Hamburger Beiträge
zur Angewandten Mathematik

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

Francesco Ludovici und Winnifried Wollner

Nr. 2014-05
March 2014

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

FRANCESCO LUDOVICI AND WINNIFRIED WOLLNER*

Abstract. This article is concerned with the discretization of parabolic optimization problems subject to pointwise in time constraints on mean values of the derivative of the state variable. Central component of the analysis are a priori error estimates for the dG(0)-cG(1) discretization of the parabolic partial differential equation (PDE) in the $L^\infty(0,T;H_0^1(\Omega))$-norm, together with corresponding estimates in $L^1(0,T;H^{-1}(\Omega))$ for the adjoint PDE. These results are then utilized to show convergence orders for the discrete approximation towards the solution of the parabolic optimization problem.

Key words. parabolic optimization, gradient state constraints, pointwise in time constraints, space-time a priori error

AMS subject classifications. 49M25, 65M12, 65M15, 65M60

1. Introduction.

We are concerned with optimization problems governed by parabolic partial differential equations (PDEs). For clarity of the presentation, we confine ourselves to the case of the heat-equation with homogeneous Dirichlet boundary conditions and a control acting distributed in the domain. The most important feature is the consideration of pointwise in time constraints on weighted mean-values of the spatial gradient of the solution of the PDE. Consideration of such constraints is motivated by bounds on average stresses in glass cooling processes, [10,31,32], and steel cooling, see, e.g., [38] and the references therein.

To be precise, for a time interval $I = (0,T)$ and a domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, let $u = u(t,x)$ be the state variable, i.e., the solution of the parabolic PDE and $q = q(t)$ the control variable. We consider the following model problem

$$\text{Minimize } \frac{1}{2}\int_I\int_\Omega (u(x,t) - u_d(x,t))^2 dxdt + \frac{\alpha}{2}\int_I q(t)^2 dt,$$

where u and q are coupled by the parabolic PDE

$$\partial_t u(t,x) - \Delta u(t,x) = q(t)g(x)$$

with suitable boundary conditions and initial data. Additionally, box constraints on the control variable and, most importantly, state constraints of the form

$$\int_\Omega |\nabla u(x,t)|^2 \omega(x)dx \le b \qquad \forall t \in [0,T]$$

are considered. The precise formulation of the problem is presented in Section 2.

The following a priori error analysis is inspired by the work of [26] where the authors extended the technique developed in [27] and [28], to the case of constraints on mean values of the state. However, due to the consideration of derivatives of the state, the analysis of the problem at hand is severely more involved. Indeed, the error estimate for the optimal control problem requires, at any level of discretization, error

*Fachbereich Mathematik, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany.[francesco.ludovici|winnifried.wollner]@uni-hamburg.de

estimates for the state equation in $L^\infty(I, H_0^1(\Omega))$, instead of $L^\infty(I, L^2(\Omega))$ considered in [26], which are not present in the literature. Namely, for the temporal discretization we will show

$$\|u - u_k\|_{L^\infty(I,V)} \le Ck\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}}\Big(\|f\|_{L^\infty(I,V)} + \|u_0\|_{H^3(\Omega)}\Big)$$

in Theorem 4.8. The corresponding estimate for the spatial discretization error

$$\|u_{kh} - u_k\|_{L^\infty(I,V)} \le Ch(\|f\|_{L^2(I,V)} + \|u_0\|_{H^2(\Omega)})$$

will be provided in Theorem 4.12.

In addition, as usual in the presence of pointwise state constraints, the associated Lagrange multiplier is a Borel measure in $C(\bar{I})^*$, leading to low regularity of the adjoint state.

The literature on gradient state constraints for parabolic problem has only few contributions. To the best of the authors knowledge, an a priori analysis in the case of gradient state constraints has not yet been considered. Integral gradient constraint pointwise in time are considered in [25], where existence and optimality conditions are discussed. In [8], a Pontryagin's principle is obtained using Ekeland's variational principle. In [33], second order sufficient conditions are discussed in a setting including integral gradient constraints.

Integral state constraints, involving the state, but not its derivative, were considered in [15] and [2]. In both cases, second order sufficient conditions were investigated; the former in presence of a non-linearity in the boundary condition, the latter in presence of a cubic non-linearity in the differential equation. For state constraints of integral and mixed type in the semilinear case, we mention also [4, 9].

State constraints pointwise in space and time are discussed in several publications. Regarding the linear case, in the recent publication [16] a priori error estimates in the L^2-norm are derived. A Lavrentiev-type regularization was considered in [29] for both distributed and boundary controls. For a discussion of the variational discretization approach in the parabolic case, we refer to [12]. The papers [1, 5, 21, 34] deal with semilinear differential equations and second order sufficient conditions.

Gradient constraints for elliptic problem have recently received more attention than the parabolic case. Optimality conditions have been derived on smooth domains in [6, 7], the case of nonsmooth polygonal domains was considered in [39] Algorithmically, barrier methods where considered in [35], while penalty methods are considered in [19] for smooth domains and [40] for nonsmooth polygonal domains. A priori error estimates have been derived in [11, 18, 30], and [40] for nonsmooth domains.

This paper is structured as follows: In Section 2, we define the model problem, introduce some notations and state the necessary optimality conditions. The time and space discretization of the problem is presented in Section 3. In Section 4, we provide stability estimates for the state equation and for additional auxiliary problems. At any level of discretization, we derive error estimate for the state equation in the $L^\infty(I, H_0^1(\Omega))$-norm. In Section 5, we assemble the results providing the rate of convergence for the optimal control problem.

2. The Problem. Let $\Omega \subset \mathbb{R}^d$, with $d \in \{2, 3\}$, be a convex bounded domain with C^2-boundary, $I = (0, T)$ a given time interval and abbreviate $V := H_0^1(\Omega)$,

$H := L^2(\Omega)$. We consider the linear parabolic PDE

$$\begin{aligned} \partial_t u - \Delta u = f \qquad & \text{in } I \times \Omega, \\ u = 0 \qquad & \text{on } I \times \partial\Omega, \\ u = u_0 \quad & \text{in } \{0\} \times \Omega, \end{aligned} \tag{2.1}$$

with right hand side $f = f(t,x) = q(t)g(x) = \sum_{i=1}^m q_i(t)g_i(t)$, $q_i \in L^2(I)$, $g_i \in V$, $u_0 \in \ H^2(\Omega) \cap V$. The splitting of the right hand side is motivated by practical considerations in the context of industrial applications, where the control q acts in time only and g represents the control action.

The regularity of the data ensures the existence of a weak solution for (2.1) in the space

$$U = \{u \in L^2(I, H^2(\Omega)) \cap L^\infty(I, V), u_t \in L^2(I, H)\}, \tag{2.2}$$

see [14, Chapter 7, Theorem 5]. The thus defined control-to-state map $S : L^2(I)^m \to U$, which associates to any given $q \in L^2(I)^m$ the solution $Sq = u(q)$ to (2.1), is continuous.

In the following, $(\cdot,\cdot)_I$ denotes the standard inner product in $L^2(I,H)$, i.e., $(\cdot,\cdot)_I = \int_I (\cdot,\cdot)dt$ with associated norm $\|\cdot\|_I$, while $(\cdot,\cdot)$ and $\|\cdot\|$ is used for H. Additional notation is introduced at the end of the section.

The following optimal control problem with tracking-type objective is then considered

$$\begin{aligned} \underset{(q,u)\in Q_{ad}\times U}{\text{Minimize}}\ & J(q,u) = \frac{1}{2}\|u-u_d\|_I^2 + \frac{\alpha}{2}\|q\|_{L^2(I)}^2, \\ & \text{subject to (2.1) and} \\ & (|\nabla u|^2, \omega) \leq b \quad \forall t \in \bar{I} \end{aligned} \tag{2.3}$$

with prescribed temperature profile $u_d \in L^2(I,H)$, weighting function $\omega \in L^\infty(\Omega)$, $b \in \mathbb{R}$, and admissible control set

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

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

Assumption 2.1. *We assume the following regularity condition to hold:*

$$\exists \tilde{q} \in Q_{ad} \ \text{ such that } \ (|\nabla u(\tilde{q})|^2, \omega) - b < -\epsilon < 0, \tag{2.4}$$

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

The regularity condition ensures the existence of a feasible point for (2.3), justifying the well-posedness of the problem using standard arguments.

Proposition 2.2. *The optimal control problem* (2.3) *admits a unique solution* $(\bar{q},\bar{u}) \in Q \times U$, *where* $Q = L^\infty(I)$.

Proof. The problem can be formulated in the setting of the distributed optimal control problem analyzed in [25], where well-posedness is showed. The additional regularity of the control is a consequence of the box-control constraint. □

Remark 2.3. *We observe that there holds the embedding* $U \hookrightarrow C(\bar{I}, V)$, *see [22, Theorem 3.1, Chapter 1]. This is what we need to treat the state constraint. Indeed, defining* $G(u) := (|\nabla u|^2, \omega)$ *we have* $G\colon U \to C(\bar{I})$.

In a next step, we formulate the necessary optimality conditions for the optimal control problem.

THEOREM 2.4. *Given Assumption 2.1 the pair* $(\bar{q}, \bar{u}) \in Q_{ad} \times U$ *is optimal for* (2.3) *if and only if it is feasible and there exists a Lagrange multiplier* $\bar{\mu} \in C(\bar{I})^*$ *and an adjoint state* $\bar{z} \in L^2(I \times \Omega)$ *satisfying the following system of optimality conditions:*

$$(\partial_t \bar{u}, \varphi)_I + (\nabla \bar{u}, \nabla \varphi)_I = (\bar{q} g, \varphi)_I + (u_0, \varphi(0)) \qquad \forall \varphi \in U, \tag{2.5a}$$

$$(\partial_t \varphi, \bar{z})_I + (\nabla \varphi, \nabla \bar{z})_I = (\bar{u} - u_d, \varphi)_I + \langle \bar{\mu}, 2(\nabla \bar{u} \nabla \varphi, \omega) \rangle \qquad \forall \varphi \in U, \tag{2.5b}$$

$$\alpha(\bar{q}, q - \bar{q})_{L^2(I)} + (\bar{z}, (q - \bar{q})g)_I \geq 0 \qquad \forall q \in Q_{ad}, \tag{2.5c}$$

$$\langle b - G(\bar{u}), \bar{\mu} \rangle = 0, \; \bar{\mu} \geq 0, \tag{2.5d}$$

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

Proof. The continuity of the control-to-state map S ensures that the reduced cost functional $j(q) = J(q, Sq)$ is well-defined.

Denoting with $K = \{v \in C(\bar{I}) \,|\, v \leq 0, \text{ a.e. in } \bar{I}\}$ the closed convex cone of non positive continuous functions, we observe that Assumption 2.1 corresponds to $G(S\tilde{q}) \in \operatorname{int} K$. Thus, we can formulate problem (2.3) in the abstract setting

$$\min j(q)$$
$$\text{s.t. } G(Sq) \in K \;\; q \in Q_{ad},$$

and the claim follows by standard argument, see, e.g., [37, Chapter 6] together with the solvability of the adjoint equation, see, e.g., [25, Lemma 3]. ∎

REMARK 2.5. *Indeed, for the solution of the adjoint equation* (2.5b) *there holds the additional regularity* $\bar{z} \in L^\infty(I, H^{-1}(\Omega))$, *see [25, Appendix 1].*

We conclude the section with some notation for continuous and discrete negative norms following [36]. For a nonnegative integer s, we introduce the space

$$\dot{H}^s(\Omega) = \{v \in H^s(\Omega) \,|\, \Delta^j v = 0 \text{ on } \partial\Omega, \text{ for } j < s/2\},$$

and the iterated solution operators for Poisson's problem

$$-\Delta^{-1} \,:\, H^{-1}(\Omega) \to \dot{H}^1(\Omega),$$
$$-\Delta^{-1} \,:\, L^2(\Omega) \to \dot{H}^2(\Omega),$$
$$-\Delta^{-2} \,:\, H^{-1}(\Omega) \to \dot{H}^3(\Omega).$$

Then, the semi-norm

$$|\cdot|_{-s} = (-\Delta^{-s} \cdot, \cdot)^{1/2}$$

is equivalent to the usual negative norm on the space $\dot{H}^s(\Omega)$, see [36, Lemma 5.1]. As a consequence, we can define the following equivalent norms on $H^{-s}(\Omega)$ and $L^2(I, H^{-s}(\Omega))$

$$\|\cdot\|_{H^{-1}(\Omega)} := \|\nabla \Delta^{-1} \cdot\|, \qquad \|\cdot\|_{L^2(I, H^{-1}(\Omega))} := \|\nabla \Delta^{-1} \cdot\|_I,$$
$$\|\cdot\|_{H^{-2}(\Omega)} := \|\Delta^{-1} \cdot\|, \qquad \|\cdot\|_{L^2(I, H^{-2}(\Omega))} := \|\Delta^{-1} \cdot\|_I,$$
$$\|\cdot\|_{H^{-3}(\Omega)} := \|\nabla \Delta^{-2} \cdot\|, \qquad \|\cdot\|_{L^2(I, H^{-3}(\Omega))} := \|\nabla \Delta^{-2} \cdot\|_I.$$

Denoting by V_h the standard conforming finite element space of piecewise linear functions, which will be introduced in detail in Section 3.2, we define the inverse of the discrete Laplacian

$$-\Delta_h \,:\, H^{-1}(\Omega) \to V_h,$$

which associates to any $f \in H^{-1}(\Omega)$ an element $v_h \in V_h$ given by

$$(\nabla v_h, \nabla \varphi_h) = f(\varphi_h), \ \forall \varphi_h \in V_h.$$

We introduce the discrete semi-norm

$$|\cdot|_{-s,h} = (-\Delta_h^{-s} \cdot, \cdot)^{1/2},$$

which is equivalent to the continuous semi-norm modulo a small constant, see [36, Lemma 5.3].

Throughout the article, we denote by C a generic constant.

REMARK 2.6. *From the definition of $\dot{H}^s(\Omega)$, we observe that the norm $\|\cdot\|_{H^{-s}}$ corresponds to the norm of $(H^s(\Omega) \cap H_0^1(\Omega))^*$ when $s = 1, 2$. While for $s = 3$, we have the additional condition $\Delta v = 0$.*

3. Discretization. In this section, we briefly describe the discretization in time and space of the state equation together with the corresponding optimality conditions for the semidiscrete and discrete optimal control problem.

The problem is discretized using the so called dG(0)cG(1) method, continuous in space and discontinuous in time Galerkin method. We refer to [13] and [36] for more details.

For the discretization of the control variable, we use the variational approach, going back to [20], implying that the control variable is discretized implicitly by the optimality conditions. Here this means that q_k is piecewise constant.

3.1. Time Discretization. To discretize the problem in time, let t_i be such that $0 = t_0 < t_1 < ... < t_{N-1} < t_N = T$. Then, the intervals $I_n = (t_{n-1}, t_n]$ for $n = 1, ..., N$ and $I_0 = \{0\}$ give a partition of $\bar{I}$. The length of the interval I_n is k_n and we set $k = \max_n k_n$. Further, we assume the existence of strictly positive constants $a, c, \tilde{k}$ such that the following technical conditions hold:

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

Denoting with $\mathcal{P}_0(I_n, V)$ the space of piecewise constant polynomials on I_n with values in V, we introduce the semidiscrete state and trial space

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

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., $(\cdot, \cdot)_{I_n} = \int_{I_n} (\cdot, \cdot) dt$.

Since our functions are piecewise constant on each interval, we can simplify standard notation in our case. It is

$$\varphi_{n+1} = \varphi_n^+ = \lim_{t \to 0^+} \varphi(t_n + t), \ \varphi_n = \lim_{t \to 0^+} \varphi(t_n - t), \ [\varphi]_n = \varphi_{n+1} - \varphi_n,$$

for functions $\varphi \in U_k$. Then, for $u_k, \varphi \in U_k$ we introduce the bilinear form

$$B(u_k, \varphi) = \sum_{n=1}^{N} (\partial_t u_k, \varphi)_{I_n} + (\nabla u_k, \nabla \varphi)_I + \sum_{n=2}^{N} ([u_k]_{n-1}, \varphi_n) + (u_{k,1}, \varphi_1), \qquad (3.1)$$

and the semidiscrete state equation reads: for $q \in Q$, find $u_k(q) \in U_k$ such that

$$B(u_k(q), \varphi) = (qg, \varphi)_I + (u_0, \varphi_1) \qquad (3.2)$$

holds for any $\varphi \in U_k$. In particular, we observe that $G(u_k)$ is constant on each I_n, i.e., $G(u_k) \in U_k(\mathbb{R})$.

REMARK 3.1. *Utilizing that the solution $u(q)$ of (2.1) is in $C(I, L^2(\Omega))$ it is clear, that $u(q)$ satisfies (3.2) as well. Thus, there holds the orthogonality relation $B(u(q) - u_k(q), \varphi) = 0$, $\forall \varphi \in U_k$.*

After this preparation, we state the semidiscrete optimal control problem:

$$\begin{aligned} \underset{(q_k,u_k)\in Q_{ad}\times U_k}{\text{Minimize}} \quad & J(q_k, u_k) = \frac{1}{2}\|u_k - u_d\|_I^2 + \frac{\alpha}{2}\|q_k\|_{L^2(I)}^2 \\ & \textit{subject to } (3.2) \textit{ and} \\ & G(u_k)\,|_{I_n} \le b, \; n = 1, ..., N. \end{aligned} \tag{3.3}$$

REMARK 3.2. *Indeed, given Assumption 2.1, the above problem (3.3) satisfies a regularity condition once k is sufficiently small. To see this, we note that (2.4) asserts the existence of $\epsilon > 0$ such that $G(u(\tilde{q})) \le b - \epsilon$, from which*

$$G(u_k(\tilde{q})) = G(u(\tilde{q})) + G(u_k(\tilde{q}) - u(q)) \le b - \epsilon + \|\omega\|_{L^\infty(\Omega)}\|u(\tilde{q}) - u_k(\tilde{q})\|_{L^\infty(I,V)}.$$

By the $L^\infty(I, V)$ convergence of the semi-discretization provided in Theorem 4.8, we conclude that $G(u_k(\tilde{q})) < b$ once k is sufficiently small.

PROPOSITION 3.3. *The semidiscrete optimal control problem (3.3) admits a unique solution $(\bar{q}_k, \bar{u}_k) \in Q_{ad} \times U_k$ once k is sufficiently small.*

Proof. The well-posedness of the problem follows by standard argument, utilizing that by Remark 3.2 there exists a feasible point once k is small enough. □

THEOREM 3.4. *Given Assumption 2.1 for the semi-discretized solution operator, the pair $(\bar{q}_k, \bar{u}_k) \in Q_{ad} \times U_k$ is optimal for (3.3) if and only if it is feasible and there exists a Lagrange multiplier $\bar{\mu}_k \in C(\bar{I})^*$ and an adjoint state $\bar{z}_k \in U_k$ satisfying the following system of optimality conditions*

$$B(\bar{u}_k, \varphi) = (\bar{q}_k g, \varphi)_I + (u_0, \varphi_0^+) \qquad \forall \varphi \in U_k, \tag{3.4a}$$
$$B(\varphi, \bar{z}_k) = (\bar{u}_k - u_d, \varphi)_I + \langle \bar{\mu}_k, 2(\nabla \bar{u}_k \nabla \varphi, \omega)\rangle \qquad \forall \varphi \in U_k, \tag{3.4b}$$
$$\alpha(\bar{q}_k, q - \bar{q}_k)_{L^2(I)} + (\bar{z}_k, (q - \bar{q}_k)g)_I \ge 0 \qquad \forall q \in Q_{ad}, \tag{3.4c}$$
$$\langle b - G(\bar{u}_k), \bar{\mu}_k\rangle = 0, \tag{3.4d}$$

where the Lagrange multiplier $\bar{\mu}_k$ is given by

$$\langle \bar{\mu}_k, v\rangle = \sum_{n=1}^{N} \frac{\mu_{k,n}}{k_n} \int_{I_n} v(t)dt, \;\; \forall v \in C(\bar{I}) \cup U_k(\mathbb{R}) \tag{3.5}$$

with $\mu_{k,n} \in \mathbb{R}^+$ for any $n = 1, ..., N$.

Proof. The proof moves along the same line of the continuous case with the difference given by the presence of the finitely many state constraints $G(u_k)\,|_{I_n} \le b, n = 1, ..., N$. This lead to a different definition of the closed convex cone $\{v \in \mathbb{R}^N \,|\, v_n \le 0,\, n = 1, ..., N\}$. As a consequence, we have the existence of Lagrange multipliers $\mu_{k,n} \in \mathbb{R}_+$ for all $n = 1, ..., N$, associated to the subintervals I_n. Then, we have $\bar{\mu}_k \in C(\bar{I})^*$ by construction in (3.5). □

REMARK 3.5. *The boundedness of the optimal pair $(\bar{q}_k, \bar{u}_k)$ and Lagrange multiplier $\bar{\mu}_k$ follows by standard argument. In particular, in view of Theorem 4.8 and Assumption 2.1, one has $J(\bar{q}_k, \bar{u}_k) \le J(\tilde{q}, u_k(\tilde{q})) \le c$, implying $\|\bar{u}_k\|_I + \|\bar{q}_k\|_{L^2(I)} \le c$. Exploiting the adjoint equation (3.4b), the variational inequality (3.8c) and the positivity of $\bar{\mu}_k$ one obtains $\|\bar{\mu}_k\|_{C(\bar{I})^*} \le c$, compare to [26, Lemma 6.2.].*

3.2. Space Discretization. We consider a family $\mathcal{T}_h$ of subdivisions 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 considered 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 [3] 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 $v \in V_h$. Then, the discrete state and trial space are given by

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

The discrete state equation reads: for $q \in Q$ find $u_{kh} \in U_{k,h}$ such that

$$B(u_{kh}(q), \varphi) = (qg, \varphi)_I + (u_0, \varphi_0^+), \tag{3.6}$$

for any $\varphi \in U_{k,h}$.

Then, the discrete optimal control problem is given by

$$\begin{aligned} \underset{(q_{kh},u_{kh})\in Q_{ad}\times U_{k,h}}{\text{Minimize}} \quad & J(q_{kh}, u_{kh}) = \frac{1}{2}\|u_{kh} - u_d\|_I^2 + \frac{\alpha}{2}\|q_{kh}\|_{L^2(I)}^2 \\ & \textit{subject to } (3.6) \textit{ and} \\ & G(u_{kh})\,|_{I_n} \le b,\; n = 1, ..., N. \end{aligned} \tag{3.7}$$

Using similar arguments as in the semidiscrete case, the regularity condition for the discrete problem is a consequence of Assumption 2.1 once k, h are sufficiently small utilizing the discretization error of the state equation shown in Theorem 4.12, compare Remark 3.2. In particular, Assumption 2.1 provides the existence of a feasible point for (3.7) once k, h are sufficiently small. Then, the existence of a unique optimal pair $(\bar{q}_{kh}, \bar{u}_{kh}) \in Q_{ad} \times U_{k,h}$ of (3.7) follows by standard arguments.

The optimality conditions are given in the following.

THEOREM 3.6. *Given Assumption 2.1 for the discretized solution operator, the pair $(\bar{u}_{kh}, \bar{q}_{kh}) \in U_{k,h} \times Q_{ad}$ is optimal for (3.3) if and only if it is feasible and there exists a Lagrange multiplier $\bar{\mu}_{kh} \in C(\bar{I})^*$ and an adjoint state $\bar{z}_{kh} \in U_{k,h}$ satisfying the following system of optimality conditions*

$$B(\bar{u}_{kh}, \varphi) = (\bar{q}_{kh} g, \varphi)_I + (u_0, \varphi_0^+) \qquad \forall \varphi \in U_{k,h}, \tag{3.8a}$$

$$B(\varphi, \bar{z}_{kh}) = (\bar{u}_{kh} - u_d, \varphi)_I + \langle \bar{\mu}_{kh}, 2(\nabla \bar{u}_{kh} \nabla \varphi, \omega)\rangle \qquad \forall \varphi \in U_{k,h}, \tag{3.8b}$$

$$\alpha(\bar{q}_{kh}, q - \bar{q}_{kh})_{L^2(I)} + (\bar{z}_{kh}, (q - \bar{q}_{kh})g)_I \ge 0 \qquad \forall q \in Q_{ad}, \tag{3.8c}$$

$$\langle b - G(\bar{u}_{kh}), \bar{\mu}_{kh}\rangle = 0, \tag{3.8d}$$

where the Lagrange multiplier $\bar{\mu}_{kh}$ is given by

$$\langle \bar{\mu}_{kh}, v\rangle = \sum_{n=1}^{N} \frac{\mu_{kh,n}}{k_n} \int_{I_n} v(t)dt,\ \forall v \in C(\bar{I}) \cup U_k(\mathbb{R}) \tag{3.9}$$

with $\mu_{kh,n} \in \mathbb{R}^+$ for any $n = 1, ..., N$.

Proof. The proof follows by standard arguments, compare to Theorem 3.4 ◻

REMARK 3.7. *The boundedness of $(\bar{q}_{kh}, \bar{u}_{kh})$ and $\bar{\mu}_{kh}$ independent of the discretization parameters follows as in Remark 3.5, compare to [26, Lemma 6.5].*

4. The State Equation . In this section, we derive the $L^\infty(I, V)$ error estimate for the state equation. The derivation employs a duality technique for parabolic equations requiring at any level of discretization the introduction of the homogeneous (uncontrolled) backward counterpart of the state equation, see, e.g., [24], [36].

In the following subsection, we analyze the stability of the continuous backward problem together with an additional auxiliary problem. In Section 4.2 and 4.3, we inspect the temporal and spatial error, respectively.

4.1. Continuous Auxiliary Solutions. For a given $w_T \in V^* := H^{-1}(\Omega)$, we consider the problem to find $w \in W := W(I) = L^2(I, L^2(\Omega)) \cap H^1\big(I, (\dot{H}^2)^*\big)$ such that

$$\begin{aligned} -(\varphi, \partial_t w)_I + (\nabla\varphi, \nabla w)_I &= 0, \\ w(T) &= w_T, \end{aligned} \tag{4.1}$$

for any $\varphi \in L^2(I, H^2(\Omega)) \cap H^1(I, L^2(\Omega))$, see [23, Chapter 4, Section 8] and Lemma 4.2 below.

REMARK 4.1. *We observe that, thanks to $[L^2(\Omega), (\dot{H}^2)^*]_{1/2} = V^*$, compare to [22, Chapter 1, Theorem 12.5], there holds the embedding $W \hookrightarrow C(\bar{I}, V^*)$, see [22, Chapter 1, Theorem 3.1].*

Further, we need an additional backward continuous problem on the truncated time interval $\widehat{I} = (0, \hat{t})$, where $\hat{t} \in I_N$. Find $\widehat{w} \in W(\widehat{I})$ such that

$$\begin{aligned} -(\varphi, \partial_t \widehat{w})_{\widehat{I}} + (\nabla\varphi, \nabla \widehat{w})_{\widehat{I}} &= 0, \\ \widehat{w}(\hat{t}\,) &= w_T, \end{aligned} \tag{4.2}$$

for any $\varphi \in L^2(\widehat{I}, H^2(\Omega)) \cap H^1(\widehat{I}, L^2(\Omega))$.

We start the analysis with a regularity result for the solution of (4.1). The following regularity result extends the well-known energy estimates for linear parabolic equations with homogeneous Dirichlet data. This result is already present in some classical books, see, e.g., [23]; we include it to keep the exposition self-contained.

LEMMA 4.2. *Let $w \in W$ be the solution of* (4.1). *Then, there holds*

$$\|w\|_I + \max_{t \in \bar{I}} \|w(t)\|_{H^{-1}(\Omega)} \le C \|w_T\|_{H^{-1}(\Omega)}. \tag{4.3}$$

Proof. We test (4.1) with $\varphi = -\Delta^{-1} w$, obtaining

$$(\Delta^{-1} w, \partial_t w)_I - (\nabla \Delta^{-1} w, \nabla w)_I - (w(T) - w_T, \Delta^{-1} w(T)) = 0. \tag{4.4}$$

We note that (4.4) holds also pointwise almost everywhere on I and $w(T) = w_T$. Then for a.e. $t \in I$

$$(\Delta^{-1} w(t), \partial_t w(t)) - (\nabla \Delta^{-1} w(t), \nabla w(t)) = 0. \tag{4.5}$$

We reformulate the first term, using the relation $\partial_t w = -\Delta w$, obtaining

$$(\Delta^{-1} w(t), \partial_t w(t)) = -\|w(t)\|^2.$$

For the second, we see, by analogous arguments,

$$\begin{aligned} -(\nabla \Delta^{-1} w(t), \nabla w(t)) &= -(\Delta^{-1} w(t), \partial_t w(t)) \\ &= (\nabla \Delta^{-1} w(t), \nabla \Delta^{-1} \partial_t w(t)). \end{aligned}$$

Then, observing that the time derivative interchanges with ∇ and Δ^{-1}, we have

$$(\nabla\Delta^{-1}w(t), \nabla\Delta^{-1}\partial_t w(t)) = \frac{1}{2}\frac{d}{dt}\|\nabla\Delta^{-1}w(t)\|^2.$$

Thus, it follows from (4.5) that

$$\frac{d}{dt}\|\nabla\Delta^{-1}w(t)\|^2 = 2\|w(t)\|^2. \tag{4.6}$$

Integrating (4.6) over (t,T) and defining $\eta(t) = \|\nabla\Delta^{-1}w(t)\|^2$,$\psi(t) = \|w(t)\|^2$, we obtain

$$\eta(t) + 2\int_t^T \psi(s)\,ds = \eta(T).$$

Noting that both η and ψ are nonnegative, this shows the assertion. □

In a next step, we derive an error estimate for the solution of (4.1) and (4.2). In Section 4.2, we use this estimate to investigate the error at the nodal points of the time discretization.

Lemma 4.3. *Let w and $\widehat{w}$ be solutions of* (4.1) *and* (4.2)*, respectively. Then the error satisfies*

$$\|w-\widehat{w}\|_{L^1(\widehat{I},H^{-1}(\Omega))} + \|w(0)-\widehat{w}(0)\|_{H^{-3}(\Omega)} \le Ck\Big(\log\frac{T}{k}+1\Big)^{\frac{1}{2}}\|w_T\|_{H^{-1}(\Omega)}. \tag{4.7}$$

Proof. In a first step, we derive the equation for the error $\varepsilon := \widehat{w} - w$. Then, subtracting (4.1) from (4.2), integrating only on $\widehat{I}$, we obtain

$$-(\varphi, \partial_t\varepsilon)_{\widehat{I}} + (\nabla\varphi, \nabla\varepsilon)_{\widehat{I}} = 0$$

for any $\varphi \in L^2(\widehat{I}, H^2(\Omega)) \cap H^1(\widehat{I}, L^2(\Omega))$.

Integration by parts in the second term gives

$$-(\varphi, \partial_t\varepsilon)_{\widehat{I}} - (\Delta\varphi, \varepsilon)_{\widehat{I}} = 0. \tag{4.8}$$

The proof is now divided in two parts corresponding to the two terms in the left-hand side of (4.7). We start estimating $\|\varepsilon(0)\|_{H^{-3}(\Omega)}$.

(i) Testing (4.8) with $\varphi = -\Delta^{-3}\varepsilon$, we have

$$(\Delta^{-3}\varepsilon, \partial_t\varepsilon)_{\widehat{I}} + (\Delta^{-2}\varepsilon, \varepsilon)_{\widehat{I}} = 0. \tag{4.9}$$

Observing that

$$(\Delta^{-3}\varepsilon, \partial_t\varepsilon)_{\widehat{I}} = \int_{\widehat{I}} \partial_t(\Delta^{-3}\varepsilon, \varepsilon)dt - (\partial_t(\Delta^{-3}\varepsilon), \varepsilon)_{\widehat{I}},$$

we rewrite (4.9) as

$$\begin{aligned} -(\Delta^{-3}\varepsilon(0), \varepsilon(0)) - (\partial_t(\Delta^{-3}\varepsilon), \varepsilon)_{\widehat{I}} + (\Delta^{-2}\varepsilon, \varepsilon)_{\widehat{I}} \\ = -(\Delta^{-3}\varepsilon(\hat{t}), w_T - w(\hat{t})). \end{aligned} \tag{4.10}$$

We consider each term in the last equation separately. Integration by parts in space gives

$$-(\Delta^{-3}\varepsilon(0), \varepsilon(0)) = (\nabla\Delta^{-2}\varepsilon(0), \nabla\Delta^{-2}\varepsilon(0)). \tag{4.11}$$

Further, the relation $\partial_t \varepsilon = -\Delta\varepsilon$ and Δ^{-1} being self-adjoint implies

$$-(\partial_t(\Delta^{-3}\varepsilon), \varepsilon)_{\widehat{I}} = (\Delta^{-2}\varepsilon, \varepsilon)_{\widehat{I}} = (\Delta^{-1}\varepsilon, \Delta^{-1}\varepsilon)_{\widehat{I}}. \tag{4.12}$$

To estimate the right-hand side of (4.10), we observe that

$$\varepsilon(\hat{t}\,) = w_T - w(\hat{t}\,).$$

Then, with the help of the Cauchy-Schwarz inequality, the Fubini-Tonelli theorem, and Lemma 4.2, we have

$$\begin{aligned}
-(\Delta^{-3}\varepsilon(\hat{t}\,), w_T - w(\hat{t}\,)) &= \|\nabla\Delta^{-2}\varepsilon(\hat{t}\,)\|^2 \\
&= \int_\Omega \Big(\int_{\hat{t}}^T \nabla\Delta^{-2}\partial_t w(t)dt \Big)^2 dx \\
&= \int_\Omega \Big(\int_{\hat{t}}^T -\nabla\Delta^{-1} w(t)dt \Big)^2 dx \\
&\leq k \int_\Omega \Big(\int_{\hat{t}}^T |\nabla\Delta^{-1} w(t)|^2 dt \Big) dx \\
&= k \int_{\hat{t}}^T \int_\Omega |\nabla\Delta^{-1} w(t)|^2 dx dt \\
&\leq Ck^2 \|\nabla\Delta^{-1} w_T\|^2.
\end{aligned} \tag{4.13}$$

Combining (4.10) with the relations (4.11), (4.12) and the estimate of the right-hand side (4.13), we conclude

$$\|\nabla\Delta^{-2}\varepsilon(0)\|^2 + 2\|\Delta^{-1}\varepsilon\|_{\widehat{I}}^2 \leq Ck^2 \|\nabla\Delta^{-1} w_T\|^2. \tag{4.14}$$

(ii) To derive an estimate for $\|\varepsilon\|_{L^1(I,H^{-1}(\Omega))}$, we set $\varphi = \tau\Delta^{-2}\varepsilon$ in (4.8), where $\tau(t) = \max(\hat{t} - t, k)$ for $t \in \widehat{I}$.
Assuming that the following relation has already been derived

$$\|\sqrt{\tau}\nabla\Delta^{-1}\varepsilon\|_{\widehat{I}}^2 \leq Ck^2 \|\nabla\Delta^{-1} w_T\|^2. \tag{4.15}$$

It follows

$$\begin{aligned}
\|\varepsilon\|^2_{L^1(\widehat{I},H^{-1}(\Omega))} &\leq \|\sqrt{\tau}^{-1}\|^2_{L^2(\widehat{I})} \|\sqrt{\tau}\varepsilon\|^2_{L^2(\widehat{I},H^{-1}(\Omega))} \\
&\leq Ck^2 \Big(\log\frac{T}{k} + 1 \Big) \|w_T\|^2_{H^{-1}(\Omega)}.
\end{aligned}$$

Therefore, we focus in the derivation of (4.15). Inserting $\varphi = \tau\Delta^{-2}\varepsilon$ in (4.8), it follows

$$-(\tau\Delta^{-2}\varepsilon, \partial_t\varepsilon)_{\widehat{I}} - (\tau\Delta^{-1}\varepsilon, \varepsilon)_{\widehat{I}} = 0. \tag{4.16}$$

We reformulate the first term on the left-hand side using the relation

$$-(\tau\Delta^{-2}\varepsilon, \partial_t\varepsilon)_{\widehat{I}} = -\frac{1}{2}\int_{\widehat{I}} \frac{\partial}{\partial t}\Big(\tau(\Delta^{-2}\varepsilon(t), \varepsilon(t))\Big)dt + \frac{1}{2}\int_{\widehat{I}} \tau^{'}(\Delta^{-2}\varepsilon(t), \varepsilon(t))$$

where τ' denotes the first derivative of τ with respect to t. The second term in (4.16) is handled by

$$-(\tau\Delta^{-1}\varepsilon, \varepsilon)_{\widehat{I}} = \|\sqrt{\tau}\nabla\Delta^{-1}\varepsilon\|_{\widehat{I}}^2.$$

Then, observing that $-\tau^{'} \leq 1$ and $\varepsilon(\hat{t}) = w_T - w(\hat{t}\,)$, we obtain from (4.16)

$$\frac{\hat{t}}{2}\|\Delta^{-1}\varepsilon(0)\|^2 + \|\sqrt{\tau}\nabla\Delta^{-1}\varepsilon\|^2_{\hat{I}} \leq \frac{\|\Delta^{-1}\varepsilon\|^2_{\hat{I}}}{2} + \frac{k}{2}(\Delta^{-2}\varepsilon(\hat{t}\,), w_T - w(\hat{t}\,)). \tag{4.17}$$

In the next step, we estimate the second term in the right-hand side of the previous expression. Thanks to (4.13) and Lemma 4.2, it follows

$$\begin{aligned} k(\Delta^{-2}\varepsilon(\hat{t}\,), w_T - w(\hat{t}\,)) &= -k(\nabla\Delta^{-2}\varepsilon(\hat{t}\,), \nabla\Delta^{-1}(w_T - w(\hat{t}\,)) \\ &\leq k\|\nabla\Delta^{-2}\varepsilon(\hat{t}\,)\|\|\nabla\Delta^{-1}(w_T - w(\hat{t}\,)\| \\ &\leq Ck^2\|\nabla\Delta^{-1}w_T\|^2. \end{aligned}$$

Then, from (4.17) and thanks to (4.14) we conclude

$$\hat{t}\|\Delta^{-1}\varepsilon(0)\|^2 + 2\|\sqrt{\tau}\nabla\Delta^{-1}\varepsilon\|^2_{\hat{I}} \leq Ck^2\|\nabla\Delta^{-1}w_T\|^2.$$

This establishes (4.15) as required.

□

We conclude the section with a time weighted stability result for the solution of (4.1). This estimate will be used later in the derivation of the temporal error in the interior of the time interval. A similar technique has been used in [26, Theorem 4.4], see also [13, Lemma 1].

LEMMA 4.4. *Let $w \in W$ be solution of (4.1). Then there holds*

$$\int_I (T-t)\|\partial_t w(t)\|^2_{H^{-1}(\Omega)}dt \leq C\|w_T\|^2_{H^{-1}(\Omega)}, \tag{4.18}$$

$$\int_{I\setminus I_N} \|\partial_t w(t)\|_{H^{-1}(\Omega)}dt \leq C\Big(\log\frac{T}{k}\Big)^{\frac{1}{2}}\|w_T\|_{H^{-1}(\Omega)}. \tag{4.19}$$

Proof. We start with the first relation. The choice $\varphi = (T-t)\Delta^{-1}\partial_t w$ in (4.1) leads to

$$-\int_I (T-t)(\Delta^{-1}\partial_t w, \partial_t w)dt + \int_I (T-t)(\nabla\Delta^{-1}\partial_t w, \nabla w)dt = 0. \tag{4.20}$$

We observe that

$$-\int_I (T-t)(\Delta^{-1}\partial_t w, \partial_t w) = \int_I (T-t)\|\nabla\Delta^{-1}\partial_t w\|^2 dt,$$

and

$$\begin{aligned} \int_I (T-t)(\nabla\Delta^{-1}\partial_t w, \nabla w) &= -\int_I (T-t)(\partial_t w, w)dt \\ &= -\frac{1}{2}\|w\|^2_I - \frac{1}{2}\int_I \frac{d}{dt}((T-t)\|w(t)\|^2)dt. \end{aligned}$$

Then, from (4.20) we conclude

$$\int_I (T-t)\|\nabla\Delta^{-1}\partial_t w\|^2 dt + \frac{T}{2}\|w(0)\|^2 = \frac{1}{2}\|w\|^2_I \leq C\|\nabla\Delta^{-1}w_T\|^2$$

where in the last step we used Lemma 4.2.

The second relation directly follows from (4.18) by means of Cauchy-Schwarz inequality. In fact, recalling that $k \neq T$, there holds

$$\begin{aligned}\int_{I\setminus I_N} \|\partial_t w(t)\|_{H^{-1}(\Omega)} dt &\leq \Big(\int_{I\setminus I_N} (T-t)^{-1} dt\Big)^{\frac{1}{2}} \Big(\int_{I\setminus I_N} (T-t)\|\partial_t w(t)\|^2_{H^{-1}(\Omega)} dt\Big)^{\frac{1}{2}} \\ &\leq C\Big(\log\frac{T}{k}\Big)^{\frac{1}{2}} \Big(\int_I (T-t)\|\partial_t w(t)\|^2_{H^{-1}(\Omega)} dt\Big)^{\frac{1}{2}}. \\ &\leq C\Big(\log\frac{T}{k}\Big)^{\frac{1}{2}} \|\nabla\Delta^{-1} w_T\|.\end{aligned}$$

□

4.2. Temporal Discretization Error Estimates. We now focus on the derivation of the $L^\infty(I,V)$ error estimate for the temporal discretization error which will be given in Theorem 4.8.

In a first step, we introduce the semidiscrete counterpart of (4.1). For a given $w_T \in H^{-1}(\Omega)$, find $w_k \in U_k(V^*)$ such that

$$B(\varphi, w_k) = (\varphi_N, w_T) \tag{4.21}$$

for any $\varphi \in U_k(V)$.

As in [26, Lemma 5.2], we introduce a projection operator $\pi_k : C(\bar{I} \setminus I_N, V^*) \to U_k(V^*)$ onto the semi-discrete space, defined by the relation

$$\pi_k w\,|_{I_n} = w(t_{n-1}), \tag{4.22}$$

and establish a system of equations for the error between (4.22), when applied to the solution of (4.1), and the solution of (4.21).

Lemma 4.5. *For the error $\varepsilon = \pi_k w - w_k$ between the solutions w given by* (4.1) *and w_k given by* (4.21) *with the same initial value w_T, it holds for any $n = 1, \ldots, N$ and $\varphi \in \mathcal{P}_0(I_n, H^2(\Omega) \cap V)$*

$$(\nabla\varphi, \nabla\varepsilon_k)_{I_n} - (\varphi_n, [\varepsilon_k]_n) = \int_{I_n} (t_n - t)(\Delta\varphi, \partial_t w(t)) dt. \tag{4.23}$$

Proof. The relation can be obtained by Galerkin orthogonality and the definition of the semidiscrete projection, as in [26, Lemma 5.2]. □

Remark 4.6. *We observe that, thanks to the embedding $U \hookrightarrow C(\bar{I}, V)$, the application of the semidiscrete projector to elements of the continuous state space is well-posed. Further, we can extend π_k to U_k by letting $\pi_k\big|_{U_k} = \mathrm{Id}_{U_k}$.*

The following result will be used in Theorem 4.8 for the error estimate in the interior of the time interval.

Lemma 4.7. *Let w and w_k be solutions of* (4.1) *and* (4.21)*, respectively. Then the corresponding error satisfies*

$$\|w - w_k\|_{L^1(I,H^{-1}(\Omega))} + \|w(0) - w_{k,1}\|_{H^{-3}(\Omega)} \leq Ck\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}} \|w_T\|_{H^{-1}(\Omega)}. \tag{4.24}$$

Proof. We recall the abbreviation $\varepsilon_k = \pi_k w - w_k$ and separate the proof in to two parts corresponding to the norms in the left-hand side of the assertion.

(i) We observe, that

$$\|w(0) - w_{k,1}\|_{H^{-3}(\Omega)} = \|\pi_k w(0) - w_{k,1}\|_{H^{-3}(\Omega)} = \|\varepsilon_{k,1}\|_{H^{-3}(\Omega)}. \tag{4.25}$$

We set $\varphi = -\Delta^{-3}\varepsilon_k$ in (4.23) and obtain, using integration by parts and $\partial_t w = -\Delta w$,

$$(\Delta^{-2}\varepsilon_k, \varepsilon_k)_{I_n} + (\Delta^{-3}\varepsilon_{k,n}, [\varepsilon_k]_n) = \int_{I_n} (t_n - t)(\Delta^{-1}\varepsilon_k, w(t))dt. \tag{4.26}$$

Then, noticing that

$$(\Delta^{-3}\varepsilon_{k,n}, [\varepsilon_k]_n) = -(\nabla\Delta^{-2}\varepsilon_{k,n}, [\nabla\Delta^{-2}\varepsilon_k]_n),$$

and thanks to the equality

$$-(\varphi_n, [\varphi]_n) = \frac{1}{2}(-\|\varphi_{n+1}\|^2 + \|[\varphi]_n\|^2 + \|\varphi_n\|^2) \quad \forall\phi \in U_k, \tag{4.27}$$

we obtain from (4.26)

$$\begin{aligned} \|\Delta^{-1}\varepsilon_k\|_{I_n}^2 + &\frac{1}{2}(\|\nabla\Delta^{-2}\varepsilon_{k,n}\|^2 - \|\nabla\Delta^{-2}\varepsilon_{k,n+1}\|^2) \\ &\leq \int_{I_n} (t_n - t)(\Delta^{-1}\varepsilon_k, w(t))dt. \end{aligned} \tag{4.28}$$

Summation over $n = 1, ..., N$,Lemma 4.2, and $\varepsilon_{k,N+1} = 0$ gives

$$\begin{aligned} \|\Delta^{-1}\varepsilon_k\|_I^2 + \|\nabla\Delta^{-2}\varepsilon_{k,1}\|^2 &\leq Ck^2\|w\|_I^2 \\ &\leq Ck^2\|\nabla\Delta^{-1}w_T\|^2. \end{aligned}$$

We conclude

$$\|\nabla\Delta^{-2}\varepsilon_{k,1}\|^2 \leq Ck^2\|\nabla\Delta^{-1}w_T\|^2. \tag{4.29}$$

(ii) To estimate the $L^1(I, H^{-1}(\Omega))$-norm, we utilize the splitting

$$\begin{aligned} \|w - w_k\|_{L^1(I,H^{-1}(\Omega))} \leq &\|\varepsilon_k\|_{L^1(I,H^{-1}(\Omega))} \\ &+ \|w - \pi_k w\|_{L^1(I,H^{-1}(\Omega))}. \end{aligned} \tag{4.30}$$

In this part, we derive an upper bound for $\|\varepsilon_k\|_{L^1(I,H^{-1}(\Omega))}$. Before starting, we note that it will be sufficient to show

$$\sum_{n=1}^{N} \tau_{k,n}\|\nabla\Delta^{-1}\varepsilon_k\|_{I_n}^2 \leq Ck^2\|\nabla\Delta^{-1}w_T\|^2. \tag{4.31}$$

Then, the required estimate follows by

$$\begin{aligned} \|\varepsilon_k\|_{L^1(I,H^{-1}(\Omega))}^2 &\leq \sum_{n=1}^{N} k_n\tau_{k,n}^{-1} \sum_{n=1}^{N} \tau_{k,n}\|\nabla\Delta^{-1}\varepsilon_k\|_{I_n}^2 \\ &\leq Ck^2\Big(\log\frac{T}{k} + 1\Big)\|w_T\|_{H^{-1}(\Omega)}^2. \end{aligned}$$

Therefore, we show that (4.31) holds. Testing (4.23) with $\varphi := \tau_{k,n}\Delta^{-2}\varepsilon_k$, where $\tau_{k,n} := T - t_{n-1}$, we have

$$\begin{aligned} -(\tau_{k,n}\Delta^{-1}\varepsilon_k, \varepsilon_k)_{I_n} - (\tau_{k,n}\Delta^{-2}\varepsilon_{k,n}, [\varepsilon_k]_n) \\ = \int_{I_n} (t_n - t)(\tau_{k,n}\Delta^{-1}\varepsilon_k, \partial_t w(t))dt. \end{aligned} \tag{4.32}$$

For the left-hand of (4.32) side, we have, using (4.27),

$$-(\tau_{k,n}\Delta^{-1}\varepsilon_k, \varepsilon_k)_{I_n} = \tau_{k,n}\|\nabla\Delta^{-1}\varepsilon_k\|^2_{I_n},$$
$$-(\tau_{k,n}\Delta^{-2}\varepsilon_{k,n}, [\varepsilon_k]_n) = \frac{1}{2}\tau_{k,n}(-\|\Delta^{-1}\varepsilon_{k,n+1}\|^2 + \|\Delta^{-1}\varepsilon_{k,n}\|^2 + \|[\Delta^{-1}\varepsilon_k]_n\|^2).$$

We estimate the right-hand side of (4.32) as follows

$$\begin{aligned} &\int_{I_n} (t_n - t)(\tau_{k,n}\Delta^{-1}\varepsilon_k, \partial_t w(t))dt \\ &\quad = -\int_{I_n} (t_n - t)(\tau_{k,n}\nabla\Delta^{-1}\varepsilon_k, \nabla\Delta^{-1}\partial_t w(t))dt \\ &\quad \le \frac{\tau_{k,n}}{2}\|\nabla\Delta^{-1}\varepsilon_k\|^2_{I_n} + \frac{\tau_{k,n}}{2}\int_{I_n} (t_n - t)^2\|\nabla\Delta^{-1}\partial_t w(t)\|^2 dt. \end{aligned}$$

Combining the previously derived relations, using the equality $\tau_{k,n} = \tau_{k,n+1} + k_n$ for the term $\|\Delta^{-1}\varepsilon_{k,n+1}\|^2$, it follows

$$\begin{aligned} &\tau_{k,n}\|\nabla\Delta^{-1}\varepsilon_k\|^2_{I_n} + \tau_{k,n}\|\Delta^{-1}\varepsilon_{k,n}\|^2 - \tau_{k,n+1}\|\Delta^{-1}\varepsilon_{k,n+1}\|^2 \\ &\qquad \le k_n\|\Delta^{-1}\varepsilon_{k,n+1}\|^2 + \tau_{k,n}\int_{I_n} (t_n - t)^2 \ \|\nabla\Delta^{-1}\partial_t w(t)\|^2 dt. \end{aligned}$$

Summing over $n = 1, ..., N$, using $\varepsilon_{k,N+1} = 0$ and recalling that $k_n \le \tilde{k}k_{n+1}$, we have

$$\begin{aligned} &\sum_{n=1}^{N} \tau_{k,n}\|\nabla\Delta^{-1}\varepsilon_k\|^2_{I_n} + T\|\Delta^{-1}\varepsilon_{k,1}\|^2 \\ &\qquad \le \tilde{k}\|\Delta^{-1}\varepsilon_k\|^2_I \\ &\qquad\quad + \sum_{n=1}^{N} \tau_{k,n}\int_{I_n} (t_n - t)^2\|\nabla\Delta^{-1}\partial_t w(t)\|^2 dt. \end{aligned} \tag{4.33}$$

We note that for $t \in I_n$ and $n = 1, ..., N-1$ it holds $\tau_{k,n} \le (1 + \tilde{k})(T - t)$, while $\tau_{k,N} = k_N$. This observation suggests the following splitting for the second term in the right-hand side of (4.33)

$$\begin{aligned} &\sum_{n=1}^{N} \tau_{k,n}\int_{I_n} (t_n - t)^2\|\nabla\Delta^{-1}\partial_t w(t)\|^2 dt \\ &\quad \le \sum_{n=1}^{N-1} k_n^2\int_{I_n} \tau_{k,n}\|\nabla\Delta^{-1}\partial_t w(t)\|^2 dt + k_N^2\int_{I_N} (T - t)\|\nabla\Delta^{-1}\partial_t w(t)\|^2 dt \\ &\quad \le (1 + \tilde{k})k^2\int_I (T - t)\|\nabla\Delta^{-1}\partial_t w(t)\|^2 dt. \end{aligned}$$

In conclusion, inserting the above bound in (4.33), using Lemma 4.4 and the estimate on $\|\Delta^{-1}\varepsilon_k\|_I^2$ from (4.29), we obtain (4.31).

(iii) We continue and estimate the remaining term on the left of (4.30). In view of (4.19), we introduce the splitting

$$\|w-\pi_k w\|_{L^1(I,H^{-1}(\Omega))} = \int_{I\setminus I_N} \|w-\pi_k w\|_{H^{-1}(\Omega)} dt + \int_{I_N} \|w-\pi_k w\|_{H^{-1}(\Omega)} dt.$$

To estimate the first term we observe that Δ^{-1}, being independent of t, interchanges with π_k. As a consequence, it holds

$$\begin{aligned}\int_{I\setminus I_N} \|w-\pi_k w\|_{H^{-1}(\Omega)} dt &= \int_{I\setminus I_N} \|\nabla\Delta^{-1}(w-\pi_k w)\| dt,\\ &= \int_{I\setminus I_N} \|\nabla(\Delta^{-1}w-\pi_k\Delta^{-1}w)\| dt.\end{aligned}$$

We note that the projection operator π_k is an interpolation operator acting at the nodal points in time. By standard transformation arguments, we assert

$$\int_{I\setminus I_N} \|w-\pi_k w\|_{H^{-1}(\Omega)} dt \le ck \int_{I\setminus I_N} \|\partial_t w(t)\|_{H^{-1}(\Omega)} dt, \tag{4.34}$$

compare [36, Equation (12.10)].

For the second term, it clearly holds

$$\int_{I_N} \|w-\pi_k w\|_{H^{-1}(\Omega)} dt \le ck \max_{t\in I_N} \|w\|_{H^{-1}(\Omega)}.$$

Then, Lemma 4.2 and Lemma 4.4 give the desired bound.

□

After this preparation, we are ready to show the temporal discretization error estimate.

THEOREM 4.8. *Let $u\in U$ and $u_k \in U_k$ be solution of* (2.1) *and* (3.2)*, respectively, with $f(x,t) = q(t)g(x) \in L^\infty(I,V)$ and $u_0 \in \dot{H}^3(\Omega)$. Then for the semi-discretization error it holds*

$$\|u-u_k\|_{L^\infty(I,V)} \le Ck\Big(\log\frac{T}{k}+1\Big)^{\frac{1}{2}}\Big(\|f\|_{L^\infty(I,V)} + \|u_0\|_{H^3(\Omega)}\Big). \tag{4.35}$$

Proof. Defining $\xi_k = u - u_k$, on each time interval I_n, $n = 1, ..., N$, we consider the following splitting of the error

$$\|\xi_k\|_{L^\infty(I_n,V)} \le \|u(\cdot)-u(t_n)\|_{L^\infty(I_n,V)} + \|u(t_n)-u_k(\cdot)\|_{L^\infty(I_n,V)}. \tag{4.36}$$

We estimate the two terms in the right-hand side separately on each time interval I_n. Then, summing over $n = 1, .., N$ the resulting estimates gives the assertion.

With no loss of generality, we focus on the last time interval I_N denoting by $\hat{t} \in I_N$ a generic fixed time. For a generic I_n, the proof follows by similar arguments considering (4.1) on $I = (0, t_n)$ and (4.2) on $\widehat{I} = (0, \hat{t})$ for $\hat{t} \in (t_{n-1}, t_n]$, noting that $0 \le \log(t_n/k) \le \log(T/k)$.

(i) We start the analysis with the interpolation error $u(\hat{t}) - u(t_N)$. We consider the solutions w and $\hat{w}$ to (4.1) and (4.2), respectively, with terminal value w_T to be specified later. Integration by parts in time of (4.1) and (4.2) leads to

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

for any $\varphi \in U$.
In particular, setting $\varphi = u$ it follows from the state equation (2.1) that

$$-(u(T), w(T)) + (u(0), w(0)) + (f, w)_I = 0, \tag{4.37}$$
$$-(u(\hat{t}), \hat{w}(\hat{t})) + (u(0), \hat{w}(0)) + (f, \hat{w})_{\hat{I}} = 0. \tag{4.38}$$

Since by definition $w(T) = w(\hat{t}) = w_T$, subtracting (4.37) from (4.38) yields

$$(u(\hat{t}) - u(T), w_T) = (u(0), \hat{w}(0) - w(0)) + (f, \hat{w} - w)_{\hat{I}} - (f, w)_{I \setminus \hat{I}}. \tag{4.39}$$

We observe that the choice $w_T = -\Delta(u(\hat{t}) - u(T))$ and integration by parts for the left-hand side of (4.39) gives $\|\nabla(u(\hat{t}) - u(T))\|^2$.
Therefore, we have

$$\begin{aligned}
\|\nabla(u(\hat{t}) - u(T))\|^2 &= (u(0), \hat{w}(0) - w(0)) + (f, \hat{w} - w)_{\hat{I}} - \int_{\hat{t}}^{T} (f(t), w(t)) dt \\
&\leq \Big(\|\hat{w} - w\|_{L^1(\hat{I}, H^{-1}(\Omega))} + \|\hat{w}(0) - w(0)\|_{H^{-3}(\Omega)} \\
&\quad + k\|w\|_{L^\infty(I, H^{-1}(\Omega))} \Big) \Big(\|f\|_{L^\infty(I,V)} + \|u_0\|_{H^3(\Omega)} \Big) \\
&\leq Ck \log\left(\frac{T}{k} + 1\right)^{\frac{1}{2}} \|w_T\|_{H^{-1}(\Omega)} (\|f\|_{L^\infty(I,V)} + \|u_0\|_{H^3(\Omega)}).
\end{aligned}$$

In the last step we used Lemma 4.2 to bound $\|w\|_{L^\infty(I,H^{-1}(\Omega))}$ and Lemma 4.3 for the remaining terms.
Then, observing that it holds $\|w_T\|_{H^{-1}(\Omega)} = \|\nabla(u(\hat{t}) - u(T))\|$ due to the choice of the terminal value, we conclude

$$\|\nabla(u(\hat{t}) - u(T))\| \leq Ck \log\left(\frac{T}{k} + 1\right)^{\frac{1}{2}} (\|f\|_{L^\infty(I,V)} + \|u_0\|_{H^3(\Omega)}). \tag{4.40}$$

(ii) We consider w and w_k solutions of (4.1) and (4.21), respectively, with terminal value $w_T = -\Delta(u(t) - u(T))$.
This choice gives

$$B(\varphi, w) = B(\varphi, w_k) = (\varphi_N, -\Delta(u(T) - u_k(t))) = (\nabla \varphi_N, \nabla(u(T) - u_{k,N}))$$

for any $\varphi \in U_k$.
In particular, by means of Galerkin orthogonality and (3.2), we have

$$\begin{aligned}
\|\nabla(u(T) - u_{k,N})\|^2 &= B(u - u_k, w) = B(u, w - w_k) \\
&= (f, w - w_k)_I + (u_0, w(0) - w_k(0)) \\
&\leq (\|w - w_k\|_{L^1(I,H^{-1}(\Omega))} + \|w(0) - w_k(0)\|_{H^{-3}(\Omega)}) \\
&\times (\|f\|_{L^\infty(I,V)} + \|u_0\|_{H^3(\Omega)}).
\end{aligned}$$

Then, thanks to Lemma 4.7 we obtain

$$\|\nabla(u(T) - u_{k,N})\| \leq Ck \log\left(\frac{T}{k} + 1\right)^{\frac{1}{2}} (\|f\|_{L^\infty(I,V)} + \|u_0\|_{H^3(\Omega)}). \quad (4.41)$$

In conclusion, inserting (4.40) and (4.41) in (4.36), we have shown the desired estimate for the interval I_N. □

We conclude the section with a stability result for the solutions of the auxiliary problems

LEMMA 4.9. *For w_k solution of* (4.21)*, there holds*

$$\|w_k\|_I + \|w_{k,1}\|_{H^{-1}(\Omega)} \leq c\|w_T\|_{H^{-1}(\Omega)} \quad (4.42)$$

Proof. Using integration by parts, the bilinear form (3.1) is formulated as

$$B(\varphi, w_k) = -\sum_{n=1}^{N} (\varphi, \partial_t w_k)_{I_n} + (\nabla\varphi, \nabla w_k)_I - \sum_{n=1}^{N-1} (\varphi_n, [w_k]_n) + (\varphi_N, w_{k,N}). \quad (4.43)$$

Then, observing $w_{k,N} = w_T$, we rewrite (4.21) for any I_n, $n = 1, ..., N-1$, as

$$(\nabla\varphi, \nabla w_k)_{I_n} - (\varphi_n, [w_k]_n) = 0, \quad \forall \varphi \in \mathcal{P}_0(I_n, V_h). \quad (4.44)$$

By the choice $\varphi = -\Delta^{-1} w_k$, we have

$$\|w_k\|_{I_n}^2 - (\nabla\Delta^{-1} w_{k,n}, [\nabla\Delta^{-1} w_k]_n) = 0.$$

The second term in the left-hand side can be expressed by the identity (4.27) leading to

$$\|w_k\|_{I_n}^2 + \|\nabla\Delta^{-1} w_{k,n}\|^2 - \|\nabla\Delta^{-1} w_{k,n+1}\|^2 \leq 0.$$

The assertion follows summing over $n = 1, ..., N-1$. □

4.3. Spatial Discretization Error Estimates. In this section, we derive the spatial $L^\infty(I, V)$ error estimate with a series of lemmas culminating in the main result, namely Theorem 4.12.

We introduce the discrete counterpart of (4.21) to find $w_{kh} \in U_{k,h}(V^*)$ such that

$$B(\varphi, w_{kh}) = (\varphi_N, w_T), \quad \forall \varphi \in U_{k,h}, \quad (4.45)$$

where $w_T \in H^{-1}(\Omega)$.

Further, for given $v_0 \in \dot{H}^2$ we consider the forward problems to find $v_k \in U_k$ such that

$$B(v_k, \varphi) = (v_0, \varphi_1), \quad \forall \varphi \in U_k, \quad (4.46)$$

and to find $v_{kh} \in U_{k,h}$ such that

$$B(v_{kh}, \varphi) = (v_0, \varphi_1), \quad \forall \varphi \in U_{k,h}. \quad (4.47)$$

When needed, we will require additional regularity on the initial data v_0.

LEMMA 4.10. *Let w_k, w_{kh} be solutions of* (4.21) *and* (4.45)*, respectively. Then, there holds*

$$\|w_{k,1} - w_{kh,1}\|_{H^{-2}(\Omega)} \leq Ch\|w_T\|_{H^{-1}(\Omega)}. \tag{4.48}$$

Proof. By definition of the norm, we have

$$\|w_{k,1} - w_{kh,1}\|_{H^{-2}(\Omega)} \simeq \sup_{\psi \in \dot{H}^2(\Omega)} \frac{(w_{k,1} - w_{kh,1}, \psi)}{\|\psi\|_{H^2(\Omega)}}, \tag{4.49}$$

therefore we provide an upper bound of the numerator in terms of $\|\psi\|_{H^2(\Omega)}$ and $\|w_T\|_{H^{-1}(\Omega)}$.

To obtain $(w_{k,1} - w_{kh,1}, \psi)$, we pick the test functions in the auxiliary problems so that the backward and forward problems have same left-hand side. Namely, for a fixed $\psi \in \dot{H}^2(\Omega)$, we consider $v_0 = \psi$ in (4.46), and (4.47). Then, we set $\varphi = w_k$ in (4.46), $\varphi = w_{kh}$ in (4.47) and $\varphi = v_k$ in (4.21), $\varphi = v_{kh}$ in (4.45), obtaining

$$\begin{aligned}(\psi, w_{k,1}) &= B(v_k, w_k) = (v_{k,N}, w_T),\\ (\psi, w_{kh,1}) &= B(v_{kh}, w_{kh}) = (v_{kh,N}, w_T).\end{aligned}$$

Using Galerkin orthogonality we have

$$(\psi, w_{k,1} - w_{kh,1}) = B(v_k - v_{kh}, w_k - w_{kh}) = (v_{k,N} - v_{kh,N}, w_T),$$

from which

$$\begin{aligned}(\psi, w_{k,1} - w_{kh,1}) &= -(\nabla(v_{k,N} - v_{kh,N}), \nabla\Delta^{-1} w_T)\\ &\leq \|\nabla(v_{k,N} - v_{kh,N})\|\|\nabla\Delta^{-1} w_T\|\end{aligned}$$

follows.

By standard interpolation and inverse estimates, there holds

$$\begin{aligned}\|\nabla(v_{k,N} - v_{kh,N})\| &\leq \|\nabla(v_{k,N} - \mathcal{I}_h v_{k,N})\| + \|\nabla(\mathcal{I}_h(v_{k,N}) - v_{kh,N}))\|\\ &\leq ch\|\nabla^2 v_{k,N}\| + ch^{-1}(\|\mathcal{I}_h v_{k,N} - v_{k,N}\| + \|v_{k,N} - v_{kh,N}\|)\\ &\leq c(h\|\Delta v_{k,N}\| + h^{-1}\|v_{k,N} - v_{kh,N}\|)\end{aligned}$$

where the last estimates follows, e.g., from [17, Theorem 3.1.3.1] For the first term in the right-hand side, we use [26, Theorem 4.6] and obtain

$$T\|\nabla\Delta v_{k,N}\|^2 + \|\Delta v_{k,N}\|^2 + \|\nabla\Delta v_k\|_I^2 + \sum_{n=2}^{N} \frac{t_{n-1}}{k_n}\|[\Delta v_k]_{n-1}]\|^2 \leq C\|\Delta v_0\|^2. \tag{4.50}$$

For the second term, there holds

$$\|v_{k,N} - v_{kh,N}\| \leq Ch^2\|\Delta v_0\| \tag{4.51}$$

due to [26, Lemma 5.7]. Combining this, we assert

$$\|\nabla(v_{k,N} - v_{kh,N})\| \leq ch(\|\Delta v_{k,N}\| + \|\Delta v_0\|) \leq ch\|\Delta v_0\|.$$

This implies

$$\begin{aligned}(\psi, w_{k,1} - w_{kh,1}) &\le ch\|\Delta v_0\| \|\nabla \Delta^{-1} w_T\| \\ &\le ch\|v_0\|_{H^2(\Omega)} \|\nabla \Delta^{-1} w_T\|,\end{aligned}$$

from which, recalling that $\psi = v_0$, we obtain the assertion

$$\|w_{k,1} - w_{kh,1}\|_{H^{-2}(\Omega)} \le Ch\|w_T\|_{H^{-1}(\Omega)}.$$

□

LEMMA 4.11. *Let w_k, w_{kh} be solutions of* (4.21) *and* (4.45)*, respectively. Then, there holds*

$$\|w_k - w_{kh}\|_{L^2(I,H^{-1}(\Omega))} \le ch\|w_T\|_{H^{-1}(\Omega)}. \tag{4.52}$$

Proof. We introduce the L^2-projection in space $P_h : V \to V_h$ and, noting that $P_h w_{kh} = w_{kh}$, we split the error as

$$\eta_h := w_k - w_{kh} = w_k - P_h w_k + P_h \eta_h. \tag{4.53}$$

For the first part of the error, standard error estimates for the L^2-projection give

$$\|\nabla \Delta^{-1}(w_k - P_h w_k)\|_I \le ch\|w_k\|_I, \tag{4.54}$$

and we obtain the desired estimate by virtue of Lemma 4.9. Hence, we are left with the estimate of $P_h \eta_h$.

Subtracting (4.45) from (4.21) and using (4.43), we have

$$(\nabla \varphi, \nabla \eta_h)_{I_n} - (\varphi_n, [\eta_h]_n) = 0, \ \ \forall \varphi \in \mathcal{P}_0(I_n, V_h), \ n = 1, ..., N-1. \tag{4.55}$$

Thanks to (4.53) and the definition of P_h, this can written as

$$(\nabla \varphi, \nabla P_h \eta_h)_{I_n} - (\varphi_n, [P_h \eta_h]_n) = (\nabla \varphi, \nabla (P_h w_k - w_k))_{I_n}. \tag{4.56}$$

Then, we set $\varphi = \Delta_h^{-2} P_h \eta_h$ to obtain

$$\begin{aligned}\|\nabla \Delta_h^{-1} P_h \eta_h\|_{I_n}^2 &- \big(\Delta_h^{-1} P_h \eta_{h,n}, [\Delta_h^{-1} P_h \eta_h]_n\big) \\ &= \big(\nabla \Delta_h^{-1} P_h \eta_h, \nabla \Delta_h^{-1}(P_h w_k - w_k)\big)_{I_n} \\ &\le \frac{\|\nabla \Delta_h^{-1} P_h \eta_h\|_{I_n}^2}{2} + \frac{\|\nabla \Delta_h^{-1}(P_h w_k - w_k)\|_{I_n}^2}{2}.\end{aligned}$$

For the second term on the left-hand side we use (4.27) and, noting that the jump term is positive, we have

$$\|\nabla \Delta_h^{-1} P_h \eta_h\|_{I_n}^2 + \|\Delta_h^{-1} P_h \eta_{h,n}\|^2 - \|\Delta_h^{-1} P_h \eta_{h,n+1}\|^2 \le \|\nabla \Delta_h^{-1}(P_h w_k - w_k)\|_{I_n}^2$$

Adding these inequalities for $n = 1, ..., N-1$, we obtain

$$\sum_{n=1}^{N-1} \|\nabla \Delta_h^{-1} P_h \eta_h\|_{I_n}^2 + \|\Delta_h^{-1} P_h \eta_{h,1}\|^2 - \|\Delta_h^{-1} P_h \eta_{h,N}\|^2 \le \sum_{n=1}^{N-1} \|\nabla \Delta_h^{-1}(P_h w_k - w_k)\|_{I_n}^2.$$

We observe that $P_h\eta_{h,N} = 0$. As a consequence, the third term in the left-hand is zero and we can extend the sum up to the last time interval. Then, observing that the second term in the left-hand side is positive, thanks to the splitting (4.53) we obtain

$$\|\nabla\Delta_h^{-1}\eta_h\|_I^2 \leq 2\|\nabla\Delta_h^{-1}(P_h w_k - w_k)\|_I^2.$$

Then, using (4.54) and recalling the equivalence between the discrete negative norm $\|\nabla\Delta_h^{-1}\cdot\|$ and the continuous one $\|\nabla\Delta^{-1}\cdot\|$, we get

$$\|\nabla\Delta^{-1}\eta_h\|_I^2 \leq \|\nabla\Delta^{-1}(P_h w_k - w_k)\|_I^2 \leq ch^2\|w_k\|_I^2.$$

The proof is concluded by virtue of Lemma 4.9. ∎

After this preparation, we conclude the section combining the preceding results to obtain the $L^\infty(I,V)$ error estimate in space.

THEOREM 4.12. *Let $u_k \in U_k$ and $u_{kh} \in U_{k,h}$ be solution of* (3.2) *and* (3.6)*, respectively, with $f(x,t) = q(t)g(x) \in L^\infty(I,V)$ and $u_0 \in \dot{H}^2(\Omega)$ Then for the discretization error in space it holds*

$$\|u_{kh} - u_k\|_{L^\infty(I,V)} \leq Ch(\|f\|_{L^2(I,V)} + \|u_0\|_{H^2(\Omega)}). \tag{4.57}$$

Proof. We observe that both u_k, u_{kh} are constant on I_n for any $n = 1,...,N$, hence we can equivalently show the estimate on I_n and with no loss of generality we consider I_N only.

Considering $w_k \in U_k$, $w_{kh} \in U_{k,h}$ solutions of (4.21) and (4.45), respectively, with $w_T = -\Delta_h(u_{k,N} - u_{kh,N})$, by means of the duality argument and Galerkin orthogonality there holds

$$\begin{aligned}\|\nabla(u_{k,N} - u_{kh,N})\|^2 &= B(u_k - u_{kh}, w_k) = B(u_k, w_k - w_{kh}) \\ &= (f, w_k - w_{kh})_I + (u_0, w_{k,1} - w_{kh,1}) \\ &\leq \big(\|w_k - w_{kh}\|_{L^2(I,H^{-1}(\Omega))} + \|w_{k,1} - w_{kh,1}\|_{H^{-2}(\Omega)}\big) \\ &\quad \times \big(\|f\|_{L^2(I,V)} + \|u_0\|_{H^2(\Omega)}\big).\end{aligned}$$

Using Lemma 4.10 and Lemma 4.11, we conclude

$$\|\nabla(u_{k,N} - u_{kh,N})\|^2 \leq ch\|w_T\|_{H^{-1}(\Omega)}\big(\|f\|_{L^2(I,V)} + \|u_0\|_{H^2(\Omega)}\big).$$

The assertion follows observing that by our choice of w_T it holds

$$\|w_T\|_{H^{-1}(\Omega)} = \|\nabla(u_{k,N} - u_{kh,N})\|.$$

∎

5. Error Analysis for the Optimization Problem. This section is concerned with the estimate of the error between the solution $(\bar{q},\bar{u})$ of the continuous optimal control problem (2.3) and the solution $(\bar{q}_{kh},\bar{u}_{kh})$ of the discretized problem (3.7). We recall that by virtue of the variational discretization approach the optimal control problem is already fully discretized with the dG(0)-cG(1) method.

We analyze the error arising from the time and space discretization of the problem separately. The main result of this article, i.e., the error for the space-time discretization of the optimization problem, is shown at the end of the section.

In both cases, we first derive the error estimate emphasizing its dependency on the $L^\infty(I,V)$ error estimates for the state equation derived in Theorem 4.8 and in Theorem 4.12.

We start the analysis with the temporal error.

THEOREM 5.1. *Let* $(\bar q, \bar u)$ *and* $(\bar q_k, \bar u_k)$ *be the optimal solutions of* (2.3) *and* (3.3), *respectively. Then, there holds*

$$\begin{aligned}\|\bar u - \bar u_k\|_I^2 &+ \alpha\|\bar q - \bar q_k\|_{L^2(I)}^2 \\ &\le C\Big(\|\bar u - u_k(\bar q)\|_{L^\infty(I,V)} + \|u(\bar q_k) - \bar u_k\|_{L^\infty(I,V)}\Big).\end{aligned}$$

Proof. We test (2.5c) with $q = \bar q_k$ and (3.4a) with $q = \bar q$, obtaining

$$\begin{aligned}0 &\le \alpha(\bar q, \bar q_k - \bar q) + ((\bar q_k - \bar q)g, \bar z)_I,\\ 0 &\le \alpha(\bar q_k, \bar q - \bar q_k) + ((\bar q - \bar q_k)g, \bar z_k)_I.\end{aligned}$$

Adding the inequalities above, we have

$$\begin{aligned}\alpha\|\bar q - \bar q_k\|^2 &= -\alpha(\bar q - \bar q_k, \bar q_k - \bar q)\\ &\le ((\bar q_k - \bar q)g, \bar z - \bar z_k)_I\\ &= \underbrace{((\bar q_k - \bar q)g, \bar z)_I}_{(a)} + \underbrace{((\bar q - \bar q_k)g, \bar z_k)_I}_{(b)}.\end{aligned}\tag{5.1}$$

We now consider the two terms separately.

(a) We consider (2.5a) with right-hand side given by $q = \bar q_k - \bar q$ and we set $\varphi = \bar z$. Then, in a first step we obtain

$$((\bar q_k - \bar q)g, \bar z)_I = (\partial_t(u(\bar q_k) - \bar u), \bar z)_I + (\nabla(u(\bar q_k) - \bar u), \nabla \bar z)_I.\tag{5.2}$$

With the choice $\varphi = u(\bar q_k) - \bar u$ in the adjoint equation (2.5b), we have from (5.2) that

$$((\bar q_k - \bar q)g, \bar z)_I = (\bar u - u_d, u(\bar q_k) - \bar u)_I + 2\langle\bar\mu, (\nabla\bar u\nabla(u(\bar q_k) - \bar u), \omega)\rangle.\tag{5.3}$$

We consider the second term on the right-hand side. In view of the complementary slackness condition (2.5d), the positivity of $\bar\mu$, the boundedness of $\nabla u(\bar q_k)$ and $\nabla \bar u_k$ in $L^\infty(I,H)$, we have

$$\begin{aligned}2\langle\bar\mu, (\nabla\bar u\nabla(u(\bar q_k) - \bar u), \omega)\rangle &= \langle\bar\mu, \big(|\nabla\bar u|^2 + |\nabla u(\bar q_k)|^2, \omega\big)\rangle - 2\langle\bar\mu, (|\nabla \bar u|^2, \omega)\rangle\\ &= \langle\bar\mu, (|\nabla u(\bar q_k)|^2 - |\nabla\bar u|^2, \omega)\rangle\\ &= \langle\bar\mu, (|\nabla u(\bar q_k)|^2 - |\nabla\bar u_k|^2 + |\nabla\bar u_k|^2 - |\nabla\bar u|^2, \omega)\rangle\\ &\le \langle\bar\mu, (|\nabla u(\bar q_k)|^2 - |\nabla\bar u_k|^2, \omega)\rangle + \langle\bar\mu, b - G(\bar u)\rangle\\ &\le \|\bar\mu\|_{C(\bar I)^*}\|\omega\|_{L^\infty(\Omega)}\||\nabla u(\bar q_k)|^2 - |\nabla\bar u_k|^2\|_{L^\infty(I,H)}\\ &\le c\|(|\nabla u(\bar q_k)| - |\nabla\bar u_k|)(|\nabla u(\bar q_k)| + |\nabla\bar u_k|)\|_{L^\infty(I,H)}\\ &\le c\||\nabla u(\bar q_k)| - |\nabla\bar u_k|\|_{L^\infty(I,H)}\\ &\le c\|\nabla(u(\bar q_k) - \bar u_k)\|_{L^\infty(I,H)}\\ &= c\|u(\bar q_k) - \bar u_k\|_{L^\infty(I,V)}\end{aligned}$$

Therefore, we obtain

$$((\bar q_k - \bar q)g, \bar z)_I \le (\bar u - u_d, u(\bar q_k) - \bar u)_I + c\|u(\bar q_k) - \bar u_k\|_{L^\infty(I,V)}.\tag{5.4}$$

(b) We proceed along the same lines of the previous case using the semi-discrete state and adjoint equation. We consider (3.4a) with right-hand side $q = \bar{q} - \bar{q}_k$ and we set $\varphi = \bar{z}_k$. Then, through the choice $\varphi = u_k(\bar{q}) - \bar{u}_k$ in (3.4b), we have

$$\begin{aligned}((\bar{q} - \bar{q}_k)g, \bar{z}_k)_I &= B(u_k(\bar{q}) - \bar{u}_k, \bar{z}_k) \\ &= (\bar{u}_k - u_d, u_k(\bar{q}) - \bar{u}_k)_I + 2\langle \bar{\mu}_k, (\nabla \bar{u}_k \nabla (u_k(\bar{q}) - \bar{u}_k), \omega)\rangle.\end{aligned}$$

Estimating the second term in the right-hand side as in case (a), using the uniform boundedness of $\|\bar{\mu}_k\|_{C(\bar{I})^*}$, we have

$$((\bar{q} - \bar{q}_k)g, \bar{z}_k)_I \leq (\bar{u}_k - u_d, u_k(\bar{q}) - \bar{u}_k)_I + c\|u_k(\bar{q}) - \bar{u}\|_{L^\infty(I,V)}. \tag{5.5}$$

We go back to (5.1) inserting (5.4),(5.5) to obtain

$$\begin{aligned}\alpha\|\bar{q} - \bar{q}_k\|^2_{L^2(I)} \leq{}& (\bar{u} - u_d, u(\bar{q}_k) - \bar{u})_I + (\bar{u}_k - u_d, u_k(\bar{q}) - \bar{u}_k)_I \\ &+ c(\|u(\bar{q}_k) - \bar{u}_k\|_{L^\infty(I,V)}\|u_k(\bar{q}) - \bar{u}\|_{L^\infty(I,V)}).\end{aligned} \tag{5.6}$$

Now, we note that

$$\|\bar{u} - \bar{u}_k\|_I^2 = (\bar{u} - u^d, \bar{u} - \bar{u}_k) - (\bar{u}_k - u^d, \bar{u} - \bar{u}_k).$$

Adding this to (5.6), we obtain

$$\begin{aligned}\|\bar{u} - \bar{u}_k\|_I^2 + \alpha\|\bar{q} - \bar{q}_k\|^2 &\leq (\bar{u} - u_d, u(\bar{q}_k) - \bar{u}_k)_I + (\bar{u}_k - u_d, u_k(\bar{q}) - \bar{u})_I \\ &\quad + c(\|u(\bar{q}_k) - \bar{u}_k\|_{L^\infty(I,V)}\|u_k(\bar{q}) - \bar{u}\|_{L^\infty(I,V)}) \\ &\leq c\Big(\|u(\bar{q}_k) - \bar{u}_k\|_I + \|u_k(\bar{q}) - \bar{u}\|_I \\ &\qquad + \|u(\bar{q}_k) - \bar{u}_k\|_{L^\infty(I,V)}\|u_k(\bar{q}) - \bar{u}\|_{L^\infty(I,V)}\Big).\end{aligned}$$

The assertion follows, since the $\|\cdot\|_I$-norm can be bounded by the $\|\cdot\|_{L^\infty(I,V)}$-norm.
□

COROLLARY 5.2. *Under the assumptions of Theorem* 4.8 *and Theorem* 5.1*, the following estimate holds*

$$\|\bar{u} - \bar{u}_k\|_I^2 + \alpha\|\bar{q} - \bar{q}_k\|^2_{L^2(I)} \leq ck\Big(\log\frac{T}{k} + 1\Big)^{\frac{1}{2}}. \tag{5.7}$$

Proof. The claim directly follows from the previous Theorem inserting the $L^\infty(I,V)$ estimate of Theorem 4.8 together with the assumed regularity of f and u_0. □

In a second step, we consider the error arising from the space discretization.

THEOREM 5.3. *Let* $(\bar{q}_k, \bar{u}_k)$ *and* $(\bar{q}_{kh}, \bar{u}_{kh})$ *be the optimal solutions of* (3.3) *and* (3.7)*, respectively. Then, there holds*

$$\begin{aligned}&\|\bar{u}_k - \bar{u}_{kh}\|_I^2 + \alpha\|\bar{q}_k - \bar{q}_{kh}\|^2_{L^2(I)} \\ &\qquad \leq C\Big(\|\bar{u}_k - u_{kh}(\bar{q}_k)\|_{L^\infty(I,V)} + \|u_k(\bar{q}_{kh}) - \bar{u}_{kh}\|_{L^\infty(I,V)}\Big).\end{aligned}$$

Proof. The proof moves along the same lines of the error estimate for the semi-discrete case in Theorem 5.1.

In particular, testing (3.4c) with $q = \bar{q}_{kh}$ and (3.8c) with $q = \bar{q}_k$, we can add the resulting inequalities to get

$$\alpha\|\bar{q}_k - \bar{q}_{kh}\|_{L^2(I)} \leq (\bar{z}_k - \bar{z}_{kh}, (\bar{q}_{kh} - \bar{q}_k)g)_I \\ = \underbrace{(\bar{z}_k, (\bar{q}_{kh} - \bar{q}_k)g)_I}_{(a)} + \underbrace{(\bar{z}_{kh}, (\bar{q}_k - \bar{q}_{kh})g)_I}_{(b)}. \tag{5.8}$$

We now consider the two terms separately.

(a) As in the semi-discrete case, the idea is to express (a) in term of the semi-discrete state equation and then in term of the semi-discrete adjoint equation equalizing the common term $B(\cdot,\cdot)$.
First, we consider (3.4a) with right-hand side $q_k = \bar{q}_{kh} - \bar{q}_k$ and we set $\varphi = \bar{z}_k$. Then, with the choice $\varphi = u_k(\bar{q}_{kh}) - \bar{u}_k$ in (3.4b), we have

$$((\bar{q}_{kh} - \bar{q}_k)g, \bar{z}_k)_I = (\bar{u}_k - u_d, u_k(\bar{q}_{kh}) - \bar{u}_k)_I + \langle\bar{\mu}_k, 2(\nabla\bar{u}_k\nabla(u_k(\bar{q}_{kh}) - \bar{u}_k), \omega)\rangle.$$

The proof now proceeds exactly as in Theorem 4.1, yielding

$$(\bar{z}_k, (\bar{q}_{kh} - \bar{q}_k))_I \leq (\bar{u}_k - u_d, u_k(\bar{q}_{kh}) - \bar{u}_k)_I \\ + c\|\bar{\mu}_k\|_{C(\bar{I})^*}\|\omega\|_{L^\infty(\Omega)}\|u_k(\bar{q}_{kh}) - \bar{u}_{kh}\|_{L^\infty(I,V)}. \tag{5.9}$$

(b) We now use the discrete state equation (3.8a) with $\varphi = \bar{z}_{kh}$ and right-hand side given by $q_{kh} = \bar{q}_k - \bar{q}_{kh}$, together with the discrete adjoint equation (3.8b) with $\varphi = u_{kh}(\bar{q}_k) - \bar{u}_{kh}$. This setting leads to

$$(\bar{z}_{kh}, (\bar{q}_k - \bar{q}_{kh})g)_I \leq (\bar{u}_{kh} - u_d, u_{kh}(\bar{q}_k) - \bar{u}_{kh})_I \\ + c\|\bar{\mu}_{kh}\|_{C(\bar{I})^*}\|\omega\|_{L^\infty(\Omega)}\|\bar{u}_k - u_{kh}(\bar{q}_k)\|_{L^\infty(I,V)}. \tag{5.10}$$

We insert (5.9),(5.10) in (5.8) and, as in Theorem 5.1, we manipulate the resulting inequality by adding and subtracting $\bar{u}_k$, $\bar{u}_{kh}$. Then, uniform boundedness of $\bar{\mu}_k$ and $\bar{\mu}_{kh}$, see Remark 3.7 concludes the proof. □

COROLLARY 5.4. *Under the assumptions of Theorem* 4.12 *and Theorem* 5.3*, the following estimate holds*

$$\|\bar{u}_k - \bar{u}_{kh}\|_I^2 + \alpha\|\bar{q}_k - \bar{q}_{kh}\|_{L^2(I)}^2 \leq ch. \tag{5.11}$$

Proof. The thesis follows from the previous Theorem, inserting the error estimate from Theorem 4.12 together with the assumed regularity of u_0 and f. □

We conclude the error analysis for the optimal control problem with the main result of this article. The following Theorem is obtained combining the error estimate for the time discretization with the error estimate for the space discretization derived above.

THEOREM 5.5. *Let* $(\bar{u}, \bar{q}) \in U \times Q_{ad}$ *be the optimal solution of the continuous problem* (2.3) *and* $(\bar{u}_{kh}, \bar{q}_{kh}) \in U_{k,h} \times Q_{ad}$ *be the optimal solution of the discrete problem* (3.7)*. Then, there holds the following error estimate*

$$\|\bar{u} - \bar{u}_{kh}\|_I^2 + \alpha\|\bar{q} - \bar{q}_{kh}\|_{L^2(I)}^2 \leq C\left(k\left(\log\frac{T}{k} + 1\right)^{\frac{1}{2}} + h\right). \tag{5.12}$$

Acknowledgment. The authors would like to acknowledge some helpful discussions with Boris Vexler during the preparation of the manuscript.

REFERENCES

[1] N. Arada and J.-P. Raymond, *Dirichlet boundary control of semilinear parabolic equations. II. Problems with pointwise state constraints*, Appl. Math. Optim., 45 (2002), pp. 145–167.

[2] 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.

[3] 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.

[4] E. Casas, *Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations*, SIAM J. Control Optim., 35 (1997), pp. 1297–1327.

[5] 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.

[6] E. Casas and L. A. Fernández, *Corrigendum: Optimal control of semilinear elliptic equations with pointwise constraints on the gradient of the state*, Appl. Math. Optim., 28 (1993), pp. 337–339.

[7] ———, *Optimal control of semilinear elliptic equations with pointwise constraints on the gradient of the state*, Appl. Math. Optim., 27 (1993), pp. 35–56.

[8] E. Casas, M. Mateos, and J.-P. Raymond, *Pontryagin's principle for the control of parabolic equations with gradient state constraints*, Nonlinear Anal., 46 (2001), pp. 933–956.

[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] D. Clever and J. Lang, *Optimal control of radiative heat transfer in glass cooling with restrictions on the temperature gradient*, Optimal Control Appl. Meth., 33 (2012), pp. 157–175.

[11] K. Deckelnick, A. Günther, and M. Hinze, *Finite element approximation of elliptic control problems with constraints on the gradient*, Numer. Math., 111 (2009), pp. 335–350.

[12] 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.

[13] K. Eriksson, C. Johnson, and V. Thomée, *Time discretization of parabolic problems by the discontinuous Galerkin method*, RAIRO Modél. Math. Anal. Numér., 19 (1985), pp. 611–643.

[14] L. C. Evans, *Partial Differential Equations*, vol. 19 of Grad. Stud. Math., AMS, Providence, second ed., 2010.

[15] 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.

[16] 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.

[17] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Monographs and studies in Mathematics, Pitman, Boston, 1. ed., 1985.

[18] A. Günther and M. Hinze, *Elliptic control problems with gradient constraints—variational discrete versus piecewise constant controls*, Comput. Optim. Appl., 49 (2011), pp. 549–566.

[19] M. Hintermüller and K. Kunisch, *PDE-constrained optimization subject to pointwise constraints on the control, the state, and its derivative*, SIAM J. Optim., 20 (2009), pp. 1133–1156.

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

[21] 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.

[22] J. L. Lions and E. Magenes, *Non-Homogeneous Boundary Value Problems and Applications, Vol. 1*, Grundlehren der mathematischen Wissenschaften, 1972.

[23] ———, *Non-Homogeneous Boundary Value Problems and Applications, Vol. 2*, Grundlehren der mathematischen Wissenschaften, 1972.

[24] M. Luskin and R. Rannacher, *On the smoothing property of the Galerkin method for parabolic equations*, SIAM J. Numer. Anal., 19 (1982), pp. 93–113.

[25] U. Mackenroth, *On parabolic distributed optimal control problems with restrictions on the*

gradient, Appl. Math. Optim., 10 (1983), pp. 69–95.

[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] ———, *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] I. Neitzel and F. Tröltzsch, *On regularization methods for the numerical solution of parabolic control problems with pointwise state constraints*, ESAIM Control Optim. Calc. Var., 15 (2009), pp. 426–453.

[30] C. Ortner and W. Wollner, *A priori error estimates for optimal control problems with pointwise constraints on the gradient of the state*, Numer. Math., 118 (2011), pp. 587–600.

[31] R. Pinnau, *Analysis of optimal boundary control for radiative heat transfer modeled by the* SP_1*-system*, Commun. Math. Sci., 5 (2007), pp. 951–969.

[32] R. Pinnau and G. Thömmes, *Optimal boundary control of glass cooling processes*, Math. Methods Appl. Sci., 27 (2004), pp. 1261–1281.

[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] J. P. Raymond and H. Zidani, *Pontryagin's principle for state-constrained control problems governed by parabolic equations with unbounded controls*, SIAM J. Control Optim., 36 (1998), pp. 1853–1879.

[35] A. Schiela and W. Wollner, *Barrier methods for optimal control problems with convex nonlinear gradient state constraints*, SIAM J. Optim., 21 (2011), pp. 269–286.

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

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

[38] A. Unger and F. Tröltzsch, *Fast solution of optimal control problems in the selective cooling of steel*, ZAMM Z. Angew. Math. Mech., 81 (2001), pp. 447–456.

[39] W. Wollner, *Optimal control of elliptic equations with pointwise constraints on the gradient of the state in nonsmooth polygonal domains*, SIAM J. Control Optim., 50 (2012), pp. 2117–2129.

[40] W. Wollner, *A priori error estimates for optimal control problems with constraints on the gradient of the state on nonsmooth polygonal domains*, in Control and Optimization with PDE Constraints, vol. 164 of Internat. Ser. Numer. Math., Springer Basel, 2013, pp. 193–215.