A DISCONTINUOUS GALERKIN TIME-STEPPING SCHEME FOR THE VELOCITY TRACKING PROBLEM*

EDUARDO CASAS† AND KONSTANTINOS CHRYSAFINOS‡

Abstract. The velocity tracking problem for the evolutionary Navier-Stokes equations in 2d is studied. The controls are of distributed type and they are submitted to bound constraints. First and second order necessary and sufficient conditions are proved. A fully-discrete scheme based on discontinuous (in time) Galerkin approach combined with conforming finite element subspaces in space, is proposed and analyzed. Provided that the time and space discretization parameters, τ and h respectively, satisfy $\tau \leq Ch^2$, then L^2 error estimates of order $O(h)$ are proved for the difference between the locally optimal controls and their discrete approximations.

Key words. evolution Navier-Stokes equations, optimal control, a priori error estimates, discontinuous Galerkin methods

AMS subject classifications. 49J20, 49K20, 35K55, 65N30

1. Introduction. In this paper we prove some error estimates for the numerical approximation of a distributed optimal control problem governed by the evolution Navier–Stokes equations, with pointwise control constraints. More precisely, we consider the following problem:

$$\text{(P)} \quad \begin{cases} \min J(\mathbf{u}) \\ \mathbf{u} \in \mathcal{U}_{ad} \end{cases}$$

where

$$\begin{aligned} J(\mathbf{u}) &= \frac{1}{2}\int_0^T \int_\Omega |\mathbf{y}_\mathbf{u}(t,x) - \mathbf{y}_d(t,x)|^2 \,\mathrm{d}x\mathrm{d}t + \frac{\gamma}{2}\int_\Omega |\mathbf{y}_\mathbf{u}(T,x) - \mathbf{y}_\Omega(x)|^2 \,\mathrm{d}x \\ &+ \frac{\lambda}{2}\int_0^T \int_\Omega |\mathbf{u}(t,x)|^2 \,\mathrm{d}x\mathrm{d}t. \end{aligned}$$

Here $\mathbf{y}_\mathbf{u}$ denotes the solution of the 2d evolution Navier-Stokes equations

$$\begin{cases} \mathbf{y}_t - \nu\Delta\mathbf{y} + (\mathbf{y}\cdot\nabla)\mathbf{y} + \nabla p = \mathbf{f} + \mathbf{u} \ \text{ in } \Omega_T = (0,T)\times\Omega, \\ \operatorname{div}\mathbf{y} = 0 \ \text{ in } \Omega_T, \ \mathbf{y}(0) = \mathbf{y}_0 \ \text{ in } \Omega, \ \mathbf{y} = 0 \ \text{ on } \Sigma_T = (0,T)\times\Gamma, \end{cases} \tag{1.1}$$

and $\mathcal{U}_{ad}$ is the set of feasible controls

$$\mathcal{U}_{ad} = \{\mathbf{u} \in L^2(0,T;\mathbf{L}^2(\Omega)) : \alpha_j \leq u_j(t,x) \leq \beta_j \ \text{ a.e. } (t,x) \in \Omega_T, \ j = 1,2\}$$

*The first author was partially supported by the Spanish Ministry of Science and Innovation under projects MTM2008-04206 and "Ingenio Mathematica (i-MATH)" CSD2006-00032 (Consolider Ingenio 2010).

†Departmento de Matemática Aplicada y Ciencias de la Computación, E.T.S.I. Industriales y de Telecomunicación, Universidad de Cantabria, Av. Los Castros s/n, 39005 Santander, Spain (eduardo.casas@unican.es).

‡Department of Mathematics, School of Applied Mathematics and Physical Sciences, National Technical University of Athens, Zografou Campus, Athens 15780, Greece (chrysafinos@math.ntua.gr)

where $-\infty \le \alpha_j < \beta_j \le +\infty$, $j = 1, 2$. Hereinafter, we will denote $\alpha = (\alpha_1, \alpha_2)$ and $\beta = (\beta_1, \beta_2)$.

The scope of the above optimal control problem is to match the velocity vector field to a given target field, by influencing the behavior of the system through a control function. The control function is of distributed type and satisfies certain constraints. This is achieved by minimizing the standard tracking type functional, while the parameter $\lambda > 0$ denotes a penalty parameter, which is typically small compared to the actual size of the data. The terminal term has been included in order to obtain more effective approximations near the end point of the time interval. For related discussion and references regarding the computational significance of the above optimal control problem we refer the reader to [13].

The analysis of such optimal control problems is well understood. However, when it comes to the approximation and to the numerical analysis of such problems the existing literature is quite limited. This is due to the fact that the regularity of solutions of Navier-Stokes equations, within the optimal setting is very limited, which creates additional difficulties in analyzing suitable schemes for optimal control problems. Standard techniques developed for the numerical analysis of the uncontrolled Navier-Stokes equations can not be directly applied in the optimal control setting. In addition, optimal control problems constrained to nonlinear evolutionary pdes with control constraints typically exhibit fine properties and hence require special techniques involving both first and second order necessary and sufficient conditions.

Our work analyzes a numerical scheme based on the discontinuous time-stepping Galerkin scheme for the piecewise constant time combined with standard conforming finite element subspaces for the discretization in space. The main result of our work is to derive space-time error estimates, under suitable regularity assumptions on the data by utilizing ideas from [4] developed for the stationary Navier-Stokes, together with a detailed error analysis of the uncontrolled state and adjoint equations of the underlying scheme. A key part of our work is to show that the discrete problem possesses similar regularity properties to the continuous one. To our best knowledge our estimates are new. For some related earlier work, we refer the reader to [1], [14], [15], [16], [17], [24], [25], [29] and the reference cited therein.

The discontinuous Galerkin time-stepping schemes are known to perform well in a variety of problems whose solutions satisfy low regularity properties. For earlier work on discontinuous time-stepping schemes within the context of optimal control problems we refer the reader to [21], [22] for error estimates for an optimal control problem for the heat equation, with and without control constraints respectively, and to [6] for a convergence result for a semilinear parabolic optimal control problem. For general results related to discontinuous time step schemes for linear parabolic uncontrolled pdes, we refer the reader to [8, 9, 10, 11, 27] (see also references within). Finally in the recent work of [7], discontinuous time-stepping schemes of arbitrary order for the Navier-Stokes equations in 2d and 3d where examined. Further results concerning the analysis and numerical analysis of the uncontrolled Navier-Stokes can be found in the classical works of [12], [18], [19], [26]. For several issues related to the analysis and numerics of optimal control problems we refer the reader to [28] (see also references within).

2. Assumptions and preliminary results. Ω is a bounded open and convex subset in $\mathbb{R}^2$, Γ being its boundary. The outward unit normal vector to Γ at a point $x \in \Gamma$ is denoted by $\mathbf{n}(x)$. Given $0 < T < +\infty$, we denote $\Omega_T = (0, T) \times \Omega$ and

$\Sigma_T = (0,T)\times\Gamma$. Let us introduce some function spaces and operators. Throughout the following lines we fix the notation for Sobolev spaces: $\mathbf{H}^1(\Omega) = H^1(\Omega;\mathbb{R}^2)$, $\mathbf{H}_0^1(\Omega) = H_0^1(\Omega;\mathbb{R}^2)$, $\mathbf{H}^{-1}(\Omega) = (\mathbf{H}_0^1(\Omega))'$ and $\mathbf{W}^{s,p}(\Omega) = W^{s,p}(\Omega;\mathbb{R}^2)$ for $1 \le p \le \infty$ and $s > 0$. We also consider the spaces of integrable functions

$$L_0^2(\Omega) = \{w \in L^2(\Omega) : \int_\Omega w(x)\,\mathrm{d}x = 0\};$$

$\mathbf{L}^p(\Omega) = L^p(\Omega;\mathbb{R}^2)$ and, for a given Banach space X, $L^p(0,T;X)$ will denote the integrable functions defined in $(0,T)$ and taking values in X endowed with the usual norm. Following Lions and Magenes [20, Vol. 1] we put

$$H^{2,1}(\Omega_T) = \left\{ y \in L^2(\Omega_T) : \frac{\partial y}{\partial x_i}, \frac{\partial^2 y}{\partial x_i x_j}, \frac{\partial y}{\partial t} \in L^2(\Omega_T),\ 1 \le i,j \le 2 \right\}$$

and

$$\|y\|_{H^{2,1}(\Omega_T)} = \left\{ \int_{\Omega_T} \left(|y|^2 + \left|\frac{\partial y}{\partial t}\right|^2 \right) \mathrm{d}x\mathrm{d}t + \sum_{i=1}^{2} \int_{\Omega_T} \left|\frac{\partial y}{\partial x_i}\right|^2 \mathrm{d}x\mathrm{d}t + \sum_{i,j=1}^{2} \int_{\Omega_T} \left|\frac{\partial^2 y}{\partial x_i x_j}\right|^2 \mathrm{d}x\mathrm{d}t \right\}^{1/2}.$$

In [20, Vol. 1] it is proved that every element of $H^{2,1}(\Omega_T)$, after a modification over a zero measure set, is a continuous function from $[0,T] \longrightarrow H^1(\Omega)$. We also set $\mathbf{H}^{2,1}(\Omega_T) = H^{2,1}(\Omega_T) \times H^{2,1}(\Omega_T)$.

We introduce the usual spaces of divergence-free vector fields:

$$\mathbf{Y} = \{\mathbf{y} \in \mathbf{H}_0^1(\Omega) : \operatorname{div}\mathbf{y} = 0 \text{ in } \Omega\}$$

$$\mathbf{H} = \{\mathbf{y} \in \mathbf{L}^2(\Omega) : \operatorname{div}\mathbf{y} = 0 \text{ in } \Omega \ \text{ and } \ \mathbf{y}\cdot\mathbf{n} = 0 \text{ on } \Gamma\}.$$

Along this paper, we will assume that $\mathbf{f}, \mathbf{u} \in L^2(0,T;\mathbf{L}^2(\Omega))$ and $\mathbf{y}_0 \in \mathbf{Y}$. A solution of (1.1) will be sought in the space

$$\mathbf{W}(0,T) = \{\mathbf{y} \in L^2(0,T;\mathbf{Y}) : \mathbf{y}_t \in L^2(0,T;\mathbf{Y}^*)\}.$$

It is well known that $\mathbf{W}(0,T) \subset C_w([0,T],\mathbf{H})$, where $C_w([0,T],\mathbf{H})$ is the space of weakly continuous functions $\mathbf{y} : [0,T] \longrightarrow \mathbf{H}$.

It is well known that (2.1) has a unique solution in $\mathbf{W}(0,T)$. Once the velocity $\mathbf{y}$ is obtained, then the existence of a pressure $p \in \mathcal{D}(\Omega_T)$ is proved in such a way that the first equation of (1.1) holds in a distribution sense. Thanks to the regularity assumed on $\mathbf{f}$, $\mathbf{y}_0$ and Ω, then some extra regularity is proved for $(\mathbf{y}, p)$. Indeed, we have that $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T],\mathbf{Y})$ and $p \in L^2(0,T;H^1(\Omega))$, the pressure being unique up to an additive constant; see, for instance, Ladyzhenskaya [18], Lions [19], Temam [26].

Let us introduce the weak formulation of (1.1). To this end we define the bilinear and trilinear forms $a : \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \longrightarrow \mathbb{R}$ and $c : \mathbf{L}^4(\Omega) \times \mathbf{H}^1(\Omega) \times \mathbf{H}^1(\Omega) \longrightarrow \mathbb{R}$ by

$$a(\mathbf{y},\mathbf{z}) = \nu \int_\Omega (\nabla\mathbf{y} : \nabla\mathbf{z})\,\mathrm{d}x = \nu \sum_{i,j=1}^{2} \int_\Omega \partial_{x_i} y_j \, \partial_{x_i} z_j \,\mathrm{d}x$$

and

$$c(\mathbf{y},\mathbf{z},\mathbf{w}) = \frac{1}{2}\left[\hat{c}(\mathbf{y},\mathbf{z},\mathbf{w}) - \hat{c}(\mathbf{y},\mathbf{w},\mathbf{z})\right] \text{ with } \hat{c}(\mathbf{y},\mathbf{z},\mathbf{w}) = \sum_{i,j=1}^{2}\int_\Omega \mathbf{y}_j\left(\frac{\partial \mathbf{z}_i}{\partial x_j}\right)\mathbf{w}_i\,\mathrm{d}x.$$

The following weak formulation is frequently used: we seek $\mathbf{y} \in \mathbf{W}(0,T)$ such that for a.e. $t \in (0,T)$,

$$\left\{\begin{array}{l}(\mathbf{y}_t,\mathbf{w}) + a(\mathbf{y},\mathbf{w}) + c(\mathbf{y},\mathbf{y},\mathbf{w}) = (\mathbf{f}+\mathbf{u},\mathbf{w}) \quad \forall\, \mathbf{w} \in \mathbf{Y}\\ \mathbf{y}(0) = \mathbf{y}_0.\end{array}\right. \tag{2.1}$$

Above $(\cdot\,,\cdot)$ denotes the scalar product in $\mathbf{L}^2(\Omega)$. This notation will be frequently used along the paper and $\|\cdot\|$ will denote the associated norm. Any other norm will be indicated by a subscript.

Returning back to the control problem (P), we note that the mapping associating to each control $\mathbf{u} \in L^2(0,T;\mathbf{L}^2(\Omega))$ the corresponding state $\mathbf{y}_\mathbf{u} \in \mathbf{H}^{2,1}(\Omega_T)\cap C([0,T],\mathbf{Y})$ is well defined, and continuous. Therefore, by assuming that $\mathbf{y}_d \in L^2(0,T;\mathbf{L}^2(\Omega))$ and $\mathbf{y}_\Omega \in \mathbf{L}^2(\Omega)$, we have that the cost functional $J : L^2(0,T;\mathbf{L}^2(\Omega)) \longrightarrow \mathbb{R}$ is well defined and continuous. Since the problem (P) is not convex, we will deal in the next sections with global and local solutions. A control $\bar{\mathbf{u}} \in \mathcal{U}_{ad}$ is said a local solution of (P) if there exists $\varepsilon > 0$ such that $J(\bar{\mathbf{u}}) \le J(\mathbf{u})$ for every $\mathbf{u} \in \mathcal{U}_{ad} \cap B_\varepsilon(\bar{\mathbf{u}})$, where $B_\varepsilon(\bar{\mathbf{u}})$ denote the open ball of $L^2(0,T;\mathbf{L}^2(\Omega))$ centered at $\bar{\mathbf{u}}$ and radius ε. The following regularity assumption will be assumed for the data defining J:

$$\lambda > 0, \quad \gamma \ge 0, \quad \mathbf{y}_d \in L^2(0,T;\mathbf{L}^2(\Omega)) \text{ and } \mathbf{y}_\Omega \in Y. \tag{2.2}$$

Under these assumptions, the proof of the existence of at least one solution of (P) is standard.

Before finishing this section, let us state some properties of the trilinear form c that we will use later. The proof can be found in many books; see [18], [19] or [26].

Lemma 2.1. *The trilinear form satisfies*

$$\begin{aligned}&c(\mathbf{y},\mathbf{w},\mathbf{z}) = \hat{c}(\mathbf{y},\mathbf{z},\mathbf{w}) = -\hat{c}(\mathbf{y},\mathbf{w},\mathbf{z}) \quad \forall \mathbf{y} \in \mathbf{Y} \text{ and } \forall \mathbf{z},\mathbf{w} \in \mathbf{H}_0^1(\Omega),\\ &c(\mathbf{y},\mathbf{z},\mathbf{w}) = -c(\mathbf{y},\mathbf{w},\mathbf{z}) \quad \forall \mathbf{y} \in \mathbf{L}^4(\Omega) \text{ and } \forall \mathbf{z},\mathbf{w} \in \mathbf{H}^1(\Omega),\\ &c(\mathbf{y},\mathbf{w},\mathbf{w}) = 0 \quad \forall \mathbf{y} \in \mathbf{L}^4(\Omega) \text{ and } \forall \mathbf{w} \in \mathbf{H}^1(\Omega).\end{aligned}$$

Moreover, the following inequalities hold

$$\begin{aligned}&|c(\mathbf{y},\mathbf{z},\mathbf{w})| \le \|\mathbf{y}\|_{\mathbf{L}^p(\Omega)}\|\nabla\mathbf{z}\|_{\mathbf{L}^2(\Omega)}\|\mathbf{w}\|_{\mathbf{L}^q(\Omega)}, \quad (1/p)+(1/q) = (1/2),\\ &|c(\mathbf{y},\mathbf{z},\mathbf{w})| \le \|\mathbf{y}\|_{\mathbf{L}^4(\Omega)}\|\nabla\mathbf{z}\|_{\mathbf{L}^2(\Omega)}\|\mathbf{w}\|_{\mathbf{L}^4(\Omega)}.\end{aligned}$$

By using the interpolation inequality

$$\|z\|_{L^4(\Omega)} \le 2^{1/4}\|z\|_{L^2(\Omega)}^{1/2}\|\nabla z\|_{L^2(\Omega)}^{1/2} \quad \forall z \in H_0^1(\Omega), \tag{2.3}$$

(see [26, Lema 3.3, page 91]) we obtain $\forall \mathbf{y},\mathbf{w} \in \mathbf{H}_0^1(\Omega)$ and $\forall \mathbf{z} \in \mathbf{H}^1(\Omega)$

$$|c(\mathbf{y},\mathbf{z},\mathbf{w})| \le C\|\mathbf{y}\|_{\mathbf{L}^2(\Omega)}^{1/2}\|\nabla\mathbf{y}\|_{\mathbf{L}^2(\Omega)}^{1/2}\|\nabla\mathbf{z}\|_{\mathbf{L}^2(\Omega)}\|\mathbf{w}\|_{\mathbf{L}^2(\Omega)}^{1/2}\|\nabla\mathbf{w}\|_{\mathbf{L}^2(\Omega)}^{1/2}. \tag{2.4}$$

3. Optimality conditions. In this section, we are going to prove first- and second-order optimality conditions for a local solution of problem (P) . To this end, we begin analyzing the differentiability of the control-to-state mapping. We denote by $G : L^2(0,T;\mathbf{L}^2(\Omega)) \to \mathbf{H}^{2,1}(\Omega_T)\cap C([0,T];\mathbf{Y})$ the mapping associating to each control $\mathbf{u}$ the corresponding state $G(\mathbf{u}) = \mathbf{y_u}$ solution of (2.1). The next theorem establishes the differentiability of G and provide the first and second derivatives, which are crucial to derive the optimality conditions.

THEOREM 3.1. *The mapping* $G : L^2(0,T;\mathbf{L}^2(\Omega)) \to \mathbf{H}^{2,1}(\Omega_T)\cap C([0,T];\mathbf{Y})$ *is of class* C^∞. *Moreover, for any* $\mathbf{u},\mathbf{v} \in L^2(0,T;\mathbf{L}^2(\Omega))$, *if we denote* $\mathbf{y_u} = G(\mathbf{u})$, $\mathbf{z_v} = G'(\mathbf{u})\mathbf{v}$ *and* $\mathbf{z_{vv}} = G''(\mathbf{u})\mathbf{v}^2$, *then* $\mathbf{z_v}$ *and* $\mathbf{z_{vv}}$ *are the unique solutions of the following equations* $\forall \mathbf{w}\in\mathbf{Y}$

$$\left\{\begin{array}{l}(\mathbf{z}_{\mathbf{v},t},\mathbf{w}) + a(\mathbf{z_v},\mathbf{w}) + c(\mathbf{z_v},\mathbf{y_u},\mathbf{w}) + c(\mathbf{y_u},\mathbf{z_v},\mathbf{w}) = (\mathbf{v},\mathbf{w})\\ \mathbf{z_v}(0) = 0,\end{array}\right. \tag{3.1}$$

$$\left\{\begin{array}{l}(\mathbf{z}_{\mathbf{vv},t},\mathbf{w}) + a(\mathbf{z_{vv}},\mathbf{w}) + c(\mathbf{z_{vv}},\mathbf{y_u},\mathbf{w}) + c(\mathbf{y_u},\mathbf{z_{vv}},\mathbf{w}) + 2c(\mathbf{z_v},\mathbf{z_v},\mathbf{w}) = 0\\ \mathbf{z_{vv}}(0) = 0.\end{array}\right. \tag{3.2}$$

The proof of this theorem is based on the following result.

THEOREM 3.2 (Casas [3]). *Let* $\mathbf{f} \in L^2(0,T;\mathbf{L}^2(\Omega))$, $\mathbf{g} \in L^8(0,T;\mathbf{L}^4(\Omega))$, *with* $\operatorname{div}\mathbf{g} = 0$ *in* Ω_T, $\mathbf{e} \in L^\infty(0,T;\mathbf{Y})\cap L^{3/2}(0,T;\mathbf{H}^2(\Omega))$ *and* $\mathbf{z}_0 \in \mathbf{Y}$. *Then there exists a unique element* $\mathbf{z} \in \mathbf{H}^{2,1}(\Omega_T)\cap C([0,T],\mathbf{Y})$ *solution of the following problem*

$$\left\{\begin{array}{l}(\mathbf{z}_t,\mathbf{w}) + a(\mathbf{z},\mathbf{w}) + c(\mathbf{g},\mathbf{z},\mathbf{w}) + c(\mathbf{z},\mathbf{e},\mathbf{w}) = (\mathbf{f},\mathbf{w}) \ \forall \mathbf{w}\in\mathbf{Y}\\ \mathbf{z}(0) = \mathbf{z}_0.\end{array}\right. \tag{3.3}$$

Moreover, there exists an increasing function $\eta : [0,+\infty) \longrightarrow [0,+\infty)$ *depending only on* Ω *and* ν *such that*

$$\begin{aligned}\|\mathbf{z}\|_{\mathbf{H}^{2,1}(\Omega_T)} \le \eta\big(&\|\mathbf{z}_0\|_{\mathbf{Y}} + \|\mathbf{f}\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + \|\mathbf{g}\|_{L^8(0,T;\mathbf{L}^4(\Omega))}\\ &+ \|\mathbf{e}\|_{L^\infty(0,T;\mathbf{Y})} + \|\mathbf{e}\|_{L^{3/2}(0,T;\mathbf{H}^2(\Omega))}\big).\end{aligned} \tag{3.4}$$

Proof of Theorem 3.1. Let us denote by

$$F : \left[\mathbf{H}^{2,1}(\Omega_T)\cap C([0,T],\mathbf{Y})\right] \times L^2(0,T;\mathbf{L}^2(\Omega)) \longrightarrow L^2(0,T;\mathbf{L}^2(\Omega)) \times \mathbf{Y}$$

the mapping given by

$$F(\mathbf{y},\mathbf{u}) = (\mathbf{y}_t - \nu\Delta\mathbf{y} + B\mathbf{y} - \mathbf{f} - \mathbf{u}, \mathbf{y}(0) - \mathbf{y}_0),$$

where $B : \mathbf{Y} \longrightarrow \mathbf{Y}^*$ is defined as follows

$$\langle B\mathbf{y},\mathbf{w}\rangle = c(\mathbf{y},\mathbf{y},\mathbf{w}). \tag{3.5}$$

Then, B is well defined and continuous; recall Lemma 2.1. Moreover, for every $\mathbf{y} \in \mathbf{H}^2(\Omega)\cap\mathbf{Y}$ we have

$$\begin{aligned}|\langle B\mathbf{y},\mathbf{w}\rangle| &\le C\|\mathbf{y}\|_{\mathbf{L}^4(\Omega)}\|\mathbf{y}\|_{\mathbf{W}^{1,4}(\Omega)}\|\mathbf{w}\|_{\mathbf{L}^2(\Omega)}\\ &\le C\|\mathbf{y}\|_{\mathbf{Y}}\|\mathbf{y}\|_{\mathbf{H}^2(\Omega)}\|\mathbf{w}\|_{\mathbf{L}^2(\Omega)}.\end{aligned}$$

Thus we have that $B\mathbf{y} \in L^2(0,T;\mathbf{L}^2(\Omega))$ for every $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap \mathbf{Y}$.

Also it is immediate to check that F is of class C^∞ and

$$\frac{\partial F}{\partial \mathbf{y}}(\mathbf{y},\mathbf{u}) \cdot \mathbf{z} = (\mathbf{z}_t - \nu\Delta\mathbf{z} + B'(\mathbf{y})\mathbf{z}, \mathbf{z}(0))\,,$$

where

$$\langle B'(\mathbf{y})\mathbf{z}, \mathbf{w}\rangle = c(\mathbf{y},\mathbf{z},\mathbf{w}) + c(\mathbf{z},\mathbf{y},\mathbf{w}) \quad \forall \mathbf{w} \in \mathbf{Y}.$$

Applying Theorem 3.2, with $\mathbf{g} = \mathbf{e} = \mathbf{y}$, we deduce that $\dfrac{\partial F}{\partial \mathbf{y}}(\mathbf{y},\mathbf{u})$ is an isomorphism from $\mathbf{H}^{2,1}(\Omega_T) \cap C([0,T],\mathbf{Y})$ onto $L^2(0,T,\mathbf{L}^2(\Omega)) \times \mathbf{Y}$. Now if $\mathbf{y_u}$ is the solution of (2.1), then $F(\mathbf{y_u},\mathbf{u}) = (0,0)$. Therefore, we can apply the implicit function theorem to deduce that $G : L^2(0,T;\mathbf{L}^2(\Omega)) \longrightarrow \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T],\mathbf{Y})$ is C^∞. Moreover, from the identity $F(\mathbf{y_u},\mathbf{u}) = F(G(\mathbf{u}),\mathbf{u}) = (0,0)$ for every $\mathbf{u}$, we get for all $\forall \mathbf{v} \in L^2(0,T;\mathbf{L}^2(\Omega))$

$$\frac{\partial F}{\partial \mathbf{y}}(\mathbf{y_u},\mathbf{u})G'(\mathbf{u})\mathbf{v} + \frac{\partial F}{\partial \mathbf{u}}(\mathbf{y_u},\mathbf{u})\mathbf{v} = (0,0)$$

and

$$\begin{aligned}&\frac{\partial F}{\partial \mathbf{y}}(\mathbf{y_u},\mathbf{u})G''(\mathbf{u})\mathbf{v}^2 + \frac{\partial^2 F}{\partial \mathbf{y}^2}(\mathbf{y_u},\mathbf{u})(G'(\mathbf{u})\mathbf{v}, G'(\mathbf{u})\mathbf{v})\\ &+2\frac{\partial^2 F}{\partial \mathbf{y}\partial \mathbf{u}}(\mathbf{y_u},\mathbf{u})(G'(\mathbf{u})\mathbf{v},\mathbf{v}) + \frac{\partial^2 F}{\partial \mathbf{u}^2}(\mathbf{y_u},\mathbf{u})\mathbf{v}^2 = (0,0).\end{aligned}$$

Then setting $\mathbf{z_v} = G'(\mathbf{u}) \cdot \mathbf{v}$ and $\mathbf{z_{vv}} = G''(\mathbf{u})\mathbf{v}^2$, we obtain (3.1) and (3.2) from the above two identities. □

As a consequence of Theorem 3.1 we get the differentiability of the cost functional J.

THEOREM 3.3. *The cost functional* $J : L^2(0,T;\mathbf{L}^2(\Omega)) \longrightarrow \mathbb{R}$ *is of class* C^∞ *and for every* $\mathbf{u},\mathbf{v} \in L^2(0,T;\mathbf{L}^2(\Omega))$ *we have*

$$J'(\mathbf{u})\mathbf{v} = \int_0^T \int_\Omega (\boldsymbol{\varphi}_\mathbf{u} + \lambda\mathbf{u})\mathbf{v}\,\mathrm{d}x\mathrm{d}t \tag{3.6}$$

and

$$J''(\mathbf{u})\mathbf{v}^2 = \int_0^T \int_\Omega (|\mathbf{z_v}|^2 - 2(\mathbf{z_v}\cdot\nabla)\mathbf{z_v}\boldsymbol{\varphi}_\mathbf{u})\mathrm{d}x\mathrm{d}t + \gamma\int_\Omega |\mathbf{z_v}(T)|^2\,\mathrm{d}x + \lambda\int_0^T\int_\Omega |\mathbf{v}|^2\mathrm{d}x\mathrm{d}t, \tag{3.7}$$

where $\mathbf{z_v} = G'(\mathbf{u})\mathbf{v}$ *is the solution of* (3.1) *and* $\boldsymbol{\varphi}_\mathbf{u} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T],\mathbf{Y})$ *is the unique element satisfying for every* $\mathbf{w} \in \mathbf{Y}$

$$\left\{\begin{array}{l} -(\boldsymbol{\varphi}_{\mathbf{u},t},\mathbf{w}) + a(\boldsymbol{\varphi}_\mathbf{u},\mathbf{w}) + c(\mathbf{w},\mathbf{y_u},\boldsymbol{\varphi}_\mathbf{u}) + c(\mathbf{y_u},\mathbf{w},\boldsymbol{\varphi}_\mathbf{u}) = (\mathbf{y_u} - \mathbf{y}_d,\mathbf{w}),\\ \boldsymbol{\varphi}_\mathbf{u}(T) = \gamma(\mathbf{y}(T) - \mathbf{y}_\Omega). \end{array}\right. \tag{3.8}$$

Proof. First of all, let us observe that the equation (3.8) is the adjoint of (3.1). Since (3.1) has a unique solution in $\mathbf{H}^{2,1}(\Omega_T) \cap C([0,T],\mathbf{Y})$ for any $\mathbf{v} \in L^2(0,T;\mathbf{L}^2(\Omega))$, then arguing by transposition we can prove the existence and uniqueness of the solution $\boldsymbol{\varphi}_{\mathbf{u}}$ of (3.8), as well as the regularity $\boldsymbol{\varphi}_{\mathbf{u}} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T],\mathbf{Y})$. Now, the differentiability property of J is a consequence of Theorem 3.1 and the chain rule. Moreover, we have

$$J'(\mathbf{u})\mathbf{v} = \int_0^T \int_\Omega (\mathbf{y_u} - \mathbf{y}_d)\mathbf{z_v}\,\mathrm{d}x\mathrm{d}t + \gamma \int_\Omega (\mathbf{y_u}(T) - \mathbf{y}_\Omega) z_{\mathbf{v}}(T)\,\mathrm{d}x + \lambda \int_0^T \int_\Omega \mathbf{u}\cdot\mathbf{v}\,\mathrm{d}x\mathrm{d}t$$

and

$$J''(\mathbf{u})\mathbf{v}^2 = \int_0^T \int_\Omega [|\mathbf{z_v}|^2 + (\mathbf{y_u} - \mathbf{y}_d)\mathbf{z_{vv}}]\,\mathrm{d}x\mathrm{d}t$$
$$+\gamma \int_\Omega [|\mathbf{z_v}(T)|^2 + (\mathbf{y_u}(T) - \mathbf{y}_\Omega)\mathbf{z_{vv}}(T)]\,\mathrm{d}x + \lambda \int_0^T \int_\Omega |\mathbf{v}|^2\,\mathrm{d}x\mathrm{d}t,$$

where $\mathbf{z_v}$ and $\mathbf{z_{vv}}$ are defined by (3.1) and (3.2). Then, using (3.8) and making an integration by parts we get (3.6) and (3.7) from the above relations. □

We are now able to prove the optimality conditions. We start with the first-order necessary conditions.

THEOREM 3.4. *Let us assume that* $\bar{\mathbf{u}}$ *is a local solution of problem* (P), *then there exist* $\bar{\mathbf{y}}$ *and* $\bar{\boldsymbol{\varphi}}$ *belonging to* $\mathbf{H}^{2,1}(\Omega_T) \cap C([0,T],\mathbf{Y})$ *such that*

$$\text{(3.9)} \qquad \begin{cases} (\bar{\mathbf{y}}_t, \mathbf{w}) + a(\bar{\mathbf{y}}, \mathbf{w}) + c(\bar{\mathbf{y}}, \bar{\mathbf{y}}, \mathbf{w}) = (\mathbf{f} + \bar{\mathbf{u}}, \mathbf{w}) \quad \forall\, \mathbf{w} \in \mathbf{Y}, \\ \bar{\mathbf{y}}(0) = \mathbf{y}_0, \end{cases}$$

$$\text{(3.10)} \qquad \begin{cases} -(\bar{\boldsymbol{\varphi}}_t, \mathbf{w}) + a(\bar{\boldsymbol{\varphi}}, \mathbf{w}) + c(\mathbf{w}, \bar{\mathbf{y}}, \bar{\boldsymbol{\varphi}}) + c(\bar{\mathbf{y}}, \mathbf{w}, \bar{\boldsymbol{\varphi}}) = (\bar{\mathbf{y}}_{\mathbf{u}} - \mathbf{y}_d, \mathbf{w}) \ \forall \mathbf{w} \in \mathbf{Y}, \\ \bar{\boldsymbol{\varphi}}(T) = \gamma(\bar{\mathbf{y}}(T) - \mathbf{y}_\Omega), \end{cases}$$

$$\text{(3.11)} \qquad \int_0^T \int_\Omega (\bar{\boldsymbol{\varphi}} + \lambda\bar{\mathbf{u}})(\mathbf{u} - \bar{\mathbf{u}})\,\mathrm{d}x\mathrm{d}t \geq 0 \quad \forall \mathbf{u} \in \mathcal{U}_{ad}.$$

Moreover, the regularity property $\bar{\mathbf{u}} \in \mathbf{H}^1(\Omega_T) \cap C([0,T],\mathbf{H}^1(\Omega)) \cap L^2(0,T;\mathbf{W}^{1,p}(\Omega))$ *holds for all* $1 \leq p < +\infty$.

Proof. Since $\mathcal{U}_{ad}$ is convex, any local solution $\bar{\mathbf{u}}$ satisfies the condition $J'(\bar{\mathbf{u}})(\mathbf{u}-\bar{\mathbf{u}}) \geq 0$ for every $\mathbf{u} \in \mathcal{U}_{ad}$. Then, it is enough to use the expression of the derivative given by (3.6) and take $\bar{\mathbf{y}} = \mathbf{y}_{\bar{\mathbf{u}}}$ and $\bar{\boldsymbol{\varphi}} = \boldsymbol{\varphi}_{\bar{\mathbf{u}}}$ to deduce (3.9)-(3.11). The regularity of $\bar{\mathbf{u}}$ follows from (3.11) as usual, we simply observe that (3.11) implies that

$$\text{(3.12)} \qquad \bar{u}_j(t,x) = \mathrm{Proj}_{[\alpha_j,\beta_j]}\left(-\frac{1}{\lambda}\bar{\varphi}_j(t,x)\right) \quad \text{for a.a. } (t,x) \in \Omega_T, \quad j = 1,2,$$

which leads to the desired regularity of $\bar{u}_j$, $j = 1,2$. □

To write the second-order conditions we need to define the cone of critical directions. To this end, let us introduce the function

$$\text{(3.13)} \qquad \bar{\mathbf{d}} = \bar{\boldsymbol{\varphi}} + \lambda\bar{\mathbf{u}}.$$

Now we set

$$\text{(3.14)} \qquad \mathcal{C}_{\bar{\mathbf{u}}} = \{\mathbf{v} \in L^2(0,T;\mathbf{L}^2(\Omega)) : \mathbf{v} \text{ satisfies } (3.15) - (3.17)\},$$

$$v_j(t,x) \geq 0 \text{ if } -\infty < \alpha_j = \bar u_j(t,x), \tag{3.15}$$
$$v_j(t,x) \leq 0 \text{ if } \bar u_j(t,x) = \beta_j < +\infty, \qquad j = 1,2, \tag{3.16}$$
$$v_j(t,x) = 0 \text{ if } \bar d_j(t,x) \neq 0. \tag{3.17}$$

Let us notice that

$$\begin{aligned} &J'(\bar{\mathbf{u}})\mathbf{v} = \int_0^T\int_\Omega \bar{\mathbf{d}}(t,x)\cdot\mathbf{v}(t,x)\,\mathrm{d}x\mathrm{d}t,\\ &\bar{\mathbf{d}}(t,x)\cdot\mathbf{v}(t,x) = 0 \text{ for a.a. } (t,x)\in\Omega_T \text{ and } \forall \mathbf{v}\in\mathcal{C}_{\bar{\mathbf{u}}}. \end{aligned} \tag{3.18}$$

We also deduce as usual from (3.11)

$$\begin{cases} \bar u_j(t,x) = \alpha_j & \Rightarrow \bar d_j(t,x) \geq 0,\\ \bar u_j(t,x) = \beta_j & \Rightarrow \bar d_j(t,x) \leq 0,\\ \alpha_j < \bar u_j(t,x) < \beta_j & \Rightarrow \bar d_j(t,x) = 0, \end{cases} \text{ and } \begin{cases} \bar d_j(t,x) > 0 \Rightarrow \bar u_j(t,x) = \alpha_j,\\ \bar d_j(t,x) < 0 \Rightarrow \bar u_j(t,x) = \beta_j, \end{cases} \tag{3.19}$$

a.e. in Ω_T and $j = 1,2$.

THEOREM 3.5. *Let $\bar{\mathbf{u}}$ be a local solution of problem* (P), *then $J''(\bar{\mathbf{u}})\mathbf{v}^2 \geq 0$ for all $\mathbf{v}\in\mathcal{C}_{\bar{\mathbf{u}}}$.*

Proof. We sketch the proof in the case where $-\infty < \alpha_j < \beta_j < \infty$ for $j = 1$ and 2. The modifications for the other cases are obvious. The reader is referred to [4] a more detailed proof in the case of steady-state Navier-Stokes equations. Take $\mathbf{v}\in\mathcal{C}_{\bar{\mathbf{u}}}$, and for $\varepsilon < \min\{(\beta_j-\alpha_j)/2 :\ 1\leq j\leq 2\}$ define

$$v_{j,\varepsilon}(t,x) = \begin{cases} 0 & \text{if } \alpha_j < \bar u_j(t,x) < \alpha_j+\varepsilon,\\ 0 & \text{if } \beta_j-\varepsilon < \bar u_j(t,x) < \beta_j,\\ \mathrm{Proj}_{[-\frac{1}{\varepsilon},\frac{1}{\varepsilon}]}(v_j(t,x)) & \text{otherwise.} \end{cases}$$

Then, it is easy to check that $\mathbf{v}_\varepsilon\in\mathcal{C}_{\bar{\mathbf{u}}}$ for every $\varepsilon > 0$ and $\mathbf{v}_\varepsilon\to\mathbf{v}$ strongly in $L^2(0,T;\mathbf{L}^2(\Omega))$. Moreover, $\bar{\mathbf{u}}+\rho\mathbf{v}_\varepsilon\in\mathcal{U}_{ad}$ for every $0<\rho<\varepsilon^2$. Making a second order Taylor expansion of J at $\bar{\mathbf{u}}$ and taking into account that $\bar{\mathbf{u}}$ is a local minimum, for $\rho<\varepsilon^2$ small enough there exists $0<\theta_\rho<\rho$ such that

$$0 \leq J(\bar{\mathbf{u}}+\rho\mathbf{v}_\varepsilon) - J(\bar{\mathbf{u}}) = \rho J'(\bar{\mathbf{u}})\mathbf{v}_\varepsilon + \frac{\rho^2}{2}J''(\bar{\mathbf{u}}+\theta_\rho\mathbf{v}_\varepsilon)\mathbf{v}_\varepsilon^2.$$

Since $\mathbf{v}_\varepsilon\in\mathcal{C}_{\bar{\mathbf{u}}}$, (3.18) implies that $J'(\bar{\mathbf{u}})\mathbf{v}_\varepsilon = 0$. Therefore, the above inequality leads to $J''(\bar{\mathbf{u}}+\theta_\rho\mathbf{v}_\varepsilon)\mathbf{v}_\varepsilon^2 \geq 0$. Now we must take the limit as $\rho\to 0$ to get $J''(\bar{\mathbf{u}})\mathbf{v}_\varepsilon^2\geq 0$. Next, it is enough to take the limit as $\varepsilon\to 0$. To do this, let us recall the expression of $J''(\bar{\mathbf{u}})$ provided by (3.7)

$$\begin{aligned} &J''(\bar{\mathbf{u}})\mathbf{v}_\varepsilon^2\\ &= \int_0^T\int_\Omega [|\mathbf{z}_{\mathbf{v}_\varepsilon}|^2 - 2(\mathbf{z}_{\mathbf{v}_\varepsilon}\cdot\nabla)\mathbf{z}_{\mathbf{v}_\varepsilon}\cdot\bar{\boldsymbol\varphi}]\,\mathrm{d}x\mathrm{d}t + \gamma\int_\Omega |\mathbf{z}_{\mathbf{v}_\varepsilon}(T)|^2\,\mathrm{d}x + \lambda\int_0^T\int_\Omega |\mathbf{v}_\varepsilon|^2\mathrm{d}x\mathrm{d}t\\ &\longrightarrow \int_0^T\int_\Omega [|\mathbf{z}_{\mathbf{v}}|^2 - 2(\mathbf{z}_{\mathbf{v}}\cdot\nabla)\mathbf{z}_{\mathbf{v}}\cdot\bar{\boldsymbol\varphi}]\,\mathrm{d}x\mathrm{d}t + \gamma\int_\Omega |\mathbf{z}_{\mathbf{v}}(T)|^2\,\mathrm{d}x + \lambda\int_0^T\int_\Omega |\mathbf{v}|^2\mathrm{d}x\mathrm{d}t\\ &= J''(\bar{\mathbf{u}})\mathbf{v}^2 \quad \text{as } \varepsilon\to 0. \end{aligned}$$

□

Finally, we formulate the sufficient conditions for optimality.

THEOREM 3.6. *Let us assume that* $\bar{\mathbf{u}} \in \mathcal{U}_{ad}$ *satisfies*

$$J'(\bar{\mathbf{u}})(\mathbf{u}-\bar{\mathbf{u}}) \geq 0 \quad \forall \mathbf{u} \in \mathcal{U}_{ad}, \tag{3.20}$$

$$J''(\bar{\mathbf{u}})\mathbf{v}^2 > 0 \quad \forall \mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}} \setminus \{0\}, \tag{3.21}$$

then there exist $\varepsilon > 0$ *and* $\delta > 0$ *such that*

$$J(\bar{\mathbf{u}}) + \frac{\delta}{2}\|\mathbf{u}-\bar{\mathbf{u}}\|^2_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq J(\mathbf{u}) \quad \forall \mathbf{u} \in \mathcal{U}_{ad} \cap B_\varepsilon(\bar{\mathbf{u}}), \tag{3.22}$$

where $B_\varepsilon(\bar{\mathbf{u}})$ *is the* $L^2(0,T;\mathbf{L}^2(\Omega))$*-ball of center* $\bar{\mathbf{u}}$ *and radius* ε.

Proof. The proof follows by contradiction. Indeed, let us suppose that the theorem is false, then there exists a sequence $\{\mathbf{u}_k\}_{k=1}^\infty \subset \mathcal{U}_{ad}$ such that

$$\|\bar{\mathbf{u}}-\mathbf{u}_k\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq \frac{1}{k} \quad \text{and} \quad J(\bar{\mathbf{u}}) + \frac{1}{2k}\|\bar{\mathbf{u}}-\mathbf{u}_k\|^2_{L^2(0,T;\mathbf{L}^2(\Omega))} > J(\mathbf{u}_k). \tag{3.23}$$

Now, we define

$$\rho_k = \|\bar{\mathbf{u}}-\mathbf{u}_k\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \quad \text{and} \quad \mathbf{v}_k = \frac{1}{\rho_k}(\mathbf{u}_k-\bar{\mathbf{u}}). \tag{3.24}$$

Then, taking a subsequence if necessary, we can assume that $\mathbf{v}_k \rightharpoonup \mathbf{v}$ weakly in $L^2(0,T;\mathbf{L}^2(\Omega))$. The proof is divide into three steps.

Step I - $\mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}}$. We have to prove that $\mathbf{v}$ satisfies (3.15)-(3.17). First, we observe that the set of elements of $L^2(0,T;\mathbf{L}^2(\Omega))$ satisfying (3.15) and (3.16) is closed and convex. From the definition of $\mathbf{v}_k$, it is obvious that each $\mathbf{v}_k$ satisfies (3.15) and (3.16), therefore its weak limit also does it. Let us prove (3.17). From (3.23) and using the mean value theorem we get for some $0 < \theta_k < 1$

$$J(\bar{\mathbf{u}}) + \frac{\rho_k^2}{2k} > J(\mathbf{u}_k) = J(\bar{\mathbf{u}} + \rho_k\mathbf{v}_k) = J(\bar{\mathbf{u}}) + \rho_k J'(\bar{\mathbf{u}} + \theta_k\rho_k\mathbf{v}_k)\mathbf{v}_k,$$

hence

$$J'(\bar{\mathbf{u}} + \theta_k\rho_k\mathbf{v}_k) < \frac{\rho_k}{2k} \to 0. \tag{3.25}$$

Let us prove that $J'(\bar{\mathbf{u}} + \theta_k\rho_k\mathbf{v}_k) \to J'(\bar{\mathbf{u}})\mathbf{v}$. To this end, we set $\mathbf{u}_{\theta_k} = \bar{\mathbf{u}} + \theta_k\rho_k\mathbf{v}_k$. From (3.23) and (3.24) we know that $\mathbf{u}_{\theta_k} \to \bar{\mathbf{u}}$ in $L^2(0,T;\mathbf{L}^2(\Omega))$ strongly. Therefore, its associated state $\mathbf{y}_{\theta_k}$ and adjoint state $\boldsymbol{\varphi}_{\theta_k}$ converge strongly to $\bar{\mathbf{y}}$ and $\bar{\boldsymbol{\varphi}}$ in $\mathbf{H}^{2,1}(\Omega_T) \cap C([0,T],\mathbf{Y})$, then with (3.6) we have

$$J'(\bar{\mathbf{u}} + \theta_k\rho_k\mathbf{v}_k) = \int_0^T \int_\Omega (\boldsymbol{\varphi}_{\theta_k} + \lambda\mathbf{u}_{\theta_k})\mathbf{v}_k \,\mathrm{d}x\mathrm{d}t \to \int_0^T \int_\Omega (\bar{\boldsymbol{\varphi}} + \lambda\bar{\mathbf{u}})\mathbf{v} \,\mathrm{d}x\mathrm{d}t = J'(\bar{\mathbf{u}})\mathbf{v}.$$

Then, (3.25) implies that $J'(\bar{\mathbf{u}})\mathbf{v} \leq 0$. But, (3.15), (3.16) and (3.19) (recall that (3.20) holds and this implies (3.19)) lead to the identities $\bar{d}_j(t,x)v_j(t,x) \geq 0$ for almost every $(t,x) \in \Omega_T$, then

$$\sum_{j=1}^2 \int_0^T \int_\Omega |\bar{d}_j v_j| \,\mathrm{d}x\mathrm{d}t = \int_0^T \int_\Omega \bar{\mathbf{d}} \cdot \mathbf{v} \,\mathrm{d}x\mathrm{d}t = \int_0^T \int_\Omega (\bar{\boldsymbol{\varphi}} + \lambda\bar{\mathbf{u}})\mathbf{v} \,\mathrm{d}x\mathrm{d}t = J'(\bar{\mathbf{u}})\mathbf{v} \leq 0,$$

which proves that $\mathbf{v}$ satisfies (3.17).

Step II - $\mathbf{v} = 0$. We will prove that $J''(\bar{\mathbf{u}})\mathbf{v}^2 \leq 0$, then according to (3.21), this is only possible if $\mathbf{v} = 0$. Using again (3.23) and (3.24) and making a Taylor expansion, we get for some $0 < \theta_k < 1$

$$J(\bar{\mathbf{u}}) + \rho_k J'(\bar{\mathbf{u}})\mathbf{v}_k + \frac{\rho_k^2}{2} J''(\bar{\mathbf{u}} + \theta_k\rho_k\mathbf{v}_k)\mathbf{v}_k^2 < J(\bar{\mathbf{u}}) + \frac{\rho_k^2}{2k}.$$

From this inequality and (3.20) we get $J''(\bar{\mathbf{u}})\mathbf{v}^2 \leq 0$ as follows

$$J''(\bar{\mathbf{u}} + \theta_k\rho_k\mathbf{v}_k)\mathbf{v}_k^2 < \frac{1}{k} \Rightarrow J''(\bar{\mathbf{u}})\mathbf{v}^2 \leq \liminf_{k\to\infty} J''(\bar{\mathbf{u}} + \theta_k\rho_k\mathbf{v}_k)\mathbf{v}_k^2 \leq 0. \tag{3.26}$$

Let us prove the above pass to the limit. Again, we set $\mathbf{u}_{\theta_k} = \bar{\mathbf{u}} + \theta_k\rho_k\mathbf{v}_k$, $\mathbf{y}_{\theta_k} = G(\mathbf{u}_{\theta_k})$ and $\boldsymbol{\varphi}_{\theta_k}$ the associated adjoint state. We also denote $\mathbf{z}_{\theta_k} = G'(\mathbf{u}_{\theta_k})\mathbf{v}_k$. Now, from (3.7) we have

$$\begin{aligned} J''(\mathbf{u}_{\theta_k})\mathbf{v}_{\theta_k}^2 &= \int_0^T\int_\Omega (|\mathbf{z}_{\theta_k}|^2 - 2(\mathbf{z}_{\theta_k}\cdot\nabla)\mathbf{z}_{\theta_k}\boldsymbol{\varphi}_{\theta_k})\mathrm{d}x\mathrm{d}t \\ &+\gamma\int_\Omega |\mathbf{z}_{\theta_k}(T)|^2\,\mathrm{d}x + \lambda\int_0^T\int_\Omega |\mathbf{v}_k|^2\mathrm{d}x\mathrm{d}t. \end{aligned} \tag{3.27}$$

It is easy to pass to the limit in the first to integrals by using that $\mathbf{z}_{\theta_k} \rightharpoonup \mathbf{z}_\mathbf{v}$ weakly in $H^{2,1}(\Omega_T)$ and $\boldsymbol{\varphi}_{\theta_k} \to \bar{\boldsymbol{\varphi}}$ strongly in $H^{2,1}(\Omega_T)$. In the last integral we use the lower semicontinuity with respect to the weak topology of $L^2(0,T;\mathbf{L}^2(\Omega))$.

Step III - Final contradiction. Since $\mathbf{v} = 0$, we get that $\mathbf{z}_{\theta_k} \rightharpoonup 0$ weakly in $H^{2,1}(\Omega_T)$. Then (3.26), (3.27) and the identity $\|\mathbf{v}_k\|_{L^2(0,T;\mathbf{L}^2(\Omega))} = 1$ allow us to conclude

$$0 \geq \liminf_{k\to\infty} J''(\mathbf{u}_{\theta_k})\mathbf{v}_{\theta_k}^2$$

$$= \liminf_{k\to\infty}\left\{\int_0^T\int_\Omega (|\mathbf{z}_{\theta_k}|^2 - 2(\mathbf{z}_{\theta_k}\cdot\nabla)\mathbf{z}_{\theta_k}\boldsymbol{\varphi}_{\theta_k})\mathrm{d}x\mathrm{d}t + \gamma\int_\Omega |\mathbf{z}_{\theta_k}(T)|^2\,\mathrm{d}x\right\} + \lambda = \lambda,$$

which is a contradiction. □

REMARK 3.7. *The gap between the necessary optimality conditions provided by Theorems 3.4 and 3.5 and the sufficient ones given in Theorem 3.6 is minimal, the same than we have in finite dimensional optimization problems. This problem does not suffer from the typical two-norm discrepancy arising usually in infinite dimensional optimization problems. This is due to the C^2-differentiability of J with respect to the $L^2(0,T;\mathbf{L}^2(\Omega))$-norm, thanks to a certain compactness with respect to $\mathbf{u}$ in the first two integrals defining J and the fact that the last one is the square of the norm of the control. On the other hand it is well known that the condition $J''(\bar{\mathbf{u}})\mathbf{v}^2 > 0$ for every non zero $\mathbf{v}$ belonging to the cone of critical directions is not a sufficient optimality condition, in general, in infinite dimensional optimization problems. An inequality of type $J''(\bar{\mathbf{u}})\mathbf{v}^2 \geq \delta\|\mathbf{v}\|^2_{L^2(0,T;\mathbf{L}^2(\Omega))}$ is required in the infinite dimensional case. In finite dimension, both conditions are equivalent, but this is not the usual case for infinite dimension. However, in our problem we can prove that both conditions are also* *equivalent. Indeed, let us observe that* (3.22) *implies that* $\bar{\mathbf{u}}$ *is a local solution of the problem*

$$(\mathrm{P}_\varepsilon)\quad \begin{cases} \min J_\delta(\mathbf{u}) = J(\mathbf{u}) - \dfrac{\delta}{2}\|\mathbf{u}-\bar{\mathbf{u}}\|^2_{L^2(0,T;\mathbf{L}^2(\Omega))} \\ \mathbf{u}\in\mathcal{U}_{ad}\cap B_\varepsilon(\bar{\mathbf{u}}) \end{cases}$$

Therefore, we can apply Theorem 3.5 and obtain that $J''_\delta(\bar{\mathbf{u}})\mathbf{v}^2 \geq 0$ *for every* $\mathbf{v}\in\mathcal{C}_{\bar{\mathbf{u}}}$*. It is enough to notice that* $J''_\delta(\bar{\mathbf{u}})\mathbf{v}^2 = J''(\bar{\mathbf{u}})\mathbf{v}^2 - \delta\|\mathbf{v}\|^2_{L^2(0,T;\mathbf{L}^2(\Omega))}$ *to conclude that* (3.20)-(3.21) *imply*

$$J''(\bar{\mathbf{u}})\mathbf{v}^2 \geq \delta\|\mathbf{v}\|^2 \quad \forall \mathbf{v}\in\mathcal{C}_{\bar{\mathbf{u}}}. \tag{3.28}$$

4. Numerical approximation of the control problem. In this section we consider the complete discretization of the control problem (P). To this end, we consider a family of triangulations $\{\mathcal{T}_h\}_{h>0}$ of $\bar{\Omega}$, defined in the standard way, e.g. in [2, Chapter 3.3]. With each element $T\in\mathcal{T}_h$, we associate two parameters h_T and ϱ_T, where h_T denotes the diameter of the set T and ϱ_T is the diameter of the largest ball contained in T. Define the size of the mesh by $h = \max_{T\in\mathcal{T}_h} h_T$. We also assume that the following regularity assumptions on the triangulation are satisfied.

(i) – There exist two positive constants $\varrho_{\mathcal{T}}$ and $\delta_{\mathcal{T}}$ such that

$$\frac{h_T}{\varrho_T} \leq \varrho_{\mathcal{T}} \quad\text{and}\quad \frac{h}{h_T}\leq \delta_{\mathcal{T}} \quad \forall T\in\mathcal{T}_h \text{ and } \forall h>0.$$

(ii) – Define $\overline{\Omega}_h = \cup_{T\in\mathcal{T}_h} T$, and let Ω_h and Γ_h denote its interior and its boundary, respectively. We assume that the vertices of $\mathcal{T}_h$ placed on the boundary Γ_h are points of Γ.

Since Ω is convex, from the last assumption we have that Ω_h is also convex. Moreover, we know that

$$|\Omega\setminus\Omega_h| \leq Ch^2; \tag{4.1}$$

see, for instance, [23, estimate (5.2.19)].

On the mesh $\mathcal{T}_h$ we consider two finite dimensional spaces $\mathbf{Z}_h\subset\mathbf{H}^1_0(\Omega)$ and $Q_h\subset L^2_0(\Omega)$ formed by piecewise polynomials in Ω_h and vanishing in $\Omega\setminus\Omega_h$. We make the following assumptions on these spaces.

(A1) If $\mathbf{z}\in\mathbf{H}^{1+l}(\Omega)\cap\mathbf{H}^1_0(\Omega)$, then

$$\inf_{\mathbf{z}_h\in\mathbf{Z}_h}\|\mathbf{z}-\mathbf{z}_h\|_{\mathbf{H}^s(\Omega_h)} \leq Ch^{l+1-s}\|\mathbf{z}\|_{\mathbf{H}^{1+l}(\Omega)}, \quad \text{for } 0\leq l\leq 1 \text{ and } s=0,1. \tag{4.2}$$

(A2) If $q\in H^l(\Omega)\cap L^2_0(\Omega)$, then

$$\inf_{q_h\in Q_h}\|q-q_h\|_{L^2(\Omega_h)} \leq Ch\|q\|_{H^1(\Omega)}. \tag{4.3}$$

(A3) The subspaces $\mathbf{Z}_h$ and Q_h satisfy the inf-sup condition: $\exists c>0$ such that

$$\inf_{q_h\in Q_h}\sup_{\mathbf{z}_h\in\mathbf{Z}_h}\frac{b(\mathbf{z}_h,q_h)}{\|\mathbf{z}_h\|_{\mathbf{H}^1(\Omega_h)}\|q_h\|_{L^2(\Omega_h)}} \geq c, \tag{4.4}$$

where $b : \mathbf{H}^1(\Omega) \times L^2(\Omega) \longrightarrow \mathbb{R}$ is defined by

$$b(\mathbf{z}, q) = \int_\Omega q(x) \operatorname{div} \mathbf{z}(x)\, \mathrm{d}x.$$

These assumptions are satisfied by the usual finite elements considered in the discretization of Navier-Stokes equations: "Taylor-Hood", P1-Bubble finite element, and some others; see [12, Chapter 2].

We also consider a subspace $\mathbf{Y}_h$ of $\mathbf{Z}_h$ defined by

$$\mathbf{Y}_h = \{\mathbf{y}_h \in \mathbf{Z}_h : b(\mathbf{y}_h, q_h) = 0 \ \forall q_h \in Q_h\}$$

and we set

$$\mathbf{U}_h = \{\mathbf{u}_h \in \mathbf{L}^2(\Omega_h) : \mathbf{u}_{h|_T} \equiv \mathbf{u}_T \in \mathbb{R}^2\}.$$

We proceed now with the discretization in time. Let us consider a grid of points $0 = t_0 < t_1 < \ldots < t_{N_\tau} = T$. We denote $\tau_n = t_n - t_{n-1}$. We make the following assumption

$$\exists \varrho_0 > 0 \text{ such that } \tau = \max_{1 \le n \le N_\tau} \tau_n < \varrho_0 \tau_n \quad \forall 1 \le n \le N_\tau \text{ and } \forall \tau > 0. \tag{4.5}$$

Given a triangulation $\mathcal{T}_h$ of Ω and a grid of points $\{t_n\}_{n=0}^{N_\tau}$ of $[0, T]$, we set $\sigma = (\tau, h)$. Finally, we consider the following spaces

$$\begin{aligned}
\mathcal{Y}_\sigma &= \{\mathbf{y}_\sigma \in L^2(0, T; \mathbf{Y}_h) : \mathbf{y}_{\sigma|_{(t_{n-1}, t_n)}} \in \mathbf{Y}_h \ \text{ for } \ 1 \le n \le N_\tau\},\\
\mathcal{Q}_\sigma &= \{q_\sigma \in L^2(0, T; Q_h) : q_{\sigma|_{(t_{n-1}, t_n)}} \in Q_h \ \text{ for } \ 1 \le n \le N_\tau\},\\
\mathcal{U}_\sigma &= \{\mathbf{u}_\sigma \in L^2(0, T; \mathbf{U}_h) : \mathbf{u}_{\sigma|_{(t_{n-1}, t_n)}} \in \mathbf{U}_h \ \text{ for } \ 1 \le n \le N_\tau\}.
\end{aligned}$$

We have that the functions of $\mathcal{Y}_\sigma$, $\mathcal{Q}_\sigma$ and $\mathcal{U}_\sigma$ are piecewise constant in time. We will look for the discrete controls in the space $\mathcal{U}_\sigma$. An element of this space can be written in the form

$$\mathbf{u}_\sigma = \sum_{n=1}^{N_\tau} \sum_{T \in \mathcal{T}_h} \mathbf{u}_{n,T} \chi_n \chi_T, \quad \text{with} \ \ \mathbf{u}_{n,T} \in \mathbb{R}^2, \tag{4.6}$$

where χ_n and χ_T are the characteristic functions of (t_{n-1}, t_n) and T, respectively. Therefore, the dimension of $\mathcal{U}_\sigma$ is $2N_\tau N_h$, where N_h is the number of triangles in $\mathcal{T}_h$. In $\mathcal{U}_\sigma$ we consider the convex subset

$$\mathcal{U}_{\sigma,ad} = \mathcal{U}_\sigma \cap \mathcal{U}_{ad} = \{\mathbf{u}_\sigma \in \mathcal{U}_\sigma : \mathbf{u}_{n,T} \in [\alpha_1, \beta_1] \times [\alpha_2, \beta_2]\}.$$

On the other hand, the elements of $\mathcal{Y}_\sigma$ can be written in the form

$$\mathbf{y}_\sigma = \sum_{n=1}^{N_\tau} \mathbf{y}_{n,h} \chi_n, \quad \text{with} \ \ \mathbf{y}_{n,h} \in \mathbf{Y}_h, \tag{4.7}$$

where χ_n is as above. For every discrete state $\mathbf{y}_\sigma$ we will fix $\mathbf{y}_\sigma(t_n) = \mathbf{y}_{n,h}$, so that $\mathbf{y}_\sigma$ is continuous on the left. In particular, we have $\mathbf{y}_\sigma(T) = \mathbf{y}_\sigma(t_{N_\tau}) = y_{N_\tau,h}$.

To define the discrete control problem we have to consider the numerical discretization of the state equation (1.1) or equivalently (2.1). We achieve this goal by using a discontinuous time-stepping Galerkin method, with piecewise constants in time and conforming finite element spaces in space. For any $\mathbf{u} \in L^2(0,T;\mathbf{L}^2(\Omega))$ the discrete state equation is given by

$$(4.8)\quad \begin{cases} \text{For } n = 1,\dots,N_\tau, \quad \text{and } \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \left(\dfrac{\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}}{\tau_n}, \mathbf{w}_h\right) + a(\mathbf{y}_{n,h}, \mathbf{w}_h) + c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_h) = (\mathbf{f}_n + \mathbf{u}_n, \mathbf{w}_h), \\ \mathbf{y}_{0,h} = \mathbf{y}_{0h}, \end{cases}$$

where

$$(4.9)\qquad (\mathbf{f}_n, \mathbf{w}_h) = \frac{1}{\tau_n}\int_{t_{n-1}}^{t_n} (\mathbf{f}(t), \mathbf{w}_h)\mathrm{d}t, \quad (\mathbf{u}_n, \mathbf{w}_h) = \frac{1}{\tau_n}\int_{t_{n-1}}^{t_n} (\mathbf{u}(t), \mathbf{w}_h)\mathrm{d}t$$

and

$$(4.10)\quad \mathbf{y}_{0h} \in \mathbf{Y}_h \ \text{ with } \ \|\mathbf{y}_0 - \mathbf{y}_{0h}\|_{\mathbf{L}^2(\Omega_h)} \le Ch \ \text{ and } \ \|\mathbf{y}_{0h}\|_{\mathbf{H}^1(\Omega_h)} \le C \ \ \forall h > 0.$$

The above scheme is essentially an implicit Euler in time / conforming in space scheme, and can be easily extended to higher order polynomial in time discretizations; see e.g. [27] and references within. For stability and error estimates under suitable regularity assumptions for high order discontinuous time-stepping schemes we refer the reader to [7]. Here, we focus on the lowest case of polynomial approximation in time, due to the low regularity imposed by the nature of our optimal control problem

We will prove later that for any $\mathbf{u} \in L^2(0,T;\mathbf{L}^2(\Omega))$, (4.8) has a unique solution $\mathbf{y}_\sigma(\mathbf{u}) \in \mathcal{Y}_\sigma$. A key feature of the proposed scheme is that the regularity properties of the discrete solution mimics the continuous problem. Then, we can define the discrete control problem as follows

$$(\mathrm{P}_\sigma) \qquad \begin{cases} \min J_\sigma(\mathbf{u}_\sigma) \\ \mathbf{u}_\sigma \in \mathcal{U}_{\sigma,ad} \end{cases}$$

where

$$J_\sigma(\mathbf{u}_\sigma) = \frac{1}{2}\int_0^T \int_{\Omega_h} |\mathbf{y}_\sigma(\mathbf{u}_\sigma) - \mathbf{y}_d|^2 \mathrm{d}x\mathrm{d}t$$
$$+\frac{\gamma}{2}\int_{\Omega_h} |\mathbf{y}_\sigma(T) - \mathbf{y}_{\Omega_h}|^2 \mathrm{d}x + \frac{\lambda}{2}\int_0^T \int_{\Omega_h} |\mathbf{u}_\sigma|^2 \mathrm{d}x\mathrm{d}t,$$

and

$$(4.11)\quad \mathbf{y}_{\Omega_h} \in \mathbf{Y}_h \ \text{ with } \ \|\mathbf{y}_\Omega - \mathbf{y}_{\Omega_h}\|_{\mathbf{L}^2(\Omega_h)} \le Ch \ \text{ and } \ \|\mathbf{y}_{\Omega_h}\|_{\mathbf{H}^1(\Omega_h)} \le C \ \ \forall h > 0.$$

The study of the control problem is divided in several subsections. First, we analyze the discrete state equation (4.8); then we study the discrete adjoint state equation; the third step is the proof of the convergence of (P_σ); and finally we prove the error estimates for the discretization.

4.1. Analysis of the discrete state equation. By a standard argument, using the identity $c(\mathbf{z},\mathbf{w},\mathbf{w}) = 0$ $\forall \mathbf{z} \in \mathbf{L}^4(\Omega)$ and $\forall \mathbf{w} \in \mathbf{H}^1(\Omega)$ (Lemma 2.1) and the Brower's fixed-point theorem, we can easily prove that (4.8) has at least one solution. In this section, we will prove that the solution is unique under some restrictions on $\sigma = (\tau, h)$. For the moment, let us denote $\mathbf{y} = \mathbf{y_u} = G(\mathbf{u})$ and $\mathbf{y}_\sigma \in \mathcal{Y}_\sigma$ a solution of (4.8) We are going to prove some error estimates for $\mathbf{y} - \mathbf{y}_\sigma$. To this end, we need to introduce some projection operators.

DEFINITION 4.1. *We define the projection operator* $P_h : \mathbf{L}^2(\Omega) \longrightarrow \mathbf{Y}_h$ *by*

$$(P_h\mathbf{y}, \mathbf{w}_h) = (\mathbf{y}, \mathbf{w}_h) \quad \forall \mathbf{w}_h \in \mathbf{Y}_h.$$

We also define $P_\sigma : C([0,T], \mathbf{L}^2(\Omega)) \longrightarrow \mathcal{Y}_\sigma$ *by* $(P_\sigma\mathbf{y})_{n,h} = P_h\mathbf{y}(t_n)$ *for every* $1 \le n \le N_\tau$.

LEMMA 4.2. *There exists a constant* $C > 0$ *independent of* σ *such that for every* $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T]; \mathbf{Y})$

$$\|\mathbf{y} - P_\sigma\mathbf{y}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \le C\left\{\tau\|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h^2\|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))}\right\}. \tag{4.12}$$

Proof. From Assumptions (A1)–(A3) and using (4.2) with $s = 0$ and $l = 1$ (see also [12, Chapter II]), the definition of P_σ, and the stability of P_h we get

$$\|\mathbf{y} - P_\sigma\mathbf{y}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} = \left\{\sum_{n=1}^{N_\tau}\int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - P_h\mathbf{y}(t_n)\|^2\,\mathrm{d}t\right\}^{1/2}$$
$$\le \left\{\sum_{n=1}^{N_\tau}\int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - P_h\mathbf{y}(t)\|^2\,\mathrm{d}t\right\}^{1/2} + \left\{\sum_{n=1}^{N_\tau}\int_{t_{n-1}}^{t_n} \|P_h\mathbf{y}(t) - P_h\mathbf{y}(t_n)\|^2\,\mathrm{d}t\right\}^{1/2}$$
$$\le Ch^2\left\{\sum_{n=1}^{N_\tau}\int_{t_{n-1}}^{t_n} \|\mathbf{y}(t)\|^2_{\mathbf{H}^2(\Omega)}\,\mathrm{d}t\right\}^{1/2} + \left\{\sum_{n=1}^{N_\tau}\int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}(t_n)\|^2\,\mathrm{d}t\right\}^{1/2}$$
$$\le Ch^2\|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + \left\{\sum_{n=1}^{N_\tau}\int_{t_{n-1}}^{t_n}(t_n - t)\int_{t_{n-1}}^{t_n} \|\mathbf{y}'(s)\|^2\,\mathrm{d}s\mathrm{d}t\right\}^{1/2}$$
$$\le C\left\{h^2\|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + \tau\|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\right\}.$$

□

DEFINITION 4.3. *The operator* $\Pi_h : \mathbf{Y} \longrightarrow \mathbf{Y}_h$ *is defined by*

$$a(\Pi_h\mathbf{y}, \mathbf{w}_h) = a(\mathbf{y}, \mathbf{w}_h) \quad \forall \mathbf{w}_h \in \mathbf{Y}_h.$$

For any element $\mathbf{y} \in C([0,T]; \mathbf{Y})$, *we define* $\mathbf{y}_h \in C([0,T], \mathbf{Y}_h)$ *by* $\mathbf{y}_h(t) = \Pi_h(\mathbf{y}(t))$.

The next lemma is an immediate consequence of Assumptions (A1)–(A3); see again [12, Chapter II].

LEMMA 4.4. *There exists a constant* $C > 0$ *independent of* σ *such that*

$$\|\mathbf{y} - \Pi_h\mathbf{y}\|_{\mathbf{H}^s(\Omega_h)} \le Ch^{2-s}\|\mathbf{y}\|_{\mathbf{H}^2(\Omega)} \quad \forall \mathbf{y} \in \mathbf{H}^2(\Omega) \cap \mathbf{Y} \quad \textit{and} \quad s = 0, 1. \tag{4.13}$$

As a consequence of the previous two lemmas we have the following result.

LEMMA 4.5. *There exists a constant $C > 0$ independent of σ such that for every* $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T];\mathbf{Y})$

$$\|\mathbf{y} - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \le C\left\{\frac{\tau}{h}\|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h\|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))}\right\}. \tag{4.14}$$

Proof. From Lemma 4.4 we get

$$\begin{aligned}&\|\mathbf{y} - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \le \|\mathbf{y} - \mathbf{y}_h\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} + \|\mathbf{y}_h - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\\ &\le Ch\|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + \|\mathbf{y}_h - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}.\end{aligned}$$

Let us estimate the last term. Using the definition of P_σ, an inverse inequality, (4.12) and (4.13) we obtain

$$\begin{aligned}&\|\mathbf{y}_h - P_\sigma \mathbf{y}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} = \left\{\sum_{n=1}^{N_\tau}\int_{t_{n-1}}^{t_n}\|\mathbf{y}_h(t) - P_h\mathbf{y}(t_n)\|^2_{\mathbf{H}^1(\Omega_h)}\,\mathrm{d}t\right\}^{1/2}\\ &\le \frac{C}{h}\left\{\sum_{n=1}^{N_\tau}\int_{t_{n-1}}^{t_n}\|\mathbf{y}_h(t) - P_h\mathbf{y}(t_n)\|^2_{\mathbf{L}^2(\Omega_h)}\,\mathrm{d}t\right\}^{1/2}\\ &\le \frac{C}{h}\left\{\|\mathbf{y} - \mathbf{y}_h\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} + \|\mathbf{y} - P_\sigma\mathbf{y}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\right\}\\ &\le C\left\{\frac{\tau}{h}\|\mathbf{y}'\|_{L^2(0,T;L^2(\Omega))} + h\|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))}\right\}.\end{aligned}$$

□

Before proving the error estimates for $\mathbf{y} - \mathbf{y}_\sigma$, we need to establish the corresponding estimates for the Stokes problem. Let us formulate this result as follows.

LEMMA 4.6. *Let $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T],\mathbf{Y})$ be the solution of* (2.1) *and let $\hat{\mathbf{y}}_\sigma \in \mathcal{Y}_\sigma$ satisfy*

$$\begin{cases}\text{For } n = 1,\ldots,N_\tau, \quad \text{and} \quad \forall \mathbf{w}_h \in \mathbf{Y}_h,\\ \left(\dfrac{\hat{\mathbf{y}}_{n,h} - \hat{\mathbf{y}}_{n-1,h}}{\tau_n}, \mathbf{w}_h\right) + a(\hat{\mathbf{y}}_{n,h}, \mathbf{w}_h) = (\hat{\mathbf{f}}_n, \mathbf{w}_h),\\ \hat{\mathbf{y}}_{0,h} = \mathbf{y}_{0h},\end{cases} \tag{4.15}$$

where

$$\begin{aligned}(\hat{\mathbf{f}}_n, \mathbf{w}_h) &= \frac{1}{\tau_n}\int_{t_{n-1}}^{t_n}\{a(\mathbf{y}(t), \mathbf{w}_h) + (\mathbf{y}'(t), \mathbf{w}_h)\}\,\mathrm{d}t\\ &= \frac{1}{\tau_n}\int_{t_{n-1}}^{t_n} a(\mathbf{y}(t), \mathbf{w}_h)\,\mathrm{d}t + \left(\frac{\mathbf{y}(t_n) - \mathbf{y}(t_{n-1})}{\tau_n}, \mathbf{w}_h\right).\end{aligned}$$

Then, (4.15) *has a unique solution $\hat{\mathbf{y}}_\sigma \in \mathcal{Y}_\sigma$: moreover, the following properties hold*

1 - $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ is bounded in $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$.

2 - There exists a constant $C > 0$ independent of σ such that

$$\begin{aligned}&\max_{1\le n\le N_\tau}\|\mathbf{y}(t_n) - \mathbf{y}_\sigma(t_n)\| + \|\mathbf{y} - \hat{\mathbf{y}}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\\ &\le C\left\{\frac{\tau}{h}\|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h\|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + h\|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)}\right\}.\end{aligned} \tag{4.16}$$

Proof. Since $\hat{\mathbf{y}}$ depends linearly on the data $(\hat{\mathbf{f}}, \mathbf{y}_{0h})$ and the number of unknowns $\{\mathbf{y}_{n,h}\}_{n=1}^{N_\tau}$ is equal to the number of equations, it is enough to prove that $\hat{\mathbf{y}} = \mathbf{0}$ is the unique solution corresponding to $(\hat{\mathbf{f}}, \mathbf{y}_{0h}) = (\mathbf{0}, \mathbf{0})$. This follows in the usual way by multiplying the n-th equation of (4.15) by $\hat{\mathbf{y}}_{n,h}$ and ordering the terms in the way

$$\frac{1}{2}\|\hat{\mathbf{y}}_{n,h}\|^2 - \frac{1}{2}\|\hat{\mathbf{y}}_{n-1,h}\|^2 + \frac{1}{2}\|\hat{\mathbf{y}}_{n,h} - \hat{\mathbf{y}}_{n-1,h}\|^2 + \tau_n a(\hat{\mathbf{y}}_{n,h}, \hat{\mathbf{y}}_{n,h}) = 0.$$

Now, making the sum of these identities, we get

$$\frac{1}{2}\|\hat{\mathbf{y}}_{N_\tau,h}\|^2 + \int_0^T a(\hat{\mathbf{y}}_\sigma, \hat{\mathbf{y}}_\sigma)\, dt = 0,$$

then $\hat{\mathbf{y}}_\sigma = \mathbf{0}$.

Now, we prove that $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ is bounded in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$. To this end, we introduce the discrete Laplacian $\Delta_h \hat{\mathbf{y}}_{n,h} \in \mathbf{Y}_h$ defined by

$$(\Delta_h \hat{\mathbf{y}}_{n,h}, \mathbf{w}_h) = a(\hat{\mathbf{y}}_{n,h}, \mathbf{w}_h) \quad \forall \mathbf{w}_h \in \mathbf{Y}_h. \tag{4.17}$$

Now we take $\mathbf{w}_h = \Delta_h \hat{\mathbf{y}}_{n,h}$ in (4.15) and we use (4.17) to get

$$a\left(\hat{\mathbf{y}}_{n,h}, \frac{\hat{\mathbf{y}}_{n,h} - \hat{\mathbf{y}}_{n-1,h}}{\tau_n}\right) + \|\Delta_h \hat{\mathbf{y}}_{n,h}\|^2 = (\hat{\mathbf{f}}_n, \Delta_h \hat{\mathbf{y}}_{n,h}),$$

which can be written

$$\begin{aligned}&\frac{1}{2}a(\hat{\mathbf{y}}_{n,h}, \hat{\mathbf{y}}_{n,h}) - \frac{1}{2}a(\hat{\mathbf{y}}_{n-1,h}, \hat{\mathbf{y}}_{n-1,h}) + \frac{1}{2}a(\hat{\mathbf{y}}_{n,h} - \hat{\mathbf{y}}_{n-1,h}, \hat{\mathbf{y}}_{n,h} - \hat{\mathbf{y}}_{n-1,h})\\ &+\tau_n\|\Delta_h \hat{\mathbf{y}}_{n,h}\|^2 = \tau_n(\hat{\mathbf{f}}_n, \Delta_h \hat{\mathbf{y}}_{n,h}).\end{aligned} \tag{4.18}$$

Since $\mathbf{y}(t) \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega)$ for almost every t and $\Delta_h \hat{\mathbf{y}}_{n,h} \in \mathbf{H}_0^1(\Omega_h)$, we can make an integration by parts and deduce from the definition of $\hat{\mathbf{f}}_n$ and from Young's inequality

$$\begin{aligned}\tau_n(\hat{\mathbf{f}}_n, \Delta_h \hat{\mathbf{y}}_{n,h}) &= -\int_{t_{n-1}}^{t_n} (\Delta \mathbf{y}, \Delta_h \hat{\mathbf{y}}_{n,h})\, \mathrm{d}t + \int_{t_{n-1}}^{t_n} (\mathbf{y}'(t), \Delta_h \hat{\mathbf{y}}_{n,h})\, dt\\ &\le \int_{t_{n-1}}^{t_n} \|\mathbf{y}\|_{\mathbf{H}^2(\Omega)}\|\Delta_h \hat{\mathbf{y}}_{n,h}\|\, \mathrm{d}t + \int_{t_{n-1}}^{t_n} \|\mathbf{y}'(t)\|_{\mathbf{L}^2(\Omega)}\|\Delta_h \hat{\mathbf{y}}_{n,h}\|\, \mathrm{d}t\\ &\le \int_{t_{n-1}}^{t_n} \|\mathbf{y}\|^2_{\mathbf{H}^2(\Omega)}\, \mathrm{d}t + \int_{t_{n-1}}^{t_n} \|\mathbf{y}'(t)\|^2_{\mathbf{L}^2(\Omega)}\, \mathrm{d}t + \frac{\tau_n}{2}\|\Delta_h \hat{\mathbf{y}}_{n,h}\|^2.\end{aligned}$$

Substituting the above inequality in (4.18), it follows

$$\begin{aligned}&\frac{1}{2}a(\hat{\mathbf{y}}_{n,h}, \hat{\mathbf{y}}_{n,h}) - \frac{1}{2}a(\hat{\mathbf{y}}_{n-1,h}, \hat{\mathbf{y}}_{n-1,h}) + \frac{1}{2}a(\hat{\mathbf{y}}_{n,h} - \hat{\mathbf{y}}_{n-1,h}, \hat{\mathbf{y}}_{n,h} - \hat{\mathbf{y}}_{n-1,h})\\ &+\frac{\tau_n}{2}\|\Delta_h \hat{\mathbf{y}}_{n,h}\|^2 \le \int_{t_{n-1}}^{t_n} \|\mathbf{y}\|^2_{\mathbf{H}^2(\Omega)}\, \mathrm{d}t + \int_{t_{n-1}}^{t_n} \|\mathbf{y}'(t)\|^2_{\mathbf{L}^2(\Omega)}\, \mathrm{d}t.\end{aligned}$$

Making the addition from $n = 1$ until $n = k$ for any $1 \le k \le N_\tau$, we get that

$$\|\hat{\mathbf{y}}_{k,h}\|_{\mathbf{H}^1(\Omega_h)} \le \|\mathbf{y}_{0h}\|^2_{\mathbf{H}^1(\Omega_h)} + 2\int_0^T \|\mathbf{y}\|^2_{\mathbf{H}^2(\Omega)}\, \mathrm{d}t + 2\int_0^T \|\mathbf{y}'(t)\|^2_{\mathbf{L}^2(\Omega)}\, \mathrm{d}t, \quad 1 \le k \le N_\tau,$$

which proves the boundedness of $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ in $L^\infty(0,T;\mathbf{H}^1(\Omega))$.

Finally, we prove (4.16). Let us set $\hat{\mathbf{e}} = \mathbf{y} - \hat{\mathbf{y}}_\sigma$, $\hat{\mathbf{d}} = \mathbf{y} - P_\sigma\mathbf{y}$ and $\hat{\mathbf{e}}_\sigma = P_\sigma\mathbf{y} - \hat{\mathbf{y}}_\sigma$, with P_σ given in Definition 4.1. Thus, we have $\hat{\mathbf{e}} = \hat{\mathbf{d}} + \hat{\mathbf{e}}_\sigma$. Taking $(P_\sigma\mathbf{y})_{0,h} = P_h\mathbf{y}(0)$, the equation (4.15) can be written

$$(4.19) \qquad (\hat{\mathbf{e}}(t_n) - \hat{\mathbf{e}}(t_{n-1}), \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} a(\hat{\mathbf{e}}(t), \mathbf{w}_h)\,\mathrm{d}t = 0 \quad \forall \mathbf{w}_h \in \mathbf{Y}_h,\ 1 \le n \le N_\tau$$

where we have taken into account the second equality defining $(\hat{\mathbf{f}}_n, \mathbf{w}_h)$. Setting $\hat{\mathbf{e}} = \hat{\mathbf{d}} + \hat{\mathbf{e}}_\sigma$ in the above identity we obtain

$$\begin{aligned}&(\hat{\mathbf{e}}_{n,h} - \hat{\mathbf{e}}_{n-1,h}, \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} a(\hat{\mathbf{e}}_{n,h}, \mathbf{w}_h)\,\mathrm{d}t \\ &= -(\hat{\mathbf{d}}(t_n) - \hat{\mathbf{d}}(t_{n-1}), \mathbf{w}_h) - \int_{t_{n-1}}^{t_n} a(\hat{\mathbf{d}}(t), \mathbf{w}_h)\,\mathrm{d}t.\end{aligned}$$

From the definition of P_σ we get

$$(\hat{\mathbf{d}}(t_n) - \hat{\mathbf{d}}(t_{n-1}), \mathbf{w}_h) = (\mathbf{y}(t_n) - P_h\mathbf{y}(t_n), \mathbf{w}_h) - (\mathbf{y}(t_{n-1}) - P_h\mathbf{y}(t_{n-1}), \mathbf{w}_h) = 0.$$

Therefore,

$$\begin{aligned}&(\hat{\mathbf{e}}_{n,h} - \hat{\mathbf{e}}_{n-1,h}, \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} a(\hat{\mathbf{e}}_{n,h}, \mathbf{w}_h)\,\mathrm{d}t \\ &= -\int_{t_{n-1}}^{t_n} a(\hat{\mathbf{d}}(t), \mathbf{w}_h)\,\mathrm{d}t.\end{aligned}$$

Now, taking $\mathbf{w}_h = \hat{\mathbf{e}}_{n,h}$ we obtain

$$\begin{aligned}&\frac{1}{2}\|\hat{\mathbf{e}}_{n,h}\|^2 - \frac{1}{2}\|\hat{\mathbf{e}}_{n-1,h}\|^2 + \frac{1}{2}\|\hat{\mathbf{e}}_{n,h} - \hat{\mathbf{e}}_{n-1,h}\|^2 + \nu\|\nabla\hat{\mathbf{e}}_{n,h}\|^2_{L^2(t_{n-1},t_n;\mathbf{L}^2(\Omega_h))} \\ &\le \nu\|\nabla\hat{\mathbf{d}}\|_{L^2(t_{n-1},t_n;\mathbf{L}^2(\Omega))}\|\nabla\hat{\mathbf{e}}_{n,h}\|_{L^2(t_{n-1},t_n;\mathbf{L}^2(\Omega_h))} \\ &\le \frac{\nu}{2}\|\nabla\hat{\mathbf{d}}\|^2_{L^2(t_{n-1},t_n;\mathbf{L}^2(\Omega_h))} + \frac{\nu}{2}\|\nabla\hat{\mathbf{e}}_{n,h}\|^2_{L^2(t_{n-1},t_n;\mathbf{L}^2(\Omega_h))},\end{aligned}$$

hence

$$\begin{aligned}&\frac{1}{2}\|\hat{\mathbf{e}}_{n,h}\|^2 + \frac{1}{2}\|\hat{\mathbf{e}}_{n,h} - \hat{\mathbf{e}}_{n-1,h}\|^2 + \frac{\nu}{2}\|\nabla\hat{\mathbf{e}}_{n,h}\|^2_{L^2(t_{n-1},t_n;\mathbf{L}^2(\Omega_h))} \\ &\le \frac{1}{2}\|\hat{\mathbf{e}}_{n-1,h}\|^2 + \frac{\nu}{2}\|\nabla\hat{\mathbf{d}}\|^2_{L^2(t_{n-1},t_n;\mathbf{L}^2(\Omega_h))}.\end{aligned}$$

Once again, making the addition from $n = 1$ to k, for $1 \le k \le N_\tau$ and using that $\hat{\mathbf{d}} = \mathbf{y} - P_\sigma\mathbf{y}$ we obtain

$$\begin{aligned}&\|\hat{\mathbf{e}}_{k,h}\|^2 + \sum_{n=1}^{k}\|\hat{\mathbf{e}}_{k,h} - \hat{\mathbf{e}}_{k-1,h}\|^2 + \nu\|\nabla\hat{\mathbf{e}}_{k,h}\|^2_{L^2(0,t_n;\mathbf{L}^2(\Omega_h))} \\ &\le \|\mathbf{y}_0 - \mathbf{y}_{0h}\|^2 + \nu\|\mathbf{y} - P_\sigma\mathbf{y}\|^2_{L^2(0,T;\mathbf{H}^1(\Omega_h))},\end{aligned}$$

which leads to (4.16) with the help of (4.10), (4.14) and the fact that $\mathbf{y} - \hat{\mathbf{y}}_\sigma = (\mathbf{y} - P_\sigma\mathbf{y}) + \hat{\mathbf{e}}$. ∎

Remark 4.7. *From* (4.16) *we can deduce an error estimate in* $L^\infty(0,T;\mathbf{L}^2(\Omega_h))$. *Indeed, let us assume that* $t_{n-1} < t < t_n$ *for some* $1 \leq \leq N_\tau$, *then*

$$\|\mathbf{y}(t) - \hat{\mathbf{y}}_\sigma(t)\| \leq \|\mathbf{y}(t) - \mathbf{y}(t_n)\| + \|\mathbf{y}(t_n) - \hat{\mathbf{y}}_\sigma(t_n)\|.$$

The second term of the right hand side of the inequality has been estimated in (4.16). *Let us study the first term. For any* $\mathbf{w} \in \mathbf{L}^2(\Omega)$

$$|(\mathbf{y}(t) - \mathbf{y}(t_n), \mathbf{w})| = \left|\int_t^{t_n} (\mathbf{y}'(s), \mathbf{w})\,\mathrm{d}s\right| \leq \int_t^{t_n} \|\mathbf{y}'(s)\|\,\mathrm{d}s\|\mathbf{w}\|$$
$$\leq \sqrt{\tau}\|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\|\mathbf{w}\|,$$

which implies

$$\|\mathbf{y}(t) - \mathbf{y}(t_n)\| \leq \sqrt{\tau}\|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))}.$$

Finally, this estimate and (4.16) *infer*

$$\|\mathbf{y} - \hat{\mathbf{y}}_\sigma\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))}$$
$$\leq C\left\{\left(\frac{\tau}{h} + \sqrt{\tau}\right)\|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h\|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + h\|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)}\right\}.$$

The discrete solution of the linear Stokes problem, will subsequently play the role of a global in time projection, which facilitates the derivation of error estimates under the restricted regularity assumptions of the control problem (see also [7]). Finally, we obtain the result concerning the discrete state equation (4.8).

Theorem 4.8. *Given* $\mathbf{u} \in L^2(0,T;\mathbf{L}^2(\Omega))$, *let* $\mathbf{y} \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T];\mathbf{Y})$ *be the solution of* (2.1) *and let* $\mathbf{y}_\sigma \in \mathcal{Y}_\sigma$ *be any solution of* (4.8), *then there exists a constant* $C > 0$ *independent of* $\mathbf{u}$, $\mathbf{y}$ *and* σ *such that*

$$\max_{1\leq n\leq N_\tau} \|\mathbf{y}(t_n) - \mathbf{y}_\sigma(t_n)\| + \|\mathbf{y} - \mathbf{y}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}$$
$$\leq C\left\{\frac{\tau}{h}\|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h\|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + h\|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)}\right\}. \tag{4.20}$$

Moreover, if there exists a constant $C_0 > 0$ *such that* $\tau \leq C_0 h^2$ *for every* $\sigma = (\tau, h)$, *then* $\{\mathbf{y}_\sigma\}_\sigma$ *is bounded in* $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$ *and* (4.8) *has a unique solution.*

Proof. Let us define $\mathbf{e} = \mathbf{y} - \mathbf{y}_\sigma = (\mathbf{y} - \hat{\mathbf{y}}_\sigma) + (\hat{\mathbf{y}}_\sigma - \mathbf{y}_\sigma) = \hat{\mathbf{e}} + \mathbf{e}_\sigma$, where $\hat{\mathbf{y}}_\sigma$ is the solution of (4.15). First we observe that (2.1) and (4.8) imply

$$(\mathbf{e}(t_n) - \mathbf{e}(t_{n-1}), \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} a(\mathbf{e}(t), \mathbf{w}_h)\,\mathrm{d}t$$
$$= \int_{t_{n-1}}^{t_n} \{c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_h) - c(\mathbf{y}(t), \mathbf{y}(t), \mathbf{w}_h)\}\,\mathrm{d}t.$$

Using the decomposition $\mathbf{e} = \hat{\mathbf{e}} + \mathbf{e}_\sigma$, invoking (4.19) and setting $\mathbf{w}_h = \mathbf{e}_{n,h}$ it follows

$$(\mathbf{e}_{n,h} - \mathbf{e}_{n-1,h}, \mathbf{e}_{n,h}) + \int_{t_{n-1}}^{t_n} a(\mathbf{e}_{n,h}, \mathbf{e}_{n,h})\,\mathrm{d}t$$
$$= \int_{t_{n-1}}^{t_n} \{c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}, \mathbf{e}_{n,h}) - c(\mathbf{y}(t), \mathbf{y}(t), \mathbf{e}_{n,h})\}\,\mathrm{d}t. \tag{4.21}$$

Let us estimate the righthand side. First we observe that

$$\begin{aligned}&c(\mathbf{y}_{n,h},\mathbf{y}_{n,h},\mathbf{e}_{n,h})-c(\mathbf{y}(t),\mathbf{y}(t),\mathbf{e}_{n,h})\\&=c(\mathbf{y}_{n,h}-\hat{\mathbf{y}}_{n,h},\mathbf{y}_{n,h},\mathbf{e}_{n,h})-c(\mathbf{y}(t)-\hat{\mathbf{y}}_{n,h},\mathbf{y}(t),\mathbf{e}_{n,h})-c(\hat{\mathbf{y}}_{n,h},\mathbf{y}(t)-\mathbf{y}_{n,h},\mathbf{e}_{n,h}).\end{aligned}$$

Recalling that $\mathbf{e}_{n,h}=\hat{\mathbf{y}}_{n,h}-\mathbf{y}_{n,h}$, $\mathbf{e}(t)=\mathbf{y}(t)-\mathbf{e}_{n,h}$ for every $t\in(t_{n-1},t_n)$, and $\hat{\mathbf{e}}(t)=\mathbf{y}(t)-\hat{\mathbf{y}}_{n,h}$, we have

$$\begin{aligned}&c(\mathbf{y}_{n,h},\mathbf{y}_{n,h},\mathbf{e}_{n,h})-c(\mathbf{y}(t),\mathbf{y}(t),\mathbf{e}_{n,h})\\&=-c(\mathbf{e}_{n,h},\mathbf{y}_{n,h},\mathbf{e}_{n,h})-c(\hat{\mathbf{e}}(t),\mathbf{y}(t),\mathbf{e}_{n,h})-c(\hat{\mathbf{y}}_{n,h},\mathbf{e}(t),\mathbf{e}_{n,h}).\end{aligned}$$

From the identity $c(\mathbf{w}_h,\mathbf{e}_{n,h},\mathbf{e}_{n,h})=0$ for any $\mathbf{w}_h\in\mathbf{L}^4(\Omega_h)$ and using that $\mathbf{e}_{n,h}=\hat{\mathbf{y}}_{n,h}-\mathbf{y}_{n,h}$ and $\mathbf{e}(t)=\hat{\mathbf{e}}(t)+\mathbf{e}_{n,h}$ for all $t\in(t_{n-1},t_n)$, we get

$$c(\mathbf{e}_{n,h},\mathbf{y}_{n,h},\mathbf{e}_{n,h})=c(\mathbf{e}_{n,h},\hat{\mathbf{y}}_{n,h},\mathbf{e}_{n,h})\text{ and }c(\hat{\mathbf{y}}_{n,h},\mathbf{e}(t),\mathbf{e}_{n,h})=c(\hat{\mathbf{y}}_{n,h},\hat{\mathbf{e}}(t),\mathbf{e}_{n,h}).$$

Therefore we get

$$\begin{aligned}&c(\mathbf{y}_{n,h},\mathbf{y}_{n,h},\mathbf{e}_{n,h})-c(\mathbf{y}(t),\mathbf{y}(t),\mathbf{e}_{n,h})=\\&-c(\mathbf{e}_{n,h},\hat{\mathbf{y}}_{n,h},\mathbf{e}_{n,h})-c(\hat{\mathbf{e}}(t),\mathbf{y}(t),\mathbf{e}_{n,h})-c(\hat{\mathbf{y}}_{n,h},\hat{\mathbf{e}}(t),\mathbf{e}_{n,h}).\end{aligned}$$

The above identity and (4.21) lead to

$$\begin{aligned}&\frac{1}{2}\|\mathbf{e}_{n,h}\|^2-\frac{1}{2}\|\mathbf{e}_{n-1,h}\|^2+\frac{1}{2}\|\mathbf{e}_{n,h}-\mathbf{e}_{n-1,h}\|^2+\nu\int_{t_{n-1}}^{t_n}\|\nabla\mathbf{e}_{n,h}\|^2\,\mathrm{d}t\\&\le\int_{t_{n-1}}^{t_n}\{|c(\mathbf{e}_{n,h},\hat{\mathbf{y}}_{n,h},\mathbf{e}_{n,h})|+|c(\hat{\mathbf{e}}(t),\mathbf{y}(t),\mathbf{e}_{n,h})|+|c(\hat{\mathbf{y}}_{n,h},\hat{\mathbf{e}}(t),\mathbf{e}_{n,h})|\}\,\mathrm{d}t.\end{aligned}\tag{4.22}$$

It remains to estimate the last three terms. For the first we use that $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ is bounded in $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$ (see Lemma 4.6) and (2.3), then

$$\begin{aligned}&\int_{t_{n-1}}^{t_n}|c(\mathbf{e}_{n,h},\hat{\mathbf{y}}_{n,h},\mathbf{e}_{n,h})|\mathrm{d}t\le C\int_{t_{n-1}}^{t_n}\|\mathbf{e}_{n,h}\|\|\nabla\mathbf{e}_{n,h}\|\,\mathrm{d}t\\&\le\frac{C\tau_n}{\nu}\|\mathbf{e}_{n,h}\|^2+\frac{\nu}{4}\int_{t_{n-1}}^{t_n}\|\nabla\mathbf{e}_{n,h}\|^2\,\mathrm{d}t.\end{aligned}$$

For the second term we use that $\mathbf{y}\in L^\infty(0,T;\mathbf{H}^1(\Omega))$

$$\begin{aligned}&\int_{t_{n-1}}^{t_n}|c(\hat{\mathbf{e}}(t),\mathbf{y}(t),\mathbf{e}_{n,h})|\,\mathrm{d}t\le C\int_{t_{n-1}}^{t_n}\|\nabla\hat{\mathbf{e}}(t)\|\|\nabla\mathbf{e}_{n,h}\|\,\mathrm{d}t\\&\le\frac{C}{\nu}\int_{t_{n-1}}^{t_n}\|\nabla\hat{\mathbf{e}}(t)\|^2\,\mathrm{d}t+\frac{\nu}{4}\int_{t_{n-1}}^{t_n}\|\nabla\mathbf{e}_{n,h}\|^2\,\mathrm{d}t.\end{aligned}$$

Finally, using again the boundedness of $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ in $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$, we get the same estimate as the last one for the third term. Putting all these estimates in (4.22) we obtain

$$\begin{aligned}&(1-C\tau_n)\|\mathbf{e}_{n,h}\|^2+\|\mathbf{e}_{n,h}-\mathbf{e}_{n-1,h}\|^2+\frac{\nu}{2}\int_{t_{n-1}}^{t_n}\|\nabla\mathbf{e}_{n,h}\|^2\,\mathrm{d}t\\&\le\|\mathbf{e}_{n-1,h}\|^2+C\|\hat{\mathbf{e}}\|^2_{L^2(t_{n-1},t_n;\mathbf{H}^1(\Omega))}.\end{aligned}$$

Then, using the discrete Grönwall inequality and the fact that $\mathbf{e}_{0,h} = 0$, we get

$$\|\mathbf{e}_{n,h}\|^2 + \frac{\nu}{2}\int_0^{t_n} \|\nabla \mathbf{e}_{n,h}\|^2 \,\mathrm{d}t \le C\|\hat{\mathbf{e}}\|^2_{L^2(0,T;\mathbf{H}^1(\Omega))} \quad \forall 1 \le n \le N_\tau.$$

Then, this inequality along with (4.16) and the identity $\mathbf{y} - \mathbf{y}_\sigma = \hat{\mathbf{e}} + \mathbf{e}_\sigma$ prove (4.20). The proof of the boundedness of $\{\mathbf{y}_\sigma\}_\sigma$ in $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$ is an easy consequence of the previous results. Indeed, first we recall that $\{\hat{\mathbf{y}}_\sigma\}_\sigma$ is bounded in $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$ (Lemma 4.6). Using the inequality

$$\|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \le \|\mathbf{y}_\sigma - \hat{\mathbf{y}}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} + \|\hat{\mathbf{y}}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))},$$

it is obvious that it is enough to prove the boundedness of the first term. From an inverse inequality (see, for instance, [2, Section 4.5]), the estimates (4.16) and (4.20) and the inequality $\tau \le C_0 h^2$ we get

$$\begin{aligned}
&\|\mathbf{y}_\sigma - \hat{\mathbf{y}}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \le \frac{C}{h}\|\mathbf{y}_\sigma - \hat{\mathbf{y}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\\
&\le \frac{C}{h}\left\{\|\mathbf{y}_\sigma - \mathbf{y}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} + \|\mathbf{y} - \hat{\mathbf{y}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\right\} \le C \quad \forall \sigma.
\end{aligned}$$

To conclude the proof, we have to show the uniqueness of a solution of (4.8). Let us assume that $\mathbf{y}^1_\sigma, \mathbf{y}^2_\sigma \in \mathcal{Y}_\sigma$ are two solutions of (4.8). Then we set $\mathbf{y}_\sigma = \mathbf{y}^2_\sigma - \mathbf{y}^1_\sigma$ and we will prove that $\mathbf{y}_\sigma = 0$. Subtracting the equations (4.8) for $\mathbf{y}^2_\sigma$ and $\mathbf{y}^1_\sigma$ and setting $\mathbf{w}_h = \mathbf{y}_{n,h}$ we get

$$\left(\frac{\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}}{\tau_n}, \mathbf{y}_{n,h}\right) + a(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}) = c(\mathbf{y}^1_{n,h}, \mathbf{y}^1_{n,h}, \mathbf{y}_{n,h}) - c(\mathbf{y}^2_{n,h}, \mathbf{y}^2_{n,h}, \mathbf{y}_{n,h}).$$

Since $c(\mathbf{y}^2_{n,h}, \mathbf{y}_{n,h}, \mathbf{y}_{n,h}) = 0$, then $c(\mathbf{y}^2_{n,h}, \mathbf{y}^2_{n,h}, \mathbf{y}_{n,h}) = c(\mathbf{y}^2_{n,h}, \mathbf{y}^1_{n,h}, \mathbf{y}_{n,h})$, therefore

$$c(\mathbf{y}^1_{n,h}, \mathbf{y}^1_{n,h}, \mathbf{y}_{n,h}) - c(\mathbf{y}^2_{n,h}, \mathbf{y}^2_{n,h}, \mathbf{y}_{n,h}) = -c(\mathbf{y}_{n,h}, \mathbf{y}^1_{n,h}, \mathbf{y}_{n,h}).$$

Using this in the above identity and the boundedness of $\{\mathbf{y}^1_\sigma\}_\sigma$ in $L^\infty(0,T:\mathbf{H}^1(\Omega_h))$, we deduce

$$\begin{aligned}
&\frac{1}{2}\|\mathbf{y}_{n,h}\|^2 - \frac{1}{2}\|\mathbf{y}_{n-1,h}\|^2 + \frac{1}{2}\|\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}\|^2 + \nu\tau_n\|\nabla \mathbf{y}_{n,h}\|^2\\
&= -\tau_n c(\mathbf{y}_{n,h}, \mathbf{y}^1_{n,h}, \mathbf{y}_{n,h}) \le \tau_n \|\nabla \mathbf{y}^1_{n,h}\|\|\mathbf{y}_{n,h}\|\|\nabla \mathbf{y}_{n,h}\|\\
&\le \frac{\nu\tau_n}{2}\|\nabla \mathbf{y}_{n,h}\|^2 + \nu\|\nabla \mathbf{y}^1_{n,h}\|^2\tau_n\|\mathbf{y}_{n,h}\|^2,
\end{aligned}$$

hence

$$(1 - C\tau_n)\|\mathbf{y}_{n,h}\|^2 + \|\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}\|^2 + \nu\tau_n\|\nabla \mathbf{y}_{n,h}\|^2 \le \|\mathbf{y}_{n-1,h}\|^2.$$

Using once again the discrete Grönwall inequality and the fact that $\mathbf{y}_{0,h} = 0$, we deduce that $\mathbf{y}_\sigma = 0$. ☐

Remark 4.9. *Arguing as in Remark 4.7, we deduce from* (4.20)

$$\begin{aligned}
&\|\mathbf{y} - \mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))}\\
&\le C\left\{\left(\frac{\tau}{h} + \sqrt{\tau}\right)\|\mathbf{y}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))} + h\|\mathbf{y}\|_{L^2(0,T;\mathbf{H}^2(\Omega))} + h\|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)}\right\}.
\end{aligned}$$

Remark 4.10. *Looking at the proof of the uniqueness of a discrete solution, one realizes that if there exists a family* $\{\mathbf{y}_\sigma\}_\sigma$ *of solutions of equations* (4.8) *that is bounded in* $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$*, then* (4.8) *has a unique solution even if the hypothesis* $\tau \le C_0h^2$ *is not assumed. Indeed, only the boundedness of* $\{\mathbf{y}^1_\sigma\}_\sigma$ *in* $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$ *was required, this assumption was not necessary for* $\{\mathbf{y}^2_\sigma\}_\sigma$*.*

Hereinafter, we will assume

$$\exists C_0 > 0 \text{ such that } \tau \le C_0 h^2 \quad \forall \sigma = (\tau, h). \tag{4.23}$$

We establish a corollary of Theorem 4.8 that will be useful later.

Corollary 4.11. *Let* $\mathbf{u}, \mathbf{v} \in L^2(0,T;\mathbf{L}^2(\Omega))$ *such that*

$$\max\{\|\mathbf{u}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}, \|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\} \le M.$$

Let $\mathbf{y}_u \in \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T];\mathbf{Y})$ *be the solution of* (2.1) *and* $\mathbf{y}_\sigma(\mathbf{v}) \in \mathcal{Y}_\sigma$ *the solution of the discrete equation* (4.8) *corresponding to the control* $\mathbf{v}$*. Then, there exists a constant* $C_M > 0$ *such that*

$$\begin{aligned} &\|\mathbf{y_u} - \mathbf{y}_\sigma(\mathbf{v})\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))} + \|\mathbf{y_u} - \mathbf{y}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \\ &\le C_M \left\{ h + \|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \right\} \end{aligned} \tag{4.24}$$

Moreover, if $\mathbf{u}_\sigma \in \mathcal{U}_\sigma$ *for every* σ *and* $\mathbf{u}_\sigma \rightharpoonup \mathbf{u}$ *weakly in* $L^2(0,T;\mathbf{L}^2(\Omega))$*, then*

$$\left\{ \begin{array}{l} \|\mathbf{y_u} - \mathbf{y}_\sigma(\mathbf{u}_\sigma)\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \to 0, \\ \|\mathbf{y_u} - \mathbf{y}_\sigma(\mathbf{u}_\sigma)\|_{L^p(0,T;\mathbf{L}^2(\Omega_h))} \to 0 \quad \forall 1 \le p < +\infty, \\ \|\mathbf{y_u}(T) - \mathbf{y}_\sigma(\mathbf{u}_\sigma)(T)\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \to 0. \end{array} \right. \tag{4.25}$$

Proof. From (4.20) and (4.23), we get

$$\begin{aligned} &\|\mathbf{y_u} - \mathbf{y}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \le \|\mathbf{y_u} - \mathbf{y_v}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} + \|\mathbf{y_v} - \mathbf{y}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \\ &\le \|G(\mathbf{u}) - G(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega))} + Ch, \end{aligned}$$

where C depends on $\|\mathbf{y}_\Omega\|_{\mathbf{H}^1(\Omega)}$ and $\|\mathbf{y_v}\|_{\mathbf{H}^{2,1}(\Omega_T)}$. The last term can be estimated by $\|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}$. On the other hand, using that $G : L^2(0,T;\mathbf{L}^2(\Omega)) \longrightarrow \mathbf{H}^{2,1}(\Omega_T) \cap C([0,T];\mathbf{Y})$ is of class C^∞, we can apply the mean value theorem to get (4.24), with C_M depending on M. Using Remark 4.9, we can repeat the same argument to get the estimate in the $L^\infty(0,T;\mathbf{L}^2(\Omega_h))$.

To prove (4.25) we set $\mathbf{y_u} - \mathbf{y}_\sigma(\mathbf{u}_\sigma) = (\mathbf{y_u} - \mathbf{y}_{\mathbf{u}_\sigma}) + (\mathbf{y}_{\mathbf{u}_\sigma} - \mathbf{y}_\sigma(\mathbf{u}_\sigma))$. From the well known properties of equation (2.1) and the boundedness of $\{\mathbf{f} + \mathbf{u}_\sigma\}_\sigma$ in $L^2(0,T;\mathbf{L}^2(\Omega))$, we have that $\|\mathbf{y}_{\mathbf{u}_\sigma}\|_{\mathbf{H}^{2,1}(\Omega_T)} \le C$. Furthermore, any subsequence of $\{\mathbf{y}_{\mathbf{u}_\sigma}\}_\sigma$ weakly convergent in $\mathbf{H}^{2,1}(\Omega_T)$, converges to $\mathbf{y_u}$. This is easily proved by passing to the limit in (2.1). Then, we have that $\mathbf{y}_{\mathbf{u}_\sigma} \rightharpoonup \mathbf{y_u}$ weakly in $\mathbf{H}^{2,1}(\Omega_T)$. From the compactness of the embeddings $\mathbf{H}^{2,1}(\Omega_T) \subset L^2(0,T;\mathbf{H}^1(\Omega))$ and $\mathbf{H}^{2,1}(\Omega_T) \subset L^p(0,T;\mathbf{L}^2(\Omega))$ $(1 \le p < +\infty)$ and the compactness of the trace $\mathbf{H}^{2,1}(\Omega_T) \hookrightarrow \mathbf{L}^2(\partial\Omega_T)$ we obtain

$$\|\mathbf{y_u} - \mathbf{y}_{\mathbf{u}_\sigma}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} + \|\mathbf{y_u} - \mathbf{y}_{\mathbf{u}_\sigma}\|_{L^p(0,T;\mathbf{L}^2(\Omega_h))} + \|\mathbf{y_u}(T) - \mathbf{y}_{\mathbf{u}_\sigma}(T)\|_{\mathbf{L}^2(\Omega_h)} \to 0.$$

On the other hand, from (4.20) and (4.23) we get

$$\|\mathbf{y}_{\mathbf{u}_\sigma}(T) - \mathbf{y}_\sigma(\mathbf{u}_\sigma)(T)\| + \|\mathbf{y}_{\mathbf{u}_\sigma} - \mathbf{y}_\sigma(\mathbf{u}_\sigma)\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))} \to 0,$$

and with Remark 4.9

$$\|\mathbf{y}_{\mathbf{u}_\sigma} - \mathbf{y}_\sigma(\mathbf{u}_\sigma)\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))} \to 0,$$

which combined with the established convergences imply (4.25). □

We finish this section by studying the differentiability of the relation control - discrete state.

Theorem 4.12. *The mapping* $G_\sigma : L^2(0,T;\mathbf{L}^2(\Omega)) \longrightarrow \mathcal{Y}_\sigma$, *defined by* $G_\sigma(\mathbf{u}) = \mathbf{y}_\sigma(\mathbf{u})$ *solution of* (4.8), *is of class* C^∞. *Moreover,* $\mathbf{z}_\sigma(\mathbf{v}) = G'_\sigma(\mathbf{u})\mathbf{v}$ *is the unique solution of the problem*

$$\text{(4.26)}\qquad \left\{\begin{array}{l} \text{For } n = 1,\dots,N_\tau, \quad \text{and } \ \forall \mathbf{w}_h \in \mathbf{Y}_h, \\ \left(\dfrac{\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}}{\tau_n}, \mathbf{w}_h\right) + a(\mathbf{z}_{n,h}, \mathbf{w}_h) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_h) \\ \qquad\qquad +c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \mathbf{w}_h) = \frac{1}{\tau_n}\int_{t_{n-1}}^{t_n} (\mathbf{v}(t), \mathbf{w}_h)\,\mathrm{d}t, \\ \mathbf{z}_{0,h} = 0, \end{array}\right.$$

where we have set $\mathbf{y}_\sigma = \mathbf{y}_\sigma(\mathbf{u})$.

Proof. Let us consider the mapping

$$F_\sigma : \mathcal{Y}_\sigma \times L^2(0,T;\mathbf{L}^2(\Omega)) \longrightarrow \mathcal{Y}'_\sigma, \quad F_\sigma(\mathbf{y}_\sigma, \mathbf{u}) = \mathbf{g}_\sigma,$$

where $\mathbf{g}_\sigma$ is defined by

$$\langle \mathbf{g}_\sigma, \mathbf{w}_\sigma\rangle = \sum_{n=1}^{N_\tau} \{(\mathbf{y}_{n,h} - \mathbf{y}_{n-1,h}, \mathbf{w}_{n,h}) + \tau_n[a(\mathbf{y}_{n,h}, \mathbf{w}_{n,h}) + c(\mathbf{y}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_{n,h})]\}$$
$$-\sum_{n=1}^{N_\tau} \int_{t_{n-1}}^{t_n} (\mathbf{f}(t) + \mathbf{u}(t), \mathbf{w}_{n,h})\,\mathrm{d}t \quad \forall \mathbf{w}_\sigma \in \mathcal{Y}_\sigma,$$

where we set $\mathbf{y}_{0,h} = \mathbf{y}_{0h}$. Obviously, F_σ is of class C^∞ and $\hat{\mathbf{g}}_\sigma = \frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma, \mathbf{u})\mathbf{z}_\sigma$ is defined by

$$\langle \hat{\mathbf{g}}, \mathbf{w}_\sigma\rangle = \sum_{n=1}^{N_\tau} \{(\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}, \mathbf{w}_{n,h}) + \tau_n a(\mathbf{z}_{n,h}, \mathbf{w}_{n,h})\}$$
$$+\sum_{n=1}^{N_\tau} \{\tau_n[c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \mathbf{w}_{n,h}) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \mathbf{w}_{n,h})]\}, \quad \text{with } \mathbf{z}_{0,h} = 0.$$

On the other hand, $F_\sigma(G_\sigma(\mathbf{u}), \mathbf{u}) = F_\sigma(\mathbf{y}_\sigma(\mathbf{u}), \mathbf{u}) = 0$ for every $\mathbf{u} \in L^2(0,T;\mathbf{L}^2(\Omega))$. The proof is a consequence of the implicit function theorem, we only need to prove that

$$\frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma(\mathbf{u}), \mathbf{u}) : \mathcal{Y}_\sigma \longrightarrow \mathcal{Y}'_\sigma$$

is an isomorphism for every $\mathbf{u}$. In fact, we will prove that $\frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma, \mathbf{u})$ is an isomorphism for every $(\mathbf{y}_\sigma, \mathbf{u}) \in \mathcal{Y}_\sigma \times L^2(0,T;\mathbf{L}^2(\Omega))$. Since $\frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma, \mathbf{u})$ is a linear mapping between two finite dimension spaces of the same dimension, it is enough to prove that it is injective. Let us assume that $\frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma, \mathbf{u})\mathbf{z}_\sigma = 0$ for some $\mathbf{z}_\sigma \in \mathcal{Y}_\sigma$, We will prove that $\mathbf{z}_\sigma = 0$. Applying $\frac{\partial F_\sigma}{\partial \mathbf{y}_\sigma}(\mathbf{y}_\sigma, \mathbf{u})\mathbf{z}_\sigma \in \mathcal{Y}'_\sigma$ to $\mathbf{z}_\sigma$ and using that $c(\mathbf{y}_{n,h}, \mathbf{z}_{n,h}, \mathbf{z}_{n,h}) = 0$, we get

$$\sum_{n=1}^{N_\tau} \{(\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}, \mathbf{z}_{n,h}) + \tau_n[a(\mathbf{z}_{n,h}, \mathbf{z}_{n,h}) + c(\mathbf{z}_{n,h}, \mathbf{y}_{n,h}, \mathbf{z}_{n,h})]\} = 0.$$

Hence,

$$\begin{aligned}
&\sum_{n=1}^{N_\tau} \left\{ \frac{1}{2}\|\mathbf{z}_{n,h}\|^2 - \frac{1}{2}\|\mathbf{z}_{n-1,h}\|^2 + \frac{1}{2}\|\mathbf{z}_{n,h} - \mathbf{z}_{n-1,h}\|^2 + \nu\tau_n\|\nabla \mathbf{z}_{n,h}\|^2 \right\} \\
&\le \sum_{n=1}^{N_\tau} \tau_n \|\nabla \mathbf{y}_{n,h}\| \|\mathbf{z}_{n,h}\| \|\nabla \mathbf{z}_{n,h}\| \\
&\le \frac{1}{\nu}\|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \sum_{n=1}^{N_\tau} \tau_n \|\mathbf{z}_{n,h}\|^2 + \frac{\nu}{2}\sum_{n=1}^{N_\tau} \tau_n \|\nabla \mathbf{z}_{n,h}\|^2.
\end{aligned}$$

Again, an application of the discrete Grönwall inequality and the fact that $\mathbf{z}_{0,h} = 0$ imply that $\mathbf{z}_\sigma = 0$. □

4.2. Analysis of the discrete adjoint state equation. Along this section, as well as in the rest of the paper, the condition (4.23) is assumed. As a consequence of Theorem 4.12 and applying the chain rule, we get that $J_\sigma : L^2(0, T; \mathbf{L}^2(\Omega)) \longrightarrow \mathbb{R}$ is of class C^∞ and we have a first expression of its derivative as follows

$$\begin{aligned}
J'_\sigma(\mathbf{u})\mathbf{v} = &\int_0^T \int_{\Omega_h} (\mathbf{y}_\sigma - \mathbf{y}_d)\mathbf{z}_\sigma \,\mathrm{d}x\mathrm{d}t \\
&+\gamma \int_{\Omega_h} (\mathbf{y}_\sigma(T) - \mathbf{y}_d)\mathbf{z}_\sigma(T)\,\mathrm{d}x + \lambda \int_0^T \int_{\Omega_h} \mathbf{u}\mathbf{v}\,\mathrm{d}x\mathrm{d}t,
\end{aligned}$$

where $\mathbf{y}_\sigma = \mathbf{y}_\sigma(\mathbf{u}) = G_\sigma(\mathbf{u})$ and $\mathbf{z}_\sigma = G'_\sigma(\mathbf{u})\mathbf{v}$ is the solution of (4.26). As usual in control theory, we have to introduce the adjoint state to simplify the expression of this derivative. To this end we consider the discrete adjoint state equation: we look for $\boldsymbol{\varphi}_\sigma \in \mathcal{Y}_\sigma$ such that

$$\text{(4.27)} \qquad \begin{cases}
\text{for } n = N_\tau, \dots, 1, \quad \text{and} \quad \forall \mathbf{w}_h \in \mathbf{Y}_h, \\
\left(\dfrac{\boldsymbol{\varphi}_{n,h} - \boldsymbol{\varphi}_{n+1,h}}{\tau_n}, \mathbf{w}_h\right) + a(\boldsymbol{\varphi}_{n,h}, \mathbf{w}_h) + c(\mathbf{w}_h, \mathbf{y}_{n,h}, \boldsymbol{\varphi}_{n,h}) \\
+c(\mathbf{y}_{n,h}, \mathbf{w}_h, \boldsymbol{\varphi}_{n,h}) = \frac{1}{\tau_n}\int_{t_{n-1}}^{t_n} (\mathbf{y}_{n,h} - \mathbf{y}_d(t), \mathbf{w}_h)\,\mathrm{d}t, \\
\boldsymbol{\varphi}_{N_\tau+1,h} = \gamma(\mathbf{y}_{N_\tau,h} - \mathbf{y}_{\Omega_h}).
\end{cases}$$

Observe that in the above system, first we compute $\boldsymbol{\varphi}_{N_\tau,h}$ from $\boldsymbol{\varphi}_{N_\tau+1,h} = \gamma(\mathbf{y}_{N_\tau,h} - \mathbf{y}_{\Omega_h})$ and then we descend in n until $n = 1$. Unlike the discrete states $\mathbf{y}_\sigma$, we will set for the discrete adjoint states $\boldsymbol{\varphi}_\sigma(t_{n-1}) = \boldsymbol{\varphi}_{n,h}$ for every $1 \le n \le N_\tau$.

System (4.27) corresponds to the discretization of the backward equation (3.8). Using that $\{\mathbf{y}_\sigma\}_\sigma$ is bounded in $L^\infty(0, T; \mathbf{H}^1(\Omega_h))$ (Theorem (4.8)), then we can proceed in the same way as we did in the proof of Theorem 4.12 to obtain the existence and uniqueness of a solution of (4.27). Below we check that this is actually the discrete adjoint state equation. To this end we use (4.26) and (4.27) to show that

$$
\begin{aligned}
&\int_0^T\int_{\Omega_h}(\mathbf{y}_\sigma-\mathbf{y}_d)\mathbf{z}_\sigma\,\mathrm{d}x\mathrm{d}t\\
&=\sum_{n=1}^{N_\tau}\int_{t_{n-1}}^{t_n}(\mathbf{y}_{n,h}-\mathbf{y}_d(t),\mathbf{z}_{n,h})\,\mathrm{d}t=\sum_{n=1}^{N_\tau}(\boldsymbol{\varphi}_{n,h}-\boldsymbol{\varphi}_{n+1,h},\mathbf{z}_{n,h})\\
&+\sum_{n=1}^{N_\tau}\tau_n[a(\boldsymbol{\varphi}_{n,h},\mathbf{z}_{n,h})+c(\mathbf{z}_{n,h},\mathbf{y}_{n,h},\boldsymbol{\varphi}_{n,h})+c(\mathbf{y}_{n,h},\mathbf{z}_{n,h},\boldsymbol{\varphi}_{n,h})]\\
&=\sum_{n=1}^{N_\tau}(\mathbf{z}_{n,h}-\mathbf{z}_{n-1,h},\boldsymbol{\varphi}_{n,h})-(\boldsymbol{\varphi}_{N_\tau+1,h},\mathbf{z}_{N_\tau,h})+(\boldsymbol{\varphi}_{1,h},\mathbf{z}_{0,h})\\
&+\sum_{n=1}^{N_\tau}\tau_n[a(\mathbf{z}_{n,h},\boldsymbol{\varphi}_{n,h})+c(\mathbf{z}_{n,h},\mathbf{y}_{n,h},\boldsymbol{\varphi}_{n,h})+c(\mathbf{y}_{n,h},\mathbf{z}_{n,h},\boldsymbol{\varphi}_{n,h})]\\
&=\int_0^T\int_{\Omega_h}\mathbf{v}\boldsymbol{\varphi}_\sigma\,\mathrm{d}x\mathrm{d}t-\gamma\int_{\Omega_h}(\mathbf{y}_\sigma(T)-\mathbf{y}_{\Omega_h})\mathbf{z}_\sigma(T)\,\mathrm{d}x,
\end{aligned}
$$

where we have used that $\boldsymbol{\varphi}_{N_\tau+1,h}=\gamma(\mathbf{y}_{N_\tau,h}-\mathbf{y}_{\Omega_h})=\gamma(\mathbf{y}_\sigma(T)-\mathbf{y}_{\Omega_h})$ and $\mathbf{z}_{0,h}=0$. From the obtained identity and the expression of $J'_\sigma(\mathbf{u})\mathbf{v}$ given above we conclude

$$
J'_\sigma(\mathbf{u})\mathbf{v}=\int_0^T\int_{\Omega_h}(\boldsymbol{\varphi}_\sigma+\lambda\mathbf{u})\mathbf{v}\,\mathrm{d}x\mathrm{d}t. \tag{4.28}
$$

The next theorem states the error estimates in the approximation of the adjoint state equation.

THEOREM 4.13. *Given* $\mathbf{u}\in L^2(0,T;\mathbf{L}^2(\Omega))$, *let* $\mathbf{y}=\mathbf{y}_\mathbf{u}$ *be the associated state, solution of* (2.1), $\boldsymbol{\varphi}$ *the associated adjoint state, solution of* (3.8), $\mathbf{y}_\sigma=\mathbf{y}_\sigma(\mathbf{u})$ *the associated discrete state, solution of* (4.8), *and* $\boldsymbol{\varphi}_\sigma$ *the associated discrete adjoint state, solution of* (4.27). *Then,* $\{\boldsymbol{\varphi}_\sigma\}_\sigma$ *is bounded in* $L^\infty(0,T;\mathbf{H}^1(\Omega_h))$ *and there exists a constant* $C>0$ *independent of* σ *and* $\mathbf{u}$ *such that*

$$
\begin{aligned}
&\|\boldsymbol{\varphi}-\boldsymbol{\varphi}_\sigma\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))}+\|\boldsymbol{\varphi}-\boldsymbol{\varphi}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\\
&\le Ch\left\{\|\mathbf{u}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}+\|\mathbf{y}_0\|_{\mathbf{H}^1(\Omega)}+\|\mathbf{y}_\Omega\|_{\mathbf{H}^1(\Omega)}\right\}.
\end{aligned} \tag{4.29}
$$

Proof. Let us consider the projection operator $R_\sigma:C([0,T];\mathbf{L}^2(\Omega))\longrightarrow\mathcal{Y}_\sigma$ by $(R_\sigma\mathbf{w})_{n,h}=P_h\mathbf{w}(t_{n-1})$ for $1\le n\le N_\tau$, with P_h given in Definition 4.1. As for the discrete adjoint states, we fix $(R_\sigma\mathbf{w})(t_{n-1})=(R_\sigma\mathbf{w})_{n,h}$. Analogously to (4.12) and (4.14), we have the estimates for every $\mathbf{w}\in\mathbf{H}^{2,1}(\Omega_T)\cap C([0,T],\mathbf{Y})$

$$
\|\mathbf{w}-R_\sigma\mathbf{w}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\le C\left\{\tau\|\mathbf{w}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))}+h^2\|\mathbf{w}\|_{L^2(0,T;\mathbf{H}^2(\Omega))}\right\}, \tag{4.30}
$$

$$
\|\mathbf{w}-R_\sigma\mathbf{w}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\le C\left\{\frac{\tau}{h}\|\mathbf{w}'\|_{L^2(0,T;\mathbf{L}^2(\Omega))}+h\|\mathbf{w}\|_{L^2(0,T;\mathbf{H}^2(\Omega))}\right\}. \tag{4.31}
$$

We set $\boldsymbol{\epsilon}=\boldsymbol{\varphi}-\boldsymbol{\varphi}_\sigma=(\boldsymbol{\varphi}-R_\sigma\boldsymbol{\varphi})+(R_\sigma\boldsymbol{\varphi}-\boldsymbol{\varphi}_\sigma)=\boldsymbol{\eta}+\boldsymbol{\epsilon}_\sigma$. According to our notation above fixed, we have $\boldsymbol{\eta}(t_n)=\boldsymbol{\varphi}(t_n)-(R_\sigma\boldsymbol{\varphi})(t_n)=\boldsymbol{\varphi}(t_n)-(R_\sigma\boldsymbol{\varphi})_{n+1,h}=\boldsymbol{\varphi}(t_n)-P_h\boldsymbol{\varphi}(t_n)$, for $0\le n\le N_\tau-1$. Also we have $\boldsymbol{\epsilon}_\sigma(t_n)=\boldsymbol{\epsilon}_{n+1,h}$, $0\le n\le N_\tau-1$.

Setting $(R_\sigma \mathbf{w})_{N_\tau+1,h} = P_h \mathbf{w}(T)$ and recalling that $\boldsymbol{\varphi}_{N_\tau+1,h} = \gamma(\mathbf{y}_{N_\tau,h} - \mathbf{y}_{\Omega_h})$, then the previous identities are also well defined for $n = N_\tau$. Then, (3.8) and (4.27) lead to the identities, $n = N_\tau, \dots, 1$,

$$
\begin{aligned}
&(\boldsymbol{\epsilon}(t_{n-1}) - \boldsymbol{\epsilon}(t_n), \mathbf{w}_h) + \int_{t_{n-1}}^{t_n} a(\boldsymbol{\epsilon}(t), \mathbf{w}_h)\,\mathrm{d}t \\
&+ \int_{t_{n-1}}^{t_n} [c(\mathbf{y}(t), \mathbf{w}_h, \boldsymbol{\varphi}(t)) + c(\mathbf{w}_h, \mathbf{y}(t), \boldsymbol{\varphi}(t))]\,\mathrm{d}t \\
&- \int_{t_{n-1}}^{t_n} [c(\mathbf{y}_{n,h}, \mathbf{w}_h, \boldsymbol{\varphi}_{n,h}) + c(\mathbf{w}_h, \mathbf{y}_{n,h}, \boldsymbol{\varphi}_{n,h})]\,\mathrm{d}t = \int_{t_{n-1}}^{t_n} (\mathbf{y}(t) - \mathbf{y}_{n,h}, \mathbf{w}_h)\mathrm{d}t.
\end{aligned}
$$

Now, writing $\boldsymbol{\epsilon} = \boldsymbol{\eta} + \boldsymbol{\epsilon}_\sigma$ and taking into account that

$$
(\boldsymbol{\eta}(t_n), \mathbf{w}_h) = (\boldsymbol{\varphi}(t_n) - P_h\boldsymbol{\varphi}(t_n), \mathbf{w}_h) = 0 \quad \forall \mathbf{w}_h \in \mathbf{Y}_h \text{ and for } 0 \le n \le N_\tau,
$$

we obtain for $\mathbf{w}_h = \boldsymbol{\epsilon}_{n,h}$

$$
\begin{aligned}
&(\boldsymbol{\epsilon}_{n,h} - \boldsymbol{\epsilon}_{n+1,h}, \boldsymbol{\epsilon}_{n,h}) + \int_{t_{n-1}}^{t_n} a(\boldsymbol{\epsilon}_{n,h}, \boldsymbol{\epsilon}_{n,h})\,\mathrm{d}t \\
&= \int_{t_{n-1}}^{t_n} (\mathbf{y}(t) - \mathbf{y}_{n,h}, \boldsymbol{\epsilon}_{n,h})\,\mathrm{d}t - \int_{t_{n-1}}^{t_n} a(\boldsymbol{\eta}(t), \boldsymbol{\epsilon}_{n,h})\,\mathrm{d}t \\
&+ \int_{t_{n-1}}^{t_n} [c(\mathbf{y}_{n,h}, \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}_{n,h}) + c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_{n,h}, \boldsymbol{\varphi}_{n,h})]\,\mathrm{d}t \\
&- \int_{t_{n-1}}^{t_n} [c(\mathbf{y}(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}(t)) + c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}(t), \boldsymbol{\varphi}(t))]\,\mathrm{d}t,
\end{aligned}
$$

hence

$$
\begin{aligned}
&\frac{1}{2}\|\boldsymbol{\epsilon}_{n,h}\|^2 - \frac{1}{2}\|\boldsymbol{\epsilon}_{n+1,h}\|^2 + \frac{1}{2}\|\boldsymbol{\epsilon}_{n,h} - \boldsymbol{\epsilon}_{n+1,h}\|^2 + \nu \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2 \mathrm{d}t \\
&\le \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|\|\boldsymbol{\epsilon}_{n,h}\|\,\mathrm{d}t + \nu \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\eta}(t)\|\|\nabla \boldsymbol{\epsilon}_{n,h}\|\,\mathrm{d}t \\
&+ \int_{t_{n-1}}^{t_n} [c(\mathbf{y}_{n,h}, \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}_{n,h}) - c(\mathbf{y}(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}(t))]\,\mathrm{d}t \\
&+ \int_{t_{n-1}}^{t_n} [c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_{n,h}, \boldsymbol{\varphi}_{n,h}) - c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}(t), \boldsymbol{\varphi}(t))]\,\mathrm{d}t. \qquad (4.32)
\end{aligned}
$$

Let us estimate the right hand side of (4.32).

$$
\begin{aligned}
&\int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|\|\boldsymbol{\epsilon}_{n,h}\|\,\mathrm{d}t + \nu \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\eta}(t)\|\nabla \boldsymbol{\epsilon}_{n,h}\|\,\mathrm{d}t \\
&\le \frac{1}{2}\int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|^2\,\mathrm{d}t + \frac{\tau_n}{2}\|\boldsymbol{\epsilon}_{n,h}\|^2 \\
&+ \frac{2}{\nu}\int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\eta}(t)\|^2\,\mathrm{d}t + \frac{\nu}{8}\int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2\,\mathrm{d}t.
\end{aligned}
$$

Now we proceed with the second term

$$
\begin{aligned}
&\left|\int_{t_{n-1}}^{t_n} [c(\mathbf{y}_{n,h}, \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}_{n,h}) - c(\mathbf{y}(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}(t))] \,\mathrm{d}t\right| \\
&= \left|\int_{t_{n-1}}^{t_n} [c(\mathbf{y}_\sigma(t) - \mathbf{y}(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}(t)) - c(\mathbf{y}_\sigma(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\epsilon}(t))] \,\mathrm{d}t\right| \\
&= \left|\int_{t_{n-1}}^{t_n} [c(\mathbf{y}_\sigma(t) - \mathbf{y}(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\varphi}(t)) - c(\mathbf{y}_\sigma(t), \boldsymbol{\epsilon}_{n,h}, \boldsymbol{\eta}(t))] \,\mathrm{d}t\right| \\
&\leq \|\boldsymbol{\varphi}\|_{L^\infty(0,T;\mathbf{H}^1(\Omega))} \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|_{\mathbf{H}^1(\Omega_h)} \|\nabla \boldsymbol{\epsilon}_{n,h}\| \,\mathrm{d}t \\
&+ \|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \int_{t_{n-1}}^{t_n} \|\boldsymbol{\eta}(t)\|_{\mathbf{H}^1(\Omega_h)} \|\nabla \boldsymbol{\epsilon}_{n,h}\| \,\mathrm{d}t \\
&\leq C \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|^2_{\mathbf{H}^1(\Omega_h)} \,\mathrm{d}t + \frac{\nu}{8} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2 \,\mathrm{d}t \\
&+ C \int_{t_{n-1}}^{t_n} \|\boldsymbol{\eta}(t)\|^2_{\mathbf{H}^1(\Omega_h)} \,\mathrm{d}t + \frac{\nu}{8} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2 \,\mathrm{d}t.
\end{aligned}
$$

For the last term of (4.32), we first observe that

$$
\begin{aligned}
&c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_{n,h}, \boldsymbol{\varphi}_{n,h}) - c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}(t), \boldsymbol{\varphi}(t)) \\
&= -[c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}(t) - \mathbf{y}_\sigma(t), \boldsymbol{\varphi}(t)) + c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_\sigma(t), \boldsymbol{\eta}(t)) + c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_\sigma(t), \boldsymbol{\epsilon}_{n,h})].
\end{aligned}
$$

The first two terms can be estimated in a similar way to the previous one

$$
\begin{aligned}
&\left|\int_{t_{n-1}}^{t_n} [c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}(t) - \mathbf{y}_\sigma(t), \boldsymbol{\varphi}(t)) + c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_\sigma(t), \boldsymbol{\eta}(t))] \,\mathrm{d}t\right| \\
&\leq C \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|^2_{\mathbf{H}^1(\Omega_h)} \,\mathrm{d}t + \frac{\nu}{8} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2 \,\mathrm{d}t \\
&+ C \int_{t_{n-1}}^{t_n} \|\boldsymbol{\eta}(t)\|^2_{\mathbf{H}^1(\Omega_h)} \,\mathrm{d}t + \frac{\nu}{8} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2 \,\mathrm{d}t.
\end{aligned}
$$

Finally,

$$
\begin{aligned}
&\left|\int_{t_{n-1}}^{t_n} c(\boldsymbol{\epsilon}_{n,h}, \mathbf{y}_\sigma(t), \boldsymbol{\epsilon}_{n,h}) \,\mathrm{d}t\right| \leq \sqrt{2} \|\mathbf{y}_\sigma\|_{L^\infty(0,T;\mathbf{H}^1(\Omega_h))} \int_{t_{n-1}}^{t_n} \|\boldsymbol{\epsilon}_{n,h}\| \|\nabla \boldsymbol{\epsilon}_{n,h}\| \,\mathrm{d}t \\
&\leq C\tau_n \|\boldsymbol{\epsilon}_{n,h}\|^2 + \frac{\nu}{8} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2.
\end{aligned}
$$

Collecting all the estimates, we infer from (4.32)

$$
\begin{aligned}
&(1 - C\tau_n)\|\boldsymbol{\epsilon}_{n,h}\|^2 + \|\boldsymbol{\epsilon}_{n,h} - \boldsymbol{\epsilon}_{n+1,h}\|^2 + \frac{3\nu}{4} \int_{t_{n-1}}^{t_n} \|\nabla \boldsymbol{\epsilon}_{n,h}\|^2 \mathrm{d}t \\
&\|\boldsymbol{\epsilon}_{n+1,h}\|^2 + C \left\{ \int_{t_{n-1}}^{t_n} \|\boldsymbol{\eta}(t)\|^2_{\mathbf{H}^1(\Omega_h)} \,\mathrm{d}t + \int_{t_{n-1}}^{t_n} \|\mathbf{y}(t) - \mathbf{y}_\sigma(t)\|^2 \,\mathrm{d}t \right\}.
\end{aligned}
$$

To conclude the proof it is enough to use the discrete Grönwall inequality along with (4.11), (4.20), (4.23), (4.31) and the fact that the $\mathbf{H}^{2,1}(\Omega_T)$ norm of $\boldsymbol{\varphi}$ can be estimated by the $L^2(0,T;\mathbf{L}^2(\Omega))$ norm of $\mathbf{y}-\mathbf{y}_d$ and the $\mathbf{H}^1(\Omega)$ norm of $\mathbf{y}_\Omega$, and the $L^2(0,T;\mathbf{L}^2(\Omega))$ norm of $\mathbf{y}$ is estimated by the $L^2(0,T;\mathbf{L}^2(\Omega))$ norm of $\mathbf{u}$. □

As a consequence of the previous theorem we have the following result analogous to Corollary 4.11.

Corollary 4.14. *Let* $\mathbf{u},\mathbf{v}\in L^2(0,T;\mathbf{L}^2(\Omega))$ *such that*

$$\max\{\|\mathbf{u}\|_{L^2(0,T;\mathbf{L}^2(\Omega))},\|\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\}\le M.$$

Let $\varphi_u\in\mathbf{H}^{2,1}(\Omega_T)\cap C([0,T];\mathbf{Y})$ *be the solution of* (3.8) *and* $\boldsymbol{\varphi}_\sigma(\mathbf{v})\in\mathcal{Y}_\sigma$ *the solution of the discrete equation* (4.27) *corresponding to the control* $\mathbf{v}$. *Then, there exists a constant* $C_M>0$ *such that*

$$\begin{aligned}&\|\boldsymbol{\varphi}_{\mathbf{u}}-\boldsymbol{\varphi}_\sigma(\mathbf{v})\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))}+\|\boldsymbol{\varphi}_{\mathbf{u}}-\boldsymbol{\varphi}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\\ &\le C_M\left\{h+\|\mathbf{u}-\mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\right\}\end{aligned}\tag{4.33}$$

Proof. First we observe that (4.29) implies

$$\begin{aligned}&\|\boldsymbol{\varphi}_{\mathbf{u}}-\boldsymbol{\varphi}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\le\|\boldsymbol{\varphi}_{\mathbf{u}}-\boldsymbol{\varphi}_{\mathbf{v}}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\\ &+\|\boldsymbol{\varphi}_{\mathbf{v}}-\boldsymbol{\varphi}_\sigma(\mathbf{v})\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\le\|\boldsymbol{\varphi}_{\mathbf{u}}-\boldsymbol{\varphi}_{\mathbf{v}}\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}+Ch,\end{aligned}$$

where C depends on $\|\mathbf{u}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}$. We proceed analogously to get the estimate for $\|\boldsymbol{\varphi}_{\mathbf{u}}-\boldsymbol{\varphi}_\sigma(\mathbf{v})\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))}$. Now, we estimate $\boldsymbol{\varphi}_{\mathbf{u}}-\boldsymbol{\varphi}_{\mathbf{v}}$ in $L^2(0,T;\mathbf{H}^1(\Omega))$ and $L^\infty(0,T;\mathbf{L}^2(\Omega))$, respectively. Let us set $\boldsymbol{\varphi}=\boldsymbol{\varphi}_{\mathbf{u}}-\boldsymbol{\varphi}_{\mathbf{v}}$, then subtracting the equations satisfied by $\boldsymbol{\varphi}_{\mathbf{u}}$ and $\boldsymbol{\varphi}_{\mathbf{v}}$, we get

$$\begin{aligned}&-(\boldsymbol{\varphi}_t,\mathbf{w})+a(\boldsymbol{\varphi},\mathbf{w})=(\mathbf{y}_{\mathbf{u}}-\mathbf{y}_{\mathbf{v}},\mathbf{w})\\ &+c(\mathbf{w},\mathbf{y}_{\mathbf{v}},\boldsymbol{\varphi}_{\mathbf{v}})+c(\mathbf{y}_{\mathbf{v}},\mathbf{w},\boldsymbol{\varphi}_{\mathbf{v}})-c(\mathbf{w},\mathbf{y}_{\mathbf{u}},\boldsymbol{\varphi}_{\mathbf{u}})-c(\mathbf{y}_{\mathbf{u}},\mathbf{w},\boldsymbol{\varphi}_{\mathbf{u}}).\end{aligned}$$

Taking $\mathbf{w}=\boldsymbol{\varphi}$ and using the identities

$$\begin{aligned}&c(\boldsymbol{\varphi},\mathbf{y}_{\mathbf{v}},\boldsymbol{\varphi}_{\mathbf{v}})-c(\boldsymbol{\varphi},\mathbf{y}_{\mathbf{u}},\boldsymbol{\varphi}_{\mathbf{u}})=c(\boldsymbol{\varphi},\mathbf{y}_{\mathbf{v}}-\mathbf{y}_{\mathbf{u}},\boldsymbol{\varphi}_{\mathbf{v}})-c(\boldsymbol{\varphi},\mathbf{y}_{\mathbf{u}},\boldsymbol{\varphi})\\ &c(\mathbf{y}_{\mathbf{v}},\boldsymbol{\varphi},\boldsymbol{\varphi}_{\mathbf{v}})-c(\mathbf{y}_{\mathbf{u}},\boldsymbol{\varphi},\boldsymbol{\varphi}_{\mathbf{u}})=c(\mathbf{y}_{\mathbf{v}}-\mathbf{y}_{\mathbf{u}},\boldsymbol{\varphi},\boldsymbol{\varphi}_{\mathbf{v}}),\end{aligned}$$

we deduce by integration in the interval (t,T) and the equality $\boldsymbol{\varphi}(T)=\boldsymbol{\varphi}_{\mathbf{u}}(T)-\boldsymbol{\varphi}_{\mathbf{v}}(T)=\gamma(\mathbf{y}_{\mathbf{u}}(T)-\mathbf{y}_{\mathbf{v}}(T))$

$$\begin{aligned}&\frac{1}{2}\|\boldsymbol{\varphi}(t)\|^2-\gamma^2\frac{1}{2}\|\mathbf{y}_{\mathbf{u}}(T)-\mathbf{y}_{\mathbf{v}}(T)\|^2+\nu\int_t^T\|\nabla\boldsymbol{\varphi}(s)\|^2\,\mathrm{d}s\\ &\le\int_t^T\|\mathbf{y}_{\mathbf{u}}(s)-\mathbf{y}_{\mathbf{v}}(s)\|\|\boldsymbol{\varphi}(s)\|\,\mathrm{d}s\\ &+C\int_t^T\|\left\{\|\boldsymbol{\varphi}(s)\|^{1/2}\|\nabla\boldsymbol{\varphi}(s)\|^{1/2}\|\nabla\mathbf{y}_{\mathbf{v}}(s)-\nabla\mathbf{y}_{\mathbf{u}}(s)\|\|\boldsymbol{\varphi}_{\mathbf{v}}(s)\|_{\mathbf{H}^1(\Omega)}\right\}\mathrm{d}s\\ &+C\int_t^T\|\boldsymbol{\varphi}(s)\|\|\nabla\boldsymbol{\varphi}(s)\|\|\nabla\mathbf{y}_{\mathbf{u}}\|\,\mathrm{d}s\\ &+C\int_t^T\|\mathbf{y}_{\mathbf{v}}(s)-\mathbf{y}_{\mathbf{u}}(s)\|_{\mathbf{H}^1(\Omega)}\|\nabla\boldsymbol{\varphi}(s)\|\|\boldsymbol{\varphi}_{\mathbf{v}}(s)\|_{\mathbf{H}^1(\Omega)}\,\mathrm{d}s.\end{aligned}$$

Since $\mathbf{y_u}, \boldsymbol{\varphi}_\mathbf{v} \in L^\infty(0,T;\mathbf{H}^1(\Omega))$, with norms estimated by a constant depending on M, we infer from the above inequality

$$\begin{aligned}
&\frac{1}{2}\|\boldsymbol{\varphi}(t)\|^2 - \gamma^2\frac{1}{2}\|\mathbf{y_u}(T) - \mathbf{y_v}(T)\|^2 + \nu\int_t^T \|\nabla\boldsymbol{\varphi}(s)\|^2 \,\mathrm{d}s \\
&\leq \frac{1}{2}\int_t^T \|\mathbf{y_u}(s) - \mathbf{y_v}(s)\|^2 \,\mathrm{d}s + \frac{1}{2}\int_t^T \|\boldsymbol{\varphi}(s)\|^2 \,\mathrm{d}s \\
&+C\int_t^T \Big\{\|\boldsymbol{\varphi}(s)\|^2 + \|\mathbf{y_v}(s) - \mathbf{y_u}(s)\|^2_{\mathbf{H}^1(\Omega)}\Big\} \,\mathrm{d}s + \frac{\nu}{2}\int_t^T \|\nabla\boldsymbol{\varphi}(s)\|^2 \,\mathrm{d}s.
\end{aligned}$$

On the other hand, we have

$$\begin{aligned}
&\|\mathbf{y_u} - \mathbf{y_v}\|_{\mathbf{H}^{2,1}(\Omega_T)} = \|G(\mathbf{u}) - G(\mathbf{v})\|_{\mathbf{H}^{2,1}(\Omega_T)} \\
&\leq \sup_{0\leq\rho\leq 1}\|G'(\mathbf{u} + \rho(\mathbf{v} - \mathbf{u}))\|\|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \leq C\|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))},
\end{aligned}$$

where C depends on M. The last two inequalities lead to

$$\begin{aligned}
&\|\boldsymbol{\varphi}(t)\|^2 + \nu\int_t^T \|\nabla\boldsymbol{\varphi}(s)\|^2 \,\mathrm{d}s \\
&\leq C\left\{\|\mathbf{y_u}(T) - \mathbf{y_v}(T)\|^2 + \|\mathbf{y_u} - \mathbf{y_v}\|^2_{L^2(0,T;\mathbf{H}^1(\Omega))} + \int_t^T \|\boldsymbol{\varphi}(s)\|^2 \,\mathrm{d}s\right\} \\
&\leq C\left\{\|\mathbf{y_u} - \mathbf{y_v}\|^2_{\mathbf{H}^{2,1}(\Omega_T)} + \int_t^T \|\boldsymbol{\varphi}(s)\|^2 \,\mathrm{d}s\right\} \\
&\leq C\left\{\|\mathbf{u} - \mathbf{v}\|^2_{L^2(0,T;\mathbf{L}^2(\Omega))} + \int_t^T \|\boldsymbol{\varphi}(s)\|^2 \,\mathrm{d}s\right\} \quad \forall t \in [0,T].
\end{aligned}$$

Now the Grönwall inequality implies

$$\|\boldsymbol{\varphi}(t)\| \leq C\|\mathbf{u} - \mathbf{v}\|_{L^2(0,T;\mathbf{L}^2(\Omega))} \quad \forall t \in [0,T],$$

which also implies with the aid of the previous estimates

$$\nu\int_0^T \|\nabla\boldsymbol{\varphi}(s)\|^2 \,\mathrm{d}s \leq C\|\mathbf{u} - \mathbf{v}\|^2_{L^2(0,T;\mathbf{L}^2(\Omega))},$$

which concludes the proof. □

4.3. Convergence of the discrete control problem. In this section we analyze the convergence of the solutions of control problems (P$_\sigma$) towards solutions of the continuous problem (P). Since these problems are not convex, we will also address the issue of the approximation of local solutions of problem (P). It is clear that every problem (P$_\sigma$) has at least one solution because it consists of the minimization of a continuous and coercive function on a nonempty closed subset of a finite dimensional space. The next theorem proves the convergence of these discrete solutions to solutions of problem (P).

THEOREM 4.15. *For every $\sigma = (\tau, h)$ let $\bar{\mathbf{u}}_\sigma$ be a global solution of problem* (P$_\sigma$), *then the sequence $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ is bounded in $L^2(0,T;\mathbf{L}^2(\Omega))$ and there exist subsequences,* *denoted in the same way, converging to a point* $\bar{\mathbf{u}}$ *weakly in* $L^2(0,T;\mathbf{L}^2(\Omega))$. *Any of these limit points is a solution of problem* (P). *Moreover, we have*

$$\lim_{\sigma\to 0}\|\bar{\mathbf{u}}-\bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}=0 \quad \text{and} \quad \lim_{\sigma\to 0}J_\sigma(\bar{\mathbf{u}}_\sigma)=J(\bar{\mathbf{u}}). \tag{4.34}$$

REMARK 4.16. *Strictly speaking, it is not correct to claim that sequence* $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ *is bounded in* $L^2(0,T;\mathbf{L}^2(\Omega))$ *because* $\bar{\mathbf{u}}_\sigma$ *is only defined in* $(0,T)\times\Omega_h$, *with* $\Omega_h\subset\Omega$, *for* $\sigma=(\tau,h)$. *We will prove that* $\|\mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\le C$ *for some constant independent of* σ. *Now, if we take any element* $\mathbf{v}\in L^2(0,T;\mathbf{L}^2(\Omega))$ *and we extend every* $\bar{\mathbf{u}}_\sigma$ *to* $(0,T)\times\Omega$ *by setting* $\bar{\mathbf{u}}_\sigma(t,x)=\mathbf{v}(t,x)$ *for every* $(t,x)\in(0,T)\times(\Omega\setminus\Omega_h)$, *then we have that these extensions constitute a sequence of bounded functions in* $L^2(0,T;\mathbf{L}^2(\Omega))$ *and every weak limit point is a solution of* (P), *it does not matter the choice of* $\mathbf{v}$. *This is a consequence of the property* (4.1). *The theorem should be understood in this sense.*

Proof. Let $\tilde{\mathbf{u}}$ be a solution of problem (P) and let us take $\mathbf{u}_\sigma\in\mathcal{U}_\sigma$ defined by

$$\mathbf{u}_\sigma=\sum_{n=1}^{N_\tau}\sum_{T\in\mathcal{T}_h}\mathbf{u}_{n,T}\chi_n\chi_T,\quad \text{with}\quad \mathbf{u}_{n,T}=\frac{1}{\tau_n|T|}\int_{t_{n-1}}^{t_n}\int_T\tilde{\mathbf{u}}(t,x)\,\mathrm{d}x\mathrm{d}t. \tag{4.35}$$

Then, $\mathbf{u}_\sigma$ is the $L^2(0,T;\mathbf{L}^2(\Omega_h))$ projection of $\tilde{\mathbf{u}}$ on $\mathcal{Y}_\sigma$. From our assumptions (A1)-(A3), we have that $\|\tilde{\mathbf{u}}-\mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\to 0$ when $\sigma\to 0$. Using Corollary 4.11, we deduce easily that $J_\sigma(\mathbf{u}_\sigma)\to J(\tilde{\mathbf{u}})$. On the other hand, it is immediate that $\mathbf{u}_\sigma\in\mathcal{U}_{\sigma,ad}$ for every σ, then the optimality of $\bar{\mathbf{u}}_\sigma$ and the definition of J_σ lead to

$$\frac{\lambda}{2}\|\bar{\mathbf{u}}_\sigma\|^2\le J_\sigma(\bar{\mathbf{u}}_\sigma)\le J_\sigma(\mathbf{u}_\sigma)\le C\quad\forall\sigma.$$

Therefore, we deduce the existence of subsequences weakly convergent. Let $\bar{\mathbf{u}}$ be one of these limit points. Obviously the property $\bar{\mathbf{u}}\in\mathcal{U}_{ad}$ holds. Moreover, using again Corollary 4.11 and the convexity of the cost functional in the third term involving the control, we have

$$\inf(\mathrm{P})\ \le J(\bar{\mathbf{u}})\le\liminf_{\sigma\to 0}J_\sigma(\bar{\mathbf{u}}_\sigma)\le\limsup_{\sigma\to 0}J_\sigma(\bar{\mathbf{u}}_\sigma)\le\limsup_{\sigma\to 0}J_\sigma(\mathbf{u}_\sigma)=J(\tilde{\mathbf{u}})=\inf(\mathrm{P})$$

which implies that $\bar{\mathbf{u}}$ is a solution of (P) as well as the convergence $J_\sigma(\bar{\mathbf{u}}_\sigma)\to J(\bar{\mathbf{u}})$. From this convergence along with the convergence properties of $\mathbf{y}_{\bar{\mathbf{u}}_\sigma}\to\mathbf{y}_{\bar{\mathbf{u}}}$ given in Corollary 4.11, we get $\|\bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\to\|\bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega))}$. Invoking once again (4.1), we obtain the strong convergence of $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ to $\bar{\mathbf{u}}$ stated in (4.34). □

The next theorem is important from a practical point of view because it states that every strict local minimum of problem (P) can be approximated by local minima of problems (P_σ).

THEOREM 4.17. *Let* $\bar{\mathbf{u}}$ *be a strict local minimum of* (P), *then there exists a sequence* $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ *of local minima of problems* (P_σ) *such that* (4.34) *holds.*

Proof. Let $\bar{\mathbf{u}}$ be a strict local minimum of (P), then there exists $\varepsilon>0$ such that $\bar{\mathbf{u}}$ is the unique solution of

$$(\mathrm{P}_\varepsilon)\qquad \min_{\mathbf{u}\in\mathcal{U}_{ad}\cap\bar{B}_\varepsilon(\bar{\mathbf{u}})}J(\mathbf{u}),$$

where $B_\varepsilon(\bar{\mathbf{u}})$ is a ball in $L^2(0,T;\mathbf{L}^2(\Omega))$. Let us extend all the elements of $\mathcal{U}_\sigma$ to $(0,T)\times\Omega$ by taking $\mathbf{u}_\sigma(t,x) = \bar{\mathbf{u}}(t,x)$ for any $(t,x)\in(0,T)\times(\Omega\setminus\Omega_h)$. Let us consider the discrete problems

$$(\mathrm{P}_{\varepsilon,\sigma}) \qquad \min_{\mathbf{u}_\sigma\in\mathcal{U}_{\sigma,ad}\cap\bar{B}_\varepsilon(\bar{\mathbf{u}})} J_\sigma(\mathbf{u}_\sigma).$$

For every σ sufficiently small, the problem $(\mathrm{P}_{\varepsilon,\sigma})$ has at least one solution. Indeed, the only delicate point is to check that $\mathcal{U}_{\sigma,ad}\cap\bar{B}_\varepsilon(\bar{\mathbf{u}})$ is not empty. To this end, we define $\mathbf{u}_\sigma\in\mathcal{U}_{\sigma,ad}$ as in (4.35), with $\tilde{\mathbf{u}}$ replaced by $\bar{\mathbf{u}}$. Then, $\|\bar{\mathbf{u}}-\mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\to 0$, therefore $\mathbf{u}_\sigma\in\mathcal{U}_{\sigma,ad}\cap\bar{B}_\varepsilon(\bar{\mathbf{u}})$ for any σ sufficiently small. Let $\bar{\mathbf{u}}_\sigma$ be a solution of (P_ε). Then we can argue as in the proof of Theorem 4.15 to deduce that any subsequence of $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ converges strongly in $L^2(0,T;\mathbf{L}^2(\Omega))$ to a solution of (P_ε). Since this problem has a unique solution, we have $\|\bar{\mathbf{u}}-\bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega))}\to 0$ for the whole sequence as $\sigma\to 0$. This implies that the constraint $\bar{\mathbf{u}}_\sigma\in\bar{B}_\varepsilon(\bar{\mathbf{u}})$ is not active for σ small, and hence $\bar{\mathbf{u}}_\sigma$ is a local solution of (P_σ) and (4.34) is fulfilled. □

4.4. Error estimates. We still assume that (4.23) holds. In this section $\bar{\mathbf{u}}$ will denote a local solution of problem (P) and for every σ, $\bar{\mathbf{u}}_\sigma$ denotes a local solution of (P_σ) such that $\|\bar{\mathbf{u}}-\bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\to 0$; see Theorems 4.15 and 4.17. Hereinafter, all the elements $\mathbf{u}\in\mathcal{U}_\sigma$ are extended to $(0,T)\times\Omega$ by setting $\mathbf{u}(t,x)=\bar{\mathbf{u}}(t,x)$ for $(t,x)\in(0,T)\times(\Omega\setminus\Omega_h)$. We will also denote by $\bar{\mathbf{y}}$ and $\bar{\boldsymbol{\varphi}}$ the state and adjoint state associated to $\bar{\mathbf{u}}$, and $\bar{\mathbf{y}}_\sigma$ and $\bar{\boldsymbol{\varphi}}_\sigma$ will denote the discrete state and adjoint state associated to $\bar{\mathbf{u}}_\sigma$. The goal of this section is to prove the following theorem.

THEOREM 4.18. *Suppose that* (3.21) *holds. Then, there exists a constant* $C>0$ *independent of* σ *such that*

$$\|\bar{\mathbf{u}}-\bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\le Ch, \tag{4.36}$$

$$\|\bar{\mathbf{y}}-\bar{\mathbf{y}}_\sigma\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))}+\|\bar{\mathbf{y}}-\bar{\mathbf{y}}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\le Ch, \tag{4.37}$$

$$\|\bar{\boldsymbol{\varphi}}-\bar{\boldsymbol{\varphi}}_\sigma\|_{L^\infty(0,T;\mathbf{L}^2(\Omega_h))}+\|\bar{\boldsymbol{\varphi}}-\bar{\boldsymbol{\varphi}}_\sigma\|_{L^2(0,T;\mathbf{H}^1(\Omega_h))}\le Ch. \tag{4.38}$$

The estimates (4.37) and (4.38) are an immediate consequence of (4.36), (4.24) and (4.33). We only have to prove (4.36). To this end, we proceed by contradiction and we assume that it is false. This implies that

$$\limsup_{\sigma\to 0}\frac{1}{h}\|\bar{\mathbf{u}}-\bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}=+\infty,$$

therefore, there exists a sequence of σ such that

$$\lim_{\sigma\to 0}\frac{1}{h}\|\bar{\mathbf{u}}-\bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}=+\infty. \tag{4.39}$$

We will obtain a contradiction for this sequence. We need some lemmas. The first one is concerned with the projection of $\bar{\mathbf{u}}$ on $\mathcal{U}_\sigma$ given by the formulas (4.35) and denoted in the sequel by $\mathbf{u}_\sigma$. Let us recall that according to Theorem 3.4, the regularity $\bar{\mathbf{u}}\in\mathbf{H}^1(\Omega_T)$ holds for any local minimum.

LEMMA 4.19. *There exists a constant* $C>0$ *independent of* σ *such that*

$$\|\bar{\mathbf{u}}-\mathbf{u}_\sigma\|_{\mathbf{H}^1(\Omega_{Th})^*}+h\|\bar{\mathbf{u}}-\mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\le Ch^2\|\bar{\mathbf{u}}\|_{\mathbf{H}^1(\Omega_T)}, \tag{4.40}$$

where $\Omega_{Th} = (0,T) \times \Omega_h$.

Proof. The estimate in the $L^2(0,T;\mathbf{L}^2(\Omega_h))$ norm is well know. Let us check the estimate in the $\mathbf{H}^1(\Omega_{Th})^*$ norm. Let $\mathbf{v} \in \mathbf{H}^1(\Omega_{Th})$ be any element and take $\mathbf{v}_\sigma$ as the projection according to the expression (4.35). From the definition of the projection we have

$$\int_0^T \int_{\Omega_h} \mathbf{v}(\bar{\mathbf{u}} - \mathbf{u}_\sigma)\,\mathrm{d}t\mathrm{d}x = \int_0^T \int_{\Omega_h} (\mathbf{v} - \mathbf{v}_\sigma)(\bar{\mathbf{u}} - \mathbf{u}_\sigma)\,\mathrm{d}t\mathrm{d}x$$
$$\leq \|\mathbf{v} - \mathbf{v}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\|\bar{\mathbf{u}} - \mathbf{u}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \leq Ch^2\|\bar{\mathbf{u}}\|_{\mathbf{H}^1(\Omega_T)}\|\mathbf{v}\|_{\mathbf{H}^1(\Omega_{Th})},$$

which proves the lemma. □

Since $\bar{\mathbf{u}}_\sigma$ is a local minimum of (P_σ) , J_σ is a C^∞ mapping and $\mathbf{u}_\sigma \in \mathcal{U}_{\sigma,ad}$, then $J'_\sigma(\bar{\mathbf{u}}_\sigma)(\mathbf{u}_\sigma - \bar{\mathbf{u}}_\sigma) \geq 0$. This inequality can be rewritten in the form

$$J'(\bar{\mathbf{u}}_\sigma)(\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma) + [J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}}_\sigma)](\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma)$$
$$+[J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\mathbf{u}_\sigma - \bar{\mathbf{u}}) + J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}) \geq 0.$$

On the other hand, since $\bar{\mathbf{u}}_\sigma \in \mathcal{U}_{ad}$, then $J'(\bar{\mathbf{u}})(\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}) \geq 0$. Adding this inequality to the last one, we obtain

$$[J'(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}) \leq [J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}}_\sigma)](\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma)$$
$$+[J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\mathbf{u}_\sigma - \bar{\mathbf{u}}) + J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}). \tag{4.41}$$

This inequality is crucial in the proof. First, we get an estimate from below for the left hand side, then we estimate from above the three terms of the right hand side.

LEMMA 4.20. *Suppose that that* (4.36) *is false and let* $\delta > 0$ *be given by Remark 3.7. Then, there exists* σ_0 *such that*

$$\frac{1}{2}\min\{\delta,\lambda\}\|\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}\|^2_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \leq [J'(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}) \quad \text{if } |\sigma| < |\sigma_0|. \tag{4.42}$$

Proof. In this proof, we follow the steps of [5, Lemma 7.2]. Let us take a sequence $\{\bar{\mathbf{u}}_\sigma\}_\sigma$ satisfying (4.39). By applying the mean value theorem we get for some $\hat{\mathbf{u}}_\sigma = \bar{\mathbf{u}} + \theta_h(\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}})$

$$[J'(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}) = J''(\hat{\mathbf{u}}_\sigma)(\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}})^2. \tag{4.43}$$

Let us take

$$\rho_\sigma = \|\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}\|_{L^2((0,T;\mathbf{L}^2(\Omega_h))} \quad \text{and} \quad \mathbf{v}_\sigma = \frac{1}{\rho_\sigma}(\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}).$$

Taking a subsequence if necessary, we can assume that $\mathbf{v}_\sigma \rightharpoonup \mathbf{v}$ in $L^2((0,T;\mathbf{L}^2(\Omega_h))$. Let us prove that $\mathbf{v}$ belongs to the critical cone $\mathcal{C}_{\bar{\mathbf{u}}}$ defined in (3.14). First of all remark that every $\mathbf{v}_\sigma$ satisfies the sign conditions (3.15)-(3.16), hence $\mathbf{v}$ also does. Let us prove that $v_j(t,x) = 0$ if $d_j(t,x) \neq 0$, $\bar{\mathbf{d}}$ being defined by (3.13). Let us denote by

$$J'_\sigma(\bar{\mathbf{u}}_\sigma) = \bar{\mathbf{d}}_\sigma = \bar{\varphi}_\sigma + \lambda\bar{\mathbf{u}}_\sigma;$$

see (4.28). Note that $\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \to 0$ and (4.33) imply the convergence $\|\bar{\mathbf{d}} - \bar{\mathbf{d}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \to 0$. Now, we have

$$\int_0^T \int_\Omega \bar{\mathbf{d}}\mathbf{v}\,\mathrm{d}x\mathrm{d}t = \lim_{\sigma\to 0}\int_0^T\int_{\Omega_h} \bar{\mathbf{d}}_\sigma \mathbf{v}_\sigma\,\mathrm{d}x\mathrm{d}t$$
$$\lim_{\sigma\to 0}\frac{1}{\rho_\sigma}\left\{\int_0^T\int_{\Omega_h}\bar{\mathbf{d}}_\sigma(\mathbf{u}_\sigma - \bar{\mathbf{u}})\,\mathrm{d}x\mathrm{d}t + \int_0^T\int_{\Omega_h}\bar{\mathbf{d}}_\sigma(\bar{\mathbf{u}}_\sigma - \mathbf{u}_\sigma)\,\mathrm{d}x\mathrm{d}t\right\}.$$

From (4.39), (4.40) and the inequality $J'_\sigma(\bar{\mathbf{u}}_\sigma)(\mathbf{u}_\sigma - \bar{\mathbf{u}}_\sigma) \geq 0$ we conclude that

$$\int_0^T\int_\Omega \bar{\mathbf{d}}\mathbf{v}\,\mathrm{d}x\mathrm{d}t \leq \lim_{\sigma\to 0}\frac{1}{\rho_\sigma}\int_0^T\int_{\Omega_h}\bar{\mathbf{d}}_\sigma(\mathbf{u}_\sigma - \bar{\mathbf{u}})\,\mathrm{d}x\mathrm{d}t$$
$$\leq \lim_{\sigma\to 0}\frac{Ch}{\|\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}\|_{L^2((0,T;\mathbf{L}^2(\Omega_h))}} = 0.$$

Since $\mathbf{v}$ satisfies the sign conditions (3.15)-(3.16), then $d_j(t,x)v_j(t,x) \geq 0$; hence the above inequality implies that (3.17) holds as well, then $\mathbf{v} \in \mathcal{C}_{\bar{\mathbf{u}}}$. Now, from the definition of $\mathbf{v}_\sigma$, (3.7) and (3.28) we get

$$\lim_{\sigma\to 0} J''(\hat{\mathbf{u}}_\sigma)(\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}})^2$$
$$= \lim_{\sigma\to 0}\left\{\int_0^T\int_\Omega(|\mathbf{z}_{\mathbf{v}_\sigma}|^2 - 2(\mathbf{z}_{\mathbf{v}_\sigma}\cdot\nabla)\mathbf{z}_{\mathbf{v}_\sigma}\boldsymbol{\varphi}_{\hat{\mathbf{u}}_\sigma})\mathrm{d}x\mathrm{d}t + \gamma\int_\Omega|\mathbf{z}_{\mathbf{v}_\sigma}(T)|^2\,\mathrm{d}x + \lambda\right\}$$
$$= \int_0^T\int_\Omega(|\mathbf{z}_\mathbf{v}|^2 - 2(\mathbf{z}_\mathbf{v}\cdot\nabla)\mathbf{z}_\mathbf{v}\bar{\boldsymbol{\varphi}})\mathrm{d}x\mathrm{d}t + \gamma\int_\Omega|\mathbf{z}_\mathbf{v}(T)|^2\,\mathrm{d}x + \lambda$$
$$= J''(\bar{\mathbf{u}})\mathbf{v}^2 + \lambda\left(1 - \|\mathbf{v}\|^2_{L^2((0,T;\mathbf{L}^2(\Omega))}\right) \geq \lambda + (\delta - \lambda)\|\mathbf{v}\|^2_{L^2((0,T;\mathbf{L}^2(\Omega))}.$$

Taking into account that $\|v\|_{L^2((0,T;\mathbf{L}^2(\Omega))} \leq 1$, these inequalities lead to

$$\lim_{\sigma\to 0} J''(\hat{\mathbf{u}}_\sigma)\mathbf{v}_\sigma^2 \geq \min\{\delta,\lambda\} > 0,$$

which proves the existence of σ_0, with $|\sigma| > 0$, such that

$$J''(\hat{\mathbf{u}}_\sigma)\mathbf{v}_\sigma^2 \geq \frac{1}{2}\min\{\delta,\lambda\} \quad \forall|\sigma| < |\sigma_0|.$$

From this inequality, the definition of $\mathbf{v}_\sigma$ and (4.43) we deduce (4.42). □

With (4.41) and (4.42) we obtain

$$\frac{1}{2}\min\{\delta,\lambda\}\|\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}\|^2_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \leq [J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}}_\sigma)](\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma)$$
$$+[J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\mathbf{u}_\sigma - \bar{\mathbf{u}}_\sigma) + J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}}). \tag{4.44}$$

Let us estimate the three terms of the right hand sided. From (3.6) and (4.28) along with the fact that $\bar{\mathbf{u}} = \bar{\mathbf{u}}_\sigma$ in $(0,T)\times(\Omega\setminus\Omega_h)$ we have

$$|[J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}}_\sigma)](\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma)| \leq \|\boldsymbol{\varphi}_{\bar{\mathbf{u}}_\sigma} - \bar{\boldsymbol{\varphi}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}.$$

Taking $\mathbf{u} = \mathbf{v} = \bar{\mathbf{u}}_\sigma$ in (4.33), the previous inequality leads to

$$|[J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}}_\sigma)](\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma)| \leq Ch\|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))}. \tag{4.45}$$

For the second term of (4.44) we use again (4.33) with $\mathbf{u} = \bar{\mathbf{u}}$ and $\mathbf{v} = \bar{\mathbf{u}}_\sigma$, as well as (4.40)

$$
\begin{aligned}
&|[J'_\sigma(\bar{\mathbf{u}}_\sigma) - J'(\bar{\mathbf{u}})](\mathbf{u}_\sigma - \bar{\mathbf{u}}_\sigma)| \\
&\le \left\{ \|\bar{\boldsymbol{\varphi}}_\sigma - \bar{\boldsymbol{\varphi}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} + \|\bar{\mathbf{u}}_\sigma - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \right\} \|\mathbf{u}_\sigma - \bar{\mathbf{u}}\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \\
&\le C \left\{ h + \|\bar{\mathbf{u}} - \bar{\mathbf{u}}_\sigma\|_{L^2(0,T;\mathbf{L}^2(\Omega_h))} \right\} h. \qquad (4.46)
\end{aligned}
$$

Last, we estimate the third term using again (4.40)

$$
|J'(\bar{\mathbf{u}})(\mathbf{u}_\sigma - \bar{\mathbf{u}})| \le \|\bar{\boldsymbol{\varphi}} + \lambda \bar{\mathbf{u}}\|_{\mathbf{H}^1(\Omega_{Th})} \|\mathbf{u}_\sigma - \bar{\mathbf{u}}\|_{\mathbf{H}^1(\Omega_{Th})^*} \le Ch^2. \qquad (4.47)
$$

Finally (4.36) follows from (4.44)-(4.47) with the help of Young's inequality, which contradicts (4.39)

REFERENCES

[1] F. Abergel and R. Temam. On some control problems in fluid mechanics. *Theoret. Comput. Fluid Dynamics*, 1:303–325, 1990.

[2] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, New York, Berlin, Heidelberg, 1994.

[3] E. Casas. An optimal control problem governed by the evolution Navier-Stokes equations. In S.S. Sritharan, editor, *Optimal Control of Viscous Flows*, Philadelphia, 1998. Frontiers in Applied Mathematics, SIAM.

[4] E. Casas, M. Mateos, and J.-P. Raymond. Error estimates for the numerical approximation of a distributed control problem for the steady-state navier-stokes equations. *SIAM J. on Control & Optim.*, 46(3):952–982, 2007.

[5] E. Casas and J.-P. Raymond. Error estimates for the numerical approximation of Dirichlet boundary control for semilinear elliptic equations. *SIAM J. on Control & Optim.*, 45(5):1586–1611, 2006.

[6] K. Chrysafinos. Convergence of discontinuous Galerkin approximations of an optimal control problem associated to semilinear parabolic pde's. *ESAIM M^2AN*, 44(1):189–206, 2010.

[7] K. Chrysafinos anf N.J. Walkington. Discontinuous Galerkin approximations of the Stokes and Navier-Stokes equations. *Math. Comput.*, 79(272):2135–2167, 2010.

[8] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.*, 28(1):43–77, 1991.

[9] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $l_\infty(l^2)$ and $l_\infty(l_\infty)$. *SIAM J. Numer. Anal.*, 32(3):706–740, 1995.

[10] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems IV: Nonlinear problems. *SIAM J. Numer. Anal.*, 32(6):1729–1749, 1995.

[11] D. Estep and S. Larsson. The discontinuous Galerkin method for semilinear parabolic equations. *RAIRO Modél. Math. Anal. Numér.*, 27:35–54, 1993.

[12] P.Girault and P .A.Raviart. *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms*. Springer-Verlag, Berlin, Heidelberg, New York, Tokio, 1986.

[13] M.D. Gunzburger. *Perspectives in flow control and optimization*. SIAM, Advances in Design and Control, Philadelphia, 2003.

[14] M.D. Gunzburger. *Flow Control*. Springer-Verlag, New York, 1995.

[15] M.D. Gunzburger and S. Manservisi. The velocity tracking problem for navier-stokes flows with bounded distributed control. *SIAM J. Control Optim.*, 37(6):1913–1945, 1999.

[16] M.D. Gunzburger and S. Manservisi. Analysis and approximation of the velocity tracking problem for Navier-Stokes flows with distributed control. *SIAM J. Numer. Anal.*, 37:1481–1512, 2000.

[17] M. Hinze and K. Kunisch. Second order methods for optimal control of time-dependent fluid flow. *SIAM J. Control Optim.*, 40(3):925–946, 2001.

[18] O.A. Ladyzhenskaya. *The Mathematical Theory of Viscous Incompressible Flow*. Gordon and Breach, New York, second edition edition, 1969. English translation.

[19] J.L. Lions. *Quelques Méthodes de Résolution des Problèmes aux Limites non Linéaires*. Dunod, Paris, 1969.

[20] J.L. Lions and E. Magenes. *Problèmes aux Limites non Homogènes.* Dunod, Paris, 1968.
[21] D. Meidner and B. Vexler. A priori error estimates for the space-time finite element discretization of parabolic optimal control problems. Part I: Problems without control constraints. *SIAM J. Control and Optim.*, 47(3):1150–1177, 2008.
[22] D. Meidner and B. Vexler. A priori error estimates for the space-time finite element discretization of parabolic optimal control problems. Part II: Problems with control constraints. *SIAM J. Control and Optim.*, 47(3):1301–1329, 2008.
[23] P.A. Raviart and J.M. Thomas. *Introduction à L'analyse Numérique des Equations aux Dérivées Partielles.* Masson, Paris, 1983.
[24] A. Rösch. Error estimates for parabolic optimal control problems with control constraints. *Z. Anal. Anwendungen*, 23:353–376, 2004.
[25] S.S. Sritharan. *Optimal control of viscous flow.* SIAM, Philadelphia, 1998.
[26] R. Temam. *Navier-Stokes Equations.* North-Holland, Amsterdam, 1979.
[27] V. Thomée. *Galerkin finite element methods for parabolic problems.* Spinger-Verlag, Berlin, 1997.
[28] F. Tröltzsch. *Optimal Control of Partial Differential Equations*, volume 112 of *Graduate Studies in Mathematics.* American Mathematical Society, Philadelphia, 2010.
[29] F. Tröltzsch and D. Wachsmuth. Second-order suficcient optimality conditions for the optimal control of Navier-Stokes equations. *ESAIM: COCV*, 12:93–119, 2006.