SIAM J. CONTROL OPTIM.
Vol. 54, No. 5, pp. 2403–2435

A PRIORI ERROR ESTIMATES FOR THREE DIMENSIONAL PARABOLIC OPTIMAL CONTROL PROBLEMS WITH POINTWISE CONTROL*

DMITRIY LEYKEKHMAN† AND BORIS VEXLER‡

Abstract. In this paper we provide an a priori error analysis for parabolic optimal control problems with a pointwise (Dirac-type) control in space on three-dimensional domains. The two-dimensional case was treated in [D. Leykekhman and B. Vexler, *SIAM J. Numer. Anal.*, 51 (2013), pp. 2797–2821]; however, the three-dimensional case is technically much more involved. To approximate the problem we use standard continuous piecewise linear elements in space and the piecewise constant discontinuous Galerkin method in time. Despite low regularity of the state equation, we establish $\mathcal{O}(\sqrt{k}+h)$ order of convergence rate for the control in the L^2 norm. This result improves almost twice the previously known estimate in [W. Gong, M. Hinze, and Z. Zhou, *SIAM J. Control Optim.*, 52 (2014), pp. 97–119] and in addition does not require any relationship between the time step k and the mesh size h. The main technical tools are discrete maximal parabolic regularity results and the best approximation-type estimate for the finite element error in the $L^\infty(\Omega; L^2(I))$ norm.

Key words. optimal control, pointwise control, parabolic problems, finite elements, discontinuous Galerkin, error estimates, pointwise error estimates

AMS subject classifications. 49M25, 49K20, 65N30, 65N15

DOI. 10.1137/15M1028042

1. Introduction.

In this paper we provide numerical analysis for the following optimal control problem:

$$\min_{q,u} J(q,u) := \frac{1}{2}\int_0^T \|u(t)-\widehat{u}(t)\|^2_{L^2(\Omega)}dt + \frac{\alpha}{2}\int_0^T |q(t)|^2 dt \tag{1.1}$$

subject to the second order parabolic equation

$$u_t(t,x) - \Delta u(t,x) = q(t)\delta_{x_0}(x), \quad (t,x) \in I\times\Omega, \tag{1.2a}$$

$$u(t,x) = 0, \quad (t,x)\in I\times\partial\Omega, \tag{1.2b}$$

$$u(0,x) = 0, \quad x\in\Omega, \tag{1.2c}$$

and the pointwise control constraints

$$q_a \le q(t) \le q_b \quad \text{a.e. in } I. \tag{1.3}$$

Here $I=(0,T)$ is the time interval, $\Omega\subset\mathbb{R}^3$ is a convex polyhedral domain, $x_0\in\Omega$ fixed, and δ_{x_0} is the Dirac delta function. The parameter α is assumed to be positive and the desired state $\widehat{u}$ fulfills $\widehat{u}\in L^\infty(I;L^3(\Omega))$. The control bounds $q_a, q_b \in \mathbb{R}\cup\{\pm\infty\}$ fulfill $q_a<q_b$. The precise functional-analytic setting is discussed in the next section. This setup is a model problem for pointwise control, where, for simplicity,

*Received by the editors June 26, 2015; accepted for publication (in revised form) June 20, 2016; published electronically September 21, 2016.
http://www.siam.org/journals/sicon/54-5/M102804.html
†Department of Mathematics, University of Connecticut, Storrs, CT 06269 (leykekhman@math.uconn.edu).
‡Chair of Optimal Control, Faculty of Mathematics, Technical University of Munich, Boltzmannstraße 3, 85748 Garching b. München, Germany (vexler@ma.tum.de).

we choose the heat equation as the state equation and consider the case of only one point source. However, all our results extend directly to more general self-adjoint elliptic operators with smooth coefficients (instead of $-\Delta$) and $l \geq 1$ point sources $\sum_{i=1}^{l} q_i(t)\delta_{x_i}$ on the right-hand side of (1.2).

There are several publications dealing with optimal control problems of this type, starting with [34]. The main difficulty in the pointwise control problem is the low regularity of the state variable. We refer to [3, 12, 16] for pointwise control of parabolic equations and to [14, 39] for pointwise control of Burgers-type equations. Moreover, pointwise control problems are strongly related to measure valued formulations of parabolic sparse control problems; see [8, 9, 10, 11, 28].

For the discretization of the problem under consideration, we consider standard continuous piecewise linear finite elements in space and the piecewise constant discontinuous Galerkin method in time. This is a special case ($r = 0$, $s = 1$) of the so-called cG(s)dG(r) discretization; see, e.g., [19] for analysis of the method for parabolic problems and, e.g., [37, 38] for error estimates in the context of optimal control problems. The lowest order discontinuous Galerkin method dG(0) is a variation of the backward Euler method, while higher order discontinuous Galerkin methods coincide with the subdiagonal Padé method for homogeneous problems. Throughout, we will denote by h the spatial mesh size and by k the time step; see section 3 for details.

Although numerical analysis for elliptic problems with rough right-hand sides was considered in a number of papers [4, 6, 7, 18, 26, 29, 43], there are few papers that consider parabolic problems with rough sources. We are only aware of the paper [23], where $L^2(I; L^2(\Omega))$ error estimates are considered. Based on the results of this paper, suboptimal error estimates of order $\mathcal{O}(k^{\frac{1}{2}} + h)$ for two-dimensional and of order $\mathcal{O}(k^{\frac{1}{2}} + h^{\frac{1}{2}})$ for three-dimensional control problems with pointwise controls were derived in [24]. In both results a restrictive assumption $k = \mathcal{O}(h^d)$, with d being the dimension d of the domain Ω, is used. This assumption is not natural for the method, especially in three space dimensions, since the dG(0) time discretization is a variation of the backward Euler method and unconditionally stable. For the two-dimensional problem, we improved this estimate in [30] to almost $\mathcal{O}(k + h^2)$ up to a logarithmic term and avoided any coupling between the discretization parameters k and h. The proof in [30] was based on a novel best-approximation-type result with respect to the $L^\infty(\Omega; L^2(I))$ norm and was restricted to the two-dimensional case.

In this paper we provide a corresponding best-approximation-type result with respect to the $L^\infty(\Omega; L^2(I))$ norm for a convex polyhedral domain $\Omega \subset \mathbb{R}^3$. To avoid any coupling between k and h we use recently established discrete maximal parabolic regularity results from [31]. Moreover, we exploit different regularity of the state and the adjoint equations. It turns out that the error analysis for the adjoint equation with respect to the $L^\infty(\Omega; L^2(I))$ norm and for the state variable with respect to the $L^2(I; L^{\frac{3}{2}}(\Omega))$ norm leads to the best possible results for the optimal control problem under consideration. This strategy allows us to prove the error estimate for the optimal control $\bar{q}$ and its discrete counterpart $\bar{q}_{kh}$ of order $\mathcal{O}(k^{\frac{1}{2}} + h)$ up to a logarithmic term and without any coupling condition between k and h; see Theorem 4.3 below. This improves the three-dimensional result from [24] almost twice. The main technical tools in our analysis are global and interior pointwise error estimates with respect to the $L^\infty(\Omega; L^2(I))$ norm (see Theorems 3.3 and 3.4, respectively), that extend our findings from [30] from the two-dimensional to the three-dimensional case. To our knowledge, these are the first pointwise in space fully discrete best approximation results for three-dimensional parabolic problems; cf. also a recent paper [33]. We refer to [20, 40] for two-dimensional pointwise error estimates.

Throughout the paper we use the usual notation for the Lebesgue and Sobolev spaces. We denote by $(\cdot,\cdot)_\Omega$ the inner product in $L^2(\Omega)$ and by $(\cdot,\cdot)_{J\times\Omega}$ with some subinterval $J \subset I$ the inner product in $L^2(J; L^2(\Omega))$.

The rest of the paper is organized as follows. In section 2 we discuss the functional analytic setting of the optimal control problem, state the optimality conditions, and provide precise regularity results for the optimal control, the optimal state, and the adjoint state. In section 3 we discuss the discretization scheme and state the global and interior best-approximation-type results with respect to the $L^\infty(\Omega; L^2(I))$ norm, which are proven in sections 6 and 7, respectively. Section 4 is devoted to the proof of our main result for the optimal control problem. In section 5 we provide some auxiliary elliptic results required for the proof of Theorem 3.3.

2. Continuous problem and regularity. In order to state the functional analytic setting for the optimal control problem, we first introduce an auxiliary problem

$$\begin{aligned} v_t(t,x) - \Delta v(t,x) &= f(t,x), && (t,x) \in I \times \Omega, \\ v(t,x) &= 0, && (t,x) \in I \times \partial\Omega, \\ v(0,x) &= 0, && x \in \Omega, \end{aligned} \tag{2.1}$$

with a right-hand side $f \in L^r(I; L^p(\Omega))$ for some $1 < p, r < \infty$. In what follows we will use the following maximal regularity result for the solution of this equation.

LEMMA 2.1. *Let $f \in L^r(I; L^p(\Omega))$ with $1 < p, r < \infty$. Then (2.1) possesses a unique solution v with $v_t, \Delta v \in L^r(I; L^p(\Omega))$ and there exist constants $c_r \le \frac{c\,r^2}{r-1}$ and $c_p > 0$ independent of f and v such that*

$$\|v_t\|_{L^r(I;L^p(\Omega))} + \|\Delta v\|_{L^r(I;L^p(\Omega))} \le c_r\, c_p\, \|f\|_{L^r(I;L^p(\Omega))}.$$

For $1 < p \le 2$ the solution v is in $W^{1,r}(I; L^p(\Omega)) \cap L^r(I; W^{2,p}(\Omega))$ and the estimate

$$\|v_t\|_{L^r(I;L^p(\Omega))} + \|\nabla^2 v\|_{L^r(I;L^p(\Omega))} \le c_r\, c_p\, \|f\|_{L^r(I;L^p(\Omega))}$$

holds.

Proof. The first result is a direct application of maximal parabolic regularity; see, e.g., [17]. The dependence of the constant c_r on r can be found, e.g., in [5, Theorem 3.2, p. 28]. The second result follows then by elliptic $W^{2,p}(\Omega)$ regularity, which holds for all $1 < p \le 2$ for convex domains Ω; see [21, Corollary 1]. ∎

In what follows sequel we will need an exact form of the constant in the following embedding.

LEMMA 2.2. *For $p > 3$ the space $W^{1,p}(\Omega)$ is continuously embedded into $C(\bar\Omega)$ and there holds*

$$\|v\|_{C(\bar\Omega)} \le c \left(\frac{p-1}{p-3}\right)^{1-\frac{1}{p}} \|v\|_{W^{1,p}(\Omega)}$$

for all $v \in W^{1,p}(\Omega)$, where the constant c is independent of v and p.

Proof. The embedding is well known. The exact form of the constant can be traced, for example, from the proof of [1, Theorem 8.1]; cf. also [13]. ∎

Remark 2.3. For our purpose it will be enough to use the following estimate

$$\|v\|_{C(\bar\Omega)} \le \frac{c}{\frac{3}{2} - p'} \|v\|_{W^{1,p}(\Omega)} \tag{2.2}$$

for all $3 < p$ and $\frac{1}{p} + \frac{1}{p'} = 1$.

Moreover we will use that $W^{t,s}(\Omega) \hookrightarrow C(\bar{\Omega})$ if $t - \frac{3}{s} > 0$. For $t \geq 1$ we can use the embedding $W^{t,s}(\Omega) \hookrightarrow W^{1,p}(\Omega)$ with

$$t - \frac{3}{s} = 1 - \frac{3}{p}$$

and therefore the estimate

$$\|v\|_{C(\bar{\Omega})} \leq \frac{c}{t - \frac{3}{s}} \|v\|_{W^{t,s}(\Omega)}. \tag{2.3}$$

We will also need the following interior regularity result. Here, and in what follows, we will denote an open ball of radius d centered at x_0 by $B_d = B_d(x_0)$.

LEMMA 2.4. *Let $\overline{B}_{2d} \subset \Omega$. If $f \in L^r(I; L^2(\Omega)) \cap L^r(I; L^p(B_{2d}))$ for some $1 < p \leq 6$ and $1 < r < \infty$, then the solution of (2.1) fulfills $v \in L^r(I; W^{2,p}(B_d)) \cap W^{1,r}(I; L^p(B_d))$ and there exist constants $c_r \leq \frac{c\, r^2}{r-1}$ and $c_p > 0$ independent of d such that*

$$\begin{aligned} &\|v_t\|_{L^r(I;L^p(B_d))} + \|v\|_{L^r(I;W^{2,p}(B_d))} \\ &\qquad \leq c_r^2\, c_p \left(\|f\|_{L^r(I;L^p(B_{2d}))} + d^{-3\left(\frac{1}{2}-\frac{1}{p}\right)} \|f\|_{L^r(I;L^2(\Omega))} \right). \end{aligned}$$

Proof. Let $v \in L^r(I; H^2(\Omega)) \cap W^{1,r}(I; L^2(\Omega))$ be the solution of (2.1). We define

$$\bar{v}(t) = \frac{1}{|B_{2d}|} \int_{B_{2d}} v(t,x) dx.$$

By the Cauchy–Schwarz inequality we have

$$|\bar{v}_t(t)| \leq \frac{1}{|B_{2d}|} |B_{2d}|^{\frac{1}{2}} \|v_t(t)\|_{L^2(B_{2d})} = |B_{2d}|^{-\frac{1}{2}} \|v_t(t)\|_{L^2(B_{2d})} \leq c\, d^{-\frac{3}{2}} \|v_t(t)\|_{L^2(B_{2d})}. \tag{2.4}$$

We consider a smooth cutoff function $\omega \in [0,1]$ with the following properties,

$$\omega(x) \equiv 1, \quad x \in B_d(x_0), \tag{2.5a}$$

$$\omega(x) \equiv 0, \quad x \in \Omega \setminus B_{2d}(x_0), \tag{2.5b}$$

$$|\nabla \omega| \leq c d^{-1}, \quad |\nabla^2 \omega| \leq c d^{-2}, \tag{2.5c}$$

and set $\tilde{v} = (v - \bar{v})\omega$. There holds

$$\Delta \tilde{v} = \omega \Delta v + \nabla v \cdot \nabla \omega + (v - \bar{v}) \Delta \omega$$

and therefore $\tilde{v}$ satisfies the following equation:

$$\tilde{v}_t - \Delta \tilde{v} = g, \quad \tilde{v}(0, \cdot) = 0,$$

on B_{2d} with homogeneous Dirichlet boundary conditions, where

$$\begin{aligned} g &= (v_t - \Delta v)\omega - \nabla v \cdot \nabla \omega - (v - \bar{v})\Delta \omega - \bar{v}_t \omega \\ &= f\omega - \nabla v \cdot \nabla \omega - (v - \bar{v})\Delta \omega - \bar{v}_t \omega. \end{aligned}$$

We have

$$\begin{aligned} \|g\|_{L^r(I;L^p(B_{2d}))} \leq c \Big(&\|f\|_{L^r(I;L^p(B_{2d}))} + d^{-1} \|\nabla v\|_{L^r(I;L^p(B_{2d}))} \\ &+ d^{-2} \|v - \bar{v}\|_{L^r(I;L^p(B_{2d}))} + \|\bar{v}_t\|_{L^r(I;L^p(B_{2d}))} \Big). \end{aligned}$$

Using the Hölder inequality, the Sobolev embedding $H^2(\Omega) \hookrightarrow W^{1,6}(\Omega)$, and Lemma 2.1, we have

$$\begin{aligned}\|\nabla v\|_{L^r(I;L^p(B_{2d}))} &\le c\,d^{\frac{6-p}{2p}}\|\nabla v\|_{L^r(I;L^6(B_{2d}))} \le c\,d^{\frac{6-p}{2p}}\|v\|_{L^r(I;H^2(\Omega))}\\ &\le c_r\,d^{\frac{6-p}{2p}}\|f\|_{L^r(I;L^2(\Omega))}.\end{aligned}$$

Similarly, using additionally the Poincaré inequality, we obtain

$$\|v-\bar v\|_{L^r(I;L^p(B_{2d}))} \le c\,d\|\nabla v\|_{L^r(I;L^p(B_{2d}))} \le c_r\,d^{1+\frac{6-p}{2p}}\|f\|_{L^r(I;L^2(\Omega))}.$$

Further, by (2.4) we have

$$\|\bar v_t\|_{L^r(I;L^p(B_{2d}))} \le c\,d^{\frac{3}{p}-\frac{3}{2}}\|v_t\|_{L^r(I;L^2(B_{2d}))} \le c_r\,d^{-3\left(\frac12-\frac1p\right)}\|f\|_{L^r(I;L^2(\Omega))}. \tag{2.6}$$

Using maximal parabolic regularity for $\tilde v$ (see Lemma 2.1), we obtain

$$\begin{aligned}&\|\tilde v_t\|_{L^r(I;L^p(B_{2d}))} + \|\Delta\tilde v\|_{L^r(I;L^p(B_{2d}))}\\ &\quad\le c_r\,c_p\,\|g\|_{L^r(I;L^p(B_{2d}))}\\ &\quad\le c_r^2\,c_p\left(d^{-3\left(\frac12-\frac1p\right)}\|f\|_{L^r(I;L^2(\Omega))} + \|f\|_{L^r(I;L^p(B_{2d}))}\right),\end{aligned}$$

and due to the fact that B_{2d} has a smooth boundary we also have

$$\|\tilde v\|_{L^r(I;W^{2,p}(B_{2d}))} \le c_p\|\Delta\tilde v\|_{L^r(I;L^p(B_{2d}))}$$

for any $1 < p \le 6$, where the exact form of the constant c_p for the elliptic $W^{2,p}$ regularity estimate can be traced, for example, from [22, Theorem 9.9]. Observing that $\nabla^2 v = \nabla^2\tilde v$ on B_d we obtain the desired estimate for $\|v\|_{L^r(I;W^{2,p}(B_d))}$. The estimate for $\|v_t\|_{L^r(I;L^p(B_d))}$ follows from the fact that $v_t = \tilde v_t + \bar v_t$ on B_d, estimate (2.6), and by the triangle inequality. This completes the proof. ∎

In the following we will utilize a corresponding interior elliptic result. The proof follows the lines of the proof of Lemma 2.4.

LEMMA 2.5. *Let $\overline{B}_{2d} \subset \Omega$ and $f \in L^2(\Omega)\cap L^p(B_{2d})$ for some $1 < p \le 6$. Let w be the solution of the Poisson equation*

$$\begin{aligned}-\Delta w &= f \quad \text{in}\,\Omega,\\ w &= 0 \quad \text{on}\,\partial\Omega.\end{aligned}$$

Then there exists a constant $c_p > 0$ such that

$$\|w\|_{W^{2,p}(B_d)} \le c_p\left(\|f\|_{L^p(B_{2d})} + d^{-3\left(\frac12-\frac1p\right)}\|f\|_{L^2(\Omega)}\right).$$

To introduce a weak solution of the state equation (1.2) we use the method of transposition; cf. [35]. For a given control $q \in Q = L^2(I)$, we denote by $u = u(q) \in L^2(I;L^2(\Omega))$ a weak solution of (1.2), if for all $\varphi \in L^2(I;L^2(\Omega))$ there holds

$$(u,\varphi)_{I\times\Omega} = \int_I w(t,x_0)q(t)\,dt,$$

where $w \in L^2(I;H^2(\Omega)\cap H^1_0(\Omega))\cap H^1(I;L^2(\Omega))$ is the weak solution of the adjoint equation

$$\begin{aligned}-w_t(t,x) - \Delta w(t,x) &= \varphi(t,x), && (t,x)\in I\times\Omega,\\ w(t,x) &= 0, && (t,x)\in I\times\partial\Omega,\\ w(T,x) &= 0, && x\in\Omega.\end{aligned} \tag{2.7}$$

The existence of this weak solution $u = u(q)$ follows by the Riesz representation theorem using the embedding $L^2(I; H^2(\Omega)) \hookrightarrow L^2(I; C(\bar{\Omega}))$.

The following regularity result for the state equation is based again on the maximal parabolic regularity.

THEOREM 2.6. *Let $q \in L^r(I)$ for some $1 < r < \infty$ be given and $u = u(q)$ be the solution of the state equation (1.2). Then for all $1 < s < \frac{3}{2}$ and $1 < p < 3$*

$$u \in L^r(I; W_0^{1,s}(\Omega)), \quad u \in L^r(I; L^p(\Omega)), \quad \textit{and} \quad u_t \in L^r(I; W^{-1,s}(\Omega))$$

with the corresponding estimates

$$\|u\|_{L^r(I;W_0^{1,s}(\Omega))} + \|u_t\|_{L^r(I;W^{-1,s}(\Omega))} \le \frac{c_r}{\frac{3}{2} - s}\|q\|_{L^r(I)} \tag{2.8}$$

and

$$\|u\|_{L^r(I;L^p(\Omega))} \le \frac{c_r}{3 - p}\|q\|_{L^r(I)}, \tag{2.9}$$

where the constant $c_r \le \frac{cr^2}{r-1}$ is independent of s and p. Moreover, the state u fulfills the following weak formulation,

$$\langle u_t, \varphi\rangle + (\nabla u, \nabla \varphi) = \int_I \varphi(t, x_0) q(t)\, dt \quad \textit{for all} \ \ \varphi \in L^{r'}(I; W_0^{1,s'}(\Omega)),$$

where $\frac{1}{s'} + \frac{1}{s} = 1$, $\frac{1}{r'} + \frac{1}{r} = 1$, and $\langle \cdot, \cdot\rangle$ is the duality product between $L^r(I; W^{-1,s}(\Omega))$ and $L^{r'}(I; W_0^{1,s'}(\Omega))$.

Proof. For $s < \frac{3}{2}$ we have $s' > 3$ and by (2.2) $W_0^{1,s'}(\Omega)$ is embedded into $C(\bar{\Omega})$ with

$$\|w\|_{C(\bar{\Omega})} \le \frac{c}{\frac{3}{2} - s}\|w\|_{W_0^{1,s'}(\Omega)} \quad \text{for all} \ \ w \in W_0^{1,s'}(\Omega).$$

Therefore, for $q \in L^r(I)$ the right-hand side $q(t)\delta_{x_0}$ of the state equation can be identified with an element in $L^r(I; W^{-1,s}(\Omega))$ with

$$\|q\delta_{x_0}\|_{L^r(I;W^{-1,s}(\Omega))} \le \|q\|_{L^r(I)}\|\delta_{x_0}\|_{W^{-1,s}(\Omega)} \le \frac{c}{\frac{3}{2} - s}\|q\|_{L^r(I)}.$$

Using the result from [17, Theorem 5.1] on maximal parabolic regularity we obtain

$$u \in L^r(I; W_0^{1,s}(\Omega)) \quad \text{and} \quad u_t \in L^r(I; W^{-1,s}(\Omega))$$

and the estimate (2.8) holds. The dependence of the constant c_r on r follows again from [5, Theorem 3.2, p. 28]. The estimate (2.9) holds then by the embedding $W^{1,s}(\Omega) \hookrightarrow L^p(\Omega)$ with $p = \frac{3s}{3-s}$. Given this regularity the corresponding weak formulation is fulfilled by a density argument (cf. also [28, Theorem 2.5]). ∎

In the following, we will use interpolation and embedding results summarized in the following lemma.

LEMMA 2.7.

(a) *Let X and $\mathcal{D}$ be Banach spaces with $\mathcal{D} \subset X$ dense in X. Let $1 < r < \infty$, $\theta \in (0, 1 - \frac{1}{r})$. Then the following embedding holds:*

$$W^{1,r}(I; X) \cap L^r(I; \mathcal{D}) \hookrightarrow C^\beta(\bar{I}; (X, \mathcal{D})_{\theta,1}) \quad \textit{if} \ \ 0 \le \beta \le 1 - \frac{1}{r} - \theta,$$

where $(X, \mathcal{D})_{\theta,1}$ denotes the real interpolation space between X and $\mathcal{D}$.

(b) *Let X and $\mathcal{D}$ be Banach spaces with $\mathcal{D} \subset X$ dense in X. Let $1 \le r_0, r_1 < \infty$, $\theta \in (0,1)$. Then the following embedding holds*

$$L^{r_0}(I;X) \cap L^{r_1}(I;\mathcal{D}) \hookrightarrow L^r(\bar{I};[X,\mathcal{D}]_\theta) \quad \textit{for} \quad \frac{1}{r} = \frac{1-\theta}{r_0} + \frac{\theta}{r_1},$$

where $[X,\mathcal{D}]_\theta$ denotes the complex interpolation space between X and $\mathcal{D}$.

(c) *Let X and $\mathcal{D}$ be Banach spaces with $\mathcal{D} \subset X$ dense in X. Let $\theta \in (0,1)$. There holds the following relation between the real and complex interpolation spaces,*

$$(X,\mathcal{D})_{\theta,1} \hookrightarrow [X,\mathcal{D}]_\theta \hookrightarrow (X,\mathcal{D})_{\theta,\infty}.$$

(d) *Let $\theta \in (0,1)$ and $1 < s < \infty$. There holds the following interpolation result,*

$$(W^{-1,s}(\Omega), W^{1,s}(\Omega))_{\theta,1} \hookrightarrow W^{2\theta-1,s}(\Omega).$$

(e) *Let $D \subset \Omega$ be a subdomain with smooth boundary, let $\theta \in (0,1)$, and $2 \le p < \infty$. Then the following interpolation result holds*

$$[L^p(D), W^{2,p}(D)]_\theta \hookrightarrow W^{2\theta,p}(D).$$

Proof. The embedding (a) can be found, e.g., in [2] for $\beta < 1 - \frac{1}{r} - \theta$. The case $\beta = 1 - \frac{1}{r} - \theta$ is shown in [15, Lemma 3.4]. The embedding (b) is given in [44, Chapter 1.18.4]. For the standard embedding (c) we refer, e.g., to [2]. For the interpolation result (d) we refer, e.g., to [27, Corollary A.28]. The standard interpolation result (e) can be found, e.g., in [44, Chapter 2.4.2]. ∎

As the next step we introduce the reduced cost functional $j\colon Q \to \mathbb{R}$ on the control space $Q = L^2(I)$ by

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

where J is the cost functional in (1.1) and $u(q)$ is the weak solution of the state equation (1.2) as defined above. The optimal control problem can then be equivalently reformulated as

$$\min\, j(q), \quad q \in Q_{\text{ad}}, \tag{2.10}$$

where the set of admissible controls is defined according to (1.3) by

$$Q_{\text{ad}} = \{\, q \in Q \mid q_a \le q(t) \le q_b \text{ a.e. in } I \,\}. \tag{2.11}$$

For simplicity of the presentation, we assume $0 \in [q_a, q_b]$ in the following, but all our results hold also without this assumption.

By standard arguments, this optimization problem possesses a unique solution $\bar{q} \in Q = L^2(I)$ with the corresponding state $\bar{u} = u(\bar{q}) \in L^2(I; L^p(\Omega))$ for $p < 3$; see Theorem 2.6 for the regularity of $\bar{u}$. Due to the fact that this optimal control problem is convex, the solution $\bar{q}$ is equivalently characterized by the optimality condition

$$j'(\bar{q})(\delta q - \bar{q}) \ge 0 \quad \text{for all } \delta q \in Q_{\text{ad}}. \tag{2.12}$$

The (directional) derivative $j'(q)(\delta q)$ for given $q, \delta q \in Q$ can be expressed as

$$j'(q)(\delta q) = \int_I (\alpha q(t) + z(t,x_0))\, \delta q(t)\, dt, \tag{2.13}$$

where $z = z(q) \in L^2(I; H^2(\Omega) \cap H^1_0(\Omega)) \cap H^1(I; L^2(\Omega))$ is the solution of the adjoint equation

$$
\begin{aligned}
-z_t(t,x) - \Delta z(t,x) &= u(t,x) - \widehat{u}(t,x), && (t,x) \in I \times \Omega,\\
z(t,x) &= 0, && (t,x) \in I \times \partial\Omega,\\
z(T,x) &= 0, && x \in \Omega,
\end{aligned}
\tag{2.14}
$$

and $u = u(q)$ on the right-hand side of (2.14) is the solution of the state equation (1.2). The adjoint solution, which corresponds to the optimal control $\bar{q}$, is denoted by $\bar{z} = z(\bar{q})$.

The optimality condition (2.12) is a variational inequality, which can be equivalently formulated using the pointwise projection

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

The resulting optimality condition reads

$$\bar{q} = P_{Q_{\mathrm{ad}}}\left(-\frac{1}{\alpha}\bar{z}(\cdot, x_0)\right). \tag{2.15}$$

A standard regularity result on the optimal control $\bar{q}$, the optimal state $\bar{u}$, and the optimal adjoint state $\bar{z}$ is summarized in the following lemma.

Lemma 2.8. *Let $\bar{q}$ be the optimal control, $\bar{u} = u(\bar{q})$ be the optimal state, and $\bar{z} = z(\bar{q})$ be the optimal adjoint state. Then $\bar{z} \in L^2(I; H^2(\Omega) \cap H^1_0(\Omega)) \cap H^1(I; L^2(\Omega))$ and the following estimate holds:*

$$\alpha^{\frac{1}{2}}\|\bar{q}\|_{L^2(I)} + \|\bar{u}\|_{L^2(I;L^2(\Omega))} + \|\nabla^2 \bar{z}\|_{L^2(I;L^2(\Omega))} + \|\bar{z}_t\|_{L^2(I;L^2(\Omega))} \le c\|\widehat{u}\|_{L^2(I;L^2(\Omega))}.$$

Proof. By the optimality of $(\bar{q}, \bar{u})$ we obtain

$$\|\bar{u} - \widehat{u}\|^2_{L^2(I;L^2(\Omega))} + \alpha\|\bar{q}\|^2_{L^2(I)} = 2J(\bar{q}, \bar{u}) \le 2J(0,0) = \|\widehat{u}\|^2_{L^2(I;L^2(\Omega))}.$$

Hence,

$$\alpha^{\frac{1}{2}}\|\bar{q}\|_{L^2(I)} \le \|\widehat{u}\|_{L^2(I;L^2(\Omega))} \quad \text{and} \quad \|\bar{u}\|_{L^2(I;L^2(\Omega))} \le 2\|\widehat{u}\|_{L^2(I;L^2(\Omega))}.$$

By Lemma 2.1 applied to the adjoint equation (2.14) we obtain

$$\|\nabla^2 \bar{z}\|_{L^2(I;L^2(\Omega))} + \|\bar{z}_t\|_{L^2(I;L^2(\Omega))} \le c\|\bar{u} - \widehat{u}\|_{L^2(I;L^2(\Omega))} \le c\|\widehat{u}\|_{L^2(I;L^2(\Omega))}.$$

This completes the proof. ∎

By a bootstrapping argument, we obtain the following additional regularity result for the optimal control $\bar{q}$, the optimal state $\bar{u}$, and the optimal adjoint state $\bar{z}$.

Theorem 2.9. *Let $\bar{q}$ be the optimal control, $\bar{u} = u(\bar{q})$ the optimal state, $\bar{z} = z(\bar{q})$ the optimal adjoint state, and $d > 0$ be such that $\overline{B}_{2d} = \overline{B}_{2d}(x_0) \subset \Omega$. Then, there exists a constant $c_r \le \frac{cr^2}{r-1}$ and constant c_d depending on d and independent of r, s, and p such that*

(a) *$\bar{u} \in L^r(I; W^{1,s}_0(\Omega)) \cap W^{1,r}(I; W^{-1,s}(\Omega))$ for all $1 < r < \infty$ and $1 < s < \frac{3}{2}$ with the corresponding estimate*

$$\|\bar{u}\|_{L^r(I;W^{1,s}_0(\Omega))} + \|\bar{u}\|_{W^{1,r}(I;W^{-1,s}(\Omega))} \le \frac{c_r\, c_d \alpha^{-\frac{5}{2}}}{\frac{3}{2} - s}\|\widehat{u}\|_{L^3(I;L^3(\Omega))};$$

(b) $\bar{u} \in L^r(I; L^p(\Omega))$ *for all* $1 < r < \infty$ *and* $2 \le p < 3$ *with the corresponding estimate*

$$\|\bar{u}\|_{L^r(I;L^p(\Omega))} \le \frac{c_r\, c_d\, \alpha^{-\frac{5}{2}}}{3-p} \|\widehat{u}\|_{L^3(I;L^3(\Omega))};$$

(c) $\bar{u} \in C^\gamma(\bar{I}; L^{\frac{3}{2}}(\Omega))$ *for all* $\gamma < \frac{1}{2}$ *with the corresponding estimate*

$$\|\bar{u}\|_{C^\gamma(\bar{I};L^{\frac{3}{2}}(\Omega))} \le \frac{c_d \alpha^{-\frac{5}{2}}}{\left(\frac{1}{2}-\gamma\right)^2} \|\widehat{u}\|_{L^3(I;L^3(\Omega))};$$

(d) $\bar{z} \in L^r(I; W^{2,p}(B_d)) \cap W^{1,r}(I; L^p(B_d))$ *for all* $1 < r < \infty$ *and* $2 \le p < 3$ *with the corresponding estimate*

$$\|\bar{z}\|_{L^r(I;W^{2,p}(B_d))} + \|\bar{z}\|_{W^{1,r}(I;L^p(B_d))} \le \frac{c_r^3\, c_d\, \alpha^{-\frac{5}{2}}}{3-p} \|\widehat{u}\|_{L^\infty(I;L^3(\Omega))};$$

(e) $\bar{z} \in C^\gamma(\bar{I}; C(\overline{B}_d))$ *and* $\bar{q} \in C^\gamma(\bar{I})$ *for all* $\gamma < \frac{1}{2}$ *with the corresponding estimate*

$$\alpha\|\bar{q}\|_{C^\gamma(\bar{I})} + \|\bar{z}\|_{C^\gamma(\bar{I};C(\overline{B}_d))} \le \frac{c_d \alpha^{-\frac{5}{2}}}{(\frac{1}{2}-\gamma)^5} \|\widehat{u}\|_{L^\infty(I;L^3(\Omega))}.$$

Proof. The proof consists of four steps. In the first step we prove $\bar{q} \in L^3(I)$; in the second step we show $\bar{q} \in L^\infty(I)$. In the third step we provide the regularity results (a), (b), and (c) for $\bar{u}$; in the last step we show the regularity results for the adjoint state $\bar{z}$ and, finally, $\bar{q} \in C^\gamma(\bar{I})$ for all $\gamma < \frac{1}{2}$. If the control bounds q_a, q_b are not equal to $\pm\infty$, i.e., $q_a, q_b \in \mathbb{R}$, then the first two steps can be omitted, since then $\bar{q} \in L^\infty(I)$ holds trivially.

Step 1: To show $\bar{q} \in L^3(I)$, we first observe that by Lemma 2.8 we have $\|\bar{q}\|_{L^2(I)} \le \alpha^{-\frac{1}{2}} \|\widehat{u}\|_{L^2(I;L^2(\Omega))}$. Then by the estimate (2.9) from Theorem 2.6 we obtain $\bar{u} \in L^2(I; L^p(\Omega))$ for all $1 < p < 3$ and

$$\|\bar{u}\|_{L^2(I;L^p(\Omega))} \le \frac{c}{3-p} \|\bar{q}\|_{L^2(I)} \le \frac{c\alpha^{-\frac{1}{2}}}{3-p} \|\widehat{u}\|_{L^2(I;L^2(\Omega))}.$$

Then, applying Lemma 2.4 to the adjoint equation (2.14) we obtain $\bar{z} \in L^2(I; W^{2,p}(B_d)) \cap H^1(I; L^p(B_d))$ and

$$\begin{aligned}
&\|\bar{z}\|_{L^2(I;W^{2,p}(B_d))} + \|\bar{z}\|_{H^1(I;L^p(B_d))} \\
&\quad \le c\|\bar{u} - \widehat{u}\|_{L^2(I;L^p(B_{2d}))} + c\, d^{-3\left(\frac{1}{2}-\frac{1}{p}\right)} \|\bar{u} - \widehat{u}\|_{L^2(I;L^2(\Omega))} \\
&\quad \le c_d \|\bar{u} - \widehat{u}\|_{L^2(I;L^p(\Omega))} \le \frac{c_d \alpha^{-\frac{1}{2}}}{3-p} \|\widehat{u}\|_{L^2(I;L^p(\Omega))} \\
&\quad \le \frac{c_d \alpha^{-\frac{1}{2}}}{3-p} \|\widehat{u}\|_{L^2(I;L^3(\Omega))}.
\end{aligned}$$

Using the fact that $H^1(I; L^p(B_d)) \hookrightarrow L^t(I; L^p(B_d))$ for all $1 \le t \le \infty$ we obtain by Lemma 2.7(b) and (e)

$$\begin{aligned}
L^2(I; W^{2,p}(B_d)) \cap H^1(I; L^p(B_d)) &\hookrightarrow L^2(I; W^{2,p}(B_d)) \cap L^t(I; L^p(B_d)) \\
&\hookrightarrow L^3(I; [L^p(B_d), W^{2,p}(B_d)]_\theta) \\
&\hookrightarrow L^3(I; W^{2\theta,p}(B_d)) \quad \text{for all } 0 \le \theta < \frac{2}{3}.
\end{aligned}$$

It is possible to choose $\theta \in (0, \frac{2}{3})$ and $p \in (1,3)$, e.g., $\theta = \frac{3}{5}$, $p = \frac{8}{3}$, such that

$$2\theta - \frac{3}{p} > 0$$

and therefore $W^{2\theta,p}(B_d) \hookrightarrow C(\overline{B}_d)$. Hence, $\bar z \in L^3(I; C(\overline{B}_d))$. By the optimality condition (2.15) we obtain $\bar q \in L^3(I)$ and

$$\|\bar q\|_{L^3(I)} \le \alpha^{-1} \|\bar z\|_{L^3(I;C(\overline{B}_d))} \le c_d \alpha^{-\frac{3}{2}} \|\widehat u\|_{L^2(I;L^3(\Omega))}.$$

Step 2: Using the previous estimate, we obtain again by the estimate (2.9) from Theorem 2.6 that $\bar u \in L^3(I; L^p(\Omega))$ for all $1 < p < 3$ and

$$\|\bar u\|_{L^3(I;L^p(\Omega))} \le \frac{c}{3-p} \|\bar q\|_{L^3(I)} \le \frac{c\alpha^{-\frac{3}{2}}}{3-p} \|\widehat u\|_{L^2(I;L^3(\Omega))}.$$

Then applying Lemma 2.4 to the adjoint equation we obtain $\bar z \in L^3(I; W^{2,p}(B_d)) \cap W^{1,3}(I; L^p(B_d))$ and

$$\begin{aligned}
&\|\bar z\|_{L^3(I;W^{2,p}(B_d))} + \|\bar z\|_{W^{1,3}(I;L^p(B_d))} \\
&\quad \le c\|\bar u - \widehat u\|_{L^3(I;L^p(B_{2d}))} + c\, d^{-3\left(\frac{1}{2}-\frac{1}{p}\right)} \|\bar u - \widehat u\|_{L^3(I;L^2(\Omega))} \\
&\quad \le c_d \|\bar u - \widehat u\|_{L^3(I;L^p(\Omega))} \le \frac{c_d \alpha^{-\frac{3}{2}}}{3-p} \|\widehat u\|_{L^3(I;L^3(\Omega))}.
\end{aligned}$$

In contrast to the procedure in the first step, we use here the embedding (a) from Lemma 2.7 and obtain, together with (c) and (e) from Lemma 2.7,

$$\bar z \in C^\beta(\bar I, (L^p(B_d), W^{2,p}(B_d))_{\theta,1}) \hookrightarrow C^\beta(\bar I, [L^p(B_d), W^{2,p}(B_d)]_\theta) \hookrightarrow C^\beta(\bar I, W^{2\theta,p}(B_d))$$

with $\beta = \frac{2}{3} - \theta$ for all $0 \le \theta < \frac{2}{3}$. It is possible to choose $\theta \in (0, \frac{2}{3})$ and $p \in (1,3)$, e.g., $\theta = \frac{3}{5}$, $p = \frac{8}{3}$, such that

$$2\theta - \frac{3}{p} > 0$$

and therefore $W^{2\theta,p}(B_d) \hookrightarrow C(\overline{B}_d)$. Hence, $\bar z \in L^\infty(I; C(\overline{B}_d))$. By the optimality condition (2.15) we obtain $\bar q \in L^\infty(I)$ and

$$\|\bar q\|_{L^\infty(I)} \le c_d \alpha^{-\frac{5}{2}} \|\widehat u\|_{L^3(I;L^3(\Omega))}. \tag{2.16}$$

Step 3: By the above result (2.16) and by Theorem 2.6 we obtain $\bar u \in L^r(I; W_0^{1,s}(\Omega)) \cap W^{1,r}(I; W^{-1,s}(\Omega))$ for all $1 < s < \frac{3}{2}$ and all $1 < r < \infty$ and the desired estimate (a):

$$\begin{aligned}
\|\bar u\|_{L^r(I;W_0^{1,s}(\Omega))} + \|\bar u\|_{W^{1,r}(I;W^{-1,s}(\Omega))} &\le \frac{c_r}{\frac{3}{2}-s} \|\bar q\|_{L^r(I)} \le \frac{c_r}{\frac{3}{2}-s} \|\bar q\|_{L^\infty(I)} \\
&\le \frac{c_r\, c_d \alpha^{-\frac{5}{2}}}{\frac{3}{2}-s} \|\widehat u\|_{L^3(I;L^3(\Omega))}.
\end{aligned}$$

Similarly, again by Theorem 2.6 we obtain $\bar u \in L^r(I; L^p(\Omega))$ for all $1 < p < 3$ and the desired estimate (b).

Using the interpolation and embedding results (a) and (d) from Lemma 2.7 we obtain

$$L^r(I;W_0^{1,s}(\Omega)) \cap W^{1,r}(I;W^{-1,s}(\Omega)) \hookrightarrow C^\beta(\bar I;W^{2\theta-1,s}(\Omega))$$

which holds for $0 < \beta = 1 - \frac{1}{r} - \theta$. We choose

$$\theta_s = \frac{1}{2} + \frac{\frac{3}{2} - s}{s}$$

and obtain by the Sobolev embedding $W^{2\theta-1,s}(\Omega) \hookrightarrow L^{\frac{3}{2}}(\Omega)$. For $\frac{6}{5} < s < \frac{3}{2}$ we choose

$$r_s = \frac{s}{\frac{3}{2} - s}$$

which fulfills

$$\beta_s = 1 - \frac{1}{r_s} - \theta_s = \frac{1}{2} - \frac{2\left(\frac{3}{2} - s\right)}{s} \in \left(0, \frac{1}{2}\right),$$

and therefore

$$\bar u \in C^{\beta_s}(\bar I; L^{\frac{3}{2}}(\Omega)) \quad \text{for} \;\; \beta_s = \frac{1}{2} - \frac{2\left(\frac{3}{2} - s\right)}{s}$$

with

$$\|\bar u\|_{C^{\beta_s}(\bar I;L^{\frac{3}{2}}(\Omega))} \le \frac{c_{r_s}\, c_d \alpha^{-\frac{5}{2}}}{\frac{3}{2} - s} \|\widehat u\|_{L^3(I;L^3(\Omega))} \le \frac{c_d \alpha^{-\frac{5}{2}}}{\left(\frac{3}{2} - s\right)^2} \|\widehat u\|_{L^3(I;L^3(\Omega))}.$$

For each $\gamma < \frac{1}{2}$ we express s in terms of $\gamma = \beta_s$ and obtain

$$\|\bar u\|_{C^{\gamma}(\bar I;L^{\frac{3}{2}}(\Omega))} \le \frac{c_d \alpha^{-\frac{5}{2}}}{\left(\frac{1}{2} - \gamma\right)^2} \|\widehat u\|_{L^3(I;L^3(\Omega))}.$$

Step 4: Using already proven estimate (b) and applying Lemma 2.4 to the adjoint equation (2.14) we obtain $\bar z \in L^r(I;W^{2,p}(B_d)) \cap W^{1,r}(I;L^p(B_d))$ and

$$\begin{aligned}
&\|\bar z\|_{L^r(I;W^{2,p}(B_d))} + \|\bar z\|_{W^{1,r}(I;L^p(B_d))} \\
&\quad \le c_r^2 \|\bar u - \widehat u\|_{L^r(I;L^p(B_{2d}))} + c_r^2\, d^{-3\left(\frac{1}{2} - \frac{1}{p}\right)} \|\bar u - \widehat u\|_{L^r(I;L^2(\Omega))} \\
&\quad \le c_r^2\, c_d \|\bar u - \widehat u\|_{L^r(I;L^p(\Omega))} \le \frac{c_r^3\, c_d \alpha^{-\frac{5}{2}}}{3 - p} \|\widehat u\|_{L^\infty(I;L^3(\Omega))},
\end{aligned}$$

which gives (d). As in Step 2, we use here the embedding (a) from Lemma 2.7 and obtain together with (c) and (e) from Lemma 2.7

$$\bar z \in C^\beta(\bar I, (L^p(B_d), W^{2,p}(B_d))_{\theta,1}) \hookrightarrow C^\beta(\bar I, [L^p(B_d), W^{2,p}(B_d)]_\theta) \hookrightarrow C^\beta(\bar I, W^{2\theta,p}(B_d))$$

for all $0 < \beta = 1 - \frac{1}{r} - \theta$. We choose

$$\theta_p = \frac{1}{2} + \frac{3 - p}{p}$$

resulting in

$$2\theta_p - \frac{3}{p} = \frac{3 - p}{p} > 0.$$

Therefore by embedding (2.3) we have $\bar z \in C^\beta(\bar I; C(\overline{B}_d))$ with

$$\|\bar z\|_{C^\beta(\bar I; C(\overline{B}_d))} \le \frac{c}{3-p}\|\bar z\|_{C^\beta(\bar I, W^{2\theta,p}(B_d))} \le \frac{c_r^3 c_d \alpha^{-\frac{5}{2}}}{(3-p)^2}\|\widehat u\|_{L^\infty(I;L^3(\Omega))}$$

for $\beta = \frac{1}{2} - \frac{1}{r} - \frac{3-p}{p}$ and r such that $\beta > 0$. For $\frac{12}{5} < p < 3$ we choose $r_p = \frac{p}{3-p}$ and obtain

$$\beta_p = \frac{1}{2} - \frac{2\,(3-p)}{p} \in \left(0, \frac{1}{2}\right)$$

and

$$\|\bar z\|_{C^{\beta_p}(\bar I; C(\overline{B}_d))} \le \frac{c_d \alpha^{-\frac{5}{2}}}{(3-p)^5}\|\widehat u\|_{L^\infty(I;L^3(\Omega))}.$$

For each $\gamma < \frac{1}{2}$ we express p in terms of $\gamma = \beta_p$ and obtain

$$\|\bar z\|_{C^\gamma(\bar I; C(\overline{B}_d))} \le \frac{c_d \alpha^{-\frac{5}{2}}}{(\frac{1}{2}-\gamma)^5}\|\widehat u\|_{L^\infty(I;L^3(\Omega))}.$$

Using again the optimality condition (2.15) and the stability of the projection operator $P_{Q_{\text{ad}}}$ in C^γ we obtain $\bar q \in C^\gamma(I)$ for all $0 \le \gamma < \frac{1}{2}$ with

$$\|\bar q\|_{C^\gamma(\bar I)} \le \frac{c_d \alpha^{-\frac{7}{2}}}{(\frac{1}{2}-\gamma)^5}\|\widehat u\|_{L^\infty(I;L^3(\Omega))}.$$

The last two inequalities result in the desired estimate (e). ∎

3. Discretization and best approximation results. For the discretization of the problem under consideration we introduce a partition of $\bar I = [0, T]$ into subintervals $I_m = (t_{m-1}, t_m]$ of length $k_m = t_m - t_{m-1}$, where $0 = t_0 < t_1 < \cdots < t_{M-1} < t_M = T$. The maximal time step is denoted by $k = \max_m k_m$. We impose the following conditions on the time mesh (cf. [36]):

(i) There are constants $c, \gamma > 0$ independent of k such that

$$\min_{m=1,2,\ldots,M} k_m \ge c k^\gamma.$$

(ii) There is a constant $\kappa > 0$ independent of k such that for all $m = 1, 2, \ldots, M-1$

$$\kappa^{-1} \le \frac{k_m}{k_{m+1}} \le \kappa.$$

(iii) It holds $k \le \frac{1}{4}\min(T, 1)$.

These assumptions allow for a large class of time meshes, also with a strong mesh grading.

The semidiscrete space X_k^0 of piecewise constant functions in time is defined by

$$X_k^0 = \{v_k \in L^2(I; H_0^1(\Omega)) : \ v_k|_{I_m} \in \mathcal{P}_0(I; H_0^1(\Omega)), \ m = 1, 2, \ldots, M\},$$

where $\mathcal{P}_0(I; V)$ is the space of constant functions in time with values in V. We will employ the following notation for functions in X_k^0:

(3.1)
$$v_m^+ = \lim_{\varepsilon\to 0^+} v(t_m+\varepsilon) := v_{m+1}, \quad v_m^- = \lim_{\varepsilon\to 0^+} v(t_m-\varepsilon) = v(t_m) := v_m, \quad [v]_m = v_m^+ - v_m^-.$$

Let $\mathcal{T}$ denote a quasi-uniform triangulation of Ω with mesh size h, i.e., $\mathcal{T} = \{\tau\}$ is a partition of Ω into tetrahedrons (cells) τ of diameter h_τ such that for $h = \max_\tau h_\tau$,

$$\operatorname{diam}(\tau) \le h \le c|\tau|^{\frac{1}{3}} \quad \forall \tau \in \mathcal{T}$$

holds. Let V_h be the set of all functions in $H_0^1(\Omega)$ that are linear on each τ, i.e., V_h is the usual space of linear finite elements, and let $i_h \colon C_0(\Omega) \to V_h$ be the usual nodal interpolant.

To obtain the fully discrete approximation we consider the space-time finite element space

$$X_{k,h}^{0,1} = \{v_{kh} \in X_k^0 :\ v_{kh}|_{I_m} \in \mathcal{P}_0(I; V_h),\ m = 1, 2, \dots, M\}. \tag{3.2}$$

We will also need the following semidiscrete interpolant $\pi_k \colon C(\bar{I}; H_0^1(\Omega)) \to X_k^0$ defined by

$$\pi_k v|_{I_m}(\cdot) = v(t_m, \cdot), \quad m = 1, 2, \dots, M, \tag{3.3}$$

and the semidiscrete L^2 projection $P_k \colon L^2(I; H^1(\Omega)) \to X_k^0$ defined by

$$P_k v|_{I_m}(\cdot) = \frac{1}{k_m} \int_{I_m} v(t, \cdot)\, dt, \quad m = 1, 2, \dots, M. \tag{3.4}$$

We note that, for any Banach space X, e.g., $X = L^\infty(D)$ with $D \subset \Omega$, we have stability of the projection $P_k \colon L^2(I; H^1(\Omega)) \cap L^2(I; X) \to X_k^0 \cap L^2(I; X)$ with respect to the $L^2(I; X)$ norm. There holds

$$\begin{aligned} \|P_k v\|_{L^2(I;X)}^2 &= \sum_{m=1}^M k_m \left\| \frac{1}{k_m} \int_{I_m} v(t)\, dt \right\|_X^2 \le \sum_{m=1}^M \frac{1}{k_m} \left(\int_{I_m} \|v(t)\|_X \, dt \right)^2 \\ &\le \sum_{m=1}^M \int_{I_m} \|v(t)\|_X^2 \, dt = \|v\|_{L^2(I;X)}^2, \end{aligned} \tag{3.5}$$

where we have used the Cauchy–Schwarz inequality.

To introduce the cG(1)dG(0) discretization we define the following bilinear form

$$B(v, \varphi) = \sum_{m=1}^M \langle v_t, \varphi \rangle_{I_m \times \Omega} + (\nabla v, \nabla \varphi)_{I \times \Omega} + \sum_{m=2}^M ([v]_{m-1}, \varphi_{m-1}^+)_\Omega + (v_0^+, \varphi_0^+)_\Omega, \tag{3.6}$$

where $\langle \cdot, \cdot \rangle_{I_m \times \Omega}$ is the duality product between $L^2(I_m; W^{-1,s}(\Omega))$ and $L^2(I_m; W_0^{1,s'}(\Omega))$. We note that the first sum vanishes for $v \in X_k^0$. Rearranging the terms in (3.6), we obtain an equivalent (dual) expression of B:

$$B(v, \varphi) = -\sum_{m=1}^M \langle v, \varphi_t \rangle_{I_m \times \Omega} + (\nabla v, \nabla \varphi)_{I \times \Omega} - \sum_{m=1}^{M-1} (v_m^-, [\varphi_k]_m)_\Omega + (v_M^-, \varphi_M^-)_\Omega. \tag{3.7}$$

For the solution v of the auxiliary equation (2.1) we consider its dG(0) semidiscrete (in time) approximation $v_k \in X_k^0$ and its cG(1)dG(0) fully discrete approximation $v_{kh} \in X_{k,h}^{0,1}$ defined as

$$B(v_k, \varphi_k) = (f, \varphi_k)_{I \times \Omega} \quad \text{for all} \;\; \varphi_k \in X_k^0 \tag{3.8}$$

and

$$B(v_{kh}, \varphi_{kh}) = (f, \varphi_{kh})_{I\times\Omega} \quad \text{for all } \ \varphi_{kh} \in X^{0,1}_{k,h}. \tag{3.9}$$

Since this method leads to a consistent discretization, we have the following Galerkin orthogonality relations,

$$B(v - v_k, \varphi_k) = 0 \quad \text{for all } \ \varphi_k \in X^0_k$$

and

$$B(v - v_{kh}, \varphi_{kh}) = 0 \quad \text{for all } \ \varphi_{kh} \in X^{0,1}_{k,h}.$$

In the following we will use the following semidiscrete and fully discrete maximal parabolic regularity results from [31].

THEOREM 3.1 (see [31, Theorem 2]). *Let $f \in L^r(I; L^p(\Omega))$, $1 \le r, p \le \infty$, and let $v_k \in X^0_k$ be the solution of* (3.8). *There exists a constant c independent of k and f such that*

$$\|\Delta v_k\|_{L^r(I;L^p(\Omega))} \le c \ln\frac{T}{k}\|f\|_{L^r(I;L^p(\Omega))}.$$

THEOREM 3.2 (see [31, Theorem 11]). *Let $f \in L^r(I; L^p(\Omega))$, $1 \le r, p \le \infty$, and let $v_{kh} \in X^{0,1}_{k,h}$ be the solution of* (3.9). *There exists a constant c independent of k and f such that*

$$\|\Delta_h v_{kh}\|_{L^r(I;L^p(\Omega))} \le c \ln\frac{T}{k}\|f\|_{L^r(I;L^p(\Omega))},$$

where $\Delta_h \colon V_h \to V_h$ is the discrete Laplace operator which is defined later in (5.6).

In the following we establish global and interior best-approximation-type results in the $L^\infty(\Omega; L^2(I))$ norm. These results constitute the main technical tools for proving our main result.

THEOREM 3.3 (global best approximation). *Let v, v_k, and v_{kh} satisfy* (2.1), (3.8)*, and* (3.9)*, respectively. There exists a constant c independent of k, h, and v such that for any $1 \le p \le \infty$ the estimates*

$$\|v - v_{kh}\|^2_{L^\infty(\Omega;L^2(I))} \le c\left(\ln\frac{T}{k}\right)^2 |\ln h|^2 \inf_{\chi\in X^{0,1}_{k,h}} \left(\|v-\chi\|^2_{L^2(I;L^\infty(\Omega))} + h^{-\frac{6}{p}}\|\pi_k v - \chi\|^2_{L^2(I;L^p(\Omega))}\right)$$

and

$$\|v_k - v_{kh}\|^2_{L^\infty(\Omega;L^2(I))} \le c\left(\ln\frac{T}{k}\right)^2 |\ln h|^2 \inf_{\chi\in X^{0,1}_{k,h}} \|v_k-\chi\|^2_{L^2(I;L^\infty(\Omega))}$$

hold.

The proof of this theorem is given in section 6. Note, that the norm on the left-hand side of these estimates is $L^\infty(\Omega; L^2(I))$ and on the right-hand side is $L^2(I; L^\infty(\Omega))$. Therefore, we call this result a *best-approximation-type* estimate.

For the error at point x_0 we are able to obtain a sharper result, which shows a more localized dependence of the error at a point. For elliptic problems a similar result was obtained in [41]. As before, we denote by $B_d = B_d(x_0)$ a ball of radius d centered at x_0.

THEOREM 3.4 (interior best approximation). *Let v, v_k, and v_{kh} satisfy* (2.1), (3.8), *and* (3.9), *respectively, and let $d > 4h$. There exists a constant c independent of h, k, d, and v such that for any $1 \le p \le \infty$ the estimates*

$$\begin{aligned}\int_0^T &|(v - v_{kh})(t, x_0)|^2 dt \\ &\le c \left(\ln \frac{T}{k}\right)^2 |\ln h|^2 \inf_{\chi \in X_{k,h}^{0,1}} \Big\{ \|v - \chi\|^2_{L^2(I;L^\infty(B_d(x_0)))} + h^{-\frac{6}{p}} \|\pi_k v - \chi\|^2_{L^2(I;L^p(B_d(x_0)))} \\ &\qquad + d^{-3}\Big(\|v - \chi\|^2_{L^2(I;L^2(\Omega))} + \|\pi_k v - \chi\|^2_{L^2(I;L^2(\Omega))} \\ &\qquad + h^2 \|\nabla(v - \chi)\|^2_{L^2(I;L^2(\Omega))}\Big)\Big\}\end{aligned} \tag{3.10}$$

and

$$\begin{aligned}\int_0^T &|(v_k - v_{kh})(t, x_0)|^2 dt \\ &\le c \left(\ln \frac{T}{k}\right)^2 |\ln h|^2 \inf_{\chi \in X_{k,h}^{0,1}} \Big\{ \|v_k - \chi\|^2_{L^2(I;L^\infty(B_d(x_0)))} \\ &\qquad + d^{-3}\Big(\|v_k - \chi\|^2_{L^2(I;L^2(\Omega))} \\ &\qquad + h^2 \|\nabla(v_k - \chi)\|^2_{L^2(I;L^2(\Omega))}\Big)\Big\}\end{aligned} \tag{3.11}$$

hold.

The proof of this theorem is given in section 7.

Remark 3.5. In [30, Theorem 3.5] we formulated the corresponding two-dimensional best-approximation-type result. There, the term $h^2\|\nabla(v - \chi)\|^2_{L^2(I;L^2(\Omega))}$ was forgotten due to a small mistake in the proof. Since both terms, $\|v - \chi\|^2_{L^2(I;L^2(\Omega))}$ and $h^2\|\nabla(v-\chi)\|^2_{L^2(I;L^2(\Omega))}$, can be estimated by the exact same term, this additional term has no influence on further results in [30]. The proof for the three-dimensional case presented below in section 7 works also in the two-dimensional case.

4. Discretization of the optimal control problem. In this section, we describe the discretization of the optimal control problem (1.1)–(1.2) and prove an estimate for the error $\|\bar q - \bar q_{kh}\|_{L^2(I)}$ between the continuous and the discrete optimal control.

We start with the discretization of the state equation. For a given control $q \in Q$ we define the corresponding discrete state $u_{kh} = u_{kh}(q) \in X_{k,h}^{0,1}$ by

$$B(u_{kh}, \varphi_{kh}) = \int_0^T q(t)\varphi_{kh}(t, x_0)\, dt \quad \text{for all } \varphi_{kh} \in X_{k,h}^{0,1}. \tag{4.1}$$

Using the weak formulation for $u = u(q)$ from Theorem 2.6, we obtain that this discretization is consistent, i.e., the Galerkin orthogonality holds:

$$B(u - u_{kh}, \varphi_{kh}) = 0 \quad \text{for all } \varphi_{kh} \in X_{k,h}^{0,1}.$$

Note, that the jump terms involving u vanish due to the fact that $u \in C(\bar{I}; W^{-1,s}(\Omega))$ and $\varphi_{kh,m} \in W^{1,\infty}(\Omega)$.

As on the continuous level, we define the discrete reduced cost functional $j_{kh}\colon Q \to \mathbb{R}$ by

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

where J is the cost functional in (1.1). The discretized optimal control problem is then given as

$$\min\, j_{kh}(q), \quad q \in Q_{\mathrm{ad}}, \tag{4.2}$$

where Q_{ad} is the set of admissible controls (2.11). We note that the control variable q is not explicitly discretized (cf. [30] and [25]), but the optimal control is computable as a piecewise constant function; see the discussion below. With standard arguments, one proves the existence of a unique solution $\bar{q}_{kh} \in Q_{\mathrm{ad}}$ of (4.2). Due to convexity of the problem, the following condition is necessary and sufficient for optimality:

$$j'_{kh}(\bar{q}_{kh})(\delta q - \bar{q}_{kh}) \geq 0 \quad \text{for all} \ \ \delta q \in Q_{\mathrm{ad}}. \tag{4.3}$$

As on the continuous level, the directional derivative $j'_{kh}(q)(\delta q)$ for given $q, \delta q \in Q$ can be expressed as

$$j'_{kh}(q)(\delta q) = \int_I (\alpha q(t) + z_{kh}(t, x_0))\, \delta q(t)\, dt, \tag{4.4}$$

where $z_{kh} = z_{kh}(q) \in X^{0,1}_{k,h}$ is the solution of the discrete adjoint equation

$$B(\varphi_{kh}, z_{kh}) = (u_{kh}(q) - \widehat{u}, \varphi_{kh})_{I\times\Omega} \quad \text{for all} \ \ \varphi_{kh} \in X^{0,1}_{k,h}. \tag{4.5}$$

The discrete adjoint state corresponding to the discrete optimal control $\bar{q}_{kh}$ is denoted by $\bar{z}_{kh} = z_{kh}(\bar{q}_{kh})$. The variational inequality (4.3) is equivalent to the following pointwise projection formula (cf. (2.15)),

$$\bar{q}_{kh} = P_{Q_{\mathrm{ad}}}\left(-\frac{1}{\alpha}\bar{z}_{kh}(\cdot, x_0)\right).$$

Due to the fact that $\bar{z}_{kh} \in X^{0,1}_{k,h}$, we have that $\bar{z}_{kh}(\cdot, x_0)$ is piecewise constant and therefore, by the projection formula, $\bar{q}_{kh}$ is also piecewise constant. Therefore, the optimization problem (4.2) with a nondiscretized control variable is equivalent to the corresponding optimal control problem, where the control is searched for in the space of piecewise constant functions.

To prove an estimate for the error $\|\bar{q} - \bar{q}_{kh}\|_{L^2(I)}$, we first need estimates for the error in the state and in the adjoint variables corresponding to the optimal control $\bar{q}$. Due to the structure of the optimality conditions, we will have to estimate the error $\|\bar{z}(\cdot, x_0) - \hat{z}_{kh}(\cdot, x_0)\|_{L^2(I)}$, where $\bar{z} = z(\bar{q})$ and $\hat{z}_{kh} = z_{kh}(\bar{q})$. Note, that $\hat{z}_{kh}$ is not the Galerkin projection of $\bar{z}$ due to the fact that the right-hand side of the adjoint equation (2.14) involves $\bar{u} = u(\bar{q})$ and the right-hand side of the discrete adjoint equation for $\hat{z}_{kh}$ involves $\hat{u}_{kh} = u_{kh}(\bar{q})$; see the details below. To obtain an estimate of optimal order, we will first estimate the error $\bar{u} - \hat{u}_{kh}$ with respect to the $L^2(I; L^{\frac{3}{2}}(\Omega))$ norm. In the two-dimensional case, the appropriate choice is the $L^2(I; L^1(\Omega))$ norm; see [30]. Note that an $L^2(I \times \Omega)$ estimate would not lead to an optimal result.

THEOREM 4.1. *Let $\bar{q}$ be the optimal control and let $\bar{u} = u(\bar{q})$ be the optimal state. Let, moreover, $\hat{u}_{kh} = u_{kh}(\bar{q}) \in X^{0,1}_{k,h}$ be the solution of the discrete state equation* (4.1) *with $q = \bar{q}$. Then the following estimate holds:*

$$\|\bar{u} - \hat{u}_{kh}\|_{L^2(I;L^{\frac{3}{2}}(\Omega))} \le c_d \alpha^{-\frac{1}{2}} \left(\ln \frac{T}{k}\right)^3 \left(\alpha^{-2} k^{\frac{1}{2}} + h|\ln h|\right) \|\widehat{u}\|_{L^3(I;L^3(\Omega))},$$

where c_d is a constant depending on the radius $d > 4h$ of the largest ball centered at x_0 that is contained in Ω.

Proof. First we introduce $\hat{u}_k \in X^0_k$, which is the semidiscrete approximation of $\bar{u}$ defined by

$$B(\bar{u} - \hat{u}_k, \varphi_k) = 0 \quad \text{for all} \ \ \varphi_k \in X^0_k.$$

By the triangle inequality, we have

$$\|\bar{u} - \hat{u}_{kh}\|_{L^2(I;L^{\frac{3}{2}}(\Omega))} \le \|\bar{u} - \hat{u}_k\|_{L^2(I;L^{\frac{3}{2}}(\Omega))} + \|\hat{u}_k - \hat{u}_{kh}\|_{L^2(I;L^{\frac{3}{2}}(\Omega))}.$$

We estimate both terms on the right-hand side separately. Further, we decompose the error $e_k := \bar{u} - \hat{u}_k$ as

$$e_k := \bar{u} - \hat{u}_k = (\bar{u} - \pi_k \bar{u}) + (\pi_k \bar{u} - \hat{u}_k) =: \xi_k + \eta_k.$$

A direct consequence of the semidiscrete maximal parabolic regularity result from Theorem 3.1 is the almost best approximation with respect to the $L^2(I; L^{\frac{3}{2}}(\Omega))$ norm; see [31, Theorem 9]. It says that

$$\|e_k\|_{L^2(I;L^{\frac{3}{2}}(\Omega))} \le c \ln \frac{T}{k} \|\xi_k\|_{L^2(I;L^{\frac{3}{2}}(\Omega))}.$$

To estimate the interpolation error ξ_k we use the regularity result $\bar{u} \in C^\gamma(\bar{I}; L^{\frac{3}{2}}(\Omega))$ from Theorem 2.9(c). We note that, the pointwise interpolant $\pi_k \bar{u}$ is well-defined. We choose

$$\gamma = \gamma_k = \frac{1}{2} - |\ln k|^{-1}$$

and therefore by the estimate (c) from Theorem 2.9

$$\|\xi_k\|_{L^2(I;L^{\frac{3}{2}}(\Omega))} \le c k^{\gamma_k} \|\bar{u}\|_{C^{\gamma_k}(\bar{I};L^{\frac{3}{2}}(\Omega))} \le c_d k^{\frac{1}{2}} |\ln k|^2 \alpha^{-\frac{5}{2}} \|\widehat{u}\|_{L^3(I;L^3(\Omega))}$$

resulting in

$$\|e_k\|_{L^2(I;L^{\frac{3}{2}}(\Omega))} \le c_d \alpha^{-\frac{5}{2}} k^{\frac{1}{2}} \left(\ln \frac{T}{k}\right)^3 \|\widehat{u}\|_{L^3(I;L^3(\Omega))}, \tag{4.6}$$

where we used $|\ln k| \le c \ln \frac{T}{k}$. It remains to estimate $e_h := \hat{u}_k - \hat{u}_{kh}$. To this end we obtain by duality

$$\|e_h\|_{L^2(I;L^{\frac{3}{2}}(\Omega))} = \sup_{\substack{\psi \in L^2(I;L^3(\Omega)) \\ \|\psi\|_{L^2(I;L^3(\Omega))} \le 1}} (e_h, \psi).$$

For such fixed $\psi \in L^2(I; L^3(\Omega))$ we introduce $y_k \in X_k$ satisfying

$$B(\varphi_k, y_k) = (\psi, \varphi_k) \quad \text{for all} \ \ \varphi_k \in X^0_k, \tag{4.7}$$

and its fully discrete analog $y_{kh} \in X^{0,1}_{k,h}$ by

$$B(\varphi_{kh}, y_{kh}) = (\psi, \varphi_{kh}) \quad \text{for all } \varphi_{kh} \in X^{0,1}_{k,h}.$$

Thus applying the estimate (3.11) from Theorem 3.4, we obtain

$$
\begin{aligned}
(e_h, \psi) &= (\hat{u}_k, \psi) - (\hat{u}_{kh}, \psi) \\
&= B(\hat{u}_k, y_k) - B(\hat{u}_{kh}, y_{kh}) = \int_I \bar{q}(t)\,(y_k(t, x_0) - y_{kh}(t, x_0))\,dt \\
&\le \|\bar{q}\|_{L^2(I)} \left(\int_I (y_k(t, x_0) - y_{kh}(t, x_0))^2\,dt \right)^{\frac{1}{2}} \\
&\le c \ln\frac{T}{k} |\ln h|\, \|\bar{q}\|_{L^2(I)} \Big(\|y_k - i_h y_k\|_{L^2(I;L^\infty(B_d))} + d^{-\frac{3}{2}} \|y_k - i_h y_k\|_{L^2(I;L^2(\Omega))} \\
&\qquad + d^{-\frac{3}{2}} h \|\nabla(y_k - i_h y_k)\|_{L^2(I;L^2(\Omega))} \Big) \\
&\le c \ln\frac{T}{k} |\ln h|\, \|\bar{q}\|_{L^2(I)} \left(h\|y_k\|_{L^2(I;W^{2,3}(B_d))} + h^2 d^{-\frac{3}{2}} \|y_k\|_{L^2(I;H^2(\Omega))} \right) \\
&\le c \ln\frac{T}{k} |\ln h|\, \|\bar{q}\|_{L^2(I)} \left(h\|y_k\|_{L^2(I;W^{2,3}(B_d))} + h d^{-\frac{1}{2}} \|\psi\|_{L^2(I;L^2(\Omega))} \right),
\end{aligned}
\tag{4.8}
$$

where in the last step we used the standard semidiscrete stability result; see, e.g., [37, Theorem 4.1] and $h < cd$. Using the interior elliptic estimate from Lemma 2.5 we obtain

$$
\begin{aligned}
\|y_k\|_{L^2(I;W^{2,3}(B_d))} &\le c\|\Delta y_k\|_{L^2(I;L^3(B_{2d}))} + cd^{-\frac{1}{2}} \|\Delta y_k\|_{L^2(I;L^2(\Omega))} \\
&\le c\|\Delta y_k\|_{L^2(I;L^3(\Omega))} + cd^{-\frac{1}{2}} \|\Delta y_k\|_{L^2(I;L^2(\Omega))} \\
&\le c \ln\frac{T}{k} \|\psi\|_{L^2(I;L^3(\Omega))} + cd^{-\frac{1}{2}} \|\psi\|_{L^2(I;L^2(\Omega))},
\end{aligned}
$$

where in the last step we used the semidiscrete maximal parabolic regularity result from Theorem 3.1. Inserting this estimate into (4.8) we get

$$(e_h, \psi) \le c \left(\ln\frac{T}{k} \right)^2 d^{-\frac{1}{2}} h |\ln h|\, \|\bar{q}\|_{L^2(I)} \|\psi\|_{L^2(I;L^3(\Omega))}.$$

Using the estimate from Lemma 2.8 we obtain

$$\|\hat{u}_k - \hat{u}_{kh}\|_{L^2(I;L^{\frac{3}{2}}(\Omega))} \le c_d \alpha^{-\frac{1}{2}} \left(\ln\frac{T}{k} \right)^2 h|\ln h|\, \|\widehat{u}\|_{L^2(I;L^2(\Omega))}.$$

Combining the above estimate with (4.6), we complete the proof. ∎

In the next theorem we estimate the error in the adjoint state.

THEOREM 4.2. *Let $\bar{q}$ be the optimal control, let $\bar{u} = u(\bar{q})$ be the optimal state, and $\bar{z} = z(\bar{q})$ be the optimal adjoint state. Let moreover $\hat{u}_{kh} = u_{kh}(\bar{q}) \in X^{0,1}_{k,h}$ be the solution of the discrete state equation* (4.1) *with $q = \bar{q}$ and $\hat{z}_{kh} = z_{kh}(\bar{q})$ be the corresponding discrete adjoint state, i.e., the solution of* (4.5) *with the right-hand side $\hat{u}_{kh} - \widehat{u}$. Then the following estimate holds:*

$$\left(\int_0^T (\bar{z}(t, x_0) - \hat{z}_{kh}(t, x_0))^2\,dt \right)^{\frac{1}{2}} \le c_d c_\alpha \left(\ln\frac{T}{k} \right)^6 |\ln h|^2 \left(k^{\frac{1}{2}} + h \right) \|\widehat{u}\|_{L^\infty(I;L^3(\Omega))},$$

where c_d is a constant depending on the radius $d > 4h$ of the largest ball centered at x_0 that is contained in Ω and $c_\alpha = \max(\alpha^{-\frac{5}{2}}, 1)$.

Proof. We introduce the Galerkin projection $\tilde{z}_{kh} \in X^{0,1}_{k,h}$ of $\bar{z}$ by

$$B(\varphi_{kh}, \bar{z} - \tilde{z}_{kh}) = 0 \quad \text{for all} \;\; \varphi_{kh} \in X^{0,1}_{k,h}$$

and will estimate the errors $\bar{z} - \tilde{z}_{kh}$ and $\tilde{z}_{kh} - \hat{z}_{kh}$ separately.

Step 1: *Estimate for* $\int_0^T \left(\bar{z}(t, x_0) - \tilde{z}_{kh}(t, x_0)\right)^2 \, dt$. For this error we apply directly the estimate (3.10) from Theorem 3.4 with $p = \infty$ resulting in

(4.9)

$$\begin{aligned}
&\int_0^T |(\bar{z} - \tilde{z}_{kh})(t, x_0)|^2 dt \\
&\leq c \left(\ln \frac{T}{k}\right)^2 |\ln h|^2 \inf_{\chi \in X^{0,1}_{k,h}} \Bigg\{ \|\bar{z} - \chi\|^2_{L^2(I; L^\infty(B_d(x_0)))} + \|\pi_k \bar{z} - \chi\|^2_{L^2(I; L^\infty(B_d(x_0)))} \\
&\qquad + d^{-3} \Big(\|\bar{z} - \chi\|^2_{L^2(I; L^2(\Omega))} + \|\pi_k \bar{z} - \chi\|^2_{L^2(I; L^2(\Omega))} \\
&\qquad\qquad + h^2 \|\nabla(\bar{z} - \chi)\|^2_{L^2(I; L^2(\Omega))} \Big) \Bigg\}.
\end{aligned}$$

We choose $\chi = P_k i_h \bar{z}$, where P_k is the semidiscrete L^2 projection; see (3.4), and i_h is the nodewise interpolant. Note, that $i_h z$ is well-defined since $\bar{z} \in L^2(I; H^2(\Omega)) \hookrightarrow L^2(I; C(\bar{\Omega}))$. For the first interior term in (4.9) we obtain

$$\begin{aligned}
&\|\bar{z} - P_k i_h \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))} \\
&\quad \leq \|\bar{z} - P_k \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))} + \|P_k(\bar{z} - i_h \bar{z})\|_{L^2(I; L^\infty(B_d(x_0)))} \\
&\quad \leq \|\bar{z} - P_k \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))} + c\|\bar{z} - i_h \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))} \\
&\quad \leq \|\bar{z} - \pi_k \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))} + c\|\bar{z} - i_h \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))},
\end{aligned}$$

where we have used the stability of P_k with respect to the $L^2(I; L^\infty(B_d(x_0)))$ norm; see (3.5).

For the second interior term there holds

$$\begin{aligned}
&\|\pi_k \bar{z} - P_k i_h \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))} \\
&\quad \leq \|P_k(\pi_k \bar{z} - \bar{z})\|_{L^2(I; L^\infty(B_d(x_0)))} + \|P_k(\bar{z} - i_h \bar{z})\|_{L^2(I; L^\infty(B_d(x_0)))} \\
&\quad \leq c\|\bar{z} - \pi_k \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))} + c\|\bar{z} - i_h \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))}.
\end{aligned}$$

Therefore both interior terms in (4.9) can be estimated as

$$\begin{aligned}
(4.10) \quad &\|\bar{z} - \chi\|^2_{L^2(I; L^\infty(B_d(x_0)))} + \|\pi_k \bar{z} - \chi\|^2_{L^2(I; L^\infty(B_d(x_0)))} \\
&\qquad \leq c\|\bar{z} - \pi_k \bar{z}\|^2_{L^2(I; L^\infty(B_d(x_0)))} + c\|\bar{z} - i_h \bar{z}\|^2_{L^2(I; L^\infty(B_d(x_0)))}.
\end{aligned}$$

The first term on the right-hand side of (4.10) is estimated by

$$\|\bar{z} - \pi_k \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))} \leq c k^\gamma \|\bar{z}\|_{C^\gamma(\bar{I}; C(\overline{B}_d))}$$

using the regularity result (e) from Theorem 2.9. We choose $\gamma = \frac{1}{2} - c_0 |\ln k|^{-1}$ and obtain by the estimate (e) from Theorem 2.9

$$(4.11) \qquad \|\bar{z} - \pi_k \bar{z}\|_{L^2(I; L^\infty(B_d(x_0)))} \leq c_d k^{\frac{1}{2}} \left(\ln \frac{T}{k}\right)^5 \alpha^{-\frac{5}{2}} \|\hat{u}\|_{L^\infty(I; L^3(\Omega))}.$$

For the second term on the right-hand side of (4.10) we obtain

$$\|\bar z - i_h\bar z\|_{L^2(I;L^\infty(B_d(x_0)))} \le ch^{2-\frac{3}{p}}\|\bar z\|_{L^2(I;W^{2,p}(B_d(x_0)))}.$$

We choose $p = 3-|\ln h|^{-1}$ and $r=2$ in the estimate (d) from Theorem 2.9 and obtain

$$\|\bar z - i_h\bar z\|_{L^2(I;L^\infty(B_d(x_0)))} \le c_d h|\ln h|\alpha^{-\frac{5}{2}}\|\widehat u\|_{L^\infty(I;L^3(\Omega))}. \tag{4.12}$$

Inserting (4.11) and (4.12) into (4.10) we get

$$\begin{aligned}\|\bar z - \chi\|^2_{L^2(I;L^\infty(B_d(x_0)))} &+ \|\pi_k\bar z - \chi\|^2_{L^2(I;L^\infty(B_d(x_0)))}\\ &\le c_d\alpha^{-5}\left(k\left(\ln\frac{T}{k}\right)^{10} + h^2|\ln h|^2\right)\|\widehat u\|^2_{L^\infty(I;L^3(\Omega))}.\end{aligned} \tag{4.13}$$

Next we consider the global terms from (4.9). Arguing as before and exploiting the stability of P_k with respect to the $L^2(I;L^2(\Omega))$ norm (see (3.5)) we obtain

$$\begin{aligned}\|\bar z - \chi\|^2_{L^2(I;L^2(\Omega))} &+ \|\pi_k\bar z - \chi\|^2_{L^2(I;L^2(\Omega))} + h^2\|\nabla(\bar z - \chi)\|^2_{L^2(I;L^2(\Omega))}\\ &\le c\|\bar z - \pi_k\bar z\|^2_{L^2(I;L^2(\Omega))} + c\|\bar z - i_h\bar z\|^2_{L^2(I;L^2(\Omega))}\\ &\quad + ch^2\|\nabla(\bar z - i_h\bar z)\|_{L^2(I;L^2(\Omega))} + ch^2\|\nabla(\bar z - \pi_k\bar z)\|_{L^2(I;L^2(\Omega))}.\end{aligned} \tag{4.14}$$

Using that $\bar z \in L^2(I;H^2(\Omega)\cap H^1_0(\Omega))\cap H^1(I;L^2(\Omega))$ we apply standard estimates for the first three terms on the right-hand side of (4.14) and the estimate from [28, Lemma 3.13] for the last term resulting in

$$\begin{aligned}\|\bar z - \chi\|^2_{L^2(I;L^2(\Omega))} &+ \|\pi_k\bar z - \chi\|^2_{L^2(I;L^2(\Omega))} + h^2\|\nabla(\bar z - \chi)\|^2_{L^2(I;L^2(\Omega))}\\ &\le c(k^2+h^4+h^2k)\left(\|\nabla^2\bar z\|^2_{L^2(I;L^2(\Omega))} + \|\bar z_t\|^2_{L^2(I;L^2(\Omega))}\right)\\ &\le c(k^2+h^4)\|\widehat u\|^2_{L^2(I;L^2(\Omega))},\end{aligned} \tag{4.15}$$

where in the last step we have used Lemma 2.8. Inserting (4.13) and (4.15) into (4.9) we observe that the interior terms dominate and obtain

$$\begin{aligned}\int_0^T &|(\bar z - \tilde z_{kh})(t,x_0)|^2 dt\\ &\le c_d c_\alpha^2\left(\ln\frac{T}{k}\right)^2|\ln h|^2\left(k\left(\ln\frac{T}{k}\right)^{10} + h^2|\ln h|^2\right)\|\widehat u\|^2_{L^\infty(I;L^3(\Omega))}.\end{aligned} \tag{4.16}$$

Step 2: *Estimate for* $\int_0^T(\tilde z_{kh}(t,x_0) - \hat z_{kh}(t,x_0))^2\,dt$. We denote by $w_{kh} = \tilde z_{kh} - \hat z_{kh}$. By construction $w_{kh}\in X^{0,1}_{k,h}$ fulfills

$$B(\varphi_{kh}, w_{kh}) = (\bar u - \hat u_{kh}, \varphi_{kh}) \quad \text{for all } \ \varphi_{kh}\in X^{0,1}_{k,h}. \tag{4.17}$$

Using the discrete elliptic result from Lemma 5.3 below, we obtain

$$\int_0^T w_{kh}(t,x_0)^2\,dt \le c|\ln h|^{\frac{2}{3}}\|\Delta_h w_{kh}\|^2_{L^2(I;L^{\frac{3}{2}}(\Omega))}.$$

By the discrete maximal parabolic regularity result from Theorem 3.2 applied to (4.17) we have

$$\|\Delta_h w_{kh}\|_{L^2(I;L^{\frac{3}{2}}(\Omega))} \le c \ln \frac{T}{k} \|\bar{u} - \hat{u}_{kh}\|_{L^2(I;L^{\frac{3}{2}}(\Omega))}$$

and therefore

$$\int_0^T w_{kh}(t,x_0)^2\, dt \le c \left(\ln \frac{T}{k}\right)^2 |\ln h|^{\frac{2}{3}} \|\bar{u} - \hat{u}_{kh}\|^2_{L^2(I;L^{\frac{3}{2}}(\Omega))}.$$

Applying Theorem 4.1 we obtain

(4.18)
$$\int_0^T w_{kh}(t,x_0)^2\, dt \le c_d \alpha^{-1} \left(\ln \frac{T}{k}\right)^8 |\ln h|^{\frac{2}{3}} \left(\alpha^{-2} k^{\frac{1}{2}} + h|\ln h|\right)^2 \|\widehat{u}\|^2_{L^2(I;L^3(\Omega))}.$$

Combining the estimates (4.16) and (4.18) and choosing dominated terms we finally obtain

$$\left(\int_0^T \left(\bar{z}(t,x_0) - \hat{z}_{kh}(t,x_0)\right)^2\, dt\right)^{\frac{1}{2}} \le c_d c_\alpha \left(\ln \frac{T}{k}\right)^6 |\ln h|^2 \left(k^{\frac{1}{2}} + h\right) \|\widehat{u}\|^2_{L^\infty(I;L^3(\Omega))}.$$

This completes the proof. ∎

In the next theorem, we prove our main result for the optimal control problem under consideration.

THEOREM 4.3. *Let $\bar{q}$ be the solution of the optimal control problem* (2.10) *and $\bar{q}_{kh}$ be the solution of the discrete optimal discrete control* (4.2)*. Then there exists a constant c_d that depends only on the radius $d > 4h$ of the largest ball centered at x_0 that is contained in Ω, such that*

$$\|\bar{q} - \bar{q}_{kh}\|_I \le c_d \alpha^{-1} c_\alpha \left(\ln \frac{T}{k}\right)^6 |\ln h|^2 \left(k^{\frac{1}{2}} + h\right) \|\widehat{u}\|_{L^\infty(I;L^3(\Omega))}$$

with $c_\alpha = \max(\alpha^{-\frac{5}{2}}, 1)$.

Proof. Due to the quadratic structure of discrete reduced functional j_{kh}, the second derivative $j''_{kh}(q)(\cdot,\cdot)$ is independent of q and there holds

(4.19)
$$j''_{kh}(q)(p,p) \ge \alpha \|p\|^2_I \quad \text{for all } \ p \in Q.$$

Using the optimality conditions (2.12) for $\bar{q}$ and (4.3) for $\bar{q}_{kh}$ and the fact that $\bar{q}, \bar{q}_{kh} \in Q_{\text{ad}}$, we obtain

$$-j'_{kh}(\bar{q}_{kh})(\bar{q} - \bar{q}_{kh}) \le 0 \le -j'(\bar{q})(\bar{q} - \bar{q}_{kh}).$$

Using the coercivity (4.19), the representations (2.13) and (4.4) for the derivatives of j and j', we get, with $\hat{z}_{kh} = z_{kh}(\bar{q})$,

$$\begin{aligned}
\alpha\|\bar{q} - \bar{q}_{kh}\|^2_I &\le j''_{kh}(\bar{q})(\bar{q} - \bar{q}_{kh}, \bar{q} - \bar{q}_{kh}) = j'_{kh}(\bar{q})(\bar{q} - \bar{q}_{kh}) - j'_{kh}(\bar{q}_{kh})(\bar{q} - \bar{q}_{kh}) \\
&\le j'_{kh}(\bar{q})(\bar{q} - \bar{q}_{kh}) - j'(\bar{q})(\bar{q} - \bar{q}_{kh}) \\
&= -(z(\bar{q})(t,x_0) - z_{kh}(\bar{q})(t,x_0), \bar{q} - \bar{q}_{kh})_I \\
&= -(\bar{z}(t,x_0) - \hat{z}_{kh}(t,x_0), \bar{q} - \bar{q}_{kh})_I \\
&\le \left(\int_0^T |\bar{z}(t,x_0) - \hat{z}_{kh}(t,x_0)|^2\, dt\right)^{\frac{1}{2}} \|\bar{q} - \bar{q}_{kh}\|_I.
\end{aligned}$$

Applying Theorem 4.2 completes the proof. ∎

5. Elliptic estimates in weighted norms. In this section we collect some estimates for the finite element discretization of elliptic problems in weighted norms on convex polyhedral domains mainly taken from [32] . These results will be used in the following sections for the proofs of Theorems 3.3 and 3.4.

In what follows, we consider a fixed (but arbitrary) point $y \in \Omega$. Associated with this point, we introduce a smoothed Dirac delta function [42, Appendix], which we will denote by $\tilde{\delta} = \tilde{\delta}_y = \tilde{\delta}_y^h$. This function is supported in one cell, which is denoted by τ_y, and satisfies

$$(\chi, \tilde{\delta})_{\tau_y} = \chi(y) \quad \forall \chi \in \mathbb{P}^1(\tau_y).$$

In addition we also have

$$\|\tilde{\delta}\|_{W^{s,p}(\Omega)} \le ch^{-s-3(1-\frac{1}{p})}, \quad 1 \le p \le \infty, \quad s = 0, 1. \tag{5.1}$$

Thus, in particular, $\|\tilde{\delta}\|_{L^1(\Omega)} \le c$, $\|\tilde{\delta}\|_{L^2(\Omega)} \le ch^{-\frac{3}{2}}$, and $\|\tilde{\delta}\|_{L^\infty(\Omega)} \le ch^{-3}$. Next, we introduce the weight function

$$\sigma(x) = \sqrt{|x - y|^2 + K^2 h^2}, \tag{5.2}$$

where $K > 0$ is a sufficiently large constant to be chosen later. One can easily check that σ satisfies the following properties:

$$\|\sigma^{-\frac{3}{2}}\|_{L^2(\Omega)} \le c|\ln h|^{\frac{1}{2}}, \tag{5.3a}$$

$$|\nabla \sigma| \le c, \tag{5.3b}$$

$$|\nabla^2 \sigma| \le c|\sigma^{-1}|, \tag{5.3c}$$

$$\max_{x\in\tau} \sigma(x) \le c \min_{x\in\tau} \sigma(x) \quad \text{for all cells } \tau. \tag{5.3d}$$

For the finite element space V_h we will utilize the L^2 projection $P_h \colon L^2(\Omega) \to V_h$ defined by

$$(P_h v, \chi)_\Omega = (v, \chi)_\Omega \quad \forall \chi \in V_h, \tag{5.4}$$

the Ritz projection $R_h \colon H_0^1(\Omega) \to V_h$ defined by

$$(\nabla R_h v, \nabla \chi)_\Omega = (\nabla v, \nabla \chi)_\Omega \quad \forall \chi \in V_h, \tag{5.5}$$

and the usual nodal interpolant $i_h \colon C_0(\Omega) \to V_h$. Moreover we introduce the discrete Laplace operator $\Delta_h \colon V_h \to V_h$ by

$$(-\Delta_h v_h, \chi)_\Omega = (\nabla v_h, \nabla \chi)_\Omega \quad \forall \chi \in V_h. \tag{5.6}$$

The following lemma is a superapproximation result in weighted norms.

LEMMA 5.1 (see [32, Lemma 3]). *Let $v_h \in V_h$. Then, the following estimates hold for any $\alpha, \beta \in \mathbb{R}$ and K large enough:*

$$\|\sigma^\alpha (\mathrm{Id} - i_h)(\sigma^\beta v_h)\|_{L^2(\Omega)} + h\|\sigma^\alpha \nabla (\mathrm{Id} - i_h)(\sigma^\beta v_h)\|_{L^2(\Omega)} \le ch\|\sigma^{\alpha+\beta-1} v_h\|_{L^2(\Omega)}, \tag{5.7}$$

$$\|\sigma^\alpha (\mathrm{Id} - P_h)(\sigma^\beta v_h)\|_{L^2(\Omega)} + h\|\sigma^\alpha \nabla (\mathrm{Id} - P_h)(\sigma^\beta v_h)\|_{L^2(\Omega)} \le ch\|\sigma^{\alpha+\beta-1} v_h\|_{L^2(\Omega)}. \tag{5.8}$$

The next lemma describes a connection between the regularized Dirac delta function $\tilde{\delta}$ and the weight σ.

LEMMA 5.2 (see [32, Lemma 4]). *There holds*

$$\|\sigma^{\frac{3}{2}}\tilde{\delta}\|_{L^2(\Omega)} + h\|\sigma^{\frac{3}{2}}\nabla\tilde{\delta}\|_{L^2(\Omega)} + \|\sigma^{\frac{3}{2}}P_h\tilde{\delta}\|_{L^2(\Omega)} \le c. \tag{5.9}$$

On the continuous level, the solution of the Poisson equation with a right-hand side in $L^{\frac{3}{2}}(\Omega)$ is, in general, not in $L^\infty(\Omega)$. Similarly, the solution of the Poisson equation with a right-hand side in $L^1(\Omega)$ is, in general, not in $L^3(\Omega)$. The next two lemmas provide the corresponding a priori estimates on the discrete level with logarithmic terms.

LEMMA 5.3. *There is a constant $c > 0$ such that*

$$\|w_h\|_{L^\infty(\Omega)} \le c|\ln h|^{\frac{1}{3}}\|\Delta_h w_h\|_{L^{\frac{3}{2}}(\Omega)}$$

for all $w_h \in V_h$.

Proof. For a given point $y \in \Omega$ we consider the discrete Green's function $g_h \in V_h$ defined by

$$-\Delta_h g_h = P_h\tilde{\delta}$$

with $\tilde{\delta}$ being the regularized Dirac delta function defined above. In [32] we have shown (in the proof of Theorem 12) that

$$\|g_h\|_{L^3(\Omega)} \le c|\ln h|^{\frac{1}{3}}.$$

Using this estimate, we obtain for an arbitrary $w_h \in V_h$

$$\begin{aligned} w_h(y) &= (\nabla w_h, \nabla g_h) \\ &= (-\Delta_h w_h, g_h) \le \|\Delta_h w_h\|_{L^{\frac{3}{2}}(\Omega)}\|g_h\|_{L^3(\Omega)} \le c|\ln h|^{\frac{1}{3}}\|\Delta_h w_h\|_{L^{\frac{3}{2}}(\Omega)} \end{aligned}$$

and thus

$$\|w_h\|_{L^\infty(\Omega)} \le c|\ln h|^{\frac{1}{3}}\|\Delta_h w_h\|_{L^{\frac{3}{2}}(\Omega)}. \tag{5.10}$$ ∎

LEMMA 5.4. *There is a constant $c > 0$ such that*

$$\|v_h\|_{L^3(\Omega)} \le c|\ln h|^{\frac{1}{3}}\|\Delta_h v_h\|_{L^1(\Omega)}$$

for all $v_h \in V_h$.

Proof. Let $v_h \in V_h$ be arbitrary. We consider $w_h \in V_h$ solving

$$-\Delta_h w_h = P_h(v_h|v_h|)$$

and obtain using the previous lemma

$$\begin{aligned} \|v_h\|^3_{L^3(\Omega)} &= (v_h, v_h|v_h|) = (\nabla v_h, \nabla w_h) = (-\Delta_h v_h, w_h) \\ &\le \|\Delta_h v_h\|_{L^1(\Omega)}\|w_h\|_{L^\infty(\Omega)} \le c|\ln h|^{\frac{1}{3}}\|\Delta_h v_h\|_{L^1(\Omega)}\|v_h|v_h|\|_{L^{\frac{3}{2}}(\Omega)} \\ &= c|\ln h|^{\frac{1}{3}}\|\Delta_h v_h\|_{L^1(\Omega)}\|v_h\|^2_{L^3(\Omega)}. \end{aligned}$$ ∎

The next lemma is a three-dimensional version of Lemma 2.4 in [40], that says that the L^1 norm of jumps across element boundaries of any piecewise linear function can be controlled by the properly weighted discrete H^2 norm.

LEMMA 5.5 (see [32, Lemma 6]). *There exists a constant $c > 0$ independent of h, such that for any $v_h \in V_h$,*

$$\sum_{\tau\in\mathcal{T}} \|[\![\partial_n v_h]\!]\|_{L^1(\partial\tau)} \le c|\ln h|^{\frac{1}{2}} \left(\|\sigma^{\frac{3}{2}}\Delta_h v_h\|_{L^2(\Omega)} + \|\sigma^{\frac{1}{2}}\nabla v_h\|_{L^2(\Omega)}\right),$$

where $[\![\partial_n v_h]\!]$ denotes the jump of the normal derivatives of v_h across the boundary of τ.

6. Proof of Theorem 3.3. Let $v_{kh} \in X^{0,1}_{k,h}$ be the solution of (3.9) and $y \in \Omega$ be an arbitrary but fixed point. To establish the first estimate from Theorem 3.3, it is sufficient to establish

$$\int_0^T |v_{kh}(t,y)|^2 dt \le c\left(\ln\frac{T}{k}\right)^2 |\ln h|^2 \left(\|v\|^2_{L^2(I;L^\infty(\Omega))} + h^{-\frac{6}{p}}\|\pi_k v\|^2_{L^2(I;L^p(\Omega))}\right) \tag{6.1}$$

for some constant c independent of h, k, and y. Then using that the cG(1)dG(0) method is invariant on $X^{0,1}_{k,h}$, by replacing v and v_{kh} with $v-\chi$ and $v_{kh}-\chi$ for any $\chi \in X^{0,1}_{k,h}$, by taking the supremum over y, and using the triangle inequality we will obtain Theorem 3.3. In order for this argument to be valid, we have to be careful and make sure that only norms of v are used that can be applied to functions from $X^{0,1}_{k,h}$.

To obtain (6.1) we use a duality argument. To this end we define g to be the solution to the following backward parabolic problem

$$\begin{aligned} -g_t(t,x) - \Delta g(t,x) &= v_{kh}(t,y)\tilde{\delta}_y(x), && (t,x) \in I\times\Omega,\\ g(t,x) &= 0, && (t,x)\in I\times\partial\Omega,\\ g(T,x) &= 0, && x\in\Omega, \end{aligned} \tag{6.2}$$

where $\tilde{\delta}_y$ is the smoothed Dirac delta introduced in section 5. Let $g_{kh} \in X^{0,1}_{k,h}$ be the corresponding cG(1)dG(0) solution defined by

$$B(\varphi_{kh}, g_{kh}) = (v_{kh}(t,y)\tilde{\delta}_y, \varphi_{kh})_{I\times\Omega}, \quad \forall \varphi_{kh}\in X^{0,1}_{k,h}. \tag{6.3}$$

Then, using that the cG(1)dG(0) method is consistent, we have

$$\int_0^T |v_{kh}(t,y)|^2 dt = B(v_{kh}, g_{kh}) = B(v, g_{kh}) = (\nabla v, \nabla g_{kh})_{I\times\Omega} - \sum_{m=1}^M (v_m, [g_{kh}]_m)_\Omega, \tag{6.4}$$

where we have used the dual expression (3.7) for the bilinear form $B(\cdot,\cdot)$ and the fact that the last term in (3.7) can be included in the sum by setting $g_{kh,M+1} = 0$ and defining consequently $[g_{kh}]_M = -g_{kh,M}$. The first sum in (3.7) vanishes due to $g_{kh} \in X^{0,1}_{k,h}$. For each t, integrating by parts elementwise and using that g_{kh} is linear in the space variable, by the Hölder's inequality we have

$$(\nabla v(t), \nabla g_{kh}(t))_\Omega = \frac{1}{2}\sum_\tau (v(t), [\![\partial_n g_{kh}(t)]\!])_{\partial\tau} \le c\|v(t)\|_{L^\infty(\Omega)} \sum_\tau \|[\![\partial_n g_{kh}(t)]\!]\|_{L^1(\partial\tau)}, \tag{6.5}$$

where $[\![\partial_n g_{kh}(t)]\!]$ denotes the jumps of the normal derivatives across the element faces. By Lemma 5.5, we obtain

$$(\nabla v(t), \nabla g_{kh}(t))_\Omega \le c|\ln h|^{\frac{1}{2}}\|v(t)\|_{L^\infty(\Omega)} \left(\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}(t)\|_{L^2(\Omega)} + \|\sigma^{\frac{1}{2}}\nabla g_{kh}(t)\|_{L^2(\Omega)}\right).$$

Integrating in time and using the Cauchy–Schwarz inequality we get

$$
\begin{aligned}
(6.6)\quad &(\nabla v, \nabla g_{kh})_{I\times\Omega} \\
&\le c|\ln h|^{\frac12}\|v\|_{L^2(I;L^\infty(\Omega))}\left(\int_0^T \left(\|\sigma^{\frac32}\Delta_h g_{kh}\|^2_{L^2(\Omega)} + \|\sigma^{\frac12}\nabla g_{kh}\|^2_{L^2(\Omega)}\right)dt\right)^{\frac12}.
\end{aligned}
$$

To estimate the term involving the jumps in (6.4), we first use the Hölder's inequality and the inverse estimate for some $1 \le p \le \infty$ to obtain

$$
\begin{aligned}
\sum_{m=1}^M (v_m, [g_{kh}]_m)_\Omega &\le c\sum_{m=1}^M k_m^{\frac12}\|v_m\|_{L^p(\Omega)} k_m^{-\frac12} h^{-\frac3p}\|[g_{kh}]_m\|_{L^1(\Omega)} \\
&\le ch^{-\frac3p}\left(\sum_{m=1}^M k_m\|v_m\|^2_{L^p(\Omega)}\right)^{\frac12}\left(\sum_{m=1}^M k_m^{-1}\|[g_{kh}]_m\|^2_{L^1(\Omega)}\right)^{\frac12}.
\end{aligned}
$$

Inserting the weight function σ and using (5.3a), we have

$$
\|[g_{kh}]_m\|_{L^1(\Omega)} \le \|\sigma^{-\frac32}\|_{L^2(\Omega)}\|\sigma^{\frac32}[g_{kh}]_m\|_{L^2(\Omega)} \le c|\ln h|^{\frac12}\|\sigma^{\frac32}[g_{kh}]_m\|_{L^2(\Omega)}.
$$

With the semidiscrete interpolant π_k defined in (3.3), we obtain

(6.7)

$$
\sum_{m=1}^M (v_m, [g_{kh}]_m)_\Omega \le ch^{-\frac3p}|\ln h|^{\frac12}\|\pi_k v\|_{L^2(I;L^p(\Omega))}\left(\sum_{m=1}^M k_m^{-1}\|\sigma^{\frac32}[g_{kh}]_m\|^2_{L^2(\Omega)}\right)^{\frac12}.
$$

Inserting (6.6) and (6.7) in (6.4) we obtain

(6.8)

$$
\begin{aligned}
&\int_0^T |v_{kh}(t,y)|^2 dt \\
&\quad\le c|\ln h|^{\frac12}\left(\|v\|^2_{L^2(I;L^\infty(\Omega))} + h^{-\frac6p}\|\pi_k v\|^2_{L^2(I;L^p(\Omega))}\right)^{\frac12} \\
&\quad\quad\times\left(\int_0^T\left(\|\sigma^{\frac32}\Delta_h g_{kh}\|^2_{L^2(\Omega)} + \|\sigma^{\frac12}\nabla g_{kh}\|^2_{L^2(\Omega)}\right)dt + \sum_{m=1}^M k_m^{-1}\|\sigma^{\frac32}[g_{kh}]_m\|^2_{L^2(\Omega)}\right)^{\frac12}.
\end{aligned}
$$

To complete the proof of the theorem we need to show that

$$
\begin{aligned}
(6.9)\quad &\int_0^T\left(\|\sigma^{\frac32}\Delta_h g_{kh}\|^2_{L^2(\Omega)} + \|\sigma^{\frac12}\nabla g_{kh}\|^2_{L^2(\Omega)}\right)dt + \sum_{m=1}^M k_m^{-1}\|\sigma^{\frac32}[g_{kh}]_m\|^2_{L^2(\Omega)} \\
&\quad\le c\left(\ln\frac{T}{k}\right)^2|\ln h|\int_0^T |v_{kh}(t,y)|^2 dt.
\end{aligned}
$$

The above result will follow from a series of lemmas. The first lemma treats the term $\|\sigma^{\frac32}\Delta_h g_{kh}\|^2_{L^2(I;L^2(\Omega))}$.

LEMMA 6.1. *For any $\varepsilon > 0$ there exists c_ε such that*

$$
\begin{aligned}
\int_0^T \|\sigma^{\frac32}\Delta_h g_{kh}\|^2_{L^2(\Omega)} dt &\le c_\varepsilon\int_0^T\left(|v_{kh}(t,y)|^2 + \|\sigma^{\frac12}\nabla g_{kh}\|^2_{L^2(\Omega)}\right)dt \\
&\quad + \varepsilon\sum_{m=1}^M k_m^{-1}\|\sigma^{\frac32}[g_{kh}]_m\|^2_{L^2(\Omega)}.
\end{aligned}
$$

Proof. We use the fact that (6.3) can be rewritten on each time level as

$$(\nabla\varphi_{kh},\nabla g_{kh})_{I_m\times\Omega}-(\varphi_{kh,m},[g_{kh}]_m)_\Omega=(v_{kh}(t,y)\tilde{\delta}_y,\varphi_{kh})_{I_m\times\Omega},$$

or, equivalently, as

$$-k_m\Delta_h g_{kh,m}-[g_{kh}]_m=k_m v_{kh,m}(y)P_h\tilde{\delta}_y. \tag{6.10}$$

We multiply this equation by $\varphi=-\sigma^3\Delta_h g_{kh}$ and obtain

$$\begin{aligned}\int_{I_m}\|\sigma^{\frac32}\Delta_h g_{kh}\|^2_{L^2(\Omega)}dt&=-([g_{kh}]_m,\sigma^3\Delta_h g_{kh,m})_\Omega-(v_{kh}(t,y)P_h\tilde{\delta}_y,\sigma^3\Delta_h g_{kh})_{I_m\times\Omega}\\&=-([\sigma^3 g_{kh}]_m,\Delta_h g_{kh,m})_\Omega-(v_{kh}(t,y)P_h\tilde{\delta}_y,\sigma^3\Delta_h g_{kh})_{I_m\times\Omega}\\&=([\nabla(\sigma^3 g_{kh})]_m,\nabla g_{kh,m})_\Omega+([\nabla(P_h-\mathrm{Id})\sigma^3 g_{kh}]_m,\nabla g_{kh,m})_\Omega\\&\quad-(v_{kh}(t,y)P_h\tilde{\delta}_y,\sigma^3\Delta_h g_{kh})_{I_m\times\Omega}=J_1+J_2+J_3.\end{aligned}$$

We have

$$J_1=3(\sigma^2\nabla\sigma[g_{kh}]_m,\nabla g_{kh,m})_\Omega+(\sigma^{\frac32}[\nabla g_{kh}]_m,\sigma^{\frac32}\nabla g_{kh,m})_\Omega=J_{11}+J_{12}.$$

By the Cauchy–Schwarz inequality and using (5.3b) we get

$$J_{11}\le c\|\sigma^{\frac32}[g_{kh}]_m\|_{L^2(\Omega)}\|\sigma^{\frac12}\nabla g_{kh,m}\|_{L^2(\Omega)}.$$

Using the identity

$$([w_{kh}]_m,w_{kh,m})_\Omega=\frac12\|w_{kh,m+1}\|^2_{L^2(\Omega)}-\frac12\|w_{kh,m}\|^2_{L^2(\Omega)}-\frac12\|[w_{kh}]_m\|^2_{L^2(\Omega)}, \tag{6.11}$$

we have

$$J_{12}=\frac12\|\sigma^{\frac32}\nabla g_{kh,m+1}\|^2_{L^2(\Omega)}-\frac12\|\sigma^{\frac32}\nabla g_{kh,m}\|^2_{L^2(\Omega)}-\frac12\|\sigma^{\frac32}[\nabla g_{kh}]_m\|^2_{L^2(\Omega)}.$$

Using Young's inequality for J_{11} and neglecting $-\frac12\|\sigma^{\frac32}[\nabla g_{kh}]_m\|^2_{L^2(\Omega)}$ in J_{12} we obtain

$$\begin{aligned}J_1&\le\frac12\|\sigma^{\frac32}\nabla g_{kh,m+1}\|^2_{L^2(\Omega)}-\frac12\|\sigma^{\frac32}\nabla g_{kh,m}\|^2_{L^2(\Omega)}\\&\quad+c_\varepsilon k_m\|\sigma^{\frac12}\nabla g_{kh,m}\|^2_{L^2(\Omega)}+\frac{\varepsilon}{4k_m}\|\sigma^{\frac32}[g_{kh}]_m\|^2_{L^2(\Omega)}.\end{aligned} \tag{6.12}$$

For J_2 we get, by the superapproximation estimate (5.8) from Lemma 5.1,

$$\begin{aligned}J_2&\le\|\sigma^{-\frac12}[\nabla(P_h-\mathrm{Id})\sigma^3 g_{kh}]_m\|_{L^2(\Omega)}\|\sigma^{\frac12}\nabla g_{kh,m}\|_{L^2(\Omega)}\\&\le\|\sigma^{\frac32}[g_{kh}]_m\|_{L^2(\Omega)}\|\sigma^{\frac12}\nabla g_{kh,m}\|_{L^2(\Omega)}\\&\le c_\varepsilon k_m\|\sigma^{\frac12}\nabla g_{kh,m}\|^2_{L^2(\Omega)}+\frac{\varepsilon}{4k_m}\|\sigma^{\frac32}[g_{kh}]_m\|^2_{L^2(\Omega)}.\end{aligned} \tag{6.13}$$

To estimate J_3 we use Lemma 5.2, which states that $\|\sigma^{\frac32}P_h\tilde{\delta}\|_{L^2(\Omega)}\le c$ and obtain

$$\begin{aligned}J_3&\le\int_{I_m}|v_{kh}(t,y)|\,\|\sigma^{\frac32}P_h\tilde{\delta}\|_{L^2(\Omega)}\|\sigma^{\frac32}\Delta_h g_{kh}\|_{L^2(\Omega)}\,dt\\&\le c\int_{I_m}|v_{kh}(t,y)|^2dt+\frac12\int_{I_m}\|\sigma^{\frac32}\Delta_h g_{kh}\|^2_{L^2(\Omega)}\,dt.\end{aligned} \tag{6.14}$$

Using the estimates (6.12), (6.13), and (6.14) we have

$$\begin{aligned}\int_{I_m} \|\sigma^{\frac{3}{2}}\Delta_h g_{kh}\|^2_{L^2(\Omega)} dt \le c_\varepsilon \int_{I_m} \left(|v_{kh}(t,y)|^2 + \|\sigma^{\frac{1}{2}}\nabla g_{kh}\|^2_{L^2(\Omega)}\right) dt \\ + \frac{\varepsilon}{k_m}\|\sigma^{\frac{3}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)} + \|\sigma^{\frac{3}{2}}\nabla g_{kh,m+1}\|^2_{L^2(\Omega)} \\ - \|\sigma^{\frac{3}{2}}\nabla g_{kh,m}\|^2_{L^2(\Omega)}.\end{aligned}$$

Summing over m and using that $g_{kh,M+1} = 0$ we obtain the lemma. ∎

The next lemma treats the term involving the jumps.

LEMMA 6.2. *There exists a constant c such that*

$$\sum_{m=1}^{M} k_m^{-1}\|\sigma^{\frac{3}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)} \le c\int_0^T \left(\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}\|^2_{L^2(\Omega)} + |v_{kh}(t,y)|^2\right) dt.$$

Proof. We test (6.10) with $\varphi = \sigma^3[g_{kh}]_m$ and obtain

$$\|\sigma^{\frac{3}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)} = -(\Delta_h g_{kh}, \sigma^3[g_{kh}]_m)_{I_m\times\Omega} - (v_{kh}(t,y)P_h\tilde{\delta}, \sigma^3[g_{kh}]_m)_{I_m\times\Omega}. \tag{6.15}$$

The first term on the right-hand side of (6.15) can be easily estimated by using Young's inequality as

$$-(\Delta_h g_{kh}, \sigma^3[g_{kh}]_m)_{I_m\times\Omega} \le ck_m\int_{I_m}\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}\|^2_{L^2(\Omega)} dt + \frac{1}{4}\|\sigma^{\frac{3}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)}.$$

For the last term on the right-hand side of (6.15) we use again the inequality (5.9) from Lemma 5.2, which states that $\|\sigma^{\frac{3}{2}}P_h\tilde{\delta}\|_{L^2(\Omega)} \le c$ and obtain

$$(v_{kh}(t,y)P_h\tilde{\delta}, \sigma^3[g_{kh}]_m)_{I_m\times\Omega} \le ck_m\int_{I_m}|v_{kh}(t,y)|^2 dt + \frac{1}{4}\|\sigma^{\frac{3}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)}.$$

Combining the above two estimates we obtain

$$\|\sigma^{\frac{3}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)} \le ck_m\int_{I_m}\left(\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}\|^2_{L^2(\Omega)} + |v_{kh}(t,y)|^2\right) dt.$$

Summing over m we obtain the lemma. ∎

In the next lemma we treat the term $\|\sigma^{\frac{1}{2}}\nabla g_{kh}\|_{L^2(I;L^2(\Omega))}$.

LEMMA 6.3. *There exists a constant c such that*

$$\int_0^T \|\sigma^{\frac{1}{2}}\nabla g_{kh}\|^2_{L^2(\Omega)} dt \le c\int_0^T \left(\|\sigma^{-\frac{1}{2}} g_{kh}\|^2_{L^2(\Omega)} + |v_{kh}(t,y)|^2\right) dt.$$

Proof. We test (6.10) with $\varphi = \sigma g_{kh}$ and obtain

$$-k_m(\Delta_h g_{kh,m}, \sigma g_{kh,m}) - ([g_{kh}]_m, \sigma g_{kh,m}) = k_m v_{kh,m}(y)(P_h\tilde{\delta}_y, \sigma g_{kh,m}).$$

There holds

$$\begin{aligned}&- k_m(\Delta_h g_{kh,m}, \sigma g_{kh,m}) \\ &\quad = -k_m(\Delta_h g_{kh,m}, P_h(\sigma g_{kh,m})) = k_m(\nabla g_{kh,m}, \nabla P_h(\sigma g_{kh,m})) \\ &\quad = k_m(\nabla g_{kh,m}, \nabla(\sigma g_{kh,m})) + k_m(\nabla g_{kh,m}, \nabla(P_h - \mathrm{Id})(\sigma g_{kh,m})) \\ &\quad = k_m\|\sigma^{\frac{1}{2}}\nabla g_{kh,m}\|^2_{L^2(\Omega)} + k_m(\nabla g_{kh,m}, g_{kh,m}\nabla\sigma) \\ &\qquad + k_m(\nabla g_{kh,m}, \nabla(P_h - \mathrm{Id})(\sigma g_{kh,m})).\end{aligned}$$

Therefore, we obtain

$$\begin{aligned} k_m\|\sigma^{\frac{1}{2}}\nabla g_{kh,m}\|^2_{L^2(\Omega)} &= -k_m(\nabla g_{kh,m}, g_{kh,m}\nabla\sigma) - k_m(\nabla g_{kh,m}, \nabla(P_h - \mathrm{Id})(\sigma g_{kh,m})) \\ &\quad + ([g_{kh}]_m, \sigma g_{kh,m}) + k_m v_{kh,m}(y)(P_h\tilde{\delta}_y, \sigma g_{kh,m}) \\ &= J_1 + J_2 + J_3 + J_4. \end{aligned}$$

For J_1 we get by (5.3b)

$$\begin{aligned} J_1 &\le ck_m\|\sigma^{\frac{1}{2}}\nabla g_{kh,m}\|_{L^2(\Omega)}\|\sigma^{-\frac{1}{2}}g_{kh,m}\|_{L^2(\Omega)} \\ &\le \frac{1}{4}k_m\|\sigma^{\frac{1}{2}}\nabla g_{kh,m}\|^2_{L^2(\Omega)} + ck_m\|\sigma^{-\frac{1}{2}}g_{kh,m}\|^2_{L^2(\Omega)}. \end{aligned}$$

For J_2 we use the estimate (5.8) from Lemma 5.1 and obtain

$$\begin{aligned} J_2 &\le k_m\|\sigma^{\frac{1}{2}}\nabla g_{kh,m}\|_{L^2(\Omega)}\|\sigma^{-\frac{1}{2}}\nabla(P_h - \mathrm{Id})(\sigma g_{kh,m})\|_{L^2(\Omega)} \\ &\le \frac{1}{4}k_m\|\sigma^{\frac{1}{2}}\nabla g_{kh,m}\|^2_{L^2(\Omega)} + ck_m\|\sigma^{-\frac{1}{2}}g_{kh,m}\|^2_{L^2(\Omega)}. \end{aligned}$$

For J_3 we use the identity (6.11) and get

$$J_3 = \frac{1}{2}\|\sigma^{\frac{1}{2}}g_{kh,m+1}\|^2_{L^2(\Omega)} - \frac{1}{2}\|\sigma^{\frac{1}{2}}g_{kh,m}\|^2_{L^2(\Omega)} - \frac{1}{2}\|\sigma^{\frac{1}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)}.$$

For J_4 we again use the inequality (5.9) from Lemma 5.2 and obtain

$$J_4 \le ck_m v_{kh,m}(y)\|\sigma^{-\frac{1}{2}}g_{kh,m}\|_{L^2(\Omega)} \le ck_m|v_{kh,m}(y)|^2 + ck_m\|\sigma^{-\frac{1}{2}}g_{kh,m}\|^2_{L^2(\Omega)}.$$

Summing over m and neglecting the term with the jumps we obtain the desired estimate. ∎

To complete the proof of Theorem 3.3 it remains to estimate $\|\sigma^{-\frac{1}{2}}g_{kh}\|_{I\times\Omega}$.

LEMMA 6.4. *There exists a constant $c > 0$, such that*

$$\int_0^T \|\sigma^{-\frac{1}{2}}g_{kh}\|^2_{L^2(\Omega)}\,dt \le c\left(\ln\frac{T}{k}\right)^2 |\ln h| \int_0^T |v_{kh}(t,y)|^2\,dt.$$

Proof. First we observe

$$\|\sigma^{-\frac{1}{2}}g_{kh,m}\|^2_{L^2(\Omega)} \le \|\sigma^{-1}\|_{L^3(\Omega)}\,\|g^2_{kh,m}\|_{L^{\frac{3}{2}}(\Omega)} \le c|\ln h|^{\frac{1}{3}}\,\|g_{kh,m}\|^2_{L^3(\Omega)}$$

and therefore

$$\int_0^T \|\sigma^{-\frac{1}{2}}g_{kh}\|^2_{L^2(\Omega)}\,dt \le c|\ln h|^{\frac{1}{3}}\int_0^T \|g_{kh}\|^2_{L^3(\Omega)}\,dt \le c|\ln h|\int_0^T \|\Delta_h g_{kh}\|^2_{L^1(\Omega)}\,dt,$$

where in the last step we have used the estimate from Lemma 5.4 pointwise in time. Using the fully discrete maximal parabolic estimate from Theorem 3.2 with respect to the $L^2(I;L^1(\Omega))$ norm, we obtain

$$\begin{aligned} \int_0^T \|\Delta_h g_{kh}\|^2_{L^1(\Omega)}\,dt &\le c\left(\ln\frac{T}{k}\right)^2 \|\tilde{\delta}\|^2_{L^1(\Omega)}\int_0^T |v_{kh}(t,y)|^2\,dt \\ &\le c\left(\ln\frac{T}{k}\right)^2 \int_0^T |v_{kh}(t,y)|^2\,dt. \end{aligned}$$ ∎

We proceed with the proof of Theorem 3.3 starting with (6.8). Using Lemmas 6.1, 6.2, and then 6.3 we get

$$\begin{aligned}
&\int_0^T \left(\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}\|^2_{L^2(\Omega)} + \|\sigma^{\frac{1}{2}}\nabla g_{kh}\|^2_{L^2(\Omega)}\right) dt + \sum_{m=1}^M k_m^{-1}\|\sigma^{\frac{3}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)}\\
&\le c\int_0^T \left(\|\sigma^{\frac{1}{2}}\nabla g_{kh}\|^2_{L^2(\Omega)} + |v_{kh}(t,y)|^2\right) dt \qquad (6.16)\\
&\le c\int_0^T \left(\|\sigma^{-\frac{1}{2}} g_{kh}\|^2_{L^2(\Omega)} + |v_{kh}(t,y)|^2\right) dt.
\end{aligned}$$

Then we estimate using Lemma 6.4

$$\int_0^T \|\sigma^{-\frac{1}{2}} g_{kh}\|^2_{L^2(\Omega)}\, dt \le c\left(\ln\frac{T}{k}\right)^2 |\ln h| \int_0^T |v_{kh}(t,y)|^2\, dt. \qquad (6.17)$$

From (6.8) we arrive at (6.1) and this completes the proof of the first estimate in Theorem 3.3. The second estimate is established following the lines of the proof, replacing v by v_k at all places and using $\pi_k v_k = v_k$.

7. Proof of Theorem 3.4. To obtain the interior estimate we introduce a smooth cutoff function ω with the properties that

$$\omega(x) \equiv 1, \quad x \in B_{d/2}, \qquad (7.1a)$$

$$\omega(x) \equiv 0, \quad x \in \Omega \setminus B_d, \qquad (7.1b)$$

$$|\nabla\omega| \le cd^{-1}, \quad |\nabla^2\omega| \le cd^{-2}, \qquad (7.1c)$$

where $B_d = B_d(x_0)$ is a ball of radius d centered at x_0. We set $y = x_0$ in the definitions of the regularized Dirac delta function $\tilde\delta$ and of the weight σ.

As in the proof of Theorem 3.3 we obtain by (6.4) that

$$\int_0^T |v_{kh}(t,x_0)|^2 dt = B(v_{kh}, g_{kh}) = B(v, g_{kh}) = B(\omega v, g_{kh}) + B((1-\omega)v, g_{kh}), \qquad (7.2)$$

where g_{kh} is the solution of (6.3). The first term can be estimated using the global result from Theorem 3.3. To this end we introduce the solution $\tilde v_{kh} \in X^{0,1}_{k,h}$ defined by

$$B(\tilde v_{kh} - \omega v, \varphi_{kh}) = 0 \quad \text{for all } \varphi_{kh} \in X^{0,1}_{k,h}.$$

There holds

$$\begin{aligned}
B(\omega v, g_{kh}) = B(\tilde v_{kh}, g_{kh}) &= \int_0^T v_{kh}(t,x_0)\tilde v_{kh}(t,x_0)\, dt\\
&\le \frac{1}{2}\int_0^T |v_{kh}(t,x_0)|^2\, dt + \frac{1}{2}\int_0^T |\tilde v_{kh}(t,x_0)|^2\, dt.
\end{aligned}$$

Subtracting the first term from the left-hand side and applying Theorem 3.3 for the second term, we obtain

$$\begin{aligned}
\int_0^T \tilde v_{kh}(t,x_0)^2\, dt &\le c\left(\ln\frac{T}{k}\right)^2 |\ln h|^2 \left(\|\omega v\|^2_{L^2(I;L^\infty(\Omega))} + h^{-\frac{6}{p}}\|\pi_k(\omega v)\|^2_{L^2(I;L^p(\Omega))}\right)\\
&\le c\left(\ln\frac{T}{k}\right)^2 |\ln h|^2 \left(\|v\|^2_{L^2(I;L^\infty(B_d))} + h^{-\frac{6}{p}}\|\pi_k v\|^2_{L^2(I;L^p(B_d))}\right).
\end{aligned}$$

This results in

(7.3)
$$\int_0^T |v_{kh}(t,x_0)|^2 dt \le c\left(\ln\frac{T}{k}\right)^2 |\ln h|^2 \left(\|v\|^2_{L^2(I;L^\infty(B_d))} + h^{-\frac{6}{p}}\|\pi_k v\|^2_{L^2(I;L^p(B_d))}\right) + B((1-\omega)v, g_{kh}).$$

It remains to estimate the term $B((1-\omega)v, g_{kh})$. Using the dual expression (3.7) of the bilinear form B we obtain

(7.4)
$$B((1-\omega)v, g_{kh}) = (\nabla((1-\omega)v), \nabla g_{kh})_{I_m\times\Omega} - \sum_{m=1}^{M}((1-\omega)v_m, [g_{kh}]_m)_\Omega = J_1 + J_2.$$

To estimate J_1 we define $\psi = (1-\omega)v$ and proceed using the Ritz projection R_h defined by (5.5). There holds

$$\begin{aligned}(\nabla\psi(t), \nabla g_{kh}(t))_\Omega &= (\nabla R_h\psi(t), \nabla g_{kh}(t))_\Omega = -(R_h\psi(t), \Delta_h g_{kh}(t))_\Omega \\ &= -(R_h\psi(t), \Delta_h g_{kh}(t))_{B_{d/4}} - (R_h\psi(t), \Delta_h g_{kh}(t))_{\Omega\setminus B_{d/4}} \\ &\le \|R_h\psi(t)\|_{L^\infty(B_{d/4})}\|\Delta_h g_{kh}(t)\|_{L^1(B_{d/4})} \\ &\quad + \|\sigma^{-\frac{3}{2}}R_h\psi(t)\|_{L^2(\Omega\setminus B_{d/4})}\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}(t)\|_{L^2(\Omega)}.\end{aligned}$$

Using the estimate

$$\|\Delta_h g_{kh}(t)\|_{L^1(B_{d/4})} \le \|\sigma^{-\frac{3}{2}}\|_{L^2(\Omega)}\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}(t)\|_{L^2(B_{d/4})} \le c|\ln h|^{\frac{1}{2}}\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}(t)\|_{L^2(\Omega)}$$

we obtain

(7.5)
$$(\nabla\psi(t), \nabla g_{kh}(t))_\Omega \le c|\ln h|^{\frac{1}{2}}\left(\|R_h\psi(t)\|^2_{L^\infty(B_{d/4})} + \|\sigma^{-\frac{3}{2}}R_h\psi(t)\|^2_{L^2(\Omega\setminus B_{d/4})}\right)^{\frac{1}{2}}\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}(t)\|_{L^2(\Omega)}.$$

By the interior pointwise error estimates from Theorem 5.1 in [41], we have

$$\begin{aligned}\|R_h\psi(t)\|_{L^\infty(B_{d/4})} &\le c|\ln h|\|\psi(t)\|_{L^\infty(B_{d/2})} + Cd^{-\frac{3}{2}}\|R_h\psi(t)\|_{L^2(B_{d/2})} \\ &= Cd^{-\frac{3}{2}}\|R_h\psi(t)\|_{L^2(B_{d/2})},\end{aligned}$$

since the support of $\psi = (1-\omega)v$ is contained in $\Omega\setminus B_{d/2}$. On $\Omega\setminus B_{d/4}$ there holds $\sigma \ge d/4$ and therefore

$$\|\sigma^{-\frac{3}{2}}R_h\psi(t)\|_{L^2(\Omega\setminus B_{d/4})} \le cd^{-\frac{3}{2}}\|R_h\psi(t)\|_{L^2(\Omega\setminus B_{d/4})}.$$

Inserting the last two estimates into (7.5) we get

$$(\nabla\psi(t), \nabla g_{kh}(t))_\Omega \le cd^{-\frac{3}{2}}|\ln h|^{\frac{1}{2}}\|R_h\psi(t)\|_{L^2(\Omega)}\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}(t)\|_{L^2(\Omega)}.$$

Using a standard elliptic estimate and recalling $\psi = (1-\omega)v$ we have

$$\begin{aligned}\|R_h\psi(t)\|_{L^2(\Omega)} &\le \|\psi(t)\|_{L^2(\Omega)} + \|\psi(t) - R_h\psi(t)\|_{L^2(\Omega)} \\ &\le \|\psi(t)\|_{L^2(\Omega)} + ch\|\nabla\psi(t)\|_{L^2(\Omega)} \\ &\le \|v(t)\|_{L^2(\Omega)} + ch\|(1-\omega)\nabla v(t) - \nabla\omega v(t)\|_{L^2(\Omega)} \\ &\le c\|v(t)\|_{L^2(\Omega)} + ch\|\nabla v(t)\|_{L^2(\Omega)},\end{aligned}$$

where in the last step we used $|\nabla\omega| \le cd^{-1} \le ch^{-1}$. This results in

$$(\nabla\psi(t), \nabla g_{kh}(t))_\Omega \le cd^{-\frac{3}{2}}|\ln h|^{\frac{1}{2}} \left(\|v(t)\|_{L^2(\Omega)} + h\|\nabla v(t)\|_{L^2(\Omega)}\right) \|\sigma^{\frac{3}{2}}\Delta_h g_{kh}(t)\|_{L^2(\Omega)}.$$

Therefore, we get

$$J_1 \le cd^{-\frac{3}{2}}|\ln h|^{\frac{1}{2}} \left(\|v\|_{L^2(I;L^2(\Omega))} + h\|\nabla v\|_{L^2(I;L^2(\Omega))}\right) \|\sigma^{\frac{3}{2}}\Delta_h g_{kh}\|_{L^2(I;L^2(\Omega))}. \tag{7.6}$$

For J_2 we obtain

$$\begin{aligned} J_2 &\le \sum_{m=1}^{M} \|\sigma^{-\frac{3}{2}}(1-\omega)v_m\|_{L^2(\Omega)} k_m^{\frac{1}{2}} k_m^{-\frac{1}{2}} \|\sigma^{\frac{3}{2}}[g_{kh}]_m\|_{L^2(\Omega)} \\ &\le c\left(\sum_{m=1}^{M} d^{-3}k_m\|(1-\omega)v_m\|^2_{L^2(\Omega)}\right)^{\frac{1}{2}} \left(\sum_{m=1}^{M} k_m^{-1}\|\sigma^{\frac{3}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)}\right)^{\frac{1}{2}} \quad (7.7)\\ &\le cd^{-\frac{3}{2}}\|\pi_k v\|_{L^2(I;L^2(\Omega))} \left(\sum_{m=1}^{M} k_m^{-1}\|\sigma^{\frac{3}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)}\right)^{\frac{1}{2}}, \end{aligned}$$

where we used that $\operatorname{supp}(1-\omega)v_m \subset \Omega \setminus B_{d/2}$ and $\sigma \ge d/2$ on this set as well as the definition of π_k (3.3). Inserting the estimate (7.6) for J_1 and estimate (7.7) for J_2 into (7.4) we obtain

$$\begin{aligned} &B((1-\omega)v, g_{kh}) \\ &\quad\le cd^{-\frac{3}{2}}|\ln h|^{\frac{1}{2}} \left(\|v\|^2_{L^2(I;L^2(\Omega))} + h^2\|\nabla v\|^2_{L^2(I;L^2(\Omega))} + \|\pi_k v\|^2_{L^2(I;L^2(\Omega))}\right)^{\frac{1}{2}} \\ &\quad\quad\times \left(\|\sigma^{\frac{3}{2}}\Delta_h g_{kh}\|^2_{L^2(I;L^2(\Omega))} + \sum_{m=1}^{M} k_m^{-1}\|\sigma^{\frac{3}{2}}[g_{kh}]_m\|^2_{L^2(\Omega)}\right)^{\frac{1}{2}}. \end{aligned}$$

Using the estimates (6.16) and (6.17), we get

$$\begin{aligned} &B((1-\omega)v, g_{kh}) \\ &\quad\le cd^{-\frac{3}{2}}\ln\frac{T}{k}|\ln h| \left(\|v\|^2_{L^2(I;L^2(\Omega))} + h^2\|\nabla v\|^2_{L^2(I;L^2(\Omega))} + \|\pi_k v\|^2_{L^2(I;L^2(\Omega))}\right)^{\frac{1}{2}} \\ &\quad\quad\times \left(\int_0^T |v_{kh}(t,y)|^2\,dt\right)^{\frac{1}{2}}. \end{aligned}$$

Inserting this inequality into (7.3) we obtain

$$\begin{aligned} &\int_0^T |v_{kh}(t,x_0)|^2 dt \\ &\quad\le c\left(\ln\frac{T}{k}\right)^2 |\ln h|^2 \left(\|v\|^2_{L^2(I;L^\infty(B_d))} + h^{-\frac{6}{p}}\|\pi_k v\|^2_{L^2(I;L^p(B_d))}\right) \\ &\quad\quad + cd^{-3}\left(\ln\frac{T}{k}\right)^2 |\ln h|^2 \left(\|v\|^2_{L^2(I;L^2(\Omega))} + h^2\|\nabla v\|^2_{L^2(I;L^2(\Omega))} + \|\pi_k v\|^2_{L^2(I;L^2(\Omega))}\right). \end{aligned}$$

Using that the dG(0)cG(1) method is invariant on $X^{0,1}_{k,h}$, by replacing v and v_{kh} with $v-\chi$ and $v_{kh}-\chi$ for any $\chi \in X^{0,1}_{k,h}$ (cf. the discussion at the beginning of the proof of Theorem 3.3), we obtain the first estimate in Theorem 3.4. The second estimate is established following the lines of the proof, replacing v by v_k at all places and using $\pi_k v_k = v_k$.

Acknowledgments. The authors would like to thank Dominik Meidner and Konstantin Pieper for the careful reading of the manuscript and providing valuable suggestions that helped to improve the presentation of the paper.

REFERENCES

[1] H. W. Alt, *Linear Functional Analysis. An Application Oriented Introduction*, 6th ed., Springer, Berlin, 2011.
[2] H. Amann, *Nonautonomous parabolic equations involving measures*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI), 306 (2003), pp. 16–52.
[3] M. Amouroux and J.-P. Babary, *On the optimal pointwise control and parametric optimization of distributed parameter systems*, Internat. J. Control, 28 (1978), pp. 789–807.
[4] T. Apel, O. Benedix, D. Sirch, and B. Vexler, *A priori mesh grading for an elliptic problem with Dirac right-hand side*, SIAM J. Numer. Anal., 49 (2011), pp. 992–1005.
[5] A. Ashyralyev and P. E. Sobolevskiĭ, *Well-Posedness of Parabolic Difference Equations*, Oper. Theory Adv. Appl. 69, Birkhäuser Verlag, Basel, 1994.
[6] I. Babuška, *Error-bounds for finite element method*, Numer. Math., 16 (1970/1971), pp. 322–333.
[7] E. Casas, *L^2 estimates for the finite element method for the Dirichlet problem with singular data*, Numer. Math., 47 (1985), pp. 627–632.
[8] E. Casas, C. Clason, and K. Kunisch, *Parabolic control problems in measure spaces with sparse solutions*, SIAM J. Control Optim., 51 (2013), pp. 28–63.
[9] E. Casas and K. Kunisch, *Parabolic control problems in space-time measure spaces*, ESAIM Control Optim. Calc. Var., 22 (2016), pp. 355–370.
[10] E. Casas, B. Vexler, and E. Zuazua, *Sparse initial data identification for parabolic PDE and it's finite element approximations*, Math. Control Relat. Fields, 5 (2015), pp. 377–399.
[11] E. Casas and E. Zuazua, *Spike controls for elliptic and parabolic PDEs*, Systems Control Lett., 62 (2013), pp. 311–318.
[12] I. Chryssoverghi, *Approximate methods for optimal pointwise control of parabolic systems*, Systems Control Lett., 1 (1981), pp. 216–219.
[13] A. Cianchi, *Sharp Morrey-Sobolev inequalities and the distance from extremals*, Trans. Amer. Math. Soc., 360 (2008), pp. 4335–4347.
[14] E. J. Dean and P. Gubernatis, *Pointwise control of Burgers' equation—a numerical approach*, Comput. Math. Appl., 22 (1991), pp. 93–100.
[15] K. Disser, J. Rehberg, and A. ter Elst, *Hölder estimates for parabolic operators on domains with rough boundary*, preprint, arXiv:1503.07035, 2015.
[16] J. Droniou and J.-P. Raymond, *Optimal pointwise control of semilinear parabolic equations*, Nonlinear Anal., 39 (2000), pp. 135–156.
[17] J. Elschner, J. Rehberg, and G. Schmidt, *Optimal regularity for elliptic transmission problems including C^1 interfaces*, Interfaces Free Bound., 9 (2007), pp. 233–252.
[18] K. Eriksson, *Finite element methods of optimal order for problems with singular data*, Math. Comp., 44 (1985), pp. 345–360.
[19] K. Eriksson, C. Johnson, and V. Thomée, *Time discretization of parabolic problems by the discontinuous Galerkin method*, ESAIM Math. Model. Numer. Anal., 19 (1985), pp. 611–643.
[20] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems* II: *Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal., 32 (1995), pp. 706–740.
[21] S. J. Fromm, *Potential space estimates for Green potentials in convex domains*, Proc. Amer. Math. Soc., 119 (1993), pp. 225–233.
[22] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, 2nd ed., Classics Math., Springer, Berlin, 2001.
[23] W. Gong, *Error estimates for finite element approximations of parabolic equations with measure data*, Math. Comp., 82 (2013), pp. 69–98.
[24] W. Gong, M. Hinze, and Z. Zhou, *A priori error analysis for finite element approximation of parabolic optimal control problems with pointwise control*, SIAM J. Control Optim., 52 (2014), pp. 97–119.
[25] M. Hinze, *A variational discretization concept in control constrained optimization: The linear-quadratic case*, Comput. Optim. Appl., 30 (2005), pp. 45–61.
[26] T. Köppl and B. Wohlmuth, *Optimal a priori error estimates for an elliptic problem with Dirac right-hand side*, SIAM J. Numer. Anal., 52 (2014), pp. 1753–1769.

[27] 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.
[28] K. Kunisch, K. Pieper, and B. Vexler, *Measure valued directional sparsity for parabolic optimal control problems*, SIAM J. Control Optim., 52 (2014), pp. 3078–3108.
[29] D. Leykekhman, D. Meidner, and B. Vexler, *Optimal error estimates for finite element discretization of elliptic optimal control problems with finitely many pointwise state constraints*, Comput. Optim. Appl., 55 (2013), pp. 769–802.
[30] 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.
[31] D. Leykekhman and B. Vexler, *Discrete maximal parabolic regularity for Galerkin finite element methods*, published online, doi:10.1007/s00211-016-0821-2.
[32] D. Leykekhman and B. Vexler, *Finite element pointwise results on convex polyhedral domains*, SIAM J. Numer. Anal., 54 (2016), pp. 561–587.
[33] D. Leykekhman and B. Vexler, *Pointwise best approximation results for Galerkin finite element solutions of parabolic problems*, SIAM J. Numer. Anal., 54 (2016), pp. 1365–1384.
[34] J.-L. Lions, *Some Aspects of the Optimal Control of Distributed Parameter Systems*, Reg. Conf. Ser. Appl. Math. 6, SIAM, Philadelphia, 1972.
[35] J.-L. Lions and E. Magenes, *Non-homogeneous boundary value problems and applications, Vol.* II, Grundlehren Math. Wiss. 182, Springer, New York, 1972.
[36] 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.
[37] D. Meidner and B. Vexler, *A priori error estimates for space-time finite element discretisation of parabolic optimal control problems Part* I: *Problems without control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1150–1177.
[38] D. Meidner and B. Vexler, *A priori error estimates for space-time finite element discretisation of parabolic optimal control problems Part* II: *Problems with control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1301–1329.
[39] A. M. Ramos, R. Glowinski, and J. Periaux, *Pointwise control of the Burgers equation and related Nash equilibrium problems: Computational approach*, J. Optim. Theory Appl., 112 (2002), pp. 499–516.
[40] R. Rannacher, *L^∞-stability estimates and asymptotic error expansion for parabolic finite element equations*, in Extrapolation and Defect Correction (1990), Bonner Math. Schriften 228, University of Bonn, Bonn, 1991, pp. 74–94.
[41] A. H. Schatz and L. B. Wahlbin, *Interior maximum norm estimates for finite element methods*, Math. Comp., 31 (1977), pp. 414–442.
[42] A. H. Schatz and L. B. Wahlbin, *Interior maximum-norm estimates for finite element methods.* II, Math. Comp., 64 (1995), pp. 907–928.
[43] R. Scott, *Finite element convergence for singular data*, Numer. Math., 21 (1973), pp. 317–327.
[44] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, 2nd ed.,Johann Ambrosius Barth, Heidelberg, 1995.