SIAM J. CONTROL OPTIM.
Vol. 47, No. 3, pp. 1150–1177

A PRIORI ERROR ESTIMATES FOR SPACE-TIME FINITE ELEMENT DISCRETIZATION OF PARABOLIC OPTIMAL CONTROL PROBLEMS PART I: PROBLEMS WITHOUT CONTROL CONSTRAINTS*

DOMINIK MEIDNER† AND BORIS VEXLER‡

Abstract. In this paper we develop a priori error analysis for Galerkin finite element discretizations of optimal control problems governed by linear parabolic equations. The space discretization of the state variable is done using usual conforming finite elements, whereas the time discretization is based on discontinuous Galerkin methods. For different types of control discretizations we provide error estimates of optimal order with respect to both space and time discretization parameters. The paper is divided into two parts. In the first part we develop some stability and error estimates for space-time discretization of the state equation and provide error estimates for optimal control problems without control constraints. In the second part of the paper, the techniques and results of the first part are used to develop a priori error analysis for optimal control problems with pointwise inequality constraints on the control variable.

Key words. optimal control, parabolic equations, error estimates, finite elements

AMS subject classifications. 49N10, 49M25, 65M15, 65M60

DOI. 10.1137/070694016

1. Introduction.

In this paper we develop a priori error analysis for space-time finite element discretizations of parabolic optimization problems. We consider the following linear-quadratic optimal control problem for the state variable u and the control variable q:

$$\text{Minimize } J(q,u) = \frac{1}{2}\int_0^T\int_\Omega (u(t,x)-\hat{u}(t,x))^2\,dx\,dt + \frac{\alpha}{2}\int_0^T\int_\Omega q(t,x)^2\,dx\,dt \tag{1.1a}$$

subject to

$$\begin{aligned} \partial_t u - \Delta u &= f + q \qquad \text{in } (0,T)\times\Omega, \\ u(0) &= u_0 \qquad \text{in } \Omega, \end{aligned} \tag{1.1b}$$

combined with either homogeneous Dirichlet or homogeneous Neumann boundary conditions on $(0,T)\times\partial\Omega$. A precise formulation of this problem including a functional analytic setting is given in the next section.

While the a priori error analysis for finite element discretization of optimal control problems governed by elliptic equations is discussed in many publications (see, e.g., [12, 15, 1, 16, 22, 4]), there are only a few published results on this topic for parabolic problems; see [20, 28, 17, 19, 24].

*Received by the editors June 8, 2007; accepted for publication (in revised form) November 27, 2007; published electronically March 19, 2008. The first author was supported by the German research Foundation DFG through the International Research Training Group 710 “Complex Processes: Modeling, Simulation, and Optimization.” The second author was partially supported by the Austrian Science Fund FWF project P18971-N18 “Numerical analysis and discretization strategies for optimal control problems with singularities.”
http://www.siam.org/journals/sicon/47-3/69401.html

†Institut für Angewandte Mathematik, Ruprecht-Karls-Universität Heidelberg, INF 294, 69120 Heidelberg, Germany (dominik.meidner@iwr.uni-heidelberg.de).
‡Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenberger Straße 69, 4040 Linz, Austria (boris.vexler@oeaw.ac.at).

In this paper, we will use discontinuous finite element methods for time discretization of the state equation (1.1b), as proposed, e.g., in [7, 10]. The spatial discretization will be based on usual H^1-conforming finite elements. In [2] this type of discretization is shown to allow for a natural translation of the optimality conditions from the continuous to the discrete level. This gives rise to exact computation of the derivatives required in the optimization algorithms on the discrete level. In [21] a posteriori error estimates for this type of discretization are derived and an adaptive algorithm is developed.

Throughout, we will use a general discretization parameter σ consisting of three discretization parameters $\sigma = (k, h, d)$, where k corresponds to the time discretization of the state variable, h to the space discretization of the state variable, and d to the discretization of the control variable q, respectively. The space and time discretizations of the control variable may differ from the discretizations of the state. Therefore, the discretization parameter d consists of the discretization parameters k_d and h_d for the time and space discretizations of the control variable. In this paper we will derive a priori error estimates of optimal order with respect to all discretization parameters, where the influences of different parts of the discretization are clearly separated. Moreover, the temporal and spatial regularity properties of the solution to the continuous problem (1.1) are separated as well.

For the discretization error between the solution of the continuous optimization problem $(\bar{q}, \bar{u})$ and the solution of the discretized problem $(\bar{q}_\sigma, \bar{u}_\sigma)$, we will prove error estimates of the following structure:

$$(1.2)\quad \|\bar{q}-\bar{q}_\sigma\|_{L^2((0,T)\times\Omega)} \le C_1(\bar{u},\bar{z})\, k^{r+1} + C_2(\bar{u},\bar{z})\, h^{s+1} + C_3(\bar{q})\, k_d^{r_d+1} + C_4(\bar{q})\, h_d^{s_d+1}\,,$$

where r, r_d are the highest degrees of the polynomials in the time discretization of the state and the control variable, respectively, and s, s_d are the highest degree of the polynomials in the space discretization of the control and the state variable. The constants $C_1(\bar{u}, \bar{z})$ and $C_2(\bar{u}, \bar{z})$ depend on the temporal and the spatial regularity of the optimal state $\bar{u}$ and the corresponding adjoint state $\bar{z}$, respectively; cf. Theorem 6.1. The temporal and spatial regularity of the optimal control $\bar{q}$ determines the constants $C_3(\bar{q})$ and $C_4(\bar{q})$, respectively.

In [19] a similar result is proved for the case $r = 0$, $s = 1$, and under the assumption $k \approx h^2$. We would like to emphasize that the discretization parameters k, h, k_d, h_d in estimate (1.2) can be chosen independently of each other.

The purpose of this paper is twofold. The first goal is to derive a priori error estimates for optimal control problem (1.1) of the above structure. The second goal is to provide techniques which will be used in the second part of the paper for derivation of a priori error estimates for problems involving pointwise inequality constraints on the control variable.

The paper is organized as follows. In the next section we recall the function analytic setting and optimality conditions for the optimal control problem under consideration. In section 3 the space-time finite element discretization is presented. Based on stability estimates developed in section 4, we provide a priori error analysis for the state equation in section 5. The main result on the error analysis for the considered optimal control problem is given in section 6. In this section error estimates for the error in the control, state, and adjoint variables are developed. In the last section we present a numerical example illustrating our results.

2. Optimization. In this section we briefly discuss the precise formulation of the optimization problem under consideration. Furthermore, we recall theoretical results on existence, uniqueness, and regularity of optimal solutions as well as optimality conditions.

To set up a weak formulation of the state equation (1.1b), we introduce the following notation: For a convex polygonal domain $\Omega \subset \mathbb{R}^n$, $n \in \{ 2, 3 \}$, we denote V to be either $H^1(\Omega)$ or $H^1_0(\Omega)$ depending on the prescribed type of boundary conditions (homogeneous Neumann or homogeneous Dirichlet). Together with $H = L^2(\Omega)$, the Hilbert space V and its dual V^* build a Gelfand triple $V \hookrightarrow H \hookrightarrow V^*$. Here and in what follows, we employ the usual notion for Lebesgue and Sobolev spaces.

For a time interval $I = (0, T)$ we introduce the state space

$$X \coloneqq \big\{ v \;\big|\; v \in L^2(I, V) \text{ and } \partial_t v \in L^2(I, V^*) \big\}$$

and the control space

$$Q = L^2(I, L^2(\Omega)).$$

In addition, we use the following notation for the inner products and norms on $L^2(\Omega)$ and $L^2(I, L^2(\Omega))$:

$$(v, w) \coloneqq (v, w)_{L^2(\Omega)}, \qquad (v, w)_I \coloneqq (v, w)_{L^2(I, L^2(\Omega))},$$
$$\|v\| \coloneqq \|v\|_{L^2(\Omega)}, \qquad \|v\|_I \coloneqq \|v\|_{L^2(I, L^2(\Omega))}.$$

In this setting, a standard weak formulation of the state equation (1.1b) for given control $q \in Q$, $f \in L^2(I, H)$, and $u_0 \in V$ reads as follows: Find a state $u \in X$ satisfying

$$\begin{aligned} (\partial_t u, \varphi)_I + (\nabla u, \nabla \varphi)_I &= (f + q, \varphi)_I \qquad \forall \varphi \in X, \\ u(0) &= u_0. \end{aligned} \tag{2.1}$$

For simplicity of notation, we skip here and throughout the paper the dependence of the solution variable on x and t.

For this formulation of the state equation, we recall the following result on existence and regularity.

Proposition 2.1. *For fixed control $q \in Q$, $f \in L^2(I, H)$, and $u_0 \in V$ there exists a unique solution $u \in X$ of problem* (2.1). *Moreover, the solution exhibits the improved regularity*

$$u \in L^2(I, H^2(\Omega) \cap V) \cap H^1(I, L^2(\Omega)) \hookrightarrow C(\bar{I}, V).$$

It holds the stability estimate

$$\|\partial_t u\|_I + \|\nabla^2 u\|_I \le C\big\{ \|f + q\|_I + \|\nabla u_0\| \big\}.$$

Proof. The proof of existence and uniqueness is given, e.g., in [18] and [29]. The improved regularity relies on the fact that Ω is polygonal and convex and is proved, e.g., in [11]. The embedding of $L^2(I, H^2(\Omega) \cap V) \cap H^1(I, L^2(\Omega))$ into $C(\bar{I}, V)$ can be found, for instance, in [6]. ∎

The weak formulation of the optimal control problem (1.1) is given as follows:

$$\text{Minimize } J(q, u) \coloneqq \frac{1}{2}\|u - \hat{u}\|_I^2 + \frac{\alpha}{2}\|q\|_I^2 \text{ subject to (2.1) and } (q, u) \in Q \times X, \tag{2.2}$$

where $\hat{u} \in L^2(I, H)$ is a given desired state and $\alpha > 0$ is the regularization parameter.

Proposition 2.2. *For given $f, \hat{u} \in L^2(I, H)$, $u_0 \in V$, and $\alpha > 0$, the optimal control problem* (2.2) *admits a unique solution $(\bar{q}, \bar{u}) \in Q \times X$. The optimal control $\bar{q}$ possesses the regularity*

$$\bar{q} \in L^2(I, H^2(\Omega)) \cap H^1(I, L^2(\Omega)).$$

Proof. For existence and uniqueness we refer to [18]. First order necessary optimality conditions and Proposition 2.1 imply the stated regularity of the optimal control. ∎

The existence result for the state equation in Proposition 2.1 ensures the existence of a control-to-state mapping $q \mapsto u = u(q)$ defined through (2.1). By means of this mapping we introduce the reduced cost functional $j : Q \to \mathbb{R}$:

$$j(q) := J(q, u(q)).$$

The optimal control problem (2.2) can then be equivalently reformulated as follows:

$$\text{Minimize } j(q) \text{ subject to } q \in Q. \tag{2.3}$$

The first order necessary optimality condition for (2.3) reads as

$$j'(\bar{q})(\delta q) = 0 \quad \forall \delta q \in Q. \tag{2.4}$$

Due to the linear-quadratic structure of the optimal control problem this condition is also sufficient for optimality.

Utilizing the adjoint state equation for $z = z(q) \in X$ given by

$$\begin{aligned} -(\varphi, \partial_t z)_I + (\nabla \varphi, \nabla z)_I &= (\varphi, u(q) - \hat{u})_I \quad \forall \varphi \in X, \\ z(T) &= 0, \end{aligned} \tag{2.5}$$

the first derivative of the reduced cost functional can be expressed as

$$j'(q)(\delta q) = (\alpha q + z(q), \delta q)_I. \tag{2.6}$$

3. Discretization.

In this section we describe the space-time finite element discretization of the optimal control problem (2.2).

3.1. Semidiscretization in time.

At first, we present the semidiscretization in time of the state equation by discontinuous Galerkin methods. We consider a partitioning of the time interval $\bar{I} = [0, T]$ as

$$\bar{I} = \{0\} \cup I_1 \cup I_2 \cup \cdots \cup I_M \tag{3.1}$$

with subintervals $I_m = (t_{m-1}, t_m]$ of size k_m and time points

$$0 = t_0 < t_1 < \cdots < t_{M-1} < t_M = T.$$

We define the discretization parameter k as a piecewise constant function by setting $k\big|_{I_m} = k_m$ for $m = 1, 2, \ldots, M$. Moreover, we denote by k the maximal size of the time steps, i.e., $k = \max k_m$.

The semidiscrete trial and test space is given as

$$X_k^r = \left\{ v_k \in L^2(I, V) \,\middle|\, v_k\big|_{I_m} \in \mathcal{P}_r(I_m, V),\ m = 1, 2, \ldots, M \right\}.$$

Here, $\mathcal{P}_r(I_m, V)$ denotes the space of polynomials up to order r defined on I_m with values in V. On X_k^r we use the notation

$$(v, w)_{I_m} := (v, w)_{L^2(I_m, L^2(\Omega))} \quad \text{and} \quad \|v\|_{I_m} := \|v\|_{L^2(I_m, L^2(\Omega))}.$$

To define the discontinuous Galerkin (dG(r)) approximation using the space X_k^r we employ the following definitions for functions $v_k \in X_k^r$:

$$v_{k,m}^+ := \lim_{t \to 0^+} v_k(t_m + t), \quad v_{k,m}^- := \lim_{t \to 0^+} v_k(t_m - t) = v_k(t_m), \quad [v_k]_m := v_{k,m}^+ - v_{k,m}^-$$

and define the bilinear form $B(\cdot, \cdot)$ for $u_k, \varphi \in X_k^r$ by

$$B(u_k, \varphi) := \sum_{m=1}^{M} (\partial_t u_k, \varphi)_{I_m} + (\nabla u_k, \nabla \varphi)_I + \sum_{m=2}^{M} ([u_k]_{m-1}, \varphi_{m-1}^+) + (u_{k,0}^+, \varphi_0^+). \tag{3.2}$$

Then, the dG(r) semidiscretization of the state equation (2.1) for a given control $q \in Q$ reads as follows: Find a state $u_k = u_k(q) \in X_k^r$ such that

$$B(u_k, \varphi) = (f + q, \varphi)_I + (u_0, \varphi_0^+) \quad \forall \varphi \in X_k^r. \tag{3.3}$$

The existence and uniqueness of solutions to (3.3) can be shown by using Fourier analysis; see [27] for details.

Remark 3.1. Using a density argument, it is possible to show that the exact solution $u = u(q) \in X$ also satisfies the identity

$$B(u, \varphi) = (f + q, \varphi)_I + (u_0, \varphi_0^+) \quad \forall \varphi \in X_k^r.$$

Thus, we have here the property of Galerkin orthogonality

$$B(u - u_k, \varphi) = 0 \quad \forall \varphi \in X_k^r,$$

although the dG(r) semidiscretization is a nonconforming Galerkin method ($X_k^r \not\subset X$).

The semidiscrete optimization problem for the dG(r) time discretization has the following form:

$$\text{Minimize } J(q_k, u_k) \text{ subject to (3.3) and } (q_k, u_k) \in Q \times X_k^r. \tag{3.4}$$

PROPOSITION 3.2. *The semidiscrete optimal control problem* (3.4) *admits for* $\alpha > 0$ *a unique solution* $(\bar{q}_k, \bar{u}_k) \in Q \times X_k^r$.

Proof. The proof is done by translating standard arguments from the proof in the continuous case and by employing the continuity of the mapping $q \mapsto u_k(q)$ provided by the stability estimates derived in the next section (cf. Theorem 4.3). ∎

Note that the optimal control $\bar{q}_k$ is searched for in the continuous space Q and the subscript k indicates the usage of the semidiscretized state equation.

Similar to the continuous case, we introduce the semidiscrete reduced cost functional $j_k : Q \to \mathbb{R}$:

$$j_k(q) := J(q, u_k(q))$$

and reformulate the semidiscrete optimal control problem (3.4) as follows:

$$\text{Minimize } j_k(q_k) \text{ subject to } q_k \in Q.$$

The first order necessary optimality condition reads as

$$j_k'(\bar{q}_k)(\delta q) = 0 \quad \forall \delta q \in Q, \tag{3.5}$$

and the derivative of j_k can be expressed as

$$j_k'(q)(\delta q) = (\alpha q + z_k(q), \delta q)_I. \tag{3.6}$$

Here, $z_k = z_k(q) \in X_k^r$ denotes the solution of the semidiscrete adjoint equation

$$B(\varphi, z_k) = (\varphi, u_k(q) - \hat{u})_I \quad \forall \varphi \in X_k^r. \tag{3.7}$$

Note that by using integration by parts in time, the bilinear form $B(\cdot,\cdot)$ defined by (3.2) can equivalently be expressed as

$$B(\varphi, z_k) = -\sum_{m=1}^{M} (\varphi, \partial_t z_k)_{I_m} + (\nabla \varphi, \nabla z_k)_I - \sum_{m=1}^{M-1} (\varphi_m^-, [z_k]_m) + (\varphi_M^-, z_{k,M}^-). \tag{3.8}$$

3.2. Discretization in space. To define the finite element discretization in space, we consider two or three dimensional shape-regular meshes; see, e.g., [5]. A mesh consists of quadrilateral or hexahedral cells K, which constitute a nonoverlapping cover of the computational domain Ω. The corresponding mesh is denoted by $\mathcal{T}_h = \{K\}$, where we define the discretization parameter h as a cellwise constant function by setting $h\big|_K = h_K$ with the diameter h_K of the cell K. We use the symbol h also for the maximal cell size, i.e., $h = \max h_K$.

On the mesh $\mathcal{T}_h$ we construct a conforming finite element space $V_h \subset V$ in a standard way:

$$V_h^s = \left\{ v \in V \big| v\big|_K \in \mathcal{Q}_s(K) \text{ for } K \in \mathcal{T}_h \right\}.$$

Here, $\mathcal{Q}_s(K)$ consists of shape functions obtained via (bi-/tri-)linear transformations of polynomials in $\widehat{\mathcal{Q}_s}(\widehat{K})$ defined on the reference cell $\widehat{K} = (0,1)^n$, where

$$\widehat{\mathcal{Q}_s}(\widehat{K}) = \operatorname{span} \left\{ \prod_{j=1}^{n} x_j^{\alpha_j} \middle| \alpha_j \in \mathbb{N}_0,\ \alpha_j \le s \right\}.$$

Remark 3.3. The definition of V_h^s can be extended to the case of triangular meshes in the obvious way.

To obtain the fully discretized versions of the time discretized state equation (3.3), we utilize the space-time finite element space

$$X_{k,h}^{r,s} = \left\{ v_{kh} \in L^2(I, V_h^s) \middle| v_{kh}\big|_{I_m} \in \mathcal{P}_r(I_m, V_h^s) \right\} \subset X_k^r.$$

Remark 3.4. Here, the spatial mesh and, therefore, also the space V_h^s is fixed for all time intervals. We refer to [25] for a discussion of treatment of different meshes $\mathcal{T}_h^m$ for each of the subintervals I_m.

The so-called cG(s)dG(r) discretization of the state equation for given control $q \in Q$ has the following form: Find a state $u_{kh} = u_{kh}(q) \in X_{k,h}^{r,s}$ such that

$$B(u_{kh}, \varphi) = (f + q, \varphi)_I + (u_0, \varphi_0^+) \quad \forall \varphi \in X_{k,h}^{r,s}. \tag{3.9}$$

Remark 3.5. The notation cG(s)dG(r) is taken from [7] and describes a method with conforming (continuous) discretization in space of order s and discontinuous discretization in time of order r.

Then, the corresponding optimal control problem is given as follows:

$$\text{Minimize } J(q_{kh}, u_{kh}) \text{ subject to (3.9) and } (q_{kh}, u_{kh}) \in Q \times X_{k,h}^{r,s}, \tag{3.10}$$

and by means of the discrete reduced cost functional $j_{kh}\colon Q \to \mathbb{R}$,

$$j_{kh}(q) := J(q, u_{kh}(q)),$$

it can be reformulated as follows:

$$\text{Minimize } j_{kh}(q_{kh}) \text{ subject to } q_{kh} \in Q.$$

The uniquely determined optimal solution of (3.10) is denoted by $(\bar{q}_{kh}, \bar{u}_{kh}) \in Q \times X_{k,h}^{r,s}$.

The optimal control $\bar{q}_{kh} \in Q$ fulfills the first order optimality condition

$$j'_{kh}(\bar{q}_{kh})(\delta q) = 0 \quad \forall \delta q \in Q, \tag{3.11}$$

where $j'_{kh}(q)(\delta q)$ is given by

$$j'_{kh}(q)(\delta q) = (\alpha q + z_{kh}(q), \delta q)_I \tag{3.12}$$

with the discrete adjoint solution $z_{kh} = z_{kh}(q) \in X_{k,h}^{r,s}$ of

$$B(\varphi, z_{kh}) = (\varphi, u_{kh}(q) - \hat{u})_I \quad \forall \varphi \in X_{k,h}^{r,s}. \tag{3.13}$$

3.3. Discretization of the controls. To obtain the fully discrete optimal control problem we restrict the control space Q to a finite dimensional subspace $Q_d \subset Q$. The optimal control problem on this level of discretization is given as follows:

$$\text{Minimize } J(q_\sigma, u_\sigma) \text{ subject to (3.9) and } (q_\sigma, u_\sigma) \in Q_d \times X_{k,h}^{r,s}. \tag{3.14}$$

The unique optimal solution of (3.14) is denoted by $(\bar{q}_\sigma, \bar{u}_\sigma) \in Q_d \times X_{k,h}^{r,s}$, where the subscript σ collects the discretization parameters k, h, and d. The optimality condition is given using the discrete reduced cost functional j_{kh} introduced in the previous section by

$$j'_{kh}(\bar{q}_\sigma)(\delta q) = 0 \quad \forall \delta q \in Q_d. \tag{3.15}$$

Most of our results presented below hold true independently of the choice of the control discretization; see Theorem 6.1. However, we present here some possibilities for construction of the discrete control space Q_d, which will play a role in the discussion of the error in the state and adjoint variables (see section 6.2) and which will be employed for the numerical example in section 7.

For the construction of Q_d it is possible to use spatial and temporal meshes, which are different from those employed for the discretization of the state variable. However, for simplicity of notation we will use the same time-partitioning (3.1). Using a spatial mesh $\mathcal{T}_{h_d}$ we consider two corresponding finite element spaces:

$$V_{h_d}^{s_d} = \left\{ v \in C(\bar{\Omega}) \middle| v\big|_K \in \mathcal{Q}_{s_d}(K) \text{ for } K \in \mathcal{T}_{h_d} \right\}$$

and

$$\widetilde{V}_{h_d}^{s_d} = \left\{ v \in L^2(\Omega) \big| v\big|_K \in \mathcal{Q}_{s_d}(K) \text{ for } K \in \mathcal{T}_{h_d} \right\}.$$

The space $V_{h_d}^{s_d}$ consists of continuous cellwise polynomial functions of order s_d, whereas the functions in the space $\widetilde{V}_{h_d}^{s_d}$ are discontinuous. Using these spaces we define two possibilities for the choice of Q_d.

The first possibility is similar to the construction of the state space $X_{k,h}^{r,s}$, consisting of functions which are continuous in space and discontinuous in time, and results in the following definition:

$$Q_d = \left\{ v_{kh} \in L^2(I, V_{h_d}^{s_d}) \Big| v_{kh}\big|_{I_m} \in \mathcal{P}_{r_d}(I_m, V_{h_d}^{s_d}) \right\}.$$

We will refer to this control discretization as cG(s_d)dG(r_d). If the control mesh $\mathcal{T}_{h_d}$ coincides with the state mesh $\mathcal{T}_h$ and one chooses the same order of polynomials ($r = r_d$, $s = s_d$), then the state space $X_{k,h}^{r,s}$ coincides with the control space Q_d in case of homogeneous Neumann boundary conditions and is a subspace of it, i.e., $X_{k,h}^{r,s} \subset Q_d$ in the presence of homogeneous Dirichlet boundary conditions. In this case one can show (cf. the discussion in section 6) that $\bar{q}_{kh} = \bar{q}_\sigma$. This means that a complete discretization of the optimal control problem is achieved already after discretization of the state equation; cf. [16].

For the second possibility we employ the space $\widetilde{V}_{h_d}^{s_d}$ of discontinuous cellwise polynomials and obtain the following definition:

$$Q_d = \left\{ v_{kh} \in L^2(I, \widetilde{V}_{h_d}^{s_d}) \Big| v_{kh}\big|_{I_m} \in \mathcal{P}_{r_d}(I_m, \widetilde{V}_{h_d}^{s_d}) \right\}.$$

We will refer to this control discretization as dG(s_d)dG(r_d). The special choice $s_d = 0$ leads to cellwise constant discretization in space.

4. Stability estimates for the state and adjoint equations.

The first step in proving the desired a priori estimate is to prove stability estimates for the solution of the semidiscrete (3.3) and the fully discretized (3.9) state equation. Throughout this section we discuss the uncontrolled situation and set therefore $q = 0$.

In the following theorem we provide a stability estimate for semidiscretization in time, which has a structure similar to the estimate on the continuous level given in Proposition 2.1. A comparable estimate is shown in [8, 9] for the case $f = 0$.

THEOREM 4.1. *For the solution $u_k \in X_k^r$ of the* dG(r) *semidiscretized state equation* (3.3) *with right-hand side $f \in L^2(I, H)$, initial condition $u_0 \in V$, and $q = 0$, the stability estimate*

$$\sum_{m=1}^{M} \|\partial_t u_k\|_{I_m}^2 + \|\Delta u_k\|_I^2 + \sum_{m=1}^{M} k_m^{-1} \|[u_k]_{m-1}\|^2 \le C\left\{ \|f\|_I^2 + \|\nabla u_0\|^2 \right\}$$

holds. The constant C depends only on the polynomial degree r and the domain Ω. The jump term $[u_k]_0$ at $t = 0$ is defined as $u_{k,0}^+ - u_0$.

Proof. By means of the definition $[u_k]_0 = u_{k,0}^+ - u_0$, the solution $u_k \in X_k^r$ of (3.3) fulfills for all $\varphi \in \mathcal{P}_r(I_m, V)$ the following system of equations:

$$(4.1) \quad (\partial_t u_k, \varphi)_{I_m} + (\nabla u_k, \nabla \varphi)_{I_m} + ([u_k]_{m-1}, \varphi_{m-1}^+) = (f, \varphi)_{I_m}, \quad m = 1, 2, \dots, M.$$

The proof of the desired estimate consist of three steps—one for each term of the left-hand side of (4.1). The steps are based on consecutively testing with $\varphi = -\Delta u_k$, $\varphi = (t - t_{m-1})\partial_t u_k$, and $\varphi = [u_k]_{m-1}$.

Step (i). At first, we want to choose $\varphi = -\Delta u_k$. For applying integration by parts in space to (4.1), it is necessary to prove $\Delta u_k\big|_{I_m} \in \mathcal{P}_r(I_m, H)$. This assertion follows immediately from applying elliptic regularity theory (cf. [11]) to the transformed time stepping equation

$$(\nabla u_k, \nabla\varphi)_{I_m} = (f - \partial_t u_k, \varphi)_{I_m} - ([u_k]_{m-1}, \varphi^+_{m-1}).$$

The fact that $u_k\big|_{I_m}$ is polynomial in time with values in $V \subseteq H$ implies that the right-hand side is in H for almost all $t \in I_m$. Thus, $\Delta u_k\big|_{I_m}$ is also in H for almost all $t \in I_m$, and since $u_k\big|_{I_m}$ is polynomial with respect to time, this yields $\Delta u_k\big|_{I_m} \in \mathcal{P}_r(I_m, H)$.

Consequently, it is feasible to integrate (4.1) by parts in space to obtain the formulation

$$(\partial_t u_k, \varphi)_{I_m} - (\Delta u_k, \varphi)_{I_m} + ([u_k]_{m-1}, \varphi^+_{m-1}) = (f, \varphi)_{I_m}, \quad m = 1, 2, \ldots, M. \tag{4.2}$$

The arising boundary terms vanish for both homogeneous Neumann and homogeneous Dirichlet boundary conditions.

Since there are no spatial derivatives on the test function φ anymore, formulation (4.2) holds not only for all $\varphi \in \mathcal{P}_r(I_m, V)$ but by the density of V in H also for all $\varphi \in \mathcal{P}_r(I_m, H)$. Hence, we may choose $\varphi = -\Delta u_k$ as a test function and get, by applying integration by parts in space a second time,

$$(\partial_t \nabla u_k, \nabla u_k)_{I_m} + (\Delta u_k, \Delta u_k)_{I_m} + ([\nabla u_k]_{m-1}, \nabla u^+_{k,m-1}) = (f, -\Delta u_k)_{I_m}.$$

Again, the arising boundary terms vanish due to the prescribed homogeneous boundary conditions of Neumann or Dirichlet type.

By means of the identities

$$(\partial_t v, v)_{I_m} = \frac{1}{2}\|v^-_m\|^2 - \frac{1}{2}\|v^+_{m-1}\|^2, \tag{4.3a}$$

$$([v]_{m-1}, v^+_{m-1}) = \frac{1}{2}\|v^+_{m-1}\|^2 + \frac{1}{2}\|[v]_{m-1}\|^2 - \frac{1}{2}\|v^-_{m-1}\|^2, \tag{4.3b}$$

we achieve

$$\frac{1}{2}\|\nabla u^-_{k,m}\|^2 + \frac{1}{2}\|[\nabla u_k]_{m-1}\|^2 - \frac{1}{2}\|\nabla u^-_{k,m-1}\|^2 + \|\Delta u_k\|^2_{I_m} = (f, -\Delta u_k)_{I_m}.$$

Summation of the equations for $m = 1, 2, \ldots, M$ leads to

$$\frac{1}{2}\|\nabla u^-_{k,M}\|^2 + \frac{1}{2}\sum_{m=1}^{M}\|[\nabla u_k]_{m-1}\|^2 + \|\Delta u_k\|^2_I = (f, -\Delta u_k)_I + \frac{1}{2}\|\nabla u_0\|^2.$$

Using Young's inequality for the right-hand side, we obtain the first intermediary result

$$\|\Delta u_k\|^2_I \le \|f\|^2_I + \|\nabla u_0\|^2. \tag{4.4}$$

Step (ii). To bound the time derivative $\partial_t u_k$, we will use the inverse estimate

$$\|v_k\|_{I_m}^2 \leq C k_m^{-1} \int_{I_m} (t - t_{m-1}) \|v_k\|^2 \, dt, \tag{4.5}$$

which holds true for all $v_k \in \mathcal{P}_r(I_m, V)$. To obtain this estimate one transforms both sides of the inequality into the reference time interval $[0, 1]$, uses equivalence of norms for finite dimensional spaces, and transforms the inequality back into the real interval I_m.

We choose $\varphi = (t - t_{m-1})\partial_t u_k$ and obtain, utilizing the fact that $\varphi_{m-1}^+ = 0$,

$$\int_{I_m} (t - t_{m-1}) \|\partial_t u_k\|^2 \, dt = \int_{I_m} (t - t_{m-1})(f + \Delta u_k, \partial_t u_k) \, dt$$

$$\leq \left(\int_{I_m} (t - t_{m-1}) \|f + \Delta u_k\|^2 \, dt \right)^{\frac{1}{2}} \left(\int_{I_m} (t - t_{m-1}) \|\partial_t u_k\|^2 \, dt \right)^{\frac{1}{2}}.$$

The inverse estimate (4.5) yields, by means of Hölder's inequality,

$$\|\partial_t u_k\|_{I_m}^2 \leq C k_m^{-1} \int_{I_m} (t - t_{m-1}) \|f + \Delta u_k\|^2 \, dt \leq C \big\{ \|f\|_{I_m}^2 + \|\Delta u_k\|_{I_m}^2 \big\}.$$

Then, (4.4) implies the second intermediary result

$$\sum_{m=1}^{M} \|\partial_t u_k\|_{I_m}^2 \leq C \big\{ \|f\|_I^2 + \|\nabla u_0\|^2 \big\}. \tag{4.6}$$

Step (iii). It remains to estimate the jump terms. To this end, we choose $\varphi = [u_k]_{m-1}$ constant in time on I_m and obtain

$$\|[u_k]_{m-1}\|^2 = (f + \Delta u_k - \partial_t u_k, [u_k]_{m-1})_{I_m}$$

$$\leq \frac{k_m}{2} \|f + \Delta u_k - \partial_t u_k\|_{I_m}^2 + \frac{1}{2k_m} \|[u_k]_{m-1}\|_{I_m}^2.$$

Since $[u_k]_{m-1}$ is constant in time, we have $\|[u_k]_{m-1}\|_{I_m}^2 = k_m \|[u_k]_{m-1}\|^2$. This implies

$$k_m^{-1} \|[u_k]_{m-1}\|^2 \leq \|f + \Delta u_k - \partial_t u_k\|_{I_m}^2.$$

The results (4.4) and (4.6) yield the remaining estimate

$$\sum_{m=1}^{M} k_m^{-1} \|[u_k]_{m-1}\|^2 \leq C \big\{ \|f\|_I^2 + \|\nabla u_0\|^2 \big\}. \qquad \square$$

The result of the previous theorem will also be applied to dual (adjoint) equations. Let $g \in L^2(I, H)$ be a given right-hand side and $z_T \in V$ a given terminal condition; then the corresponding semidiscretized dual equation is given by

$$B(\varphi, z_k) = (\varphi, g)_I + (\varphi_M^-, z_T) \quad \forall \varphi \in X_k^r. \tag{4.7}$$

Note that the semidiscrete adjoint solution defined in (3.7) can be obtained by setting $g = u_k(q) - \hat{u}$ and $z_T = 0$.

COROLLARY 4.2. *For the solution $z_k \in X_k^r$ of the semidiscrete dual equation (4.7) with right-hand side $g \in L^2(I, H)$ and terminal condition $z_T \in V$, the estimate from Theorem 4.1 reads as*

$$\sum_{m=1}^{M} \|\partial_t z_k\|_{I_m}^2 + \|\Delta z_k\|_I^2 + \sum_{m=1}^{M} k_m^{-1} \|[z_k]_m\|^2 \le C\big\{\|g\|_I^2 + \|\nabla z_T\|^2\big\}.$$

Here, the jump term $[z_k]_M$ at $t = T$ is defined as $z_T - z_{k,M}^-$.

Proof. Let $z_k \in X_k^r$ be the solution of (4.7). Then formula (3.8) implies that it also fulfills for all $\varphi \in \mathcal{P}_r(I_m, V)$ the following system of equations:

$$-(\varphi, \partial_t z_k)_{I_m} + (\nabla\varphi, \nabla z_k)_{I_m} - (\varphi_m^-, [z_k]_m) = (g, \varphi)_{I_m}, \quad m = 1, 2, \ldots, M.$$

Based on this representation, all steps of the proof of Theorem 4.1 can be repeated similarly to obtain the stated result. ◻

For proving a priori estimates for the control problem (2.2), we will additionally need stability estimates for the $L^2(I, H)$-norm of the solution $\|u_k\|_I$ and of its gradient $\|\nabla u_k\|_I$, which are given in the following theorem.

THEOREM 4.3. *For the solution $u_k \in X_k^r$ of the* dG(r) *semidiscretized state equation (3.3) with right-hand side $f \in L^2(I, H)$, initial condition $u_0 \in V$, and $q = 0$, the stability estimate*

$$\|u_k\|_I^2 + \|\nabla u_k\|_I^2 \le C\big\{\|f\|_I^2 + \|\nabla u_0\|^2 + \|u_0\|^2\big\}$$

holds true with a constant C that depends only on the polynomial degree r, the domain Ω, and the final time T.

Remark 4.4. In the case of homogeneous Dirichlet boundary conditions, the estimate can be proved by means of Poincaré's inequality with a constant independent of T.

Proof. The proof is done using a duality argument: Let $\tilde{z} \in X$ be the solution of

$$-(\varphi, \partial_t \tilde{z})_I + (\nabla\varphi, \nabla\tilde{z})_I = (\varphi, u_k)_I \quad \forall \varphi \in X$$

together with the terminal condition $\tilde{z}(T) = \tilde{z}_T = 0$. Thus, due to Remark 3.1, which applies similarly to dual or adjoint equations, $\tilde{z}$ also fulfills

$$B(\varphi, \tilde{z}) = (\varphi, u_k)_I \quad \forall \varphi \in X_k^r.$$

By means of this equality, we write

$$\begin{aligned}\|u_k\|_I^2 &= B(u_k, \tilde{z}) \\ &= \sum_{m=1}^{M} (\partial_t u_k, \tilde{z})_{I_m} + (\nabla u_k, \nabla\tilde{z})_I + \sum_{m=2}^{M} ([u_k]_{m-1}, \tilde{z}(t_{m-1})) + (u_{k,0}^+, \tilde{z}(0)).\end{aligned}$$

Using the setting $[u_k]_0 = u_{k,0}^+ - u_0$ we obtain

$$\|u_k\|_I^2 = \sum_{m=1}^{M} (\partial_t u_k, \tilde{z})_{I_m} + (\nabla u_k, \nabla\tilde{z})_I + \sum_{m=1}^{M} ([u_k]_{m-1}, \tilde{z}(t_{m-1})) + (u_0, \tilde{z}(0)),$$

from which we get, with integration by parts in space and Hölder's inequality,

$$\|u_k\|_I^2 \le \left(\sum_{m=1}^M \|\partial_t u_k\|_{I_m}^2\right)^{\frac{1}{2}} \|\tilde{z}\|_I + \|\Delta u_k\|_I \|\tilde{z}\|_I + \left(\sum_{m=1}^M k_m^{-1} \|[u_k]_{m-1}\|^2\right)^{\frac{1}{2}} \left(\sum_{m=1}^M k_m \|\tilde{z}(t_{m-1})\|^2\right)^{\frac{1}{2}} + \|u_0\| \|\tilde{z}(0)\|.$$

The stability estimate for the continuous solution $\tilde{z} \in X$

$$\max_{t \in \bar{I}} \|\tilde{z}(t)\| \le C \|u_k\|_I,$$

which makes use of the continuity of the mapping $u_k \mapsto \tilde{z}$ (cf. [18]) and the continuous embedding of X into $C(\bar{I}, H)$, implies

$$\|u_k\|_I \le C\sqrt{T} \left(\sum_{m=1}^M \|\partial_t u_k\|_{I_m}^2\right)^{\frac{1}{2}} + C\sqrt{T} \|\Delta u_k\|_I + C\sqrt{T} \left(\sum_{m=2}^M k_m^{-1} \|[u_k]_{m-1}\|^2\right)^{\frac{1}{2}} + C\|u_0\|,$$

from which the desired estimate for $\|u_k\|_I^2$ follows by application of Theorem 4.1.

To prove the estimate for $\|\nabla u_k\|_I^2$, we proceed similarly to the proof of Theorem 4.1 and test (4.1) with $\varphi = u_k$. We obtain for $m = 1, 2, \ldots, M$

$$(\partial_t u_k, u_k)_{I_m} + (\nabla u_k, \nabla u_k)_{I_m} + ([u_k]_{m-1}, u_{k,m-1}^+) = (f, u_k)_{I_m}.$$

The identities (4.3) lead to

$$\frac{1}{2}\|u_{k,m}^-\|^2 + \frac{1}{2}\|[u_k]_{m-1}\|^2 - \frac{1}{2}\|u_{k,m-1}^-\|^2 + \|\nabla u_k\|_{I_m}^2 = (f, u_k)_{I_m}.$$

After summing up these equations for $m = 1, 2, \ldots, M$ and by application of Young's inequality, we have

$$\|\nabla u_k\|_I^2 \le \frac{1}{2}\big\{\|f\|_I^2 + \|u_k\|_I^2 + \|u_0\|^2\big\}.$$

Insertion of the already proved estimate for $\|u_k\|_I^2$ completes the proof. ∎

COROLLARY 4.5. *For the solution $z_k \in X_k^r$ of the semidiscrete dual equation* (4.7) *with right-hand side $g \in L^2(I, H)$ and terminal condition $z_T \in V$, the estimate from Theorem* 4.3 *reads as*

$$\|z_k\|_I^2 + \|\nabla z_k\|_I^2 \le C\big\{\|g\|_I^2 + \|\nabla z_T\|^2 + \|z_T\|^2\big\}.$$

Proof. The proof is done similarly to the proof of Theorem 4.3. ∎

All the estimates proved in this section hold true also for the fully discrete cG(s)dG(r) solutions u_{kh}, $z_{kh} \in X_{k,h}^{r,s}$ almost without any changes. Only two differences have to be regarded: We have to replace the continuous Laplacian Δ by a discrete analogue $\Delta_h \colon V_h^s \to V_h^s$ defined by

$$(\Delta_h u, \varphi) = -(\nabla u, \nabla \varphi) \quad \forall \varphi \in V_h^s,$$

and the jump terms $[u_{kh}]_0$ and $[z_{kh}]_M$ are given here by means of the spatial L^2-projection $\Pi_h\colon V \to V_h^s$ as

$$[u_{kh}]_0 = u_{kh,0}^+ - \Pi_h u_0 \qquad \text{and} \qquad [z_{kh}]_M = \Pi_h z_T - z_{kh,M}^-.$$

Here, $z_{kh} \in X_{k,h}^{r,s}$ is the solution of the fully discretized dual equation with given right-hand side $g \in L^2(I,H)$ and terminal condition $z_T \in V$ given by

$$B(\varphi, z_{kh}) = (\varphi, g)_I + (\varphi_M^-, z_T) \quad \forall \varphi \in X_{k,h}^{r,s}. \tag{4.8}$$

For convenience of the reader, we state here the estimates for the fully discrete solutions.

THEOREM 4.6. *For the solution $u_{kh} \in X_{k,h}^{r,s}$ of the discrete state equation* (3.9) *with right-hand side $f \in L^2(I,H)$, initial condition $u_0 \in V$, and $q = 0$, the stability estimate*

$$\sum_{m=1}^{M} \|\partial_t u_{kh}\|_{I_m}^2 + \|\Delta_h u_{kh}\|_I^2 + \sum_{m=1}^{M} k_m^{-1} \|[u_{kh}]_{m-1}\|^2 \le C\big\{\|f\|_I^2 + \|\nabla \Pi_h u_0\|^2\big\}$$

holds. The constant C depends only on the polynomial degree r and the domain Ω. The jump term $[u_{kh}]_0$ at $t = 0$ is defined as $u_{kh,0}^+ - \Pi_h u_0$. Furthermore, the estimate

$$\|u_{kh}\|_I^2 + \|\nabla u_{kh}\|_I^2 \le C\big\{\|f\|_I^2 + \|\nabla \Pi_h u_0\|^2 + \|\Pi_h u_0\|^2\big\}$$

holds true with a constant C that depends only on the polynomial degree r, the domain Ω, and the final time T.

COROLLARY 4.7. *For the solution $z_{kh} \in X_{k,h}^{r,s}$ of the discrete dual equation* (4.8) *with right-hand side $g \in L^2(I,H)$ and terminal condition $z_T \in V$, the estimates from Theorem* 4.6 *read as*

$$\sum_{m=1}^{M} \|\partial_t z_{kh}\|_{I_m}^2 + \|\Delta_h z_{kh}\|_I^2 + \sum_{m=1}^{M} k_m^{-1} \|[z_{kh}]_m\|^2 \le C\big\{\|g\|_I^2 + \|\nabla \Pi_h z_T\|^2\big\}$$

and

$$\|z_{kh}\|_I^2 + \|\nabla z_{kh}\|_I^2 \le C\big\{\|g\|_I^2 + \|\nabla \Pi_h z_T\|^2 + \|\Pi_h z_T\|^2\big\}.$$

Here, the jump term $[z_{kh}]_M$ at $t = T$ is defined as $\Pi_h z_T - z_{kh,M}^-$.

5. Analysis of the discretization error for the state equation. The goal of this section is to prove an a priori error estimate for the discretization error of the (uncontrolled) state equation. Due to the choice of the control space $Q = L^2(I, L^2(\Omega))$, we will need error estimates for the error in the state (and adjoint) variable with respect to the norm of $L^2(I, L^2(\Omega))$; cf. the discussion in section 6. Similar error estimates with respect to the $L^\infty(I, L^2(\Omega))$-norm can be found in [8, 9], and with respect to the $L^2(I, H^1(\Omega))$-norm in [13].

Let $u \in X$ be the solution of the state equation (2.1) for $q = 0$, $u_k \in X_k^r$ be the solution of the corresponding semidiscretized equation (3.3), and $u_{kh} \in X_{k,h}^{r,s}$ be the solution of the fully discretized state equation (3.9). To separate the influences of the space and time discretizations, we split the total discretization error $e := u - u_{kh}$ into its temporal part $e_k := u - u_k$ and its spatial part $e_h := u_k - u_{kh}$. The temporal discretization error will be estimated in the following subsection, and the spatial discretization error is treated in section 5.2.

Throughout this section we will assume that the solutions $u \in X$ and $u_k \in X_k^r$ possess the regularity $\partial_t^{r+1} u \in L^2(I, L^2(\Omega))$ and $\nabla^{s+1} u_k \in L^2(I, L^2(\Omega))$. Note that Proposition 2.1 and Theorem 4.1 ensure this assumption for $s = 1$ and $r = 0$ for convex polygonal domains Ω. Better regularity results ($r > 0$, $s > 1$) usually require stronger assumptions on the domain Ω and additional compatibility relations.

5.1. Analysis of the temporal discretization error. In this section, we will prove the following error estimate for the temporal discretization error e_k.

THEOREM 5.1. *For the error $e_k := u - u_k$ between the continuous solution $u \in X$ of* (2.1) *and the* dG(r) *semidiscretized solution $u_k \in X_k^r$ of* (3.3)*, we have the error estimate*

$$\|e_k\|_I \leq C k^{r+1} \|\partial_t^{r+1} u\|_I,$$

where the constant C is independent of the size of the time steps k.

For clarity of presentation, we divide the proof of this theorem into several steps, which are discussed in the following lemmas.

Before doing so, we define a semidiscrete projection $\pi_k \colon C(\bar{I}, V) \to X_k^r$ for $m = 1, 2, \ldots, M$ by $\pi_k u\big|_{I_m} \in \mathcal{P}_r(I_m, V)$ and

$$(\pi_k u - u, \varphi)_{I_m} = 0 \qquad \forall \varphi \in \mathcal{P}_{r-1}(I_m, V) \quad \text{for } r > 0, \tag{5.1a}$$

$$\pi_k u(t_m^-) = u(t_m^-). \tag{5.1b}$$

In the case $r = 0$, $\pi_k u$ is defined solely by condition (5.1b). The projection π_k is well-defined by these conditions; see, for instance, [27] or [26]. By Proposition 2.1 the solution u of (2.1) belongs to $C(\bar{I}, V)$, and therefore π_k is applicable to the state u.

To shorten the notation in the following analysis, we introduce the abbreviations

$$\eta_k := u - \pi_k u \quad \text{and} \quad \xi_k := \pi_k u - u_k$$

and split the error e_k as

$$e_k = \eta_k + \xi_k.$$

LEMMA 5.2. *For the projection error η_k defined above, the identity*

$$B(\eta_k, \varphi) = (\nabla \eta_k, \nabla \varphi)_I$$

holds for all $\varphi \in X_k^r$.

Proof. By means of (3.8), we have

$$B(\eta_k, \varphi) = -\sum_{m=1}^{M} (\eta_k, \partial_t \varphi)_{I_m} + (\nabla \eta_k, \nabla \varphi)_I - \sum_{m=1}^{M-1} (\eta_{k,m}^-, [\varphi]_m) + (\eta_{k,M}^-, \varphi_{k,M}^-).$$

The term $(\eta_k, \partial_t \varphi)_{I_m}$ vanishes due to (5.1a), and $\eta_{k,m}^- = 0$ for all m due to (5.1b). This completes the proof. ∎

LEMMA 5.3. *The temporal discretization error $e_k = u - u_k$ is bounded by the projection error η_k with respect to the $L^2(I, L^2(\Omega))$-norm, that is,*

$$\|e_k\|_I \leq C \|\eta_k\|_I.$$

Proof. We define $\tilde{z}_k \in X_k^r$ to be the solution of

$$B(\varphi, \tilde{z}_k) = (\varphi, e_k)_I \quad \forall \varphi \in X_k^r.$$

Thus, we obtain by Galerkin orthogonality the relation $B(\xi_k + \eta_k, \tilde{z}_k) = 0$ (cf. Remark 3.1) which implies

$$\|e_k\|_I^2 = (\xi_k, e_k)_I + (\eta_k, e_k)_I = B(\xi_k, \tilde{z}_k) + (\eta_k, e_k)_I = -B(\eta_k, \tilde{z}_k) + (\eta_k, e_k)_I.$$

Using Lemma 5.2 and integration by parts in space, and the stability estimate from Corollary 4.2, it follows that

$$-B(\eta_k, \tilde{z}_k) = -(\nabla \eta_k, \nabla \tilde{z}_k)_I = (\eta_k, \Delta \tilde{z}_k)_I \leq \|\eta_k\|_I \|\Delta \tilde{z}_k\|_I \leq C \|\eta_k\|_I \|e_k\|_I.$$

Note that the arising boundary terms vanish for both homogeneous Neumann and homogeneous Dirichlet boundary conditions. This leads, by means of Cauchy's inequality, to the desired assertion. ∎

LEMMA 5.4. *For the projection error $\eta_k = u - \pi_k u$ the following estimate holds:*

$$\|\eta_k\|_{I_m} \leq C k_m^{r+1} \|\partial_t^{r+1} u\|_{I_m}.$$

Proof. Similarly to [27], the proof is done by standard arguments utilizing the Bramble–Hilbert lemma from [3]. ∎

After these preparations, we are able to give the proof of Theorem 5.1.

Proof of Theorem 5.1. From the Lemmas 5.3 and 5.4 we directly obtain

$$\|e_k\|_I^2 \leq C \|\eta_k\|_I^2 = C \sum_{m=1}^M \|\eta_k\|_{I_m}^2 \leq C \sum_{m=1}^M k_m^{2r+2} \|\partial_t^{r+1} u\|_{I_m}^2 \leq C k^{2r+2} \|\partial_t^{r+1} u\|_I^2,$$

which implies the stated result. ∎

5.2. Analysis of the spatial discretization error. In this section we give a proof of the following result.

THEOREM 5.5. *For the error $e_h := u_k - u_{kh}$ between the* dG(r) *semidiscretized solution $u_k \in X_k^r$ of* (3.3) *and the fully* cG(s)dG(r) *discretized solution $u_{kh} \in X_{k,h}^{r,s}$ of* (3.9)*, we have the error estimate*

$$\|e_h\|_I \leq C h^{s+1} \|\nabla^{s+1} u_k\|_I,$$

where the constant C is independent of the mesh size h and the size of the time steps k.

Similar to the previous subsection, the proof is divided into several steps which are collected in the following lemmas.

We define the projection $\pi_h : X_k^r \to X_{k,h}^{r,s}$ by means of the spatial L^2-projection $\Pi_h : V \to V_h^s$ pointwise in time as

$$(\pi_h u_k)(t) = \Pi_h u_k(t).$$

For the solutions of the semidiscrete and fully discretized state equations $u_k \in X_k^r$ and $u_{kh} \in X_{k,h}^{r,s}$, and for $\tilde{z}_k \in X_k^r$ being the solution of the dual equation (4.7) with right-hand side $g = e_h$ and terminal condition $\tilde{z}_T = 0$, we use the abbreviations

$$\eta_h := u_k - \pi_h u_k, \quad \xi_h := \pi_h u_k - u_{kh}, \quad \text{and} \quad \eta_h^* := \tilde{z}_k - \pi_h \tilde{z}_k,$$

and split the error e_h as

$$e_h = \eta_h + \xi_h.$$

LEMMA 5.6. *For the projection errors η_h and η_h^* defined above, the identities*

$$B(\eta_h, \varphi) = (\nabla \eta_h, \nabla \varphi)_I \qquad \text{and} \qquad B(\varphi, \eta_h^*) = (\nabla \varphi, \nabla \eta_h^*)_I$$

hold for all $\varphi \in X_{k,h}^{r,s}$.

Proof. As in the proof of Lemma 5.2 we obtain

$$\begin{aligned} B(\eta_h, \varphi) &= -\sum_{m=1}^{M} (\eta_h, \partial_t \varphi)_{I_m} + (\nabla \eta_h, \nabla \varphi)_I - \sum_{m=1}^{M-1} (\eta_{h,m}^-, [\varphi]_m) + (\eta_{h,M}^-, \varphi_M^-) \\ &= (\nabla \eta_h, \nabla \varphi)_I \end{aligned}$$

by means of the definition of π_h. The assertion for $B(\varphi, \eta_h^*)$ follows directly when employing representation (3.2) instead of (3.8). $\square$

LEMMA 5.7. *For the error ξ_h and the projection error η_h, the estimate*

$$\|\nabla \xi_h\|_I \leq \|\nabla \eta_h\|_I$$

holds.

Proof. As in [13], we have for all $v \in X_k^r$ by (3.2) and (3.8)

$$\begin{aligned} B(v, v) &= \sum_{m=1}^{M} (\partial_t v, v)_{I_m} + (\nabla v, \nabla v)_I + \sum_{m=1}^{M-1} ([v]_m, \quad v_m^+) + (v_0^+, v_0^+), \\ B(v, v) &= -\sum_{m=1}^{M} (v, \partial_t v)_{I_m} + (\nabla v, \nabla v)_I + \sum_{m=1}^{M-1} (-v_m^-, [v]_m) + (v_M^-, v_M^-). \end{aligned}$$

We arrive at

$$B(v, v) \geq (\nabla v, \nabla v)_I$$

by adding these two identities. Utilizing the Galerkin orthogonality of the space discretization, we obtain

$$\|\nabla \xi_h\|_I^2 = (\nabla \xi_h, \nabla \xi_h)_I \leq B(\xi_h, \xi_h) = -B(\eta_h, \xi_h) = -(\nabla \eta_h, \nabla \xi_h)_I \leq \|\nabla \eta_h\|_I \|\nabla \xi_h\|_I.$$

Division by $\|\nabla \xi_h\|_I$ leads to the asserted result. $\square$

LEMMA 5.8. *For the projection errors η_h and η_h^* we have the intermediary result*

$$B(\eta_h, \eta_h^*) \leq \|\nabla \eta_h\|_I \|\nabla \eta_h^*\|_I + C \|\eta_h\|_I \|e_h\|_I.$$

Proof. Since $\pi_h \tilde{z}_k \in X_{k,h}^{r,s}$, it holds by (3.8) and the definition of π_h that

$$B(\eta_h, \eta_h^*) = -\sum_{m=1}^{M} (\eta_h, \partial_t \tilde{z}_k)_{I_m} + (\nabla \eta_h, \nabla \eta_h^*)_I - \sum_{m=1}^{M-1} (\eta_{h,m}^-, [\tilde{z}_k]_m) + (\eta_{h,M}^-, \tilde{z}_{k,M}^-).$$

Using $\tilde{z}_T = 0$, we subtract the term $(\eta_{h,M}^-, \tilde{z}_T)$ and obtain by means of the definition $[\tilde{z}_k]_M = \tilde{z}_T - \tilde{z}_{k,M}^-$

$$B(\eta_h, \eta_h^*) = -\sum_{m=1}^{M} (\eta_h, \partial_t \tilde{z}_k)_{I_m} + (\nabla \eta_h, \nabla \eta_h^*)_I - \sum_{m=1}^{M} (\eta_{h,m}^-, [\tilde{z}_k]_m). \tag{5.2}$$

Now, we separately treat the three terms on the right-hand side above: For the term containing spatial derivatives, we have immediately

$$(\nabla \eta_h, \nabla \eta_h^*)_I \leq \|\nabla \eta_h\|_I \|\nabla \eta_h^*\|_I. \tag{5.3}$$

For the term containing the time derivatives, we achieve by Cauchy's inequality and with the stability estimate from Corollary 4.2

$$-\sum_{m=1}^{M} (\eta_h, \partial_t \tilde{z}_k)_{I_m} \leq \|\eta_h\|_I \left(\sum_{m=1}^{M} \|\partial_t \tilde{z}_k\|_{I_m}^2 \right)^{\frac{1}{2}} \leq C \|\eta_h\|_I \|e_h\|_I. \tag{5.4}$$

For the jump terms, we obtain again by Cauchy's inequality

$$-\sum_{m=1}^{M} (\eta_{h,m}^-, [\tilde{z}_k]_m) \leq \left(\sum_{m=1}^{M} k_m \|\eta_{h,m}^-\|^2 \right)^{\frac{1}{2}} \left(\sum_{m=1}^{M} k_m^{-1} \|[\tilde{z}_k]_m\|^2 \right)^{\frac{1}{2}}.$$

Utilizing the inverse estimate (cf. [8])

$$k_m \|\eta_{h,m}^-\|^2 \leq C \|\eta_h\|_{I_m}^2,$$

which holds true for polynomials in time, and the stability estimate from Corollary 4.2, we finally obtain

$$-\sum_{m=1}^{M} (\eta_{h,m}^-, [\tilde{z}_k]_m) \leq C \|\eta_h\|_I \|e_h\|_I. \tag{5.5}$$

We complete the proof by inserting the three estimates (5.3), (5.4), and (5.5) into (5.2). ∎

We are now prepared to give the proof of Theorem 5.5.

Proof of Theorem 5.5. The solution $\tilde{z}_k \in X_k^r$ defined above satisfies

$$B(\varphi, \tilde{z}_k) = (\varphi, e_h)_I \quad \forall \varphi \in X_k^r.$$

Due to Galerkin orthogonality, which is applicable for $\pi_h \tilde{z}_k \in X_{k,h}^{r,s}$, the identity

$$\|e_h\|_I^2 = B(e_h, \tilde{z}_k) = B(e_h, \tilde{z}_k - \pi_h \tilde{z}_k) = B(\xi_h, \eta_h^*) + B(\eta_h, \eta_h^*)$$

is fulfilled. For the first term we obtain, using Lemma 5.6 and Lemma 5.7,

$$B(\xi_h, \eta_h^*) = (\nabla \xi_h, \nabla \eta_h^*)_I \leq \|\nabla \xi_h\|_I \|\nabla \eta_h^*\|_I \leq \|\nabla \eta_h\|_I \|\nabla \eta_h^*\|_I.$$

This yields, together with Lemma 5.8,

$$\|e_h\|_I^2 \leq 2 \|\nabla \eta_h\|_I \|\nabla \eta_h^*\|_I + C \|\eta_h\|_I \|e_h\|_I. \tag{5.6}$$

Due to the definition of π_h, well-known a priori estimates for the spatial L^2-projection Π_h can be employed to directly obtain estimates for η_h and η_h^*. We have

$$\|\eta_h\|_I \leq Ch^{s+1}\|\nabla^{s+1}u_k\|_I, \quad \|\nabla\eta_h\|_I \leq Ch^s\|\nabla^{s+1}u_k\|_I, \quad \|\nabla\eta_h^*\|_I \leq Ch\|\nabla^2\tilde{z}_k\|_I.$$

These estimates applied to (5.6) lead to

$$\|e_h\|_I^2 \leq Ch^{s+1}\|\nabla^{s+1}u_k\|_I\big\{\|\nabla^2\tilde{z}_k\|_I + \|e_h\|_I\big\}.$$

Due to the fact that the domain Ω is polygonal and convex, elliptic regularity theory yields

$$\|\nabla^2 z_k\|_I \leq C\|\Delta z_k\|_I,$$

and we obtain the stated result by the stability estimate from Corollary 4.2. □

6. Error analysis for the optimal control problem. In this section, we prove the main results of this article, namely, an estimate of the error between the solution $(\bar{q}, \bar{u})$ of the continuous optimal control problem (2.2) and the solution $(\bar{q}_\sigma, \bar{u}_\sigma)$ of the discretized problem (3.14).

Throughout this section, we will indicate the dependence of the state and the adjoint state on the specific control $q \in Q$ by the notation introduced in section 2 and section 3, that is, $u(q)$, $z(q)$ on the continuous level, $u_k(q), z_k(q)$ on the semidiscrete and $u_{kh}(q), z_{kh}(q)$ on the discrete level.

6.1. Error in the control variable. In this section we analyze the error with respect to the control variable and prove the following result.

THEOREM 6.1. *The error between the solution $\bar{q} \in Q$ of the continuous optimization problem* (2.2) *and the solution $\bar{q}_\sigma \in Q_d$ of the discrete optimization problem* (3.14) *can be estimated as*

$$\begin{aligned}\|\bar{q} - \bar{q}_\sigma\|_I \leq \frac{C}{\alpha}k^{r+1}\big\{\|\partial_t^{r+1}u(\bar{q})\|_I + \|\partial_t^{r+1}z(\bar{q})\|_I\big\} \\ + \frac{C}{\alpha}h^{s+1}\big\{\|\nabla^{s+1}u_k(\bar{q})\|_I + \|\nabla^{s+1}z_k(\bar{q})\|_I\big\} + \Big(2 + \frac{C}{\alpha}\Big)\inf_{p_d \in Q_d}\|\hat{q} - p_d\|_I,\end{aligned}$$

where $\hat{q} \in Q$ can be chosen either as the continuous solution $\bar{q}$ or as the solution $\bar{q}_{kh}$ of the purely state discretized problem (3.10)*. The constants C are independent of the mesh size h, the size of the time steps k, and the choice of the discrete control space $Q_d \subset Q$.*

We first discuss the infimum term appearing on the right-hand side of the error estimate above. Thereby, we make use of the two possible formulations of this term for $\hat{q} = \bar{q}$ or $\hat{q} = \bar{q}_{kh}$: From the optimality conditions (3.11) for the optimal control problem (3.10) obtained after the discretization of the state equation in space and time, we get

$$(\bar{q}_{kh}, \delta q)_I = \frac{1}{\alpha}(z_{kh}(\bar{q}_{kh}), \delta q)_I \quad \forall \delta q \in Q,$$

and therefore $\bar{q}_{kh} = \frac{1}{\alpha}z_{kh}(\bar{q}_{kh}) \in X_{k,h}^{r,s} \subset Q$. Thus, if Q_d is chosen such that $Q_d \supset X_{k,h}^{r,s}$, the term

$$\inf_{p_d \in Q_d}\|\bar{q}_{kh} - p_d\|_I$$

vanishes. In this case, the solution $\bar{q}_\sigma$ of the fully discretized optimal control problem (3.14) coincides with the solution $\bar{q}_{kh}$; cf. [16]. Consequently, it is reasonable to discretize the control at most as fine as the adjoint state. The same conclusion can be drawn by inspection of the a posteriori error estimates developed in [21].

If the discrete control space Q_d does not fulfill the condition $Q_d \supset X_{k,h}^{r,s}$, it is desirable to choose $\hat{q} = \bar{q}$ in the above theorem to obtain an estimate for the infimum term. For both choices of the space Q_d described in section 3.3 we obtain the following estimate using interpolation theory:

$$\inf_{p_d \in Q_d} \|\bar{q} - p_d\|_I \leq Ck^{r_d+1}\|\partial_t^{r_d+1}\bar{q}\|_I + Ch_d^{s_d+1}\|\nabla^{s_d+1}\bar{q}\|_I.$$

Here, h_d is the discretization parameter corresponding to the spatial mesh employed for the control discretization.

The proof of Theorem 6.1 makes use of the assertions of the following lemmas and will be given at the end of this section.

Lemma 6.2. *Let $q \in Q$ be a given control. The error between the continuous state $u = u(q) \in X$ determined by* (2.1) *and the discrete state $u_{kh} = u_{kh}(q) \in X_{k,h}^{r,s}$ determined by* (3.9) *can be estimated as*

$$\|u(q) - u_{kh}(q)\|_I \leq Ck^{r+1}\|\partial_t^{r+1}u(q)\|_I + Ch^{s+1}\|\nabla^{s+1}u_k(q)\|_I.$$

For the error between the continuous adjoint state $z = z(q) \in X$ determined by (2.5) *and the discrete adjoint state $z_{kh} = z_{kh}(q) \in X_{k,h}^{r,s}$ determined by* (3.13), *the following estimate holds:*

$$\|z(q) - z_{kh}(q)\|_I \leq Ck^{r+1}\big\{\|\partial_t^{r+1}u(q)\|_I + \|\partial_t^{r+1}z(q)\|_I\big\} + Ch^{s+1}\big\{\|\nabla^{s+1}u_k(q)\|_I + \|\nabla^{s+1}z_k(q)\|_I\big\}.$$

Proof. The estimate for the error in terms of the state variable is immediately obtained by Theorems 5.1 and 5.5 since for $q \in Q$ the right-hand side $f + q$ of the state equation (2.1) is in $L^2(I, H)$ and thus fulfills the assumptions of Proposition 2.1.

For estimating the error in z, we introduce additionally the solutions $\tilde{z}_k \in X_k^r$ and $\tilde{z}_{kh} \in X_{k,h}^{r,s}$ which solve

$$B(\varphi, \tilde{z}_k) = (\varphi, u(q) - \hat{u})_I \qquad \forall \varphi \in X_k^r \qquad \text{and}$$
$$B(\varphi, \tilde{z}_{kh}) = (\varphi, u_k(q) - \hat{u})_I \qquad \forall \varphi \in X_{k,h}^{r,s}.$$

Since the adjoint solution $z(q) \in X$ is determined by (2.5), we may apply Theorem 5.1 to obtain

$$\|z(q) - \tilde{z}_k\|_I \leq Ck^{r+1}\|\partial_t^{r+1}z(q)\|_I.$$

Correspondingly, due to the definition of the semidiscrete adjoint solution $z_k(q) \in X_k^r$ by (3.7), Theorem 5.5 yields the estimate

$$\|z_k(q) - \tilde{z}_{kh}\|_I \leq Ch^{s+1}\|\nabla^{s+1}z_k(q)\|_I.$$

Using (3.7) for $z_k(q)$, we obtain that the difference $\tilde{z}_k - z_k(q)$ solves

$$B(\varphi, \tilde{z}_k - z_k(q)) = (\varphi, u(q) - u_k(q))_I \quad \forall \varphi \in X_k^r.$$

Then, the stability estimate from Corollary 4.2 yields

$$\|\tilde{z}_k - z_k(q)\|_I \le C\|u(q) - u_k(q)\|_I.$$

Similarly, using (3.13) for $z_{kh}(q)$, we obtain for the difference $\tilde{z}_{kh} - z_{kh}(q)$ the identity

$$B(\varphi, \tilde{z}_{kh} - z_{kh}(q)) = (\varphi, u_k(q) - u_{kh}(q))_I \quad \forall \varphi \in X_{k,h}^{r,s},$$

and the stability estimate from Corollary 4.7 implies

$$\|\tilde{z}_{kh} - z_{kh}(q)\|_I \le C\|u_k(q) - u_{kh}(q)\|_I.$$

Finally, the triangle inequality and the error estimates from Theorems 5.1 and 5.5 for the error in the state variable lead to the proposed result. □

LEMMA 6.3. *For given controls q, $r \in Q$, the difference between the derivatives of the continuous reduced functional j and the discrete reduced functional j_{kh} can be estimated by*

$$|j'(q)(r) - j'_{kh}(q)(r)| \le \|z(q) - z_{kh}(q)\|_I \|r\|_I.$$

Proof. The representations (2.6) and (3.12) for j' and j'_{kh}, respectively, imply directly the assertion

$$|j'(q)(r) - j'_{kh}(q)(r)| = |(z(q) - z_{kh}(q), r)_I| \le \|z(q) - z_{kh}(q)\|_I \|r\|_I. \qquad \square$$

LEMMA 6.4. *The derivatives of the discrete reduced functional j_{kh} are Lipschitz continuous on Q. That is, for arbitrary p, q, $r \in Q$, the estimate*

$$|j'_{kh}(q)(r) - j'_{kh}(p)(r)| \le (C + \alpha)\|q - p\|_I \|r\|_I$$

holds true.

Proof. By means of (3.12), we have

$$\begin{aligned} |j'_{kh}(q)(r) - j'_{kh}(p)(r)| &\le \alpha|(q - p, r)_I| + |(z_{kh}(q) - z_{kh}(p), r)_I| \\ &\le \alpha\|q - p\|_I \|r\|_I + \|z_{kh}(q) - z_{kh}(p)\|_I \|r\|_I. \end{aligned}$$

Since $z_{kh}(q) - z_{kh}(p)$ solves

$$B(\varphi, z_{kh}(q) - z_{kh}(p)) = (\varphi, u_{kh}(q) - u_{kh}(p))_I \quad \forall \varphi \in X_{k,h}^{r,s},$$

and $u_{kh}(q) - u_{kh}(p)$ satisfies

$$B(u_{kh}(q) - u_{kh}(p), \varphi) = (q - p, \varphi)_I \quad \forall \varphi \in X_{k,h}^{r,s},$$

the stability estimates for z_{kh} from Corollary 4.7 and for u_{kh} from Theorem 4.6 yield

$$\|z_{kh}(q) - z_{kh}(p)\|_I \le C\|u_{kh}(q) - u_{kh}(p)\|_I \le C\|q - p\|_I,$$

which implies the desired result. □

With the aid of these preliminary results, we now prove Theorem 6.1.

Proof of Theorem 6.1. To obtain the asserted result, we split the error to be estimated in two different ways:

$$\|\bar{q} - \bar{q}_\sigma\|_I \le \|\bar{q} - p_d\|_I + \|p_d - \bar{q}_\sigma\|_I, \tag{6.1}$$

$$\|\bar{q} - \bar{q}_\sigma\|_I \le \|\bar{q} - \bar{q}_{kh}\|_I + \|\bar{q}_{kh} - p_d\|_I + \|p_d - \bar{q}_\sigma\|_I. \tag{6.2}$$

Here, p_d is an arbitrary element of Q_d and $\bar{q}$, $\bar{q}_{kh}$, and $\bar{q}_\sigma$ are the optimal solutions on the different levels of discretization.

Due to the linear-quadratic structure of the optimal control problem under consideration, we have for all $p, r \in Q$,

$$j''_{kh}(p)(r,r) \geq \alpha \|r\|_I^2,$$

and $j''_{kh}(p)$ does not depend on p. This implies, for arbitrary $p, p_d \in Q_d$,

$$\alpha\|p_d - \bar{q}_\sigma\|_I^2 \leq j''_{kh}(p)(p_d - \bar{q}_\sigma, p_d - \bar{q}_\sigma) = j'_{kh}(p_d)(p_d - \bar{q}_\sigma) - j'_{kh}(\bar{q}_\sigma)(p_d - \bar{q}_\sigma).$$

Since $\bar{q}$, $\bar{q}_{kh}$, and $\bar{q}_\sigma$ are the optimal solutions of the continuous, semidiscrete, and discrete optimization problems, we have by (2.4), (3.11), and (3.15),

$$j'_{kh}(\bar{q}_\sigma)(p_d - \bar{q}_\sigma) = j'_{kh}(\bar{q}_{kh})(p_d - \bar{q}_\sigma) = j'(\bar{q})(p_d - \bar{q}_\sigma) = 0.$$

Using these identities, we obtain for the separation (6.1) (which we use to prove the theorem in the case $\hat{q} = \bar{q}$) the estimate

$$\begin{aligned}\alpha\|p_d - \bar{q}_\sigma\|_I^2 &\leq j'_{kh}(p_d)(p_d - \bar{q}_\sigma) - j'(\bar{q})(p_d - \bar{q}_\sigma)\\ &= j'_{kh}(p_d)(p_d - \bar{q}_\sigma) - j'_{kh}(\bar{q})(p_d - \bar{q}_\sigma) + j'_{kh}(\bar{q})(p_d - \bar{q}_\sigma) - j'(\bar{q})(p_d - \bar{q}_\sigma).\end{aligned}$$

By means of Lemmas 6.3 and 6.4, we achieve

$$\alpha\|p_d - \bar{q}_\sigma\|_I^2 \leq (C+\alpha)\|p_d - \bar{q}\|_I \|p_d - \bar{q}_\sigma\|_I + \|z(\bar{q}) - z_{kh}(\bar{q})\|_I \|p_d - \bar{q}_\sigma\|_I.$$

Using (6.1) we get the estimate

$$\|\bar{q} - \bar{q}_\sigma\|_I \leq \frac{1}{\alpha}\|z(\bar{q}) - z_{kh}(\bar{q})\|_I + \left(2 + \frac{C}{\alpha}\right)\|\bar{q} - p_d\|_I. \tag{6.3}$$

To use separation (6.2) for proving the theorem in the case $\hat{q} = \bar{q}_{kh}$, we estimate alternatively by means of Lemma 6.4,

$$\alpha\|p_d - \bar{q}_\sigma\|_I^2 \leq j'_{kh}(p_d)(p_d - \bar{q}_\sigma) - j'_{kh}(\bar{q}_{kh})(p_d - \bar{q}_\sigma) \leq (C+\alpha)\|p_d - \bar{q}_{kh}\|_I \|p_d - \bar{q}_\sigma\|_I.$$

In the same manner as before, we can estimate $\|\bar{q} - \bar{q}_{kh}\|_I$ by Lemma 6.3 as

$$\begin{aligned}\alpha\|\bar{q} - \bar{q}_{kh}\|_I^2 &\leq j''_{kh}(p)(\bar{q} - \bar{q}_{kh}, \bar{q} - \bar{q}_{kh})\\ &= j'_{kh}(\bar{q})(\bar{q} - \bar{q}_{kh}) - j'_{kh}(\bar{q}_{kh})(\bar{q} - \bar{q}_{kh})\\ &= j'_{kh}(\bar{q})(\bar{q} - \bar{q}_{kh}) - j'(\bar{q})(\bar{q} - \bar{q}_{kh})\\ &\leq \|z(\bar{q}) - z_{kh}(\bar{q})\|_I \|\bar{q} - \bar{q}_{kh}\|_I.\end{aligned}$$

Then, the two latter estimates imply

$$\|\bar{q} - \bar{q}_\sigma\|_I \leq \frac{1}{\alpha}\|z(\bar{q}) - z_{kh}(\bar{q})\|_I + \left(2 + \frac{C}{\alpha}\right)\|\bar{q}_{kh} - p_d\|_I. \tag{6.4}$$

Finally, the inequalities (6.3) and (6.4) prove the assertion by means of the estimate for $\|z(\bar{q}) - z_{kh}(\bar{q})\|_I$ from Lemma 6.2. ∎

To concretize the result of Theorem 6.1, we consider the following choice of discretizations: The state space is discretized by the cG(1)dG(0) method, that is, we consider the case when $r = 0$ and $s = 1$. Using for simplicity the same triangulation of the spatial domain ($h_d = h$) and the same distribution of the time steps ($k_d = k$) as for the discretization of the state, we discuss the following two possibilities for the control discretization (cf. section 3.3):

1. cG(1)dG(0) *discretization, i.e., cellwise (bi-/tri-)linear in space and piecewise constant in time*: In this case the infimum term in the estimate in Theorem 6.1 vanishes, since $Q_d \supset X_{k,h}^{r,s}$; see the above discussion. Thus, Theorem 6.1 implies that the discretization error is of order

$$\|\bar{q} - \bar{q}_\sigma\|_I = \mathcal{O}(k + h^2).$$

2. dG(0)dG(0) *discretization, i.e., cellwise constant in space and piecewise constant in time:* In this case the infimum term of the error estimation from Theorem 6.1 has to be taken into account, leading to the discretization error of order

$$\|\bar{q} - \bar{q}_\sigma\|_I = \mathcal{O}(k + h).$$

Note that the regularity of the optimal solutions required for these estimates is ensured by Propositions 2.1 and 2.2 for the continuous solutions q, u, and z, and by Theorem 4.1 and Corollary 4.2 for the time-discrete solutions u_k and z_k. A numerical validation of these estimates will be given in section 7.

6.2. Error in the state and in the adjoint variable. In this subsection we prove error estimates for the state and adjoint state variables. That is, we consider the discretization errors

$$\|\bar{u} - \bar{u}_\sigma\|_I = \|u(\bar{q}) - u_{kh}(\bar{q}_\sigma)\|_I \qquad \text{and} \qquad \|\bar{z} - \bar{z}_\sigma\|_I = \|z(\bar{q}) - z_{kh}(\bar{q}_\sigma)\|_I.$$

By means of the stability estimates derived in section 4, one simply obtains the following result.

THEOREM 6.5. *Let $(\bar{q}, \bar{u})$ be the solution of the continuous optimal control problem* (2.2) *and $\bar{z} = z(\bar{q})$ be the corresponding adjoint state. Let, moreover, $(\bar{q}_\sigma, \bar{u}_\sigma)$ be the solution of the discrete optimal control problem* (3.14) *with the corresponding discrete adjoint state $\bar{z}_\sigma = z_{kh}(\bar{q}_\sigma)$. Then, the following estimates hold:*

$$\text{(i)} \quad \|\bar{u} - \bar{u}_\sigma\|_I \leq \|u(\bar{q}) - u_{kh}(\bar{q})\|_I + C\|\bar{q} - \bar{q}_\sigma\|_I,$$

$$\text{(ii)} \quad \|\bar{z} - \bar{z}_\sigma\|_I \leq \|z(\bar{q}) - z_{kh}(\bar{q})\|_I + C\|\bar{q} - \bar{q}_\sigma\|_I.$$

Proof. Using the fact that $\bar{u} = u(\bar{q})$ and $\bar{u}_\sigma = u_{kh}(\bar{q}_\sigma)$, we have

$$\|\bar{u} - \bar{u}_\sigma\|_I \leq \|u(\bar{q}) - u_{kh}(\bar{q})\|_I + \|u_{kh}(\bar{q}) - u_{kh}(\bar{q}_\sigma)\|_I. \tag{6.5}$$

By means of the stability result from Theorem 4.6, we obtain

$$\|u_{kh}(\bar{q}) - u_{kh}(\bar{q}_\sigma)\|_I \leq C\|\bar{q} - \bar{q}_\sigma\|_I.$$

This proves the first assertion. The second assertion follows in the same way utilizing the stability of the adjoint state given by Corollary 4.7. ∎

Employing the discretization of the control by cG(1)dG(0), the above theorem leads to the optimal order of convergence using Lemma 6.2 and Theorem 6.1. That is, we have

$$\|u(\bar{q}) - u_{kh}(\bar{q})\|_I = \mathcal{O}(k + h^2), \qquad \|z(\bar{q}) - z_{kh}(\bar{q})\|_I = \mathcal{O}(k + h^2),$$

$$\|\bar{q} - \bar{q}_\sigma\|_I = \mathcal{O}(k + h^2),$$

and thus

$$\|\bar{u} - \bar{u}_\sigma\|_I = \mathcal{O}(k + h^2) \quad \text{and} \quad \|\bar{z} - \bar{z}_\sigma\|_I = \mathcal{O}(k + h^2).$$

However, in the case of dG(0)dG(0) discretization, this theorem does not lead to the optimal order of convergence: In this case, we have indeed as before,

$$\|u(\bar{q}) - u_{kh}(\bar{q})\|_I = \mathcal{O}(k + h^2) \quad \text{and} \quad \|z(\bar{q}) - z_{kh}(\bar{q})\|_I = \mathcal{O}(k + h^2)$$

since the discretization of the state space is unaffected by the discretization of the controls, but we have only

$$\|\bar{q} - \bar{q}_\sigma\|_I = \mathcal{O}(k + h)$$

due to the first order discretization of the control space. This would lead to $\mathcal{O}(k + h)$ for the state and the adjoint variable.

Utilizing a more detailed analysis, we can prove also in this case the optimal order of convergence $\mathcal{O}(k + h^2)$ for the errors $\|\bar{u} - \bar{u}_\sigma\|_I$ and $\|\bar{z} - \bar{z}_\sigma\|_I$.

For both choices of the space Q_d described in section 3.3 the following results hold.

THEOREM 6.6. *Let $(\bar{q}, \bar{u})$ be the solution of the continuous optimal control problem* (2.2) *and $\bar{z} = z(\bar{q})$ be the corresponding adjoint state. Let, moreover, $(\bar{q}_\sigma, \bar{u}_\sigma)$ be the solution of the discrete optimal control problem* (3.14) *with the corresponding discrete adjoint state $\bar{z}_\sigma = z_{kh}(\bar{q}_\sigma)$. In addition we assume $r = r_d$, i.e., the same discretization of the state and control variable in time. Then, the following estimates hold:*

$$\text{(i)} \quad \|\bar{u} - \bar{u}_\sigma\|_I \leq \|u(\bar{q}) - u_{kh}(\bar{q})\|_I + Ch_d\Big(1 + \frac{1}{\alpha}\Big)\|\bar{q} - \pi_d\bar{q}\|_I + \frac{C}{\alpha}\|z(\bar{q}) - z_{kh}(\bar{q})\|_I,$$

$$\text{(ii)} \quad \|\bar{z} - \bar{z}_\sigma\|_I \leq Ch_d\Big(1 + \frac{1}{\alpha}\Big)\|\bar{q} - \pi_d\bar{q}\|_I + C\Big(1 + \frac{1}{\alpha}\Big)\|z(\bar{q}) - z_{kh}(\bar{q})\|_I,$$

where $\pi_d\colon Q \to Q_d$ is the space-time L^2-projection on Q_d.

Proof. For proving (i) we split the error $\|\bar{u} - \bar{u}_\sigma\|_I$ as follows:

$$\|\bar{u} - \bar{u}_\sigma\|_I \leq \|u(\bar{q}) - u_{kh}(\bar{q})\|_I + \|u_{kh}(\bar{q}) - u_{kh}(\pi_d\bar{q})\|_I + \|u_{kh}(\pi_d\bar{q}) - u_{kh}(\bar{q}_\sigma)\|_I. \tag{6.6}$$

The second term on the right-hand side of (6.6) is estimated using the following duality argument: Let $\tilde{z}_{kh} \in X_{k,h}^{r,s}$ be the solution of

$$B(\varphi, \tilde{z}_{kh}) = (\varphi, u_{kh}(\bar{q}) - u_{kh}(\pi_d\bar{q}))_I \quad \forall\varphi \in X_{k,h}^{r,s}.$$

By means of the discrete state equation (3.9) for $u_{kh}(\bar{q})$ and $u_{kh}(\pi_d\bar{q})$, we obtain

$$\|u_{kh}(\bar{q}) - u_{kh}(\pi_d\bar{q})\|_I^2 = B(u_{kh}(\bar{q}) - u_{kh}(\pi_d\bar{q}), \tilde{z}_{kh}) = (\bar{q} - \pi_d\bar{q}, \tilde{z}_{kh})_I.$$

Since π_d is the L^2-projection, we have

$$\|u_{kh}(\bar{q}) - u_{kh}(\pi_d\bar{q})\|_I^2 = (\bar{q} - \pi_d\bar{q}, \tilde{z}_{kh} - \pi_d\tilde{z}_{kh})_I \leq \|\bar{q} - \pi_d\bar{q}\|_I\|\tilde{z}_{kh} - \pi_d\tilde{z}_{kh}\|_I. \tag{6.7}$$

Using the fact that $r = r_d$ and that therefore the same time discretization is employed for the control and state variable, the space-time L^2-projection π_d applied to $\tilde{z}_{kh}$ can be expressed as spatial L^2-projection $\Pi_{h_d}\tilde{z}_{kh}$.

Applying an interpolation estimate and the stability estimate from Corollary 4.7 we obtain

$$\|\tilde{z}_{kh} - \pi_d \tilde{z}_{kh}\|_I = \|\tilde{z}_{kh} - \Pi_{h_d} \tilde{z}_{kh}\|_I \leq C h_d \|\nabla \tilde{z}_{kh}\|_I \leq C h_d \|u_{kh}(\bar{q}) - u_{kh}(\pi_d \bar{q})\|_I.$$

Plugging this estimate into (6.7) yields

$$\|u_{kh}(\bar{q}) - u_{kh}(\pi_d \bar{q})\|_I \leq C h_d \|\bar{q} - \pi_d \bar{q}\|_I. \tag{6.8}$$

For the third term in (6.6) we obtain, using Theorem 4.6,

$$\|u_{kh}(\pi_d \bar{q}) - u_{kh}(\bar{q}_\sigma)\|_I \leq C \|\pi_d \bar{q} - \bar{q}_\sigma\|_I.$$

For estimating the term $\|\pi_d \bar{q} - \bar{q}_\sigma\|_I$ we proceed as in the proof of Theorem 6.1 for the term $\|p_d - \bar{q}_\sigma\|_I$:

$$\alpha \|\pi_d q - q_\sigma\|_I^2 \leq j'_{kh}(\pi_d q)(\pi_d q - q_\sigma) - j'(q)(\pi_d q - q_\sigma).$$

Using representation (3.12) of j'_{kh} and (2.6) of j' we have

$$\alpha \|\pi_d \bar{q} - \bar{q}_\sigma\|_I^2 \leq \alpha(\pi_d \bar{q} - \bar{q}, \pi_d \bar{q} - \bar{q}_\sigma)_I + (z_{kh}(\pi_d \bar{q}) - z(\bar{q}), \pi_d \bar{q} - \bar{q}_\sigma)_I.$$

Since $\pi_d \bar{q} - \bar{q}_\sigma \in Q_d$, the term $(\pi_d \bar{q} - \bar{q}, \pi_d \bar{q} - \bar{q}_\sigma)_I$ vanishes, and due to Corollary 4.7, we end up with

$$\begin{aligned}
\alpha \|\pi_d \bar{q} - \bar{q}_\sigma\|_I &\leq \|z_{kh}(\pi_d \bar{q}) - z(\bar{q})\|_I \\
&\leq \|z_{kh}(\pi_d \bar{q}) - z_{kh}(\bar{q})\|_I + \|z_{kh}(\bar{q}) - z(\bar{q})\|_I \\
&\leq C \|u_{kh}(\pi_d \bar{q}) - u_{kh}(\bar{q})\|_I + \|z_{kh}(\bar{q}) - z(\bar{q})\|_I,
\end{aligned}$$

which implies, by using (6.8), the estimate

$$\begin{aligned}
\|u_{kh}(\pi_d \bar{q}) - u_{kh}(\bar{q}_\sigma)\|_I &\leq \frac{C}{\alpha} \|u_{kh}(\pi_d \bar{q}) - u_{kh}(\bar{q})\|_I + \frac{C}{\alpha} \|z_{kh}(\bar{q}) - z(\bar{q})\|_I \\
&\leq \frac{C}{\alpha} h_d \|\bar{q} - \pi_d \bar{q}\|_I + \frac{C}{\alpha} \|z_{kh}(\bar{q}) - z(\bar{q})\|_I.
\end{aligned} \tag{6.9}$$

Plugging (6.8) and (6.9) into (6.6) we complete the proof of (i). The assertion (ii) follows using (6.8), (6.9), and the following estimate exploiting the stability result from Corollary 4.7:

$$\begin{aligned}
&\|\bar{z} - \bar{z}_\sigma\|_I \\
&\quad \leq \|z(\bar{q}) - z_{kh}(\bar{q})\|_I + \|z_{kh}(\bar{q}) - z_{kh}(\pi_d \bar{q})\|_I + \|z_{kh}(\pi_d \bar{q}) - z_{kh}(\bar{q}_\sigma)\|_I \\
&\quad \leq \|z(\bar{q}) - z_{kh}(\bar{q})\|_I + C \|u_{kh}(\bar{q}) - u_{kh}(\pi_d \bar{q})\|_I + C \|u_{kh}(\pi_d \bar{q}) - u_{kh}(\bar{q}_\sigma)\|_I. \qquad \square
\end{aligned}$$

For the case of dG(0)dG(0) discretization of the control space with $h_d = h$ and $k_d = k$ this theorem leads to the improved (optimal) order of convergence

$$\|\bar{u} - \bar{u}_\sigma\|_I = \mathcal{O}(k + h^2) \quad \text{and} \quad \|\bar{z} - \bar{z}_\sigma\|_I = \mathcal{O}(k + h^2).$$

7. Numerical results. In this section, we are going to validate the a priori error estimates for the error in the control, state, and adjoint state numerically. To this end, we consider the following concretion of the model problem (2.2) with known analytical exact solution on $\Omega \times I = (0,1)^2 \times (0, 0.1)$ and homogeneous Dirichlet boundary conditions. The right-hand side f, the desired state $\hat{u}$, and the initial condition u_0 are given in terms of the eigenfunctions

$$w_a(t, x_1, x_2) \coloneqq \exp(a\pi^2 t) \sin(\pi x_1) \sin(\pi x_2), \quad a \in \mathbb{R},$$

of the operator $\pm\partial_t - \Delta$ as

$$\begin{aligned} f(t, x_1, x_2) &\coloneqq -\pi^4 w_a(T, x_1, x_2), \\ \hat{u}(t, x_1, x_2) &\coloneqq \frac{a^2 - 5}{2 + a} \pi^2 w_a(t, x_1, x_2) + 2\pi^2 w_a(T, x_1, x_2), \\ u_0(x_1, x_2) &\coloneqq \frac{-1}{2 + a} \pi^2 w_a(0, x_1, x_2). \end{aligned}$$

For this choice of data and with the regularization parameter α chosen as $\alpha = \pi^{-4}$, the optimal solution triple $(\bar{q}, \bar{u}, \bar{z})$ of the optimal control problem (2.2) is given by

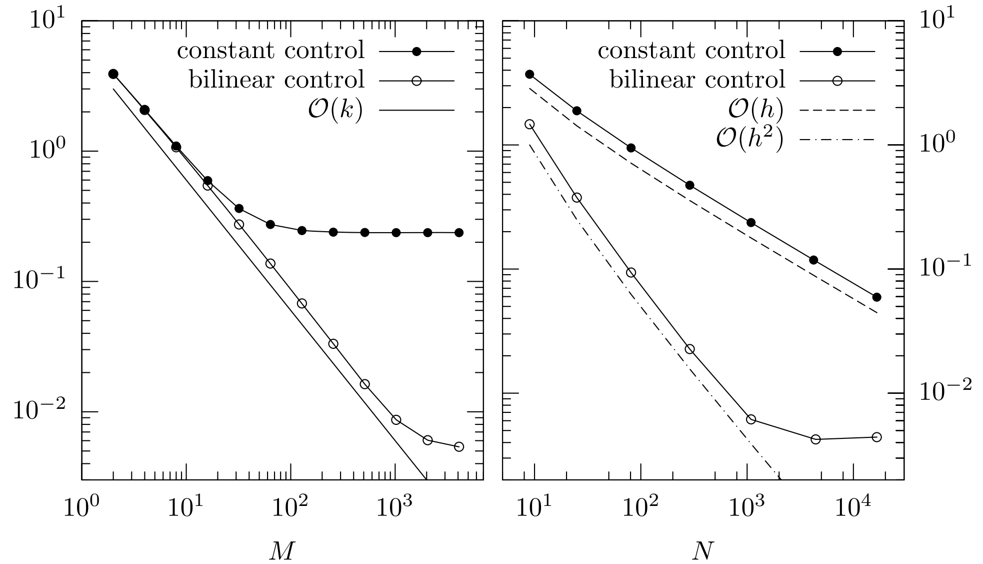
$$\begin{aligned} \bar{q}(t, x_1, x_2) &\coloneqq -\pi^4 \{w_a(t, x_1, x_2) - w_a(T, x_1, x_2)\}, \\ \bar{u}(t, x_1, x_2) &\coloneqq \frac{-1}{2 + a} \pi^2 w_a(t, x_1, x_2), \\ \bar{z}(t, x_1, x_2) &\coloneqq w_a(t, x_1, x_2) - w_a(T, x_1, x_2). \end{aligned}$$

We are going to validate the estimates developed in the previous section by separating the discretization errors. That is, we consider at first the behavior of the error for a sequence of discretizations with decreasing size of the time steps and a fixed spatial triangulation with $N = 1089$ nodes. Second, we examine the behavior of the error under refinement of the spatial triangulation for $M = 2048$ time steps.

The state discretization is chosen as cG(1)dG(0), i.e., $r = 0$, $s = 1$. For the control discretization we use the same temporal and spatial meshes as for the state variable and present the result for two choices of the discrete control space Q_d: cG(1)dG(0) and dG(0)dG(0). For the following computations, we choose the free parameter a to be $-\sqrt{5}$. For this choice the right-hand side f and the desired state $\hat{u}$ do not depend on time which avoids side effects introduced by numerical quadrature.

The optimal control problems are solved by the optimization library RoDoBo [23] and the finite element toolkit Gascoigne [14] using a conjugate gradient method applied to the reduced problem (3.14).

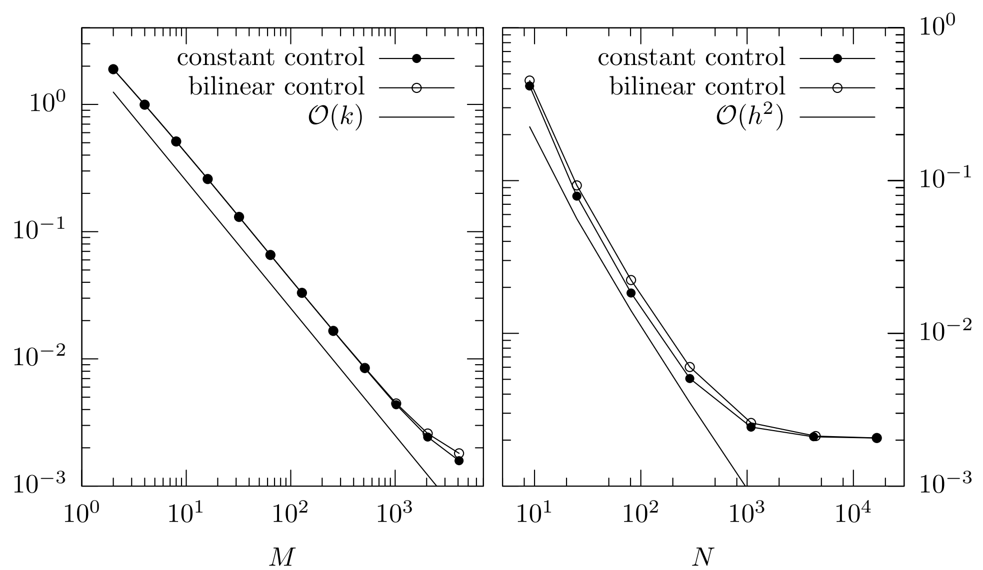
Figure 7.1(a) depicts the development of the error under refinement of the temporal step size k. Up to the spatial discretization error it exhibits the proven convergence order $\mathcal{O}(k)$ for both kinds of spatial discretization of the control space. For piecewise constant control (dG(0)dG(0) discretization), the discretization error is already reached at 128 time steps, whereas in the case of bilinear control (cG(1)dG(0) discretization), the number of time steps could be increased up to $M = 4096$ until reaching the spatial accuracy.



(a) Refinement of the time steps for $N = 1089$ spatial nodes

(b) Refinement of the spatial triangulation for $M = 2048$ time steps

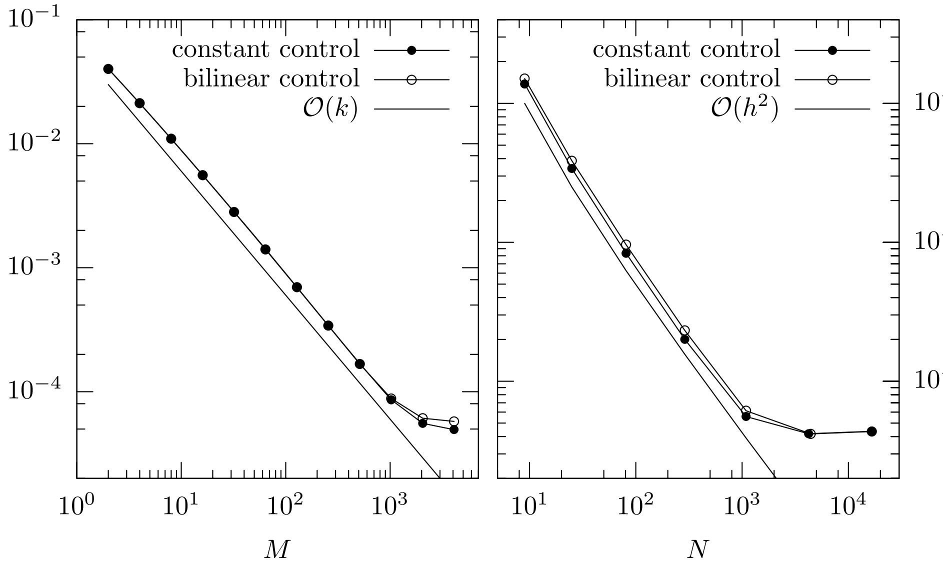
FIG. 7.1. *Discretization error* $\|\bar{q} - \bar{q}_\sigma\|_I$.



(a) Refinement of the time steps for $N = 1089$ spatial nodes

(b) Refinement of the spatial triangulation for $M = 2048$ time steps

FIG. 7.2. *Discretization error* $\|\bar{u} - \bar{u}_\sigma\|_I$.



(a) Refinement of the time steps for $N = 1089$ spatial nodes

(b) Refinement of the spatial triangulation for $M = 2048$ time steps

FIG. 7.3. *Discretization error* $\|\bar{z} - \bar{z}_\sigma\|_I$.

In Figure 7.1(b) the development of the error in the control variable under spatial refinement is shown. The expected order $\mathcal{O}(h)$ for piecewise constant control (dG(0)dG(0) discretization) and $\mathcal{O}(h^2)$ for bilinear control (cG(1)dG(0) discretization) is observed.

Figures 7.2 and 7.3 show the errors in the state and the adjoint variables, $\|\bar{u}-\bar{u}_\sigma\|_I$ and $\|\bar{z} - \bar{z}_\sigma\|_I$, for separate refinement of the time and space discretization. Thereby, we observe convergence of order $\mathcal{O}(k + h^2)$ regardless of the type of spatial discretization used for the controls. This is consistent with the results proven in the previous section.

REFERENCES

[1] N. Arada, E. Casas, and F. Tröltzsch, *Error estimates for a semilinear elliptic optimal control problem*, Comput. Optim. Appl., 23 (2002), pp. 201–229.

[2] R. Becker, D. Meidner, and B. Vexler, *Efficient numerical solution of parabolic optimization problems by finite element methods*, Optim. Methods Softw., 22 (2007), pp. 813–833.

[3] J. H. Bramble and S. R. Hilbert, *Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation*, SIAM J. Numer. Anal., 7 (1970), pp. 112–124.

[4] E. Casas, M. Mateos, and F. Tröltzsch, *Error estimates for the numerical approximation of boundary semilinear elliptic control problems*, Comput. Optim. Appl., 31 (2005), pp. 193–220.

[5] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics Appl. Math. 40, SIAM, Philadelphia, 2002.

[6] R. Dautray and J.-L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology: Evolution Problems* I, Vol. 5, Springer-Verlag, Berlin, 1992.

[7] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson, *Computational Differential Equations*, Cambridge University Press, Cambridge, UK, 1996.

[8] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems* I: *A linear model problem*, SIAM J. Numer. Anal., 28 (1991), pp. 43–77.
[9] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems* II: *Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., 32 (1995), pp. 706–740.
[10] K. Eriksson, C. Johnson, and V. Thomée, *Time discretization of parabolic problems by the discontinuous Galerkin method*, M2AN Math. Model. Numer. Anal., 19 (1985), pp. 611–643.
[11] L. C. Evans, *Partial Differential Equations*, Grad. Stud. Math. 19, AMS, Providence, 2002.
[12] R. Falk, *Approximation of a class of optimal control problems with order of convergence estimates*, J. Math. Anal. Appl., 44 (1973), pp. 28–47.
[13] M. Feistauer and K. Švadlenka, *Space-Time Discontinuous Galerkin Method for Solving Nonstationary Convection-Diffusion-Reaction Problems*, Preprint MATH-knm-2005/2, Charles University, Prague, 2005.
[14] *GASCOIGNE: The Finite Element Toolkit,* http://www.gascoigne.uni-hd.de/.
[15] T. Geveci, *On the approximation of the solution of an optimal control problem governed by an elliptic equation*, M2AN Math. Model. Numer. Anal., 13 (1979), pp. 313–328.
[16] M. Hinze, *A variational discretization concept in control constrained optimization: The linear-quadratic case*, Comput. Optim. Appl., 30 (2005), pp. 45–61.
[17] I. Lasiecka and K. Malanowski, *On discrete-time Ritz-Galerkin approximation of control constrained optimal control problems for parabolic systems*, Control Cybern., 7 (1978), pp. 21–36.
[18] J.-L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Grundlehren Math. Wiss. 170, Springer, Berlin, 1971.
[19] K. Malanowski, *Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal-control problems*, Appl. Math. Optim., 8 (1981), pp. 69–95.
[20] R. S. McNight and W. E. Bosarge, Jr., *The Ritz–Galerkin procedure for parabolic control problems*, SIAM J. Control Optim., 11 (1973), pp. 510–524.
[21] D. Meidner and B. Vexler, *Adaptive space-time finite element methods for parabolic optimization problems*, SIAM J. Control Optim., 46 (2007), pp. 116–142.
[22] C. Meyer and A. Rösch, *Superconvergence properties of optimal control problems*, SIAM J. Control Optim., 43 (2004), pp. 970–985.
[23] *RoDoBo: A C++ Library for Optimization with Stationary and Nonstationary PDEs,* http://rodobo.uni-hd.de/.
[24] A. Rösch, *Error estimates for parabolic optimal control problems with control constraints*, Z. Anal. Anwendungen, 23 (2004), pp. 353–376.
[25] M. Schmich and B. Vexler, *Adaptivity with dynamic meshes for space-time finite element discretizations of parabolic equations*, SIAM J. Sci. Comput., 30 (2008), pp. 369–393.
[26] D. Schötzau, *hp-DGFEM for Parabolic Evolution Problems*, Ph.D. thesis, Swiss Federal Institute of Technology, Zürich, 1999.
[27] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer Ser. Comput. Math. 25, Springer-Verlag, Berlin, 2006.
[28] R. Winther, *Error estimates for a Galerkin approximation of a parabolic control problem*, Ann. Math. Pura Appl. (4), 117 (1978), pp. 173–206.
[29] J. Wloka, *Partielle Differentialgleichungen: Sobolevräume und Randwertaufgaben*, B. G. Teubner, Stuttgart, 1982.