POD–Galerkin Model Order Reduction for Parametrized Time Dependent Linear Quadratic Optimal Control Problems in Saddle Point Formulation

note finali coverpage

(Article begins on next page)

15 May 2024

POD-Galerkin Model Order Reduction for Parametrized Time Dependent Linear Quadratic Optimal Control Problems in Saddle Point Formulation

Maria Strazzullo$^\sharp$, Francesco Ballarin$^\sharp$, and Gianluigi Rozza$^\sharp$

$^\sharp$*mathlab, Mathematics Area, International School for Advanced Studies (SISSA), Via Bonomea 265, I-34136 Trieste, Italy*

Abstract

In this work we recast parametrized time dependent optimal control problems governed by partial differential equations in a saddle point formulation and we propose reduced order methods as an effective strategy to solve them. Indeed, on one hand parametrized time dependent optimal control is a very powerful mathematical model which is able to describe several physical phenomena; on the other hand, it requires a huge computational effort. Reduced order methods are a suitable approach to have rapid and accurate simulations. We rely on POD-Galerkin reduction over the physical and geometrical parameters of the optimality system in a space-time formulation. Our theoretical results and our methodology are tested on two examples: a boundary time dependent optimal control for a Graetz flow and a distributed optimal control governed by time dependent Stokes equations. With these two experiments the convenience of the reduced order modelling is further extended to the field of time dependent optimal control.

Keywords: reduced order methods, proper orthogonal decomposition, time dependent parametrized optimal control problems, time dependent PDEs state equations, saddle point formulation.
AMS: 49J20, 76N25, 35Q35

1   Introduction

Parametrized optimal control problems (OCP($\boldsymbol{\mu}$)s) governed by parametrized partial differential equations (PDE($\boldsymbol{\mu}$)s) play a ubiquitous role in several applications, yet are very challenging to analyse theoretically and simulate numerically. In a parametrized setting, where a parameter $\boldsymbol{\mu} \in \mathscr{P} \subset \mathbb{R}^d$ could represent physical or geometrical features, OCP($\boldsymbol{\mu}$)s can be helpful in order to describe and simulate different configurations of several physical and natural phenomena. Indeed, optimal control framework is very versatile and it has been exploited in many contexts and fields: from shape optimization, see e.g. [18, 23, 36], to fluid dynamics, see e.g. [16, 37, 39, 15], from biomedical applications [6, 32, 54, 60] to environmental ones [41, 42, 53, 54]. Even if OCP($\boldsymbol{\mu}$)s are a very powerful tool widespread in many research fields, their complexity and computational demanding simulations still limit their applicability. Furthermore, the required computational efforts get even larger if the optimization is constrained to a time dependent PDE($\boldsymbol{\mu}$). Surely, time optimization makes the mathematical model more complete and it arises in many applications, see e.g. [25, 27, 33, 50, 51, 52].

The computational effort required for OCP($\boldsymbol{\mu}$)s simulations becomes unbearable and not viable when the time dependent optimal control depends on a variety of physical and/or geometrical parameters: indeed, in this context, many configurations are studied for several values of $\boldsymbol{\mu} \in \mathscr{P}$, increasing the necessary time in order to understand the phenomena behaviour with respect to different parameters. A rapid and suitable approach to manage this drawback is to rely on reduced order methods (ROMs), which allow us to solve the parametrized optimality system in a low dimensional framework, reducing the computational costs, see for example [24, 40, 44, 45]. Literature is quite complete with respect to the applications of ROM techniques for parametrized steady OCP($\boldsymbol{\mu}$)s, the interested reader may refer to the following far-from-exhaustive list [4, 5, 17, 20, 28, 29, 30, 31, 37, 38, 42], where reduced optimal control is treated and analyzed for several state equations and managed with different approaches and methodologies.

We seek therefore to extend the consolidated knowledge about ROM for steady OCP($\boldsymbol{\mu}$)s to time dependent OCP($\boldsymbol{\mu}$)s.

Our purpose is to adapt the formulation presented in [37, 38] for steady OCP($\boldsymbol{\mu}$)s to quadratic optimization models constrained to linear time dependent PDE($\boldsymbol{\mu}$)s. We propose two simple test cases in order to validate our theoretical framework:

- a time dependent boundary optimal control problem for Graetz flow;

- a time dependent distributed optimal control problem for Stokes equations.

Both the proposed examples have geometrical and physical parametrization.
The main merit of this work is to recast linear quadratic time dependent OCP($\boldsymbol{\mu}$)s in the very well known and general framework of saddle point problems. To the best of our knowledge, another element of novelty is the numerical simulation and subsequent reduction of a time dependent OCP($\boldsymbol{\mu}$) governed by Stokes equations. The proposed techniques have some limitations in practical applications: indeed, we rely on direct solvers for very complex linear systems. This limits our capability to increase the resolutions both in time and space for simulations. This issue could be overcome using more computational resources and through preconditioned Krylov solvers and multigrid approaches: we refer the interested reader to [9, 48, 51, 52]. We highlight that the proposed iterative approaches may only lighten the computational effort needed for the linear system at hand during the so-called offline phase of the ROM. Due to this, we are aware that the numerical simulations of time dependent OCP($\boldsymbol{\mu}$)s in the all-at-once saddle point framework could still be unfeasible for actual real-time applications because of the extremely large CPU time required by the offline stage. Nonetheless, this paper aims at showing how ROMs can recover the behaviour of full order simulations with a substantial reduction of computational time. Indeed, we validate our methodology thanks to several experiments for academic test cases. This is only a first step towards the study of new approaches for the application of ROMs to more general and complicated time dependent OCP($\boldsymbol{\mu}$)s in saddle point formulation.

This work is outlined as follows. In section 2, we rely on theoretical formulation for time dependent optimal control presented in [26] and [55], then, we prove the well-posedness of linear quadratic parabolic time dependent OCP($\boldsymbol{\mu}$)s in a saddle point framework verifying the standard hypotheses of Brezzi's theorem [11, 12]. A brief introduction of the full order Finite Element (FE) approximation precedes the discretization of the optimality system, treated exploiting an all-at-once approach proposed in [51, 52, 59]. Section 3 introduces the reduced order approximation following [24, 29] and how to apply Proper Orthogonal Decomposition (POD) sampling algorithm for OCP($\boldsymbol{\mu}$)s in saddle point framework with a brief mention of aggregated reduced space strategy used in [17, 37, 38] and affine decomposition, see e.g. [24]. In section 4, numerical

2

results for geometrical and physical parametrization of a boundary optimal control problem for Graetz flow are shown: it is a time dependent version of an OCP($\boldsymbol{\mu}$) presented in [38]. In section 5, we analyze the saddle point structure of time dependent OCP($\boldsymbol{\mu}$)s governed by Stokes equations. We prove the well-posedness of the mixed problem and we briefly show the discretization techniques used extending the strategies already described in sections 2 and 3. In section 6 we will present a distributed optimal control problem governed by time dependent Stokes equations having consistent results with respect to those in [25, 52]. Conclusions follow in section 7.

2 Problem Formulation and Full Order Discretization for Linear Quadratic Parabolic Time Dependent OCP($\boldsymbol{\mu}$)s

In this section parabolic time dependent OCP($\boldsymbol{\mu}$)s are presented in saddle point formulation. To the best of our knowledge, the saddle point theoretical analytic framework in the context of time dependent OCP($\boldsymbol{\mu}$)s is a novelty element, even if it is quite a standard approach for stationary linear state equations [29, 37, 38, 46]. First of all, the well-posedness of the saddle point structure is proved in a space-time formulation. Then, we will introduce the full order discretized problem: the high fidelity approximation is presented following the all-at-once structure exploited in [25, 48, 51, 52].

2.1 Saddle Point Structure of Linear Quadratic Parabolic Time Dependent OCP($\boldsymbol{\mu}$)s: Theoretical Analysis

This section aims at recasting OCP($\boldsymbol{\mu}$)s governed by linear time dependent PDE($\boldsymbol{\mu}$)s in a saddle point formulation. In the systems at hand the parameter $\boldsymbol{\mu} \in \mathscr{P} \subset \mathbb{R}^p$ could be physical or geometrical. In the following analysis we assume a physical parametrized PDE($\boldsymbol{\mu}$), but the reached conclusions do not change for the case of geometrical parametrization.

Let us consider $\Omega \subset \mathbb{R}^n$ an open and bounded regular domain and the time interval $(0, T)$. Let us indicate with Γ_D and Γ_N the portions of the boundary $\partial\Omega$ where Dirichlet and Neumann boundary conditions are applied, respectively. Let Y and H be two separable Hilbert spaces which verify $Y \hookrightarrow H \hookrightarrow Y^*$ and let us consider other two Hilbert spaces U and $Z \supseteq Y$. Let us define the Hilbert spaces $\mathcal{Y} = L^2(0, T; Y)$, $\mathcal{Y}^* = L^2(0, T; Y^*)$, $\mathcal{U} = L^2(0, T; U)$ and $\mathcal{Z} := L^2(0, T; Z) \supseteq \mathcal{Y}$. Finally, we define the space $\mathcal{Y}_t := \left\{ y \in \mathcal{Y} \text{ s.t. } \frac{\partial y}{\partial t} \in \mathcal{Y}^* \right\}$. The problem to be solved is the following: for a given $\boldsymbol{\mu} \in \mathscr{P}$ find the pair $(y, u) \in \mathcal{Y}_t \times \mathcal{U}$ which solves

$$\min_{(y,u) \in \mathcal{Y}_t \times \mathcal{U}} \frac{1}{2} \int_0^T \int_\Omega (y - y_d(\boldsymbol{\mu}))^2 \, d\Omega dt + \frac{\alpha}{2} \int_0^T \int_\Omega u^2 \, d\Omega dt \tag{1}$$

under a constraint of the form

$$\begin{cases} \dfrac{\partial y}{\partial t} + \mathcal{D}_a(\boldsymbol{\mu}) y = \mathcal{D}_c(\boldsymbol{\mu}) u + f & \text{in } \Omega \times (0, T), \\ \dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N \times (0, T), \\ y = g & \text{on } \Gamma_D \times (0, T), \\ y(0) = y_0 & \text{in } \Omega, \end{cases} \tag{2}$$

where $\alpha > 0$ is a fixed constant, $\mathcal{D}_a(\boldsymbol{\mu})$ and $\mathcal{D}_c(\boldsymbol{\mu})$ are general parametrized differential operators associated to state and control, respectively, and $y_d(\boldsymbol{\mu})$ is an observation in the space $\mathcal{Z}$. Furthermore, for the sake of simplicity we assume that $\|y_0\|_H^2 = 0$ and $\mathcal{Z} = \mathcal{Y}_t$. Problems of this kind have been covered extensively in classical references as [34, 55], for example, where the topic of

optimal control is treated in its entirety. Basically, we are adding a parametrization framework to this very versatile mathematical model, in order to study different physical and/or geometrical configurations. Indeed, we underline that we are actually looking for solutions which are parameter dependent, but for the sake of notation we write $y := y(\boldsymbol{\mu})$ and $u := u(\boldsymbol{\mu})$: depending on the case, we will omit the parameter dependence from now on.

Moreover, we suppose that for every $\boldsymbol{\mu} \in \mathscr{P}$ and every control variable u there exists an unique solution $y := y(u)$. The time dependent state equation (2) can be expressed in the following weak formulation:

$$\begin{cases} \int_0^T \left\langle \dfrac{\partial y}{\partial t}, q \right\rangle_{Y^*, Y} dt + \int_0^T a(y, q; \boldsymbol{\mu}) \, dt = \int_0^T c(u, q; \boldsymbol{\mu}) \, dt + \int_0^T \langle G(\boldsymbol{\mu}), q \rangle_{Y^*, Y} \, dt & \forall q \in \mathcal{Y}_t, \\ y(0) = y_0 & \text{in } \Omega, \end{cases} \quad (3)$$

where $a : \mathcal{Y}_t \times \mathcal{Y}_t \to \mathbb{R}$ and $c : \mathcal{U} \times \mathcal{Y}_t \to \mathbb{R}$ are the bilinear forms associated to $\mathcal{D}_a(\boldsymbol{\mu})$ and $\mathcal{D}_c(\boldsymbol{\mu})$, respectively, while $G \in \mathcal{Y}^*$ collects forcing and boundary terms of the state equation. In our applications it will always be the case $c(u, q; \boldsymbol{\mu}) = \int_\Omega uq \, d\Omega$. We further remark that we consider $q \in \mathcal{Y}_t$ rather than $q \in \mathcal{Y}$ as it will be convenient to restrict q to $\mathcal{Y}_t$ for a proper definition of the adjoint variable. Let us define the following bilinear forms

$$m : \mathcal{Y}_t \times \mathcal{Y}_t \to \mathbb{R} \qquad m(y, z) = \int_0^T \int_\Omega yz \, d\Omega dt, \qquad (4)$$

$$n : \mathcal{U} \times \mathcal{U} \to \mathbb{R} \qquad n(u, v) = \int_0^T \int_\Omega uv \, d\Omega dt, \qquad (5)$$

which represent the objective for the state variable and for the control variable in the whole time interval, respectively. Thus, let us define the functional

$$J(y, u; \boldsymbol{\mu}) = \frac{1}{2} m(y - y_d(\boldsymbol{\mu}), y - y_d(\boldsymbol{\mu})) + \frac{\alpha}{2} n(u, u). \qquad (6)$$

The OCP($\boldsymbol{\mu}$) reads as follows: given $\boldsymbol{\mu} \in \mathscr{P}$, find the solution of

$$\min_{(y, u) \in \mathcal{Y}_t \times \mathcal{U}} J(y, u; \boldsymbol{\mu}) \qquad \text{such that (3) is satisfied.} \qquad (7)$$

In order to set the problem in a mixed formulation, we need to define the state-control product space $\mathcal{X} = \mathcal{Y}_t \times \mathcal{U}$. Given $x = (y, u)$ and $w = (z, v)$ elements of $\mathcal{X}$, the scalar product of $\mathcal{X}$ is defined by $(x, w)_\mathcal{X} = (y, z)_{\mathcal{Y}_t} + (u, v)_\mathcal{U}$, that induces the norm $\|\cdot\|_\mathcal{X}$, since $(\cdot, \cdot)_{\mathcal{Y}_t}$ and $(\cdot, \cdot)_\mathcal{U}$ define the following norms, respectively:

$$\|y\|_{\mathcal{Y}_t}^2 = \|y\|_\mathcal{Y}^2 + \left\| \frac{\partial y}{\partial t} \right\|_{\mathcal{Y}^*}^2 = \int_0^T \|y\|_Y^2 \, dt + \int_0^T \left\| \frac{\partial y}{\partial t} \right\|_{Y^*}^2 \, dt, \qquad \|u\|_\mathcal{U}^2 = \int_0^T \|u\|_U^2 \, dt.$$

Moreover, in the following, we need the forms:

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

$$\mathcal{B} : \mathcal{X} \times \mathcal{Y}_t \to \mathbb{R}, \qquad \mathcal{B}(w, q; \boldsymbol{\mu}) = \int_0^T \left\langle \frac{\partial z}{\partial t}, q \right\rangle dt + \int_0^T a(z, q; \boldsymbol{\mu}) \, dt - \int_0^T c(v, q; \boldsymbol{\mu}) \, dt, \qquad \forall w \in \mathcal{X}, \forall q \in \mathcal{Y}_t,$$

$$F(\boldsymbol{\mu}) \in \mathcal{X}^*, \qquad \int_0^T \langle F(\boldsymbol{\mu}), w \rangle \, dt = m(y_d(\boldsymbol{\mu}), z). \qquad\qquad\qquad \forall w \in \mathcal{X}.$$

We underline that for the analysis and for sake of notation, the problem we propose has a distributed control, but it can be extended also for boundary control if one defines:

$$\|u\|_{\mathcal{U}}^2 = \int_0^T \|u\|_{U(\Gamma_c)}^2 \, dt,$$

where $\Gamma_c \subseteq \partial\Omega$ is the boundary portion where the control is applied. An application of boundary control will be presented in Section 4.

As we did for the state and the control variables, for the sake of notation we define $p := p(\boldsymbol{\mu})$. In order to build the optimality system, first of all we construct the following Lagrangian functional

$$\mathscr{L}(y, u, p; \boldsymbol{\mu}) = \mathcal{J}((y, u); \boldsymbol{\mu}) + \mathcal{B}((y, u), p; \boldsymbol{\mu}) + \int_0^T \langle G(\boldsymbol{\mu}), p \rangle \, dt, \tag{8}$$

where

$$\mathcal{J}((y, u), \boldsymbol{\mu}) = \frac{1}{2}\mathcal{A}(x, x; \boldsymbol{\mu}) - \int_0^T \langle F(\boldsymbol{\mu}), x \rangle \, dt, \tag{9}$$

recalling that $x = (y, u)$. Then, we perform a differentiation by the adjoint variable p, the state variable y and the control u. The minimization of (8) is equivalent to find the solution of the following system: given $\boldsymbol{\mu} \in \mathscr{P}$, find $(x, p) \in \mathcal{X} \times \mathcal{Y}_t$ such that

$$\begin{cases} D_y\mathscr{L}(y, u, p; \boldsymbol{\mu})[z] = 0 & \forall z \in \mathcal{Y}_t, \\ D_u\mathscr{L}(y, u, p; \boldsymbol{\mu})[v] = 0 & \forall v \in \mathcal{U}, \\ D_p\mathscr{L}(y, u, p; \boldsymbol{\mu})[q] = 0 & \forall q \in \mathcal{Y}_t. \end{cases} \tag{10}$$

The framework introduced is totally general and it also holds for nonlinear time dependent problems.

We now focus on the case of **parabolic linear governing equations**. In this case, the system of equations related to (10) could be written in the following form:

$$\begin{cases} \mathcal{A}(x, w) + \mathcal{B}(w, p; \boldsymbol{\mu}) = \int_0^T \langle F(\boldsymbol{\mu}), w \rangle \, dt & \forall w \in \mathcal{X}, \\ \mathcal{B}(x, q; \boldsymbol{\mu}) = \int_0^T \langle G(\boldsymbol{\mu}), q \rangle \, dt & \forall q \in \mathcal{Y}_t. \end{cases} \tag{11}$$

As one can see, the classical structure of saddle point formulation that characterizes steady linear quadratic OCP($\boldsymbol{\mu}$)s is preserved also in the linear time dependent case. We now want to provide the well-posedness of the problem (11) through the fulfillment of [10, Proposotion 1.7], a Proposotion based on Brezzi's theorem [11, 12]:

Proposition 1. *Assume that the Brezzi's theorem is verified, i.e.*

1. $\mathcal{A}(\cdot, \cdot)$ *is continuous and weakly coercive on the kernel of* $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$, *that we indicate with* $\mathcal{X}_0 \subset \mathcal{X}$, *i.e. for every* $x, w \in \mathcal{X}_0$ *it holds:*

$$\inf_{w \in \mathcal{X}_0\backslash\{0\}} \sup_{x \in \mathcal{X}_0\backslash\{0\}} \frac{\mathcal{A}(x, w)}{\|x\|_{\mathcal{X}_0}\|w\|_{\mathcal{X}_0}} > 0 \quad and \quad \inf_{x \in \mathcal{X}_0\backslash\{0\}} \sup_{w \in \mathcal{X}_0\backslash\{0\}} \frac{\mathcal{A}(x, w)}{\|x\|_{\mathcal{X}_0}\|w\|_{\mathcal{X}_0}} > 0. \tag{12}$$

2. $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ *is continuous and satisfies the inf-sup condition, i.e. for every* $x \in \mathcal{X}$ *and* $q \in \mathcal{Y}_t$ *the following inequality is verified:*

$$\inf_{q \in \mathcal{Y}_t\backslash\{0\}} \sup_{x \in \mathcal{X}\backslash\{0\}} \frac{\mathcal{B}(x, q; \boldsymbol{\mu})}{\|x\|_{\mathcal{X}}\|q\|_{\mathcal{Y}_t}} = \beta(\boldsymbol{\mu}) > 0. \tag{13}$$

5

Let us assume further that the bilinear form $\mathcal{A}(\cdot, \cdot)$ is symmetric, non-negative, and coercive on $\mathcal{X}_0$, then the minimization of the functional (9) *constrained to equation* (2) *and the resolution of the saddle point problem* (11) *are equivalent.*

In order to prove that the hypotheses over $\mathcal{A}(\cdot, \cdot)$ and $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ hold for OCP($\boldsymbol{\mu}$)s governed by linear time dependent state equations, first of all, we require that:

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

(i) $|c(u, q; \boldsymbol{\mu})| \leq c_c(\boldsymbol{\mu}) \|u\|_U \|q\|_Y \qquad \forall u \in \mathcal{U}$ *and* $\forall q \in \mathcal{Y}_t$;

(ii) $|a(y, q; \boldsymbol{\mu})| \leq c_a(\boldsymbol{\mu}) \|y\|_Y \|q\|_Y \qquad \forall y, q \in \mathcal{Y}_t$;

(iii) $a(y, y; \boldsymbol{\mu}) \geq M_a(\boldsymbol{\mu}) \|y\|_Y^2 \qquad \forall y \in \mathcal{Y}_t$;

(iv) $n(\cdot, \cdot)$ *is symmetric, continuous and such that* $n(u, u) \geq \gamma_n \|u\|_\mathcal{U}^2$;

(v) $m(\cdot, \cdot)$ *is symmetric, continuous and positive definite.*

Furthermore, we will exploit the following inequalities in order to assert the inequalities (12) and (13):

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

$$\left\| \frac{\partial y}{\partial t} \right\|_{\mathcal{Y}^*} \leq \|y\|_{\mathcal{Y}_t}, \quad \|y\|_\mathcal{Y} \leq \|y\|_{\mathcal{Y}_t}, \quad \|y\|_{\mathcal{Y}_t} \leq \|x\|_\mathcal{X} \quad \text{and} \quad \|u\|_\mathcal{U} \leq \|x\|_\mathcal{X}; \qquad (14)$$

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:

$$\|y\|_{\mathcal{Y}_t} \leq k(\boldsymbol{\mu})(\|f\|_{L^2(0,T;Y)} + \|y_0\|_Y). \qquad (15)$$

Two more ingredients used to guarantee the *inf-sup condition* (13) are the following two lemmas.

Lemma 1. *Let q be a function in $\mathcal{Y}_t$, then the following inequality holds:*

$$\frac{\|q\|_\mathcal{Y}^2}{\|q\|_{\mathcal{Y}_t}^2} \geq \frac{1}{6}.$$

Proof. We divide the proof in two cases: first of all, let us assume $\left\| \frac{\partial q}{\partial t} \right\|_{\mathcal{Y}^*} \leq \|q\|_\mathcal{Y}$ then $\|q\|_{\mathcal{Y}_t}^2 \leq 2\|q\|_\mathcal{Y}^2$. This leads to the following inequality:

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

We now want to prove a similar inequality for $q \in \mathcal{Y}_t$ such that $\left\| \frac{\partial q}{\partial t} \right\|_{\mathcal{Y}^*} > \|q\|_\mathcal{Y}$. This assumption allows us to assert that:

$$2\|q\|_\mathcal{Y}^2 < \|q\|_{\mathcal{Y}_t}^2 \qquad \Rightarrow \qquad \frac{1}{2\|q\|_\mathcal{Y}^2} > \frac{1}{\|q\|_{\mathcal{Y}_t}^2}, \qquad (16)$$

and, by definition of $\|q\|_{\mathcal{Y}_t}$,

$$\|q\|_{\mathcal{Y}_t}^2 + 2\left\| \frac{\partial q}{\partial t} \right\|_{\mathcal{Y}^*}^2 \geq 3\left\| \frac{\partial q}{\partial t} \right\|_{\mathcal{Y}^*}^2 \Rightarrow -\frac{1}{\|q\|_{\mathcal{Y}_t}^2 + 2\left\| \frac{\partial q}{\partial t} \right\|_{\mathcal{Y}^*}^2} \geq -\frac{1}{3\left\| \frac{\partial q}{\partial t} \right\|_{\mathcal{Y}^*}^2}. \qquad (17)$$

Then, we can prove that

$$
\begin{aligned}
\frac{\|q\|_{\mathcal{Y}}^2}{\|q\|_{\mathcal{Y}_t}^2} &= \frac{2\|q\|_{\mathcal{Y}}^2}{2\|q\|_{\mathcal{Y}}^2 + 2\left\|\frac{\partial q}{\partial t}\right\|_{\mathcal{Y}^*}^2} \underbrace{\geq}_{\text{for (16)}} \frac{2\|q\|_{\mathcal{Y}}^2}{\|q\|_{\mathcal{Y}_t}^2 + 2\left\|\frac{\partial q}{\partial t}\right\|_{\mathcal{Y}^*}^2} \\[2mm]
&= \frac{2\|q\|_{\mathcal{Y}}^2 + \left\|\frac{\partial q}{\partial t}\right\|_{\mathcal{Y}^*}^2 - \left\|\frac{\partial q}{\partial t}\right\|_{\mathcal{Y}^*}^2}{\|q\|_{\mathcal{Y}_t}^2 + 2\left\|\frac{\partial q}{\partial t}\right\|_{\mathcal{Y}^*}^2} \\[2mm]
&\geq \frac{\min\{1,2\}\|q\|_{\mathcal{Y}_t}^2}{\max\{1,2\}\|q\|_{\mathcal{Y}_t}^2} - \frac{\left\|\frac{\partial q}{\partial t}\right\|_{\mathcal{Y}^*}^2}{\|q\|_{\mathcal{Y}_t}^2 + 2\left\|\frac{\partial q}{\partial t}\right\|_{\mathcal{Y}^*}^2} \\[2mm]
&\underbrace{\geq}_{\text{for (17)}} \frac{1}{2} - \frac{\left\|\frac{\partial q}{\partial t}\right\|_{\mathcal{Y}^*}^2}{3\left\|\frac{\partial q}{\partial t}\right\|_{\mathcal{Y}^*}^2} = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}.
\end{aligned}
$$

$\square$

Let us now prove a second lemma needed to show the well-posedness of problem (11).

Lemma 2. *Given a function $v \in \mathcal{Y}_t$, there exists $\bar{y} \in \mathcal{Y}_t$ which verifies:*

$$
\int_0^T \left\langle \frac{\partial \bar{y}}{\partial t}, q \right\rangle dt + \int_0^T a(\bar{y}, q; \boldsymbol{\mu}) \, dt = \int_0^T a(v, q; \boldsymbol{\mu}) \, dt \qquad \forall q \in \mathcal{Y}_t, \tag{18}
$$

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

$$
\|\bar{y}\|_{\mathcal{Y}_t} \leq \bar{k}(\boldsymbol{\mu})\|v\|_{\mathcal{Y}}. \tag{19}
$$

Proof. The existence of the solution $\bar{y}$ is actually proposed in the proof of property (A.3) of the Theorem 5.1 in [49], where the existence of $\bar{y} \in \mathcal{Y}_t$ is guaranteed for a given $v \in \mathcal{Y} \supset \mathcal{Y}_t$ and for every initial condition.

Now, for $t \in (0, T)$ let us consider the linear operator $\mathcal{D}_a(\boldsymbol{\mu}) : Y \to Y^*$ defined by $\langle \mathcal{D}_a v, q \rangle_{H^*, H}$. We define $C_{\mathcal{D}_a}(\boldsymbol{\mu}) = \|\mathcal{D}_a\|_{H^*}$ which is finite since $\mathcal{D}_a$ is a continuous operator. Furthermore, since $\bar{y}$ verifies (18), from the standard inequality (15) we can derive the inequality

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

Since $Y \hookrightarrow H$, it holds $\|q\|_H \leq \bar{C}\|q\|_Y$, then, setting $\bar{k}(\boldsymbol{\mu}) = k(\boldsymbol{\mu})C_{\mathcal{D}_a}\bar{C}$, we can prove the thesis and

$$
\|\bar{y}\|_{\mathcal{Y}_t} \leq \bar{k}(\boldsymbol{\mu})\|v\|_{\mathcal{Y}}. \tag{20}
$$

$\square$

We now have all the ingredients which will help us proving the well-posedness of the saddle point system (11). Indeed, the following theorem provides conditions (12) and (13) for linear parabolic time dependent OCP($\boldsymbol{\mu}$)s, i.e. the existence and the uniqueness of an optimal solution for the minimization problem defined by (11).

Theorem 1. *The saddle point problem* (11) *satisfies the hypotheses of Proposition 1 under the Assumptions 1, i.e. it has a unique optimal solution.*

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

$$|\mathcal{A}(x, w)| \leq \|y\|_{\mathcal{Y}_t} \|z\|_{\mathcal{Y}_t} + \alpha \|u\|_{\mathcal{U}} \|v\|_{\mathcal{U}}$$
$$\leq \max\{1, \alpha\} \|x\|_{\mathcal{X}} \|w\|_{\mathcal{X}}.$$

Indeed, the above inequality follows from continuity of the bilinear forms $m(\cdot, \cdot)$ and $n(\cdot, \cdot)$, which can be shown as follows:

$$\left| \int_0^T \int_\Omega yz \, d\Omega dt \right| \underbrace{\leq}_{\text{Cauchy}} \left(\int_0^T \|y\|_Y \|z\|_Y \, dt \right) \underbrace{\leq}_{\text{Holder}} \sqrt{\int_0^T \|y\|_Y^2 \, dt} \sqrt{\int_0^T \|z\|_Y^2 \, dt} = \|y\|_{\mathcal{Y}} \|z\|_{\mathcal{Y}}$$
$$\underbrace{\leq}_{\text{II}} \|y\|_{\mathcal{Y}_t} \|z\|_{\mathcal{Y}_t}.$$

The same argument can be used for $n(\cdot, \cdot)$, since

$$\left| \int_0^T \int_\Omega uv \, d\Omega dt \right| \underbrace{\leq}_{\text{Cauchy}} \left(\int_0^T \|u\|_U \|v\|_U \, dt \right) \underbrace{\leq}_{\text{Holder}} \sqrt{\int_0^T \|u\|_U^2 \, dt} \sqrt{\int_0^T \|v\|_U^2 \, dt} = \|u\|_{\mathcal{U}} \|v\|_{\mathcal{U}}.$$

Thanks to the hypothesis (iv) and (v), $\mathcal{A}(\cdot, \cdot)$ is symmetric, positive definite and continuous. We can now prove the coercivity of $\mathcal{A}$ on $\mathcal{X}_0$. If $x \in \mathcal{X}_0$, then it holds $\int_0^T \left\langle \frac{\partial y}{\partial t}, q \right\rangle dt + a(y, q; \boldsymbol{\mu}) = c(u, q; \boldsymbol{\mu})$ and then $\|y\|_{\mathcal{Y}_t} \leq k(\boldsymbol{\mu})(\|u\|_{\mathcal{U}} + \|y_0\|_H) = k(\boldsymbol{\mu}) \|u\|_{\mathcal{U}}$ by assumption over the initial condition of y. Then, it holds:

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

Let us prove on the continuity of $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$. Exploiting the continuity of $a(\cdot, \cdot; \boldsymbol{\mu})$ and $c(\cdot, \cdot; \boldsymbol{\mu})$, it holds:

$$|\mathcal{B}(x, q; \boldsymbol{\mu})| \leq \int_0^T \left| \left\langle \frac{\partial y}{\partial t}, q \right\rangle \right| dt + \int_0^T c_a(\boldsymbol{\mu}) \|y\|_Y \|q\|_Y \, dt + \int_0^T c_c(\boldsymbol{\mu}) \|u\|_U \|q\|_Y \, dt$$
$$\underbrace{\leq}_{\text{Cauchy + Hölder}} \left\| \frac{\partial y}{\partial t} \right\|_{\mathcal{Y}^*} \|q\|_{\mathcal{Y}_t} + c_a(\boldsymbol{\mu}) \|y\|_{\mathcal{Y}_t} \|q\|_{\mathcal{Y}_t} + c_c(\boldsymbol{\mu}) \|u\|_{\mathcal{U}} \|q\|_{\mathcal{Y}_t}$$
$$\underbrace{\leq}_{\text{I}} \max\{1, c_a(\boldsymbol{\mu}), c_c(\boldsymbol{\mu})\} \|x\|_{\mathcal{X}} \|q\|_{\mathcal{Y}_t},$$

where $c_a(\boldsymbol{\mu})$ and $c_c(\boldsymbol{\mu})$ are the continuity constants of $a(\cdot, \cdot; \boldsymbol{\mu})$ and $c(\cdot, \cdot; \boldsymbol{\mu})$. respectively. Now we focus on the fulfillment of the *inf-sup condition* for the bilinear form $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$. First of all, let us consider $q \in \mathcal{Y}_t$ and $\bar{y} \in \mathcal{Y}_t$ the solution of the problem (18) presented in Lemma 2

with $v \equiv q$. Then,

$$
\sup_{0 \neq x \in \mathcal{X}} \frac{\mathcal{B}(x, q; \boldsymbol{\mu})}{\|x\|_{\mathcal{X}} \|q\|_{\mathcal{Y}_t}} = \sup_{0 \neq (y, u)} \frac{\int_0^T \left\langle \frac{\partial y}{\partial t}, q \right\rangle dt + \int_0^T a(y, q; \boldsymbol{\mu}) \, dt - \int_0^T c(u, q; \boldsymbol{\mu}) \, dt}{\|x\|_{\mathcal{X}} \|q\|_{\mathcal{Y}_t}}
$$

$$
\underbrace{\geq}_{x = (\bar{y}, 0)} \frac{\int_0^T \left\langle \frac{\partial \bar{y}}{\partial t}, q \right\rangle dt + \int_0^T a(\bar{y}, q; \boldsymbol{\mu}) \, dt}{\|\bar{y}\|_{\mathcal{Y}_t} \|q\|_{\mathcal{Y}_t}} \geq \frac{\int_0^T a(q, q; \boldsymbol{\mu}) \, dt}{\|\bar{y}\|_{\mathcal{Y}_t} \|q\|_{\mathcal{Y}_t}}
$$

$$
\underbrace{\geq}_{\text{Lemma 2}} \frac{M_a(\boldsymbol{\mu}) \|q\|_{\mathcal{Y}}^2}{\bar{k}(\boldsymbol{\mu}) \|q\|_{\mathcal{Y}_t}^2} \underbrace{\geq}_{\text{Lemma 1}} \frac{M_a(\boldsymbol{\mu})}{6\bar{k}(\boldsymbol{\mu})} > 0.
$$

Since we have proved the inequality for all $q \in \mathcal{Y}_t$, it holds:

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

$\square$ ∎

The theorem just presented guarantees the existence and uniqueness of an optimal solution for time dependent OCP($\boldsymbol{\mu}$)s.

We now propose a brief overview on the FE approach, that we exploit as full order approximation in order to build the reduced order model.

For the spatial discretization, first of all we define the triangulation $\mathcal{T}^{\mathcal{N}}$ of Ω. We can now define discretized spatial spaces as $Y^{\mathcal{N}} = Y \cap \mathscr{X}_r^{\mathcal{N}}$ and $U^{\mathcal{N}} = U \cap \mathscr{X}_r^{\mathcal{N}}$, where

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

The space $\mathbb{P}^r$ is the space of all the polynomials of degree at most equal to r and K is a triangular element of $\mathcal{T}^{\mathcal{N}}$. Then, the function spaces considered are: $\mathcal{Y}^{\mathcal{N}} = L^2(0, T; Y^{\mathcal{N}})$, $\mathcal{Y}^{*\mathcal{N}} = L^2(0, T; Y^{*\mathcal{N}})$ and $\mathcal{U}^{\mathcal{N}} = L^2(0, T; U^{\mathcal{N}})$. The full order state variable will be considered in $\mathcal{Y}_t^{\mathcal{N}} := \left\{ y \in \mathcal{Y}^{\mathcal{N}} \ s.t. \ \frac{\partial y}{\partial t} \in \mathcal{Y}^{*\mathcal{N}} \right\}$. As we did in section 2.1, for the mixed formulation, we exploit the product space $\mathcal{X}^{\mathcal{N}} = \mathcal{Y}_t^{\mathcal{N}} \times \mathcal{U}^{\mathcal{N}} \subset \mathcal{X}$. The Galerkin FE discretization of the saddle point problem (11) reads as follows: given $\boldsymbol{\mu} \in \mathscr{P}$, find $(x^{\mathcal{N}}, p^{\mathcal{N}}) \in \mathcal{X}^{\mathcal{N}} \times \mathcal{Y}^{\mathcal{N}}$ such that

$$
\begin{cases}
\mathcal{A}(x^{\mathcal{N}}, w^{\mathcal{N}}) + \mathcal{B}(w^{\mathcal{N}}, p^{\mathcal{N}}; \boldsymbol{\mu}) = \int_0^T \langle F(\boldsymbol{\mu}), w^{\mathcal{N}} \rangle \, dt & \forall v^{\mathcal{N}} \in \mathcal{X}^{\mathcal{N}}, \\
\mathcal{B}(x^{\mathcal{N}}, q^{\mathcal{N}}; \boldsymbol{\mu}) = \int_0^T \langle G(\boldsymbol{\mu}), q^{\mathcal{N}} \rangle \, dt & \forall q^{\mathcal{N}} \in \mathcal{Y}_t^{\mathcal{N}},
\end{cases}
\tag{21}
$$

where $x^{\mathcal{N}} := x^{\mathcal{N}}(\boldsymbol{\mu})$ and $p^{\mathcal{N}} := p^{\mathcal{N}}(\boldsymbol{\mu})$ are parameter dependent solutions. Following the same strategies used in Section 2.1, it can be shown that Brezzi's Theorem [11, 12] holds at the discrete level. Moreover one can prove the coercivity of the form $\mathcal{A}(\cdot, \cdot)$: then, the minimization problem (21) is well-posed [10].

In this section we introduced the general formulation for linear time dependent OCP($\boldsymbol{\mu}$)s and we recast it in a saddle point system of the form (11). Then, we provided a proof for Theorem 1 which led to a unique optimal solution, after having introduced all the assumptions and lemmas needed in order to have Proposition 1 verified. Furthermore, we introduced FE element approximation as our full order discretization technique.

In the next section, the full order problem will be shown in space-time formulation following the approaches already presented for parabolic PDEs($\boldsymbol{\mu}$) in [21, 56, 57, 58].

2.2 Algebraic System and All-at-Once Approach

In this section we introduce a discretized version of time dependent OCP($\boldsymbol{\mu}$)s. Practically, we follow the all-at-once space-time discretization already proposed in [25, 51, 52].

We now want to consider the full order discretization of the OCP($\boldsymbol{\mu}$) introduced in (7). Let us suppose to have discretized the spaces considered in Section 2.1 with FE technique in order to solve the full order optimality system (21). We aim at showing how the saddle point structure is reflected also in the algebraic formulation of the problem. We recall that the solution (y, u, p) solves the following optimality system

$$
\begin{cases}
y - \dfrac{\partial p}{\partial t} + \mathcal{D}_a^*(p) = y_d & \text{in } \Omega \times (0, T), \\
\alpha u - \mathcal{D}_c^\star(p) = 0 & \text{in } \Omega \times (0, T), \\
\dfrac{\partial y}{\partial t} + \mathcal{D}_a(y) - \mathcal{D}_c(u) = g & \text{in } \Omega \times (0, T), \\
y(0) = y_0 & \text{in } \Omega, \\
p(T) = 0 & \text{in } \Omega, \\
\text{boundary conditions} & \text{on } \partial\Omega \times (0, T),
\end{cases}
\tag{22}
$$

where $\mathcal{D}_a$, $\mathcal{D}_c$, $\mathcal{D}_a^\star$ and $\mathcal{D}_c^\star$ are the differential operators associated to $a(\cdot, \cdot; \boldsymbol{\mu})$ and $c(\cdot, \cdot; \boldsymbol{\mu})$ and their adjoint bilinear form, respectively. The time integral has been approximated exploiting the composite rectangle quadrature rule formula. The time interval is divided in N_t sub-intervals of length Δt.

Remark 1. *The applied time discretization is actually equivalent to a classical implicit Euler approach [19, 57]. We underline that the state equation is discretized forward in time with a backward Euler method. The adjoint equation will be discretized backward in time using the forward Euler method, which is equivalent to backward Euler with respect to time $T - t$, for $t \in (0, T)$.*

Let us begin our analysis from the state equation. Let us define $\boldsymbol{y} = [y_1, \ldots, y_{N_t}]^T$ and $\boldsymbol{u} = [u_1, \ldots, u_{N_t}]^T$ and $\boldsymbol{p} = [p_1, \ldots, p_{N_t}]^T$, where $y_i \in Y^\mathcal{N}$, $u_i \in U^\mathcal{N}$ and $p_i \in Y^\mathcal{N}$ for $1 \leq i \leq N_t$ are the vectors of all the discrete variables at each time step. With y_i, u_i and p_i we indicate the row vectors containing the coefficients of the FE discretization for state, control and adjoint, respectively.

The vector representing the initial condition for the state variable is $\boldsymbol{y}_0 = [y_0, 0, \ldots, 0]^T$. The vector $\boldsymbol{g} = [g_1, \ldots, g_{N_t}]^T$ corresponds to the forcing term. Finally, the vector $\boldsymbol{y}_d = [y_{d_1}, \ldots, y_{d_{N_t}}]^T$ is our discretized desired state.

Let D_a and D_c be the matrices associated to the bilinear forms applied to the basis functions of the FE spaces $Y^\mathcal{N} := \text{span}\{\phi_i, 1 \leq i \leq \mathcal{N}\}$ and $U^\mathcal{N} := \text{span}\{\psi_i, 1 \leq i \leq \mathcal{N}\}$, i.e. $D_{a_{ij}} = a(\phi_i, \phi_j; \boldsymbol{\mu})$ and $D_{c_{ij}} = c(\psi_i, \phi_j; \boldsymbol{\mu})$ for $i, j = 1, \ldots, \mathcal{N}$, respectively. For the sake of notation, the dependence of the matrix on the parameter $\boldsymbol{\mu}$ is dropped.

The state equation to be solved is

$$
My_k + \Delta t D_a y_k - \Delta t D_c u_k = My_{k-1} + g_k \Delta t \qquad \text{for } k \in \{1, 2, \ldots, N_t\},
$$

where M is the mass matrix relative to the FE discretization.

In order to solve the system in an all-at-once approach we can write:

$$
\begin{bmatrix}
M + \Delta t D_a & & & & \\
-M & M + \Delta t D_a & & & \\
& -M & M + \Delta t D_a & & \\
& & \ddots & \ddots & \\
& & & -M & M + \Delta t D_a
\end{bmatrix}
\begin{bmatrix}
y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{N_t}
\end{bmatrix}
$$

$$
\underbrace{\phantom{\begin{bmatrix} M \end{bmatrix}}}_{\mathcal{K}}
$$

$$
-\Delta t
\underbrace{
\begin{bmatrix}
D_c & 0 & \cdots & & \\
& D_c & & & \\
& & D_c & & \\
& & & \ddots & \\
& & & & D_c
\end{bmatrix}
}_{\mathcal{C}}
\begin{bmatrix}
u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{N_t}
\end{bmatrix}
=
\begin{bmatrix}
My_0 + \Delta t g_1 \\
0 + \Delta t g_2 \\
0 + \Delta t g_3 \\
\vdots \\
0 + \Delta t g_{N_t}
\end{bmatrix}.
$$

Then, the above system could be written in the following form:

$$
\mathcal{K}\boldsymbol{y} - \Delta t \mathcal{C}\boldsymbol{u} = \mathcal{M}\boldsymbol{y}_0 + \Delta t \boldsymbol{g},
$$

where $\mathcal{M}$ is a block diagonal matrix in $\mathbb{R}^{\mathcal{N} \cdot N_t} \times \mathbb{R}^{\mathcal{N} \cdot N_t}$ which diagonal entries are $[M, \cdots, M]$. In the same way we can analyze the adjoint equation: we have to solve the first equation of the optimality system (22) at each time step as follows:

$$
Mp_k = Mp_{k+1} + \Delta t(-My_k - D_a^T p_k + My_{d_k}) \qquad \text{for } k \in \{N_t - 1, N_t - 2, \ldots, 1\}.
$$

As we did for the state equation, one could use an all-at-once strategy and consider the following system:

$$
\underbrace{
\begin{bmatrix}
M + \Delta t D_a^T & -M & & & \\
& M + \Delta t D_a^T & -M & & \\
& & \ddots & \ddots & \\
& & & M + \Delta t D_a^T & -M \\
& & & & M + \Delta t D_a^T
\end{bmatrix}
}_{\mathcal{K}^T}
\begin{bmatrix}
p_1 \\ p_2 \\ p_3 \\ \vdots \\ p_{N_t}
\end{bmatrix}
+
\begin{bmatrix}
\Delta t M y_1 \\ \Delta t M y_2 \\ \Delta t M y_3 \\ \vdots \\ \Delta t M y_{N_t}
\end{bmatrix}
=
\begin{bmatrix}
\Delta t M y_{d_1} \\ \Delta t M y_{d_2} \\ \Delta t M y_{d_3} \\ \vdots \\ \Delta t M y_{d_{N_t}}
\end{bmatrix}.
$$

Then, the adjoint system to be solved is:

$$
\mathcal{K}^T \boldsymbol{p} + \Delta t \mathcal{M}\boldsymbol{y} = \Delta t \mathcal{M}\boldsymbol{y}_d.
$$

Now, if we consider the optimality equation given by the differentiation of (8) with respect to the control variable, at every time step we have to solve the equation $\alpha \Delta t M u_k - \Delta t D_c^T p_k = 0$. In a vector notation we have $\alpha \Delta t \mathcal{M}\boldsymbol{u} - \Delta t \mathcal{C}^T \boldsymbol{p} = 0$. In the end, the final system considered and solved through an one shot approach is the following:

$$
\begin{bmatrix}
\Delta t \mathcal{M} & 0 & \mathcal{K}^T \\
0 & \alpha \Delta t \mathcal{M} & -\Delta t \mathcal{C}^T \\
\mathcal{K} & -\Delta t \mathcal{C} & 0
\end{bmatrix}
\begin{bmatrix}
\boldsymbol{y} \\ \boldsymbol{u} \\ \boldsymbol{p}
\end{bmatrix}
=
\begin{bmatrix}
\Delta t \mathcal{M}\boldsymbol{y}_d \\
0 \\
\mathcal{M}\boldsymbol{y}_0 + \Delta t \boldsymbol{g}
\end{bmatrix}.
\tag{23}
$$

Now, let us denote with

$$
A = \begin{bmatrix} \Delta t \mathcal{M} & 0 \\ 0 & \alpha \Delta t \mathcal{M} \end{bmatrix}, \quad
B = \begin{bmatrix} \mathcal{K} & -\Delta t \mathcal{C} \end{bmatrix}, \quad
F = \begin{bmatrix} \Delta t \mathcal{M}\boldsymbol{y}_d \\ 0 \end{bmatrix} \quad \text{and} \quad G = \mathcal{M}\boldsymbol{y}_0 + \Delta t \boldsymbol{g}.
$$

Then, the system (23) can be written as follows:

$$\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{p} \end{bmatrix} = \begin{bmatrix} F \\ G \end{bmatrix}. \tag{24}$$

The global dimension of the full order model presented is $\mathcal{N}_{\text{tot}} = 3 \times N_t \times \mathcal{N}$. We underline that the saddle point structure does not depend on the used time approximation scheme: indeed, this framework can be extended to more general time discretization. We managed to solve the linear system (23) through a direct approach. Nevertheless, iterative methods based on Krylov solvers and Schur preconditioning are a very common choice in the solution of saddle point structures [9] and they have been applied in the optimal control framework in [48, 52], for example.

Now that we have introduced the FE discretization as our full order approximation, we can exploit it in order to build a reduced system, as it will be specified in the following section.

3 POD-Galerkin ROM Applied to Parabolic Time Dependent OCP($\boldsymbol{\mu}$)s

This section aims at introducing ROM approximation for time dependent OCP($\boldsymbol{\mu}$)s into the framework of saddle point problems by applying the techniques already used for the steady case in [37, 38]. After a general introduction of the main ideas behind ROM applicability, we will briefly introduce the affine assumption over the bilinear and linear forms involved in the problem formulation. Then, in Section 3.3 we will describe the POD-Galerkin algorithm, see for example [7, 13, 14, 24] as references, and we will extend it to parametrized time dependent OCP($\boldsymbol{\mu}$)s. In the end, we will show how to guarantee the well-posedness of reduced OCP($\boldsymbol{\mu}$)s thanks to the aggregated space strategy [17, 37, 38].

3.1 Reduced Problem Formulation

In Section 2.1, we showed that linear quadratic time dependent OCP($\boldsymbol{\mu}$)s could be seen as saddle point problems of the form (11). We recall to the reader that our state-control variable is $x(\boldsymbol{\mu}) = (y(\boldsymbol{\mu}), u(\boldsymbol{\mu})) \in \mathcal{X}$: in this case we highlight the parameter dependency which is of crucial importance in the understanding of the reduced formulation concepts. We assume that our parametric solution defines a smooth *solution manifold*

$$\mathscr{M} = \{(x(\boldsymbol{\mu}), p(\boldsymbol{\mu})) \mid \boldsymbol{\mu} \in \mathscr{P}\}.$$

The FE approximation is reflected also in the solution manifold: indeed one can define a approximated solution manifold:

$$\mathscr{M}^{\mathcal{N}} = \{(x^{\mathcal{N}}(\boldsymbol{\mu}), p^{\mathcal{N}}(\boldsymbol{\mu})) \mid \boldsymbol{\mu} \in \mathscr{P}\}.$$

Once again, we assume that also the approximated manifold has a smooth dependence with respect to $\boldsymbol{\mu}$. Reduced order methodology wants to recover the structure of $\mathscr{M}^{\mathcal{N}}$ through basis functions derived from properly chosen full order solutions $x^{\mathcal{N}}(\boldsymbol{\mu})$ and $p^{\mathcal{N}}(\boldsymbol{\mu})$ called snapshots. In other words, we construct the reduced bases exploiting FE solutions evaluated for some specific $\boldsymbol{\mu} \in \mathscr{P}$. Let us assume to have already built $\mathcal{X}_N \subset \mathcal{X}^{\mathcal{N}} \subset \mathcal{X}$ and $\mathcal{Y}_{tN} \subset \mathcal{Y}_t^{\mathcal{N}} \subset \mathcal{Y}_t$ as reduced product state-control space and reduced adjoint space, respectively. We have all the ingredients to define the reduced problem as follows: given $\boldsymbol{\mu} \in \mathscr{P}$, find $(x_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in \mathcal{X}_N \times \mathcal{Y}_{tN}$ such that

$$\begin{cases} \mathcal{A}(x_N(\boldsymbol{\mu}), w_N) + \mathcal{B}(w_N, p_N(\boldsymbol{\mu}); \boldsymbol{\mu}) = \displaystyle\int_0^T \langle F(\boldsymbol{\mu}), w_N \rangle \, dt & \forall w_N \in \mathcal{X}_N, \\ \mathcal{B}(x_N(\boldsymbol{\mu}), q_N; \boldsymbol{\mu}) = \displaystyle\int_0^T \langle G(\boldsymbol{\mu}), q_N \rangle \, dt & \forall q_N \in \mathcal{Y}_{tN}. \end{cases} \tag{25}$$

Also at the reduced level, in order to assert the well-posedness of the reduced saddle point problem (25), Brezzi's theorem has to be verified. In Section 3.4, we will show how to recover the existence and the uniqueness of the reduced minimizing solution $(x_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in \mathcal{X}_N \times \mathcal{Y}_{tN}$.

3.2 Affine Assumption: Offline–Online decomposition

Let us briefly underline the crucial hypothesis which guarantees efficient applicability of reduced order methods: affine assumption. A problem in a saddle point framework (11) is affinly decomposed with respect to the parameter if the involved bilinear forms and functionals can be recast as:

$$\mathcal{A}(x, w; \boldsymbol{\mu}) = \sum_{q=1}^{Q_\mathcal{A}} \Theta_\mathcal{A}^q(\boldsymbol{\mu}) \mathcal{A}^q(x, w), \qquad \mathcal{B}(w, p; \boldsymbol{\mu}) = \sum_{q=1}^{Q_\mathcal{B}} \Theta_\mathcal{B}^q(\boldsymbol{\mu}) \mathcal{B}^q(w, p),$$
$$\langle G(\boldsymbol{\mu}), p \rangle = \sum_{q=1}^{Q_G} \Theta_G^q(\boldsymbol{\mu}) \langle G^q, p \rangle, \qquad \langle F(\boldsymbol{\mu}), w \rangle = \sum_{q=1}^{Q_F} \Theta_F^q(\boldsymbol{\mu}) \langle F^q, w \rangle, \tag{26}$$

for some finite $Q_\mathcal{A}, Q_\mathcal{B}, Q_G, Q_F$, where $\Theta_\mathcal{A}^q, \Theta_\mathcal{B}^q, \Theta_G^q, \Theta_F^q$ are $\boldsymbol{\mu}$−dependent smooth functions, whereas $\mathcal{A}^q, \mathcal{B}^q, G^q, F^q$ are $\boldsymbol{\mu}$−independent bilinear forms and functionals.

Thanks to this assumption, the solving process of our OCP($\boldsymbol{\mu}$) can be divided in two different phases: an **offline** stage where the reduced spaces are derived and the $\boldsymbol{\mu}$−independent quantities are assembled. This stage could be very expensive but it is performed only once. Then, an **online** stage follows and all the $\boldsymbol{\mu}$−dependent quantities are assembled and the whole reduced system is solved. The online phase is performed for every new parameter evaluation in order to study different physical and/or geometrical configurations.

If the problem does not fulfill the affine assumption, the empirical interpolation method can be exploited in order to recover it, as presented in [8] or in [24, Chapter 5].

3.3 POD Algorithm for OCP($\boldsymbol{\mu}$)s

In this section we introduce the POD-Galerkin strategy which we exploited in order to build the reduced spaces needed for time dependent OCP($\boldsymbol{\mu}$)s, following [7, 13, 14, 24].

The first step of the the POD-Galerkin approach is to choose a discrete subset of parameters $\mathscr{P}_h \subset \mathscr{P}$. Thanks to this new finite dimensional parametric set, we can define a specific solution manifold:

$$\mathscr{M}^\mathcal{N}(\mathscr{P}_h) = \{(x^\mathcal{N}(\boldsymbol{\mu}), p^\mathcal{N}(\boldsymbol{\mu})) \mid \boldsymbol{\mu} \in \mathscr{P}_h\},$$

which cardinality is $N_{\max} = |\mathscr{P}_h|$ and which satisfies the inclusion $\mathscr{M}^\mathcal{N}(\mathscr{P}_h) \subset \mathscr{M}^\mathcal{N}$ since $\mathscr{P}_h \subset \mathscr{P}$. When the finite parameter space $\mathscr{P}_h$ is large enough, the manifold $\mathscr{M}^\mathcal{N}(\mathscr{P}_h)$ can be a reliable representation of the discrete manifold $\mathscr{M}^\mathcal{N}$. The POD-Galerkin approach compresses the redundant information contained in the snapshots of $\mathscr{M}^\mathcal{N}(\mathscr{P}_h)$.

We exploited a *partitioned approach*, i.e. the POD algorithm has been applied separately for state, control and adjoint variables. In the end, the POD algorithm provides N−dimensional reduced spaces that minimize the quantities:

$$\sqrt{\frac{1}{N_{max}} \sum_{\boldsymbol{\mu} \in \mathscr{P}_h} \min_{z_N \in \mathcal{Y}_{tN}} \|y^\mathcal{N}(\boldsymbol{\mu}) - z_N\|_{\mathcal{Y}_t}^2}, \quad \sqrt{\frac{1}{N_{max}} \sum_{\boldsymbol{\mu} \in \mathscr{P}_h} \min_{v_N \in \mathcal{U}_N} \|u^\mathcal{N}(\boldsymbol{\mu}) - v_N\|_\mathcal{U}^2}, \quad \sqrt{\frac{1}{N_{max}} \sum_{\boldsymbol{\mu} \in \mathscr{P}_h} \min_{q_N \in \mathcal{Y}_{tN}} \|p^\mathcal{N}(\boldsymbol{\mu}) - q_N\|_{\mathcal{Y}_t}^2}.$$
$$\tag{27}$$

We are going to introduce the POD-Galerkin procedure for the state solution $y(\boldsymbol{\mu})$. The same strategy has been used for control and adjoint variables as well.

Let us consider ordered parameters $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_{N_{max}} \in \mathscr{P}_h$ and the resulting ordered FE solutions

$y^{\mathcal{N}}(\boldsymbol{\mu}_1), \ldots, y^{\mathcal{N}}(\boldsymbol{\mu}_{N_{max}})$. Furthermore, we define the correlation matrix $\mathbf{C}^y \in \mathbb{R}^{N_{max} \times N_{max}}$ of snapshots of the state variable, i.e.:

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

We look for the N-largest eigenvalue-eigenvector pairs (λ_n^y, v_n^y), which solve the following equations:

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

with $\|v_n^y\| = 1$. Let us order the eigenvalues $\lambda_1^y \geq \lambda_2^y \geq \cdots \geq \lambda_N^y$ from the largest to the smallest. This order reflects on the basis functions $\{\xi_1^y, \ldots, \xi_N^y\}$ of the reduced space $\mathcal{Y}_{tN} = \text{span}\{\xi_1^y, \ldots, \xi_N^y\}$. The basis satisfies:

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

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

Remark 2. *We underline that, even if we applied the POD algorithm separately for the different variables, we have not separated time instances, i.e. each snapshot still contains the solution at all temporal steps. In this way, the reduced basis functions comply with the space-time formulation introduced in (21), and the resulting POD-Galerkin ROM is a space-time reduced order model.*

3.4 Aggregated Spaces Approach

As we underlined in Section 2.1, the adjoint variable $p(\boldsymbol{\mu})$ is considered in the state space $\mathcal{Y}_t$, in order to ensure the well-posedness of the whole $\text{OCP}(\boldsymbol{\mu})$. It is well known that a POD for state and adjoint variables will not lead necessarily to the same reduced space approximation. Indeed, let us assume to have applied the POD algorithm with the same value of $N_{\max}$ and retaining the first N eigenvalues for all the involved variables, as described in Section 3.3: the procedure provides reduced spaces for state, control and adjoint variables as

$$\begin{aligned}
\mathcal{Y}_{t_N} &= \text{span}\{y^{\mathcal{N}}(\boldsymbol{\mu}^n)\ n = 1, \ldots, N\}, \\
\mathcal{U}_N &= \text{span}\{u^{\mathcal{N}}(\boldsymbol{\mu}^n)\ n = 1, \ldots, N\}, \\
\mathcal{Q}_{t_N} &= \text{span}\{p^{\mathcal{N}}(\boldsymbol{\mu}^n)\ n = 1, \ldots, N\},
\end{aligned}$$

respectively. As already done for the continuous and full order versions of the problem, we define the product space $\mathcal{X}_N = \mathcal{Y}_{t_N} \times \mathcal{U}_N$. Once the reduced spaces are available, it remains to prove if the Brezzi's theorem is still valid, i.e. if the reduced saddle point problem (25) admits an unique solution. The continuity of the bilinear forms $\mathcal{A}(\cdot, \cdot)$ and $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$ are directly inherited from the FE approximation, as well as the coercivity of $\mathcal{A}(\cdot, \cdot)$ over the kernel $\mathcal{X}_0^N = \{w_N \in \mathcal{X}_N : \mathcal{B}(w_N, q; \boldsymbol{\mu}) = 0, \forall q \in \mathcal{Q}_{t_N}\}$, even if the state and the adjoint reduced spaces do not coincide. Although, it is not guaranteed the fulfillment of the *reduced inf-sup condition*. Indeed, only the following inequality holds:

$$\inf_{0 \neq q \in \mathcal{Y}_{t_N}} \sup_{0 \neq x \in \mathcal{X}_N} \frac{\mathcal{B}(x, q; \boldsymbol{\mu})}{\|x\|_{\mathcal{X}} \|q\|_{\mathcal{Y}_t}} \geq \beta_N(\boldsymbol{\mu}) > 0, \tag{28}$$

i.e. the *reduced inf-sup condition* (28) in verified only when $\mathcal{Y}_{t_N} \equiv \mathcal{Q}_{t_N}$, which is not the case for a standard POD approach.

In order to avoid this inconvenience, we exploit the aggregated spaces technique as presented in [17, 37, 38]. The main feature of this approach is to define a common space for state and adjoint variables, given by

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

The space Z_N is then used to describe both the reduced state variable $y_N(\boldsymbol{\mu})$ and the reduced adjoint variable $p_N(\boldsymbol{\mu})$. The new product space is now $\mathcal{X}_N = Z_N \times \mathcal{U}_N$, where the control space is of the standard form

$$\mathcal{U}_N = \text{span} \{u^{\mathcal{N}}(\boldsymbol{\mu}^n), \ n = 1, \dots, N\}.$$

This choice will lead to a global dimension $N_{\text{tot}} = 5N$. Thanks to this strategy, the reduced optimality system is well-posed since all the hypotheses of Brezzi's theorem hold with the *reduced inf-sup* of the following form

$$\inf_{0 \neq q \in Z_N} \sup_{0 \neq x \in \mathcal{X}_N} \frac{\mathcal{B}(x, q; \boldsymbol{\mu})}{\|x\|_{\mathcal{X}} \|q\|_{\mathcal{Y}_t}} \geq \beta_N(\boldsymbol{\mu}) > 0. \tag{29}$$

Now we have all the necessary notions needed in order to show some applications of ROM for time dependent $\text{OCP}(\boldsymbol{\mu})$s. In the next Section we will show how advantageous reduced modelling could be in this very costly context.

4 Numerical Results: Time Dependent OCP($\boldsymbol{\mu}$) for Graetz flows

In this section we are going to present a numerical example in order to validate the performances of POD-Galerkin method for time dependent $\text{OCP}(\boldsymbol{\mu})$s: we will apply our methodology to a time dependent version of the test case proposed in [38].

The proposed test case deals with a time dependent $\text{OCP}(\boldsymbol{\mu})$ governed by a Graetz flow with a control over the boundary $\Gamma_C = ([1, 1 + \mu_3] \times \{0\}) \cup ([1, 1 + \mu_3] \times \{1\})$, which is represented in Figure 4. In this case, μ_3 is a geometrical parameter which stretches the length of $\Omega_2(\mu_3) := [1, 1 + \mu_3] \times [0.2, 0.8]$ and $\Omega_3(\mu_3) := [1, 1 + \mu_3] \times [0, 0.2] \cup [1, 1 + \mu_3] \times [0.8, 1]$ as Figure 4 shows. We now introduce the other two parameters: μ_1 which represents the diffusivity coefficient of the system and μ_2 which is the desired profile solution we want to reach in $\Omega_3(\mu_3)$. The parameter $\boldsymbol{\mu} = [\mu_1, \mu_2, \mu_3]$ is considered in $\mathscr{P} = [1/20, 1/6] \times [1, 3] \times [1/2, 3]$.

Let us specify the function spaces Y and U needed. For this test case we define $y \in \mathcal{Y}_t$ where $Y = H^1_{\Gamma_D}(\Omega(\mu_3))$ and $u \in \mathcal{U}$ where $U = L^2(\Gamma_C(\mu_3))$. Furthermore, let $\mathcal{X}$ be the product space $\mathcal{Y}_t \times \mathcal{U}$. All the data of the test case are recap in Table 1. The problem we consider reads: given $\boldsymbol{\mu} \in \mathscr{P}$, find the state-control variable $(y, u) \in \mathcal{X}$ which solves:

$$\min_{(y,u)\in\mathcal{X}} J(y, u; \boldsymbol{\mu}) = \min_{(y,u)\in\mathcal{X}} \frac{1}{2} \int_0^T \int_{\Omega_3} (y - y_d(\boldsymbol{\mu}))^2 dx dt + \frac{\alpha}{2} \int_0^T \int_{\Gamma_C} u^2 dx dt \tag{30}$$

constrained to the equation

$$\begin{cases} \dfrac{\partial y}{\partial t} + \mu_1 \Delta y + x_2(1 - x_2)\dfrac{\partial y}{\partial x_1} = 0 & \text{in } \Omega(\mu_3) \times (0, T), \\ y = 1 & \text{on } \Gamma_D(\mu_3) \times (0, T), \\ \mu_1 \dfrac{\partial y}{\partial n} = u & \text{on } \Gamma_C(\mu_3) \times (0, T), \\ \mu_1 \dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N(\mu_3) \times (0, T), \\ y = y_0 & \text{in } \Omega(\mu_3) \times \{0\}, \end{cases} \tag{31}$$

where x_1 and x_2 are the spatial components, y_0 is the null function in the domain which respects the boundary conditions and $y_d(\boldsymbol{\mu}) \equiv \mu_2$. As already presented in Section 2.1, we applied a Lagrangian approach and the optimize-then-discretize technique in order to recover the following optimality system: given $\boldsymbol{\mu} \in \mathscr{P}$, find $((y, u), p) \in \mathcal{X} \times \mathcal{Y}_t$ such that

$$
\begin{cases}
y - \dfrac{\partial p}{\partial t} + \mu_1 \Delta p - x_2(1 - x_2)\dfrac{\partial p}{\partial x_1} = y_d(\boldsymbol{\mu}) & \text{in } \Omega(\mu_3) \times (0, T), \\
p = 0 & \text{on } \Gamma_D(\mu_3) \times (0, T) \\
\mu_1 \dfrac{\partial p}{\partial n} = 0 & \text{on } \Gamma_N(\mu_3) \times (0, T), \\
p = 0 & \text{in } \Omega(\mu_3) \times \{T\}, \\
\alpha u = p & \text{in } \Gamma_C(\mu_3) \times (0, T), \\
\dfrac{\partial y}{\partial t} + \mu_1 \Delta y + x_2(1 - x_2)\dfrac{\partial y}{\partial x_1} = 0 & \text{in } \Omega(\mu_3) \times (0, T), \\
y = 1 & \text{on } \Gamma_D \times (0, T), \\
\mu_1 \dfrac{\partial y}{\partial n} = u & \text{on } \Gamma_C(\mu_3) \times (0, T), \\
\mu_1 \dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N(\mu_3) \times (0, T), \\
y = y_0 & \text{in } \Omega(\mu_3) \times \{0\}.
\end{cases}
\tag{32}
$$

The problem has been solved exploiting the following strategy: first of all, we traced back the original problem into the reference domain presented in Figure 4. The reference domain corresponds to $\mu_3 = 1$. The full order discretization is performed as described in Section 2.2: we used $\Delta t = 1/6$ over the interval $(0, T) = (0, 5)$, with a resulting number of time steps $N_t = 30$. Moreover, for the space discretization, we used $\mathbb{P}^1$ elements for all the variables involved with $\mathcal{N} = 3487$, working with function spaces of the form $\mathcal{X}^{\mathcal{N}} := \mathcal{Y}_t^{\mathcal{N}} \times \mathcal{U}^{\mathcal{N}}$ and $\mathcal{Y}_t^{\mathcal{N}}$, for every time step of state-control and adjoint variables, respectively. The total number of degree of freedom of the full order approximation is $\mathcal{N}_{\text{tot}} = 3 \times N_t \times \mathcal{N} = 313'830$. We built the reduced spaces applying the POD-Galerkin approach as presented in Section 3.3. For all the variables we choose $N_{\max} = 70$ snapshots. The basis functions were obtained retaining the first $N = 35$ eigenvectors of the snapshots correlation matrix: from now on we will define N as the *basis number*. In order to guarantee the well-posedness of the reduced saddle point problem arising from the constrained optimization, we exploited aggregated space technique described in Section 3.4: it led to a total reduced dimension of $N_{\text{tot}} = 5N = 175$. The basis considered are sufficient and they well represent the optimal solution as one can observe from the average relative error[1] plots in Figure 1 and the solution plots in Figure 2. The FE simulations and the reduced simulations are compared in Figure 2: they match for different time instances. In Figure 1, we show the average relative errors over a testing set of 50 parameters, uniformly distributed: as expected, it decreases for a high basis number N, with a minimum value of 10^{-4} for all the variables. In Figure 1, we also present the average relative error between the FE and the ROM values of the functional $J(\cdot, \cdot; \boldsymbol{\mu})$, i.e. our *output*. We specify that the effort needed for solving an offline phase for time dependent OCP($\boldsymbol{\mu}$)s drastically increases for high values of $N_{\max}$: still, our choice for the number of snapshots gave us a good overview of the whole parametrized system in a reasonable amount of time. Let us now focus on the computational time involved both in the full and in the reduced order simulations. In Figure 3, the *speedup* index for this test case is presented. The speedup represents how many reduced order simulations can be performed in the time of a single full order FE element simulation. It reaches a maximum value of $1, 8 \cdot 10^5$, while

---

[1]The error for state, control and adjoint variables are presented in the following norms: $\|y^{\mathcal{N}} - y_N\|_{H^1}^2$, $\|u^{\mathcal{N}} - u_N\|_{L^2}^2$ and $\|p^{\mathcal{N}} - p_N\|_{H^1}^2$, respectively.

the lowest values associated to an increasing value of N are not below 10^5. The speedup index underlines how convenient is ROM system for repeated parametric instances of time dependent OCP($\boldsymbol{\mu}$)s, since the very expensive formulation of the whole system (23) is projected in a low dimensional framework which recovers the evolution of the optimality system and permits to study several configurations in the online phase.

Table 1: Data for the OCP($\boldsymbol{\mu}$) governed by a Graetz flow.

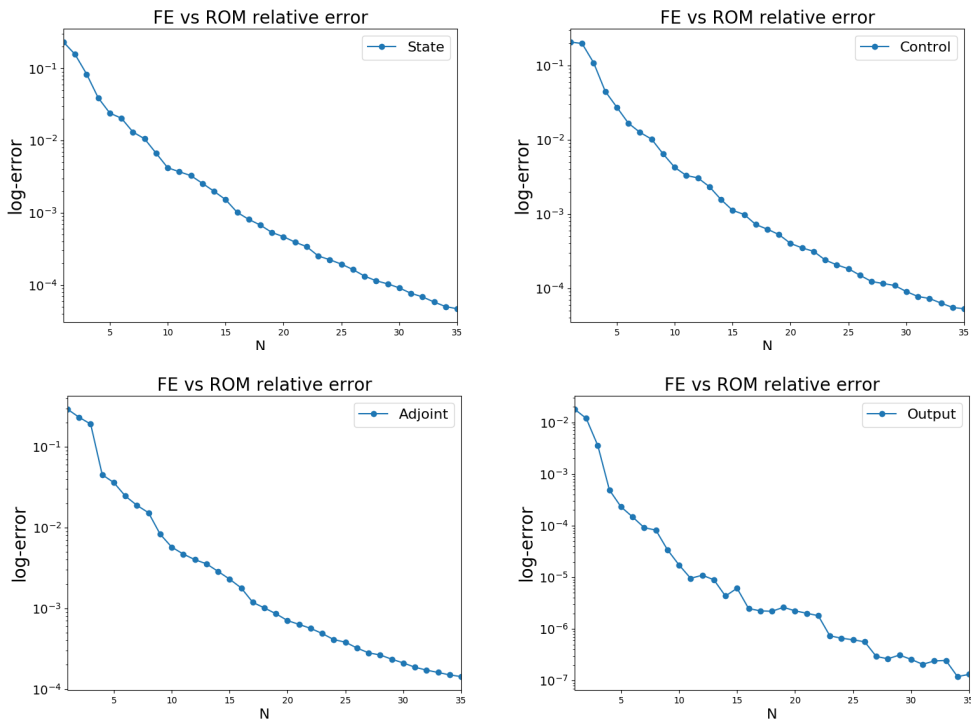| Data	Values
$\mathscr{P}$	$[1/20, 1/6] \times [1, 3] \times [1/2, 3]$
$(\mu_1, \mu_2, \mu_3, \alpha)$	$(1/12, 2, 5/2, 10^{-2})$
$N_{\max}$	70
Basis Number N	35
Sampling Distribution	Uniform
$\mathcal{N}_{\text{tot}}$	$313'830$
N_t	30
ROM System Dimension	175

Figure 1: FE vs ROM relative errors with respect to the basis number N in logarithmic scale. Top left: state error; top right: control error; bottom left: adjoint error; bottom right: output error.
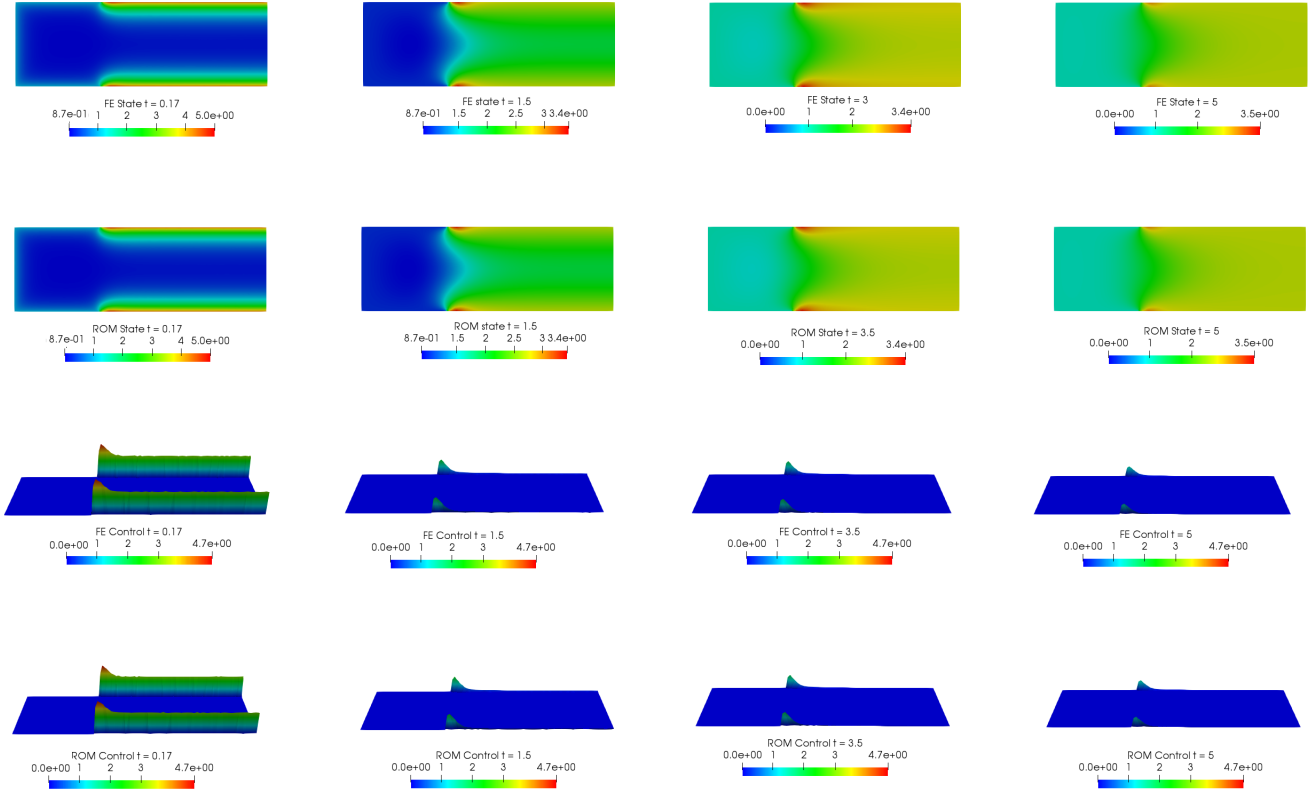
Figure 2: FE vs ROM simulations of state and control variables for $\boldsymbol{\mu} = [1/12, 2, 2.5]$ and $\alpha = 10^{-2}$. First row: FE state for $t = 0.17, 1.5, 3.5, 5$; second row: ROM state for $t = 0.17, 1.5, 3.5, 5$; third row: FE control for $t = 0.17, 1.5, 3.5, 5$; fourth row: ROM control for $t = 0.17, 1.5, 3.5, 5$. In order to better visualize the boundary control, we represent it in a third dimension.
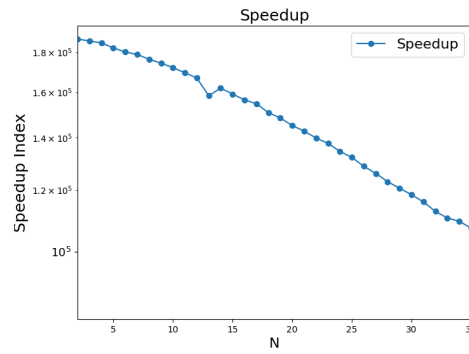

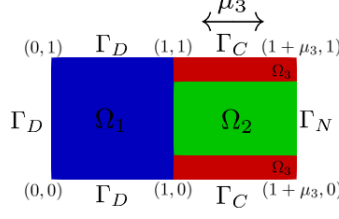**Figure 3:** Speedup analysis in logarithmic scale shown with respect to the basis number N.

Figure 4: Considered reference domain and subdomains for $\mu_3 = 1$.

5 Linear Quadratic Time Dependent OCP($\boldsymbol{\mu}$)s governed by Stokes Equations

In this section, we introduce a distributed OCP($\boldsymbol{\mu}$) governed by time dependent Stokes equations. As we did for the parabolic case in section 2.1, first of all, we provide the proof of well-posedness of this specific linear quadratic OCP($\boldsymbol{\mu}$) in a saddle point framework. Then, we will adapt the space-time discretization and the aggregated POD-Galerkin technique of Sections 2.2 and 3.4, respectively, to this more complex problem.

5.1 Problem Formulation

First of all, we would like to specify the nature of the parametrized problem: in our applications we consider $\boldsymbol{\mu} = [\mu_1, \mu_2]$. The parameters μ_1 and μ_2 characterize the physics and the geometry of the governing equation, respectively. Let us consider the parametrized domain $\Omega(\mu_2) \subset \mathbb{R}^2$. We indicate with $\Gamma_D(\mu_2)$ the portion of the domain where Dirichlet boundary conditions are applied, while $\Gamma_N(\mu_2)$ is the part of $\partial\Omega$ characterized by Neumann conditions. Let us specify the function spaces involved in this example: let us introduce the space $V = H^1_{\Gamma_D(\mu_2)}(\Omega(\mu_2))$, the space $P = L^2(\Omega(\mu_2))$, $\mathcal{V}_t = \{y \in \mathcal{V} := L^2(0, T; V) \text{ s.t. } \frac{\partial y}{\partial t} \in L^2(0, T; V^*)\}$ and $\mathcal{P} = L^2(0, T; P)$. Our parameter dependent state variable $(y, p) \in \mathcal{Y}_t := \mathcal{V}_t \times \mathcal{P}$ is controlled by the variable $u \in \mathcal{U} := L^2(0, T; L^2(\Omega(\mu_2)))$. Following the structure described in Section 2.1, we will exploit the velocity-pressure-control variable $x \in \mathcal{X} := \mathcal{Y}_t \times \mathcal{U}$, where $x := x(\boldsymbol{\mu})$ represents the triplet $(y, p, u) := (y(\boldsymbol{\mu}), p(\boldsymbol{\mu}), u(\boldsymbol{\mu}))$. The norm of the space $\mathcal{X}$ is $\|x\|^2_{\mathcal{X}} = \|y\|^2_{\mathcal{V}_t} + \|p\|^2_{\mathcal{P}} + \|u\|^2_{\mathcal{U}}$. In order to build the OCP($\boldsymbol{\mu}$) we need to define the adjoint variable $(\lambda, \xi) \in \mathcal{Y}_t$. From now on, the pair (λ, ξ) will be indicated with σ. For a given $\boldsymbol{\mu}$ in a suitable parameter space $\mathscr{P} \subset \mathbb{R}^2$, we want to find $(y, p, u) \in \mathcal{X}$ which solves:

$$\min_{(y,p,u) \in \mathcal{X}} \frac{1}{2} \int_0^T \int_\Omega (y - y_d)^2 \, d\Omega dt + \frac{\alpha}{2} \int_0^T \int_\Omega u^2 \, d\Omega dt \tag{33}$$

constrained to the time dependent Stokes equations:

$$\begin{cases} \dfrac{\partial y}{\partial t} - \mu_1 \Delta y + \nabla p = u & \text{in } \Omega(\mu_2) \times (0, T), \\ \text{div}(y) = 0 & \text{in } \Omega(\mu_2) \times (0, T), \\ \dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N(\mu_2) \times (0, T), \\ y = g & \text{on } \Gamma_D(\mu_2) \times (0, T), \\ y(0) = y_0 & \text{in } \Omega(\mu_2) \times \{0\}, \end{cases} \tag{34}$$

where $y_d \in \mathcal{V}_t$ is a desired velocity profile defined on the whole domain. Also in this application, $\|y_0\|_{L^2(\Omega(\mu_2))} = 0$. Again, for the sake of notation, we have omitted the parameter dependency of the variables.

In order to build the optimality system, first of all we consider the linear state equation in its weak formulation, after a lifting procedure. Let $w = (z, q) \in \mathcal{Y}_t$ be a test function: again we consider $w \in \mathcal{Y}_t$ rather than $w \in \mathcal{Y}$ in order to guarantee a future proper definition of the adjoint variable. As we already did in Section 2.1, we can specify the following forms:

$$\mathcal{A} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}, \quad \mathcal{A}(x, w) = m(y, z) + \alpha n(u, v), \qquad\qquad \forall x, w \in \mathcal{X},$$

$$\mathcal{B} : \mathcal{X} \times \mathcal{Y}_t \to \mathbb{R}, \quad \mathcal{B}(x, w; \boldsymbol{\mu}) = \int_0^T \int_\Omega \frac{\partial y}{\partial t} z \, d\Omega dt + \int_0^T a((y, p), (z, q); \boldsymbol{\mu}) \, dt - \int_0^T c(u, z) \, dt, \quad \forall w \in \mathcal{Y}_t,$$

$$F(\boldsymbol{\mu}) \in \mathcal{X}^*, \quad \langle F(\boldsymbol{\mu}), w \rangle = \int_\Omega y_d(\boldsymbol{\mu}) z \, d\Omega, \qquad\qquad \forall w \in \mathcal{Y}_t,$$

$$G \in \mathcal{Y}_t^*, \quad \langle G(\boldsymbol{\mu}), w \rangle = 0, \qquad\qquad \forall w \in \mathcal{Y}_t,$$

where

$$m : \mathcal{Y}_t \times \mathcal{Y}_t \to \mathbb{R}, \qquad m(y, z) = \int_0^T \int_\Omega yz \, d\Omega dt,$$

$$n : \mathcal{U} \times \mathcal{U} \to \mathbb{R}, \qquad n(u, v) = \int_0^T \int_\Omega uv \, d\Omega dt,$$

$$a : \mathcal{Y}_t \times \mathcal{Y}_t \to \mathbb{R}, \qquad a((y, p), (z, q); \boldsymbol{\mu}) = \mu_1 \int_\Omega \nabla y \cdot \nabla z \, d\Omega - \int_\Omega p \operatorname{div}(z) \, d\Omega - \int_\Omega \operatorname{div}(y) q \, d\Omega,$$

$$c : \mathcal{U} \times \mathcal{Y} \to \mathbb{R}, \qquad c(u, q) = \int_\Omega uq \, d\Omega.$$

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

$$\mathsf{a} : \mathcal{V}_t \times \mathcal{V}_t \to \mathbb{R}, \qquad \mathsf{a}(y, z; \boldsymbol{\mu}) = \int_0^T \langle \frac{\partial y}{\partial t}, z \rangle \, dt + \mu_1 \int_0^T \int_\Omega \nabla y \cdot \nabla z \, d\Omega dt, \tag{35}$$

$$\mathsf{b} : \mathcal{V}_t \times \mathcal{P} \to \mathbb{R}, \qquad \mathsf{b}(z, p) = -\int_0^T \int_\Omega p \operatorname{div}(z) \, d\Omega dt. \tag{36}$$

We are now able to define the functional of the form (9) and, then, the Lagrangian functional:

$$\mathscr{L}((y, p), u, (\lambda, \xi); \boldsymbol{\mu}) = \mathcal{J}((y, u); \boldsymbol{\mu}) + \mathcal{B}(x, \sigma; \boldsymbol{\mu}), \tag{37}$$

where $\mathcal{J}(\cdot, \cdot; \boldsymbol{\mu})$ is defined as in (9). After the differentiation of the functional (37) with respect to the variables (y, p, u, λ, ξ), we obtain the optimality system of the form (10):

$$\begin{cases} D_y \mathscr{L}((y, p), u, (\lambda, \xi); \boldsymbol{\mu})[z] = 0 & \forall z \in \mathcal{V}_t, \\ D_p \mathscr{L}((y, p), u, (\lambda, \xi); \boldsymbol{\mu})[q] = 0 & \forall q \in \mathcal{P}, \\ D_u \mathscr{L}((y, p), u, (\lambda, \xi); \boldsymbol{\mu})[v] = 0 & \forall v \in \mathcal{U}, \\ D_\lambda \mathscr{L}((y, p), u, (\lambda, \xi); \boldsymbol{\mu})[\chi] = 0 & \forall \chi \in \mathcal{V}_t. \\ D_\xi \mathscr{L}((y, p), u, (\lambda, \xi); \boldsymbol{\mu})[\tau] = 0 & \forall \tau \in \mathcal{P}. \end{cases} \tag{38}$$

Since the time dependent Stokes equations (34) are linear, the weak formulation of the optimality system (38) can be recast in a saddle point problem of the form (11), already introduced in

Section 2.1.

Given a parameter $\boldsymbol{\mu} \in \mathscr{P}$, we claim that solving the optimality system (38) is equivalent to find the solution $(x, \sigma) \in \mathcal{X} \times \mathcal{Y}_t$ of

$$
\begin{cases}
\mathcal{A}(x, w) + \mathcal{B}(w, \sigma; \boldsymbol{\mu}) = \displaystyle\int_0^T \langle F(\boldsymbol{\mu}), w \rangle \, dt & \forall w \in \mathcal{X}, \\
\mathcal{B}(x, \zeta; \boldsymbol{\mu}) = 0 & \forall \zeta \in \mathcal{Y}_t.
\end{cases}
\tag{39}
$$

We assert that the problem (39) is well-posed. In order to prove the claim, we need the following two lemmas. The first one will be exploited for the *inf-sup condition* of the form $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$. Indeed, the well-posedness of time dependent Stokes equations in mixed formulation implies the second hypothesis of the Brezzi's theorem for the whole state equation $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$.

Lemma 3. *Time dependent Stokes equations with $\|y_0\|_{L^2(\Omega(\mu_2))} = 0$ in their saddle point structure verifies Brezzi's theorem.*

Proof. The continuity of $\mathsf{a}(\cdot, \cdot; \boldsymbol{\mu})$ and of $\mathsf{b}(\cdot, \cdot)$ is obvious. Moreover, the inf-sup condition for the bilinear form $\mathsf{b}(\cdot, \cdot)$ follows from [22, Theorem 4.7 and Proposition 2.2].
It remains to prove the weakly coercivity of $\mathsf{a}(\cdot, \cdot; \boldsymbol{\mu})$ over the kernel of $\mathsf{b}(\cdot, \cdot)$. In our case, the bilinear form $\mathsf{a}(\cdot, \cdot; \boldsymbol{\mu})$ is actually coercive over the whole space $\mathcal{V}_t$ since

$$
\mathsf{a}(y, y; \boldsymbol{\mu}) = \int_0^T \langle \frac{\partial y}{\partial t}, y \rangle \, dt + \mu_1 \int_0^T \int_\Omega \nabla y \cdot \nabla y \, d\Omega dt = \frac{1}{2} \|y(T)\|^2_{L^2(\Omega(\mu_2))} + \mu_1 \|y\|^2_{\mathcal{V}}
$$

$$
\geq \mu_1 \|y\|^2_{\mathcal{V}} \frac{\|y\|^2_{\mathcal{V}_t}}{\|y\|^2_{\mathcal{V}_t}} \underbrace{\geq}_{\text{for Lemma 1}} \frac{\mu_1}{6} \|y\|^2_{\mathcal{V}_t}.
$$

$\square$

A second lemma is needed to demonstrate the coercivity of the bilinear form $\mathcal{A}(\cdot, \cdot)$. Let us indicate the kernel of $\mathcal{B}$ as $\mathcal{X}_0$, as we did in the parabolic case: we are going to prove a norm equivalence which will be used in the proof of the well-posedness of (39).

Lemma 4. *On the space $\mathcal{X}_0$, the norm $\|\cdot\|^2_{\mathcal{X}}$ is equivalent to $\|\cdot\|^2_{\mathcal{V}_t} + \|\cdot\|^2_{\mathcal{U}}$.*

Proof. Let us consider $x = (y, p, u)$ in the kernel of $\mathcal{B}$. First of all, it is obvious that $\|\cdot\|^2_{\mathcal{X}} \geq \|\cdot\|^2_{\mathcal{V}_t} + \|\cdot\|^2_{\mathcal{U}}$. Then, it remains to prove that there exists a positive constant $C_e(\boldsymbol{\mu})$ such that

$$
\|\cdot\|^2_{\mathcal{X}} \leq C_e(\boldsymbol{\mu})(\|\cdot\|^2_{\mathcal{V}_t} + \|\cdot\|^2_{\mathcal{U}}).
\tag{40}
$$

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

$$
\mathsf{b}(w, p) = c(u, w) - \mu_1 \int_0^T \int_\Omega \nabla y \cdot \nabla w \, d\Omega dt - \int_0^T \langle \frac{\partial y}{\partial t}, w \rangle \, dt \qquad \forall w \in \mathcal{V}_t.
$$

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

$$
\mathsf{b}(w, p) \leq \|u\|_{\mathcal{U}} \|w\|_{\mathcal{V}} + \mu_1 \|y\|_{\mathcal{V}} \|w\|_{\mathcal{V}} + \left\| \frac{\partial y}{\partial t} \right\|_{\mathcal{V}^*} \|w\|_{\mathcal{V}}
$$

$$
\leq \min\{1, \mu_1\}(\|y\|_{\mathcal{V}_t} + \|u\|_{\mathcal{U}})\|w\|_{\mathcal{V}_t}.
$$

Now, since the above inequality does not depend from w and p, from lemma 3 we know that

$$
\bar{\beta} \|p\|_{\mathcal{P}} \|w\|_{\mathcal{V}_t} \leq \inf_{0 \neq p \in \mathcal{P}} \sup_{0 \neq w \in \mathcal{V}_t} \mathsf{b}(w, p) \leq \min\{1, \mu_1\}(\|y\|_{\mathcal{V}_t} + \|u\|_{\mathcal{U}})\|w\|_{\mathcal{V}_t}.
$$

Moreover, applying Young's inequality, we have that

$$\|p\|_{\mathcal{P}}^2 \leq \frac{\min\{1,\mu_1\}^2}{\bar{\beta}^2}(\|y\|_{\mathcal{V}_t} + \|u\|_{\mathcal{U}})^2 \leq 2\frac{\min\{1,\mu_1\}^2}{\bar{\beta}^2}(\|y\|_{\mathcal{V}_t}^2 + \|u\|_{\mathcal{U}}^2).$$

Setting $\bar{C}_e(\boldsymbol{\mu}) = \left(2\frac{\min\{1,\mu_1\}^2}{\bar{\beta}^2}\right)$ the inequality (40) is verified with $C_e(\boldsymbol{\mu}) = \min\{1, \bar{C}_e(\boldsymbol{\mu})\}$. $\quad\square$

Thanks to these two lemmas, we are now able to demonstrate the following theorem, which guarantees existence and uniqueness of the optimal solution for OCP($\boldsymbol{\mu}$)s governed by time dependent Stokes equations.

Theorem 2. *The problem* (39) *is equivalent to the minimization problem* (33) *under the constraint* (34). *Moreover, the saddle point problem* (39) *admits a unique solution.*

Proof. In order to prove existence and uniqueness of the solution of problem (39) and its equivalence with the minimization problem (33)-(34), we exploit Proposition 1, i.e. we show that:

- $\mathcal{A}(\cdot,\cdot)$ is a symmetric positive definite continuous bilinear form coercive on $\mathcal{X}_0$;

- $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ is continuous and satisfies the *inf-sup condition*.

The form $\mathcal{A}(\cdot,\cdot)$ is trivially symmetric and positive definite thanks to its definition. The continuity of the bilinear forms $\mathcal{A}(\cdot,\cdot)$ and $\mathcal{B}(\cdot,\cdot;\boldsymbol{\mu})$ can be derived using the same techniques already presented in theorem 1. Let us now focus on the coercivity of $\mathcal{A}(\cdot,\cdot)$. For every $x \in \mathcal{X}_0$ we want to prove that there exists a positive constant C such that

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

Since norm of $\|\cdot\|_{\mathcal{X}}^2$ is equivalent to $\|\cdot\|_{\mathcal{V}_t}^2 + \|\cdot\|_{\mathcal{U}}^2$ as we proved in Lemma 4, it is sufficient to show that

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

When $x \in \mathcal{X}_0$ it means that the pair (y,p) verifies the Stokes equations with the control as forcing term. It is well known that the velocity solution of Stokes equations with $\|y_0\|_{L^2(\Omega(\mu_2))} = 0$ and control as forcing term admits a $k(\boldsymbol{\mu}) > 0$ such that $\|y\|_{\mathcal{V}_t} \leq k(\boldsymbol{\mu})\|u\|_{\mathcal{U}}$, see [43, Chapter 13]. Then for $x \in \mathcal{X}_0$

$$\mathcal{A}(x,x) \geq \|y\|_{L^2(0,T;H)}^2 + \alpha\|u\|_{\mathcal{U}}^2 \geq \frac{\alpha}{2}\|u\|_{\mathcal{U}}^2 + \frac{\alpha}{2}\|u\|_{\mathcal{U}}^2 \geq \frac{\alpha}{2k(\boldsymbol{\mu})^2}\|y\|_{\mathcal{V}_t}^2 + \frac{\alpha}{2}\|u\|_{\mathcal{U}}^2$$

$$\geq \min\left\{\frac{\alpha}{2k(\boldsymbol{\mu})^2}, \frac{\alpha}{2}\right\}(\|y\|_{\mathcal{V}_t}^2 + \|u\|_{\mathcal{U}}^2).$$

It remains to show that

$$\inf_{0 \neq q \in \mathcal{Y}_t} \sup_{0 \neq x \in \mathcal{X}} \frac{\mathcal{B}(x,q;\boldsymbol{\mu})}{\|x\|_{\mathcal{X}}\|q\|_{\mathcal{Y}_t}} > \beta(\boldsymbol{\mu}) > 0. \tag{41}$$

First of all, let us consider the bilinear form $\mathsf{A} : \mathcal{Y}_t \times \mathcal{Y}_t \to \mathbb{R}$ defined as $\mathsf{A}((y,p),(w,q);\boldsymbol{\mu}) = \mathsf{a}(y,w) + \mathsf{b}(w,p) + \mathsf{b}(v,q)$. We can now proceed as presented in [37, Appendix A.1]. Thanks to lemma 3, we infer that the operator $\mathsf{A}(\cdot,\cdot;\boldsymbol{\mu})$ is a well-posed mixed problem, then the Babuška inf-sup constant $\beta_B(\boldsymbol{\mu})$ is well defined as

$$\beta_B(\boldsymbol{\mu}) = \inf_{0 \neq (y,p) \in \mathcal{Y}_t} \sup_{0 \neq (w,q) \in \mathcal{Y}_t} \frac{\mathsf{A}((y,p),(w,q);\boldsymbol{\mu})}{\|(y,p)\|_{\mathcal{Y}_t}\|(w,q)\|_{\mathcal{Y}_t}} = \inf_{0 \neq (w,q) \in \mathcal{Y}_t} \sup_{0 \neq (y,p) \in \mathcal{Y}_t} \frac{\mathsf{A}((y,p),(w,q);\boldsymbol{\mu})}{\|(y,p)\|_{\mathcal{Y}_t}\|(w,q)\|_{\mathcal{Y}_t}},$$
$$(42)$$

see the classical reference [3]. Then we have the following inequalities:

$$\sup_{0 \neq x \in \mathcal{X}} \frac{\mathcal{B}(x,\sigma;\boldsymbol{\mu})}{\|x\|_{\mathcal{X}}\|\sigma\|_{\mathcal{Y}_t}} \underset{x=(y,p,0)}{\geq} \sup_{0 \neq (y,p) \in \mathcal{Y}_t} \frac{\mathsf{A}((y,p),(w,q);\boldsymbol{\mu})}{\|(y,p)\|_{\mathcal{Y}_t}\|(w,q)\|_{\mathcal{Y}_t}}$$

$$\geq \inf_{0 \neq (w,q) \in \mathcal{Y}_t} \sup_{0 \neq (y,p) \in \mathcal{Y}_t} \frac{\mathsf{A}((y,p),(w,q);\boldsymbol{\mu})}{\|(y,p)\|_{\mathcal{Y}_t}\|(w,q)\|_{\mathcal{Y}_t}} \geq \beta_B(\boldsymbol{\mu})$$

Since the above inequality does not depend on the choice of $\sigma \in \mathcal{Y}_t$, then the inequality (41) is verified and the OCP($\boldsymbol{\mu}$) is well-posed.

∎ □

In this section we generalized the saddle point theory of OCP($\boldsymbol{\mu}$)s governed by parabolic state equations to OCP($\boldsymbol{\mu}$)s under Stokes equations constraint. After the introduction of the minimization problem, we provided a proof for the well-posedness of the OCP($\boldsymbol{\mu}$) in the saddle point framework. The next Section will focus on the full order discretization of the OCP($\boldsymbol{\mu}$) proposed.

5.2 Finite Element Discretization and Time Discretization for OCP($\boldsymbol{\mu}$)s governed by Stokes Equations: all-at-once approach

In this section we introduce the discretized version of time dependent OCP($\boldsymbol{\mu}$)s governed by Stokes equations following the properties of the schemes already presented in [25, 52].
In order to expose the discrete formulation of the optimality system (38), we need its strong formulation. The minimization of the functional (33) constrained to time dependent Stokes equations (34) is equivalent to find the pair $(x,\sigma) \in \mathcal{X} \times \mathcal{Y}_t$ such that the state, the adjoint and the optimality equations read as follows:

$$\begin{cases} \dfrac{\partial y}{\partial t} - \mu_1 \Delta y + \nabla p = u & \text{in } \Omega(\mu_2) \times (0,T), \\ \text{div}(y) = 0 & \text{in } \Omega(\mu_2) \times (0,T), \\ y(t) = g(t) & \text{on } \Gamma_D(\mu_2) \times (0,T), \\ \dfrac{\partial y}{\partial n} = 0 & \text{on } \Gamma_N(\mu_2) \times (0,T), \\ y(0) = y_0 & \text{in } \Omega(\mu_2) \times \{0\}, \\ y - \dfrac{\partial \lambda}{\partial t} - \mu_1 \Delta \lambda + \nabla \xi = y_d & \text{in } \Omega(\mu_2) \times (0,T), \\ \text{div}(\lambda) = 0 & \text{in } \Omega(\mu_2) \times (0,T), \\ \lambda(t) = 0 & \text{on } \partial\Omega(\mu_2) \times (0,T), \\ \lambda(T) = 0 & \text{in } \Omega(\mu_2) \times \{T\}, \\ \alpha u = \lambda & \text{in } \Omega(\mu_2) \times (0,T), \\ \text{boundary conditions} & \text{on } \partial\Omega \times (0,T). \end{cases} \quad (43)$$

In order to discretize in time and in space the system (43), we will exploit the technique already presented in Section 2.2 for parabolic time dependent OCP($\boldsymbol{\mu}$)s.

Let us define $\boldsymbol{y} = [y_1, \ldots, y_{N_t}]^T, \boldsymbol{p} = [p_1, \ldots, p_{N_t}]^T$, $\boldsymbol{u} = [u_1, \ldots, u_{N_t}]^T$ and $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_{N_t}]^T$, $\boldsymbol{\xi} = [\xi_1, \ldots, \xi_{N_t}]^T$, i.e. the vectors where y_i, u_i, p_i, λ_i and ξ_i for $1 \leq i \leq N_t$ are row vectors representing the coefficients of the FE discretization of each time instance. Also in this case we divided the time interval in N_t sub-intervals of length Δt. The vector representing the initial condition for the state variable is $\boldsymbol{y}_0 = [y_0, 0, \ldots, 0]^T$. The discretized desired state and the forcing term are given by $\boldsymbol{y}_d = [y_{d_1}, \ldots, y_{d_{N_t}}]^T$ and $\boldsymbol{g} = [g_1, \ldots, g_{N_t}]^T$, respectively.

In order to create the time steps, we used the rectangle composite quadrature formula following the same arguments as already presented in Remark 1. The discretized state equation reads as follows:

$$\begin{cases} My_k + \mu_1 \Delta t K y_k + \Delta t D^T p_k = \Delta t M u_k + My_{k-1} + \Delta t g_k & \text{for } k \in \{1, 2, \ldots, N_t\}, \\ Dy_k = 0 & \text{for } k \in \{1, 2, \ldots, N_t\}, \end{cases} \quad (44)$$

where M and K are the mass and the stiffness matrices relative to the FE discretization, respectively. Moreover, D is the differential operator representing to the continuity equation. Then, in order to solve the system in an all-at-once approach we have to solve the following system:

$$\mathcal{K}[\boldsymbol{y}, \boldsymbol{p}] - \Delta t \mathcal{C} \boldsymbol{u} = \mathcal{M} \boldsymbol{y}_0 + \Delta t \boldsymbol{g},$$

where:

$$\mathcal{K} = \begin{bmatrix} M - \mu\Delta t K & 0 & \cdots & 0 & \Delta t D^T & 0 & \cdots & & 0 \\ D & 0 & & & & & & & \\ -M & M - \mu\Delta t K & 0 & \cdots & 0 & \Delta t D^T & 0 & \cdots & 0 \\ 0 & D & & & & & & & \\ & & \ddots & & \ddots & & & & \\ & & & \ddots & \ddots & & & & \\ & & & & \ddots & \ddots & & & \\ & & & & & \ddots & & \ddots & \\ 0 & \cdots & & 0 & -M & M - \mu\Delta t K & 0 & \cdots & 0 & \Delta t D^T \\ 0 & \cdots & & & & D & & 0 & \cdots & 0 \end{bmatrix}$$

and $\mathcal{C} = \mathcal{M} \in \mathbb{R}^{\mathcal{N} \cdot N_t} \times \mathbb{R}^{\mathcal{N} \cdot N_t}$ is the block diagonal matrix which diagonal entries are $[M, \cdots, M]$. Then, we discretize the optimality equation, which has the following form:

$$\alpha \Delta t M u_k - \Delta t M \lambda_k = 0 \qquad \text{for } k \in \{1, 2, \ldots, N_t\}.$$

This means that the system considered is of the form:

$$\alpha \Delta t \mathcal{M} \boldsymbol{u} - \Delta t \mathcal{C}^T \boldsymbol{\lambda} = 0.$$

Finally, we are able to treat the adjoint equation. We use a forward Euler as we did for the parabolic case in Section 2.2. Thanks to this approach, it holds:

$$\begin{cases} M\lambda_k = M\lambda_{k+1} + \Delta t(-My_k + \mu_1 \Delta t K \lambda_k - \Delta t D^T \xi_k + My_{d_k}) & \text{for } k \in \{N_t - 1, N_t - 2, \ldots, 1\}, \\ D\lambda_k = 0 & \text{for } k \in \{N_t - 1, N_t - 2, \ldots, 1\}. \end{cases} \quad (45)$$

In the end, the adjoint equation has the following form:

$$\Delta t \mathcal{M} \boldsymbol{y} + \mathcal{K}^T[\boldsymbol{\lambda}, \boldsymbol{\xi}] = \Delta t \mathcal{M} \boldsymbol{y}_d.$$

Then, the final parametrized system to be solve trough a one shot approach is the of the form presented in (23), i.e.:

$$
\begin{bmatrix} \Delta t \mathcal{M} & 0 & \mathcal{K}^T \\ 0 & \alpha \Delta t \mathcal{M} & -\Delta t \mathcal{C}^T \\ \mathcal{K} & -\Delta t \mathcal{C} & 0 \end{bmatrix} \begin{bmatrix} (\boldsymbol{y}, \boldsymbol{p}) \\ \boldsymbol{u} \\ (\boldsymbol{\lambda}, \boldsymbol{\xi}) \end{bmatrix} = \begin{bmatrix} \Delta t \mathcal{M} \boldsymbol{y}_d \\ 0 \\ \mathcal{M} \boldsymbol{y}_0 + \Delta t \boldsymbol{g} \end{bmatrix}.
\tag{46}
$$

Now, calling

$$
A = \begin{bmatrix} \Delta t \mathcal{M} & 0 \\ 0 & \alpha \Delta t \mathcal{M} \end{bmatrix}, \quad B = \begin{bmatrix} \mathcal{K} & -\Delta t \mathcal{C} \end{bmatrix}, \quad F = \begin{bmatrix} \Delta t \mathcal{M} \boldsymbol{y}_d \\ 0 \end{bmatrix} \quad \text{and} \quad G = \mathcal{M} \boldsymbol{y}_0 + \Delta t \boldsymbol{g}
$$

the system (23) can be written as follows:

$$
\begin{bmatrix} A & B^T \\ B & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ \boldsymbol{\chi} \end{bmatrix} = \begin{bmatrix} F \\ G \end{bmatrix},
\tag{47}
$$

where $\boldsymbol{x} = ((\boldsymbol{y}, \boldsymbol{p}), \boldsymbol{u})$ and $\boldsymbol{\chi} = (\boldsymbol{\lambda}, \boldsymbol{\xi})$.

As in the parabolic case, we used a direct approach in order to find the solution to the saddle point system (47), even if iterative solvers are widely exploited for this kind problem: the interested reader may refer to [25, 48, 52].

We now move towards ROM version of the system, focusing on supremizers aggregated space technique [37] which guarantees the well-posedness of the reduced optimality system.

5.3 Reduced basis method: Supremizer Stabilization and Aggregated Spaces

In this section, we exploit a POD-Galerkin strategy in order to recast the full order optimality system in a reduced framework. The technique used is the same already presented in Section 3.3: we remark that in this case we apply the POD algorithm for each space-time FE variables $v^{\mathcal{N}}, p^{\mathcal{N}}, u^{\mathcal{N}}, \lambda^{\mathcal{N}}$ and $\xi^{\mathcal{N}}$, which contain all the temporal instances.

Let us suppose to have built the reduced spaces with the POD approach, then the reduced optimality system reads: given $\boldsymbol{\mu} \in \mathscr{P}$, find $(x_N(\boldsymbol{\mu}), \sigma_N(\boldsymbol{\mu})) \in \mathcal{X}_N \times \mathcal{Y}_N$ such that

$$
\begin{cases} \mathcal{A}(x_N(\boldsymbol{\mu}), w_N) + \mathcal{B}(w_N, \sigma_N(\boldsymbol{\mu}); \boldsymbol{\mu}) = \langle F(\boldsymbol{\mu}), w_N \rangle & \forall w_N \in \mathcal{X}_N, \\ \mathcal{B}(x_N(\boldsymbol{\mu}), \zeta_N; \boldsymbol{\mu}) = \langle G(\boldsymbol{\mu}), \zeta_N \rangle & \forall \zeta_N \in \mathcal{Y}_N. \end{cases}
\tag{48}
$$

Anyway, in this case we are managing a nested saddle point structure, since we are dealing with Stokes equations as constraints. Indeed, the state equation problem is formulated as: given $\boldsymbol{\mu} \in \mathscr{P}$, find $(y_N(\boldsymbol{\mu}), p_N(\boldsymbol{\mu})) \in \mathcal{V}_{t_N} \times \mathcal{P}_N$ such that

$$
\begin{cases} \mathsf{a}(y_N(\boldsymbol{\mu}), z_N; \boldsymbol{\mu}) + \mathsf{b}(z_N, p_N(\boldsymbol{\mu})) = 0 & \forall z_N \in \mathcal{V}_{t_N}, \\ \mathsf{b}(y_N(\boldsymbol{\mu}), \zeta_N) = 0 & \forall \zeta_N \in \mathcal{P}_N, \end{cases}
\tag{49}
$$

where the bilinear forms $\mathsf{a}(\cdot, \cdot; \boldsymbol{\mu})$ and $\mathsf{b}(\cdot, \cdot)$ are defined as (35) and (36), respectively, and $\mathcal{V}_{t_N}$ and $\mathcal{P}_N$ are the reduced spaces obtained trough POD algorithm over velocity and pressure variables. It is well known that, in order to guarantee the well-posedness of the reduced Stokes state equations (49), the *reduced inf-sup stability condition* is required for the bilinear form $\mathsf{b}(\cdot, \cdot)$ and it is not directly inherited from the FE approximation. In other words, we have to build our reduced spaces for velocity and pressure variables such that there exists a $\bar{\beta}_N > 0$ which verifies

$$
\inf_{p_N \in \mathcal{P}_N} \sup_{y_N \in \mathcal{V}_{t_N}} \frac{\mathsf{b}(y_N, p_N; \boldsymbol{\mu})}{\|y_N\|_{\mathcal{V}_t} \|p_N\|_{\mathcal{P}}} \geq \bar{\beta}_N(\boldsymbol{\mu}) > 0 \qquad \forall \boldsymbol{\mu} \in \mathscr{P}.
\tag{50}
$$

To ensure the inequality (50) at the reduced level, we follow the strategy of pressure supremizers [47]: let us consider the supremizer operator $T_p^{\boldsymbol{\mu}} : \mathcal{P}^{\mathcal{N}} \to \mathcal{V}_t^{\mathcal{N}}$ defined as follows:

$$(T_p^{\boldsymbol{\mu}} s, \phi)_{\mathcal{V}_t} = \mathsf{b}(\phi, s; \boldsymbol{\mu}), \qquad \forall \phi \in \mathcal{V}_t^{\mathcal{N}}. \tag{51}$$

Then, we enrich the reduced velocity space with supremizers and we build a new space as follows:

$$\mathcal{V}_{tN}^{\boldsymbol{\mu}} = \mathrm{span}\{y^{\mathcal{N}}(\boldsymbol{\mu}^n), \ T_p^{\boldsymbol{\mu}^n} p^{\mathcal{N}}(\boldsymbol{\mu}^n), \ n = 1, \dots, N\}.$$

If we exploit this new reduced space $\mathcal{V}_{tN}^{\boldsymbol{\mu}}$ for velocity, then (50) is verified.

Once proved the stability of the Stokes equations, we can focus on the whole OCP($\boldsymbol{\mu}$) governed by this particular state equation. In order verify the *reduced inf-sup condition* for the bilinear form $\mathcal{B}(\cdot, \cdot; \boldsymbol{\mu})$, we used the aggregated spaces strategy already presented in Section 3.4. As shown in [37], we define the aggregated spaces for both the state and adjoint pressure variables as

$$\mathcal{P}_N = \mathrm{span}\{p^{\mathcal{N}}(\boldsymbol{\mu}^n), \xi^{\mathcal{N}}(\boldsymbol{\mu}^n), \ n = 1, \dots, N\}, \tag{52}$$

while, for the state and the adjoint velocity variables, we consider

$$\mathcal{V}_{tN}^{\boldsymbol{\mu}^n} = \mathrm{span} \ \{y^{\mathcal{N}}(\boldsymbol{\mu}^n), T_p^{\boldsymbol{\mu}^n} p^{\mathcal{N}}(\boldsymbol{\mu}^n), \lambda^{\mathcal{N}}(\boldsymbol{\mu}^n), T_p^{\boldsymbol{\mu}^n} \xi^{\mathcal{N}}(\boldsymbol{\mu}^n), \ n = 1, \dots, N\}. \tag{53}$$

Let us now define the following *aggregated* space for state and adjoint velocity-pressure variables

$$\mathcal{Y}_N = \mathcal{V}_{tN}^{\boldsymbol{\mu}^n} \times \mathcal{P}_N.$$

Finally, the control space is:

$$\mathcal{U}_N = \mathrm{span} \ \{u^{\mathcal{N}}(\boldsymbol{\mu}^n), \ n = 1, \dots, N\}. \tag{54}$$

Considering the product space $\mathcal{X}_N = \mathcal{Y}_N \times \mathcal{U}_N$, the well-posedness of the reduced optimality system (48) in its nested saddle point structure is verified. The stability of these techniques is numerically demonstrated in [7, 20, 37, 44, 46]. Thanks to this new reduced spaces formulation, the dimension of the state and the adjoint spaces is $6N$, whereas the control space has dimension N, for a total reduced dimension $N_{\mathrm{tot}} = 13N$. In the next section, we are going to test this methodology in a numerical simulation, showing that ROM online phase could be of great advantage, despite the increasing value of N_{tot}.

6 Numerical Results: Time Dependent OCP($\boldsymbol{\mu}$) for a Cavity Viscous Flow

In this section we propose a geometrical and physical parametrized version of the test case already presented in [25, 52]: it is a time dependent OCP($\boldsymbol{\mu}$) for cavity viscous flow problem. The parametrization presented has been studied also in [2], where several stabilization techniques are proposed. In this case, we deal with the consistent FE pair $\mathbb{P}^2 - \mathbb{P}^1$ for velocity and pressure variables, respectively: then, no stabilization is needed except for the supremizer techniques described in Section 5.3. For the discretization of the control variable, we exploited $\mathbb{P}^2$ polynomials for its FE approximation. We call $\mathcal{N}_y$, $\mathcal{N}_p$ and $\mathcal{N}_u$ the FE dimensions for velocity, pressure and control space, respectively. Our space discretization leads to $\mathcal{N}_y = \mathcal{N}_u = 4554$ and $\mathcal{N}_p = 591$. Then we define the global FE dimension as $\mathcal{N} = 2\mathcal{N}_y + 2\mathcal{N}_p + \mathcal{N}_u$. The OCP($\boldsymbol{\mu}$) is based on the minimization of the functional (33) with (34) as state equation. We now provide the information about the numerical test, which are resumed in Table 2. First of all, let us specify the role of the parameters: μ_1 is physical parameter describing the diffusivity action of

the system, while μ_2 changes the geometry of the problem, stretching the length of the domain. In order to deal with the geometrical parametrization, we traced back $\Omega(\mu_2) \subset \mathbb{R}^2$ into a two dimensional square $(0,1) \times (0,1)$ as reference domain, where $\Gamma_{IN} = (0,1) \times \{1\}$ is the inlet boundary and $\Gamma_D = \partial\Omega \setminus \Gamma_{IN}$ is characterized by homogeneous Dirichlet boundary conditions. The reference domain structure is presented in Figure 6. Since we considered Dirichlet boundary conditions all over the domain, we assumed $p \in \mathcal{P} = L^2(0,T; L_0^2(\Omega(\mu_2)))^2$. The aim of the OCP($\boldsymbol{\mu}$) proposed is to make the state velocity y the most similar to a target velocity y_d, given $\boldsymbol{\mu} \in \mathscr{P} = [10^{-3}, 10^{-1}] \times [0.5, 2.5]$. The target velocity profile y_d is defined as the FE solution of the uncontrolled time dependent Stokes equations with Dirichlet boundary conditions given by the constant velocity components $(1,0)$ in Γ_{IN} and homogeneous Dirichlet boundary conditions in Γ_D, for $\mu_1 = 1$ fixed. The target velocity has been simulated in $(0,T) = (0,1)$. For the OCP($\boldsymbol{\mu}$) test case, we consider a different time dependent inlet boundary condition: the state velocity profile is $y = \left(1 + \frac{1}{2}\cos(4\pi t - \pi), 0\right)$ over Γ_{IN} for $t \in (0,T) = (0,1)$. As already specified, homogeneous Dirichlet boundary conditions are considered elsewhere. The distributed control u has the role to reduce the impact of the periodic inlet over the system. The FE discretization is performed as presented in Section 5.2. For the time discretization we used $\Delta t = 0.05$, resulting in $N_t = 20$ time instances. The total number of degree of freedom is $\mathcal{N}_{\text{tot}} = N_t \times \mathcal{N} = 296'880$. In order to reduce the dimension of the FE system, we applied the POD-Galerkin algorithm described in Section 3.3. For all the variables involved, we choose $N_{\max} = 70$ snapshots for the correlation matrix and we only retain the first $N = 25$ eigenvectors as basis functions of our low-dimensional spaces. Once again, we will call N the *basis number*. As already specified in Section 4, the choice of the number of snapshots $N_{\max}$ is strongly related to the complexity of the offline phase of the time dependent optimality system, which turned out to be very expensive to be solved. The total dimension of the reduced system is $N_{\text{tot}} = 13N = 325$, tanking into account supremizer aggregated space technique for the nested saddle point problem given by the Stokes equations and the optimal control structure, as already specified in Section 5.3. A comparison between FE and ROM state velocity and state pressure profiles is shown in Figure 5: the ROM simulations recover the behaviour of the FE solutions for different time instances. The accuracy of the method is also underlined by the plots in Figure 8, where the average relative error[3] plots over a testing set of 35 uniformly distributed parameters are presented: the greater is the value of N the better are the results, as expected. The relative error is about 10^{-3} for all the involved variables. In Figure 8, we also show the relative error between the value of the functional *output* (33) evaluated for the FE solution and for the ROM solution: it reaches the very low value of 10^{-8}. Also in this test case, ROM is a strategy which reaches accurate results in a small computational time. In Figure 7 we show the *speedup* index for this test case: the POD-Galerkin approach could be very useful for time dependent OCP($\boldsymbol{\mu}$)s governed by Stokes equations, since it guarantees a good approximation of a quite complex system in low-dimensional framework. The number of reduced problem one can perform in a single FE simulation decreases with respect to the value of N, but still it is never below the value of $6 \cdot 10^4$. Also in this case, from the high speedup values, we can conclude that ROM is a good approach

---

[2] The space $L_0^2(\Omega(\mu_2))$ is made by functions $p \in L^2(\Omega(\mu_2))$ which satisfy $\int_\Omega p \, d\Omega = 0$. In the reduced model the aggregated basis associated to the state and adjoint pressure was built in order to satisfy this constraint, i.e. the reduced adjoint variable ξ_N verifies $\int_\Omega \xi_N \, d\Omega = 0$.

[3]The error for state velocity and pressure, control and adjoint velocity and pressure variables are: $\|y^{\mathcal{N}} - y_N\|_{H^1}^2$, $\|p^{\mathcal{N}} - p_N\|_{L^2}^2$, $\|u^{\mathcal{N}} - u_N\|_{L^2}^2$, $\|\lambda^{\mathcal{N}} - \lambda_N\|_{H^1}^2$ and $\|\xi^{\mathcal{N}} - \xi_N\|_{L^2}^2$, respectively. We underline that in order to make the FE and ROM adjoint pressures comparable we define $\xi^{\mathcal{N}} := \bar{\xi}^{\mathcal{N}} - \int_\Omega \bar{\xi}^{\mathcal{N}} dx$, where $\bar{\xi}^{\mathcal{N}}$ is the actual truth solution.

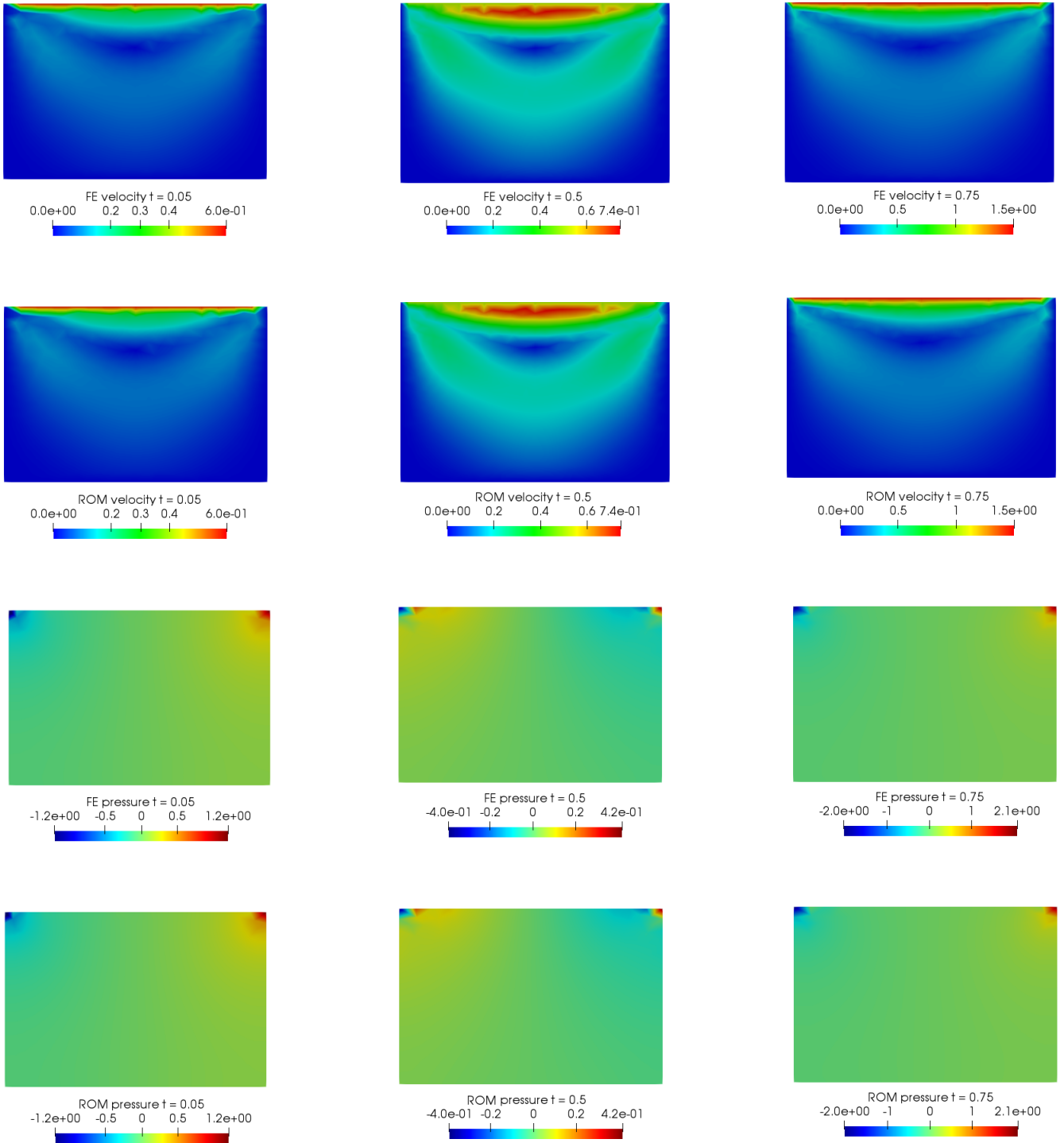to manage several repeated simulations for different values of the parameter considered.

Figure 5: FE vs ROM simulations of state velocity and pressure variables for $\boldsymbol{\mu} = [1/2, 10^{-2}]$ and $\alpha = 10^{-2}$. First row: FE state velocity for $t = 0.05, 0.5, 0.75$; second row: ROM state velocity for $t = 0.05, 0.5, 0.75$; third row: FE state pressure for $t = 0.05, 0.5, 0.75$; fourth row: ROM state pressure for $t = 0.05, 0.5, 0.75$.
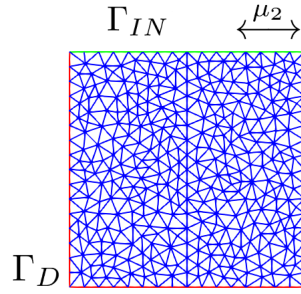
Figure 6: Domain and boundary considered.

Table 2: Data for the OCP($\boldsymbol{\mu}$) governed by a Stokes equation.

Data	Values
$\mathscr{P}$	$[0.5, 2.5] \times [10^{-3}, 10^{-1}]$
(μ_1, μ_2, α)	$(10^{-2}, 3/2, 10^{-2})$
$N_{\max}$	70
Basis Number N	25
Sampling Distribution	Uniform
$\mathcal{N}_{\text{tot}}$	296$'$880
N_t	20
ROM System Dimension	325
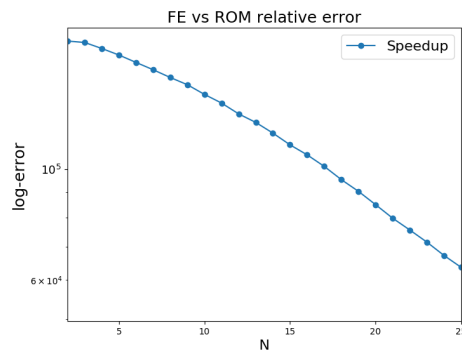
Figure 7: Speedup analysis in logarithmic scale shown with respect to the reduced basis dimension N.
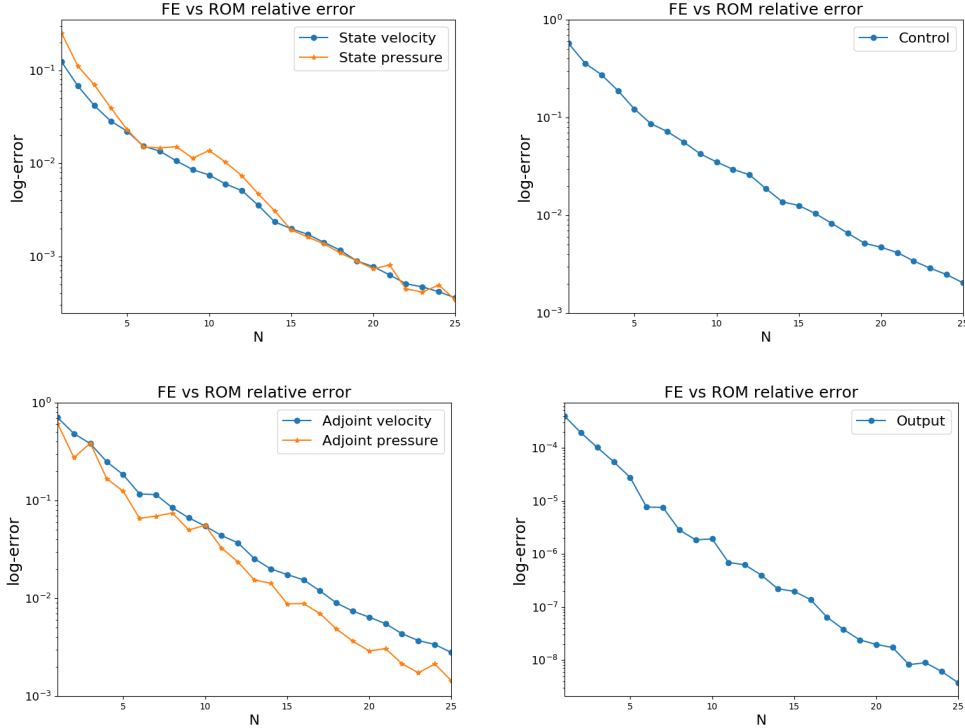
Figure 8: FE vs ROM errors in logarithmic scale. Top left: state error (velocity and pressure), top right: control error, bottom left: adjoint error (velocity and pressure), bottom right: output error.

7 Conclusions and Perspectives

In this work we have exploited ROM in OCP($\boldsymbol{\mu}$)s dealing with time dependent linear state equations. First of all we showed how the saddle point structure for linear quadratic OCP($\boldsymbol{\mu}$) is maintained also for the time dependent case. Then, we have exploited a POD-Galerkin method as sampling strategy for the projection of the systems in a low dimensional framework in order to solve them in an accurate and fast way. The results have been tested, on one side, for a boundary OCP($\boldsymbol{\mu}$) governed by a Graetz flow, on the other side, the test case has been performed over an OCP($\boldsymbol{\mu}$) constrained to time dependent Stokes equations. To the best of our knowledge, the main novelty of this work is in the POD-Galerkin reduction used for this all-at-once time dependent optimal control problems recast in saddle point formulation: it is a very versatile approach due to the great speedup index, for both the state equations analyzed.

Some improvements of this work will follow. First of all, we would like to stress that the used all-at-once approach is not the most appropriate. Indeed, it leads to a very costly formulation and complex linear systems to deal with. This issue can be overcome thanks to more computing power and/or preconditioned iterative strategies as proposed in [9, 48, 51, 52]. Some improvements are needed in this direction, since the direct solution of such a system affects not only the resolutions of full order solutions and, consequently, of the reduced ones, but also the time needed for the offline phase. Indeed, the proposed all-at-once formulation is the major reason for an expensive basis construction. Even if we accomplished our goal of showing how ROMs can be very effective in lowering the computational time needed for online parametrized simulations, we pay a huge amount of computational resources for the offline stage. We propose our methodology as a starting point which needs a deeper analysis and improvements in order to reach better performances: overcaming this issue will pave the way to applications of ROMs

to more complicated OCP($\boldsymbol{\mu}$)s.

Moreover, we are moving towards nonlinear state equations in order to recover a complete optimal control model that could have impact in several fields of applications. Time dependent nonlinear OCP($\boldsymbol{\mu}$)s could be a way to simulate more and more realistic physical phenomena. Reduced order modelling, most of all in its nonlinear time dependent formulation, could actually be a suitable and versatile approach to be used, in order to drastically reduce the computational costs in real time contexts. Another step forward could be adding a parameter stochastic dependency and uncertainty quantification in the optimality system. For sure, another improvement is based on the development of proper error estimators. The extension of classical estimates for steady linear quadratic OCP($\boldsymbol{\mu}$)s to time dependent state equations will be the topic of future investigations.

Acknowledgements

References

[1] RBniCS - reduced order modelling in FEniCS. http://mathlab.sissa.it/rbnics, 2015.

[2] S. Ali, F. Ballarin, and G. Rozza. Stabilized reduced basis methods for parametrized steady Stokes and Navier-Stokes equations, 2018.

[3] I. Babuška. Error-bounds for finite element method. *Numerische Mathematik*, 16(4):322–333, Jan 1971.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1âĂŞ137, 2005.

[10] P. B. Bochev and M. D. Gunzburger. *Least-squares finite element methods*, volume 166. Springer-Verlag, New York, 2009.

[11] D. Boffi, F. Brezzi, and M. Fortin. *Mixed finite element methods and applications*, volume 44. Springer-Verlag, Berlin and Heidelberg, 2013.

[12] 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.

[13] J. Burkardt, M. Gunzburger, and H.C. 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.

[14] 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.

[15] 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.

[16] L. Dedè. Optimal flow control for Navier-Stokes equations: Drag minimization. *International Journal for Numerical Methods in Fluids*, 55(4):347–366, 2007.

[17] 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.

[18] M. C. Delfour and J. Zolésio. *Shapes and geometries: metrics, analysis, differential calculus, and optimization*, volume 22. SIAM, Philadelphia, 2011.

[19] 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.

[20] 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.

[21] S. Glas, A. Mayerhofer, and K. Urban. *Two Ways to Treat Time in Reduced Basis Methods*, pages 1–16. Springer International Publishing, Cham, 2017.

[22] R. Guberovic, C. Schwab, and R. Stevenson. Space-time variational saddle point formulations of stokes and navierâĂŞstokes equations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 48(3):875–894, 2014.

[23] J. Haslinger and R. A. E. Mäkinen. *Introduction to shape optimization: theory, approximation, and computation*. SIAM, Philadelphia, 2003.

[24] J. S. Hesthaven, G. Rozza, and B. Stamm. Certified reduced basis methods for parametrized partial differential equations. *SpringerBriefs in Mathematics*, 2015, Springer, Milano.

[25] M.L. Hinze, M. Köster, and S. Turek. A hierarchical space-time solver for distributed control of the Stokes equation. Technical Report, SPP1253-16-01, 2008.

[26] M.L. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23. Springer Science & Business Media, Antwerp, 2008.

[27] L. Iapichino, S. Trenz, and S. Volkwein. Reduced-order multiobjective optimal control of semilinear parabolic problems. In Bülent Karasözen, Murat Manguoğlu, Münevver Tezer-Sezgin, Serdar Göktepe, and Ömür Uğur, editors, *Numerical Mathematics and Advanced Applications ENUMATH 2015*, pages 389–397, Cham, 2016. Springer International Publishing.

[28] L. Iapichino, S. Ulbrich, and S. Volkwein. Multiobjective pde-constrained optimization using the reduced-basis method. *Adv. Comput. Math.*, 43(5):945âĂŞ972, October 2017.

[29] 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.

[30] 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.

[31] K. Kunisch and S. Volkwein. Proper orthogonal decomposition for optimality systems. *ESAIM: Mathematical Modelling and Numerical Analysis*, 42(1):1–23, 2008.

[32] 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.

[33] G. Leugering, P. Benner, S. Engell, A. Griewank, H. Harbrecht, M. Hinze, R. Rannacher, and S. Ulbrich. *Trends in PDE constrained optimization*. Springer, New York, 2014.

[34] J. L. Lions. *Optimal Control of System Governed by Partial Differential Equations*, volume 170. Springer-Verlagr, Berlin and Heidelberg, 1971.

[35] A. Logg, K.A. Mardal, and G. Wells. *Automated Solution of Differential Equations by the Finite Element Method.* Springer-Verlag, Berlin, 2012.

[36] B. Mohammadi and O. Pironneau. *Applied shape optimization for fluids.* Oxford University Press, New York, 2010.

[37] 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.

[38] 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.

[39] M. Pošta and T. Roubíček. Optimal control of Navier–Stokes equations by Oseen approximation. *Computers & Mathematics With Applications*, 53(3):569–581, 2007.

[40] 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.

[41] A. Quarteroni, G. Rozza, L. Dedè, and A. Quaini. Numerical approximation of a control problem for advection-diffusion processes. In *IFIP Conference on System Modeling and Optimization*, pages 261–273, Ceragioli F., Dontchev A., Futura H., Marti K., Pandolfi L. (eds) System Modeling and Optimization. CSMO 2005. vol 199. Springer, Boston, 2005.

[42] A. Quarteroni, G. Rozza, and A. Quaini. Reduced basis methods for optimal control of advection-diffusion problems. In *Advances in Numerical Mathematics*, pages 193–216. RAS and University of Houston, Moscow, 2007.

[43] A. Quarteroni and A. Valli. *Numerical approximation of partial differential equations*, volume 23. Springer Science & Business Media, Berlin and Heidelberg, 2008.

[44] G. Rozza, D.B.P. Huynh, and A. Manzoni. Reduced basis approximation and a posteriori error estimation for Stokes flows in parametrized geometries: Roles of the inf-sup stability constants. *Numerische Mathematik*, 125(1):115–152, 2013.

[45] G. Rozza, D.B.P. Huynh, and A.T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: Application to transport and continuum mechanics. *Archives of Computational Methods in Engineering*, 15(3):229–275, 2008.

[46] 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.

[47] 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.

[48] 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.

[49] C. Schwab and R. Stevenson. Space-time adaptive wavelet methods for parabolic evolution problems. *Mathematics of Computation*, 78(267):1293–1318, 2009.

[50] Z. K. Seymen, H. Yücel, and B. Karasözen. Distributed optimal control of time-dependent diffusion–convection–reaction equations using space–time discretization. *Journal of Computational and Applied Mathematics*, 261:146–157, 2014.

[51] M. Stoll and A. Wathen. All-at-once solution of time-dependent PDE-constrained optimization problems. 2010.

[52] M. Stoll and A. Wathen. All-at-once solution of time-dependent Stokes control. *J. Comput. Phys.*, 232(1):498–515, January 2013.

[53] 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.

[54] 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. Submitted, 2019.

[55] F. Tröltzsch. Optimal control of partial differential equations. *Graduate studies in mathematics*, 112, Verlag, Wiesbad, 2010.

[56] 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.

[57] 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.

[58] 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.

[59] 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.

[60] Z. Zainib, F. Ballarin, G. Rozza, P. Triverio, L. Jiménez-Juan, and S. Fremes. Reduced order methods for parametric optimal flow control in coronary bypass grafts, towards patient-specific data assimilation. Submitted, 2019.