A POSTERIORI ERROR ESTIMATES BY RECOVERED GRADIENTS IN PARABOLIC FINITE ELEMENT EQUATIONS

D. LEYKEKHMAN AND L. B. WAHLBIN

Abstract. This paper considers a posteriori error estimates by averaged gradients in second order parabolic problems. Fully discrete schemes are treated. The theory from the elliptic case as to when such estimates are asymptotically exact, on an element, is carried over to the error on an element at a given time. The basic principle is that the elliptic theory can be extended to the parabolic problems provided the time-step error is smaller than the space-discretization error. Numerical illustrations confirming the theoretical results are given. Our results are not practical in the sense that various constants can not be estimated realistically. They are conceptual in nature.

1. Introduction

Averaging techniques are widely used in order to gauge errors. In practice, they are attractive because they do not depend on the particular problem at hand, as residual methods do, and they are thus easy to implement as a post-processing to existing codes. If they work, they give error estimators, not merely error indicators; indeed, sometimes they are asymptotically exact. In the last few years research has been done to explain the success of averaging methods analytically for meshes which are not uniform. Most of this analytical work is done for elliptic problems, cf. Bank and Xu [1], [2], Carey [4], Hoffmann, Schatz, Wahlbin, and Wittum [7], and references therein.

The method of lines is a popular way to solve initial boundary value problems. Using this method the problem is first discretized in space by some method and the resulting system of ODEs is then approximated in time. A natural question to ask is whether the averaging technique error estimators work as well for time dependent problems.

This paper gives a mixed answer in the case of a parabolic initial boundary value problem. We show that if the time discretization error is sufficiently small compared to the space discretization error, then the elliptic averaging techniques work as well for parabolic problems. In fact, this caveat is natural: The basic averaging methods were developed for elliptic problems, and for the semidiscrete (time-continuous and fictitious) problems it is well-known that the elliptic projection of the solution is superclose to the semidiscrete solution with respect to gradients. Thus, it seems plausible that the elliptic theory will essentially apply provided the time-discretization error is suitably controlled. These general observations were made by Ziukas and Wiberg in [14] and nicely illustrated numerically by them. On the

2000 *Mathematics Subject Classification.* 65M60, 65M20, 65M15.

Key words and phrases. a posteriori, finite element, fully discrete, pointwise estimates, parabolic second order equation.

other hand if the time discretization error dominates the spacial one then there is no guarantee that one will benefit from using the averaging operators.

To reiterate, in contrast to the elliptic case, we now have two main sources of the error, one from the space discretization and one from the time discretization. A main technical difficulty is to unravel these two errors in a sufficiently distinct and separate fashion as to be able to apply the fine-tuned elliptic results at the single element level also to the parabolic case. For further references on analysis of parabolic a posteriori error estimates we refer the reader to the introduction in Lakkis and Makridakis [8].

To avoid misunderstanding about the purpose of this paper, we wish to say that it gives no practical clues as how to choose spatial mesh-lengths or time-steps in order to secure realistically an error a posteriori. The theory is asymptotic and replete with constants that cannot be estimated in advance or a posteriori. Our development for the parabolic case is based on the elliptic case and even there no such practical guides are available. We hope that our analysis leads to conceptual understanding. The plan of the paper is as follows. In Section 2 we state our main result. Section 3 is occupied with preliminary technical results and, as far as novel mathematics is concerned, forms the major part of the paper. The proof of the main result in Section 4 then follows essentially as in the elliptic case. The final Section 5 gives numerical illustrations. Familiarity with the elliptic case would be helpful to the reader.

2. Statement of the result

Let Ω be a bounded domain in $\mathbb{R}^N, N \geq 1$, with a sufficiently smooth boundary. Consider the second order parabolic initial boundary value problem

$$\begin{aligned} u_t(x,t) + Au(x,t) &= f(x,t), \quad x \in \Omega,\ t > 0, \\ u(x,0) &= v(x), \quad x \in \Omega, \\ \frac{\partial u(x,t)}{\partial n} &= 0, \quad x \in \partial\Omega,\ t > 0. \end{aligned} \tag{2.1}$$

For simplicity we assume $A = -\Delta + 1$, where $\Delta = \sum \partial^2/\partial x_j^2$ is the Laplacian and 1 is the identity operator.

Remark 1. The form of the operator A is not essential and can be replaced by a general uniformly elliptic second order operator. This will however make the corresponding proofs more technical. The analysis heavily relies on the Green's functions estimates for the elliptic and parabolic problems. We need the smoothness of Ω in order to guarantee such Green's functions estimates. The assumption on the regularity of f connected to the time discretization scheme used. In the forthcoming analysis $f \in L^2(\Omega)$ is sufficient.

For the spatial finite element approximation of this problem, let $\mathcal{T}_h = \{\tau_h\}$, $0 < h < 1$, be a sequence of triangulations of Ω, $\overline{\Omega} = \bigcup_{\tau \in \mathcal{T}_h} \overline{\tau}$, with the elements τ mutually disjoint. Furthermore, we assume that each τ that does not meet $\partial\Omega$ is an N-dimensional straight simplex, while those τ that meet $\partial\Omega$ are allowed to have lower-dimensional curved faces in common with $\partial\Omega$. The partitions are face-to-face so that simplices meet only in full lower-dimensional faces or not at all. The triangulations are assumed to be globally quasi-uniform, i.e. (if necessary after a renormalization of h),

$$diam\ \tau \leq h \leq C(meas\ \tau)^{1/N}, \quad \forall \tau \in \mathcal{T}_h.$$

Our finite element spaces are then the $\mathcal{C}^0$ simplicial Lagrange elements

$$S_h = S_h^{l-1}(\Omega) = \{\chi \in \mathcal{C}(\overline{\Omega}) : \chi\,|_\tau \in \Pi^{l-1}(\tau)\},$$

where $\Pi^{l-1}(D)$ denotes the set of polynomials of total degree $\leq l-1$ on D. Thus the usual inverse and approximation properties hold, which is really all we shall use. Of course, all the results from the literature that we use are valid under those assumptions.

We consider a single step fully discrete solution U^n which is defined by

$$U^{n+1} = r(kA_h)U^n + k\big(Q_{kh}P_h f\big)(t_n), \quad U^0 = R_h v, \tag{2.2}$$

where

$$Q_{kh}P_h f(t) = \sum_{i=1}^{m} p_i(kA_h)P_h f(t+\tau_i k).$$

In the formulas above we have used the following notation:

- k is a uniform time step, and $t_n = nk$ for any nonnegative integer n,
- $r(\lambda)$ and $p_i(\lambda)$ are rational functions defined on the spectrum of kA_h, and we further assume r is A-stable and $r(\infty) = 0$,
- $A_h : S_h \to S_h$ is defined by
$$(A_h v, \chi) = (\nabla v, \nabla \chi) + (v, \chi), \text{ for } \chi \in S_h,$$
- $P_h : L_2(\Omega) \to S_h$ is the L_2 projection defined by
$$(P_h v, \chi) = (v, \chi), \text{ for } \chi \in S_h,$$
- elliptic projection $R_h : H^1(\Omega) \to S_h$ is given by
$$(\nabla R_h v, \nabla \chi) + (R_h v, \chi) = (\nabla v, \nabla \chi) + (v, \chi), \text{ for } \chi \in S_h,$$
- τ_i are distinct real numbers from the interval $[0, 1]$.

To avoid the effect of unnatural boundary conditions and possible order reduction (cf. Brenner, Crouzeix, and Thomée [3]), we assume that the method (2.2) is both accurate and strictly accurate of order q, i.e. $r(\lambda)$ and $p_i(\lambda)$ satisfy the following properties as $\lambda \to 0$:

$$r(\lambda) = e^{-\lambda} + O(\lambda^{q+1}), \tag{2.3}$$

$$\sum_{i=1}^{m} \tau_i^l p_i(\lambda) = \frac{l!}{(-\lambda)^{l+1}} \left(e^{-\lambda} - \sum_{j=0}^{l} \frac{(-\lambda)^j}{j!} \right) + O(\lambda^{q-l}), \quad \text{for } 0 \leq l \leq q, \tag{2.4}$$

and

$$\sum_{i=1}^{m} \tau_i^l p_i(\lambda) = \frac{l!}{(-\lambda)^{l+1}} \left(r(\lambda) - \sum_{j=0}^{l} \frac{(-\lambda)^j}{j!} \right), \quad \text{for } 0 \leq l \leq q-1. \tag{2.5}$$

For the construction of such schemes, see [3] or Thomée [12], page 129. In fact, looking at these constructions, it is natural to assume also that the rational functions p_i vanish at infinity,

$$p_i(\lambda) = O\left(\frac{1}{\lambda}\right), \quad \text{as} \quad \lambda \to \infty. \tag{2.6}$$

Let d_H be a domain in Ω with a diameter comparable to H, where $H \geq h$. We assume that our recovered gradient operators $G_H : \mathcal{C}(d_H) \to \mathcal{C}(d_H)^N$ satisfy the following two conditions. First is a smoothing type estimate,

$$\|G_H v\|_{L_\infty(d_H)} \leq C_G H^{-1} \|v\|_{L_\infty(d_H)}, \ \forall v \in W^1_\infty(\bar{d}_H), \tag{2.7}$$

and then is the approximation property,

$$\|\nabla v - G_H v\|_{L_\infty(d_H)} \leq C_G H^r \|v\|_{W^{r+1}_\infty(d_H)}, \ \forall v \in \mathcal{C}^{r+1}_\infty(\bar{d}_H). \tag{2.8}$$

Several examples of such operators are given in [7].

Our local error estimator is constructed as follows. Take any $\tau \in \mathcal{T}_h$, $\tau \subset d_H$ and set

$$\mathcal{E}(\tau) = \|\nabla U^n - G_H U^n\|_{L_\infty(\tau)}.$$

This is to be an estimator for $\|\nabla e^n\|_{L_\infty(\tau)}$, $e^n = U^n - u(t_n)$, the real gradient error on the element. The parameter $0 < \varepsilon < 1$, in the statement of the theorem below, is for you to choose, essentially to determine how close to pure approximation theory you want to be. With ε near 1, you are closest. Note that Alternative 1 ("nondegeneracy") or Alternative 2 ("degeneracy") below are tied to your choice of ε. Here is now our main result:

Theorem 2.1. *Let $r \geq 3$ and $0 < \varepsilon < 1$. Let U^n be the one step fully discrete approximate solution to (2.1) at time $t_n = kn$ given by (2.2), which is both accurate and strictly accurate of order q, and $r(z)$ is A-stable with $r(\infty) = 0$ and the p_i satisfy (2.6). Then, provided $h^\varepsilon \leq 1/2$, $k^q \ll h^{r+1-\varepsilon}$, and t_n is bounded, there exist constants C_1 and C_2 independent of h and k such that for each $\tau \in \mathcal{T}_h$, and with*

$$m = C_1(h/H + (H/h)^r h^\varepsilon),$$

we have one of the following two alternatives:

Alternative 1. *Suppose on the element τ, the function u satisfies the nondegeneracy condition*

$$|u(t_n)|_{W^r_\infty(\tau)} \geq h^{1-\varepsilon} \left(\|u(t_n)\|_{W^{r+1}_\infty(\Omega)} + \sup_{s \leq t_n} \|u_t(s)\|_{W^r_\infty(\Omega)} \right) := h^{1-\varepsilon}[u; t_n]. \tag{2.9}$$

In this case

$$\|\nabla u(t_n) - G_H U^n\|_{L_\infty(\tau)} \leq m \|\nabla e^n\|_{L_\infty(\tau)}, \tag{2.10}$$

$$\mathcal{E}(\tau) \leq (1 + m) \|\nabla e^n\|_{L_\infty(\tau)}, \tag{2.11}$$

and, if $m < 1$,

$$\frac{1}{1+m} \mathcal{E}(\tau) \leq \|\nabla e^n\|_{L_\infty(\tau)} \leq \frac{1}{1-m} \mathcal{E}(\tau). \tag{2.12}$$

If $H = H(h)$ is chosen so that $m \to 0$ as $h \to 0$, the estimator is asymptotically exact.

Alternative 2. *Suppose (2.9) does not hold, i.e.*

$$|u(t_n)|_{W^r_\infty(\tau)} < h^{1-\varepsilon}[u; t_n]. \tag{2.13}$$

In this case

$$\|\nabla e^n\|_{L_\infty(\tau)} \leq C_2 h^{r-\varepsilon}[u; t_n], \tag{2.14}$$

$$\|\nabla u(t_n) - G_H U^n(t_n)\|_{L_\infty(\tau)} \leq m h^{r-\varepsilon}[u; t_n], \tag{2.15}$$

and

$$\mathcal{E}(\tau) \leq (m + C_2) h^{r-\varepsilon}[u; t_n]. \tag{2.16}$$

Thus in the "degenerate" Alternative 2, both the real error and the error estimator are very small, but there may be no connection between them. (cf. our numerical illustration in Section 5, Example 4).

Remark 2. To keep this paper short, we only consider approximation degree $r \geq 3$ and recovered gradient operators G_H that satisfy (2.7) and (2.8). As for the piecewise linear case, $r = 2$, it could have been included using the elliptic results of [11]. The quadrature weighted Zienkiewicz-Zhu (ZZ) operator, which fails to satisfy the smoothing property (2.7), could also been included using the elliptic results of [4]. However, the most widely used form of the ZZ operator, the unweighted one, still eludes serious analysis except for meshes with certain degrees of uniformity where superconvergence phenomena can be used in the analysis.

3. Preliminary Results

In this section we collect results which are essential in order to prove the main result of the paper. We will use the notation

$$\ell_h = |\log h|$$

throughout the paper.

The first result is the Green's function estimate for the continuous problem (2.1).

Lemma 3.1. *If $f \equiv 0$, the solution of (2.1) may be represented in terms of a Green's function $G(x, y; t, 0), t > 0, x, y \in \Omega$, as*

$$u(x,t) = \int_\Omega G(x, y; t, 0)\, v(y)\, dy.$$

Assume that the boundary $\partial\Omega$ is sufficiently smooth. Then for any integer l_0 and multi-integer l, there exist constants C and $c > 0$ such that for the Green's function $G(x, y; t, s), 0 \leq s < t$, and $x, y \in \Omega$, we have

$$|D_t^{l_0} D_x^l G(x, y; t, s)| \leq C(|x - y| + \sqrt{t - s})^{-(N+2l_0+|l|)} e^{-c\frac{|x-y|^2}{t-s}}.$$

A proof is given in Èĭdel'man and Ivasišen [5].

The next result that we need is Lemma 3.3 in [7].

Lemma 3.2. *There exists a constant $\widehat{C}_1$ independent of h such that*

$$h^{r-1}|u(t_n)|_{W^r_\infty(\tau)} - h^r \|u(t_n)\|_{W^{r+1}_\infty(\tau)} \leq \widehat{C}_1 \|\nabla e^n\|_{L_\infty(\tau)}.$$

In the next lemma we will show stability and smoothing type estimates of the homogeneous solution operator $E_h(t)$ for the semicontinuous problem in the W^1_∞ norm.

Lemma 3.3. *Let $E_h(t) = e^{-A_h t}$. There exists a constant C independent of h such that*

$$\|E_h(t)v\|_{W^1_\infty(\Omega)} + t\ell_h^{-1}\|E_h'(t)v\|_{W^1_\infty(\Omega)} \leq C\|v\|_{W^1_\infty(\Omega)}, \quad \forall v \in S_h,\ t \geq 0.$$

Proof. Let $u(t) = e^{-At}v$ and $u_h(t) = e^{-A_h t}v$, where $v \in S_h$. By the triangle inequality,

$$\|u_h(t)\|_{W^1_\infty(\Omega)} \le \|u_h(t) - u(t)\|_{W^1_\infty(\Omega)} + \|u(t)\|_{W^1_\infty(\Omega)}.$$

With a suitable $\chi \in S_h$, using the triangle and inverse inequalities, we have

$$\begin{aligned}
\|u_h(t) - u(t)\|_{W^1_\infty(\Omega)} &\le \|u_h(t) - \chi\|_{W^1_\infty(\Omega)} + \|\chi - u(t)\|_{W^1_\infty(\Omega)} \\
&\le \frac{C}{h}\|u_h(t) - \chi\|_{L_\infty(\Omega)} + \|\chi - u(t)\|_{W^1_\infty(\Omega)} \\
&\le \frac{C}{h}\|u_h(t) - u(t)\|_{L_\infty(\Omega)} + \frac{C}{h}\|u(t) - \chi\|_{L_\infty(\Omega)} + \|\chi - u(t)\|_{W^1_\infty(\Omega)} \\
&\le \frac{C}{h}\|u_h(t) - u(t)\|_{L_\infty(\Omega)} + C\|u(t)\|_{W^1_\infty(\Omega)}.
\end{aligned}$$

Thus,

$$\|u_h(t)\|_{W^1_\infty(\Omega)} \le \frac{C}{h}\|u_h(t) - u(t)\|_{L_\infty(\Omega)} + C\|u(t)\|_{W^1_\infty(\Omega)}.$$

By Corollary 2.4 in [10], $\|u_h(t)-u(t)\|_{L_\infty(\Omega)} \le Ch\|u\|_{W^1_\infty(Q_t)}$, where $Q_t = \Omega\times[0,t]$. Hence,

$$\|u_h(t)\|_{W^1_\infty(\Omega)} \le C\|u\|_{W^1_\infty(Q_t)}.$$

Next we will show $\|u\|_{W^1_\infty(Q_t)} \le C\|v\|_{W^1_\infty(\Omega)}$.

Given any initial value v the corresponding solution u to the homogeneous problem can be represented using the Green's function as

$$u(x,t) = \int_\Omega G(x,y;t,0)v(y)dy,$$

and hence

$$D_x u(x,t) = \int_\Omega D_x G(x,y;t,0)v(y)dy.$$

It is easy to see that if $v \equiv 1$, then $u(x,t) = e^{-t}$, and

$$\int_\Omega D_x G(x,y;t,0)dy = 0.$$

Thus

$$D_x u(x,t) = \int_\Omega D_x G(x,y;t,0)v(y)dy = \int_\Omega D_x G(x,y;t,0)(v(y) - v(x))dy. \tag{3.1}$$

By the Mean Value Theorem and the Green's function estimate in Lemma 3.1, we have

$$|D_x u(x,t)| \le C\|v\|_{W^1_\infty(\Omega)} \int_\Omega \frac{e^{-c\frac{|x-y|^2}{t}}|y-x|}{(|x-y| + \sqrt{t})^{N+1}}dy \le C\|v\|_{W^1_\infty(\Omega)},$$

which proves the stability of $E_h(t)$.

Next we will show the smoothing property.

Let $w_h(t) = tu_{h,t}(t)$ and $w(t) = tu_t(t)$. Then

$$w_{h,t} + A_h w_h = u_{h,t}, \quad w_h(0) = 0,$$

and

$$w_t + Aw = u_t, \quad w(0) = 0.$$

Let $\eta := w_h - R_h w$. Since $\frac{\partial w}{\partial n} \mid_{\partial\Omega}= 0$, we have $P_h A w = A_h R_h w$. Thus $\eta \in S_h$ satisfies

$$\eta_t + A_h\eta = (P_h - R_h)w_t + (u_h - P_h u)_t$$
$$\eta(0) = 0.$$

By Duhamel's principle

$$\eta(t) = \int_0^t E_h(t-s)\left[(P_h - R_h)w + (u_h - P_h u)\right]_t(s)ds.$$

Integrating by parts and using that $v \in S_h$ and $w(0) = 0$, we have

$$\eta(t) = \int_0^t E_h'(t-s)\big[(P_h - R_h)w + (u_h - P_h u)\big](s)ds + (P_h - R_h)w(t) + (u_h - P_h u)(t)$$
$$= I_1 + I_2 + I_3.$$

By Theorem 2.1 in [10], $\|E_h'(t)\|_{L_\infty} \leq \frac{C}{t+h^2}$. Also using Corollary 2.4 in [10] for $(u_h - P_h u)(t)$ and the already proved stability result, we have

$$\|I_1\|_{L_\infty(\Omega)} \leq \int_0^t \frac{C}{t-s+h^2}\left[h\|w\|_{W^1_\infty(Q_t)} + h\|u\|_{W^1_\infty(Q_t)}\right]ds$$
$$\leq C\ell_h h(\|w\|_{W^1_\infty(Q_t)} + \|u\|_{W^1_\infty(Q_t)}) \leq C\ell_h h(\|w\|_{W^1_\infty(Q_t)} + \|v\|_{W^1_\infty(\Omega)}).$$

Similarly

$$\|I_2\|_{L_\infty(\Omega)} + \|I_3\|_{L_\infty(\Omega)} \leq C\ell_h h(\|w\|_{W^1_\infty(Q_t)} + \|v\|_{W^1_\infty(\Omega)}).$$

Thus, by an inverse estimate,

$$\|\eta(t)\|_{W^1_\infty(\Omega)} \leq Ch^{-1}\|\eta(t)\|_{L_\infty(\Omega)} \leq C\ell_h(\|w\|_{W^1_\infty(Q_t)} + \|v\|_{W^1_\infty(\Omega)}).$$

Since $w_h = \eta + R_h w$ and we have stability of R_h in the W^1_∞ norm, the only thing that is left is to estimate $\|w(t)\|_{W^1_\infty(\Omega)}$. From (3.1),

$$D_t D_x u(x,t) = \int_\Omega D_t D_x G(x,y;t,0)(v(y) - v(x))dy$$
$$\leq C\|v\|_{W^1_\infty(\Omega)} \int_\Omega \frac{e^{-c\frac{|x-y|^2}{t}}|y-x|}{(|x-y| + \sqrt{t})^{N+3}}dy \leq \frac{C}{t}\|v\|_{W^1_\infty(\Omega)}.$$

Hence, $\|w\|_{W^1_\infty(Q_t)} \leq C\|v\|_{W^1_\infty(\Omega)}$ and we have the smoothing property

$$\|E_h'(t)v\|_{W^1_\infty(\Omega)} \leq C\frac{\ell_h}{t}\|v\|_{W^1_\infty(\Omega)}.$$

This completes the proof of Lemma 3.3. □

Remark 3. Actually, in our present work, we only use the stability part of Lemma 3.3. We have included the smoothing part for completeness and possible future references.

Our next lemma is an intermediate derivatives estimate on the finite element spaces. The second derivatives on the finite element spaces are interpreted in the sense of A_h; a logarithmic factor appears due to this.

Lemma 3.4. *There exists a constant C independent of h, such that for all $0 < \delta \leq 1$ and $\chi \in S_h$,*

$$\|\chi\|_{W^1_\infty(\Omega)} \leq C\ell_h\left((h+\delta)\|A_h\chi\|_{L_\infty(\Omega)} + \frac{1}{\delta+h}\|\chi\|_{L_\infty(\Omega)}\right).$$

Proof. Using an inverse estimate, for any $2 < p < \infty$,

$$\|\chi\|_{W^1_\infty(\Omega)} \le Ch^{-N/p}\|\chi\|_{W^1_p(\Omega)}, \text{ for any } \chi \in S_h. \tag{3.2}$$

Now with $T = A^{-1}$ and $T_h = A_h^{-1}P_h$, using approximation properties of the elliptic projection and elliptic regularity theory,

$$\begin{aligned}\|\chi\|_{W^1_p(\Omega)} &= \|T_hA_h\chi\|_{W^1_p(\Omega)} \le \|(T_h - T)A_h\chi\|_{W^1_p(\Omega)} + \|TA_h\chi\|_{W^1_p(\Omega)}\\ &\le Ch\|TA_h\chi\|_{W^2_p(\Omega)} + \|TA_h\chi\|_{W^1_p(\Omega)} \le Chp\|A_h\chi\|_{L_\infty(\Omega)} + \|TA_h\chi\|_{W^1_p(\Omega)}.\end{aligned}$$

From (3.2), then

$$\|\chi\|_{W^1_\infty(\Omega)} \le Ch^{-N/p}\left(Chp\|A_h\chi\|_{L_\infty(\Omega)} + \|TA_h\chi\|_{W^1_p(\Omega)}\right). \tag{3.3}$$

For any $0 < \delta \le 1$, by standard intermediate derivatives estimates and again elliptic regularity,

$$\begin{aligned}\|TA_h\chi\|_{W^1_p(\Omega)} &\le C\left(\delta\|TA_h\chi\|_{W^2_p(\Omega)} + \frac{1}{\delta}\|TA_h\chi\|_{L_p(\Omega)}\right)\\ &\le C\left(\delta p\|A_h\chi\|_{L_\infty(\Omega)} + \frac{1}{\delta}\|TA_h\chi\|_{L_p(\Omega)}\right),\end{aligned}$$

so that using (3.3),

$$\|\chi\|_{W^1_\infty(\Omega)} \le Ch^{-N/p}\left((\delta + h)p\|A_h\chi\|_{L_\infty(\Omega)} + \frac{1}{\delta}\|TA_h\chi\|_{L_p(\Omega)}\right). \tag{3.4}$$

Finally,

$$\begin{aligned}\|TA_h\chi\|_{L_p(\Omega)} &\le \|(T - T_h)A_h\chi\|_{L_p(\Omega)} + \|\chi\|_{L_p(\Omega)}\\ &\le Ch^2p\|A_h\chi\|_{L_\infty(\Omega)} + \|\chi\|_{L_\infty(\Omega)} \le Cp\|\chi\|_{L_\infty(\Omega)}.\end{aligned}$$

Thus, from (3.4), choosing $p = \ell_h$,

$$\|\chi\|_{W^1_\infty(\Omega)} \le C\ell_h\left((\delta + h)\|A_h\chi\|_{L_\infty(\Omega)} + \frac{1}{\delta}\|\chi\|_{L_\infty(\Omega)}\right).$$

Of course, if $\delta \le h$, then by an inverse estimate we have

$$\|\chi\|_{W^1_\infty(\Omega)} \le \frac{C}{h}\|\chi\|_{L_\infty(\Omega)}.$$

This completes the proof of the lemma. □

The next result shows a localized property of the gradient of the error of the fully discrete solution.

Proposition 3.5. *Let u satisfy (2.1) and let U^n be the fully discrete solution computed at time $t_n = nk$ by (2.2). Then, for any $0 < \varepsilon < 1$, there exists a constant C_ε independent of u, k and h, such that for any $x \in \Omega$,*

$$\begin{aligned}\left|\nabla e^n(x)\right| &\le C_\varepsilon h^{r-1}\left(\sum_{|\alpha|=r}|D^\alpha_x u(x,t_n)| + h^{1-\varepsilon}\|u(t_n)\|_{W^{r+1}_\infty(\Omega)}\right)\\ &\quad + C\ell_h h^r\sqrt{t_n}\sup_{s\le t_n}\|u_t(s)\|_{W^r_\infty(\Omega)} + \boldsymbol{P}_1,\end{aligned}$$

where

$$\boldsymbol{P}_1 = Ck^q\int_0^{t_n}\left(\left\|Au^{(q)}(s)\right\|_{W^1_\infty(\Omega)} + \left\|u^{(q+1)}(s)\right\|_{W^1_\infty(\Omega)}\right)ds.$$

Proof. Using the triangle inequality we have

$$\left|\nabla\big(u(x,t_n)-U^n(x)\big)\right| \le \left|\nabla\big(u(x,t_n)-R_hu(x,t_n)\big)\right| + \left|\nabla\big(R_hu(x,t_n)-U^n(x)\big)\right|.$$

From Theorem 4.2 in [9], for any $0<\varepsilon<1$,

$$\left|\nabla\big(u(x,t_n)-R_hu(x,t_n)\big)\right| \le C_\varepsilon h^{r-1}\left(\sum_{|\alpha|=r}|D_x^\alpha u(x,t_n)| + h^{1-\varepsilon}\ \|u(t_n)\|_{W_\infty^{r+1}(\Omega)}\right).$$

In order to show the desired bound for $\left|\nabla\big(R_hu(x,t_n)-U^n(x)\big)\right|$, we will follow the proof of Theorem 9.6 in [12]. First we note that $w_h = R_hu$ satisfies

$$w_{h,t} + A_hw_h = R_hu_t + P_hAu = P_h(f+\rho_t) := g_h,$$

where $\rho = R_hu - u$.

Let

$$W^{n+1} = r(kA_h)W^n + k\big(Q_{kh}g_h\big)(t_n),\ W^0 = R_hv,$$

be the corresponding fully discrete solution to w_h. Since by Lemma 3.3, $E_h(t)$ is stable in W_∞^1 norm, from Theorem 1 in [3] with $p_0 = q$, we have

$$\big\|W^n - w_h(t_n)\big\|_{W_\infty^1(\Omega)} \le Ck^q\int_0^{t_n}\left(\big\|A_hw_h^{(q)}(s)\big\|_{W_\infty^1(\Omega)} + \|w_h^{(q+1)}(s)\|_{W_\infty^1(\Omega)}\right)ds.$$

Since $A_hR_h = P_hA$ and the fact that R_h and P_h are stable in W_∞^1 norm, we have

$$\big\|W^n - w_h(t_n)\big\|_{W_\infty^1(\Omega)} \le Ck^q\int_0^{t_n}\left(\big\|Au^{(q)}(s)\big\|_{W_\infty^1(\Omega)} + \big\|u^{(q+1)}(s)\big\|_{W_\infty^1(\Omega)}\right)ds.$$

It remains to consider $Z^n = U^n - W^n$, which satisfies

$$Z^{n+1} = r(kA_h)Z^n + k\big(Q_{kh}P_h\rho_t\big)(t_n), \quad Z^0 = 0.$$

Thus,

(3.5)

$$\big\|Z^n\big\|_{W_\infty^1(\Omega)} \le k\big\|Q_{kh}P_h\rho_t(t_{n-1})\big\|_{W_\infty^1(\Omega)} + k\sum_{j=0}^{n-2}\big\|r(kA_h)^{n-1-j}Q_{kh}P_h\rho_t(t_j)\big\|_{W_\infty^1(\Omega)}.$$

We start by bounding the sum on the right. First we will show the following smoothing type estimate

$$\|r^n(kA_h)\chi\|_{W_\infty^1(\Omega)} \le C\frac{\ell_h}{\sqrt{kn}}\|\chi\|_{L_\infty(\Omega)},\ \text{for any } \chi\in S_h. \tag{3.6}$$

We note that, by the result of [10], A_h satisfies a resolvent estimate in L_∞ norm which is logarithm free. Hence from Theorem 2.2 in Hansbo [6],

$$\|A_hr^n(kA_h)\chi\|_{L_\infty(\Omega)} \le \frac{C}{kn}\|\chi\|_{L_\infty(\Omega)}, \tag{3.7}$$

and since $r(\lambda)$ is A-stable, by Theorem 8.2 in [12],

$$\|r^n(kA_h)\chi\|_{L_\infty(\Omega)} \le C\|\chi\|_{L_\infty(\Omega)}.$$

Taking $\delta = \sqrt{nk+h^2}$ in Lemma 3.4 proves (3.6).

Using (3.6), the boundedness of Q_{kh} and of P_h in L_∞ norm, we obtain

$$k\sum_{j=0}^{n-2}\big\|r(kA_h)^{n-1-j}Q_{kh}P_h\rho_t(t_j)\big\|_{W^1_\infty(\Omega)} \le C\ell_h k\sum_{j=0}^{n-2}\frac{\sup_{t_j\le s\le t_{j+1}}\big\|\rho_t(s)\big\|_{L_\infty(\Omega)}}{\sqrt{k(n-1-j)}}$$
$$\le C\ell_h h^r\sqrt{k}\sup_{s\le t_n}\|u_t(s)\|_{W^r_\infty(\Omega)}\sum_{j=0}^{n-2}\frac{1}{\sqrt{n-1-j}} \le C\ell_h h^r\sqrt{t_n}\sup_{s\le t_n}\|u_t(s)\|_{W^r_\infty(\Omega)}.$$

To bound the first term on the right in (3.5), we recall (2.6), which implies similarly to (3.6) that

$$k\big\|Q_{kh}P_h\rho_t(t_{n-1})\big\|_{W^1_\infty(\Omega)} \le C\ell_h\sqrt{k}\sup_{t_{n-1}\le s\le t_n}\big\|P_h\rho_t(s)\big\|_{L_\infty(\Omega)}$$
$$\le C\ell_h\sqrt{k}h^r\sup_{t_{n-1}\le s\le t_n}\big\|u_t(s)\big\|_{W^r_\infty(\Omega)}.$$

Thus we have shown the desired bound for $\|Z^n\|_{W^1_\infty(\Omega)}$, and the proof of the proposition is complete. $\square$

We shall also need the similar result for function values.

Proposition 3.6. *There exists a constant C_ε independent of u, k and h, such that for any $x\in\Omega$,*

$$\big|e^n(x)\big| \le Ch^r\left(\sum_{|\alpha|=r}|D^\alpha_x u(x,t_n)| + h^{1-\varepsilon}\|u(t_n)\|_{W^{r+1}_\infty(\Omega)}\right)$$
$$+ Ch^{r+1}\log(n+1)\sup_{s\le t_n}\|u_t(s)\|_{W^r_\infty(\Omega)} + \boldsymbol{P}_2,$$

where

$$\boldsymbol{P}_2 = Ck^q\int_0^{t_n}\left(\big\|Au^{(q)}(s)\big\|_{L_\infty(\Omega)} + \big\|u^{(q+1)}(s)\big\|_{L_\infty(\Omega)}\right)ds.$$

Proof. Similarly to the proof of the previous proposition, from the triangle inequality we have

$$\big|u(x,t_n)-U^n(x)\big| \le \big|u(x,t_n)-R_hu(x,t_n)\big| + \big|R_hu(x,t_n)-U^n(x)\big|.$$

By Theorem 4.1 in [9], for any $0<\varepsilon<1$,

$$\big|u(x,t_n)-R_hu(x,t_n)\big| \le Ch^r\left(\sum_{|\alpha|=r}|D^\alpha_x u(x,t_n)| + h^{1-\varepsilon}\ \|u(t_n)\|_{W^{r+1}_\infty(\Omega)}\right).$$

To bound $\big|R_hu(x,t_n)-U^n(x)\big|$, we again proceed analogously to the proof of the previous proposition. Thus we obtain

$$\big\|W^n-w_h(t_n)\big\|_{L_\infty(\Omega)} \le Ck^q\int_0^{t_n}\left(\big\|Au^{(q)}(s)\big\|_{L_\infty(\Omega)} + \big\|u^{(q+1)}(s)\big\|_{L_\infty(\Omega)}\right)ds,$$

where $w_h = R_hu$ and W^n is its fully discrete solution.

Now $Z^n = U^n - W^n$ with $\rho = R_hu - u$ satisfies

$$Z^n = k\sum_{j=0}^{n-1}r(kA_h)^{n-j-1}Q_{kh}P_h\rho_t(t_j).$$

Thus,

(3.8)

$$\begin{aligned}\big\|Z^n\big\|_{L_\infty(\Omega)} &\le k\big\|Q_{kh}P_h\rho_t(t_{n-1})\big\|_{L_\infty(\Omega)} + k\sum_{j=0}^{n-2}\big\|r(kA_h)^{n-1-j}Q_{kh}P_h\rho_t(t_j)\big\|_{L_\infty(\Omega)}\\ &= I_1 + I_2.\end{aligned}$$

To bound the sum, we use that Q_{kh} and A_h commute and the operator identity $A_h^{-1}P_h = R_hA^{-1}$. Hence we can write

$$I_2 = k\sum_{j=0}^{n-2}\big\|A_hr(kA_h)^{n-1-j}Q_{kh}R_hA^{-1}\rho_t(t_j)\big\|_{L_\infty(\Omega)}.$$

Recalling (3.7),

$$\|A_hr^n(kA_h)\chi\|_{L_\infty(\Omega)} \le \frac{C}{kn}\|\chi\|_{L_\infty(\Omega)},$$

and using the boundedness of Q_{kh} and R_h in L_∞ norm for $r\ge 3$, we have

$$I_2 \le C\sum_{j=0}^{n-2}\frac{\sup_{t_j\le s\le t_{j+1}}\big\|A^{-1}\rho_t(s)\big\|_{L_\infty(\Omega)}}{n-1-j}.$$

Next we will estimate $\|A^{-1}\rho_t(s)\|_{L_\infty(\Omega)}$. Setting up a duality argument, we write

$$\|A^{-1}(R_h-I)v\|_{L_\infty} = \sup_{\|\psi\|_{L_1}=1}\big(A^{-1}(R_h-I)v,\psi\big).$$

Using that A^{-1} is self adjoint and the definition of the elliptic projection, we have for any $\chi\in S_h$ and fixed ψ,

$$\begin{aligned}&\big(A^{-1}(R_h-I)v,\psi\big) = \big((R_h-I)v,A^{-1}\psi\big) = \big((R_h-I)v,AA^{-2}\psi\big)\\ &= \big(\nabla(R_h-I)v,\nabla(A^{-2}\psi-\chi)\big) + \big((R_h-I)v,A^{-2}\psi-\chi\big)\\ &\le \|R_hv-v\|_{W^1_\infty(\Omega)}\|A^{-2}\psi-\chi\|_{W^1_1(\Omega)}.\end{aligned}$$

By approximation theory, we find that

$$\|A^{-1}\rho_t(s)\|_{L_\infty(\Omega)} \le Ch^{r-1}\|u_t(s)\|_{W^r_\infty(\Omega)}Ch^2\|A^{-2}\psi\|_{W^3_1(\Omega)} \le Ch^{r+1}\|u_t(s)\|_{W^r_\infty(\Omega)}.$$

Hence

(3.9)
$$I_2 \le Ch^{r+1}\log(n+1)\sup_{s\le t_n}\|u_t(s)\|_{W^r_\infty(\Omega)}.$$

For the first term on the right of (3.8), we proceed similarly using also (2.6):

(3.10)
$$\begin{aligned}I_1 &\le k\big\|Q_{kh}P_h\rho_t(t_{n-1})\big\|_{L_\infty(\Omega)} = k\big\|A_hQ_{kh}R_hA^{-1}\rho_t(t_{n-1})\big\|_{L_\infty(\Omega)}\\ &\le C\sup_{t_{n-1}\le s\le t_n}\big\|A^{-1}\rho_t(s)\big\|_{L_\infty(\Omega)} \le Ch^{r+1}\sup_{t_{n-1}\le s\le t_n}\big\|u_t(s)\big\|_{W^r_\infty(\Omega)}.\end{aligned}$$

Hence from (3.8), (3.9), and (3.10),

$$\|Z^n\|_{L_\infty(\Omega)} \le Ch^{r+1}\log(n+1)\sup_{s\le t_n}\|u_t(s)\|_{W^r_\infty(\Omega)}$$

and the proof of the proposition is complete. □

4. Proof of the Main Results

Using our preparation from the previous section, we can now prove Theorem 2.1. The exact meaning of $k^q << h^{r+1-\varepsilon}$ is the assumption that the terms $\mathbf{P}_1$ and $\mathbf{P}_2$ can be dropped from the estimates in Propositions 3.5 and 3.6.

Proof. Recall the notation

$$[u;t_n] := \|u(t_n)\|_{W^{r+1}_\infty(\Omega)} + \sup_{s\le t_n} \|u_t(s)\|_{W^r_\infty(\Omega)}.$$

We assume that the final time t_n is bounded, and the time discretization is sufficiently fine, i.e. $k^q \ll h^{r+1-\varepsilon}$. If we thus neglect in Propositions 3.5 and 3.6 all terms involving the time discretization step-size k, we have

$$\big|\nabla e^n(x)\big| \le \widehat{C} h^{r-1}\left(\sum_{|\alpha|=r} |D^\alpha_x u(x,t_n)| + h^{1-\varepsilon}[u;t_n]\right). \tag{4.1}$$

and

$$\big|e^n(x)\big| \le \widehat{C} h^{r}\left(\sum_{|\alpha|=r} |D^\alpha_x u(x,t_n)| + h^{1-\varepsilon}[u;t_n]\right). \tag{4.2}$$

Next we will show (2.10). By the triangle inequality, (2.7), and (2.8), we have

$$\begin{aligned}\big|\nabla u(t_n)(x) - G_H U^n(x)\big| &\le \big|\nabla u(t_n)(x) - G_H u(t_n)(x)\big| + \big|G_H e^n(x)\big| \\ &\le C_G H^r \|u(t_n)\|_{W^{r+1}_\infty(\Omega)} + C_G H^{-1}\|e^n\|_{L_\infty(d_H)} \\ &\le C_G H^r [u;t_n] + C_G H^{-1}\|e^n\|_{L_\infty(d_H)}.\end{aligned} \tag{4.3}$$

Let $\overline{x} \in d_H$ be the point such that $\|e^n\|_{L_\infty(d_H)} = |e^n(\overline{x})|$. From (4.2) and the Mean Value Theorem,

$$\begin{aligned}\big|e^n(\overline{x})\big| &\le \widehat{C} h^r\left(\sum_{|\alpha|=r} |D^\alpha_x u(\overline{x},t_n)| + h^{1-\varepsilon}[u;t_n]\right) \\ &\le \widehat{C} h^r\Big(|u(t_n)|_{W^r_\infty(\tau)} + (H + h^{1-\varepsilon})[u;t_n]\Big).\end{aligned} \tag{4.4}$$

Combining the last two equations and using that $H \ge h$, we have

$$\begin{aligned}\big|\nabla u(t_n)(x) - G_H U^n(x)\big| &\le C_G H^r[u;t_n] + \widehat{C} C_G \frac{h^r}{H}\left(|u(t_n)|_{W^r_\infty(\tau)} + (H + h^{1-\varepsilon})[u;t_n]\right) \\ &\le \widehat{C}_2 h^{r-1}\left(\frac{h}{H}|u(t_n)|_{W^r_\infty(\tau)} + (\frac{H^r}{h^{r-1}} + h + \frac{h^{2-\varepsilon}}{H})[u;t_n]\right) \\ &\le \widehat{C}_2 h^{r-1}\left(\frac{h}{H}|u(t_n)|_{W^r_\infty(\tau)} + (\frac{H^r}{h^r}h^\varepsilon + \frac{h}{H})h^{1-\varepsilon}[u;t_n]\right) \\ &\le \widehat{C}_2 h^{r-1}\widehat{m}\left(|u(t_n)|_{W^r_\infty(\tau)} + h^{1-\varepsilon}[u;t_n]\right),\end{aligned} \tag{4.5}$$

where $\widehat{m} = (H/h)^r h^\varepsilon + h/H$.

Using this we shall now consider the two alternatives involved in Theorem 2.1.

Alternative 1. In this case we have

$$|u(t_n)|_{W^r_\infty(\tau)} \ge h^{1-\varepsilon}[u;t_n],$$

so that by (4.5),

$$\big|\nabla u(t_n)(x) - G_H U^n(x)\big| \le 2\widehat{C}_2 h^{r-1} \widehat{m} |u(t_n)|_{W_\infty^r(\tau)}. \tag{4.6}$$

By Lemma 3.2 and taking h such that $h^\varepsilon < \frac{1}{2}$,

$$h^{r-1}\left(|u(t_n)|_{W_\infty^r(\tau)} - \frac{h^{1-\varepsilon}}{2}[u;t_n] \right) \le \widehat{C}_1 \|\nabla e^n\|_{L_\infty(\tau)}.$$

Combining the above estimate with (4.6), we have

$$\big|\nabla u(t_n)(x) - G_H U^n(x)\big| \le 4\widehat{C}_1\widehat{C}_2\widehat{m}\|\nabla e^n\|_{L_\infty(\tau)}.$$

This shows (2.10) with $m = 4\widehat{C}_1\widehat{C}_2\widehat{m}$.

The estimate (2.11) is a simple consequence of the triangle inequality,

$$\mathcal{E}(\tau) \le \big|\nabla u(t_n)(x) - G_H U^n(x)\big| + |\nabla e^n(x)| \le (1+m)\|\nabla e^n\|_{L_\infty(\tau)}.$$

Also, if $m < 1$,

$$\|\nabla e^n\|_{L_\infty(\tau)} \le \big|\nabla u(t_n)(x) - G_H U^n(x)\big| + \mathcal{E}(\tau) \le m\|\nabla e^n\|_{L_\infty(\tau)} + \mathcal{E}(\tau),$$

which gives (2.12).

Alternative 2. In this case, from (4.1) we have

$$\|\nabla e^n\|_{L_\infty(\tau)} \le 2\widehat{C}h^{r-\varepsilon}[u;t_n]$$

and from (4.5),

$$\|\nabla u(t_n) - G_H U^n\|_{L_\infty(\tau)} \le 2\widehat{C}_2\widehat{m}h^{r-\varepsilon}[u;t_n].$$

By the triangle inequality then,

$$\mathcal{E}(\tau) \le (m + C_2)h^{r-\varepsilon}[u;t_n].$$

This shows (2.14), (2.15), and (2.16) with $m = 2\widehat{C}_2\widehat{m}$ and $C_2 = 2\widehat{C}$. Thus the proof of Theorem 2.1 is complete. □

5. Numerical illustrations

To illustrate the theoretical results above, we consider a simple one dimensional problem

$$\begin{aligned} u_t(x,t) - (1+x)u_{xx}(x,t) &= f(x,t), \text{ for } 0 \le x \le 1,\ t \ge 0, \\ u_x(0,t) &= 0,\ t \ge 0, \\ u_x(1,t) &= 0,\ t \ge 0, \\ u(x,0) &= 0,\ 0 \le x \le 1. \end{aligned}$$

In all examples the finite elements are quadratic Lagrange elements. The meshes are quasi-uniform but not uniform. In fact, they were constructed starting with a uniform mesh of size h and then perturbing points randomly. Also, in all examples we report at time $t = 1$ on the efficiency index, $\mathcal{E}(\tau)/\|\nabla e\|_{L_\infty(\tau)}$, where τ is the element containing $x = 1/2$. "Patch size" stands for the number of neighboring elements on each side of τ in the patch d_H, i.e. patch size 1 means there are three elements in the patch d_H, patch size 2 means there are five elements and so on.

Below we compare the performances of three different methods.

Method 1: (Local interpolation). Put

$$G_H u_h = \frac{\partial}{\partial x} I_H^3 u_h,$$

where $I_H^3 : \mathcal{C}(d_H) \to \mathbb{P}^3(d_H)$ is the local Lagrange interpolant onto cubic polynomials on d_H. To avoid phenomena of superconvergence, the interpolation points are chosen away from the known first derivative superconvergent points.

Method 2: (Local L_2 projection). Put

$$G_H u_h = \frac{\partial}{\partial x} P_H^3 u_h,$$

where $P_H^3 : L_2(d_H) \to \mathbb{P}^3(d_H)$ is the local L_2 projection onto cubic polynomials on d_H, i.e.

$$(P_H^3 v, \chi)_{d_H} = (v, \chi)_{d_H}, \text{ for } \chi \in \Pi^3(d_H).$$

Method 3: (Local L_2 over-projection). Put

$$G_H u_h = \frac{\partial}{\partial x} P_H^4 u_h,$$

where $P_H^4 : L_2(d_H) \to \mathbb{P}^4(d_H)$ is the local L_2 projection onto quartic polynomials on d_H.

$$(P_H^4 v, \chi)_{d_H} = (v, \chi)_{d_H}, \text{ for } \chi \in \Pi^4(d_H).$$

After this set-up, we now give our numerical illustrations.

Example 1. In the first example we take f such that the exact solution is

$$u(x,t) = t \cos(\pi x).$$

For the time discretization, we use the Backward Euler method with uniform time steps. Since the Backward Euler method is exact in this case, the error only depends on the space discretization.

Table 1. (40 elements)

patch size	Method 1	Method 2	Method 3
1	1.2101	1.0345	1.0320
2	0.9390	1.0306	1.0181
3	0.9451	1.0525	1.0348
4	0.9531	1.1135	1.0838
5	1.0404	1.2292	1.1878

Table 2. (1280 elements)

patch size	Method 1	Method 2	Method 3
1	1.1641	1.0269	1.0313
2	0.9573	1.0149	1.0179
3	0.9551	1.0078	1.0076
4	0.9540	1.0054	1.0053
5	0.9589	1.0037	1.0038

As we see from the tables above, the efficiency indices are close to 1 as predicted by the theory.

Example 2. This time we take the exact solution to be

$$u(x,t) = t^3 \cos(\pi x).$$

For the time discretization we use the Backward Euler method with uniform time steps.

Table 3. (40 space intervals and 160 time steps)

patch size	Method 1	Method 2	Method 3
1	0.4145	0.3684	0.3673
2	0.3462	0.3684	0.3654
3	0.3478	0.3783	0.3712
4	0.3509	0.3978	0.3872
5	0.3714	0.4366	0.4182

Table 4. (40 space intervals and 1280 time steps)

patch size	Method 1	Method 2	Method 3
1	0.9435	0.8385	0.8361
2	0.7880	0.8385	0.8318
3	0.7915	0.8610	0.8450
4	0.7984	0.9052	0.8813
5	0.8449	0.9934	0.9521

Table 5. (40 space intervals and 10240 time steps)

patch size	Method 1	Method 2	Method 3
1	1.1233	0.9983	0.9955
2	0.9381	0.9983	0.9903
3	0.9423	1.0250	1.0060
4	0.9505	1.0777	1.0493
5	1.0058	1.1826	1.1335

As we see from the tables above, these results are in agreement with the theory. The localization effect become evident only when time step is sufficiently small. Thus in Table 3, $k = 1/160$ and $h^2 \approx 1/1600$, hence the time discretization error dominates and we see that the performance of the error estimator is poor. In Table 4, $k \approx h^2$ and the performance is reasonable. Finally, in Table 5, $k \ll h^2$ and the performance is very good.

Example 3. Same problem as in Example 2 with the only difference that for the time discretization we now use the Crank-Nicolson method with uniform time steps. Although Theorem 2.1 does not cover the Crank-Nicolson method, standard arguments, (cf. Theorem 2.3 in Thomée, Xu, and Zhang [13]), can be adapted to our problem up to dimension 2. Since the time discretization is of second order, we see good results already with 40 time steps.

Table 6. (40 space intervals and 20 time steps)

patch size	Method 1	Method 2	Method 3
1	1.0404	0.8991	0.9070
2	0.8310	0.8983	0.8939
3	0.8336	0.9157	0.9039
4	0.8405	0.9751	0.9527
5	0.8905	1.0806	1.0423

Table 7. (40 space intervals and 40 time steps)

patch size	Method 1	Method 2	Method 3
1	1.1468	0.9910	0.9998
2	0.9159	0.9902	0.9853
3	0.9188	1.0093	0.9963
4	0.9265	1.0749	1.0501
5	0.9816	1.1912	1.1489

For the element $\tau \ni 1/2$ we have

$$\|\cos(\pi x)\|_{L_\infty(\tau)} \approx \frac{\pi h}{2} \quad \text{and} \quad \|\sin(\pi x)\|_{L_\infty(\tau)} = 1.$$

If $u(x,t) = t^m \cos(\pi x)$ then at $t_n = 1$ we have,

$$\begin{aligned}
\|u(t_n)\|_{L_\infty(\tau)} &\approx \frac{\pi h}{2} \\
|u(t_n)|_{W^1_\infty(\tau)} &= \pi \\
|u(t_n)|_{W^2_\infty(\tau)} &\approx \frac{\pi^3 h}{2} \\
|u(t_n)|_{W^3_\infty(\tau)} &= \pi^3 \\
|u(t_n)|_{W^4_\infty(\tau)} &\approx \frac{\pi^5 h}{2}.
\end{aligned}$$

Thus, for the Examples 1-3 we are in Alternative 1 if

$$\pi^3 \geq h^{1-\varepsilon}\left(4\frac{\pi h}{2} + 4\pi + 4\frac{\pi^3 h}{2} + 4\pi^3 + \frac{\pi^5 h}{2}\right),$$

which is the case for $h = 1/40$ with $\varepsilon = 1/2$, for example.

Example 4. In the last example we take the exact solution

$$u(x,t) = t^3 \cos(2\pi x).$$

We use the Backward Euler method with uniform time steps.

Table 8. (40 space intervals and 10240 time steps)

patch size	Method 1	Method 2	Method 3
1	0.0014	0.0025	0.0007
2	0.0049	0.0071	0.0005
3	0.0097	0.0130	0.0006
4	0.0162	0.0237	0.0007
5	0.0277	0.0320	0.0010

As we see from the table above, the performance of the error estimator is very poor even for a very small time step. This is in agreement with the theory because we are in the case of Alternative 2 since

$$\frac{\partial^3 u}{\partial x^3} = -(2\pi)^3 t^3 \sin(2\pi x)$$

vanishes at $x = 1/2$. A more detailed analysis similar to the above shows that for $\tau \ni 1/2$ we have

$$8\pi^4 h \leq h^{1-\varepsilon}\left(4 + 8\pi^2 h + 16\pi^2 + 32\pi^4 h + 16\pi^4\right)$$

for any ε.

Acknowledgments: This research was partly supported by an NSF VIGRE Fellowship and U.S. National Science Foundation, under grant DMS 0310539.

References

1. R.E. Bank and J. Xu, *Asymptotically exact a posteriori error estimators, part 1: Grids with superconvergence*, SIAM J. Numer. Anal., **41** (2003), 2294–2312.
2. R.E. Bank and J. Xu, *Asymptotically exact a posteriori error estimators, part 2: General unstructured grids*, SIAM J. Numer. Anal., **41** (2003), 2313–2332.
3. P. Brenner, M. Crouzeix, and V. Thomée, *Single step methods for inhomogeneous linear differential equations in Banach space*, RAIRO Anal. Numer. **16** (1982), 5–26.
4. V. Carey, *A posteriori error estimation via recovered gradients*, Ph.D. thesis, Cornell University, 2005.
5. S.D. Èĭdel'man and S.D. Ivasišen, *Investigation of the Green's matrix for a homogeneous parabolic boundary value problem*, Trans. Moscow Math. Soc. **23** (1970), 179–242.
6. A. Hansbo, *Nonsmooth data error estimates for damped single step methods for parabolic equations in Banach space*, Calcolo, **36** (1999), 75–101.
7. W. Hoffmann, A.H. Schatz, L.B. Wahlbin, and G. Wittum, *Asymptotically exact a posteriori estimators for the pointwise gradient error on each element in irregular meshes. Part 1: A smooth problem and globally quasi-uniform meshes*, Math. Comp. **70** (2001), 897–909.
8. O. Lakkis and C. Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp. **75** (2006), 1627–1658.
9. A.H. Schatz, *Pointwise error estimates and asymptotic error expansion inequalities for the finite element method on irregular grids: Part 1*, Math. Comp. **67** (1998), 877–899.
10. A.H. Schatz, V. Thomée, and L.B. Wahlbin, *Stability, analyticity, and almost best approximation in maximum-norm for parabolic finite element equations*, Comm. Pure Appl. Math. **51** (1998), 1349–1385.
11. A.H. Schatz and L.B. Wahlbin, *Asymptotically exact a posteriori estimators for the pointwise gradient error on each element in irregular meshes. Part II: The piecewise linear case.* , Math. Comp. **73** (2004), 517–523.
12. V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer Series in Computational Mathematics No.25, Springer-Verlag, Berlin, 1997.
13. V. Thomée, J.-C. Xu, and N.-Y. Zhang, *Superconvergence of the gradient in piecewise linear finite-element approximation to a parabolic problem*, SIAM J. Numer. Anal., **26** (1989), 553–573.
14. S. Ziukas and N.-E. Wiberg, *Adaptive procedure with superconvergent patch recovery for linear parabolic problems*, Finite element methods, superconvergence, post-processing, and a posteriori estimates, Marcel Dekker, Inc., New York-Basel-Hong Kong, 1998, 303–314.

Department of Mathematics, University of Connecticut, Storrs, CT 06269
E-mail address: leykekhman@math.uconn.edu

Department of Mathematics, Cornell University, Ithaca, NY 14853
E-mail address: wahlbin@math.cornell.edu