GRADIENT RECOVERY IN ADAPTIVE FINITE ELEMENT METHODS FOR PARABOLIC PROBLEMS

OMAR LAKKIS AND TRISTAN PRYER

Abstract. We derive energy-norm a posteriori error bounds, using gradient recovery (ZZ) estimators to control the spatial error, for fully discrete schemes for the linear heat equation. This appears to be the first completely rigorous derivation of ZZ estimators for *fully discrete* schemes for evolution problems, without any restrictive assumption on the timestep size. An essential tool for the analysis is the elliptic reconstruction technique.

Our theoretical results are backed with extensive numerical experimentation aimed at (a) testing the practical sharpness and asymptotic behaviour of the error estimator against the error, and (b) deriving an adaptive method based on our estimators.

An extra novelty provided is an implementation of a coarsening error "preindicator", with a complete implementation guide in ALBERTA in the appendix.

1. Introduction

Gradient recovery a posteriori error estimators have been widely used since their dissemination in the engineering and scientific computation community by [ZZ87], for which we will often refer to them shortly as *ZZ estimators*. Since their introduction they have constituted the most serious rival to *residual estimators* introduced earlier on in [BR78]. The key to ZZ estimator's success is their implementation's simplicity, mild dependence upon the problem's data, and striking superconvergence and asymptotic exactness properties. On the other hand, residual estimators, which are the main competitor to ZZ estimators, are a bit more involved in implementation and cost more to compute, but they are easier to handle from the mathematical analysis view-point in deriving rigorous upper and lower bounds. This situation has led to most of the theoretical results for evolution equations being obtained in the last two decades via residual estimators; we refer to [LM06] for a review. Meanwhile rigorous mathematical work on recovery estimators has progressed, especially in the last decade, but mostly for stationary elliptic equations, e.g., [AO00, BX03a, BX03b, CB02, FV06, LZ99, Pic03, XZ04]. In contrast, very little progress was made on evolution problems, where an exception is [LW06], where the main analytic difficulty comes from the singularly perturbed nature of the elliptic problems arising from time-stepping procedures.

The aim of our work is to bridge the gap between the practical use of ZZ estimators in adaptivity for evolution equations, studied by [ZW98, Pic03], and the rather mature error control theory via recovery for stationary equations. We focus on the model problem provided by the linear heat equation. [LW06], who are to our knowledge the only researchers to have explored this issue in depth, while obtaining satisfactory error bounds for spatially discrete schemes, must assume unrealistically small time-steps for the fully discrete case. In this paper we push one step forward by thoroughly analysing the fully discrete backward Euler schemes.

Date: October 30, 2018.

T.P. was supported by his EPSRC D.Phil. scholarship grant. O.L. was partially supported by a Nuffield Young Researcher's Grant.

Namely, we provide reliable error bounds. The efficiency and asymptotic exactness of the bounds is dealt with computationally.

Our main analytical tool to tackle the fully discrete scheme's difficulties is the *elliptic reconstruction* in the fully discrete context, studied in [LM06], which provides a way to take advantage of elliptic a posteriori error estimates based on gradient recovery following the exposition of [AO00].

The elliptic reconstruction technique, introduced under this name by [MN03], involves the *error's splitting* into two parts, a *parabolic error* and an *elliptic error*, through the use of the *elliptic reconstruction* of the discrete solution, defined in (3.1). This allows to utilise existing elliptic a posteriori estimators for the elliptic part and standard parabolic energy estimates to control the second part. Despite this technique being initially introduced to derive sharp bounds for lower order spatial error norms, such as $\mathrm{L}_2(\Omega)$ [MN03, LM06, LM07] and $\mathrm{L}_\infty(\Omega)$ [DLM09], we apply it here as an analysis tool in an energy-norm framework, where a direct approach may lead to a highly complicated analysis for the fully discrete scheme. In fact, the single most interesting feature of the elliptic reconstruction, is that *the parabolic error's energy norm term is of higher order (with respect to the spatial mesh-size parameter) than the elliptic error*, as seen in [LM06]. In this paper we show, rigorously, that the *full energy error can be accounted for only by the elliptic error*, as long as data and time-step are resolved well enough (cf. Lemma 3.3). This crucial observation is also used to obtain residual a posteriori estimates for nonconforming methods in [GL08]. Note that, it is part of the adaptive methods practitioner's folklore to employ heuristic versions of this argument. By way of example, we quote "the [full parabolic] discretisation in energy norm can be bounded by the [elliptic error] estimator" from [ZW98].

Although we treat the case of the Laplace operator, for simplicity, in this paper, our results can be extended to cover more general elliptic operators, even time-dependent ones, by using appropriate elliptic gradient recovery techniques, described by [FV06], and a more careful time-step analysis, as in [GL08].

The paper is organised as follows. In §2 we introduce the model problem, and its discretisations via conforming finite elements in space and backward Euler in time and we review the known results, about recovery estimators for elliptic problems, that will be used in the sequel. In §3 we describe the elliptic reconstruction technique and illustrate its use for the spatially semidiscrete problem. This paves the way to tackle the fully discrete problem in §4, where our main results are stated. In §5, using numerical tests, we study the practical behaviour of the estimators and in §6 we explore the adaptive schemes based on our estimators.

As we have used the finite element toolbox ALBERTA, written and documented by [SS05], for the tests, we have taken the opportunity to implement a *coarsening preindicator*, previously unavailable and (for space's sake) fully described in the Appendix A. This estimator predicts the "information loss" error that will occur under coarsening of the mesh at each timestep of the adaptive method and is crucial in an adaptive code to control information loss during coarsening.

2. Set up

2.1. **The model problem.** Let $\Omega \subset \mathbb{R}^d$ be a bounded polyhedral domain and consider the (generalised or weak) Laplace operator denoted by

$$\begin{array}{rccl} \mathscr{A}: & \mathrm{H}_0^1(\Omega) & \to & \mathrm{H}^{-1}(\Omega) \\ & u & \mapsto & \mathscr{A}u := -\Delta u := -\operatorname{div}\nabla u = -\sum_{i=1}^d \partial_i^2 u. \end{array} \tag{2.1}$$

We denote by $\mathrm{L}_2(\Omega)$ the space of square summable functions on Ω, with inner product and norm respectively defined by

$$\langle f, g\rangle := \int_\Omega f(\mathbf{x})g(\mathbf{x})\,\mathrm{d}x \text{ and } \|f\| := \langle f, f\rangle^{1/2}. \tag{2.2}$$

We will use the standard [Cia78, Eva98,] Sobolev spaces

$$\mathrm{H}^1(\Omega) := \{\phi \in \mathrm{L}_2(\Omega) :\ \nabla\phi \in \mathrm{L}_2(\Omega)\}, \tag{2.3}$$

$$\mathrm{H}^1_0(\Omega) := \left\{\phi \in \mathrm{H}^1(\Omega) :\ \phi|_{\partial\Omega} = 0\right\} \tag{2.4}$$

$$\text{and } \mathrm{H}^{-1}(\Omega) := \mathrm{dual}\big(\mathrm{H}^1_0(\Omega)\big). \tag{2.5}$$

Let $T > 0$, the model parabolic problem consists in finding a function $u \in \mathrm{L}_2(0,T;\mathrm{H}^1_0(\Omega))$ and $\partial_t u \in \mathrm{L}_2(0,T;\mathrm{H}^{-1}(\Omega))$ such that

$$\begin{aligned} \partial_t u(t) + \mathscr{A}u(t) &= f(\cdot,t), \text{ for all } t \in (0,T], \\ u(\boldsymbol{x},0) &= u_0(\boldsymbol{x}), \text{ for } \boldsymbol{x} \in \overline{\Omega}, \\ u(\boldsymbol{x},t) &= 0, \text{ for } (\boldsymbol{x},t) \in \partial\Omega \times (0,T]. \end{aligned} \tag{2.6}$$

We consider the case where $u_0 \in \mathrm{L}_2(\Omega)$ and $f \in \mathrm{L}_2(0,T;\mathrm{L}_2(\Omega))$ for which the problem (2.6) admits a unique solution [Eva98,].

Problem (2.6) is understood in the following weak form

$$\begin{aligned} \langle \partial_t u(t), \phi\rangle + a(u(t),\phi) &= \langle f(t),\phi\rangle \quad \forall\, \phi \in \mathrm{H}^1_0(\Omega),\, t \in (0,T] \\ u(\cdot,0) &= u_0(\cdot), \end{aligned} \tag{2.7}$$

where $\langle\cdot,\cdot\rangle$ is defined in (2.2) and $a(\phi,\psi) := \langle\nabla\phi,\nabla\psi\rangle$. The form $a(\cdot,\cdot)$ is clearly bounded and coercive, i.e.,

$$a(\phi,\phi) \geq \alpha\,\|\phi\|_1^2 \quad \forall\, \phi \in \mathrm{H}^1_0(\Omega), \tag{2.8}$$

where $\alpha = (1 + C_{\mathrm{P}}^2)^{-1}$ and C_{P} is the Poincaré constant. The bilinear form defines an inner product on $\mathrm{H}^1_0(\Omega)$ and hence we can denote the energy norm $\|\cdot\|_a^2 := a(\cdot,\cdot)$.

These observations justify our use of $\|\cdot\|_a$ (instead of $\|\cdot\|_{\mathrm{H}^1(\Omega)}$) as the norm of $\mathrm{H}^1_0(\Omega)$ to be with the implied dual norm on $\mathrm{H}^{-1}(\Omega)$ in (2.5).

2.2. **Spatial discretisation.** Let $\mathscr{T}$ be a conforming, not necessarily quasiuniform, triangulation of Ω, i.e., (1) $K \in \mathscr{T}$ means K is an open simplex (triangle for $d = 2$ or tetrahedron for $d = 3$), (2) for any $K, J \in \mathscr{T}$ we have that $\overline{K} \cap \overline{J}$ is a full subsimplex (i.e., it is either $\emptyset$, a vertex, an edge, a face, or the whole of $\overline{K}$ and $\overline{J}$) of both $\overline{K}$ and $\overline{J}$ and (3) $\bigcup_{K\in\mathscr{T}} \overline{K} = \overline{\Omega}$. The shape regularity of $\mathscr{T}$ is defined as the number

$$\mu(\mathscr{T}) := \inf_{K\in\mathscr{T}} \frac{\rho_K}{h_K}, \tag{2.9}$$

where ρ_K is the radius of the largest ball contained inside K and h_K is the longest side of K. An indexed family of triangulations $\{\mathscr{T}^n\}_n$ is called *shape regular* if

$$\mu := \inf_n \mu(\mathscr{T}^n) > 0. \tag{2.10}$$

We will use henceforth the usual convention where $h : \Omega \to \mathbb{R}$ denotes the *mesh-size function* of $\mathscr{T}$, i.e.,

$$h(x) := h_{\mathscr{T}}(x) := \max_{\overline{K}\ni x} h_K, \text{ and } h_n := h_{\mathscr{T}^n}. \tag{2.11}$$

With a triangulation $\mathscr{T}$ as described above, and an integer $p \geq 1$ considered fixed in the sequel, we may consider the *finite element space*

$$\mathbb{V} := \left\{\Phi \in \mathrm{H}^1_0(\Omega) :\ \Phi|_K \in \mathbb{P}^p \quad \forall\, K \in \mathscr{T}\right\}; \tag{2.12}$$

and $\mathbb{P}^k$ denotes the linear space of polynomials in d variables of degree no higher than a positive integer k. The *spatially discrete finite element solution* in $\mathbb{V}$, is the function $U : [0,T] \to \mathbb{V}$ such that

$$\begin{aligned} \langle \partial_t U, \Phi \rangle + a(U,\Phi) &= \langle f, \Phi \rangle \quad \forall\, \Phi \in \mathbb{V}, \\ U(\boldsymbol{x},0) = U^0 &:= \Pi^{\mathbb{V}} u_0(\boldsymbol{x}) \quad \forall\, \boldsymbol{x} \in \Omega, \end{aligned} \tag{2.13}$$

where $\Pi^{\mathbb{V}} : \mathrm{L}_2(\Omega) \to \mathbb{V}$ is a suitable projector (or an interpolator if the data u_0 is in a higher regularity subspace of $\mathrm{L}_2(\Omega)$, e.g., $\mathscr{T}$-wise continuous) and $\langle \cdot, \cdot \rangle$ is the same as in (2.7) and (2.2).

We will often write the scheme (2.13) in its *pointwise form*

$$\partial_t U + AU = P_0 f \text{ and } U(0) = U^0, \tag{2.14}$$

where the finite dimensional space operator $A : \mathbb{V} \to \mathbb{V}$ is the *discrete Laplacian* defined, through the Riesz representation in $\mathbb{V}$, by

$$\langle AV, \Phi \rangle = a(V,\Phi) \quad \forall\, \Phi \in \mathbb{V}, \tag{2.15}$$

and $P_0 : \mathrm{L}_2(\Omega) \to \mathbb{V}$ is the $\mathrm{L}_2(\Omega)$-projection operator such that, for each $v \in \mathrm{L}_2(\Omega)$,

$$\langle P_0 v, \Phi \rangle = \langle v, \Phi \rangle \quad \forall\, \Phi \in \mathbb{V}. \tag{2.16}$$

The pointwise form is convenient as it allows for a more compact notation.

2.3. **Fully discrete scheme.** Subdivide the time interval $[0,T]$ into a partition of N consecutive adjacent subintervals whose endpoints are denoted $t_0 = 0 < t_1 < \ldots < t_N = T$. The n-th timestep is defined as $\tau_n := t_n - t_{n-1}$. We will consistently use the shorthand $F^n(\cdot) := F(\cdot, t_n)$ for a generic time function F.

The backward Euler method consists in finding a sequence of functions, $U^n \in \mathbb{V}^n$, such that for each $n = 1, \ldots, N$ we have:

$$\begin{aligned} \frac{1}{\tau_n} \left\langle U^n - \Lambda^n U^{n-1}, \Phi \right\rangle + a(U^n, \Phi) &= \langle f^n, \Phi \rangle \quad \forall\, \Phi \in \mathbb{V}^n, \\ U^0 &= \Pi^0 u_0, \end{aligned} \tag{2.17}$$

where $\Lambda^{\mathbb{V}} : \mathrm{C}^0(\Omega) \to \mathbb{V}$ denotes the Lagrange interpolation operator, $\Lambda^n := \Lambda^{\mathbb{V}^n}$, and Π^0 is defined as $\Pi^{\mathbb{V}}$.

Note our nonrestrictive use of the Lagrange interpolator as a "data-transfer" operator from a finite element space to the next. We do this to reflect exactly what we do in practical computations (where interpolation is faster than averaging). All our analysis applies, however to a different data-transfer operator, including the $\mathrm{L}_2(\Omega)$ projector, if necessary.

As with the semidiscrete scheme the fully discrete scheme can be written in a pointwise form as follows:

$$\frac{U^n - \Lambda^n U^{n-1}}{\tau_n} + A^{\mathbb{V}^n} U^n = P_0^n f^n \text{ and } U^0 = \Pi^0 u_0, \tag{2.18}$$

where $A^n = A^{\mathbb{V}^n}$ and $P_0^n = P_0^{\mathbb{V}^n}$ (cf. (2.15)).

2.4. **Recovery a posteriori estimators.** The stationary elliptic problem corresponding to a steady state of the evolution equation (2.6) is,

$$\text{given } g \in \mathrm{L}_2(\Omega), \text{ find } w \in \mathrm{H}_0^1(\Omega) \text{ such that } \mathscr{A} w = g, \tag{2.19}$$

where the operator is understood in a generalised sense and the solution is a weak one. The finite element discretisation of the elliptic problem (2.19) consists in

$$\text{finding } W \in \mathbb{V} \text{ such that } a(W,\Phi) = \langle g, \Phi \rangle \quad \forall\, \Phi \in \mathbb{V}. \tag{2.20}$$

We shall henceforth denote by w and W the solutions of (2.19) and (2.20).

From the literature on elliptic a posteriori estimation [AO00, BR78, Cia78, Ver96, BX03a, ZZ87,] there is a variety of ways to compute upper and lower bounds for the error in some functional space $\mathcal{X}$ (e.g., $\mathrm{H}_0^1(\Omega)$, $\mathrm{L}_2(\Omega)$ and $\mathrm{L}_\infty(\Omega)$). For instance, a generic upper a posteriori $\mathcal{X}$-norm error bound takes the form

$$\|w - W\|_{\mathcal{X}} \leq \mathscr{E}[W, g, \mathcal{X}, \mathbb{V}]\,, \tag{2.21}$$

where $\mathscr{E}$ is an appropriate *estimator functional*.

One way of providing an estimator functional consists, for example, in starting by applying a *gradient postprocessing operator (postprocessor)*, say G, to the approximate solution W. And then proving that $\|GW - \nabla W\|$ is equivalent to the error $\|\nabla w - \nabla W\|$. *Gradient recovery operators* form a subclass of gradient postprocessors.

Recovery operators can be built in a variety of ways such as local weighted averaging (where the gradient is sampled from neighbouring elements) [Pic03,], discrete $\mathrm{L}_2(\Omega)$-projection (using least squares fitting) [ZZ87,] or global $\mathrm{L}_2(\Omega)$-projection (where a full discrete problem is solved) [BX03a,]. In our numerical results we use local weighted averaging, defined explicitly in (5.4), but our theoretical results can be applied with any choice of recovery operator that provide upper and lower bounds for the elliptic problem. The fundamental idea behind these approaches is to build an approximation GW of ∇w which is more regular than the piecewise discontinuous gradient ∇W; the extra regularity is aimed at obtaining a higher approximation order.

2.5. **Definition** (gradient recovery operator, from [AO00])**.** A *gradient recovery (ZZ) operator* on $\mathbb{V}$ is a linear operator $G : \mathbb{V} \to \mathbb{V}^d$ which enjoys the following properties:

Consistency: we have, with $\mathit{\Lambda}^{\mathbb{V}} : \mathrm{C}^0(\Omega) \to \mathbb{V}$ denoting the Lagrange interpolator,

$$G(\mathit{\Lambda}^{\mathbb{V}} v)\big|_K = \nabla v|_K \quad \forall\, v \in \mathbb{P}^{p+1},\, K \in \mathscr{T}\,. \tag{2.22}$$

Local bound: there exists a $C_{\mathrm{ZZ}} > 0$ such that

$$\|GV\|_{\mathrm{L}_\infty(K)} \leq C_{\mathrm{ZZ}}\, \|\nabla V\|_{\mathrm{L}_\infty(\hat{K})} \quad \forall\, V \in \mathbb{V},\, K \in \mathscr{T}\,, \tag{2.23}$$

where $\hat{K}$ is the *patch* generated by K (the union of all $L \in \mathscr{T}$ such that $\overline{L} \cap \overline{K} \neq \emptyset$).

For simplicity, we assume that the operator is in a mesh-local relation with ∇V noting, nonetheless, that global methods such as the global $\mathrm{L}_2(\Omega)$-projection proposed by [BX03a, BX03b] exist and can be included in our discussion.

Under certain regularity assumptions recovery estimators are shown to be asymptotically exact. For instance, [Zlá77] shows that if $w \in \mathrm{H}^{s+1}(\Omega)$, with reference to (2.19) and (2.20), its approximation $W \in \mathbb{V}$ satisfies the following *superconvergence property*:

$$\big\|\nabla(W - \mathit{\Lambda}^{\mathbb{V}} w)\big\| = \mathrm{O}(h^{p+\zeta}) \text{ for some } \zeta \in (0, 1]\,. \tag{2.24}$$

A review of superconvergence results is given by [KN87]. If (2.24) is satisfied then the recovered gradient also satisfies the following superconvergence property [AO00,]:

$$\|\nabla w - GW\| = \mathrm{O}(h^{p+\zeta}) \text{ for some } \zeta \in (0, 1]\,. \tag{2.25}$$

The reach of Zlámal's result is appreciated by stating the following consequence.

2.6. **Lemma** (gradient recovery a posteriori estimate from [AO00])**.** *Let $\mathbb{V}$ be the finite element space defined in (2.12) and $G : \mathbb{V} \to \mathbb{V}^d$ a gradient recovery operator*

according to §2.5. If w, W are the solutions of (2.19) and (2.20), respectively, and (2.25) *holds then the recovery operator is asymptotically exact, in the sense that*

$$\lim_{h_{\mathscr{T}}\to 0} \frac{\|\nabla W - GW\|}{\|\nabla(W-w)\|} = 1. \tag{2.26}$$

Thus, there exist $\delta_0 \geq 0$, such that $\delta_0(h) \to 0$ as $h \to 0$ and

$$(1-\delta_0)\,\|\nabla W - GW\| \leq \|\nabla[W-w]\| \leq (1+\delta_0)\,\|\nabla W - GW\| \tag{2.27}$$

for all partitions $\mathscr{T}$ of Ω satisfying $h_{\mathscr{T}} < h_0$.

2.7. **Remark** (recovery in absence of regularity)**.** Lacking Zlámal's regularity assumption, recovery-based estimators are empirically observed to be efficient, reliable estimators, even on meshes with low shape-regularity [Car04,].

For more details about recovery-based estimators we refer to the available literature [BX03a, BX03b, XZ04, LZ99, AO00, FV06,].

2.8. **Definition** (gradient recovery a posteriori estimator functional)**.** Lemma 2.6 then justifies the use of the *recovery estimator* in the $\mathrm{H}^1_0(\Omega)$-norm (and by equivalence the energy norm) by defining, for the rest of this paper, the *gradient recovery a posteriori estimator functional*

$$\mathscr{E}[V] := \mathscr{E}\left[V, \mathrm{H}^1_0(\Omega), \mathbb{V}\right] := \|GV - \nabla V\|\,, \text{ for } V \in \mathbb{V}, \tag{2.28}$$

where G is a gradient recovery operator as defined in §2.5.

2.9. **Assumption** (elliptic a posteriori error estimates)**.** We will consider henceforth the blanket assumption that for a fixed h_0, there are some $c_0 < C_0$, such that for any $\mathbb{V}$ with mesh-size $h < h_0$, for w and W solutions of (2.19) and (2.20), respectively and $\mathscr{E}$ defined in 2.8 the following bounds are true

$$c_0\mathscr{E}[W] \leq \|\nabla[W-w]\| \leq C_0\mathscr{E}[W]\,. \tag{2.29}$$

Optionally, we will assume *asymptotic exactness*, in which case

$$C_0 \leq 1 + B(h_0) \text{ and } c_0 \geq 1 + \beta(h_0), \tag{2.30}$$

for some continuous functions B and β that vanish at 0.

Assumptions (2.29) and (2.30) are true, modulo hierarchic *oscillation terms* of the data function g in (2.20). These assumptions are thus justified, for example, when g is in a finite dimensional space, for example $g \in \mathbb{V}$ as we shall assume in the sequel, by isolating the bulk of the oscillations in data-approximation terms. For a thorough discussion of the oscillation in the context of recovery, we refer to [FV06].

The lower bound is not needed for the theory to be developed herein, as we will prove only upper bounds. Nonetheless, this property is required for the efficiency of the parabolic estimators in practical situations.

3. Semidiscrete scheme

To make the link between the parabolic problem and the elliptic recovered gradient estimates we utilise the elliptic reconstruction technique [MN03, LM06,]. To make the discussion more accessible, we first do this for the spatially (semi)discrete scheme. We divide the error into two parts—one called elliptic error the other parabolic error—via the *elliptic reconstruction of the discrete solution*. Because the elliptic error can be directly bounded under the blanket Assumption 2.9, it is enough to show that the full error can be bounded in terms of the elliptic error only. This result is in accordance with the fact that the parabolic error on uniform meshes is of higher h-order in the energy norm with respect to the elliptic (and thus the full) error, as observed by [LM06]. The main result of this section is summarised in Theorem 3.6.

3.1. **Definition** (elliptic reconstruction)**.** The *elliptic reconstruction operator* is defined as $\mathscr{R} : \mathbb{V} \to \mathrm{H}^1_0(\Omega)$ such that

$$\mathscr{A}[\mathscr{R}V] = AV, \tag{3.1}$$

where A is the discrete elliptic operator defined in (2.15). In weak form, equation (3.1) reads

$$a(\mathscr{R}V, \Phi) = \langle AV, \Phi\rangle \quad \forall\, \Phi \in \mathrm{H}^1_0(\Omega), \tag{3.2}$$

and it is well defined in virtue of the elliptic problem's well Sydney's. We will refer to the function $\mathscr{R}V$ as the *elliptic reconstruction* of V, while the elliptic reconstruction operator $\mathscr{R}$ will be called the *reconstruction operator* (or just the *reconstructor*) from $\mathbb{V}$.

If $U(t)$ denotes the solution of (2.14) at time t, we shall indicate by $\omega(t)$ its reconstruction $\mathscr{R}U(t)$.

Thus, posing $g(t) := AU(t)$, we then see $U(t)$ is the finite element solution corresponding to the elliptic problem of finding $\omega(t) \in \mathrm{H}^1_0(\Omega)$ such that $\mathscr{A}\omega(t) = g$.

3.2. **The error and its splitting.** For the whole of this section we shall consider u to be the solution of (2.6), understood in the weak sense, and U its semidiscrete approximation given by (2.14). The corresponding *semidiscrete error* is defined by

$$e(t) := U(t) - u(t), \tag{3.3}$$

and can be split, using the elliptic reconstruction $\omega = \mathscr{R}U$ defined in §3.1, as follows:

$$e(t) = (\omega(t) - u(t)) - (\omega(t) - U(t)) =: \rho(t) - \epsilon(t). \tag{3.4}$$

We shall refer to ϵ and ρ here defined as the *elliptic (reconstruction) error* and the *parabolic error* respectively. Using this notation we have the estimate

$$\|\nabla[U - u](t)\| \le \|\nabla\rho(t)\| + \|\nabla\epsilon(t)\|, \tag{3.5}$$

where, following the remarks made in Definition 3.1 and Assumption 2.9, the elliptic error can be bounded by the computable elliptic a posteriori estimator functional $\mathscr{E}$:

$$\|\epsilon(t)\|_a = \|\nabla\epsilon(t)\| \le C_0\mathscr{E}[U(t)]. \tag{3.6}$$

It is therefore sufficient to bound the error's energy norm using the elliptic error's energy norm.

3.3. **Lemma** (elliptic energy bound for parabolic semidiscrete error)**.** *If e, ϵ are defined as in §3.2 then, for each $t \in [0, T]$, we have*

$$\|e(t)\|^2 + \int_0^t \|e(s)\|_a^2 \,\mathrm{d}s \le \|e(0)\|^2 + \int_0^t \|\epsilon(s)\|_a^2 + 2\,\langle P_0 f(s) - f(s), e(s)\rangle \,\mathrm{d}s. \tag{3.7}$$

Proof From the the exact problem (2.6), the semidiscrete scheme (2.14), and the splitting (3.4)

$$\partial_t e + \mathscr{A}\rho = \partial_t[U - u] + \mathscr{A}[\omega - u] = \partial_t U + AU - \partial_t u - \mathscr{A}u = P_0 f - f. \tag{3.8}$$

Testing with e we obtain

$$\langle \partial_t e, e\rangle + a(\rho, e) = \langle P_0 f - f, e\rangle \tag{3.9}$$

and thus

$$\frac{1}{2}\,\mathrm{d}_t\|e\|^2 + \|e\|_a^2 = \langle P_0 f - f, e\rangle - a(\epsilon, e). \tag{3.10}$$

Integration from 0 to t yields

$$\|e(t)\|^2 + 2\int_0^t \|e\|_a^2 = \|e(0)\|^2 + 2\int_0^t \langle P_0 f - f, e\rangle - 2\int_0^t a(\epsilon, e) \quad \forall t \in [0, T]. \tag{3.11}$$

Hence, by Young's inequality on $a(\epsilon, e)$, we have

$$\|e(t)\|^2 + 2\int_0^t \|e\|_a^2 \leq \|e(0)\|^2 + 2\int_0^t \langle P_0 f - f, e\rangle + \int_0^t \|e\|_a^2 + \int_0^t \|\epsilon\|_a^2, \quad (3.12)$$

whereby the claim is verified. □

3.4. **Remark** (proliferation of $\sqrt{2}$ syndrome)**.** Let $a, b, c \geq 0$ such that $a^2 \leq c^2 + ab$, then, by Young's inequality, it follows that $a^2 \leq 2c^2 + b^2$. Note however that the factor "2" in $2c^2$ is not needed, in that we also have that $a \leq c + b$. If $1 > a \sim c \gg b > 0$, then the first bound provides $a/c \approx \sqrt{2}$ whereas the second bound gives $a/c \approx 1$, which is tighter.

The following result, which generalises $a \leq c + b$, is extremely simple yet useful in avoiding this "proliferation of $\sqrt{2}$ syndrome" from repeated usage of Young's inequality.

3.5. **Proposition** (L_2 simplification rule)**.** *If $\mathbf{a}, \mathbf{b} \in \mathbb{R}^N$, $N \in \mathbb{N}$, $c \in \mathbb{R}$ and $f, g \in \mathrm{L}_2(D)$, for some measurable domain D, are such that*

$$|\mathbf{a}|^2 + \|f\|^2 \leq c^2 + \mathbf{a}^\intercal \mathbf{b} + \int_D fg, \quad (3.13)$$

then

$$\left(|\mathbf{a}|^2 + \|f\|^2\right)^{1/2} \leq |c| + \left(|\mathbf{b}|^2 + \|g\|^2\right)^{1/2}, \quad (3.14)$$

where all the vector norms are Euclidean, and the function norms $\mathrm{L}_2(D)$.

Proof

Denote by $\boldsymbol{\alpha} := (|\mathbf{a}|, \|f\|)$ and $\boldsymbol{\beta} := (|\mathbf{b}|, \|g\|)$.

If $|\boldsymbol{\alpha}| \leq |\boldsymbol{\beta}|$ then (3.14) is trivially satisfied. Otherwise we have $|\boldsymbol{\alpha}| > |\boldsymbol{\beta}|$ whereby (3.13) and the Cauchy–Bunyakovskiy–Schwarz inequality imply that

$$\begin{aligned} |\boldsymbol{\alpha}|^2 &\leq c^2 + |\mathbf{a}|\,|\mathbf{b}| + \|f\|\,\|g\| + |\boldsymbol{\beta}|(|\boldsymbol{\alpha}| - |\boldsymbol{\beta}|) \\ &\leq c^2 + 2\,|\boldsymbol{\alpha}|\,|\boldsymbol{\beta}| - |\boldsymbol{\beta}|^2. \end{aligned} \quad (3.15)$$

Hence $(|\boldsymbol{\alpha}| - |\boldsymbol{\beta}|)^2 \leq c^2$, and thereby

$$|\boldsymbol{\alpha}| \leq |c| + |\boldsymbol{\beta}|, \quad (3.16)$$

as claimed. □

3.6. **Theorem** (a posteriori semidiscrete error estimate)**.** *With u and U as defined by (2.6) and (2.13), respectively, and an estimator functional $\mathscr{E}$ as defined in (2.28), we have*

$$\begin{aligned} &\left(\|U(t) - u(t)\|^2 + \int_0^t \|U - u\|_a^2\right)^{1/2} \\ &\qquad \leq \|U(0) - u(0)\| + C_0\,\|\mathscr{E}[U]\|_{\mathrm{L}_2[0,T]} + 2\,\|P_0 f - f\|_{\mathrm{L}_2(0,T;\mathrm{H}^{-1}(\Omega))}. \end{aligned} \quad (3.17)$$

Proof Using Lemma 3.3 we have

$$\|e(t)\|^2 + \int_0^t \|e\|_a^2 \leq \|e(0)\|^2 + \int_0^t \|\epsilon\|_a^2 + 2\int_0^t \left\langle P_0^\forall f - f, e\right\rangle. \quad (3.18)$$

Using Proposition 3.5, we obtain

$$\left(\|e(t)\|^2 + \int_0^t \|e\|_a^2\right)^{1/2} \leq \left(\|e(0)\|^2 + \int_0^t \|\epsilon\|_a^2\right)^{1/2} + 2\left(\int_0^t \|P_0 f - f\|_{\mathrm{H}^{-1}(\Omega)}^2\right)^{1/2}. \quad (3.19)$$

Assumption (2.29) and the discussion in §3.2 ensure then that

$$\left(\|e(t)\|^2+\int_0^t\|e\|_a^2\right)^{1/2}$$
$$\leq\left(\|e(0)\|^2+C_0^2\int_0^t\mathscr{E}[U]^2\right)^{1/2}+2\left(\int_0^t\|P_0f-f\|^2_{\mathrm{H}^{-1}(\Omega)}\right)^{1/2},\quad(3.20)$$

which implies the claim. □

3.7. **Remark** (short versus long integration times)**.** The bound for the pointwise in time $\mathrm{L}_2(\Omega)$ error, $\|e(t)\|$, appearing on the left-hand side of (3.17), is tight only for very short times. For example, it is well-known that on a uniform mesh of size $h\to 0$ on a convex domain Ω the energy term $\left(\int_0^t\|e\|_a^2\right)^{1/2}$ is $\mathrm{O}(h^p)$, while $\|e(t)\|$ is $\mathrm{O}(h^{p+1})$.

3.8. **Remark** (dealing with the $\mathrm{H}^{-1}(\Omega)$ norm)**.** In practise the $\mathrm{H}^{-1}(\Omega)$ norm can be well approximated as shown by Lemma 3.9, so, in the lack of a priori information, the last term in (3.17) may be replaced using the Poincaré inequality

$$2\,\|P_0f-f\|_{\mathrm{L}_2(0,T;\mathrm{H}^{-1}(\Omega))}\leq 2C_{\mathrm{P}}(\Omega)\,\|P_0f-f\|_{\mathrm{L}_2(\Omega\times(0,T))}\,.\quad(3.21)$$

It is also possible to obtain bounds by using the Cauchy–Bunyakovskiy–Schwarz inequality for $\mathrm{L}_2(\Omega)$ on the term $\langle P_0f-f,e\rangle$—rather than the $(\mathrm{H}^{-1},\mathrm{H}^1_0)$ duality—and "absorb" the resulting $\sup_{[0,t]}\|e\|$ into the first term on the right hand side of (3.17). However, whenever possible, we shy away from this procedure as it incurs in artificially higher constants and a $\mathrm{L}_1\,[0,T]$ accumulation on the right-hand side while the energy term on the left-hand side accumulates like $\mathrm{L}_2\,[0,T]$. This time-accumulation disparity between the error and the estimator is likely to result in an error–estimator ratio bound that has the order of $\sqrt{T}$, that is, although having the right order of convergence, the estimator will overestimate the error over long integration times.

We show now how to practically approximate the $\mathrm{H}^{-1}(\Omega)$ norm of an arbitrary given function $v\in\mathrm{L}_2(\Omega)$.

3.9. **Lemma** (computing the $\mathrm{H}^{-1}(\Omega)$ norm)**.** *Let $v\in\mathrm{L}_2(\Omega)$, consider the functions $\psi\in\mathrm{H}^1_0(\Omega)$ and $\Psi\in\mathbb{V}$ such that*

$$\mathscr{A}\psi=v\ \text{and}\ A\Psi=P_0v,\quad(3.22)$$

where A and P_0 are the discrete Laplacian and the $\mathrm{L}_2(\Omega)$ projection on $\mathbb{V}$, respectively. Then, recalling our convention whereby $\|v\|_{\mathrm{H}^1_0(\Omega)}=\|\nabla v\|$ we have

$$\|v\|^2_{\mathrm{H}^{-1}(\Omega)}=\|\psi\|^2_{\mathrm{H}^1_0(\Omega)}=\|\psi-\Psi\|^2_{\mathrm{H}^1_0(\Omega)}+\|\Psi\|^2_{\mathrm{H}^1_0(\Omega)}\,.\quad(3.23)$$

Proof With ψ and Ψ as given in (3.22) we have $\Phi\in\mathbb{V}$

$$\langle\mathscr{A}\psi-A\Psi\,|\,\Phi\rangle=\langle v-P_0v,\Phi\rangle=0,\quad(3.24)$$

i.e., that $\psi-\Psi$ is Galerkin-orthogonal to $\mathbb{V}$. Also, we have

$$\|v\|_{\mathrm{H}^{-1}(\Omega)}=\|\psi\|_{\mathrm{H}^1_0(\Omega)}\,.\quad(3.25)$$

Indeed, on the one hand

$$\begin{aligned}\|v\|_{\mathrm{H}^{-1}(\Omega)}&:=\sup_{\phi\in\mathrm{H}^1_0(\Omega)}\frac{\langle v,\phi\rangle}{\|\phi\|_{\mathrm{H}^1_0(\Omega)}}=\sup_{\phi\in\mathrm{H}^1_0(\Omega)}\frac{\langle\nabla v,\nabla\phi\rangle}{\|\phi\|_{\mathrm{H}^1_0(\Omega)}}\\&\leq\sup_{\phi\in\mathrm{H}^1_0(\Omega)}\frac{\|\psi\|_{\mathrm{H}^1_0(\Omega)}\,\|\phi\|_{\mathrm{H}^1_0(\Omega)}}{\|\phi\|_{\mathrm{H}^1_0(\Omega)}}=\|\psi\|_{\mathrm{H}^1_0(\Omega)}\,,\end{aligned}\quad(3.26)$$

and, on the other hand

$$\|v\|_{\mathrm{H}^{-1}(\Omega)} := \sup_{\phi\in\mathrm{H}^1_0(\Omega)} \frac{\langle\nabla\psi,\nabla\phi\rangle}{\|\phi\|_{\mathrm{H}^1_0(\Omega)}} \geq \frac{\langle\nabla\psi,\nabla\psi\rangle}{\|\psi\|_{\mathrm{H}^1_0(\Omega)}} = \|\psi\|_{\mathrm{H}^1_0(\Omega)}\,. \tag{3.27}$$

By the above, Galerkin-orthogonality and Pythagoras's Theorem, we have

$$\|v\|^2_{\mathrm{H}^{-1}(\Omega)} = \|\psi\|^2_{\mathrm{H}^1_0(\Omega)} = \|\psi-\Psi\|^2_{\mathrm{H}^1_0(\Omega)} + \|\Psi\|^2_{\mathrm{H}^1_0(\Omega)}\,. \tag{3.28}$$

□

3.10. **Remark** (the $\mathrm{H}^{-1}(\Omega)$ norm approximation)**.** The next-to-last term $\|\psi-\Psi\|_{\mathrm{H}^1_0(\Omega)}$ is the error of a function and its Ritz projection. This can be easily estimated with a fully computable a posteriori estimator functional $\mathscr{E}$ such that

$$\|\psi-\Psi\|_{\mathrm{H}^1_0(\Omega)} \leq \mathscr{E}[\Psi, v, \mathbb{V}] = \mathrm{O}(h^r_{\mathbb{V}}), \tag{3.29}$$

where $h_{\mathbb{V}}$ is the "mesh-size" of the space $\mathbb{V}$.

Hence the $\mathrm{H}^{-1}(\Omega)$ can be computed using the relation:

$$\|v\|^2_{\mathrm{H}^{-1}(\Omega)} = \|\Psi\|^2_{\mathrm{H}^1_0(\Omega)} + \zeta[\Psi,v]^2, \text{ where } \zeta[\Psi,v] \leq \mathscr{E}[\Psi]\,. \tag{3.30}$$

The term $\|\Psi\|_{\mathrm{H}^1_0(\Omega)}$ is clearly computable, by computing Ψ, which involves one $\mathrm{L}_2(\Omega)$-projection, one stiffness matrix inversion and one (discrete) energy norm computation. Furthermore

$$\|v\|^2_{\mathrm{H}^{-1}(\Omega)} = \|\Psi\|^2_{\mathrm{H}^1_0(\Omega)} + \mathrm{O}(h^{2r}_{\mathbb{V}}). \tag{3.31}$$

Hence, if Ψ is finite with respect to the mesh-size $h_{\mathbb{V}}$, i.e., $\|\Psi\| = \mathrm{O}(h^0_{\mathbb{V}})$, then we can approximate the $\mathrm{H}^{-1}(\Omega)$ of a function with as much precision as the finite element method allows it for the energy norm. On the other hand if Ψ is small, precisely, $\Psi = \mathrm{O}(h^s_{\mathbb{V}})$ with $s>0$ (implying that v is small as well), then this result has to be handled with more care for the error to be of some order of $h_{\mathbb{V}}$ higher than the computed quantity.

3.11. **Remark** (sharper versions of Theorem 3.6)**.** The error estimate (3.17) can be tightened further to

$$\left(\frac{1}{2}\|e(t)\|^2 + \int_0^t \|e\|^2_a\right)^{1/2} \leq \frac{1}{\sqrt{2}}\|e(0)\| + \left(\int_0^t \|P_0 f - f\|_{\mathrm{H}^{-1}(\Omega)} + C_0^2\mathscr{E}[U]^2\right)^{1/2}. \tag{3.32}$$

But this estimate becomes noticeably better only when one of the terms $\|e(0)\|$ or $\|P_0 f - f\|_{\mathrm{H}^{-1}(\Omega)}$ dominates the $\mathscr{E}[U]$ term, which should not be allowed to happen. So there is no need to lengthen the discussion by insisting on such tight bounds, as long as it is possible to obtain the elliptic a posteriori estimate constant C_0 in the leading term on the right-hand side.

4. Fully Discrete scheme

The main result of this section—and the paper—is the a posteriori error bound, stated in Theorem 4.6, on the error between the approximate solution U of the fully discrete problem (2.18) and that of the exact problem (2.6).

The analysis in this section follows narrowly the one we performed in §3, albeit with the complications that the fully discrete scheme imports. We will first extend the discrete solution sequence to a continuous-time function. Then we derive an error identity on which we mimic the energy techniques of §3 to bound the error's energy norm in terms of some residual terms and the elliptic error's energy norm, which is finally controlled via gradient recovery estimators.

4.1. **Time extension of the discrete solution.** Recalling the fully discrete scheme (2.18), the fully discrete solution is the sequence of finite element functions $U^n \in \mathbb{V}^n$ defined at each discrete time t_n, $n = 0, \dots, N$. Define the piecewise linear (affine) extension

$$U(t) := \sum_{n=0}^{N} U^n l_n(t), \tag{4.1}$$

where we use the one-dimensional piecewise linear continuous Lagrange basis functions , defined for $t \geq 0$, as

$$l_n(t) := \begin{cases} (t - t_{n-1})/\tau_n, & \text{for } t \in (t_{n-1}, t_n] \quad (\text{and } n > 0)\ , \\ (t_{n+1} - t)/\tau_{n+1}, & \text{for } t \in (t_n, t_{n+1}] \\ 0, & \text{otherwise.} \end{cases} \tag{4.2}$$

We warn the reader that we use the same symbol, U, to indicate the fully discrete solution's extension to $[0, T]$, as the one we used for its semidiscrete counterpart in § 3.

4.2. **Elliptic reconstruction and error splitting.** Next we define the elliptic reconstruction, needed for the following analysis, similarly to that of the semidiscrete scheme (cf. (3.1)). For each $n \in [0 : N]$, with the discrete elliptic operator A^n as in 2.3, we define the corresponding elliptic reconstruction operator $\mathscr{R}^n : \mathbb{V}^n \to \mathrm{H}_0^1(\Omega)$, for each $V \in \mathbb{V}^n$, by solving for $\mathscr{R}^n V$ the elliptic problem

$$\mathscr{A}\mathscr{R}^n V = A^n V, \tag{4.3}$$

which can be read in the weak form as

$$a(\mathscr{R}^n V, \Phi) = \langle A^n V, \Phi \rangle \quad \forall\, \Phi \in \mathrm{H}_0^1(\Omega). \tag{4.4}$$

We denote

$$\omega^n := \mathscr{R}^n U^n, \text{ for each } n = 0, \dots, N, \tag{4.5}$$

and this sequence's piecewise linear extension by $\omega : [0, T] \to \mathrm{H}_0^1(\Omega)$, i.e.,

$$\omega(t) := \sum_{n=0}^{N} \omega^n l_n(t). \tag{4.6}$$

As in the semidiscrete analysis we introduce symbols for the *full error* $e := U - u$, the *elliptic error* $\epsilon := \omega - U$ and the *parabolic error* $\rho := \omega - u$, whereby

$$e = \rho - \epsilon, \tag{4.7}$$

and, based on the Assumption 2.9,

$$\begin{aligned} \|\epsilon(t)\|_a &\leq C_0 \mathscr{E}\left[U^n l_n(t) + U^{n-1} l_{n-1}(t)\right] \\ &\leq C_0 \left(\mathscr{E}[U^n]\, l_n(t) + \mathscr{E}\left[U^{n-1}\right] l_{n-1}(t)\right) \text{ for } t \in [t_{n-1}, t_n]\,. \end{aligned} \tag{4.8}$$

The last step is guaranteed by the linearity of the operators G and ∇, hence the homogeneity $\mathscr{E}[\lambda V] = |\lambda|\, \|GV - \nabla V\|$, and by the triangle inequality .

4.3. **Lemma** (parabolic error identity)**.** *For each $n = 1, \dots, N$ and each $t \in (t_{n-1}, t_n)$ we have*

$$\partial_t e(t) + \mathscr{A}\rho(t) = \left(\Lambda^n U^{n-1} - U^{n-1}\right)/\tau_n + \mathscr{A}[\omega(t) - \omega^n] + P_0^n f^n - f(t). \tag{4.9}$$

Proof By the definition of U, (4.1), for each $n = 1, \dots, N$ and $t \in (t_{n-1}, t_n)$ we have

$$\partial_t U(t) = U^n l_n'(t) + U^{n-1} l_{n-1}'(t) = \left(U^n - U^{n-1}\right)/\tau_n \tag{4.10}$$

and using the fully discrete scheme (2.18), we have

$$\begin{aligned}\partial_t U(t) + \mathscr{A}\omega^n &= \left(\Lambda^n U^{n-1} - U^{n-1}\right)/\tau_n + \left(U^n - \Lambda^n U^{n-1}\right)/\tau_n + A^n U^n \\ &= \left(\Lambda^n U^{n-1} - U^{n-1}\right)/\tau_n + P_0^n f^n.\end{aligned} \tag{4.11}$$

Hence

$$\partial_t U(t) + \mathscr{A}\omega(t) = \left(\Lambda^n U^{n-1} - U^{n-1}\right)/\tau_n + \mathscr{A}[\omega(t) - \omega^n] + P_0^n f^n \tag{4.12}$$

and, using the exact PDE (2.6), we get

$$\begin{aligned}\partial_t e(t) + \mathscr{A}\rho(t) &= \partial_t U(t) + \mathscr{A}\omega(t) - \partial_t u(t) - \mathscr{A}u(t) \\ &= \left(\Lambda^n U^{n-1} - U^{n-1}\right)/\tau_n + \mathscr{A}[\omega(t) - \omega^n] + P_0^n f^n - f(t),\end{aligned} \tag{4.13}$$

as stated. □

4.4. **Definition** (a posteriori error indicators)**.** The notation we introduce here will be valid for the rest of the article.

elliptic error indicator via recovery:

$$\varepsilon_n := \mathscr{E}\left[U^n, \mathrm{H}_0^1(\Omega), \mathbb{V}^n\right] = C_0 \left\|\nabla U^n - G^n[U^n]\right\|, \tag{4.14}$$

with the functional $\mathscr{E}$ as defined in §2.8, and[1]

$$\tilde{\varepsilon}_n^2 := \frac{1}{3}\left(\varepsilon_n^2 + \varepsilon_{n-1}^2 + \varepsilon_n\varepsilon_{n-1}\right) \le \frac{1}{2}\left(\varepsilon_n^2 + \varepsilon_{n-1}^2\right). \tag{4.15}$$

time-discretisation error indicators:

$$\theta_n := \frac{1}{\sqrt{3}} \begin{cases} \left\|P_0^n f^n - \Lambda^n \partial U^n - \left(P_0^{n-1} f^{n-1} - \Lambda^{n-1}\partial U^{n-1}\right)\right\|_{\mathrm{H}^{-1}(\Omega)} & \text{for } n \ge 2, \\ \left\|P_0^1 f^1 - \Lambda^1 \partial U^1 - A^0 U^0\right\|_{\mathrm{H}^{-1}(\Omega)} & \text{for } n = 1, \end{cases} \tag{4.16}$$

where $\partial U^n := \left(U^n - U^{n-1}\right)/\tau_n$, (cf. Lemma 3.9), also possible to use in its alternative (faster to compute but not as sharp) version

$$\tilde{\theta}_n := C_\mu \left\|U^{n-1} - U^n\right\|_a, \tag{4.17}$$

where C_μ is dependent on the shape regularity μ of the family of triangulations defined in (2.10).

mesh-change (coarsening) error indicators: a main mesh-change indicator

$$\gamma_n := \tau_n^{-1} \left\|\Lambda^n U^{n-1} - U^{n-1}\right\|_{\mathrm{H}^{-1}(\Omega)}, \tag{4.18}$$

and a *higher order* mesh-change indicator

$$\tilde{\gamma}_n := {C_\mu}' \begin{cases} \left\|\hat{h}_n\left(P_0^n f^n - \Lambda^n \partial U^n - P_0^{n-1} f^{n-1} + \Lambda^{n-1}\partial U^{n-1}\right)\right\|, & n \ge 2, \\ \left\|\hat{h}_1\left(P_0^1 f^1 - \Lambda^1 \partial U^1 - A^0 U^0\right)\right\|, & n = 1, \end{cases} \tag{4.19}$$

where $\hat{h}_n(\boldsymbol{x}) = \max\left\{h_{n-1}(\boldsymbol{x}), h_n(\boldsymbol{x})\right\}$ for $\boldsymbol{x} \in \Omega$ and a constant C'_μ.

data approximation error indicator:

$$\beta_n := \tau_n^{-1} \int_{t_{n-1}}^{t_n} \left\|P_0^n f^n - f(t)\right\|_{\mathrm{H}^{-1}(\Omega)} \,\mathrm{d}t. \tag{4.20}$$

4.5. **Remark** (computing $\mathrm{H}^{-1}(\Omega)$ norms)**.** Clearly the $\mathrm{H}^{-1}(\Omega)$ norms appearing in Definition 4.4 cannot be computed in practise. The corresponding indicators can be replaced by upper bounds using the (dual) Poincaré inequality

$$\left\|\phi\right\|_{\mathrm{H}^{-1}(\Omega)} \le C_{\mathrm{P}} \left\|\phi\right\|. \tag{4.21}$$

Other alternatives, as described in Remark 3.8 are possible but will not be described here.

[1] In the numerical experiments we use $(\varepsilon_n^2 + \varepsilon_{n-1}^2)/2$ instead of $\tilde{\varepsilon}_n$.

4.6. **Theorem** (a posteriori estimate for fully discrete scheme)**.** *Let the sequence* $(U^n)_{n\in[0:N]}$, $U^n \in \mathbb{V}^n$, *be the solution of the fully discrete problem* (2.17) *and* U *its piecewise linear time-extension as in* (4.1). *Let* u *be the exact solution of the exact problem* (2.6) *then*

$$\left(\frac{\left\|U^N - u(T)\right\|^2}{2} + \int_0^T \|U(t) - u(t)\|_a^2 \,\mathrm{d}t\right)^{1/2} \leq \frac{\|U(0) - u(0)\|}{\sqrt{2}} + \eta_N \tag{4.22}$$

where the (global) error estimator *is given by the following discrete* $\mathrm{L}_2(0,T)$ *summation of the error indicators defined in §4.4:*

$$\eta_N^2 = \sum_{n=1}^N \left(\tilde{\varepsilon}_n + \gamma_n + \beta_n + \theta_n\right)^2 \tau_n. \tag{4.23}$$

Proof The proof shadows that of Lemma 3.3 and Theorem 3.6, but we must take into account the complications arising from the time discretisation. For the reader's convenience we divide it into steps.

Step 1. Using the notation from Lemma 4.3 and identity (4.9) therein we have that

$$\begin{aligned}\partial_t e(t) + \mathscr{A} e(t) = \mathscr{A}\epsilon(t) + \left(\Lambda^n U^{n-1} - U^{n-1}\right)/\tau_n \\ + \mathscr{A}[\omega(t) - \omega^n] + P_0^n f^n - f(t).\end{aligned} \tag{4.24}$$

Testing this with e we obtain

$$\begin{aligned}\frac{1}{2}\,\mathrm{d}_t \|e(t)\|^2 + \|e(t)\|_a^2 = a(\epsilon(t), e(t)) + \left\langle \left(\Lambda^n U^{n-1} - U^{n-1}\right)/\tau_n, e(t)\right\rangle \\ + \left\langle \mathscr{A}[\omega(t) - \omega^n], e(t)\right\rangle + \left\langle P_0^n f^n - f(t), e(t)\right\rangle,\end{aligned} \tag{4.25}$$

for all $t \in (t_{n-1}, t_n)$ and each $n = 1, \ldots, N$. Integrating over $[0,T]$ gives

$$\begin{aligned}\left\|e^N\right\|^2/2 + \int_0^T \|e(t)\|_a^2 \,\mathrm{d}t = \left\|e^0\right\|^2/2 + \int_0^T a(\epsilon(t), e(t)) \,\mathrm{d}t \\ + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \left\langle \left(\Lambda^n U^{n-1} - U^{n-1}\right)/\tau_n, e(t)\right\rangle \\ + a(\omega(t) - \omega^n, e(t)) + \left\langle P_0^n f^n - f(t), e(t)\right\rangle \mathrm{d}t \\ =: \mathcal{B}_1 + \mathcal{B}_2 + \mathcal{B}_3 + \mathcal{B}_4 + \left\|e^0\right\|^2/2.\end{aligned} \tag{4.26}$$

We proceed by bounding each of the terms $\mathcal{B}_j$, $j = 1, \ldots, 4$, appearing in the right-hand side of (4.26).

Step 2. The first term to be bounded in (4.26) yields the spatial discretisation error indicator as follows:

$$\begin{aligned}\mathcal{B}_1 = \int_0^T a(\epsilon(t), e(t)) \,\mathrm{d}t = \sum_{n=1}^N \int_{t_{n-1}}^{t_n} a(\epsilon(t), e(t)) \,\mathrm{d}t \\ \leq \sum_{n=1}^N \left(\int_{t_{n-1}}^{t_n} \|\epsilon\|_a^2\right)^{1/2} \left(\int_{t_{n-1}}^{t_n} \|e\|_a^2\right)^{1/2} \leq \sum_{n=1}^N \tilde{\varepsilon}_n \tau_n^{1/2} \left(\int_{t_{n-1}}^{t_n} \|e\|_a^2\right)^{1/2}\end{aligned} \tag{4.27}$$

where we have used (4.15) and in view of (4.8) and (4.15), we may write

$$\int_{t_{n-1}}^{t_n} \|\epsilon\|_a^2 \leq \varepsilon_{n-1}^2 \int_{t_{n-1}}^{t_n} l_{n-1}^2 + 2\varepsilon_{n-1}\varepsilon_n \int_{t_{n-1}}^{t_n} l_{n-1} l_n + \varepsilon_n^2 \int_{t_{n-1}}^{t_n} l_n^2 = \tilde{\varepsilon}_n^2 \tau_n. \tag{4.28}$$

The second term in (4.26) contains mesh-change term which we bound as follows:

$$\begin{aligned}\mathcal{B}_2 &= \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n}\left\langle \left(\Lambda^n U^{n-1} - U^{n-1}\right)/\tau_n, e(t)\right\rangle \,\mathrm{d}t \\ &\leq \sum_{n=1}^{N}\left\|\Lambda^n U^{n-1} - U^{n-1}\right\|_{\mathrm{H}^{-1}(\Omega)}\tau_n^{-1}\int_{t_{n-1}}^{t_n}\|e(t)\|_a \,\mathrm{d}t \\ &\leq \sum_{n=1}^{N}\gamma_n\tau_n^{1/2}\left(\int_{t_{n-1}}^{t_n}\|e\|_a^2\right)^{1/2}\end{aligned} \tag{4.29}$$

where γ_n is defined by (4.18).

Similarly the data error term is bounded as follows

$$\mathcal{B}_4 = \int_0^T \langle P_0^n f^n - f(t), e(t)\rangle \,\mathrm{d}t \leq \sum_{n=1}^{N}\beta_n\tau_n^{1/2}\left(\int_{t_{n-1}}^{t_n}\|e\|_a^2\right)^{1/2}, \tag{4.30}$$

where β_n is defined in (4.20).

Step 3. The third term in (4.26) yields a time discretisation term and is a bit more involved to estimate. Using the definition of ω^n, ω and $\mathscr{R}^n$, given in (4.3) and (4.6), we observe that

$$\begin{aligned}\mathcal{B}_3 &= \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} a(\omega - \omega^n, e(t))\,\mathrm{d}t \\ &= \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} a\big(l_{n-1}(t)\mathscr{R}^{n-1}U^{n-1} + l_n(t)\mathscr{R}^n U^n - \mathscr{R}^n U^n, e(t)\big)\,\mathrm{d}t \\ &= \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} l_{n-1}(t)a\big(\mathscr{R}^{n-1}U^{n-1} - \mathscr{R}^n U^n, e(t)\big)\,\mathrm{d}t \\ &= \sum_{n=1}^{N}\int_{t_{n-1}}^{t_n} l_{n-1}(t)\left\langle A^{n-1}U^{n-1} - A^n U^n, e(t)\right\rangle \,\mathrm{d}t \\ &\leq \sum_{n=1}^{N}\left\|A^{n-1}U^{n-1} - A^n U^n\right\|_{\mathrm{H}^{-1}(\Omega)}\left(\int_{t_{n-1}}^{t_n} l_{n-1}^2\right)^{1/2}\left(\int_{t_{n-1}}^{t_n}\|e\|_a^2\right)^{1/2} \\ &\leq \sum_{n=1}^{N}\theta_n\tau_n^{1/2}\left(\int_{t_{n-1}}^{t_n}\|e\|_a^2\right)^{1/2},\end{aligned} \tag{4.31}$$

where in the last passage we use the discrete scheme (2.18) for the substitution

$$A^n U^n = \left(\Lambda^n U^{n-1} - U^n\right)/\tau_n + P_0^n f^n \text{ for } n \geq 1. \tag{4.32}$$

Step 4. Grouping together (4.26), (4.27), (4.29), (4.30) and (4.31), we obtain

$$\begin{aligned}&\left\|e^N\right\|^2/2 + \int_0^T \|e(t)\|_a^2 \,\mathrm{d}t \\ &\qquad\leq \left\|e^0\right\|^2/2 + \sum_{n=1}^{N}\left(\tilde{\varepsilon}_n + \gamma_n + \beta_n + \theta_n\right)\tau_n^{1/2}\left(\int_{t_{n-1}}^{t_n}\|e\|_a^2\right)^{1/2}.\end{aligned} \tag{4.33}$$

Using an L_2 simplification (cf. §3.5), we conclude that

$$\left(\frac{\left\|e^N\right\|^2}{2} + \int_0^T \|e(t)\|_a^2 \,\mathrm{d}t\right)^{1/2} \leq \frac{\left\|e^0\right\|}{\sqrt{2}} + \left(\sum_{n=1}^{N}\left(\tilde{\varepsilon}_n + \gamma_n + \beta_n + \theta_n\right)^2\tau_n\right)^{1/2}. \tag{4.34}$$

Referring to the notation in (4.1) and Definition 4.4, we obtain the result. □

4.7. **Remark** (the alternative time indicator)**.** Assuming there is no mesh change from time t_{n-1} to time t_n, then the discrete Laplacians A^{n-1} and A^n, defined in (2.15), coincide. Thus the time discretisation error indicator θ_n, which is part of the estimator η_N in Theorem 4.6, can be written as

$$\theta_n = \frac{1}{\sqrt{2}} \left\| A^n U^n - A^{n-1} U^{n-1} \right\|_{\mathrm{H}^{-1}(\Omega)} = \frac{\tau_n}{\sqrt{2}} \left\| A^n \partial_t U \right\|_{\mathrm{H}^{-1}(\Omega)}. \tag{4.35}$$

In the form given in (4.16) and using the dual Poincaré inequality (4.21), this indicator is easily bounded.

A more precise, but slightly more expensive, calculation can be done using Lemma 3.9. The same idea, will be used in the next result where we show that the indicator θ_n is equivalent, up to higher order terms, to the alternative time indicator $\tilde{\theta}_n$, defined in (4.17), which requires only an energy norm computation. This alternative time indicator, which is more common in energy estimates [Pic98, e.g.], $\tilde{\theta}_n$ is also more "natural", as it measures the time derivative in the energy norm as opposed to the H^{-1} norm of the time derivative of AU. Due to mesh-change effects, this simpler indicator comes at the (affordable) price of having to add the higher order mesh change term $\tilde{\gamma}_n$ to the otherwise simpler γ_n.

4.8. **Theorem** (alternative time estimator)**.** *With the same assumptions and notation of Theorem 4.6 we have*

$$\left(\frac{\left\| U^N - u(T) \right\|^2}{2} + \int_0^T \left\| U(t) - u(t) \right\|_a^2 \, \mathrm{d}t \right)^{1/2} \leq \frac{\left\| U(0) - u(0) \right\|}{\sqrt{2}} + \tilde{\eta}_N \tag{4.36}$$

where the (alternative global) error estimator *is given by the following discrete* $\mathrm{L}_2(0,T)$ *summation of the error indicators defined in §4.4:*

$$\tilde{\eta}_N^2 := \sum_{n=1}^{N} \left(\tilde{\varepsilon}_n + \gamma_n + \tilde{\gamma}_n + \beta_n + \tilde{\theta}_n \right)^2 \tau_n. \tag{4.37}$$

Proof We proceed similarly to the proof of Theorem 4.6, in steps. The notation is the same and steps 1 and 2 are identical.

Step 3. This step starts similarly to its homologue in the proof of Theorem 4.6 by observing that

$$\mathcal{B}_3 = \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} l_{n-1}(t) \left\langle A^{n-1} U^{n-1} - A^n U^n, e(t) \right\rangle \, \mathrm{d}t. \tag{4.38}$$

The function $A^{n-1}U^{n-1} - A^n U^n$ belongs to $\mathbb{V}^n + \mathbb{V}^{n-1}$, but in general it is in neither of $\mathbb{V}^n$ nor $\mathbb{V}^{n-1}$). Thus, to proceed, we use the $\mathrm{L}_2(\Omega)$-projection and the Clément–Scott–Zhang interpolator denoted respectively by

$$\check{P}^n : \mathrm{L}_2(\Omega) \to \mathbb{V}^n + \mathbb{V}^{n-1} \text{ and } \hat{I}\!I^n : \mathrm{L}_2(\Omega) \to \mathbb{V}^n \cap \mathbb{V}^{n-1}. \tag{4.39}$$

We recall that the operators $\hat{I}\!I^n$ and $\check{P}^n$ are both known [SZ90, Car02, resp.] to enjoy the following stability properties for all $n = 0, \ldots, N$:

$$\left\| \hat{I}\!I^n \phi \right\|_a \leq C_{1,\mu} \left\| \phi \right\|_a \quad \forall\, \phi \in \mathrm{H}^1(\Omega), \tag{4.40}$$

$$\left\| \check{P}^n \phi \right\|_a \leq C_{2,\mu} \left\| \phi \right\|_a \quad \forall\, \phi \in \mathrm{H}^1(\Omega), \tag{4.41}$$

where μ is the shape-regularity of the triangulation family $\{\mathscr{T}^n\}_{n=0,\ldots,N}$ defined in (2.10). Furthermore, the following interpolation inequality is valid [LM06, §B.3]

$$\left\| \left(\psi - \hat{I}\!I^n \psi \right) / \hat{h}_n \right\| \leq C_{3,\mu} \left\| \psi \right\|_a \quad \forall\, \psi \in \mathrm{H}_0^1(\Omega),\ n = 1, \ldots, N, \tag{4.42}$$

where $\hat{h}_n := \max\{h_n, h_{n-1}\}$.

Step 4. Using these operators, we derive that

$$\begin{aligned}
\mathcal{B}_3 &= \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \left\langle A^{n-1}U^{n-1} - A^nU^n, \check{P}^n e(t)\right\rangle l_{n-1}(t)\,\mathrm{d}t \\
&= \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big(\left\langle A^{n-1}U^{n-1} - A^nU^n, \check{P}^n e(t) - \hat{\Pi}^n\check{P}^n e(t)\right\rangle \\
&\qquad + \left\langle A^{n-1}U^{n-1} - A^nU^n, \hat{\Pi}^n\check{P}^n e(t)\right\rangle \Big) l_{n-1}(t)\,\mathrm{d}t \\
&\le \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \Big(\left\|\hat{h}_n\big(A^{n-1}U^{n-1} - A^nU^n\big)\right\| \left\|\hat{h}_n^{-1}\big(\check{P}^n e(t) - \hat{\Pi}^n\check{P}^n e(t)\big)\right\| \\
&\qquad + a\big(U^{n-1} - U^n, \hat{\Pi}^n\check{P}^n e(t)\big)\Big) l_{n-1}(t)\,\mathrm{d}t.
\end{aligned} \tag{4.43}$$

Using inequalities (4.40), (4.41) and (4.42), we get the bound

$$\begin{aligned}
\mathcal{B}_3 &\le \sum_{n=1}^{N} \int_{t_{n-1}}^{t_n} \big(C_{3,\mu} \left\|\hat{h}_n(A^{n-1}U^{n-1} - A^nU^n)\right\| \left\|\check{P}^n e(t)\right\|_a \\
&\qquad + C_{1,\mu} \left\|U^{n-1} - U^n\right\|_a \left\|\check{P}^n e(t)\right\|_a \big) l_{n-1}(t)\,\mathrm{d}t \\
&\le \sum_{n=1}^{N} \Big(C_{3,\mu} \left\|\hat{h}_n(A^{n-1}U^{n-1} - A^nU^n)\right\| + C_{1,\mu} \left\|U^{n-1} - U^n\right\|_a \Big) \\
&\qquad \times C_{2,\mu} \int_{t_{n-1}}^{t_n} \|e(t)\|_a \, l_{n-1}(t)\,\mathrm{d}t \\
&\le \sum_{n=1}^{N} \Big(\tilde{\gamma}_n + \tilde{\theta}_n\Big) \left(\int_{t_{n-1}}^{t_n} \|e\|_a^2 \right)^{1/2}
\end{aligned} \tag{4.44}$$

by taking $C_\mu := C_{1,\mu}C_{2,\mu}/3$, $C_\mu{}' := C_{3,\mu}C_{2,\mu}/3$ in (4.17) and (4.19) for the last step.
We may now conclude exactly like the last step in the proof of Theorem 4.6, albeit with θ_n replaced by $\tilde{\gamma}_n + \tilde{\theta}_n$. □

5. Computer experiments: convergence rates

In this section and in §6 we study the numerical behaviour of the error indicators and estimators and compare this behaviour with the true error on three model problems. The C code that we used includes the adaptive FEM library ALBERTA [SS05,]. The quadrature formal error is made negligible with respect to other error by using overkill quadrature formulas (exact on polynomials of degree 17 and less).

5.1. **Benchmark problems.** Consider three benchmark problems, the solution of which is known. Namely, take $d = 2$, each problem's data f, u_0 is then chosen such that the exact solution to 2.6 is given by:

$$u(\boldsymbol{x}, t) = \sin(\pi t) \exp\Big(-10\,|\boldsymbol{x}|^2\Big), \tag{5.1}$$

$$u(\boldsymbol{x}, t) = \sin(20\pi t) \exp\Big(-10\,|\boldsymbol{x}|^2\Big), \tag{5.2}$$

$$u(\boldsymbol{x}, t) = t \sin\frac{2\arctan(x_2/x_1)}{3} \boldsymbol{x}^{2/3} \exp\left(\frac{-1}{1-|\boldsymbol{x}|^2}\right), \tag{5.3}$$

The domain Ω for Problems (5.1) and (5.2) is the square $S := (-1,1)\times(-1,1)$. Problem (5.3), whose solution's gradient is singular at the origin, is considered on the L shaped domain $\Omega = S \setminus [0,1]\times[-1,0]$. The benchmark problems (5.1) and

(5.2) have been chosen such that they can be compared with previous numerical studies [LM06,].

For all Problems (5.1)–(5.3), we take zero initial condition, $u_0 = 0$ to avoid dealing with the initial adaptivity which is a side issue here.

The solution (5.2) has a time dominant discretisation error, while (5.3) was constructed to have a dominant spatial error. It is the product of a linear function in time, a well known solution to Laplace's equation producing the spatial singularity and a mollifier.

Problem (5.1) is used to test asymptotic behaviour of the indicators under uniform space-time refinements further in §5.5. Problems (5.3) and (5.2) will be used to test the adaptive strategies in §6.

5.2. **Gradient recovery implementation.** The recovery operator, G^n, is obtained by taking the discontinuous gradients of the numerical solution at the super convergent sampling points [AO00,] (and references therein). The recovery operator used here is built in the following way: fixing $V \in \mathbb{V}^n$, for each degree of freedom $\boldsymbol{x}$, we define

$$G^n[V](\boldsymbol{x}) := \frac{\sum_{K\in\mathscr{T}^n:\boldsymbol{x}\in K}|K|\,\nabla V|_K(\boldsymbol{x})}{\sum_{K\in\mathscr{T}^n:\boldsymbol{x}\in K}|K|}, \tag{5.4}$$

This defines a unique piecewise polynomial field $G^n[V] \in \mathbb{V}^d$. (Note that formula (5.4) is non trivial for only for those DOF that are are on the boundary of an element; for the internal DOF, that arise in using $\mathbb{P}^p$ elements for $p \geq 3$, it is not necessary to calculate anything.)

5.3. **Definition** (experimental order of convergence)**.** Given two sequences $a(i)$ and $h(i) \searrow 0$, $i = l, \ldots,$, we define experimental order of convergence (EOC) to be the local slope of the $\log a(i)$ vs. $\log h(i)$ curve, i.e.,

$$\mathrm{EOC}(a, h; i) := \frac{\log(a(i+1)/a(i))}{\log(h(i+1)/h(i))}. \tag{5.5}$$

5.4. **Definition** (effectivity index)**.** The main tool deciding the quality of an estimator is the effectivity index (EI) which is the ratio of the error and the estimator, i.e.,

$$\mathrm{EI}(t_n) := \eta_n / \|U - u\|_{\mathrm{L}_2(0,t_n;\mathrm{H}^1_0(\Omega))}. \tag{5.6}$$

If $\mathrm{EI}(t_n) \to 1$ as $\sup_{x,n} h_n(x) \to 0$ then we say the estimator is *asymptotically exact.*

5.5. **Indicator's numerical asymptotic behaviour.** In the following convergence rate tests we discuss the practical realisation of Theorems 4.6 and 4.8, to which we refer for notation.

We use a uniform timestep and uniform meshes that are fixed with respect to time. Hence for each test we have $\mathbb{V}^n = \mathbb{V}^0 = \mathbb{V}$ and $\tau_n = \tau(h)$ for all $n = 1, \ldots, N$. For each test we fix the polynomial degree p and two parameters k, c and then compute a sequence of solutions with $h = h(i) = 2^{-i/2}$, and $\tau = ch^k$ for a sequence of refinement levels $i = l, \ldots, L$.

Due to the finite element space invariance in time, the coarsening indicator γ_n vanishes and is thus not computed (this indicator will be discussed in §6).

The initial value being zero makes the initial error $U(0) - u(0)$ zero. Thus we do not need to calculate this term in the estimator.

For all solutions the boundary values are not exactly zero, but of a negligible value, hence little interpolation error is committed here (nonetheless some care is taken when dealing with very small errors). Finally, the data approximation error term, β_n, though important for highly oscillatory data, will not be studied here given the regularity of our data.

Therefore, what we compute on a space-time uniform mesh are the indicators ε_n and θ_n (or $\tilde{\theta}_n$,γ_n), defined in §4.4, and the corresponding *cumulative indicators* $(E_n)_{n=1,\dots,N}$ and $(\Theta_n)_{n=1,\dots,N}$ defined by:

$$\begin{aligned} E_m^2 &:= \sum_{n=1}^{m} \left(\varepsilon_n^2 + \varepsilon_{n-1}^2\right) \tau_n/2 \qquad \text{(for space)}\,, \\ \text{and } \Theta_m^2 &:= \sum_{n=1}^{m} \theta_n^2 \tau_n \text{ or } \sum_{n=1}^{m} \left(\tilde{\theta}_n^2 + \tilde{\gamma}_n^2\right) \tau_n \qquad \text{(for time)}\,. \end{aligned} \tag{5.7}$$

From the Theorems 4.6 and 4.8, we know that

$$\|U^n - u(t_n)\|^2 \leq E_m^2 + \Theta_m^2 + \sum_{n=1}^{m} \beta_n^2 \tau_n. \tag{5.8}$$

Our results and the comments are reported in the captions of figures.

In Figures 1–4 we visualise the results and comment them, for Problem (5.1) for conforming finite elements of polynomial degree $p = 1, \dots, 4$, respectively. Having fixed p, k, c such that $\tau = ch^k$, for each level i, we plot Θ_m and E_m, $\|U - u\|_{\mathrm{L}_2(0,t_m;\mathrm{H}_0^1(\Omega))}$, their experimental order of convergence and the effectivity index $\mathrm{EI}(t_m)$ versus (discrete) time $t_m = 0, \dots, T$. The conclusion is that the estimator is sharp and reliable, but to achieve asymptotic exactness (or close) the time indicator must be made smaller than the space indicator by taking $\tau \ll h^p$. In all these tests we used the first form for Θ_m appearing in (5.7).

In Figure 5 we summarise a comparison between the two time indicators θ_n and $\tilde{\theta}_n$, showing that the latter yields a much sharper bound, but with the added cost of having to compute the higher order term $\tilde{\gamma}_n$.

6. Computer experiments: adaptive schemes

We present now an adaptive algorithm based on the error indicators defined in §4.4. As with many adaptive methods for time-dependent problems [Pic98, SS05, CJ04,], we perform space and time adaptivity separately. Adaptivity is controlled via the indicators η_n and η_n (or $\tilde{\eta}_n$)—see Theorems 4.6 and 4.8—which are kept under a given tolerance tol.

Namely, at each timestep $t_{n-1} \to t_n$, we use adaptive schemes for elliptic problems as to minimise the indicators $\tilde{\varepsilon}_n$ and β_n. There are different strategies to perform the timestep adaptivity, all geared towards minimising θ_n (or $\operatorname{var} \theta_n$). Finally, the coarsening estimator γ_n is minimised by precomputing it and performing only one coarsening operation at the beginning of each timestep.

Note that it is not in the scope of this paper to prove any rigorous result about the adaptive algorithm and, based on heuristic arguments only, we use it for illustration purposes.

6.1. **Space adaptivity via maximum strategy.**

At each timestep an elliptic problem is solved. For linear elliptic problems, convergence of adaptive schemes is reasonably well understood [MNS02, BDD04,] so we follow the criteria given therein, namely the Maximum Strategy.

The algorithm we used can be pseudocoded as follows.

6.2. **Space Adapt.**

Require: $(U^{\mathrm{old}}, \mathbb{V}^{\mathrm{old}}, \mathrm{tol}_\varepsilon, k_{\max}, t, \tau, \xi, \mathrm{tol}_\gamma)$
Ensure: $(U^{\mathrm{new}}, \mathbb{V}^{\mathrm{new}})$ solution of (2.17)
  procedure Coarsening
    $\boldsymbol{\gamma} = (\gamma^K)_{K \in \mathscr{T}} :=$ Coarsening Preindicator$(U\mathrm{old}, \mathbb{V}^{\mathrm{old}})$ (cf. §A.12).
    $\mathscr{T} :=$ Mesh$(\mathbb{V}^{\mathrm{old}})$

find $\mathscr{C} \subset \mathscr{T}$ such that $\sum_{K\in\mathscr{C}} (\gamma^K)^2 \le \mathrm{tol}_\gamma^2$
$\mathscr{T} := \mathsf{Coarsen}(\mathscr{T}, \mathscr{C})$ using [SS05, §1.1.2–1.1.3]
end procedure
procedure MAXIMUM STRATEGY REFINEMENT [SS05]
$k := 0$
compute ε_n using (4.14)
$\mathscr{R} := \emptyset$ ▷ refinement set
while $\varepsilon_n > \mathrm{tol}_\varepsilon$ and $k \le k_{\max}$ **do**
for $K \in \mathscr{T}^n$ **do**
if $\varepsilon_{K,n}^2 \ge \xi \max_{L\in\mathscr{T}^n} \varepsilon_{L,n}^2$ **then**
$\mathscr{R} := \{K\} \cup \mathscr{R}$ ▷ mark K for refinement
end if
end for
$\mathscr{T} := \mathsf{Refine}(\mathscr{T}, \mathscr{R})$ using [SS05, §1.1.1] ▷ hence update $(U^{\mathrm{old}}, \mathbb{V})$
set $\Lambda^n U^{n-1} := U^{\mathrm{old}}$, $\tau_n = \tau$, $t_n = t$ and solve for U^n in (2.18)
$U := U^n$
compute ε_n using (4.14)
$k := k + 1$
end while
end procedure
return $(U, \mathbb{V})$

6.3. **Coarsening.** In time-dependent problems mesh coarsening , which is not to be confused with the coarsening needed in proving optimal complexity for adaptive schemes [BDD04,], is used to reduce DOF that become redundant in time.

Mesh coarsening is a delicate procedure and should be used sparingly as to avoid needless overhead computing time. In Algorithm 6.2, coarsening is performed only once, at the beginning, for each time-step.

The coarsening strategy we propose is based on *predicting the effect of a possible removal of degrees of freedom.* The reason for this is that in ALBERTA (and many other finite element codes) upon coarsening, all DOF-dependent vectors (encoding finite element function coefficients) are "coarsened" via interpolation. This makes it possible to compute the effect of coarsening, and the coarsening estimator γ_n defined in (4.18), *before* mesh-change occurs. The details of this procedure are discussed in § A.

6.4. **Timestep control.** Timestep control can be achieved using two different strategies.

An *implicit timestep control* strategy used is ready implemented in ALBERTA [SS05,] using Algorithm 6.2 upon each timestep.

Here we propose an *explicit timestep control* strategy which we have implemented in ALBERTA. The reason for this is that the implicit strategy, though better in terms of timestep determination, is very time-consuming as it requires the repeated solution of the timestep. In contrast, the explicit strategy has a rougher—nonetheless still satisfactory— control over the timestep, but it is much faster. The conclusion is that the ideal control should be a smart implicit/explicit-switching algorithm.

The explicit strategy can be described as follows.

6.5. **Explicit Timestep Adapt.**

Require: $(\tau_0, t_0, T, \mathscr{T}^0, u^0, \mathrm{tol}_\varepsilon, k_{\max}, \xi, \mathrm{tol}_\gamma, \mathrm{tol}_{\theta,\min}, \mathrm{tol}_\theta)$
Ensure: $(\tau_n, \mathbb{V}^n, U^n)_{n=1,\dots,N}$ satisfying (2.17) and possibly $\int_0^T \|U-u\|^2 \le \mathrm{tol}^2$
$(U^0, \mathbb{V}^0) = \mathsf{Initial\ Space\ Adapt}(\mathscr{T}^0, u^0, k_{\max}, \xi, \kappa)$ ▷ data interpolation
$n := 1$

$\tau_n := \tau_{n-1}$
$t_n := t_{n-1} + \tau_n$
while $t_n \leq T$ **do**
  $(U^n, \mathbb{V}^n) :=$ Space Adapt$(U^{n-1}, \mathbb{V}^{n-1}, \mathrm{tol}_\varepsilon, k_{\max}, \tau_n, t_n, \xi, \mathrm{tol}_\gamma)$
  compute θ_n
  if $\theta_n > \mathrm{tol}_\theta$ **then**
    $\tau_{n+1} := \tau_n/\sqrt{2}$
  else if $\theta_n \leq \mathrm{tol}_{\theta,\min}$ **then**
    $\tau_{n+1} := \sqrt{2}\tau_n$
  end if
  $t_{n+1} := t_n + \tau_{n+1}$
  $n := n + 1$
end while
return $(U^n)_{n=1,\ldots,N}$,

where the *global tolerance* tol is given by the relation

$$\mathrm{tol}^2 = T\left(\mathrm{tol}_\theta^2 + \mathrm{tol}_\varepsilon^2 + \mathrm{tol}_\gamma^2\right). \tag{6.1}$$

Note that this algorithm does not guarantee reaching a tolerance, unlike more sophisticated ones found in the literature [CJ04, e.g.], but it guarantees termination in reasonable CPU times.

6.6. **Numerical results.** In Tables 1–3 we compare the implicit timestep control strategy described by algorithm 6.5 with a uniform timestep scheme. For the uniform strategy we take a stationary mesh in time and set $\tau = 0.04h^2$. We calculate the error for various numerical simulations using differing values of h using the uniform strategy and set those values as tolerances for the adaptive scheme varying ξ appropriately.

Each column displays results for either the uniform strategy or the adaptive strategy using various thresholds. These columns are further subdivided into two, the first containing $\sum_{n=1}^{N} \dim \mathbb{V}^n$ (i.e., the total number of degrees of freedom from all meshes over time) which we denote DOF and the second containing CPU time (secs) for all model problems (5.1)–(5.3).

	Uniform		Adaptive					
			$\xi = 0.65$		$\xi = 0.70$		$\xi = 0.75$	
tol	DOF's	CPU	DOF's	CPU	DOF's	CPU	DOF's	CPU
0.573	232,290	3	24,080	4	22,792	5	22,240	4
0.295	3,489,090	49	42,042	8	39,414	8	38,630	6
0.149	54,097,020	598	82,172	15	77,932	15	76,452	16
0.0625	OOM	OOM	206,709	39	195,810	37	191,650	37

TABLE 1. Explicit timestep control with various spatial maximum strategy thresholds for Problem (5.1). The adaptive method clearly saves DOF and CPU time over the uniform method.

	Uniform		Adaptive					
			$\xi = 0.65$		$\xi = 0.7$		$\xi = 0.75$	
tol	DOF's	CPU	DOF's	CPU	DOF's	CPU	DOF's	CPU
0.296	3,489,090	47	12,092	5	11,430	5	11,498	5
0.21	13,940,289	196	17,038	7	16,140	8	16,201	7
0.104	54,097,020	602	106,188	32	100,058	29	22,597	10
0.03125	OOM	OOM	513,694	120	460,637	118	449,568	115

TABLE 2. Explicit timestep control with various spatial maximum strategy thresholds for spatial-error dominant Problem (5.3). Adaptivity saves DOF and CPU.

	Uniform		Adaptive			
			$\xi = 0.7$		$\xi = 0.75$	
tol	DOF's	CPU	DOF's	CPU	DOF's	CPU
1.000	925,809	12	159,070	43	127,610	58
0.569	3,489,090	49	237,960	142	204,376	180
0.295	54,097,020	605	471,733	755	471,542	920
0.149	OOM	OOM	940,618	1410	940,138	1850

TABLE 3. Implicit timestep control with various spatial maximum strategy thresholds for spatial-error dominant Problem (5.2). Adaptivity saves DOF (even better than explicit control) but the CPU time grows very quickly due to overhead.

	Uniform		Adaptive			
			$\xi = 0.7$		$\xi = 0.75$	
tol	DOF's	CPU	DOF's	CPU	DOF's	CPU
1.000	925,809	12	135,788	5	127,004	4
0.569	3,489,090	49	198,628	7	194,311	8
0.295	54,097,026	605	397,716	15	395,876	16
0.149	OOM	OOM	2,177,666	79	2,079,081	76

TABLE 4. Explicit timestep control with various spatial maximum strategy thresholds for time-error dominant Problem (5.2)

6.7. **Remark** (implicit timestep control on fast oscillating solutions)**.** We take note of the CPU times from the results for Problem (5.2). These show that implicit timestep control is undesirable for fast oscillating functions. This is because the timestep searching becomes computationally inefficient. Numerical simulations for an explicit timestep control strategy is given in Table 4. This algorithm is described in detail in the ALBERTA manual [SS05, §1.5.4] The results show although for a method with low tolerance we use more degrees of freedom we make a substantial gain on the CPU time.

We then fix a value of ξ and compare an adaptive strategy with uniform for a single value of tol. This is to illustrate how the number of degrees of freedom of the mesh change over time, and how the implicit timestep control affects the timestep size for all test problems in Figures 6.

6.8. **Incompatible data singular solution.** We close the paper by testing the adaptive algorithm on an example with incompatible initial and boundary conditions, which is the type of situation where adaptivity is really needed in practise. Consider problem (2.6) with $\Omega = (0,1) \times (0,1)$, $f = 0$ and $u_0 = 1$. The initial conditions are thus incompatible with the homogeneous Dirichlet boundary valid for all positive times. The exact solution u, though singular at all points of $\partial\Omega \times \{0\}$, can be readily evaluated "by hand" and may be represented in terms of Fourier series of the Laplacian's eigenvalues. Namely, we have

$$u(\boldsymbol{x}, t) = \sum_{m,n=1}^{\infty} C_{m,n} \exp(-\left(m^2 + n^2\right) \pi^2 t) \sin(m\pi x_1) \sin(n\pi x_2), \text{ for } t > 0, \quad (6.2)$$

where the constant $C_{m,n}$ is given by

$$C_{m,n} = \frac{4}{nm\pi^2} (1 - \cos(m\pi) - \cos(n\pi) + \cos(n\pi)\cos(m\pi)). \quad (6.3)$$

Since the solution (6.2) is an infinite Fourier series it cannot be computed exactly, but its rapid decay allows to truncate early with machine-epsilon precision.

In order to generate a reference tolerance, which is common for the uniform and the adaptive scheme we couple $h = 0.05\tau$ and run the uniform refinement code. We use then the error computed as a tolerance for the adaptive scheme, results of this are shown in Figure 7. In Figure 8 we visualise the adapted FE mesh for Problem (6.2) at various times.

Acknowledgement

O.L.'s research was partially supported by a Nuffield Young Researcher's Grant.
T.P.'s research is fully supported by his EPSRC D.Phil. scholarship grant.
Both authors wish to thank Alan Demlow for an interesting exchange of ideas.

Appendix A. Coarsening error preindicator implementation

We describe next a practical implementation of the *coarsening error preindicator* (we use this term to emphasise the fact that this indicator can be computed a priori, as opposed to the other indicators involved in the adaptive strategy). Since we used ALBERTA for our computations, this section relies substantially on the principles described in the manual [SS05]. We briefly describe these principles in the next paragraph, in order to expose the main idea behind the coarsening preindicator.

A.1. **Refinement, coarsening and interpolation in ALBERTA.** Mathematically, a simplicial mesh (or partition, or triangulation) is a set of disjoint open simplexes, the union of the closure of which is $\overline{\Omega}$. A mesh into a new mesh is refined by *bisecting* a subset of its simplexes, following a special procedure which ensures mesh conformity (e.g., no hanging nodes) and does not deteriorate shape-regularity (on fully fitted polygonal domains). A mesh is thus represented as a binary tree, where each node represents a simplex. The children of each simplex are thus the 2 subsimplexes obtained by bisection. Hence, from a coding view-point, refinement means growing the binary tree.

The inverse of refinement is coarsening. Thus coarsening a mesh in ALBERTA consists in removing pairs of sibling simplexes (both marked for coarsening) and produces the new—coarsened—mesh where the pairs of siblings are replaced by their parent.

The coarsening preindicator is a real number defined on each simplex, of the triangulation to be coarsened. This estimator can in fact be *precomputed* with respect to coarsening. This is in contrast with usual *a posteriori* error estimators which can be postcomputed only (i.e., after the discrete solution has been computed). To clarify this point, let us focus on the particular situation of interest. Let U^{n-1} be the solution from the previous timestep; $U^{n-1} \in \mathbb{V}^{n-1}$, the finite element space with respect to mesh $\mathscr{T}^{n-1}$. The error due to coarsening appears in the term

$$U^{n-1} - \Lambda^n U^{n-1}. \tag{A.1}$$

This term is nonzero only when simplexes are coarsened.

Furthermore, we assume that the new mesh $\mathscr{T}^n$ is a refinement of $\mathscr{T}_0^n$, which is a coarsening of the old mesh $\mathscr{T}^{n-1}$:

$$\mathscr{T}^{n-1} \xrightarrow{\text{coarsen}} \mathscr{T}_0^n \xrightarrow{\text{refine}} \cdots \xrightarrow{\text{refine}} \mathscr{T}^n \tag{A.2}$$

If Λ_0^n is the Lagrange interpolator onto the finite element space $\mathbb{V}_0^n$, relative to the new coarse mesh $\mathscr{T}_0^n$, it is not very difficult to predict $\Lambda_0^n U^{n-1}$ without actually computing it. Therefore this term can be predicted from (a) the simplexes of $\mathscr{T}^{n-1}$ marked for coarsening which leads to $\mathscr{T}_0^n$ and (b) the values of U^{n-1}.

Note that since $\mathscr{T}_0^n$ is subsequently *refined but not coarsened* to produce $\mathscr{T}^n$, as depicted in (A.2), then the additional coarsening error will be zero. Namely, if Λ^n

denotes the Lagrange interpolant onto $\mathbb{V}^n$, the finite element space over $\mathscr{T}^n$, which is a refinement of $\mathscr{T}_0^n$, then $\Lambda^n U^{n-1} = \Lambda_0^n U^{n-1}$, and thus

$$U^{n-1} - \Lambda^n U^{n-1} = U^{n-1} - \Lambda_0^n U^{n-1}. \tag{A.3}$$

The coarsening strategy therefore consists in choosing a subset of simplexes of $\mathscr{T}^{n-1}$ which minimises term $\left\|U^{n-1} - \Lambda_0^n U^{n-1}\right\|$ *before* producing the new coarse mesh $\mathscr{T}_0^n$.

The rest of this section describes how $U^{n-1} - \Lambda_0^n U^{n-1}$ can be precomputed.

A.2. **Notation.** Let K be an element of the new coarse mesh $\mathscr{T}_0^n$ resulting from the coarsening of its two children which we denote by $K^\pm$. (Note that K^+ and K^- correspond to `child[0]` and `child[1]` of K in the ALBERTA manual [SS05].) Define the *fine space*

$$\mathbb{Y} := \left\{ \Phi|_K : \Phi \in \mathbb{V}^{n-1} \right\}. \tag{A.4}$$

Likewise define the *coarse space* $\mathbb{X}$ to be the local finite element space, i.e.,

$$\mathbb{X} := \left\{ \Phi|_K : \Phi \in \mathbb{V}_0^n \right\}; \tag{A.5}$$

simply put we just have $\mathbb{X} = \mathbb{P}^p$. We introduce also the *fine spaces* $\mathbb{Y}^\pm$, defined like $\mathbb{Y}$, but restricting functions over $K^\pm$, respectively (so functions in $\mathbb{Y}^\pm$ are in fact the same as $\mathbb{X} = \mathbb{P}^p$, albeit with different domains).

Denote by $\{\boldsymbol{x}_0, \dots, \boldsymbol{x}_L\}$ and $\left\{\boldsymbol{x}_0^\pm, \dots, \boldsymbol{x}_L^\pm\right\}$ the set of Lagrange degrees of freedom on the simplex K and its children $K^\pm$, respectively. We indicate with $\left\{\pi^0, \dots, \pi^L\right\}$ and $\left\{\pi_\pm^0, \dots, \pi_\pm^L\right\}$ the corresponding Lagrange polynomial bases of $\mathbb{X}$ and $\mathbb{Y}^\pm$, respectively, whereby

$$\pi^i(\boldsymbol{x}_j) = \pi_\pm^i(\boldsymbol{x}_j^\pm) = \delta_j^i. \tag{A.6}$$

For short we will write these bases as column vectors $\boldsymbol{\pi} = \left(\pi^0, \dots, \pi^L\right)^\intercal$, etc. We also define the (local) *coarse-on-fine matrixes* by

$$\mathbf{A}^\pm := \left(\boldsymbol{\pi}(\boldsymbol{x}_0^\pm) \ \dots \ \boldsymbol{\pi}(\boldsymbol{x}_L^\pm)\right) = \left(\pi^i(\boldsymbol{x}_j^\pm)\right)_{i,j=0,\dots,L}. \tag{A.7}$$

These matrixes are closely related to ALBERTA's *refine-interpolation* matrix [SS05, matrix A (1.5) in §1.4.4].

A.3. **Proposition** (coarse-on-fine matrix properties)**.** *The matrixes $\mathbf{A}^+$ and $\mathbf{A}^-$ are independent of K, K^+, K^- and*

$$\boldsymbol{\pi}|_{K^\pm} = \mathbf{A}^\pm \boldsymbol{\pi}_\pm. \tag{A.8}$$

Proof Fix $i = 0, \dots, L$. Because π^i is a polynomial and $\left\{\pi_+^0, \dots, \pi_+^L\right\}$ is a polynomial basis, it follows that

$$\pi^i = \sum_{j=0}^{L} a_j^i \pi_+^j, \tag{A.9}$$

for some vector $\left(a_0^i, \dots, a_L^i\right)$. Applying π^i to $\boldsymbol{x}_j^+$, and recalling (A.6), we obtain

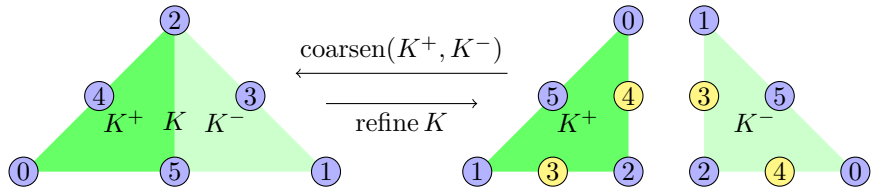
$$a_j^i = \pi^i(\boldsymbol{x}_j^+), \tag{A.10}$$

and hence

$$\pi^i = \left[\mathbf{A}^+ \boldsymbol{\pi}_+\right]^i. \tag{A.11}$$

□

A.4. **Example** (quadratic elements in 2 dimensions)**.** To make the discussion more accessible, we will illustrate it as we go with the concrete situation where $p = 2$ (quadratic elements) and $d = 2$. Following the ALBERTA conventions the relation between the coarse and fine triangles is given by the following diagram.



In this case, the coarse-on-fine matrixes are computed as follows:

$$\mathbf{A}^+ = \begin{bmatrix} 1 & & 3/8 & -1/8 \\ & & -1/8 & -1/8 \\ 1 & & & \\ & & & 1/2 \\ & & & 1/2 \quad 1 \\ & 1 & 3/4 & 1/4 \end{bmatrix}, \quad \mathbf{A}^- = \begin{bmatrix} & & -1/8 & -1/8 \\ 1 & & -1/8 & 3/8 \\ & 1 & & \\ & & 1/2 & 1 \\ & & 1/2 & \\ & 1 & 1/4 & 3/4 \end{bmatrix} \tag{A.12}$$

A.5. **Degrees of freedom and global–local relations.** Denote by U the generic finite element function in the old space $\mathbb{V}^{n-1}$ and let $V := \Lambda_0^n U$. Then we have

$$U = \mathbf{u}^\intercal \boldsymbol{\Psi} \text{ and } V = \mathbf{v}^\intercal \boldsymbol{\Phi}, \tag{A.13}$$

where $\boldsymbol{\Psi} = \left(\Psi^0, \dots, \Psi^N\right)^\intercal$ and $\boldsymbol{\Phi} = \left(\Phi^0, \dots, \Phi^M\right)^\intercal$, are the columns of nodal Lagrange piecewise polynomial bases of $\mathbb{V}^{n-1}$ and $\mathbb{V}_0^n$, respectively, and $\mathbf{u}$ and $\mathbf{v}$ are the corresponding vectors of DOF values.

There are $L+1$ degrees of freedom (DOF) per simplex, e.g., $L = 5$ for $p = 2 = d$. The simplex K in $\mathscr{T}_0^n$ comes with a *local-to-global* index relation $g = g_K^{\mathscr{T}_0^n} : [0 : L] \to [0 : M]$ whereby

$$\Phi^{g(i)}\Big|_K = \pi^i \quad \forall\, j = 0, \dots, L. \tag{A.14}$$

It follows that the finite element function V is locally represented on K by

$$Y := V|_K = \sum_{i=0}^{L} v_{g(i)} \pi^i =: \mathbf{y}^\intercal \boldsymbol{\pi}. \tag{A.15}$$

Similarly we have $g^\pm = g_{K^\pm}^{\mathscr{T}^{n-1}} : [0 : L] \to [0 : N]$ such that

$$Y^\pm := U|_{K^\pm} = \sum_{j=0}^{L} u_{g^\pm(j)} \pi_\pm^j =: \mathbf{y}^{\pm\intercal} \boldsymbol{\pi}_\pm. \tag{A.16}$$

The relation between the DOF coefficients $\mathbf{u}$ and $\mathbf{v}$ will be described next.

A.6. **Local fine–coarse DOF relations.** Some degrees of freedom—that is those depicted in yellow or bright—are removed during coarsening. The others, which are kept, have their local index change. This information is fully encoded in the *fine-to-coarse* index maps $c^\pm : D^\pm \to C^\pm$ where

$$D^\pm := \left\{ j = 0, \dots, L : \ \boldsymbol{x}_j^\pm \in \{\boldsymbol{x}_0, \dots, \boldsymbol{x}_L\} \right\}. \tag{A.17}$$

and

$$C^\pm := c^\pm(D^\pm) \subseteq [0 : L]\,. \tag{A.18}$$

A basic property of the fine-to-coarse maps is that

$$C^+ \cup C^- = [0 : L]\,, \tag{A.19}$$

but C^+ and C^- need not be disjoint (in fact, for conforming methods these are never disjoint). The fine-to-coarse maps $c^\pm$ are injective and we denote their inverses, the *coarse-to-fine* maps, by $d^\pm : C^\pm \to D^\pm$.

In the example above, $p = 2 = d$, the fine-to-coarse maps $c^\pm : D^\pm \to [0:5]$, satisfy $D^+ = D^- = \{0,1,2,5\}$ (though D^+ and D^- do not generally coincide, as seen for $p = 3$, $d = 2$, e.g.) and evaluated by the schedule

$$\begin{array}{rcccccc} j = & 0 & 1 & 2 & 3 & 4 & 5, \\ c^+(j) = & 2 & 0 & 5 & - & - & 4, \\ c^-(j) = & 1 & 2 & 5 & - & - & 3. \end{array} \tag{A.20}$$

It follows that $C^+ = \{0,2,4,5\}$ and $C^- = \{1,2,3,5\}$ and

$$\begin{array}{rcccccc} i = & 0 & 1 & 2 & 3 & 4 & 5, \\ d^+(i) = & 1 & - & 0 & - & 5 & 2, \\ d^-(i) = & - & 0 & 1 & 5 & - & 2. \end{array} \tag{A.21}$$

A.7. **Remark** (redundancy of the coarse-to-fine maps)**.** The coarse-to-fine maps $c^\pm$ and their inverses $d^\pm$ are partially redundant with $\mathbf{A}^\pm$. Namely, if $j \in D^\pm$, then $j = d^\pm(i)$ and $i = c^\pm(j)$, for some $i = 0, \dots, L$. By definition of $c^\pm$ it follows that $\boldsymbol{x}_j^\pm = \boldsymbol{x}_i$. Therefore

$$[\mathbf{A}^\pm]_j^k = \pi^k(\boldsymbol{x}_j^\pm) = \pi^k(\boldsymbol{x}_i) = \delta_i^k. \tag{A.22}$$

We have thus proved the following result that will be used to compress $\mathbf{A}^\pm$ in the sequel.

A.8. **Proposition** (redundant coarse-on-fine columns)**.** *If $j \in D^\pm$, then $\mathbf{A}^\pm$'s j-th column is described by*

$$[\mathbf{A}^\pm]_j^k = \delta^k_{c^\pm(j)}. \tag{A.23}$$

A.9. **Precomputing the coarsening error.** The coarsening error is the difference between U, to which we have access via $\mathbf{u}$, and its interpolation on the locally coarser mesh V, to which we have no direct access. Working locally at the coarsening-marked element K^+ (and similarly for K^-), all we need is to compute $V|_{K^+}$ and subtract it from $U|_{K^+}$.

Recalling that in ALBERTA $V = \Lambda_0^n U$ is built by simply "dropping" the coefficients of the DOF removed by coarsening we have

$$\mathbf{y}^\intercal \boldsymbol{\pi} = Y = V|_K = \sum_{i \in C^+} u_{g_+(d^+(i))} \pi^i + \sum_{i \in C^- \smallsetminus C^+} u_{g_-(d^-(i))} \pi^i, \tag{A.24}$$

that is, for $j = 0, \dots, L$, we set

$$v_{g(i)} := y_i := \begin{cases} u_{g_+(d^+(i))} = y^+_{d_+(i)} & \text{if } i \in C^+ \\ u_{g_-(d^-(i))} = y^-_{d_-(i)} & \text{otherwise .} \end{cases} \tag{A.25}$$

(Note that the vector $\mathbf{y}$ is the same for the two siblings $K^\pm$ and needs to be calculated only once.) Following the example with $p = 2 = d$, we see that

$$\begin{aligned} \mathbf{y} &= \left(y_1^+, y_0^-, y_0^+, y_5^-, y_5^+, y_2^+\right)^\intercal \\ &= \left(y_1^+, y_0^-, y_1^-, y_5^-, y_5^+, y_2^-\right)^\intercal. \end{aligned} \tag{A.26}$$

To conclude we rewrite the coarse basis, $\boldsymbol{\pi}$, in terms of the fine one, $\boldsymbol{\pi}_+$, using Proposition A.3 as follows:

$$V|_{K^+} = Y|_{K^+} = \mathbf{y}^\intercal\, \boldsymbol{\pi}|_{K^+} = \mathbf{y}^\intercal \mathbf{A}^+ \boldsymbol{\pi}_+. \tag{A.27}$$

Thus the coarsening error on K^+ is calculated as

$$[U - V]|_{K^+} = \mathbf{y}^{+\intercal} \boldsymbol{\pi}_+ - \mathbf{y}^\intercal \mathbf{A}^+ \boldsymbol{\pi}_+ = \boldsymbol{\pi}_+{}^\intercal \left(\mathbf{y}^+ - \mathbf{A}^{+\intercal} \mathbf{y}\right) = \sum_{j=0}^{L} \left(y_j^+ - \mathbf{y}^\intercal [\mathbf{A}^+]_j\right) \pi_j^+. \tag{A.28}$$

Recalling Proposition A.8, if $j \in D^+$ we have

$$\mathbf{y}^\intercal[\mathbf{A}^+]_j = \sum_{i=0}^{L} y_k \delta^k_{c^+(j)} = y_{c^+(j)} = y^+_j, \tag{A.29}$$

and thus the coefficient for π^+_j is 0, and it needs not be calculated. Proceeding similarly on K^- we may summarise the findings as follows.

A.10. **Theorem** (coarsening error calculation)**.** *Let $U \in \mathbb{V}^{n-1}$ with the notation of §A.5, to calculate the coarsening error that would result from coarsening the elements $K^+, K^- \in \mathscr{T}^{n-1}$ into $K \in \mathscr{T}^n$*

(1) calculate $\mathbf{y}$ following (A.25) using the coarse-to-fine map d^+ defined in §A.6,
(2) obtain the error using

$$\begin{aligned} [U - \Lambda^n_0 U]|_{K^+} &= \sum_{j \in [0:L] \setminus D^+} \left(y^+_j - \mathbf{y}^\intercal[\mathbf{A}^+]_j\right) \pi^j_+, \\ [U - \Lambda^n_0 U]|_{K^-} &= \sum_{j \in [0:L] \setminus D^-} \left(y^-_j - \mathbf{y}^\intercal[\mathbf{A}^-]_j\right) \pi^j_-. \end{aligned} \tag{A.30}$$

A.11. **Remark.** Note that the j-th coefficient of the coarsening error's local DOF vector is zero when $j \in D^\pm$, respectively. So the calculation needs to be carried out only for those $j \notin D^\pm$.

Also, the coefficients for the DOF that are common to K^+ and K^- must be equal, so they can be in fact computed once.

For example in the case of quadratic elements in $d = 2$ we have

$$\begin{aligned} Y^+ - Y|_{K^+} =& \pi^3_+ \left(y^+_3 - \frac{3}{8} y^+_1 + \frac{1}{8} y^-_0 - \frac{3}{4} y^+_2 \right) \\ &+ \pi^4_+ \left(y^+_4 + \frac{1}{8} y^+_1 + \frac{1}{8} y^-_0 - \frac{1}{4} y^+_2 - \frac{1}{2} y^+_5 - \frac{1}{2} y^-_5 \right), \\ Y^- - Y|_{K^-} =& \pi^3_- \left(y^-_3 + \frac{1}{8} y^+_1 + \frac{1}{8} y^-_0 - \frac{1}{2} y^-_5 - \frac{1}{2} y^+_5 - \frac{1}{4} y^+_2 \right) \\ &+ \pi^4_- \left(y^-_4 + \frac{1}{8} y^+_1 - \frac{3}{8} y^-_0 - \frac{3}{4} y^+_2 \right) \end{aligned} \tag{A.31}$$

A.12. **Coarsening error algorithm.** As seen in §A.9, the information needed for the coarsening error computation for Lagrange finite elements of degree p in dimension d, is contained in the coarse-on-fine matrixes $\mathbf{A}^\pm$ defined by (A.7) and the fine-to-coarse maps, $d^\pm$, and their domains $C^\pm$ defined in A.6. This information is independent of the particular pair of simplex siblings $K^\pm$ and their parent K and can be included in the code via given index permutations and efficient matrix-vector multiplication.

With this information at hand and the notation previously introduced in this section, we formulate an ALBERTA-implementable algorithm to precompute the coarsening error on all seimplexes.

Coarsening Preindicator.
Require: $(U = \mathbf{u}^\intercal \mathbf{\Phi}, \mathbb{V}, \mathscr{T})$
Ensure: $\boldsymbol{\gamma} = (\gamma_K : K \in \mathscr{T})$
  for all $K \in \mathscr{T}$ **do**
    if childorder$(K) = 0$[2] **then**

[2]The element information in ALBERTA is quite local and to determine whether an element is left or right child is not trivial. In ALBERTA 1.2 this can be done utilising `EL->index` which provides a global indexing of elements. Testing the `EL_INFO->parent->child[0]->index` against

```
        D := D^+, D' := D^-, c := c^+, c' := c^-, A := A^+
    else
        D := D^-, D' := D^+, c := c^-, c' := c^+, A := A^-
    end if
    K' := sibling K
    initialise two local DOF vectors y and r
    for all j ∈ D do
        y_{c(j)} = u_{g_K(j)}
    end for
    for all j ∈ D' do
        y_{c'(j)} = u_{g_{K'}(j)}
    end for
    for all j ∉ D ∪ D' do
        r_j = u_{g_K(j)} − y^T [A]_j
    end for
    γ_K = 0
    for all i ∉ D ∪ D' do
        for all j ∉ D ∪ D' do
            γ_K = γ_K + r_i r_j ⟨Φ_i, Φ_j⟩_K
        end for
    end for
end for
```

A.13. **Coarsening preindicator matrixes.** To close, we provide here the information needed to implement Algorithm A.12 for Lagrange piecewise $\mathbb{P}^p$ finite elements in dimension $d = 2$. (For dimension 3 the situation is complicated by the "types" of tetrahedrons, whereby the matrixes $A^{\pm}$ and the maps $c^{\pm}$ may depend on the type and is not covered in this appendix.)

A.14. $\mathbb{P}^1$ **elements.** The coarse-on-fine matrixes (omitting 0 entries for clarity) are given by

$$\mathbf{A}^+ = \begin{bmatrix} & 1 & 1/2 \\ & & 1/2 \\ 1 & & \end{bmatrix}, \mathbf{A}^- = \begin{bmatrix} & & 1/2 \\ 1 & & 1/2 \\ & 1 & \end{bmatrix}, \tag{A.32}$$

the fine-to-coarse maps and the coarse-to-fine maps are respectively given by

$$\begin{array}{rcccc} i = & 0 & 1 & 2, & \\ c^+(i) = & 2 & 0 & -, & \text{and} \\ c^-(i) = & 1 & 2 & -, & \end{array} \qquad \begin{array}{rcccc} i = & 0 & 1 & 2, & \\ d^+(i) = & 1 & - & 0, & \\ d^-(i) = & - & 0 & 1. & \end{array} \tag{A.33}$$

A.15. $\mathbb{P}^2$ **elements.** See the worked example in §A.

`EL->index` gives the correct child order of K. In ALBERTA 2.0 `EL->index` is unavailable so we check the global index of DOF for both parent and children.

A.16. $\mathbb{P}^3$ **elements.** The coarse-on-fine matrixes are given by

$$\mathbf{A}^+ = \begin{bmatrix} 1 & -1/16 & 5/16 & & 1/16 & & -1/16 \\ & -1/16 & 1/16 & & 1/16 & & 1/16 \\ 1 & & & & & & \\ & & & & -1/4 & & -1/8 \\ & & & & 1/2 & & \\ & & & & 1/2 & 1 & \\ & & & & -1/4 & 1 & 3/8 \\ & 9/16 & 15/16 & 1 & -1/16 & & 3/16 \\ & 9/16 & -5/16 & & -1/16 & & -3/16 \\ & & & 1 & 1/2 & & 3/4 \end{bmatrix} \tag{A.34}$$

and

$$\mathbf{A}^- = \begin{bmatrix} & -1/16 & 1/16 & & 1/16 & & & -1/16 \\ 1 & -1/16 & 1/16 & & 5/16 & & & \\ 1 & & & & & & & 1/16 \\ & & -1/4 & & & 1 & & \\ & & 1/2 & & & & 1 & -1/8 \\ & & 1/2 & & & & & -3/16 \\ & & -1/4 & & & & & 3/16 \\ & 9/16 & -1/16 & & -5/16 & & & 3/8 \\ & 9/16 & -1/16 & & 15/16 & 1 & & \\ & & 1/2 & 1 & & & & 3/4 \end{bmatrix} \tag{A.35}$$

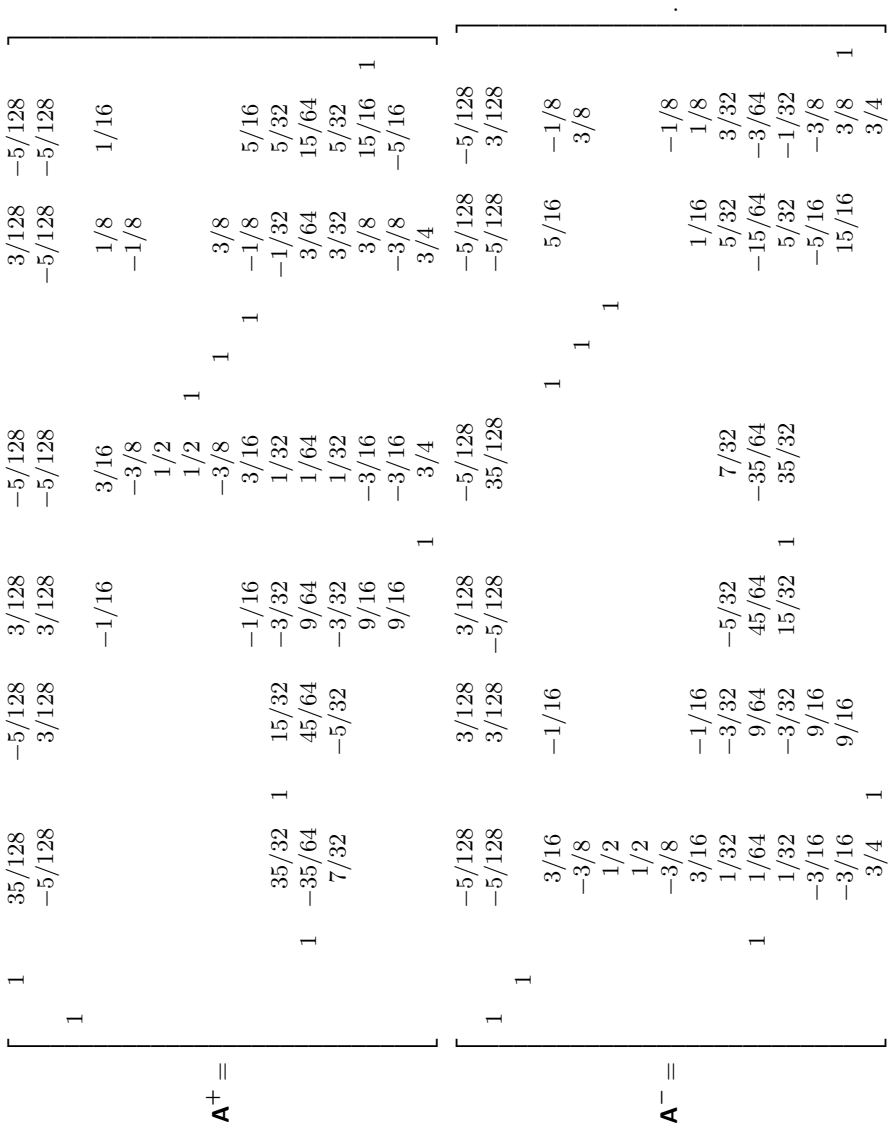
the fine-to-coarse maps

$$\begin{array}{rcccccccccc} i = & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9, \\ c^+(i) = & 2 & 0 & - & - & 7 & 9 & - & 5 & 6 & -, \\ c^-(i) = & 1 & 2 & - & - & 9 & 8 & - & 3 & 4 & -. \end{array} \tag{A.36}$$

and the coarse-to-fine maps

$$\begin{array}{rcccccccccc} i = & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9, \\ d^+(i) = & 1 & - & 0 & - & - & 7 & 8 & 4 & - & 5, \\ d^-(i) = & - & 0 & 1 & 7 & 8 & - & - & - & 5 & 4. \end{array} \tag{A.37}$$

A.17. $\mathbb{P}^4$ **elements.** The coarse-on-fine matrixes are given by



The fine-to-coarse maps are given by

$$\begin{array}{rcccccccccccccccc} i = & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14, \\ c^+(i) = & 2 & 0 & 10 & - & 9 & - & - & 14 & - & 6 & 7 & 8 & - & - & 12, \\ c^-(i) = & 1 & 2 & 10 & - & 14 & - & - & 11 & - & 3 & 4 & 5 & - & - & 13. \end{array}$$

and the coarse-to-fine maps by

$$\begin{array}{rcccccccccccccccc} i = & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14, \\ d^+(i) = & 1 & - & 0 & - & - & - & 9 & 10 & 11 & 4 & 2 & - & 14 & - & 7, \\ d^-(i) = & - & 0 & 1 & 9 & 10 & 11 & - & - & - & - & 2 & 7 & - & 14 & 4. \end{array}$$

References

[AO00] Mark Ainsworth and J. Tinsley Oden, *A posteriori error estimation in finite element analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2000. MR MR1885308 (2003b:65001)

[BDD04] Peter Binev, Wolfgang Dahmen, and Ron DeVore, *Adaptive finite element methods with convergence rates*, Numer. Math. **97** (2004), no. 2, 219–268. MR MR2050077 (2005d:65222)

[BR78] Ivo Babuška and Werner C. Rheinboldt, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal. **15** (1978), no. 4, 736–754. MR 58 #3400

[BX03a] Randolph E. Bank and Jinchao Xu, *Asymptotically exact a posteriori error estimators. I. Grids with superconvergence*, SIAM J. Numer. Anal. **41** (2003), no. 6, 2294–2312 (electronic). MR MR2034616 (2004k:65194)

[BX03b] ———, *Asymptotically exact a posteriori error estimators. II. General unstructured grids*, SIAM J. Numer. Anal. **41** (2003), no. 6, 2313–2332 (electronic). MR MR2034617 (2004m:65212)

[Car02] Carsten Carstensen, *Merging the Bramble-Pasciak-Steinbach and the Crouzeix-Thomée criterion for H^1-stability of the L^2-projection onto finite element spaces*, Math. Comp. **71** (2002), no. 237, 157–163 (electronic). MR MR1862993 (2002i:65125)

[Car04] ———, *All first-order averaging techniques for a posteriori finite element error control on unstructured grids are efficient and reliable*, Math. Comp. **73** (2004), no. 247, 1153–1165 (electronic). MR MR2047082 (2005e:65173)

[CB02] Carsten Carstensen and Sören Bartels, *Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. I. Low order conforming, nonconforming, and mixed FEM*, Math. Comp. **71** (2002), no. 239, 945–969 (electronic). MR MR1898741 (2003e:65212)

[Cia78] Philippe G. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Co., Amsterdam, 1978, Studies in Mathematics and its Applications, Vol. 4. MR 58 #25001

[CJ04] Zhiming Chen and Feng Jia, *An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems*, Math. Comp. **73** (2004), no. 247, 1167–1193 (electronic). MR MR2047083 (2005e:65131)

[DLM09] Alan Demlow, Omar Lakkis, and Charalambos Makridakis, *A posteriori error estimates in the maximum norm for parabolic problems*, SIAM Journal on Numerical Analysis **47** (2009), no. 3, 2157–2176.

[Eva98] Lawrence C. Evans, *Partial differential equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998. MR MR1625845 (99e:35001)

[FV06] Francesca Fierro and Andreas Veeser, *A posteriori error estimators, gradient recovery by averaging, and superconvergence*, Numer. Math. **103** (2006), no. 2, 267–298. MR MR2222811 (2007a:65178)

[GL08] Emmanuil Georgoulis and Omar Lakkis, *A posteriori error control for discontinuous Galerkin methods for parabolic problems*, preprint submitted to Journal, under revision 0804.4262, arXiv.org, 2008.

[KN87] Michal Křížek and Pekka Neittaanmäki, *On superconvergence techniques*, Acta Appl. Math. **9** (1987), no. 3, 175–198. MR MR900263 (88h:65208)

[LM06] Omar Lakkis and Charalambos Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp. **75** (2006), no. 256, 1627–1658 (electronic). MR MR2240628 (2007e:65122)

[LM07] Omar Lakkis and Charalambos Makridakis, *A posteriori error control for parabolic problems via elliptic reconstruction and duality*, Technical Report 0709.0916, arXiv.org, September 2007.

[LW06] Dmitriy Leykekhman and Lars Wahlbin, *A posteriori error estimates by recovered gradients in parabolic finite element equations*, Tech. report, University of Texas, Austin, 2006, Preprint (submitted to Math. Comp.).

[LZ99] Bo Li and Zhimin Zhang, *Analysis of a class of superconvergence patch recovery techniques for linear and bilinear finite elements*, Numer. Methods Partial Differential Equations **15** (1999), no. 2, 151–167. MR MR1674357 (99m:65201)

[MN03] Charalambos Makridakis and Ricardo H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal. **41** (2003), no. 4, 1585–1594 (electronic). MR MR2034895 (2004k:65157)

[MNS02] Pedro Morin, Ricardo H. Nochetto, and Kunibert G. Siebert, *Convergence of adaptive finite element methods*, SIAM Rev. **44** (2002), no. 4, 631–658 (electronic) (2003), Revised reprint of "Data oscillation and convergence of adaptive FEM" [SIAM J. Numer. Anal. **38** (2000), no. 2, 466–488 (electronic); MR1770058 (2001g:65157)]. MR MR1980447

[Pic98] Marco Picasso, *Adaptive finite elements for a linear parabolic problem*, Comput. Methods Appl. Mech. Engrg. **167** (1998), no. 3-4, 223–237. MR 2000b:65188

[Pic03] ———, *An anisotropic error indicator based on Zienkiewicz-Zhu error estimator: application to elliptic and parabolic problems*, SIAM J. Sci. Comput. **24** (2003), no. 4, 1328–1355 (electronic). MR MR1976219 (2004e:65124)

[SS05] Alfred Schmidt and Kunibert G. Siebert, *Design of adaptive finite element software*, Lecture Notes in Computational Science and Engineering, vol. 42, Springer-Verlag, Berlin, 2005, The finite element toolbox ALBERTA, With 1 CD-ROM (Unix/Linux). MR MR2127659

[SZ90] L. Ridgway Scott and Shangyou Zhang, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp. **54** (1990), no. 190, 483–493. MR MR1011446 (90j:65021)

[Ver96] RÃ¼diger VerfÃ¼rth, *A review of* a posteriori *error estimation and adaptive mesh-refinement techniques*, Wiley-Teubner, Chichester-Stuttgart, 1996.

[XZ04] Jinchao Xu and Zhimin Zhang, *Analysis of recovery type a posteriori error estimators for mildly structured grids*, Math. Comp. **73** (2004), no. 247, 1139–1152 (electronic). MR MR2047081 (2005f:65141)

[Zlá77] Miloš Zlámal, *Some superconvergence results in the finite element method*, Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975), Springer, Berlin, 1977, pp. 353–362. Lecture Notes in Math., Vol. 606. MR MR0488863 (58 #8365)

[ZW98] S. Ziukas and N.-E. Wiberg, *Adaptive procedure with superconvergent patch recovery for linear parabolic problems*, Finite element methods (Jyväskylä, 1997), Lecture Notes in Pure and Appl. Math., vol. 196, Dekker, New York, 1998, pp. 303–314. MR MR1602726

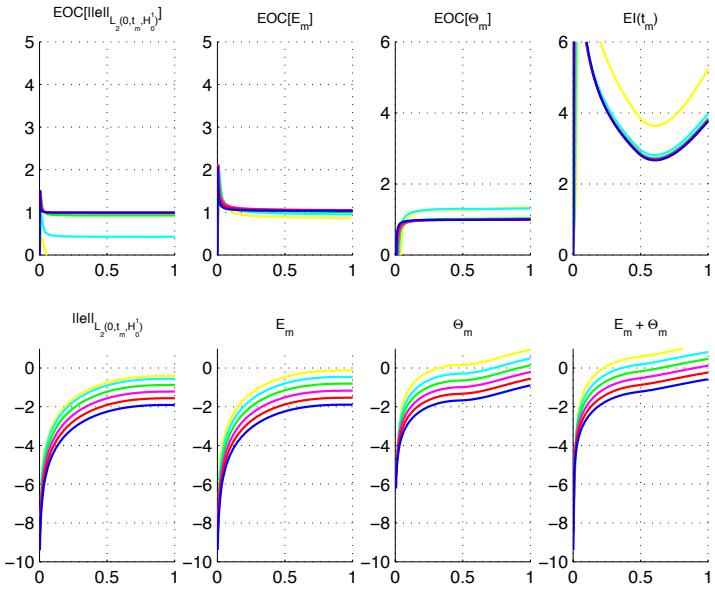
[ZZ87] Olgierd C. Zienkiewicz and J. Z. Zhu, *A simple error estimator and adaptive procedure for practical engineering analysis*, Internat. J. Numer. Methods Engrg. **24** (1987), no. 2, 337–357. MR MR875306 (87m:73055)

Department of Mathematics, University of Sussex, Falmer near Brighton, England, UK-BN1 9RF

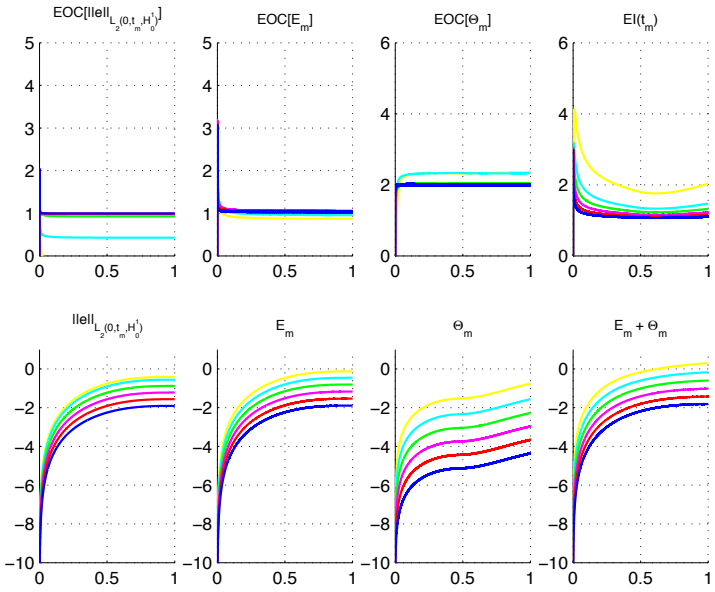
E-mail address: o.lakkis@sussex.ac.uk
URL: http://www.maths.sussex.ac.uk/Staff/OL

E-mail address: t.m.pryer@sussex.ac.uk
URL: http://www.sussex.ac.uk/maths/profile131964.html

FIGURE 1. Numerical Results for Problem (5.1) with $\mathbb{P}^1$ and $h = h(i) = 2^{-i/2}$, $i = 4, \ldots, 9$ (details in §5.5).

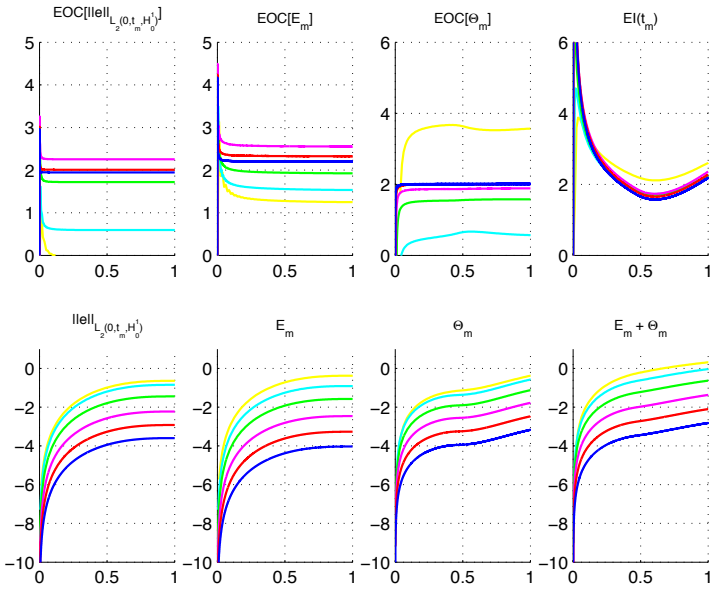


(a) Mesh-size is h and timestep $\tau = 0.1\,h$. On top we plot the EOC's of the single cumulative indicators E and Θ. Below we plot their logs. Both indicators have EOC $\to 1$, but the cumulative time error indicator Θ_m is dominant. The estimator is reliable and sharp, but not asymptotically exact and results in EI $\gg 1$.

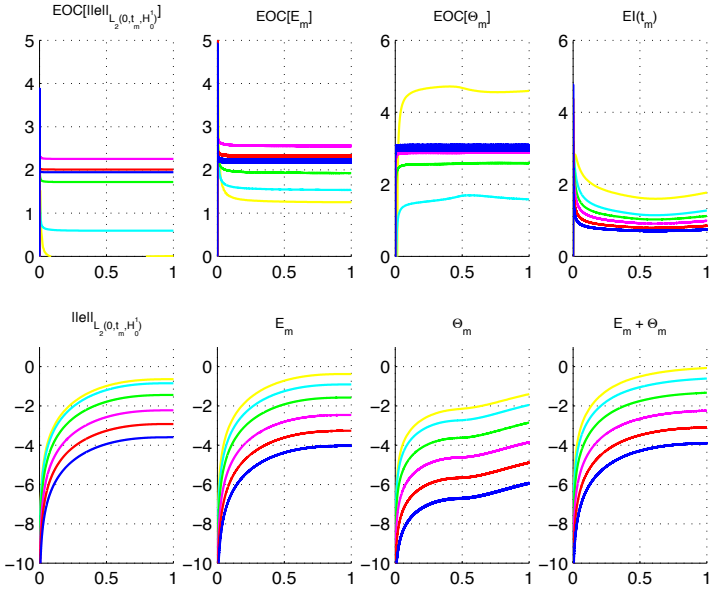


(b) Timestep is $\tau = 0.1\,h^2$. This choice leads to EOC$[\Theta_m] \to 2$ and EOC$[E_m] \approx 1$, i.e., the time indicator Θ_m is of higher order than the spatial indicator E_m which leads the estimator's order. Thus we obtain asymptotic exactness EI $\to 1$, as expected from ZZ estimators for $p = 1$.

FIGURE 2. Numerical Results for (5.1) with $\mathbb{P}^2$ elements and $h = h(i) = 2^{-i/2}$ with $i = 3, \ldots, 8$. We compute the same quantities as in Figure 1.

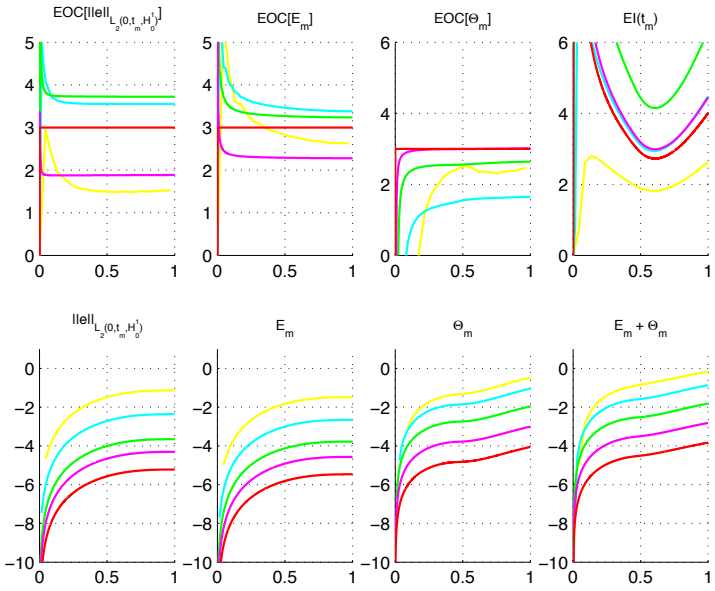


(a) Timestep $\tau = 0.1\,h^2$. The cumulative time error indicator Θ_m is dominant with $\mathrm{EOC}[\Theta_m] \to 2$, but $\mathrm{EI} \gg 1$.

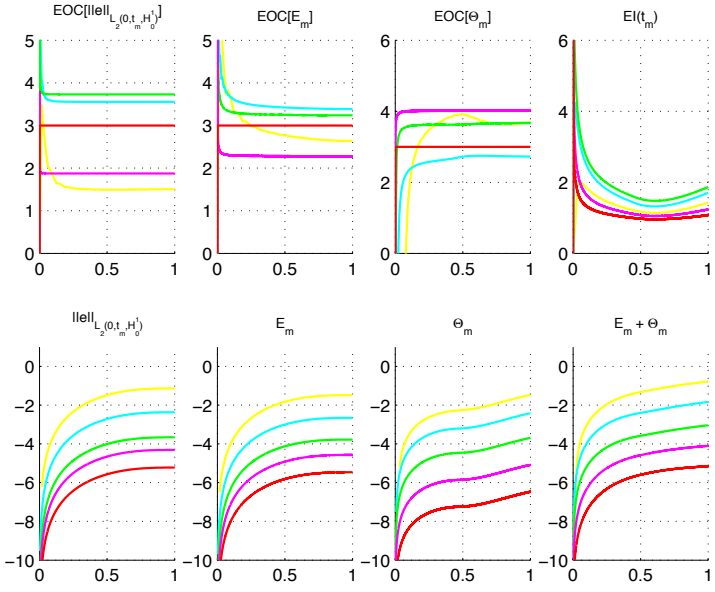


(b) Timestep is $\tau = 0.1\,h^3$, with In the bottom set of results the spatial is dominant ($\mathrm{EOC} \approx 2$) showing the estimator is sharp and reliable for higher order polynomials as well, and close to asymptotically exact (EI just smaller than 1).

FIGURE 3. Numerical Results for (5.1) with $\mathbb{P}^3$ elements for mesh-sizes $h(i) = 2^{-i/2}$, $i = 2, \ldots, 6$.. We compute the same quantities as in Figures 1 and 2.

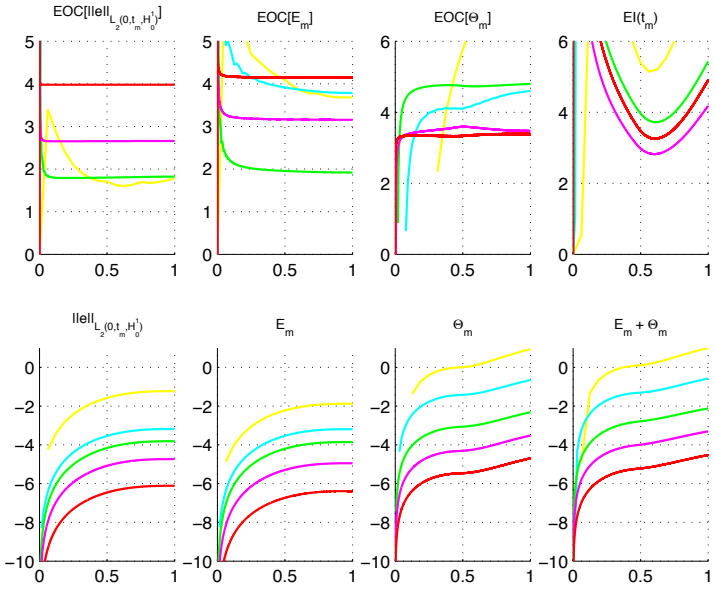


(a) Timestep is $\tau = 0.1\,h^3$. Again, the time indicator is dominant and EOC$[\Theta_m] \to 3$, but EI $\gg 1$.

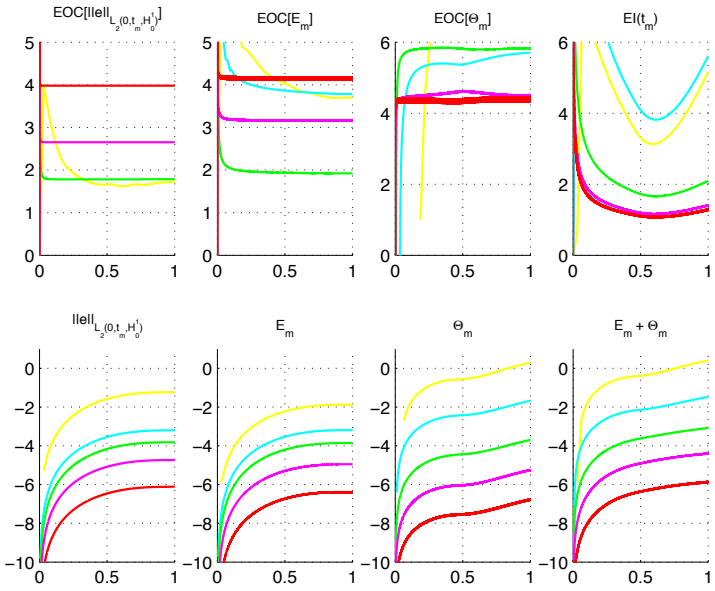


(b) Timestep is $\tau = 0.1\,h^4$. The elliptic error is dominant (EOC$[E_m] \to 3$) and the estimator is sharp and reliable with very good EI.

FIGURE 4. Results for (5.1) with $\mathbb{P}^4$ elements and $h(i) = 2^{-i/2}$, $i = 2, \ldots, 6$. We compute the same time accumulation quantities as in Figures 1–3.



(a) Mesh-size is $\tau = 0.1\, h^4$. Again, the time indicator is dominant with order $\mathrm{EOC}[\Theta_m] \to 4$) and a quite good EI in this case.



(b) Mesh-size is $\tau = 0.1h^5$. The spatial error is dominant and $\mathrm{EOC}[E_m] \to 4$. Effectivity index improves slightly over previous case.

FIGURE 5. For each $m = 1, \ldots, N$ we plot values and EOC's of two alternative time indicators $\left(\sum_{n=1}^{m} \tau_n \tilde{\theta}_n^2\right)^{1/2}$ (above) and $\left(\sum_{n=1}^{m} \tau_n \theta_n^2\right)^{1/2}$(below) and the alternative mesh-change indicator $\sum_{n=1}^{m} \tau_n \tilde{\gamma}_n^2$ (above-right). All quantities are plotted against time. We took a uniform timestep $\tau = 0.1\, h$ and mesh-size $h = 2^{-i}$, $i = 4, \ldots, 9$. The numerical results show (1) that the two time indicators are equivalent in order, as expected, and (2) that the term $\sum_{n=1}^{m} \tau_n \tilde{\gamma}_n^2$ is indeed a higher order term and can be safely ignored in most practical schemes. The indicators $\tilde{\theta}_n$ have a better effectivity index.

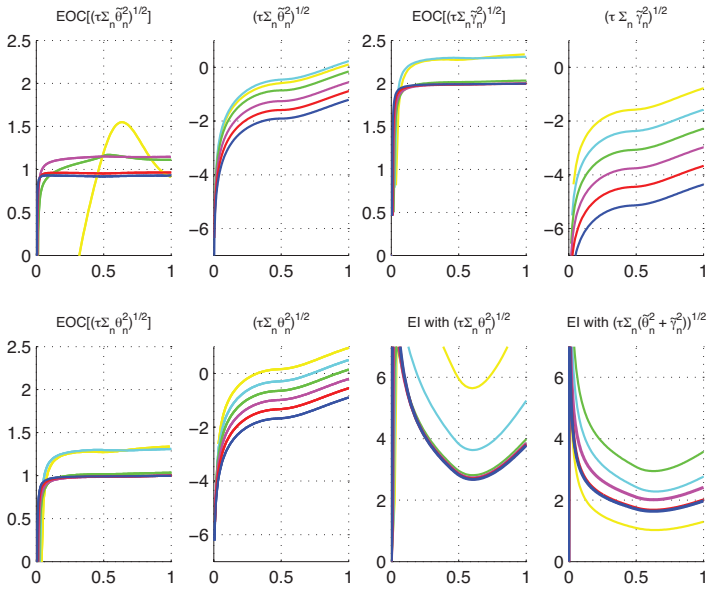
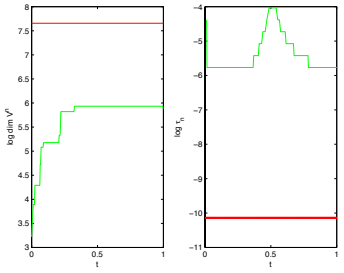
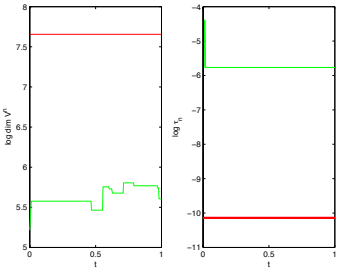


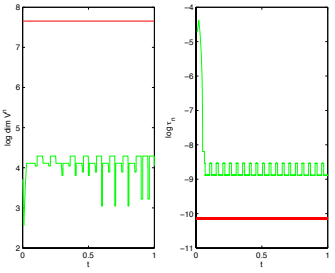
FIGURE 6. Adaptive (green) against uniform (red) degrees of freedom and timestep sizes. In each pair of graphs we plot the (log of) the DOF against time on the left, and the timestep against time on the right.



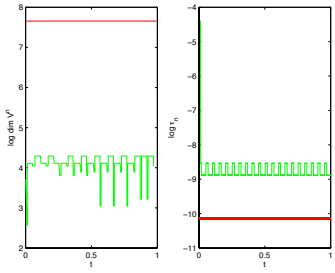
(a) Implicit timestep control for Problem (5.1). The explicit timestep control yields the same results (but is much more CPU efficient), thus it is not shown.



(b) Implicit timestep control for Problem (5.2), where the spatial error dominates. The explicit timestep control yields the same meshes and time-steps, thus not shown.



(c) Explicit timestep control for Problem (5.3), where the time discretisation error dominates. Interesting when compared with Figure 7(d).



(d) Implicit timestep control for Problem (5.3). Comparing with Figure 7(c) shows that the implicit timestep control yields more efficient timestep and meshes, but at a much higher CPU cost (cf. Tables 3 and 4).

FIGURE 7. Implicit timestep control for Problem (6.2).

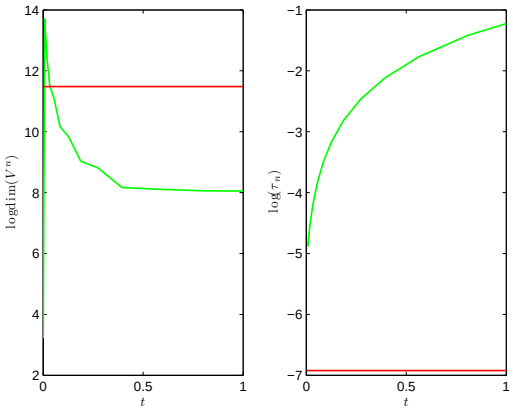
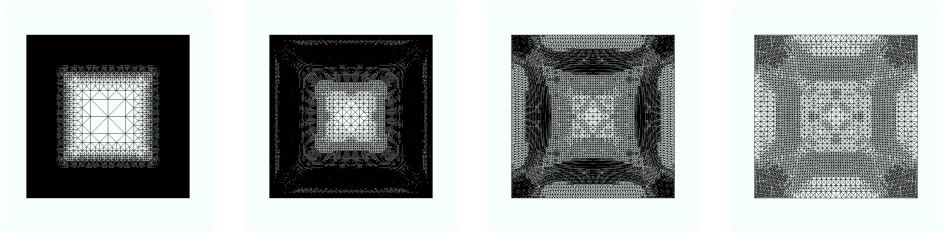


FIGURE 8. The adaptive scheme for (6.2) using implicit timestep control.



(a) Mesh at time $t_n = 0.007544$ with $\dim(\mathbb{V}^n) = 894,677$

(b) Mesh at time $t_n = 0.033302$ with $\dim(\mathbb{V}^n) = 98,773$

(c) Mesh at time $t_n = 0.127492$ with $\dim(\mathbb{V}^n) = 18,613$

(d) Mesh at time $t_n = 0.393893$ with $\dim(\mathbb{V}^n) = 3,525$