ENERGY NORM ERROR ESTIMATES FOR FINITE ELEMENT DISCRETIZATION OF PARABOLIC PROBLEMS

HERBERT EGGER

Department of Mathematics, TU Darmstadt, Germany

Abstract. We consider the discretization of parabolic initial boundary value problems by finite element methods in space and a Runge-Kutta time stepping scheme. Order optimal a-priori error estimates are derived in an energy-norm under natural smoothness assumptions on the solution and without artificial regularity conditions on the parameters and the domain. The key steps in our analysis are the careful treatment of time derivatives in the $H^{-1}$-norm and the the use of an $L^2$-projection in the error splitting instead of the Ritz projector. This allows us to restore the optimality of the estimates with respect to smoothness assumptions on the solution and to avoid artificial regularity requirements for the problem, usually needed for the analysis of the Ritz projector, which limit the applicability of previous work. The wider applicability of our results is illustrated for two irregular problems, for which previous results can either not be applied or yield highly sub-optimal estimates.

Keywords: parabolic problems, Galerkin discretization, finite element method, backward Euler scheme, energy-norm estimates

AMS-classification: 35K20, 65L06, 65L70, 65M15, 65M20, 65M60

1. Introduction

Let $\Omega \subset \mathbb{R}^d$ be a bounded domain and $T > 0$ denote a time horizon. We consider the numerical solution of parabolic initial boundary value problems of the form

$$u'(t) + A(t)u(t) = f(t), \quad \text{in } \Omega, \quad 0 < t < T, \quad (1a)$$

$$u(t) = 0, \quad \text{on } \partial\Omega, \quad 0 < t < T, \quad (1b)$$

$$u(0) = u_0, \quad \text{in } \Omega. \quad (1c)$$

Here $u'$ is the time derivative and $Au = -\text{div}(a\nabla u) + b \cdot \nabla u + cu$ is a second order differential operator with coefficients $a, b, c \in L^\infty$ and $a(x, t) \geq a > 0$, such that $A(t)$ is uniformly elliptic for every point in time. It is well known, that for any $u_0 \in L^2(\Omega)$ and any $f \in L^2(0, T; H^{-1}(\Omega))$ there exists a unique weak solution in the energy-space

$$W(0, T) = \{u \in L^2(0, T; H^1_0(\Omega)) : u' \in L^2(0, T; H^{-1}(\Omega))\}, \quad (2)$$

which can be bounded a-priori by

$$\|u\|_{W(0,T)} = \|u\|_{L^2(0,T;H^1_0(\Omega))} + \|u'\|_{L^2(0,T;H^{-1}(\Omega))} \leq C \left( \|f\|_{L^2(0,T;H^{-1}(\Omega))} + \|u_0\|_{L^2(\Omega)} \right). \quad (3)$$

Moreover, the constant $C$ only depends on the domain and the bounds for the coefficients. We will refer to $\| \cdot \|_{W(0,T)}$ as the energy-norm of the problem. If the coefficients and the
data are sufficiently smooth, the solution of (1a)-(1d) can be expected to be more regular:
For instance, if one assumes that $a, b, c \in W^{1,\infty}$, then
\begin{equation}
\|u\|_{L^2(0,T;H^1_0(\Omega))} + \|u'\|_{L^2(0,T;H^{-1}(\Omega))} \leq C\left(\|f'\|_{L^2(0,T;H^{-1}(\Omega))} + \|f(0) - A(0)u_0\|_{L^2(\Omega)}\right),
\end{equation}
whenever the right hand side is bounded. If, in addition, also the domain is sufficiently regular and $u_0 \in H^1_0(\Omega)$, then
\begin{equation}
\|u\|_{L^2(0,T;H^2(\Omega))} + \|u'\|_{L^2(0,T;L^2(\Omega))} \leq C\left(\|f\|_{L^2(0,T;L^2(\Omega))} + \|u_0\|_{H^1_0(\Omega)}\right).
\end{equation}
We refer to [9] for details and proofs of these estimates and further results.

Let us emphasize at this point that the two estimates (5) and (6) also characterize the basic regularity spaces for parabolic problems. In particular, time derivatives typically lack two order of spatial regularity compared to the solution itself.

In this paper, we study the numerical solution of (1a)-(1d) by finite element discretization in space and Runge-Kutta time stepping schemes. Such approximations have been investigated intensively in the literature, see e.g. [6, 12, 14, 18, 23, 25]; we refer to [20] for a comprehensive collection of results and further references. Our main goal here is to derive order optimal a-priori error estimates in the energy-norm $\| \cdot \|_{W_{0,T}}$. These can be obtained under minimal regularity assumptions on the coefficients and the domain, and we only require natural smoothness conditions for the solution. To keep the notation simple, we consider here in detail only the lowest order approximation by piecewise linear finite elements in space and the backward Euler method in time. The generalization of our arguments to approximations of higher order is however straightforward.

To put our results into perspective, let us shortly recall some of the standard results for the Galerkin approximation of parabolic problems [20, 23, 25] and compare them with the main contributions of our manuscript:

As a first step, we will consider the semi-discretization in space by piecewise linear finite elements. The standard lowest order error estimate for parabolic problems then reads
\begin{equation}
\|u - u_h\|_{L^\infty(0,T;L^2(\Omega))} \leq Ch\left(\|u_0\|_{H^1(\Omega)} + \|u'\|_{L^1(0,T;H^1(\Omega))}\right).
\end{equation}
Here $u_h$ is the finite element approximation and $h$ the mesh size. This estimate amounts to [20] Theorem 1.2 with $r = 1$. Assuming sufficient regularity of the domain, the parameters, and the solution, also second order convergence can be obtained here. The basic step in the proof of this error estimate is the decomposition of the error into
\begin{equation}
R_h = (u - R_hu) + (R_hu - u_h),
\end{equation}
where $R_h$ denotes the Ritz projection associated to the operator $A$; see [24, 25]. The proof of the error bound then relies on a discrete energy estimate for the semi-discrete problem and certain properties of the Ritz projection, in particular, on a bound
\begin{equation}
\|u - R_hu\|_{L^2(\Omega)} \leq Ch\|u\|_{H^1(\Omega)},
\end{equation}
for the $L^2$-norm error, which can be obtained by the usual duality arguments [11, 17]. Without some additional regularity assumptions on the parameters and the domain, the estimate (9) for the Ritz projection is however not valid, and therefore the validity of the bound (7) for irregular problems cannot be granted by the proofs given in [20, 23, 25]. Also note that the estimate (7) is somewhat sub-optimal concerning the regularity requirements for the solution: In fact, it should suffice to assume that $u \in L^\infty(0,T;H^1(\Omega))$, which is already valid if $u \in L^2(0,T;H^2(\Omega)) \cap H^1(0,T;L^2(\Omega))$. Let us emphasize that no spatial
regularity for the time derivative seems to be required. Also, in view of the a-priori estimate \( (6) \), this latter condition would be a natural regularity assumption.

As a replacement for \( (7) \), we will derive an error estimate in the energy-norm \( ||·||_{W(0,T)} \), which arises in the a-priori estimate for the solution. We will show that

\[
\|u - u_h\|_{W(0,T)} = \|u - u_h\|_{L^2(0,T;H^1(\Omega))} + \|u' - u_h'\|_{L^2(0,T;H^{-1}(\Omega))}
\leq Ch \left( \|u\|_{L^2(0,T;H^2(\Omega))} + \|u'\|_{L^2(0,T;L^2(\Omega))} \right). \tag{10}
\]

By continuous embedding, one has \( \|u - u_h\|_{L^\infty(0,T;\Omega)} \leq C\|u - u_h\|_{W(0,T)} \), which yields the corresponding estimate also for the error in the \( L^2 \)-norm at every point in time. One can easily see that the estimate \( (10) \) is optimal with respect to the approximation properties of the finite element spaces and also with respect to the smoothness requirements on the solution, which are natural for the problem under investigation. Also note that we require no additional regularity of the domain or the parameters for the proof of this result.

As a second step, we will then investigate the time discretization by the backward Euler method. The standard error estimate for the fully discrete approximation reads

\[
\max_{0 \leq n \leq T} \|u(t^n) - u^n_h\|_{L^2(\Omega)} \leq Ch \left( \|u_0\|_{H^1(\Omega)} + \|u_t\|_{L^1(0,T;H^1(\Omega))} + \tau \|u_t\|_{L^1(0,T;L^2(\Omega))} \right), \tag{11}
\]

see [20] Theorem 1.5] with \( r = 1 \). Here \( \tau = t^n - t^{n-1} \) is the time-step size, and \( u^n_h \) denotes the \( n \)-th iterate of the Euler method. Again, we observe a certain sub-optimality concerning the regularity requirements for the solution. Moreover, the proof given in [20], see also [23], is only valid under additional restrictive regularity assumptions on the parameters and the domain, which strongly limit the applicability of the results.

As a replacement for \( (11) \), we will derive the energy-norm estimate

\[
\|u - \tilde{u}_h\|_{W(0,T)} = \|u - \tilde{u}_h\|_{L^2(0,T;H^1(\Omega))} + \|u' - \tilde{u}'_h\|_{L^2(0,T;H^{-1}(\Omega))}
\leq C \left( h\|u\|_{L^2(0,T;H^2(\Omega))} + h\|u'\|_{L^2(0,T;L^2(\Omega))} \right.
+ \tau \|u'\|_{L^2(0,T;H^1(\Omega))} + \tau \|u''\|_{L^2(0,T;H^{-1}(\Omega))} \bigg). \tag{12}
\]

Here \( \tilde{u}_h \) denotes the function obtained from \( u^n_h \) by piecewise linear interpolation in time. The bound for \( u^n_h \) again follows easily by embedding of \( W(0,T) \) into \( L^\infty(0,T;L^2(\Omega)) \). As before, the estimate \( (12) \) can be seen to be optimal with respect to the approximation properties of the discretization, and, in view of the a-priori bounds \( (5) \) and \( (6) \), the regularity requirements for the solution seem natural. Moreover, no artificial regularity of the domain and the coefficients will be required for the proof of this estimate.

One key step in the derivation of our results will be the careful estimation of time derivatives in the \( H^{-1} \)-norm. This seems natural in view of the definition of the energy-norm and its role in the a-priori estimates. The importance of such estimates has already been observed in the context of a-posteriori error estimation \[15\]. Let us also mention \[13, 26\], where \( H^{-1} \)-Galerkin methods for the solution of parabolic initial boundary value problems have been investigated.

A second difference to previous investigations is, that we use here a somewhat different error splitting as usually employed, namely

\[
u - u_h = (u - \pi_h u) + (\pi_h u - u_h), \tag{13}\]

where \( \pi_h \) is the \( L^2 \)-projection onto the finite element space. The \( L^2 \)-projection \( \pi_h \) has important advantages in comparison to the Ritz-projector \( R_h \): First, the estimate corresponding to \( (3) \) and further approximation properties can be proven without regularity assumptions on the domain or the parameters. In addition, \( \pi_h \) commutes with the time
derivative, even for problems with time dependent parameters. At several places in our analysis, we will rely on the $H^1$-stability of the $L^2$-projection \[3\]. Moreover, we use approximation error estimates for the $L^2$-projection in various norms, in particular including estimates in the $H^{-1}$-norm.

The remainder of the manuscript is organized as follows: In Section 2, we formally present the problem to be investigated and we discuss the basic assumptions on the domain and the coefficients. In Section 3, we introduce the finite element spaces and summarize some estimates for the $L^2$-projection that are required later in our analysis. Section 4 is then concerned with the derivation of the estimate (10) for the semi-discretization. In Section 5, we investigate the time discretization by the backward Euler method and we derive the second estimate (12). For illustration of the wider applicability of our results, we discuss in Section 6 two irregular test problems, for which, due to lack of regularity, the standard bounds (7) and (14) cannot be applied directly. Due to the weaker requirements for our estimates (10) and (12), the optimal convergence in the energy-norm can however still be guaranteed also theoretically. We close with a short discussion of our results and highlight some possibilities for generalization and future investigations.

2. Preliminaries

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be some bounded domain and $T > 0$. We use standard notation for function spaces, see e.g. [9]; in particular, $H^1_0(\Omega)$ denotes the sub-space of functions in $H^1(\Omega)$ with vanishing traces on $\partial\Omega$, and $H^{-1}(\Omega) = (H^1_0(\Omega))^\prime$ is the space of bounded linear functional on $H^1_0(\Omega)$; the corresponding duality product is denoted by $\langle \cdot, \cdot \rangle_{H^{-1}(\Omega) \times H^1_0(\Omega)}$. Of particular importance for our analysis is the energy-space

$$W(0,T) = \{ u \in L^2(0,T; H^1_0(\Omega)) : u' \in L^2(0,T; H^{-1}(\Omega)) \},$$

(14)

which is equipped with the norm $\|u\|_{W(0,T)} = \|u\|_{L^2(0,T; H^1(\Omega))} + \|u'\|_{L^2(0,T; H^{-1}(\Omega))}$. For later reference, let us recall the following embedding result [9 Sec. 5.9].

Lemma 1. Let $u \in W(0,T)$. Then $u \in C([0,T]; L^2(\Omega))$ and

$$\sup_{0 \leq t \leq T} \|u(t)\|_{L^2(\Omega)} \leq C \|u\|_{W(0,T)}.$$

As a consequence, all estimates derived in the energy-norm $\| \cdot \|_{W(0,T)}$ automatically yield corresponding bounds in $\| \cdot \|_{L^\infty(0,T; L^2(\Omega))}$, i.e., pointwise in time. Here and below, we denote by $C$ some generic constant which may have different values in different occasions.

Let us now turn to the initial-boundary value problem (1a)-(1c). We assume that the operator $A$ has the form $Au = -\text{div}(a \nabla u) + b \cdot \nabla u + cu$. In order to prove well-posedness of the initial boundary value problem, we assume that the parameters satisfy

(A1) $a, c \in L^\infty(\Omega \times (0,T))$ and $b \in L^\infty(\Omega \times (0,T))^d$,

(A2) $a(x,t) \geq \underline{a}$ for some constant $\underline{a} > 0$ and a.e. $(x,t) \in \Omega \times (0,T)$.

Because of the first assumption, $A$ defines a bounded linear operator from $L^2(0,T; H^1_0(\Omega))$ to $L^2(0,T; H^{-1}(\Omega))$. For a.e. $t \in (0,T)$ we then define an associated bilinear form by

$$a(u,v;t) = \int_\Omega a(x,t)\nabla u(x) \nabla v(x) + b(x,t) \cdot \nabla u(x)v(x) + c(x,t)u(x)v(x) \, dx.$$

(15)

The weak formulation of the initial boundary value problem (1a)-(1c) now reads
Problem 2 (Weak formulation).

Given \( f \in L^2(0,T;H^{-1}(\Omega)) \) and \( u_0 \in L^2(\Omega) \), find \( u \in W(0,T) \) such that \( u(0) = u_0 \) and

\[
\langle u'(t), v \rangle_{H^{-1}(\Omega) \times H_0^1(\Omega)} + a(u(t), v; t) = \langle f(t), v \rangle_{H^{-1}(\Omega) \times H_0^1(\Omega)}
\]

holds for all test functions \( v \in H_0^1(\Omega) \) and for a.e. \( t \in (0,T) \).

Under assumptions (A1)-(A2), Problem 2 has a unique solution \( u \in W(0,T) \) and there holds \( \|u\|_{W(0,T)} \leq C (\|u_0\|_{L^2(\Omega)} + \|f\|_{L^2(0,T;H^{-1}(\Omega))}) \) with a constant \( C \) that depends only on the bounds for the coefficients and on the domain; see [9] for details. The proof relies on a Gronwall argument and the following properties of the bilinear form.

Lemma 3 (Continuity and Garding inequality).

Let (A1) and (A2) hold. Then there exist constants \( C, \alpha, \eta > 0 \) such that

\[
a(u, v; t) \leq C \|u\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)} \quad \text{and} \quad a(u, u; t) + \eta \|u\|_{L^2(\Omega)}^2 \geq \alpha \|u\|_{H^1(\Omega)}^2,
\]

and these estimates hold uniformly for a.e. \( t \in (0,T) \) and all \( u, v \in H^1(\Omega) \).

Proof. The continuity follows from the Cauchy-Schwarz inequality. Using the lower and upper bounds for the coefficients, we get

\[
a(u, u; t) \geq \alpha \|\nabla u\|_{L^2(\Omega)}^2 + \|b(t)\|_{L^\infty} \|\nabla u\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} + \|c(t)\|_{L^\infty} \|u\|_{L^2(\Omega)}^2.
\]

The Garding inequality then follows by Young’s inequality and choosing the coefficients, for instance, as \( \alpha = \eta/2 \) and \( \eta = \alpha/2 + \|b\|_{L^\infty}/(2\alpha) + \|c\|_{L^\infty} \).

3. Properties of the \( L^2 \)-projection onto finite element spaces

For the semi-discretization in space, we will employ a standard finite element method. To avoid technical difficulties, we assume that \( \Omega \) is polyhedral and that it can be partitioned into a set \( T_h = \{T\} \) of simplicial elements \( T \). More precisely, we require that

- (A3) \( T_h \) is a regular simplicial partition of \( \Omega \), i.e., the intersection of two different elements either empty, or a vertex, an entire edge, (an entire face) of both elements;
- (A4) \( T_h \) is locally quasi-uniform, i.e., there exists a \( \gamma > 0 \) such that the diameter \( h_T \) of an element \( T \) and the radius \( r_T \) of the largest ball that can be inscribed in \( T \) are related by \( \gamma h_T \leq r_T \leq h_T \) for all elements \( T \).

Given such a mesh \( T_h \), we consider the standard finite element space

\[
V_h = \{v \in H_0^1(\Omega) : v|_T \in P_1(T) \text{ for all } T \in T_h\},
\]

of piecewise linear continuous functions that vanish on the boundary. Furthermore, we denote by \( \pi_h : L^2(\Omega) \rightarrow V_h \) the \( L^2 \)-orthogonal projection defined by

\[
(\pi_h u, v_h)_{L^2(\Omega)} = (u, v_h)_{L^2(\Omega)} \quad \text{for all } v_h \in V_h.
\]

Obviously, \( \|\pi_h u\|_{L^2(\Omega)} \leq \|u\|_{L^2(\Omega)} \), i.e., the \( L^2 \)-projection is stable (a bounded linear operator) on \( L^2(\Omega) \). Under the assumption (A3)-(A4), it is however also stable on \( H_0^1(\Omega) \).

Lemma 4 (\( H^1 \)-stability). Let (A3)-(A4) hold. Then

\[
\|\pi_h u\|_{H^1(\Omega)} \leq C \|u\|_{H^1(\Omega)} \quad \text{for all } u \in H_0^1(\Omega),
\]

and the constant \( C \) only depends on the domain and the regularity constants of the mesh.

For globally quasi-uniform meshes, the result follows directly from the Bramble-Hilbert Lemma and an inverse inequality. The proof for locally quasi-uniform meshes has been given in [4]; see [3] for generalizations including higher order approximations.

We will also require the following approximation error estimates.
Lemma 5. Let (A3)-(A4) hold. Then
\[ \| u - \pi_h u \|_{H^s(\Omega)} \leq Ch^{k-s}\| u \|_{H^k(\Omega)} \]
for all \( u \in H^k(\Omega) \) with \( 0 \leq k \leq 2 \) and \(-1 \leq s \leq \min\{1, k\}\).

Proof. For completeness, we sketch the main steps. The case \( s = 0 \) and \( 0 \leq k \leq 2 \) is well known and follows from the Bramble-Hilbert lemma and scaling arguments. To show the estimate for \( s = 1 \) and \( 1 \leq k \leq 2 \), let us denote by \( \pi_h^1 : H^1_0(\Omega) \rightarrow V_h \) the \( H^1 \)-orthogonal projection defined by
\[ (\pi_h^1 u, v_h)_{H^1(\Omega)} = (u, v_h)_{H^1(\Omega)} \quad \text{for all } v_h \in V_h. \]
Recall that \( \| \pi_h^1 u - u \|_{H^1(\Omega)} \leq C'h^{k-1}\| u \|_{H^k(\Omega)} \) for \( 1 \leq k \leq 2 \), which is the usual finite element error estimate [5]. We can then proceed by
\[ \| u - \pi_h u \|_{H^1(\Omega)} \leq \| u - \pi_h \pi_h^1 u \|_{H^1(\Omega)} + \| \pi_h (u - \pi_h^1 u) \|_{H^1(\Omega)} \]
\[ \leq (1 + C')\| u - \pi_h^1 u \|_{H^1(\Omega)} \leq (1 + C')Ch^k\| u \|_{H^k(\Omega)}, \]
where we used the projection property and the \( H^1 \)-stability of \( \pi_h \), and the approximation properties of \( \pi_h^1 \) in the last two steps. Now assume that \( u \in H^k(\Omega) \) with \( 0 \leq k \leq 2 \). Then
\[ \| u - \pi_h u \|_{H^{-1}(\Omega)} = \sup_{v \in H_0^1(\Omega)} (u - \pi_h u, v)_{L^2(\Omega)} / \| v \|_{H^1(\Omega)} \]
\[ = \sup_{v \in H_0^1(\Omega)} (u, v - \pi_h v)_{L^2(\Omega)} / \| v \|_{H^1(\Omega)} \leq Ch^{k+1}\| u \|_{H^k(\Omega)}. \]
Here we used the approximation property of \( \pi_h \) for \( s = 0 \). This yields the estimate for \( s = -1 \) and \( 0 \leq k \leq 2 \) and completes the proof. \( \square \)

Corresponding estimates for real valued \( s \) and \( k \) can be obtained by interpolation arguments. Also note that the dependence on \( h \) in the above estimates can be localized.

4. Galerkin semi-discretization

As a first step in the approximation process, let us investigate the semi-discretization in space by finite elements. Proceeding in a standard fashion, we define

Problem 6 (Semi-discretization).

Find \( u_h \in H^1(0, T; V_h) \) such that \( u_h(0) = \pi_h u_0 \) and
\[ (u_h'(t), v_h)_{L^2(\Omega)} + a(u_h(t), v_h; t) = \langle f(t), v_h \rangle_{H^{-1}(\Omega) \times H_0^1(\Omega)} \]
for all test functions \( v_h \in V_h \) and a.e. \( t \in (0, T) \).

The first term could also be written as \( \langle u_h'(t), v_h \rangle_{H^{-1}(\Omega) \times H_0^1(\Omega)} \). By choosing a basis for \( V_h \), the semi-discrete problem yields an ordinary differential equation, and existence and uniqueness follow by the Picard-Lindelöf theorem. More precisely, we have

Lemma 7 (Well-posedness and discrete a-priori estimate).

Let (A1)-(A4) hold. Then Problem 6 has a unique solution \( u_h \in H^1(0, T; V_h) \) and
\[ \| u_h \|_{W(0, T)} \leq C(\| u_0 \|_{L^2(\Omega)} + \| f \|_{L^2(0, T; H^{-1}(\Omega))}). \]
The constant \( C \) in this estimate depends only on the bounds for the coefficients, the domain, and the constants characterizing the regularity of the mesh.
Proof. It remains to verify the a-priori estimate. By testing the discrete variational problem in the usual way with \( v_h = u_h(t) \), one can show that

\[
\frac{1}{2} \frac{d}{dt} \| u_h(t) \|_{L^2(\Omega)}^2 + \frac{\alpha}{2} \| u_h(t) \|_{H^1(\Omega)}^2 \leq \eta \| u \|_{L^2(\Omega)}^2 + \frac{1}{2\alpha} \| f(t) \|_{L^2(\Omega)}^2.
\]

Here we only used the Gårding inequality and some elementary manipulations. Integrating over time and applying a Gronwall argument then yields the energy estimate

\[
\| u \|_{L^\infty(0,T;L^2(\Omega))}^2 + \| u \|_{L^2(0,T;H^1(\Omega))}^2 \leq C \left( \| u_0 \|_{L^2(\Omega)}^2 + \| f \|_{L^2(0,T;H^{-1}(\Omega))}^2 \right).
\]

To obtain the remaining bound for the time derivative, observe that the \( H^{-1} \)-norm of a discrete function can be expressed as

\[
\| u'_h(t) \|_{H^{-1}(\Omega)} = \sup_{v \in H^1_0(\Omega)} \frac{(u'_h(t), v)_{L^2(\Omega)}}{\| v \|_{H^1(\Omega)}} \]

Using the discrete variational problem with \( v_h = \pi_h v \) further yields

\[
(u'_h(t), \pi_h v)_{L^2(\Omega)} = (f_h(t), \pi_h v)_{H^{-1}(\Omega) \times H^1_0(\Omega)} - a(u_h(t), \pi_h v; t) \leq (\| f_h(t) \|_{H^{-1}(\Omega)} + C \| u_h(t) \|_{H^1(\Omega)}) \| \pi_h v \|_{H^1(\Omega)}.
\]

Since \( \pi_h \) is bounded on \( H^1_0(\Omega) \), we further have \( \| \pi_h v \|_{H^1(\Omega)} \leq C \| v \|_{H^1(\Omega)} \). Using these estimates in the expression for the norm of the time derivative finally yields

\[
\| u_h(t) \|_{H^{-1}(\Omega)} \leq C'(\| f_h(t) \|_{H^{-1}(\Omega)} + \| u_h(t) \|_{H^1(\Omega)}).
\]

The result then follows by integration over time and the energy estimate derived before. \( \square \)

As a direct consequence of the definition and the similarity of the continuous and the semi-discrete variational problems, we obtain

**Lemma 8** (Galerkin orthogonality).

Let \( u \) and \( u_h \) denote the solutions of Problems 2 and 2 respectively. Then

\[
\langle u(t) - u_h(t), v_h \rangle_{H^{-1}(\Omega) \times H^1_0(\Omega)} + a(u(t) - u_h(t), v_h; t) = 0
\]

for almost every \( t \in (0, T) \) and all \( v_h \in V_h \). Moreover, \( \pi_h(u(0) - u_h(0)) = 0 \).

We can now turn to the error analysis of the semi-discretization. To this end, we divide the error into an approximation error and a discrete error as

\[
u(t) - u_h(t) = (u(t) - \pi_h u(t)) + (\pi_h u(t) - u_h(t)) = (i) + (ii).
\]

Due to the approximation properties of \( \pi_h \), the first term can be bounded readily by

**Lemma 9** (Approximation error). Let (A3)-(A4) hold. Then

\[
\| u - \pi_h u \|_{W(0,T)} \leq C h \left( \| u \|_{L^2(0,T;H^2(\Omega))} + \| u' \|_{L^2(0,T;L^2(\Omega))} \right),
\]

and \( C \) only depends on the domain and the shape-regularity constants of the mesh.

Using Galerkin orthogonality and the discrete stability of the method, the discrete error component can then be bounded as usual by the approximation error as well.

**Lemma 10** (Discrete error). Let (A1)-(A4) hold. Then

\[
\| \pi_h u - u_h \|_{W(0,T)} \leq C \| u - \pi_h u \|_{W(0,T)}.
\]

Moreover, the constant \( C \) only depends on the domain, the bounds for the coefficients, and the shape-regularity of the mesh.
Proof. The discrete error \( e_h = \pi_h u(t) - u_h(t) \) satisfies
\[
(e_h(t), v_h)_{L^2(\Omega)} + a(e_h(t), v_h; t) = (\pi_h u'(t) - u'(t), v_h)_{L^2(\Omega)} + a(\pi_h u(t) - u(t), v_h; t)
\]
and also \( e_h(0) = \pi_h u(t) - u_h(t) = 0 \). Using the continuity of the bilinear form, we have
\[
\|\tilde{f}(t)\|_{H^{-1}(\Omega)} \leq C_a \|u(t) - \pi_h u(t)\|_{H^1(\Omega)}.
\]
By application of Lemma 4 to the equation for the discrete error, we thus obtain
\[
\|e_h\|_{W(0,T)} \leq C\|\tilde{f}\|_{L^2(0,T;H^{-1}(\Omega))} \leq C'\|u - \pi_h u\|_{W(0,T)},
\]
and this already proves the assertion of the lemma. \( \square \)

By combining the previous two lemmas, we readily obtain our first main result.

**Theorem 11** (Energy-norm error estimate for the semi-discretization).

Let (A1)-(A4) hold, and let \( u \) and \( u_h \) denote the solutions of Problems 2 and 6, respectively. Then
\[
\|u - u_h\|_{W(0,T)} \leq C\|u\|_{L^2(0,T;H^2(\Omega))} + \|u'\|_{L^2(0,T;L^2(\Omega))},
\]
and the constant \( C \) only depends on the domain, the bounds for the parameters, and the shape regularity of the mesh.

Let us emphasize that the regularity requirements for the solution are natural. In view of approximation properties of finite elements, they are also almost necessary to guarantee the approximation order one in the energy-norm. Also note that no additional regularity of the domain or the coefficients was required for our proofs.

5. Time stepping

Let us now turn to the time discretization of the semi-discrete Problem 6 by the backward Euler method. For ease of presentation, we only consider here uniform time steps of size \( \tau = T/N \), and therefore set \( t^n = n\tau \). To allow for evaluation of the coefficients and the data at individual points in time, we further assume that
\begin{align*}
\text{(A5)} \quad & a, b, c \in W^{1,\infty}(0, T; L^\infty(\Omega)), \\
\text{(A6)} \quad & f \in H^1(0, T; H^{-1}(\Omega)).
\end{align*}
The condition (A5) could be relaxed by using a Discontinuous-Galerkin method for the time discretization. In addition, we require that the time step is sufficiently small, i.e.,
\[
\text{(A7)} \quad \tau < 1/\eta,
\]
where \( \eta \) is the constant from the Gårding inequality [16]. This condition could be avoided by treating the lower order terms in the bilinear form in an explicit manner. To facilitate the notation, we will use in the sequel
\[
\partial_t u_{h}^{n+1} = \frac{1}{\tau}(u_{h}^{n+1} - u_{h}^{n})
\]
to denote the discrete time derivatives at \( t^{n+1} \). Applying the backward Euler scheme for the time discretization of Problem 6 then leads to

**Problem 12** (Fully discrete scheme).

Set \( u_{h}^{0} := \pi_h u_0 \) and find \( u_{h}^{n} \in V_h \) for \( 1 \leq n \leq N \), such that
\[
(\partial_t u_{h}^{n+1}, v_h)_{L^2(\Omega)} + a(u_{h}^{n+1}, v_h; t^{n+1}) = (f(t^{n+1}), v_h)_{H^{-1}(\Omega) \times H^1(\Omega)}
\]
holds for all test functions \( v_h \in V_h \).
The equation (19) amounts to an implicit time-stepping scheme and that the snapshots $u^n_h$ can be computed recursively. The assumption (A7) guarantees that the elliptic problem for each time-step is uniquely solvable. More precisely, we have

**Lemma 13** (Discrete well-posedness and a-priori estimates). Let (A1)-(A7) hold. Then Problem 12 has a unique solution $\{u^n_h\}_{0 \leq n \leq N}$, and

$$\sum_{n=1}^N \tau (\|\partial_t u^n_h\|_{H^{-1}(\Omega)}^2 + \|u^n_h\|_{H^1(\Omega)}^2) \leq C (\|u_0\|_{L^2(\Omega)} + \|f\|_{H^1(0,T;H^{-1}(\Omega))}).$$

The norm on the left hand side is a discrete version of the energy-norm $\|\cdot\|_{W(0,T)}$.

**Proof.** Testing (19) with $u^{n+1}_h$ and proceeding as in the proof of Lemma 7 yields

$$\frac{1}{2\tau} \|u^{n+1}_h\|^2_{L^2(\Omega)} + \frac{\alpha}{2} \|u^{n+1}_h\|^2_{H^1(\Omega)} \leq \frac{1}{2\tau} \|u^n_h\|^2_{L^2(\Omega)} + \eta \|u^{n+1}_h\|^2_{L^2(\Omega)} + \frac{1}{2\alpha} \|f(t^{n+1})\|^2_{H^1(\Omega)}.$$

Via a Gronwall argument, we then obtain the discrete energy estimate

$$\max_{1 \leq n \leq N} \|u^n_h\|^2_{L^2(\Omega)} + \sum_n \tau \|u^n_h\|^2_{H^1(\Omega)} \leq C (\|u_0\|^2_{H^1(\Omega)} + \sum_n \tau \|f(t^n)\|^2_{H^1(\Omega)}).$$

Due to condition (A5), the last term can be estimated via Taylor expansion by

$$\sum_n \tau \|f(t^{n+1})\|^2_{H^1(\Omega)} \leq C \left( \|f\|^2_{L^2(0,T;H^1(\Omega))} + \tau^2 \|f'\|^2_{L^2(0,T;H^1(\Omega))} \right).$$

In order to derive the estimate for $\partial_t u^n_h$, recall that

$$\|\partial_t u^{n+1}_h\|_{H^{-1}(\Omega)} = \sup_{v \in H^1_0(\Omega)} \langle \partial_t u^{n+1}_h, \pi_h v \rangle_{L^2(\Omega)} / \|v\|_{H^1(\Omega)}.$$

Using $v = \pi_h v$ as a test function in the discrete scheme (19), we further obtain

$$\langle \partial_t u^{n+1}_h, \pi_h v \rangle_{L^2(\Omega)} = \langle f(t^{n+1}), \pi_h v \rangle_{H^{-1}(\Omega) \times H^1_0(\Omega)} - a(u^{n+1}_h, \pi_h v; t^{n+1}) \leq C \|f(t^{n+1})\|_{H^{-1}(\Omega)} + C \|u^{n+1}_h\|_{H^1(\Omega)} \|\pi_h v\|_{H^1(\Omega)}.$$

Using the stability of the projection $\pi_h$, this allows to estimate the discrete time derivative by known terms. The assertion then follows from the previous estimates. \qed

Let us now turn to the derivation of error estimates for the full discretization defined in Problem 12. Similar as in the previous section, we use an error decomposition into an approximation error and a discrete error by

$$u(t^n) - u^n_h = (u(t^n) - \pi_h u(t^n)) + (\pi_h u(t^n) - u^n_h) = (i) + (ii).$$

The first component can be bounded by the approximation estimates for $\pi_h$ as follows.

**Lemma 14** (Approximation error). Let (A3)-(A4) hold. Then

$$\sum_{n=1}^N \tau (\|\partial_t u(t^n) - \partial_t \pi_h u(t^n)\|^2_{H^{-1}(\Omega)} + \|u(t^n) - \pi_h u(t^n)\|^2_{H^1(\Omega)}) \leq C (h^2 \|u\|_{L^2(0,T;H^2(\Omega))} + h^2 \|u'\|^2_{L^2(0,T;L^2(\Omega))} + \tau^2 \|u'\|^2_{L^2(0,T;H^1(\Omega))}).$$

**Proof.** Using Taylor expansion in time and the properties of the projection, we obtain

$$\sum_n \tau \|\partial_t u(t^n) - \partial_t \pi_h u(t^n)\|^2_{H^1(\Omega)} \leq C \|u' - \pi_h u'\|^2_{L^2(0,T;H^{-1}(\Omega))} \leq C' h^2 \|u'\|^2_{L^2(0,T;L^2(\Omega))}.$$
In a similar manner, we obtain for the second term
\[ \sum_n \tau \| u(t^n) - \pi_h u(t^n) \|^2_{H^1(\Omega)} \]
\[ \leq C (\| u - \pi_h u \|^2_{L^2(0,T;H^1(\Omega))} + \tau^2 \| u' - \pi_h u' \|^2_{L^2(0,T;H^1(\Omega))}) \]
\[ \leq C' (h^2 \| u \|^2_{L^2(0,T;H^2(\Omega))} + \tau^2 \| u' \|^2_{L^2(0,T;H^1(\Omega))}). \]

The assertion then follows by summing up the two contributions.

The discrete stability of the scheme (19) allows us to bound the discrete error as follows.

**Lemma 15 (Discrete error).** Let (A1)-(A7) hold. Then
\[ \sum_{n=1}^N \tau (\| \partial_t \pi_h u(t^n) - \partial_t u^n \|^2_{H^{-1}(\Omega)} + \| \pi_h u(t^n) - u^n \|^2_{H^1(\Omega)}) \]
\[ \leq C (h^2 \| u \|^2_{L^2(0,T;H^2(\Omega))} + \tau^2 \| u' \|^2_{L^2(0,T;H^1(\Omega))} + \tau^2 \| u'' \|^2_{L^2(0,T;H^{-1}(\Omega))}). \]

**Proof.** Let \( e_h^n = \partial_t \pi_h u(t^n) - u^n \) denote the discrete error. Then
\[ \partial_t (e_h^{n+1}, v_h)_{L^2(\Omega)} + a(e_h^{n+1}, v_h; t^{n+1}) = (\partial_t u(t^{n+1}) - u'(t^{n+1}), v_h)_{L^2(\Omega)} + a(\pi_h u(t^{n+1}) - u(t^{n+1}), v_h; t^{n+1}). \]

With similar arguments as in the proof of Lemma 13, we then obtain
\[ \max_{1 \leq n \leq N} \| e_h^n \|^2_{L^2(\Omega)} + \sum_n \tau \| e_h^n \|^2_{H^1(\Omega)} \]
\[ \leq C \sum_n \tau (\| \partial_t u(t^n) - u'(t^n) \|^2_{H^{-1}(\Omega)} + \| \pi_h u(t^n) - u(t^n) \|^2_{H^1(\Omega)}). \]

Using Taylor expansion, the first term on the right hand side can be further bounded by
\[ \sum_n \tau \| \partial_t u(t^n) - u'(t^n) \|^2_{H^{-1}(\Omega)} \leq \frac{\tau^2}{2} \| u'' \|^2_{L^2(0,T;H^{-1}(\Omega))}. \]

In a similar way, we obtain for the second term
\[ \sum_n \tau \| \pi_h u(t^{n+1}) - u(t^{n+1}) \|^2_{H^1(\Omega)} \]
\[ \leq C (\| \pi_h u - u \|^2_{L^2(0,T;H^1(\Omega))} + \tau^2 \| \pi_h u' - u' \|^2_{L^2(0,T;H^1(\Omega))}) \]
\[ \leq C' (h^2 \| u \|^2_{L^2(t^n,t^{n+1};H^2(\Omega))} + \tau^2 \| u' \|^2_{L^2(t^n,t^{n+1};H^1(\Omega))}). \]

Here we employed Taylor expansions and the approximation and stability properties of the projection \( \pi_h \). By combination of the two estimates we obtain
\[ \max_{1 \leq n \leq N} \| \pi_h u(t^n) - u^n \|^2_{L^2(\Omega)} + \sum_n \tau \| \pi_h u(t^n) - u^n \|^2_{H^1(\Omega)} \]
\[ \leq C (h^2 \| u \|^2_{L^2(0,T;H^2(\Omega))} + \tau^2 \| u' \|^2_{L^2(0,T;H^1(\Omega))} + \tau^2 \| u'' \|^2_{L^2(0,T;H^{-1}(\Omega))}). \]

Using the characterization of the \( H^{-1} \)-norm for finite element functions, the definition of the discrete scheme (19), Galerkin-orthogonality, and some basic manipulations, we can further bound the discrete time derivative terms by
\[ \sum_n \tau \| \partial_t \pi_h u(t^n) - u^n \|^2_{H^{-1}(\Omega)} \]
\[ \leq C \sum_n \tau (\| \partial_t u(t^n) - u'(t^n) \|^2_{H^{-1}(\Omega)} + \| u(t^n) - \pi_h u(t^n) \|^2_{H^1(\Omega)} + \| u^n - \pi_h u(t^n) \|^2_{H^1(\Omega)}). \]

The first and second term on the right hand side can be estimated as above, and the third term is an approximation error which has already been bounded.
Summing up the two estimates for approximation error and the discrete error yields

**Lemma 16** (Discrete energy-norm error estimate). Let (A1)-(A7) hold. Then

\[
\sum_{n=1}^{N} \tau \left( \| \partial_t u(t^{n+1}) - \partial_t u^n_h \|_{H^{-1}(\Omega)}^2 + \| u(t^{n+1}) - u^n_h \|_{H^1(\Omega)}^2 \right) \\
\leq C \left( h^2 \| u \|_{L^2(0,T;H^2(\Omega))}^2 + \| u' \|_{L^2(0,T;L^2(\Omega))}^2 \right) + \tau^2 \| u'' \|_{L^2(0,T;H^1(\Omega))}^2 + \tau^2 \| u'' \|_{L^2(0,T;H^{-1}(\Omega))}^2).
\]

To finally obtain an estimate in the energy-norm \( \| \cdot \|_{W(0,T)} \), let us denote by

\[
\tilde{u}_h(t) = \frac{t^{n+1} - t^n}{t^{n+1} - t^n} u^n_h + \frac{t - t^n}{t^{n+1} - t^n} u^{n+1}_h, \quad t^n \leq t \leq t^{n+1},
\]

the linear interpolant of the fully discrete approximations \( u^n_h \) in time. Using the definition of \( \tilde{u}_h \) and simple manipulations, we now obtain the second main result of this paper.

**Theorem 17** (Error estimate for the full discretization). Let (A1)-(A7) hold and \( u \) be sufficiently smooth. Then

\[
\| u - \tilde{u}_h \|_{W(0,T)} \leq C \left( h^2 \| u \|_{L^2(0,T;H^2(\Omega))}^2 + \| u' \|_{L^2(0,T;L^2(\Omega))}^2 \right) + \tau^2 \| u'' \|_{L^2(0,T;H^1(\Omega))}^2 + \tau^2 \| u'' \|_{L^2(0,T;H^{-1}(\Omega))}^2).
\]

Like in the previous estimates, the constant \( C \) here only depends on the domain, on the bounds for the coefficients, and on the shape regularity of the mesh.

Let us emphasize that in view of the bounds (5) and (6), the regularity requirements for the solution are reasonable and that the rates are optimal with respect to the approximation properties of the discretization. Also note, that no additional smoothness assumptions for the domain or on the spatial regularity of the parameters were required.

### 6. Numerical Tests

For illustration of the benefits of our results, let us shortly present two irregular test problems, for which the standard error estimates (7) and (11) cannot be applied directly, while our estimates (10) and (12) still provide order optimal error estimates.

#### 6.1. Lack of smoothness

Let us consider the one-dimensional heat equation

\[
\partial_t u(x,t) = u_{xx}(x,t) + f(x,t), \quad 0 < x < 1, \quad 0 < t < 1,
\]

with homogeneous initial and boundary conditions. Here \( \partial_t u = u' \) is used for the time derivative synonymously. We assume that the exact solution has the form

\[
u(x,t) = \sum_{n \geq 1} u_n \sin(n \pi x) \sin(n^2 \pi t) \quad \text{with} \quad u_n = (1 + n^2)^{-5/4 - \varepsilon}.
\]

The parameter \( \varepsilon > 0 \) is assumed to be small. The special form of the solution allows to compute various norms of \( u \) analytically and, in particular, to show that

\[u \in L^2(0,T;H^2(\Omega)) \quad \text{and} \quad u' \in L^2(0,T;L^2(\Omega)).\]

At the same time, one can easily check that

\[u \notin L^\infty(0,T;H^{1+3\varepsilon}(\Omega)) \quad \text{and} \quad u' \notin L^1(0,T;H^{3\varepsilon}(\Omega)).\]

Due to lack of regularity, the standard a-priori estimate (4) for the semi-discretization therefore cannot be applied, in contrast to our energy-norm estimate (10), which yields

\[\| u - u_h \|_{L^\infty(0,T;L^2(\Omega))} \leq C \| u - u_h \|_{W(0,T)} \leq C' h.
\]
Note that, since \( u \notin L^\infty(0, T; H^{1+\beta}(\Omega)) \), the rate of convergence for the error in the norm of \( L^\infty(0, T; L^2(\Omega)) \) cannot be improved substantially here. For a solution of the form

\[
u(x, t) = \sum_{n \geq 1} u_n \sin(n \pi x) \sin(n \pi t) \quad \text{with} \quad u_n = (1 + n^2)^{-5/4 - \varepsilon},
\]

one can see in a similar manner that

\[
u \in L^2(0, T; H^2(\Omega)), \quad \nu' \in L^2(0, T; H^1(\Omega)), \quad \text{and} \quad \nu'' \in L^2(0, T; H^{-1}(\Omega)).
\]

However, \( \nu'' \notin L^2(0, T; L^2(\Omega)) \), and therefore the standard estimate \[11\] for the full discretization cannot be applied due to lack of regularity. On the other hand, our error estimate \[12\] for the full discretization still allows to guarantee

\[
\|u(t^n) - u_h(t^n)\|_{L^2(\Omega)} \leq \|u - \tilde{u}_h\|_{L^\infty(L^2(\Omega))} \leq C\|u - \tilde{u}_h\|_{W(0, T)} \leq C(h + \tau).
\]

Note that by interpolation one has at least \( u \in L^\infty(0, T; H^{3/2}) \) here, so the rate of convergence for the error \( \|u(t) - u_h(t)\|_{L^2(\Omega)} \) in terms of the meshsize may possibly be improved. For sufficiently smooth solutions, the estimates of \[23, 25, 20\] would in fact predict the optimal rate \( \|u(t) - u_h(t)\|_{L^2(\Omega)} \leq C(h^2 + \tau) \) here.

### 6.2. Discontinuous parameters

As a second test case, we consider a thermal diffusion problem on a square covered by an inhomogeneous medium. The governing system reads

\[
\partial_t u(x, t) = \nabla(a(x) \nabla u(x, t)) + f(x, t), \quad x \in (-1, 1)^2, \quad 0 < t < 1/2.
\]

As before, we prescribe homogeneous initial and boundary conditions. Moreover, we assume that the diffusion parameter has the form

\[
a(x) = \begin{cases} 
1, & x_1 \cdot x_2 > 0, \\
\varepsilon, & x_1 \cdot x_2 < 0,
\end{cases}
\]

where \( \varepsilon \) is some small positive constant. It is well-known \[16\], that the associated elliptic operator \( Lu = -\nabla(a \nabla u) \) for such a problem is rather irregular. More precisely: for every \( \beta > 0 \), one can choose an \( \varepsilon > 0 \), such that \( L \) is not an isomorphism from \( H^1_0(\Omega) \) to \( H^{-1+\beta}(\Omega) \). In particular, the maximal value of \( \beta \), such that

\[
\|R_h u - u\|_{L^2(\Omega)} \leq C \|u\|_{H^{1+\beta}(\Omega)}
\]

holds for arbitrary \( u \in H^1_0(\Omega) \) can be made arbitrarily small. Note that the standard estimate \[7\] does not apply directly here, since only \[21\] holds instead of \[9\]. A generalization of the error estimate to non-smooth problems however still allows to guarantee

\[
\|u - u_h\|_{L^\infty(0, T; L^2(\Omega))} \leq C h^{2\beta}.
\]

provided that the solution \( u \) is sufficiently smooth; see \[20\] for details. Since \( \beta \) can be arbitrarily small in general, this estimate is highly unsatisfactory. Our energy-norm estimate \[10\] for the semi-discrete approximation however still applies and yields

\[
\|u - u_h\|_{L^\infty(0, T; L^2(\Omega))} \leq C \|u - u_h\|_{W(0, T)} \leq C' h,
\]

provided that the solution has the required smoothness. This results hold regardless of the spatial regularity of the diffusion parameter \( a(\cdot) \). For illustration of the validity of this theoretical result, we also provide some results of numerical tests. To verify the convergence rate for the semi-discretization, we integrated the semi-discrete problem numerically in time with a very accurate time stepping scheme. The resulting errors obtained on a sequence of uniformly refined meshes are summarized in Table \[1\]. As predicted by the theory, we observe convergence of the energy-norm error with first order. The numerical tests actually yield a better convergence rate for the error in the norm of \( L^\infty(0, T; L^2(\Omega)) \),
Table 1. Errors $e_1 = \|u - u_h\|_{W(0,T)}$ and $e_2 = \|u - u_h\|_{L^\infty(0,T;L^2(\Omega))}$ obtained on a sequence of uniformly refined meshes with meshsize $h$.

| $h$   | 0.50000 | 0.25000 | 0.12500 | 0.06250 | 0.03125 | rate  |
|-------|---------|---------|---------|---------|---------|-------|
| $e_1$ | 1.65010 | 0.82062 | 0.39475 | 0.19318 | 0.09710 | 1.03  |
| $e_2$ | 0.37632 | 0.11211 | 0.02928 | 0.00741 | 0.00186 | 1.93  |

which in fact is the optimal one from an approximation point of view. We however cannot give a full explanation for this observation yet.

7. Summary

Various results concerning the numerical analysis of Galerkin approximations for parabolic problems are available in the literature; see [20] for a comprehensive overview and further references. This paper contributes to this active field with providing error estimates in the energy-norm $\|u\|_{W(0,T)} = \|u\|_{L^2(0,T;H^1)} + \|u_t\|_{L^2(0,T;H^{-1})}$, which seems to be a natural choice from an analytic point of view, but which has not been studied intensively in previous works.

In this manuscript, we considered only low order discretizations of a simple model problem. The general approach is however applicable to much wider class of problems and discretization schemes. Moreover, our arguments may also be fruitful for the derivation of a-priori error estimates in other norms [13, 19, 22, 24, 26] and for the derivation of a-posteriori error estimates [7, 8, 10, 15].

Our main motivation to consider the energy-norm, was to overcome a sub-optimality of the standard estimates (7) and (11) concerning the regularity requirements for the solution, for the domain, and for the parameters. This sub-optimality is partly due to a loose handling of time derivatives in the estimates, and, on the other hand, stems from the use of the Ritz projector in the error decomposition, which requires duality arguments and regularity of the underlying elliptic problem. We therefore utilize here the $L^2$-projection in our error splitting and carefully estimate time derivatives in the $H^{-1}$-norm.

In our presentation, we focused on a-priori error estimates in the energy-norm $\|\cdot\|_{W(0,T)}$, and we could establish optimal convergence rates under minimal regularity assumptions. By continuous embedding, we could also obtain estimates pointwise in time with the same convergence rates. In our numerical experiments, we observed a better convergence of the error in the norm of $L^\infty(0,T;L^2)$ for a particular problem, which does not follow directly from our results. For sufficiently regular problems, this better convergence is well explained by the standard results [20, 25, 23]. A justification for the case of certain irregular problems is however missing.

Acknowledgements

The author would like to gratefully acknowledge support by the German Research Foundation (DFG) via grants GSC 233, IRTG 1529, and TRR 154.

References

[1] J. P. Aubin. Behaviour of the error of the approximate solution of boundary value problems for linear elliptic operators by Galerkin’s and finite difference methods. Ann. Scuola Norm. Sup. Pisa, 21:599–637, 1967.
[2] I. Babuska and S. Ohnimus. A posteriori error estimation for the semidiscrete finite elements method of parabolic differential equations. *Comput. Methods Appl. Mech. Engrg.*, 190:4691–4712, 2001.

[3] R. E. Bank and H. Yserentant. On the $H^1$-stability of the $L^2$-projection onto finite element spaces. *Numer. Math.*, 126:361–381, 2014.

[4] J. H. Bramble, J. E. Pasciak, and O. Steinbach. On the stability of the $L^2$ projection in $H^1(\Omega)$. *Math. Comp.*, 71:147–156, 2002.

[5] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, 2008.

[6] T. Dupont. $L^2$ error estimates for projection methods for parabolic equations in approximating domains. In *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pages 313–352. Academic Press, New York, San Francisco, London, 1974.

[7] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.*, 28:43–77, 1991.

[8] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L^\infty(L^2)$ and $L^\infty(L^\infty)$. *SIAM J. Numer. Anal.*, 32:706–740, 1995.

[9] L. C. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Rhode Island, 1998.

[10] Emmanuil H. Georgoulis, Omar Lakkis, and Juha M. Virtanen. A posteriori error control for discontinuous Galerkin methods for parabolic problems. *SIAM J. Numer. Anal.*, 49:427–458, 2011.

[11] J. Douglas Jr. and T. Dupont. Galerkin methods for parabolic equations. *SIAM J. Numer. Anal.*, 7:575–626, 1970.

[12] J. Douglas Jr. and T. Dupont. $H^{-1}$ Galerkin methods for problems involving several space variables. In J. J. H. Miller, editor, *Topics in Numerical Analysis III*, pages 125–141. Academic Press, London, 1977.

[13] J. Douglas Jr., T. Dupont, and M. F. Wheeler. $H^1$-Galerkin methods for the Laplace and heat equations. In C. de Boor, editor, *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pages 383–416. Academic Press, New York, 1974.

[14] Ch. Makridakis and R. H. Nochetto. Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Numer. Anal.*, 41:1585–1594, 2003.

[15] S. Nicu and A. M. Sandig. General interface problems I-II. *Math. Meth. Appl. Sci.*, 17:395–450, 1994.

[16] J. A. Nitsche. Ein Kriterium für die Quasioptimalität des Ritzschen Verfahrens. *Numer. Math.*, 11:346–348, 1968.

[17] V. Thomee. Some convergence results for Galerkin methods for parabolic boundary value problems. In C. de Boor, editor, *Mathematical Aspects of Finite Elements*, pages 55–88. Academic Press, New York, 1974.

[18] V. Thomee. Negative norm estimates and superconvergence in Galerkin methods for parabolic problems. *Math. Comp.*, 34:93–113, 1980.

[19] V. Thomee. *Galerkin Finite Element Methods for Parabolic Problems*. 2nd ed. Number 25 in *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2006.

[20] V. Thomee and L. B. Wahlbin. On Galerkin methods in semilinear parabolic problems. *SIAM J. Numer. Anal.*, 12:378–389, 1975.

[21] V. Thomee and L. B. Wahlbin. Maximum-norm stability and error estimates in Galerkin methods for parabolic equations in one space variable. *Numer. Math.*, 41:345–371, 1983.

[22] R. S. Varga. *Functional Analysis and Approximation Theory in Numerical Analysis*. CBMS-NSF Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, 1971.

[23] M. F. Wheeler. $L^\infty$ estimates of optimal order for Galerkin methods for one dimensional second order parabolic and hyperbolic equations. *SIAM J. Numer. Anal.*, 10:908–913, 1973.

[24] M. F. Wheeler. A priori $L^2$ error estimates for Galerkin approximations to parabolic partial differential equations. *SIAM J. Numer. Anal.*, 10:723–759, 1973.

[25] M. F. Wheeler. An $H^{-1}$ Galerkin method for a parabolic problem in a single space variable. *SIAM J. Numer. Anal.*, 12:803–817, 1975.