Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn
Wegelerstraße 6 • 53115 Bonn • Germany
phone +49 228 73-3427 • fax +49 228 73-7527
www.ins.uni-bonn.de

F. Ludovici, I. Neitzel, W. Wollner

A priori error estimates for state constrained semilinear parabolic optimal control problems

INS Preprint No. 1605

February 2016

A priori error estimates for state constrained semilinear parabolic optimal control problems

FRANCESCO LUDOVICI, IRA NEITZEL, AND WINNIFRIED WOLLNER

Abstract. We consider the finite element discretization of semilinear parabolic optimization problems subject to pointwise in time constraints on mean values of the state variable. In contrast to many results in numerical analysis of optimization problems subject to semilinear parabolic equations, we assume weak second order sufficient conditions. Relying on the resulting quadratic growth condition of the continuous problem, we derive rates of convergence as temporal and spatial mesh sizes tend to zero.

1. Introduction

This paper is concerned with optimal control problems governed by semilinear parabolic partial differential equations (PDEs) subject to pointwise in time constraints on mean values in space of the solution to the PDE. We derive convergence rates for a space-time discretization of the problem based on conforming finite elements in space and a discontinuous Galerkin discretization method in time. The control variable is an $\mathbb{R}^m$-vector-valued function depending on time, acting distributed in the domain. The inequality constraint on the solution of the PDE is imposed pointwise in time and averaged in space. Extending the result of [26] to the case of semilinear parabolic PDEs, we consider, for a time interval $I = (0, T)$ and a domain $\Omega \subset \mathbb{R}^2$ the following problem

$$\text{Minimize } J(q,u) := \frac{1}{2}\int_I\int_\Omega (u(t,x) - u_{\mathrm{d}}(t,x))^2 \mathrm{d}x\mathrm{d}t + \frac{\alpha}{2}\int_I q(t)^T q(t)\mathrm{d}t, \tag{1a}$$

where the state $u(t,x)$ and the control $q(t) = (q_i)_{i=1}^m$ are coupled by the semilinear parabolic PDE

$$\begin{aligned} \partial_t u(t,x) - \Delta u(t,x) + d(t,x,u(t,x)) &= \sum_{i=1}^m q_i(t) g_i(x) && \text{in } I \times \Omega, \\ u(t,x) &= 0 && \text{on } I \times \partial\Omega, \\ u(0,x) &= u_0 && \text{in } \{0\} \times \Omega, \end{aligned} \tag{1b}$$

with a monotone and smooth nonlinearity d. Further, we consider control constraints

$$q_{min} \leq q(t) \leq q_{max} \qquad \text{a.e. in } I, \tag{1c}$$

and, for a given weighting function $\omega(x)$, state constraints

$$\int_\Omega u(t,x)\omega(x)\mathrm{d}x \leq 0 \qquad \forall t \in [0,T]. \tag{1d}$$

The precise formulation of the problem will be given in the next section.

Date: February 26, 2016.

This class of problems is a simplified model motivated by applications from industrial processes like cooling/heating in steel manufacturing, or tumor therapy in biomathematics. For an extended overview of the possible applications we refer, e.g., to [14, 21]. In many of these applications the control variable depends on finitely many parameters with fixed spatial influence but varying in time. Further, especially in cooling processes and material optimization, bounds on the state variable and its derivatives are prescribed to avoid material failure and to preserve product quality.

Despite all these interesting applications, the literature on a priori error estimates for semilinear parabolic optimal control, even without state constraints, has only a few contributions. Error estimates were derived in [23, 31] in a setting including bilateral control constraints; in the latter the authors also discussed several control discretization approaches. Error estimates were obtained in [12, 13] for a problem without control and state constraints.

The lack of results for semilinear parabolic problems in the presence of state constraints is also explained by the sparsity of results for the corresponding linear theory. Only recently, error estimates for the space-time discretization of the state equation in the $L^\infty(I, L^2(\Omega))$ and $L^\infty(I, H_0^1(\Omega))$-norm were derived in [26] and [25], respectively. Indeed, estimates in these norms are necessary for the consideration of constraints pointwise in time on the mean value of the state variable and its first derivative. We would also like to point out the new result [24] where a pointwise (quasi)-best-approximation result in $L^\infty(I \times \Omega)$ for the discretization of linear parabolic PDEs has been derived.

Confining ourselves to the linear parabolic setting, error estimates in the L^2-norm for pointwise in time and space state constraints are derived in [18], while [15] is concerned with the variational discretization approach. For control constraints, we refer to [27, 28].

Less sparse is the literature on state constrained semilinear elliptic problems. We refer the reader to [8, 20, 30] and the references therein.

Recently, more attention was devoted to the study of second-order optimality conditions for state constrained parabolic optimal control problems. A well-written survey on the state-of-the art can be found in [11].

For the case at hand, we will rely on second-order sufficient conditions (SSCs) that were introduced in [6]. The authors, inspired by techniques from nonlinear optimization in finite-dimensional spaces, obtained SSCs that are very close to the necessary ones. Their analysis was limited to the one dimensional case, i.e., $\Omega \subset \mathbb{R}^1$, and has been extended in [14] to domains of arbitrary dimensions considering, as in our case, $\mathbb{R}^m$-vector valued controls functions depending on time only. Due to the nature of the problem, the resulting cone of critical directions can be recasted also from the theory of semi-infinite optimization [2].

Seminal papers for the theory of SSCs in presence of integral state constrains are [5, 17]. The former deals with boundary controls and handles the state constraints using Ekeland's principle. The latter considers a nonlinearity in the boundary conditions and uses concepts of semi-group theory to cope with the limitations on the dimension of the domain. More recently, [22] has overcome the limitation in the dimension using concepts of maximal parabolic regularity. For other contributions to the theory of SSCs in presence of state constraints, we refer to [3, 9, 33].

In this paper one of our main aims is to use weak SSCs for the continuous problem as derived in [6] in order to prove discretization error estimates. For elliptic state-constrained problems, it is known that the proof of convergence requires the quadratic growth condition for the continuous problem, only. Once the quadratic growth for the continuous problem is given, it does not matter if the growth is given due to weak or strong SSCs, see [30] in the elliptic case.

In the parabolic setting, if one wants to achieve a clear separation of the spatial and temporal errors, the numerical analysis has previously been done in two steps introducing an intermediate time-discrete problem, cf., [25, 27, 28] for convex or [31] for non-convex problems. As a consequence of this approach, a quadratic growth condition has to hold for the time-discrete problem in order to prove an error estimate for the fully-discrete problem. Rather than relying on an additional assumption for each time-discrete problem, SSCs can be transfered from the continuous to the semidiscrete level if one uses a rather strong SSC, see [31]. In contrast, weak SSCs have not been shown to be stable with respect to time discretization; and it is not clear at all if this is possible without further assumptions.

In this paper, in favor of the more general weak SSCs, we will derive error estimates without the use of an intermediate auxiliary problem. Our main result is the error estimate of Theorem 33, namely the convergence

$$\|\bar{q}-\bar{q}_{kh}\|^2_{L^2(I,\mathbb{R}^m)} \leq c\left(k\Big(\log\frac{T}{k}+1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k}+1\Big)\right)$$

which coincides with the orders obtained for convex problem in [26].

Our technique allows to derive an estimate for the error between the continuous and semidiscrete solution of order

$$\|\bar{q}-\bar{q}_{k}\|^2_{L^2(I,\mathbb{R}^m)} \leq ck\Big(\log\frac{T}{k}+1\Big)^{\frac{1}{2}},$$

only depending on the time step size k. The price we pay for not transfering the SSC is that the analogous error estimate

$$\|\bar{q}_k-\bar{q}_{kh}\|^2_{L^2(I,\mathbb{R}^m)} \leq ch^2\Big(\log\frac{T}{k}+1\Big)$$

can not be shown.

The paper is organized as follows: in Section 2, we give a precise definition of the model problem sketched in (1), introduce the operators and functionals involved in the analysis and state first and second order optimality conditions. Section 3 is devoted to the time and space discretization of the problem. In Section 4, we derive estimates in the $L^\infty(I, L^2(\Omega))$-norm for the discretization error between the solution of the continuous, semidiscrete and discrete state equation, extending techniques from [26] for linear parabolic problems to the case at hand. The core of the paper is Section 5. Extending techniques for the elliptic case presented in [30] to the parabolic case, we derive the rate of convergence for the optimal control problem.

2. Assumptions and analytic setting

In this section, we discuss the precise analytic setting of the problem, introduce the main assumptions, and fix the notation.

The spatial domain $\Omega \subset \mathbb{R}^2$ is a convex, bounded domain with C^2-boundary $\partial\Omega$ and $I = (0, T)$ is a given time interval. Further, for $i = 1, ..., m$, we consider controls $q_i \in L^2(I)$ and fixed functions $g_i \in L^\infty(\Omega)$. We assume that the desired state satisfies $u_d \in L^2(I \times \Omega)$ and the initial data satisfies $u_0 \in H_0^1(\Omega) \cap H^2(\Omega) \cap C(\overline{\Omega})$. The state constraint is denoted by $F(u) := (u, \omega)$, where $\omega \in L^\infty(\Omega)$ is a weighting function.

In the following, we set $V := H_0^1(\Omega)$, $H := L^2(\Omega)$; $(\cdot, \cdot)_I$ denotes the standard inner product in $L^2(I, H)$, i.e., $(\cdot, \cdot)_I = \int_I (\cdot, \cdot)\mathrm{d}t$ with associated norm $\|\cdot\|_I$, while $(\cdot, \cdot)$ and $\|\cdot\|$ is used for $L^2(\Omega)$. Throughout the paper, c will denote a generic constant independent of the discretization parameters, that may take different values at each appearance.

Before discussing the problem in detail, we impose the following usual assumptions on the nonlinearity, see, e.g., [35, Chapter 5, Assumption 5.6].

Assumption 1. *The nonlinearity $d(t, x, u)\colon I \times \Omega \times \mathbb{R}$ is assumed to satisfy the following:*

- **(i)** *For all $u \in \mathbb{R}$, the nonlinearity is measurable with respect to $(t, x) \in I \times \Omega$. Further, for almost every $(t, x) \in I \times \Omega$ it is twice continuously differentiable with respect to u.*
- **(ii)** *For $u = 0$, there is $c > 0$ such that $d(t, x, u)$ satisfies, together with its derivatives up to order two, the boundedness condition*

$$\|d(\cdot, \cdot, 0)\|_{L^\infty(I\times\Omega)} + \|\partial_u d(\cdot, \cdot, 0)\|_{L^\infty(I\times\Omega)} + \|\partial_u^2 d(\cdot, \cdot, 0)\|_{L^\infty(I\times\Omega)} \le c.$$

Further, each of these satisfy a local Lipschitz condition with respect to u, i.e., for any $M > 0$ there exist a constant $L(M) > 0$ such that for any $|u_j| \le M$ $j = 1, 2$ there holds

$$\|\partial_u^i d(\cdot, \cdot, u_1) - \partial_u^i d(\cdot, \cdot, u_2)\|_{L^\infty(I\times\Omega)} \le L(M)|u_1 - u_2|,$$

for every $i = 0, 1, 2$ and almost every $(t, x) \in I \times \Omega$.
- **(iii)** *For all $u \in \mathbb{R}$ and for almost every $(t, x) \in I \times \Omega$, there holds the monotonicity condition*

$$\partial_u d(t, x, u) \ge 0.$$

When no confusion arises, we shorten the notation for the semilinearity from $d(\cdot, \cdot, u)$ to $d(u)$.

We now focus on the well-posedness of the state equation (1b). We introduce the Hilbert space

$$W(0, T) = \{u \in L^2(I, V), \partial_t u \in L^2(I, V^*)\},$$

and the space of admissible controls

$$Q_{\mathrm{ad}} = \{q \in L^2(I, \mathbb{R}^m) \,|\, q_{min} \le q(t) \le q_{max}, \text{ a.e. in } I\}$$

with $q_{min} < q_{max} \in \mathbb{R}^m$.

Denoting with V^* the dual space of V, we recall that the triplet

$$V \hookrightarrow H \hookrightarrow V^*$$

forms a Gelfand triple. Then, for $u, \varphi \in W(0, T)$ we define a bilinear form

$$b(u, \varphi) = (\partial_t u, \varphi)_I + (\nabla u, \nabla\varphi)_I + (u(0), \varphi(0))$$

and the weak formulation of (1b) reads: for given $q \in L^2(I, \mathbb{R}^m)$ and initial data $u_0 \in V \cap H^2(\Omega) \cap C(\overline{\Omega})$, find $u \in W(0,T)$ satisfying

$$b(u, \varphi) + (d(\cdot, \cdot, u), \varphi)_I = (qg, \varphi)_I + (u_0, \varphi(0)), \ \forall \varphi \in W(0,T). \tag{2}$$

It is well known that (2) admits a unique solution $u \in W(0,T) \cap C(\overline{I} \times \overline{\Omega})$, see, e.g., [35, Theorem 5.5]. Further, thanks to the monotonicity assumption on $d(\cdot, \cdot, u)$, the solution u of (2) satisfies the additional regularity

$$u \in L^2(I, V \cap H^2(\Omega)) \cap L^\infty(I \times \Omega) \cap H^1(I, L^2(\Omega)), \tag{3}$$

and the following stability estimates hold, cf. [31, Proposition 2.1], justifying the use of the L^2 inner-product in the notation of (2).

Proposition 2. *Let $u \in W(0,T)$ be the solution of (2) for given data q, g, u_0 and d. Then, there holds*

$$\begin{aligned} &\|u\|_{L^\infty(I\times\Omega)} \le c\big(\|qg\|_{L^\infty(I\times\Omega)} + \|u_0\|_{L^\infty(\Omega)} + \|d(0)\|_{L^\infty(I\times\Omega)}\big), \\ &\|u\|_{L^2(I,V)} + \|u\|_{L^\infty(I,V)} + \|\partial_t u\|_I \le c\big(\|qg\|_I + \|u_0\|_V + \|d(0)\|_I\big). \end{aligned} \tag{4}$$

Remark 3. *We observe that the regularity of $u \in W(0,T)$ is enough to treat the state constraint. Indeed, there holds the embedding $W(0,T) \hookrightarrow C(I,H)$ and we have $F\colon W(0,T) \to C(\overline{I})$ where*

$$F(u)(t) := \int_\Omega u(t,x)\omega(x)\mathrm{d}x.$$

On the other hand, we require more regularity for the solution of (2), because stability estimates in the norms of $L^\infty(I \times \Omega)$, $L^\infty(I,H)$ will come into play to ensure Lipschitz continuity for the control-to-state map. Further, we note that $u_0 \in V \cap C(\overline{\Omega})$ is enough to ensure well-posedness of the problem. The assumption $u_0 \in H^2(\Omega)$ is posed to use results from [26, 28], where this regularity is required to fully exploit the approximation property of the discontinuous Galerkin method.

Thanks to (1c), we can regard the control variable q as an element of $L^\infty(I, \mathbb{R}^m)$. Then the following definitions are justified. We introduce the control-to-state map

$$S\colon L^\infty(I, \mathbb{R}^m) \to W(0,T) \cap C(I \times \Omega),$$

associating to any given q the solution $u(q) := S(q)$ of (2). We denote the concatenation of the control-to-state map and the state constraint F by

$$G = (F \circ S)\colon L^\infty(I, \mathbb{R}^m) \to C(\overline{I}).$$

In the subsequent analysis, we will need G to be of class C^2. This is indeed the case, see [6].

In order to formulate the optimal control problem in reduced form, we introduce the set of feasible controls

$$Q_{\text{feas}} = \{q \in Q_{\text{ad}} \,|\, G(q) \le 0\}.$$

Then, (1) reads

$$\min j(q) := J(q, S(q)) \qquad \text{s.t. } q \in Q_{\text{feas}}. \tag{$\mathbb{P}$}$$

Proposition 4. *Assuming the existence of a feasible point, problem ($\mathbb{P}$) admits at least one solution $(\bar{q}, \bar{u}) \in L^\infty(I, \mathbb{R}^m) \times \big(W(0,T) \cap C(\overline{I} \times \bar{\Omega}) \cap H^1(I,H)\big)$, where $\bar{u} = S(\bar{q})$.*

Proof. The existence of a feasible point will be justified in the next section after introducing a Slater type regularity condition. The additional regularity of the control is a consequence of the box-control constraints. Then, the assertion follows by standard arguments see, e.g., [35, Theorem 5.7]. ☐

We remark that the problem at hand is non-convex due to the presence of the nonlinear term in the state equation. As a consequence, it is suitable to consider local solutions as defined below.

Definition 5. *A control $\bar{q} \in Q_{feas}$ is a local solution in the sense of $L^2(I, \mathbb{R}^m)$ if there exists some $\epsilon > 0$ such that there holds*

$$j(\bar{q}) \leq j(q)$$

for all $q \in Q_{feas}$ with $\|q - \bar{q}\|_{L^2(I,\mathbb{R}^m)} \leq \epsilon$.

We conclude the section with well-known differentiability properties of the operators and functionals involved in the analysis, referring to [35, Chapter 5] for details.

Lemma 6. *The map $S\colon L^\infty(I, \mathbb{R}^m) \to W(0,T) \cap L^\infty(I \times \Omega)$ is of class C^2 from $L^\infty(I, \mathbb{R}^m)$ to $W(0,T)$. For $p \in L^\infty(I, \mathbb{R}^m)$ and for all $\varphi \in W(0,T)$, its first derivative $S^{'}(q)p =: v_p$ in direction p is the solution of*

$$b(v_p, \varphi) + (\partial_u d(\cdot, \cdot, u(q))v_p, \varphi)_I = (pg, \varphi)_I. \tag{5}$$

For $p_1, p_2 \in L^\infty(I, \mathbb{R}^m)$ and for all $\varphi \in W(0,T)$, the second derivative $S^{''}(q)p_1p_2 =: v_{p_1p_2}$ in the directions p_1, p_2 is the solution of

$$b(v_{p_1p_2}, \varphi) + (\partial_u d(\cdot, \cdot, u_q)v_{p_1p_2}, \varphi)_I = -(\partial_{uu} d(\cdot, \cdot, u_q)v_{p_1}v_{p_2}, \varphi)_I, \tag{6}$$

where v_{p_1}, v_{p_2} are given by (5).

For S and its first derivative the following Lipschitz properties hold.

Lemma 7. *For $p, q_1, q_2 \in Q_{ad}$, there exists a constant $c > 0$ such that*

$$\|S(q_1) - S(q_2)\|_I \leq c\|q_1 - q_2\|_{L^2(I,\mathbb{R}^m)}, \tag{7a}$$

$$\|S(q_1) - S(q_2)\|_{L^\infty(I,H)} \leq c\|q_1 - q_2\|_{L^2(I,\mathbb{R}^m)}, \tag{7b}$$

$$\|S(q_1) - S(q_2)\|_{L^\infty(I,V)} \leq c\|q_1 - q_2\|_{L^2(I,\mathbb{R}^m)}, \tag{7c}$$

$$\|S^{'}(q_1)p - S^{'}(q_2)p\|_I \leq c\|q_1 - q_2\|_{L^2(I,\mathbb{R}^m)}\|p\|_{L^2(I,\mathbb{R}^m)}, \tag{7d}$$

$$\|S^{'}(q_1)p - S^{'}(q_2)p\|_{L^\infty(I,H)} \leq c\|q_1 - q_2\|_{L^2(I,\mathbb{R}^m)}\|p\|_{L^2(I,\mathbb{R}^m)}. \tag{7e}$$

Proof. The claim follows from [31, Lemma 2.3] where, for $q_1 g, q_2 g \in L^\infty(I \times \Omega)$, it is shown that

$$\|S(q_1) - S(q_2)\|_I \leq c\|g(q_1 - q_2)\|_I.$$

Adapting the argument to the time dependent nature of the control variable for the case at hand, we deduce

$$\|S(q_1) - S(q_2)\|_I \leq c\|g\|_{L^\infty(\Omega)}\|q_1 - q_2\|_{L^2(I,\mathbb{R}^m)}.$$

Similarly, we deduce (7b), (7d), compare with the proof of [31, Lemma 2.3].

To show (7e), we consider $\xi := S^{'}(q_1)p - S^{'}(q_2)p$ and define $\tilde{u} := S^{'}(q_2)p$. We note that, for any $\varphi \in W(0,T)$, ξ fulfills

$$b(\xi, \varphi) + \big(\partial_u d(u(q_1))\xi, \varphi\big)_I = -\big(\partial_u d(u(q_1))\tilde{u} - \partial_u d(u(q_2))\tilde{u}, \varphi\big)_I. \tag{8}$$

It is clear that, due to the boundedness of $\partial_u d(\cdot)$, for $S^{'}(q)p$ there hold the same stability estimates as for $S(q)$, compare with (4). Then, by means of such stability estimate in $L^\infty(I,H)$ in combination with the Lipschitz continuity of $\partial_u d(\cdot)$, we obtain

$$\begin{aligned}\|\xi\|_{L^\infty(I,H)} &\leq c\|\big(\partial_u d(u(q_1)) - \partial_u d(u(q_2))\big)\tilde{u}\|_I \\ &\leq c\|u(q_1) - u(q_2)\|_{L^4(I\times\Omega)}\|\tilde{u}\|_{L^4(I\times\Omega)} \\ &\leq c\|u(q_1) - u(q_2)\|_{L^\infty(I,V)}\|\tilde{u}\|_{L^\infty(I,V)} \\ &\leq c\|q_1 - q_2\|_I\|p\|_I,\end{aligned}$$

where we used the embedding $L^\infty(I,V) \hookrightarrow L^4(I\times\Omega)$. □

Corollary 8. *The functional $j(q)\colon L^\infty(I,\mathbb{R}^m) \to \mathbb{R}$ is of class C^2 in $L^\infty(I,\mathbb{R}^m)$ and for $q, p, p_1, p_2 \in L^\infty(I,\mathbb{R}^m)$ there holds*

$$j^{'}(q)(p) = \int_\Omega \sum_{i=1}^m (\alpha q_i(t) + z_0(q)g_i(x))p_i(t)\,\mathrm{d}t,$$

$$j^{''}(q)p_1p_2 = \int_{\Omega_I} (v_{p_1}v_{p_2} + \alpha p_1p_2 - z_0(q)\partial_{uu}d(x,t,u(q))v_{p_1}v_{p_2})\,\mathrm{d}t\mathrm{d}x,$$

where $z_0(q) \in W(0,T)$ is the adjoint state associated with q and j, defined, for all $\varphi \in W(0,T)$ as the unique solution of

$$b(\varphi, z) + (\partial_u d(\cdot,\cdot,u(q))z, \varphi)_I = (u_q - u_d, \varphi)_I, \tag{9}$$

and $v_{p_i}, i = 1,2$ is defined as (5).

Remark 9. *As observed in [35, Section 5.7.4], when the control appears quadratically in the cost functional and linearly in the state equation, then the reduced cost functional is of class C^2 not only in $L^\infty(I,\mathbb{R}^m)$ but also in $L^2(I,\mathbb{R}^m)$, see also [6, Remark 2.8]. In particular, this allows the introduction of a quadratic growth condition without two-norm discrepancy.*

2.1. **Optimality conditions.** In this section, we discuss the optimality conditions for our optimal control problem. In a first step, we state standard first-order necessary conditions in KKT form. Then, we introduce second-order sufficient conditions using the approach developed in [7] for semilinear elliptic problems. Their analysis was extended to semilinear parabolic problems in [6] for the one-dimensional case. Indeed, in [6] control functions are from $L^2(I\times\Omega)$. As a consequence, the control-to-state map is not in general twice continuously differentiable from $L^2(I\times\Omega)$ to $C(\bar{I}\times\bar{\Omega})$, however, it is always the case in the one-dimensional setting. This restriction has been circumvented in [14], considering, as in this paper, controls depending on time only.

We rely on the following linearized Slater's regularity condition.

Assumption 10. *Given a local solution $\bar{q}$ of $(\mathbb{P})$, we assume the existence of $q_\gamma \in Q_{ad}$ such that*

$$G(\bar{q}) + G^{'}(\bar{q})(q_\gamma - \bar{q}) \leq -\gamma < 0, \tag{10}$$

for some $\gamma \in \mathbb{R}^+$.

Based on the Slater condition, we obtain first order necessary optimality conditions in KKT form, see, e.g., [3].

Theorem 11. *Let $\bar{q} \in Q_{feas}$ be a local solution of* $(\mathbb{P})$ *such that Assumption 10 is satisfied, and let $\bar{u}$ be the associated state. Then, there exists a Lagrange multiplier $\bar{\mu} \in C(\bar{I})^*$ and an adjoint state $\bar{z} \in L^2(I, V)$ such that*

$$b(\bar{u}, \varphi) + (d(\cdot, \cdot, \bar{u}), \varphi)_I = (\bar{q}g, \varphi)_I + (u_0, \varphi(0)) \qquad \forall \varphi \in W(0,T), \tag{11a}$$

$$b(\varphi, \bar{z}) + (\varphi, \partial_u d(\cdot, \cdot, \bar{u})\bar{z}) = (\bar{u} - u_d, \varphi)_I + \langle \bar{\mu}, F(\varphi) \rangle \qquad \forall \varphi \in W(0,T), \tag{11b}$$

$$\alpha(\bar{q}, q - \bar{q})_{L^2(I)} + (\bar{z}, (q - \bar{q})g)_I \geq 0 \qquad \forall q \in Q_{ad}, \tag{11c}$$

$$\langle F(\bar{u}), \bar{\mu} \rangle = 0, \ \bar{\mu} \geq 0, \ F(\bar{u}) \leq 0 \tag{11d}$$

where we used the linearity of $F(\cdot)$, and $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $C(\bar{I})^*$ and $C(\bar{I})$.

To discuss SSCs, we introduce the Hamiltonian function $H \colon \mathbb{R} \times \Omega \times \mathbb{R} \times \mathbb{R} \times R \to \mathbb{R}$ given by

$$H(q, u, z) = H(t, x, q, u, z) = \frac{1}{2}(u - u_d)^2 + \frac{\alpha}{2}q^2 + z\Big(\sum_{i=1}^{m} q_i g_i - d(u)\Big),$$

suppressing the first two arguments t, x in the exposition. Moreover, the reduced Lagrangian function is given by

$$\mathcal{L}(q, \mu) = j(q) + \langle \mu, G(u) \rangle.$$

Remark 12. *For better readability, at each $(t, x) \in (I \times \Omega)$, we denote by $\bar{H}, \bar{\mathcal{L}}$ the Hamiltonian and Lagrangian function when evaluated at $(\bar{q}, \bar{u}, \bar{z})$. We note that $\frac{\partial H}{\partial q}, \frac{\partial^2 H}{\partial q^2}$ are, respectively, an $\mathbb{R}^m$-vector and an $\mathbb{R}^{m \times m}$-matrix. When referring to the i-th component and the (i,j)-entry, we abbreviate $\partial_q H_i, \partial_q^2 H_{i,j}$, respectively.*

We now give the cone of critical directions associated with $\bar{q} \in Q_{\text{feas}}$, following [6]. Introducing the conditions

$$p_i(t) = \begin{cases} \geq 0 & \text{if } \bar{q}_i = q_{min}, \\ \leq 0 & \text{if } \bar{q}_i = q_{max}, \\ = 0 & \text{if } \int_\Omega \partial_q \bar{H}_i \mathrm{d}x \neq 0, \end{cases} \qquad \text{for all } i = 1, ..., m \tag{12}$$

$$F(v_p) = \frac{\partial F}{\partial u}(\bar{u}) v_p \leq 0 \text{ if } F(\bar{u}) = 0, \tag{13}$$

$$\int_\Omega F(v_p) d\bar{\mu} = 0, \tag{14}$$

where v_p is defined by (5), the cone of critical direction is given by

$$C_{\bar{q}} = \{p \in L^2(I, \mathbb{R}^m) \,|\, p \text{ satisfies (12), (13), (14)}\}. \tag{15}$$

After this preparation, we postulate the following second-order sufficient condition.

Assumption 13. *Let $\bar{q} \in Q_{feas}$ fulfill, together with the associated state $\bar{u}$, the adjoint state $\bar{z}$, and Lagrange multipliers $\bar{\mu}$, the first-order optimality conditions* (11). *Then, we assume*

$$\frac{\partial^2 \bar{\mathcal{L}}}{\partial^2 q} p^2 > 0 \qquad \forall p \in C_{\bar{q}} \setminus \{0\}. \tag{16}$$

Remark 14. *Comparing the second-order sufficient condition of Assumption 13 with the one of [6], we observe that the assumption*

$$\partial_q^2 \bar{H}_{i,i} \geq \xi \qquad \forall t \in I \setminus E_i^\nu, \forall i = 1, ..., m,$$

where

$$E_i^\nu = \Big\{ t \in I \,\Big|\, \big| \int_\Omega \partial_q \bar{H}_i \mathrm{d}x \big| \geq \nu \Big\}$$

is the set of sufficiently active control constraints and ξ, ν are positive constants, is implicitly satisfied in our setting. Indeed, since the control appears quadratically in the cost functional and linearly in the state equation, it trivially follows $\partial_q^2 \bar{H}_{i,i} = \alpha \mathbb{I} > 0$, where $\mathbb{I}$ denotes the identity operator.

With the second-order conditions at hand, we obtain the following quadratic growth condition.

Theorem 15. *Let $\bar{q} \in Q_{feas}$ satisfy the first order necessary optimality conditions* (11) *and let Assumption 13 hold. Then, there exist constants $\delta, \eta > 0$ such that*

$$j(q) \geq j(\bar{q}) + \delta \|q - \bar{q}\|^2_{L^2(I,\mathbb{R}^m)} \tag{17}$$

for any $q \in Q_{feas}$ with $\|q - \bar{q}\|_{L^2(I,\mathbb{R}^m)} \leq \eta$.

Proof. The proof is by contradiction and moves along the lines of [6, Theorem 4.1]. The approach there has been extended to our setting in [14, Theorem 5]. The only difference is the C^2-differentiability in $L^2(I, \mathbb{R}^m)$ of the reduced cost functional j, see Remark 9, leading to the quadratic growth condition (17) without two-norm discrepancy. □

3. Discretization

We briefly describe the discretization in time and space of our problem. We use the $dG(0)cG(1)$ method, discontinuous in time and continuous in space Galerkin method, referring to [34] for additional details.

The control variable is discretized implicitly by the optimality conditions through the variational discretization approach, attributed to [19].

3.1. Time discretization.

We consider a partitioning of $\overline{I}$ consisting of time intervals $I_n = (t_{n-1}, t_n]$ for $n = 1, ..., N$ and $I_0 = \{0\}$, where the times t_i are such that $0 = t_0 < t_1 < ... < t_{N-1} < t_N = T$. The length of the interval I_n is k_n and we set $k = \max_n k_n$ imposing that $k < T$. Further, we assume the existence of strictly positive constants $a, b, \tilde{k}$ such that the following technical conditions hold:

$$\min_{n>0} k_n \geq a k^b, \qquad\qquad \tilde{k}^{-1} \leq \frac{k_n}{k_{n+1}} \leq \tilde{k} \quad \forall n > 0.$$

We denote with $\mathcal{P}_0(I_n, V)$ the space of piecewise constant polynomials on I_n with values in V. The semidiscrete state and trial space is

$$U_k = U_k(V) = \big\{ \varphi_k \in L^2(I, V) \,\big|\, \varphi_{k,n} = \varphi_k|_{I_n} \in \mathcal{P}_0(I_n, V),\, n = 1, ..., N \big\},$$

with inner product $(\cdot,\cdot)_{I_n}$ and norm $\|\cdot\|_{I_n}$ given by the restriction of the usual inner product and norm of $L^2(I, H)$ onto the interval I_n, i.e., $(\cdot,\cdot)_{I_n} = \int_{I_n} (\cdot,\cdot)\mathrm{d}t$.

Our functions are piecewise constant on each interval. Thus, we can simplify standard notation and for functions $\varphi_k \in U_k$ we write

$$\varphi_{k,n+1} = \varphi_{k,n}^{+} = \lim_{t\to 0^+} \varphi_k(t_n + t) = \lim_{t\to 0^+} \varphi_k(t_{n+1} - t), \qquad [\varphi_k]_n = \varphi_{k,n+1} - \varphi_{k,n}.$$

For $u_k, \varphi_k \in U_k$, the semidiscrete bilinear form is defined as

$$\begin{aligned} B(u_k, \varphi) &= \sum_{n=1}^{N} (\partial_t u_k, \varphi)_{I_n} + (\nabla u_k, \nabla\varphi)_I + \sum_{n=2}^{N} ([u_k]_{n-1}, \varphi_n) + (u_{k,1}, \varphi_1) \\ &= (\nabla u_k, \nabla\varphi)_I + \sum_{n=2}^{N} ([u_k]_{n-1}, \varphi_n) + (u_{k,1}, \varphi_1), \end{aligned}$$

and the semidiscrete state equation reads: given $q \in L^2(I, \mathbb{R}^m)$, $u_0 \in H^2(\Omega) \cap V$, find $u_k = u_k(q) \in U_k$ such that

$$B(u_k, \varphi_k) + (d(\cdot,\cdot,u_k), \varphi_k)_I = (qg, \varphi_k)_I + (u_0, \varphi_{k,1}), \ \forall \varphi_k \in U_k. \tag{18}$$

Using standard arguments, one can show that (18) admits a unique solution in U_k, see [31, Theorem 3.1] and the references therein. Further, in the subsequent analysis we will heavily rely on the fact that the solution is also uniformly bounded in $L^\infty(I \times \Omega)$ independent of k.

Proposition 16. *For the solution $u_k \in U_k$ of (18) the following stability estimates hold*

$$\|u_k\|_{L^\infty(I\times\Omega)} \le c\big(\|qg\|_{L^p(I\times\Omega)} + \|u_0\|_{L^\infty} + \|d(\cdot,\cdot,0)\|_{L^p(I\times\Omega)}\big) \tag{19}$$

$$\|u_k\|_{L^\infty(I,V)} \le c\big(\|qg\|_I^2 + \|u_0\|_V + \|d(\cdot,\cdot,0)\|_I\big), lo \tag{20}$$

where $p > 2$.

Proof. For the boundedness of $\|u_k\|_{L^\infty(I\times\Omega)}$ we refer to [31, Theorem 3.1]. The estimate (20) follows by [31, Theorem 3.2]. □

As for the continuous case, we now introduce the semidiscrete control-to-state map

$$S_k \colon L^\infty(I, \mathbb{R}^m) \to U_k,$$

associating to any given q the solution $u_k(q) := S_k(q)$ of (18). The state constraint is

$$F_k := (\cdot, w) \colon U_k \to U_k(\mathbb{R}),$$

the concatenation of the control-to-state map with F_k is denoted by

$$G_k = (F_k \circ S_k) \colon L^\infty(I, \mathbb{R}^m) \to U_k(\mathbb{R}),$$

and the set of feasible controls by

$$Q_{k,\text{feas}} = \{q \in Q_{\text{ad}} \,|\, G_k(q) \le 0\}.$$

Then, the semidiscrete version of $(\mathbb{P})$ reads

$$\min j_k(q) := J(q, S_k(q)) \qquad \text{s.t. } q \in Q_{k,\text{feas}}. \tag{$\mathbb{P}_k$}$$

Arguing as in the continuous case, we have that S_k and G_k are of class C^2.

Lemma 17. *The operator $S_k\colon L^\infty(I,\mathbb{R}^m) \to U_k$ is of class C^2. For $u_k = S_k(q)$ and $p \in L^\infty(I,\mathbb{R}^m)$, its first derivative $S_k^{'}(q)p := v_{k,p}$, in direction p, is the solution of*

$$B(v_{k,p},\varphi_k) + (\partial_u d(\cdot,\cdot,u_k)v_{k,p},\varphi_k)_I = (pg,\varphi_k)_I,\ \forall \varphi_k \in U_k. \tag{21}$$

For $p_1,p_2 \in L^\infty(I,\mathbb{R}^m)$, its second derivative $S_k^{''}(q)p_1p_2 = v_{k,p_1p_2}$, in the directions p_1,p_2, is the solution of

$$B(v_{k,p_1p_2},\varphi_k) + (\partial_u d(\cdot,\cdot,u_k)v_{k,p_1p_2},\varphi_k)_I = -(\partial_{uu} d(\cdot,\cdot,u_k)v_{k,p_1}v_{k,p_2},\varphi_k)_I, \tag{22}$$

for all $\varphi_k \in U_k$.

Similarly to S, also for S_k and its first derivative there holds a Lipschitz property.

Lemma 18. *For $q_1,q_2,p \in L^\infty(I,\mathbb{R}^m)$ there holds*

$$\|S_k(q_1) - S_k(q_2)\|_I \leq c\|q_1-q_2\|_{L^2(I,\mathbb{R}^m)}, \tag{23a}$$

$$\|S_k^{'}(q_1)p - S_k^{'}(q_2)p\|_I \leq c\|q_1-q_2\|_{L^2(I,\mathbb{R}^m)}\|p\|_{L^2(I,\mathbb{R}^m)}. \tag{23b}$$

$$\|S_k^{'}(q_1)p - S_k^{'}(q_2)p\|_{L^\infty(I,H)} \leq c\|q_1-q_2\|_{L^2(I,\mathbb{R}^m)}\|p\|_{L^2(I,\mathbb{R}^m)}. \tag{23c}$$

Proof. The proof is analogous to the one of Lemma 7 utilizing [31, Lemma 3.1]. □

3.2. **Space discretization.** We consider a family $\mathcal{T}_h$ of subdivisions of Ω consisting of closed triangles or quadrilaterals (tetrahedral or hexahedral in dimension three) T which are affine equivalent to their reference elements. The union of these elements $\Omega_h = \operatorname{int}\big(\bigcup_{T\in\mathcal{T}_h} T\big)$ is assumed to be such that the vertices on $\partial\Omega_h$ are located on $\partial\Omega$. We assume the family $\mathcal{T}_h$ to be quasi-uniform and shape regular in the sense of [4] denoting by h_T the diameter of T and $h := \max_{T\in\mathcal{T}_h} h_T$. Then, we define the conforming finite element space $V_h \subset V$ as the space of piecewise linear functions with respect to $\mathcal{T}_h$ with the canonical extension $v\big|_{\Omega\setminus\Omega_h} \equiv 0$ for any $v \in V_h$. Moreover, we assume that the sequence of spatial meshes is such that the L^2-projection onto V_h is stable with respect to the H^1-norm, for conditions ensuring this stability see, e.g., [1]. Then, the discrete state and trial spaces are given by

$$U_{kh} = U_{kh}(V_h) = \big\{\varphi_{kh} \in L^2(I,V_h)\,\big|\,\varphi_{kh,n} = \varphi_{kh}|_{I_n} \in \mathcal{P}_0(I_n,V_h),\, n = 1,...,N\big\},$$

and the discrete state equation reads: for $q \in L^\infty(I,\mathbb{R}^m)$, find $u_{kh} = u_{kh}(q) \in U_{kh}$ such that

$$B(u_{kh},\varphi_{kh}) + (d(\cdot,\cdot,u_{kh}),\varphi_{kh})_I = (qg,\varphi_{kh})_I + (u_0,\varphi_{kh,1}),\ \forall \varphi_{kh} \in U_{kh}. \tag{24}$$

Just as in the semidiscrete case, we have the following stability estimates, see [31, Theorem 4.1]. We remark again that the uniform boundedness of u_{kh} independent of the discretization parameters k,h will play a crucial role.

Proposition 19. *For the solution $u_{kh} \in U_{kh}$ of (24) the following stability estimates holds*

$$\|u_{kh}\|_{L^\infty(I\times\Omega)} \leq c\big(\|qg\|_{L^p(I\times\Omega)} + \|\Pi_h u_0\|_{L^\infty(\Omega)} + \|d(\cdot,\cdot,0)\|_{L^p(I\times\Omega)}\big) \tag{25}$$

$$\|u_{kh}\|_{L^\infty(I,V)} \leq c\big(\|qg\|_I^2 + \|\Pi_h u_0\|_V + \|d(\cdot,\cdot,0)\|_I\big), \tag{26}$$

where $p > 2$ and $\Pi_h\colon V \to V_h$ is the L^2-projection in space.

Next, we introduce the discrete control-to-state map $S_{kh}\colon L^\infty(I,\mathbb{R}^m) \to U_{kh}$, the discrete state constraint $F_{kh} : U_{kh} \to U_{kh}(\mathbb{R})$, the C^2-functional $G_{kh} = (F_{kh} \circ S_{kh})$, and the set of feasible controls $Q_{kh,\text{feas}} = \{q \in Q_{\text{ad}} \,|\, G_{kh}(q) \le 0\}$.

The discrete problem reads

$$(\mathbb{P}_{kh}) \qquad \min j_{kh}(q) := J(q, S_{kh}(q)) \qquad \text{s.t. } q \in Q_{kh,\text{feas}}.$$

Similar to the semidiscrete case, first and second derivatives of the discrete control-to-state map S_{kh} are defined via (21), (22), respectively, with test functions from U_{kh}. Further, for S_{kh} and its first derivative there holds the Lipschitz property analog to Lemma 18, compare with [31, Lemma 4.1].

We formulate now standard KKT optimality conditions for problem $(\mathbb{P}_{kh})$. These conditions will be justified after the introduction of an auxiliary problem in Section 5. In particular, we will show in Lemma 30 that, for k, h small enough, the Slater point for (1) is also a Slater point for $(\mathbb{P}_{kh})$.

Theorem 20. *Let $\bar{u}_{kh} \in Q_{kh,feas}$ be a local solution of $(\mathbb{P}_{kh})$ with $\bar{u}_{kh} \in U_{kh}$ the associated state. Then, under Assumption 10, for k, h sufficiently small there exists a Lagrange multiplier $\bar{\mu}_{kh} \in U_{kh}(\mathbb{R})^* \cap C(\bar{I})^*$ and an adjoint state $\bar{z}_{kh} \in U_{kh}$ such that*

$$\begin{aligned}
&B(\bar{u}_{kh}, \varphi) + (d(\cdot,\cdot,\bar{u}_{kh}), \varphi)_I = (\bar{q}_{kh} g, \varphi)_I + (u_0, \varphi_{kh,1}) && \forall \varphi \in U_{kh},\\
&B(\varphi, \bar{z}_{kh}) + (\varphi, \partial_u d(\cdot,\cdot,\bar{u}_{kh})\bar{z}_{kh}) = (\bar{u} - u_d, \varphi)_I + \langle \bar{\mu}_{kh}, F_{kh}(\varphi)\rangle && \forall \varphi \in U_{kh},\\
&\alpha(\bar{q}_{kh}, q - \bar{q}_{kh})_{L^2(I)} + (\bar{z}_{kh}, (q - \bar{q}_{kh})g)_I \ge 0 && \forall q \in Q_{kh,feas},\\
&\langle F_{kh}(\bar{u}_{kh}), \bar{\mu}_{kh}\rangle = 0, \ \bar{\mu} \ge 0,
\end{aligned}$$

where $\langle \cdot,\cdot\rangle$ denotes the duality pairing between $U_{kh}(\mathbb{R})^*$ and $U_{kh}(\mathbb{R})$. Further, the Lagrange multiplier can be represented as an element of $C(\bar{I})^*$ by

$$\langle v, \bar{\mu}_{kh}\rangle = \sum_{n=1}^{N} \frac{\mu_{kh,n}}{k_n} \int_{I_n} v(t)\mathrm{d}t, \ \forall v \in C(\bar{I}) \cup U_{kh}(\mathbb{R}).$$

4. The state equation

In this section, we are interested in the derivation of $L^\infty(I,H)$ error estimates for the solutions of the continuous, semidiscrete and discrete state equation. The technique behind these estimates is based on a duality argument requiring, at any level of discretization, the introduction of auxiliary linearized problems. This approach has been used in [26] for a linear parabolic state equation. We now intend to extend it to the semilinear parabolic case adapting an idea of [32] for semilinear elliptic equations.

4.1. **Error estimates for the temporal discretization.**

In a first step, we introduce the backward uncontrolled linearized counterpart of the state equation. For a given fixed $q \in L^\infty(I,\mathbb{R}^m)$, we consider u, u_k solution of (2), (18), respectively, and we define

$$\tilde{d} = \begin{cases} \frac{d(u(t,x)) - d(u_k(t,x))}{u(t,x) - u_k(t,x)} & \text{if } u(t,x) \ne u_k(t,x)\\ 0 & \text{else.}\end{cases}$$

Then, we consider

$$\begin{aligned} -(\varphi, \partial_t w)_I + (\nabla\varphi, \nabla w)_I + (\varphi, \tilde{d} w)_I &= 0,\\ w(T) &= w_T, \end{aligned} \tag{27}$$

for any $\varphi \in W(0,T) \cap H^1(I,H)$, with $w_T \in H$.

Denoting by $\hat{I} = (0,\hat{t})$, $\hat{t} \in (0,T)$, a truncated time interval, we introduce

$$\begin{aligned} -(\varphi, \partial_t \hat{w})_{\hat{I}} + (\nabla\varphi, \nabla\hat{w})_{\hat{I}} + (\varphi, \tilde{d}\hat{w})_{\hat{I}} &= 0, \\ w(\hat{t}) &= w_T. \end{aligned} \tag{28}$$

Further, the semidiscrete counterpart of (27), for any $\varphi_k \in U_k$, reads

$$B(\varphi_k, w_k) + (\varphi_k, \tilde{d}w_k)_I = (\varphi_{k,N}, w_T). \tag{29}$$

Before starting, we observe that, for any $\varphi_k \in U_k$, the following relations hold

$$B(u-u_k, \varphi_k) = -(d(u) - d(u_k), \varphi_k)_I = -((u-u_k)\tilde{d}, \varphi_k)_I \tag{30}$$

$$B(\varphi_k, w - w_k) = -(\varphi_k, (w-w_k)\tilde{d})_I. \tag{31}$$

In the following analysis, we will need negative norm estimates for the error between the solutions of (27), (28), and (29). These estimates will be used to derive the error at the time nodal points and inside the time intervals I_n. Their derivation follows exactly as in [26, Lemma 5.1, Lemma 5.2], with minor changes due to the presence of the linearization $\tilde{d}$ of the semilinear term, and therefore it is omitted. The crucial point is the boundedness of $\tilde{d}$ in $L^\infty(I \times \Omega)$ which follows from the Lipschitz continuity of $d(\cdot)$ and the regularity of $u, u_k \in L^\infty(I \times \Omega)$.

For the convenience of the reader, the analog to [26, Lemma 5.1, Lemma 5.2] in our case reads as follows.

Lemma 21. *For the error between the solutions w, $\hat{w}$, and w_k of* (27), (28), *and* (29), *respectively, there holds*

$$\begin{aligned} \|w - \hat{w}\|_{L^1(\hat{I},H)} + \|w(0) - \hat{w}(0)\|_{H^{-2}(\Omega)} &\leq ck\Big(\log\frac{T}{k}\Big)^{\frac{1}{2}} \|w_T\|, \\ \|w - w_k\|_{L^1(I,H)} + \|w(0) - w_{k,1}\|_{H^{-2}(\Omega)} &\leq ck\Big(\log\frac{T}{k}\Big)^{\frac{1}{2}} \|w_T\|. \end{aligned}$$

With these estimates at hand, we are ready to derive the main result of the section.

Theorem 22. *For given $qg \in L^\infty(I,H)$ and $u_0 \in H^2(\Omega) \cap V$, let $u \in U$ and $u_k \in U_k$ be the solution of* (2) *and* (18), *respectively. Then, there holds*

$$\|u - u_k\|_{L^\infty(I,H)} \leq ck\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} \Big(\|qg\|_{L^\infty(I,H)} + \|u_0\|_{H^2(\Omega)} + \|d(0)\|_{L^\infty(I\times\Omega)}\Big).$$

Proof. Let $e_k = u - u_k$ denote the error arising from the dG(0)-time-discretization. In every time interval, we split the error into

$$\|e_k\|_{L^\infty(I_n,H)} \leq \underbrace{\|u(\cdot) - u(t_n)\|_{L^\infty(I_n,H)}}_{(a_1)} + \underbrace{\|u(t_n) - u_k(\cdot)\|_{L^\infty(I_n,H)}}_{(a_2)}, \tag{32}$$

and we analyze the two terms $(a_1), (a_2)$ separately. Then, taking the maximum over all $n = 1, ..., N$, we obtain the assertion. Without loss of generality, we consider the last time interval I_N. For an arbitrary time interval I_n, we consider (27) on $I = (0, t_n)$, (28) on $\hat{I} = (0,\hat{t})$ for $\hat{t} \in (t_{n-1}, t_n]$, and the proof follows mutatis mutandis, observing that $0 \leq \log(t_n/k) \leq \log(T/k)$.

(a_1) For a generic fixed time $\hat{t} \in I_N$, we start the derivation considering the interpolation error $u(\hat{t}) - u(t_N)$.

Consider the solutions w and $\hat{w}$ to (27) and (28) on $\hat{I} = (0, \hat{t})$, respectively, with terminal value $w_T = u(\hat{t}) - u(t_N)$. Integration by parts in time of (27) and (28) gives

$$\begin{aligned} -(\varphi(T), w(T)) + (\varphi(0), w(0)) + (\partial_t \varphi, w)_I + (\nabla\varphi, \nabla w)_I + (\varphi, \tilde{d}w)_I = 0, \\ -(\varphi(\hat{t}), \hat{w}(\hat{t})) + (\varphi(0), \hat{w}(0)) + (\partial_t \varphi, \hat{w})_{\hat{I}} + (\nabla\varphi, \nabla \hat{w})_{\hat{I}} + (\varphi, \tilde{d}\hat{w})_{\hat{I}} = 0, \end{aligned}$$

for any $\varphi \in W(0,T) \cap H^1(I, H)$.

In particular, setting $\varphi = u$, the state equation (2) yields

$$\begin{aligned} -(u(T), w(T)) + (u(0), w(0)) + (qg, w)_I - (d(u), w)_I + (u, \tilde{d}w)_I = 0, \\ -(u(\hat{t}), \hat{w}(\hat{t})) + (u(0), \hat{w}(0)) + (qg, \hat{w})_{\hat{I}} - (d(u), \hat{w})_{\hat{I}} + (u, \tilde{d}\hat{w})_{\hat{I}} = 0. \end{aligned}$$

By definition $w(T) = w(\hat{t}) = w_T$, subtracting the equalities above, we get

$$\begin{aligned} (u(\hat{t}) - u(T), w_T) =& (u(0), \hat{w}(0) - w(0)) + (qg, \hat{w} - w)_{\hat{I}} - (qg, w)_{I \setminus \hat{I}} \\ & + \underbrace{\big(u, (\hat{w} - w)\tilde{d}\big)_{\hat{I}}}_{(b_1)} \underbrace{-(u, \tilde{d}w)_{I \setminus \hat{I}}}_{(b_2)} \\ & + \underbrace{(d(u), w - \hat{w})_{\hat{I}}}_{(b_3)} + \underbrace{(d(u), w)_{I \setminus \hat{I}}}_{(b_4)}. \end{aligned} \tag{33}$$

We analyze the terms separately.

(b_1) Due to the stability in $L^\infty(I \times \Omega)$ of the solutions of (2), (18) and the Lipschitz continuity of d, we observe that $\|\tilde{d}\|_{L^\infty(I\times\Omega)} \le c$. Therefore,

$$\big(u, (\hat{w} - w)\tilde{d}\big)_{\hat{I}} \le c\|u\|_{L^\infty(I,H)} \|\hat{w} - w\|_{L^1(\hat{I},H)}.$$

(b_2) Exploiting again the boundedness of $\tilde{d}$ in $L^\infty(I \times \Omega)$, and $|T - \hat{t}| \le k$, we have

$$\begin{aligned} -(u, \tilde{d}w)_{I\setminus\hat{I}} &\le \Big| \int_{\hat{t}}^T (u, \tilde{d}w) \mathrm{d}t \Big| \\ &\le ck\|u\|_{L^\infty(I,H)} \|w\|_{L^\infty(I,H)}. \end{aligned}$$

(b_3) The Lipschitz property of $d(u)$ and the boundedness of $d(0)$ in $L^\infty(\hat{I}, H)$ yield

$$\begin{aligned} (d(u), w - \hat{w})_{\hat{I}} &= (d(u) - d(0), w - \hat{w})_{\hat{I}} + (d(0), w - \hat{w})_{\hat{I}} \\ &\le \|d(u) - d(0)\|_{L^\infty(\hat{I},H)} \|w - \hat{w}\|_{L^1(\hat{I},H)} \\ &\quad + \|d(0)\|_{L^\infty(\hat{I},H)} \|w - \hat{w}\|_{L^1(\hat{I},H)} \\ &\le c\big(\|u\|_{L^\infty(\hat{I},H)} + \|d(0)\|_{L^\infty(\hat{I},H)}\big) \|w - \hat{w}\|_{L^1(\hat{I},H)}. \end{aligned}$$

(b_4) Using the same argument as for (b_3), we conclude

$$\begin{aligned} (d(u), w)_{I\setminus\hat{I}} &= (d(u) - d(0), w)_{I\setminus\hat{I}} + (d(0), w)_{I\setminus\hat{I}} \\ &\le ck\big(\|u\|_{L^\infty(I\times\Omega)} + \|d(0)\|_{L^\infty(I\times\Omega)}\big) \|w\|_{L^\infty(I,H)}. \end{aligned}$$

Going back to (33), but now with the value $w_T = u(\hat{t}) - u(T)$ we obtain

$$\|u(\hat{t}) - u(T)\|^2 \le c\Big(\|w - \hat{w}\|_{L^1(\hat{I},H)} + \|(w-\hat{w})(0)\|_{H^{-2}(\Omega)} + k\|w\|_{L^\infty(I,H)}\Big) \cdot \Big(\|qg\|_{L^\infty(I,H)} + \|u_0\|_{H^2(\Omega)} + \|d(0)\|_{L^\infty(I\times\Omega)} + \|u\|_{L^\infty(I,H)} + \|u\|_{L^\infty(I\times\Omega)}\Big).$$

Using the stability of the solution w of (27), i.e., $\|w\|_{L^\infty(I,H)} \le c\|w_T\|$, see, e.g., [26, Theorem 5.3], Proposition 2, Lemma 21 and after dividing by $\|w_T\| = \|u(\hat{t}) - u(T)\|$, we conclude

$$\|u(\hat{t}) - u(T)\| \le ck \log\Big(\frac{T}{k} + 1\Big)^{\frac{1}{2}} \Big(\|q\|_{L^\infty(I,\mathbb{R}^m)}\|g\|_H + \|u_0\|_{H^2(\Omega)} + \|d(0)\|_{L^\infty(I\times\Omega)}\Big). \tag{34}$$

(a_2) To obtain the error of the $dG(0)$-discretization inside the time interval I_N, we set $w_T = u(t_N) - u_{k,N} = u(T) - u_{k,N}$ in (27) and in (29). Then, for any $\phi \in U_k + (L^2(I,V) \cap H^1(I,H))$ it holds

$$B(\varphi, w) + (\varphi, \tilde{d}w)_I = B(\varphi, w_k) + (\varphi, \tilde{d}w_k)_I = (\varphi_N, u(T) - u_{k,N}).$$

In particular, testing the relation above with $\varphi = u - u_k$ and making use of (30) and (31), we have

$$\begin{aligned}
\|u(T) - u_{k,N}\|^2 &= B(u-u_k, w) + (u-u_k, \tilde{d}w)_I \\
&= B(u-u_k, w-w_k) - ((u-u_k)\tilde{d}, w_k)_I + ((u-u_k)\tilde{d}, w)_I \\
&= B(u, w-w_k) + (u_k, (w-w_k)\tilde{d})_I + ((u-u_k)\tilde{d}, w-w_k)_I \\
&= (qg, w-w_k)_I + (u_0, w(0) - w_k(0)) \underbrace{-(d(u), w-w_k)_I}_{(c_1)} \\
&\quad + \underbrace{(u_k, (w-w_k)\tilde{d})_I}_{(c_2)} + \underbrace{((u-u_k)\tilde{d}, w-w_k)_I}_{(c_3)},
\end{aligned}$$

where in the last step we used (2).

We consider the three terms $(c_1) - (c_3)$ separately.

(c_1) Observing that $L^\infty(I,V) \hookrightarrow L^\infty(I,H)$, the stability result (4) of the solution u of (2), the Lipschitz continuity of $d(\cdot)$, and the boundedness of $d(0)$ in $L^\infty(I,H)$ yield

$$\begin{aligned}
-(d(u), w-w_k)_I &\le \Big(\|d(u) - d(0)\|_{L^\infty(I,H)} + \|d(0)\|_{L^\infty(I,H)}\Big)\|w-w_k\|_{L^1(I,H)} \\
&\le c\Big(\|u\|_{L^\infty(I,H)} + \|d(0)\|_{L^\infty(I,H)}\Big)\|w-w_k\|_{L^1(I,H)} \\
&\le c\Big(\|qg\|_I + \|u_0\|_V + \|d(0)\|_{L^\infty(I,H)}\Big)\|w-w_k\|_{L^1(I,H)}.
\end{aligned}$$

(c_2) The boundedness of $\tilde{d}$ in $L^\infty(I\times\Omega)$ and the stability result of the semidiscrete equation of Proposition 16 yield

$$\begin{aligned}
(u_k, \tilde{d}(w-w_k))_I &\le \|u_k\|_{L^\infty(I,H)}\|w-w_k\|_{L^1(I,H)} \\
&\le c\Big(\|qg\|_I + \|u_0\|_V + \|d(0)\|_I\Big)\|w-w_k\|_{L^1(I,H)}.
\end{aligned}$$

(c_3) From the Lipschitz continuity of $d(\cdot)$, as well as the definition and boundedness of $\tilde{d}$, it follows

$$\begin{aligned}(\tilde{d}(u-u_k), w-w_k)_I &= (d(u)-d(u_k), w-w_k)_I \\ &= (d(u)-d(0), w-w_k)_I + (d(0)-d(u_k), w-w_k)_I \\ &\le c\Big(\|u\|_{L^\infty(I,H)} + \|u_k\|_{L^\infty(I,H)}\Big)\|w-w_k\|_{L^1(I,H)} \\ &\le c\Big(\|qg\|_I + \|u_0\|_V + \|d(0)\|_I\Big)\|w-w_k\|_{L^1(I,H)},\end{aligned}$$

where in the last step, we used the stability of the solutions u, u_k of (2) and (18), from Proposition 2 and Proposition 16, respectively.

Summing up, for the error inside the time interval, we obtain

$$\begin{aligned}\|u(T)-u_{k,N}\|^2 \le c\Big(&\|w-w_k\|_{L^I(I,H)} + \|w(0)-w_k(0)\|_{H^{-2}(\Omega)}\Big) \\ &\cdot\Big(\|qg\|_{L^\infty(I,H)} + \|u_0\|_{H^2(\Omega)} + \|d(0)\|_{L^\infty(I,H)}\Big).\end{aligned} \tag{35}$$

In conclusion, combining (34) with (35) and thanks to Lemma 21, we obtain the assertion dividing by $\|w_T\| = \|u(T)-u_{k,N}\|$. □

4.2. **Error estimates for the spatial discretization.** We develop error estimates for the spatial discretization of the problem using similar steps as in the semidiscrete case. The linearization of d now reads

$$\hat{d} = \begin{cases}\frac{d(u_k(t,x))-d(u_{kh}(t,x))}{u_k(t,x)-u_{kh}(t,x)} & \text{if } u_k(t,x) \neq u_{kh}(t,x) \\ 0 & \text{else.}\end{cases}$$

We remark that, thanks to the Lipschitz continuity of $d(\cdot)$, the linearized term $\hat{d}$ is bounded in $L^\infty(I\times\Omega)$.

We introduce the discrete counterpart of (27) with $\hat{d}$ instead of $\tilde{d}$. Find $w_{kh} \in U_{kh}$ such that

$$B(\varphi_{kh}, w_{kh}) + (\varphi_{kh}, \hat{d}w_{kh})_I = (\varphi_N, w_T), \tag{36}$$

for any $\varphi_{kh} \in U_{kh}$, with $w_T \in H$.

We also consider the auxiliary problem (29) with $\hat{d}$ instead of $\tilde{d}$, namely, find $w_k \in U_k$ such that

$$B(\varphi_k, w_k) + (\varphi_k, \hat{d}w_k)_I = (\varphi_{k,N}, w_T), \tag{37}$$

for any $\varphi_k \in U_k$.

We observe that for any $\varphi_{kh} \in U_{kh}$ the following relations hold

$$B(u_k-u_{kh}, \varphi_{kh}) = -(d(u_k)-d(u_{kh}), \varphi_{kh})_I = -((u_k-u_{kh})\hat{d}, \varphi_{kh})_I \tag{38}$$

$$B(\varphi_{kh}, w_k-w_{kh}) = -(\varphi_{kh}, (w_k-w_{kh})\hat{d})_I. \tag{39}$$

As for the error in the dG(0)-semidiscretization, also here we will employ a duality argument requiring estimates for the error between the solutions of (37) and (36). The proof follows the lines of [26, Lemma 5.8 and Lemma 5.9] with the obvious modifications due to the presence of $\hat{d}$. For the convenience of the reader, we collect such estimates from [26] below.

Lemma 23. *For the error between the solutions w_k, w_{kh} of (37), (36), respectively, there holds*

$$\|w_{k,1} - w_{kh,1}\|_{H^{-2}(\Omega)} + T\|w_{k,1} - w_{kh,1}\| \le ch^2\|w_T\|. \tag{40}$$

Theorem 24. *For given $qg \in L^\infty(I,H)$ and $u_0 \in H^2(\Omega)\cap V$, let $u_k \in U_k$ and $u_{kh} \in U_{kh}$ be the solutions of (18) and (24), respectively. Then, there holds*

$$\|u_k - u_{kh}\|_{L^\infty(I,H)} \le ch^2\Big(\log\frac{T}{k}+1\Big)\Big(\|qg\|_{L^\infty(I,H)} + \|u_0\|_{H^2(\Omega)} + \|d(0)\|_{L^\infty(I\times\Omega)}\Big).$$

Proof. Since both u_k, u_{kh} are constant on each time interval I_n, we can equivalently show the estimate on a single time interval I_n and with no loss of generality we consider the last time interval only. For an arbitrary time interval I_n, we consider (36) and (37) on $I=(0,t_n)$ and, noting that $0 \le \log(t_n/k) \le \log(T/k)$, the proof follows mutatis mutandis.

Proceeding as in the proof of Theorem 22, we set $w_T = u_{k,N} - u_{kh,N}$ in (36) and (37). Then, using (38) and (39), we have

$$\begin{aligned}\|u_{k,N} - u_{kh,N}\|^2 &= B(u_k - u_{kh}, w_k) + (u_k - u_{kh}, \hat{d}w_k)_I\\ &= B(u_k - u_{kh}, w_k - w_{kh}) - (\hat{d}(u_k - u_{kh}), w_{kh})_I + (\hat{d}(u_k - u_{kh}), w_k)_I\\ &= B(u_k, w_k - w_{kh}) + (u_{kh}, \hat{d}(w_k - w_{kh}))_I + (\hat{d}(u_k - u_{kh}), w_k - w_{kh})_I\\ &= (qg, w_k - w_{kh})_I + (u_0, w_{k,1} - w_{kh,1}) \underbrace{-(d(u_k), w_k - w_{kh})_I}_{(a_1)}\\ &\quad + \underbrace{(u_{kh}, \hat{d}(w_k - w_{kh}))_I}_{(a_2)} + \underbrace{(\hat{d}(u_k - u_{kh}), w_k - w_{kh})_I}_{(a_3)},\end{aligned}$$

where in the last step we used (18). We analyze the three terms separately.

(a_1) The Lipschitz continuity of $d(\cdot)$ and the boundedness of $d(0)$ in $L^\infty(I,H)$, give

$$\begin{aligned}-(d(u_k), w_k - w_{kh})_I &\le c\Big(\|d(u_k) - d(0)\|_{L^\infty(I,H)} + \|d(0)\|_{L^\infty(I,H)}\Big)\\ &\qquad \cdot \|w_k - w_{kh}\|_{L^1(I,H)}\\ &\le c\Big(\|u_k\|_{L^\infty(I,H)} + \|d(0)\|_{L^\infty(I,H)}\Big)\|w_k - w_{kh}\|_{L^1(I,H)}.\end{aligned}$$

(a_2) Recalling that $\hat{d}$ is bounded, we have

$$(u_{kh}, \hat{d}(w_k - w_{kh}))_I \le c\|u_{kh}\|_{L^\infty(I,H)}\|w_k - w_{kh}\|_{L^1(I,H)}.$$

(a_3) For the last term, we rely again on the Lipschitz continuity of $d(\cdot)$ to conclude

$$\begin{aligned}(\hat{d}(u_k - u_{kh}), w_k - w_{kh})_I &= (d(u_k) - d(u_{kh}), w_k - w_{kh})_I\\ &\le c\Big(\|u_k\|_{L^\infty(I,H)} + \|u_{kh}\|_{L^\infty(I,H)}\Big)\|w_k - w_{kh}\|_{L^1(I,H)}.\end{aligned}$$

We now combine the previous inequalities and, thanks to the stability estimates (20) and (26), we obtain

$$\begin{aligned}\|u_{k,N} - u_{kh,N}\|^2 &\le c\Big(\|w_k - w_{kh}\|_{L^1(I,H)} + \|w_{k,1} - w_{kh,1}\|_{H^{-2}(\Omega)}\Big)\\ &\qquad \cdot \Big(\|qg\|_{L^\infty(I,H)} + \|u_0\|_{H^2(\Omega)} + \|d(0)\|_{L^\infty(I\times\Omega)}\Big).\end{aligned}$$

Noting that the L^2-estimate in (40) remains true on shorter intervals, it follows with $\tau_{n,k} = T - t_{n-1}$

$$\begin{aligned}\|w_k - w_{kh}\|_{L^1(I,H)} &\le \sum_{n=1}^{N} k_n \tau_{k,n}^{-1} \max_{n=1,\dots,N} \left(\tau_{k,n}\|w_{k,n} - w_{kh,n}\|\right) \\ &\le ch^2\Big(\log\frac{T}{k} + 1\Big)\|w_T\|.\end{aligned}$$

and, using Lemma 23, dividing by $\|w_T\| = \|u_{k,N} - u_{kh,N}\|$, we obtain the assertion. □

5. Convergence analysis

In this section, we focus on the main result of this paper. We show that for any local solution $\bar{q}$ of the continuous problem satisfying KKT-conditions and SSCs, there exists a sequence of local solutions $\bar{q}_{kh}$ of $(\mathbb{P}_{kh})$ converging to $\bar{q}$. To analyze the errors induced by the discretization, we use the so called *two-way feasibility* argument, see, e.g., [16, 29]. In this method the linearized Slater point q_γ from Assumption 10 is used to construct sequences of controls (competitors) which are feasible for the continuous and discrete problem, respectively. If the problem is linear, these sequences of feasible competitors can be used in the first order necessary and sufficient conditions to obtain convergence of the discrete problem. In the semilinear case, due to the presence of the linearized term, the complementary slackness condition cannot be used as in the linear setting. Therefore, the feasible controls have to be used in combination with second order information; in particular, in the quadratic growth condition (17) arising from the second order sufficient conditions. This approach has been used in the recent paper [30] for the semilinear elliptic case in combination with a localization argument as in [10]. We now intend to extend that approach to our semilinear parabolic optimal control problem with state constraints.

In the following analysis, we will introduce auxiliary problems in a neighborhood of the optimal local solution $\bar{q}$. To this end, we denote with $r > 0$ a radius, to be chosen conveniently later, and we define

$$\begin{aligned}Q^r &:= \{q \in Q_{\text{ad}} \,|\, \|q - \bar{q}\|_{L^2(I,\mathbb{R}^m)} \le r\}, \\ Q^r_{\text{feas}} &:= \{q \in Q^r \,|\, G(q) \le 0\}.\end{aligned}$$

Then, the continuous auxiliary problem reads

$$(\mathbb{P}^r) \qquad \min\, j(q) := J(q, S(q)) \quad \text{s.t. } q \in Q^r_{\text{feas}}.$$

Due to the SSCs, for r sufficiently small, the unique global solution of $(\mathbb{P}^r)$ coincides with the selected local solution $\bar{q}$ of $(\mathbb{P})$. The value of introducing the auxiliary problem lies in the fact that a discretization of $(\mathbb{P}^r)$, following later, will provide a sequence of solutions converging to the selected local optimum.

Assumption 25. *We assume that q_γ satisfying Slater's regularity condition* (10)*, is close enough to $\bar{q} \in Q_{feas}$, meaning*

$$\|q_\gamma - \bar{q}\|_{L^2(I,\mathbb{R}^m)} \le \frac{r}{2}. \tag{42}$$

The fact that q_γ is in a neighborhood of $\bar{q}$ is a reasonable assumption. Indeed, as observed in [30, Section 2], given any Slater point q_γ with parameter γ one can construct a Slater point $q_\gamma^r = \bar q + t(q_\gamma - \bar q)$ close to $\bar q$ with a parameter $\gamma(r) = t\gamma \simeq r\gamma$ with $t = \min\{1, r/2\|q_\gamma - \bar q\|\}$. Hence, one has that (10) holds with γ replaced by $t\gamma$. Further, after showing that $(\mathbb{P}_{kh})$ admits local solutions and thanks to the following Lemma 30, we will see that it is reasonable to assume that (42) holds also for the discrete problem $(\mathbb{P}_{kh})$, namely

$$\|q_\gamma - \bar q_{kh}\|_{L^2(I,\mathbb{R}^m)} \le \frac{r}{2}. \tag{43}$$

We will now define three constants c_1, c_2, c_3, independent of the discretization parameter k, h. These constants are given by

$$\sup_{q\in B_{\frac{r}{2}}(\bar q)} \|(\omega, u_{kh}(q) - u(q))\|_{L^\infty(I)} \le c_1\Big(k\Big(\ln\frac{T}{k}+1\Big)^{\frac12} + h^2\Big(\ln\frac{T}{k}+1\Big)\Big),$$

$$\sup_{q\in B_{\frac{r}{2}}(\bar q)} \|G^{''}(q_k)\|_{\mathcal{L}(L^2(I,\mathbb{R}^m)^2;L^\infty(I))}, \sup_{q\in B_{\frac{r}{2}}(\bar q)} \|G^{''}_{kh}(q_k)\|_{\mathcal{L}(L^2(I,\mathbb{R}^m)^2;L^\infty(I))} \le c_2,$$

$$\sup_{q\in B_{\frac{r}{2}}(\bar q)} \|(G'_{kh}(q) - G'(\bar q))(q_\gamma - \bar q)\|_{L^\infty(I)} \le c_3\Big(k\Big(\ln\frac{T}{k}+1\Big)^{\frac12} + h^2\Big(\ln\frac{T}{k}+1\Big) + \frac{r^2}{2}\Big),$$

where $B_{\frac{r}{2}}(\bar q)$ denotes an $L^2(I,\mathbb{R}^m)$ ball centered in $\bar q$ with radius $\frac{r}{2}$.

Remark 26. *To see that these constants are independent of k, h proceed as follows*

- *For the constant c_1, we notice that this error can be estimated by the discretization errors obtained by Theorem 22, 24 noting that by the proof of these theorems the constant in the error estimates remains bounded on $B_{\frac{r}{2}}(\bar q)$.*
- *The constant c_2 is a consequence of G being a C^2 functional together with a discretization error bound for G''_{kh}.*
- *For the constant c_3, we notice that*

  $$F(\varphi) = F_{kh}(\varphi) = \int_\Omega \varphi(t,x)\omega(x)\mathrm{d}x, \quad \varphi \in W(0,T)\cup U_{kh}$$

  is linear and consequently the error satisfies

  $$\begin{aligned}(G'_{kh}(q) - G'(\bar q))(q_\gamma - \bar q) &= F_{kh}\big(S^{'}_{kh}(q)(q_\gamma - \bar q)\big) - F\big(S^{'}(\bar q)(q_\gamma - \bar q)\big)\\ &= \Big(\omega, \big(S^{'}_{kh}(q) - S^{'}(\bar q)\big)(q_\gamma - \bar q)\Big)\\ &= \Big(\omega, \big(S^{'}_{kh}(q) - S^{'}(q) + S^{'}(q) - S^{'}(\bar q)\big)(q_\gamma - \bar q)\Big)\\ &\le c\Big(\|\big(S^{'}_{kh}(q) - S^{'}(q)\big)(q_\gamma - \bar q)\|_{L^\infty(I,H)}\\ &\qquad + \|q - \bar q\|_{L^2(I,\mathbb{R}^m)}\|q_\gamma - \bar q\|_{L^2(I,\mathbb{R}^m)}\Big),\end{aligned}$$

  where in the last step we used the stability of S", i.e., (7e). *The remaining term is a discretization error that can be estimated by [26, Corollary 5.5, 5.11]. Namely, we have*

$$\begin{aligned}\|\big(S^{'}_{kh}(\bar q^r_{kh}) - S^{'}(\bar q^r_{kh})\big)(q_\gamma - \bar q)\|_{L^\infty(I,H)} \le c\Big(k\Big(\log\frac{T}{k}+1\Big)^{\frac12} + h^2\Big(\log\frac{T}{k}+1\Big)\Big)\cdot\\ \cdot\big(\|g\|_{L^\infty(\Omega)}\|q_\gamma - \bar q\|_{L^\infty(I,\mathbb{R}^m)}\big).\end{aligned}$$

By virtue of the control constraints, we have $\|q_\gamma - \bar{q}\|_{L^\infty(I,\mathbb{R}^m)} \leq |q_{max} - q_{min}|$. Then, thanks to (42) *and the feasibility of $\bar{q}^r_{kh}$ for* ($\mathbb{P}^r_{kh}$)*, we conclude*

$$|(G'_{kh}(q) - G'(\bar{q}))(q_\gamma - \bar{q})| \leq c_3 \left(k\Big(\log \frac{T}{k} + 1\Big)^{\frac{1}{2}} + h^2 \Big(\log \frac{T}{k} + 1 \Big) + \frac{r^2}{2} \right).$$

Moreover, by the above arguments it is clear that c_1–c_3 remain bounded as $r \to 0$.

As we have seen in the discussion after Assumption 25 it holds $\gamma(r) \simeq r\gamma$. Hence there exists $\tilde{r} \leq r$ such that

$$-\gamma(\tilde{r}) + \Big(c_2 + \frac{c_3}{2}\Big)\tilde{r}^2 \leq -\frac{3}{4}\gamma(\tilde{r}). \tag{44}$$

We can now summarize our requirements on r. Throughout the rest of the paper we rely on the following.

Assumption 27. *Let the radius $r > 0$ be small enough such that* (44) *holds and the quadratic growth condition* (17) *holds for elements in Q^r_{feas}. Namely,*

$$j(q) \geq j(\bar{q}) + \delta \|q - \bar{q}\|^2_{L^2(I,\mathbb{R}^m)}$$

for any $q \in Q^r_{feas}$.

After this preparation, for $G_{kh} = F_{kh} \circ S_{kh}$, we introduce the discrete auxiliary problem ($\mathbb{P}^r_{kh}$)

$$(\mathbb{P}^r_{kh}) \qquad \begin{aligned} &\min\ j_{kh}(q_{kh}) := J(S_{kh}(q_{kh}, q_{kh})) \\ &\text{s.t. } q_{kh} \in Q^r_{kh,\text{feas}} := \{q_{kh} \in Q^r \mid G_{kh}(q_{kh}) \leq 0\}. \end{aligned}$$

We remark again that the control is not discretized, the index k, h is taken only to clarify the association to the problem ($\mathbb{P}^r_{kh}$).

In a first step, we construct feasible competitors for ($\mathbb{P}^r_{kh}$).

Proposition 28. *Let $\bar{q}$ be a local solution of* ($\mathbb{P}$) *and q_γ be the Slater's point from Assumption 10. Let*

$$t(k,h) = \frac{c_1(k(\log(T/k) + 1)^{1/2} + h^2(\log(T/k) + 1))}{c_4 r^2 - \gamma}$$

be given with c_4 such that $0 < c_4 r^2 - \gamma < \gamma/2$. Then, the sequence of controls defined by

$$q_{t(k,h)} = \bar{q} + t(k,h)(q_\gamma - \bar{q}) \tag{45}$$

is feasible for ($\mathbb{P}^r_{kh}$)*, for k, h sufficiently small such that $0 < t(k,h) < 1$.*

Proof. To verify the feasibility of $q_{t(k,h)}$ we use a Taylor's expansion argument. The definition of $q_{t(k,h)}$ suggests to expand $G(q_{t(k,h)})$ at $\bar{q}$, obtaining

$$G(q_{t(k,h)}) = G(\bar{q}) + G^{'}(\bar{q})(q_{t(k,h)} - \bar{q}) + \frac{1}{2} G^{''}(q_\zeta)(q_{t(k,h)} - \bar{q})^2,$$

where q_ζ is a convex combination of $q_{t(k,h)}$ and $\bar{q}$.

We insert this expansion in the following calculations

$$
\begin{aligned}
G_{kh}(q_{t(k,h)}) &= G_{kh}(q_{t(k,h)}) - G(q_{t(k,h)}) + G(q_{t(k,h)}) \\
&= G_{kh}(q_{t(k,h)}) - G(q_{t(k,h)}) + G(\bar{q}) + G^{'}(\bar{q})(q_{t(k,h)} - \bar{q}) \\
&\quad + \frac{1}{2}G^{''}(q_\zeta)(q_{t(k,h)} - \bar{q})^2 \\
&= G_{kh}(q_{t(k,h)}) - G(q_{t(k,h)}) + G(\bar{q}) + t(k,h)G(\bar{q}) - t(k,h)G(\bar{q}) \\
&\quad + t(k,h)G^{'}(\bar{q})(q_\gamma - \bar{q}) + \frac{1}{2}G^{''}(q_\zeta)(q_{t(k,h)} - \bar{q})^2 \\
&= \underbrace{G_{kh}(q_{t(k,h)}) - G(q_{t(k,h)})}_{(a_1)} \\
&\quad + \underbrace{(1 - t(k,h))G(\bar{q}) + t(k,h)(G(\bar{q}) + G^{'}(\bar{q})(q_\gamma - \bar{q}))}_{(a_2)} \\
&\quad + \underbrace{\frac{1}{2}G^{''}(q_\zeta)(q_{t(k,h)} - \bar{q})^2}_{(a_3)}.
\end{aligned}
$$

(a_1) By definition of c_1 it holds

$$
\begin{aligned}
G_{kh}(q_{t(k,h)}) - G(q_{t(k,h)}) &= (u_{kh}(q_{t(k,h)}) - u(q_{t(k,h)}), \omega(x))_I \\
&\le c_1\Big(k\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k} + 1\Big)\Big).
\end{aligned}
$$

(a_2) This part is handled thanks to the feasibility of $\bar{q}$ for $(\mathbb{P})$ and Slater's regularity condition of Assumption 10. Indeed, for k, h sufficiently small, such that $0 < t(k,h) < 1$, we have

$$
\begin{aligned}
(1 - t(k,h))G(\bar{q}) &\le 0, \\
t(k,h)(G(\bar{q}) + G^{'}(\bar{q})(q_\gamma - \bar{q})) &\le -t(k,h)\gamma,
\end{aligned}
$$

from which we obtain

$$
(a_2) \le -t(k,h)\gamma.
$$

(a_3) By definition of c_2 it follows

$$
G^{''}(q_\zeta)(q_{t(k,h)} - \bar{q})^2 \le c_2 t(k,h)^2 \|q_\gamma - \bar{q}\|^2_{L^2(I,\mathbb{R}^m)} \le c_2 t(k,h)^2 \frac{r^2}{4}.
$$

Combining the three parts and using the definition of $t(k,h)$, we have

$$
\begin{aligned}
G_{kh}(q_{t(k,h)}) &\le c_1\Big(k\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k} + 1\Big)\Big) + t(k,h)(c_2 t(k,h)\frac{r^2}{4} - \gamma) \\
&= t(k,h)(c_4 r^2 - \gamma) + t(k,h)(c_2 t(k,h)\frac{r^2}{4} - \gamma) \\
&= t(k,h)\Big(c_4 r^2 - 2\gamma + c_2 t(k,h) r^2\Big).
\end{aligned}
$$

Hence, for h, k are sufficiently small such that $0 < t(k,h) < 1$, we obtain from (44) and the definition of c_4 that

$$\begin{aligned} G_{kh}(q_{t(k,h)}) &\le t(k,h)\Big(c_4 r^2 - 2\gamma + c_2 r^2\Big) \\ &\le (c_4 - \gamma) + (c_2 r^2 - \gamma) \\ &\le \frac{\gamma}{2} - \frac{3}{4}\gamma \\ &\le -\frac{1}{4}\gamma < 0 \end{aligned}$$

and the feasibility of $q_{t(k,h)}$ is verified. □

The proposition above in particular ensures that $Q^r_{kh,\text{feas}}$ is not empty once k, h are small enough.

Corollary 29. *For k, h sufficiently small, there exists at least one global solution $\bar{q}^r_{kh} \in Q^r_{kh,feas}$ of $(\mathbb{P}^r_{kh})$.*

In a second step, we show that the linearized regularity condition of Assumption 10 continues to hold in the discrete setting.

Lemma 30. *Under Assumption 10, for k, h small enough it holds*

$$G_{kh}(\bar{q}^r_{kh}) + G'_{kh}(\bar{q}^r_{kh})(q_\gamma - \bar{q}^r_{kh}) \le -\frac{1}{2}\gamma < 0 \quad \text{on } \overline{I}. \tag{46}$$

Proof. In view of Assumption 10, we add and subtract $G(\bar{q}), G_{kh}(\bar{q})$, $G'(\bar{q})(q_\gamma - \bar{q})$ to obtain

$$\begin{aligned} G_{kh}(\bar{q}^r_{kh}) + G'(\bar{q}^r_{kh})(q_\gamma - \bar{q}^r_{kh}) &= G(\bar{q}) + G'(\bar{q})(q_\gamma - \bar{q}) + G_{kh}(\bar{q}^r_{kh}) \\ &\quad + G'(\bar{q}^r_{kh})(q_\gamma - \bar{q}^r_{kh}) - G(\bar{q}) - G'(\bar{q})(q_\gamma - \bar{q}) \\ &\le -\gamma + \underbrace{G_{kh}(\bar{q}^r_{kh}) + G'_{kh}(\bar{q}^r_{kh})(\bar{q} - \bar{q}^r_{kh}) - G_{kh}(\bar{q})}_{(b_1)} \\ &\quad + \underbrace{G_{kh}(\bar{q}) - G(\bar{q})}_{(b_2)} + \underbrace{\big(G'_{kh}(\bar{q}^r_{kh}) - G'(\bar{q})\big)(q_\gamma - \bar{q})}_{(b_3)}. \end{aligned}$$

(b_1) Taylor expansion of $G_{kh}(\bar{q})$ at $\bar{q}^r_{kh}$ reads

$$G_{kh}(\bar{q}) = G_{kh}(\bar{q}^r_{kh}) + G'_{kh}(\bar{q}^r_{kh})(\bar{q} - \bar{q}^r_{kh}) + \frac{1}{2}G''_{kh}(q_\zeta)(\bar{q} - \bar{q}^r_{kh})^2,$$

with q_ζ convex combination of $\bar{q}$ and $\bar{q}^r_{kh}$, yielding

$$(b_1) = -\frac{1}{2}G''_{kh}(q_\zeta)(\bar{q} - \bar{q}^r_{kh})^2 \le c_2\|\bar{q} - \bar{q}^r_{kh}\|^2_{L^2(I,\mathbb{R}^m)} \le c_2 r^2,$$

where we used G_{kh} being a C^2-functional together with the feasibility of $\bar{q}^r_{kh}$ for $(\mathbb{P}^r_{kh})$.

(b_2) By definition of c_1 it holds

$$\begin{aligned} G_{kh}(\bar{q}_k) - G(\bar{q}) &= \int_\Omega \big(u_{kh}(\bar{q}) - u(\bar{q})\big)\omega(x)\mathrm{d}x \\ &\le c_1\Big(k\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k} + 1\Big)\Big). \end{aligned}$$

(b_3) By definition of c_3 it follows

$$(G'_{kh}(\bar{q}^r_{kh}) - G'(\bar{q}))(q_\gamma - \bar{q}) \leq c_3\Big(k\Big(\ln\frac{T}{k}+1\Big)^{\frac{1}{2}} + h^2\Big(\ln\frac{T}{k}+1\Big) + \frac{r^2}{2}\Big)$$

In conclusion, for k, h sufficiently small and thanks to (44), the three estimates for $(b_1), (b_2), (b_3)$ yield

$$\begin{aligned}
G_{kh}(\bar{q}^r_{kh}) + G^{'}(\bar{q}^r_{kh})(q_\gamma - \bar{q}^r_{kh}) &\leq -\gamma + c_2 r^2 + c_1\Big(k\Big(\log\frac{T}{k}+1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k}+1\Big)\Big)\\
&\quad + c_3\Big(k\Big(\log\frac{T}{k}+1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k}+1\Big) + \frac{r^2}{2}\Big)\\
&\leq -\gamma + (c_2 + \frac{c_3}{2})r^2 + (c_1+c_3)\cdot\\
&\quad \cdot\Big(k\Big(\log\frac{T}{k}+1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k}+1\Big)\Big)\\
&\leq -\frac{3}{4}\gamma + (c_1+c_3)\Big(k\Big(\log\frac{T}{k}+1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k}+1\Big)\Big)\\
&\leq -\frac{1}{2}\gamma.
\end{aligned}$$

□

We now introduce the feasible competitors for the continuous auxiliary problem $(\mathbb{P}^r)$. The proof is similar to Proposition 28, therefore we highlight only the main arguments.

Proposition 31. *Let $\bar{q}^r_{kh}$ be a global optimum for $(\mathbb{P}^r_{kh})$ and q_γ be the Slater's point from Assumption 10. Further, let*

$$\tau(k,h) = \frac{c_1(k(\log(T/k)+1)^{1/2} + h^2(\log(T/k)+1)}{c_4 r^2 - \gamma}$$

be given with a constant c_4 such that $0 < c_4 r^2 - \gamma < \gamma/2$. Then, the sequence of controls defined by

$$q_{\tau(k,h)} = \bar{q}^r_{kh} + \tau(k,h)(q_\gamma - \bar{q}^r_{kh}) \tag{47}$$

is feasible for $(\mathbb{P}^r)$, for k, h sufficiently small.

Proof. Analogously to Proposition 28, with the help of Taylor expansion, this time of $G_{kh}(q_{\tau(k,h)})$, at $\bar{q}^r_{kh}$, we obtain

$$\begin{aligned}
G(q_{\tau(k,h)}) = &\underbrace{G(q_{\tau(k,h)}) - G_{kh}(q_{\tau(k,h)})}_{(b_1)} + \\
&\underbrace{(1-\tau(k,h)G_{kh}(\bar{q}^r_{kh})) + \tau(k,h)(G_{kh}(\bar{q}^r_{kh}) + G^{'}_{kh}(\bar{q}^r_{kh})(q_\gamma - \bar{q}^r_{kh}))}_{(b_2)}\\
&+ \underbrace{\frac{1}{2}G^{''}_{kh}(q_\zeta)(q_{\tau(k,h)} - \bar{q}^r_{kh})}_{(b_3)}\\
\leq\ &\tau(k,h)\Big(c_4 r^2 - 2\gamma + c_2\tau(k,h) r^2\Big),
\end{aligned}$$

where we used Theorems 22 and 24 for (b_1), the feasibility of $\bar{q}^r_{kh}$ and discrete Slater condition of Lemma 30 for (b_2), G_{kh} being a C^2-functional together with (43) for (b_3).

Then, once k, h are sufficiently small such that $0 < \tau(k,h) < 1$, in addition to the prerequisite of Lemma 30, the claim follows as in Proposition 28. $\square$

With these results at hands, we now show that global solutions of $(\mathbb{P}^r_{kh})$ converge to the considered local solution of $(\mathbb{P})$.

Proposition 32. *Let k, h be small enough, such that Propositions 28 and 31 hold. Let $\bar{q}$ be a local solution for $(\mathbb{P})$ satisfying the assumptions of Theorem 11 and Assumption 13, and let $\bar{q}^r_{kh}$ be a global solution of $(\mathbb{P}^r_{kh})$. Then it holds the error estimate*

$$\|\bar{q} - \bar{q}^r_{kh}\|^2_{L^2(I,\mathbb{R}^m)} \le c\Big(k\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k} + 1\Big)\Big). \tag{48}$$

Proof. Let $q_{t(k,h)}$ and $q_{\tau(k,h)}$ be defined as in Proposition 28 and Proposition 31, respectively, and let k, h be small enough such that $0 < t(k,h), \tau(k,h) < 1$. We have

$$\|\bar{q} - \bar{q}^r_{kh}\|_{L^2(I,\mathbb{R}^m)} \le \|\bar{q} - q_{\tau(k,h)}\|_{L^2(I,\mathbb{R}^m)} + \|q_{\tau(k,h)} - \bar{q}^r_{kh}\|_{L^2(I,\mathbb{R}^m)}.$$

For the second term we have

$$\|q_{\tau(k,h)} - \bar{q}^r_{kh}\|_{L^2(I,\mathbb{R}^m)} \le c\Big(k\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}}\Big),$$

since, thanks to Proposition 31, $q_{\tau(k,h)}$ converges strongly in $L^2(I,\mathbb{R}^m)$ to $\bar{q}^r_{kh}$ with order $\tau(k,h)$. Therefore, we are left with the first term.

The competitor $q_{\tau(k,h)}$ is feasible for $(\mathbb{P}^r)$ and, using the quadratic growth condition (17), we obtain

$$\begin{aligned}
\delta\|\bar{q} - q_{\tau(k,h)}\|^2_{L^2(I,\mathbb{R}^m)} &\le j(q_{\tau(k,h)}) - j(\bar{q}) \\
&= j(q_{\tau(k,h)}) - j_{kh}(\bar{q}^r_{kh}) + j_{kh}(\bar{q}^r_{kh}) - j_{kh}(q_{t(k,h)}) \\
&\quad + j_{kh}(q_{t(k,h)}) - j(\bar{q}) \\
&\le \underbrace{j(q_{\tau(k,h)}) - j_{kh}(\bar{q}^r_{kh})}_{(d_1)} + \underbrace{j_{kh}(q_{t(k,h)}) - j(\bar{q})}_{(d_2)},
\end{aligned}$$

where in the last step we have used that $q_{t(k,h)} \in Q^r_{kh,\text{feas}}$ and $\bar{q}^r_{kh}$ is a global optimum for $(\mathbb{P}^r_{kh})$.

We now analyze the two terms separately.

(d_1) With simple algebraic manipulations and the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
j(q_{\tau(k,h)}) - j_{kh}(\bar{q}^r_{kh}) \le& \frac{1}{2}\|u(q_{\tau(k,h)}) + u_{kh}(\bar{q}^r_{kh}) - 2u_d\|_I \|u(q_{\tau(k,h)}) - u_{kh}(\bar{q}^r_{kh})\|_I \\
&+ \frac{\alpha}{2}\|q_{\tau(k,h)} + \bar{q}^r_{kh}\|_{L^2(I,\mathbb{R}^m)}\|q_{\tau(k,h)} - \bar{q}^r_{kh}\|_{L^2(I,\mathbb{R}^m)}.
\end{aligned}$$

Then, by means of the stability of the solution u and u_{kh} of (2) and (24), respectively, together with the boundedness of Q_{ad}, and with the help of the Cauchy-Schwarz inequality, we get

$$
\begin{aligned}
j(q_{\tau(k,h)}) - j_{kh}(\bar{q}^r_{kh}) &\leq c\Big(\|u(q_{\tau(k,h)}) - u(\bar{q}^r_{kh})\|_I + \|u(\bar{q}^r_{kh}) - u_{kh}(\bar{q}^r_{kh})\|_I \\
&\qquad + \|q_{\tau(k,h)} - \bar{q}^r_{kh}\|_{L^2(I,\mathbb{R}^m)}\Big) \\
&\leq c\Big(\|u(\bar{q}^r_{kh}) - u_{kh}(\bar{q}^r_{kh})\|_I + \|q_{\tau(k,h)} - \bar{q}^r_{kh}\|_{L^2(I,\mathbb{R}^m)}\Big),
\end{aligned}
$$

where in the last step we have used (7a).

The first term is a discretization error that can be estimated by [31, Theorems 3.3 and 4.2] together with the regularity of the solution of (2), obtaining

$$
\|u(\bar{q}^r_{kh}) - u_{kh}(\bar{q}^r_{kh})\|_I \leq c(k + h^2).
$$

The estimate for the second term, $\|q_{\tau(k,h)} - \bar{q}^r_{kh}\|$, follows directly from Proposition 31. Summing up, we conclude

$$
\begin{aligned}
j(q_{\tau(k,h)}) - j_k(q^r_k) &\leq c\Big(k + h^2 + k\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k} + 1\Big)\Big) \\
&\leq c\Big(k\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k} + 1\Big)\Big).
\end{aligned}
$$

(d_2) We proceed exactly as for (d_1).

$$
\begin{aligned}
j_{kh}(q_{t(k,h)}) - j(\bar{q}) &\leq \frac{1}{2}\|u_{kh}(q_{t(k,h)}) + u(\bar{q}) - 2u_d\|_I \|u_{kh}(q_{t(k,h)}) - u(\bar{q})\|_I \\
&\quad + \frac{\alpha}{2}\|q_{t(k,h)} + \bar{q}\|_{L^2(I,\mathbb{R}^m)}\|q_{t(k,h)} - \bar{q}\|_{L^2(I,\mathbb{R}^m)} \\
&\leq c\Big(\|u_{kh}(q_{t(k,h)}) - u(q_{t(k,h)})\|_I + \|q_{t(k,h)} - \bar{q}\|_{L^2(I,\mathbb{R}^m)}\Big) \\
&\leq c\Big(k\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k} + 1\Big)\Big).
\end{aligned}
$$

Combining (d_1) with (d_2), we have the assertion. □

It is readily seen that, for k, h small enough, global solutions of $(\mathbb{P}^r_{kh})$ are local solutions of $(\mathbb{P}_{kh})$, as the constraint $\|\bar{q} - \bar{q}^r_{kh}\|_{L^2(I,\mathbb{R}^m)} \leq r$ is not active. In particular, this ensures the existence of a sequence $\bar{q}_{kh}$, of local solutions to $(\mathbb{P}_{kh})$, converging to $\bar{q}$. We formalize this in the main result of the paper

Theorem 33. *Let $\bar{q}$ be a local solution of $(\mathbb{P})$ satisfying the assumptions of Theorem 11 and Assumption 13. Then, for k, h sufficiently small, there exists a sequence $(\bar{q}_{kh})$ of local solution of $(\mathbb{P}_{kh})$ converging to $\bar{q}$ as $k, h \to 0$. Further, there holds the error estimate*

$$
\|\bar{q} - \bar{q}_{kh}\|^2_{L^2(I,\mathbb{R}^m)} \leq c\left(k\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} + h^2\Big(\log\frac{T}{k} + 1\Big)\right). \tag{49}
$$

Remark 34. *Making use of the same technique one can derive an error estimate arising from the discretization in time only. As one might expect, this reads*

$$
\|\bar{q} - \bar{q}_k\|^2_{L^2(I,\mathbb{R}^m)} \leq ck\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}}
$$

where $\bar{q}_k$ is a suitable sequence of local solutions to $(\mathbb{P}_k)$. On the other hand, the error between the semidiscrete and discrete solution will only satisfy the estimate *in Theorem 33. In order to obtain*

$$\|\bar{q}_k - \bar{q}_{kh}\|^2_{L^2(I,\mathbb{R}^m)} \leq ch^2\Big(\log\frac{T}{k}+1\Big),$$

a quadratic growth condition in the solution of $\mathbb{P}_k^r$ needs to hold uniformly in k. To this end, the SSCs need to be transfered to the semidiscrete setting. The difficulty in this procedure lies in the convergence of the critical directions. As it is not clear how this can be shown, one would resolve to utilize a strong SSC; thereby avoiding the need for certain critical directions.

Acknowledgments

The authors are grateful for the support of their former host institutions. To this end, I. Neitzel acknowledges the support of the Technische Universität München and F. Ludovici and W. Wollner the support of the Universität Hamburg.

References

[1] R. E. Bank and H. Yserentant, *On the H^1-stability of the L_2-projection onto finite element spaces*, Numer. Math., 126 (2014), pp. 361–381.
[2] J. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer Ser. Oper. Res. Financ. Eng., Springer-Verlag, New York, first ed., 2010.
[3] J. F. Bonnans and P. Jaisson, *Optimal control of a parabolic equation with time-dependent state constraints*, SIAM J. Control Optim., 48 (2010), pp. 4550–4571.
[4] S. C. Brenner and R. L. Scott, *The Mathematical Theory of Finite Element Methods*, vol. 15 of Texts Appl. Math., Springer, New York, third ed., 2008.
[5] E. Casas, *Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations*, SIAM J. Control Optim., 35 (1997), pp. 1297–1327.
[6] E. Casas, J. C. de los Reyes, and F. Tröltzsch, *Sufficient second-order optimality conditions for semilinear control problems with pointwise state constraints*, SIAM J. Optim., 19 (2008), pp. 616–643.
[7] E. Casas and M. Mateos, *Second order optimality conditions for semilinear elliptic control problems with finitely many state constraints*, SIAM J. Control Optim., 40 (2002), pp. 1431–1454 (electronic).
[8] E. Casas and M. Mateos, *Uniform convergence of the FEM. Applications to state constrained control problems*, Comput. Appl. Math., 21 (2002), pp. 67–100.
[9] E. Casas, J.-P. Raymond, and H. Zidani, *Pontryagin's principle for local solutions of control problems with mixed control-state constraints*, SIAM J. Control Optim., 39 (2000), pp. 1182–1203.
[10] E. Casas and F. Tröltzsch, *Error estimates for the finite-element approximation of a semilinear elliptic control problem*, Control Cybernet., 31 (2002), pp. 695–712.
[11] E. Casas and F. Tröltzsch, *Second Order Optimality Conditions and Their Role in PDE Control*, Jahresber. Dtsch. Math.-Ver., 117 (2015), pp. 3–44.
[12] K. Chrysafinos, *Convergence of discontinuous Galerkin approximations of an optimal control problem associated to semilinear parabolic PDE's*, M2AN Math. Model. Numer. Anal., 44 (2010), pp. 189–206.
[13] K. Chrysafinos and E. N. Karatzas, *Symmetric error estimates for discontinuous Galerkin approximations for an optimal control problem associated to semilinear parabolic PDE's*, Discrete Contin. Dyn. Syst. Ser. B, 17 (2012), pp. 1473–1506.
[14] J. C. de Los Reyes, P. Merino, J. Rehberg, and F. Tröltzsch, *Optimality conditions for state-constrained PDE control problems with time-dependent controls*, Control Cybernet., 37 (2008), pp. 5–38.
[15] K. Deckelnick and M. Hinze, *Variational discretization of parabolic control problems in the presence of pointwise state constraints*, J. Comput. Math., 29 (2011), pp. 1–15.
[16] R. S. Falk, *Approximation of a class of optimal control problems with order of convergence estimates*, J. Math. Anal. Appl., 44 (1973), pp. 28–47.

[17] H. Goldberg and F. Tröltzsch, *Second-order sufficient optimality conditions for a class of nonlinear parabolic boundary control problems*, SIAM J. Control Optim., 31 (1993), pp. 1007–1025.

[18] W. Gong and M. Hinze, *Error estimates for parabolic optimal control problems with control and state constraints*, Comput. Optim. Appl., 56 (2013), pp. 131–151.

[19] M. Hinze, *A variational discretization concept in control constrained optimization: the linear-quadratic case*, Comput. Optim. Appl., 30 (2005), pp. 45–61.

[20] M. Hinze and C. Meyer, *Stability of semilinear elliptic optimal control problems with pointwise state constraints*, Comput. Optim. Appl., 52 (2012), pp. 87–114.

[21] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich, *Optimization with PDE Constraints*, Mathematical Modelling: Theory and Applications, Springer Netherlands, 2010.

[22] K. Krumbiegel and J. Rehberg, *Second order sufficient optimality conditions for parabolic optimal control problems with pointwise state constraints*, SIAM J. Control Optim., 51 (2013), pp. 304–331.

[23] D. Leykekhman and B. Vexler, *Optimal a priori error estimates of parabolic optimal control problems with pointwise control*, SIAM J. Numer. Anal., 51 (2013), pp. 2797–2821.

[24] D. Leykekhman and B. Vexler, *Pointwise best approximation results for galerkin finite element solutions of parabolic problems*, Tech. Report arXiv:1508.01165, 2016.

[25] F. Ludovici and W. Wollner, *A priori error estimates for a finite element discretization of parabolic optimization problems with pointwise constraints in time on mean values of the gradient of the state*, SIAM J. Control Optim., 53 (2015), pp. 745–770.

[26] D. Meidner, R. Rannacher, and B. Vexler, *A priori error estimates for finite element discretizations of parabolic optimization problems with pointwise state constraints in time*, SIAM J. Control Optim., 49 (2011), pp. 1961–1997.

[27] D. Meidner and B. Vexler, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems. I. Problems without control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1150–1177.

[28] D. Meidner and B. Vexler, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems. II. Problems with control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1301–1329.

[29] C. Meyer, *Error estimates for the finite-element approximation of an elliptic control problem with pointwise state and control constraints*, Control Cybernet., 37 (2008), pp. 51–83.

[30] I. Neitzel, J. Pfefferer, and A. Rösch, *Finite element discretization of state-constrained elliptic optimal control problems with semilinear state equation*, SIAM J. Control Optim., 53 (2015), pp. 874–904.

[31] I. Neitzel and B. Vexler, *A priori error estimates for space-time finite element discretization of semilinear parabolic optimal control problems*, Numer. Math., 120 (2012), pp. 345–386.

[32] R. H. Nochetto, *Sharp L^∞-error estimates for semilinear elliptic problems with free boundaries*, Numer. Math., 54 (1988), pp. 243–255.

[33] J.-P. Raymond and F. Tröltzsch, *Second order sufficient optimality conditions for nonlinear parabolic control problems with state constraints*, Discrete Contin. Dynam. Systems, 6 (2000), pp. 431–450.

[34] V. Thomèe, *Galerkin Finite Element Methods for Parabolic Problems*, vol. 25 of Springer Ser. Comput. Math., Springer, Berlin, second ed., 2006.

[35] F. Tröltzsch, *Optimal Control of Partial Differential Equations. Theory, Methods and Applications.*, vol. 112 of Grad. Stud. Math., AMS, 2010.

(F. Ludovici) Fachbereich Mathematik, Technische Universität Darmstadt, Dolivostr. 15, 64293 Darmstadt, Germany
E-mail address: ludovici@mathematik.tu-darmstadt.de

(I. Neitzel) Institut für Numerische Simulation, Rheinische Friedrich-Wilhelms-Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany
E-mail address: neitzel@ins.uni-bonn.de

(W. Wollner) Fachbereich Mathematik, Technische Universität Darmstadt, Dolivostr. 15, 64293 Darmstadt, Germany
E-mail address: wollner@mathematik.tu-darmstadt.de