SIAM J. NUMER. ANAL.
Vol. 56, No. 4, pp. 2178–2202

DISCRETE MAXIMAL PARABOLIC REGULARITY FOR GALERKIN FINITE ELEMENT METHODS FOR NONAUTONOMOUS PARABOLIC PROBLEMS*

DMITRIY LEYKEKHMAN† AND BORIS VEXLER‡

Abstract. The main goal of the paper is to establish time semidiscrete and space-time fully discrete maximal parabolic regularity for the lowest order time discontinuous Galerkin solution of linear parabolic equations with time-dependent coefficients. Such estimates have many applications. As one of the applications we establish best approximations type results with respect to the $L^p(0,T;L^2(\Omega))$ norm for $1 \le p \le \infty$.

Key words. parabolic problems, maximal parabolic regularity, discrete maximal parabolic regularity, finite elements, discontinuous Galerkin methods, optimal error estimates, time-dependent coefficients, nonautonomous problems

AMS subject classifications. 65N30, 65N15

DOI. 10.1137/17M114100X

1. Introduction.

Let Ω be a Lipschitz domain in $\mathbb{R}^d$, $d \ge 1$, and $I = (0,T)$. We consider the following second order parabolic partial differential equation with time-dependent coefficients,

$$\begin{aligned} \partial_t u(t,x) + A(t,x)u(t,x) &= f(t,x), & (t,x) &\in I\times\Omega,\\ u(t,x) &= 0, & (t,x) &\in I\times\partial\Omega,\\ u(0,x) &= u_0(x), & x &\in \Omega, \end{aligned} \tag{1.1}$$

with the right-hand side $f \in L^p(I;L^2(\Omega))$ for some $1 \le p \le \infty$ and $u_0 \in L^2(\Omega)$, where the time-dependent elliptic operator is given by the formal expression

$$A(t,x)u(t,x) = -\sum_{i,j=1}^{d} \partial_j(a_{ij}(t,x)\partial_i u(t,x)) \tag{1.2}$$

with $a_{ij}(t,x) \in L^\infty(I\times\Omega)$ for $i,j = 1,\dots,d$ satisfying $a_{ij} = a_{ji}$ and the uniform ellipticity property

$$\sum_{i,j=1}^{d} a_{ij}(t,x)\xi_i\xi_j \ge \alpha|\xi|^2 \quad \forall \xi \in \mathbb{R}^d \quad \text{and a.e. } (t,x) \in I\times\Omega \tag{1.3}$$

for some constant $\alpha > 0$. We also assume that the coefficients $a_{ij}(t,x)$ are continuous in t for almost all $x \in \Omega$ and that the following condition holds:

$$|a_{ij}(t_1,x) - a_{ij}(t_2,x)| \le \omega(|t_1 - t_2|), \quad 1 \le i,j \le d, \tag{1.4}$$

---

*Received by the editors July 27, 2017; accepted for publication (in revised form) May 16, 2018; published electronically July 19, 2018.
http://www.siam.org/journals/sinum/56-4/M114100.html
Funding: The work of the first author was partially supported by NSF grant DMS-1115288.
†Department of Mathematics, University of Connecticut, Storrs, CT 06269 (dmitriy.leykekhman@uconn.edu).
‡Chair of Optimal Control, Center for Mathematical Sciences,Technical University of Munich, 85748 Garching by Munich, Germany (vexler@ma.tum.de).

for all $t_1, t_2 \in \bar{I}$ and almost all $x \in \Omega$, where $\omega\colon [0,T] \to [0,\infty)$ is a nondecreasing function such that

$$\frac{w(t)}{t^{\frac{3}{2}}} \text{ is nonincreasing on } (0,T] \text{ and } \quad \int_0^T \frac{\omega(t)}{t^{\frac{3}{2}}} dt < \infty; \tag{1.5}$$

see [17] for a similar assumption. This assumption is fulfilled, for example, if a_{ij} is Hölder continuous with exponent $\frac{1}{2} + \varepsilon$ in time and uniformly continuous in space.

The maximal parabolic regularity for $u_0 \equiv 0$ says that there exists a constant C such that for $f \in L^p(I; L^s(\Omega))$,

$$\|\partial_t u\|_{L^p(I;L^s(\Omega))} + \|A(\cdot)u\|_{L^p(I;L^s(\Omega))} \leq C \, \|f\|_{L^p(I;L^s(\Omega))}, \quad 1 < s, p < \infty. \tag{1.6}$$

For time-independent coefficients the above result is well understood [4, 6, 7, 18, 19], but for time-dependent coefficients it is still an active area of research [3, 8, 9, 17]. The maximal parabolic regularity is an important analytical tool and has a number of applications, especially to nonlinear problems and optimal control problems when sharp regularity results are required (cf., e.g., [21, 26, 27, 29, 32]).

The main goal of this paper is to establish similar maximal parabolic regularity results for time semidiscrete discontinuous Galerkin solutions as well as for fully discrete Galerkin approximations. Such results are very useful, for example, in a priori error estimates and essential in obtaining error estimates where the spatial mesh size h and the time steps k are independent of each other (cf. [30, 31]).

Previously in [33] we established the corresponding discrete maximal parabolic regularity for discontinuous Galerkin time schemes of arbitrary order for autonomous problems. The extension to nonautonomous problems is not straightforward, especially for the critical values of $s, p = 1$ or $s, p = \infty$. In this paper, we investigate the maximal parabolic regularity for $s = 2$ and arbitrary $1 \leq p \leq \infty$ for the lowest order time discontinuous Galerkin (dG(0)) methods, which can be considered as a modified backward Euler (BE) method. The main difference between dG(0) and BE methods lies in the way the time-dependent coefficients and the right-hand side are approximated. In the dG(0) formulation they are approximated by averages over each subinterval I_m (see the details below) while in BE methods they are evaluated at time nodes. As a result, dG(0) approximations are weakly consistent, i.e., they satisfy the Galerkin orthogonality relation; see section 2 for details.

Parabolic problems with time-dependent coefficients are important, for example, for analyzing quasi-linear problems. Over the years, there has been a considerable number of publications devoted to various numerical methods for problems with time-dependent coefficients [5, 10, 12, 20, 22, 23, 24, 35, 36, 38, 39]. The publication [5] is the most relevant to our presentation since it treats discontinuous Galerkin methods. However, none of the above publications addresses the question of the discrete maximal parabolic regularity and the techniques used in those papers are not immediately applicable for establishing such results even for $p = 2$.

Time discrete maximal parabolic regularity (sometimes called well-posedness or coercivity property in the literature) has been investigated in a number of publications for various time schemes [1, 2, 14, 15, 16, 25, 33]. However, all the above-mentioned publications are dealing with the autonomous case. We are only aware of the publications [28, 34], where the discrete maximal parabolic regularity is established for problems with time-dependent coefficients for the BE [34] and for more general Runge–Kutta schemes [28]. Although the results in [28, 34] are similar in nature, there are significant differences:

- Our result is valid for the limiting cases $p = 1$ and $p = \infty$, whereas the results from [28, 34] are established for $1 < p < \infty$. As far as we can see, the arguments there cannot be naturally extended to the critical values of $p = 1$ and $p = \infty$ even with the expense of the logarithmic term since the proofs there require Gronwall's inequality. The inclusion of these critical cases is important for the derivation of best approximation type estimates in the $L^\infty(I; L^2(\Omega))$ norm (cf. Theorem 4.5), which are often required in the context of optimal control problems (see, e.g., [37]) and which are not available in the literature for nonautonomous problems.
- Our discrete maximal parabolic regularity estimates are given with respect to the $L^p(I; L^2(\Omega))$ norm, whereas the corresponding results from [28, 34] are formulated in terms of discrete time values describing a weighted $l^p(0, T; L^2(\Omega))$ norm. We require the results for the integrated norm to prove best approximation type error estimates (cf. Theorem 4.5) by applying a duality argument.

Moreover, the results in [28, 34] require $a_{ij}(t, x) \in W^1_\infty(I \times \Omega)$, smoothness of Ω, and treat only uniform time steps, but they are valid in $L^s(\Omega)$ norms in space for $1 < s < \infty$. Our results, on the other hand, require only a Hölder continuity of $a_{ij}(t, x)$ in t and L^∞ in space, allow Ω to be merely Lipschitz, and treat variable time steps, but are valid only in $L^2(\Omega)$ norm in space. We also want to mention that we went through some technical obstacles in order to incorporate variable time steps. In the case of uniform time steps many arguments can be significantly simplified.

Our presentation is inspired by [17], where the maximal parabolic regularity was established for continuous problems for $s = 2$ and $1 < p < \infty$ with rather weak assumptions on A. In particular, we show that for the dG(0) method the semidiscrete solution u_k on any time level m for $u_0 = 0$ and $f \in L^\infty(I; L^2(\Omega))$ satisfies

$$\|A_{k,m}u_{k,m}\|_{L^\infty(I_m;L^2(\Omega))} + \left\|\frac{[u_k]_{m-1}}{k_m}\right\|_{L^2(\Omega)} \le C \ln\frac{T}{k}\|f\|_{L^\infty(I;L^2(\Omega))}. \tag{1.7}$$

For $p = 1$ with $u_0 \in L^2(\Omega)$ and $f \in L^1(I; L^2(\Omega))$ we also obtain

$$\sum_m \left(\|A_{k,m}u_{k,m}\|_{L^1(I_m;L^2(\Omega))} + \|[u_k]_{m-1}\|_{L^2(\Omega)}\right) \le C \ln\frac{T}{k}\left(\|f\|_{L^1(I;L^2(\Omega))} + \|u_0\|_{L^2(\Omega)}\right), \tag{1.8}$$

where k_m is the time step on subinterval I_m and $A_{k,m}$ is the average of $A(t)$ on I_m. (See section 2 for a detailed description.) In contrast to the continuous estimate (1.6), the above estimates include the limiting cases of $p = \infty$ and $p = 1$, which explains the logarithmic factor in (1.7) and (1.8).

The corresponding results also hold for the fully discrete approximation u_{kh}. Thus in particular for $1 \le p \le \infty$ and $u_0 = 0$, we establish

$$\left[\sum_m \left(\|A_{kh,m}u_{kh,m}\|^p_{L^p(I_m;L^2(\Omega))} + k_m\left\|\frac{[u_{kh}]_{m-1}}{k_m}\right\|^p_{L^2(\Omega)}\right)\right]^{\frac{1}{p}} \le C \ln\frac{T}{k}\|f\|_{L^p(I;L^2(\Omega))} \tag{1.9}$$

with corresponding changes for $p = \infty$. We would like to point out that the above fully discrete result is valid on rather general meshes and does not require the mesh to be quasi-uniform or even shape regular, only admissible (no hanging nodes).

As an application of the discrete maximal parabolic regularity we show that if the coefficients $a_{ij}(t, x)$ are sufficiently regular (see Assumption 1) and Ω convex we obtain symmetric error estimate

$$\|u-u_k\|_{L^p(I;L^2(\Omega))} \le C\ln\frac{T}{k}\|u-\pi_k u\|_{L^p(I;L^2(\Omega))}, \quad 1\le p<\infty,$$

where π_k is an interpolation into the space of piecewise constant in time functions defined in (4.5). For $p=\infty$ we can establish even the best approximation type result

$$\|u-u_k\|_{L^\infty(I;L^2(\Omega))} \le C\ln\frac{T}{k}\|u-\chi\|_{L^\infty(I;L^2(\Omega))}$$

for any χ in the subspace of piecewise constant in time functions; see Theorem 4.2. The corresponding fully discrete versions are

$$\|u-u_{kh}\|_{L^p(I;L^2(\Omega))} \le C\ln\frac{T}{k}\left(\|u-\pi_k u\|_{L^p(I;L^2(\Omega))}+\|u-R_h u\|_{L^p(I;L^2(\Omega))}\right)$$

and

$$\|u-u_{kh}\|_{L^\infty(I;L^2(\Omega))} \le C\ln\frac{T}{k}\left(\|u-\chi\|_{L^\infty(I;L^2(\Omega))}+\|u-R_h u\|_{L^\infty(I;L^2(\Omega))}\right)$$

for any χ in the subspace and $R_h(t)$ being the Ritz projection corresponding to $A(t)$. The rate of convergence depends of course on the regularity of u.

The rest of the paper is organized as follows. In section 2 we introduce the discontinuous Galerkin method and some notation. Section 3, which is the central piece of the paper, consists of several parts. First we write the dG(0) approximate solution u_k in a convenient form. Then we introduce a transform function w_k that satisfies a similar equation, but with transform operators. Then in a series of lemmas we show that the resulting operators are bounded in certain norms. Finally in sections 3.2 and 3.3 we establish semidiscrete and fully discrete maximal parabolic regularity in $L^p(I;L^2(\Omega))$ norms, respectively. We conclude our paper with section 4, where we show how the above discrete maximal parabolic regularity results can be used to establish symmetric and best approximation type error estimates for the problems on convex domains with coefficients satisfying some additional assumptions.

2. Preliminaries. First, we introduce the bilinear form $a\colon \mathbb{R}\times H_0^1(\Omega)\times H_0^1(\Omega)\to\mathbb{R}$ defined by

$$a(t;u,v)=\int_\Omega\sum_{i,j=1}^d a_{ij}(t,x)\partial_i u(x)\partial_j v(x)dx. \tag{2.1}$$

From $\{a_{ij}(t,x)\}_{i,j=1}^d\subset L^\infty(I\times\Omega)$ one can see that for each $t\in I$ the bilinear form $a(t;\cdot,\cdot)$ is bounded

$$a(t;u,v)\le C\|u\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)}, \tag{2.2}$$

from the uniform ellipticity assumption (1.3), it is coercive

$$a(t;u,u)\ge\alpha\|u\|^2_{H^1(\Omega)}, \tag{2.3}$$

and from (1.4) it follows that

$$|a(t_1;u,v)-a(t_2;u,v)|\le C\omega(|t_1-t_2|)\|u\|_{H^1(\Omega)}\|v\|_{H^1(\Omega)}. \tag{2.4}$$

In view of the homogeneous Dirichlet boundary conditions the H^1 norm is equivalent to the H^1 seminorm. For each $t \in \bar{I}$ this bilinear form defines an operator $A(t)\colon H_0^1(\Omega) \to H^{-1}(\Omega)$ by

$$\langle A(t)u, v\rangle = a(t,u,v) \quad \forall u, v \in H_0^1(\Omega),$$

where $\langle\cdot,\cdot\rangle$ is the duality pairing between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$ spaces.

To introduce the time discontinuous Galerkin discretization for the problem, we partition $I = (0,T]$ 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 and minimal time steps are denoted by $k = \max_m k_m$ and $k_{\min} = \min_m k_m$, respectively. We impose the following conditions on the time mesh.

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

$$k_{\min} \geq ck^{\beta}.$$

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

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

(iii) It holds that $k \leq \frac{1}{4}T$.

Similar assumptions are made, e.g., in [37]. The semidiscrete space X_k^0 of piecewise constant functions in time is defined by

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

where $\mathcal{P}_0(V)$ is the space of constant functions in time with values in a Banach space V. We will employ the notation

$$v_m^+ := \lim_{t\to 0^+} v(t_m + t), \quad v_-^+ := \lim_{t \to 0^+} v(t_m - t), \quad \text{and} \quad [v]_m = v_m^+ - v_m^-$$

if these limits exist. For a function v_k from X_k^0 we denote $v_{k,m} := v_k|_{I_m}$ resulting in

$$v_{k,m}^+ = v_{k,m+1}, \quad v_{k,m}^- = v_{k,m}, \quad \text{and} \quad [v_k]_m = v_{k,m+1} - v_{k,m}$$

for $m = 1, 2, \ldots, M-1$.

Next we define the following bilinear form:

(2.5)

$$B(u,\varphi) = \sum_{m=1}^{M} \langle \partial_t u, \varphi\rangle_{I_m\times\Omega} + \sum_{i,j=1}^{d} (a_{ij}\partial_i u, \partial_j \varphi)_{I\times\Omega} + \sum_{m=2}^{M} ([u]_{m-1}, \varphi_{m-1}^+)_\Omega + (u_0^+, \varphi_0^+)_\Omega,$$

where $(\cdot,\cdot)_\Omega$ and $(\cdot,\cdot)_{I_m\times\Omega}$ are the usual L^2 space and space-time inner-products, and $\langle\cdot,\cdot\rangle_{I_m\times\Omega}$ is the duality product between $L^2(I_m; H^{-1}(\Omega))$ and $L^2(I_m; H_0^1(\Omega))$. Rearranging the terms in (2.5), we obtain an equivalent (dual) expression for B,

(2.6)

$$B(u,\varphi) = -\sum_{m=1}^{M} \langle u, \partial_t \varphi\rangle_{I_m\times\Omega} + \sum_{i,j=1}^{d} (a_{ij}\partial_i u, \partial_j\varphi)_{I\times\Omega} - \sum_{m=1}^{M-1} (u_m^-, [\varphi]_m)_\Omega + (u_M^-, \varphi_M^-)_\Omega.$$

We note that for $u_k, \varphi_k \in X_k^0$ the bilinear form (2.5) simplifies to

$$B(u_k,\varphi_k) = \sum_{i,j=1}^{d} (a_{ij}\partial_i u_k, \partial_j\varphi_k)_{I\times\Omega} + \sum_{m=2}^{M} ([u_k]_{m-1}, \varphi_{k,m})_\Omega + (u_{k,1}, \varphi_{k,1})_\Omega$$

and

$$B(u_k,\varphi_k)=\sum_{i,j=1}^{d}(a_{ij}\partial_i u_k,\partial_j\varphi_k)_{I\times\Omega}-\sum_{m=1}^{M-1}(u_{k,m}^-,[\varphi_k]_m)_\Omega+(u_{k,M}^-,\varphi_{k,M}^-)_\Omega.$$

The dG(0) semidiscrete (in time) approximation $u_k\in X_k^0$ of (1.1) is defined as

$$B(u_k,\varphi_k)=(f,\varphi_k)_{I\times\Omega}+(u_0,\varphi_{k,1})_\Omega\quad\forall\varphi_k\in X_k^0, \tag{2.7}$$

and by the construction we have the following Galerkin orthogonality:

$$B(u-u_k,\varphi_k)=0\quad\forall\varphi_k\in X_k^0. \tag{2.8}$$

To rewrite the dG(0) method as a time-stepping scheme we introduce the following notation. We define $f_k\in X_k^0$ by

$$f_{k,m}=\frac{1}{k_m}\int_{I_m}f(t)dt,\quad m=1,2,\dots,M \tag{2.9}$$

and $A_{k,m}\colon H_0^1(\Omega)\to H^{-1}(\Omega)$ by

$$A_{k,m}=\frac{1}{k_m}\int_{I_m}A(t)dt,\quad m=1,2,\dots,M. \tag{2.10}$$

Thus, the dG(0) solution u_k satisfies

$$\begin{aligned}u_{k,1}+k_1A_{k,1}u_{k,1}&=u_0+k_1f_{k,1},\\ u_{k,m}+k_mA_{k,m}u_{k,m}&=u_{k,m-1}+k_mf_{k,m},\quad m=2,3,\dots,M.\end{aligned} \tag{2.11}$$

To use results known for the autonomous problems we rewrite this formula for some fixed $2\le m\le M$ as

$$\begin{aligned}u_{k,1}+k_1A_{k,m}u_{k,1}&=u_0+k_1f_{k,1}+k_1(A_{k,m}-A_{k_1})u_{k,1},\\ u_{k,l}+k_lA_{k,m}u_{k,l}&=u_{k,l-1}+k_lf_{k,l}+k_l(A_{k,m}-A_{k_l})u_{k,l},\quad l=2,3,\dots,m.\end{aligned}$$

Then using the representation formula for the dG(0) solution in the autonomous case (cf. the proof of Theorem 2.1 in [33]), we obtain the following representation:

$$u_{k,m}=\sum_{l=1}^{m-1}k_lR_{m,l}(A_{k,m}-A_{k,l})u_{k,l}+\sum_{l=1}^{m}k_lR_{m,l}f_{k,l}+R_{m,1}u_0,\quad m=1,2,\dots,M, \tag{2.12}$$

where

$$R_{m,l}=\prod_{j=1}^{m-l+1}r(k_{m+1-j}A_{k,m})\quad\text{and}\quad r(z)=\frac{1}{1+z}. \tag{2.13}$$

Throughout the paper we use a convention $\sum_{l=1}^{0}=0$. Next we define three operators $Q\colon X_k^0\to X_k^0$, $L\colon X_k^0\to X_k^0$, and $D\colon L^2(\Omega)\to X_k^0$ by

$$(Qg_k)_m=\sum_{l=1}^{m-1}k_lA_{k,m}R_{m,l}(A_{k,m}-A_{k,l})A_{k,l}^{-1}g_{k,l}\quad\text{for } g_k\in X_k^0, \tag{2.14}$$

$$(Lf_k)_m = \sum_{l=1}^{m} k_l A_{k,m} R_{m,l} f_{k,l}, \tag{2.15}$$

and

$$(Du_0)_m = A_{k,m} R_{m,1} u_0. \tag{2.16}$$

Thus, for $v_k \in X_k^0$ defined by

$$v_{k,l} = A_{k,l} u_{k,l} \quad \text{for} \quad l = 1, 2, \ldots, M,$$

we have

$$v_k = Q v_k + L f_k + D u_0. \tag{2.17}$$

3. Maximal parabolic regularity for time discretization.

3.1. Estimate for the transformed operator. Let $\mu > 0$ be a sufficiently large number to be chosen later. Define $w_{k,m}$ by

$$w_{k,m} = \prod_{l=1}^{m} (1 + \mu k_l)^{-1} u_{k,m}, \quad m = 1, 2, \ldots, M.$$

Thus using (2.11) we obtain

$$\begin{aligned} (1 + \mu k_1) w_{k,1} + k_1 (1 + \mu k_1) A_{k,1} w_{k,1} &= u_0 + k_1 f_{k,1}, \\ \prod_{l=1}^{m} (1 + \mu k_l) w_{k,m} + k_m \prod_{l=1}^{m} (1 + \mu k_l) A_{k,m} w_{k,m} &= \prod_{l=1}^{m-1} (1 + \mu k_l) w_{k,m-1} + k_m f_{k,m} \end{aligned}$$

for $m = 2, \ldots, M$. Dividing both sides of the last equation by $\prod_{l=1}^{m-1}(1 + \mu k_l)$, we obtain

$$(1 + k_m \mu) w_{k,m} + k_m (1 + k_m \mu) A_{k,m} w_{k,m} = w_{k,m-1} + \prod_{l=1}^{m-1} (1 + \mu k_l)^{-1} k_m f_{k,m}.$$

Hence, we can rewrite (2.11) as

$$\begin{aligned} w_{k,1} + k_1 \widetilde{A}_{k,1} w_{k,1} &= u_0 + k_1 \widetilde{f}_{k,1}, \\ w_{k,m} + k_m \widetilde{A}_{k,m} w_{k,m} &= w_{k,m-1} + k_m \widetilde{f}_{k,m}, \quad m = 2, \ldots, M, \end{aligned}$$

where

$$\widetilde{A}_{k,m} = (1 + k_m \mu) A_{k,m} + \mu \,\mathrm{Id}, \quad \widetilde{f}_{k,m} = \prod_{l=1}^{m-1} (1 + \mu k_l)^{-1} f_{k,m}, \quad m = 1, 2, \ldots, M. \tag{3.1}$$

Here we use a convention $\prod_{l=1}^{0} = 1$. Similarly to (2.17), for $\tilde{v}_k \in X_k^0$ defined by

$$\tilde{v}_{k,l} = \widetilde{A}_{k,l} w_{k,l} \quad \text{for} \quad l = 1, \ldots, M,$$

we have

$$\tilde{v}_k = \widetilde{Q} \tilde{v}_k + \widetilde{L} \widetilde{f}_k + \widetilde{D} u_0, \tag{3.2}$$

where similarly to the definitions of Q, L, and D above,

$$\widetilde{R}_{m,l} = \prod_{j=1}^{m-l+1} r(k_{m+1-j}\widetilde{A}_{k,m}), \tag{3.3}$$

$$(\widetilde{Q}g_k)_m = \sum_{l=1}^{m-1} k_l \widetilde{A}_{k,m}\widetilde{R}_{m,l}(\widetilde{A}_{k,m} - \widetilde{A}_{k,l})\widetilde{A}_{k,l}^{-1} g_{k,l} \tag{3.4}$$

and

$$(\widetilde{L}\widetilde{f}_k)_m = \sum_{l=1}^{m} k_l \widetilde{A}_{k,m}\widetilde{R}_{m,l}\widetilde{f}_{k,l}, \quad (\widetilde{D}u_0)_m = \widetilde{A}_{k,m}\widetilde{R}_{m,1}u_0. \tag{3.5}$$

Using the ellipticity of $A_{k,m}$ we obtain the following resolvent estimate for $\widetilde{A}_{k,m}$. For a given $\gamma \in (0, \pi/2)$ we define

$$\Sigma_\gamma = \{z \in \mathbb{C} : |\arg(z)| \le \gamma\}. \tag{3.6}$$

Moreover we introduce the complex spaces $\mathbb{H} = L^2(\Omega) + iL^2(\Omega)$ and $\mathbb{V} = H_0^1(\Omega) + iH_0^1(\Omega)$.

LEMMA 3.1. *For any $\gamma > 0$, there exists a constant C independent of k and μ such that*

$$\|(z - \widetilde{A}_{k,m})^{-1}v\|_{L^2(\Omega)} \le \frac{C}{|z| + \mu}\|v\|_{L^2(\Omega)} \quad \forall z \in \mathbb{C} \setminus \Sigma_\gamma, \quad \forall v \in \mathbb{H}$$

and

$$\|\widetilde{A}_{k,m}(z - \widetilde{A}_{k,m})^{-1}v\|_{L^2(\Omega)} \le C\|v\|_{L^2(\Omega)}, \quad \forall z \in \mathbb{C} \setminus \Sigma_\gamma, \quad \forall v \in \mathbb{H}.$$

Proof. For an arbitrary $v \in \mathbb{H}$ we define

$$g = -(z - \widetilde{A}_{k,m})^{-1}v \in \mathbb{V},$$

or equivalently

$$-z(g,\varphi) + (1 + k_m\mu)(A_{k,m}g,\varphi) + \mu(g,\varphi) = (v,\varphi) \quad \forall \varphi \in \mathbb{V}, \tag{3.7}$$

whose existence and uniqueness follow from the Fredholm alternative. In this proof $(\cdot,\cdot)$ denotes the Hermitian inner product, i.e., $(v,\varphi) = \int_\Omega v\bar{\varphi}\, dx$.

Taking $\varphi = g$ we obtain

$$-z\|g\|^2_{L^2(\Omega)} + (1 + k_m\mu)(A_{k,m}g, g) + \mu\|g\|^2_{L^2(\Omega)} = (v, g). \tag{3.8}$$

Since $\gamma \le |\arg z| \le \pi$ and $\alpha\|g\|^2_{H^1(\Omega)} \le (A_{k,m}g, g)$ and is real, this equation is of the form

$$e^{i\delta}a + b = c \quad \text{with} \quad a, b > 0, \quad \gamma \le |\delta| \le \pi,$$

and by multiplying it by $e^{-\frac{i\delta}{2}}$ and taking real parts, we have

$$a + b \le \left(\cos\left(\frac{\delta}{2}\right)\right)^{-1} |(v,g)| \le \left(\sin\left(\frac{\gamma}{2}\right)\right)^{-1} |(v,g)| = C_\gamma|(v,g)|.$$

From (3.8) we therefore conclude

$$(|z| + \mu)\,\|g\|^2_{L^2(\Omega)} + \alpha(1 + k_m\mu)\|g\|^2_{H^1(\Omega)} \le C_\gamma\|g\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)} \quad \text{for } z \in \mathbb{C} \setminus \Sigma_\gamma.$$

Thus, we have

$$\|g\|_{L^2(\Omega)} \leq \frac{C_\gamma}{|z|+\mu}\|v\|_{L^2(\Omega)},$$

which establishes the first result. The second result follows from the identity

$$\widetilde{A}_{k,m}(z-\widetilde{A}_{k,m})^{-1} = -\operatorname{Id} + z(z-\widetilde{A}_{k,m})^{-1}. \qquad \square$$

LEMMA 3.2. *There exists a constant C independent of k and μ such that*

$$\|(\widetilde{A}_{k,m})^{-1}v\|_{L^2(\Omega)} \leq \frac{1}{\mu}\|v\|_{L^2(\Omega)}$$

and

$$\|(\widetilde{A}_{k,m})^{-1}v\|_{H^1(\Omega)} \leq \frac{C}{\sqrt{\mu}(1+k_m\mu)^{\frac{1}{2}}}\|v\|_{L^2(\Omega)}.$$

Proof. For an arbitrary $v \in L^2(\Omega)$, we define

$$g = (\widetilde{A}_{k,m})^{-1}v,$$

or equivalently

$$(\widetilde{A}_{k,m}g,\varphi) = (1+k_m\mu)(A_{k,m}g,\varphi) + \mu(g,\varphi) = (v,\varphi) \quad \forall \varphi \in H^1_0(\Omega). \tag{3.9}$$

Taking $\varphi = g$ and using the coercivity (2.3), we obtain

$$\alpha(1+k_m\mu)\|g\|^2_{H^1(\Omega)} + \mu\|g\|^2_{L^2(\Omega)} \leq \|v\|_{L^2(\Omega)}\|g\|_{L^2(\Omega)}. \tag{3.10}$$

From the estimate above, we immediately conclude that

$$\|g\|_{L^2(\Omega)} \leq \frac{1}{\mu}\|v\|_{L^2(\Omega)} \tag{3.11}$$

and using (3.11), we also obtain

$$\alpha(1+k_m\mu)\|g\|^2_{H^1(\Omega)} \leq \|v\|_{L^2(\Omega)}\|g\|_{L^2(\Omega)} \leq \frac{1}{\mu}\|v\|^2_{L^2(\Omega)},$$

from which the second estimate of the lemma follows. $\square$

We will also require the following result that estimates the difference $\widetilde{A}_{k,m} - \widetilde{A}_{k,l}$.

LEMMA 3.3. *There exists a constant C independent of k and μ such that for $m \geq l$*

$$\|(\widetilde{A}_{k,m}-\widetilde{A}_{k,l})v\|_{H^{-1}(\Omega)} \leq C\left((1+\mu\min\{k_l,k_m\})\omega(t_m-t_{l-1}) + \mu|k_m-k_l|\right)\|v\|_{H^1(\Omega)}.$$

Proof. By duality we have

$$\|(\widetilde{A}_{k,m}-\widetilde{A}_{k,l})v\|_{H^{-1}(\Omega)} = \sup_{\substack{w\in H^1_0(\Omega),\\ \|w\|_{H^1(\Omega)}\leq 1}} ((\widetilde{A}_{k,m}-\widetilde{A}_{k,l})v,w)_\Omega.$$

For each such w, we have

$$((\widetilde{A}_{k,m}-\widetilde{A}_{k,l})v,w)_\Omega = \mu\left((k_mA_{k,m}-k_lA_{k,l})v,w\right)_\Omega + ((A_{k,m}-A_{k,l})v,w)_\Omega = J_1+J_2.$$

Using the definitions of $A_{k,m}$ and $A_{k,l}$, changing variables, and using (1.4) and that ω is nondecreasing, we have

$$
\begin{aligned}
J_2 &= \left(\Big(\frac{1}{k_m}\int_{t_{m-1}}^{t_m} A(t)dt - \frac{1}{k_l}\int_{t_{l-1}}^{t_l} A(t)dt\Big)v, w\right)_\Omega \\
&= \int_0^1 \Big((A(sk_m + t_{m-1}) - A(sk_l + t_{l-1}))v, w\Big)_\Omega ds \\
&\le \int_0^1 \omega(|sk_m + t_{m-1} - sk_l - t_{l-1}|)\|v\|_{H^1(\Omega)}\|w\|_{H^1(\Omega)}ds \\
&\le \omega(t_m - t_{l-1})\|v\|_{H^1(\Omega)}\|w\|_{H^1(\Omega)}.
\end{aligned}
\tag{3.12}
$$

To estimate J_1 we use

$$k_m A_{k,m} - k_l A_{k,l} = k_m(A_{k,m} - A_{k,l}) + (k_m - k_l)A_{k,l}$$

or

$$k_m A_{k,m} - k_l A_{k,l} = k_l(A_{k,m} - A_{k,l}) + (k_m - k_l)A_{k,m}.$$

Then using (3.12), we obtain

$$J_1 \le C\mu\left(\min\{k_l, k_m\}\omega(t_m - t_{l-1})\|v\|_{H^1(\Omega)}\|w\|_{H^1(\Omega)} + |k_m - k_l|\|v\|_{H^1(\Omega)}\|w\|_{H^1(\Omega)}\right).$$

Combining the estimates for J_1 and J_2, we obtain the lemma. ∎

LEMMA 3.4. *There exists a constant C independent of k and μ such that*

$$\|\widetilde{A}_{k,m}\widetilde{R}_{m,l}v\|_{L^2(\Omega)} \le \frac{C}{t_m - t_{l-1}}\|v\|_{L^2(\Omega)} \quad \forall v \in L^2(\Omega).$$

Moreover, for $m - l \ge 1$ there holds

$$\|\widetilde{A}_{k,m}^2\widetilde{R}_{m,l}v\|_{L^2(\Omega)} \le \frac{C}{(t_m - t_{l-1})^2}\|v\|_{L^2(\Omega)} \quad \forall v \in L^2(\Omega).$$

Proof. First we observe that each term $\widetilde{R}_{m,l}v$ can be thought of as $m-l+1$ time steps of the dG(0) method of the homogeneous problem

$$\partial_t u + \widetilde{A}_{k,m}u = 0$$

with the initial condition $u(t_{l-1}) = v$. Then, using Lemma 3.1 the first estimate follows from [11] (cf. also [33, Theorem 1]). To prove the second estimate we use a representation

$$\widetilde{A}_{k,m}^2\widetilde{R}_{m,l}v = g(\widetilde{A}_{k,m})v$$

with the function

$$g(\lambda) = \lambda^2\prod_{j=l}^{m} r(k_j\lambda).$$

Using the fact that the spectrum $\sigma(\widetilde{A}_{k,m})$ of $\widetilde{A}_{k,m}$ is real and positive we obtain by Parseval's identity (cf. [40, Chapter 7])

$$\|\widetilde{A}_{k,m}^2\widetilde{R}_{m,l}v\|_{L^2(\Omega)} \le \sup_{\lambda\in\sigma(\widetilde{A}_{k,m})}|g(\lambda)|\;\|v\|_{L^2(\Omega)}.$$

To estimate $|g(\lambda)|$ we proceed similarly to the proof of Theorem 5.1 in [11] and observe

$$\prod_{j=l}^{m}(1+k_j\lambda) \geq 1+\lambda\sum_{j=l}^{m}k_j+\frac{\lambda^2}{2}\left(\sum_{\substack{i,j=l\\ i\neq j}}^{m}k_ik_j\right) = 1+\lambda(t_m-t_{l-1})+\frac{\lambda^2}{2}\left(\sum_{\substack{i,j=l\\ i\neq j}}^{m}k_ik_j\right).$$

Let $k_{\max} = \max_{l\leq j\leq m} k_j$ and first consider the case $k_{\max} < (t_m - t_{l-1})/2$. We have

$$(t_m-t_{l-1})^2 = \left(\sum_{j=l}^{m}k_j\right)^2 = \sum_{j=l}^{m}k_j^2+\sum_{\substack{i,j=l\\ i\neq j}}^{m}k_ik_j \leq k_{\max}(t_m-t_{l-1})+\sum_{\substack{i,j=l\\ i\neq j}}^{m}k_ik_j,$$

and with the assumption $k_{\max} < (t_m - t_{l-1})/2$,

$$\sum_{\substack{i,j=l\\ i\neq j}}^{m}k_ik_j \geq \frac{(t_m-t_{l-1})^2}{2}.$$

This results in

$$\prod_{j=l}^{m}(1+k_j\lambda) \geq 1+\lambda(t_m-t_{l-1})+\frac{\lambda^2}{4}(t_m-t_{l-1})^2$$

and therefore we have

$$|g(\lambda)| \leq \frac{\lambda^2}{1+\lambda(t_m-t_{l-1})+\frac{\lambda^2}{4}(t_m-t_{l-1})^2} \leq \frac{4}{(t_m-t_{l-1})^2},$$

which proves the assertion in this case. In the case $k_{\max} \geq (t_m - t_{l-1})/2$ let $l \leq m_0 \leq m$ be such that $k_{m_0} = k_{\max}$. Due to $m-l \geq 1$ we can choose m_0' as either $m_0 - 1$ or $m_0 + 1$ such that $l \leq m_0' \leq m$. Then we obtain

$$\begin{aligned}\sup_{\lambda\in\sigma(\widetilde{A}_{k,m})}|\lambda^2\prod_{j=l}^{m}r(k_j\lambda)| &\leq \sup_{\lambda\in\sigma(\widetilde{A}_{k,m})}\left|\frac{\lambda^2}{(1+k_{m_0}\lambda)(1+k_{m_0'}\lambda)}\right|\\ &\leq \frac{1}{k_{m_0}\,k_{m_0'}} \leq \frac{C}{k_{\max}^2} \leq \frac{C}{(t_m-t_{l-1})^2},\end{aligned}$$

where we have used our assumption (ii) on the time steps. This completes the proof for this case. ∎

Lemma 3.5. *There exists a constant C independent of k and μ such that for $m-l \geq 1$,*

$$\|\widetilde{A}_{k,m}\widetilde{R}_{m,l}v\|_{H^1(\Omega)} \leq \frac{C}{(t_m-t_{l-1})^{\frac{3}{2}}(1+\mu k_m)^{\frac{1}{2}}}\|v\|_{L^2(\Omega)} \quad \forall v\in L^2(\Omega).$$

Proof. By the coercivity of the operator A for any $w \in H_0^1(\Omega)$, we have

$$(\widetilde{A}_{k,m}w,w) \geq (1+\mu k_m)(A_{k,m}w,w) \geq (1+\mu k_m)\alpha\|w\|_{H^1(\Omega)}^2.$$

Thus, with $w = \widetilde{A}_{k,m}\widetilde{R}_{m,l}v$, we have by the previous lemma

$$\begin{aligned}(1+\mu k_m)\alpha\|\widetilde{A}_{k,m}\widetilde{R}_{m,l}v\|^2_{H^1(\Omega)} &\le (\widetilde{A}_{k,m}\widetilde{A}_{k,m}\widetilde{R}_{m,l}v, \widetilde{A}_{k,m}\widetilde{R}_{m,l}v)\\ &\le \|\widetilde{A}^2_{k,m}\widetilde{R}_{m,l}v\|_{L^2(\Omega)}\|\widetilde{A}_{k,m}\widetilde{R}_{m,l}v\|_{L^2(\Omega)}\\ &\le \frac{C}{(t_m - t_{l-1})^3}\|v\|_{L^2(\Omega)}.\end{aligned}$$

This completes the proof. ∎

LEMMA 3.6. *There exists a constant C independent of k and μ such that for $m - l \ge 1$*

$$\|\widetilde{A}_{k,m}\widetilde{R}_{m,l}v\|_{L^2(\Omega)} \le \frac{C}{(t_m - t_{l-1})^{\frac{3}{2}}(1+\mu k_m)^{\frac{1}{2}}}\|v\|_{H^{-1}(\Omega)} \quad \forall v \in H^{-1}(\Omega).$$

Proof. By duality

$$\|\widetilde{A}_{k,m}\widetilde{R}_{m,l}v\|_{L^2(\Omega)} = \sup_{\substack{w\in L^2(\Omega),\\ \|w\|_{L^2(\Omega)}\le 1}} (\widetilde{A}_{k,m}\widetilde{R}_{m,l}v, w)_\Omega.$$

Since A is a symmetric operator, we have $\widetilde{A}_{k,m} = \widetilde{A}^*_{k,m}$ and as a result $\widetilde{R}_{m,l} = \widetilde{R}^*_{m,l}$. Moreover $\widetilde{A}_{k,m}$ and $\widetilde{R}_{m,l}$ commute. Thus,

$$(\widetilde{A}_{k,m}\widetilde{R}_{m,l}v, w)_\Omega = (v, \widetilde{A}^*_{k,m}\widetilde{R}^*_{m,l}w)_\Omega \le \|v\|_{H^{-1}(\Omega)}\|\widetilde{A}^*_{k,m}\widetilde{R}^*_{m,l}w\|_{H^1(\Omega)}.$$

Since $\widetilde{A}^*_{k,m} = \widetilde{A}_{k,m}$, by Lemma 3.5, we obtain

$$\|\widetilde{A}^*_{k,m}\widetilde{R}^*_{m,l}w\|_{H^1(\Omega)} \le \frac{C}{(t_m - t_{l-1})^{\frac{3}{2}}(1+\mu k_m)^{\frac{1}{2}}}\|w\|_{L^2(\Omega)},$$

which establishes the lemma. ∎

Combining the above lemmas we obtain the following result.

LEMMA 3.7. *There exist constants C_1 and C_2 independent of μ and k such that for any $g_k \in X^0_k$*

$$\begin{aligned}&\|(\widetilde{Q}g_k)_m\|_{L^2(\Omega)}\\ &\quad\le \max_{1\le j\le m}\|g_{k,j}\|_{L^2(\Omega)}\left(\frac{C_1}{\sqrt{\mu}} + C_2\sqrt{\mu}\sum_{l=1}^{m-1}\frac{k_l|k_m - k_l|}{(t_m - t_{l-1})^{\frac{3}{2}}}\right), \quad m = 1,\dots,M,\end{aligned}$$

and

$$\sum_{m=1}^{M} k_m\|(\widetilde{Q}g_k)_m\|_{L^2(\Omega)} \le \left(\sum_{l=1}^{M} k_l\|g_{k,l}\|_{L^2(\Omega)}\right)\left(\frac{C_1}{\sqrt{\mu}} + C_2\sqrt{\mu}\max_{1\le l\le M}\sum_{m=l+1}^{M}\frac{k_m|k_m - k_l|}{(t_m - t_{l-1})^{\frac{3}{2}}}\right),$$

where $\widetilde{Q}$ is the operator defined in (3.4).

Proof. Using that

$$(\widetilde{Q}g_k)_m = \sum_{l=1}^{m-1} k_l\widetilde{A}_{k,m}\widetilde{R}_{m,l}(\widetilde{A}_{k,m} - \widetilde{A}_{k,l})\widetilde{A}^{-1}_{k,l}g_{k,l},$$

we have

$$\|(\widetilde{Q}g_k)_m\|_{L^2(\Omega)} \leq \sum_{l=1}^{m-1} k_l \|\widetilde{A}_{k,m}\widetilde{R}_{m,l}\|_{H^{-1}\to L^2}\|\widetilde{A}_{k,m}-\widetilde{A}_{k,l}\|_{H^1\to H^{-1}}\|(\widetilde{A}_{k,l})^{-1}\|_{L^2\to H^1}\|g_{k,l}\|_{L^2(\Omega)}.$$

Combining estimates from Lemmas 3.2, 3.3, and 3.6 and using that

$$\frac{(1+\mu\min\{k_l,k_m\})}{(1+\mu k_l)^{\frac{1}{2}}(1+\mu k_m)^{\frac{1}{2}}} \leq 1$$

for any $m = 1, 2, \dots, M$ we have

$$\|(\widetilde{Q}g_k)_m\|_{L^2(\Omega)} \leq C\sum_{l=1}^{m-1}\left(\frac{k_l}{\sqrt{\mu}}\frac{\omega(t_m-t_{l-1})}{(t_m-t_{l-1})^{\frac{3}{2}}} + \sqrt{\mu}k_l\frac{|k_m-k_l|}{(t_m-t_{l-1})^{\frac{3}{2}}}\right)\|g_{k,l}\|_{L^2(\Omega)}. \tag{3.13}$$

From the condition (1.5) and properties of the Riemann sums, we obtain

$$\sum_{l=1}^{m-1} k_l\frac{\omega(t_m-t_{l-1})}{(t_m-t_{l-1})^{\frac{3}{2}}} \leq \int_0^{t_m}\frac{\omega(t_m-s)}{(t_m-s)^{\frac{3}{2}}}\,ds \leq C,$$

and taking the maximum over l of $\|g_{k,l}\|_{L^2(\Omega)}$, we obtain the first estimate of the lemma.

From (3.13) we also obtain

$$\sum_{m=1}^{M} k_m\|(\widetilde{Q}g_k)_m\|_{L^2(\Omega)} \leq C\sum_{m=1}^{M} k_m \sum_{l=1}^{m-1}\left(\frac{k_l}{\sqrt{\mu}}\frac{\omega(t_m-t_{l-1})}{(t_m-t_{l-1})^{\frac{3}{2}}} + \frac{\sqrt{\mu}k_l|k_m-k_l|}{(t_m-t_{l-1})^{\frac{3}{2}}}\right)\|g_{k,l}\|_{L^2(\Omega)}.$$

Changing the order of summation we have

$$\sum_{m=1}^{M} k_m\|(\widetilde{Q}g_k)_m\|_{L^2(\Omega)} \leq \sum_{l=1}^{M-1} k_l\|g_{k,l}\|_{L^2(\Omega)}\sum_{m=l+1}^{M}\left(\frac{k_m}{\sqrt{\mu}}\frac{\omega(t_m-t_{l-1})}{(t_m-t_{l-1})^{\frac{3}{2}}} + \frac{\sqrt{\mu}k_m|k_m-k_l|}{(t_m-t_{l-1})^{\frac{3}{2}}}\right).$$

Again similar to the above, by (1.5) we have

$$\sum_{m=l+1}^{M} k_m\frac{\omega(t_m-t_{l-1})}{(t_m-t_{l-1})^{\frac{3}{2}}} \leq \int_{t_{l-1}}^{T}\frac{w(s-t_{l-1})}{(s-t_{l-1})^{\frac{3}{2}}}\,ds \leq C.$$

Taking maximum over l in the sum $\sum_{m=l+1}^{M}\frac{k_m|k_m-k_l|}{(t_m-t_{l-1})^{\frac{3}{2}}}$ completes the proof. ∎

PROPOSITION 3.8. *There exists $\mu > 0$ sufficiently large and $\delta_0 > 0$ such that for $k - k_{\min} \leq \delta_0$ the following estimates hold for all $g_k \in X_k^0$:*

$$\|(\widetilde{Q}g_k)_m\|_{L^2(\Omega)} \leq \frac{3}{4}\max_{1\leq l\leq m}\|g_{k,l}\|_{L^2(\Omega)}, \quad m = 1, 2, \dots, M, \tag{3.14}$$

and

$$\sum_{m=1}^{M} k_m \|(\widetilde{Q} g_k)_m\|_{L^2(\Omega)} \leq \frac{3}{4} \sum_{l=1}^{M} k_l \|g_{k,l}\|_{L^2(\Omega)}. \tag{3.15}$$

Proof. Using the first estimate from Lemma 3.7 and choosing $\mu = 4C_1^2$ we obtain

$$\|(\widetilde{Q} g_k)_m\|_{L^2(\Omega)} \leq \max_{1 \leq j \leq m} \|g_{k,j}\|_{L^2(\Omega)} \left(\frac{1}{2} + C_2 \sqrt{\mu} \sum_{l=1}^{m-1} \frac{k_l |k_m - k_l|}{(t_m - t_{l-1})^{\frac{3}{2}}} \right).$$

Using $|k_m - k_l| \leq t_m - t_{l-1}$ we get for some $0 < \varepsilon < 1$

$$\begin{aligned} C_2 \sqrt{\mu} \sum_{l=1}^{m-1} \frac{k_l |k_m - k_l|}{(t_m - t_{l-1})^{\frac{3}{2}}} &\leq 2C_1 C_2 \sum_{l=1}^{m-1} \frac{k_l |k_m - k_l|^{\frac{1}{2} - \varepsilon}}{(t_m - t_{l-1})^{1-\varepsilon}} \\ &\leq 2C_1 C_2 (k - k_{\min})^{\frac{1}{2} - \varepsilon} \sum_{l=1}^{m-1} \frac{k_l}{(t_m - t_{l-1})^{1-\varepsilon}}. \end{aligned}$$

Using the properties of the Riemann sums we can estimate

$$\sum_{l=1}^{m-1} \frac{k_l}{(t_m - t_{l-1})^{1-\varepsilon}} \leq \int_0^{t_{m-1}} \frac{ds}{t_m - s} \leq C_\varepsilon.$$

Choosing, for example, $\varepsilon = \frac{1}{4}$ we get with $C_3 = 2C_1 C_2 C_\varepsilon$

$$\|(\widetilde{Q} g_k)_m\|_{L^2(\Omega)} \leq \max_{1 \leq j \leq m} \|g_{k,j}\|_{L^2(\Omega)} \left(\frac{1}{2} + C_3 (k - k_{\min})^{\frac{1}{4}} \right).$$

The estimate (3.14) follows then with the choice $\delta_0 = \frac{1}{(4C_3)^4}$. The estimate (3.15) follows from Lemma 3.7 similarly. ∎

Remark 3.9. The condition $k - k_{\min} \leq \delta_0$ trivially holds in the case of uniform time steps. For nonuniform time steps it is sufficient to assume $k \leq \frac{1}{2}\delta_0$.

The above proposition shows that under certain conditions, the operator $\mathrm{Id} - \widetilde{Q}$ is invertible with a bounded inverse with respect to both the $L^\infty(I; L^2(\Omega))$ and $L^1(I; L^2(\Omega))$ norms on X_k^0. This is the central piece of our argument. In order to obtain a discrete maximal parabolic regularity, we will also require estimates for $\widetilde{L}$ and $\tilde{D}$, which we will show next.

LEMMA 3.10. *For the operator $\widetilde{L}$ defined in* (3.5) *there exists a constant C independent of k such that for all $f_k \in X_k^0$ the following estimates hold:*

$$\|(\widetilde{L} f_k)_m\|_{L^2(\Omega)} \leq C \ln \frac{T}{k} \max_{1 \leq l \leq m} \|f_{k,l}\|_{L^2(\Omega)} \tag{3.16}$$

and

$$\sum_{m=1}^{M} k_m \|(\widetilde{L} f_k)_m\|_{L^2(\Omega)} \leq C \ln \frac{T}{k} \sum_{l=1}^{M} k_l \|f_{k,l}\|_{L^2(\Omega)}. \tag{3.17}$$

Proof. From the definition of $\widetilde{L}$ and Lemma 3.4 we obtain

$$\begin{aligned}\|(\widetilde{L}f_k)_m\|_{L^2(\Omega)} &\le \sum_{l=1}^{m} k_l \|\widetilde{A}_{k,m}\widetilde{R}_{m,l}\|_{L^2\to L^2}\|f_{k,l}\|_{L^2(\Omega)}\\ &\le C \max_{1\le l\le m}\|f_{k,l}\|_{L^2(\Omega)} \sum_{l=1}^{m}\frac{k_l}{t_m-t_{l-1}} \le C\ln\frac{T}{k}\max_{1\le l\le m}\|f_{k,l}\|_{L^2(\Omega)},\end{aligned}$$

where in the last step we used that

$$\sum_{l=1}^{m}\frac{k_l}{t_m-t_{l-1}} \le 1+\int_0^{t_{m-1}}\frac{dt}{t_m-t} = 1+\ln\frac{t_m}{k_m} \le C\ln\frac{T}{k}.$$

This completes the proof of (3.16). The estimate (3.17) can be shown similarly by changing the order of summation. ∎

LEMMA 3.11. *For the operator $\widetilde{D}$ defined in* (3.5) *there exists a constant C independent of k such that for all $u_0 \in L^2(\Omega)$ the following estimate holds:*

$$\sum_{m=1}^{M} k_m\|(\widetilde{D}u_0)_m\|_{L^2(\Omega)} \le C\ln\frac{T}{k}\|u_0\|_{L^2(\Omega)}. \tag{3.18}$$

If in addition $u_0 \in H_0^1(\Omega)$ with $A_{k,m}u_0 \in L^2(\Omega)$ for all $m = 1, 2, \ldots, M$, then

$$\max_{1\le m\le M}\|(\widetilde{D}u_0)_m\|_{L^2(\Omega)} \le C\mu\max_{1\le m\le M}\|A_{k,m}u_0\|_{L^2(\Omega)}. \tag{3.19}$$

Proof. There holds by Lemma 3.4

$$\|(\widetilde{D}u_0)_m\|_{L^2(\Omega)} \le \frac{C}{t_m}\|u_0\|_{L^2(\Omega)}.$$

This results in

$$\sum_{m=1}^{M} k_m\|(\widetilde{D}u_0)_m\|_{L^2(\Omega)} \le C\|u_0\|_{L^2(\Omega)}\sum_{m=1}^{M}\frac{k_m}{t_m} \le C\ln\frac{T}{k}\|u_0\|_{L^2(\Omega)}.$$

To prove (3.19) we use the fact that

$$\|(\widetilde{D}u_0)_m\|_{L^2(\Omega)} \le C\|\widetilde{A}_{k,m}u_0\|_{L^2(\Omega)} \le C\mu\|A_{k,m}u_0\|_{L^2(\Omega)}.$$ ∎

3.2. Semidiscrete in time maximal parabolic regularity. Combining the above results we can finally establish the maximal parabolic regularity with respect to the $L^\infty(I; L^2(\Omega))$ and the $L^1(I; L^2(\Omega))$ norms in the following two theorems.

THEOREM 3.12 (semidiscrete maximal parabolic regularity for $p = \infty$). *Let $f \in L^\infty(I; L^2(\Omega))$, and let $u_0 \in H_0^1(\Omega)$ with $A_{k,m}u_0 \in L^2(\Omega)$ for all $m = 1, 2, \ldots, M$. Let moreover u_k be the dG(0) semidiscrete solution to* (2.7). *There exists $\mu > 0$ sufficiently large and $\delta_0 > 0$ such that for $k - k_{\min} \le \delta_0$*

$$\begin{aligned}&\max_{1\le m\le M}\left(\|A_{k,m}u_{k,m}\|_{L^2(\Omega)} + \left\|\frac{[u_k]_{m-1}}{k_m}\right\|_{L^2(\Omega)}\right)\\ &\qquad\le Ce^{\mu T}\left(\ln\frac{T}{k}\|f\|_{L^\infty(I;L^2(\Omega))} + \mu T\max_{1\le m\le M}\|A_{k,m}u_0\|_{L^2(\Omega)}\right).\end{aligned}$$

Proof. Recalling the definitions of $\tilde{v}_k$ and w_k, namely,

$$\tilde{v}_{k,m} = \widetilde{A}_{k,m} w_{k,m} \quad \text{and} \quad u_{k,m} = \prod_{l=1}^{m}(1+\mu k_l) w_{k,m},$$

we have

$$\widetilde{A}_{k,m} u_{k,m} = \prod_{l=1}^{m}(1+\mu k_l)\widetilde{A}_{k,m} w_{k,m} = \prod_{l=1}^{m}(1+\mu k_l)\tilde{v}_{k,m}. \tag{3.20}$$

By (3.2) we have

$$\max_{1\le m\le M}\|\tilde{v}_{k,m}\|_{L^2(\Omega)} \le \max_{1\le m\le M}\left(\|(\widetilde{Q}\tilde{v}_k)_m\|_{L^2(\Omega)} + \|(\widetilde{L}\tilde{f}_k)_m\|_{L^2(\Omega)} + \|(\widetilde{D}u_0)_m\|_{L^2(\Omega)}\right).$$

For the first term we use estimate (3.14) from Proposition 3.8, for the second one estimate (3.16) from Lemma 3.10, and for the third one estimate (3.19) from Lemma 3.11. This results in

$$\begin{aligned}\max_{1\le m\le M}\|\tilde{v}_{k,m}\|_{L^2(\Omega)} \le \frac{3}{4}\max_{1\le m\le M}\|\tilde{v}_{k,m}\|_{L^2(\Omega)} + C\ln\frac{T}{k}\max_{1\le m\le M}\|\tilde{f}_{k,m}\|_{L^2(\Omega)}& \\ + C\mu T\max_{1\le m\le M}\|A_{k,m}u_0\|_{L^2(\Omega)}&.\end{aligned}$$

Absorbing the first term on the right-hand side we obtain

$$\max_{1\le m\le M}\|\tilde{v}_{k,m}\|_{L^2(\Omega)} \le C\ln\frac{T}{k}\|\widetilde{f}_k\|_{L^\infty(I;L^2(\Omega))} + C\mu T\max_{1\le m\le M}\|A_{k,m}u_0\|_{L^2(\Omega)}.$$

Now using that

$$\|\widetilde{f}_k\|_{L^\infty(I;L^2(\Omega))} \le C\|f\|_{L^\infty(I;L^2(\Omega))},$$

we obtain

$$\max_{1\le m\le M}\|\widetilde{A}_{k,m}u_{k,m}\|_{L^2(\Omega)} \le Ce^{\mu T}\left(\ln\frac{T}{k}\|f\|_{L^\infty(I;L^2(\Omega))} + \mu T\max_{1\le m\le M}\|A_{k,m}u_0\|_{L^2(\Omega)}\right),$$

where we used that $\prod_{l=1}^{m}(1+\mu k_l) \le e^{\mu t_m}$. Since $\tilde{A}_{k,m}$ is invertible for each m, using Lemma 3.2 we also have

$$\|u_{k,m}\|_{L^2(\Omega)} = \|\widetilde{A}_{k,m}^{-1}\widetilde{A}_{k,m}u_{k,m}\|_{L^2(\Omega)} \le \frac{1}{\mu}\|\widetilde{A}_{k,m}u_{k,m}\|_{L^2(\Omega)}.$$

Thus from (3.20) and the definition of $\widetilde{A}_{k,m}$, namely,

$$\widetilde{A}_{k,m} = (1+k_m\mu)A_{k,m} + \mu\,\mathrm{Id},$$

and by the triangle inequality and the estimates above we obtain

$$\|A_{k,m}u_{k,m}\|_{L^2(\Omega)} \le \mu\|u_{k,m}\|_{L^2(\Omega)} + \|\widetilde{A}_{k,m}u_{k,m}\|_{L^2(\Omega)} \le 2\|\widetilde{A}_{k,m}u_{k,m}\|_{L^2(\Omega)}$$

and therefore

$$\max_{1\le m\le M}\|A_{k,m}u_{k,m}\|_{L^2(\Omega)} \le Ce^{\mu T}\left(\ln\frac{T}{k}\|f\|_{L^\infty(I;L^2(\Omega))} + \mu T\max_{1\le m\le M}\|A_{k,m}u_0\|_{L^2(\Omega)}\right).$$

The estimate for the second term in the statement of the theorem follows from the observation that (2.11) is just

$$\frac{[u_k]_{m-1}}{k_m} = -A_{k,m}u_{k,m} + f_{k,m}.$$

∎

Next we establish the maximal parabolic regularity in $L^1(I;L^2(\Omega))$ norm.

THEOREM 3.13 (semidiscrete maximal parabolic regularity for $p=1$). *Let $f \in L^1(I;L^2(\Omega))$ and $u_0 \in L^2(\Omega)$. Let moreover u_k be the dG(0) semidiscrete solution to* (2.7)*. There exists $\mu > 0$ sufficiently large and $\delta_0 > 0$ such that for $k - k_{\min} \le \delta_0$ there holds*

$$\sum_{m=1}^{M} \left(k_m \|A_{k,m}u_{k,m}\|_{L^2(\Omega)} + \|[u_k]_{m-1}\|_{L^2(\Omega)}\right) \le Ce^{\mu T} \ln \frac{T}{k} \left(\|f\|_{L^1(I;L^2(\Omega))} + \|u_0\|_{L^2(\Omega)}\right).$$

Proof. The proof is very similar to the proof of the previous theorem. By (3.2) we have

$$\sum_{m=1}^{M} k_m \|\tilde{v}_{k,m}\|_{L^2(\Omega)} \le \sum_{m=1}^{M} k_m \left(\|(\widetilde{Q}\tilde{v}_k)_m\|_{L^2(\Omega)} + \|(\widetilde{L}\tilde{f}_k)_m\|_{L^2(\Omega)} + \|(\widetilde{D}u_0)_m\|_{L^2(\Omega)}\right).$$

For the first term we use estimate (3.15) from Proposition 3.8, for the second one estimate (3.17) from Lemma 3.10, and for the third one estimate (3.18) from Lemma 3.11. This results in

$$\sum_{m=1}^{M} k_m \|\tilde{v}_{k,m}\|_{L^2(\Omega)} \le \frac{3}{4} \sum_{m=1}^{M} k_m \|\tilde{v}_{k,m}\|_{L^2(\Omega)} + C \ln \frac{T}{k} \sum_{m=1}^{M} k_m \|\tilde{f}_{k,m}\|_{L^2(\Omega)} + C \ln \frac{T}{k} \|u_0\|_{L^2(\Omega)}.$$

The rest of the proof goes along the lines of the proof of the previous theorem. ∎

COROLLARY 3.14. *For $u_0 = 0$ interpolating between the results of Theorems* 3.12 *and* 3.13 *we obtain the discrete maximal parabolic regularity for $1 \le p < \infty$, namely,*

$$\left[\sum_{m=1}^{M} \left(\|A_{k,m}u_{k,m}\|^p_{L^p(I_m;L^2(\Omega))} + k_m \left\|\frac{[u_k]_{m-1}}{k_m}\right\|^p_{L^2(\Omega)}\right)\right]^{\frac{1}{p}} \le Ce^{\mu T} \ln \frac{T}{k} \|f\|_{L^p(I;L^2(\Omega))}.$$

3.3. Fully discrete maximal parabolic regularity. In this section, we consider the fully discrete approximation (1.1). We will establish the corresponding results for fully discrete approximations.

Let Ω be a polygonal/polyhedral domain and let $\mathcal{T}$ denote an admissible triangulation of Ω, i.e., $\mathcal{T} = \{\tau\}$ is a conformal partition of Ω into simplices (line segments, triangles, tetrahedrons, etc.) τ of diameter h_τ. Let $h = \max_\tau h_\tau$ and V_h be the set of all functions in $H^1_0(\Omega)$ that are polynomials of degree $r \ge 1$ on each τ, i.e., V_h is the usual space of continuous finite elements. We would like to point out that we do not make any assumptions on shape regularity or quasi-uniformity of the meshes. To obtain the fully discrete approximation we consider the space-time finite element space

$$X^{0,r}_{k,h} = \{v_{kh} :\ v_{kh}|_{I_m} \in \mathcal{P}_0(I_m;V_h),\ m = 1,2,\dots,M\}. \tag{3.21}$$

We define a fully discrete analogue $u_{kh} \in X^{0,r}_{k,h}$ of u_k introduced in (2.7) by

$$B(u_{kh},\varphi_{kh}) = (f,\varphi_{kh})_{I\times\Omega} + (u_0,\varphi^+_{kh})_\Omega \quad \forall \varphi_{kh} \in X^{0,r}_{k,h}. \tag{3.22}$$

Moreover, we introduce two operators $A_h(t)\colon V_h \to V_h$ defined by

$$(A_h(t)v_h,\chi)_\Omega = \sum_{i,j=1}^{d} (a_{ij}(t,\cdot)\partial_i v_h, \partial_j \chi)_\Omega \quad \forall \chi \in V_h \tag{3.23}$$

and the orthogonal L^2 projection $P_h\colon V_h \to V_h$ defined by

$$(P_h v_h, \chi)_\Omega = (v_h,\chi)_\Omega \quad \forall \chi \in V_h.$$

Similarly to $A_{k,m}$ in (2.10) we also define $A_{kh,m}\colon X^{0,r}_{k,h} \to X^{0,r}_{k,h}$

$$A_{kh,m} = \frac{1}{k_m}\int_{I_m} A_h(t)dt, \quad m = 1,2,\dots,M. \tag{3.24}$$

With the help of the above operators, the fully discrete approximation $u_{kh} \in X^{0,r}_{k,h}$ defined in (3.22) satisfies

$$\begin{aligned} u_{kh,1} + k_1 A_{kh,1}u_{kh,1} &= P_h u_0 + k_1 P_h f_{k,1},\\ u_{kh,m} + k_m A_{kh,m}u_{kh,m} &= u_{kh,m-1} + k_m P_h f_{k,m}, \quad m = 2,3,\dots,M, \end{aligned} \tag{3.25}$$

where $f_{k,m}$ is defined in (2.9). Hence the same formulas for u_k, namely, (2.12)–(2.17), also hold for u_{kh} with the difference that A_k is replaced by A_{kh} and f_k by $P_h f_k$. The analysis from section 3 of the paper translates almost immediately to the fully discrete setting since all arguments are energy based arguments and $\|P_h\|_{L^2(\Omega)\to L^2(\Omega)} \le 1$ on any mesh. Thus we directly establish the following results.

THEOREM 3.15 (fully discrete maximal parabolic regularity for $p=\infty$). *Let the conditions of Theorem 3.12 be fulfilled and let u_{kh} be fully discrete solution to (1.1) defined by (3.22) on any conformal triangulation of Ω. Then there exists a constant C independent of k and h such that the following estimate holds:*

$$\begin{aligned} &\max_{1\le m\le M}\left(\|A_{kh,m}u_{kh,m}\|_{L^2(\Omega)} + \left\|\frac{[u_{kh}]_{m-1}}{k_m}\right\|_{L^2(\Omega)}\right)\\ &\qquad\le Ce^{\mu T}\left(\ln\frac{T}{k}\|f\|_{L^\infty(I;L^2(\Omega))} + \mu T \max_{1\le m\le M}\|A_{kh,m}P_h u_0\|_{L^2(\Omega)}\right). \end{aligned}$$

The corresponding result for the $L^1(I;L^2(\Omega))$ norm is formulated in the following theorem.

THEOREM 3.16 (fully discrete maximal parabolic regularity for $p=1$). *Under the conditions of Theorem 3.15 there exists a constant C independent of k and h such that*

$$\begin{aligned} &\sum_{m=1}^{M}\left(k_m\|A_{kh,m}u_{kh,m}\|_{L^2(\Omega)} + \|[u_{kh}]_{m-1}\|_{L^2(\Omega)}\right)\\ &\qquad\le Ce^{\mu T}\ln\frac{T}{k}\left(\|f\|_{L^1(I;L^2(\Omega))} + \|u_0\|_{L^2(\Omega)}\right). \end{aligned}$$

COROLLARY 3.17. *For $u_0 = 0$ interpolating between Theorems 3.15 and 3.16 we obtain discrete maximal parabolic regularity for $1 \le p < \infty$*

$$\begin{aligned} &\left[\sum_{m=1}^{M}\left(\|A_{kh,m}u_{kh,m}\|^p_{L^p(I_m;L^2(\Omega))} + k_m\left\|\frac{[u_{kh}]_{m-1}}{k_m}\right\|^p_{L^2(\Omega)}\right)\right]^{\frac{1}{p}}\\ &\qquad\le Ce^{\mu T}\ln\frac{T}{k}\|f\|_{L^p(I;L^2(\Omega))}. \end{aligned}$$

4. Applications to error estimates. In this section we illustrate how the discrete maximal parabolic results from the previous section can be applied to error estimates. For the rest of the section we assume that Ω is convex and in addition the following assumption holds.

Assumption 1.
$$a_{ij}(t,\cdot) \in W^{1,\infty}(\Omega) \quad \forall t \in \bar{I},$$
and
$$L := \max_{1\le i,j\le d} \sup_{t\in\bar{I}} \|a_{ij}(t)\|_{W^{1,\infty}} < \infty.$$

4.1. Time semidiscrete error estimates. Using the convexity of Ω and Assumption 1 we establish the following preliminary result.

LEMMA 4.1. *There exists a constant C independent of k such that*
$$\sup_{t\in I_m} \|A(t)A_{k,m}^{-1}\|_{L^2\to L^2} \le C, \quad m = 1, 2, \dots, M.$$

Proof. Take an arbitrary $v \in L^2(\Omega)$ and set $w = A_{k,m}^{-1}v$, where $A_{k,m}$ is an elliptic operator with coefficients $a_{k,m}^{ij}$ defined by
$$a_{k,m}^{ij}(x) = \frac{1}{k_m}\int_{I_m} a_{ij}(t,x)dt. \tag{4.1}$$
By Assumption 1 we have $a_{k,m}^{ij} \in W^{1,\infty}(\Omega)$ with $\|a_{k,m}^{ij}\|_{W^{1,\infty}} \le L$ for all $1 \le i,j \le d$. From Theorems 2.2.2.3 and 3.2.1.2 in [13], we can conclude that $w \in H^2(\Omega)$ and
$$\|w\|_{H^2(\Omega)} \le C\|v\|_{L^2(\Omega)}, \tag{4.2}$$
where the constant C depends on Ω and L only. Again by Assumption 1, $A(t)$ is a bounded operator from $H^2(\Omega)\cap H_0^1(\Omega)$ to $L^2(\Omega)$ and as a result
$$\|A(t)A_{k,m}^{-1}v\|_{L^2(\Omega)} = \|A(t)w\|_{L^2(\Omega)} \le C\|w\|_{H^2(\Omega)} \le C\|v\|_{L^2(\Omega)}.$$
Taking supremum over v concludes the proof. □

As a first application of the discrete maximum regularity we establish semidiscrete best approximation result in the case of $p = \infty$.

THEOREM 4.2. *Let the coefficients $a_{ij}(t,x)$ satisfy Assumption* 1 *and let u be the solution to* (1.1) *with $u \in C(\bar{I}; L^2(\Omega))$ and u_k be the dG(0) semidiscrete solution to* (2.7)*. Then under the conditions of Theorem* 3.13 *there exists a constant C independent of k such that*
$$\|u-u_k\|_{L^\infty(I;L^2(\Omega))} \le Ce^{\mu T}\ln\frac{T}{k}\inf_{\chi\in X_k^0}\|u-\chi\|_{L^\infty(I;L^2(\Omega))}.$$

Proof. Let $\tilde{t} \in (0,T]$ be an arbitrary but fixed point in time. Without loss of generality we assume $\tilde{t} \in I_M = (t_{M-1},T]$. We consider the following dual problem:
$$\begin{aligned} \partial_t g(t,x) - A(t,x)g(t,x) &= \tilde{\theta}(t)u_k(\tilde{t},x), && (t,x)\in I\times\Omega,\\ g(t,x) &= 0, && (t,x)\in I\times\partial\Omega,\\ g(T,x) &= 0, && x\in\Omega, \end{aligned} \tag{4.3}$$

where $\tilde{\theta} \in C^\infty(0,T)$ is the regularized Delta function in time with the properties $\mathrm{supp}(\tilde{\theta}) \subset I_M$, $\|\tilde{\theta}\|_{L^1(I_M)} \le C$ and

$$(\tilde{\theta}, \varphi_k)_{I_M} = \varphi_k(\tilde{t}) \quad \forall \varphi_k \in X_k^0.$$

Let $g_k \in X_k^0$ be a dG(0) approximation of g, i.e., $B(g - g_k, \varphi_k) = 0$ for any $\varphi_k \in X_k^0$. Then

$$\begin{aligned}
\|u_k(\tilde{t})\|_{L^2(\Omega)}^2 &= (u_k(\cdot,\cdot), \tilde{\theta}(\cdot) u_k(\tilde{t},\cdot))_{I\times\Omega} \\
&= B(u_k, g) = B(u_k, g_k) = B(u, g_k) \\
&= \sum_{m=1}^M (u, A(t) g_{k,m})_{I_m\times\Omega} - \sum_{m=1}^M (u(t_m), [g_k]_m)_\Omega := J_1 + J_2.
\end{aligned}$$

Using the Hölder inequality in time, the Cauchy–Schwarz inequality in space, Theorem 3.13, and Lemma 4.1, we have

$$\begin{aligned}
J_1 &\le \|u\|_{L^\infty(I;L^2(\Omega))} \sum_{m=1}^M k_m \sup_{t\in I_m} \|A(t) A_{k,m}^{-1} A_{k,m} g_{k,m}\|_{L^2(\Omega)} \\
&\le C\|u\|_{L^\infty(I;L^2(\Omega))} \sum_{m=1}^M k_m \sup_{t\in I_m} \|A(t) A_{k,m}^{-1}\|_{L^2\to L^2} \|A_{k,m} g_{k,m}\|_{L^2(\Omega)} \\
&\le C e^{\mu T} \ln\frac{T}{k} \|u\|_{L^\infty(I;L^2(\Omega))} \|\tilde{\theta}\|_{L^1(I)} \|u_k(\tilde{t})\|_{L^2(\Omega)} \\
&\le C e^{\mu T} \ln\frac{T}{k} \|u\|_{L^\infty(I;L^2(\Omega))} \|u_k(\tilde{t})\|_{L^2(\Omega)}.
\end{aligned}$$

Similarly, we obtain

$$\begin{aligned}
J_2 &\le \sum_{m=1}^M \|u(t_m)\|_{L^2(\Omega)} \|[g_k]_m\|_{L^2(\Omega)} \\
&\le \|u\|_{L^\infty(I;L^2(\Omega))} \sum_{m=1}^M \|[g_k]_m\|_{L^2(\Omega)} \\
&\le C e^{\mu T} \ln\frac{T}{k} \|u\|_{L^\infty(I;L^2(\Omega))} \|u_k(\tilde{t})\|_{L^2(\Omega)}.
\end{aligned}$$

Canceling by $\|u_k(\tilde{t})\|_{L^2(\Omega)}$ and taking the supremum over $\tilde{t}$, we establish

$$\|u_k\|_{L^\infty(I;L^2(\Omega))} \le C e^{\mu T} \ln\frac{T}{k} \|u\|_{L^\infty(I;L^2(\Omega))}. \tag{4.4}$$

Using that the dG(0) method is invariant on X_k^0, by replacing u and u_k with $u-\chi$ and $u_k - \chi$ for any $\chi \in X_k^0$, and using the triangle inequality we obtain the theorem. ∎

For $1 \le p < \infty$, we can obtain the following result, which is similar to the result that was obtained for time independent coefficients in [33] for the $L^p(I; L^s(\Omega))$ norm. To state the result we define a projection π_k for $u \in C(I, L^2(\Omega))$ with $\pi_k u|_{I_m} \in P_0(I_m; L^2(\Omega))$ for $m = 1, 2, \dots, M$ on each subinterval I_m by

$$\pi_k u(t) = u(t_m), \quad t \in I_m. \tag{4.5}$$

THEOREM 4.3. *Let the coefficients* $a_{ij}(t,x)$ *satisfy Assumption* 1 *and let* u *be the solution to* (1.1) *with* $u \in C(\bar{I}; L^2(\Omega))$ *and* u_k *be the dG*(0) *semidiscrete solution to* (2.7)*. Then under the conditions of Theorem* 3.13 *there exists a constant* C *independent of* k *such that*

$$\|u - u_k\|_{L^p(I;L^2(\Omega))} \le Ce^{\mu T} \ln \frac{T}{k} \|u - \pi_k u\|_{L^p(I;L^2(\Omega))}, \quad 1 \le p < \infty,$$

where the projection π_k *is defined above in* (4.5)*.*

Proof. The proof uses the result of Corollary 3.14 and goes along the lines of the proof of Theorem 9 in [33] and Theorem 4.2 above. ∎

4.2. Applications to fully discrete error estimates. Similarly to the semidiscrete case, as an application of the fully discrete maximal parabolic regularity, we show optimal convergence rates for the dG(0)cG(r) solution. As in the semidiscrete case, first using the convexity of Ω and Assumption 1 and in addition that the triangulation $\mathcal{T}$ is quasi-uniform we establish the space discrete version of Lemma 4.1. Thus, for the rest of the paper we assume the following.

Assumption 2. There exists a constant C independent of h such that

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

where $d = 2, 3$ is the dimension on Ω. For the results below we will require one Ritz projection $R_h(t)\colon H_0^1(\Omega) \to V_h$, which is for every $t \in \bar{I}$ defined by

$$\sum_{i,j=1}^{d} (a_{ij}(t)\partial_i(R_h(t)v), \partial_j\chi)_\Omega = \sum_{i,j=1}^{d} (a_{ij}(t)\partial_i v, \partial_j\chi)_\Omega \quad \forall \chi \in V_h, \tag{4.6}$$

and another Ritz projection $R_{kh,m}\colon H_0^1(\Omega) \to V_h$, which is for every $m = 1, 2, \dots, M$ defined by

$$\sum_{i,j=1}^{d} \left(a_{k,m}^{ij}\partial_i R_{kh,m}v, \partial_j\chi\right)_\Omega = \sum_{i,j=1}^{d} \left(a_{k,m}^{ij}\partial_i v, \partial_j\chi\right)_\Omega \quad \forall \chi \in V_h, \tag{4.7}$$

where $a_{k,m}^{ij}$ are defined in (4.1).

LEMMA 4.4. *There exists a constant* C *independent of* k *and* h *such that*

$$\sup_{t \in I_m} \|A_h(t)A_{kh,m}^{-1}\|_{L^2 \to L^2} \le C, \quad m = 1, 2, \dots, M.$$

Proof. Take an arbitrary $v \in L^2(\Omega)$ and define $w_h = A_{kh,m}^{-1}P_h v$. In addition we also define $w = A_{k,m}^{-1}v$. Notice that $R_{kh,m}w = w_h$. By the definition of A_h in (3.23),

$$(A_h(t)w_h, \varphi_h)_\Omega = \sum_{i,j=1}^{d} (a_{ij}(t)\partial_i w_h, \partial_j\varphi_h)_\Omega \quad \forall \varphi_h \in V_h.$$

Put $z_h(t) = A_h(t)w_h$. Then adding and subtracting w, we have

$$\begin{aligned}\|z_h(t)\|^2_{L^2(\Omega)} = \|A_h(t)w_h\|^2_{L^2(\Omega)} &= \sum_{i,j=1}^{d} (a_{ij}(t)\partial_i w_h, \partial_j z_h(t))_\Omega \\ &= \sum_{i,j=1}^{d} (a_{ij}(t)\partial_i w, \partial_j z_h(t))_\Omega \\ &\quad + \sum_{i,j=1}^{d} (a_{ij}(t)\partial_i (w_h - w), \partial_j z_h(t))_\Omega := J_1 + J_2.\end{aligned}$$

To estimate J_1 we integrate by parts and use the Cauchy–Schwarz inequality

$$J_1 = -\sum_{i,j=1}^{d} (\partial_j(a_{ij}(t)\partial_i w), z_h(t))_\Omega \le L\|w\|_{H^2(\Omega)}\|z_h(t)\|_{L^2(\Omega)}.$$

Using that $w_h = R_{kh,m}w$, the standard elliptic error estimates, and the inverse inequality, we obtain

$$\begin{aligned}J_2 \le \sup_{i,j} \|a_{ij}(t)\|_{L^\infty(\Omega)}\|w - w_h\|_{H^1(\Omega)}\|z_h(t)\|_{H^1(\Omega)} &\le Ch\|w\|_{H^2(\Omega)}\|z_h(t)\|_{H^1(\Omega)} \\ &\le C\|w\|_{H^2(\Omega)}\|z_h(t)\|_{L^2(\Omega)}.\end{aligned}$$

Combining the estimates for J_1 and J_2 we can conclude that

$$\|z_h(t)\|_{L^2(\Omega)} \le C\|w\|_{H^2(\Omega)} \le C\|v\|_{L^2(\Omega)}, \tag{4.8}$$

where we used that the definition of w is identical to the definition of w in Lemma 4.1 and from (4.2) we know that $\|w\|_{H^2(\Omega)} \le C\|v\|_{L^2(\Omega)}$. Taking supremum over v concludes the proof. ∎

Similar to the semidiscrete case, we also establish a corresponding result for $p = \infty$ in the fully discrete case.

THEOREM 4.5. *Let the coefficients $a_{ij}(t,x)$ satisfy Assumption* 1 *and let u be the solution to* (1.1) *with $u \in C(\bar{I}; L^2(\Omega))$ and u_{kh} be the dG(0)cG(r) solution for $r \ge 1$ on a quasi-uniform triangulation $\mathcal{T}$ with the coefficients $a_{ij}(t,x)$ satisfying Assumption* 1. *Then under the assumption of Theorem* 3.15 *there exists a constant C independent of k and h such that for $1 \le p < \infty$,*

$$\begin{aligned}&\|u - u_{kh}\|_{L^\infty(I;L^2(\Omega))} \\ &\quad \le Ce^{\mu T} \ln\frac{T}{k}\left(\min_{\chi \in X^{0,r}_{k,h}} \|u - \chi\|_{L^\infty(I;L^2(\Omega))} + \|u - R_h u\|_{L^\infty(I;L^2(\Omega))}\right).\end{aligned}$$

Proof. As in the proof of Theorem 4.2, let $\tilde{t} \in (0,T]$ be an arbitrary but fixed point in time. Without loss of generality we assume $\tilde{t} \in I_M = (t_{M-1}, T]$. Consider the following dual problem:

$$\begin{aligned}\partial_t g(t,x) - A(t,x)g(t,x) &= \tilde{\theta}(t)u_{kh}(\tilde{t},x), && (t,x) \in I \times \Omega, \\ g(t,x) &= 0, && (t,x) \in I \times \partial\Omega, \\ g(T,x) &= 0, && x \in \Omega,\end{aligned} \tag{4.9}$$

where $\tilde{\theta} \in C^\infty(0,T)$ is the regularized Delta function in time with properties $\operatorname{supp}(\tilde{\theta}) \subset I_M$, $\|\tilde{\theta}\|_{L^1(I_M)} \le C$ and

$$(\tilde{\theta}, \varphi_{kh})_{I_M} = \varphi_{kh}(\tilde{t}) \quad \forall \varphi_{kh} \in X^{0,r}_{k,h}.$$

Let g_{kh} be dG(0)cG(r) approximation of g, i.e., $B(g - g_{kh}, \varphi_{kh}) = 0$ for any $\varphi_{kh} \in X^{0,r}_{k,h}$. Then

$$\begin{aligned}
\|u_{kh}(\tilde{t})\|^2_{L^2(\Omega)} &= (u_{kh}, \tilde{\theta} u_{kh}(\tilde{t}))_{I\times\Omega} \\
&= B(u_{kh}, g) = B(u_{kh}, g_{kh}) = B(u, g_{kh}) \\
&= \sum_{m=1}^{M} (R_h(t)u, A_h(t) g_{kh,m})_{I_m\times\Omega} - \sum_{m=1}^{M} (u(t_m), [g_{kh}]_m)_\Omega = J_1 + J_2.
\end{aligned}$$

Using the Hölder inequality in time, the Cauchy–Schwarz inequality in space, Lemma 4.4, and Theorem 3.16, we have

$$\begin{aligned}
J_1 &\le \|R_h u\|_{L^\infty(I;L^2(\Omega))} \sum_{m=1}^{M} k_m \sup_{t\in I_m} \|A_h(t) A^{-1}_{kh,m} A_{kh,m} g_{k,m}\|_{L^2(\Omega)} \\
&\le \|R_h u\|_{L^\infty(I;L^2(\Omega))} \sum_{m=1}^{M} k_m \sup_{t\in I_m} \|A_h(t) A^{-1}_{kh,m}\|_{L^2\to L^2} \|A_{kh,m} g_{k,m}\|_{L^2(\Omega)} \\
&\le C e^{\mu T} \ln\frac{T}{k} \|R_h u\|_{L^\infty(I;L^2(\Omega))} \|\tilde{\theta}\|_{L^1(I)} \|u_{kh}(\tilde{t},\cdot)\|_{L^2(\Omega)} \\
&\le C e^{\mu T} \ln\frac{T}{k} \|R_h u\|_{L^\infty(I;L^2(\Omega))} \|u_{kh}(\tilde{t})\|_{L^2(\Omega)}.
\end{aligned}$$

Exactly as in the estimate of J_2 in Theorem 4.5, we obtain

$$J_2 \le C e^{\mu T} \ln\frac{T}{k} \|u\|_{L^\infty(I;L^2(\Omega))} \|u_{kh}(\tilde{t})\|_{L^2(\Omega)}.$$

Thus canceling $\|u_{kh}(\tilde{t})\|_{L^2(\Omega)}$ and taking supremum over $\tilde{t}$, we establish

$$\|u_{kh}\|_{L^\infty(I;L^2(\Omega))} \le C e^{\mu T} \ln\frac{T}{k} \left(\|R_h u\|_{L^\infty(I;L^2(\Omega))} + \|u\|_{L^\infty(I;L^2(\Omega))} \right). \tag{4.10}$$

Using that the dG(0)cG(r) method is invariant on $X^{0,r}_{k,h}$, by replacing u and u_{kh} with $u - \chi$ and $u_{kh} - \chi$ for any $\chi \in X^{0,r}_{k,h}$, and using the triangle inequality we obtain the theorem. ∎

THEOREM 4.6. *Let the coefficients $a_{ij}(t,x)$ satisfy Assumption* 1 *and let u be the solution to* (1.1) *with $u \in C(\bar{I}; L^2(\Omega))$ and u_{kh} be the dG(0)cG(r) solution for $r \ge 1$ on a quasi-uniform triangulation $\mathcal{T}$ with the coefficients $a_{ij}(t,x)$ satisfying Assumption* 1. *Then under the assumption of Theorem* 3.15 *there exists a constant C independent of k and h such that for $1 \le p < \infty$,*

$$\|u - u_{kh}\|_{L^p(I;L^2(\Omega))} \le C \ln\frac{T}{k} \left(\|u - \pi_k u\|_{L^p(I;L^2(\Omega))} + \|u - R_h u\|_{L^p(I;L^2(\Omega))} \right),$$

where the projection π_k is defined in (4.5) *and R_h in* (4.6).

Proof. The proof uses the result of Corollary (3.17) and goes along the lines of the proof of Theorem 12 in [33] and Theorem 4.5 above. ∎

For sufficiently regular solutions, combining the above two theorems and using the approximation theory we immediately obtain an optimal order convergence result.

COROLLARY 4.7. *Under the assumptions of Theorem* 4.6 *and the regularity* $u \in W^{1,p}(I; L^2(\Omega) \cap L^p(I; H^{r+1}(\Omega))$ *for some* $1 \leq p \leq \infty$, *there exists a constant* C *independent of* k *and* h *such that*

$$\|u - u_{kh}\|_{L^p(I;L^2(\Omega))} \leq C \ln \frac{T}{k} \left(k\|u\|_{W^{1,p}(I;L^2(\Omega))} + h^{r+1}\|u\|_{L^p(I;H^{r+1}(\Omega))} \right).$$

Remark 4.8. The results of Theorems 4.2, 4.3, 4.5, and 4.6 also hold for the elliptic operator of the form $A(t,x) = b(t)A(x)$, where $b(t) \in C^{\frac{1}{2}+\varepsilon}(\bar{I})$ and $A(x)$ is the second order elliptic operator with bounded coefficients. In view of the uniform ellipticity condition (1.3), we have $b(t) \geq b_0 > 0$ for some $b_0 \in \mathbb{R}^+$, and as a result Lemmas 4.1 and 4.4 trivially hold without any additional assumptions, such as Assumptions 1 and 2. Thus, the results of the theorems hold for nonconvex polygonal/polyhedral domains and on graded meshes.

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

REFERENCES

[1] A. ASHYRALYEV, S. PISKAREV, AND L. WEIS, *On well-posedness of difference schemes for abstract parabolic equations in $L^p([0,T];E)$ spaces*, Numer. Funct. Anal. Optim., 23 (2002), pp. 669–693.

[2] A. ASHYRALYEV AND P. E. SOBOLEVSKIĬ, *Well-Posedness of Parabolic Difference Equations*, Oper. Theory Adv. Appl. 69, Birkhäuser Verlag, Basel, 1994.

[3] P. AUSCHER AND M. EGERT, *On non-autonomous maximal regularity for elliptic operators in divergence form*, Arch. Math. (Basel), 107 (2016), pp. 271–284.

[4] P. CANNARSA AND V. VESPRI, *On maximal L^p regularity for the abstract Cauchy problem*, Boll. Unione Mat. Ital., 5 (1986), pp. 165–175.

[5] K. CHRYSAFINOS AND N. J. WALKINGTON, *Error estimates for the discontinuous Galerkin methods for parabolic equations*, SIAM J. Numer. Anal., 44 (2006), pp. 349–366.

[6] T. COULHON AND X. T. DUONG, *Maximal regularity and kernel bounds: Observations on a theorem by Hieber and Prüss*, Adv. Differential Equations, 5 (2000), pp. 343–368.

[7] R. DENK, M. HIEBER, AND J. PRÜSS, *R-boundedness, Fourier multipliers and problems of elliptic and parabolic type*, Mem. Amer. Math. Soc., 166 (2003).

[8] D. DIER, *Non-autonomous maximal regularity for forms of bounded variation*, J. Math. Anal. Appl., 425 (2015), pp. 33–54.

[9] K. DISSER, A. F. M. TER ELST, AND J. REHBERG, *On maximal parabolic regularity for non-autonomous parabolic operators*, J. Differential Equations, 262 (2017), pp. 2039–2072.

[10] M. DOBROWOLSKI, *Optimale Konvergenz der Methode der finiten Elemente bei parabolischen Anfangs-, Randwertaufgaben*, Bonner Mathematische Schriften 112, Universität Bonn, Mathematisches Institut, Bonn, 1978.

[11] K. ERIKSSON, C. JOHNSON, AND S. LARSSON, *Adaptive finite element methods for parabolic problems.* VI. *Analytic semigroups*, SIAM J. Numer. Anal., 35 (1998), pp. 1315–1325.

[12] C. GONZÁLEZ AND C. PALENCIA, *Stability of time-stepping methods for abstract time-dependent parabolic problems*, SIAM J. Numer. Anal., 35 (1998), pp. 973–989.

[13] P. GRISVARD, *Elliptic Problems in Nonsmooth Domains*, Classics Appl. Math. 69, SIAM, Philadelphia, 2011.

[14] D. GUIDETTI, *Backward Euler scheme, singular Hölder norms, and maximal regularity for parabolic difference equations*, Numer. Funct. Anal. Optim., 28 (2007), pp. 307–337.

[15] D. GUIDETTI AND S. PISKAREV, *Stability of Rothe's scheme and maximal regularity for parabolic equations in $C^\theta(\overline{\Omega})$*, in Progress in Partial Differential Equations, Vol. 1, Pitman Res. Notes Math. Ser. 383, Longman, Harlow, 1998, pp. 167–180.

[16] D. GUIDETTI AND S. PISKAREV, *Stability of the Crank–Nicolson scheme and maximal regularity for parabolic equations in $C^\theta(\overline{\Omega})$ spaces*, Numer. Funct. Anal. Optim., 20 (1999), pp. 251–277.

[17] B. H. Haak and E. M. Ouhabaz, *Maximal regularity for non-autonomous evolution equations*, Math. Ann., 363 (2015), pp. 1117–1145.

[18] R. Haller-Dintelmann and J. Rehberg, *Maximal parabolic regularity for divergence operators including mixed boundary conditions*, J. Differential Equations, 247 (2009), pp. 1354–1396.

[19] M. Hieber and J. Prüss, *Heat kernels and maximal L^p-L^q estimates for parabolic evolution equations*, Comm. Partial Differential Equations, 22 (1997), pp. 1647–1669.

[20] D. Hipp, M. Hochbruck, and A. Ostermann, *An exponential integrator for non-autonomous parabolic problems*, Electron. Trans. Numer. Anal., 41 (2014), pp. 497–511.

[21] D. Hömberg, C. Meyer, J. Rehberg, and W. Ring, *Optimal control for the thermistor problem*, SIAM J. Control Optim., 48 (2009/10), pp. 3449–3481.

[22] M. Y. Huang and V. Thomée, *Some convergence estimates for semidiscrete type schemes for time-dependent nonselfadjoint parabolic equations*, Math. Comp., 37 (1981), pp. 327–346.

[23] O. A. Karakashian, *On Runge–Kutta methods for parabolic problems with time-dependent coefficients*, Math. Comp., 47 (1986), pp. 77–101.

[24] S. L. Keeling, *Galerkin/Runge–Kutta discretizations for parabolic equations with time-dependent coefficients*, Math. Comp., 52 (1989), pp. 561–586.

[25] B. Kovács, B. Li, and C. Lubich, *A-stable time discretizations preserve maximal parabolic regularity*, SIAM J. Numer. Anal., 54 (2016), pp. 3600–3624.

[26] 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.

[27] 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.

[28] P. C. Kunstmann, B. Li, and C. Lubich, *Runge–Kutta time discretization of nonlinear parabolic equations studied via discrete maximal parabolic regularity*, Found. Comput. Math., 2017.

[29] 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.

[30] D. Leykekhman and B. Vexler, *Global and interior pointwise best approximation results for the gradient of Galerkin solutions for parabolic problems*, SIAM J. Numer. Anal., 55 (2017), pp. 2025–2049.

[31] 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.

[32] D. Leykekhman and B. Vexler, *A priori error estimates for three dimensional parabolic optimal control problems with pointwise control*, SIAM J. Control Optim., 54 (2016), pp. 2403–2435.

[33] D. Leykekhman and B. Vexler, *Discrete maximal parabolic regularity for Galerkin finite element methods*, Numer. Math., 135 (2017), pp. 923–952.

[34] B. Li and W. Sun, *Maximal regularity of fully discrete finite element solutions of parabolic equations*, SIAM J. Numer. Anal., 55 (2017), pp. 521–542.

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

[36] M. Luskin and R. Rannacher, *On the smoothing property of the Crank–Nicolson scheme*, Appl. Anal., 14 (1982/83), pp. 117–135.

[37] 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.

[38] P. Sammon, *Fully discrete approximation methods for parabolic problems with nonsmooth initial data*, SIAM J. Numer. Anal., 20 (1983), pp. 437–470.

[39] P. H. Sammon, *Approximations for Parabolic Equations with Time-Dependent Coeficients*, Ph.D. Thesis, Cornell University, 1978.

[40] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer Ser. Comput. Math. 25, Springer-Verlag, Berlin, 2006.