J. Math. Fluid Mech. (2021) 23:56

https://doi.org/10.1007/s00021-021-00577-z

Journal of Mathematical Fluid Mechanics

Optimal Control for the Navier–Stokes Equation with Time Delay in the Convection: Analysis and Finite Element Approximations

Gilbert Peralta and John Sebastian Simon

Communicated by S. Turek

Abstract. A distributed optimal control problem for the 2D incompressible Navier–Stokes equation with delay in the convection term is studied. The delay corresponds to the non-instantaneous effect of the motion of a fluid parcel on the mass transfer, and can be realized as a regularization or stabilization to the Navier–Stokes equation. The existence of optimal controls is established, and the corresponding first-order necessary optimality system is determined. A semi-implicit discontinuous Galerkin scheme with respect to time and conforming finite elements for space is considered. Error analysis for this numerical scheme is discussed and optimal convergence rates are proved. The fully discrete problem is solved by the Barzilai-Borwein gradient method. Numerical examples for the velocity-tracking and vorticity minimization problems based on the Taylor-Hood elements are presented.

Mathematics Subject Classification. 49J20, 49K20, 49M15.

Keywords. Delayed Navier–Stokes equation, Distributed control, Optimality conditions, Finite elements, Discontinuous Galerkin method, Error estimates.

1. Introduction

Optimal flow control problems remain a very active field of research due to their wide variety of applications in physics and engineering. These include combustion, optimal mixing, shape design, kinetic energy regulation and turbulence minimization to name a few. Rigorous mathematical analysis of such problems, as well as their realization to efficient numerical methods, are among the main themes in the past decades. The pioneering work of Abergel and Temam [1] served as an impetus in the study of optimal control problems for time-dependent fluid flows, where first-order necessary optimality conditions were established. Gradient-based algorithms approximating the controls were also suggested. Since then, there are numerous papers extending this work, see for instance [10–12, 20, 21, 31, 32, 38, 53] and the references therein. We also refer to the earlier works of Fursikov [25–27]. Recent developments also include thermodynamic effects, multi-phase flows, phase transitions, and the interaction with either elastic or rigid bodies.

The current paper is dedicated to the analysis and numerical approximations to a distributed optimal control problem for time-dependent incompressible fluid flows governed by the two-dimensional Navier–Stokes equation with delay in the convection. A very short account for control problems of partial differential equations with delay was presented in [40, Section 18.1], and recent work that dealt with numerical aspects is given in [43]. In both cases the delay appears linear in the state. In our work, on the other hand, the delayed term is bilinear in nature, for which the history acts as a convective force for the fluid flow. This leads to different characteristics of the control. For instance, velocity-tracking problems at the terminal time have controls with limited regularity.

Let us now state the precise formulation of the optimal control problem. Given a fixed final time $T > 0$ and an open, bounded and connected domain $\Omega \subset \mathbb{R}^2$ that is either of class C^2 or a convex polygonal with boundary Γ, we consider the following infinite-dimensional optimization problem:

Published online: 05 May 2021

$$
\left\{
\begin{aligned}
&\min_{q\in L^2(0,T;L^2(\Omega)^2)} J(u,q) := \frac{\alpha_{\Omega_T}}{2}\int_0^T\!\!\int_\Omega |u-u_d|^2\,\mathrm{d}x\,\mathrm{d}t + \frac{\alpha_T}{2}\int_\Omega |u(T)-u_T|^2\,\mathrm{d}x\\
&\qquad\qquad + \frac{\alpha_R}{2}\int_0^T\!\!\int_\Omega |\nabla\times u|^2\,\mathrm{d}x\,\mathrm{d}t + \frac{\alpha}{2}\int_0^T\!\!\int_\Omega |q|^2\,\mathrm{d}x\,\mathrm{d}t\\
&\text{subject to the state equation}\\
&\qquad \partial_t u - \nu\Delta u + \operatorname{div}(u^r\otimes u) + \nabla p = f + q \quad \text{in } \Omega_T := (0,T)\times\Omega,\\
&\qquad \operatorname{div} u = 0 \quad \text{in } (-r,T)\times\Omega,\\
&\qquad u = 0 \quad \text{on } \Gamma_T := (0,T)\times\Gamma,\\
&\qquad u(0) = u_0 \quad \text{in } \Omega,\\
&\qquad u = z \quad \text{in } \Omega_r := (-r,0)\times\Omega.
\end{aligned}
\right.
\tag{P}
$$

The unknown state variables $u : (0,T)\times\Omega \to \mathbb{R}^2$ and $p : (0,T)\times\Omega \to \mathbb{R}$ represent the velocity field and the pressure of the fluid. The given functions $f : (0,T)\times\Omega \to \mathbb{R}^2$, $u_0 : \Omega \to \mathbb{R}$ and $z : (-r,0)\times\Omega \to \mathbb{R}^2$ are the external forces, initial velocity and initial history, respectively. A no-slip condition for the velocity on the boundary is imposed. For the state equation in (P), the constant $\nu > 0$ is the fluid viscosity and the fluid density has been normalized to 1 for the sake of simplicity. Also, $\alpha > 0$ and $\alpha_{\Omega_T}, \alpha_T, \alpha_R \geq 0$ are given constants, where at least one of the latter three parameters is positive in order to have a nontrivial solution to (P).

We use the customary notation $u^r(t,x) := u(t-r,x)$ for the delay of velocity with respect to time, where $0 < r < T$ is a fix delay parameter. The convection term $(u^r\cdot\nabla)u$ corresponds to the non-instantaneous transfer of momentum on the fluid bulk. As pointed out in [41], if there is a time delay r in "following the fluid", then the material derivative is given by $\frac{Du}{Dt} = \partial_t u + (u^r\cdot\nabla)u$, where the directional derivative of u is taken with respect to the delayed velocity field u^r. Due to the incompressibility assumption $\operatorname{div} u^r = 0$, the convective term $(u^r\cdot\nabla)u$ coincides with $\operatorname{div}(u^r\otimes u)$. Here, the tensor product $v\otimes w : \Omega \to \mathbb{R}^{2\times 2}$ of two vector valued functions $v, w : \Omega \to \mathbb{R}^2$ has the components $(v\otimes w)_{ij} := v_i w_j$ for $i,j = 1,2$. For works on the Navier–Stokes with delay, we refer the reader to [7,9,28,29,51,52]. The delay in the convective term can be considered as a regularization or stabilization to the Navier–Stokes equation.

In the cost functional J, the first two integrals correspond to a velocity tracking problem, where u_d and u_T are the desired velocity profiles in the space-time domain and space domain at the terminal time, respectively. These intend to minimize the kinetic energy, or a fraction of it, of the difference between the optimal state to the desired target. The third integral aims to minimize the turbulence of the fluid flow, where $\nabla\times u = \partial_{x_2}u_1 - \partial_{x_1}u_2$ is the curl of the fluid velocity $u = (u_1,u_2)$. Finally, the fourth integral is a Tikhonov regularization term leading to coercivity of J, and it also measures the cost of the control. The general rule of thumb here is that the smaller the value of α, the more the controls are going to be *expensive*.

One of the goals of the paper is to establish the well-posedness and regularity of the solutions to the state equation and as well as the associated linearized and adjoint problems. Although the results are analogous to the case without delay, this has to be done *ab initio* in order to have a clear understanding on how the initial history enters in the analysis. In fact, we shall see that the delay impedes further regularity on the optimal control. To be precise, if $\alpha_T > 0$ then even for compatible initial datum and initial history in the state equation, the adjoint state does not enjoy the same compatibility at the terminal time, see Theorem 3.4. Nonetheless, the results here will be useful in the error analysis for the finite-dimensional approximations. The differentiability properties of the so-called control-to-state operator will be established from the implicit function theorem, deviating from those that were presented in [1,53].

The other goal is to analyze a semi-implicit scheme for (P) based on discontinuous-in-time Galerkin and finite element methods. It will be shown that in terms of the space-time L^2-norm, the errors between the continuous and discrete optimal solutions have the order of convergence $\mathcal{O}((\alpha_{\Omega_T}+\alpha+1)h+\alpha_R+\alpha_T)h)$, see Theorem 4.21 and Corollary 4.22. This is with the stability condition $\tau = \mathcal{O}(h^2)$ for the temporal and spatial step sizes τ and h, respectively, a typical condition for explicit or semi-implicit schemes to parabolic problems. Note that in the uncontrolled case of the Navier–Stokes equation without the delay, the condition $\tau = \mathcal{O}(h^\gamma)$ for some $\gamma > 0$ on the time step and mesh size was imposed in [33–37,42] when explicit and semi-implicit schemes are applied in the convection term. However, the methods presented in these papers are not applicable to the current problem due to the limited time-regularity of the controls. If $\alpha_R = \alpha_T = 0$, then we obtain the expected optimal quadratic order of convergence. To establish this convergence rate, we shall utilize Aubin–Nitsche-type duality arguments. In addition, error estimates for the control, state, and adjoint variables in terms of the norms of the function spaces $L^2(0,T;H_0^1(\Omega)^2)$ and $L^\infty(0,T;L^2(\Omega)^2)$ will be proved.

The associated finite-dimensional optimization problem will be solved by the gradient method of Barzilai and Borwein [6]. This particular choice is based on its simplicity, efficiency, and applicability to large-scale optimization problems. As an application, we consider examples on the velocity-tracking and vorticity minimization problems with local controls.

This paper is organized as follows: In Sect. 2, we establish the well-posedness and regularity of solutions of the state, linearized state and adjoint equations. The existence and regularity of the optimal controls will be discussed in Sect. 3. Section 4 deals with the proposed numerical scheme for the optimal control problem. Finally, numerical experiments based on the two commonly utilized finite elements for the Navier–Stokes equation, the mini-finite and Taylor-Hood elements, will be presented in Sect. 5.

2. Analysis of the State, Linearized State and Adjoint Equations

The existence and uniqueness of solutions to the state equation in the optimal control problem (P) can be established through a standard spectral Galerkin method. There are two possible directions that one may pursue. One such approach is to successively consider intervals of length equal to the delay and show well-posedness using the fact that the state equation is an Oseen equation at each subinterval. Alternatively, one can proceed by following the classical strategy for the nonlinear Navier–Stokes equation. For the sake of completeness and clarity, especially the required regularity and compatibility conditions on the initial history, we discuss in detail the latter approach.

2.1. Preliminaries

Let us introduce the function spaces and notations that will be used throughout the paper. The dual space of a Banach space Z will be denoted by Z^* and $\langle z^*, z\rangle_{Z^*,Z}$ represents the duality pairing between $z^* \in Z^*$ and $z \in Z$. The set of all bounded linear operators from a Banach space U into a Banach space Z is denoted by $\mathcal{L}(U,Z)$ and $\mathcal{L}(U) := \mathcal{L}(U,U)$. We follow standard notations for the Lebesgue space $L^p(\Omega)$ and Sobolev space $H^r(\Omega)$ for $1 \le p \le \infty$ and $r \in \mathbb{R}$, and denote the corresponding norms by $\|\cdot\|_{L^p}$ and $\|\cdot\|_{H^r}$, see [2] for more details. The closure in $H^r(\Omega)$ of the set $C_0^\infty(\Omega)$ consisting of all infinitely differentiable functions that vanish on a neighborhood of Γ will be denoted by $H_0^r(\Omega)$. All throughout the paper, we use the abbreviations

$$X := L^2(\Omega)^2, \quad W := H_0^1(\Omega)^2, \quad M := L^2(\Omega)/\mathbb{R}, \quad Y := H^1(\Omega) \cap M.$$

The solenoidal functions with no-slip boundary condition will be denoted by

$$H := \{u \in X : \operatorname{div} u = 0 \text{ in } \Omega,\ u \cdot n = 0 \text{ on } \Gamma\}, \quad V := W \cap H,$$

where n is the unit outward vector normal to Γ. These are Hilbert spaces with respect to the inner products in X and W. The embedding $V \subset H$ is dense, continuous, and compact.

Let $A : D(A) \subset H \to H$ be the Stokes operator defined by $Au = -P\Delta u$ for $u \in D(A)$, where $P : X \to H$ is the Leray projection operator associated with the Helmholtz decomposition $X = H \oplus \nabla L^2(\Omega)$. Since $\Omega \subset \mathbb{R}^2$ is either a convex polygonal domain or of class C^2, then $D(A) = V \cap H^2(\Omega)^2$, see [39] and

[49, Lemma III.2.1]. Equipped with the inner product $(u,v)_{D(A)} = (Au, Av)_H$, $D(A)$ becomes a Hilbert space. Moreover, the norms $\|\cdot\|_{H^2}$ and $\|\cdot\|_{D(A)}$ are equivalent in $D(A)$.

It is well-known that A is a self-adjoint positive operator with dense domain and compact inverse. As a consequence, H has an orthonormal basis $\{\varphi_n\}_{n=1}^\infty$ consisting of eigenfunctions of A with an associated sequence of eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \cdots$ where $\lambda_k \to \infty$ as $k \to \infty$. Each $u \in H$ admits the unique Fourier expansion $u = \sum_{k=1}^\infty (u, \varphi_k)_H \varphi_k$. The domain and the action of the linear operator A can be written as follows:

$$D(A) = \left\{ u \in H : \sum_{k=1}^\infty (1+\lambda_k^2)|(u,\varphi_k)_H|^2 < \infty \right\}, \qquad Au = \sum_{k=1}^\infty \lambda_k (u,\varphi_k)_H \varphi_k.$$

Let V^k be the linear span of $\{\varphi_j\}_{j=1}^k$ and $P_{V^k} : H \to H$ be the orthogonal projection of H onto V^k, that is, $P_{V^k} u := \sum_{j=1}^k (u,\varphi_j)_H \varphi_j$. For each k, it holds that $\|P_{V^k}\|_{\mathcal{L}(H)} \leq 1$ and $\|P_{V^k}u - u\|_H \to u$ as $k \to \infty$ whenever $u \in H$. Given $u \in D(A)$, we have $Au \in H$ and so $\|P_{V^k}u\|_{D(A)} = \|AP_{V^k}u\|_H = \|P_{V^k}Au\|_H \leq \|Au\|_H$ and $\|P_{V^k}u - u\|_{D(A)} = \|P_{V^k}Au - Au\|_H \to 0$ as $k \to \infty$. In particular, $\|P_{V^k}\|_{\mathcal{L}(D(A))} \leq 1$ for every k.

The square root $A^{1/2}$ of A is well-defined and $D(A^{1/2}) = V$. The following spectral representations for the space V and its corresponding norm hold

$$V = \left\{ u \in H : \sum_{k=1}^\infty \lambda_k |(u,\varphi_k)_H|^2 < \infty \right\}, \qquad \|u\|_V = \left(\sum_{k=1}^\infty \lambda_k |(u,\varphi_k)_H|^2 \right)^{1/2}.$$

From the orthonormality of the basis, it follows from these representations that $\|P_{V^k}\|_{\mathcal{L}(V)} \leq 1$ for every k and each $u \in V$ satisfies $\|P_{V^k}u - u\|_V \to 0$ as $k \to \infty$.

We will work in the Bochner spaces $L^p(I, Z)$ for $1 \leq p \leq \infty$ from an interval $I = (a,b)$ into a real Hilbert space Z, and $C(\bar{I}, Z)$ the space of continuous functions from $\bar{I}$ into Z equipped with their usual norms $\|u\|_{C(\bar{I},Z)} = \sup_{t\in\bar{I}} \|u(t)\|_Z$, $\|v\|_{L^\infty(I,Z)} = \operatorname{ess\,sup}_{t\in I} \|v(t)\|_Z$ and

$$\|w\|_{L^p(I,Z)} = \left(\int_I \|w(t)\|_Z^p \, \mathrm{d}t \right)^{1/p} \qquad (1 \leq p < \infty).$$

The space $W^{k,p}(I,Z)$ is the set of all $u \in L^p(I,Z)$ having distributional derivatives $\partial_t^j u \in L^p(I,Z)$ for $0 \leq j \leq k$, while $C^k(\bar{I}, Z)$ is the space of all functions $u : \bar{I} \to X$ such that $\partial_t^j u \in C(\bar{I}, X)$ for every $0 \leq j \leq k$. We shall write $H^k(I,Z)$ for $W^{k,2}(I,Z)$.

Consider the Banach space $W^p(I) := \{u \in L^2(I,V) : \partial_t u \in L^p(I,V^*)\}$ with the norm

$$\|u\|_{W^p(I)} := \|u\|_{L^2(I,V)} + \|\partial_t u\|_{L^p(I,V^*)}.$$

Then the embeddings $W^p(I) \subset C(\bar{I}, V^*)$ for $1 \leq p \leq \infty$ and $W^2(I) \subset C(\bar{I}, H)$ are continuous. Moreover, by the well-known Aubin-Lions-Simon Lemma [46], $W^p(I) \subset L^p(I,H)$ is compact for $1 < p < \infty$. Let $V^2(I) := L^2(I,V) \cap L^\infty(I,H)$ and $V^{2,1}(I) := \{w \in V^2(I) : \partial_t w \in V^2(I)\} = H^1(I,V) \cap W^{1,\infty}(I,H)$ be endowed with the graph norms

$$\|v\|_{V^2(I)} := \|v\|_{L^2(I,V)} + \|v\|_{L^\infty(I,H)}, \qquad \|w\|_{V^{2,1}(I)} := \|w\|_{V^2(I)} + \|\partial_t w\|_{V^2(I)}.$$

It holds that $W^2(I) \subset V^2(I)$ and $V^{2,1}(I) \subset V^2(I)$ continuously. By interpolation theory, we also have the continuous embedding $H^{2,1}(I) := L^2(I, D(A)) \cap H^1(I,H) \subset C(\bar{I}, V)$. Furthermore, $H^{2,1}(I) \subset L^2(I,V)$ is compact.

For the nonlinear convection term, we define the trilinear form $b : W \times W \times W \to \mathbb{R}$ by

$$b(u,v,w) := ((u \cdot \nabla)v, w)_X. \tag{2.1}$$

It follows from the divergence theorem that $b(u,v,w) = -\,b(u,w,v) - ((\operatorname{div} u)v, w)_X$ for each $u,v,w \in W$. In particular, it holds that $b(u,v,w) = -b(u,w,v)$ and $b(u,v,v) = 0$ for every $u \in V$ and $v,w \in W$. In writing the strong form of the adjoint equation, the following equation $b(u,v,w) = ((\nabla v)^\top w, u)_X$ for every $u \in V$ and $v,w \in W$ will be utilized.

For $u \in V$ and $v \in V$, the distributional divergence $\operatorname{div}(u \otimes v) \in V^*$ of $u \otimes v$ is defined by

$$\langle \operatorname{div}(u \otimes v), w\rangle_{V^*,V} := -(u \otimes v, \nabla w) = -b(u, w, v) \qquad \forall w \in V.$$

Let us recall in the following lemma the standard estimates for the trilinear form b, see [18,24,49] for instance. In fact, these estimates follow from the Hölder, Gagliardo–Nirenberg, Agmon and Poincaré inequalities.

Lemma 2.1. *The trilinear form b satisfies the following estimates:*

(a) $|b(u, v, w)| \leq c\|u\|_H^{1/2}\|u\|_V^{1/2}\|v\|_H^{1/2}\|v\|_V^{1/2}\|w\|_V$ *for every* $u, v, w \in V$,
(b) $|b(u, v, w)| \leq c\|u\|_H^{1/2}\|u\|_V^{1/2}\|v\|_V^{1/2}\|Av\|_H^{1/2}\|w\|_H$ *for every* $u \in V$, $v \in D(A)$, $w \in H$,
(c) $|b(u, v, w)| \leq c\|u\|_H\|v\|_V\|Aw\|_H$ *for every* $u \in H$, $v \in V$, $w \in D(A)$,
(d) $|b(u, v, w)| \leq c\|Au\|_H\|v\|_V\|w\|_H$ *for every* $u \in D(A)$, $v \in V$, $w \in H$,

for some constant $c > 0$ *independent of* u, v *and* w.

In what follows, the time, history and *future* domains will be denoted by

$$I := (0, T), \quad I_r := (-r, 0), \quad I^r := (T, T + r), \quad J_r := (-r, T), \quad J^r := (0, T + r) \tag{2.2}$$

where we take without of loss of generality that $r < T$. We shall use $c > 0$ and $\mathfrak{c} > 0$ to denote generic constants and continuous functions, respectively, whose values may differ on each line. To emphasize the dependence on other quantities, we will put a subscript on c or $\mathfrak{c}$.

2.2. Analysis of the State Equation

In this subsection, we study the existence and uniqueness of weak solutions to the state equation and provide the regularity of the solutions under suitable smoothness and compatibility of the initial data and history.

Given $u_0 \in H$, $z \in V^2(I_r)$ and $f \in L^2(I, V^*)$, a function $u \in W^2(I)$ is called a *weak solution* of

$$\begin{cases} \partial_t u - \nu\Delta u + \operatorname{div}(u^r \otimes u) + \nabla p = f \ \text{ in } \Omega_T, \\ \operatorname{div} u = 0 \ \text{ in } \Omega_T, \quad u = 0 \ \text{ in } \Gamma_T, \quad u(0) = u_0 \ \text{ in } \Omega, \quad u = z \ \text{ in } \Omega_r, \end{cases} \tag{2.3}$$

if the following variational equation holds

$$\langle \partial_t u(t), \varphi\rangle_{V^*,V} + \nu(\nabla u(t), \nabla\varphi) - b(u^r(t), \varphi, u(t)) = \langle f(t), \varphi\rangle_{V^*,V} \quad \forall \varphi \in V \tag{2.4}$$

for a.e. $t \in I$, $u(0) = u_0$ in H and $u = z$ in $V^2(I_r)$.

The point-wise value $u(0)$ is well-defined since $W^2(I) \subset C(\bar{I}, H)$. Now, we write an equivalent and convenient formulation of (2.3) as an abstract evolution equation. First, let us extend the definition of the Stokes operator $A : L^2(I, V) \to L^2(I, V^*)$ to the time-dependent case by

$$\langle Av, \varphi\rangle_{L^2(I,V^*),L^2(I,V)} := \int_I \langle Av(t), \varphi(t)\rangle_{V^*,V}\,\mathrm{d}t = \int_I (\nabla v(t), \nabla\varphi(t))_{X^2}\,\mathrm{d}t.$$

Given $z \in V^2(I_r)$, let $B_z : W^2(I) \to L^2(I, V^*)$ be the operator defined by

$$\langle B_z(u), \varphi\rangle = -\int_0^r b(z^r(t), \varphi(t), u(t))\,\mathrm{d}t - \int_r^T b(u^r(t), \varphi(t), u(t))\,\mathrm{d}t.$$

An application of the Hölder inequality and Lemma 2.1(a) yields

$$\|B_z(u)\|_{L^2(I,V^*)} \leq c(\|z\|_{V^2(I_r)} + \|u\|_{V^2(I)})\|u\|_{L^2(I,V)}. \tag{2.5}$$

Then $u \in W^2(I)$ is a weak solution of (2.3) if and only if it satisfies the differential equation

$$\begin{cases} \partial_t u + \nu Au + B_z(u) = f & \text{in } L^2(I, V^*), \\ u(0) = u_0 & \text{in } H. \end{cases} \tag{2.6}$$

Take note here that the history was included in the definition of the nonlinear operator B_z and treated as a coefficient of the evolution equation. With regards to the existence of weak solutions, we have the following theorem. Here, the regularity of the initial history is different from the one provided in [51] for the three-dimensional case.

Theorem 2.2. *Given $u_0 \in H$, $z \in V^2(I_r)$ and $f \in L^2(I, V^*)$, the evolution equation (2.6) has a unique solution $u \in W^2(I)$ and there exists a constant $c > 0$ such that*

$$\|u\|_{V^2(I)} \leq c(\|u_0\|_H + \|f\|_{L^2(I,V^*)}) \tag{2.7}$$

$$\|\partial_t u\|_{L^2(I,V^*)} \leq c((1 + \|z\|_{V^2(I_r)} + \|u\|_{V^2(I)})\|u\|_{L^2(I,V)} + \|f\|_{L^2(I,V^*)}). \tag{2.8}$$

Proof. The proof is based on the spectral Galerkin method, which we provide for the sake of the reader. Take the approximations $u_{0k} := P_{V^k} u_0 \in V^k$ and $z_k := P_{V^k} z \in L^\infty(I_r, V^k)$ for the initial data and initial history. Consider the ansatz $u_k(t,x) = \sum_{j=1}^k \alpha_j(t)\varphi_j(x)$, where $\alpha_j \in H^1(I)$ for $j = 1, \ldots, n$, to the following system of nonlinear delay differential equations

$$\begin{cases} \partial_t u_k + P_{V^k}^*(\nu A u_k + B_{z_k}(u_k)) = P_{V^k}^* f \quad \text{in } L^2(I, V^k), \\ u_k(0) = u_{0k} \quad \text{in } V^k, \qquad u_k = z_k \quad \text{in } L^\infty(I_r, V^k). \end{cases} \tag{2.9}$$

Here, we have extended the projection operators P_{V^k} into the time-dependent case in the obvious way so that $P_{V^k} : L^2(I, V^k) \to L^2(I, V)$. Thus, for the adjoint operator, we have $P_{V^k}^* : L^2(I, V^*) \to L^2(I, V^k)$, where $L^2(I, V^k)^*$ was identified with $L^2(I, V^k)$.

According to the classical Cauchy–Lipschitz theory of delay differential equations, the above system admits a unique solution $u_k \in H^1(0, t_k; V^k)$ for some $0 < t_k \leq T$. The a priori estimates below shows that $t_k = T$. Indeed, taking the inner product of the first equation (2.9) with u_k in H and applying the Young inequality, we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u_k(t)\|_H^2 + \nu\|u_k(t)\|_V^2 \leq \frac{1}{2\nu}\|f(t)\|_{V^*}^2 + \frac{\nu}{2}\|u_k(t)\|_V^2.$$

Integrating over $[0,t]$ for $t \in (0, t_k)$, we deduce that

$$\|u_k(t)\|_H^2 + \nu\int_0^t \|u_k(s)\|_V^2 \,\mathrm{d}s \leq \|u_{0k}\|_H^2 + \frac{1}{\nu}\int_0^t \|f(s)\|_{V^*}^2 \,\mathrm{d}s.$$

Since $\|u_{0k}\|_H \leq \|u_0\|_H$, $\|z_k\|_{V^2(I_r)} \leq \|z\|_{V^2(I_r)}$, and by the virtue of the Gronwall Lemma, there exists a constant $c > 0$ such that

$$\|u_k\|_{V^2(I)} = \|u_k\|_{L^\infty(I,H)} + \|u_k\|_{L^2(I,V)} \leq c(\|u_0\|_H + \|f\|_{L^2(I,V^*)}). \tag{2.10}$$

Using a classical continuation argument, this implies that (2.9) has a solution on the whole interval I. Since $\|P_{V^k}^*\|_{\mathcal{L}(L^2(I,V^*))} \leq 1$ for each k, we obtain from (2.5) that

$$\|\partial_t u_k\|_{L^2(I,V^*)} \leq c((1 + \|z_k\|_{V^2(I_r)} + \|u_k\|_{V^2(I)})\|u_k\|_{L^2(I,V)} + \|f\|_{L^2(I,V^*)}). \tag{2.11}$$

From the a priori estimates (2.10) and (2.11), the sequences $\{u_k\}_{k=1}^\infty$ and $\{u_k^r\}_{k=1}^\infty$ are bounded in $W^2(I)$ and $V^2(J_r)$, respectively. Therefore, one can take a subsequence, denoted by the same indices for simplicity, so that $u_k \rightharpoonup u$ in $L^2(I, V)$, $Au_k \rightharpoonup Au$ in $L^2(I, V^*)$, $\partial_t u_k \rightharpoonup \partial_t u$ in $L^2(I, V^*)$ and $u_k^r \overset{*}{\rightharpoonup} u^r$ in $L^\infty(J_r, H)$ for some $u \in W^2(I) \cap V^2(J_r)$. By the compactness of $W^2(I) \subset L^2(I, H)$, a further subsequence can be extracted in such a way that $u_k \to u$ in $L^2(I, H)$. In particular, $u_k^r \to u^r$ in $L^2(J_r, H)$ and $u = z$ in $L^2(I_r, H)$, since $z_k \to z$ in $L^2(I_r, H)$.

Let us now pass to the limit. Take $\varphi \in L^4(I,V)$. From Lemma 2.1(a) and the Hölder inequality, we obtain, by letting $k \to \infty$, that

$$\begin{aligned}
&|\langle B_{z_k}(u_k) - B_z(u), \varphi\rangle_{L^2(I,V^*),L^2(I,V)}| \le \int_I |(u^r \otimes u - u_k^r \otimes u_k, \nabla\varphi)_X|\,\mathrm{d}t \\
&\le \int_I |((u^r - u_k^r)\otimes u, \nabla\varphi)_X|\,\mathrm{d}t + \int_I |(u_k^r \otimes (u - u_k), \nabla\varphi)_X|\,\mathrm{d}t \\
&\le c\|u_k^r - u^r\|_{L^2(I,H)}^{1/2}\|u_k^r - u\|_{L^2(I,V)}^{1/2}(\|u\|_{L^2(I,V)} + \|u_k^r\|_{L^2(I,V)})\|\varphi\|_{L^4(I,V)} \to 0.
\end{aligned}$$

By the density of $L^4(I,V)$ in $L^2(I,V)$ and the boundedness of $\{B_{z_k}(u_k)\}_{k=1}^\infty$ in $L^2(I,V^*)$, this implies $B_{z_k}(u_k) \rightharpoonup B_z(u)$ in $L^2(I,V^*)$. Therefore,

$$\partial_t u_k + P_{V^k}^*(\nu A u_k + B_{z_k}(u_k) - f) \rightharpoonup \partial_t u + \nu A u + B_z(u) - f \quad \text{in } L^2(I,V^*).$$

By the continuity of the map $\varphi \mapsto \varphi(0)$ from $W^2(I)$ to H, we get $u_k(0) \rightharpoonup u(0)$ in H. Since $u_k(0) = u_{0k} \to u_0$ in H, we obtain that $u(0) = u_0$. Thus u is a solution to (2.6). The a priori estimates (2.7) and (2.8) follows from taking the limit inferior of (2.10) and (2.11), respectively, and utilizing the lower semicontinuity of the norm with respect to weak and weak-star topologies.

For the uniqueness of the solution, it is enough to observe that on the interval $[0,r]$, (2.6) is a linearized Navier–Stokes equation, whose uniqueness of solution follows from standard results. We then apply this to the next interval $[r,2r]$ to conclude that the solution of (2.6) is unique on the interval $[0,2r]$. Continuing this procedure leads to the uniqueness of solution to (2.6) on the whole time interval I. □

Under appropriate conditions on the initial history, we recover the same regularity as in the case of Navier–Stokes equation without delay. This property is reflected on the existence of strong solutions as shown in the theorem below.

Theorem 2.3. *Suppose that $u_0 \in V$, $z \in L^\infty(I_r,V)$ and $f \in L^2(I,X)$. Then the solution of (2.6) satisfies $u \in H^{2,1}(I)$ and there exists a unique $p \in L^2(I,Y)$ such that*

$$\partial_t u + \nu A u + B_z(u) + \nabla p = f \qquad \text{in } L^2(I,X). \tag{2.12}$$

Furthermore, there exists a continuous function $\mathfrak{c} > 0$ such that

$$\|u\|_{H^{2,1}(I)} + \|p\|_{L^2(I,Y)} \le \mathfrak{c}(\|u_0\|_V, \|z\|_{L^\infty(I_r,V)}, \|f\|_{L^2(I,X)}). \tag{2.13}$$

In particular, it holds that $u \in C(\bar{I},V)$.

Proof. Let us adopt the notations in the proof of Theorem 2.2. Taking the inner product of (2.9) with Au_k in H and invoking the Young inequality, one has

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u_k(t)\|_V^2 + \nu\|Au_k(t)\|_H^2 + b(u_k^r(t), u_k(t), Au_k(t)) \le \frac{1}{\nu}\|f(t)\|_X^2 + \frac{\nu}{4}\|Au_k(t)\|_H^2. \tag{2.14}$$

The trilinear term in (2.14) can be estimated according to Lemma 2.1(b) and the Young inequality as follows

$$|b(u_k^r(t), u_k(t), Au_k(t))| \le \frac{1}{\nu}\|u_k^r(t)\|_H^2\|u_k^r(t)\|_V^2\|u_k(t)\|_V^2 + \frac{\nu}{4}\|Au_k(t)\|_H^2.$$

Using this in (2.14), and integrating over $[0,t]$ in the resulting estimate, we obtain

$$\|u_k(t)\|_V^2 + \nu\int_0^t \|Au_k(s)\|_H^2\,\mathrm{d}s \le \|u_{0k}\|_V^2 + \int_0^t c\|u_k^r(s)\|_H^2\|u_k^r(s)\|_V^2\|u_k(s)\|_V^2 + \frac{2}{\nu}\|f(s)\|_X^2\,\mathrm{d}s.$$

From $\|u_{0k}\|_V \le \|u_0\|_V$, $\|z_k\|_{V^2(I_r)} \le c\|z\|_{L^\infty(I_r,V)}$, the Gronwall Lemma and (2.10), there exists a continuous function $\mathfrak{c} > 0$ such that

$$\|u_k\|_{L^\infty(I,V)} + \|u_k\|_{L^2(I,D(A))} \le \mathfrak{c}(\|u_0\|_V, \|z\|_{L^\infty(I_r,V)}, \|f\|_{L^2(I,X)}). \tag{2.15}$$

Using (2.9), Lemma 2.1(b), $\|P^*_{V^k}\|_{\mathcal{L}(L^2(I,H))} \le 1$ and the continuity of the embeddings $D(A) \subset V \subset H$, we have

$$\|\partial_t u_k\|_{L^2(I,H)} \le c((1 + \|u_k^r\|_{L^\infty(I,V)})\|u_k\|_{L^2(I,D(A))} + \|f\|_{L^2(I,X)}). \tag{2.16}$$

Thus, $\{u_k\}_{k=1}^\infty$ is bounded in $H^{2,1}(I)$, hence after extraction of an appropriate subsequence, the weak limit of this sequence in $H^{2,1}(I)$ is the solution of (2.6). Passing to the limit inferior as $k \to \infty$ in the a priori estimates (2.15) and (2.16), we deduce (2.13) but without the pressure term. On the other hand, the existence and uniqueness of $p \in L^2(I, Y)$ satisfying (2.12) follows from de Rham's Theorem and it holds that

$$\|p\|_{L^2(I,Y)} \le c(\|\partial_t u\|_{L^2(I,H)} + (1 + \|u^r\|_{L^\infty(I,V)})\|u\|_{L^2(I,D(A))} + \|f\|_{L^2(I,X)}).$$

Therefore, we have (2.13) and this completes the proof of the theorem. □

Theorem 2.4. *If $u_0 \in D(A)$ and $z \in V^{2,1}(I_r) \cap L^2(I_r, D(A))$ satisfy the compatibility condition $z(0) = u_0$ and the source term $f \in H^1(I, V^*)$ satisfies $f(0) \in X$, then the solution of (2.6) satisfies $u \in V^{2,1}(J_r) \cap L^2(J_r, D(A))$. Moreover, there is a continuous function $\mathfrak{c} > 0$ such that*

$$\|u\|_{V^{2,1}(J_r)\cap L^2(J_r,D(A))} \le \mathfrak{c}(\|u_0\|_{D(A)}, \|z\|_{V^{2,1}(I_r)\cap L^2(I_r,D(A))}, \|f\|_{H^1(I,V^*)}, \|f(0)\|_X).$$

In particular, we have $u \in C(\bar{J}_r, V)$.

Proof. We follow the notations in the proofs of the previous theorems. First, let us note that the compatibility condition $z(0) = u_0$ is carried out in the finite dimensional approximation, that is, $z_k(0) = P_{V^k} z(0) = P_{V^k} u_0 = u_{0k}$ for every k. This compatibility, together with the regularity $z_k \in W^{1,\infty}(I_r, V^k)$ and $f \in H^1(I, V^*)$, implies that $u_k \in H^2(I, V^k) \cap H^1(J_r, V^k)$ according to classical regularity results for delay differential equations. Therefore, we may differentiate the system (2.9). Doing so, we see that $y_k := \partial_t u_k$ satisfies the delay differential equation

$$\begin{cases} \partial_t y_k + P^*_{V^k}(\nu A y_k + B_{z_k}(y_k) + B_{y_k}(u_k)) = P^*_{V^k}\partial_t f & \text{in } L^2(I, V^k), \\ y_k(0) = \partial_t u_k(0) \ \text{ in } V^k, \qquad y_k = \partial_t z_k \ \text{ in } L^\infty(I_r, V^k). \end{cases}$$

Taking the inner product of the differential equation with y_k in H gives us

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|y_k(t)\|_H^2 + \nu\|y_k(t)\|_V^2 - b(y_k^r(t), y_k(t), u_k(t)) \le \frac{1}{\nu}\|\partial_t f(t)\|_{V^*}^2 + \frac{\nu}{4}\|y_k(t)\|_V^2.$$

From Lemma 2.1(a), we have

$$|b(y_k^r(t), y_k(t), u_k(t))| \le \frac{c}{\nu}\|y_k^r(t)\|_V^2\|u_k(t)\|_V^2 + \frac{\nu}{4}\|y_k(t)\|_V^2.$$

Substituting this estimate to the previous one, integrating over $[0, t]$ and then applying the Gronwall-type Lemma 7.1 with $\phi(t) = \frac{1}{2}\|y_k(t)\|_H^2$, $\varphi(t) = \frac{\nu}{2}\|y_k(t)\|_V^2$, $\psi(t) = \frac{1}{\nu}\|\partial_t f(t)\|_{V^*}^2$, $\alpha(t) = \beta(t) = 0$, $\gamma(t) = \frac{c}{\nu}\|u_k(t)\|_V^2$ and $a = \frac{1}{2}\|y_k(0)\|_H^2$, there is a continuous function $\mathfrak{c} > 0$ such that

$$\|y_k\|_{V^2(I)} \le \mathfrak{c}(\|y_k(0)\|_H, \|u_k\|_{L^\infty(I,V)}, \|\partial_t z_k\|_{V^2(I_r)}, \|\partial_t f\|_{L^2(I,V^*)}). \tag{2.17}$$

Note that the sequences $\{u_k\}_{k=1}^\infty$ and $\{z_k\}_{k=1}^\infty$ are bounded in $L^\infty(I, V)$ and $V^{2,1}(I_r)$, respectively, and $u_k \overset{*}{\rightharpoonup} u$ in $L^\infty(I, V)$ and $z_k \overset{*}{\rightharpoonup} z$ in $V^{2,1}(I_r)$. Thus, it remains to estimate $\|y_k(0)\|_H$ to establish the boundedness of $\{y_k\}_{k=1}^\infty$ in $V^2(I)$. Indeed, setting $t = 0$ in (2.9) and then taking the norm of the resulting equation in X, we obtain that

$$\begin{aligned} \|y_k(0)\|_H &\le c(\|Au_{0k}\|_H + \|(u_k^r(0)\cdot\nabla)u_{0k}\|_X + \|f(0)\|_X) \\ &\le c(\|Au_0\|_H + \|z\|_{H^1(I_r,V)}\|Au_0\|_H + \|f(0)\|_X). \end{aligned}$$

As a result, $\{u_k\}_{k=1}^\infty$ is bounded in $V^{2,1}(J_r)$, and for a subsequence $u_k \rightharpoonup u$ in $V^{2,1}(J_r)$. The a priori estimate in the statement of the theorem follows by taking the sum of (2.13) with the inequality obtained by passing to the limit inferior in (2.17) and $\|u_k\|_{V^{2,1}(I_r)\cap L^2(I_r,D(A))} \le \|z\|_{V^{2,1}(I_r)\cap L^2(I_r,D(A))}$. From the compatibility condition, we have $u \in H^{2,1}(J_r) \subset C(\bar{J}_r, V)$. □

2.3. Analysis of the Linearized State Equation

In this subsection, we study the linearization of the state equation (2.6) at a given element $u \in W^2(I)$. Suppose that $z \in V^2(I_r)$. From the quadratic nature of B_z, one can verify immediately that $B_z \in C^\infty(W^2(I), L^2(I,V^*))$. Moreover, the action of the Fréchet derivative of B_z at $u \in W^2(I)$ in the direction $h \in W^2(I)$ is given by

$$\begin{aligned}&\langle B_z'(u)h, \varphi\rangle_{L^2(I,V^*),L^2(I,V)}\\&\quad= -\int_0^r b(z^r(t),\varphi(t),h(t))\,\mathrm{d}t - \int_r^T b(u^r(t),\varphi(t),h(t))\,\mathrm{d}t - \int_r^T b(h^r(t),\varphi(t),u(t))\,\mathrm{d}t.\end{aligned}$$

Likewise, the action of the second Fréchet derivative of B_z at u in the directions $h_1, h_2 \in W^2(I)$ is

$$\langle B_z''(u)[h_1,h_2], \varphi\rangle_{L^2(I,V^*),L^2(I,V)} = -\int_r^T b(h_1^r(t),\varphi(t),h_2(t))\,\mathrm{d}t - \int_r^T b(h_2^r(t),\varphi(t),h_1(t))\,\mathrm{d}t.$$

Take note that $B_z''(u)$ is independent on u, hence the derivatives of B_z beyond order 3 vanish.

Theorem 2.5. *Given $z \in V^2(I_r)$, $u \in W^2(I)$, $v_0 \in H$ and $f \in L^2(I,V^*)$, the linearized state equation*

$$\begin{cases} \partial_t v + \nu A v + B_z'(u)v = f & \text{in } L^2(I,V^*),\\ v(0) = v_0 & \text{in } H,\end{cases} \tag{2.18}$$

has a unique solution $v \in W^2(I)$. Moreover, there exists a continuous function $\mathfrak{c} > 0$ such that

$$\|v\|_{W^2(I)} \le \mathfrak{c}(\|u\|_{W^2(I)}, \|z\|_{V^2(I_r)})(\|v_0\|_H + \|f\|_{L^2(I,V^*)}). \tag{2.19}$$

Proof. The proof is similar to the one provided in Theorem 2.2. For this reason, the details are omitted to avoid repetition. □

Remark 2.6. The solution of the linearized state equation (2.18) can be regarded as the *weak solution* of

$$\begin{cases} \partial_t v - \nu\Delta v + \operatorname{div}(u^r \otimes v) + \operatorname{div}(v^r \otimes u) + \nabla\varpi = f \ \text{ in } \Omega_T,\\ \operatorname{div} v = 0 \ \text{ in } \Omega_T, \quad v = 0 \ \text{ in } \Gamma_T, \quad v(0) = v_0 \ \text{ in } \Omega, \quad v = 0 \ \text{ in } \Omega_r,\end{cases} \tag{2.20}$$

where $u = z$ in Ω_r and ϖ is the corresponding linearized pressure. Sufficient conditions for the weak solution v to be in $H^{2,1}(I)$ and for the existence of the pressure $\varpi \in L^2(I,Y)$ are provided in the following theorem.

Theorem 2.7. *If $z \in L^\infty(I_r,V)$, $u \in H^{2,1}(I)$, $v_0 \in V$ and $f \in L^2(I,X)$, then the solution of (2.18) satisfies $v \in H^{2,1}(I)$ and there exists a unique $\varpi \in L^2(I,Y)$ such that*

$$\partial_t v + \nu A v + B_z'(u)v + \nabla\varpi = f \quad \text{in } L^2(I,X). \tag{2.21}$$

Moreover, there is a continuous function $\mathfrak{c} > 0$ such that

$$\|v\|_{H^{2,1}(I)} + \|\varpi\|_{L^2(I,Y)} \le \mathfrak{c}(\|u\|_{H^{2,1}(I)}, \|z\|_{L^\infty(I_r,V)})(\|v_0\|_V + \|f\|_{L^2(I,X)}). \tag{2.22}$$

Proof. The proof is similar to the case of the state equations, see Theorem 2.3, hence we only derive the necessary a priori estimates. Moreover, we drop the indices k in the associated approximating spectral Galerkin system. Using the test function Av and the antisymmetry of b, we get

$$\begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|v(t)\|_V^2 + \nu\|Av(t)\|_H^2 + b(u^r(t),v(t),Av(t)) + b(v^r(t),v(t),Au(t))\\&\qquad\le \frac{1}{2\nu}\|f(t)\|_X^2 + \frac{\nu}{2}\|Av(t)\|_H^2.\end{aligned} \tag{2.23}$$

For the trilinear terms on the left hand side, one can estimate them from above with the help of Lemma 2.1(b) and the Cauchy–Schwarz inequality as follows

$$|b(u^r(t), v(t), Av(t))| \le c\|u^r(t)\|_H^2\|u^r(t)\|_V^2\|v(t)\|_V^2 + \frac{\nu}{4}\|Av(t)\|_H^2$$
$$|b(v^r(t), u(t), Av(t))| \le c(\|u(t)\|_V^2\|v^r(t)\|_V^2 + \|Au(t)\|_H^2\|v^r(t)\|_V^2) + \frac{\nu}{8}\|Av(t)\|_H^2.$$

Using these estimates in the energy inequality (2.23) and then applying the Gronwall-type Lemma 7.1 with $\phi(t) = \frac{1}{2}\|v(t)\|_V^2$, $\varphi(t) = \frac{\nu}{8}\|Av(t)\|_H^2$, $\psi(t) = \frac{1}{2\nu}\|f(t)\|_X^2$, $\alpha(t) = c\|u^r(t)\|_H^2\|u^r(t)\|_V^2$, $\beta(t) = c(\|u(t)\|_V^2 + \|Au(t)\|_H^2)$, $\gamma(t) = 0$ and $a = \frac{1}{2}\|v(0)\|_V^2$, we obtain

$$\|v\|_{L^\infty(I,V)} + \|v\|_{L^2(I,D(A))} \le \mathfrak{c}(\|u\|_{H^{2,1}(I)}, \|z\|_{L^\infty(I_r,V)})(\|v_0\|_V + \|f\|_{L^2(I,X)})$$

for some continuous function $\mathfrak{c} > 0$. Hence, the time derivative of v can be estimated by

$$\begin{aligned}\|\partial_t v\|_{L^2(I,H)} &\le c(\|Av\|_{L^2(I,H)} + \|B_z'(u)v\|_{L^2(I,H)} + \|f\|_{L^2(I,X)})\\ &\le c((1 + \|z\|_{L^\infty(I_r,V)} + \|u\|_{L^\infty(I,V)})\|Av\|_{L^2(I,H)} + \|f\|_{L^2(I,X)}).\end{aligned}$$

In the second inequality, we used Lemma 2.1(b). The last two inequalities imply that $v \in H^{2,1}(I)$ and (2.3), but without the term involving ϖ. However, the existence of a unique pressure $\varpi \in L^2(I,Y)$ satisfying (2.21) and (2.22) can be established as in the proof of Theorem 2.3. □

Define the map $L_z : W^2(I) \to \mathcal{L}(W^2(I), L^2(I,V^*))$ according to

$$L_z(u)v = \partial_t v + \nu Av + B_z'(u)v \qquad \forall u, v \in W^2(I).$$

Also, let $N_z : W^2(I) \to \mathcal{L}(W^2(I), L^2(I,V^*) \times H)$ be given by

$$N_z(u)v = (L_z(u)v, v(0)).$$

It follows from Theorem 2.5 that the linear operator $N_z(u) \in \mathcal{L}(W^2(I), L^2(I,V^*) \times H)$ is an isomorphism for each $u \in W^2(I)$. In particular, $L_z(u) \in \mathcal{L}(W_0^2(I), L^2(I,V^*))$ is an isomorphism, where $W_0^2(I) := \{v \in W^2(I) : v(0) = 0\}$, and from (2.19) we obtain that

$$\|L_z(u)^{-1}\|_{\mathcal{L}(L^2(I,V^*),W_0^2(I))} \le \mathfrak{c}(\|u\|_{W^2(I)}, \|z\|_{V^2(I_r)}). \tag{2.24}$$

If $z \in L^\infty(I_r, V)$, then $L_z : H^{2,1}(I) \to \mathcal{L}(H_0^{2,1}(I), L^2(I,X))$ according to Theorem 2.7, and

$$\|L_z(u)^{-1}\|_{\mathcal{L}(L^2(I,X),H_0^{2,1}(I))} \le \mathfrak{c}(\|u\|_{H^{2,1}(I)}, \|z\|_{L^\infty(I_r,V)}) \tag{2.25}$$

where $H_0^{2,1}(I) := \{v \in H^{2,1}(I) : v(0) = 0\}$.

For the rest of the paper, $Q := L^2(I,X)$ will denote the space of controls. Given fixed initial data $u_0 \in H$, history $z \in V^2(I_r)$ and source term $f \in L^2(I,V^*)$, let us define the nonlinear operator $F : W^2(I) \times Q \to L^2(I,V^*) \times H$ by

$$F(u,q) = (\partial_t u + \nu Au + B_z(u) - f - q, u(0) - u_0).$$

If $q \in Q \subset L^2(I,V^*)$, then there exists a unique $u \in W^2(I)$ such that $F(u,q) = 0$ by Theorem 2.2. Conversely, if $F(u,q) = 0$ then u is the solution of (2.6) with f replaced by $f + q$. In this way, we define the so-called *control-to-state operator* $S : Q \to W^2(I)$ by $S(q) = u$ if and only if $F(u,q) = 0$. Note that F, and hence S, depends on the triple (u_0, z, f), however, we shall not explicitly write this dependence for simplicity of notation.

Theorem 2.8. *Let $u_0 \in H$, $z \in V^2(I_r)$ and $f \in L^2(I,V^*)$. Then $S \in C^\infty(Q, W^2(I))$. The action of the first and second Fréchet derivatives of S at q in the directions $g \in Q$ and $(g_1,g_2) \in Q \times Q$ are given by*

$$S'(q)g = L_z(S(q))^{-1}g$$
$$S''(q)[g_1,g_2] = -L_z(S(q))^{-1}B_z''(u)[S'(q)g_1, S'(q)g_2].$$

Proof. One can easily see that $F \in C^\infty(W^2(I) \times Q, L^2(I,V^*) \times H)$. Let $q^\star \in Q$ and $u^\star := S(q^\star) \in W^2(I)$, so that $F(u^\star, q^\star) = 0$. We have $\frac{\partial}{\partial u} F(u^\star, q^\star) = N_z(u^\star) \in \mathcal{L}(W^2(I), L^2(I,V^*) \times H)$, which is an isomorphism according to the above discussion. Therefore, by the Implicit Function Theorem [54, Section 4.7], there exist open neighborhoods $O_{q^\star} \subset Q$ and $O_{u^\star} \subset W^2(I)$ of $q^\star$ and $u^\star$, respectively, and a map $\tilde{S} \in C^\infty(O_{q^\star}, O_{u^\star})$ such that $F(\tilde{S}(q), q) = 0$ for every $q \in O_{q^\star}$. This implies that $\tilde{S} = S$ in $O_{q^\star}$ by the definition of S. Since $q^\star \in Q$ is arbitrary, one obtains that $S \in C^\infty(Q, W^2(I))$.

Applying the chain rule to $F(S(q), q) = 0$, we have that $S'(q)g = -L_z(S(q))^{-1} \frac{\partial}{\partial q} F(S(q), q)g = L_z(S(q))^{-1} g$. This implies that $L_z(S(q))S'(q)g = g$. Setting $g = g_1$, taking the derivative in the direction of g_2, and then invoking the chain rule once more to the resulting equation, one has

$$L_z'(S(q))[S'(q)g_1, S'(q)g_2] + L_z(S(q))S''(q)[g_1, g_2] = 0.$$

The result for the second derivative now follows from $L_z'(u)[v_1, v_2] = -B_z''(u)[v_1, v_2]$ for $u \in W^2(I)$ and $v_1, v_2 \in W_0^2(I)$. Here, we note that $L_z' : W^2(I) \to \mathcal{L}(W^2(I), \mathcal{L}(W_0^2(I), L^2(I,V^*)))$, where the latter space is isometrically isomorphic to $\mathcal{L}(W^2(I) \times W_0^2(I), L^2(I,V^*))$. □

Remark 2.9. If $u_0 \in V$, $z \in L^\infty(I_r, V)$ and $f \in L^2(I,X)$ then $S \in C^\infty(Q, H^{2,1}(I))$. This is a consequence of the fact that $N_z(u) \in \mathcal{L}(H^{2,1}(I), Q \times V)$ is an isomorphism for every $u \in H^{2,1}(I)$.

In terms of the strong formulation, $S'(q)g = v$ if and only v is the *weak solution* of (2.20) with $u = S(q)$, $f = g$ and $v_0 = 0$. Also, $S''(q)[g_1, g_2] = y$ if and only if y is the *weak solution* of

$$\begin{cases} \partial_t y - \nu \Delta y + \operatorname{div}(u^r \otimes y) + \operatorname{div}(y^r \otimes u) + \nabla \varrho = -\operatorname{div}(v_1^r \otimes v_2) - \operatorname{div}(v_2^r \otimes v) & \text{in } \Omega_T, \\ \operatorname{div} y = 0 \ \text{ in } \Omega_T, \quad y = 0 \ \text{ in } \Gamma_T, \quad y(0) = 0 \ \text{ in } \Omega, \quad y = 0 \ \text{ in } \Omega_r, \end{cases} \tag{2.26}$$

where $u = S(q)$, $v_1 = S'(q)g_1$, $v_2 = S'(q)g_2$ and $u = z$ in Ω_r.

Theorem 2.10. *If $u_0 \in H$, $z \in V^2(I_r)$ and $f \in L^2(I,V^*)$, then the map $S : Q \to W^2(I)$ is weak-weak continuous, that is, $q_k \rightharpoonup q$ in Q implies $S(q_k) \rightharpoonup S(q)$ in $W^2(I)$. Moreover, if $u_0 \in V$, $z \in L^\infty(I_r, V)$ and $f \in L^2(I,X)$, then $S : Q \to H^{2,1}(I)$ is weak-weak continuous.*

Proof. Since Q and $W^2(I)$ are both reflexive and separable, any closed ball in these spaces is metrizable. Hence, with respect to the weak topologies, continuity is equivalent to weak sequential continuity [22, page 426]. Suppose that $q_k \rightharpoonup q$ in Q. Then $\{q_k\}_{k=1}^\infty$ is bounded in Q and $\{S(q_k)\}_{k=1}^\infty$ is bounded in $W^2(I)$ by Theorem 2.2. Therefore, up to a subsequence, $S(q_k) \rightharpoonup u$ in $W^2(I)$ for some $u \in W^2(I)$. Following the passage of limit in the proof of Theorem 2.2, it can be deduced that $S(q) = u$. Since u is uniquely determined, this implies that the whole sequence $\{S(q_k)\}_{k=1}^\infty$ must converge weakly to u in $W^2(I)$. Indeed, this follows from the fact that every subsequence of $\{S(q_k)\}_{k=1}^\infty$ has a subsequence that converges weakly to $S(q)$. With the help of Theorem 2.3, the proof of the second statement can be handled in a similar manner. □

2.4. Analysis of the Adjoint Equation

In this subsection, we study the adjoint problem corresponding to (2.18). According to the discussions in the previous subsection, $S'(q) = L_z(S(q))^{-1} \in \mathcal{L}(L^2(I,V^*), W_0^2(I))$. In particular, $S'(q)^* = L_z(S(q))^{-*} \in \mathcal{L}(W_0^2(I)^*, L^2(I,V))$.

The action of the adjoint $B_z'(u)^* : L^2(I,V) \to C_0^\infty(I,V)^*$ of the linear operator $B_z'(u) : C_0^\infty(I,V) \to L^2(I,V^*)$, where $u \in W^2(I)$ and $z \in V^2(I_r)$, is given by

$$\begin{aligned} \langle B_z'(u)^* w, \varphi \rangle_{C_0^\infty(I,V)^*, C_0^\infty(I,V)} = &- \int_0^r b(z^r(t), w(t), \varphi(t)) \,\mathrm{d}t - \int_r^T b(u^r(t), w(t), \varphi(t)) \,\mathrm{d}t \\ &- \int_0^{T-r} b(\varphi(t), w^{-r}(t), u^{-r}(t)) \,\mathrm{d}t \end{aligned}$$

for $w \in L^2(I,V)$ and $\varphi \in C_0^\infty(I,V)$. Here, we recall that $u^{-r}(t) := u(t+r)$ and $w^{-r}(t) := w(t+r)$. If $u \in W^2(I)$ and $z \in V^2(I_r)$, then from Lemma 2.1(a) and the Hölder inequality

$$|\langle B_z'(u)^* w, \varphi\rangle_{C_0^\infty(I,V)^*, C_0^\infty(I,V)}| \le c(\|z\|_{V^2(I_r)} + \|u\|_{W^2(I)})\|w\|_{L^2(I,V)}\|\varphi\|_{L^4(I,V)}.$$

By density of $C_0^\infty(I,V)$ in $L^4(I,V)$, it follows that $B_z'(u)^* w \in L^{4/3}(I,V^*)$.

To treat the tracking part at the final time, let us define the linear map $e_T : H \to W_0^2(I)^*$ by

$$\langle e_T v, w\rangle_{W_0^2(I)^*, W_0^2(I)} := (v, w(T))_H, \qquad v \in H,\ w \in W_0^2(I).$$

This operator is bounded since $\|e_T v\|_{W_0^2(I)^*} \le c\|v\|_H$ for every $v \in H$, where $c > 0$ is the constant corresponding to the continuous embedding $W_0^2(I) \subset C(\bar{I}, H)$.

Theorem 2.11. *Assume that $z \in V^2(I_r)$, $u \in W^2(I)$, $g_d \in L^2(I,V^*)$ and $w_T \in H$. Then the function $w := L_z(u)^{-*}(g_d + e_T w_T) \in W^{4/3}(I)$ is precisely the unique solution of*

$$\begin{cases} -\partial_t w + \nu A w + B_z'(u)^* w = g_d & \text{in } L^{4/3}(I,V^*), \\ w(T) = w_T & \text{in } H, \end{cases} \tag{2.27}$$

and there exists a continuous function $\mathfrak{c} > 0$ such that

$$\|w\|_{W^{4/3}(I)} \le \mathfrak{c}(\|u\|_{W^2(I)}, \|z\|_{V^2(I_r)})(\|w_T\|_H + \|g_d\|_{L^2(I,V^*)}). \tag{2.28}$$

Proof. The continuity of the embedding $L^2(I,V^*) \subset W_0^2(I)^*$ and the boundedness of e_T imply that $g_d + e_T w_T \in W_0^2(I)^*$ and there holds $\|g_d + e_T w_T\|_{W_0^2(I)^*} \le c(\|g_d\|_{L^2(I,V^*)} + \|w_T\|_H)$. Observe that the equation $w = L_z(u)^{-*}(g_d + e_T w_T)$ is equivalent to the variational problem

$$\langle \partial_t v + \nu A v + B_z'(u)v, w\rangle_{L^2(I,V^*),L^2(I,V)} = \langle g_d, v\rangle_{L^2(I,V^*),L^2(I,V)} + (v(T), w_T)_H \tag{2.29}$$

for every $v \in W_0^2(I)$. Using (2.24) and $\|L_z(u)^{-*}\|_{\mathcal{L}(W_0^2(I)^*, L^2(I,V))} = \|L_z(u)^{-1}\|_{\mathcal{L}(L^2(I,V^*), W_0^2(I))}$, one obtains that

$$\|w\|_{L^2(I,V)} \le \mathfrak{c}(\|u\|_{W^2(I)}, \|z\|_{V^2(I_r)})(\|w_T\|_H + \|g_d\|_{L^2(I,V^*)}). \tag{2.30}$$

Taking $v \in C_0^\infty(I,V)$ in (2.29), we see that $\partial_t w = \nu A w + B_z'(u)^* w - g_d$ in $C_0^\infty(I,V)^*$. Since $B_z'(u)^* w \in L^{4/3}(I,V^*)$ and $\nu A w - g_d \in L^2(I,V^*) \subset L^{4/3}(I,V^*)$, it follows that $\partial_t w \in L^{4/3}(I,V^*)$ and it satisfies the estimate

$$\|\partial_t w\|_{L^{4/3}(I,V^*)} \le c((1 + \|z\|_{V^2(I_r)} + \|u\|_{W^2(I)})\|w\|_{L^2(I,V)} + \|g_d\|_{L^2(I,V^*)}). \tag{2.31}$$

Therefore, $w \in W^{4/3}(I)$ and (2.28) is verified by the previous estimates (2.30) and (2.31).

To demonstrate the terminal condition $w(T) = w_T$, we shall proceed by a density argument. Since $C^1(\bar{I},V)$ is dense in $W^{4/3}(I,V)$ [45, Lemma 7.2], there is a sequence $\{w_k\}_{k=1}^\infty$ in $C^1(\bar{I},V)$ such that $w_k \to w$ in $W^{4/3}(I,V)$. Given $\varphi \in V$ and $\chi \in C^1(\bar{I})$ such that $\chi(0) = 0$ and $\chi(T) = 1$, we have $\chi\varphi \in W_0^2(I) \cap L^4(I,V)$, and by invoking the continuity of the map $\psi \mapsto \psi(T)$ from $W^{4/3}(I)$ into V^*, we deduce by partial integration that

$$\begin{aligned} \langle \partial_t(\chi\varphi), w\rangle_{L^2(I,V^*),L^2(I,V)} &= \lim_{k\to\infty}((\varphi, w_k(T))_H - (\chi\varphi, \partial_t w_k)_{L^2(I,X)}) \\ &= \langle w(T), \varphi\rangle_{V^*,V} - \langle \partial_t w, \chi\varphi\rangle_{L^4(I,V^*),L^4(I,V)}. \end{aligned}$$

Since $\varphi \in V$ is arbitrary, it follows from the first equation in (2.27) and (2.29) with $v = \chi\varphi$ that $w(T) = w_T$.

Conversely, if w satisfies (2.27), then (2.29) holds for every $v \in W_0^2(I) \cap L^4(I,V)$, and hence for every $v \in W_0^2(I)$ by density of $W_0^2(I) \cap L^4(I,V)$ in $W_0^2(I)$ as well as the continuity of $L_z(u) : W_0^2(I) \to L^2(I,V^*)$. Therefore, $w = L_z(u)^{-*}(g_d + e_T w_T)$ and the proof of the theorem is now complete. □

Remark 2.12. The function $w = L_z(u)^{-*}(g_d + e_T w_T)$ can be viewed as the unique *weak solution* of the backward-in-time linear system with homogeneous *future* data

$$\begin{cases} -\partial_t w - \nu\Delta w - (u^r \cdot \nabla)w - (\nabla w^{-r})^\top u^{-r} + \nabla\pi = g_d \text{ in } \Omega_T, \\ \operatorname{div} w = 0 \text{ in } \Omega_T, \quad w = 0 \text{ in } \Gamma_T, \quad w(T) = w_T \text{ in } \Omega, \quad w = 0 \text{ in } \Omega_{T+r}, \end{cases} \tag{2.32}$$

where $\Omega_{T+r} := (T, T+r) \times \Omega$, $u = z$ in Ω_r, and π can be regarded as the associated *adjoint pressure*. Since the convection term in the state equation was written in divergence form, the above dual problem is not the usual form compared to the one in the literature for the Navier–Stokes equation without delay, specifically, the term involving $-(\nabla w^{-r})^\top u^{-r}$. However, in view of the weak formulations, these representations of the adjoint equation are equivalent for $r = 0$.

We now show that the weak solution of the adjoint problem (2.27) enjoys additional regularity, provided that of course the initial data and initial history also satisfy appropriate regularity and compatibility conditions.

Corollary 2.13. *Suppose that $w_T \in H$, $z \in L^\infty(I_r, V)$, $u \in H^{2,1}(I)$ and $g_d \in L^2(I, V^*)$. Then the solution of (2.27) satisfies $w \in W^2(I)$ and for some continuous function $\mathfrak{c} > 0$ it holds that*

$$\|w\|_{W^2(I)} \le \mathfrak{c}(\|u\|_{H^{2,1}(I)}, \|z\|_{L^\infty(I_r,V)})(\|w_T\|_H + \|g_d\|_{L^2(I,V^*)}). \tag{2.33}$$

Proof. Using the continuity of $H^{2,1}(I) \subset L^\infty(I, V)$, we obtain $B_z'(u)^* w \in L^2(I, V^*)$ and

$$\|B_z'(u)^* w\|_{L^2(I,V^*)} \le \mathfrak{c}(\|u\|_{H^{2,1}(I)}, \|z\|_{L^\infty(I_r,V)})\|w\|_{L^2(I,V)}.$$

Therefore, from the proof of the previous theorem, we deduce that $w \in W^2(I)$, and the estimate (2.33) follows immediately from this estimate along with (2.28). □

Theorem 2.14. *If $w_T \in V$, $z \in L^\infty(I_r, V)$, $u \in H^{2,1}(I)$ and $g_d \in L^2(I, X)$, then the solution of (2.27) satisfies $w \in H^{2,1}(I)$. There exists a unique $\pi \in L^2(I, Y)$ such that*

$$-\partial_t w + \nu A w + B_z'(u)^* w + \nabla\pi = g_d \quad \text{in } L^2(I, X) \tag{2.34}$$

and there is a continuous function $\mathfrak{c} > 0$ such that

$$\|w\|_{H^{2,1}(I)} + \|\pi\|_{L^2(I,Y)} \le \mathfrak{c}(\|u\|_{H^{2,1}(I)}, \|z\|_{L^\infty(I_r,V)})(\|w_T\|_V + \|g_d\|_{L^2(I,X)}). \tag{2.35}$$

If in addition, $w_T = 0$, $u \in V^{2,1}(I) \cap L^2(I, D(A))$, $z \in V^{2,1}(I_r) \cap L^2(I_r, D(A))$, $z(0) = u_0$, $g_d \in H^1(I, V^*)$ and $g_d(0) \in X$, then $w \in V^{2,1}(J^r) \cap L^2(J^r, D(A))$. In particular, $w \in C(\bar{J}^r, V)$.

Proof. By uniqueness, the solution of (2.27) coincides with the one that can be constructed from the spectral Galerkin method. Therefore, to prove the above regularity, we can do the same strategy as in the case of linearized state equation. For this reason, we shall only formally derive the necessary a priori estimates. Let us set $u = z$ in Ω_r and $w = 0$ in Ω_{T+r}. We apply the test function Aw so that

$$\begin{aligned} -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|w(t)\|_V^2 + \nu\|Aw(t)\|_H^2 - b(u^r(t), w(t), Aw(t)) \\ - b(Aw(t), w^{-r}(t), u^{-r}(t)) \le \frac{1}{2\nu}\|g_d(t)\|_X^2 + \frac{\nu}{2}\|Aw(t)\|_H^2. \end{aligned} \tag{2.36}$$

The trilinear terms is estimated from above using Lemma 2.1(b) and (c) according to

$$|b(u^r(t), w(t), Aw(t))| \le c\|u^r(t)\|_H^2\|u^r(t)\|_V^2\|w(t)\|_V^2 + \frac{\nu}{4}\|Aw(t)\|_H^2$$
$$|b(Aw(t), w^{-r}(t), u^{-r}(t))| \le c\|Au^{-r}(t)\|_H^2\|w^{-r}(t)\|_V^2 + \frac{\nu}{8}\|Aw(t)\|_H^2.$$

Using these estimates in (2.36), integrating over $[0, t]$ and then applying a backward-in-time version of the Gronwall-type Lemma 7.1 to the resulting inequality, it follows that

$$\|w\|_{L^\infty(I,V)} + \|Aw\|_{L^2(I,H)} \le \mathfrak{c}(\|u\|_{H^{2,1}(I)}, \|z\|_{L^\infty(I_r,V)})(\|w_T\|_V + \|g_d\|_{L^2(I,X)}).$$

From this, an estimate for the time derivative of w is now available

$$\begin{aligned}\|\partial_t w\|_{L^2(I,H)} &\le \|Aw\|_{L^2(I,H)} + \|B_z'(u)^*w\|_{L^2(I,H)} + \|g_d\|_{L^2(I,X)}\\ &\le c((1+\|z\|_{L^\infty(I_r,V)} + \|u\|_{H^{2,1}(I)})\|Aw\|_{L^2(I,H)} + \|g_d\|_{L^2(I,X)}).\end{aligned}$$

Therefore, $w \in H^{2,1}(I)$ and the a priori estimate (2.35) without the *dual* pressure π is satisfied. Again, the existence of the dual pressure can be reasoned out as in the proof of Theorem 2.3. For the last statement, thanks to the compatibility of the homogenous terminal data and history, we can adapt the proof of Theorem 2.4 by differentiating the spectral Galerkin system approximating (2.6). Afterwards, one can utilize sequential compactness arguments to deduce that $w \in V^{2,1}(J^r) \cap L^2(J^r, D(A))$. □

Corollary 2.15. *Let $w_T \in V$, $z \in L^\infty(I_r, V)$, $g_d \in L^2(I,X)$ and $w_i = L_z(u_i)^{-*}(g_d + e_T w_T)$ for $i = 1,2$. Then there exists a constant $c > 0$ depending continuously on the norms of w_T, z, g_d, u_1 and u_2 in their indicated spaces such that $\|w_1 - w_2\|_{H^{2,1}(I)} \le c\|u_1 - u_2\|_{H^{2,1}(I)}$.*

Proof. The difference of w_1 and w_2 satisfies

$$w_1 - w_2 = L_z(u_1)^{-*}(B_z'(u_1)^*w_2 - B_z'(u_2)^*w_2) = L_z(u_1)^{-*}B_z'(u_1-u_2)^*w_2.$$

Thus, Theorem 2.14 gives us

$$\begin{aligned}\|w_1 - w_2\|_{H^{2,1}(I)} &\le c\|B_z'(u_1-u_2)^*w_2\|_{L^2(I,X)}\\ &\le c\|u_1-u_2\|_{H^{2,1}(I)}\|w_2\|_{H^{2,1}(I)} \le c\|u_1-u_2\|_{H^{2,1}(I)}\end{aligned}$$

where $c > 0$ is a constant as described by the corollary. □

Given $u_d \in L^2(I,X)$ and $u_T \in H$, let us define the *control-to-adjoint state* operator $D : Q \to W^{4/3}(I)$ by

$$D(q) = L_z(S(q))^{-*}(\alpha_{\Omega_T}(S(q) - u_d) + \alpha_R \nabla\times(\nabla\times S(q)) + \alpha_T e_T(S(q)(T) - u_T)). \tag{2.37}$$

In other words, $w = D(q)$ if and only if w is the weak solution of

$$\begin{cases} -\partial_t w + \nu Aw + B_z'(u)^*w = \alpha_{\Omega_T}(u-u_d) + \alpha_R\nabla\times(\nabla\times u) & \text{in } L^{4/3}(I,V^*),\\ w(T) = \alpha_T(u(T) - u_T) & \text{in } H,\end{cases} \tag{2.38}$$

where $u = S(q)$ and $\langle \nabla\times(\nabla\times u), \varphi\rangle_{L^2(I,V^*),L^2(I,V)} := (\nabla\times v, \nabla\times\varphi)_{L^2(I,L^2(\Omega))}$ for $\varphi \in L^2(I,V)$. The map D is locally bounded by Theorems 2.2 and 2.11.

Theorem 2.16. *The following properties of the control-to-adjoint state operator D hold:*

(a) If $u_0 \in V$, $z \in L^\infty(I_r,V)$, $f \in L^2(I,X)$, $u_d \in L^2(I,X)$ and $w_T \in H$, then $D : Q \to W^2(I)$ is locally bounded. If in addition, $w_T \in V$ then $D : Q \to H^{2,1}(I)$ is locally bounded.

(b) If $u_0 \in D(A)$, $z \in V^{2,1}(I_r) \cap L^2(I_r, D(A))$, $f \in H^1(I,V^*)$, $u_d \in H^1(I,X)$, $f(0) \in X$, $z(0) = z_0$ and $\alpha_T = 0$, then $D : Q \to V^{2,1}(J^r) \cap L^2(J^r, D(A))$ is locally bounded.

Proof. Part (a) is a direct consequence of Theorem 2.3, Corollary 2.13 and Theorem 2.14. On the other hand, (b) follows from (a), Theorem 2.4, $\nabla\times(\nabla\times u) \in H^1(I,V^*)$ since $u \in H^1(I,V)$, $\nabla\times(\nabla\times u_0) \in X$ and the last statement of Theorem 2.14. □

Corollary 2.17. *Let $u_0 \in V$, $z \in L^\infty(I_r,V)$, $f \in L^2(I,X)$, $u_d \in L^2(I,X)$ and $w_T \in V$. Given $q_1 \in Q$ and $q_2 \in Q$, there exists a constant $c > 0$ depending continuously on the norms of u_0, z, f, u_d, w_T, q_1 and q_2 in their indicated spaces such that $\|D(q_1) - D(q_2)\|_{H^{2,1}(I)} \le c\|q_1 - q_2\|_Q$.*

Proof. Let $u_1 = S(q_1)$, $u_2 = S(q_2)$, $u = u_1 - u_2$ and $q = q_1 - q_2$. The difference $D(q_1) - D(q_2)$ can be written as

$$\begin{aligned}D(q_1) - D(q_2) &= L_z(u_1)^{-*}(\alpha_{\Omega_T}u + \alpha_R\nabla\times(\nabla\times u) + \alpha_T e_T u(T))\\ &\quad + (L_z(u_1)^{-*} - L_z(u_2)^{-*})(\alpha_{\Omega_T}(u_2 - u_d) + \alpha_R\nabla\times(\nabla\times u_2) + \alpha_T e_T(u_2(T) - u_T))).\end{aligned}$$

Let d_1 and d_2 denote the terms on the right hand side. In the following, $c > 0$ will be a constant with the stated dependence on the given data. By the mean-value theorem and Remark 2.9, $\|u\|_{H^{2,1}(I)} \leq c\|q\|_Q$ and thus $\|u\|_{L^2(I,X)} + \|\nabla\times(\nabla\times u)\|_{L^2(I,L^2(\Omega))} + \|u(T)\|_V \leq c\|q\|_Q$. This inequality along with (2.35) yields $\|d_1\|_{H^{2,1}(I)} \leq c\|q\|_Q$. On the other hand, from the stability estimate in Corollary 2.15, we obtain that $\|d_2\|_{H^{2,1}(I)} \leq c\|u\|_{H^{2,1}(I)} \leq c\|q\|_Q$. The desired estimate now follows from the triangle inequality. □

3. Analysis of the Optimal Control Problem

In this section, we address the well-posedness of the optimal control problem (P) and establish the first order necessary and second order sufficient conditions for local optimality. Let us introduce the reduced cost functional $j : Q \to \mathbb{R}$ by $j(q) = J(S(q), q)$, where $S : Q \to W^2(I)$ is the control-to-state operator defined in Sect. 2.3. The optimization problem (P) is then equivalent to the unconstrained formulation

$$\min_{q\in Q} j(q). \tag{P}$$

Theorem 3.1. *Let $\alpha > 0$ and $\alpha_{\Omega_T}, \alpha_T, \alpha_R \geq 0$. Assume that $u_0 \in H$, $z \in V^2(I_r)$, $f \in L^2(I, V^*)$, $u_d \in L^2(I, X)$ and $u_T \in H$. Then (P) has a global solution $q^\star \in Q$, that is, $j(q^\star) \leq j(q)$ for every $q \in Q$.*

Proof. The proof follows a standard weak sequential argument in [40,50], which we outline for the sake of convenience. Let $\{q_k\}_{k=1}^\infty$ be a minimizing sequence, that is, $j(q_k) \to j^\star$, where $j^\star$ is the infimum of j. Thus $\frac{\alpha}{2}\|q_k\|_Q^2 < j^\star + 1$ for sufficiently large indices k. Hence, $\{q_k\}_{k=1}^\infty$ is bounded in Q, and therefore up to a subsequence, $q_k \rightharpoonup q^\star$ in Q. Let $u_k = S(q_k)$ and $u^\star = S(q^\star)$. From the weak-weak continuity of S, see Theorem 2.10, it follows that $u_k \rightharpoonup u^\star$ in $W^2(I)$. The continuity of the map $\varphi \mapsto \varphi(T)$ from $W^2(I)$ into H implies that $u_k(T) \rightharpoonup u^\star(T)$ in H. Also, we have $\nabla \times u_k \rightharpoonup \nabla \times u$ in $L^2(I, L^2(\Omega))$. Due to the lower semicontinuity of the norm in the weak topology, $j^\star \leq j(q^\star) = J(u^\star, q^\star) \leq \liminf_{k\to\infty} J(u_k, q_k) = \liminf_{n\to\infty} j(q_k) = j^\star$. Thus (P) has at least one solution $q^\star \in Q$ such that $j(q^\star) \leq j(q)$ for every $q \in Q$. □

Since j is a sum of squared-norms and $S \in C^\infty(Q, W^2(I))$, it follows from the chain rule that $j \in C^\infty(Q, \mathbb{R})$. Given $q \in Q$ and $g \in Q$, if $u = S(q)$, $v = S'(q)g$ and $w = D(q)$ denote the respective solutions of the state, linearized state and adjoint equations, then by standard arguments one can deduce the following representations for the first and second order directional derivatives of j

$$j'(q)g = (w + \alpha q, g)_Q \tag{3.1}$$

$$\begin{aligned} j''(q)[g, g] = {} & \alpha_{\Omega_T}\|v\|_{L^2(I,H)}^2 + \alpha_R\|\nabla\times v\|_{L^2(I,L^2(\Omega))}^2 + \alpha_T\|v(T)\|_H^2 \\ & - 2\langle B_z''(u)[v, v], w\rangle_{L^2(I,V^*),L^2(I,V)} + \alpha\|g\|_Q^2. \end{aligned} \tag{3.2}$$

We would like to point out that the regularity of the state is different from the regularity of the adjoint state if $\alpha_T > 0$. Also, unless $u(T) = u_T$, which is unlikely in practice, we do not have the compatibility condition for the terminal data and dual history in the adjoint equation. Thus, a presence of delay in the state equation impedes further smoothness with respect to time on the adjoint state, and hence on the control.

A control $q^\star \in Q$ is said to be a *local solution* to (P) if there exists $\delta > 0$ such that $j(q^\star) \leq j(q)$ for every $q \in Q$ with $\|q - q^\star\|_Q < \delta$. For the first and second order necessary condition, the following theorem can be established by classical arguments of unconstrained optimization in Hilbert spaces, we refer to [38,53] for the details.

Theorem 3.2. *Suppose that the assumptions of the previous theorem are fulfilled and let $q^\star$ be a local solution to (P) and $u^\star = S(q^\star)$. Then $q^\star = -\alpha^{-1}D(q^\star) \in W^{4/3}(I)$. Furthermore, $j''(q^\star)[g, g] \geq 0$ for every $g \in Q$.*

We now formulate a second order sufficient condition with the assumption that the optimal state is close enough to the desired data. This is typically observed numerically when the penalty parameter $\alpha > 0$ is chosen to be sufficiently small.

Theorem 3.3. *Let $u_0 \in H$, $z \in V^2(I_r)$ and $f \in L^2(I, V^*)$. Suppose that $q^\star \in Q$ satisfies $j'(q^\star)q = 0$ for every $q \in Q$. Then there exist constants $\eta > 0$ and $\mu = \mu_{\alpha,\eta} > 0$ such that if $u^\star = S(q^\star)$ and*

$$\sqrt{\alpha_{\Omega_T}}\|u^\star - u_d\|_{L^2(I,X)} + \sqrt{\alpha_T}\|u^\star(T) - u_T\|_H + \sqrt{\alpha_R}\|\nabla \times u^\star\|_{L^2(I,X)} < \eta \tag{3.3}$$

then $j''(q^\star)[q, q] \geq \mu\|q\|_Q^2$ for every $q \in Q$. In particular, $q^\star$ is a strict local solution of (P).

Proof. From the characterization of the second derivative of j in (3.2), if we let $v = S'(q^\star)q$ and $w^\star = D(q^\star)$, then for some $c_1 > 0$ we have $j''(q^\star)[q, q] \geq \alpha\|q\|_Q^2 - c_1\|v\|_{W_0^2(I)}^2\|w^\star\|_{L^2(I,V)}$. Note that $\|v\|_{W_0^2(I)}^2 \leq c_2\|q\|_Q^2$ where $c_2 = \|S'(q^\star)\|_{\mathcal{L}(Q,W_0^2(I))}^2$. Assuming that (3.3) holds, then we get from (2.33) that $\|w^\star\|_{L^2(I,V)} \leq c_2\eta$ for a constant $c_2 > 0$ independent of η. Thus, we take $\eta > 0$ small enough so that $\mu := \alpha - c_1c_2c_3\eta > 0$. In this case, we obtain $j''(q^\star)[q, q] \geq \mu\|q\|_Q^2$ for every $q \in Q$. The fact that $q^\star$ is a strict local solution follows from the coercivity of $j''(q^\star)$. □

We close this section by stating the improved regularity for the optimal state and optimal control under additional regularity and compatibility of the initial data and history.

Theorem 3.4. *An optimal triple $(q^\star, u^\star, w^\star)$ for* (P)*, where $q^\star$ is a local solution, $u^\star = S(q^\star)$ and $w^\star = D(q^\star)$, satisfies the following:*

(a) If $u_0 \in V$, $z \in L^\infty(I_r, V)$, $f \in L^2(I, X)$, $u_d \in L^2(I, X)$ and $w_T \in H$, then $u^\star \in H^{2,1}(I)$ and $w^\star$, $q^\star \in W^2(I)$. In addition, if $w_T \in V$, then $w^\star$, $q^\star \in H^{2,1}(I)$.

(b) If $u_0 \in D(A)$, $z \in V^{2,1}(I_r) \cap L^2(I_r, D(A))$, $f \in H^1(I, V^*)$, $u_d \in H^1(I, X)$, $f(0) \in X$ and $z(0) = u_0$, then $u^\star \in V^{2,1}(J_r) \cap L^2(J_r, D(A))$. If in addition $\alpha_T = 0$, then $w^\star \in V^{2,1}(J^r) \cap L^2(J^r, D(A))$ and $q^\star \in V^{2,1}(I) \cap L^2(I, D(A))$.

Proof. Apply Theorems 2.3, 2.4 and 2.16. □

4. Galerkin Finite Element Discretization

The goal of this section is to present a numerical scheme for the finite-dimensional approximation of solutions to the control problem (P). In the forthcoming discussion, the following assumption for the regularity and compatibility of the initial data, history and target data will be imposed:

(A1) $u_0 \in D(A)$, $u_T \in D(A)$, $z \in H^{2,1}(I_r)$, $u_d \in H^{2,1}(I)$, $f \in H^{2,1}(I)$ and $z(0) = u_0$.

From the results of the previous section, (A1) implies that $u = S(q) \in H^{2,1}(J_r)$, $v = S'(q)g \in H^{2,1}(J_r)$ and $w = D(q) \in H^{2,1}(I)$ for every $q \in Q$ and $g \in Q$.

4.1. Finite Element Spaces and Approximation Operators

Let $\mathcal{K}_h = \{K_h\}$ for $h > 0$ be a family of triangulations of a convex polygonal domain Ω parametrized by the mesh size h, that is, the length of the largest triangle edge in the subdivision. Let W_h and M_h be finite dimensional subspaces W and M, respectively, and define

$$V_h := \{u_h \in W_h : (\operatorname{div} u_h, p_h)_{L^2} = 0 \ \forall p_h \in M_h\}.$$

The following assumptions on these finite dimensional subspaces will be considered:

(A2) There exist finite element approximation operators $\mathit{\Pi}_h : D(A) \to W_h$ and $\mathit{\Gamma}_h : Y \to M_h$ such that, for some constant $c > 0$, we have $\|u - \mathit{\Pi}_h u\|_X + h\|u - \mathit{\Pi}_h u\|_W \leq ch^2\|Au\|_H$ and $\|p - \mathit{\Gamma}_h p\|_M \leq ch\|p\|_Y$ for every $u \in D(A)$ and $p \in Y$.

(A3) We have the inverse estimate $\|u_h\|_W \leq ch^{-1}\|u_h\|_X$ for every $u_h \in W_h$.

(A4) The pair (W_h, M_h) satisfies the uniform discrete inf-sup condition

$$\inf_{u_h \in W_h \setminus \{0\}} \sup_{p_h \in M_h \setminus \{0\}} \frac{(\operatorname{div} u_h, p_h)_{L^2}}{\|u_h\|_W\|p_h\|_M} \geq c > 0.$$

These assumptions are satisfied for the mini-element [3] and the Taylor-Hood finite element spaces [30], and if the family of triangulations is shape-regular. The latter means that there is a constant $c > 0$ such that, if e_{K_h} and δ_{K_h} denote the respective largest edge and diameter of the largest ball contained in K_h, then $h/e_{K_h} \leq c$ and $e_{K_h}/\delta_{K_h} \leq c$. In other words, the smallest interior angle of each triangle should not tend to zero as $h \to 0$.

Let $P_h : X \to V_h$ be the L^2-projection onto V_h, that is, for $u \in X$ let $(P_h u, v_h)_X = (u, v_h)_X$ for all $v_h \in V_h$. It is well known that the operator P_h satisfies the stability and error estimates $\|P_h w\|_X \leq c\|w\|_X$, $\|u - P_h u\|_X + h\|u - P_h u\|_W \leq ch\|u\|_W$ and $\|v - P_h v\|_X + h\|v - P_h v\|_W \leq ch^2\|Av\|_H$ whenever $w \in X$, $u \in W$ and $v \in D(A)$. For further details, the reader is referred to [17,30].

We extend the above projections to the time-dependent case according to $\Pi_h : L^2(I, D(A)) \to L^2(I, W_h)$, $\Gamma_h : L^2(I, Y) \to L^2(I, M_h)$ and $P_h : L^2(I, X) \to L^2(I, V_h)$. For instance, $(P_h u)(t) = P_h(u(t))$ for a.e. $t \in I$. Similarly, these operators will be considered in the history interval I_r.

To simplify the exposition, let us assume that $T = n_0 r$ for some positive integer n_0 and take a uniform temporal step size. Partition the history interval $\bar{I}_r = [-r, 0]$ with grid size $\tau = r/N_r$ into $-r = t_{-N_r} < \cdots < t_{-1} < t_0 = 0$, where $t_{-j} = -j\tau$ for $j = 0, 1, \ldots, N_r$. Likewise, subdivide the time domain $\bar{I} = [0, T]$ with grid size τ into $0 = t_0 < t_1 < \cdots < t_{N_\tau} = T$, where $t_j = j\tau$ for $j = 0, 1, \ldots, N_\tau$ and $N_\tau = N_r n_0$. For each $j = -N_r, \ldots, N_\tau$, let $I_j = (t_{j-1}, t_j]$. In the case of the adjoint equation, we likewise partition $\bar{I}^r = [T, T + r]$ into intervals $I_j = (t_{j-1}, t_j]$ for $j = N_\tau + 1, \cdots, N_\tau + N_r$. We denote by $\sigma := (\tau, h)$ for the pair of temporal and spatial mesh sizes. Note that $|\sigma|^2 \leq T^2 + \mathrm{diam}(\Omega)^2$, where $\mathrm{diam}(\Omega)$ is the diameter of Ω. As mentioned in the introduction, we take the following stability condition:

(A5) There exists $c > 0$ such that $\tau \leq ch^2$ for every $\sigma = (\tau, h)$.

All throughout this section, the assumptions (A1)–(A5) shall be implicitly imposed.

For the temporal discretization, we consider a discontinuous Galerkin scheme. This is a variant of the backward Euler method, where time-evaluation is replaced by time-averaging. In this direction, we denote

$$\mathcal{P}_\tau(I, Z) = \left\{ v_\sigma \in L^2(I, Z) : v_\sigma = \sum_{k=1}^{N_\tau} v_h^k \mathbb{1}_{I_k},\ v_h^k \in Z,\ k = 1, \ldots, N_\tau \right\}$$

$$\mathcal{P}_\tau(I_r, Z) = \left\{ z_\sigma \in L^2(I_r, Z) : z_\sigma = \sum_{k=0}^{N_r - 1} z_h^{-k} \mathbb{1}_{I_{-k}}\ \ z_h^{-k} \in Z,\ k = 0, \ldots, N_r - 1 \right\}$$

the space of piecewise constant functions on I and I_r with values in a Hilbert space Z, respectively, corresponding to the above partitions of I and I_r. Here, $\mathbb{1}_{I_k}$ is the indicator function of the interval I_k. Given $v_\sigma \in \mathcal{P}(I, Z)$, we write $v_h^k := v_\sigma|_{I_k}$ for each $k = 1, \ldots, N_\tau$, and similarly for the elements of $\mathcal{P}_\tau(I_r, Z)$. By definition, $\mathcal{P}_\tau(I, Z) \subset L^\infty(I, Z)$ and for every $v_\sigma \in \mathcal{P}_\tau(I, Z)$ there holds

$$\|v_\sigma\|^2_{L^\infty(I,Z)} = \max_{1 \leq k \leq N_\tau} \|v_h^k\|^2_Z, \qquad \|v_\sigma\|^2_{L^2(I,Z)} = \sum_{k=1}^{N_\tau} \tau \|\nabla v_h^k\|^2_Z. \tag{4.1}$$

Proposition 4.1. *Let $s_k \in I_k$ for each $k = 1, \ldots, N_\tau$. For every $u \in H^1(I, X)$, we have*

$$\left(\sum_{k=1}^{N_\tau} \int_{I_k} \|u - u(s_k)\|^2_X \,\mathrm{d}t \right)^{1/2} \leq \tau \|\partial_t u\|_{L^2(I,X)},$$
$$\max_{1 \leq k \leq N_\tau} \|u - u(s_k)\|_{L^\infty(I_k, X)} \leq \sqrt{\tau} \|\partial_t u\|_{L^2(I,X)}.$$

Proof. For each $t \in I_k$, let t_k^+ and t_k^- denote largest and smallest between t and s_k, respectively. By the Cauchy–Schwarz inequality

$$\|u(t) - u(s_k)\|^2_X \leq |s_k - t| \int_{t_k^-}^{t_k^+} \|\partial_t u(s)\|^2_X \,\mathrm{d}s \leq \tau \int_{I_k} \|\partial_t u(s)\|^2_X \,\mathrm{d}s. \tag{4.2}$$

Note that point-wise time evaluation is admissible since $H^1(I,X) \subset C(\bar{I},X)$. Integrating (4.2) over the interval I_k and then taking the sum over all $k = 1,\dots,N_\tau$, we get

$$\sum_{k=1}^{N_\tau} \int_{I_k} \|u(t) - u(s_k)\|_X^2 \, \mathrm{d}t \le \tau^2 \|\partial_t u\|_{L^2(I,X)}^2.$$

Taking square roots yields the first inequality. Getting the supremum over all $t \in I_k$ in (4.2) and then the maximum over all $1 \le k \le N_\tau$, one obtains the second inequality. □

Proposition 4.2. *Let $s_k \in I_k$ for each $k = 1,\dots,N_\tau$. There exists a constant $c > 0$ such that for every $u \in H^{2,1}(I)$, the following error estimates hold:*

$$\max_{1\le k\le N_\tau} \|u - P_h u(s_k)\|_{L^\infty(I_k,X)} \le c(\sqrt{\tau} + h)\|u\|_{H^{2,1}(I)}$$

$$\left(\sum_{k=1}^{N_\tau} \int_{I_k} \|u - P_h u(s_k)\|_X^2 \, \mathrm{d}t\right)^{1/2} + h\left(\sum_{k=1}^{N_\tau} \int_{I_k} \|\nabla u - \nabla P_h u(s_k)\|_X^2 \, \mathrm{d}t\right)^{1/2} \le c(\tau + h^2)\|u\|_{H^{2,1}(I)}.$$

Proof. Suppose that $u \in H^{2,1}(I)$. On the interval I_k, write $u - P_h u(s_k) = (u - P_h u) + P_h(u - u(s_k))$. Using the boundedness of $P_h : X \to X$, Proposition 4.1 and the continuity of $H^{2,1}(I) \subset C(\bar{I}, V)$

$$\begin{aligned}\|u - P_h u(s_k)\|_{L^\infty(I_k,X)} &\le \|u - P_h u\|_{L^\infty(I_k,X)} + \|P_h(u - u(s_k))\|_{L^\infty(I_k,X)} \\ &\le c(h\|u\|_{L^\infty(I,V)} + \sqrt{\tau}\|\partial_t u\|_{L^2(I,H)}).\end{aligned}$$

Getting the maximum over all indices $1 \le k \le N_\tau$ proves the first estimate. On the other hand,

$$\begin{aligned}\sum_{k=1}^{N_\tau} \int_{I_k} \|u - P_h u(s_k)\|_X^2 \, \mathrm{d}t &\le 2\sum_{k=1}^{N_\tau} \int_{I_k} \|u - P_h u\|_X^2 \, \mathrm{d}t + 2\sum_{k=1}^{N_\tau} \int_{I_k} \|P_h(u - u(s_k))\|_X^2 \, \mathrm{d}t \\ &\le c(h^4\|Au\|_{L^2(I,H)}^2 + \tau^2\|\partial_t u\|_{L^2(I,H)}^2).\end{aligned}$$

Applying the inverse estimate (A3), we obtain that

$$h^2 \sum_{k=1}^{N_\tau} \int_{I_k} \|\nabla u - \nabla P_h u(s_k)\|_X^2 \, \mathrm{d}t \le c \sum_{k=1}^{N_\tau} \int_{I_k} \|u - P_h u(s_k)\|_X^2 \, \mathrm{d}t.$$

Taking the sum of the last two inequalities and then the square roots prove the second estimate. □

The special cases where $s_k = t_{k-1}$ or $s_k = t_k$ will be utilized in our analysis. If one wishes to use the approximation operator Π_h instead of the projection operator P_h, which is typical in practice, then due to the limited regularity of the initial history and source term, time-evaluation is not applicable anymore and has to be replaced by time-averaging. This approach was introduced in [5] for hereditary control problems with ordinary differential equations. Define the linear operator $R_\tau : L^2(I,X) \to \mathcal{P}_\tau(I,X)$ by

$$R_\tau u = \sum_{k=1}^{N_\tau} \left(\frac{1}{\tau}\int_{I_k} u(t) \, \mathrm{d}t\right) \mathbb{1}_{I_k}$$

with the obvious modification when I is replaced by I_r. Then, $\|R_\tau u\|_{L^2(I,X)} \le \|u\|_{L^2(I,X)}$ and $\|R_\tau w\|_{L^\infty(I,X)} \le \|w\|_{L^\infty(I,X)}$ for every $u \in L^2(I,X)$ and $w \in L^\infty(I,X)$. It can also be shown that there is a constant $c > 0$ such that $\|R_\tau v - v\|_{L^2(I,X)} \le \tau\|v\|_{H^1(I,X)}$ and $\|R_\tau v - v\|_{L^\infty(I,X)} \le \sqrt{\tau}\|v\|_{H^1(I,X)}$ for every $v \in H^1(I,X)$. With these, the following error estimates can be established by adapting the proof of Proposition 4.2.

Proposition 4.3. *Consider the operator $\Pi_h R_\tau : L^2(I, D(A)) \to \mathcal{P}_\tau(I, W_h)$. Then there exists a constant $c > 0$ such that for every $u \in H^{2,1}(I)$ there holds*

$$\|\Pi_h R_\tau u - u\|_{L^\infty(I,X)} \le c(\sqrt{\tau} + h)\|u\|_{H^{2,1}(I)}$$

$$\|\Pi_h R_\tau u - u\|_{L^2(I,X)} + h\|\nabla \Pi_h R_\tau u - \nabla u\|_{L^2(I,X)} \le c(\tau + h^2)\|u\|_{H^{2,1}(I)}.$$

These estimates are also valid when $\Pi_h R_\tau$ is replaced by $P_h R_\tau : L^2(I, D(A)) \to \mathcal{P}_\tau(I, V_h)$.

4.2. Fully Discrete Optimal Control Problem

Here, we present the full space-time discretization of the optimal control problem (P). Given a control $q \in Q$ and discrete initial data $(u_{0h}, z_\sigma) \in W_h \times \mathcal{P}_\tau(I_r, W_h)$, consider the following discrete problem: Find $u_\sigma = \sum_{k=1}^{N_\tau} u_h^k \mathbb{1}_{I_k} \in \mathcal{P}_\tau(I, V_h)$ such that for $k = 1, \ldots, N_\tau$

$$\begin{cases} (d_\tau u_h^k, \varphi_h)_X + \nu(\nabla u_h^k, \nabla \varphi_h)_X \\ \qquad = b(u_h^{k-N_r}, \varphi_h, u_h^{k-1}) + \dfrac{1}{\tau}\displaystyle\int_{I_k} (f_\sigma(t) + q(t), \varphi_h)_X \,\mathrm{d}t \quad \forall \varphi_h \in V_h, \\ u_h^j = z_h^j \quad \text{for } j = 1 - N_r, \ldots, 0, \end{cases} \tag{4.3}$$

where $d_\tau u_h^1 := \tau^{-1}(u_h^1 - u_{0h})$ and $d_\tau u_h^k := \tau^{-1}(u_h^k - u_h^{k-1})$ for $k = 2, \ldots, N_\tau$. Notice that the initial data u_{0h} is only applied at the first time step, while the history z_σ was utilized in the trilinear form b. This is an implicit-explicit scheme in the sense that the Laplacian and delay are discretized implicitly, while the convection term is discretized explicitly. More precisely, on each interval I_k, the following approximations were adapted:

$$\int_{I_k} (\nabla u(t), \nabla \varphi_h) \,\mathrm{d}t \approx \tau(\nabla u(t_k), \nabla \varphi_h)$$

$$\int_{I_k} b(u^r(t), \varphi_h, u(t)) \,\mathrm{d}t \approx \tau b(u^r(t_k), \varphi_h, u(t_{k-1})).$$

The scheme (4.3) is one of the simplest possible discretization of the state equation in (P), where it is possible to prove stability and error estimates. Observe that the matrices of the corresponding linear system to (4.3) are the same at every time step. This will also be the case of the discrete adjoint problem below. For gradient-based optimization algorithms, where one has to solve both the discrete state and adjoint equations in order to get a directional derivative, this is advantageous. For instance, one can pre-factorize the matrix before doing the primal and dual solvers for efficiency.

As with the continuous case, we will set $u_\sigma = z_\sigma$ on Ω_r so that $u_\sigma \in L^2(J_r, W_h)$, that is,

$$u_\sigma = \sum_{k=0}^{1-N_r} z_h^{-k} \mathbb{1}_{I_{-k}} + \sum_{k=1}^{N_\tau} u_h^k \mathbb{1}_{I_k}. \tag{4.4}$$

We note that (4.3) is similar to the one presented in [51], the main difference is that the trilinear term there is $b(u_h^{k-1-N_r}, \varphi_h, u_h^{k-1})$, that is, the convection is discretized explicitly. Also, spatial discretization was not considered. The one given in (4.3) is a natural choice based on (4.4), where the solution and the history of the continuous problem are evaluated at the right endpoint of the interval I_k for each $k = 1 - N_r, \ldots, N_\tau$. This is also suitable in the case when the initial data and history are not compatible. Discontinuous Galerkin time-schemes of arbitrary order for the 2D and 3D Navier–Stokes equation can be found in [16].

With regard to the initial data, history and source term, we shall take the approximations $u_{0h} = \Pi_h u_0 \in W_h$, $z_\sigma = R_\tau \Pi_h z \in \mathcal{P}_\tau(I, W_h)$ and $f_\sigma = R_\tau \Pi_h f \in \mathcal{P}_\tau(I_r, W_h)$. Thanks to (A1), (A2), (A5) and Proposition 4.3, there is a constant $c > 0$ independent on σ such that

$$\|u_{0h} - u_0\|_X + h\|u_{0h} - u_0\|_W \le ch^2 \|Au_0\|_H \tag{4.5}$$

$$\|z_\sigma - z\|_{L^2(I_r,X)} + h(\|z_\sigma - z\|_{L^\infty(I_r,X)} + \|z_\sigma - z\|_{L^2(I_r,W)}) \le ch^2 \|z\|_{H^{2,1}(I_r)} \tag{4.6}$$

$$\|f_\sigma - f\|_{L^2(I,X)} + h(\|f_\sigma - f\|_{L^\infty(I,X)} + \|f_\sigma - f\|_{L^2(I,W)}) \le ch^2 \|f\|_{H^{2,1}(I)}. \tag{4.7}$$

The existence and uniqueness of $u_\sigma \in \mathcal{P}_\tau(I, V_h)$ satisfying (4.3) follows immediately from the Lax-Milgram Lemma and by induction. In fact, it is enough to observe that the finite-dimensional square system for each k is injective. Let us define the *discrete control-to-state operator* $S_\sigma : Q \to \mathcal{P}_\tau(I, V_h)$ by $S_\sigma(q) = u_\sigma$ if and only if u_σ is the solution of (4.3). By the discrete inf-sup condition (A4), (4.3) is equivalent to the problem of finding a pair $(u_\sigma, p_\sigma) \in \mathcal{P}_\tau(I, W_h) \times \mathcal{P}_\tau(I, M_h)$ satisfying the following system of mixed problems:

$$\begin{cases} (d_\tau u_h^k,\varphi_h)_X + \nu(\nabla u_h^k, \nabla\varphi_h)_X - (\operatorname{div}\varphi_h, p_h^k)_{L^2} + (\operatorname{div} u_h^k, \rho_h)_{L^2} \\ \qquad = b(u_h^{k-N_r}, \varphi_h, u_h^{k-1}) + \dfrac{1}{\tau}\displaystyle\int_{I_k} (f_\sigma(t) + q(t), \varphi_h)_X \, \mathrm{d}t \quad \forall(\varphi_h, \rho_h) \in W_h \times M_h, \\ u_h^j = z_h^j \quad \text{for } j = 1 - N_r, \ldots, 0, \end{cases} \tag{4.8}$$

for each $k = 1, \ldots, N_\tau$, see [8]. The scheme is thus conforming with respect to W but not with V, since $W_h \subset W$ while $V_h \not\subset V$. Whereas (4.3) is used in the analysis, the more practical mixed problem (4.8) is the one that is utilized in the numerical implementation.

Remark 4.4. The discrete solution operator S_σ is invariant under left composition by $P_h R_\tau$, that is, $S_\sigma = S_\sigma P_h R_\tau$ as a map from Q into $\mathcal{P}_\tau(I, V_h)$.

We shall take $Q_\sigma := \mathcal{P}_\tau(I, W_h) \subset Q$ as the discretization of the control space Q. The fully discrete optimal control problem is then given by

$$\min_{q_\sigma \in Q_\sigma} j_\sigma(q_\sigma) \tag{P$_\sigma$}$$

where $j_\sigma : \mathcal{P}_\tau(I, V_h) \to \mathbb{R}$ is the discrete analogue of j given by

$$\begin{aligned} j_\sigma(q_\sigma) &= \frac{1}{2}\int_I \alpha_{\Omega_T}\|u_\sigma - u_{d\sigma}\|_X^2 + \alpha_R\|\nabla\times u_\sigma\|_{L^2}^2 \, \mathrm{d}t + \frac{\alpha_T}{2}\|u_\sigma(T) - u_{Th}\|_X^2 + \frac{\alpha}{2}\|q_\sigma\|_Q^2 \\ &= \frac{\tau}{2}\sum_{k=1}^{N_\tau}(\alpha_{\Omega_T}\|u_h^k - u_{dh}^k\|_X^2 + \alpha_R\|\nabla\times u_h^k\|_{L^2}^2) + \frac{\alpha_T}{2}\|u_h^{N_\tau} - u_{Th}\|_X^2 + \frac{\alpha\tau}{2}\sum_{k=1}^{N_\tau}\|q_h^k\|_X^2 \end{aligned}$$

with $u_\sigma = S_\sigma(q_\sigma) \in \mathcal{P}_\tau(I, V_h)$, $u_{d\sigma} = \mathit{\Pi}_h R_\tau u_d \in \mathcal{P}_\tau(I, W_h)$ and $u_{Th} = \mathit{\Pi}_h u_T \in W_h$. Again, from (A1), (A2), (A5) and Proposition 4.3, the following error estimates for the target data hold for some constant $c > 0$ independent on σ:

$$\|u_{Th} - u_T\|_X + h\|u_{Th} - u_T\|_W \le ch^2\|Aw_T\|_H \tag{4.9}$$

$$\|u_{d\sigma} - u_d\|_{L^2(I,X)} + h(\|u_{d\sigma} - u_d\|_{L^\infty(I,X)} + \|u_{d\sigma} - u_d\|_{L^2(I,W)}) \le ch^2\|u_d\|_{H^{2,1}(I)}. \tag{4.10}$$

The existence of a global solution to the finite-dimensional optimization problem (P$_\sigma$) is immediate since j_σ is continuous and coercive, that is, $j_\sigma(q_\sigma) \to \infty$ as $\|q_\sigma\|_Q \to \infty$. We will see later that replacing Q_σ by $\mathcal{P}_\tau(I, V_h)$ leads to an equivalent problem, see Remark 4.13 below.

4.3. Error Estimates for the Discrete State Equation

In this subsection we analyze the discrete state equation (4.3) and prove error estimates. First, let us establish the local boundedness of solutions.

Theorem 4.5. *Let $q \in Q$. If $u_\sigma = S_\sigma(q)$, then there exists a continuous function $\mathfrak{c} > 0$ independent of σ such that*

$$\|u_\sigma\|_{L^\infty(I,X)} + \|u_\sigma\|_{L^2(I,W)} \le \mathfrak{c}(\|q\|_Q, \|f\|_{L^2(I,X)}, \|u_0\|_H, \|z\|_{L^\infty(I_r,H)\cap L^2(I_r,V)}). \tag{4.11}$$

Proof. Let $\widehat{u}_h^0 := u_{0h}$ and $\widehat{u}_h^k := u_h^k$ for $k = 1, \ldots, N_\tau$ so that $d_\tau u_h^k = \tau^{-1}(u_h^k - \widehat{u}_h^{k-1})$. Taking the test function $\varphi_h = 2\tau u_h^k$ in (4.3) and using the equation $2(v - w, v)_X = \|v\|_X^2 - \|w\|_X^2 + \|v - w\|_X^2$, we obtain for $k = 1, \ldots, N_\tau$ that

$$\begin{aligned} &\|u_h^k\|_X^2 - \|\widehat{u}_h^{k-1}\|_X^2 + \|u_h^k - \widehat{u}_h^{k-1}\|_X^2 + 2\nu\tau\|\nabla u_h^k\|_X^2 \\ &\qquad \le 2\tau b(u_h^{k-N_r}, u_h^k, u_h^{k-1}) + 2\int_{I_k}(f_\sigma(t) + q(t), u_h^k)_X \, \mathrm{d}t. \end{aligned} \tag{4.12}$$

We can estimate the second term on right hand side using (A2) and the Young and Poincaré inequalities by

$$2\int_{I_k}(f_\sigma(t)+q(t),u_h^k)_X\,\mathrm{d}t \le \frac{c}{\nu}\int_{I_k}\|f(t)\|_X^2+\|q(t)\|_X^2\,\mathrm{d}t+\frac{\nu\tau}{2}\|\nabla u_h^k\|_X^2. \tag{4.13}$$

Given $\varepsilon>0$, by the Hölder, Gagliardo–Nirenberg and Young inequalities, the trilinear term is estimated from above as follows:

$$\begin{aligned}&2\tau b(u_h^{k-N_r},u_h^k,u_h^{k-1})\\&\quad\le \frac{\nu\tau}{2}\|\nabla u_h^k\|_X^2+\varepsilon\tau\|u_h^{k-N_r}\|_X^2\|\nabla u_h^{k-1}\|_X^2+\frac{c\tau}{\varepsilon}\|\nabla u_h^{k-N_r}\|_X^2\|u_h^{k-1}\|_X^2.\end{aligned} \tag{4.14}$$

Plugging the estimates (4.13) and (4.14) in (4.12), and then taking the sum over all $k=1,\dots,j$ with $1\le j\le N_\tau$, yields the following:

$$\begin{aligned}\|u_h^j\|_X^2+\sum_{k=1}^{j}\nu\tau\|\nabla u_h^k\|_X^2 &\le \|u_{0h}\|_X^2+\varepsilon\tau\|z_h^{1-N_r}\|_X^2\|\nabla z_h^0\|_X^2+\frac{c\tau}{\varepsilon}\|\nabla z_h^{1-N_r}\|_X^2\|z_h^0\|_X^2\\&\quad+\frac{c}{\nu}\sum_{k=1}^{j}\int_{I_k}\|f(t)\|_X^2\,\mathrm{d}t+\frac{c}{\nu}\sum_{k=1}^{j}\int_{I_k}\|q(t)\|_X^2\,\mathrm{d}t+\varepsilon\sum_{k=1}^{j-1}\tau\|u_h^{k+1-N_r}\|_X^2\|\nabla u_h^k\|_X^2\\&\quad+\frac{c}{\varepsilon}\sum_{k=1}^{j-1}\tau\|\nabla u_h^{k+1-N_r}\|_X^2\|u_h^k\|_X^2.\end{aligned} \tag{4.15}$$

From (A2), (4.1) and the boundedness of $R_\tau: V^2(I_r)\to V^2(I_r)$, there is some positive constant c independent on σ such that the sum of the first three terms on the right hand side of (4.15) can be estimated by

$$\begin{aligned}&\|u_{0h}\|_X^2+\varepsilon\tau\|z_h^{1-N_r}\|_X^2\|\nabla z_h^0\|_X^2+\frac{c\tau}{\varepsilon}\|\nabla z_h^{1-N_r}\|_X^2\|z_h^0\|_X^2\\&\qquad\le c\left(\|u_0\|_H^2+\left(1+\frac{1}{\varepsilon^2}\right)\varepsilon\|z\|_{L^\infty(I_r,H)}^2\|z\|_{L^2(I_r,V)}^2\right).\end{aligned} \tag{4.16}$$

We shall prove by induction that for each $\ell=1,\dots,n_0$, there is a continuous function $\mathfrak{c}_\ell>0$ such that if $\mathfrak{c}_\ell(q,f,u_0,z):=\mathfrak{c}_\ell(\|q\|_Q,\|f\|_{L^2(I,X)},\|u_0\|_H,\|z\|_{L^\infty(I_r,H)\cap L^2(I_r,V)})$, then

$$\max_{1\le k\le \ell N_r}\|u_h^k\|_X^2+\sum_{k=1}^{\ell N_r}\frac{\nu\tau}{2}\|\nabla u_h^k\|_X^2\le \mathfrak{c}_\ell(q,f,u_0,z). \tag{4.17}$$

Consider the case $\ell=1$. From (4.6), we get for some constant $c_0>0$ that $\|u_h^k\|_X=\|z_h^k\|_X\le c_0\|z\|_{L^\infty(I_r,H)}$ for each $1-N_r\le k\le 0$. Taking $\varepsilon=\varepsilon_1:=\nu/(2c_0^2\|z\|_{L^\infty(I_r,H)}^2+1)$, we have

$$\varepsilon\sum_{k=1}^{j-1}\tau\|u_h^{k+1-N_r}\|_X^2\|\nabla u_h^k\|_X^2\le \varepsilon c_0^2\|z\|_{L^\infty(I_r,H)}^2\sum_{k=1}^{j-1}\tau\|\nabla u_h^k\|_X^2\le \sum_{k=1}^{j-1}\frac{\nu\tau}{2}\|\nabla u_h^k\|_X^2 \tag{4.18}$$

each for $1\le j\le N_r$. Thus, we obtain from (4.15), (4.16) with $\varepsilon=\varepsilon_1$, and (4.18) that

$$\begin{aligned}\|u_h^j\|_X^2+\sum_{k=1}^{j}\frac{\nu\tau}{2}\|\nabla u_h^k\|_X^2 &\le c\left(\|u_0\|_H^2+\left(1+\frac{1}{\varepsilon_1^2}\right)\frac{\nu}{2c_0^2}\|z\|_{L^2(I_r,V)}^2\right)\\&\quad+\frac{c}{\nu}\int_I\|f(t)\|_X^2\,\mathrm{d}t+\frac{c}{\nu}\int_I\|q(t)\|_X^2\,\mathrm{d}t+\frac{c}{\varepsilon_1}\sum_{k=1}^{j-1}\tau\|\nabla u_h^{k+1-N_r}\|_X^2\|u_h^k\|_X^2.\end{aligned} \tag{4.19}$$

Let $c(\varepsilon_1, q, f, u_0, z)$ denote the sum of the first three terms on the right hand side. Applying the discrete Gronwall Lemma 7.2 to the previous inequality, we have

$$\max_{1\le k\le N_r} \|u_h^k\|_X^2 + \sum_{k=1}^{N_r} \frac{\nu\tau}{2}\|\nabla u_h^k\|_X^2 \le e^{\gamma_1} c(\varepsilon_1, q, f, u_0, z),$$

where

$$\gamma_1 = \frac{c}{\varepsilon_1}\sum_{k=1}^{N_r-1} \tau\|\nabla u_h^{k+1-N_r}\|_X^2 \le \frac{c}{\varepsilon_1}\|z\|_{L^2(I_r,V)}^2.$$

The last two inequalities verify (4.17) for $\ell = 1$.

Now, let us assume that (4.17) is satisfied for ℓ. Then by taking

$$\varepsilon = \varepsilon_{\ell+1} := \nu(2\max\{c_0^2\|z\|_{L^\infty(I_r,H)}^2, \mathfrak{c}_\ell(q, f, u_0, z)\} + 1)^{-1} < \varepsilon_1$$

we obtain (4.18) for every $1 \le j \le (\ell+1)N_r$. Thus, adapting the above procedure leads to

$$\max_{1\le k\le(\ell+1)N_r} \|u_h^k\|_X^2 + \sum_{k=1}^{(\ell+1)N_r} \frac{\nu\tau}{2}\|\nabla u_h^k\|_X^2 \le e^{\gamma_{\ell+1}} c(\varepsilon_{\ell+1}, q, f, u_0, z),$$

where $c(\varepsilon_{\ell+1}, q, f, u_0, z)$ is the sum of the first three terms on the right hand side of (4.19), with ε_1 replaced by $\varepsilon_{\ell+1}$. Here, the constant $\gamma_{\ell+1}$ is bounded by the induction hypothesis, that is,

$$\begin{aligned}\gamma_{\ell+1} &= \frac{c}{\varepsilon_\ell}\sum_{k=1}^{N_r-1} \tau\|\nabla u_h^{k+1-N_r}\|_X^2 + \frac{c}{\varepsilon_\ell}\sum_{k=N_r}^{\ell N_r-1} \tau\|\nabla u_h^{k+1-N_r}\|_X^2\\ &\le \frac{c}{\varepsilon_\ell}\left(\|z\|_{L^2(I_r,V)}^2 + \frac{2}{\nu}\mathfrak{c}_\ell(q, f, u_0, z)\right).\end{aligned}$$

Hence, (4.17) is also valid when ℓ is replaced by $\ell + 1$, completing the induction step. Therefore, the inequality (4.11) follows from (4.1) and (4.17) with $\ell = n_0 = T/r$ since $N_\tau = n_0 N_r$. ☐

Remark 4.6. Note that the function $\mathfrak{c}$ in the previous theorem is not uniform in terms of the delay parameter r. In fact, according to the proof, $\mathfrak{c}$ depends on the ratio T/r, so that in particular $\mathfrak{c} \to \infty$ as $r \to 0$.

We are now in position to prove error estimates for the solutions between the continuous and discrete state problems. Let us start with the errors in the norms of $L^\infty(I, X)$ and $L^2(I, W)$.

Theorem 4.7. *Let $q \in Q$, $u = S(q)$ and $u_\sigma = S_\sigma(q)$. Then there is a constant $c > 0$ depending continuously on $\|q\|_Q$, $\|f\|_{H^{2,1}(I)}$, $\|Au_0\|_H$ and $\|z\|_{H^{2,1}(J_r)}$ but independent on σ such that*

$$\max_{1\le k\le N_\tau} \|u(t_k) - u_h^k\|_X + \|u - u_\sigma\|_{L^\infty(J_r,X)} + \|u - u_\sigma\|_{L^2(J_r,W)} \le ch. \tag{4.20}$$

Proof. Since $u - u_\sigma = z - z_\sigma$ in Ω_r, it is enough to prove (4.20) with J_r replaced by I according to (4.6). Define $\bar{u}_\sigma \in L^2(J_r, V_h)$ such that $\bar{u}_h^k = P_h u(t_k)$ for each $k = 1 - N_r, \ldots, N_\tau$ and split the error into $u - u_\sigma = (u - \bar{u}_\sigma) + (\bar{u}_\sigma - u_\sigma)$. From Proposition 4.2, it holds that

$$\|u - \bar{u}_\sigma\|_{L^\infty(I,X)} + \|u - \bar{u}_\sigma\|_{L^2(I,W)} \le ch. \tag{4.21}$$

In particular, this estimate implies

$$\max_{1\le k\le N_\tau} \|u(t_k) - \bar{u}_h^k\|_X \le \max_{1\le k\le N_\tau} \|u - \bar{u}_\sigma\|_{L^\infty(I_k,X)} \le ch. \tag{4.22}$$

Also, by (4.6), Propositions 4.1 and 4.3

$$\|z_\sigma - \bar{u}_\sigma\|_{L^2(I_r,W)} \le \|z_\sigma - z\|_{L^2(I_r,W)} + \|u - \bar{u}_\sigma\|_{L^2(I_r,W)} \le ch. \tag{4.23}$$

The next step is to derive analogous estimates for the error term $e_\sigma := \bar{u}_\sigma - u_\sigma$. Taking a test function $\varphi_h \in V_h \subset W$ in the Eq. (2.12) and then integrating over I_k, we obtain that

$$(P_h u(t_k) - P_h u(t_{k-1}), \varphi_h)_X + \int_{I_k} \nu(\nabla u(t), \nabla\varphi_h)_X \,\mathrm{d}t - \int_{I_k} (\operatorname{div}\varphi_h, p(t))_{L^2} \,\mathrm{d}t$$
$$= \int_{I_k} b(u^r(t), \varphi_h, u(t)) \,\mathrm{d}t + \int_{I_k} (f(t) + q(t), \varphi_h)_X \,\mathrm{d}t$$

for every $k = 1, \dots, N_\tau$. Let us rewrite the above variational equation as follows:

$$(d_\tau \bar{u}_h^k, \varphi_h)_X + \nu(\nabla\bar{u}_h^k, \nabla\varphi_h)_X = b(\bar{u}_h^{k-N_r}, \varphi_h, \bar{u}_h^{k-1})$$
$$+ \frac{1}{\tau}\int_{I_k} (f_\sigma(t) + q(t), \varphi_h) \,\mathrm{d}t + \frac{1}{\tau} R_h^k(\varphi_h) \qquad \forall\varphi_h \in V_h, \tag{4.24}$$

where $d_\tau \bar{u}_h^1 = \tau^{-1}(\bar{u}_h^1 - P_h u_0)$, $d_\tau \bar{u}_h^k = \tau^{-1}(\bar{u}_h^k - \bar{u}_h^{k-1})$ for $k = 2, \dots, N_\tau$ and the remainder term $R_h^k(\varphi_h)$ is given by

$$R_h^k(\varphi_h) = \int_{I_k} \nu(\nabla\bar{u}_h^k - \nabla u(t), \nabla\varphi_h)_X \,\mathrm{d}t + \int_{I_k} (\operatorname{div}\varphi_h, p(t) - \Gamma_h p(t))_{L^2} \,\mathrm{d}t$$
$$+ \int_{I_k} b(u^r(t), \varphi_h, u(t) - \bar{u}_h^{k-1}) \,\mathrm{d}t + \int_{I_k} b(u^r(t) - \bar{u}_h^{k-N_r}, \varphi_h, \bar{u}_h^{k-1}) \,\mathrm{d}t$$
$$+ \int_{I_k} (f(t) - f_\sigma(t), \varphi_h)_X \,\mathrm{d}t.$$

Here, we used the fact that $(\operatorname{div}\varphi_h, \Gamma_h p)_{L^2(I_k, L^2(\Omega))} = 0$ since $\varphi_h \in V_h$.

Let $\widehat{e}_h^0 := P_h u_0 - u_{0h}$ and $\widehat{e}_h^k := e_h^k = \bar{u}_h^k - u_h^k$ for $k \neq 0$. Subtracting (4.3) from (4.24) and then choosing the test function $\varphi_k = 2\tau e_h^k$, we get the following recurrence relation for the errors:

$$\|e_h^k\|_X^2 - \|\widehat{e}_h^{k-1}\|_X^2 + \|e_h^k - \widehat{e}_h^{k-1}\|_X^2 + 2\nu\tau\|\nabla e_h^k\|_X^2$$
$$= 2\tau b(e_h^{k-N_r}, e_h^k, \bar{u}_h^{k-1}) + 2\tau b(u_h^{k-N_r}, e_h^k, e_h^{k-1}) + 2R_h^k(e_h^k) \tag{4.25}$$

for each $k = 1, \dots, N_\tau$. Using the Gagliardo–Nirenberg and Young inequalities, we can estimate the trilinear terms in (4.25) as follows:

$$2\tau b(e_h^{k-N_r}, e_h^k, \bar{u}_h^{k-1}) \le \frac{\nu\tau}{3}\|\nabla e_h^k\|_X^2 + c\tau\|\nabla e_h^{k-N_r}\|_X^2 \|\nabla\bar{u}_h^{k-1}\|_X^2 \tag{4.26}$$

$$2\tau b(u_h^{k-N_r}, e_h^k, e_h^{k-1}) \le \frac{\nu\tau}{3}\|\nabla e_h^k\|_X^2 + \frac{c\tau}{\varepsilon}\|\nabla u_h^{k-N_r}\|_X^2 \|e_h^{k-1}\|_X^2 + \varepsilon\tau\|u_h^{k-N_r}\|_X^2 \|\nabla e_h^{k-1}\|_X^2. \tag{4.27}$$

From Theorem 4.5 and (4.6), there is a constant $c_0 > 0$ independent on σ such that $\|u_h^{k-N_r}\|_X \le c_0$ for every $k = 1, \dots, N_\tau$. Also, note that $\|\nabla\bar{u}_h^k\|_X = \|\nabla P_h u(t_k)\|_X \le c\|u\|_{L^\infty(J_r, V)}$ for every $k = 1 - N_r, \dots, N_\tau$. By the Cauchy–Schwarz inequality, the remainder term $2R_h^k(e_h^k)$ satisfies

$$2R_h^k(e_h^k) \le \frac{\nu\tau}{3}\|\nabla e_h^k\|_X^2 + c\int_{I_k} \|\nabla P_h u(t_k) - \nabla u(t)\|_X^2 \,\mathrm{d}t + c\int_{I_k} \|p(t) - \Gamma_h p(t)\|_M^2 \,\mathrm{d}t$$
$$+ c\|u^r\|_{L^\infty(I_k, V)}^2 \int_{I_k} \|\nabla u(t) - \nabla P_h u(t_{k-1})\|_X^2 \,\mathrm{d}t + c\int_{I_k} \|f_\sigma(t) - f(t)\|_X^2 \,\mathrm{d}t$$
$$+ c\|u\|_{L^\infty(I_k, V)}^2 \int_{I_k} \|\nabla u^r(t) - \nabla P_h u^r(t_k)\|_X^2 \,\mathrm{d}t =: \frac{\nu\tau}{3}\|\nabla e_h^k\|_X^2 + \int_{I_k} \rho_h^k(t) \,\mathrm{d}t. \tag{4.28}$$

Taking into account (4.26)-(4.28) in (4.25) and choosing $\varepsilon = \nu/(2c_0^2)$, we have

$$\|e_h^k\|_X^2 - \|\widehat{e}_h^{k-1}\|_X^2 + \nu\tau\|\nabla e_h^k\|_X^2 \le c\tau\|\nabla e_h^{k-N_r}\|_X^2 \|u\|_{L^\infty(J_r, V)}^2$$
$$+ \frac{2cc_0^2}{\nu}\tau\|\nabla u_h^{k-N_r}\|_X^2 \|e_h^{k-1}\|_X^2 + \frac{\nu\tau}{2}\|\nabla e_h^{k-1}\|_X^2 + \int_{I_k} \rho_h^k(t) \,\mathrm{d}t.$$

Getting the sum of the previous inequality over all $k=1,\dots,j$ with $1\le j\le N_\tau$, we deduce the existence of a constant $c>0$ independent on σ such that

$$\|e_h^j\|_X^2+\frac{\nu}{2}\sum_{k=1}^{j}\tau\|\nabla e_h^k\|_X^2\le\|\widehat{e}_h^0\|_X^2+c\tau\|\nabla z_h^{1-N_r}\|_X^2\|e_h^0\|_X^2$$
$$+c\sum_{k=1}^{j-1}\tau\|\nabla u_h^{k+1-N_r}\|_X^2\|e_h^k\|_X^2+c\sum_{k=1}^{j}\tau\|\nabla e_h^{k-N_r}\|_X^2+c\sum_{k=1}^{j}\int_{I_k}\rho_h^k(t)\,\mathrm{d}t. \tag{4.29}$$

For the first two terms on the right hand side of (4.29), we have $\|\widehat{e}_h^0\|_X^2=\|\varPi_h u_0-P_h u_0\|_X^2\le ch^4$ by (4.5), and $\tau\|\nabla z_h^{1-N_r}\|_X^2\|e_h^0\|_X^2\le c\|z\|_{L^\infty(I_r,V)}^2\|z_\sigma-\bar{u}_\sigma\|_{L^2(I_r,X)}^2\le ch^2$ by (4.23). To treat the fifth term, we invoke (A2), (4.7), Proposition 4.1 and Proposition 4.3, so that for each $j=1,\dots,N_\tau$

$$\sum_{k=1}^{j}\int_{I_k}\rho_h^k(t)\,\mathrm{d}t\le\int_I\rho_h^k(t)\,\mathrm{d}t\le ch^2.$$

Utilizing these estimates in (4.29) and applying the discrete Gronwall Lemma 7.2, we get that

$$\max_{1\le k\le \ell N_r}\|e_h^k\|_X^2+\sum_{k=1}^{\ell N_r}\nu\tau\|\nabla e_h^k\|_X^2\le ce^{\gamma_\ell}\left(h^2+\sum_{k=1}^{\ell N_r}\tau\|\nabla e_h^{k-N_r}\|_X^2\right) \tag{4.30}$$

for every $\ell=1,\dots,n_0$, where the constant γ_ℓ is given by

$$\gamma_\ell=c\sum_{k=1}^{\ell N_r-1}\tau\|\nabla u_h^{k+1-N_r}\|_X^2=c\sum_{k=1}^{N_r-1}\tau\|\nabla z_h^{k+1-N_r}\|_X^2+c\sum_{k=N_r}^{\ell N_r-1}\tau\|\nabla u_h^{k+1-N_r}\|_X^2\le c$$

due to (4.6) and Theorem 4.5.

For $\ell=1$ in the sum on the right hand side of (4.30), one obtains from (4.23) that

$$\sum_{k=1}^{N_r}\tau\|\nabla e_h^{k-N_r}\|_X^2=\|\nabla z_\sigma-\nabla\bar{u}_\sigma\|_{L^2(I_r,X)}^2\le ch^2.$$

Using this estimate along with an induction argument on the inequality (4.30), we have

$$\max_{1\le k\le N_\tau}\|\bar{u}_h^k-u_h^k\|_X+\|e_\sigma\|_{L^\infty(I,X)}+\|e_\sigma\|_{L^2(I,W)}\le ch. \tag{4.31}$$

Combining the error estimates (4.21), (4.22) and (4.31) leads to (4.20) with I in place of J_r. □

Let us now prove an error estimate in terms of the norm in $L^2(I,X)$. First, let us state the following lemma for the error estimate of the temporal shift by τ. Since the proof follows the ideas in Proposition 4.1, the details are omitted.

Lemma 4.8. *There exists a constant $c>0$ independent on τ such that for every $u\in H^{2,1}(J_r)$ and $w\in H^{2,1}(J^r)$ we have*

$$\|u-u^\tau\|_{L^2(I,X)}+\sqrt{\tau}\|u-u^\tau\|_{L^\infty(I,X)}\le c\tau\|u\|_{H^{2,1}(J_r)}$$
$$\|w-w^{-\tau}\|_{L^2(I,X)}+\sqrt{\tau}\|w-w^{-\tau}\|_{L^\infty(I,X)}\le c\tau\|w\|_{H^{2,1}(J^r)}.$$

Theorem 4.9. *Let $q\in Q$, $u=S(q)$ and $u_\sigma=S_\sigma(q)$. There exists a constant $c>0$ depending continuously on $\|q\|_Q$, $\|f\|_{H^{2,1}(I)}$, $\|Au_0\|_H$ and $\|z\|_{H^{2,1}(J_r)}$ but independent of σ such that*

$$\|u-u_\sigma\|_{L^2(I,X)}\le ch^2. \tag{4.32}$$

Proof. We shall proceed by an Aubin–Nitsche-type duality argument. We recall the reader of our stability condition (A5) that will be frequently used in the proof. Given $g \in L^2(I,X)$ such that $\|g\|_{L^2(I,X)} \le 1$, Theorem 2.14 implies the existence of a unique weak solution $(w,\pi) \in H^{2,1}(J^r) \times L^2(I,Y)$ to the following dual problem:

$$\begin{cases} -\partial_t w - \nu\Delta w - (u^r\cdot\nabla)w - (\nabla w^{-r})^\top u^{-r} + \nabla\pi = g \text{ in } \Omega_T, \\ \operatorname{div} w = 0 \ \text{ in } \Omega_T, \quad w = 0 \ \text{ in } \Gamma_T, \quad w(T) = 0 \ \text{ in } \Omega, \quad w = 0 \ \text{ in } \Omega_{T+r}. \end{cases} \tag{4.33}$$

Furthermore, there is a constant $c > 0$ independent on g, w and π such that

$$\|w(0)\|_H + \|w\|_{H^{2,1}(J^r)} + \|\pi\|_{L^2(I,Y)} \le c. \tag{4.34}$$

Let $w_\sigma \in \mathcal{P}_\tau(J^r, V_h)$ be given by $w_h^k = P_h w(t_{k-1})$ for each $k = 1,\dots,N_\tau + N_r$. Applying (A2) and Proposition 4.1, one can see that

$$\|w - w_\sigma\|_{L^2(I,X)} + h\|w - w_\sigma\|_{L^2(I,W)} + h\|\pi - \mathit{\Gamma}_h\pi\|_{L^2(I,M)} \le ch^2. \tag{4.35}$$

Moreover, we have $\|\nabla w_\sigma\|_{L^\infty(I,X)} \le c\|\nabla w\|_{L^\infty(I,X)} \le c\|w\|_{H^{2,1}(I)} \le c$ by the continuity of the embedding $H^{2,1}(I) \subset C(\bar{I}, V)$ and boundedness of $P_h : L^\infty(I,W) \to L^\infty(I,W)$.

Denote the error by $e_\sigma := u - u_\sigma \in L^2(J_r, W_h) \subset L^2(I,W)$. Multiplying (4.33) with the test function e_σ, integrating over Ω_T and applying Green's identity yield

$$\begin{aligned} \int_I (g(t), e_\sigma(t))_X \,\mathrm{d}t = &- \int_I (e_\sigma(t), \partial_t w(t))_X \,\mathrm{d}t + \int_I \nu(\nabla e_\sigma(t), \nabla w(t))_X \,\mathrm{d}t \\ &- \int_I b(u^r(t), w(t), e_\sigma(t))\,\mathrm{d}t - \int_I b(e_\sigma^r(t), w(t), u(t))\,\mathrm{d}t \\ &+ \int_{[0,r]} b(e_\sigma^r(t), w(t), u(t))\,\mathrm{d}t + \int_I (\operatorname{div} e_\sigma(t), \pi(t) - \mathit{\Gamma}_h\pi(t))_{L^2}\,\mathrm{d}t \\ := &-J_1 + J_2 - J_3 - J_4 + J_5 + J_6. \end{aligned} \tag{4.36}$$

In the last term we used the fact that $(\operatorname{div} e_\sigma(t), \mathit{\Gamma}_h\pi(t))_{L^2} = -(\operatorname{div} u_\sigma(t), \mathit{\Gamma}_h\pi(t))_{L^2} = 0$ for almost every $t \in I$ since $u \in L^2(I,V)$, $u_\sigma \in L^2(I,V_h)$ and $\mathit{\Gamma}_h\pi \in L^2(I,M_h)$. The last two terms in (4.36) can be immediately estimated from above by

$$|J_6| \le c\|\nabla e_\sigma\|_{L^2(I,X)}\|\pi - \mathit{\Gamma}_h\pi\|_{L^2(I,M)} \le ch^2 \tag{4.37}$$

$$|J_5| \le c\|z - z_\sigma\|_{L^2(I_r,X)}\|\Delta w\|_{L^2(I,H)}\|\nabla u\|_{L^\infty(I,X)} \le ch^2 \tag{4.38}$$

thanks to (4.6), (4.20) and (4.35).

Let us consider the first integral in (4.36). Integrating by parts with respect to time

$$\begin{aligned} -J_1 &= -\int_I (u(t), \partial_t w(t))_X \,\mathrm{d}t + \sum_{k=1}^{N_\tau} (u_h^k, w_h^{k+1} - w_h^k)_X \\ &= (u_0 - u_{0h}, P_h w(0))_X + \int_I (\partial_t u(t), w(t))_X \,\mathrm{d}t - \sum_{k=1}^{N_\tau} (u_h^k - \widehat{u}_h^{k-1}, w_h^k)_X =: J_7 + J_8, \end{aligned} \tag{4.39}$$

where $\widehat{u}_h^{k-1}$ is defined as in the proof of Theorem 4.5. According to (4.5) and (4.34), it holds that $|J_7| = |(u_0 - u_{0h}, P_h w(0))_X| \le ch^2$. On the other hand, from the continuous and discrete state equations satisfied by u and u_σ, respectively, we obtain that

$$\begin{aligned} J_8 = &\int_I (f(t) - f_\sigma(t), w(t))_X \,\mathrm{d}t + \int_I (f_\sigma(t) + q(t), w(t) - w_\sigma(t))_X \,\mathrm{d}t \; - \int_I \nu(\nabla u(t), \nabla w(t))_X \,\mathrm{d}t \\ &+ \int_I b(u^r(t), w(t), u(t))\,\mathrm{d}t + \int_I \nu(\nabla u_\sigma(t), \nabla w_\sigma(t))_X \,\mathrm{d}t - \int_I b(u_\sigma^r(t), w_\sigma(t), u_\sigma^\tau(t))\,\mathrm{d}t \\ =: &\; J_9 + J_{10} - J_{11} + J_{12} + J_{13} - J_{14}. \end{aligned} \tag{4.40}$$

Here, we used the fact that $u_\sigma^\tau|_{I_k} = u_\sigma^\tau(t_k) = u_\sigma(t_{k-1}) = u_h^{k-1}$, and similarly $u_\sigma^r|_{I_k} = u_h^{k-N_r}$. From (4.7), (4.34), (4.35), we have $|J_9| + |J_{10}| \le ch^2$. Collecting what we have obtained in (4.36)-(4.40) gives us

$$\int_I (g(t), e_\sigma(t))_X \,\mathrm{d}t \le ch^2 + (J_2 - J_{11} + J_{13}) + (J_{12} - J_{14} - J_3 - J_4). \tag{4.41}$$

The remaining task is to derive estimates on the terms inside the parentheses. The first group in (4.41) can be easily done with the aid of Green's identity, (4.20) and (4.35) so that

$$\begin{aligned} J_2 - J_{11} + J_{13} &= \int_I \nu(\nabla e_\sigma(t), \nabla w(t) - \nabla w_\sigma(t))_X \,\mathrm{d}t - \int_I \nu(\Delta u(t), w(t) - w_\sigma(t))_X \,\mathrm{d}t \\ &\le c\|\nabla e_\sigma\|_{L^2(I,X)}\|\nabla w - \nabla w_\sigma\|_{L^2(I,X)} + c\|\Delta u\|_{L^2(I,X)}\|w - w_\sigma\|_{L^2(I,X)} \le ch^2. \end{aligned} \tag{4.42}$$

Due to the explicit discretization of the convection term, we need to shift by τ the third argument involving the trilinear terms in J_{12}, J_3 and J_4. This is to match the form of the term J_{14}. In this direction, we split J_{12} as follows

$$\begin{aligned} J_{12} &= \int_I b(u^r(t), w(t), u(t) - u^\tau(t)) \,\mathrm{d}t \\ &\quad + \int_I b(u^r(t), w(t) - w_\sigma(t), u^\tau(t)) \,\mathrm{d}t + \int_I b(u^r(t), w_\sigma(t), u^\tau(t)) \,\mathrm{d}t. \end{aligned} \tag{4.43}$$

The first two terms on the right hand side satisfy

$$\int_I |b(u^r(t), w(t), u(t) - u^\tau(t))| \,\mathrm{d}t \le c\|\nabla u^r\|_{L^\infty(I,X)}\|\Delta w\|_{L^2(I,H)}\|u - u^\tau\|_{L^2(I,X)}$$
$$\int_I |b(u^r(t), w(t) - w_\sigma(t), u^\tau(t))| \,\mathrm{d}t \le c\|\nabla u^r\|_{L^\infty(I,X)}\|\Delta u^\tau\|_{L^2(I,H)}\|w - w_\sigma\|_{L^2(I,X)}.$$

In the second inequality, we used the anti-symmetry of b with respect to the second and third arguments. Using these inequalities in (4.43) and applying (4.35), Lemma 4.8 and (A5), we have

$$J_{12} \le ch^2 + \int_I b(u^r(t), w_\sigma(t), u^\tau(t)) \,\mathrm{d}t. \tag{4.44}$$

Next, the trilinear term J_3 can be equivalently written as

$$\begin{aligned} J_3 &= \int_I b(u^r(t) - u^{r-\tau}(t), w(t), e_\sigma(t)) \,\mathrm{d}t + \int_I b(u^{r-\tau}(t), w(t) - w^{-\tau}(t), e_\sigma(t)) \,\mathrm{d}t \\ &\quad - \int_{I_1} b(u^r(t)s, w(t), e_\sigma^\tau(t)) \,\mathrm{d}t + \int_I b(u^r(t), w(t), e_\sigma^\tau(t)) \,\mathrm{d}t. \end{aligned} \tag{4.45}$$

Employing the Hölder inequality and the properties of the trilinear form b to the first three terms on the right hand side of (4.45), we deduce that

$$\int_{I_1} |b(u^r(t), w(t), e_\sigma^\tau(t))| \,\mathrm{d}t \le c\sqrt{\tau}\|\nabla u^r\|_{L^\infty(I_1,X)}\|\Delta w\|_{L^2(I_1,X)}\|e_\sigma^\tau\|_{L^\infty(I_1,X)}$$
$$\int_I |b(u^r(t) - u^{r-\tau}(t), w(t), e_\sigma(t))| \,\mathrm{d}t \le c\|u^r - u^{r-\tau}\|_{L^\infty(I,X)}\|\Delta w\|_{L^2(I,X)}\|\nabla e_\sigma\|_{L^2(I,X)}$$
$$\int_I |b(u^{r-\tau}(t), w(t) - w^{-\tau}(t), e_\sigma(t))| \,\mathrm{d}t \le c\|\Delta u^{r-\tau}\|_{L^2(I,X)}\|w - w^{-\tau}\|_{L^\infty(I,X)}\|\nabla e_\sigma\|_{L^2(I,X)}.$$

Again, we utilized in the third inequality the anti-symmetry of b. Plugging these estimates in (4.45) and then applying Lemma 4.8, (4.20), (4.35) and (A5), we get

$$J_3 \ge -ch^2 + \int_I b(u^r(t), w(t), e_\sigma^\tau(t)) \,\mathrm{d}t. \tag{4.46}$$

Finally, for the term J_4 we use Lemma 4.8 and (4.20) once again to deduce that

$$\begin{aligned} J_4 &= \int_I b(e_\sigma^r(t), w(t), u(t) - u^\tau(t))\,\mathrm{d}t + \int_I b(e_\sigma^r(t), w(t), u^\tau(t))\,\mathrm{d}t \\ &\geq -c\|\nabla e_\sigma^r\|_{L^2(I,X)}\|\Delta w\|_{L^2(I,H)}\|u - u^\tau\|_{L^2(I,X)} + \int_I b(e_\sigma^r(t), w(t), u^\tau(t))\,\mathrm{d}t \\ &\geq -ch^2 + \int_I b(e_\sigma^r(t), w(t), u^\tau(t))\,\mathrm{d}t. \end{aligned} \tag{4.47}$$

Hence, if one applies the estimates (4.44), (4.46) and (4.47), then after some rearrangement of the trilinear terms the following estimate holds:

$$\begin{aligned} J_{12} - J_{14} - J_3 - J_4 &\leq ch^2 + \int_I b(e_\sigma^r(t), w_\sigma(t) - w(t), u^\tau(t))\,\mathrm{d}t \\ &\quad + \int_I b(u^r(t), w_\sigma(t) - w(t), e_\sigma^\tau(t))\,\mathrm{d}t + \int_I b(e_\sigma^r(t), w_\sigma(t), e_\sigma^\tau(t))\,\mathrm{d}t. \end{aligned} \tag{4.48}$$

In virtue of the Hölder inequality, the terms on the right hand side of (4.48) satisfy

$$\begin{aligned} \int_I b(e_\sigma^r(t), w_\sigma(t), e_\sigma^\tau(t))\,\mathrm{d}t &\leq c\|\nabla e_\sigma^r\|_{L^2(I,X)}\|\nabla w_\sigma\|_{L^\infty(I,X)}\|\nabla e_\sigma^\tau\|_{L^2(I,X)} \\ \int_I b(u^r(t), w_\sigma(t) - w(t), e_\sigma^\tau(t))\,\mathrm{d}t &\leq c\|\nabla u^r\|_{L^\infty(I,X)}\|\nabla w_\sigma - \nabla w\|_{L^2(I,X)}\|\nabla e_\sigma^\tau\|_{L^2(I,X)} \\ \int_I b(e_\sigma^r(t), w_\sigma(t) - w(t), u^\tau(t))\,\mathrm{d}t &\leq c\|\nabla e_\sigma^r\|_{L^2(I,X)}\|\nabla w_\sigma - \nabla w\|_{L^2(I,X)}\|\nabla u^\tau\|_{L^\infty(I,X)}. \end{aligned}$$

Substituting these in (4.48) and then recalling the previous estimates (4.20) and (4.35), we arrive at $J_{12} - J_{14} - J_3 - J_4 \leq ch^2$. This inequality, together with those in (4.41) and (4.42), implies $(e_\sigma, g)_{L^2(I,X)} \leq ch^2$ for every $g \in L^2(I,X)$ with $\|g\|_{L^2(I,X)} \leq 1$. Therefore, (4.32) holds true by duality. The proof of the theorem is now complete. □

4.4. Error Estimates for the Discrete Linearized State Equation

In this subsection, we consider the discretization of the linearized state equation. In fact, the resulting scheme will be obtained by taking the derivative of the discrete solution operator. It is easy to see that the map S_σ belongs to $C^\infty(Q, \mathcal{P}_\tau(I, V_h))$. Moreover, given a direction $g \in Q$, we have $v_\sigma = S'_\sigma(q)g \in \mathcal{P}_\tau(I, V_h)$ if and only if $v_\sigma = \sum_{k=1}^{N_\tau} v_h^k \mathbb{1}_{I_k}$ is the solution of

$$\begin{cases} (\bar{d}_\tau v_h^k, \varphi_h)_X + \nu(\nabla v_h^k, \nabla \varphi_h)_X = b(u_h^{k-N_r}, \varphi_h, v_h^{k-1}) \\ \qquad + b(v_h^{k-N_r}, \varphi_h, u_h^{k-1}) + \dfrac{1}{\tau}\displaystyle\int_{I_k} (g(t), \varphi_h)_X\,\mathrm{d}t \quad \forall \varphi_h \in V_h, \\ v_h^j = 0 \quad \text{for } j = 1 - N_r, \ldots, 0. \end{cases} \tag{4.49}$$

for every $k = 1, \ldots, N_\tau$, where $u_\sigma = S_\sigma(q)$ and $\bar{d}_\tau v_h^k := \tau^{-1}(v_h^k - v_h^{k-1})$. Similarly, if $g \in Q$ then $y_\sigma = S''_\sigma(q)[g, g] \in \mathcal{P}_\tau(I, V_h)$ if and only if for each $k = 1, \ldots, N_\tau$

$$\begin{cases} (\bar{d}_\tau y_h^k, \varphi_h)_X + \nu(\nabla y_h^k, \nabla \varphi_h)_X = b(u_h^{k-N_r}, \varphi_h, y_h^{k-1}) \\ \qquad + b(y_h^{k-N_r}, \varphi_h, u_h^{k-1}) + 2b(v_h^{k-N_r}, \varphi_h, v_h^{k-1}) \quad \forall \varphi_h \in V_h, \\ y_h^j = 0 \quad \text{for } j = 1 - N_r, \ldots, 0, \end{cases} \tag{4.50}$$

where $v_\sigma = S'_\sigma(q)g$ and $u_\sigma = S_\sigma(q)$. In the succeeding discussions, we deal with the error estimates for the solutions of (2.20) and (4.49).

Theorem 4.10. *Let $q \in Q$, $g \in Q$, $u_\sigma = S_\sigma(q)$ and $v_\sigma = S'_\sigma(q)g$. Then there exists a continuous function $\mathfrak{c} > 0$ such that for every σ, we have*

$$\|v_\sigma\|_{L^\infty(I,X)} + \|v_\sigma\|_{L^2(I,W)} \leq \mathfrak{c}(\|u_\sigma\|_{L^\infty(J_r,X)\cap L^2(J_r,W)})\|g\|_Q. \tag{4.51}$$

Proof. Taking the test function $\varphi_h = 2\tau v_h^k$ in (4.49) and using the Young inequality, we obtain for $k = 1, \ldots, N_\tau$ that

$$\begin{aligned}
&\|v_h^k\|_X^2 - \|v_h^{k-1}\|_X^2 + \|v_h^k - v_h^{k-1}\|_X^2 + 2\nu\tau\|\nabla v_h^k\|_X^2 \\
&\quad \leq 2\tau b(u_h^{k-N_r}, v_h^k, v_h^{k-1}) + 2\tau b(v_h^{k-N_r}, v_h^k, u_h^{k-1}) + \int_{I_k} (g, v_h^k)_X \, \mathrm{d}t \\
&\quad \leq \nu\tau\|\nabla v_h^k\|_X^2 + \varepsilon\tau\|u_h^{k-N_r}\|_X^2\|\nabla v_h^{k-1}\|_X^2 + \frac{c\tau}{\varepsilon}\|\nabla u_h^{k-N_r}\|_X^2\|v_h^{k-1}\|_X^2 \\
&\qquad + \frac{c\tau}{\varepsilon}\|v_h^{k-N_r}\|_X^2\|\nabla u_h^{k-1}\|_X^2 + \varepsilon\tau\|\nabla v_h^{k-N_r}\|_X^2\|u_h^{k-1}\|_X^2 + \frac{c}{\nu}\int_{I_k} \|g(t)\|_X^2 \, \mathrm{d}t.
\end{aligned}$$

Let us choose $0 < \varepsilon < \nu(4\|u_\sigma\|_{L^\infty(J_r,X)} + 1)^{-1}$. Thus, $\varepsilon\|u_h^{k-N_r}\|_X^2 \leq \nu/4$ and $\varepsilon\|u_h^{k-1}\|_X^2 \leq \nu/4$ for every $k = 1, \ldots, N_\tau$. Given $1 \leq j \leq N_\tau$, take the sum of the above inequality over all $k = 1, \ldots, j$ and use $v_h^0 = 0$ so that

$$\begin{aligned}
\|v_h^j\|_X^2 + \sum_{k=1}^{j} \frac{\nu\tau}{2}\|\nabla v_h^k\|_X^2 &\leq \frac{c}{\nu}\sum_{k=1}^{j}\int_{I_k} \|g(t)\|_X^2 \, \mathrm{d}t \\
&\quad + \frac{c}{\varepsilon}\sum_{k=1}^{j} \tau\|\nabla u_h^{k-N_r}\|_X^2\|v_h^{k-1}\|_X^2 + \frac{c}{\varepsilon}\sum_{k=1}^{j} \tau\|\nabla u_h^{k-1}\|_X^2\|v_h^{k-N_r}\|_X^2.
\end{aligned}$$

By the discrete Gronwall Lemma 7.2, recalling that $v_h^k = 0$ for every $k = 1 - N_r, \ldots, 0$, we have

$$\max_{1\leq k\leq N_\tau} \|v_h^k\|_X^2 + \sum_{k=1}^{N_\tau} \frac{\nu\tau}{2}\|\nabla v_h^k\|_X^2 \leq ce^\gamma\|g\|_Q^2, \tag{4.52}$$

where the constant γ is given by

$$\gamma = \frac{c}{\varepsilon}\sum_{k=1}^{N_\tau} \tau\|\nabla u_h^{k-N_r}\|_X^2 + \frac{c}{\varepsilon}\sum_{k=1}^{N_\tau} \tau\|\nabla u_h^{k-1}\|_X^2 \leq c\|u_\sigma\|_{L^2(J_r,W)}^2. \tag{4.53}$$

The inequalities (4.52) and (4.53) imply (4.51). □

Theorem 4.11. *Let $q \in Q$, $g \in Q$, $v = S'(q)g$ and $v_\sigma = S'_\sigma(q)g$. Then there is a constant $c > 0$ depending continuously on $\|q\|_Q$, $\|g\|_Q$, $\|f\|_{H^{2,1}(I)}$, $\|Au_0\|_H$ and $\|z\|_{H^{2,1}(J_r)}$ so that for every σ*

$$\max_{1\leq k\leq N_\tau} \|v(t_k) - v_h^k\|_X + \|v - v_\sigma\|_{L^\infty(J_r,X)} + \|v - v_\sigma\|_{L^\infty(J_r,W)} \leq ch. \tag{4.54}$$

Furthermore, we also have

$$\|v - v_\sigma\|_{L^2(I,X)} \leq ch^2. \tag{4.55}$$

Proof. Since both v and v_σ vanish on I_r, we only need to prove (4.54) with J_r replaced by I. Let $u = S(q)$, $u_\sigma = S_\sigma(q)$ and $\bar{u}_\sigma \in \mathcal{P}_\tau(J_r, W_h)$ with $\bar{u}_h^k = P_h u(t_k)$ for each $k = 1 - N_r, \ldots, N_\tau$. Also, define $\bar{v}_\sigma \in L^2(J_r, V_h)$ by $\bar{v}_h^k = P_h v(t_k)$ for $k = 1 - N_r, \ldots, N_\tau$. Split the error according to $v - v_\sigma = (v - \bar{v}_\sigma) + (\bar{v}_\sigma - v_\sigma)$. By construction of $\bar{v}_\sigma$, we immediately have

$$\max_{1\leq k\leq N_\tau} \|v(t_k) - \bar{v}_h^k\|_X + \|v - \bar{v}_\sigma\|_{L^\infty(I,X)} + \|v - \bar{v}_\sigma\|_{L^2(I,W)} \leq ch. \tag{4.56}$$

Following the proof of Theorem 4.7, now with the linearized state problem (2.20) and its discrete version (4.49), one obtains that the error term $\eta_\sigma := \bar{v}_\sigma - v_\sigma \in \mathcal{P}_\tau(I, V_h)$ satisfies the variational equation

$$
\begin{aligned}
(\bar{d}_\tau \eta_h^k, \varphi_h)_X + \nu(\nabla \eta_h^k, \nabla\varphi_h)_X &= b(e_h^{k-N_r}, \varphi_h, \eta_h^{k-1}) + b(\eta_h^{k-N_r}, \varphi_h, e_h^{k-1}) \\
&\quad + b(e_h^{k-N_r}, \varphi_h, v_h^{k-1}) + b(u_h^{k-N_r}, \varphi_h, \eta_h^{k-1}) + \frac{1}{\tau} R_h^k(\varphi_h) \qquad \forall \varphi_h \in V_h,
\end{aligned} \tag{4.57}
$$

where $e_\sigma = \bar{u}_\sigma - u_\sigma$ and the remainder term $R_h^k(\varphi_h)$ is now given by

$$
\begin{aligned}
R_h^k(\varphi_h) &= \int_{I_k} \nu(\nabla P_h v(t_k) - \nabla v(t), \nabla\varphi_h)_X \,\mathrm{d}t + \int_{I_k} d(\varphi_h, \varpi(t) - \mathit{\Gamma}_h \varpi(t)) \,\mathrm{d}t \\
&\quad + \int_{I_k} b(v^r(t), \varphi_h, u(t) - P_h u(t_{k-1})) \,\mathrm{d}t + \int_{I_k} b(v^r(t) - P_h v^r(t_k), \varphi_h, P_h u(t_{k-1})) \,\mathrm{d}t \\
&\quad + \int_{I_k} b(u^r(t), \varphi_h, v(t) - P_h v(t_{k-1})) \,\mathrm{d}t + \int_{I_k} b(u^r(t) - P_h u^r(t_k), \varphi_h, P_h v(t_{k-1})) \,\mathrm{d}t.
\end{aligned} \tag{4.58}
$$

If we take $\varphi_h = 2\tau\eta_k$ in (4.57) and apply the Gagliardo–Nirenberg and Young inequalities, then

$$
\begin{aligned}
&\|\eta_h^k\|_X^2 - \|\eta_h^{k-1}\|_X^2 + 2\nu\tau \|\nabla \eta_h^k\|_X^2 \le \nu\tau\|\nabla\eta_h^k\|_X^2 + \varepsilon\tau \|e_h^{k-N_r}\|_X^2 \|\nabla \eta_h^{k-1}\|_X^2 \\
&\quad + \frac{c\tau}{\varepsilon}\|\nabla e_h^{k-N_r}\|_X^2\|\eta_h^{k-1}\|_X^2 + c\tau\|\eta_h^{k-N_r}\|_X^2\|\nabla e_h^{k-1}\|_X^2 + c\tau\|\nabla\eta_h^{k-N_r}\|_X^2\|e_h^{k-1}\|_X^2 \\
&\quad + c\tau\|e_h^{k-N_r}\|_X^2\|\nabla v_h^{k-1}\|_X^2 + c\tau\|\nabla v_h^{k-N_r}\|_X^2\|e_h^{k-1}\|_X^2 + \varepsilon\tau\|u_h^{k-N_r}\|_X^2\|\nabla\eta_h^{k-1}\|_X^2 \\
&\quad + \frac{c\tau}{\varepsilon}\|\nabla u_h^{k-N_r}\|_X^2\|\eta_h^{k-1}\|_X^2 + c|R_h^k(\eta_h^k)|.
\end{aligned}
$$

Choose $\varepsilon > 0$ such that $0 < \varepsilon < \nu(4\|u_\sigma\|_{L^\infty(J_r,X)} + 1)^{-1}$ and $0 < \varepsilon < \nu(4\|e_\sigma\|_{L^\infty(J_r,X)} + 1)^{-1}$. Since v_σ and η_σ are bounded in $L^\infty(J_r, X) \cap L^2(J_r, W)$ and $\|e_\sigma\|_{L^\infty(J_r,X)} + \|e_\sigma\|_{L^2(I,W)} \le ch$, the above inequality implies that, after taking the sum over all $k = 1, \dots, j$,

$$
\begin{aligned}
\|\eta_h^j\|_X^2 + \sum_{k=1}^{N_\tau} \frac{\nu\tau}{2}\|\nabla\eta_h^k\|_X^2 &\le ch^2 + c\sum_{k=1}^{j} |R_h^k(\eta_h^k)| \\
&\quad + \frac{c}{\varepsilon}\sum_{k=1}^{j} \tau(\|\nabla e_h^{k-N_r}\|_X^2 + \|\nabla u_h^{k-N_r}\|_X^2)\|\eta_h^{k-1}\|_X^2
\end{aligned} \tag{4.59}
$$

The term $R_h^k(\eta_h^k)$ can be treated in the same way as with the one given in proof of Theorem 4.7, and by doing this process we get

$$
c\sum_{k=1}^{j} |R_h^k(\eta_h^k)| \le ch^2 + \sum_{k=1}^{j} \frac{\nu\tau}{4}\|\nabla\eta_h^k\|_X^2, \qquad j = 1, \dots, N_\tau. \tag{4.60}
$$

Plugging (4.60) in (4.59) and then applying the discrete Gronwall Lemma

$$
\max_{1\le k\le N_\tau} \|\bar{v}_h^k - v_h^k\|_X + \|\eta_\sigma\|_{L^\infty(I,X)} + \|\eta_\sigma\|_{L^2(I,W)} \le ch. \tag{4.61}
$$

Taking the sum (4.56) and (4.61) proves the desired a priori estimate (4.54). Finally, (4.55) can be established by a duality argument as in Theorem 4.9 with the same adjoint problem. □

4.5. Error Estimates for the Discrete Adjoint Equation

Let $u_\sigma = S_\sigma(q)$ be the solution of the discrete state equation (4.3). The fully discretized adjoint problem that we consider is the following: Find $w_\sigma = \sum_{k=1}^{N_\tau} w_h^k \mathbb{1}_{I_k} \in \mathcal{P}_\tau(I, V_h)$ such that for $k = N_\tau, \ldots, 1$

$$\begin{cases} (d_{-\tau} w_h^k, \psi_h)_X + \nu(\nabla\psi_h, \nabla w_h^k)_X = b(u_h^{k+1-N_r}, w_h^{k+1}, \psi_h) + b(\psi_h, w_h^{k+N_r}, u_h^{k-1+N_r}) \\ \qquad = \alpha_{\Omega_T}(u_h^k - u_{dh}^k, \psi_h)_X + \alpha_R(\nabla\times u_h^k, \nabla\times\psi_h)_{L^2} \qquad \forall \psi_h \in V_h, \\ w_h^j = 0, \quad \text{for } j = N_\tau+1, \ldots, N_\tau+N_r. \end{cases} \tag{4.62}$$

where the forward difference operator $d_{-\tau}$ is defined by

$$d_{-\tau} w_h^k := \begin{cases} \tau^{-1}(w_h^{N_\tau} - \alpha_T(u_h^{N_\tau} - u_{Th})) & \text{if } k = N_\tau, \\ \tau^{-1}(w_h^k - w_h^{k+1}) & \text{if } k = N_\tau - 1, \ldots, 1. \end{cases}$$

The existence and uniqueness of solution to this problem follows immediately since the corresponding bilinear form is coercive for each k. Let us introduce the *discrete control-to-adjoint state operator* $D_\sigma : Q \to \mathcal{P}_\tau(I, V_h)$ by $D_\sigma(q) = w_\sigma$ if and only if w_σ is the solution of (4.62). The following lemma is the discrete versions of (3.1) and (3.2).

Lemma 4.12. *The action of the first and second derivatives of $j_\sigma : Q \to \mathbb{R}$ at $q \in Q$ in a direction $g \in Q$ are given by*

$$j_\sigma'(q)g = \int_I (w_\sigma + \alpha q, g)_X \, \mathrm{d}t$$

$$j_\sigma''(q)[g,g] = \int_I \alpha_{\Omega_T}\|v_\sigma\|_X^2 + \alpha_R\|\nabla\times v_\sigma\|_{L^2}^2 + 2b(v_\sigma^r, w_\sigma, v_\sigma^\tau)\,\mathrm{d}t + \alpha_T\|v_\sigma(T)\|_X^2 + \alpha\|g\|_Q^2$$

where $v_\sigma = S_\sigma'(q)g$ and $w_\sigma = D_\sigma(q)$.

Proof. From the chain rule and the fact that v_σ and w_σ are constants with respect to time on each subinterval I_k, it follows that

$$\begin{aligned} j_\sigma'(q)g = \tau \sum_{k=1}^{N_\tau} [\alpha_{\Omega_T}(u_h^k - u_{dh}^k, v_h^k)_X + \alpha_R(\nabla\times u_h^k, \nabla\times v_h^k)_{L^2}] \\ + \alpha_T(u_h^{N_\tau} - u_{Th}, v_h^{N_\tau})_X + \alpha(q,g)_Q. \end{aligned}$$

Let $I_\sigma := j_\sigma'(q)g - \alpha(q,g)_Q$. Taking the test function $\psi_h = \tau v_h^k$ in the discrete adjoint problem (4.62) and getting the sum over all $k = 1, \ldots, N_\tau$, we have

$$\begin{aligned} I_\sigma &= (w_h^{N_\tau}, v_h^{N_\tau})_X + \sum_{k=1}^{N_\tau-1}(w_h^k - w_h^{k+1}, v_h^k)_X + \sum_{k=1}^{N_\tau} \nu\tau(\nabla v_h^k, \nabla w_h^k)_X \\ &\quad - \sum_{k=1}^{N_\tau} \tau[b(u_h^{k+1-N_r}, w_h^{k+1}, v_h^k) + b(v_h^k, w_h^{k+N_r}, u_h^{k-1+N_r})] \\ &= \sum_{k=1}^{N_\tau}[(v_h^k - v_h^{k-1}, w_h^k)_X + \nu\tau(\nabla v_h^k, \nabla w_h^k)_X] - \sum_{k=1}^{N_\tau}\tau[b(u_h^{k-N_r}, w_h^k, v_h^{k-1}) + b(v_h^{k-N_r}, w_h^k, u_h^{k-1})] \\ &= \sum_{k=1}^{N_\tau}\int_{I_k}(g(t), w_h^k)_X\,\mathrm{d}t = \int_I (w_\sigma(t), g(t))_X\,\mathrm{d}t \end{aligned}$$

since $v_h^k = 0$ for $k = 1 - N_\tau, \ldots, 0$ and $w_h^k = 0$ for $k = N_\tau+1, \ldots, N_\tau+N_r$. This proves the case of the first directional derivative. The second directional derivative can be handled in a similar way using (4.50). □

Remark 4.13. If $q_\sigma^\star$ is a solution to (P_σ), then $q_\sigma^\star = -\alpha^{-1} D_\sigma(q_\sigma^\star) \in \mathcal{P}_\tau(I, V_h)$. Thus, ($\mathrm{P}_\sigma$) is equivalent to the minimization problem $\min_{q_\sigma \in \mathcal{P}_\tau(I,V_h)} j_\sigma(q_\sigma)$. Indeed, for a solution $q_\sigma^\star$ of (P_σ), it holds that

$$\min_{q_\sigma \in \mathcal{P}_\tau(I,V_h)} j_\sigma(q_\sigma) \le j_\sigma(q_\sigma^\star) = \min_{q_\sigma \in \mathcal{P}_\tau(I,W_h)} j_\sigma(q_\sigma).$$

The reverse inequality follows from the fact that $\mathcal{P}_\tau(I, V_h) \subset \mathcal{P}_\tau(I, W_h)$.

We have the following local boundedness of the discrete control-to-adjoint operator D_σ. The proof is omitted since it is similar to the discrete linearized problem (4.49), see Theorem 4.10.

Theorem 4.14. *Let $q \in Q$, $u_\sigma = S_\sigma(q)$ and $w_\sigma = D_\sigma(q)$. Then there exists a continuous function $\mathfrak{c} > 0$ such that for every $\sigma > 0$ we have*

$$\begin{aligned}
&\|w_\sigma\|_{L^\infty(I,X)} + \|w_\sigma\|_{L^2(I,W)} \\
&\le \mathfrak{c}(\|u_\sigma\|_{L^\infty(I_r,X)\cap L^2(I,W)})(\|u_\sigma - u_{d\sigma}\|_{L^2(I,X)} + \|u_h^{N_\tau} - u_{Th}\|_X + \|\nabla \times u_\sigma\|_{L^2(I,L^2(\Omega))}).
\end{aligned}$$

In the sequel, we shall derive error estimates by following the procedure already developed in the previous subsections. Recall that we do not have the compatibility of the terminal data for the adjoint problem if $\alpha_T > 0$.

Theorem 4.15. *Let $q \in Q$, $w = D(q)$ and $w_\sigma = D_\sigma(q)$. There exists a constant $c > 0$ depending continuously on $\|q\|_Q$, $\|f\|_{H^{2,1}(I)}$, $\|Au_0\|_H$, $\|z\|_{H^{2,1}(J_r)}$, $\|Au_T\|_H$ and $\|u_d\|_{H^{2,1}(I)}$ such that for every σ, we have the error estimate*

$$\max_{1\le k\le N_\tau} \|w(t_{k-1}) - w_h^k\|_X + \|w - w_\sigma\|_{L^\infty(I,X)} + \|w - w_\sigma\|_{L^2(I,W)} \le ch. \tag{4.63}$$

Proof. Let $\bar{w}_\sigma \in L^2(J^r, V_h)$ with $\bar{w}_h^k = P_h w(t_{k-1})$ for $k = 1, \dots, N_\tau$ and $\bar{w}_{h,k} = 0$ for $k = N_\tau + 1, \dots, N_\tau + N_r$. Also, let $\bar{u}_\sigma \in L^2(J_r, W_h)$ be as in the proof of Theorem 4.7. As usual, split the error by $w - w_\sigma = (w - \bar{w}_\sigma) + (\bar{w}_\sigma - w_\sigma)$. By Proposition 4.1, we have

$$\max_{1\le k\le N_\tau} \|w(t_{k-1}) - \bar{w}_h^k\|_X + \|w - \bar{w}_\sigma\|_{L^\infty(I,X)} + \|w - \bar{w}_\sigma\|_{L^2(I,W)} \le ch. \tag{4.64}$$

Multiplying the continuous adjoint problem (2.38) by the test function $\psi_h \in V_h \subset W$ and then integrating over $I_k \times \Omega$, we deduce that

$$\begin{aligned}
&(P_h w(t_{k-1}) - P_h w(t_k), \psi_h)_X + \int_{I_k} \nu(\nabla\psi_h, \nabla w(t))_X \,\mathrm{d}t - \int_{I_k} (\operatorname{div}\psi_h, \pi(t))\,\mathrm{d}t \\
&\quad = \int_{I_k} b(u^r(t), w(t), \psi_h)\,\mathrm{d}t + \int_{I_k} b(\psi_h, w^{-r}(t), u^{-r}(t))\,\mathrm{d}t \\
&\quad + \int_{I_k} \alpha_{\Omega_T}(u(t) - u_d(t), \psi_h)_X \,\mathrm{d}t + \int_{I_k} \alpha_R(\nabla\times u(t), \nabla\times\psi_h)_{L^2}\,\mathrm{d}t
\end{aligned}$$

for each $k = N_\tau, \dots, 1$. This equation can be rewritten as follows:

$$\begin{aligned}
&(d_{-\tau}\bar{w}_h^k, \psi_h)_X + \nu(\nabla\psi_h, \nabla\bar{w}_h^k)_X \\
&\qquad = b(\bar{u}_h^{k+1-N_r}, \bar{w}_h^{k+1}, \psi_h) + b(\psi_h, \bar{w}_h^{k+N_r}, \bar{u}_h^{k-1+N_r}) + \frac{1}{\tau} R_h^k(\psi_h)
\end{aligned} \tag{4.65}$$

where $d_{-\tau}\bar{w}_h^{N_\tau} := \tau^{-1}(\bar{w}_h^{N_\tau} - \alpha_T(u(T) - u_T))$ and $d_{-\tau}\bar{w}_h^k := \tau^{-1}(\bar{w}_h^k - \bar{w}_h^{k+1})$ for $k = N_\tau - 1, \ldots, 1$, and the residual term R_h^k is given by

$$\begin{aligned}
R_h^k(\psi_h) &= \int_{I_k} \nu(\nabla\bar{w}_h^k - \nabla w(t), \nabla\psi_h)_X \,\mathrm{d}t - \int_{I_k} (\operatorname{div}\psi_h, \pi(t) - \Gamma_h\pi(t))\,\mathrm{d}t \\
&\quad + \int_{I_k} b(u^r(t), w(t), \psi_h) - b(\bar{u}_h^{k+1-N_r}, \bar{w}_h^{k+1}, \psi_h)\,\mathrm{d}t \\
&\quad + \int_{I_k} b(\psi_h, w^{-r}(t), u^{-r}(t)) - b(\psi_h, \bar{w}_h^{k+N_r}, \bar{u}_h^{k-1+N_r})\,\mathrm{d}t \\
&\quad + \int_{I_k} \alpha_{\Omega_T}((u(t) - \bar{u}_h^k) - (u_d(t) - u_{dh}^k), \psi_h)_X \,\mathrm{d}t + \int_{I_k} \alpha_R(\nabla\times(u(t) - \bar{u}_h^k), \nabla\times\psi_h)_{L^2}\,\mathrm{d}t.
\end{aligned}$$

Let $e_\sigma := \bar{u}_\sigma - u_\sigma$ and $\eta_\sigma := \bar{w}_\sigma - w_\sigma$. Taking the difference of (4.62) and (4.65), we get

$$\begin{aligned}
(d_{-\tau}\eta_h^k, \psi_h)_X + \nu(\nabla\psi_h, \nabla\eta_h^k)_X &= b(e_h^{k+1-N_r}, \eta_h^{k+1}, \psi_h) + b(\psi_h, \eta_h^{k+N_r}, e_h^{k-1+N_r}) \\
&\quad + b(u_h^{k+1-N_r}, \eta_h^{k+1}, \psi_h) + b(\psi_h, w_h^{k+N_r}, e_h^{k-1+N_r}) + \frac{1}{\tau}R_h^k(\psi_h)
\end{aligned} \tag{4.66}$$

where $d_{-\tau}\eta_h^{N_\tau} := \tau^{-1}(\bar{w}_h^{N_\tau} - w(T) - \alpha_T[(u(T) - u_h^{N_\tau}) - (u_T - u_{Th})])$ and $d_{-\tau}\eta_h^k := \tau^{-1}(\eta_h^k - \eta_h^{k+1})$ for $k = N_\tau - 1, \ldots, 1$. Observe that (4.66) has the same form as (4.57), however, the superscripts are now in descending order. Nevertheless, one can adapt the same methods and use the backward version of the discrete Gronwall Lemma 7.2. For this reason, it suffices to derive an estimate for the term $R_h^k(\eta_h^k)$. Let $I_{h,k}$ and $J_{h,k}$ denote the third and fourth integrals in $R_h^k(\eta_h^k)$. By (A2), Proposition 4.2 and Theorem 4.7, for each $\varepsilon > 0$ we have

$$\sum_{k=1}^{N_\tau} |R_h^k(\eta_h^k) - I_{h,k} - J_{h,k}| \le \varepsilon\sum_{k=1}^{N_\tau} \nu\tau\|\nabla\eta_h^k\|_X^2 + c_\varepsilon h^2. \tag{4.67}$$

We are now going to estimate $I_{h,k}$ and $J_{h,k}$. To this end, let us write these terms by $I_{h,k} = I_{h,k}^a + I_{h,k}^b$ and $I_{h,k} = J_{h,k}^a + J_{h,k}^b$, where

$$\begin{aligned}
I_{h,k}^a + I_{h,k}^b &:= \int_{I_k} b(u^r(t) - \bar{u}_h^{k+1-N_r}, w(t), \eta_h^k)\,\mathrm{d}t + \int_{I_k} b(\bar{u}_h^{k+1-N_r}, w(t) - \bar{w}_h^{k+1}, \eta_h^k)\,\mathrm{d}t \\
J_{h,k}^a + J_{h,k}^b &:= \int_{I_k} b(\eta_h^k, w^{-r}(t) - \bar{w}_h^{k+N_r}, u^{-r}(t))\,\mathrm{d}t + \int_{I_k} b(\eta_h^k, \bar{w}_h^{k+N_r}, \bar{u}_h^{k-1+N_r} - u^{-r}(t))\,\mathrm{d}t.
\end{aligned}$$

By the Gagliardo–Nirenberg and Hölder inequalities, $I_{h,k}^a$ and $I_{h,k}^b$ satisfy the following estimates:

$$\begin{aligned}
|I_{h,k}^a| &\le \varepsilon\nu\tau\|\nabla\eta_h^k\|_X^2 + c_\varepsilon\|\nabla w\|_{L^\infty(I_k,X)}^2 \int_{I_k} \|\nabla u^r(t) - \nabla P_h u^r(t_{k+1})\|_X^2\,\mathrm{d}t \\
|I_{h,k}^b| &\le \varepsilon\nu\tau\|\nabla\eta_h^k\|_X^2 + c_\varepsilon\|\nabla u^r\|_{L^\infty(I_k,X)}^2 \int_{I_k} \|\nabla w(t) - \nabla P_h w(t_k)\|_X^2\,\mathrm{d}t, \quad k = N_\tau - 1, \ldots, 1, \\
|I_{h,N_\tau}^b| &\le \varepsilon\nu\tau\|\nabla\eta_h^k\|_X^2 + c_\varepsilon\tau\|\nabla u^r\|_{L^\infty(I_k,X)}^2\|\nabla w\|_{L^\infty(I_{N_\tau},X)}^2.
\end{aligned}$$

On the other hand, $J_{h,k} = 0$ if $k = N_\tau, \ldots, N_\tau - N_r + 1$, while for $k = N_\tau - N_r, \ldots, 1$ we have

$$\begin{aligned}
|J_{h,k}^a| &\le \varepsilon\nu\tau\|\nabla\eta_h^k\|_X^2 + c_\varepsilon\|\nabla u^{-r}\|_{L^\infty(I_k,X)}^2 \int_{I_k} \|\nabla w^{-r}(t) - \nabla w^{-r}(t_{k-1})\|_X^2\,\mathrm{d}t \\
|J_{h,k}^b| &\le \varepsilon\nu\tau\|\nabla\eta_h^k\|_X^2 + c_\varepsilon\|\nabla w^{-r}\|_{L^\infty(I_k,X)}^2 \int_{I_k} \|\nabla u^{-r}(t) - \nabla u^{-r}(t_{k-1})\|_X^2\,\mathrm{d}t.
\end{aligned}$$

Combining the previous estimates, taking the sum over all $k = 1, \ldots, N_\tau$ and then using (4.67) lead to the following estimate:

$$\sum_{k=1}^{N_\tau} |R_h^k(\eta_h^k)| \leq \sum_{k=1}^{N_\tau} |R_h^k(\eta_h^k) - I_{h,k} - J_{h,k}| + \sum_{k=1}^{N_\tau} |I_{h,k} + J_{h,k}| \leq 5\varepsilon \sum_{k=1}^{N_\tau} \nu\tau \|\nabla \eta_h^k\|_X^2 + c_\varepsilon h^2. \tag{4.68}$$

If we take $\psi_h = \eta_h^k$ in (4.66), follow the remarks mentioned above and then take $\varepsilon > 0$ small enough so that the first term on the right hand side of (4.68) can be absorbed to the left, then one would eventually arrive at the inequality

$$\max_{1\leq k\leq N_\tau} \|\bar{w}_h^k - w_h^k\|_X + \|\eta_\sigma\|_{L^\infty(I,X)} + \|\eta_\sigma\|_{L^2(I,W)} \leq ch. \tag{4.69}$$

Therefore, we can see that (4.63) holds true due to (4.64) and (4.69). □

Theorem 4.16. *Under the assumptions of the Theorem 4.15, there exists a constant $c > 0$ independent on σ such that*

$$\|w - w_\sigma\|_{L^2(I,X)} \leq c(\alpha_{\Omega_T} h + \alpha_R + \alpha_T)h. \tag{4.70}$$

Proof. We again proceed with a duality argument, but now taking advantage of the error estimates on the solutions of the state and linearized state equations. Let $g \in L^2(I,X)$ with $\|g\|_{L^2(I,X)} \leq 1$. Then, according to (3.1) and Lemma 4.12, we have

$$\begin{aligned}\int_I (w(t) - w_\sigma(t), g(t))_X \,\mathrm{d}t &= j'(q)g - j'_\sigma(q)g \\ &= \alpha_{\Omega_T} \int_I (u(t) - u_d(t), v(t))_X - (u_\sigma(t) - u_{d\sigma}(t), v_\sigma(t))_X \,\mathrm{d}t \\ &\quad + \alpha_R \int_I (\nabla \times u(t), \nabla \times v(t))_{L^2} - (\nabla \times u_\sigma(t), \nabla \times v_\sigma(t))_{L^2} \,\mathrm{d}t \\ &\quad + \alpha_T((u(T) - u_T, v(T))_X - (u_\sigma(T) - \mathit{\Pi}_h u_T, v_\sigma(T))_X),\end{aligned}$$

where $v = S'(q)g$ and $v_\sigma = S'_\sigma(q)g$. From the Cauchy–Schwarz inequality and the error estimates (4.32) and (4.55), we deduce that

$$\begin{aligned}&\int_I |(u(t), v(t))_X - (u_\sigma(t), v_\sigma(t))_X| \,\mathrm{d}t \\ &\qquad \leq \|u - u_\sigma\|_{L^2(I,X)} \|v\|_{L^2(I,X)} + \|u_\sigma\|_{L^2(I,X)} \|v - v_\sigma\|_{L^2(I,X)} \leq ch^2 \\ &\int_I |(u_d(t), v(t))_X - (u_{d\sigma}(t), v_\sigma(t))_X| \,\mathrm{d}t \\ &\qquad \leq \|u_d - u_{d\sigma}\|_{L^2(I,X)} \|v\|_{L^2(I,X)} + \|u_{d\sigma}\|_{L^2(I,X)} \|v - v_\sigma\|_{L^2(I,X)} \leq ch^2.\end{aligned}$$

Using similar decompositions along with the error estimates (4.20) and (4.54), we get

$$\begin{aligned}\int_I |(\nabla \times u(t), \nabla \times v(t))_{L^2} - (\nabla \times u_\sigma(t), \nabla \times v_\sigma(t))_{L^2}| \,\mathrm{d}t &\leq ch \\ |(u(T), v(T))_X - (u_\sigma(T), v_\sigma(T))_X| &\leq ch.\end{aligned}$$

Also, (A2) and (4.54) imply that $|(u_T, v(T))_X - (\mathit{\Pi}_h u_T, v_\sigma(T))_X| \leq c(h^2 + h) \leq ch$. Thus, we obtain that $(w - w_\sigma, g)_{L^2(I,X)} \leq c(\alpha_{\Omega_T} h + \alpha_R + \alpha_T)h$ for every $g \in L^2(I,X)$ such that $\|g\|_{L^2(I,X)} \leq 1$, and this results into the desired inequality (4.70) by duality. □

Theorem 4.17. *If $q_\sigma \to q$ in Q and $g_\sigma \rightharpoonup g$ in Q as $\sigma \to 0$, then*

$$\lim_{\sigma\to 0}(j_\sigma''(q_\sigma)[g_\sigma, g_\sigma] - \alpha\|g_\sigma\|_Q^2) = j''(q)[g,g] - \alpha\|g\|_Q^2 \tag{4.71}$$

$$\liminf_{\sigma\to 0} j_\sigma''(q_\sigma)[g_\sigma, g_\sigma] \geq j''(q)[g,g]. \tag{4.72}$$

Proof. Let $v = S'(q)g$, $w = D(q)$, $v_\sigma = S_\sigma'(q_\sigma)g_\sigma$ and $w_\sigma = D_\sigma(q_\sigma)$. By assumption, $\{q_\sigma\}_\sigma$ and $\{g_\sigma\}_\sigma$ are bounded in Q. This implies that both $\{v_\sigma\}_\sigma$ and $\{w_\sigma\}_\sigma$ are bounded in $L^\infty(I, X) \cap L^2(I, W)$ by Theorems 4.5, 4.10 and 4.14 .

Let $\tilde{v}_\sigma = S'(q_\sigma)g_\sigma$ and $\widehat{v}_\sigma = S'(q)g_\sigma$, and consider the decomposition $v_\sigma - v = (v_\sigma - \tilde{v}_\sigma) + (\tilde{v}_\sigma - \widehat{v}_\sigma) + (\widehat{v}_\sigma - v)$. By Theorem 4.11, $v_\sigma - \tilde{v}_\sigma \to 0$ in $L^2(I, W)$ and $v_\sigma(T) - \tilde{v}_\sigma(T) \to 0$ in X. Using the mean value theorem and $q_\sigma \to q$ in Q, we have $\tilde{v}_\sigma - \widehat{v}_\sigma \to 0$ in $H^{2,1}(I)$. As a particular case, $\tilde{v}_\sigma - \widehat{v}_\sigma \to 0$ in $L^2(I, W)$ and $\tilde{v}_\sigma(T) - \widehat{v}_\sigma(T) \to 0$ in X up to a subsequence due to the compactness of the embeddings $H^{2,1}(I) \subset L^2(I, W)$ and $V \subset H$. For the third difference, $\widehat{v}_\sigma - v \rightharpoonup 0$ in $H^{2,1}(I)$ by the continuity of the linear map $S'(q) : Q \to H^{2,1}(I)$ and $g_\sigma \rightharpoonup g$ in Q . By compactness of the said embeddings once more, we get $\widehat{v}_\sigma - v \to 0$ in $L^2(I, W)$ and $\widehat{v}_\sigma(T) - v(T) \to 0$ in H up to a subsequence. Therefore, $v_\sigma \to v$ in $L^2(I, W)$ and $v_\sigma(T) \to v(T)$ in X.

Let $\tilde{w}_\sigma = D(q_\sigma)$. We get from Theorem 4.15 that $w_\sigma - \tilde{w}_\sigma \to 0$ in $L^2(I, W)$. On the other hand, the stability estimate in Corollary 2.17 implies that $\tilde{w}_\sigma \to w$ in $H^{2,1}(I)$. Hence, $w_\sigma \to w$ in $L^2(I, W)$ up to a subsequence. We can now pass to the limit in the trilinear term in $j_\sigma''(q_\sigma)[g_\sigma, g_\sigma]$, that is,

$$\begin{aligned}
&\lim_{\sigma\to 0}\int_I b(v_\sigma^r(t), w_\sigma(t), v_\sigma^\tau(t)) - b(v(t), w(t), v(t))\,\mathrm{d}t \\
&\quad = \lim_{\sigma\to 0}\int_I b(v_\sigma^r(t) - v^r(t), w_\sigma(t), v_\sigma^\tau(t))\,\mathrm{d}t + \lim_{\sigma\to 0}\int_I b(v^r(t), w_\sigma(t) - w(t), v_\sigma^\tau(t))\,\mathrm{d}t \\
&\quad + \lim_{\sigma\to 0}\int_I b(v^r(t), w(t), v_\sigma^\tau(t) - v^\tau(t))\,\mathrm{d}t + \lim_{\sigma\to 0}\int_I b(v^r(t), w(t), v^\tau(t) - v(t))\,\mathrm{d}t = 0.
\end{aligned}$$

This implies the limit (4.71). To obtain the inequality (4.72), we just need to rewrite $j_\sigma''(q_\sigma)[g_\sigma, g_\sigma] = (j_\sigma''(q_\sigma)[g_\sigma, g_\sigma] - \alpha\|g_\sigma\|_Q^2) + \alpha\|g_\sigma\|_Q^2$, apply (4.71), take the limit inferior and then apply the lower semi-continuity of the norm. □

4.6. Error Estimates for the Discrete Optimal Control Problem

The goal of this subsection is to derive error estimates between the solutions of the optimal control problems (P) and (P_σ). Here, we follow the strategy developed in [11,12]. Let us start with the following lemma with a sequence of controls that converges weakly.

Lemma 4.18. *Let $\{q_\sigma\}_\sigma \subset Q_\sigma$ and $q \in Q$ be such that $q_\sigma \rightharpoonup q$ in Q as $\sigma \to 0$. If $u = S(q)$ and $u_\sigma = S_\sigma(q_\sigma)$, then*

$$\lim_{\sigma\to 0}(\|u - u_\sigma\|_{L^2(I,W)} + \|u(T) - u_\sigma(T)\|_X) = 0.$$

Proof. Let $\bar{u}_\sigma = S(q_\sigma)$ and consider $u - u_\sigma = (u - \bar{u}_\sigma) + (\bar{u}_\sigma - u_\sigma)$. Since q_σ is bounded in Q, it follows from Theorem 4.7 that $\|\bar{u}_\sigma - u_\sigma\|_{L^2(I,W)} + \|\bar{u}_\sigma(T) - u_\sigma(T)\|_X \to 0$. Recall from Theorem 2.10 that $S : Q \to H^{2,1}(I)$ is weak-weak continuous, hence $\bar{u}_\sigma \rightharpoonup u$ in $H^{2,1}(I)$. By compactness arguments, one can show up to a subsequence that $\|u - \bar{u}_\sigma\|_{L^2(I,W)} + \|u(T) - \bar{u}_\sigma(T)\|_X \to 0$, and the conclusion follows from the triangle inequality. □

Theorem 4.19. *For each $\sigma = (\tau, h)$ satisfying* (A5)*, let $q_\sigma^\star$ be a global solution of* (P_σ)*. Then $\{q_\sigma^\star\}_\sigma$ is bounded in Q. Any weak limit point $\bar{q} \in Q$ of $\{q_\sigma^\star\}_\sigma$ as $\sigma \to 0$ is a solution to* (P) *and it satisfies $q_\sigma^\star \to \bar{q}$ in Q and $j_\sigma(q_\sigma^\star) \to j(\bar{q})$.*

Proof. Suppose that $\widehat{q} \in Q$ is a solution of (P) and let $q_\sigma = R_\tau P_h \widehat{q}$. Since $\widehat{q} \in H^{2,1}(I)$, it follows from Proposition 4.3 that $\|q_\sigma - \widehat{q}\|_Q \to 0$. From the previous lemma, if $u_\sigma = S_\sigma(u_\sigma)$ and $\widehat{u} = S(\widehat{q})$, then $u_\sigma \to \widehat{u}$ in $L^2(I, W)$ and $u_\sigma(T) \to \widehat{u}(T)$ in X. By construction, $u_{d\sigma} \to u_d$ in $L^2(I, X)$ and $u_{Th} \to u_T$ in X. These imply that $j_\sigma(q_\sigma) \to j(\widehat{q})$. According to the optimality of $q^\star_\sigma$ with respect to j_σ, we have $\|q^\star_\sigma\|^2_Q \le \frac{2}{\alpha} j_\sigma(q^\star_\sigma) \le \frac{2}{\alpha} j_\sigma(q_\sigma)$, where the right hand side is bounded for $\sigma > 0$. Hence, $\{q^\star_\sigma\}_\sigma$ has a weak limit point in Q.

Let $\bar{q} \in Q$ be any such weak limit point. Again, invoking the previous lemma and using the same argument as above, it follows that $j_\sigma(q^\star_\sigma) \to j(\bar{q})$, and if $\bar{u} = S(\bar{q})$ then $u^\star_\sigma \to \bar{u}$ in $L^2(I, W)$ and $u^\star_\sigma(T) \to \bar{u}(T)$ in X. From the definitions of j and j_σ, these imply that $\|q^\star_\sigma\|_Q \to \|\bar{q}\|_Q$, and therefore $q^\star_\sigma \to \bar{q}$ in Q. Moreover, by the optimality of $q^\star_\sigma$ with respect to j_σ once more, one has

$$j(\bar{q}) = \lim_{\sigma\to 0} j_\sigma(q^\star_\sigma) \le \lim_{\sigma\to 0} j_\sigma(q_\sigma) = j(\widehat{q}),$$

and this shows that $\bar{q}$ is a solution of (P). □

Corollary 4.20. *If $q^\star \in Q$ is a strict local solution of (P), then there exists a sequence $\{q^\star_\sigma\}_\sigma \subset Q_\sigma$ of solutions to (P_σ) such that $q^\star_\sigma \to q^\star$ in Q and $j_\sigma(q^\star_\sigma) \to j(q^\star)$.*

Proof. By assumption, there is $\delta > 0$ such that $q^\star$ is the only minimizer of j in $B_\delta(q^\star) := \{q \in Q : \|q - q^\star\| < \delta\}$. From the proof of the previous theorem with $\widehat{q} = q^\star$, there is a sequence $\{q^\star_\sigma\}_\sigma \subset Q_\sigma$ of solutions to (P_σ) in $B_r(q^\star)$ such that $q^\star_\sigma \to \bar{q}$ and $j_\sigma(q^\star_\sigma) \to j(\bar{q})$ for some $\bar{q} \in B_r(q^\star)$. Since $q^\star$ is the only solution of j in $B_r(q^\star)$, it follows that $\bar{q} = q^\star$. □

Let us now prove error estimates for strict local minimizers having coercive second derivatives.

Theorem 4.21. *Suppose that $j'(q^\star)g = 0$ for every $g \in Q$ and there is a constant $\mu > 0$ such that $j''(q^\star)[q, q] \ge \mu\|q\|^2_Q$ for every $q \in Q$, so that $q^\star$ is a strict local solution of* (P). *For each $\sigma = (\tau, h)$ satisfying* (A5)*, let $q^\star_\sigma \in Q_\sigma$ be a solution to (P_σ) with $q^\star_\sigma \to q^\star$ in Q. Then for some constant $c > 0$ independent on σ, we have*

$$\|q^\star - q^\star_\sigma\|_Q \le c[(\alpha_{\Omega_T} + \alpha + 1)h + \alpha_T + \alpha_R]h. \tag{4.73}$$

Proof. The existence of the sequence $\{q^\star_\sigma\}_\sigma \subset Q_\sigma$ is guaranteed by the previous corollary. Let $\bar{q}^\star_\sigma := R_\tau P_h q^\star \in Q_\sigma$. Since $q^\star = -\alpha^{-1} D(q^\star) \in H^{2,1}(I)$, we have $\|q^\star_\sigma - \bar{q}^\star_\sigma\|_Q \le ch^2$ by Proposition 4.3. Let $e_\sigma := \bar{q}^\star_\sigma - q^\star_\sigma \in Q_\sigma$ and $g_\sigma := e_\sigma/\|e_\sigma\|$. Up to a subsequence, $g_\sigma \rightharpoonup g$ in Q for some $g \in Q$. From the mean-value theorem and the local optimality of $q^\star_\sigma$, there is a $\lambda_\sigma \in (0, 1)$ such that if $q_\sigma := \lambda_\sigma \bar{q}^\star_\sigma + (1 - \lambda_\sigma)q^\star_\sigma$, then

$$j'_\sigma(\bar{q}^\star_\sigma)e_\sigma = j'_\sigma(\bar{q}^\star_\sigma)e_\sigma - j'_\sigma(q^\star_\sigma)e_\sigma = j''_\sigma(q_\sigma)[e_\sigma, e_\sigma].$$

Since $q_\sigma \to q^\star$ in Q, $g_\sigma \rightharpoonup g$ in Q, $\|g_\sigma\|_Q = 1$ and $\|g\|_Q \le 1$, (4.71) gives us

$$\lim_{\sigma\to 0}(j''_\sigma(q_\sigma)[g_\sigma, g_\sigma] - \alpha) = j''(q^\star)[g, g] - \alpha\|g\|^2_Q \ge \mu - \alpha.$$

Hence, there exists $\sigma_0 > 0$ such that $j''_\sigma(q_\sigma)[e_\sigma, e_\sigma] \ge \frac{\mu}{2}\|e_\sigma\|^2_Q$ whenever $|\sigma| < \sigma_0$. By the local optimality of $q^\star$ and $Q_\sigma \subset Q$, we have

$$\frac{\mu}{2}\|e_\sigma\|^2_Q \le j'_\sigma(\bar{q}^\star_\sigma)e_\sigma = (j'_\sigma(\bar{q}^\star_\sigma)e_\sigma - j'_\sigma(q^\star)e_\sigma) + (j'_\sigma(q^\star)e_\sigma - j'(q^\star)e_\sigma). \tag{4.74}$$

Note that $j'_\sigma(\bar{q}^\star_\sigma)e_\sigma - j'_\sigma(q^\star)e_\sigma = (D_\sigma(\bar{q}^\star_\sigma) - D_\sigma(q^\star), e_\sigma)_{L^2(I,X)} + \alpha(\bar{q}^\star_\sigma - q^\star, e_\sigma)_Q$ by Lemma 4.12. From the invariance property mentioned in Remark 4.4, one has $S_\sigma(q^\star) = S_\sigma(R_\tau P_h q^\star) = S_\sigma(\bar{q}^\star_\sigma)$, and consequently $D_\sigma(\bar{q}^\star_\sigma) = D_\sigma(q^\star)$. Thus,

$$|j'_\sigma(\bar{q}^\star_\sigma)e_\sigma - j'_\sigma(q^\star)e_\sigma| \le \alpha\|\bar{q}^\star_\sigma - q^\star\|_Q\|e_\sigma\|_Q. \tag{4.75}$$

Similarly, from $j'_\sigma(q^\star)e_\sigma - j'(q^\star)e_\sigma = (D_\sigma(q^\star) - D(q^\star), e_\sigma)_{L^2(I,X)}$, we obtain

$$|j'_\sigma(q^\star)e_\sigma - j'(q^\star)e_\sigma| \le \|D_\sigma(q^\star) - D(q^\star)\|_{L^2(I,X)}\|e_\sigma\|_Q. \tag{4.76}$$

Invoking (4.75) and (4.76) in (4.74) yields

$$\begin{aligned}\|q^\star - q^\star_\sigma\|_Q &\le \|q^\star - \bar{q}^\star_\sigma\|_Q + \|e_\sigma\|_Q \\ &\le \|\bar{q}^\star_\sigma - q^\star\|_Q + \frac{2}{\mu}(\alpha\|\bar{q}^\star_\sigma - q^\star\|_Q + \|D_\sigma(q^\star) - D(q^\star)\|_{L^2(I,X)}).\end{aligned}$$

Utilizing $\|q^\star_\sigma - \bar{q}^\star_\sigma\|_Q \le ch^2$ and Theorem 4.16, we finally obtain (4.73). □

Corollary 4.22. *Consider the framework of Theorem 4.21. If $u^\star = S(q^\star)$, $u^\star_\sigma = S_\sigma(q^\star_\sigma)$, $w^\star = D(q^\star)$ and $w^\star_\sigma = D_\sigma(q^\star_\sigma)$ are the optimal states and adjoint states corresponding to $q^\star$ and $q^\star_\sigma$, then there exists $c > 0$ such that for each σ, we have*

$$\|u^\star - u^\star_\sigma\|_{L^2(I,X)} + \|w^\star - w^\star_\sigma\|_{L^2(I,X)} \le c[(\alpha_{\Omega_T} + \alpha + 1)h + \alpha_T + \alpha_R]h.$$

Furthermore, there is a constant $c > 0$ not depending on σ such that

$$\|q^\star - q^\star_\sigma\|_{L^\infty(I,X)\cap L^2(I,W)} + \|u^\star - u^\star_\sigma\|_{L^\infty(I,X)\cap L^2(I,W)} + \|w^\star - w^\star_\sigma\|_{L^\infty(I,X)\cap L^2(I,W)} \le ch.$$

Proof. From the mean value theorem, (4.73) and Theorem 4.9, we have

$$\begin{aligned}\|u^\star - u^\star_\sigma\|_{L^2(I,X)} &\le \|S(q^\star) - S(q^\star_\sigma)\|_{L^2(I,X)} + \|S(q^\star_\sigma) - S_\sigma(q^\star_\sigma)\|_{L^2(I,X)} \\ &\le c\|q^\star - q^\star_\sigma\|_Q + ch^2 \le c[(\alpha_{\Omega_T} + \alpha + 1)h + \alpha_T + \alpha_R]h.\end{aligned}$$

For the adjoint states, we just need to apply Theorem 4.16, Corollary 2.17 and (4.73) so that

$$\begin{aligned}\|w^\star - w^\star_\sigma\|_{L^2(I,X)} &\le \|D(q^\star) - D(q^\star_\sigma)\|_{L^2(I,X)} + \|D(q^\star_\sigma) - D_\sigma(q^\star_\sigma)\|_{L^2(I,X)} \\ &\le c[(\alpha_{\Omega_T} + \alpha + 1)h + \alpha_T + \alpha_R]h.\end{aligned}$$

These prove the first error estimate.

The proof of the second error estimate will be carried out with the help of Theorem 4.7 and Theorem 4.15. In the case of the state variables, using the continuity of the embedding $H^{2,1}(I) \subset L^\infty(I,X) \cap L^2(I,W)$ and $S \in C^\infty(Q, H^{2,1}(I))$ one obtains that

$$\begin{aligned}\|u^\star - u^\star_\sigma\|_{L^\infty(I,X)\cap L^2(I,W)} &\le \|S(q^\star) - S(q^\star_\sigma)\|_{H^{2,1}(I)} + \|S(q^\star_\sigma) - S_\sigma(q^\star_\sigma)\|_{L^\infty(I,X)\cap L^2(I,W)} \\ &\le c\|q^\star - q^\star_\sigma\|_Q + ch \le ch.\end{aligned}$$

The case of the adjoint variables can be handled in a similar fashion thanks to Corollary 2.17. Finally, for the controls, we use $q^\star - q^\star_\sigma = -\alpha^{-1}(w^\star - w^\star_\sigma)$ and apply the error estimate for the adjoint states. □

5. Approximation of the Optimal Control Problem

This section deals with the specific aspects of computing numerically a solution to the fully discrete optimal control problem (P_σ). We shall focus the discussion in the case of the triangular Taylor-Hood element, however, the adaptation to the mini-finite element can be carried out in a similar fashion. The resulting finite-dimensional optimization problem will be an approximation of (P_σ). This is due to the variational crimes committed from the use of numerical quadrature and the addition of an artificial compressibility penalty term for the elimination of the pressure. The analysis of additional errors due to these processes is beyond the scope of the current paper. Nevertheless, artificial compressibility penalizations to the uncontrolled Navier–Stokes equation without delay has been examined in the series of papers [13–15]. There are several available softwares for the implementation of the finite element method, however, we shall do our own assembly of the finite element matrices.

5.1. Penalized Finite-Dimensional Approximation

Let $\mathcal{K}_h = \{K_h\}$ be a shape-regular triangulation of a convex polygonal domain Ω, $\{x_{h,s}\}_{s=1}^{n_{ph}}$ be the corresponding set of nodes and $\{x_{h,i}\}_{i=n_{ph+1}}^{n_h}$ be the set of edge midpoints of the triangles. Denote by $\mathbb{P}^k(K_h)$ the space of all polynomials on K_h with degree at most k. Consider the finite-dimensional spaces

$$M_h := \{\rho_h \in C(\bar{\Omega}) : \rho_h|_K \in \mathbb{P}^1(K_h)\ \forall K_h \in \mathcal{K}_h\}$$
$$Y_h := \{\phi_h \in C(\bar{\Omega}) : \phi_h|_K \in \mathbb{P}^2(K_h)\ \forall K_h \in \mathcal{K}_h\}$$

and set $W_h := Y_h \times Y_h$. Let $\{\phi_{h,i}\}_{i=1}^{n_h}$ be the Lagrange nodal basis for Y_h such that $\phi_{h,i}(x_{h_j}) = \delta_{ij}$ for every $i,j = 1,\dots,n_h$, where δ_{ij} is the Kronecker delta symbol. Similarly, let $\{\rho_{h,s}\}_{s=1}^{n_{ph}}$ be the Lagrange nodal basis for M_h with $\rho_{h,s}(x_{h,\ell}) = \delta_{s\ell}$ for each $s,\ell = 1,\dots,n_{ph}$. For $i = 1,\dots,n_h$, let us define $\varphi_{h,i} := [\phi_{h,i}, 0]^\top$ and $\varphi_{h,n_h+i} := [0, \phi_{h,i}]^\top$, so that $\{\varphi_{h,i}\}_{i=1}^{2n_h}$ forms a basis for W_h.

Given $K_h \in \mathcal{K}_h$, let $T_{K_h}\xi = A_{K_h}\xi + b_{K_h}$, where $A_{K_h} \in \mathbb{R}^{2\times 2}$ and $b_{K_h} \in \mathbb{R}^2$, be the affine transformation from the reference triangle K_{ref}, having vertices at $(0,0)$, $(1,0)$ and $(0,1)$, to the physical triangle K_h. Suppose that $\{\xi_\ell\}_{\ell=1}^g$ and $\{\omega_\ell\}_{\ell=1}^g$ are Gaussian quadrature nodes and weights on K_{ref}. To simplify the formulas, let $\omega_{K_h,\ell} := \omega_\ell |\det A_{K_h}|$, $\rho_{K_h,s} := \rho_{h,s} \circ T_{K_h}$, $\phi_{K_h,i} := \phi_{h,i} \circ T_{K_h}$, and $\varphi_{K_h,i} := \varphi_{h,i} \circ T_{K_h}$.

In evaluating the integrals for the contributions of the basis functions in the finite element matrices, we shall apply the transformation formula from K_h to K_{ref} and then use Gaussian quadrature. The entries of the mass matrix $\tilde{N}_h \in \mathbb{R}^{n_h \times n_h}$ and stiffness matrix $\tilde{A}_h \in \mathbb{R}^{n_h \times n_h}$ can be calculated as follows:

$$\int_\Omega \phi_{h,i}\phi_{h,j}\,\mathrm{d}x \approx \sum_{\ell=1}^{g} \sum_{K_h \in \mathcal{K}_h} \omega_{K_h,\ell}\phi_{K_h,i}(\xi_\ell)\phi_{K_h,j}(\xi_\ell) =: [\tilde{N}_h]_{i,j}$$
$$\int_\Omega \nabla\phi_{h,i} \cdot \nabla\phi_{h,j}\,\mathrm{d}x \approx \sum_{\ell=1}^{g} \sum_{K_h \in \mathcal{K}_h} \omega_{K_h,\ell} A_{K_h}^{-\top}\nabla_\xi \phi_{K_h,i}(\xi_\ell) \cdot A_{K_h}^{-\top}\nabla_\xi \phi_{K_h,j}(\xi_\ell) =: [\tilde{A}_h]_{i,j}$$

for $i,j = 1,\dots,n_h$. These corresponds to a component of the velocity field. The full mass and stiffness matrices are given by $N_h := \tilde{N}_h \otimes I_2$ and $A_h := \tilde{A}_h \otimes I_2$, respectively. Here, $\otimes$ is the tensor product and I_2 is the 2×2 identity matrix.

Write the first and second components of the differential operator $A_{K_h}^{-\top}\nabla_\xi$ by $(A_{K_h}^{-\top}\nabla_\xi)^{x_1}$ and $(A_{K_h}^{-\top}\nabla_\xi)^{x_2}$, respectively. From this, the entries of the discrete divergence matrix $B_h = [B_h^{x_1}\ B_h^{x_2}] \in \mathbb{R}^{n_{ph} \times 2n_h}$ are computed, for $i = 1,\dots,n_h$ and $s = 1,\dots,n_{ph}$, according to

$$\int_\Omega (\partial_{x_a}\phi_{h,i})\rho_{h,s}\,\mathrm{d}x \approx \sum_{\ell=1}^{g} \sum_{K_h \in \mathcal{K}_h} \omega_{K_h,\ell}(A_{K_h}^{-\top}\nabla_\xi)^{x_a}\phi_{K_h,i}(\xi_\ell)\rho_{K_h,s}(\xi_\ell) =: [B_h^{x_a}]_{s,i}, \quad a = 1,2.$$

Given an approximation $u_h = \sum_{m=1}^{2n_h} u_{h,m}\varphi_{h,m} \in W_h$ of $u \in W$, where $u_{h,m} \in \mathbb{R}$, the entries of the associated convection matrix $C_h(u_h) \in \mathbb{R}^{2n_h \times 2n_h}$ and *dual* convection matrix $D_h(u_h) \in \mathbb{R}^{2n_h \times 2n_h}$ will be determined as follows:

$$\int_\Omega (u \cdot \nabla)\varphi_{h,j} \cdot \varphi_{h,i}\,\mathrm{d}x$$
$$\approx \sum_{\ell=1}^{g} \sum_{K_h \in \mathcal{K}_h} \sum_{m=1}^{2n_h} \omega_{K_h,\ell} u_{h,m} (\varphi_{K_h,m}(\xi_\ell) \cdot A_{K_h}^{-\top}\nabla_\xi)\varphi_{K_h,j}(\xi_\ell) \cdot \varphi_{K_h,i}(\xi_\ell) =: [C_h(u_h)]_{i,j}$$
$$\int_\Omega (\nabla u)^\top \varphi_{h,i} \cdot \varphi_{h,j}\,\mathrm{d}x$$
$$\approx \sum_{\ell=1}^{g} \sum_{K_h \in \mathcal{K}_h} \sum_{m=1}^{2n_h} \omega_{K_h,\ell} u_{h,m} (A_{K_h}^{-\top}\nabla_\xi \varphi_{K_h,m}(\xi_\ell))^\top \varphi_{K_h,i}(\xi_\ell) \cdot \varphi_{K_h,j}(\xi_\ell) =: [D_h(u_h)]_{i,j}$$

for each $i, j = 1, \ldots, 2n_h$. Finally, the matrix $R_h \in \mathbb{R}^{2N_h \times 2N_h}$ corresponding to vorticity has the following entries for each $i, j = 1, \ldots, 2n_h$

$$\int_\Omega (\nabla \times \varphi_{h,i})(\nabla \times \varphi_{h,j})\,\mathrm{d}x \approx \sum_{\ell=1}^{g} \sum_{K_h \in \mathcal{K}_h} \omega_{K_h,\ell}(A_{K_h}^{-\top}\nabla_\xi \times \varphi_{K_h,i}(\xi_\ell))(A_{K_h}^{-\top}\nabla_\xi \times \varphi_{K_h,j}(\xi_\ell)) =: [R_h]_{i,j}.$$

In what follows, a Gaussian quadrature of order 6 having $g = 12$ nodes on the reference element will be applied. This is sufficient since this order is the highest possible degree of a polynomial that can appear in the above integrals with quadratic basis functions, in particular, for the matrices $C_h(u_h)$ and $D_h(u_h)$. Further practical aspects in matrix assembly can be found in [23,44].

We shall identify $\mathcal{P}_\tau(I, W_h)$ to $\mathbb{R}^{N_\tau \times 2n_h}$ and $\mathcal{P}_\tau(I_r, W_h)$ to $\mathbb{R}^{N_r \times 2n_h}$. Moreover, we set $q_\sigma = \{q_h^k\}_{k=1}^{N_\tau}$ where $q_h^k \in \mathbb{R}^{2n_h}$, and similar representation for the elements of $\mathcal{P}_\tau(I_r, W_h)$. The discrete state equation we consider is the following: Given $q_\sigma = \{q_h^k\}_{k=1}^{N_\tau} \in \mathbb{R}^{N_\tau \times 2n_h}$, $u_{0h} \in \mathbb{R}^{2n_h}$ and $z_\sigma = \{z_h^j\}_{j=0}^{1-N_r} \in \mathbb{R}^{N_r \times 2n_h}$, seek $u_\sigma = \{u_h^k\}_{k=1}^{N_\tau} \in \mathbb{R}^{N_\tau \times 2n_h}$ such that $u_\sigma|_\Gamma = 0$ and for each $k = 1, \ldots, N_\tau$

$$\begin{cases} (\nu A_h + \tau^{-1} N_h + \varepsilon_p^{-1} B_h^\top B_h) u_h^k = C_h(u_h^{k-N_r}) u_h^{k-1} + \tau^{-1} N_h \widehat{u}_h^{k-1} + N_h f_h^k + N_h q_h^k \\ u_h^j = z_h^j \qquad j = 1 - N_r, \ldots, 0, \end{cases} \tag{5.1}$$

where $\widehat{u}_h^0 = u_{0h}$ and $\widehat{u}_h^{k-1} = u_h^{k-1}$ for $k > 1$. Here, the discrete incompressibility condition $B_h u_h = 0$ has been replaced by $B_h u_h + \varepsilon_p p_h^k = 0$. Note that the additional error induced from this penalization is of order $\mathcal{O}(\varepsilon_p)$, see [8, Section 4.3] for instance. In the numerical experiments below, we take the penalty parameter $\varepsilon_p = 10^{-10}$.

As long as the spatial mesh size h is sufficiently small, the matrix $\nu A_h + \tau^{-1} N_h + \varepsilon_p^{-1} B_h^\top B_h$ is symmetric and positive definite, hence the solvability of the linear system (5.1) at each time step is guaranteed. For details regarding this matter, we refer to [48] and [17, Theorem 4.1.2].

For the discrete adjoint equation, we consider the following discretization: Seek $w_\sigma = \{w_h^k\}_{k=1}^{N_\tau} \in \mathbb{R}^{N_\tau \times 2n_h}$ such that $w_\sigma|_\Gamma = 0$ and for each $k = N_\tau, \ldots, 1$

$$\begin{cases} (\nu A_h + \tau^{-1} N_h + \varepsilon_p^{-1} B_h^\top B_h) w_h^k = C_h(u_h^{k+1-N_r})^\top w_h^{k+1} + D_h(w_h^{k+N_r}) u_h^{k-1+N_r} \\ \qquad\qquad + \tau^{-1} N_h \widehat{w}_h^{k+1} + \alpha_{\Omega_T}(N_h u_h^k - N_h u_{dh}^k) + \alpha_R R_h u_h^k \\ w_h^j = 0 \qquad j = N_\tau + 1, \ldots, N_\tau + N_r, \end{cases} \tag{5.2}$$

where $\widehat{w}_h^{N_\tau+1} = \alpha_T(u_h^{N_\tau} - u_{Th})$ and $\widehat{w}_h^{k+1} = w_h^{k+1}$ for $k < N_\tau$. Note that (5.1) and (5.2) are the respective perturbed versions of the mixed problems (4.8) and the one corresponding to (4.62). The convection matrix $C_h(u_h^{k-N_r})$ has to be assembled at each time step in (5.1). We do not store these matrices for (5.2), but instead re-assemble them in addition to that of $D_h(w_h^{k+N_r})$. Therefore, efficient schemes are necessary in the construction of these matrices.

In terms of the above finite element matrices, the discrete cost functional can be computed using the box-rule as follow:

$$\begin{aligned} j_{\sigma,\varepsilon_p}(q_\sigma) := \frac{\alpha_{\Omega_T}\tau}{2} \sum_{k=1}^{N_\tau} (u_h^k - u_{dh}^k)^\top N_h (u_h^k - u_{dh}^k) + \frac{\alpha_T}{2}(u_h^{N_\tau} - u_{Th})^\top N_h (u_h^{N_\tau} - u_{Th}) \\ + \frac{\alpha_R \tau}{2} \sum_{k=1}^{N_\tau} u_h^{k\top} R_h u_h^k + \frac{\alpha\tau}{2} \sum_{k=1}^{N_\tau} q_h^{k\top} N_h q_h^k, \end{aligned}$$

where $u_\sigma = \{u_h^k\}_{k=1}^{N_\tau} \in \mathbb{R}^{N_\tau \times 2n_h}$ is the solution of (5.1). One can now formulate the penalization for (P_σ) as the finite-dimensional optimization problem:

$$\min_{q_\sigma \in \mathcal{P}_\tau(I, W_h)} j_{\sigma,\varepsilon_p}(q_\sigma). \tag{$\mathrm{P}_{\sigma,\varepsilon_p}$}$$

To seek for a solution to (P_{σ,ε_p}), the gradient method of Barzilai and Borwein (BB) in [6] will be utilized. The algorithm is terminated once there is no significant change in the cost values and that the derivative is close to zero. More precisely, if the relative error of the successive cost values and the gradient norm is less than a prescribed tolerance $0 < \epsilon_{\text{tol}} \ll 1$, that is, if the condition

$$\max\left\{ \frac{|j_{\sigma,\varepsilon_p}(q_\sigma^{(\ell)}) - j_{\sigma,\varepsilon_p}(q_\sigma^{(\ell-1)})|}{j_{\sigma,\varepsilon_p}(q_\sigma^{(\ell)})}, \|\alpha q_\sigma^{(\ell)} + w_\sigma^{(\ell)}\|_{L^2(I,X)} \right\} < \epsilon_{\text{tol}} \tag{5.3}$$

is satisfied, where $q_\sigma^{(\ell)}$ and $w_\sigma^{(\ell)}$ are the control and adjoint state at the ℓth iteration. In each of the experiments below, the control will be initialized to zero and the second point of the gradient method will be determined by steepest descent. In such a case, we look for solutions in a neighborhood of the null control. The analysis of the Barzilai-Borwein gradient method has been extended recently to the infinite-dimensional setting in [4] for strictly convex quadratic problems.

5.2. Experimental Order of Convergence

In this subsection we verify numerically the order of convergences presented in the previous section. Following the procedure in [43], we shall manufacture a reference numerical solution to the optimal control problem (P_{σ,ε_p}).

For the computational domain, we take the unit square $\Omega = (0,1) \times (0,1)$ and put $T = 1$, $r = 0.5$, $\nu = 1$, $\alpha_{\Omega_T} = \alpha_T = \alpha_R = 1$ and $\alpha = 10^{-1}$. Consider the functions

$$\begin{aligned} p(t,x_1,x_2) &= \sin(\pi t)(\cos(2\pi x_2) - \cos(2\pi x_1)) \\ u(t,x_1,x_2) &= \cos(\pi t)[(1-\cos(2\pi x_1))\sin(2\pi x_2),\ \sin(2\pi x_1)(\cos(2\pi x_2)-1)]^\top. \end{aligned} \tag{5.4}$$

We regard $u^\star_{\text{ref},\sigma} := \sum_{k=1}^{N_\tau} \mathit{\Pi}_h u(t_k)\mathbb{1}_{I_k}$ as our reference optimal state, with $p^\star_{\text{ref},\sigma} := \sum_{k=1}^{N_\tau} \mathit{\Pi}_h p(t_k)\mathbb{1}_{I_k}$ as the associated pressure. Here, $\mathit{\Pi}_h$ is the nodal Lagrange interpolation operator. For the history, we put $z = u$ in $I_r \times \Omega$ and is discretized by time-averaging and nodal Lagrange interpolation in space, see Sect. 4.2.

We consider $u_{d\sigma} = -u^\star_{\text{ref},\sigma}$ and $u_{Th} = -u^\star_{\text{ref},\sigma}(T)$ as the desired velocities. From these, the solution $w^\star_{\text{ref},\sigma}$ of (5.2) is computed and then $q^\star_{\text{ref},\sigma} = -\alpha^{-1}w^\star_{\text{ref},\sigma}$ is taken as the reference optimal control. In order for $(u^\star_{\text{ref},\sigma}, w^\star_{\text{ref},\sigma}, q^\star_{\text{ref},\sigma})$ to be a *solution* of (P_{σ,ε_p}), we set the forcing function $f_\sigma = \{f_h^k\}_{k=1}^{N_\tau} \in \mathbb{R}^{N_\tau \times 2n_h}$ in such a way that $u^\star_{\text{ref},\sigma}$ and $q^\star_{\text{ref},\sigma}$ satisfy the discrete state equation (5.1). To investigate the order of convergence, the step sizes $\sigma_k = (\tau_k, h_k) = (2^{-2k}\cdot 10^{-1}, 2^{1/2-k}\cdot 10^{-1})$ for $k = 0,1,2,3$ will be utilized. For these pairs of time steps and mesh sizes, we have $\tau_k = 5h_k^2$, so that the stability condition (A5) is satisfied.

The mesh generation, matrix assemblies, sparse linear solvers, and visualizations were done in Python 3.7.6 (Python Software Foundation, https://www.python.org/) on a 2.3 GHz Intel Core i5 with 8GB RAM. The repository containing the source codes as well as the iteration histories can be downloaded at https://github.com/grperalta/nsedelay. An LU factorization of the system matrix was obtained via the built-in-function splu in the SciPy package. In the factorization, a column permutation for sparsity preservation through a minimum degree ordering on the symmetric structure of the system matrix was used. The linear systems were solved using the UMFPACK option.

The order of convergences are presented in Table 1. For instance, in the case of the controls, we compute

$$\text{eoc}_k := \frac{\ln(\|q^\star_{\text{ref},\sigma_{k-1}} - q^\star_{\sigma_{k-1}}\|_{L^2(I,X)} / \|q^\star_{\text{ref},\sigma_k} - q^\star_{\sigma_k}\|_{L^2(I,X)})}{\ln(h_{k-1}/h_k)}, \qquad k = 1,2,3. \tag{5.5}$$

In the stopping criterion (5.3), we used $\epsilon_{\text{tol}} = 10^{-6}$.

As to be expected, the Taylor-Hood finite element performs better than the mini-element, however, at the expense of additional computing time. In the case of the mini-element, one can observe more or

Table 1. Experimental order of convergence (EOC) for the errors between the optimal solutions with reference to the norms of $L^2(I, L^2(\Omega)^2)$, $L^2(I, H_0^1(\Omega)^2)$ and $L^\infty(I, L^2(\Omega)^2)$ using the triangular mini-finite and Taylor-Hood elements

EOC with Mini-Finite (P1-bubble/P1) Element						
k	$\lVert u^\star_{\mathrm{ref},\sigma_k} - u^\star_{\sigma_k}\rVert_{L^2(I,X)}$	eoc_k	$\lVert w^\star_{\mathrm{ref},\sigma_k} - w^\star_{\sigma_k}\rVert_{L^2(I,X)}$	eoc_k	$\lVert q^\star_{\mathrm{ref},\sigma_k} - q^\star_{\sigma_k}\rVert_{L^2(I,X)}$	eoc_k
0	$5.509948 \cdot 10^{-2}$	–	$4.568061 \cdot 10^{-2}$	–	$3.816914 \cdot 10^{-1}$	–
1	$1.420196 \cdot 10^{-2}$	1.956	$1.081773 \cdot 10^{-2}$	2.078	$1.032727 \cdot 10^{-1}$	1.886
2	$3.570072 \cdot 10^{-3}$	1.992	$2.663922 \cdot 10^{-3}$	2.022	$2.633360 \cdot 10^{-2}$	1.971
3	$8.924269 \cdot 10^{-4}$	2.000	$6.625920 \cdot 10^{-4}$	2.007	$6.606980 \cdot 10^{-3}$	1.995
k	$\lVert u^\star_{\mathrm{ref},\sigma_k} - u^\star_{\sigma_k}\rVert_{L^2(I,W)}$	eoc_k	$\lVert w^\star_{\mathrm{ref},\sigma_k} - w^\star_{\sigma_k}\rVert_{L^2(I,W)}$	eoc_k	$\lVert q^\star_{\mathrm{ref},\sigma_k} - q^\star_{\sigma_k}\rVert_{L^2(I,W)}$	eoc_k
0	$1.815098 \cdot 10^{0}$	–	$1.098782 \cdot 10^{0}$	–	$7.395482 \cdot 10^{0}$	–
1	$9.192071 \cdot 10^{-1}$	0.982	$4.094915 \cdot 10^{-1}$	1.424	$3.541439 \cdot 10^{0}$	1.062
2	$4.610771 \cdot 10^{-1}$	0.995	$1.819815 \cdot 10^{-1}$	1.170	$1.745271 \cdot 10^{0}$	1.021
3	$2.307294 \cdot 10^{-1}$	0.999	$8.783346 \cdot 10^{-2}$	1.051	$8.688125 \cdot 10^{-1}$	1.006
k	$\lVert u^\star_{\mathrm{ref},\sigma_k} - u^\star_{\sigma_k}\rVert_{L^\infty(I,X)}$	eoc_k	$\lVert w^\star_{\mathrm{ref},\sigma_k} - w^\star_{\sigma_k}\rVert_{L^\infty(I,X)}$	eoc_k	$\lVert q^\star_{\mathrm{ref},\sigma_k} - q^\star_{\sigma_k}\rVert_{L^\infty(I,X)}$	eoc_k
0	$7.936219 \cdot 10^{-2}$	–	$7.936217 \cdot 10^{-2}$	–	$5.777529 \cdot 10^{-1}$	–
1	$2.036783 \cdot 10^{-2}$	1.962	$2.036772 \cdot 10^{-2}$	1.962	$1.661747 \cdot 10^{-1}$	1.798
2	$5.114179 \cdot 10^{-3}$	1.994	$5.090099 \cdot 10^{-3}$	2.001	$4.575560 \cdot 10^{-2}$	1.861
3	$1.277605 \cdot 10^{-3}$	2.001	$1.268801 \cdot 10^{-3}$	2.004	$1.201757 \cdot 10^{-2}$	1.929
EOC with Taylor-Hood (P2/P1) Element						
k	$\lVert u^\star_{\mathrm{ref},\sigma_k} - u^\star_{\sigma_k}\rVert_{L^2(I,X)}$	eoc_k	$\lVert w^\star_{\mathrm{ref},\sigma_k} - w^\star_{\sigma_k}\rVert_{L^2(I,X)}$	eoc_k	$\lVert q^\star_{\mathrm{ref},\sigma_k} - q^\star_{\sigma_k}\rVert_{L^2(I,X)}$	eoc_k
0	$7.163321 \cdot 10^{-3}$	–	$7.933576 \cdot 10^{-3}$	–	$7.231624 \cdot 10^{-2}$	–
1	$1.698009 \cdot 10^{-3}$	2.077	$1.779201 \cdot 10^{-3}$	2.156	$1.735902 \cdot 10^{-2}$	2.059
2	$4.188071 \cdot 10^{-4}$	2.019	$4.326775 \cdot 10^{-4}$	2.040	$4.300022 \cdot 10^{-3}$	2.013
3	$1.043778 \cdot 10^{-4}$	2.004	$1.074448 \cdot 10^{-4}$	2.010	$1.072851 \cdot 10^{-3}$	2.003
k	$\lVert u^\star_{\mathrm{ref},\sigma_k} - u^\star_{\sigma_k}\rVert_{L^2(I,W)}$	eoc_k	$\lVert w^\star_{\mathrm{ref},\sigma_k} - w^\star_{\sigma_k}\rVert_{L^2(I,W)}$	eoc_k	$\lVert q^\star_{\mathrm{ref},\sigma_k} - q^\star_{\sigma_k}\rVert_{L^2(I,W)}$	eoc_k
0	$5.932207 \cdot 10^{-2}$	–	$6.148310 \cdot 10^{-2}$	–	$5.529714 \cdot 10^{-1}$	–
1	$1.307225 \cdot 10^{-2}$	2.182	$1.334016 \cdot 10^{-2}$	2.204	$1.299808 \cdot 10^{-1}$	2.089
2	$3.130027 \cdot 10^{-3}$	2.062	$3.208756 \cdot 10^{-3}$	2.056	$3.188526 \cdot 10^{-2}$	2.027
3	$7.733619 \cdot 10^{-4}$	2.017	$7.944229 \cdot 10^{-4}$	2.014	$7.932280 \cdot 10^{-3}$	2.007
k	$\lVert u^\star_{\mathrm{ref},\sigma_k} - u^\star_{\sigma_k}\rVert_{L^\infty(I,X)}$	eoc_k	$\lVert w^\star_{\mathrm{ref},\sigma_k} - w^\star_{\sigma_k}\rVert_{L^\infty(I,X)}$	eoc_k	$\lVert q^\star_{\mathrm{ref},\sigma_k} - q^\star_{\sigma_k}\rVert_{L^\infty(I,X)}$	eoc_k
0	$1.031764 \cdot 10^{-2}$	–	$1.045768 \cdot 10^{-2}$	–	$1.045769 \cdot 10^{-1}$	–
1	$2.467746 \cdot 10^{-3}$	2.064	$2.493781 \cdot 10^{-3}$	2.068	$2.493852 \cdot 10^{-2}$	2.068
2	$6.088774 \cdot 10^{-4}$	2.019	$6.149193 \cdot 10^{-4}$	2.020	$6.148528 \cdot 10^{-3}$	2.020
3	$1.517279 \cdot 10^{-4}$	2.005	$1.531064 \cdot 10^{-4}$	2.006	$1.531007 \cdot 10^{-3}$	2.006

less a quadratic reduction with respect to the norms of $L^2(I, L^2(\Omega)^2)$ and $L^\infty(I, L^2(\Omega)^2)$, while a linear reduction in $L^2(I, H_0^1(\Omega)^2)$. For the Taylor-Hood element, we have a quadratic order in these spaces, which is better than the one predicted in Corollary 4.22, at least for this example. We also observe the mesh-independence of the gradient method, that is, the number of gradient iterations is independent on the considered spatial mesh sizes and temporal step sizes.

Let us discuss the order of convergence as the Tikhonov regularization parameter α decreases. For this example, we shall employ the Taylor-Hood finite element with the step size $\sigma_3 = (\tau_3, h_3) = (2^{-6} \cdot 10^{-1}, 2^{1/2-3} \cdot 10^{-1})$. Denote by $u^\star_{\mathrm{ref},\alpha_k}$, $w^\star_{\mathrm{ref},\alpha_k}$ and $q^\star_{\mathrm{ref},\alpha_k}$ the reference optimal state, adjoint state and control associated with the parameter $\alpha_k = 10^{-k}$ for $k = 0, 1, 2, 3$. The results are summarized in Table 2. Here, the required number of gradient iterations to reach the desired tolerance are given by 3, 6, 12 and 36. Observe that the reduction rate of the errors in the optimal state and optimal adjoint state are nearly the same. However, the errors for the optimal controls are increasing and we have $\|q^\star_{\mathrm{ref},\alpha_k} - q^\star_{\alpha_k}\| \approx \alpha_k^{-1}\|w^\star_{\mathrm{ref},\alpha_k} - w^\star_{\alpha_k}\|$ under the norms considered above. Nevertheless, the latter approximation is consistent with the optimality condition $\alpha q^\star + w^\star = 0$ relating the optimal control and the optimal adjoint, see Theorem 3.2. Further investigation is needed to obtain a precise representation of the order of convergence, or at least a suitable bound, for the optimal control as $\alpha \to 0$.

5.3. Velocity Tracking with Local Control

Let us consider the domain $\Omega = (0, 3) \times (0, 1)$ with the control region $\omega = (0.5, 2.5) \times (0.25, 0.75)$. In this situation, the control space is given by $L^2(I, L^2(\omega)^2)$ and the optimal control and adjoint state are related by $q^\star = -\alpha^{-1} w^\star \mathbb{1}_\omega$, where $\mathbb{1}_\omega$ is the indicator function on ω. For the parameters in the optimal control problem (P), we take $T = 1$, $r = 0.5$, $\nu = 0.01$, $\alpha_{\Omega_T} = \alpha_T = 1$, $\alpha_R = 0$ and $\alpha = 10^{-3}$. A uniform triangulation of Ω with 5124 nodes and 9922 triangles will be employed, with the corresponding mesh size $h \approx 0.034779$. Here, the Taylor-Hood finite element is implemented. For the temporal grid, the chosen step size is $\tau = 0.01$. With these discretizations, the degrees of freedom for the velocity field is $4033800 \approx \mathcal{O}(10^6)$.

The solution of the steady Stokes equation with artificial compressibility and a random source is taken as the initial data. More precisely, we take $u_{0h} = (\nu A_h + \varepsilon_p^{-1} B_h^\top B_h)^{-1} N_h f_h$, where $u_{0h}|_\Gamma = 0$ and $-10 \le f_h \le 10$ in (5.1), for which $\|u_{0h}\|_{L^\infty} \approx 0.775$. Also, $z = 0.5u$ is the initial history, where u is the function defined by (5.4). The solution of the uncontrolled Navier–Stokes equation without the delay in the convection term will be the chosen target state. To facilitate better performance, an alternating step length strategy was employed in the BB gradient algorithm, see [4,19] for instance. For the sake of the reader, the specific method utilized here is presented below. The gradient method converges, under the stopping criterion (5.3) with $\epsilon_{\mathrm{tol}} = 10^{-5}$, after 123 iterations with $j_{\sigma,\varepsilon_p}(q^\star_\sigma) \approx 3.646 \cdot 10^{-3}$ and $\|\alpha q^\star_\sigma + w^\star_\sigma\|_{L^2(I\times\omega)^2} \approx 6.580 \cdot 10^{-6}$. One can observe from Fig. 2b the non-monotone property of the BB gradient method. We observe a fast convergence initially, followed by little changes in the cost values, while the norm of the derivative still oscillates until it reached the required tolerance, see (b) and (d) of Fig. 2. This is a typical characteristic of gradient methods.

The optimal solution at $t = 0.1, 0.5, 1.0$ are given in Fig. 1. A quadric interpolation was rendered on the image data for better visualization. The magnitudes of the velocity field for the Navier–Stokes flow without delay, which is the target state, and with delay are depicted in parts (a) and (b) of Fig. 1, respectively. At $t = 0.1$, we somewhat have a turbulent flow from the random force for the Stokes flow, which is then stabilized due to viscosity. The formation of vortices at $t = 0.5$ and $t = 1.0$ in Fig. 1c is due to the profile of the initial history that acts as a convective force on the fluid.

Comparing (a) and (c) in Fig. 1, one can see that the optimal velocity nearly matches the target on the region where the control is applied. This is a common feature for tracking-type problems with local controls, for which the influence of the control is more significant on the region where it is applied or at least near to it. From Fig. 2a, we can see that from the start up to approximately before $t = 0.5$, the space-time L^2-error between the optimal and target velocities increases. We can also observe from

TABLE 2. Experimental order of convergence for the errors between the optimal solutions with decreasing Tikhonov regularization parameters $\alpha_k = 10^{-k}$ for $k = 0, 1, 2, 3$

EOC with Taylor-Hood (P2/P1) element for decreasing Tikhonov regularization						
k	$\lVert u^\star_{\text{ref},\alpha_k} - u^\star_{\alpha_k}\rVert_{L^2(I,X)}$	eoc_k	$\lVert w^\star_{\text{ref},\alpha_k} - w^\star_{\alpha_k}\rVert_{L^2(I,X)}$	eoc_k	$\lVert q^\star_{\text{ref},\alpha_k} - q^\star_{\alpha_k}\rVert_{L^2(I,X)}$	eoc_k
0	$1.218162 \cdot 10^{-4}$	–	$1.254202 \cdot 10^{-4}$	–	$1.252901 \cdot 10^{-4}$	–
1	$1.043778 \cdot 10^{-4}$	0.067	$1.074447 \cdot 10^{-4}$	0.067	$1.072851 \cdot 10^{-3}$	–
2	$4.377482 \cdot 10^{-5}$	0.377	$4.424767 \cdot 10^{-5}$	0.385	$4.421008 \cdot 10^{-3}$	–
3	$6.696397 \cdot 10^{-6}$	0.815	$6.721729 \cdot 10^{-6}$	0.818	$6.788698 \cdot 10^{-3}$	–
k	$\lVert u^\star_{\text{ref},\alpha_k} - u^\star_{\alpha_k}\rVert_{L^2(I,W)}$	eoc_k	$\lVert w^\star_{\text{ref},\alpha_k} - w^\star_{\alpha_k}\rVert_{L^2(I,W)}$	eoc_k	$\lVert q^\star_{\text{ref},\alpha_k} - q^\star_{\alpha_k}\rVert_{L^2(I,W)}$	eoc_k
0	$8.978485 \cdot 10^{-4}$	–	$9.230519 \cdot 10^{-4}$	–	$9.220613 \cdot 10^{-4}$	–
1	$7.733619 \cdot 10^{-4}$	0.065	$7.944229 \cdot 10^{-4}$	0.065	$7.932280 \cdot 10^{-3}$	–
2	$3.437169 \cdot 10^{-4}$	0.352	$3.442176 \cdot 10^{-4}$	0.363	$3.438689 \cdot 10^{-2}$	–
3	$8.627997 \cdot 10^{-5}$	0.600	$7.176810 \cdot 10^{-5}$	0.681	$7.214384 \cdot 10^{-2}$	–
k	$\lVert u^\star_{\text{ref},\alpha_k} - u^\star_{\alpha_k}\rVert_{L^\infty(I,X)}$	eoc_k	$\lVert w^\star_{\text{ref},\alpha_k} - w^\star_{\alpha_k}\rVert_{L^\infty(I,X)}$	eoc_k	$\lVert q^\star_{\text{ref},\alpha_k} - q^\star_{\alpha_k}\rVert_{L^\infty(I,X)}$	eoc_k
0	$1.772242 \cdot 10^{-4}$	–	$1.791175 \cdot 10^{-4}$	–	$1.791386 \cdot 10^{-4}$	–
1	$1.517279 \cdot 10^{-4}$	0.067	$1.531064 \cdot 10^{-4}$	0.068	$1.531007 \cdot 10^{-3}$	–
2	$6.408268 \cdot 10^{-5}$	0.374	$6.301082 \cdot 10^{-5}$	0.386	$6.308845 \cdot 10^{-3}$	–
3	$1.935223 \cdot 10^{-5}$	0.520	$1.021446 \cdot 10^{-5}$	0.790	$1.620508 \cdot 10^{-2}$	–

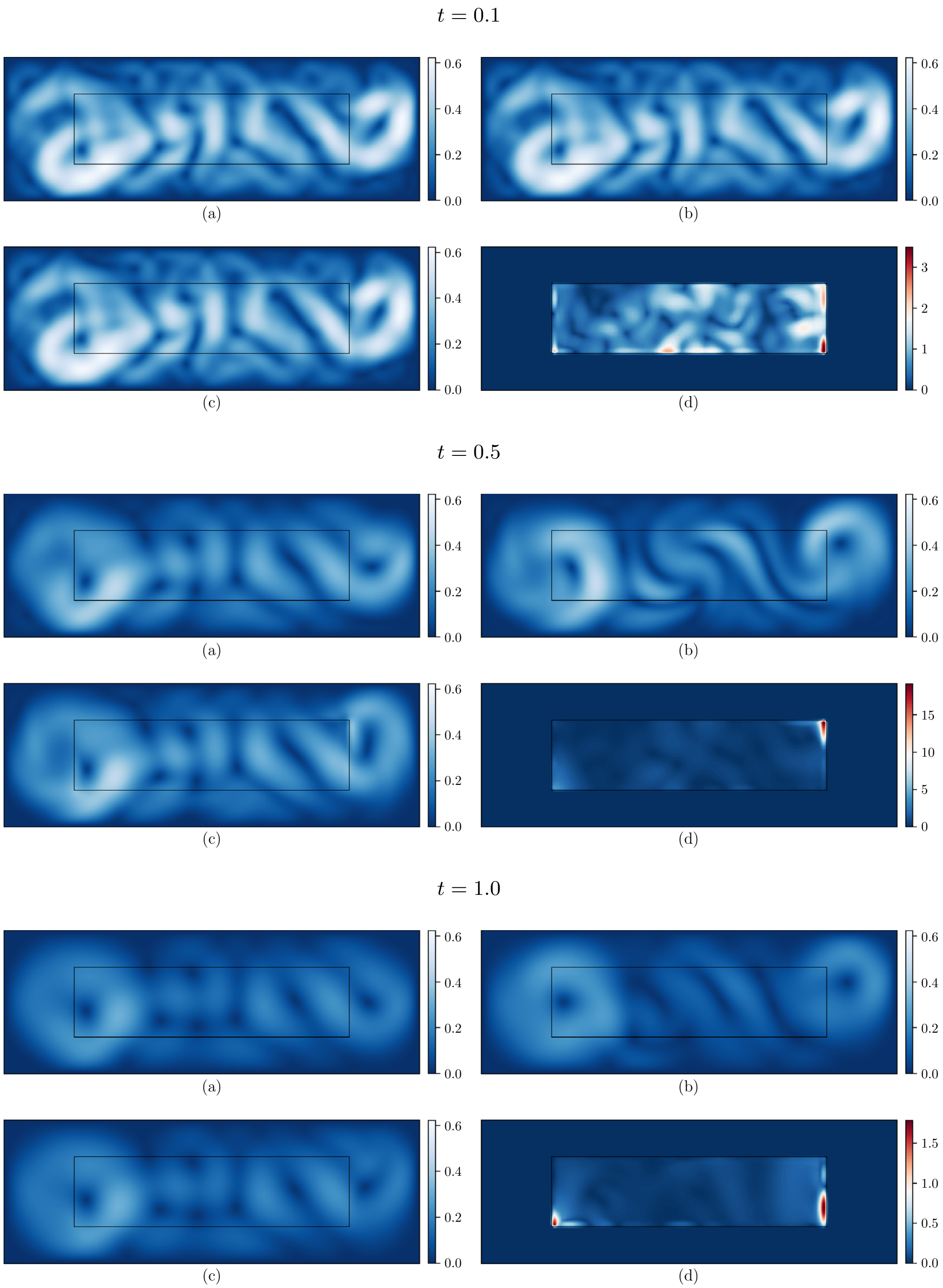

Fig. 1. Magnitudes of the target velocity (**a**), uncontrolled velocity (**b**), optimal velocity (**c**) and optimal control (**d**) at $t = 0.1$, $t = 0.5$ and $t = 1.0$ in the velocity-tracking problem. Bounding boxes represent the location of the control region ω

Algorithm: Barzilai-Borwein Gradient Method with Alternating Step Size Selection

Set $q_\sigma^{(0)} \leftarrow 0$ and compute $g_\sigma^{(0)} \leftarrow \alpha q_\sigma^{(0)} + w_\sigma^{(0)} \mathbb{1}_\omega$.
Put $q_\sigma^{(1)} \leftarrow q_\sigma^{(0)} - g_\sigma^{(0)}$, calculate $g_\sigma^{(1)} \leftarrow \alpha q_\sigma^{(1)} + w_\sigma^{(1)} \mathbb{1}_\omega$ and set $\ell \leftarrow 1$.
while $\max\{|j_{\sigma,\varepsilon_p}(q_\sigma^{(\ell)}) - j_{\sigma,\varepsilon_p}(q_\sigma^{(\ell-1)})|/j_{\sigma,\varepsilon_p}(q_\sigma^{(\ell)}), \|g_\sigma^{(\ell)}\|_{L^2(I\times\omega)^2}\} \geq \epsilon_{\text{tol}}$ **do**

Choose step size $s_\ell \leftarrow \begin{cases} (q_\sigma^{(\ell)} - q_\sigma^{(\ell-1)})^\top (g_\sigma^{(\ell)} - g_\sigma^{(\ell-1)})/|g_\sigma^{(\ell)} - g_\sigma^{(\ell-1)}|^2 & \text{if } \ell \text{ is odd,} \\ |q_\sigma^{(\ell)} - q_\sigma^{(\ell-1)}|^2/(q_\sigma^{(\ell)} - q_\sigma^{(\ell-1)})^\top (g_\sigma^{(\ell)} - g_\sigma^{(\ell-1)}) & \text{if } \ell \text{ is even.} \end{cases}$

Update the control $q_\sigma^{(\ell+1)} \leftarrow q_\sigma^{(\ell)} - s_\ell g_\sigma^{(\ell)}$.
Update the gradient $g_\sigma^{(\ell+1)} \leftarrow \alpha q_\sigma^{(\ell+1)} + w_\sigma^{(\ell+1)} \mathbb{1}_\omega$.
$\ell \leftarrow \ell + 1$

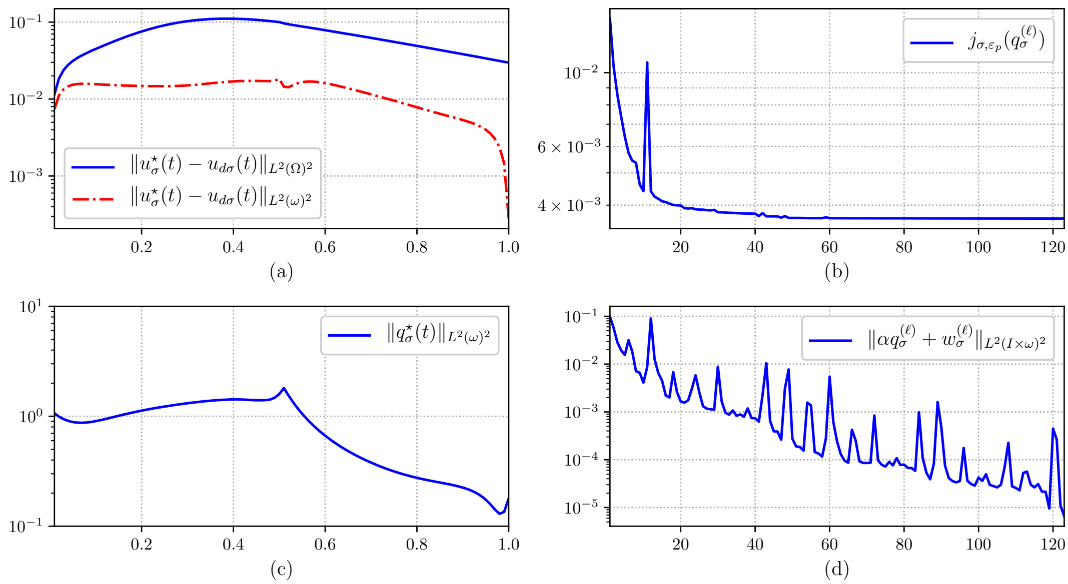

FIG. 2. The norms between the difference of the computed optimal and desired velocities in Ω and ω (**a**), and the computed optimal control (**c**) as functions of time. Parts (**b**) and (d) show the behavior of the cost values and gradient norms as functions of the number of iterations in the BB method

Fig. 2c that during this period the control is exerted in an increasing magnitude, which shows a very different control implementation as compared to the results in [32], where the authors considered the Navier–Stokes flow without delay.

After the time delay, the residual norm in the control region decreases, and the rate is faster near the final time. This trend can also be observed for the optimal control, see Fig. 2c, with the exception that there is an increase due to the tracking term at the final time in the objective functional. In this scenario, the flow is dominated by diffusion and convection has little effect. The magnitude of the control is relatively larger on the edges of ω, which is very natural if one wishes to steer the flow to a desired target that is outside of ω.

The results discussed above can be improved by choosing a smaller regularization parameter α. Larger magnitudes for the optimal control will be expected for this process. In general, this would require more gradient iterations. A higher resolution of the temporal and spatial mesh will also lead to better results, especially the tracking part outside of ω.

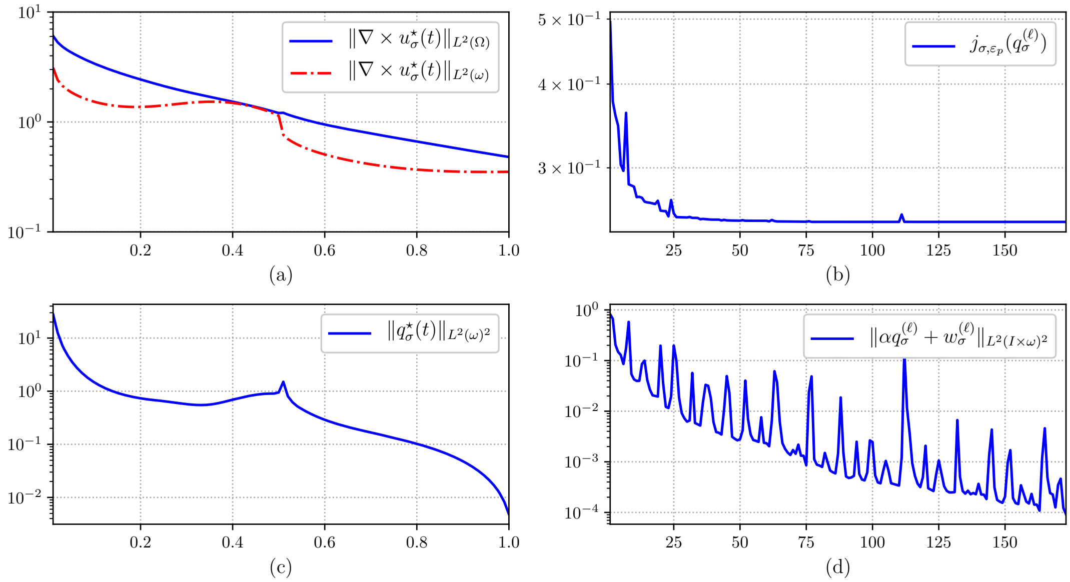

FIG. 3. The norm of the curl of computed optimal state in Ω and ω as functions of time (**a**) in the vorticity minimization problem. Here, **b**–**d** have the same descriptions as in Fig. 2

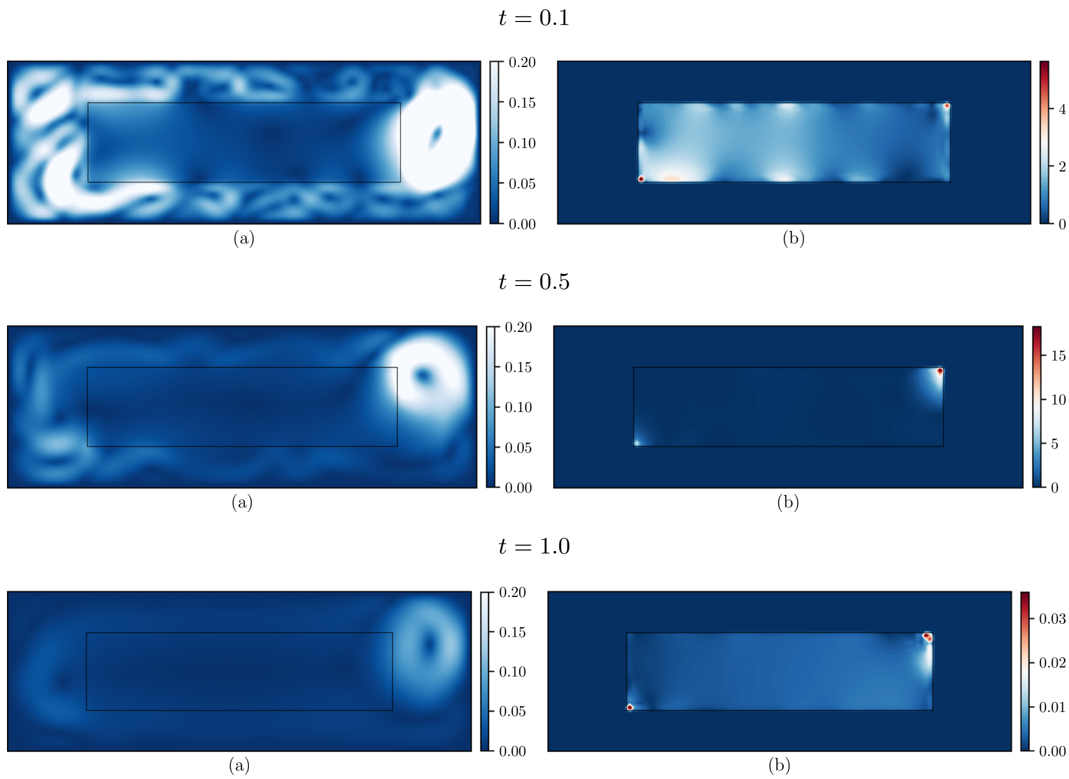

FIG. 4. Magnitudes of the optimal velocity (**a**) and optimal control (**b**) at $t = 0.1$, $t = 0.5$ and $t = 1.0$ in the vorticity minimization problem. Bounding boxes represent the location of the control region ω

5.4. Vorticity Minimization with Local Control

Let us consider the problem of minimizing the vorticity of the fluid flow. The set-up in the previous subsection will be used, but with the following modifications: $\alpha_{\Omega_T} = \alpha_T = 0$, $\alpha_R = 0.1$ and $\epsilon_{\mathrm{tol}} = 10^{-4}$. For this problem, the BB method terminated after 173 iterations for which $j_{\sigma,\varepsilon_p}(q_\sigma^\star) \approx 2.4911 \cdot 10^{-1}$ and $\|\alpha q_\sigma^\star + w_\sigma^\star\|_{L^2(I\times\omega)^2} \approx 9.2342 \cdot 10^{-5}$. Further information is provided in Fig. 3. Again, we have the non-monotonic property of the BB gradient algorithm. Notice also that the cost values and gradient norms in (b) and (d) behave analogously as in the previous subsection.

In Fig. 3c, one can see that the time-evolution for the norms of the optimal control share similar characteristics as in the velocity-tracking problem, that is, more effort is required for $t < r$. The behavior at $t = r$ can be attributed to the non-compatibility of the initial data and history.

Snapshots at $t = 0.1, 0.5, 1.0$ of the magnitudes of the computed optimal state and control are given in Fig. 4. At the early stages, the control forces the flow inside ω to stabilize, while creating vortices surrounding it. Most of the activity of the flow is then outside of the control region. In order to minimize the vorticity, the control needs to exert more work near the boundary of ω until the flow outside is dissipated.

6. Conclusion

We showed the existence and regularity of solutions to a distributed optimal control problem for the 2D Navier–Stokes equation with delay in the convection. A full discretization of the control problem based on the discontinuous Galerkin method and mixed finite elements has been studied, and optimal convergence rates were established using duality arguments. Finally, numerical examples were provided to validate the order of convergence and to demonstrate the effectiveness of the theoretical results.

Further analysis is needed to understand the behavior of the optimal solutions as the Tikhonov regularization parameter $\alpha \to 0$. Mixed methods based on quadrilateral finite elements and/or higher-order time advancing schemes through multi-step methods are possible extensions of the numerical scheme proposed in this paper. Under suitable regularity conditions on the domain, initial data, and initial history, together with an appropriate stability condition for the spatial and temporal mesh sizes analogous to (A5), higher convergence rates may be obtained. For instance, the biquadratic-linear (Q2/P1) velocity-pressure element in [30,47] could lead to an optimal convergence rate $\mathcal{O}(h^3)$ with respect to the space-time L^2-norm.

Acknowledgements. This work was supported in part by the One U.P. Faculty Grant 2019-101374 for G.P. and by the Philippine Department of Science and Technology - Accelerated Science and Technology Human Resource Development Program (DOST-ASTHRDP) for J.S.S. The authors are grateful to the anonymous referees for the comments and suggestions that lead to the improvement of the manuscript.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

7. Appendix

In the following, an extension of the Gronwall Lemma that is needed in the analysis of the state and linearized state equations is presented. The reader is reminded on the notation of various time intervals in (2.2).

Lemma 7.1. *Suppose that $a \geq 0$, $\phi \in L^\infty(I_r) \cap L^1(I)$, $\varphi \in L^1(J_r)$, α, β, $\psi \in L^1(I)$, and $\gamma \in L^\infty(I)$ are nonnegative and for a.e. $t \in I$ it holds that*

$$\phi(t) + \int_0^t \varphi(s)\,\mathrm{d}s \leq a + \int_0^t \alpha(s)\phi(s) + \beta(s)\phi^r(s) + \gamma(s)\varphi^r(s) + \psi(s)\,\mathrm{d}s. \tag{7.1}$$

Then $\phi \in L^\infty(I)$ and there exists a continuous function $\mathfrak{c} > 0$ such that

$$\|\phi\|_{L^\infty(I)} + \|\varphi\|_{L^1(I)} \leq \mathfrak{c}_{T,r,\alpha,\beta,\gamma}(a + \|\phi\|_{L^\infty(I_r)} + \|\varphi\|_{L^1(I_r)} + \|\psi\|_{L^1(I)})$$

where $\mathfrak{c}_{T,r,\alpha,\beta,\gamma} := \mathfrak{c}(T, r, \|\alpha\|_{L^1(I)}, \|\beta\|_{L^1(I)}, \|\gamma\|_{L^\infty(I)})$.

Proof. Let N be the largest positive integer such that $(N-1)r < T \leq Nr$, and set $I_n := [0, nr]$ for $n = 1, \dots, N$. For each n, we shall demonstrate by induction that

$$\|\phi\|_{L^\infty(I_n)} + \|\phi\|_{L^1(I_n)} \leq \mathfrak{c}_{T,r,\alpha,\beta,\gamma}(a + \|\phi\|_{L^\infty(I_r)} + \|\varphi\|_{L^1(I_r)} + \|\psi\|_{L^1(I_n)}). \tag{7.2}$$

Let us verify this for $n = 1$. Using the assumption (7.1) restricted to $t \in I_1$, we can apply the usual Gronwall Lemma so that

$$\|\phi\|_{L^\infty(I_1)} \leq (a + \|\beta\|_{L^1(I)}\|\phi\|_{L^\infty(I_r)} + \|\gamma\|_{L^\infty(I)}\|\varphi\|_{L^1(I_r)} + \|\psi\|_{L^1(I_1)})e^{\|\alpha\|_{L^1(I)}}. \tag{7.3}$$

On the other hand, (7.1) also yields the following estimate

$$\begin{aligned}\|\varphi\|_{L^1(I_1)} \leq a &+ \|\alpha\|_{L^1(I)}\|\phi\|_{L^\infty(I_1)} + \|\beta\|_{L^1(I)}\|\phi\|_{L^\infty(I_r)} \\ &+ \|\gamma\|_{L^\infty(I)}\|\varphi\|_{L^1(I_r)} + \|\psi\|_{L^1(I_1)}.\end{aligned} \tag{7.4}$$

Substituting (7.3) in the second term of the right hand side in (7.4) and then adding the resulting inequality with (7.3) prove (7.2) for $n = 1$.

Now, suppose that (7.2) holds for $n = k$. For $t \in I_{k+1}$, we obtain from (7.1) that

$$\begin{aligned}\phi(t) + \int_0^t \varphi(s)\,\mathrm{d}s \leq a &+ \|\beta\|_{L^1(I)} \max\{\|\phi\|_{L^\infty(I_k)}, \|\phi\|_{L^\infty(I_r)}\} \\ &+ \|\gamma\|_{L^\infty(I)}(\|\varphi\|_{L^1(I_k)} + \|\varphi\|_{L^1(I_r)}) + \int_0^t \alpha(s)\phi(s) + \psi(s)\,\mathrm{d}s.\end{aligned}$$

Thus, applying the Gronwall Lemma once more, one has the estimate

$$\begin{aligned}\|\phi\|_{L^\infty(I_{k+1})} \leq (a &+ \|\beta\|_{L^1(I)} \max\{\|\phi\|_{L^\infty(I_k)}, \|\phi\|_{L^\infty(I_r)}\} \\ &+ \|\gamma\|_{L^\infty(I)}(\|\varphi\|_{L^1(I_k)} + \|\varphi\|_{L^1(I_r)}) + \|\psi\|_{L^1(I_{k+1})})e^{\|\alpha\|_{L^1(I)}}\end{aligned}$$

and as a consequence it follows that

$$\begin{aligned}\|\varphi\|_{L^1(I_{k+1})} \leq a &+ \|\beta\|_{L^1(I)} \max\{\|\phi\|_{L^\infty(I_k)}, \|\phi\|_{L^\infty(I_r)}\} \\ &+ \|\gamma\|_{L^\infty(I)}(\|\varphi\|_{L^1(I_k)} + \|\varphi\|_{L^1(I_r)}) + \|\alpha\|_{L^1(I)}\|\phi\|_{L^\infty(I_{k+1})} + \|\psi\|_{L^1(I_{k+1})}.\end{aligned}$$

The last two inequalities along with the induction hypothesis imply (7.2) for $n = k + 1$. This completes the proof of the induction step. □

Next, we recall the following discrete version of the Gronwall Lemma, see [34] for instance. This is utilized in the error analysis of the fully-discrete optimal control problem.

Lemma 7.2. *Let $n \in \mathbb{N}$, $a \geq 0$, $\{a_k\}_{k=1}^n$, $\{b_k\}_{k=1}^n$, and $\{c_k\}_{k=1}^{n-1}$ be nonnegative sequences with*

$$a_j + \sum_{k=1}^{j} b_k \leq a + \sum_{k=1}^{j-1} c_k a_k \quad \textit{for all } j = 1, \dots, n.$$

Then it holds that

$$\max_{1 \leq k \leq l} a_k + \sum_{k=1}^{l} b_k \leq a \exp\Big(\sum_{k=1}^{l-1} c_k\Big) \quad \textit{for all } l = 1, \dots, n.$$

References

[1] Abergel, F., Temam, R.: On some control problems in fluid mechanics. Theor. Comput. Fluid Dyn. **1**, 303–325 (1990). https://doi.org/10.1007/BF00271794

[2] Adams, R.A.: Sobolev Spaces. Academic Press, New York (1975)

[3] Arnold, D.N., Brezzi, F., Fortin, M.: A stable finite element for the Stokes equation. Calcolo **21**, 337–344 (1984). https://doi.org/10.1007/BF02576171

[4] Azmi, B., Kunisch, K.: Analysis and performance of the Barzilai–Borwein step-size rules for optimization problems in Hilbert spaces. J. Optim. Theory Appl. **185**, 819–844 (2020)

[5] Banks, H.T., Burns, J.A.: Hereditary control problems: numerical methods based on averaging approximations. SIAM J. Control Optim. **16**, 169–208 (1978). https://doi.org/10.1137/0316013

[6] Barzilai, J., Borwein, J.: Two-point step size gradient method. IMA J. Numer. Anal. **8**(1), 141–148 (1988). https://doi.org/10.1093/imanum/8.1.141

[7] Bessaih, H., Garrido-Atienza, M.J., Schmalfuß, B.: On 3d Navier-Stokes equations: regularization and uniqueness by delays. Physica D Nonlinear Phenomena **376**–**377**, 228–237 (2018). https://doi.org/10.1016/j.physd.2018.03.004

[8] Boffi, D., Brezzi, F., Fortin, M.: Mixed Finite Element Methods and Applications, Vol. 44 of Springer Series in Computational Mathematics. Springer, Heidelberg (2013)

[9] Caraballo, T., Real, J.: Navier–Stokes equations with delays. Proc. R. Soc. Lond. A **457**, 2441–2453 (2001). https://doi.org/10.1098/rspa.2001.0807

[10] Casas, E.: Optimality conditions for some control problems of turbulent flow, In: Gunzburger M.D. (eds) Flow Control. The IMA Volumes in Mathematics and its Applications, Vol 68. Springer, New York, NY, pp. 127–147 (1995). https://doi.org/10.1007/978-1-4612-2526-3_6

[11] Casas, E., Chrysafinos, K.: A discontinuous Galerkin time-stepping scheme for the velocity tracking problem. SIAM J. Numer. Anal. **50**(5), 2281–2306 (2012). https://doi.org/10.1137/110829404

[12] Casas, E., Mateos, M., Raymond, J.-P.: Error estimates for the numerical approximation of a distributed control problem for the steady-state Navier–Stokes equations. SIAM J. Control Optim. **46**(3), 952–982 (2007). https://doi.org/10.1137/060649999

[13] Chrysafinos, K.: Analysis and finite element error estimates for the velocity tracking problem for Stokes flows via penalized formulation. ESAIM Control Optim. Calc. Var. **10**(4), 574–592 (2004). https://doi.org/10.1051/cocv:2004021

[14] Chrysafinos, K.: Error estimates for the velocity tracking problem for Navier–Stokes flows based on the artificial compressibility formulation. Numer. Func. Anal. Opt. **26**(7–8), 773–812 (2005). https://doi.org/10.1080/01630560500377311

[15] Chrysafinos, K.: Error estimates for time-discretizations for the velocity tracking problem for the Navier–Stokes flows by penalty methods. Discrete Cont. Dyn. B **6**(5), 1077–1096 (2006). https://doi.org/10.3934/dcdsb.2006.6.1077

[16] Chrysafinos, K., Walkington, N.J.: Discontinuous Galerkin approximations of the Stokes and Navier–Stokes equations. Math. Comput. **79**, 2135–2167 (2010). https://doi.org/10.1090/S0025-5718-10-02348-3

[17] Ciarlet, P. G.: The Finite Element Method for Elliptic Problems, Vol. 4 of Studies in Mathematics and Its Applications. North-Holland, Amsterdam (1978)

[18] Constantin, P., Foias, C.: Navier–Stokes Equations. Chicago Lectures in Mathematics. The University of Chicago Press, Chicago (1988)

[19] Dai, Y.-H., Fletcher, R.: Projected Barzilai–Borwein methods for large-scale box constrained quadratic programming. Numer. Math. **100**, 21–47 (2005). https://doi.org/10.1007/s00211-004-0569-y

[20] Deckelnick, K., Hinze, M.: Error estimates in space and time for tracking-type control of the instationary Stokes system. In: Desch, W., Kappel, F., Kunisch, K. (eds.) Control and Estimation of Distributed Parameter Systems, pp. 87–103. Birkhäuser, Basel (2003). https://doi.org/10.1007/978-3-0348-8001-5_6

[21] Deckelnick, K., Hinze, M.: Semidiscretization and error estimates for distributed control of the instationary Navier-Stokes equations. Numer. Math. **97**, 297–320 (2004). https://doi.org/10.1007/s00211-003-0507-4

[22] Dunford, N., Schwartz, J.Y.: Linear Operators, vol. I. General Theory. Intersci. Publ, New York (1958)

[23] Ern, A., Guermond, J..-L.: Theory and Practice of Finite Elements. Springer, New York (2004). https://doi.org/10.1007/978-1-4757-4355-5

[24] Foias, C., Manley, O., Rosa, R., Temam, R.: Navier–Stokes Equations and Turbulence. Cambridge University Press, Encyclopedia of Mathematics and its Applications (2001)

[25] Fursikov, A.V.: On some control problems and results related to the unique solution of mixed problems by the three-dimensional Navier–Stokes and Euler equations. Dokl. Akad. Nauk SSSR **252**(5), 1066–1070 (1980)

[26] Fursikov, A.V.: Control problems and theorems concerning on the unique solvability of a mixed boundary value problem for the three-dimensional Navier-Stokes and Euler equations. Mat. USSR-Sb. **43**(2), 251–273 (1982). https://doi.org/10.1070/sm1982v043n02abeh002447

[27] Fursikov, A.V.: Properties of solutions of certain extremum problems related to the Navier–Stokes and Euler equations. Mat. USSR-Sb. **46**(3), 323–351 (1983). https://doi.org/10.1070/sm1983v046n03abeh002796

[28] Gao, X., Gao, H.: Existence and uniqueness of weak solutions to stochastic 3d Navier–Stokes equations with delays. Appl. Math. Lett. **95**, 158–164 (2019). https://doi.org/10.1016/j.aml.2019.03.037

[29] Guzzo, S.M., Planas, G.: On a class of three dimensional Navier-Stokes equations with bounded delay. Discrete Contin. Dyn. Syst. Ser. B **16**(1), 225–238 (2011). https://doi.org/10.3934/dcdsb.2011.16.225

[30] Girault, V., Raviart, P..-A.: Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms, Vol. 5 of Springer Series in Computational Mathematics. Springer, Berlin (1986). https://doi.org/10.1007/978-3-642-61623-5
[31] Gunzburger, M.D., Maservisi, S.: The velocity tracking problem for Navier–Stokes flows with bounded distributed controls. SIAM J. Control Optim. **37**(6), 1913–1945 (1999). https://doi.org/10.1137/S0363012998337400
[32] Gunzburger, M.D., Maservisi, S.: Analysis and approximation of the velocity tracking problem for Navier-Stokes flows with distributed control. SIAM J. Numer. Anal. **37**(5), 1481–1512 (2000). https://doi.org/10.1137/S0036142997329414
[33] He, Y.: Two-level method based on finite element and Crank–Nicolson extrapolation for the time-dependent Navier–Stokes equations. SIAM J. Numer. Anal. **41**(4), 1263–1285 (2003). https://doi.org/10.1137/S0036142901385659
[34] He, Y., Li, K.: Convergence and stability of finite element nonlinear Galerkin method for the Navier–Stokes equations. Numer. Math. **79**, 77–106 (1998). https://doi.org/10.1007/s002110050332
[35] He, Y., Li, K.: Nonlinear Galerkin method and two-step method for the Navier–Stokes equations. Numer. Methods Partial Differ. Equ. **12**(3), 283–305 (1996)
[36] He, Y., Sun, W.: Stability and convergence of the Crank–Nicolson/Adams–Bashforth scheme for the time-dependent Navier–Stokes equations. SIAM J. Numer. Anal. **45**(2), 837–869 (2007). https://doi.org/10.1137/050639910
[37] Heywood, J.G., Rannacher, R.: Finite-element approximations of the nonstationary Navier–Stokes problem. Part IV: error estimates for second-order time discretization. SIAM J. Numer. Anal. **27**(2), 353–384 (1990). https://doi.org/10.1137/0727022
[38] Hinze, M., Kunisch, K.: Second order methods for optimal control of time-dependent fluid flow. SIAM J. Control Optim. **40**(3), 925–946 (2001). https://doi.org/10.1137/S0363012999361810
[39] Kellog, R.B., Osborn, J.E.: A regularity result for the Stokes problem in a convex polygon. J. Func. Anal. **21**(4), 397–431 (1976). https://doi.org/10.1016/0022-1236(76)90035-5
[40] Lions, J.L.: Optimal Control of Systems Governed by Partial Differential Equations. Grundlehren der mathematischen Wissenschaften, vol. 170. Springer, New York (1971)
[41] Liu, W.: Asymptotic behavior of solutions of time-delayed Burgers' equation. DCDS Ser. B **2**(1), 47–56 (2002). https://doi.org/10.3934/dcdsb.2002.2.47
[42] Marion, M., Temam, R.: Navier–Stokes Equations: Theory and Approximation, Handb. Numer. Anal. VI, North-Holland, Amsterdam, pp. 503–688 (1998). https://doi.org/10.1016/S1570-8659(98)80010-0
[43] Peralta, G., Kunisch, K.: Analysis and finite element discretization for optimal control of a linear fluid-structure interaction problem with delay. IMA J. Numer. Anal. **40**(1), 140–206 (2020). https://doi.org/10.1093/imanum/dry070
[44] Quarteroni, A., Valli, A.: Numerical Approximations of Partial Differential Equations. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-85268-1
[45] Roubíček, T.: Nonlinear Partial Differential Equations with Applications, 2nd edition. Vol. 23 of Springer Series in Computational Mathematics, Birkhäuser, Basel (2013). https://doi.org/10.1007/978-3-0348-0513-1
[46] Simon, J.: Compact sets in $L^p(0,T;B)$. Ann. Mat. Pur. Appl. **146**, 65–96 (1986). https://doi.org/10.1007/BF01762360
[47] Stenberg, R.: Analysis of mixed finite element methods for the Stokes problem: a unified approach. Math. Comput. **42**, 9–23 (1984). https://doi.org/10.1090/S0025-5718-1984-0725982-9
[48] Strang, G., Fix, G.: An Analysis of the Finite Element Method. Prentice-Hall Series in Automatic Computation, Prentice-Hall, Englewood Cliffs, NJ (1973)
[49] Temam, R.: Navier–Stokes Equations: Theory and Numerical Analysis, 2nd edn. North Holland, Amsterdam (1979)
[50] Tröltzsch, F.: Optimal Control of Partial Differential Equations: Theory, Methods and Applications. Graduate Studies in Mathematics, vol. 112. American Mathematical Society, Providence, RI (2010)
[51] Varnhorn, W.: Time delay and finite differences for the non-stationary non-linear Navier–Stokes equations. Math. Methods Appl. Sci. **15**(2), 89–108 (1992). https://doi.org/10.1002/mma.1670150204
[52] Varnhorn, W.: A fractional step method for regularized Navier-Stokes equations, In: J.G. Heywood, K. Masuda, R. Rautmann R. and V.A Solonnikov V.A. (eds) The Navier-Stokes Equations II: Theory and Numerical Methods, Lecture Notes in Mathematics, vol 1530, Springer, Berlin, Heidelberg (1992). https://doi.org/10.1007/BFb0090343
[53] Wachsmuth, D.: Optimal control of the unsteady Navier-Stokes equations. PhD Thesis, Technischen Universität Berlin (2006). https://depositonce.tu-berlin.de/bitstream/11303/1767/1/Dokument_27.pdf
[54] Zeidler, E.: Applied Functional Analysis and its Applications I: Fixed Point Theorems. Springer, New York (1986)

Gilbert Peralta and John Sebastian Simon
Department of Mathematics and Computer Science
University of the Philippines Baguio
Governor Pack Road
Baguio
Philippines
e-mail: grperalta@up.edu.ph

John Sebastian Simon
Division of Mathematical and Physical Sciences Graduate
School of Natural Science and Technology
Kanazawa University
Kakuma Kanazawa
Japan
e-mail: john.simon@stu.kanazawa-u.ac.jp

(accepted: March 21, 2021)