IMA Journal of Numerical Analysis (2013) **33**, 1245–1264
doi:10.1093/imanum/drs057
Advance Access publication on March 21, 2013

A posteriori $L^\infty(L^2)$-error bounds for finite element approximations to the wave equation

EMMANUIL H. GEORGOULIS*
Department of Mathematics, University of Leicester, University Road, Leicester LE1 7RH, UK
Corresponding author: emmanuil.georgoulis@le.ac.uk

OMAR LAKKIS
Department of Mathematics, University of Sussex, Falmer near Brighton, BN1 9RF, UK
o.lakkis@sussex.ac.uk

AND

CHARALAMBOS MAKRIDAKIS
Department of Applied Mathematics, University of Crete, GR-71409 Heraklion, Greece
and
Institute for Applied and Computational Mathematics, Foundation for Research and Technology-Hellas, Vasilika Vouton P.O. Box 1527, GR-71110 Heraklion, Greece
makr@tem.uoc.gr

[Received on 26 June 2012; revised on 25 November 2012]

We address the error control of Galerkin discretization (in space) of linear second-order hyperbolic problems. More specifically, we derive *a posteriori* error bounds in the $L^\infty(L^2)$ norm for finite element methods for the linear wave equation, under minimal regularity assumptions. The theory is developed for both the space-discrete case and for an implicit fully discrete scheme. The derivation of these bounds relies crucially on carefully constructed space and time reconstructions of the discrete numerical solutions, in conjunction with a technique introduced by Baker (1976, Error estimates for finite element methods for second-order hyperbolic equations. *SIAM J. Numer. Anal.*, **13**, 564–576) in the context of *a priori* error analysis of Galerkin discretization of the wave problem in weaker-than-energy spatial norms.

Keywords: *a posteriori* bounds; wave equation; implicit time stepping; reconstruction.

1. Introduction

In computing approximate solutions of evolution initial boundary value problems, mesh adaptivity plays an important role in that it drives variable resolution requirements, thereby contributing a reduction in computational cost. Adaptive strategies are often based on *a posteriori* error estimates, that is, computable quantities that estimate the error of the finite element method measured in a suitable norm (or other functionals of interest).

A posteriori error bounds are well developed for stationary boundary value problems (e.g., Dörfler, 1996; Verfürth, 1996; Ainsworth & Oden, 2000; Babuška & Strouboulis, 2001; Carstensen & Bartels, 2002; Stevenson, 2007; Cascon *et al.*, 2008; and references therein). Adaptivity and error estimation for parabolic problems has also been an active area of research for the last two decades (e.g., Eriksson & Johnson, 1995; Picasso, 1998; Houston & Süli, 2001; Makridakis & Nochetto, 2003; Verfürth, 2003; Bernardi & Verfürth, 2004; Bergam *et al.*, 2005; Lakkis & Makridakis, 2006; and references therein).

Surprisingly, there has been considerably less work on the error control of finite element methods for second-order hyperbolic problems, despite the substantial amount of research in the design of finite element methods for the wave problem (e.g., Baker, 1976; Baker & Bramble, 1979; Baker *et al.*, 1979; Baker & Dougalis, 1980; Dougalis & Serbin, 1981; Johnson, 1993; Makridakis, 1992; Bamberger *et al.*, 1990; Cohen *et al.*, 1993; Bécache *et al.*, 2000; Karakashian & Makridakis, 2005; and references therein). *A posteriori* bounds for standard implicit time-stepping finite element approximations to the linear wave equation are proposed and analysed (but only in very specific situations) by Adjerid (2002). Also, Bernardi & Süli (2005) derive rigorous *a posteriori* bounds, using energy arguments, for finite element methods with first-order implicit time stepping. Moreover, Bernardi & Süli (2005) propose an adaptive algorithm based on the *a posteriori* bounds derived therein. Goal-oriented error estimation for wave problems (via duality techniques) is also available (Bangerth & Rannacher, 1999, 2001), while some earlier work on *a posteriori* estimates for first-order hyperbolic systems is studied in the time semidiscrete setting (Makridakis & Nochetto, 2006), as well as in the fully discrete one (Johnson, 1993; Süli, 1996, 1999).

In this work, we derive *a posteriori* bounds in the $L^\infty(L^2)$ norm of the error. The theory is developed for both the space-discrete case, as well as for the practically relevant case of an implicit fully discrete scheme. The derivation of these bounds relies crucially on *reconstruction* techniques, used earlier for parabolic problems (Makridakis & Nochetto, 2003; Akrivis *et al.*, 2006; Lakkis & Makridakis, 2006). Another key tool in our analysis is the special testing procedure according to Baker (1976), who used it in the *a priori* error analysis of Galerkin discretization of the wave problem in weaker-than-energy spatial norms. It is expected (although it is not considered here) that the novel space-time reconstruction presented here could also be applicable to spatially nonlinear second-order hyperbolic problems and to different, possibly nonconforming, spatial discretizations. Moreover, it is also possible to combine the abstract results presented here with a wide class of *a posteriori* error estimators for elliptic problems.

While, for the proof of *a posteriori* bounds for the semidiscrete case, the *elliptic reconstruction* previously considered in Makridakis & Nochetto (2003) and Lakkis & Makridakis (2006) suffices, the fully discrete analysis necessitates the careful introduction of a novel space-time reconstruction, satisfying a crucial *local vanishing-moment property* in time. Our approach is based on the one-field formulation of the wave equation and, thus, nontrivial three-point time reconstructions are required. A further challenge presented by the wave equation is the special treatment of deriving bounds for the 'elliptic error' of the reconstruction framework, to obtain practically implementable residual estimators. The derived *a posteriori* estimators are formally of optimal order, that is, of the same order as the error on uniform space and time meshes.

The *a posteriori* bounds proposed in this work could be used within an adaptive algorithm, such as the one presented in Bernardi & Süli (2005). However, this is an important task in its own right and will be considered elsewhere.

The rest of this work is organized as follows. In Section 2, we present the model problem and the necessary basic definitions, along with the finite element methods for the wave equations considered in this work. In Section 3, we consider the case of *a posteriori* bounds for the space-discrete problem. In Section 4, we derive abstract *a posteriori* error bounds for the fully discrete implicit finite element method, while in Section 5, the case of *a posteriori* bounds of residual type is presented. In Section 6, we draw some final concluding remarks.

2. Preliminaries

2.1 *Model problem and notation*

We denote by $L^p(\omega)$, $1 \leqslant p \leqslant +\infty$, for $\omega \subset \mathbb{R}^d$ open, the Lebesgue spaces, with corresponding norms $\|\cdot\|_{L^p(\omega)}$. The norm of $L^2(\omega)$, denoted by $\|\cdot\|_\omega$, corresponds to the $L^2(\omega)$ inner product $\langle\cdot,\cdot\rangle_\omega$. We denote by $H^s(\omega)$, the Hilbertian Sobolev space of order $s \geqslant 0$ of real-valued functions defined on $\omega \subset \mathbb{R}^d$ (see, e.g., Adams & Fournier, 2003 for definitions and basic properties); in particular, $H^1_0(\omega)$ signifies the space of functions in $H^1(\omega)$ that vanish on the boundary $\partial\omega$ (boundary values are taken in the sense of traces). Negative-order Sobolev spaces $H^{-s}(\omega)$ for $s > 0$ are defined through duality. In the case $s = 1$, the definition of $\langle\cdot,\cdot\rangle_\omega$ is extended to the standard duality pairing between $H^{-1}(\omega)$ and $H^1_0(\omega)$. For $1 \leqslant p \leqslant +\infty$, we also define the spaces $L^p(0,T,X)$, with X being a real separable Banach space with norm $\|\cdot\|_X$, consisting of all measurable functions $v : (0,T) \to X$ for which

$$
\begin{aligned}
\|v\|_{L^p(0,T;X)} &:= \left(\int_0^T \|v(t)\|_X^p \,\mathrm{d}t\right)^{1/p} < +\infty \quad \text{for } 1 \leqslant p < +\infty,\\
\|v\|_{L^\infty(0,T;X)} &:= \operatorname*{ess\,sup}_{0\leqslant t\leqslant T} \|v(t)\|_X < +\infty \quad \text{for } p = +\infty.
\end{aligned} \tag{2.1}
$$

Let $\Omega \subset \mathbb{R}^d$ be a bounded open polygonal domain with Lipschitz boundary $\partial\Omega$. For brevity, the standard inner product on $L^2(\Omega)$ will be denoted by $\langle\cdot,\cdot\rangle$ and the corresponding norm by $\|\cdot\|$.

For time $t \in (0,T]$, we consider the linear second-order hyperbolic initial boundary value problem of finding $u \in L^2(0,T;H^1_0(\Omega))$, with $u_t \in L^2(0,T;L^2(\Omega))$ and $u_{tt} \in L^2(0,T;H^{-1}(\Omega))$ such that

$$
u_{tt} - \nabla\cdot(a\nabla u) = f \quad \text{in } (0,T)\times\Omega, \tag{2.2}
$$

where $f \in L^2(0,T;L^2(\Omega))$ and $a \in C(\bar\Omega)$ is a scalar-valued function, with $0 < \alpha_{\min} \leqslant a \leqslant \alpha_{\max}$, such that

$$
\begin{aligned}
u(x,0) &= u_0(x) \quad \text{on } \Omega\times\{0\},\\
u_t(x,0) &= u_1(x) \quad \text{on } \Omega\times\{0\},\\
u(0,t) &= 0 \quad \text{on } \partial\Omega\times(0,T],
\end{aligned} \tag{2.3}
$$

where $u_0 \in H^1_0(\Omega)$ and $u_1 \in L^2(\Omega)$. For existence and uniqueness results for this problem, we refer, for example, to Raviart & Thomas (1983, Chapter 8).

We identify a function $v \in \Omega\times[0,T] \to \mathbb{R}$ with the function $v : [0,T] \to H^1_0(\Omega)$ and we use $v(t)$ to indicate $v(\cdot,t)$.

2.2 *Finite element method*

Let $\mathcal{T}$ be a shape-regular subdivision of Ω into disjoint open simplicial or quadrilateral elements. Each element $\kappa \in \mathcal{T}$ is constructed via mappings $F_\kappa : \hat\kappa \to \kappa$, where $\hat\kappa$ is the reference simplex or reference square, so that $\bar\Omega = \bigcup_{\kappa\in\mathcal{T}} \bar\kappa$ (see, e.g., Ciarlet, 1978).

For a non-negative integer p, we denote by $\mathcal{P}_p(\hat\kappa)$ either the set of all polynomials on $\hat\kappa$ of degree p or less when $\hat\kappa$ is the simplex, or the set of polynomials of at most degree p in each variable when $\hat\kappa$ is the reference square (or cube). We consider p fixed and use the finite element space

$$V_h := \{v \in H_0^1(\Omega) : v|_\kappa \circ F_\kappa \in \mathcal{P}_p(\hat{\kappa}), \kappa \in \mathcal{T}\}. \tag{2.4}$$

Further, we denote by $\Gamma := \bigcup_{\kappa \in \mathcal{T}}(\partial\kappa \setminus \partial\Omega)$, that is, the union of all $(d-1)$-dimensional element edges (or faces) e in Ω associated with the subdivision $\mathcal{T}$ excluding the boundary. We introduce the mesh-size function $h : \Omega \to \mathbb{R}$, defined by $h(x) = \operatorname{diam}\kappa$ if $x \in \kappa$ and $h(x) = \operatorname{diam}(e)$ if $x \in e$ when e is an edge.

The semidiscrete finite element method for the initial boundary value problems (2.2, 2.3) consists in finding $U \in L^2(0,T;V_h)$ such that

$$\langle U_{tt}, V\rangle + a(U,V) = \langle f, V\rangle \quad \forall V \in L^2(0,T;V_h), \tag{2.5}$$

where the bilinear form a is defined for each $z, v \in H_0^1(\Omega)$ by

$$a(z,v) = \int_\Omega a\nabla z \cdot \nabla v \,\mathrm{d}x \tag{2.6}$$

and the corresponding energy norm is defined for $v \in H_0^1(\Omega)$ by

$$\|v\|_a = \|\sqrt{a}\nabla v\|. \tag{2.7}$$

To introduce the fully discrete implicit scheme approximating (2.2, 2.3), we consider a subdivision of the time interval $(0,T]$ into subintervals $(t^{n-1}, t^n]$, $n = 1,\ldots,N$, with $t^0 = 0$ and $t^N = T$, and we define $k_n := t^n - t^{n-1}$, the local time step. Associated with the time subdivision, let $\mathcal{T}^n$, $n = 0,\ldots,N$ be a sequence of meshes which are assumed to be *compatible* (see, e.g., Lakkis & Makridakis, 2006 for a precise definition of mesh compatibility in this context) in the sense that for any two consecutive meshes $\mathcal{T}^{n-1}$ and $\mathcal{T}^n$, $\mathcal{T}^n$ can be obtained from $\mathcal{T}^{n-1}$ by locally coarsening some of its elements and then locally refining some (possibly other) elements. The finite element space corresponding to $\mathcal{T}^n$ will be denoted by V_h^n.

We consider the fully discrete scheme for the wave problems (2.2) and (2.3):

$$\begin{aligned} &\text{for each } n = 1,\ldots,N, \text{ find } U^n \in V_h^n \text{ such that} \\ &\langle \partial^2 U^n, V\rangle + a(U^n, V) = \langle f^n, V\rangle \quad \forall V \in V_h^n, \end{aligned} \tag{2.8}$$

where $f^n := f(t^n, \cdot)$, the backward second and first finite differences

$$\partial^2 U^n := \frac{\partial U^n - \partial U^{n-1}}{k_n}, \tag{2.9}$$

with

$$\partial U^n := \begin{cases} \dfrac{U^n - U^{n-1}}{k_n} & \text{for } n = 1,2,\ldots,N, \\ V^0 := \pi^0 u_1 & \text{for } n = 0, \end{cases} \tag{2.10}$$

where $U^0 := \pi^0 u_0$ and $\pi^0 : L^2(\Omega) \to V_h^0$ is a suitable projection onto the finite element space (e.g., the orthogonal L^2-projection operator).

Denoting by A^n the stiffness matrix for the mesh $\mathcal{T}^n$, and by $\underline{U}^n$ the respective coefficient vector for U^n, the implicit method reads as follows: find $\underline{U}^n \in \mathbb{R}^{\dim V_h^n}$ such that

$$\frac{1}{k_n}\left(\frac{\underline{U}^n - \underline{U}^{n-1}}{k_n} - \frac{\underline{U}^{n-1} - \underline{U}^{n-2}}{k_{n-1}}\right) + A^n\underline{U}^n = \underline{f}^n,$$

with $\underline{f}^n := (f_i^n)_i$ and $f_i^n := (f^n, \phi_i^n)$, $i = 1, \ldots, \dim V_h^n$ for ϕ_i^n such that $V_h^n = \mathrm{span}\{\phi_i^n : i = 1, \ldots, \dim V_h^n\}$.

3. *A posteriori* error bounds for the semidiscrete problem

Here, we derive an *a posteriori* error bound for the error $\|u - U\|_{L^\infty(0,T;L^2(\Omega))}$ between the exact solution of (2.2) and (2.3) and that of the semidiscrete scheme (2.5).

Definition 3.1 (Elliptic reconstruction and error splitting) Let U be the (semidiscrete) finite element solution to problem (2.5). Let also $\Pi : L^2(\Omega) \to V_h$ be the orthogonal L^2-projection operator onto the finite element space V_h. We define the *elliptic reconstruction* $w = w(t) \in H_0^1(\Omega)$, $t \in [0, T]$ of U to be the solution of the elliptic problem

$$a(w, v) = \langle g, v\rangle \quad \forall v \in H_0^1(\Omega), \tag{3.1}$$

where

$$g := AU - \Pi f + f \tag{3.2}$$

and $A : V_h \to V_h$ is the *discrete elliptic operator* defined by

$$\langle Aq, \chi\rangle = a(q, \chi) \quad \forall q, \chi \in V_h. \tag{3.3}$$

We decompose the error as

$$e := U - u = \rho - \epsilon \quad \text{where } \epsilon := w - U \ \text{ and } \ \rho := w - u. \tag{3.4}$$

Lemma 3.2 (Error relation) With reference to the notation in (3.4), we have

$$\langle e_{tt}, v\rangle + a(\rho, v) = 0 \quad \forall v \in H_0^1(\Omega). \tag{3.5}$$

Proof. We have

$$\begin{aligned}\langle e_{tt}, v\rangle + a(\rho, v) &= \langle U_{tt}, v\rangle + a(w, v) - \langle u_{tt}, v\rangle - a(u, v)\\ &= \langle U_{tt}, v\rangle + a(w, v) - \langle f, v\rangle\\ &= \langle U_{tt}, \Pi v\rangle + a(w, v) - \langle f, v\rangle\\ &= -a(U, \Pi v) + a(w, v) + \langle \Pi f - f, v\rangle = 0,\end{aligned} \tag{3.6}$$

observing the identity $a(U, \Pi v) - \langle \Pi f - f, v\rangle = a(w, v)$ due to the construction of w. □

THEOREM 3.3 (Abstract semidiscrete error bound) With the notation introduced in (3.4), the following error bound holds:

$$\begin{aligned}\|e\|_{L^\infty(0,T;L^2(\Omega))} \leqslant \|\epsilon\|_{L^\infty(0,T;L^2(\Omega))} + \sqrt{2}(\|u_0 - U(0)\| + \|\epsilon(0)\|) \\ + 2\int_0^T \|\epsilon_t\| + C_{a,T}\|u_1 - U_t(0)\|,\end{aligned} \tag{3.7}$$

where $C_{a,T} := \min\{2T, \sqrt{2C_\Omega/\alpha_{\min}}\}$, where C_Ω is the constant of the Poincaré–Friedrichs inequality $\|v\|^2 \leqslant C_\Omega\|\nabla v\|^2$ for $v \in H_0^1(\Omega)$.

Proof. We use a testing procedure due to Baker (1976). Let $\tilde{v} : [0, T] \times \Omega \to \mathbb{R}$ with

$$\tilde{v}(t, \cdot) = \int_t^\tau \rho(s, \cdot)\,\mathrm{d}s, \quad t \in [0, T], \tag{3.8}$$

for some fixed $\tau \in [0, T]$. Clearly $\tilde{v} \in H_0^1(\Omega)$ as $\rho \in H_0^1(\Omega)$. Also, we observe that

$$\tilde{v}(\tau, \cdot) = 0,\ \nabla\tilde{v}(\tau, \cdot) = 0 \text{ and } \tilde{v}_t(t, \cdot) = -\rho(t, \cdot) \ \text{ a.e. in } [0, T]. \tag{3.9}$$

Set $v = \tilde{v}$ in (3.5), integrate between 0 and τ with respect to the variable t and integrate by parts the first term on the left-hand side to obtain

$$-\int_0^\tau \langle e_t, \tilde{v}_t\rangle + \langle e_t(\tau), \tilde{v}(\tau)\rangle - \langle e_t(0), \tilde{v}(0)\rangle + \int_0^\tau a(\rho, \tilde{v}) = 0. \tag{3.10}$$

Using (3.9), we have

$$\int_0^\tau \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\rho(t)\|^2 - \int_0^\tau \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}a(\tilde{v}(t), \tilde{v}(t)) = \int_0^\tau \langle \epsilon_t, \rho\rangle + \langle e_t(0), \tilde{v}(0)\rangle, \tag{3.11}$$

which implies

$$\frac{1}{2}\|\rho(\tau)\|^2 - \frac{1}{2}\|\rho(0)\|^2 + \frac{1}{2}a(\tilde{v}(0), \tilde{v}(0)) = \int_0^\tau \langle \epsilon_t, \rho\rangle + \langle e_t(0), \tilde{v}(0)\rangle. \tag{3.12}$$

Hence, we deduce

$$\frac{1}{2}\|\rho(\tau)\|^2 - \frac{1}{2}\|\rho(0)\|^2 + \frac{1}{2}a(\tilde{v}(0), \tilde{v}(0)) \leqslant \max_{0\leqslant t\leqslant T}\|\rho(t)\| \int_0^\tau \|\epsilon_t\| + \|e_t(0)\|\,\|\tilde{v}(0)\|. \tag{3.13}$$

Now, we select τ such that $\|\rho(\tau)\| = \max_{0\leqslant t\leqslant T}\|\rho(t)\|$ (this is possible due to the continuity of u in the time variable under the data and domain regularity assumptions above; see, e.g., Raviart & Thomas, 1983, Chapter 8) and we present two alternative, but complementary, ways to complete the proof.

In the first way, we start by observing that, for this τ, we have $\|\tilde{v}(0)\| \leqslant \|\rho(\tau)\|$, which gives

$$\frac{1}{4}\|\rho(\tau)\|^2 - \frac{1}{2}\|\rho(0)\|^2 \leqslant \left(\int_0^\tau \|\partial_t \epsilon\| + \tau\|e_t(0)\|\right)^2. \tag{3.14}$$

Using the bound $\|\rho(0)\| \leqslant \|e(0)\| + \|\epsilon(0)\|$, $e(0) = U(0) - u_0$ and $e_t(0) = U_t(0) - u_1$, and (3.14) for τ as above, we conclude that

$$\begin{aligned}\|e\|_{L^\infty(0,T;L^2(\Omega))} \leqslant &\|\epsilon\|_{L^\infty(0,T;L^2(\Omega))} + \|\rho\|_{L^\infty(0,T;L^2(\Omega))}\\ \leqslant &\|\epsilon\|_{L^\infty(0,T;L^2(\Omega))} + \sqrt{2}\,(\|u_0 - U(0)\| + \|\epsilon(0)\|)\\ &+ 2\left(\int_0^T \|\epsilon_t\| + T\|u_1 - U_t(0)\|\right).\end{aligned} \tag{3.15}$$

The second alternative, described next, consists in a different treatment of the last term on the right-hand side of (3.13). The Poincaré–Friedrichs inequality and the positivity of the diffusion coefficient a imply $\|\tilde{v}(0)\|^2 \leqslant C_\Omega \alpha_{\min}^{-1}\|\tilde{v}(0)\|_a^2$ for some constant C_Ω depending on the domain Ω only. Combining this bound with (3.13), we arrive at

$$\frac{1}{2}\|\rho(\tau)\|^2 - \frac{1}{2}\|\rho(0)\|^2 \leqslant \max_{0\leqslant t\leqslant T}\|\rho(t)\| \int_0^\tau \|\epsilon_t\| + \frac{1}{2}C_\Omega\alpha_{\min}^{-1}\|e_t(0)\|^2, \tag{3.16}$$

which implies

$$\begin{aligned}\|e\|_{L^\infty(0,T;L^2(\Omega))} \leqslant &\|\epsilon\|_{L^\infty(0,T;L^2(\Omega))} + \sqrt{2}\,(\|u_0 - U(0)\| + \|\epsilon(0)\|)\\ &+ 2\int_0^T \|\epsilon_t\| + \sqrt{2C_\Omega/\alpha_{\min}}\|u_1 - U_t(0)\|.\end{aligned} \tag{3.17}$$

Taking the minimum of the bounds (3.15) and (3.17) yields the result. □

Remark 3.4 (Short and long integration times) The use of two alternative arguments in the last step of the proof of Lemma 3.2 improves the ‘reliability constant’ $C_{a,T}$ that works for both the short-time and the long-time integration regimes.

Remark 3.5 (Completing the *a posteriori* estimation) To obtain a practical *a posteriori* bound, we need to estimate the norms involving the elliptic error ϵ. By construction, the elliptic reconstruction w is the exact solution to the elliptic boundary value problem (3.1) whose finite element solution is U. Indeed, inserting $v = V \in V_h$ in (3.1), we have

$$a(w, V) = \langle AU - \Pi f + f, V\rangle = a(U, V), \tag{3.18}$$

which implies the Galerkin orthogonality property $a(w - U, V) = 0$. Therefore, by construction, ϵ is the error of the finite element method on V_h for the elliptic problem

$$-\nabla\cdot(a\nabla w) = g, \tag{3.19}$$

with homogeneous Dirichlet boundary conditions, with g defined by (3.2).

DEFINITION 3.6 For every element face $e \subset \Gamma$, we define the *jump* across e of a field $\mathbf{w}$, defined in an open neighbourhood of e, by

$$[\![\mathbf{w}]\!](x) = \lim_{\delta\to 0}(\mathbf{w}(x+\delta\mathbf{n}_e) - \mathbf{w}(x-\delta\mathbf{n}_e))\cdot\mathbf{n}_e \tag{3.20}$$

for $x \in e$, where $\mathbf{n}_e$ denotes one of the two normal vectors to e (the definition of jump is independent of the choice).

THEOREM 3.7 (Elliptic *a posteriori* residual bounds) Let $z \in H_0^1(\Omega)$ be the solution to the elliptic problem

$$-\nabla\cdot(a\nabla z) = r, \tag{3.21}$$

$r \in L^2(\Omega)$ and Ω convex, and let $Z \in V_h$ be the finite element approximation of z satisfying

$$a(Z,V) = \langle r, V\rangle \quad \forall V \in V_h. \tag{3.22}$$

Then, there exists a positive constant C_{el}, independent of $\mathcal{T}$, h, z and Z, so that

$$\|z - Z\|^2 \leqslant C_{\mathrm{el}}\,\mathcal{E}(Z,r,\mathcal{T}), \tag{3.23}$$

where

$$\mathcal{E}(Z,r,\mathcal{T}) := \left(\sum_{\kappa\in\mathcal{T}}\left(\|h^2(r+\nabla\cdot(a\nabla Z))\|_\kappa^2 + \sum_{e\subset\Gamma}\|h^{3/2}[\![a\nabla Z]\!]\|_e^2\right)\right)^{1/2}. \tag{3.24}$$

Such results (some with various extra assumptions) are generally available in the literature. We refer to Verfürth (1996, Remark 2.4) and Ainsworth & Oden (2000, Theorem 2.7) for proofs of Theorem 3.7, and to references therein for similar approaches.

COROLLARY 3.8 (Semidiscrete residual-type *a posteriori* error bound) Assume that the hypotheses of Theorems 3.3 and 3.7 hold. Assume further that f is differentiable with respect to time. Then the following error bound holds:

$$\begin{aligned}\|e\|_{L^\infty(0,T;L^2(\Omega))} \leqslant C_{\mathrm{el}}\|\mathcal{E}(U,g,\mathcal{T})\|_{L^\infty(0,T)} + 2C_{\mathrm{el}}\int_0^T \mathcal{E}(U_t,g_t,\mathcal{T}) + \sqrt{2}C_{\mathrm{el}}\mathcal{E}(U(0),g(0),\mathcal{T})\\ + \sqrt{2}\|u_0 - U(0)\| + C_{a,T}\|u_1 - U_t(0)\|.\end{aligned} \tag{3.25}$$

Proof. Using (3.18), $\|\epsilon\|$ and $\|\epsilon_t\|$ can be bounded from above using (3.23). □

REMARK 3.9 A bound of the form (3.23) is only required to hold for Corollary 3.8 to be valid. Therefore, other available *a posteriori* bounds for elliptic problems can be also used; see, for example, Verfürth (1996), Ainsworth & Oden (2000), and references therein.

4. *A posteriori* error bounds for the fully discrete problem

The analysis of Section 3 is now extended to the case of a fully discrete implicit scheme with the aid of a novel three-point space-time reconstruction, satisfying a crucial *vanishing-moment property* in the time variable.

DEFINITION 4.1 (Space-time reconstruction) Let U^n, $n=0,\dots,N$ be the fully discrete solution computed by method (2.8), $\Pi^n : L^2(\Omega) \to V_h^n$ be the orthogonal L^2 projection and $A^n : V_h^n \to V_h^n$ be the discrete operator defined by

$$\text{for } q \in V_h^n,\ \langle A^n q, \chi \rangle = a(q, \chi) \quad \forall \chi \in V_h^n. \tag{4.1}$$

We define the *elliptic reconstruction* $w^n \in H_0^1(\Omega)$, of U^n to be the solution of the elliptic problem

$$a(w^n, v) = \langle g^n, v\rangle \quad \forall v \in H_0^1(\Omega), \tag{4.2}$$

with

$$g^n := A^n U^n - \Pi^n f^n + \bar{f}^n, \tag{4.3}$$

where $\bar{f}^0(\cdot) := f(0,\cdot)$ and $\bar{f}^n(\cdot) := k_n^{-1} \int_{t^{n-1}}^{t^n} f(t,\cdot)\, \mathrm{d}t$ for $n = 1,\dots,N$. Finally, we need to define the *elliptic reconstruction* $\partial w^0 \in H_0^1(\Omega)$, of V^0 to be the solution of the elliptic problem

$$a(\partial w^0, v) = \langle \partial g^0, v\rangle \quad \forall v \in H_0^1(\Omega), \tag{4.4}$$

with

$$\partial g^0 := A^0\, V^0 - \Pi^0 f^0 + f^0. \tag{4.5}$$

The *time reconstruction* $U : [0,T] \times \Omega \to \mathbb{R}$ of $\{U^n\}_{n=0}^N$ is defined by

$$U(t) := \frac{t - t^{n-1}}{k_n} U^n + \frac{t^n - t}{k_n} U^{n-1} - \frac{(t - t^{n-1})(t^n - t)^2}{k_n} \partial^2 U^n \tag{4.6}$$

for $t \in (t^{n-1}, t^n]$, $n = 1,\dots,N$, with $\partial^2 U^n$ given in (2.9), noting that ∂U^0 is well defined in (2.9). We note that U is a C^1 function in the time variable, with $U(t^n) = U^n$ and $U_t(t^n) = \partial U^n$ for $n = 0,1,\dots,N$.

We shall also use the time-continuous elliptic reconstruction w, defined by

$$w(t) := \frac{t - t^{n-1}}{k_n} w^n + \frac{t^n - t}{k_n} w^{n-1} - \frac{(t - t^{n-1})(t^n - t)^2}{k_n} \partial^2 w^n, \tag{4.7}$$

noting that ∂w^0 is well defined. By construction, this is also a C^1 function in the time variable.

We decompose the error as follows:

$$e := U - u = \rho - \epsilon, \text{ where } \epsilon := w - U \text{ and } \rho := w - u. \tag{4.8}$$

REMARK 4.2 (Notation overload) In this section, we use symbols, for example, U, w, e, ϵ, ρ, which are also used in Section 3, but with a slightly different meaning. Indeed, these are now fully discrete constructs, corresponding in aim and meaning, but different to their semidiscrete counterparts. It is hoped that this overload of notation should not create any confusion.

PROPOSITION 4.3 (Fully discrete error relation) For $t \in (t^{n-1}, t^n]$, $n = 1, \ldots, N$ we have

$$\langle e_{tt}, v\rangle + a(\rho, v) = \langle (I - \Pi^n)U_{tt}, v\rangle + \mu^n \langle \partial^2 U^n, \Pi^n v\rangle + a(w - w^n, v) + \langle \bar{f}^n - f, v\rangle \tag{4.9}$$

for all $v \in H_0^1(\Omega)$, with I being the identity mapping in $L^2(\Omega)$, and

$$\mu^n(t) := -6k_n^{-1}(t - t^{n-1/2}), \tag{4.10}$$

where $t^{n-1/2} := \frac{1}{2}(t^n + t^{n-1})$.

Proof. Noting that $U_{tt}(t) = (1 + \mu^n(t))\partial^2 U^n$ for $t \in (t^{n-1}, t^n]$, $n = 1, \ldots, N$, and the identity $a(U^n, \Pi^n v) - \langle \Pi^n f^n - \bar{f}^n, v\rangle = a(w^n, v)$, we deduce

$$\begin{aligned}\langle e_{tt}, v\rangle + a(\rho, v) &= \langle U_{tt}, v\rangle + a(w, v) - \langle f, v\rangle \\ &= \langle (I - \Pi^n)U_{tt}, v\rangle + \langle U_{tt}, \Pi^n v\rangle + a(w, v) - \langle f, v\rangle \\ &= \langle (I - \Pi^n)U_{tt}, v\rangle + \mu^n(t)\langle \partial^2 U^n, \Pi^n v\rangle - a(U^n, \Pi^n v) + a(w, v) + \langle \Pi^n f^n - f, v\rangle \\ &= \langle (I - \Pi^n)U_{tt}, v\rangle + \mu^n(t)\langle \partial^2 U^n, \Pi^n v\rangle + a(w - w^n, v) + \langle \bar{f}^n - f, v\rangle.\end{aligned} \tag{4.11}$$

□

REMARK 4.4 (Vanishing-moment property) The particular form of the remainder $\mu^n(t)$ satisfies the vanishing-moment property

$$\int_{t^{n-1}}^{t^n} \mu^n(t)\,\mathrm{d}t = 0, \tag{4.12}$$

which appears to be of crucial importance for the optimality of the *a posteriori* bounds presented below.

DEFINITION 4.5 (*A posteriori* error indicators) We define in list form the error indicators that will form error estimator the fully discrete bounds in Theorem 4.6:

- Mesh change indicator $\eta_1(\tau) := \eta_{1,1}(\tau) + \eta_{1,2}(\tau)$, with

$$\eta_{1,1}(\tau) := \sum_{j=1}^{m-1} \int_{t^{j-1}}^{t^j} \|(I - \Pi^j)U_t\| + \int_{t^{m-1}}^{\tau} \|(I - \Pi^m)U_t\| \tag{4.13}$$

  and

$$\eta_{1,2}(\tau) := \sum_{j=1}^{m-1} (\tau - t^j)\|(\Pi^{j+1} - \Pi^j)\partial U^j\| + \tau\|(I - \Pi^0)V^0(0)\|; \tag{4.14}$$

- evolution error indicator

$$\eta_2(\tau) := \int_0^\tau \|\mathcal{G}\|, \tag{4.15}$$

  where $\mathcal{G} : (0, T] \to \mathbb{R}$ with $\mathcal{G}|_{(t^{j-1}, t^j]} := \mathcal{G}^j$, $j = 1, \ldots, N$ and

$$\mathcal{G}^j(t) := \frac{(t^j - t)^2}{2}\partial g^j - \left(\frac{(t^j - t)^4}{4k_j} - \frac{(t^j - t)^3}{3}\right)\partial^2 g^j - \gamma_j, \tag{4.16}$$

  with g^j as in Definition 4.1 and $\gamma_j := \gamma_{j-1} + (k_j^2/2)\partial g^j + (k_j^3/12)\partial^2 g^j$, $j = 1, \ldots, N$, with $\gamma_0 = 0$;

- data error indicator

$$\eta_3(\tau) := \frac{1}{2\pi}\sum_{j=1}^{m-1}\left(\int_{t^{j-1}}^{t^j} k_j^3\|\bar{f}^j - f\|^2\right)^{1/2} + \left(\int_{t^{m-1}}^{\tau} k_m^3\|\bar{f}^m - f\|^2\right)^{1/2}; \tag{4.17}$$

- time reconstruction error indicator

$$\eta_4(\tau) := \frac{1}{2\pi}\sum_{j=1}^{m-1}\left(\int_{t^{j-1}}^{t^j} k_j^3\|\mu^j\partial^2 U^j\|^2\right)^{1/2} + \left(\int_{t^{m-1}}^{\tau} k_m^3\|\mu^m\partial^2 U^m\|^2\right)^{1/2}. \tag{4.18}$$

THEOREM 4.6 (Abstract fully discrete error bound) Recalling the notation of Definition 4.1 and the indicators of Definition 4.5, we have the bound

$$\begin{aligned}\|e\|_{L^\infty(0,t^N;L^2(\Omega))} &\leqslant \|\epsilon\|_{L^\infty(0,t^N;L^2(\Omega))} + \sqrt{2}\,(\|u_0 - U(0)\| + \|\epsilon(0)\|)\\ &\quad + 2\left(\int_0^{t^N}\|\epsilon_t\| + \sum_{i=1}^{4}\eta_i(t^N)\right) + C_{a,N}\|u_1 - V^0\|,\end{aligned} \tag{4.19}$$

where $C_{a,N} := \min\{2t^N, \sqrt{2C_\Omega/\alpha_{\min}}\}$.

The proof of Theorem 4.6 is the content of the remainder of this section.

Next, we set $v = \tilde{v}$ in (4.9) with $\tilde{v}$ defined by (3.8) where ρ is defined as in (4.8) (i.e., the fully discrete ρ), assuming that $t^{m-1} < \tau \leqslant t^m$ for some integer m with $1 \leqslant m \leqslant N$. We integrate the resulting equation with respect to t between 0 and τ, to arrive at

$$\int_0^\tau \langle e_{tt}, \tilde{v}\rangle + \int_0^\tau a(\rho, \tilde{v}) = \mathcal{I}_1(\tau) + \mathcal{I}_2(\tau) + \mathcal{I}_3(\tau) + \mathcal{I}_4(\tau), \tag{4.20}$$

where

$$\begin{aligned}\mathcal{I}_1(\tau) &:= \sum_{j=1}^{m-1}\int_{t^{j-1}}^{t^j}\langle (I - \Pi^j)U_{tt}, \tilde{v}\rangle + \int_{t^{m-1}}^{\tau}\langle (I - \Pi^m)U_{tt}, \tilde{v}\rangle,\\ \mathcal{I}_2(\tau) &:= \sum_{j=1}^{m-1}\int_{t^{j-1}}^{t^j} a(w - w^j, \tilde{v}) + \int_{t^{m-1}}^{\tau} a(w - w^m, \tilde{v})\\ \mathcal{I}_3(\tau) &:= \sum_{j=1}^{m-1}\int_{t^{j-1}}^{t^j}\langle \bar{f}^j - f, \tilde{v}\rangle + \int_{t^{m-1}}^{\tau}\langle \bar{f}^m - f, \tilde{v}\rangle,\\ \mathcal{I}_4(\tau) &:= \sum_{j=1}^{m-1}\int_{t^{j-1}}^{t^j}\mu^j\langle \partial^2 U^j, \Pi^j\tilde{v}\rangle + \int_{t^{m-1}}^{\tau}\mu^m\langle \partial^2 U^m, \Pi^m\tilde{v}\rangle.\end{aligned} \tag{4.21}$$

In Lemmas 4.7–4.9 and 4.11, we will derive bounds of the form

$$\mathcal{I}_i(\tau) \leqslant \eta_i(\tau)\max_{0\leqslant t\leqslant T}\|\rho(t)\|, \tag{4.22}$$

for $i = 1, 2, 3, 4$. With the help of these, we will conclude the proof of Theorem 4.6 at the end of this section.

LEMMA 4.7 (Mesh change error estimate) Under the assumptions of Theorem 4.6 and with the notation (4.21), we have

$$\mathcal{I}_1(\tau) \leqslant \eta_1(\tau) \max_{0\leqslant t\leqslant T} \|\rho(t)\|. \tag{4.23}$$

Proof. Observing that the projections $\Pi^j, j = 1, \ldots, N$ commute with time differentiation, we integrate by parts with respect to t, arriving at

$$\begin{aligned}\mathcal{I}_1(\tau) = &\sum_{j=1}^{m-1} \int_{t^{j-1}}^{t^j} \langle (I - \Pi^j)U_t, \rho\rangle + \int_{t^{m-1}}^{\tau} \langle (I - \Pi^m)U_t, \rho\rangle \\ &+ \sum_{j=1}^{m-1} \langle (\Pi^{j+1} - \Pi^j)U_t(t^j), \tilde{v}(t^j)\rangle - \langle (I - \Pi^0)U_t(0), v(0)\rangle.\end{aligned} \tag{4.24}$$

The first two terms on the right-hand side of (4.24) are bounded by

$$\max_{0\leqslant t\leqslant T} \|\rho(t)\| \left(\sum_{j=1}^{m-1} \int_{t^{j-1}}^{t^j} \|(I - \Pi^j)U_t\| + \int_{t^{m-1}}^{\tau} \|(I - \Pi^m)U_t\| \right). \tag{4.25}$$

Recalling the definition of $\tilde{v}$ and that of $U(t^j) = \partial U^j, j = 0, 1, \ldots, N$, we can bound the last two terms on the right-hand side of (4.24) by

$$\max_{0\leqslant t\leqslant T} \|\rho(t)\| \left(\sum_{j=1}^{m-1} (\tau - t^j)\|(\Pi^{j+1} - \Pi^j)\partial U^j\| + \tau\|(I - \Pi^0)V^0(0)\| \right). \tag{4.26}$$

□

LEMMA 4.8 (Evolution error bound) Under the assumptions of Theorem 4.6 and with the notation (4.21), we have

$$\mathcal{I}_2(\tau) \leqslant \eta_2(\tau) \max_{0\leqslant t\leqslant T} \|\rho(t)\|. \tag{4.27}$$

Proof. First, we observe the identity

$$w - w^j = -(t^j - t)\partial w^j + \left(k_j^{-1}(t^j - t)^3 - (t^j - t)^2\right) \partial^2 w^j \tag{4.28}$$

on each $(t^{j-1}, t^j], j = 2, \ldots, m$. Hence, from Definition 4.1, we deduce

$$a(w - w^j, \tilde{v}) = \langle -(t^j - t)\partial g^j + (k_j^{-1}(t^j - t)^3 - (t^j - t)^2)\partial^2 g^j, \tilde{v}\rangle. \tag{4.29}$$

The integral of the first component in the inner product on the right-hand side of (4.29) with respect to t between t^{j-1} and t^j is then given by $\mathcal{G}$. The choice of constants in $\mathcal{G}$ implies that $\mathcal{G}$ is continuous on $t^j, j = 1, 2, \ldots, N$ and $\mathcal{G}(0) = 0$.

Hence, integrating by parts on each interval $(t^{j-1}, t^j]$, $j = 1, \ldots, m$, we obtain

$$\mathcal{I}_2(\tau) = \int_0^\tau \langle \mathcal{G}, \rho \rangle, \tag{4.30}$$

which now implies the result. □

Lemma 4.9 (Data approximation error bound) Under the assumptions of Theorem 4.6 and with the notation (4.21), we have

$$\mathcal{I}_3(\tau) \leqslant \eta_3(\tau) \max_{0\leqslant t\leqslant T} \|\rho(t)\|. \tag{4.31}$$

Proof. We begin by observing that

$$\int_{t^{j-1}}^{t^j} (\bar{f}^j - f) = 0 \tag{4.32}$$

for all $j = 1, \ldots, m-1$. Hence, we have

$$\sum_{j=1}^{m-1} \int_{t^{j-1}}^{t^j} \langle \bar{f}^j - f, \tilde{v} \rangle = \sum_{j=1}^{m-1} \int_{t^{j-1}}^{t^j} \langle \bar{f}^j - f, \tilde{v} - \bar{\tilde{v}}^j \rangle, \tag{4.33}$$

where $\bar{\tilde{v}}^j(\cdot) := k_j^{-1} \int_{t^{j-1}}^{t^j} \tilde{v}(t, \cdot)\,\mathrm{d}t$. Using the inequality

$$\int_{t^{j-1}}^{t^j} \|\tilde{v} - \bar{\tilde{v}}^j\|^2 \leqslant \frac{k_j^2}{4\pi^2} \int_{t^{j-1}}^{t^j} \|\tilde{v}_t\|^2 \tag{4.34}$$

and recalling that $\tilde{v}_t = \rho$, we have

$$\begin{aligned}
\sum_{j=1}^{m-1} \int_{t^{j-1}}^{t^j} \langle \bar{f}^j - f, \tilde{v} \rangle &\leqslant \sum_{j=1}^{m-1} \left(\int_{t^{j-1}}^{t^j} \|\bar{f}^j - f\|^2 \right)^{1/2} \left(\int_{t^{j-1}}^{t^j} \|\tilde{v} - \bar{\tilde{v}}^j\|^2 \right)^{1/2} \\
&\leqslant \frac{1}{2\pi} \sum_{j=1}^{m-1} \left(\int_{t^{j-1}}^{t^j} \|\bar{f}^j - f\|^2 \right)^{1/2} \left(\int_{t^{j-1}}^{t^j} k_j^2 \|\rho\|^2 \right)^{1/2} \\
&\leqslant \frac{1}{2\pi} \sum_{j=1}^{m-1} \left(\int_{t^{j-1}}^{t^j} k_j^3 \|\bar{f}^j - f\|^2 \right)^{1/2} \max_{0\leqslant t\leqslant T} \|\rho(t)\|.
\end{aligned} \tag{4.35}$$

For the remaining term in $\mathcal{I}_3$, we first observe that

$$\int_{t^{m-1}}^{\tau} \|\tilde{v}\|^2 \,\mathrm{d}t \leqslant \int_{t^{m-1}}^{\tau} k_m \int_t^\tau \|\rho\|^2 \,\mathrm{d}s\,\mathrm{d}t \leqslant k_m^3 \max_{0\leqslant s\leqslant T} \|\rho(t)\|^2, \tag{4.36}$$

which implies

$$\int_{t^{m-1}}^{\tau} \langle \bar{f}^m - f, \tilde{v} \rangle \leqslant \left(\int_{t^{m-1}}^{\tau} k_m^3 \|\bar{f}^m - f\|^2 \right)^{1/2} \max_{0\leqslant t\leqslant T} \|\rho(t)\|. \tag{4.37}$$

Recalling η_3 from Definition 4.5, we conclude the proof. □

REMARK 4.10 (The order of the data approximation indicator) The choice of the particular combination of functions involving the right-hand-side data f in the definition of g^n in the elliptic reconstruction results in the property (4.32). When f is differentiable, we have $\eta_3(\tau) = \mathcal{O}(k^2)$ as $k := \max_{1\leqslant j\leqslant m} k_j \to 0$, and the convergence is of second order with respect to the maximum time step. In this case, η_3 is, therefore, a higher-order term.

LEMMA 4.11 (Time-reconstruction error bound) Under the assumptions of Theorem 4.6 and with the notation (4.21), we have

$$\mathcal{I}_4(\tau) \leqslant \eta_4(\tau) \max_{0\leqslant t\leqslant T} \|\rho(t)\|. \tag{4.38}$$

Proof. The method of bounding $\mathcal{I}_4(\tau)$ is similar to that of Lemma 4.9, so we shall highlight only the differences.

Recalling the vanishing-moment property (4.12) and noting that $\partial^2 U^j$ is piecewise constant in time, we have

$$\sum_{j=1}^{m-1} \int_{t^{j-1}}^{t^j} \mu^j \langle \partial^2 U^j, \Pi^j \tilde{v} \rangle = \sum_{j=1}^{m-1} \int_{t^{j-1}}^{t^j} \mu^j \langle \partial^2 U^j, \Pi^j (\tilde{v} - \bar{\tilde{v}}^j) \rangle, \tag{4.39}$$

where $\bar{\tilde{v}}^j(\cdot) = k_j^{-1} \int_{t^{j-1}}^{t^j} \tilde{v}(t,\cdot)\,\mathrm{d}t$. Hence, since Π^j commutes with time integration, we obtain

$$\begin{aligned} \sum_{j=1}^{m-1} \int_{t^{j-1}}^{t^j} \mu^j \langle \partial^2 U^j, \Pi^j (\tilde{v} - \bar{\tilde{v}}^j) \rangle &\leqslant \frac{1}{2\pi} \sum_{j=1}^{m-1} \left(\int_{t^{j-1}}^{t^j} \|\mu^j \partial^2 U^j\|^2 \right)^{1/2} \left(\int_{t^{j-1}}^{t^j} k_j^2 \|\Pi^j \rho\|^2 \right)^{1/2} \\ &\leqslant \frac{1}{2\pi} \sum_{j=1}^{m-1} \left(\int_{t^{j-1}}^{t^j} k_j^3 \|\mu^j \partial^2 U^j\|^2 \right)^{1/2} \max_{0\leqslant t\leqslant T} \|\rho(t)\|. \end{aligned} \tag{4.40}$$

For the remaining term in $\mathcal{I}_4$, upon using an argument similar to (4.36), we have

$$\int_{t^{m-1}}^{\tau} \langle \mu^m \partial^2 U^m, \Pi^m \tilde{v} \rangle \leqslant \left(\int_{t^{m-1}}^{\tau} k_m^3 \|\mu^m \partial^2 U^m\|^2 \right)^{1/2} \max_{0\leqslant t\leqslant T} \|\rho(t)\|. \tag{4.41}$$

Recalling the definition of η_4 in Definition 4.5, we conclude. □

Starting from (4.20), integrating by parts the first term on the left-hand side, and using the properties of $\tilde{v}$, we arrive to

$$\int_0^\tau \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\rho\|^2 - \int_0^\tau \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t} a(\tilde{v},\tilde{v}) = \int_0^\tau \langle \epsilon_t, \rho \rangle + \langle e_t(0), \tilde{v}(0) \rangle + \sum_{i=1}^4 \mathcal{I}_i(\tau), \tag{4.42}$$

which implies

$$\frac{1}{2}\|\rho(\tau)\|^2 - \frac{1}{2}\|\rho(0)\|^2 + \frac{1}{2} a(\tilde{v}(0),\tilde{v}(0)) = \int_0^\tau \langle \epsilon_t, \rho \rangle + \langle e_t(0), \tilde{v}(0) \rangle + \sum_{i=1}^4 \mathcal{I}_i(\tau). \tag{4.43}$$

Hence, we deduce

$$\frac{1}{2}\|\rho(\tau)\|^2 - \frac{1}{2}\|\rho(0)\|^2 + \frac{1}{2}a(\tilde{v}(0),\tilde{v}(0)) \leqslant \max_{0\leqslant t\leqslant T}\|\rho(t)\|\left(\int_0^\tau \|\epsilon_t\| + \sum_{i=1}^4 \eta_i(\tau)\right) + \|e_t(0)\|\,\|\tilde{v}(0)\|. \tag{4.44}$$

We select $\tau = \hat{\tau}$ such that $\|\rho(\hat{\tau})\| = \max_{0\leqslant t\leqslant t^N}\|\rho(t)\|$. First, observing that $\|\tilde{v}(0)\| \leqslant \tau\|\rho(\hat{\tau})\|$, gives

$$\frac{1}{4}\|\rho(\tau)\|^2 - \frac{1}{2}\|\rho(0)\|^2 \leqslant \left(\int_0^\tau \|\epsilon_t\| + \sum_{i=1}^4 \eta_i(\tau) + \tau\|e_t(0)\|\right)^2. \tag{4.45}$$

Using the bound $\|\rho(0)\| \leqslant \|e(0)\| + \|\epsilon(0)\|$ and observing that $e(0) = \hat{U}(0) - u(0) = U^0 - u_0$ and that $e_t(0) = \hat{U}_t(0) - u_t(0) = V^0 - u_1$, we arrive at

$$\|e\|_{L^\infty(0,t^N;L^2(\Omega))} \leqslant \|\epsilon\|_{L^\infty(0,t^N;L^2(\Omega))} + \sqrt{2}\left(\|u_0 - U^0\| + \|\epsilon(0)\|\right) + 2\left(\int_0^{t^N} \|\epsilon_t\| + \sum_{i=1}^4 \eta_i(t^N) + t^N\|u_1 - V^0\|\right). \tag{4.46}$$

The second way is completely analogous to the proof of the semidiscrete case. Combining the bounds above suffices to conclude the proof of Theorem 4.6.

5. Fully discrete *a posteriori* estimates of residual type

To arrive at a practical *a posteriori* bound for the fully discrete scheme from Theorem 4.6, the quantities involving the elliptic error ϵ should be estimated in an *a posteriori* fashion: this is the content of Lemmas 5.1 and 5.3, where residual-type *a posteriori* estimates are used.

LEMMA 5.1 (Estimation of the elliptic error) With the notation introduced in Definition 4.1, we have

$$\|\epsilon\|_{L^\infty(0,t^N;L^2(\Omega))} + \sqrt{2}\|\epsilon(0)\| \leqslant \delta_1(t^N) + \sqrt{2}C_{\mathrm{el}}\mathcal{E}^0, \tag{5.1}$$

where

$$\delta_1(t^N) := \max\left\{\frac{8k_1}{27}C_{\mathrm{el}}\mathcal{E}(V^0,\partial g^0,\mathcal{T}^0), \left(\frac{35}{27} + \frac{31}{27}\max_{1\leqslant j\leqslant N}\frac{k_j}{k_{j-1}}\right)\max_{0\leqslant j\leqslant N}\left(C_{\mathrm{el}}\mathcal{E}^j + C_\Omega\alpha_{\min}^{-1}\|\bar{f}^j - f^j\|\right)\right\}, \tag{5.2}$$

with $\mathcal{E}^j := \mathcal{E}(U^j, A^jU^j - \Pi^jf^j + f^j, \mathcal{T}^j)$, $j = 0, 1, \ldots, N$.

Proof. For $t \in (t^{j-1}, t^j]$, $j = 1, \ldots, N$, we have

$$\epsilon(t) = \frac{t - t^{j-1}}{k_j}(w^j - U^j) + \frac{t^j - t}{k_j}(w^{j-1} - U^{j-1}) - \frac{(t - t^{j-1})(t^j - t)^2}{k_j}(\partial^2 w^j - \partial^2 U^j), \tag{5.3}$$

from which, we can deduce

$$\|\epsilon(t)\| \leqslant \max\left\{\left(\frac{35}{27} + \frac{31}{27}\max_{1\leqslant j\leqslant N}\frac{k_j}{k_{j-1}}\right)\max_{0\leqslant j\leqslant N}\|w^j - U^j\|, \frac{8k_1}{27}\|\partial w^0 - V^0\|\right\}, \tag{5.4}$$

noting that

$$\max_{t\in(t^{j-1},t^j]}\frac{(t-t^{j-1})(t^j-t)^2}{k_j} = \frac{4k_j^2}{27}. \tag{5.5}$$

It remains to estimate the terms $\|w^j - U^j\|$ and $\|\partial w^0 - V^0\|$. To this end, recalling the notation of Definition 4.1, we define $w_*^j \in H_0^1(\Omega)$ to be the solution of the elliptic problem

$$a(w_*^j, v) = \langle A^j U^j - \Pi^j f^j + f^j, v\rangle \quad \forall v \in H_0^1(\Omega), \tag{5.6}$$

for $j = 0, 1, \ldots, N$. Note that, due to the fact that $\bar{f}^0 = f^0$, we have $w_*^0 = w^0$. By construction, we have $a(w_*^j, V) = \langle A^j U^j - \Pi^j f^j + f^j, V\rangle = a(U^j, V)$ for all $V \in V_h^j$, $j = 0, 1, \ldots, N$. Hence, U^j is the finite element solution (in V_h^j) of the elliptic boundary value problem (5.6). In view of Theorem 3.7, this implies that

$$\|w_*^j - U^j\| \leqslant C_{\mathrm{el}}\mathcal{E}^j \tag{5.7}$$

for $j = 0, \ldots, N$. Similarly, by construction, we have $a(\partial w^0, V) = \langle A^0 V^0 - \Pi^0 f^0 + f^0, V\rangle = a(V^0, V)$ for all $V \in V_h^0$. Hence,

$$\|\partial w^0 - \partial U^0\| \leqslant C_{\mathrm{el}}\mathcal{E}(V^0, \partial g^0, \mathcal{T}^0). \tag{5.8}$$

Moreover, since $w^j - w_*^j$ is the solution of an elliptic problem with $\bar{f}^j - f^j$, standard elliptic stability results yield

$$\|w^j - w_*^j\| \leqslant C_\Omega \alpha_{\min}^{-1}\|\bar{f}^j - f^j\| \tag{5.9}$$

for $j = 1, \ldots, N$. Finally, using the triangle inequality

$$\|w^j - U^j\| \leqslant \|w^j - w_*^j\| + \|w_*^j - U^j\|, \tag{5.10}$$

along with the bounds (5.9), (5.8) and (5.7), now implies the result. □

REMARK 5.2 The bound (5.1) contains both the *elliptic estimators* $\mathcal{E}(\cdot,\cdot,\cdot)$ and the data oscillation terms $\|\bar{f}^j - f^j\|$ which are, in general, of first order with respect to the time step. The data oscillation terms are expected to dominate the data error indicator η_3 (cf. Remark 4.10). On the other hand, if the numerical scheme (2.8) is altered so that $f^j = \bar{f}^j$ (as in, e.g., Baker, 1976), then the data oscillation terms in (5.1) vanish. Similar remarks apply to the result of Lemma 5.3.

For each $n = 1, \ldots, N$, we denote by $\hat{\mathcal{T}}^n$ the finest common coarsening of $\mathcal{T}^n$ and $\mathcal{T}^{n-1}$, and by $\hat{V}_h^n := V_h^n \cap V_h^{n-1}$ the corresponding finite element space, along with the orthogonal L^2-projection operator $\hat{\Pi}^n : L^2(\Omega) \to \hat{V}_h^n$.

LEMMA 5.3 (Estimation of the time derivative of the elliptic error) With the notation introduced in Definition 4.1, we have

$$\int_0^{t^N} \|\epsilon_t\| \leqslant \delta_2(t^N), \tag{5.11}$$

where

$$\delta_2(t^N) := \frac{2}{3}\sum_{j=0}^{N}(2k_j + k_{j+1})\left(C_{\mathrm{el}}\mathcal{E}_\partial^j + C_\Omega \alpha_{\min}^{-1}\|\partial f^j - \partial \bar{f}^j\|\right), \tag{5.12}$$

with

$$\mathcal{E}_\partial^j := \mathcal{E}(\partial U^j, \partial(A^j U^j) - \partial(\Pi^j f^j) + \partial f^j, \hat{\mathcal{T}}^j), \quad j = 0, 1, \ldots, N. \tag{5.13}$$

Proof. For $t \in (t^{j-1}, t^j]$, $j = 1, \ldots, N$, we have

$$\epsilon_t = \partial w^j - \partial U^j - k_j^{-1}(t^j - t)(t^j - 2t^{j-1} + t)(\partial^2 w^j - \partial^2 U^j), \tag{5.14}$$

from which we deduce

$$\int_{t^{j-1}}^{t^j} \|\epsilon_t\| \leqslant \frac{4k_j}{3}\|\partial w^j - \partial U^j\| + \frac{2k_j}{3}\|\partial w^{j-1} - \partial U^{j-1}\|, \tag{5.15}$$

noting that

$$\int_{t^{j-1}}^{t^j} k_j^{-2}(t^j - t)(t^j - 2t^{j-1} + t)\,\mathrm{d}t = \frac{2k_j}{3}. \tag{5.16}$$

Combining (5.15) for $j = 1, \ldots, N$, we arrive at

$$\int_0^{t^N} \|\epsilon_t\| \leqslant \frac{2}{3}\sum_{j=0}^{N}(2k_j + k_{j+1})\|\partial w^j - \partial U^j\|, \tag{5.17}$$

with $k_0 = 0$ and $k_{N+1} = 0$.

It remains to estimate the terms $\|\partial w^j - \partial U^j\|$. To this end, recalling the definition of the functions $w_*^j \in H_0^1(\Omega)$ from the proof of Lemma 5.1 and, since $\hat{V}_h^j := V_h^j \cap V_h^{j-1}$, we have $a(w_*^j, V) = a(U^j, V)$ for all $V \in \hat{V}_h^j$ and $a(w_*^{j-1}, V) = a(U^{j-1}, V)$ for all $V \in \hat{V}_h^j$, for $j = 1, \ldots, N$. Therefore, we deduce

$$a(\partial w_*^j, V) = a(\partial U^j, V) \quad \forall\, V \in \hat{V}_h^j \tag{5.18}$$

for $j = 1, \ldots, N$; that is, ∂U^j is the finite element solution in $\hat{V}_h^j$ of the boundary value problem

$$a(\partial w_*^j, V) = \langle \partial(A^j U^j) - \partial(\Pi^j f^j) + \partial f^j, v\rangle \quad \forall v \in H_0^1(\Omega). \tag{5.19}$$

In view of Theorem 3.7, this implies that

$$\|\partial w_*^j - \partial U^j\| \leqslant C_{\mathrm{el}}\mathcal{E}_\partial^j \tag{5.20}$$

for $j=1,\ldots,N$. We also recall that, by construction, we have $a(\partial w^0, V) = a(V^0, V)$ for all $V \in V_h^0$. Hence, (5.8) also holds.

Moreover, since

$$a(\partial w^j, V) = \langle \partial (A^j U^j) - \partial (\Pi^j f^j) + \partial \bar{f}^j, v\rangle \quad \forall v \in H_0^1(\Omega), \tag{5.21}$$

$j=1,\ldots,N$ (cf. Definition 4.1). As in (5.9), elliptic stability implies

$$\|\partial w^j - \partial w_*^j\| \leqslant C_\Omega \alpha_{\min}^{-1} \|\partial \bar{f}^j - \partial f^j\| \tag{5.22}$$

for $j=1,\ldots,N$ and, using the triangle inequality

$$\|\partial w^j - \partial U^j\| \leqslant \|\partial w^j - \partial w_*^j\| + \|\partial w_*^j - \partial U^j\|, \tag{5.23}$$

along with the bounds (5.22), (5.8) and (5.20), now implies the result. □

THEOREM 5.4 (Fully discrete residual-type *a posteriori* bound) With the same hypotheses and notation as in Theorems 4.6 and 3.7, we have the bound

$$\begin{aligned}\|e\|_{L^\infty(0,t^N;L^2(\Omega))} \leqslant \delta_1(t^N) + \sqrt{2} C_{\mathrm{el}} \mathcal{E}^0 + \sqrt{2}\|u_0 - U(0)\| \\ + 2\delta_2(t^N) + 2\sum_{i=1}^{4} \eta_i(t^N) + C_{a,N}\|u_1 - V^0\|,\end{aligned} \tag{5.24}$$

where δ_1 and $\mathcal{E}^0$ are defined in Lemma 5.1, δ_2 is defined in Lemma 5.3 and η_i, $i=1,2,3,4$ from Definition 4.5.

Proof. Combining Theorem 4.6 with the bounds derived for ϵ in Lemma 5.1 and ϵ_t in Lemma 5.3, we arrive at an *a posteriori* error bound. □

6. Final remarks

The design and implementation of adaptive algorithms for the wave equation based on rigorous *a posteriori* error estimators is a largely unexplored subject, despite the importance of these problems in the modelling of a number of physical phenomena. To this end, this work presents rigorous *a posteriori* error bounds in the $L^\infty(L^2)$ norm for second-order linear hyperbolic initial/boundary value problems. The use of a novel space-time reconstruction technique, which hinges on the one-field formulation of the problem, appears to be generic and it is expected to be applicable to second-order hyperbolic problems with possibly nonlinear spatial operators or other with spatial discretizations. Although the case of residual-type estimators has been demonstrated above, it is evident that Theorem 4.6 can be combined with a variety of other *a posteriori* estimators for elliptic problems. The derived bounds appear to be of optimal order, although no efficiency bounds are presented; this would be an interesting direction for further research. It is worth noting, however, that some of the terms appearing in the *a posteriori* bound presented above are completely analogous to some of the terms in the respective *a posteriori* bounds from Bernardi & Süli (2005), which are, in turn, shown to be efficient. The numerical implementation of the proposed bounds in the context of adaptive algorithm design for second-order hyperbolic problems deserves special attention and will be considered elsewhere.

Funding

E.H.G. acknowledges the support of the Nuffield Foundation, UK and of the Foundation for Research and Technology-Hellas, Heraklion, Greece. O.L. acknowledges the partial support of the Royal Society UK and of the Foundation for Research and Technology-Hellas, Heraklion, Greece, where the initial steps of this work were taken. C.M. acknowledges the support of the London Mathematical Society, the Universities of Leicester and Sussex, UK, and was supported in part by the European Union grant No. MEST-CT-2005-021122.

References

Adams, R. A. & Fournier, J. J. F. (2003) *Sobolev Spaces: Pure and Applied Mathematics*, vol. 140, 2nd edn. Amsterdam: Elsevier/Academic Press, pp. xiv+305.

Adjerid, S. (2002) A posteriori finite element error estimation for second-order hyperbolic problems. *Comput. Methods Appl. Mech. Eng.*, **191**, 4699–4719.

Ainsworth, M. & Oden, J. T. (2000) *A Posteriori Error Estimation in Finite Element Analysis. Pure and Applied Mathematics*. New York: Wiley-Interscience, pp. xx+240.

Akrivis, G., Makridakis, C. & Nochetto, R. H. (2006) A posteriori error estimates for the Crank–Nicolson method for parabolic equations. *Math. Comp.*, **75**, 511–531 (electronic).

Babuška, I. & Strouboulis, T. (2001) *The Finite Element Method and Its Reliability. Numerical Mathematics and Scientific Computation*. New York: Clarendon Press, pp. xii+802.

Baker, G. A. (1976) Error estimates for finite element methods for second order hyperbolic equations. *SIAM J. Numer. Anal.*, **13**, 564–576.

Baker, G. A. & Bramble, J. H. (1979) Semidiscrete and single step fully discrete approximations for second order hyperbolic equations. *RAIRO Anal. Numér.*, **13**, 75–100.

Baker, G. A. & Dougalis, V. A. (1980) On the L^∞-convergence of Galerkin approximations for second-order hyperbolic equations. *Math. Comp.*, **34**, 401–424.

Baker, G. A., Dougalis, V. A. & Serbin, S. M. (1979) High order accurate two-step approximations for hyperbolic equations. *RAIRO Anal. Numér.*, **13**, 201–226.

Bamberger, A., Joly, P. & Roberts, J. E. (1990) Second-order absorbing boundary conditions for the wave equation: a solution for the corner problem. *SIAM J. Numer. Anal.*, **27**, 323–352.

Bangerth, W. & Rannacher, R. (1999) Finite element approximation of the acoustic wave equation: error control and mesh adaptation. *East-West J. Numer. Math.*, **7**, 263–282.

Bangerth, W. & Rannacher, R. (2001) Adaptive finite element techniques for the acoustic wave equation. *J. Comput. Acoust.*, **9**, 575–591.

Bécache, E., Joly, P. & Tsogka, C. (2000) An analysis of new mixed finite elements for the approximation of wave propagation problems. *SIAM J. Numer. Anal.*, **37**, 1053–1084 (electronic).

Bergam, A., Bernardi, C. & Mghazli, Z. (2005) A posteriori analysis of the finite element discretization of some parabolic equations. *Math. Comp.*, **74**, 1117–1138 (electronic).

Bernardi, C. & Süli, E. (2005) Time and space adaptivity for the second-order wave equation. *Math. Models Methods Appl. Sci.*, **15**, 199–225.

Bernardi, C. & Verfürth, R. (2004) *A posteriori* error analysis of the fully discretized time-dependent Stokes equations. *Math. Model. Numer. Anal.*, **38**, 437–455.

Carstensen, C. & Bartels, S. (2002) 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**, 945–969 (electronic).

Cascon, J. M., Kreuzer, C., Nochetto, R. H. & Siebert, K. G. (2008) Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, **46**, 2524–2550.

Ciarlet, P. G. (1978) *The Finite Element Method for Elliptic Problems*, vol. 4. Amsterdam: North-Holland, pp. xix+530.

Cohen, G., Joly, P. & Tordjman, N. (1993) Construction and analysis of higher order finite elements with mass lumping for the wave equation. *Second International Conference on Mathematical and Numerical Aspects of Wave Propagation, Newark, DE, 1993*. Philadelphia, PA: SIAM, pp. 152–160.

Dörfler, W. (1996) A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, **33**, 1106–1124.

Dougalis, V. A. & Serbin, S. M. (1981) On the efficiency of some fully discrete Galerkin methods for second-order hyperbolic equations. *Comput. Math. Appl.*, **7**, 261–279.

Eriksson, K. & Johnson, C. (1995) Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$. *SIAM J. Numer. Anal.*, **32**, 706–740.

Houston, P. & Süli, E. (2001) Adaptive Lagrange–Galerkin methods for unsteady convection–diffusion problems. *Math. Comp.*, **70**, 77–106.

Johnson, C. (1993) Discontinuous Galerkin finite element methods for second order hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, **107**, 117–129.

Karakashian, O. & Makridakis, C. (2005) Convergence of a continuous Galerkin method with mesh modification for nonlinear wave equations. *Math. Comp.*, **74**, 85–102 (electronic).

Lakkis, O. & Makridakis, C. (2006) Elliptic reconstruction and *a posteriori* error estimates for fully discrete linear parabolic problems. *Math. Comp.*, **75**, 1627–1658 (electronic).

Makridakis, C. G. (1992) On mixed finite element methods for linear elastodynamics. *Numer. Math.*, **61**, 235–260.

Makridakis, C. & Nochetto, R. H. (2003) Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Numer. Anal.*, **41**, 1585–1594 (electronic).

Makridakis, C. & Nochetto, R. H. (2006) A posteriori error analysis for higher order dissipative methods for evolution problems. *Numer. Math.*, **104**, 489–514.

Picasso, M. (1998) Adaptive finite elements for a linear parabolic problem. *Comput. Methods Appl. Mech. Eng.*, **167**, 223–237.

Raviart, P.-A. & Thomas, J.-M. (1983) *Introduction à l'analyse numérique des équations aux dérivées partielles*. Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree]. Paris: Masson, p. 224.

Stevenson, R. (2007) Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, **7**, 245–269.

Süli, E. (1996) *A Posteriori Error Analysis and Global Error Control for Adaptive Finite Volume Approximations of Hyperbolic Problems. Numerical Analysis 1995 (Dundee, 1995)*. Pitman Research Notes in Mathematics Series, vol. 344. Harlow: Longman, pp. 169–190.

Süli, E. (1999) *A posteriori* error analysis and adaptivity for finite element approximations of hyperbolic problems. *An Introduction to Recent Developments in Theory and Numerics for Conservation Laws (Freiburg/Littenweiler, 1997)*. Lecture Notes of Computer Science and Engineering, vol. 5. Berlin: Springer, pp. 123–194.

Verfürth, R. (1996) *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Chichester/Stuttgart: Wiley/Teubner, pp. vi+128.

Verfürth, R. (2003) *A posteriori* error estimates for finite element discretizations of the heat equation. *Calcolo*, **40**, 195–212.