A Priori Error Analysis for Space-Time Finite Element Discretization of Parabolic Optimal Control Problems

D. Meidner and B. Vexler

Abstract In this article we discuss a priori error estimates for Galerkin finite element discretizations of optimal control problems governed by linear parabolic equations and subject to inequality control constraints. 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 taking into account the spatial and the temporal regularity of the optimal solution. For the treatment of the control discretization we discuss different approaches extending techniques known from the elliptic case. For detailed proofs and numerical results we refer to [18, 19].

1 Introduction

A priori error analysis for finite element discretizations of optimization problems governed by partial differential equations is an active area of research. While many publications are concerned with elliptic problems, see, e.g., [1, 7, 11, 12, 13, 20], there are only few published results on this topic for parabolic problems, see [14, 16, 17, 21, 23]. In this paper we consider an optimal control problem governed by the heat equation. For the discretization of the state equation we employ a space-time finite element discretization and discuss different approaches for the discretization of the control variable. Extending techniques for treating inequality constraints on the control variable known for elliptic problems, we derive a priori error estimates

Dominik Meidner
Lehrstuhl für Mathematische Optimierung, Technische Universität München, Fakultät für Mathematik, Boltzmannstraße 3, Garching b. München, Germany, e-mail: meidner@ma.tum.de

Boris Vexler
Lehrstuhl für Mathematische Optimierung, Technische Universität München, Fakultät für Mathematik, Boltzmannstraße 3, Garching b. München, Germany, e-mail: vexler@ma.tum.de

taking into account spatial and temporal regularity of the solutions. The detailed proofs and numerical results illustrating our considerations are presented in [18, 19].

For a convex, polygonal spatial domain $\Omega \subset \mathbb{R}^n$ ($n = 2,3$) and a time interval $I = (0,T)$ we consider the state space X and the control space Q given as

$$X := \left\{ v \,\middle|\, v \in L^2(I,V) \text{ and } \partial_t v \in L^2(I,V^*) \right\}, \quad V = H_0^1(\Omega), \quad Q = L^2(I,L^2(\Omega)).$$

The inner product of $L^2(\Omega)$ and the corresponding norm are denoted by $(\cdot,\cdot)$ and $\|\cdot\|$. The inner product and the norm of Q are denoted by $(\cdot,\cdot)_I$ and $\|\cdot\|_I$. The state equation in a weak form for the state variable $u = u(q) \in X$, the control variable $q \in Q$ and the initial condition $u_0 \in V$ reads:

$$(\partial_t u, \varphi)_I + (\nabla u, \nabla \varphi)_I + (u(0), \varphi(0)) = (f + q, \varphi)_I + (u_0, \varphi(0)) \quad \forall \varphi \in X. \tag{1}$$

The optimal control problem under consideration is formulated as follows:

$$\text{Minimize } J(q,u) := \frac{1}{2}\|u - \hat{u}\|_I^2 + \frac{\alpha}{2}\|q\|_I^2 \text{ subject to (1) and } (q,u) \in Q_{\text{ad}} \times X, \tag{2}$$

where $\hat{u} \in L^2(I,L^2(\Omega))$ is a given desired state and $\alpha > 0$ is the regularization parameter. The admissible set Q_{ad} describes box constraints on the control variable:

$$Q_{\text{ad}} := \{\, q \in Q \mid q_a \le q(t,x) \le q_b \quad \text{a.e. in } I \times \Omega \,\}.$$

Throughout we consider two situations: the problem without control constraints, i.e., $q_a = -\infty$, $q_b = +\infty$, $Q_{\text{ad}} = Q$, and the problem with control constraints, i.e., $q_a, q_b \in \mathbb{R}$ with $q_a < q_b$. In both cases the optimal control problem (2) is known to possess a unique solution $(\bar{q}, \bar{u})$, see, e.g., [15]. This solution is characterized by the optimality system involving an adjoint equation for the adjoint variable $z = z(q) \in X$ satisfying

$$-(\varphi, \partial_t z)_I + (\nabla \varphi, \nabla z)_I + (z(T), \varphi(T)) = (\varphi, u(q) - \hat{u})_I \quad \forall \varphi \in X. \tag{3}$$

The optimal solution $(\bar{q}, \bar{u})$ and the corresponding adjoint state $\bar{z} = z(\bar{q})$ exhibit for any $p < \infty$ when $n = 2$ and $p \le 6$ when $n = 3$ the following regularity properties (see [19, Proposition 2.3]):

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

For the time discretization we use the lowest order discontinuous Galerkin method dG(0), see [9, 10], which is a variant of the implicit Euler scheme. We exploit a time partitioning

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

with corresponding time intervals $I_m := (t_{m-1}, t_m]$ of length k_m and a semidiscrete space

$$X_k^0 := \left\{ v_k \in L^2(I,V) \;\middle|\; v_k\big|_{I_m} \in \mathscr{P}_0(I_m,V),\ m = 1,2,\dots,M \right\}.$$

Here, the discretization parameter k is defined as the maximum of all step sizes k_m. The dG(0) semidiscretization of the state equation (1) for given control $q \in Q$ reads: Find a state $u_k = u_k(q) \in X_k^0$ such that

$$B(u_k,\varphi) = (f+q,\varphi)_I + (u_0,\varphi_0^+) \quad \forall \varphi \in X_k^0, \tag{4}$$

where the bilinear form $B(\cdot,\cdot)$ is defined as

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

with $[u_k]_m$ denoting the jump of the function u_k at t_m. The semidiscrete optimization problem for the dG(0) time discretization has the form:

$$\text{Minimize } J(q_k,u_k) \text{ subject to (4) and } (q_k,u_k) \in Q_{\text{ad}} \times X_k^0. \tag{5}$$

The unique solution of this problem is denoted by $(\overline{q}_k,\overline{u}_k)$ and is characterized using a semidiscrete adjoint solution $z_k = z_k(q) \in X_k^0$ determined by:

$$B(\varphi,z_k) = (\varphi,u_k(q)-\hat{u})_I \quad \forall \varphi \in X_k^0. \tag{6}$$

By stability estimates (see [18, Theorems 4.1, 4.3, Corollaries 4.2, 4.5]) we deduce that the semidiscrete optimal state $\overline{u}_k$ and adjoint state $\overline{z}_k = z_k(\overline{q})$ have the regularity

$$\overline{u}_k\big|_{I_m}, \overline{z}_k\big|_{I_m} \in L^2(I_m,H^2(\Omega)) \cap H^1(I_m,L^2(\Omega)),\ m = 1,2,\dots,M$$

uniformly in k.

For the spatial discretization we use a usual conforming finite element space $V_h \subset V$ consisting of cellwise (bi-/tri-)linear functions over a quasi-uniform mesh $\mathscr{T}_h$, see, e.g., [3] for standard definitions. Then, the so called cG(1)dG(0) discretization of the state equation for given control $q \in Q$ has the form: Find a state $u_{kh} = u_{kh}(q) \in X_{k,h}^{0,1}$ such that

$$B(u_{kh},\varphi) = (f+q,\varphi)_I + (u_0,\varphi_0^+) \quad \forall \varphi \in X_{k,h}^{0,1}, \tag{7}$$

where $X_{k,h}^{0,1}$ is defined similar to the semidiscrete space X_k^0 as

$$X_{k,h}^{0,1} := \left\{ v_{kh} \in L^2(I,V_h) \;\middle|\; v_{kh}\big|_{I_m} \in \mathscr{P}_0(I_m,V_h),\ m = 1,2,\dots,M \right\} \subset X_k^0.$$

The corresponding optimal control problem is given as

$$\text{Minimize } J(q_{kh},u_{kh}) \text{ subject to (7) and } (q_{kh},u_{kh}) \in Q_{\text{ad}} \times X_{k,h}^{0,1}. \tag{8}$$

As above, the unique solution of this problem is denoted by $(\bar{q}_{kh}, \bar{u}_{kh})$ and can be characterized using the corresponding adjoint solution $z_{kh} = z_{kh}(q) \in X_{k,h}^{0,1}$ determined by

$$B(\varphi, z_{kh}) = (\varphi, u_{kh}(q) - \hat{u})_I \quad \forall \varphi \in X_{k,h}^{0,1}. \tag{9}$$

To obtain a fully discrete optimal control problem we consider a subspace $Q_d \subset Q$ of the control space and come up with the following problem:

$$\text{Minimize } J(q_\sigma, u_\sigma) \text{ subject to (7) and } (q_\sigma, u_\sigma) \in Q_{d,\text{ad}} \times X_{k,h}^{0,1}, \tag{10}$$

where $Q_{d,\text{ad}} = Q_d \cap Q_{\text{ad}}$. The solution of this problem is denoted by $(\bar{q}_\sigma, \bar{u}_\sigma)$ where the discretization parameter σ consists of all discretization parameters, i.e., $\sigma = (k, h, d)$. In the following sections we will consider different choices of the discrete control space Q and provide error estimates for the error $\|\bar{q} - \bar{q}_\sigma\|_I$.

2 Error analysis for problems without control constraints

In this section, we investigate the optimal control problem (2) in the situation when no constraints are imposed on the control, i.e., in the case of $Q_{\text{ad}} = Q$. The control space Q is discretized in time as the state space X, i.e., by piecewise constant polynomials on the subintervals I_m. The space discretization of Q is done either by piecewise (bi-/tri-)linear finite elements (as employed also for discretizing the state space) or by cellwise constant polynomials.

For this combination of state and control discretizations, the following estimate for the error in the control variable holds (cf. [18, Theorem 6.1]):

Theorem 1. *The error between the the solution $\bar{q} \in Q$ of the continuous optimization problem* (2) *and the solution $\bar{q}_\sigma \in Q_d$ of the discrete optimization problem* (10) *can be estimated as*

$$\begin{aligned}\|\bar{q} - \bar{q}_\sigma\|_I \le \frac{C}{\alpha} k \big\{ \|\partial_t u(\bar{q})\|_I + \|\partial_t z(\bar{q})\|_I \big\} \\ + \frac{C}{\alpha} h^2 \big\{ \|\nabla^2 u_k(\bar{q})\|_I + \|\nabla^2 z_k(\bar{q})\|_I \big\} + \left(2 + \frac{C}{\alpha}\right) \inf_{p_d \in Q_d} \|\bar{q} - p_d\|_I.\end{aligned}$$

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 $Q_d \subset Q$.

By applying standard interpolation estimates to the infimum term we obtain the optimal asymptotic orders of convergence of $\|\bar{q} - \bar{q}_\sigma\|_I = \mathcal{O}(k + h^2)$ for the piecewise (bi-/tri-)linear space discretization of the control space Q and $\|\bar{q} - \bar{q}_\sigma\|_I = \mathcal{O}(k + h)$ for the cellwise constant space discretization of the controls.

3 Error analysis for problems with control constraints

In this section, we provide estimates of the error in terms of the control variable for the constrained optimal control problem (2) with different choices of the discrete control space Q_d.

3.1 Cellwise constant discretization

Like in the unconstrained case, the controls are discretized with respect to time by piecewise constant polynomials. The space discretization is done here by cellwise constant polynomials. For proving the desired error estimate, we extend the techniques presented in [8] for elliptic problems to the case of parabolic optimal control problems. This demands the introduction of the solution $\bar{q}_d$ of the purely control discretized problem:

$$\text{Minimize } J(q_d, u_d) \text{ subject to (1) and } (q_d, u_d) \in Q_{d,\text{ad}} \times X\,. \tag{11}$$

By means of this problem, the following estimate for the error in the control variable can be proven (cf. [19, Corollary 5.3]):

Theorem 2. *Let* $\bar{q} \in Q_{ad}$ *be the solution of the optimal control problem* (2), $\bar{q}_\sigma \in Q_{d,ad}$ *be the solution of the discretized problem* (10)*, where the cellwise constant discretization for the control variable is employed. Let moreover* $\bar{q}_d \in Q_{d,ad}$ *be the solution of the purely control discretized problem* (11)*. Then the following estimate holds:*

$$\begin{aligned}\|\bar{q} - \bar{q}_\sigma\|_I \leq \frac{C}{\alpha} k \left\{ \|\partial_t \bar{q}\|_I + \|\partial_t u(\bar{q}_d)\|_I + \|\partial_t z(\bar{q}_d)\|_I \right\} \\ + \frac{C}{\alpha} h \left\{ \|\nabla \bar{q}\|_I + \|\nabla z(\bar{q}_d)\|_I + h \left(\|\nabla^2 u_k(\bar{q}_d)\|_I + \|\nabla^2 z_k(\bar{q}_d)\|_I \right) \right\} = \mathcal{O}(k+h).\end{aligned}$$

We note, that all terms in the above estimate which depend on discretization parameters k and d are uniformly bounded with respect to these parameters. Therefore the constant in $\mathcal{O}(k+h)$ is independent of all discretization parameters and depends only on problem data.

3.2 Cellwise linear discretization

A better convergence result for the error can be achieved by discretizing the controls with piecewise (bi-/tri-)linear finite elements in space instead of using only piecewise constant trial functions. In this section, we treat this discretization combined with the already known time discretization by piecewise constant polynomials. The desired estimate is obtained by adapting the techniques described in [4, 6] to parabolic problems.

The analysis for this configuration is based on an assumption on the structure of the active sets. For each time interval I_m we group the cells K of the mesh $\mathscr{T}_h$ depending on the value of $\bar{q}_k$ on K into three sets $\mathscr{T}_h = \mathscr{T}_{h,m}^1 \cup \mathscr{T}_{h,m}^2 \cup \mathscr{T}_{h,m}^3$ with $\mathscr{T}_{h,m}^i \cap \mathscr{T}_{h,m}^j = \emptyset$ for $i \neq j$. The sets are chosen as follows:

$$\begin{aligned}
\mathscr{T}_{h,m}^1 &:= \{\, K \in \mathscr{T}_h \mid \bar{q}_k(t_m,x) = q_a \text{ or } \bar{q}_k(t_m,x) = q_b \text{ for all } x \in K \,\} \\
\mathscr{T}_{h,m}^2 &:= \{\, K \in \mathscr{T}_h \mid q_a < \bar{q}_k(t_m,x) < q_b \text{ for all } x \in K \,\} \\
\mathscr{T}_{h,m}^3 &:= \mathscr{T}_h \setminus (\mathscr{T}_{h,m}^1 \cup \mathscr{T}_{h,m}^2)
\end{aligned}$$

Hence, the set $\mathscr{T}_{h,m}^3$ consists of the cells which contain the free boundary between the active and the inactive sets for the time interval I_m.

Assumption 1. *We assume that there exists a positive constant C independent of k, h, and m such that*

$$\sum_{K \in \mathscr{T}_{h,m}^3} |K| \leq Ch.$$

We note, that this assumption is fulfilled, if the boundary of active sets consists of a finite number of rectifiable curves. Similar assumptions are used in [2, 20].

Under this assumption, the following estimate for the error in the control variable can be proven (cf. [19, Corollary 5.8]):

Theorem 3. *Let $\bar{q} \in Q_{ad}$ be the solution of the optimal control problem* (2) *and $\bar{q}_\sigma \in Q_{d,ad}$ be the solution of the discrete problem* (10), *where the cellwise (bi-/tri-)linear discretization for the control variable is employed. Then, if Assumption 1 is fulfilled, the following estimate holds for $n < p \leq \infty$ provided $z_k(\bar{q}_k) \in L^2(I, W^{1,p}(\Omega))$:*

$$\begin{aligned}
\|\bar{q} - \bar{q}_\sigma\|_I \leq \frac{C}{\alpha} k \big\{ \|\partial_t u(\bar{q})\|_I + \|\partial_t z(\bar{q})\|_I \big\} + \frac{C}{\alpha}\Big(1 + \frac{1}{\alpha}\Big) \big\{ h^2 \|\nabla^2 u_k(\bar{q}_k)\|_I \\
+ h^2 \|\nabla^2 z_k(\bar{q}_k)\|_I + h^{\frac{3}{2} - \frac{1}{p}} \|\nabla z_k(\bar{q}_k)\|_{L^2(I, L^p(\Omega))} \big\} = \mathscr{O}(k + h^{\frac{3}{2} - \frac{1}{p}}).
\end{aligned}$$

For elliptic optimal control problems and cellwise (bi-/tri-)linear discretization of the control space, the convergence order $\mathscr{O}(h^{\frac{3}{2}})$ can be obtained, see [2, 5, 22]. Especially in two space dimensions a corresponding estimate follows from the above theorem for the parabolic problem. In this case a uniform bound for $z_k(\bar{q}_k) \in L^2(I, W^{1,p}(\Omega))$ for all $p < \infty$ can be obtained leading to $\mathscr{O}(k + h^{\frac{3}{2} - \varepsilon})$ for each $\varepsilon > 0$.

3.3 Variational approach

In this section we prove an estimate for the error $\|\bar{q} - \bar{q}_\sigma\|_I$ in the case of no control discretization, cf. [13]. In this case we choose $Q_d = Q$ and thus, $Q_{d,\mathrm{ad}} = Q_{\mathrm{ad}}$. This implies $\bar{q}_\sigma = \bar{q}_{kh}$. However, $\bar{q}_{kh}$ is in general not a finite element function corresponding to the spatial mesh $\mathscr{T}_h$. This fact requires more care for the computation of $\bar{q}_{kh}$, see [13] for details. On the other hand, this approach provides the optimal order of convergence (cf. [19, Corollary 5.11]):

Theorem 4. *Let $\bar{q} \in Q_{ad}$ be the solution of optimization problem* (2) *and $\bar{q}_{kh} \in Q_{ad}$ be the solution of the discretized problem* (8)*. Then the following estimate holds:*

$$\|\bar{q} - \bar{q}_{kh}\|_I \le \frac{C}{\alpha} k \big\{ \|\partial_t u(\bar{q})\|_I + \|\partial_t z(\bar{q})\|_I \big\} + \frac{C}{\alpha} h^2 \big\{ \|\nabla^2 u_k(\bar{q})\|_I + \|\nabla^2 z_k(\bar{q})\|_I \big\} = \mathscr{O}(k + h^2).$$

3.4 Post-processing strategy

In this section, we extend the post-processing techniques initially proposed in [20] to the parabolic case. We discretize the control by piecewise constants in time and space as in Subsection 3.1. To improve the quality of the approximation, we additionally employ the post-processing step

$$\tilde{q}_\sigma := P_{Q_{\mathrm{ad}}} \left(-\frac{1}{\alpha} z_{kh}(\bar{q}_\sigma) \right), \tag{12}$$

which makes use of the projection

$$P_{Q_{\mathrm{ad}}} \colon Q \to Q_{\mathrm{ad}}, \quad P_{Q_{\mathrm{ad}}}(r)(t,x) = \max(q_a, \min(q_b, r(t,x))).$$

We obtain the following error estimate (cf. [19, Corollary 5.17]):

Theorem 5. *Let $\bar{q} \in Q_{ad}$ be the solution of the optimal control problem* (2) *and $\tilde{q}_\sigma \in Q_{ad}$ be given by means of* (12) *employing the adjoint state $z_{kh}(\bar{q}_\sigma)$ related to the solution $\bar{q}_\sigma$ of the discrete problem* (10)*, where the cellwise constant discretization for the control variable is employed. Let, moreover, Assumption 1 be fulfilled and $n < p \le \infty$. Then, it holds*

$$\begin{aligned}\|\bar{q} - \tilde{q}_\sigma\|_I \le \frac{C}{\alpha}\Big(1 + \frac{1}{\alpha}\Big) k \big\{ \|\partial_t u(\bar{q})\|_I + \|\partial_t z(\bar{q})\|_I \big\} \\ + \frac{C}{\alpha}\Big(1 + \frac{1}{\alpha}\Big) h^2 \big\{ \|\nabla^2 u_k(\bar{q}_k)\|_I + \frac{1}{\alpha} \|\nabla z_k(\bar{q}_k)\|_I + \Big(1 + \frac{1}{\alpha}\Big) \|\nabla^2 z_k(\bar{q}_k)\|_I \big\} \\ + \frac{C}{\alpha^2}\Big(1 + \frac{1}{\alpha}\Big) h^{2-\frac{1}{p}} \|\nabla z_k(\bar{q}_k)\|_{L^2(I, L^p(\Omega))} = \mathscr{O}\big(k + h^{2-\frac{1}{p}}\big)\end{aligned}$$

provided that $z_k(\bar{q}_k) \in L^2(I, W^{1,p}(\Omega))$.

Since in two space dimensions the adjoint solution $z_k(\bar{q}_k)$ is uniformly bounded in $L^2(I, W^{1,p}(\Omega))$ with respect to k for all $p < \infty$, the presented estimate leads to the almost optimal order of convergence $\mathscr{O}(k + h^{2-\varepsilon})$ for each $\varepsilon > 0$.

References

1. Arada, N., Casas, E., Tröltzsch, F.: Error estimates for a semilinear elliptic optimal control problem. Comput. Optim. Appl. **23**, 201–229 (2002)
2. Becker, R., Vexler, B.: Optimal control of the convection-diffusion equation using stabilized finite element methods. Numer. Math. **106**(3), 349–367 (2007)
3. Brenner, S., Scott, R.L.: The mathematical theory of finite element methods. Springer Verlag, Berlin Heidelberg New York (2002)
4. Casas, E.: Using piecewise linear functions in the numerical approximation of semilinear elliptic control problems. Adv. Comput. Math. **26**(1-3), 137–153 (2007)
5. Casas, E., Mateos, M.: Error estimates for the numerical approximation of boundary semilinear elliptic control problems. Continuous piecewise linear approximations. In: Systems, control, modeling and optimization, pp. 91–101. Springer, New York (2006). IFIP Int. Fed. Inf. Process., 202
6. Casas, E., Mateos, M.: Error estimates for the numerical approximation of neumann control problems. Comput. Optim. Appl. (2007). DOI 10.1007/s10589-007-9056-6. Published electronically
7. Casas, E., Mateos, M., Tröltzsch, F.: Error estimates for the numerical approximation of boundary semilinear elliptic control problems. Comput. Optim. Appl. **31**(2), 193–220 (2005)
8. Casas, E., Tröltzsch, F.: Error estimates for linear-quadratic elliptic control problems. In: V. Barbu, I. Lasiecka, D. Tiba, C. Varsan (eds.) Analysis and Optimization of Differential Systems, pp. 89–100. Kluwer Academic Publishers, Boston (2003). Proceeding of International Working Conference on Analysis and Optimization of Differential Systems
9. Eriksson, K., Estep, D., Hansbo, P., Johnson, C.: Computational Differential Equations. Cambridge University Press, Cambridge (1996)
10. Eriksson, K., Johnson, C., Thomée, V.: Time discretization of parabolic problems by the discontinuous Galerkin method. M2AN Math. Model. Numer. Anal. **19**, 611–643 (1985)
11. Falk, R.: Approximation of a class of optimal control problems with order of convergence estimates. J. Math. Anal. Appl. **44**, 28–47 (1973)
12. Geveci, T.: On the approximation of the solution of an optimal control problem governed by an elliptic equation. M2AN Math. Model. Numer. Anal. **13**, 313–328 (1979)
13. Hinze, M.: A variational discretization concept in control constrained optimization: The linear-quadratic case. Comput. Optim. and Appl. **30**(1), 45–61 (2005)
14. Lasiecka, I., Malanowski, K.: On discrete-time Ritz-Galerkin approximation of control constrained optimal control problems for parabolic systems. Control Cybern. **7**(1), 21–36 (1978)
15. Lions, J.L.: Optimal Control of Systems Governed by Partial Differential Equations, *Grundlehren Math. Wiss.*, vol. 170. Springer, Berlin (1971)
16. Malanowski, K.: Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal-control problems. Appl. Math. Optim. **8**, 69–95 (1981)
17. McNight, R.S., Bosarge jr., W.E.: The Ritz-Galerkin procedure for parabolic control problems. SIAM J. Control Optim. **11**(3), 510–524 (1973)
18. Meidner, D., Vexler, B.: A priori error estimates for space-time finite element approximation of parabolic optimal control problems. Part I: Problems without control constraints. SIAM J. Control Optim. **47**(3), 1150–1177 (2008)
19. Meidner, D., Vexler, B.: A priori error estimates for space-time finite element approximation of parabolic optimal control problems. Part II: Problems with control constraints. SIAM J. Control Optim. **47**(3), 1301–1329 (2008)
20. Meyer, C., Rösch, A.: Superconvergence properties of optimal control problems. SIAM J. Control Optim. **43**(3), 970–985 (2004)
21. Rösch, A.: Error estimates for parabolic optimal control problems with control constraints. Zeitschrift für Analysis und ihre Anwendungen ZAA **23**(2), 353–376 (2004)
22. Rösch, A.: Error estimates for linear-quadratic control problems with control constraints. Optim. Methods Softw. **21**(1), 121–134 (2006)
23. Winther, R.: Error estimates for a Galerkin approximation of a parabolic control problem. Ann. Math. Pura Appl. (4) **117**, 173–206 (1978)