Optimal a priori error estimates of parabolic optimal control problems with a moving point control

Dmitriy Leykekhman and Boris Vexler

Abstract In this paper we consider a parabolic optimal control problem with a Dirac type control with moving point source in two space dimensions. We discretize the problem with piecewise constant functions in time and continuous piecewise linear finite elements in space. For this discretization we show optimal order of convergence with respect to the time and the space discretization parameters modulo some logarithmic terms. Error analysis for the same problem was carried out in the recent paper [17], however, the analysis there contains a serious flaw. One of the main goals of this paper is to provide the correct proof. The main ingredients of our analysis are the global and local error estimates on a curve, that have an independent interest.

1 Introduction

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

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

subject to the second order parabolic equation

$$u_t(t,x) - \Delta u(t,x) = q(t)\delta_{\gamma(t)}, \qquad (t,x) \in I\times\Omega, \tag{2a}$$

$$u(t,x) = 0, \qquad (t,x) \in I\times\partial\Omega, \tag{2b}$$

$$u(0,x) = 0, \qquad x \in \Omega \tag{2c}$$

Dmitriy Leykekhman
Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA, e-mail: dmitriy.leykekhman@uconn.edu

Boris Vexler
Technical University of Munich, Chair of Optimal Control, Center for Mathematical Sciences, Boltzmannstraße 3, 85748 Garching by Munich, Germany, e-mail: vexler@ma.tum.de

and subject to pointwise control constraints

$$q_a \le q(t) \le q_b \quad \text{a.e. in } I. \tag{3}$$

Here $I=(0,T)$, $\Omega\subset\mathbb{R}^2$ is a convex polygonal domain and $\delta_{\gamma(t)}$ is the Dirac delta function at point $x_t=\gamma(t)$ at each t. We will assume:

Assumption 1 • $\gamma\in C^1(\bar I)$ *and* $\max_{t\in\bar I}|\gamma'(t)|\le C_\gamma$.

Assumption 2 • $\gamma(t)\subset\overline{\Omega}_0\subset\subset\Omega_1$*, for any* $t\in I$*, with* $\overline{\Omega}_1\subset\subset\Omega$.

The parameter α is assumed to be positive and the desired state $\hat u$ fulfills $\hat u\in L^2(I;L^\infty(\Omega))$. 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.

For the discretization, we consider the standard continuous piecewise linear finite elements in space and piecewise constant discontinuous Galerkin method in time. This is a special case ($r=0$, $s=1$) of so called dG(r)cG(s) discretization, see e.g. [14] for the analysis of the method for parabolic problems and e.g. [25, 26] for error estimates in the context of optimal control problems. Throughout, we will denote by h the spatial mesh size and by k the size of time steps, see Section 3 for details.

The main result of the paper is the following.

Theorem 1. *Let $\bar q$ be optimal control for the problem* (1)-(2) *and $\bar q_{kh}$ be the optimal dG(0)cG(1) solution. Then there exists a constant C independent of h and k such that*

$$\|\bar q-\bar q_{kh}\|_{L^2(I)}\le C\left(|\ln h|^3(k+h^2)+C_\gamma|\ln h|k\right)\left(\|\bar q\|_{L^2(I)}+\|\hat u\|_{L^2(I;L^\infty(\Omega))}\right).$$

We would also like to point out that in addition to the optimal order estimate, modulo logarithmic terms, our analysis does not require any relationship between the sizes of the space discretization h and the time steps k.

The problem with fixed location of the point source (i.e. with $\delta_{x_0}(x)$ for some fixed $x_0\in\Omega$) starting with the work of Lions [23], was investigated in a number of publications, see [2, 3, 10, 12, 28] for the continuous problem and [16, 21, 22] for the finite element approximation and error estimates. There is also a closely related problem of measured valued controls, which received a lot of attention lately [5, 6, 7, 8, 20].

The problem with moving Dirac was considered in [9, 27] on a continuous level. The error analysis was carried out in the recent paper [17]. However, the analysis there contains a serious flaw. The last inequality in the estimate (3.33) in [17] is not correct. One of the main goals of this paper is to provide the correct proof. The main ingredients of our analysis are the global and local error estimates on a curve, Theorem 2 and Theorem 3, respectively. These results are new and have an independent interest.

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

The rest of the paper is organized as follows. In Section 2 we discuss the functional analytic setting of the problem, state the optimality system and prove regularity results for the state and for the adjoint state. In Section 3 we establish important global and local best approximation results along the curve for the heat equation. Finally in Section 4 we prove our main result.

2 Optimal control problem and regularity

In order to state the functional analytic setting for the optimal control problem, we first introduce the auxiliary problem

$$\begin{aligned} v_t(t,x) - \Delta v(t,x) &= f(t,x), && (t,x) \in I \times \Omega,\\ v(t,x) &= 0, && (t,x) \in I \times \partial\Omega,\\ v(0,x) &= 0, && x \in \Omega, \end{aligned} \tag{4}$$

with a right-hand side $f \in L^2(I;L^p(\Omega))$ for some $1 < p < \infty$. This equation possesses a unique solution

$$v \in L^2(I;H^1_0(\Omega)) \cap H^1(I;H^{-1}(\Omega)).$$

Due to the convexity of the polygonal domain Ω the solution v possesses an additional regularity for $p = 2$:

$$v \in L^2(I;H^2(\Omega) \cap H^1_0(\Omega)) \cap H^1(I;L^2(\Omega)),$$

with the corresponding estimate

$$\|v\|_{L^2(I;H^2(\Omega))} + \|v_t\|_{L^2(I;L^2(\Omega))} \le C\|f\|_{L^2(I;L^2(\Omega))}, \tag{5}$$

see, e.g., [15]. From the Sobolev embedding $H^2(\Omega) \hookrightarrow W^{1,s}(\Omega)$ for any $s < \infty$ in two space dimensions and the previous lemma we can establish the following result for $s > 2$,

$$\|v\|_{L^2(I;W^{1,s}(\Omega))} \le Cs\|v\|_{L^2(I;H^2(\Omega))} \le Cs\|f\|_{L^2(I;L^2(\Omega))}. \tag{6}$$

The exact form of the constant can be traced, for example, from the proof of [1, Thm. 10.8]. In addition, there holds the following regularity result (see [21]).

Lemma 1. *If $f \in L^2(I;L^p(\Omega))$ for an arbitrary $p > 1$, then $v \in L^2(I;C(\Omega))$ and*

$$\|v\|_{L^2(I;C(\Omega))} \le C_p\|f\|_{L^2(I;L^p(\Omega))},$$

where $C_p \sim \frac{1}{p-1}$, as $p \to 1$.

We will also need the following local regularity result (see [21]).

Lemma 2. *Let $\Omega_0 \subset\subset \Omega_1 \subset\subset \Omega$ and $f \in L^2(I;L^2(\Omega)) \cap L^2(I;L^p(\Omega_1))$ for some $2 \le p < \infty$. Then $v \in L^2(I;W^{2,p}(\Omega_0)) \cap H^1(I;L^p(\Omega_0))$ and there exists a constant C*

independent of p such that

$$\|v_t\|_{L^2(I;L^p(\Omega_0))} + \|v\|_{L^2(I;W^{2,p}(\Omega_0))} \le Cp(\|f\|_{L^2(I;L^p(\Omega_1))} + \|f\|_{L^2(I;L^2(\Omega))}).$$

To introduce a weak solution of the state equation (2) we use the method of transposition, (cf. [24]). For a given control $q \in Q = L^2(I)$ we denote by $u = u(q) \in L^2(I;L^p(\Omega))$ with $2 \le p < \infty$ a weak solution of (2), if for all $\varphi \in L^2(I;L^{p'}(\Omega))$ with $\frac{1}{p} + \frac{1}{p'} = 1$ there holds

$$\langle u, \varphi \rangle_{L^2(I;L^p(\Omega)),L^2(I;L^{p'}(\Omega))} = \int_I w(t,\gamma(t))q(t)\,dt,$$

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

$$\begin{aligned} -w_t(t,x) - \Delta w(t,x) &= \varphi(t,x), & (t,x) &\in I \times \Omega, \\ w(t,x) &= 0, & (t,x) &\in I \times \partial\Omega, \\ w(T,x) &= 0, & x &\in \Omega. \end{aligned} \tag{7}$$

The existence of this weak solution $u = u(q)$ follows by duality using the embedding $L^2(I;W^{2,p'}(\Omega)) \hookrightarrow L^2(I;C(\Omega))$ for $p' > 1$. Using Lemma 1 we can prove additional regularity for the state variable $u = u(q)$.

Proposition 2.1 *Without lose of generality we assume $2 \le p < \infty$. Let $q \in Q = L^2(I)$ be given and $u = u(q)$ be the solution of the state equation (2). Then $u \in L^2(I;L^p(\Omega))$ for any $p < \infty$ and the following estimate holds for $p \to \infty$ with a constant C independent of p,*

$$\|u\|_{L^2(I;L^p(\Omega))} \le Cp\|q\|_{L^2(I)}.$$

Proof. To establish the result we use a duality argument. There holds

$$\|u\|_{L^2(I;L^p(\Omega))} = \sup_{\|\varphi\|_{L^2(I;L^{p'}(\Omega))}=1} (u,\varphi)_{I\times\Omega}, \quad \text{where} \quad \frac{1}{p} + \frac{1}{p'} = 1.$$

Let w be the solution to (7) for $\varphi \in L^2(I;L^{p'}(\Omega))$ with $\|\varphi\|_{L^2(I;L^{p'}(\Omega))} = 1$. From Lemma 1, $w \in L^2(I;C(\Omega))$ and the following estimate holds

$$\|w\|_{L^2(I;C(\Omega))} \le \frac{C}{p'-1}\|\varphi\|_{L^2(I;L^{p'}(\Omega))} = \frac{C}{p'-1} \le Cp, \text{ as } p \to \infty.$$

Thus,

$$\|u\|_{L^2(I;L^p(\Omega))} = \sup_{\|\varphi\|_{L^2(I;L^{p'}(\Omega))}=1} (u,\varphi)_{I\times\Omega}$$
$$= \int_I q(t)w(t,\gamma(t))\,dt \le \|q\|_{L^2(I)}\|w\|_{L^2(I;C(\Omega))} \le Cp\|q\|_{L^2(I)}.$$

Remark 1. We would like to note that the above regularity requires only Assumption 2 on γ. Higher regularity of γ is needed for optimal order error estimates only.

A further regularity result for the state equation follows from [13].

Proposition 2.2 *Let $q \in Q = L^2(I)$ be given and $u = u(q)$ be the solution of the state equation* (2)*. Then for each $1 < s < 2$ there holds*

$$u \in L^2(I;W_0^{1,s}(\Omega)) \quad \text{and} \quad u_t \in L^2(I;W^{-1,s}(\Omega)).$$

Moreover, the state u fulfills the following weak formulation

$$\langle u_t,\varphi\rangle + (\nabla u,\nabla\varphi) = \int_I q(t)\varphi(t,\gamma(t))\,dt \quad \text{for all } \varphi \in L^2(I;W_0^{1,s'}(\Omega)),$$

where $\frac{1}{s'} + \frac{1}{s} = 1$ and $\langle\cdot,\cdot\rangle$ is the duality product between $L^2(I;W^{-1,s}(\Omega))$ and $L^2(I;W_0^{1,s'}(\Omega))$.

Proof. For $s < 2$ we have $s' > 2$ and therefore $W_0^{1,s'}(\Omega)$ is embedded into $C(\bar{\Omega})$. Therefore the right-hand side $q(t)\delta_{\gamma(t)}$ of the state equation can be identified with an element in $L^2(I;W^{-1,s}(\Omega))$. Using the result from [13, Theorem 5.1] on maximal parabolic regularity and exploiting the fact that $-\Delta\colon W_0^{1,s}(\Omega) \to W^{-1,s}(\Omega)$ is an isomorphism, see [19], we obtain

$$u \in L^2(I;W_0^{1,s}(\Omega)) \quad \text{and} \quad u_t \in L^2(I;W^{-1,s}(\Omega)).$$

Given the above regularity the corresponding weak formulation is fulfilled by a standard density argument.

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 function in (1) and $u(q)$ is the weak solution of the state equation (2) as defined above. The optimal control problem can then be equivalently reformulated as

$$\min j(q), \quad q \in Q_{\text{ad}}, \tag{8}$$

where the set of admissible controls is defined according to (3) by

$$Q_{\text{ad}} = \{q \in Q \mid q_a \le q(t) \le q_b \text{ a. e. in } I\}. \tag{9}$$

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 all $p < \infty$, see

Proposition 2.1 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

$$j'(\bar{q})(\partial q - \bar{q}) \geq 0 \quad \text{for all } \partial q \in Q_{\text{ad}}. \tag{10}$$

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

$$j'(q)(\partial q) = \int_I (\alpha q(t) + z(t, \gamma(t)))\, \partial q(t)\, dt,$$

where $z = z(q)$ is the solution of the adjoint equation

$$-z_t(t,x) - \Delta z(t,x) = u(t,x) - \hat{u}(t,x), \qquad (t,x) \in I \times \Omega, \tag{11a}$$

$$z(t,x) = 0, \qquad (t,x) \in I \times \partial\Omega, \tag{11b}$$

$$z(T,x) = 0, \qquad x \in \Omega, \tag{11c}$$

and $u = u(q)$ on the right-hand side of (11a) is the solution of the state equation (2). The adjoint solution, which corresponds to the optimal control $\bar{q}$ is denoted by $\bar{z} = z(\bar{q})$.

The optimality condition (10) is a variational inequality, which can be equivalently formulated using the projection

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

The resulting condition reads:

$$\bar{q}(t) = P_{Q_{\text{ad}}}\left(-\frac{1}{\alpha}\bar{z}(t, \gamma(t))\right). \tag{12}$$

In the next proposition we provide regularity results for the solution of the adjoint equation.

Proposition 2.3 *Let $q \in Q$ be given, let $u = u(q)$ be the corresponding state fulfilling* (2) *and let $z = z(q)$ be the corresponding adjoint state fulfilling* (11)*. Then,*

(a) $z \in L^2(I; H^2(\Omega) \cap H^1_0(\Omega)) \cap H^1(I; L^2(\Omega))$ and the following estimate holds

$$\|\nabla^2 z\|_{L^2(I;L^2(\Omega))} + \|z_t\|_{L^2(I;L^2(\Omega))} \leq C(\|q\|_{L^2(I)} + \|\hat{u}\|_{L^2(I;L^2(\Omega))}).$$

(b) If $\Omega_0 \subset\subset \Omega$, then $z \in L^2(I; W^{2,p}(\Omega_0)) \cap H^1(I; L^p(\Omega_0))$ for all $2 \leq p < \infty$ and the following estimate holds

$$\|\nabla^2 z\|_{L^2(I;L^p(\Omega_0))} + \|z_t\|_{L^2(I;L^p(\Omega_0))} \leq Cp^2(\|q\|_{L^2(I)} + \|\hat{u}\|_{L^2(I;L^\infty(\Omega))}).$$

Proof. (a) The right-hand side of the adjoint equation fulfills $u - \hat{u} \in L^2(I; L^p(\Omega))$ for all $1 < p < \infty$, see Proposition 2.1. Due to the convexity of the domain Ω we directly obtain $z \in L^2(I; H^2(\Omega) \cap H^1_0(\Omega)) \cap H^1(I; L^2(\Omega))$ and the estimate

$$\|\nabla^2 z\|_{L^2(I;L^2(\Omega))} + \|z_t\|_{L^2(I;L^2(\Omega))} \le C\|u-\hat{u}\|_{L^2(I;L^2(\Omega))}.$$

The result from Proposition 2.1 leads directly to the first estimate.
(b) From Lemma 2 for $p \ge 2$ we have

$$\|\nabla^2 z\|_{L^2(I;L^p(\Omega_0))} + \|z_t\|_{L^2(I;L^p(\Omega_0))} \le Cp\|u-\hat{u}\|_{L^2(I;L^p(\Omega))}.$$

Hence, by the triangle inequality and Proposition 2.1 we obtain

$$\|u-\hat{u}\|_{L^2(I;L^p(\Omega))} \le C\left(p\|q\|_{L^2(I)} + \|\hat{u}\|_{L^2(I;L^\infty(\Omega))}\right).$$

That completes the proof.

3 Discretization and the best approximation type results

3.1 Space-time discretization and notation

For discretization of the problem under the consideration we introduce a partitions of $I=[0,T]$ into subintervals $I_m=(t_{m-1},t_m]$ of length $k_m=t_m-t_{m-1}$, where $0=t_0<t_1<\cdots<t_{M-1}<t_M=T$. We assume that

$$k_{m+1} \le \kappa k_m, \quad m=1,\dots,M-1, \quad \text{for some} \quad \kappa>0. \tag{13}$$

The maximal time step is denoted by $k=\max_m k_m$. The semidiscrete space X_k^0 of piecewise constant functions in time is defined by

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

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

$$v_m^+ = \lim_{\varepsilon\to 0^+} v(t_m+\varepsilon) := v_{m+1}, \quad v_m^- = \lim_{\varepsilon\to 0^+} v(t_m-\varepsilon) = v(t_m) := v_m, \quad [v]_m = v_m^+ - v_m^-. \tag{14}$$

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

$$\text{diam}(\tau) \le h \le C|\tau|^{\frac{1}{2}}, \quad \forall \tau \in \mathscr{T}$$

hold. 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 continuous piecewise linear finite elements. We will require the modified Clément interpolant $i_h\colon L^1(\Omega)\to V_h$ and the L^2-projection $P_h\colon L^2(\Omega)\to V_h$ defined by

$$(P_h v,\chi)_\Omega = (v,\chi)_\Omega, \quad \forall \chi \in V_h. \tag{15}$$

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

$$X_{k,h}^{0,1} = \{v_{kh} \in X_k^0 :\ v_{kh}|_{I_m} \in \mathscr{P}_0(I_m;V_h),\ m=1,2,\dots,M\}. \tag{16}$$

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

$$\pi_k v|_{I_m} = v(t_m), \quad m=1,2,\dots,M, \tag{17}$$

and the fully discrete projection $\pi_{kh}\colon C(\bar{I};L^1(\Omega)) \to X_{k,h}^{0,1}$ defined by $\pi_{kh} = i_h \pi_k$.

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

$$B(v,\varphi) = \sum_{m=1}^{M} \langle v_t, \varphi\rangle_{I_m\times\Omega} + (\nabla v, \nabla\varphi)_{I\times\Omega} + \sum_{m=2}^{M} ([v]_{m-1}, \varphi_{m-1}^+)_\Omega + (v_0^+, \varphi_0^+)_\Omega, \tag{18}$$

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, we obtain an equivalent (dual) expression for B:

$$B(v,\varphi) = -\sum_{m=1}^{M} \langle v, \varphi_t\rangle_{I_m\times\Omega} + (\nabla v, \nabla\varphi)_{I\times\Omega} - \sum_{m=1}^{M-1} (v_m^-, [\varphi_k]_m)_\Omega + (v_M^-, \varphi_M^-)_\Omega. \tag{19}$$

In the two following theorems we establish global and local best approximation type results along the curve for the error between the solution v of the auxiliary equation (4) and its dG(0)cG(1) approximation $v_{kh} \in X_{k,h}^{0,1}$ defined as

$$B(v_{kh}, \varphi_{kh}) = (f, \varphi_{kh})_{I\times\Omega} \quad \text{for all } \ \varphi_{kh} \in X_{k,h}^{0,1}. \tag{20}$$

Since dG(0)cG(1) method is a consistent discretization we have the following Galerkin orthogonality relation:

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

3.2 Discretization of the curve and the weight function

To define fully discrete optimization problem we will also require a discretization of the curve γ. We define $\gamma_k = \pi_k\gamma$ by

$$\gamma_k|_{I_m} = \gamma(t_m) := \gamma_{k,m} \in \Omega_0, \quad m=1,2,\dots,M, \tag{21}$$

i.e., γ_k is a piecewise constant approximation of γ. Next we introduce a weight function

$$\sigma(t,x) = \sqrt{|x-\gamma(t)|^2 + h^2} \tag{22}$$

and a discrete piecewise constant in time approximation

$$\sigma_k(t,x) = \sqrt{|x-\gamma_k(t)|^2+h^2}. \tag{23}$$

Define

$$\sigma_{k,m} := \sigma_k|_{I_m} = \sigma_k(t_m,x) = \sigma(t_m,x). \tag{24}$$

One can easily check that σ and σ_k satisfy the following properties for any $(t,x) \in I\times\Omega$,

$$\|\sigma^{-1}(t,\cdot)\|_{L^2(\Omega)}, \|\sigma_k^{-1}(t,\cdot)\|_{L^2(\Omega)} \le C|\ln h|^{\frac{1}{2}}, \quad t\in\bar{I}, \tag{25a}$$

$$|\nabla\sigma(t,x)|,\ |\nabla\sigma_k(t,x)| \le C, \tag{25b}$$

$$|\nabla^2\sigma_k(t,x)| \le C|\sigma_k^{-1}(t,x)|, \tag{25c}$$

$$|\sigma_t(t,x)| \le |\nabla\sigma(t,x)|\cdot|\gamma'(t)| \le CC_\gamma, \tag{25d}$$

$$\max_{x\in\tau}\sigma(x,t) \le C\min_{x\in\tau}\sigma(x,t), \quad \forall\tau\in\mathcal{T}. \tag{25e}$$

3.3 Global error estimate along the curve

In this section we prove the following global approximation result.

Theorem 2 (Global best approximation). *Assume v and v_{kh} satisfy* (4) *and* (20) *respectively. Then there exists a constant C independent of k and h such that for any $1\le p\le\infty$,*

$$\int_I |(v-v_{kh})(t,\gamma_k(t))|^2 dt \le C|\ln h|^2 \times \inf_{\chi\in X_{k,h}^{0,1}}\left(\|v-\chi\|^2_{L^2(I;L^\infty(\Omega))} + h^{-\frac{4}{p}}\|\pi_k v-\chi\|^2_{L^2(I;L^p(\Omega))}\right).$$

Proof. To establish the result we use a duality argument. First, we introduce a smoothed Delta function, which we will denote by $\tilde{\delta}_{\gamma_k}$. This function on each I_m is defined as $\tilde{\delta}_{\gamma_{k,m}}$ and supported in one cell, which we denote by τ_m^0, i.e.

$$(\chi,\tilde{\delta}_{\gamma_{k,m}})_{\tau_m^0} = \chi(\gamma_{k,m}) = \chi(\gamma(t_m)), \quad \forall\chi\in\mathbb{P}^1(\tau_m^0), \quad m=1,2,\dots,M.$$

In addition we also have (see [31, Appendix])

$$\|\tilde{\delta}_{\gamma_k}\|_{W_p^s(\Omega)} \le Ch^{-s-2(1-\frac{1}{p})}, \quad 1\le p\le\infty, \quad s=0,1. \tag{26}$$

Thus in particular $\|\tilde{\delta}_{\gamma_k}\|_{L^1(\Omega)}\le C$, $\|\tilde{\delta}_{\gamma_k}\|_{L^2(\Omega)}\le Ch^{-1}$, and $\|\tilde{\delta}_{\gamma_k}\|_{L^\infty(\Omega)}\le Ch^{-2}$.

We define g to be a solution to the following backward parabolic problem

$$\begin{aligned} -g_t(t,x) - \Delta g(t,x) &= v_{kh}(t,\gamma_k(t))\tilde{\delta}_{\gamma_k}(x), && (t,x) \in I \times \Omega, \\ g(t,x) &= 0, && (t,x) \in I \times \partial\Omega, \\ g(T,x) &= 0, && x \in \Omega. \end{aligned} \tag{27}$$

There holds

$$\begin{aligned} \int_{I\times\Omega} v_{kh}(t,\gamma_k(t))\tilde{\delta}_{\gamma_k}(x)\varphi_{kh}(t,x)\,dt\,dx &= \sum_{m=1}^{M}\int_{I_m} v_{kh}(t,\gamma_k(t))\left(\int_\Omega \tilde{\delta}_{\gamma_k}(x)\varphi_{kh}(t,x)dx\right)dt \\ &= \sum_{m=1}^{M}\int_{I_m} v_{kh}(t,\gamma_k(t))\varphi_{kh}(t,\gamma_k(t))dt \\ &= \int_I v_{kh}(t,\gamma_k(t))\varphi_{kh}(t,\gamma_k(t))dt. \end{aligned}$$

Let $g_{kh} \in X^{0,1}_{k,h}$ be dG(0)cG(1) solution defined by

$$B(\varphi_{kh}, g_{kh}) = (v_{kh}(t,\gamma_k(t))\tilde{\delta}_{\gamma_k}, \varphi_{kh})_{I\times\Omega}, \quad \forall \varphi_{kh} \in X^{0,1}_{k,h}. \tag{28}$$

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

$$\begin{aligned} \int_0^T |v_{kh}(t,\gamma_k(t))|^2 dt &= B(v_{kh}, g_{kh}) = B(v, g_{kh}) \\ &= (\nabla v, \nabla g_{kh})_{I\times\Omega} - \sum_{m=1}^{M}(v_m, [g_{kh}]_m)_\Omega, \end{aligned} \tag{29}$$

where we have used the dual expression (19) for the bilinear form B and the fact that the last term in (19) 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 (19) 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 spacial variable, by the Hölder's inequality we have

$$(\nabla v, \nabla g_{kh})_\Omega = \frac{1}{2}\sum_\tau (v, [\![\partial_n g_{kh}]\!])_{\partial\tau} \le C\|v\|_{L^\infty(\Omega)} \sum_\tau \|[\![\partial_n g_{kh}]\!]\|_{L^1(\partial\tau)}, \tag{30}$$

where $[\![\partial_n g_{kh}]\!]$ denotes the jumps of the normal derivatives across the element faces.

From Lemma 2.4 in [29] we have

$$\sum_\tau \|[\![\partial_n g_{kh}]\!]\|_{L^1(\partial\tau)} \le C|\ln h|^{\frac{1}{2}}\left(\|\sigma_k \Delta_h g_{kh}\|_{L^2(\Omega)} + \|\nabla g_{kh}\|_{L^2(\Omega)}\right),$$

where $\Delta_h \colon V_h \to V_h$ is the discrete Laplace operator, defined by

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

To estimate the term involving the jumps in (29), we first use the Hölder's inequality and the inverse estimate to obtain

$$\sum_{m=1}^{M}(v_m,[g_{kh}]_m)_\Omega \le C\sum_{m=1}^{M} k_m^{\frac{1}{2}}\|v_m\|_{L^p(\Omega)}k_m^{-\frac{1}{2}}h^{-\frac{2}{p}}\|[g_{kh}]_m\|_{L^1(\Omega)}. \tag{31}$$

Now we use the fact that the equation (28) can be rewritten on the each time level as

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

or equivalently as

$$-k_m\Delta_h g_{kh,m}-[g_{kh}]_m=k_m v_{kh,m}(\gamma_{k,m})P_h\tilde{\delta}_{\gamma_{k,m}}, \tag{32}$$

where P_h is the L^2-projection, see (15). From (32) by the triangle inequality, we obtain

$$\|[g_{kh}]_m\|_{L^1(\Omega)}\le k_m\|\Delta_h g_{kh,m}\|_{L^1(\Omega)}+k_m\|P_h\tilde{\delta}_{\gamma_{k,m}}\|_{L^1(\Omega)}|v_{kh,m}(\gamma_{k,m})|.$$

Using that the L^2-projection is stable in L^1-norm (cf. [11]), we have

$$\|P_h\tilde{\delta}_{\gamma_{k,m}}\|_{L^1(\Omega)}\le C\|\tilde{\delta}_{\gamma_{k,m}}\|_{L^1(\Omega)}\le C.$$

Inserting the above estimate into (31) and using (25a), we obtain

$$\begin{aligned}
&\sum_{m=1}^{M}(v_m,[g_{kh}]_m)_\Omega \le Ch^{-\frac{2}{p}}\sum_{m=1}^{M}k_m^{\frac{1}{2}}\|v_m\|_{L^p(\Omega)}k_m^{\frac{1}{2}}\left(\|\Delta_h g_{kh,m}\|_{L^1(\Omega)}+|v_{kh,m}(\gamma_{k,m})|\right)\\
&\le Ch^{-\frac{2}{p}}\left(\sum_{m=1}^{M}k_m\|v_m\|^2_{L^p(\Omega)}\right)^{\frac{1}{2}}\left(\sum_{m=1}^{M}k_m\|\Delta_h g_{kh,m}\|^2_{L^1(\Omega)}+k_m|v_{kh,m}(\gamma_{k,m})|^2\right)^{\frac{1}{2}}\\
&\le Ch^{-\frac{2}{p}}\|\pi_k v\|_{L^2(I;L^p(\Omega))}\left(\int_0^T|\ln h|\|\sigma_k\Delta_h g_{kh}\|^2_{L^2(\Omega)}+|v_{kh}(t,\gamma_k(t))|^2dt\right)^{\frac{1}{2}}.
\end{aligned}$$

Combining (29) and (30) with the above estimates we have

$$\begin{aligned}
&\int_0^T|v_{kh}(t,\gamma_k(t))|^2dt\le C|\ln h|^{\frac{1}{2}}\left(\|v\|_{L^2(I;L^\infty(\Omega))}+h^{-\frac{2}{p}}\|\pi_k v\|_{L^2(I;L^p(\Omega))}\right)\times\\
&\left(\int_0^T\|\sigma_k\Delta_h g_{kh}\|^2_{L^2(\Omega)}+\|\nabla g_{kh}\|^2_{L^2(\Omega)}+|v_{kh}(t,\gamma_k(t))|^2dt\right)^{\frac{1}{2}}.
\end{aligned} \tag{33}$$

To complete the proof of the theorem it is sufficient to show

$$\int_0^T\left(\|\sigma_k\Delta_h g_{kh}\|^2_{L^2(\Omega)}+\|\nabla g_{kh}\|^2_{L^2(\Omega)}\right)dt\le C|\ln h|\int_0^T|v_{kh}(t,\gamma_k(t))|^2dt. \tag{34}$$

Then from (33) and (34) it would follow that

$$\int_0^T|v_{kh}(t,\gamma_k(t))|^2dt\le C|\ln h|^2\left(\|v\|^2_{L^2(I;L^\infty(\Omega))}+h^{-\frac{4}{p}}\|\pi_k v\|^2_{L^2(I;L^p(\Omega))}\right).$$

Then using that the dG(0)cG(1) method is invariant on $X^{0,1}_{k,h}$, by replacing v an v_{kh} with $v-\chi$ and $v_{kh}-\chi$ for any $\chi \in Xkh$, we obtain Theorem 2.

The estimate (34) will follow from the series of lemmas. The first lemma treats the term $\|\sigma_k \Delta_h g_{kh}\|^2_{L^2(I;L^2(\Omega))}$.

Lemma 3. *For any $\varepsilon > 0$ there exists C_ε such that*

$$\int_0^T \|\sigma_k \Delta_h g_{kh}\|^2_{L^2(\Omega)} dt \le C_\varepsilon \int_0^T \left(|v_{kh}(t,\gamma_k(t))|^2 + \|\nabla g_{kh}\|^2_{L^2(\Omega)} \right) dt$$
$$+ \varepsilon \sum_{m=1}^M k_m^{-1} \|\sigma_{k,m}[g_{kh}]_m\|^2_{L^2(\Omega)},$$

where σ_k and $\sigma_{k,m}$ are defined in (23) *and* (24)*, respectively.*

Proof. The equation (28) for each time interval I_m can be rewritten as (32). Multiplying (32) with $\varphi = -\sigma_k^2 \Delta_h g_{kh}$ and integrating over $I_m \times \Omega$, we have

$$\begin{aligned}
&\int_{I_m} \|\sigma_{k,m}\Delta_h g_{kh}\|^2_{L^2(\Omega)} dt \\
&\quad = -([g_{kh}]_m, \sigma^2_{k,m}\Delta_h g_{kh,m})_\Omega - (v_{kh}(t,\gamma_{k,m}) P_h \tilde{\delta}_{\gamma_k}, \sigma^2_{k,m}\Delta_h g_{kh})_{I_m\times\Omega} \\
&\quad = -(P_h(\sigma^2_{k,m}[g_{kh}]_m), \Delta_h g_{kh,m})_\Omega - (v_{kh}(t,\gamma_{k,m}) P_h \tilde{\delta}_{\gamma_k}, \sigma^2_{k,m}\Delta_h g_{kh})_{I_m\times\Omega} \\
&\quad = (\nabla(\sigma^2_{k,m}[g_{kh}]_m), \nabla g_{kh,m})_\Omega + (\nabla(P_h - I)(\sigma^2_{k,m}[g_{kh}]_m), \nabla g_{kh,m})_\Omega \\
&\qquad - (v_{kh}(t,\gamma_{k,m}) P_h \tilde{\delta}_{\gamma_k}, \sigma^2_{k,m}\Delta_h g_{kh})_{I_m\times\Omega} \\
&\quad = J_1 + J_2 + J_3.
\end{aligned}$$

We have

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

By the Cauchy-Schwarz inequality and using (25b) we get

$$J_{11} \le C\|\sigma_{k,m}[g_{kh}]_m\|_{L^2(\Omega)} \|\nabla g_{kh,m}\|_{L^2(\Omega)}.$$

On the other hand we have

$$\begin{aligned}
J_{12} &= ([\sigma_k \nabla g_{kh}]_m, \sigma_{k,m}\nabla g_{kh,m})_\Omega + ((\sigma_{k,m} - \sigma_{k,m+1})\nabla g_{kh,m+1}, \sigma_{k,m}\nabla g_{kh,m})_\Omega \\
&= J_{121} + J_{122}.
\end{aligned}$$

Using the identity

$$([w_{kh}]_m, w_{kh,m})_\Omega = \frac{1}{2}\|w_{kh,m+1}\|^2_{L^2(\Omega)} - \frac{1}{2}\|w_{kh,m}\|^2_{L^2(\Omega)} - \frac{1}{2}\|[w_{kh}]_m\|^2_{L^2(\Omega)}, \quad (35)$$

we have

$$J_{121} = \frac{1}{2}\|\sigma_{k,m+1}\nabla g_{kh,m+1}\|^2_{L^2(\Omega)} - \frac{1}{2}\|\sigma_{k,m}\nabla g_{kh,m}\|^2_{L^2(\Omega)} - \frac{1}{2}\|[\sigma_k \nabla g_{kh}]_m\|^2_{L^2(\Omega)}.$$

By the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} J_{122} &\le \|(\sigma_{k,m}-\sigma_{k,m+1})\nabla g_{kh,m+1}\|_{L^2(\Omega)}\|\sigma_{k,m}\nabla g_{kh,m}\|_{L^2(\Omega)}\\ &\le CC_\gamma k_m\|\nabla g_{kh,m+1}\|_{L^2(\Omega)}\|\sigma_{k,m}\nabla g_{kh,m}\|_{L^2(\Omega)}, \end{aligned}$$

where in the last step we used that from (25d)

$$|\sigma_{k,m}(x)-\sigma_{k,m+1}(x)| = |\sigma(t_m,x)-\sigma(t_{m+1},x)| \le Ck_m|\sigma_t(\tilde{t},x)| \le CC_\gamma k_m,$$

for some $\tilde{t}\in I_m$. Using the Young's inequality for J_{11}, neglecting $-\frac{1}{2}\|[\sigma_k\nabla g_{kh}]_m\|^2_{L^2(\Omega)}$, and using the assumption on the time steps $k_m \le \kappa k_{m+1}$ and that $\sigma_k \le C$, we obtain

$$\begin{aligned} J_1 \le \frac{1}{2}\|\sigma_{k,m+1}\nabla g_{kh,m+1}\|^2_{L^2(\Omega)} &- \frac{1}{2}\|\sigma_{k,m}\nabla g_{kh,m}\|^2_{L^2(\Omega)} + \frac{\varepsilon}{k_m}\|\sigma_{k,m}[g_{kh}]_m\|^2_{L^2(\Omega)}\\ &+ C_\varepsilon k_m\|\nabla g_{kh,m}\|^2_{L^2(\Omega)} + Ck_{m+1}\|\nabla g_{kh,m+1}\|^2_{L^2(\Omega)}. \end{aligned} \tag{36}$$

To estimate J_2, first by the Cauchy-Schwarz inequality and the approximation theory we have

$$J_2 = \sum_\tau(\nabla(P_h-I)(\sigma^2_{k,m}[g_{kh}]_m),\nabla g_{kh,m})_\tau \le Ch\sum_\tau\|\nabla^2(\sigma^2_{k,m}[g_{kh}]_m)\|_{L^2(\tau)}\|\nabla g_{kh,m}\|_{L^2(\tau)}.$$

Using that g_{kh} is piecewise linear we have

$$\nabla^2(\sigma^2[g_{kh}]_m) = \nabla^2(\sigma^2)[g_{kh}]_m + \nabla(\sigma^2)\cdot\nabla[g_{kh}]_m \quad \text{on } \tau.$$

There holds $\partial_{ij}(\sigma^2) = 2(\partial_i\sigma)(\partial_j\sigma) + 2\sigma\partial_{ij}\sigma$ and $\nabla(\sigma^2) = 2\sigma\nabla\sigma$. Thus by the properties of σ (25b) and (25c), we have

$$|\nabla^2(\sigma^2)| \le C \quad \text{and} \quad |\nabla(\sigma^2)| \le C\sigma.$$

Same estimates hold for σ_k. Using these estimates, the fact that $h \le \sigma_k$ and the inverse inequality (in view of (25e) the inverise inequality is valid with σ inside the norm), we obtain

$$\begin{aligned} J_2 &\le C\sum_\tau\Big(h\|[g_{kh}]_m\|_{L^2(\tau)} + h\|\sigma_{k,m}\nabla[g_{kh}]_m\|_{L^2(\tau)}\Big)\|\nabla g_{kh,m}\|_{L^2(\tau)}\\ &\le C\sum_\tau\Big(\|\sigma_{k,m}[g_{kh}]_m\|_{L^2(\tau)} + C_{inv}\|\sigma_{k,m}[g_{kh}]_m\|_{L^2(\tau)}\Big)\|\nabla g_{kh,m}\|_{L^2(\tau)}\\ &\le C\sum_\tau\|\sigma_{k,m}[g_{kh}]_m\|_{L^2(\tau)}\|\nabla g_{kh,m}\|_{L^2(\tau)}\\ &\le C_\varepsilon k_m\|\nabla g_{kh,m}\|^2_{L^2(\Omega)} + \frac{\varepsilon}{k_m}\|\sigma_{k,m}[g_{kh}]_m\|^2_{L^2(\Omega)}. \end{aligned} \tag{37}$$

To estimate J_3 we first notice that

$$\|\sigma_k P_h\tilde{\delta}_{\gamma_k}\|_{L^2(\Omega)} \le C. \tag{38}$$

The proof is identical to the proof of (3.21) in [21].

By the Cauchy-Schwarz inequality, (38), and the Young's inequality, we obtain

$$J_3 \le C \int_{I_m} |v_{kh}(t,\gamma_k)|^2 dt + \frac{1}{2}\int_{I_m} \|\sigma_{k,m}\Delta_h g_{kh,m}\|^2_{L^2(\Omega)} dt. \tag{39}$$

Using the estimates (36), (37), and (39) we have

$$\begin{aligned}\int_{I_m} \|\sigma_{k,m}\Delta_h g_{kh}\|^2_{L^2(\Omega)} dt \le C_\varepsilon \int_{I_m} \left(|v_{kh}(t,\gamma_k(t))|^2 + \|\nabla g_{kh}\|^2_{L^2(\Omega)}\right) dt \\ + CC_\gamma \int_{I_{m+1}} \|\nabla g_{kh}\|^2_{L^2(\Omega)} dt + \frac{\varepsilon}{k_m}\|\sigma_{k,m}[g_{kh}]_m\|^2_{L^2(\Omega)} \\ + \frac{1}{2}\|\sigma_{k,m+1}\nabla g_{kh,m+1}\|^2_{L^2(\Omega)} - \frac{1}{2}\|\sigma_{k,m}\nabla g_{kh,m}\|^2_{L^2(\Omega)}.\end{aligned}$$

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

The second lemma treats the term involving jumps.

Lemma 4. *There exists a constant C such that*

$$\sum_{m=1}^{M} k_m^{-1}\|\sigma_{k,m}[g_{kh}]_m\|^2_{L^2(\Omega)} \le C\int_0^T \left(\|\sigma_k\Delta_h g_{kh}\|^2_{L^2(\Omega)} + |v_{kh}(t,\gamma_k(t)|^2\right) dt.$$

Proof. We test (32) with $\varphi|_{I_m} = \sigma^2_{k,m}[g_{kh}]_m$ and obtain

$$\begin{aligned}\|\sigma_{k,m}[g_{kh}]_m\|^2_{L^2(\Omega)} = &-(\Delta_h g_{kh}, \sigma^2_{k,m}[g_{kh}]_m)_{I_m\times\Omega} \\ &-(v_{kh}(t,\gamma_k(t))P_h\tilde{\delta}_{\gamma_k}, \sigma^2_{k,m}[g_{kh}]_m)_{I_m\times\Omega}.\end{aligned} \tag{40}$$

The first term on the right hand side of (40) using the Young's inequality can be estimated as

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

The last term on the right hand side of (40) can easily be estimated using (38) as

$$(v_{kh}(t,\gamma_{k,m})P_h\tilde{\delta}_{\gamma_k}, \sigma^2_{k,m}[g_{kh}]_m)_{I_m\times\Omega} \le Ck_m \int_{I_m} |v_{kh}(t,\gamma_k(t))|^2 dt + \frac{1}{4}\|\sigma_{k,m}[g_{kh}]_m\|^2_{L^2(\Omega)}.$$

Combining the above two estimates we obtain

$$\|\sigma_{k,m}[g_{kh}]_m\|^2_{L^2(\Omega)} \le Ck_m \int_{I_m} \left(\|\sigma_{k,m}\Delta_h g_{kh}\|^2_{L^2(\Omega)} + |v_{kh}(t,\gamma_k(t))|^2\right) dt.$$

Summing over m we obtain the lemma.

Lemma 5. *There exists a constant C such that*

$$\|\nabla g_{kh}\|^2_{L^2(I\times\Omega)} \le C|\ln h|\int_0^T |v_{kh}(t,\gamma_k(t))|^2 dt.$$

Proof. Adding the primal (18) and the dual (19) representation of the bilinear form $B(\cdot,\cdot)$ one immediately arrives at

$$\|\nabla v\|^2_{L^2(I\times\Omega)} \le B(v,v) \quad \text{for all} \;\; v\in X^0_k,$$

see e.g., [25]. Applying this inequality together with the discrete Sobolev inequality, see [4, Lemma 4.9.2], results in

$$\begin{aligned}
\|\nabla g_{kh}\|^2_{L^2(I\times\Omega)} &\le B(g_{kh},g_{kh}) = (v_{kh}(t,\gamma_k(t))\tilde{\delta}_{\gamma_k}, g_{kh})_{I\times\Omega}\\
&= \int_0^T v_{kh}(t,\gamma_k(t)) g_{kh}(t,\gamma_k(t))\,dt\\
&\le \left(\int_0^T |v_{kh}(t,\gamma_k(t))|^2\,dt\right)^{\frac{1}{2}} \left(\int_0^T |g_{kh}(t,\gamma_k(t))|^2\,dt\right)^{\frac{1}{2}}\\
&\le \left(\int_0^T |v_{kh}(t,\gamma_k(t))|^2\,dt\right)^{\frac{1}{2}} \|g_{kh}\|_{L^2(I;L^\infty(\Omega))}\\
&\le c|\ln h|^{\frac{1}{2}} \left(\int_0^T |v_{kh}(t,\gamma_k(t))|^2\,dt\right)^{\frac{1}{2}} \|\nabla g_{kh}\|_{L^2(I\times\Omega)}.
\end{aligned}$$

This gives the desired estimate.

We proceed with the proof of Theorem 2. From Lemma 3, Lemma 4, and Lemma 5. It follows that

$$\begin{aligned}
\int_0^T \left(\|\sigma_k\Delta_h g_{kh}\|^2_{L^2(\Omega)} + \|\nabla g_{kh}\|^2_{L^2(\Omega)}\right)dt \le C_\varepsilon|\ln h|\int_0^T |v_{kh}(t,\gamma_k(t))|^2 dt\\
+C\varepsilon\int_0^T \|\sigma_k\Delta_h g_{kh}\|^2_{L^2(\Omega)} dt.
\end{aligned}$$

Taking ε sufficiently small we have (34). From (33) we can conclude that

$$\int_0^T |v_{kh}(t,\gamma_k(t))|^2 dt \le C|\ln h|^2\left(\|v\|^2_{L^2(I;L^\infty(\Omega))} + h^{-\frac{4}{p}}\|\pi_k v\|^2_{L^2(I;L^p(\Omega))}\right),$$

for some constant C independent of h and k. Using that 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 $\chi\in X^{0,1}_{k,h}$, we obtain

$$\int_0^T |(v_{kh}-\chi)(t,\gamma_k(t))|^2 dt \le C|\ln h|^2\left(\|v-\chi\|^2_{L^2(I;L^\infty(\Omega))} + h^{-\frac{4}{p}}\|\pi_k v-\chi\|^2_{L^2(I;L^p(\Omega))}\right).$$

By the triangle inequality and the above estimate we deduce

$$\int_0^T |(v-v_{kh})(t,\gamma_k(t))|^2 dt \le \int_0^T |(v_{kh}-\chi)(t,\gamma_k(t))|^2 dt + \int_0^T |(v-\chi)(t,\gamma_k(t))|^2 dt$$
$$\le C|\ln h|^2 \left(\|v-\chi\|^2_{L^2(I;L^\infty(\Omega))} + h^{-\frac{4}{p}} \|\pi_k v - \chi\|^2_{L^2(I;L^p(\Omega))} \right).$$

Taking the infimum over χ, we obtain Theorem 2.

3.4 Interior error estimate

To obtain optimal error estimates we will also require the following interior result.

Theorem 3 (Interior approximation). *Let $B_{d,m} := B_d(\gamma(t_m))$ denote a ball of radius d centered at $\gamma(t_m)$. Assume v and v_{kh} satisfy* (4) *and* (20) *respectively and let $d > 4h$. Then there exists a constant C independent of h, k and d such that for any $1 \le p \le \infty$*

$$\int_0^T |(v-v_{kh})(t,\gamma_k(t))|^2 dt$$
$$\le C|\ln h|^2 \inf_{\chi \in X^{0,1}_{k,h}} \left\{ \sum_{m=1}^M \left(\|v-\chi\|^2_{L^2(I_m;L^\infty(B_{d,m}))} + h^{-\frac{4}{p}} \|\pi_k v - \chi\|^2_{L^2(I_m;L^p(B_{d,m}))} \right) \right.$$
$$\left. + d^{-2} \left(\|v-\chi\|^2_{L^2(I;L^2(\Omega))} + \|\pi_k v - \chi\|^2_{L^2(I;L^2(\Omega))} + h^2 \|\nabla(v-\chi)\|^2_{L^2(I;L^2(\Omega))} \right) \right\}. \tag{41}$$

Proof. To obtain the interior estimate we introduce a smooth cut-off function ω in space and piecewise constant in time, such that $\omega_m := \omega|_{I_m}$,

$$\omega_m(x) \equiv 1, \quad x \in B_{d/2,m} \tag{42a}$$
$$\omega_m(x) \equiv 0, \quad x \in \Omega \setminus B_{d,m} \tag{42b}$$
$$|\nabla \omega_m| \le C d^{-1}, \quad |\nabla^2 \omega_m| \le C d^{-2},. \tag{42c}$$

As in the proof of Theorem 2 we obtain by (29) that

$$\int_0^T |v_{kh}(t,\gamma_k(t))|^2 dt = B(v_{kh}, g_{kh}) = B(v, g_{kh}) = B(\omega v, g_{kh}) + B((1-\omega)v, g_{kh}), \tag{43}$$

where g_{kh} is the solution of (28). Note that ωv is discontinuous in time. The first term can be estimated using the global result from Theorem 2. 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

$$B(\omega v, g_{kh}) = B(\tilde{v}_{kh}, g_{kh}) = \int_0^T v_{kh}(t,\gamma_k(t))\tilde{v}_{kh}(t,\gamma_k(t))\,dt$$
$$\leq \frac{1}{2}\int_0^T |v_{kh}(t,\gamma_k(t))|^2\,dt + \frac{1}{2}\int_0^T |\tilde{v}_{kh}(t,\gamma_k(t))|^2\,dt.$$

Applying Theorem 2 for the second term, we have

$$\int_0^T |\tilde{v}_{kh}(t,\gamma_k(t))|^2 dt \leq C|\ln h|^2 \left(\|\omega v\|^2_{L^2(I;L^\infty(\Omega))} + h^{-\frac{4}{p}} \|\pi_k(\omega v)\|^2_{L^2(I;L^p(\Omega))} \right)$$
$$\leq C|\ln h|^2 \sum_{m=1}^{M} \left(\|v\|^2_{L^2(I_m;L^\infty(B_{d,m}))} + h^{-\frac{4}{p}} \|\pi_k v\|^2_{L^2(I_m;L^p(B_{d,m}))} \right).$$

From (43), canceling $\frac{1}{2}\int_0^T |v_{kh}(t,\gamma_k(t))|^2\,dt$ and using the above estimate, we obtain

$$\begin{aligned}\int_0^T |v_{kh}(t,\gamma_k(t))|^2 dt &\leq B((1-\omega)v, g_{kh}) \\ &+ C|\ln h|^2 \sum_{m=1}^{M} \left(\|v\|^2_{L^2(I_m;L^\infty(B_{d,m}))} + h^{-\frac{4}{p}} \|\pi_k v\|^2_{L^2(I_m;L^p(B_{d,m}))} \right).\end{aligned} \tag{44}$$

It remains to estimate the term $B((1-\omega)v, g_{kh})$. Using the dual expression (19) of the bilinear form B we obtain

$$\begin{aligned}B((1-\omega)v, g_{kh}) &= \sum_{m=1}^{M} \left((\nabla((1-\omega_m)v), \nabla g_{kh})_{I_m\times\Omega} - ((1-\omega_m)v_m, [g_{kh}]_m)_\Omega \right) \\ &= J_1 + J_2.\end{aligned} \tag{45}$$

To estimate J_1 we define $\psi = (1-\omega)v$ and proceed using the Ritz projection $R_h\colon H^1_0(\Omega) \to V_h$ defined by

$$(\nabla R_h v, \nabla \chi)_\Omega = (\nabla v, \nabla \chi)_\Omega, \quad \forall \chi \in V_h. \tag{46}$$

There holds

$$\begin{aligned}(\nabla\psi, \nabla g_{kh})_{I_m\times\Omega} &= (\nabla R_h\psi, \nabla g_{kh})_{I_m\times\Omega} = -(R_h\psi, \Delta_h g_{kh})_{I_m\times\Omega} \\ &= -(R_h\psi, \Delta_h g_{kh})_{I_m\times B_{d/4,m}} - (R_h\psi, \Delta_h g_{kh})_{I_m\times\Omega\setminus B_{d/4,m}} \\ &\leq \|R_h\psi\|_{L^2(I_m;L^\infty(B_{d/4,m}))} \|\Delta_h g_{kh}\|_{L^2(I_m;L^1(B_{d/4,m}))} \\ &+ \|\sigma_{k,m}^{-1} R_h\psi\|_{L^2(I_m\times\Omega\setminus B_{d/4,m})} \|\sigma_{k,m}\Delta_h g_{kh}\|_{L^2(I_m\times\Omega)}.\end{aligned}$$

Using the estimate

$$\begin{aligned}\|\Delta_h g_{kh}\|_{L^2(I_m;L^1(B_{d/4,m}))} &\leq \|\sigma_{k,m}^{-1}\|_{L^2(\Omega)} \|\sigma_{k,m}\Delta_h g_{kh}\|_{L^2(I_m\times B_{d/4,m})} \\ &\leq C|\ln h|^{\frac{1}{2}} \|\sigma_{k,m}\Delta_h g_{kh}\|_{L^2(I_m\times\Omega)},\end{aligned}$$

where in the last step we used (25a), we obtain

$$(\nabla\psi,\nabla g_{kh})_{I_m\times\Omega} \le C|\ln h|^{\frac{1}{2}}\left(\|R_h\psi\|_{L^2(I_m;L^\infty(B_{d/4,m}))}+\|\sigma_{k,m}^{-1}R_h\psi\|_{L^2(I_m\times\Omega\setminus B_{d/4,m})}\right)\\ \times\|\sigma_{k,m}\Delta_h g_{kh}\|_{L^2(I_m\times\Omega)}. \tag{47}$$

By the interior pointwise error estimates from Theorem 5.1 in [30], we have for each $t\in I_m$,

$$\begin{aligned}\|R_h\psi(t)\|_{L^\infty(B_{d/4,m})} &\le c|\ln h|\|\psi(t)\|_{L^\infty(B_{d/2,m})}+Cd^{-1}\|R_h\psi(t)\|_{L^2(B_{d/2,m})}\\ &= Cd^{-1}\|R_h\psi(t)\|_{L^2(B_{d/2,m})},\end{aligned}$$

since the support of $\psi_m=(1-\omega_m)v$ is contained in $\Omega\setminus B_{d/2,m}$. On $\Omega\setminus B_{d/4,m}$ there holds $\sigma_{k,m}\ge d/4$ and therefore for each $t\in I_m$,

$$\|\sigma_{k,m}^{-1}R_h\psi(t)\|_{L^2(\Omega\setminus B_{d/4,m})}\le Cd^{-1}\|R_h\psi(t)\|_{L^2(\Omega\setminus B_{d/4,m})}.$$

Inserting the last two estimates into (47) we get

$$(\nabla\psi,\nabla g_{kh})_{I_m\times\Omega}\le Cd^{-1}|\ln h|^{\frac{1}{2}}\|R_h\psi\|_{L^2(I_m\times\Omega)}\|\sigma_{k,m}\Delta_h g_{kh}\|_{L^2(I_m\times\Omega)}.$$

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

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

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

$$(\nabla\psi,\nabla g_{kh})_{I_m\times\Omega}\le Cd^{-1}|\ln h|^{\frac{1}{2}}\left(\|v\|_{L^2(I_m\times\Omega)}+h\|\nabla v\|_{L^2(I_m\times\Omega)}\right)\|\sigma_{k,m}\Delta_h g_{kh}\|_{L^2(I_m\times\Omega)}.$$

Therefore, we get

$$J_1\le cd^{-1}|\ln h|^{\frac{1}{2}}\left(\|v\|_{L^2(I;L^2(\Omega))}+h\|\nabla v\|_{L^2(I;L^2(\Omega))}\right)\|\sigma_k\Delta_h g_{kh}\|_{L^2(I;L^2(\Omega))}. \tag{48}$$

For J_2 we obtain

$$
\begin{aligned}
J_2 &\le \sum_{m=1}^{M} \|\sigma_m^{-1}(1-\omega_m)v_m\|_{L^2(\Omega)} k_m^{\frac{1}{2}} k_m^{-\frac{1}{2}} \|\sigma_{k,m}[g_{kh}]_m\|_{L^2(\Omega)} \\
&\le C\left(\sum_{m=1}^{M} d^{-2} k_m \|(1-\omega_m)v_m\|_{L^2(\Omega)}^2\right)^{\frac{1}{2}} \left(\sum_{m=1}^{M} k_m^{-1}\|\sigma_{k,m}[g_{kh}]_m\|_{L^2(\Omega)}^2\right)^{\frac{1}{2}} \qquad (49)\\
&\le C d^{-1}\|\pi_k v\|_{L^2(I;L^2(\Omega))}\left(\sum_{m=1}^{M} k_m^{-1}\|\sigma_{k,m}[g_{kh}]_m\|_{L^2(\Omega)}^2\right)^{\frac{1}{2}},
\end{aligned}
$$

where we used that $supp(1-\omega_m)v_m \subset \Omega \setminus B_{d/2,m}$ and $\sigma_{k,m} \ge d/2$ on this set as well as the definition of π_k (17). Inserting the estimate (48) for J_1 and the estimate (49) for J_2 into (45) we obtain

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

Using the estimate (34) and Lemma 4

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

Inserting this inequality into (44) we obtain

$$
\begin{aligned}
\int_0^T |v_{kh}(t,\gamma_k(t))|^2 dt \le C|\ln h|^2 &\left(\sum_{m=1}^{M} \|v\|_{L^2(I_m;L^\infty(B_{d,m}))}^2 + h^{-\frac{4}{p}}\|\pi_k v\|_{L^2(I_m;L^p(B_{d,m}))}^2\right) \\
&+ C d^{-2}|\ln h|^2 \left(\|v\|_{L^2(I;L^2(\Omega))}^2 + h^2\|\nabla v\|_{L^2(I;L^2(\Omega))}^2 + \|\pi_k v\|_{L^2(I;L^2(\Omega))}^2\right).
\end{aligned}
$$

Using that the dG(0)cG(1) method is invariant on $X_{k,h}^{0,1}$, by replacing v and v_{kh} with $v-\chi$ and $v_{kh}-\chi$ for any $\chi \in X_{k,h}^{0,1}$, we obtain the estimate in Theorem 3.

4 Discretization of the optimal control problem

In this section we describe the discretization of the optimal control problem (1)-(2) and prove our main result, Theorem 1. We start with 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

$$B(u_{kh},\varphi_{kh}) = \int_0^T q(t)\varphi_{kh}(t,\gamma_k(t))\,dt \quad \text{for all } \varphi_{kh} \in X^{0,1}_{k,h}. \tag{50}$$

Using the weak formulation for $u = u(q)$ from Proposition 2.2 we obtain the perturbed Galerkin orthogonality,

$$B(u-u_{kh},\varphi_{kh}) = \int_0^T q(t)\left(\varphi_{kh}(t,\gamma(t)) - \varphi_{kh}(t,\gamma_k(t))\right)dt \quad \text{for all } \varphi_{kh} \in X^{0,1}_{k,h}. \tag{51}$$

Note, that the jump terms involving u vanish due to the fact that

$$u \in H^1(I;W^{-1,s}(\Omega)) \hookrightarrow C(I;W^{-1,s}(\Omega))$$

and $\varphi_{kh,m} \in W^{1,\infty}(\Omega)$.

Similarly to the continuous problem, 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 function in (1). The discretized optimal control problem is then given as

$$\min\, j_{kh}(q), \quad q \in Q_{\mathrm{ad}}, \tag{52}$$

where Q_{ad} is the set of admissible controls (9). We note, that the control variable q is not explicitly discretized, cf. [18]. With standard arguments one proves the existence of a unique solution $\bar{q}_{kh} \in Q_{\mathrm{ad}}$ of (52). Due to convexity of the problem, the following condition is necessary and sufficient for the optimality,

$$j'_{kh}(\bar{q}_{kh})(\partial q - \bar{q}_{kh}) \geq 0 \quad \text{for all } \partial q \in Q_{\mathrm{ad}}. \tag{53}$$

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

$$j'_{kh}(q)(\partial q) = \int_I (\alpha q(t) + z_{kh}(t,\gamma_k(t)))\,\partial q(t)\,dt,$$

where $z_{kh} = z_{kh}(q)$ is the solution of the discrete adjoint equation

$$B(\varphi_{kh},z_{kh}) = (u_{kh}(q) - \hat{u},\varphi_{kh})_{I\times\Omega} \quad \text{for all } \varphi_{kh} \in X^{0,1}_{k,h}. \tag{54}$$

The discrete adjoint state, which corresponds to the discrete optimal control $\bar{q}_{kh}$ is denoted by $\bar{z}_{kh} = z(\bar{q}_{kh})$. The variational inequality (53) is equivalent to the following pointwise projection formula, cf. (12),

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

or

$$\bar{q}_{kh,m} = P_{Q_{\mathrm{ad}}}\left(-\frac{1}{\alpha}\bar{z}_{kh,m}(\gamma_{k,m})\right),$$

on each I_m. Due to the fact that $\bar{z}_{kh} \in X^{0,1}_{k,h}$, we have $\bar{z}_{kh}(t,\gamma_k(t))$ is piecewise constant and therefore by the projection formula also $\bar{q}_{kh}$ is piecewise constant. As a result no explicit discretization of the control variable is required.

To prove Theorem 1 we first need estimates for the error in the state and in the adjoint variables for a given (fixed) control q. Due to the structure of the optimality conditions, we will have to estimate the error $\|z(\cdot,\gamma(\cdot)) - z_{kh}(\cdot,\gamma_k(\cdot))\|_I$, where $z = z(q)$ and $z_{kh} = z_{kh}(q)$. Note, that z_{kh} is not the Galerkin projection of z due to the fact that the right-hand side of the adjoint equation (11) involves $u = u(q)$ and the right-hand side of the discrete adjoint equation (54) involves $u_{kh} = u_{kh}(q)$. To obtain an estimate of optimal order, we will first estimate the error $u - u_{kh}$ with respect to the $L^2(I;L^1(\Omega))$ norm. Note, that an L^2 estimate would not lead to an optimal result.

Theorem 4. *Let $q \in Q$ be given and let $u = u(q)$ be the solution of the state equation* (2) *and $u_{kh} = u_{kh}(q) \in X^{0,1}_{k,h}$ be the solution of the discrete state equation* (50)*. Then there holds the following estimate*

$$\|u - u_{kh}\|_{L^2(I;L^1(\Omega))} \leq \left(C|\ln h|^2(k+h^2) + C_\gamma|\ln h|k\right)\|q\|_I.$$

Proof. We denote by $e = u - u_{kh}$ the error and consider the following auxiliary dual problem

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

where

$$b(t,x) = \operatorname{sgn}(e(t,x))\|e(t,\cdot)\|_{L^1(\Omega)} \in L^2(I;L^\infty(\Omega))$$

and the corresponding discrete solution $w_{kh} \in X^{0,1}_{k,h}$ defined by

$$B(\varphi_{kh}, w - w_{kh}) = 0, \quad \forall \varphi_{kh} \in X^{0,1}_{k,h}.$$

Using (51) for $e = u - u_{kh}$ and the Galerkin orthogonality for $w - w_{kh}$ we obtain,

$$
\begin{aligned}
\int_0^T \|e(t,\cdot)\|^2_{L^1(\Omega)}\,dt &= (e,\operatorname{sgn}(e)\|e(t,\cdot)\|_{L^1(\Omega)})_{I\times\Omega}\\
&= (e,b)_{I\times\Omega}\\
&= B(e,w)\\
&= B(e,w-w_{kh})+B(e,w_{kh})\\
&= B(u,w-w_{kh})+B(e,w_{kh})\\
&= \int_0^T q(t)(w-w_{kh})(t,\gamma(t))dt+\int_0^T q(t)(w_{kh}(t,\gamma(t))-w_{kh}(t,\gamma_k(t)))dt \qquad (55)\\
&= \int_0^T q(t)(w(t,\gamma(t))-w_{kh}(t,\gamma_k(t)))dt\\
&\le \|q\|_I\left(\int_0^T |w(t,\gamma(t))-w_{kh}(t,\gamma_k(t))|^2 dt\right)^{\frac{1}{2}}\\
&\le \|q\|_I\left(\int_0^T \left(|w(t,\gamma(t))-w(t,\gamma_k(t))|^2+|(w-w_{kh})(t,\gamma_k(t))|^2\right)dt\right)^{\frac{1}{2}}.
\end{aligned}
$$

Using the local estimate from Theorem 3 with $B_{d,m}\subset\Omega_1$ for any $m=1,\dots,M$, where $\Omega_0\subset\subset\Omega_1\subset\subset\Omega$, we obtain

$$
\begin{aligned}
\int_0^T &|(w-w_{kh})(t,\gamma_k(t))|^2 dt\\
&\le C|\ln h|^2\int_0^T\left(\|w-\chi\|^2_{L^\infty(\Omega_1)}+h^{-\frac{4}{p}}\|\pi_k w-\chi\|^2_{L^p(\Omega_1)}\right)dt\\
&\quad + C|\ln h|^2\int_0^T\left(\|w-\chi\|^2_{L^2(\Omega)}+h^2\|\nabla(w-\chi)\|^2_{L^2(\Omega)}+\|\pi_k w-\chi\|^2\right)dt\\
&= J_1+J_2+J_3+J_4+J_5.
\end{aligned}
$$

We take $\chi=i_h\pi_k w$, where i_h is the modified Clément interpolant and π_k is the projection defined in (17). Thus, by the triangle inequality, approximation theory, inverse inequality and the stability of the Clément interpolant in L^p norm, we have

$$
\begin{aligned}
J_1 &\le C|\ln h|^2\int_0^T\left(\|w-i_h w\|^2_{L^\infty(\Omega_1)}+\|i_h(w-\pi_k w)\|^2_{L^\infty(\Omega_1)}\right)dt\\
&\le C|\ln h|^2\int_0^T\left(h^{4-\frac{4}{p}}\|w\|^2_{W^{2,p}(\Omega_1)}+h^{-\frac{4}{p}}\|i_h(w-\pi_k w)\|^2_{L^p(\Omega_1)}\right)dt\\
&\le Ch^{-\frac{4}{p}}|\ln h|^2(h^4+k^2)\int_0^T\left(\|w\|^2_{W^{2,p}(\Omega_1)}+\|w_t\|^2_{L^p(\Omega_1)}\right)dt.
\end{aligned}
$$

J_2 can be estimated similarly since for $\chi=i_h\pi_k w$ by the triangle inequality we have

$$
\|\pi_k w-i_h\pi_k w\|_{L^p(\Omega)}\le\|\pi_k w-w\|_{L^p(\Omega)}+\|w-i_h w\|_{L^p(\Omega)}+\|i_h(w-\pi_k w)\|_{L^p(\Omega)}.
$$

As a result

$$J_1+J_2 \le Ch^{-\frac{4}{p}}|\ln h|^2(h^4+k^2)\int_0^T \left(\|w\|^2_{W^{2,p}(\Omega_1)}+\|w_t\|^2_{L^p(\Omega_1)}\right)dt.$$

Using Lemma 2, we obtain

$$\int_0^T \left(\|w\|^2_{W^{2,p}(\Omega_1)}+\|w_t\|^2_{L^p(\Omega_1)}\right)dt \le Cp^2\|b\|^2_{L^2(I;L^p(\Omega))} \le Cp^2\|e\|^2_{L^2(I;L^1(\Omega))}, \quad (56)$$

and hence

$$J_1+J_2 \le Ch^{-\frac{4}{p}}|\ln h|^2(h^4+k^2)p^2\|e\|^2_{L^2(I;L^1(\Omega))}. \quad (57)$$

For the terms J_3 and J_4 we obtain using an L^2-estimate from [25]

$$\begin{aligned} J_3+J_4 &\le C|\ln h|^2(h^4+k^2)\left(\|\nabla^2 w\|^2_{L^2(I;L^2(\Omega))}+\|w_t\|^2_{L^2(I;L^2(\Omega))}\right) \\ &\le C|\ln h|^2(h^4+k^2)\|b\|^2_{L^2(I;L^2(\Omega))} \\ &\le C|\ln h|^2(h^4+k^2)\|e\|^2_{L^2(I;L^1(\Omega))}. \end{aligned}$$

J_5 can be estimated similarly since by the triangle inequality

$$\|\pi_k w - i_h\pi_k w\|_{L^2(I\times\Omega)} \le \|\pi_k w - w\|_{L^2(I\times\Omega)} + \|w - i_h\pi_k w\|_{L^2(I\times\Omega)}.$$

On the other hand using that $w\in L^2(I;W^{2,p}(\Omega_0))$ for $p>2$ and that $W^{2,p}(\Omega_0))\hookrightarrow C^1(\Omega_0)$ for $p>2$, and using Assumption 1, we have

$$\begin{aligned} \int_0^T |w(t,\gamma(t))-w(t,\gamma_k(t))|^2 dt &\le \int_0^T \|w(t,\cdot)\|^2_{C^1(\Omega_0)}|\gamma(t)-\gamma_k(t)|^2 dt \\ &\le C\|\gamma-\gamma_k\|^2_{C^0(I)}\int_0^T \|w(t,\cdot)\|^2_{W^{2,p}(\Omega_0)}dt \\ &\le CC_\gamma^2 k^2\|w\|^2_{L^2(I;W^{2,p}(\Omega_0))} \\ &\le CC_\gamma^2 k^2 p^2\|b\|^2_{L^2(I;L^p(\Omega))} \\ &\le CC_\gamma^2 k^2 p^2\|e\|^2_{L^2(I;L^1(\Omega))}, \end{aligned}$$

where in the last two steps we used (56). Combining the estimate for J_1, J_2, J_3, J_4, J_5 and the above estimate and inserting them into (55) we obtain:

$$\|e\|_{L^2(I;L^1(\Omega))} \le \left(C|\ln h|(ph^{-\frac{2}{p}}+1)(h^2+k)+C_\gamma pk\right)\|q\|_{L^2(I)}.$$

Setting $p=|\ln h|$ completes the proof.

In the following theorem we provide an estimate of the error in the adjoint state for fixed control q.

Theorem 5. *Let $q\in Q$ be given and let $z=z(q)$ be the solution of the adjoint equation* (11) *and $z_{kh}=z_{kh}(q)\in X^{0,1}_{k,h}$ be the solution of the discrete adjoint equation* (54)*. Then there holds the following estimate*

$$\left(\int_0^T |z(t,\gamma(t)) - z_{kh}(t,\gamma_k(t))|^2 \, dt \right)^{\frac{1}{2}} \\ \le C \left(|\ln h|^3 (k+h^2) + C_\gamma |\ln h| k \right) \left(\|q\|_{L^2(I)} + \|\hat{u}\|_{L^2(I;L^\infty(\Omega))} \right).$$

Proof. First by the triangle inequality

$$\int_0^T |z(t,\gamma(t)) - z_{kh}(t,\gamma_k(t))|^2 \, dt \le \int_0^T |z(t,\gamma(t)) - z(t,\gamma_k(t))|^2 \, dt \\ + \int_0^T |(z - z_{kh})(t,\gamma_k(t))|^2 \, dt.$$

Using Proposition 2.3 and the assumptions on γ, we have similarly to Theorem 4

$$\begin{aligned} \int_0^T |z(t,\gamma(t)) - z(t,\gamma_k(t))|^2 \, dt &\le C \|\gamma - \gamma_k\|^2_{C^0(I)} \int_0^T \|z(t,\cdot)\|^2_{C^1(\Omega_0)} \, dt \\ &\le C C_\gamma^2 k^2 \int_0^T \|z(t,\cdot)\|^2_{W^{2,p}(\Omega_0)} \, dt \\ &\le C C_\gamma^2 p k^2 \left(\|q\|^2_{L^2(I)} + \|\hat{u}\|^2_{L^2(I\times\Omega)} \right). \end{aligned}$$

Setting $p = |\ln h|$, we obtain

$$\left(\int_0^T |z(t,\gamma(t)) - z(t,\gamma_k(t))|^2 \, dt \right)^{\frac{1}{2}} \le C C_\gamma |\ln h| k \left(\|q\|_{L^2(I)} + \|\hat{u}\|_{L^2(I\times\Omega)} \right). \quad (58)$$

Next, we introduce an intermediate adjoint state $\widetilde{z}_{kh} \in X^{0,1}_{k,h}$ defined by

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

where $u = u(q)$ and therefore $\widetilde{z}_{kh}$ is the Galerkin projection of z. By the local best approximation result of Theorem 3 for any $\chi \in X^{0,1}_{k,h}$ we have

$$\begin{aligned} \int_0^T |(z - \widetilde{z}_{kh})(t,\gamma_k(t))|^2 \, dt &\le C |\ln h|^2 \int_0^T \left(\|z - \chi\|^2_{L^\infty(\Omega_1)} + h^{-\frac{4}{p}} \|\pi_k z - \chi\|^2_{L^p(\Omega_1)} \right) dt \\ &\quad + C |\ln h|^2 \int_0^T \left(\|z - \chi\|^2_{L^2(\Omega)} + h \|\nabla (z - \chi)\|^2_{L^2(\Omega)} + \|\pi_k z - \chi\|^2_{L^2(\Omega)} \right) dt \\ &= J_1 + J_2 + J_3 + J_4 + J_5. \end{aligned}$$

The terms J_1, J_2, J_3, J_4 and J_5 can be estimated the same way as in the proof of Theorem 4 using the regularity result for the adjoint state z from Proposition 2.3. This results in

$$\int_0^T |(z - \widetilde{z}_{kh})(t,\gamma_k(t)|^2 \, dt \le C |\ln h|^2 (p h^{-\frac{2}{p}} + 1)^2 (h^2 + k)^2 \left(\|q\|^2_{L^2(I)} + \|\hat{u}\|^2_{L^2(I;L^\infty(\Omega))} \right).$$

Setting $p = |\ln h|$ and taking square root, we obtain

$$\left(\int_0^T |(z-\widetilde{z}_{kh})(t,\gamma_k(t))|^2\,dt\right)^{\frac{1}{2}} \le C|\ln h|^2(h^2+k)\left(\|q\|_{L^2(I)}+\|\hat{u}\|_{L^2(I;L^\infty(\Omega))}\right). \tag{59}$$

It remains to estimate the corresponding error between $\widetilde{z}_{kh}$ and z_{kh}. We denote $e_{kh} = \widetilde{z}_{kh} - z_{kh} \in X^{0,1}_{k,h}$. Then we have

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

As in the proof of Lemma 5 we use the fact that

$$\|\nabla v\|^2_{L^2(I\times\Omega)} \le B(v,v)$$

holds for all $v \in X^{0,1}_{k,h}$. Applying this inequality together with the discrete Sobolev inequality, see [4], results in

$$\begin{aligned}
\|e_{kh}\|^2_{L^2(I;L^\infty(\Omega))} &\le C|\ln h|\|\nabla e_{kh}\|^2_{L^2(I\times\Omega)}\\
&\le C|\ln h|B(e_{kh},e_{kh})\\
&= C|\ln h|(u-u_{kh},e_{kh})_{I\times\Omega}\\
&\le C|\ln h|\|u-u_{kh}\|_{L^2(I;L^1(\Omega))}\|e_{kh}\|_{L^2(I;L^\infty(\Omega))}.
\end{aligned}$$

Therefore

$$\|e_{kh}\|_{L^2(I;L^\infty(\Omega))} \le C|\ln h|\|u-u_{kh}\|_{L^2(I;L^1(\Omega))}.$$

Using Theorem 4 we obtain

$$\|e_{kh}\|_{L^2(I;L^\infty(\Omega))} \le C\left(|\ln h|^3(k+h^2)+C_\gamma|\ln h|k\right)\|q\|_{L^2(I)}.$$

Combining this estimate with (59) we complete the proof.

Using the result of Theorem 5 we proceed with the proof of Theorem 1.

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

$$j''_{kh}(q)(p,p) \ge \alpha\|p\|^2_{L^2(I)} \quad \text{for all } p \in Q. \tag{60}$$

Using optimality conditions (10) for $\bar{q}$ and (53) 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 (60) we get

$$\begin{aligned}
\alpha\|\bar q-\bar q_{kh}\|^2_{L^2(I)} &\le j''_{kh}(\bar q)(\bar q-\bar q_{kh},\bar q-\bar q_{kh})_I\\
&= j'_{kh}(\bar q)(\bar q-\bar q_{kh}) - j'_{kh}(\bar q_{kh})(\bar q-\bar q_{kh})\\
&\le j'_{kh}(\bar q)(\bar q-\bar q_{kh}) - j'(\bar q)(\bar q-\bar q_{kh})\\
&= (z(\bar q)(t,\gamma(t)) - z_{kh}(\bar q)(t,\gamma_k(t)),\bar q-\bar q_{kh})_I\\
&\le \left(\int_0^T |z(\bar q)(t,\gamma(t)) - z_{kh}(\bar q)(t,\gamma_k(t))|^2\,dt\right)^{\frac12}\|\bar q-\bar q_{kh}\|_{L^2(I)}.
\end{aligned}$$

Applying Theorem 5 completes the proof.

References

1. H. W. ALT, *Linear functional analysis*, Universitext, Springer-Verlag London, Ltd., London, 2016. An application-oriented introduction, Translated from the German edition by Robert Nürnberg.
2. 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.
3. H. T. BANKS, ed., *Control and estimation in distributed parameter systems*, vol. 11 of Frontiers in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.
4. S. C. BRENNER AND L. R. SCOTT, *The mathematical theory of finite element methods*, vol. 15 of Texts in Applied Mathematics, Springer, New York, third ed., 2008.
5. 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.
6. E. CASAS AND K. KUNISCH, *Parabolic control problems in space-time measure spaces*, ESAIM Control Optim. Calc. Var., 22 (2016), pp. 355–370.
7. E. CASAS, B. VEXLER, AND E. ZUAZUA, *Sparse initial data identification for parabolic PDE and its finite element approximations*, Math. Control Relat. Fields, 5 (2015), pp. 377–399.
8. E. CASAS AND E. ZUAZUA, *Spike controls for elliptic and parabolic PDEs*, Systems Control Lett., 62 (2013), pp. 311–318.
9. C. CASTRO AND E. ZUAZUA, *Unique continuation and control for the heat equation from an oscillating lower dimensional manifold*, SIAM J. Control Optim., 43 (2004/05), pp. 1400–1434 (electronic).
10. I. CHRYSSOVERGHI, *Approximate methods for optimal pointwise control of parabolic systems*, Systems Control Lett., 1 (1981/82), pp. 216–219.
11. M. CROUZEIX AND V. THOMÉE, *The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces*, Math. Comp., 48 (1987), pp. 521–532.
12. J. DRONIOU AND J.-P. RAYMOND, *Optimal pointwise control of semilinear parabolic equations*, Nonlinear Anal., 39 (2000), pp. 135–156.
13. 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.
14. K. ERIKSSON, C. JOHNSON, AND V. THOMÉE, *Time discretization of parabolic problems by the discontinuous Galerkin method*, RAIRO Modél. Math. Anal. Numér., 19 (1985), pp. 611–643.
15. L. C. EVANS, *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2010.
16. 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.

17. W. GONG AND N. YAN, *Finite element approximations of parabolic optimal control problems with controls acting on a lower dimensional manifold*, SIAM J. Numer. Anal., 54 (2016), pp. 1229–1262.
18. M. HINZE, *A variational discretization concept in control constrained optimization: the linear-quadratic case*, Comput. Optim. Appl., 30 (2005), pp. 45–61.
19. D. JERISON AND C. E. KENIG, *The inhomogeneous Dirichlet problem in Lipschitz domains*, J. Funct. Anal., 130 (1995), pp. 161–219.
20. 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.
21. 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.
22. ———, *A priori error estimates for three dimensional parabolic optimal control problems with pointwise control*, SIAM J. Control Optim., 54 (2016), pp. 2403–2435.
23. J.-L. LIONS, *Optimal control of systems governed by partial differential equations.*, Translated from the French by S. K. Mitter. Die Grundlehren der mathematischen Wissenschaften, Band 170, Springer-Verlag, New York-Berlin, 1971.
24. J.-L. LIONS AND E. MAGENES, *Non-homogeneous boundary value problems and applications. Vol. II*, Springer-Verlag, New York-Heidelberg, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 182.
25. D. MEIDNER AND B. VEXLER, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems. I. Problems without control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1150–1177.
26. ———, *A priori error estimates for space-time finite element discretization of parabolic optimal control problems. II. Problems with control constraints*, SIAM J. Control Optim., 47 (2008), pp. 1301–1329.
27. P. A. NGUYEN AND J.-P. RAYMOND, *Control problems for convection-diffusion equations with control localized on manifolds*, ESAIM Control Optim. Calc. Var., 6 (2001), pp. 467–488 (electronic).
28. ———, *Pointwise control of the Boussinesq system*, Systems Control Lett., 60 (2011), pp. 249–255.
29. R. RANNACHER, *L^∞-stability estimates and asymptotic error expansion for parabolic finite element equations*, in Extrapolation and defect correction (1990), vol. 228 of Bonner Math. Schriften, Univ. Bonn, Bonn, 1991, pp. 74–94.
30. A. H. SCHATZ AND L. B. WAHLBIN, *Interior maximum norm estimates for finite element methods*, Math. Comp., 31 (1977), pp. 414–442.
31. ———, *Interior maximum-norm estimates for finite element methods. II*, Math. Comp., 64 (1995), pp. 907–928.