Analysis and finite element discretization for optimal control of a linear fluid–structure interaction problem with delay

GILBERT PERALTA*
Department of Mathematics and Computer Science, University of the Philippines Baguio,
Governor Pack Road, Baguio 2600, Philippines
*Corresponding author: grperalta@up.edu.ph

AND

KARL KUNISCH
Institute for Mathematics and Scientific Computing, University of Graz, Heinrichstrasse 36,
Graz 8010, Austria and RICAM Institute, Austrian Academy of Sciences, Altenbergerstrasse 69,
Linz 4040, Austria
karl.kunisch@uni-graz.at

[Received on 18 October 2017; revised on 21 August 2018]

An optimal control problem for a linearized fluid–structure interaction model with a delay term in the structural damping is analyzed. A distributed control acting on the fluid domain, structure domain or both is considered. The necessary optimality conditions are derived both for rough and smooth initial data. A parabolic regularization of the problem and its convergence are investigated. Finite element discretization for the regularized problem and error estimates are provided. Piecewise linear elements with bubble functions for the fluid and a discontinuous Galerkin scheme for the spatial and temporal discretizations are utilized respectively. Numerical experiments illustrating the theoretical results are given.

Keywords: fluid–structure interaction; delay; optimal control; space-time finite elements; error estimates.

1. Introduction

This paper is dedicated to the analysis and numerical approximation of optimal control problems for linearized models describing the interaction of an incompressible fluid and a structure. For the physical configuration we consider the setting where the structure is completely submerged in the fluid domain. An example of this situation is microbubble suspension in a fluid in ultrasound imaging (Dayton et al., 2002). In this study we assume that the interface between the fluid and the solid is static, a reasonable assumption in the case of small yet rapid oscillations for the structure. Although this assumption is somewhat restricted and limits applicability to more realistic problems, the current work is a contribution toward nonlinear fluid–structure interaction (FSI) models.

FSI models have been studied in past years from both analytical and computational perspectives. These works deal with the well-posedness for linear problems (Du et al., 2003; Avalos & Trigianni, 2007, 2009b), nonlinear problems (Barbu et al., 2007; Kukavica et al., 2011; Kukavica & Tuffaha, 2012; Ignatova et al., 2014, 2017), asymptotic stability (Lasiecka & Lu, 2011), interior feedback stabilization (Avalos & Trigianni, 2013; Lu, 2013a,b), boundary feedback stabilization...
(Avalos & Trigianni, 2008, 2009a, 2013; Lasiecka & Lu, 2012), regularity (Avalos et al., 2008; Barbu et al., 2008), numerical analysis and approximations (Causin et al., 2005; Burman & Fernandez, 2009; Guidoboni et al., 2009; Astronio & Grandmont, 2010; Lukáˇcova-Medvid’ová et al., 2013; Avalos & Toundykov, 2016), Bolza control problems (Lasiecka & Tuffaha, 2009a,b) and optimal control (Failer et al., 2016) to name a few. Specifically, the authors in Failer et al. (2016) considered an optimal control problem for an unsteady linear FSI problem and derived optimality conditions based on the adjoint equations of a symmetric formulation of the state equations. This strategy is advantageous in the implementation of gradient-based optimization algorithms.

We denote by \( \Omega_f \), \( \Omega_s \), \( \Gamma_s \) and \( \Gamma_f \) the domain occupied by the fluid, the solid, the interface between the two and the boundary of the wall for the fluid, respectively. Since the structure lies entirely in the fluid, \( \Gamma_f \) and \( \Gamma_s \) have no points in common. The entire domain of the interaction model will be denoted by \( \Omega = \Omega_f \cup \Omega_s \cup \Gamma_s \). We suppose that \( \Omega \subset \mathbb{R}^2 \) is a polygonal domain and \( \Omega_s \) is a sufficiently smooth domain. The optimal control problem we are interested in is

\[
\min_{q \in Q} J(u, w, q) = G(u, w) + \frac{\alpha}{2} \|q\|_Q^2
\]  

subject to state equation

\[
\begin{align*}
  u_t - \Delta u + \nabla p &= B_f q & \text{in } Q_f = I \times \Omega_f, \\
  \text{div } u &= 0 & \text{in } Q_f, \\
  u &= 0 & \text{on } \Sigma_f = I \times \Gamma_f, \\
  u &= w_t & \text{on } \Sigma_s = I \times \Gamma_s, \\
  w_{tt} - \Delta w + w + \mu w_t (\cdot - r) &= B_s q & \text{in } Q_s = I \times \Omega_s, \\
  \partial_v w - \partial_v u + pv &= 0 & \text{on } \Sigma_s, \\
  u(0) &= u_0 & \text{in } \Omega_f, \\
  w(0) &= w_0, \quad w_t(0) = v_0 & \text{in } \Omega_s, \\
  w_t &= z_0 & \text{in } Q_r = (-r,0) \times \Omega_s,
\end{align*}
\]  

where \( I = (0, T) \), \( B_f : Q \to L^2(Q_f) \), \( B_s : Q \to L^2(Q_s) \), \( \mu \in \mathbb{R} \), \( r > 0 \), \( \alpha > 0 \), and \( T > 0 \) is a given time horizon. Furthermore, \( Q \) denotes the Hilbert space of controls. Here \( v \) is the unit normal on \( \Gamma_s \) outward to \( \Omega_s \), hence will be inward with respect to \( \Omega_f \), and \( \partial_v \) denotes the normal derivative. For simplicity of exposition, the fluid viscosity, fluid density and structure density are normalized to 1.

The state equation (1.2) will be understood in the weak sense, which will be specified concretely in the succeeding section. In this equation, \( u : Q_f \to \mathbb{R}^2 \), \( p : Q_f \to \mathbb{R} \) and \( w : Q_s \to \mathbb{R}^2 \) represent the fluid velocity field, the pressure in the fluid and the structural displacement, respectively. We have a no-slip
boundary condition for the fluid on the wall $\Gamma_f$ and the continuity of velocities and normal stresses for the fluid and structure on the interface $\Gamma_s$. The retarded term $\mu w_t(-r)$ in (1.2) represents a delay term in the structural damping which, from a physical point of view, may occur due to material properties of the structure. The constants $\mu$ and $r$ represent the strength and extent of the delay, respectively. The stability of system (1.2) without delay and similar versions of it have been studied in Avalos & Trigianni (2013) with internal mechanical dissipation and in Avalos & Trigianni (2008, 2009a) with interface mechanical dissipation and zero internal static damping. For the stability of nonlinear FSI models without delay we refer to Lasiecka & Lu (2011, 2012); Lu (2013a,b).

Well-posedness and stability of a linear FSI model with delay have been considered in Peralta (2016). We reiterate here that due to the transport phenomenon induced by the delay term, oscillations occur that may result in instability. This has already been observed in the case of wave equations; see for instance Datko (1988); Datko et al. (1985); Nicaise & Pignotti (2006). In particular, the presence of delays may lead to solutions that have either constant or increasing energy as time progresses. Optimal control then serves as a useful tool to stabilize the system by minimizing its energy. In this work we shall consider the finite horizon case. Infinite time horizon problems will be a topic for future work. For optimal control of parabolic problems with delay the necessary optimality conditions were discussed in Lions (1969, Section 18.1).

With regards to the function $G$ appearing in the cost $J$ we consider a functional keeping track of the total or a part of the energy of the system, namely

$$G(u, w) = \frac{1}{2} \int_I \left( \gamma_f \|u - u_d\|^2_{\Omega_f} + \gamma_s1 \|w_t - v_d\|^2_{\Omega_s} + \gamma_s2 \|w - w_d\|^2_{\Omega_s} + \gamma_s3 \|
abla w - \nabla w_d\|^2_{\Omega_s} \right) dt,$$  \hspace{1cm} (1.3)

for given desired velocity fields $u_d$ and $v_d$ and displacement $w_d$, where $\gamma_f, \gamma_{si} \geq 0$ for $i = 1, 2, 3$. Different treatments in the analysis for $\gamma_s3 > 0$ and $\gamma_s3 = 0$ will be needed. This is reflected in the regularity requirements on the source terms appearing in the fluid and structure equations. If one requires that the energy on the time interval $I$ be minimized then we just take the desired states to be zero. The consideration for the cost functional (1.3) is motivated in the context where $T$ is large, which relates to stabilization.

We would like to point out that the above problem has been studied in Failer et al. (2016) without the retarded term, with $\gamma_s1 = \gamma_s3 = 0$ and under smooth initial data satisfying appropriate compatibility conditions. In this work we shall study the theoretical aspects of the optimal control problem with rough initial data having finite energy. The authors in Failer et al. (2016) reformulated the variational equations for the FSI problem in a symmetric form, and as mentioned earlier, this approach is advantageous in the numerical analysis and computation of the optimal control problem. Also, a formal Lagrangian approach to the original weak formulation of the state equation leads to an adjoint equation with new coupling conditions on the interface, while the symmetric formulation leads to an adjoint equation that is again an FSI problem.

Nevertheless, our optimality conditions established from a more direct method are equivalent to the one obtained from their symmetric formulation. The approach we follow in this paper will be more transparent in identifying the strong form of the adjoint equations, from which we will see that it is also a linear FSI problem but with nonlocal-in-time terms on the right-hand side of the structure equation. Regularity results and a priori error estimates for the primal states can be then applied to the associated adjoint states.
The second aim of this paper is to study a parabolic regularization of the above optimal control problem where the state equation (1.2) is replaced by

\[
\begin{align*}
  u_\varepsilon^t - \Delta u_\varepsilon + \nabla p_\varepsilon &= B_f q_\varepsilon & \text{in } Q_f, \\
  \text{div } u_\varepsilon &= 0 & \text{in } Q_f, \\
  u_\varepsilon &= 0 & \text{on } \Sigma_f, \\
  u_\varepsilon &= w_\varepsilon^t & \text{on } \Sigma_s, \\
  w_{tt}^\varepsilon - \Delta w_\varepsilon - \varepsilon \Delta w_\varepsilon^t + w_\varepsilon + \mu w_\varepsilon^t (\cdot - r) &= B_s q_\varepsilon & \text{in } Q_s, \\
  \partial_n w_\varepsilon^t + \varepsilon \partial_n w_\varepsilon^t - \partial_n u_\varepsilon^t + p_\varepsilon v &= 0 & \text{on } \Sigma_s, \\
  u_\varepsilon^t (0) &= u_0 & \text{in } \Omega_f, \\
  w_\varepsilon^t (0) &= w_0 & \text{in } \Omega_s, \\
  w_\varepsilon^t &= z_0 & \text{in } Q_r,
\end{align*}
\]

with a regularization parameter or strong damping coefficient \( \varepsilon > 0 \). This regularization strategy is widely used for hyperbolic problems; see Kröner et al. (2011); Lions (1971) for example. As a result, better convergence rates for the discretization errors will be obtained. From a physical point of view the stress for the structure is proportional not only to the strain but also to the strain rate (Caroll & Showalter, 1976). This changes the nature of the FSI model from a coupled parabolic–hyperbolic system to a coupled parabolic–parabolic system. System (1.4) without delay has been studied in Zhang (2017), where it was shown that the associated semigroup generator is analytic and exponentially stable.

It will be shown that the optimal solution of the regularized problem converges to the optimal solution of the original problem as the parameter \( \varepsilon \) tends to 0. Due to the strong damping on the wave equation, this problem possesses solutions that have better regularity properties, and we shall utilize this information to propose and analyze a numerical method approximating the optimal control. Moreover, we prove \textit{a priori} error estimates for the control, state and adjoint variables.

We shall use piecewise linear elements for the discretization of the structure and control and mini finite elements for the fluid velocity and pressure (see Arnold et al., 1984). For the mini finite element, extra degrees of freedom are used at the barycenters of each triangle in the spatial mesh. The corresponding shape function is commonly called a bubble function. This is one of the simplest and most economical finite element methods to implement for the Stokes equation that has the appropriate approximation properties and fulfills the discrete inf–sup condition, a necessary criterion to derive \textit{a priori} estimates. Recall that for the linear Stokes equation, linear elements both for the fluid velocity and the pressure are not sufficient since it may produce the so-called checkerboard-like instability that leads to the failure of the inf–sup condition (Ern & Guermond, 2004, Section 4.2.3). The use of bubble functions in the mini element can be viewed as a Galerkin/least squares approximation for the P1–P1 element; see Quarteroni & Valli (2008, Section 9.4) for the details. Moreover, the proposed numerical scheme preserves the continuity on the interface of the fluid and structure velocities at the discrete level.

For the temporal discretization we shall employ a discontinuous Galerkin scheme. For this type of scheme, it turns out that the history will be discretized through an averaging method, which is reminiscent of the methods proposed in Banks & Burns (1978) for optimal control problems of delay differential equations. The full space-time discretization will then be a linear discrete time-delay system.
Depending on the value of $\gamma_3$ in (1.1), we obtain either a linear or a quadratic order of convergence with respect to the spatial discretization and a linear order with respect to time. Galerkin discretizations are favorable schemes both in theory and numerics because the two approaches, discretize-then-optimize and optimize-then-discretize, are equivalent; see Meidner & Vexler (2008) in the case of parabolic problems.

The numerical scheme presented in this paper is a strongly coupled algorithm. The unknowns of the linear systems are the fluid velocity, fluid pressure, structure displacement and velocity. By a suitable substitution and penalization the structure displacement and the fluid pressure will be eliminated in the system; thus, the fluid and structure velocities remain as the degrees of freedom. In principle, this is a monolithic approach for the FSI algorithm. For linear two-dimensional problems, this is an affordable method and has advantages from the stability point of view. However, for nonlinear and three-dimensional problems such an approach is computationally expensive and appropriate solvers and preconditioners are important.

Alternative approaches based on partitioned schemes have been proposed in past years to circumvent the disadvantages of monolithic approaches. For such an approach the fluid and structure variables are computed separately in their respective domains and they are coupled through the interface boundary conditions. Partitioned schemes gained significance due to their numerical and storage efficiency, modularity and scalability. However, instabilities may occur due to the so-called mass-added effect, which means that the mass of the structure increases due to the surrounding fluid as it vibrates. This is typical in hemodynamics where the densities of the blood and the arterial wall tissue are comparable. Nevertheless, appropriate operator splitting schemes have been developed to overcome such instabilities. For more details on partitioned schemes and mass-added effects we refer the reader to Baek & Karniadakis (2012); Banks et al. (2017); Bukač et al. (2013, 2015); Causin et al. (2005); Fernandez (2011); Guidoboni et al. (2009); Li et al. (2016); Lukáčova-Medvid’ová et al. (2013) and the references therein. It will be good future work to extend the current paper to partitioned schemes for nonlinear two-dimensional and three-dimensional problems with either a static or a dynamic interface, specifically, to investigate the form of adjoint equations and to analyze the corresponding discrete optimal control problem.

Now, as a motivation, let us take into consideration the influence of delay on the optimal control problem (1.1), (1.3) subject to the state equation (1.4). In the following we shall use the setup of the first example in Section 8 with control acting in the structure domain. Using the numerical scheme described above and discussed in detail in Section 6, we computed the optimal control by neglecting the delay ($r = 0$) and then utilized it as a control to the dynamics with delay $r = 1$. While the residuals on the fluid velocity and structure stress are comparable in size we can observe from Fig. 1 that there is a clear difference between the structural displacement and velocity when the optimal control obtained by neglecting delay is applied to the state equation with delay. Therefore, if there is a priori knowledge that time delay is present in the state equation then one should utilize this information to improve the results of the optimal control formulation.

This paper will be organized as follows. In Section 2 we will discuss the well-posedness and regularity of solutions for the state equations (1.2) and (1.4). The necessary optimality conditions for the associated optimal control problems will be tackled in Section 3. We present equivalent symmetric formulations for the state and adjoint equations in Section 4. In Section 5 a semidiscretization for the symmetric formulation as well as a priori error estimates will be established. Full space-time discretization of these equations is the concern of Section 6. Section 7 will deal with the error analysis for the optimal controls of the discretized and continuous problems. Finally, numerical experiments illustrating the theoretical results will be provided in Section 8.
2. Analysis of the state equations

In this section we present the weak formulation of the state equation (1.2) and its parabolic counterpart (1.4). The plan is to study the case where there is no delay first and then use successive substitutions in the presence of the delay term. Without delay the existence and uniqueness of weak solutions has been established in the literature using different methods such as semigroup theory, the Galerkin method and monotone operators. Nevertheless, we sketch the Galerkin method for clarity and completeness of presentation. The product space approach, in which an auxiliary state is introduced to keep track of the history, serves as a good theoretical framework both in the analysis and approximations for delay differential equations. This is also an appropriate tool for partial differential equations with delay from a theoretical point of view; however, this is not practical from the computational perspective since it blows up the number of degrees of freedom. With this concern we will use successive substitution instead, and we will see later that this approach is compatible with the discontinuous Galerkin scheme with respect to time. The main drawback of the method of successive substitution is the additional storage of the history components, which is unavoidable when dealing with a dynamics that depends on the history of the state.

2.1 Well-posedness of the state equations

The Lebesgue and Sobolev spaces on a domain $O$ in $\mathbb{R}^d$ will be denoted by $L^p(O)$ and $W^{k,p}(O)$, respectively, and we let $H^k(O) = W^{k,2}(O)$. The corresponding norms will be denoted by

$$
\|u\|_O = \|u\|_{L^2(O)}, \quad \|u\|_{k,p,O} = \|u\|_{W^{k,p}(O)}, \quad \|u\|_{k,O} = \|u\|_{k,2,O}.
$$

Simplifying notation, the product of $m$ copies of a Banach space $X$ will be denoted again by $X$ instead of $X^m$. We shall use the abbreviations $L^p(X) = L^p(I,X)$, $W^{s,p}(X) = W^{s,p}(I,X)$, $H^s(X) = W^{s,k}(X)$ and
$C^k(X) = C^k(\mathcal{I}, X)$ for $p \in [1, \infty]$, $s \in \mathbb{R}$ and $k \in \mathbb{N} \cup \{0\}$. A subscript $r$ will be used if instead of the interval $I$ we have $(-r, 0)$, for example $L^p_r(X) = L^p((-r, 0), X)$ and $H^s_r(X) = H^s((-r, 0), X)$. The indicator function of a set $O$ is denoted by $\chi_O$. We use $\langle \cdot, \cdot \rangle$ to denote the pairing between a Banach space and its dual.

With respect to the fluid the typical solenoidal function spaces for the Stokes equation will be used, namely

\[ H_f = \{ u \in L^2(\Omega_f) : \text{div} \ u = 0 \text{ in } \Omega_f, \ u \cdot \nu = 0 \text{ on } \Gamma_f \}, \]

\[ V_f = \{ u \in H^1(\Omega_f) : \text{div} \ u = 0 \text{ in } \Omega_f, \ u = 0 \text{ on } \Gamma_f \}, \]

with the corresponding norms

\[ \| u \|_{H_f} = \| u \|_{\Omega_f}, \quad \| u \|_{V_f} = \| \nabla u \|_{\Omega_f}. \]

On the other hand, for the structure displacement and velocity, we take the function spaces $H_s = L^2(\Omega_s)$ and $V_s = H^1(\Omega_s)$ equipped with their usual norms. Also, we introduce the spaces

\[ V = \{ \xi \in H^1_0(\Omega) : \text{div} \ \xi = 0 \text{ on } \Omega_f \}, \]

\[ \tilde{V}_f = \{ u \in H^1_0(\Omega_f) : \text{div} \ u = 0 \text{ in } \Omega_f \}, \]

endowed with the respective norms

\[ \| \xi \|_V = \| \nabla \xi \|_\Omega, \quad \| u \|_{\tilde{V}_f} = \| \nabla u \|_{\Omega_f}. \]

For the pressure we use the space $M = L^2(\Omega_f)$. This choice of function space for the pressure stems from the Neumann-type boundary condition on the interface. Notice here that the pressure does not necessarily have average zero, in contrast to the Stokes equation with no-slip boundary condition where one has to impose the zero average condition to obtain the uniqueness of the pressure term. Furthermore, we shall use $W = H^1_0(\Omega)$ as the space for test functions for the variational formulation of the state equations with regular data.

The duals of $V$ and $V_f$ with respect to the pivot spaces $L^2(\Omega)$ and $H_f$ will be denoted by $V'$ and $V'_f$, respectively. Similarly, $V'_s$ and $\tilde{V}'_f$ are the duals of $V_s$ and $\tilde{V}_f$ with respect to the pivot spaces $H_s$ and $H_f$, respectively.

Let us begin with the definition of weak solutions for (1.2) without the retarded term. In the following discussion we consider a nonhomogeneous boundary condition in relation to the normal stresses on the interface $\Gamma_s$; compare with Lasiecka & Tuffaha (2009a). This type of boundary condition appears naturally in the formulation of the adjoint equations in the optimality system for the case $\gamma_{s3} > 0$. 

DEFINITION 2.1 Let \( u_0 \in H_f, v_0 \in H_s, w_0 \in V_s, f \in L^2(V_f'), \sigma \in L^2(H_s) \) and \( \beta \in L^2(H^{-1/2}(\Gamma_s)) \). A pair \( (u, w) \in [L^\infty(H_f) \cap L^2(V_f') \cap H^1(\tilde{V}_f)] \times [L^\infty(V_f') \cap W^{1,\infty}(H_s) \cap H^2(H^{-1}(\Omega_s))] \) is called a weak solution of

\[
\begin{aligned}
  u_t - \Delta u + \nabla p &= f & \text{in } Q_f, \\
  \text{div } u &= 0 & \text{in } Q_f, \\
  u &= 0 & \text{on } \Sigma_f, \\
  u &= w_t & \text{on } \Sigma_s, \\
  w_t - \Delta w + w &= \sigma & \text{in } Q_s, \\
  \partial_n w - \partial_n u + p v &= \beta & \text{on } \Sigma_s, \\
  u(0) &= u_0 & \text{in } \Omega_f, \\
  w(0) &= w_0, & w_t(0) &= v_0 & \text{in } \Omega_s,
\end{aligned}
\]

if the initial conditions in (2.1) hold, \( u|_{\Gamma_s} = w|_{\Gamma_s} \) in \( L^2(H^{1/2}(\Gamma_s)) \) and for almost every \( t \in I \) we have

\[
\langle u_t, \varphi \rangle + \langle \nabla u, \nabla \varphi \rangle_{\Omega_f} + \langle w_t, \varphi \rangle + \langle w, \varphi \rangle_{1,\Omega_s} = \langle f, \varphi \rangle + \langle \beta, \varphi \rangle + \langle \sigma, \varphi \rangle_{\Omega_s}
\]

(2.2)

for every \( \varphi \in V \).

Equation (2.2) should be understood in the sense of distributions, that is,

\[
- \int_I (u(t), \varphi) \varphi'(t) \, dt + \int_I (\nabla u(t), \nabla \varphi)_{\Omega_f} \varphi(t) \, dt - \int_I (w_t(t), \varphi) \varphi'(t) \, dt \\
+ \int_I (w(t), \varphi)_{1,\Omega_s} \varphi(t) \, dt = \int_I \{ \langle f(t), \varphi \rangle + \langle \beta(t), \varphi \rangle + \langle \sigma(t), \varphi \rangle_{\Omega_s} \} \varphi(t) \, dt
\]

for every \( \varphi \in C_0^\infty(I) \). Note that \( w_t|_{\Gamma_s} \) is a well-defined element in \( H^{-1}(H^{1/2}(\Gamma_s)) \). According to the definition, we have \( u \in C(\tilde{V}_f') \) and \( w \in C^1(H^{-1}(\Omega_s)) \). Consequently, the pointwise values of \( u, w \) and \( w_t \) at \( t = 0 \) are well defined.

We would like to point out here that our definition of a weak solution is adapted from Lions (1969, Section I.9); see also Du et al. (2003) in the case of more regular data. A different notion of weak solution is given in Barbu et al. (2007) where the test functions are not necessarily coupled at the interface. The crucial point in that formulation is the hidden regularity of the normal trace on the interface of solutions for the wave equation, namely \( \partial_n w \in L^2(H^{-1/2}(\Gamma_s)) \), obtained from a microlocal analysis argument. Under additional regularity assumptions on the data the weak solution enjoys additional regularity as well and coincides with the notion of strong solution; see Theorem 2.3 below. Moreover, the variational form (2.2) is a natural setup for strongly coupled algorithms while the one given in Barbu et al. (2007) is suitable for partitioned algorithms. Since our numerical scheme is written in terms of the global velocity field we shall utilize definition (2.2).
Theorem 2.2  System (2.1) has a unique weak solution and there exists a constant $C > 0$ independent of the solution and the data such that

$$
\|u\|_{L^\infty(H_f)} + \|w\|_{L^\infty(V_f)} \leq C(\|f\|_{L^2(V_f)} + \|\beta\|_{H^1/2(\Gamma_f)} + \|\sigma\|_{H^2} + \|u_0\|_{\Omega_f} + \|v_0\|_{\Omega_f} + \|w_0\|_{1,\Omega_f}).
$$

(2.3)

Moreover, the components of the weak solution satisfy

$$(u, w) \in C(H_f) \times [C(V_f) \cap C^1(\Gamma_f)].$$

(2.4)

Proof. For the proof of the existence and uniqueness of a weak solution $(u, w) \in [L^\infty(H_f) \cap L^2(V_f)] \times [L^\infty(V_f) \cap W^{1,\infty}(H_f)]$ by the Faedo–Galerkin method and the a priori estimate (2.3) without the norms in $H^1(\tilde{\Omega}_f)$ and $H^2(H^{-1}(\Omega_s))$ we refer the reader to Lions (1969, Section I.9). By choosing $\varphi \in \tilde{V}_f$ and extending it by zero outside $\Omega_f$ we can see from (2.2) that $u \in H^1(\tilde{\Omega}_f)$. Similarly, by taking $\varphi \in H^1_0(\Omega_s)$ and extending it by zero outside $\Omega_s$, we have $w \in H^2(H^{-1}(\Omega_s))$. In particular, the following estimates hold:

$$
\|u_t\|_{L^2(\tilde{\Omega}_f)} \leq C(\|u\|_{L^2(V_f)} + \|f\|_{L^2(V_f)}),
$$

$$
\|w_{tt}\|_{L^2(H^{-1}(\Omega_s))} \leq C(\|w\|_{L^2(V_f)} + \|\sigma\|_{H^2}).
$$

Finally, the continuity of weak solutions with respect to time can be shown by following the methods in Lions & Magenes (1972, Chapter 3).

The existence and uniqueness of the pressure can be established under additional assumptions on the data.

Theorem 2.3  Suppose that $u_0 \in V_f, w_0, v_0 \in V_s$ satisfy $\Delta u_0 \in L^2(\Omega_f), \Delta w_0 \in H_s, u_0 = v_0$ on $\Gamma_s$ and $\partial_n w_0 = \partial_n u_0 - \rho_0 v_0$ on $\Gamma_s$ for some $p_0 \in H^1(\Omega_f)$. If $f \in H^1(V_f)$, $\sigma \in H^1(H_s)$ and $\beta \in H^1(H^{-1/2}(\Gamma_f))$ then the weak solution of (2.1) satisfies

$$
u \in W^{1,\infty}(H_f) \cap H^1(V_f), \quad w \in W^{1,\infty}(V_s) \cap W^{2,\infty}(H_s),
$$

and there exists a unique $p \in L^2(M)$ such that

$$(u, \varphi)_{\Omega_f} + (\nabla u, \nabla \varphi)_{\Omega_f} - (p, \text{div} \varphi)_{\Omega_f} + (w, \varphi)_{\Omega_s} + (w, \varphi)_{1,\Omega_s} = (f, \varphi) + \langle \beta, \varphi \rangle + \langle \sigma, \varphi \rangle_{\Omega_s}
$$

for every $\varphi \in W$ and for a.e. $t \in I$. Moreover, there exists $C > 0$ independent of the solution and the data such that

$$
\|u\|_{W^{1,\infty}(H_f) \cap H^1(V_f)} + \|w\|_{W^{1,\infty}(V_s) \cap W^{2,\infty}(H_s)} + \|p\|_{L^2(M)}
$$

$$
\leq C(\|f\|_{H^1(V_f)} + \|\beta\|_{H^1(H^{-1/2}(\Gamma_f))} + \|\sigma\|_{H^1(H_s)} + \|u_0\|_{1,\Omega_f} + \|\Delta u_0\|_{\Omega_f})
$$

$$
+ C(\|v_0\|_{1,\Omega_s} + \|\Delta w_0\|_{\Omega_s} + \|w_0\|_{1,\Omega_s} + \|p_0\|_{1,\Omega_f}).
$$


Proof. The proof of this result can be seen in Du et al. (2003) using an inf–sup condition (see Theorem 2.14 below) with the additional assumptions that \( u_0 \in H^2(\Omega_s), w_0 \in H^2(\Omega_s), f \in H^1(H_f) \) and \( \beta = 0 \). However, the arguments can be adjusted so that the results are still valid under the requirements for \( u_0, w_0, f \) and \( \beta \) stated in the theorem. \( \square \)

In the framework of the previous theorem, it can be shown that the boundary condition \( \partial_v w = \partial_v u - pv + \beta \) on \( \Gamma_s \) is satisfied in \( L^2(H^{-1/2}(\Gamma_s)) \). We refer the reader to Barbu et al. (2008) for a proof of this remark.

Theorem 2.4 Assume that \( u_0 \in V_f \cap H^2(\Omega_f), w_0 \in H^2(\Omega_s) \) and \( v_0 \in V_s \) satisfy the compatibility conditions \( u_0 = v_0 \) on \( \Gamma_s \) and \( \partial_v w_0 = \partial_v u_0 - p_0 v_0 \) on \( \Gamma_s \) for some \( p_0 \in H^1(\Omega_f) \). Let \( f \in H^1(V_f') \cap L^2(L^2(\Omega_f)), \sigma \in H^1(Q_s) \) and \( \beta \in H^1(H^{-1/2}(\Gamma_s)) \cap L^2(H^{1/2}(\Gamma_s)) \). If \( \Omega_s \) is a sufficiently smooth domain then the weak solution of (2.1) satisfies

\[
\begin{align*}
    u & \in L^2(H^2(\Omega_f)), \quad w \in L^2(H^2(\Omega_s)), \quad p \in L^2(H^1(\Omega_f)),
\end{align*}
\]

and there exists \( C > 0 \) independent of the solution and the data such that

\[
\begin{align*}
    \|u\|_{L^2(H^2(\Omega_f))} + \|w\|_{L^2(H^2(\Omega_s))} + \|p\|_{L^2(H^1(\Omega_f))} \\
    \leq C \left( \|f\|_{H^1(V_f') \cap L^2(L^2(\Omega_f))} + \|\beta\|_{H^1(H^{-1/2}(\Gamma_s)) \cap L^2(H^{1/2}(\Gamma_s))} + \|\sigma\|_{H^1(Q_s)} \right) \\
    + C \left( \|u_0\|_{2,\Omega_f} + \|v_0\|_{1,\Omega_s} + \|w_0\|_{2,\Omega_s} + \|p_0\|_{1,\Omega_f} \right). \tag{2.6}
\end{align*}
\]

Proof. A proof of this theorem in the case \( f = 0, \sigma = 0 \) and \( \beta = 0 \) can be found in Barbu et al. (2008), which can be adapted to the nonhomogeneous case with the above regularity assumptions. For completeness we give the proof. First, let us prove regularity away from the interface. For this purpose we define \( \Omega^\delta_f = \{ x \in \Omega_f : \text{dist}(x, \Gamma_s) > \delta \} \) for \( f = f \) or \( s \), whenever \( \delta > 0 \) is sufficiently small. Let \( \chi_f \in C^\infty_c(\Omega^\delta_f \cup \Gamma_f) \) be a cutoff function such that \( \chi_f = 1 \) on \( \Omega^\delta_f \cup \Gamma_f \). Multiplying the fluid equation by \( \chi_f \) leads to

\[
\begin{align*}
    -\Delta \tilde{u} + \nabla \tilde{p} &= \tilde{f} & \text{in } Q_f, \\
    \text{div } \tilde{u} &= \nabla \chi_f \cdot u & \text{in } Q_f, \\
    \tilde{u} &= 0 & \text{on } \Sigma_f \cup \Sigma_s,
\end{align*}
\]

where \( \tilde{u} = \chi_f u \) and \( \tilde{f} = \chi_f (f - u_t) - [\Delta, \chi_f] u + [\nabla, \chi_f] p \). The commutators \( [\Delta, \chi_f] \) and \( [\nabla, \chi_f] \) are of orders 1 and 0, respectively, and hence from Theorem 2.3 we have \( \tilde{f} \in L^2(L^2(\Omega_f)) \). For the above Stokes problem we have the compatibility condition

\[
\int_{\Omega_f} \nabla \chi_f \cdot u \, dx = \int_{\Gamma_f \cup \Gamma_s} \chi_f u \cdot v \, dx = 0.
\]
It follows from the regularity theory for the Stokes equation in Kellog & Osborn (1976); Temam (2001) that $\bar{u} \in L^2(H^2(\Omega_f))$ and thus $u \in L^2(H^2(\Omega_f^\delta))$ for every sufficiently small $\delta > 0$ and

$$\|u\|_{L^2(H^2(\Omega_f^\delta))} \leq C(\|u\|_{H^1(\Gamma_f^\delta \cap \Gamma_f^\delta)} + \|p\|_{L^2(\Gamma_f)} + \|f\|_{L^2(\Omega_f^\delta)}). \quad (2.7)$$

In a similar way we multiply the wave equation by a cutoff function $\chi \in C^\infty_0(\Omega_s)$ such that $\chi = 1$ in $\Omega_s^\delta$ to obtain

$$\begin{cases} -\Delta \tilde{w} + \tilde{w} = \tilde{\sigma}, & \text{in } Q_s, \\ \tilde{w} = 0 & \text{on } \Sigma_s, \end{cases}$$

where $\tilde{w} = \chi_s w$ and $\tilde{\sigma} = \chi_s(\sigma - w_t - \mu w_t \cdot - r) - [\Delta, \chi_s]w$. With a similar argument to above, it holds that $\tilde{\sigma} \in L^2(\Omega_s^\delta)$. From the regularity of elliptic equations in Lions & Magenes (1972) we have $\tilde{w} \in L^2(\Omega_s^\delta)$, and so $w \in L^2(\Omega_s^\delta)$ for every sufficiently small $\delta > 0$ and

$$\|w\|_{L^2(H^2(\Omega_s^\delta))} \leq C(\|w\|_{H^2(\Gamma_f^\delta \cap \Gamma_f^\delta)} + \|z_0\|_{L^2(\Omega_s)} + \|\sigma\|_{L^2(\Omega_f^\delta)}). \quad (2.8)$$

The right-hand sides of (2.7) and (2.8) are finite in the light of Theorem 2.3.

It remains to prove regularity on a neighborhood of the interface $\Gamma_s$. As an intermediate step, let us consider the case where the FSI domain is $\Omega^* = [-1, 1]^2$. More precisely, let

$$\Omega_f^* = (-1, 1) \times (0, 1), \quad \Omega_s^* = (-1, 1) \times (-1, 0), \quad \Gamma_s^* = (-1, 1) \times \{0\}.$$

Suppose that $u \in H^1(L^2(\Omega_f^*)), w \in W^{2,\infty}(L^2(\Omega_s^*)) \cap W^{1,\infty}(H^1(\Omega_s^*)), p \in L^2(L^2(\Omega_s^*)), f \in L^2(L^2(\Omega_f^*)), g \in L^2(H^1(\Omega_s^*)), \sigma \in L^2(H^1(\Omega_s^*))$ and $\beta \in L^2(H^{1/2}(\Gamma_s^*))$ satisfy the following equations in the sense of distributions:

$$\begin{cases} u_t - \Delta u + \nabla p = f & \text{in } I \times \Omega_f^*, \\ \text{div } u = g & \text{in } I \times \Omega_f^*, \\ u = 0 & \text{on } I \times (\partial \Omega_f^* \setminus \Gamma_s^*), \\ u = w_t & \text{on } I \times \Gamma_s^*, \\ w_{tt} - \Delta w + w = \sigma & \text{in } I \times \Omega_s^*, \\ \partial_y w - \partial_y u + pe_y = \beta & \text{on } I \times \Gamma_s^*, \\ u(0) = u_0 & \text{in } \Omega_f^*, \\ w(0) = w_0, \quad w_t(0) = v_0 & \text{in } \Omega_s^*, \end{cases} \quad (2.9)$$

where $e_y = (0, 1)^T$. 
For $\eta > 0$, let $\rho_\eta$ be a standard mollifier with respect to the first spatial variable $x$ and let $u^n = u * \rho_\eta$, $p^n = p * \rho_\eta$ and $w^n = w * \rho_\eta$, where $*$ denotes convolution. Analogous definitions for $f^n$, $g^n$, $\beta^n$ and $\sigma^n$ will be utilized. Extending the functions outside $\Omega^*$ by zero, mollifying the differential equations in (2.9) and then applying $\partial_x$, we obtain the following estimate by multiplying the fluid equation by $\partial_x u^n$ and the structure equation by $\partial_x w^n$:.

\[
\frac{1}{2} \frac{d}{dt} \left( \| \partial_x u^n(t) \|_{L^2(\Omega^*_t)}^2 + \| \partial_x w^n(t) \|_{L^2(\Omega^*_t)}^2 + \| \partial_t w^n(t) \|_{L^2(\Omega^*_t)}^2 + \| \nabla (\partial_x u^n(t)) \|_{L^2(\Omega^*_t)}^2 \right) \leq C \left( \| f^n(t) \|_{L^2(\Omega^*_t)}^2 + \| \partial_x p^n(t), \partial_x g^n(t), \partial_x w^n(t) \|_{L^2(\Omega^*_t)}^2 \right) + C \left( \| \partial_x \sigma(t), \partial_x w^n(t) \|_{L^2(\Omega^*_t)}^2 + \frac{1}{2} \| \partial_x u^n(t) \|_{L^2(\Omega^*_t)}^2 \right). \tag{2.10}
\]

Here $\eta > 0$ is small enough so that the supports of the involved functions do not exceed the square $\Omega^*$. According to the trace theorem and Poincaré inequality,

\[
\left\| (\partial_x \beta^n(t), \partial_x u^n(t)) \right\|_{L^2(\Gamma^*_t)} \leq \| \partial_x \beta^n(t) \|_{L^2(\Gamma^*_t)} + \| \partial_x u^n(t) \|_{L^2(\Omega^*_t)} \leq C \| \beta^n(t) \|_{L^2(\Gamma^*_t)} + \epsilon \| \partial_x u^n(t) \|_{L^2(\Omega^*_t)} \tag{2.11}
\]

for each $\epsilon > 0$. Also, for each $\epsilon > 0$ it holds that

\[
\left\| (\partial_x p^n(t), \partial_x g^n(t)) \right\|_{L^2(\Omega^*_t)} \leq \epsilon \| \partial_x p^n(t) \|_{L^2(\Omega^*_t)} + C \| g^n(t) \|_{L^2(\Omega^*_t)}. \tag{2.12}
\]

Integrating (2.10) over the time interval $I$, applying Gronwall’s lemma and using (2.11) and (2.12), we infer that

\[
\| \partial_x u^n \|_{L^\infty(L^2(\Omega^*_T))} + \| \partial_x w^n \|_{L^\infty(L^2(\Omega^*_T))} + \| \partial_t w^n \|_{L^\infty(H^1(\Omega^*_T))} + \| \nabla \partial_x u^n \|_{L^2(L^2(\Omega^*_T))} \leq C (\| f \|_{L^2(L^2(\Omega^*_T))} + \| g \|_{L^2(H^1(\Omega^*_T))} + \| \beta \|_{L^2(H^1(\Gamma^*_T))} + \| \sigma \|_{L^2(H^1(\Omega^*_T))} + c_0^*) \]

\[+ 2\epsilon \| \partial_x p^n \|_{L^2(L^2(\Omega^*_T))}, \tag{2.13}\]

where $c_0^* = \| u_0 \|_{L^2(\Omega^*_T)} + \| v_0 \|_{L^2(\Omega^*_T)} + \| w_0 \|_{L^2(\Omega^*_T)}$. Here we used the fact that convolution is uniformly bounded for sufficiently small $\eta > 0$. From the equation $\nabla (\partial_x p^n) = \partial_x f^n + \Delta (\partial_x u^n) - \partial_x u^n$ we have $\nabla (\partial_x p^n) \in L^2(H^{-1}(\Omega^*_T))$ and

\[
\| \nabla (\partial_x p^n) \|_{L^2(H^{-1}(\Omega^*_T))} \leq C \left( \| u_0 \|_{L^2(L^2(\Omega^*_T))} + \| \nabla (\partial_x u^n) \|_{L^2(L^2(\Omega^*_T))} + \| f \|_{L^2(L^2(\Omega^*_T))} \right). \tag{2.14} \]
It follows from Nečas (1976, Chapter 3, Lemma 7.1) and (2.14) that
\[
\|\partial_x u^0\|_{L^2(H^1(\Omega_f^s))} + \|\partial_x w^0\|_{L^2(H^1(\Omega_f^s))} + \|\partial_y p^0\|_{L^2(L^2(\Omega_f^s))} \leq C\|f\|_{L^2(L^2(\Omega_f^s))} + C(\|g\|_{L^2(H^1(\Omega_f^s))} + \|\beta\|_{L^2(H^{1/2}(\Gamma_f^s))} + \|u_t\|_{L^2(L^2(\Omega_f^s))} + \|\sigma\|_{L^2(H^1(\Omega_f^s))} + \ell_0^s).
\]  
Combining (2.13) and (2.15), and then choosing \(\ell\) small enough, we deduce after neglecting some non-negative terms on the left-hand side that
\[
\|\partial_x u^0\|_{L^2(H^1(\Omega_f^s))} + \|\partial_x w^0\|_{L^2(H^1(\Omega_f^s))} + \|\partial_y p^0\|_{L^2(L^2(\Omega_f^s))} \leq C\|f\|_{L^2(L^2(\Omega_f^s))} + C(\|g\|_{L^2(H^1(\Omega_f^s))} + \|\beta\|_{L^2(H^{1/2}(\Gamma_f^s))} + \|u_t\|_{L^2(L^2(\Omega_f^s))} + \|\sigma\|_{L^2(H^1(\Omega_f^s))} + \ell_0^s).
\]  
Thus, \((\partial_x u^0)_\eta, (\partial_x w^0)_\eta\) and \((\partial_y p^0)_\eta\) are bounded in \(L^2(H^1(\Omega_f^s))\), \(L^2(H^1(\Omega_f^s))\) and \(L^2(L^2(\Omega_f^s))\), respectively, and hence up to a subsequence each converges weakly to some element in the corresponding spaces as \(\eta \to 0\). However, we know that \(\partial_x u^0 \to \partial_x u\) in \(L^2(L^2(\Omega_f^s))\), \(\partial_x w^0 \to \partial_x w\) in \(L^2(L^2(\Omega_f^s))\) and \(\partial_y p^0 \to \partial_y p\) in \(L^2(H^{-1}(\Omega_f^s))\). From the uniqueness of weak limits, it follows that \(\partial_x u \in L^2(H^1(\Omega_f^s))\), \(\partial_x w \in L^2(H^1(\Omega_f^s))\) and \(\partial_y p \in L^2(L^2(\Omega_f^s))\). Passing to the limit inferior in (2.16) and using the weak lower semicontinuity of the norm we have
\[
\|\partial_x u\|_{L^2(H^1(\Omega_f^s))} + \|\partial_x w\|_{L^2(H^1(\Omega_f^s))} + \|\partial_y p\|_{L^2(L^2(\Omega_f^s))} \leq C\|f\|_{L^2(L^2(\Omega_f^s))} + C(\|g\|_{L^2(H^1(\Omega_f^s))} + \|\beta\|_{L^2(H^{1/2}(\Gamma_f^s))} + \|u_t\|_{L^2(L^2(\Omega_f^s))} + \|\sigma\|_{L^2(H^1(\Omega_f^s))} + \ell_0^s).
\]  
The stated regularities of \(u\) and \(p\) with respect to space can now be obtained from the fluid equation. Indeed, if \(u = (u_1, u_2)^T\) then we have \(\partial_x u_1 = u_1_t - \partial_y u_1 + \partial_y p - f_1 \in L^2(L^2(\Omega_f^s))\), \(\partial_y u_2 = \partial_y g - \partial_y u_1 \in L^2(L^2(\Omega_f^s))\), therefore \(u \in L^2(H^1(\Omega_f^s))\), and consequently \(p \in L^2(H^1(\Omega_f^s))\). Using the same argument for \(w\) in the wave equation yields \(w \in L^2(H^2(\Omega_f^s))\). Together with (2.17), these also imply that
\[
\|\partial_x u\|_{L^2(H^1(\Omega_f^s))} + \|\partial_y w\|_{L^2(H^1(\Omega_f^s))} + \|\partial_y p\|_{L^2(L^2(\Omega_f^s))} \leq C(\|f\|_{L^2(L^2(\Omega_f^s))} + \|g\|_{L^2(H^1(\Omega_f^s))} + \|\beta\|_{L^2(H^{1/2}(\Gamma_f^s))} + \|u_t\|_{L^2(L^2(\Omega_f^s))} + \|\sigma\|_{L^2(H^1(\Omega_f^s))} + \ell_0^s).
\]  
With respect to the original domain, one considers a partition of unity in a neighborhood of the interface \(\Gamma_s\) and performs a transformation of variables from each patch to the square domain \(\Omega_f^s\). Then using the same technique as for the Stokes equation in \(C^2\)-domains, see for instance Sohr (2001, pp. 119–123), one can obtain \(H^2\)-spatial regularity both for \(u\) and \(w\) on each patch. This in turn implies \(H^2\)-regularity of \(u\) and \(w\) in a neighborhood of \(\Gamma_s\), and upon combining this with the earlier interior regularity we obtain the desired regularity result. Finally, estimate (2.6) can be established by taking the sum of estimates (2.17) and (2.18) obtained on each of the patches and estimates (2.7) and (2.8) obtained from interior regularity.
Remark 2.5 In the proof of the above theorem we refer to Barbu et al. (2008) for a similar flattening of the boundary. In that paper the authors used Melrose–Sjöstrand coordinates (Melrose & Sjöstrand, 1978) for the transformation. This microlocal strategy requires that the domain $\Omega_s$ is sufficiently smooth.

We now consider the state equation (1.2) including the delay term.

Definition 2.6 Let $u_0 \in H_f$, $v_0 \in H_s$, $w_0 \in V_s$, $z_0 \in L^2(Q_r)$, $f \in L^2(V_f')$, $\sigma \in L^2(H_s)$ and $\beta \in L^2(H^{-1/2}(\Gamma^*_s))$. A pair $(u, w) \in [L^\infty(H_f) \cap L^2(V_f) \cap H^1(V_f')] \times [L^\infty(V_s) \cap W^{1,\infty}(H_s) \cap H^2(H^{-1}(\Omega_s))]$ is called a weak solution of

\[
\begin{align*}
    &u_t - \Delta u + \nabla p = f & \text{in } Q_f, \\
    &\text{div } u = 0 & \text{in } Q_f, \\
    &u = 0 & \text{on } \Sigma_f, \\
    &u = w_t & \text{on } \Sigma_s, \\
    &w_{tt} - \Delta w + w + \mu w_t(\cdot - \tau) = \sigma & \text{in } Q_s, \\
    &\partial_\nu w - \partial_\nu u + p
\nu = \beta & \text{on } \Sigma_s, \\
    &u(0) = u_0 & \text{in } \Omega_f, \\
    &w(0) = w_0, \quad w_t(0) = v_0 & \text{in } \Omega_s, \\
    &w_t = z_0 & \text{in } Q_r,
\end{align*}
\]

if the initial conditions in (2.19) hold, $u|_{\Gamma^*_s} = w_t|_{\Gamma^*_s}$ in $L^2(H^{1/2}(\Gamma^*_s))$ and for almost every $t \in I$ we have

\[
(u_t, \varphi) + (\nabla u, \nabla \varphi)_{Q_f} + (w_{tt}, \varphi) + (w_t, \varphi)_{\Omega_s} + (\mu w_t(\cdot - \tau), \varphi)_{\Omega_s} = \langle f, \varphi \rangle + \langle \beta, \varphi \rangle + \langle \sigma, \varphi \rangle_{\Omega_s}
\]

for every $\varphi \in V$ in the sense of distributions.

According to the definition we have $w_t \in L^2(-\tau, T; H_s)$.

Theorem 2.7 System (2.19) has a unique weak solution and there exists a constant $C > 0$ independent of the solution and the data such that

\[
\|u\|_{L^\infty(H_f) \cap L^2(V_f) \cap H^1(V_f')} + \|w\|_{L^\infty(V_s) \cap W^{1,\infty}(H_s) \cap H^2(H^{-1}(\Omega_s))} \leq C(\|f\|_{L^2(V_f')} + \|\beta\|_{L^2(H^{-1/2}(\Gamma^*_s))} + \|\sigma\|_{L^2(H_s)} + \|u_0\|_{\Omega_f} + \|v_0\|_{\Omega_s} + \|w_0\|_{1,\Omega_s} + \|z_0\|_{Q_r}).
\]

Furthermore, (2.4) is satisfied.

Proof. Apply Theorem 2.2 successively on the intervals $[0, r], [r, 2r]$ and so on. \qed
Theorem 2.8  (i) Under the assumptions of Theorem 2.3, and if \( z_0 \in H^1_0(H_s) \) with \( z_0(0) = v_0 \), then the weak solution of (2.19) satisfies \( u \in W^{1,\infty}(H_f) \cap H^1(V_f) \), \( w \in W^{1,\infty}(V_s) \cap W^{2,\infty}(H_s) \) and there exists a unique \( p \in L^2(M) \) such that

\[
(u_t, \varphi)_{\Omega_f} + (\nabla u, \nabla \varphi)_{\Omega_f} - (p, \text{div} \varphi)_{\Omega_f} + (w_t, \varphi)_{\Omega_s} + (w, \varphi)_{1,\Omega_s} + (\mu w_i(\cdot - r), \varphi)_{\Omega_s} = (f, \varphi) + (\beta, \varphi) + (\sigma, \varphi)_{\Omega_s},
\]

for each \( \varphi \in W \) and for a.e. \( t \in I \). Moreover, there exists \( C > 0 \) independent of the solution and the data such that

\[
\|u\|_{W^{1,\infty}(H_f) \cap H^1(V_f)} + \|w\|_{W^{1,\infty}(V_s) \cap W^{2,\infty}(H_s)} + \|p\|_{L^2(M)} \leq C(\|f\|_{H^1(V_f)} + \|\beta\|_{H^1(\Omega_s)} + \|\sigma\|_{H^1(H_s)} + \|u_0\|_{1,\Omega_f} + \|w_0\|_{1,\Omega_s} + \|z_0\|_{H^1(H_s)} + \|p_0\|_{1,\Omega_f}).
\]

(ii) In the framework of Theorem 2.4, and if \( z_0 \in H^1(\Omega_s) \) and \( z_0(0) = v_0 \), then the weak solution of (2.19) satisfies (2.5) and there exists \( C > 0 \) such that

\[
\|u\|_{L^2(H^2(\Omega_f))} + \|w\|_{L^2(\Omega_s)} + \|p\|_{L^2(H^1(\Omega_f))} \leq C(\|f\|_{H^1(V_f)} + \|\beta\|_{H^1(\Omega_s)} + \|\sigma\|_{H^1(H_s)} + \|u_0\|_{2,\Omega_f} + \|w_0\|_{2,\Omega_s} + \|z_0\|_{1,\Omega_s} + \|p_0\|_{1,\Omega_f}).
\]

Proof. The regularity and compatibility assumptions on \( z_0 \) and \( v_0 \) imply that \( w_i(\cdot - r) \in H^1(H_s) \) and \( w_i(\cdot - r) \in H^1(\Omega_s) \) in (i) and (ii), respectively. With this information, (i) and (ii) now follow from Theorems 2.3 and 2.4, respectively, with \( \sigma \) replaced by \( \sigma - \mu w_i(\cdot - r) \).

In considering the cost functional \( G \) with \( \gamma_{k,3} > 0 \) we shall need the concept of a very weak solution to (2.19) with \( \beta = 0 \) to characterize the necessary optimality conditions, or more precisely its corresponding dual version. Here we present the definition for the primal problem using the method of transposition. The definition below is obtained by multiplying the strong equations by appropriate test functions, integrating over the space-time domain and passing all time and space derivatives to the test functions.

Definition 2.9  Let \( u_0 \in H_f, v_0 \in H_s, w_0 \in V_s, z_0 \in L^1_1(H_s), f \in L^2(V_f), \sigma_1 \in L^1(H_s) \) and \( \sigma_2 \in W^{1,1}(V_s) \). A pair \( (u, w) \in L^2(H_f) \times H^1(H_s) \) is called a very weak solution of (2.19) with \( \sigma = \sigma_1 + \sigma_2 \) and \( \beta = 0 \) if for every \( (g, \kappa) \in L^2(H_f) \times L^2(H_s) \) it holds that \( w(0) = w_0 \) and

\[
\int_I (u, g)_{\Omega_f} + (w_t, \kappa)_{\Omega_s} \, dt = \int_I (f, \varphi) - (\sigma_1, \psi_1)_{\Omega_s} + (\sigma_2, \psi) \, dt + (\sigma_2(0), \psi(0))_{\Omega_s} + (u_0, \varphi(0))_{\Omega_f} + (v_0, \psi_1(0))_{\Omega_s} - (w_0, \psi(0))_{1,\Omega_s} + \mu \int_{-r}^0 (z_0(\theta), \psi_1(\theta + r))_{\Omega_s} \, d\theta,
\]

(2.21)
where \((\varphi, \psi)\) is the weak solution of

\[
\begin{align*}
-\varphi_t - \Delta \varphi + \nabla \pi &= g \quad \text{in } Q_f, \\
\text{div} \varphi &= 0 \quad \text{in } Q_f, \\
\varphi &= 0 \quad \text{on } \Sigma_f, \\
\varphi &= -\psi_t \quad \text{on } \Sigma_s, \\
\psi_{tt} - \Delta \psi + \psi - \mu \psi_t (\cdot + r) &= \kappa \quad \text{in } Q_s, \\
\partial_t \psi - \partial_t \varphi + \pi \nu &= 0 \quad \text{on } \Sigma_s, \\
\varphi(T) &= 0 \quad \text{in } \Omega_f, \\
\psi(T) &= 0, \quad \psi_t(T) = 0 \quad \text{in } \Omega_s, \\
\psi_t &= 0 \quad \text{in } (T, T + r) \times \Omega_s.
\end{align*}
\] (2.22)

Note that after time reversal, (2.22) can be written in the form of (2.19), which justifies the notion of a weak solution for (2.22).

**Theorem 2.10** System (2.19) with \(\sigma = \sigma_1 + \sigma_2\) and \(\beta = 0\) admits a unique very weak solution and there is a constant \(C > 0\) independent of the solution and the data such that

\[
\|u\|_{L^2(H_f)} + \|w\|_{H^1(H_s)} \leq C(\|f\|_{L^2(V_f)} + \|\sigma_1\|_{L^1(H_s)} + \|\sigma_2\|_{W^{1,1}(V_s)} + \|u_0\|_{\Omega_f} + \|v_0\|_{\Omega_s} + \|w_0\|_{1,\Omega_s} + \|z_0\|_{H^1(H_s)}).
\] (2.23)

**Proof.** Reversing the time \(t \to T - t\) in (2.22), Theorem 2.2 implies that given \((g, \kappa) \in L^2(H_f) \times L^2(H_s)\), system (2.22) has a unique weak solution \((\varphi, \psi) \in [C(H_f) \cap L^2(V_f)] \times [C(V_s) \cap C^1(H_s)]\) and for some constant \(C > 0\) it holds that

\[
\|\varphi\|_{C(H_f) \cap L^2(V_f)} + \|\psi\|_{C(V_s) \cap C^1(H_s)} \leq C(\|g\|_{L^2(H_f)} + \|\kappa\|_{L^2(H_s)}).
\]

This implies that the mapping of \((g, \kappa)\) to the right-hand side of (2.21) defines a linear functional on \(L^2(H_f) \times L^2(H_s)\). Therefore, according to the Riesz representation theorem, there exists a unique pair \((u, v) \in L^2(H_f) \times L^2(H_s)\) that satisfies (2.21) with \(w_t\) replaced by \(v\). Also, if we denote by \(\tilde{C}\) the constant on the right-hand side of (2.23) then

\[
\|u\|_{L^2(H_f)} + \|v\|_{L^2(H_s)} \leq \tilde{C}.
\] (2.24)

If we define \(w(t) = w_0 + \int_0^t v(s) \, ds\) then it follows that the pair \((u, w)\) is a very weak solution of (2.19) with \(\beta = 0\). Thus, (2.23) follows from (2.24). Uniqueness can be shown in a standard manner. \(\square\)
2.2 Well-posedness of the regularized state equations

In this subsection we study the well-posedness of the parabolic regularization (1.4) of the state equations. Due to the presence of strong damping on the wave equation, less stringent regularity assumptions are needed on the source terms to obtain smooth solutions. We start with the definition of weak solutions.

**Definition 2.11** Suppose that \( u_0 \in H_f, v_0 \in H_s, w_0 \in V_s, z_0 \in L^2(Q_s), f \in L^2(V'_f), \sigma \in L^2(V'_f) \) and \( \beta \in L^2(H^{-1/2}(\Gamma^i)). \) A pair \((u^\varepsilon, w^\varepsilon)\) in \([L^\infty(H_f) \cap L^2(V'_f) \cap H^1(V'_f)] \times [L^\infty(V_s) \cap W^{1,\infty}(H_s) \cap H^2(H^{-1}(\Omega_s))]\) such that \( u^\varepsilon \chi_{\Omega_f} + w^\varepsilon \chi_{\Omega_s} \in L^2(V) \cap H^1(V') \) is called a weak solution of

\[
\begin{aligned}
\dot{u}^\varepsilon - \Delta u^\varepsilon + \nabla p^\varepsilon &= f & \text{in } Q_f, \\
\nabla \cdot u^\varepsilon &= 0 & \text{in } Q_f, \\
\nabla u^\varepsilon &= 0 & \text{on } \Sigma_f, \\
\dot{w}^\varepsilon - \Delta w^\varepsilon - \varepsilon \Delta w^\varepsilon + w^\varepsilon + \mu w^\varepsilon (-r) &= \sigma & \text{in } Q_s, \\
\partial_n w^\varepsilon + \varepsilon \partial_n w^\varepsilon - \partial_n u^\varepsilon + p^\varepsilon v &= \beta & \text{on } \Sigma_s, \\
u^\varepsilon(0) &= u_0 & \text{in } \Omega_f, \\
w^\varepsilon(0) &= w_0, \quad w^\varepsilon_t(0) &= v_0 & \text{in } \Omega_s, \\
w^\varepsilon_t &= z_0 & \text{in } Q_r,
\end{aligned}
\]

if the initial conditions in (2.25) are satisfied and for almost every \( t \in I \) it holds that

\[
(u^\varepsilon, \varphi) + (\nabla u^\varepsilon, \nabla \varphi)_{\Omega_f} + (w^\varepsilon_t, \varphi)_{\Omega_s} + \varepsilon(\nabla w^\varepsilon, \nabla \varphi)_{\Omega_s} + (w^\varepsilon, \varphi)_{1, \Omega_s}
\]

\[+ (\mu w^\varepsilon (-r), \varphi)_{\Omega_s} = (f, \varphi) + (\beta, \varphi) + (\sigma, \varphi), \quad (2.26)\]

for every \( \varphi \in V \) in the sense of distributions.

We point out that the criterion \( u^\varepsilon \chi_{\Omega_f} + w^\varepsilon \chi_{\Omega_s} \in L^2(V) \) implies \( w^\varepsilon \in H^1(V_s) \), and in particular, the continuity of the velocities \( u^\varepsilon = w^\varepsilon \) on the interface \( \Gamma^i \) in the sense of \( L^2(H^{1/2}(\Gamma^i)) \). Moreover, \( u^\varepsilon \in C(V'_f) \) and \( w^\varepsilon \in C^1(1/H_s) \), and hence the pointwise values of \( u^\varepsilon, w^\varepsilon \) and \( w^\varepsilon_t \) at \( t = 0 \) are well defined. Well-posedness of (2.25) with \( f = 0, \sigma = 0, \beta = 0 \) and without delay via semigroup theory is discussed in Zhang (2017).

**Theorem 2.12** System (2.25) admits a unique weak solution. Moreover, \( u^\varepsilon \in C(\bar{H}_f), w^\varepsilon \in C(V_s) \cap C^1(\bar{H}_s) \), and there exists \( C_\varepsilon > 0 \) with \( C_\varepsilon \to \infty \) as \( \varepsilon \to 0 \) such that

\[
\|u^\varepsilon\|_{L^\infty(H_t) \cap L^2(V'_f)} + \|w^\varepsilon\|_{W^{1,\infty}(H_s) \cap L^\infty(V_s)} + \|\nabla w^\varepsilon_t\|_{L^2(H_s)}
\]

\[
\leq C_\varepsilon (\|f\|_{L^2(V'_f)} + \|\beta\|_{L^2(H^{-1/2}(\Gamma^i))} + \|\sigma\|_{L^2(V'_f)} + \|u_0\|_{\Omega_f} + \|v_0\|_{\Omega_s} + \|w_0\|_{1, \Omega_s} + \|z_0\|_{Q_r}).
\]

\( \text{(27)} \)
\[ \frac{d}{dt} \left( \| u^\varepsilon(t) \|_{\Omega_j}^2 + \| w^\varepsilon(t) \|_{\Omega_s}^2 + \| w^\varepsilon(t) \|_{1,\Omega_s}^2 \right) + \| \nabla u^\varepsilon(t) \|_{\Omega_j}^2 + \varepsilon \| \nabla w^\varepsilon(t) \|_{1,\Omega_s}^2 \]  

(2.28)

Absorbing the terms involving \( \nabla u^\varepsilon \) and \( \nabla w^\varepsilon \) on the right-hand side of (2.28) to the left, and then applying Gronwall’s lemma to the resulting estimate, we obtain (2.27) without the initial history \( z_0 \).

**Remark 2.13** If \( \sigma \in L^2(H_s) \) then a modification of (2.28) yields that the constant on the right-hand side of (2.27) can be taken to be independent of \( \varepsilon \).

We now prove the existence of a pressure term similar to Theorem 2.8. We point out that there is no need for compatibility conditions on the initial data due to the regularizing effect of the strong damping term in the wave equation.

**Theorem 2.14** Assume that \( u_0 \in \mathcal{V}_f, v_1, w_0 \in \mathcal{V}_s, z_0 \in L^2(Q_f), f \in L^2(Q_f), \sigma \in L^2(Q_s) \) and \( \beta \in H^1(H^{-1/2}(I_r)) \). Then the unique weak solution of (2.25) satisfies in addition \( u^\varepsilon \in H^1(H_f) \cap L^\infty(V_f) \) and \( w^\varepsilon \in H^2(H_s) \cap W^{1,\infty}(V_s) \). There is a unique \( p^\varepsilon \in L^2(M) \) such that

\[ (u^\varepsilon, \varphi)_{\Omega_f} + (\nabla u^\varepsilon, \nabla \varphi)_{\Omega_f} - (p^\varepsilon, \text{div} \varphi)_{\Omega_f} + (w^\varepsilon, \varphi)_{\Omega_s} + \varepsilon (\nabla w^\varepsilon, \nabla \varphi)_{\Omega_s} + (w^\varepsilon, \varphi)_{1,\Omega_s} + (\mu w^\varepsilon (- \cdot r), \varphi)_{\Omega_s} = (f, \varphi)_{\Omega_f} + (\beta, \varphi) + (\sigma, \varphi)_{\Omega_s} \]

(2.29)

for each \( \varphi \in W \) and for a.e. \( t \in I \). Moreover, for some \( C_\varepsilon > 0 \) we have

\[ \| u^\varepsilon \|_{H^1(H_f) \cap L^\infty(V_f)} + \| w^\varepsilon \|_{H^2(H_s) \cap W^{1,\infty}(V_s)} + \| p^\varepsilon \|_{L^2(M)} \leq C_\varepsilon \| f \|_{L^2(Q_f)} \]

\[ + C_\varepsilon \left( \| \beta \|_{H^1(H^{-1/2}(I_r))} + \| \sigma \|_{L^2(Q_s)} + \| u_0 \|_{1,\Omega_f} + \| v_0 \|_{1,\Omega_s} + \| w_0 \|_{1,\Omega_s} + \| z_0 \|_{Q_r} \right). \]

**Proof.** From the remarks in the proof of the preceding theorem, it is enough to consider the case \( \mu = 0 \). Testing the fluid and structural equations with \( u^\varepsilon_t \) and \( w^\varepsilon_t \), respectively, provides the energy identity

\[ \frac{d}{dt} \left( \| \nabla u^\varepsilon(t) \|_{\Omega_j}^2 + \| w^\varepsilon(t) \|_{\Omega_s}^2 + \varepsilon \| \nabla w^\varepsilon(t) \|_{\Omega_s}^2 \right) + \| u^\varepsilon_t(t) \|_{\Omega_j}^2 + \| w^\varepsilon_t(t) \|_{\Omega_s}^2 = (f(t), u^\varepsilon(t))_{\Omega_f} + (\beta(t), u^\varepsilon(t)) + (\sigma(t), w^\varepsilon(t))_{\Omega_s}. \]
Using the Cauchy–Schwarz inequality on the right-hand side of this estimate, multiplying the resulting inequality by \( \eta \) and then taking the sum with (2.28) we have

\[
\frac{d}{dt} \left( \| u^f(t) \|_{L^2}^2 + \| w^f(t) \|_{W}^2 + \| w^f(t) \|_{1}^2 \right)
\]

(2.30)

\[
+ \eta \frac{d}{dt} \left( \| \nabla u^f(t) \|_{L^2} + (w^f(t), w^f(t))_{1,\Omega} + \varepsilon \| \nabla w^f(t) \|_{L^2}^2 \right)
\]

\[
+ \frac{\eta}{2} \| u^f(t) \|_{L^2}^2 + \frac{\eta}{2} \| w^f(t) \|_{W}^2 + \frac{\eta}{2} \| \nabla u^f(t) \|_{L^2}^2 + \frac{1}{2} (\varepsilon - 2\eta) \| \nabla w^f(t) \|_{L^2}^2
\]

\[
\leq C_{\varepsilon, \eta} (\| f(t) \|_{L^2}^2 + \| \sigma(t) \|_{L^2}^2 + \| u^f(t) \|_{L^2}^2 + \| w^f(t) \|_{L^2}^2) + \langle \beta(t), u^f(t) \rangle.
\]

Integrating by parts with respect to \( t \) and using the Poincaré inequality we obtain

\[
\int_0^t \langle \beta(s), u^f(s) \rangle \, ds = \langle \beta(t), u^f(t) \rangle - \langle \beta(0), u_0 \rangle - \int_0^t \langle \beta_s(s), u^f(s) \rangle \, ds
\]

\[
\leq C_{\eta} (\| u_0 \|_{1,\Omega}^2 + \| \beta \|_{H^1(H^{-1/2}(I))]}) + \eta \| \nabla u^f(t) \|_{L^2}^2
\]

\[
+ \eta \int_0^t \| \nabla u^f(s) \|_{L^2}^2 \, ds.
\]

(2.31)

We integrate (2.30) over \([0, t]\), use Young’s inequality, Gronwall’s lemma, invoke (2.31), take the supremum over all \( t \in I \) and neglect some non-negative terms on the left-hand side to obtain

\[
\eta \| \nabla u^f \|_{L^\infty(H^j)} + \left( 1 - \frac{\eta}{2\varepsilon} \right) \| \nabla w^f \|_{L^\infty(H^j)} + \frac{\varepsilon}{2} \| \nabla w^f \|_{L^\infty(H^j)} + \frac{\eta}{2} \| u^f \|_{L^2(H^j)} + \frac{\eta}{2} \| w^f \|_{L^2(H^j)}
\]

\[
\leq C_{\eta, \varepsilon} (\| f \|_{L^2(H^j)} + \| \beta \|_{H^1(H^{-1/2}(I))]}) + \| \sigma \|_{L^2(H^j)} + \| w \|_{1,\Omega}^2 + \| v \|_{1,\Omega}^2)
\]

for \( \eta > 0 \) small enough. From this we infer that \( u^f \in L^2(H^j) \), \( u^f \in L^\infty(V_j) \), \( w^f \in L^2(H^j) \) and \( w^f \in L^\infty(V_j) \).

For the existence of the pressure we use the following inf–sup condition as in Du et al. (2004):

\[
\inf_{q \in M \setminus \{0\}} \sup_{\varphi \in H^1(\Omega) \setminus \{0\}} \frac{b(\varphi, q)}{\| \varphi \|_{L^2(\Omega)} \| q \|_{L^2(\Omega)}} \geq c > 0,
\]

(2.32)

where \( b : H^1_0(\Omega) \times M \to \mathbb{R} \) is the bilinear form defined by

\[
b(\varphi, q) = -\int_{\Omega} q \text{ div } \varphi \, dx.
\]

(2.33)
For each $t \in I$, consider the linear functional $\ell_t : H^1_0(\Omega) \to \mathbb{R}$ defined by

$$
\langle \ell_t, \varphi \rangle = (f, \varphi)_{\Omega_f} + \langle \beta, \varphi \rangle + (\sigma, \varphi)_{\Omega_s} - (u^\varepsilon_t, \varphi)_{\Omega_f} - (\nabla u^\varepsilon, \nabla \varphi)_{\Omega_f} - (w^\varepsilon_t, \varphi)_{\Omega_f} - \varepsilon(\nabla w^\varepsilon, \nabla \varphi)_{\Omega_s} - (w^\varepsilon, \varphi)_{1,\Omega_s} - (\mu w^\varepsilon_t (- \cdot \cdot r), \varphi)_{\Omega_s}.
$$

Define the linear operator $B : H^1_0(\Omega) \to M'$ by $\langle B\varphi, q \rangle = b(\varphi, q)$ whose kernel is $V$. Since $(u^\varepsilon, w^\varepsilon)$ is a weak solution of (2.25) satisfying $t \in I$ theorem is similar to the proof of Theorem 2.4 and therefore the details are omitted. This result will be immediately. □

For each $B$ such that $p^\varepsilon \in L^2(Q_f)$ and $w^\varepsilon_t \in L^2(Q_s)$, $\ell_t$ lies in the polar set $\{ \ell \in H^{-1}(\Omega) : \langle \ell, \varphi \rangle = 0 \ \forall \ \varphi \in V \}$ of $V$. From Girault & Raviart (1986, Lemma 4.1) there exists a unique $p^\varepsilon(t) \in M$ such that $B'p^\varepsilon(t) = \ell_t$, where $B'$ is the dual of $B$, that is, $b(\varphi, p^\varepsilon(t)) = \langle \ell_t, \varphi \rangle$ for every $\varphi \in H^1_0(\Omega)$ and $t \in I$. Moreover,

$$
c\|p^\varepsilon(t)\|_{\Omega_f} \leq \|B'\|_{\Omega_f} H^{-1}(\Omega).
$$

Since $t \mapsto \ell_t \in L^2(H^{-1}(\Omega))$ it follows that $p^\varepsilon \in L^2(M)$. The estimate for the pressure follows immediately. □

One may also prove regularity of solutions for the state equation (2.25). The proof of the following theorem is similar to the proof of Theorem 2.4 and therefore the details are omitted. This result will be our basis in deriving error estimates for the numerical approximations in the succeeding sections.

**Theorem 2.15** Assume that $u_0 \in V_f$, $w_0 \in H^2(\Omega_f)$, $v_0 \in V_s$, $z_0 \in L^2(Q_s)$, $f \in L^2(L^2(\Omega_f))$, $\sigma \in L^2(H_s)$, $\beta \in H^1(H^{-1/2}(\Gamma_s)) \cap L^2(H^{1/2}(\Gamma_s))$ and $\Omega_s$ is a $C^2$-domain. Then the weak solution of (2.25) satisfies

$$
u^\varepsilon \in L^2(H^2(\Omega_f)), \quad w^\varepsilon \in H^1(H^2(\Omega_s)), \quad p^\varepsilon \in L^2(H^1(\Omega_s)),
$$

and for some constant $C_\varepsilon > 0$ we have the *a priori* estimate

$$
\|u^\varepsilon\|_{L^2(H^2(\Omega_f))} + \|w^\varepsilon\|_{H^1(H^2(\Omega_s))} + \|p^\varepsilon\|_{L^2(H^1(\Omega_s))} \leq C_\varepsilon (\|f\|_{L^2(L^2(\Omega_f))} + \|\beta\|_{H^1(H^{-1/2}(\Gamma_s)) \cap L^2(H^{1/2}(\Gamma_s))} + \|\sigma\|_{L^2(H_s)})
$$

$$+ C_\varepsilon (\|u_0\|_{1,\Omega_f} + \|v_0\|_{1,\Omega_s} + \|w_0\|_{2,\Omega_s} + \|z_0\|_{\Omega_s}).
$$

In addition, suppose further that $u_0 \in H^2(\Omega_f)$, $v_0 \in H^2(\Omega_s)$, $z_0 \in H^r(\Omega_s)$, $f \in H^1(L^2(\Omega_f))$, $\sigma \in H^1(H_s)$ and the compatibility conditions $u_0 = v_0$ on $\Gamma_s$, $\partial_v w_0 + \varepsilon \partial_v v_0 = \partial_v u_0 - p_0 v$ on $\Gamma_s$ for some $p_0 \in H^1(\Omega_f)$, and $z_0(0) = v_0$ hold. Then $u^\varepsilon \in W^{1,\infty}(H_f) \cap H^1(V_f)$, $w^\varepsilon \in W^{1,\infty}(V_s) \cap W^{2,\infty}(H_s)$ and there exists a constant $C_\varepsilon > 0$ such that

$$
\|u^\varepsilon\|_{W^{1,\infty}(H_f) \cap H^1(V_f)} + \|w^\varepsilon\|_{W^{1,\infty}(V_s) \cap W^{2,\infty}(H_s)} \leq C_\varepsilon (\|f\|_{H^1(L^2(\Omega_f))} + \|\beta\|_{H^1(H^{-1/2}(\Gamma_s)) \cap L^2(H^{1/2}(\Gamma_s))} + \|\sigma\|_{H^1(H_s)})
$$

$$+ C_\varepsilon (\|u_0\|_{2,\Omega_f} + \|v_0\|_{2,\Omega_s} + \|w_0\|_{2,\Omega_s} + \|z_0\|_{H^1(\Omega_s)}).
$$

We end this section by showing the convergence of the regularized state equation to the original one as the parameter $\varepsilon$ tends to 0.
Theorem 2.16  Let \( u_0 \in H_f, v_0 \in H_s, w_0 \in V_s, \) \( z_0 \in L^2(Q_s), f \in L^2(V_f), \) \( \beta \in L^2(H^{-1/2}(\Gamma_s)) \) and \( \sigma \in L^2(H_s) \). The solution of the regularized problem (2.25) tends to the solution of problem (2.19) as \( \varepsilon \to 0 \) in the following sense:

\[
\|u^\varepsilon - u\|_{C(I,H_f) \cap L^2(I,V_f)} + \|w^\varepsilon - w\|_{C(I,H_s) \cap C(I,V_x)} \to 0.
\] (2.35)

Proof.  We adapt the proof in Lions (1971, pp. 352–353) for hyperbolic equations. Let us assume for the moment that the initial conditions and the source terms satisfy the conditions of Theorem 2.8(i). Let \( \xi^\varepsilon = u^\varepsilon \chi_{\Omega_f} + w^\varepsilon \chi_{\Pi_s} \) and \( \xi = u \chi_{\Omega_f} + w \chi_{\Pi_s} \). From the assumptions on the initial data we have \( \xi \in L^2(I,V) \). Let \( N \) be the positive integer such that \( N \varepsilon < T \leq (N + 1)\varepsilon, \) \( J_n = [(n - 1)\varepsilon, n\varepsilon] \) for \( 0 \leq n \leq N \) and \( J_{N+1} = [N\varepsilon, T] \). There exists \( C > 0 \) such that for each \( t \in I \) it holds that

\[
\|u^\varepsilon(t)\|_{L^2_f}^2 + \|w^\varepsilon(t)\|_{L^2_s}^2 + \|w^\varepsilon(t)\|_{H^1_s}^2 + \int_0^t (1 - \varepsilon)\|\nabla u^\varepsilon(s)\|_{L^2_f}^2 ds + \varepsilon \int_0^t \|\nabla \xi^\varepsilon(s)\|_{L^2}^2 ds
\]
\[
\leq C \int_0^t \|f(s)\|_{L^2_f}^2 + \|\beta(s)\|_{H^{-1/2}(\Gamma_s)}^2 + \|\sigma(s)\|_{L^2_s}^2 ds
\]
\[
+ \int_0^t \frac{1}{2} \|\nabla u^\varepsilon(s)\|_{L^2_f}^2 + C\|w^\varepsilon(s)\|_{L^2_s}^2 + \|u_0\|_{L^2_f}^2 + \|w_0\|_{L^2_s}^2 + \|z_0\|_{Q_s}^2.
\]

Applying Gronwall’s inequality, we see that \( u^\varepsilon, w^\varepsilon, w^\varepsilon_t \) and \( \sqrt{\varepsilon} \xi^\varepsilon \) are bounded in \( L^\infty(I,H_f) \cap L^2(I,V_f) \), \( L^\infty(I,V_s), L^\infty(I,H_s) \) and \( L^2(I,V) \), respectively. Thus, for a subsequence and some \( \tilde{u} \in L^\infty(I,H_f) \cap L^2(I,V_f) \) and \( \tilde{w} \in L^\infty(I,V_s) \cap W^{1,\infty}(I,H_s) \) we have

\[
\begin{align*}
\{ & u^\varepsilon \to \tilde{u} & \text{ in } & L^2(I,V_f), \\
& w^\varepsilon \to \tilde{w} & \text{ in } & L^2(I,V_s), \\
& w^\varepsilon_t \to \tilde{w}_t & \text{ in } & L^2(I,H_s), \\
& \varepsilon \xi^\varepsilon \to 0 & \text{ in } & L^2(I,V).
\end{align*}
\] (2.36)

Passing to the limit \( \varepsilon \to 0 \) in the weak form of the regularized problem leads to \( \xi_t^\varepsilon \to \tilde{\xi}_t \) in \( L^2(I,V') \), where \( \tilde{\xi} = \tilde{u} \chi_{\Omega_f} + \tilde{w} \chi_{\Pi_s} \). Therefore, we have \( \xi^\varepsilon(0) \to \tilde{\xi}(0) \) in \( V' \) and \( w^\varepsilon(0) \to \tilde{w}(0) \) in \( H_s \). Thus, \( \tilde{u}(0) = u_0, \tilde{w}_t(0) = v_0 \) and \( \tilde{w}(0) = w_0 \). Passing to the limit in the weak formulation of \( (u^\varepsilon, w^\varepsilon) \) and using (2.36), we see that \( (\tilde{u}, \tilde{w}) \) is a weak solution to (2.19), and by uniqueness we have \( (\tilde{u}, \tilde{w}) = (u, w) \).

Now let us prove strong convergence, and for this purpose we define

\[
N_\varepsilon(t) = \|u^\varepsilon(t) - u(t)\|_{L^2_f}^2 + \|w^\varepsilon(t) - w_t(t)\|_{L^2_s}^2 + \|w^\varepsilon(t) - w(t)\|_{H^1_s}^2
\]
\[
+ 2\varepsilon \int_0^t \|\nabla u^\varepsilon(s) - \nabla u(s)\|_{L^2_s}^2 ds + 2\mu \int_0^t \|w^\varepsilon_t(s) - w_t(s)\|_{L^2_s}^2 ds
\]
\[
+ 2\varepsilon \int_0^t \|\nabla w^\varepsilon(s)\|_{L^2_s}^2 ds.
\] (2.37)
In this section we discuss the optimal control problems (1.1) and (1.2), its regularization (1.1)–(1.4) and define the control-to-state operator

\[ S_j(q) = \int_0^t (\nabla w_i(s), \nabla v_i(s))_\Omega^2 \, ds \]

For each \( t \in J_1 \) we have \( w_i^t(t-r) = w_i(t-r) \) since (2.1) and (2.19) have the same initial history. In particular, the integral involving the delay term in (2.37) vanishes for every \( t \in J_1 \). From (2.38) and the first and fourth lines of (2.36) it follows that

\[ \|N_\varepsilon\|_{L^\infty(I)} \to 0 \quad \text{as} \quad \varepsilon \to 0. \]  

(2.39)

Using (2.37) for \( t \in J_1 \) and (2.39), we can see that (2.35) is satisfied with \( I \) replaced by \( J_1 \). Now using this information and estimating the delay term in (2.37) for \( t \in J_2 \) by Young’s inequality, we obtain that (2.35) holds on the interval \( J_2 \) as well. Continuing this process on the intervals \( J_3, \ldots, J_{N+1} \), we obtain (2.35), however, under the additional regularity assumptions on the data. For initial data and source terms that merely satisfy those in the statement of the theorem, one can proceed by a standard density argument and apply Remark 2.13. \( \square \)

3. Analysis of the optimal control problems

In this section we discuss the optimal control problems (1.1) and (1.2), its regularization (1.1)–(1.4) and provide the necessary optimality conditions, which are also sufficient due to the linear–quadratic structure of the problem. Consider the optimal control problem

\[ \min_{q \in Q} J(u, w, q) = G(u, w) + \frac{\alpha}{2} \|q\|^2_Q \quad \text{subject to} \quad (1.2), \]  

(3.1)

where \( G \) is given by (1.3), \( \gamma_f, \gamma_s \geq 0 \) for \( i = 1, 2, 3 \) and \( \alpha > 0 \). Further, \( u_d \in L^2(H_f) \), \( v_d \in L^2(H_s) \) and either \( w_d \in L^2(V_s) \) if \( \gamma_{s3} > 0 \) or \( w_d \in L^2(H_s) \) if \( \gamma_{s3} = 0 \). For the rest of this paper \( Q \) will be a Hilbert space of control, \( B_f : Q \to L^2(Q_f) \) and \( B_s : Q \to L^2(Q_s) \) are bounded linear operators. The following theorem can be shown using standard methods in linear–quadratic optimal control problems and thus the details are omitted; see Lions (1971); Tröltzsch (2010).

Theorem 3.1 Suppose that \( u_0 \in H_f, v_0 \in H_s, w_0 \in V_s \) and \( z_0 \in L^2(Q_s) \). Then the optimal control problem (3.1) has a unique solution.

In what follows we derive the necessary optimality conditions to (3.1). First, let us consider the case where \( \gamma_{s3} = 0 \). Introduce the Hilbert space \( X_0 = L^2(H_f) \times L^2(H_s) \times L^2(H_s) \) with the weighted norm

\[ \|(u, v, w)\|_{X_0}^2 = \gamma_f \|u\|_{L^2(H_f)}^2 + \gamma_s \|v\|_{L^2(H_s)}^2 + \gamma_s \|w\|_{L^2(H_s)}^2 \]

and define the control-to-state operator \( S_0 : Q \to X_0 \) by \( S_0 q = (u(q), w_i(q), w(q)) \), where \((u(q), w(q))\) is the weak solution of (1.2) for a given control \( q \in Q \). One can easily see that \( S \) is affine and continuous. Define the reduced cost functional \( j_0 : Q \to \mathbb{R} \) by

\[ j_0(q) = \frac{1}{2} \|S_0 q - z_d\|_{X_0}^2 + \frac{\alpha}{2} \|q\|^2_Q. \]
where \( z_d = (u_d, v_d, w_d) \). The derivative of \( j_0 \) at \( q \) in the direction of \( \delta q \in Q \) is given by

\[
J_0(q)\delta q = (S_0q - z_d, S_0\delta q)_{X_0} + \alpha(q, \delta q)_Q.
\]  

(3.2)

If \( q^* \) is the solution to (3.1) then we must have \( J_0(q^*)\delta q = 0 \) for every \( \delta q \in Q \).

**Theorem 3.2** Suppose that \( \gamma_{s3} = 0 \). The optimal solution \( q^* \) to (3.1) is characterized by the following necessary conditions.

(i) The optimal state \( (u^*, w^*) = (u(q^*), w(q^*)) \) is the weak solution of (2.19) with \( f = B_f q^* \) and \( \sigma = B_s q^* \).

(ii) The adjoint state \( (\varphi^*, \psi^*) = (\varphi(q^*), \psi(q^*)) \) is the weak solution of (2.22) with \( g = \gamma_f (u^* - u_d) \) and \( \kappa = \gamma_{s1}(w^* - v_d) + \int_1^T \gamma_{s2}(w^* - w_d)(s) \, ds \).

(iii) \( q^* = -\frac{1}{\alpha}(B_f^* \varphi^* - B_s^* \psi^*) \).

**Proof.** The proof is based on a density argument. The idea is to approximate the data and control so that the corresponding global velocity field is an admissible test function. Take a sequence \( (z^*_n)_n \) in \( L^2(L^2(\Omega)) \) such that \( f_n := z^*_n|_{\Omega_f} \in H^1(V_f) \), \( \sigma_n := z^*_n|_{\Omega_s} \in H^1(H_s) \) and \( z^*_n \to B_f q^* \chi_{\Omega_f} + B_s q^* \chi_{\Omega_s} \) in \( L^2(L^2(\Omega)) \). Likewise, take a sequence \( (u_{0n}, w_{0n}, v_{0n}, z_{0n})_n \) of data satisfying the conditions in Theorem 2.8(i) so that \( (u_{0n}, w_{0n}, v_{0n}, z_{0n}) \to (u_0, w_0, v_0, z_0) \) in \( H_f \times V_s \times H_s \times L^2(Q_r) \). Let \( (u_n, w_n) \) be the weak solution of (2.19) with \( f = f_n, \sigma = \sigma_n \) and initial data \( (u_{0n}, w_{0n}, v_{0n}, z_{0n}) \). From Theorem 2.8(i), \( u_n \in W^{1,\infty}(H_f), u_{0n} \in W^{\infty}(I, H_f) \) and the continuity of the solution operator implies that

\[
(u_n, w_n) \to (u^*, w^*) \text{ in } L^2(H_f) \times L^2(H_s).
\]

(3.3)

On the other hand, take sequences \( (u_{dn})_n \subset H^1(H_f) \) and \( (v_{dn})_n \subset H^1(H_s) \) so that \( u_{dn} \to u_d \) in \( L^2(H_f) \) and \( v_{dn} \to v_d \) in \( L^2(H_s) \). If \( (\varphi_n, \psi_n) \) is the weak solution of (2.22) with right-hand sides \( g = \gamma_f (u_n - u_d) \) and \( \kappa = \gamma_{s1}(w_{dn} - v_d) + \int_1^T \gamma_{s2}(w^* - w_d)(s) \, ds \) then from Theorem 2.8(i) once more we also have \( \varphi_n \in W^{1,\infty}(H_f), \psi_n \in W^{\infty}(H_s) \), and we obtain from (3.3) that

\[
(\varphi_n, \psi_n) \to (\varphi^*, \psi^*) \text{ in } L^2(H_f) \times L^2(H_s).
\]

(3.4)

Given \( q \in Q \), let \( (u, w) \) be the solution of (1.2). We also consider a sequence \( (z_n)_n \) in \( L^2(L^2(\Omega)) \) such that \( z_n|_{\Omega_f} \in H^1(V_f), z_n|_{\Omega_s} \in H^1(H_s) \) and \( z_n \to B_f q \chi_{\Omega_f} + B_s q \chi_{\Omega_s} \) in \( L^2(L^2(\Omega)) \). Denote by \( (u^n, w^n) \) the solution of (2.19) with source terms \( f = z_n|_{\Omega_f}, g = z_n|_{\Omega_s} \) and with the same initial data as with \( (u_n, w_n) \). Set \( \delta u_n = u_n - u^n \) and \( \delta w_n = w_n - w^n \). Using \( \delta u_n \chi_{\Omega_f} + \delta w_n \chi_{\Omega_s} \in L^2(V) \) as a test function in the weak formulation of the approximated adjoint state \( (\varphi_n, \psi_n) \) in (2.22), we obtain

\[
\int_I \gamma_f (u_n - u_{dn})_{\Omega_f} + \gamma_{s1}(w_{dn} - v_{dn})_{\Omega_s} \, dt + \int_I \int_T \gamma_{s2}(w^* - w_d)_{\Omega_s} \, ds \, dt
\]

\[
= \int_I (\varphi_n, \psi_n)_{\Omega_f} + (\nabla \varphi_n, \nabla u_n)_{\Omega_f} \, dt
\]

\[
+ \int_I (\psi_n, \delta w_n)_{\Omega_s} + (\psi_n, \delta w_n)_{1,\Omega_s} - (\mu \psi_n (\cdot + r), \delta w_n)_{\Omega_s} \, dt.
\]

(3.5)
Integrating by parts with respect to time and using $\psi_n(T) = \delta w_n(0) = 0$, we have

$$\int_I (\psi_n, \delta w_n)_{1, \Omega}, dt = - \int_I (\psi_n, \delta w_n)_{1, \Omega}, dr.$$  \hspace{1cm} (3.6)

In a similar fashion we obtain

$$\int_I \int_I^T (w^* - w_d, \delta w_n)_{\Omega}, ds dt = \int_I (w^* - w_d, \delta w_n)_{\Omega}, dr.$$  \hspace{1cm} (3.7)

On the other hand, since $\psi_n(\theta) = 0$ for $\theta \in (T, T + r)$ and $\delta w_n(\theta) = 0$ for $\theta \in (-r, 0)$,

$$\int_I (\psi_n(\cdot + r), \delta w_n)_{\Omega}, dr = \int_I (\psi_n, \delta w_n(\cdot - r))_{\Omega}, dr.$$  \hspace{1cm} (3.8)

Substituting (3.6)–(3.8) in (3.5), integrating by parts with respect to time and using the fact that $(\varphi_n, \psi_n)$ and $(\delta u_n, \delta w_n)$ vanish at $t = T$ and $t = 0$, respectively, yields

$$\int_I \gamma_f (u_n - u_{dn}, \delta u_n)_{\Omega_f} + \gamma_s^1 (w_n - v_{dn}, \delta w_n)_{\Omega_s} + \gamma_s^2 (w^* - w_d, \delta w_n)_{\Omega_s}, dr$$

$$= \int_I (\varphi_n, \delta u_n)_{\Omega_f} + (\nabla \varphi_n, \nabla \delta u_n)_{\Omega_f} - (\psi_n, \delta w_n(\cdot - r))_{\Omega_s}, dr$$

$$- \int_I (\psi_n, \delta w_n)_{1, \Omega_s} - (\psi_n, \delta w_n(\cdot - r))_{\Omega_s}, dr$$

$$= \int_I (\varphi_n, z^* - z_n)_{\Omega_f} - (\psi_n, z^*_n - z_n)_{\Omega_s}, dr.$$

Passing to the limit $n \rightarrow \infty$ and using (3.3) and (3.4), we have

$$\int_I \gamma_f (u^* - u_d, u^* - u)_{\Omega_f} + \gamma_s^1 (w^*_r - v_d, w^*_r - w)_{\Omega_s} + \gamma_s^2 (w^* - w_d, w^* - w)_{\Omega_s}, dt$$

$$= \int_I (\varphi^*, B_f (q^* - q))_{\Omega_f} - (\psi^*_f, B_s (q^* - q))_{\Omega_s}, dt.$$

However, the left-hand side is equal to $(Sq^* - z_d, S(q^* - q))_{\chi_0}$ and thus $(B_f^* \varphi^* - B_s^* \psi^* + \alpha q^* - q)_{\Omega} = 0$ for every $q \in \mathcal{Q}$, which implies (iii). \hspace{1cm} \Box

Next we consider the case where $\gamma_{s3} > 0$. In this case we define the space $X = L^2(H_f) \times L^2(H_s) \times L^2(V_s)$ endowed with the weighted norm

$$\|(u, v, w)\|_X^2 = \gamma_f \|u\|_{\mathcal{Q}_f}^2 + \gamma_s^1 \|v\|_{\mathcal{Q}_s}^2 + \gamma_s^2 \|w\|_{\mathcal{Q}_s}^2 + \gamma_{s3} \|\nabla w\|_{\mathcal{Q}_s}^2,$$
and the control-to-state operator $S : Q \to X$ by $S q = (u, w, w)$, where $(u, w)$ is the weak solution of (1.2) and $z_d = (u_d, v_d, w_d)$. Introduce the reduced cost functional $j : Q \to \mathbb{R}$ by

$$j(q) = \frac{1}{2} \| S q - z_d \|_X^2 + \frac{\alpha}{2} \| q \|_Q^2.$$ 

The directional derivative of the reduced cost is given by

$$j'(q) \delta q = (S q - z_d, S \delta q) + \alpha(\delta q, \delta q)_Q, \quad \delta q \in Q.$$

(3.9)

Define the bounded linear operator $\tilde{\Delta} : V_s \to V'_s$ as

$$\langle \tilde{\Delta} \varphi, \psi \rangle = -\langle \nabla \varphi, \nabla \psi \rangle_{\Omega_s}, \quad \forall \varphi, \psi \in V_s.$$

**Theorem 3.3** Let $\gamma_{s3} > 0$ and $q^*$ be the unique minimizer of the problem (3.1). Then $q^*$ is characterized by the following necessary conditions.

(i) The optimal state $(u^*, w^*) = (u(q^*), w(q^*))$ is the weak solution of (2.19) with $f = B f q^*$ and $\sigma = B_s q^*$.

(ii) The adjoint state $(\varphi^*, \psi^*) = (\varphi(q^*), \psi(q^*))$ is the very weak solution of (2.22) with $g = \gamma_f (u^* - u_d)$ and

$$\kappa = \gamma_{s1} (w_t^* - v_d) + \int_t^T \gamma_{s2} (w^*(s) - w_d(s)) \, ds - \int_t^T \gamma_{s3} \tilde{\Delta} (w^*(s) - w_d(s)) \, ds.$$

(iii) $q^* = -\frac{1}{\alpha} (B_f \varphi^* - B_s \psi^*_t)$.

**Proof.** Observe that $g \in L^2(H_f)$ and we can decompose $\kappa = \kappa_1 + \kappa_2$ where $\kappa_1 = \gamma_{s1} (w_t^* - v_d) \in L^2(H_s)$ and $\kappa_2 = \int_t^T \gamma_{s2} (w^*(s) - w_d(s)) \, ds - \int_t^T \gamma_{s3} \tilde{\Delta} (w^*(s) - w_d(s)) \, ds \in H^1(V_s)$. Let $(\varphi^*, \psi^*)$ be as described in the statement of the theorem. Due to the homogeneous terminal data, dual history of the adjoint equations and the facts that $\kappa_2(T) = 0$ and $\kappa_2 = -\gamma_{s2} (w_t^* - w_d) + \gamma_{s3} \tilde{\Delta} (w^* - w_d)$, we have for every $(f, \sigma) \in L^2(H_f) \times L^2(H_s)$,

$$\int_I (\varphi^*, f)_{\Omega_f} - (\psi^*_t, \sigma)_{\Omega_s} \, dt = \int_I \gamma_f (u^* - u_d, u)_{\Omega_f} \, dt + \int_I \gamma_{s1} (w_t^* - v_d, w)_{\Omega_s} \, dt + \gamma_{s2} (w^* - w_d, w)_{\Omega_s} + \gamma_{s3} (\nabla w^* - \nabla w_d, w)_{\Omega_s} \, dt,$$

where $(u, w)$ is the solution of (2.19) with $\beta = 0$ and homogeneous initial data and history. Given $q \in Q$ we choose $f = B f (q^* - q)$ and $\sigma = B_s (q^* - q)$ to obtain

$$\int_I (\varphi^*, B_f (q^* - q))_{\Omega_f} - (\psi^*_t, B_s (q^* - q))_{\Omega_s} \, dt = \int_I (S q^* - z_d, S (q^* - q))_{\chi} \, dt.$$

Comparing this with (3.9) and using the condition $j'(q^*) (q^* - q) = 0$ for every $q \in Q$, we see that $B_f \varphi^* - B_s \psi^*_t + \alpha q^* = 0$ and hence (iii). \qed
With additional regularity on the data and on the desired states, one may use the weak formulation of the adjoint equation instead of the very weak one.

**Theorem 3.4** Suppose that \((u_0, v_0, w_0, z_0)\) satisfies the conditions of Theorem 2.8(ii), \(u_d \in H^1(V) \cap L^2(H_g), v_d \in H^1(H)\) and \(w_d \in L^2(H^2(\Omega_g)).\) If the unique minimizer \(q^*\) of (3.1) satisfies \(B_f q^* \in H^1(V)\) and \(B_q q^* \in H^1(Q_g)\) then \(q^* = -\frac{1}{\alpha}(B_f \varphi^* - B_q \psi^*),\) where \((\varphi^*, \psi^*)\) is the weak solution of (2.22) with \(g = g_f(u^* - u_d),\)

\[
\kappa = g_{s1}(w^* - v_d) + \int_t^T g_{s2}(w^*(s) - w_d(s)) - g_{s3} \Delta(w^*(s) - w_d(s)) \, ds,
\]

and the boundary conditions

\[
\partial_v \psi - \partial_v \varphi + \pi v = \int_t^T g_{s3} \partial_v (w^*(s) - w_d(s)) \, ds,
\]

where \((u^*, w^*)\) is the corresponding optimal state.

**Proof.** Let \(\beta\) denote the right-hand side of (3.10). The above assumptions imply that \(u^* \in H^1(V)\) and \(w^* \in L^2(H^2(\Omega_g)) \cap H^2(\Omega_g).\) As a consequence we have \(g \in H^1(V) \cap L^2(\Omega_g), \kappa \in H^1(\Omega_g)\) and \(\beta \in H^1(\Omega_g).\) Let \((\varphi^*, \psi^*)\) be described as above. The equation \(q^* = -\frac{1}{\alpha}(B_f \varphi^* - B_q \psi^*)\) can be derived using a similar argument to the proof of Theorem 3.2 and using the equation

\[
\int_I (\kappa, \delta w)_{\Omega_i} + (\beta, \delta w)_{\Omega_i} \, dt = \int_I g_{s1}(w^* - v_d, \delta w)_{\Omega_i} \, dt
\]

\[
+ \int_I g_{s2}(w^* - w_d, \delta w)_{\Omega_i} + g_{s3}(\nabla w^* - \nabla w_d, \nabla \delta w)_{\Omega_i} \, dt.
\]

We would like to point out that since the states are smooth one may proceed directly without the use of an approximation argument.

Now we consider the regularized optimal control problem

\[
\min_{q^* \in Q} J(u^\varepsilon, w^\varepsilon, q^*) = G(u^\varepsilon, w^\varepsilon) + \frac{\alpha}{2} \|q^*\|^2_Q \quad \text{subject to (1.4),}
\]

where \(G\) is given by (1.3) and \(g_i, g_{s1} \geq 0\) for \(i = 1, 2, 3\) and \(\alpha > 0.\) Let \(S_e : Q \rightarrow X\) be the control-to-state operator \(S_e(q^*) = (u^\varepsilon(q^*), w^\varepsilon(q^*), w^\varepsilon(q^*))\) where \((u^\varepsilon, w^\varepsilon)\) is the solution of (1.4). Define the reduced cost functional \(J_e : Q \rightarrow \mathbb{R}\) by

\[
J_e(q^*) = \frac{1}{2} \|S_e q^* - z_d\|^2_X + \frac{\alpha}{2} \|q^*\|^2_Q,
\]

where \(z_d = (u_d, v_d, w_d).\) Again, existence follows from standard techniques for linear–quadratic optimal control problems.
Theorem 3.5 Assume that $u_0 \in H_f, w_0 \in V_s, v_0 \in H_s, z_0 \in L^2(Q_r), u_d \in L^2(H_f), v_d \in L^2(H_s)$, and either $w_d \in L^2(H_s)$ if $\gamma_{d3} = 0$ or $w_d \in L^2(V_s)$ if $\gamma_{d3} > 0$. Then (3.11) has a unique minimizer.

We now prove the convergence of the optimal controls and optimal states for (3.1) and (3.11).

Theorem 3.6 If $q^*$ and $q^*_\varepsilon$ are the solutions to (3.1) and (3.11), respectively, then

$$
\|q^*_\varepsilon - q^*\|_Q \to 0 \text{ as } \varepsilon \to 0,
$$

(3.12)

$$
j_{\varepsilon}(q^*_\varepsilon) \to j(q^*) \text{ as } \varepsilon \to 0.
$$

(3.13)

Furthermore, if $(u^*, w^*)$ and $(u^*_\varepsilon, w^*_\varepsilon)$ are the corresponding optimal states then

$$
\|u^*_\varepsilon - u^*\|_{L^2(V_f)} + \|w^*_\varepsilon - w^*\|_{L^2(V_s) \cap H^1(H_s)} \to 0 \text{ as } \varepsilon \to 0.
$$

(3.14)

Proof. We adapt the proof in Lions (1971, pp. 354–355). First, by optimality of $q^*_\varepsilon$, we have $j_{\varepsilon}(q^*_\varepsilon) \leq j_{\varepsilon}(q^*)$. Taking the limit superior and using Theorem 2.16 we have

$$
\limsup_{\varepsilon \to 0} j_{\varepsilon}(q^*_\varepsilon) \leq \limsup_{\varepsilon \to 0} j_{\varepsilon}(q^*) = j(q^*).
$$

Since $\|q^*_\varepsilon\|_Q^2 \leq \frac{2}{\alpha} j_{\varepsilon}(q^*)$ the sequence $(q^*_\varepsilon)_\varepsilon$ is bounded in $Q$, so that for a subsequence we have $q^*_\varepsilon \rightharpoonup \tilde{q}$ in $Q$ for some $\tilde{q} \in Q$. Applying a similar proof to that of Theorem 2.16 yields the weak convergence

$$
(u^*_\varepsilon, w^*_\varepsilon) \rightharpoonup (\tilde{u}, \tilde{w}) = (u(\tilde{q}), w(\tilde{q})) \text{ in } L^2(V_s) \times [L^2(V_s) \cap H^1(H_s)].
$$

(3.15)

By weak lower semicontinuity of the norm we have

$$
j(q^*) \leq j(\tilde{q}) \leq \liminf_{\varepsilon \to 0} j_{\varepsilon}(q^*_\varepsilon).
$$

Thus, we have $j_{\varepsilon}(q^*_\varepsilon) \to j(q^*)$, and as a consequence it holds that $q^* = \tilde{q}$ by uniqueness of the solution to (3.1). This proves (3.13).

To prove (3.12), let $(\tilde{u}^\varepsilon, \tilde{w}^\varepsilon)$ and $(\tilde{u}, \tilde{w})$ be the weak solutions to (1.2) and (1.4) with $q^\varepsilon = 0$ and $q = 0$, respectively. Define $z^* = (u^*, w^*_t, w^*)$, $z^*_\varepsilon = (u^*_\varepsilon, w^*_t, w^*_\varepsilon)$, $\tilde{z}_\varepsilon = (\tilde{u}_\varepsilon, \tilde{w}^\varepsilon)$ and $\tilde{z} = (\tilde{u}, \tilde{w})$. Then $\tilde{z}_\varepsilon \to \tilde{z}$ in $X$ as $\varepsilon \to 0$ from Theorem 2.16. Passing to the limit $\varepsilon \to 0$ and using (3.13) and (3.15) we obtain

$$
\frac{1}{2} \|z^*_\varepsilon - \tilde{z}\|_X^2 + \frac{\alpha}{2} \|q^*_\varepsilon\|_Q^2 = j_{\varepsilon}(q^*_\varepsilon) - (z^*_\varepsilon - \tilde{z}_\varepsilon - z_d)_X + j_{\varepsilon}(0)
$$

$$
\to j(q^*) - (z^* - z_d, \tilde{z} - z_d)_X + j(0) = \frac{1}{2} \|z^* - \tilde{z}\|_X^2 + \frac{\alpha}{2} \|q^*\|_Q^2.
$$
The latter is a norm equivalent to the norm in \( Q \) since the solution operators \( S_\varepsilon \) and \( S \) are affine and continuous, and \( z_\varepsilon^*, \tilde{z}_\varepsilon, z^* \) and \( \tilde{z} \) have the same initial data and history. Together with the weak convergence of \( q_\varepsilon^* \) to \( q^* \) in \( Q \), this proves (3.12). Finally, (3.14) is a consequence of (3.12) and of arguments similar to those in the proof of Theorem 2.16.

The necessary optimality conditions corresponding to (3.11) can be established as in the previous discussions. For completeness we present them below.

**Theorem 3.7** Suppose that \( u_0 \in V_f, w_0 \in V_s, v_0 \in V_s, z_0 \in L^2(Q_r), u_d \in L^2(H_f), v_d \in L^2(H_s), \) and either \( w_d \in L^2(H_s) \) if \( \gamma_{s3} = 0 \) or \( w_d \in L^2(V_s) \) if \( \gamma_{s3} > 0 \). Then the unique minimizer \( q_\varepsilon^* \) of (3.11) is characterized by the following optimality conditions.

(i) The optimal state \((u_\varepsilon^*, w_\varepsilon^*) = (u^\varepsilon(q_\varepsilon^*), w^\varepsilon(q_\varepsilon^*))\) is the weak solution of (2.25).

(ii) The adjoint state \((\varphi^\varepsilon, \psi^\varepsilon, \varphi_\varepsilon^*, \psi_\varepsilon^*) = (\varphi^\varepsilon(q_\varepsilon^*), \psi^\varepsilon(q_\varepsilon^*))\) is the weak solution of

\[
\begin{align*}
-\varphi_t^\varepsilon - \Delta \varphi^\varepsilon + \nabla \pi^\varepsilon = g_\varepsilon & \quad \text{in } Q_f, \\
\text{div } \varphi_\varepsilon = 0 & \quad \text{in } Q_f, \\
\varphi_\varepsilon = 0 & \quad \text{on } \Sigma_f, \\
\varphi^\varepsilon = -\psi^\varepsilon_\varepsilon & \quad \text{on } \Sigma_s, \\
\psi_\varepsilon^\varepsilon - \Delta \psi^\varepsilon + \varepsilon \Delta \varphi_\varepsilon^\varepsilon + \psi^\varepsilon - \mu \varphi^\varepsilon(-r) = \kappa_\varepsilon & \quad \text{in } Q_s, \\
\partial_v \varphi^\varepsilon - \varepsilon \partial_v \varphi_\varepsilon^\varepsilon - \partial_v \psi^\varepsilon + \pi^\varepsilon \nu = 0 & \quad \text{on } \Sigma_s, \\
\varphi^\varepsilon(T) = 0 & \quad \text{in } \Omega_f, \\
\psi^\varepsilon(T) = 0 & \quad \text{in } \Omega_s, \\
\psi_\varepsilon^\varepsilon = 0 & \quad \text{in } (T, T + r) \times \Omega_s,
\end{align*}
\]

where \( g_\varepsilon = \gamma_f(u^*_\varepsilon - u_d) \) and \( \kappa_\varepsilon = \gamma_{s1}(w^*_\varepsilon - v_d) + J_f^T \gamma_{s2}(w^*_\varepsilon(s) - w_d(s)) - \gamma_{s3} \tilde{A}(w^*_\varepsilon - w_d)(s) \),

(iii) \( q_\varepsilon^* = -\frac{1}{\alpha}(B_f^* \varphi^\varepsilon_\varepsilon - B_s^* \psi^\varepsilon_\varepsilon). \)

We close this section by proving regularity of the optimal controls to (3.11) under the choice of the control space \( Q = L^2(I \times \Omega) \). For regularity results pertaining to problem (3.1) without delay, with \( Q = L^2(I) \) or \( Q = L^2(\Omega) \), \( \gamma_{s1} = \gamma_{s3} = 0 \) and under particular control operators \( B_f \) and \( B_s \) we refer to Failer et al. (2016).

**Corollary 3.8** Suppose that the initial data and desired states satisfy the condition of Theorem 3.7, and in addition, it holds that \( w_0 \in H^2(\Omega_s) \) and \( w_d \in L^2(H^2(\Omega_s)). \) Let \( Q = L^2(L^2(\Omega)), B_f q = q \chi_{\Omega_f} \) and \( B_s q = q \chi_{\Omega_s} \) for each \( q \in Q \). Then the optimal solution \( q^*_\varepsilon \) to (3.11) satisfies \( q^*_\varepsilon \in L^2(V) \cap H^1(L^2(\Omega)), q^*_\varepsilon|_{I \times \Omega_f} \in L^2(H^2(\Omega_f)) \) and \( q^*_\varepsilon|_{I \times \Omega_s} \in L^2(H^2(\Omega_s)). \)
Proof. Let \( g_ε \) be the function given in Theorem 3.7. According to the choice of the control space and control operators, \( B_f^* \varphi \) and \( B_q^* \psi \) are the extensions of \( \varphi \in L^2(Ω_f) \) and \( \psi \in L^2(Ω_s) \) by zero outside \( Ω_f \) and \( Ω_s \), respectively. Using Green’s identity on the variational form of the fifth equation of (3.16) the structural component for the adjoint problem can be rewritten as

\[
\begin{aligned}
\psi_{tt}^ε - \Delta \psi_t^ε + \varepsilon \Delta \psi_t^ε + \psi^ε - \mu \psi_t^ε(-r) &= \kappa_ε \quad \text{in } Q_s, \\
\partial_t \psi^ε - \varepsilon \partial_t \psi_t^ε - \partial_t \varphi^ε + \pi^ε v &= \beta_ε \quad \text{on } \Sigma_s,
\end{aligned}
\] (3.17)

where

\[
\beta_ε = \int_I T \gamma_s t \partial_t(w^ε(s) - w_d(s)) \, ds,
\]

\[
\kappa_ε = \gamma_s 1(w_{et} - v_d) + \int_I T \gamma_s 2(w^ε(s) - w_d(s)) - \gamma_s \Delta(w^ε - w_d)(s) \, ds.
\]

Note that \( g_ε \in L^2(\Omega_f) \), \( \kappa_ε \in L^2(H_s) \) and from Theorem 2.15 we have \( \beta_ε \in H^1(\Omega_f) \). As a consequence, \( \varphi^ε \in H^1(\Omega_f) \cap L^2(\Omega_s) \) and \( \psi^ε \in H^1(\Omega_s) \), by the dual version of the first part of Theorem 2.15. The result now follows from the identity \( q^ε = -\frac{1}{\alpha}(\varphi^ε \chi_{\Omega_f} - \psi^ε \chi_{\Omega_s}) \).

The above corollary can be adjusted to the case where the control is acting only on a subset of the fluid domain \( \Omega_f \) and \( B_f q = \chi_{\omega_f} q \) where \( \omega_f \subset \Omega_f \) and \( B_s = 0 \) or on a subset of the structure domain \( \Omega_s \) and \( B_s q = \chi_{\omega_s} q \) where \( \omega_s \subset \Omega_s \).

4. Symmetric formulations of the state and adjoint equations

Inspired by the work in Failer et al. (2016) we shall rewrite the variational equations for (1.4) and the associated adjoint system in symmetric form. The advantage of this formulation is that the nonlocal-in-time terms appearing on the right-hand side of the adjoint equation will be eliminated, leading to a straightforward application of time-advancing schemes. In the present and succeeding sections we will use the control space \( Q = L^2(I \times \Omega) \) and the control operators \( B_f q = q \chi_{\Omega_f} \) and \( B_f q = q \chi_{\Omega_s} \) for \( q \in Q \). From now on we drop the superscripts \( \varepsilon > 0 \) in the notation. For the rest of the paper we assume that the initial data and the desired states satisfy the following regularity conditions.

(A) It holds that \( u_0 \in V_f \cap H^2(\Omega_f) \), \( w_0, v_0 \in H^2(\Omega_s) \), \( z_0 \in H^1(V_s) \cap L^2(H^2(\Omega_s)) \), \( u_d \in H^1(\Omega_f) \cap L^2(H^2(\Omega_f)) \), \( v_d, w_d \in H^1(\Omega_s) \cap L^2(H^2(\Omega_s)) \) and \( w_d \in H^1(\Omega_s) \cap L^2(H^2(\Omega_s)) \). Furthermore, the compatibility conditions \( u_0 = v_0 \) on \( \Gamma_s \), \( \partial_v w_0 + \varepsilon \partial_v v_0 = \partial_u u_0 - p_0 v \) for some \( p_0 \in H^1(\Omega_f) \), \( z_0(0) = v_0 \) and \( \int_{\Gamma_s} z_0(\theta) \cdot v \, dx = 0 \) for every \( \theta \in I_\varepsilon \) are satisfied.

Hypothesis (A) implies that Corollary 3.8 is applicable, and in particular, the wave component of the adjoint equation is equivalent to (3.17). In order to have a unified treatment both for the state and
adjoint equations we consider the following system:

\[
\begin{cases}
    u_t - \Delta u + \nabla p = f_1 & \text{in } Q_f, \\
    \text{div } u = 0 & \text{in } Q_f, \\
    u = 0 & \text{on } \Sigma_f, \\
    u = w_i & \text{on } \Sigma_s, \\
    w_{tt} - \Delta w - \varepsilon \Delta w_t + w + \mu w_t (\cdot - r) = f_2 + \int_0^t (g_1 - \Delta g_2) \, ds & \text{in } Q_s, \\
    \partial_t w + \varepsilon \partial_t w_t - \partial_t u + pv = \int_0^t \partial_t g_2 \, ds & \text{on } \Sigma_s, \\
    u(0) = u_0 & \text{in } \Omega_f, \\
    w(0) = w_0, \quad w_t(0) = v_0 & \text{in } \Omega_s, \\
    \prescript{}{w}{t} = z_0 & \text{in } \Omega_r.
\end{cases}
\]

Indeed, by taking \( f_1 = q\chi_{\Omega_f}, f_2 = q\chi_{\Omega_s} \) and \( g_1 = g_2 = 0 \) we recover the state equation (1.2). On the other hand, replacing \((u, w, p)\) by \((\varphi, \psi, \pi)\), choosing \( f_1 = \gamma_1(u(T - \cdot) - u_d(T - \cdot)), f_2 = \gamma_2(u(T - \cdot) - v_d(T - \cdot)) \), \( g_1 = \gamma_2(w(T - \cdot) - w_d(T - \cdot)) \) and \( g_2 = \gamma_3(w(T - \cdot) - w_d(T - \cdot)) \) and then reversing time via \( t \mapsto T - t \) we obtain the adjoint equations (3.16) and (3.17). With regard to the general system (4.1) we suppose at the very least the following regularity conditions on the source terms.

(B) It holds that \( f_1 \in L^2(L^2(\Omega_f)), f_2 \in L^2(H_s), g_1 \in L^2(H_s) \) and \( g_2 \in L^2(H^2(\Omega_s)) \).

Observe that under the above choices of the source terms for the state and adjoint equations, hypothesis (B) is satisfied according to Theorem 2.15, Corollary 3.8 and assumption (A).

The variational form of (4.1) is given by

\[
\begin{align*}
    (u_t, \varphi)_{\Omega_f} + (\nabla u, \nabla \varphi)_{\Omega_f} - (p, \text{div } \varphi)_{\Omega_f} \\
    + (w_{tt}, \varphi)_{\Omega_s} + \varepsilon (\nabla w_t, \nabla \varphi)_{\Omega_s} + (w, \varphi)_{1, \Omega_s} + (\mu w_t (\cdot - r), \varphi)_{\Omega_s} \\
    = (f_1, \varphi)_{\Omega_f} + (f_2, \varphi)_{\Omega_s} + \int_0^t (g_1, \varphi)_{\Omega_s} + (\nabla g_2, \nabla \varphi)_{\Omega_s} \, dt \quad \forall \varphi \in W, \\
    (\rho, \text{div } u) = 0 & \quad \forall \rho \in M, \\
    u(0) = u_0, \quad w(0) = w_0, \quad w_t(0) = v_0, \quad w_t = z_0 & \quad \text{in } Q_r.
\end{align*}
\]

Define \( v \) to be the weak solution of the following elliptic boundary value problem with parameter \( t \in I \):

\[-\Delta v(t) + v(t) = g_2(t) - g_1(t) \quad \text{in } Q_s, \quad \partial_t v = 0 \quad \text{on } \Sigma_s.\]

By elliptic regularity theory we have \( v \in L^2(H^2(\Omega_s)) \). It can be easily verified that the triple \((u, w, p)\) satisfies (4.2) if and only if the quadruple \((u, w, \zeta, p)\), where

\[
\zeta(t) = w(t) + \int_0^t (v(s) - g_2(s)) \, ds,
\]

(4.3)
satisfies the following system:

\[

d \begin{cases}
    (u_t, \psi)_{\Omega_t} + (\nabla u, \nabla \psi)_{\Omega_t} - (p, \text{div } \psi)_{\Omega_t} + (w_\mu, \psi)_{\Omega_s} + \varepsilon(\nabla w_t, \nabla \psi)_{\Omega_s} \\
    + (\zeta, \phi)_{1,\Omega_s} + (\mu w_t(-r), \phi)_{\Omega_s} = (f_1, \phi)_{\Omega_t} + (f_2, \phi)_{\Omega_s} \quad \forall \phi \in W, \\
    (\xi_t, \psi)_{1,\Omega_s} - (w_t, \psi)_{1,\Omega_s} = -(g_1, \psi)_{\Omega_s} - (\nabla g_2, \nabla \psi)_{\Omega_s} \quad \forall \psi \in V_s, \\
    (\rho, \text{div } u) = 0 \quad \forall \rho \in M, \\
    u(0) = u_0, \ w(0) = w_0, \ w_t(0) = v_0, \ \zeta(0) = w_0, \ \xi = z_0 \text{ in } Q_r.
\end{cases}
\]

(4.4)

We would like to point out that in relation to the state equation (1.2) we have \( \xi = w \).

Introducing the global velocity vector field \( \xi := u \chi_{\Omega_t} + w_t \chi_{\Omega_s} \) and its corresponding initial data \( \xi_0 := u_0 \chi_{\Omega_t} + v_0 \chi_{\Omega_s} \in W \), (4.4) can be rewritten as

\[
\begin{align*}
    (\xi_t, \psi)_{\Omega_t} + a_{\varepsilon}(\xi, \psi) + b(\psi, p) + a_s(\xi, \psi) + (\mu \xi(-r), \psi)_{\Omega_s} \\
    = (f_1, \psi)_{\Omega_t} + (f_2, \psi)_{\Omega_s} \quad \forall \psi \in W, \\
    a_s(\xi_t, \psi) - a_s(\xi, \psi) = -(g_1, \psi)_{\Omega_s} - (\nabla g_2, \nabla \psi)_{\Omega_s} \quad \forall \psi \in V_s, \\
    b(\xi, \rho) = 0 \quad \forall \rho \in M, \\
    \xi(0) = \xi_0, \ \zeta(0) = w_0, \ \xi = z_0 \text{ in } Q_r,
\end{align*}
\]

(4.5)

where \( a_{\varepsilon} : W \times W \to \mathbb{R} \) and \( a_s : V_s \times V_s \to \mathbb{R} \) are the bilinear forms

\[
    a_s(\xi, \psi) = (\nabla \xi, \nabla \psi)_{\Omega_s} + \varepsilon(\nabla \xi, \nabla \psi)_{\Omega_s}, \quad a_s(w, \psi) = (w, \psi)_{1,\Omega_s},
\]

and \( b \) is the bilinear form defined in (2.33). Equation (4.5) is in fact analogous to the symmetric form given in Failer et al. (2016) with \( \mu = 0, g_2 = 0 \) and under certain particular choices of control space and control operators. From (4.3), \( \xi \) and \( \zeta \) are related according to

\[
    \zeta = \xi + v - g_2.
\]

(4.6)

In particular, if \( g_1 = g_2 \) then \( \zeta = \xi - g_2 \) since \( v = 0 \).

It follows from assumptions (A) and (B), Theorem 2.15 and (4.3) that \( \xi \in H^1(L^2(\Omega)) \cap L^2(H^1(\Omega)) \cap H^2(\Omega_s) \) and \( \zeta \in H^1(H^2(\Omega_s)) \). Furthermore, we have the \textit{a priori} estimate

\[
    \| \xi \|_{H^1(L^2(\Omega)) \cap L^2(H^1(\Omega)) \cap H^2(\Omega_s)) + \| \xi \|_{H^1(H^2(\Omega_s))} \\
    \leq C(\| f_1 \|_{L^2(\Omega_s)} + \| f_2 \|_{L^2(H^1)} + \| g_1 \|_{L^2(H^1)}) \\
    + C(\| g_2 \|_{L^2(H^2(\Omega_s))} + \| \xi_0 \|_{2,\Omega_s} + \| \xi_0 \|_{2,\Omega_s} + \| w_0 \|_{2,\Omega_s} + \| z_0 \|_{H^1(\Omega_s)}).
\]

(4.7)

In the succeeding sections we provide a space-time discretization of (4.5). In order to have an error estimate independent of the mesh for space and time we apply the method of lines, that is, we first discretize the problem in space in Section 5 and then further discretize in time in Section 6.
5. Semidiscretization for the symmetric formulation

In this section we discuss a semidiscretization of equation (4.5) for a fixed $\varepsilon > 0$.

5.1 Finite element spaces

Let $\{K_h\}_{h>0}$ be a family of triangulations of $\Omega$. We suppose that $\Omega$ is a convex polygonal domain and $\Omega_s$ is sufficiently smooth and convex. For each triangle $K \in K_h$, let $\varrho_K$ denote the diameter of $K$ and $\vartheta_K$ be the diameter of the largest ball contained in $K$. The mesh size of the triangulation is given by the parameter $h = \max_{K \in K_h} \varrho_K$. Assume that the family of triangulations is quasi-uniform, that is, there exist $C_\varrho, C_\vartheta > 0$ such that

$$\frac{h}{\varrho_K} \leq C_\varrho, \quad \frac{\vartheta_K}{\varrho_K} \leq C_\vartheta \quad \forall K \in K_h, \forall h > 0.$$ 

Since $\Omega$ is polygonal we have $\Omega = \bigcup_{K \in K_h} K$. Let $\Omega_{sh}$ denote the union of all triangles in $K_h$ that lie entirely in $\Omega_s$, and let $\Omega_{fh}$ be its interior and boundary, respectively. Set $\Omega_{fh} = \Omega \setminus \Omega_{sh}$. We assume that the barycentric coordinates of each triangle in $\Omega_{fh}$ with an edge in $\Gamma_{sh}$ lie in $\Omega_f$. This is satisfied if $h$ is small enough. Suppose that there exists $h_0 > 0$ such that for each $0 < h < h_0$ the nodes in $\Omega_{fh}$ are also nodes in $\Omega_{fh_0}$.

By convexity of $\Omega_s$ we have $\Omega_{sh} \subset \Omega_s$, $\Omega_f \subset \Omega_{fh}$ and $\omega_h := \Omega_s \setminus \Omega_{sh} = \Omega_{fh} \setminus \Omega_f$. Furthermore, we suppose that the vertices of $K_h$ on $\Gamma_{fh} = \Gamma_f$ and $\Gamma_{sh}$ are also points on the fluid boundaries $\Gamma_f$ and $\Gamma_s$, respectively, and there exists $C_s > 0$ such that

$$\sup_{x \in \Gamma_{sh}} \inf_{y \in \Gamma_s} |x - y| \leq C_s h^2, \quad (5.1)$$

which is satisfied as soon as $\Omega_s$ is a $C^2$-domain. From (5.1), it follows that there exists $C > 0$ such that

$$\|\xi\|_{k,\omega_h} \leq Ch \|\xi\|_{k+1,\Omega_s}, \quad \forall \xi \in H^{k+1}(\Omega_s), \quad k = 0, 1. \quad (5.2)$$

We refer to Raviart & Thomas (1983, pp. 118–119) for the proofs of (5.1) and (5.2). Inequality (5.2) plays a crucial role in the proof of the stability and error estimates, specifically the errors due to the variational crimes induced by the discretization of the curved interface $\Gamma_f$.

In the following, we shall discretize in space using a cG(1) approximation scheme, that is, a continuous Galerkin method using piecewise linear functions with bubble functions for the fluid velocity. More precisely, we consider a P1-bubble/P1 (mini element) approximation for the fluid velocity and P1 approximation for the states corresponding to the solid.

Let $P_1$ denote the set of polynomials in two variables of degree at most 1. For each triangle $K \in T_h$, let $\lambda_{K,1}, \lambda_{K,2}$ and $\lambda_{K,3}$ be the corresponding shape functions and $B_K$ the linear span of the bubble function $27\lambda_{K,1}\lambda_{K,2}\lambda_{K,3}$. Define

$$W_h = \left\{ \xi_h \in C(\partial) : \xi_h|_K \in (P_1 \oplus B_K)^2 \text{ if } K \subset \Omega_{fh}, \xi_h|_K \in P_1^2 \text{ if } K \subset \Omega_{sh}, \xi_h|_{\Gamma_f} = 0 \right\}.$$
Observe that we have a conforming scheme with respect to the global velocity field $\xi$ since $W_h \subset W$. We also define the restrictions of the functions in the discretized fluid and structure domains as

$$W_{fh} = \{ \xi|_{\Omega_{fh}} : \xi \in W_h \}, \quad W_{sh} = \{ \xi|_{\Omega_{sh}} : \xi \in W_h \}.$$  

For the discretization of the pressure we consider the space

$$M_h = \{ p_h \in C(\Omega_{fh}) : p_h|_K \in P_1 \}.$$  

With this the restrictions of functions in $M_h$ to $\Omega_f$ lies in $M$.

We also define the space of discretely divergence-free elements of $W_h$ in $\Omega_{fh}$,

$$X_h = \{ \xi_h \in W_h : b_h(\xi_h,p_h) = 0 \quad \forall p_h \in M_h \},$$  

where $b_h : W_h \times M_h \rightarrow \mathbb{R}$ is the discrete version of the bilinear form (2.33) given by

$$b_h(\xi_h,p_h) = -\int_{\Omega_{fh}} p_h \text{div} \xi_h \, dx.$$  

Note that according to the above assumptions on the nodes in $\Omega_{fh}$ and $\Omega_{fh0}$, each piecewise linear function in $W_{fh0}$ is also an element of $W_{fh}$. By using the arguments of the proof in Du et al. (2003, Theorem 3.5), it follows that the pair $(W_h,M_h)$ satisfies the inf–sup condition

$$\inf_{q_h \in M_h \setminus \{0\}} \sup_{\varphi_h \in W_h \setminus \{0\}} \frac{b_h(\varphi_h,q_h)}{\|\varphi_h\|_{1,\Omega} \|q_h\|_{\Omega_{fh}}} \geq c > 0. \quad (5.3)$$  

It is well known that the finite element spaces $W_{fh}$ and $W_{sh}$ and $M_h$ satisfy the following approximation properties:

$$\inf_{\varphi_h \in W_{sh}} \|\varphi - \varphi_h\|_{j,\Omega_{sh}} \leq Ch^{2-j}\|\varphi\|_{2,\Omega_{sh}} \quad \forall \varphi \in H^2(\Omega_{sh}), \quad j \in [0,2],$$  

$$\inf_{\varphi_h \in W_{fh}} \|\varphi - \varphi_h\|_{j,\Omega_{fh}} \leq Ch^{2-j}\|\varphi\|_{2,\Omega_{fh}} \quad \forall \varphi \in H^2(\Omega_{fh}), \quad j \in [0,2],$$  

$$\inf_{p_h \in M_h} \|p - p_h\|_{j,\Omega_{fh}} \leq Ch^{1-j}\|p\|_{1,\Omega_{fh}} \quad \forall p \in H^1(\Omega_{fh}), \quad j \in [0,1].$$  

From the quasi-uniformity of the triangulations the following inverse estimate holds:

$$\|\varphi_h\|_{1,\Omega} \leq Ch^{-1}\|\varphi_h\|_{\Omega} \quad \forall \varphi_h \in W_h. \quad (5.4)$$  

Observe that the discretized fluid domain is slightly larger than the continuous one, that is, $\Omega_f \subset \Omega_{fh}$. In our analysis we artificially extend functions in $\Omega_f$ to all of $\Omega$. Define an extension operator $E : H^k(\Omega_f) \rightarrow H^k(\Omega)$ for $k = 1, 2$ such that for some constant $C > 0$,

$$\|E\varphi\|_{k,\Omega} \leq C\|\varphi\|_{k,\Omega_f} \quad \forall \varphi \in H^k(\Omega_f). \quad (5.5)$$
For example, given $\varphi \in H^k(\Omega_f)$ with $k = 1$ or $k = 2$, let $\tilde{\varphi}$ be the harmonic extension of $\varphi|_{\Gamma_s}$ to $\Omega_s$ and set $E\varphi = \varphi \chi_{\Omega_f} + \tilde{\varphi} \chi_{\Omega_s}$. Estimate (5.5) then follows from elliptic theory.

The approximation properties of $W_{f_h}$ and $W_{s_h}$ carry over to $W_h$, which we prove below (see also Du et al., 2004).

**Theorem 5.1** There exists a constant $C > 0$ such that for every $\xi \in W$ satisfying $\xi|_{\Omega_f} \in H^2(\Omega_f)$ and $\xi|_{\Omega_s} \in H^2(\Omega_s)$ we have

$$\inf_{\xi_h \in W_h} \|\xi - \xi_h\|_{j,\Omega} \leq C h^{2-j} (\|\xi\|_{2,\Omega_f} + \|\xi\|_{2,\Omega_s}), \quad j = 0, 1.$$ 

Similarly, there exists $C > 0$ such that for every $p \in H^1(\Omega_f)$ it holds that

$$\inf_{p_h \in M_h} \|E p - p_h\|_{\Omega_{f_h}} \leq C h \|p\|_{1,\Omega_f}.$$ 

**Proof.** Let $i_h : C(\overline{\Omega}) \to W_h$, $i_{f_h} : C(\overline{\Omega}_{f_h}) \to W_{f_h}$ and $i_{s_h} : C(\overline{\Omega}_{s_h}) \to W_{s_h}$ be the nodal Lagrange interpolation operators. According to the assumption on the nodes of the barycenters of the triangles in $\Omega_{f_h}$ that intersect the discretized interface $\Gamma_{s_h}$ we have $(i_{f_h} \xi)|_{\Omega_{f_h}} = i_{f_h} \xi$ and $(i_{s_h} \xi)|_{\Omega_{s_h}} = i_{s_h} \xi$. From the triangle inequality,

$$\|\xi - i_h \xi\|_{j,\Omega} \leq \|E \xi - i_{f_h} \xi\|_{j,\Omega_{f_h}} + \|\xi - i_{s_h} \xi\|_{j,\Omega_{s_h}} + \|\xi - E \xi\|_{j,\omega_h}, \quad j = 0, 1.$$ 

Using interpolation error estimates, (5.2) and (5.5) we see that $\|\xi - i_h \xi\|_{j,\Omega} \leq C h^{2-j} (\|\xi\|_{2,\Omega_f} + \|\xi\|_{2,\Omega_s})$. Since $i_h \xi \in W_h$ the first part of the theorem follows. The second statement can be shown in an analogous way. \hfill \Box

In the following we construct an extension of functions defined in $\Omega_{s_h}$ to $\Omega_s$.

**Lemma 5.2** There exists an extension operator $E_h : W_{s_h} \to V_s$ such that $E_h \psi_h = \psi_h$ on $\Omega_{s_h}$, and for some $C > 0$ there holds $\|E_h \psi_h\|_{1,\Omega_s} \leq C \|\psi_h\|_{1,\Omega_{s_h}}$ for every $h > 0$ and $\psi_h \in W_{s_h}$.

**Proof.** Let $\hat{K}$ denote a reference element of each triangulation $K_h$. Given an element $K \in K_h$ such that $K \subset \overline{\Omega}_{s_h}$ and with an edge $e_K$ on the discretized interface $\Gamma_{s_h}$, denote by $\omega_K$ the region bounded by this edge and $\Gamma_s$. Then there exists a unique $\hat{K} \in K_h$ such that $\omega_K \subset \hat{K}$. Let $F_K$ and $F_{\hat{K}}$ be affine transformations mapping $\hat{K}$ onto $K$ and $\hat{K}$, respectively. Choose $F_K$ and $F_{\hat{K}}$ in such a way that exactly one edge of $\hat{K}$ is mapped onto the common edge $e_K$ of $K$ and $\hat{K}$, and moreover preserves orientation. On $\Omega_{s_h}$ we define $E_h \psi_h = \psi_h$, and on each $\omega_K$ we define

$$E_h \psi_h = \psi_h \circ F_K \circ F_{\hat{K}}^{-1}.$$
By construction, $E_h \psi_h \in V_s$. Moreover, if $A_K$ and $\tilde{A}_K$ are the matrices of transformations corresponding to $F_K$ and $\tilde{F}_K$, respectively, then standard results in finite element analysis imply the existence of a constant $C > 0$ independent of $h$ such that

$$
\|E_h \psi_h\|_{1,0_K} \leq \|E_h \psi_h\|_{1,\tilde{K}} \leq C \|A_K^{-1}\| \|A_K\| \left|\det A_K\right|^{\frac{1}{2}} \left|\det \tilde{A}_K\right|^{-\frac{1}{2}} \|\psi_h\|_{1,K}
$$

where the last inequality is due to the quasi-uniformity of $\mathcal{K}_h$. Doing this on each of the residual regions in $\omega_h$ we obtain $\|E_h \psi_h\|_{1,\Omega_s} \leq (C + 1)\|\psi_h\|_{1,\Omega_{sh}}$ for every $h > 0$ and $\psi_h \in W_{sh}$.

The above lemma implies that the $H^1$-norms of $E_h \psi_h$ and $\psi_h$ are equivalent independent of $h$. If a function defined on $\Omega_{sh}$ is integrated over the slightly larger domain $\Omega_s$ or a subset of it, we mean precisely its extension through $E_h$.

### 5.2 Semidiscretization in space for the symmetric formulation

Let $Q_{rh} = I_r \times \Omega_{sh}$. For the semidiscretization of (4.5) we consider the following: given $\xi_{0h} \in W_h$, $w_{0h} \in W_{sh}$, $z_{0h} \in L^2(W_{sh})$, $f_1 \in L^2(L^2(\Omega_f))$, $f_2, g_1 \in L^2(H_s)$ and $g_2 \in L^2(V_s)$ find a triplet $(\xi_h, \zeta_h, p_h) \in H^1(W_h) \times H^1(W_{sh}) \times L^2(M_h)$ such that

$$
\begin{align*}
(\xi_h, \psi_h) + a_{\varepsilon h}(\xi_h, \varphi_h) + b_h(\varphi_h, p_h) + a_{sh}(\zeta_h, \varphi_h) + (\mu \varepsilon_h (- r), \varphi_h)_{\Omega_{sh}} &= (f_1, \varphi_h)_{\Omega_h} + (f_2, \psi_h)_{\Omega_{sh}} \quad \forall \varphi_h \in W_h, \\
(a_{sh}(\zeta_h, \psi_h), \psi_h) - a_{sh}(\xi_h, \psi_h) &= -(g_1, \varphi_h)_{\Omega_{sh}} - (\nabla g_2, \nabla \psi_h)_{\Omega_{sh}} \quad \forall \psi_h \in W_{sh}, \\
b(\xi_h, \rho_h) &= 0 \quad \forall \rho_h \in M_h, \\
(\xi_{0h}, \psi_{0h}) &= (w_{0h}, \xi_{0h})_{\Omega_{sh}} = z_{0h} \quad \text{in } Q_{rh},
\end{align*}
$$

for a.e. $t \in I$. Here $a_{\varepsilon h} : W_h \times W_h \to \mathbb{R}$ and $a_{sh} : W_{sh} \times W_{sh} \to \mathbb{R}$ are the discrete versions of $a_\varepsilon$ and $a_s$ given by

$$
a_{\varepsilon h}(\xi_h, \varphi_h) = (\nabla \xi_h, \nabla \varphi_h)_{\Omega_h} + \varepsilon(\nabla \xi_h, \nabla \varphi_h)_{\Omega_{sh}}, \quad a_{sh}(\zeta_h, \psi_h) = (\zeta_h, \psi_h)_{1,\Omega_{sh}}.
$$

Let $P_h : L^2(\Omega) \to X_h$, $P_{fh} : L^2(\Omega_f) \to L^2(\Omega_{fh})$ and $P_{sh} : L^2(\Omega_s) \to L^2(\Omega_{sh})$ be the projection operators defined as follows: given $\xi \in L^2(\Omega)$, $u \in L^2(\Omega_f)$ and $w \in L^2(\Omega_s)$ let $P_h \xi$, $P_{fh} u$ and $P_{sh} \xi$ be the solutions of

$$
\begin{align*}
(P_h \xi, \varphi_h)_{\Omega} &= (\xi, \varphi_h)_{\Omega} \quad \forall \varphi_h \in X_h, \\
(P_{fh} u, u_h)_{\Omega_{fh}} &= (u, u_h)_{\Omega_{fh}} \quad \forall u_h \in L^2(\Omega_{fh}), \\
(P_{sh} \xi, \zeta_h)_{\Omega_{sh}} &= (\xi, \zeta_h)_{\Omega_{sh}} \quad \forall \zeta_h \in L^2(\Omega_{sh}).
\end{align*}
$$
Also, we define the Ritz–Galerkin projection operators $R_{sh} : V_s \rightarrow W_{sh}$ and $\tilde{R}_{sh} : V_s \rightarrow W_{sh}$ as follows: given $\xi \in V_s$, let $R_{sh}\xi$ and $\tilde{R}_{sh}\xi$ be the finite element solutions of

$$a_{sh}(R_{sh}\xi, \zeta_h) = a_{sh}(\xi, \zeta_h) \quad \forall \zeta_h \in W_{sh},$$

$$\gamma_2(\tilde{R}_{sh}\xi, \zeta_h)_{\Omega_{sh}} + \gamma_3(\nabla \tilde{R}_{sh}\xi, \nabla \zeta_h)_{\Omega_{sh}} = \gamma_2(\xi, \zeta_h)_{\Omega_{sh}} + \gamma_3(\nabla \xi, \nabla \zeta_h)_{\Omega_{sh}} \quad \forall \zeta_h \in W_{sh}.$$

From the theory of finite elements we have the error estimate

$$\| R_{sh}\xi - \xi \|_{1,\Omega_{sh}} \leq Ch \| \xi \|_{2,\Omega_s} \quad \forall \xi \in H^2(\Omega_s). \quad (5.7)$$

**Remark 5.3** Let $v_h$ be the finite element solution of the elliptic problem with parameter $t \in I$,

$$a_{sh}(v_h(t), \psi_h(t)) = (g_2(t) - g_1(t), \psi_h(t))_{\Omega_{sh}} \quad \forall \psi_h \in W_{sh}. \quad (5.8)$$

From the second equation in (5.6) we infer that

$$\zeta_{ht} = \xi_h + v_h - R_{sh}g_2, \quad \text{in } L^2(W_{sh}), \quad (5.9)$$

and this can be realized as a semidiscretization of (4.6). If $g_1 = g_2$ then $v_h = 0$ and thus

$$\zeta_{ht} = \xi_h - R_{sh}g_2. \quad (5.10)$$

In order to prove an approximation error induced by the projection $P_h$ we introduce the following Ritz–Galerkin-type approximation associated with the Stokes–Neumann operator.

**Lemma 5.4** Let $p \in H^1(\Omega_f)$ and $\xi \in W$ with $\xi|_{\Omega_f} \in H^2(\Omega_f)$ and $\xi|_{\Omega_s} \in H^2(\Omega_s)$. Then there exists a unique pair $(\xi_h, P_h) \in W_h \times M_h$ such that

$$\begin{cases}
a_{eh}(\xi_h, \varphi_h) + b_h(\varphi_h, p_h) = a_{eh}(\xi, \varphi_h) + b_h(\varphi_h, Ep) \quad \forall \varphi_h \in W_h, \\
b_h(\xi_h, \rho_h) = 0 \quad \forall \rho_h \in M_h,
\end{cases} \quad (5.11)$$

and there is a constant $C > 0$ such that

$$\| \xi - \xi_h \|_{\Omega_f} + h \| \xi - \xi_h \|_{1,\Omega_f} + h \| Ep - p_h \|_{\Omega_{sh}} \leq Ch^2 (\| \xi \|_{2,\Omega_f} + \| \xi \|_{2,\Omega_s} + \| p \|_{1,\Omega_f}). \quad (5.12)$$

**Proof.** Consider the auxiliary mixed variational problem

$$\begin{cases}
a_{eh}(\tilde{\xi}_h, \varphi_h) + b_h(\varphi_h, p_h) = a_{eh}(\xi, \varphi_h) + b_h(\varphi_h, Ep) \quad \forall \varphi_h \in W_h, \\
b_h(\tilde{\xi}_h, \rho_h) = b_h(\xi, \rho_h) \quad \forall \rho_h \in M_h,
\end{cases} \quad (5.13)$$

Thanks to the discrete inf–sup condition (5.3) the finite-dimensional mixed problems (5.11) and (5.13) possess unique solutions; see Girault & Raviart (1986, Theorem II.1.1). Define the error terms $e_h =
\( \xi - \xi_h \) and \( \pi_h = \mathbf{E} - p_h \). We split these according to \( e_h = \xi - \xi_h \) and \( \pi_h = \mathbf{E} - \pi_h \), where \( \mathbf{E} = \xi - \xi_h \), \( \mathbf{E} - \pi_h = \mathbf{E} - \pi_h \), and \( \pi_h = \pi_h - p_h \). We have the \textit{a priori} estimate

\[
\| \xi - \xi_h\|_{1,\Omega} + \| \pi_h\|_{\Omega_h} \leq C \inf_{\varphi_h \in W_h} \| \xi - \varphi_h\|_{1,\Omega} + C \inf_{\rho_h \in M_h} \| \mathbf{E} - \rho_h\|_{\Omega_h}.
\]

On the other hand, by stability of solutions, (5.2) and \( \text{div} \xi = 0 \) on \( \Omega_f \) we have

\[
\| \xi - \xi_h\|_{1,\Omega} + \| \pi_h\|_{\Omega_h} \leq C \| \text{div} \xi\|_{\omega_h} \leq Ch \| \xi\|_{2,\Omega_f}.
\]

Consequently, by the triangle inequality, we obtain

\[
\| e_h\|_{1,\Omega} + \| \pi_h\|_{\Omega_h} \leq C \left\{ h\| \xi\|_{2,\Omega_f} + \inf_{\varphi_h \in W_h} \| \xi - \varphi_h\|_{1,\Omega} + \inf_{\rho_h \in M_h} \| \mathbf{E} - \rho_h\|_{\Omega_h} \right\}.
\]

Therefore, from the approximation properties in Theorem 5.1, we infer that

\[
\| e_h\|_{1,\Omega} + \| \pi_h\|_{\Omega_h} \leq Ch\left( \| \xi\|_{2,\Omega_f} + \| \xi\|_{2,\Omega_f} + \| p\|_{1,\Omega_f} \right).
\]

To prove estimate (5.12) for the \( L^2 \)-norm we use a standard Aubin–Nitsche trick. For each \( g \in L^2(\Omega) \) let \( (z_g, \pi_g) \in W \times M \) be the solution of the dual mixed problem

\[
\begin{aligned}
a_e(\varphi, z_g) + b(\varphi, \pi_g) &= (g, \varphi)_{\Omega} \quad \forall \varphi \in W, \\
b(z_g, \rho) &= 0 \quad \forall \rho \in M.
\end{aligned}
\]

We have a unique solution for this problem thanks to (2.32). Using the same argument as in the unsteady case it can be shown that \( z_g|_{\Omega_f} \in H^2(\Omega_f) \), \( z_g|_{\Omega_s} \in H^2(\Omega_s) \) and \( \pi_g \in H^1(\Omega_f) \). Moreover, we have the \textit{a priori} estimate

\[
\| z_g\|_{2,\Omega_f} + \| z_g\|_{2,\Omega_s} + \| \pi_g\|_{1,\Omega_f} \leq C\| g\|_{\Omega}.
\]

We rewrite the first equation in (5.16) as

\[
a_e(\varphi, z_g) + b(\varphi, E\pi_g) = (g, \varphi)_{\Omega} - (E\pi_g, \text{div} \varphi)_{\omega_h} + (1 - \varepsilon)(\nabla \varphi, \nabla z_g)_{\omega_h}.
\]

Observe that \( b(\varphi, \pi_g) = -\pi_h \text{div} z_g_{\omega_h} \) and \( b(\varphi, \rho_h) = -\rho_h \text{div} \xi_{\omega_h} \) for every \( \rho_h \in M_h \) and \( a_e(\varphi, \rho_h) = -\pi_h \text{div} (\varphi, \pi_h) \) for all \( \varphi, \rho_h \in W_h \). Taking \( \varphi = \xi_h \) in (5.18) we obtain

\[
(g, e_h)_{\Omega} = a_e(e_h, z_g)_{\Omega} + b(e_h, E\pi_g) + (E\pi_g, \text{div} e_h)_{\omega_h} - (1 - \varepsilon)(\nabla e_h, \nabla z_g)_{\omega_h}
\]

\[
= a_e(e_h, z_g - \varphi_h - \pi_h)_{\Omega} + b(z_g - \varphi_h, \pi_h) + b(e_h, E\pi_g - \rho_h) + \ell(\rho_h),
\]

where \( \ell(\rho_h) = (\pi_h, \text{div} z_g)_{\omega_h} + (E\pi_g, \text{div} e_h)_{\omega_h} - (\rho_h, \text{div} \xi)_{\omega_h} - (1 - \varepsilon)(\nabla e_h, \nabla z_g)_{\omega_h} \). Choosing \( \rho_h \in M_h \) to be the linear interpolant of \( E\pi_g \) we have, using (5.17),

\[
\| E\pi_g - \rho_h\|_{\Omega_h} \leq Ch \| \pi_g\|_{1,\Omega_f} \leq Ch \| g\|_{\Omega}.
\]
On the other hand, invoking (5.2), (5.15) and (5.17), we have

$$|\ell(\rho_h)| \leq Ch^2(\|\xi\|_{2,\Omega_f} + \|\xi\|_{2,\Omega_s} + \|p\|_{1,\Omega_f})\|g\|_{\Omega_f}.$$ (5.21)

Taking the infimum over all \(\varphi_h \in W_h\) and then the supremum over all \(g \in L^2(\Omega) \setminus \{0\}\) in (5.19),

$$\|e_h\|_{\Omega_f} \leq C \sup_{g \in L^2(\Omega) \setminus \{0\}} \frac{|\ell(\rho_h)|}{\|g\|_{\Omega_f}} + C(\|e_h\|_{1,\Omega_f} + \|\pi_h\|_{\Omega_f})$$

$$\times \sup_{g \in L^2(\Omega) \setminus \{0\}} \frac{1}{\|g\|_{\Omega_f}} \left[ \inf_{\varphi_h \in W_h} \left( \|z_g - \varphi_h\|_{1,\Omega_f} + \|\pi_g - \rho_h\|_{\Omega_f} \right) \right].$$ (5.22)

Using (5.20), (5.21) and Theorem 5.1 in (5.22), we deduce (5.12) involving the \(L^2\)-norm. \(\square\)

**Theorem 5.5** Suppose that \(\xi \in W\) satisfies \(\xi|_{\Omega_f} \in H^2(\Omega_f)\) and \(\xi|_{\Omega_s} \in H^2(\Omega_s)\). Then there exists a constant \(C > 0\) such that

$$\|P_h\xi - \xi\|_{j,\Omega} \leq C h^{2-j} \left( \|\xi\|_{2,\Omega_f} + \|\xi\|_{2,\Omega_s} \right), \quad j = 0, 1.$$ (5.23)

**Proof.** Take \(p \in H^1(\Omega_f)\) with \(\|p\|_{1,\Omega_f} \leq \|\xi\|_{2,\Omega_f}\) and let \((\xi_h, p_h)\) be the solution of (5.11) corresponding to the pair \((\xi, p)\). Notice that \(\xi_h \in X_h\). The approximation property of a projection operator gives us \(\|P_h\xi - \xi\|_{\Omega_f} \leq \|\xi_h - \xi\|_{\Omega_f}\), and hence from Lemma 5.4 we have (5.23) for \(j = 0\). According to inverse estimate (5.4) we obtain

$$\|P_h\xi - \xi\|_{1,\Omega} \leq C h^{-1} \|P_h\xi - \xi_h\|_{\Omega} + \|\xi_h - \xi\|_{1,\Omega}$$

$$\leq C h^{-1} (\|P_h\xi - \xi\|_{\Omega} + \|\xi - \xi_h\|_{\Omega}) + \|\xi_h - \xi\|_{1,\Omega}.$$ (5.23)

Thus, from Lemma 5.4 and (5.23) with \(j = 0\) we have (5.23) with \(j = 1\). \(\square\)

With regard to the approximation of the initial data we consider

$$\xi_{0h} = P_h\xi_0, \quad w_{0h} = R_{sh}w_0, \quad z_{0h}(\theta) = P_h(\bar{u}_0(\theta)\chi_{\Omega_f} + z_0(\theta)\chi_{\Omega_s})|_{\Omega_{sh}},$$ (5.24)

where, for each \(\theta \in I_f\), \((\bar{u}_0(\theta), \bar{p}_0(\theta)) \in V_f \times (M/\mathbb{R})\) is the solution of the Stokes equation

\[
\begin{cases}
-\Delta \bar{u}_0(\theta) + \nabla \bar{p}_0(\theta) = -\Delta u_0 + \nabla p_0 & \text{in } \Omega_f, \\
\text{div } \bar{u}_0(\theta) = 0 & \text{in } \Omega_f, \\
\bar{u}_0(\theta) = 0 & \text{in } \Gamma_f, \\
\bar{u}_0(\theta) = z_0(\theta) & \text{in } \Gamma_s.
\end{cases}
\]

From hypothesis (A) we have \(\bar{u}_0 \in H^1_0(V_f) \cap L^2(\Omega_f)\) and \(\bar{u}_0(0) = u_0\). The choice of the approximation for the initial history implies that the compatibility condition \(z_0(0) = v_0 = \xi_0|_{\Omega_s}\) on
the continuous level is carried out to the discrete level, that is, \( z_{0h}(0) = P_h \xi_0 |_{\Omega_h} = \xi_{0h} |_{\Omega_h} \). From the theory of (A), (5.7) and the approximation properties for \( W_h \) and \( W_{sh} \) we have the stability estimate

\[
\| \xi_{0h} \|_{1,\Omega} + \| w_{0h} \|_{1,\Omega} + \| z_{0h} \|_{1,\Omega} + \| z_{0h} \|_{H^1(\Omega)} \leq C(\| \xi_0 \|_{1,\Omega} + \| w_0 \|_{1,\Omega} + \| z_0 \|_{H^1(V)})
\]

(5.25)

and the error estimate

\[
\| \xi_0 - \xi_{0h} \|_{\Omega} + h\| \xi_0 - \xi_{0h} \|_{1,\Omega} + h\| w_0 - w_{0h} \|_{1,\Omega} + \| z_0 - z_{0h} \|_{L^2(\Omega)} + \| z_0 - z_{0h} \|_{L^2(\Omega)})
\leq C h^2 (\| \xi_0 \|_{2,\Omega} + \| \xi_0 \|_{2,\Omega} + \| w_0 \|_{2,\Omega} + \| z_0 \|_{L^2(\Omega)}).
\]

(5.26)

An alternative and more practical choice of approximations for the initial data and history based on interpolation will be provided in the next section.

We now prove the existence and stability of the semidiscrete problem (5.6) in the following theorem.

**Theorem 5.6** Let \( f_1 \in L^2(L^2(\Omega_{sh})), f_2 \in L^2(L^2(\Omega_{sh})), g_1 \in L^2(L^2(\Omega_{sh})), g_2 \in L^2(H^1(\Omega_{sh})), \) \( \xi_{0h} \in X_h, w_{0h} \in W_{sh} \) and \( z_{0h} \in L^2(W_{sh}). \) Then there exists a triplet \((\xi_h, \psi_h, p_h) \in H^1(X_h) \times H^1(W_{sh}) \times L^2(M_h)\) satisfying (5.6). Moreover, there exists a constant \( C > 0 \) independent of \( h \), the source terms and initial data such that

\[
\| \xi_h \|_{H^1(L^2(\Omega))} + \| \psi_h \|_{H^1(L^2(\Omega))} + \| p_h \|_{L^2(\Omega)} + \| \xi_{0h} \|_{H^1(\Omega_{sh})} + \| w_{0h} \|_{1,\Omega} + \| z_{0h} \|_{Q_{rh}}
\leq C(\| f_1 \|_{L^2(L^2(\Omega_{sh}))} + \| f_2 \|_{L^2(L^2(\Omega_{sh}))} + \| g_1 \|_{L^2(L^2(\Omega_{sh}))} + \| g_2 \|_{L^2(H^1(\Omega_{sh}))})
\]

\[
+ C(\| \xi_{0h} \|_{1,\Omega} + \| w_{0h} \|_{1,\Omega} + \| z_{0h} \|_{Q_{rh}}).
\]

(5.27)

**Proof.** First, we consider the auxiliary problem: find \((\xi_h, \psi_h) \in H^1(X_h) \times H^1(W_{sh})\) such that

\[
\begin{cases}
(\xi_h, \psi_h) = (f, h)_{\Omega} + a_{sh}(\xi_h, \psi_h) + a_{sh}(\xi_h, \psi_h) + (\mu \xi_h (-r), \psi_h)_{\Omega_{sh}} \\
= (f_1, h)_{\Omega_{sh}} + (f_2, h)_{\Omega_{sh}} \forall \psi_h \in X_h, \\
\xi_{0h}(0) = \xi_{0h}, \quad \psi_{0h}(0) = w_{0h}, \quad \xi_h = z_{0h} \text{ in } Q_{rh}.
\end{cases}
\]

(5.28)

Expanding \( \xi_h \) and \( \psi_h \) in terms of the finite element bases for \( X_h \) and \( W_{sh} \), respectively, (5.28) admits a unique solution \((\xi_h, \psi_h) \in H^1(X_h) \times H^1(W_{sh})\) according to the theory of delay differential equations. The existence of the semidiscrete pressure \( p_h \) satisfying the first equation in (5.6) now follows from the discrete inf–sup condition (5.3), along with the same argument as in the proof of Theorem 2.14. The \textit{a priori} estimate (5.27) can be proved as in the continuous case using \((\varphi_h, \psi_h) = (\xi_h, \psi_h)\) and \((\varphi_h, \psi_h) = (\xi_h, \psi_h)\) as test functions. \( \square \)

**Theorem 5.7** Let \( f_1 \in H^1(L^2(\Omega_{sh})), f_2 \in H^1(L^2(\Omega_{sh})), g_1 \in L^2(L^2(\Omega_{sh})), g_2 \in L^2(H^1(\Omega_{sh})), \) \( \xi_{0h}, w_{0h}, \) and \( z_{0h} \) be given as in (5.24). Then \( \xi_h \in H^2(X_h) \) and there exists a constant \( C > 0 \) independent of \( h \).
such that
\[
\|\xi_h\|_{H^1_0(\Omega)} \leq C(\|f_1\|_{L^2(\Omega)} + \|f_2\|_{L^2(\Omega)}) + \|g_1\|_{L^2(\Omega)} + \|g_2\|_{L^2(\Omega)}) + C(\|\xi_0\|_{2,\Omega} + \|\xi_0\|_{2,\Omega} + \|p_0\|_{1,\Omega} + \|w_0\|_{2,\Omega} + \|z_0\|_{H^1_0(\Omega)}).
\]

**Proof.** Let us denote by \(\tilde{C}\) the term on the right-hand side of (5.29). The fact that \(\xi_h \in H^2(X_h)\) is a consequence of the compatibility condition \(z_{0h}(0) = \xi_{0h}|_{\Omega_{sh}}\), since this implies that \(\xi_h(-r) \in H^1(W_{sh})\). Differentiating the first equation in (5.28) with respect to \(t\), taking the test functions \(\varphi_h = \partial_t \xi_h\) and \(\psi_h = \partial_r \xi_h\), integrating over \(I\) and using (5.25) and (5.27) we have
\[
\|\partial_t \xi_h\|_{L^2(\Omega)} \leq \tilde{C} + C\|\partial_t \xi_h(0)\|_{\Omega}.
\]

Hence, it remains to bound the term \(\|\partial_t \xi_h(0)\|_{\Omega}\).

Evaluating the first equation in (5.28) at \(t = 0\) and by taking \(\varphi_h = \partial_t \xi_h(0)\) we have
\[
\|\partial_t \xi_h(0)\|^2 + a_{sh}(\xi_{0h}, \partial_t \xi_h(0)) + a_{wh}(w_{0h}, \partial_t \xi_h(0)) + (\mu z_{0h}(-r), \partial_t \xi_h(0))_{\Omega_{sh}}
= (f_1(0), \partial_t \xi_h(0))_{\Omega_{sh}} + (f_2(0), \partial_r \xi_h(0))_{\Omega_{sh}}.
\]

Using Young’s inequality and the boundedness of the projection \(P_h\), we infer that
\[
|(f_1(0), \partial_t \xi_h(0))_{\Omega_{sh}}| + |(f_2(0), \partial_t \xi_h(0))_{\Omega_{sh}}| + |(\mu z_{0h}(-r), \partial_t \xi_h(0))_{\Omega_{sh}}| \leq \tilde{C}C_\rho + \rho \|\partial_t \xi_h(0)\|^2_{\Omega}
\]
for every \(\rho > 0\). To estimate the remaining terms in (5.31) we rewrite
\[
a_{sh}(\xi_{0h}, \partial_t \xi_h(0)) + a_{wh}(w_{0h}, \partial_t \xi_h(0))
= a_{sh}(\xi_0 - \xi_{0h}, \partial_t \xi_h(0)) + a_{sh}(w_{0h} - w_0, \partial_t \xi_h(0)) + a_{s}(\xi_0, \partial_t \xi_h(0))
+ a_{s}(\xi_0, \partial_t \xi_h(0)) - (1 - \epsilon)(\nabla \xi_0, \nabla \partial_t \xi_h(0))_{\omega_h} - (w_0, \partial_t \xi_h(0))_{\omega_h}
\]

Using Green’s identity and the condition \(\partial_t w_0 \geq \partial_t u_0 = \partial_t u_0 - p_0 v\) in hypothesis (A) we obtain
\[
a_{s}(\xi_0, \partial_t \xi_h(0)) + a_{s}(\xi_0, \partial_t \xi_h(0))
= -((\Delta \xi_0, \partial_t \xi_h(0))_{\Omega} - (\epsilon \Delta \xi_0 + \Delta w_0 - w_0, \partial_t \xi_h(0))_{\Omega} + b(\partial_t \xi_h(0), p_0).
\]

Now, according to (5.2), the inverse estimates (5.4), (5.26) and (5.33) it can be shown from (5.32) that for every \(\rho > 0\) there exists \(C_\rho > 0\) such that
\[
|a_{sh}(\xi_{0h}, \partial_t \xi_h(0)) + a_{wh}(w_{0h}, \partial_t \xi_h(0))|
\leq C_{\rho}(\|\xi_0\|_{2,\Omega} + \|\xi_0\|_{2,\Omega} + \|p_0\|_{1,\Omega} + \|w_0\|_{2,\Omega})^2 + \rho \|\partial_t \xi_h(0)\|^2_{\Omega}
\]
Combining the above estimates and taking \(\rho > 0\) small enough we see that \(\|\partial_t \xi_h(0)\|_{\Omega}\) is bounded by \(\tilde{C}\). Plugging this information into (5.30) we deduce (5.29).
Theorem 5.8 Suppose that \( f_1 \in H^1(L^2(\Omega_\varnothing)) \), \( f_2 \in H^1(L^2(\Omega_\varnothing)) \), \( g_1 \in L^2(V_\varnothing) \) and \( g_2 \in L^2(H^2(\Omega_\varnothing)) \). Let \((\xi, \zeta)\) and \((\xi_\varnothing, \zeta_\varnothing)\) be the solutions of (4.5) and (5.6), respectively. Then there exists a constant \( C > 0 \) independent of \( h \) such that

\[
\|\xi - \xi_\varnothing\|_{L^\infty(L^2(\Omega))} + \|\zeta - \zeta_\varnothing\|_{H^1(H^1(\Omega_\varnothing))} \leq Ch.
\] (5.34)

Proof. Let us introduce the discretization errors

\[
e_h = \xi - \xi_\varnothing, \quad \eta_h = \zeta - \zeta_\varnothing, \quad r_h = Ep - p_h.
\]

We split these according to \( e_h = \tilde{e}_h + \hat{e}_h \) and \( \eta_h = \tilde{\eta}_h + \hat{\eta}_h \) where \( \tilde{e}_h = \xi - P_h \xi, \hat{e}_h = P_h \xi - \xi_\varnothing, \tilde{\eta}_h = \zeta - R_{sh} \zeta \) and \( \hat{\eta}_h = R_{sh} \zeta - \zeta_\varnothing \). The approximation properties of \( P_h \) and \( R_{sh} \) along with the estimate for \( \xi \) and \( \zeta \) given in (4.7) imply that

\[
\|\tilde{e}_h\|_{L^2(L^2(\Omega))} + \|\tilde{\eta}_h\|_{H^1(H^1(\Omega_\varnothing))} \leq Ch.
\] (5.35)

Taking the difference of the weak formulations between the continuous and semidiscrete problems (4.5) and (5.6), one obtains

\[
\begin{cases}
(e_h, \varphi_\varnothing)_{\Omega} + a_{sh}(e_h, \varphi_\varnothing) + b_h(\varphi_\varnothing, r_h) + a_{sh}(\eta_h, \varphi_\varnothing) \\
\quad + (\mu e_h(\cdot - r), \varphi_\varnothing)_{\Omega_{sh}} = \ell_1(\varphi_\varnothing) \quad \forall \varphi_\varnothing \in W_{\varnothing}, \\
(a_{sh}(\eta_\varnothing, \psi_\varnothing) - a_{sh}(e_h, \psi_\varnothing)) = \ell_2(\psi_\varnothing) \quad \forall \psi_\varnothing \in W_{sh}, \\
b(e_h, \rho_\varnothing) = 0 \quad \forall \rho_\varnothing \in M_{\varnothing}, \\
e_h(0) = \xi_0 - \xi_{\varnothing 0}, \eta_h(0) = w_0 - w_{\varnothing 0}, e_h = z_0 - z_{\varnothing 0} \text{ in } Q_{rh},
\end{cases}
\] (5.36)

for a.e. \( t \in I \), where \( \ell_1 \) and \( \ell_2 \) are the errors due to the variational crimes, namely

\[
\ell_1(\varphi_\varnothing) = (1 - \varepsilon)(\nabla \xi - \nabla \varphi_\varnothing)_{\varnothing 0} - (Ep, \text{div } \varphi_\varnothing)_{\varnothing 0} - (w, \varphi_\varnothing)_{1,\varnothing_0}, \\
\ell_2(\psi_\varnothing) = -(\xi_t, \psi_\varnothing)_{1,\varnothing_0} + (\xi, \psi_\varnothing)_{1,\varnothing_0} - (g_1, \psi_\varnothing)_{\varnothing 0} - (\nabla g_2, \nabla \psi_\varnothing)_{\varnothing 0}.
\] (5.37)

Choosing \( \varphi_\varnothing = \hat{e}_h \) and \( \psi_\varnothing = \hat{\eta}_h \) in (5.36) we have

\[
\frac{1}{2} \frac{d}{dt} (\|e_h\|_{\Omega}^2 + \|\eta_h\|_{1,\Omega_{sh}}^2) + a_{sh}(e_h, e_h) = (e_{ht}, \hat{e}_h)_{\Omega} + a_{sh}(e_h, \hat{e}_h) - b_h(\hat{e}_h, r_h) \\
- a_{sh}(\eta_h, \hat{e}_h) - (\mu e_h(\cdot - r), \hat{e}_h)_{\Omega_{sh}} + a_{sh}(\eta_\varnothing, \hat{\eta}_h) + a_{sh}(e_h, \hat{\eta}_h) + \ell_1(\hat{e}_h) + \ell_2(\hat{\eta}_h).
\]

Since \( \hat{e}_h \) takes values in \( X_{\varnothing} \) it follows that \( (e_{ht}, \hat{e}_h)_{\Omega} = (e_{ht} - \hat{e}_{ht}, \hat{e}_h)_{\Omega} = \frac{1}{2} \frac{d}{dt} \|\hat{e}_h\|_{\Omega}^2 \) and \( b_h(\hat{e}_h, r_h) = b_h(\hat{e}_h, Ep - \rho_\varnothing) \) for every \( \rho_\varnothing \in M_{\varnothing} \). Likewise, we note that \( a_{sh}(\eta_\varnothing, \hat{\eta}_h) = \frac{1}{2} \frac{d}{dt} \|\hat{\eta}_h\|_{1,\Omega_{sh}}^2 \). Using the
equations \( \hat{e}_h = e_h - \hat{e}_h, \hat{\eta}_h = \eta_h - \hat{\eta}_h \), (5.2) and Young’s inequality we have the following estimates:

\[
|b_h(\hat{e}_h, r_h)| \leq C_e (\|Ep - \rho_h\|_{L^2(\Omega_h)}^2 + \|\hat{e}_h\|_{L^2(\Omega)}^2) + \tilde{\ell} \|e_h\|_{L^1(\Omega)},
\]

\[
|\langle \mu e_h(-r), \hat{e}_h \rangle_{L^2(\Omega)}| \leq C (\|e_h(-r)\|_{L^2(\Omega)}^2 + \|\hat{e}_h\|_{L^2(\Omega)}^2 + \|e_h\|_{L^2(\Omega)}^2),
\]

\[
|a_{sh}(\eta_h, \hat{e}_h)| + |a_{sh}(e_h, \hat{\eta}_h)| \leq C_d (\|\hat{e}_h\|_{1,\Omega_{sh}} + \|\hat{\eta}_h\|_{1,\Omega_{sh}} + \|\eta_h\|_{1,\Omega_{sh}}) + \tilde{\ell} \|e_h\|_{L^1(\Omega)}. \]

A similar strategy applied to the terms involving \( \ell_1 \) and \( \ell_2 \) and the fact that \( \xi \in H^1(\Omega) \) and \( \xi \in L^2(\Omega) \) gives

\[
|\ell_1(\hat{e}_h)| + |\ell_2(\hat{\eta}_h)| \leq C_d (h^2 + \|\hat{\eta}_h\|_{1,\Omega_{sh}}^2 + \|\hat{e}_h\|_{1,\Omega}^2 + \|\eta_h\|_{1,\Omega_{sh}}^2 + \|\hat{\eta}_h\|_{1,\Omega_{sh}}^2 + \|\hat{\eta}_h\|_{1,\Omega_{sh}}^2) + \tilde{\ell} \|e_h\|_{L^1(\Omega)}. \]

Combining these estimates, applying the approximation property of \( M_h \) and then choosing \( \tilde{\ell} \) small enough we deduce the estimate

\[
\frac{d}{dt}(\|e_h\|_{L^2(\Omega)}^2 + \|\eta_h\|_{1,\Omega_{sh}}^2) + \|\nabla e_h\|_{L^2(\Omega)}^2 \leq C \frac{d}{dt}(\|\hat{e}_h\|_{L^2(\Omega)}^2 + \|\hat{\eta}_h\|_{1,\Omega_{sh}}^2) + C(h^2 + \|e_h(-r)\|_{L^2(\Omega)}^2 + \|e_h\|_{L^2(\Omega)}^2 + \|\hat{e}_h\|_{L^2(\Omega)}^2 + \|\hat{\eta}_h\|_{1,\Omega_{sh}}^2 + \|\eta_h\|_{1,\Omega_{sh}}^2 + \|\hat{\eta}_h\|_{1,\Omega_{sh}}^2). \]

Utilizing (5.26), (5.35) and Gronwall’s lemma on this inequality we conclude that

\[
\|e_h\|_{L^\infty(L^2(\Omega)) \cap L^2(H^1(\Omega))} + \|\eta_h\|_{L^\infty(H^1(\Omega_{sh}))} \leq C h. \tag{5.38} \]

Taking \( \psi_h = \hat{\eta}_h \) in the second equation in (5.36) and rearranging terms yields

\[
a_{sh}(\hat{\eta}_h, \hat{\eta}_h) = -a_{sh}(\hat{\eta}_h, \hat{\eta}_h) + a_{sh}(e_h, \hat{\eta}_h) + \ell_2(\hat{\eta}_h). \]

Applying Young’s inequality to this equation and using (5.35) and (5.38) we deduce that

\[
\|\hat{\eta}_h\|_{H^1(\Omega_{sh})} \leq C h. \]

This estimate together with (5.35) and (5.38) implies (5.34).

Using a standard duality argument we shall establish a second order of convergence with respect to \( h \) under the \( L^2 \)-norm, which is optimal using linear elements.

**Theorem 5.9**  With the assumptions of Theorem 5.8 there exists a constant \( C > 0 \) independent of \( h \) such that

\[
\|\xi - \xi_h\|_{L^2(\Omega)}^2 + \|\xi - \xi_h\|_{L^2(\Omega_{sh})} \leq C h^2. \tag{5.39} \]
Proof. We shall use the same notation as in the proof of the previous theorem. Let \((f, g) \in L^2(L^2(\Omega)) \times L^2(L^2(\Omega_{sh}))\) and \((y, \vartheta, \pi)\) be the weak solution of

\[
\begin{align*}
-(y, \varphi)_\Omega + a_e(y, \varphi) + b(\varphi, \pi) + a_s(\vartheta, \varphi) + (\mu y(\cdot + r), \varphi)_{\Omega_s} &= (f, \varphi)_\Omega, \quad \forall \varphi \in W, \\
-a_s(\vartheta, \psi) - a_s(y, \psi) &= -(g, \psi)_{\Omega_{sh}}, \quad \forall \psi \in V_s, \\
b(y, \rho) &= 0, \quad \forall \rho \in M, \\
y(T) = 0, \vartheta(T) = 0, \quad y = 0 \text{ in } (T, T + r) \times \Omega_s.
\end{align*}
\]

(5.40)

The dual version of (4.7) can be adapted to the solution of (5.40) so that

\[
\begin{align*}
\|y\|_{L^\infty(L^2(\Omega)) \cap L^2(H^1_0(\Omega) \cap H^2(\Omega))} + \|\vartheta\|_{H^1(L^2(\Omega_s))} + \|\pi\|_{L^2(H^1(\Omega))} & \leq C(\|f\|_{L^2(L^2(\Omega))} + \|g\|_{L^2(L^2(\Omega_{sh}))}).
\end{align*}
\]

(5.41)

Denote by \((y_h, \pi_h)\) the solution of the problem in Lemma 5.4 corresponding to the pair \((y, \pi)\) and let \(\vartheta_h = R_{sh}(\vartheta).\) Then we obtain from (5.7), (5.12) and (5.41) the inequality

\[
\begin{align*}
\|y - y_h\|_{L^2(L^2(\Omega))} + \|y - y_h\|_{L^2(H^1_0(\Omega))} + h\|E\varpi - \pi_h\|_{L^2(M_h)} + \|\vartheta - \vartheta_h\|_{H^1(L^2(\Omega_{sh}))} & \leq Ch^2(\|f\|_{L^2(L^2(\Omega))} + \|g\|_{L^2(L^2(\Omega_{sh}))}).
\end{align*}
\]

(5.42)

Taking \(\varphi = e_h \) and \(\psi = \eta_h,\) or more precisely \(\psi = E_h\eta_h,\) in (5.40) and integrating by parts we obtain

\[
\begin{align*}
\int (e_h, f)_\Omega + (\eta_h, g)_{\Omega_{sh}} dt &= \int (e_h(0), y(0))_\Omega + \int (e_h(t), y(t))_\Omega + a_e(e_h, y) dt + b(e_h, \pi) dt + \int a_s(e_h, \vartheta) dt + (\mu y(\cdot + r), e_h)_{\Omega_s} dt - a_s(\vartheta(0), \eta_h(0)) - \int a_s(\vartheta, \eta_h) dt - a_s(y, \eta_h) dt.
\end{align*}
\]

(5.43)

If we take \(\varphi = y_h\) in (5.36) then we have for every \(\rho_h \in M_h,\)

\[
\begin{align*}
\int (e_h(t), y(t))_\Omega dt &= \int (e_h(t), y - y_h(t))_\Omega + (e_h(t), y_h(t)) dt \\
&= \int (e_h(t), y - y_h(t))_\Omega dt - \int a_s(e_h(t), y_h(t)) + b_h(y_h(t), E\varpi - \rho_h(t)) + a_{sh}(\eta_h(t), y_h(t)) dt \\
&\quad - \int (y_h(t), \mu e_h(\cdot - r))_{\Omega_{sh}} dt - \ell_1(y_h(t)) dt.
\end{align*}
\]

(5.44)

Similarly, if we take \(\psi = \vartheta_h\) in (5.36) then we obtain

\[
\begin{align*}
\int a_{sh}(\vartheta, \eta_{h(t)}) dt &= \int a_{sh}(\vartheta - \vartheta_h, \eta_{h(t)}) dt + a_{sh}(\vartheta_h, \eta_{h(t)}) dt \\
&= \int a_{sh}(\vartheta - \vartheta_h, \eta_{h(t)}) dt + \int a_{sh}(\vartheta_h, e_h(t)) + \ell_2(\vartheta_h) dt.
\end{align*}
\]

(5.45)
where \( \ell_1 \) and \( \ell_2 \) are given by (5.37). Substituting (5.44) and (5.45) in (5.43) and applying the Galerkin orthogonality property \( a_{sh}(\eta_h(0), \vartheta(0)) = a_{sh}(\eta_h(0), \vartheta(0) - \vartheta_h(0)) \) will then result in the following equation:

\[
\int_I (e_h, f)_{\Omega} + (\eta_h, g)_{\Omega_{sh}} \, dt = (e_h(0), y(0))_{\Omega} + \int_I (e_h, y - y_h)_{\Omega} + a_{sh}(e_h, y - y_h) + b_h(e_h, \Theta) \, dt \\
+ \int_I b_h(y_h, \Theta) + a_{sh}(\varrho - \varrho_h, \eta_h) \, dt \\
+ \int_I \mu(\varrho + r) - \eta_h(\varrho + r), e_h)_{\Omega_{sh}} \, dt + \int_I \mu(e_h, \eta_h(\varrho + r))_{\Omega_{sh}} \, dt \\
+ \int_I \ell_1(y_h) - \ell_2(\varrho_h) + \ell_3(\eta_h) + \ell_4(e_h) \, dt - a_{sh}(\eta_h(0), \vartheta(0) - \vartheta_h(0)) \\
- (\vartheta(0), \eta_h(0))_{1,oh},
\]

(5.46)

where \( \ell_3(\eta_h) \) and \( \ell_4(e_h) \) are the residuals due to the variational crimes given by

\[
\ell_3(\eta_h) = -(\vartheta, \eta_h)_{1,oh}, \\
\ell_4(e_h) = (\text{div} e_h, \Theta)_{oh} + b_h(e_h, \pi_h) - (1 - \epsilon)(\nabla e_h, \nabla y)_{oh} + (e_h, \varrho)_{1,oh}
\]

and \( I_r = (-r, 0) \). From the definitions of \( y_h, \pi_h \) and \( \vartheta_h \) we have the Galerkin orthogonality \( a_{eh}(e_h, y - y_h) + b_h(e_h, \Theta - \pi_h) = 0 \) and \( a_{sh}(\varrho - \varrho_h, \eta_h) = 0 \).

We are now going to estimate the remaining terms on the right-hand side of (5.46). In what follows, we shall frequently use (5.41) and (5.42). From (5.2), (5.34) and the boundedness of \( e_{ht} \) in \( L^2(L^2(\Omega)) \), which is guaranteed according to (4.7) and (5.27), we have

\[
(e_h(0), y(0))_{\Omega} + \int_I (e_h, y - y_h)_{\Omega} \, dt + \int_I \mu(\varrho + r) - \eta_h(\varrho + r), e_h)_{\Omega_{sh}} \, dt \\
+ a_{sh}(\eta_h(0), \vartheta(0) - \vartheta(h(0)) \leq Ch^2(\|f\|_{L^2(L^2(\Omega))} + \|g\|_{L^2(L^2(\Omega_{sh}))}).
\]

(5.47)

Using the approximation property of \( M_h \) and (5.2) we obtain

\[
\int_I b_h(y_h, \Theta - \rho_h) \, dt = \int_I b_h(y - y_h, \Theta - \rho_h) + (\text{div} y, \Theta - \rho_h)_{oh} \, dt \\
\leq Ch^2(\|f\|_{L^2(L^2(\Omega))} + \|g\|_{L^2(L^2(\Omega_{sh}))}).
\]

(5.48)

The error estimate for the initial history in (5.26) implies that

\[
\int_{I_r} \mu(e_h, \eta_h(\varrho + r))_{\Omega_{sh}} \, dt \leq Ch^2(\|f\|_{L^2(L^2(\Omega))} + \|g\|_{L^2(L^2(\Omega_{sh}))}).
\]

(5.49)
By writing $\ell_1(y_h) = \ell_1(y_h - y) + \ell_1(y)$ and $\ell_2(\vartheta_h) = \ell_2(\vartheta_h - \vartheta) + \ell_2(\vartheta)$ it can be verified using analogous methods to the proof of the previous theorem that

$$\int t \ell_1(y_h) - \ell_2(\xi_h) + \ell_4(e_h) \, dt \leq C h^2(\|f\|_{L^2(L^2(\Omega))} + \|g\|_{L^2(L^2(\Omega_{sh}))}).$$

(5.50)

Finally, from (5.2), (5.34) and Lemma 5.2 we have

$$(\vartheta(0), \eta_h(0))_{1,sh} + \int t \ell_3(\eta_h) \, dt \leq C h^2(\|f\|_{L^2(L^2(\Omega))} + \|g\|_{L^2(L^2(\Omega_{sh}))}).$$

(5.51)

Consequently, using estimates (5.47)–(5.51) in (5.46), we obtain

$$\int t(e_h f)_\Omega + (\eta_h,g)_{\Omega_{sh}} \, dt \leq C h^2(\|f\|_{L^2(L^2(\Omega))} + \|g\|_{L^2(L^2(\Omega_{sh}))}),$$

from which the $L^2$-estimate (5.39) follows by duality.

\[\square\]

6. Full discretization of the symmetric formulation

In this section we discuss the full space-time discretization of the variational problem (4.5). This is done by discretizing the semidiscrete problem (5.6) with respect to time. We use a discontinuous Galerkin approach for the time discretization, which can be viewed as an implicit Euler scheme.

First, let us set some notation and assumptions. Let $\tau_r = \frac{T}{n_r}$ for a given positive integer $n_r$ and $I_j = [-r + (\ell - 1)\tau_r, -r + \ell \tau_r]$ for $1 \leq \ell \leq n_r - 1$ and $I_{n_r} = [-\tau_r, 0]$. With regard to the discretization of the history we consider the space

$$Z_{t,sh} = \{z_{t,sh} \in L^2_r(W_{sh}) : z_{t,sh} = \sum_{j=1}^{n_r} z_{j,h} x_{I_j}, z_{j,h} \in W_{sh}\}.$$  

We also consider a possibly finer time-step size compared to the history interval, namely $\tau = \frac{\tau_r}{N_{r}}$ for some integer $N_{r}$. For simplicity of presentation we suppose that $T = N_{r}r$ for some integer $N_{r}$. Note that if $T$ is not a multiple of $r$ then the length of the last subinterval in the temporal mesh will be less than $\tau$. Nonetheless, the succeeding analysis can be applied to this case. Thus, we have $T = N_{r}\tau$ with $N_{r} = N_{r}n_{r}n_{r}$. Let $J_0 = \{0\}$ and $J_\ell = ((\ell - 1)k, \ell k]$ for $1 \leq \ell \leq N_{r}$. For the spatio-temporal discretization of the state equations we use the spaces

$$W_{kh} = \{\xi_{kh} \in L^2(W_h) : \xi_{kh} = \sum_{j=0}^{N_{r}} \xi_{j,h} x_{I_j}, \xi_{j,h} \in W_h\},$$

$$M_{kh} = \{p_{kh} \in L^2(M_h) : p_{kh} = \sum_{j=0}^{N_{r}} p_{j,h} x_{I_j}, p_{j,h} \in M_h\}.$$  

We also define $W_{skh} = \{\xi_{kh} \in \Omega_{sh} : \xi_{kh} \in W_{kh}\}.$
Let $X$ be a given Hilbert space. For each $\varphi \in L^2(X)$ and $z \in L^2_r(X)$ we define $\Pi_k \varphi \in L^2(X)$ and $\Pi_r z \in L^2_r(X)$ according to

$$\Pi_k \varphi = \sum_{j=1}^{N_r} \varphi_j x_j, \quad \varphi_j = \frac{1}{\tau} \int_{J_j} \varphi(s) \, ds,$$

$$\Pi_r z = \sum_{\ell=1}^{n_r} z_\ell x_{\ell}, \quad z_\ell = \frac{1}{\tau_r} \int_{I_{\ell}} z(\theta) \, d\theta.$$

Recall that there exists a constant $C > 0$ independent of $\tau$ and $\tau_r$ such that for each $\varphi \in H^1(X)$, $z \in H^1_r(X)$, $1 \leq j \leq N_r$ and $1 \leq \ell \leq n_r$ we have

$$\|\Pi_k \varphi - \varphi\|_{L^2(J_j, X)} \leq C \|\varphi\|_{H^1(J_j, X)}, \quad (6.1)$$

$$\|\Pi_r z - z\|_{L^2(I_{\ell}, X)} \leq C \tau_r \|z\|_{H^1(I_{\ell}, X)}, \quad (6.2)$$

The full discretization of the history $z_0$ will be $\Pi_r z_{0h} \in Z_{eh}$.

For the space-time discretization of (4.5) we consider the approximations

$$\xi_{kh} = \sum_{k=0}^{N_r} \xi_{k,h} x_k \in W_{kh}, \quad \xi_{kh} = \sum_{k=0}^{N_r} \xi_{k,h} x_k \in W^{sh}_{kh}, \quad p_{kh} = \sum_{k=0}^{N_r} p_{k,h} x_k \in M_{kh},$$

where for $k = 1, \ldots, N_r$,

$$\frac{1}{\tau} (\xi_{k,h} - \xi_{k-1,h} , \varphi_h)_{\Omega} + a_{sh}(\xi_{k,h}, \varphi_h) + b_h(\varphi_h, p_{k,h}) + a_{sh}(\xi_{k,h}, \varphi_h)$$

$$+ (\mu \xi_{k-1,n_r,h} , \varphi_h)_{\Omega_0} = \frac{1}{\tau} \int_{J_k} (f_1, \varphi_h)_{\Omega_{kh}} + (f_2, \varphi_h)_{\Omega_{sh}} \, dt \quad \forall \varphi_h \in W_h,$$

$$\frac{1}{\tau} a_{sh}(\xi_{k,h} - \xi_{k-1,h} , \psi_h) - a_{sh}(\xi_{k,h} , \psi_h)$$

$$= - \frac{1}{\tau} \int_{J_k} (g_1, \psi_h)_{\Omega_{sh}} + (V g_2, V \psi_h)_{\Omega_{sh}} \, dt \quad \forall \psi_h \in W_{sh},$$

$$b_h(\xi_{k,h} , \rho_h) = 0 \quad \forall \rho_h \in M_h,$$

$$\xi_{0,h} = \xi_{0h}, \quad \xi_{0,h} = w_{0h}, \quad \xi_{j,n_r,h} = z_{0h} \big|_{I_{n_r-j+1}} \text{ for } 1 \leq j \leq n_r, 0 \leq k \leq n_r - 1.$$

The existence and uniqueness of $(\xi_{k,h}, w_{kh}, p_{kh}) \in X_h \times W_{sh} \times M_h$ for each $k = 1, \ldots, N_r$ satisfying (6.3) follows from the discrete inf–sup condition (5.3) and an induction argument.

**Remark 6.1** From the second equation in (6.3) one can see that

$$\frac{1}{\tau} (\xi_{k,h} - \xi_{k-1,h}) = \xi_{k,h} + \frac{1}{\tau} \int_{J_k} (v_h - R_{sh} g_2) \, dt,$$
where \( v_h \) is given by (5.8), and in particular, if \( g_1 = g_2 \) we have

\[
\frac{1}{\tau} (\xi_{h,k} - \xi_{h,k-1}) = \xi_{k,h} - \frac{1}{\tau} \int_{J_k} R_{sh} g_2 \, dt.
\]

These are in fact the backward Euler discretizations of ordinary differential equations (5.9) and (5.10), respectively.

Let us prove the stability of the scheme (6.3). We shall use the abbreviations \([\xi]_{kh} = \xi_{k,h} - \xi_{k-1,h}\) and \([\zeta]_{kh} = \zeta_{k,h} - \zeta_{k-1,h}\) for the jumps. Also, let \( J_\ell = J_{n_r, \ell+1} \cap \cdots \cap J_{n_r, (\ell+1)} = (\ell r, (\ell + 1) r) \) for \( 0 \leq \ell \leq N_r - 1 \) and

\[
N_h = \| f_1 \|_{L^2(\Omega_h)} + \| f_2 \|_{L^2(\Omega_{sh})} + \| g_1 \|_{L^2(\Omega_{sh})} + \| \triangledown g_2 \|_{L^2(\Omega_{sh})} + \| \xi_{0h} \|_{\Omega} + \| w_{0h} \|_{1, \Omega_{sh}} + \| \zeta_{0h} \|_{\Omega_h}.
\]

**Theorem 6.2** If \( (\xi_{k,h}, \zeta_{k,h}, p_{k,h}) \) is the solution of (6.3) then there exists a constant \( C > 0 \) independent of \( h \) and \( \tau \) such that

\[
\sum_{k=1}^{N_r} (\| [\xi]_{kh} \|_{\Omega}^2 + \| [\xi]_{kh} \|_{1, \Omega_{sh}}^2) + \tau a_{kh}(\xi_{k,h}, \xi_{k,h}) + \max_{1 \leq k \leq N_r} (\| \xi_{k,h} \|_{\Omega}^2 + \| \xi_{k,h} \|_{1, \Omega_{sh}}^2) \leq CN_h^2. \tag{6.4}
\]

In particular, we have \( \| \xi_{kh} \|_{L^\infty(\Omega)} + \| \zeta_{kh} \|_{L^\infty(H^1(\Omega))} \leq CN_h \).

**Proof.** Choosing \( \varphi_h = \xi_{k,h} \) and \( \psi_h = \zeta_{k,h} \) as test functions in (6.3), using \( b_h(\xi_{k,h}, p_{k,h}) = 0 \) and then taking the sum we obtain the identity

\[
\frac{1}{2\tau} [\| \xi_{k,h} \|_{\Omega}^2 - \| \xi_{k-1,h} \|_{\Omega}^2 + \| [\xi]_{kh} \|_{\Omega}^2 + \| \xi_{k,h} \|_{1, \Omega_{sh}}^2 - \| \xi_{k-1,h} \|_{1, \Omega_{sh}}^2 + \| [\xi]_{kh} \|_{1, \Omega_{sh}}^2] + a_{kh}(\xi_{k,h}, \zeta_{k,h}) + (\mu \xi_{k-1-n_r, h}, \xi_{k,h})_{\Omega_{sh}} = R_k, \tag{6.5}
\]

where \( R_k = \frac{1}{\tau} \int_{J_k} (f_1, \xi_{k,h})_{\Omega_{sh}} + (f_2, \xi_{k,h})_{\Omega_{sh}} - (g_1, \zeta_{k,h})_{\Omega_{sh}} - (\triangledown g_2, \triangledown \zeta_{k,h})_{\Omega_{sh}} \, dt. \) This term can be estimated using Young's inequality:

\[
|R_k| \leq \frac{1}{8\tau} (\| \xi_{k,h} \|_{\Omega}^2 + \| \zeta_{k,h} \|_{1, \Omega_{sh}}^2) + C(\| f_1 \|_{L^2(\Omega_{sh})} + \| f_2 \|_{L^2(\Omega_{sh})} + \| g_1 \|_{L^2(\Omega_{sh})} + \| \triangledown g_2 \|_{L^2(\Omega_{sh})}. \tag{6.6}
\]

For the term associated with delay we use \( \tau = \frac{\tau}{n_r} \) to obtain

\[
|\mu \xi_{k-1-n_r, h}, \xi_{k,h})_{\Omega_{sh}} | \leq \frac{1}{8\tau} (\| \xi_{k,h} \|_{\Omega_{sh}}^2 + \frac{C\tau}{n_r} \| \xi_{k-1-n_r, h} \|_{\Omega_{sh}}^2. \tag{6.7}
\]
Considering indices \(1 \leq k \leq n_{r_{\tau}}\) observe that
\[
\sum_{k=1}^{n_{r_{\tau}}} n_{r_{\tau}} \left\| \xi_{k-n_{r_{\tau}}-1,h} \right\|_{\Omega_{sh}}^2 = n_{\tau} \left\| \Pi_{r}z_{0h} \right\|_{Q_{sh}}^2 \leq Cn_{\tau} \left\| z_{0h} \right\|_{Q_{sh}}^2 ,
\] (6.8)
and therefore by taking the sum over all such indices we obtain from (6.5)–(6.8),
\[
\sum_{k=1}^{n_{r_{\tau}}} \left( \left\| \xi_{k,h} \right\|_{\Omega}^2 + \left\| \xi_{k,h} \right\|_{1,\Omega_{sh}}^2 + \tau a_{eh}(\xi_{k,h},\xi_{k,h}) \right) + \max_{1 \leq k \leq n_{r_{\tau}}} \left\{ \left( \left\| \xi_{k,h} \right\|_{\Omega}^2 + \left\| \xi_{k,h} \right\|_{1,\Omega_{sh}}^2 \right) \right\}
\leq C \left( \left\| f_{1} \right\|_{J_{h} \times \Omega_{sh}}^2 + \left\| f_{2} \right\|_{J_{h} \times \Omega_{sh}}^2 + \left\| g_{1} \right\|_{J_{h} \times \Omega_{sh}}^2 + \left\| \nabla g_{2} \right\|_{J_{h} \times \Omega_{sh}}^2 \right)
+ C \left( \left\| \xi_{0h} \right\|_{2,\Omega_{sh}}^2 + \left\| w_{0h} \right\|_{1,\Omega_{sh}}^2 + \left\| z_{0h} \right\|_{Q_{sh}}^2 \right) .
\] (6.9)

Suppose that \(\ell > 0\). Taking the sum of (6.5) for \(n_{r_{\tau}} \ell + 1 \leq k \leq n_{r_{\tau}} (\ell + 1)\), invoking (6.6) and (6.7) and then reindexing yields
\[
\sum_{k=n_{r_{\tau}} \ell + 1}^{n_{r_{\tau}} (\ell + 1)} \left( \left\| \xi_{k,h} \right\|_{\Omega}^2 + \left\| \xi_{k,h} \right\|_{1,\Omega_{sh}}^2 + \tau a_{eh}(\xi_{k,h},\xi_{k,h}) \right) + \max_{n_{r_{\tau}} \ell + 1 \leq k \leq n_{r_{\tau}} (\ell + 1)} \left\{ \left( \left\| \xi_{k,h} \right\|_{\Omega}^2 + \left\| \xi_{k,h} \right\|_{1,\Omega_{sh}}^2 \right) \right\}
\leq C \left( \left\| f_{1} \right\|_{J_{\ell} \times \Omega_{sh}}^2 + \left\| f_{2} \right\|_{J_{\ell} \times \Omega_{sh}}^2 + \left\| g_{1} \right\|_{J_{\ell} \times \Omega_{sh}}^2 + \left\| \nabla g_{2} \right\|_{J_{\ell} \times \Omega_{sh}}^2 \right)
+ C \left( \left\| \xi_{n_{r_{\tau}} \ell} \right\|_{2,\Omega_{sh}}^2 + \left\| w_{n_{r_{\tau}} \ell} \right\|_{1,\Omega_{sh}}^2 + \sum_{k=n_{r_{\tau}} (\ell - 1)}^{n_{r_{\tau}} \ell - 1} \frac{\tau_{r}}{n_{\tau}} \left\| \xi_{k,h} \right\|_{Q_{sh}}^2 \right) .
\] (6.10)
The summation on the right-hand side of (6.10) can be estimated by
\[
\sum_{k=n_{r_{\tau}} (\ell - 1)}^{n_{r_{\tau}} \ell - 1} \frac{\tau_{r}}{n_{\tau}} \left\| \xi_{k,h} \right\|_{Q_{sh}}^2 \leq r \max_{n_{r_{\tau}} (\ell - 1) \leq k \leq n_{r_{\tau}} \ell - 1} \left\| \xi_{k,h} \right\|_{Q_{sh}}^2 .
\] (6.11)
Using (6.9)–(6.11) along with an induction argument proves (6.4). \(\Box\)

Next we establish an error estimate for the semidiscrete and fully discrete problems (5.6) and (6.3). For the proof we introduce a projection operator \(r_{k} : H^{1}(X) \rightarrow L^{2}(X)\) for a given Hilbert space \(X\) as follows: given \(\varphi \in H^{1}(X)\) define \(r_{k}\varphi = \sum_{j=0}^{N_{k}} \varphi(t_{j}) \chi_{j}\). Note that for some \(C > 0\) independent of \(\varphi \in H^{1}(X)\) and \(k\) we have
\[
\left\| r_{k}\varphi - \varphi \right\|_{L^{2}(J_{k},X)} \leq C_{\tau} \left\| \varphi \right\|_{H^{1}(J_{k},X)} , \quad 1 \leq k \leq N_{\tau} .
\] (6.12)
Theorem 6.3 Let \((\hat{\xi}_{kh}, \hat{\zeta}_{kh})\) and \((\hat{\xi}_h, \hat{\zeta}_h)\) be the solutions of (6.3) and (5.28), respectively. For each \(\tau_0 > 0\) there is a constant \(C = C(\tau_0) > 0\) independent of \(h, \tau_r\) and \(\tau \in (0, \tau_0)\) such that
\[
\|\hat{\xi}_{kh} - \hat{\xi}_h\|_{L^2(I, H^1(\Omega))} + \|\hat{\zeta}_{kh} - \hat{\zeta}_h\|_{L^2(I, H^1(\Omega_{sh}))} \leq C(\tau + \tau_r).
\]

Proof. Let \(e_{kh} = \hat{e}_h - \hat{e}_h\) and \(\eta_{kh} = \hat{\eta}_h - \hat{\eta}_h\) be the discretization errors. Separate these into \(e_{kh} = \hat{e}_{kh} + \hat{e}_{kh}\) and \(\eta_{kh} = \hat{\eta}_{kh} - \hat{\eta}_{kh}\), where \(\hat{e}_{kh} = \hat{e}_h - r_k \hat{\xi}_h\), \(\hat{e}_{kh} = r_k \hat{\xi}_h - \hat{\xi}_{kh}\), \(\hat{\eta}_{kh} = \hat{\zeta}_h - r_k \hat{\zeta}_h\) and \(\hat{\eta}_{kh} = r_k \hat{\zeta}_h - \hat{\zeta}_{kh}\). Observe that \(\hat{e}_{kh}\) and \(\hat{\eta}_{kh}\) are constant on each interval \(J_k\), with values in \(X_h\) and \(W_{sh}\), respectively. We extend the function \(e_{kh}\) to \(I_r\) according to \(e_{kh} = z_{Oh} - \Pi_{I_r} z_{Oh}\).

Integrating the semidiscrete problem (5.28) over \(J_k\), subtracting it from (6.3) and taking \(\varphi_h = \hat{e}_{kh}|_{J_k}\) and \(\psi_h = \hat{\eta}_{kh}|_{J_k}\) we have
\[
(\hat{e}_{kh}|_{J_k} - \hat{e}_{kh}|_{J_{k-1}}, \hat{e}_{kh}|_{J_k})_\Omega + \int_{J_k} \{a_{eh}(e_{kh}, \hat{e}_{kh}) + a_{eh}(\eta_{kh}, \hat{e}_{kh})\} dt
+ \int_{J_k} (\mu e_{kh}(-r), \hat{e}_{kh})_{\Omega_{sh}} dt + a_{eh}(\hat{\eta}_{kh}|_{J_k} - \hat{\eta}_{kh}|_{J_{k-1}}, \hat{\eta}_{kh}|_{J_k}) - \int_{J_k} a_{eh}(e_{kh}, \hat{\eta}_{kh}) dt = 0.
\]

Let us estimate each of the integrals in this equation. Using (6.12) we obtain
\[
\int_{J_k} a_{eh}(e_{kh}, \hat{e}_{kh}) dt = \int_{J_k} a_{eh}(\hat{e}_{kh}, \hat{e}_{kh}) dt + \int_{J_k} a_{eh}(\hat{e}_{kh}, \hat{e}_{kh}) dt
\geq (1 - \rho) \int_{J_k} a_{eh}(\hat{e}_{kh}, \hat{e}_{kh}) dt - C_\rho \tau^2 \|\hat{\xi}_h\|_{H^1(I, H^1(\Omega))}^2.
\]

Also, since \(a_{eh}(\eta_{kh}, \hat{e}_{kh}) - a_{eh}(e_{kh}, \hat{\eta}_{kh}) = a_{eh}(\hat{\eta}_{kh}, \hat{e}_{kh}) - a_{eh}(\hat{e}_{kh}, \hat{\eta}_{kh})\) we have
\[
\int_{J_k} a_{eh}(\eta_{kh}, \hat{e}_{kh}) - a_{eh}(e_{kh}, \hat{\eta}_{kh}) dt \geq - \rho \int_{J_k} \{a_{eh}(\hat{\eta}_{kh}, \hat{\eta}_{kh}) + a_{eh}(\hat{e}_{kh}, \hat{e}_{kh})\} dt
- C_\rho \tau^2 (\|\hat{\xi}_h\|_{H^1(I, H^1(\Omega_{sh}))}^2 + \|\hat{\zeta}_h\|_{H^1(I, H^1(\Omega_{sh}))}^2).
\]

With regard to the delay term we estimate it as
\[
\int_{J_k} (\mu e_{kh}(\cdot - r), \hat{e}_{kh})_{\Omega_{sh}} dt \geq - \frac{\rho}{\tau} \int_{J_k} \|\hat{e}_{kh}\|_{\Omega_{sh}}^2 dt - C_\rho \tau \int_{J_k} \|e_{kh}(\cdot - r)\|_{\Omega_{sh}}^2 dt.
\]

Set \(\tilde{J}_{k-1} = I_r\). Taking the sum of (6.14) over all \(n_n, n_r, \ell + 1 \leq k \leq n_n, n_r (\ell + 1)\) for \(0 \leq \ell \leq N_r - 1\), using estimates (6.15)–(6.17) and Theorem 5.7, we deduce that
\[
\sup_{t \in \tilde{J}_\ell} (1 - \rho) \int_{J_{\ell}} \|\hat{\xi}_{kh}^2\|_{\Omega}^2 + (1 - \rho) \int_{J_{\ell}} \|\hat{\eta}_{kh}^2\|_{\Omega_{sh}}^2 + (1 - 2\rho) \int_{J_{\ell}} \|\hat{e}_{kh}\|_{\Omega_{sh}}^2 dt
\leq C_\rho \tau^2 + C(\|\hat{e}_{kh}(\ell r)\|_{\Omega}^2 + \|\hat{\eta}_{kh}(\ell r)\|_{\Omega_{sh}}^2) + C_\rho \tau \int_{J_{\ell-1}} \|e_{kh}\|_{\Omega_{sh}}^2 dt.
\]

(6.18)
Consider the case $\ell = 0$. Note that $\hat{\xi}_{kh}(0) = 0$, $\hat{\eta}_{kh}(0) = 0$ and from (6.2),

$$
\int_{J_{-1}}^{} \| e_{kh} \|_{\Omega_{sh}}^2 \, dt = \int_{-r}^{0} \| z_{0h} - \Pi_{h}z_{0h} \|_{\Omega_{sh}}^2 \, dt \leq C_n \tau_r^2 \| z_{0h} \|_{H^1_\ell(\Omega_{sh})}^2.
$$

Choosing $\rho < \frac{1}{\tau_0}$ in (6.18), using (6.19) and $n_r \tau_r = r$ leads to the estimate

$$
\sup_{t \in J_0} (\| \hat{\xi}_{kh} \|_{L^2(\tilde{J}, H^1_\ell(\Omega)))} + \| \hat{\eta}_{kh} \|_{L^2(\tilde{J}, H^1_\ell(\Omega_{sh}))}) \leq C \tau (\| \xi \|_{H^1_\ell(\tilde{J}, H^1(\Omega)))} + \| \xi \|_{H^1_\ell(\tilde{J}, H^1(\Omega_{sh}))})
$$

for each $0 \leq \ell \leq N_r - 1$. Therefore, from (6.20), (6.21) and Theorem 5.7,

$$
\| e_{kh} \|_{L^2(\tilde{J}, H^1_\ell(\Omega)))} + \| \eta_{kh} \|_{L^2(\tilde{J}, H^1_\ell(\Omega_{sh}))} \leq C (\tau + \tau_r).
$$

Continuing this process, and using an induction argument, one can infer that

$$
\| e_{kh} \|_{L^2(\tilde{J}, H^1_\ell(\Omega)))} + \| \eta_{kh} \|_{L^2(\tilde{J}, H^1_\ell(\Omega_{sh}))} \leq C (\tau + \tau_r)
$$

for each $0 \leq \ell \leq N_r - 1$. Taking the sum of (6.22) over all such indices $\ell$ and noting that $N_r = \frac{T}{r}$ we obtain the error estimate (6.13).

Combining Theorems 5.9 and 6.2 we obtain the following error estimate between the solutions of the continuous and the fully discrete problems. In the succeeding discussions we assume that $\tau \in (0, \tau_0)$ for a given fixed $\tau_0 > 0$.

**Corollary 6.4** Suppose that the conditions of Theorem 5.9 hold. Let $(\xi, \zeta)$ and $(\xi_{h,k}, \zeta_{h,k})$ be the solutions of (4.5) and (6.3), respectively. Then there exists $C > 0$ independent of $h$, $\tau_r$ and $\tau$ such that

$$
\| \xi - \xi_{kh} \|_{L^2(\tilde{J}, \Omega))} + \| \zeta - \zeta_{kh} \|_{L^2(\tilde{J}, \Omega_{sh}))} + h \| \xi - \zeta_{kh} \|_{L^2(\tilde{J}, H^1(\Omega_{sh}))} \leq C (\tau + \tau_r + h^2).
$$

**Remark 6.5** Instead of (5.24) we can use $\tilde{\xi}_{0h} = i_k \tilde{\xi}_0$, $\tilde{w}_{0h} = R_{ik}w_0$ and $z_{0h}(\theta) = \Pi_{r}i_{sh}z_0(\theta)$ as the approximation of the initial data and history. The order $O(\tau + \tau_r + h^2)$ given in Corollary 6.4 is preserved by applying the stability estimate in Theorem 6.2, along with interpolation error estimates.
To end this section we shall write the full space-time discretization of the state and adjoint equations for future reference. Recall from Section 4 that the weak formulation of the state equation is equivalent to

\[
\begin{align*}
&\left(\xi, \varphi\right)_\Omega + a_s(\xi, \varphi) + b(\varphi, p) + a_s(\xi, \varphi) + (\mu \xi \cdot - r, \varphi)_\Omega = (q, \varphi)_\Omega \quad \forall \varphi \in W, \\
a_s(w, \psi) - a_s(\xi, \psi) = 0 \quad \forall \psi \in V_s, \\
b(\xi, \rho) = 0 \quad \forall \rho \in M, \\
\xi(0) = \xi_0, \quad w(0) = w_0, \\
\xi = z_0 \quad \text{in } Q_r,
\end{align*}
\]

(6.23)

and for the adjoint equation its weak formulation is equivalent to

\[
\begin{align*}
&-(y, \varphi)_\Omega + a_s(y, \varphi) + b(\varphi, \pi) + a_s(\varphi, \varphi) + (\mu y \cdot + r, \varphi)_\Omega \\
&= \gamma_f(\xi - u_d, \varphi)_\Omega + \gamma_{s1}(\xi - v_d, \varphi)_\Omega \quad \forall \varphi \in W, \\
&-a_s(\varphi, \varphi) - a_s(y, \varphi) = -\gamma_{s2}(w - w_d, \psi)_\Omega - \gamma_{s3}(\nabla w - \nabla w_d, \nabla \psi)_\Omega \quad \forall \psi \in V_s, \\
b(y, \rho) = 0 \quad \forall \rho \in M, \\
y(T) = 0, \quad \vartheta(T) = 0, \\
\vartheta(\theta) = 0, \quad \theta \in (T, T + r),
\end{align*}
\]

(6.24)

where we recall that \(\xi = u \chi_{\Omega_{\ell_j}} + w \chi_{\Omega_{\Omega_j}}\). Therefore, the dG(0)–cG(1) space-time discretizations of (6.23) and (6.24) are given by

\[
\begin{align*}
&\frac{1}{\tau}(\xi_{k,h} - \xi_{k-1,h}, \varphi_{h})_\Omega + a_{sh}(\xi_{k,h}, \varphi_{h}) + b_h(\varphi_{h}, p_{k,h}) + a_{sh}(w_{k,h}, \varphi_{h}) \\
&+ (\mu \xi_{k-1-n_{\Omega_j}, h}, \varphi_{h})_\Omega = \frac{1}{\tau} \int_{J_k} (q, \varphi_{h})_\Omega \, dt \quad \forall \varphi_{h} \in W_h, \\
&\frac{1}{\tau} a_{sh}(w_{k,h} - w_{k-1,h}, \psi_{h}) - a_{sh}(\xi_{k,h}, \psi_{h}) = 0 \quad \forall \psi_{h} \in W_{sh}, \\
b_h(\xi_{k,h}, \rho_{h}) = 0 \quad \forall \rho_{h} \in M_h, \\
\xi_{0,h} = \hat{\xi}_{0,h}, \quad w_{0,h} = \omega_{0,h}, \\
\xi_{-n_{\Omega_j}+1,h}\Omega_{\Omega_{sh}} = \Pi_{r=0}^{n_{\Omega_j} \Omega_{sh}} \text{ for } 1 \leq j \leq n_r, \quad 0 \leq \ell \leq n_{\tau} - 1
\end{align*}
\]

(6.25)

for \(k = 1, \ldots, N_{\tau}\), and

\[
\begin{align*}
&-\frac{1}{\tau}(y_{k,h} - y_{k-1,h}, \varphi_{h})_\Omega + a_{sh}(y_{k-1,h}, \varphi_{h}) + b_h(\varphi_{h}, \pi_{k-1,h}) + a_{sh}(\vartheta_{k-1,h}, \varphi_{h}) \\
&+ \mu (y_{k+n_{\Omega_j}, h}, \varphi_{h})_\Omega = \frac{1}{\tau} \int_{J_k} \gamma_f(\xi - u_d, \varphi_{h})_\Omega + \gamma_{s1}(\xi - v_d, \varphi_{h})_\Omega \, dt \quad \forall \varphi_{h} \in W_h \\
&-\frac{1}{\tau} a_{sh}(\vartheta_{k,h} - \vartheta_{k-1,h}, \psi_{h}) - a_{sh}(y_{k-1,h}, \psi_{h}) \\
&= -\frac{1}{\tau} \int_{J_k} \gamma_{s2}(w - w_d, \psi_{h})_\Omega + \gamma_{s3}(\nabla w - \nabla w_d, \nabla \psi_{h})_\Omega \, dt, \quad \forall \psi_{h} \in W_{sh}, \\
b_h(y_{k-1,h}, \rho_{h}) = 0, \quad \forall \rho_{h} \in M_h, \\
y_{N_{\tau},h} = 0, \quad \vartheta_{N_{\Omega_j}, h} = 0, \\
y_{\ell,h}|_{\Omega_{\Omega_{sh}}} = 0 \text{ for } N_{\tau} + 1 \leq \ell \leq N_{\tau} + n_{\tau} \text{, }
\end{align*}
\]

(6.26)

for \(k = N_{\tau}, \ldots, 1\), respectively.
7. Error analysis for the optimal control problem

Given a control \( q \in Q \) we denote by \((\xi(q), w(q))\) and \((y(q), \theta(q))\) the solutions of the state and adjoint equations of (6.23) and (6.24), respectively. Likewise, we let \((\xi_{kh}(q), w_{kh}(q))\) be the solution of (6.25) and \((y_{kh}(q), \theta_{kh}(q))\) the solution of (6.26), where instead of \( \xi \) and \( w \) we have \( \xi_{kh}(q) \) and \( w_{kh}(q) \). We would like to recall for the reader our hypotheses on the initial and desired data in (A).

**Lemma 7.1** Given \( q \in H^1(L^2(\Omega)) \) there is \( C > 0 \) independent of \( q, h, \tau_r \) and \( \tau \) such that

\[
\|\xi(q) - \xi_{kh}(q)\|_{L^2(L^2(\Omega))} + \|w(q) - w_{kh}(q)\|_{L^2(L^2(\Omega_{sh}))} + h\|w(q) - w_{kh}(q)\|_{L^2(H^1(\Omega_{sh}))} \leq C(\tau + \tau_r + h^2),
\]

\( (7.1) \)

\[
\|y(q) - y_{kh}(q)\|_{L^2(L^2(\Omega))} + \|\theta(q) - \theta_{kh}(q)\|_{L^2(L^2(\Omega_{sh}))} + h\|\theta(q) - \theta_{kh}(q)\|_{L^2(H^1(\Omega_{sh}))} \leq C(\tau + \tau_r + h^{\kappa(Y_{s3})}),
\]

\( (7.2) \)

where

\[
\kappa(Y_{s3}) = \begin{cases} 1 & \text{if } Y_{s3} > 0, \\ 2 & \text{if } Y_{s3} = 0. \end{cases}
\]

**Proof.** Estimate (7.1) immediately follows by applying Corollary 6.4 to (6.23) and (6.25). To prove the second estimate let us introduce the pair \( (\tilde{y}_{kh}, \tilde{\theta}_{kh}) \) solving system (6.26), where \( \xi = \xi(q) \) and \( w = w(q) \). Since \( \xi - u_d \in H^1(H_f), \xi - v_d \in H^1(H_s) \) and \( w - w_d \in L^2(H^2(\Omega_s)) \) we can apply Corollary 6.4 to conclude that

\[
\|y(q) - \tilde{y}_{kh}\|_{L^2(L^2(\Omega))} + \|\xi(q) - \tilde{\xi}_{kh}\|_{L^2(L^2(\Omega_{sh}))} + h\|\xi(q) - \tilde{\xi}_{kh}\|_{L^2(H^1(\Omega_{sh}))} \leq C(\tau + \tau_r + h^2).
\]

\( (7.4) \)

On the other hand, the stability estimate in Theorem 6.2 implies

\[
\|\tilde{y}_{kh} - y_{kh}(q)\|_{L^2(L^2(\Omega))} + \|\tilde{\xi}_{kh} - \xi_{kh}(q)\|_{L^2(H^1(\Omega_{sh}))} \leq C\|\xi(q) - \xi_{kh}(q)\|_{L^2(L^2(\Omega))} + CY_{s2}\|w(q) - w_{kh}(q)\|_{L^2(L^2(\Omega_{sh}))} + CY_{s3}\|\nabla w(q) - \nabla w_{kh}(q)\|_{L^2(L^2(\Omega_{sh}))} \leq C(\tau + \tau_r + h^{\kappa(Y_{s3})}).
\]

\( (7.5) \)

Therefore, (7.2) follows from (7.1), (7.4) and (7.5). \( \square \)

7.1 Semidiscrete optimal control problem

For the discretization of the desired states we choose

\[
u_{dkh} = \Pi_k P_{fh} u_d, \quad v_{dkh} = \Pi_k P_{sh} v_d, \quad w_{dkh} = \begin{cases} \Pi_k \tilde{R}_{sh} w_d & \text{if } Y_{s3} > 0, \\ \Pi_k P_{sh} w_d & \text{if } Y_{s3} = 0. \end{cases}
\]

\( (7.6) \)
Let $G_{kh}$ be the discrete analog of $G$ given by

$$G_{kh}(\xi_{kh}, w_{kh}) := \frac{\gamma_f}{2} \int_I \|\xi_{kh} - u_{dkh}\|^2_{\Omega_{kh}} \, dt + \frac{\gamma_s}{2} \int_I \|\xi_{kh} - v_{dkh}\|^2_{\Omega_{kh}} \, dt$$

$$+ \frac{\gamma_s}{2} \int_I \|w_{kh} - w_{dkh}\|^2_{\Omega_{kh}} \, dt + \frac{\gamma_s}{2} \int_I \|\nabla w_{kh} - \nabla w_{dkh}\|^2_{\Omega_{kh}} \, dt.$$ 

Consider the semidiscrete optimal control problem

$$\min_{q \in Q} J_{kh}(\xi_{kh}, w_{kh}, q) = G_{kh}(\xi_{kh}, w_{kh}) + \frac{\alpha}{2} \|q\|^2_Q \quad \text{subject to (6.25)}. \quad (7.7)$$

Take note here that the control has not been discretized yet. The complete discretization of the optimal control problem will be discussed below. Nevertheless, the above problem admits a unique optimal control that we denote by $\bar{q}_{kh} \in Q$. Define the reduced cost functional $j_{kh}: Q \rightarrow \mathbb{R}$ by

$$j_{kh}(q) = J_{kh}(\xi_{kh}(q), w_{kh}(q), q).$$

The derivative of $j_{kh}$ is given by

$$j'_{kh}(q) = (y_{kh}(q) + \alpha q, \delta q)_Q \quad \forall q, \delta q \in Q, \quad (7.8)$$

where $y_{kh}(q) = \sum_{k=1}^{N_k} y_{k-1,h} X_k$ and $(y_{k-1,h})_{k=1}^{N_k}$ is the solution of (6.26) with $\xi, w, u_d, v_d$ and $w_d$ replaced by their discrete counterparts $\xi_{kh}, w_{kh}, u_{dkh}, v_{dkh}$ and $w_{dkh}$, respectively. The proof of (7.8) is analogous to the one given below for the fully discrete optimal control problem. Hence, the details are omitted to avoid repetition.

7.2 Discrete optimal control problem

We now consider the optimal control problem where the control space is also discretized. For the discretization of the control space we take $Q_{kh} = \mathcal{W}_{kh}$ and consider the fully discrete optimal control problem

$$\min_{q_{kh} \in Q_{kh}} J_{kh}(\xi_{kh}, w_{kh}, q_{kh}) = G_{kh}(\xi_{kh}, w_{kh}) + \frac{\alpha}{2} \|q_{kh}\|^2_{Q_{kh}} \quad \text{subject to (6.25) with } q = q_{kh}. \quad (7.9)$$

Denote by $q^*_{kh}$ the optimal solution to this problem. We prove in the following subsection that the derivative of the reduced cost

$$j_{kh}(q_{kh}) = J_{kh}(\xi_{kh}(q_{kh}), w_{kh}(q_{kh}), q_{kh})$$

is given by

$$j'_{kh}(q_{kh}) \delta q_{kh} = (y_{kh}(q_{kh}) + \alpha q_{kh}, \delta q_{kh})_Q \quad \forall q_{kh}, \delta q_{kh} \in Q_{kh}, \quad (7.10)$$

where $y_{kh}(q_{kh})$ is the solution of (6.26) with $\xi, w, u_d, v_d$ and $w_d$ replaced by $\xi_{kh}, w_{kh}, u_{dkh}, v_{dkh}$ and $w_{dkh}$, respectively, and $(\xi_{kh}, w_{kh})$ is the solution of (6.25) with $q = q_{kh}$.
Remark 7.2. From the choice of the discretization of the control space it follows that the optimal controls of (7.7) and (7.9) are related according to $q_{kh}^* = \Pi_k P_h \tilde{q}_{kh}$.

In the following we have the Lipschitz estimates for the derivatives of $j$ and $j_{kh}$. The proofs are similar to the one given in Meidner & Vexler (2008) and thus omitted.

Lemma 7.3 There exists a constant $C > 0$ such that for every $q, \tilde{q}, \delta q \in Q$ we have

$$|j'(q)\delta q - j_{kh}'(q)\delta q| \leq \|y(q) - y_{kh}(q)\|_Q \|\delta q\|_Q,$$

$$|j_{kh}'(q)\delta q - j_{kh}'(\tilde{q})\delta q| \leq C\|q - \tilde{q}\|_Q \|\delta q\|_Q.$$

Now, we state and prove the main result of this section.

Theorem 7.4 Let $q^*$ and $q_{kh}^*$ be the respective solutions of the continuous and discrete optimal control problems (3.11) and (7.9). Then there exists a constant $C > 0$ independent of $h$, $\tau_r$ and $\tau$ such that

$$\|q^* - q_{kh}^*\|_Q \leq C(\tau + \tau_r + h^\kappa(\gamma_3)), \tag{7.11}$$

where $\kappa(\gamma_3)$ is given by (7.3). Moreover, if $(\xi^*, w^*)$ and $(\xi_{kh}^*, w_{kh}^*)$ are the corresponding states and $(y^*, \vartheta^*)$ and $(y_{kh}^*, \vartheta_{kh}^*)$ are the adjoint states then

$$\|\xi^* - \xi_{kh}^*\|_{L^2(L^2(\Omega))} + \|w^* - w_{kh}^*\|_{L^2(L^2(\Omega_d))} + h\|w^* - w_{kh}^*\|_{L^2(H^1(\Omega_d))} \leq C(\tau + \tau_r + h^\kappa(\gamma_3)), \tag{7.12}$$

$$\|y^* - y_{kh}^*\|_{L^2(L^2(\Omega))} + \|\vartheta^* - \vartheta_{kh}^*\|_{L^2(L^2(\Omega_d))} + h\|\vartheta^* - \vartheta_{kh}^*\|_{L^2(H^1(\Omega_d))} \leq C(\tau + \tau_r + h^\kappa(\gamma_3)). \tag{7.13}$$

Proof. Let $\tilde{q}_{kh}^* = \Pi_k P_h q^*$. Recall that $q_{kh}^*$, $\tilde{q}_{kh}$ and $q^*$ are the solutions of fully discrete (7.9), semidiscrete (7.7) and continuous (3.11) optimal control problems, respectively. By optimality we have

$$j'_{kh}(\tilde{q}_{kh}) (\tilde{q}_{kh} - q_{kh}^*) = j'_{kh}(q_{kh}) (\tilde{q}_{kh} - q_{kh}^*) = j' (q^*) (\tilde{q}_{kh} - q_{kh}^*) = 0. \tag{7.14}$$

According to the linear–quadratic nature of the optimal control problems we have

$$j''_{kh}(q)(\delta q, \delta p) = \gamma_f(\xi_{kh}(\delta q), \xi_{kh}(\delta p))L^2(L^2(\Omega_d)) + \gamma_{s1}(\xi_{kh}(\delta q), \xi_{kh}(\delta p))L^2(L^2(\Omega_d))$$

$$+ \gamma_{s2}(w_{kh}(\delta q), w_{kh}(\delta p))L^2(L^2(\Omega_d)) + \gamma_{s3}(\vartheta_{kh}(\delta q), \vartheta_{kh}(\delta p))L^2(L^2(\Omega_d))$$

$$+ \alpha(\delta q, \delta p) Q$$

for every $q, \delta q, \delta p \in Q$, and in particular $j''_{kh}(q)$ is independent of $q$. Thus, from (7.14),

$$\alpha \|\tilde{q}_{kh}^* - q_{kh}^*\|_{Q}^2 \leq j''_{kh}(q_{kh}^*) (\tilde{q}_{kh} - q_{kh}^*) (\tilde{q}_{kh} - q_{kh}^*)$$

$$= j'_{kh}(\tilde{q}_{kh}) (\tilde{q}_{kh} - q_{kh}^*) - j'_{kh}(q_{kh}) (\tilde{q}_{kh} - q_{kh}^*)$$

$$= j'_{kh}(\tilde{q}_{kh}) (\tilde{q}_{kh} - q_{kh}^*) - j'_{kh}(q^*) (\tilde{q}_{kh} - q_{kh}^*) + j_{kh}(q^*) (\tilde{q}_{kh} - q_{kh}^*) - j' (q^*) (\tilde{q}_{kh} - q_{kh}^*)$$
and therefore, from Lemma 7.3, we have the estimate

\[ \| \tilde{q}_{kh} - q_{kh}^* \|_Q \leq C_{\alpha} \{ \| \tilde{q}_{kh}^* - q_{kh}^* \|_Q + \| y(q^*) - y_{kh}(q^*) \|_Q \}. \]

Consequently, from the triangle inequality, we obtain

\[ \| q^* - q_{kh}^* \|_Q \leq C \{ \| \tilde{q}_{kh}^* - q_{kh}^* \|_Q + \| y(q^*) - y_{kh}(q^*) \|_Q \}. \] (7.15)

Applying interpolation error estimates the regularity of the optimal control \( q^* \in H^1(L^2(\Omega)) \cap L^2(H^1_0(\Omega)) \cap L^2(\Omega) \) of the continuous problem (see Corollary 3.8), the uniform boundedness of \( \Pi_k \) and (6.1) we have

\[ \| \tilde{q}_{kh}^* - q_{kh}^* \|_Q \leq \| \Pi_k \| \| P_h q^* - q^* \|_Q + \| \Pi_k q_{kh}^* - q_{kh}^* \|_Q \leq C(h^2 + \tau). \] (7.16)

From (7.15), (7.16) and Lemma 7.1 we deduce the error estimate (7.11). The error estimate (7.12) can be derived by writing \( \xi^* - \xi_{kh}^* = \left( \xi(q^*) - \xi_{kh}(q^*) \right) + \left( \xi_{kh}(q^*) - \xi_{kh}(q_{kh}^*) \right) \) and \( w^* - w_{kh}^* = \left( w(q^*) - w_{kh}(q^*) \right) + \left( w_{kh}(q^*) - w_{kh}(q_{kh}^*) \right) \), applying Lemma 7.1, the stability estimate in Theorem 6.2 and (7.11). Analogous decomposition can be done for the adjoint state to obtain (7.13).

**Remark 7.5** Instead of (7.6) one may choose interpolation for the approximation of the desired states, that is,

\[ u_{dkh} = \Pi_{k_j} u_d, \quad v_{dkh} = \Pi_{k_j} v_d, \quad w_{dkh} = \Pi_{k_j} w_d. \] (7.17)

The order \( O(\tau + \tau_r + h^{\epsilon/2}) \) is preserved using interpolation error estimates and Theorem 6.2.

### 7.3 Numerical solution

We prove (7.10) by rewriting (7.9) in algebraic form. This will be also useful in setting up the linear system for the implementation of the numerical scheme. Consider a triangulation \( T_h \) in \( \{ T_h \}_{h > 0} \) discussed in Section 5. Let \( \{ x_{h,l} \}_{l=1}^{n_{sh}} \), \( \{ x_{h,l} \}_{l=n_{sh}+1}^{m_{sh}} \), \( \{ x_{h,l} \}_{l=m_{sh}+1}^{n_{fh}} \) and \( \{ x_{h,l} \}_{l=n_{fh}+1}^{m_{fh}} \) be the interior nodes of \( T_h \) in \( \Omega_{sh} \), the nodes on the discretized interface \( \Gamma_{sh} \), the interior nodes in \( \Omega_{fh} \) together with the nodes on the boundary \( \Gamma_j \) and the barycenters of the triangles in \( \Omega_{fh} \), respectively. Let \( \psi_{h,l} \) for \( 1 \leq l \leq n_{fh} \) be the piecewise linear function in \( \Omega \) and \( \psi_{h,l} \) for \( n_{fh} + 1 \leq l \leq m_{fh} \) be the bubble function in \( \Omega \) such that \( \psi_{h,l}(x_{h,j}) = \delta_{lj} \) for \( 1 \leq j, l \leq m_{fh} \).
The nodal bases for the finite element spaces \( W_h, W_{sh} \) and \( M_h \) are given by the scalar-valued basis functions \( \{ \varphi_{h,l} \}_{l=1}^{m_h} \), \( \{ \varphi_{sh,l} \}_{l=m_{sh}+1}^{m_{sh}+n_{sh}} \) and \( \{ \varphi_{h,l} \}_{l=m_{sh}+1}^{m_{sh}} \) with the underlying fields \( \mathbb{R}^2, \mathbb{R}^2 \) and \( \mathbb{R} \), respectively. The approximate solutions of (6.25) and the control \( q = q_{kh} \) can be expressed as

\[
\xi_{kh} = \sum_{k=0}^{N_h} \sum_{l=1}^{m_h} \xi_{k,h,l} \chi_{k,h}\varphi_{h,l}, \quad \psi_{kh} = \sum_{k=0}^{N_h} \sum_{l=1}^{m_h} \psi_{k,h,l} \chi_{k,h}\varphi_{h,l}, \quad q_{kh} = \sum_{k=1}^{N_h} \sum_{l=1}^{m_h} q_{k,h,l} \chi_{k,h}\varphi_{h,l}, \quad p_{kh} = \sum_{k=1}^{N_h} \sum_{l=1}^{n_{sh}} p_{k,h,l} \chi_{k,h}\varphi_{h,l},
\]

for some \( \xi_{k,h,l}, \psi_{k,h,l}, q_{k,h,l}, p_{k,h,l} \in \mathbb{R}^2 \) and \( p_{k,h,l} \in \mathbb{R} \). Let \( \xi_{k,h} = (\xi_{k,h,l})_{l=1}^{m_h} \in \mathbb{R}^{2m_h} \), where we arrange the vectors in such a way that the first components of \( \xi_{k,h,l} \) for \( 1 \leq l \leq m_h \) are located on the first half of \( \xi_{kh} \) and the second components on the second half. We shall use the same notation \( \xi_{kh} \) for the vector \( (\xi_{k,h})_{k=1}^{N_h} \in \mathbb{R}^{2N_hm_h} \). Similar notation will be utilized for the other variables \( w_{kh} \in \mathbb{R}^{2N_hm_{sh}}, p_{kh} \in \mathbb{R}^{2N_hm_{sh}}, q_{kh} \in \mathbb{R}^{2N_hm_{sh}} \) and the discretized desired states \( u_{dkh} \in \mathbb{R}^{2N_h(s_{m_{sh}}-m_{sh})}, v_{dkh} \in \mathbb{R}^{2N_hm_{sh}} \) and \( w_{dkh} \in \mathbb{R}^{2N_hm_{sh}} \). We set the functions \( u_{dkh} \) and \( v_{dkh} \) on the nodes outside \( \Omega_{sh} \) and \( \Omega_{fh} \), respectively, to zero so that \( u_{dkh}, v_{dkh} \in \mathbb{R}^{2N_hm_{sh}} \).

Consider the following mass and stiffness matrices for the fluid and structure:

\[
(\tilde{M}_{sh})_{ij} = (\varphi_{h,i}^* \varphi_{h,j})_{\Omega_{sh}}, \quad (\tilde{A}_{sh})_{ij} = (\nabla \varphi_{h,i}^* \nabla \varphi_{h,j})_{\Omega_{sh}}, \\
(\tilde{M}_{fsh})_{ij} = (\varphi_{sh,i}^* \varphi_{sh,j})_{\Omega_{sh}}, \quad (\tilde{A}_{sh})_{ij} = (\nabla \varphi_{sh,i}^* \nabla \varphi_{sh,j})_{\Omega_{sh}} + \varepsilon (\nabla \varphi_{sh,i}^* \nabla \varphi_{sh,j})_{\Omega_{sh}}, \\
(\tilde{M}_{fh})_{ij} = (\varphi_{h,i}^* \varphi_{h,j})_{\Omega_{fh}}, \quad (B_{sh})_{ik} = -\partial_j \varphi_{h,i}^* \varphi_{h,k} \Omega_{fh}, \quad (B_{yh})_{ik} = -\partial_j \varphi_{h,i}^* \varphi_{h,k} \Omega_{fh},
\]

for \( 1 \leq i, j \leq m_h \) and \( m_{sh} + 1 \leq k \leq n_{sh} \). Let

\[
B_h = [B_{sh}, B_{yh}]^T, \quad \tilde{C}_h = [\gamma_2 (M_{sh})_{ij} + \gamma_3 (A_{sh})_{ij}]_{1 \leq i, j \leq m_{sh}}, \\
\tilde{D}_h = [(M_{sh})_{ij} + (A_{sh})_{ij}]_{1 \leq i, j \leq m_{sh}}, \quad \tilde{D}_{h1} = [(M_{sh} + A_{sh})_{ij}]_{1 \leq i \leq m_{sh}, 1 \leq j \leq m_{sh}}.
\]

Furthermore, we define the matrices

\[
A_{sh} = \tilde{A}_{sh} \otimes I_2, \quad M_{sh} = \tilde{M}_{sh} \otimes I_2, \quad C_h = \tilde{C}_h \otimes I_2, \\
M_{fh} = \tilde{M}_{fh} \otimes I_2, \quad M_{fsh} = \tilde{M}_{fsh} \otimes I_2, \quad D_h = \tilde{D}_h \otimes I_2, \quad D_{h1} = \tilde{D}_{h1} \otimes I_2,
\]

where \( \otimes \) is the Kronecker tensor product and \( I_2 \) is the 2 \times 2 identity matrix. Observe that these matrices are symmetric except for \( D_{h1} \).
The discrete optimal control problem (7.9) can now be rewritten equivalently as

$$\begin{align*}
\min_{q_{kh} \in \mathbb{R}^{2N_{\tau}m_{fh}}} J_{kh}(\xi_{k,h}, w_{kh}, q_{kh}) &= \frac{1}{2} \sum_{k=1}^{N_{\tau}} \{y_f(\xi_{k,h} - u_{dk,h})^T M_{fh}(\xi_{k,h} - u_{dk,h}) \\
&+ \gamma_s(\xi_{k,h} - v_{dk,h})^T M_{sh}(\xi_{k,h} - v_{dk,h}) + (w_{k,h} - w_{dk,h})^T C_h(w_{k,h} - w_{dk,h})\} \\
&+ \frac{\alpha}{2} \sum_{k=1}^{N_{\tau}} q_{kh}^T M_{sh} q_{kh}
\end{align*}$$

subject to the following linear discrete time delay system:

$$\begin{bmatrix}
(\frac{1}{\tau} M_{sh} + A_{eh}) & D_{h1} & B_h \\
D_{h1}^T & -\frac{1}{\tau} D_h & O \\
B_h^T & O & O
\end{bmatrix}
\begin{bmatrix}
\xi_{k,h} \\
w_{k,h} \\
p_{k,h}
\end{bmatrix}
= \begin{bmatrix}
-\mu M_{sh} \xi_{k-1,n_{n_{rk}},h} + \frac{1}{\tau} M_{sh} \xi_{k-1,h} + M_{sh} q_{kh} \\
-\frac{1}{\tau} D_h w_{k-1,h} \\
0
\end{bmatrix},
$$

\quad k = 1, \ldots, N_{\tau},

with initial data $\xi_{0,h}, w_{0,h}$ and initial history $\xi_{j,h}$ for $j = -n_{n_{rk}}, \ldots, -1$. The derivative of the reduced cost $j_{kh}(q_{kh}) = J_{kh}(\xi_{k,h}(q_{kh}), w_{kh}(q_{kh}), q_{kh})$ in the direction $\delta q_{kh} \in \mathbb{R}^{2N_{\tau}m_{fh}}$ is given by

$$j'_{kh}(q_{kh}) \delta q_{kh} = \frac{\tau}{2} \sum_{k=1}^{N_{\tau}} \{y_f(\delta \xi_{k,h}^T M_{fh}(\delta \xi_{k,h} - u_{dk,h}) + \gamma_s(\delta \xi_{k,h}^T M_{sh}(\delta \xi_{k,h} - v_{dk,h}) \\
+ \delta w_{k,h}^T C_h(\delta w_{k,h} - w_{dk,h}) + \alpha \delta q_{k,h}^T M_{sh} q_{kh}\}.
$$

where $(\delta \xi_{kh}, \delta w_{kh}, \delta p_{kh})$ is the solution of (7.19) with $q_{kh}$ replaced by $\delta q_{kh}$. We will show that this is equivalent to

$$j'_{kh}(q_{kh}) \delta q_{kh} = \frac{\tau}{2} \sum_{k=1}^{N_{\tau}} \{\delta q_{k,h}^T M_{sh} y_{k-1,h}(q_{kh}) + \alpha \delta q_{k,h}^T M_{sh} q_{kh}\},
$$

where $y_{k-1,h} = y_{k-1,h}(q_{kh})$ for $k = N_{\tau}, \ldots, 1$ is the solution of

$$\begin{bmatrix}
(\frac{1}{\tau} M_{sh} + A_{eh}) & D_{h1} & B_h \\
D_{h1}^T & -\frac{1}{\tau} D_h & O \\
B_h^T & O & O
\end{bmatrix}
\begin{bmatrix}
y_{k-1,h} \\
y_{k-1,h} \\
y_{k-1,h}
\end{bmatrix}
= \begin{bmatrix}
-\mu M_{sh} y_{k-n_{n_{rk}},h} + \frac{1}{\tau} M_{sh} y_{k,h} + y_f(\xi_{k,h} - u_{dk,h}) + \gamma_s M_{sh}(\xi_{k,h} - v_{dk,h}) \\
-\frac{1}{\tau} D_h y_{k,h} + C_h(w_{k,h} - w_{dk,h}) \\
0
\end{bmatrix},
$$

with homogeneous terminal data and dual history

$$\vartheta_{N_{\tau},h} = 0, \quad y_{j,h} = 0 \quad \text{for} \quad j = N_{\tau}, \ldots, N_{\tau} + n_{n_{rk}}.
$$
We then prove that (7.21) is equivalent to (7.10). For this purpose we will abbreviate the matrix on the left-hand side of (7.22) by $A_h$. Let $\delta q_{kh} \in \mathbb{R}^{2N_r m_h}$ and $(\delta \xi_{k,h}, \delta w_{k,h}, \delta p_{k,h})$ be the solution of (7.19) with control $\delta q_{kh}$. Without loss of generality, assume that $\delta w_{0,h} = 0$ and $\delta \xi_{j,h} = 0$ for $j = -n_r, n_r, \ldots, 0$. Using this and reindexing we have

$$\sum_{k=1}^{N_r} \tau [y_{k-1,h}, \vartheta_{k-1,h}, \pi_{k-1,h}] A_h [\delta \xi_{k,h}, \delta w_{k,h}, \delta p_{k,h}]^T$$

$$= \sum_{k=1}^{N_r} \tau \left\{ -\mu y_{k-1,h}^T M_{sh} \delta \xi_{k-1-n_r n_r,h} + \frac{1}{\tau} y_{k-1,h}^T M_{fsh} \delta \xi_{k-1,h} \\
+ y_{k-1,h}^T M_{fsh} \delta q_{k,h} - \frac{1}{\tau} \vartheta_{k-1,h}^T D_h \delta w_{k-1,h} \right\}$$

$$= \sum_{k=1}^{N_r} \tau \left\{ -\mu y_{k-1,n_r n_r,h}^T M_{sh} \delta \xi_{k,h} + \frac{1}{\tau} y_{k,h}^T M_{fsh} \delta \xi_{k,h} \\
+ y_{k-1,h}^T M_{fsh} \delta q_{k,h} - \frac{1}{\tau} \vartheta_{k,h}^T D_h \delta w_{k,h} \right\}. \quad (7.24)$$

Applying the symmetry of $A_h$ and using (7.22), this is equal to

$$\sum_{k=1}^{N_r} \tau [\delta \xi_{k,h}, \delta w_{k,h}, \delta p_{k,h}] A_h [y_{k-1,h}, \vartheta_{k-1,h}, \pi_{k-1,h}]^T$$

$$= \sum_{k=1}^{N_r} \tau \left\{ -\mu \delta \xi_{k,h}^T M_{sh} y_{k+n_r n_r,h} + \frac{1}{\tau} \delta \xi_{k,h}^T M_{fsh} y_{k,h} + \frac{1}{\tau} \gamma_f \delta \xi_{k,h}^T M_{fh} (\xi_{k,h} - u_{dk,h}) \\
+ y_{k,h}^T M_{fsh} (\xi_{k,h} - v_{dk,h}) - \frac{1}{\tau} \delta w_{k,h}^T D_h \vartheta_{k,h} + \delta w_{k,h}^T C_h (w_{k,h} - w_{dk,h}) \right\}. \quad (7.25)$$

Comparing (7.24) and (7.25), using (7.20) and the symmetry of $M_{sh}, M_{fsh}$ and $D_h$ we deduce (7.21), and therefore (7.10).

System (7.22) and (7.23) is the dG(0)–cG(1) discretization of the adjoint equation for the continuous problem. Hence, for the proposed numerical scheme the two strategies discretize-then-optimize and optimize-then-discretize coincide. In other words, the discretization schemes obtained from the optimality system of the discretized problem and the one from the discretization of the optimality system for the continuous problem are the same.

The discretized state and adjoint equations will be solved by adding an artificial compressibility, that is, the matrix on the left-hand side in (7.19) and (7.22) will be replaced by

$$\begin{bmatrix}
\left( \frac{1}{\tau} M_{fsh} + A_{eh} \right) & D_{h1} & B_h \\
D_{h1}^T & -\frac{1}{\tau} D_h & O \\
-B_h^T & O & \eta I_{n_{yh} = m_h}
\end{bmatrix}$$
for small enough $\eta > 0$. Alternatively, one can replace the identity matrix $I_{nh\rightarrow msh}$ by the mass matrix associated with the finite element space $M_h$. The error between the original solution and the one obtained by this penalization is of order $O(\eta)$. We refer to Girault & Raviart (1986, Section I.4.3) and Ern & Guermond (2004, Section 4.4.4) for more details. The linear system will be reduced by eliminating the discrete pressure. To do this, we introduce $D_{h}\xi = M_{sh} + A_{sh}$ and $\xi_{k,h} = (\xi_{k,h,1})_{l=1}^{N\bar{r}}$. Then by straightforward algebra the discretized state equation can be reduced to the following system for $k = 1, \ldots, N_{r}$:

$$
\begin{align*}
\left[ \left( \frac{1}{\tau} M_{sh} + A_{sh} + \tau D_{h}\right) + \frac{1}{\eta} BB^T \right] \xi_{k,h} \\
&= -\mu M_{sh}\xi_{k-1-n_{x},h} + \frac{1}{\tau} M_{sh}\xi_{k-1,h} + M_{sh}q_{k,h} - D_{h}w_{k-1,h}, \\
w_{k,h} &= w_{k-1,h} + \tau \xi_{k,h}.
\end{align*}
$$

(7.26)

Therefore, at the $k$th time step, we can solve for $\xi_{k,h}$ first and then for $w_{k,h}$. Note that the matrix in the linear system associated to $\xi_{k,h}$ is symmetric and positive definite. Thus, Cholesky factorization or conjugate gradient methods are applicable to solve the first system in (7.26).

Similarly, the discretized adjoint equation can be reduced to the system for $k = N_{r}, \ldots, 1$,

$$
\begin{align*}
\left[ \left( \frac{1}{\tau} M_{sh} + A_{sh} + \tau D_{h}\right) + \frac{1}{\eta} BB^T \right] y_{1-h} &= -\mu M_{sh}y_{k+n_{x}n_{x}+1} + \frac{1}{\tau} M_{sh}y_{k,h} \\
&+ \gamma_{f} M_{sh}(\xi_{k-1,h} - u_{d,k,h}) + \gamma_{s1} M_{sh}(\xi_{k,h} - v_{d,k,h}) - D_{h}\partial_{k,h} + \tau C_{h}(w_{k,h} - w_{d,k,h}), \\
\partial_{k-1,h} &= \partial_{k,h} + \tau y'_{k-1,h} - \tau D_{h}C_{h}(w_{k,h} - w_{d,k,h}),
\end{align*}
$$

(7.27)

where $y'_{k-1,h} = (y'_{k-1,h,l})_{l=1}^{N\bar{r}}$. If $\gamma_{s2} = \gamma_{s3}$ then $C_{h} = \gamma_{s2} D_{h}$ and the second equation in (7.27) simplifies to $\partial_{k-1,h} = \partial_{k,h} + \tau y'_{k-1,h} - \tau \gamma_{s3}(w_{k,h} - w_{d,k,h})$, which is analogous to the second equation in (7.26). In the case where $\gamma_{s2} \neq \gamma_{s3}$, the second equation in (7.27) can be solved by writing it as the system

$$
\begin{align*}
D_{h}\lambda_{k-1,h} &= C_{h}(w_{k,h} - w_{d,k,h}), \\
\partial_{k-1,h} &= \partial_{k,h} + \tau y'_{k-1,h} - \tau \lambda_{k-1,h}.
\end{align*}
$$

8. Numerical examples

In this section we present numerical examples illustrating the theoretical results of the paper.

8.1 Example 1

For the FSI domain we consider the unit square $\Omega = (0, 1)^2$, and for the structural domain $\Omega_{s}$ we take the ball centered at $(0.3, 0.6)$ with radius 0.2025. The parameters are $T = 2$, $\mu = 2$, $\varepsilon = 0.1$, $r = 1$, $\gamma_{f} = \gamma_{s1} = \gamma_{s2} = 1$, $\gamma_{s3} = 0.01$ and $\alpha = 10^{-6}$. We consider a quasi-uniform mesh refined at the interface having a mesh size $h = 0.0671$ with 1871 nodes, 2800 triangles in the fluid domain and 820 triangles in the structure domain. The step sizes for the time and history grids are $\tau = \tau_{r} = 0.0025$. The total number of unknowns (primal, dual and control variables excluding the pressure) for the control problems acting in the whole FSI domain, in the fluid domain and in the structure domain are of orders $2.38 \cdot 10^{7}$, $2.32 \cdot 10^{7}$ and $1.7 \cdot 10^{6}$, respectively. Mass and stiffness matrices for the fluid velocity...
and pressure are assembled with the help of the formulas given in Koko (2012) and they are stored in compressed sparse column format.

We choose the following target states:

\[ u_d(t, x, y) = \cos(\pi t)(\phi_1(x, y), \phi_2(x, y))^T, \]
\[ v_d(t, x, y) = \cos(\pi t)(\phi_1(x, y)\rho(x, y), \phi_2(x, y)\rho(x, y))^T, \]
\[ w_d(t, x, y) = \pi^{-1}(1 + \sin(\pi t))(\phi_1(x, y)\rho(x, y), \phi_2(x, y)\rho(x, y))^T, \]

where

\[ \phi_1(x, y) = (1 - \cos(2\pi x))\sin(2\pi y), \]
\[ \phi_2(x, y) = \sin(2\pi x)(\cos(2\pi y) - 1), \]
\[ \rho(x, y) = 0.2025^{-2}(x - 0.3)^2 + (y - 0.6)^2, \]

and for the initial data and history we take \( u_0(x) = u_d(0, x), v_0(x) = v_d(0, x), w_0(x) = w_d(0, x) \) and \( z_0(\theta, x) = v_d(\theta, x) \). The initial data are discretized through nodal interpolation, and the time average integral for the initial history is approximated using the trapezoidal rule.
For numerical optimization we use the Barzilai–Borwein version of the gradient method in Barzilai & Borwein (1988) with an alternating steplength selection method and terminate the routine once

Fig. 3. Components $q_1$ (left) and $q_2$ (right) of the optimal controls $q = (q_1, q_2)^T$ acting in the domain $\Omega$ (first row), $\Omega_f$ (second row) and $\Omega_s$ (third row) at time $t = 2$. 
Table 1  

<table>
<thead>
<tr>
<th></th>
<th>Control in Ω</th>
<th>Control in Ω_f</th>
<th>Control in Ω_s</th>
</tr>
</thead>
<tbody>
<tr>
<td># BB iterations</td>
<td>218</td>
<td>72</td>
<td>231</td>
</tr>
<tr>
<td>Cost value</td>
<td>3.40559 · 10⁻³</td>
<td>4.35328 · 10⁻²</td>
<td>6.75255 · 10⁻²</td>
</tr>
<tr>
<td>Optimality residual</td>
<td>2.44245 · 10⁻⁶</td>
<td>5.21040 · 10⁻⁶</td>
<td>2.83850 · 10⁻⁵</td>
</tr>
</tbody>
</table>

The relative error of two successive cost function values is less than the tolerance 10⁻⁶. The second iterate of the gradient method is computed using an inexact line search with Armijo’s rule as a step-length selection criterion. The algorithm is implemented in Python 3.6.4 (Python Software Foundation, https://www.python.org/) on a 2.5 GHz Intel Core i5 with 4 GB RAM. Solutions of the linear systems for each time step of every primal and dual solve is computed using the function splu with the SuperLU option (Li et al., 1999) in the package SciPy. An LU factorization was computed beforehand and a column permutation for sparsity preservation via minimum degree ordering was utilized. The bulk of the computational time for the gradient algorithm lies in the forward and backward solve for the discrete primal and adjoint equations.

Figure 2 illustrates the time evolution of the norms for the residuals of the states to the desired states using controls acting in the entire FSI domain, in the fluid domain only or in the structure domain only. The components of the computed optimal controls at the terminal time $t = 2$ are given in Fig. 3. For controls acting either in the fluid or structure only we observe huge effort near the interface. This means that we need large amplitudes near the interface to control the fluid velocity if the control is acting only in the structure domain and similarly to control the structure displacement, stress and velocity if the control is acting only in the fluid domain.

As expected, the value of the optimal cost is smallest if the control acts on all of the domain, rather than on Ω_f and Ω_s only; see Table 1. Also, the spatial amplitudes of the control are quite large, which is consistent with the choice of a small value for $\alpha$, which means the controls are cheap. Finally, we observe that the oscillatory (in space) and periodic (in time) nature of the desired states $(u_d, v_d, w_d)$ are reflected in the nature of the optimal controls.

8.2 Example 2

We investigate the effect of the delay parameter $r$ on the optimal controls. We use the setup of the previous example and denote by $q_a$ and $q_b$ the computed optimal controls corresponding to $r = 0.2$ and $r = 1$, respectively. For $\alpha = 10^{-3}$ we can observe in Fig. 4 that the difference occurs mainly in the structural domain, which is reasonable because delay appears only in this part of the domain. The situation is also quite similar with smaller regularization $\alpha = 10^{-6}$, where the majority of the difference occurs in the solid domain.

8.3 Example 3

In this example we study the convergence rates of the optimal control and the corresponding primal and adjoint states. For the setup we take $T = 0.4$, $r = \mu = \varepsilon = 0.1$, $\gamma_f = \gamma_{s1} = \gamma_{s2} = 1$, $\gamma_{s3} = 0.001$ and $\alpha = 0.1$. We consider the same physical configuration as in Example 1 with control acting in the entire FSI domain.
Fig. 4. Difference of the optimal controls for the system with delay \( r = 0.2 \) and \( r = 1 \) and regularization \( \alpha = 10^{-3} \) (top) and \( \alpha = 10^{-6} \) (bottom).

In the absence of an explicit solution we proceed as follows: we define

\[
\begin{align*}
    u(t, x, y) &= \cos(\pi t)(\phi_1(x, y), \phi_2(x, y))^T, \\
    w(t, x, y) &= \pi^{-1}(1 + \sin(\pi t))(\phi_1(x, y)\rho(x, y), \phi_2(x, y)\rho(x, y))^T, \\
    p(t, x, y) &= 2\pi \sin(\pi t)(\cos(2\pi y) - \cos(2\pi x)),
\end{align*}
\]
where \( \phi_1, \phi_2 \) and \( \rho \) are the functions defined in Example 1.

We would like \((u, w, p)\) to be the solution of (1.4). For this purpose we add appropriate source terms on the right-hand side of (1.4). With state variables \((u, w, p)\) and desired states \((u_d, v_d, w_d) = -(u, w, w)\) we compute numerically the adjoint state \((y, \vartheta)\) using the scheme (6.26) and then use the equation \( q = -\frac{1}{\alpha}y \) to be the optimal control.

We use bisection for the mesh refinement, that is, midpoints of the edges are used as new nodes in the refined mesh. Moreover, to have a better approximation of the curved interface, each midpoint of an edge that is located on the discretized interface is projected onto \( \Gamma_s \). Up to four successive grid refinement this ensures a quadratic order reduction rate for the distance between \( \Gamma_s \) and its discretization \( \Gamma_{sh} \); see (5.1). In Figs 5 and 6 we observe an approximate order \( \mathcal{O}(h^2) \) and \( \mathcal{O}(\tau) \) (where \( \tau = \tau_i \)) with respect to spatial and temporal discretization errors, respectively, which agrees with the theoretical results presented in the previous section.

**Funding**

European Union’s H2020 research program ERC advanced grant (668998, OCLOC); Austrian Agency for International Cooperation in Education and Research (Ernst Mach grant).

**Acknowledgements**

The authors are grateful to the referees for their careful reading and valuable suggestions.

**References**


Datko, R. (1988) Not all feedback stabilized hyperbolic systems are robust with respect to small time delays in their feedback. SIAM J. Control Optim., 26, 697–713.


