SISSA
Scuola
Internazionale
Superiore di
Studi Avanzati

Mathematics Area - PhD course in
Mathematical Analysis, Modelling, and Applications

Model Order Reduction for Nonlinear and Time-Dependent Parametric Optimal Flow Control Problems

Candidate:
Maria Strazzullo

Advisor:
Prof. Gianluigi Rozza
Co-advisor:
Dr. Francesco Ballarin

Academic Year 2020-21

Alla mia famiglia

"Ne abbiamo attraversate di tempeste,
e quante prove antiche e dure."
Lode all'Inviolato
Franco Battiato

Abstract

The goal of this thesis is to provide an overview of the latest advances on reduced order methods for parametric optimal control governed by partial differential equations.
Historically, parametric optimal control problems are a powerful and elegant mathematical framework to fill the gap between collected data and model equations to make numerical simulations more reliable and accurate for forecasting purposes.
For this reason, parametric optimal control problems are widespread in many research and industrial fields. However, their computational complexity limits their actual applicability, most of all in a parametric nonlinear and time-dependent framework. Moreover, in the forecasting setting, many simulations are required to have a more comprehensive knowledge of very complex systems and this should happen in a small amount of time.
In this context, reduced order methods might represent an asset to tackle this issue. Thus, we employed space-time reduced techniques to deal with a wide range of equations. We propose a space-time proper orthogonal decomposition for nonlinear (and linear) time-dependent (and steady) problems and a space-time Greedy with a new error estimation for parabolic governing equations.
First of all, we validate the proposed techniques through many examples, from the more academic ones to a test case of interest in coastal management exploiting the Shallow Waters Equations model.
Then, we will focus on the great potential of optimal control techniques in several advanced applications. As a first example, we will show some deterministic and stochastic environmental applications, adapting the reduced model to the latter case to reach even faster numerical simulations. Another application concerns the role of optimal control in steering bifurcating phenomena arising in nonlinear governing equations. Finally, we propose a neural network-based paradigm to deal with the optimality system for parametric prediction.

Acknowledgements

First of all, I would like to thank my supervisor, Prof. Gianluigi Rozza, for his continuous support and for all the trust he placed in me. I am deeply grateful for his guidance and mentorship. He always encouraged me to give my best, being a role model for my daily life as a student.
I am profoundly thankful to my co-advisor, Dr. Francesco Ballarin, for sharing his knowledge and experience with kindness and patience. His support made our meetings fruitful learning moments and they will be a valuable asset to my future research life.
I thank the referees of my thesis, Prof. Claudio Canuto and Prof. Olga Mula, for revising this contribution.
I am also grateful to all the people I collaborated with during my studies. Among them, I thank Prof. Traian Iliescu, who kindly introduced me to a new topic with a proactive attitude.
I want to express my gratitude to all the colleagues who helped me in broadening my scientific horizons: thanks Marco Tezzele, Nicola Demo, Dr. Michele Girfoglio and Dr. Federico Pichi.
Lastly, I would like to thank the whole mathLab group for making me feel part of a prosperous scientific environment with common goals and values.

I kindly ask the reader not to blame me for using my mother tongue for more personal acknowledgments.

Inizio dalle persone che hanno rallegrato i momenti di lavoro, dai miei compagni di ufficio e di corridoio, vecchi, nuovi e ritrovati. Grazie Monica, Saddam, Laura, Ema, Ivan, Francesco, Anna Chiara, Sara, Ginevra: perché sebbene lo si pensi spesso, non si è mai davvero soli su questa barca e, se ci si lascia trasportare dal vento insieme, è tutto un po' meno spaventoso.
Un grazie speciale ad Elisa, per avermi arricchita con i suoi consigli, navigando con me per qualche tempo.
Un aiuto in mare sono state tutte le mie stelle fisse. Tra queste, Arianna e Michele, che hanno illuminato con la loro simpatia e profondità i momenti più bui. Negli astri poi, Francesca, c'è sempre stata: a volte l'ho persa, ma ogni volta che l'ho ritrovata, mi ha indicato la via della fiducia.
I momenti più spensierati di questo viaggio, quelli di mare calmo, sono pieni di risate, aneddoti e diari di bordo invidiabili. Ringrazio Luca, Fede e Matte per essere stati al mio fianco sempre e per avermi regalato i migliori ricordi. Tuttavia, può capitare di cadere in acque fredde. Quando sei in balia delle onde, ringrazi Luca che ti afferra la mano. Quando la corrente è troppo forte, ringrazi Fede e Matte che ti trascinano su di peso. Tutti lì, seduti sul ponte, ti scaldano e ti ricordano che i marinai coraggiosi sono quelli che smettono

di remare quando non si hanno più le energie. Sono dei compagni esemplari. Ognuno per motivi diversi, resterà nel bagaglio che un giorno aprirò a riva. In cuor nostro, sappiamo cosa ci lega ed è qualcosa di più forte di una tempesta. Grazie per aver salvato questa donna in mare, molte volte.
In questi momenti, mi volto indietro. Non posso non pensare ai miei amici di una vita, alla mie *pasturelle*, Alice, Betta, Gigio, Giulia, Dario, Martina. Loro ci sono sempre, mi accudiscono e mi cullano da lontano, con i pensieri. Devo ringraziarli perchè i viaggi non si fanno senza portafortuna, e loro sono il mio.
In queste acque sconfinate la mente si popola delle parole di mia madre, di mio padre e di mio fratello. Senza di loro non sarei mai diventata la donna che sono e non avrei mai remato così forte. Grazie per tutto il sostegno e l'amore che mi avete sempre dimostrato. Grazie per avermi insegnato ad emozionarmi con niente e stupirmi con meno.
Ora, che tante tempeste sono passate, sogno la riva a cui punta la prua. So che comunque vada a finire questo viaggio, a terra ci sarà Riccardo. Il mio porto, la mia spiaggia, il mio rifugio. Grazie perché mi ricordi il bello della terraferma.

Una postilla. Ringrazio tutte le donne eccezionali della mia vita. Tra queste, soprattutto, mia madre, per avermi insegnato che oltre ad essere *mamma*, si è sempre *donna*.

Trieste, September 2021

Contents

Introduction

In our daily life, we aim at describing processes to better understand the world around us. A powerful mathematical tool to describe physical phenomena consists of partial differential equations (PDEs). Whatever one is studying, it is clear that the response to some stimuli changes depending on what the features of the physical system are. Thus, the model can be more attested if it includes parameters in the formulation. Parameters can describe the physical and/or geometrical properties of the system. They reflect an input-output relation between what the focus is, i.e. the parameters, and what the event is, i.e. the parametric solution.
For this reason, we will focus on parametrized partial differential equations (PDE($\boldsymbol{\mu}$)s) as a valuable model to study several physical dynamics. Nonetheless, in such an analysis, a further modelling step can be taken to have a deeper comprehension of the experienced circumstances. For example, It is the case of forecasting modelling. Imagine that something dangerous or unpleasant is expected or predicted, say a flood, a pollutant loss in a protected area, a catastrophic event due to anthropic or natural causes, an accident in a factory. How can I act on the system to avoid an undesirable configuration? Can I change my model to be safer?
These are the roots of what is known as the *controllability theory*. A controlled PDE($\boldsymbol{\mu}$) is a parametric system on which some external variables called *controls* may act changing the usual and expected behaviour of the solution [41]. The goal of the controlled model is to steer its solution towards *a preferable configuration*.
However, is it always possible to reach an exact prescribed profile for a PDE($\boldsymbol{\mu}$)? Mathematically, this question is not only fascinating, but of interest in many applied fields. For the reader, we here report a list of contributions on this topic for linear problems [41, 54, 92, 94, 97]. It is indisputable that the controllability problem becomes more and more involved in the nonlinear setting. Indeed, the growing physical complexity related to such models leads to an increased need for control over the systems. The study of such a setting prospered in the eighties in the fluid dynamics fields thanks to J. L. Lions. His contributions paved the way to a broad literature production, see e.g. [7, 40, 93, 95, 96, 157].
However, it is natural to guess that not all the PDE($\boldsymbol{\mu}$)s are controllable: i.e. it is not possible to guarantee the existence of external variables that result in the *exact configuration* required. For example, is it possible to act on a pollutant loss to totally *stop* its spread? It is clear that, maybe an external control capable of such an action may not exists or, if any, it can be unfeasible or not physically meaningful. Thus, the controllability has been extended to the *optimal control theory*. In this setting, the goal is not to reach a configuration, but to be, somehow, *the most similar* to that peculiar profile, satisfying the underlying PDE($\boldsymbol{\mu}$).

The parametric optimal control framework is an elegant and versatile mathematical tool that has been employed in many applied fields in a *data assimilation* fashion: i.e. it is interpreted as a way to change the model to reach, eventually, the information given by collected data or some previous awareness of the system behaviour (consider, for example, historical series or synthetic knowledge given by other models). Parametric optimal control problems (OCP($\boldsymbol{\mu}$)s), are usually quite complicated and, thus, in most cases, it is not possible to provide an analytical solution for them. For this reason, scientists and researchers usually rely on approximated numerical simulations. Yet, even if OCP($\boldsymbol{\mu}$)s can be an asset to supply reliable forecasting simulations, they are computationally demanding and this limits their applicability. This issue is amplified when one is dealing with nonlinear and time-dependent equations, for which the required computational resources drastically increase and a parametric study can be unbearable to perform using standard discretization techniques. Because of this, we propose reduced order methods (ROMs) to obtain fast parametric simulations. In a reduced setting, the solution is sought for a large number of parameters. Some examples involve applications of interest in many fields such as parameter estimation, statistical analysis, real-time predictions and so on. The main idea underlying the model order reduction is to build on a previous model discretization, say the *high fidelity* model, to identify a suitable low dimensional and problem-dependent space that represents how the parametric system changes with respect to the parameters. Therefore, a reduced model is employed to simulate a parametric instance in a small amount of time without paying in accuracy with respect to the standard model. This goal is reached when the model order reduction is performed through the *offline-online paradigm*: once built the model in a (possibly costly) offline phase, each new parameter is evaluated and simulated in a fast way in an *online phase*, exploiting the previously collected, manipulated and stored offline information. For an introduction to the topic, the interested reader may refer to [10, 20, 64, 123, 125, 133, 134], for example.

Focusing on the optimal control framework, ROMs have been successfully employed in many applied fields, from fluid dynamics [42, 43, 110, 122], to heamodynamics [14, 90, 150, 168] and environmental sciences [33, 126, 127, 145, 147] and more.

This thesis proposes numerical strategies for nonlinear and time-dependent OCP($\boldsymbol{\mu}$)s that will be tested for several examples. In the reduced optimal control context, a wide production of literature concerns steady linear governing equations, see e.g. the following far-from-exhaustive list [12, 13, 44, 51, 71, 79, 80, 86, 110, 111, 127]. Thus, a natural extension to time-dependent problems followed. Indeed, building on the space-time formulation for non-parametric settings, see e.g. [65, 143, 144, 167], in [16, 70, 146, 148, 150], for example, they generalized the formulation and the standard algorithms to build reduced spaces to time-dependent OCP($\boldsymbol{\mu}$)s, highly accelerating the solution process. Few recent contributions can be found for nonlinear and time-dependent opti-

mal control in a parametric framework, see e.g. [16, 118, 147, 150, 168].
The main motivations behind this work were to show a versatile reduced order model to be adapted to several OCP($\boldsymbol{\mu}$)s, with a focus on interdisciplinary applications such as environmental ones. These are characterized by complicated model predictions based on data collection or expected solution profiles. In this context, OCP($\boldsymbol{\mu}$)s are used to make the model forecasting capabilities more and more reliable. In Figure 1 a pipeline of optimal control process is presented. As already specified, OCP($\boldsymbol{\mu}$)s, given a parametric instance, aim at reaching a configuration, represented by the collected data in this case. Mathematically, the data information is included in a functional to be minimized. The minimization process is performed through external variables, the controls, that change the classical model (the standard PDE($\boldsymbol{\mu}$)) to reach a more reliable solution to be exploited with forecasting purposes. The main contributions we will discuss in the thesis are about the *reduced space-time formulation*. The space-time approach is a very intuitive strategy already employed in several non-parametric optimal control examples, see e.g. [65, 67, 66, 89]. This discretization approach has been successfully exploited for parabolic PDE($\boldsymbol{\mu}$)s in [53, 156, 165, 166]. In the thesis, we build on these papers to propose space-time reduced strategies to be used for several applications based on OCP($\boldsymbol{\mu}$)s.
Nevertheless, in this framework, many questions arose, most of all for complicated tasks. For example, what is the best configuration to work with time-dependent problems? Is the problem well-posed? How can one make the reduced strategy versatile enough to be easily used in several contexts? Can we certify the error we are committing with respect to the standard discretization, at least for the simplest cases?

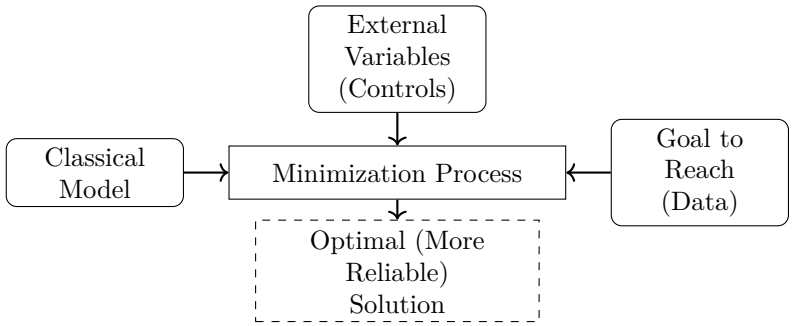


Figure 1: Graphic representation of the optimal control pipeline.

The space-time formulation reveals itself as a natural choice from this point of view. Indeed, besides its cons, one over all, the computational effort needed for the simulations, it utterly has an important strong point: a space-time problem can be treated as a steady problem and all the techniques, theorems and

well-known results deriving from literature for steady OCP($\boldsymbol{\mu}$)s can be slightly modified to apply in the time-dependent context.
For sure, if we go beyond standard model order reduction, other interesting questions can be addressed. What happens when the nonlinearity affects the uniqueness of the solution? How can I change the reduced model when dealing with stochastic controlled equations? Can artificial intelligence help in the prediction of the optimal solution?
All these topics will be discussed in this contribution and we will show how OCP($\boldsymbol{\mu}$)s can be useful even in more complicated applications. This work is original and based on findings and results obtained and presented in published papers, submitted preprints or ongoing projects carried out by myself jointly with others. Now, a detailed outline of the thesis follows.

◊ *Part I.* In this first part, we will deal with the theoretical formulation of OCP($\boldsymbol{\mu}$)s. Then, we will discuss the numerical approximation of these problems: i.e., the space-time high fidelity model is presented.

This part is divided into two Chapters. The proofs of Chapters 1 and 2 derive from an accepted work with G. Rozza and F. Ballarin [146] and from a submitted work with the same co-authors [148]. However, the presented formulation refers to several other contributions such as [16] and a published work with Z. Zainib [150].

- *Chapter 1.* We here present the continuous formulation of OCP($\boldsymbol{\mu}$)s, from steady linear problems to time-dependent and nonlinear ones. For linear time-dependent problems, we propose the proofs of the well-posedness of the systems. This Chapter represents the backbone of all the numerical results we are going to show in the thesis.
- *Chapter 2.* In this Chapter we will discuss the high fidelity approximation, proposing the space-time discretization both for steady and time-dependent problems. Also in this case, we will focus on the well-posedness of this specific finite dimensional setting.

◊ *Part II.* The second part of this contribution is dedicated to reduced strategies for OCP($\boldsymbol{\mu}$)s. Besides the techniques and the algorithms we relied on, we show some first numerical results to validate them. This part consists of three Chapters.

We want to stress that Chapter 4 presents the numerical results of the already cited works [16, 146], building on the submitted preprint [147]. Chapter 5 is based on the results obtained in [148].

- *Chapter 3.* In this Chapter we discuss the main ideas behind model order reduction and the basic assumptions to deal with an efficient and well-posed reduced system. The discussion applies both to steady and to time-dependent problems, covering a broad class of equations.

- *Chapter 4*. This Chapter is about space-time Proper Orthogonal Decomposition. We point out the versatility and the effectiveness of such an approach through some preliminary results for linear time-dependent OCP($\boldsymbol{\mu}$)s. They will be the stepping stones for a more complicated problem in environmental sciences: a nonlinear time-dependent application for coastal safeguard management.
- *Chapter 5*. A further step in the model reduction for OCP($\boldsymbol{\mu}$)s is here considered: the space-time Greedy. Unlike the data compression algorithms (such as POD), in this Chapter, we propose an iterative procedure to build the reduced spaces based on a posteriori error estimator between the high fidelity approximation and the reduced one. The error certification for linear problems is proved both for time-dependent and for steady cases and the strategy is validated through some numerical tests.

◊ *Part III*. The last part of the thesis is dedicated to the application of OCP($\boldsymbol{\mu}$)s to several examples of interest, from environmental sciences to heamodynamics. We will deal with stochastic equations and bifurcating nonlinear systems. Furthermore, for the first time in the thesis, artificial-intelligence-based approaches are studied and employed.

This part is divided into three Chapters. Chapter 6 is based on two works: an accepted paper with R. Mosetti, collaborator affiliated to the National Institute of Oceanography and Applied Geophysics (OGS) [145] and a submitted preprint in collaboration with G. Carere and R. Stevenson, from University of Amsterdam [33]. Chapter 7 entirely describes the submitted preprint in collaboration with F. Pichi [118]. Finally, Chapter 8 describes a part of an ongoing project carried out with N. Demo [46].

- *Chapter 6*. In this Chapter we will focus on the application of reduced techniques to problems of actual interest in environmental sciences: a pollutant release control in the Gulf of Trieste, Italy, and a forecasting oceanographic model for the North Atlantic Ocean dynamic. The models will be discussed both in their deterministic and stochastic versions. In the latter case, we will describe and exploit tailored reduction techniques for stochastic inputs and we will show some heuristic conclusions on how much convenient is to use these specific strategies with respect to standard model order reduction.
- *Chapter 7*. The topic of this Chapter concerns the role of optimal control on bifurcating phenomena arising in nonlinear PDE($\boldsymbol{\mu}$)s. We tested several control actions to understand how the features related to the loss of uniqueness might change under the influence of external variables.
- *Chapter 8*. The last Chapter is about the use of artificial intelligence and machine learning to deal with OCP($\boldsymbol{\mu}$)s. Indeed, we propose

physics informed neural network strategies to supply fast parametric predictions in an optimal control context. We here maintain the fashion of offline-online decomposition for parametric systems, but with a different flavour.

Finally, some conclusions and perspectives follow.

In this thesis we are not going to present the results obtained from a collaboration with M. Girfoglio and T. Iliescu from Virginia Tech [149]: the preprint is about model order reduction for convection dominated Navier-Stokes equations and, since unrelated to the main topic of this thesis, it has been omitted from this contribution.

We would like to remark that all the numerical results shown in this thesis have been coded with multiphenics [1] and RBniCS libraries [2], developed at SISSA mathLab. They are python-based libraries, built on FEniCS [98]. RBniCS is conceived to simulate parametric systems in a model order reduction framework, while multiphenics aims at easily defining multiphysics and coupled problems. We acknowledge the developers, the contributors and the maintainers of all the cited libraries.

Part I

High Fidelity Parametric Optimal Control Problems

Chapter 1

An Introduction to Parametric Optimal Control Problems

This Chapter deals with optimal control problems (OCP($\boldsymbol{\mu}$)s) governed by parametrized partial differential equations (PDE($\boldsymbol{\mu}$)s). Our goal is to provide an overview on their formulation, from nonlinear time-dependent systems to linear steady ones. First of all, in Section 1.1, we present the Lagrangian formalism [59, 67, 72] as a suitable approach to solve OCP($\boldsymbol{\mu}$)s making no distinction on the governing equation. Then, we will analyse specific features related to the linear governing equations and the well-posedness of such systems [111, 146] in Section 1.2. Finally, we propose an alternative formulation for the OCP($\boldsymbol{\mu}$)s following [148]: this will be the topic of Section 1.3.

1.1 Lagrangian approach for OCP($\boldsymbol{\mu}$)s

In the following Section, we provide the analytical formulation for nonlinear time-dependent OCP($\boldsymbol{\mu}$)s at the continuous level. After stating the problem setting, we will briefly describe the Lagrangian formalism [59, 67, 72] to tackle the minimization of quadratic cost functionals constrained to nonlinear time-dependent PDE($\boldsymbol{\mu}$)s. The proposed description focuses on Hilbert spaces, for the sake of clarity. However, all the concepts still hold in the Banach spaces framework.

1.1.1 Problem Formulation

Let us define as $\Omega \subset \mathbb{R}^d$, with $d = 2, 3$ a spatial domain where a physical event occurs in a time interval, say $[0, T]$. We suppose that the phenomenon we are interested in is described by a nonlinear time-dependent PDE($\boldsymbol{\mu}$). To represent and study such an evolution, we rely on the function spaces

$$\mathcal{Y}_0 = \{y \in L^2(0, T; Y) \text{ such that } y_t \in L^2(0, T; Y^*) \text{ with } y(0) = 0\} \tag{1.1}$$

and

$$\mathcal{Q} = L^2(0, T; Y), \tag{1.2}$$

for a Hilbert space Y. Let us assume to be provided by a nonlinear time-dependent *state equation* $G : \mathcal{Y}_0 \to \mathcal{Q}^*$ that reads

$$G(y; \boldsymbol{\mu}) = f. \tag{1.3}$$

Here, $y := y(\boldsymbol{\mu}) \in \mathcal{Y}_0$ is the *state variable*, i.e. the physical quantity we are focusing on, while $f := f(\boldsymbol{\mu}) \in \mathcal{Q}^*$ is an external forcing term. The parameter $\boldsymbol{\mu}$ is considered in the parameter space $\mathcal{D} \subset \mathbb{R}^D$, of dimension $D \geq 1$. The value of $\boldsymbol{\mu}$ may represent physical and/or geometrical features of the system solution. With the notation $\mathcal{L}(\cdot, \cdot)$, we indicate the space of the continuous linear functions between two function spaces. The state equation we consider is described by a PDE($\boldsymbol{\mu}$) of the following form:

$$G(y; \boldsymbol{\mu}) = y_t + E_{n\ell}(y; \boldsymbol{\mu}) + E_\ell(y; \boldsymbol{\mu}), \tag{1.4}$$

where $E_\ell \in \mathcal{L}(\mathcal{Y}_0, \mathcal{Q}^*)$ and $E_{n\ell}$ are the linear and the nonlinear contributions to the system, respectively. The time evolution of the solution is represented by the time derivative y_t. Now, let us suppose that the state variable y must reach a *desired profile* $y_\mathrm{d} := y_\mathrm{d}(\boldsymbol{\mu}) \in \mathcal{Y}_\mathrm{obs} \supseteq \mathcal{Y}_0$, where $\mathcal{Y}_\mathrm{obs} := L^2(0, T; Y_\mathrm{obs})$ with $Y_\mathrm{obs} \supset Y$. The observation in taken along all the time interval in a physical *observation domain* $\Omega_\mathrm{obs} \subset \Omega$. To steer the state solution towards the desired profile, we introduce a *control variable* $u := u(\boldsymbol{\mu}) \in \mathcal{U}$, where $\mathcal{U} = L^2(0, T; U)$ with U another Hilbert space defined over the *control domain* $\Omega_u \subset \Omega$. When $\Omega_u = \Omega$, we say that the OCP($\boldsymbol{\mu}$) is *distributed*. Namely, the variable u will act on the system and will change the behaviour of the state to achieve the data solution y_d. Theoretically, this action is represented by the *controlled equation* $\mathcal{E}(y, u; \boldsymbol{\mu})$, where $\mathcal{E} : \mathcal{Y}_0 \times \mathcal{U} \to \mathcal{Q}^*$ reads:

$$\mathcal{E}(y, u; \boldsymbol{\mu}) := \; G(y; \boldsymbol{\mu}) - C(u) - f = 0. \tag{1.5}$$

Here, $C \in \mathcal{L}(\mathcal{U}, \mathcal{Q}^*)$ is the *control operator* that influences the system in order to change its original state variable. The goal of an OCP($\boldsymbol{\mu}$) is reached though the following constrained minimization problem: given a $\boldsymbol{\mu} \in \mathcal{D}$ find the pair $(y, u) \in \mathcal{Y}_0 \times \mathcal{U}$ which solves

$$\min_{y \in \mathcal{Y}_{0\mathrm{ad}} \subset \mathcal{Y}_0, u \in \mathcal{U}_\mathrm{ad} \subset \mathcal{U}} J(y, u; y_\mathrm{d}) \text{ subject to } \mathcal{E}(y, u; \boldsymbol{\mu}) = 0, \tag{1.6}$$

where $J : \mathcal{Y}_0 \times \mathcal{U} \times \mathcal{Y}_\mathrm{obs} \to \mathbb{R}$ is the following *cost functional*:

$$J(y, u; y_\mathrm{d}) := \frac{1}{2}\|y - y_\mathrm{d}\|^2_{\mathcal{Y}_\mathrm{obs}} + \frac{\alpha}{2}\|u\|^2_{\mathcal{U}}, \tag{1.7}$$

with $\alpha \in (0, 1]$ *penalization parameter*. The value of α changes the role of the control variable: indeed, a larger value of α do not let the variable u to highly affect the original system, while the contrary happens when α is small. The minimization problem (1.6) has a solution when [67, Section 1.5.2]:

(i) $\mathcal{U}_\mathrm{ad} \subset \mathcal{U}$ is convex, bounded and closed;

(ii) $\mathcal{Y}_{0\mathrm{ad}} \subset \mathcal{Y}_0$ is convex and closed;

(iii) for every $\boldsymbol{\mu} \in \mathcal{D}$, the controlled system $\mathcal{E}(y, u; \boldsymbol{\mu}) = 0$ has a bounded solution map $u \in \mathcal{U} \mapsto y(u) \in \mathcal{Y}_0$;

(iv) for a given $\boldsymbol{\mu} \in \mathcal{D}$, the map $(y, u, \boldsymbol{\mu}) \in \mathcal{Y}_0 \times \mathcal{U} \to \mathcal{E}(y, u; \boldsymbol{\mu}) \in \mathcal{Q}^*$ is weakly continuous with respect to (w.r.t.) the first two arguments;

(v) for a given $y_\mathrm{d} \in \mathcal{Y}_\mathrm{obs}$, the objective functional $J(y, u; y_\mathrm{d})$ is weakly lower semicontinous w.r.t. the first two arguments.

Here, we considered $\mathcal{Y}_{0\mathrm{ad}}$ and $\mathcal{U}_\mathrm{ad}$ as *feasible sets* for the optimal solutions. However, from now on, we will indicate them with $\mathcal{Y}_0$ and $\mathcal{U}$, since these are the spaces considered in the numerical tests of this contribution. The OCP($\boldsymbol{\mu}$) can be solved through a Lagrangian argument. To this end, let us consider the arbitrary *adjoint variable* $z := z(\boldsymbol{\mu}) \in \mathcal{Y}_T \subset \mathcal{Q}$, where

$$\mathcal{Y}_T = \Big\{ z \in L^2(0,T;Y) \ \text{ s.t. } \ z_t \in L^2(0,T;Y^*) \text{ such that } z(T) = 0 \Big\}. \tag{1.8}$$

The employment of the adjoint variable will recast the problem (1.6) in an unconstrained minimization framework. For the sake of notation, we define the *global variable* $\boldsymbol{X} := \boldsymbol{X}(\boldsymbol{\mu}) = (y(\boldsymbol{\mu}), u(\boldsymbol{\mu}), z(\boldsymbol{\mu})) \in \mathbb{X}_{0,T} := \mathcal{Y}_0 \times \mathcal{U} \times \mathcal{Y}_T$ and thus, we consider the *Lagrangian functional* $\mathscr{L} : \mathbb{X}_{0,T} \times \mathcal{Y}_\mathrm{obs} \to \mathbb{R}$

$$\mathscr{L}(\boldsymbol{X}; y_\mathrm{d}, \boldsymbol{\mu}) := J(y, u; y_\mathrm{d}) + \langle z, \mathcal{E}(y, u; \boldsymbol{\mu}) \rangle_{\mathcal{Q}\mathcal{Q}^*}, \tag{1.9}$$

where $\langle \cdot, \cdot \rangle_{\mathcal{Q}\mathcal{Q}^*}$ is the duality pairing of the spaces $\mathcal{Q}$ and $\mathcal{Q}^*$. Furthermore we assume that:

(vi) $\mathcal{U}$ is nonempty;

(vii) $J : \mathcal{Y}_0 \times \mathcal{U} \times \mathcal{Y}_\mathrm{obs} \to \mathbb{R}$ and $\mathcal{E} : \mathcal{Y}_0 \times \mathcal{U} \to \mathcal{Q}^*$ are continuously Fréchet differentiable w.r.t. the first two arguments;

(viii) given $\boldsymbol{\mu} \in \mathcal{D}$, the controlled system $\mathcal{E}(y, u; \boldsymbol{\mu}) = 0$ has a unique solution $y = y(u) \in \mathcal{Y}_0$ for all $u \in \mathcal{U}$;

(ix) given $\boldsymbol{\mu} \in \mathcal{D}$, $D_y \mathcal{E}(y, u; \boldsymbol{\mu}) \in \mathcal{L}(\mathcal{Y}_0, \mathcal{Q}^*)$ has a bounded inverse for all $u \in \mathcal{U}$.

In assumption (ix), with D_y we indicate the Fréchet derivative w.r.t. the state variable and we will exploit this notation for u and z as well, with D_u and D_z, respectively. For a given $\boldsymbol{\mu} \in \mathcal{D}$, assumptions (vi) - (ix) assure the existence of an adjoint variable $z \in \mathcal{Y}_T$ related to the solution pair $(y, u) \in \mathcal{Y}_0 \times \mathcal{U}$ of (1.6), which satisfies the following *optimality system* [67]:

$$\begin{cases} D_y \mathscr{L}(\boldsymbol{X}; y_\mathrm{d}, \boldsymbol{\mu})[\omega] = 0 & \forall \omega \in \mathcal{Q}, \\ D_u \mathscr{L}(\boldsymbol{X}; y_\mathrm{d}, \boldsymbol{\mu})[\kappa] = 0 & \forall \kappa \in \mathcal{U}, \\ D_z \mathscr{L}(\boldsymbol{X}; y_\mathrm{d}, \boldsymbol{\mu})[\zeta] = 0 & \forall \zeta \in \mathcal{Q}, \end{cases} \tag{1.10}$$

which in strong form reads:

$$\begin{cases} y + D_y\mathcal{E}(y,u;\boldsymbol{\mu})^*(z) = y_{\mathrm{d}}, \\ \alpha u - C^*(z) = 0, \\ \mathcal{E}(y,u;\boldsymbol{\mu}) = 0. \end{cases} \tag{1.11}$$

In the system, $D_y\mathcal{E}(y,u;\boldsymbol{\mu})^* \in \mathcal{L}(\mathcal{Q},\mathcal{Y}_0^*)$ represents the adjoint operator of the Fréchet differencial of $\mathcal{E}(y,u;\boldsymbol{\mu})$ w.r.t. the state variable and it has the following form:

$$D_y\mathcal{E}(y,u;\boldsymbol{\mu})^*(z) = -z_t + D_yE_{n\ell}(y;\boldsymbol{\mu})^*z + E_\ell(y;\boldsymbol{\mu})^*z. \tag{1.12}$$

We remark that the dual variable z is considered in $\mathcal{Y}_T$ in order to guarantee a proper definition of the backward time evolution of the form $-z_t$ in the expression of the operator $D_y\mathcal{E}(y,u;\boldsymbol{\mu})^*$. For the sake of clarity, we underline that we have chosen to work with the Hilbert space Y also for the adjoint variable, rather then another Hilbert space, say Z. The reason of this choice is postponed in the thesis, see Section 1.2. In the same fashion, with $C^* \in \mathcal{L}(\mathcal{Y}_0,\mathcal{U}^*)$ we indicate the adjoint of the control operator. For the sake of notation, it can be useful to recast the optimality system (1.10) in the following compact form: given $\boldsymbol{\mu} \in \mathcal{D}$, find $\boldsymbol{X} \in \mathbb{X}_{0,T}$ such that

$$\mathcal{G}(\boldsymbol{X};\boldsymbol{\mu}) = \mathcal{F}, \tag{1.13}$$

with

$$\mathcal{G}(\boldsymbol{X};\boldsymbol{\mu}) := \begin{bmatrix} y + D_y\mathcal{E}(y,u;\boldsymbol{\mu})^*(z) \\ \alpha u - C^*(z) \\ G(y,\boldsymbol{\mu}) - C(u) \end{bmatrix} \quad \text{and} \quad \mathcal{F} := \begin{bmatrix} y_{\mathrm{d}} \\ 0 \\ f \end{bmatrix}.$$

A crucial step to solve problem (1.13) is to analyse it in variational form. To this end, let us define the form $\boldsymbol{G} : \mathbb{X}_{0,T} \times \mathbb{X}_{\mathrm{test}} \to \mathbb{R}$ and a function space $\mathbb{X}_{\mathrm{test}} \subseteq \mathcal{Q} \times \mathcal{U} \times \mathcal{Q}$ such as

$$\boldsymbol{G}(\boldsymbol{X},\boldsymbol{\Xi};\boldsymbol{\mu}) = \langle \mathcal{G}(\boldsymbol{X};\boldsymbol{\mu}),\boldsymbol{\Xi}\rangle_{\mathbb{X}_{0,T}{}^*,\mathbb{X}_{\mathrm{test}}} \quad \forall \boldsymbol{X} \in \mathbb{X}_{0,T}, \forall \boldsymbol{\Xi} \in \mathbb{X}_{\mathrm{test}}. \tag{1.14}$$

Thus, the optimality system (1.10) is equivalent to the variational formulation of (1.11) and reads: for a given $\boldsymbol{\mu} \in \mathcal{D}$, find the solution $\boldsymbol{X} \in \mathbb{X}_{0,T}$ of

$$\boldsymbol{G}(\boldsymbol{X},\boldsymbol{\Xi};\boldsymbol{\mu}) = \langle \mathcal{F},\boldsymbol{\Xi}\rangle_{\mathbb{X}_{0,T}{}^*,\mathbb{X}_{\mathrm{test}}} \quad \forall \boldsymbol{\Xi} \in \mathbb{X}_{\mathrm{test}}. \tag{1.15}$$

For now, we will always restrict to the case of *well-posedness* of the governing equation, i.e. we will assume its local invertibility, assured by hypotheses (viii) and (ix), that will guarantee the existence and the uniqueness of the state y for a given u. Moreover, we will work with the assumption of existence and uniqueness of the global optimal solution $\boldsymbol{X}$. In other words, we consider the optimality system $\mathcal{G}$ to be injective and surjective for every $\boldsymbol{\mu} \in \mathcal{D}$. While, this is provable for linear equations, it might not be the case for nonlinear controlled

equations. The analysis of this issue is postponed in Chapter 7, where we will describe the loss of uniqueness of the solution related to bifurcation theory, as presented in [118]. In the next Section, we will focus our attention on the specific case of linear problems, both steady and time-dependent, underlining the *saddle point* structure arising from the optimality system.

1.2 Saddle Point Structure for Linear Problems

This Section deals with linear governing equations and the saddle point structure related to the optimization problem (1.6) under linear constraints. We will start with the simplest case of steady equations (Section 1.2.1) and, then, we will analyse the time-dependent case (Sections 1.2.2 and 1.2.3). For both the classes of problems, we will state the well-posedness through the Brezzi Theorem [28], a classical choice in this setting. The literature is quite complete w.r.t. the saddle point framework arising from the Lagrangian formalism applied to steady linear constraints: an overview on the topic can be found in [26, 67]. We will present the main of these works and, moreover, we will, therefore, extend them to time-dependent ones [146], building on the setting proposed in [65, 143, 144].

1.2.1 Linear Steady OCP($\boldsymbol{\mu}$)s

Let us consider steady linear quadratic problems. All the concepts defined in Section 1.1 are still valid, however they have to be simplified in the following way:

- the state variable is $y \in Y$, the control variable $u \in U$ and the desired state $y_{\mathrm{d}} \in Y_{\mathrm{obs}}$. Namely, since no time evolution is present, we focus only on the space domain Ω. The definitions of control and observation domains still hold. The forcing term f is now considered in Y^*.
- The cost functional $J : Y \times U \times Y_{\mathrm{obs}} \to \mathbb{R}$ has the following form

$$J(y, u; y_{\mathrm{d}}) := \frac{1}{2}\|y - y_{\mathrm{d}}\|^2_{Y_{\mathrm{obs}}} + \frac{\alpha}{2}\|u\|^2_U, \tag{1.16}$$

  constrained to

$$\mathcal{E}(y, u; \boldsymbol{\mu}) := E_\ell(y; \boldsymbol{\mu}) - C(u) - f = 0, \tag{1.17}$$

  where $\mathcal{E} : Y \times U \to Y^*$ is the steady linear controlled equation.

Here, we used the same notations of the general problem of Section 1.1.1, however, there is no risk of misinterpretation, since we are only simplifying the context, retaining the general problem structure. Furthermore, also in this case,

we omit the explicit $\boldsymbol{\mu}$-dependency of the involved quantities. The steady minimization problem reads: given a parameter $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, u) \in Y \times U$ which minimizes (1.16) and solves (1.17). Once again, we rely on the Lagrangian formalism to build an optimality system based on the differentiation w.r.t. the problem variables. The main goal of this Section is to highlight the saddle point structure arising in this specific linear optimization setting. To this end, we need to define the following quantities:

- the *state bilinear form* $a : Y \times Y \to \mathbb{R}$, i.e. the weak form of the state operator $E_\ell(y; \boldsymbol{\mu})$;
- the *control bilinear form* $c : U \times Y \to \mathbb{R}$, i.e. the weak form of the control operator $C(u)$.

In this way, the weak formulation of the whole controlled state equation reads:

$$a\left(y, \zeta; \boldsymbol{\mu}\right) - c\left(u, \zeta; \boldsymbol{\mu}\right) - \left\langle f\left(\boldsymbol{\mu}\right), \zeta \right\rangle = 0, \quad \forall \zeta \in Y. \tag{1.18}$$

We now recast the functional (1.16) as

$$J(y, u; y_{\mathrm{d}}) = \frac{1}{2} m(y - y_{\mathrm{d}}, y - y_{\mathrm{d}}; \boldsymbol{\mu}) + \frac{\alpha}{2} n(u, u; \boldsymbol{\mu}). \tag{1.19}$$

where $m : Y \times Y \to \mathbb{R}$ and $n : U \times U \to \mathbb{R}$ are two bilinear forms describing the L^2-products over the observation and the control domains, respectively. The parameter dependence might arise if a geometrical parametrization is present: we postpone this topic in the next Section. From now on, in the linear cases, we will limit our analysis to the case of the following assumptions being verified.

Assumptions 1 *the bilinear forms $c(\cdot, \cdot; \boldsymbol{\mu})$, $a(\cdot, \cdot; \boldsymbol{\mu})$, $n(\cdot, \cdot; \boldsymbol{\mu})$ and $m(\cdot, \cdot; \boldsymbol{\mu})$ verify the properties*

(a) $a(\cdot, \cdot; \boldsymbol{\mu})$ is continuous and coercive of constants $c_a(\boldsymbol{\mu})$ and $\gamma_a(\boldsymbol{\mu})$, respectively;

(b) $c(\cdot, \cdot; \boldsymbol{\mu})$ is continuous of constant $c_c(\boldsymbol{\mu})$;

(c) $n(\cdot, \cdot; \boldsymbol{\mu})$ is symmetric, continuous and coercive of constants $c_n(\boldsymbol{\mu})$ and $\gamma_n(\boldsymbol{\mu})$, respectively;

(d) $m(\cdot, \cdot; \boldsymbol{\mu})$ is symmetric, continuous and positive definite.

We remark that hypotheses (a) and (b) ensure the existence of a unique $y \in Y$, solution to (1.18), for a given $u \in \mathcal{U}$ and $\boldsymbol{\mu} \in \mathcal{D}$. As did in Section 1.1.1, we define the adjoint variable $z \in Y$ and the steady Lagrangian functional as

$$\mathscr{L}(y, u, z; \boldsymbol{\mu}) := J(y, u; y_{\mathrm{d}}) + \left\langle z, \mathcal{E}\left(y, u; \boldsymbol{\mu}\right) \right\rangle_{YY^*} \tag{1.20}$$

The optimization problem relies on the solution of the optimality system given by the differentiation of (1.20) w.r.t. the state, the control and the adjoint variables, respectively, building a steady linear analogous of system (1.11), where $\boldsymbol{X} := (y, u, z) \in \mathbb{X} := Y \times U \times Y$ and

$$D_y \mathcal{E}(y, u; \boldsymbol{\mu})^*(z) = E_\ell(y; \boldsymbol{\mu})^* z, \tag{1.21}$$

due to the linearity of the governing equation.
We can now bring to the light the saddle point structure of the optimality system. Indeed, we define the *state-control variable* $x = (y, u) \in X := Y \times U$ and the bilinear forms $\mathcal{A} : X \times X \to \mathbb{R}$ and $\mathcal{B} : X \times Y \to \mathbb{R}$ and the functional $\mathcal{H} \in Y^*$ as

$$\begin{aligned} &\mathcal{A}(x, \xi; \boldsymbol{\mu}) = m(y, \omega; \boldsymbol{\mu}) + \alpha n(u, \kappa; \boldsymbol{\mu}) && \forall \xi := (\omega, \kappa) \in X, \\ &\mathcal{B}(x, \zeta; \boldsymbol{\mu}) = a(y, \zeta; \boldsymbol{\mu}) - c(u, \zeta) && \forall \zeta \in Y, \\ &\langle H(\boldsymbol{\mu}), \xi \rangle = m(y_{\mathrm{d}}, \omega; \boldsymbol{\mu}) && \forall \xi := (\omega, \kappa) \in X. \end{aligned} \tag{1.22}$$

From simple computations, exploiting the definitions (1.22), the optimality system deriving from the differentiation of the Lagrangian Functional (1.20) can be recast in mixed formulation as follows: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(x, z) \in X \times Y$ which verifies

$$\begin{cases} \mathcal{A}(x, \xi; \boldsymbol{\mu}) + \mathcal{B}(\xi, z; \boldsymbol{\mu}) = \langle H(\boldsymbol{\mu}), \xi \rangle & \forall \xi \in X, \\ \mathcal{B}(x, \zeta; \boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), \zeta \rangle & \forall \zeta \in Y, \end{cases} \tag{1.23}$$

where $F \in Y^*$ encodes the forcing term and the boundary conditions of the controlled state equation. The well-posedness of the system (1.23) relies on the following Theorem.

Theorem 1 (Brezzi Theorem) *The parametric problem (1.23) admits a unique solution $(x, z) \in X \times Y$ if:*

- *$\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ is continuous and weakly coercive on the set X_0 which denotes the kernel of $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$, i.e, for a positive constant $\beta_{\mathcal{A}}(\boldsymbol{\mu})$, it holds*

$$\inf_{x \in X_0 \setminus \{0\}} \sup_{\xi \in X_0 \setminus \{0\}} \frac{\mathcal{A}(x, \xi; \boldsymbol{\mu})}{\|x\|_X \|\xi\|_X} \geq \beta_{\mathcal{A}}(\boldsymbol{\mu}) > 0,$$

and

$$\inf_{\xi \in X_0 \setminus \{0\}} \sup_{x \in X_0 \setminus \{0\}} \frac{\mathcal{A}(x, \xi; \boldsymbol{\mu})}{\|x\|_X \|\xi\|_X} \geq 0.$$

- *$\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ is continuous and inf-sup stable of positive constant $\beta_{\mathcal{B}}(\boldsymbol{\mu})$, i.e.*

$$\inf_{\zeta \in Y \setminus \{0\}} \sup_{x \in X \setminus \{0\}} \frac{\mathcal{B}(x, \zeta; \boldsymbol{\mu})}{\|x\|_X \|\zeta\|_Y} \geq \beta_{\mathcal{B}}(\boldsymbol{\mu}) > 0. \tag{1.24}$$

The choice of $z \in Y$ and Assumptions 1 make Theorem 1 provable: see [111] and the references therein. The goal of the next Section is to extend this structure to linear time-dependent problems, proposing an analysis of their well-posedness.

Remark 1.2.1 (Steady Nonlinear OCP($\boldsymbol{\mu}$)s) *We stress that, once defined $y, z \in Y$, $u \in U$, $y_d \in Y_{obs}$, one can deal with nonlinear steady problems and solve the minimization of (1.16) constrained to*

$$\mathcal{E}(y, u; \boldsymbol{\mu}) := E_{n\ell}(y; \boldsymbol{\mu}) + E_{\ell}(y; \boldsymbol{\mu}) - C(u) - f = 0, \tag{1.25}$$

where $\mathcal{E} : Y \times U \to Y^*$ is the steady nonlinear controlled equation. To tackle this problem, one may exploit the Lagrangian techniques building a steady version of the optimality system (1.10), where the adjoint equation will present the following term

$$D_y \mathcal{E}(y, u; \boldsymbol{\mu})^*(z) = D_y E_{n\ell}(y; \boldsymbol{\mu})^* z + E_{\ell}(y; \boldsymbol{\mu})^* z, \tag{1.26}$$

i.e. (1.12) without backward time evolution. For nonlinear problems there is not a general theory which is able to recast them into saddle point problems: it is strictly related to the state equation at hand, see for example [100, 168] for the Navier-Stokes equations. However, at the algebraic level one can recover this peculiar structure for nonlinear equations. In Section 2.2, we will show nonlinear an example for governing equations which depend quadratically w.r.t. the state y.

1.2.2 Parabolic Time-Dependent OCP($\boldsymbol{\mu}$)s

In this Section, parabolic time-dependent OCP($\boldsymbol{\mu}$)s are recast in a saddle point framework. Building on standard approaches for stationary linear state equations [79, 110, 111, 135], we provide the well-posedness of the saddle point structure in a space-time formulation. In this context, we consider the spatial state space Y verifying $Y \hookrightarrow H \hookrightarrow Y^*$, with H another Hilbert space. In the following we assume that Y, H and U are contained in $L^2(\Omega)$, a typical framework for the parabolic OCP($\boldsymbol{\mu}$)s we are dealing with. In Section 1.1.1 we have already defined the Hilbert spaces suited for the problem at hand. Now, we endow $\mathcal{Y}_0$ and $\mathcal{U}$ with the following norms, respectively:

$$\|y\|^2_{\mathcal{Y}_0} = \int_0^T \|y\|^2_Y \, dt + \int_0^T \|y_t\|^2_{Y^*} \, dt \qquad \text{and} \qquad \|u\|^2_{\mathcal{U}} = \int_0^T \|u\|^2_U \, dt.$$

The adjoint space $\mathcal{Y}_T$ is characterized by the same norm of the state space. Namely, we will use $\|\cdot\|_{\mathcal{Y}_0}$ and $\|\cdot\|_{\mathcal{Y}_T}$, interchangeably. Time-dependent OCP($\boldsymbol{\mu}$)s read as follows: for a given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, u) := (y(\boldsymbol{\mu}), u(\boldsymbol{\mu})) \in$

$\mathcal{Y}_0 \times \mathcal{U}$ which solves

$$\min_{(y,u)\in\in\mathcal{Y}_0\times\mathcal{U}} J((y,u);\boldsymbol{\mu}) = \frac{1}{2}\int_0^T m(y-y_{\mathrm{d}}, y-y_{\mathrm{d}};\boldsymbol{\mu})\,dt + \frac{\alpha}{2}\int_0^T n(u,u;\boldsymbol{\mu})\,dt, \tag{1.27}$$

governed by

$$\begin{cases} S(\boldsymbol{\mu})y_t + D_a(\boldsymbol{\mu})y = C(\boldsymbol{\mu})u + f(\boldsymbol{\mu}) & \text{in } \Omega\times(0,T), \\ y = g(\boldsymbol{\mu}) & \text{on } \Gamma_D\times(0,T), \\ \dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N\times(0,T), \\ y(0) = 0 & \text{in } \Omega. \end{cases} \tag{1.28}$$

Here,

- $D_a(\boldsymbol{\mu}) : Y \to Y^*$ is a general differential state operator,
- $S(\boldsymbol{\mu}) : Y^* \to Y^*$ is a function representing the time evolution,
- $C(\boldsymbol{\mu}) : U \to Y^*$ is an operator describing the control action,
- $f(\boldsymbol{\mu})$ denotes external sources,
- Γ_D is the portion of the boundary $\partial\Omega$ where Dirichlet boundary conditions are applied, and $g(\boldsymbol{\mu})$ represents Dirichlet data,
- Γ_N is the portion of the boundary $\partial\Omega$ where Neumann boundary conditions are applied,
- $m(\cdot,\cdot;\boldsymbol{\mu}) : Y \times Y \to \mathbb{R}$, related to the operator $M(\boldsymbol{\mu}) : Y \to Y^*$, and $n(\cdot,\cdot;\boldsymbol{\mu}) : U \times U \to \mathbb{R}$, related to the operator $N(\boldsymbol{\mu}) : U \to U^*$, are the two bilinear forms defined in Section 1.2.1.

In this setting, $\boldsymbol{\mu}$ might also represent some geometrical features: we will assume to have already traced back the problem to the *reference domain* that we will call Ω with abuse of notation, and that $S(\boldsymbol{\mu})$, $D_a(\boldsymbol{\mu})$, $C(\boldsymbol{\mu})$, $M(\boldsymbol{\mu})$, $N(\boldsymbol{\mu})$, and $f(\boldsymbol{\mu})$ encode information about the pulled back operators, see e.g. [134]. Without loss of generality, we assume $C(\boldsymbol{\mu})$ and $S(\boldsymbol{\mu})$ to be both the identity map, or its trace back. Hence, $S(\boldsymbol{\mu})$ and $C(\boldsymbol{\mu})$ are self-adjoint and, when a geometrical parameter is present, they have the following forms

$$\sum_i^{Q_S} c_S^i(\boldsymbol{\mu})\chi_{\Omega_S^i} \quad \text{and} \quad \sum_i^{Q_C} c_C^i(\boldsymbol{\mu})\chi_{\Omega_C^i}, \tag{1.29}$$

respectively, for $Q_S, Q_C \in \mathbb{N}$ and $c_S^i(\boldsymbol{\mu}), c_C^i(\boldsymbol{\mu})$ positive constants resulting from the trace back of indicator functions $\chi_{\Omega_S^i}, \chi_{\Omega_C^i}$ which verify

$$\bigcup_i^{Q_S} \Omega_S^i = \Omega \quad \text{and} \quad \bigcup_i^{Q_C} \Omega_C^i = \Omega_u.$$

This scenario complies with literature related to OCP($\boldsymbol{\mu}$)s and does not limit the applicability of the model. The state equation at hand can be recast in weak formulation as follows: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, u) \in \mathcal{Y}_0 \times \mathcal{U}$ which verifies

$$\begin{cases} \displaystyle\int_0^T s(y, \zeta; \boldsymbol{\mu})\, dt \\ \qquad + \displaystyle\int_0^T a(y, \zeta; \boldsymbol{\mu})\, dt = \int_0^T c(u, \zeta; \boldsymbol{\mu})\, dt \\ \qquad\qquad + \displaystyle\int_0^T \langle F(\boldsymbol{\mu}), \zeta \rangle_{Y^*, Y}\, dt \quad \forall \zeta \in \mathcal{Q}, \\ y(0) = 0 \qquad\qquad \text{in } \Omega, \end{cases} \tag{1.30}$$

where $a : Y \times Y \to \mathbb{R}$ and $c : U \times Y \to \mathbb{R}$ are the bilinear forms related to $D_a(\boldsymbol{\mu})$ and $C(\boldsymbol{\mu})$, respectively, already defined in Section 1.2.1. In the same Section, we also defined $F(\boldsymbol{\mu}) \in Y^*$, as a functional giving information about forcing terms and spatial boundary conditions deriving from the weak state equation. Moreover,

$$s(y, \zeta; \boldsymbol{\mu}) = \langle S(\boldsymbol{\mu}) y_t, \zeta \rangle_{Y^* Y}\,. \tag{1.31}$$

Now, the OCP($\boldsymbol{\mu}$) has the following form: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, u) \in \mathcal{Y}_0 \times \mathcal{U}$ which satisfies

$$\min_{(y,u) \in \mathcal{Y}_0 \times \mathcal{U}} J((y, u); \boldsymbol{\mu}) \qquad \text{such that (1.28) holds.} \tag{1.32}$$

Also this minimization problem can be solved through a Lagrangian approach, defining an *adjoint variable* $z := z(\boldsymbol{\mu}) \in \mathcal{Y}_T$. We stress that the adjoint variable has the same regularity in space w.r.t. the state variable. This assumption is crucial to prove the well-posedness results presented in [146]. The obtained optimality system is: for a given $\boldsymbol{\mu} \in \mathcal{D}$, find $(y, u, z) \in \mathcal{Y}_0 \times \mathcal{U} \times \mathcal{Y}_T$ such that

$$\begin{cases} M(\boldsymbol{\mu}) y \chi_{\Omega_{\text{obs}}} - S(\boldsymbol{\mu}) z_t + D_a(\boldsymbol{\mu})^* z = M(\boldsymbol{\mu}) y_{\text{d}} & \text{in } \Omega \times (0, T), \\ \alpha N(\boldsymbol{\mu}) u - C(\boldsymbol{\mu}) z \chi_{\Omega_u} = 0 & \text{in } \Omega \times (0, T), \\ S(\boldsymbol{\mu}) y_t + D_a(\boldsymbol{\mu}) y - C(\boldsymbol{\mu}) u = f(\boldsymbol{\mu}) & \text{in } \Omega \times (0, T), \\ y(0) = y_0 & \text{in } \Omega, \\ z(T) = 0 & \text{in } \Omega, \\ \text{boundary conditions} & \text{on } \partial\Omega \times (0, T), \end{cases} \tag{1.33}$$

here $D_a(\boldsymbol{\mu})^*$ is the dual operator of $D_a(\boldsymbol{\mu})$, while $\chi_{\Omega_{\text{obs}}}$ and χ_{Ω_u} are the indicator functions of the observation domain Ω_{obs} and the control domain Ω_u, respectively. We now recover the saddle point structure already presented in Section 1.2.1. First of all, as did in the steady case, we define the *state-control variable* $x = (y, u) \in \mathcal{X} := \mathcal{Y}_0 \times \mathcal{U}$ and the bilinear forms $\mathcal{A} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and $\mathcal{B} : \mathcal{X} \times \mathcal{Y}_0 \to \mathbb{R}$ and the functional $\mathcal{H} \in Y^*$ as:

$$\begin{aligned}
\mathcal{A}(x, \xi; \boldsymbol{\mu}) &= \int_0^T m(y, \omega; \boldsymbol{\mu})\, dt \\
&\qquad + \alpha \int_0^T n(u, \kappa; \boldsymbol{\mu})\, dt \qquad \forall \xi := (\omega, \kappa) \in \mathcal{X}, \\
\mathcal{B}(x, \zeta; \boldsymbol{\mu}) &= \int_0^T s(y, \zeta)\, dt + \int_0^T a(y, \zeta; \boldsymbol{\mu})\, dt \\
&\qquad - \int_0^T c(u, \zeta; \boldsymbol{\mu})\, dt \qquad \forall \zeta \in \mathcal{Y}_0, \\
\mathcal{H}(\xi; \boldsymbol{\mu}) &= \int_0^T \langle H(\boldsymbol{\mu}), \omega \rangle\ dt = \int_0^T m(y_d, \omega; \boldsymbol{\mu})\, dt \qquad \forall \omega \in \mathcal{Y}_0.
\end{aligned} \tag{1.34}$$

After having applied a Lagrangian argument, the optimality system (1.10) with the new definitions (1.34) can be written in the following form:

$$\begin{cases}
\mathcal{A}(x, \xi; \boldsymbol{\mu}) + \mathcal{B}(\xi, z; \boldsymbol{\mu}) = \mathcal{H}(\xi; \boldsymbol{\mu}) & \forall \xi \in \mathcal{X}, \\
\mathcal{B}(x, \zeta; \boldsymbol{\mu}) = \displaystyle\int_0^T \langle F(\boldsymbol{\mu}), \zeta \rangle\, dt & \forall \zeta \in \mathcal{Y}_0.
\end{cases} \tag{1.35}$$

Namely, the typical saddle point structure of steady linear quadratic OCP($\boldsymbol{\mu}$)s is preserved also in the linear time-dependent case. As in the steady case, we can prove the well-posedness of the problem at hand thanks to the Brezzi theorem [27, 28]:

Theorem 2 (Brezzi Theorem, time-dependent case) *The parametric problem (1.35) admits a unique solution $(x, z) \in \mathcal{X} \times \mathcal{Y}_T$ if:*

1. *$\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ is continuous and weakly coercive on the set $\mathcal{X}_0$ which denotes the kernel of $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$, i.e. for a positive constant $\beta_{\mathcal{A}}(\boldsymbol{\mu})$, it holds*

$$\inf_{\xi \in \mathcal{X}_0 \setminus \{0\}} \sup_{x \in \mathcal{X}_0 \setminus \{0\}} \frac{\mathcal{A}(x, \xi; \boldsymbol{\mu})}{\|x\|_{\mathcal{X}_0} \|\xi\|_{\mathcal{X}}} > \beta_{\mathcal{A}}(\boldsymbol{\mu}) > 0 \tag{1.36}$$

and

$$\inf_{x \in \mathcal{X} \setminus \{0\}} \sup_{\xi \in \mathcal{X}_0 \setminus \{0\}} \frac{\mathcal{A}(x, \xi; \boldsymbol{\mu})}{\|x\|_{\mathcal{X}_0} \|w\|_{\mathcal{X}}} > 0. \tag{1.37}$$

2. $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ *is continuous and inf-sup stable of positive constant* $\beta_{\mathcal{B}}(\boldsymbol{\mu})$ *, i.e.*

$$\inf_{\zeta\in\mathcal{Y}_T\setminus\{0\}}\sup_{x\in\mathcal{X}\setminus\{0\}}\frac{\mathcal{B}(x,\zeta;\boldsymbol{\mu})}{\|x\|_{\mathcal{X}}\|\zeta\|_{\mathcal{Y}_T}}=\beta_{\mathcal{B}}(\boldsymbol{\mu})>0. \tag{1.38}$$

Furthermore, if one assumes the bilinear form $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$ to be symmetric, non-negative, and coercive on $\mathcal{X}_0$, then the minimization of the functional (1.27) constrained to equation (1.28) and the saddle point problem (1.35) are equivalent. Our aim is to prove the Theorem 2 as did in [146]. Here, however, we propose some modifications related to the setting of the geometrical parametrization. To reach this goal, besides the Assumptions 1, we will exploit the following relations:

(I) by definition, for every $y\in\mathcal{Y}_0$ and $u\in\mathcal{U}$ it holds:

$$\|y_t\|_{\mathcal{Y}^*}\leq\|y\|_{\mathcal{Y}_0},\|y\|_{\mathcal{Y}}\leq\|y\|_{\mathcal{Y}_0},\|y\|_{\mathcal{Y}_0}\leq\|x\|_{\mathcal{X}}\text{ and }\|u\|_{\mathcal{U}}\leq\|x\|_{\mathcal{X}}; \tag{1.39}$$

(II) for y solution of a parabolic PDE($\boldsymbol{\mu}$) with forcing term f and $y(0)=y_0$, there exists $k(\boldsymbol{\mu})>0$ such that:

$$\|y\|_{\mathcal{Y}_0}\leq k(\boldsymbol{\mu})(\|f\|_{\mathcal{Q}}+\|y_0\|_Y). \tag{1.40}$$

We stress that $\mathcal{Y}_0$ and $\mathcal{Y}_T$ share the same norm: thus the inequalities in (I) also hold true for $y\in\mathcal{Y}_T$. Moreover, in order to guarantee the inf-sup stability (1.38) we need the two following Lemmas.

Lemma 1.1 *Let* ζ *be a function in* $\mathcal{Y}_0$ *(or in* $\mathcal{Y}_T$*), then the following inequality holds:*

$$\frac{\|\zeta\|_{\mathcal{Y}}^2}{\|\zeta\|_{\mathcal{Y}_0}^2}\geq\frac{1}{6}.$$

Proof. To tackle the proof, we study to separate cases.
Case 1. Let us assume $\|\zeta_t\|_{\mathcal{Y}^*}\leq\|\zeta\|_{\mathcal{Y}}$ then $\|\zeta\|_{\mathcal{Y}_0}^2\leq 2\|\zeta\|_{\mathcal{Y}}^2$. This leads to

$$\frac{\|\zeta\|_{\mathcal{Y}}^2}{\|\zeta\|_{\mathcal{Y}_0}^2}\geq\frac{\|\zeta\|_{\mathcal{Y}}^2}{2\|\zeta\|_{\mathcal{Y}}^2}=\frac{1}{2}>\frac{1}{6}.$$

Case 2. We now focus on $\zeta\in\mathcal{Y}_0$ such that $\|\zeta_t\|_{\mathcal{Y}^*}>\|\zeta\|_{\mathcal{Y}}$. This assumption results in the following relation:

$$2\|\zeta\|_{\mathcal{Y}}^2<\|\zeta\|_{\mathcal{Y}_0}^2\qquad\Rightarrow\qquad\frac{1}{2\|\zeta\|_{\mathcal{Y}}^2}>\frac{1}{\|\zeta\|_{\mathcal{Y}_0}^2}, \tag{1.41}$$

and, by definition of $\|\zeta\|_{\mathcal{Y}_0}$,

$$\|\zeta\|^2_{\mathcal{Y}_0} + 2\,\|\zeta_t\|^2_{\mathcal{Y}^*} \geq 3\,\|\zeta_t\|^2_{\mathcal{Y}^*} \Rightarrow -\frac{1}{\|\zeta\|^2_{\mathcal{Y}_0} + 2\,\|\zeta_t\|^2_{\mathcal{Y}^*}} \geq -\frac{1}{3\,\|\zeta_t\|^2_{\mathcal{Y}^*}}. \tag{1.42}$$

Then, we can prove that

$$\begin{aligned}
\frac{\|\zeta\|^2_{\mathcal{Y}}}{\|\zeta\|^2_{\mathcal{Y}_0}} &= \frac{2\|\zeta\|^2_{\mathcal{Y}}}{2\|\zeta\|^2_{\mathcal{Y}} + 2\,\|\zeta_t\|^2_{\mathcal{Y}^*}} \underbrace{\geq}_{\text{for (1.41)}} \frac{2\|\zeta\|^2_{\mathcal{Y}}}{\|\zeta\|^2_{\mathcal{Y}_0} + 2\,\|\zeta_t\|^2_{\mathcal{Y}^*}} \\
&= \frac{2\|\zeta\|^2_{\mathcal{Y}} + \|\zeta_t\|^2_{\mathcal{Y}^*} - \|\zeta_t\|^2_{\mathcal{Y}^*}}{\|\zeta\|^2_{\mathcal{Y}_0} + 2\,\|\zeta_t\|^2_{\mathcal{Y}^*}} \\
&\geq \frac{\min\{1,2\}\|\zeta\|^2_{\mathcal{Y}_0}}{\max\{1,2\}\|\zeta\|^2_{\mathcal{Y}_0}} - \frac{\|\zeta_t\|^2_{\mathcal{Y}^*}}{\|\zeta\|^2_{\mathcal{Y}_0} + 2\,\|\zeta_t\|^2_{\mathcal{Y}^*}} \\
&\underbrace{\geq}_{\text{for (1.42)}} \frac{1}{2} - \frac{\|\zeta_t\|^2_{\mathcal{Y}^*}}{3\,\|\zeta_t\|^2_{\mathcal{Y}^*}} = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}.
\end{aligned}$$

■

As previously specified, in order to prove the well-posedness of the problem at hand, we still need the next Lemma.

Lemma 1.2 *Given a function $v \in \mathcal{Y}_0$, there exists $\bar{y} \in \mathcal{Y}_0$ which verifies:*

$$\int_0^T \left\langle \bar{y}_t, \zeta \right\rangle dt + \int_0^T a(\bar{y}, \zeta; \boldsymbol{\mu})\, dt = \int_0^T a(v, \zeta; \boldsymbol{\mu})\, dt \qquad \forall \zeta \in \mathcal{Y}_0, \tag{1.43}$$

with $\bar{y}(0) = 0$. Moreover, there exists a positive constants $\bar{k}(\boldsymbol{\mu})$ such that the following inequality holds:

$$\|\bar{y}\|_{\mathcal{Y}_0} \leq \bar{k}(\boldsymbol{\mu})\|v\|_{\mathcal{Y}}. \tag{1.44}$$

Proof. For the existence of the solution $\bar{y}$ to the specific auxiliary problem (1.43) we refer to the proof of property (A.3) of the Theorem 5.1 in [139], where the existence of $\bar{y} \in \mathcal{Y}_0$ is guaranteed for a given $v \in \mathcal{Y} \supset \mathcal{Y}_0$ and for every initial condition. Thus, for $t \in (0, T)$, we consider the linear state operator $D_a(\boldsymbol{\mu}) : Y \to Y^*$ defined by $\langle D_a v, \zeta \rangle_{H^*,H}$. We call $C_{D_a}(\boldsymbol{\mu}) := \|D_a\|_{H^*}$ which is finite due to the continuity of the operator. Furthermore, since $\bar{y}$ verifies (1.43), from (1.40) we may assert

$$\|\bar{y}\|_{\mathcal{Y}_0} \leq k(\boldsymbol{\mu})\|D_a v\|_{L^2(0,T;H)} \leq k(\boldsymbol{\mu}) C_{D_a} \|v\|_{L^2(0,T;H)}.$$

Since $Y \hookrightarrow H$, it holds $\|\zeta\|_H \leq \bar{C}\|\zeta\|_Y$, then, calling $\bar{k}(\boldsymbol{\mu}) = k(\boldsymbol{\mu})C_{D_a}\bar{C}$, we can prove the statement.

■

Thanks to these two Lemmas, we are now able to prove the time-dependent Brezzi theorem hypotheses in the following well-posedness theorem.

Theorem 3 *The saddle point problem* (1.35) *satisfies the hypotheses of Theorem 2 under the Assumptions 1, and, thus, it is well-posed.*

Proof. Let us consider the continuity of $\mathcal{A}(\cdot,\cdot;\boldsymbol{\mu})$.

$$\begin{aligned}|\mathcal{A}(x,w,\boldsymbol{\mu})| &\leq c_n(\boldsymbol{\mu})\|y\|_{\mathcal{Y}_0}\|z\|_{\mathcal{Y}_0} + \alpha\|u\|_{\mathcal{U}}\|v\|_{\mathcal{U}} \\ &\leq \max\{c_n(\boldsymbol{\mu}),\alpha\}\|x\|_{\mathcal{X}}\|w\|_{\mathcal{X}}.\end{aligned}$$

Here, we exploited the continuity of

$$\int_0^T m(\cdot,\cdot;\boldsymbol{\mu})\,dt \quad \text{and} \quad \int_0^T n(\cdot,\cdot;\boldsymbol{\mu})\,dt,$$

which can be shown through simple computations, see e.g. [146] for details. Furthermore, the hypothesis (c) and (d) assure the symmetry, the positive definiteness of the form $\mathcal{A}(\cdot,\cdot,\boldsymbol{\mu})$. It remains the coercivity of $\mathcal{A}(\cdot,\cdot,\boldsymbol{\mu})$ over the kernel of the controlled state equation. If $x \in \mathcal{X}_0$, then it holds

$$\begin{aligned}\int_0^T \sum_i^{Q_S} c_S^i(\boldsymbol{\mu})\chi_{\Omega_S^i}\langle y_t,\zeta\rangle\,dt + \int_0^T a(y,\zeta;\boldsymbol{\mu})\,dt &= \int_0^T s(y,\zeta;\boldsymbol{\mu})\,dt \\ &\quad + \int_0^T a(y,\zeta;\boldsymbol{\mu})\,dt \\ &= \int_0^T c(u,\zeta;\boldsymbol{\mu})\,dt,\end{aligned}$$

and thus, without loss of generality, we can exploit (1.40) where the forcing term in weak form is represented by

$$\int_0^T c(u,\zeta;\boldsymbol{\mu})\,dt,$$

obtaining $\|y\|_{\mathcal{Y}_0} \leq k(\boldsymbol{\mu})(\|u\|_{\mathcal{U}} + \|y_0\|_H) = k(\boldsymbol{\mu})\|u\|_{\mathcal{U}}$ since $y_0 = y(0) = 0$, where, with the constant $k(\boldsymbol{\mu})$, we are hiding the contribution of the geometrical

parametrization. Then, it holds:

$$\begin{aligned}\mathcal{A}(x,x;\boldsymbol{\mu}) &= m(y,y,\boldsymbol{\mu}) + \alpha n(u,u,\boldsymbol{\mu}) \geq \frac{\alpha}{2}\|u\|^2_{\mathcal{U}} + \frac{\alpha}{2}\|u\|^2_{\mathcal{U}} \\ &\geq \frac{\alpha}{2k(\boldsymbol{\mu})^2}\|y\|^2_{\mathcal{Y}_0} + \frac{\alpha}{2}\|u\|^2_{\mathcal{U}} \geq \min\left\{\frac{\alpha}{2k(\boldsymbol{\mu})^2}, \frac{\alpha}{2}\right\}\|x\|^2_{\mathcal{X}}.\end{aligned}$$

For the sake of brevity, we do not report the computations needed to prove the continuity of $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ which relies on the continuity assumption for $a(\cdot,\cdot;\boldsymbol{\mu})$ and $c(\cdot,\cdot;\boldsymbol{\mu})$. However, the interested reader may refer once again to [146, Theorem 1]. Although, we report the proof of the fulfillment of the *inf-sup condition* for the controlled state equation. First of all, let us consider $\zeta \in \mathcal{Y}_T$ and $\bar{y} \in \mathcal{Y}_0$ the solution of the auxiliary problem (1.43) in Lemma 1.2 where $v \equiv \zeta$. Then, for $x \neq 0$,

$$\begin{aligned}\sup_{x\in\mathcal{X}} \frac{\mathcal{B}(x,\zeta;\boldsymbol{\mu})}{\|x\|_{\mathcal{X}}\|\zeta\|_{\mathcal{Y}_0}} &= \sup_{x\in\mathcal{X}} \frac{\displaystyle\int_0^T s(y,\zeta;\boldsymbol{\mu})\,dt + \int_0^T a(y,\zeta;\boldsymbol{\mu})\,dt - \int_0^T c(u,\zeta;\boldsymbol{\mu})\,dt}{\|x\|_{\mathcal{X}}\|\zeta\|_{\mathcal{Y}_0}} \\ &\underbrace{\geq}_{x=(\bar{y},0)} \frac{\displaystyle\int_0^T \sum_i^{Q_S} c_S^i(\boldsymbol{\mu})\chi_{\Omega_S^i}\left\langle \bar{y}_t, \zeta\right\rangle dt + \int_0^T a(\bar{y},\zeta;\boldsymbol{\mu})\,dt}{\|\bar{y}\|_{\mathcal{Y}_0}\|\zeta\|_{\mathcal{Y}_0}} \\ &\geq \frac{\max\left\{\displaystyle\sum_i^{Q_S} c_S^i(\boldsymbol{\mu}), 1\right\}\displaystyle\int_0^T a(\zeta,\zeta;\boldsymbol{\mu})\,dt}{\|\bar{y}\|_{\mathcal{Y}_0}\|\zeta\|_{\mathcal{Y}_0}} \underbrace{\geq}_{\text{Lemma 1.2}} \frac{M_a(\boldsymbol{\mu})\|\zeta\|^2_{\mathcal{Y}}}{\bar{k}(\boldsymbol{\mu})\|\zeta\|^2_{\mathcal{Y}_0}} \\ &\underbrace{\geq}_{\text{Lemma 1.1}} \frac{M_a(\boldsymbol{\mu})}{6\bar{k}(\boldsymbol{\mu})} > 0,\end{aligned}$$

where $M_a(\boldsymbol{\mu}) := \max\left\{\sum_i^{Q_S} c_S^i(\boldsymbol{\mu}), 1\right\}\gamma_a(\boldsymbol{\mu})$. Since we have proved the inequality for all $\zeta \in \mathcal{Y}_T$, it holds:

$$\inf_{0\neq\zeta\in\mathcal{Y}_T}\sup_{0\neq x\in\mathcal{X}} \frac{\mathcal{B}(x,\zeta;\boldsymbol{\mu})}{\|x\|_{\mathcal{X}}\|\zeta\|_{\mathcal{Y}_0}} \geq \frac{M_a(\boldsymbol{\mu})}{6\bar{k}(\boldsymbol{\mu})} := \beta_{\mathcal{B}}(\boldsymbol{\mu}) > 0.$$

■

This theorem guarantees the existence and uniqueness of the optimal solution for time-dependent OCP($\boldsymbol{\mu}$)s, governed by parabolic state equations. The next Section deals with OCP($\boldsymbol{\mu}$)s governed by time-dependent Stokes equations. Also in this case, we will show how the saddle point structure is preserved and how Brezzi Theorem assures the well-posedness of such a formulation.

1.2.3 Time-Dependent Stokes OCP($\boldsymbol{\mu}$)s

This Section focuses on distributed OCP($\boldsymbol{\mu}$)s governed by time-dependent Stokes equations and provides a proof of well-posedness in a saddle point framework, as we did for the parabolic case in Section 1.2.2. We define $V := [H^1_{\Gamma_D}(\Omega)]^2$ and $P := L^2(\Omega)$. Thus, we consider

$$\mathcal{V}_0 := \left\{ v \in L^2(0,T;V) \text{ such that } v_t \in L^2(0,T;V^*) \text{ such that } v(0) = 0 \right\}$$

and $\mathcal{P} = L^2(0,T;P)$. The state variable is the pair $y := (v,p) \in \mathcal{Y}_0 := \mathcal{V}_0 \times \mathcal{P}$. It is influenced by the control variable $u \in \mathcal{U} := L^2(0,T;U)$, where $U \doteq [L^2(\Omega)]^2$. With analogous arguments w.r.t. Section 1.2.2, we define the velocity-pressure-control variable $x \in \mathcal{X} := \mathcal{Y}_0 \times \mathcal{U}$, where $x := (y,u)$. We endow the space $\mathcal{X}$ with the following norm

$$\|x\|^2_{\mathcal{X}} = \|v\|^2_{\mathcal{V}_0} + \|p\|^2_{\mathcal{P}} + \|u\|^2_{\mathcal{U}}.$$

In this specific context, we want to minimize the cost functional

$$\frac{1}{2}\|v - v_{\mathrm{d}}\|^2_{L^2(0,T;V_{\mathrm{obs}})} + \frac{\alpha}{2}\|u\|^2_{\mathcal{U}}, \tag{1.45}$$

under the following constraint

$$\begin{cases} v_t - \mu\Delta v + \nabla p = u & \text{in } \Omega \times (0,T), \\ \mathrm{div}(v) = 0 & \text{in } \Omega \times (0,T), \\ \dfrac{\partial v}{\partial n} = 0 & \text{on } \Gamma_N \times (0,T), \\ v = g & \text{on } \Gamma_D \times (0,T), \\ v(0) = v_0 & \text{in } \Omega \times \{0\}, \end{cases} \tag{1.46}$$

where $v_d \in L^2(0,T;V_{\mathrm{obs}})$ is a desired velocity profile taken in the observation domain Ω_{obs} and $V_{\mathrm{obs}} := L^2(\Omega)$. As the reader may notice, in this setting, we are assuming no geometrical parametrization, for the sake of clarity. However, all the following arguments can be easily extended to the case of geometrical parametrization, exploiting the same techniques of Section 1.2.2. Furthermore, for the same reason, we are restricting ourselves to the case where only a parameter is present, a diffusivity coefficient μ. In the numerical results of Section 4.2.2, we will treat a more complicated problem with geometrical parametrization, also. However, we decided that presenting this simpler setting is useful to state in a clearer way the framework we are dealing with. In order to build the optimality system, we exploit a Lagrangian argument after defining the adjoint variable $z := (z_v, z_p) \in \mathcal{Y}_T$, where $\mathcal{Y}_T := \mathcal{V}_T \times \mathcal{P}$, with

$$\mathcal{V}_T := \left\{ v \in L^2(0,T;V) \text{ such that } v_t \in L^2(0,T;V^*) \text{ such that } v(T) = 0 \right\}.$$

For a given $\mu \in \mathcal{D}$, we want to find $(x, z) \in \mathcal{X} \times \mathcal{Y}_T$ which solves:

$$\begin{cases} D_v\mathscr{L}((v,p),u,(z_v,z_p);\mu)[\omega_v] = 0 & \forall \omega_v \in \mathcal{V}_0, \\ D_p\mathscr{L}((v,p),u,(z_v,z_p);\mu)[\omega_p] = 0 & \forall \omega_p \in \mathcal{P}, \\ D_u\mathscr{L}((v,p),u,(z_v,z_p);\mu)[\kappa] = 0 & \forall \kappa \in \mathcal{U}, \\ D_{z_v}\mathscr{L}((v,p),u,(z_v,z_p);\mu)[\zeta_v] = 0 & \forall \zeta_v \in \mathcal{V}_0. \\ D_{z_p}\mathscr{L}((v,p),u,(z_v,z_p);\mu)[\zeta_p] = 0 & \forall \zeta_p \in \mathcal{P}. \end{cases} \tag{1.47}$$

The aforementioned problem can be rewritten in strong form as follows:

$$\begin{cases} v_t - \mu\Delta v + \nabla p = u & \text{in } \Omega \times (0,T), \\ \operatorname{div}(v) = 0 & \text{in } \Omega \times (0,T), \\ v = g & \text{on } \Gamma_D \times (0,T), \\ \dfrac{\partial v}{\partial n} = 0 & \text{on } \Gamma_N \times (0,T), \\ v(0) = v_0 & \text{in } \Omega \times \{0\}, \\ v - z_{vt} - \mu\Delta z_v + \nabla z_p = v_d & \text{in } \Omega \times (0,T), \\ \operatorname{div}(z_v) = 0 & \text{in } \Omega \times (0,T), \\ z_v(t) = 0 & \text{on } \partial\Omega \times (0,T), \\ z_v(T) = 0 & \text{in } \Omega \times \{T\}, \\ \alpha u = z_v & \text{in } \Omega \times (0,T), \\ \text{boundary conditions} & \text{on } \partial\Omega \times (0,T). \end{cases} \tag{1.48}$$

Now, we define the bilinear forms $\mathcal{A} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and $\mathcal{B} : \mathcal{X} \times \mathcal{Y}_0 \to \mathbb{R}$ and the functional $\mathcal{H} \in \mathcal{Y}^*$ as already did in (1.34), where the state equation is defined by

$$a(y,z;\mu) = \langle v_t, z_v\rangle + \mu \int_\Omega \nabla v \cdot \nabla z_v \, d\Omega - \int_\Omega p\operatorname{div}(z_v) \, d\Omega - \int_\Omega \operatorname{div}(v) z_p \, d\Omega. \tag{1.49}$$

Moreover, let us define the bilinear forms involved in the Stokes equations as:

$$\mathsf{a} : V \times V \to \mathbb{R}, \qquad \mathsf{a}(v,z;\mu) = \langle v_t, z_v\rangle + \mu \int_\Omega \nabla v \cdot \nabla z_v \, d\Omega, \tag{1.50}$$

$$\mathsf{b} : V \times P \to \mathbb{R}, \qquad \mathsf{b}(z_v,p) = -\int_\Omega p\operatorname{div}(z_v) \, d\Omega. \tag{1.51}$$

Here, we are facing a *nested saddle point*. Indeed, not only the global optimality system presents this peculiar structure, but also the state equation itself can be written as:

$$\begin{cases} \displaystyle\int_0^T \mathsf{a}(v,\zeta_v;\mu) \, dt + \int_0^T \mathsf{b}(\zeta_v,p) \, dt = 0 & \forall \zeta_v \in \mathcal{V}_0, \\ \displaystyle\int_0^T \mathsf{b}(v,\zeta_p) \, dt = 0 & \forall \zeta_p \in \mathcal{P}, \end{cases} \tag{1.52}$$

but, due to the linearity of the state equations, also the optimality system can recast written in a saddle point framework as:

$$\begin{cases} \mathcal{A}(x,\omega;\mu)+\mathcal{B}(\omega,z;\mu)=\displaystyle\int_0^T \langle \mathcal{H}(\mu),\omega\rangle\, dt & \forall \omega\in\mathcal{X}, \\ \mathcal{B}(x,\zeta;\mu)=0 & \forall \zeta\in\mathcal{Y}_T. \end{cases} \tag{1.53}$$

We want to prove the well-posedness of the system (1.53). To do so, also in this case, we rely on two Lemmas. The first one proves the well-posedness of the mixed Stokes optimality system (1.53), that will be exploited to guarantee the second hypothesis of the Brezzi Theorem for the controlled state equation $\mathcal{B}(\cdot,\cdot;\mu)$.

Lemma 1.3 *Time-dependent Stokes equations in their saddle point structure* (1.52) *verifies the Brezzi Theorem.*

Proof. First of all, it is a matter of simple computations to prove the continuity of $\mathsf{a}(\cdot,\cdot;\mu)$ and of $\mathsf{b}(\cdot,\cdot)$. The inf-sup condition for the bilinear form $\mathsf{b}(\cdot,\cdot)$ follows from [58, Theorem 4.7 and Proposition 2.2].
It still remains to prove that $\mathsf{a}(\cdot,\cdot;\mu)$ is weakly coercive over the kernel of $\mathsf{b}(\cdot,\cdot)$. In our case, the bilinear form $\mathsf{a}(\cdot,\cdot;\mu)$ is actually coercive over $\mathcal{V}_0$ since

$$\begin{aligned} \int_0^T \mathsf{a}(v,v;\mu)\, dt &= \int_0^T \langle v_t, v\rangle\, dt + \mu \int_0^T \int_\Omega \nabla v\cdot\nabla v\, d\Omega dt \\ &= \frac{1}{2}\|v(T)\|^2_{L^2(\Omega)} + \mu\|v\|^2_{\mathcal{V}} \\ &\geq \mu\|v\|^2_{\mathcal{V}} \frac{\|v\|^2_{\mathcal{V}_0}}{\|v\|^2_{\mathcal{V}_0}} \underbrace{\geq}_{\text{for Lemma 1.1}} \frac{\mu}{6}\|v\|^2_{\mathcal{V}_0}. \end{aligned}$$

■

We still need a Lemma to prove the coercivity of the bilinear form $\mathcal{A}(\cdot,\cdot)$. Also in this case, we denote the kernel of $\mathcal{B}(\cdot,\cdot,\mu)$ with $\mathcal{X}_0$. This Lemma proves a norm equivalence which will be used in the proof of the well-posedness of the whole optimality system (1.53).

Lemma 1.4 *On the space $\mathcal{X}_0$, the norm $\|\cdot\|^2_{\mathcal{X}}$ is equivalent to $\|\cdot\|^2_{\mathcal{V}_0}+\|\cdot\|^2_{\mathcal{U}}$.*

Proof. Let us consider $x=(y,u)$ in the kernel of $\mathcal{B}(\cdot,\cdot,\mu)$. By definition, $\|\cdot\|^2_{\mathcal{X}} \geq \|\cdot\|^2_{\mathcal{V}_0}+\|\cdot\|^2_{\mathcal{U}}$. This means that we only need to prove that there exists a positive constant $C_e(\mu)$ such that

$$\|\cdot\|^2_{\mathcal{X}} \leq C_e(\mu)(\|\cdot\|^2_{\mathcal{V}_0}+\|\cdot\|^2_{\mathcal{U}}). \tag{1.54}$$

If $x \in \mathcal{X}_0$, it holds that

$$\int_0^T \mathsf{b}(\omega, p)\, dt = \int_0^T c(u, \omega)\, dt - \mu \int_0^T \int_\Omega \nabla v \cdot \nabla \omega \, d\Omega dt - \int_0^T \langle v_t, \omega \rangle \, dt \qquad \forall \omega \in \mathcal{V}_0.$$

Then, we can derive the following inequalities for all $\omega \in \mathcal{V}_0$:

$$\begin{aligned} \mathsf{b}(\omega, p) &\leq \|u\|_{\mathcal{U}} \|\omega\|_{\mathcal{V}} + \mu \|v\|_{\mathcal{V}} \|\omega\|_{\mathcal{V}} + \|v_t\|_{\mathcal{V}^*} \|\omega\|_{\mathcal{V}} \\ &\leq \min\{1, \mu\} (\|v\|_{\mathcal{V}_0} + \|u\|_{\mathcal{U}}) \|\omega\|_{\mathcal{V}_0}. \end{aligned}$$

The above inequality does not depend on the choice of ω and p. Thus, from Lemma 1.3 we have the following relation

$$\bar{\beta} \|p\|_{\mathcal{P}} \|\omega\|_{\mathcal{V}_0} \leq \inf_{0 \neq p \in \mathcal{P}} \sup_{0 \neq \omega \in \mathcal{V}_0} \mathsf{b}(\omega, p) \leq \min\{1, \mu\} (\|v\|_{\mathcal{V}_0} + \|u\|_{\mathcal{U}}) \|\omega\|_{\mathcal{V}_0}.$$

Now, thanks to Young's inequality, we obtain

$$\|p\|_{\mathcal{P}}^2 \leq \frac{\min\{1, \mu\}^2}{\bar{\beta}^2} (\|v\|_{\mathcal{V}_0} + \|u\|_{\mathcal{U}})^2 \leq 2 \frac{\min\{1, \mu\}^2}{\bar{\beta}^2} (\|v\|_{\mathcal{V}_0}^2 + \|u\|_{\mathcal{U}}^2).$$

Thus, calling $\bar{C}_e(\mu) = \left(2 \frac{\min\{1, \mu\}^2}{\bar{\beta}^2}\right)$, the constant $C_e(\mu) = \min\{1, \bar{C}_e(\mu)\}$ verifies the inequality (1.54).

■

These two Lemmas will prove the following theorem concerning the existence and uniqueness of the optimal solution for OCP($\boldsymbol{\mu}$)s governed by time-dependent Stokes equations.

Theorem 1.5 *The saddle point problem* (1.53) *admits a unique solution.*

Proof. Once again we rely on the Brezzi Theorem to prove the well-posedness of the problem at hand. By definition, $\mathcal{A}(\cdot, \cdot)$ is trivially symmetric and positive definite. Futhermore, the same arguments presented in Theorem 3 can be applied in this specific case to show the continuity of the bilinear forms $\mathcal{A}(\cdot, \cdot)$ and $\mathcal{B}(\cdot, \cdot; \mu)$. Now, we will focus on the coercivity of $\mathcal{A}(\cdot, \cdot)$ over the kernel of the controlled state equation, i.e., there exists a positive constant C such that

$$\mathcal{A}(x, x) \geq C \|x\|_{\mathcal{X}}^2,$$

for all $x \in \mathcal{X}_0$. However, since we prove the equivalence between the norms $\|\cdot\|^2_{\mathcal{X}}$ and $\|\cdot\|^2_{\mathcal{V}_0} + \|\cdot\|^2_{\mathcal{U}}$ in Lemma 1.4, we will use the following relation

$$\mathcal{A}(x,x) \geq C(\|v\|^2_{\mathcal{V}_0} + \|u\|^2_{\mathcal{U}}).$$

Indeed, $x \in \mathcal{X}_0$ translates in a pair (v,p) that is solution to the Stokes equations with the control as forcing term. This implies the following well known result: the existence of a constant $k(\mu) > 0$ such that $\|v\|_{\mathcal{V}_0} \leq k(\mu)\|u\|_{\mathcal{U}}$, see [128, Chapter 13]. Thus, for $x \in \mathcal{X}_0$, it holds:

$$\begin{aligned}\mathcal{A}(x,x) &\geq \|v\|^2_{L^2(0,T;H)} + \alpha\|u\|^2_{\mathcal{U}} \geq \frac{\alpha}{2}\|u\|^2_{\mathcal{U}} + \frac{\alpha}{2}\|u\|^2_{\mathcal{U}} \geq \frac{\alpha}{2k(\mu)^2}\|v\|^2_{\mathcal{V}_0} + \frac{\alpha}{2}\|u\|^2_{\mathcal{U}} \\ &\geq \min\left\{\frac{\alpha}{2k(\mu)^2}, \frac{\alpha}{2}\right\}(\|v\|^2_{\mathcal{V}_0} + \|u\|^2_{\mathcal{U}}).\end{aligned}$$

It remains to prove that

$$\inf_{0\neq\zeta\in\mathcal{Y}_T} \sup_{0\neq x\in\mathcal{X}} \frac{\mathcal{B}(x,\zeta;\mu)}{\|x\|_{\mathcal{X}}\|\zeta\|_{\mathcal{Y}_T}} > \beta(\mu) > 0. \tag{1.55}$$

To this end, we consider the bilinear form $\mathsf{A} : \mathcal{Y}_0 \times \mathcal{Y}_T \to \mathbb{R}$ defined as

$$\mathsf{A}((v,p),(\zeta_v,\zeta_p);\mu) = \int_0^T \mathsf{a}(v,\zeta_v)\,dt + \int_0^T \mathsf{b}(\zeta_v,p)\,dt + \int_0^T \mathsf{b}(v,\zeta_p)\,dt.$$

We now follow [110, Appendix A.1]. Lemma 1.3 shows that the mixed operator $\mathsf{A}(\cdot,\cdot;\mu)$ is invertible and then the Babuška inf-sup constant $\beta_B(\mu)$ is well defined as

$$\begin{aligned}\beta_B(\mu) &= \inf_{0\neq(v,p)\in\mathcal{Y}_0} \sup_{0\neq(\zeta_v,\zeta_p)\in\mathcal{Y}_T} \frac{\mathsf{A}((v,p),(\zeta_v,\zeta_p);\mu)}{\|(v,p)\|_{\mathcal{Y}_0}\|(\zeta_v,\zeta_p)\|_{\mathcal{Y}_0}} \\ &= \inf_{0\neq(\zeta_v,\zeta_p)\in\mathcal{Y}_T} \sup_{0\neq(v,p)\in\mathcal{Y}_0} \frac{\mathsf{A}((v,p),(\zeta_v,\zeta_p);\mu)}{\|(v,p)\|_{\mathcal{Y}_0}\|(\zeta_v,\zeta_p)\|_{\mathcal{Y}_0}},\end{aligned}$$

see the classical reference [11]. We are now able to prove the following inequalities:

$$\begin{aligned}\sup_{0\neq x\in\mathcal{X}} \frac{\mathcal{B}(x,\zeta;\mu)}{\|x\|_{\mathcal{X}}\|\zeta\|_{\mathcal{Y}_T}} &\underbrace{\geq}_{x=(v,p,0)} \sup_{0\neq(v,p)\in\mathcal{Y}_0} \frac{\mathsf{A}((v,p),(\zeta_v,\zeta_p);\mu)}{\|(v,p)\|_{\mathcal{Y}_0}\|(\zeta_v,\zeta_p)\|_{\mathcal{Y}_0}} \\ &\geq \inf_{0\neq(\zeta_v,\zeta_p)\in\mathcal{Y}_0} \sup_{0\neq(v,p)\in\mathcal{Y}_0} \frac{\mathsf{A}((v,p),(\zeta_v,\zeta_p);\mu)}{\|(v,p)\|_{\mathcal{Y}_0}\|(\zeta_v,\zeta_p)\|_{\mathcal{Y}_0}} \geq \beta_B(\mu).\end{aligned}$$

The above inequality does not depend on the choice of $\zeta \in \mathcal{Y}_0$, then the inf-sup condition (1.55) is verified and the problem is well-posed.

■

In the previous Sections, we highlighted the saddle point structure for several linear state equations. We will now explore a new framework where the control is hidden in the action of the adjoint variable: the *no-control* framework, which will be the content of the following Section.

1.3 The No-Control Framework

This Section presents an equivalent formulation for the linear OCP($\boldsymbol{\mu}$)s we are dealing with. In this framework, we exploited the relation between control and adjoint variable to get rid of an equation. The usefulness of this peculiar formulation will be clearer in Chapter 5. For now, we will discuss this formulation at the continuous level for linear time-dependent and steady linear OCP($\boldsymbol{\mu}$)s, in Section 1.3.1 and 1.3.2 proposing and alternative proof of their well-posedness, as presented in [148].

1.3.1 Parabolic Time-Dependent OCP($\boldsymbol{\mu}$)s

We recall that in Section 1.2.2, after the differentiation of the Lagrangian functional (1.9), we obtain a three equation system that, in strong form, reads as (1.33). For the sake of clarity, we will call the first equation of system (1.33) *adjoint equation*, the second one *optimality equation* and the last one as *state equation*. It is natural to consider the relation given by the optimality equation

$$\alpha N(\boldsymbol{\mu})u - C(\boldsymbol{\mu})z\chi_{\Omega_u} = 0 \text{ in } \Omega \times (0,T). \tag{1.56}$$

By definition, $C(\boldsymbol{\mu})$ and $N(\boldsymbol{\mu})$ are both the L^2-scalar product over the the control domain (or its possible trace back). Thanks to this assumption the three equation system (1.33) reads: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y,z) \in \mathcal{Y}_0 \times \mathcal{Y}_T$ such that the following system is verified

$$\begin{cases} M(\boldsymbol{\mu})y\chi_{\Omega_{\text{obs}}} - S(\boldsymbol{\mu})z_t + D_a(\boldsymbol{\mu})^* z = M(\boldsymbol{\mu})y_d & \text{in } \Omega \times (0,T), \\ S(\boldsymbol{\mu})y_t + D_a(\boldsymbol{\mu})y - \dfrac{1}{\alpha}C(\boldsymbol{\mu})z\chi_{\Omega_u} = f & \text{in } \Omega \times (0,T), \\ y(0) = y_0 & \text{in } \Omega, \\ z(T) = 0 & \text{in } \Omega, \\ \text{boundary conditions} & \text{on } \partial\Omega \times (0,T). \end{cases} \tag{1.57}$$

We call the system (1.57) the *no-control* framework. From now on, in this Section, we will focus on this formulation where the control variable is eliminated from the system, following the structure presented in [89]. This choice is opposite to classical optimality systems used, for example, in [79, 80, 110, 111, 127, 146]. However, we do not lose any information. Indeed, the control variable can be recovered in post-processing using the optimality equation (1.56). The proposed system (1.57) can be recast in a mixed variational formulation as: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, z) \in \mathcal{Y}_0 \times \mathcal{Y}_T$ such that

$$\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu}) = \big\langle \mathcal{F}(\boldsymbol{\mu}),(\omega,\zeta)\big\rangle \quad \forall (\omega,\zeta) \in \mathcal{Q}\times\mathcal{Q}. \tag{1.58}$$

with

$$\begin{aligned}\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu}) = &\int_0^T s(y,\zeta;\boldsymbol{\mu})\,dt + \int_0^T a(y,\zeta;\boldsymbol{\mu})\,dt \\ &- \frac{1}{\alpha}\int_0^T c(z,\zeta;\boldsymbol{\mu})\,dt + \int_0^T m(y,\omega;\boldsymbol{\mu})\,dt \\ &- \int_0^T s(z,\omega;\boldsymbol{\mu})\,dt + \int_0^T a(\omega,z;\boldsymbol{\mu})\,dt.\end{aligned} \tag{1.59}$$

and

$$\big\langle \mathcal{F}(\boldsymbol{\mu}),(\omega,\zeta)\big\rangle = \int_0^T m(y_d,\omega;\boldsymbol{\mu})\,dt + \int_0^T \big\langle G(\boldsymbol{\mu}),\zeta\big\rangle\,dt. \tag{1.60}$$

In order to prove the well-posedness of (1.58), we want to exploit the Nečas-Babuška theorem [108]. It is straightforward to prove the continuity of the bilinear form (1.59) thanks to the Assumptions 1 and the definition (1.31). Indeed there exists a positive constant $c_{\underline{\mathcal{B}}}(\boldsymbol{\mu})$ such that:

$$\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu}) \le c_{\underline{\mathcal{B}}}(\boldsymbol{\mu})\sqrt{\|y\|^2_{\mathcal{Y}_0} + \|z\|^2_{\mathcal{Y}_T}}\sqrt{\|\omega\|^2_{\mathcal{Q}} + \|\zeta\|^2_{\mathcal{Q}}}. \tag{1.61}$$

From now on, we will need the following relations, that are verified thanks to the assumptions $Y \subset Y_{\text{obs}}$ and $Y \subset U$ (natural consequence of the definitions of Section 1.2.2):

$$\|y\|_{Y_{\text{obs}}} \le c_{\text{obs}}\|y\|_Y, \qquad \forall y \in Y, \tag{1.62}$$

$$\|y\|_U \le c_u\|y\|_Y, \qquad \forall y \in Y. \tag{1.63}$$

where c_{obs} and c_u are positive constants. Moreover, we will use two Lemmas to prove the injectivity and surjectivity of (1.59), to recover the hypotheses of the Nečas-Babuška theorem. The first one is about the surjectivity of the adjoint form of (1.59). The proof combines strategies from [156, Proposition 2.2] that were successfully exploited for the parabolic equations and the techniques presented in [89] for distributed OCP($\boldsymbol{\mu}$)s.

Lemma 1.6 (Surjectivity of $\underline{\mathcal{B}}^*$) *There exists $\beta(\boldsymbol{\mu}) > 0$ such that $\beta_{\underline{\mathcal{B}}}(\boldsymbol{\mu}) > \beta(\boldsymbol{\mu})$, where*

$$\beta_{\underline{\mathcal{B}}}(\boldsymbol{\mu}) := \inf_{(y,z)\in(\mathcal{Y}_0\times\mathcal{Y}_T)} \sup_{(\omega,\zeta)\in(\mathcal{Q}\times\mathcal{Q})} \frac{\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu})}{\sqrt{\|y\|^2_{\mathcal{Y}_0}+\|z\|^2_{\mathcal{Y}_T}}\sqrt{\|\omega\|^2_{\mathcal{Q}}+\|\zeta\|^2_{\mathcal{Q}}}}, \quad (1.64)$$

with both $(y,z)\neq 0$ and $(\omega,\zeta)\neq 0$.

Proof. Let us consider $0\neq(y,z)\in\mathcal{Y}_0\times\mathcal{Y}_T$ and let us define

$$\zeta_y = (D_a(\boldsymbol{\mu})^*)^{-1}y_t \quad \text{and} \quad \omega_z = -D_a(\boldsymbol{\mu})^{-1}z_t. \quad (1.65)$$

Case 1. We focus first on the case $\Omega_u = \Omega_{\text{obs}}$.
In this case, $c(\cdot,\cdot,\boldsymbol{\mu}) \equiv m(\cdot,\cdot;\boldsymbol{\mu})$, since both represent the L^2-scalar product over the same domain, due to the coincidence of the control and the observation domains. Furthermore, it is clear that $(\alpha c_z y + c_{\zeta_y}\zeta_y, c_z z + \omega_z) \in \mathcal{Q}\times\mathcal{Q}$, where the positive constants c_{ζ_y} and c_z will be determined afterwards. Thus, for $(\omega,\zeta)\neq 0$, we can state the following:

$$\begin{aligned}
\sup_{(\omega,\zeta)\in(\mathcal{Q}\times\mathcal{Q})} \underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu}) &\geq \underline{\mathcal{B}}((y,z),(\alpha c_z y + c_{\zeta_y}\zeta_y, c_z z+\omega_z);\boldsymbol{\mu}) \\
&\geq \frac{\alpha}{2}c_z c_S(\boldsymbol{\mu})\|y(T)\|^2_H + \alpha c_z\gamma_a(\boldsymbol{\mu})\|y\|^2_{\mathcal{Q}} - c_z\int_0^T c(z,y;\boldsymbol{\mu})\,dt \\
&\quad + c_{\zeta_y}s(y,\zeta_y;\boldsymbol{\mu}) + c_{\zeta_y}\int_0^T a(y,\zeta_y;\boldsymbol{\mu})\,dt - \frac{c_{\zeta_y}}{\alpha}\int_0^T c(z,\zeta_y;\boldsymbol{\mu})\,dt \\
&\quad + c_z\int_0^T m(y,z;\boldsymbol{\mu})\,dt + \frac{c_S(\boldsymbol{\mu})}{2}\|z(0)\|^2_H + c_z\gamma_a(\boldsymbol{\mu})\|z\|^2_{\mathcal{Q}} \\
&\quad + \int_0^T m(y,\omega_z;\boldsymbol{\mu})\,dt - s(z,\omega_z;\boldsymbol{\mu}) + \int_0^T a(\omega_z,z;\boldsymbol{\mu})\,dt.
\end{aligned}$$

Here, we used the coercivity of $a(\cdot,\cdot;\boldsymbol{\mu})$ and the following relation

$$\begin{aligned}
\int_0^T s(w,w;\boldsymbol{\mu})\,dt &= \frac{1}{2}\int_0^T \sum_i^{Q_S} c_S^i(\boldsymbol{\mu})\chi_{\Omega_S^i}\frac{d\|w(t)\|^2_H}{dt}\,dt \\
&\geq \underbrace{\min_i\{c_s^i(\boldsymbol{\mu})\}}_{c_S(\boldsymbol{\mu})}\left(\frac{1}{2}\|w(T)\|^2_H - \frac{1}{2}\|w(0)\|^2_H\right) \quad (1.66)\\
&\geq c_S(\boldsymbol{\mu})\frac{1}{2}\|w(T)\|^2_H \qquad \forall w\in\mathcal{Y}_0,
\end{aligned}$$

which reads, analogously, $-s(w,w;\boldsymbol{\mu}) \geq c_S(\boldsymbol{\mu})\frac{1}{2}\|w(0)\|$ for all $w \in \mathcal{Y}_T$. Furthermore, we observe that

$$a(y,\zeta_y;\boldsymbol{\mu}) = \Big\langle D_a(\boldsymbol{\mu})y, (D_a(\boldsymbol{\mu})^*)^{-1}y_t \Big\rangle_{Y^*Y} = \Big\langle y, y_t \Big\rangle_{Y^*Y} = \frac{1}{2}\frac{d\|y(t)\|_H^2}{dt}, \tag{1.67}$$

and

$$a(\omega_z,z;\boldsymbol{\mu}) = -\Big\langle D_a(\boldsymbol{\mu})(D_a(\boldsymbol{\mu}))^{-1}z_t, z \Big\rangle_{Y^*Y} = -\Big\langle z_t, z \Big\rangle_{Y^*Y} = -\frac{1}{2}\frac{d\|z(t)\|_H^2}{dt}, \tag{1.68}$$

which, if integrated in time, are non negative quantities due to the initial and final time conditions for the state and the adjoint variable, respectively. Furthermore, we recall that $m(y,\omega;\boldsymbol{\mu}) = c(\omega,y;\boldsymbol{\mu})$ since $\Omega_u = \Omega_{\text{obs}}$. Exploiting the inequalities

$$s(y,\zeta_y;\boldsymbol{\mu}) = \sum_i^{Q_S} c_S^i(\boldsymbol{\mu})\chi_{\Omega_S^i} a(\zeta_y,\zeta_y;\boldsymbol{\mu}) \geq c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})\|\zeta_y\|_Y^2, \tag{1.69}$$

and

$$-s(z,\omega_z;\boldsymbol{\mu}) = \sum_i^{Q_S} c_S^i(\boldsymbol{\mu})\chi_{\Omega_S^i} a(\omega_z,\omega_z;\boldsymbol{\mu}) \geq c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})\|\omega_z\|_Y^2, \tag{1.70}$$

together with the Young's inequality and the continuity assumption of the bilinear form $m(\cdot,\cdot;\boldsymbol{\mu})$, we can state that, for $(\omega,\zeta) \neq 0$,

$$\begin{aligned}
\sup_{(\omega,\zeta)\in(\mathcal{Q}\times\mathcal{Q})} \underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu}) &\geq \alpha c_z\gamma_a(\boldsymbol{\mu})\|y\|_{\mathcal{Q}}^2 + c_{\zeta_y}c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})\|\zeta_y\|_{\mathcal{Q}}^2 \\
&\quad - \frac{c_c(\boldsymbol{\mu})c_{\zeta_y}}{\alpha}\int_0^T \|z\|_U\|\zeta_y\|_Y \, dt + c_z\gamma_a(\boldsymbol{\mu})\|z\|_{\mathcal{Q}}^2 \\
&\quad - c_m(\boldsymbol{\mu})\int_0^T \|y\|_Y\|\omega_z\|_Y \, dt + c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})\|\omega_z\|_{\mathcal{Q}}^2 \\
&\geq \frac{\alpha c_z\gamma_a(\boldsymbol{\mu})}{2}\|y\|_{\mathcal{Q}}^2 + \Big(\frac{\alpha c_z\gamma_a(\boldsymbol{\mu})}{2} - \frac{c_m(\boldsymbol{\mu})}{2\eta_2}\Big)\|y\|_{\mathcal{Q}}^2 \\
&\quad + \frac{c_{\zeta_y}c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2}\|\zeta_y\|_{\mathcal{Q}}^2 + \Big(\frac{c_{\zeta_y}c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2} - \frac{c_c(\boldsymbol{\mu})c_{\zeta_y}\eta_1}{2\alpha}\Big)\|\zeta_y\|_{\mathcal{Q}}^2 \\
&\quad + \frac{c_z\gamma_a(\boldsymbol{\mu})}{2}\|z\|_{\mathcal{Q}}^2 + \Big(\frac{c_z\gamma_a(\boldsymbol{\mu})}{2} - \frac{c_c(\boldsymbol{\mu})c_{\zeta_y}c_u(\boldsymbol{\mu})}{2\alpha\eta_1}\Big)\|z\|_{\mathcal{Q}}^2 \\
&\quad + \frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2}\|\omega_z\|_{\mathcal{Q}}^2 + \Big(\frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2} - \frac{c_m(\boldsymbol{\mu})\eta_2}{2}\Big)\|\omega_z\|_{\mathcal{Q}}^2,
\end{aligned}$$

for some positive η_1 and η_2, and $c_u(\boldsymbol{\mu})$ is the constant of (1.63). Choosing

$$\eta_1 = \frac{\alpha c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{c_c(\boldsymbol{\mu})}, \quad c_{\zeta_y} = \frac{c_z c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})^2\alpha^2}{c_c(\boldsymbol{\mu})^2 c_u(\boldsymbol{\mu})},$$

$$\eta_2 = \frac{\gamma_a(\boldsymbol{\mu})c_S(\boldsymbol{\mu})}{c_m(\boldsymbol{\mu})}, \text{ and } c_z = \frac{c_m(\boldsymbol{\mu})^2}{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})^2\alpha},$$

it holds, for $(\omega, \zeta) \neq 0$,

$$\begin{aligned}\sup_{(\omega,\zeta)\in(\mathcal{Q}\times\mathcal{Q})} \underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu}) \geq{}& \frac{c_m(\boldsymbol{\mu})^2}{2c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}\|y\|_{\mathcal{Q}}^2 \\ &+ \frac{c_m(\boldsymbol{\mu})^2 c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})\alpha}{2c_c(\boldsymbol{\mu})^2 c_u(\boldsymbol{\mu})}\|\zeta_y\|_{\mathcal{Q}}^2 \\ &+ \frac{c_m(\boldsymbol{\mu})^2}{2c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})\alpha}\|z\|_{\mathcal{Q}}^2 \\ &+ \frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2}\|\omega_z\|_{\mathcal{Q}}^2.\end{aligned}$$

We now use that

$$\begin{aligned}\|y_t\|_{Y^*} &= \|D_a(\boldsymbol{\mu})^*\zeta_y\| \\ &\leq c_a(\boldsymbol{\mu})\|\zeta_y\|_Y\end{aligned} \quad \text{and} \quad \begin{aligned}\|\zeta_t\|_{Y^*} &= \|-D_a(\boldsymbol{\mu})\omega_z\| \\ &\leq c_a(\boldsymbol{\mu})\|\omega_z\|_Y,\end{aligned} \tag{1.71}$$

thus, for $(\omega, \zeta) \neq 0$,

$$\begin{aligned}&\sup_{(\omega,\zeta)\in(\mathcal{Q}\times\mathcal{Q})} \underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu}) \\ &\geq \min\left\{\frac{c_m(\boldsymbol{\mu})^2}{2c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}, \frac{c_m(\boldsymbol{\mu})^2 c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})\alpha}{2c_c(\boldsymbol{\mu})^2 c_u(\boldsymbol{\mu})c_a(\boldsymbol{\mu})^2}, \frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2c_a(\boldsymbol{\mu})^2}\right\}(\|y\|_{\mathcal{Y}_0}^2 + \|z\|_{\mathcal{Y}_T}^2).\end{aligned}$$

Now, we will tackle the denominator of (1.64). Defining

$$\beta_a(\boldsymbol{\mu}) := \inf_{\phi\in Y\setminus\{0\}} \sup_{\psi\in Y\setminus\{0\}} \frac{a(\psi,\phi;\boldsymbol{\mu})}{\|\phi\|_Y\|\psi\|_Y}, \tag{1.72}$$

we have, being $\alpha \leq 1$,

$$\begin{aligned}
&\sqrt{\|\alpha c_z y + c_{\zeta_y}\zeta_y\|_{\mathcal{Q}}^2 + \|c_z z + \omega_z\|_{\mathcal{Q}}^2} \\
&\qquad \leq \sqrt{2(\alpha^2 c_z^2 \|y\|_{\mathcal{Q}}^2 + c_{\zeta_y}^2 \|\zeta_y\|_{\mathcal{Q}}^2 + c_z^2 \|z\|_{\mathcal{Q}}^2 + \|\omega_z\|_{\mathcal{Q}}^2)} \\
&\qquad\qquad \leq \sqrt{\begin{aligned} &2\Big(\alpha^2 c_z^2 \|y\|_{\mathcal{Q}}^2 + \frac{c_{\zeta_y}^2}{\beta_a(\boldsymbol{\mu})^2}\int_0^T \|y_t\|_{Y^*}^2 \, dt \\ &\qquad + c_z^2 \|z\|_{\mathcal{Q}}^2 + \frac{1}{\beta_a(\boldsymbol{\mu})^2}\int_0^T \|z_t\|_{Y^*}^2 \Big) \, dt \end{aligned}} \\
&\qquad\qquad\qquad \leq \sqrt{2\max\left\{c_z^2, \frac{c_{\zeta_y}^2}{\beta_a(\boldsymbol{\mu})^2}, \frac{1}{\beta_a(\boldsymbol{\mu})^2}\right\}(\|y\|_{\mathcal{Y}_0}^2 + \|z\|_{\mathcal{Y}_T}^2)}.
\end{aligned}$$

The proposed estimates do not depend on the choice of $(y, z) \in \mathcal{Y}_0 \times \mathcal{Y}_T$, then, relation (1.64) holds with

$$\beta(\boldsymbol{\mu}) = \frac{\min\left\{\frac{c_m(\boldsymbol{\mu})^2}{2c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}, \frac{c_m(\boldsymbol{\mu})^2 c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})\alpha}{2c_c(\boldsymbol{\mu})^2 c_u(\boldsymbol{\mu})c_a(\boldsymbol{\mu})^2}, \frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2c_a(\boldsymbol{\mu})^2}\right\}}{\sqrt{2\max\left\{c_z^2, \frac{c_{\zeta_y}^2}{\beta_a(\boldsymbol{\mu})^2}, \frac{1}{\beta_a(\boldsymbol{\mu})^2}\right\}}}.$$

Case 2. We now take into account the case $\Omega_u \neq \Omega_{\text{obs}}$. This assumption means that at least one of the two between the control and the observation domain is not Ω. Indeed, if both are Ω, once again, we are considering Case 1. In this proof we choose[1] $\Omega_{\text{obs}} \neq \Omega$. Also in this case, inequality (1.64) holds. Indeed, let us define $\overline{\kappa} := \overline{\kappa}(\boldsymbol{\mu}) \in \mathcal{Y}_0$ solution of the following auxiliary problem for a given $\boldsymbol{\mu} \in \mathcal{D}$, a positive constant c_y and a given $y \in \mathcal{Q}$:

$$\begin{cases}
\begin{aligned} &\int_0^T s(\overline{\kappa}, r; \boldsymbol{\mu}) \, dt \\ &\quad + \int_0^T a(\overline{\kappa}, r; \boldsymbol{\mu}) \, dt = \\ &\qquad - \int_0^T m(y, r; \boldsymbol{\mu}) \, dt + \frac{c_y}{\alpha}\int_0^T c(r, y; \boldsymbol{\mu}) \, dt \end{aligned} & \forall r \in \mathcal{Q}, \\
\overline{\kappa}(0) = 0 & \text{in } \Omega, \\
\overline{\kappa} \equiv 0 & \text{in } \Omega_{\text{obs}}.
\end{cases} \tag{1.73}$$

The parabolic problem (1.73) is well-posed: indeed $m(y, r; \boldsymbol{\mu})$ and $c(r, y; \boldsymbol{\mu})$ are continuous and the continuity and coercivity of $a(\cdot, \cdot; \boldsymbol{\mu})$ are preserved in $\Omega \backslash \Omega_{\text{obs}}$. We consider the element $(c_y y + c_{\zeta_y}\zeta_y, z + \omega_z + \overline{\kappa}) \in \mathcal{Q} \times \mathcal{Q}$. Also in this case, we

[1] The choice has been driven by the numerical experiments we will show in Section 5.2.2. However, we will propose a generalization to $\Omega_u \neq \Omega$ in Remark 1.3.1.

want to determine the value of the two constants c_y, and c_{ζ_y}. Thus, it holds, for $(\omega, \zeta) \neq 0$,

$$\begin{aligned}
\sup_{(\omega,\zeta)\in(\mathcal{Q}\times\mathcal{Q})} \underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu}) &\geq \underline{\mathcal{B}}((y,z),(c_y y + c_{\zeta_y}\zeta_y, z+\omega_z+\overline{\kappa});\boldsymbol{\mu}) \\
&\geq c_S(\boldsymbol{\mu})\frac{c_y}{2}\|y(T)\|_H^2 + c_y\gamma_a(\boldsymbol{\mu})\|y\|_{\mathcal{Q}}^2 - \frac{c_y}{\alpha}\int_0^T c(z,y;\boldsymbol{\mu})\,dt \\
&\quad + c_{\zeta_y}s(y,\zeta_y;\boldsymbol{\mu}) + c_{\zeta_y}\int_0^T a(y,\zeta_y;\boldsymbol{\mu})\,dt - \frac{c_{\zeta_y}}{\alpha}\int_0^T c(z,\zeta_y;\boldsymbol{\mu})\,dt \\
&\quad + \int_0^T m(y,z;\boldsymbol{\mu})\,dt + \frac{c_S(\boldsymbol{\mu})}{2}\|z(0)\|_H^2 + \gamma_a(\boldsymbol{\mu})\|z\|_{\mathcal{Q}}^2 \\
&\quad + \int_0^T m(y,\omega_z;\boldsymbol{\mu})\,dt - \int_0^T s(z,\omega_z;\boldsymbol{\mu})\,dt \\
&\quad + \int_0^T a(\omega_z,z;\boldsymbol{\mu})\,dt + \int_0^T m(y,\overline{\kappa};\boldsymbol{\mu})\,dt \\
&\quad - \int_0^T s(z,\overline{\kappa};\boldsymbol{\mu})\,dt + \int_0^T a(\overline{\kappa},z;\boldsymbol{\mu})\,dt
\end{aligned}$$

Now, exploiting the definition of $\overline{\kappa}$ in (1.73) we obtain:

$$\begin{aligned}
-\int_0^T s(\overline{\kappa},z;\boldsymbol{\mu})\,dt &= \int_\Omega \overline{\kappa}(0)z(0)\,d\Omega - \int_\Omega \overline{\kappa}(T)z(T)\,d\Omega + \int_0^T s(\overline{\kappa},z;\boldsymbol{\mu})\,dt \\
&= \int_0^T s(\overline{\kappa},z;\boldsymbol{\mu})\,dt,
\end{aligned}$$

and $m(y,\overline{\kappa};\boldsymbol{\mu}) = 0$, which combined with (1.66), (1.67), (1.68), (1.69), (1.70) and the hypotheses (b), together with the definition of $n(\cdot,\cdot;\boldsymbol{\mu})$ and $m(\cdot,\cdot;\boldsymbol{\mu})$,

we get the following relation for $(\omega,\zeta)\neq 0$,

$$\begin{aligned}\sup_{(\omega,\zeta)\in(\mathcal{Q}\times\mathcal{Q})}\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu}) &\geq c_y\gamma_a(\boldsymbol{\mu})\|y\|^2_{\mathcal{Q}}+c_{\zeta_y}c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})\|\zeta_y\|^2_{\mathcal{Q}}\\ &\quad-\frac{c_c(\boldsymbol{\mu})c_{\zeta_y}}{\alpha}\int_0^T\|z\|_U\|\zeta_y\|_Y\,dt+\gamma_a(\boldsymbol{\mu})\|z\|^2_{\mathcal{Q}}\\ &\quad-c_m(\boldsymbol{\mu})\int_0^T\|y\|_Y\|\omega_z\|_Y\,dt+c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})\|\omega_z\|^2_{\mathcal{Q}}\\ &\geq\frac{c_y\gamma_a(\boldsymbol{\mu})}{2}\|y\|^2_{\mathcal{Q}}+\Big(\frac{c_y\gamma_a(\boldsymbol{\mu})}{2}-\frac{c_m(\boldsymbol{\mu})}{2\eta_2}\Big)\|y\|^2_{\mathcal{Q}}\\ &\quad+\frac{c_{\zeta_y}c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2}\|\zeta_y\|^2_{\mathcal{Q}}+\Big(\frac{c_{\zeta_y}c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2}-\frac{c_c(\boldsymbol{\mu})c_{\zeta_y}\eta_1}{2\alpha}\Big)\|\zeta_y\|^2_{\mathcal{Q}}\\ &\quad+\frac{\gamma_a(\boldsymbol{\mu})}{2}\|z\|^2_{\mathcal{Q}}+\Big(\frac{\gamma_a(\boldsymbol{\mu})}{2}-\frac{c_c(\boldsymbol{\mu})c_{\zeta_y}c_u(\boldsymbol{\mu})}{2\alpha\eta_1}\Big)\|z\|^2_{\mathcal{Q}}\\ &\quad+\frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2}\|\omega_z\|^2_{\mathcal{Q}}+\Big(\frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2}-\frac{c_m(\boldsymbol{\mu})\eta_2}{2}\Big)\|\omega_z\|^2_{\mathcal{Q}},\end{aligned}$$

for some positive η_1 and η_2, resulting from the Young's inequality. Thus, with

$$\eta_1=\frac{\alpha c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{c_c(\boldsymbol{\mu})},\quad c_{\zeta_y}=\frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})^2\alpha^2}{c_c(\boldsymbol{\mu})^2c_u(\boldsymbol{\mu})},$$

$$\eta_2=\frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{c_m(\boldsymbol{\mu})},\quad\text{and}\quad c_y=\frac{c_m(\boldsymbol{\mu})^2}{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})^2\alpha},$$

and exploiting (1.71), we obtain

$$\begin{aligned}&\sup_{(\omega,\zeta)\in\mathcal{Q}\times\mathcal{Q}}\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu})\geq\\ &\min\Big\{\frac{c_m(\boldsymbol{\mu})^2}{2c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})},\frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})^3\alpha^2}{2c_c(\boldsymbol{\mu})^2c_u(\boldsymbol{\mu})c_a(\boldsymbol{\mu})^2},\frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2c_a(\boldsymbol{\mu})^2},\frac{\gamma_a(\boldsymbol{\mu})}{2}\Big\}(\|y\|^2_{\mathcal{Y}_0}+\|z\|^2_{\mathcal{Y}_T}).\end{aligned}$$

It remains to estimate the denominator (1.64). Thus, we use $r=\overline{\kappa}$ in (1.73) and the relation (1.66), to prove the following inequality:

$$\gamma_a(\boldsymbol{\mu})\|\overline{\kappa}\|^2_{\mathcal{Q}}\leq\frac{c_yc_c(\boldsymbol{\mu})c_u}{\alpha}\|\overline{\kappa}\|_{\mathcal{Q}}\|y\|_{\mathcal{Q}}\Rightarrow\|\overline{\kappa}\|_{\mathcal{Q}}\leq\frac{c_yc_c(\boldsymbol{\mu})c_u}{\alpha\gamma_a(\boldsymbol{\mu})}\|y\|_{\mathcal{Q}}.\tag{1.74}$$

This implies that

$$\begin{aligned}&\sqrt{\|\alpha c_y y+c_{\zeta_y}\zeta_y\|^2_{\mathcal{Q}}+\|z+\omega_z+\overline{\kappa}\|^2_{\mathcal{Q}}}\\ &\qquad\leq\sqrt{2(c_y^2\|y\|^2_{\mathcal{Q}}+c^2_{\zeta_y}\|\zeta_y\|^2_{\mathcal{Q}}+\|z\|^2_{\mathcal{Q}}+\|\omega_z\|^2_{\mathcal{Q}}+\|\overline{\kappa}\|^2_{\mathcal{Q}})}\\ &\qquad\leq\sqrt{2\max\left\{c_y^2,\frac{c^2_{\zeta_y}}{\beta_a(\boldsymbol{\mu})^2},\frac{1}{\beta_a(\boldsymbol{\mu})^2},\left(\frac{c_yc_c(\boldsymbol{\mu})c_u}{\alpha\gamma_a(\boldsymbol{\mu})}\right)^2\right\}(\|y\|^2_{\mathcal{Y}_0}+\|z\|^2_{\mathcal{Y}_T})}.\end{aligned}$$

Namely, the surjectivity condition (1.64) holds with

$$\beta(\boldsymbol{\mu}) = \frac{\min\left\{\frac{c_m(\boldsymbol{\mu})^2}{2c_S(\boldsymbol{\mu})\gamma_s(\boldsymbol{\mu})}, \frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})^3\alpha^2}{2c_c(\boldsymbol{\mu})^2 c_u(\boldsymbol{\mu})c_a(\boldsymbol{\mu})^2}, \frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2c_a(\boldsymbol{\mu})^2}, \frac{\gamma_a(\boldsymbol{\mu})}{2}\right\}}{\sqrt{2\max\left\{c_y^2, \frac{c_{\zeta_y}^2}{\beta_a(\boldsymbol{\mu})^2}, \frac{1}{\beta_a(\boldsymbol{\mu})^2}, \left(\frac{c_y c_c(\boldsymbol{\mu})c_u}{\alpha\gamma_a(\boldsymbol{\mu})}\right)^2\right\}}} > 0. \tag{1.75}$$

■

Besides the surjectivity inequality (1.64) for OCP($\boldsymbol{\mu}$)s governed by linear time-dependent parabolic equations, in order to guarantee the well-posedness of the whole optimality system (1.58), we need to exploit another Lemma, that combined to Lemma 1.6, will assure the existence and uniqueness of the optimal solution to (1.58).

Lemma 1.7 (Injectivity of $\underline{\mathcal{B}}^*$) *The bilinear form (1.59) satisfies the following inf-sup stability condition:*

$$\inf_{(\omega,\zeta)\in(\mathcal{Q}\times\mathcal{Q})}\sup_{(y,z)\in(\mathcal{Y}_0\times\mathcal{Y}_T)}\frac{\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu})}{\sqrt{\|y\|_{\mathcal{Y}_0}^2+\|z\|_{\mathcal{Y}_T}^2}\sqrt{\|\omega\|_{\mathcal{Q}}^2+\|\zeta\|_{\mathcal{Q}}^2}} > 0. \tag{1.76}$$

for $(\omega,\zeta)\neq 0$ and $(y,z)\neq 0$.

Proof. First of all, we devide the proof in two cases, one dealing with $\Omega_u = \Omega_{\text{obs}}$ and otherwise.
Case 1. Let us consider $\Omega_u = \Omega_{\text{obs}}$. In this specific case, for every $r, w \in Y$, the action of $m(r,w;\boldsymbol{\mu})$ and $c(w,r;\boldsymbol{\mu})$ coincide. It is clear that, for $(y,z)\neq 0$

$$\sup_{(y,z)\in(\mathcal{Y}_0\times\mathcal{Y}_T)}\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu}) \geq \underline{\mathcal{B}}((1/\alpha)\bar{y},\bar{z}),(\omega,\zeta);\boldsymbol{\mu}),$$

where $\bar{y}\in\mathcal{Y}_0$ and , $\bar{z}\in\mathcal{Y}_T$ have been properly chosen with these properties:

$$\bar{y}_t = \zeta \quad \text{and} \quad \bar{z}_t = -\omega. \tag{1.77}$$

This implies

$$\begin{aligned}\underline{\mathcal{B}}((1/\alpha)\bar{y},\bar{z}),(\omega,\zeta);\boldsymbol{\mu}) = &\frac{1}{\alpha}\int_0^T s(\bar{y},\zeta;\boldsymbol{\mu})\,dt + \int_0^T \frac{1}{\alpha}a(\bar{y},\zeta;\boldsymbol{\mu})\,dt\\ &-\frac{1}{\alpha}\int_0^T c(\bar{z},\zeta;\boldsymbol{\mu})\,dt + \frac{1}{\alpha}\int_0^T m(\bar{y},\omega;\boldsymbol{\mu})\,dt\\ &-\int_0^T s(\bar{z},\omega;\boldsymbol{\mu})\,dt + \int_0^T a(\omega,\bar{z};\boldsymbol{\mu})\,dt.\end{aligned} \tag{1.78}$$

Thanks to (1.77), we notice that:

$$s(\bar{y},\zeta;\boldsymbol{\mu}) \geq c_S(\boldsymbol{\mu})\|\zeta\|^2_{\mathcal{Q}} \quad \text{and} \quad -s(\bar{z},\omega;\boldsymbol{\mu}) \geq c_S(\boldsymbol{\mu})\|\omega\|^2_{\mathcal{Q}}. \tag{1.79}$$

Furthermore, by the definition of $c(\cdot,\cdot,\boldsymbol{\mu})$ and $m(\cdot,\cdot,\boldsymbol{\mu})$ and from the time boundary conditions for state and adjoint variables, we obtain the following relation

$$\begin{aligned} -\int_0^T c(\bar{z},\zeta;\boldsymbol{\mu})\,dt &= -\int_0^T \int_\Omega \sum_i^{Q_C} c_C(\boldsymbol{\mu})^i \chi_{\Omega_C^i}\bar{z}\bar{y}_t\,d\Omega dt = \\ &= \int_0^T \int_\Omega \sum_i^{Q_C} c_C(\boldsymbol{\mu})^i \chi_{\Omega_C^i}\bar{y}\bar{z}_t\,d\Omega dt = -\int_0^T m(\bar{y},\omega;\boldsymbol{\mu})\,dt. \end{aligned} \tag{1.80}$$

The aforementioned properties implies

$$\begin{aligned} \underline{\mathcal{B}}((1/\alpha)\bar{y},\bar{z}),(\omega,\zeta);\boldsymbol{\mu}) \geq \frac{c_S(\boldsymbol{\mu})}{\alpha}\|\zeta\|^2_{\mathcal{Q}} + \frac{1}{\alpha}\int_0^T a(\bar{y},\bar{y}_t;\boldsymbol{\mu})\,dt \\ + \|\omega\|^2_{\mathcal{Q}} - \int_0^T a(\bar{z}_t,\bar{z};\boldsymbol{\mu})\,dt. \end{aligned}$$

Here, we are assuming that the time derivative commutes with the bilinear form operators[2]. Thanks to this assumption, we finally prove the relation

$$\begin{aligned} \underline{\mathcal{B}}((1/\alpha)\bar{y},\bar{z}),(\omega,\zeta);\boldsymbol{\mu}) &\geq \frac{c_S(\boldsymbol{\mu})}{\alpha}\|\zeta\|^2_{\mathcal{Q}} + \frac{1}{2\alpha}\int_0^T \frac{d\,a(\bar{y},\bar{y};\boldsymbol{\mu})}{dt}\,dt \\ &\quad + c_S(\boldsymbol{\mu})\|\omega\|^2_{\mathcal{Q}} - \frac{1}{2}\int_0^T \frac{d\,a(\bar{z},\bar{z};\boldsymbol{\mu})}{dt}\,dt \\ &\geq \frac{c_S(\boldsymbol{\mu})}{\alpha}\|z\|^2_{\mathcal{Q}} + \frac{\gamma_a(\boldsymbol{\mu})}{2\alpha}\|y(T)\|^2_Y \\ &\quad + c_S(\boldsymbol{\mu})\|q\|^2_{\mathcal{Q}} + \frac{\gamma_a(\boldsymbol{\mu})}{2}\|z(0)\|^2_Y > 0. \end{aligned}$$

Since the aforementioned relation does not depend on the choice of ω and ζ, we have proved (1.76).

Case 2. We now consider $\Omega_u \neq \Omega_{\text{obs}}$[3], assuming $\Omega_{\text{obs}} \neq \Omega$. Thus, we consider $\bar{y}$ as (1.77) and the indicator function $\chi_{\Omega\setminus\Omega_{\text{obs}}}$, by definition. Using the same arguments exploited in Case 1:

$$\begin{aligned} \underline{\mathcal{B}}((\bar{y}\chi_{\Omega\setminus\Omega_{\text{obs}}},0),(\omega,\zeta);\boldsymbol{\mu}) &= s(\bar{y}\chi_{\Omega\setminus\Omega_{\text{obs}}},\zeta;\boldsymbol{\mu}) \\ &\quad + \int_0^T a(\bar{y}\chi_{\Omega\setminus\Omega_{\text{obs}}},\zeta;\boldsymbol{\mu})\,dt \\ &\geq \|\zeta\chi_{\Omega\setminus\Omega_{\text{obs}}}\|^2_{\mathcal{Q}} + \frac{\gamma_a(\boldsymbol{\mu})}{2}\|y(T)\|^2_Y, \end{aligned} \tag{1.81}$$

[2]This is always the case for the Hilbert spaces considered for the numerical experiment presented in Section 5.2.2.

[3]See Footnote 1.

where $m(\bar{y}\chi_{\Omega\setminus\Omega_{\text{obs}}}, \omega) = 0$ for all $\omega \in \mathcal{Q}$. Since (1.81) does not depend on the choice of the test functions, inequality (1.76) holds. ∎

Lemma 1.6 and Lemma 1.7 guarantees the hypotheses of Nečas-Babuška theorem and, thus, we can now state the following well-posedness result:

Theorem 1.8 *For a given $\boldsymbol{\mu} \in \mathcal{D}$, the problem (1.58) has a unique solution pair $(y, z) \in \mathcal{Y}_0 \times \mathcal{Y}_T$.*

Remark 1.3.1 ($\Omega_u \neq \Omega$, time-dependent OCP($\boldsymbol{\mu}$)s) *In all the proofs we assumed $\Omega_{obs} \neq \Omega$. However, Lemma 1.6 and Lemma 1.7 are still provable assuming $\Omega_u \neq \Omega$. For the sake of completeness, we outline the ideas behind the proofs that are very similar to the previous cases.*

- *Lemma 1.6. One can consider $(c_y y + c_{\zeta_y}\zeta_y + \overline{\kappa}, z + \omega_z) \in \mathcal{Q} \times \mathcal{Q}$, where $\overline{\kappa} \in \mathcal{Y}_T$ is the solution of the following backward parabolic problem: given $\boldsymbol{\mu} \in \mathcal{D}$ and $z \in \mathcal{Q}$*

$$\begin{cases} -\displaystyle\int_0^T s(\overline{\kappa}, r; \boldsymbol{\mu})\, dt \\ \qquad + \displaystyle\int_0^T a(r, \overline{\kappa}; \boldsymbol{\mu})\, dt = \\ \qquad\qquad - \displaystyle\int_0^T m(z, r; \boldsymbol{\mu})\, dt + \frac{c_y}{\alpha}\int_0^T c(r, z; \boldsymbol{\mu})\, dt & \forall r \in \mathcal{Q}, \\ \overline{\kappa}(T) = 0 & \text{in } \Omega, \\ \overline{\kappa} \equiv 0 & \text{in } \Omega_u. \end{cases} \tag{1.82}$$

  The inf-sup condition is still verified with the following $\boldsymbol{\mu}-$dependent constant:

$$\beta(\boldsymbol{\mu}) = \frac{\min\left\{\frac{c_m(\boldsymbol{\mu})^2}{2c_S(\boldsymbol{\mu})\gamma_s(\boldsymbol{\mu})}, \frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})^3\alpha^2}{2c_c(\boldsymbol{\mu})^2 c_u(\boldsymbol{\mu})c_a(\boldsymbol{\mu})^2}, \frac{c_S(\boldsymbol{\mu})\gamma_a(\boldsymbol{\mu})}{2c_a(\boldsymbol{\mu})^2}, \frac{\gamma_a(\boldsymbol{\mu})}{2}\right\}}{\sqrt{2\max\left\{c_y^2, \frac{c_{\zeta_y}^2}{\beta_a(\boldsymbol{\mu})^2}, \frac{1}{\beta_a(\boldsymbol{\mu})^2}, \left(\frac{c_m(\boldsymbol{\mu})c_{obs}}{\gamma_a(\boldsymbol{\mu})}\right)^2\right\}}}, \tag{1.83}$$

  having applied

$$\gamma_a(\boldsymbol{\mu})\|\overline{\kappa}\|_Y \leq c_m(\boldsymbol{\mu})c_{obs}\|z\|_Y. \tag{1.84}$$

  Here, we used (1.62) together with $m(\overline{\kappa}, z; \boldsymbol{\mu})$ as right hand side of (1.74).

- *Lemma 1.7. The inequality (1.76) is verified when the chosen representative is $(0, \bar{z}\chi_{\overline{\Omega}\setminus\Omega_u})$ with $\bar{z}$ as in (1.77).*

1.3.2 Steady OCP($\boldsymbol{\mu}$)s: Problem Formulation

In this Section, we provide a no-control framework analysis for steady problems too. Namely, the variables are $y, z \in Y$, and $u \in U$, while $y_{\mathrm{d}} \in Y_{\mathrm{obs}}$ and no time integration is considered. The whole optimality problem reads: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, u) \in Y \times Y$ such that

$$\underline{\mathcal{B}}_s((y,z),(\omega,\zeta);\boldsymbol{\mu}) = \langle \mathcal{F}_s(\boldsymbol{\mu}),(\omega,\zeta)\rangle \quad \forall (\omega,\zeta) \in Y \times Y. \tag{1.85}$$

where the left hand side is given by $\underline{\mathcal{B}}_s : (Y \times Y) \times (Y \times Y) \to \mathbb{R}$ with

$$\underline{\mathcal{B}}_s((y,z),(\omega,\zeta);\boldsymbol{\mu}) = a(y,\zeta;\boldsymbol{\mu}) - \frac{1}{\alpha}c(z,\zeta;\boldsymbol{\mu}) + m(y,\omega;\boldsymbol{\mu}) + a(\omega,z;\boldsymbol{\mu}),$$

while the right hand side is

$$\langle \mathcal{F}_s(\boldsymbol{\mu}),(\omega,\zeta)\rangle = m(y_d,\omega;\boldsymbol{\mu}) + \langle G(\boldsymbol{\mu}),\zeta\rangle\,.$$

As already did in the time-dependent case, we would like to use the Nečas-Babuška theorem to assure the existence and uniqueness of the optimal solution. Namely, we aim at proving a steady version of Lemma 1.6 and Lemma 1.7 exploiting arguments not so different from the ones used in the time-dependent framework. For the sake of completeness and clarity, we will report them.

Lemma 1.9 (Surjectivity of $\underline{\mathcal{B}}_s{}^*$) *For the bilinear form (1.85), the following inf-sup stability condition holds: there exists $\beta_s(\boldsymbol{\mu}) > 0$ such that $\beta_{\underline{\mathcal{B}}_s}(\boldsymbol{\mu}) > \beta_s(\boldsymbol{\mu}) > 0$ with*

$$\beta_{\underline{\mathcal{B}}_s}(\boldsymbol{\mu}) := \inf_{(y,z)\in(Y\times Y)} \sup_{(\omega,\zeta)\in(Y\times Y)} \frac{\underline{\mathcal{B}}_s((y,z),(\omega,\zeta);\boldsymbol{\mu})}{\sqrt{\|y\|_Y^2 + \|z\|_Y^2}\sqrt{\|\omega\|_Y^2 + \|\zeta\|_Y^2}}, \tag{1.86}$$

for $(y,z) \neq 0$ and $(\omega,\zeta) \neq 0$.

Proof. Case 1. Let us suppose $\Omega_u = \Omega_{\mathrm{obs}}$, as in Lemma 1.6[4]. Choosing $\omega = \alpha y$ and $\zeta = z$ leads to

$$\sup_{(\omega,\zeta)\in(Y\times Y)\setminus\{(0,0)\}} \frac{\underline{\mathcal{B}}_s((y,z),(\omega,\zeta);\boldsymbol{\mu})}{\sqrt{\|y\|_Y^2 + \|z\|_Y^2}\sqrt{\|\omega\|_Y^2 + \|\zeta\|_Y^2}} \geq \alpha\gamma_a(\boldsymbol{\mu}), \tag{1.87}$$

since, when $\Omega_u = \Omega_{\mathrm{obs}}$, the action of $c(z,y;\boldsymbol{\mu})$ and $m(y,z;\boldsymbol{\mu})$ coincide. The inf-sup condition (1.86) holds: indeed, the inequality does not depend on the choice of y and z.

[4] Once again, we postpone the analysis of case $\Omega_u \neq \Omega$ to Remark 1.3.2.

Case 2. Now, let us take into account $\Omega_u \neq \Omega_{\text{obs}}$, where $\Omega_{\text{obs}} \neq \Omega$[5]. Here, we chose $\omega = y$ and $\zeta = z + \overline{\kappa}$, where $\overline{\kappa}$ is the solution to the following equation

$$\begin{cases} a(\overline{\kappa}, r; \boldsymbol{\mu}) = -m(y, r; \boldsymbol{\mu}) + \dfrac{1}{\alpha} c(r, y; \boldsymbol{\mu}) & \forall r \in L^2(\Omega), \\ \overline{\kappa} \equiv 0 & \text{in } \Omega_{\text{obs}}. \end{cases} \tag{1.88}$$

Thanks to these assumptions, we obtain $\|\overline{\kappa}\|_Y \leq \dfrac{c_c(\boldsymbol{\mu}) c_u}{\alpha \gamma_a(\boldsymbol{\mu})} \|y\|_Y$ exploiting the same arguments already presented in Lemma 1.6. Thus, (1.86) holds with

$$\beta_s(\boldsymbol{\mu}) = \frac{\gamma_a(\boldsymbol{\mu})}{\sqrt{\left(2 \max\left\{1, \left(\frac{c_c(\boldsymbol{\mu}) c_u}{\alpha \gamma_a(\boldsymbol{\mu})}\right)^2\right\}\right)}} > 0.$$

■

It only remains to prove the steady version of (1.76).

Lemma 1.10 (Injectivity of $\underline{\mathcal{B}_s}^*$) *The bilinear form (1.59) satisfies the following inf-sup stability condition:*

$$\inf_{(\omega,\zeta)\in(Y\times Y)} \sup_{(y,z)\in(Y\times Y)} \frac{\underline{\mathcal{B}_s}((y,z),(\omega,\zeta);\boldsymbol{\mu})}{\sqrt{\|y\|_Y^2 + \|z\|_Y^2}\sqrt{\|\omega\|_Y^2 + \|\zeta\|_Y^2}} > 0, \tag{1.89}$$

for $(\omega, \zeta) \neq 0$ and $(y, z) \neq 0$.

Proof. The proof, as usual, is divided in two cases.
Case 1. Let us assume $\Omega_u = \Omega_{\text{obs}}$. The same arguments of Case 1 of Lemma 1.9 can be applied choosing $y = \alpha\omega$ and $z = \zeta$, obtaining (1.87) also in the steady case, where the supremum is considered in the space $Y \times Y$, which already proves (1.89).
Case 2. Now, we consider $\Omega_u \neq \Omega_{\text{obs}}$ and, once again, $\Omega_{\text{obs}} \neq \Omega$, without loss of generality[6]. To prove the inequality (1.89), we take $z = 0$ and $y = \omega\chi_{\Omega\setminus\Omega_{\text{obs}}}$ to obtain

$$\sup_{(y,z)\in(Y\times Y)\setminus\{(0,0)\}} \underline{\mathcal{B}_s}((y,z),(\omega,\zeta);\boldsymbol{\mu}) \geq \gamma_a(\boldsymbol{\mu}) \|\omega\chi_{\Omega\setminus\Omega_{\text{obs}}}\|_Y^2,$$

proving the estimate. ■

Moreover, the continuity of $\mathcal{B}_s^{\text{ocp}}(\cdot,\cdot;\boldsymbol{\mu})$ is directly inherited from the continuity of the various bilinear forms of the equation. Now, exploiting the continuity of $\mathcal{B}_s^{\text{ocp}}(\cdot,\cdot,\boldsymbol{\mu})$ and $\mathcal{F}_s(\boldsymbol{\mu})$ combined with Lemma 1.9 and Lemma 1.10, the Nečas-Babuška theory is verified and the following theorem holds.

[5] See Footnote 4.
[6] See Footnote 4.

Theorem 1.11 *For a given $\boldsymbol{\mu} \in \mathcal{D}$, problem (1.85) has a unique solution $(y, z) \in Y \times Y$.*

Remark 1.3.2 ($\Omega_u \neq \Omega$, steady OCP($\boldsymbol{\mu}$)s) *Also for the steady case, we considered $\Omega_{obs} \neq \Omega$ guided by the results in Section 5.2.2. We can prove Lemma 1.9 and Lemma 1.10 also assuming $\Omega_u \neq \Omega$ in a steady framework.*

- *Lemma 1.9. Let us consider $\omega = y + \overline{\kappa}$ and $\zeta = z$, where $\overline{\kappa}$ is the solution of*

$$\begin{cases} a(r, \overline{\kappa}; \boldsymbol{\mu}) = -m(r, z; \boldsymbol{\mu}) + \dfrac{1}{\alpha} c(z, r; \boldsymbol{\mu}) & \forall r \in L^2(\Omega), \\ \overline{\kappa} \equiv 0 & \text{in } \Omega_u, \end{cases}$$

  for a given $\boldsymbol{\mu} \in \mathcal{D}$ and $z \in Y$. Exploiting the continuity and the coercivity properties of the considered bilinear form, we obtain $\|\overline{\kappa}\|_Y \leq \dfrac{c_m(\boldsymbol{\mu}) c_{obs}}{\gamma_a(\boldsymbol{\mu})} \|y\|_Y$, using the strategies of Lemma 1.9. Thus, the relation (1.86) holds with

$$\beta_s(\boldsymbol{\mu}) = \frac{\gamma_a(\boldsymbol{\mu})}{\sqrt{\left(2 \max\left\{1, \left(\frac{c_m(\boldsymbol{\mu}) c_{obs}}{\alpha \gamma_a(\boldsymbol{\mu})}\right)^2\right\}\right)}} > 0.$$

- *Lemma 1.10. To prove (1.89) we simply consider $y = 0$ and $z = \zeta \chi_{\overline{\Omega} \setminus \Omega_u}$.*

In this context, the analysis had the only purpose to prove the well-posedness of the problems we will deal with along the Thesis, in order to apply a numerical discretization to simulate the optimality system for a given parameter $\boldsymbol{\mu} \in \mathcal{D}$. The next Chapter describes the concept of *high fidelity* approximation of OCP($\boldsymbol{\mu}$)s based on space-time formulations, see e.g. [53, 65, 67, 66, 156, 165, 166] and its adaptation to the problems we introduced along this Chapter.

Chapter 2

Space–time Approximation for Parametric Optimal Control Problems

In this Chapter, we will introduce space-time approximation for OCP($\boldsymbol{\mu}$)s. This discretization technique have been successfully applied to parabolic PDEs($\boldsymbol{\mu}$) in [53, 89, 156, 165, 166] and also to linear constrained optimization problems in a non-parametric setting, see e.g. [65, 66, 143, 144]. We will focus on the versatility of such an approach, that highlights the algebraic saddle point structure of discretized OCP($\boldsymbol{\mu}$)s and easily adapts to several settings, as presented in [16, 146, 150]. Along the Chapter, we might refer to the space-time approximation as the *high fidelity approximation*: the reason of this choice is postponed in Chapter 3. The structure is discussed for several governing equations. For all of them, we exploited a *optimize-then-discretize* approach: namely, the discretization is applied only once we have built the optimality system (1.10). The interested reader may refer to [49, 59] for an overview on the topic and on the alternative approaches. A brief outline of the Chapter follows. First of all, Section 2.1 will describe Galerkin approach in a space-time fashion and its well-posedness. In Section 2.2 we will focus on the algebraic structure of nonlinear time-dependent OCP($\boldsymbol{\mu}$)s. The construction will be adapted to linear problems in Section 2.3. For both these last two Sections, we will briefly describe the steady versions of the algebraic systems at hand.

2.1 The High Fidelity Approximation

This Section focuses on the numerical approximation of OCP($\boldsymbol{\mu}$)s. First, we will briefly introduce the Galerkin projection strategy and thus we will discuss the well-posedness of the OCP($\boldsymbol{\mu}$)s presented in Chapter 1 but in a finite dimensional setting. In the ROMs community, the high fidelity approximation represents a first stage of discretization, characterized by high accuracy[1]. We will employ Galerkin projection based on space-time techniques [53, 89, 156, 165, 166] to discretize the problem at hand. The detailed description of such an approximation will be postponed in the next Section. For now, let us assume to have built space-time function spaces $\mathcal{Y}^{\mathcal{N}_y}$ and $\mathcal{U}^{\mathcal{N}_u}$, finite dimensional subsets of

[1]The high fidelity discretization might lead to unbearable simulations in terms of computational time, most of all in an OCP($\boldsymbol{\mu}$)s setting, where many parameters evaluations are needed for the system of three equation (1.10). However, this issue can be tackled following the reduced techniques presented in Chapter 3.

$\mathcal{Y}$ and $\mathcal{U}$, respectively. Namely, we now want to solve the variational formulation of (1.13) exploiting a Galerkin projection into the finite dimensional space $\mathbb{X}^{\mathcal{N}} := \mathcal{Y}^{\mathcal{N}_y} \times \mathcal{U}^{\mathcal{N}_u} \times \mathcal{Y}^{\mathcal{N}_y}$. Thus, the high fidelity space-time problem reads: for a given $\boldsymbol{\mu} \in \mathcal{D}$, find the solution $\boldsymbol{X}^{\mathcal{N}} \in \mathbb{X}^{\mathcal{N}}$ such that

$$G(\boldsymbol{X}^{\mathcal{N}}, \boldsymbol{\Xi}; \boldsymbol{\mu}) = \langle \mathcal{F}, \boldsymbol{\Xi} \rangle_{\mathbb{X}^{\mathcal{N}*}, \mathbb{X}^{\mathcal{N}}} \quad \forall \boldsymbol{\Xi} \in \mathbb{X}^{\mathcal{N}}. \tag{2.1}$$

In this approximation context, we are assuming that the the test space and the solution function space coincide, which is not a restrictive hypothesis for the problems we will deal with, as we will see later on in this Chapter and in the numerical results we are going to present in this Thesis. Let us recall that $\mathcal{Y}_0 \subset \mathcal{Q}$. Thus, for general nonlinear problems, the existence is guaranteed by the fulfillment of (i)-(ix), also at the discrete level.
Now, focusing on the linear cases, we will, once again, specify the space-time saddle point structure. We first define the combined function space for state and control as $\mathcal{X}^{\mathcal{N}_x} := \mathcal{Y}^{\mathcal{N}_y} \times \mathcal{U}^{\mathcal{N}_u}$, with $\mathcal{N}_x = \mathcal{N}_y + \mathcal{N}_u$. The finite dimensional optimality system reads: for a given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(x^{\mathcal{N}_x}, z^{\mathcal{N}_y}) \in \mathcal{X}^{\mathcal{N}_x} \times \mathcal{Y}^{\mathcal{N}_y}$ such that:

$$\begin{cases} \mathcal{A}(x^{\mathcal{N}_x}, \xi; \boldsymbol{\mu}) + \mathcal{B}(\xi, z^{\mathcal{N}_y}; \boldsymbol{\mu}) = \mathcal{H}(\xi) & \forall \xi \in \mathcal{X}^{\mathcal{N}_x}, \\ \mathcal{B}(x^{\mathcal{N}_x}, \zeta; \boldsymbol{\mu}) = \displaystyle\int_0^T \langle F(\boldsymbol{\mu}), \zeta \rangle \, dt & \forall \zeta \in \mathcal{Y}^{\mathcal{N}_y}, \end{cases} \tag{2.2}$$

where the involved bilinear forms have been already defined both for OCP($\boldsymbol{\mu}$)s governed by parabolic and Stokes equations, in Sections 1.2.2 and 1.2.3, respectively. The well-posedness is assured by Theorem 2 in both cases, assuming the state space and the adjoint space coinciding, see e.g. [146]. The same argument holds for steady problems. Indeed, let us assume to have build a discretization (only in space) for Y and U. The resulting function spaces will be denoted by $Y^{N^y_{FE}}$ and $U^{N^u_{FE}}$. Furthermore, we define the state-control space $X^{N^x_{FE}}$ as $Y^{N^y_{FE}} \times U^{N^u_{FE}}$. Thus, in the steady context, discrete OCP($\boldsymbol{\mu}$)s read: for a given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(x^{N^x_{FE}}, z^{N^y_{FE}}) \in X^{N^x_{FE}} \times Y^{N^y_{FE}}$ such that:

$$\begin{cases} \mathcal{A}(x^{N^x_{FE}}, \xi; \boldsymbol{\mu}) + \mathcal{B}(\xi, z^{N^y_{FE}}; \boldsymbol{\mu}) = \langle H(\boldsymbol{\mu}), \xi \rangle & \forall \xi \in X^{N^x_{FE}}, \\ \mathcal{B}(x^{N^x_{FE}}, \zeta; \boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), \zeta \rangle & \forall \zeta \in Y^{N^y_{FE}}. \end{cases} \tag{2.3}$$

Also in this case, the Brezzi Theorem is verified using the same techniques of the continuous problem, assuming both the state and the adjoint variables in $Y^{N^y_{FE}}$, see e.g. [110, 111].
The last case we want to address is the no-control framework. In this specific setting, the proof of the well-posedness of such a problem is more involved and we would like to report the main results proposed in [148] on this topic. Nevertheless, the Theorems and the proofs we are going to present will be of utmost importance for Chapter 5.

2.1.1 Well-Posedness of the No-Control Problem

This Section deals with the well-posedness of the space-time discretize version of the optimality system (1.58), i.e. the system describing OCP($\boldsymbol{\mu}$)s governed by linear time-dependent equations in a no-control framework. Here, we want to solve a discretized version of (1.58), thus, we have to define high fidelity space-time function spaces. First of all, we take into consideration the FE function space $Y^{N_{\text{FE}}^y} = Y \cap \mathbb{K}_1$ where

$$\mathbb{K}_1 = \{v \in C^0(\overline{\Omega}) \ : \ v|_K \in \mathbb{P}^1, \ \ \forall K \in \mathcal{T}\},$$

with $\mathbb{P}^1$ set of all the polynomials of degree at most equal to 1 and with K an element of a triangulation $\mathcal{T}$ of the spatial domain Ω. Thanks to this definition, the semi-discrete function spaces

$$\mathcal{Y}^{N_{\text{FE}}^y} = \left\{ y \in L^2(0,T;Y^{N_{\text{FE}}^y}) \ \text{ s.t. } \ y_t \in L^2(0,T;(Y^{N_{\text{FE}}^y})^*) \right\},$$

and $\mathcal{Q}^{N_{\text{FE}}^y} = L^2(0,T;Y^{N_{\text{FE}}^y})$ can be defined, with N_{FE}^y representing the FE dimension for the state and the adjoint variables. After addressing the space discretization, we can deal with time approximation in order to obtain the space-time function spaces, say $\mathcal{Y}_{N_t}^{N_{\text{FE}}^y}$ and $\mathcal{Q}_{N_t}^{N_{\text{FE}}^y}$ where N_t is the dimension of the time discretization. For this specific problem formulation, we assume that the same space-time discretization technique might be employed both for state and adjoint, namely $\mathcal{Y}_{N_t}^{N_{\text{FE}}^y} \equiv \mathcal{Q}_{N_t}^{N_{\text{FE}}^y}$ [89, 148]. For the sake of notation, in the no-control formulation, we will refer to this common space as $\mathcal{Q}^{\mathcal{N}_y}$. The high fidelity optimality system is: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y^{\mathcal{N}}, \zeta^{\mathcal{N}}) \in \mathcal{Q}^{\mathcal{N}_y} \times \mathcal{Q}^{\mathcal{N}_y}$ such that

$$\underline{\mathcal{B}}((y^{\mathcal{N}}, z^{\mathcal{N}}), (\omega, \zeta); \boldsymbol{\mu}) = \left\langle \mathcal{F}(\boldsymbol{\mu}), (\omega, \zeta) \right\rangle \quad \forall (\omega, \zeta) \in \mathcal{Q}^{\mathcal{N}_y} \times \mathcal{Q}^{\mathcal{N}_y}. \tag{2.4}$$

We remark that the global dimension is $\mathcal{N} = 2\mathcal{N}_y = 2N_{FE}^y \cdot N_t$. In order to prove the well-posedness of this space-time optimality system, we exploit the Nečas-Babuška theorem, as in the continuous case. Namely, we require the discrete inf-sup stability conditions in the discretized space $\mathcal{Q}^{\mathcal{N}_y}$ to be verified together with the continuity of (1.58). The latter property is directly inherited from the continuous system. Thus, we only need to provide the following Lemma that proves the inf-sup stability condition w.r.t. the discretized spaces.

Lemma 2.1 (Discrete Surjectivity of $\underline{\mathcal{B}}^*$) *There exists $\beta^{\mathcal{N}}(\boldsymbol{\mu}) > 0$ such that $\beta_{\underline{\mathcal{B}}}^{\mathcal{N}}(\boldsymbol{\mu}) \geq \beta^{\mathcal{N}}(\boldsymbol{\mu})$, where $\beta_{\underline{\mathcal{B}}}^{\mathcal{N}}(\boldsymbol{\mu})$ is*

$$\inf_{(y,z)\in(\mathcal{Q}^{\mathcal{N}_y}\times\mathcal{Q}^{\mathcal{N}_y})} \sup_{(\omega,\zeta)\in(\mathcal{Q}^{\mathcal{N}_y}\times\mathcal{Q}^{\mathcal{N}_y})} \frac{\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu})}{\sqrt{\|y\|_{\mathcal{Q}}^2 + \|z\|_{\mathcal{Q}}^2}\sqrt{\|\omega\|_{\mathcal{Q}}^2 + \|\zeta\|_{\mathcal{Q}}^2}}, \tag{2.5}$$

for $(y,z) \neq 0$ and $(\omega,\zeta) \neq 0$.

PROOF. Case 1. Let Ω_u and Ω_{obs} coincide. Now, choosing $(\omega, \zeta) = (\alpha y, z)$, applying $c(z, y; \boldsymbol{\mu}) = m(y, z; \boldsymbol{\mu})$ as we already did in the continuous case, (1.66) and the coercivity of the state equation, we obtain, for $(\omega, \zeta) \neq 0$,

$$\begin{aligned}
\sup_{(\omega,\zeta)\in(\mathcal{Q}^{\mathcal{N}_y}\times\mathcal{Q}^{\mathcal{N}_y})} & \frac{\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu})}{\sqrt{\|y\|^2_{\mathcal{Q}}+\|z\|^2_{\mathcal{Q}}}\sqrt{\|\omega\|^2_{\mathcal{Q}}+\|\zeta\|^2_{\mathcal{Q}}}} \\
&\geq \frac{\underline{\mathcal{B}}((y,z),(\alpha y,z);\boldsymbol{\mu})}{\sqrt{\|y\|^2_{\mathcal{Q}}+\|z\|^2_{\mathcal{Q}}}\sqrt{\|\alpha y\|^2_{\mathcal{Q}}+\|\zeta\|^2_{\mathcal{Q}}}} \\
&= \frac{\alpha c_S(\boldsymbol{\mu})\|y(T)\|^2_H+\alpha\gamma_a(\boldsymbol{\mu})\|y\|^2_{\mathcal{Q}}+c_S(\boldsymbol{\mu})\|z(0)\|^2_H+\gamma(\boldsymbol{\mu})\|z\|^2_{\mathcal{Q}}}{\|y\|^2_{\mathcal{Q}}+\|z\|^2_{\mathcal{Q}}} \\
&\geq \frac{\min\{\alpha\gamma_a(\boldsymbol{\mu}),\gamma_a(\boldsymbol{\mu})\}(\|y\|^2_{\mathcal{Q}}+\|z\|^2_{\mathcal{Q}})}{\|y\|^2_{\mathcal{Q}}+\|z\|^2_{\mathcal{Q}}} \\
&\geq \alpha\gamma_a(\boldsymbol{\mu}).
\end{aligned}$$

Case 2. We now focus on the case $\Omega_u \neq \Omega_{\text{obs}}$. Furthermore, without loss of generality, we can consider at least one between the two sets different from the whole spatial domain, say Ω_{obs}[2]. In this case we define $\kappa \in \mathcal{Q}^{\mathcal{N}_y}$, solution of the auxiliary problem (1.73) with $c_y = 1$. Choosing $(\omega, \zeta) = (y, z + \kappa)$ and recalling that, for each pair of elements in $\mathcal{Q}^{\mathcal{N}_y}$, holds

$$\begin{aligned}
-\int_0^T s(\kappa, z; \boldsymbol{\mu})\, dt &= \int_\Omega \kappa(0)z(0)\, d\Omega - \int_\Omega \kappa(T)z(T)\, d\Omega + \int_0^T s(\kappa, z; \boldsymbol{\mu})\, dt \\
&= \int_0^T s(\kappa, z; \boldsymbol{\mu})\, dt,
\end{aligned}$$

[2]The arguments of Footnote 1 and Remark 1.3.1 hold also in this space-time setting.

thus, for $(\omega, \zeta) \neq 0$:

$$\begin{aligned}
\sup_{(\omega,\zeta)\in(\mathcal{Q}^{\mathcal{N}_y}\times\mathcal{Q}^{\mathcal{N}_y})} & \frac{\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu})}{\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2}\sqrt{\|\omega\|_{\mathcal{Q}}^2+\|\zeta\|_{\mathcal{Q}}^2}} \\
&\geq \frac{\underline{\mathcal{B}}((y,z),(y,z+\kappa);\boldsymbol{\mu})}{\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2}\sqrt{\|y\|_{\mathcal{Q}}^2+2(\|z\|_{\mathcal{Q}}^2+\|\kappa\|_{\mathcal{Q}}^2)}} \\
&= \frac{c_S(\boldsymbol{\mu})\|y(T)\|_H^2+\gamma_a(\boldsymbol{\mu})\|y\|_{\mathcal{Q}}^2+c_S(\boldsymbol{\mu})\|z(0)\|_H^2+\gamma(\boldsymbol{\mu})\|z\|_{\mathcal{Q}}^2}{\sqrt{2}\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2}\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2+\|\kappa\|_{\mathcal{Q}}^2}} \\
&\quad + \frac{\displaystyle\int_0^T m(y,z;\boldsymbol{\mu})\,dt-\frac{1}{\alpha}\int_0^T c(z,y;\boldsymbol{\mu})\,dt}{\sqrt{2}\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2}\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2+\|\kappa\|_{\mathcal{Q}}^2}} \\
&\quad + \frac{\displaystyle\int_0^T m(y,\kappa;\boldsymbol{\mu})\,dt+\int_0^T s(\kappa,z;\boldsymbol{\mu})\,dt+\int_0^T a(\kappa,z;\boldsymbol{\mu})\,dt}{\sqrt{2}\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2}\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2+\|\kappa\|_{\mathcal{Q}}^2}} \\
&\geq \frac{\gamma_a(\boldsymbol{\mu})(\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2)}{\sqrt{2}\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2}\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2+\|\kappa\|_{\mathcal{Q}}^2}}.
\end{aligned}$$

Furthermore, we are going to prove the existence of a constant $\bar{c} > 0$ such that:

$$\|\kappa\|_{\mathcal{Q}} \leq \bar{c}\|y\|_{\mathcal{Q}}.$$

This goal can be reached through relation (1.74), with $c_y = 1$, which reads $\bar{c} = \dfrac{c_c(\boldsymbol{\mu})c_u}{\alpha\gamma_a(\boldsymbol{\mu})}$. Thus, for $(\omega, \zeta) \neq 0$, this leads us to the following estimate:

$$\sup_{(\omega,\zeta)\in(\mathcal{Q}^{\mathcal{N}_y}\times\mathcal{Q}^{\mathcal{N}_y})} \frac{\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu})}{\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2}\sqrt{\|\omega\|_{\mathcal{Q}}^2+\|\zeta\|_{\mathcal{Q}}^2}} \geq \frac{\gamma_a(\boldsymbol{\mu})}{\sqrt{2(\max\{1,\bar{c}^2\})}}.$$

The inequality (2.1.1) does not depend on the pair $(y, z) \in \mathcal{Q}^{\mathcal{N}_y} \times \mathcal{Q}^{\mathcal{N}_y}$, thus the inf-sup condition (2.5) holds true both when $\Omega_u = \Omega_{\text{obs}}$ and otherwise.

■

We stress that in the finite dimensional case there is no need to prove a discrete equivalent of (1.76), indeed, see e.g. [11, 163], since (1.76) coincides with:

$$\inf_{(\omega,\zeta)\in(\mathcal{Q}^{\mathcal{N}_y}\times\mathcal{Q}^{\mathcal{N}_y})}\sup_{(y,z)\in(\mathcal{Q}^{\mathcal{N}_y}\times\mathcal{Q}^{\mathcal{N}_y})} \frac{\underline{\mathcal{B}}((y,z),(\omega,\zeta);\boldsymbol{\mu})}{\sqrt{\|y\|_{\mathcal{Q}}^2+\|z\|_{\mathcal{Q}}^2}\sqrt{\|\omega\|_{\mathcal{Q}}^2+\|\zeta\|_{\mathcal{Q}}^2}},$$

for $(\omega, \zeta) \neq 0$ and $(y, z) \neq 0$. Thanks to the continuity of (2.4) and Lemma 2.1 we can state the well-posedness of the discretized problem (2.4) in a space-time fashion since the Nečas-Babuška theory applies, proving the following theorem:

Theorem 2.2 *For a given $\boldsymbol{\mu} \in \mathcal{D}$, the problem (2.4) has a unique solution $(y, z) \in \mathcal{Q}^{\mathcal{N}_y} \times \mathcal{Q}^{\mathcal{N}_y}$.*

Remark 2.1.1 (Steady Case) *We remark that the well-posedness of the no-control framework in the steady setting can be proved adapting the strategies we just presented for time-dependent equations. First of all, here we are dealing with a pure FE space of dimension $2N_{FE}^y$. Steady OCP($\boldsymbol{\mu}$)s in no-control framework reads: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, z) \in Y^{N_{FE}^y} \times Y^{N_{FE}^y}$ such that*

$$\underline{\mathcal{B}}_s((y, z), (\omega, \zeta)) = \langle \mathcal{F}_s(\boldsymbol{\mu}), (\omega, \zeta) \rangle \quad \forall (\omega, \zeta) \in Y^{N_{FE}^y} \times Y^{N_{FE}^y}. \tag{2.6}$$

The following holds:

Theorem 2.3 *For a given $\boldsymbol{\mu} \in \mathcal{D}$ and for a given observation $y_d \in Y_{obs}^{N_{FE}^y}$, problem (2.6) has a unique solution $(y, z) \in Y^{N_{FE}^y} \times Y^{N_{FE}^y}$.*

Proof. Also in this case, the well-posedness is an application of the Nečas-Babuška theorem. First of all, the continuity directly derives from the continuous forms. Furthermore, Lemma 1.9 can be easily adapted in this setting defining $\beta_{\underline{\mathcal{B}}_s}^{N_{FE}^y}(\boldsymbol{\mu})$ as

$$\inf_{(y,z) \in (Y^{N_{FE}^y} \times Y^{N_{FE}^y})} \sup_{(\omega,\zeta) \in (Y^{N_{FE}^y} \times Y^{N_{FE}^y})} \frac{\underline{\mathcal{B}}_s((y, z), (\omega, \zeta); \boldsymbol{\mu})}{\sqrt{\|y\|_Y^2 + \|z\|_Y^2}\sqrt{\|\omega\|_Y^2 + \|\zeta\|_Y^2}},$$

and, thus, proving that $\beta_{\underline{\mathcal{B}}_s}^{N_{FE}^y} \geq \beta_s^{N_{FE}^y}(\boldsymbol{\mu}) > 0$ for $(y, z) \neq 0$ and $(\omega, \zeta) \neq 0$.

■

Furthermore, we would like to point out that Remark 1.3.2 is still valid in the space-time framework. The next Section focuses on the algebraic structure of the OCP($\boldsymbol{\mu}$)s that we have introduced along this contribution, moving from time-dependent nonlinear governing equations to linear steady problems.

2.2 Space–time nonlinear OCP($\boldsymbol{\mu}$)s

We are now interested in the numerical simulation of nonlinear systems of the

type of (1.13), to study several parametric features varying $\boldsymbol{\mu} \in \mathcal{D}$ in a space-time fashion. Here, we follow the structure of [16, 66, 147]. First of all, we will introduce a general algebraic framework that is suited to time-dependent nonlinear problems with quadratic nonlinearity w.r.t. the state variable[3]. Then, we will drop the time-dependency and we will simplify the setting to deal with steady nonlinear OCP($\boldsymbol{\mu}$)s as already presented in Remark 1.2.1.

2.2.1 Nonlinear Time-Dependent OCP($\boldsymbol{\mu}$)s

Our goal is to find a numerical approximation for the optimal solution $\boldsymbol{X}$ of (1.13). We briefly introduced the discretization technique in the previous Section, however here we use a more general formulation. First, for the sake of clarity, we will focus on the space discretization. Only subsequently we will describe how we treated the time approximation. As already specified in the no-control framework of Section 2.1.1, the applications presented in this contribution are FE-based. Thus, a brief introduction of FE based on general polynomial degree follows. Let $Y^{N^y_{\text{FE}}} = Y \cap \mathbb{K}_{r_y}$ and $U^{N^u_{\text{FE}}} = U \cap \mathbb{K}_{r_u}$ be the FE function spaces. Here, we define

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

where $\mathbb{P}^r$ is the space of all the polynomials of degree at most equal to r and K is an element of a triangulation $\mathcal{T}$ of the spatial domain Ω. Once employed this first stage of discretization, we can define the semi-discrete function spaces

$$\mathcal{Y}^{N^y_{\text{FE}}} = \left\{ y \in L^2(0,T;Y^{N^y_{\text{FE}}}) \ \text{ s.t. } \ y_t \in L^2(0,T;(Y^{N^y_{\text{FE}}})^*) \right\},$$

$\mathcal{Q}^{N^y_{\text{FE}}} = L^2(0,T;Y^{N^y_{\text{FE}}})$ and $\mathcal{U}^{N^u_{\text{FE}}} = L^2(0,T;U^{N^u_{\text{FE}}})$, where N^y_{FE} represents the FE dimension for the state and the adjoint variables, while N^u_{FE} is the analogous dimension of the control space. Once the space approximation is performed, we can deal with time in order to achieve the final space-time function spaces, say $\mathcal{Y}^{N^y_{\text{FE}}}_{N_t}$, $\mathcal{Q}^{N^y_{\text{FE}}}_{N_t}$ and $\mathcal{U}^{N^u_{\text{FE}}}_{N_t}$, where N_t is the number of the considered timesteps taken in the time interval $[0,T]$, i.e. the time dimension of the discrete space. Also in this case, as already did in [89, 148], we use the same discrete strategy to represent $\mathcal{Y}_0, \mathcal{Y}_T$ and $\mathcal{Q}$, namely, at the discrete level $\mathcal{Y}^{N^y_{\text{FE}}}_{N_t} \equiv \mathcal{Q}^{N^y_{\text{FE}}}_{N_t}$. The space-time function spaces will be indicated as $\mathcal{Y}^{\mathcal{N}_y} \equiv \mathcal{Q}^{\mathcal{N}_y}$ and $\mathcal{U}^{\mathcal{N}_u}$, for the state/the adjoint and the control variables, respectively. Consequently, $\mathbb{X}^{\mathcal{N}} := \mathcal{Q}^{\mathcal{N}_y} \times \mathcal{U}^{\mathcal{N}_u} \times \mathcal{Q}^{\mathcal{N}_y}$, where $\mathcal{N} = 2\mathcal{N}_y + \mathcal{N}_u$, is the global dimension of the space-time optimal variable with $\mathcal{N}_y = N^y_{\text{FE}} \cdot N_t$ and $\mathcal{N}_u = N^u_{\text{FE}} \cdot N_t$. The discrete problem in this framework translates as: given $\boldsymbol{\mu} \in \mathcal{D}$, find $\boldsymbol{X}^{\mathcal{N}} := \boldsymbol{X}^{\mathcal{N}}(\boldsymbol{\mu}) \in \mathcal{X}^{\mathcal{N}}$ such that

$$\mathcal{G}(\boldsymbol{X}^{\mathcal{N}}; \boldsymbol{\mu}) = \mathcal{F}. \tag{2.7}$$

[3]The quadratic nonlinearity assumption is guided by the numerical results we are going to present in the following Chapters. However, the construction can be adapted to more complicated nonlinear problems.

In this contribution, the time discretizations we used are Euler-based. However, this approach is consistent with the space-time formulation presented in many works in literature, see e.g. [156, 165, 166]: indeed, the backward Euler scheme in time, for example, coincides with a piecewise constants Discontinuous Galerkin discretization [48]. Although, for the sake of clarity, we will always talk about Euler's schemes. In this framework, we divide the time interval $[0, T]$ in N_t equispaced subintervals of length Δt. The generic time instance is $t_k = k\Delta t$ for $k = 0, \dots, N_t$. We can consider the variables $y_k^{\mathcal{N}}$, $u_k^{\mathcal{N}}$ and $z_k^{\mathcal{N}}$, i.e. the state, the adjoint and the control variables evaluated at t_k, respectively. They are represented by the FE basis $\{\phi^i\}_{i=1}^{N_{\text{FE}}^y}$ and $\{\psi^i\}_{i=1}^{N_{\text{FE}}^u}$ as follows

$$y_k^{\mathcal{N}} = \sum_1^{N_{\text{FE}}^y} y_k^i \phi^i, \qquad u_k^{\mathcal{N}} = \sum_1^{N_{\text{FE}}^u} u_k^i \psi^i \qquad \text{and} \qquad z_k^{\mathcal{N}} = \sum_1^{N_{\text{FE}}^y} z_k^i \phi^i. \tag{2.8}$$

Thanks to the expansions (2.8), we can consider the space-time state, control and adjoint vectors given by

$$\mathsf{y} = \begin{bmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_{N_t} \end{bmatrix}, \quad \mathsf{u} = \begin{bmatrix} \bar{u}_1 \\ \vdots \\ \bar{u}_{N_t} \end{bmatrix}, \quad \text{and} \quad \mathsf{z} = \begin{bmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_{N_t} \end{bmatrix}.$$

Here, $\bar{y}_k$, $\bar{u}_k$ and $\bar{z}_k$ are the column vectors consisting in the FE coefficients of the variables at the time instance t_k for $k = 1, \dots, N_t$. An analogous construction can be performed for the state initial condition, the desired state and the forcing term, defining

$$\mathsf{y}_0 = \begin{bmatrix} \bar{y}_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathsf{y}_d = \begin{bmatrix} \bar{y}_{d1} \\ \bar{y}_{d2} \\ \vdots \\ \bar{y}_{dN_t} \end{bmatrix}, \quad \text{and} \quad \mathsf{f} = \begin{bmatrix} \bar{f}_1 \\ \bar{f}_2 \\ \vdots \\ \bar{f}_{N_t} \end{bmatrix},$$

respectively. A separate discretization analysis of the three equations of the optimality system (1.10) follows.

The State Equation. Let us focus on a single time instance t_k. Employing the FE method to the controlled state equation (1.5), one derives the following matrix expression for the operators

$$\mathsf{E}_{n\ell}(\boldsymbol{\mu}) + \mathsf{E}_{\ell}(\boldsymbol{\mu}) - \mathsf{C}, \tag{2.9}$$

where, for now, we are omitting the time action. Furthermore, for the sake of notation, we are also omitting the y_k-dependence of $\mathsf{E}_{n\ell}(\boldsymbol{\mu})$. Calling with M_y and M_u the mass matrices for state (or adjoint) and control spaces, respectively,

it is clear that, at each timestep, exploiting backward Euler in order to simulate the evolution of the state equation, the following system has to be solved:

$$\mathsf{M}_y\bar{y}_k + \Delta t \underbrace{(\mathsf{E}_{n\ell}(\boldsymbol{\mu}) + \mathsf{E}_{\ell}(\boldsymbol{\mu}))}_{\mathsf{E}(\boldsymbol{\mu})}\bar{y}_k - \Delta t\mathsf{C}\bar{u}_k = \mathsf{M}_y\bar{y}_{k-1} + \Delta t\bar{f}_k, \tag{2.10}$$

for $k \in \{1, \dots, N_t\}$. Thus, the whole space-time system reads

$$\underbrace{\begin{bmatrix} \mathsf{M}_y + \Delta t\mathsf{E}(\boldsymbol{\mu}) & & & & & \\ -\mathsf{M}_y & \mathsf{M}_y + \Delta t\mathsf{E}(\boldsymbol{\mu}) & & & & \\ & -\mathsf{M}_y & \mathsf{M}_y + \Delta t\mathsf{E}(\boldsymbol{\mu}) & & & \\ & & \ddots & \ddots & & \\ & & & -\mathsf{M}_y & \mathsf{M}_y + \Delta t\mathsf{E}(\boldsymbol{\mu}) \end{bmatrix}}_{\mathsf{K}(\boldsymbol{\mu})} \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_3 \\ \vdots \\ \bar{y}_{N_t} \end{bmatrix}$$

$$-\Delta t \begin{bmatrix} \mathsf{C} & & & & \\ & \mathsf{C} & & & \\ & & \mathsf{C} & & \\ & & & \ddots & \\ & & & & \mathsf{C} \end{bmatrix} \begin{bmatrix} \bar{u}_1 \\ \bar{u}_2 \\ \bar{u}_3 \\ \vdots \\ \bar{u}_{N_t} \end{bmatrix} = \begin{bmatrix} \mathsf{M}_y\bar{y}_0 + \Delta t\bar{f}_1 \\ 0 + \Delta t\bar{f}_2 \\ 0 + \Delta t\bar{f}_3 \\ \vdots \\ 0 + \Delta t\bar{f}_{N_t} \end{bmatrix}.$$

The space-time state equation can be written in compact form as

$$\mathsf{K}(\boldsymbol{\mu})\mathsf{y} - \Delta t\mathsf{C}_{\mathrm{st}}\mathsf{u} = \mathsf{M}_{\mathsf{y}_{\mathrm{st}}}\mathsf{y}_0 + \Delta t\mathsf{f}, \tag{2.11}$$

where the subscript "st" indicates the all-at-once space-time matrices, namely C_{st} is the block-diagonal matrix which entries are given by C of dimension $\mathbb{R}^{\mathcal{N}_y} \times \mathbb{R}^{\mathcal{N}_u}$ and $\mathsf{M}_{\mathsf{y}_{\mathrm{st}}}$ is the block-diagonal matrix consisting of M_y on the diagonal, of dimension $\mathbb{R}^{\mathcal{N}_y} \times \mathbb{R}^{\mathcal{N}_y}$.

The Optimality Equation. Fixing a time instance t_k, the optimality equation reads

$$\alpha\Delta t\mathsf{M}_u\bar{u}_k - \Delta t\mathsf{C}^T\bar{z}_k = 0 \qquad \text{for } k \in \{1, \dots, N_t\}, \tag{2.12}$$

or, equivalently, in compact form

$$\alpha\Delta t\mathsf{M}_{\mathsf{u}\mathrm{st}}\mathsf{u} - \Delta t\mathsf{C}_{\mathrm{st}}^T\mathsf{z} = 0, \tag{2.13}$$

where $\mathsf{M}_{\mathsf{u}\mathrm{st}}$ is the block-diagonal matrix of dimension $\mathbb{R}^{\mathcal{N}_u} \times \mathbb{R}^{\mathcal{N}_u}$ which presents M_u on the diagonal.

The Adjoint Equation. The algebraic structure of the adjoint equation is a little bit more involved due to the presence of nonlinear terms still depending

on the state y. First of all, we need an explicit matrix notation for $D_y\mathcal{E}(y,u;\boldsymbol{\mu})$. Under the assumption of the quadratic nonlinearity in the state variable, the Fréchet derivative of the controlled state equation w.r.t. the state y can be written as follows: $\mathsf{E}'_{n\ell}[\mathsf{y}](\boldsymbol{\mu})+\mathsf{E}_\ell(\boldsymbol{\mu})$. Here, the linear state structure is preserved, while the nonlinear operator is linearized in $\mathsf{E}'_{n\ell}[\mathsf{y}]$, which, however, still depends on the state variable. We remark that the control operator is not involved in the formulation since it does not depend on y. Thus, employing a forward Euler method which is equivalent to an implicit scheme due the backward parabolic nature of the adjoint equation, at each time instance the equation reads:

$$\mathsf{M}_y\bar{z}_{k-1}=\mathsf{M}_y\bar{z}_k+\Delta t(-\mathsf{M}_{\text{obs}}\bar{y}_{k-1}\underbrace{-\mathsf{E}'_{n\ell}[\mathsf{y}]^T(\boldsymbol{\mu})-\mathsf{E}_\ell^T(\boldsymbol{\mu})}_{-\mathsf{E}^{\text{adj}^T}(\boldsymbol{\mu})}\bar{p}_{k-1}+\mathsf{M}_{\text{obs}}\bar{y}_{d_{k-1}}),$$

for $k\in\{N_t-1,N_t-2,\dots,1\}$, where M_{obs} is the state mass matrix restricted to the observation domain. the global space-time adjoint system has the following form:

$$\underbrace{\begin{bmatrix}\mathsf{M}_y+\Delta t\mathsf{E}^{\text{adj}^T}(\boldsymbol{\mu}) & -\mathsf{M}_y & & \\ & \ddots & \ddots & \\ & & \mathsf{M}_y+\Delta t\mathsf{E}^{\text{adj}^T}(\boldsymbol{\mu}) & -\mathsf{M}_y \\ & & & \mathsf{M}_y+\Delta t\mathsf{E}^{\text{adj}^T}(\boldsymbol{\mu})\end{bmatrix}}_{\mathsf{K}^{\text{adj}}(\boldsymbol{\mu})^T}\begin{bmatrix}\bar{z}_1\\ \bar{z}_2\\ \vdots\\ \bar{z}_{N_t}\end{bmatrix}$$

$$+\Delta t\begin{bmatrix}\mathsf{M}_{\text{obs}} & & & \\ & \mathsf{M}_{\text{obs}} & & \\ & & \ddots & \\ & & & \mathsf{M}_{\text{obs}}\end{bmatrix}\begin{bmatrix}\bar{y}_1\\ \bar{y}_2\\ \vdots\\ \bar{y}_{N_t}\end{bmatrix}=\begin{bmatrix}\Delta t\mathsf{M}_{\text{obs}}\bar{y}_{d_1}\\ \Delta t\mathsf{M}_{\text{obs}}\bar{y}_{d_2}\\ \vdots\\ \Delta t\mathsf{M}_{\text{obs}}\bar{y}_{d_{N_t}}\end{bmatrix}.$$

Thus, in a compact notation, the adjoint equation reads:

$$\mathsf{K}^{\text{adj}}(\boldsymbol{\mu})^T\mathsf{z}+\Delta t\mathsf{M}_{\text{obs}_{\text{st}}}\mathsf{y}=\Delta t\mathsf{M}_{\text{obs}_{\text{st}}}\mathsf{y}_{\text{d}}, \tag{2.14}$$

where $\mathsf{M}_{\text{obs}_{\text{st}}}\in\mathbb{R}^{\mathcal{N}_y}\times\mathbb{R}^{\mathcal{N}_y}$ is the block-diagonal matrix which entries are given by M_{obs}.

The Global System. If we collect all the aforementioned information, we can define the following all-at-once system:

$$\overbrace{\underbrace{\begin{bmatrix}\Delta t\mathsf{M}_{\text{obs}_{\text{st}}} & 0 & \mathsf{K}^{\text{adj}}(\boldsymbol{\mu})^T\\ 0 & \alpha\Delta t\mathsf{M}_{\text{u}_{\text{st}}} & -\Delta t\mathsf{C}_{\text{st}}^T\\ \mathsf{K}(\boldsymbol{\mu}) & -\Delta t\mathsf{C}_{\text{st}} & 0\end{bmatrix}}_{\mathsf{G}(\mathsf{X};\boldsymbol{\mu})}\underbrace{\begin{bmatrix}\mathsf{y}\\ \mathsf{u}\\ \mathsf{z}\end{bmatrix}}_{\mathsf{X}}}^{\mathsf{G}(\mathsf{X};\boldsymbol{\mu})}=\overbrace{\begin{bmatrix}\Delta t\mathsf{M}_{\text{obs}_{\text{st}}}\mathsf{y}_d\\ 0\\ \mathsf{M}_{\mathsf{y}_{\text{st}}}\mathsf{y}_0+\Delta t\mathsf{f}\end{bmatrix}}^{\mathsf{F}}, \tag{2.15}$$

that in residual formulation reads

$$\mathsf{R}(\mathsf{X};\boldsymbol{\mu}) := \mathsf{G}(\mathsf{X};\boldsymbol{\mu}) - \mathsf{F} = 0, \tag{2.16}$$

where $\mathsf{R}(\mathsf{X};\boldsymbol{\mu})$ will be called the *global residual* of the optimality system. We tackled the nonlinear system (2.15) with Netwon's method: namely we iteratively solve

$$\mathsf{X}^{j+1} = \mathsf{X}^{j} + \mathsf{Jac}(\mathsf{X}^{j};\boldsymbol{\mu})^{-1}(\mathsf{F} - \mathsf{G}(\mathsf{X}^{j};\boldsymbol{\mu})), \qquad j \in \mathbb{N}, \tag{2.17}$$

stopping the iterations thanks to a residual based convergence criterion. We point out that the matrix $\mathsf{K}^{\mathrm{adj}}(\boldsymbol{\mu})$ still depends on the state vector y in the term $\mathsf{E}'_{n\ell}[\mathsf{y}]^T$. Then, performing a linearization of $\mathsf{K}^{\mathrm{adj}}(\boldsymbol{\mu})[\mathsf{z}]$ w.r.t. the variable y leads to a new matrix in the formulation:

$$\mathsf{D}_{\mathsf{y}}(\mathsf{E}'_{n\ell}[\mathsf{y}^j]^T)[\mathsf{z}^j]. \tag{2.18}$$

The new term (2.18) does not depend on the state variable anymore, however still depends on the j-th evaluation of the adjoint variable. The final Jacobian matrix has the following form:

$$\mathsf{Jac}(\mathsf{X}^{j};\boldsymbol{\mu}) = \begin{bmatrix} \Delta t\mathsf{M}_{\mathrm{obsst}} + \mathsf{D}_{\mathsf{y}}(\mathsf{E}'_{n\ell}[\mathsf{y}^j]^T)[\mathsf{z}^j] & 0 & \mathsf{K}^{\mathrm{adj}}(\boldsymbol{\mu})^T \\ 0 & \alpha\Delta t\mathsf{M}_{\mathrm{ust}} & -\Delta t\mathsf{C}_{\mathrm{st}}^T \\ \mathsf{K}^{\mathrm{adj}}(\boldsymbol{\mu}) & -\Delta t\mathsf{C}_{\mathrm{st}} & 0 \end{bmatrix}, \tag{2.19}$$

Thanks to this linearization process, we can bring to light the saddle point structure arising from the constrained optimization in this dicretized setting, in total analogy with Section 1.2.2. Indeed, the Jacobian (2.19) can be written as

$$\mathsf{Jac}(\mathsf{X}^{j};\boldsymbol{\mu}) = \begin{bmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & 0 \end{bmatrix}, \tag{2.20}$$

where

$$\mathsf{A} = \begin{bmatrix} \Delta t\mathsf{M}_{\mathrm{obsst}} + \mathsf{D}_{\mathsf{y}}(\mathsf{E}'_{n\ell}[\mathsf{y}^j]^T)[\mathsf{z}^j] & 0 \\ 0 & \alpha\Delta t\mathsf{M}_{\mathrm{ust}} \end{bmatrix} \text{ and } \mathsf{B} = \begin{bmatrix} \mathsf{K}^{\mathrm{adj}}(\boldsymbol{\mu}) & -\Delta t\mathsf{C}_{\mathrm{st}} \end{bmatrix}. \tag{2.21}$$

The discrete analogous of the Brezzi Theorem relies on the invertibility of A over the kernel of B and on the *discrete Brezzi inf-sup condition* that reads:

$$\beta^{\mathcal{N}}(\boldsymbol{\mu}) := \inf_{0\neq\mathsf{z}} \sup_{0\neq\mathsf{x}} \frac{\mathsf{z}^T\mathsf{B}\mathsf{x}}{\|\mathsf{x}\|_{\mathcal{X}}\|\mathsf{z}\|_{\mathcal{Y}_0}} \geq \hat{\beta}^{\mathcal{N}}(\boldsymbol{\mu}) > 0, \tag{2.22}$$

where $\mathsf{x} = \begin{bmatrix} \mathsf{y} \\ \mathsf{u} \end{bmatrix}$. In the space-time context, the inequality (2.22) holds when the function spaces for state and adjoint coincide [89, 110, 111, 146]: this assumption is guaranteed at the discrete level too, since z lives in $\mathcal{Q}^{\mathcal{N}_y}$.

Remark 2.2.1 (Steady nonlinear OCP($\boldsymbol{\mu}$)s) *We now want to adapt the aforementioned structure to steady nonlinear problems. It is clear that, in this case, the discretization is performed only in space, namely the considered discrete variable is in* $\mathbb{X}^{\mathcal{N}} := Y^{N^y_{FE}} \times U^{N^u_{FE}} \times Y^{N^y_{FE}}$*, of dimension* $\mathcal{N} = 2N^y_{FE} + N^u_{FE}$*. The FE approximation of parametric OCP($\boldsymbol{\mu}$)s governed by nonlinear equations as defined in (1.25) reads: given* $\boldsymbol{\mu} \in \mathcal{D}$*, find* $\mathbf{X}^{\mathcal{N}} := \boldsymbol{X}^{\mathcal{N}}(\boldsymbol{\mu}) \in \mathbb{X}^{\mathcal{N}}$ *such that*

$$\mathcal{G}_s(\mathbf{X}^{\mathcal{N}}; \boldsymbol{\mu}) = \mathcal{F}_s. \tag{2.23}$$

With abuse of notation, we use here the same symbols used in Section 2.2. Indeed, we believe that this choice is not misleading, rather it will highlight the similarities with the space-time approach. Here y*,* u *and* z *indicate the column vectors which entries are given by the FE coefficients of the state, the control and the adjoint variables in their approximated spaces, respectively. The vectors of the FE coefficients for the desired state and the forcing term are denoted with* y_d *and* f*, respectively. In the FE context, we are dealing with a state equation of the form (2.9). Moreover, we call* M_y *and* M_u *the mass matrices for the state (and the adjoint) and th control, respectively. The global matrix formulation of the optimization system (2.7) is*

$$\overbrace{\begin{bmatrix} \mathsf{M}_y & 0 & \mathsf{E}'_{n\ell}[\mathsf{y}]^T + \mathsf{E}_\ell^T \\ 0 & \alpha\mathsf{M}_u & -\mathsf{C}^T \\ \mathsf{E}_{n\ell} + \mathsf{E}_\ell & -\mathsf{C} & 0 \end{bmatrix}}^{\mathsf{G}(\mathsf{X};\boldsymbol{\mu})} \underbrace{\begin{bmatrix} \mathsf{y} \\ \mathsf{u} \\ \mathsf{z} \end{bmatrix}}_{\mathsf{X}} = \overbrace{\begin{bmatrix} \mathsf{M}_y\mathsf{y}_d \\ 0 \\ \mathsf{f} \end{bmatrix}}^{\mathsf{F}}. \tag{2.24}$$

Also in the steady case, the system (2.24) can be written in residual form as in (2.16) and it can be solved with Netwon-based algorithm. As in the time-dependent setting, the arguments for the algebraic structure of $D_y\mathcal{E}(y, u; \boldsymbol{\mu})$ *applies. Indeed, the Jacobian matrix presents the saddle point structure (2.20): indeed, applying a proper definition to* A *and* B*, i.e.*

$$\mathsf{A} = \begin{bmatrix} \mathsf{M}_y + \mathsf{D}_\mathsf{y}(\mathsf{E}'_{n\ell}[\mathsf{y}^j]^T)[\mathsf{z}^j] & 0 \\ 0 & \alpha\mathsf{M}_u \end{bmatrix} \text{ and } \mathsf{B} = \begin{bmatrix} \mathsf{E}'_{n\ell}[\mathsf{y}^j] + \mathsf{E}_\ell & -\mathsf{C} \end{bmatrix}, \tag{2.25}$$

we recover (2.20).

In the next Section we will deal with linear time-dependent equations, mirroring and adapting the aforementioned contents to that simpler setting.

2.3 Space–time linear OCP($\boldsymbol{\mu}$)s

In this Section, we show the algebraic formulation of the space-time discretization related to linear OCP($\boldsymbol{\mu}$)s. The exploited structure is totally analogous

to the one presented in Section 2.2.1, however, we have decided to highlight the discretized structure we are dealing with also in this case, for the sake of clarity. First of all, we will describe the parabolic case and then we will move towards the Stokes problem, in Sections 2.3.1 and 2.3.2, respectively. Along the Sections, we will adapt the time-dependent setting to the steady and the no-control framework.

2.3.1 Parabolic Time-Dependent OCP($\boldsymbol{\mu}$)s

Let us focus on the discretization of the continuous parabolic problem described in Section 1.2.2. For the sake of completeness, we report the algebraic structure in presence of geometrical parametrization. We remark that we use the space-time strategy presented in Section 2.2.1. Namely, we consider the state and the adjoint variable $\mathcal{Q}^{\mathcal{N}_y}$ and the control in $\mathcal{U}^{\mathcal{N}_u}$. We recall that the space-time variables are y, u and z, while the desired state is y_d, the initial condition is y_0 and the forcing term is f. Here, we present the classical saddle point framework already showed in several works, see e.g. [65, 143, 144, 146, 148], adapted to the case of geometrical parametrization. In this setting, we introduce two further matrices: $\mathsf{S}(\boldsymbol{\mu})$ and $\mathsf{D}_a(\boldsymbol{\mu})$, defined as $\mathsf{S}(\boldsymbol{\mu})_{ij} = s(\phi_j, \phi_i; \boldsymbol{\mu})$ and $\mathsf{D}_a(\boldsymbol{\mu})_{ij} = a(\phi_j, \phi_i; \boldsymbol{\mu})$, for $i, j = 1, \dots, N^y_{FE}$, respectively, where $s(\cdot, \cdot; \boldsymbol{\mu})$ and $a(\cdot, \cdot; \boldsymbol{\mu})$ represent the time evolution and the state operators, respectively, as already defined in Section 1.2.2. In the following, once again, we will separately focus on the three equations of the optimality system (1.10), for the sake of clarity.

The State Equation. Here, as already specified in the previous Sections, we use a backward Euler approximation for the time evolution, i.e. for each time instance we deal with the following equation:

$$\mathsf{S}(\boldsymbol{\mu})\bar{y}_k + \Delta t \mathsf{D}_a(\boldsymbol{\mu})\bar{y}_k - \Delta t \mathsf{C}(\boldsymbol{\mu})\bar{u}_k = \mathsf{S}(\boldsymbol{\mu})\bar{y}_{k-1} + \bar{f}_k \Delta t, \tag{2.26}$$

for $k \in \{1, \dots, N_t\}$. We stress that, where the geometrical parametrization is not present, as one can notice from equation (2.10), the matrix $\mathsf{S}(\boldsymbol{\mu})$ is replaced by M_y. The algebraic system related to the state equation is

$$\underbrace{\begin{bmatrix} \mathsf{S}(\boldsymbol{\mu}) + \Delta t \mathsf{D}_a(\boldsymbol{\mu}) & & & & \\ -\mathsf{S}(\boldsymbol{\mu}) & \mathsf{S}(\boldsymbol{\mu}) + \Delta t \mathsf{D}_a(\boldsymbol{\mu}) & & & \\ & \ddots & \ddots & & \\ & & & -\mathsf{S}(\boldsymbol{\mu}) & \mathsf{S}(\boldsymbol{\mu}) + \Delta t \mathsf{D}_a(\boldsymbol{\mu}) \end{bmatrix}}_{\mathsf{K}(\boldsymbol{\mu})} \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_{N_t} \end{bmatrix}$$

$$-\Delta t \begin{bmatrix} \mathsf{C}(\boldsymbol{\mu}) & & & \\ & \mathsf{C}(\boldsymbol{\mu}) & & \\ & & \ddots & \\ & & & \mathsf{C}(\boldsymbol{\mu}) \end{bmatrix} \begin{bmatrix} \bar{u}_1 \\ \bar{u}_2 \\ \vdots \\ \bar{u}_{N_t} \end{bmatrix} = \begin{bmatrix} \mathsf{S}(\boldsymbol{\mu})\bar{y}_0 + \Delta t \bar{f}_1 \\ 0 + \Delta t \bar{f}_2 \\ \vdots \\ 0 + \Delta t \bar{f}_{N_t} \end{bmatrix}.$$

In compact form, the state equation space-time system can be written as

$$\mathsf{K}(\boldsymbol{\mu})\mathsf{y} - \Delta t \mathsf{C}(\boldsymbol{\mu})_{\mathrm{st}}\mathsf{u} = \mathsf{S}(\boldsymbol{\mu})\mathsf{y}_0 + \Delta t \mathsf{f}, \tag{2.27}$$

where $\mathsf{C}(\boldsymbol{\mu})_{\mathrm{st}}$ is the block-diagonal matrix of entries $\mathsf{C}(\boldsymbol{\mu})$ which takes into account a possible trace back of the L^2-product restricted to the control domain.

The Optimality Equation. This equation can be exactly defined as already did in (2.12) for each time instance. Furthermore, it can be recast in compact form as we did in (2.13).

The Adjoint Equation. We now focus on the discretization of the adjoint equation for each instance t_k through a forward Euler method:

$$\mathsf{S}(\boldsymbol{\mu})\bar{z}_{k-1} = \mathsf{S}(\boldsymbol{\mu})\bar{z}_k + \Delta t(-\mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs}}\bar{y}_{k-1} - \mathsf{D}_a(\boldsymbol{\mu})^T\bar{z}_{k-1} + \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs}}\bar{y}_{d_{k-1}}),$$

for $k \in \{N_t - 1, N_t - 2, \ldots, 1\}$. This equation in compact form reads:

$$\underbrace{\begin{bmatrix} \mathsf{S}(\boldsymbol{\mu}) + \Delta t \mathsf{D}_a(\boldsymbol{\mu})^T & -\mathsf{S}(\boldsymbol{\mu}) & & \\ & \ddots & \ddots & \\ & & \mathsf{S}(\boldsymbol{\mu}) + \Delta t \mathsf{D}_a(\boldsymbol{\mu})^T & -\mathsf{S}(\boldsymbol{\mu}) \\ & & & \mathsf{S}(\boldsymbol{\mu}) + \Delta t \mathsf{D}_a(\boldsymbol{\mu})^T \end{bmatrix}}_{\mathsf{K}(\boldsymbol{\mu})^T} \begin{bmatrix} \bar{z}_1 \\ \vdots \\ \bar{z}_{N_t-1} \\ \bar{z}_{N_t} \end{bmatrix}$$

$$+\Delta t \begin{bmatrix} \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs}} & & & \\ & \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs}} & & \\ & & \ddots & \\ & & & \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs}} \end{bmatrix} \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_{N_t} \end{bmatrix} = \begin{bmatrix} \Delta t \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs}}\bar{y}_{d_1} \\ \Delta t \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs}}\bar{y}_{d_2} \\ \vdots \\ \Delta t \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs}}\bar{y}_{d_{N_t}} \end{bmatrix}.$$

We recall that the matrices are, eventually, the trace back of the ones already defined in Section 2.2.1: this is the reason why we decided to explicit the dependence on the parameter $\boldsymbol{\mu}$. The space-time adjoint equation reads as:

$$\mathsf{K}(\boldsymbol{\mu})^T\mathsf{z} + \Delta t \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs_{st}}}\mathsf{y} = \Delta t \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs_{st}}}\mathsf{y}_d, \tag{2.28}$$

where $M(\boldsymbol{\mu})_{\mathrm{obs_{st}}}$ is the global observation mass matrix (the possible trace back) define din Section 2.2.1.

The Global System. Combining the aforementioned algebraic systems, we obtain the following global optimality system:

$$\overbrace{\begin{bmatrix} \Delta t \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs_{st}}} & 0 & \mathsf{K}(\boldsymbol{\mu})^T \\ 0 & \alpha \Delta t \mathsf{M}(\boldsymbol{\mu})_{\mathrm{u_{st}}} & -\Delta t \mathsf{C}(\boldsymbol{\mu})_{\mathrm{st}}^T \\ \mathsf{K}(\boldsymbol{\mu}) & -\Delta t \mathsf{C}(\boldsymbol{\mu})_{\mathrm{st}} & 0 \end{bmatrix}}^{\mathsf{G}} \underbrace{\begin{bmatrix} \mathsf{y} \\ \mathsf{u} \\ \mathsf{z} \end{bmatrix}}_{\mathsf{X}} = \overbrace{\begin{bmatrix} \Delta t \mathsf{M}_{\mathrm{obs_{st}}} \mathsf{y}_d \\ 0 \\ \mathsf{S}(\boldsymbol{\mu}) \mathsf{y}_0 + \Delta t \mathsf{f} \end{bmatrix}}^{\mathsf{F}}, \tag{2.29}$$

where $\mathsf{M}(\boldsymbol{\mu})_{\mathrm{u_{st}}}$ is the block diagonal matrix which entries are given by the possible trace back of the control mass matrix. In the following, we rewrite the linear space-time system (2.29) as:

$$\mathsf{G}\mathsf{X} = \mathsf{F}, \tag{2.30}$$

where G can be interpret in a saddle point framework recasting it as

$$\mathsf{G} = \begin{bmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & 0 \end{bmatrix}, \tag{2.31}$$

defining

$$\mathsf{A} = \begin{bmatrix} \Delta t \mathsf{M}(\boldsymbol{\mu})_{\mathrm{obs_{st}}} & 0 \\ 0 & \alpha \Delta t \mathsf{M}(\boldsymbol{\mu})_{\mathrm{u_{st}}} \end{bmatrix} \text{ and } \mathsf{B} = \begin{bmatrix} \mathsf{K}(\boldsymbol{\mu}) & -\Delta t \mathsf{C}(\boldsymbol{\mu})_{\mathrm{st}} \end{bmatrix}. \tag{2.32}$$

Namely, the saddle point structure directly arises from the optimization problem at hand, as one might expect from the continuous version of the OCP($\boldsymbol{\mu}$)s.

Remark 2.3.1 (Steady Linear OCP($\boldsymbol{\mu}$)s) *As already presented in Remark 2.2.1, it is straightforward to adapt the aforementioned structure to the steady case. Here, we focus on steady linear OCP($\boldsymbol{\mu}$)s, where, once again, $\mathbb{X}^{\mathcal{N}} = Y^{N_{FE}^y} \times U^{N_{FE}^u} \times Y^{N_{FE}^y}$ is the global space of dimension $\mathcal{N} = 2N_{FE}^y + N_{FE}^u$: namely, the variables are purely FE functions. Following the notation we introduced at the algebraic level for the system (2.29), the all-at-once steady problem reads: given $\boldsymbol{\mu} \in D$, find the vector $\mathsf{X} := [\mathsf{y}, \mathsf{u}, \mathsf{z}]^T \in \mathbb{R}^{\mathcal{N}}$ such that*

$$\mathsf{G}_s \mathsf{X} = \mathsf{F}_s. \tag{2.33}$$

Also in this case, we remark that using the same notation for the space-time variables and the purely FE variables, is not misleading, rather shows how the space-time structure is similar to standard steady one. The considered system is defined as

$$\mathsf{G}_s := \begin{bmatrix} \mathsf{M}(\boldsymbol{\mu})_{\mathsf{obs}} & 0 & \mathsf{D}_a(\boldsymbol{\mu})^T \\ 0 & \alpha \mathsf{M}(\boldsymbol{\mu})_u & -\mathsf{C}(\boldsymbol{\mu})^T \\ \mathsf{D}_a(\boldsymbol{\mu}) & -\mathsf{C}(\boldsymbol{\mu}) & 0 \end{bmatrix} \text{ and } \mathsf{F}_s := \begin{bmatrix} \mathsf{M}(\boldsymbol{\mu})_{\mathsf{obs}} \mathsf{y_d} \\ 0 \\ \mathsf{f} \end{bmatrix}. \tag{2.34}$$

where y_d *and* f *are the vectors of the FE element coefficients for the desired state and the forcing term. Once again, we can recast the problem in a saddle point formulation, where*

$$\mathsf{G}_s = \begin{bmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & 0 \end{bmatrix}, \tag{2.35}$$

with

$$\mathsf{A} = \begin{bmatrix} \mathsf{M}(\boldsymbol{\mu})_{\mathsf{obs}} & 0 \\ 0 & \alpha \mathsf{M}(\boldsymbol{\mu})_{\mathsf{u}} \end{bmatrix} \text{ and } \mathsf{B} = \begin{bmatrix} \mathsf{D}(\boldsymbol{\mu})_a & -\mathsf{C}(\boldsymbol{\mu}) \end{bmatrix}. \tag{2.36}$$

Remark 2.3.2 (The No-Control Global System) *When we are dealing with the no-control framework, a few modifications must be performed to the optimality system (2.31). First of all, exploiting (1.56), we consider only the state and the adjoint variables, namely* y *and* z*, respectively. While the discretization of the adjoint equation is totally equivalent to (2.28), for the state equation we have*

$$\mathsf{S}(\boldsymbol{\mu})\bar{y}_k + \Delta t \mathsf{D}_a(\boldsymbol{\mu})\bar{y}_k - \frac{\Delta t}{\alpha}\mathsf{C}(\boldsymbol{\mu})\bar{z}_k = \mathsf{S}(\boldsymbol{\mu})\bar{y}_{k-1} + \bar{f}_k \Delta t, \tag{2.37}$$

for $k \in \{1, \dots, N_t\}$*. These equations, if combined together, result in:*

$$\begin{bmatrix} \Delta t \mathsf{M}(\boldsymbol{\mu})_{\mathsf{obs_{st}}} & \mathsf{K}(\boldsymbol{\mu})^T \\ \mathsf{K}(\boldsymbol{\mu}) & -\frac{\Delta t}{\alpha}\mathsf{C}(\boldsymbol{\mu})_{\mathsf{st}} \end{bmatrix} \begin{bmatrix} \mathsf{y} \\ \mathsf{z} \end{bmatrix} = \begin{bmatrix} \Delta t \mathsf{M}_{\mathsf{obs_{st}}} \mathsf{y}_d \\ \mathsf{S}\mathsf{y}_0 + \Delta t \mathsf{f} \end{bmatrix}. \tag{2.38}$$

In this very specific case, the space-time optimal control is already a saddle point system. It is clear that the no-control system (2.38) has $2\mathcal{N}_y$ *as global dimension.*

Remark 2.3.3 (The Steady No-Control System) *The no-control framework can be easily adapted to steady problems. We apply a few modifications to remark 2.3.1. We define the global space as* $\mathbb{X}^{\mathcal{N}} = Y^{N^y_{FE}} \times Y^{N^y_{FE}}$ *of dimension* $\mathcal{N} = 2N^y_{FE}$*. Thus, the FE system to be solved is: given* $\boldsymbol{\mu} \in \mathcal{D}$*, find the vector* $\mathsf{X} := [\mathsf{y}, \mathsf{z}]^T \in \mathbb{R}^{2N^y_{FE}}$ *such that*

$$\mathsf{G}^{nc}_s \mathsf{X} = \mathsf{F}^{nc}_s,$$

where

$$\mathsf{G}^{nc}_s := \begin{bmatrix} \mathsf{M}(\boldsymbol{\mu})_{\mathsf{obs}} & \mathsf{D}_a(\boldsymbol{\mu})^T \\ \mathsf{D}_a(\boldsymbol{\mu}) & -\mathsf{C}(\boldsymbol{\mu}) \end{bmatrix} \quad \text{and} \quad \mathsf{F}^{nc}_s := \begin{bmatrix} \mathsf{M}(\boldsymbol{\mu})_{\mathsf{obs}} \mathsf{y_d} \\ \mathsf{f} \end{bmatrix}.$$

2.3.2 Time-Dependent OCP($\boldsymbol{\mu}$)s governed by Stokes Equations

This Section briefly describes how to deal with OCP($\boldsymbol{\mu}$)s governed by Stokes equations in a spece-time setting. Namely, we show the finite dimensional version of the optimization problem introduced in Section 1.2.3. Similarly to the

previous cases, we define the space-time variables as

$$\mathsf{v} = \begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_{N_t} \end{bmatrix} \quad \mathsf{p} = \begin{bmatrix} \bar{p}_1 \\ \vdots \\ \bar{p}_{N_t} \end{bmatrix} \quad \mathsf{u} = \begin{bmatrix} \bar{u}_1 \\ \vdots \\ \bar{u}_{N_t} \end{bmatrix} \quad \mathsf{z}^{\mathsf{v}} = \begin{bmatrix} \bar{z}_1^{\mathsf{v}} \\ \vdots \\ \bar{z}_{N_t}^{\mathsf{v}} \end{bmatrix} \text{ and } \mathsf{z}^{\mathsf{p}} = \begin{bmatrix} \bar{z}_1^{\mathsf{p}} \\ \vdots \\ \bar{z}_{N_t}^{\mathsf{p}} \end{bmatrix}, \tag{2.39}$$

to represent the state velocity, the state pressure, the control, the adjoint velocity and the adjoint pressure, respectively. Also in this case, the column vectors $\bar{v}_i, \bar{p}_i, \bar{u}_i, \bar{z}_i^{\mathsf{v}}$ and $\bar{z}_i^{\mathsf{p}}$ for $1 \leq i \leq N_t$ consist of the coefficients of the FE discretization for each time instance. Indeed, for the time t_k, the space-time variables can be expanded in the FE setting as

$$v_k^{\mathcal{N}} = \sum_1^{N_{\text{FE}}^v} v_k^i \phi^i, \quad p_k^{\mathcal{N}} = \sum_1^{N_{\text{FE}}^p} p_k^i \psi^i,$$

$$u_k^{\mathcal{N}} = \sum_1^{N_{\text{FE}}^u} u_k^i \varphi^i, \quad z_{v_k}^{\mathcal{N}} = \sum_1^{N_{\text{FE}}^y} z_{v_k}^i \phi^i \quad \text{and} \quad z_{p_k}^{\mathcal{N}} = \sum_1^{N_{\text{FE}}^y} z_{p_k}^i \psi^i.$$

In this specific setting, $\{\phi^i\}_{i=1}^{N_{\text{FE}}^v}$ are bases for $V^{N_{\text{FE}}^v}$, $\{\psi^i\}_{i=1}^{N_{\text{FE}}^p}$ are bases for $P^{N_{\text{FE}}^p}$ and $\{\varphi^i\}_{i=1}^{N_{\text{FE}}^u}$ for $U^{N_{\text{FE}}^u}$ which represent, respectively, the FE spaces for V, P and U, i.e. the function spaces for the velocity, the pressure and the control variables as defined in Section 1.2.3. For the time discretization, the time interval $[0, T]$ is divided in N_t subintervals of length Δt. The forcing term, the desired velocity and the initial condition space-time vectors are

$$\mathsf{v}_0 = \begin{bmatrix} \bar{v}_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathsf{v}_{\text{d}} = \begin{bmatrix} \bar{v_d}_1 \\ \vdots \\ \bar{v_d}_{N_t} \end{bmatrix}, \quad \text{and} \quad \mathsf{g} = \begin{bmatrix} \bar{g}_1 \\ \vdots \\ \bar{g}_{N_t} \end{bmatrix},$$

respectively. Here, in total analogy with the aforementioned variables, $\bar{v}_0$ and $\bar{v_d}_k$ and $\bar{g}_k$ and the column vectors of their FE coefficients, for $k = 1, \ldots, N_t$. We now have all the ingredients to discretize the optimality system related to Stokes equations (1.48).

The State Equation. Let us start from the state equation that reads:

$$\begin{cases} \mathsf{M}_v \bar{v}_k + \mu \Delta t \mathsf{K} \bar{v}_k + \Delta t \mathsf{D}^T \bar{p}_k = \Delta t \mathsf{M}_u \bar{u}_k + \mathsf{M}_v \bar{v}_{k-1} + \Delta t \bar{g}_k \\ \mathsf{D} \bar{y}_k = 0 \end{cases} \tag{2.40}$$

for $k \in \{1, 2, \ldots, N_t\}$, where M_v, K and M_u are the velocity mass matrix, the stiffness matrix and the control mass matrix related to the FE discretization,

respectively. Furthermore, D is the differential operator derived from the incompressibility constraint. The system in compact form can be written as:

$$\overline{\mathsf{K}} \begin{bmatrix} \mathsf{v} \\ \mathsf{p} \end{bmatrix} - \Delta t \mathsf{C}_u \mathsf{u} = \mathsf{M}_v \mathsf{v}_0 + \Delta t \mathsf{g},$$

where $\overline{\mathsf{K}}$ is:

$$\begin{bmatrix}
\mathsf{M}_v - \mu\Delta t\mathsf{K} & 0 & \cdots & 0 & \Delta t\mathsf{D}^T & 0 & \cdots & & 0 \\
\mathsf{D} & 0 & & & & & & & \\
-\mathsf{M} & \mathsf{M}_v - \mu\Delta t\mathsf{K} & 0 & \cdots & 0 & \Delta t\mathsf{D}^T & 0 & \cdots & 0 \\
0 & \mathsf{D} & & & & & & & \\
 & & \ddots & \ddots & & & & & \\
0 & \cdots & & 0 & -\mathsf{M} & \mathsf{M}_v - \mu\Delta t\mathsf{K} & 0 & \cdots & 0 \quad \Delta t\mathsf{D}^T \\
0 & \cdots & & & 0 & \mathsf{D} & & 0 \quad \cdots & 0
\end{bmatrix},$$

and $\mathsf{C}_u \in \mathbb{R}^{\mathcal{N}_u \cdot N_t} \times \mathbb{R}^{\mathcal{N}_u \cdot N_t}$ is the block-diagonal matrix which entries are given by the mass matrix of the control space, while $\overline{\mathsf{M}_v} \in \mathbb{R}^{\mathcal{N}_v \cdot N_t} \times \mathbb{R}^{\mathcal{N}_v \cdot N_t}$ is the analogous block-diagonal matrix which entries are $[\mathsf{M}_v, \cdots, \mathsf{M}_v]$.

The Optimality Equation. We now discretize the optimality equation, which has the following form:

$$\alpha \Delta t \mathsf{M}_u \bar{u}_k - \Delta t \mathsf{M}_u \bar{z}_k^{\mathsf{v}} = 0 \qquad \text{for } k \in \{1, 2, \ldots, N_t\}.$$

Thus, the global space-time optimality system is:

$$\alpha \Delta t \overline{\mathsf{M}_u} \mathsf{u} - \Delta t \mathsf{C}_u^T \mathsf{z}^{\mathsf{v}} = 0,$$

where $\overline{\mathsf{M}_u} = \mathsf{C}_u$ and $\mathsf{C}_u = \mathsf{C}_u^T$ for the Stokes case we are analysing.

The Adjoint Equation. The last step of the approximation takes into account the adjoint equation discretized with a forward Euler method as we did for the parabolic case in Section 2.3.1. The adjoint problem at each timestep reads:

$$\begin{cases} \mathsf{M}_v \bar{z}_k^{\mathsf{v}} = \mathsf{M}_v \bar{z}_{k+1}^{\mathsf{v}} + \Delta t(-\mathsf{M}_v \bar{v}_k + \mu\Delta t \mathsf{K} \bar{z}_k^{\mathsf{v}} - \Delta t \mathsf{D}^T \bar{z}_k^{\mathsf{p}} + \mathsf{M}\bar{v}_{d_k}) \\ \mathsf{D}\bar{z}_k^{\mathsf{v}} = 0. \end{cases} \tag{2.41}$$

for $k \in \{N_t - 1, N_t - 2, \ldots, 1\}$. Exploiting very familiar arguments by now, we can write the adjoint equation in compact form as

$$\Delta t \overline{\mathsf{M}_v} \mathsf{v} + \overline{\mathsf{K}}^T \begin{bmatrix} \mathsf{z}^{\mathsf{v}} \\ \mathsf{z}^{\mathsf{p}} \end{bmatrix} = \Delta t \overline{\mathsf{M}_v} \mathsf{v}_d.$$

The Global System. The whole optimality system now reads as

$$\begin{bmatrix} \Delta t\mathcal{M} & 0 & \overline{\mathsf{K}}^T \\ 0 & \alpha\Delta t\overline{\mathsf{M}_u} & -\Delta t\mathcal{C}^T \\ \overline{\mathsf{K}} & -\Delta t\mathcal{C} & 0 \end{bmatrix} \begin{bmatrix} \mathsf{y} \\ \mathsf{u} \\ \mathsf{z} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \Delta t\overline{\mathsf{M}_v}\mathsf{v}_d \\ 0 \end{bmatrix} \\ 0 \\ \begin{bmatrix} \overline{\mathsf{M}_v}\mathsf{v}_0 + \Delta t\mathsf{g} \\ 0 \end{bmatrix} \end{bmatrix}, \tag{2.42}$$

where $\mathcal{M} \in \mathbb{R}^{(N_v+N_p)\cdot N_t} \times \mathbb{R}^{N_v+N_p\cdot N_t}$ and $\mathcal{C} \in \mathbb{R}^{(N_v+N_p)\cdot N_t} \times \mathbb{R}^{N_v\cdot N_t}$ are the following block diagonal matrices

$$\mathcal{M} = \begin{bmatrix} \mathsf{M}_v & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathcal{C} = \begin{bmatrix} \mathsf{C}_u \\ 0 \end{bmatrix}, \tag{2.43}$$

while $\mathsf{y} = \begin{bmatrix} \mathsf{v} \\ \mathsf{p} \end{bmatrix}$ and and $\mathsf{z} = \begin{bmatrix} \mathsf{z}^{\mathsf{v}} \\ \mathsf{z}^{\mathsf{p}} \end{bmatrix}$. Also for the Stokes equations, it is possible to highlight a saddle point structure arised from the optimization problem. Thus, defining

$$\mathsf{A} = \begin{bmatrix} \Delta t\mathcal{M} & 0 \\ 0 & \alpha\Delta t\overline{\mathsf{M}_u} \end{bmatrix}, \quad \mathsf{B} = \begin{bmatrix} \mathsf{K} & -\Delta t\mathcal{C} \end{bmatrix}, \quad \mathsf{F} = \begin{bmatrix} \begin{bmatrix} \Delta t\overline{\mathsf{M}_v}\mathsf{v}_d \\ 0 \end{bmatrix} \\ 0 \end{bmatrix}$$

and $\mathsf{G}_\mathsf{g} = \begin{bmatrix} \overline{\mathsf{M}_v}\mathsf{v}_0 + \Delta t\mathsf{g} \\ 0 \end{bmatrix}$, the optimality system (2.42) translates as:

$$\begin{bmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & 0 \end{bmatrix} \begin{bmatrix} \mathsf{x} \\ \mathsf{z} \end{bmatrix} = \begin{bmatrix} \mathsf{F} \\ \mathsf{G}_\mathsf{g} \end{bmatrix}, \tag{2.44}$$

once defined $\mathsf{x} = \begin{bmatrix} \mathsf{y} \\ \mathsf{u} \end{bmatrix}$.

This Chapter focused on the high fidelity space-time approximation for OCP($\boldsymbol{\mu}$)s and on their saddle point algebraic structure. We have shown several governing equations, guided by the numerical results we will show later on in the thesis. We introduced:

- time-dependent nonlinear problems,
- steady nonlinear problems,
- parabolic problems (both in the standard and in the no-control framework),
- steady linear problems,
- time-dependent Stokes equations.

The main reason why we highlighted the algebraic structure of OCP($\boldsymbol{\mu}$)s relies in the issue of the computational costs needed for their simulations, most of all in a time-dependent framework. Moreover, the required computational resources grow in a parametric setting, where several instances of $\boldsymbol{\mu} \in \mathcal{D}$ are studied and many simulations must be performed in a small amount of time. Indeed, despite the versatility of the space-time structure, its dimensionality drastically increases when mesh refinement both in space and time is employed. The applications we are going to present, were solved through a direct solver, in a one-shot fashion. This strategy can lead to unbearable simulations in the high fidelity context. To tackle this problem, nevertheless, iterative algorithms based on Krylov solvers and Schur preconditioning specifically built for saddle point structures [21] can be exploited and they have been successfully exploited for OCP($\boldsymbol{\mu}$)s in [138, 144], for example. However, the proposed iterative methods might be not enough to deal with many-query or real-time applications. Thus, other techniques must be used aiming at exploiting the parametric structure of the system at hand solving it in a low dimensional framework: the ROMs, that will be the main topic of the next Chapters.

Part II

Model Order Reduction for Parametric Optimal Control Problems

Chapter 3

Reduced Order Methods for Parametric Optimal Control Problems

The aim of this Chapter is to introduce ROMs for OCP($\boldsymbol{\mu}$)s. The presented strategies and concepts are suited to a broad class of governing equations, from time-dependent nonlinear ones to steady linear problems. The ROMs are a tool to exploit the parametric structure of the system at hand in order to solve a complex parametric problem in a low dimensional framework, accelerating the solution process. Indeed, many applications depend on parameters which can represent several physical phenomena or geometrical structures and the usefulness of PDE($\boldsymbol{\mu}$) is unquestioned in this context. However, their simulation can be computationally unfeasible with standard discretization techniques. This issue is amplified for OCP($\boldsymbol{\mu}$)where the high fidelity dimension grows due to the optimality system nature. In this Chapter, we will present the ROMs in the specific setting of OCP($\boldsymbol{\mu}$)s problems. In Section 3.1, we will bring to light the main ideas and the motivation behind the employment of ROMs. Thus we will move towards specific techniques to ensure the well-posedness of OCP($\boldsymbol{\mu}$)s in a reduced saddle point setting, in Section 3.2. In this Chapter we will deal with the three-equation setting, while the topic of the no-control framework is postponed in Chapter 5.

3.1 Preliminaries

This Section focuses on the motivations and the assumptions underlying the application of ROMs to OCP($\boldsymbol{\mu}$)s. First, we will discuss how ROMs can be employed in order to reach faster simulations w.r.t. standard space-time or FE approximations. This will be the topic of Section 3.1.1. Then, in Section 3.1.2, we will show some crucial assumptions to deal with efficient well-posed reduced simulations for several parametric instances.

3.1.1 Reduced Order Approximation

In this Section, we aim at providing basic knowledge about ROMs for OCP($\boldsymbol{\mu}$)s. Along the Thesis, we tried to point out clear that OCP($\boldsymbol{\mu}$)s are very useful and complete models. OCP($\boldsymbol{\mu}$)s have been employed as a *data-assimilation* tool in several scientific fields. The importance of a parametric setting is undoubted in many contexts: indeed, a parameter $\boldsymbol{\mu} \in \mathcal{D} \subset \mathbb{R}^d$ can describe physical or geometrical configurations and the OCP($\boldsymbol{\mu}$)s are exploited to better study and

deeply understand the considered physical phenomenon one is interested in. In this specific context, the parametrization can be not only present in the physics or the geometry of the optimality system, but it can also arise in the observation and OCP($\boldsymbol{\mu}$)s might be employed to respond to the need of driving the model solution towards a parametric instance, this for several values of $\boldsymbol{\mu}$. However, solving OCP($\boldsymbol{\mu}$)s with standard discretization techniques, i.e. in a high fidelity fashion, is related to large computational costs. Indeed:

- the optimal solutions derive from the system (1.10) which consists of three equations, i.e. the state equation, the adjoint equation and the optimality equation. This is a complication w.r.t. standard solutions of PDE($\boldsymbol{\mu}$)s which relies only on one forward problem;
- OCP($\boldsymbol{\mu}$)s are usually related to time consuming activities such us inverse problems, parameter estimation, statistical analysis... In other words, they are usually combined to a large number of simulations for several values of $\boldsymbol{\mu} \in \mathcal{D}$.

Nevertheless, those optimality solutions are usually required in a small amount of time. Indeed, in many applications, an optimization process based on optimal control strategies must be performed in a *real-time* context as well as in a *many-query* one. Namely, many evaluations of the optimal solution $\boldsymbol{X}(\boldsymbol{\mu})$ are required, and, furthermore, they are needed in a small amount of time. Among the applications relying on such a paradigm we can mention forecasting modelling and management plans modelling guided by data collection. This can be of interest in many fields, from natural sciences to industrial applications. To achieve the goal of dealing with OCP($\boldsymbol{\mu}$)s models under parametric actions, we rely on ROMs. This discretization approach exploits the parametric structure of the problem builds a low dimensional subspace of the finite dimensional high fidelity space to represent the changing of the system w.r.t. $\boldsymbol{\mu} \in \mathcal{D}$. Thus, a Galerkin projection can be performed in this low dimensional framework, say the *reduced space*, to solve several parametric instances in a faster way. This procedure gives us the opportunity to obtain information on the system at hand in a small amount of time with less computational efforts. The computational time gained can be employed to better study and analyse the optimal solutions w.r.t. several parameters. For an introduction to ROMs and the strategies that can be exploited to build the reduced space, the interested reader may refer to [64, 123, 125, 133, 134]. For the sake of clarity, we will always specify the explicit $\boldsymbol{\mu}-$dependency in the involved variables to stress the role of the parametric setting. Let the value $\boldsymbol{\mu}$ vary in the parameter space $\mathcal{D}$. Thus, we can define the ensemble of the parametric solutions of the optimality system in $\mathbb{X}$ as

$$\mathbb{M} := \{\boldsymbol{X}(\boldsymbol{\mu}) := (y(\boldsymbol{\mu}), u(\boldsymbol{\mu}), z(\boldsymbol{\mu})) \mid \boldsymbol{\mu} \in \mathcal{D}\}.$$

Here, no distinction is made for steady and time-dependent problems. In the first case we define $\mathbb{X} = Y \times U \times Y$, while we will consider $\mathbb{X} = \mathcal{Y}_0 \times \mathcal{U} \times \mathcal{Y}_T$

for time-dependent OCP($\boldsymbol{\mu}$)s. We assume that the set $\mathbb{M}$, i.e. the *continuous solution manifold*, to be smooth w.r.t. the parameter $\boldsymbol{\mu}$. Now, let us suppose to have applied some kind of approximation of the continuous space $\mathbb{X}$: namely let $\mathbb{X}^{\mathcal{N}} \subset \mathbb{X}$ be an high fidelity approximation of the continuous function space. Here, we consider both steady and time-dependent problems. Indeed:

- for steady OCP($\boldsymbol{\mu}$)s, $\mathcal{N}$ will represent the global dimension of a FE space,
- for space-time OCP($\boldsymbol{\mu}$)s, the $\mathcal{N}$ will represent the global space-time dimension.

If we take into consideration this high fidelity approximation, we can define the set of all the high fidelity solutions as:

$$\mathbb{M}^{\mathcal{N}} := \{\boldsymbol{X}(\boldsymbol{\mu})^{\mathcal{N}} := (y(\boldsymbol{\mu})^{h_y}, u(\boldsymbol{\mu})^{h_u}, z(\boldsymbol{\mu})^{h_y}) \mid \boldsymbol{\mu} \in \mathcal{D}\},$$

where h_y is N^y_{FE} for the steady case and $h_y = \mathcal{N}_y$ for the time-dependent space-time approximation. The same argument applies to the control variable that is sought in a discrete function space of dimension $h_u = N^u_{FE}$ in the steady framework, while $h_u = \mathcal{N}_u$ in the space-time setting. It is clear that if $\mathcal{N}$ is large enough, $\mathbb{M}^{\mathcal{N}}$ is a good approximation of $\mathbb{M}$. The goal of ROMs is to represent the structure of the *high fidelity solution manifold* $\mathbb{M}^{\mathcal{N}}$ with a reduced space of low dimension built through the employment of a linear combination of *snapshots*, i.e. high fidelity evaluations of the optimal solution $\boldsymbol{X}(\boldsymbol{\mu})$ computed in properly chosen parameters values. The snapshots are collected and manipulated through the employment of algorithms based on data exploring techniques. In this contribution we will focus on Proper Orthogonal Decomposition (POD) [15, 30, 35, 64] and Greedy algorithm [29, 64]. Their description is postponed in Chapters 4 and 5, respectively. For now, let us assume to have build a low dimensional framework $\mathbb{X}_N \subset \mathbb{X}^{\mathcal{N}}$ able to capture the behaviour of the global solution $\boldsymbol{X}^{\mathcal{N}}(\boldsymbol{\mu})$ w.r.t. some changes in the parameter $\boldsymbol{\mu}$. Namely, $\mathbb{X}_N := \mathbb{Y}^y_N \times \mathbb{U}_N \times \mathbb{Y}^z_N$, where $\mathbb{Y}^y_N$ and $\mathbb{Y}^z_N$ are the reduced space for the state and the adjoint variables, respectively, and $\mathbb{U}_N$ is the reduced function space for the control variable. Here, we are highlighting a very important concept: even if the space-time of the FE approximation coincide for state and adjoint, the reduced function spaces can be different since they are based on the linear combination of snapshots of different problem variables. Once provided of these reduced function spaces, a standard Galerkin projection is performed in the reduced space in order to find a low dimensional optimal solution in a small amount of time not paying in accuracy w.r.t. the high fidelity optimal solution. Indeed, the reduced problem reads: given $\boldsymbol{\mu} \in \mathcal{D}$, find $\boldsymbol{X}_N(\boldsymbol{\mu}) \in \mathbb{X}_N$ such that

$$G(\boldsymbol{X}_N(\boldsymbol{\mu}), \boldsymbol{\Xi}; \boldsymbol{\mu}) = \langle \mathcal{F}, \boldsymbol{\Xi} \rangle_{\mathbb{X}^{\mathcal{N}*}, \mathbb{X}^{\mathcal{N}}} \quad \forall \boldsymbol{\Xi} \in \mathbb{X}_N. \tag{3.1}$$

Once again, we do not distinguish between steady or time-dependent problems. Indeed, the techniques that we are going to present can adapt to both the

formulations. In Chapter 2 we already pointed out that the strength of the space-time formulation relies in its similarity with the steady framework. Thus, the standard reduced techniques we are going to present apply to both the typologies of OCP($\boldsymbol{\mu}$)s. The main idea behind the ROMs application is that N is much lower than the high fidelity dimension $\mathcal{N}$. While the building phase still depends on $\mathcal{N}$, the projection phase can be performed rapidly for each new parametric evaluation. In the next Section, we will describe the *offline-online* paradigm, which assures an efficient and fast solution of the reduced system for each new value of $\boldsymbol{\mu} \in \mathcal{D}$.

3.1.2 Offline-Online Decomposition

This Section focuses on some crucial assumptions one has to guarantee in order to rely on an efficient application of the ROMs. Indeed, one of the main goals of reduced approaches is to achieve a rapid solution of the problem at hand. This aim can be reached assuming that the ROMs can be divided in two stages, an *offline phase* and an *online phase*. The construction of the bases is the crucial process of the *offline phase* together with the assembly and the storage of all the $\boldsymbol{\mu}$-independent quantities. After this procedure, given a parameter $\boldsymbol{\mu} \in \mathcal{D}$, the *reduced optimality system* (3.1) can be written as follows:

$$\begin{cases} D_y\mathscr{L}(\boldsymbol{X}_N(\boldsymbol{\mu}); y_\mathrm{d}, \boldsymbol{\mu})[\omega] = 0 & \forall \omega \in \mathbb{Y}_N^y, \\ D_u\mathscr{L}(\boldsymbol{X}_N(\boldsymbol{\mu}); y_\mathrm{d}, \boldsymbol{\mu})[\kappa] = 0 & \forall \kappa \in \mathbb{U}_N, \\ D_z\mathscr{L}(\boldsymbol{X}_N(\boldsymbol{\mu}); y_\mathrm{d}, \boldsymbol{\mu})[\zeta] = 0 & \forall \zeta \in \mathbb{Y}_N^z. \end{cases} \tag{3.2}$$

We recall that $\mathbb{Y}_N^y$, $\mathbb{U}_N$ and $\mathbb{Y}_N^z$ are the reduced spaces obtained by the manipulation of the state, the control and the adjoint snapshots, respectively. As already specified, the Galerkin projection (3.2) is preferable to be performed in a small amount of time. Let us assume that the equations of (3.2) can be written as:

$$\begin{aligned} D_y\mathscr{L}(\boldsymbol{X}_N, y_\mathrm{d}, \boldsymbol{\mu})[\omega] &= \sum_{\mathsf{q}=1}^{Q_y} \Theta_y^{\mathsf{q}}(\boldsymbol{\mu}) D_y\mathscr{L}^{\mathsf{q}}(\boldsymbol{X}_N; y_\mathrm{d})[\omega], \\ D_u\mathscr{L}(\boldsymbol{X}_N; y_\mathrm{d}, \boldsymbol{\mu})[\kappa] &= \sum_{\mathsf{q}=1}^{Q_u} \Theta_u^{\mathsf{q}}(\boldsymbol{\mu}) D_u\mathscr{L}^{\mathsf{q}}(\boldsymbol{X}_N; y_\mathrm{d})[\kappa], \\ D_z\mathscr{L}(\boldsymbol{X}_N; y_\mathrm{d}, \boldsymbol{\mu})[\zeta] &= \sum_{\mathsf{q}=1}^{Q_z} \Theta_z^{\mathsf{q}}(\boldsymbol{\mu}) D_z\mathscr{L}^{\mathsf{q}}(\boldsymbol{X}_N; y_\mathrm{d})[\zeta]. \end{aligned} \tag{3.3}$$

Here, the system can be seen as the product of $\boldsymbol{\mu}-$dependent smooth functions $\Theta_y^{\mathsf{q}}(\boldsymbol{\mu}), \Theta_u^{\mathsf{q}}(\boldsymbol{\mu}), \Theta_z^{\mathsf{q}}(\boldsymbol{\mu})$ and $\boldsymbol{\mu}-$independent forms

$$D_y\mathscr{L}^{\mathsf{q}}(\boldsymbol{X}_N, y_\mathrm{d})[\omega], D_u\mathscr{L}^{\mathsf{q}}(\boldsymbol{X}_N; y_\mathrm{d})[\kappa], \text{ and } D_z\mathscr{L}^{\mathsf{q}}(\boldsymbol{X}_N; y_\mathrm{d})[\zeta].$$

In this case, we say that the system respects the *affine decomposition* paradigm. It is now clear that, when this happens, the ROMs approach can be divided efficiently in:

- an *offline stage*, where the snapshots are manipulated and the basis functions are built. Here, all the quantities that are $\boldsymbol{\mu}-$independent are precomputed and stored. This stage, that still depends on the global dimension $\mathcal{N}$, is performed only once.

- An *online stage*, where the $\boldsymbol{\mu}-$dependent quantities are evaluated and the optimality system (3.2) is assembled and solved. This stage does not depend on the high fidelity dimension and assures the solution of the system in a smaller amount of time w.r.t. standard FE or space-time approaches.

Namely, the reduced projection might be preceded by a more expensive *offline stage*, consisting in the bases construction and the assembly of the quantities that do not depend on $\boldsymbol{\mu}$. After this possibly costly phase which, however, is performed only once, the fast projection-based *online phase* follows. In this stage, the optimality system (3.2) is solved for several evaluations of the parameter $\boldsymbol{\mu}$.

Remark 3.1.1 *In the nonlinear case, even when the structure* (3.3)*, i.e. the affinity w.r.t. the parameter* $\boldsymbol{\mu}$*, is verified, the nonlinear forms still depends on* $\boldsymbol{X}^{\mathcal{N}}(\boldsymbol{\mu})$*. Thus, both the offline and the online stages involve the assembly (and projection) of the high fidelity solution. This issue can be tackled through the employment of hyper-reduction techniques such as the Empirical Interpolation Method (EIM). The interested reader may refer to [17] or to [64, Chapter 5].*

We have underlined the usefulness of reduced approaches and how they can be applied to accelerate simulations. We stressed that this is of utmost importance most of all in a parametric optimal control context, where a system of three equations must be solved for several parametric instances. The following Section deals with the reduced structure of the problem at hand and the techniques that are usually employed to guarantee the well-posedness of the reduced optimality system (3.2).

3.2 Well-Posedness of Reduced Order OCP($\boldsymbol{\mu}$)s

This Section focuses on the structure of the reduced optimality system and on the specific techniques that assure its well-posedness in most of the PDE($\boldsymbol{\mu}$)s we have introduced in the previous Chapters. We start highlighting the reduced saddle point structure in Section 3.2.1. Then we will introduce the aggregated spaces strategy together with the supremizer stabilization for Stokes and Navier-Stokes problems in Section 3.2.2.

3.2.1 Reduced Order Saddle Point Structure

As already specified in Chapters 1 and 2, PDE($\boldsymbol{\mu}$)s constrained optimization problems lead to an optimality system that can be recast in saddle point formulation. Let us take into consideration the linear problems. At the continuous level, this holds when the constraints are linear PDE($\boldsymbol{\mu}$)s, both in the steady [21, 26, 67, 111, 110, 144] and in the time-dependent case [65, 66, 111, 110, 143, 144, 146, 150]. However, we have shown that this still holds true at the discrete level [65, 66, 143, 144, 146, 150]. For the sake of clarity, we recall the high fidelity structure of the linear problem at hand. In Chapter 2, we showed that the discretized OCP($\boldsymbol{\mu}$)s can be written as follows, for a given $\boldsymbol{\mu} \in \mathcal{D}$:

$$\mathsf{G}\mathsf{X} = \mathsf{F}, \tag{3.4}$$

where G has the following saddle point structure

$$\mathsf{G} = \begin{bmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & 0 \end{bmatrix}. \tag{3.5}$$

In order to lighten the notation, we are omitting the parameter dependence from the matrices and from the vectors of the discrete system we are analysing. The main features of the saddle point structure proposed in (3.5) is that its well-posedness strictly relies on the the inf-sup condition of the state equation represented by the matrix B, besides the invertibility of A. Namely, w.r.t. the state equation, given a $\boldsymbol{\mu} \in \mathcal{D}$, the following inequality must hold:

$$\inf_{0 \neq \mathsf{z}} \sup_{0 \neq \mathsf{x}} \frac{\mathsf{z}^T \mathsf{B} \mathsf{x}}{\|\mathsf{x}\|_{\mathbb{Y} \times \mathbb{U}} \|\mathsf{z}\|_{\mathbb{Y}}} \geq \hat{\beta}^{\mathcal{N}}(\boldsymbol{\mu}) > 0, \tag{3.6}$$

see [11, 27, 28] as references. In Chapter 2, at the discrete level, we showed that the relation (3.6) is verified thanks to the assumption of the coinciding space for the state and the adjoint variable, while A the invertibility simply derives from its features. However, as already specified, in the ROMs setting, even if the high fidelity space is the same for both the variables, the two reduced spaces $\mathbb{Y}_N^y$ and $\mathbb{Y}_N^z$ may not coincide. Indeed, the ROMs bases construction relies on snapshots manipulation and the space spanned by the state variable can be different from the one spanned by the adjoint one. Let us suppose to have applied a standard reduction algorithm that led us to basis functions for the state, the control and the adjoint variables respectively defined as:

$$\begin{aligned} \mathbb{Y}_N^y &= \text{span } \{\chi_n^y,\ n = 1, \dots, N\}, \\ \mathbb{U}_N &= \text{span}\{\chi_n^u,\ n = 1, \dots, N\}, \\ \mathbb{Y}_N^z &= \text{span } \{\chi_n^z,\ n = 1, \dots, N\}. \end{aligned}$$

Here, for the sake of clarity, we are assuming that all the three reduced spaces have all dimension N. Thus, we can build the basis matrices

$$\mathsf{Z}_{\mathsf{x}} = \begin{bmatrix} \mathsf{Z}_{\mathsf{y}} \\ \mathsf{Z}_{\mathsf{u}} \end{bmatrix}, \qquad \mathsf{Z}_{\mathsf{z}} \qquad \text{and} \qquad \mathsf{Z} = \begin{bmatrix} \mathsf{Z}_{\mathsf{x}} \\ \mathsf{Z}_{\mathsf{z}} \end{bmatrix},$$

where $\mathsf{Z}_\mathsf{y} = [\chi_1^y | \cdots | \chi_N^y] \in \mathbb{R}^{h_y \times N}$ and an analogous definition can be given for the matrices $\mathsf{Z}_\mathsf{u} \in \mathbb{R}^{h_u \times N}$ and $\mathsf{Z}_\mathsf{z} \in \mathbb{R}^{h_y \times N}$. When the spaces have been built, the optimality system (3.4) is solved in a low dimensional framework through a Galerkin projection into the reduced spaces. The reduced system, for each parametric instance $\boldsymbol{\mu}$, thus, reads:

$$\mathsf{G}_N \mathsf{X}_N = \mathsf{F}_N, \tag{3.7}$$

where

$$\mathsf{G}_N := \mathsf{Z}^T \mathsf{G}, \qquad \mathsf{X}_N := \mathsf{Z}^T \mathsf{X}, \qquad \text{and} \qquad \mathsf{F}_N := \mathsf{Z}^T \mathsf{F}.$$

Here,

$$\mathsf{G}_N = \begin{bmatrix} \mathsf{A}_N & \mathsf{B}_N^T \\ \mathsf{B}_N & 0 \end{bmatrix}, \tag{3.8}$$

with $\mathsf{A}_N = Z_\mathsf{x}^T \mathsf{A} \mathsf{Z}_\mathsf{x}$ and $\mathsf{B}_N = \mathsf{Z}_\mathsf{z}^T \mathsf{B} \mathsf{Z}_\mathsf{x}$. The just presented framework is typical of linear OCP($\boldsymbol{\mu}$)s [110, 111, 145, 146]. However, the same saddle point structure is present in the solution of nonlinear problems through Newton's method, as we have already seen at the high fidelity level in Section 2.2. Here, we report the reduced version of nonlinear OCP($\boldsymbol{\mu}$)s. Once again, we are not making difference between time-dependency or steadiness, since the algorithms we are going to show in Chapters 4 and 5 are suited to both the problem formulations. We recall that, in the nonlinear setting, we want to solve a problem like

$$\mathsf{G}(\mathsf{X})\mathsf{X} = \mathsf{F}, \tag{3.9}$$

where the right hand side still depends on the global high fidelity variable X. Employing a Galerkin projection into the reduced spaces, as we did for the linear case, we obtain the following reduced nonlinear optimality system:

$$\mathsf{G}_N(\mathsf{X}_N)\mathsf{X}_N = \mathsf{F}_N, \tag{3.10}$$

where $\mathsf{G}_N(\mathsf{X}_N) := \mathsf{Z}^T \mathsf{G}(\mathsf{Z}\mathsf{X}_N)$ and F_N is defined in analogy with the linear case. The system (3.10) inherits the nonlinearity from the high fidelity one. Also in the reduced setting we can rely on Newton's method and iteratively solve

$$\mathsf{X}_N^{j+1} = \mathsf{X}_N^j + \mathsf{Jac}_N(\mathsf{X}_N^j)^{-1}(\mathsf{F}_N - \mathsf{G}(\mathsf{X}_N^j)\mathsf{X}_N^j), \qquad j \in \mathbb{N}, \tag{3.11}$$

where the Jacobian matrix is ohf the form

$$\mathsf{Jac}_N(\mathsf{X}_N)\mathsf{X}_N = \begin{bmatrix} \mathsf{A}_N & \mathsf{B}_N^T \\ \mathsf{B}_N & 0 \end{bmatrix} \begin{bmatrix} \mathsf{x}_N \\ \mathsf{z}_N \end{bmatrix}, \tag{3.12}$$

with $\mathsf{Jac}_N(\mathsf{X}_N) = \mathsf{Z}^T \mathsf{Jac}(\mathsf{Z}\mathsf{X}_N)\mathsf{Z}$, $\mathsf{A}_N = \mathsf{Z}_\mathsf{x}^T \mathsf{A} \mathsf{Z}_\mathsf{x}$ and $\mathsf{B}_N = \mathsf{Z}_\mathsf{z}^T \mathsf{B} \mathsf{Z}_\mathsf{x}$.
These formulations imply that, also at the reduced level, we need to guarantee the fulfillment of the *reduced inf-sup condition* of the form

$$\inf_{0 \neq \mathsf{z}_N} \sup_{0 \neq \mathsf{x}_N} \frac{\mathsf{z}_N^T \mathsf{B}_N \mathsf{x}_N}{\|\mathsf{x}_N\|_{\mathbb{Y} \times \mathbb{U}} \|\mathsf{z}_N\|_\mathcal{Y}} \geq \hat{\beta}_N(\boldsymbol{\mu}) > 0. \tag{3.13}$$

It is clear that the inequality is provable when the reduced spaces for the state and the adjoint variables coincide. However, we already pointed out that this property does not directly derive from the coincidence of the spaces at the high fidelity discrete level. To assure the well-posedness of the system, then, we need to *modify* the bases functions. In the next Section, we will show the techniques we exploit to reach this goal.

3.2.2 Aggregated Spaces and Supremizer Stabilization

In this Section, we will introduce two techniques which will guarantee the fulfillment of the inf-sup stability condition for reduced OCP($\boldsymbol{\mu}$)s: the *aggregated space strategy* and the *supremizer enrichment*. The aggregated space technique has been widely exploited in the steady setting [12, 13, 44, 51, 79, 80, 110, 111, 127]. However, it can be applied also to more general problems, such as time-dependent and nonlinear ones [147, 146, 148]. As already specified in Section 3.1.2, we are assuming that standard reduced algorithms might lead to three different spaces for the state, the control and the adjoint variables. Besides, we already remarked that, if the state and the adjoint spaces coincide at the high fidelity level, they might be different at the reduced one. The aggregated space manipulates the basis functions in order to prove the inequality (3.13). The aggregated space reaches this goal building a common space for both the state and adjoint variables. Namely, the new spaces are given by

$$\mathbb{Y}_N = \text{span } \{\chi_n^y, \chi_n^z,\ n = 1, \dots, N\}, \tag{3.14}$$

$$\mathbb{U}_N = \text{span}\{\chi_n^u,\ n = 1, \dots, N\}. \tag{3.15}$$

After this procedure, the basis matrices are:

$$\mathsf{Z} = \begin{bmatrix} \mathsf{Z}_{\mathsf{x}} \\ \mathsf{Z}_{\mathsf{z}} \end{bmatrix} \qquad \text{and} \qquad \mathsf{Z}_{\mathsf{x}} = \begin{bmatrix} \mathsf{Z}_{\mathsf{y}} \\ \mathsf{Z}_{\mathsf{u}} \end{bmatrix},$$

where $\mathsf{Z}_{\mathsf{y}} \equiv \mathsf{Z}_{\mathsf{z}} = [\chi_1^y | \cdots | \chi_N^y | \chi_1^z | \cdots | \chi_N^z] \in \mathbb{R}^{h_y \times 2N}$ and $\mathsf{Z}_{\mathsf{u}} = [\chi_1^u | \cdots | \chi_N^u] \in \mathbb{R}^{h_u \times N}$. This technique is actually increasing the dimension of the global system since the global reduced dimension is $5N$. Still, exploiting aggregated ROMs for OCP($\boldsymbol{\mu}$)s will result in a computational advantage in terms of time needed for parametric simulations.

Now, we are going to introduce the supremizer enrichment. We recall that the OCP($\boldsymbol{\mu}$)s governed by Stokes equations can be recast in saddle point framework, as we said in Section 2.3.2. The same also holds for the evaluation of the Jacobian of the Navier-Stokes equations which can be recast in a saddle point due to the linearization of the nonlinear framework proposed in Section 2.2. Dealing with Stokes and Navier-Stokes equations leads to a *nested saddle point* structure. Indeed, not only the global inf-sup condition (3.13) has to be taken into account, but also the inf-sup condition related to the controlled state equation. Let us recall the main ingredients of a Stokes (or Navier-Stokes) problem. The

variables involved are $v \in \mathbb{V}$, $p \in \mathbb{P}$ and $u \in \mathbb{U}$: they represent the velocity state variable, the pressure state variable and the control variable, respectively. As we already specified, in this Chapter the notation $\mathbb{V}$ comprises both the steady and the time-dependent case. The same applies to the other function spaces $\mathbb{P}$ and $\mathbb{U}$. Now, defining the continuity matrix as

$$\mathsf{D}_{ij} = \mathsf{b}(\phi^i, \psi^j), \tag{3.16}$$

where $\{\phi^k\}_{k=1}^{N_{FE}^v}$ and $\{\psi^k\}_{k=1}^{N_{FE}^p}$ are the FE basis functions for the velocity and pressure spaces V and P and $\mathsf{b}(\cdot,\cdot)$ represents the continuity constraint of Stokes (and Navier-Stokes) equations. Thus, the following inf-sup stability condition must be verified

$$\beta^S(\boldsymbol{\mu}) := \inf_{0\neq \mathsf{p}} \sup_{0\neq \mathsf{v}} \frac{\mathsf{p}^T \mathsf{D} \mathsf{v}}{\|\mathsf{v}\|_{\mathbb{V}} \|\mathsf{p}\|_{\mathbb{P}}} \geq \hat{\beta}^S(\boldsymbol{\mu}) > 0. \tag{3.17}$$

The inf-sup stability condition (3.17) is essential in Stokes (and Navier-Stokes) problem to guarantee the solvability of the system. Indeed, the saddle point structure of the state equation is quite straightforward to visualize in the steady case, see e.g. [128], but also in the time-dependent framework as one can notice from the continuous formulation in equation (1.52) and in the discretized setting of the form (2.40). At the reduced level, once performing a Galerkin projection, we end up with $\mathsf{D}_N := \mathsf{Z}_\mathsf{p}^T \mathsf{D} \mathsf{Z}_\mathsf{v}$, where Z_v and Z_p are the basis function matrices for velocity and pressure, respectively. Now we want to understand the velocity space that will be able to guarantee the following reduced inf-sup condition:

$$\beta_N^S(\boldsymbol{\mu}) := \inf_{0\neq \mathsf{p}_N} \sup_{0\neq \mathsf{v}_N} \frac{\mathsf{p}_N^T \mathsf{D}_N \mathsf{v}_N}{\|\mathsf{v}_N\|_{\mathbb{V}} \|\mathsf{p}_N\|_{\mathbb{P}}} \geq \hat{\beta}_N^S(\boldsymbol{\mu}) > 0. \tag{3.18}$$

To build the reduced velocity space, we employed the *supremizer stabilization*, see e.g. [136] This approach is based on a supremizer operator $T^{\boldsymbol{\mu}} : \mathbb{P}^{h_p} \to \mathbb{V}^{h_v}$ as follows:

$$(T^{\boldsymbol{\mu}} s, \phi)_{\mathbb{V}} = \mathsf{b}(\phi, s; \boldsymbol{\mu}), \qquad \forall \phi \in \mathbb{V}^{h_v}. \tag{3.19}$$

Then, the reduced velocity space is enriched thanks to the the pressure supremizers as follows:

$$\mathbb{V}_N = \text{span}\{\chi_n^v,\ \chi_n^{T_p}, \chi_n^{z_v},\ \chi_n^{T_{z_p}},\ n = 1, \dots, N\},$$

where $\chi_n^{T_p}$ and $\chi_n^{T_{z_p}}$ are the supremizers basis functions obtained through the state and the adjoint pressure snapshots, respectively. Enlarging in this way the reduced space for the velocity variable will guarantee the fulfillment of the inf-sup condition (3.18). The stability of this strategy has been numerically demonstrated in [15, 51, 110, 133, 135]. Here, for the sake of clarity, we report

the final reduced spaces employed for all the considered variables:

$$\begin{aligned}
\mathbb{V}_N &= \text{span}\{\chi_n^v,\ \chi_n^{T_p}, \chi_n^{z_v},\ \chi_n^{T_{z_p}},\ n = 1, \dots, N\},\\
\mathbb{P}_N &= \text{span}\{\chi_n^p, \chi_n^{z_p},\ n = 1, \dots, N\}\\
\mathbb{U}_N &= \text{span}\{\chi_n^u,\ n = 1, \dots, N\},
\end{aligned}$$

where $\mathbb{V}_N$ and $\mathbb{P}_N$ have been exploited both for the state and the adjoint velocity and pressure variables. This reduced technique, that relies on the supremizer stabilization and the aggregated spaces, guarantees the inf-sup stability of a whole optimlaity system based on Stokes or Navier-Stokes equations. Even if the reduced dimension increases from $3N$ to $13N$, the ROMs will be still convenient if compared with the high fidelity approximation.

Remark 3.2.1 *We stress that, while in the linear context the inveritibility* A *is inherited from the continuous well-posedness of the problem*[1]*, at the nonlinear level, the definition of* A *comprises some additional terms deriving from the linearization of the adjoint equation w.r.t. the state variable and the (problem dependent) proof might be not so straightforward.*

In the next Chapter, we will show some first numerical results through the employment of the POD, aiming at showing how ROMs for OCP($\boldsymbol{\mu}$)s can be useful in many scientific contexts. Among them, we will focus on environmental sciences, proposing some OCP($\boldsymbol{\mu}$)s based on real case scenario studies, from pollutant release control to coastal management.

---

[1] The assumption of the coercivity of the bilinear form $\mathcal{A}(\cdot, \cdot; \boldsymbol{\mu})$ over the kernel of $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ is still provable with Brezzi Theorem standard argument.

Chapter 4

Space-Time Proper Orthogonal Decomposition for Parametric Optimal Control Problems

This Chapter focuses on the POD Galerkin algorithm applied to time-dependent OCP($\boldsymbol{\mu}$)s [16, 146, 147, 150] in a space-time fashion. The *space-time POD* is a very versatile technique. Its strength relies in the capability to be adapted to a broad class of governing equations. After a brief description of the algorithm in Section 4.1, we will first present some preliminary results for linear time-dependent equations in Section 4.2. Thus, in Section 4.3, we will move towards a more complicated problem: a coastal management application governed by nonlinear viscous Shallow Waters Equations (SWEs). This model is more difficult to analyse, however we will show the advantages deriving from the space-time POD, even in this complex context.
We recall that the algorithm will be described making no distinction between time-dependency and steadiness. Indeed, the technique easily adapts to steady problems and the related numerical results are postponed in Chapters 6 and 7.

4.1 The Proper Orthogonal Decomposition

In this contribution, we will mainly rely on two different strategies to perform the building phase of the reduced framework for OCP($\boldsymbol{\mu}$)s: the POD [15, 30, 35, 64] and the Greedy algorithm [51, 64, 110, 111, 136]. This Section presents the first strategy, while we will address the latter in Chapter 5. The main feature of the POD is that it can be applied to any governing state equation: indeed, it does not require, as Greedy-based techniques, an error estimator. Such an estimator, for example, is still not available for nonlinear time-dependent OCP($\boldsymbol{\mu}$)s, i.e. the type of problem we are going to treat in Section 4.3. Let us start describing the algorithm and how it manipulates the snapshots in order to build the basis functions for the state, the adjoint and the control variables.
First of all, the POD-Galerkin algorithm collects $N_{\max}$ snapshots $\{\boldsymbol{X}(\boldsymbol{\mu}_i)\}_{i=1}^{N_{\max}}$ with $\boldsymbol{\mu}^i \in \mathcal{D}_{N_{\max}} \subset \mathcal{D}$. The set $\mathcal{D}_{N_{\max}}$ will be called the *training set*. Besides this *exploratory phase*, a compressing stage is performed, where $N < N_{\max}$ basis functions are built through snapshots manipulation getting rid of the redundant features of the system at hand, retaining the most important parametric information, only. Namely, the POD wants to compress the data related to *sampled*

manifold

$$\mathbb{M}^{\mathcal{N}}_{N_{\max}} := \{\boldsymbol{X}^{\mathcal{N}}(\boldsymbol{\mu}) \mid \boldsymbol{\mu} \in \mathcal{D}_{N_{\max}}\} \subset \mathbb{M}^{\mathcal{N}}.$$

It is clear that, when $N_{\max}$ is large enough, the sampled manifold $\mathbb{M}^{\mathcal{N}}_{N_{\max}}$ is a reliable representation of the high fidelity solution ensemble $\mathbb{M}^{\mathcal{N}}$. The numerical tests we are going to show in this thesis have been reduced through a *partitioned approach*: i.e. the POD algorithm have been separately exploited for each involved variable[1]. The partitioned procedure provides (at least) three different $N-$dimensional reduced spaces which minimizes the following quantities:

$$\sqrt{\frac{1}{N_{\max}} \sum_{\boldsymbol{\mu}\in\mathcal{D}_{N_{\max}}} \min_{\omega_N \in \mathbb{Y}^y_N} \|y^{\mathcal{N}_y}(\boldsymbol{\mu}) - \omega_N\|^2_{\mathbb{Y}}},$$

$$\sqrt{\frac{1}{N_{\max}} \sum_{\boldsymbol{\mu}\in\mathcal{D}_{N_{\max}}} \min_{\kappa_N \in \mathbb{U}_N} \|u^{\mathcal{N}_u}(\boldsymbol{\mu}) - \kappa_N\|^2_{\mathbb{U}}},$$

$$\sqrt{\frac{1}{N_{\max}} \sum_{\boldsymbol{\mu}\in\mathcal{D}_{N_{\max}}} \min_{\zeta_N \in \mathbb{Y}^z_N} \|z^{\mathcal{N}_y}(\boldsymbol{\mu}) - \zeta_N\|^2_{\mathbb{Y}}}.$$

For the sake of brevity, we will describe the data compression phase only for the state variable $y(\boldsymbol{\mu})$. However, the proposed techniques identically apply to the control and the adjoint variables as well. Let us consider the set of ordered parameters $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{N_{max}} \in \mathcal{D}_{N_{\max}}$. This parametric set is related to the ordered snapshots $y^{\mathcal{N}_y}(\boldsymbol{\mu}_1), \ldots, y^{\mathcal{N}_y}(\boldsymbol{\mu}_{N_{max}})$. Let $\mathbf{C}^y \in \mathbb{R}^{N_{max}\times N_{max}}$ be the correlation matrix of the snapshots of the state variable, i.e.:

$$\mathbf{C}^y_{ml} = \frac{1}{N_{max}}(y^{\mathcal{N}_y}(\boldsymbol{\mu}_m), y^{\mathcal{N}_y}(\boldsymbol{\mu}_l))_{\mathbb{Y}}, \qquad 1 \le m, l \le N_{max}.$$

First of all, the following eigenvalue problem is solved:

$$\mathbf{C}^y x^y_n = \lambda^y_n x^y_n, \qquad 1 \le n \le N,$$

where $\|x^y_n\|_{\mathbb{Y}} = 1$. Let us sort the eigenvalues $\lambda^y_1, \ldots, \lambda^y_{N_{\max}}$ in decreasing order and let us retain the first N ones, namely $\lambda^y_1, \ldots, \lambda^y_N$ together with the related eigenvectors $x^y_1, \ldots, x^y_N$. Here, $(x^y_n)_m$ represents the *m-th* component of the state eigenvector $x^y_n \in \mathbb{R}^{N_{max}}$. Once the compression process is terminated, the following relation between basis functions and eigenvalues-eigenvectors pairs holds:

$$\chi^y_n = \frac{1}{\sqrt{\lambda^y_n}} \sum_{m=1}^{N_{max}} (x^y_n)_m y^{\mathcal{N}_y}(\boldsymbol{\mu}_m), \qquad 1 \le n \le N. \tag{4.1}$$

[1]The main motivation for the use of three different PODs is given by the better performances in the errors between reduced and high fidelity solutions. Indeed, employing a partitioned approach combined with aggregated space technique is a common choice in many works in literature. The benefits of this strategies have been experimentally shown, for example, in [13, 80, 110, 110, 145, 168].

Equation (4.1) is well-known in data-compression-based algorithms, see e.g. [64, 125].

Remark 4.1.1 *As already specified in Chapter 3, in the presented context, no distinction between time-dependent and steady problems is made. Indeed, when dealing with time-dependent OCP($\boldsymbol{\mu}$)s, the time instances are not separated in the POD procedure. Namely, the space-time problem is treated as a steady one and each snapshot carries all the information about time evolution.*

It still remains to give a criterion to decide the values of N_{max} and N. The choice can be made performing an eigenvalue analysis of the correlation matrix. Indeed, the following relation holds [64, 125]:

$$\sqrt{\frac{1}{N_{\max}}\sum_{m=1}^{N_{\max}} \|y^{\mathcal{N}_y}(\boldsymbol{\mu}_m) - \Pi_N(y^{\mathcal{N}_y}(\boldsymbol{\mu}_m))\|_{\mathbb{Y}}^2} = \sqrt{\sum_{m=N+1}^{N_{\max}} \lambda_m^y}, \tag{4.2}$$

where $\Pi_N : \mathbb{Y} \to \mathbb{Y}_N$ projects functions in $\mathbb{Y}$ onto the state reduced space $\mathbb{Y}_N$. The right hand side of (4.2) gives information on how well the space is approximated using N basis over the chosen training set. A tricky aspect to take care of is the sampling strategy for the N_{max} parameters employed in the POD. The sampling can be guided by some previous knowledge about the system. This topic will be treated in detail in Chapter 6.
As pointed out in Chapter 3, the same high fidelity representation for the state and the adjoint variables does not guarantee the coincidence of the reduced state space and the reduced adjoint space merely applying a standard POD approach. Thus, we employed aggregated space technique as described in Section 3.2.2. Namely, the final spaces are:

$$\mathbb{Y}_N = \text{span } \{\chi_n^y, \chi_n^z,\ n = 1, \dots, N\}, \tag{4.3}$$

while for the control variable we rely on the standard space construction, i.e.

$$\mathbb{U}_N = \text{span}\{\chi_n^u,\ n = 1, \dots, N\}. \tag{4.4}$$

As already pointed out in Remark 4.1.1, the proposed POD-Galerkin technique is suited to both steady and time-dependent problems. However, in the following Sections, we will focus on the time-dependent cases and we will show some first numerical results to heuristically prove the advantages of exploiting POD for time-dependent nonlinear OCP($\boldsymbol{\mu}$)s.

4.2 Some Preliminary Numerical Results

In this Section we propose some first results on space-time POD-Galerkin applied to linear time-dependent OCP($\boldsymbol{\mu}$)s. We will focus on two test cases

- (*Section 4.2.1*) a boundary control for a Graetz flow;
- (*Section 4.2.2*) a distributed control for a Stokes cavity flow problem.

The two examples are the ones presented in [146]. They aim at paving the way to the application of the POD-Galrkin approach to more complicated equations, as we will see in Section 4.3.

4.2.1 The Graetz flow OCP($\boldsymbol{\mu}$)

Here, we present results concerning a time dependent OCP($\boldsymbol{\mu}$) governed by a Graetz flow under physical and geometrical parametrization. Indeed, the parameter is $\boldsymbol{\mu} = [\mu_1, \mu_2, \mu_3] \in \mathcal{D} = [1/20, 1/6] \times [1/2, 3] \times [1, 3]$. In this specific context, μ_3 will represent the geometrical parameter, while μ_1 and μ_2 denote, respectively, the diffusivity coefficient of the system and the temperature profile y_d we want to reach in the observation domain $\Omega_d(\mu_3) := [1, 1+\mu_3] \times [0, 0.2] \cup [1, 1+\mu_3] \times [0.8, 1]$. The whole physical domain Ω is shown in Figure 4.1. The control variable acts over the boundary $\Gamma_C = ([1, 1+\mu_3]\times\{0\})\cup([1, 1+\mu_3]\times\{1\})$. The portion of the boundary $\Gamma_D = (\{0\} \times [0, 1]) \cup ([0, 1] \times \{0\}) \cup ([0, 1] \times \{1\})$ is where Dirichlet boundary conditions have been applied, while, in $\Gamma_N(\mu_3) = \{1 + \mu_3\} \times [0, 1]$ Neumann boundary conditions apply. Namely, the value of μ_3 stretches the length of the subdomains $\Omega_2(\mu_3) := [1, 1 + \mu_3] \times [0.2, 0.8]$ and $\Omega_d(\mu_3)$. For this test case, we consider $y \in \mathcal{Y}_0$ with $Y = H^1_{\Gamma_D}(\Omega(\mu_3))$ and $u \in \mathcal{U}$ where $U = L^2(\Gamma_C(\mu_3))$. The PDE-constrained minimization problem reads: given $\boldsymbol{\mu} \in \mathcal{D}$, find the state-control variable $(y, u) \in \mathcal{Y}_0 \times \mathcal{U}$ such that:

$$\min_{\mathcal{Y}_0\times\mathcal{U}} \frac{1}{2}\int_0^T \int_{\Omega_3(\mu_3)} (y - y_d(\mu_2))^2 dxdt + \frac{\alpha}{2}\int_0^T \int_{\Gamma_C(\mu_3)} u^2 dxdt, \tag{4.5}$$

constrained to the equation

$$\begin{cases} y_t + \mu_1 \Delta y + x_2(1 - x_2)\dfrac{\partial y}{\partial x_1} = 0 & in\ \ \Omega(\mu_3) \times (0, T) \\ y = 1 & on\ \ \Gamma_D(\mu_3) \times (0, T), \\ \mu_1 \dfrac{\partial y}{\partial n} = u & on\ \ \Gamma_C(\mu_3) \times (0, T), \\ \mu_1 \dfrac{\partial y}{\partial n} = 0 & on\ \ \Gamma_N(\mu_3) \times (0, T), \\ y = y_0 & in\ \ \Omega(\mu_3) \times \{0\}. \end{cases} \tag{4.6}$$

Here, x_1 and x_2 are the spatial coordinates, y_0 is the null function in Ω which respects the boundary conditions, $T = 5$ and $y_d(\mu_2) \equiv \mu_2$ is the desired state field. The first step of the solution of this OCP(μ) relies in its trace back in the reference domain that correspond to $\mu_3 = 1$.

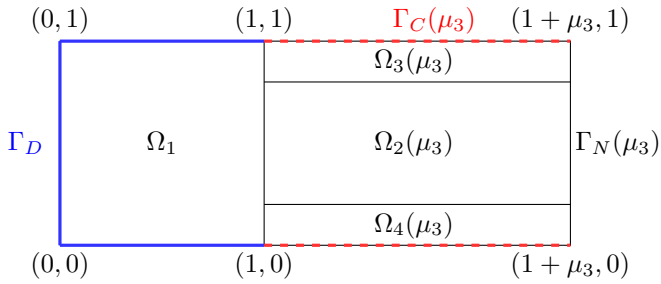


Figure 4.1: *Graetz flow, boundary control.* Domain Ω. *Observation domain:* $\Omega_d(\mu_3) = \Omega_3(\mu_3) \cup \Omega_4(\mu_3)$, *Control domain:* $\Gamma_C(\mu_3)$ (red dashed line). *Blue solid line:* Dirichlet boundary conditions. The reference domain Ω is given by $\mu_3 = 1$.

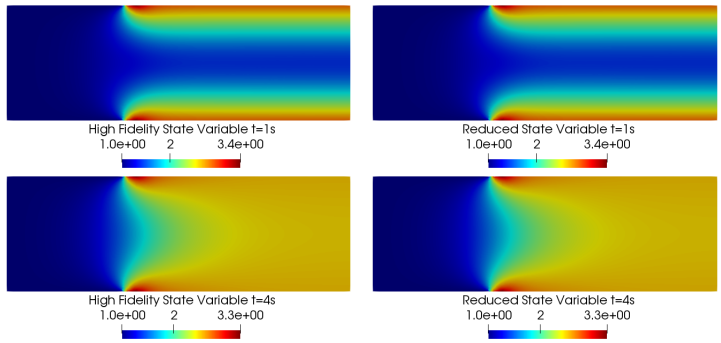


Figure 4.2: *Graetz flow, boundary control.* *Left.* FE state solutions for t=1s,4s for $\boldsymbol{\mu} = (12.0, 2.5, 2.0)$. *Right.* Reduced state solutions for t=1s,4s for $\boldsymbol{\mu} = (12.0, 2.5, 2.0)$.

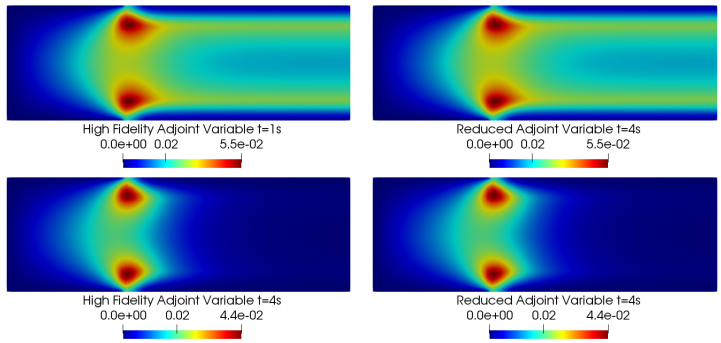


Figure 4.3: *Graetz flow, boundary control.* *Left.* FE adjoint solutions for t=1s,4s for $\boldsymbol{\mu} = (12.0, 2.5, 2.0)$. *Right.* Reduced adjoint solutions for t=1s,4s for $\boldsymbol{\mu} = (12.0, 2.5, 2.0)$.

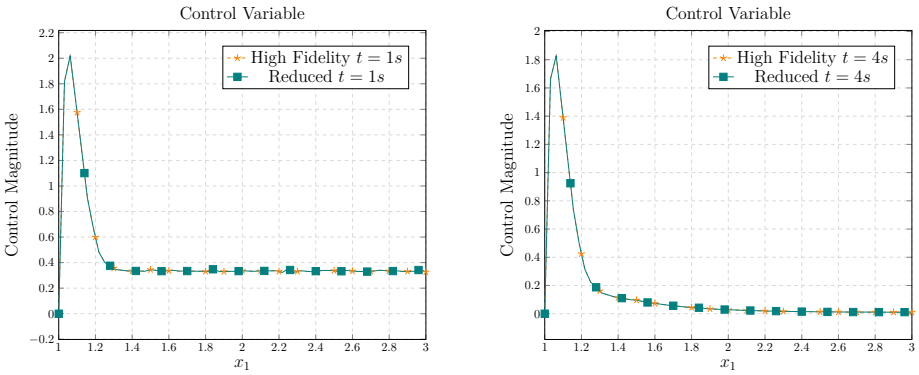


Figure 4.4: *Graetz flow, boundary control.* Comparison between high fidelity and reduced control variables for $\boldsymbol{\mu} = (12.0, 2.5, 2.0)$. *Left.* Solutions for $t = 1s$. *Right.* Solutions for $t = 4s$. In both cases, the two approaches coincide.

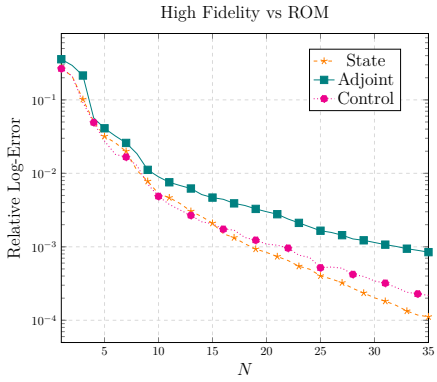


Figure 4.5: *Graetz flow, boundary control.* Averaged relative log-error for the variables.

The time step interval is given by $\Delta t = 1/6$ which divides $[0, T]$ in $N_t = 30$ subintervals. While the time discretization has been performed by means of Euler schemes, the FE spatial approximation involves $\mathbb{P}^1$ elements for all the variables. The global space-time dimension is $\mathcal{N} = 313'830$. The POD-Galerkin algorithm has been applied to compress the information of $N_{max} = 70$ snapshots[2]. We consider $N = 35$ basis functions for each variable, separately. Thus, after employing the space aggregation technique, the reduced system presents

[2] The number of collected snapshots was highly affected by the complexity of the space-time system, characterized by high computational costs.

a global dimension of $5N = 175$. The approach gives promising results, as can be noticed in Figures 4.2, 4.3 and 4.4 where some representative space-time solutions and ROM solutions are compared for different time instances and $\boldsymbol{\mu} = (12.0, 2.5, 2.0)$, for the state, the adjoint and the control variables, respectively. The space-time POD is able to recover the high fidelity fields, regardless the time value t_k taken into account. Furthermore, we studied the accuracy of the reduced model studying the behaviour of the log-error plot in Figure 4.5. The error has been averaged over a testing set of 50 uniformly distributed parameters. We notice that the POD-Galerkin approach is very effective for all the involved quantities reaching values around 10^{-4} for the state and the control variables, and around 10^{-3} for the adjoint variable. Furthermore, we conclude analysing the *speedup index*. With this terminology, we define the number of reduced systems that can be solved in the time of an high fidelity solution. In this case, the speedup is remarkable, reaching values around the order of $O(10^5)$. This is an experimental proof of the advantages given by the use of ROMs for time-dependent OCP($\boldsymbol{\mu}$)s. The next Section focuses on a classical example in fluid dynamics: an OCP($\boldsymbol{\mu}$) governed by Stokes equations. Also in this setting, we will show how convenient using ROM might be in order to study several parameterized configurations in a space-time fashion.

4.2.2 Cavity Viscous Flow OCP($\boldsymbol{\mu}$)

We now propose a time-dependent OCP($\boldsymbol{\mu}$) governed by Stokes equations in a cavity flow setting. The test case has been analysed in [146] and it is an extension of the numerical experiments presented in [65, 144] under physical and geometrical parametrization. Since we are dealing with Stokes equations, we exploited a standard FE $\mathbb{P}^2 - \mathbb{P}^1$ pair to deal with the state and adjoint variables, composed by the velocity field together with the pressure field.

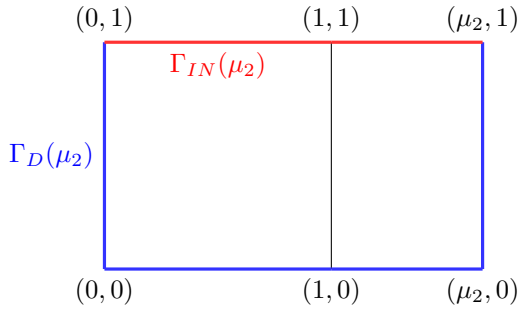


Figure 4.6: *Stokes flow, distributed control.* Domain Ω. The reference domain is given by $\mu_2 = 1$.

At the high fidelity level, the control variable has been discretized with $\mathbb{P}^2$ FE.

Recalling Section 2.3.2, with N^v_{FE}, N^p_{FE} and N^u_{FE} we indicate the FE function space dimensions for the state velocity and pressure fields (and by construction of the related adjoint variables) and for the control variable, respectively. For this specific test case, we obtain $N^v_{FE} = N^u_{FE} = 4554$ and $N^p_{FE} = 591$. The dimensions derive from a triangulation over the domain depicted in Figure 4.6. The considered problem deals with the minimization of the functional (1.45) with (1.46) as state equation: i.e. a velocity solution tracking problem governed by Stokes equations. We now summarize all the information about the numerical test. As already specified in Section 1.2.3, the value of μ_1 changes the diffusivity action of the system, while the parameter μ_2 affects the geometry of the spatial domain, stretching it. First of all, as we did for the previous test case, the domain is traced back into the two dimensional square $\Omega := (0,1) \times (0,1)$ as reference domain. The portion $\Gamma_{IN} = (0,1) \times \{1\}$ is the inlet boundary, while $\Gamma_D = \partial\Omega \backslash \Gamma_{IN}$ is where homogeneous Dirichlet boundary conditions are applied. Here, with the notation Γ_D and Γ_{IN} we refer to the traced back boundaries $\Gamma_D(\mu_2)$ and $\Gamma_{IN}(\mu_2)$, respectively. Since Dirichlet boundary conditions apply on the whole boundary $\partial\Omega$, we assumed $p \in \mathcal{P} = L^2(0,T; L^2_0(\Omega(\mu_2)))$, where $L^2_0(\Omega(\mu_2))$ is the space made by functions $p \in L^2(\Omega(\mu_2))$ which satisfy

$$\int_\Omega p \, d\Omega = 0, \tag{4.7}$$

in order to guarantee a unique solution for the pressure variable[3]. The parameter space is $\mathcal{D} := [10^{-3}, 10^{-1}] \times [1, 2.5]$, while the velocity profile we want to reach is the FE solution of the uncontrolled time-dependent Stokes equations endowed with an inlet constant velocity of components $(1,0)$ in Γ_{IN} and homogeneous Dirichlet boundary conditions in Γ_D, for $\mu_1 = 1$ fixed. The simulation gives the evolution of the target velocity in $(0,T) = (0,1)$, which is the time interval we are taking into consideration for this numerical test. We stress that for the OCP($\boldsymbol{\mu}$) at hand we used a different time-dependent inlet boundary condition, i.e. $v = (1 + 0.5\cos(4\pi t - \pi), 0)$ over Γ_{IN} for $t \in (0,T) = (0,1)$. The action of the distributed control u reduces the impact of the periodic inlet over the system. The time discretization is Euler-based with $\Delta t = 0.05$, i.e. $N_t = 20$. Thus, the global dimension of the optimality system is $\mathcal{N} = 296'880$. The dimensionality of the system has been reduced exploiting a partitioned POD-Galerkin approach over $N_{\max} = 70$ snapshots[4]. For each correlation matrix we built $N = 25$ basis functions. The global reduced dimension is $13N = 325$, taking into account supremizer stabilization combined to aggregated space technique applied to tackle the nested saddle point problem that arises in OCP($\boldsymbol{\mu}$)s governed by Stokes equations. A comparison between high fidelity and ROM state

[3] In the reduced model we employed aggregated space technique after a post-processing of the adjoint pressure variables. They have been made with null mean, since the reduced variable p_N still needs to verify the constraint (4.7). In this way, also the adjoint pressure variable has a null mean, even if not required by the problem formulation.

[4] See footnote 2.

velocity and state pressure profiles is depicted in Figures 4.7 and 4.8, respectively: the simulations coincide for different time instances for the representative parameter $\boldsymbol{\mu} = (10^{-2}, 1.5)$. For the sake of brevity, we did not show the control and the adjoint variables, however, the accuracy of the model is visible in plots in Figure 4.9, where the average relative log-error[5] w.r.t. N is presented over a testing set of 35 uniformly distributed parameters. The relative error is about 10^{-3} for all the involved variables. Also in this test case, the *speedup* reaches very good values. For $N = 25$ the value is around $6 \cdot 10^4$ thus, the POD-Galerkin approach could be very useful for this kind of system in a parametric setting.

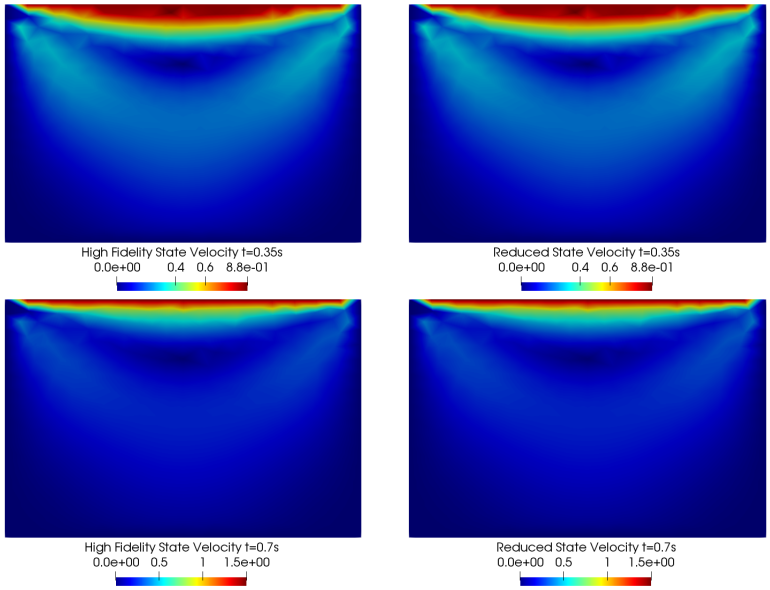


Figure 4.7: *Stokes flow, distributed control.* *Left.* FE state velocity solutions for t=0.35s,0.7s for $\boldsymbol{\mu} = (10^{-2}, 1.5)$. *Right.* Reduced state velocity solutions for t=0.35s,0.7s for $\boldsymbol{\mu} = (10^{-2}, 1.5)$.

[5]To make the high fidelity and ROM adjoint pressures comparable for the error analysis we define

$$z_p^{\mathcal{N}_y} := \bar{z}_p^{\mathcal{N}_y} - \int_\Omega \bar{z}_p^{\mathcal{N}_y} \, dx,$$

where $\bar{z}_p^{\mathcal{N}_y}$ is the high fidelity adjoint pressure, which is not required to verify the null mean constraint.

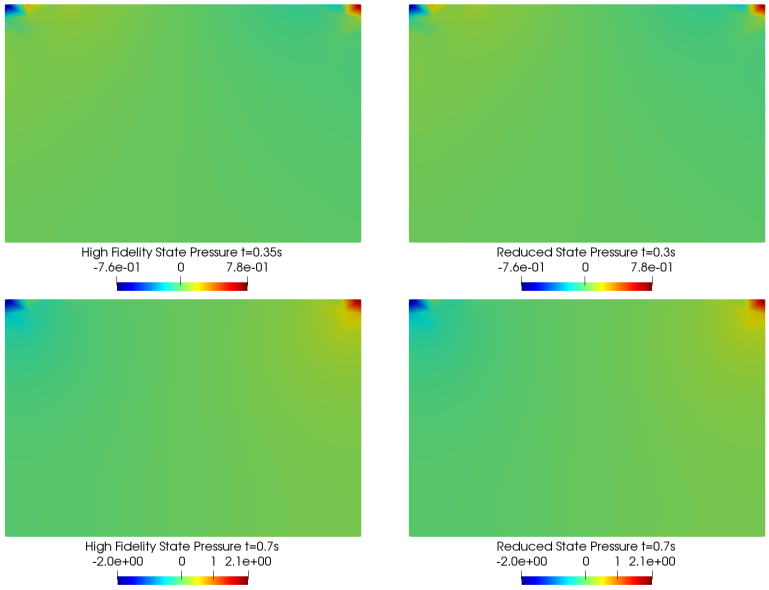


Figure 4.8: *Stokes flow, distributed control.* *Left*. FE state pressure solutions for t=0.35s,0.7s for $\boldsymbol{\mu} = (10^{-2}, 1.5)$. *Right*. Reduced state pressure solutions for t=0.35s,0.7s for $\boldsymbol{\mu} = (10^{-2}, 1.5)$.

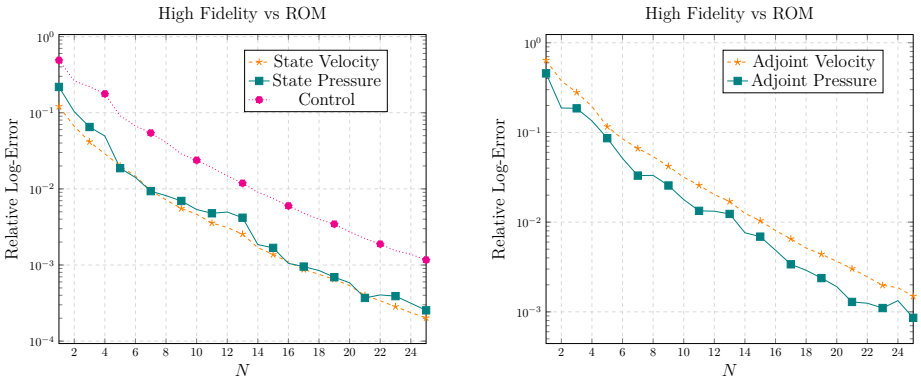


Figure 4.9: *Stokes flow, distributed control.* Averaged relative log-error for the variables.

The next Section concerns a nonlinear time-dependent problem and it will treat a coastal marine environmental application. Namely, these first results were meant to validate the space-time POD procedure in a simpler setting in order to apply it to more complicated problems, such as the one that follows.

4.3 The Nonlinear Viscous SWEs Model

We now want to apply the space-time POD-Galerkin methodology to a more involved problem: an OCP($\boldsymbol{\mu}$) governed by viscous SWEs. The model is an asset of growing impact in environmental sciences and, more specifically, in coastal management. This kind of scientific field needs to be highly supported by computational real-time simulations. The ultimate goal is to build an OCP($\boldsymbol{\mu}$) s based on such a model, able to forecast and eventually tackle possible dangerous phenomena. In this contribution, we will focus on the formulation of [147] and the numerical results presented in [16]. First of all, we are going to describe the model at hand in Section 4.3.1, remarking its usefulness in coastal sciences. Section 4.3.2 will test the space-time POD-Galerkin approach for a viscous SWE model that presents physical and geometrical parametrization.

4.3.1 Problem Formulation

The viscous SWEs are a very versatile model in the marine environmental field. Indeed, they have a remarkable impact on coastal engineering due to its capability to study not only the marine behaviour on shores and coasts, but also a wide range of marine phenomena such tsunamis waves, global currents dynamic... [34, 160]. The state equations together with its controlled version have been tackled analytically and numerically in many works, see e.g. [3, 4, 5, 6, 50, 105, 104, 103, 131, 132, 153]. However, few contributions have been appeared in a reduced setting. A reduction of the time evolution through POD has been exploited in [141, 142] for the uncontrolled state equation, while in [16, 147] a space-time POD-Galerkin approach is used to study several parametric instances in an optimal control framework, building on the standard two dimensional viscous model presented in [3, 105]. We now present the features of the parametric setting proposed by [16, 147]. Let $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4) \in \mathcal{D} \subset \mathbb{R}^4$ be a parameter that changes the physical behaviour of the system together with the geometry of the domain we are dealing with. Indeed, the physics will be represented by μ_1, μ_2 and μ_3, while the latter parameter will act on the space domain. We will describe them more specifically later on in the Section. For this specific test case, we define the following function spaces $Y_v = H^1_{\Gamma_{D_v}(\mu_4)}(\Omega(\mu_4))$, $Y_h = L^2(\Omega(\mu_4))$ and the space $U = L^2(\Omega(\mu_4))$. The notation $\Gamma_{D_v(\mu_4)}$ represents where Dirichlet boundary conditions have been imposed over the boundary $\partial\Omega(\mu_4)$. The variables of the state problem are:

- the vertically averaged velocity vector field $\boldsymbol{v}$;
- the scalar surface elevation variable h.

The authors in [16, 147] restrict themselves to the setting of constant bottom bathymetry z_b. However, z_b can represent more realistic bathymetry, such as

slopes for examples, see e.g [50, 132, 131]. Figure 4.10 shows a schematic description of the quantities we have just introduced.

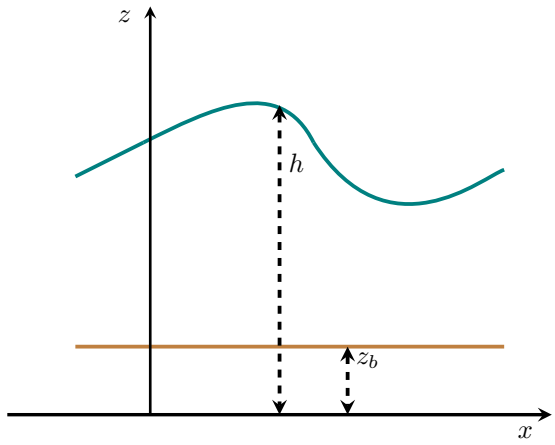


Figure 4.10: *SWEs model.* Notations: schematic representation.

The function spaces for the state velocity and the state height are, respectively:

$$\mathcal{Y}_v = \Big\{ \boldsymbol{v} \in L^2(0,T;[Y_v]^2) \text{ such that } \boldsymbol{v}_t \in L^2(0,T;[Y_v^*]^2) \Big\},$$

and

$$\mathcal{Y}_h = \Big\{ h \in L^2(0,T;Y_h) \text{ such that } h_t \in L^2(0,T;Y_h^*) \Big\}.$$

In this setting, the global state function space will be denoted by $\mathbb{Y} := \mathcal{Y}_v \times \mathcal{Y}_h$. Moreover, another function space is needed for the control variable: thus, let $\mathbb{U} := L^2(0,T;[U]^2)$ be the function space where the control variable will be sought. We present a distributed OCP($\boldsymbol{\mu}$): in this setting, you are not actually able to control the system since the forcing term represents natural quantities such as the bottom friction and the pressure on the surface of the fluid. However, the OCP($\boldsymbol{\mu}$) can be interpret as an inverse problem able to forecast the external physical conditions needed to reach a desired velocity-height observation $(\boldsymbol{v}_\mathrm{d}, h_\mathrm{d}) \in \mathbb{Y}_\mathrm{obs} := \mathcal{Q}_v \times \mathcal{Q}_h$, where $\mathcal{Q}_v := L^2(0,T;[L^2(\Omega(\mu_4))]^2)$ and $\mathcal{Q}_h := L^2(0,T;L^2(\Omega(\mu_4)))$. The constrained minimization problem reads: given

$\boldsymbol{\mu} \in \mathcal{D}$ find $(\boldsymbol{v}, h) \in \mathbb{Y}$ which minimizes $J((\boldsymbol{v}, h), \boldsymbol{u}, (\boldsymbol{v}_{\mathrm{d}}, h_{\mathrm{d}}))$, where

$$\begin{aligned} J((\boldsymbol{v}, h), \boldsymbol{u}; (\boldsymbol{v}_{\mathrm{d}}, h_{\mathrm{d}})) =& \frac{1}{2} \int_0^T \int_{\Omega(\mu_4)} (h - h_{\mathrm{d}}(\mu_3))^2 d\Omega(\mu_4) dt \\ &+ \frac{1}{2} \int_0^T \int_{\Omega(\mu_4)} (\boldsymbol{v} - \boldsymbol{v}_{\mathrm{d}}(\mu_3))^2 d\Omega(\mu_4) dt \\ &+ \frac{\alpha}{2} \int_0^T \int_{\Omega(\mu_4)} \boldsymbol{u}^2 d\Omega(\mu_4) dt, \end{aligned} \tag{4.8}$$

such that

$$\begin{cases} \boldsymbol{v}_t - \mu_1 \Delta \boldsymbol{v} + \mu_2 (\boldsymbol{v} \cdot \nabla) \boldsymbol{v} + g \nabla h - \boldsymbol{u} = 0 & \text{in } \Omega(\mu_4) \times (0, T), \\ h_t + \operatorname{div}(h\boldsymbol{v}) = 0 & \text{in } \Omega(\mu_4) \times (0, T), \\ \boldsymbol{v} = \boldsymbol{v}_0 & \text{on } \Omega(\mu_4) \times \{0\}, \\ h = h_0 & \text{on } \Omega(\mu_4) \times \{0\}, \\ \boldsymbol{v} = \boldsymbol{0} & \text{on } \partial\Omega(\mu_4) \times (0, T). \end{cases} \tag{4.9}$$

Remark 4.3.1 *Here, the velocity is such that* $\boldsymbol{v} \in L^2(0, T; [Y_v]^2)$ *and* $\boldsymbol{v}_t \in L^2(0, T; [{Y_v}^*]^2)$*, while* $h \in L^2(0, T; Y_h)$*,* $h_t \in L^2(0, T; {Y_h}^*)$ *and the control* $\boldsymbol{u} \in L^2(0, T; [U]^2)$*, where* $Y_h = U = L^2(\Omega(\mu_4))$ *and* $Y_v = H^1_{\Gamma_D}(\Omega(\mu_4))$*. Thus, since it holds*

$$Y_v \hookrightarrow U \hookrightarrow Y_v^*,$$

the first equation makes sense as a sum of terms in $L^2(0, T; [{Y_v}^*]^2)$*. Instead, the second equation can be interpret as a summation of terms in* $L^2(0, T; {Y_v}^*)$*. Indeed, exploiting standard regularity results for Navier-Stokes equations [39], h is actually in* $L^2(0, T; Y_v)$ *and thanks to the three-dimensional Sobolev embedding theorems, we have* $h\boldsymbol{v} \in L^2(0, T; [Y_h]^2)$ *and div*$(h\boldsymbol{v}) \in L^2(0, T; Y_v^*)$ *to due the inclusion* $Y_h^* \hookrightarrow Y_v^*$*.*

From now on, we will assume that the OCP($\boldsymbol{\mu}$) at hand is well-posed[6], The SWEs (4.9) represent the free surface of incompressible flows verifying hydrostatic pressure: this assumption holds when the water height h is much lower than the wavelength. This is a classical framework for coastal marine simulations.
For this specific test case, we have three physical parameters: μ_1 and μ_2 will represent diffusive and convective actions of the fluid while μ_3 will define the

[6]To the best of our knowledge the well-posedness of OCP($\boldsymbol{\mu}$)s governed by the two-dimensional SWEs has not been treated in a Lagrangian context in a parametric setting. We believe that the proof can be performed following the same techniques used in [67, Section 1.8.1] for Navier-Stokes equations.

desired solution profile. The geometry changes through μ_4. The problem can be solved exploiting the Lagrangian formulation described in 1.1.1. Here for the sake of brevity, we will report only the final optimality system in strong form. The reader interested in its variational formulation and in its detailed construction may refer to [147]. The problem reads: given $\boldsymbol{\mu} \in \mathcal{D}$, find $((\boldsymbol{v}, h), \boldsymbol{u}, (\boldsymbol{z}, q)) \in \mathbb{Y} \times \mathbb{U} \times \mathbb{Y}$ such that

$$\begin{cases} \boldsymbol{v} - \boldsymbol{z}_t - \mu_1 \Delta \boldsymbol{z} - \mu_2 (\boldsymbol{v} \cdot \nabla) \boldsymbol{z} \\ \qquad + \mu_2 (\nabla \boldsymbol{v})^T \boldsymbol{z} - h \nabla q = \boldsymbol{v}_d(\mu_3) & \text{in } \Omega(\mu_4) \times (0,T), \\ h - q_t - \boldsymbol{v} \cdot \nabla q - g \mathrm{div}(\boldsymbol{z}) = h_d(\mu_3) & \text{in } \Omega(\mu_4) \times (0,T), \\ \boldsymbol{z} = \boldsymbol{0} & \text{on } \partial\Omega(\mu_4) \times (0,T) \\ \boldsymbol{z} = \boldsymbol{0} & \text{on } \Omega(\mu_4) \times \{T\}, \\ q = 0 & \text{on } \Omega(\mu_4) \times \{T\}, \\ \alpha \boldsymbol{u} = \boldsymbol{z} & \text{in } \Omega(\mu_4) \times (0,T) \\ \boldsymbol{v}_t - \mu_1 \Delta \boldsymbol{v} + \mu_2 (\boldsymbol{v} \cdot \nabla) \boldsymbol{v} + g \nabla h = \boldsymbol{u} & \text{in } \Omega(\mu_4) \times (0,T), \\ h_t + \mathrm{div}(h \boldsymbol{v}) = 0 & \text{in } \Omega(\mu_4) \times (0,T), \\ \boldsymbol{v} = \boldsymbol{0} & \text{on } \partial\Omega(\mu_4) \times (0,T), \\ \boldsymbol{v} = \boldsymbol{v}_0 & \text{on } \Omega(\mu_4) \times \{0\}, \\ h = h_0 & \text{on } \Omega(\mu_4) \times \{0\}, \end{cases} \tag{4.10}$$

where the adjoint variable is represented by the pair $(\boldsymbol{z}, q)$.
The next Section will show the numerical results related to the optimality system (4.10) together with a brief description of the high fidelity and reduced strategies used for this specific test case.

4.3.2 Numerical Results

This Section presents the numerical results related to the application of a space-time POD-gelerkin approach on the test case discussed in [147] where the authors recast the experiments of [50] in the parametric setting addressed in Section 4.3.1. The parameter is considered in $\mathcal{D} = (0.00001, 1.) \times (0.01, 0.5) \times (0.1, 1.) \times (1., 1.5)$. Indeed, the spatial domain is given by $\Omega(\mu_4) = [0, 10\mu_4] \times [0, 10]$. For the sake of brevity, we do not report a Figure of the domain considered, since it is totally analogous to Figure 4.6, where the traced back domain is given for $\mu_4 = 1$. In this specific test case, we considered $z_b = 0$ and the wave time evolution is simulated in the time interval $[0, T]$ with $T = 0.8s$. Furthermore, let us indicate with x_1 and x_2 the spatial coordinates of the domain. We want to analyse an OCP($\boldsymbol{\mu}$) that aims at decreasing the effects of the impact of a mass of water spreading in the domain with an initial Gaussian distributed elevation under a null initial velocity: i.e.

$$\boldsymbol{v}_0 = \boldsymbol{0}, \quad \text{and} \quad h_0 = 0.2(1 + 5e^{(-(\frac{x_1}{\mu_4} - 5)^2 - (x_2 - 5)^2 + 1))}).$$

The optimality system at hand answers to this question: what are the forcing terms (say wind, bottom friction...) that let us reach a desired state $(\mu_3 \boldsymbol{v}_\mathrm{d}, \mu_3 h_\mathrm{d})$? Here, the desired state $(\boldsymbol{v}_\mathrm{d}, h_\mathrm{d})$ is the solution at time T of the uncontrolled state equation (4.9), with

$$\boldsymbol{v}_{\mathrm{d}0} = \mathbf{0}, \qquad \text{and} \qquad h_{\mathrm{d}0} = 2e^{(-(\frac{x_1}{\mu_4}-5)^2-(x_2-5)^2+1)},$$

and null external forces, i.e. $\boldsymbol{u} = \mathbf{0}$.
Focusing on the high fidelity approximation, we exploited linear polynomials for all the involved variables as proposed in [137]. The time approximation is Euler-based and we divided the time interval in $N_t = 8$ time steps[7], employing a $\Delta t = 0.1s$. The final high fidelity dimension is $\mathcal{N} = 76352$. Now, let us focus on the reduced approximation. Also in this case, we used a partitioned approach: five different POD data compression over a the correlation matrices of dimensions $N_{\max} = 100$ have been performed. We ended up with the following reduced spaces:

$$\begin{aligned}
\mathbb{Y}_N^v &= \mathrm{span}\{\chi_n^v,\ n = 1, \dots, N\},\\
\mathbb{Y}_N^h &= \mathrm{span}\{\chi_n^h,\ n = 1, \dots, N\},\\
\mathbb{U}_N &= \mathrm{span}\{\chi_n^u,\ n = 1, \dots, N\},\\
\mathbb{Y}_N^z &= \mathrm{span}\{\chi_n^z,\ n = 1, \dots, N\},\\
\mathbb{Y}_N^q &= \mathrm{span}\{\chi_n^q,\ n = 1, \dots, N\},
\end{aligned}$$

where we retained $N = 30$ basis functions for all the variables. Thus, after the application of the space aggregation, we obtain a global reduced space dimension of $9N = 270$. In this case, no supremizer stabilization is needed. The SWEs in its uncontrolled version is an hyperbolic system and the exploited space-time POD-Galerkin approach might give sub-optimal reduction results as pointed out in several papers, see e.g. [57, 152]. However, this specific test case seems not to present this case: first of all, it is a viscous SWEs model and, moreover, the OCP($\boldsymbol{\mu}$) framework changes the system in order to simulate a less convection-dominated phenomenon. The accuracy of the proposed space-time POD-Galerkin approach is tested by the relative log-errors in Figure 4.11. These errors are averaged over 20 uniformly distributed parameters: the use of $N = 30$ basis functions for each problem variable leads to values around 10^{-3} (it is slightly below for the control, the state and the adjoint elevation). For the sake of clarity, we show a comparison between some representative space-time solutions and their reduced counterpart in Figures 4.13, 4.12 and 4.14 for the

[7] As already specified in Chapter 2, the time discretization can be refined considering iterative techniques [65, 66, 143, 144], for example. Although, in this contribution, for the sake of simplicity, we always exploited a direct solver for (4.10).

state velocity, the state elevation and the control at $t = 0.1s, 0.4s, 0.8s$, respectively. The last aspect to take care of is the computational time needed for a reduced simulation.

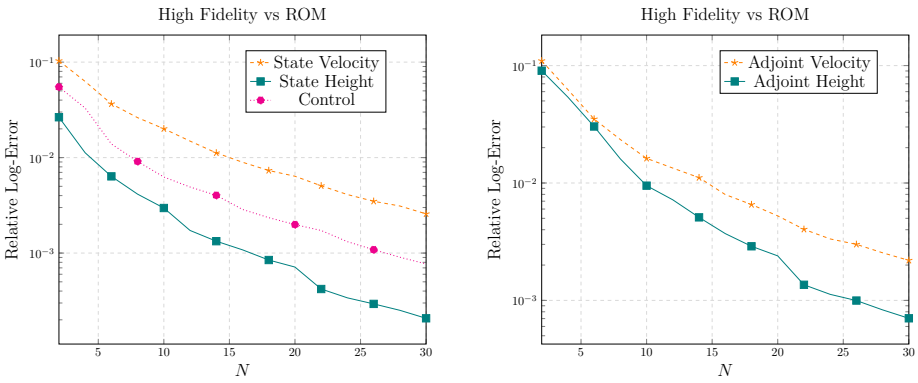


Figure 4.11: *SWEs model.* Averaged relative log-error for the variables.

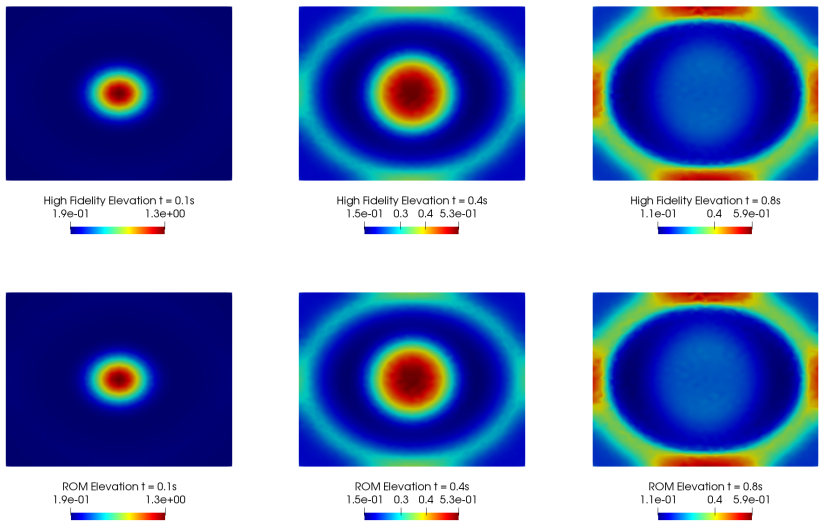


Figure 4.12: *SWEs model.* Optimal high fidelity and reduced state elevation variable with $\alpha = 0.1$ and $\boldsymbol{\mu} = (0.1, .01, .1, 1.5)$. *Top.* High fidelity solutions for $t = 0.1s, 0.4s, 0.8s$. *Bottom.* reduced solutions for $t = 0.1s, 0.4s, 0.8s$.

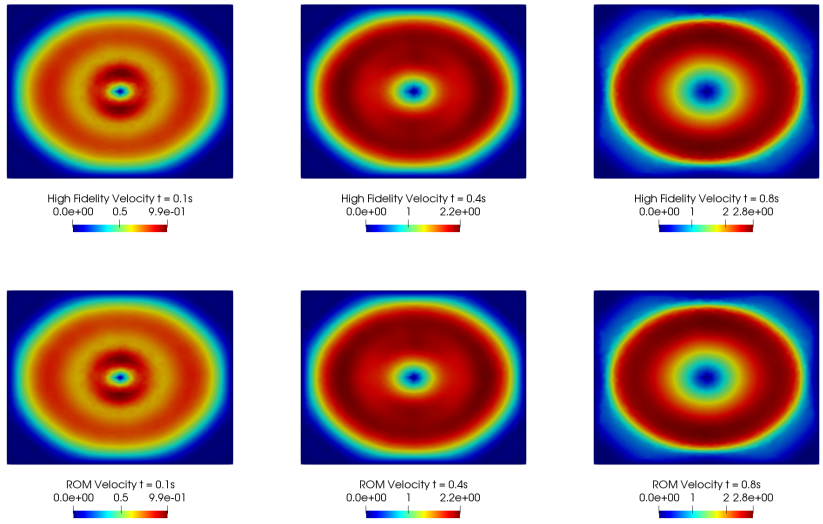


Figure 4.13: *SWEs model.* Optimal high fidelity and reduced state velocity variable with $\alpha = 0.1$ and $\boldsymbol{\mu} = (0.1, .01, .1, 1.5)$. *Top.* High fidelity solutions for $t = 0.1s, 0.4s, 0.8s$. *Bottom.* reduced solutions for $t = 0.1s, 0.4s, 0.8s$.

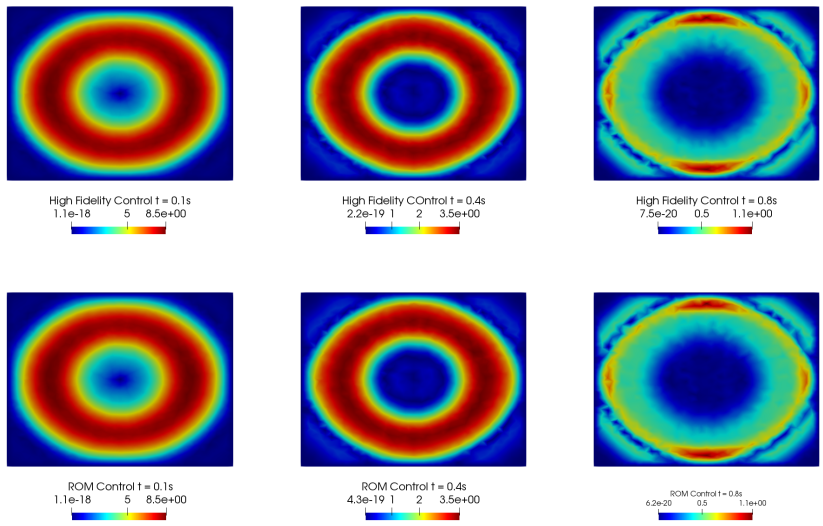


Figure 4.14: *SWEs model.* Optimal high fidelity and reduced control variable with $\alpha = 0.1$ and $\boldsymbol{\mu} = (0.1, .01, .1, 1.5)$. *Top.* High fidelity solutions for $t = 0.1s, 0.4s, 0.8s$. *Bottom.* reduced solutions for $t = 0.1s, 0.4s, 0.8s$.

In this context, the speedup is slightly affected by the value of N and it is of the order of $O(20)$, regardless the dimension of the reduced space[8].

The next Chapter will present the second typical strategy to build the reduced spaces: the Greedy algorithm. In this contribution, it will be addressed in a space-time fashion suited to parabolic state equations in the setting of OCP($\boldsymbol{\mu}$)s. We will follow the work of [148], where the authors propose a new error certification for time-dependent linear OCP($\boldsymbol{\mu}$)s. Then, we will move towards some numerical experiments to test the capability of the discussed strategy.

---

[8]The speedup had been influenced by the non-affine structure deriving from the nonlinearity of the state equation at hand. However, we are confident that better results can be reached exploiting the hyper-reduction techniques briefly introduced in Remark 3.1.1.

Chapter 5

A Certified Space-Time Reduced Basis Method for Linear Parametric Parabolic Optimal Control Problems

In this Chapter we discuss the use of a space-time Greedy approach for time-dependent linear OCP($\boldsymbol{\mu}$)s. We present the results of [148], where a new error estimator suited to both time-dependent and steady problems is provided for the no-control setting. In Section 5.1, we will recall some basic knowledge about ROMs, adapting them to the no-control framework. In the proposed setting, we will refer to ROMs also with the terminology Reduced Basis (RB). Section 5.2 shows some numerical tests that will validate the error certification for time-dependent and steady problems, in different parametric settings.

5.1 Error Certification for Space-Time Parabolic OCP($\boldsymbol{\mu}$)s

This Section focuses on the description of a Greedy-based reduction strategy and on the error certification needed in order to employ it in an efficient way. First of all, in Section 5.1.1, we will recall the main ideas about ROMs. Indeed, for the sake of clarity, we will adapt what was already introduced in Chapter 3 in the no-control framework. As a matter of fact, for the Greedy-based reduction, we will always work with this peculiar formulation, where the control variable can be obtained from the adjoint variable in postprocessing. The no-control setting will lead us to an explicit formulation of an error certification that does not depend on the high fidelity dimension of the global system. Naturally, this will lighten the building phase construction for space-time OCP($\boldsymbol{\mu}$)s.

5.1.1 Reduced Formulation for No-Control Problems

In Chapter 3 we highlighted these following main aspects of the ROMs:

- the need of representing the solution manifold $\mathbb{M}$ in a reliable way through reduced algorithms;
- the need of representing the solution manifold $\mathbb{M}$ in an efficient way though affine decomposition;

- the need of making provable the Brezzi Theorem at the reduced level thanks to space aggregation (and supremizer stabilization when dealing with a nested saddle point problem).

It is clear that in the no-control framework the fundamentals of the ROMs change due to the different formulation of the problem at hand. For the sake of clarity we will briefly report the new reduced structure we will deal with.
First of all, the solution manifold is of the following form

$$\mathbb{M} = \{(y(\boldsymbol{\mu}), z(\boldsymbol{\mu})) \mid \boldsymbol{\mu} \in \mathcal{D}\}.$$

Indeed, the global solution is only given by the state and the adjoint variables since the control variable can be recovered by the linear relation (1.56). Analogously the discrete solution manifold is

$$\mathbb{M}^{\mathcal{N}} = \{(y^{\mathcal{N}_y}(\boldsymbol{\mu}), z^{\mathcal{N}_y}(\boldsymbol{\mu})) \mid \boldsymbol{\mu} \in \mathcal{D}\}.$$

We now aim at approximating the behaviour of $\mathbb{M}^{\mathcal{N}}$ through a reduced space built by means of a Greedy approach. Here, $\mathcal{N}$ is the global dimension of the high fidelity solution, given by $\mathcal{N} := 2\mathcal{N}_y$. To work with faster simulations, we are looking for a surrogate low-dimensional space $\mathcal{Q}_N \times \mathcal{Q}_N \subset \mathcal{Q}^{\mathcal{N}_y} \times \mathcal{Q}^{\mathcal{N}_y} \subset \mathcal{Y}_0 \times \mathcal{Y}_T$. We recall that $\mathcal{Q}^{\mathcal{N}_y}$ is defined in Section 2.1.1 and represents the space-time approximation for both the state and the adjoint variables. Namely, in this context, we will specifically refer to time-dependent problems and we will mirror all the arguments for steady ones, explicitly.
The reduced optimality system we are going to solve reads: given $\boldsymbol{\mu} \in \mathcal{D}$, find the optimal pair $(y_N(\boldsymbol{\mu}), z_N(\boldsymbol{\mu})) \in \mathcal{Q}_N \times \mathcal{Q}_N$ such that

$$\underline{\mathcal{B}}((y_N(\boldsymbol{\mu}), z_N(\boldsymbol{\mu})), (\omega, \zeta); \boldsymbol{\mu}) = \langle \mathcal{F}(\boldsymbol{\mu}), (\omega, \zeta) \rangle \qquad \forall (\omega, \zeta) \in \mathcal{Q}_N \times \mathcal{Q}_N. \quad (5.1)$$

Once again, we underline that the convenience of the ROMs relies in an efficient division between an offline and an online phase, that separates the basis construction stage and the Galerkin projection for a new parametric instance. This is guaranteed when the affine decomposition holds for system (5.1), i.e. when the forms of the optimality system can be written as:

$$\underline{\mathcal{B}}((y, z), (\omega, \zeta); \boldsymbol{\mu}) = \sum_{l=1}^{Q_{\underline{\mathcal{B}}}} \Theta^l_{\underline{\mathcal{B}}}(\boldsymbol{\mu}) \underline{\mathcal{B}}^l((y, z), (\omega, \zeta)),$$

$$\langle \mathcal{F}(\boldsymbol{\mu}), (\omega, \zeta) \rangle = \sum_{l=1}^{Q_{\mathcal{F}}} \Theta^l_{\mathcal{F}}(\boldsymbol{\mu}) \langle \mathcal{F}^l, (\omega, \zeta) \rangle,$$

where $Q_{\underline{\mathcal{B}}}$ and $Q_{\mathcal{F}}$ are some integers, with $\Theta^l_{\underline{\mathcal{B}}}$, and $\Theta^l_{\mathcal{F}}$ $\boldsymbol{\mu}-$dependent smooth functions and $\underline{\mathcal{B}}^l$ and $\mathcal{F}^l$ are forms that do not depend on the parameter. We

recall that this structure allows us to assemble and store all the $\boldsymbol{\mu}-$independent quantities and then to exploit them in the online phase, where, for a given parameter, all the $\boldsymbol{\mu}$-dependent features are evaluated and the system (5.1) is assembled and solved.

5.1.2 Greedy Algorithm for OCP($\boldsymbol{\mu}$)s

This Section presents the space-time Greedy algorithm we used in order to build the reduced framework for parabolic OCP($\boldsymbol{\mu}$)s in the no-control setting. For a general introduction to this algorithm the reader may refer to [29, 64]. This Section aims at extending it in a space-time fashion for parabolic OCP($\boldsymbol{\mu}$)s. Indeed, time-dependent OCP($\boldsymbol{\mu}$)s have been tackled most of all with the POD-Greedy algorithm [61, 64] in order to compress the data information given by the several time instances. Although, another approach can be employed: namely, the space-time setting already exploited for parabolic problems in [156, 165, 166] can be adapted to time-dependent linear parabolic OCP($\boldsymbol{\mu}$)s (and we will specify it for steady ones).
The Greedy algorithm iteratively builds the basis functions of the reduced framework: the main idea is to enrich the bases with new information given by suitably chosen snapshots. At each step, a high fidelity solution of the time-dependent (or steady) optimality system is computed: namely, to build a complete $N-$dimensional reduced function space, N space-time (or steady) optimal solution must be evaluated. In order to introduce the space-time Greedy algorithm, we define the global error e between the space-time optimality solution and the reduced one, i.e.

$$e := (y^{\mathcal{N}_y} - y_N, z^{\mathcal{N}_y} - z_N). \tag{5.2}$$

The Greedy algorithm is based on an *estimate* of the norm of the global error (5.2) which is independent from the high fidelity dimension of the system[1]: to this end, we define a $\mathcal{N}-$independent quantity $\Delta_N(\boldsymbol{\mu})$ such that

$$\|e\|_{\mathcal{Q}\times\mathcal{Q}} \leq \Delta_N(\boldsymbol{\mu}). \tag{5.3}$$

Let us assume to have been provided of a such an estimator[2]. The first step of the space-time Greedy algorithm is to choose a finite set $\mathcal{D}_h \subset \mathcal{D}$ consisting in $N_{\max}$ parameters. A large cardinality of $\mathcal{D}_h$ allows a good representation of the discrete solution manifold $\mathbb{M}^{\mathcal{N}}$. As already said, the reduced space is built through an iterative procedure. Thus, we consider a tolerance τ and we define the first-step aggregated reduced space for state and adjoint variables as $\mathcal{Q}_N = \text{span}\{y^{\mathcal{N}_y}(\boldsymbol{\mu}_0), z^{\mathcal{N}_y}(\boldsymbol{\mu}_0)\}$ for an initial parametric value $\boldsymbol{\mu}_0$. The $n-$th

[1] Here, we are assuming that the space-time structure (or the pure FE discretization) is a good approximation of the continuous optimal solution, in order to directly consider the error (5.2) to represent the accuracy of the reduced model.
[2] The explicit expression of $\Delta_N(\boldsymbol{\mu})$ is postponed in the next Section.

step of the algorithm chooses the parameter

$$\boldsymbol{\mu}_n = \arg\max_{\boldsymbol{\mu}\in\mathcal{D}_h} \Delta_N(\boldsymbol{\mu}), \tag{5.4}$$

and, then, the aggregated reduced space is enriched with the snapshots evaluated in this chosen parameter $\boldsymbol{\mu}_n$, i.e.

$$\mathcal{Q}_N = \text{span}\{y^{\mathcal{N}_y}(\boldsymbol{\mu}_0), \dots, y^{\mathcal{N}_y}(\boldsymbol{\mu}_n), z^{\mathcal{N}_y}(\boldsymbol{\mu}_0) \dots, z^{\mathcal{N}_y}(\boldsymbol{\mu}_n)\}.$$

The iterative process stops when a selected parameter verifies $\Delta_N(\boldsymbol{\mu}) \leq \tau$. We assume that the last step is the $N-$th step. We remark that the aggregated space technique guarantees the well-posedness of the optimality system. Indeed, also in the no-control framework, the reduced inf-sup stability constant $\beta_N(\boldsymbol{\mu})$ is defined as:

$$\inf_{(y_N,z_N)\in\mathcal{Q}_N\times\mathcal{Q}_N} \sup_{(\omega_N,\zeta_N)\in\mathcal{Q}_N\times\mathcal{Q}_N} \frac{\underline{\mathcal{B}}((y_N,z_N),(\omega_N,\zeta_N);\boldsymbol{\mu})}{\sqrt{\|y_N\|^2_{\mathcal{Q}}+\|z_N\|^2_{\mathcal{Q}}}\sqrt{\|\omega_N\|^2_{\mathcal{Q}}+\|\zeta_N\|^2_{\mathcal{Q}}}},$$

for $(y_N, z_N) \neq 0$ and $(\omega_N, \zeta_N) \neq 0$. The problem is well-posed when there exists a positive constant $\hat{\beta}_N(\boldsymbol{\mu})$ such that $\beta_N(\boldsymbol{\mu}) > \hat{\beta}_N(\boldsymbol{\mu})$. In this specific context, the reduced framework has a global dimension of $4N$ since $\mathcal{Q}_N$ has dimension $2N$ and it is considered for both the variables. Although, we will show that using such a enriched space is still convenient in terms of the computational speedup. In the next Section we propose an explicit formulation for the error estimator $\Delta_N(\boldsymbol{\mu})$, that can be used both for time dependent and steady OCP($\boldsymbol{\mu}$)s.

5.1.3 Rigorous a posteriori error estimate

The Greedy-based approaches rely on an a posteriori error estimate in order to build a reliable reduced space. Indeed, the error estimation provides:

- a bound for the sampling strategy over the parametric space $\mathcal{D}$ in the bases construction stage;
- for every $\boldsymbol{\mu} \in \mathcal{D}$, a bound for the error between the high fidelity and the reduced solution in the online stage.

We will define an a posteriori error estimate related to the classical Brezzi and Nečas-Babuška stability analysis [11, 27, 108], as presented in [148] where seminal results in the context of steady OCP($\boldsymbol{\mu}$)s, see e.g. [110, 111], have been generalized to parabolic OCP($\boldsymbol{\mu}$)s.
Namely, the goal is to explicit a formulation for an $\mathcal{N}-$independent, and thus, fast to compute, a posteriori error bound $\Delta_N(\boldsymbol{\mu})$ such that:

$$\|e\|_{\mathcal{Q}\times\mathcal{Q}} = \sqrt{\|y^{\mathcal{N}_y} - y_N\|^2_{\mathcal{Q}} + \|z^{\mathcal{N}_y} - z_N\|^2_{\mathcal{Q}}} \leq \Delta_N(\boldsymbol{\mu}). \tag{5.5}$$

Considering the space-time high fidelity approximation, we have that the discrete inf-sup condition related to the whole optimality system (2.5) is verified and, as a consequence of the Nečas- Babuška theorem, we have the following stability estimate for the global solution:

$$\sqrt{\|y^{\mathcal{N}_y}\|^2_{\mathcal{Q}} + \|z\|^{\mathcal{N}_y\,2}_{\mathcal{Q}}} \leq \frac{1}{\beta^{\mathcal{N}}_{\underline{\mathcal{B}}}(\boldsymbol{\mu})} \|\mathcal{F}(\boldsymbol{\mu})\|_{(\mathcal{Q}\times\mathcal{Q})^*}, \tag{5.6}$$

where $\beta^{\mathcal{N}}_{\underline{\mathcal{B}}}(\boldsymbol{\mu})$ is the discrete inf-sup constant of the whole optimality system, as defined in (2.5). In the expression of the error estimator, besides the inf-sup constant, another crucial element is the dual norm of the optimality system residual $\mathcal{R} \in (\mathcal{Q}^{\mathcal{N}_y} \times \mathcal{Q}^{\mathcal{N}_y})^*$

$$\mathcal{R}((\omega,\zeta);\boldsymbol{\mu}) = \underline{\mathcal{B}}((y_N,\zeta_N),(\omega,\zeta);\boldsymbol{\mu}) - \langle \mathcal{F}(\boldsymbol{\mu}),(\omega,\zeta)\rangle, \tag{5.7}$$

for all $(\omega,\zeta) \in \mathcal{Q}^{\mathcal{N}_y} \times \mathcal{Q}^{\mathcal{N}_y}$. By definition and by the stability estimate of the Nečas-Babuška theorem, it is a matter of simple computations to show that:

$$\|e\|_{\mathcal{Q}\times\mathcal{Q}} \leq \frac{\|\mathcal{R}\|_{(\mathcal{Q}\times\mathcal{Q})^*}}{\beta^{\mathcal{N}}_{\underline{\mathcal{B}}}(\boldsymbol{\mu})} \quad \forall \boldsymbol{\mu} \in \mathcal{D},$$

since

$$\underline{\mathcal{B}}(e,(\omega,\zeta);\boldsymbol{\mu}) = \mathcal{R}((\omega,\zeta);\boldsymbol{\mu}). \tag{5.8}$$

In order to reliably and efficiently apply such an estimation, a practical way to compute the inf-sup constant $\beta^{\mathcal{N}}_{\underline{\mathcal{B}}}(\boldsymbol{\mu})$ is needed. Indeed, let us assume to have been provided by a lover bound $\beta^{LB}(\boldsymbol{\mu}) > 0$ such that $\beta^{\mathcal{N}}_{\underline{\mathcal{B}}}(\boldsymbol{\mu}) \geq \beta^{LB}(\boldsymbol{\mu})$. Then, we can estimate the error as:

$$\|e\|_{\mathcal{Q}\times\mathcal{Q}} \leq \frac{\|\mathcal{R}\|_{(\mathcal{Q}\times\mathcal{Q})^*}}{\beta^{LB}(\boldsymbol{\mu})} := \Delta_N(\boldsymbol{\mu}) \qquad \forall \boldsymbol{\mu} \in \mathcal{D}, \qquad \forall N = 1,\dots,N_{\max}. \tag{5.9}$$

In [148] an explicit form for the lower bound is discussed. We report the main results and the related proof. All the involved quantities have been defined in Chapter 1, where the parabolic problem is considered at the continuous level. However, for the sake of clarity, we will report them again:

- $\gamma_a(\boldsymbol{\mu})$: the coercivity constant related to the state equation $a(\cdot,\cdot;\boldsymbol{\mu})$;
- $c_c(\boldsymbol{\mu})$: the continuity constant of the control bilinear form $c(\cdot,\cdot;\boldsymbol{\mu})$;
- $c_m(\boldsymbol{\mu})$: the continuity constant of the bilinear form $m(\cdot,\cdot;\boldsymbol{\mu})$;
- c_u and c_{obs}: the embedding constants between Y and U and between Y and Y_{obs}, respectively defined in (1.63) and (1.62);

Theorem 5.1 *Let us suppose that a space-time OCP($\boldsymbol{\mu}$) governed by a parabolic equation is well-posed. Then, there exists a lower bound $\beta^{LB}(\boldsymbol{\mu})$ for $\beta^{\mathcal{N}}_{\underline{\mathcal{B}}}(\boldsymbol{\mu})$ of the following form:*

$$\begin{cases} \alpha\gamma_a(\boldsymbol{\mu}) & \text{for} \quad \Omega_u = \Omega_{obs}, \\ \dfrac{\gamma_a(\boldsymbol{\mu})}{\sqrt{2\max\left\{1, \left(\frac{c_c(\boldsymbol{\mu})c_u}{\alpha\gamma_a(\boldsymbol{\mu})}\right)^2\right\}}} & \Omega_u \neq \Omega_{obs} \text{ assuming } \Omega_{obs} \neq \Omega, \\ \dfrac{\gamma_a(\boldsymbol{\mu})}{\sqrt{2\max\left\{1, \left(\frac{c_m(\boldsymbol{\mu})c_{obs}}{\alpha\gamma_a(\boldsymbol{\mu})}\right)^2\right\}}} & \Omega_u \neq \Omega_{obs} \text{ assuming } \Omega_u \neq \Omega. \end{cases} \tag{5.10}$$

Proof. The statement is a consequence of Theorem 2.2 and the stability estimate (5.6), applied to the problem (5.8), where the lower bounds are given in the proofs of Lemma 2.1 and Remark 1.3.2. In this theorem, we are recovering all the cases: indeed, at least, one between control and observation domain must be different from Ω to be $\Omega_u \neq \Omega_{\text{obs}}$.

Since the quantities involved in the lower bound $\beta^{LB}(\boldsymbol{\mu})$ are all known and $\mathcal{N}-$independent, it results fast to be computed for a given parameter $\boldsymbol{\mu} \in \mathcal{D}$. Furthermore, the dual norm of the residual can be rapidly evaluated exploiting the affine assumption by means of suitable Riesz representers [110, 111, 134].

Remark 5.1.1 (The steady case) *The arguments concerning the lower bound for the error estimation can be also specified to steady linear OCP($\boldsymbol{\mu}$)s. In the steady setting, the reduced problem reads: given $\boldsymbol{\mu} \in \mathcal{D}$, find the optimal pair $(y_N(\boldsymbol{\mu}), z_N(\boldsymbol{\mu})) \in Y_N \times Y_N$ such that*

$$\underline{\mathcal{B}}_s((y_N(\boldsymbol{\mu}), z_N(\boldsymbol{\mu})), (\omega, z), \boldsymbol{\mu}) = \langle \mathcal{F}_s(\boldsymbol{\mu}), (\omega, \zeta)\rangle \qquad \forall (\omega, \zeta) \in Y_N \times Y_N, \tag{5.11}$$

with $Y_N \subset Y^{N^y_{FE}}$. Also in this case, there exists a positive inf-sup steady stability constant $\beta^{\mathcal{N}}_{\underline{\mathcal{B}}_s}(\boldsymbol{\mu})$ and exploiting the affine decomposition of the system (5.11), one can efficiently apply the Greedy algorithm thanks to this relation [111]:

$$\|e\|_{Y\times Y} \leq \frac{\|\mathcal{R}\|_{(Y\times Y)^*}}{\beta^{LB}_s(\boldsymbol{\mu})} \qquad \forall \boldsymbol{\mu} \in \mathcal{D}.$$

As a special case, we derive a novel lower bound $\beta^{LB}_s(\boldsymbol{\mu})$ for the steady framework [148]: the bound avoids the successive constraint methods [69] to approximate the inf-sup stability constant related to the optimality system (5.11). Indeed, the lower bound is easy to be computed and, applying Lemma 1.9 and Theorem 2.3,

it has the following form:

$$\begin{cases} \alpha\gamma_a(\boldsymbol{\mu}) & \text{for} \quad \Omega_u = \Omega_{obs}, \\ \dfrac{\gamma_a(\boldsymbol{\mu})}{\sqrt{2\max\left\{1, \left(\frac{c_c(\boldsymbol{\mu})c_u}{\alpha\gamma_a(\boldsymbol{\mu})}\right)^2\right\}}} & \Omega_u \neq \Omega_{obs} \text{ assuming } \Omega_{obs} \neq \Omega. \\ \dfrac{\gamma_a(\boldsymbol{\mu})}{\sqrt{2\max\left\{1, \left(\frac{c_c(\boldsymbol{\mu})c_{obs}}{\alpha\gamma_a(\boldsymbol{\mu})}\right)^2\right\}}} & \Omega_u \neq \Omega_{obs} \text{ assuming } \Omega_u \neq \Omega. \end{cases} \tag{5.12}$$

In the next Section, the error estimators we have provided will be tested both for steady and time-dependent OCP($\boldsymbol{\mu}$)s governed by Graetz flows in a distributed and boundary control settings.

5.2 OCP($\boldsymbol{\mu}$)s governed by Graetz flow

This Section validates the error estimator bounds for parabolic and steady OCP($\boldsymbol{\mu}$)s in two different parametric settings. The first one is characterized by physical parameters only, while the second one presents both physical and geometrical parametrization. Here, we report the numerical tests of [148], which are inspired by [80, 111, 146]. Moreover, both the parametric frameworks are presented in a time-dependent and steady version. We consider $Y = H_0^1(\Omega_{\Gamma_D})$, i.e. as the space of H^1 functions over Ω that vanish on Γ_D, the portion of the boundary $\partial\Omega$ where Dirichlet boundary conditions apply. The observation space is $Y_{\text{obs}} = L^2(\Omega_{\text{obs}})$. Furthermore, x_1 and x_2 denote the spatial coordinates. In both the considered setting we will compare the employment of the explicit formulation of the error estimators (5.10) w.r.t. the employment of the exact Babuška inf-sup condition (2.5). The goal is to show how convenient using such a technique would be in order to deal with a very effective reduced setting built through a lightened offline phase based on a Greedy procedure.

5.2.1 Physical Parametrization

Let us focus on OCP($\boldsymbol{\mu}$)s governed by a Graetz flow with physical parametrization only. The physical domain is Ω, depicted in Figure (5.1), where the observation domain is $\Omega_{\text{obs}} = \Omega_1 \cup \Omega_2$, with $\Omega_1 = [0.2, 0.8] \times [0.3, 0.7]$ and $\Omega_2 = [1.2, 2.5] \times [0.3, 0.7]$- In this specific case, $\Omega_u = \Omega$, i.e. the control is distributed all over the physical domain. The parameter $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)$ is considered in $\mathcal{D} = [3, 20] \times [0.5, 1.5] \times [1.5, 2.5]$: the first parameter μ_1 is the Péclet number of the governing advection-diffusion state equation, while μ_2 and μ_3 are constants representing the desired state y_{d} we want to reach in the subdomains Ω_1 and Ω_2, respectively. The problem formulation changes a heat source in order to achieve a parametric desired configuration. The optimality system reads:

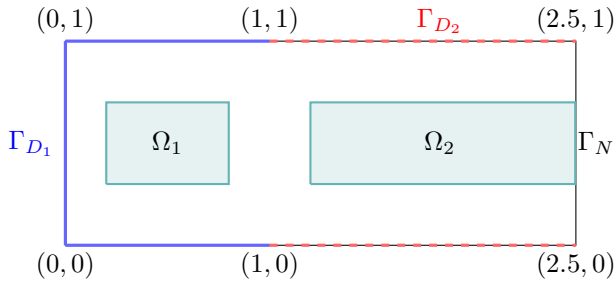


Figure 5.1: *Error Certification for parabolic problems.* Domain Ω. *Observation domain:* $\Omega_{\text{obs}} = \Omega_1 \cup \Omega_2$, *Control domain:* Ω. *Blue solid line:* first Dirichlet boundary conditions. *Red dashed line:* second Dirichlet boundary conditions.

given $\boldsymbol{\mu} \in \mathcal{D}$, find the optimal pair $(y, z) \in \mathcal{Y}_0 \times \mathcal{Y}_T$ such that

$$
\begin{cases}
y(\chi_{\Omega_1} + \chi_{\Omega_2}) - z_t \\
\qquad -\dfrac{1}{\mu_1}\Delta z - x_2(1-x_2)\dfrac{\partial z}{\partial x_1} = \mu_2\chi_{\Omega_1} + \mu_3\chi_{\Omega_2} & \text{in } \Omega \times (0,T), \\
y_t - \dfrac{1}{\mu_1}\Delta y - x_2(1-x_2)\dfrac{\partial y}{\partial x_1} - \dfrac{1}{\alpha}z = 0 & \text{in } \Omega \times (0,T), \\
y(0) = y_0 & \text{in } \Omega, \\
z(T) = 0 & \text{in } \Omega, \\
\dfrac{1}{\mu_1}\dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N \times (0,T), \\
y = 1 \text{ and } z = 0 & \text{on } \Gamma_{D_1} \times (0,T), \\
y = 2 \text{ and } z = 0 & \text{on } \Gamma_{D_2} \times (0,T),
\end{cases}
\tag{P}
$$

where y_0 is a null function verifying the boundary conditions, $\Gamma_{D_1} = \partial\Omega \cap \{(x_1, x_2) \mid x_1 \leq 1\}$ and $\Gamma_{D_2} = \partial\Omega \cap \{(x_1, x_2) \mid 1 < x_1 \leq 2.5\}$. Calling $\Gamma_D := \Gamma_{D_1} \cup \Gamma_{D_2}$, then $\Gamma_N = \partial\Omega \setminus \Gamma_D$ will represent where Neumann boundary conditions have been applied: here, n is the normal outer vector w.r.t. the portion of the boundary Γ_N. We now propose problem (P) in its steady version: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, z) \in Y \times Y$ such that

$$
\begin{cases}
y(\chi_{\Omega_1} + \chi_{\Omega_2}) - \dfrac{1}{\mu_1}\Delta z - x_2(1-x_2)\dfrac{\partial z}{\partial x_1} = \mu_2\chi_{\Omega_1} + \mu_3\chi_{\Omega_2} & \text{in } \Omega, \\
-\dfrac{1}{\mu_1}\Delta y - x_2(1-x_2)\dfrac{\partial y}{\partial x_1} - \dfrac{1}{\alpha}z = 0 & \text{in } \Omega, \\
\dfrac{1}{\mu_1}\dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N, \\
y = 1 \text{ and } \ = 0 & \text{on } \Gamma_{D_1}, \\
y = 2 \text{ and } z = 0 & \text{on } \Gamma_{D_2}.
\end{cases}
\tag{P_s}
$$

From the formulations (P) and (P_s), it is straightforward to verify the affinity assumption. As already specified, the problem is a distributed control problem, i.e. $\Omega_u = \Omega$ and thus, following the bounds in (5.10), we must consider

$$\beta^{LB}(\boldsymbol{\mu}) = \beta_s^{LB}(\boldsymbol{\mu}) = \frac{\gamma_a(\boldsymbol{\mu})}{\sqrt{2\max\left\{1, \left(\frac{c_c(\boldsymbol{\mu})c_u}{\alpha\gamma_a(\boldsymbol{\mu})}\right)^2\right\}}}. \tag{5.13}$$

However, we still need to specify the constants involved in the bound, which are problem dependent. First of all, for this specific state equation it holds:

$$\gamma_a(\boldsymbol{\mu}) := \frac{1}{\mu_1(1 + C_\Omega^2)}, \tag{5.14}$$

where C_Ω is the Poincaré constant which verifies $\|v\|_{L^2(\Omega)} \leq C_\Omega \|v\|_{H^1(\Omega)}$, for all $v \in H^1(\Omega)$, see e.g. [128]. Furthermore, for the problem at hand, $c_u = C_\Omega$. It remains to specify $c_c(\boldsymbol{\mu})$, that, actually, does not depend on $\boldsymbol{\mu}$ and, thus, we can write $c_c := c_c(\boldsymbol{\mu}) = C_\Omega$, indeed:

$$|c(z, y; \boldsymbol{\mu})| = \left|\int_\Omega zy \, d\Omega\right| \leq \|z\|_{L^2(\Omega)}\|y\|_{L^2(\Omega)} \leq C_\Omega \|z\|_U \|y\|_Y. \tag{5.15}$$

The value of C_Ω is parameter independent and it has been pre-computed solving the related eigenvalue problem. The computational costs of this procedure is affordable since the constant can be evaluated only once in the offline phase.
For both the test cases, we choose $\mathcal{D}_h$ as a set of $N_{\max} = 225$ parameters uniformly distributed in $\mathcal{D}$ and we considered $\tau = 10^{-4}$ as a tolerance for the Greedy algorithm. The performances of the error certification has been shown through an average error analysis and average effectivity analysis over 100 uniformly distributed parameters in $\mathcal{D}$. Furthermore, at the spatial discrete level, the pair (y, z) is represented by $\mathbb{P}^1 - \mathbb{P}^1$ FE pair, while the time discretization is Euler-based and performed over the time interval $[0, 5]$ with $\Delta t = 1/6$. This results into $N_t = 30$ time steps, making the tests comparable with the results of [146], presented in Section 4.2.1. We start by presenting the time-dependent results.

The time-dependent case

We test the performances of the lower bound $\beta^{LB}(\boldsymbol{\mu})$ to problem (P), with fixed $\alpha = 0.01$. The space-time Greedy algorithm reached the chosen tolerance τ after $N = 13$ steps and it results, applying aggregated spaces technique, into a reduced space of dimension $4N = 52$. The reduced dimension is much smaller w.r.t. the high fidelity one, of dimension $\mathcal{N} = 2N_{FE}^y \cdot N_t = 272'160$. This difference in the reduced and high fidelity dimensions gives impressive results if we consider the *speedup* analysis. For this specific test case, averaging over 100 parameters, a number around $2 \cdot 10^4$ reduced simulation can be performed in the time needed for an high fidelity one. In Figures 5.2 and 5.3, we show some

representative solutions for $t = 1s, 3s$, with $\alpha = 0.01$ and $\boldsymbol{\mu} = (12.0, 1.0, 2.5)$. They all match: hence, the reduced model is capable to reproduce the high fidelity solution for different time instances. We recall that the control can be recovered through relation (1.56). The next comparison is based on the employment of the Babuška inf-sup constant. Table 5.1 presents the average absolute and relative errors

$$\|e\|_{\text{abs}} := \sqrt{(\|y^{\mathcal{N}_y} - y_N\|_Y^2 + \|z^{\mathcal{N}_y} - z_N\|_Y^2)} \tag{5.16}$$

and

$$\|e\|_{\text{rel}} := \frac{\sqrt{(\|y^{\mathcal{N}_y} - y_N\|_Y^2 + \|z^{\mathcal{N}_y} - z_N\|_Y^2)}}{\sqrt{(\|y^{\mathcal{N}_y}\|_Y^2 + \|z^{\mathcal{N}_y}\|_Y^2)}}, \tag{5.17}$$

respectively, together with the effectivity $\eta := \Delta_N(\boldsymbol{\mu})/\|e\|_{\mathcal{Q}\times\mathcal{Q}}$ and the value of the error estimator[3] $\Delta_N(\boldsymbol{\mu})$.

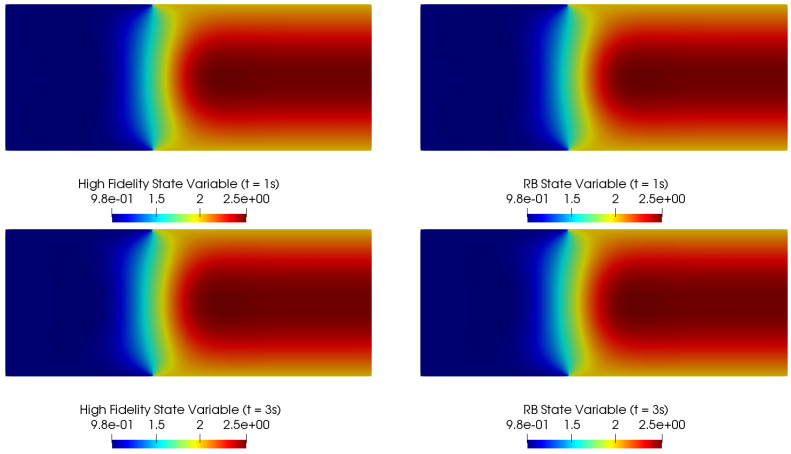


Figure 5.2: *Error Certification for parabolic problems.* Optimal high fidelity and reduced state solutions with $\alpha = 0.01$ and $\boldsymbol{\mu} = (12.0, 1.0, 2.5)$. *Left.* High fidelity state variable for $t = 1s, 3s$. *Right.* Reduced state variable for $t = 1s, 3s$.

By definition, the lower bound $\beta^{LB}(\boldsymbol{\mu})$ cannot give better results w.r.t. the employment of the exact value of $\beta_{\underline{\mathcal{B}}}^{\mathcal{N}}(\boldsymbol{\mu})$. However, using the lower bound lightens the computational costs of the offline phase. Indeed, given a parameter $\boldsymbol{\mu}$, the exact computation of the Babuška inf-sup constant takes, in average, $9.7s$, while the computation of the lower bound is performed in $0.09s$, only.

[3] With abuse of notation, $\Delta_N(\boldsymbol{\mu})$ will describe both the *exact error estimator*, given by the use of the Babuška inf-sup constant $\beta_{\underline{\mathcal{B}}}^{\mathcal{N}}(\boldsymbol{\mu})$, and the *surrogate error estimator*, derived by employing the lower bound $\beta^{LB}(\boldsymbol{\mu})$.

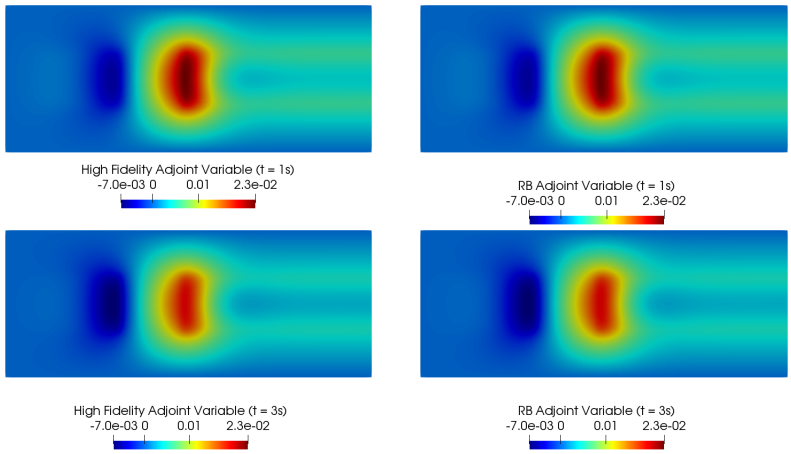


Figure 5.3: *Error Certification for parabolic problems.* Optimal high fidelity and reduced adjoint solutions with $\alpha = 0.01$ and $\boldsymbol{\mu} = (12.0, 1.0, 2.5)$. *Left.* High fidelity adjoint variable for $t = 1s, 3s$. *Right.* Reduced adjoint variable for $t = 1s, 3s$.

Table 5.1: *Error Certification for parabolic problems.* Time-dependent case: performance analysis for the problem (P) for $\alpha = 0.01$. Average error, estimators and effectivities exploiting the lower bound $\beta^{LB}(\boldsymbol{\mu})$ and the Babuška inf-sup constant $\beta^{\mathcal{N}}_{\underline{\mathcal{B}}}(\boldsymbol{\mu})$, w.r.t. N.

N	$\beta^{LB}(\boldsymbol{\mu})$				$\beta^{\mathcal{N}}_{\underline{\mathcal{B}}}(\boldsymbol{\mu})$			
	$\|e\|_{\text{rel}}$	$\|e\|_{\text{abs}}$	$\Delta_N(\boldsymbol{\mu})$	η	$\|e\|_{\text{rel}}$	$\|e\|_{\text{abs}}$	$\Delta_N(\boldsymbol{\mu})$	η
1	5.61e–1	4.37e+0	1.29e+2	2.96e+1	5.25e–1	4.46e+1	1.30e+1	2.91e+0
3	1.81e–1	5.84e–1	3.42e+1	5.86e+1	1.16e–1	5.58e–1	3.16e+0	5.67e+0
5	3.13e–2	1.58e–1	7.25e+0	4.58e+1	3.84e–2	1.99e–1	9.03e–1	4.53e+0
7	1.12e–3	4.98e–2	3.07e–1	6.17e+1	7.70e–3	3.76e–2	1.98e–1	5.26e+0
9	4.36e–2	1.33e–2	6.38e–1	4.78e+1	3.42e–3	1.24e–2	5.41e–2	4.36e+0
11	1.46e–2	4.68e–3	2.23e–1	4.76e+1	1.19e–3	4.06e–3	2.11e–2	5.21e+0
13	3.90e–4	1.35e–3	7.32e–2	5.38e+1	3.53e–4	1.26e–3	6.46e–3	5.10e+0

Nevertheless, exploiting the lower bound instead of the exact error estimator poorly affects the accuracy of the greedy algorithm: their errors are comparable. Another indicator of the effectivity of the proposed error bound is also represented by Figure 5.4, where the value of the exact inf-sup stability constant is compared with the lower bound w.r.t. the value of the parameter μ_1. The latter is the only parameter affecting the left hand side of the system and, thus, the constants we are dealing with. We point out how the value of the penalization factor α slightly changes the tightness of the provided lower bound: for lower values of α, we have, generally, a worse approximation of the Babuška inf-sup constant. This phenomenon is not new in literature, see, for example [80]. Thus, while we can observe a good representation of $\beta^{\mathcal{N}}_{\underline{\mathcal{B}}}(\boldsymbol{\mu})$ through the

bound in the left plot of Figure 5.4, we lose some precision for smaller values of the penalization parameter as depicted in the right plot of Figure 5.4, most of all for larger values of μ_1.

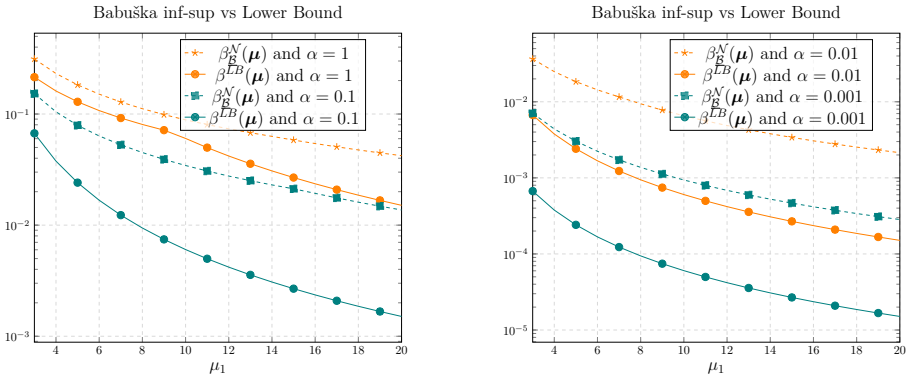


Figure 5.4: *Error Certification for parabolic problems.* Comparison of the value of the lower bound $\beta^{LB}(\boldsymbol{\mu})$ with respect to the exact Babuška inf-sup constant $\beta_{\mathcal{B}}^{\mathcal{N}}(\boldsymbol{\mu})$ for $\alpha = 1, 0.1$ (*Left*) and $\alpha = 0.01, 0.001$ (*Right*). The analysis has been performed varying the value of μ_1.

The steady case

We now show the results for the steady problem (P_s). The high fidelity dimension is $2N_{FE}^y = 9072$. The Greedy procedure has been performed with the tolerance τ, reached after picking $N = 11$ snapshots. The aggregated reduced system is really convenient to be employed due to its low dimensionality of $4N = 44$. The number of basis functions is lower if compared to the time-dependent one. This was quite expected since we are dealing with a simpler setting. Also in this case, we fixed $\alpha = 0.01$. In Figure 5.5, we show a representative solution for state and adjoint variables, obtained exploiting the lower bound $\beta_s^{LB}(\boldsymbol{\mu})$. Furthermore, an averaged performance analysis is reported in Table 5.2, where an average absolute and relative error, given by

$$\|e\|_{\text{abs}} := \sqrt{(\|y^{N_{FE}^y} - y_N\|_Y^2 + \|z^{N_{FE}^y} - z_N\|_Y^2)}, \tag{5.18}$$

and

$$\|e\|_{\text{rel}} := \frac{\sqrt{(\|y^{N_{FE}^y} - y_N\|_Y^2 + \|z^{N_{FE}^y} - z_N\|_Y^2)}}{\sqrt{(\|y^{N_{FE}^y}\|_Y^2 + \|z^{N_{FE}^y}\|_Y^2)}}, \tag{5.19}$$

are compared w.r.t. use of the lower bound and the exact Babuška inf-sup constant. The performance is also presented in terms of comparison of the

effectivity value $\eta := \Delta_N(\boldsymbol{\mu})/\|e\|_{Y\times Y}$ and the error estimator itself[4]. Also in this case, by definition, the lower bound cannot perform as well as the exact Babuška inf-sup constant, which is still preferable in terms of estimator and effectivity.

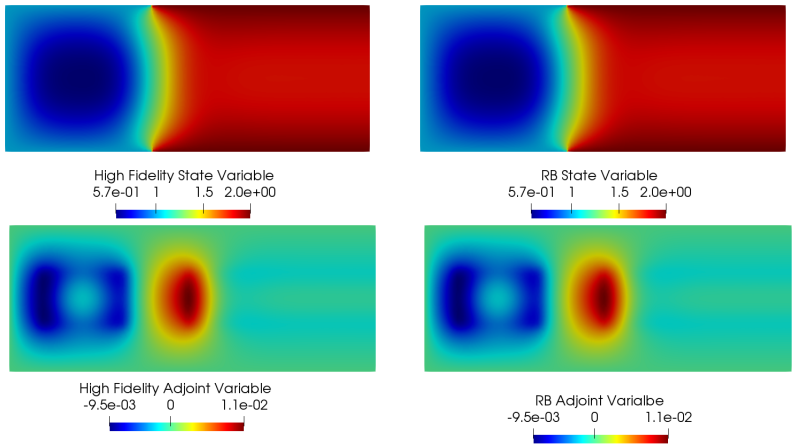


Figure 5.5: *Error Certification for parabolic problems.* Optimal high fidelity and reduced state and adjoint solutions with $\alpha = 0.01$ and $\boldsymbol{\mu} = (15.0, 0.6, 1.8)$. *Top Left*. High fidelity state variable. *Top Right*. Reduced state variable. *Bottom Left*. High fidelity adjoint variable. *Bottom Right*. Reduced adjoint variable.

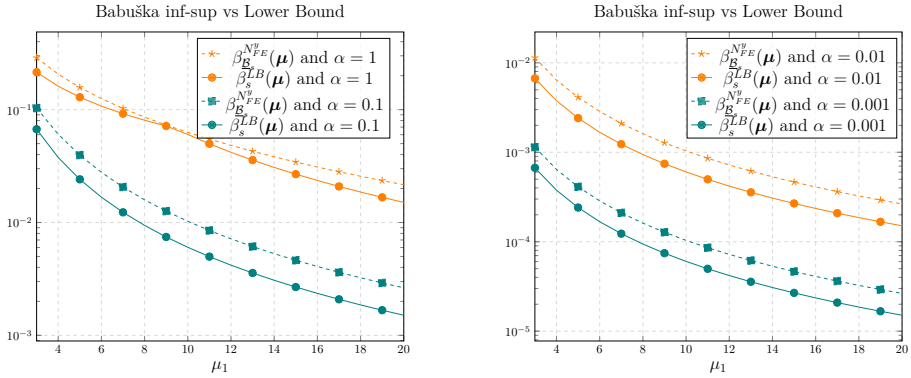


Figure 5.6: *Error Certification for parabolic problems.* Comparison of the value of the lower bound $\beta^{LB}(\boldsymbol{\mu})$ with respect to the exact Babuška inf-sup constant $\beta^{\mathcal{N}}_{\underline{\mathcal{B}}_s}(\boldsymbol{\mu})$ for $\alpha = 1, 0.1$ (*Left*) and $\alpha = 0.01, 0.001$ (*Right*). The analysis has been performed varying the value of μ_1.

[4]See footnote 3.

Table 5.2: *Error Certification for parabolic problems.* Steady case: performance analysis for the problem (P_s) for $\alpha = 0.01$. Average error, estimators and effectivities exploiting the lower bound β_s^{LB} and the Babuška inf-sup constant $\beta_{\underline{\mathcal{B}}_s}^{N_{FE}^y}(\boldsymbol{\mu})$, w.r.t. N.

N	$\beta_s^{LB}(\boldsymbol{\mu})$				$\beta_{\underline{\mathcal{B}}_s}^{N_{FE}^y}(\boldsymbol{\mu})$			
	$\|e\|_{\text{rel}}$	$\|e\|_{\text{abs}}$	$\Delta_N(\boldsymbol{\mu})$	η	$\|e\|_{\text{rel}}$	$\|e\|_{\text{abs}}$	$\Delta_N(\boldsymbol{\mu})$	η
1	7.76e–1	6.08e–1	7.35e+1	1.20e+2	7.72e–1	5.57e–1	1.52e+1	2.73e+1
3	1.23e–1	4.85e–2	4.36e+1	1.75e+2	1.71e–1	5.56e–2	1.40e+0	2.52e+1
5	3.83e–2	1.77e–2	5.61e+0	1.86e+1	4.76e–2	1.84e–2	3.91e–1	2.11e+1
7	6.28e–3	1.92e–3	9.66e–1	2.11e+2	1.07e–2	3.52e–3	1.05e–1	2.98e+1
9	1.59e–3	6.64e–4	1.44e–1	1.56e+2	5.07e–3	1.36e–3	3.01e–2	2.21e+1
11	9.30e–4	2.18e–4	7.27e–2	1.51e+2	1.45e–3	2.19e–4	6.88e–3	3.14e+1

However, the lower bound is convenient to use also in the steady case, even if it represents a simpler setting. Indeed, it gives comparable results in the average error analysis and the computational time needed for an evaluation is approximately of 0.09s, while the steady eigenvalue problem related to the exact computation of $\beta_{\underline{\mathcal{B}}_s}^{N_{FE}^y}$ takes $0.38s$ to be solved. We recall that the evaluation must be performed for all the parameters in $\mathcal{D}_h$. Furthermore, in the steady case, the penalization parameter α slightly affects the effectivity of the estimator, as plots in Figure 5.6 show. The bounds comparison is shown in terms of μ_1: the observation parameters μ_2 and μ_3 are not relevant for the computation of the Babuška inf-sup since they affect the right hand side of the problem at hand. We remark that the behaviour of the constants is comparable to the time-dependent one. This is not surprising, since we treated the problem as a space-time approximation itself that deals with time dependency as a steady structure.
Last, in terms of speedup we reach quite important results also in this simpler setting, with averaged values around $80 - 85$, independently from the value of N. In the next Section we will show how the proposed lower bound performs in a more complex test case, where also geometrical parametrization is considered.

5.2.2 Physical and Geometrical Parametrization

This Section shows numerical results for the no-control version of the boundary OCP($\boldsymbol{\mu}$) governed by a Graetz flow already presented in Section 4.2.1. The parameter dependent domain is $\Omega(\mu_3)$ represented in Figure 4.1. For the sake of clarity, we recall the structure of the domain. The observation domain is $\Omega_{\text{obs}}(\mu_3) = \Omega_3(\mu_3) \cup \Omega_4(\mu_3)$, where $\Omega_3(\mu_3) = [1, 1 + \mu_3] \times [0.8, 1]$, $\Omega_4(\mu_3) = [1, 1 + \mu_3] \times [0., 0.2]$, while $\Omega_2(\mu_3) = [1, 1 + \mu_3] \times [0.2, 0.8]$. With Ω_1 we denote the unit square, while we will call the control domain as $\Gamma_C(\mu_3) = [1, 1 + \mu_3] \times \{0\} \cup [1, 1 + \mu_3] \times \{1\}$. Also the parametric setting has already been introduced in Section 4.2.1. Here, we recall it and, moreover, we will adapt the classical optimality system to the no-control framework. The parameter we are

considering is $\boldsymbol{\mu} := (\mu_1, \mu_2, \mu_3) \in \mathcal{D} = [6.0, 20.0] \times [0.5, 3.0] \times [1.0, 3.0]$, where μ_1 represents the Péclet number, μ_2 is the desired constant value we want to reach in the observation domain, and μ_3 stretches the length of the domain. The control acts on the boundary and, namely, the problem tries to change the Neumann conditions in order to steer the solution towards the desired profile. The no-control optimality system reads: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, z) \in \mathcal{Y}_0 \times \mathcal{Y}_T$ such that

$$
\begin{cases}
y\chi_{\Omega_{\text{obs}}(\boldsymbol{\mu})} - z_t & \\
\quad -\dfrac{1}{\mu_1}\Delta z - x_2(1-x_2)\dfrac{\partial z}{\partial x_1} = \mu_2 \chi_{\Omega_{\text{obs}}(\boldsymbol{\mu})} & \text{in } \Omega(\mu_3) \times (0,T), \\
y_t - \dfrac{1}{\mu_1}\Delta y - x_2(1-x_2)\dfrac{\partial y}{\partial x_1} - \dfrac{1}{\alpha} z \chi_{\Omega_u} = 0 & \text{in } \Omega(\mu_3) \times (0,T), \\
y(0) = y_0 & \text{in } \Omega(\mu_3), \\
z(T) = 0 & \text{in } \Omega(\mu_3), \\
\dfrac{1}{\mu_1}\dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N(\mu_3) \times (0,T), \\
\dfrac{1}{\mu_1}\dfrac{\partial y}{\partial n} = u & \text{on } \Gamma_C(\mu_3) \times (0,T), \\
y = 1 \text{ and } z = 0 & \text{on } \Gamma_D \times (0,T),
\end{cases}
\tag{P^g}
$$

where $y_0 = 0$ in $\Omega(\mu_3)$ and satisfies the boundary conditions which are of Dirichlet type on $\Gamma_D = \partial\Omega(\mu_3) \cap \{(x_1, x_2) \mid x_1 \leq 1\}$ and of Neumann type on $\Gamma_N(\mu_3) = \partial\Omega(\mu_3) \setminus \Gamma_D \cup \Gamma_C(\boldsymbol{\mu})$. Analogously, the steady version of the problem has the following form: given $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, z) \in Y \times Y$ such that

$$
\begin{cases}
y - \dfrac{1}{\mu_1}\Delta z - x_2(1-x_2)\dfrac{\partial z}{\partial x_1} = \mu_2 \chi_{\Omega_{\text{obs}}(\mu_3)} & \text{in } \Omega(\mu_3), \\
-\dfrac{1}{\mu_1}\Delta y - x_2(1-x_2)\dfrac{\partial y}{\partial x_1} - \dfrac{1}{\alpha} z = 0 & \text{in } \Omega(\boldsymbol{\mu}), \\
\dfrac{1}{\mu_1}\dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N(\mu_3), \\
\dfrac{1}{\mu_1}\dfrac{\partial y}{\partial n} = u & \text{on } \Gamma_C(\mu_3), \\
y = 1 \text{ and } \ = 0 & \text{on } \Gamma_D.
\end{cases}
\tag{P_s^g}
$$

The reference domain will be referred as Ω and it is related to $\mu_3 = 1$, the interested reader can find the details in [109]. Due to the relation $\Omega_{\text{obs}}(\mu_3) \neq \Omega(\mu_3)$, we exploit the lower bound (5.13) for $\alpha = 0.07$, both for the time-dependent and the steady case, to be consistent with the test case of Section 5.2.1. We analyzed also other values of the penalization parameters and the last bound of (5.10): we postpone the analysis in Remark 5.2.1. It is clear that, by definition, the value of the reference geometrical parameter affects the constants c_u and c_{obs}, however, for the sake of notation, we will omit this dependency. In

this specific test case,

$$\gamma_a(\boldsymbol{\mu}) := \min\left\{\frac{1}{\mu_1}, \frac{1}{\mu_1\mu_3}, \frac{\mu_3}{\mu_1}, 1\right\}\frac{1}{(1+C_\Omega^2)},$$

where C_Ω is the Poincaré constant related to Ω. Furthermore, the explicit definition of $c_u = C_{\Gamma_C}$, is related to the *trace constant* that verifies $\|z\|_{\Gamma_C} \leq C_{\Gamma_C}\|z\|_{H_1(\Omega)}$, once again we refer to [128]. After tracing back the problem in a reference domain, the computation of the Poincaré and the trace constants can be performed only once, solving directly the related eigenvalue problem in Ω. Namely, the exact computation of these constants does not affect the offline performance procedure. Furthermore, following the same strategy of (5.15), we obtain that $c_c(\boldsymbol{\mu}) = C_{\Gamma_C}$. For the spatial discretization we exploited $\mathbb{P}^1 - \mathbb{P}^1$ elements while $N_t = 30$ considering the time interval $[0, T] = [0., 5.]$, i.e. $\Delta t = 1./6.$: namely, the results can be compared to Section 4.2.1. As in the previous test case, the Greedy algorithm is performed on a uniformly distributed parametric set $\mathcal{D}_h \subset \mathcal{D}$ of cardinality $N_{\max} = 225$. The tolerance has been set equal to 10^{-4}. For both the steady and the time-dependent case, the performance analysis has been carried out over 100 parameters uniformly distributed in $\mathcal{D}$.

The time-dependent case

Let us focus on equation (P^g). We fix $\alpha = 0.07$ and we apply the space-time Greedy algorithm exploiting the lower bound (5.13). The tolerance τ is reached after picking $N = 19$ snapshots. The employment of the aggregated space technique leads to a global dimension of $4N = 76$. The problem at hand needs a larger number of basis: this is due to its more complicated structure, based on boundary control framework and on the geometrical parametrization. Still, using the reduced model will be convenient compared to the high fidelity one, of dimension $\mathcal{N} = 2N_{FE}^y \times N_t = 310'980$.

Table 5.3: *Error Certification for parabolic problems.* Time-dependent case: performance analysis for the problem (P^g) for $\alpha = 0.07$. Average error, estimators and effectivities exploiting the lower bound $\beta^{LB}(\boldsymbol{\mu})$ and the Babuška inf-sup constant $\beta_{\underline{\mathcal{B}}}^{\mathcal{N}}(\boldsymbol{\mu})$, w.r.t. N.

N	$\beta^{LB}(\boldsymbol{\mu})$				$\beta_{\underline{\mathcal{B}}}^{\mathcal{N}}(\boldsymbol{\mu})$			
	$\|e\|_{\text{rel}}$	$\|e\|_{\text{abs}}$	$\Delta_N(\boldsymbol{\mu})$	η	$\|e\|_{\text{rel}}$	$\|e\|_{\text{abs}}$	$\Delta_N(\boldsymbol{\mu})$	η
1	5.61e–1	7.16e+0	3.05e+4	4.27e+3	6.62e–1	1.51e+1	7.68e+1	5.08e+1
3	2.10e–1	2.26e+0	5.44e+3	2.40e+3	2.75e–1	3.64e–1	1.96e+1	5.41e+1
5	8.66e–2	8.92e–1	1.99e+3	2.23e+3	8.01e–2	1.33e–1	3.89e+0	2.91e+1
7	3.88e–2	4.11e–1	6.79e+2	1.65e+3	4.57e–2	7.03e–2	3.89e+0	2.76e+1
9	2.46e–2	2.68e–1	4.99e+2	1.91e+3	2.17e–2	3.71e–2	1.94e+0	2.84e+1
11	1.09e–2	1.11e–1	1.95e+2	1.76e+3	9.06e–3	1.65e–2	1.05e+0	2.76e+1
13	7.16e–3	7.83e–2	1.40e+2	1.79e+3	5.98e–3	9.39e–3	4.59e–1	2.82e+1
15	4.60e–3	4.93e–2	1.08e+2	2.20e+3	4.26e–3	6.61e–3	2.65e–1	3.34e+1
17	2.48e–3	2.46e–2	4.65e+1	1.88e+3	2.36e–3	4.01e–3	1.11e–1	2.76e+1
19	1.72e–3	1.16e–2	3.14e+1	1.93e+3	1.73e–3	4.81e–3	6.97e–2	2.55e+1

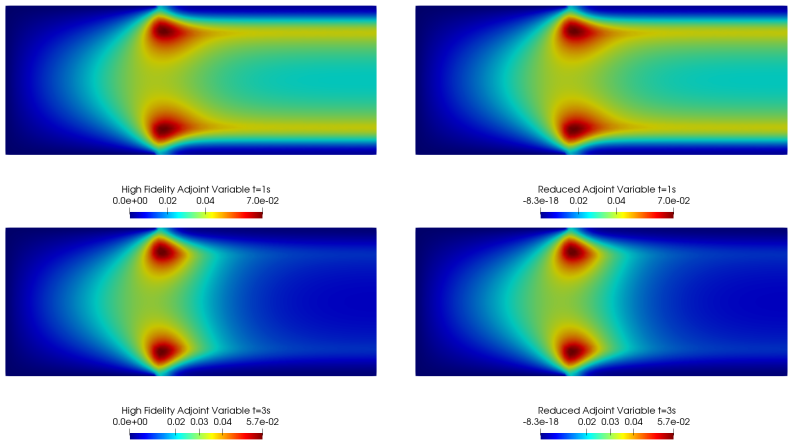


Figure 5.8: *Error Certification for parabolic problems.* Optimal high fidelity and reduced adjoint solutions with $\alpha = 0.07$ and $\boldsymbol{\mu} = (15.0, 2.5, 1.5)$. *Left.* High fidelity adjoint variable for $t = 1s, 3s$. *Right.* Reduced adjoint variable for $t = 1s, 3s$.

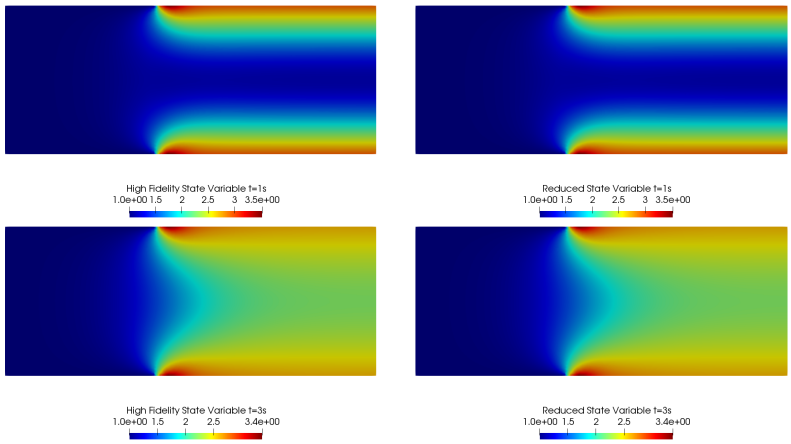


Figure 5.7: *Error Certification for parabolic problems.* Optimal high fidelity and reduced state solutions with $\alpha = 0.07$ and $\boldsymbol{\mu} = (15.0, 2.5, 1.5)$. *Left.* High fidelity state variable for $t = 1s, 3s$. *Right.* Reduced state variable for $t = 1s, 3s$.

For this reason, averaging over 100 parameters uniformly distributed in $\mathcal{D}$, we reach a quite important speedup: it is around $3 \cdot 10^4$ independently from N. Figures 5.7 and 5.8 show some representative high fidelity and reduced state and adjoint solutions for $t = 1s, 3s$, respectively, fixing $\boldsymbol{\mu} = (15.0, 2.5, 1.5)$. The reduced model solution coincide with the high fidelity one.

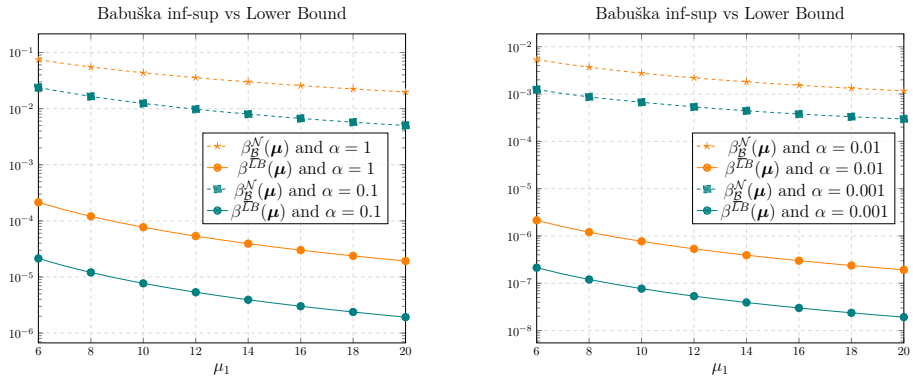


Figure 5.9: *Error Certification for parabolic problems.* Comparison of the value of the lower bound $\beta^{LB}(\boldsymbol{\mu})$ w.r.t. the exact Babuška inf-sup constant $\beta_{\underline{\mathcal{B}}}^{\mathcal{N}}(\boldsymbol{\mu})$ for $\alpha = 1, 0.1$ (*Left*) and $\alpha = 0.01, 0.001$ (*Right*). The analysis has been performed varying the value of μ_1 and fixing $\mu_3 = 2$.

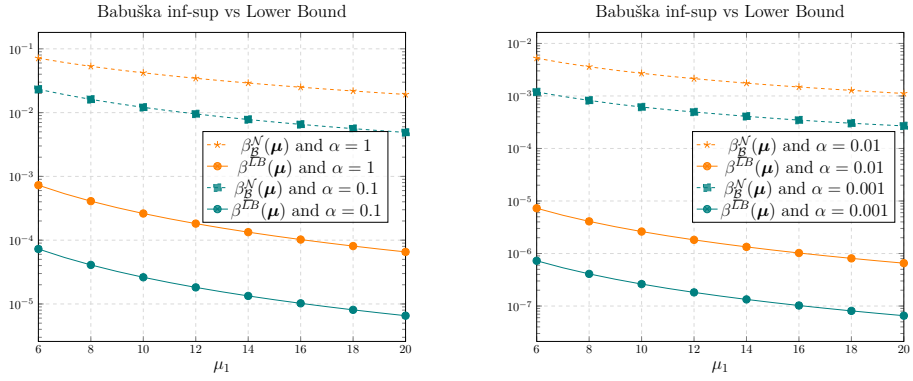


Figure 5.10: *Error Certification for parabolic problems.* Comparison of the value of the lower bound $\beta^{LB}(\boldsymbol{\mu})$ w.r.t. the exact Babuška inf-sup constant $\beta_{\underline{\mathcal{B}}}^{\mathcal{N}}(\boldsymbol{\mu})$ for $\alpha = 1, 0.1$ (*Left*) and $\alpha = 0.01, 0.001$ (*Right*). The analysis has been performed varying the value of μ_1 and fixing $\mu_3 = 1$.

Table 5.3 shows the performances of the Greedy algorithm in terms of average errors defined in (5.18) and (5.19) together with the effectivity and the error estimator[5] w.r.t. the Babuška inf-sup constant $\beta_{\underline{\mathcal{B}}}^{\mathcal{N}}(\boldsymbol{\mu})$ and the lower bound. In this case, the lower bound presents large effectivities, however, we recall that computing a single Babuška inf-sup constant takes around 8.6s, while the

[5]See footnote 3

evaluation of the value of $\beta^{LB}(\boldsymbol{\mu})$ only $0.09s$. Nonetheless, the two approaches are totally comparable in terms of the relatives error (see first and fifth columns of Table 5.3). Figures 5.9 compares the two constants w.r.t. the parameter μ_1 for several values of the penalization parameter α, with $\mu_3 = 2$ fixed. In this specific case, the two bounds worsen not only for lower values of α, but also for larger values of μ_3, as we can observe from Figure 5.10.

The steady case

We will briefly provide the results for the steady problem (P_s^g). In this case, the high fidelity dimension is $\mathcal{N} = 2N_{FE}^y = 10366$ and, after applying the Greedy algorithm, a reduced space of dimension $4N = 40$ is built. Figure 5.11 shows representative solutions for state and adjoint variables (top and bottom, respectively).

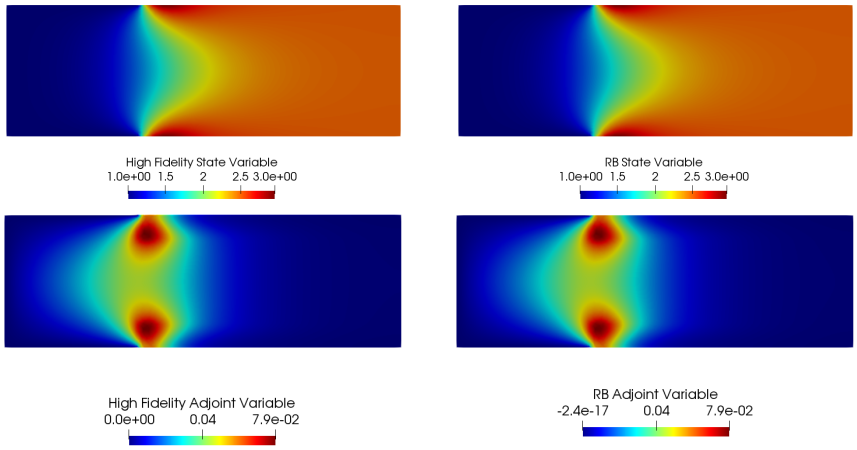


Figure 5.11: *Error Certification for parabolic problems.* Optimal high fidelity and reduced solutions with $\alpha = 0.07$ and $\boldsymbol{\mu} = (12.0, 2.5, 2.0)$. *Top Left.* High fidelity state variable. *Top Right.* Reduced state variable. *Bottom Left.* High fidelity adjoint variable. *Bottom Right.* Reduced adjoint variable.

Furthermore, averaged performance analysis is depicted in Table 5.5, where errors (5.18) and (5.19) are shown, together with an effectivity and error estimator behaviours[6]. Even if, in terms of effectivity, by definition, the Babuška inf-sup constant performs better, it pays in the offline basis construction even in this steady framework, since its exact computation, averagely, takes $0.8s$. Also in this simpler setting, the effectivity is related to the value of α as well as the value of the geometrical parameter. For the sake of brevity, we do not show the plots concerning this behaviour due to their similarity to Figure 5.9 and Figure 5.10. Moreover the speedup reaches values around 135, averagely, over a 100

[6]See footnote 3.

uniformly distributed parameters.

Remark 5.2.1 (Other α, other $\beta^{LB}(\boldsymbol{\mu})$) *In the geometrical parametritazion context we made other tests. First of all, we tried several values of α and the results reflects the plots of Figures 5.9 and 5.10. The effectivity increases when α is smaller reaching values of 10^5 with $\alpha = 0.01, 0.008$.*
Furthermore, we employed the bound

$$\beta^{LB}(\boldsymbol{\mu}) = \beta_s^{LB}(\boldsymbol{\mu}) = \frac{\gamma_a(\boldsymbol{\mu})}{\sqrt{2\max\left\{1, \left(\frac{c_m(\boldsymbol{\mu})c_{obs}}{\alpha\gamma_a(\boldsymbol{\mu})}\right)^2\right\}}}, \tag{5.20}$$

in order to compare it to the bound (5.13). Indeed, this test case allows the use of (5.20), since the control is not distributed. In this setting we need to explicit $c_m(\boldsymbol{\mu})$ and c_{obs}. It is a matter of simple computations to show that $c_m(\boldsymbol{\mu}) = C_\Omega \mu_3$. Then, the other constants C_Ω and c_{obs} can be approximated solving the related eigenvalue problem before the offline phase. Fixing $\alpha = 0.07$, the bound (5.20) gives better results w.r.t. (5.13), with lower effectivities both for the steady and the time-dependent framework. The results have been summarized in Table 5.4. The values of $\beta^{LB}(\boldsymbol{\mu})$ must be compared to Table 5.3: an order of magnitude for η is gained exploiting 5.20. The same happens for the steady constant $\beta_s^{LB}(\boldsymbol{\mu})$: indeed, compared with Table 5.5, its effectivities are around $2 \cdot 10^3$. The behaviour of the errors remains comparable w.r.t. the previous test cases. The better sharpness of (5.20) is represented in Figures 5.12, where the new lower bound and the Babuška inf-sup constant are compared for $\alpha = 0.07$ and $\mu_3 = 1$ (refer and compare to Figure 5.10). Due to its similarity with the steady case, we report here only the time-dependent case, for the sake of brevity.

Table 5.4: *Error Certification for parabolic problems.* Time-Dependent case: performance analysis for the problem (P^g) and (P_s^g) for $\alpha = 0.07$. Average error, estimators and effectivities exploiting the lower bounds $\beta^{LB}(\boldsymbol{\mu})$ and $\beta^{LB}(\boldsymbol{\mu})$ given by (5.20). (B.T.) Below tolerance τ.

N	$\beta^{LB}(\boldsymbol{\mu})$				$\beta_s^{LB}(\boldsymbol{\mu})$			
	$\|e\|_{\text{rel}}$	$\|e\|_{\text{abs}}$	$\Delta_N(\boldsymbol{\mu})$	η	$\|e\|_{\text{rel}}$	$\|e\|_{\text{abs}}$	$\Delta_N(\boldsymbol{\mu})$	η
2	4.23e–1	4.33e+0	3.689e+3	8.51e+2	2.99e–1	2.90e–1	6.53e+2	2.24e+3
4	1.45e–1	1.50e+0	1.01e+3	6.70e+2	6.36e–2	5.35e–2	1.12e+2	2.10e+3
6	5.24e–2	5.36e–1	2.67e+2	4.98e+2	3.01e–2	2.04e–2	3.99e+1	1.95e+3
8	2.80e–2	2.94e–1	1.58e+2	5.40e+2	7.60e–3	6.21e–3	2.56e+1	2.01e+3
10	1.38e–2	1.41e–1	6.83e+1	4.84e+2	B.T.			
12	9.49e–3	9.21e–2	4.32e+1	4.69e+2	B.T.			
14	5.12e–3	5.24e–2	3.01e+1	5.74e+2	B.T.			
16	3.78e–3	3.85e–2	2.12e+1	5.52e+2	B.T.			
18	2.57e–3	2.41e–2	1.52e+1	6.32e+2	B.T.			
20	1.25e–3	1.21e–2	8.42e+0	6.93e+2	B.T.			
22	9.07e–4	8.64e–3	6.30e+0	7.28e+2	B.T.			

Table 5.5: *Error Certification for parabolic problems.* Steady case: performance analysis for the problem (P_s^g). Average error, estimators and effectivities exploiting the lower bound $\beta_s^{LB}(\boldsymbol{\mu})$ and the Babuška inf-sup constant $\beta_{\underline{\mathcal{B}}_s^{N_{FE}^y}}^{N_{FE}^y}(\boldsymbol{\mu})$, w.r.t. N.

N	$\beta_s^{LB}(\boldsymbol{\mu})$				$\beta_{\underline{\mathcal{B}}_s^{N_{FE}^y}}^{N_{FE}^y}(\boldsymbol{\mu})$			
	$\|e\|_{\text{rel}}$	$\|e\|_{\text{abs}}$	$\Delta_N(\boldsymbol{\mu})$	η	$\|e\|_{\text{rel}}$	$\|e\|_{\text{abs}}$	$\Delta_N(\boldsymbol{\mu})$	η
2	3.16e–1	3.16e–1	3.25e+3	1.01e+4	5.82e–1	5.57e–1	1.01e+4	3.72e+1
4	5.69e–2	5.69e–2	4.53e+2	7.95e+3	6.67e–2	4.93e–2	7.95e+3	2.39e+1
6	1.52e–2	1.52e–2	1.11e+2	7.32e+3	2.72e–2	1.52e–2	7.32e+3	3.22e+1
8	5.23e–3	4.95e–3	3.40e+1	6.88e+3	7.33e–3	4.54e–3	6.88e+3	3.13e+1
10	2.03e–3	1.89e–3	1.25e+1	6.78e+3	4.63e–3	2.65e–3	6.78e+3	3.78e+1

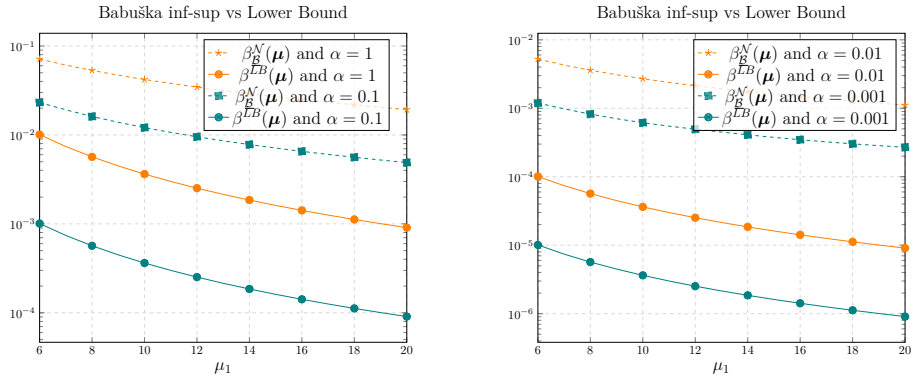


Figure 5.12: *Error Certification for parabolic problems.* Comparison of the value of the lower bound $\beta^{LB}(\boldsymbol{\mu})$ w.r.t. the exact Babuška inf-sup constant $\beta_{\underline{\mathcal{B}}}^{\mathcal{N}}(\boldsymbol{\mu})$ for $\alpha = 1, 0.1$ (*Left*) and $\alpha = 0.01, 0.001$ (*Right*). The analysis has been performed varying the value of μ_1 and fixing $\mu_2 = 1$.

This Chapter together with Chapter 4, provides a tool to deal with reduced OCP($\boldsymbol{\mu}$)s form many state equations, from steady problems to time-dependent nonlinear ones. We aimed at giving the basic ideas of ROMs and at describing the strategies that can be employed in a time-dependent framework. In the next Chapters, we will show some advanced applications in OCP($\boldsymbol{\mu}$)s which combine reduction to other mathematical fields, such as Uncertainty Quantification, bifurcation theory and automatic learning techniques.

Part III

Advanced Applications of Optimal Control Problems

Chapter 6

Parametric Optimal Control for Environmental Sciences: Deterministic and Stochastic Reduction Approaches

This Chapter focuses on OCP($\boldsymbol{\mu}$)s for environmental marine applications exploiting POD-Galerkin technique both in deterministic and stochastic contexts. The main motivations underlay in the growing demand of fast and reliable simulations in such a field. Indeed, on one side OCP($\boldsymbol{\mu}$)s can make the model more accurate: the state equation is changed through control variables in order to reach the desired profile representing a *collected data*, while, on the other side, the ROMs can be a great tool to have at one's disposal parametric simulation in real-time to improve the capabilities of analysis of the environmental model at hand. This will be the topic of Section 6.1, where the POD approach has been applied to realistic test cases presented in [145]. In Section 6.2, the setting will be generalized to the stochastic context proposing a specific algorithm to deal with random inputs as parameters, the so called weighted ROMs (w-ROMs) [37, 158]. Some numerical results will be also presented, following the scenarios discussed in [33].

6.1 Deterministic OCP($\boldsymbol{\mu}$)s

This Section concerns the application of the POD reduction technique in geographically realistic experiments for marine sciences with environmental purposes. Indeed, the marine behaviour is related to social and economic growth, biodiversity preservation, monitoring plans for possibly dangerous phenomena related to natural or anthropic factors. It is clear that, in this context, PDE($\boldsymbol{\mu}$)s play a crucial role in representing different physical features and, thus, fast simulations are needed to analyse different configurations. However, PDE($\boldsymbol{\mu}$)s can lead to results that do not represent the expected behaviour guessed from data. Hence, OCP($\boldsymbol{\mu}$)s can be interpret in a variational *data assimilation* fashion, being a support tool to improve the forecasting capabilities of the models to predict future scenarios, see e.g. [52, 78, 155]. Yet, as presented in [16, 147] and in Section 4.3, data assimilated problems are characterized by huge computational complexity that still limits their applicability, most of all if the optimization problem deals with very complicated parametric flow models. Reduced OCP($\boldsymbol{\mu}$)s have been successfully applied to several test cases of interest in this research field, see e.g. [45, 126, 127, 145, 150, 147]). In this context, ROMs can

be exploited to efficiently study different marine behaviours related to a parametric change in the physics and in the geometry of the phenomenon. In this Section, we present the results of [145] based on the application of POD-Galerkin approach for

- (Section 6.1.1) a pollutant control in the Gulf of Trieste, Italy;
- (Section 6.1.2) a nonlinear Oceanographic solution tracking governed by Quasi-Geostrophic equations.

Remark 6.1.1 *For the sake of clarity, we will briefly discuss the formulation we used for both the test cases. Indeed:*

- *the pollutant control test case is governed by a steady linear PDE($\boldsymbol{\mu}$), thus the reader may refer to Section 1.2.1 and to Remark 2.3.1 for the continuous and high fidelity frameworks, respectively.*
- *The nonlinear solution tracking will be governed by a steady nonlinear PDE($\boldsymbol{\mu}$): the formulation of this setting has been described in Remark 1.2.1 at the continuous level and in Remark 2.2.1 at the discrete one.*

Concerning the reduced procedure, we employed a partitioned POD as described in Section 4.1, together with the aggregated space technique to guarantee the unique solution of the reduced optimality system, see Section 3.2.2.

6.1.1 Pollutant Control in the Gulf of Trieste

The test case we are going to present aims at validating the POD reduction strategy of Section 4.1 in marine environment and ecosystem monitoring setting. The numerical example deals with an advection-diffusion pollutant control problem in the Gulf of Trieste, Italy. The Gulf is a physical basin which presents a very peculiar windy behaviour. Besides, it is characterized by a rich naturalistic biodiversity that is always kept in consideration in monitoring and preserving plans. A parametric analysis of the system is necessary in this context. Indeed, it is related to the marine safeguard of the Gulf and, consequently, to its impact on the city life. Trieste overlooks the sea and its economical and social activities are strictly dependent on its marine environment.
In the presented experiment, we pretended a pollutant loss in the city harbour. In this dangerous setting, it was important to work with a realistic mesh derived from satellite images of the geographic area at hand. Figure 6.1 shows the satellite area of the Gulf of Trieste and a mesh overlapping for the Trieste harbour, our zone of interest. Having this specific triangulation of the domain allowed us to work in a more physical meaningful framework. This led to reliable results that could be potentially compared to *in situ* collected data. The OCP($\boldsymbol{\mu}$) we considered aim at letting the pollutant concentration y be under

a safeguard threshold, in other words, the desired concentration y_{d}. The phenomenon is studied in a parametric setting, representing various wind actions on the surface of the geographic area of the Gulf. The problem reads: given $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3) \in \mathcal{D} = [0.5, 1] \times [0., 1.] \times [0., 1.]$ and $y_{\mathrm{d}} = 0.2 \in L^2(\Omega_{\mathrm{obs}})$ find the optimal pair $(y, u) \in H^1_{\Gamma_D}(\Omega) \times \mathbb{R}$ that solves:

$$\min_{Y \times U} \frac{1}{2} \int_{\Omega_{\mathrm{obs}}} (y - y_{\mathrm{d}})^2 d\Omega_{\mathrm{obs}} + \frac{\alpha}{2} \int_{\Omega_u} u^2 d\Omega_u, \tag{6.1}$$

under the following constraint:

$$\begin{cases} \mu_1 \Delta y + [\mu_2, \mu_3] \cdot \nabla y = Lu & \text{in} \quad \Omega, \\ \dfrac{\partial y}{\partial n} = 0 & \text{on} \quad \Gamma_N, \\ y = 0 & \text{on} \quad \Gamma_D, \end{cases} \tag{6.2}$$

where the constant $L = 10^3$ makes the system non-dimensional. Figure 6.2 represents the domain Ω and its sub-domains. In this case, in the open sea Γ_N Neumann conditions are imposed, while homogeneous Dirichlet boundary conditions apply along the coastline named Γ_D. The control variable u is the maximum quantity of pollutant allowed to be released in Ω_u (green sub-domain in Figure 6.2) to guarantee the safeguard threshold of $y_{\mathrm{d}} = 0.2$ in Ω_{obs} (red sub-domain in Figure 6.2), representing the naturalist area of Miramare, characterized by a large biodiversity of flora end fauna species. The parameter $\boldsymbol{\mu}$ describes the diffusivity and the convection current behaviour deriving from the wind blowing on the Gulf surface. In order to work in a reduced framework, we applied a partitioned POD to the variables and we aggregated the spaces. The reduced spaces have been built over $N_{max} = 100$ uniformly distributed snapshots for $\alpha = 10^{-5}$ fixed. The high fidelity dimension of the global FE space has dimension $\mathcal{N} = 5'939$. The FE approximation relies on $\mathbb{P}^1$ polynomials for all the variables. In the end, we tested the accuracy of the problem retaining $N = 20$ bases for state and adjoint variables and a single basis for the control variable, being $U = \mathbb{R}$. This choice led to a global reduced space of dimension $4N + 1 = 81$. The number of basis functions was able to recover the parametric solution. Indeed, Figure 6.3 shows the reduced and the FE solution coinciding for a representative parameter and, furthermore, the accuracy is confirmed by the average relative log-error of the variables w.r.t. the value of N, depicted in Figure 6.4. The error is averaged over 100 uniformly distributed parameters and we reach good approximations even with a small values of N (we obtain values between 10^{-8} and 10^{-9} with $N = 20$). Furthermore, the affine structure of the problem leads to high speedup index, which is around 250 for a reduced space of global dimension 81. Furthermore, we want to stress that the system has actually changed to reach the desired configuration: the left plot of Figure 6.3 shows an uncontrolled scenario where the maximum pollutant quantity is released. In order to verify the safeguard threshold y_{d}, the pollutant must

have a smaller concentration. It is clear that the solution has been steered to a safer scenario. Namely, the ROMs can be an ally in monitoring plan to prevent dangerous consequences in the field of environmental research.

Figure 6.1: *Pollutant Control.* *Left*. Satellite images of the Gulf of Trieste. *Center*. Zoom of the mesh overlapping satellite images (Trieste harbour). *Right*. The final mesh.

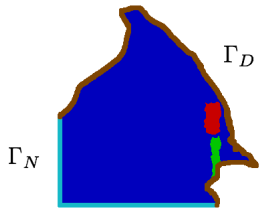


Figure 6.2: *Pollutant Control.* sub-domains and boundaries. Red: observation domain Ω_{obs}. Green: control domain Ω_u.

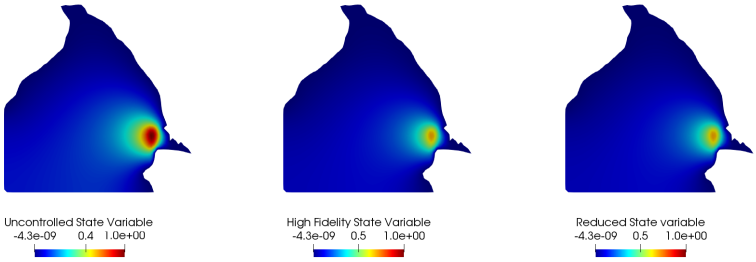


Figure 6.3: *Pollutant Control.* *Left*. Uncontrolled concentration solution. *Center and Right*. Optimal FE and ROM state pollutant concentration for $\boldsymbol{\mu} = (1., -1., 1.)$, representing the Bora wind action.

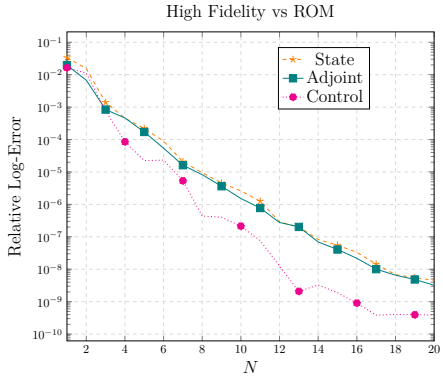


Figure 6.4: *Pollutant Control.* Averaged relative log-error for the variables.

The next Section will focus on another application of OCP($\boldsymbol{\mu}$)s in environmental sciences, where the control can be interpret as the natural features needed in order to achieve a future expected scenario.

6.1.2 Nonlinear Oceanographic Solution Tracking

In this Section, we discuss the employment of reduced optimal control strategies in order to predict several parametric instances in an Oceanographic environmental setting. Specifically, we will deal with global marine circulation weather forecasting. In this field:

- the large scale flow dynamics is represented by Quasi-Geostophic equations that describe the Ocean and Atmosphere system [34, Chapter 3]. Even if they represent a quite complex phenomenon and even if they carry out a good degree of completeness, they are not enough to satisfactory forecast the Ocean circulation behaviour. Indeed, they might lead to inaccurate results not comparable with historical or *in situ* collected data.

- The Oceanography field is characterized by a huge lack of data. Indeed, they can be complicated and expensive to collect. Thus, the small information provided is usually difficult to interpret, scattered and noisy.

Hence, the model can be not reliable while the data might be physically meaningless. In this context, we aim at adding data information thanks to the employment of an OCP($\boldsymbol{\mu}$) that reads: given $\boldsymbol{\mu} \in \mathcal{D}$, $\psi_{\mathrm{d}} \in L^2(\Omega)$, find the pair $(\psi, q) \in Y = H_0^1(\Omega) \times H_0^1(\Omega)$ and $u \in U = L^2(\Omega)$ which solves

$$\min_{Y \times U} \frac{1}{2} \int_\Omega (\psi - \psi_d)^2 \, d\Omega + \frac{\alpha}{2} \int_\Omega u^2 \, d\Omega, \tag{6.3}$$

constrained to the Quasi-Geostrophic equations

$$\begin{cases} q = \Delta\psi & \text{in } \Omega, \\ \dfrac{\partial\psi}{\partial x} = u - \mu_1 q + \mu_2 \Delta q - \mu_3 \Theta(\psi, q) & \text{in } \Omega, \\ \psi = 0 & \text{on } \partial\Omega, \\ q = 0 & \text{on } \partial\Omega, \end{cases} \tag{6.4}$$

where the nonlinear expression $\Theta : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is defined as:

$$\Theta(\psi, q) = \frac{\partial\psi}{\partial x}\frac{\partial q}{\partial y} - \frac{\partial\psi}{\partial y}\frac{\partial q}{\partial x}. \tag{6.5}$$

In other words, we are dealing with the following fourth-order controlled equation:

$$\mu_3 \Theta(\psi, \Delta\psi) + \frac{\partial\psi}{\partial x} = u - \mu_1 \Delta\psi + \mu_2 \Delta^2 \psi, \tag{6.6}$$

The action of the parameter $\boldsymbol{\mu}$ describes the North Atlantic Ocean dynamics completely, since it gives information about how the large scale Ocean circulation is affected by different phenomena, such as location and intensity variations of its gyres and its currents (typically described by μ_3). Let us recall that Ocean dynamic is related to the wind stress and atmospheric behaviour. The wind action is represented by the control variable, since it is considered as a forcing term acting on the whole geographical region depicted in the top row plots of Figure 6.5. Now, let us assume that an observation field ψ_{d} is provided. Using an optimal control framework in this context will result in a pair $(\psi, u) \in Y \times U$ such that:

- ψ is similar to the data;
- u is the forcing term needed to achieve that result.

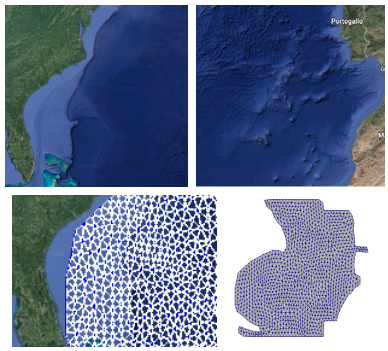


Figure 6.5: *Oceanographic Solution Tracking. (underlineTop row).* Satellite images of Florida Peninsula (left) and North Europe (Right). *Bottom Left.* Zoom of the mesh overlapping satellite images (Florida Peninsula). *Bottom Right.* Triangulation of the spatial domain Ω, representing the North Atlantic Ocean, from Florida Peninsula to North Europe.

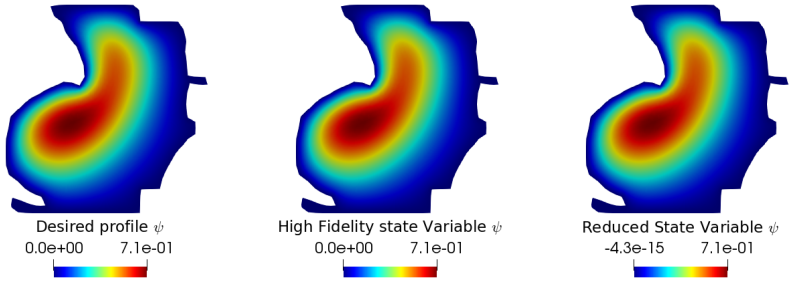


Figure 6.6: *Oceanographic Solution Tracking.* Comparison between desired state (left), FE (center) and ROM solutions (right) for $\boldsymbol{\mu} = (10^{-4}, 0.073, 0.0452)$, representing the Gulf Stream.

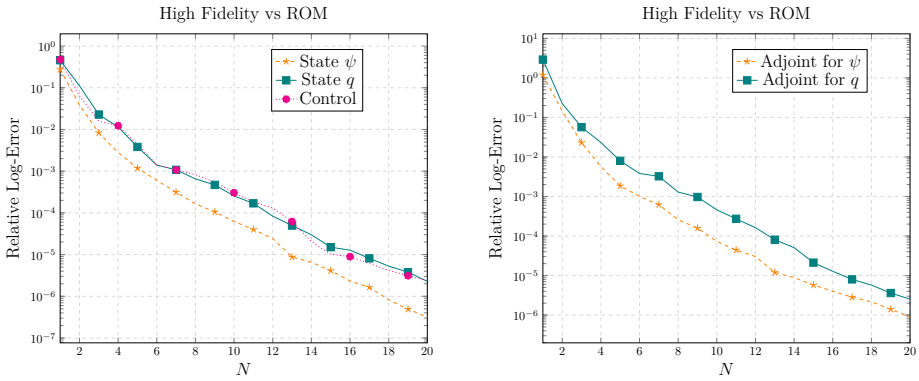


Figure 6.7: *Oceanographic Solution Tracking.* Averaged relative log-error for the variables.

Thus, on one side we are changing the system so that the data collection could be avoided, gaining in terms of time and economical costs. Furthermore, we are solving an *identification problem*: indeed, we are asking what kind of wind action u has to blow on the Atlantic region in order to reach a desirable profile. In a parametric setting, this translates in a deep analysis of seasonal phenomena and the ROMs can be a valuable tool in order to run many simulations for several parametric instances. In our test case, we would like to reproduce the Gulf Stream, represented in the left plot of Figure 6.6. The region of interest is the Florida Peninsula and the north Europe: the Gulf stream from Florida moves towards North-East to warm the European coasts. In Figure 6.5, we show the satellite image of the considered geographical area together with the used mesh. In the following we report the results related to equation (6.4),

a handier system[1] w.r.t. equation (6.6). As already introduced, μ_1 and μ_2 represent a diffusive action, while μ_3 affects the local nonlinear dynamics Θ. The Gulf stream, i.e. the desired configuration ψ_d, is given by a FE simulation obtained by the uncontrolled governing equation with a right hand side given by $-sin(\pi y)$, with $\mu_1 = 10^{-4}$, $\mu_2 = 0.07^3$ and $\mu_3 = 0.07^2$ fixed parameter. Figure 6.6 shows how effective is the optimality system for this problem: the desired solution profile, the high fidelity and the reduced solutions coincide. Furthermore, the accuracy of the reduced system is visible in Figure 6.7. It shows the average relative log-error between the FE and the reduced variables over 100 log-uniformly distributed parametric evaluations. Namely, the reduced space built with partitioned POD over $N_{max} = 100$ log-uniformly distributed snapshots, retaining $N = 20$ basis functions for each variable is able to recover all the considered fields. We recall that after the application of aggregated spaces technique, the global dimension is $9N = 180$ to be compared with the FE dimension $\mathcal{N} = 6'490$ obtained through the employment of $\mathbb{P}^1$ elements for all the variables. Despite the dimensional reduction, we reach a speedup of 2. However, this is not surprising: indeed we are not exploiting any hyper-reduction technique. As stressed in Remark 3.1.1, the issue can be overcome through the Empirical Interpolation Method (EIM) [17]. Up to now, we relied on deterministic models. However, environmental sciences indisputably undergo some *uncertainty*, not only derived from the lack of knowledge from the data collection, but also affecting the parameters of the model itself. In the next Section, we are going to discuss a tailored reduction strategy to deal with this kind of issue.

6.2 Stochastic OCP($\boldsymbol{\mu}$)s

In several scientific fields, stochastic models are used to guarantee more reliability in the simulations w.r.t. the deterministic ones. In this more complex framework, the parameters carry uncertainties and, thus, the solution can be affected by this uncertainty as well. Furthermore, the classical statistical analysis is based on Monte Carlo methods, i.e. on many simulations and this might lead to unbearable computational costs to deal with in the context of standard discretization. The aim of this Section is to introduce weighted reduced order methods (w-ROMs) to accelerate the statistical analysis process. Furthermore, we will present some of the numerical results of Section 6.1 in a stochastic fashion. After having introduced the main motivations and the needed mathematical tools in Section 6.2.1, we will introduce the weighted POD (w-POD) algorithm for OCP($\boldsymbol{\mu}$)s in Section 6.2.2. Lastly, in Section 6.2.3 we will show the numerical results presented in [33].

[1] This version of the problem does not ensure the coercivity of the state equation, that is proved in [84] for the state equation of the system (6.6).

6.2.1 Motivations and Notations

Deterministic PDE($\boldsymbol{\mu}$)s are an effective tool to model natural phenomena. However, realistic problems, such as environmental ones, usually depend on parameters that can be affected by *uncertainty* deriving from inaccurate measurement or scattered data information. To overcome this issue, stochastic PDE($\boldsymbol{\mu}$)s can be employed. In the UQ context, the main goal is to run many simulations as possible in order to compute statistics over quantities of interest. Thus, high fidelity approximation can lead unbearable computational costs when dealing with standard statistical analysis techniques based on Monte Carlo methods. This issue amplifies in the OCP($\boldsymbol{\mu}$)s setting, made by a system of three equations. Also in this UQ framework, we would like to apply a reduced approach in order to achieve the goal of many simulations in a rapid way, still being accurate. Here, we introduce w-ROMs, where, the reduced order model construction complies with some previous knowledge on the parameter distribution in order to accelerate even more the reduced simulations, see e.g. [37, 158] and the references therein. We now describe the mathematical formulation of the uncertain problem at hand. Let us consider a complete probability space $(\mathscr{A}, \mathscr{F}, P)$, where $\mathscr{A}$ is the set of the possible outcomes, $\mathscr{F}$ is a σ–algebra of events and P is a probability measure. We consider $\boldsymbol{\mu} := \boldsymbol{\mu}(\overline{\omega})$, i.e. $\boldsymbol{\mu} : (\mathscr{A}, \mathscr{F}) \to (\Gamma, \mathscr{B})$ with $\Gamma \subset \mathbb{R}^D$ compact set, $\mathscr{B}$ is the Borel measure and $\mu(\overline{\omega}) = (\mu_1(\overline{\omega}), \dots, \mu_D(\overline{\omega}))$ is a random vector. Its components are independent absolutely continuous random variables and they will describe the physical features of the considered problem. Furthermore, with $\rho : \mathbb{R}^D \to \mathbb{R}$ we denote the probability density function of $\boldsymbol{\mu}$. We now have all the ingredients to define the *stochastic OCP($\boldsymbol{\mu}$)s*. Their formulation is not different from the deterministic one besides the presence of a random input $\boldsymbol{\mu}(\overline{\omega})$ with $\overline{\omega} \in \mathscr{A}$.

Remark 6.2.1 *As already specified, we were building on the numerical results of [145], where the two test cases analysed are the ones presented in Sections 6.1.1 and 6.1.2. We dealt with steady linear (see e.g. Section 1.2.1 and Remark 2.3.1) and with steady nonlinear problems (see e.g. Remarks 1.2.1 and 2.2.1). Thus the generalization to random inputs is straightforward.*

The next Section concerns the use of a tailored POD strategy that takes into account the randomness of the system leading to even more convenient ROMs approximations.

6.2.2 Weighted POD for OCP($\boldsymbol{\mu}$)s

As already specified, stochastic equations aim at computing statistical momenta of an output of interest and they usually rely on averaging algorithms such as Monte Carlo methods. In this context, many simulations for different parameters $\boldsymbol{\mu} \in \Gamma$ are necessary to achieve reliable data analysis. Thus, the ROMs can be effective in this field. Namely, the goal of this Section is to propose w-POD

as introduced in [158]. For the sake of clarity, we briefly recall the POD ideas for OCP($\boldsymbol{\mu}$)s. In the steady framework, given a value N, the POD algorithm as described in Section 4.1 minimizes the mean square errors:

$$\int_\Gamma \|y^{N^y_{FE}}(\boldsymbol{\mu}) - y_N(\boldsymbol{\mu})\|^2_{Y^{N^y_{FE}}} d\boldsymbol{\mu}, \quad \int_\Gamma \|u^{N^u_{FE}}(\boldsymbol{\mu}) - u_N(\boldsymbol{\mu})\|^2_{U^{N^u_{FE}}} d\boldsymbol{\mu}$$

$$\text{and} \int_\Gamma \|z^{N^y_{FE}}(\boldsymbol{\mu}) - z_N(\boldsymbol{\mu})\|^2_{Y^{N^y_{FE}}} d\boldsymbol{\mu},$$

w.r.t. all the possible reduced space $Y_N \subset Y^{N^y_{FE}}$ of dimension N, for state and adjoint variables, and $U_N \subset U^{N^u_{FE}}$ for the control variable. Let us focus on the state variable. If we consider the discretized version of it, the POD minimizes over the linear subspaces of $Y^{N^y_{FE}}$ the following quantity:

$$\frac{1}{N_{max}} \sum_{i=1}^{N_{max}} \|y^{N^y_{FE}}(\boldsymbol{\mu}_i) - y_N(\boldsymbol{\mu}_i)\|^2_{Y^{N^y_{FE}}}.$$

As specified in Section 4.1, the reduced space Y_N is defined by the N eigenvectors related to the N largest eigenvalues of the correlation matrix $\mathbf{C}^y$ where

$$\mathbf{C}^y_{ml} = \frac{1}{N_{max}} (y^{N^y_{FE}}(\boldsymbol{\mu}_m), y^{N^y_{FE}}(\boldsymbol{\mu}_l))_{Y^{N^y_{FE}}}, \qquad 1 \le m, l \le N_{max}.$$

Few modifications apply in the stochastic setting, where the mean square errors are modified into the expected value

$$\begin{aligned}\mathbb{E}[\|y^{N^y_{FE}}(\boldsymbol{\mu}) - y_N(\boldsymbol{\mu})\|^2_{Y^{N^y_{FE}}}] &= \int_\Omega \|y^{N^y_{FE}}(\boldsymbol{\mu}) - y_N(\boldsymbol{\mu})\|^2_{Y^{N^y_{FE}}} dP(\omega) \\ &= \int_\Gamma \|y^{N^y_{FE}}(\boldsymbol{\mu}) - y_N(\boldsymbol{\mu})\|^2_{Y^{N^y_{FE}}} \rho(\boldsymbol{\mu}) d\boldsymbol{\mu}.\end{aligned}$$

It is clear that the numerical approximations of the aforementioned integrals are performed in the following way:

$$\sum_{i=1}^{N_{max}} \underbrace{w_i \rho(\boldsymbol{\mu})}_{W_i} \|y^{N^y_{FE}}(\boldsymbol{\mu}_i) - y_N(\boldsymbol{\mu}_i)\|^2_{Y^{N^y_{FE}}}, \tag{6.7}$$

exploiting a quadrature rule consisting of nodes $\{\boldsymbol{\mu}_i\}_{i=1}^{N_{max}}$ and of the related quadrature weights $\{w_i\}_{i=1}^{N_{max}}$. This changed structure leads to a modified correlation matrix $\hat{\mathbf{C}}^y$, of the form

$$\hat{\mathbf{C}}^y_{ml} = W_m (y^{N^y_{FE}}(\boldsymbol{\mu}_m), y^{N^y_{FE}}(\boldsymbol{\mu}_l))_{Y^{N^y_{FE}}}, \qquad 1 \le m, l \le N_{max}.$$

Namely, the *global weight* W_m depends on two factors:

- on the probability density function chosen for $\boldsymbol{\mu}$,
- on the quadrature rule used to chosen to approximate the expected value.

The choice of the distribution and the quadrature rule are problem dependent. For an overview on the possible strategies to be used, we refer to [112, 37, 151, 162, 68, 158].

We stress that the previous arguments directly apply to both the control and the adjoint variables. Thus, partitioned w-POD can be applied and aggregated space technique can be employed also in the stochastic setting.

Remark 6.2.2 *We stress that the choice of using POD was guided by the test case proposed in the following Section as presented in [33]. However, we would like to underline that a weighted version of the Greedy algorithm was originally proposed in [37].*

In the next Section we are going to present the numerical results of [33], built on [145], mirroring what already done in Sections 6.1.1 and 6.1.2 for the deterministic case.

6.2.3 Numerical Results in Environmental Sciences

We now propose the application of w-ROM in the pollutant control environmental test case proposed in Section 6.1.1. For the sake of brevity, we will discuss only this example since the same conclusions can be derived from the stochastic version of nonlinear solution tracking problems presented in Section 6.1.2. However, the interested reader may refer to [33] for a complete w-POD framework for the nonlinear example. In the framework of the OCP($\boldsymbol{\mu}$) that minimizes (6.1) constrained to (6.2), we took into consideration different quadrature rules and distributions in order to understand the role of the weights in the reduction process. In the following, we will test the performances three different quadrature rules (see e.g. for more details [151]):

- a Monte Carlo sampler with all the quadrature weights equal to $\dfrac{1}{N_{max}}$ and that samples from the given distribution $\rho(\boldsymbol{\mu})$. In this context when the distribution is uniform we are applying standard POD.
- The tensor product of three Gaussian quadrature rules.
- A Pseudo-Random sampler adapted to general distributions through the method of inversion.

The physical framework we studied is identical to the one proposed in Section 6.1.1. However, the parameter is random $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3) \in (0.5, 1) \times (-1, 1) \times$

$(-1, 1)$ in order to better model the considered sea dynamics. In all the test cases, we chose the penalization parameter as $\alpha = 10^{-7}$. Namely, we will show the performances of informed sampling w.r.t. standard reduction strategies. We aim at empirically proving that exploiting data information can be a powerful tool in real-time monitoring. In all the experiment the snapshots number will be $N_{max} = 100$. First of all, we tried several quadrature rules. The results for all the involved variables are depicted in Figure 6.8, where the aforementioned quadrature rules have been tested in terms of relative log-errors. Here, we considered $\rho(\boldsymbol{\mu})$ as uniform. Namely, we tried to isolate the role of the quadrature weights w_i. As already observed in [158], the POD procedure is slightly affected by the quadrature weights w_i.

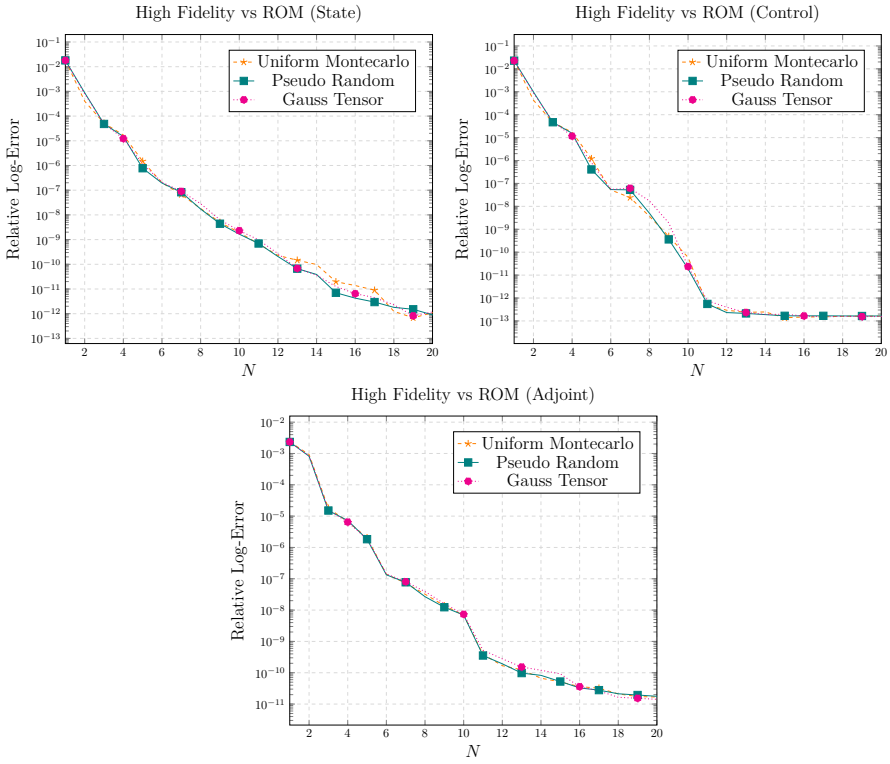


Figure 6.8: *Stochastic Pollutant Control.* Averaged relative log-error for the variables for different quadrature rules. *Top row.* State and Control variables. *Bottom row.* Adjoint variable.

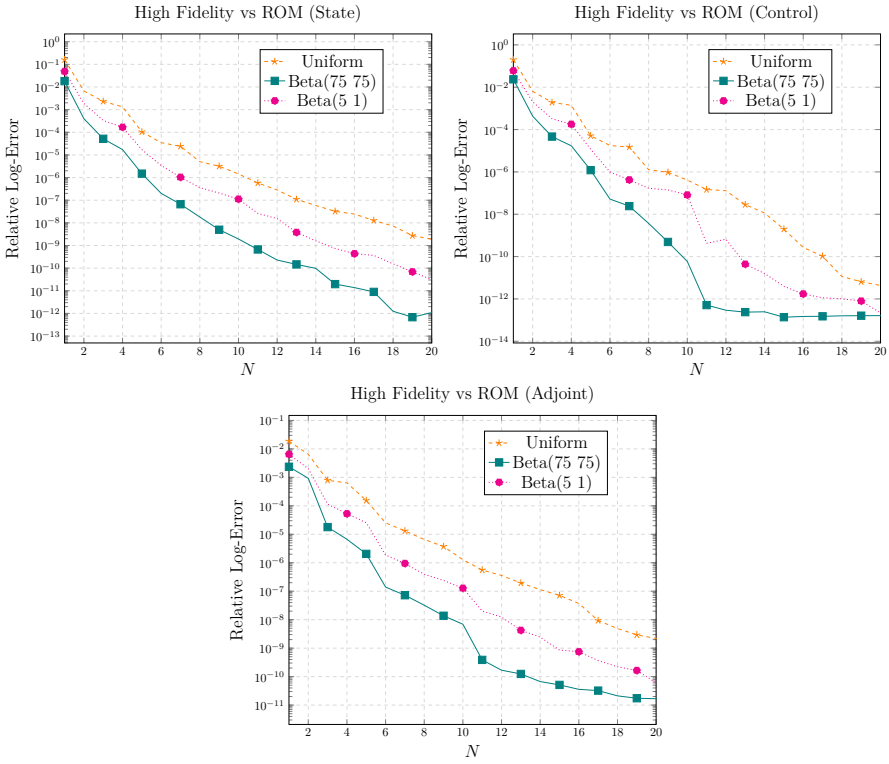


Figure 6.9: *Stochastic Pollutant Control.* Averaged relative log-error for the variables for different distribution fixed the uniform quadrature rule. *Top left.* State variable for uniform, Beta(75,75) and Beta(5,1) samplings. *Top right.* Control variable for uniform, Beta(75,75) and Beta(5,1) samplings. *Bottom row.* Adjoint variable for uniform, Beta(75,75) and Beta(5,1) samplings.

For this reason, we choose to put them all equal to $\frac{1}{N_{max}}$ focusing on the effects of the sampling strategy. For example, we compared the action of standard POD w.r.t. w-POD with the distribution $\rho(\boldsymbol{\mu})$ from which the parameters are picked to build the snapshots covariance matrix. Figure 6.9 shows the average relative log-error for the state, the adjoint and the control variables: the sampling strategy for the error analysis is the same used in the offline phase. This choice leads to satisfactory performances for all the samplings. However, it is clear that the distribution information is crucial: indeed, if we are provided by a previous knowledge on the parametric distribution a lower number of basis functions can be employed to reach more accurate results. It is the case of $\boldsymbol{\mu} \sim$

Beta(75,75). The Beta distribution is defined through the Beta function

$$B(\alpha, \beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$$

and the probability density function

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}. \tag{6.8}$$

The Beta(75,75) is very picked and informative distribution. Thus, weighting the covariance matrix through this distribution will help a lot in the accuracy of the reduced system. Indeed, the relative log-error is around 10^{-5} exploiting only $N = 4$ basis for each variable, i.e. exploiting a global reduced space of dimension $5N = 20$. This allowed us to gain two orders of magnitude in accuracy w.r.t. standard POD approaches. Furthermore, the advantages of using less bases relies in a actual gain of computational time that can be exploited to evaluate statistical quantities in UQ analysis. We remark that OCP($\boldsymbol{\mu}$)s are complicated and time consuming problems and, thus, we propose a very convenient tool to reduced the computational costs related to the optimality systems. As already said, the same conclusions are valid even in the nonlinear case as presented in [33].

This Chapter focused on how important is to employ all the information available in order to make the model more reliable and to better study the input-output structure of OCP($\boldsymbol{\mu}$) for several parameters. The next Chapter is somehow related to this need of better understanding the relation between the control action and the optimal state. We will discuss such a bound in a very complex framework: bifurcating nonlinear PDE($\boldsymbol{\mu}$)s.

Chapter 7

Parametric Optimal Control Problems to Steer Bifurcations in Nonlinear Governing Equations

This Chapter deals with OCP($\boldsymbol{\mu}$)s as a strategy to steer bifurcating solution of nonlinear PDE($\boldsymbol{\mu}$)s towards a desired configuration, as presented in [118]. Indeed, in the nonlinear context, given a parameter $\boldsymbol{\mu}$, many solutions may co-exist. Thus, OCP($\boldsymbol{\mu}$)s can be interpret not only as a tool to change the solution, but also its stability features. In Section 7.1 we provide a general introduction to the problem, while in Section 7.2 we will specify the properties of bifurcating Navier-Stokes equations (NSE). This problem will then proposed in optimal control framework in Section 7.3. Finally, some results on the application of reduced order techniques in this specific setting will be presented in Section 7.4.

7.1 OCP($\boldsymbol{\mu}$)s for Bifurcating Systems

This Section aims at introducing the problem of multiple solutions for nonlinear PDE($\boldsymbol{\mu}$)s. First of all, we will provide some motivations about the interest in the analysis of such problems in Section 7.1.1. Then, in Section 7.1.2, we will be more specific w.r.t. the problem we are dealing with, introducing the concept of stability and the role of the eigenvalue analysis in the loss of uniqueness of the solution.

7.1.1 Motivations

Nonlinear PDE($\boldsymbol{\mu}$)s are ubiquitous in many application fields, from continuum mechanics to quantum mechanics, passing through fluid dynamics. Despite their important role, given a parameter $\boldsymbol{\mu}$, a nonlinear PDE($\boldsymbol{\mu}$) may have a not unique solution. Namely, as opposed to linear cases, a nonlinear PDE($\boldsymbol{\mu}$) might loose the smooth parametric dependence and a slight change in the parameter space can suddenly alter the behaviour of the solution. We will say that this PDE($\boldsymbol{\mu}$) experiences *bifurcations* [9, 32, 140]. Bifurcations may arise in several models, such as the Von Kármán plate model for buckling [159, 18, 24, 117], the Gross-Pitaevskii equation for Bose-Einstein condensates [102, 81, 36, 116] and the NSE [124, 120, 119]. In this contribution, we will focus on the latter application.
The parameter for which the system presents multiple solutions will be called

bifurcation point and will be denoted by $\boldsymbol{\mu}^*$. To visualize the bifurcating phenomena a plot of a scalar output of the solution w.r.t. the parameter can be analysed. A classical example in NSE is the *pitchfork bifurcation diagram* [140], where, from the critical point $\boldsymbol{\mu}^*$, three different solutions arise. In this case, from the central *branch*, two symmetric branches coexist switching the stability properties of the system at hand. Namely, they inherit the stability feature from the central branch: an example of bifurcation diagram is postponed in Figure 7.4. It is clear that in this context some bifurcating configurations might be undesirable and inconvenient. Specifically, we will treat the problem of the Coanda effect [154] in sudden-expansion channel flows. It is a phenomenon related to the tendency of a fluid jet to be attracted to a nearby surface. Namely, a solution symmetric jet and an upper (or lower) wall-hugging non-symmetric jet coexist. The latter can represent an issue under the medical viewpoint as we will specify later in the Chapter. Thus, in this case, it would be of interest to steer the system towards the branch of symmetric solutions. For this reason, we are going to employ OCP($\boldsymbol{\mu}$)s to reach the convenient configuration of a straight jet, interpreting it as an *attractor* towards a desired profile somehow related to the most convenient branch configuration.

7.1.2 Problem Formulation

First of all, we remark that the continuous formulation we are dealing with is the one briefly introduced in Remark 1.2.1. However, for the sake of clarity, we will specify all the needed quantities and notations. First of all, let

$$G(y;\boldsymbol{\mu}) = f, \tag{7.1}$$

be a steady nonlinear PDE($\boldsymbol{\mu}$) we are provided with. Here, $y := y(\boldsymbol{\mu}) \in Y$ and $f \in Y^*$, following the notation we have adopted for steady problems. Namely, for $\boldsymbol{\mu} \in \mathcal{D}$, the state equation has the following form $G(y;\boldsymbol{\mu}) = E_{n\ell}(y;\boldsymbol{\mu}) + E_\ell(y;\boldsymbol{\mu})$, where $E_\ell \in \mathcal{L}(Y, Y^*)$ and $E_{n\ell}$ represent the linear and nonlinear contributions to the system, respectively. We are already familiar with the concept of OCP($\boldsymbol{\mu}$)s which are built to make y the most similar to a solution profile $y_\mathrm{d} := y_\mathrm{d}(\boldsymbol{\mu}) \in Y_\mathrm{obs} \supseteq Y$. As usual, the controlled equation will have the following form

$$\mathcal{E}(y, u;\boldsymbol{\mu}) := \; G(y;\boldsymbol{\mu}) - C(u) - f = 0.$$

The reader may refer to Chapter 1 to a more extended description of the setting we are working in. In this specific context of the bifurcating phenomena, we are trying to exploit the action of $C(u)$ to change the solution y. To this purpose, we define the following OCP($\boldsymbol{\mu}$): given a $\boldsymbol{\mu} \in \mathcal{D}$, find the pair $(y, u) \in Y \times U$ which solves

$$\min_{y\in Y, u\in U} J(y, u; y_\mathrm{d}) \text{ subject to } \mathcal{E}(y, u;\boldsymbol{\mu}) = 0, \tag{7.2}$$

where $J : Y \times U \times Y_{\text{obs}} \to \mathbb{R}$ is

$$J(y, u; y_{\text{d}}) := \frac{1}{2}\|y - y_{\text{d}}\|^2_{Y_{\text{obs}}} + \frac{\alpha}{2}\|u\|^2_U, \tag{7.3}$$

with $\alpha \in (0, 1]$ as a penalization parameter. We will later see how the penalization parameter play a very important role in the control of bifurcations. Also in this setting, the problem can be tackled through Lagrangian argument. Thus, considering $z \in Y$ as adjoint variable and calling $\boldsymbol{X} := (y, u, z)$, the optimality system reads as (1.10) in weak form or equivalently in strong form as (1.11). As already did in Section 1.1, the problem can be written in compact form as: given $\boldsymbol{\mu} \in \mathcal{D}$, find $\boldsymbol{X} \in \mathbb{X} := Y \times U \times Y$ such that

$$\mathcal{G}(\boldsymbol{X}; \boldsymbol{\mu}) = \mathcal{F}, \tag{7.4}$$

with

$$\mathcal{G}(\boldsymbol{X}; \boldsymbol{\mu}) := \begin{bmatrix} y + D_y\mathcal{E}(y, u; \boldsymbol{\mu})^*(z) \\ \alpha u - C^*(z) \\ G(y, \boldsymbol{\mu}) - C(u) \end{bmatrix} \quad \text{and} \quad \mathcal{F} := \begin{bmatrix} y_{\text{d}} \\ 0 \\ f \end{bmatrix}.$$

The nonlinear system, even considering the only contribution of the state equation, might lead to many configurations for a given parameter $\boldsymbol{\mu}$. The loss of uniqueness might transfer to the OCP($\boldsymbol{\mu}$), since its local well-posedness strictly relies on assumptions (viii) and (ix) of Section 1.1. Indeed, they fail when a bifurcation occurs. In this setting, we can define the *solution branches*: i.e. multiple parametric dynamics of problem (7.4) w.r.t. the value of the parameter $\boldsymbol{\mu}$. We denote by k the number of branches, and by $\mathbb{M}_i$, for $i = 1, \dots, k$, the set of parametric solutions for each i-th branch. Thus, the solution manifold is of the following form:

$$\mathbb{M} := \bigcup_{i=1}^{k} \{\boldsymbol{X}(\boldsymbol{\mu}) \in \mathbb{M}_i \mid \boldsymbol{\mu} \in \mathcal{D}\}. \tag{7.5}$$

In order to numerically solve the problem (7.4), we use the FE formulation introduced in Remark 2.2.1. We recall that the nonlinear problem can be treated through the employment of a Newton's method over the global residual of the optimality system

$$\mathsf{R}(\mathsf{X}; \boldsymbol{\mu}) := \mathsf{G}(\mathsf{X}; \boldsymbol{\mu}) - \mathsf{F} = 0, \tag{7.6}$$

solving

$$\mathsf{X}^{j+1} = \mathsf{X}^j + \mathsf{Jac}(\mathsf{X}^j; \boldsymbol{\mu})^{-1}(\mathsf{F} - \mathsf{G}(\mathsf{X}^j; \boldsymbol{\mu})), \qquad j \in \mathbb{N}, \tag{7.7}$$

up to a convergence criterion. Here, the matrix G and the vector F represent the left and the right hand sides of system (7.4) after a FE approximation. We recall that, at each iteration of the Newton's method the Jacobian matrix has a saddle point structure

$$\mathsf{Jac}(\mathsf{X}^j; \boldsymbol{\mu}) = \begin{bmatrix} \mathsf{A} & \mathsf{B}^T \\ \mathsf{B} & 0 \end{bmatrix}. \tag{7.8}$$

The reader interested in the detailed matrix formulation may refer, once again, to Remark 2.2.1 and to Section 2.2. We still want to address that, in order to guarantee the solvability of the system, the Brezzi inf-sup condition (2.22) must be verified. We report it below, for the sake of clarity:

$$\beta^{\mathcal{N}}(\boldsymbol{\mu}) := \inf_{0\neq \mathsf{z}} \sup_{0\neq \mathsf{x}} \frac{\mathsf{z}^T \mathsf{B}\mathsf{x}}{\|\mathsf{x}\|_{Y\times U}\|\mathsf{z}\|_Y} \geq \hat{\beta}^{\mathcal{N}}(\boldsymbol{\mu}) > 0, \tag{7.9}$$

where $\mathsf{x} = \begin{bmatrix} \mathsf{y} \\ \mathsf{u} \end{bmatrix}$. We recall that the well-posedness of the linearized system is provable once assuming the coincidence between state and adjoint spaces. The next Section specifies the features of a bifurcation problem, introducing the spectral analysis needed to describe the procedure employed for the reconstruction of the bifurcation diagram related to the application we will present later on in the Chapter.

7.1.3 Bifurcations and Stability Analysis

This Section provides tools to study the solution of general nonlinear PDE($\boldsymbol{\mu}$)s also in a optimal control framework. We remark that the well-posedness of OCP($\boldsymbol{\mu}$)s relies on the assumptions presented in Section 1.1.1. Most of all, hypotheses (vii)-(ix) ensure the applicability of the Implicit Function Theorem [9, 38] which guarantees the local invertibility of a nonlinear problem. Hence, when they are verified, if the parameter $\boldsymbol{\mu}$ slightly changes, the solution remains stable and unique. On the contrary, when it is not the case, the model bifurcates. Let us give a proper definition for a *bifurcation point* $\boldsymbol{\mu}^*$. For the sake of clarity, we will restrict ourselves to the case $f = 0$ (that is equivalent to the inclusion of the forcing term in the left hand side of the expression) [9].

Definition 7.1 A parameter value $\boldsymbol{\mu}^* \in \mathcal{D}$ is a *bifurcation point* for (7.1) from the solution $y^* := y(\boldsymbol{\mu}^*)$, if there exists a sequence $(y_n, \boldsymbol{\mu}_n) \in Y \times \mathcal{D}$, with $y_n \neq y^*$, such that

- $G(y_n; \boldsymbol{\mu}_n) = 0$
- $(y_n, \boldsymbol{\mu}_n) \to (y^*, \boldsymbol{\mu}^*)$.

In other words, in nonlinear analysis, a bifurcation phenomena is a necessary condition in the failure of the Implicit Function Theorem. Hence, the following holds.

Proposition 7.2 *A necessary condition for $\boldsymbol{\mu}^*$ to be a bifurcation point for G is that the partial derivative $D_y G(y^*; \boldsymbol{\mu}^*)$ is not invertible.*

Now that we have clear in mind that multiple configurations may arise in the nonlinear context, it still remains to handle the *stability* property. Indeed, as briefly introduced in Section 7.1.1, the stability of the unique solution can be inherited by other ones, totally changing the expected behaviour of the system. To have information about this, a common choice is to perform a spectral analysis of the problem. In analogy with the theory for ordinary differential equation (ODEs) [88, 83, 140], the stability of a solution to (7.1) can be understood by means of the sign of the eigenvalues of the linearization of (7.1) around the solution we are interested in, say $\hat{y} = y(\hat{\boldsymbol{\mu}})$. Thus, for a fixed $\hat{\boldsymbol{\mu}}$, one can solve the following eigenvalue problem

$$D_y G(\hat{y}; \hat{\boldsymbol{\mu}}) y_e = \rho_{\hat{\boldsymbol{\mu}}} y_e, \tag{7.10}$$

for the eigenvalue-eigenvector pair $(\rho_{\hat{\boldsymbol{\mu}}}, y_e)$. We remark that the solution is *stable* when a small perturbation in the parameter $\boldsymbol{\mu}$ leads to a dynamic that remains in a neighborhood of the considered solution. In this case, y will be called a *stable solution*. Now, in analogy with ODEs stability theory, when the real part of $\rho_{\hat{\boldsymbol{\mu}}}$, i.e. $\Re(\rho_{\hat{\boldsymbol{\mu}}})$, is positive, we observe an exponentially divergent dynamical behaviour, while $\Re(\rho_{\hat{\boldsymbol{\mu}}}) < 0$ produces small perturbations of the solution. Thus, in order to have a stable solution, all the eigenvalues must have negative real parts. We now generalize this concepts to OCP($\boldsymbol{\mu}$)s where the analysis is more complicated due to the presence of the adjoint equation. Indeed, the saddle point optimization structure is high indefinite and a standard sign-analysis is no longer possible, see e.g. [21, 22, 23]. Nonetheless, we defined an analogous eigenvalue problem for the optimality system (7.4), to investigate its spectral properties:

$$D_{\boldsymbol{X}} \mathcal{G}(\hat{\boldsymbol{X}}; \hat{\boldsymbol{\mu}}) \boldsymbol{X}_e = \sigma_{\hat{\boldsymbol{\mu}}} \boldsymbol{X}_e, \tag{7.11}$$

where $\hat{\boldsymbol{X}} = \boldsymbol{X}(\hat{\boldsymbol{\mu}})$ is the solution of which we are investigating the stability property and $(\sigma_{\hat{\boldsymbol{\mu}}}, \boldsymbol{X}_e)$ is the global eigenvalue-eigenvector pair. We will call (7.10) the *state eigenvalue problem* and (7.11) the *global eigenvalue problem*.
We now want to characterized the local non-invertibility of nonlinear OCP($\boldsymbol{\mu}$)s through the *continuous Babuška inf-sup stability* condition. Indeed, we say that the nonlinear optimality system is well-posed when there exists an inf-sup constant $\hat{\beta}_{Ba} > 0$ such that

$$\beta_{Ba}(\boldsymbol{\mu}) = \inf_{\boldsymbol{X} \in \mathbb{X}} \sup_{\boldsymbol{Y} \in \mathbb{X}} \frac{\langle D_{\boldsymbol{X}} \mathcal{G}[\hat{\boldsymbol{X}}](\boldsymbol{X}; \boldsymbol{\mu}), \boldsymbol{Y} \rangle_{\mathbb{X}\mathbb{X}^*}}{\|\boldsymbol{X}\|_{\mathbb{X}} \|\boldsymbol{Y}\|_{\mathbb{X}}} \geq \hat{\beta}_{Ba}(\boldsymbol{\mu}) \qquad \forall\, \boldsymbol{\mu} \in \mathcal{D}, \tag{7.12}$$

and

$$\inf_{\boldsymbol{Y} \in \mathbb{X}} \sup_{\boldsymbol{X} \in \mathbb{X}} \frac{\langle D_{\boldsymbol{X}} \mathcal{G}[\hat{\boldsymbol{X}}](\boldsymbol{X}; \boldsymbol{\mu}), \boldsymbol{Y} \rangle_{\mathbb{X}\mathbb{X}^*}}{\|\boldsymbol{X}\|_{\mathbb{X}} \|\boldsymbol{Y}\|_{\mathbb{X}}} > 0 \qquad \forall\, \boldsymbol{\mu} \in \mathcal{D}. \tag{7.13}$$

When one of the two condition fails, a bifurcation occurs. At the discrete level, since $\mathbb{X}^{\mathcal{N}} \subset \mathbb{X}$, we need to re-formulate the *discrete Babuška inf-sup stability* as

follows: the problem is well-posed if there exists a constant $\hat{\beta}^{\mathcal{N}}_{Ba}(\boldsymbol{\mu}) > 0$ such that

$$\beta^{\mathcal{N}}_{Ba}(\boldsymbol{\mu}) = \inf_{\mathsf{X}\neq 0} \sup_{\mathsf{Y}\neq 0} \frac{\mathsf{Y}^T \,\mathsf{Jac}\, \mathsf{X}}{\|\mathsf{X}\|_{\mathbb{X}^{\mathcal{N}}} \|\mathsf{Y}\|_{\mathbb{X}^{\mathcal{N}}}} \geq \hat{\beta}^{\mathcal{N}}_{Ba}(\boldsymbol{\mu}) \qquad \forall\, \boldsymbol{\mu} \in \mathcal{D}. \tag{7.14}$$

It is clear that the surjectivity condition (7.13) is no longer needed for the discrete inf-sup stability (7.14). Indeed, relation (7.14) translates with the invertibility of the Jacobian matrix. At the duscrete level, condition (7.13) would require injectivity of Jac^T, that is equivalent to (7.14) since it is a square matrix. It is clear that for $\boldsymbol{\mu}^*$ these conditions do not hold true. We now have all the ingredients to introduce the procedure we developed to numerically detect multiple solution branches. In this contribution, we consider the first component μ of the parameter $\boldsymbol{\mu} \in \mathcal{D} \subset \mathbb{R}^D$ is the one producing a bifurcating feature in the model. To visualize a single branch we fixed all the $P-1$ remaining parameters. The building blocks, combined in the procedure represented in Algorithm 1, are

- a Newton's method, as the nonlinear solver,
- a Galerkin FE method, as the discretization phase,
- a simple continuation method, as bifurcations path tracer,
- a generalized eigenvalue problem, as the stability detector.

Algorithm 1 A pseudo-code for the reconstruction of a branch

$\mathsf{X}_0 = \mathsf{X}_{guess}$ ▷ Initial guess
for $\boldsymbol{\mu}_j \in \mathcal{P}_K$ **do** ▷ Continuation loop
  $\mathsf{X}_j^{(0)} = \mathsf{X}_{j-1}$ ▷ Continuation guess
  while $||\mathsf{R}(\mathsf{X}_j^{(i)}; \boldsymbol{\mu}_j)|| > \epsilon$ **do** ▷ Newton's method
    $\mathsf{Jac}(\mathsf{X}_j^{(i)}; \boldsymbol{\mu}_j)\delta\mathsf{X} = \mathsf{R}(\mathsf{X}_j^{(i)}; \boldsymbol{\mu}_j)$ ▷ Galerkin FE method
    $\mathsf{X}_j^{(i+1)} = \mathsf{X}_j^{(i)} - \delta\mathsf{X}$
  end while
  $\mathsf{Jac}_\mathsf{y}(\mathsf{y_j}; \boldsymbol{\mu}_\mathsf{j})\mathsf{y_e} = \rho_{\boldsymbol{\mu}_\mathsf{j}} \mathsf{V_y y_e}$ ▷ State eigenproblem
  $\mathsf{Jac}(\mathsf{X_j}; \boldsymbol{\mu}_\mathsf{j})\mathsf{X_e} = \sigma_{\boldsymbol{\mu}_\mathsf{j}} \mathsf{V X_e}$ ▷ Global eigenproblem
end for

First of all, we chose the branch to approximate and an initial guess which will let the solver converge to the expected configuration behaviour. Thus, to reconstruct the chosen branch, we define a ordered finite subset of the parameter space $\mathcal{D}_K = [\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K] \subset \mathcal{D}$, where the ordered is induced by the first component of the parameter. The bifurcation behaviour of the model is followed

assigning the solution for a given parameter $\boldsymbol{\mu}_{j-1}$ as the initial guess in the nonlinear solver for the next iteration in $\boldsymbol{\mu}_j$. Here, we employed the simpler variant of a continuation method [8], with $\mathcal{D}_K$ fixed a priori. This method have been successfully applied in pitchfork like bifurcation detection, see e.g. [117, 116]. To solve the nonlinear problem at each parametric instance, we employed the Newton-Kantorovich method [38] combined with the Galerkin FE method repeatedly applied until a residual convergence criterion is satisfied with tolerance ϵ. We now clarify some notations in Algorithm 1: we denote with $\mathsf{Jac}_{\mathsf{y}}(\hat{\mathsf{y}}, \hat{\boldsymbol{\mu}})$ the Jacobian of the controlled state equation (7.1) and with V and V_{y} the scalar product matrices of the global optimization variable and of the state variable, respectively. Furthermore, we denote with $\delta\mathsf{X}$ the difference between two solutions derived from the Newton's method. In the end, we study the stability properties of the solution X_j to (7.4) for the parameter $\boldsymbol{\mu}_j$ by means of the generalized eigenproblems for the controlled state equation and the optimality system.

7.2 Bifurcating NSE: the Coanda Effect

We now take into consideration a bifurcating phenomenon deriving from NSE in a sudden-expansion channel flow problem. Let us consider the physical domain Ω depicted in Figure 7.1. In this geometrical setting, when the fluid presents high viscosity features, it has a symmetric structure w.r.t. the horizontal axis. Moreover, a pair symmetric of vortexes, the Moffatt eddies [106], arise downstream of the expansion. Now imagine to deal with a parametric viscosity μ. Lowering the value of μ, the inertial forces of the system prevail and the two eddies break their symmetry due to a non-uniform decrease of the pressure field along the vertical coordinate. Indeed, when the parametric critical value μ^* is reached, one of the vortexes expands while the other shrinks leading to an asymmetric jet. Namely, the system provides multiple solutions for the same value of $\mu < \mu^*$. In this context, we will deal with two different configurations:

- the *symmetric solution*, a physically unstable configuration with a symmetric jet flow;
- the *asymmetric solution*, a physically stable configuration with a wall-hugging jet.

These solutions, depicted in Figure 7.2, coexist for $\mu < \mu^*$ and belong to different branches intersecting μ^*: this gives rise to the pitchfork bifurcation structure as represented in Figure 7.4 (we postpone the analysis of the plot in the next Section). We now want to specify the problem properties in terms of mathematical formulation and stability analysis: this will be of utmost importance to formulate proper OCP($\boldsymbol{\mu}$)s to deal with bifurcations.

7.2.1 The Uncontrolled NSE

Let us describe the uncontrolled system we will consider from now on. The physical domain is Ω, a two-dimensional planar straight channel with a narrow inlet and a sudden expansion, represented in Figure 7.1. This domain can be interpret as a simplification of the left atrium together with the mitral valve. The inflow and outflow boundary conditions apply to $\Gamma_{\text{in}} = \{0\} \times [2.5, 5]$ and $\Gamma_{\text{out}} = \{50\} \times [0, 7.5]$, respectively. We denote with $\Gamma_{\text{wall}} = \Gamma_{\text{D}} \cup \Gamma_0$ the boundaries representing the walls, where $\Gamma_{\text{D}} = \{\{0\}\times[0, 2.5]\}\cup\{\{0\}\times[5, 7.5]\}$ and $\Gamma_0 = \partial\Omega\setminus\{\Gamma_{\text{in}}\cup\Gamma_{\text{D}}\cup\Gamma_{\text{out}}\}$. We now focus on the parametrized steady and incompressible NSE for a viscous flow in Ω.

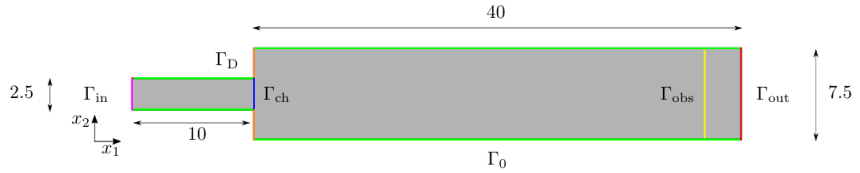


Figure 7.1: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* Domain Ω for the uncontrolled system: straight channel with a narrow inlet.

The problem reads:

$$\begin{cases} -\mu\Delta v + v \cdot \nabla v + \nabla p = 0 & \text{in } \Omega, \\ \text{div}(v) = 0 & \text{in } \Omega, \\ v = v_{\text{in}} & \text{on } \Gamma_{\text{in}}, \\ v = 0 & \text{on } \Gamma_{\text{wall}}, \\ -pn + (\mu\nabla v)n = 0 & \text{on } \Gamma_{\text{out}}, \end{cases} \tag{7.15}$$

where $v = (v_{x_1}, v_{x_2})$ and p are, respectively, the velocity of the fluid and the pressure fields normalized over a constant density and $\mu \in [0.5, 2.]$ is the kinematic viscosity. We applied Neumann boundary conditions on Γ_{out} with outer normal n and no-slip (homogeneous) Dirichlet boundary condition on Γ_{wall} while a non-homogeneous Dirichlet boundary condition v_{in} is provided at the inlet Γ_{in}. The inlet flow has the following form:

$$v_{\text{in}}(x_2) = \begin{bmatrix} 20(5 - x_2)(x_2 - 2.5) \\ 0 \end{bmatrix}.$$

We already declared that the bifurcating phenomenon depends on how much the flow behaves as viscous. As an index of this property, we introduce the dimensionless Reynolds number, representing the ratio between inertial and viscous forces. It is given by $\text{Re} = Uh/\mu$, where U and h are the characteristic

velocity (i.e., the maximum inlet velocity, $U = 31.25$) and the characteristic length of Ω (i.e., the length of the inlet section, $h = 2.5$), respectively. In the considered parameter space, the Re varies in the interval [39.0, 156.0]. Indeed, changing the value of μ will affect the Re, thus, we will refer to both indistinctly while commenting the numerical results. For this particular test case, a unique solution exists up to $\mu^* \approx 0.96$. As already specified, below that value, we experience a pitchfork bifurcation. Let $V = \left[H^1(\Omega)\right]^2$, $V_{\text{in}} = \{v \in V \mid v = v_{\text{in}} \text{ on } \Gamma_{\text{in}}, v = 0 \text{ on } \Gamma_{\text{wall}}\}$, $V_0 = \{v \in V \mid v = 0 \text{ on } \Gamma_{\text{in}} \cup \Gamma_{\text{wall}}\}$ and $P = L^2(\Omega)$ be the function spaces for velocity and pressure, respectively. Problem (7.15) can be recast in weak formulation as: given $\mu \in \mathcal{D}$, find $v \in V_{\text{in}}$ and $p \in P$ such that

$$\begin{cases} \mu \displaystyle\int_\Omega \nabla v \cdot \nabla \psi \, d\Omega + \int_\Omega (v \cdot \nabla v)\, \psi \, d\Omega - \int_\Omega p \text{div}(\psi) \, d\Omega = 0 & \forall\, \psi \in V_0, \\ \displaystyle\int_\Omega \pi \text{div}(v) \, d\Omega = 0 & \forall\, \pi \in P, \end{cases} \tag{7.16}$$

or, equivalently, as: given $\mu \in \mathcal{D}$, find $v \in V_{\text{in}}$ and $p \in P$ such that

$$\begin{cases} a(v, \psi; \mu) + s(v, v, \psi) + b(\psi, p) = 0 & \forall\, \psi \in V_0, \\ b(v, \pi) = 0 & \forall\, \pi \in P, \end{cases} \tag{7.17}$$

where

$$\begin{aligned} a(v, \psi; \mu) &= \mu \int_\Omega \nabla v \cdot \nabla \psi \, d\Omega && \forall\, v, \psi \in V, \\ b(v, p) &= -\int_\Omega \text{div}(v)\, p \, d\Omega && \forall\, v \in V,\ \forall\, p \in P, \\ s(v, \bar{v}, \psi) &= \int_\Omega (v \cdot \nabla \bar{v})\, \psi \, d\Omega && \forall\, v, \bar{v}, \psi \in V. \end{aligned} \tag{7.18}$$

7.2.2 FE Numerical Approximation

For the NSE model, the FE strategy is totally analogous to the one adopted for the Stokes problem in Section 4.2.2, the reader may refer to [128]. Namely, after building a triangulation over Ω of $2'785$ cells, we employed a $N^y_{FE} = N^v_{FE} + N^p_{FE} = 24'301$ for the state variable $y = (v, p)$. Figure 7.4 shows a bifurcation diagram with all the solution branches of the system (7.17) w.r.t. the value of μ. The value of the depicted function is given by the vertical component of the velocity v_{x_2} in the point $(x_1, x_2) = (14, 4)$. The point, for its position, easily detects that the symmetry breaks around $\mu^* \approx 0.96$, when a pitchfork bifurcation occurs. Moreover, below that value μ^*, there is a change in the stability properties of the model. Indeed, the unique symmetric solution remains stable until μ^*, when it becomes unstable, while the stability is inherited by the asymmetric solutions. Figure 7.2 shows some representative solutions for the

lower (top plots) and middle branch (bottom plots): it is clear that, for the same value $\mu = 0.5$, the solution behaviour is totally different for the two branches. In the asymmetric case, the pressure decreases in the bottom-left corner of the expansion, thus the velocity deflects in a wall-hugging profile. Finally, due to the no-slip boundary conditions, the flux gets back end exits the channel in a non-symmetric way. The stability analysis is performed employing Algorithm 1 only for the linearized uncontrolled state equation (7.1) around the solution of interest. In particular, we studied how the first $N_{eig} = 100$ eigenvalues of (7.10) varying the viscosity μ. The eigenvalues are shown in Figure 7.3 for the stable lower branch (left plot) and unstable middle branch (right plot). The zoom of the plots shows if $\Re(\rho_{\boldsymbol{\mu}})$ crosses the origin or not. Thus, from the considerations of Section 7.1.3, the wall-hugging branch inherits the stability of the solution since in the left plot of Figure 7.3 there is no crossing of negative-real part eigenvalues. Finally, we remark that, even if the asymmetric branch is the stable one, it is inconvenient in medical applications, since it represents a regurgitation that can lead to some issues in measurements. This is the main reason why we exploit OCP($\boldsymbol{\mu}$)s in this framework, which will be the topic of the next Section.

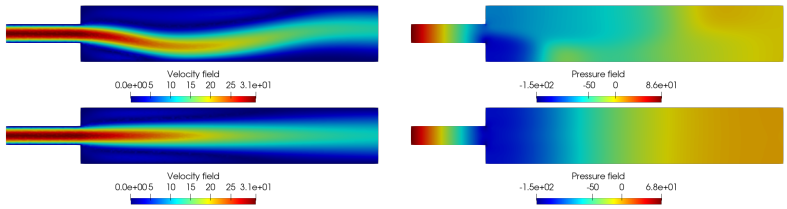


Figure 7.2: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* Representative solutions for the uncontrolled NSE for $\mu = 0.5$, velocity and pressure fields. *Top*. lower branch. *Bottom*. Middle branch.

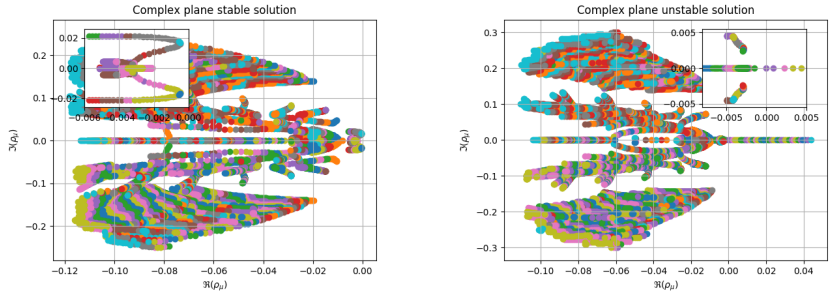


Figure 7.3: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* Eigenvalues of the state eigenproblem in the complex plane for the uncontrolled NSE. *Left*. Stable solution. *Right*. Unstable solution.

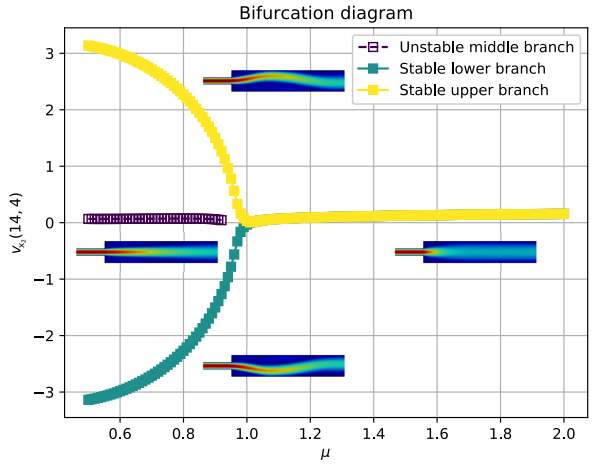
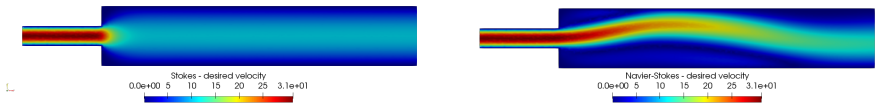


Figure 7.4: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* Bifurcation diagram for the uncontrolled NSE.

7.3 Steering bifurcations through OCP($\boldsymbol{\mu}$)s

We now deal with several OCP($\boldsymbol{\mu}$)s governed by NSE in the setting proposed in Section 7.2.1. The goal is to understand the role of different controls on the solution behaviour when a bifurcation phenomenon occurs and if OCP($\boldsymbol{\mu}$)s can be considered as attractors towards some preferable solutions. Indeed, the loss of uniqueness can represent an issue in several applications. In the context of the mitral valve regurgitation, for instance, a wall-hugging solution might lead to inaccurate experimental measurements by echocardiography.
We remark that, even if we focus on a very specific test case, the presented procedure is general and can be employed in wide variety of other nonlinear applications.
Let us consider the domain Ω shown in Figure 7.1. In this setting, we require the velocity solution $v \in V$ to be the most similar to a desired profile $v_{\mathrm{d}} \in V_{\mathrm{obs}} := [L^2(\Gamma_{\mathrm{obs}})]^2$. The *observation domain* $\Gamma_{\mathrm{obs}} = \{47\} \times [0, 7.5]$ is a line near the end of the channel. For the experiments, we employed two different velocity profiles, which are showed in Figure 7.5: we call them the *symmetric desired profile (or target)* for Figure 7.5a and the *asymmetric desired profile (or target)* for Figure 7.5b. The first is the result of a Stokes system over Ω for $\mu = 1$ with the same boundary conditions of the uncontrolled NSE presented in (7.15). The latter is the physically stable uncontrolled NSE solution for $\mu = 0.49$. Namely, the first profile aims at reaching a globally symmetric configuration with a diffusive outgoing flux, while, the latter can be exploited to achieve the opposite goal.



(a) (b)

Figure 7.5: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* Desired velocity profiles: (a) symmetric profile obtained as Stokes solution for $\mu = 1$; (b) asymmetric profile given by the physically stable Navier-Stokes solution for $\mu = 0.49$.

To steer the bifurcating behaviour towards a preferable solution, we rely on the minimization of the functional

$$J_{\mathrm{NS}}(v, u; v_{\mathrm{d}}) = \frac{1}{2}\|v - v_{\mathrm{d}}\|^2_{V_{\mathrm{obs}}} + \frac{\alpha}{2}\|u\|^2_U, \tag{7.19}$$

where $U := [L^2(\Omega_u)]^2$. The control domain $\Omega_u \subset \overline{\Omega}$ can be a part of the domain or of the boundary. In the experiments, we will analyze how the choice of Ω_u together with the penalization parameter α affects the solution behaviour of the standard NSE. First of all, we will briefly introduce the problem formulation in Section 7.3.1. Then, we will provide the analysis of several OCP($\boldsymbol{\mu}$)s, as follows.

Sec. 7.3.2. A *weak control* is built by changing the flow behaviour at the Neumann boundary. The uncontrolled bifurcating solution is slightly affected by the optimality system.

Sec. 7.3.3. A *strong control* is performed through a distributed forcing term. The classical bifurcating behaviour solution is highly affected by the control variable.

Sec. 7.3.4. We analyse the *penalization parameter* in a control framework acting at the end of the inlet channel. Different values of α will change the uncontrolled system resulting in several interesting configurations.

Sec. 7.3.5. Lastly, we will control different *boundary flux* conditions and this will drastically change the known behaviour of the uncontrolled NSE.

Finally, some final remarks and comparisons concerning the spectral analysis of all the presented test cases are discussed in Section 7.3.6.

7.3.1 OCP($\boldsymbol{\mu}$)s Governed by NSE

In this Section we are going to describe the structure of OCP($\boldsymbol{\mu}$)s constrained to NSE. Indeed, we will specify the structure of Remark 2.2.1 for this specific state equation. Here we report the continuous and the discrete problem formulation, for the sake of clarity. The formulation is similar to the Stokes problem of

Section 1.2.3, however, we here denote the variables in a differnt way, following the notation of [118]. The controlled version of the NSE in Ω reads as:

$$\begin{cases} -\mu\Delta v + v\cdot\nabla v + \nabla p = C(u) & \text{in } \Omega, \\ \operatorname{div}(v) = 0 & \text{in } \Omega, \end{cases} \tag{7.20}$$

together with some proper boundary conditions. As usual, $C : U \to V^*$ is the control operator and can represent either an external forcing term or a boundary term. We recall that if C acts on the whole domain the problem is said *distributed*. However, in this context, when C will only act in a subset of the internal domain, we will talk about *localized control*. Furthermore, we will refer to *Neumann control* and *Dirichlet control*, if $\Omega_u = \Gamma_{\text{out}}$ and $\Omega_u = \Gamma_{\text{D}}$, respectively (as represented in Figure 7.1). The weak formulation of the controlled state equation (7.20) is: given $\mu \in \mathcal{D}$, find $v \in V$, $p \in P$ and $u \in U$ such that

$$\begin{cases} a(v,\psi;\mu) + s(v,v,\psi) + b(\psi,p) = c(u,\psi) & \forall\, \psi \in V_0, \\ b(v,\pi) = 0 & \forall\, \pi \in P, \end{cases} \tag{7.21}$$

where $a(\cdot,\cdot;\mu)$, $b(\cdot,\cdot)$ and $s(\cdot,\cdot,\cdot)$ have been already defined in (7.18) while $c : U \times V \to \mathbb{R}$ is the bilinear form related to the control operator. As we specified in Chapter 1.1.1, the problem can be tackled through a Lagrangian argument and solved thanks to the following optimality system: given $\mu \in \mathcal{D}$, find $\boldsymbol{X} = ((v,p),u,(w,q)) \in \mathbb{X}$ such that

$$\begin{cases} D_v\mathscr{L}_{\text{NS}}(\boldsymbol{X}; v_{\text{d}}, \boldsymbol{\mu})[\varphi] = 0 & \forall\, \varphi \in V_0, \\ D_p\mathscr{L}_{\text{NS}}(\boldsymbol{X}; v_{\text{d}}, \boldsymbol{\mu})[\theta] = 0 & \forall\, \theta \in P, \\ D_u\mathscr{L}_{\text{NS}}(\boldsymbol{X}; v_{\text{d}}, \boldsymbol{\mu})[\tau] = 0 & \forall\, \tau \in U, \\ D_w\mathscr{L}_{\text{NS}}(\boldsymbol{X}; v_{\text{d}}, \boldsymbol{\mu})[\psi] = 0 & \forall\, \psi \in V_0, \\ D_q\mathscr{L}_{\text{NS}}(\boldsymbol{X}; v_{\text{d}}, \boldsymbol{\mu})[\pi] = 0 & \forall\, \pi \in P, \end{cases} \tag{7.22}$$

where $\mathscr{L}_{\text{NS}}(\boldsymbol{X}; v_{\text{d}}, \boldsymbol{\mu})$ is built as described in Section 1.1.1 and $\mathbb{X} := (V_{\text{in}} \times P) \times U \times (V_{\text{in}} \times P)$. The adjoint equation has the following form:

$$\begin{cases} m(v,\varphi) + a(w,\varphi;\mu) + s(\varphi,v,w) \\ \qquad\qquad + s(v,\varphi,w) + b(\varphi,q) = m(v_{\text{d}},\varphi) & \forall\, \varphi \in V_0, \\ b(w,\theta) = 0 & \forall\, \theta \in P, \end{cases} \tag{7.23}$$

and the optimality equation is

$$\alpha n(u,\tau) = c(\tau,w) \quad \forall\, \tau \in U, \tag{7.24}$$

where $m : V \times V \to \mathbb{R}$ and $n : U \times U \to \mathbb{R}$ are defined as in Section 1.2.1: they represent the L^2–product in Γ_{obs} and Ω_u, respectively. Furthermore, the

summation $s(\varphi, v, w) + s(v, \varphi, w)$ is the linearization around v of the trilinear form $s(v, v, \varphi)$. Therefore, the strong formulation for (7.23) and (7.24) reads:

$$\begin{cases} v\mathbb{I}_{\Omega_{\text{obs}}} - \mu\Delta w - v \cdot \nabla w + (\nabla v)^T w + \nabla q = v_{\text{d}}\mathbb{I}_{\Omega_{\text{obs}}} & \text{in } \Omega, \\ \text{div}(w) = 0 & \text{in } \Omega, \\ \alpha u \mathbb{I}_{\Omega_u} = C^* w & \text{in } \Omega, \end{cases} \tag{7.25}$$

where $\mathbb{I}_{\Omega_u}$ and $\mathbb{I}_{\Omega_{\text{obs}}}$ are the indicator functions of the control and observation domains, respectively. In this context, similarly to the one proposed in the Stokes framework in Section 1.2.3, considering $y := (v, p)$ and $z := (w, q)$, we recover the algebraic formulation Remark 2.2.1 once applied the same Taylor-Hood approximation $\mathbb{P}^2$-$\mathbb{P}^1$ to the state y and the adjoint variable z. Moreover, the space U is approximated by $\mathbb{P}^2$ polynomials. Now, specifying the general strategy presented in Remark 2.2.1, we define

$$\mathsf{y} = \begin{bmatrix} \mathsf{v} \\ \mathsf{p} \end{bmatrix}, \quad \mathsf{z} = \begin{bmatrix} \mathsf{w} \\ \mathsf{q} \end{bmatrix}, \quad \mathsf{M}_y = \begin{bmatrix} \mathsf{M}_v & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{and} \quad \mathsf{C} = \begin{bmatrix} \mathsf{C}_v \\ 0 \end{bmatrix}, \tag{7.26}$$

with $\mathsf{v}, \mathsf{p}, \mathsf{w}, \mathsf{q}$ the column vectors of FE coefficients for state and adjoint, velocities and pressures, respectively, and M_v and C_v the mass velocity matrix and the matrix representation of the bilinear form $c(\cdot, \cdot)$. Nonetheless, we can discretize the linearized state equation in the following way

$$\mathsf{E}'_{n\ell}[\mathsf{y}^j] + \mathsf{E}_\ell = \begin{bmatrix} \mathsf{S}[\mathsf{v}^j] & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \mathsf{K} & \mathsf{D}^T \\ \mathsf{D} & 0 \end{bmatrix} = \begin{bmatrix} \mathsf{K} + \mathsf{S}[\mathsf{v}^j] & \mathsf{D}^T \\ \mathsf{D} & 0 \end{bmatrix}, \tag{7.27}$$

where K is the stiffness matrix related to $a(\cdot, \cdot; \mu)$, D is the matrix deriving from the continuity constraint $b(\cdot, \cdot)$ and, finally, $\mathsf{S}[\mathsf{v}^j]$ is the algebraic formulation of $s(v, \cdot, \cdot) + s(\cdot, v, \cdot)$ w.r.t. FE velocity basis functions. We recall that, in the optimality system, the term $\mathsf{D}_{\mathsf{y}}(\mathsf{E}'_{n\ell}[\mathsf{y}]^T)[\mathsf{z}^j]$ defined in (2.19) arises. In this specific context, we call $s_{\text{ad}}(v, w, \varphi)$ the *adjoint operator* of the linearized trilinear form $s(v, v, \cdot)$ around the state velocity v. The related matrix formulation is $\mathsf{S}[\mathsf{v}^j]^T$. When evaluating the Jacobian, the linearization of $s_{\text{ad}}(w, v, \varphi)$ is performed both w.r.t. w and w.r.t. v. This process will lead to

$$\mathsf{D}_{\mathsf{y}^j}(\mathsf{E}'_{n\ell}[\mathsf{y}^j]^T)[\mathsf{z}^j] = \begin{bmatrix} \mathsf{D}_{\mathsf{v}}(\mathsf{S}[\mathsf{v}^j]^T)([\mathsf{w}^j]) & 0 \\ 0 & 0 \end{bmatrix}, \tag{7.28}$$

where $\mathsf{D}_{\mathsf{v}}(\mathsf{S}[\mathsf{v}^j]^T)([\mathsf{w}^j])$ is given by the form $s_{\text{ad}}(w, \cdot, \cdot)$ applied to the velocity basis functions. Finally, the whole linearized OCP($\boldsymbol{\mu}$) problem reads:

$$\mathsf{Jac}_{\text{NS}}(\mathsf{X}^j; \boldsymbol{\mu}) = \begin{bmatrix} \mathsf{M}_v + \mathsf{D}_{\mathsf{v}}(\mathsf{S}[\mathsf{v}^j]^T)[\mathsf{w}^j] & 0 & 0 & \mathsf{K}^T + \mathsf{S}[\mathsf{v}^j]^T & \mathsf{D}^T \\ 0 & 0 & 0 & \mathsf{D} & 0 \\ 0 & 0 & \alpha\mathsf{M}_u & -\mathsf{C}_v^T & 0 \\ \mathsf{K} + \mathsf{S}[\mathsf{v}^j] & \mathsf{D}^T & -\mathsf{C}_v & 0 & 0 \\ \mathsf{D} & 0 & 0 & 0 & 0 \end{bmatrix}$$

where X optimal solution coefficient vector. Now, calling

$$\mathsf{A} = \begin{bmatrix} \mathsf{M}_v + \mathsf{D}_\mathsf{v}(\mathsf{S}[\mathsf{v}^j]^T)([\mathsf{w}^j]) & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha \mathsf{M}_u \end{bmatrix} \quad \text{and} \quad \mathsf{B} = \begin{bmatrix} \mathsf{K} + \mathsf{S}[\mathsf{v}^j] & \mathsf{D}^T & -\mathsf{C}_v \\ \mathsf{D} & 0 & 0 \end{bmatrix}, \tag{7.29}$$

it is clear that the Jacobian matrix presents the saddle point structure of equation (2.20). We remark that we are dealing with a nested saddle point structure as presented for the time-dependent Stokes case in Section 2.3.2. We recall that, besides the babi ška inf-sup constant 7.14, for a given $\mu \neq \mu^*$, thanks to the Taylor-Hood discretization, the following Brezzi inf-sup condition holds, i.e.

$$\beta_{\mathrm{NS}}^{\mathcal{N}}(\mu) := \inf_{\mathsf{p}\neq 0} \sup_{\mathsf{v}\neq 0} \frac{\mathsf{p}^T \mathsf{D}\mathsf{v}}{\|\mathsf{v}\|_V \|\mathsf{p}\|_P} \geq \hat{\beta}_{\mathrm{NS}}^{\mathcal{N}}(\mu) > 0. \tag{7.30}$$

As already specified, in the next subsections we will show how different OCP($\boldsymbol{\mu}$)s change the uncontrolled problem behaviour presented in Section 7.2. For the sake of notation, from now on, we will use the terms symmetric/asymmetric to denote both the desired velocity target or and a visual representation of the obtained optimal solution. Furthermore, with *natural optimal branch* we refer to the branch obtained by means of Algorithm 1 employing a trivial initial guess. It can be either symmetric or asymmetric: it will be related to the test case. Namely, from a numerical viewpoint, the natural optimal branch is the simplest configuration to achieve by the optimality system with Algorithm 1 even after a random perturbation of the trivial guess. We remark that further branches might exist, however, they are difficult to detect in practice and they need very tailored initial guesses. These kind of branches will be named *non-natural optimal branches*[1]. We guess that the natural branches are a consequence of a *numerical stability* property related to the optimality system itself. Indeed, by definition, OCP($\boldsymbol{\mu}$)s are not physical, since they rely on an "artificial" adjoint variables to change the system behaviour.

7.3.2 Neumann Control: weak steering

The first numerical experiment we deal with is a Neumann control over the boundary Γ_{out}. Here, homogeneous Dirichlet conditions are applied to $\Gamma_{\mathrm{wall}} := \Gamma_0 \cup \Gamma_{\mathrm{D}}$. In this specific context, the optimality system reads: given $\mu \in \mathcal{D}$ find

[1] For the sake of clarity, each branch is extended to $\mu > \mu^*$ with the unique solution.

$\boldsymbol{X} \in \mathbb{X}$ such that

$$\begin{cases} v\mathbb{I}_{\Gamma_{\text{obs}}} - \mu\Delta w - v \cdot \nabla w + (\nabla v)^T w + \nabla q = v_{\text{d}}\mathbb{I}_{\Gamma_{\text{obs}}} & \text{in } \Omega, \\ \nabla \cdot w = 0 & \text{in } \Omega, \\ w = 0 & \text{on } \Gamma_{\text{in}} \cup \Gamma_{\text{wall}}, \\ -qn + (\mu\nabla w)n = 0 & \text{on } \Gamma_{\text{out}}, \\ \alpha u\mathbb{I}_{\Gamma_{\text{out}}} = w\mathbb{I}_{\Gamma_{\text{out}}} & \text{in } \Omega, \\ -\mu\Delta v + v \cdot \nabla v + \nabla p = 0 & \text{in } \Omega, \\ \nabla \cdot v = 0 & \text{in } \Omega, \\ v = v_{\text{in}} & \text{on } \Gamma_{\text{in}}, \\ v = 0 & \text{on } \Gamma_{\text{wall}}, \\ -pn + (\mu\nabla v)n = u & \text{on } \Gamma_{\text{out}}. \end{cases} \tag{7.31}$$

The target velocity v_{d} will always be considered as symmetric. Namely, we want to find the best Neumann boundary condition framework to reach the profile depicted in Figure 7.5a. We analysed several values of the penalization parameter, say $\alpha = 1, 0.1, 0.001, 0.0001$, where, we recall, the greater is the value of α the lower is the strength of the control. Figure 7.6 shows some representative solutions for $\alpha = 0.01$ and $\mu = 0.5$, for state velocity and pressure variables. In this case, the natural optimal branch features asymmetric solutions (Figure 7.6, top). A further non-natural optimal branch made up by symmetric solutions (Figure 7.6, bottom) has been observed. Some other qualitative results of natural optimal and non-natural optimal branches are depicted in Figures 7.7 and 7.8, respectively. From the results, we assert that the Neumann control *weakly* affects the standard system behaviour: indeed, it does not steer the system towards the desired symmetric profile once the bifurcation happened. Thus, the configurations are very similar to the uncontrolled case (refer, for example, to Figure 7.4).

Let us analyse the left plot of Figure 7.7: it depicts the velocity profile magnitude over Γ_{obs} for the highest value of the Reynolds number when following the natural optimal branch. The obtained velocity (marked by an orange line) is indeed different from the desired profile (denoted by a blue line), especially for what concerns peak values. However, we observe that the Neumann control straightens the flux near the end of the channel (compare the orange line to the green line, which represents the uncontrolled asymmetric profile), even when high Reynolds numbers are considered. The resulting profile is similar to the uncontrolled symmetric velocity (red line). The right plot of Figure 7.7 represents the control variable for some values of μ following the natural optimal branch. As expected, the control is stronger for $\mu < \mu^*$ (i.e., when the wall-hugging phenomenon occurs and, thus, a straightening action is necessary), while it remains low in magnitude for $\mu > \mu^*$.

Similarly, the left plot of Figure 7.8 shows the velocity profile magnitude over Γ_{obs} for the highest value of the Reynolds number when considering the non-

natural optimal branch (which is symmetric in this case). The controlled symmetric profile (orange line) coincides with the uncontrolled symmetric profile (red line). Furthermore, from the right plot of Figure 7.8, we can observe that the control variable around the critical μ^* (e.g., $\mu = 1$ and $\mu = 0.95$) is asymmetric: this is due to the need of counteracting the physically stable uncontrolled wall-hugging behaviour. We further remark that, compared to the natural optimal branch, the control variable of the non-natural optimal branch has a lower magnitude. Table 7.1 shows the value of the cost functional (7.19) for several values of μ (rows) and α (columns), following either the natural optimal or non-natural optimal branches. In the first column, we also report the value of the *uncontrolled functional*, i.e. (7.19) evaluated for the uncontrolled velocity v. In all the cases, the non-natural branch presents lower values of the functional compared to the natural branch since the first one is clearly closer to the symmetric target (compare e.g. for $\mu = 0.05$ and $\alpha = 0.001$ the left panels of Figures 7.7 and 7.8). However, when dealing with the natural branch, the optimal control procedure influences the most the cost functional values: for instance, for $\mu = 0.05$ and $\alpha = 0.001$, the cost functional is decreased by 6% on the non-natural branch and by 55% on the natural one w.r.t. the corresponding uncontrolled configuration. Again, this is not unexpected since from Figures 7.7 and 7.8 it is clear that in the natural configuration the control action prevails. We remark that larger values of μ present a negligible value of the cost functional, since the target velocity almost coincides with the uncontrolled velocity. From this preliminary analysis, we can deduce that the optimality performance may vary between the different configurations, thus, it is of utmost importance to deeply analyse all the branches to detect the solution that best recover the desired velocity profile. To study the stability of the optimal solutions, we performed the eigenvalues analysis described in Algorithm 1. We can derive several information from the Figure 7.9: it represents the global eigenvalue problem for the natural branch, against the parameter μ such that $\Re(\sigma_\mu) = [-0.01, 0.01]$. We plot the first $N_{eig} = 100$ eigenvalues of the linearized system (7.11) around the global optimal solution, using a Krylov-Schur algorithm. The plot shows two eigenvalues (highlighted with blue markers) approaching $\Re(\sigma_\mu) = 0$: in this case, we will say that the *shears phenomenon* is occurring. Moreover, the lower is α, the larger is the number of positive eigenvalues, while the negative eigenvalues are lowering except for the negative shear eigenvalue. Furthermore, it is clear form plots 7.9c and 7.9d that positive real eigenvalues cluster in the value of α. In addition, a single eigenvalue (denoted by red markers) approaching zero is clearly visible. The conclusion we can draw from the global eigenvalue analysis is that a much controlled system (i.e. smaller values of α) highly affects the concentration of negative eigenvalues. Unfortunately, no information about the physical stability can be derived by the performed global eigenvalue analysis. Indeed, similar eigenvalues behaviour is observed for both the natural and non-natural branches (only the former being shown here for the sake of brevity).

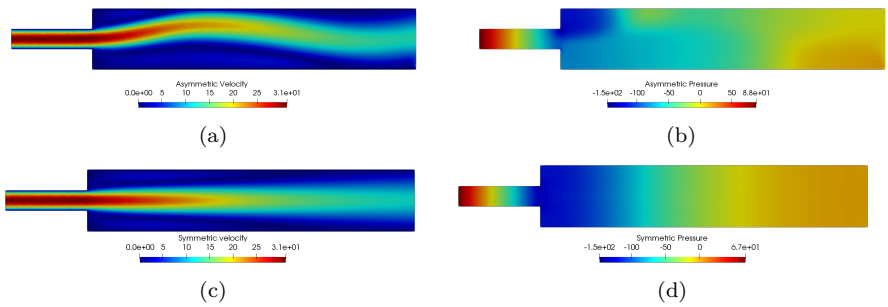


Figure 7.6: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Neumann control). Optimal solutions with $\alpha = 0.01$ and $\mu = 0.5$, belonging to the natural optimal (panels (a) and (b) for state velocity and pressure, respectively) and the non-natural optimal (panels (c) and (d)) branches.

Table 7.1: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* Comparison of the functional value for Neumann control w.r.t. stable and unstable uncontrolled solutions. (Nat.) Natural optimal branch. (n-Nat.) Non-natural optimal branch.

μ	Stable	Unstable	Nat.	n-Nat.	Nat.	n-Nat.	Nat.	n-Nat.	Nat.	n-Nat.
	Uncontrolled		$\alpha = 1$		$\alpha = 0.1$		$\alpha = 0.01$		$\alpha = 0.001$	
2	5.14e–9	5.14e–9	5.13e–9	5.13e–9	5.13e–9	5.13e–9	5.13e–9	5.13e–9	5.07e–9	5.07e–9
1.5	4.38e–6	4.38e–6	4.38e–6	4.38e–6	4.38e–6	4.38e–6	4.37e–6	4.37e–6	4.28e–6	4.28e–6
1	4.10e–3	4.10e–3	4.10e–3	4.10e–3	4.10e–3	4.10e–3	4.08e–3	4.10e–3	3.92e–3	3.92e–3
0.9	3.33e–2	1.63e–2	3.33e–2	1.63e–2	3.30e–2	1.63e–2	3.15e–2	1.63e–2	2.93e–2	1.55e–2
0.8	2.08e–1	6.52e–2	2.07e–1	6.52e–2	2.04e–1	6.51e–2	1.88e–1	6.51e–2	1.70e–1	6.15e–2
0.7	1.01e+0	2.59e–1	1.01e+0	2.59e–1	9.80e–1	2.59e–1	8.63e–1	2.59e–1	7.67e–1	2.43e–1
0.6	4.48e+0	1.70e+0	4.44e+0	1.02e+0	4.15e+0	1.02e+0	3.33e+0	1.02e+0	2.91e+0	9.57e–1
0.5	1.88e+1	3.92e+0	1.83e+1	3.92e+0	1.50e+1	3.92e+0	9.61e+0	3.92e+0	8.54e+0	3.68e+0

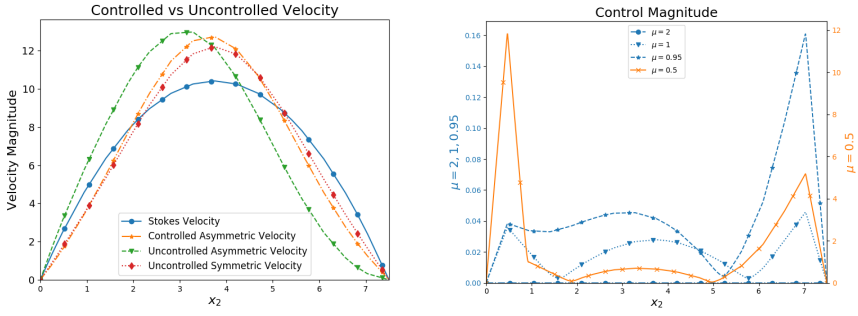


Figure 7.7: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* *Left.* (Neumann control). Comparison of velocity profiles in the controlled and uncontrolled cases for $\alpha = 0.01$, $\mu = 0.5$ on Γ_{obs} w.r.t. the desired profile when following the natural optimal branch. *Right.* representation of control variable evolution for $\alpha = 0.01$, $\mu = 2, 1, 0.95, 0.5$ over Γ_{out} when following the natural optimal branch.

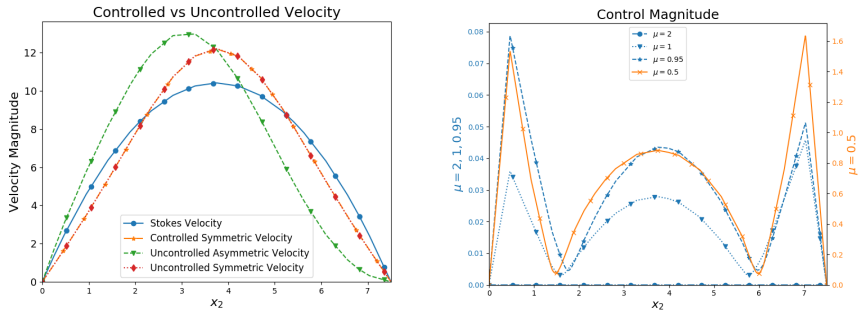


Figure 7.8: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Neumann control). *Left.* Comparison of velocity profiles in the controlled and uncontrolled cases for $\alpha = 0.01$, $\mu = 0.5$ on Γ_{obs} w.r.t. the desired profile when following the non-natural optimal branch. The lines marked by "Controlled Symmetric Velocity" and "Uncontrolled Symmetric Velocity" overlap. *Right.* representation of control variable evolution for $\alpha = 0.01$, $\mu = 2, 1, 0.95, 0.5$ over Γ_{out} when following the non-natural optimal branch.

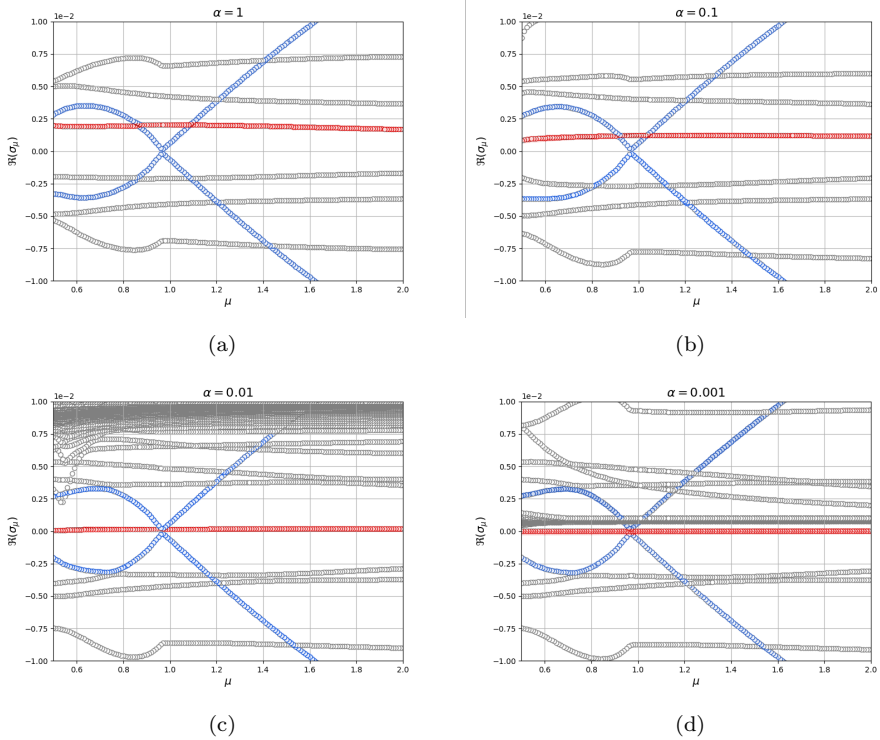


Figure 7.9: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Neumann control). Spectral analysis of Neumann control with $\alpha = 1, 0.1, 0.01, 0.001$.

Therefore, even for the other experiments, we will discuss only the numerical stability represented by natural optimal solution. Summing up, the Neumann control does not fully steer the uncontrolled solutions towards the desired symmetric configuration. However, this will be the case of the following test case, where a stronger control effect is analysed.

7.3.3 Distributed Control: strong steering

This Section deals with a distributed control in $\Omega_u \equiv \Omega$. Also in this case, we consider $\Gamma_{\text{wall}} = \Gamma_0 \cup \Gamma_{\text{D}}$. Given $\mu \in \mathcal{D}$, the optimal solution $\boldsymbol{X} \in \mathbb{X}$ solves the following system:

$$\begin{cases} v\mathbb{I}_{\Gamma_{\text{obs}}} - \mu\Delta w - v \cdot \nabla w + (\nabla v)^T w + \nabla q = v_{\text{d}}\mathbb{I}_{\Gamma_{\text{obs}}} & \text{in } \Omega, \\ \text{div}(w) = 0 & \text{in } \Omega, \\ w = 0 & \text{on } \Gamma_{\text{in}} \cup \Gamma_{\text{wall}}, \\ -qn + (\mu\nabla w)n = 0 & \text{on } \Gamma_{\text{out}}, \\ \alpha u = w & \text{in } \Omega, \\ -\mu\Delta v + v \cdot \nabla v + \nabla p = u & \text{in } \Omega, \\ \text{div}(v) = 0 & \text{in } \Omega, \\ v = v_{\text{in}} & \text{on } \Gamma_{\text{wall}}, \\ v = 0 & \text{on } \Gamma_0, \\ -pn + (\mu\nabla v)n = 0 & \text{on } \Gamma_{\text{out}}. \end{cases} \tag{7.32}$$

For this experiment we expect a stronger action of the control variable that has the possibility to deeply affect the original system. To show this feature, we will steer the system towards either symmetric or asymmetric desired profiles v_{d}:

- *Symmetric target.* First of all, let us focus on Figures 7.10a and 7.10b where two representative solutions of the control variable, defined on the whole domain, are shown. The plots have been obtained for $\mu = 2$ and $\mu = 0.5$, when following the natural optimal branch, which is composed of symmetric solutions. The strong action of the control lead to a more diffusive velocity field w.r.t. the uncontrolled symmetric profile, as represented in the left plot of Figure 7.13, corresponding to the velocity solution slice on the observation domain for $\mu = 0.5$: in this case the controlled velocity (orange line) and the symmetric target (blue line) almost overlap. The right plot of Figure 7.13 shows that a slightly asymmetric control is only required near the critical value μ^* (also compare with Figures 7.10a and 7.10b for the cases $\mu = 2$ and $\mu = 0.5$). Nonetheless, as expected, the control action is higher when the Re value increases. Indeed, for $\mu = 2$ the control exclusively acts in the proximity of Γ_{obs} with a maximum magnitude of $1.8{\cdot}10^{-4}$, while for $\mu = 0.5$ it reaches a value of 1.6 of magnitude

and it is significant all over the domain.
A further non-natural optimal branch exists, and is made of asymmetric solutions. However, it is hardly detectable by Algorithm 1 unless one exploits very tailored guesses in a small neighborhood of μ^*.

- *Asymmetric target.* In contrast with all the previous test cases, here we want to reach an asymmetric target for all $\mu \in \mathcal{D}$. Figures 7.10c and 7.10d show two representative control solutions for $\mu = 2$ and $\mu = 0.5$ when following the natural optimal branch. In this case, it features asymmetric solutions. The action of the control is also visible in the left plot of Figure 7.14, obtained for $\mu = 2$. The control, as expected, pushes the flux over Γ_{obs} towards the domain wall (orange line), avoiding the symmetric profile of the uncontrolled velocity (green line). Namely, also for this other test, the distributed control drives the solution towards the desired velocity. For this reason, the control magnitude is larger when $\mu > \mu^*$, i.e. when the uncontrolled configuration is symmetric. Indeed, in Figure 7.10c the maximum magnitude for the control is 7 and it is reached for $\mu = 2$ in the upper part of the domain. On the contrary, Figure 7.10d shows how it lowers to 10^{-11} for $\mu = 0.5$: indeed, the stable velocity solution does not need any control effect being already asymmetric. These deductions comply with the right plot of Figure 7.14 w.r.t. several values of μ.

Also in this case, a non-natural optimal branch (this time featuring symmetric solutions) can be found with some numerical effort. Table 7.2 shows a comparison of the values of the cost functional (7.19) for the reached natural branch for both symmetric and asymmetric targets. Several values of μ (rows) and α (columns) have been taken into consideration and compered to the uncontrolled case. Once again, the functional is lower for smaller α. In the symmetric target case, the distributed control steers the solution towards the desired profile and this is visible for $\mu = 0.05$ and $\alpha = 0.01$, values for which the functional is decreased by a 90% w.r.t. its uncontrolled counterpart. If we lower α, say $\alpha = 0.001$, the cost functional is almost decreased by 99%. An analogous argument holds for the asymmetric target: in that case, the maximum action of the control variable is given for low Re. Indeed, for $\mu = 2.$ the functional decreases of the 77.5% for $\alpha = 0.01$. The percentage grows up to a 97% if we employ $\alpha = 0.001$. We here stress that no control action is needed for $\mu = 0.5 \approx 0.49$, i.e. it is the parameter value for which the asymmetric v_{d} was computed. This translates in very low values of (7.19), even below the machine precision. The spectral analysis of this optimality system is depicted in Figure 7.12: in particular, Figures 7.12a ($\alpha = 1$) and 7.12b ($\alpha = 0.1$) are related to the symmetric target when following the corresponding natural optimal branch, while Figures 7.12c ($\alpha = 1$) and 7.12d ($\alpha = 0.1$) consider the asymmetric target when following its natural optimal branch. It is clear that the behaviour between the top and bottom panels of Figure 7.12 is similar. Thus, we will only focus on the

role of α in this context. We remark that, as in the previous experiment, no stability information about the optimal solution is given by means of the global eigenvalue problem. Indeed, the remaining (i.e., non-natural optimal) branches present very similar patterns to the ones in Figure 7.12. Plotting the eigenvalues for $\alpha = 1$ in $\Re(\sigma_\mu) = [-0.01, 0.01]$ and for $\alpha = 0.1$ in $\Re(\sigma_\mu) = [-0.005, 0.005]$, we can observe a predominance of positive eigenvalues even for larger values of the penalization parameter. Also in this case, if α is smaller, less negative eigenvalues are visible.

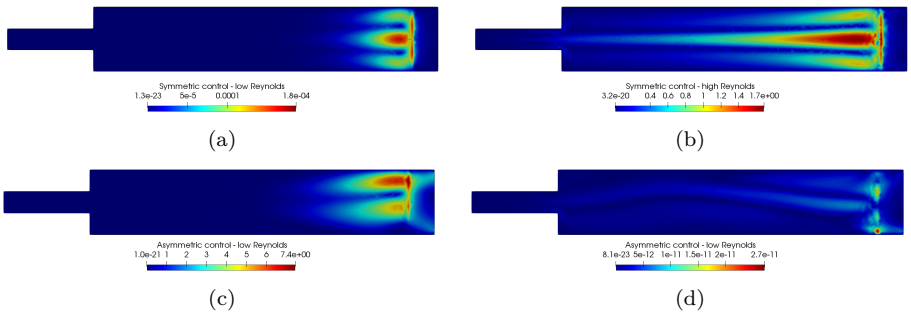


Figure 7.10: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Distributed control). Optimal control profiles for $\alpha = 0.01$. Left: $\mu = 2$ in (a) and (c); right: $\mu = 0.5$ in (b) and (d). *Top*. symmetric target in (a) and (b); *Bottom*. asymmetric target in (c) and (d).

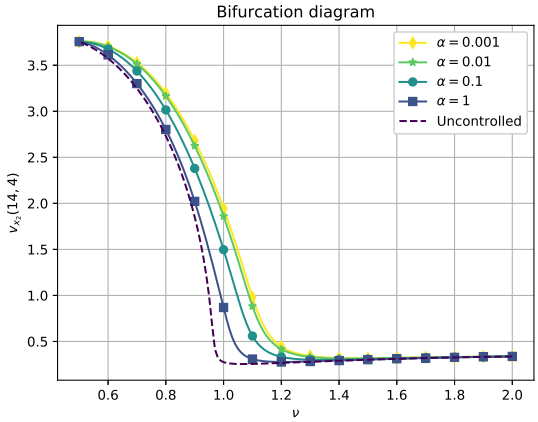


Figure 7.11: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Distributed control). Bifurcation diagram (upper branch only) for controlled state velocity obtained with $\alpha = 1, 0.1, 0.01, 0.001$ and asymmetric target, compared to the uncontrolled velocity.

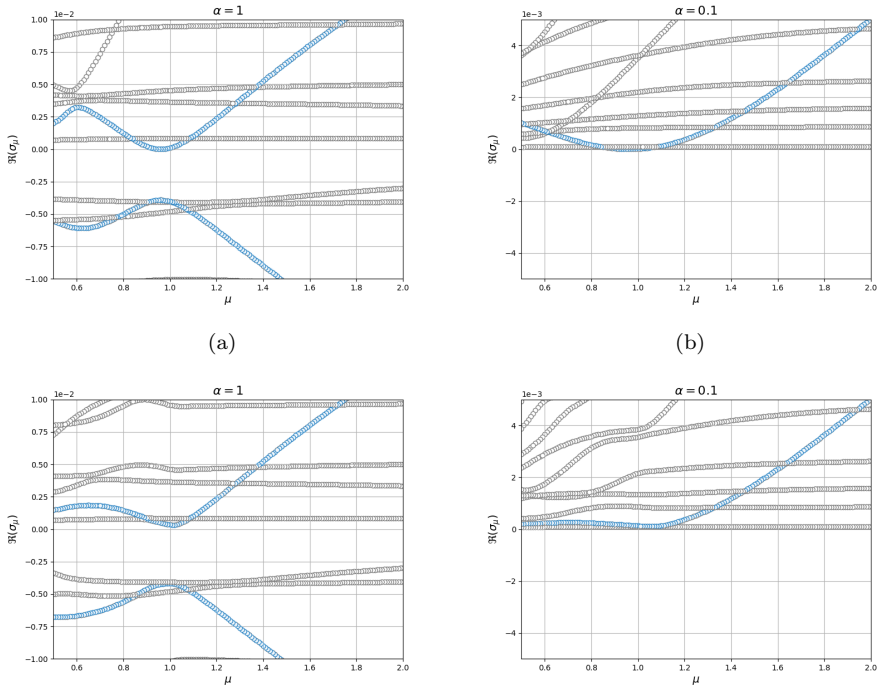


Figure 7.12: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Distributed control). Spectral analysis with $\alpha = 1, 0.1$ (left to right) for the natural optimal branch with symmetric (top) and asymmetric (bottom) targets.

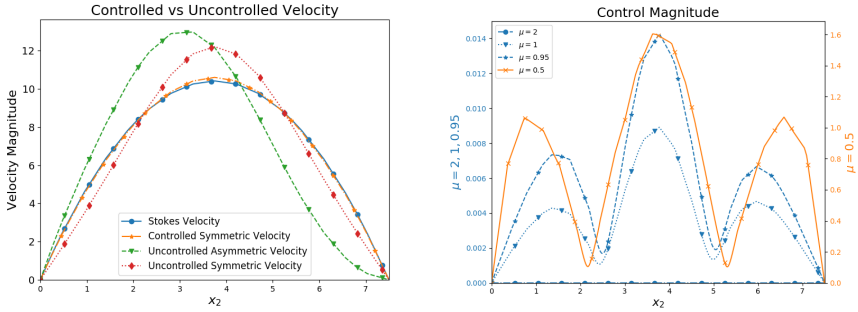


Figure 7.13: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Distributed control). *Left.* Comparison of velocity profiles in the controlled and uncontrolled cases for $\alpha = 0.01$, $\mu = 0.5$ on Γ_{obs} w.r.t. the symmetric desired profile when following the natural optimal branch. *Right.* Representation of control variable evolution for $\alpha = 0.01$, $\mu = 2, 1, 0.95, 0.5$ for $x_1 = 45$ when following the natural optimal branch.

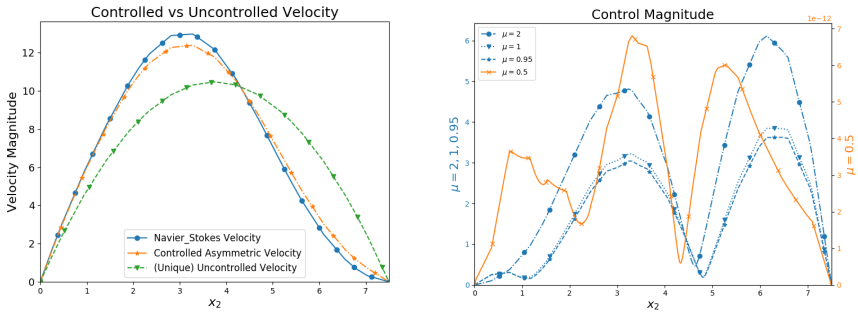


Figure 7.14: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Distributed control). *Left.* Comparison of velocity profiles in the controlled and uncontrolled cases for $\alpha = 0.01$, $\mu = 2$. on Γ_{obs} w.r.t. the asymmetric desired profile when following the natural optimal branch. *Right.* Representation of control variable evolution for $\alpha = 0.01$, $\mu = 2, 1, 0.95, 0.5$ for $x_1 = 45$ when following the natural optimal branch.

Table 7.2: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* Comparison of the functional value for distributed control. (Sym.) Natural optimal branch for symmetric target. (Asym.) Natural optimal branch for asymmetric target. (Sym.-U.) Unstable uncontrolled solution with symmetric target. (Asym.-S.) Stable uncontrolled solution with asymmetric target. (B.M.E.) Below machine epsilon.

μ	Sym.-U.	Asym.-S.	Sym.	Asym.	Sym.	Asym.	Sym.	Asym.	Sym.	Asym.
	Uncontrolled		$\alpha = 1$		$\alpha = 0.1$		$\alpha = 0.01$		$\alpha = 0.001$	
2	5.14e–9	1.88e+1	5.06e–9	1.81e+1	4.51e–9	1.36e+1	2.22e–9	4.23e+0	4.04e–10	5.66e–1
1.5	4.38e–6	1.88e+1	4.29e–6	1.77e+1	3.61e–6	1.20e+1	1.46e–6	3.09e+0	2.28e–7	3.87e–1
1	4.10e–3	1.86e+1	3.95e–3	1.67e+1	2.99e–3	9.15e+0	9.14e–4	1.86e+0	1.23e–4	2.17e–1
0.9	1.63e–2	1.84e+1	1.56e–2	1.54e+1	1.14e–2	7.88e+0	3.26e–3	1.50e+0	4.26e–4	1.73e–1
0.8	6.52e–2	1.54e+1	6.21e–2	1.31e+1	4.36e–2	6.06e+0	1.14e–2	1.08e+0	1.45e–3	1.22e–1
0.7	2.59e–1	1.15e+1	2.45e–1	9.28e+0	1.63e–1	3.68e+0	3.93e–2	6.16e–1	4.81e–3	6.90e–2
0.6	1.70e+0	5.34e+0	9.54e–1	3.76e+0	5.94e–1	1.24e+0	1.28e–1	2.00e–1	1.70e–2	2.22e–2
0.5	3.92e+0	B.M.E.	3.59e+0	B.M.E.	2.04e+0	B.M.E.	3.92e–1	B.M.E.	4.47e–2	B.M.E.

Moreover, we noticed that, no matter the used α, the shears phenomenon does not happen. Even if a small trace of the shears structure is still present for $\alpha = 1$ (highlighted in blue) in Figures 7.12a and 7.12c, for $\alpha = 0.1$ the structure completely breaks. Figure 7.12b and 7.12d only depict one eigenvalue (representing the top of the shears, and marked in blue) approaching $\Re(\sigma_\mu) = 0$ without crossing it. Let us denote with μ^{**} the point for which the upper shears curve is the closest to the axis $\Re(\sigma_\mu) = 0$. It will give us further information on the bifurcating phenomenon. Indeed, from Figure 7.12, $\mu^{**} \approx \mu^* = 0.96$ for the symmetric target, regardless of α, while, employing asymmetric target, we noticed that $\mu^{**} \in [1.0, 1.2]$ with a mild dependence on α. This feature is highlighted by Figure 7.11, representing the bifurcation diagram for the controlled solution with asymmetric target. Namely, we can state that μ^{**} indicates where

the bifurcation of the controlled system occurs. We conclude that this optimality system steers the state solution towards a desired branch affecting the location of the bifurcation point, eventually. The role of α in the optimization process will be clarified in the next Section.

7.3.4 Channel Control: the α effect

This Section aims at describing how the value of α can affect the natural convergence towards a symmetric target over Γ_{obs}. In this setting, the problem reads: given $\mu \in \mathcal{D}$ find the optimal solution $\boldsymbol{X} \in \mathbb{X}$ that verifies:

$$\begin{cases} v\mathbb{I}_{\Gamma_{\text{obs}}} - \mu\Delta w - v\cdot\nabla w + (\nabla v)^T w + \nabla q = v_{\text{d}}\mathbb{I}_{\Gamma_{\text{obs}}} & \text{in } \Omega, \\ \text{div}(w) = 0 & \text{in } \Omega, \\ w = 0 & \text{on } \Gamma_{\text{in}} \cup \Gamma_{\text{wall}}, \\ -qn + (\mu\nabla w)n = 0 & \text{on } \Gamma_{\text{out}}, \\ \alpha u\mathbb{I}_{\Gamma_{\text{ch}}} = w\mathbb{I}_{\Gamma_{\text{ch}}} & \text{in } \Omega, \\ -\mu\Delta v + v\cdot\nabla v + \nabla p = u\mathbb{I}_{\Gamma_{\text{ch}}} & \text{in } \Omega, \\ \text{div}(v) = 0 & \text{in } \Omega, \\ v = v_{\text{in}} & \text{on } \Gamma_{\text{in}}, \\ v = 0 & \text{on } \Gamma_{\text{wall}}, \\ -pn + (\mu\nabla v)n = 0 & \text{on } \Gamma_{\text{out}}, \end{cases} \tag{7.33}$$

where the control variable is defined at the end of inlet channel, i.e. $\Omega_u = \Gamma_{\text{ch}}$, as depicted in Figure 7.1, and Γ_{wall} is, once again, $\Gamma_0 \cup \Gamma_{\text{D}}$. Namely, we are interpreting the control as a forcing term influencing how the flow enters in the expansion channel. Figure 7.15 shows the adjoint velocity and pressure profiles obtained for $\mu = 0.5$ for $\alpha = 1$ and $\alpha = 0.01$. For the highest value of the penalization parameter, following Algorithm 1, the natural optimal branch features a wall-hugging behaviour, while for smaller values of α the control variable drives the velocity towards a symmetric flux (see the left panels of Figures 7.16 and 7.17).

Therefore, in this case, the natural optimal branch is highly influenced by the penalization parameter. If we focus on the right plots of Figures 7.16 and 7.17, we notice that the control is very sensitive around μ^*. This is confirmed by its asymmetric configuration both for the wall-hugging solution and the straight one. We remark that we were capable to detect two solutions using different initial guesses in the continuation method for $\alpha = 1, 0.1, 0.01$, showing symmetric and asymmetric features coexisting for values of $\mu < \mu^*$. However, the smaller was α, the bigger was the effort needed to numerically solve this non-natural branch detection task. In the case of $\alpha = 0.001$ the control action was so strong that we were not able to actually recover the whole optimal non-natural branch.

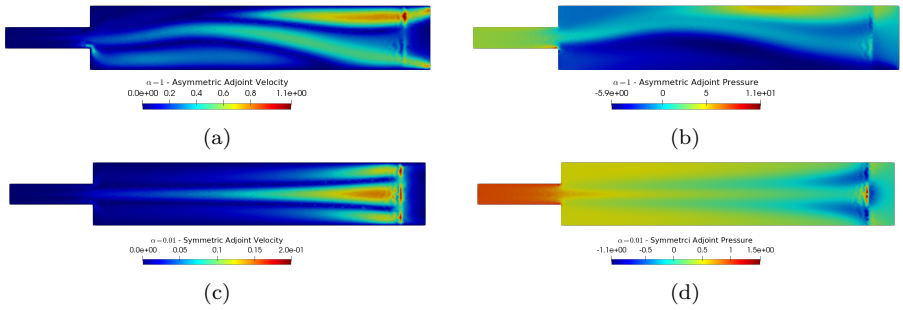


Figure 7.15: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Channel control). Two optimal solutions for adjoint velocity and pressure for $\mu = 0.5$: $\alpha = 1$ in (a) and (b), and $\alpha = 0.01$ in (c) and (d), respectively.

Table 7.3: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* Comparison of the functional value for channel control w.r.t. stable and unstable uncontrolled solutions. *Headers:* (Nat.) Natural optimal branch. (n-Nat.) Non-natural optimal branch. *Trailing cell characters*: (s) the solution has symmetric profile. (a) The solution has asymmetric profile. (nat-C.) Converging to natural branch despite tailored guess. (non-C.) Non-converging Newton's solver for tailored guess.

μ	Stable	Unstable	Nat.	n-Nat.	Nat.	n-Nat.	Nat.	n-Nat.	Nat.	n-Nat.
	Uncontrolled		$\alpha = 1$		$\alpha = 0.1$		$\alpha = 0.01$		$\alpha = 0.001$	
2	5.14e–9	5.14e–9	5.14e–9s	5.14e–9s	5.14e–9s	5.14e–9s	5.14e–9s	5.14e–9s	5.07e–9s	5.14e–9s
1.5	4.38e–6	4.38e–6	4.38e–6s	4.38e–6s	4.38e–6s	4.38e–6s	4.38e–6s	4.38e–6s	4.28e–6s	4.38e–6s
1	4.10e–3	4.10e–3	4.10e–3s	4.10e–3s	4.10e–3s	4.10e–3s	4.08e–3s	4.10e–3s	3.92e–3s	4.10e–3s
0.9	3.33e–2	1.63e–2	3.33e–2a	1.63e–2s	1.63e–1s	3.33e–2a	1.63e–1s	nat-C.	2.93e–2s	non-C.
0.8	2.08e–1	6.52e–2	2.08e–1a	6.52e–2s	6.52e–2s	2.07e–1a	6.52e–2s	2.04e–1a	6.51e–2s	nat-C.
0.7	1.01e+0	2.59e–1	1.01e+0a	2.59e–1s	2.59e–1s	1.01e+0a	2.59e–1s	9.76e–1a	2.24e–1s	nat-C.
0.6	4.48e+0	1.70e+0	4.48e+0a	1.02e+0s	1.02e+0s	4.43e+0a	1.02e+0s	4.03e+0a	9.90e–1s	nat-C.
0.5	1.88e+1	3.92e+0	1.87e+1a	3.92e+0s	3.92e+1s	non-C.	3.87e+0s	non-C.	3.50e+0s	nat-C.

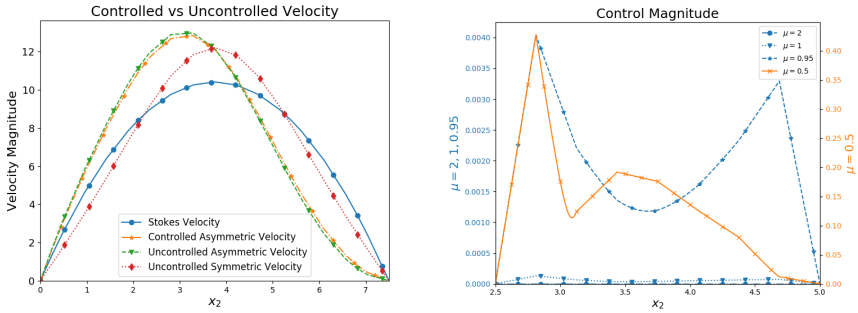


Figure 7.16: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Channel control). *Left.* Comparison of velocity profiles in the controlled and uncontrolled cases for $\alpha = 1$, $\mu = 0.5$ on Γ_{obs} w.r.t. the symmetric desired profile when following the natural optimal branch. *Right.* Representation of control variable evolution for $\alpha = 1$, $\mu = 2, 1, 0.95, 0.5$ at $x_1 = 10$ when following the natural optimal branch.

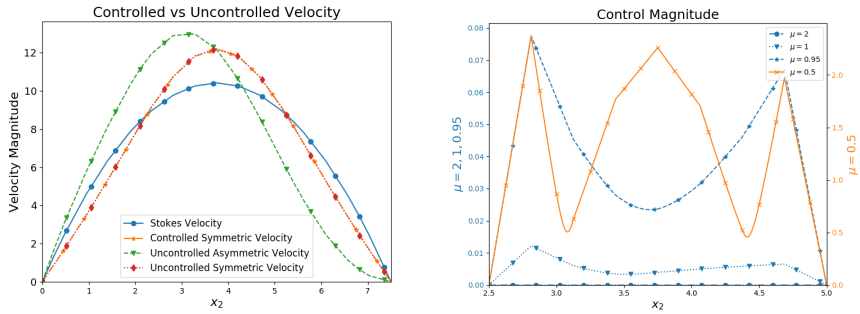


Figure 7.17: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Channel control). *Left.* Comparison of velocity profiles in the controlled and uncontrolled cases for $\alpha = 0.01$, $\mu = 0.5$ on Γ_{obs} w.r.t. the symmetric desired profile when following the natural optimal branch. *Right.* Representation of control variable evolution for $\alpha = 0.01$, $\mu = 2, 1, 0.95, 0.5$ for $x_1 = 10$ when following the natural optimal branch.

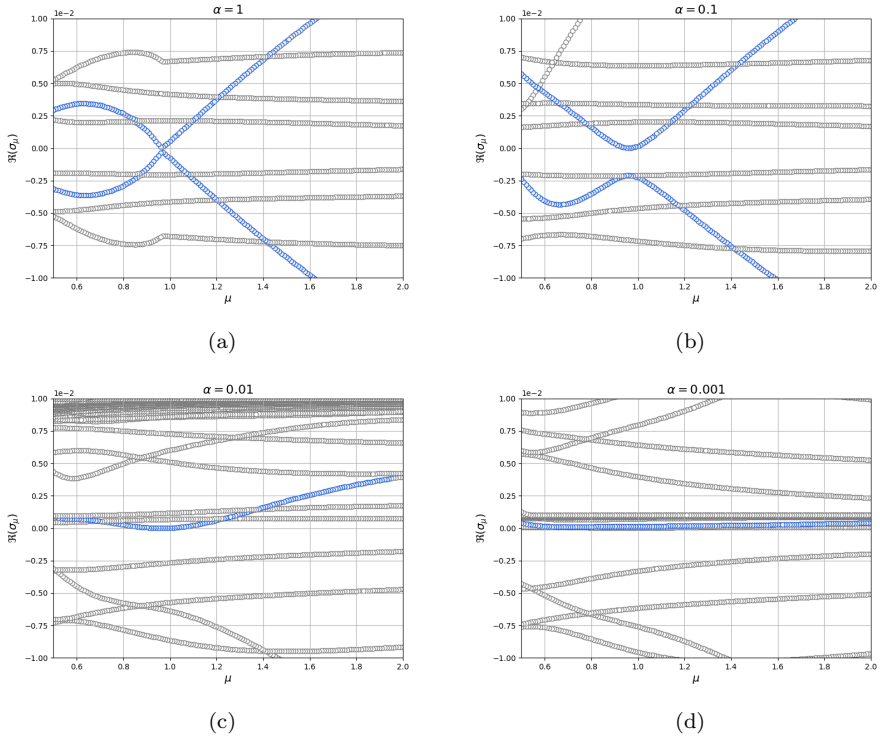


Figure 7.18: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Channel control). Spectral analysis with $\alpha = 1, 0.1, 0.01, 0.001$ when following the natural optimal branch.

Indeed, either the Newton's solver did not converge (this happens also for $\alpha = 0.1, 0.01$ and $\mu = 0.5$, compare Table 7.3) or converged to the symmetric natural branch. Let us comment on Table 7.3 and on the role of the penalization parameter. As the other test cases of Sections 7.3.2 and 7.3.3, the straight configuration gives better results in minimizing the functional w.r.t. the asymmetric solution thanks to its similarity with the symmetric target we fixed. For this test case, the decrease of the functional is not remarkable w.r.t. the uncontrolled system. Namely, even if the optimal solution might present straight natural convergence to the symmetric v_{d}, the parabolic profile on Γ_{obs} is not achieved (for example, the reader can compare the functional value for $\mu = 0.5$ and $\alpha = 0.001$ w.r.t. the uncontrolled symmetric solution: it only decreases of a 10%). Finally, we present the global eigenvalue analysis in Figure 7.18 in the range $\Re(\sigma_\mu) = [-0.01, 0.01]$ when following the natural optimal branch. For $\alpha = 1$, the shears phenomenon happens, while the structure is broken for other values of the penalization parameter. As expected, lower values of α lead to a positive-dominated eigenvalues ensemble. Moreover, a clustering around the value of α can be observed in plots 7.18c and 7.18d. We remark that the same properties are preserved in the non-natural branch analysis.
The next Section will deal, once again, with the role of the penalization parameter, presenting, however, very peculiar results.

7.3.5 Dirichlet Control: flux action

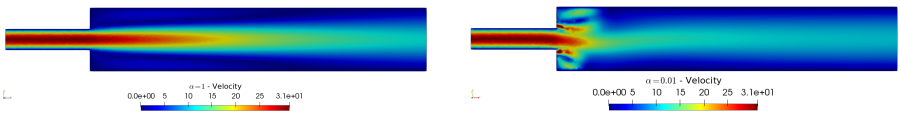
In this final example, we build a Dirichlet control over the boundary $\Omega_u \equiv \Gamma_{\mathrm{D}}$. Also for this test case, we consider the symmetric v_{d} over the line Γ_{obs}. Here, we set $\Gamma_{\mathrm{wall}} = \Gamma_0$. The problem to be solved reads: given $\mu \in \mathcal{D}$ find $\boldsymbol{X} \in \mathbb{X}$ such that

$$
\begin{cases}
v\mathbb{I}_{\Gamma_{\mathrm{obs}}} - \mu\Delta w - v\cdot\nabla w \\
\qquad\qquad +(\nabla v)^T w \\
\qquad\qquad\quad +\nabla q = v_{\mathrm{d}}\mathbb{I}_{\Gamma_{\mathrm{obs}}} & \text{in } \Omega, \\
\mathrm{div}(w) = 0 & \text{in } \Omega, \\
w = 0 & \text{on } \Gamma_{\mathrm{in}} \cup \Gamma_{\mathrm{D}} \cup \Gamma_{\mathrm{wall}}, \\
-qn + (\mu\nabla w)n = 0 & \text{on } \Gamma_{\mathrm{out}}, \\
\alpha u\mathbb{I}_{\Gamma_{\mathrm{D}}} = w\mathbb{I}_{\Gamma_{\mathrm{D}}} & \text{in } \Omega, \\
-\mu\Delta v + v\cdot\nabla v + \nabla p = 0 & \text{in } \Omega, \\
\mathrm{div}(v) = 0 & \text{in } \Omega, \\
v = v_{\mathrm{in}} & \text{on } \Gamma_{\mathrm{in}}, \\
v = u & \text{on } \Gamma_{\mathrm{D}}, \\
v = 0 & \text{on } \Gamma_{\mathrm{wall}}, \\
-pn + (\mu\nabla v)n = 0 & \text{on } \Gamma_{\mathrm{out}}.
\end{cases}
\tag{7.34}
$$

Namely, in this formulation, we are letting the flux to freely enter or exit from the boundary Γ_{D}. This will highly affect the solution behaviour. First of all, we stress that the control aims at straighten the solution, due to the choice of v_{d}. This feature is well represented by Figure 7.19a and the left plot of Figure 7.20. The symmetric profile is reached even for large values of α and, for lower values of α, the optimal velocity on Γ_{obs} is parabolic. This property is also highlighted from the functional values in Table 7.4, where, once again, we compare the uncontrolled behaviour of the functional (7.19) w.r.t. several optimal solutions changing μ (rows) and α (columns). As expected, the functional largely decreases for smaller values of the penalization parameter. For example, for $\mu = 0.5$, the functional only lowers of 18% for $\alpha = 0.01$, while we reach an 82% of decrease percentage for $\alpha = 0.001$. Furthermore, for $\alpha = 0.001$, the system has an interesting and unexpected profile, shown in Figure 7.19b. The flux presents a new bifurcating asymmetric configuration for a low value of μ. This asymmetry is related to the high influence of the control that not only allows the flow to exit from Γ_{D} (to avoid the asymmetric recirculation of the wall-hugging solution), but it also adds flux near the channel, in order to achieve the parabolic velocity profile over the observation domain, as it is represented in the right plot of Figure 7.20.

Table 7.4: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* Comparison of the functional value for Dirichlet control w.r.t. the stable and unstable uncontrolled solutions.

μ	Stable	Unstable	Controlled Solution			
	Uncontrolled		$\alpha = 1$	$\alpha = 0.1$	$\alpha = 0.01$	$\alpha = 0.001$
2	5.14e–9	5.14e–9	4.98e–9	4.83e–9	4.79e–9	4.79e–9
1.5	4.38e–6	4.38e–6	4.24e–6	4.10e–6	4.07e–6	4.06e–6
1	4.10e–3	4.10e–3	3.94e–3	3.78e–3	3.74e–3	3.72e–3
0.9	3.33e–2	1.63e–2	1.56e–2	1.49e–2	1.47e–2	1.45e–2
0.8	2.08e–1	6.52e–2	6.20e–2	5.88e–2	5.78e–2	5.46e–2
0.7	1.01e+0	2.69e–1	2.44e–1	2.29e–1	2.21e–1	1.82e–1
0.6	4.48e+0	1.70e+0	9.49e–1	8.73e–1	8.09e–1	3.57e–1
0.5	1.88e+1	3.92e+0	3.58e+0	3.21e+0	2.41e+0	4.73e–1



(a) (b)

Figure 7.19: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Dirichlet control). Two optimal velocity solutions for $\mu = 0.5$, with $\alpha = 1$ and $\alpha = 0.001$, left and right, respectively.

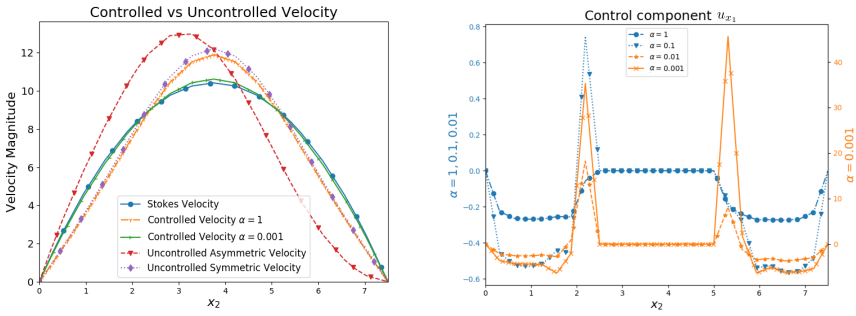


Figure 7.20: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* *Left.* Comparison of velocity profiles in the controlled and uncontrolled cases for $\alpha = 1, 0.01$, $\mu = 0.5$ on Γ_{obs} w.r.t. the symmetric desired profile when following the natural optimal branch. *Right.* Representation of control variable evolution for $\alpha = 1, 0.1, 0.01, 0.001$ and $\mu = 0.5$ at $x_1 = 10$ when following the natural optimal branch.

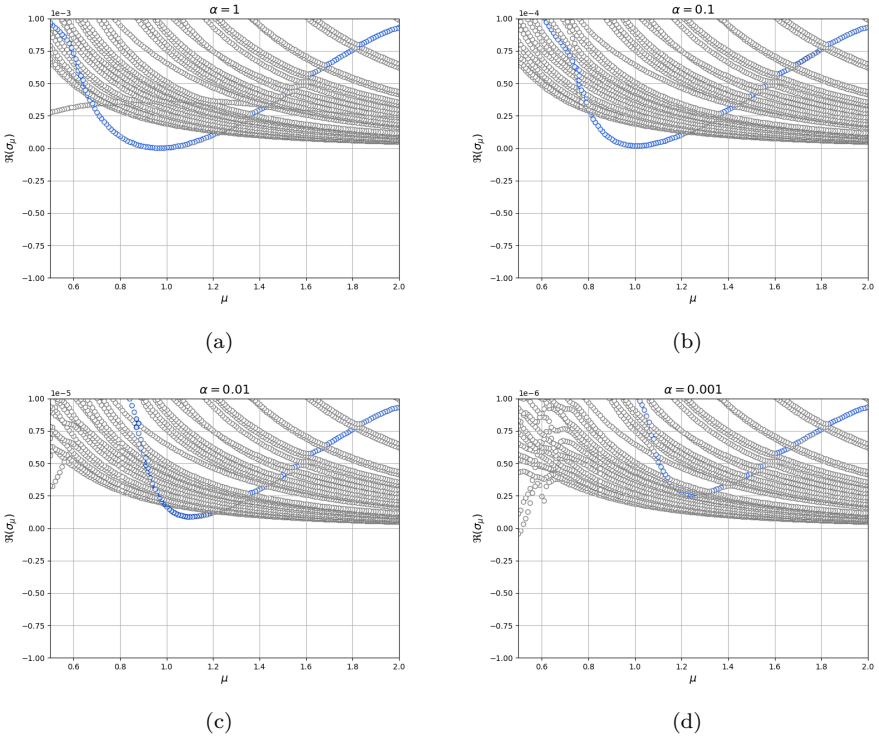


Figure 7.21: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Dirichlet control). Spectral analysis with $\alpha = 1, 0.1, 0.01, 0.001$.

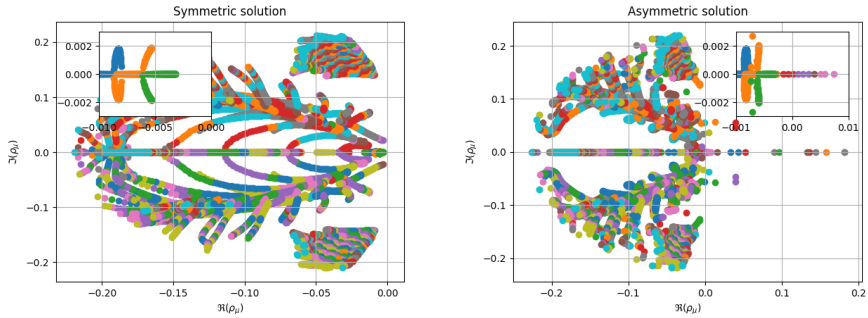


Figure 7.22: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Dirichlet control). Eigenvalues of the state eigenproblem in the complex plane for the Dirichlet optimal control: symmetric and asymmetric solutions, left and right, respectively.

In Figure 7.21 the results on eigenvalues analysis are presented. We show close-ups starting with $\Re(\sigma_\mu) = [-0.001, 0.001]$ for $\alpha = 1$ in the top-right image. Then, the vertical interval is restricted following the order of the lower values of α in the remaining panels. The arguments we discussed for the other test cases apply also in this context. For example, a stronger control action leads to larger number of positive eigenvalues. As we did in the distributed case of Section 7.3.3, we define the value μ^{**} as the parameter for which the top curve of the shears (marked in blue) approaches $\Re(\sigma_\mu) = 0$. As α becomes lower and lower, the curve is moved away from $\Re(\sigma_\mu) = 0$. Thus, in strong controlled problems, such a point μ^{**} might not exist. The previous experimental settings have shown that the top shear structure approaching to $\Re(\sigma_\mu) = 0$ are typically related to a bifurcation phenomenon. Thus, we guess that the standard pitchfork bifurcating configuration is not occurring here. Yet, other type of bifurcating phenomena might be arose. Indeed, the system seems to be featuring a *different bifurcation*, presented in Figure 7.19b. The plot shows an eigenvalue crossing the line $\Re(\sigma_\mu) = 0$ for the global eigenproblem for $\alpha = 0.001$. This led us to analyse the the state eigenproblem of Algorithm 1 (Figure 7.22) around the optimal solution. In this new controlled framework, the symmetric profile does never cross the origin, while it happens when one is dealing with the asymmetric solution for $\alpha = 0.001$ represented in Figure 7.19b. Namely, properly modifying the boundary conditions, the controlled straight profile can be interpret as a physical stable solution. Lastly, we remark that in the right plot of Figure 7.22 a couple of complex and conjugate eigenvalue are crossing the imaginary axis, as in the Hopf bifurcation scenario [124, 140]. We can thus affirm that the Dirichlet control problem deeply changes the classical behaviour of the solution.

Remark 7.3.1 (Lagrange multipliers) *Numerically, the optimality system* (7.34) *has been solved through the employment of Lagrangian multipliers. Namely,*

the condition $v = u$ on Γ_D has been weakly imposed in integral form

$$\int_{\Gamma_D} v\lambda \, ds = \int_{\Gamma_D} u\lambda \, ds \quad \forall \lambda \in V. \tag{7.35}$$

This equation reflects in system (7.21), since the term

$$\int_{\Gamma_D} \phi(\nabla \cdot vn + pn) \, ds \quad \forall \phi \in V, \tag{7.36}$$

appears. This numerical choice leads to new terms in the adjoint equations. Moreover, we weakly imposed the adjoint boundary condition $w = 0$ on Γ_D with another multiplier. The reason of this latter decision will be explained in Section 7.4.

7.3.6 Comparative Eigenvalue Analysis

This Section focuses on all the common observations and the results obtained by means of the global eigenvalue analysis over the four presented OCP($\boldsymbol{\mu}$)s. A list of the similarities among them follows:

- the *eigenvalues cluster* around the value of α. This behaviour arises from the optimality equation 7.24. This feature is well represented in Figures 7.9c, 7.9d, 7.18c and 7.18d;

- the *predominance of positive eigenvalues* over the negative ones. In all the test cases the control action tends to lower the negative eigenvalues. The stronger it is, the more remarkable is this process, as represented in Figures 7.12b and 7.21;

- the *shears effect* for low controlled systems that do not highly affect the standard uncontrolled system solution. It is the case of Neumann control in Figure 7.9a and of the channel control for $\alpha = 1$ as shown in Figure 7.18a. For the other cases, this feature might be visible for higher values of the penalization parameter. However, if the control is strong enough, the structure is completely broken;

- the μ^{**} *identification*. We observed that the shears (or their top curve, if broken) approach the real line in μ^{**}, i.e. where the bifurcating phenomenon of the controlled system is happening. This argument often holds regardless of α, see for example Figures 7.12b, 7.18b, 7.18c. The detection of μ^{**} can be even feasible in strongly controlled system, such as the Dirichlet optimal control, see Figure 7.21. In some cases, we observed a shift of the μ^{**} w.r.t. the uncontrolled critical point μ^*.

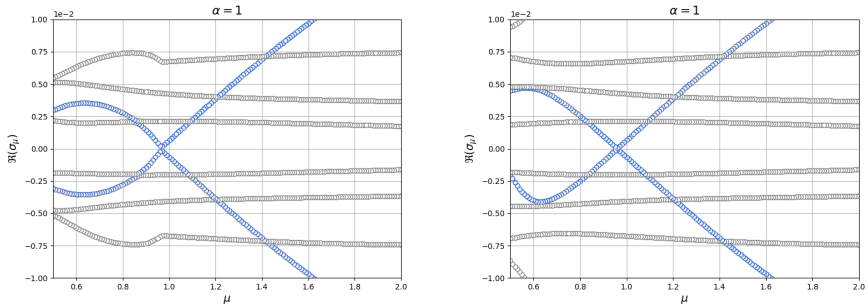


Figure 7.23: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* *Left.* Asymmetric velocity with Neumann control for $\alpha = 1$. *Right.* Symmetric velocity with distributed control for $\alpha = 1$.

To gain more information on the problem, we perform an eigenvalue analysis dealing with only the state and the adjoint equations around the optimal solutions. For all the test cases, shears occurs. However, we noticed that their structure is symmetric when the solution shows the wall-hugging property, while it is slightly asymmetric when the state flow is straight. We believe that this behaviour derives from the different reaction of state and adjoint variables to bifurcations. Indeed, when reaching the desired straight flux, the behaviour of the state variable should be preserved for all μ. This happens thanks to the adjoint problem, related to the control variable: it re-balances the flux, resulting in an asymmetric contribution reflected in the structure of the shears. Now that we have deeply analysed the theme of bifurcations in nonlinear OCP($\boldsymbol{\mu}$)s, in the next Section will present some numerical results dealing with the reduction of this kind of problems.

7.4 ROMs for bifurcating OCP($\boldsymbol{\mu}$)s

In this Section we are going to present how ROMs strategies as presented in Chapter 3 behaves w.r.t. the setting of nonlinear bifurcating OCP($\boldsymbol{\mu}$)s. We here combined the reduction techniques for OCP($\boldsymbol{\mu}$)s discussed in Part II together with reduction strategies for bifurcating system as presented in [62, 63, 116, 117, 119, 121]. We employed a *branch-wise approach*, i.e., for every bifurcating solution branch $\mathbb{M}_i$, we build a different reduced model. This strategy is remarkably effective to reach accurate ROM solutions. However, in [117, 116], for example, a global reduction approach was pursued for other applications with remarkable results. Before showing the numerical results we would like to list the main features of the used approach:

- for each branch we performed a partitioned POD as described in Section 4.1;

- the aggregated space techniques was applied in order to guarantee the well-posedness of the linearized reduced system for $\mu \neq \mu^*$ as discussed in Section 3.2.2;
- the supremizers are needed in order to deal with the nested saddle point structure of the problem.

Namely, we are working with an enlarged reduced space: this is the key point in order to have a numerically stable system at the reduced level when we are not analysing the case of the critical parameter μ^*. As already specified in Remark 3.1.1, the nonlinear problem still depends on the high fidelity solution. In this setting, no hyper-reduction technique is applied following [118].

7.4.1 Numerical Results

We now present the numerical results deriving from the reduction of the four controlled test cases described in Section 7.3.1. For each numerical experiment, we extract information from $N_{max} = 51$ snapshots evaluated for equidistant parameters in $\mathcal{D} = [0.5, 2]$. For all the test cases, we chose to retain $N = 20$ bases, except for the Dirichlet test case were 12 bases have been employed for each variable. We remark that for this specific test case we need additional basis functions for the two Lagrangian multiplier variables defined in Remark 7.3.1: the final reduced space is, thus, of dimension $15N$ and this justifies the use of a slightly smaller reduced framework for the Dirichlet problem. Then, we performed and error analysis over 151 equidistant values of $\mu \in \mathcal{D}$. The reliability of the ROM approach has been evaluated through

- an average error over the parameter space w.r.t. an increasing value of basis functions;
- a μ-dependent error computed for the fixed value $N = 12$ for the Dirichlet problem and $N = 20$ for all the other cases.

From the two error analyses we can deduce different features of the problem at hand. Indeed, the average error tests how the reliability changes w.r.t. the behaviour of the solution. The symmetric profile appears to always be the best approximated. See for example the Neumann and the Channel control, which average errors are depicted in the left plots of Figures 7.24 and 7.28, respectively. Their asymmetric counterparts, i.e. the right panels of Figures 7.24 and 7.28, show how recovering the Stokes-like profile for $\mu > \mu^*$ and the wall-hugging feature for the lower values of μ is a more difficult task w.r.t. the symmetric case. Nonetheless, the provided accuracy is satisfactory for practical applications for both the cases. This holds true for all the involved variables, however, the state ones are the best described by the reduced system. As we already said, the control is very sensitive to the parameter and, thus, it is the

most challenging to approximate. Indeed, the control variable features a sort of *on-off* behaviour and this affects the quality of reduced representation. For example, when dealing with Stokes target v_{d}, the control is *off* for larger values of μ, but, when $\mu \sim \mu^*$, it starts to drastically grow in magnitude and it also presents qualitatively differences w.r.t. the previous values of μ. This property is directly inherited by the adjoint variable, by definition. This is visible from the right panels of Figures 7.25 and 7.31, where higher values of error correspond to higher values of μ. Due to the same reason, for Channel and Dirichlet test cases with low Re, we chose to plot the absolute errors for the control variable since its magnitude was essentially zero and in this way we prevented a meaningless relative error. This issue does not occur in the Distributed control case, see for example the right plot of Figure 7.26: a remarkable errors decay for all the variables can be noticed.

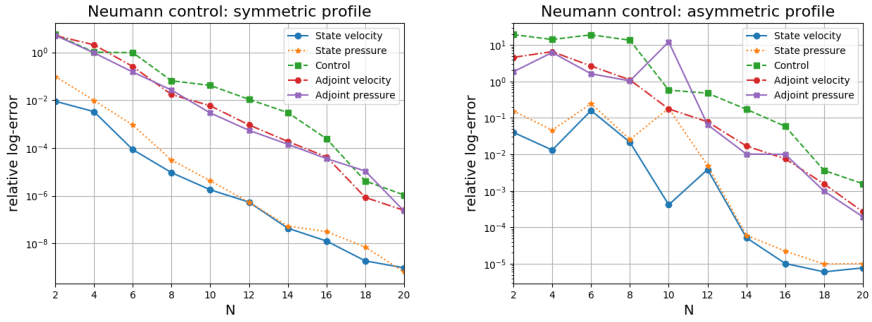


Figure 7.24: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena* (Neumann control). Average error over μ with $N = 20$ and $\alpha = 0.01$. *Left*. Symmetric profile. *Right*. Symmetric profile.

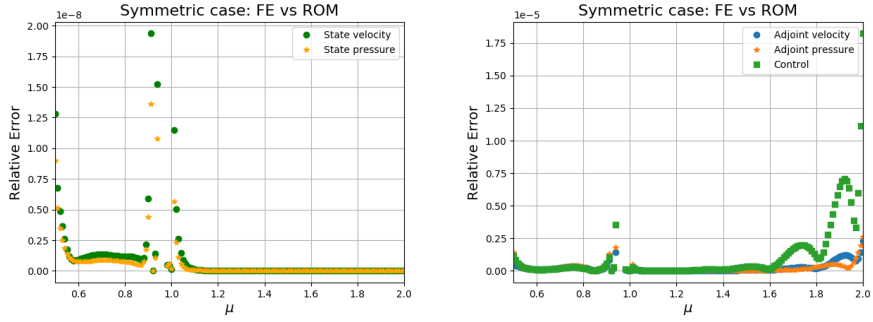


Figure 7.25: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Neumann control). The μ-dependent error with $N = 20$ and $\alpha = 0.01$. *Left*. Symmetric profile for state variable. *Right*. Symmetric profile for adjoint variable..

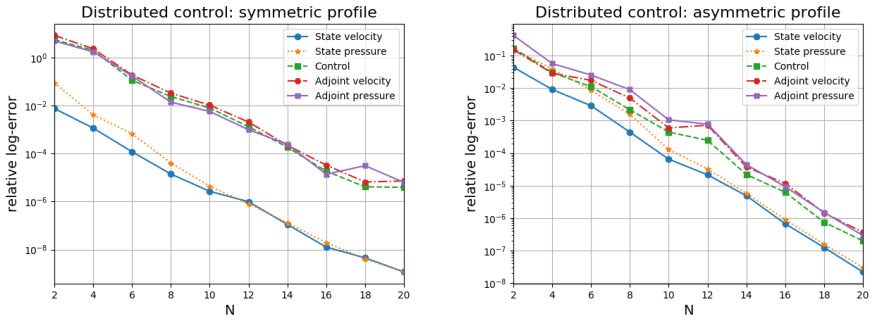


Figure 7.26: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Distributed control). Average error over μ with $N = 20$ and $\alpha = 0.01$. *Left.* Symmetric profile. *Right.* Asymmetric profile.

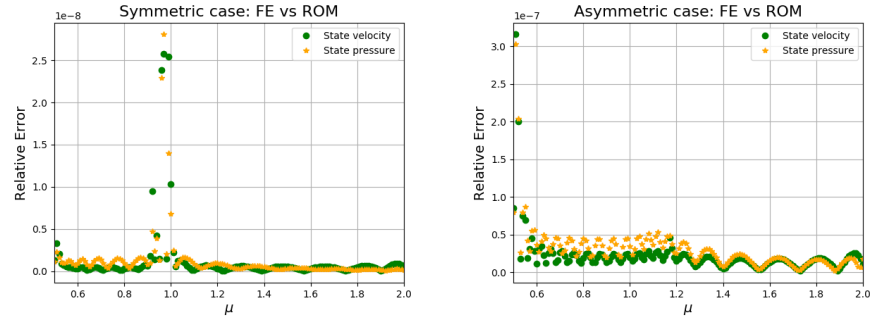


Figure 7.27: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Distributed control). The μ-dependent error with $N = 20$ and $\alpha = 0.01$. *Left.* Symmetric profile. *Right.* Asymmetric profile.

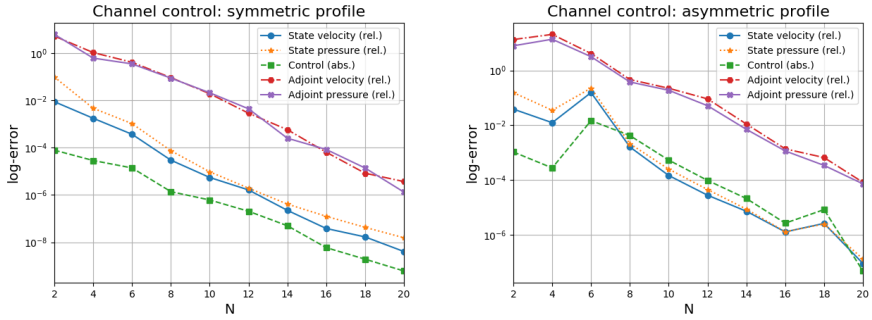


Figure 7.28: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena* (Channel control) Average error with $N = 20$ over μ. *Left.* Symmetric profile ($\alpha = 1$). *Right.* Asymmetric profile ($\alpha = 0.01$).

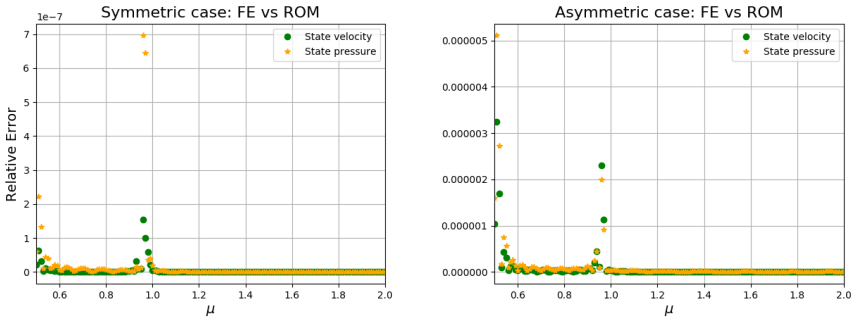


Figure 7.29: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Channel control). The μ-dependent error for $N = 20$ *Left*. Symmetric state profile. *Right*. Asymmetric state profile.

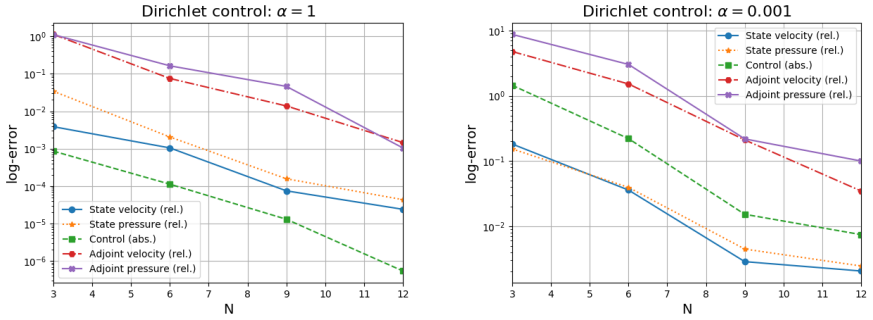


Figure 7.30: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Dirichlet control). Average error over μ with $N = 12$. *Left*. Profile for $\alpha = 1$. *Right*. Profile for $\alpha = 0.001$.

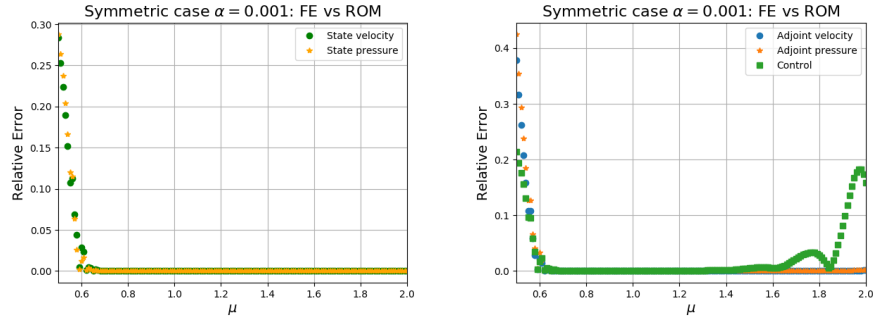


Figure 7.31: *OCP($\boldsymbol{\mu}$)s for bifurcating phenomena.* (Dirichlet control). The μ-dependent error for $\alpha = 0.001$ with $N = 12$. *Left*. State variable. *Right*. Adjoint and control variables.

The Dirichlet test case for $\alpha = 0.001$ is very challenging to represent: the reduced model suffers the complex dynamic and the new bifurcation behaviour. Indeed, while the other test cases had an average error are ranging between 10^{-5} and 10^{-8}, here we reach an error of almost 10^{-3} for the controlled state. Yet, this accuracy is still acceptable for many practical applications. This feature is highlighted also in the μ-dependent errors depicted in Figure 7.31. In both the pictures, we observe an increment of the error for high Re. This happens also for the other test cases, see, for example, the right plot of Figure 7.27 and both the plots of Figure 7.29. However, the Dirichlet case is the worst behaving under this point of view. Nonetheless, the μ-dependent error gives an *a posteriori* information about the bifurcation point. Indeed, it is well-known that in order to achieve a good ROM representation, regularity on the parametric dependence is needed [64]. For this reason, we can notice an error peak around μ^*. It happens, for example, for $\mu^* \sim 0.96$ for Neumann, Distributed and Channel control. It is clear that this property can be very useful when no previous knowledge about bifurcations of the system is provided. Thus, ROMs are not only useful to faster solve a very complicated time consuming system, but also to detect problematic parameters related to the bifurcating nature of the problem at hand, since these will be the worst approximated. The same feature arises in the Dirichlet control case at the left end of the parametric domain $\mathcal{D}$: this is due to the new configuration observed in Figure 7.19b for high Reynolds.

This Chapter aimed at showing how OCP($\boldsymbol{\mu}$)s can be useful to prevent some configurations and dynamics in nonlinear analysis. We here conclude the application of standard model order reduction techniques for OCP($\boldsymbol{\mu}$)s. Indeed, a final Chapter concerning the use of non-intrusive techniques based on artificial intelligence for OCP($\boldsymbol{\mu}$)s follows.

Chapter 8

Physics Informed Neural Networks for Optimal Control Problems

In this Chapter we extend the concept of physics informed supervised machine learning strategies to OCP($\boldsymbol{\mu}$)s in real-time and many-query applications. Indeed, following [46], we will provide a physics informed learning paradigm to reach accurate optimal simulations in a small amount of time. In Section 8.1, we will give some motivations about the employment of these approaches together with the standard formulation of physics informed neural networks (PINNs). While, in Section 8.2 we exploit the physical model to build tailored neural networks (NNs) to accelerate the training phase of the process. Then, these techniques will be combined and tested in an optimal control framework.

8.1 Main Ideas Behind PINNs

Machine Learning (ML) represents a research topic of growing impact. In the latest years, it has been successfully exploited in several fields of applications, see for example [47, 85, 91, 161]. This massive interest in ML is directly inherited by the availability of a huge amount of data, nowadays.

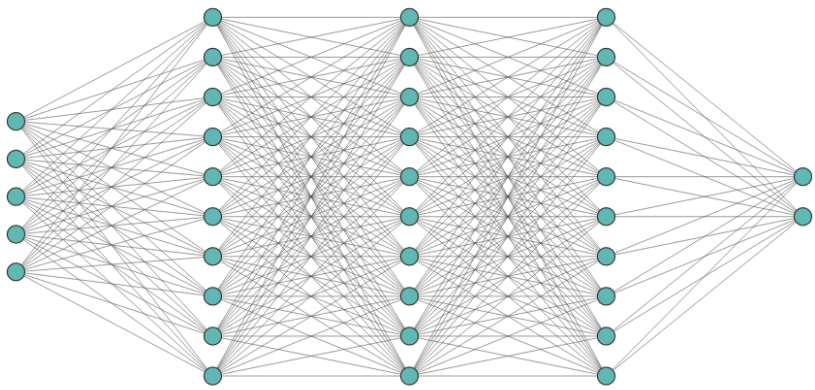


Figure 8.1: *PINNs for OCP($\boldsymbol{\mu}$)s.* Example of a feedforward NN, with a 5-neurons input layer, three 10-neurons hidden layers, 2-neurons output layer.

The classical structure used in this context is the NN, i.e. a computational

architecture able to learn a configuration from some observed data. A NN consists of a set of *neurons* connected by a weighted synaptic structure. The data information travels among this architecture and, after a training phase, from an input, the NN provides an output. In this contribution we will only consider *feedforward NNs*, where the neurons are aligned into one input layer, some hidden layers and an output layer. Each layer is connected to the next layer and the information travels to the output layer, as represented in the example depicted in Figure 8.1. Namely, a training procedure over the weights of the connections is performed through some rules and parameters (such as learning rate and activation functions) which help in reaching a low prediction error represented by the *loss function*. This only was a brief overview of NNs and their structure: the interested reader can refer, e.g., to [25, 55] for more details. Even if NNs are indisputably useful in many contexts, they may suffer of some lack of accuracy w.r.t. some expected behaviour in fields where classical PDE($\boldsymbol{\mu}$)s-based models have been applied. This issue can be tackled though physics informed strategies. After introduced some motivations in Section 8.1.1, we will discuss how to enforce physical features to the output of the NN in Section 8.1.2.

8.1.1 Motivations

Techniques based on ML are undeniably an asset to the improvement of numerical simulations. However, their accuracy, most of all for complicated systems, is strictly related to the amount of data to train at one's disposal. This, may represent an issue since data can be expensive to collect and they are usually characterized by a scattered and incomplete nature. These features might lead to *non–physical* and inconsistent results w.r.t. some previous knowledge related to mathematical modelling. Lately, to solve this kind of problem, a new technique has been conceived: the Physics Informed Neural Networks (PINNs). The idea of PINNs is as simple as effective: if any, add some prior physical knowledge to the Neural Network (NN) and this will eventually result in more accurate and robust predictions. Historically, the mathematical model related to PINNs were (non-parametric) PDEs. The reader may refer to the seminal paper [129] for an overview on the topic of guessing a PDE solution thanks to this physics informed structure. The promising results reached in this work led to a series of interesting extensions and applications. Here, we propose an incomplete list of PINN-based works [31, 76, 77, 82, 99, 101, 114, 130, 164], with many variations and applications.

In this contribution, we want to extend the physics informed paradigm to OCP($\boldsymbol{\mu}$)s, presenting the main results of [46], where they deal with parametric setting both in the optimal control framework and in the forward one. The authors, propose PINN-based strategies in order to reach more robustness for NN predictions in many-query and real-time contexts. Strategies based on NNs comply with the ROM strategy of offline-online decomposition. Indeed, after a possibly long training phase of the net (offline phase), each new paramet-

ric instance can be guessed by the NN in a very small amount of time (online phase). However, it is natural that more complex problems, such as OCP($\boldsymbol{\mu}$)s, are characterized by a more time consuming training phase. In the literature, some techniques have been designed to accelerate this stage: adaptive activation functions [73, 74] and Prior Dictionary PINNs [115] are some examples. We extend the concept of physics informed networks though two approaches:

- using augmented inputs called the *extra features* that will speed up the convergence to a loss minimum;
- chancing the architecture of the NN itself, guided by the physical model one is dealing with.

In the next Section, we generalized the standard PINN formulation to the parametric setting we have always used along this thesis.

8.1.2 PINNs Formulation

Here, we introduce a different notation for PDE($\boldsymbol{\mu}$)s w.r.t. the one used in the other Chapters, due to the slightly different setting we are working in. Thus, let $G : \mathbb{Y} \to \mathbb{Q}^*$

$$G(w(\mathbf{x}, \boldsymbol{\mu})) = f(\mathbf{x}, \boldsymbol{\mu}), \tag{8.1}$$

be a PDE($\boldsymbol{\mu}$) together with some boundary conditions. Here, $w := w(\mathbf{x}, \boldsymbol{\mu}) \in \mathbb{Y}$ is the physical variable we want to approximate through a NN. As usual, $\mathbb{Y}$ and $\mathbb{Q}$ are two Hilbert spaces. Also in this Chapter, the domain is denoted by $\Omega \subset \mathbb{R}^d$ and $f(\boldsymbol{\mu}) \in \mathbb{Q}^*$ is an external forcing term. In this context, we are making no distinction on the features of the problem: it can be steady, time-dependent, linear, nonlinear. The parameter $\boldsymbol{\mu}$, in this Chapter, will only represent physical features. Furthermore, we will deal only with steady linear problems, however in [46] the reader may find some time-dependent and nonlinear examples too. Thus, $\mathbf{x}$ is a spatial input in $\mathbb{R}^d$. As already specified, the goal is to find a surrogate of the solution to (8.1). Thus, taking inspiration from the classical reference [129], we adapt the PINN concept to a parametric setting. First of all, we define the parametric residual of (8.1) as:

$$r(w(\mathbf{x}, \boldsymbol{\mu})) := G(w(\mathbf{x}, \boldsymbol{\mu})) - f(\mathbf{x}, \boldsymbol{\mu}). \tag{8.2}$$

Let us suppose to have at one's disposal some values of the boundary conditions at, say, N_b points $\{\mathbf{x}_i^b\}_{i=1}^{N_b}$ and, moreover, a bunch of N_p internal points $\{\mathbf{x}_i^p\}_{i=1}^{N_p}$ and $N_{\boldsymbol{\mu}}$ values of the parameter $\boldsymbol{\mu} \in \mathcal{D}$, i.e. $\{\boldsymbol{\mu}_i\}_{i=1}^{N_{\boldsymbol{\mu}}}$. We now build a NN that exploits the residual information. This NN, called $\hat{w}$, takes the domain point $\mathbf{x}$ as input together with the parameter $\boldsymbol{\mu}$ and gives as an output a prediction of the value of the solution $w(\mathbf{x}, \boldsymbol{\mu})$. Thus, the terminology *physics informed* is appropriate since the value of the residual is used to achieve a predicted solution which is physically consistent. Indeed, as already specified in Section 8.1, the

NN aims at reducing the value of a loss function that, in this specific case, is given by the following mean squared error:

$$MSE^{\boldsymbol{\mu}} := \frac{1}{N_{\boldsymbol{\mu}}} \sum_{i=1}^{N_{\boldsymbol{\mu}}} (MSE_b^{\boldsymbol{\mu}_i} + MSE_p^{\boldsymbol{\mu}_i}). \tag{8.3}$$

The loss presents two contributions:

- the *boundary loss* given by

$$MSE_b^{\boldsymbol{\mu}_i} := \frac{1}{N_b} \sum_{k=1}^{N_b} |\hat{w}(\mathbf{x}_k^b, \boldsymbol{\mu}_i) - w(\boldsymbol{\mu}_i)_k^b|^2, \tag{8.4}$$

  based on the boundary values of the solution $w(\boldsymbol{\mu}_i)_k^b$ for $i = k, \dots, N_b$ at the points $\{\mathbf{x}_k^b\}_{k=1}^{N_b}$;

- and the *residual loss* is:

$$MSE_p^{\boldsymbol{\mu}_i} := \frac{1}{N_p} \sum_{k=1}^{N_p} |r(\hat{w}(\mathbf{x}_k^p, \boldsymbol{\mu}_i))|^2. \tag{8.5}$$

Namely:

- the first contribution tries to learn the boundary conditions,
- while the residual loss enforces the physical model bahaviour to the prediction.

In [46], the parameters are taken into consideration as a problem input to employ the PINN in a context where many evaluations are needed to better study the phenomenon one is analysing in a small amount of time. Indeed, after an acceptable training phase time, we can exploit the PINN for a new parameter $\boldsymbol{\mu}$ to reliably predict the parametric instance one is interested in.

8.2 Physics Informed OCP($\boldsymbol{\mu}$)s

This Section shows the performances of the physics informed paradigm in the context of OCP($\boldsymbol{\mu}$)s. First of all, we extend the arguments of Section 8.1.2 to systems of multiple equations. However, in order to deal with this kind of problems, we had to conceive some tailored techniques to reduce the computational time for the training phase and to increase the accuracy of the prediction. Thus, in Section 8.2.2 we will describe *extra features* that are meant to lighten the training phase, while Section 8.2.3 will focus on the physics informed architecture (PI-Arch) building process to reach a more accurate solution. Finally, some numerical results are presented in Section 8.2.4.

8.2.1 PINNs for Multiple Equations Problems

In order to deal with OCP($\boldsymbol{\mu}$)s, we should understand how to use the PINN structure in the case of systems of multiple equations. Namely, the problem under investigation is of the following form:

$$\mathcal{G}(w(\mathbf{x}, \boldsymbol{\mu})) = \mathcal{F}(\mathbf{x}, \boldsymbol{\mu}), \tag{8.6}$$

where the n-dimensional solution is $w(\mathbf{x}, \boldsymbol{\mu}) := (w_1(\mathbf{x}, \boldsymbol{\mu}), \dots, w_n(\mathbf{x}, \boldsymbol{\mu}))$, while the system of equation is:

$$\mathcal{G}(w(\mathbf{x}, \boldsymbol{\mu})) := \begin{bmatrix} G_1(w(\mathbf{x}, \boldsymbol{\mu})) \\ \vdots \\ G_n(w(\mathbf{x}, \boldsymbol{\mu})) \end{bmatrix} \quad \text{and} \quad \mathcal{F}(\mathbf{x}, \boldsymbol{\mu}) := \begin{bmatrix} f_1(\mathbf{x}, \boldsymbol{\mu}) \\ \vdots \\ f_n(\mathbf{x}, \boldsymbol{\mu}) \end{bmatrix}. \tag{8.7}$$

In this context, in order to apply the classical PINN strategy to (8.6), we need to define a new mean square error loss, that complies with the presented multi-dimensional structure. Hence, we can refer to the $j-$th residual as:

$$r_j(w(\mathbf{x}, \boldsymbol{\mu})) := G_j(w(\mathbf{x}, \boldsymbol{\mu})) - f_j((\mathbf{x}, \boldsymbol{\mu})). \tag{8.8}$$

Now, let us assume to have been provided of a set of indexes $J = \{1, \dots, m\}$ related to the number of boundary conditions. Moreover, let us call $I = \{1, \dots, n\}$ the set of indices indicating the number of equations. Now, for a given parameter $\boldsymbol{\mu} \in \mathcal{D}$ and given the boundary collocation points $\{\mathbf{x}_i^b\}_{i=1}^{N_b^l}$ and boundary values $\{w(\boldsymbol{\mu})_k^b\}_{k=1}^{N_b^l}$ for $l \in J$ we consider:

- the new boundary loss as

$$MSE_b^{\boldsymbol{\mu}_i} := \sum_{l \in J} \frac{1}{N_b^l} \sum_{k=1}^{N_b^l} |\hat{w}(\mathbf{x}_k^b, \boldsymbol{\mu}_i) - w(\boldsymbol{\mu}_i)_k^b|^2; \tag{8.9}$$

- and the new residual loss as

$$MSE_p^{\boldsymbol{\mu}_i} := \sum_{j \in I} \frac{1}{N_p} \sum_{k=1}^{N_p} |r_j(\hat{w}(\mathbf{x}_k, \boldsymbol{\mu}_i))|^2. \tag{8.10}$$

Also in this case $\hat{w}$ is a NN built to approximate the global solution w minimizing the global loss defined as did for (8.3). Thus, the most naive approach consists in solving the problem by means of a standard PINN approach as presented in Section 8.1.2, once adapted the loss. However, in the next Section, we will discuss a strategy that gives better performances in terms of loss convergence.

8.2.2 Extra Features

As already specified in Section 8.1, a NN is a structure where nonlinear information travels through connections to predict an output. It is clear that, the greater is the structure and, thus, the number of functions defined on it, the more and more complex the output of the network model will be. It is well-known [75, 77] that increasing the *depth* of the NN, i.e. the number of hidden layers, allows the net to predict highly nonlinear behaviours. However, working with a huge number of hidden layers can lead to an unbearable training phase. A way out can be represented by the employment of an augmented input that allows to contain the depth of the NN together with an acceptable training complexity. In the numerical example we are going to present, the inputs are the spatial coordinates and the parameter value $\boldsymbol{\mu}$, i.e. $\mathbf{x} = \begin{bmatrix} x_1 & \dots & x_d & \mu_1 & \dots & \mu_D \end{bmatrix}$. Now, the model knowledge about the problem can be used to extend the input information through the employment of one (or more) kernel function (functions[1]) as augmented first-layer input data. We will call these functions the *extra features* from now on, and they will be denoted with $\{k_i(\mathbf{x})\}_{i=1}^{N_f}$. Thus, the final input will be of the following form:

$$\mathbf{x}_{\text{extra}} = \begin{bmatrix} \mathbf{x} & k_1(\mathbf{x}) & \boldsymbol{\mu} & \dots & k_{N_f}(\mathbf{x}) \end{bmatrix} \tag{8.11}$$

of dimension $d+D+N_f$. Exploiting a proper set of extra features might lead to an easier and smaller NN to train, since they highlight the correlation between the input and the expected output. Of course, this augmented strategy is highly problem dependent and it has to be properly tuned. However, we propose few tips to select the features:

- the analytical formulation of the term $f(\mathbf{x}, \boldsymbol{\mu})$ might represent a good asset to the learning procedure since the output is usually strongly related to the forcing term of the equation.

- Simple kernel functions that satisfy the boundary conditions might improve the training procedure leading to a fast convergence of the NN to a physically meaningful solution.

In this contribution we will only focus on *fixed* extra features, however, in [46] they explore also *learnable* adaptive functions that guarantee an improvement both in terms of accuracy and computational time needed for the training phase. In the next Section, we will explore the possibility to exploit model information not only in the loss function definition, but also in the design phase of the NN, building a *physics informed architecture* (PI-Arch) that is able to better tackle PDE($\boldsymbol{\mu}$)s with more than one variable.

[1] which should be differentiable.

8.2.3 The PI-Arch strategy

The idea behind this Section is to extend the PINN paradigm in order to build tailored NN architectures that can increase the accuracy of the prediction, building on the techniques for multi–fidelity approaches used in [19, 60, 107]. We postpone the comments on the advantages of such a strategy in Section 8.2.4, while, now, we are going to provide a general description of it. Let us suppose of having been provided of a system of multiple equations as (8.6) with n equations and m boundary conditions. Also in this case, the loss is (8.3), with the same definition for $MSE_p^{\boldsymbol{\mu}}$ and $MSE_b^{\boldsymbol{\mu}}$ as (8.10) and (8.9). We want to change the architecture of $\hat{w}$ exploiting the information of the system at hand to reach a more reliable output prediction through a tailored (combination of) NN (NNs). Let us consider $H_1 \subset I$. Here, a first NN takes the inputs and gives as output only a part of the problem variables, say $\{\hat{w}_{h_1}(\mathbf{x}, \mu)\}_{h_1 \in H_1}$. We call this set the *one-level output* (1-out). Now, the 1-out (or a part of it) together with, eventually, the initial input (or a part of it) is interpreted as an input for a new NN that predicts another set of variables, say $\{\hat{w}_{h_2}(\mathbf{x}, \mu)\}_{h_2 \in H_2}$, where $H_2 \subset I \setminus H_1$. This will be the *two-level output* (2-out). These predicted variables combined with the 1-out (or a part of it) and the initial input (or part of it) leads to the *three-level output* (3-out) based on the indices $H_3 \subset I \setminus \{H_1 \cup H_2\}$, and so on. The process is repeated for $k-$times: the final output is $\{\hat{w}_{h_k}(\mathbf{x}, \mu)\}_{h_k \in H_k}$ (k-out), with

$$\bigcup_k H_k = I.$$

Moreover, let us suppose that we are able to extrapolate a direct or indirect relation between the variables from the equations, say Υ, between the outputs. Say that we can connect 1-out to the 2-out through

$$\Upsilon(\text{1-out}) \approx \text{2-out}. \tag{8.12}$$

Once provided of this information, thus we can

- replace the NN between 1-out and 2-out with Υ(1-out), if we are confident enough of the relation,
- or build a NN to approximate Υ(1-out).

Figure 8.2 shows an example of such an architecture. This technique helps in isolating between the various outputs of the NN model. Thus, we proposed to exploit some prior knowledge, that, otherwise, must be learned by the system though a longer training phase and with a larger number of hidden layers. In this sense, this concept totally complies with the physics informed paradigm and this is the reason why we will refer to this strategy as *physics informed architecture* (PI-Arch). The next Section shows how the proposed physics-based technique can be effective in the prediction of solution to OCP($\boldsymbol{\mu}$)s.

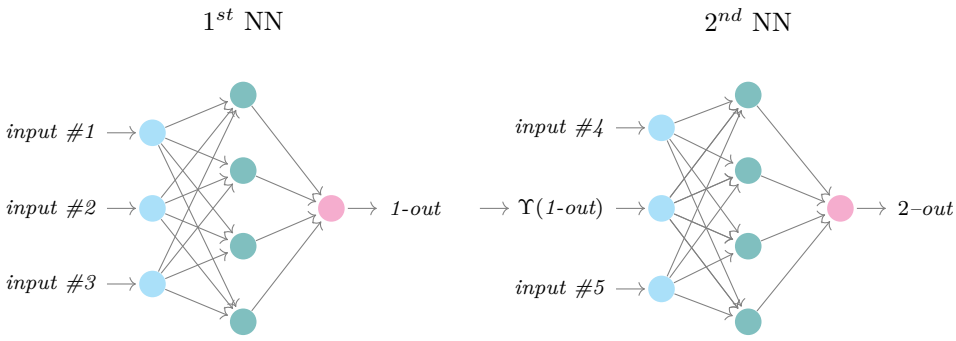


Figure 8.2: *PINNs and OCP($\boldsymbol{\mu}$)s.* PI-Arch example.

8.2.4 Numerical Results

We tested the proposed methodology with a steady OCP($\boldsymbol{\mu}$) defined in the physical domain $\Omega = [-1,1]\times[-1,1]$. The problem reads: given $\boldsymbol{\mu} := (\mu_1, \mu_2) \in [0.5,3] \times [0.01,1]$ find $(y(\mathbf{x},\boldsymbol{\mu}), u(\mathbf{x},\boldsymbol{\mu})) \in H_0^1(\Omega) \times L^2(\Omega) := Y \times U$ such that verifies

$$\min_{(y(\mathbf{x},\boldsymbol{\mu}),u(\mathbf{x},\boldsymbol{\mu}))\in Y\times U} \frac{1}{2}\|y(\mathbf{x},\boldsymbol{\mu}) - \mu_1\|^2_{L^2(\Omega)} + \frac{\mu_2}{2}\|u(\mathbf{x},\boldsymbol{\mu})\|^2_{L^2(\Omega)},$$
$$\text{constrained to} \tag{8.13}$$
$$\begin{cases} -\Delta y(\mathbf{x},\boldsymbol{\mu}) = u(\mathbf{x},\boldsymbol{\mu}) & \text{in } \Omega, \\ y(\mathbf{x},\boldsymbol{\mu}) = 0 & \text{on } \partial\Omega. \end{cases}$$

Namely, the desired state is $y_{\mathrm{d}}(\mathbf{x},\boldsymbol{\mu}) \equiv \mu_1$ on the whole physical domain. Moreover, it is the first time in this contribution, that we are dealing with a parametric penalization term, i.e. $\mu_2 = \alpha$. Namely, we want to change the system in order to reach the value μ_1 under a different action of the control variable given by several instances of μ_2. The Lagrangian technique described in Section 1.1.1 applied to problem 8.13 leads to the following optimality system

$$\begin{cases} y(\mathbf{x},\boldsymbol{\mu}) - \Delta z(\mathbf{x},\boldsymbol{\mu}) = \mu_1 & \text{in } \Omega, \\ z(\mathbf{x},\boldsymbol{\mu}) = 0 & \text{on } \partial\Omega, \\ \mu_2 u(\mathbf{x},\boldsymbol{\mu}) = z(\mathbf{x},\boldsymbol{\mu}) & \text{in } \Omega, \\ -\Delta y(\mathbf{x},\boldsymbol{\mu}) = u(\mathbf{x},\boldsymbol{\mu}) & \text{in } \Omega, \\ y(\mathbf{x},\boldsymbol{\mu}) = 0 & \text{on } \partial\Omega. \end{cases} \tag{8.14}$$

Here, we employed the PI-Arch technique to build $\hat{w}$. The NN will take $(\mathbf{x},\boldsymbol{\mu})$ as input and will give the three predicted variables $(y(\mathbf{x},\boldsymbol{\mu}), u(\mathbf{x},\boldsymbol{\mu}), z(\mathbf{x},\boldsymbol{\mu}))$.

In this setting, we first tried the standard PINN strategy proposed in Section 8.2.1. The NN consisted in 3 hidden layers. The first two layers presented 40 neurons, while the last one had only 20 neurons. We exploited an ADAM optimizer with a Softplus activation function. The learning rate was fixed at 0.002. We used the following number of collocation points: $N_p = 900$, $N_b = 200$ and $N_{\boldsymbol{\mu}} = 50$. Furthermore, in order to help the convergence towards the right boundary conditions, we used

$$k_1(x_0, x_1) := (1 - x_0^2)(1 - x_1^2). \tag{8.15}$$

as extra feature. Figure 8.4 shows the prediction of the three variables after 10000 epochs training. The top row depicts the state, the control and the adjoint variables for $\boldsymbol{\mu} = (3, 1)$. While the results for this parameter were quite promising, we noticed an issue for the evaluation of $\boldsymbol{\mu} = (3, 0.01)$, as one may notice from the bottom row of the plot. Namely, the standard PINN approach did not recover the optimality equation relation $\mu_2 u(\mathbf{x}, \boldsymbol{\mu}) = z(\mathbf{x}, \boldsymbol{\mu})$, and the behaviour of the adjoint variable $z(\mathbf{x}, \boldsymbol{\mu})$ and of the control variable is completely different. Hence, we decided to directly use the optimality equation in a PI-Arch structure. Figure 8.3 is a scheme of the NN structure we used for this specific application. A first NN, from the input $(\mathbf{x}, \boldsymbol{\mu}, k_1(x_0, x_1))$, gives an approximation of the control variable and the state variable (the 1-out). Then, the control variable is an input to another NN that, together with the value μ_2, defines the adjoint variable as $\Upsilon(1\text{-out}) = \mu_2 u(\mathbf{x}, \boldsymbol{\mu})$. The advantages related to this tailored structure are represented in Figure 8.5. Indeed, the use of $\Upsilon(1\text{-out})$ in $\hat{w}$ increases the accuracy of the prediction: on one side, we still remained precise in the approximation of the case $\boldsymbol{\mu} = (3, 1)$, and, on the other, we well quite accurate also for smaller values of μ_2, i.e. for $\boldsymbol{\mu} = (3, 0.01)$. Now, the adjoint and the control variables are completely consistent w.r.t. the model. Furthermore, we tested the accuracy of the PI-Arch structure comparing the prediction of $\hat{w}$ in the domain point $(x_0, x_1) = (0, 0)$ to the FE solution in $\mu_1 = 1, 2, 3$ and $\mu_2 = 1, 0.1, 0.01$.

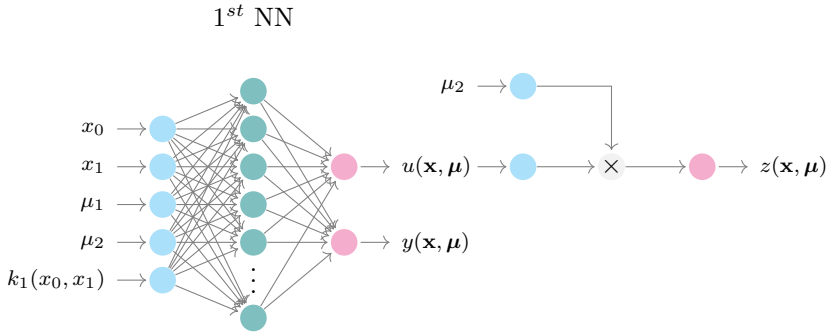


Figure 8.3: *PINNs and OCP($\boldsymbol{\mu}$)s.* PI-Arch used to solve problem (8.13).

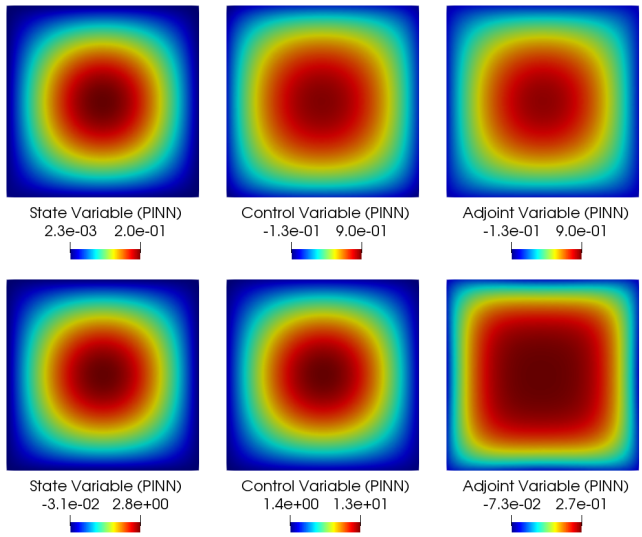


Figure 8.4: *PINNs and OCP($\boldsymbol{\mu}$)s.* *Top row.* PINN approximation for $\boldsymbol{\mu} = (3, 1)$. *Bottom row.* PINN approximation for $\boldsymbol{\mu} = (3, 0.01)$. *Left.* State variable. *Center.* Control variable. *Right.* Adjoint variable.

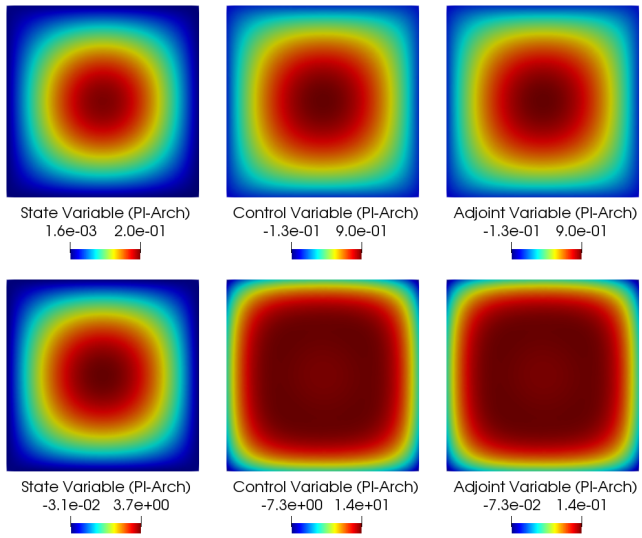


Figure 8.5: *PINNs and OCP($\boldsymbol{\mu}$)s.* *Top row.* PI-Arch approximation for $\boldsymbol{\mu} = (3, 1)$. *Bottom row.* PI-Arch approximation for $\boldsymbol{\mu} = (3, 0.01)$. *Left.* State variable. *Center.* Control variable. *Right.* Adjoint variable.

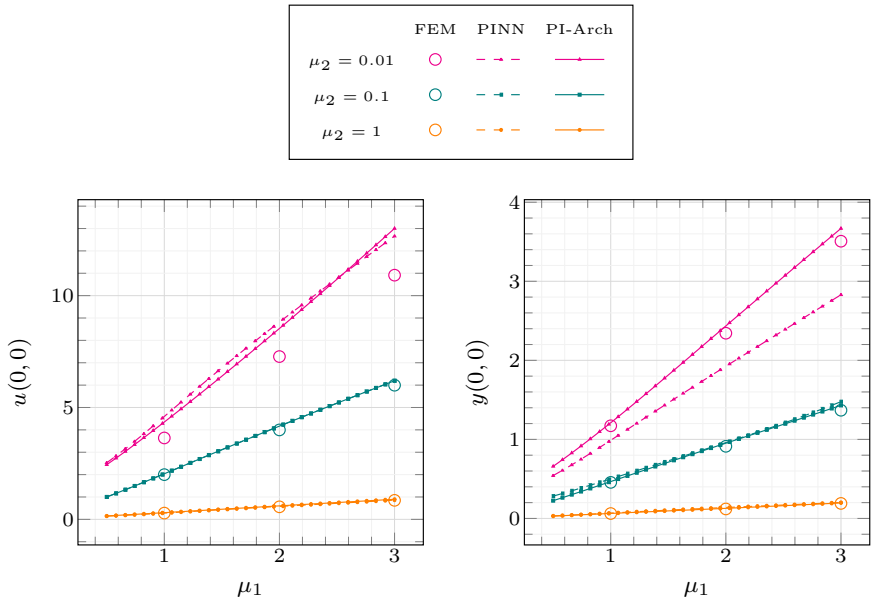


Figure 8.6: *PINNs and OCP($\boldsymbol{\mu}$)s.* Prediction for $\mu_2 = 1, 0.1, 0.01$ with respect to μ_1 in $(x_0, x_1) = (0, 0)$ compared to the FE approximation for $\mu_1 = 1, 2, 3$. *Left.* Control variable. *Right.* State variable.

The results are depicted in Figure 8.6, where also a comparison w.r.t. the standard PINN model is shown. First of all, we remark that exploiting the PI-Arch structure improves the prediction of the state variable for $\mu_2 = 0.01$ reaching values that are closer to the FE approximation. Moreover, the prediction of the state variable is quite accurate for all the values of μ_2 (right plot). However, the prediction is slightly worse when dealing with the control for $\mu_2 = 0.01$ (left plot). Nonetheless, we believe that the results are promising. Indeed, this is a first step towards the application of such strategies to more complex OCP($\boldsymbol{\mu}$)s. Furthermore, we would like to highlight that a more tailored PI-Arch and few modifications of the structure combined with a possibly longer training phase might allows us to reach more accurate results. However, the tuning of the $\hat{w}$ goes beyond the purposes of [46] where the authors aim at providing a NN-based tool to use in a many-query and real-time context for many PDE($\boldsymbol{\mu}$)s.

This Chapter concludes the set of advanced applications related to OCP($\boldsymbol{\mu}$)s. Here, we exploited a totally different approach based on a ML framework. In the following Chapter, the reader will find some comments and perspectives about what we discussed in this contribution.

Conclusions and Perspectives

In this work, we have proposed ROMs as reliable and fast tools to deal with parametric nonlinear and time-dependent OCP($\boldsymbol{\mu}$)s. We have shown a space-time approach that can be applied to several controlled PDE($\boldsymbol{\mu}$)s. Furthermore, we adapted this general context also to steady problems, proposing several applications, moving from environmental sciences to artificial-intelligence-based parametric prediction. The optimal control formulation has been described both at the continuous and at the discrete level. As already specified, we tackled several kind of equations: time-dependent nonlinear problems, steady nonlinear problems, parabolic problems (both in the standard and in the no-control framework), steady linear problems.
Moved by the computational complexity of such systems, we employed model order reduction to reduced the costs of the parametric analysis of the optimal control framework. We proposed two algorithms, the space-time POD and the space-time Greedy. Both the strategies have been validated through numerical examples in the linear and nonlinear setting.
Nonetheless, we focused on different types of applications: optimal controls for environmental sciences in a deterministic and stochastic fashion, a control over the Coanda effect deriving from bifurcations arising in Navier-Stokes equations and a neural network approach for parametric prediction. Let us summarize some of the main features emerged in this work.

- ◊ In the space-time setting for the classical three-equation system, the saddle point structure for linear quadratic OCP($\boldsymbol{\mu}$)s is maintained and, thus, the well-posedness of the problem can be proved through standard arguments based on Brezzi theorem. This resulted into a natural application of a partitioned POD to the space-time variables, not separating the time instances. This approach is versatile (we treat a time-dependent problem as a steady one) and features large values of speedup.

- ◊ The aforementioned reduced space-time framework has been exploited in more complex systems, like the nonlinear time-dependent viscous SWEs with physical and geometrical parameters. Also in this case we showed that model order reduction could be a suitable tool to rapidly simulate marine dynamics.

- ◊ We proposed the *no-control* control framework as an alternative approach to deal with parabolic problems. The well-posedness results paved the way to the definition of a new error estimator that allowed us to employ a space-time Greedy strategy. This algorithm, due to its iterative and adaptive nature, lightened the costs of the construction of the bases. One of the main advantages of this error estimator relies on its explicit expression,

made by known a priori quantities that are easy to compute. Furthermore, it is also valid for steady elliptic problems and this avoids the use of expensive approximations algorithm as did, for example, in [111].

Besides these merits, the proposed space-time formulation might be not appropriate in actual application due to the costs of the related offline phase. Indeed, some improvements are needed in this direction: the direct solution of the space-time optimality system might affect the accuracy of the high fidelity solution and, consequently, of the reduced one. Moreover, a huge amount of computational resources has been paid in the offline simulations and this issue has been only partially solved through the employment of the space-time Greedy approach. Nonetheless, we accomplished our goal of showing how ROMs are an effective strategy to lower the computational time needed for space-time parametric simulations and we think that this is a first step towards a deeper analysis and further developments of such a research line interpreting it as a stepping stone for more complicated applications.
The last part of thesis dealt with applications with different focuses and interests.

- ◊ The first field we wanted to analyse was marine and ecological sciences. The two shown examples have been treated both in a deterministic and in a stochastic way. The presented reduced approaches led to promising results in terms of speedup and parametric analysis in a data assimilation framework. Moreover, we proposed the weighted ROMs as an accelerating factor, remarking the concept of introducing parameter information (such as parameter distribution) into the reduced model to make the simulations even faster.

- ◊ We discussed the role of OCP($\boldsymbol{\mu}$)s in bifurcating systems as a first step towards a stability analysis of controlled nonlinear problems. Moved by the will of changing the solution features to reach a preferable configuration, we tried several numerical results. They covered a wide spectrum of behaviours: a particular example is given by the Dirichlet case. The test cases might offer many improvements and developments on the role of OCP($\boldsymbol{\mu}$)s as attractors.

- ◊ Lastly, we interpret PINN as an extension of the offline-online paradigm in an artificial-intelligence-based context. Indeed, the training phase can replace the basis functions construction stage: after this possibly long training, the model rapidly predicts several $\boldsymbol{\mu}$-dependent solutions. We remark that we exploited several schemes to accelerate the training process.

The presented results have room for improvement and many questions remain open and of interest for future developments.

We would like to tackle more efficiently the issue related to the offline costs of the space-time formulation. As already specified in Chapter 2, it can be overcome using more computational resources and through *ad hoc* preconditioning and multigrid approaches [21, 138, 143, 144]. However, we remark that the proposed iterative approaches might only moderately decrease the computational effort needed for the space-time simulation. For this reason, the employment of data-assisted techniques can be of assistance in this setting. In an ongoing collaboration with E. Donadini and M. Tezzele, we are moving the first steps to investigate how Dynamic Mode Decomposition [87] can help in predict time dynamics in the optimal control framework.
Another interesting question may arise from the SWEs example. Indeed, even if the proposed test case represents standard laboratory-reproducible scenarios, it is rather academic. Yet, investigating a more complex example might present challenging features due to the hyperbolic nature of the state equation, for instance:

- *instability.* Some numerical oscillations might occur and they can transfer (or not) to the optimal control framework.
- *Difficult reduced reconstruction.* This property is well known even for simpler one-dimensional test cases. Indeed, the ROMs struggle in recovering advection-dominated hyperbolic dynamics [56, 113].

These topics are part of an ongoing master project carried out with F. Zoccolan. We aim at addressing the issue of convection dominated stochastic OCP($\boldsymbol{\mu}$)s and we will try to understand whether the stabilizing schemes are needed for all the equations of the optimality system. Furthermore, this line of research is strictly related to environmental sciences, since, both the random inputs and the stabilized model can be employed in more realistic applications.
Concerning marine ecosystem control, in collaboration with G. Karniadakis, we would like to exploit OCP($\boldsymbol{\mu}$)s for the prediction and management of Massachusetts Bay water acidification. The project fits with the idea of OCP($\boldsymbol{\mu}$)s as a tool of actual interest in many interdisciplinary fields.
Moreover, many questions also arose in the bifurcating OCP($\boldsymbol{\mu}$)s: it is clear that control variables were changing the system at hand. This feature can be used in several applications since nonlinear dynamics characterizes many research and industrial fields. We presented results concerning the pitchfork bifurcation diagram, however, in the Dirichlet case a new type of phenomenon was occurring (possibly and Hopf bifurcation) and, thus, a deeper analysis is needed in that setting.
Last, focusing on the presented physics informed strategy, we would like, first of all, to tackle more complex problems in the OCP($\boldsymbol{\mu}$)s framework. Moreover, all the discussed physics informed techniques are highly problem dependent. A line of research to explore might concern a better choice for extra features and PI-Arch structures.

In all the proposed applications, we wanted to highlight how optimal control is a tool suited in many scientific contexts to bridge the gap between data and mathematical modelling. Additionally, model order reduction can help in many scenarios. Thus, exploiting OCP($\boldsymbol{\mu}$)s is a natural choice to add data information in the system without spoiling its physical meaning. We believe that the proposed techniques enhanced with these improvements and developments would make OCP($\boldsymbol{\mu}$)s even more widespread in several fields of science and industry.

References

[1] multiphenics - easy prototyping of multiphysics problems in fenics, https://mathlab.sissa.it/multiphenics.

[2] RBniCS – reduced order modelling in FEniCS, http://mathlab.sissa.it/rbnics.

[3] V. I. Agoshkov, D. Ambrosi, V. Pennati, A. Quarteroni, and F. Saleri. Mathematical and numerical modelling of shallow water flow. *Computational Mechanics*, 11(5):280–299, Sep 1993.

[4] V. I. Agoshkov, E. Ovchinnikov, A. Quarteroni, and F. Saleri. Recent developments in the numerical simulation of shallow water equations II: Temporal discretization. *Mathematical Models and Methods in Applied Sciences*, 4(04):533–556, 1994.

[5] V. I. Agoshkov, A. Quarteroni, and F. Saleri. Recent developments in the numerical simulation of shallow water equations I: Boundary conditions. *Applied Numerical Mathematics*, 15(2):175–200, 1994.

[6] V. I. Agoshkov, F. Saleri, and E. Miglio. An optimal control approach for 1D-2D shallow water equations coupling. In *Communications to SIMAI Congress*, volume 1, 2007.

[7] A. Agrachev and A. Sarychev. Navier–Stokes equations: controllability by means of low modes forcing. *Journal of Mathematical Fluid Mechanics*, 7(1):108–152, 2005.

[8] E. Allgower and K. Georg. *Introduction to Numerical Continuation Methods*. Society for Industrial and Applied Mathematics, 2003.

[9] A. Ambrosetti and G. Prodi. *A Primer of Nonlinear Analysis*. Cambridge University Press, Cambridge, 1993.

[10] A. C. Antoulas, C. A. Beattie, and S. Güğercin. *Interpolatory methods for model reduction*. SIAM, 2020.

[11] I. Babuška. Error-bounds for finite element method. *Numerische Mathematik*, 16(4):322–333, Jan 1971.

[12] E. Bader, M. Kärcher, M. A. Grepl, and K. Veroy. Certified reduced basis methods for parametrized distributed elliptic optimal control problems with control constraints. *SIAM Journal on Scientific Computing*, 38(6):A3921–A3946, 2016.

[13] E. Bader, M. Kärcher, M. A. Grepl, and K. Veroy-Grepl. A certified reduced basis approach for parametrized linear-quadratic optimal control problems with control constraints. *IFAC-PapersOnLine*, 48(1):719–720, 2015.

[14] F. Ballarin, E. Faggiano, A. Manzoni, A. Quarteroni, G. Rozza, S. Ippolito, C. Antona, and R. Scrofani. Numerical modeling of hemodynamics scenarios of patient-specific coronary artery bypass grafts. *Biomechanics and Modeling in Mechanobiology*, 16(4):1373–1399, Aug 2017.

[15] 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.

[16] F. Ballarin, G. Rozza, and M. Strazzullo. Reduced order methods for parametric flow control problems and applications. Submitted, 2020, https://arxiv.org/abs/2011.12101.

[17] 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.

[18] L. Bauer and E. L. Reiss. Nonlinear buckling of rectangular plates. *Journal of the Society for Industrial and Applied Mathematics*, 13(3):603–626, 1965.

[19] A. G. Baydin, B. A. Pearlmutter, A. A. Radul, and J. M. Siskind. Automatic differentiation in machine learning: a survey. *Journal of machine learning research*, 18, 2018.

[20] P. Benner, M. Ohlberger, A. Cohen, and K. Willcox. *Model reduction and approximation: theory and algorithms*. SIAM, 2017.

[21] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[22] M. Benzi and V. Simoncini. On the eigenvalues of a class of saddle point matrices. *Numer. Math*, 103:173–196, 2006.

[23] M. Benzi and A. J. Wathen. *Some Preconditioning Techniques for Saddle Point Problems*, pages 195–211. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[24] M. S. Berger. On Von Kármán's equations and the buckling of a thin elastic plate, I the clamped plate. *Communications on Pure and Applied Mathematics*, 20(4):687–719, 1967.

[25] C. M. Bishop. Pattern recognition. *Machine learning*, 128(9), 2006.

[26] P. B. Bochev and M. D. Gunzburger. *Least-squares finite element methods*, volume 166. Springer-Verlag, New York, 2009.

[27] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44. Springer-Verlag, Berlin and Heidelberg, 2013.

[28] F. Brezzi. On the existence, uniqueness and approximation of saddle-point problems arising from lagrangian multipliers. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique*, 8(R2):129–151, 1974.

[29] A. Buffa, Y. Maday, A. T. Patera, C. Prud'homme, and G. Turinici. A priori convergence of the greedy algorithm for the parametrized reduced basis method. *ESAIM: Mathematical modelling and numerical analysis*, 46(3):595–603, 2012.

[30] J. Burkardt, M. Gunzburger, and H. Lee. POD and CVT-based reduced-order modeling of Navier–Stokes flows. *Computer Methods in Applied Mechanics and Engineering*, 196(1-3):337–355, 2006.

[31] S. Cai, Z. Wang, S. Wang, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks for heat transfer problems. *Journal of Heat Transfer*, 143(6):060801, 2021.

[32] G. Caloz and J. Rappaz. Numerical analysis for nonlinear and bifurcation problems, 1997.

[33] G. Carere, M. Strazzullo, F. Ballarin, G. Rozza, and R. Stevenson. Weighted POD-reduction for parametrized optimal control problems with random inputs applied to environmental sciences. Submitted, 2021, https://arxiv.org/abs/2103.00632.

[34] 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.

[35] D. Chapelle, A. Gariah, P. Moireau, and J. Sainte-Marie. A Galerkin strategy with proper orthogonal decomposition for parameter-dependent problems: Analysis, assessments and applications to parameter estimation. *ESAIM: Mathematical Modelling and Numerical Analysis*, 47(6):1821–1843, 2013.

[36] E. Charalampidis, P. Kevrekidis, and P. Farrell. Computing stationary solutions of the two-dimensional Gross–Pitaevskii equation with deflated continuation. *Communications in Nonlinear Science and Numerical Simulation*, 54:482 – 499, 2018.

[37] P. Chen, A. Quarteroni, and G. Rozza. A weighted reduced basis method for elliptic partial differential equations with random input data. *SIAM Journal on Numerical Analysis*, 51(6):3163–3185, 2013.

[38] P. Ciarlet. *Linear and Nonlinear Functional Analysis with Applications*. Other Titles in Applied Mathematics. Society for Industrial and Applied Mathematics, 2013.

[39] P. Constantin and C. Foias. *Navier-Stokes equations*. University of Chicago Press, 1988.

[40] J. M. Coron. On the controllability of the 2-D incompressible Navier-Stokes equations with the Navier slip boundary conditions. *ESAIM: Control, Optimisation and Calculus of Variations*, 1:35–75, 1996.

[41] J. M. Coron. *Control and nonlinearity*. Number 136. American Mathematical Soc., 2007.

[42] 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.

[43] L. Dedè. Optimal flow control for Navier-Stokes equations: Drag minimization. *International Journal for Numerical Methods in Fluids*, 55(4):347–366, 2007.

[44] 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.

[45] 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.

[46] N. Demo, M. Strazzullo, and G. Rozza. An extended physics informed neural network for preliminary analysis of parametric optimal control problems. In preparation, 2021.

[47] C. H. Ding and I. Dubchak. Multi-class protein fold recognition using support vector machines and neural networks. *Bioinformatics*, 17(4):349–358, 2001.

[48] K. Eriksson and C. Johnson. Error estimates and automatic time step control for nonlinear parabolic problems, I. *SIAM Journal on Numerical Analysis*, 24(1):12–23, 1987.

[49] 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.

[50] S. Ferrari and F. Saleri. A new two-dimensional shallow water model including pressure effects and slow varying bottom topography. *ESAIM: Mathematical Modelling and Numerical Analysis*, 38(2):211–234, 2004.

[51] 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.

[52] M. Ghil and P. Malanotte-Rizzoli. Data assimilation in meteorology and oceanography. *Advances in geophysics*, 33:141–266, 1991.

[53] S. Glas, A. Mayerhofer, and K. Urban. *Two Ways to Treat Time in Reduced Basis Methods*, pages 1–16. Springer International Publishing, Cham, 2017.

[54] R. Glowinski. Ensuring well-posedness by analogy; stokes problem and boundary control for the wave equation. *Journal of Computational Physics*, 103(2):189 – 221, 1992.

[55] I. Goodfellow, Y. Bengio, and A. Courville. *Deep learning*. MIT press, 2016.

[56] C. Greif and K. Urban. Decay of the kolmogorov n-width for wave problems. *Applied Mathematics Letters*, 96:216–222, 2019.

[57] S. Grimberg, C. Farhat, and N. Youkilis. On the stability of projection-based model order reduction for convection-dominated laminar and turbulent flows. *Journal of Computational Physics*, 419:109681, 2020.

[58] R. Guberovic, C. Schwab, and R. Stevenson. Space-time variational saddle point formulations of Stokes and Navier–Stokes equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 48(3):875–894, 2014.

[59] M. D. Gunzburger. *Perspectives in flow control and optimization*, volume 5. SIAM, Philadelphia, 2003.

[60] M. Guo, A. Manzoni, M. Amendt, P. Conti, and J. S. Hesthaven. Multi-fidelity regression using artificial neural networks: efficient approximation of parameter-dependent output quantities. *arXiv preprint arXiv:2102.13403*, 2021.

[61] B. Haasdonk and M. Ohlberger. Reduced basis method for finite volume approximations of parametrized linear evolution equations. *ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 42(2):277–302, 2008.

[62] M. Hess, A. Alla, A. Quaini, G. Rozza, and M. Gunzburger. A localized reduced-order modeling approach for PDEs with bifurcating solutions. *Computer Methods in Applied Mechanics and Engineering*, 351:379 – 403, 2019.

[63] M. Hess, A. Quaini, and G. Rozza. Reduced basis model order reduction for Navier-Stokes equations in domains with walls of varying curvature. *International Journal of Computational Fluid Dynamics*, 34(2):119–126, 2020.

[64] J. S. Hesthaven, G. Rozza, and B. Stamm. Certified reduced basis methods for parametrized partial differential equations. *SpringerBriefs in Mathematics*, 2015, Springer, Milano.

[65] M. Hinze, M. Köster, and S. Turek. A hierarchical space-time solver for distributed control of the Stokes equation. *Technical Report, SPP1253-16-01*, 2008.

[66] M. Hinze, M. Köster, and S. Turek. A space-time multigrid method for optimal flow control. In *Constrained optimization and optimal control for partial differential equations*, pages 147–170. Springer, 2012.

[67] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23. Springer Science & Business Media, Antwerp, 2008.

[68] M. Holtz. *Sparse Grid Quadrature in High Dimensions with Applications in Finance and Insurance*. Springer, 2010.

[69] D. Huynh, D. Knezevic, Y. Chen, J. S. Hesthaven, and A. Patera. A natural-norm successive constraint method for inf-sup lower bounds. *Computer Methods in Applied Mechanics and Engineering*, 199(29-32):1963–1975, 2010.

[70] L. Iapichino, S. Trenz, and S. Volkwein. Reduced-order multiobjective optimal control of semilinear parabolic problems. In B. Karasözen, M. Manguoğlu, M. Tezer-Sezgin, S. Göktepe, and Ö. Uğur, editors, *Numerical Mathematics and Advanced Applications ENUMATH 2015*, pages 389–397, Cham, 2016. Springer International Publishing.

[71] L. Iapichino, S. Ulbrich, and S. Volkwein. Multiobjective pde-constrained optimization using the reduced-basis method. *Adv. Comput. Math.*, 43(5):945–972, Oct. 2017.

[72] K. Ito and K. Kunisch. *Lagrange multiplier approach to variational problems and applications*, volume 15. SIAM, Philadelphia, 2008.

[73] A. D. Jagtap, K. Kawaguchi, and G. Em Karniadakis. Locally adaptive activation functions with slope recovery for deep and physics-informed neural networks. *Proceedings of the Royal Society A*, 476(2239):20200334, 2020.

[74] A. D. Jagtap, K. Kawaguchi, and G. E. Karniadakis. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *Journal of Computational Physics*, 404:109136, 2020.

[75] A. D. Jagtap, K. Kawaguchi, and G. E. Karniadakis. Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *Journal of Computational Physics*, 404:109136, 2020.

[76] A. D. Jagtap, E. Kharazmi, and G. E. Karniadakis. Conservative physics-informed neural networks on discrete domains for conservation laws: Applications to forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering*, 365:113028, 2020.

[77] X. Jin, S. Cai, H. Li, and G. E. Karniadakis. Nsfnets (navier-stokes flow nets): Physics-informed neural networks for the incompressible navier-stokes equations. *Journal of Computational Physics*, 426:109951, 2021.

[78] E. Kalnay. *Atmospheric modeling, data assimilation and predictability.* Cambridge university press, 2003, Cambridge.

[79] M. Kärcher and M. A. Grepl. A certified reduced basis method for parametrized elliptic optimal control problems. *ESAIM: Control, Optimisation and Calculus of Variations*, 20(2):416–441, 2014.

[80] M. Kärcher, Z. Tokoutsi, M. A. Grepl, and K. Veroy. Certified reduced basis methods for parametrized elliptic optimal control problems with distributed controls. *Journal of Scientific Computing*, 75(1):276–307, 2018.

[81] P. Kevrekidis, D. Frantzeskakis, and R. Carretero-González. *The Defocusing Nonlinear Schrödinger Equation.* Society for Industrial and Applied Mathematics, Philadelphia, PA, 2015.

[82] E. Kharazmi, Z. Zhang, and G. E. Karniadakis. hp-vpinns: Variational physics-informed neural networks with domain decomposition. *Computer Methods in Applied Mechanics and Engineering*, 374:113547, 2021.

[83] H. Kielhöfer. *Bifurcation Theory: An Introduction with Applications to PDEs.* Applied Mathematical Sciences. Springer New York, 2006.

[84] T. 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.

[85] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105, 2012.

[86] K. Kunisch and S. Volkwein. Proper orthogonal decomposition for optimality systems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 42(1):1–23, 2008.

[87] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor. *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016.

[88] Y. Kuznetsov. *Elements of Applied Bifurcation Theory*. Applied Mathematical Sciences. Springer New York, 2004.

[89] U. Langer, O. Steinbach, F. Tröltzsch, and H. Yang. Unstructured space-time finite element methods for optimal control of parabolic equations. 04 2020.

[90] T. Lassila, A. Manzoni, A. Quarteroni, and G. Rozza. A reduced computational and geometrical framework for inverse problems in hemodynamics. *International Journal for Numerical Methods in Biomedical Engineering*, 29(7):741–776, 2013.

[91] X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan. All-optical machine learning using diffractive deep neural networks. *Science*, 361(6406):1004–1008, 2018.

[92] J. L. Lions. *Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles*. Dunod, 1968.

[93] J. L. Lions. *Quelques méthodes de résolution des problemes aux limites non linéaires*. Dunod, 1969.

[94] J. L. Lions. Exact controllability, stabilization and perturbations for distributed systems. *SIAM review*, 30(1):1–68, 1988.

[95] J. L. Lions and E. Zuazua. Contrôlabilité exacte des approximations de Galerkin des équations de Navier-Stokes. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 324(9):1015–1021, 1997.

[96] J. L. Lions and E. Zuazua. Exact boundary controllability of Galerkin's approximations of Navier-Stokes equations. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze*, 26(4):605–621, 1998.

[97] J. L. Lions and E. Zuazua. A generic uniqueness result for the Stokes system and its control theoretical consequences. *Partial differential equations and applications: Collected Papers in Honor of Carlo Pucci*, page 221, 2017.

[98] A. Logg, K. Mardal, and G. Wells. *Automated Solution of Differential Equations by the Finite Element Method*. Springer-Verlag, Berlin, 2012.

[99] M. Mahmoudabadbozchelou, M. Caggioni, S. Shahsavari, W. H. Hartt, G. Em Karniadakis, and S. Jamali. Data-driven physics-informed constitutive metamodeling of complex fluids: A multifidelity neural network (mfnn) framework. *Journal of Rheology*, 65(2):179–198, 2021.

[100] A. Manzoni, A. Quarteroni, and S. Salsa. Mox-report no . 13 / 2019. A saddle point approach to an optimal boundary control problem for steady Navier-Stokes equations. page 252, 2019.

[101] X. Meng, Z. Li, D. Zhang, and G. E. Karniadakis. Ppinn: Parareal physics-informed neural network for time-dependent pdes. *Computer Methods in Applied Mechanics and Engineering*, 370:113250, 2020.

[102] S. Middelkamp, P. Kevrekidis, D. Frantzeskakis, R. Carretero-González, and P. Schmelcher. Emergence and stability of vortex clusters in Bose-Einstein condensates: A bifurcation approach near the linear limit. *Physica D: Nonlinear Phenomena*, 240(18):1449 – 1459, 2011.

[103] E. Miglio, S. Perotto, and F. Saleri. Model coupling techniques for free-surface flow problems: Part I. *Nonlinear Analysis: Theory, Methods & Applications*, 63(5-7):e1885–e1896, 2005.

[104] E. Miglio, S. Perotto, and F. Saleri. Model coupling techniques for free-surface flow problems: Part II. *Nonlinear Analysis: Theory, Methods & Applications*, 63(5-7):e1897–e1908, 2005.

[105] E. Miglio, A. Quarteroni, and F. Saleri. Finite element approximation of quasi-3D shallow water equations. *Computer Methods in Applied Mechanics and Engineering*, 174(3-4):355–369, 1999.

[106] H. K. Moffatt. Viscous and resistive eddies near a sharp corner. *Journal of Fluid Mechanics*, 18(1):1–18, 1964.

[107] M. Motamed. A multi-fidelity neural network surrogate sampling method for uncertainty quantification. *International Journal for Uncertainty Quantification*, 10(4), 2020.

[108] J. Nečas. Les méthodes directes en théorie des équations elliptiques. 1967.

[109] F. Negri. Reduced basis method for parametrized optimal control problems governed by PDEs. *Master thesis, Politecnico di Milano*, 2011.

[110] 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.

[111] 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.

[112] E. Novak and K. Ritter. High dimensional integration of smooth functions over cubes. *Numerische Mathematik*, 1(75):79–97, 1996.

[113] M. Ohlberger and S. Rave. Reduced basis methods: Success, limitations and future challenges. *arXiv preprint arXiv:1511.02021*, 2015.

[114] G. Pang, M. D'Elia, M. Parks, and G. E. Karniadakis. npinns: nonlocal physics-informed neural networks for a parametrized nonlocal universal laplacian operator. algorithms and applications. *Journal of Computational Physics*, 422:109760, 2020.

[115] W. Peng, W. Zhou, J. Zhang, and W. Yao. Accelerating physics-informed neural network training with prior dictionaries. *arXiv preprint arXiv:2004.08151*, 2020.

[116] F. Pichi, A. Quaini, and G. Rozza. A reduced order modeling technique to study bifurcating phenomena: Application to the Gross–Pitaevskii equation. *SIAM Journal on Scientific Computing*, 42(5):B1115–B1135, 2020.

[117] F. Pichi and G. Rozza. Reduced basis approaches for parametrized bifurcation problems held by non-linear Von Kármán equations. *Journal of Scientific Computing*, 81(1):112–135, 2019.

[118] F. Pichi, M. Strazzullo, F. Ballarin, and G. Rozza. Driving bifurcating parametrized nonlinear pdes by optimal control strategies: Application to Navier-Stokes equations and model reduction. Submitted, 2020, https://arxiv.org/abs/2010.13506.

[119] M. Pintore, F. Pichi, M. Hess, G. Rozza, and C. Canuto. Efficient computation of bifurcation diagrams with a deflated approach to reduced basis spectral element method. *Advances in Computational Mathematics*, 47(1):1–39, 2021.

[120] G. Pitton, A. Quaini, and G. Rozza. Computational reduction strategies for the detection of steady bifurcations in incompressible fluid-dynamics: applications to Coanda effect in cardiology. *Journal of Computational Physics*, 344:534–557, 2017.

[121] G. Pitton and G. Rozza. On the application of reduced basis methods to bifurcation problems in incompressible fluid dynamics. *Journal of Scientific Computing*, 73(1):157–177, 2017.

[122] M. Pošta and T. Roubíček. Optimal control of Navier–Stokes equations by Oseen approximation. *Computers & Mathematics With Applications*, 53(3):569–581, 2007.

[123] C. Prud'Homme, D. V. Rovas, K. Veroy, L. Machiels, Y. Maday, A. Patera, and G. Turinici. Reliable real-time solution of parametrized partial differential equations: Reduced-basis output bound methods. *Journal of Fluids Engineering*, 124(1):70–80, 2002.

[124] A. Quaini, R. Glowinski, and S. Canic. Symmetry breaking and preliminary results about a Hopf bifurcation for incompressible viscous flow in an expansion channel. *International Journal of Computational Fluid Dynamics*, 30(1):7–19, 2016.

[125] A. Quarteroni, A. Manzoni, and F. Negri. *Reduced basis methods for partial differential equations: an introduction*, volume 92. Springer, 2015.

[126] A. Quarteroni, G. Rozza, L. Dedè, and A. Quaini. Numerical approximation of a control problem for advection-diffusion processes. In *Ceragioli F., Dontchev A., Futura H., Marti K., Pandolfi L. (eds) System Modeling and Optimization. International Federation for Information Processing, CSMO Conference on System Modeling and Optimization*, pages vol 199, 261–273. Springer, Boston, 2005.

[127] A. Quarteroni, G. Rozza, and A. Quaini. Reduced basis methods for optimal control of advection-diffusion problems. In *Advances in Numerical Mathematics*, number CMCS-CONF-2006-003, pages 193–216. RAS and University of Houston, 2007.

[128] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*, volume 23. Springer Science & Business Media, Berlin and Heidelberg, 2008.

[129] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.

[130] M. Raissi, A. Yazdani, and G. E. Karniadakis. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science*, 367(6481):1026–1030, 2020.

[131] M. Ricchiuto, R. Abgrall, and H. Deconinck. Application of conservative residual distribution schemes to the solution of the shallow water equations on unstructured meshes. *Journal of Computational Physics*, 222(1):287–331, 2007.

[132] M. Ricchiuto and A. Bollermann. Stabilized residual distribution for shallow water simulations. *Journal of Computational Physics*, 228(4):1071–1115, 2009.

[133] G. Rozza, D. 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.

[134] G. Rozza, D. Huynh, and A. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics. *Archives of Computational Methods in Engineering*, 15(3):229–275, 2008.

[135] G. Rozza, A. Manzoni, and F. Negri. Reduction strategies for PDE-constrained oprimization problems in Haemodynamics. pages 1749–1768, ECCOMAS, September 2012, Congress Proceedings, Vienna, Austria.

[136] 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.

[137] F. Saleri and E. Miglio. Geometric multiscale approach by optimal control for shallow water equations. In *Applied And Industrial Mathematics In Italy II*, pages 537–548. World Scientific, 2007.

[138] 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.

[139] C. Schwab and R. Stevenson. Space-time adaptive wavelet methods for parabolic evolution problems. *Mathematics of Computation*, 78(267):1293–1318, 2009.

[140] R. Seydel. *Practical Bifurcation and Stability Analysis*. Interdisciplinary Applied Mathematics. Springer New York, 2009.

[141] R. Ştefănescu and I. M. Navon. POD/DEIM nonlinear model order reduction of an ADI implicit shallow water equations model. *Journal of Computational Physics*, 237:95–114, 2013.

[142] R. Ştefănescu, A. Sandu, and I. M. Navon. Comparison of POD reduced order strategies for the nonlinear 2D shallow water equations. *International Journal for Numerical Methods in Fluids*, 76(8):497–521, 2014.

[143] M. Stoll and A. Wathen. All-at-once solution of time-dependent PDE-constrained optimization problems. 2010.

[144] M. Stoll and A. Wathen. All-at-once solution of time-dependent Stokes control. *J. Comput. Phys.*, 232(1):498–515, Jan. 2013.

[145] M. Strazzullo, F. Ballarin, R. Mosetti, and G. Rozza. Model reduction for parametrized optimal control problems in environmental marine sciences and engineering. *SIAM Journal on Scientific Computing*, 40(4):B1055–B1079, 2018.

[146] M. Strazzullo, F. Ballarin, and G. Rozza. POD-Galerkin model order reduction for parametrized time dependent linear quadratic optimal control problems in saddle point formulation. *Journal of Scientific Computing*, 83(55), 2020.

[147] M. Strazzullo, F. Ballarin, and G. Rozza. POD-Galerkin model order reduction for parametrized nonlinear time dependent optimal flow control: an application to Shallow Water Equations. Submitted, 2020, https://arxiv.org/abs/2003.09695.

[148] M. Strazzullo, F. Ballarin, and G. Rozza. A Certified Reduced Basis method for linear parametrized parabolic optimal control problems in space-time formulation. Submitted, 2021, https://arxiv.org/abs/2103.00460.

[149] M. Strazzullo, M. Girfoglio, F. Ballarin, T. Iliescu, and G. Rozza. Consistency of the full and reduced order models for evolve-filter-relax regularization of convection-dominated, marginally-resolved flows. In preparation, 2021.

[150] M. Strazzullo, Z. Zainib, F. Ballarin, and G. Rozza. Reduced order methods for parametrized nonlinear and time dependent optimal flow control problems: towards applications in biomedical and environmental sciences. *Numerical Mathematics and Advanced Applications ENUMATH 2019*, 2021.

[151] T. Sullivan. *Introduction to Uncertainty Quantification*. Springer, 2015.

[152] T. Taddei and L. Zhang. Space-time registration-based model reduction of parameterized one-dimensional hyperbolic PDEs. *ESAIM: Mathematical Modelling and Numerical Analysis*, 55(1):99 – 130, 2021.

[153] S. Takase, K. Kashiyama, S. Tanaka, and T. E. Tezduyar. Space–time SUPG formulation of the shallow-water equations. *International Journal for Numerical Methods in Fluids*, 64(10-12):1379–1394, 2010.

[154] D. J. Tritton. *Physical fluid dynamics*. Springer Science & Business Media, 2012.

[155] 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.

[156] K. Urban and A. T. Patera. A new error bound for reduced basis approximation of parabolic partial differential equations. *Comptes Rendus Mathematique*, 350(3-4):203–207, 2012.

[157] R. Vázquez, E. Trélat, and J. M. Coron. Control for fast and stable laminar-to-high-Reynolds-numbers transfer in a 2D Navier-Stokes channel flow. *Discrete & Continuous Dynamical Systems - Series B, American Institute of Mathematical Sciences*, 10(4):925–956, 2008.

[158] L. Venturi, F. Ballarin, and G. Rozza. A weighted pod method for elliptic pdes with random inputs. *Journal of Scientific Computing*, 81(1):136–153, 2019.

[159] T. Von Kármán. Festigkeitsprobleme im Maschinenbau. *Encyclopädie der Mathematischen Wissenschaften*, 4, 1910.

[160] C. B. Vreugdenhil. *Numerical methods for shallow-water flow*, volume 13. Springer Science & Business Media, 2013.

[161] D. Wang and W. Liao. Modeling and control of magnetorheological fluid dampers using neural networks. *Smart materials and structures*, 14(1):111, 2004.

[162] G. W. Wasilkowski. Explicit cost bounds of algorithms for multivariate tensor product problems. *Journal of Complexity*, 1(11):1–56, 1995.

[163] J. Xu and L. Zikatanov. Some observations on Babuška and Brezzi theories. *Numerische Mathematik*, 94(1):195–202, 2003.

[164] L. Yang, X. Meng, and G. E. Karniadakis. B-pinns: Bayesian physics-informed neural networks for forward and inverse pde problems with noisy data. *Journal of Computational Physics*, 425:109913, 2021.

[165] M. Yano. A space-time Petrov–Galerkin certified reduced basis method: Application to the Boussinesq equations. *SIAM Journal on Scientific Computing*, 36(1):A232–A266, 2014.

[166] M. Yano, A. T. Patera, and K. Urban. A space-time hp-interpolation-based certified reduced basis method for Burgers' equation. *Mathematical Models and Methods in Applied Sciences*, 24(09):1903–1935, 2014.

[167] F. Yilmaz and B. Karasözen. An all-at-once approach for the optimal control of the unsteady Burgers equation. *Journal of Computational and Applied Mathematics*, 259:771 – 779, 2014. Recent Advances in Applied and Computational Mathematics: ICACM-IAM-METU.

[168] Z. Zainib, F. Ballarin, S. Fremes, P. Triverio, L. Jiménez-Juan, and G. Rozza. Reduced order methods for parametric optimal flow control in coronary bypass grafts, toward patient-specific data assimilation. *International Journal for Numerical Methods in Biomedical Engineering*, page e3367, 2020.