[1].

If one substitutes $y(u)$ to the problem (1.1.2) obtains:

$$\min_{u\in U} \hat{J}(u) \quad \text{subject to } \mathcal{E}(y(u),u) = 0 \quad u \in U_{ad}. \tag{1.1.3}$$

This formulation is usually known as the *reduced problem*[2] and $\hat{J}(u) := J(y(u), u)$ is the so called *reduced functional*. Under assumptions 1.1 the problems (1.1.3) and (1.1.2) are equivalent. The minimization of $\hat{J}(u)$ is essentially based on the reduced functional derivative $\hat{J}'(u)$. There are typically two ways to represent $\hat{J}'(u)$: sensitivities analysis and adjoint approach. We will focus on the latter (for sensitivity method theory and examples see [35, 29]). For our purposes, an expression for $y'(u)$ is needed. It can be derived by differentiating $\mathcal{E}(y(u), u) = 0$ with respect to u:

$$\mathcal{E}_y(y(u),u)y'(u) + \mathcal{E}_u(y(u),u) = 0 \Rightarrow y'(u) = -\mathcal{E}_y(y(u),u)^{-1}\mathcal{E}_u(y(u),u). \tag{1.1.4}$$

Let us exploit this result to compute

$$\langle \hat{J}'(u), s\rangle_{U^*U} = \langle J_y(y(u),u), y'(u)s\rangle_{Y^*Y} + \langle J_u(y(u),u), s\rangle_{U^*U}$$

$$\langle y'(u)^* J_y(y(u),u), s\rangle_{Y^*Y} + \langle J_u(y(u),u), s\rangle_{U^*U},$$

where the apex * indicates the dual variable. From the previous expression one can deduce

$$\hat{J}'(u) = y'(u)^* J_y(y(u),u) + J_u(y(u),u).$$

We can deduce that the vector $y'(u)^* J_y(y(u), u)$ is required to compute reduce functional derivative. From (1.1.4)

$$y'(u)^* J_y(y(u),u) = -\mathcal{E}_u(y(u),u)^*(\mathcal{E}_y(y(u),u)^{-1})^* J_y(y(u),u).$$

[1] This assumption ensures that the state solution operator $u \mapsto y(u)$ is continuously differentiable.

[2] From chapter 3 on, with reduced problem we will indicate the Reduced Basis approximation of a problem.

This is equivalent to

$$y'(u)^* J_y(y(u),u) = \mathcal{E}_u(y(u),u)^* p(u),$$

where $p(u) \in Z^*$ is the *adjoint state* that solves the following *adjoint equation*:

$$\mathcal{E}_y(y(u),u)^* p(u) = -J_y(y(u),u). \tag{1.1.5}$$

So finally we obtain a new representation for the reduced functional derivative:

$$\hat{J}'(u) = J_u(y(u),u) + \mathcal{E}_u(y(u),u)^* p(u). \tag{1.1.6}$$

So $\hat{J}'(u)$ can be computed through two steps:

1. find the adjoint state $p = p(u) \in Z^*$ solving the adjoint equation (1.1.5),
2. compute $\hat{J}'(u)$ via (1.1.6).

1.1.3 Lagrangian Representation and First Order Necessary Conditions

Now we are going to analyse a different way to derive the adjoint equation deduced in the subsection (1.1.2). To reach our goal let us define $\mathscr{L} : Y \times U \times Z^* \to \mathbb{R}$ known as the *Lagrangian Functional*

$$\mathscr{L}(y,u,p) = J(y,u) + \langle p, \mathcal{E}(y,u)\rangle_{Z^*Z},$$

where $p \in Z^*$. We assume that hypotheses 1.1 are valid. We know that $u \mapsto y(u)$ uniquely. Substituting $y(u)$ in the Lagrangian functional the following essential equality is reached:

$$\mathscr{L}(y(u),u,p) = J(y(u),u) + \langle p, \underbrace{\mathcal{E}(y(u),u)}_{=0}\rangle_{Z^*Z} = J(y(u),u) = \hat{J}(u). \tag{1.1.7}$$

Differentiating one obtains

$$\langle \hat{J}'(u), s\rangle_{U^*U} = \langle \mathscr{L}_y(y(u),u,p), y'(u)s\rangle_{Y^*Y} + \langle \mathscr{L}_u(y(u),u,p), s\rangle_{U^*U}.$$

To have an explicit description of the reduced functional derivative, we want to find a special $p = p(u) \in Z^*$ such that

$$\mathscr{L}_y(y(u),u,p) = 0. \tag{1.1.8}$$

This equality implies

$$\begin{aligned}\langle \mathscr{L}_y(y(u),u,p), w\rangle_{Y^*Y} &= \langle J_y(y(u),u), w\rangle_{Y^*Y} + \langle p, \mathcal{E}_y(y(u),u)w\rangle_{Z^*Z} \\ &= \langle J_y(y(u),u) + \mathcal{E}_y(y(u),u)^* p, w\rangle_{Y^*Y} \qquad \forall w \in Y.\end{aligned}$$

Therefore we choose $p = p(u) \in Z^*$ such that

$$\mathcal{E}_y(y(u),u)^* p = -J_y(y(u),u), \tag{1.1.9}$$

obtaining the adjoint equation (notice how (1.1.9) and (1.1.5) coincide). For this particular choice of $p = p(u) \in Z^*$ the adjoint derivative representation can be deduced easily:

$$\hat{J}'(u) \underset{(1.1.8)}{=} \mathscr{L}_u(y(u),u,p) = J_u(y(u),u) + \mathcal{E}_u(y(u),u,p)^* p \tag{1.1.10}$$

Finally we have a direct representation for $\hat{J}'(u)$ totally equivalent to the one analysed in subsection 1.1.2, proved by the equality between (1.1.10) and (1.1.6).
Thanks to this description, we are able to handle some results about necessary optimality conditions. It holds the following theorem (see [35], Theorem 1.48)

Theorem 1.2 (Minimum Principle)**.** *Let us suppose that assumptions 1.1 are verified. If $u_\star$ is a local solution of the reduced problem (1.1.3) then the following inequality holds:*

$$\langle \hat{J}'(u_\star), v - u_\star \rangle_{U^*U} \geq 0, \qquad \forall v \in U_{ad}. \tag{1.1.11}$$

Exploiting the derived formulation for $\hat{J}'(u_\star)$, this statement follows:

Corollary 1.2.1. *Let $(y_\star, u_\star) \in Y \times U_{ad}$ be an optimal solution on the reduced problem (1.1.3). Suppose that assumptions (1.1) hold. Then there exists an adjoint state $p_\star \in Z^*$ such that the following conditions are verified:*

$$\begin{cases} \mathcal{E}(y_\star, u_\star) = 0, \\ \mathcal{E}_y(y_\star, u_\star)^* p_\star = -J_y(y_\star, u_\star) \\ \langle J_u(y_\star, u_\star) + \mathcal{E}_u(y_\star, u_\star)^* p_\star, v - u_\star \rangle_{U^*U} \geq 0 \qquad \forall v \in U_{ad}. \end{cases} \tag{1.1.12}$$

Using Lagrangian notation the system (1.1.12) is equivalent to:

$$\begin{cases} \mathscr{L}_p(y_\star, u_\star, p_\star) = 0 \\ \mathscr{L}_y(y_\star, u_\star, p_\star) = 0 \\ \langle \mathscr{L}_u(y_\star, u_\star, p_\star), v - u_\star \rangle_{U^*U} \geq 0 \qquad \forall v \in U_{ad}. \end{cases} \tag{1.1.13}$$

System (1.1.12) can be written in a weak form with respect to J and $\mathcal{E}$:

$$\begin{cases} \langle \mathcal{E}(y_\star, u_\star), q \rangle_{ZZ^*} = 0, & \forall q \in Z^* \\ \langle \mathcal{E}_y(y_\star, u_\star)^* p_\star + J_y(y_\star, u_\star), z \rangle_{Y^*Y} = 0, & \forall z \in Y \\ \langle J_u(y_\star, u_\star) + \mathcal{E}_u(y_\star, u_\star)^* p_\star, v - u_\star \rangle_{U^*U} \geq 0, & \forall v \in U_{ad}. \end{cases} \tag{1.1.14}$$

When $U_{ad} = U$ the latter inequality of (1.1.14) becomes[3]

$$\langle J_u(y_\star, u_\star) + \mathcal{E}_u(y_\star, u_\star)^* p_\star, v \rangle_{U^*U} = 0, \qquad \forall v \in U.$$

This conditions will be essential in the following applications of this work (from now on we will omit the star pedix for the optimal variable).

1.2 Saddle-Point Formulation

In the previous section we constructed optimality system for the general OCP formulation. When linear quadratic control problems are considered, then the optimality conditions theory leads to a saddle-point structure. Our problem is recasting in a mixed variational framework. This different approach is more usual than the classical Lagrangian method in order to treat linear quadratic optimal controls (i.e. in [33, 53, 64, 66]).
First we will introduce a general saddle-point system setting, focusing on existence and uniqueness results (see [57, 6, 5]). In a second analysis a connection between general saddle-point theory and constrained optimization problems is established. Finally some examples of distributed control will be shown.

[3] In this case (an unconstrained control problem), equations (1.1.12) can be seen as the *Euler Lagrange system* for the Lagrangian functional. Indeed an optimal solution for the OCP (1.1.13) represent a stationary point of $\mathscr{L}(\cdot,\cdot,\cdot)$:

$$\nabla \mathscr{L}(y_\star, u_\star, p_\star)[z, v, q] = 0 \qquad \forall (z, v, q) \in Y \times U \times Z^*.$$

1.2.1 Generic Problem Formulation and Existence Result

Let X and Q be two Hilbert Spaces, respectively endowed with the norms $\|\cdot\|_X, \|\cdot\|_Q$. The dual spaces X^* and Q^* are considered. Let us introduce two continuous bilinear forms $\mathcal{A}(\cdot,\cdot) : X \times X \to \mathbb{R}$ and $\mathcal{B}(\cdot,\cdot) : X \times Q \to \mathbb{R}$. Now, consider the functionals $F : X \to \mathbb{R}$ and $G : Q \to \mathbb{R}$ and the following *saddle-point* problem: find $(x, p) \in X \times Q$ such that

$$\begin{cases} \mathcal{A}(x,v) + \mathcal{B}(v,p) = \langle F, v\rangle & \forall v \in X, \\ \mathcal{B}(x,q) = \langle G, q\rangle & \forall q \in Q. \end{cases} \tag{1.2.1}$$

Now let us define the linear operators $A : X \to X^*$ and $B : X \to Q^*$ respectively derived from $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{B}(\cdot,\cdot)$ verifying the following relations:

$$\begin{aligned} \langle Aw, v\rangle_{X^*X} &= \mathcal{A}(w,v) && \forall w, v \in X, \\ \langle Bw, q\rangle_{Q^*Q} &= \mathcal{B}(w,q) && \forall w \in X, \forall q \in Q. \end{aligned}$$

Let $B^t : Q \to X^*$ be the transpose operator of B obtained by:

$$\langle Bw, q\rangle_{Q^*Q} = \langle w, B^t q\rangle_{X,X^*} \qquad \forall w \in X, \forall q \in Q.$$

So the system (1.2.1) can also be interpreted as:

$$\begin{cases} Ax + B^t p = F & \text{in } X^*, \\ Bx = G & \text{in } Q^*. \end{cases} \tag{1.2.2}$$

Now let us consider

$$X_0 = \{w \in X \mid \mathcal{B}(w,q) = 0, \ \ \forall q \in Q\} = \ker(B),$$

subspace of X. The existence and uniqueness of the solution of this saddle-point problem derive from the following well-known theorem (see [6] for the proof).

Theorem 1.3 (Brezzi)**.** *Assume that the Hilbert spaces X and Q, the functionals $F \in X^*$ and $G \in Q^*$, and the bilinear forms $\mathcal{A}(\cdot,\cdot) : X \times X \to \mathbb{R}$ and $\mathcal{B}(\cdot,\cdot) : X \times Q \to \mathbb{R}$ are given. Assume that the bilinear forms $\mathcal{A}(\cdot,\cdot) : X \times X \to \mathbb{R}$ and $\mathcal{B}(\cdot,\cdot) : X \times Q \to \mathbb{R}$ satisfy:*

1. *$\mathcal{A}(\cdot,\cdot)$ and $\mathcal{B}(\cdot,\cdot)$ are continuous, i.e. there exist $\gamma_{\mathcal{A}}, \gamma_{\mathcal{B}} > 0$ such that:*

$$|\mathcal{A}(w,v)| \le \gamma_{\mathcal{A}} \|w\|_X \|v\|_X \ \ \forall w, v \in X$$
$$\textit{and}$$
$$|\mathcal{B}(w,q)| \le \gamma_{\mathcal{B}} \|w\|_X \|q\|_Q \quad \forall w \in X, \forall q \in Q;$$

2. *$\mathcal{A}(\cdot,\cdot)$ is weakly coercive on X_0, i.e. there exist $\alpha_0 > 0$ such that:*

$$\inf_{w \in X_0} \sup_{v \in X_0} \frac{\mathcal{A}(w,v)}{\|v\|_X \|w\|_X} \ge \alpha_0 > 0 \quad \textit{and} \quad \inf_{v \in X_0} \sup_{w \in X_0} \frac{\mathcal{A}(w,v)}{\|v\|_X \|w\|_X} > 0;$$

3. *$\mathcal{B}(\cdot,\cdot)$ satisfies the inf-sup condition*

$$\beta = \inf_{q \in Q} \sup_{w \in X} \frac{\mathcal{B}(w,q)}{\|w\|_X \|q\|_Q} \ge \beta_0 > 0.$$

Then there exists a unique solution $(x, p) \in X \times Q$ to the problem (1.2.1) for all $F \in X^*$ and $G \in Q^*$. Moreover the following a priori estimates hold:

$$\|x\|_X \le \frac{1}{\alpha_0}\Big[\|f\|_{X^*} + \frac{\alpha_0 + \gamma_{\mathcal{A}}}{\beta_0}\|g\|_{Q^*}\Big],$$
$$\|p\|_Q \le \frac{1}{\beta_0}\Big[\Big(1 + \frac{\gamma_{\mathcal{A}}}{\alpha_0}\Big)\|f\|_{X^*} + \frac{\gamma_{\mathcal{A}}(\alpha_0 + \gamma_{\mathcal{A}})}{\alpha_0 + \beta_0}\|g\|_{Q^*}\Big].$$

1.2.2 Saddle-Point Structure for Optimization Problems

In subsection (1.2.1) we have analysed the generic structure of a saddle-point problem. Now we want to focus on the relation between this formulation and constrained linear-quadratic OCPs.
Let us consider $\Omega \subset \mathbb{R}^d$ an open and bounded domain with Lipschitz boundary $\Gamma = \partial\Omega$. Let Y and U be the Hilbert spaces for state and control variable respectively. The Hilbert observation space will be indicated with $Z \supset Y$. Taking into account another Hilbert space Q, the linear constraint equation is defined by:

$$a(y, q) = c(u, q) + \langle G, q\rangle \qquad \forall q \in Q, \tag{1.2.3}$$

where $a(\cdot,\cdot) : Y \times Q \to \mathbb{R}$ represents the state operator, $c(\cdot,\cdot) : U \times Q \to \mathbb{R}$ describes the role of the control and $G \in Q^*$ is acting as a forcing term. Given a constant $\alpha > 0$, the quadratic objective functional is given by:

$$J(y, u) = \frac{1}{2}m(y - y_d, y - y_d) + \frac{\alpha}{2}n(u, u), \tag{1.2.4}$$

where $y_d \in Z$ is an observation function, $m : Z \times Z \to \mathbb{R}$ is a bilinear form that defines the objective and $n : U \times U \to \mathbb{R}$ is a bilinear form representing a penalization term for the control variable. So an OCP problem can be formalized as follows:

$$\min_{(y,u)\in Y\times U} J(y, u) \qquad \text{such that } (y, u) \in Y \times U \text{ satisfies (1.2.3).} \tag{1.2.5}$$

Our aim is to recast the problem in a saddle-point framework. In order to reach this new formulation, let us define $X = Y \times U$. Being $\underline{x} = (y, u) \in X$ and $\underline{w} = (z, v) \in X$, we can endow X with the scalar product $(\underline{x}, \underline{w})_X = (y, z)_Y + (u, v)_U$ and with the norm $\|\cdot\|_X = \sqrt{(\cdot,\cdot)_X}$. Now let us consider the bilinear form $\mathcal{A}(\cdot,\cdot) : X \times X \to \mathbb{R}$ defined as

$$\mathcal{A}(\underline{x}, \underline{w}) = m(y, z) + \alpha n(u, v) \qquad \forall \underline{x}, \underline{w} \in X,$$

and the bilinear form $\mathcal{B}(\cdot,\cdot) : X \times Q \to \mathbb{R}$ as:

$$\mathcal{B}(\underline{x}, q) = a(z, q) - c(v, q) \qquad \forall \underline{w} \in X,\ \forall q \in Q.$$

Finally, let $F \in X^*$ be

$$\langle F, \underline{w}\rangle = m(y_d, z) \qquad \forall \underline{w} \in X,$$

and define a new functional as follows:

$$\mathcal{J}(\underline{x}) = \frac{1}{2}\mathcal{A}(\underline{x}, \underline{x}) - \langle F, \underline{x}\rangle.$$

Thanks to these relations we can give a new formulation to the objective functional

$$J(y, u) = \mathcal{J}(\underline{x}) + \mathcal{M}(y_d),$$

where $\mathcal{M}(y_d) = \frac{1}{2}m(y_d, y_d)$ is a constant term that does not give any contribution to the minimization of $J(\cdot,\cdot)$. For these reasons, it is now possible give a new formulation to the problem (1.2.5): find

$$\min_{\underline{x}\in X} \mathcal{J}(\underline{x}) \qquad \text{such that} \qquad \mathcal{B}(\underline{x}, q) = \langle G, q\rangle \qquad \forall q \in Q. \tag{1.2.6}$$

The constrained optimization problem (1.2.6) can be recast into an unconstrained optimization problem by defining the Lagrangian functional $\mathcal{L}(\cdot,\cdot) : X \times Q \to \mathbb{R}$ as

$$\mathscr{L}(\underline{x}, p) = \mathcal{J}(\underline{x}) + \mathcal{B}(\underline{x}, p) - \langle G, p\rangle. \tag{1.2.7}$$

By deriving with respect to $(\underline{x}, p) \in X \times Q$ we can build a saddle-point structure for the optimality system: find $(\underline{x}, p) \in X \times Q$ such that

$$\begin{cases} \mathcal{A}(\underline{x}, \underline{w}) + \mathcal{B}(\underline{w}, p) = \langle F, \underline{w}\rangle & \forall \underline{w} \in X, \\ \mathcal{B}(\underline{x}, q) = \langle G, q\rangle & \forall q \in Q, \end{cases} \tag{1.2.8}$$

where $p \in Q$ is the adjoint variable of the constraint equation. Existence and uniqueness of the solution are guaranteed under the assumptions of theorem 1.3.
What we have shown in this subsection is justified by the following theorem (see [5], Porposition 1.7):

Theorem 1.4. *Assume that the hypotheses of the (Brezzi) theorem 1.3 hold. Furthermore, let $\mathcal{A}(\cdot,\cdot)$ be a symmetric, nonnegative and coercive bilinear form on X_0 with coercivity constant $\alpha_0 > 0$, i.e.*

$$\mathcal{A}(\underline{x}, \underline{w}) = \mathcal{A}(\underline{w}, \underline{x}), \quad \mathcal{A}(\underline{x}, \underline{x}) \geq 0 \quad \forall \underline{x}, \underline{w} \in X, \quad \mathcal{A}(\underline{x}, \underline{x}) \geq \alpha_0 ||\underline{x}||_X^2 \quad \forall \underline{x} \in X_0.$$

Then the problem (1.2.8) is equivalent to the following constrained minimization problem:

$$\begin{cases} \min\limits_{\underline{x} \in X} \mathcal{J}(\underline{x}) = \dfrac{1}{2}\mathcal{A}(\underline{x}, \underline{x}) - \langle F, \underline{x}\rangle \\ \text{subject to } \mathcal{B}(\underline{x}, q) = \langle G, q\rangle \qquad \forall q \in Q. \end{cases} \tag{1.2.9}$$

1.3 Linear Quadratic Optimal Control Problems: Theory and Examples

In this section we will provide a theoretical formalization of linear quadratic OCPs. This kind of control problems have several applications and are wide spread and studied. This category includes either elliptic coercive problems (e.g. [54, 59, 58]) or Stokes problems (e.g [64, 53, 61]). We will be able to interpret the linear quadratic optimality system under the results of section 1.2, noticing its saddle-point structure (as a reference, see [35]). The theoretical approach will be enriched by some examples of distributed control problems with elliptic and Stokes governing equations.

1.3.1 Linear Quadratic Optimal Control Theory

Let us consider a generic linear quadratic unconstrained OCP[4]:

$$\begin{gathered} \min_{(y,u)\in Y\times U} J(y,u) = \frac{1}{2}\|\mathcal{Q}y - y_d\|_H^2 + \frac{\alpha}{2}\|u\|_U^2 \\ \text{subjected to } Ay + Cu = f, \end{gathered} \tag{1.3.1}$$

where Y, U, H are Hilbert spaces, $y_d \in H$, $f \in Y^*$. We are assuming that: $A \in \mathcal{L}(Y, Y^*)$ with a bounded inverse $A^{-1} \in \mathcal{L}(Y^*, Y)$, $C \in \mathcal{L}(U, Y^*)$ and $\mathcal{Q} \in \mathcal{L}(Y, H)$. We will refer to $\mathcal{Q}$ as the observation operator. Under this assumptions theorem 1.1 holds, moreover if

[4]The problem has a linear state equation, a quadratic objective functional.

$\alpha > 0$, the solution is unique (see theorem 1.43 of [35]). Hilbert spaces are reflexive, so the equality $Y = Y^{**}$ holds. Furthermore, we suppose $U^* = U$ and $H^* = H$. Let us set $\mathcal{E}(y,u) = Ay + Cu - f$, from which $\mathcal{E}_y(y,u) = A$ and $\mathcal{E}_u(y,u) = C$.
Furthermore:

$$\langle J_y(y,u), s\rangle_{Y^*Y} = (\mathcal{Q}y - y_d, \mathcal{Q}s)_H = \langle \mathcal{Q}^*(\mathcal{Q}y - yd), s\rangle_{Y^*Y},$$

where $\mathcal{Q}^* \in \mathcal{L}(H^*, Y^*)$. It also holds:

$$\langle J_u(y,u), w\rangle_{U^*U} = \alpha(u,w)_U,$$

Let $p \in Q = Y$ be the adjoint variable. Thus the variational formulation of the optimality system reads:

$$\begin{cases} \langle Ay + Cu - f, q\rangle_{Y^*Y} = 0 & \forall q \in Q, \\ \langle A^*p + \mathcal{Q}^*(\mathcal{Q}y - y_d), z\rangle_{Y^*Y} = 0 & \forall z \in Y, \\ \langle \alpha u + C^*p, v\rangle_{U^*U} = 0 & \forall v \in U, \end{cases} \tag{1.3.2}$$

or, equivalently:

$$\begin{cases} Ay + Cu = f, \\ A^*p = -\mathcal{Q}^*(\mathcal{Q}y - y_d), \\ \alpha u + C^*p = 0. \end{cases} \tag{1.3.3}$$

Remark 1.3.1. Another way to reach this same result is trough the Lagrangian functional

$$\mathscr{L}(y,u,p) = \frac{1}{2}(\mathcal{Q}y - y_d, \mathcal{Q}y - y_d)_H + \frac{\alpha}{2}(u,v)_U + \langle p, Ay + Cu - f\rangle_{YY^*}.$$

Thanks to the assumptions made, the optimality system (1.3.2) can be built from the derivative of $\mathscr{L}(\cdot,\cdot,\cdot)$ with respect to the three variables $(y,u,p) \in Y \times U \times Y$:

$$\begin{cases} \langle \mathscr{L}_p(y,u,p), q\rangle_{Y^*Y} = 0, \\ \langle \mathscr{L}_y(y,u,p), z\rangle_{Y^*Y} = 0, \\ \langle \mathscr{L}_u(y,u,p), v\rangle_{U^*U} = 0. \end{cases} \tag{1.3.4}$$

1.3.2 Linear Quadratic Optimal Control Examples

In this section we will present some illustrative examples of distributed control. The first one is governed by a Laplace equation, the second one by an advection-diffusion state equation and the last one is described by a governing Stokes equation.

Example 1.3.2.1 (Distribuited OCP governed by Laplace equation)**.** In this example we will show a distributed linear quadratic OCP governed by the Laplace equation. Let us consider an open, bounded domain $\Omega \subset \mathbb{R}^d$, where $d = 1, 2, 3$. The boundary $\partial\Omega$ is supposed to be sufficiently regular (Lipschitz). The mathematical formalization of the problem reads:

$$\begin{aligned} \min_{(y,u)\in Y\times U} J(y,u) &= \frac{1}{2}\int_\Omega (y - y_d)^2 d\Omega + \frac{\alpha}{2}\int_\Omega u^2 d\Omega, \\ \text{such that } &\begin{cases} -\Delta y = f + u & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases} \end{aligned} \tag{1.3.5}$$

where y_d and f are two given functions of $L^2(\Omega)$, $Y = H_0^1(\Omega)$, $U = L^2(\Omega)$, $Y^* = H^{-1}(\Omega)$ and $H = L^2(\Omega)$. Thanks to the boundary conditions and integration by parts, the weak formulation of the state equation reads: find $y \in Y$ such that

$$a(y, q) = c(u, q) + (f, q)_{L^2}, \qquad \forall q \in Y,$$

where $a(\cdot, \cdot) : Y \times Y \to \mathbb{R}$ and $c(\cdot, \cdot) : U \times Y \to \mathbb{R}$ are bilinear forms defined by:

$$a(y, q) = \int_\Omega \nabla y \cdot \nabla q \, d\Omega,$$
$$c(u, q) = \int_\Omega uq \, d\Omega.$$

For fixed $u \in U$, the existence and uniqueness of the solution of the state equation is guaranteed by the Lax-Milgram lemma[5]. We can refer to the optimal control formulation introduced in (1.3.1) thanks to the operators related to the bilinear forms. The bilinear form $a(\cdot, \cdot)$ induces the operator $A \in (Y, Y^*)$ satisfying $\langle Ay, q\rangle_{Y^*Y} = a(y, q)$, whereas to the bilinear form $c(\cdot, \cdot)$ is associated to the operator $C \in \mathcal{L}(U, Y^*)$ such that $\langle Cu, q\rangle_{Y^*Y} = c(u, q)$. In this particular case the observation operator $\mathcal{Q} \in \mathcal{L}(Y, H)$ is the identity, indeed $\mathcal{Q}y = y$. Under this operational framework, the state equation can be read as $Ay - Cu - f = 0$. Let us compute the dual operators A^*, C^* and Q^*. It is quite simple, since the forms are symmetric. Let us begin with the operator A induced by the bilinear form $a(\cdot, \cdot)$. In general find the adjoint operator of $A \in \mathcal{L}(Y, Y^*)$ means to find $A^* \in \mathcal{L}(Y^{**}, Y^*) = \mathcal{L}(Y, Y^*)$ such that:

$$\langle As, q\rangle_{Y^*Y} = \langle s, A^*q\rangle_{YY^*} \qquad \forall s, q \in Y. \tag{1.3.6}$$

So, integrating by parts, thanks to the divergence theorem and under the assumptions $q = 0$ on $\partial\Omega$ and $s = 0$ on $\partial\Omega$. we reach

$$\begin{aligned}\langle As, q\rangle_{Y^*Y} &= -\int_{\partial\Omega} \nabla sq \cdot \mathbf{n} + \int_\Omega \nabla s \cdot \nabla q \, d\Omega \\ &\underset{q\in Y}{=} \int_\Omega \nabla s \cdot \nabla q \, d\Omega = a(s, q) = a(q, s) = \langle Aq, s\rangle_{Y^*Y}.\end{aligned}$$

For this reason $A = A^*$ and, similarly, $C = C^*$ and $\mathcal{Q} = \mathcal{Q}^*$. So, taken the adjoint variable $p \in Y$, we can build the adjoint equation:

$$\begin{cases} -\Delta p = -(y - y_d) & \text{in } \Omega, \\ p = 0 & \text{on } \partial\Omega, \end{cases}$$

and optimality system as seen in (1.3.3):

$$\begin{cases} Ap = -(y - y_d) \\ \alpha u - Cp = 0 \\ Ay - Cu - f = 0. \end{cases} \tag{1.3.7}$$

[5] We illustrate the lemma without proof. As a reference we propose [60, Chapter 5].

Lemma 1.1. *Let V be and Hilbert space, let $a(\cdot, \cdot) : V \times V \to$ be and $F(\cdot) : V \to \mathbb{R}$ be a continuous coercive bilinear form and a continuous linear functional respectively. So there exists a unique solution for the following problem: find $u \in V$ such that*

$$a(u, v) = F(v) \qquad \forall v \in V.$$

We can express it in the equivalent weak form:

$$\begin{cases} \langle Ay - Cu - f, q\rangle_{Y^*Y} = 0 & \forall q \in Y, \\ \langle Ap, z\rangle_{Y^*Y} = -(y - y_d, z)_{L^2} & \forall z \in Y, \\ (\alpha u, v)_{L^2} = \langle Cp, v\rangle_{U^*U} & \forall v \in U. \end{cases} \tag{1.3.8}$$

or, in the notation of the bilinear forms:

$$\begin{cases} a(y, q) = c(u, q) + (f, q)_{L^2} & \forall q \in Y, \\ a(p, z) = -(y - y_d, z)_{L^2} & \forall z \in Y, \\ (\alpha u, v)_{L^2} = c(p, v) & \forall v \in U. \end{cases} \tag{1.3.9}$$

All these optimality systems can be obtained considering the derivative with respect to (y, u, p) of the Lagrangian functional introduced in remark 1.3.1:

$$\mathscr{L}(y, u, p) = \frac{1}{2}(y - y_d, y - y_d)_{L^2} + \frac{\alpha}{2}(u, u)_{L^2} + a(y, p) - c(u, p) - (f, q)_{L^2}.$$

Now we underline the saddle-point structure of the optimality system (1.3.9). It can be rewritten in the following way:

$$\begin{cases} (y, z)_{L^2} & +a(z, p) & & = & (y_d, z)_{L^2} & \forall z \in Y, \\ & +\alpha(u, v)_{L^2} & -c(v, p) & = & 0 & \forall v \in U, \\ a(y, q) & -c(u, q) & & = & (f, q)_{L^2} & \forall q \in Y. \end{cases}$$

Example 1.3.2.2 (Distribuited advection-diffusion OCP)**.** Let us consider a bit more complex example of distributed linear quadratic OCP. In this case the system is governed by an advection-diffusion equation. The problem is formalized as follows:

$$\min_{(y,u)} J(y, u) = \frac{1}{2}\int_{\Omega_{OBS}} (y - y_d)^2 \, d\Omega_{OBS} + \frac{\alpha}{2}\int_{\Omega_{OBS}} u^2 \, d\Omega_{OBS},$$

$$\text{such that } \begin{cases} -\mathrm{div}(\nu\nabla y) + \boldsymbol{\beta}\cdot\nabla y = f + u & \text{in } \Omega, \\ y = 0 & \text{on } \Gamma_D, \\ \nu\dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N. \end{cases} \tag{1.3.10}$$

Let Ω be an open, bounded and regular domain, with Lipschitz boundary $\partial\Omega$ that verifies $\Gamma_D \cup \Gamma_N = \partial\Omega$ and $\Gamma_D \cap \Gamma_N = \emptyset$. The observation domain $\Omega_{OBS} \subset \Omega$ is open. The control term is $u \in L^2(\Omega)$. The source term $f \in L^2(\Omega)$ is given. The given diffusivity term $\nu = \nu(x) > 0$ in Ω. The diffusivity term is considered in $L^\infty(\Omega)$. Also the advective field $\boldsymbol{\beta} = \boldsymbol{\beta}(x)$ in $L_2(\Omega) \times L_2(\Omega)$ is given. We impose homogeneous Dirichlet boundary conditions on the inlet boundary of the advection field $\Gamma_D = \{x \in \partial\Omega \ : \ \boldsymbol{\beta}\cdot\mathbf{n}(x) < 0\}$, where $\mathbf{n}(x)$ is the unit outward normal vector on $\partial\Omega$, whereas we impose homogeneous Neumann conditions on the outlet boundary of the advection field Γ_N. We consider $Y = H^1_{\Gamma_D}(\Omega) = \{y \in H^1(\Omega) \ : \ y_{|\Gamma_D} = 0\}$, $U = L^2(\Omega)$, $H = L^2(\Omega)$, $Y^* = H^{-1}(\Omega)$ and $y_d \in L^2(\Omega)$ is given. Let us begin our analysis from the weak formulation of the state equation. It reads: find $y \in Y$ such that

$$a(y, q) = c(u, q) + (f, q)_{L^2}, \qquad \forall q \in Y,$$

where the bilinear forms $a(\cdot,\cdot) : Y \times Y \to \mathbb{R}$ and $c(\cdot,\cdot) : U \times Y \to \mathbb{R}$ are defined respectively as

$$a(y,q) = \int_\Omega (\nu \nabla y \cdot \nabla q + \boldsymbol{\beta} \cdot \nabla y q)\, d\Omega,$$
$$c(u,q) = \int_\Omega uq\, d\Omega,$$

thanks to integration by parts and boundary conditions.
For existence an uniqueness of the state equation we refer to [57, Chapter 12].
To compute the adjoint operators it is useful to rewrite $a(\cdot,\cdot) = a_1(\cdot,\cdot) + a_2(\cdot,\cdot)$ where:

$$a_1(s,q) = -\int_{\Gamma_D} \nu \nabla s q \cdot \mathbf{n} - \int_{\Gamma_N} \nu \nabla s q \cdot \mathbf{n} + \int_\Omega \nu \nabla s \cdot \nabla q\, d\Omega$$
$$a_2(s,q) = \int_\Omega \boldsymbol{\beta} \cdot \nabla s q\, d\Omega.$$

Let us indicate by $A_1 \in \mathcal{L}(Y, Y^*)$ and $A_2 \in \mathcal{L}(Y, Y^*)$ the linear operators induced by $a_1(\cdot,\cdot)$ and $a_2(\cdot,\cdot)$ respectively. The linear operator $A \in L(Y, Y^*)$ induced by the bilinear form $a(\cdot,\cdot)$ (notice that this time $a(\cdot,\cdot)$ is not symmetric) is given by the sum of A_1 and A_2. As in example (1.3.2.1) we want to verify:

$$\langle As, y \rangle_{Y^*Y} = \langle s, A^* y \rangle_{YY^*}.$$

Let us act on A_1. Assume that $q = 0$ on $\partial\Omega$. Furthermore, we also know that $\nu \dfrac{\partial s}{\partial n} = 0$ on $\partial\Omega$. This reduces

$$\langle A_1 s, q \rangle_{Y^*Y} = a_1(s,q) = \int_\Omega \nu \nabla s \cdot \nabla q\, d\Omega.$$

Now, integrating by parts and using the divergence theorem, we obtain:

$$\begin{aligned}\langle A_1 s, q \rangle_{Y^*Y} &= \int_{\Gamma_D} s \nabla q \cdot \mathbf{n} + \int_{\Gamma_N} s \nabla q \cdot \mathbf{n} - \int_\Omega s \Delta q\, d\Omega \\ &\underset{s \in H^1_{\Gamma_D}(\Omega)}{=} \int_{\Gamma_N} s \nabla q \cdot \mathbf{n} - \int_\Omega s \Delta q\, d\Omega.\end{aligned}$$

Let us focus on A_2. Integrating by parts and applying the divergence theorem, we reach the following equality:

$$\begin{aligned}\langle A_2 s, q \rangle_{Y^*Y} &= \int_{\Gamma_D} \boldsymbol{\beta} q s \cdot \mathbf{n} + \int_{\Gamma_N} \boldsymbol{\beta} q s \cdot \mathbf{n} - \int_\Omega \operatorname{div}(\boldsymbol{\beta} q) s\, d\Omega \\ &\underset{s \in H^1_{\Gamma_D}(\Omega)}{=} \int_{\Gamma_N} \boldsymbol{\beta} q s \cdot \mathbf{n} - \int_\Omega \operatorname{div}(\boldsymbol{\beta} q) s\, d\Omega.\end{aligned}$$

Finally,

$$\begin{aligned}\langle As, q \rangle_{Y^*Y} &= \langle A_1 s, q \rangle_{Y^*Y} + \langle A_2 s, q \rangle_{Y^*Y} \\ &= \int_{\Gamma_N} s(\nabla q + \boldsymbol{\beta} q) \cdot \mathbf{n} - \int_\Omega s \Delta q\, d\Omega - \int_\Omega \operatorname{div}(\boldsymbol{\beta} q) s\, d\Omega.\end{aligned}$$

To obtain the adjoint equality (1.3.6) we have to assume a new boundary condition for the adjoint problem, that is $\dfrac{\partial q}{\partial n} + \boldsymbol{\beta} \cdot \mathbf{n} q = 0$. So the adjoint equation has the following form:

$$\begin{cases} -\mathrm{div}(\nu\nabla p + \boldsymbol{\beta} p) = -\chi_{OBS}(y - y_d) & \text{in } \Omega, \\ p = 0 & \text{on } \Gamma_D, \\ \dfrac{\partial p}{\partial n} + \boldsymbol{\beta}\cdot\mathbf{n} p = 0 & \text{on } \Gamma_N, \end{cases} \tag{1.3.11}$$

where $p \in Y$ is the adjoint variable and χ_{OBS} the characteristic function of Ω_{OBS}. Let $C \in \mathcal{L}(U, Y^*)$ be the linear operator induced by $c(\cdot,\cdot)$ and $\mathcal{Q} \in \mathcal{L}(Y, H)$ the identical observation operator. In this case we can exploit the symmetric assumption of example (1.3.2.1) and $C = C^*$ and $\mathcal{Q} = \mathcal{Q}^*$. Let us define

$$a^*(s,q) = \int_\Omega (\nu\nabla s\cdot\nabla q) - \int_\Omega \mathrm{div}(\boldsymbol{\beta} q)s\, d\Omega;$$

so the optimality system can reads:

$$\begin{cases} a(y,q) = c(u,q) + (f,q)_{L^2} & \forall q \in Y, \\ a^*(z,p) = -(y - y_d, z)_{L^2(\Omega_{OBS})} & \forall z \in Y, \\ (\alpha u, v)_{L^2(\Omega_{OBS})} = c(v,p) & \forall v \in U. \end{cases} \tag{1.3.12}$$

Example 1.3.2.3 (Distribuited OCP governed by Stokes equation)**.** In this final example we consider distributed OCP with a Stokes state equation. The problem is formulated in the following way:

$$\min_{(\mathbf{v},p,\mathbf{u})} J(\mathbf{v},p,\mathbf{u}) = \frac{1}{2}\int_\Omega |\mathbf{v} - \mathbf{v}_d|^2 d\Omega + \frac{\alpha}{2}\int_\Omega |\mathbf{u}|^2 d\Omega,$$
$$\text{such that } \begin{cases} -\nu\Delta\mathbf{v} + \nabla p = \mathbf{u} & \text{in } \Omega, \\ \mathrm{div}(\mathbf{v}) = 0 & \text{in } \Omega, \\ \mathbf{v} = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.3.13}$$

The domain $\Omega \subset \mathbb{R}^2$ is open, bounded and regular. The given constant ν represents the kinematic viscosity, $\mathbf{v} \in V := H_0^1(\Omega)\times H_0^1(\Omega)$ is the velocity field and

$$p \in P := L_0^2(\Omega) = \left\{ r \in L^2(\Omega) \ : \ \int_\Omega r = 0 \right\}$$

represents the pressure. In this case we consider $V^* = H^{-1}(\Omega)\times H^{-1}(\Omega)$ and $P^* = P$. Now let us consider $Y = V \times P$ as the space of the state variable $y = (\mathbf{v}, p)$ and $U = L^2(\Omega)\times L^2(\Omega)$ is the space of the control variable $\mathbf{u}$. The dual space is $Y^* = V^* \times P^*$. In order to build the weak formulation of the state equation, we define the bilinear forms $a(\cdot,\cdot) : V\times V \to \mathbb{R}$, $b(\cdot,\cdot) : V\times P \to \mathbb{R}$ and $c(\cdot,\cdot) : V\times V \to \mathbb{R}$ as follows:

$$a(\mathbf{v},\boldsymbol{\phi}) = \nu\int_\Omega \nabla\mathbf{v}\cdot\nabla\boldsymbol{\phi}\, d\Omega, \quad b(\mathbf{v},p) = -\int_\Omega \mathrm{div}(\mathbf{v})p\, d\Omega, \quad c(\mathbf{u},\boldsymbol{\phi}) = \int_\Omega \mathbf{u}\cdot\boldsymbol{\phi}\, d\Omega.$$

So the weak formulation for the state equation is: find $y = (\mathbf{v}, p)$ such that

$$\begin{cases} a(\mathbf{v},\boldsymbol{\phi}) + b(\boldsymbol{\phi}, p) = c(\mathbf{u},\boldsymbol{\phi}) & \forall \boldsymbol{\phi} \in V \\ b(\mathbf{v},\xi) = 0 & \forall \xi \in P. \end{cases} \tag{1.3.14}$$

For fixed $\boldsymbol{u} \in U$, the state equation admits a unique solution. Indeed the Stokes problem (1.3.14) can be seen as a saddle-point problem satisfying the assumptions of (Brezzi)

Theorem 1.3 (for the proof see [57, Chapter 16]). In this case the Lagrangian functional derivative is used to derive the optimality system of the control problem. Let us define the Lagrangian functional:

$$\mathscr{L}(\mathbf{v}, p, \mathbf{u}, \mathbf{w}, q) = \frac{1}{2}(\mathbf{v} - \mathbf{v}_d, \mathbf{v} - \mathbf{v}_d)_{L^2} + \frac{\alpha}{2}(\mathbf{u}, \mathbf{u})_{L^2} + a(\mathbf{v}, \mathbf{w}) + b(\mathbf{w}, p) - c(\mathbf{u}, \mathbf{w}) + b(\mathbf{v}, q).$$

To obtain the optimality system we assume the derivatives of $\mathscr{L}(\cdot, \cdot, \cdot)$ with respect to $(\mathbf{v}, p, \mathbf{u}, \mathbf{w}, q) \in V \times P \times U \times V \times P$ must vanish. For this kind of problem the system (1.3.4) reads:

$$\begin{cases} a(\mathbf{v}, \boldsymbol{\phi}) + b(\boldsymbol{\phi}, p) = c(\mathbf{u}, \boldsymbol{\phi}) & \forall \boldsymbol{\phi} \in V, \\ b(\mathbf{v}, \xi) = 0 & \forall \xi \in P, \\ a(\boldsymbol{\psi}, \mathbf{w}) + b(\boldsymbol{\psi}, q) = (\mathbf{v} - \mathbf{v}_d, \boldsymbol{\psi})_{L^2} & \forall \boldsymbol{\psi} \in V, \\ b(\mathbf{w}, \pi) = 0 & \forall \pi \in P, \\ \alpha(\mathbf{u}, \boldsymbol{\tau})_{L^2} = c(\boldsymbol{\tau}, \mathbf{w}) & \forall \boldsymbol{\tau} \in U. \end{cases} \tag{1.3.15}$$

We report below the nested saddle-point structure of this control problem:

$$\begin{cases} (\mathbf{v}, \boldsymbol{\psi})_{L^2} & & & a(\boldsymbol{\psi}, \mathbf{w}) & +b(\boldsymbol{\psi}, q) & = & (\mathbf{v}_d, \boldsymbol{\psi})_{L^2} & \forall \boldsymbol{\psi} \in V, \\ & & & b(\mathbf{w}, \pi) & & = & 0 & \forall \pi \in P, \\ & & +\alpha(\mathbf{u}, \boldsymbol{\tau})_{L^2} & -c(\boldsymbol{\tau}, \mathbf{w}) & & = & 0 & \forall \boldsymbol{\tau} \in U, \\ a(\mathbf{v}, \boldsymbol{\phi}) & +b(\boldsymbol{\phi}, p) & -c(\mathbf{u}, \boldsymbol{\phi}) & & & = & 0 & \forall \boldsymbol{\phi} \in V, \\ b(\mathbf{v}, \xi) & & & & & = & 0 & \forall \xi \in P. \end{cases}$$

Chapter 2

Numerical Approximation and Methods for Optimal Control Problems

In this chapter, our purpose is to illustrate what is a numerical approximation for an optimal control problem (OCP) and what are the methods that can be used to solve the optimality conditions system. While in the previous chapter we performed a very theoretical analysis of control problems, in this chapter we will focus on the numerical features of an OCP. We will translate the notions presented in Chapter 1 under the field of *Computational Fluid Dynamics* (CFD). As we have specified in the previous chapter, Fluid Control is an old field of research that fascinated scientists in many fields and from any time. Now, with the development of new computational technologies, fluid control is studied under this new point of view. For the last 30 years, there has been a large use of computational method to understand fluid behaviour under optimization using sophisticated and very expensive simulations. These kind of simulations need a consistent numerical approximation and a smart resolution method for the optimality system. Numerical methods for fluids are essential in many engineering applications. The need of control simulations and, consequently, their numerical approximations and the related solving methods, arise in hydrodynamics, physiological flow studies, aerodynamics, shape optimization, geophysical sciences and environmental engineering (e.g see [14, 42, 49, 64, 18, 72, 73, 51, 71, 48, 58]). In this work we will follow a classical Galerkin method based on a Finite Element discretization to reach a good approximation of our fluid state and control variables. In order to transform the control problem into a *discrete control problem* there are essentially two discretization techniques, that now we will briefly introduce (see [57, 22, 29]):

1. ***discretize-then-optimize***. In this approach we discretize the state equation and, subsequently, we solve the control. This method is represented by the following pattern

$$\text{MODEL} \rightarrow \text{DISCRETIZATION} \rightarrow \text{CONTROL};$$

2. ***optimize-then-discretize***. Through this second strategy, first the continuous control problem is formalized and then we proceed with the discretization of the equations of the optimality system. The process is described by the next scheme:

$$\text{MODEL} \rightarrow \text{CONTROL} \rightarrow \text{DISCRETIZATION}.$$

Now let us talk about solving methods. It is convenient refer to the abstract OCP formulation:

$$\min_{(y,u)} J(y,u) \qquad \text{subject to } \mathcal{E}(y,u) = 0,$$

where y and u are respectively the state and the control variable constrained to a general state equation $\mathcal{E}(y,u) = 0$. To solve an optimization problem of this kind there are usually two ways:

1. ***iterative method***. It is based on the minimization of the reduced functional $\hat{J}(u) = J(y(u),u)$ where $y(u)$ derives from the solution operator $u \mapsto y(u)$ (i.e. see [35, 70, 57]);

2. ***one-shot method***. Based on the direct resolution of the optimality system (i.e. see [67, 69]).

Let us present what we are going to treat in this Chapter. In section 2.1 we will deeply analyse the discretization techniques usually exploited in control problems. The discretization theory for saddle-point problems is described. Some examples in linear quadratic OCPs are shown. Section 2.2 is about the numerical methods to solve an OCP. The iterative method is shortly treated, while the one shot approach is presented more extensively. Finally, in section 2.3, some numerical example of distributed control are provided. They are solved by the one-shot method.

2.1 Discretization Techniques for Optimal Control Problems

The computational study of control problems is based on simulations and on the interpretation of their results. At the base of simulations there is a *discrete optimal control problem*. A control problem is a very complex issue to treat. One of the main characteristic of an OCP is the interaction of the various components of the system: to discretize a control problem can unexpectedly be a difficult task.
The aim of this section is to introduce the concept of discretization of a control problem. First of all the two approaches *discretize-then-optimize* and *optimize-then-discretize* are clarified (as a reference see [57, 53, 64]). Then we will analyse Galerkin approximation, stabilization and convergence of the saddle-point formulation (see [57, paragraph 16.3.3]) and some applications in linear quadratic OCPs (see [57, 58]).

2.1.1 Two approaches: discretize-then-optimize or optimize-then-discretize

Let us focus our attention on how to appropriate discretize an OCP. Let $y \in Y$ and $u \in U$ be the state and the control variable, respectively. Let us indicate our state equation as $\mathcal{E}(y,u) = 0$. To better fix the ideas, we will describe our control problem in the following way: find $u \in U_{ad} \subset U$ such that

$$J(y(u),u) \leq J(y(v),v) \qquad \forall v \in U_{ad}, \tag{2.1.1}$$

where $J : Y \times U \to \mathbb{R}$ is a prescribed cost functional and the minimization is subjected to $\mathcal{E}(y(u),u) = 0$. There are at least two methods to follow in order to discretize and to solve numerically the problem (2.1.1).

1. **Discretize-then-Optimise** (see [57, 22, 29])
   In this kind of approach, first U_{ad} and the state equation are discretized, obtaining respectively the discrete space $U_{ad,h}$ and the new discrete state equation:

$$\mathcal{E}_h(y_h(u_h), u_h) = 0, \tag{2.1.2}$$

   where h is a parameter that indicates the dimension of the mesh elements. We suppose that for $h \to 0$ the discrete problem converges to the continuous one. In this way, if the induction of $U_{ad,h}$ and if the equation (2.1.2) is correct, we expect to obtain a discrete state $y_h(u_h)$ from all the discrete admissible control u_h. So we can formulate the problem (2.1.1) in a discrete version: find $u \in U_{ad,h}$ such that

$$J(y_h(u_h), u_h) \leq J(y_h(v_h), v_h) \qquad \forall v_h \in U_{ad,h}, \tag{2.1.3}$$

   subject to $\mathcal{E}_h(y_h(u_h), u_h) = 0$.
   This a natural way to discretize and solve a control problem: first the state equation is discretized and then we obtain the discretized control model.

2. **Optimise-then-Discretize** (see [57, 22, 29])
   There is another way to proceed. We can consider the state equation $\mathcal{E}(y, u) = 0$ and the problem (2.1.1) to characterize the optimal state and control variables in terms of the optimality system. As introduced in chapter 1, we know that the continuous optimality system is of the form:

$$\begin{cases} \mathcal{E}(y, u) = 0, \\ \mathcal{E}_y(y, u)^* p = -J_y(y, u), \\ \mathcal{E}_u(y, u)^* p = -J_u(y, u). \end{cases} \tag{2.1.4}$$

   where $p = p(u)$ is the adjoint variable associated to y and u. At this point we can discretize and solve numerically (2.1.4).

The two different approaches do not always lead to the same results: for some problems is preferable the first strategy (usually optimal design problems), for others the second one. Which way to use depends substantially from the specific control problem we are considering: i.e. the first approach is to prefer in *optimal design* problems, while could lead to erroneous results in optimal control problems involving vibrations and waves (i.e see [36, 76, 49, 77]).

2.1.2 Galerkin Approximation, Stability and Convergence for Saddle-Point Problems.

We are going to analyse the Galerkin approximation for the saddle-point structure, that is common to all the linear quadratic OCPs. Let us recall the structure of a saddle-point problem [53, 54]. Let X and Q be two Hilbert spaces, respectively endowed with the norms $||\cdot||_X$ and $||\cdot||_Q$. Let us consider X^* and Q^*, the continuous bilinear forms $\mathcal{A}(\cdot,\cdot) : X \times X \to \mathbb{R}$ and $\mathcal{B}(\cdot,\cdot) : X \times U \to \mathbb{R}$ and the two linear functionals $F \in X^*$ and $G \in Q^*$. The saddle-point problem formulation (as introduced in (1.2.1)) reads: find $(x, p) \in X \times Q$ such that

$$\begin{cases} \mathcal{A}(x, v) + \mathcal{B}(v, p) = \langle F, v \rangle & \forall v \in X, \\ \mathcal{B}(x, q) = \langle G, q \rangle & \forall q \in Q. \end{cases} \tag{2.1.5}$$

We can introduce the Galerkin approximation of the abstract problem (2.1.5). Let $X^{\mathcal{N}}$ and $Q^{\mathcal{N}}$ be two finite dimensional subspaces of the spaces X and Q, respectively. In this work these discrete spaces are considered as Finite Element spaces, but the dissertation has value for a general discrete space.
The Galerkin Finite Element approximation of the problem (2.1.5) is: find $(x^{\mathcal{N}}, p^{\mathcal{N}}) \in X^{\mathcal{N}} \times Q^{\mathcal{N}}$ such that

$$\begin{cases} \mathcal{A}(x^{\mathcal{N}}, v^{\mathcal{N}}) + \mathcal{B}(v^{\mathcal{N}}, p^{\mathcal{N}}) = \langle F, v^{\mathcal{N}} \rangle & \forall v^{\mathcal{N}} \in X^{\mathcal{N}}, \\ \mathcal{B}(x^{\mathcal{N}}, q^{\mathcal{N}}) = \langle G, q^{\mathcal{N}} \rangle & \forall q^{\mathcal{N}} \in Q^{\mathcal{N}}. \end{cases} \tag{2.1.6}$$

Similarly to the continuous case we can define the space:

$$X_0^{\mathcal{N}} = \{w^{\mathcal{N}} \in X^{\mathcal{N}} \ : \ \mathcal{B}(w^{\mathcal{N}}, q^{\mathcal{N}}) = 0, \quad \forall q \in Q^{\mathcal{N}}\}.$$

Even if $Q^{\mathcal{N}} \subset Q$ and $X^{\mathcal{N}} \subset X$, in general $X_0^{\mathcal{N}} \nsubseteq X_0$. Now we want to provide a discrete version of the (Brezzi) theorem (1.3) (for the proof see [57], Cap.16), that gives us a result well-posedness of the problem (2.1.6).

Theorem 2.1 (Brezzi)**.** *Assume that the Hilbert spaces X and Q, the functionals $F \in X^*$ and $G \in Q^*$, and the bilinear forms $\mathcal{A}(\cdot,\cdot) : X \times X \to \mathbb{R}$ and $\mathcal{B}(\cdot,\cdot) : X \times Q \to \mathbb{R}$ are given. Let $X^{\mathcal{N}}$ and $Q^{\mathcal{N}}$ be two finite dimensional subspaces of X and Q respectively. Furthermore, assume that $\mathcal{A}(\cdot,\cdot)$ is continuous on $X^{\mathcal{N}} \times X^{\mathcal{N}}$ and $\mathcal{B}(\cdot,\cdot)$ is continuous on $X^{\mathcal{N}} \times Q^{\mathcal{N}}$. Assume that the bilinear form $\mathcal{A}(\cdot,\cdot)$ is weakly coercive on $X_0^{\mathcal{N}}$, i.e.*

$$\inf_{x^{\mathcal{N}} \in X_0^{\mathcal{N}}} \sup_{w^{\mathcal{N}} \in X_0^{\mathcal{N}}} \frac{\mathcal{A}(x^{\mathcal{N}}, w^{\mathcal{N}})}{||x^{\mathcal{N}}||_X ||w^{\mathcal{N}}||_X} \geq \alpha^{\mathcal{N}} \quad \textit{and} \quad \inf_{w^{\mathcal{N}} \in X_0^{\mathcal{N}}} \sup_{x^{\mathcal{N}} \in X_0^{\mathcal{N}}} \frac{\mathcal{A}(x^{\mathcal{N}}, w^{\mathcal{N}})}{||x^{\mathcal{N}}||_X ||w^{\mathcal{N}}||_X} > 0.$$

Moreover, suppose that $\mathcal{B}(\cdot,\cdot)$ satisfies the discrete inf-sup condition

$$\inf_{q^{\mathcal{N}} \in Q^{\mathcal{N}}} \sup_{w^{\mathcal{N}} \in X^{\mathcal{N}}} \frac{\mathcal{B}(w^{\mathcal{N}}, q^{\mathcal{N}})}{||w^{\mathcal{N}}||_X ||q^{\mathcal{N}}||_Q} \geq \beta^{\mathcal{N}} > 0.$$

Then, for all $h > 0$, the problem (2.1.6) has a unique solution $(x^{\mathcal{N}}, p^{\mathcal{N}})$. Furthermore, the following inequalities hold:

$$||x^{\mathcal{N}}||_X \leq \frac{1}{\alpha^{\mathcal{N}}} \left[||F||_X^* + \frac{\alpha^{\mathcal{N}} + \gamma_{\mathcal{A}}}{\beta^{\mathcal{N}}} ||G||_{Q^*} \right],$$
$$||p^{\mathcal{N}}||_Q \leq \frac{1}{\beta^{\mathcal{N}}} \left[\left(1 + \frac{\gamma_{\mathcal{A}}}{\alpha^{\mathcal{N}}}\right) ||F||_X^* + \frac{\gamma_{\mathcal{A}}(\alpha^{\mathcal{N}} + \gamma_{\mathcal{A}})}{\beta^{\mathcal{N}} + \alpha^{\mathcal{N}}} ||G||_{Q^*} \right].$$

Finally, if $(x, p) \in X \times Q$ denotes the unique solution of the problem (2.1.5), the following error estimate holds:

$$||x - x^{\mathcal{N}}||_X + ||p - p^{\mathcal{N}}||_Q \leq C \left(\inf_{w^{\mathcal{N}} \in X^{\mathcal{N}}} ||x - w^{\mathcal{N}}||_X + \inf_{q^{\mathcal{N}} \in Q^{\mathcal{N}}} ||p - q^{\mathcal{N}}||_Q \right),$$

where $C := C(\alpha^{\mathcal{N}}, \beta^{\mathcal{N}}, \gamma_{\mathcal{A}}, \gamma_{\mathcal{B}})$, so C is independent from $\mathcal{N}$.

Remark 2.1.1. In subsection 1.2.2 we have analysed how a linear quadratic OCP can be read in a saddle-point framework. So the general Brezzi's Theorem 2.1 also holds for the linear quadratic OCPs. This is the reason why the well posedness of a discrete linear quadratic OCP depends on the fulfillment of its assumptions.

Let us focus on the algebraic structure of the system associated to (2.1.6). The dimension of $X^{\mathcal{N}}$ and $Q^{\mathcal{N}}$ are respectively indicated with $\mathcal{N}_X$ and $\mathcal{N}_Q$. Let us define the basis of the finite spaces $X^{\mathcal{N}}$ and $Q^{\mathcal{N}}$ with:

$$\{\varphi_j \in X^{\mathcal{N}}\}_{j=1}^{\mathcal{N}_X} \qquad \{\psi_k \in Q^{\mathcal{N}}\}_{k=1}^{\mathcal{N}_Q}.$$

Now we can rewrite the solution $(x^{\mathcal{N}}, p^{\mathcal{N}}) \in X^{\mathcal{N}} \times Q^{\mathcal{N}}$ as:

$$\left(x^{\mathcal{N}} = \sum_{j=1}^{\mathcal{N}_X} x_j \varphi_j(x),\ p^{\mathcal{N}} = \sum_{k=1}^{\mathcal{N}_Q} p_k \psi_k(x)\right).$$

if the basis functions are chosen as test functions for the problem (2.1.6), one can define $A \in \mathbb{R}^{\mathcal{N}_X \times \mathcal{N}_X}, B \in \mathbb{R}^{\mathcal{N}_Q \times \mathcal{N}_X}, \mathbf{F} \in \mathbb{R}^{\mathcal{N}_X}$ and $\mathbf{G} \in \mathbb{R}^{\mathcal{N}_Q}$ as follows:

$$A_{ij} = \mathcal{A}(\varphi_i, \varphi_j), \quad B_{ml} = \mathcal{B}(\varphi_l, \psi_m), \quad \mathbf{F}_k = \langle F, \varphi_k \rangle, \quad \mathbf{G}_s = \langle G, \psi_s \rangle.$$

From those quantities, we can build the following linear system, with a block structure:

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \mathbf{G} \end{pmatrix}, \tag{2.1.7}$$

where $(\mathbf{x})_i = x_i$ and $(\mathbf{p})_k = p_k$. Thanks to the Galerkin approximation, the OCP problem (1.2.9) has the following algebraic formulation:

$$\text{minimize } \frac{1}{2}\mathbf{x}^T A \mathbf{x} - \mathbf{F}^T \mathbf{x} \qquad \text{subject to } B\mathbf{x} = \mathbf{G}. \tag{2.1.8}$$

2.1.3 Approximation of Linear Quadratic OCPs Governed by Elliptic Coercive State Equation

Our purpose is to study the Galerkin approximation of linear quadratic control governed by elliptic coercive state equations. We want to describe them in the framework proposed in section 1.2.1 and analyse the discretize model illustrated in subsection 2.1.2. As usual we consider Y and U, the state and the control space, respectively. The state and the control variables will be indicated with y and u. In this case $U = U(\omega)$, where $\omega \subset \Omega$ or $\omega \subset \partial\Omega$. Furthermore, we assume that Ω is a open, bounded domain with a regular boundary $\partial\Omega$ such that $\Gamma_N \cap \Gamma_D = \emptyset$, while $\Gamma_N \cup \Gamma_D = \partial\Omega$. Moreover, let $Q = Y$ be the adjoint space and let H be the observation space (we can observe the total domain or the boundary, or a part of them).
Let us define the OCP problem as follows:

$$\begin{aligned} \min_{(y,u)\in Y\times U} J(y,u) &= \frac{1}{2}||\mathcal{Q}y - y_d||_H^2 + \frac{\alpha}{2} n(u,u) \\ a(y,q) &= c(u,q) + \langle G, q\rangle_{Q^*Q} \qquad \forall q \in Q = Y, \end{aligned} \tag{2.1.9}$$

where, $n(\cdot,\cdot) : U \times U \to \mathbb{R}$ is a bilinear form linked to the penalization of the control variable, $\alpha > 0$ is a given constant, $y_d \in H$ and $\mathcal{Q} \in \mathcal{L}(Y, H)$.
In order to recast the problem (2.1.9) in a saddle-point formulation (2.1.5) we have to make the following assumptions.

Assumptions 2.1. Suppose:

1. $a(\cdot,\cdot) : Y \times Y \to \mathbb{R}$ is continuous and strongly coercive, i.e. there exist two constants $\gamma_a > 0$ and $\alpha_a > 0$ such that:

$$|a(s,r)| \leq \gamma_a ||s||_Y ||r||_Y \quad \forall s, r \in Y \quad \text{and} \quad a(s,s) \geq \alpha_a ||s||_Y^2 \quad \forall s \in Y,$$

2. the bilinear form $c(\cdot,\cdot) : U \times Y \to \mathbb{R}$ is symmetric and continuous, i.e. there exists a constant $\gamma_c > 0$ such that:

$$|c(v,q)| \leq \gamma_c ||v||_U ||q||_Y \qquad \forall v \in U \ , \ \forall q \in Y,$$

3. the bilinear form $n(\cdot,\cdot) : U \times U \to \mathbb{R}$ is symmetric, coercive and continuous, i.e. there exist two constants $\gamma_n > 0$ and $\alpha_n > 0$ such that:

$$n(v,w) \leq \gamma_n ||v||_U ||w||_U \quad \forall v, w \in U \quad \text{and} \quad n(v,v) \geq \alpha_n ||v||_U^2 \quad \forall v \in U.$$

Let us define $X = Y \times U$ endowed with the scalar product $(\underline{x}, \underline{w})_X = (y,z)_Y + (u,v)_U$, where $\underline{x} = (y,u) \in X$ and $\underline{w} = (z,v) \in X$. Let $\mathcal{A}(\cdot,\cdot) : X \times X \to \mathbb{R}$ be a bilinear form defined as follows

$$\mathcal{A}(\underline{x}, \underline{w}) = (\mathcal{Q}y, \mathcal{Q}z)_H + \alpha n(u,v).$$

Furthermore, let us consider $\mathcal{B}(\cdot,\cdot) : X \times Y \to \mathbb{R}$ a bilinear form such that

$$\mathcal{B}(\underline{w}, q) := a(z,q) - c(v,q).$$

To recast the OCP problem (2.1.9) we also need the *Riesz isomorphism* $\Lambda_H : H \to H^*$ and the *adjoint operator* of $\mathcal{Q}$ that we will indicate as $\mathcal{Q}^* \in \mathcal{L}(H^*, Y^*)$. Finally, we can define $\underline{F} = (\mathcal{Q}^* \Lambda_H y_d, 0) \in X^*$ and reach this new formulation:

$$\begin{cases} \min\limits_{\underline{x} \in X} \mathcal{J}(\underline{x}) = \dfrac{1}{2}\mathcal{A}(\underline{x}, \underline{x}) - \langle \underline{F}, \underline{x} \rangle \\ \mathcal{B}(\underline{x}, q) = \langle G, q \rangle \qquad\qquad\qquad\qquad \forall q \in Y. \end{cases} \tag{2.1.10}$$

our purpose is to establish an equivalence between the problems (2.1.9),(1.2.9) and the following one:

$$\begin{cases} \mathcal{A}(\underline{x}, \underline{w}) + \mathcal{B}(\underline{w}, p) = \langle \underline{F}, \underline{w} \rangle \quad \forall \underline{w} \in X \\ \mathcal{B}(\underline{x}, q) = \langle G, q \rangle \qquad\qquad\qquad\quad \forall q \in Y. \end{cases} \tag{2.1.11}$$

To do that, we have to prove the hypotheses of theorem 1.4, exploiting assumptions 2.1.

Lemma 2.1. *The bilinear forms $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{B}(\cdot,\cdot)$ verify the hypotheses of Theorem 1.4.*

Proof. The bilinear form $\mathcal{A}(\cdot,\cdot)$ is trivially symmetric and nonnegative.

1. The bilinear form $\mathcal{A}(\cdot,\cdot)$ is continuous on $X \times X$. Indeed:

$$\begin{aligned} |\mathcal{A}(\underline{x}, \underline{w})| &\leq ||\mathcal{Q}y||_H ||\mathcal{Q}z||_H + \alpha ||u||_U ||v||_U \leq ||\mathcal{Q}||^2 ||y||_H ||z||_H + \alpha ||u||_U ||v||_U \\ &\leq (||\mathcal{Q}||^2 + \alpha) ||\underline{x}||_X ||\underline{w}||_X. \end{aligned}$$

2. The bilinear form $\mathcal{B}(\cdot,\cdot)$ is continuous on $X \times Y$:

$$\begin{aligned}|\mathcal{B}(\underline{w},q)| &\leq |a(z,q)| + |c(v,q)| \leq \gamma_a \|z\|_Y \|q\|_Y + \gamma_c \|v\|_U \|q\|_Y \\ &\leq (\gamma_a + \gamma_c)\|w\|_X \|q\|_Y.\end{aligned}$$

3. The bilinear form $\mathcal{A}(\cdot,\cdot)$ is strongly coercive on

$$X_0 = \{\underline{w} \in X \ : \ \mathcal{B}(\underline{w},q) = 0, \ \ \forall q \in Y\}$$

. Let us consider $\underline{w} \in X_0 \Rightarrow \mathcal{B}(\underline{w},q) = 0 \Rightarrow a(z,q) = c(v,q) \ \ \forall q \in Y$. So it holds:

$$\alpha_a \|z\|_Y^2 \leq a(z,z) = c(v,z) \leq \gamma_c \|v\|_U \|z\|_Y \Rightarrow \|v\|_U \geq \frac{\alpha_a}{\gamma_c}\|z\|_Y.$$

This leads to:

$$\begin{aligned}\mathcal{A}(\underline{w},\underline{w}) &= ||\mathcal{Q}z||_H^2 + \alpha\|v\|_U^2 = ||\mathcal{Q}z||_H^2 + \frac{\alpha}{2}\|v\|_U^2 + \frac{\alpha}{2}\|v\|_U^2 \\ &\geq ||\mathcal{Q}z||_H^2 + \frac{\alpha\alpha_a^2}{2\gamma_c^2}\|z\|_Y^2 + \frac{\alpha}{2}\|v\|_U^2 \\ &\geq \alpha_0(\|z\|_Y^2 + \|v\|_U^2) = \alpha_0\|w\|_X^2,\end{aligned}$$

where $\alpha_0 = \dfrac{\alpha}{2}\max\left\{1, \dfrac{\alpha_a^2}{\gamma_c^2}\right\}$.

4. The bilinear form $\mathcal{B}(\cdot,\cdot)$ verifies the inf-sup condition:

$$\begin{aligned}\sup_{\substack{\underline{w}\in X,\\ \underline{w}\neq 0}} \frac{\mathcal{B}(\underline{w},q)}{\|\underline{w}\|_X} &= \sup_{\substack{(z,v)\in Y\times U,\\ (z,v)\neq(0,0)}} \frac{a(z,q) - c(v,q)}{\sqrt{\|z\|_Y^2 + \|v\|_U^2}} \\ &\underset{(z,v)=(q,0)}{\geq} \frac{a(q,q)}{\|q\|_Y} \underset{Y=Q}{\geq} \alpha_a\|q\|_Q > 0. \quad \blacksquare\end{aligned}$$

From assumptions 2.1, Theorem 1.3, Theorem 1.4 and Lemma 2.1, the following Proposition holds.

Proposition 2.1. *The OCP (1.2.9) has a unique solution given by $(\underline{x},p) \in X\times Y$, solution of the saddle-point problem (2.1.11).*

Now we are allowed to analyse the Galerkin approximation of the saddle-point problem (2.1.11). So, we can consider the two finite dimensional subspace $X^{\mathcal{N}} \subset X$ and $Q^{\mathcal{N}} = Y^{\mathcal{N}} \subset Y$. Specifically, $X^{\mathcal{N}} = Y^{\mathcal{N}} \times U^{\mathcal{N}}$, with $Y^{\mathcal{N}} \subset Y$ and $U^{\mathcal{N}} \subset U$. The discrete version of problem (2.1.9) reads: find $(\underline{x}^{\mathcal{N}}, p^{\mathcal{N}}) \in X^{\mathcal{N}} \times Y^{\mathcal{N}}$ such that:

$$\begin{cases}\mathcal{A}(\underline{x}^{\mathcal{N}},\underline{w}^{\mathcal{N}}) + \mathcal{B}(\underline{w}^{\mathcal{N}},p^{\mathcal{N}}) = \langle F,\underline{w}^{\mathcal{N}}\rangle & \forall \underline{w}^{\mathcal{N}} \in X^{\mathcal{N}}, \\ \mathcal{B}(\underline{x}^{\mathcal{N}},q^{\mathcal{N}}) = \langle G,q^{\mathcal{N}}\rangle & \forall q^{\mathcal{N}} \in Q^{\mathcal{N}}.\end{cases} \tag{2.1.12}$$

Our proposal is to verify the hypothesis of Theorem 2.1: it guarantees the well-posedness of the problem (2.1.12).

Lemma 2.2. *Thanks to the assumption that $Q^{\mathcal{N}} = Y^{\mathcal{N}}$, the bilinear forms $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{B}(\cdot,\cdot)$ satisfy the hypothesis of theorem 2.1.*

Proof. The main issue is to prove that assumptions 2.1 are also valid in the discrete version of the OCP. Let us focus our attention on the bilinear form $a(\cdot,\cdot)$: the continuity on $Y^{\mathcal{N}} \times Q^{\mathcal{N}} (= Y^{\mathcal{N}})$ is naturally inherited, while a general $Q^{\mathcal{N}}$ could lead to the loss of the strong coercivity. Thanks to the assumption $Q^{\mathcal{N}} = Y^{\mathcal{N}}$, the strong coercivity on $Y^{\mathcal{N}}$ follows from the strong coercivity on Y. It analogously holds for $c(\cdot,\cdot)$ and $n(\cdot,\cdot)$. The final result follow the same arguments of lemma 2.1. ∎

The last result and Theorem 2.1 prove the next Proposition.

Proposition 2.2. *The saddle-point problem (2.1.12) has a unique solution* $(\underline{x}^{\mathcal{N}}, p^{\mathcal{N}}) \in X^{\mathcal{N}} \times Y^{\mathcal{N}}$.

2.1.4 Approximation of OCPs Governed by Stokes State Equations

The aim of this subsection is to analyse an OCP problem governed by Stokes equation. We will focus on its saddle-point structure and on its Galerkin approximation. We will specifically face this kind of problem

$$\min_{(\mathbf{v},p,\mathbf{u})} J(\mathbf{v},p,\mathbf{u}) = \frac{1}{2}\int_\Omega |\mathbf{v}-\mathbf{v}_d|^2 d\Omega + \frac{\alpha}{2}\int_\Omega |\mathbf{u}|^2 d\Omega,$$
$$\text{such that } \begin{cases} -\nu\Delta\mathbf{v} + \nabla p = \mathbf{u} & \text{in } \Omega, \\ \operatorname{div}(\mathbf{v}) = 0 & \text{in } \Omega, \\ \mathbf{v} = 0 & \text{on } \partial\Omega. \end{cases} \tag{2.1.13}$$

Let Ω be a bounded, open regular domain and let $\partial\Omega$ be its boundary. Let us specify functional spaces and variables involved. We consider $V = H_0^1(\Omega) \times H_0^1(\Omega)$ and $\mathbf{v}, \mathbf{v}_d \in V$ as the velocity variable. The variable $p \in P = L^2(\Omega)$ represent the state pressure. The control space is $U = L^2(\Omega) \times L^2(\Omega)$ and the control variable $\mathbf{u} \in U$. The state variable is $(\mathbf{v}, p) \in Y = V \times P$, whereas the adjoint state variable is indicated with $(\mathbf{w}, q) \in Q = Y$. As in Example 1.3.2.3, we reach weak formulation:

$$\begin{cases} a(\mathbf{v},\boldsymbol{\phi}) + b(\boldsymbol{\phi},p) = (\mathbf{u},\boldsymbol{\phi})_{L^2} & \forall \boldsymbol{\phi} \in V \\ b(\mathbf{v},\xi) = 0 & \forall \xi \in P, \end{cases} \tag{2.1.14}$$

where

$$a(\mathbf{v},\boldsymbol{\phi}) = \nu\int_\Omega \nabla\mathbf{v}\cdot\nabla\boldsymbol{\phi}\, d\Omega, \qquad b(\mathbf{v},p) = -\int_\Omega p\operatorname{div}(\mathbf{v})\, d\Omega.$$

Let us define the usual product space $X = Y \times U$, and let $\mathbf{x} = ((\mathbf{v},p),\mathbf{u})$ and $\boldsymbol{\lambda} = ((\boldsymbol{\psi},\pi),\boldsymbol{\tau})$ elements of X. We consider the following bilinear form $\mathcal{A}(\cdot,\cdot) : X \times X \to \mathbb{R}$:

$$\mathcal{A}(\mathbf{x},\boldsymbol{\lambda}) := \int_\Omega \mathbf{v}\cdot\boldsymbol{\psi}\, d\Omega + \alpha\int_\Omega \mathbf{u}\cdot\boldsymbol{\tau}\, d\Omega. \tag{2.1.15}$$

Now let $(\boldsymbol{\phi},\xi) \in Q = Y$. We define the bilinear form $\mathcal{B}(\cdot,\cdot) : X \times Y \to \mathbb{R}$ as follows:

$$\mathcal{B}(\mathbf{x},(\boldsymbol{\phi},\xi)) := a(\mathbf{v},\boldsymbol{\phi}) + b(\boldsymbol{\phi},p) + b(\mathbf{v},\xi) - c(\mathbf{u},\boldsymbol{\phi}). \tag{2.1.16}$$

Given $\mathbf{F} = ((\mathbf{v}_d,0),\mathbf{0})$ we can reformulate the problem (2.1.13):

$$\begin{cases} \min\limits_{\mathbf{x}\in X} \mathcal{J}(\mathbf{x}) = \dfrac{1}{2}\mathcal{A}(\mathbf{x},\mathbf{x}) - \langle \mathbf{F},\mathbf{x}\rangle, \\ \mathcal{B}(\mathbf{x},(\boldsymbol{\phi},\xi)) = 0 & \forall (\mathbf{w},q) \in Y. \end{cases} \tag{2.1.17}$$

Our purpose is to prove assumption of Theorem 1.4 to reach the equivalence between the OCP (2.1.13) and the saddle-point framework: find $(\mathbf{x}, (\mathbf{w}, q)) \in X \times Y$ such that

$$\begin{cases} \mathcal{A}(\mathbf{x}, \boldsymbol{\lambda}) + \mathcal{B}(\boldsymbol{\lambda}, (\mathbf{w}, q)) = \langle \mathbf{F}, \boldsymbol{\lambda} \rangle & \forall \lambda \in X, \\ \mathcal{B}(\mathbf{x}, (\boldsymbol{\phi}, \xi)) = 0 & \forall (\boldsymbol{\phi}, \xi) \in Y. \end{cases} \tag{2.1.18}$$

Notice that the state Stokes equation is a mixed variational problem and so the system (2.1.14) has the features of a weakly coercive problem. Let us define $A(\cdot, \cdot) : Y \times Y \to \mathbb{R}$ in the following way:

$$\mathbf{A}((\mathbf{v}, p), (\boldsymbol{\phi}, \xi)) = a(\mathbf{v}, \boldsymbol{\phi}) + b(\boldsymbol{\phi}, p) + b(\mathbf{v}, \xi).$$

Since $a(\cdot, \cdot)$ and $b(\cdot, \cdot)$ fulfill the hypotheses of theorem 1.3, the continuity and the weak coercivity of $\mathbf{A}(\cdot, \cdot)$ can be shown (see [5]). We can exploit the Nečas - Babuška theorem to ensure the uniqueness of the solution for the state equation. Let us enunciate the it:

Theorem (Nečas)**.** *Let us consider two Hilbert spaces V and W, let $F(\cdot)$ be a continuous linear functional on W. Let $A(\cdot, \cdot) : V \times W \to \mathbb{R}$ be a bilinear form verifying:*

1. *continuity, i.e. there exists $C_A > 0$ such that*

$$|A(u, w)| \leq C_A \|u\|_V \|w\|_W \qquad \forall u \in V, \forall w \in W,$$

2. *weak coercivity, i.e. there exists a constant $\beta > 0$ such that:*

$$\inf_{v \in V} \sup_{w \in W} \frac{A(v, w)}{\|v\|_V \|w\|_X} \geq \beta \qquad \text{and} \qquad \inf_{w \in W} \sup_{v \in V} \frac{A(v, w)}{\|v\|_V \|w\|_X} > 0.$$

Then, the problem

$$A(u, w) = F(w) \qquad \forall w \in W$$

has a unique solution and it also holds:

$$\|u\|_V \leq \frac{1}{\beta} \|F\|_{W^*}.$$

What is important for our purpose is the weak coercivity of $\mathbf{A}(\cdot, \cdot)$, in particular there exists $\beta_A > 0$ such that:

$$\inf_{(\mathbf{v}, p) \in Y} \sup_{(\boldsymbol{\phi}, \xi) \in Y} \frac{\mathbf{A}((\mathbf{v}, p), (\boldsymbol{\phi}, \xi))}{\|(\mathbf{v}, p)\|_Y \|(\boldsymbol{\phi}, \xi)\|_Y} \geq \beta_A > 0.$$

We have all the ingredients to prove the following statement:

Lemma 2.3. *The bilinear forms $\mathcal{A}(\cdot, \cdot)$ and $\mathcal{B}(\cdot, \cdot)$ verify the hypotheses of Theorem 1.4.*

Proof. Let us consider let $\mathbf{x} = ((\mathbf{v}, p), \mathbf{u}), \boldsymbol{\lambda} = ((\boldsymbol{\psi}, \pi), \boldsymbol{\tau}) \in X$. To reach our proposal, the following scalar products are used:

$$\begin{aligned} &((\mathbf{v}, p), (\boldsymbol{\psi}, \pi))_Y = (\mathbf{v}, \boldsymbol{\psi})_{H^1} + (p, \pi)_{L^2}, \\ &(\mathbf{u}, \boldsymbol{\tau})_U = (\mathbf{u}, \boldsymbol{\tau})_{L^2}, \\ &(\mathbf{x}, \boldsymbol{\lambda}) = ((\mathbf{v}, p), (\boldsymbol{\psi}, \pi))_Y + (\mathbf{u}, \boldsymbol{\tau})_U. \end{aligned}$$

The bilinear form $\mathcal{A}(\cdot, \cdot)$ is trivially symmetric and nonnegative.

1. The bilinear form $\mathcal{A}(\cdot,\cdot)$ is continuous on $X \times X$. Indeed:

$$\begin{aligned}|\mathcal{A}(\mathbf{x},\boldsymbol{\lambda})| &\leq ||\mathbf{v}||_{H^1}||\boldsymbol{\phi}||_{H^1} + \alpha||\mathbf{u}||_{L^2}||\boldsymbol{\tau}||_{L^2} \\ &\leq (1+\alpha)||\mathbf{x}||_X||\boldsymbol{\lambda}||_X.\end{aligned}$$

2. The bilinear form $\mathcal{B}(\cdot,\cdot)$ is continuous on $X \times X$. We can affirm that thanks to the continuity of $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ and exploiting Poincaré inequality, which reads as follows: let $\Omega \subset \mathbb{R}^n$ be a bounded domain, then there exists a constant $C > 0$ such that

$$\|v\|_{L^2(\Omega)} \leq C||\nabla v||_{L^2(\Omega)} \qquad \forall v \in H_0^1(\Omega), \tag{2.1.19}$$

moreover it holds:

$$\|v\|_{L^2(\Omega)} \leq \hat{C}||v||_{H^1(\Omega)}.$$

So, we can reach our goal as follows:

$$\begin{aligned}|\mathcal{B}(\mathbf{x},(\boldsymbol{\phi},\xi))| &= |a(\mathbf{v},\boldsymbol{\phi}) + b(\boldsymbol{\phi},p) + b(\mathbf{v},\xi) - c(\mathbf{u},\boldsymbol{\phi})| \\ &\leq \gamma_a\|v\|_{H^1}\|\boldsymbol{\phi}\|_{H^1} + \gamma_b\|\boldsymbol{\phi}\|_{H^1}\|p\|_{L^2} + \gamma_b\|\mathbf{v}\|_{H^1}\|\xi\|_{L^2} + \|u\|_{L^2}\|\boldsymbol{\phi}\|_{L^2} \\ &\leq \gamma_a\|v\|_{H^1}\|\boldsymbol{\phi}\|_{H^1} + \gamma_b\|\boldsymbol{\phi}\|_{H^1}\|p\|_{L^2} + \gamma_b\|\mathbf{v}\|_{H^1}\|\xi\|_{L^2} + C\|u\|_{L^2}\|\boldsymbol{\phi}\|_{H^1} \\ &\leq (\gamma_a + 2\gamma_b + C)\|\mathbf{x}\|_X\|(\boldsymbol{\phi},\xi)\|_Y.\end{aligned}$$

3. The bilinear form $\mathcal{A}(\cdot,\cdot)$ is strongly coercive on

$$X_0 = \{\mathbf{x} \in X \ : \ \mathcal{B}(\mathbf{x},(\boldsymbol{\phi},\xi)) = 0, \ \ \forall(\boldsymbol{\phi},\xi) \in Y\}.$$

Notice that $\mathbf{x} = ((\mathbf{v},p),\mathbf{u}) \in X_0$ if and only if

$$a(\mathbf{v},\boldsymbol{\phi}) + b(\boldsymbol{\phi},p) + b(\mathbf{v},\xi) = (\mathbf{u},\boldsymbol{\phi})_{L^2} \qquad \forall(\boldsymbol{\phi},\xi) \in Y.$$

Let us consider $(\boldsymbol{\phi},\xi) = (\mathbf{v},\xi)$. To continue in our purpose, we need to refer to Cauchy-Schwarz inequality: let V be an inner product space, then

$$|(v,w)_V| \leq \|v\|_V\|w\|_V \qquad \forall v,w \in V.$$

So, thanks to the Poincaré and the Cauchy-Schwarz inequalities it holds:

$$\nu\|\nabla\mathbf{v}\|_{L^2}^2 + 2b(\mathbf{v},p) = (\mathbf{u},\mathbf{v})_{L^2} \leq C\|\mathbf{u}\|_{L^2}\|\nabla\mathbf{v}\|_{L^2},$$

and we can derive:

$$\|\mathbf{u}\|_{L^2} \geq \frac{\nu}{C}\|\nabla\mathbf{v}\|_{L^2} + \frac{2b(\mathbf{v},p)}{C\|\nabla\mathbf{v}\|_{L^2}}. \tag{2.1.20}$$

Now let us prove the strong coercivity of $\mathcal{A}(\cdot,\cdot)$ on X_0, assuming that $\beta_b > 0$ is the

inf-sup constant of $b(\cdot,\cdot)$:

$$\begin{aligned}
\mathcal{A}(\mathbf{x},\mathbf{x}) &= \|\mathbf{v}\|_{L^2}^2 + \alpha\|\mathbf{u}\|_{L^2}^2 = \|\mathbf{v}\|_{L^2}^2 + \frac{\alpha}{2}\|\mathbf{u}\|_{L^2}^2 + \frac{\alpha}{2}\|\mathbf{u}\|_{L^2}^2 \\
&\geq \|\mathbf{v}\|_{L^2}^2 + \frac{\alpha}{2}\frac{\nu^2}{C^2}\|\nabla\mathbf{v}\|_{L^2}^2 + \frac{\alpha}{2}\frac{4b(\mathbf{v},p)^2}{C^2\|\nabla\mathbf{v}\|_{L^2}^2} + \frac{\alpha}{2}\|\mathbf{u}\|_{L^2}^2 \\
&\geq \underbrace{\min\left\{1, \frac{\alpha}{2}\frac{\nu^2}{C^2}\right\}}_{\hat{c}}\|\mathbf{v}\|_{H^1}^2 + \frac{2\alpha\beta_b^2\|\mathbf{v}\|_{H^1}^2\|p\|_{L^2}^2}{C^2(\|\nabla\mathbf{v}\|_{L^2}^2 + \|\mathbf{v}\|_{L^2}^2)} + \frac{\alpha}{2}\|\mathbf{u}\|_{L^2}^2 \\
&\geq \hat{c}\|\mathbf{v}\|_{H^1}^2 + \frac{2\alpha\beta_b^2\|p\|_{L^2}^2}{C^2} + \frac{\alpha}{2}\|\mathbf{u}\|_{L^2}^2 \\
&\geq \underbrace{\min\left\{c_0, \frac{2\alpha\beta_b^2}{C^2}, \frac{\alpha}{2}\right\}}_{\alpha_0}\left[\|\mathbf{v}\|_{H^1}^2 + \|p\|_{L^2}^2 + \|\mathbf{u}\|_{L^2}^2\right] \\
&= \alpha_0\|\mathbf{x}\|_X^2 \qquad \forall \mathbf{x} \in X_0.
\end{aligned}$$

4. The bilinear form $\mathcal{B}(\cdot,\cdot)$ verifies the inf-sup condition:

$$\begin{aligned}
\sup_{\mathbf{x}\in X,\, \mathbf{x}\neq 0} \frac{\mathcal{B}(\mathbf{x},(\mathbf{w},q))}{\|\mathbf{x}\|_X} &= \sup_{(\mathbf{v},p),\mathbf{u})\in X,\, (\mathbf{v},p),\mathbf{u})\neq 0} \frac{\mathbf{A}((\mathbf{v},p),(\mathbf{w},q)) - (\mathbf{u},\mathbf{w})_{L^2}}{\sqrt{\|(\mathbf{v},p)\|_Y^2 + \|u\|_U^2}} \\
&\underset{\mathbf{u}=\mathbf{0}}{\geq} \sup_{(\mathbf{v},p)\in Y,\, (\mathbf{v},p)\neq 0} \frac{\mathbf{A}((\mathbf{v},p),(\mathbf{w},q))}{\sqrt{\|(\mathbf{v},p)\|_Y^2}} \geq \beta_A\|(\mathbf{w},q)\|_Y. \qquad \blacksquare
\end{aligned}$$

From theroems 1.3 and 1.4 and Lemma 2.3 we can state the following:

Proposition 2.3. *The OCP problem (2.1.13) has a unique solution given by* $(\mathbf{x},(\mathbf{w},q)) \in X \times Y$ *solution of the saddle-point problem (2.1.18).*

Now we are we can consider the Galerkin approximation of the saddle-point problem (2.1.18). So, we can consider the two finite dimensional subspace $X^{\mathcal{N}} \subset X$ and $Q^{\mathcal{N}} = Y^{\mathcal{N}} \subset Y$. Specifically, $X^{\mathcal{N}} = Y^{\mathcal{N}} \times U^{\mathcal{N}}$, with $Y^{\mathcal{N}} \subset Y$ and $U^{\mathcal{N}} \subset U$. The Galerkin approximation of the problem (2.1.18) reads: find $(\mathbf{x}^{\mathcal{N}},(\mathbf{w}^{\mathcal{N}},q^{\mathcal{N}})) \in X^{\mathcal{N}} \times Y^{\mathcal{N}}$ such that

$$\begin{cases}
\mathcal{A}(\mathbf{x}^{\mathcal{N}},\boldsymbol{\lambda}^{\mathcal{N}}) + \mathcal{B}(\boldsymbol{\lambda}^{\mathcal{N}},(\mathbf{w}^{\mathcal{N}},q^{\mathcal{N}})) = \langle \mathbf{F},\boldsymbol{\lambda}^{\mathcal{N}}\rangle & \forall \boldsymbol{\lambda}^{\mathcal{N}} \in X^{\mathcal{N}}, \\
\mathcal{B}(\mathbf{x}^{\mathcal{N}},(\boldsymbol{\phi}^{\mathcal{N}},\xi^{\mathcal{N}})) = 0 & \forall (\boldsymbol{\phi}^{\mathcal{N}},\xi^{\mathcal{N}}) \in Y^{\mathcal{N}}.
\end{cases} \tag{2.1.21}$$

Our proposal is to verify the hypothesis of Theorem 2.1: it guarantees the well-posedness of the problem (2.1.21). A necessary condition is that $Y^{\mathcal{N}} \subset Y$ must be inf-sup stable for the Stokes system (2.1.14), i.e. let $V^{\mathcal{N}} \subset V$ and $P^{\mathcal{N}} \subset P$ be the discrete spaces of velocity and pressure, respectively and $Y^{\mathcal{N}} = V^{\mathcal{N}} \times P^{\mathcal{N}}$, then $V^{\mathcal{N}}$ and $P^{\mathcal{N}}$ has to verify the inf-sup condition:

$$\inf_{\xi^{\mathcal{N}}\in P^{\mathcal{N}}} \sup_{\boldsymbol{\phi}^{\mathcal{N}}\in V^{\mathcal{N}}} \frac{b(\boldsymbol{\phi},\xi)}{\|\boldsymbol{\phi}\|_{H^1}\|\xi\|_{L^2}} \geq \beta_b^{\mathcal{N}} > 0. \tag{2.1.22}$$

Lemma 2.4. *Thanks to the assumption of the inf-sup stability of $Y^{\mathcal{N}}$ and assuming that $Q^{\mathcal{N}} = Y^{\mathcal{N}}$, the bilinear forms $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{B}(\cdot,\cdot)$ satisfy the hypotheses of theorem (2.1).*

Proof. The continuity of the bilinear forms $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ is inherited from the original functional space. Also the coercivity properties of $a(\cdot,\cdot)$ derives directly from the parent spaces. Assuming $Q^{\mathcal{N}} = Y^{\mathcal{N}}$ and the inf-sup stability of $Y^{\mathcal{N}}$ it is possible to state that there exists an $\beta_A^{\mathcal{N}} > 0$ such that:

$$\inf_{(\mathbf{v}^{\mathcal{N}},p^{\mathcal{N}})\in Y^{\mathcal{N}}} \sup_{(\boldsymbol{\phi}^{\mathcal{N}},\xi^{\mathcal{N}})\in Y^{\mathcal{N}}} \frac{\mathbf{A}((\mathbf{v}^{\mathcal{N}},p^{\mathcal{N}}),(\boldsymbol{\phi}^{\mathcal{N}},\xi^{\mathcal{N}}))}{\|(\mathbf{v}^{\mathcal{N}},p^{\mathcal{N}})\|_Y \|(\boldsymbol{\phi}^{\mathcal{N}},\xi^{\mathcal{N}})\|_Y} \geq \beta_A^{\mathcal{N}} > 0.$$

Once considered these features, we can proceed as we did in lemma 2.3.

From Theorem 2.1 and Lemma 2.4 we can derive the following result.

Proposition 2.4. *Assuming that hypothesis of Lemma 2.4 are verified, then there exists a unique solution $(\mathbf{x}^{\mathcal{N}},(\mathbf{w}^{\mathcal{N}},q^{\mathcal{N}}))\in X^{\mathcal{N}}\times Y^{\mathcal{N}}$ for the saddle-point problem (2.1.21).*

2.2 Numerical Resolution: One-Shot Approach

In this section numerical methods to solve OCP are discussed. To better understand what we are going to analyse, let us consider an OCP in abstract form. As usual, the state variable is indicated by y, the control variable by u. The general problem is the following:

$$\min_{(y,u)} J(y,u) \qquad \text{subject to } \mathcal{E}(y,u)=0, \tag{2.2.1}$$

where $\mathcal{E}(y,u)=0$ is a generic state equation. As specified in the introduction of this chapter, there are two ways to numerically solve a OCP: the **iterative method** (i.e. see [35, 70, 57]) or the **one-shot method** (i.e. see [67, 69]). We will focus our attention on the latter one. This method has been used for all the applications treated in this work.
We are going to analyse the application of this method to linear quadratic OCPs, specifically. Two examples are proposed: the first is an OCP governed by the Laplace equation, the second is an OCP with Stokes state equations.

2.2.1 One-Shot Approach for Linear Quadratic OCPs

Suppose that we are facing a Galerkin approximation of the linear quadratic OCP (1.2.9). As specified in Remark 2.1.1, the discrete version of a linear quadratic OCP reads as:

$$\text{minimize } \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \mathbf{F}^T\mathbf{x} \qquad \text{subject to } B\mathbf{x}=\mathbf{G}. \tag{2.2.2}$$

The optimality conditions of problem (2.1.8) is of the form already described in (2.1.7):

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}\begin{pmatrix}\mathbf{x}\\ \mathbf{p}\end{pmatrix} = \begin{pmatrix}\mathbf{F}\\ \mathbf{G}\end{pmatrix}. \tag{2.2.3}$$

Example 2.2.1.1 (OCP governed by the Laplace equation)**.** In this simple example a one-shot approach is applied to an OCP for the Laplace equation on an open, bounded, regular domain Ω. The problem is the same analysed in Example 1.3.2.1. Let $Y = H_0^1(\Omega)$ be the state space, $U = L^2(\Omega)$ the control space. The state and the control variables are $y \in Y$ and $u \in U$, respectively. The weak formulation of this problems has the following form:

$$\begin{aligned} \min_{(y,u)\in Y\times U} J(y,u) &= \frac{1}{2}\int_\Omega (y-y_d)^2\,d\Omega + \frac{\alpha}{2}\int_\Omega u^2\,d\Omega \\ \text{such that } a(y,q) &= (u,q)_{L^2} + (f,q)_{L^2} \qquad \forall q\in Y, \end{aligned} \tag{2.2.4}$$

where $a(s,r) = \displaystyle\int_\Omega \nabla s\cdot\nabla r\,d\Omega$. In this particular case the weak adjoint equation is:

$$a(z,p) = -(y-y_d,z)_{L^2} \qquad \forall z\in Y.$$

Let $\{\mathcal{T}^{\mathcal{N}}\}$ be a triangulation over Ω, that is, we can consider a discrete domain

$$\Omega^{\mathcal{N}} = \text{int}\left(\bigcup_{K\in\mathcal{T}^{\mathcal{N}}} K\right),$$

where K is a triangle of $\mathcal{T}^{\mathcal{N}}$. In a Finite Element approximation $Y^{\mathcal{N}} = Y\cap X^r_{\mathcal{N}}$ and $U^{\mathcal{N}} = U\cap X^r_{\mathcal{N}}$, where

$$X^r_{\mathcal{N}} = \{v^{\mathcal{N}}\in C^0(\overline{\Omega})\ :\ v^{\mathcal{N}}|_K\in\mathbb{P}_r,\ \ \forall K\in\mathcal{T}^{\mathcal{N}}\}.$$

and $\mathbb{P}_r$ represents the space of polynomials of degree at most equal to r. Now let us consider $Y^{\mathcal{N}}\subset Y$ and $U^{\mathcal{N}}\subset U$ as the FE discretization of the state and the control space, respectively. The discretization of the OCP problem (2.2.4) reads:

$$\begin{aligned} \min_{(y^{\mathcal{N}},u^{\mathcal{N}})\in Y^{\mathcal{N}}\times U^{\mathcal{N}}} J(y^{\mathcal{N}},u^{\mathcal{N}}) &= \frac{1}{2}\int_\Omega (y^{\mathcal{N}}-y_d)^2\,d\Omega + \frac{\alpha}{2}\int_\Omega (u^{\mathcal{N}})^2\,d\Omega \\ \text{such that } a(y^{\mathcal{N}},q^{\mathcal{N}}) &= (u^{\mathcal{N}},q^{\mathcal{N}})_{L^2} + (f,q^{\mathcal{N}})_{L^2} \qquad \forall q\in Y^{\mathcal{N}}. \end{aligned} \tag{2.2.5}$$

Let $\mathbf{y}$ and $\mathbf{u}$ be the coefficient of $y^{\mathcal{N}}$ and $u^{\mathcal{N}}$ expressed in terms of the nodal basis for $Y^{\mathcal{N}}$ and $U^{\mathcal{N}}$, respectively. We can analyse the algebraic formulation of the discrete problem (2.2.5):

$$\begin{aligned} \min_{(y^{\mathcal{N}},u^{\mathcal{N}})\in Y^{\mathcal{N}}\times U^{\mathcal{N}}} J(\mathbf{y},\mathbf{u}) &= \frac{1}{2}\mathbf{y}^T M\mathbf{y} - \mathbf{y}^T M\mathbf{y}_d + \frac{\alpha}{2}\mathbf{u}^T M\mathbf{u} + \frac{1}{2}\mathbf{y}_d M\mathbf{y}_d \\ \text{such that } K\mathbf{y} &= M\mathbf{u}+\mathbf{f}. \end{aligned} \tag{2.2.6}$$

where K is the stiffness matrix derived from the bilinear form $a(\cdot,\cdot)$ and M is the mass matrix associated to the functional.
In this framework, it is simple to obtain the discretized adjoint equation:

$$K^T\mathbf{p} = -M(\mathbf{y}-\mathbf{y}_d).$$

We now show the connection between this specific example and the general formulation (2.2.3):

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}\begin{pmatrix} \mathbf{x} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \mathbf{G} \end{pmatrix},$$

where

$$A = \begin{pmatrix} M & 0 \\ 0 & \alpha M \end{pmatrix},\quad B = \begin{pmatrix} K & -M \end{pmatrix},\quad \mathbf{x} = \begin{pmatrix} \mathbf{y} & \mathbf{p} \end{pmatrix}^T,\quad \mathbf{G} = \mathbf{f},\quad \mathbf{F} = \begin{pmatrix} M\mathbf{y}_d & \mathbf{0} \end{pmatrix}^T.$$

Now we are able to build the optimality system exploiting the block structure:

$$\begin{pmatrix} M & 0 & K^T \\ 0 & \alpha M & -M \\ K & -M & 0 \end{pmatrix}\begin{pmatrix} \mathbf{y} \\ \mathbf{u} \\ \mathbf{p} \end{pmatrix} = \begin{pmatrix} M\mathbf{y}_d \\ \mathbf{0} \\ \mathbf{f} \end{pmatrix}.$$

Example 2.2.1.2 (OCP governed by the Stokes equations)**.** This example faces a one-shot approach applied to an OCP governed by Stokes equations on an open, bounded, regular domain Ω. This case recalls the example 1.3.2.3: $v \in V := H_0^1(\Omega) \times H_0^1(\Omega)$ is the velocity field,

$$p \in P := L_0^2(\Omega) = \left\{ r \in L^2(\Omega) \ : \ \int_\Omega r = 0 \right\}$$

represents the pressure and $\mathbf{u} \in U := L^2(\Omega) \times L^2(\Omega)$ is the distributed control variable. The main difference is the presence of the desired pressure $p_d \in P$. The constants $\alpha > 0$ and $\delta > 0$ are penalization terms in the cost functional.

$$\begin{aligned} \min_{(\mathbf{v},p,\mathbf{u})} J(\mathbf{v},p,\mathbf{u}) &= \frac{1}{2}\int_\Omega |\mathbf{v}-\mathbf{v}_d|^2 d\Omega + \frac{\delta}{2}\int_\Omega |p-p_d|^2 d\Omega + \frac{\alpha}{2}\int_\Omega |\mathbf{u}|^2 d\Omega, \\ \text{such that } &\begin{cases} -\nu\Delta\mathbf{v} + \nabla p = \mathbf{u} + \mathbf{f} & \text{in } \Omega, \\ \operatorname{div}(\mathbf{v}) = 0 & \text{in } \Omega, \\ \mathbf{v} = \mathbf{0} & \text{on } \partial\Omega. \end{cases} \end{aligned} \tag{2.2.7}$$

Now we can consider the FE discretization of the spaces: $V^{\mathcal{N}} \subset V$, $P^{\mathcal{N}} \subset P$, $U^{\mathcal{N}} \subset U$. We are going to indicate with $(\boldsymbol{v},\boldsymbol{p},\boldsymbol{u})$ the coefficients of the discrete variables expressed in terms of the nodal basis for $V^{\mathcal{N}}, P^{\mathcal{N}}, U^{\mathcal{N}}$. The discretization of the OCP (2.2.7) is given by:

$$\begin{aligned} \min_{(y^{\mathcal{N}},u^{\mathcal{N}})\in Y^{\mathcal{N}}\times U^{\mathcal{N}}} J(\boldsymbol{v},\boldsymbol{p},\boldsymbol{u}) &= \frac{1}{2}\boldsymbol{v}^T M_v \boldsymbol{v} - \boldsymbol{v}^T M_v \boldsymbol{v}_d + \frac{\delta}{2}\boldsymbol{p}^T M_p \boldsymbol{p} \ - \delta \boldsymbol{p}^T M_p \boldsymbol{p}_d + \\ &\quad + \frac{\alpha}{2}\boldsymbol{u}^T M_v \boldsymbol{u} + \frac{1}{2}\boldsymbol{v}_d^T M_v \boldsymbol{v}_d + \frac{\delta}{2}\boldsymbol{p}_d^T M_p \boldsymbol{p}_d, \\ \text{such that } &\begin{cases} A_S\boldsymbol{v} + B_S^T\boldsymbol{p} = M_{\boldsymbol{v}}\boldsymbol{u} + \boldsymbol{f} \\ B_S\boldsymbol{v} = \mathbf{0}, \end{cases} \end{aligned} \tag{2.2.8}$$

where A_S is the stiffness matrix associated to the Laplace operator, B_S is the matrix representing the divergence operator. Here, M_v and M_p are the mass velocity matrix and the mass pressure matrix, respectively. To build the optimality system, let $\boldsymbol{w}$ and $\boldsymbol{q}$ be the adjoint velocity and the adjoint pressure, respectively. The optimality system reads:

$$\begin{pmatrix} M_v & 0 & 0 & A_S & B_S^T \\ 0 & \delta M_p & 0 & B_S & 0 \\ 0 & 0 & \alpha M_v & -M_v & 0 \\ A_S & B_S^T & -M_v & 0 & 0 \\ B_S & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{v} \\ \boldsymbol{p} \\ \boldsymbol{u} \\ \boldsymbol{w} \\ \boldsymbol{q} \end{pmatrix} = \begin{pmatrix} M_v\boldsymbol{v}_d \\ \delta M_p\boldsymbol{p}_d \\ \mathbf{0} \\ \mathbf{f} \\ 0 \end{pmatrix}. \tag{2.2.9}$$

The system can be expressed through the *aggregated variables* $\mathbf{V}, \mathbf{U}, \mathbf{W}$. $\mathbf{V}$ is the aggregated velocity-pressure variable, $\mathbf{U}$ represents the control variable, whereas $\mathbf{W}$ is the aggregated adjoint variable. The optimality system can be formulated in the following way:

$$\begin{pmatrix} M & 0 & K \\ 0 & \alpha M_v & -E^T \\ K & -E & 0 \end{pmatrix} \begin{pmatrix} \mathbf{V} \\ \mathbf{U} \\ \mathbf{W} \end{pmatrix} = \begin{pmatrix} \mathbf{f}_a \\ \mathbf{0} \\ \mathbf{f}_s \end{pmatrix}, \tag{2.2.10}$$

where

$$M = \begin{pmatrix} M_v & 0 \\ 0 & \delta M_p \end{pmatrix}, \quad K = \begin{pmatrix} A_S & B_S^T \\ B_S & 0 \end{pmatrix}, \quad E = \begin{pmatrix} M_v \\ 0 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} \mathbf{v} \\ \mathbf{p} \end{pmatrix}, \quad \mathbf{f}_a = \begin{pmatrix} M_v\boldsymbol{v}_d \\ \delta M_p\boldsymbol{p}_d \end{pmatrix}. \quad \mathbf{f}_s = \begin{pmatrix} \mathbf{f} \\ 0 \end{pmatrix}.$$

The system (2.2.10) can be rewritten under a saddle-point formulation as:

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} \mathbf{X} \\ \mathbf{W} \end{pmatrix} = \begin{pmatrix} \mathbf{F} \\ \mathbf{G} \end{pmatrix}, \tag{2.2.11}$$

where

$$A = \begin{pmatrix} M & 0 \\ 0 & \alpha M_v \end{pmatrix}, \qquad B = \begin{pmatrix} K & -E \end{pmatrix}, \qquad \mathbf{X} = \begin{pmatrix} \mathbf{V} \\ \mathbf{U} \end{pmatrix}.$$

2.3 Numerical Results

In this section we will illustrate the numerical results associated to the examples presented in subsection 2.2.1. The two experiments deal with an OCP governed by a Laplace equation (test case proposed in [16]) and an OCP with Stokes state equations (a slightly modified test case proposed in [78]). The simulations have been implemented in FEniCS (see [45], for further informations one can refer to *https://fenicsproject.org*), exploiting the one-shot method. For the approximation a FE Galerkin optimize-then-discretize method is used.

2.3.1 OCP Governed by Laplace Equation

We show numerical results for the OCP introduced in Example 2.2.1.1, focusing on a solution tracking. Let us consider $\Omega = (0,1)^2$. The control is distributed over Ω. The problem has the following strong formulation:

$$\begin{aligned} &\min_{(y,u)\in Y\times U} J(y,u) = \frac{1}{2}\int_\Omega (y-y_d)^2\, d\Omega + \frac{\alpha}{2}\int_\Omega u^2\, d\Omega, \\ &\text{such that } \begin{cases} -\Delta y = u + f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \end{cases} \end{aligned} \tag{2.3.1}$$

where $U = L^2(\Omega)$, $Y = H_0^1(\Omega)$ and $f = 0$. In Figure 2.3.1.1 a plot of the desired state $y_d = 10x_1(1-x_1)x_2(1-x_2)$ is given. For the FE discretization the space $\mathbb{P}_1$ is used.

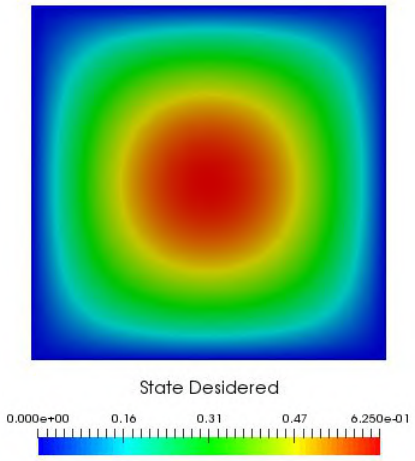

Figure 2.3.1.1: Desired state y_d.

We solved the OCP for $\alpha = 10^{-5}$. The Optimal state and the control variable are presented in figure 2.3.1.2. In this case the objective functional reaches $J = 2.32 \cdot 10^{-4}$. Then, we have repeated the same experiment with $\alpha = 10^{-2}$ as a penalization term. This time, the cost functional is $J = 4.39 \cdot 10^{-2}$. The numerical results for the optimal state and control are reported in figure 2.3.1.3.
Finally, in figure 2.3.1.4 the difference between the optimal state and the desired state is shown, respectively for $\alpha = 10^{-2}$ and $\alpha = 10^{-5}$. In the first case the maximum absolute value reached is $4.9 \cdot 10^{-1}$, while in the second case we have a $5.58 \cdot 10^{-3}$.

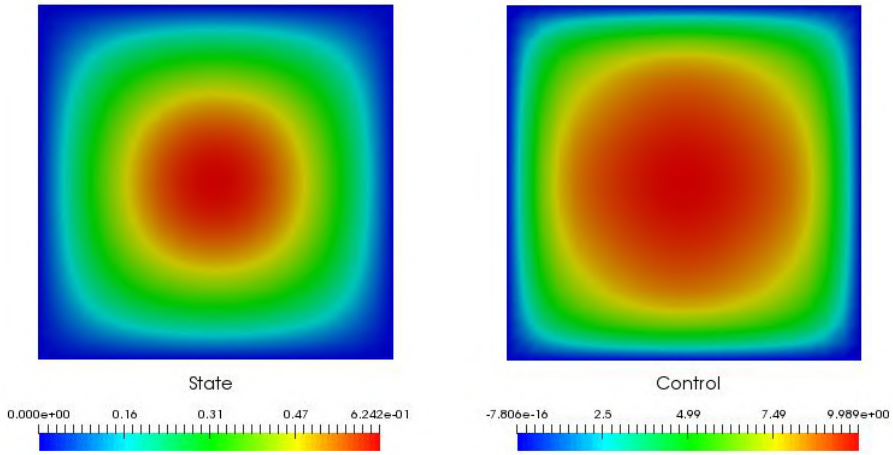

Figure 2.3.1.2: *Left*: optimal state; *right*: control. The penalization term is $\alpha = 10^{-5}$, the functional $J = 2.32 \cdot 10^{-4}$

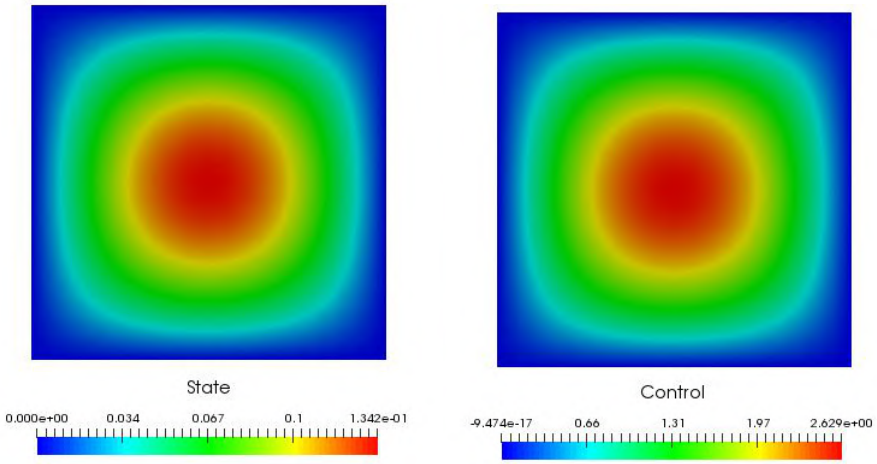

Figure 2.3.1.3: *Left*: optimal state; *right*: control. The penalization term is $\alpha = 10^{-2}$. the functional $J = 4.39 \cdot 10^{-2}$.

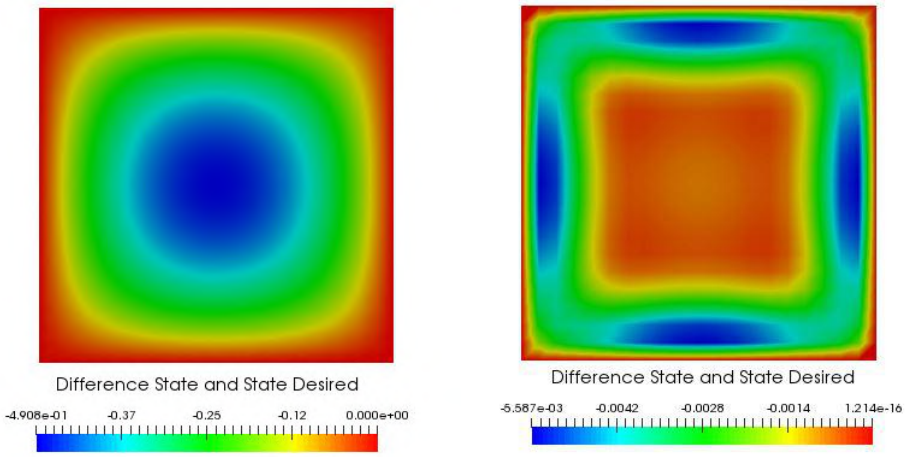

Figure 2.3.1.4: *Left*: difference between optimal state and state desired, $\alpha = 10^{-2}$; *right*: difference between optimal state and state desired, $\alpha = 10^{-5}$.

2.3.2 OCP Governed by Stokes Equations

We show numerical results for the OCP introduced in Example 2.2.1.2, focusing only on a velocity tracking. Let us consider $\Omega = (0,1)^2$. The control is distributed over Ω. The problem has the following strong formulation:

$$\min_{(y,u)\in Y\times U} J(\mathbf{v},p,\mathbf{u}) = \frac{1}{2}\int_\Omega |\mathbf{v}-\mathbf{v}_s|^2\, d\Omega + \frac{\alpha}{2}\int_\Omega |\mathbf{u}|^2\, d\Omega,$$
$$\text{such that } \begin{cases} -\nu\Delta\mathbf{v} + \nabla p = \mathbf{u} & \text{in } \Omega, \\ \operatorname{div}(\mathbf{v}) = 0 & \text{in } \Omega, \\ \mathbf{v} = \mathbf{0} & \text{on } \partial\Omega, \end{cases} \tag{2.3.2}$$

where

$$\mathbf{v}_d = \left(10\frac{\partial}{\partial x_2}(\varphi(x_1)\varphi(x_2)), 10\frac{\partial}{\partial x_1}(\varphi(x_1)\varphi(x_2))\right),$$

with $\varphi : (0,1) \to (0,1)$ is defined as $\varphi(z) = (1-\cos(0.8\pi z))(1-z)^2$. The top left plot of figure 2.3.2.1 shows the desired velocity state of our problem. Let us consider $\alpha = 10^{-4}$ as penalization for velocity. From the discretization we used a Taylor-Hood pair, i.e. continuous piecewise quadratic polynomials for the velocity and continuous piecewise linear polynomials for the pressure. The optimality system formulated in (2.2.9) has the following particular form:

$$\begin{pmatrix} M_v & 0 & 0 & A_S & B_S^T \\ 0 & 0 & 0 & B_S & 0 \\ 0 & 0 & \alpha M_v & -M_v & 0 \\ A_S & B_S^T & -M_v & 0 & 0 \\ B_S & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{v} \\ \boldsymbol{p} \\ \boldsymbol{u} \\ \boldsymbol{w} \\ \boldsymbol{q} \end{pmatrix} = \begin{pmatrix} M_v\boldsymbol{v}_d \\ 0 \\ \mathbf{0} \\ \mathbf{0} \\ 0 \end{pmatrix}.$$

A plot of optimal velocity, optimal pressure and control is given in Figure 2.3.2.1, while in figure 2.3.2.2 the difference between the state and the desired state is shown.

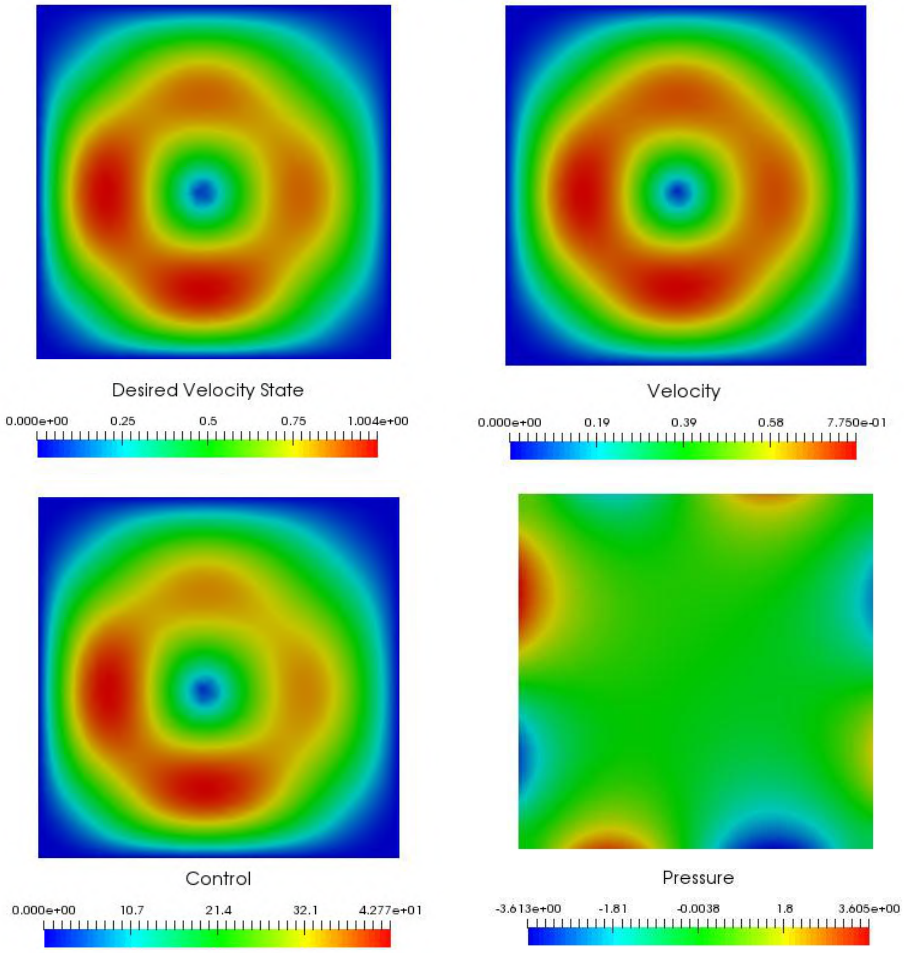

Figure 2.3.2.1: *Top left*: velocity desired; *top right*:optimal velocity; *bottom left*: control; *bottom right*: optimal pressure . The penalization terms are $\alpha = 10^{-4}$ and $\delta = 0$. The functional is $J = 4.02 \cdot 10^{-2}$.

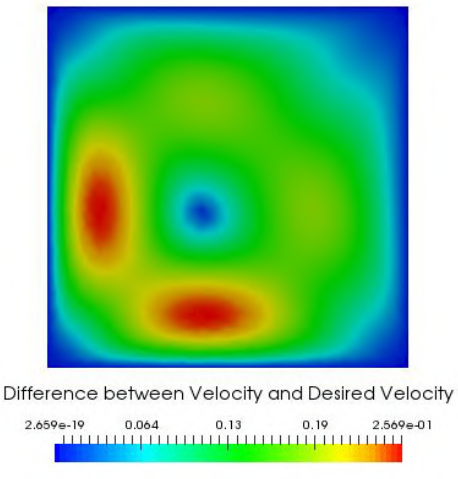

Figure 2.3.2.2: Difference between optimal velocity and desired velocity. The penalization term is $\alpha = 10^{-4}$.

Chapter 3

Reduced Basis Method for Parametrized PDEs

The aim of this chapter is to introduce the reduced basis (RB) approximation for parametrized PDEs. The interest in this efficient resolution method for parametrized PDEs arises in very different contexts (i.e. see [7, 15, 47]). Reduced Order methods act on parametrized problems that have to be evaluated many times. This kind of problems are usually very expensive in terms of computational costs. A huge number of engineering issues depends on *input* parameters: they can represent physical properties or geometrical variables. The traditional discretization techniques could not afford the computational issue of repeated resolution of these very complex problems. The RB approximation aims at reducing the computational cost of parametrized simulations.
To understand how RB methods work, we have to introduce $\mathcal{M}$, the solution manifold, in other words the set of the solutions of the parametrized PDE under the variation of the parameters. RB methods want to approximate every particular solution using only few basis functions, say N, called *reduced basis*. Let $\mathcal{N}$ be the dimension of a classical approximation space. The RB approximation is based on two different stages:

1. ***offline stage***: it is a (potentially) costly phase, where the solution manifold is explored to build a reduced basis capable to describe with a sufficient accuracy any particular solution of $\mathcal{M}$. Computationally one has to solve N problem with $\mathcal{N}$ degree of freedom;

2. ***online stage***: it consists into a Galerkin projection onto the reduced basis space, for a particular parameter value. The computational cost of this phase is independent from $\mathcal{N}$.

For RB methods, our principal theoretical references are [57, 34].
In the first Section we will introduce the abstract formulation of a parametrized PDE and we will specify the important concept of affine decomposition. In Section 3.2 the RB approximation is described, with the description of offline and online stage, respectively. Section 3.3 will be dedicated to the description of Proper Orthogonal Decomposition method (from now on POD). Section 3.4 is dedicated to an introduction to the Empirical Interpolation Method (as a reference see [34, Chapter 5], [2]) to deal with nonaffine parametric dependence. In Section 3.5 an oceanographic application modeled by Quasi-Geostrophic equations in the parametric RB framework is introduced and analysed. For RB methods, our main theoretical references are [57, 34].

3.1 Parametrized PDEs

In this section we will introduce the concept of parametrized PDEs. We will focus on the elliptic case, since in this chapter the RB method dissertation will be developed in this particular framework. We have seen how this kind of problem covers a wide class of scientific phenomena and engineering applications. Let $P(\boldsymbol{\mu})$ be our problem, with $\boldsymbol{\mu} = [\mu_1, \dots, \mu_p] \in \mathcal{P} \subset \mathbb{R}^p$, $p \geq 1$, where $\mathcal{P}$ represents our parameter space. As we have already specified in the introduction to the chapter, $\boldsymbol{\mu}$ can represent physical features of the model or geometrical variables.
The abstract formulation of a weak parametrized PDE problem and its classical approximation is presented. Then the fundamental assumption of affine decomposition is discussed (i.e. see [57, 34]).

3.1.1 Parametrized Weak Formulation

We are going to introduce a framework for a stationary parametrized problem. Let $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$ be regular physical domain. The parameter space will be indicated with $\mathcal{P} \subset \mathbb{R}^p$, whereas V is a suitable Hilbert space. Let us consider the parametrized V-continuous functionals $f : V \times \mathcal{P} \to \mathbb{R}$ and $\ell : V \times \mathcal{P} \to \mathbb{R}$ and the parametrized V-bilinear form $a : V \times V \times \mathcal{P} \to \mathbb{R}$. The parametrized weak formulation of the problem reads: given $\boldsymbol{\mu} \in \mathcal{P}$, find $u(\boldsymbol{\mu}) \in V$ such that:

$$a(u(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v; \boldsymbol{\mu}), \qquad v \in V, \tag{3.1.1}$$

and evaluate the output of interest $s : \mathcal{P} \to \mathbb{R}$

$$s(\boldsymbol{\mu}) = \ell(u(\boldsymbol{\mu}); \boldsymbol{\mu}). \tag{3.1.2}$$

In this chapter we are assuming that the problem is compliant, that is:

1. $\ell(\cdot; \boldsymbol{\mu}) = f(\cdot; \boldsymbol{\mu}), \ \ \forall \boldsymbol{\mu} \in \mathcal{P}$,
2. the bilinear form $a(\cdot, \cdot; \boldsymbol{\mu})$ is symmetric for all $\boldsymbol{\mu} \in \mathcal{P}$.

The Hilbert space V is endowed with a inner product and with a norm $|| \cdot ||_V$:

$$\begin{aligned} &(w, v)_{\boldsymbol{\mu}} = a(w, v; \boldsymbol{\mu}), && \forall w, v, \in V, \\ &\|w\|_{\boldsymbol{\mu}} = \sqrt{(w, w)_{\boldsymbol{\mu}}} = \sqrt{a(w, w; \bar{\boldsymbol{\mu}})}, && \forall w \in V. \end{aligned}$$

The well posedness of the abstract problem (3.1.1) is guaranteed by Lax-Milgram Lemma. We want the problem to be well posed for all the values of the space of parameters. So, in addition to the bilinearity of $a(\cdot, \cdot; \boldsymbol{\mu})$ and and the linearity $f(\cdot, \boldsymbol{\mu})$, we require:

1. $a(\cdot, \cdot, \boldsymbol{\mu})$ is coercive and continuous for every $\boldsymbol{\mu} \in \mathcal{P}$ with respect to the norm $|| \cdot ||_V$, i.e. there exists a positive constant $\alpha(\boldsymbol{\mu}) \geq \alpha_0 > 0$ and $\gamma(\boldsymbol{\mu}) < \infty$ such that:

$$a(v, v; \boldsymbol{\mu}) \geq \alpha(\boldsymbol{\mu}) \|v\|_V^2 \quad \text{and} \quad a(w, v; \boldsymbol{\mu}) \leq \gamma(\boldsymbol{\mu}) \|x\|_V \|v\|_V \quad \forall w, v \in V. \tag{3.1.3}$$

2. $f(\cdot, \boldsymbol{\mu})$ is continuous for all $\boldsymbol{\mu} \in \mathcal{P}$ with respect to the norm $|| \cdot ||_V$, i.e. there exists a constant $\delta(\boldsymbol{\mu}) \leq \delta_0 < \infty$ such that:

$$f(v; \boldsymbol{\mu}) \leq \delta(\boldsymbol{\mu}) \|v\|_V, \qquad \forall v \in V. \tag{3.1.4}$$

Let us specify the coercivity constant and the continuity constant of $a(\cdot,\cdot;\boldsymbol{\mu})$ over V. They are respectively defined as:

$$\alpha(\boldsymbol{\mu}) = \inf_{v\in V} \frac{a(v,v;\boldsymbol{\mu})}{\|v\|_V^2}, \qquad \gamma(\boldsymbol{\mu}) = \sup_{w\in V}\sup_{v\in V} \frac{a(w,v;\boldsymbol{\mu})}{\|w\|_V\|v\|_V}, \qquad \forall \boldsymbol{\mu}\in\mathcal{P}. \tag{3.1.5}$$

Thanks to these hypotheses the problem (3.1.1) admits unique solution.

3.1.2 The Truth Problem

We will present the abstract formulation of the discretized version of the problem (3.1.1). This framework will be indicated as the *truth approximation*. The dissertation has sense for every choice of discrete space for Galerkin method, but in this particular case we will refer to a FE discretization. Let us consider $V^{\mathcal{N}} \subset V$, an $\mathcal{N}$-dimensional approximation space.
The discrete version of the problem (3.1.1) reads: given $\boldsymbol{\mu}\in\mathcal{P}$, find $u^{\mathcal{N}}(\boldsymbol{\mu})\in V^{\mathcal{N}}$ such that

$$a(u^{\mathcal{N}}(\boldsymbol{\mu}),v;\boldsymbol{\mu}) = f(v,\boldsymbol{\mu}), \qquad \forall v\in V^{\mathcal{N}}. \tag{3.1.6}$$

Naturally, the dimension of the solution is $\mathcal{N}$ and the stiffness matrix has dimension $\mathcal{N}\times\mathcal{N}$. The order of operation needed to find $u(\boldsymbol{\mu})$ is $\mathcal{O}(\mathcal{N}^\alpha)$, with $\alpha\geq 1$, so, for great values of $\mathcal{N}$ the resolution process can be computationally costly.
For problem (3.1.6) is possible to specify the coercivity constant and the continuity constant defined respectively as:

$$\gamma^{\mathcal{N}}(\boldsymbol{\mu}) = \sup_{w^{\mathcal{N}}\in V^{\mathcal{N}}}\sup_{v^{\mathcal{N}}\in V^{\mathcal{N}}} \frac{a(w^{\mathcal{N}},v^{\mathcal{N}};\boldsymbol{\mu})}{\|w^{\mathcal{N}}\|_V\|v^{\mathcal{N}}\|_V}, \qquad \forall \boldsymbol{\mu}\in\mathcal{P}, \tag{3.1.7}$$

and

$$\alpha^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{v^{\mathcal{N}}\in V^{\mathcal{N}}} \frac{a(v^{\mathcal{N}},v^{\mathcal{N}};\boldsymbol{\mu})}{\|v^{\mathcal{N}}\|_V^2}, \qquad \forall \boldsymbol{\mu}\in\mathcal{P}. \tag{3.1.8}$$

The discrete problem (3.1.6) is well posed since $\alpha^{\mathcal{N}}(\boldsymbol{\mu})\geq\alpha(\boldsymbol{\mu})>0$ and $\gamma^{\mathcal{N}}(\boldsymbol{\mu})\leq\gamma(\boldsymbol{\mu})$, for all $\boldsymbol{\mu}\in\mathcal{P}$.

3.1.3 Affine Decomposition

To ensure the efficiency of the RB method, one has to verify the so called *affine decomposition*. This assumption as we will see in Section 3.2.2, is essential to guarantee an adequate Offline-Online procedure. We are assuming that the bilinear form $a(\cdot,\cdot;\boldsymbol{\mu})$ and the linear form $f(\cdot;\boldsymbol{\mu})$ are affine in the parameter $\boldsymbol{\mu}$, that is: there exist Q_a and Q_f such that the forms can be rewritten in the following way:

$$a(w,v;\boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \Theta_a^q(\boldsymbol{\mu}) a_q(w,v), \qquad \forall w,v\in V, \qquad \forall\boldsymbol{\mu}\in\mathcal{P}, \tag{3.1.9}$$

$$f(v;\boldsymbol{\mu}) = \sum_{q=1}^{Q_f} \Theta_f^q(\boldsymbol{\mu}) f^q(v) \qquad \forall v\in V, \qquad \forall\boldsymbol{\mu}\in\mathcal{P}, \tag{3.1.10}$$

where the forms

$$a_q : V\times V\to\mathbb{R}, \qquad f_q : V\to\mathbb{R}$$

are independent from the value of the parameter $\boldsymbol{\mu}$, while the coefficients

$$\Theta_a^q : \mathcal{P}\to\mathbb{R}, \qquad \Theta_f^q : \mathcal{P}\to\mathbb{R},$$

are $\boldsymbol{\mu}$-dependent scalar quantities.

3.2 Reduced Basis Method

In this section we will introduce the RB approximation method and its main features. A deeper analysis of Offline-Online decomposition is presented. For this theoretical part we refer to [34, 57]. Finally, the POD algorithm (see [41, 57]) is discussed and is exploited to build the reduce basis functions. The proposed analysis is based on elliptic coercive problems, but the framework can be extended to non coercive PDEs. In this more complex case the well posedness is fulfilled in the general sense of *if-sup* stability. A particolar example is the application to Stokes equations (the reader interested in it, could look at [65, 63, 25]).
Let us assume that a given FE approximation is used for our abstract problem (3.1.1), as previously specified. RB methods compute an approximation of $u^{\mathcal{N}}(\boldsymbol{\mu})$ using the space spanned by chosen solutions of the truth problem (3.1.6). We are now ready to describe the reduction method and all its features and characteristics.

3.2.1 Solution Manifold and Problem Formulation

Let us recall the abstract exact problem (3.1.1): find $u(\boldsymbol{\mu}) \in V$ such that

$$a(u(\boldsymbol{\mu}), v; \boldsymbol{\mu}) = f(v; \boldsymbol{\mu}), \qquad \forall v \in V.$$

We will refer to $u(\boldsymbol{\mu})$ as *exact solution*. In the introduction to this chapter we have shortly introduced the concept of *solution manifold*, that is the set of all the solutions of the parametric problem (3.1.1) varying the parameters. It will be indicated with:

$$\mathcal{M} = \{u(\boldsymbol{\mu}) \mid \boldsymbol{\mu} \in \mathcal{P}\} \subset V.$$

Refering to the new truth formulation, one can define the discrete solution manifold as

$$\mathcal{M}^{\mathcal{N}} = \{u^{\mathcal{N}}(\boldsymbol{\mu}) \mid \boldsymbol{\mu} \in \mathcal{P}\} \subset V^{\mathcal{N}},$$

based on the parametric truth solutions under the variation of the parameter in $\mathcal{P}$. As we said, let us suppose that $V^{\mathcal{N}}$ is a given FE approximation space.
Our goal is to describe the RB spaces construction for this kind of parametric problem[1]. Given a positive integer N_{max}, we can define a succession of hierarchical RB spaces V_N^{RB} for $N = 1, \dots, N_{max}$, that is:

$$V_1^{RB} \subset V_2^{RB} \subset \cdots \subset V_{N_{max}}^{RB} \subset V^{\mathcal{N}}.$$

These assumptions are fundamental property for the (memory) efficiency of the RB approximation. To define V_N^{RB}, given $N \in \{1, \dots, N_{max}\}$ we have to introduce the sample

$$S_N = \{\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_N\}.$$

The elements $\boldsymbol{\mu}_n \in \mathcal{P}$, with $1 \leq n \leq N$ are chosen trough an appropriate algorithm. We can now consider the *snapshots* $u^{\mathcal{N}}(\boldsymbol{\mu}_n) \in V^{\mathcal{N}}$. Now, we are able to build the RB spaces as follows:

$$V_N^{RB} = \text{span } \{u^{\mathcal{N}}(\boldsymbol{\mu}_n),\ 1 \leq n \leq N\}. \tag{3.2.1}$$

Notice that for construction the spaces are hierarchical, and also the samples have a similar nested structure:

$$S_1 = \{\boldsymbol{\mu}_1\} \subset S_2 = \{\boldsymbol{\mu}_1, \boldsymbol{\mu}_2\} \subset \cdots \subset S_{N_{max}}.$$

[1]The option proposed is very common and in literature is known as *Lagrange RB spaces*. There are other approaches based on Taylor space or Hermite spaces (i.e. see [55, 38] respectively).

We are now able to introduce the reduced problem in a Galerkin projection formulation onto the RB spaces. It is a new approximated problem and it reads: given $\boldsymbol{\mu} \in \mathcal{P}$, find $u_N^{RB}(\boldsymbol{\mu}) \in V_N^{RB} \subset V^{\mathcal{N}}$ such that

$$a(u_N^{RB}(\boldsymbol{\mu}), v_N^{RB}; \boldsymbol{\mu}) = f(v_N^{RB}; \boldsymbol{\mu}) \qquad \forall v_N^{RB} \in V_N^{RB}. \tag{3.2.2}$$

From now on we will remove the apex ^{RB} for the sake of notation, so $V_N^{RB} = V_N$ and $u_N^{RB}(\boldsymbol{\mu}) = u_N(\boldsymbol{\mu})$. The next step is to find a well-conditioned basis for V_N. The basis functions are built thanks to a Gram-Schmidt process on the snapshots $u_N(\boldsymbol{\mu}_n)$, $1 \leq n \leq N_{max}$ in the inner product $(\cdot, \cdot)_V$. In this way what we obtain is a set of orthonormalized basis functions $\{\zeta_n\}_{n=1}^{N_{max}}$, that is:

$$(\zeta_n, \zeta_m)_V = \delta_{nm},$$

where $1 \leq n, m \leq N_{max}$ and δ_{nm} is the Kronecker symbol.
Naturally, the basis functions and the snapshots of the truth solutions verify the next property:

$$V_N = \text{span } \{\zeta_1, \cdots, \zeta_N\} = \text{span } \{u^{\mathcal{N}}(\boldsymbol{\mu}_1), \cdots, u^{\mathcal{N}}(\boldsymbol{\mu}_N)\}.$$

We can now express the reduced solution in terms of the reduced basis $\{\zeta_n\}_{n=1}^{N_{max}}$ in the following way:

$$u_N(\mu) = \sum_{j=1}^{N} u_N^j(\boldsymbol{\mu})\zeta_j. \tag{3.2.3}$$

Substituting the latter expression (3.2.3) in the reduced problem (3.2.2) and choosing $v_N = \zeta_i$, $1 \leq i \leq N$, one obtains the following algebraic reduced system

$$\sum_{j=1}^{N} a(\zeta_j, \zeta_i; \boldsymbol{\mu}) u_N^j(\boldsymbol{\mu}) = f(\zeta_i; \boldsymbol{\mu}), \qquad 1 \leq i \leq N, \tag{3.2.4}$$

that has the coefficients $u_N^j(\boldsymbol{\mu})$ as unknowns. The problem (3.2.4) can be expressed in matrix form as

$$A_N(\boldsymbol{\mu})\mathbf{u}_N(\boldsymbol{\mu}) = \mathbf{f}_N(\boldsymbol{\mu}), \tag{3.2.5}$$

where $(\mathbf{u}_N(\boldsymbol{\mu}))_j = u_N^j(\boldsymbol{\mu})$ and $(\mathbf{f}_N)_i = f(\zeta_i; \boldsymbol{\mu})$, whereas the matrix A_N has as entries:

$$(A_N(\boldsymbol{\mu}))_{ij} = a(\zeta_i, \zeta_j; \boldsymbol{\mu}).$$

3.2.2 Offline-Online procedure

The system (3.2.4) usually has low dimension $N \times N$, but its formulation is linked to the FE approximation space in the basis functions $\{\zeta_j\}_{j=1}^{N_{max}}$. If one assembles the RB stiffness matrix $A_N(\boldsymbol{\mu})$ for every value of the parameter $\boldsymbol{\mu}$, the evaluation process will remain very expensive in terms of computational cost. Thanks to the affinity assumption, the formation of the matrix $A_N(\boldsymbol{\mu})$ can be decoupled in two phases: the Offline and the Online stages, that allow to efficiently solve the system (3.2.5) for each new value of the parameter $\boldsymbol{\mu}$. Specifically, using the expressions (3.1.9) and (3.1.10), the system (3.2.4) takes the following form:

$$\sum_{j=1}^{N} \left(\sum_{q=1}^{Q_a} \Theta_a^q(\boldsymbol{\mu}) a_q(\zeta_j, \zeta_i) \right) u_N^j(\boldsymbol{\mu}) = \sum_{q=1}^{Q_f} \Theta_f^q(\boldsymbol{\mu}) f^q(\zeta_i) \qquad 1 \leq i \leq N.$$

The latter problem can be rewritten in matrix form:

$$\left(\sum_{q=1}^{Q_a}\Theta_a^q(\boldsymbol{\mu})A_N^q\right)u_N(\boldsymbol{\mu})=\sum_{q=1}^{Q_f}\Theta_f^q(\boldsymbol{\mu})\mathbf{f}_N^q, \tag{3.2.6}$$

where

$$(A_N^q)_{ij}=a_q(\zeta_i,\zeta_j),\qquad (\mathbf{f}_N^q)_i=f_q(\zeta_i).$$

It is clear that a RB approximation requires a $\boldsymbol{\mu}$-independent costly Offline phase and a very efficient Online phase, $\boldsymbol{\mu}$-dependent. The first procedure is needed only once, while the second process is applied at every new evaluation of a different parameter $\boldsymbol{\mu}\in\mathcal{P}$:

- in the Offline stage, first of all the snapshots $u^{\mathcal{N}}(\boldsymbol{\mu}_n)$ for $1\le n\le N_{max}$ are computed. Then a Gram-Schmidt orthonormalization for $1\le n\le N_{max}$ is applied obtaining the equivalent basis $\{\zeta_i\}_{i=1}^{N_{max}}$. After this preliminary phase, we are ready to assemble and store the following structures:

$$f_q(\zeta_n),\qquad 1\le n\le N_{max},\qquad 1\le q\le Q_f, \tag{3.2.7}$$

  and

$$a_q(\zeta_n,\zeta_m),\qquad 1\le n,m\le N_{max},\quad 1\le q\le Q_a. \tag{3.2.8}$$

  The computational cost depends on $\mathcal{N}, Q_a$ and N_{max}.

- In the Online stage, we exploit the structures computed in the previous step to build

$$\sum_{q=1}^{Q_a}\Theta_a^q(\boldsymbol{\mu})a_q(\zeta_j,\zeta_i)\qquad\text{and}\qquad\sum_{q=1}^{Q_f}\Theta_f^q(\boldsymbol{\mu})f_q(\zeta_i),\qquad 1\le i,j\le N,$$

  and then solve the resulting linear system of dimension $N\times N$ to obtain $u_N^j(\boldsymbol{\mu})$ with $1\le j\le N$. The operation count depends on N, Q_a and Q_f, but it is independent from $\mathcal{N}$. To be more specific we need $O(Q_aN^2)$ to assemble the stiffness matrix, $O(Q_fN)$ to assemble the output vector and finally $O(N^3)$ are the operations needed to solve reduced linear system (3.2.5).

Remark 3.2.1. Let us present what is the relation between the RB quantities and the corresponding FE approximations. Let us introduce $\{\psi_s\}_{s=1}^{\mathcal{N}}$, basis for the FE space $V^{\mathcal{N}}$. Notice that $\zeta_i\in V^{\mathcal{N}}$, so they be expressed in terms of the FE basis, that is

$$\zeta_i=\sum_{s=1}^{\mathcal{N}}\zeta_{is}\psi_s,\qquad 1\le i\le N_{max}.$$

Moreover

$$a_q(\zeta_i,\zeta_j)=\sum_{s=1}^{\mathcal{N}}\sum_{r=1}^{\mathcal{N}}\zeta_{is}a(\psi_s,\psi_r)\zeta_{jr}\qquad\text{and}\qquad\mathbf{f}_N^q=\sum_{r=1}^{\mathcal{N}}\zeta_{jr}f_q(\psi_r). \tag{3.2.9}$$

Let $Z=[\zeta_1\cdots\zeta_N]\in\mathbb{R}^{\mathcal{N}\times N}$ be the *basis matrix*, for $1\le N\le N_{max}$. Then the equations in (3.2.9) can be expressed in a matrix form:

$$A_N^q=Z^TA_{\mathcal{N}}^qZ,\qquad\text{and}\qquad\mathbf{f}_N^q=Z^T\mathbf{f}_{\mathcal{N}}^q,$$

where $(A_{\mathcal{N}}^q)_{ij}=a_q(\psi_i,\psi_j)$ and $(\mathbf{f}_{\mathcal{N}}^q)_i=f_q(\psi_i)$.

3.2.3 Proper Orthogonal Decomposition (POD)

The main issue treated is this subsection is to understand how to generate reduce basis spaces. There are essentially two classical approaches to reach this goal: one is the so called *greedy algorithm* (see [34, Subsection 3.2.2]), the other one is the *proper orthogonal decomposition* (POD). We will focus our analysis on the latter of the two algorithms.
To apply a POD, a discrete and finite-dimensional subset $\mathcal{P}_h \subset \mathcal{P}$ is needed. For this specific set of parameter one can define

$$\mathcal{M}^{\mathcal{N}}(\mathcal{P}_h) = \{u^{\mathcal{N}}(\boldsymbol{\mu}) \mid \boldsymbol{\mu} \in \mathcal{P}_h\}.$$

The cardinality of $\mathcal{M}^{\mathcal{N}}(\mathcal{P}_h)$ is $M = |\mathcal{P}_h|$. Naturally it holds $\mathcal{M}^{\mathcal{N}}(\mathcal{P}_h) \subset \mathcal{M}^{\mathcal{N}}$ since $\mathcal{P}_h \subset \mathcal{P}$. When $\mathcal{P}_h$ is fine enough, $\mathcal{M}^{\mathcal{N}}(\mathcal{P}_h)$ is a good approximation of the discrete manifold $\mathcal{M}^{\mathcal{N}}$. From now on we will refer to $\mathcal{M}^{\mathcal{N}}(\mathcal{P}_h)$ as $V_{\mathcal{M}}$. The algorithm of POD is based on two processes:

1. sampling the parameter space $\mathcal{P}_h$ to compute the truth solutions at the chosen parameters,
2. a compression phase, where one discards the redundant information.

The N-space resulting from the POD algorithm minimize the following quantity:

$$\sqrt{\frac{1}{M}\sum_{\boldsymbol{\mu}\in\mathcal{P}_h} \inf_{v_N \in V_N} \|u^{\mathcal{N}}(\boldsymbol{\mu}) - v_N\|_V^2} \tag{3.2.10}$$

over the N-dimensional reduced spaces V_N of $V_{\mathcal{M}} = \text{span}\,\{u^{\mathcal{N}}(\boldsymbol{\mu}) \mid \boldsymbol{\mu} \in \mathcal{P}_h\}$.
Let us introduce an ordering on the parameters $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_M \in \mathcal{P}_h$. This induce an ordering on the truth solutions $u^{\mathcal{N}}(\boldsymbol{\mu}_1), \dots, u^{\mathcal{N}}(\boldsymbol{\mu}_M)$. To reach our goal of constructing the POD-space, we define the symmetric and linear operator $\mathbf{C} : V_{\mathcal{M}} \to V_{\mathcal{M}}$ as

$$\mathbf{C}(v^{\mathcal{N}}) = \frac{1}{M}\sum_{m=1}^{M} (v^{\mathcal{N}}, u^{\mathcal{N}}(\boldsymbol{\mu}_m))u^{\mathcal{N}}(\boldsymbol{\mu}_m), \qquad v^{\mathcal{N}} \in V_{\mathcal{M}}.$$

Let us consider the eigenvalues $\lambda_n \in \mathbb{R}$ and the corresponding eigenfunctions $\xi_n \in V_{\mathcal{M}}$, with $\|\xi_n\|_V = 1$, linked to the operator $\mathbf{C}$ verifying

$$(\mathbf{C}(\xi_n), u^{\mathcal{N}}(\boldsymbol{\mu}_m)) = \lambda_n(\xi_n, u^{\mathcal{N}}(\boldsymbol{\mu}_m)), \qquad 1 \le m \le M. \tag{3.2.11}$$

Let us assume that the eigenvalues satisfy $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_M$. The orthogonal POD basis functions are given by $\xi_1, \dots, \xi_M$, and they span $V_{\mathcal{M}}$. We can take into consideration the first N functions $\xi_1, \dots, \xi_N$ that satisfy the criterion 3.2.10. They span the space V_{POD}. One can define the projection $P_N : V \to V_{POD}$ as follows:

$$(P_N[f], \xi_n)_V = (f, \xi_n)_V, \qquad 1 \le n \le N,$$

and is given by

$$P_N[f] = \sum_{n=1}^{N} (f, \xi_n)_V \xi_n.$$

Let us apply the projection to all elements of $\mathcal{M}^{\mathcal{N}}(\mathcal{P}_h)$, so it holds:

$$\sqrt{\frac{1}{M}\sum_{m=1}^{M} \|u^{\mathcal{N}}(\boldsymbol{\mu}_m) - P_N[u^{\mathcal{N}}(\boldsymbol{\mu}_m)]\|_V^2} = \sqrt{\sum_{m=N+1}^{M} \lambda_m}.$$

Remark 3.2.2. All what we have introduced in this subsection can be seen under an algebraic point of view. Let us consider $u^{\mathcal{N}}(\boldsymbol{\mu}_m)$ for $m = 1, \dots, M$ and construct the correlation matrix $\mathbf{C} \in \mathbb{R}^{M \times M}$ as

$$\mathbf{C}_{mq} = \frac{1}{M}(u^{\mathcal{N}}(\boldsymbol{\mu}_m), u^{\mathcal{N}}(\boldsymbol{\mu}_q))_V, \qquad 1 \leq m, q \leq M.$$

Then, solve the N-largest eigenvalue-eigenvector (λ_n, v_n) problem:

$$\mathbf{C}v_n = \lambda_n v_n, \qquad 1 \leq n \leq N,$$

with $\|v_n\| = 1$. Giving a descending order to the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$, the orthogonal basis functions $\{\xi_1, \dots, \xi_N\}$ satisfy $V_{POD} = \text{span}\,\{\xi_1, \dots, \xi_N\}$. The basis is given by:

$$\xi_n = \frac{1}{\sqrt{M}} \sum_{m=1}^{M} (v_n)_m u^{\mathcal{N}}(\boldsymbol{\mu}_m), \qquad 1 \leq n \leq N,$$

where $(v_n)_m$ is *m-th* component of the eigenvector $v_n \in \mathbb{R}^M$.

3.3 The Empirical Interpolation Method (EIM)

In this section will analyse the Empirical Interpolation Method (EIM). As said in the previous section, the efficiency of the RB method is strictly linked to the affinity assumption on the forms, that is, $\forall \boldsymbol{\mu} \in \mathcal{P}$:

$$a(w, v; \boldsymbol{\mu}) = \sum_{q=1}^{Q_a} \Theta_a^q(\boldsymbol{\mu}) a_q(w, v), \quad \text{and} \quad f(v; \boldsymbol{\mu}) = \sum_{q=1}^{Q_f} \Theta_f^q(\boldsymbol{\mu}) f_q(v), \tag{3.3.1}$$

In many cases the affine assumptions are not verified and one has to use some numerical techniques to recover them: EIM is an approach to approximate the non-affine structure in a suitable way. The following introduction to EIM algorithm has as references [34, Chapter 5],[2].

3.3.1 EIM Description

Let us suppose to have a function $g(\cdot, \cdot) : X \times \mathcal{P}_{EIM} \to \mathbb{R}$, where X is a Banach space and $\mathcal{P}_{EIM}$ is a parameter space. The EIM procedure aims at approximating this kind of functions. The goal is reached through an interpolation operator I_Q, that interpolates the function of interest at some specific interpolation points $x_1, \dots, x_Q \in \Omega$ as a linear combination of some appropriate basis functions $\{h_1, \dots, h_Q\}$. The peculiarity of these basis functions is that they are part of the set $\{g(\cdot, \boldsymbol{\mu})\}_{\boldsymbol{\mu} \in \mathcal{P}_{EIM}}$. Indeed, they are build as a linear combination of specific snapshots $g_{\boldsymbol{\mu}_1}, \dots, g_{\boldsymbol{\mu}_Q}$, where the Q parameters $\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_Q \in \mathcal{P}_{EIM}$ are chosen through a suitable algorithm.
Assume that $g(\cdot, \boldsymbol{\mu}) \in C^0(\overline{\Omega}) \subset X$. Let $\boldsymbol{\mu} \in \mathcal{P}_{EIM}$ be a chosen parameter, the interpolation operator $I_Q[g(\cdot, \boldsymbol{\mu})]$ applied to $g_{\boldsymbol{\mu}}(\cdot, \boldsymbol{\mu})$ reads as:

$$I_Q[g(\cdot, \boldsymbol{\mu})](x) = \sum_{q=1}^{Q} a_q(\boldsymbol{\mu}) h_q(x), \qquad x \in \Omega. \tag{3.3.2}$$

The interpolation operator has to verify the following equality:

$$I_Q[g(\cdot, \boldsymbol{\mu})](x_j) = g(x_j, \boldsymbol{\mu}), \qquad j = 1, \dots, Q, \tag{3.3.3}$$

and the interpolation is given by the following linear system:

$$\sum_{q=1}^{Q} a_q(\boldsymbol{\mu}) h_q(x_j) = g(x_j, \boldsymbol{\mu}), \qquad j = 1, \dots, Q.$$

Let us express the previous system with $\mathbf{T}\mathrm{a}_{\boldsymbol{\mu}} = g_{\boldsymbol{\mu}}$, where:

$$\mathbf{T}_{ij} = h_j(x_i), \qquad (\mathrm{a}_{\boldsymbol{\mu}})_j = a_j(\boldsymbol{\mu}), \qquad (g_{\boldsymbol{\mu}})_i = g(x_i, \boldsymbol{\mu}), \qquad i, j = 1, \dots, Q. \tag{3.3.4}$$

There are these main issues to be analysed:

1. build the basis functions $\{h_1, \dots, h_Q\}$
2. determine the interpolation points $x_1, \dots, x_Q$
3. prove that the interpolation matrix $\mathbf{T}_{ij} = h_j(x_i)$ is invertible (the interpolation system has unique solution).

The basis functions and the interpolation points are given by greedy algorithm. The algorithm chooses $g_{\boldsymbol{\mu}}$ that is the worst well approximated by the current interpolation operator. Analogously, the interpolation point is chosen as the one that maximize the corresponding error function (if the reader is interested in a deeper description of the algorithmic procedure, see [34, page 53]). As already said, the basis functions are linear combination of some specific $g_{\boldsymbol{\mu}_1}, \dots, g_{\boldsymbol{\mu}_Q}$ that are able to approximate quite well every $g_{\boldsymbol{\mu}}$. The basis functions and the $\{h_1, \dots, h_Q\}$ and the snapshots $\{g_{\boldsymbol{\mu}_1}, \dots, g_{\boldsymbol{\mu}_Q}\}$ span the same space:

$$V_Q = \text{span } \{h_1, \dots, h_Q\} = \text{span } \{g_{\boldsymbol{\mu}_1}, \dots, g_{\boldsymbol{\mu}_Q}\}.$$

Even if they both generate the space V_Q, it is preferable use $\{h_1, \dots, h_Q\}$ as basis functions since the following properties hold:

$$\mathbf{T}_{ii} = h_i(x_i) = 1, \quad 1 \le i \le Q \quad \text{and} \quad \mathbf{T}_{ij} = h_j(x_i) = 0, \quad 1 \le i < j \le Q.$$

Thanks to this choice the basis functions and the interpolation points satisfy (as specified in [2]):

1. $\{h_1, \dots, h_Q\}$ are linearly independent,
2. matrix $\mathbf{T}$ is invertible (lower triangular with unity diagonal),
3. EIM is well posed in X until convergence is not reached.

Moreover, the interpolation operator I_Q is the identity restricted to the space V_Q. Indeed it holds:

$$I_Q[g(x, \boldsymbol{\mu}_i)] = g(x, \boldsymbol{\mu}_i) \qquad i = 1, \dots, Q, \qquad \forall x \in \Omega,$$

and

$$I_Q[g(x_i, \boldsymbol{\mu})] = g(x_i, \boldsymbol{\mu}) \qquad i = 1, \dots, Q, \qquad \forall \boldsymbol{\mu} \in \mathcal{P}_{EIM}.$$

3.3.2 EIM and RB

EIM can be a very useful instrument in reduced context: it allows to apply the RB methods to a wider class of problems, that do not verify the affine assumption 3.3.1. Our aim is to underline how EIM can be used to face this kind of issue. Let us assume that the bilinear form of our problem is:

$$a(w, v; \boldsymbol{\mu}) = \int_\Omega g(x; \boldsymbol{\mu}) b(w, v; x) \, d\Omega,$$

where $b(w, v; x)$ is bilinear in w in v for any $x \in \Omega$ and the function g depends non-trivially on $\boldsymbol{\mu}$, i.e. there is not an affine decomposition of the form:

$$g(x; \boldsymbol{\mu}) = \sum_{q=1}^{Q_a} a_q(\boldsymbol{\mu}) h_q(x), \qquad \boldsymbol{\mu} \in \mathcal{P}, x \in \Omega.$$

In this case one can apply EIM to approximate $g(x; \boldsymbol{\mu})$ and to recover an affine decomposition formulation:

$$g(x; \boldsymbol{\mu}) \approx \sum_{q=1}^{Q_a} a_q(\boldsymbol{\mu}) h_q(x), \qquad \boldsymbol{\mu} \in \mathcal{P}, x \in \Omega.$$

Thanks to this approximation

$$a(w, v; \boldsymbol{\mu}) \approx \sum_{q=1}^{Q_a} a_q(\boldsymbol{\mu}) \int_\Omega h_q(x) b(w, v; x) \, d\Omega.$$

This technique can be used to obtain an affine decomposition for the right hand side $f(v; \boldsymbol{\mu})$, in a similar way.

Remark 3.3.1. Let us analyse the EIM algorithm under an algebraic point of view. Assume that a discrete representation of the domain Ω and of the parameter space $\mathcal{P}_{EIM}$ is given by $\Omega_M = \{x_1, \dots, x_M\}$ and $\mathcal{P}_{EIM}^N = \{\boldsymbol{\mu}_1, \dots, \boldsymbol{\mu}_N\}$, respectively. Consider the following matrix representing the function g at the variation of the parameters:

$$\mathbf{G}_{ij} = g(x_i, \boldsymbol{\mu}_j), \qquad 1 \le i \le M, \qquad 1 \le j \le N.$$

Suppose that a set of basis vectors $\mathbf{H}_Q = [h_1, \dots, h_Q]$ and interpolation indices $i_1, \dots, i_Q$. The discrete version of the interpolation operator $I_Q : \mathbb{R}^Q \to \mathbb{R}^M$ applied to $g \in \mathbb{R}^Q$ is given by

$$I_Q[g] = \mathbf{H}_Q \mathrm{a}_g,$$

where a_g is such that $\mathbf{T}\mathrm{a}_g = g$. The interpolation matrix $\mathbf{T}$ is defined as:

$$\mathbf{T}_{lm} = (H_Q)_{i_l m}, \qquad l, m = 1, \dots, Q.$$

In the following we will present the EIM algorithm.
Let us consider $q = 1$, and a given tolerance $\mathbf{t}$. While $\mathbf{err} < \mathbf{t}$ do:

1. First of all, pick the index

$$j_q = \arg\max_{j=1,\dots,M} \|\mathbf{G}_{:j} - I_{q-1}[\mathbf{G}_{:j}]\|_{\ell^p},$$

where $\mathbf{G}_{:j}$ represent the $j-$th column of the matrix $\mathbf{G}$. Then we consider the interpolation index:

$$i_q = \arg\max_{i=1,\dots,N} |\mathbf{G}_{ij_q} - (I_{q-1}[\mathbf{G}_{:j_q}])_i|.$$

2. Define the basis vector by:

$$h_q = \frac{\mathbf{G}_{:j_q} - I_{q-1}[\mathbf{G}_{:j_q}]}{\mathbf{G}_{i_q j_q} - (I_{q-1}[\mathbf{G}_{:j_q}])_{i_q}}.$$

3. Finally compute the error in the following way:

$$\mathbf{err} = \max_{j=1,\dots,M} \|\mathbf{G}_{:j} - I_{q-1}[\mathbf{G}_{:j}]\|_{\ell^P}$$

and set $q := q + 1$.

In some cases one wants to obtain continuous approximation of $g(x, \boldsymbol{\mu})$ for all $(x, \boldsymbol{\mu}) \in \Omega \times \mathcal{P}_{EIM}$. The basis functions h_q are found as in 2., but all has to be interpret in a continuous context. Let us define the following quantities:

$$\mathbf{S}_{:q} = \mathrm{a}_{(\mathbf{G}_{:j_q})}, \qquad \text{from} \qquad I_{q-1}[\mathbf{G}_{:j_q}] = \mathbf{H}_{q-1}\mathrm{a}_{(\mathbf{G}_{:j_q})},$$

and

$$\mathbf{S}_{qq} = \mathbf{G}_{i_q j_q} - (I_{q-1}[\mathbf{G}_{:j_q}])_{i_q}.$$

then, the continuous basis functions can be constructed thanks the recursive formula:

$$h_q(x) = \frac{g(x, \boldsymbol{\mu}_{i_q}) - \sum_{j=1}^{q-1} \mathbf{S}_{:j_q} h_j(x)}{\mathbf{S}_{qq}}.$$

3.4 EIM-POD Galerkin Based Reduction Method Applied to Quasi-Geostrophic Equations

This section introduces a POD-Galerkin reduced order method approach applied to a climatological problem describing large scale Ocean-wind circulation. We solve a stationary quasi-geostrophic linear model[2] for constant density flows under the influence of Earth rotation. A two-dimensional square domain $\Omega = [0, 1] \times [0, 1]$ is considered. It represents a large portion of Ocean surface. The parametrized equations have the following strong formulation [3]:

$$\begin{cases} q = \Delta\psi & \text{in } \Omega, \\ \dfrac{\partial \psi}{\partial x} + \mu_1 q - \mu_2 \Delta q = -\sin(\mu_3 y + \mu_4) & \text{in } \Omega, \\ q = 0 & \text{on } \partial\Omega, \\ \psi = 0 & \text{on } \partial\Omega. \end{cases}$$

The parameter vector $\boldsymbol{\mu} = [\mu_1, \mu_2, \mu_3, \mu_4]$ is in the space $\mathcal{P} = [10^{-4}, 1]^4$. The components μ_1 and μ_2 are dissipative coefficients, while μ_3 and μ_4 change the forcing term, representing wind action on the Ocean surface. Let us define $V = H_0^1(\Omega)$. For the sake of notation, our parameter $\boldsymbol{\mu}$, will be indicated simply with μ.

[2] The non-linear version will be considered in future. The problem will be deeply analysed in the last chapter.

[3] It is the so called *Stream function* formulation, from which one can derive the currents velocity components (u, v) thanks to the following relations: $u = -\frac{\partial \psi}{\partial y}$, $v = \frac{\partial \psi}{\partial x}$.

The weak problem formulation aims at finding, for some μ, solutions $\psi(\mu)$ and $q(\mu)$ in V such that:

$$\begin{cases} \int_\Omega q(\mu)\varphi + \int_\Omega \nabla\psi(\mu)\cdot\nabla\varphi = 0 & \forall\varphi, p \in V, \\ \int_\Omega \dfrac{\partial\psi(\mu)}{\partial x}p + \mu_1\int_\Omega q(\mu)p + \mu_2\int_\Omega \nabla q(\mu)\cdot\nabla p = -\sin(\mu_3 y + \mu_4) & \forall\varphi, p \in V. \end{cases}$$

We endow V with the inner product:

$$(v, w)_V = \int_\Omega \nabla v \cdot \nabla w \qquad \forall v, w \in V.$$

The parametrization on the forcing term leads to a nonaffine formulation. The strategy to reduce the problem is a EIM-POD Galerkin algorithm (as a reference see [34, Chapter 3] and [34, Chapter 5]).

3.4.1 EIM-POD Galerkin Algorithm

As we learnt through this chapter, the efficiency of a reduced basis approach is closely linked to the affine assumption. When it is not verified, it is essential to lead the problem to an approximate affine formulation (thanks to EIM algorithm). Now we have a new problem to reduce via POD Galerkin algorithm. This is what we call EIM-POD Galerkin. Let us summarily analyse EIM action on a μ dependent function $f : \Omega \times \mathcal{P} \to \mathbb{R}$, like in our case the forcing term $-\sin(\mu_3 y + \mu_4)$. We know that EIM procedure interpolates $f := f(x, \mu)$ at $\{x_1, \ldots, x_Q\} \in \Omega$ on basis functions $\{h_1, \ldots, h_Q\}$ that are linear combination of f evaluated in Q specific parameters.
Let us define the interpolant

$$I_Q[f](x) = \sum_{q=1}^{Q} a_q(\mu)h_q(x).$$

After EIM approximation f is replaced by $I_Q[f](x)$. This kind of procedure restores the affine assumption and so one can treat the new problem with a POD Galerkin reduced basis method. Our specific system is nonaffine and gives a two-component solution. This issue could be faced in two ways:

1. perform single EIM-POD Galerkin for all the solution components (monolithic),
2. perform a EIM-POD Galerkin for each solution component (partitioned).

In the next section we will compare the two choises in terms of: error between reduced and finite element solution and POD eigenvalues decay.

3.4.2 Numerical Results

In this section we show some numerical results in order to provide reasons to split EIM-POD algorithm for the variables q and ψ. In all experiments we used:

- 20 reduced basis functions,
- 21 EIM basis,

- training set and test set of 100 points,
- log-uniform distribution for physical parameters μ_1 and μ_2,
- uniform distribution[4] for forcing parameters μ_3 and μ_4,
- a logarithmic y-axes scale for all the plots.

μ_3 and μ_4 Fixed

In the first case considered the two parameters acting on the forcing term are fixed, while μ_1 and μ_2 can vary in $[10^{-4}, 1]$. We take $\mu = [\mu_1, \mu_2, \pi, 0]$ as a parameter vector. This particular formulation guarantees the affine assumption, so we do not need EIM algorithm. In Figure 3.4.2.1 we present monolithic error[5], the partitioned error[6] and eigenvalues decay normalized with respect to the component largest eigenvalue. Even if both methods have a good final result, one can notice that the partitioned error decays more rapidly than the other one and that POD eigenvalues have the same trend, especially for ψ. Splitting the POD method could be an advantage for affine problem version.

μ_1 and μ_2 Fixed

Let us focus on a different kind of problem for which physical parameters are fixed, while forcing parameters can change in the usual range. Our vector parameter is $\mu = [0, 0.07^3, \mu_3, \mu_4]$. The experiment can be useful to understand how the two methods react to the loss of affine assumption. Here an EIM algorithm is fundamental. In Figure 3.4.2.2 we can visualize how the errors grow if we take a number of basis function greater than nine. It is the sign that EIM-POD, at that point, adds noise to the system (the eigenvalues are practically zero from basis nine on). As in previous experiment, splitting is a better choice if one uses few reduced basis functions.

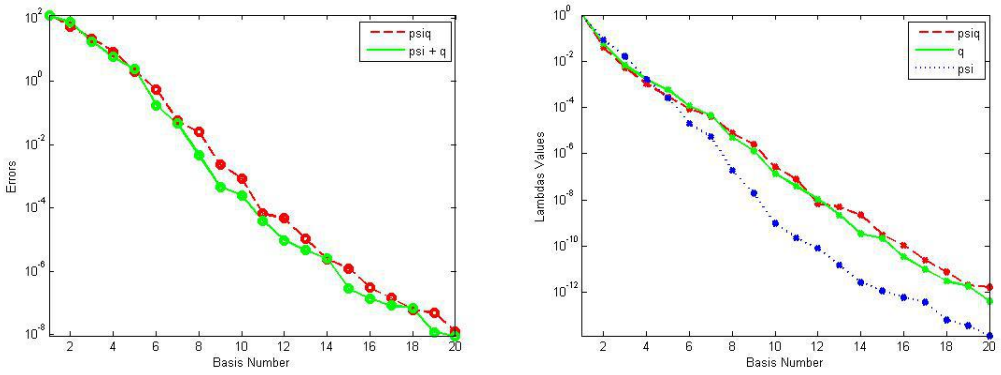

Figure 3.4.2.1: On the left we have the monolithic error (red) and partitioned error (green) plots, on the right POD eigenvalues plots for fixed μ_3 and μ_4 (red for monolithic version, green and blue for partitioned version on component q and ψ respectively).

[4]We chose a different sampling distribution because μ_1 and μ_2 usually assume small values to have a real physical meaning: we wanted to replicate this in EIM-POD algorithm.

[5]The norm for the error computation in $||\cdot||_{\mathbb{V}\times\mathbb{V}}$ is defined as $||v||^2_{\mathbb{V}\times\mathbb{V}}$, where v is the aggregated state variable.

[6]The norm for the error computation in $||\cdot||_{\mathbb{V}\times\mathbb{V}}$ is defined as $||(\psi, q)||^2_{\mathbb{V}\times\mathbb{V}} = ||\psi||^2_{\mathbb{V}} + ||q||^2_{\mathbb{V}}$.

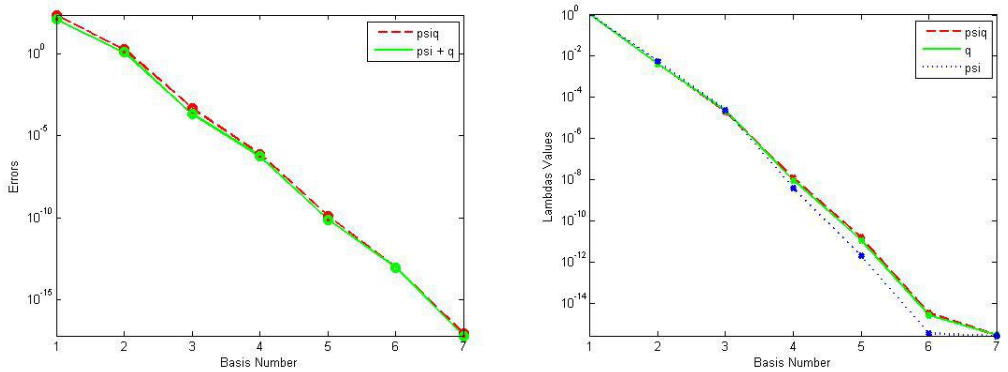

Figure 3.4.2.2: On the left we have the monolithic error (red) and partitioned error (green) plots, on the right POD eigenvalues plots for fixed μ_1 and μ_2 (red for monolithic version, green and blue for partitioned version on component q and ψ respectively).

No Fixed Parameters

The last case has the abstract form described in the introduction to this section. Let $\mu = [\mu_1, \mu_2, \mu_3, \mu_4]$ be our vector parameter varying in $\mathcal{P} = [10^{-4}, 1]^4$. The problem is more complex than before: we have to take in consideration that the reduction issue adds up to a not affine formulation. Figure 3.4.2.3 shows errors and POD eigenvalues. In this case, we reached a good result either via splitting or not, so the method are comparable. The errors trend is the same for both methods.

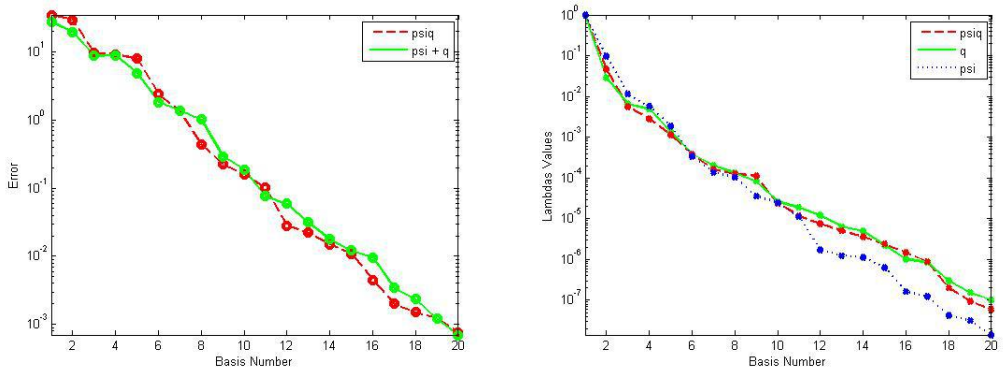

Figure 3.4.2.3: On the left we have the monolithic error (red) and partitioned error (green) plots, on the right POD eigenvalues plots for no fixed parameters (red for monolithic version, green and blue for partitioned version on component q and ψ respectively).

In the following figures, the difference between the truth state and the reduced state for the variable ψ are shown. In figure 3.4.2.4 the case with only μ_1 and μ_2 fixed is analysed. Figure 3.4.2.5 represents the pointwise error in the specific case of only μ_3 and μ_4 fixed, while in figure 3.4.2.6 the no fixed parameters problem is shown. The last two figures refer to $\boldsymbol{\mu} = [0., 0.07^3, 0.5, 0.1]$.

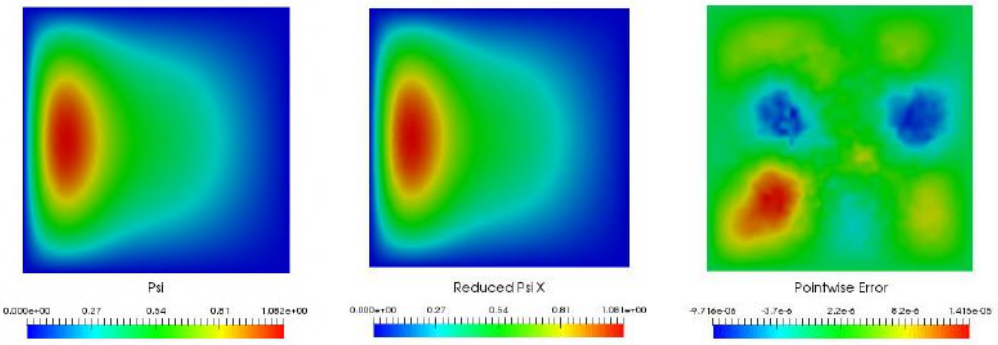

Figure 3.4.2.4: *Left*: truth solution, *center*: reduced solution, *right*: difference between state and desired state.

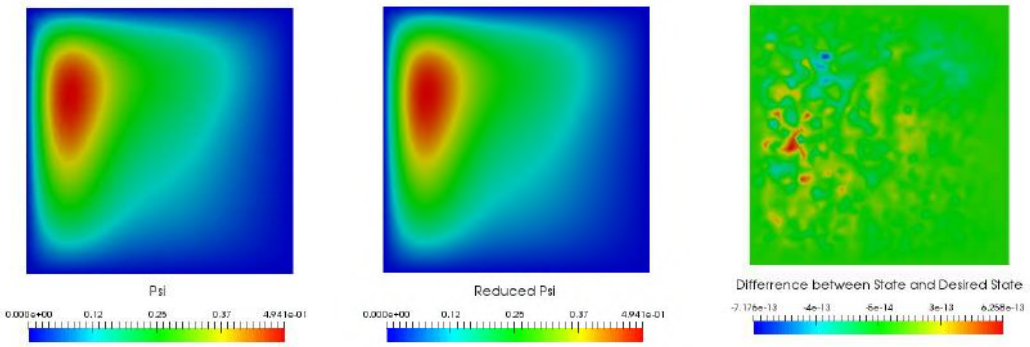

Figure 3.4.2.5: *Left*: truth solution, *center*: reduced solution, *right*: difference between state and desired state.

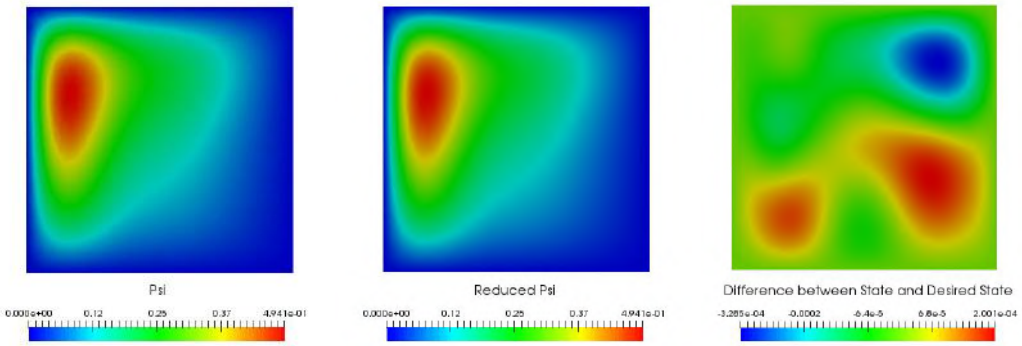

Figure 3.4.2.6: *Left*: truth solution, *center*: reduced solution, *right*: difference between state and desired state.

Chapter 4

Reduced Basis Method for Parametrized Optimal Control Problems

This chapter aims at generalizing the reduced basis techniques described in the previous Chapter to parametrized optimal control problems. As we have already specified, solving a parametric control problem is a very costly operation: we will use RB methods to solve lower dimensional approximate problems compared to the full order discretization. RB can be very useful since parametrized control applications are various (i.e. see [47, 59, 64, 53]) and all of them are computationally demanding. Our analysis focuses on linear quadratic optimal control problems governed by elliptic state equations and by Stokes equations.
In Chapter 1 we had introduced the saddle-point formulation for linear quadratic control problems. Naturally, we can extend this concept and define a *parametrized saddle-point formulation* for Optimal Control Problem (OCP($\boldsymbol{\mu}$)), where $\boldsymbol{\mu}$ is a parameter defined in finite parameter space $\mathcal{P}$.
Let us indicate, as usual, the state space and the control space with Y and U, respectively. Let Q be the adjoint space. We assume that the adjoint space and the state space will coincide, i.e. $Y = Q$. Let $X = Y \times U$ be the aggregated space of state and control. The generalized continuous parametric formulation of a linear quadratic control reads as follows:

$$\begin{cases} \min\limits_{x \in X} \mathcal{J}(x, \boldsymbol{\mu}) = \dfrac{1}{2}\mathcal{A}(x, x; \boldsymbol{\mu}) - \langle F(\boldsymbol{\mu}), x \rangle \\ \text{subject to } \mathcal{B}(x, q; \boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), q \rangle \qquad \forall q \in Q. \end{cases} \tag{4.0.1}$$

Analogously to what we have presented in the first chapter, it will be proved that the optimality system of the ($OCP(\boldsymbol{\mu})$) (4.2.1) has the following form:

$$\begin{cases} \mathcal{A}(x, w; \boldsymbol{\mu}) + \mathcal{B}(w, p; \boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), w \rangle & \forall w \in X, \\ \mathcal{B}(x, q; \boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), q \rangle & \forall q \in Q. \end{cases} \tag{4.0.2}$$

The goal of the RB methods is to find a low order approximated solution of the problem (4.0.2). Let $X_N \subset X$ and $Q_N \subset Q$ be the RB approximation spaces for the aggregated of state and control space and the adjoint space, respectively. In other words, one wants to solve: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(x_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in X_N \times Q_N$ such that

$$\begin{cases} \mathcal{A}(x_N, w_N; \boldsymbol{\mu}) + \mathcal{B}(w_N, p_N; \boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), w_N \rangle & \forall w_N \in X_N, \\ \mathcal{B}(x_N, q_N; \boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), q \rangle & \forall q_N \in Q_N. \end{cases} \tag{4.0.3}$$

In this chapter we are going to analyse the existence, uniqueness and stability of a RB solution for the problem formulation (4.0.3). In Section 4.1 we will focus on OCP($\boldsymbol{\mu}$) with elliptic state equation and their RB approximation. Section 4.2 presents the RB approximation of an OCP($\boldsymbol{\mu}$) governed by Stokes equations. Finally, in the last section, some numerical results of linear quadratic parametrized control problems are shown.

4.1 Reduced OCP($\boldsymbol{\mu}$) Governed by Elliptic State Equations

In this section we are going to present a theoretical framework to have RB methods applied to linear quadratic parametrized OCP($\boldsymbol{\mu}$) with elliptic state equation. As we have already anticipeted in the introduction to this chapter, the structure of a parametric saddle-point problem is exploited.
The main references that we have followed are [54, 52]. First of all, the whole theoretical problem structure is described, than the truth approximation, the RB approximation and the technique of aggregated spaces are analysed.

4.1.1 Problem Formulation

We have already described the framework of an Elliptic OCP in Section 1.2.2. The next step is to generalize the concept to parametrized control problems. Let Ω be a open, bounded and regular domain. As usual, let us indicate the state space with Y, the adjoint space with Q and with U the control space (remember that the control can be on a portion of Ω, in other words, on the boundary or on a subset of the spatial domain). The assumption $Q = Y$ is made. Furthermore, the observation space is H, with $y_d(\boldsymbol{\mu}) \in H$, and $G(\boldsymbol{\mu}) \in Q^*$.
We can consider the weak formulation of the OCP($\boldsymbol{\mu}$):

$$\begin{aligned} \min_{(y,u)\in Y\times U} J(y,u) &= \frac{1}{2} m(y - y_d(\boldsymbol{\mu}), y - y_d(\boldsymbol{\mu}); \boldsymbol{\mu}) + \frac{\alpha}{2} n(u,u;\boldsymbol{\mu}) \\ \text{such that } a(y,q;\boldsymbol{\mu}) &= c(u,q;\boldsymbol{\mu}) + \langle G(\boldsymbol{\mu}), q\rangle \qquad \forall q \in Q. \end{aligned} \tag{4.1.1}$$

Let us specify the hypotheses on the various forms of the problem: the bilinear form $a(\cdot,\cdot;\boldsymbol{\mu}) : Y \times Q \to \mathbb{R}$ is continuous over $Y \times Q$, that is

$$\gamma_a(\boldsymbol{\mu}) = \sup_{y\in Y}\sup_{q\in Q} \frac{a(y,q;\boldsymbol{\mu})}{\|y\|_Y \|q\|_Q} < +\infty \qquad \forall \boldsymbol{\mu} \in \mathcal{P}, \tag{4.1.2}$$

where $\mathcal{P}$ is the parameter space. Moreover, the bilinear form $a(\cdot,\cdot;\boldsymbol{\mu})$ is coercive over $Y = Q$, in other words, there exists $\alpha_0 > 0$ such that:

$$\alpha_a(\boldsymbol{\mu}) = \inf_{y\in Y} \frac{a(y,y;\boldsymbol{\mu})}{\|y\|_Y^2} = \inf_{q\in Q} \frac{a(q,q;\boldsymbol{\mu})}{\|q\|_Q^2} \geq \alpha_{a0} \qquad \forall \boldsymbol{\mu} \in \mathcal{P}. \tag{4.1.3}$$

Furthermore, the bilinear forms $c(\cdot,\cdot;\boldsymbol{\mu}) : U \times Q \to \mathbb{R}$ is symmetric and continuous, that is

$$\gamma_c(\boldsymbol{\mu}) = \sup_{u\in U}\sup_{q\in Q} \frac{c(u,q;\boldsymbol{\mu})}{\|u\|_U \|q\|_Q} < +\infty \qquad \forall \boldsymbol{\mu} \in \mathcal{P}, \tag{4.1.4}$$

and the bilinear form $n(\cdot,\cdot;\boldsymbol{\mu}) : U \times U \to \mathbb{R}$ is symmetric, continuous over $U \times U$ and coercive over U, that is

$$\gamma_n(\boldsymbol{\mu}) = \sup_{u\in U}\sup_{v\in U} \frac{n(u,v;\boldsymbol{\mu})}{\|u\|_U \|v\|_U} < +\infty \qquad \forall \boldsymbol{\mu} \in \mathcal{P}, \tag{4.1.5}$$

and

$$\alpha_n(\boldsymbol{\mu}) = \inf_{u \in U} \frac{n(u,u;\boldsymbol{\mu})}{\|u\|_U^2} \geq \alpha_{n0} > 0 \qquad \forall \boldsymbol{\mu} \in \mathcal{P}. \tag{4.1.6}$$

We have to consider other assumptions onto the bilinear form $m(\cdot,\cdot;\boldsymbol{\mu})$: it has to be symmetric, continuous and positive in the norm induced by the observation space H.
Another property that we have to guarantee is that the forms are affine in $\boldsymbol{\mu} \in \mathcal{P}$: this hypothesis is crucial for Offline-Online stages. We have to require that for some finite $Q_a, Q_c, Q_n, Q_m, Q_G, Q_{y_d}$, the forms can be expressed as follows:

$$\begin{aligned}
a(y,w;\boldsymbol{\mu}) &= \sum_{q=1}^{Q_a} \Theta_a^q(\boldsymbol{\mu}) a^q(y,w), & c(u,w;\boldsymbol{\mu}) &= \sum_{q=1}^{Q_c} \Theta_c^q(\boldsymbol{\mu}) c^q(u,w) \\
m(y,z;\boldsymbol{\mu}) &= \sum_{q=1}^{Q_m} \Theta_m^q(\boldsymbol{\mu}) m^q(y,z) & n(u,v;\boldsymbol{\mu}) &= \sum_{q=1}^{Q_n} \Theta_n^q(\boldsymbol{\mu}) n^q(u,v) \\
\langle G(\boldsymbol{\mu}), w\rangle &= \sum_{q=1}^{Q_G} \Theta_G^q(\boldsymbol{\mu}) \langle G^q, w\rangle, & y_d(\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{y_d}} \Theta_{y_d}^q(\boldsymbol{\mu}) y_d^q,
\end{aligned} \tag{4.1.7}$$

where $a^q(\cdot,\cdot), c^q(\cdot,\cdot), m^q(\cdot,\cdot), n^q(\cdot,\cdot), G^q, y_d^q$ are independent from the parameters, while $\Theta_a^q, \Theta_c^q, \Theta_m^q, \Theta_n^q, \Theta_G^q, \Theta_{y_d}^q$ are smooth functions depending on $\boldsymbol{\mu}$.
The problem (4.1.1) can be formulated in a saddle-point framework. Let us define $X = Y \times U$. Let $x = (y,u)$ and $w = (z,v)$ be two elements of X, $p, q \in Q$. We can define the bilinear forms $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu}) : X \times X \to \mathbb{R}$ and $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu}) : X \times Q \to \mathbb{R}$ as follows:

$$\begin{aligned}
\mathcal{A}(x,w;\boldsymbol{\mu}) &= m(y,z;\boldsymbol{\mu}) + \alpha n(u,v;\boldsymbol{\mu}) && \forall x, w \in X, \\
\mathcal{B}(w,q;\boldsymbol{\mu}) &= a(z,q;\boldsymbol{\mu}) - c(v,q;\boldsymbol{\mu}) && \forall w \in X \text{ and } \forall q \in Q.
\end{aligned}$$

As we have already seen in subsection 1.2.2, defining $F(\boldsymbol{\mu}) = m(y_d(\boldsymbol{\mu}),\cdot) \in X^*$, given $\boldsymbol{\mu} \in \mathcal{P}$, one can recast the problem (4.1.1) as:

$$\begin{cases}
\min\limits_{x \in X} \mathcal{J}(x,\boldsymbol{\mu}) = \dfrac{1}{2}\mathcal{A}(x,x;\boldsymbol{\mu}) - \langle F(\boldsymbol{\mu}), x\rangle \\
\text{subject to } \mathcal{B}(x,q;\boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), q\rangle \qquad \forall q \in Q.
\end{cases} \tag{4.1.8}$$

Now, the assumptions made on the linear and the bilinear forms allow to fulfill the hypotheses of Theorem 1.4, i.e. the problem (4.1.8) is equivalent to the following one: given a parameter $\boldsymbol{\mu} \in \mathcal{P}$, find the solutions $(x(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \in X \times Q$ that verify:

$$\begin{cases}
\mathcal{A}(x(\boldsymbol{\mu}), w;\boldsymbol{\mu}) + \mathcal{B}(w, p(\boldsymbol{\mu});\boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), w\rangle & \forall w \in X, \\
\mathcal{B}(x(\boldsymbol{\mu}), q;\boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), q\rangle & \forall q \in Q.
\end{cases} \tag{4.1.9}$$

First of all, let us notice that the affine assumption (4.1.7), allow us to express the bilinear and the linear forms of the problem 4.1.9 in the following way:

$$\begin{aligned}
\mathcal{A}(x,w;\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathcal{A}}} \Theta_{\mathcal{A}}^q(\boldsymbol{\mu}) \mathcal{A}^q(y,w), & \mathcal{B}(w,p;\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathcal{B}}} \Theta_{\mathcal{B}}^q(\boldsymbol{\mu}) \mathcal{B}^q(w,p), \\
\langle G(\boldsymbol{\mu}), s\rangle &= \sum_{q=1}^{Q_G} \Theta_G^q(\boldsymbol{\mu}) \langle G^q, s\rangle, & \langle F(\boldsymbol{\mu}), w\rangle &= \sum_{q=1}^{Q_F} \Theta_F^q(\boldsymbol{\mu}) \langle F^q, w\rangle,
\end{aligned} \tag{4.1.10}$$

with $\Theta_{\mathcal{A}}, \Theta_{\mathcal{B}}, \Theta_G, \Theta_F$ are $\boldsymbol{\mu}-$dependent smooth functions, and the linear and the bilinear forms $\mathcal{A}^q(\cdot,\cdot), \mathcal{B}^q(\cdot,\cdot), G^q, F^q$ are $\boldsymbol{\mu}-$independent.
Moreover we have to guarantee the following assumptions.

Assumptions 4.1. The bilinear forms $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ and $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ have to verify:

1. the bilinear form $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ must be symmetric and nonnegative over X;
2. the bilinear $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ is continuous over $X \times X$, i.e. it holds:

$$\gamma_{\mathcal{A}}(\boldsymbol{\mu}) = \sup_{x\in X}\sup_{w\in X} \frac{\mathcal{A}(x,w;\boldsymbol{\mu})}{\|x\|_X\|w\|_X} < +\infty, \qquad \boldsymbol{\mu}\in\mathcal{P};$$

3. Let us define

$$X_0 = \{w \in X \text{ such that } \mathcal{B}(w,q;\boldsymbol{\mu}) = 0, \ \ \forall q \in Q\}.$$

   The bilinear $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ is coercive over X_0, i.e. there exists a constant $\alpha_{\mathcal{A}0} > 0$ such that

$$\alpha(\boldsymbol{\mu}) = \inf_{x\in X_0} \frac{\mathcal{A}(x,x;\boldsymbol{\mu})}{\|x\|_X^2} \geq \alpha_{\mathcal{A}0}, \qquad \forall \boldsymbol{\mu}\in\mathcal{P};$$

4. the bilinear form $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ is continuous over $X \times Q$, i.e.

$$\gamma_{\mathcal{B}}(\boldsymbol{\mu}) = \sup_{w\in X}\sup_{q\in Q} \frac{\mathcal{B}(w,q;\boldsymbol{\mu})}{\|w\|_X\|q\|_Q} < +\infty, \qquad \forall \boldsymbol{\mu}\in\mathcal{P};$$

5. the bilinear form $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ verifies the inf-sup condition over $X \times Q$, in other words, there exists a constant $\beta_{\mathcal{B}0}$ such that

$$\beta(\boldsymbol{\mu}) = \inf_{q\in Q}\sup_{w\in X} \frac{\mathcal{B}(w,q;\boldsymbol{\mu})}{\|w\|_X\|q\|_Q} \geq \beta_{\mathcal{B}0} > 0, \qquad \forall \boldsymbol{\mu}\in\mathcal{P}.$$

As we have seen in chapter 1, the hypotheses made on the bilinear forms of the state equation, guarantee the fulfillment of assumptions 4.1

4.1.2 Full Order Approximation

We are going to adapt the concepts introduced in subsection 2.1.2 to an OCP($\boldsymbol{\mu}$). Let $\{\mathcal{T}^{\mathcal{N}}\}$ be a triangulation of the domain Ω. Let us define the following space

$$X^r_{\mathcal{N}} = \{v^{\mathcal{N}} \in C^0(\overline{\Omega}) \ : \ v^{\mathcal{N}}|_K \in \mathbb{P}_r, \ \ \forall K \in \mathcal{T}^{\mathcal{N}}\}.$$

and $\mathbb{P}_r$ represents the space of polynomials of degree at most equal to r. Now we can define $Y^{\mathcal{N}} = Y \cap X^r_{\mathcal{N}}$, $Q^{\mathcal{N}} = Y^{\mathcal{N}}$ and $U^{\mathcal{N}} = U \cap X^r_{\mathcal{N}}$. By construction it holds that $Y^{\mathcal{N}} \subset Y$, $U^{\mathcal{N}} \subset U$ and $X^{\mathcal{N}} = Y^{\mathcal{N}} \times U^{\mathcal{N}} \subset X$ and $Q^{\mathcal{N}} \subset Q$. Naturally, the $\mathcal{N}$ denotes the dimension of the product space $X^{\mathcal{N}} \times Q^{\mathcal{N}}$, in other words $\mathcal{N} = \mathcal{N}_X + \mathcal{N}_Q$, where $\mathcal{N}_X = \mathcal{N}_Y + \mathcal{N}_U$.
The Galerkin Finite Element problem for an OCP($\boldsymbol{\mu}$) reads: given $\boldsymbol{\mu}\in\mathcal{P}$, find $(x^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}(\boldsymbol{\mu})) \in X^{\mathcal{N}} \times Q^{\mathcal{N}}$ such that

$$\begin{cases} \mathcal{A}(x^{\mathcal{N}}(\boldsymbol{\mu}), w^{\mathcal{N}};\boldsymbol{\mu}) + \mathcal{B}(w^{\mathcal{N}}, p^{\mathcal{N}}(\boldsymbol{\mu});\boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), w^{\mathcal{N}}\rangle, & \forall w^{\mathcal{N}} \in X^{\mathcal{N}}, \\ \mathcal{B}(x^{\mathcal{N}}(\boldsymbol{\mu}), q^{\mathcal{N}};\boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), q^{\mathcal{N}}\rangle & \forall q^{\mathcal{N}} \in Q^{\mathcal{N}}. \end{cases} \tag{4.1.11}$$

We will refer to the problem (4.1.11) as the *truth problem*. As we have seen for the parameter independent formulation of the problem (see Lemma 2.2), the assumption $Y^{\mathcal{N}} = Q^{\mathcal{N}}$ provides the bilinear form $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ to fulfill the continuity over $X^{\mathcal{N}} \times X^{\mathcal{N}}$ and the coercivity over the space

$$X_0^{\mathcal{N}} = \{w^{\mathcal{N}} \in X^{\mathcal{N}} \text{ such that } \mathcal{B}(w^{\mathcal{N}}, q^{\mathcal{N}}; \boldsymbol{\mu}) = 0,\ \forall q^{\mathcal{N}} \in Q^{\mathcal{N}}\},$$

that is

$$\gamma_{\mathcal{A}}^{\mathcal{N}}(\boldsymbol{\mu}) = \sup_{x^{\mathcal{N}} \in X^{\mathcal{N}}} \sup_{w^{\mathcal{N}} \in X^{\mathcal{N}}} \frac{\mathcal{A}(x^{\mathcal{N}}, w^{\mathcal{N}}; \boldsymbol{\mu})}{\|x^{\mathcal{N}}\|_X \|w^{\mathcal{N}}\|_X} < +\infty, \qquad \forall \boldsymbol{\mu} \in \mathcal{P};$$

and

$$\alpha^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{x^{\mathcal{N}} \in X_0^{\mathcal{N}}} \frac{\mathcal{A}(x^{\mathcal{N}}, x^{\mathcal{N}}; \boldsymbol{\mu})}{\|x^{\mathcal{N}}\|_X^2} \geq \alpha_{\mathcal{A}0} > 0, \qquad \forall \boldsymbol{\mu} \in \mathcal{P}.$$

The bilinear form $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ verifies of continuity and of inf-sup stability over the space $X^{\mathcal{N}} \times Q^{\mathcal{N}}$, in other words:

$$\gamma_{\mathcal{B}}^{\mathcal{N}} = \sup_{w^{\mathcal{N}} \in X^{\mathcal{N}}} \sup_{q^{\mathcal{N}} \in Q^{\mathcal{N}}} \frac{\mathcal{B}(w^{\mathcal{N}}, q^{\mathcal{N}}; \boldsymbol{\mu})}{\|w^{\mathcal{N}}\|_X \|q^{\mathcal{N}}\|_Q} < +\infty, \qquad \forall \boldsymbol{\mu} \in \mathcal{P};$$

and there exists a constant $\beta_{\mathcal{B}0} > 0$ such that

$$\beta^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{q^{\mathcal{N}} \in Q^{\mathcal{N}}} \sup_{w^{\mathcal{N}} \in X^{\mathcal{N}}} \frac{\mathcal{B}(w^{\mathcal{N}}, q^{\mathcal{N}}; \boldsymbol{\mu})}{\|w^{\mathcal{N}}\|_X \|q^{\mathcal{N}}\|_Q} \geq \beta_{\mathcal{B}0}, \qquad \forall \boldsymbol{\mu} \in \mathcal{P}.$$

Thanks to Lemma 2.2 and the fulfillment of the hypotheses of Proposition 2.2 the discrete Finite Element problem (4.1.11) is well-posed.

Remark 4.1.1. Let us express the concepts just presented into an algebraic framework. Following the same steps made in Chapter 2, an OCP($\boldsymbol{\mu}$) governed by elliptic state equation leads to the following linear system:

$$\begin{pmatrix} A(\boldsymbol{\mu}) & B^T(\boldsymbol{\mu}) \\ B(\boldsymbol{\mu}) & 0 \end{pmatrix} \begin{pmatrix} x^{\mathcal{N}}(\boldsymbol{\mu}) \\ p^{\mathcal{N}}(\boldsymbol{\mu}) \end{pmatrix} = \begin{pmatrix} \mathbf{F}(\boldsymbol{\mu}) \\ \mathbf{G}(\boldsymbol{\mu}) \end{pmatrix}. \tag{4.1.12}$$

The affine decomposition assumption is naturally inherited by the matrices $A(\boldsymbol{\mu})$ and $B(\boldsymbol{\mu})$ associated to the bilinear forms, and by the matrices of the linear forms $\mathbf{F}(\boldsymbol{\mu})$ and $\mathbf{G}(\boldsymbol{\mu})$. In other words, the following equalities hold:

$$\begin{aligned} A(\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathcal{A}}} \Theta_{\mathcal{A}}^q(\boldsymbol{\mu}) A^q, & B(\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathcal{B}}} \Theta_{\mathcal{B}}^q(\boldsymbol{\mu}) B^q, \\ \mathbf{G}(\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathbf{G}}} \Theta_{G}^q(\boldsymbol{\mu}) \mathbf{G}^q, & \mathbf{F}(\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathbf{F}}} \Theta_{F}^q(\boldsymbol{\mu}) \mathbf{F}^q. \end{aligned} \tag{4.1.13}$$

4.1.3 Reduced Basis Approximation

We are going to extend the notions of reduced basis methods for parametric PDE to OPC($\boldsymbol{\mu}$) problems. The path to reach our goal is substantially the same discussed in Chapter 3: the idea is to use a new approximated space that has a basis composed by well-chosen solutions $(x^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}(\boldsymbol{\mu}))$ of the problem (4.1.11). An assumption has to be made: the discrete solution has a smooth dependence on $\boldsymbol{\mu}$. This hypothesis allow the

parametric manifold $\mathcal{M}$ to be smooth and to be approximated by *full order snapshots*, solutions of (4.1.11).
Let us take $N \in \{1, \dots, N_{max}\}$ and a set of parameters $S_N = \{\boldsymbol{\mu}^1, \dots, \boldsymbol{\mu}^N\}$ and the Finite Element solutions $\{(x^{\mathcal{N}}(\boldsymbol{\mu}^n), p^{\mathcal{N}}(\boldsymbol{\mu}^n))\}_{n=1}^N$. Let us define the reduced spaces for state, control and adjoint variable, respectively given by:

$$\begin{aligned} Y_N &= \text{span } \{\zeta_n := y^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}, \\ U_N &= \text{span } \{\lambda_n := u^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}, \\ Q_N &= \text{span } \{\xi_n := p^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}. \end{aligned}$$

Let us define $X_N = Y_N \times U_N$ in order to recast our problem into a saddle-point formulation. The OCP($\boldsymbol{\mu}$) problem reads: given $\boldsymbol{\mu} \in \mathcal{P}$ find $(x_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in X_N \times Q_N$ such that

$$\begin{cases} \mathcal{A}(x_N(\boldsymbol{\mu}), w_N; \boldsymbol{\mu}) + \mathcal{B}(w_N, p_N(\boldsymbol{\mu}); \boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), w_N \rangle, & \forall w_N \in X_N, \\ \mathcal{B}(x_N(\boldsymbol{\mu}), q_N; \boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), q_N \rangle & \forall q_N \in Q_N. \end{cases} \tag{4.1.14}$$

What we have to prove is the well-posedness of the RB approximation. The crucial point is to prove the coercivity of $\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ over the space

$$X_0^N = \{w_N \in X_N : \mathcal{B}(w_N, q_N; \boldsymbol{\mu}) = 0,\ \ \forall q_N \in Q_N\}$$

and the fulfillment of the inf-sup condition for the bilinear form $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$, since the continuity property derives directly from the Finite Element spaces.
Let us be more specific on the inf-sup condition of $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$: there exists a constant $\beta_0 > 0$ such that

$$\beta_N(\boldsymbol{\mu}) = \inf_{q_N \in Q_N} \sup_{w_N \in X_N} \frac{\mathcal{B}(w_N, q_N; \boldsymbol{\mu})}{\|w_N\|_X \|q_N\|_Q} \geq \beta_0, \qquad \forall \boldsymbol{\mu} \in \mathcal{P}. \tag{4.1.15}$$

To prove the hypothesis 4.1.15 we can follow the techniques already used in Lemma 2.1. So, by definition, we have:

$$\sup_{\substack{w_N \in X_N, \\ w_N \neq 0}} \frac{\mathcal{B}(w_N, q_N)}{\|w_N\|_X} = \sup_{\substack{(z_N, v_N) \in Y_N \times U_N, \\ (z_N, v_N) \neq (0,0)}} \frac{a(z_N, q_N) - c(v_N, q_N)}{\sqrt{\|z_N\|_Y^2 + \|v_N\|_U^2}} \underset{(z_N, v_N) = (q_N, 0)}{\geq} \frac{a(q_N, q_N)}{\|q_N\|_Y},$$

for all $q_N \in Q_N$. Notice that to exploit the coercivity of $a(\cdot, \cdot; \boldsymbol{\mu})$ in the last inequality we have to suppose $Y_N = Q_N$. For the reduced spaces, this assumption do not derive directly from the Finite Element approximation, since RB basis are linked on the parametric problem and they are not generic functions as in Finite Element case. A way to recover the inf-sup stability condition 4.1.15 is to define the *enriched RB spaces* Y_N, U_N and Q_N. To ensure the stability of the RB method we have to build what is known as *aggregated space* of state and adjoint variables

$$Z_N = \text{span } \{\zeta_n := y^{\mathcal{N}}(\boldsymbol{\mu}^n),\ \xi_n := p^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}.$$

Now, let us choose $Y_N = Z_N$, $X_N = Z_N \times U_N$ and $Q_N = Z_N$. So, the new RB approximated problem reads as: find $(x_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in X_N \times Q_N$ such that

$$\begin{cases} \mathcal{A}(x_N(\boldsymbol{\mu}), w_N; \boldsymbol{\mu}) + \mathcal{B}(w_N, p_N(\boldsymbol{\mu}); \boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), w_N \rangle, & \forall w_N \in X_N, \\ \mathcal{B}(x_N(\boldsymbol{\mu}), q_N; \boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), q_N \rangle & \forall q_N \in Q_N. \end{cases} \tag{4.1.16}$$

Exploiting the new spaces formulation we can fulfill the inf-sup condition 4.1.15, since $Y_N = Q_N = Z_N$.

Lemma 4.1. *The bilinear form $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ verifies the inf-sup condition 4.1.15. Furthermore it holds*

$$\beta_N(\boldsymbol{\mu}) \geq \alpha^{\mathcal{N}}(\boldsymbol{\mu}).$$

Proof. Let us compute

$$\begin{aligned}
\beta_N(\boldsymbol{\mu}) &= \inf_{q_N\in Z_N}\sup_{w_N\in X_N}\frac{\mathcal{B}(w_N,q_N;\boldsymbol{\mu})}{\|w_N\|_X\|q_N\|_Q} = \\
&= \inf_{q_N\in Z_N}\sup_{(z_N,v_N)\in Z_N\times U_N}\frac{a(z_N,q_N;\boldsymbol{\mu})-c(v_N,q_N;\boldsymbol{\mu})}{||(z_N,v_N)||_X||q_N||_Q} \\
&\geq \inf_{\substack{z_N=q_N\\ v_N=0}}\inf_{q_N\in Z_N}\frac{a(q_N,q_N;\boldsymbol{\mu})}{||q_N||_Q} = \alpha_N(\boldsymbol{\mu}) \geq \alpha^{\mathcal{N}}(\boldsymbol{\mu}) > 0.
\end{aligned}$$

Notice that the identification $z_N = q_N$ is now allowed, since they are both in Z_N.

Proposition 4.1. *The reduced basis saddle-point problem (4.1.16) has unique solution $(x_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in X_N \times Q_N$ for all $\boldsymbol{\mu} \in \mathcal{P}$.*

Proof. To prove the proposition is sufficient to require the fulfillment of the hypotheses of Brezzi's Theorem 2.1. As said previously the continuity of the bilinear and the linear forms over RB spaces is inherited from FE spaces. The fulfillment of the inf-sup condition 4.1.15 is proved in Lemma 4.1. The coercivity property of the bilinear form $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ can be proved miming the arguments used in Lemma 2.1 and Lemma **??**.

4.1.4 Algebraic Formulation of the Enriched RB Approximation

We will now introduce the algebraic structure of a RB approximated problem (4.1.16), with enriched space for state and adjoint variables. Let $\{\tau_j\}_{j=1}^{2N} = \{\zeta_j\}_{j=1}^{N} \cup \{\xi_j\}_{j=1}^{N}$ be the basis functions for the space Z_N, i.e.

$$Z_N = \text{span } \{\tau_j,\ j = 1,\dots,2N\}.$$

We can know write our state, control and adjoint variables in the following way:

$$y_N(\boldsymbol{\mu}) = \sum_{j=1}^{2N} y_{N_j}(\boldsymbol{\mu})\tau_j, \qquad u_N(\boldsymbol{\mu}) = \sum_{j=1}^{N} u_{N_j}(\boldsymbol{\mu})\lambda_j, \qquad p_N(\boldsymbol{\mu}) = \sum_{j=1}^{2N} p_{N_j}(\boldsymbol{\mu})\tau_j.$$

Let us define the product space $X_N = Z_N \times U_N$. One can generate X_N from the functions $\{\sigma_j\}_{j=1}^{3N}$ where

$$\sigma_j = \begin{cases}(\tau_j, 0), & j = 1,\dots,2N\\ (0,\lambda_{j-2N}) & j = 2N+1,\dots,3N,\end{cases}$$

in other words

$$x_N(\boldsymbol{\mu}) = (y_N(\boldsymbol{\mu}), u_n(\boldsymbol{\mu})) = \left(\sum_{j=1}^{2N} y_{N_j}(\boldsymbol{\mu})\tau_j, \sum_{j=1}^{N} u_{N_j}(\boldsymbol{\mu})\lambda_j\right).$$

In algebraic formulation, the reduced problem (4.1.16) reads:

$$\begin{cases}
\displaystyle\sum_{j=1}^{3N}\sum_{q=1}^{Q_{A_N}}\Theta^q_{A_N}(\boldsymbol{\mu})A^q_{N_{ij}}x_{N_j}(\boldsymbol{\mu}) + \sum_{k=1}^{2N}\sum_{q=1}^{Q_{B_N}}\Theta^q_{B_N}(\boldsymbol{\mu})B^q_{N_{ki}}p_{N_k}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{F_N}}\Theta^q_{F_N}(\boldsymbol{\mu})F^q_{N_i}, & 1\leq i\leq 3N\\
\displaystyle\sum_{j=1}^{3N}\sum_{q=1}^{Q_{B_N}}\Theta^q_{B_N}(\boldsymbol{\mu})B^q_{N_{ij}}x_{N_j}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{G_N}}\Theta^q_{G_N}(\boldsymbol{\mu})G^q_{N_i}, & 1\leq i\leq 2N
\end{cases} \tag{4.1.17}$$

where the matrices linked to bilinear and linear forms are given by

$$\begin{aligned} A^q_{N_{ij}} &= \mathcal{A}^q(\sigma_i, \sigma_j) && 1 \le i, j \le 3N \\ B^q_{N_{ij}} &= \mathcal{B}^q(\sigma_j, \tau_i) && 1 \le j \le 3N,\ 1 \le i \le 2N \\ F^q_{N_i} &= \langle F^q, \sigma_i \rangle && 1 \le i \le 3N \\ G^q_{N_i} &= \langle G^q, \tau_i \rangle && 1 \le i \le 2N. \end{aligned}$$

If one denotes with

$$A_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{A_N}} \Theta^q_{A_N}(\boldsymbol{\mu}) A^q_N, \qquad B_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{B_N}} \Theta^q_{B_N}(\boldsymbol{\mu}) B^q_N,$$
$$\mathbf{G}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{\mathbf{G}_N}} \Theta^q_{\mathbf{G}_N}(\boldsymbol{\mu}) G^q_N, \qquad \mathbf{F}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{\mathbf{F}_N}} \Theta^q_{\mathbf{F}_N}(\boldsymbol{\mu}) F^q_N,$$

the linear system (4.1.17) can be written as

$$\begin{pmatrix} A_N(\boldsymbol{\mu}) & B_N^T(\boldsymbol{\mu}) \\ B_N(\boldsymbol{\mu}) & 0 \end{pmatrix} \begin{pmatrix} x_N(\boldsymbol{\mu}) \\ p_N(\boldsymbol{\mu}) \end{pmatrix} = \begin{pmatrix} \mathbf{F}_N(\boldsymbol{\mu}) \\ \mathbf{G}_N(\boldsymbol{\mu}) \end{pmatrix}. \tag{4.1.18}$$

Let $Z_z = (\tau_1, \dots, \tau_N) \in \mathbb{R}^{\mathcal{N}} \times \mathbb{R}^{2N}$ and $Z_u = (\lambda_1, \dots, \lambda_N) \in \mathbb{R}^{\mathcal{N}} \times \mathbb{R}^N$ be the *basis matrix*. Thanks to these definitions we can build the following matrices:

$$Z_x = \begin{pmatrix} Z_z & 0 \\ 0 & Z_u \end{pmatrix} \in \mathbb{R}^{2\mathcal{N}} \times \mathbb{R}^{3N}, \qquad Z = \begin{pmatrix} Z_z & 0 & 0 \\ 0 & Z_u & 0 \\ 0 & 0 & Z_z \end{pmatrix} \in \mathbb{R}^{3\mathcal{N}} \times \mathbb{R}^{5N}.$$

So we can express the left hand side matrix of the system (4.1.18) in the following way, underlining how it is linked to the Finite Element space:

$$\begin{pmatrix} A_N(\boldsymbol{\mu}) & B_N^T(\boldsymbol{\mu}) \\ B_N(\boldsymbol{\mu}) & 0 \end{pmatrix} = \begin{pmatrix} Z_x^T A Z_x & Z_x^T B Z_z \\ Z_z^T B Z_x & 0 \end{pmatrix}, \tag{4.1.19}$$

the new matrix formulation is still symmetric and the dimension of the system is $5N \times 5N$.

4.2 Reduced OCP($\boldsymbol{\mu}$) Governed by Stokes State Equations

In this section we are going to provide a reduced basis approach to deal with linear quadratic OCP($\boldsymbol{\mu}$) governed by Stokes state equations. It is very important to understand how to face this kind of problem, since Stokes equation are exploited in several and different applications (i.e see [64, 61, 25]). The structure of the section follows the structure used in section 4.1: we will use the RB framework and we will study its stability. The main theoretical references are [53, 52].

4.2.1 Problem Formulation

We are going to introduce the parametric version of a linear quadratic control problem governed by Stokes state equation. Let us consider the open, bounded and regular domain

$\Omega \subset \mathbb{R}^2$. Let $H_0^1(\Omega) \times H_0^1(\Omega) \subset V \subset H^1(\Omega) \times H^1(\Omega)$ be the velocity space. The pressure space is given by

$$P := L_0^2(\Omega) = \left\{ r \in L^2(\Omega) \ : \ \int_\Omega r = 0 \right\}.$$

Then, the state space is $Y = V \times P$. As adjoint space we consider $Q = Y$, whereas the control space is given by $U = L^2(\Omega) \times L^2(\Omega)$. The state variable will be indicated with $y = (\mathbf{v}, p) \in V \times P$. The observation space taken into consideration is $H = L^2(\Omega)$.
The OCP($\boldsymbol{\mu}$) reads as:

$$\min_{(y,u)\in Y\times U} J(y,u;\boldsymbol{\mu}) = \frac{1}{2} m(\mathbf{v}-\mathbf{v}_d, \mathbf{v}-\mathbf{v}_d;\boldsymbol{\mu}) + \frac{\alpha}{2} n(\mathbf{u},\mathbf{u};\boldsymbol{\mu})$$
$$\text{subject to} \begin{cases} a(\mathbf{v},\boldsymbol{\phi},\boldsymbol{\mu}) + b(\boldsymbol{\phi},p;\boldsymbol{\mu}) = c(\mathbf{u},\boldsymbol{\phi};\boldsymbol{\mu}) + \langle F(\boldsymbol{\mu}),\boldsymbol{\phi}\rangle & \forall \boldsymbol{\phi} \in V \\ b(\mathbf{v},\xi;\boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}),\xi\rangle & \forall \xi \in P, \end{cases}$$

where $F(\boldsymbol{\mu}) \in V^*$, $G(\boldsymbol{\mu}) \in P^*$ and $\mathbf{v}_d \in H$. The bilinear forms $a(\cdot,\cdot;\boldsymbol{\mu})$, $b(\cdot,\cdot;\boldsymbol{\mu})$ and $c(\cdot,\cdot;\boldsymbol{\mu})$ are the parametric equivalent of the forms introduced in subsection 2.1.4. Let us remind the hypotheses made over the bilinear forms:

1. the bilinear form $c(\cdot,\cdot;\boldsymbol{\mu}) : U \times V \to \mathbb{R}$ must be symmetric and bounded over $U \times V$,

2. the bilinear form $n(\cdot,\cdot;\boldsymbol{\mu}) : U \times U \to \mathbb{R}$ must be symmetric, bounded over $U \times U$ and coercive over U,

3. the bilinear form $m(\cdot,\cdot;\boldsymbol{\mu}) : H \times H \to \mathbb{R}$ must be symmetric, continuous and positive in the norm induced by the space H.

The OCP($\boldsymbol{\mu}$) can be recast as follows:

$$\begin{aligned} &\min_{(y,u)\in Y\times U} J(y,u;\boldsymbol{\mu}) = \frac{1}{2} m(\mathbf{v}-\mathbf{v}_d, \mathbf{v}-\mathbf{v}_d;\boldsymbol{\mu}) + \frac{\alpha}{2} n(\mathbf{u},\mathbf{u};\boldsymbol{\mu}) \\ &\text{such that } \mathbf{A}((\mathbf{v},p),(\boldsymbol{\phi},\xi),\boldsymbol{\mu}) = \langle \mathbf{G}(\boldsymbol{\mu}),(\boldsymbol{\phi},\xi)\rangle + \mathbf{C}(\mathbf{u},(\boldsymbol{\phi},\xi);\boldsymbol{\mu}) \quad \forall (\boldsymbol{\phi},\xi) \in Q, \end{aligned} \tag{4.2.1}$$

where the bilinear form $\mathbf{A}(\cdot,\cdot;\boldsymbol{\mu}) = Y \times Q \to \mathbb{R}$ is defined as

$$\mathbf{A}((\mathbf{v},p),(\boldsymbol{\phi},\xi),\boldsymbol{\mu}) = a(\mathbf{v},\boldsymbol{\phi},\boldsymbol{\mu}) + b(\boldsymbol{\phi},p;\boldsymbol{\mu}) + b(\mathbf{v},\xi;\boldsymbol{\mu}), \tag{4.2.2}$$

while the functional $\mathbf{G}(\boldsymbol{\mu}) \in Q^*$ is given by

$$\langle \mathbf{G}(\boldsymbol{\mu}),(\boldsymbol{\phi},\xi)\rangle = \langle F(\boldsymbol{\mu}),\boldsymbol{\phi}\rangle + \langle G(\boldsymbol{\mu}),\xi\rangle, \tag{4.2.3}$$

and the bilinear form $\mathbf{C}(\cdot,\cdot;\boldsymbol{\mu}) : U \times Q \to \mathbb{R}$ is

$$\mathbf{C}(\mathbf{u},(\boldsymbol{\phi},\xi);\boldsymbol{\mu}) = c(\mathbf{u},\boldsymbol{\phi};\boldsymbol{\mu}). \tag{4.2.4}$$

Naturally, to have an efficient RB approximation of the problem, we have to require the

forms to be affine in the parameter $\boldsymbol{\mu}$: for some $Q_{\mathbf{A}}, Q_c, Q_m, Q_n, Q_{\mathbf{G}}$ it holds:

$$\begin{aligned}
\mathbf{A}((\mathbf{v},p),(\boldsymbol{\phi},\xi);\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathbf{A}}} \Theta_{\mathbf{A}}^q(\boldsymbol{\mu}) A^q((\mathbf{v},p),(\boldsymbol{\phi},\xi)),\\
\mathbf{C}(\mathbf{u},(\boldsymbol{\phi},\xi);\boldsymbol{\mu}) &= \sum_{q=1}^{Q_c} \Theta_c^q(\boldsymbol{\mu}) \mathbf{C}^q(\mathbf{u},(\boldsymbol{\phi},\xi)),\\
m(\mathbf{v},\boldsymbol{\psi};\boldsymbol{\mu}) &= \sum_{q=1}^{Q_m} \Theta_m^q(\boldsymbol{\mu}) m^q(\mathbf{v},\boldsymbol{\psi}),\\
n(\mathbf{u},\boldsymbol{\tau};\boldsymbol{\mu}) &= \sum_{q=1}^{Q_n} \Theta_n^q(\boldsymbol{\mu}) n^q(\mathbf{u},\boldsymbol{\tau}),\\
\langle \mathbf{G}(\boldsymbol{\mu}),(\boldsymbol{\phi},\xi)\rangle &= \sum_{q=1}^{Q_{\mathbf{G}}} \Theta_{\mathbf{G}}^q(\boldsymbol{\mu}) \langle \mathbf{G}^q, w\rangle
\end{aligned} \tag{4.2.5}$$

where $\Theta_{\mathbf{A}}^q, \Theta_c^q, \Theta_m^q, \Theta_n^q, \Theta_{\mathbf{G}}^q$ are smooth $\boldsymbol{\mu}$-dependent functions, and $\mathbf{A}^q(\cdot,\cdot)$, $\mathbf{C}^q(\cdot,\cdot), m^q(\cdot,\cdot), n^q(\cdot,\cdot), \mathbf{G}^q$ are continuous linear and bilinear parameter independent forms. Our goal is to recast the OCP($\boldsymbol{\mu}$) (4.2.1) in a saddle-point formulation. Let us introduce the space $X = Y \times U$, the variables $x = ((\mathbf{v},p),\mathbf{u})$, $\zeta = ((\boldsymbol{\psi},\pi),\boldsymbol{\tau}) \in X$ and $(\boldsymbol{\phi},\xi) \in Q$. Let $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu}) : X \times X \to \mathbb{R}$ be a bilinear form defined as:

$$\mathcal{A}(x,\zeta;\boldsymbol{\mu}) = m(\mathbf{v},\boldsymbol{\psi};\boldsymbol{\mu}) + \alpha n(\mathbf{u},\boldsymbol{\tau};\boldsymbol{\mu}), \qquad \forall x,\zeta \in X,$$

Let $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu}) : X \times Q \to \mathbb{R}$ be a bilinear form given by:

$$\mathcal{B}(x,(\boldsymbol{\phi},\xi),\boldsymbol{\mu}) = \mathbf{A}((\mathbf{v},p),(\boldsymbol{\phi},\xi);\boldsymbol{\mu}) - \mathbf{C}(\mathbf{u},(\boldsymbol{\phi},\xi),\boldsymbol{\mu}), \forall x \in X,\ \forall (\boldsymbol{\phi},\xi) \in Q.$$

As usual, $\mathbf{F}(\boldsymbol{\mu}) = m(\mathbf{v}_d,\cdot) \in X^*$. Thanks these new linear and bilinear forms the problem (4.2.1) reads: given $\boldsymbol{\mu} \in \mathcal{P}$

$$\begin{cases} \min\limits_{x\in X} \mathcal{J}(x;\boldsymbol{\mu}) = \dfrac{1}{2}\mathcal{A}(x,x;\boldsymbol{\mu}) - \langle \mathbf{F}(\boldsymbol{\mu}),x\rangle \\ \text{subject to } \mathcal{B}(x,(\boldsymbol{\phi},\xi);\boldsymbol{\mu}) = \langle \mathbf{G}(\boldsymbol{\mu}),(\boldsymbol{\phi},\xi)\rangle \qquad \forall (\boldsymbol{\phi},\xi) \in Q. \end{cases} \tag{4.2.6}$$

The assumptions made over the linear and the bilinear forms of the original problem (4.2.1) fulfill the hypotheses of Brezzi's Theorem 1.3 and Theorem 1.4: this guarantees the equivalence between the problem (4.2.6) and the following saddle-point formulation: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(x(\boldsymbol{\mu}),(\mathbf{w}(\boldsymbol{\mu}),q(\boldsymbol{\mu}))) \in X \times Q$ such that

$$\begin{cases} \mathcal{A}(x(\boldsymbol{\mu}),\zeta;\boldsymbol{\mu}) + \mathcal{B}(\zeta,(\mathbf{w}(\boldsymbol{\mu}),q(\boldsymbol{\mu}));\boldsymbol{\mu}) = \langle \mathbf{F}(\boldsymbol{\mu}),\zeta\rangle & \forall \zeta \in X, \\ \mathcal{B}(x(\boldsymbol{\mu}),(\boldsymbol{\phi},\xi);\boldsymbol{\mu}) = \langle \mathbf{G}(\boldsymbol{\mu}),(\boldsymbol{\phi},\xi)\rangle & \forall (\boldsymbol{\phi},\xi) \in Q. \end{cases} \tag{4.2.7}$$

The bilinear forms $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ and $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ satisfy the following properties:

1. the bilinear form $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ is symmetric and non-negative over the space X;

2. the bilinear form $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ is continuous over $X \times X$: i.e.

$$\gamma_{\mathcal{A}}(\boldsymbol{\mu}) = \sup_{x\in X}\sup_{\zeta\in X} \frac{\mathcal{A}(x,\zeta;\boldsymbol{\mu})}{\|x\|_X \|\zeta\|_X} < +\infty, \qquad \forall \boldsymbol{\mu} \in \mathcal{P};$$

3. let us define the space

$$X_0 = \{\zeta \in X \ : \ \mathcal{B}(\zeta, (\boldsymbol{\phi}, \xi), \boldsymbol{\mu}) = 0, \ \ \forall (\boldsymbol{\phi}, \xi) \in Q\} \subset X.$$

The bilinear form $\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ is coercive over X_0, i.e. there exists a constant $\alpha_0 > 0$ such that

$$\alpha(\boldsymbol{\mu}) = \inf_{x \in X_0} \frac{\mathcal{A}(x, x; \boldsymbol{\mu})}{\|x\|_X^2} \geq \alpha_0, \qquad \forall \boldsymbol{\mu} \in \mathcal{P};$$

4. the bilinear form $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ is continuous over $X \times Q$: i.e.

$$\gamma_{\mathcal{B}}(\boldsymbol{\mu}) = \sup_{\zeta \in X} \sup_{(\boldsymbol{\phi}, \xi) \in Q} \frac{\mathcal{B}(\zeta, (\boldsymbol{\phi}, \xi); \boldsymbol{\mu})}{\|\zeta\|_X \|(\boldsymbol{\phi}, \xi)\|_Q} < +\infty, \qquad \forall \boldsymbol{\mu} \in \mathcal{P};$$

5. the bilinear form $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ verifies the inf-sup condition over $X \times Q$, i.e. there exists a constant $\beta_0 > 0$ such that

$$\beta(\boldsymbol{\mu}) = \inf_{(\boldsymbol{\phi}, \xi) \in Q} \sup_{\zeta \in X} \frac{\mathcal{B}(\zeta, (\boldsymbol{\phi}, \xi); \boldsymbol{\mu})}{\|\zeta\|_X \|(\boldsymbol{\phi}, \xi)\|_Q} \geq \beta_0, \qquad \forall \boldsymbol{\mu} \in \mathcal{P}. \tag{4.2.8}$$

Naturally, the affine decomposition assumption is inherited from the affine assumptions (4.2.5). Indeed, for some $Q_{\mathcal{A}}, Q_{\mathcal{B}}, Q_{\mathbf{G}}, Q_{\mathbf{F}}$ it holds:

$$\begin{aligned}
\mathcal{A}(x, \zeta; \boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathcal{A}}} \Theta_{\mathcal{A}}^q(\boldsymbol{\mu}) \mathcal{A}^q(x, \zeta), \qquad \mathcal{B}(x, (\boldsymbol{\phi}, \xi); \boldsymbol{\mu}) = \sum_{q=1}^{Q_{\mathcal{B}}} \Theta_{\mathcal{B}}^q(\boldsymbol{\mu}) \mathcal{B}^q(x, (\boldsymbol{\phi}, \xi)), \\
\langle \mathbf{G}(\boldsymbol{\mu}), (\boldsymbol{\phi}, \xi) \rangle &= \sum_{q=1}^{Q_{\mathbf{G}}} \Theta_{\mathbf{G}}^q(\boldsymbol{\mu}) \langle \mathbf{G}^q, w \rangle, \qquad \langle \mathbf{F}(\boldsymbol{\mu}), \zeta \rangle = \sum_{q=1}^{Q_{\mathbf{F}}} \Theta_{\mathbf{F}}^q(\boldsymbol{\mu}) \langle \mathbf{F}^q, \zeta \rangle,
\end{aligned} \tag{4.2.9}$$

where $\Theta_{\mathcal{A}}, \Theta_{\mathcal{B}}, \Theta_{\mathbf{G}}, \Theta_{\mathbf{F}}$ are smooth $\boldsymbol{\mu}-$dependent functions and $\mathcal{A}^q(\cdot, \cdot), \mathcal{B}^q(\cdot, \cdot), \mathbf{G}^q, \mathbf{F}^q$ are $\boldsymbol{\mu}-$independent bilinear and linear forms. As we underlined among this chapter and among the previous one, these assumptions are fundamental in order to apply efficiently RB methods.

4.2.2 Full Order Approximation

In this subsection we will provide a Finite Element approximation of the saddle-point problem (4.2.7). In an analogy with the theory discussed in subsection 4.1.2, let $\{\mathcal{T}^{\mathcal{N}}\}$ be a triangulation of the domain Ω. Let us define the following space

$$X_{\mathcal{N}}^r = \{v^{\mathcal{N}} \in C^0(\overline{\Omega}) \ : \ v^{\mathcal{N}}|_K \in \mathbb{P}_r, \ \ \forall K \in \mathcal{T}^{\mathcal{N}}\}.$$

where $\mathbb{P}_r$ represents the space of polynomials of degree at most equal to r. Now we can define $Y^{\mathcal{N}} = Y \cap X_{\mathcal{N}}^r$, $Q^{\mathcal{N}} = Y^{\mathcal{N}}$ and $U^{\mathcal{N}} = U \cap X_{\mathcal{N}}^r$. By construction it holds that $Y^{\mathcal{N}} \subset Y$, $U^{\mathcal{N}} \subset U$ and $X^{\mathcal{N}} = Y^{\mathcal{N}} \times U^{\mathcal{N}} \subset X$ and $Q^{\mathcal{N}} \subset Q$. Moreover, let us suppose that $Y^{\mathcal{N}} = Q^{\mathcal{N}}$. Naturally, as before, the $\mathcal{N}$ denotes the dimension of the product space $X^{\mathcal{N}} \times Q^{\mathcal{N}}$, in other words $\mathcal{N} = \mathcal{N}_X + \mathcal{N}_Q$, where $\mathcal{N}_X = \mathcal{N}_Y + \mathcal{N}_U$. The *truth* Galerkin approximation of the problem (4.2.7) reads as: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(x^{\mathcal{N}}(\boldsymbol{\mu}), (\mathbf{w}^{\mathcal{N}}(\boldsymbol{\mu}), q^{\mathcal{N}}(\boldsymbol{\mu}))) \in X^{\mathcal{N}} \times Q^{\mathcal{N}}$ such that

$$\begin{cases}
\mathcal{A}(x^{\mathcal{N}}(\boldsymbol{\mu}), \zeta^{\mathcal{N}}; \boldsymbol{\mu}) + \mathcal{B}(\zeta^{\mathcal{N}}, (\mathbf{w}^{\mathcal{N}}(\boldsymbol{\mu}), q^{\mathcal{N}}(\boldsymbol{\mu})); \boldsymbol{\mu}) = \langle \mathbf{F}(\boldsymbol{\mu}), \zeta^{\mathcal{N}} \rangle & \forall \zeta^{\mathcal{N}} \in X^{\mathcal{N}}, \\
\mathcal{B}(x^{\mathcal{N}}(\boldsymbol{\mu}), (\boldsymbol{\phi}^{\mathcal{N}}, \xi^{\mathcal{N}}); \boldsymbol{\mu}) = \langle \mathbf{G}(\boldsymbol{\mu}), (\boldsymbol{\phi}^{\mathcal{N}}, \xi^{\mathcal{N}}) \rangle & \forall (\boldsymbol{\phi}^{\mathcal{N}}, \xi^{\mathcal{N}}) \in Q^{\mathcal{N}}.
\end{cases} \tag{4.2.10}$$

As we did in Lemma 2.4, assuming that $Q^{\mathcal{N}} = Y^{\mathcal{N}}$, the bilinear forms $\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ and $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ verify the following properties:

1. The bilinear form $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ is coercive over $X^{\mathcal{N}} \times X^{\mathcal{N}}$, i.e.

$$\gamma_{\mathcal{A}}^{\mathcal{N}}(\boldsymbol{\mu}) = \sup_{x^{\mathcal{N}} \in X^{\mathcal{N}}} \sup_{\zeta^{\mathcal{N}} \in X^{\mathcal{N}}} \frac{\mathcal{A}(x^{\mathcal{N}}, \zeta^{\mathcal{N}}; \boldsymbol{\mu})}{\|x^{\mathcal{N}}\|_X \|\zeta^{\mathcal{N}}\|_X} < +\infty, \qquad \forall \boldsymbol{\mu} \in \mathcal{P};$$

2. let us define the set

$$X_0^{\mathcal{N}} = \{\zeta \in X^{\mathcal{N}} \ : \ \mathcal{B}(\zeta^{\mathcal{N}}, (\boldsymbol{\phi}^{\mathcal{N}} \xi^{\mathcal{N}}); \boldsymbol{\mu}),\ \forall (\boldsymbol{\phi}^{\mathcal{N}}, \xi^{\mathcal{N}}) \in Q^{\mathcal{N}}\}.$$

   The bilinear form $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ is coercive over $X_0^{\mathcal{N}}$, in other words:

$$\alpha^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{x^{\mathcal{N}} \in X_0^{\mathcal{N}}} \frac{\mathcal{A}(x^{\mathcal{N}}, x^{\mathcal{N}}; \boldsymbol{\mu})}{\|x^{\mathcal{N}}\|_X^2} \geq \alpha(\boldsymbol{\mu}) \geq \alpha_0 > 0 \qquad \forall \boldsymbol{\mu} \in \mathcal{P};$$

3. the bilinear form $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ is continuous over $X^{\mathcal{N}} \times Q^{\mathcal{N}}$, that is:

$$\gamma_{\mathcal{B}}^{\mathcal{N}}(\boldsymbol{\mu}) = \sup_{\zeta^{\mathcal{N}} \in X^{\mathcal{N}}} \sup_{(\boldsymbol{\phi}^{\mathcal{N}}, \xi^{\mathcal{N}}) \in Q^{\mathcal{N}}} \frac{\mathcal{B}(\zeta^{\mathcal{N}}, (\boldsymbol{\phi}^{\mathcal{N}}, \xi^{\mathcal{N}}); \boldsymbol{\mu})}{\|\zeta^{\mathcal{N}}\|_X \|(\boldsymbol{\phi}^{\mathcal{N}}, \xi^{\mathcal{N}})\|_Q} < +\infty, \qquad \forall \boldsymbol{\mu} \in \mathcal{P};$$

4. the bilinear form $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ verifies the inf-sup stability over $X^{\mathcal{N}} \times Q^{\mathcal{N}}$, i.e. there exists a constant β_0 such that

$$\beta^{\mathcal{N}}(\boldsymbol{\mu}) = \inf_{(\boldsymbol{\phi}^{\mathcal{N}}, \xi^{\mathcal{N}}) \in Q^{\mathcal{N}}} \sup_{\zeta^{\mathcal{N}} \in X^{\mathcal{N}}} \frac{\mathcal{B}(\zeta^{\mathcal{N}}, (\boldsymbol{\phi}^{\mathcal{N}}, \xi^{\mathcal{N}}); \boldsymbol{\mu})}{\|\zeta^{\mathcal{N}}\|_X \|(\boldsymbol{\phi}^{\mathcal{N}}, \xi^{\mathcal{N}})\|_Q} \geq \beta_0, \qquad \forall \boldsymbol{\mu} \in \mathcal{P}. \tag{4.2.11}$$

In particular, as we did in Lemma 2.4, it holds that $\beta^{\mathcal{N}}(\boldsymbol{\mu}) \geq \tilde{\beta}^{\mathcal{N}}(\boldsymbol{\mu})$, where $\tilde{\beta}^{\mathcal{N}}(\boldsymbol{\mu})$ is the Babuška constant of the bilinear form $\mathbf{A}(\cdot,\cdot;\boldsymbol{\mu})$. For these reasons, the Finite Element (FE) approximation (4.2.10) is well-posed, thanks to Proposition 2.4.

Remark 4.2.1. Let us recall the algebraic formulation of the linear system associated to a linear quadratic OCP($\boldsymbol{\mu}$) governed by Stokes equations. Following the same steps made in Chapter 2, the system reads:

$$\begin{pmatrix} A(\boldsymbol{\mu}) & B^T(\boldsymbol{\mu}) \\ B(\boldsymbol{\mu}) & 0 \end{pmatrix} \begin{pmatrix} \mathbf{X}^{\mathcal{N}}(\boldsymbol{\mu}) \\ \mathbf{W}^{\mathcal{N}}(\boldsymbol{\mu}) \end{pmatrix} = \begin{pmatrix} \mathbf{F}(\boldsymbol{\mu}) \\ \mathbf{G}(\boldsymbol{\mu}) \end{pmatrix}, \tag{4.2.12}$$

where

$$\mathbf{X} = \begin{pmatrix} \mathbf{V} \\ \mathbf{U} \end{pmatrix}$$

denotes state and control aggregated variables, whereas $\mathbf{W}$ is the adjoint variable related to the velocity and the pressure variables. Notice that the affine decomposition 4.2.9 is inherited from the original space, in other words it holds:

$$\begin{aligned} A(\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathcal{A}}} \Theta_{\mathcal{A}}^q A^q, & B(\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathcal{B}}} \Theta_{\mathcal{B}}^q B^q, \\ \mathbf{G}(\boldsymbol{\mu}) &= \sum_{q=1}^{Q_G} \Theta_{\mathbf{G}}^q \mathbf{G}^q, & \mathbf{F}(\boldsymbol{\mu}) &= \sum_{q=1}^{Q_{\mathbf{F}}} \Theta_F^q \mathbf{F}^q. \end{aligned} \tag{4.2.13}$$

4.2.3 Reduced Basis Approximation

As underlined in the previous subsection, the aim of RB methods is to use a new approximated space that has a basis composed by well-chosen solutions $(x^{\mathcal{N}}(\boldsymbol{\mu}), (\mathbf{w}^{\mathcal{N}}(\boldsymbol{\mu}), q^{\mathcal{N}}(\boldsymbol{\mu})))$ of the problem (4.2.10). We already know that the discrete solution has a smooth dependence on $\boldsymbol{\mu}$. This allows the parametric manifold $\mathcal{M}$ to be smooth and to be approximated by *snapshot* solutions of (4.2.10).
In the specific case of OCP($\boldsymbol{\mu}$) governed by Stokes equation, one has to take into account a *nested saddle-point* problem: we have a first saddle-point structure given by the state equations, and the other one given by the linear quadratic optimization. We will proceed in the following way:

1. the well-posedness of the Stokes problem will be guaranteed by an enriched velocity space with *supremizer solutions* (i.e. see [62, 65, 53, 63]),
2. then we will focus on the stability of the whole OCP($\boldsymbol{\mu}$) problem, in other words, we have to ensure the fulfillment of the RB inf-sup condition on the bilinear form $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$. As we did for the OCP($\boldsymbol{\mu}$) governed by elliptic equations, we will exploit *aggregated spaces* for state and adjoint variables, under the assumption $Y_N = Q_N$.

RB Stability of the State Equation

Let us analyse the RB stability for the state equations. For a given $N \in \{1, \dots, N_{max}\}$, let us consider the sample set $S_N = \{\boldsymbol{\mu}^1, \dots, \boldsymbol{\mu}^N\}$ and the relative Finite Element solutions $\{(\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}^n), p^{\mathcal{N}}(\boldsymbol{\mu}^n)),\ n = 1, \dots, N\}$. A first (naive) reduced space for pressure can be defined:

$$P_N = \text{span}\ \{p^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}. \tag{4.2.14}$$

To ensure the stability of the reduced state equation we have to follow the strategy of the pressure supremizer operator $T_p^{\boldsymbol{\mu}} : P^{\mathcal{N}} \to V^{\mathcal{N}}$ defined as follows:

$$(T_p^{\boldsymbol{\mu}} s, \boldsymbol{\phi})_V = b(\boldsymbol{\phi}, s; \boldsymbol{\mu}), \qquad \forall \boldsymbol{\phi} \in V^{\mathcal{N}}. \tag{4.2.15}$$

We are now ready to build a reduced enriched velocity space:

$$V_N^{\boldsymbol{\mu}} = \text{span}\{\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}^n),\ T_p^{\boldsymbol{\mu}} p^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}.$$

By the reduced Galerkin projection onto $V_N^{\boldsymbol{\mu}} \times P_N$, we can formulate a reduce state equation. The new problem reads: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(\mathbf{v}_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in V_N^{\boldsymbol{\mu}} \times P_N$ such that

$$\begin{cases} a(\mathbf{v}_N(\boldsymbol{\mu}), \boldsymbol{\phi}; \boldsymbol{\mu}) + b(\boldsymbol{\phi}, p_N(\boldsymbol{\mu}); \boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), \boldsymbol{\phi} \rangle & \forall \boldsymbol{\phi} \in V_N^{\boldsymbol{\mu}}, \\ b(\mathbf{v}_N(\boldsymbol{\mu}), \pi; \boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), \pi \rangle & \forall \pi \in P_N. \end{cases} \tag{4.2.16}$$

Thanks to the inclusions $V_N^{\boldsymbol{\mu}} \subset V^{\mathcal{N}}$, $P_N \subset P^{\mathcal{N}}$, the bilinear form $a(\cdot,\cdot;\boldsymbol{\mu})$ remains continuous over $V_N^{\boldsymbol{\mu}} \times V_N^{\boldsymbol{\mu}}$ and coercive over $V_N^{\boldsymbol{\mu}}$, whereas the bilinear form $b(\cdot,\cdot;\boldsymbol{\mu})$ is continuous over $V_N^{\boldsymbol{\mu}} \times P_N$. Furthermore, using the enriched velocity space, the bilinear form $b(\cdot,\cdot;\boldsymbol{\mu})$ verifies the RB inf-sup condition: in [65] is proved that

$$\beta_N(\boldsymbol{\mu}) = \inf_{\pi \in P_N} \sup_{\boldsymbol{\phi} \in V_N^{\boldsymbol{\mu}}} \frac{b(\boldsymbol{\phi}, \pi; \boldsymbol{\mu})}{\|\pi\|_P \|\boldsymbol{\phi}\|_V} \geq \beta^{\mathcal{N}}(\boldsymbol{\mu}) \geq \beta_0 > 0, \qquad \forall \boldsymbol{\mu} \in \mathcal{P}.$$

For these reasons the reduced problem (4.2.16) is well-posed since they verify the hypotheses of the Brezzi's Theorem 2.1.

RB stability of the global control problem

Once proved the stability of the saddle-point state equation, we will focus on the stability of the OCP($\boldsymbol{\mu}$) saddle-point structure. In order to reach our goal, we must exploit new approximated spaces. Let us define the aggregated spaces for the pressure variable

$$P_N = \text{span}\{p^{\mathcal{N}}(\boldsymbol{\mu}^n), q^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}, \tag{4.2.17}$$

and for the velocity variable

$$V_N^{\boldsymbol{\mu}} = \text{span}\ \{\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}^n), T_p^{\boldsymbol{\mu}} p^{\mathcal{N}}(\boldsymbol{\mu}^n), \mathbf{w}^{\mathcal{N}}(\boldsymbol{\mu}^n), T_p^{\boldsymbol{\mu}} q^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}. \tag{4.2.18}$$

Finally, the control space is defined as follows:

$$U_N = \text{span}\ \{\mathbf{u}^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1, \dots, N\}. \tag{4.2.19}$$

Let us consider the following *aggregated* space for state and adjoint variables

$$Z_N = V_N^{\boldsymbol{\mu}} \times P_N,$$

this space will be use both for the state space and the adjoint space, i.e. $Q_N = Y_N = Z_N$. Considering the product space $X_N = Y_N \times U_N$, the new problem formulation reads: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(x(\boldsymbol{\mu})_N, (\mathbf{w}_N(\boldsymbol{\mu}), q_N(\boldsymbol{\mu}))) \in X_N \times Q_N$ such that

$$\begin{cases} \mathcal{A}(x_N(\boldsymbol{\mu}), \zeta_N; \boldsymbol{\mu}) + \mathcal{B}(\zeta_N, (\mathbf{w}_N(\boldsymbol{\mu}), q_N(\boldsymbol{\mu})); \boldsymbol{\mu}) = \langle \mathbf{F}(\boldsymbol{\mu}), \zeta_N \rangle & \forall \zeta_N \in X_N, \\ \mathcal{B}(x_N(\boldsymbol{\mu}), (\boldsymbol{\phi}_N, \xi_N); \boldsymbol{\mu}) = \langle \mathbf{G}(\boldsymbol{\mu}), (\boldsymbol{\phi}_N, \xi_N) \rangle & \forall (\boldsymbol{\phi}_N, \xi_N) \in Q_N. \end{cases} \tag{4.2.20}$$

Now we have to guarantee the well-posedness of the RB approximation. The continuity property of the bilinear forms $\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ and $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ is automatically inherited from the Finite Element spaces. In particular we want to prove the fulfillment of the coercivity of the bilinear form $\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ over the space

$$X_0^N = \{\zeta_N \in X_N\ :\ \mathcal{B}(\zeta_N, (\boldsymbol{\phi}_N, \xi_N); \boldsymbol{\mu}),\ \ \forall (\boldsymbol{\phi}_N, \xi_N) \in Q_N\},$$

and the fulfillment of the RB inf-sup condition of the bilinear form $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$: in other words we have to show that there exists $\beta_0 > 0$ such that

$$\beta_N(\boldsymbol{\mu}) = \inf_{(\boldsymbol{\phi}_N, \xi_N) \in Q_N} \sup_{\zeta_N \in X_N} \frac{\mathcal{B}(\zeta_N, (\boldsymbol{\phi}_N, \xi_N); \boldsymbol{\mu})}{\|\zeta_N\|_X \|(\boldsymbol{\phi}_N, \xi_N)\|_Q} \geq \beta_0, \qquad \forall \boldsymbol{\mu} \in \mathcal{P}. \tag{4.2.21}$$

Lemma 4.2. *The bilinear form $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ verify the inf-sup condition (4.2.21).*

Proof. First of all, thanks to the enrichment of the velocity space with the supremizer solutions and the fact that $Y_N = Q_N$, we are able to prove that there exists a constant $\tilde{\beta}_N^0$ such that

$$\begin{aligned} \tilde{\beta}_N(\boldsymbol{\mu}) &= \inf_{(\mathbf{v}_N, p_N) \in Y_N} \sup_{(\boldsymbol{\phi}_N, \xi_N) \in Q_N} \frac{\mathbf{A}((\mathbf{v}_N, p_N), (\boldsymbol{\phi}_N, \xi_N); \boldsymbol{\mu})}{\|(\mathbf{v}_N, p_N)\|_Y \|(\boldsymbol{\phi}_N, \xi_N)\|_Q} \\ &= \inf_{(\boldsymbol{\phi}_N, \xi_N) \in Q_N} \sup_{(\mathbf{v}_N, p_N) \in Y_N} \frac{\mathbf{A}((\mathbf{v}_N, p_N), (\boldsymbol{\phi}_N, \xi_N); \boldsymbol{\mu})}{\|(\mathbf{v}_N, p_N)\|_Y \|(\boldsymbol{\phi}_N, \xi_N)\|_Q} \geq \tilde{\beta}_N^0. \end{aligned}$$

Now we are going to exploit the weak coercivity of the bilinear form $\mathbf{A}(\cdot,\cdot;\boldsymbol{\mu})$ to prove the inf-sup stability of $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$. Indeed:

$$\begin{aligned}\sup_{x_N\in X_N}\frac{\mathcal{B}(x_N,(\boldsymbol{\phi}_N,\xi_N);\boldsymbol{\mu})}{\|x_N\|_X} &= \sup_{\substack{((\mathbf{v}_N,p_N),\mathbf{u})\in Y_N\times U_N,\\((\mathbf{v}_N,p_N),\mathbf{u})\neq 0}}\frac{\mathbf{A}((\mathbf{v}_N,p_N),(\boldsymbol{\phi}_N,\xi_N);\boldsymbol{\mu})-c(\mathbf{u},\boldsymbol{\phi};\boldsymbol{\mu})}{\sqrt{\|(\mathbf{v}_N,p_N)\|_Y^2\|\mathbf{u}\|_U^2}}\\ &\underset{\mathbf{u}=0}{\geq}\sup_{\substack{((\mathbf{v}_N,p_N),0)\in Y_N\times U_N,\\((\mathbf{v}_N,p_N),0)\neq 0}}\frac{\mathbf{A}((\mathbf{v}_N,p_N),(\boldsymbol{\phi}_N,\xi_N);\boldsymbol{\mu})}{\|(\mathbf{v}_N,p_N)\|_Y}\\ &\geq \tilde{\beta}_N^0\|(\boldsymbol{\phi}_N,\xi_N)\|_Y=\tilde{\beta}_N^0\|(\boldsymbol{\phi}_N,\xi_N)\|_Q.\end{aligned}$$

Proposition 4.2. *The RB saddle-point problem (4.2.20) has an unique solution for all* $\boldsymbol{\mu}\in\mathcal{P}$.

Proof. We have to verify the hypotheses of the Brezzi's Theorem 2.1. We have already underlined that the continuity properties of linear and bilinear forms are naturally inherited from the Finite Element spaces. In Lemma 4.2, we have already proved the RB inf-sup condition over the bilinear form $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$. The coercivity of the bilinear form $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ can be proved as we did in Lemma 2.3 and in Lemma 2.4, under the assumption of $Q_N=Y_N$ and the supremizer enrichment of the velocity space.

Remark 4.2.2. Notice that, by the definition of the supremizer $T_p^{\boldsymbol{\mu}}$, the reduced velocity space $V_N^{\boldsymbol{\mu}}$ (and therefore, all the spaces deriving from it, like Y_N and Q_N) depends on $\boldsymbol{\mu}$. The supremizer enrichment adds $2Q_bN$ basis functions to the original dimension of the reduced velocity space and leads to a less efficient application of RB methods. To avoid this inconvenient we follow the strategy presented in [53, 26, 63]: we substitute the space $V_N^{\boldsymbol{\mu}}$ with the following one

$$V_N=\text{span}\,\{\mathbf{v}^{\mathcal{N}}(\boldsymbol{\mu}^n),T_p^{\boldsymbol{\mu}^n}p^{\mathcal{N}}(\boldsymbol{\mu}^n),\mathbf{w}^{\mathcal{N}}(\boldsymbol{\mu}^n),T_p^{\boldsymbol{\mu}^n}q^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n=1,\dots,N\}. \tag{4.2.22}$$

Adding only $2N$ supremizer snapshots allow to have an efficient decoupling of the Offline and the Online stages. The stability of this technique is numerically demonstrated in [26, 63]. Thanks to this new space formulation, the dimension of the state and the adjoint space is $6N$, whereas the control space has dimension N (see [1] as a reference).

4.2.4 Algebraic Formulation of The Enriched RB Approximation

Let us introduce the algebraic formulation associated to the enriched reduced problem (4.2.20). Let us consider V_N as the space introduced in Remark 4.2.2. The aggregated space for state and adjoint variable is

$$Z_N=V_N\times P_N,\qquad Y_N=Q_N=Z_N,$$

with $\{z_j\}_{j=1}^{6N}$ as basis functions. The control space U_N is generated by $\{\lambda_j\}_{j=1}^N$. Finally, we construct $X_N=Y_N\times U_N=\text{span}\,\{\sigma_j,\ j=1,\dots,7N\}$, where the functions σ_n are defined as follows:

$$\sigma_j=\begin{cases}(z_j,0) & j=1,\dots,6N,\\ (0,\lambda_{j-6N}) & j=6N+1,\dots,7N.\end{cases}$$

Now we can express state, control and adjoint solution of (4.2.20)

$$x_N(\boldsymbol{\mu}) = \sum_{j=1}^{7N} X_{N_j}(\boldsymbol{\mu})\sigma_j, \qquad (\mathbf{w}_N(\boldsymbol{\mu}), q_N(\boldsymbol{\mu})) = \sum_{j=1}^{6N} W_{N_j}(\boldsymbol{\mu})z_j.$$

Then, the linear system associated to the RB problem (4.2.20) has the following explicit form

$$\begin{cases} \displaystyle\sum_{j=1}^{7N}\sum_{q=1}^{Q_{A_N}} \Theta^q_{A_N}(\boldsymbol{\mu})A^q_{N_{ij}}X_{N_j}(\boldsymbol{\mu}) + \sum_{k=1}^{6N}\sum_{q=1}^{Q_{B_N}} \Theta^q_{B_N}(\boldsymbol{\mu})B^q_{N_{ki}}W_{N_k}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{F_N}} \Theta^q_{F_N}(\boldsymbol{\mu})F^q_{N_i}, & 1 \le i \le 7N \\ \displaystyle\sum_{j=1}^{7N}\sum_{q=1}^{Q_{B_N}} \Theta^q_{B_N}(\boldsymbol{\mu})B^q_{N_{ij}}X_{N_j}(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{G_N}} \Theta^q_{G_N}(\boldsymbol{\mu})G^q_{N_i}, & 1 \le i \le 6N \end{cases} \tag{4.2.23}$$

where the matrices linked to bilinear and linear forms are given by

$$\begin{aligned} A^q_{N_{ij}} &= \mathcal{A}(\sigma_i, \sigma_j) && 1 \le i, j \le 7N \\ B^q_{N_{ij}} &= \mathcal{B}(\sigma_j, z_i) && 1 \le j \le 7N,\ 1 \le i \le 6N \\ F^q_{N_i} &= \langle F^q, \sigma_i \rangle && 1 \le i \le 7N \\ G^q_{N_i} &= \langle G^q, z_i \rangle && 1 \le i \le 6N. \end{aligned}$$

Denoting with

$$A_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{A_N}} \Theta^q_{A_N}(\boldsymbol{\mu})A^q_N, \qquad B_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{B_N}} \Theta^q_{B_N}(\boldsymbol{\mu})B^q_N,$$

$$\mathbf{G}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{\mathbf{G}_N}} \Theta^q_{\mathbf{G}_N}(\boldsymbol{\mu})G^q_N, \qquad \mathbf{F}_N(\boldsymbol{\mu}) = \sum_{q=1}^{Q_{\mathbf{F}_N}} \Theta^q_{\mathbf{F}_N}(\boldsymbol{\mu})F^q_N,$$

the linear system (4.2.23) can be written as

$$\begin{pmatrix} A_N(\boldsymbol{\mu}) & B_N^T(\boldsymbol{\mu}) \\ B_N(\boldsymbol{\mu}) & 0 \end{pmatrix} \begin{pmatrix} \mathbf{X}_N(\boldsymbol{\mu}) \\ \mathbf{W}_N(\boldsymbol{\mu}) \end{pmatrix} = \begin{pmatrix} \mathbf{F}_N(\boldsymbol{\mu}) \\ \mathbf{G}_N(\boldsymbol{\mu}) \end{pmatrix}. \tag{4.2.24}$$

This matrix formulation is still symmetric and the dimension of the system is $13N \times 13N$.

4.3 Numerical Results

In this section we are going to present three numerical examples to test the performance of RB methods on parametric optimal control problems. The first two tests are the parametric version of the examples proposed in Chapter 2. The last example is the reduced parametric adaptation of an advection-diffusion OCP($\boldsymbol{\mu}$) proposed in [57, subsection 17.11.2]. For the simulations we have used RBniCS library (for information visit the following website: *http://mathlab.sissa.it/rbnics*), and a *one-shot* approach to solve linear systems. To build the reduced basis we tried two strategies:

1. perform single POD Galerkin for all the solution components, i.e. monolithic approach,
2. perform a POD Galerkin for each solution component, i.e. partitioned approach, exploiting aggregated spaces.

The results given by the monolithic and the partitioned approach have been compared.

4.3.1 OCP(μ) Governed by Laplace Equation

The first test problem is based on the example proposed in subsection 2.3.1. Let us consider $\Omega = (0,1)^2$. The parametrized version of the distributed control problem (2.3.1) reads as follows:

$$\begin{aligned} \min_{(y,u)\in Y\times U} J(y,u) &= \frac{1}{2}\int_\Omega (y-y_d)^2\,d\Omega + \frac{\alpha}{2}\int_\Omega u^2\,d\Omega, \\ \text{such that } &\begin{cases} -\mu\Delta y = u+f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega. \end{cases} \end{aligned} \tag{4.3.1}$$

Let us briefly recall the main features of the problem: $U = L^2(\Omega)$, $Y = H_0^1(\Omega)$, $f = 0$ and $Q = Y$. The desired state $y_d = 10x_1(1-x_1)x_2(1-x_2)$ is given. The parameter μ represents the diffusivity constant and the parameter space is $\mathcal{P} = [0.5, 1]$. The weak formulation of the state equation reads:

$$a(y,q;\mu) = c(u,q) \qquad \forall q \in Q,$$

where the bilinear form $a : Y \times Q \to \mathbb{R}$ and $c : U \times Q \to \mathbb{R}$ are the defined as follows:

$$a(z,q;\mu) = \mu\int_\Omega \nabla z\cdot\nabla q\,d\Omega, \qquad c(v,q) = \int_\Omega vq\,d\Omega.$$

Furthermore, the bilinear forms $m : Y \times Y \to \mathbb{R}$ and $n : U \times U \to \mathbb{R}$ are given by

$$m(y,z) = \int_\Omega yz\,d\Omega, \qquad n(u,v) = \int_\Omega uv\,d\Omega.$$

Let us recast the problem in a saddle-point formulation: let $X = Y \times U$ be the product space of state and the control spaces. Let us consider two elements of X, i.e. $x = (y,u), w = (z,v)$ and $p, q \in Q$ and define the bilinear forms

$$\begin{aligned} \mathcal{A}(x,w) &= m(y,z) + \alpha n(u,v), \\ \mathcal{B}(w,q;\mu) &= a(z,q;\mu) - c(v,q), \end{aligned}$$

and the linear functional:

$$\langle F, w\rangle = \int_\Omega y_d z\,d\Omega.$$

Let us underline the affine structure (very intuitive in this simple case): with $Q_\mathcal{A} = Q_F = 1$ and $Q_\mathcal{B} = 2$ the affine decomposition of the problem is given by

$$\begin{aligned} &\Theta_\mathcal{A}^1 = 1 && \mathcal{A}^1(x,w) = \mathcal{A}(x,w) \\ &\Theta_F^1 = 1 && \langle F^1, w\rangle = \langle F^1, w\rangle \\ &\Theta_\mathcal{B}^1 = \mu && \mathcal{B}^1(w,q) = \int_\Omega \nabla z\cdot\nabla q\,d\Omega \\ &\Theta_\mathcal{B}^2 = -1 && \mathcal{B}^2(w,q) = \int_\Omega vq\,d\Omega. \end{aligned}$$

In the following we present a specific experiment. We used a POD algorithm with 20 basis functions, and a training set of 100 points. The parameter is chosen trough an uniform distribution. The penalization term is $\alpha = 10^{-5}$. In the following we are going to show the results given by the specific diffusivity parameter $\mu = 1$. In figure 4.3.1.1 full order state solution and reduced state solution are shown, with a pointwise error plot: notice that the

maximum value of the error is of the order of 10^{-8}. Analogously, in figure 4.3.1.2 we have a comparison between the full order control variable and the reduced control variable.

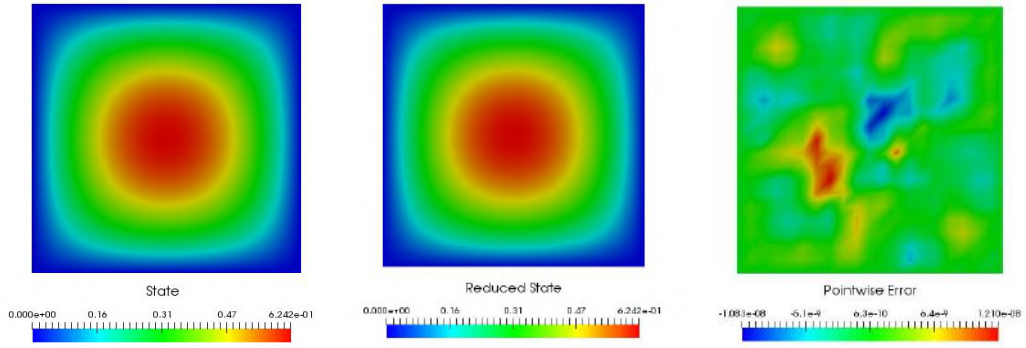

Figure 4.3.1.1: *Left*: full order optimal state; *center*: reduced optimal state, *right*: pointwise error.

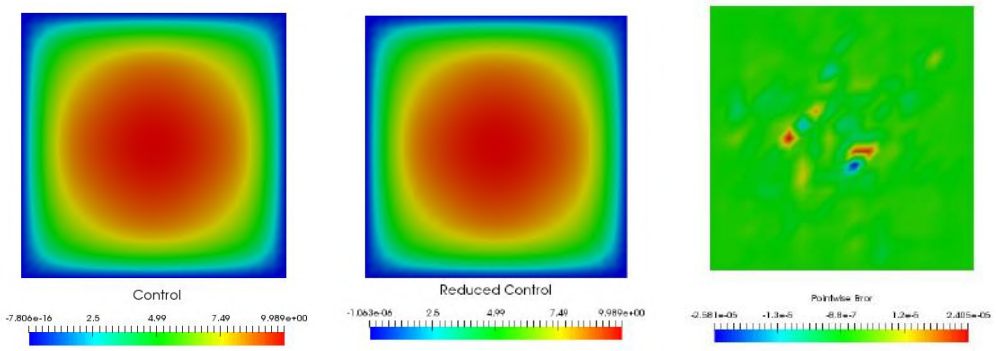

Figure 4.3.1.2: *Left*: full order control variable; *center*: reduced control variable, *right*: pointwise error.

Let us focus on the error analysis. In figure 4.3.1.3 we notice that, as we expect, increasing the reduced basis number, the error decreases for all the tree variables. In the bottom right plot a comparison between the monolithic error and the partitioned error of the reduced basis approximation with respect to the full order approximation[1] is presented: splitting the POD algorithm for the different variables and using the aggregated space formulation is more convenient and leads to better results.

[1]Let us call $W = Y \times U \times Q$. For $u \in W$, the monolithic error is given by an aggregated error of the type $\|u\|_W^2$, while the partitioned error is given by $\|u\|_W^2 = \|y\|_Y^2 + \|u\|_U^2 + \|q\|_Q^2$.

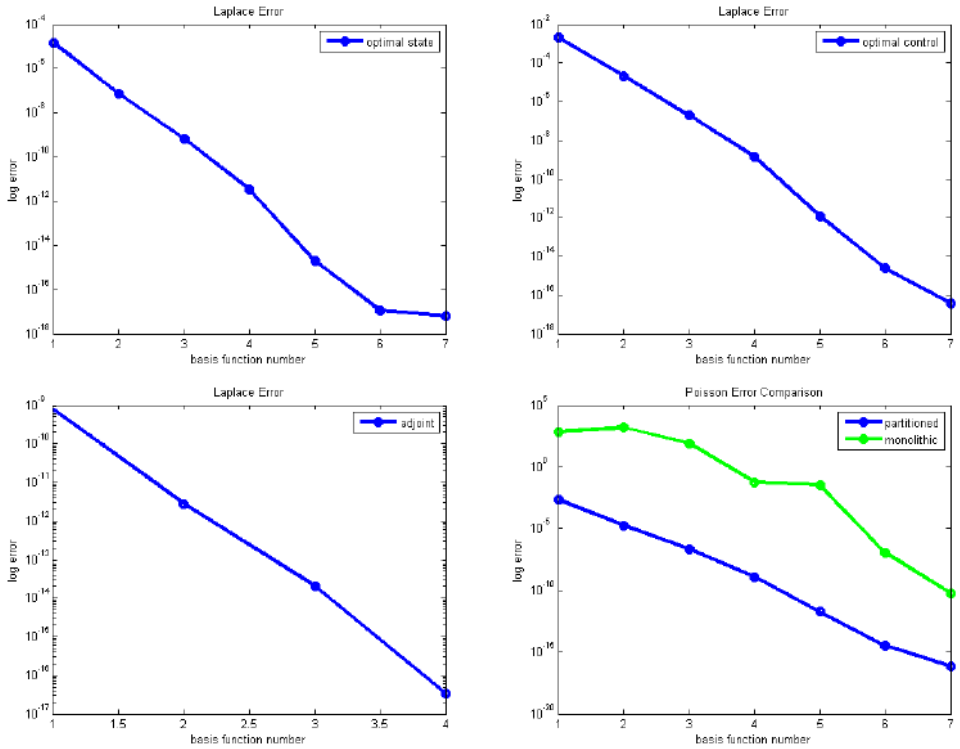

Figure 4.3.1.3: *Top left*: reduced optimal state error trend; *top right*: reduced control error trend. *bottom left*: reduced adjoint state error trend; *bottom right*: comparison between monolithic and partitioned approaches.

Table 4.1: Speed up analysis for Laplace problem

Basis Number	1	2	3	4	5	6	7	8	9	10
Speed up	62	62	61	61	60	60	59	59	58	57
Basis Number	11	12	13	14	15	16	17	18	19	20
Speed up	57	58	56	57	55	55	54	53	52	52

In the following a comparison between the dimensions of full order system and the reduced one is shown.

$\mathcal{N} \times \mathcal{N}$	1323×1323
$5N \times 5N$	100×100

As we underlined among the whole chapter, reduced basis methods are a good technique to let costly problems be solved in a small time with respect to the time necessary to a full order solving system. This can be easily noticed in Table 4.1: the reduced system dimension is drastically lower with respect the full order system dimension. The reduced space dimension is indicated with N, while the full order dimension with $\mathcal{N}$, as usual. To characterize the computational performances of the RB methods we used the *speed up* index: it measures how many reduced order systems can be afforded in the time needed for a full order system to be solved. Naturally, the time of the reduced system resolution increases if the basis number increases, but the advantages are remarkable even with a

large reduced space dimension. Let us indicate with J_t the cost functional associated to the full order solution and J_r the cost functional of the reduced problem: for the particular value $\mu = 1$ we have $J_t = J_r = 2.3286 \cdot 10^{-4}$.

4.3.2 OCP($\boldsymbol{\mu}$) Governed by Stokes Equation

The second test is based on the numerical example proposed 2.3.2. As spatial domain we consider $\Omega = (0,1)^2$. The control is distributed over Ω. The spaces are those specified in subsection 4.2.1: in this particular case, since we have homogeneous Dirichlet boundary conditions, $V = H_0^1(\Omega) \times H_0^1(\Omega)$. Let $Y = V \times P$ be the state space. Let us focus on two examples.

Affine Target Function

The parametrized version of the problem (2.3.2) reads as follows: for a given $\boldsymbol{\mu} = [\mu_1, \mu_2]$, find $(y(\boldsymbol{\mu}), u(\boldsymbol{\mu})) \in Y \times U$ solution of

$$\min_{(y,u)\in Y\times U} J(\mathbf{v}, p, \mathbf{u}) = \frac{1}{2}\int_\Omega |\mathbf{v} - \mu_2 \mathbf{v}_d|^2 \, d\Omega + \frac{\alpha}{2}\int_\Omega |\mathbf{u}|^2 \, d\Omega,$$
$$\text{such that } \begin{cases} -\mu_1 \Delta \mathbf{v} + \nabla p = \mathbf{u} & \text{in } \Omega, \\ \operatorname{div}(\mathbf{v}) = 0 & \text{in } \Omega, \\ \mathbf{v} = \mathbf{0} & \text{on } \partial\Omega, \end{cases} \tag{4.3.2}$$

where

$$\mathbf{v}_d = \left(\frac{\partial}{\partial x_2}(\varphi(x_1)\varphi(x_2)), -\frac{\partial}{\partial x_1}(\varphi(x_1)\varphi(x_2)) \right),$$

with $\varphi : (0,1) \to (0,1)$ is defined as $\varphi(z) = (1 - \cos(0.8\pi z))(1-z)^2$. The parameter μ_1 represents the diffusivity and μ_2 chances the intensity of the desired velocity field. The parameter space is $\mathcal{P} = [0.5, 1] \times [5, 10]$. In this particular example the weak formulation of the state equation is:

$$\begin{cases} a(\mathbf{v}, \boldsymbol{\phi}; \mu) + b(\boldsymbol{\phi}, p) = c(\mathbf{u}, \boldsymbol{\phi}), & \forall \boldsymbol{\phi} \in V, \\ b(\mathbf{v}, \xi) = 0, & \forall \xi \in P, \end{cases} \tag{4.3.3}$$

where the bilinear forms $a : V \times V \to \mathbb{R}$, $b : V \times P \to \mathbb{R}$ and $c : U \times V \to \mathbb{R}$ are given by:

$$a(\mathbf{v}, \boldsymbol{\phi}; \boldsymbol{\mu}) = \mu_1 \int_\Omega \nabla \mathbf{v} \cdot \nabla \boldsymbol{\phi} \, d\Omega,$$
$$b(\boldsymbol{\phi}, p) = -\int_\Omega \operatorname{div}(\boldsymbol{\phi}) p \, d\Omega,$$
$$c(\mathbf{u}, \boldsymbol{\phi}) = \int_\Omega \mathbf{u} \cdot \boldsymbol{\phi} d\Omega.$$

Now we want to recast the problem (4.3.2) into a saddle-point formulation. To reach this goal we build the bilinear forms $\mathbf{A} : Y \times Y \to \mathbb{R}$ and $\mathbf{C} : U \times Q \to \mathbb{R}$ as follows

$$\mathbf{A}((\mathbf{v}, p), (\boldsymbol{\phi}, \xi); \boldsymbol{\mu}) = a(\mathbf{v}, \boldsymbol{\phi}; \boldsymbol{\mu}) + b(\boldsymbol{\phi}, p) + b(\mathbf{v}, \xi),$$
$$\mathbf{C}(\mathbf{u}, \boldsymbol{\phi}) = c(\mathbf{u}, \boldsymbol{\phi}).$$

Let us consider the bilinear forms $m : Y \times Y \to \mathbb{R}$ ans $n : U \times U \to \mathbb{R}$ given by:

$$m((\mathbf{v}, p), (\boldsymbol{\psi}, \pi)) = \int_\Omega \mathbf{v}\boldsymbol{\psi} \, d\Omega$$
$$n(\mathbf{u}, \boldsymbol{\tau}) = \int_\Omega \mathbf{u} \cdot \boldsymbol{\tau} \, d\Omega.$$

To have a problem of the form presented in (4.2.7) let us define $X = Y \times U$ and let $x = ((\mathbf{v}, p), \mathbf{u})$ and $\zeta = ((\boldsymbol{\psi}, \pi), \boldsymbol{\tau})$ be two elements of X, while $(\boldsymbol{\phi}, \xi) \in Q$. We have to consider the following bilinear forms $\mathcal{A} : X \times X \to \mathbb{R}$ and $\mathcal{B} : X \times Q \to \mathbb{R}$ given by

$$\mathcal{A}(x, \zeta) = m((\mathbf{v}, p), (\boldsymbol{\psi}, \pi)) + \alpha n(\mathbf{u}, \boldsymbol{\tau}),$$
$$\mathcal{B}(x, (\boldsymbol{\phi}, \xi); \boldsymbol{\mu}) = \mathbf{A}((\mathbf{v}, p), (\boldsymbol{\phi}, \xi); \boldsymbol{\mu}) - \mathbf{C}(\mathbf{u}, \boldsymbol{\phi}).$$

and the linear form $F(\boldsymbol{\mu}) : X \to \mathbb{R}$

$$\langle F(\boldsymbol{\mu}), \zeta \rangle = \mu_2 \int_\Omega \mathbf{v}_d \boldsymbol{\psi} \, d\Omega.$$

The affine structure can be underlined (as the previous example, very intuitive): with $Q_\mathcal{A} = 1$, $Q_\mathcal{B} = 2$ and $Q_F = 1$ the affine decomposition of the problem is given by

$$\Theta^1_\mathcal{A} = 1 \qquad \mathcal{A}^1(x, \zeta) = \mathcal{A}(x, \zeta),$$
$$\Theta^1_\mathcal{B} = \mu_1 \qquad \mathcal{B}^1(x, (\boldsymbol{\phi}, \xi)) = \int_\Omega \nabla \mathbf{v} \cdot \nabla \boldsymbol{\phi} \, d\Omega,$$
$$\Theta^2_\mathcal{B} = 1 \qquad \mathcal{B}^2(x, (\boldsymbol{\phi}, \xi)) = -\int_\Omega \mathrm{div}(\boldsymbol{\phi}) p \, d\Omega - \int_\Omega \mathrm{div}(\mathbf{v}) \xi \, d\Omega - \int_\Omega \mathbf{u} \cdot \boldsymbol{\phi} \, d\Omega,$$
$$\Theta^1_F = \mu_2 \qquad \langle F^1, \zeta \rangle = \int_\Omega \mathbf{v}_d \boldsymbol{\psi} \, d\Omega.$$

Let us describe a specific test experiment. In this case we took as parameter $\boldsymbol{\mu} = (1, 10)$ and as penalization term $\alpha = 10^{-4}$. To build the reduced basis we directly decide to use a partitioned POD of 10 basis functions on a training set of 50 points. In figure 4.3.2.1 and in figure 4.3.2.2 we can notice how rapidly the error decays and how the reduced solution is similar to the full order one.
Table 4.2 highlights how much RB methods are useful for a Stokes problem. Even in this case the dimensionality is drastically reduced: we recall that $\mathcal{N}$ is the full order dimension, while N is the reduced one. To solve a full order system a time $t_t = 48,78s$ is needed, while $t_r = 2.38 \cdot 10^{-2} s$, where t_r represents the reduced problem time resolution. This convention is adopted for all the following examples. Let J_t and J_r be the functionals associated to the Finite Element problem and the reduced problem, respectively. In this experiment $J_t = 5.5760 \cdot 10^{-2}$ and $J_r = 5.5764 \cdot 10^{-2}$.

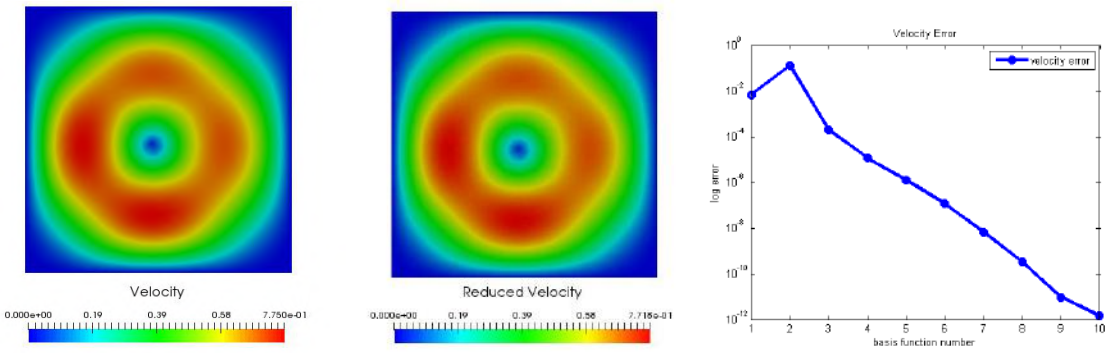

Figure 4.3.2.1: *Left*: full order state variable; *center*: reduced eim state variable, *right*: error.

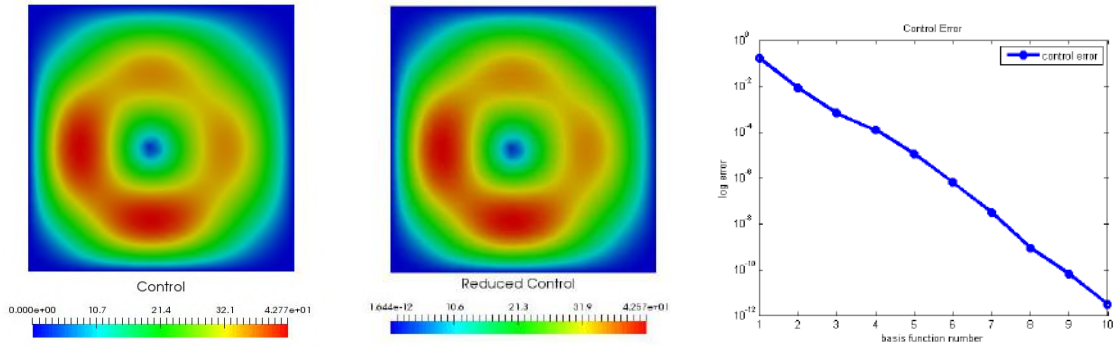

Figure 4.3.2.2: *Left*: full order control variable; *center*: reduced eim control variable, *right*: error.

Table 4.2: Speed up analysis for Stokes problem

Basis Number	1	2	3	4	5	6	7	8	9	10
Speed up	1731	2529	2408	2479	2314	2352	2252	2192	2012	1895

In the following a comparison between the dimentions of full order system and the reduced on is shown.

$\mathcal{N} \times \mathcal{N}$	23085×23085
$13N \times 13N$	130×130

Non-Affine Target Function

Let us analyse a particular non-affine case. We consider the problem (2.3.2), with a little variant on the target function $\mathbf{v}_d$. The new formulation (4.3.4) reads as follows: for a given $\boldsymbol{\mu} = [\mu_1, \mu_2, \mu_3, \mu_4]$, find $(y(\boldsymbol{\mu}), u(\boldsymbol{\mu})) \in Y \times U$ such that:

$$\min_{(y,u)\in Y\times U} J(\mathbf{v}, p, \mathbf{u}) = \frac{1}{2}\int_\Omega |\mathbf{v} - \mu_2 \mathbf{v}_d(\boldsymbol{\mu})|^2 \, d\Omega + \frac{\alpha}{2}\int_\Omega |\mathbf{u}|^2 \, d\Omega,$$
$$\text{such that } \begin{cases} -\mu_1 \Delta \mathbf{v} + \nabla p = \mathbf{u} & \text{in } \Omega, \\ \operatorname{div}(\mathbf{v}) = 0 & \text{in } \Omega, \\ \mathbf{v} = \mathbf{0} & \text{on } \partial\Omega, \end{cases} \tag{4.3.4}$$

where

$$\mathbf{v}_d = \left(\frac{\partial}{\partial x_2}(\varphi(x_1)\varphi(x_2)), -\frac{\partial}{\partial x_1}(\varphi(x_1)\varphi(x_2)) \right),$$

with $\varphi : (0,1) \to (0,1)$ is defined as $\varphi(z) = (1 - \cos(\mu_3\mu_4 z))(1-z)^2$. The target function is not affine in the parameters and the problem looses the affine assumption. To recover it we used a partitioned EIM-POD strategy (see Chapter 3), that allows us to approximate $\varphi(x)$ in an affine formulation in order to apply efficiently RB methods. The parameter $\boldsymbol{\mu}$ is in the parameter space $\mathcal{P} = [0.5, 1] \times [5, 10] \times [0.5, 0.8] \times [0, 3.5]$. Let us discuss some results: $\boldsymbol{\mu} = [1, 10, 0.8, \pi]$ is taken, we applied a POD reduction with 10 basis on a training set of 50 points using an uniform distribution on the parameters. The EIM approximation

was made exploiting 11 basis functions (they are sufficient, as we can see from the table 4.3).

Table 4.3: EIM Error $\varphi(x)$

EIM Basis Number	Error
1	3.97492e-05
2	1.56411e-06
3	1.92805e-08
4	2.08603e-10
5	2.87039e-12
6	9.38846e-14

In figure 4.3.2.3 and 4.3.2.4 the full order state and control expected are compared to their EIM-POD approximation, respectively. Table 4.4 represents the computational advantage of using RB spaces rather than full order approximation.

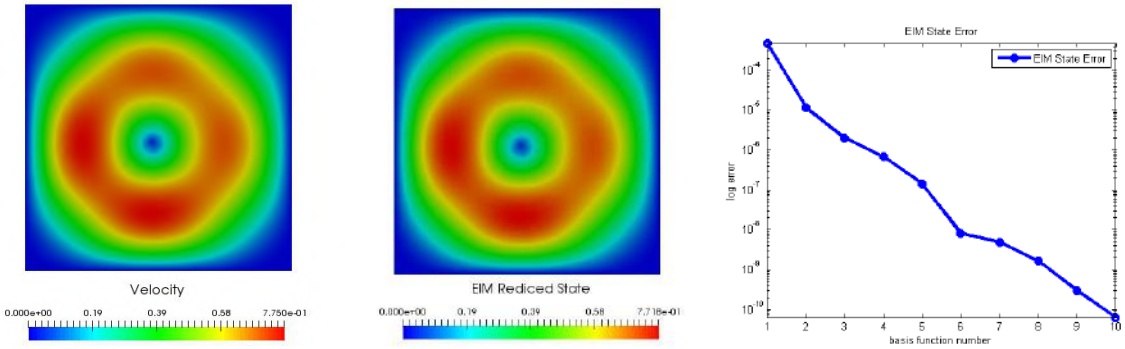

Figure 4.3.2.3: *Left*: full order state variable; *center*: reduced state variable, *right*: error.

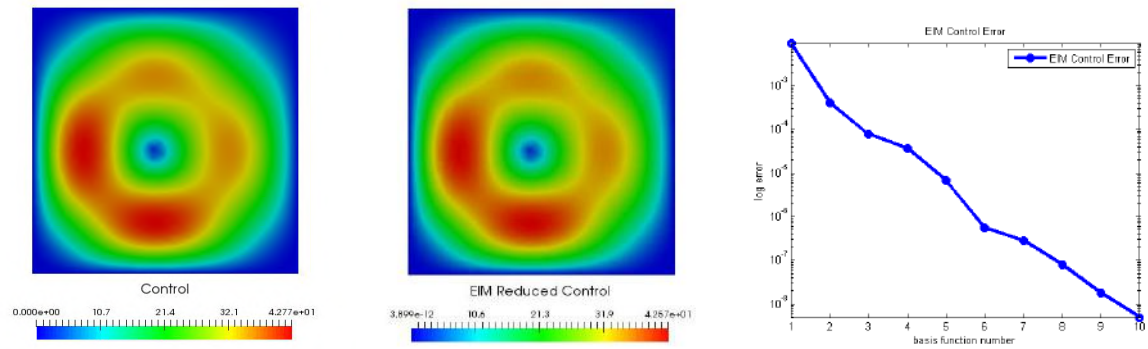

Figure 4.3.2.4: *Left*: full order control variable; *center*: reduced control variable, *right*: error.

Table 4.4: Speed up analysis for EIM Stokes problem

Basis Number	1	2	3	4	5	6	7	8	9	10
Speed up	922	1636	2079	1960	1641	1634	1602	1404	752	792

In the following a comparison between the dimensions of full order system and the reduced one is shown.

$\mathcal{N} \times \mathcal{N}$	23085×23085
$13N \times 13N$	130×130

The time resolution for the full order system and the reduced one are, respectively, $t_t = 53.59s$ and $t_r = 1.17s$. As we did for the other examples, we indicate with J_t the cost functional related to the full order problem and J_r the one related to the reduced problem. What we reach is $J_t = 5.55758 \cdot 10^{-2}$ and $J_r = 5.55759 \cdot 10^{-2}$.

4.3.3 An Environmental Preliminary Application: Thermal Pollution into a River

In this subsection we will apply RB methods to solve an OCP($\boldsymbol{\mu}$) dealing with an environmental application. We will parametrized a control problem governed by advection-diffusion state equation. The example we are going to face is inspired to the one proposed in [57, Subsection 17.11.2] and deals with the issue of the thermal pollution of a river. Thermal pollution can be very dangerous since it can change the natural habitat of an ecological specie, causing the loss of biodiversity. Advection-diffusion control models are very useful to avoid those kinds of issues and they are exploited in order to prevent ecological problems when a new industrial system must be designed.
The theoretical base of the general OCP governed by advection-diffusion is discussed in Example 1.3.2.2. As usual, we will indicate with Y and U the state and the control spaces, respectively. The adjoint space is $Q = Y$. Let us adapt the structure theoretically proposed in 1.3.2.2 to our specific parametric version of the control problem. First of all, we introduce the spatial domain used: in figure 4.3.3.1 three subdomains are highlighted

1. Ω_1: the portion of the domain where the forcing term is defined;
2. Ω_2: the portion of the domain where the control variable is defined;
3. Ω_{OBS}: the portion of the domain where the observation is made.

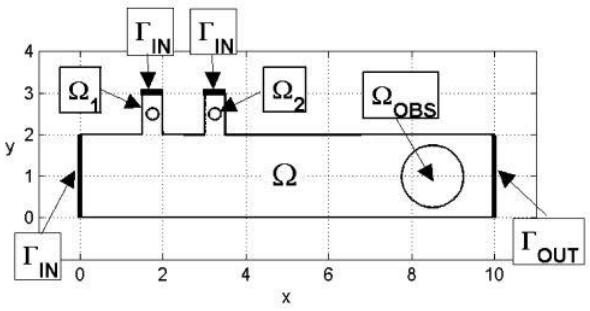

Figure 4.3.3.1: River pollution domain description, from [57, Subsection 17.11.2]

.

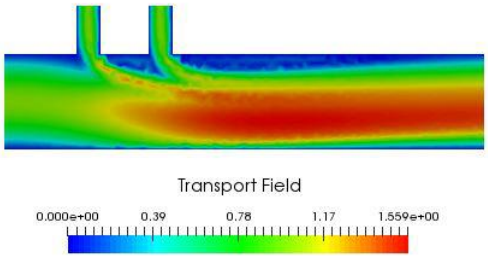

Figure 4.3.3.2: Transport field.

In this particular example we are going to reach a desired temperature profile, avoiding alterations the industrial system emission rate. The non dimensional OCP($\boldsymbol{\mu}$) problems reads as follows:

$$\begin{aligned} \min_{(y,u)} J(y,u) &= \int_{\Omega_{OBS}} (y - y_d)^2 \, d\Omega_{OBS} + \alpha \int_{\Omega_{OBS}} (u - \mu_3)^2 \, d\Omega_{OBS}, \\ \text{such that } &\begin{cases} -\text{div}(\mu_1 \nabla y) + \boldsymbol{\beta} \cdot \nabla y = \mu_2 \chi_1 + u\chi_2 & \text{in } \Omega, \\ y = 0 & \text{on } \Gamma_{IN}, \\ \mu_1 \dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N. \end{cases} \end{aligned} \tag{4.3.5}$$

where Γ_{IN} in specified in figure 4.3.3.1 and $\Gamma_N = \partial\Omega \setminus \Gamma_{IN}$, $y \in Y = H^1_{\Gamma_D}(\Omega)$ represents the temperature profile, $u \in U = \mathbb{R}$ is the control variable, whereas $y_d \in L^2(\Omega)$ is the desired temperature profile. With χ_1 and χ_2 we will indicate the characteristic functions associated to Ω_1 and Ω_2. The parameter $\boldsymbol{\mu} = [\mu_1, \mu_2, \mu_3]$ is considered in the parameter space $\mathcal{P} = [0.01, 0.1] \times [5, 10] \times [5, 10]$, where μ_1 represents the diffusivity coefficient, μ_2 a source term and μ_3 the desired emission rate. The penalization term is $\alpha = 10^{-3}$. We have to specify that $\boldsymbol{\beta}$ is a transport field (see figure 4.3.3.2) given by the solution of Navier-Stokes equation in the domain Ω with the following boundary conditions:

- a parabolic velocity profile on Γ_{IN} with 1 as maximal value;
- $\dfrac{\partial \boldsymbol{\beta}}{\partial \mathbf{n}} = 0$ on Γ_{OUT};
- *no-slip* conditions on $\partial\Omega \setminus (\Gamma_{IN} \cup \Gamma_{OUT})$;
- Reynolds number $\mathbb{R}e = 500$.

The problem can be rewritten in weak formulation as follows:

$$a(y, q; \boldsymbol{\mu}) = c(u, q), \qquad \forall q \in Q,$$

where the bilinear forms $a : Y \times Q \to \mathbb{R}$ and $c : U \times Q \to \mathbb{R}$ are defined as follows:

$$\begin{aligned} a(y, q; \boldsymbol{\mu}) &= \mu_1 \int_\Omega \nabla y \cdot \nabla q \, d\Omega + \int_\Omega \boldsymbol{\beta} \cdot \nabla y q \, d\Omega \\ c(u, q) &= u \int_\Omega q \, d\Omega. \end{aligned}$$

Moreover, the bilinear forms $m : Y \times Y \to \mathbb{R}$ and $n : U \times U \to \mathbb{R}$ are given by

$$m(y, z) = 2 \int_\Omega yz; d\Omega, \qquad n(u, v) = 2 \int_\Omega uv \, d\Omega.$$

To recast the problem in a saddle-point formulation, as usual, define $X = Y \times U$ and consider $x = (y, u)$, $w = (z, v) \in X$ and $q \in Q$. Furthermore we can define the linear forms as:

$$\langle F(\boldsymbol{\mu}), w\rangle = 2\int_\Omega y_d z \, d\Omega + 2\mu_3 \int_\Omega v \, d\Omega, \qquad \langle G(\boldsymbol{\mu}), q\rangle = \mu_2 \int_\Omega q \, d\Omega,$$

and the bilinear forms

$$\begin{aligned} \mathcal{A}(x, w : \boldsymbol{\mu}) &= m(y, z) + \alpha n(u, v), \\ \mathcal{B}(w, q; \boldsymbol{\mu}) &= a(y, q; \boldsymbol{\mu}) - c(u, q). \end{aligned}$$

In this way we have recast the problem (4.3.5) in the saddle-point formulation typical of linear quadratic OCP($\boldsymbol{\mu}$)s. We can underline the affine structure of the problem: with $Q_\mathcal{A} = 1$, $Q_\mathcal{B} = 2$, $Q_F = 2$ and $Q_G = 1$ the affine decomposition of the problem is given by

$$\begin{aligned} &\Theta^1_\mathcal{A} = 1 && \mathcal{A}^1(x, w) = \mathcal{A}(x, w), \\ &\Theta^1_\mathcal{B} = \mu_1 && \mathcal{B}^1(x, q) = \int_\Omega \nabla y \cdot \nabla q \, d\Omega, \\ &\Theta^2_\mathcal{B} = 1 && \mathcal{B}^2(x, q) = \int_\Omega \boldsymbol{\beta} \cdot \nabla y q \, d\Omega - v \int_\Omega q \, d\Omega, \\ &\Theta^1_G = \mu_2 && \langle G^1, q\rangle = \int_\Omega q \, d\Omega, \\ &\Theta^1_F = \mu_3 && \langle F^1, w\rangle = 2\int_\Omega v \, d\Omega, \\ &\Theta^2_F = 1 && \langle F^2, w\rangle = 2\int_\Omega y_d z \, d\Omega, \end{aligned}$$

Let us show some numerical results for specific values of the parameters. For the construction of the reduced basis a partitioned POD technique was used, exploiting 50 reduced basis with a training set of 100 points. The parameters were chosen trough uniform distribution. In the experiment proposed the parameter was fixed to $\boldsymbol{\mu} = [0.01, 10, 10]$. The penalization term is $\alpha = 10^{-3}$ and $y_d = 0$. In figure 4.3.3.3, the full order optimal temperature profile and the reduced optimal temperature profile are shown, the bottom figure represents the pointwise error, which has 10^{-7} as maximum value. The error trends with respect the full order approximation are presented in figure 4.3.3.4. The top left, top right and bottom left plots describe the optimal state, the control variable and the adjoint variable error, respectively. The bottom right plot represents the comparison between the monolithic approach and the partitioned approach (for the error definitions, see footnote 1) in terms of error norms: notice that for few reduced basis functions the two method are comparable, while the partitioned method gives a consistent improvement from the 15^{th} reduced basis and on. Let us analyse how much RB is convenient computationally speaking: in Table 4.5 the temporal improvement from basis number 1 to basis number 50 is presented (the comparison between full order and reduced order is shown through the usual *speed up index*); notice how the dimension of the reduced system is lower than the full order one, this lead to the following results $t_t = 1.7s$ and $t_r = 3.56 \cdot 10^{-2}$. Finally, let us analyse the cost functionals. Let J_t and J_r have the usual interpretation. In this experiment we obtain $J_t = J_r = 4.1836 \cdot 10{-}2$.

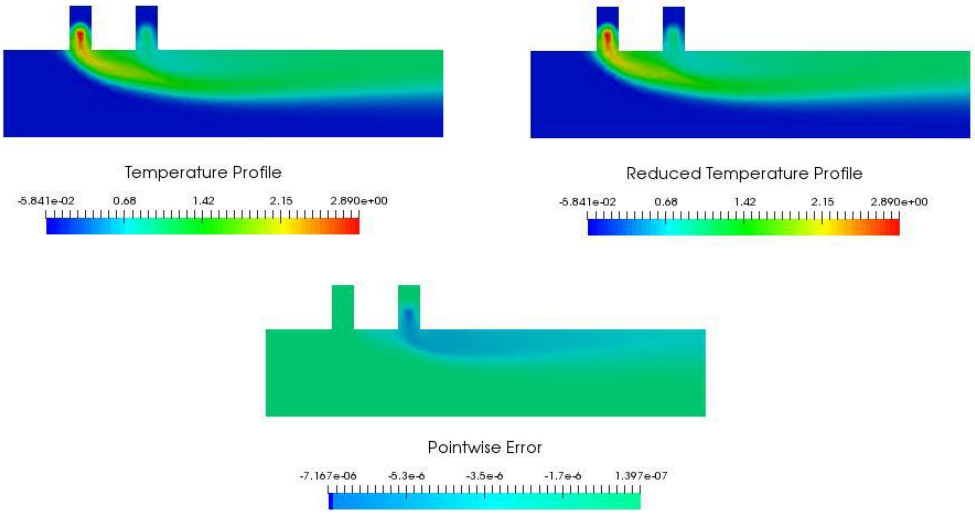

Figure 4.3.3.3: *Top left*: full order temperature profile, *top right*: reduced temperature profile, *bottom*: pointwise error.

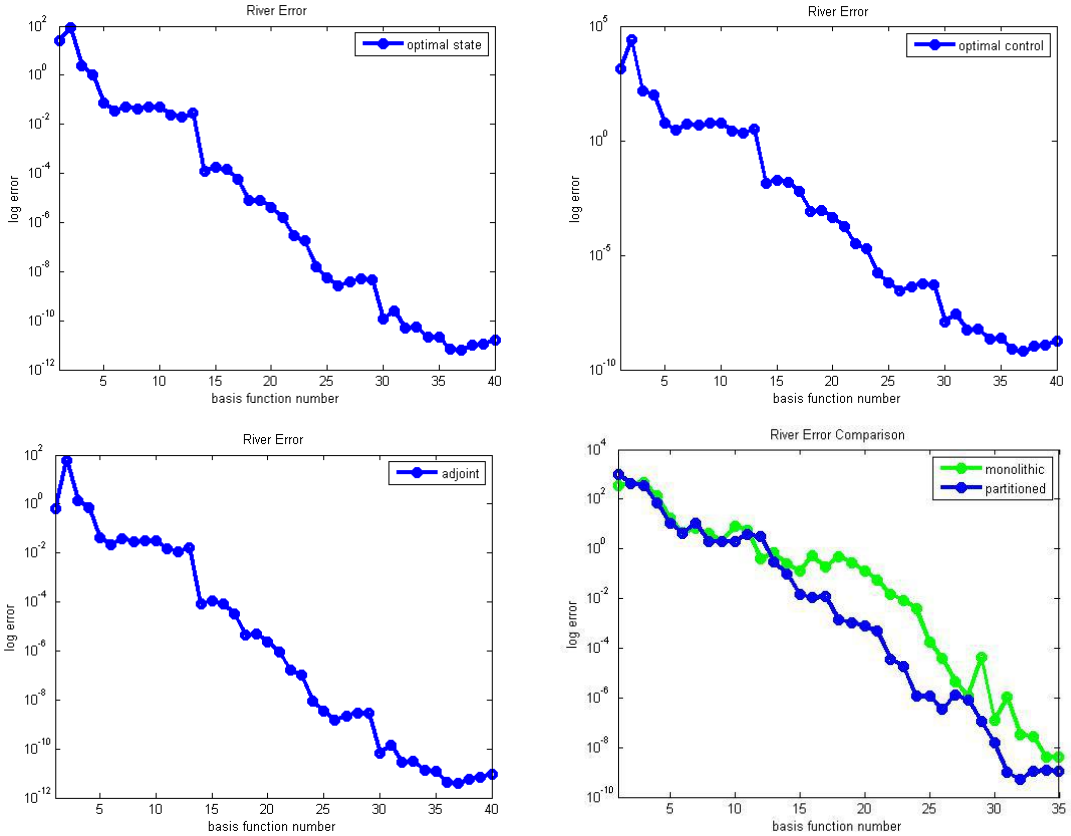

Figure 4.3.3.4: *Top left*: reduced optimal state error trend; *top right*: reduced control error trend. *bottom left*: reduced adjoint state error trend; *bottom right*: comparison between monolithic and partitioned approaches.

Table 4.5: Speed up analysis for river problem

Basis Number	1	2	3	4	5	6	7	8	9	10
Speed up	75	88	73	66	59	72	66	68	71	77
Basis Number	11	12	13	14	15	16	17	18	19	20
Speed up	67	74	69	64	58	72	65	66	68	49
Basis Number	21	22	23	24	25	26	27	28	29	30
Speed up	73	63	67	72	78	67	76	63	44	65
Basis Number	31	32	33	34	35	36	37	38	39	40
Speed up	68	65	51	64	56	60	39	55	56	54
Basis Number	41	42	43	44	45	46	47	48	49	50
Speed up	13	53	47	48	46	46	43	42	38	45

In the following a comparison between the dimensions of full order system and the reduced one is shown.

$\mathcal{N} \times \mathcal{N}$	2975×2975
$5N \times 5N$	250×250

Chapter 5

Reduced Basis Applications in Environmental and Engineering Marine Sciences

In this chapter we are going to apply the RB numerical methods to environmental control problems. Specifically, we will focus our attention on marine sciences: thematically and geographically speaking, large scale dynamic of Atlantic Ocean and pollutant control on the area of the Gulf of Trieste are analysed. As we have already underlined among this work, one can be interested in predicting some quantities and in controlling some physical or geometrical features in many contexts: from civil environmental engineering to ecological and climatological sciences. The theory of optimal control problems can be widely exploited in these topics and applications: in the last years environmental studies have grown and have been object of interest of many specific fields. Mathematical modeling and computational analysis are very important tools to study and face environmental issues. Optimal control applications can be easily adapted to various marine environmental problems: in forecasting (for example see [39, 74]), in preventing human activity effects and in monitoring pollutant quantities and substances (i.e. see [16, 18]) and in eco-friendly industrial planning and designing (i.e. see [48]), just for citing few possible applications.
Reduced basis methods can be a very efficient way to study all these phenomena: as we already specified in Chapter 4, control problems are hugely demanding under the computational point of view. RB methods can be a useful and powerful instrument to rapidly solve problems with a new low-dimensional formulation. Furthermore, environmental problems are characterized by a great use of parameters. They could describe physical features of the model or geometrical characteristics of the domain that we are considering. For these reasons, RB methods are an appropriate resource in this particular field of research. There are many works that verify how RB approach is a suitable technique to solve optimal control problems linked to environmental issues (i.e. see [16, 58, 59]).
In this thesis we will essentially focus on two thematic fields.

1. **Forecasting and studying general Ocean circulation models**
   Ocean circulation models are a wide studied topic and they have fascinated scientists in many fields of knowledge. In the last years, their analysis is growing and improving (i.e. see [71, 74, 11]) since they are related to serious issues as global warming, gulf current weakening and other problems linked to anthropic causes. They are analysed in the hope of forecasting catastrophic events in order to prevent them. One way to have more realistic forecasting models is to exploit data assimilation technique

(i.e. see [10, 9, 3, 28]): during the simulations this approach allow to train the model adding information from experimental data. In this chapter we focused on the study of the PDEs governing the Ocean circulation model and then adapt a parametric state solution tracking problem to them (this is the base of assimilation analysis). Numerically, the Ocean circulation models are deeply studied at the best of our knowledge: the numerical simulation technique is strongly linked to finite differences approach (i.e. see [73, 11]). In this work we have not only used a Finite Element approximation for the full order solution, but we will exploit RB methods to show their advantages for the specific case.

2. **Environmental engineering application of optimal control problem governed by advection-diffusion equations**
In the second application we focused on advection-diffusion pollutant optimal control problems. We have already treated the topic: theoretically in chapter 1, numerically in chapter 4, respectively. In this Chapter we are going to adapt the concept previously studied to our geographical reality: the Gulf of Trieste. Optimal control problems can be a great instrument to avoid environmental and ecological changes. The Gulf of Trieste is a physical basin particularly windy and it has very peculiar *flora* and *fauna* population (i.e. see [68, 50]). Moreover its analysis is important and interesting since it has a great impact on Trieste community: the city of Trieste overlooks the sea and depends on the Gulf and on its structures from harbours to tourist infrastructures. For these reasons it needs to be monitored and kept under control and, maybe, redesigned.

Finally, let us introduce how the chapter will be organized. It will substantially consist in two sections. In the first section the Ocean circulation model will be discussed, then we will introduce an OCP($\boldsymbol{\mu}$) state tracking problem, analysing it under a saddle-point formulation. Finally, some numerical results will be shown on different domains. In Section 5.2, RB methods for advection-diffusion control problems are discussed. We will apply them to a pollutant control test case and then on a OCP($\boldsymbol{\mu}$) with the Gulf of Trieste as reference domain.

5.1 Reduced Basis Application to Ocean Circulation Model

As we said in the introduction to the Chapter, in this section we will study the dynamics of general Ocean circulation. This kind of large scale analysis is what is behind long time scale forecasting models. Ocean dynamics is strictly linked to the wind action. Large scale circulation, indeed, can be considered as a coupled system of atmosphere and Ocean. It satisfy a balance between pressure gradient forces and the effects of Earth's rotation (physically described by the Coriolis' effect). In Oceanography, this phenomenon is called *geostrophic equlibrium*. The model that we are going to use adds quantities (time derivatives, diffusive effects...) to the equilibrium and then it constitutes the subject of the *quasi-geostrophic theory*. In this section we are going to describe the PDEs governing this large scale dynamics (for the theoretical physical topics, we refer to [11]), then we will introduce a state solution tracking control problem governed by quasi-geostrophic equations that will be solved with RB methods and then compared with Finite Element approximation.

5.1.1 General Governing Equation

In this subsection we are going to present the general non-linear quasi-geostrophic equation.
Quasi-geostrophic equation describes the homogeneous wind-driven Oceanic circulation. In other words it explains the Ocean gyres phenomena due to wind stress action under a single fluid layer[1] assumption. The non-dimensional state equation has the following form:

$$\Big(\frac{\delta_I}{L}\Big)^2 \mathcal{F}(\psi, \Delta\psi) + \frac{\partial \psi}{\partial x} = f - \frac{\delta_S}{L}\Delta\psi + \Big(\frac{\delta_M}{L}\Big)^3 \Delta^2\psi, \tag{5.1.1}$$

where, given a suitable space V, the non-linearity of the expression is given by $\mathcal{F}(\cdot,\cdot) : V \times V \to \mathbb{R}$ defined as:

$$\mathcal{F}(\psi, q) = \frac{\partial \psi}{\partial x}\frac{\partial q}{\partial y} - \frac{\partial \psi}{\partial y}\frac{\partial q}{\partial x}.$$

Moreover,

$$\Delta^2\psi = \frac{\partial^4 \psi}{\partial x^4} + \frac{\partial^4 \psi}{\partial y^4} + 2\frac{\partial^4 \psi}{\partial y^2 \partial x^2}.$$

As specified in [11, Section 3.2], the forcing term is linked to the action of the wind by the following relation:

$$f = \hat{\mathbf{k}} \cdot \operatorname{rot} \boldsymbol{\tau},$$

where $\hat{\mathbf{k}}$ is the third reference spatial unit vector, whereas $\boldsymbol{\tau}$ represents the wind stress. Considering an open, bounded and regular domain $\Omega \subset \mathbb{R}^2$ we have to impose *no-slip* conditions over ψ and $\Delta\psi$, in other words:

$$\psi = 0 \qquad \text{and} \qquad \Delta\psi = 0 \qquad \text{on } \partial\Omega.$$

We have already specified that quasi-geostrophic equation describes large scale Oceanic dynamics. The division for the constant L gives non dimensionality to the system and $L = \mathcal{O}(10^6)$.
The quasi-geostrophic equation is linked to geophysical Navier-Stokes equations: more specifically, the first equation is the *stream function* formulation of the following system of equations

$$\begin{cases} \varepsilon(\mathbf{u}\cdot\nabla)u - (1+\varepsilon y)v = -\dfrac{\partial p}{\partial x} + \dfrac{\varepsilon}{\mathbb{R}e}\Big(\dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2}\Big) + f_1 & \text{in } \Omega, \\ \varepsilon(\mathbf{u}\cdot\nabla)v + (1+\varepsilon y)u = -\dfrac{\partial p}{\partial y} + \dfrac{\varepsilon}{\mathbb{R}e}\Big(\dfrac{\partial^2 v}{\partial x^2} + \dfrac{\partial^2 v}{\partial y^2}\Big) + f_2 & \text{in } \Omega, \\ \operatorname{div}(\mathbf{u}) = 0 & \text{on } \partial\Omega, \\ \mathbf{u} = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.1.2}$$

where $\mathbf{u} = (u, v) \in H_0^1(\Omega) \times H_0^1(\Omega)$ is a velocity field with scale $\mathcal{O}(10^{-2} m/s)$, $\mathbb{R}e$ is the Reynolds number that verifies

$$\frac{1}{\mathbb{R}e} = \mathcal{O}\Big(\frac{\delta_M}{L}\Big)^3$$

and $\varepsilon = \mathcal{O}(10^{-4})$. The terms $-(1+\varepsilon y)v$ and $(1+\varepsilon y)u$ represent the Coriolis' effect for the first velocity component and the second velocity component, respectively. The reader

[1] We are assuming that the whole fluid has the same density and so stratification effects are neglected.

interested in the derivation of the quasi-geostrophic equation from the geophysical Navier-Stokes equations can refer to [11, Chapter 3]. The following relation links the velocity field and the *stream function*:

$$\mathbf{u} = (u, v) = \Big(-\frac{\partial \psi}{\partial y}, \frac{\partial \psi}{\partial x} \Big). \tag{5.1.3}$$

Let us analyse the different dynamics depending on the parameters $\delta_I, \delta_S, \delta_M$. The existence and uniqueness of the solution is non-trivially proved in [19]. As spatial domain we firstly considered $\Omega = [0,1] \times [0,1]$, that represents the Ocean surface in this example. The various regimes depend on the relative amplitude of δ_M and δ_I. The behaviour of the solution of the quasi-geostrophic equation under the forcing term $f = -\sin(\pi y)$ is shown in three cases (the simulation are based on a Finite Element discretization[2], as a reference on stability of the method see [40]):

1. the first case shown is a linear case, corresponding to $\delta_I = 0$ and $\delta_M = 7 \cdot 10^4$. As we can see in the left of figure 5.1.1.1, the solution is characterized by an intensification on the western boundary, a phenomenon very known in literature and also observed in nature (i.e. Gulf Stream, as reported in [74]);

2. in the center of figure 5.1.1.1 the solution of a moderate amount of non-linearity effect is presented. In this case we have chosen $\delta_I = \delta_M = 7 \cdot 10^4$. Notice that the gyre moves northward;

3. the last configuration corresponds to a highly non-linear system, where $\delta_I >> \delta_M$, in the specific $\delta_M = 7 \cdot 10^3$ and $\delta_I = 7 \cdot 10^4$. What we can observe is the intensification of the northward movement of the circulation gyre[3].

In figure 5.1.1.2, we can observe the same behaviour described for the squared domain, but on a mesh that simulate the North Atlantic Ocean.

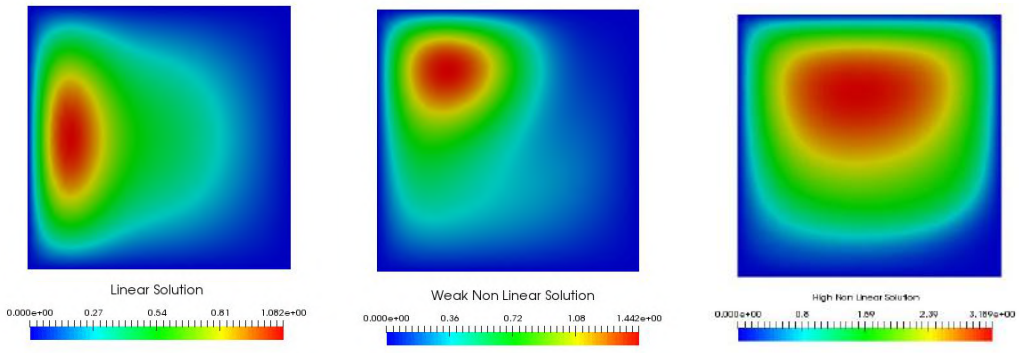

Figure 5.1.1.1: *Left*: linear solution; *center*: weak nonlinear solution, *right*: high nonlinear solution.

[2]The weak formulation of the quasi-geostrophic equation was already discussed in subsection 3.4 and it will be better presented in the following subsection.

[3]The highly non-linear configuration is reached as a steady state of the time dependent problem:

$$\frac{\partial \Delta \psi}{\partial t} + \Big(\frac{\delta_I}{L}\Big)^2 \mathcal{F}(\psi, \Delta\psi) + \frac{\partial \psi}{\partial x} = f - \frac{\delta_S}{L}\Delta\psi + \Big(\frac{\delta_M}{L}\Big)^3 \Delta^2 \psi. \tag{5.1.4}$$

Indeed, the problem (5.1.2) is unstable when $\delta_I >> \delta_M$. To avoid this inconvenient there are several techniques that can be used, like the ones proposed in the articles [12, 51]. See the appendix Perspectives for a deeper analysis.

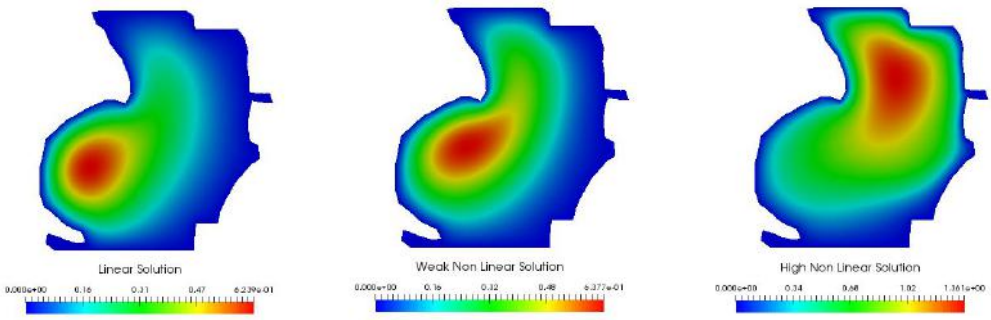

Figure 5.1.1.2: *Left*: linear solution; *center*: weak nonlinear solution, *right*: high nonlinear solution.

The mesh used in figure 5.1.1.2 is obtained trough *FreeFem++* (as a reference see [32] and visit the link *http://www.freefem.org/*) from Google Earth images, and then, thanks to *Gmsh* (see as a reference [27], and visit *http://gmsh.info/*), imported into FEniCS for Finite Element simulation (in this case we refer to [45], for further informations one can refer to *https://fenicsproject.org*). In figure 5.1.1.3 we can give a taste of the work needed to obtain a physical mesh for the North Atlantic Ocean.

Figure 5.1.1.3: *Left*: mesh and state solution; *center*: western Atlantic Ocean, *right*: eastern Atlantic Ocean.

Remark 5.1.1. To make the model more realistic one can add the influence of bathymetry in the dynamics equations. In the quasi-geostrophic case it is simple to consider the underwater depth effect in the equation (we always refer to [11, Chapter 3]). Let us suppose that the sea floor can be defined by a smooth function $h : \Omega \to \mathbb{R}$. The new state equation to be examined is:

$$\Big(\frac{\delta_I}{L}\Big)^2 \mathcal{F}(\psi, \Delta\psi + h(x,y)) + \frac{\partial \psi}{\partial x} = f + \Big(\frac{\delta_M}{L}\Big)^3 \Delta^2 \psi - \Big(\frac{\delta_S}{L}\Big)\Delta\psi.$$

To describe the Atlantic floor[4], for example, one can use

$$h(x,y) = e^{-\frac{x}{\frac{\delta_I}{L}}} + g(x,y),$$

where $g(x,y)$ is a Gaussian function representing the mid-Atlantic Ridge, whereas the exponential function represents a rapid coastal decay. We notice that bathymetry did not

[4]Atlantic domain experiments are analysed in figure 5.1.1.2

cause considerable changes: for this reason we decided to use a no-bathymetry simplified model, without loss of generality.

5.1.2 Linear Optimal Control Problem Formulation

As we have already declared in the introduction to this Chapter, we are going to build an OCP($\boldsymbol{\mu}$) governed by quasi-geostrophic equation. We will focus on the linear case: consider this simple case is not restrictive, since the behaviour of the large scale circulation usually follows a linear trend.
The optimal control problem that we will face is a state solution tracking. This procedure is the base of data assimilation technique, much exploited in forecasting models (i.e. see [10, 9, 3, 28]). Let us define the problem:

$$\min_{(\psi,u)\in Y\times U} J(\psi,u) = \frac{1}{2}\int_\Omega (\psi-\psi_d)^2\,d\Omega + \frac{\alpha}{2}\int_\Omega u^2\,d\Omega$$
$$\text{such that} \quad \begin{cases} \dfrac{\partial\psi}{\partial x} = u - \mu_1\Delta\psi + \mu_2\Delta^2\psi & \text{in } \Omega, \\ \psi = 0 & \text{on } \partial\Omega, \\ \Delta\psi = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.1.5}$$

where $\psi \in Y$ is our state variable, $u \in U$ is the forcing term to be controlled, where Y and U are two suitable functions spaces, that will be lately specified. The parameter $\boldsymbol{\mu} = (\mu_1, \mu_2)$ represents the diffusivity action of the system and the parameter space is $\mathcal{P} = [10^{-4}, 1] \times [10^{-4}, 1]$, whereas the penalization term is $\alpha = 10^{-5}$.
Now, let us rewrite the problem in the following way:

$$\min_{((\psi,q),u)\in Y\times U} J((\psi,q),u) = \frac{1}{2}\int_\Omega (\psi-\psi_d)^2\,d\Omega + \frac{\alpha}{2}\int_\Omega u^2\,d\Omega$$
$$\text{such that} \quad \begin{cases} q = \Delta\psi & \text{in } \Omega, \\ \dfrac{\partial\psi}{\partial x} = u - \mu_1 q + \mu_2\Delta q & \text{in } \Omega, \\ \psi = 0 & \text{on } \partial\Omega, \\ q = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.1.6}$$

where the spaces are defined as $Y = H_0^1(\Omega) \times H_0^1(\Omega)$ and $U = L^2(\Omega)$. The aim of the problems (5.1.5) and (5.1.6) is to let the solution ψ be the most similar to $\psi_d \in L^2(\Omega)$, the desired state variable[5].
Let us introduce the weak formulation of the state equation. It can be expressed in the following way:

$$a((\psi,q),(\phi,p);\boldsymbol{\mu}) = c(u,\phi) \qquad \forall \phi, p \in H_0^1(\Omega), \tag{5.1.7}$$

where $a : Y \times Y \to \mathbb{R}$ and $c : U \times Y \to \mathbb{R}$ are given by:

$$a((\psi,q),(\phi,p);\boldsymbol{\mu}) = \int_\Omega \frac{\partial\psi}{\partial x}\phi\,d\Omega + \mu_2\int_\Omega \nabla q\cdot\nabla\phi\,d\Omega + \\ + \mu_1\int_\Omega q\phi\,d\Omega + \int_\Omega qp\,d\Omega + \int_\Omega \nabla q\cdot\nabla p\,d\Omega,$$
$$c(u,\phi) = \int_\Omega u\phi\,d\Omega.$$

[5]The function ψ_d in a data assimilation framework represents real data that allow to update model parameters in order to be more precise in the forecasting previsions.

The definition of the bilinear form $a(\cdot,\cdot)$ derives from integration by parts, Divergence Theorem and the belonging of the variables (ψ, q) and the functions (ϕ, p) to the space $H_0^1(\Omega)\times H_0^1(\Omega)$. The equation (5.1.7) derives from the addition of the two weak equations:

$$\begin{cases} \displaystyle\int_\Omega qp\ d\Omega + \int_\Omega \nabla\psi\cdot\nabla p\ d\Omega = 0 & \forall(\phi,p)\in Y, \\ \displaystyle\int_\Omega \frac{\partial\psi}{\partial x}\phi\ d\Omega + \mu_2\int_\Omega \nabla q\cdot\nabla\phi\ d\Omega + \mu_1\int_\Omega q\phi\ d\Omega - \int_\Omega u\phi\ d\Omega = 0 & \forall(\phi,p)\in Y. \end{cases} \tag{5.1.8}$$

It is easy to prove the consistence of the weak formulation (5.1.7) with respect to the strong formulation. Indeed:

$$\int_\Omega qp\ d\Omega + \int_\Omega \nabla\psi\cdot\nabla p\ d\Omega + \int_\Omega \frac{\partial\psi}{\partial x}\phi\ d\Omega + \mu_2\int_\Omega \nabla q\cdot\nabla\phi + \mu_1\int_\Omega q\phi\ d\Omega - \int_\Omega u\phi\ d\Omega = 0$$

applying again integration by parts, Divergence Theorem and assuming that both state variables and (p,ϕ) vanish on the boundary of the domain, the previous equation can be written as

$$\int_\Omega (q-\Delta\psi)p\ d\Omega + \int_\Omega \Big(\frac{\partial\psi}{\partial x} + \mu_1 q - \mu_2\Delta q - u\Big)\phi\ d\Omega = 0.$$

Since p and ϕ are two arbitrary functions of $H_0^1(\Omega)$, we obtain the following system:

$$\begin{cases} q - \Delta\psi = 0, \\ \dfrac{\partial\psi}{\partial x} + \mu_1 q - \mu_2\Delta q - u = 0, \end{cases} \tag{5.1.9}$$

notice that the system (5.1.9) coincides with the strong formulation of the state equation in the problem (5.1.6).
Next step is to recast the optimal control problem into a saddle-point framework. Let us recall that the state space is $Y = H_0^1(\Omega)\times H_0^1(\Omega)$, the control space is $L^2(\Omega)$ and for the adjoint space we require that $Y = Q$. Let us define the product space $X = Y\times U$. Let $x = ((\psi,q),u)$ and $w = ((\chi,t),v)$ be two elements of X, whereas let $s = (\phi,p)$ be an element of Q. The problem (5.1.6) can be rewritten in the following general parametric way: given $\boldsymbol{\mu}\in\mathcal{P}$ find $(x(\boldsymbol{\mu}),p(\boldsymbol{\mu}))\in X\times Q$ such that

$$\begin{cases} \mathcal{A}(x(\boldsymbol{\mu}),w;\boldsymbol{\mu}) + \mathcal{B}(w,p(\boldsymbol{\mu}),\boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}),w\rangle & \forall w\in X \\ \mathcal{B}(x(\boldsymbol{\mu}),s;\boldsymbol{\mu}) = 0 & \forall s\in Q, \end{cases} \tag{5.1.10}$$

for some suitable bilinear forms $\mathcal{A}: X\times X\to\mathbb{R}$ and $\mathcal{B}: X\times Q\to\mathbb{R}$ and linear form $F: X\to\mathbb{R}$. To build these two bilinear forms we need not only $a(\cdot,\cdot;\boldsymbol{\mu})$ and $c(\cdot,\cdot)$, but also the bilinear forms $m: Y\times Y\to\mathbb{R}$ and $n: U\times U\to\mathbb{R}$ defined as

$$m(\psi,\chi) = \int_\Omega \psi\chi\ d\Omega$$
$$n(u,v) = \int_\Omega uv\ d\Omega,$$

and the linear form $F: X\to\mathbb{R}$ as:

$$\langle F,w\rangle = \int_\Omega \psi_d\chi\ d\Omega.$$

Let us define the bilinear form $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ and $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ as follows:

$$\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu}) = \mathcal{A}(x,w) = m((\psi,q),(\chi,t)) + \alpha n(u,v)$$
$$\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu}) = a((\psi,q),(\phi,p)) - c(v,\phi).$$

As we have specified in subsection 4.1.1, there are some assumptions to be made on the bilinear forms $a(\cdot,\cdot), c(\cdot,\cdot), m(\cdot,\cdot)$ and $n(\cdot,\cdot)$ to guarantee the consistence of the saddle point formulation, as well as, existence and uniqueness of the solution. The following are verified for every given $\boldsymbol{\mu}$:

- $a : Y \times Q \to \mathbb{R}$ is continuous over $Y \times Q$,
- $c : U \times Q \to \mathbb{R}$ is continuous over $U \times Q$,
- $n : U \times U \to \mathbb{R}$ is continuous, symmetric over $U \times U$ and coercive over U,
- $m : Y \times Y \to \mathbb{R}$ is symmetric, continuous and positive in the norm of the observation space.

In this specific case, for the decoupled version of the state equation used in the problem (5.1.6), the bilinear form $a(\cdot,\cdot;\boldsymbol{\mu})$ is not coercive[6]. Although this fact, we decided to reduce the system and study the results.
Before using reducing techniques, one has to define the discretized version of problem (5.1.10). As we already said, a Finite Element discretization is chosen. The discrete version of this specific OCP($\boldsymbol{\mu}$) problem reads: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(x^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}(\boldsymbol{\mu})) \in X^{\mathcal{N}} \times Q^{\mathcal{N}}$ such that

$$\begin{cases} \mathcal{A}(x^{\mathcal{N}}(\boldsymbol{\mu}), w^{\mathcal{N}}) + \mathcal{B}(w^{\mathcal{N}}, p^{\mathcal{N}}(\boldsymbol{\mu});\boldsymbol{\mu}) = \langle F, w^{\mathcal{N}}\rangle, & \forall w^{\mathcal{N}} \in X^{\mathcal{N}}, \\ \mathcal{B}(x^{\mathcal{N}}(\boldsymbol{\mu}), s^{\mathcal{N}};\boldsymbol{\mu}) = 0 & \forall s^{\mathcal{N}} \in Q^{\mathcal{N}}. \end{cases} \tag{5.1.11}$$

To build the reduced space we exploit a partitioned POD algorithm. Moreover, to ensure stability to the RB approximation, we used aggregated space for state and adjoint variables, defining the space

$$Z_N = \text{span } \{\zeta_n := (\psi^{\mathcal{N}}(\boldsymbol{\mu}^n), q^{\mathcal{N}}(\boldsymbol{\mu})),\ \xi_n := p^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1,\dots,N\}.$$

The control space has the usual form

$$U_N = \text{span } \{\lambda_n := u^{\mathcal{N}}(\boldsymbol{\mu}^n),\ n = 1,\dots,N\}.$$

Now, let us choose $Y_N = Z_N, X_N = Z_N \times U_N$ and $Q_N = Z_N$. So, the new RB approximated problem reads as: given $\boldsymbol{\mu} \in \mathcal{P}$ find $(x_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in X_N \times Q_N$ such that

$$\begin{cases} \mathcal{A}(x_N(\boldsymbol{\mu}), w_N) + \mathcal{B}(w_N, p_N(\boldsymbol{\mu});\boldsymbol{\mu}) = \langle F, w_N\rangle, & \forall w_N \in X_N, \\ \mathcal{B}(x_N(\boldsymbol{\mu}), s_N;\boldsymbol{\mu}) = 0 & \forall s_N \in Q_N. \end{cases} \tag{5.1.12}$$

As in the previous chapters, N is the reduced space dimension and the system we are going to solve is of the type proposed in subsection (4.1.4)

$$\begin{pmatrix} A_N(\boldsymbol{\mu}) & B_N^T(\boldsymbol{\mu}) \\ B_N(\boldsymbol{\mu}) & 0 \end{pmatrix} \begin{pmatrix} x_N(\boldsymbol{\mu}) \\ p_N(\boldsymbol{\mu}) \end{pmatrix} = \begin{pmatrix} F_N(\boldsymbol{\mu}) \\ 0 \end{pmatrix}. \tag{5.1.13}$$

[6]The coercivity of the problem can be recovered analysing the not decoupled system (5.1.5) as in proved in [40].

The affinity assumption must be guaranteed for the efficiency of the reduced problem. Let us highlight it. With $Q_{\mathcal{A}} = 1$, $Q_{\mathcal{B}} = 2$ and $Q_F = 1$ the affine decomposition of the problem is given by

$$\begin{aligned}
&\Theta^1_{\mathcal{A}} = 1 && \mathcal{A}^1(x,w) = \mathcal{A}(x,w),\\
&\Theta^1_{\mathcal{B}} = \mu_1 && \mathcal{B}^1(x,s) = \int_\Omega q\phi \, d\Omega,\\
&\Theta^2_{\mathcal{B}} = \mu_2 && \mathcal{B}^2(x,s) = \int_\Omega \nabla q \cdot \nabla \phi \, d\Omega,\\
&\Theta^2_{\mathcal{B}} = 1 && \mathcal{B}^3(x,s) = \int_\Omega \frac{\partial \psi}{\partial x}\phi \, d\Omega + \int_\Omega qp \, d\Omega + \int_\Omega \nabla q \cdot \nabla p \, d\Omega - \int_\Omega u\phi \, d\Omega,\\
&\Theta^1_{F} = 1 && \langle F^1, w\rangle = \langle F, w\rangle.
\end{aligned}$$

Let us analyse some results. We build a reduce space with 50 basis functions, thanks to a partitioned POD algorithm applied to a training set of 100 points. The sampling of the parameters is based on a log-uniform distribution for the two component of $\boldsymbol{\mu}$.

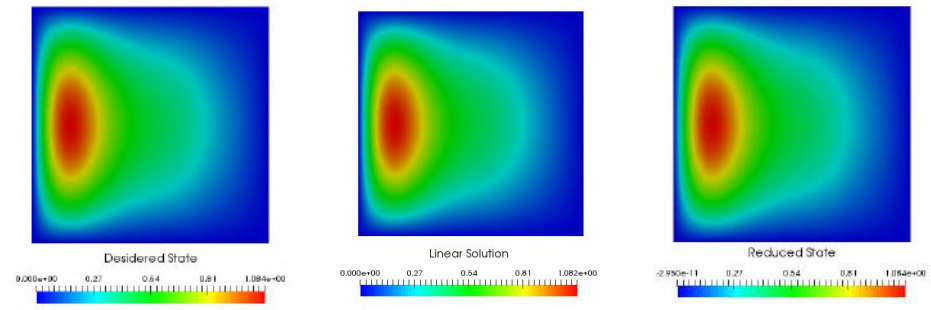

Figure 5.1.2.1: *Left*: desired state; *center*: full order solution, *right*: reduced solution.

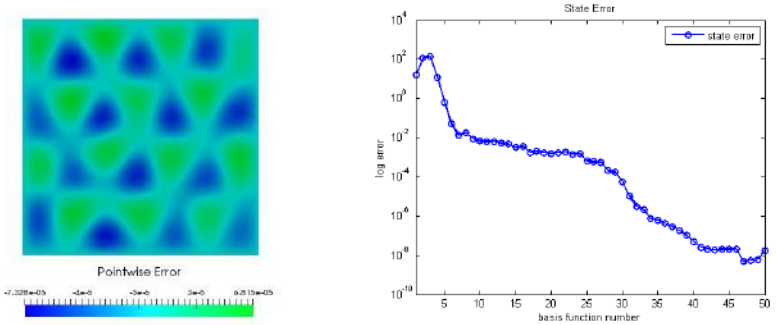

Figure 5.1.2.2: *Left*: pointwise error; *right*: error decay.

We treat the specific case with $\boldsymbol{\mu} = (0, 0.07^3)$ and the penalization term $\alpha = 10^{-5}$. First of all, in figure 5.1.2.1 we can notice that the state solution matches the desired state and the full order state solution. The pointwise error between full order state and the reduced one is shown in figure 5.1.2.2, together with error[7] decay. Lets us focus on how much is computationally convenient exploit RB methods rather than full order Finite Element

[7] of the norm. Let us define y_t the Finite Element solution and y_r the reduced solution. The state error norm is given by $\|y_t - y_r\|^2_Y$.

approximation. Geophysical simulations can last days, so it is very important to use an efficient technique in order save time and computational resources. As we did in subsection 4.3, the *speed up index* is used (we recall that it measures how many reduced problems can be solved in the time needed for a full order resolution). The performance of RB problem formulation is shown in Table 5.1. In the table the dimension of the full order system and of the reduced order system are reported: as we can notice, the reduced system is characterized by a low dimensionality with respect to the Finite Element system. As we did in the previous Chapter, we indicate with t_t and t_r, the full order and the reduced order time resolution, respectively. What we obtain is $t_t = 5.59s$ and $t_r = 2.38 \cdot 10^{-1}s$. Another quantity helps us to understand how efficient the RB methods are: the value of the cost functional. Let J_t and J_r the cost functionals related to the full order control problem and to the reduced order control problem. We obtain the following equality $J_t = J_r = 3.4465 \cdot 10^{-6}$.

Table 5.1: Speed up analysis for Quasi-Geostrophic equation on the square domain

Basis Number	1	2	3	4	5	6	7	8	9	10
Speed up	379	201	203	183	134	224	227	181	179	178
Basis Number	11	12	13	14	15	16	17	18	19	20
Speed up	212	174	187	185	156	199	133	150	140	137
Basis Number	21	22	23	24	25	26	27	28	29	30
Speed up	130	129	100	107	102	77	72	79	79	74
Basis Number	31	32	33	34	35	36	37	38	39	40
Speed up	71	62	64	58	54	53	49	43	46	40
Basis Number	41	42	43	44	45	46	47	48	49	50
Speed up	40	38	30	28	27	24	21	21	17	23

In the following a comparison between the dimensions of full order system and the reduced one is shown.

$\mathcal{N} \times \mathcal{N}$	5935×5935
$9N \times 9N$	450×450

With the same parameters and the same POD inputs, we have solved the OCP($\boldsymbol{\mu}$) problem on the mesh representing the North Atlantic Ocean. In figure 5.1.2.3 the desired state, the full order state solution and the reduced one are presented. They match: this hypothesis is supported by figure 5.1.2.4, on the left is presented a pointwise error of the difference between full order approximation and reduced approximation (the maximum value reached is $3.057 \cdot 10^{-6}$); on the right the state error norm decay (see footnote 7) is shown. Also in this case, Table 5.2 represents the computational advantages of using reduced problem: the *speed up index* shows how is convenient to exploit RB techniques and how many computational time can be saved with respect to the Finite Element approach. In this example we obtain $t_t = 6.07s$ and $t_r = 2.03 \cdot 10^{-1}s$, where t_t and t_r must be interpreted has the previous examples. Let us give to J_t and J_r the usual characterization. In this case we have $J_t = 7.3098 \cdot 10^{-6}$ and $J_r = 7.2668 \cdot 10^{-6}$.

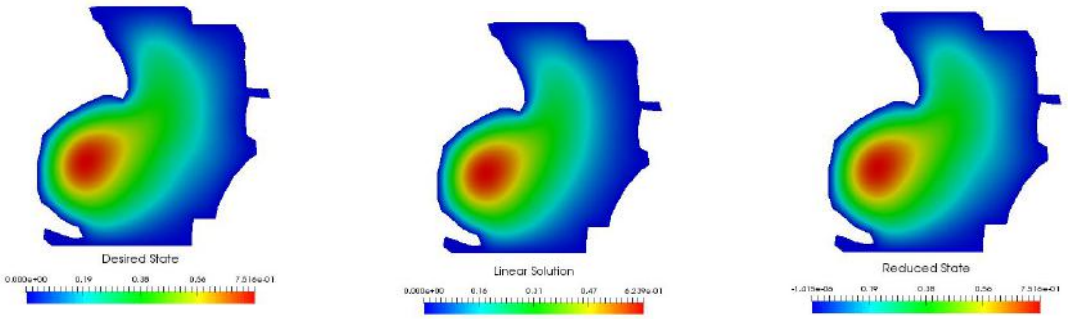

Figure 5.1.2.3: *Left*: desired state; *right*: full order solution, *bottom*: reduced solution.

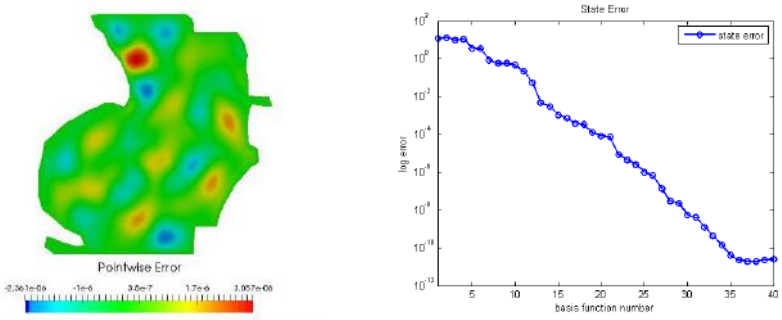

Figure 5.1.2.4: *Left*: pointwise error; *right*: error decay.

Table 5.2: Speed up analysis for Quasi-Geostrophic equation on the Atlantic Ocean mesh

Basis Number	1	2	3	4	5	6	7	8	9	10
Speed up	1049	369	437	309	395	223	315	385	247	192
Basis Number	11	12	13	14	15	16	17	18	19	20
Speed up	263	342	253	253	207	210	205	193	180	199
Basis Number	21	22	23	24	25	26	27	28	29	30
Speed up	132	151	95	66	33	75	90	107	66	54
Basis Number	31	32	33	34	35	36	37	38	39	40
Speed up	48	54	80	78	62	48	53	52	51	44
Basis Number	41	42	43	44	45	46	47	48	49	50
Speed up	29	41	35	34	31	28	25	21	17	23

In the following a comparison between the dimensions of full order system and the reduced one is shown.

$\mathcal{N} \times \mathcal{N}$	6490×6490
$9N \times 9N$	450×450

Conclusions

In Oceanography, a numerical tool capable to save computational resources is really needed and important. Even if the parametrization is usually linked to the physics of the problem, and not to the geometry of the domain, Oceanographic simulations could be a great effort to face.
A solution tracking with a given desired state was presented: as we already specified in the introduction to the chapter, this kind of optimal control problem is at the base of assimilation of experimental data. A complete data assimilation problem can be time dependent, non-linear and it could have many parameters to handle and then, it is usually very demanding. For this reasons RB methods are a valuable approach for these of physical systems.
Let us focus on the environmental aspects of the specific example proposed:

1. First of all, the state tracking solution gives us information on the Ocean currents, on their magnitude, position, on the gyres, etc. Globally speaking, the analysis of the dynamic of the Atlantic Ocean allows to understand many climatological phenomena governing the North Hemisphere.

2. Shifting our attention to the control variable, we can affirm that not only the currents can be forecast, but also the wind stress and the *fetch*, that is the portion of domain where the wind blows. It is important and interesting for the global understanding of the general Wind-Ocean Circulation dynamics and their climatological effects.

Both currents and wind stress are linked to a much more complex environmental interest, as the global warming and the global Ocean circulation. Even if the problem formulation is straightforward, with distributed control over the whole domain and homogeneous Dirichlet boundary conditions, the effects and the dynamics are very complex and linked to many geophysical phenomena. Concluding: this climatological environmental issue can be really demanding since involves a lot of physical variables to take into account. RB methods can be a very suitable, viable and powerful tool to reduce the computational effort of these complex models and could help a lot in this growing field of knowledge that attracts the interest of many scientific subjects.

5.2 Reduced Basis Application to Gulf of Trieste

In this section we will apply our knowledge on RB methods to an environmental control problem on the Gulf of Trieste. The problem aims at limiting the impact of a pollutant tracers on touristic and natural areas. The OCP($\boldsymbol{\mu}$) is governed by advection-diffussion state equation, theoretically discussed in chapter 1, and already generalized to parametric version in the RB application of subsection 4.3.3. The section is structured as follows: first of all, we will briefly recall the general problem formulation of an OCP($\boldsymbol{\mu}$): from the theoretical formulation, to the Finite Element approximation and finally, to RB saddle point formulation. Then, a test example of pollutant control is presented: it is adapted from [59]. Finally, we will show the application on the Gulf of Trieste.

5.2.1 General Problem Formulation

The aim of this subsection is to put together all the knowledge on parametric advection-diffusion optimal control problems. As usual, Y is the state space, U the control space

and $Y = Q$ the adjoint space. Let us recall the general weak formulation for a OCP($\boldsymbol{\mu}$) governed by advection-diffusion state equation: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(y(\boldsymbol{\mu}), u(\boldsymbol{\mu})) \in Y \times U$

$$\begin{aligned} \min_{(y,u)\in Y\times U} J(y,u) &= \frac{1}{2} m(y - y_d(\boldsymbol{\mu}), y - y_d(\boldsymbol{\mu}); \boldsymbol{\mu}) + \frac{\alpha}{2} n(u,u;\boldsymbol{\mu}) \\ \text{such that } a(y,q;\boldsymbol{\mu}) &= c(u,q;\boldsymbol{\mu}) + \langle G(\boldsymbol{\mu}), q\rangle \qquad \forall q \in Q. \end{aligned} \tag{5.2.1}$$

In this case, the expression $a(y, q; \boldsymbol{\mu}) = c(u, q; \boldsymbol{\mu})$ represents the parametrized version of advection-diffusion state equation with no forcing term, depending on the problem that we are considering. Let χ_{Ω_u} be the characteristic function of $\Omega_u \subset \Omega$, the portion of the domain where the control variable is defined. The strong parametric formulation of the state equation is:

$$\begin{cases} -\text{div}(\nu(\boldsymbol{\mu})\nabla y) + \boldsymbol{\beta}(\boldsymbol{\mu}) \cdot \nabla y = u\chi_{\Omega_u} & \text{in } \Omega, \\ y = 0 & \text{on } \Gamma_D, \\ \nu(\boldsymbol{\mu}) \dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N, \end{cases}$$

then the bilinear forms $a(\cdot,\cdot) : Y \times Q \to \mathbb{R}$ and $c(\cdot,\cdot) : U \times Q \to \mathbb{R}$ are defined, respectively, as

$$\begin{aligned} a(y, q, \boldsymbol{\mu}) &= \int_\Omega (\nu(\boldsymbol{\mu})\nabla y \cdot \nabla q + \boldsymbol{\beta}(\boldsymbol{\mu}) \cdot \nabla y q)\, d\Omega, \\ c(u, q) &= \int_{\Omega_u} uq\, d\Omega. \end{aligned}$$

As usual $G(\boldsymbol{\mu}) \in Q^*$ represents forcing terms and boundary conditions, whereas $m(\cdot,\cdot) : Y \times Y \to \mathbb{R}$ and $n(\cdot,\cdot) : U \times U \to \mathbb{R}$ are the bilinear forms associated to the cost functional. Let us recall the properties of the spaces and of the boundary conditions, already discussed in subsection 4.3.3

- Ω is an open, bounded and regular domain, with Lipschitz boundary $\partial\Omega = \Gamma_D \cup \Gamma_N$ and $\Gamma_D \cap \Gamma_N = \emptyset$,
- $U = L^2(\Omega)$,
- $Y = H^1_{\Gamma_D}(\Omega) = \{y \in H^1(\Omega) \ : \ y_{|\Gamma_D} = 0\}$,
- $Y = Q$,
- $y_d \in L^2(\Omega)$ is given,
- the diffusivity term verifies $\nu(\boldsymbol{\mu}) > 0$ in Ω,
- advective field $\boldsymbol{\beta} = \boldsymbol{\beta}(\boldsymbol{\mu})$ in $L_2(\Omega) \times L_2(\Omega)$ is given,
- we impose homogeneous Dirichlet boundary conditions on Γ_D,
- we impose homogeneous Neumann conditions on Γ_N.

As we specified in example 1.3.2.2, it is possible to formulate the optimal control system and find the solutions. Then one can apply POD technique on a Finite Element discretization to have a RB approximation and solve the OCP($\boldsymbol{\mu}$) problem (as a reference see [59]). We recall that the following properties are verified:

- $a : Y \times Q \to \mathbb{R}$ is coercive over Q for some specific values of $\boldsymbol{\mu}$ [8],
- $a : Y \times Q \to \mathbb{R}$ is continuous over $Y \times Q$, $\forall \boldsymbol{\mu} \in \mathcal{P}$,
- $c : U \times Q \to \mathbb{R}$ is continuous over $U \times Q$,$\forall \boldsymbol{\mu} \in \mathcal{P}$,
- $n : U \times U \to \mathbb{R}$ is continuous, symmetric over $U \times U$ and coercive over U, $\forall \boldsymbol{\mu} \in \mathcal{P}$,
- $m : Y \times Y \to \mathbb{R}$ is symmetric, continuous and positive in the norm of the observation space, $\forall \boldsymbol{\mu} \in \mathcal{P}$.

We decided to use a partitioned POD and to recast the problem (5.2.1) in a saddle-point formulation. In the examples, we will propose the structure already used in subsection 4.3.3. Analytically, we want to structure the problem in the following general way: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(x(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \in X \times Q$ such that

$$\begin{cases} \mathcal{A}(x, w; \boldsymbol{\mu}) + \mathcal{B}(w, p; \boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), w \rangle & \forall w \in X, \\ \mathcal{B}(x, q; \boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), q \rangle & \forall q \in Q, \end{cases} \tag{5.2.2}$$

where $X = Y \times U$, $x = (y, u), w = (z, v) \in X$. In this specific case $\mathcal{A} : X \times X \to \mathbb{R}$ and $\mathcal{B} : X \times Q \to \mathbb{R}$ are defined as follows:

$$\begin{aligned} \mathcal{A}(x, w; \boldsymbol{\mu}) &= \mathcal{A}(x, w) = m(y, z) + \alpha n(u, v), \\ \mathcal{B}(w, q; \boldsymbol{\mu}) &= a(z, q; \boldsymbol{\mu}) - c(v, q). \end{aligned}$$

In all the applications we implemented we had $G(\boldsymbol{\mu}) \equiv 0$, that is that we have no forcing terms and homogeneous boundary conditions, whereas

$$\langle F(\boldsymbol{\mu}), w \rangle = \langle F, w \rangle = m(y_d, z).$$

For the discretized version of the problem a Finite Element approximation can be used. Analogously to the system (5.1.11), the discrete version of advection-diffusion the $OCP(\boldsymbol{\mu})$ problem reads: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(x^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}(\boldsymbol{\mu})) \in X^{\mathcal{N}} \times Q^{\mathcal{N}}$ such that

$$\begin{cases} \mathcal{A}(x^{\mathcal{N}}(\boldsymbol{\mu}), w^{\mathcal{N}}) + \mathcal{B}(w^{\mathcal{N}}, p^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu}) = \langle F, w^{\mathcal{N}} \rangle, & \forall w^{\mathcal{N}} \in X^{\mathcal{N}}, \\ \mathcal{B}(x^{\mathcal{N}}(\boldsymbol{\mu}), q^{\mathcal{N}}; \boldsymbol{\mu}) = 0 & \forall q^{\mathcal{N}} \in Q^{\mathcal{N}}. \end{cases} \tag{5.2.3}$$

The reduced version of the problem is built trough a POD algorithm in order to obtain the following system: given $\boldsymbol{\mu} \in \mathcal{P}$ find $(x_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in X_N \times Q_N$ such that

$$\begin{cases} \mathcal{A}(x_N(\boldsymbol{\mu}), w_N) + \mathcal{B}(w_N, p_N(\boldsymbol{\mu}); \boldsymbol{\mu}) = \langle F, w_N \rangle, & \forall w_N \in X_N, \\ \mathcal{B}(x_N(\boldsymbol{\mu}), q_N; \boldsymbol{\mu}) = 0 & \forall q_N \in Q_N. \end{cases} \tag{5.2.4}$$

In the following subsections we will present two specific applications and we will give more informations on how the POD reduction was exploited.

[8] As we have seen in Theorem 1.2.9, to guarantee the the stability of the saddle-point formulation we have to require the coercivity of the bilinear form $a(\cdot, \cdot)$, that we have only for chosen parameters, see [20, Section 3.5] and [57, Chapter 12].

5.2.2 Pollutant Control Test

In this subsection we are going to apply RB approach to an OCP($\boldsymbol{\mu}$) governed by advection-diffusion equation. This test is preliminary to the example proposed in subsection 5.2.3. The problem studied is a parametric marine adaptation of the one in [57, Subsection 17.13.3], that involves three real control variables.
First of all, let us present the domain considered for this particular example.

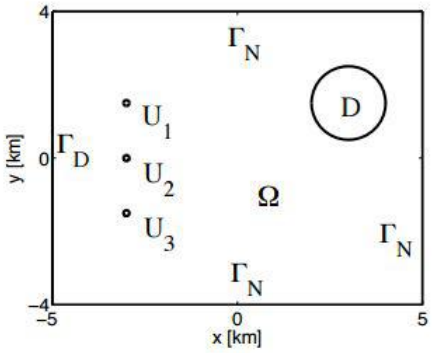

Figure 5.2.2.1: Domain considered for the pollutant control test, from [57, Subsection 17.13.3].

In figure 5.2.2.1 we can notice:

1. the observation domain $\Omega_{OBS} := D$, where the pollutant threshold is evaluated;
2. $\Omega_1 := U_1$, $\Omega_2 := U_2$ and $\Omega_3 := U_3$ are the three pollutant areas linked to the control variables u_1, u_2 and u_3, respectively;
3. on Γ_D homogeneous Dirichlet boundary conditions are applied;
4. on Γ_N homogeneous Neumann boundary conditions are applied.

Following the general formulation presented in subsection 5.2.1, we can build a specific problem that reads: given $\boldsymbol{\mu} \in \mathcal{P}$ find $(y(\boldsymbol{\mu}), u(\boldsymbol{\mu})) \in Y \times U$, such that

$$\min_{(y,u)\in Y\times U} J(y,u) = \frac{1}{2}\int_{\Omega_{OBS}} (y-y_d)^2 \, d\Omega_{OBS} + \frac{\alpha_1}{2}\int_{\Omega_1} u_1^2 \, d\Omega_1 + \frac{\alpha_2}{2}\int_{\Omega_2} u_2^2 \, d\Omega_2 + \frac{\alpha_3}{2}\int_{\Omega_3} u_3^2 \, d\Omega_3$$
$$\text{such that } a(y,q;\boldsymbol{\mu}) = c(u,q), \qquad \forall q \in Q.$$

where the state y is the pollutant concentration and $y_d \in \mathbb{R}$ is the desired concentration of pollutant, that usually represents a safety threshold. Let $u = [u_1, u_2, u_3]$ be an element of $U \equiv \mathbb{R}^3$, the control space. The bilinear form $a : Y \times Q \to \mathbb{R}$ and $c : U \times Q \to \mathbb{R}$ are defined as follows

$$a(y,q,\boldsymbol{\mu}) = \int_\Omega (\mu_1 \nabla y \cdot \nabla q + \mu_2 \boldsymbol{\beta} \cdot \nabla y q) \, d\Omega,$$
$$c(u,q) = u_1 \int_{\Omega_1} q \, d\Omega_1 + u_2 \int_{\Omega_2} q \, d\Omega_2 + u_3 \int_{\Omega_3} q \, d\Omega_3.$$

The first component of the parameter $\boldsymbol{\mu} = [\mu_1, \mu_2] \in \mathcal{P} = [0.1, 1.] \times [0.1, 3.]$ represents the diffusivity action, while μ_2 is a constant that changes the intensity of advection transport

field $\boldsymbol{\beta}$. To recast the problem in a saddle-point framework (5.2.2), we have to define the bilinear forms $m : Y \times Y \to \mathbb{R}$ and $n : U \times U \to \mathbb{R}$ as:

$$m(y, z) = \int_{\Omega_{OBS}} yz \, d\Omega_{OBS},$$

$$n(u, v) = \alpha_1 \int_{\Omega_1} u_1 v_1 \, d\Omega_1 + \alpha_2 \int_{\Omega_2} u_2 v_2 \, d\Omega_2 + \alpha_3 \int_{\Omega_3} u_3 v_3 \, d\Omega_3.$$

The next step is defining the product space of state and control variables, i.e. $X = Y \times U$. Let us consider $x = (y, u), w = (z, v) \in X$ and $q \in Q$. For this specific test case the bilinear forms $\mathcal{A} : X \times X \to \mathbb{R}$ and $\mathcal{B} : X \times Q \to \mathbb{R}$ are defined in the following way:

$$\mathcal{A}(x, w) = m(y, z) + n(u, v),$$
$$\mathcal{B}(w, q; \boldsymbol{\mu}) = a(z, q; \boldsymbol{\mu}) - c(v, q).$$

The linear form $F : X \to \mathbb{R}$ is defined as follows:

$$\langle F, w \rangle = y_d \int_{\Omega_{OBS}} z \, d\Omega_{OBS}.$$

In this way we can build the saddle point system (5.2.2), introduced in subsection 5.2.1, and then, after a Finite Element approximation (5.2.3), the reduced version of the saddle-point problem is structured, corresponding to the formulation (5.2.4), thanks to a partitioned POD algorithm with 50 basis functions generated on a training set of 100 points. The parameters have been sampled with an uniform distribution. To build the reduced problem, we used the space Z_N for both state and adjoint variable, where

$$Z_N = \text{span } \{\zeta_n := y^{\mathcal{N}}(\boldsymbol{\mu}^n), \ \xi_n := p^{\mathcal{N}}(\boldsymbol{\mu}^n), \ n = 1, \dots, N\}.$$

The control space remains $\mathbb{R}^3$. To be sure that the RB method is efficient, we have to guarantee the affinity hypotheses. Let us underline the affine structure of this specific OCP($\boldsymbol{\mu}$). With $Q_{\mathcal{A}} = 1$, $Q_{\mathcal{B}} = 2$ and $Q_F = 1$ the affine decomposition of the problem is given by

$$\begin{aligned}
&\Theta^1_{\mathcal{A}} = 1 && \mathcal{A}^1(x, w) = \mathcal{A}(x, w), \\
&\Theta^1_{\mathcal{B}} = \mu_1 && \mathcal{B}^1(x, q) = \int_\Omega \nabla y \cdot \nabla q \, d\Omega, \\
&\Theta^2_{\mathcal{B}} = \mu_2 && \mathcal{B}^2(x, q) = \int_\Omega \boldsymbol{\beta} \cdot \nabla y q \, d\Omega \\
&\Theta^1_F = 1 && \langle F^1, w \rangle = \langle F, w \rangle.
\end{aligned}$$

Let us show some numerical results corresponding to the specific choice of $\boldsymbol{\mu} = (1., 2.5)$. The desired concentration has the value $yd = 100$, and the maximum concentration for the three control variables is $U = 8 \cdot 10^5$. The transport field has been considered constant: $\boldsymbol{\beta} = \left(\cos\left(\frac{\pi}{30}\right), \sin\left(\frac{\pi}{30}\right)\right)$. The control variables have the following values:

$$u_1 = 8.8838 \cdot 10^3, \qquad u_2 = 7.4169 \cdot 10^3, \qquad u_3 = 6.9423 \cdot 10^4.$$

Figure 5.2.2.2 shows a comparison between the uncontrolled concentration of pollutant and the controlled solution of the full order and the reduced order approximation, respectively. Let us analyse the performance of the RB method with respect to the full order approximation. In figure 5.2.2.3 the pointwise error difference between the *truth* approximation and the reduced one is represented. The maximum value reached is of the order of 10^{-7}.

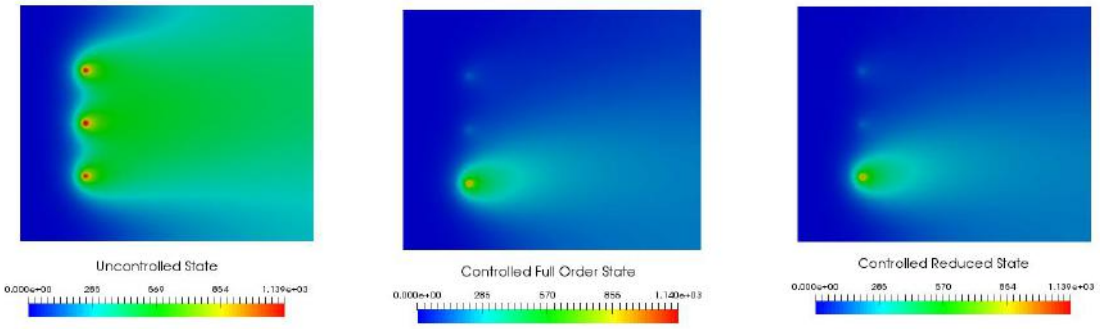

Figure 5.2.2.2: *Left:* uncontrolled state; *center*: full order controlled state; *right*:reduced order controlled state.

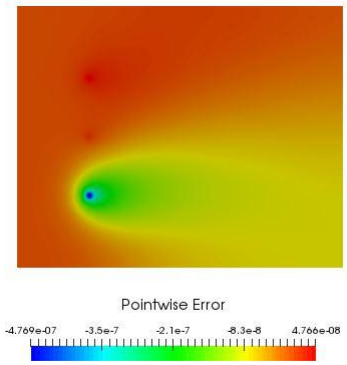

Figure 5.2.2.3: Pointwise error that compares full order solution and reduced order solution.

Figure 5.2.2.4 shows the decay of the error norms comparison between reduced solutions and full order solutions (see footnote 7 and extend the definition to all the variables). As usual we report the speed up index in Table 5.3: it is very convenient to use RB methods since they save computational time. The reduced system has a lower dimension with respect to the full order one and this allows the system to be more affordable, computationally speaking. Let us compare time of resolution: $t_t = 5.26s$, while $t_r = 6.95 \cdot 10^{-2}s$. The functional of the uncontrolled problem is $J = 5.93919 \cdot 10^4$. When we add control conditions the *truth* cost functional J_t and the cost functional associated to the reduced problem J_r reach the same value $1,02442 \cdot 10^3$.

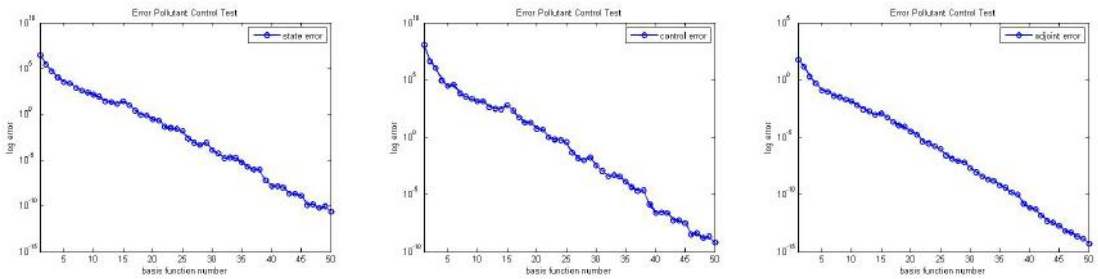

Figure 5.2.2.4: *left:* state error; *center*: control error; *right*: adjoint error.

Table 5.3: Speed up analysis for pollutant control test

Basis Number	1	2	3	4	5	6	7	8	9	10
Speed up	953	998	990	987	956	958	930	933	894	892
Basis Number	11	12	13	14	15	16	17	18	19	20
Speed up	868	846	824	800	775	755	742	724	708	698
Basis Number	21	22	23	24	25	26	27	28	29	30
Speed up	685	665	649	633	611	609	589	562	534	522
Basis Number	31	32	33	34	35	36	37	38	39	40
Speed up	514	493	483	458	436	424	411	400	388	376
Basis Number	41	42	43	44	45	46	47	48	49	50
Speed up	369	330	338	324	311	303	290	273	275	269

In the following a comparison between the dimensions of full order system and the reduced one is shown.

$\mathcal{N} \times \mathcal{N}$	6823×6823
$4N + 3 \times 4N + 3$	203×203

5.2.3 Pollutant control on the Gulf of Trieste

In this subsection we are going to apply partitioned POD method to a pollutant control on the Gulf of Trieste. We have simulated a pollutant loss from a marine accident. First of all, we had to build the physical domain. As we did for the Ocean, we had to create the mesh from Google Earth image thanks to *FreeFem++* (as a reference see [32] and visit the link *http://www.freefem.org/*) and, as in the Atlantic Ocean example, thanks to *Gmsh* (see as a reference [27], and visit *http://gmsh.info/*) we imported it into FEniCS (see [45] and for further informations one can refer to *https://fenicsproject.org*). To have an idea of the global process see figure 5.2.3.1.

Figure 5.2.3.1: *Left*: mesh; *right*: gulf of Trieste, *bottom*: subdomains considered: in red Ω_{OBS} and in green Ω_u.

The problem is formulated in the following way: let us define the state and the control spaces Y, U and the adjoint space $Y = Q$ as we did in subsection 5.2.1, the non dimensional

OCP($\boldsymbol{\mu}$) reads: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(y(\boldsymbol{\mu}), u(\boldsymbol{\mu})) \in Y \times U$ such that:

$$\begin{aligned} &\min_{(y,u)\in Y\times U} J(y,u) = \frac{1}{2}\int_{\Omega_{OBS}} (y - y_d)^2 \, d\Omega_{OBS} \\ &\text{such that } a(y,q;\boldsymbol{\mu}) = c(u,q), \qquad \forall q \in Q. \end{aligned} \tag{5.2.5}$$

where the state y is the pollutant concentration and $y_d = 0.2 \in \mathbb{R}$ represents the safe concentration of pollutant. The control variable is $u \in \mathbb{R}$. The bilinear forms $a : Y \times Q \to \mathbb{R}$ and $c : U \times Q \to \mathbb{R}$ are defined as:

$$\begin{aligned} a(y,q,\boldsymbol{\mu}) &= \int_\Omega (\nu(\boldsymbol{\mu})\nabla y \cdot \nabla q + \boldsymbol{\beta}(\boldsymbol{\mu}) \cdot \nabla y q) \, d\Omega, \\ c(u,q) &= L\, u \int_{\Omega_u} q \, d\Omega_u. \end{aligned}$$

where $\nu(\boldsymbol{\mu}) \equiv \mu_1$ represent the diffusivity action of the state equation, $\boldsymbol{\beta}(\boldsymbol{\mu}) = [\beta_1(\mu_2), \beta_2(\mu_3)]$ is the transport field and $\boldsymbol{\mu} = [\mu_1, \mu_2, \mu_3]$ represents our parameter. The constant $L = 10^3$, multiplied for the control variable $u \in \mathbb{R}$, make the system non dimensional. For the transport field we decided to take into consideration only constant functions in the proximity of the observation domain [9]. They will have the following form:

$$\beta_1(\mu_2) \equiv \mu_2, \qquad \beta_2(\mu_3) \equiv \mu_3.$$

The parameter space considered is $\mathcal{P} = [0.5, 1] \times [-1, 1] \times [-1, 1]$. Let us spend some words about the choice of the subdomains. The right plot of the figure 5.2.3.1 shows them: in green we have the zone of the domain where the pollutant is (in our mathematical formulation it is represented by Ω_u); the red part of the plot represents the observation domain Ω_{OBS}, positioned along the swimming touristic area and Miramare natural area. This particular zone is of great interest for two reasons:

1. for the peculiar ecological *flora* and *fauna* marine population,
2. and because it is an area crowded by Trieste citizens inhabitants and from many tourists.

These argumentations encourage us in the choice of Ω_{OBS}.
The boundary conditions are specified in the bulleted list of subsection 5.2.1. The coasts are considered in Γ_D, while the open sea represents Γ_N. To recast all the problem in the framework presented in (5.2.2) we have to define the bilinear forms $m : Y \times Y \to \mathbb{R}$ and $n : U \times U \to \mathbb{R}$ as follows:

$$m(y,z) = \int_{\Omega_{OBS}} yz \, d\Omega_{OBS},$$

$$n(u,v) \equiv 0.$$

As usual let $X = Y \times U$ the product space of state and control variables. Let $x = (y, u)$ and $w = (z, v)$ be elements of X, whereas q an element of Q. In this particular example, the bilinear forms $\mathcal{A} : X \times X \to \mathbb{R}$ and $\mathcal{B} : X \times Q \to \mathbb{R}$ are defined in the following way:

$$\begin{aligned} \mathcal{A}(x,w) &= m(y,z), \\ \mathcal{B}(w,q;\boldsymbol{\mu}) &= a(z,q;\boldsymbol{\mu}) - c(v,q). \end{aligned}$$

[9] They will be sufficient to simulate the most interesting configurations for the transport field action on the Gulf of Trieste, and we are not taken into consideration the dynamics at the boundary.

The linear form $F : X \to \mathbb{R}$ reads:

$$\langle F, w \rangle = y_d \int_{\Omega_{OBS}} z \, d\Omega_{OBS}.$$

Now we have all we need to apply a Finite element discretization on the system (5.2.1) and obtaining the discrete system (5.2.3) and a new OCP($\boldsymbol{\mu}$) that reads: given $\boldsymbol{\mu} \in \mathcal{P}$, find $(x^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}(\boldsymbol{\mu})) \in X^{\mathcal{N}} \times Q^{\mathcal{N}}$ such that

$$\begin{cases} \mathcal{A}(x^{\mathcal{N}}(\boldsymbol{\mu}), w^{\mathcal{N}}) + \mathcal{B}(w^{\mathcal{N}}, p^{\mathcal{N}}(\boldsymbol{\mu}); \boldsymbol{\mu}) = \langle F, w^{\mathcal{N}} \rangle, & \forall w^{\mathcal{N}} \in X^{\mathcal{N}}, \\ \mathcal{B}(x^{\mathcal{N}}(\boldsymbol{\mu}), q^{\mathcal{N}}; \boldsymbol{\mu}) = 0 & \forall q^{\mathcal{N}} \in Q^{\mathcal{N}}. \end{cases} \tag{5.2.6}$$

As we have already mentioned before, to reach a RB approximation we used a partitioned POD algorithm with $N = 50$ reduced basis functions and a training set of 100 point. To sample the parameters we used an uniform distribution on the space of the parameters. We decided to apply the technique of aggregated space and to compare it to a monolithic POD approach. Then, the state and the adjoint spaces will be approximated with the same space defined as:

$$Z_N = \text{span } \{\zeta_n := y^{\mathcal{N}}(\boldsymbol{\mu}^n), \ \xi_n := p^{\mathcal{N}}(\boldsymbol{\mu}^n), \ n = 1, \dots, N\}.$$

The reduced control space U_N is $\mathbb{R}$. Then, supposing $Y = Z_N$ and $Q = Z_N$, whereas $U = U_N$, the reduced problem reads:

$$\begin{cases} \mathcal{A}(x_N(\boldsymbol{\mu}), w_N) + \mathcal{B}(w_N, p_N(\boldsymbol{\mu}); \boldsymbol{\mu}) = \langle F, w_N \rangle, & \forall w_N \in X_N, \\ \mathcal{B}(x_N(\boldsymbol{\mu}), q_N; \boldsymbol{\mu}) = 0 & \forall q_N \in Q_N. \end{cases} \tag{5.2.7}$$

We can underline the affine structure of the problem: with $Q_{\mathcal{A}} = 1$, $Q_{\mathcal{B}} = 3$ and $Q_F = 1$ the affine decomposition of the problem is given by

$$\begin{aligned} &\Theta^1_{\mathcal{A}} = 1 && \mathcal{A}^1(x, w) = \mathcal{A}(x, w), \\ &\Theta^1_{\mathcal{B}} = \mu_1 && \mathcal{B}^1(x, q) = \int_\Omega \nabla y \cdot \nabla q \, d\Omega, \\ &\Theta^2_{\mathcal{B}} = \mu_2 && \mathcal{B}^2(x, q) = \int_\Omega \frac{\partial y}{\partial x_1} q \, d\Omega, \\ &\Theta^3_{\mathcal{B}} = \mu_3 && \mathcal{B}^3(x, q) = \int_\Omega \frac{\partial y}{\partial x_2} q \, d\Omega \\ &\Theta^1_F = 1 && \langle F^1, w \rangle = \langle F, w \rangle. \end{aligned}$$

Let us discuss some numerical results. First of all we solved an uncontrolled advection-diffusion problem for a maximum value of pollutant $u_{max} = 1$. We have studied the optimal control problem in three classical configurations:

1. neutral wind condition,
2. Bora blowing condition,
3. Scirocco blowing condition.

Naturally, the configuration is given by the values that the parameters μ_2, μ_3 may assume. Our model involves only the surface of the Gulf (the depth can be totally neglected since it is dozens meters deep). In the following we will present the solutions of the three configurations.

Neutral Configuration

The first experiment we analysed is the case with no wind blowing. This condition is given by the transport field $\boldsymbol{\beta}(\boldsymbol{\mu}) = [0, 0]$. In this particular case the OCP($\boldsymbol{\mu}$) is governed by a Laplace equation and has no advection term. This is a purely diffusive control problem. The diffusion term is given by $\mu_1 = 1$. In figure 5.2.3.2 we have the comparison between the uncontrolled concentration of pollutant and the controlled concentration in the Finite Element approximation and the reduced solution. In the bottom right of the figure a pointwise error is presented: the maximum value reached is of the order of 10^{-12}. The uncontrolled functional is $J = 1.7579 \cdot 10^{-3}$. Let us indicate with J_t the cost functional related to the Finite Element approximation, while J_r is the cost functional value derived from the reduced problem formulation (this convention will be used in the other two experiments, also). In this neutral case they have the following value: $J_t = J_r = 5.1320 \cdot 10^{-5}$. Our control variable is $u = 7.6901 \cdot 10^{-1}$.

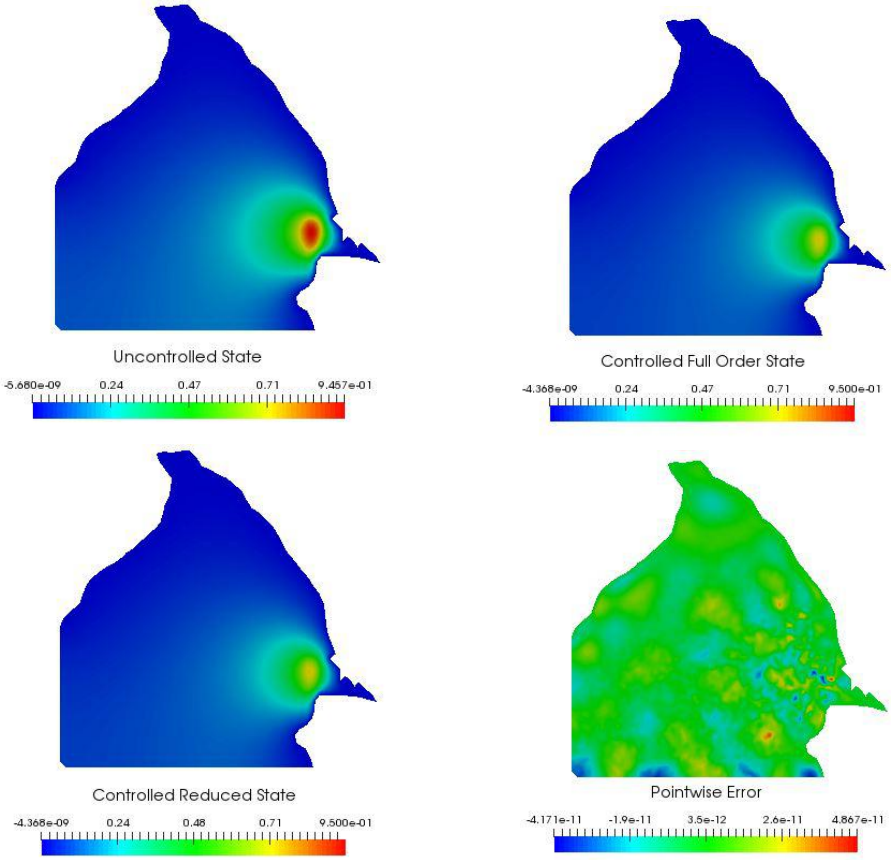

Figure 5.2.3.2: No wind configuration. *Top left*: uncontrolled state; *top right*: full order controlled state, *bottom left*: reduced controlled state, *bottom right*: pointwise error.

Bora Configuration

The second experiment shows how the physical results change under the action of wind. The peculiar wind blowing on the city of Trieste is the *Bora*. It is a cold wind that blows from East to North-West. To simulate his action we decided to take a constant transport field $\boldsymbol{\beta}(\boldsymbol{\mu}) = [-1, 1]$, that exactly simulate the water transport due to Bora blowing. In

this sense we expect that the pollutant has a minor concentration on the source domain, since the Bora effect leads the current to remove waters from the observation and source domains. For these reasons a lower value is obtained for J, J_t and J_r: the pollutant in Ω_{OBS} is less than the concentration observed in the previous configuration and in the Scirocco configuration. In Figure 5.2.3.3 we have the solution plots referred to uncontrolled state, controlled full order state and reduced controlled state. In the bottom right we find the pointwise error describing the difference between full order and reduced order pollutant concentration: the maximum value is of the order of 10^{-11}.

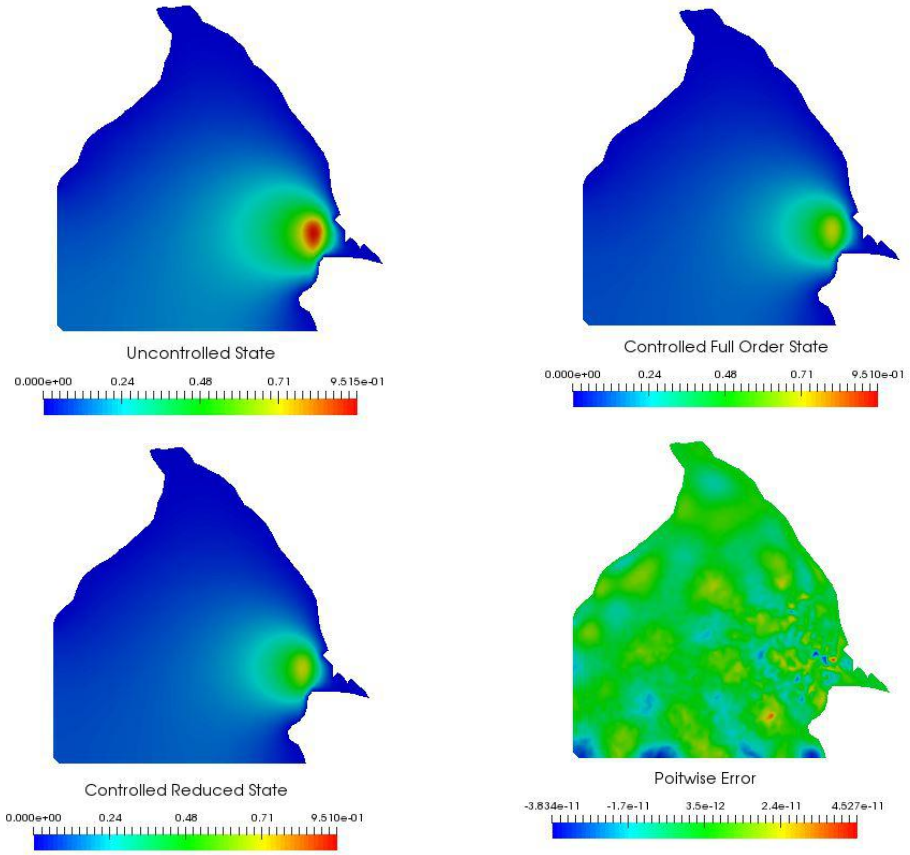

Figure 5.2.3.3: Bora configuration. *Top left*: uncontrolled state; *top right*: full order controlled state, *bottom left*: reduced controlled state, *bottom right*: pointwise error.

The uncontrolled functional assume the same value than before: $J = 1.7579 \cdot 10^{-3}$, while J_t and J_r reach both $4.9167 \cdot 10^{-5}$. The value of our control variable is $u = 7.3698 \cdot 10^{-1}$.

Scirocco Configuration

The last experiment presents the result of this control problem under the action of Scirocco. It is a warm wind that comes from South-East. To simulate its action we decided to take a constant transport field $\boldsymbol{\beta}(\boldsymbol{\mu}) = [1, -1]$. The net water transport is direct toward South and so this do not allow the dispersion of pollutant in the open Adriatic sea, as Bora does. What we expect is an higher value of the control variable u and and higher value of the cost functional, since more pollutant is present in the observation domain. In Figure 5.2.3.4 uncontrolled state, controlled full order state and reduced controlled state are shown. As in the previous two experiments, in the bottom right we find the pointwise error describing

the difference between full order and reduced order solutions: the maximum value is of the order of 10^{-11}.

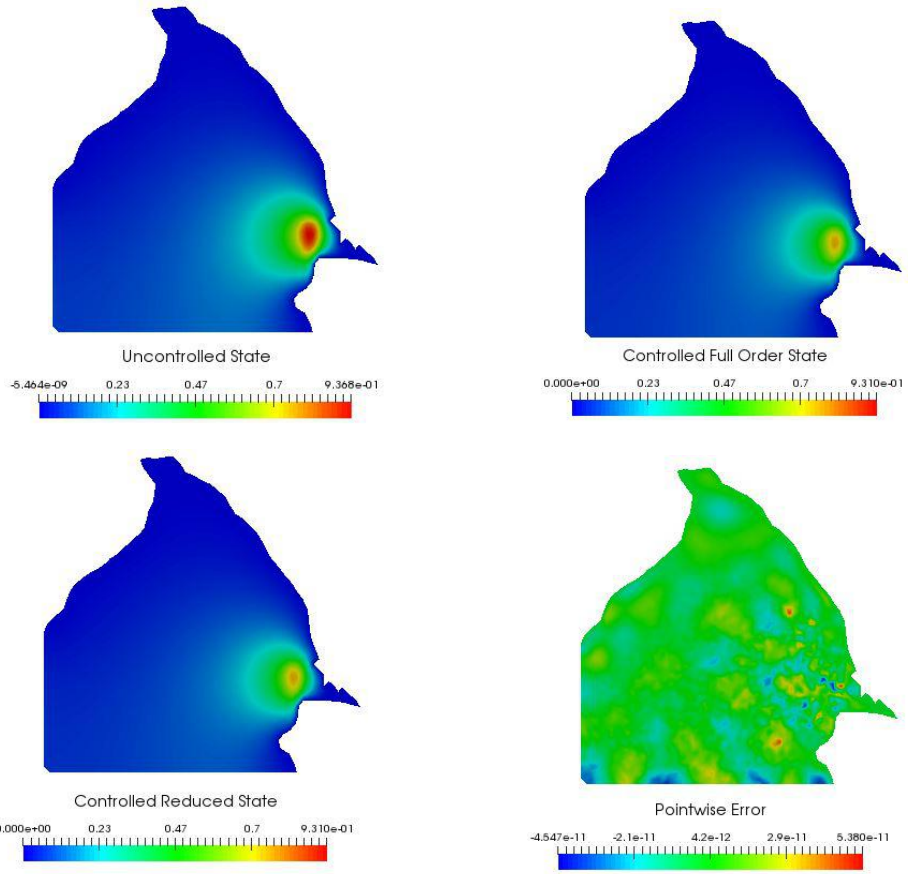

Figure 5.2.3.4: Scirocco configuration. *Top left*: uncontrolled state; *top right*: full order controlled state, *bottom left*: reduced controlled state, *bottom right*: pointwise error.

The uncontrolled functional assume the same value than before: $J = 1.7579 \cdot 10^{-3}$, while J_t and J_r reach both $5.3417 \cdot 10^{-5}$. The value of our control variable is $u = 8.0800 \cdot 10^{-1}$. Now let us analyse the error norm comparison between full order solutions and the reduced solutions. The plot in figure 5.2.3.5 shows the error norm decay for the state, control and adjoint variables. They are related to the choice $\boldsymbol{\mu} = [1, -1, 1]$, that is Bora configuration. Even if we used 50 basis functions to build the reduced space, we can notice that few of them were sufficient to reach a good approximation of the full order solutions. The bottom right plot shows the comparison between the monolithic error version and the partitioned error with respect to the full order and reduced state variable. Some comments on Table 5.4.: the *speed up index* remains considerably high. solving the reduced system is very convenient in this case, since the full order system is characterized by high dimensionality derived by the mesh. The time of resolution of the full order and the reduced order systems are $t_t = 2.79s$ and $t_r = 2.41 \cdot 10^{-2}s$, respectively.

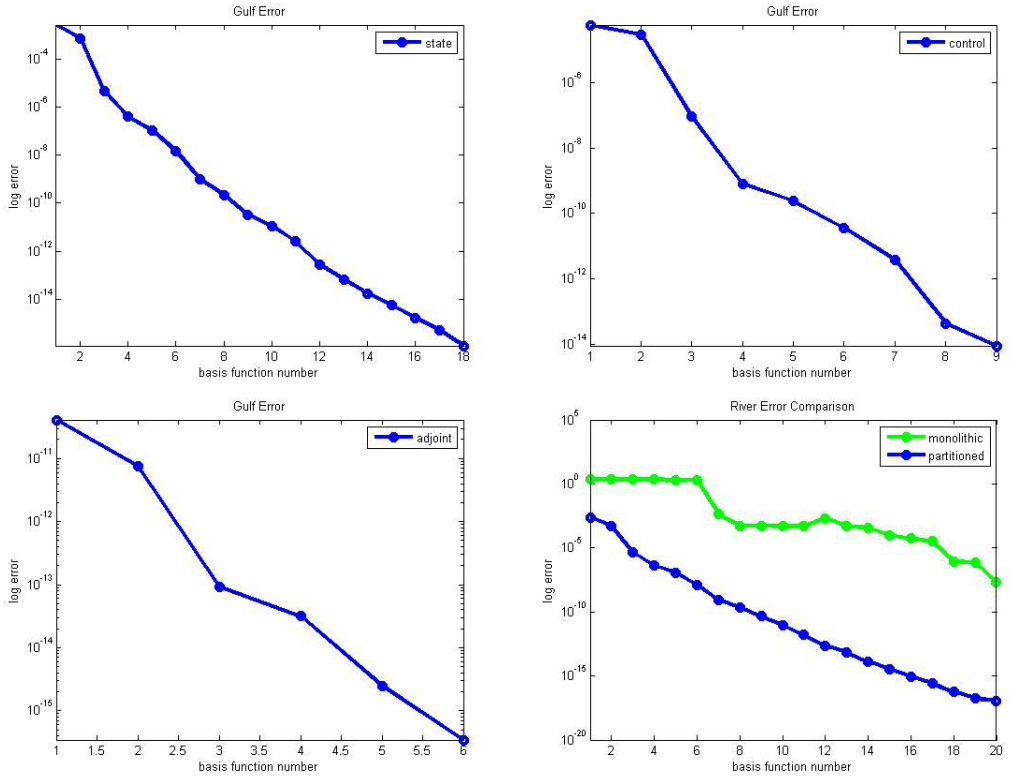

Figure 5.2.3.5: Errors. *Top left*: state error; *top right*: control error , *bottom left*: adjoint error, *bottom right*: monolithic and partitioned error comparison.

Table 5.4: Speed up analysis for Gulf of Trieste example

Basis Number	1	2	3	4	5	6	7	8	9	10
Speed up	361	380	369	366	364	362	352	348	346	350
Basis Number	11	12	13	14	15	16	17	18	19	20
Speed up	336	334	333	328	317	315	309	303	301	296
Basis Number	21	22	23	24	25	26	27	28	29	30
Speed up	294	288	282	277	271	264	267	257	251	245
Basis Number	31	32	33	34	35	36	37	38	39	40
Speed up	240	234	231	218	216	204	201	202	195	188
Basis Number	41	42	43	44	45	46	47	48	49	50
Speed up	183	175	167	161	157	152	147	147	134	135

In the following a comparison between the dimensions of full order system and the reduced one is shown.

$\mathcal{N} \times \mathcal{N}$	5639×5639
$4N + 1 \times 4N + 1$	201×201

Conclusions

The interest in parametric analysis of the gulf of Trieste is twofold:

1. it has a very peculiar dynamic, due to the particular wind circulation of the North-Eastern zone of Italy. This means that we have to consider different scenarios, typical of the region. Naturally, a model that has to simulate various solutions needs several parameters and then RB methods can be a useful and powerful instrument to handle them;

2. it is peculiar for urban and natural resources. Trieste is a city that arises on the seaside and it is linked to its gulf under many aspects: tourism, economy, industry, ecology and biodiversity, etc. For this reasons the study of the gulf can lead to several important results, involving different aspects of citizens' life and natural environment.

Concluding, pollutant substances are, obviously, dangerous (i.e. see [21]): the unhealthy effects could damage not only the peculiar *flora* and *fauna* of the Gulf, but, indirectly, also human beings: Trieste is a city that overlooks the seaside and many of its activities depends on it. As we have seen, many variables have to be taken into considerations, the morphology of the Gulf, the structure of the city and the weather conditions, very peculiar in this particular zone. In this specific example we focused on physical parametrization, but in this case one can consider also geometrical parametrization and study different phenomena. Reduced basis methods can be very versatile in this context, where many parameters are involved and query optimal control problems could be formulated. In this particular field of application, as in Oceanographic example, simulations can be computationally demanding and costly: RB techniques could be a good and viable way to reduce this issue.

Perspectives

In this appendix we are going to show some results on parametrized optimal control problems governed by PDEs (OCP($\boldsymbol{\mu}$)), that could be better developed in future.
The natural prosecution of this study, is to extend the concept proposed among the chapters of this master thesis in more general frameworks. The initial purposes of this work was to analyse not only linear quadratic optimal control problems, but also non-linear and time dependent OCP($\boldsymbol{\mu}$). Naturally, these two extensions are of great importance in engineering applications, as we can see in [19, 30, 23, 14, 13, 55], for example.
In Oceanographic applications, while the non-linearity could be neglected, time dependency is really important since the research is interested in temporal evolution of the dynamics and of the phenomena related to it. In the small scale contexts, the non-linearity has a major influence: fluids dynamics is usually modeled by Navier-Stokes equations. One of the main focus of scientific investigation in this field is linked on the the effects of non-linearity related to Reynolds number.
Among this work we bumped into OCP($\boldsymbol{\mu}$) with non-linear state equations. In the following sections we will describes some numerical results linked to the little steps made in non-linearity and time dependency, respectively.

Non-Linearity

Among this work, we tried to implement and study also OCP($\boldsymbol{\mu}$) governed by non-linear PDEs. In this section we are going to show some results obtained in this direction. The reduction was not implemented, but we ran some high fidelity simulations.
Initially, our attention focused on steady quasi-geostrophic non-linear equation and steady Navier Stokes equation, exploited in the geophysical version. We were able also to build state solution tracking non-linear OCP($\boldsymbol{\mu}$) governed by quasi-geostrophic equation. The results that we are going to show are derived from finite element approximation.

Steady Quasi-Geostrophic Equation: State Equation and Control

We recall the non-linear version of the steady quasi-geostrophic equation, already introduced in subsection 5.1.1. Considering $\Omega = [0,1] \times [0,1]$, Ocean circulation is described by the following non-linear PDE.

$$\begin{cases} \left(\frac{\delta_I}{L}\right)^2 \mathcal{F}(\psi, \Delta\psi) + \frac{\partial \psi}{\partial x} = f - \frac{\delta_S}{L}\Delta\psi + \left(\frac{\delta_M}{L}\right)^3 \Delta^2\psi & \text{in } \Omega, \\ \psi = 0 & \text{on } \partial\Omega, \\ \Delta\psi = 0 & \text{on } \partial\Omega, \end{cases} \tag{1}$$

where the non-linearity is defined by:

$$\mathcal{F}(\psi, q) = \frac{\partial \psi}{\partial x}\frac{\partial q}{\partial y} - \frac{\partial \psi}{\partial y}\frac{\partial q}{\partial x},$$

where ψ is in suitable function space Y, $L = 10^6$ is a dimensional parameter, δ_I is the non-linearity parameter, while δ_M, δ_S are diffusivity parameters and the forcing term is $f = -\sin(\pi y)$. As we did in subsection 5.1.1, we can impose $q = \Delta\psi$ and the system becomes

$$\begin{cases} q = \Delta\psi & \text{in } \Omega, \\ \left(\frac{\delta_I}{L}\right)^2 \mathcal{F}(\psi, q) + \frac{\partial \psi}{\partial x} = u - \frac{\delta_S}{L} q + \left(\frac{\delta_M}{L}\right)^3 \Delta q & \text{in } \Omega, \\ \psi = 0 & \text{on } \partial\Omega, \\ q = 0 & \text{on } \partial\Omega. \end{cases} \tag{2}$$

For the simulations we had to build the weak formulation of the problem. Let us consider $\psi, q \in Y = H_0^1(\Omega) \times H_0^1(\Omega)$. Thanks to this assumptions and exploiting integration by part and divergence theorem we can reach the following weak formulation for the problem (2): $\forall (\phi, p) \in V$, find $(\psi, q) \in V$ such that verify

$$\begin{cases} \int_\Omega qp + \int_\Omega \nabla\psi \cdot \nabla p = 0 \\ \left(\frac{\delta_I}{L}\right)^2 \int_\Omega (\psi \nabla q \times \hat{\mathbf{k}}) \cdot \nabla\phi + \int_\Omega \frac{\partial \psi}{\partial x}\phi + \left(\frac{\delta_M}{L}\right)^3 \int_\Omega \nabla q \cdot \nabla\phi + \frac{\delta_S}{L} \int_\Omega q\phi = f \end{cases}$$

where $f = -\sin(\pi y)$ and exploiting the relation (see [11, Appendix A])

$$\mathcal{F}(\psi, q) = \operatorname{div}(\psi \nabla q \times \hat{k}).$$

We studied the dynamics of this PDE in a framework of weak non-linearity: in other words, when the diffusivity and the non-linear parameters are comparable. Indeed, when $\delta_I \gg \delta_M$, the system is unstable and needs a stabilization. We propose some numerical results, some of them have been already introduced in subsection 5.1.1. In figure 1 three configuration are shown, for different choice of parameters. The diffusivity parameter $\delta_S = 0$ in all the experiments, while:

- on the left we choose $\delta_I = 0$ and $\delta_M = 7 \cdot 10^4$;
- the plot in the center shows the non-linear solution for $\delta_I = \delta_M = 7 \cdot 10^4$;
- on the right the unstable result for $\delta_I = 7 \cdot 10^4$ and $\delta_M = 7 \cdot 10^3$.

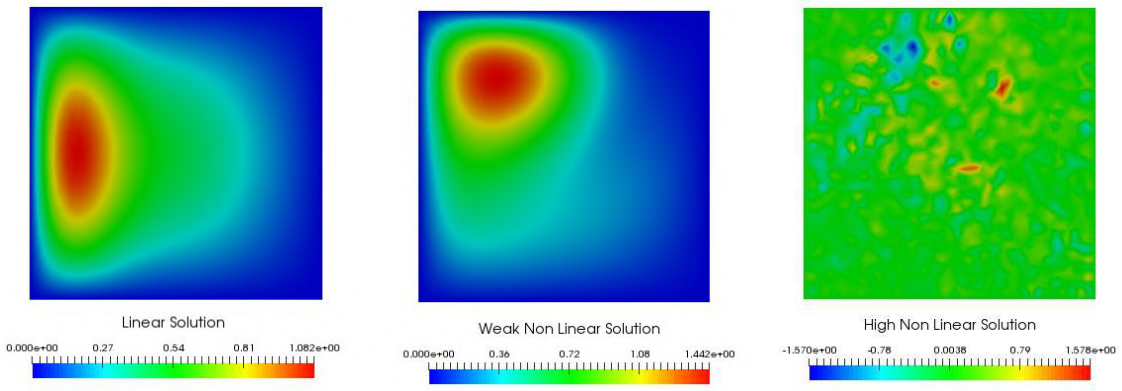

Figure 1: *Left*: linear solution; *center*: weak nonlinear solution, *right*: high nonlinear solution.

The next step was to implement a non-linear state tracking control problem. The problem formulation is totally similar to (5.1.5). We already know that the state space is Y, the control space is U and the adjoint space is $Q = Y$. The OCP split version reads:

$$\min_{(\psi,u)\in Y\times U} J(\psi,u) = \frac{1}{2}\int_\Omega (\psi-\psi_d)^2\,d\Omega + \frac{\alpha}{2}\int_\Omega u^2\,d\Omega$$

$$\text{such that} \quad \begin{cases} q = \Delta\psi & \text{in } \Omega, \\ \dfrac{\partial\psi}{\partial x} + \Big(\dfrac{\delta_I}{L}\Big)^2\mathcal{F}(\psi,q) + \Big(\dfrac{\delta_S}{L}\Big)q - \Big(\dfrac{\delta_M}{L}\Big)^3\Delta q = u & \text{in } \Omega, \\ \psi = 0 & \text{on } \partial\Omega, \\ q = 0 & \text{on } \partial\Omega. \end{cases} \tag{3}$$

Let us define $(\psi_{adj}, q_{adj}) \in Q$ as the adjoint variables. Thanks to Lagrangian approach we have built the following optimality system.

$$\begin{cases} a_{adj}((\psi_{adj},q_{adj}),(\chi,t)) = -(\psi-\psi_d,\chi)_{L^2(\Omega)} & \forall(\chi,t)\in Y, \\ (\alpha u, v)_{L^2(\Omega)} = c(v,(\psi_{adj},q_{adj})) & \forall v\in U, \\ a((\psi,q),(\phi,p)) = c(u,(\phi,p)) & \forall(\phi,q)\in Q, \end{cases} \tag{4}$$

where $a : Y\times Q \to \mathbb{R}$ is defined as the sum of the weak formulation of the right hands sides of state equation in the constrain of (3)[10]. The bilinear form $c : U\times Q\to\mathbb{R}$ is defined as

$$c(u,(\phi,p)) = \int_\Omega u\phi\,d\Omega.$$

The bilinear form $a_{adj} : Y\times Q\to\mathbb{R}$, thanks to integration by parts, divergence's theorem and thanks to the hypotheses made on the functions, is defined as follows:

$$\begin{aligned} a_{adj}((\psi_{adj},q_{adj}),(\chi,t)) = &-\int_\Omega \frac{\partial\psi_{adj}}{\partial x}\chi\,d\Omega + \Big(\frac{\delta_M}{L}\Big)^3\int_\Omega \nabla q_{adj}\cdot\nabla\chi\,d\Omega + \\ &+\Big(\frac{\delta_S}{L}\Big)\int_\Omega q_{adj}\chi\,d\Omega + \Big(\frac{\delta_I}{L}\Big)^2\int_\Omega \mathcal{F}(\chi,q)\psi_{adj}\,d\Omega + \\ &+\Big(\frac{\delta_I}{L}\Big)^2\int_\Omega \mathcal{F}(\psi,\chi)q_{adj}\,d\Omega + \\ &+\int_\Omega q_{adj}t\,d\Omega + \int_\Omega \nabla q_{adj}\cdot\nabla t\,d\Omega, \end{aligned}$$

Let us show some numerical results in the case of weak non-linearity. We use a Finite Element discretization $\mathbb{P}^1-\mathbb{P}^1$ both for state and adjoint variables, to simulate the control problem. In figure 2 a comparison between the desired state, the Finite Element solution. The right plot shows the pointwise error of the difference between desired state and state obtained: the maximum value reached is $1.703\cdot10^{-2}$, whereas the value of the cost functional is $J = 1.0960\cdot10^{-5}$.

[10]The consistency with the strong formulation of the state equation can be shown as we did in subsection 5.1.1.

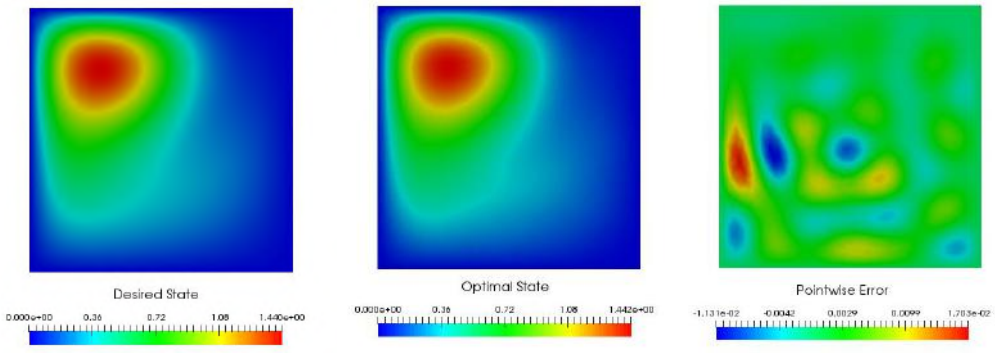

Figure 2: *Left*: desired state; *center*: control problem solution, *right*: pointwise error.

Steady Geophysical Navier Stokes State Equation

In order to understand what is the link between Quasi-Geostrophic equation and the geophysical Navier Stokes dynamics, we decided to simulate it and to compare the results obtained with these two different approaches. We recall the strong formulation of the PDE (5.1.2), describing large scale fluid motion under the influence of Earth's rotation. In this particular example, we considered $\delta_S = 0$. Our purpose is to find $u \in V = H_0^1(\Omega) \times H_0^1(\Omega)$ and $p \in P = L_0^2(\Omega)$ where

$$L_0^2(\Omega) = \left\{ r \in L^2(\Omega) \ : \ \int_\Omega r = 0 \right\},$$

such that:

$$\begin{cases} \left(\frac{\delta_I}{L}\right)^2 (\mathbf{u}\cdot\nabla)u - (1+y)u_2 = -\frac{\partial p}{\partial x} + \left(\frac{\delta_M}{L}\right)^3 \left(\frac{\partial^2 u_1}{\partial x^2} + \frac{\partial^2 u_1}{\partial y^2}\right) + f_1 & \text{in } \Omega, \\ \left(\frac{\delta_I}{L}\right)^2 (\mathbf{u}\cdot\nabla)v + (1+y)u_1 = -\frac{\partial p}{\partial y} + \left(\frac{\delta_M}{L}\right)^3 \left(\frac{\partial^2 u_2}{\partial x^2} + \frac{\partial^2 u_2}{\partial y^2}\right) + f_2 & \text{in } \Omega, \\ \operatorname{div}(\mathbf{u}) = 0 & \text{on } \partial\Omega, \\ \mathbf{u} = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.2.8}$$

in this case, $\mathbf{u} = (u_1, u_2)$ and the forcing term $\mathbf{f} = (f_1, f_2)$ is linked to the wind stress and is taken as $\mathbf{f} = \Big(-\frac{1}{\pi}\cos(\pi y), 0 \Big)$. For the simulations we have considered a Finite Element approximation, with a $\mathbb{P}^2 - \mathbb{P}^1$ discretization for both state and adjoint variables. The weak formulation of the problem reads: find $(\mathbf{u}, p)$ such that:

$$\begin{cases} a(\mathbf{u},\mathbf{v}) + b(\mathbf{v},p) = \mathbf{f} & \forall \mathbf{v} \in V, \\ b(\mathbf{u},q) = 0 & \forall q \in P, \end{cases} \tag{5.2.9}$$

where the bilinear forms $a : V \times V \to \mathbb{R}$ and $b : V \times P \to \mathbb{R}$ are defined as follows:

$$\begin{aligned} a(\mathbf{u},\mathbf{v}) =& \left(\frac{\delta_I}{L}\right)^2 \int_\Omega (\mathbf{u}\cdot\nabla)\mathbf{u}\, d\Omega - \int_\Omega (1+y)u_2 v_1 \, d\Omega \ + \\ &+ \int_\Omega (1+y)u_1 v_2 \, d\Omega + \left(\frac{\delta_M}{L}\right)^3 \int_\Omega \nabla\mathbf{u}\cdot\nabla\mathbf{v}\, d\Omega, \\ b(\mathbf{v},p) =& -\int_\Omega \operatorname{div}(\mathbf{v}) p \, d\Omega. \end{aligned}$$

Let us show some results. As in the quasi-geostrophic case, the simulation is very unstable for $\delta_I \gg \delta_M$. A stabilization is needed. Figure 3 shows the velocity fields in two configurations: in the left we have the velocity solution for $\delta_I = 0$ and in the center the weak non-linear velocity solution corresponding to $\delta_I = \delta_M = 7 \cdot 10^4$.

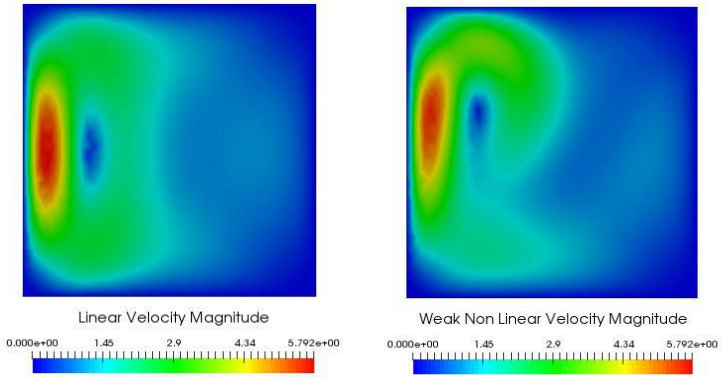

Figure 3: *Left*: linear velocity; *right*: weak non-linear velocity.

As we expect, in the linear case we have a thickening of the current toward the west boundary. When we add the non-linear effect, the velocity fields moves Northward.

Time Dependency

Time dependency is another topic of great importance in science and engineering simulations. In our work, we wanted to exploit it in Oceanographic application: one could focus on the study of the evolution of the quasi-geostrophic equations and the builing of a time dependent tracking control problem to be inserted in a data assimilation context. In the following subsection we will show time dependent evolution of quasi-geostrophic equation and of geophysical Nevier-Stokes.

Time Dependent Quasi-Geostrophic Equation State Equation

One of the major spark of this master thesis was the study of data assimilation model. In other words we aim at simulating a time dependent dynamics and at some time steps making an optimization on the solution in order to modify the parameters considered: in this way a forecasting model could be more precise and reliable. The first step we did in this direction was to simulate the time evolution of the quasi-geostrophic equation.
Let consider the usual domain $\Omega = [0,1] \times [0,1]$. The split time dependent quasi-geostrophic PDEs read as follows: find $(\psi, q) \in V = H_0^1(\Omega) \times H_0^1(\Omega)$ such that

$$\begin{cases} q = \Delta\psi & \text{in } \Omega, \\ \dfrac{\partial q}{\partial t} + \left(\dfrac{\delta_I}{L}\right)^2 \mathcal{F}(\psi, q) + \dfrac{\partial \psi}{\partial x} - \left(\dfrac{\delta_M}{L}\right)^3 \Delta^2 q + \dfrac{\delta_S}{L} q = -\sin(\pi y) & \text{in } \Omega, \\ q = 0 & \text{on } \partial\Omega, \\ \psi = 0 & \text{on } \partial\Omega. \end{cases} \tag{5.2.10}$$

As we did in the previous section, we can derive the weak formulation. It reads: find $(\psi, q) \in V$ such that for every $(\phi, p) \in V$:

$$\begin{cases} \int_\Omega qp + \int_\Omega \nabla\psi \cdot \nabla p = 0 \\ \int_\Omega \frac{\partial q}{\partial t}\phi + \left(\frac{\delta_I}{L}\right)^2 \int_\Omega (\psi \nabla q \times \hat{\mathbf{k}}) \cdot \nabla\phi + \int_\Omega \frac{\partial \psi}{\partial x}\phi + \\ \qquad + \left(\frac{\delta_M}{L}\right)^3 \int_\Omega \nabla q \cdot \nabla\phi + \frac{\delta_S}{L} \int_\Omega q\phi = f. \end{cases} \tag{5.2.11}$$

As is the previous examples, we exploited a $\mathbb{P}^1 - \mathbb{P}^1$ Finite Element approximation for the state variable. For the time evolution we used an Implicit Euler Method, with initial conditions $q(t = 0, x) = 0$. Let us show some results for different configurations.

Linear case

Let us take $\Omega = [0, 1] \times [0, 1]$. We are going to present the results obtained with the following parameters: $\delta_I = 0$ and $\delta_M = 7 \cdot 10^4$. The time interval is $[0, 60]$ and the $dt = 10$ as time increment[11]. In figure 4 the time evolution at $t = 10, t = 30$ and $t = 60$, from left to right.

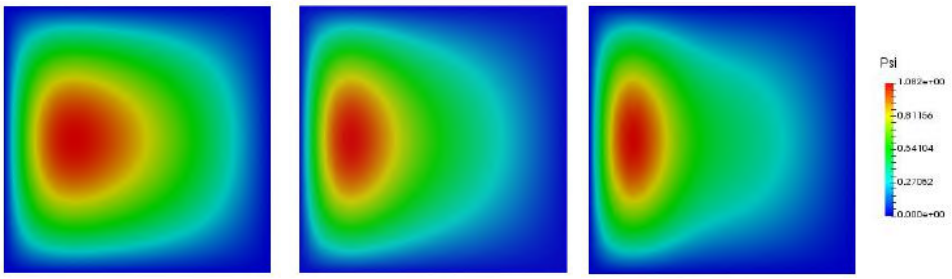

Figure 4: Time evolution of linear case on the square domain.

The same parameters are were used in the simulation on the Atlantic Ocean. The time interval of the simulation is $[0, 100]$, $dt = 10$. In figure 5 linear time evolution is presented for $t = 10$, $t = 60$ and $t = 100$, from left to right.

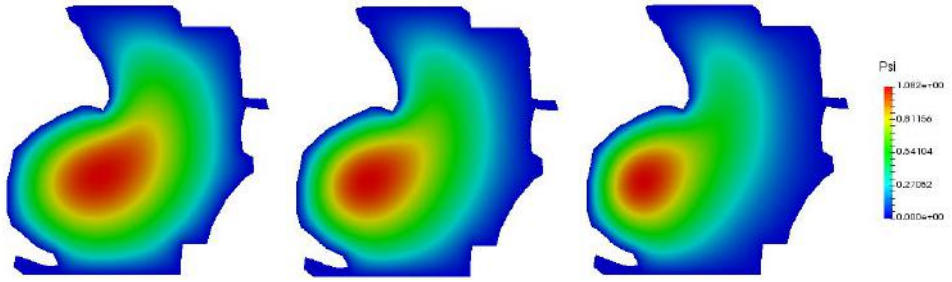

Figure 5: Time evolution of linear case on North Atlantic Ocean.

[11] A time increment of $dt = 10$ means 4 month of dynamics evolution.

Non-linear case

Let us take $\Omega = [0,1] \times [0,1]$. In this paragraph we will show the results for an high non-linear case: $\delta_I = 7 \cdot 10^4$ and $\delta_M = 7 \cdot 10^3$. The time interval is $[0, 130]$ and the $dt = 10$. In figure 6 the snapshots proposed are related to $t = 10, 60, 130$, from left to right.

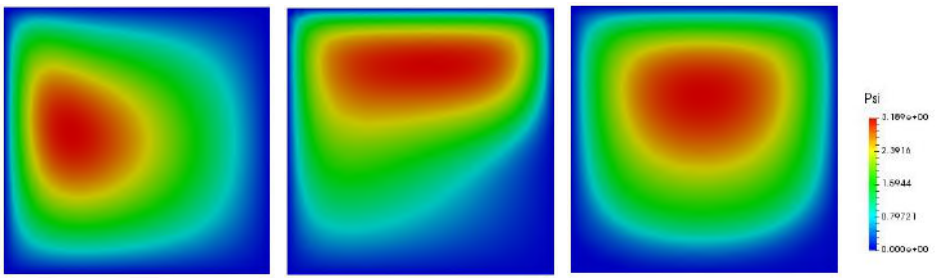

Figure 6: Time evolution of non-linear case on the square domain.

We simulated the same parameters on the North Atlantic mesh. The time interval in this case is $[0, 100]$, and the plots in figure 7 represent the ψ solutions at time $t = 10, 60, 100$, from left to right.

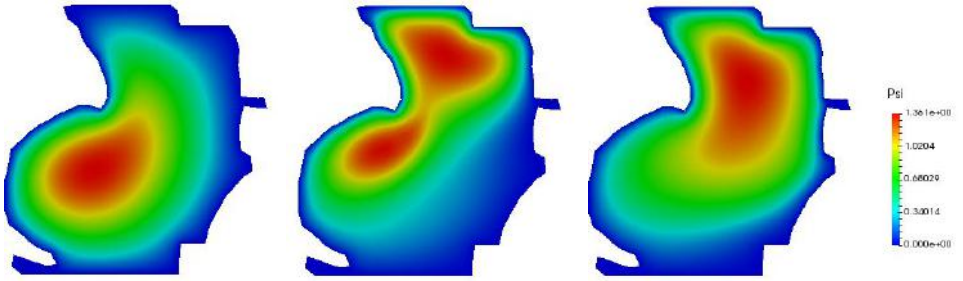

Figure 7: Time evolution of non-linear case on North Atlantic Ocean.

Time Dependent Geophysical Navier Stokes State Equation

The time dependence was studied also in the case of geophysical Navier Stokes equations. The unsteady formulation of the equations for $\delta_S = 0$ is the following: find $u \in V = H_0^1(\Omega) \times H_0^1(\Omega)$ and $p \in P = L_0^2(\Omega)$ such that

$$\begin{cases} \dfrac{\partial u_1}{\partial t} + \left(\dfrac{\delta_I}{L}\right)^2 (\mathbf{u} \cdot \nabla) u_1 - (1+y) u_2 = -\dfrac{\partial p}{\partial x} + \left(\dfrac{\delta_M}{L}\right)^3 \left(\dfrac{\partial^2 u_1}{\partial x^2} + \dfrac{\partial^2 u_1}{\partial y^2}\right) + f_1 & \text{in } \Omega \\ \dfrac{\partial u_2}{\partial t} + \left(\dfrac{\delta_I}{L}\right)^2 (\mathbf{u} \cdot \nabla) u_2 + (1+y) u_1 = -\dfrac{\partial p}{\partial y} + \left(\dfrac{\delta_M}{L}\right)^3 \left(\dfrac{\partial^2 u_2}{\partial x^2} + \dfrac{\partial^2 u_2}{\partial y^2}\right) + f_2 & \text{in } \Omega \\ \operatorname{div}(\mathbf{u}) = 0 & \text{on } \partial\Omega \\ \mathbf{u} = 0 & \text{on } \partial\Omega, \end{cases} \tag{5.2.12}$$

where $\mathbf{u} = (u_1, u_2)$. Also in this case we used a Finite Element discretization. The scheme used was a $\mathbb{P}^2 - \mathbb{P}^1$, for velocity variable and pressure, respectively. To simulate the time evolution, we exploit and implicit Euler method, with $\mathbf{u}(t = 0, x) = 0$. Let us show some result in linear and wek non-linear framework.

Linear case

As we did for quasi-geostrophic case, we analysed the case of linear solution with $\delta_I = 0$ and $\delta_M = 7 \cdot 10^4$. First of all we considered the square domain $\Omega = [0,1] \times [0,1]$. The time interval is [0,60] and $dt = 10$. From left to right, the plots in figure 8 show velocity field solutions for $t = 10, 30, 60$.

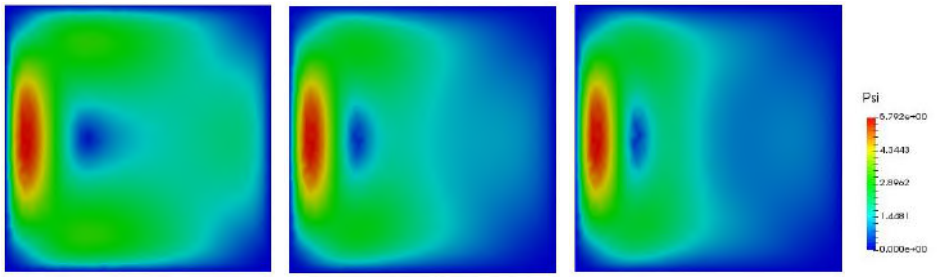

Figure 8: Linear geophysical Navier Stokes velocity on the squared domain.

The same physics and time parameters were used to understand linear time evolution in the North Atlantic Ocean. In figure 9 velocity plots are shown: from left to right $t = 10, 30, 60$ with $dt = 10$.

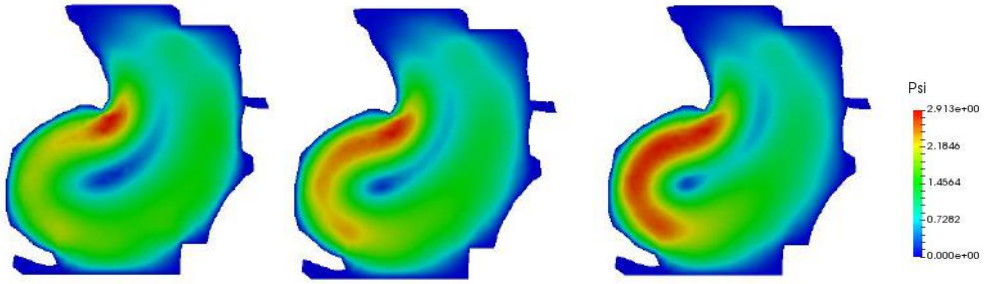

Figure 9: Linear geophysical Navier Stokes velocity on North Atlantic Ocean.

Non-linear case

Although we used no stabilization techniques, we were able to handle the weak linear case corresponding to the parameters $\delta_I = \delta_M = 7 \cdot 10^4$. The simulation in this case is made on the time interval $[0, 100]$, with the usual $dt = 10$. Figure 10 shows the evolution for $t = 10, 60, 100$.

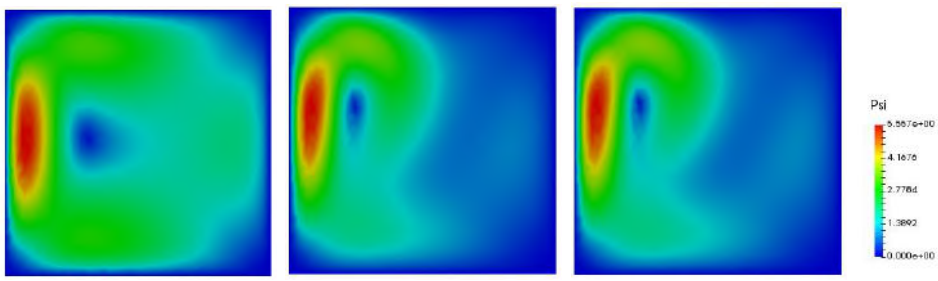

Figure 10: non-linear geophysical Navier Stokes velocity on the squared domain.

In figure 11 the results of the same experiment run on the North Atlantic Ocean mesh is

proposed, for the same temporal steps.

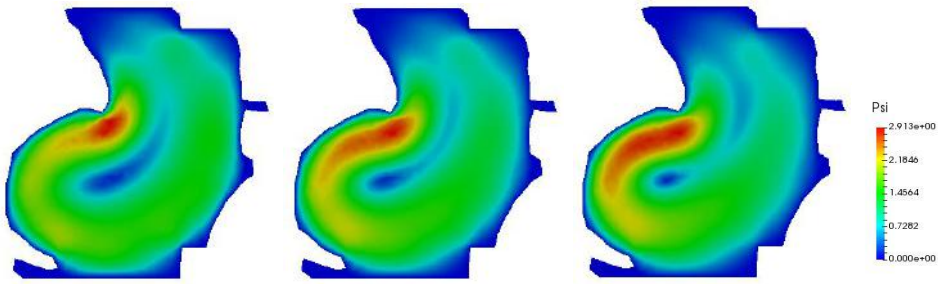

Figure 11: Non-linear geophysical Navier Stokes velocity on the Atlantic Ocean.

Conclusions and Future Developments

Among this appendix, we have presented the first results obtained in order to extend what we have discussed among this work to non-linear time dependent optimal control problems. At the end, let us expose purposes and intentions that we want to improve and complete as a natural development of this master thesis.

1. First of all, a deeper analysis of non-linear cases is required to build $OCP(\boldsymbol{\mu})$ governed by a general state equation. Then, we plan to deal with non-linear control problems, solve them with RB methods and compare this approach to the other discretization techniques, as Finite Element discretization.

2. Next step would involve the development of time dependent optimal control problems. They are of great importance in climatological applications, in order to forecast and predict future scenarios. Time dependency will make the problems more computational demanding. In this sense another objective is to apply model order reduction to save computational resources.

3. Naturally, another objective is to mix these two first points to build non-linear time dependent optimal control problems.

4. Finally, we would like to exploit what we previously underlined in environmental applications: in particular, this knowledge can be exploited in order to build a real data assimilation model. The other step is to reduce the problem and to compare RB performance with full order resolution.

Bibliography

[1] F. Ballarin, A. Manzoni, A. Quarteroni, and G. Rozza. Supremizer stabilization of POD–galerkin approximation of parametrized steady incompressible Navier–Stokes equations. *International Journal for Numerical Methods in Engineering*, 102(5):1136–1161, 2015.

[2] M. Barrault, Y. Maday, N. C. Nguyen, and A. T. Patera. An Empirical Interpolation Method: application to efficient reduced-basis discretization of partial differential equations. *Comptes Rendus Mathematique*, 339(9):667–672, 2004.

[3] D. W Behringer, M. Ji, and A. Leetmaa. An improved coupled model for ENSO prediction and implications for Ocean initialization. Part I: The Ocean data assimilation system. *Monthly Weather Review*, 126(4):1013–1021, 1998.

[4] T. R. Bewley. Flow control: new challenges for a new renaissance. *Progress in Aerospace sciences*, 37(1):21–58, 2001.

[5] P. B. Bochev and M. D. Gunzburger. *Least-squares finite element methods*, volume 166. Springer-Verlag, New York, 2009.

[6] D. Boffi, F. Brezzi, M. Fortin, et al. *Mixed finite element methods and applications*, volume 44. Springer-Verlag, Berlin and Heidelberg, 2013.

[7] S. Boyaval, C. Le Bris, Y. Maday, N. C. Nguyen, and A. Patera. A reduced basis approach for variational problems with stochastic parameters: Application to heat conduction with variable robin coefficient. *Computer Methods in Applied Mechanics and Engineering*, 198(41):3187–3206, 2009.

[8] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer Science & Business Media, New York, 2010.

[9] J. A. Carton, G. Chepurin, X. Cao, and B. Giese. A simple Ocean data assimilation analysis of the global upper Ocean 1950–95. Part I: Methodology. *Journal of Physical Oceanography*, 30(2):294–309, 2000.

[10] J. A. Carton and B. S. Giese. A reanalysis of ocean climate using Simple Ocean Data Assimilation (SODA). *Monthly Weather Review*, 136(8):2999–3017, 2008.

[11] F. Cavallini and F. Crisciani. *Quasi-geostrophic theory of Oceans and atmosphere: topics in the dynamics and thermodynamics of the Fluid Earth*, volume 45. Springer Science & Business Media, New York, 2013.

[12] S. Danilov, G. Kivman, and J. Schröter. A finite-element ocean model: principles and evaluation. *Ocean Modelling*, 6(2):125–150, 2004.

[13] J. C. de los Reyes and F. Tröltzsch. Optimal control of the stationary Navier-Stokes equations with mixed control-state constraints. *SIAM Journal on Control and Optimization*, 46(2):604–629, 2007.

[14] L. Dedè. Optimal flow control for Navier-Stokes equations: Drag minimization. *International Journal for Numerical Methods in Fluids*, 55(4):347–366, 2007.

[15] L. Dedè. Reduced basis method and a posteriori error estimation for parametrized linear-quadratic optimal control problems. *SIAM Journal on Scientific Computing*, 32(2):997–1019, 2010.

[16] L. Dedè. Adaptive and reduced basis method for optimal control problems in environmental applications. PhD thesis, Politecnico di Milano, 2008. Available at http://mox.polimi.it.

[17] M. C. Delfour and J. Zolésio. *Shapes and geometries: metrics, analysis, differential calculus, and optimization*, volume 22. SIAM, Philadelphia, 2011.

[18] J. I. Díaz (Editor). *Ocean Circulation and Pollution Control - A mathematical and Numerical Investigation.* Springer, Berlin and Heidelberg, 2004.

[19] J. A. Dutton. The nonlinear quasi-geostrophic equation: existence and uniqueness of solutions on a bounded domain. *Journal of the Atmospheric Sciences*, 31(2):422–433, 1974.

[20] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*, volume 159. Springer Science & Business Media, New York, 2013.

[21] J. Faganeli, M. E. Hines, M. Horvat, I. Falnoga, and S. Covelli. Methylmercury in the Gulf of Trieste (Northern Adriatic Sea): from microbial sources to seafood consumers. *Food Technology and Biotechnology*, 52(2):188, 2014.

[22] E. Fernández Cara and E. Zuazua Iriondo. Control theory: History, mathematical achievements and perspectives. *Boletín de la Sociedad Española de Matemática Aplicada, 26, 79-140.*, 2003.

[23] A. V. Fursikov, M. D. Gunzburger, and L. Hou. Boundary value problems and optimal boundary control for the Navier–Stokes system: the two-dimensional case. *SIAM Journal on Control and Optimization*, 36(3):852–894, 1998.

[24] M. Gad-el Hak, A. Pollard, and J. Bonnet. *Flow control: fundamentals and practices*, volume 53. Springer-Verlag, Berlin and Heidelberg, 2003.

[25] A. L. Gerner and K. Veroy. Reduced basis a posteriori error bounds for the Stokes equations in parametrized domains: a penalty approach. *Mathematical Models and Methods in Applied Sciences*, 21(10):2103–2134, 2011.

[26] A. L. Gerner and K. Veroy. Certified reduced basis methods for parametrized saddle point problems. *SIAM Journal on Scientific Computing*, 34(5):A2812–A2836, 2012.

[27] C. Geuzaine and J. F. Remacle. Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009.

[28] M. Ghil and P. Malanotte-Rizzoli. Data assimilation in meteorology and oceanography. *Advances in geophysics*, 33:141–266, 1991.

[29] M. D. Gunzburger. *Perspectives in flow control and optimization*, volume 5. SIAM, Philadelphia, 2003.

[30] M. D. Gunzburger, L. Hou, and Th. P. Svobodny. Analysis and finite element approximation of optimal control problems for the stationary Navier-Stokes equations with distributed and Neumann controls. *Mathematics of Computation*, 57(195):123–151, 1991.

[31] J. Haslinger and R. A. E. Makinene. *Perspectives in flow control and optimization.* SIAM, Philadelphia, 2003.

[32] F. Hecht. New development in FreeFem++. *J. Numer. Math.*, 20(3-4):251–265, 2012.

[33] R. Herzog and E. Sachs. Preconditioned conjugate gradient method for optimal control problems with control and state constraints. *SIAM Journal on Matrix Analysis and Applications*, 31(5):2291–2317, 2010.

[34] J. S. Hesthaven, G. Rozza, and B. Stamm. Certified reduced basis methods for parametrized partial differential equations. *SpringerBriefs in Mathematics*, 2015, Milano.

[35] M.L. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23. Springer Science & Business Media, Antwerp, 2008.

[36] J. A. Infante and E. Zuazua. Boundary observability for the space semi-discretizations of the 1–d wave equation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 33(2):407–438, 1999.

[37] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications*, volume 15. SIAM, Philadelphia, 2008.

[38] K. Ito and SS. Ravindran. A reduced-order method for simulation and control of fluid flows. *Journal of computational physics*, 143(2):403–425, 1998.

[39] E. Kalnay. *Atmospheric modeling, data assimilation and predictability.* Cambridge university press, 2003, Cambridge.

[40] T.Y. Kim, T: Iliescu, and E. Fried. B-spline based finite-element method for the stationary quasi-geostrophic equations of the Ocean. *Computer Methods in Applied Mechanics and Engineering*, 286:168–191, 2015.

[41] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza. Model order reduction in fluid dynamics: challenges and perspectives. In *Reduced Order Methods for modeling and computational reduction*, pages 235–273. Springer, 2014.

[42] G. Leugering, P. Benner, S. Engell, A. Griewank, H. Harbrecht, M. Hinze, R. Rannacher, and S. Ulbrich. *Trends in PDE constrained optimization.* Springer, New York, 2014.

[43] J. L. Lions. *Optimal Control of System Governed by Partial Differential Equations*, volume 170. Springer-Verlagr, Berlin and Heidelberg, 1971.

[44] J. L. Lions. *Some aspects of the optimal control of distributed parameter systems*. SIAM, Philadelphia, 1972.

[45] A. Logg, K.A. Mardal, and G. Wells. *Automated Solution of Differential Equations by the Finite Elemnt Mehod*. Springer-Verlag, Berlin, 2012.

[46] I. Martini, G. Rozza, and B. Haasdonk. Reduced basis approximation and a-posteriori error estimation for the coupled Stokes-Darcy system. *Advances in Computational Mathematics*, 41(5):1131–1157, 2015.

[47] R. Milani, A. Quarteroni, and G. Rozza. Reduced basis method for linear elasticity problems with many parameters. *Computer Methods in Applied Mechanics and Engineering*, 197(51):4812–4829, 2008.

[48] D. Modesto, S. Fernández-Méndez, and A. Huerta. Elliptic harbor wave model with perfectly matched layer and exterior bathymetry effects. *Journal of Waterway, Port, Coastal, and Ocean Engineering*, 2016.

[49] B. Mohammadi and O. Pironneau. *Applied shape optimization for fluids*. Oxford University Press, New York, 2010.

[50] R. Mosetti, C. Fanara, M. Spoto, and E. Vinzi. Innovative strategies for marine protected areas monitoring: the experience of the Istituto Nazionale di Oceanografia e di Geofisica Sperimentale in the Natural Marine Reserve of Miramare, Trieste-Italy. In *OCEANS, 2005. Proceedings of MTS/IEEE*, pages 92–97. IEEE, 2005.

[51] D. Nechaev, J. Schröter, and M. Yaremchuk. A diagnostic stabilized finite-element ocean circulation model. *Ocean Modelling*, 5(1):37–63, 2003.

[52] F. Negri. Reduced basis method for parametrized optimal control problems governed by PDEs. *Mester thesis, Politecnico di Milano*, 2010-2011.

[53] F. Negri, A. Manzoni, and G. Rozza. Reduced basis approximation of parametrized optimal flow control problems for the Stokes equations. *Computers & Mathematics with Applications*, 69(4):319–336, 2015.

[54] F. Negri, G. Rozza, A. Manzoni, and A. Quarteroni. Reduced basis method for parametrized elliptic optimal control problems. *SIAM Journal on Scientific Computing*, 35(5):A2316–A2340, 2013.

[55] T. Porsching. Estimation of the error in the reduced basis method solution of nonlinear equations. *Mathematics of Computation*, 45(172):487–496, 1985.

[56] M. Pošta and T. Roubíček. Optimal control of Navier–Stokes equations by Oseen approximation. *Computers & Mathematics With Applications*, 53(3):569–581, 2007.

[57] A. Quarteroni. *Numerical Models for Differential Problems*. Springer, Milano, 2014.

[58] A. Quarteroni, G. Rozza, L. Dedè, and A. Quaini. Numerical approximation of a control problem for advection-diffusion processes. In *IFIP Conference on System Modeling and Optimization*, pages 261–273, Ceragioli F., Dontchev A., Futura H., Marti K., Pandolfi L. (eds) System Modeling and Optimization. CSMO 2005. vol 199. Springer, Boston, 2005.

[59] A. Quarteroni, G. Rozza, and A. Quaini. Reduced basis methods for optimal control of advection-diffusion problems. In *Advances in Numerical Mathematics*, pages 193–216. RAS and University of Houston, Moscow, 2007.

[60] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*, volume 23. Springer Science & Business Media, Berlin and Heidelberg, 2008.

[61] A. Rösch and B. Vexler. Optimal control of the Stokes equations: A priori error analysis for finite element discretization with postprocessing. *SIAM Journal on Numerical Analysis*, 44(5):1903–1920, 2006.

[62] D. V. Rovas. *Reduced-basis output bound methods for parametrized partial differential equations*. PhD thesis, Massachusetts Institute of Technology, 2002.

[63] G. Rozza, DB. P. Huynh, and 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):115–152, 2013.

[64] G. Rozza, A. Manzoni, and F. Negri. Reduction strategies for PDE-constrained oprimization problems in Haemodynamics. ECCOMAS, Congress Proceedings, Vienna, September 2012.

[65] G. Rozza and K. Veroy. On the stability of the reduced basis method for Stokes equations in parametrized domains. *Computer methods in applied mechanics and engineering*, 196(7):1244–1260, 2007.

[66] J. Schöberl and W. Zulehner. Symmetric indefinite preconditioners for saddle point problems with applications to PDE-constrained optimisation problems. *SIAM Journal on Matrix Analysis and Applications*, 29(3):752–773, 2007.

[67] V. Schulz and I. Gherman. One-shot methods for aerodynamic shape optimization. In *MEGADESIGN and MegaOpt-German Initiatives for Aerodynamic Simulation and Optimization in Aircraft Design*, pages 207–220. Springer, 2009.

[68] T. Shiganova and A. Malej. Native and non-native ctenophores in the Gulf of Trieste, Northern Adriatic Sea. *Journal of Plankton Research*, 31(1):61–71, 2009.

[69] S. Taasan. *One shot methods for optimal control of distributed parameter systems 1: Finite dimensional control*. 1991.

[70] F. Tröltzsch. Optimal control of partial differential equations. *Graduate studies in mathematics*, 112, Verlag, Wiesbad, 2010.

[71] E. Tziperman and W. C. Thacker. An optimal-control/adjoint-equations approach to studying the Oceanic general circulation. *Journal of Physical Oceanography*, 19(10):1471–1485, 1989.

[72] G. Veronis. Wind-driven Ocean circulation Part 1. Linear theory and perturbation analysis. In *Deep Sea Research and Oceanographic Abstracts*, volume 13, pages 17–29. Elsevier, 1966.

[73] G. Veronis. Wind-driven Ocean circulation Part 2. Numerical solutions of the nonlinear problem. In *Deep Sea Research and Oceanographic Abstracts*, volume 13, pages 31–55. Elsevier, 1966.

[74] H. Yang, G. Lohmann, W. Wei, M.i Dima, M. Ionita, and J. Liu. Intensification and poleward shift of subtropical western boundary currents in a warming climate. *Journal of Geophysical Research: Oceans*, 121(7):4928–4945, 2016.

[75] K. Yosida. Functional analysis. Springer, Berlin and Heidelberg, 1995.

[76] O. C. Zienkiewicz. Achievements and some unsolved problems of the finite element method. *International Journal for Numerical Methods in Engineering*, 47(1-3):9–28, 2000.

[77] E. Zuazua. Propagation, observation, control and numerical approximation of waves. *SIAM Review*, 47(2):197–243, 2005.

[78] W. Zulehner. Nonstandard norms and robust estimates for saddle point problems. *SIAM Journal on Matrix Analysis and Applications*, 32(2):536–560, 2011.

Ringraziamenti

Giunta al termine del lavoro desidero rendere grazie a tutte le personalità che hanno reso possibile la stesura di questa tesi magistrale.

In primis, ringrazio il Professor Gianluigi Rozza per avermi dato la possibilità di realizzare questo progetto sotto la sua supervisione: ha sempre trovato il modo e il tempo non solo di seguire il mio lavoro, ma anche di indirizzarmi e sostenermi, con grande professionalità, premura e comprensione.

Grazie anche al Professor Renzo Mosetti. La sua disponibile e gentile supervisione è stata essenziale per la buona riuscita del lavoro. Le discussioni in ambito Oceanografico ed ambientale sono state fondamentali per la resa finale dell'interdisciplinarietà del progetto.

Un ringraziamento particolare va al Dr. Francesco Ballarin: grazie per tutte le volte che mi hai aiutato, grazie per tutti i consigli che mi hai dato, grazie per tutte le tue dritte su FEniCS e RBniCS, per avermi ascoltata, per aver sciolto i miei dubbi, piccoli o grandi che fossero, per avermi sempre accolta col sorriso e tanta voglia di fare.

Un grazie speciale va a tutto il gruppo *mathLab*. Lavorare con voi è stato un piacere: mi sono sentita parte di una squadra formidabile e affiatata. Grazie di avermi accolta, nonostante la mia tesi sia un umile contributo rispetto a tutti i vostri importanti progetti. Un ultimo grazie va all'intera sezione di Oceanografia dell'OGS, a tutte le persone che vi collaborano, che mi hanno sostenuta e spronata. Siete un *team* coeso e amichevole: non mi avete mai fatta sentire fuori luogo, mai un *pesce fuor d'acqua*.