Sussex Research

Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems

Omar Lakkis, Charalambos Makridakis

Publication date
01-10-2006

Licence
This work is made available under the Copyright not evaluated licence and should only be used in accordance with that licence. For more information on the specific terms, consult the repository record for this item.

Citation for this work (American Psychological Association 7th edition)
Lakkis, O., & Makridakis, C. (2006). *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems* (Version 1). University of Sussex. https://hdl.handle.net/10779/uos.23311874.v1

Published in
Mathematics of Computation

Link to external publisher version
https://doi.org/10.1090/S0025-5718-06-01858-8

Copyright and reuse:
This work was downloaded from Sussex Research Open (SRO). This document is made available in line with publisher policy and may differ from the published version. Please cite the published version where possible. Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners unless otherwise stated. For more information on this work, SRO or to report an issue, you can contact the repository administrators at sro@sussex.ac.uk. Discover more of the University's research at https://sussex.figshare.com/

Sussex Research Online

Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems

Article (Unspecified)

Lakkis, Omar and Makridakis, Charalambos (2006) Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems. Mathematics of Computation, 75 (256). pp. 1627-1658. ISSN 0025-5718

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/1230/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk

MATHEMATICS OF COMPUTATION
Volume 75, Number 256, October 2006, Pages 1627–1658
S 0025-5718(06)01858-8
Article electronically published on May 26, 2006

ELLIPTIC RECONSTRUCTION AND A POSTERIORI ERROR ESTIMATES FOR FULLY DISCRETE LINEAR PARABOLIC PROBLEMS

OMAR LAKKIS AND CHARALAMBOS MAKRIDAKIS

Abstract. We derive a posteriori error estimates for fully discrete approximations to solutions of linear parabolic equations. The space discretization uses finite element spaces that are allowed to change in time. Our main tool is an appropriate adaptation of the elliptic reconstruction technique, introduced by Makridakis and Nochetto. We derive novel a posteriori estimates for the norms of $\mathrm{L}_\infty(0,T;\mathrm{L}_2(\Omega))$ and the higher order spaces, $\mathrm{L}_\infty(0,T;\mathrm{H}^1(\Omega))$ and $\mathrm{H}^1(0,T;\mathrm{L}_2(\Omega))$, with optimal orders of convergence.

1. Introduction

Adaptive mesh refinement methods for variational problems have been the object of intense study in recent years. The main objective of these methods is to reduce the computational cost in the numerical approximation of PDE solutions. Their usefulness is especially apparent when the exact solution has strong, geometrically localized variations or exhibits singularities. A posteriori estimates have proved to be a particularly successful mathematical tool in devising efficient adaptive versions of many numerical schemes. In addition, a posteriori estimates provide a new point of view in the theoretical investigation of a scheme's behavior. This is especially important for problems where "reasonable discretizations" do not always perform as expected.

In the context of finite element methods (FEM), the theory of a posteriori estimates for linear stationary problems is by now rather mature [2, 30, and references]. The situation for nonlinear and time dependent problems, however, has not yet been as thoroughly explored. Even for the linear parabolic equation, in spite of important advances made in the early '90s [14, 15] and subsequent ones [6, 9, 27], many issues have yet to be tackled. Such issues are, for instance, the derivation of optimal order estimates in various norms via the energy or other direct methods, the use of nonresidual based estimators to control the elliptic part of the error, and estimates for various time discretization methods.

In this paper we address some of these issues in the context of fully discrete linear parabolic problems where mesh modification in time might occur—as it is natural to expect in adaptive schemes for time-dependent problems. Our main tool

Received by the editor December 26, 2003 and, in revised form, May 23, 2005.
2000 *Mathematics Subject Classification.* Primary 65N30.
The first author was supported by the E.U. RTN *Hyke* HPRN-CT-2002-00282 and the EU's *MCWave Marie Curie Fellowship* HPMD-CT-2001-00121 during the preparation of this work at FORTH in Heraklion of Crete, Greece.

©2006 American Mathematical Society
Reverts to public domain 28 years from publication

in deriving the estimates is an appropriate adaptation to the fully discrete case of the *elliptic reconstruction* technique introduced by Makridakis and Nochetto for the model problem of semidiscrete finite element approximations [24]. A main characteristic of this approach, which contrasts with other direct techniques found in the literature, is that we can virtually use any available a posteriori estimates for elliptic equations to control the main part of the spatial error. Thus one can take full advantage of a well-established theory, instead of trying to adapt the estimates case by case. This follows from the fact that, in deriving the estimates, instead of comparing directly the exact solution with the numerical one, we construct an appropriate auxiliary function that fulfills two fundamental properties: (i) we know how to estimate its difference to the numerical solution via known a posteriori results, and (ii) it satisfies a variant of the original PDE with a right-hand side that can be controlled a posteriori in an optimal way.

In this paper, we combine the elliptic reconstruction technique with *a posteriori energy estimates* for the parabolic equation. Although residual-based a posteriori estimates using energy methods have been established [6, 9, 27], it is not immediately known how to use elliptic estimates other than the residual-based ones in them. For comparison's sake, in this paper we will derive residual based energy estimates; but we emphasize the fact that the techniques presented in this paper can be relatively easily used to derive estimators for parabolic problems where the "elliptic part" of the error is controlled by nonresidual type estimators, such as estimators based on the solution of local subproblems, for instance.

The main new results in this paper are *optimal order* a posteriori estimates, via energy techniques, in the following spaces (anticipating the notation that we will introduce in §1.1):

(a) $\mathrm{L}_\infty(0,T;\mathrm{L}_2(\Omega))$,

(b) $\mathrm{L}_\infty(0,T;\mathrm{H}_0^1(\Omega))$ and $\mathrm{H}^1(0,T;\mathrm{L}_2(\Omega))$.

In particular, we successfully address the open problem of obtaining optimal a posteriori error bounds via energy methods in $\mathrm{L}_\infty(0,T;\mathrm{L}_2(\Omega))$ for fully discrete schemes. As a by-product we also recover known results, such as optimal order estimates for the $\mathrm{L}_2(0,T;\mathrm{H}_0^1(\Omega))$ norm.

Energy methods have been used by other authors for Backward Euler fully discrete approximations to parabolic problems [27, 6, 9]. While these results are of optimal order in $\mathrm{L}_2(0,T;\mathrm{H}^1(\Omega))$, they are not so in $\mathrm{L}_\infty(0,T;\mathrm{L}_2(\Omega))$. Picasso [27] and Zhiming Chen and Feng Jia [9] manage to bound, at each time step, the spatial indicators by the error (lower bound); their technique is based on the "bubble functions" technique introduced by Verfürth for elliptic problems [30]. Also Bergam, Bernardi and Mghazli [6] have established lower bounds, and their estimators have the additional feature of decoupling the space and the time discretization errors.

We stress that for our analysis we do not require any extraneous conditions on the variation of the meshes and the corresponding finite element spaces, between successive time-steps, that may be hard or impossible to enforce in practical computations. In fact, our analysis allows us to even obtain estimates in the *higher order energy norms* (as mentioned in (1) above). The direct approach, which works well in deriving the lower order energy estimates [27, 9], will not work here, unless one imposes severe mesh conditions or makes quite strong a priori assumptions. Our approach permits to override this difficulty; see §4.10 for a technical discussion about this point.

Note that the elliptic reconstruction approach is not restricted by the method used for the stability analysis—in this case it being the energy method. Indeed, in a different paper [20] we show how the elliptic reconstruction can be used in the context of *duality techniques for parabolic equations*, that were first employed by Eriksson and Johnson [14] to derive a posteriori error estimates for the discontinuous Galerkin schemes. The results therein are of optimal order, up to a logarithmic factor, in $\mathrm{L}_\infty(0,T;\mathrm{L}_2(\Omega))$, and the Backward Euler scheme—which we study here—is the lowest order member in this class of methods.

Various other a posteriori estimates for semidiscrete and fully discrete approximations to linear and nonlinear parabolic problems in various norms are found in the literature [1, 4, 5, 11, 16, 25, 26, 31, 32]. In particular, Babuška, Feistauer, and Šolín [4] have derived estimates in $\mathrm{L}_2(0,T;\mathrm{L}_2(\Omega))$; see also Babuška et al. [1, 5]. Verfürth [31, 32] showed estimates in $\mathrm{L}_r(0,T;\mathrm{L}_\rho(\Omega))$, with $1<r,\rho<\infty$ for certain fully discrete approximations of certain quasilinear parabolic equations. Lakkis and Nochetto [21] used ad-hoc geometric energy norms to derive conditional a posteriori estimates for quasilinear equations such as the mean curvature flow of graphs. De Frutos and Novo [11] proved a posteriori error estimates of the p-version of space discrete schemes for parabolic equations; a similar function to the elliptic reconstruction and its improved approximation properties is used by García-Archilla and Titi [17]. Finally, for applications of suitable reconstructions to time discretizations of various type, we refer to [3, 23].

1.1. **Problem setting and notation.** Let us now focus our discourse by introducing the fully discrete scheme which will be the object of our analysis.

We start with the exact problem. Let Ω be a bounded domain of the Euclidean space $\mathbb{R}^d$, $d\in\mathbb{Z}^+$ and $T\in\mathbb{R}^+$. We assume throughout the paper that Ω is a convex polygonal domain, noting that our results could be extended to cover certain nonconvex domains, such as domains with reentrant corners in $d=2$, by using weighted a posteriori estimates for elliptic problems [22]. Since the difficulties in the analysis below in the case of other boundaries are mainly coming from the elliptic part of the error, the reader interested in a posteriori error estimates for curved boundaries is referred to Dörfler and Rumpf [12]. We will consider the problem of finding a finite element approximation of the solution $u\in\mathrm{L}_\infty(0,T;\mathrm{H}_0^1(\Omega))$, with $\partial_t u\in\mathrm{L}_2(0,T;\mathrm{L}_2(\Omega))$, to the linear parabolic problem

$$\begin{aligned}&\langle\partial_t u,\phi\rangle + a\left(u,\phi\right)=\langle f,\phi\rangle\,,\quad \forall\phi\in\mathrm{H}_0^1(\Omega),\\ &\text{and } u(0)=g,\end{aligned}\tag{1}$$

where $f\in\mathrm{L}_2(\Omega\times(0,T))$ and $g\in\mathrm{H}_0^1(\Omega)$, and a is a bilinear form on $\mathrm{H}_0^1(\Omega)$ defined by

$$a\left(v,\psi\right):=\langle\boldsymbol{A}\nabla v,\nabla\psi\rangle\,,\ \forall v,\psi\in\mathrm{H}_0^1(\Omega),\tag{2}$$

where "∇" denotes the spatial gradient and the matrix $\boldsymbol{A}\in\mathrm{L}_\infty(\Omega)^{d\times d}$ is such that

$$a\left(\psi,\phi\right)\le\beta\left|\psi\right|_1\left|\phi\right|_1\,,\quad\forall\phi,\psi\in\mathrm{H}_0^1(\Omega),\tag{3}$$

$$a\left(\phi,\phi\right)\ge\alpha\left|\phi\right|_1^2\,,\quad\forall\phi\in\mathrm{H}_0^1(\Omega),\tag{4}$$

with $\alpha,\beta\in\mathbb{R}^+$. Whenever not stated explicitly, we assume that the data $f,g,\boldsymbol{A}$ and the solution u of the above problem are sufficiently regular for our purposes.

Here and subsequently, for a given Lebesgue measurable set $D \subset \mathbb{R}^d$, we use the common notation

$$\langle \phi, \psi \rangle_D := \int_D \phi(\boldsymbol{x})\psi(\boldsymbol{x})\,\mathrm{d}\mu(\boldsymbol{x}), \tag{5}$$

$$\|\phi\|_D := \|\phi\|_{\mathrm{L}_2(D)} := \langle \phi, \phi \rangle_D^{1/2}, \tag{6}$$

$$|\phi|_{k,D} := \left\| \mathrm{D}^k \phi \right\|_D, \text{ for } k \in \mathbb{Z}^+ \quad (\text{with } \mathrm{D}^1\phi := \nabla\phi, \text{ etc.}), \tag{7}$$

$$\|\phi\|_{k,D} := \left(\|\phi\|_D^2 + \sum_{j=1}^{k} |\phi|_{j,D}^2 \right)^{1/2}, \text{ for } k \in \mathbb{Z}^+, \tag{8}$$

where $\mathrm{d}\mu(\boldsymbol{x})$ is either the Lebesgue measure element $\mathrm{d}\boldsymbol{x}$, if D is has positive such measure, or the $(d-1)$-dimensional (Hausdorff) measure $\mathrm{d}s(\boldsymbol{x})$, if D has zero Lebesgue measure. In many instances, in order to compress notation and when there is no danger of engendering confusion, we drop altogether the "differential" symbol from integrals; this applies also to integrals in time. We use the standard function spaces $\mathrm{L}_2(D)$, $\mathrm{H}^k(D)$, $\mathrm{H}_0^k(D)$ and denote by $\mathrm{H}^{-1}(D)$ the dual space of $\mathrm{H}_0^1(D)$. We omit the subscript D whenever $D = \Omega$. We denote the Poincaré constant relative to Ω by $C_{2,1}$ and, in view of the Poincaré inequality, we consider $|\cdot|_1$ to be the norm on $\mathrm{H}_0^1(\Omega)$. The *energy norm* $|\cdot|_a$ is defined through

$$|\phi|_a := a(\phi, \phi)^{1/2}, \ \forall \phi \in \mathrm{H}_0^1(\Omega). \tag{9}$$

It is equivalent to the norm $|\cdot|_1$ on the space $\mathrm{H}_0^1(\Omega)$, in view of (3) and (4). In particular, we will often use the following inequality:

$$|\phi|_1 \le \alpha^{-1/2} |\phi|_a, \ \forall \phi \in \mathrm{H}_0^1(\Omega). \tag{10}$$

In order to discretize the time variable in (1), we introduce the partition $0 = t_0 < t_1 < \cdots < t_N = T$ of $[0, T]$. Let $I_n := (t_{n-1}, t_n]$, and we denote by $\tau_n := t_n - t_{n-1}$ the time steps. We will consistently use the following "superscript convention": whenever a function depends on time, e.g., $f(\boldsymbol{x}, t)$, and the time is fixed to be $t = t_n$, $n \in [0:N]$, we denote it by $f^n(\boldsymbol{x})$. Moreover, we often drop the space dependence explicitly, e.g., we write $f(t)$ and f^n with reference to the functions in the previous sentence.

We use a standard FEM to discretize the space variable. Let $(\mathscr{T}_n)_{n\in[0:N]}$ be a family of conforming triangulations of the domain Ω [8, 10]. These triangulations are allowed to change arbitrarily from a timestep to the next, as long as they maintain some very mild *compatibility* requirements. Our use of the term "compatibility" is precisely defined in Appendix A; it is an extremely mild requirement which is easily implemented in practice.

For each given triangulation $\mathscr{T}_n$, we denote by h_n its meshsize function defined as

$$h_n(\boldsymbol{x}) = \operatorname{diam}(K), \text{ where } K \in \mathscr{T}_n \text{ and } \boldsymbol{x} \in K, \tag{11}$$

for all $\boldsymbol{x} \in \Omega$. We also denote by $\mathscr{S}_n$ the set of *internal sides* of $\mathscr{T}_n$. These are edges in $d = 2$ or faces in $d = 3$ that are contained in the interior of Ω. The *interior mesh of edges*, Σ_n, is then defined as the union of all internal sides $\bigcup_{E\in\mathscr{S}_n} E$. We associate with these triangulations the *finite element spaces*:

$$\tilde{\mathbb{V}}^n := \{\phi \in H^1(\Omega) : \forall K \in \mathscr{T}_n : \phi|_K \in \mathbb{P}^\ell\} \quad \text{and} \quad \mathbb{V}^n := \tilde{\mathbb{V}}^n \cap \mathrm{H}_0^1(\Omega), \tag{12}$$

where $\mathbb{P}^\ell$ is the space of polynomials in d variables of degree at most $\ell \in \mathbb{Z}^+$. Given two successive compatible triangulations $\mathscr{T}_{n-1}$ and $\mathscr{T}_n$, we define $\hat{h}_n := \max(h_n, h_{n-1})$ (see Appendix A and B). We will also use the sets $\hat{\Sigma}_n := \Sigma_n \cap \Sigma_{n-1}$ and $\check{\Sigma}_n := \Sigma_n \cup \Sigma_{n-1}$. Similarly, in §4, we will use $\hat{\hat{\Sigma}}_n := \Sigma_n \cap \Sigma_{n-1} \cap \Sigma_{n+1}$, $\check{\check{\Sigma}}_n := \Sigma_n \cup \Sigma_{n-1} \cup \Sigma_{n+1}$, $\hat{\hat{h}}_n := \max_{i\in[-1:1]} h_{n+i}$.

1.2. **Definition** (fully discrete scheme)**.** The standard backward Euler–Galerkin method for the discretization of problem (1) associated with the finite element spaces $\mathbb{V}^n$, leads to the following recursive *fully discrete scheme*:

$$\begin{aligned} &\text{let } U^0 := I^0 u(0), \\ &\text{for each } n \in [1:N] \text{ find } U^n \in \mathbb{V}^n \text{ such that} \\ &\qquad \tau_n^{-1}\left\langle U^n - U^{n-1}, \phi_n\right\rangle + a\left(U^n, \phi_n\right) = \left\langle f^n, \phi_n\right\rangle, \quad \forall \phi_n \in \mathbb{V}^n. \end{aligned} \tag{13}$$

Here the operator I^0 is some suitable interpolation or projection operator from $\mathrm{H}_0^1(\Omega)$, or $\mathrm{L}_2(\Omega)$, into $\mathbb{V}^0$.

In the sequel we shall use the continuous piecewise linear interpolant in time of the sequence (t_n, U^n) which we denote by $U(t)$ for $t \in [0,T]$ (see §2.4 for the precise definition).

1.3. **A posteriori estimates and reconstruction operators.** Our method consists in associating with U an auxiliary function $\omega : [0,T] \to \mathrm{H}_0^1(\Omega)$, in such a way that the *total error*

$$e := U - u \tag{14}$$

can be split as

$$e = \rho - \epsilon, \tag{15}$$

where

$$\epsilon := \omega - U, \quad \rho := \omega - u \tag{16}$$

satisfy the following two properties:

1. The error ϵ is easily controlled by a posteriori quantities of optimal order.
2. The error ρ satisfies a modification of the original PDE whose right-hand side depends on ϵ and U. This right-hand side can be bounded a posteriori in an optimal way.

In order to successfully apply this idea we shall choose the function ω to be a suitable reconstruction of U. The choice of this reconstruction is dictated by the elliptic operator at hand, and the precise definition of this *elliptic reconstruction* process is given in §2.2. We reap the benefits of our choice of ω by deriving *optimal order estimators* for the error measured not only in $\mathrm{L}_2(0,T;\mathrm{H}^1(\Omega))$, but also in $\mathrm{L}_\infty(0,T;\mathrm{L}_2(\Omega))$, as well as in $\mathrm{L}_\infty(0,T;\mathrm{H}_0^1(\Omega))$ and $\mathrm{H}^1(0,T;\mathrm{L}_2(\Omega))$. All the effects of mesh modification will be reflected in the right-hand side of the equation for ρ. In addition, our choosing ω as the elliptic reconstruction will have the effect of separating the spatial approximation error from the time approximation as much as possible. We show that the spatial approximation is embodied in ϵ which will be referred to as the *elliptic reconstruction error*, whereas the time approximation error information is conveyed by ρ, a fact that motivates the name *main parabolic*

error for this term. The splitting (15)–(16) of the error is already apparent in the spatially discrete case [24].

The PDE satisfied by ρ can be written in the following variational form.

1.4. **Lemma** (Main parabolic error equation)**.** *For each $n \in [1:N]$, and for each $\phi \in \mathrm{H}_0^1(\Omega)$,*

$$\begin{aligned} \langle \partial_t \rho, \phi \rangle + a(\rho, \phi) &= \langle \partial_t \epsilon, \phi \rangle + a(\omega - \omega^n, \phi) \\ &\quad + \langle P_0^n f^n - f, \phi \rangle + \tau_n^{-1} \left\langle P_0^n U^{n-1} - U^{n-1}, \phi \right\rangle \ \text{ on } I_n. \end{aligned} \tag{17}$$

Here P_0^n denotes the L^2-projection into $\tilde{\mathbb{V}}^n$.

A proof of this result is given in §2.6.

1.5. **Deriving the estimates: A general overview.** Identity (17) along with the properties of the elliptic reconstruction will allow us to obtain a posteriori error estimates in different norms. We start from the aforementioned splitting of the error

$$\|e(t)\|_X \le \|\epsilon(t)\|_X + \|\rho(t)\|_X, \tag{18}$$

where X is any suitable space of functions on Ω. The choice of X depends on the applications that are in mind, and on the ability to bound the terms on the right-hand side. In this respect, the following two observations are fundamental.

1. The first term, $\|\epsilon(t)\|_X = \|\omega(t) - U(t)\|_X$, can be bounded by appropriate a posteriori error estimates for elliptic problems. To see why this is possible we first observe that, at a fixed time t_n, the elliptic reconstruction, $\omega^n = \omega(t_n)$, is defined so that it is the exact solution of the elliptic problem

$$\text{find } v \in \mathrm{H}_0^1(\Omega) \text{ s.t. } -\operatorname{div}(\boldsymbol{A}\nabla v) = A^n U^n, \tag{19}$$

where $A^n U^n$ is the result of the *discrete elliptic operator* A^n on U^n with respect to the subspace $\mathbb{V}^n$ (see §2.1 for the detailed definitions).

Second, we observe that U^n is the finite element solution in $\mathbb{V}^n$ of the same elliptic problem (19). Note that, while $A^n v$ is not straightforward to compute for a general $v \in \mathbb{V}^n$, it becomes so when $v = U^n$. In fact, the "right-hand side" $A^n U^n$ of the elliptic problem can be expressed via known terms, by using the discrete parabolic equation (see (41) further), as follows

$$A^n U^n = P_0^n f^n - \left(U^n - P_0^n U^{n-1}\right) \tau_n^{-1}. \tag{20}$$

Therefore, to obtain an a posteriori estimate for this term, it is enough to assume that a posteriori error estimates for problem (19) are provided for the X-norm, through estimator functions that depend on U^n, $A^n U^n$, the triangulation parameters and the polynomial degree—many such estimator functions are available from standard a posteriori error analysis for elliptic problems [2, 7, 8, 30, e.g.]. Note that $\|\epsilon(t)\|_X$ exclusively contains spatial error effects which motivates our choice of the name "elliptic reconstruction error" for ϵ.

2. The second term, $\|\rho(t)\|_X = \|\omega(t) - u(t)\|_X$, is estimated by an appropriate use of (17) and techniques inspired from the analysis of parabolic PDE's. In this paper we will illustrate the use of energy methods, leaving the study of duality methods to a subsequent paper [20]. The right-hand side suggests that the resulting estimators will include quantities measuring the space error, the time error, the

variation of f, and the effect of mesh changes with respect to n. Since all the time effects are included in this term, we refer to ρ as the "main parabolic error".

1.6. **Comparison with the direct approach.** Identity (17) can be appreciated if we compare it with the error equation that one obtains from a direct comparison of u and U. In the direct approach the error relation is given by

$$\langle \partial_t e, \phi \rangle + a(e, \phi) = \langle \partial_t U, \phi \rangle + a(U^n, \phi) - \langle f, \phi \rangle + a(U - U^n, \phi). \tag{21}$$

Using the fact that U is the solution of the fully discrete scheme, one sees that

$$\begin{aligned} \langle \partial_t e, \phi \rangle + a(e, \phi) = {} & \langle \partial_t U, \phi - \phi_n \rangle + a(U^n, \phi - \phi_n) \\ & - \langle f^n, \phi - \phi_n \rangle \\ & - \langle f - f^n, \phi \rangle + a(U - U^n, \phi), \text{ for } \phi_n \in \mathbb{V}^n. \end{aligned} \tag{22}$$

A comparison between (17) and (22) demonstrates the two main differences in the corresponding approaches. Equation (22) has all the information from the numerical scheme "built-in", in particular it satisfies the Galerkin orthogonality property; therefore the error (and stability) analysis is dictated by the choices of both ϕ and ϕ_n. In the case of elliptic reconstruction, the Galerkin orthogonality property *is not used explicitly* in the analysis. The fact that U is a solution of the discrete scheme is used only *implicitly* through the reconstruction ω: it is used first to estimate ϵ and $\partial_t \epsilon$ (this estimate comes "for free" from the elliptic a posteriori theory precisely because of the definition of ω); second, it is used to derive (17), which then allows us to estimate $\|\rho\|_X$ and $\|\partial_t \rho\|_X$, in terms of time and data approximation estimators, and spatial estimators dictated only by ϵ.

A second important difference between these two approaches is the presence of the term $a(U^n, \phi - \phi_n)$, which is suboptimal in $\mathrm{L}_2(\Omega)$, in (22). Because of this term, the direct approach fails to provide optimal order a posteriori estimators in $\mathrm{L}_\infty(0, T; \mathrm{L}_2(\Omega))$, [9, 27]; the same problem can appear also in a different context [6]. On another hand, it is interesting to note that, the presence of this term is also the reason why the direct approach fails to lead satisfactory a posteriori results in the higher order energy norms; see §4.10 for the details.

Concerning the time discretization error, note that the term $a(U - U^n, \phi)$ in (22) is very similar to the term $a(\omega - \omega^n, \phi)$ in (17).

1.7. **Outline.** The rest of the paper is organized as follows. We introduce the necessary discrete and continuous operators in order to define the reconstruction ω in §2 and state some of its basic properties needed in the sequel.

In §3 we provide the a posteriori analysis in $\mathrm{L}_\infty(0, T; \mathrm{L}_2)$. The estimators in Theorem 3.2 are of optimal order and residual type. We carry out the analysis with a particular class of "elliptic" estimators in mind. As mentioned earlier, other choices are possible, but in order to use them, the arguments related to the terms involving mesh-change effects must be appropriately adapted.

Next, in §4, we show a posteriori estimates of optimal order in the higher order energy norms $\mathrm{L}_\infty(0, T; \mathrm{H}_0^1(\Omega))$ and $\mathrm{H}^1(0, T; \mathrm{L}_2(\Omega))$. This case is of particular interest as a simplified situation for a class of nonlinear degenerate parabolic problems where lower order energy estimates are not available [21]. In this section, we also discuss briefly why the direct approach cannot be applied successfully in this case (§4.10).

Finally, in §5 we complement the theory with some numerical experimentation. In particular, we show that the estimators derived in §3 have the expected, optimal, experimental order of convergence in the numerical test.

In the Appendix we collect some useful facts about the concept of compatible triangulations in §A, elliptic regularity inequalities in §B.1, interpolation operators and inequalities in §§B.2–B.3, and our convention on the constants labeling in §B.4.

We refer the reader interested in the practical aspects of our estimators to a forthcoming paper in which detailed numerical experiments, including comparisons with estimators derived by duality methods, as well as numerical investigation of the effect of mesh modification via a posteriori estimators, are included [20].

2. The elliptic reconstruction: Definition and preliminaries

We introduce basic tools and the elliptic reconstruction. Although the definitions in the first part of this section are independent of the time discretization and could be applied to any finite element space, we still use the space $\mathbb{V}^n$ defined in the Introduction.

2.1. **Definition** (Representation of the elliptic operator, discrete elliptic operator, projections)**.** Suppose a function $v \in \mathbb{V}^n$. The bilinear form can then be represented as

$$a(v,\phi) = \sum_{K\in\mathscr{T}_n} \langle -\operatorname{div}(\boldsymbol{A}\nabla v), \phi\rangle_K + \sum_{E\in\mathscr{S}_n} \langle J[v],\phi\rangle_E, \ \forall \phi \in \mathrm{H}_0^1(\Omega), \tag{23}$$

where $J[v]$ is the *spatial jump of the field* $\boldsymbol{A}\nabla v$ *across an element side* $E \in \mathscr{S}_n$ defined as

$$J[v]|_E(\boldsymbol{x}) = [\![\boldsymbol{A}\nabla v]\!]_E(\boldsymbol{x}) := \lim_{\varepsilon\to 0} (\boldsymbol{A}\nabla v(\boldsymbol{x}+\varepsilon\boldsymbol{\nu}_E) - \boldsymbol{A}\nabla v(\boldsymbol{x}-\varepsilon\boldsymbol{\nu}_E))\cdot\boldsymbol{\nu}_E, \tag{24}$$

where $\boldsymbol{\nu}_E$ is a choice (which does not influence this definition) between the two possible normal vectors to E at the point $\boldsymbol{x}$.

Since we use the representation (23) quite often, we now introduce a practical notation that makes it shorter and thus easier to manipulate in convoluted computations. For a finite element function, $v \in \mathbb{V}^n$ (or more generally for any Lipschitz continuous function v that is $\mathrm{C}^2(\operatorname{int}(K))$, for each $K \in \mathscr{T}_n$), denote by $A_{\mathrm{el}}v$ the *regular part* of the distribution $-\operatorname{div}(\boldsymbol{A}\nabla v)$, which is defined as a piecewise continuous function such that

$$\langle A_{\mathrm{el}}v,\phi\rangle = \sum_{K\in\mathscr{T}_n} \langle -\operatorname{div}(\boldsymbol{A}\nabla v),\phi\rangle, \quad \forall\phi\in\mathrm{H}_0^1(\Omega). \tag{25}$$

The operator A_{el} is sometimes referred to as the *elementwise elliptic operator*, as it is the result of the application of $-\operatorname{div}(\boldsymbol{A}\nabla\cdot)$ only on the interior of each element $K \in \mathscr{T}_n$. This observation justifies our subscript in the notation. We shall write the representation (23) in the shorter form,

$$a(v,\phi) = \langle A_{\mathrm{el}}v,\phi\rangle + \langle J[v],\phi\rangle_{\Sigma_n}, \quad \forall\phi\in\mathrm{H}_0^1(\Omega). \tag{26}$$

Let us now recall some more basic definitions that we will be using. The *discrete elliptic operator* associated with the bilinear form a and the finite element space $\mathbb{V}^n$ is the operator $A^n : \tilde{\mathbb{V}}^n \to \tilde{\mathbb{V}}^n$ defined by

$$\langle A^n v,\phi_n\rangle = a(v,\phi_n), \quad \forall\phi_n\in\tilde{\mathbb{V}}^n, \tag{27}$$

for $v \in \mathbb{V}^n$. The L_2-*projection operator* is defined as the operator $P_0^n : \mathrm{L}_2(\Omega) \to \tilde{\mathbb{V}}^n$ such that

$$\langle P_0^n v, \phi_n \rangle = \langle v, \phi_n \rangle, \quad \forall \phi_n \in \tilde{\mathbb{V}}^n, \tag{28}$$

for $v \in \mathrm{L}_2(\Omega)$; and the *elliptic projection operator* $P_1^n : \mathrm{H}_0^1(\Omega) \to \mathbb{V}^n$ is defined by

$$a\left(P_1^n v, \phi_n\right) = a\left(v, \phi_n\right), \quad \forall \phi_n \in \mathbb{V}^n. \tag{29}$$

The elliptic reconstruction, which we define next, is a partial right inverse of the elliptic projection [24]. (We note that similar operators have been introduced by Heywood and Rannacher [18, 19] and García-Archilla and Titi [17].)

2.2. **Definition** (Elliptic reconstruction)**.** We define the *elliptic reconstruction operator* associated with the bilinear form a and the finite element space $\mathbb{V}^n$ to be the unique operator $\mathscr{R}^n : \mathbb{V}^n \to \mathrm{H}_0^1(\Omega)$ such that

$$a\left(\mathscr{R}^n v, \phi\right) = \langle A^n v, \phi \rangle, \quad \forall \phi \in \mathrm{H}_0^1(\Omega), \tag{30}$$

for a given $v \in \mathbb{V}^n$. The function $\mathscr{R}^n v$ is referred to as the *elliptic reconstruction.*

The single most crucial property of $\mathscr{R}^n$ is that $v - \mathscr{R}^n v$ is orthogonal, with respect to a, to $\mathbb{V}^n$:

$$a\left(v - \mathscr{R}^n v, \phi_n\right) = 0, \quad \forall \phi_n \in \mathbb{V}^n. \tag{31}$$

From this property and recalling that Ω is assumed to be a convex polygonal, we obtain the following result whose proof uses standard techniques in a posteriori error estimates for elliptic problems [2, 7, 30].

2.3. **Lemma** (Elliptic reconstruction error estimates)**.** *For any $v \in \mathbb{V}^n$ the following estimates hold true:*

$$|\mathscr{R}^n v - v|_1 \le \frac{C_{3,1}}{\alpha} \|(A_{\mathrm{el}} v - A^n v) h_n\| + \frac{C_{5,1}}{\alpha} \left\| J[v] h_n^{1/2} \right\|_{\Sigma_n}, \tag{32}$$

$$\|\mathscr{R}^n v - v\| \le C_{6,2} \left\|(A_{\mathrm{el}} v - A^n v) h_n^2\right\| + C_{10,2} \left\| J[v] h_n^{3/2} \right\|_{\Sigma_n}, \tag{33}$$

where the constants $C_{k,j}$ are defined in Appendix B.

2.4. **Definition** (Discrete time extensions and derivatives)**.** Given any discrete function of time—that is, a sequence of values associated with each time node t_n—e.g., (U^n), we associate to it the continuous function of time defined by the Lipschitz continuous piecewise linear interpolation, e.g.,

$$U(t) := l_{n-1}(t) U^{n-1} + l_n(t) U^n, \text{ for } t \in I_n \text{ and } n \in [1:N]; \tag{34}$$

where the functions l are the hat functions defined by

$$l_n(t) := \frac{t - t_{n-1}}{\tau_n} \mathbf{1}_{I_n}(t) - \frac{t - t_{n+1}}{\tau_{n+1}} \mathbf{1}_{I_{n+1}}(t), \text{ for } t \in [0, T] \text{ and } n \in [0:N], \tag{35}$$

with $\mathbf{1}_X$ denoting the characteristic function of the set X. The *time-dependent elliptic reconstruction* of U is the function

$$\omega(t) := l_{n-1}(t) \mathscr{R}^{n-1} U^{n-1} + l_n(t) \mathscr{R}^n U^n, \text{ for } t \in I_n \text{ and } n \in [1:N]. \tag{36}$$

We observe that ω is a Lipschitz continuous function of time.

We introduce the following definitions whose purpose is to make notation more compact:

(a) *Discrete (backward) time derivative*

$$\partial U^n := \frac{U^n - U^{n-1}}{\tau_n}, \tag{37}$$

Note that $\partial U^n = \partial_t U(t)$, for all $t \in I_n$, hence we can think of ∂U^n as being the value of a discrete function at t_n. We thus define ∂U as the piecewise linear extension of $(\partial U^n)_n$, as we did with U.

(b) *Discrete (centered) second time derivative*

$$\partial^2 U^n := \frac{\partial U^{n+1} - \partial U^n}{\tau_n}. \tag{38}$$

(c) *Averaged (*L_2*-projected) discrete time derivative*

$$\overline{\partial} U^n := P_0^n \partial U^n = \frac{U^n - P_0^n U^{n-1}}{\tau_n}, \quad \forall n \in [1:N]\,. \tag{39}$$

The reason we introduce this notation is that ∂U^n, in general, does not belong to the current finite element space, $\mathbb{V}^n$, whereas $\overline{\partial} U^n$ does.

(d) The L_2*-projection of* f^n

$$\overline{f}^n := P_0^n f^n. \tag{40}$$

Since this is a discrete function of time, consistent with notation (34), we denote by $\overline{f}$ the piecewise linear interpolation of $(\overline{f}^n)_n$.

2.5. *Remark* (Pointwise form). The discrete elliptic operator A^n can be employed to write the fully discrete scheme (13) in the following *pointwise form*:

$$\overline{\partial} U^n(\boldsymbol{x}) + A^n U^n(\boldsymbol{x}) = \overline{f}^n(\boldsymbol{x}),\ \forall \boldsymbol{x} \in \Omega. \tag{41}$$

Indeed, in view of $\overline{\partial} U^n + A^n U^n - \overline{f}^n \in \tilde{\mathbb{V}}^n$, (13), and (27), we have

$$\begin{aligned} \left\langle A^n U^n + \overline{\partial} U^n - \overline{f}^n, \phi \right\rangle &= \left\langle A^n U^n + \overline{\partial} U^n - \overline{f}^n, P_0^n \phi \right\rangle \\ &= a\left(U^n, P_0^n \phi\right) + \left\langle \tau_n^{-1}(U^n - U^{n-1}) - f^n, P_0^n \phi \right\rangle = 0, \end{aligned} \tag{42}$$

for any $\phi \in \mathrm{H}_0^1(\Omega)$. Thus the function $\overline{\partial} U^n + A^n U^n - \overline{f}^n$ must be zero.

2.6. **Proof of Lemma 1.4.** The definitions in Definition 2.4 and (41) yield

$$\begin{aligned} \langle \partial U^n + A^n U^n - P_0^n f^n, \phi \rangle - \tau_n^{-1} \left\langle P_0^n U^{n-1} - U^{n-1}, \phi \right\rangle \\ = \left\langle \overline{\partial} U^n + A^n U^n - \overline{f}^n, \phi \right\rangle = 0, \end{aligned} \tag{43}$$

for each $\phi \in \mathrm{H}_0^1(\Omega)$ and $n \in [1:N]$. In view of the elliptic reconstruction definition we obtain

$$0 = \langle \partial U^n, \phi \rangle + a\left(\omega^n, \phi\right) - \langle P_0^n f^n, \phi \rangle - \tau_n^{-1} \left\langle P_0^n U^{n-1} - U^{n-1}, \phi \right\rangle. \tag{44}$$

On the other hand (1) implies

$$\langle \partial_t \rho, \phi \rangle + a\left(\rho, \phi\right) = \langle \partial_t \omega, \phi \rangle + a\left(\omega, \phi\right) - \langle f, \phi \rangle\,, \tag{45}$$

from which we subtract equation (44) and obtain (17). □

2.7. **Definition** (Residuals)**.** The residuals constitute the building blocks of the a posteriori estimators used in this paper. We associate with equations (1) and (41) two residual functions: the *inner residual*, which is defined as

$$\begin{aligned} R^0 &:= A_{\mathrm{el}}U^0 - A^0U^0, \\ R^n &:= A_{\mathrm{el}}U^n - A^nU^n = A_{\mathrm{el}}U^n - \overline{f}^n + \overline{\partial}U^n, \text{ for } n \in [1:N], \end{aligned} \tag{46}$$

and the *jump residual*, which is defined as

$$J^n := J[U^n]. \tag{47}$$

We note that, with Definition 2.1 in mind, the inner residual terms can be written in the following, more familiar but also more cumbersome, fashion:

$$\langle R^n, \phi\rangle = \sum_{K\in\mathscr{T}_n} \left\langle -\operatorname{div}(\boldsymbol{A}\nabla U^n) - P_0^n f(t_n) + \frac{U^n - P_0^n U^{n-1}}{\tau_n}, \phi \right\rangle_K . \tag{48}$$

3. A posteriori error estimates in the $\mathrm{L}_\infty(\mathrm{L}_2)$ and $\mathrm{L}_2(\mathrm{H}^1)$ norms

We start by introducing the following error estimators that are local in time. The full estimators, that will appear in Theorem 3.2, are accumulations in time of these local estimators. The accumulations, which can be of L_1, L_2 or L_∞ type, are anticipated by the first subscript in the estimators.

3.1. **Definition** ($\mathrm{L}_\infty(\mathrm{L}_2)$ and $\mathrm{L}_2(\mathrm{H}^1)$ error estimators)**.** We introduce, for $n \in [0:N]$, the *elliptic reconstruction error estimators*

$$\varepsilon_{\infty,n} := C_{6,2}\left\|h_n^2 R^n\right\| + C_{10,2}\left\|h_n^{3/2}J^n\right\|_{\Sigma_n}, \tag{49}$$

$$\varepsilon_{2,n} := \frac{C_{3,1}}{\alpha}\left\|h_n R^n\right\| + \frac{C_{5,1}}{\alpha}\left\|h_n^{1/2}J^n\right\|_{\Sigma_n}, \tag{50}$$

and, for $n \in [1:N]$, the *space error estimator*

$$\eta_{1,n} := C_{6,2}\left\|\hat{h}_n^2\partial R^n\right\| + C_{10,2}\left\|\hat{h}_n^{3/2}\partial J^n\right\|_{\hat{\Sigma}_n} + C_{14,2}\left\|\hat{h}_n^{3/2}\partial J^n\right\|_{\tilde{\Sigma}_n\setminus\hat{\Sigma}_n}, \tag{51}$$

the *data approximation error estimators* for *space* and *time* respectively

$$\gamma_{2,n} := \frac{C_{3,1}}{\sqrt{\alpha}}\left\|h_n(P_0^n - \mathrm{I})\left(f^n + \frac{U^{n-1}}{\tau_n}\right)\right\|, \ \beta_{1,n} := \frac{1}{\tau_n}\int_{t_{n-1}}^{t_n}\left\|f^n - f(t)\right\|\,\mathrm{d}t, \tag{52}$$

and the *time error estimator*

$$\theta_{1,n} := \begin{cases} \frac{1}{2}\left\|\partial(\overline{f}^n - \overline{\partial}U^n)\right\|\tau_n & \text{for } n \in [2:N], \\ \frac{1}{2}\left\|\overline{f}^1 - \overline{\partial}U^1 - A^0U^0\right\| & \text{for } n = 1. \end{cases} \tag{53}$$

We refer to Appendix B for an explanation of the constants $C_{k,j}$ involved here.

3.2. **Theorem** ($\mathrm{L}_\infty(\mathrm{L}_2)$ and $\mathrm{L}_2(\mathrm{H}^1)$ a posteriori error estimates)**.** *For each $m \in [1:N]$ the following error estimates hold:*

$$\max_{t\in[0,t_m]} \|u(t)-U(t)\| \le \left\|\mathscr{R}^0 U^0 - u(0)\right\| + \max_{n\in[0:m]} \varepsilon_{\infty,n} + 4\left(\mathscr{E}_{1,m}^2 + \mathscr{E}_{2,m}^2\right)^{1/2}, \tag{54}$$

$$\begin{aligned}\left(\int_0^{t_m} |u(t)-U(t)|_1^2\right)^{1/2} &\le \left\|\mathscr{R}^0 U^0 - u(0)\right\| + \left(\sum_{n=1}^m \left(\varepsilon_{2,n}^2 + \varepsilon_{2,n-1}^2\right)\tau_n\right)^{1/2} \\ &\quad + 4\left(\mathscr{E}_{1,m}^2 + \mathscr{E}_{2,m}^2\right)^{1/2},\end{aligned} \tag{55}$$

where

$$\mathscr{E}_{1,m} := \sum_{n=1}^m \left(\theta_{1,n} + \beta_{1,n} + \eta_{1,n}\right)\tau_n, \tag{56}$$

$$\mathscr{E}_{2,m}^2 := \sum_{n=1}^m \gamma_{2,n}^2 \tau_n. \tag{57}$$

Proof. Following the general strategy of §1.5 the error is decomposed as follows:

$$\|U(t)-u(t)\| = \|e(t)\| \le \|\epsilon(t)\| + \|\rho(t)\|. \tag{58}$$

To bound the first term, which is the elliptic reconstruction error, we apply Lemma 2.3. For the estimate in (54) we use (33) as follows:

$$\begin{aligned}\|\epsilon(t)\| &= \left\|l_{n-1}(t)\epsilon^{n-1} + l_n(t)\epsilon^n\right\| \le \max\left(\|\epsilon^n\|, \left\|\epsilon^{n-1}\right\|\right) \\ &\le \max_{n\in[0:m]} \|\epsilon^n\| \le \max_{n\in[0:m]} \varepsilon_{\infty,n},\end{aligned}$$

for all $t \in I_n$ and all $n \in [1:m]$. In an analogous way, we use (32) to obtain the estimate of the elliptic reconstruction error for (55).

The second term on the right-hand side of (58), which is the main parabolic error, will be estimated via Lemma 3.3 which we establish next. □

3.3. **Lemma** ($\mathrm{L}_\infty(\mathrm{L}_2)$ a posteriori estimate for the main parabolic error)**.** *For each $m \in [1:N]$, the following estimate holds:*

$$\left(\max_{[0,t_m]} \|\rho(t)\|^2 + 2\int_0^{t_m} |\rho(t)|_a^2 \,\mathrm{d}t\right)^{1/2} \le \left\|\rho^0\right\| + 4\left(\mathscr{E}_{1,m}^2 + \mathscr{E}_{2,m}^2\right)^{1/2}. \tag{59}$$

We divide the proof of this result in several steps which constitute the paragraphs §§3.4–3.8.

3.4. **The basic estimate.** To obtain $\mathrm{L}_\infty(\mathrm{L}_2)$ and $\mathrm{L}_2(\mathrm{H}^1)$ estimates we employ standard energy techniques. We replace ϕ in (17) by the main parabolic error $\rho = \omega - u$, and we integrate in time; thus we have

$$\begin{aligned}&\frac{1}{2}\|\rho^m\|^2 - \frac{1}{2}\left\|\rho^0\right\|^2 + \int_0^{t_m} |\rho(t)|_a^2 \,\mathrm{d}t \\ &\quad\le \sum_{n=1}^m \int_{t_{n-1}}^{t_n} |\langle \partial_t \epsilon(t), \rho(t)\rangle| + |a(\omega(t)-\omega^n, \rho(t))| \\ &\qquad + \left|\left\langle P_0^n f^n - f^n + \tau_n^{-1}\left(P_0^n U^{n-1} - U^{n-1}\right), \rho(t)\right\rangle\right| + |\langle f^n - f(t), \rho(t)\rangle| \,\mathrm{d}t \\ &\quad =: \sum_{n=1}^m \left(\mathscr{I}_n^1 + \mathscr{I}_n^2 + \mathscr{I}_n^3 + \mathscr{I}_n^4\right) =: \mathscr{I}_m.\end{aligned} \tag{60}$$

If we denote by $t_*^m \in [0, t_m]$ the time for which

$$\max_{t\in[0,t_m]} \|\rho(t)\| = \|\rho(t_m^*)\| =: \|\rho_*^m\|, \tag{61}$$

we deduce that

$$\frac{1}{2}\|\rho_*^m\|^2 - \frac{1}{2}\|\rho_0\|^2 + \int_0^{t_m^*} |\rho|_a^2 \le \mathscr{I}_m. \tag{62}$$

Consequently we have

$$\frac{1}{2}\|\rho_*^m\|^2 + \int_0^{t_m} |\rho|_a^2 \le \frac{1}{2}\|\rho_0\|^2 + 2\mathscr{I}_m. \tag{63}$$

We proceed by estimating each of the summands $\mathscr{I}_n^i$ appearing on the right-hand side of (60).

3.5. **Time error estimate.** In order to bound $\mathscr{I}_n^2$ in (60), which accounts for the time discretization error, we use directly the elliptic reconstruction definition (30) as follows:

$$\begin{aligned}
\mathscr{I}_n^2 &= \int_{t_{n-1}}^{t_n} |a(\omega(t) - \omega^n, \rho(t))| \,\mathrm{d}t \\
&= \int_{t_{n-1}}^{t_n} \left|a\left(l_{n-1}(t)\mathscr{R}^{n-1}U^{n-1} + l_n(t)\mathscr{R}^n U^n - \mathscr{R}^n U^n, \rho(t)\right)\right| \,\mathrm{d}t \\
&= \int_{t_{n-1}}^{t_n} l_{n-1}(t) \left|a\left(\mathscr{R}^{n-1}U^{n-1} - \mathscr{R}^n U^n, \rho(t)\right)\right| \,\mathrm{d}t \\
&= \int_{t_{n-1}}^{t_n} l_{n-1}(t) \left|\left\langle A^{n-1}U^{n-1} - A^n U^n, \rho(t)\right\rangle\right| \,\mathrm{d}t.
\end{aligned}$$

Therefore

$$\mathscr{I}_n^2 \le \int_{t_{n-1}}^{t_n} l_{n-1}(t) \left\|A^{n-1}U^{n-1} - A^n U^n\right\| \|\rho(t)\| \,\mathrm{d}t, \tag{64}$$

which leads to

$$\sum_{n=1}^m \mathscr{I}_n^2 \le \|\rho_*^m\| \sum_{n=1}^m \theta_{1,n}\tau_n. \tag{65}$$

3.6. **Spatial error estimate.** To estimate the term $\mathscr{I}_n^1$ on the right-hand side of (60), which measures the space error and mesh change, we will exploit the orthogonality property of the elliptic reconstruction (31). Observe that for each $n \in [1:N]$ we have

$$\begin{aligned}
\mathscr{I}_n^1 &= \int_{t_{n-1}}^{t_n} |\langle \partial_t \epsilon(t), \rho(t)\rangle| \,\mathrm{d}t \\
&= \tau_n^{-1} \int_{t_{n-1}}^{t_n} \left|\left\langle \mathscr{R}^n U^n - \mathscr{R}^{n-1}U^{n-1} - U^n + U^{n-1}, \rho(t)\right\rangle\right| \,\mathrm{d}t.
\end{aligned} \tag{66}$$

Since $\mathscr{R}^n U^n - U^n$ is orthogonal to $\mathbb{V}^n$ with respect to $a(\cdot,\cdot)$, the first term inside the brackets is orthogonal to $\mathbb{V}^n \cap \mathbb{V}^{n-1}$. We can therefore use standard residual-based a posteriori estimation techniques. Let $\psi : [0,T] \to \mathrm{H}_0^1(\Omega)$ be such that

$$a(\chi, \psi(t)) = \langle \rho(t), \chi\rangle, \quad \forall \chi \in \mathrm{H}_0^1(\Omega),\ t \in [0,T]. \tag{67}$$

By (31), Definition 2.1, and the use of the interpolation operator $\hat{\Pi}^n$ defined in §B.3, it follows that

$$\begin{aligned}
&\langle \mathscr{R}^n U^n - \mathscr{R}^{n-1}U^{n-1} - U^n + U^{n-1}, \rho(t)\rangle \\
&\quad = a\left(\mathscr{R}^n U^n - \mathscr{R}^{n-1}U^{n-1} - U^n + U^{n-1}, \psi(t)\right) \\
&\quad = a\left(\mathscr{R}^n U^n - \mathscr{R}^{n-1}U^{n-1} - U^n + U^{n-1}, \psi(t) - \hat{\Pi}^n\psi(t)\right) \\
&\quad = \left\langle A^n U^n - A^{n-1}U^{n-1} - A_{\mathrm{el}}(U^n - U^{n-1}), \psi(t) - \hat{\Pi}^n\psi(t)\right\rangle \\
&\qquad + \left\langle J^n - J^{n-1}, \psi(t) - \hat{\Pi}^n\psi(t)\right\rangle_{\check{\Sigma}_n},
\end{aligned} \tag{68}$$

for each $t \in I_n$. Using the pointwise form of the fully discrete scheme (41) we can rewrite these terms in a more compact form,

$$\begin{aligned}
A^n U^n - A^{n-1}U^{n-1} - A_{\mathrm{el}}(U^n - U^{n-1}) &= \left((\overline{\partial} + A_{\mathrm{el}})(U^n - U^{n-1}) - \overline{f}^n + \overline{f}^{n-1}\right) \\
&= \tau_n\left(\partial\left((\overline{\partial} + A_{\mathrm{el}})U^n - \overline{f}^n\right)\right) = \tau_n \partial R^n,
\end{aligned} \tag{69}$$

on each interval I_n. Here $R^n := (\overline{\partial} + A_{\mathrm{el}})U^n - \overline{f}^n$ is the *internal residual* function at time t_n. Likewise we have

$$J^n - J^{n-1} = \tau_n \partial J^n. \tag{70}$$

Whence, with $j \in \mathbb{Z}^+$ being at our disposal, in view of the interpolation inequalities in §B.2 we may conclude that

$$\begin{aligned}
\mathscr{I}_n^1 \le &\left(\int_{t_{n-1}}^{t_n} |\psi(t)|_j\right) \\
&\times \left(C_{3,j}\left\|\hat{h}_n^j \partial R^n\right\| + C_{5,j}\left\|\hat{h}_n^{j-1/2}\partial J^n\right\|_{\hat{\Sigma}_n} + C_{7,j}\left\|\hat{h}_n^{j-1/2}\partial J^n\right\|_{\check{\Sigma}_n\setminus\hat{\Sigma}_n}\right).
\end{aligned} \tag{71}$$

Since $\ell \ge 1$, we may take $j = 2$ in (71) and use the elliptic regularity (127) to get

$$\mathscr{I}_n^1 \le \max_{t\in I_n} \|\rho(t)\|\, \eta_{1,n}\tau_n, \tag{72}$$

where $\eta_{1,n}$ is given by (51). Hence

$$\sum_{n=1}^m \mathscr{I}_n^1 \le \|\rho_*^m\| \sum_{n=1}^m \eta_{1,n}\tau_n. \tag{73}$$

3.7. **Data approximation and mesh change estimates.** We now bound the term $\mathscr{I}_n^3$ in (60). Here we exploit the orthogonality of the L_2-projection. Since $\mathbb{V}^n \subset \ker(P_0^n - \mathrm{I})$ we have

$$\begin{aligned}
\mathscr{I}_n^3 &= \int_{t_{n-1}}^{t_n} \left|\left\langle (P_0^n - \mathrm{I})(f^n + \tau_n^{-1}U^{n-1}), \rho(t) - \Pi^n\rho(t)\right\rangle\right| \,\mathrm{d}t \\
&\le \int_{t_{n-1}}^{t_n} \left|\left\langle h_n(P_0^n - \mathrm{I})(f^n + \tau_n^{-1}U^{n-1}), h_n^{-1}(\rho(t) - \Pi^n\rho(t))\right\rangle\right| \,\mathrm{d}t \\
&\le \alpha^{-1/2}C_{3,1}\tau_n^{1/2}\left\|h_n(P_0^n - \mathrm{I})(f^n + \tau_n^{-1}U^{n-1})\right\| \left(\int_{t_{n-1}}^{t_n} |\rho(t)|_a \,\mathrm{d}t\right)^{1/2}.
\end{aligned} \tag{74}$$

We can thus conclude that

$$\sum_{n=1}^{m} \mathscr{I}_n^3 = \sum_{n=1}^{m} \left(\int_{t_{n-1}}^{t_n} |\rho|_a^2 \right)^{1/2} \gamma_{2,n} \tau_n^{1/2}. \tag{75}$$

We conclude this paragraph by estimating the fourth term on the right-hand side of (60) in a simple way as follows:

$$\mathscr{I}_n^4 \leq \left(\max_{t \in I_n} \|\rho(t)\| \right) \int_{t_{n-1}}^{t_n} \|f^n - f(t)\| \, \mathrm{d}t. \tag{76}$$

Thus

$$\sum_{n=1}^{m} \mathscr{I}_4 \leq \|\rho_*^m\| \sum_{n=1}^{m} \beta_{1,n} \tau_n. \tag{77}$$

3.8. **Proof of Lemma 3.3: last step.** What remains to be done in order to conclude the proof is to appropriately combine the results from §§3.5–3.7 with inequalities (60) and (63). We can write

$$\begin{aligned} \frac{1}{2} \|\rho_*^m\|^2 + \int_0^{t_m} |\rho|_a^2 \leq & \frac{1}{2} \|\rho_0\|^2 + 2 \|\rho_*^m\| \sum_{n=1}^{m} (\theta_{1,n} + \beta_{1,n} + \eta_{1,n}) \tau_n \\ & + 2 \sum_{n=1}^{m} \left(\int_{t_{n-1}}^{t_n} |\rho|_a^2 \right)^{1/2} \gamma_{2,n} \tau_n^{1/2}. \end{aligned} \tag{78}$$

We can now apply the elementary fact that, for $\boldsymbol{a} = (a_0, \ldots, a_m)$, $\boldsymbol{b} = (b_0, \ldots, b_m) \in \mathbb{R}^{m+1}$, and $c \in \mathbb{R}$, if

$$|\boldsymbol{a}|^2 \leq c^2 + \boldsymbol{a} \cdot \boldsymbol{b}, \tag{79}$$

then

$$|\boldsymbol{a}| \leq |c| + |\boldsymbol{b}|. \tag{80}$$

In particular, in reference to (78), we take

$$a_0 = \frac{1}{\sqrt{2}} \|\rho_*^m\|, \quad a_n = \left(\int_{t_{n-1}}^{t_n} |\rho|_a^2 \right)^{1/2}, \quad c = \frac{1}{\sqrt{2}} \left\|\rho^0\right\|, \tag{81}$$

$$b_0 = 2\sqrt{2} \sum_{n=1}^{m} (\theta_{1,n} + \beta_{1,n} + \eta_{1,n}) \tau_n, \quad b_n = 2\gamma_{2,n} \tau_n^{1/2}, \tag{82}$$

for $n \in [1:m]$, and obtain (59), which concludes the proof of the lemma. □

3.9. *Remark* (Relation to the semidiscrete case). The spatial error estimators containing η's should be compared with the ones corresponding to the (space) semidiscrete scheme given by

$$\left(\left\| h^3 \partial_t ((\partial_t - A_{\mathrm{el}}) u_h - f) \right\|^2 + \left\| h^{5/2} \partial_t J[u_h] \right\|_{\Sigma_h}^2 \right)^{1/2}, \tag{83}$$

where one triangulation $\mathscr{T}_h$ is given for all (continuous) time $t \in [0, T]$ [24, Equation (4.4)].

3.10. *Remark* (Mesh change). We interpret the presence of the coarsest meshsize $\hat{h}_n$ in the estimator as a reflection of the discrepancy between the finite element spaces $\mathbb{V}^{n-1}$ and $\mathbb{V}^n$, which might be different in general. Mesh change, a delicate issue in evolution problems, can lead to nonconvergent schemes despite the global meshsize going to zero. This happens in an example by T. Dupont [13, §4] for which

$$\max_n \sup_\Omega h_n \to 0 \quad \left(\text{but } \max_n \sup_\Omega \hat{h}_n \not\to 0 \right), \tag{84}$$

yet the discrete solution does not converge to the exact solution.

4. A posteriori error estimates in higher order norms

In this section we derive estimates in the seminorms corresponding to the spaces $\mathrm{H}^1(0,T;\mathrm{L}_2(\Omega))$ and $\mathrm{L}_\infty(0,T;\mathrm{H}^1(\Omega))$. The exposition of this section parallels that of §3. We start by introducing a posteriori error estimators that are local in time and which will be used in the subsequent main result. We warn the reader that although some symbols for error estimators in this section are the same as those of §3, the estimators themselves are changed; this notation is valid only in this section.

4.1. **Definition** (Error estimators for the $\mathrm{L}_\infty(\mathrm{H}^1)$and $\mathrm{H}^1(\mathrm{L}_2)$ seminorms)**.** We introduce the *elliptic reconstruction error estimator*

$$\varepsilon_{\infty,n} := \frac{C_{3,1}}{\alpha} \|h_n R^n\| + \frac{C_{5,1}}{\alpha} \left\| h_n^{1/2} J^n \right\|, \tag{85}$$

$$\varepsilon_{2,n} := C_{6,2} \left\| \hat{h}_n^2 \partial R^n \right\| + C_{10,2} \left\| \hat{h}_n^{3/2} \partial J^n \right\|_{\hat{\Sigma}_n} + C_{14,2} \left\| \hat{h}_n^{3/2} \partial J^n \right\|_{\check{\Sigma}_n \setminus \hat{\Sigma}_n}, \tag{86}$$

the *space error estimator*

$$\eta_{2,n} := C_{6,2} \left\| \hat{h}_n^2 \partial R^n \right\| + C_{10,2} \left\| \hat{h}_n^{3/2} \partial J^n \right\|_{\hat{\Sigma}_n} + C_{14,2} \left\| \hat{h}_n^{3/2} \partial J^n \right\|_{\check{\Sigma}_n \setminus \hat{\Sigma}_n}, \tag{87}$$

the *data space approximation error estimators*

$$\gamma_{1,n} := \alpha^{-1/2} C_{3,1} \left\| \hat{h}_n \partial \left((P_0^n - \mathrm{I})(f^n - \tau_n U^{n-1}) \right) \right\|, \tag{88}$$

$$\gamma_{\infty,n} := \alpha^{-1/2} C_{3,1} \left\| h_n (P_0^n - \mathrm{I})(f^n - \tau_n U^{n-1}) \right\|, \tag{89}$$

and the *data time approximation error estimator* and the *time error estimator*

$$\beta_{2,n} := \left(\frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \|f^n - f(t)\|^2 \, \mathrm{d}t \right)^{1/2}, \quad \theta_{2,n} = \frac{1}{\sqrt{3}} \left\| \partial_t (\overline{f} - \overline{\partial} U) \right\| \tau_n. \tag{90}$$

We refer to Appendix B for the definition of constants $C_{k,j}$ involved above.

4.2. **Theorem** ($\mathrm{L}_\infty(\mathrm{H}^1) \cap \mathrm{H}^1(\mathrm{L}_2)$ a posteriori estimates)**.** *Suppose the exact solution u of* (1) *satisfies*

$$\partial_t u \in \mathrm{L}_2(0,T;\mathrm{L}_2(\Omega)), \tag{91}$$

$$\partial_t u(t) \in \mathrm{H}_0^1(\Omega) \text{ for a.e. } t \in [0,T], \tag{92}$$

$$\nabla u \in \mathrm{L}_2(0,T;\mathrm{H}^1(\Omega)). \tag{93}$$

Then the following a posteriori error estimates hold:

$$\left(\int_0^{t_m} \|\partial_t (U(t)-u(t))\|^2 \,\mathrm{d}t\right)^{1/2} \le \left|\mathscr{R}^0 U^0 - u(0)\right|_a + 4\left(\mathscr{E}_{1,m}^2 + \mathscr{E}_{2,m}^2\right)^{1/2} + \varepsilon'_m \tag{94}$$

$$\max_{t\in[0,t_m]} |U(t)-u(t)|_a \le \left|\mathscr{R}^0 U^0 - u(0)\right|_a + 4\left(\mathscr{E}_{1,m}^2 + \mathscr{E}_{2,m}^2\right)^{1/2} + \varepsilon_m, \tag{95}$$

where

$$\mathscr{E}_{1,m} := 2 \max_{n\in[1:m]} \gamma_{\infty,m} + \sum_{n=2}^{m} \gamma_{1,n}\tau_n,$$

$$\mathscr{E}_{2,m}^2 := \sum_{n=1}^{m} \left(\theta_{2,n}^2 + \beta_{2,n}^2 + \eta_{2,n}^2\right)\tau_n,$$

$$\varepsilon_m := \max_{n\in[0:m]} \varepsilon_{\infty,n}, \quad \textit{and} \quad (\varepsilon'_m)^2 := \sum_{n=1}^{m} \varepsilon_{2,n}^2.$$

Proof. Following the general strategy of §1.5 the error is decomposed as follows:

$$|U(t)-u(t)|_Y = |e(t)|_Y \le |\epsilon(t)|_Y + |\rho(t)|_Y\,, \tag{96}$$

where Y is either $\mathrm{H}^1(0,t_m;\mathrm{L}_2(\Omega))$ or $\mathrm{L}_\infty(0,t_m;\mathrm{H}_0^1(\Omega))$ and $|\cdot|_Y$ is the corresponding seminorm.[1] The first term on the left-hand side of (96), which is the elliptic reconstruction error, can be estimated in the residual-based context via Lemma 2.3, as to obtain

$$|\epsilon(t)|_Y \le \begin{cases} \varepsilon'_m, & \text{if } Y = \mathrm{H}^1(0,t_m;\mathrm{L}_2(\Omega)), \\ \varepsilon_m, & \text{if } Y = \mathrm{L}_\infty(0,t_m;\mathrm{H}_0^1(\Omega)). \end{cases} \tag{97}$$

The second term on the left-hand side of (96) is estimated with the help of Lemma 4.3, which we state and prove next. □

4.3. **Lemma** ($\mathrm{L}_\infty(\mathrm{H}^1)\cap\mathrm{H}^1(\mathrm{L}_2)$ estimates for the main parabolic error)**.** *For each* $m\in[1:N]$ *the following a posteriori estimate is valid:*

$$\left(\max_{t\in[0,t_m]} |\rho(t)|_a^2 + 2\int_0^{t_m} \|\partial_t\rho\|^2\right)^{1/2} \le \left|\rho^0\right|_a + 4\left(\mathscr{E}_{1,m}^2 + \mathscr{E}_{2,m}^2\right)^{1/2}, \tag{98}$$

with reference to the notation of Theorem 4.2.

[1] In this section we deliberately use $|\cdot|_a$ instead of $|\cdot|_1$ as the norm for $\mathrm{H}_0^1(\Omega)$ in order to keep the exposition clear. The changes to replace $|\cdot|_a$ by $|\cdot|_1$ are straightforward.

As in §3, the proof of this lemma is subdivided into several steps that constitute §§4.5–4.9. Since the arguments are very similar to those of §3, we condense the discourse and stress only the main differences. To motivate the proof, we first discuss the following semidiscrete case.

4.4. **The spatially discrete case.** The estimates of Theorem 4.2 are based on the energy estimate in higher order norms for problem (1) which reads

$$\left(|u(t)|_a^2 + 2\int_0^t \|\partial_t u\|^2\right)^{1/2} \leq |u(0)|_a + \sqrt{2}\,\|f\|_{\mathrm{L}_2(0,t;\mathrm{L}_2(\Omega))}\,. \tag{99}$$

This estimate can be obtained by testing the PDE with $\partial_t u$ and integrating in time. For this case more regularity of u is required than in the lower order norms case of §3. Sufficient regularity requirements on u are given by conditions (91)–(93).

We stress that in some particular situations these stronger energy norms can play an important role. For instance, an estimate that is based on nonlinear quantities similar to these higher order norms has been derived from the error related to the mean curvature flow of graphs [21]; in that situation there is no reasonable way to obtain estimates by testing the solution. The only approach that works is testing with the time derivative of the solution.

Let us now turn our discussion toward the use of this energy estimate in the semidiscrete case; namely only spatially discrete, with $\mathbb{V}^h$ as a finite element space. The semidiscrete case is simpler than, and motivated by, the more involved fully discrete case which we will deal with in the next paragraphs. The semidiscrete case has been extensively studied by Makridakis and Nochetto for the usual (lower order) norms [24]. We further simplify our discussion by also assuming that $f \in \mathbb{V}^h$.

The starting point of the error estimate is, as in §3, the semidiscrete analog of (17) which is given by

$$\langle \partial_t \rho, \phi\rangle + a\,(\rho,\phi) = \langle \partial_t \epsilon, \phi\rangle\,, \quad \forall \phi \in \mathrm{H}_0^1(\Omega). \tag{100}$$

(Compare with [24, Equation (3.2)].) Taking this identity with $\phi = \partial_t \rho$—which is why we need to assume the extra regularity properties (91)–(93)—and integrating in time we obtain

$$\int_0^T \|\partial_t \rho\|^2 + \frac{1}{2}\,|\rho(T)|_a^2 - \frac{1}{2}\,|\rho(0)|_a^2 = \int_0^T \langle \partial_t \epsilon, \partial_t \rho\rangle\,. \tag{101}$$

We now have the choice to control the right-hand side in two different ways.

(a) We use a straightforward $\mathrm{L}_2(0,T;\mathrm{L}_2(\Omega))$ estimate that leads to

$$\begin{aligned}\int_0^T \langle \partial_t \epsilon, \partial_t \rho\rangle &= \left(\int_0^T \|\partial_t \epsilon\|^2\right)^{1/2}\left(\int_0^T \|\partial_t \rho\|^2\right)^{1/2}\\ &= \left(\int_0^T \mathscr{E}[\mathbb{V}^h, A_h \partial_t u_h; \mathrm{L}_2(\Omega)]^2\right)^{1/2}\left(\int_0^T \|\partial_t \rho\|^2\right)^{1/2},\end{aligned} \tag{102}$$

where $\mathscr{E}[\cdot]$ is an *elliptic error estimator function* [24]. This could be, for instance, but not necessarily so, the residual based estimators of Lemma 2.3.

(b) An alternative estimate that is often useful when quadratic or higher finite elements are employed or when only the energy norm $|\rho|_a$ can be controlled—in a

nonlinear setting, for instance—involves an integration by parts in time as follows:

(103)
$$\begin{aligned}\int_0^T &\langle \partial_t \epsilon, \partial_t \rho\rangle - \langle \partial_t \epsilon(0), \rho(0)\rangle \\ &= \langle \partial_t \epsilon(T), \rho(T)\rangle - \int_0^T \left\langle \partial_t^2 \epsilon, \rho\right\rangle \\ &\le C \max_{[0,T]} |\rho|_a \left(\max_{[0,T]} \mathscr{E}[\mathbb{V}^h, A_h \partial_t u_h; \mathrm{H}^{-1}(\Omega)] + \int_0^T \mathscr{E}[\mathbb{V}^h, A_h \partial_t^2 u_h; \mathrm{H}^{-1}(\Omega)] \right).\end{aligned}$$

The extra time differentiation will not affect the order of convergence of the right-hand side because the elliptic reconstruction error ε is purely elliptic in nature. We note, however, that in order to make estimate (103) rigorous, it is necessary to impose extra time-regularity assumptions on the approximate solution and to have an $\mathrm{H}^{-1}(\Omega)$-norm elliptic error estimator function $\mathscr{E}[\cdot,\cdot;\mathrm{H}^{-1}(\Omega)]$ available. Such estimators can be obtained with optimal order, using the duality technique, under the assumption that the domain is smooth. This is an issue that is beyond the scope of this paper, so we will limit our analysis to the first alternative.

4.5. **The basic estimate for the fully discrete case.** We now proceed with the proof of Lemma 4.3. The first step consists of taking $\phi = \partial_t \rho$ in identity (17) and integrating by parts in time as follows:

$$\begin{aligned}\int_0^{t_m} \|\partial_t \rho\|^2 + \frac{1}{2}|\rho^m|_a^2 &= \frac{1}{2}|\rho^0|_a^2 \\ &+ \sum_{n=1}^m \int_{t_{n-1}}^{t_n} \langle \partial_t \epsilon, \partial_t \rho(t)\rangle + a(\omega(t) - \omega^n, \partial_t \rho(t)) \\ &+ \left\langle P_0^n f^n - f^n + \tau_n^{-1}(P_0^n U^{n-1} - U^{n-1}), \partial_t \rho(t)\right\rangle \\ &+ \langle f^n - f(t), \partial_t \rho(t)\rangle \,\mathrm{d}t \\ &=: \frac{1}{2}|\rho(0)|_a^2 + \sum_{n=1}^m (\mathscr{I}_n^1 + \mathscr{I}_n^2 + \mathscr{I}_n^3 + \mathscr{I}_n^4) =: \mathscr{I}_m.\end{aligned} \tag{104}$$

Introduce $t_m^* \in [0, t_m]$ such that

$$|\rho_*^m|_a := |\rho(t_m^*)|_a = \max_{t\in[0,t_m]} |\rho(t)|_a\,. \tag{105}$$

Let $m^* \in [1:m]$ be the index for which $t_m^* \in I_{m^*}$. We can write

$$\int_0^{t_m^*} \|\partial_t \rho\|^2 + \frac{1}{2}|\rho_*^m|_a^2 = \frac{1}{2}|\rho(0)|_a^2 + \sum_{n=1}^{m^*} \left(\mathscr{J}_n^1 + \mathscr{J}_n^2 + \mathscr{J}_n^3 + \mathscr{J}_n^4\right), \tag{106}$$

where

$$\mathscr{J}_n^i = \begin{cases} \mathscr{I}_n^i & \text{for } n \in [1:m^*-1]\,, \\ \int_{t_{m^*-1}}^{t_m^*} I_n^i(t)\,\mathrm{d}t & \text{for } n = m^*, \end{cases} \tag{107}$$

with I_n^i being the same integrand as that of $\mathscr{I}_n^i$. To prove the lemma we must bound $\sum_{n=1}^m \mathscr{I}_n^k$ and $\sum_{n=1}^{m^*} \mathscr{J}_n^k$ for each $k \in [1:4]$.

4.6. **Time error estimate.** We estimate the term due to time discretization. By the definition of elliptic reconstruction (30) we have

$$\begin{aligned}\mathscr{I}_n^2 &\le \int_{t_{n-1}}^{t_n} \left|a\left(\omega(t)-\omega^n, \partial_t\rho(t)\right)\right| \,\mathrm{d}t \\ &= \int_{t_{n-1}}^{t_n} l_{n-1}(t) \left|\left\langle A^{n-1}U^{n-1} - A^nU^n, \partial_t\rho(t)\right\rangle\right| \,\mathrm{d}t \\ &\le \left(\int_{t_{n-1}}^{t_n} \left\|\partial_t\rho(t)\right\|^2 \,\mathrm{d}t\right)^{1/2} \sqrt{1/3}\left\|A^{n-1}U^{n-1} - A^nU^n\right\| \tau_n^{1/2}.\end{aligned}$$

It follows that

$$\sum_{n=1}^m \mathscr{I}_n^2 \le \sum_{n=1}^m \left(\int_{t_{n-1}}^{t_n} \left\|\partial_t\rho\right\|^2\right)^{1/2} \theta_{2,n}\tau_n^{1/2}. \tag{108}$$

The same bound applies to $\sum_{n=1}^{m^*} \mathscr{J}_n^2$.

4.7. **Spatial error estimate.** The spatial error estimator term can be bounded in a similar way to the one in §3.6. First introduce the auxiliary function $\psi : [0,T] \to \mathrm{H}_0^1(\Omega)$ such that

$$a\left(\chi, \psi(t)\right) = \left\langle \partial_t\rho(t), \chi\right\rangle, \ \forall \chi \in \mathrm{H}_0^1(\Omega). \tag{109}$$

Noting that $\partial_t\epsilon$ is a piecewise constant function of time, and in view of (31), (128)–(130) (with $j=2$), and (127), we can write

$$\begin{aligned}\mathscr{I}_n^1 &= \int_{t_{n-1}}^{t_n} a\left(\partial_t\epsilon^n, \psi(t)\right) \,\mathrm{d}t \\ &= \int_{t_{n-1}}^{t_n} a\left(\partial_t\epsilon^n, \psi(t) - \hat{\Pi}^n\psi(t)\right) \,\mathrm{d}t \\ &= \int_{t_{n-1}}^{t_n} \left\langle \partial R^n, \psi(t) - \hat{\Pi}^n\psi(t)\right\rangle \\ &\quad + \left\langle \partial J^n, \psi(t) - \hat{\Pi}^n\psi(t)\right\rangle_{\check{\Sigma}_n} \,\mathrm{d}t \\ &\le \left(\int_{t_{n-1}}^{t_n} \left\|\partial_t\rho(t)\right\|^2\right)^{1/2} \eta_{2,n}\tau_n^{1/2},\end{aligned} \tag{110}$$

where $\eta_{2,n}$ is defined in (87). Thus, upon summing in time, we conclude that

$$\sum_{n=1}^m \mathscr{I}_n^1 \le \sum_{n=1}^m \left(\int_{t_{n-1}}^{t_n} \left\|\partial_t\rho(t)\right\|^2\right)^{1/2} \eta_{2,n}\tau_n^{1/2}. \tag{111}$$

The same estimate holds for $\sum_{n=1}^{m^*} \mathscr{J}_n^1$.

4.8. **Data approximation and mesh change estimates.** We conclude the estimates in this section by bounding the last two terms in (104) regarding data approximation and mesh changes.

The data space approximation error can be bounded as follows:

$$
\begin{aligned}
\sum_{n=1}^{m} \mathscr{I}_n^3 &= \sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} \left\langle (P_0^n - \mathrm{I})(f^n - \tau_n U^{n-1}), \partial_t \rho \right\rangle \\
&= \sum_{n=1}^{m} \left\langle (P_0^n - \mathrm{I})(f^n - \tau_n U^{n-1}), \rho^n - \rho^{n-1} \right\rangle \\
&= \sum_{n=1}^{m-1} \left\langle (P_0^n - \mathrm{I})(f^n - \tau_n U^{n-1}) - (P_0^{n+1} - \mathrm{I})(f^{n+1} - \tau_{n+1} U^n), \rho^n \right\rangle \\
&\quad + \left\langle (P_0^m - \mathrm{I})(f^m - \tau_m U^{m-1}), \rho^m \right\rangle - \left\langle (P_0^1 - \mathrm{I})(f^1 - \tau_1 U^0), \rho^0 \right\rangle \\
&= \sum_{n=2}^{m} \left\langle \partial \left((P_0^n - \mathrm{I})(f^n - \tau_n U^{n-1}) \right), \rho^{n-1} - \hat{\Pi}^n \rho^{n-1} \right\rangle \tau_n \\
&\quad + \left\langle (P_0^m - \mathrm{I})(f^m - \tau_m U^{m-1}), \rho^m - \Pi^m \rho^m \right\rangle \\
&\quad - \left\langle (P_0^1 - \mathrm{I})(f^1 - \tau_1 U^0), \rho^0 - \Pi^0 \rho^0 \right\rangle .
\end{aligned} \tag{112}
$$

Owing to (128) and (10) we may thus conclude that

$$
\sum_{n=1}^{m} \mathscr{I}_n^3 \le |\rho_*^m|_a \left(\gamma_{\infty,m} + \sum_{n=2}^{m} \gamma_{1,n} \tau_n + \gamma_{\infty,1} \right), \tag{113}
$$

where the estimators $\gamma_{i,n}$ are defined in (88) and (89) for $i = 1$ and ∞, respectively. Likewise we obtain the following bound:

$$
\sum_{n=1}^{m^*} \mathscr{J}_n^3 \le |\rho_*^m|_a \left(\gamma_{\infty,m^*} + \sum_{n=2}^{m^*} \gamma_{1,n} \tau_n + \gamma_{\infty,1} \right). \tag{114}
$$

The last term in (104) is handled in a straightforward way, as in §3.7, and the following bound is readily derived:

$$
\sum_{n=1}^{m} \mathscr{I}_n^4 \le \sum_{n=1}^{m} \left(\int_{t_{n-1}}^{t_n} \|\partial_t \rho\|^2 \right)^{1/2} \beta_{2,n} \tau_n^{1/2}. \tag{115}
$$

Also here, the same bound applies to $\sum_{n=1}^{m^*} \mathscr{J}_n^4$.

4.9. **Proof of Lemma 4.3: Last step.** As in §3.8, we appropriately collect the results from the preceding paragraphs, and we use (104) and (106). We thus have

$$
\begin{aligned}
\frac{1}{2} |\rho_*^m|_a^2 + \int_0^{t_m} \|\partial_t \rho\|^2 &\le \frac{1}{2} |\rho_0|_a^2 + 2 |\rho_*^m|_a \mathscr{E}_{1,m} \\
&\quad + 2 \sum_{n=1}^{m} \left(\int_{t_{n-1}}^{t_n} \|\partial_t \rho\|^2 \right)^{1/2} (\theta_{2,n} + \beta_{2,n} + \eta_{2,n}) \tau_n^{1/2}.
\end{aligned} \tag{116}
$$

We can now proceed by using the same elementary fact used in §3.8, with

$$
a_0 = \frac{1}{\sqrt{2}} |\rho_*^m|_a , \quad a_n = \left(\int_{t_{n-1}}^{t_n} \|\partial_t \rho\|^2 \right)^{1/2}, \quad c = \frac{1}{\sqrt{2}} \left|\rho^0\right|_a , \tag{117}
$$

$$
b_0 = 2\sqrt{2} \mathscr{E}_{1,m}, \quad b_n = 2 (\theta_{2,n} + \beta_{2,n} + \eta_{2,n}) \tau_n^{1/2}, \tag{118}
$$

for $n \in [1 : m]$. Some simple manipulations yield estimate (98). □

4.10. **The difficulty with the direct approach.** We conclude this section by exhibiting the main problem with the direct approach to derive energy estimates in higher norms.

To see this we go back to equation (22) which we now take with $\phi = \partial_t e$ and integrate in time. We are thus required to estimate the term

$$\sum_{n=1}^{m} \int_{t_{n-1}}^{t_n} a\left(U^n, \partial_t e - \Pi_n \partial_t e\right). \tag{119}$$

Since the only practical way to proceed seems to be by decreasing the number of derivatives acting upon $e - \Pi e$—integration by parts in space being of no help here—we perform a "summation by parts" in time as follows:

$$\begin{aligned}
&\sum_{n=1}^{m} \left(a\left(U^n, e^n - \Pi^n e^n\right) - a\left(U^n, e^{n-1} - \Pi^n e^{n-1}\right)\right) \\
&\quad = a\left(U^n, e^m - \Pi^m e^m\right) - a\left(U^0, e^0 - \Pi^1 e^0\right) \\
&\qquad + \sum_{n=1}^{m-1} a\left(U^n - U^{n+1}, e^n - \Pi^n e^n\right) \\
&\qquad - \sum_{n=1}^{m-1} a\left(U^n, (\Pi^n - \Pi^{n+1}) e^n\right).
\end{aligned}$$

The difficulty, which should be apparent now, is how to control the last term. There seems to be no practical way to do this without imposing strong assumptions on $(\Pi^n - \Pi^{n+1})e^n$. Note that this term vanishes if there is no mesh change.

5. Numerical results

We present the results of a series of numerical experiments to exemplify some of the practical aspects of the a posteriori estimates of Theorem 3.2. The main goal here is to approximate the asymptotic behavior of the various estimators and compare this behavior with that of the norms.

5.1. **Benchmark solutions.** We perform the numerical experiment by approximating in each case either one of the following two exact solutions:

$$u(\boldsymbol{x}, t) = \begin{cases} \sin(\pi t)\exp(-10\,|\boldsymbol{x}|^2) & \text{(slow)}, \\ \frac{1}{10}\sin(20\pi t)\exp(-10\,|\boldsymbol{x}|^2) & \text{(fast)}. \end{cases} \tag{120}$$

These solutions are used as benchmarks for the problem with $\boldsymbol{A} = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$. The right-hand side f of the problem is thus easily calculated by applying the parabolic operator to each u. Therefore the exact errors are computable, and we can compare them with the error estimators. The domain on which we compute this solution is the square $[-1, 1]^2$, and the time interval is $[0, 1]$. Note that the boundary conditions are not exactly zero but of the order of 10^{-6} so that special care has to be taken with very small numbers. The initial conditions are exactly zero in both cases, and there is no initial error to be computed.

TABLE 1.

simulation	problem	ℓ	k	h_1	τ_1	I (runs)	Figure
1	slow	1	2	0.5	0.04	6	1
2	fast	1	1	0.25	0.01	5	2
3	slow	1	3	0.125	0.08	4	3
4	fast	2	2	0.125	0.02	4	4

5.2. **Choice of parameters.** Since we are interested in understanding the asymptotic behavior of the estimators, we conduct tests on uniform meshes with uniform timestep. For each numerical experiment we choose a sequence of meshsizes $(h(i) : i \in [1 : I])$, to which we couple a sequence of stepsizes $(\tau(i) : i \in [1 : I])$, $\tau(i) = c_0 h(i)^k$, with k equal either 1 or 2 and I, the number of runs, ranging from 4 to 6 in each of the 4 cases that we run. Table 1 summarizes the choice of the various parameters for each of the 4 simulations.

5.3. **Computed quantities.** Note that we report numerical results for the estimates of §3 only. For each simulation and for each run $i \in [1 : I]$, we calculate the following quantities:

- the error norms
$$\|e\|_{\mathrm{L}_\infty(0,t_m;\mathrm{L}_2(\Omega))} \text{ and } |e|_{\mathrm{L}_2(0,t_m;\mathrm{H}^1(\Omega))},$$
- the reconstruction error estimators
$$\max_0^m \varepsilon_{\infty,n} \text{ and } (\tau(i)\sum_1^m \varepsilon_{2,n}^2)^{1/2},$$
- the space estimator
$$\tau(i)\sum_1^m \eta_{1,n},$$
- and the time estimator
$$\tau(i)\sum_1^m \theta_{1,n},$$

for each time $t_m \in [0 = t_0 : \tau(i) : t_N = 1]$. Of course, all the errors and estimators depend on the run i, but for the sake of conciseness we do not add this index. We deliberately ignore the estimators for data approximation (and mesh change), $\beta_{1,n}$ and $\gamma_{2,n}$, as our examples are designed so as to make these estimators either negligible or comparable with respect to one of those calculated. Indeed, we take the mesh and stepsizes small enough as to resolve the data, and we keep the mesh unchanged across timesteps. Therefore, $\gamma_{2,n}$ can be shown to be of order h^2, and $\beta_{1,n}$ to be of order τ since the function f is smooth enough [20].

For each computed norm or estimator we look at its *experimental order of convergence (EOC)*. The EOC is defined as follows: for a given finite sequence of uniform triangulations $\{\mathscr{T}_{h(i)}\}_{i=1,\dots,I}$ of meshsize $h(i)$, the EOC of a corresponding sequence of some triangulation-dependent quantity $E(i)$ (like an error or an

estimator) is itself a sequence defined by

$$\mathrm{EOC}\, E(i) = \frac{\log(E(i+1)/E(i))}{\log(h(i+1)/h(i))}. \tag{121}$$

Since the timesteps $\tau(i)$ are coupled to $h(i)$, this is well defined.

Finally we look at the *inverse effectivity index* for each error-estimator pair, defined by

$$\frac{\max_{n\in[0:m]} \|e(t_m)\|}{\max_{n\in[0:m]} \varepsilon_{\infty,n} + \tau(i)\sum_{n=1}^{m}(\eta_{1,n}+\theta_{1,n})}, \tag{122}$$

$$\frac{\left(\tau \sum |e(t_m)|_1^2\right)^{1/2}}{\left(\tau \sum_{n=0}^{m} \varepsilon_{2,n}^2\right)^{1/2} + \tau(i)\sum_{n=1}^{m}(\eta_{1,n}+\theta_{1,n})}. \tag{123}$$

The initial estimator is zero in this case, and the data approximation and the mesh change estimators are dropped from this study.

5.4. *Remark* (Effectivity index). Note that we prefer using the inverse rather than the straight effectivity index; the reason for this choice is twofold. First, from a practical aspect, since the effectivity index tends to be quite high in an initial transient time, its inverse is nicer to visualize. Second, and more importantly, since we are interested in obtaining a numerical realization for the constant C appearing in the estimates of the type $\|e\| \le C\mathscr{E}$, where e is the error and $\mathscr{E}$ the estimator, $\|e\|/\mathscr{E}$ is a more straightforward indicator for C than $\mathscr{E}/\|e\|$, which is the (straight) effectivity index.

We also observe that we take all the constants involved in the estimators, including the interpolation constants, to be equal to 1. This, of course, is not true, and a fine tuning of constants should be performed, but since our purpose here is mainly to check that the asymptotic behavior of the error and the estimator is the optimal one, the (inverse) effectivity index is to be understood only qualitatively in this paper. Note also that the time estimator, involving $\theta_{1,n}$, is in the denominator of (122) and (123).

Finally, we point out that from the theory, the inverse effectivity index must have an upper bound that is independent of the problem at hand (i.e., of the solution). In each numerical simulation, though, this upper bound is not necessarily reached (in fact this happens only in the worst-case scenario), so different problems (i.e., different solutions of u) can lead to different inverse effectivity indexes in practice.

5.5. **Conclusions.** The main conclusion of our numerical tests is that the estimators have the optimal rate of convergence which matches that of the error's norm. It is important to note that, in order to exhibit the optimality of the estimators for different norms, a different coupling of the meshsize h and the stepsize τ must be chosen: for $\mathbb{P}^1$ elements, it is necessary to take $\tau \approx h^2$ to see that the $\mathrm{L}_\infty(\mathrm{L}_2)$ error norm has EOC 2 while the $\mathrm{L}_2(\mathrm{H}^1)$ norm has EOC 1 (Figure 1). If the coupling $\tau \approx h$ is taken, for a problem where the time discretization error dominates, such as (120 fast), then both errors have EOC 1 (Figure 2).

The same observations are valid for tests with $\mathbb{P}^2$ elements, albeit the couplings are $\tau \approx h^3$ and $\tau \approx h^2$, respectively, in this case (see Figures 3 and 4).

In a different article [20] we conduct a more thorough numerical experimentation, where mesh changes and data approximation effects are included.

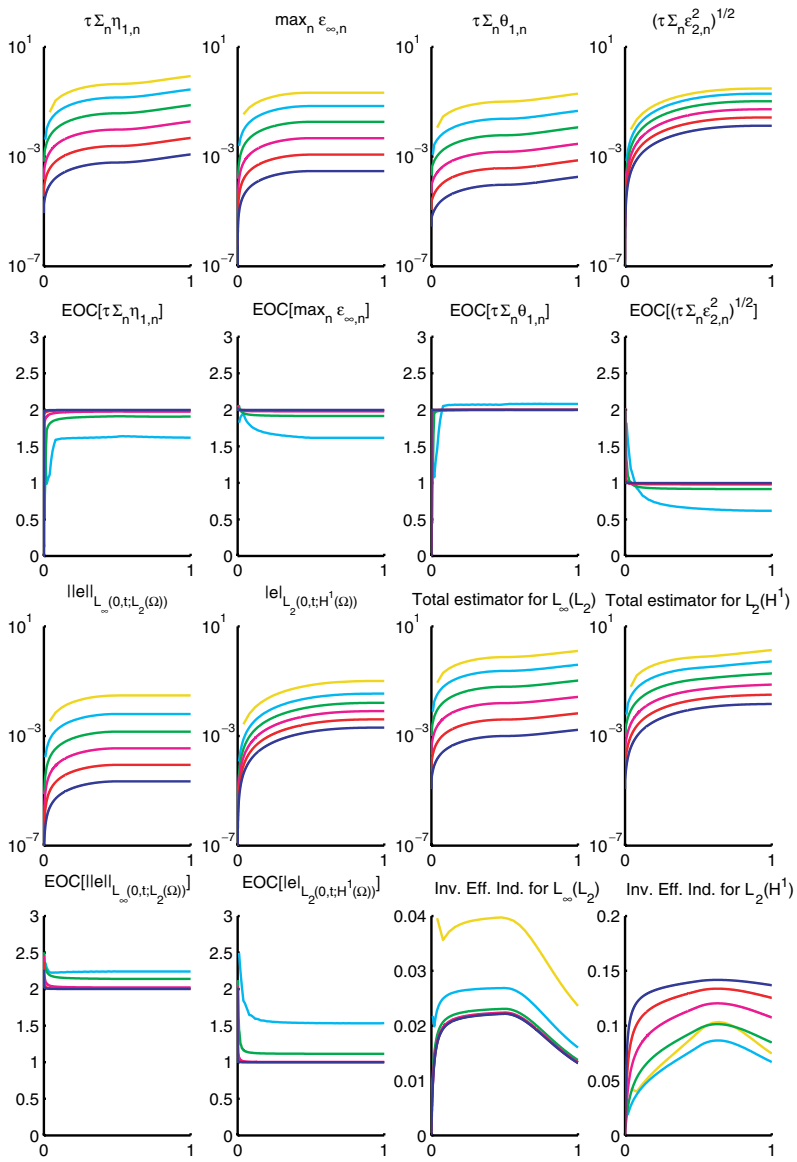

FIGURE 1. Numerical results for a problem with an exact solution ((120) slow) with $\mathbb{P}^1$ elements and $\tau \approx h^2$. The abscissa represents time which ranges in $[0,1]$. In the topmost row we plot the various estimators, and in the second row we show the corresponding EOC's. Note that $\max \varepsilon_{\infty,n}$ has EOC 2 whereas $\left(\tau \sum \varepsilon_{2,n}^2\right)^{1/2}$ has EOC 1. These are the leading terms in the total estimators (the 3 and 4 plots in the 3rd row) and match in EOC, respectively, the $\mathrm{L}_\infty(\mathrm{L}_2)$ error and the $\mathrm{L}_2(\mathrm{H}^1)$ error, as shown in the first 2 plots of the 3rd and 4th rows. Thus (54) and (55) are seen to be sharp and optimal. The last two plots in the 4th row are the inverse effectivity indexes for each norm.

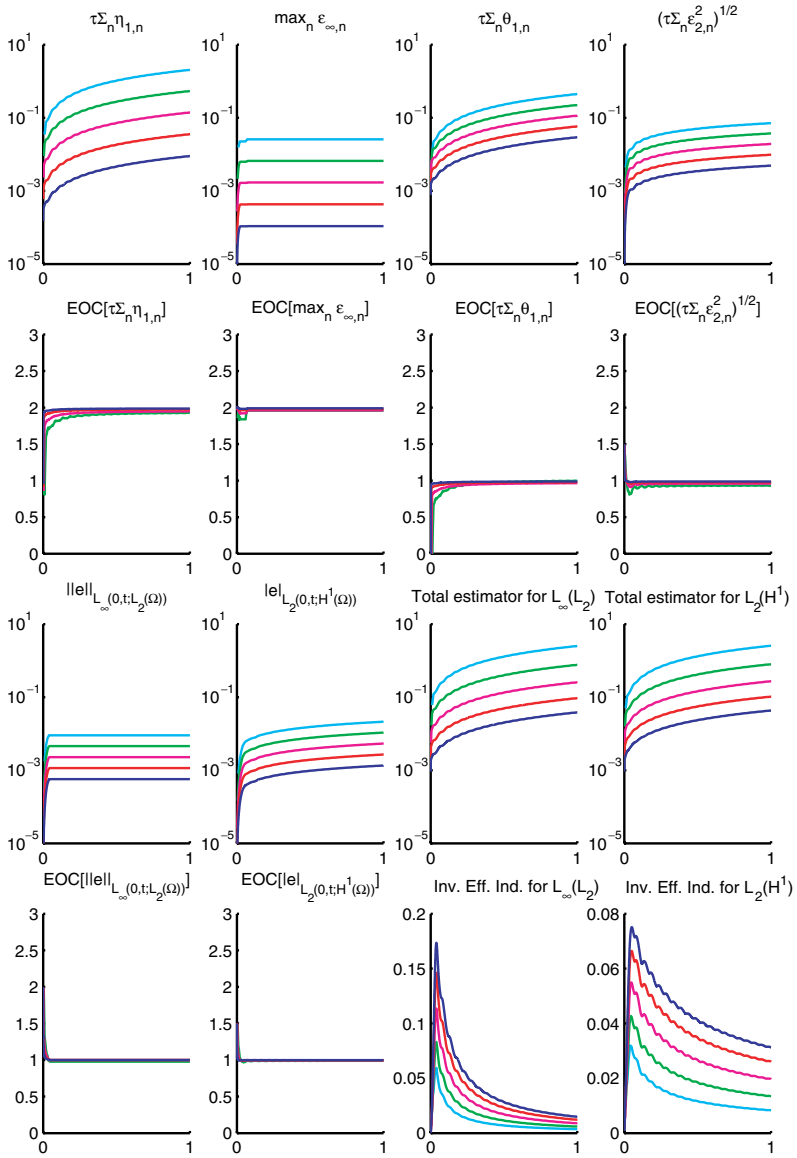

FIGURE 2. Simulation with $\mathbb{P}^1$ elements and $\tau \approx h$. Dominant time discretization error is created by taking the problem with a fast time-oscillating exact solution ((120) fast). The abscissa represents time which ranges in $[0, 1]$. In the topmost row we plot the various estimators, and in the second row we show the corresponding EOC's. Note that $\tau \sum \theta_{\infty,n}$ has EOC 1—reflecting the fact that the error due to time discretization is of order 1—as opposed to 2 in the previous example. This is now a leading term in both total estimators (plots 3 and 4 in the 3rd row) which both have EOC 1. This EOC matches that of the $\mathrm{L}_\infty(\mathrm{L}_2)$ error and of the $\mathrm{L}_2(\mathrm{H}^1)$ error, as shown in plots 1 and 2 of the 3rd and 4th rows. Thus (54) and (55) are both sharp. The last two plots in the 4th row are the inverse effectivity indexes for each norm.

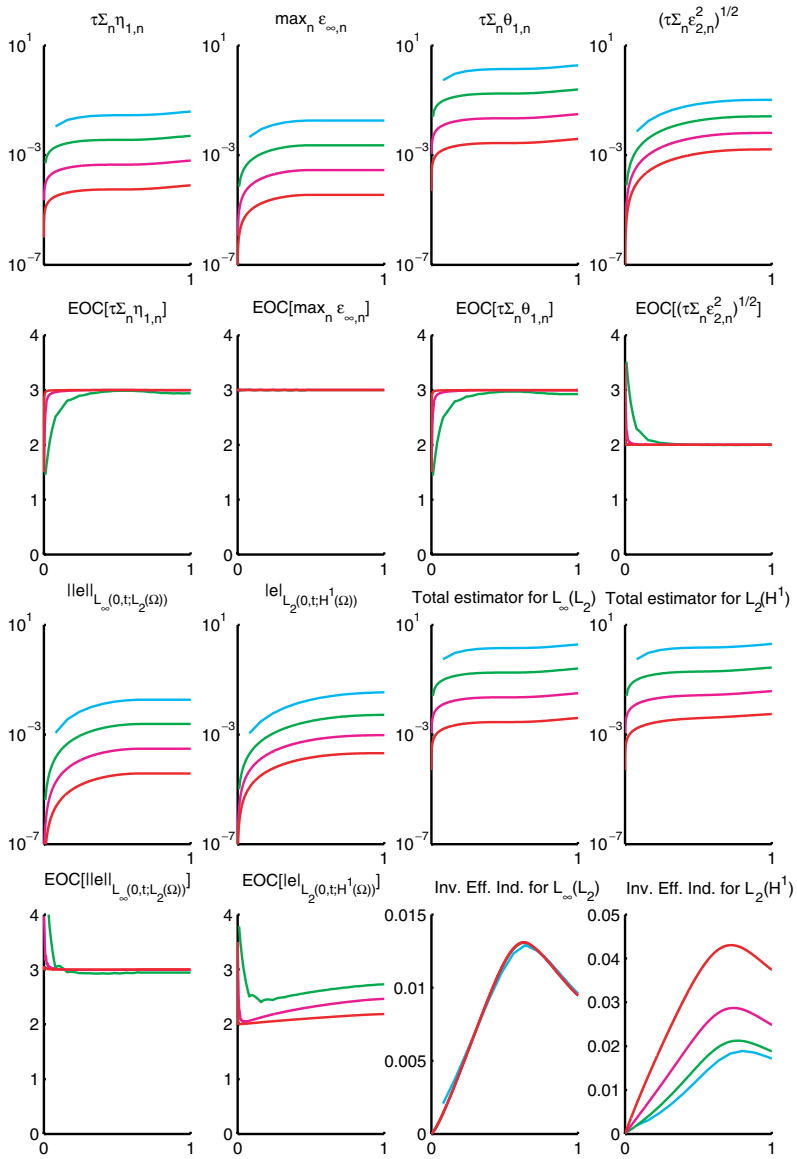

FIGURE 3. Numerical results for a problem with an exact solution ((120) slow) with $\mathbb{P}^2$ elements and $\tau \approx h^3$; i.e., with an error due to space discretization dominant. The abscissa represents time which ranges in $[0, 1]$. In the topmost row we plot the various estimators, and in the second row we show the corresponding EOC's. Note that $\max \varepsilon_{\infty,n}$ has EOC 3, whereas $\left(\tau \sum \varepsilon_{2,n}^2\right)^{1/2}$ has EOC 2. These are the leading terms in the total estimators (plots 3 and 4 in the 3rd row) and match in EOC, respectively, the $\mathrm{L}_\infty(\mathrm{L}_2)$ error and of the $\mathrm{L}_2(\mathrm{H}^1)$ error, as shown in plots 1 and 2 of the 3rd and 4th rows. Here the estimates (54) and (55) are both sharp and optimal. The last two plots in the 4th row are the inverse effectivity indexes for each norm.

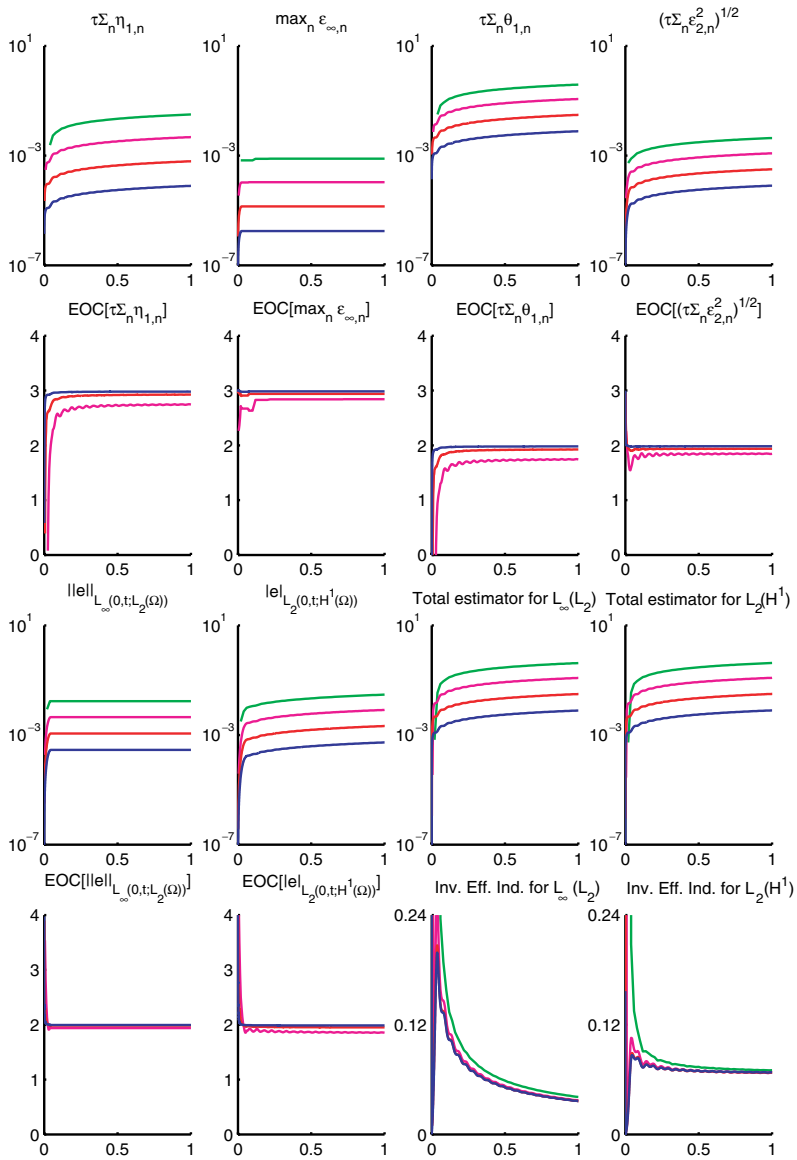

FIGURE 4. Simulation with $\mathbb{P}^2$ elements and coupling $\tau \approx h^2$. The time discretization error is dominant because the exact solution is ((120) fast). The abscissa represents time which ranges in $[0, 1]$. In the topmost row we plot the various estimators, and in the second row we show the corresponding EOC's. Note that $\tau \sum \theta_{\infty,n}$ has EOC 2, as opposed to 3 in the previous example, because of the different coupling of the mesh and step sizes. The time estimator is now a leading term in both total estimators (plots 3 and 4 in the 3rd row), each having EOC 2. This EOC matches that of the $\mathrm{L}_\infty(\mathrm{L}_2)$ error and of the $\mathrm{L}_2(\mathrm{H}^1)$ error, as shown in plots 1 and 2 of the 3rd and 4th rows. Thus (54) and (55) are both sharp. The last two plots in the 4th row are the inverse effectivity indexes for each norm.

Appendix A. Compatible triangulations

Each triangulation $\mathscr{T}_n$, for $n \in [1:N]$, is a refinement of a *macro-triangulation* $\mathscr{M}$ which is a triangulation of the domain Ω that satisfies the same conformity and shape-regularity [8] assumptions made on its refinements in §1. A refinement procedure is *admissible* if it satisfies the following criteria:

1. the refined triangulation is conforming;
2. the shape-regularity of an arbitrarily deep refinement depends only on the shape-regularity of the macro-triangulation $\mathscr{M}$;
3. if $\mathscr{T}$ and $\mathscr{T}'$ are both refinements, then for any two elements $K \in \mathscr{T}$ and $K' \in \mathscr{T}'$,

$$K \cap K' = \emptyset \quad \text{or} \quad K \subset K' \quad \text{or} \quad K' \subset K. \tag{124}$$

Refinement procedures that satisfy these criteria exist. For example, the refinement by bisection described in the ALBERT manual [28], which is known to work for the space dimensions $d = 1, 2, 3, 4$, is admissible for simplex triangulations. All the refinements by bisection of the macro-triangulation $\mathscr{M}$ can be stored in a single binary tree whose nodes represent a simplex.

We say that two triangulations are *compatible* if they are refinements of the same macro-triangulation. A set of compatible triangulations can be endowed with a partial order relation: namely, given two compatible triangulations $\mathscr{T}$ and $\mathscr{T}'$ we write $\mathscr{T} \leq \mathscr{T}'$ if $\mathscr{T}'$ is a refinement of $\mathscr{T}$. This partial ordering permits us to define in the natural way the *coarsest common refinement* of $\mathscr{T}$ and $\mathscr{T}'$, which we denote by $\mathscr{T} \vee \mathscr{T}'$, and the *finest common coarsening*, which we denote by $\mathscr{T} \wedge \mathscr{T}'$. An immediate property of these definitions is

$$\hat{h} = \max(h, h') \quad \text{and} \quad \check{h} = \min(h, h'), \tag{125}$$

where $h, h', \hat{h}, \check{h}$ denote the meshsize of $\mathscr{T}, \mathscr{T}', \mathscr{T} \vee \mathscr{T}', \mathscr{T} \wedge \mathscr{T}'$, respectively.

Appendix B. Inequalities

B.1. **Elliptic regularity.** The a posteriori estimates based in the $\mathrm{L}_2(\Omega)$ norms are based on the duality argument of Aubin and Nitzsche and the elliptic regularity. We therefore assume the coefficient matrix $\boldsymbol{A}$ defining the bilinear form a to be regular enough and Ω to be a convex polygonal as to ensure that there exists a constant $C_{2,2}$ such that if $\phi \in \mathrm{L}_2(\Omega)$ and $\psi \in \mathrm{H}_0^1(\Omega)$ are functions related by the (dual) elliptic problem

$$a(\chi, \psi) = \langle \phi, \chi \rangle, \quad \forall \chi \in \mathrm{H}_0^1(\Omega), \tag{126}$$

then

$$\psi \in \mathrm{H}^2(\Omega) \quad \text{and} \quad |\psi|_2 \leq C_{2,2} \|\phi\|. \tag{127}$$

B.2. **Interpolation inequalities.** We will use the Clément-type interpolation operator $\Pi^n : \mathrm{H}_0^1(\Omega) \to \mathbb{V}^n$ introduced by Scott and Zhang [29] which, under the

needed regularity assumptions of ψ and finite element polynomial degree ℓ, satisfies the following interpolation inequalities for $j \leq \ell + 1$:

$$\left\| h_n^{-j} (\psi - \Pi^n \psi) \right\| \leq C_{3,j} \left| \psi \right|_j , \tag{128}$$

$$\left\| h_n^{1/2-j} (\psi - \Pi^n \psi) \right\|_{\Sigma_n} \leq C_{5,j} \left| \psi \right|_j , \tag{129}$$

where the constants $C_{3,j}$ and $C_{5,j}$ depend only on the shape-regularity of the family of triangulations.

B.3. **Interpolation and mesh change.** Since the triangulations $\mathscr{T}_n$ and $\mathscr{T}_{n-1}$ can be different when adaptive mesh refinement strategies are employed, we introduce $\hat{\Pi}^n$, the Clément-Scott-Zhang interpolator relative to the *finest common coarsening* of $\mathscr{T}_n$ and $\mathscr{T}_{n-1}$, $\hat{\mathscr{T}}_n := \mathscr{T}_n \wedge \mathscr{T}_{n-1}$, whose meshsize is given by $\hat{h}_n := \max(h_n, h_{n-1})$. Then the following inequality holds:

$$\left\| \hat{h}_n^{1/2-j} (\psi - \hat{\Pi}^n \psi) \right\|_{\Sigma_n \cup \Sigma_{n-1} \setminus \Sigma_n \cap \Sigma_{n-1}} \leq C_{7,j} \left| \psi \right|_j , \tag{130}$$

where the constant $C_{7,j}$ depends on the shape-regularity of the triangulations and on the number of refinement steps (bisections) necessary to pass from $\mathscr{T}_n$ to $\mathscr{T}_{n-1}$.

B.4. **Combining constants.** We often use a combination of the constants introduced in this Appendix or throughout the paper. Since many constants appearing in theorems are products of basic constants, our convention is that whenever a constant $C_{k,j}$ appears with the index k being a nonprime integer, then $C_{k,j} = C_{i,j} C_{l,j}$ where $k = il$. If the index k is a prime, then the constant is a "basic" one and is defined in the text. E.g., $C_{6,2} = C_{3,2} C_{2,2}$, etc.

Acknowledgments

We would like to thank Ricardo Nochetto and Andreas Veeser for the stimulating discussions, and both referees for their constructive criticism that lead to important improvements.

References

[1] S. Adjerid, J. E. Flaherty, and I. Babuška. A posteriori error estimation for the finite element method-of-lines solution of parabolic problems. *Math. Models Methods Appl. Sci.*, 9(2):261–286, 1999. MR1674560 (2000a:65117)

[2] M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis.* Wiley-Interscience [John Wiley & Sons], New York, 2000. MR1885308 (2003b:65001)

[3] G. Akrivis, C. Makridakis, and R. H. Nochetto. A posteriori error estimates for the Crank–Nicolson method for parabolic equations. *Math. Comp*, 75(254):511–531, 2006. MR2196979

[4] I. Babuška, M. Feistauer, and P. Šolín. On one approach to a posteriori error estimates for evolution problems solved by the method of lines. *Numer. Math.*, 89(2):225–256, 2001. MR1855826 (2002f:65131)

[5] I. Babuška and S. Ohnimus. A posteriori error estimation for the semidiscrete finite element method of parabolic differential equations. *Comput. Methods Appl. Mech. Engrg.*, 190(35-36):4691–4712, 2001. MR1840797 (2002d:65093)

[6] A. Bergam, C. Bernardi, and Z. Mghazli. A posteriori analysis of the finite element discretization of some parabolic equations. *Math. Comp.*, 74(1117–1138), 2005. MR2136996

[7] D. Braess. *Finite elements*. Cambridge University Press, Cambridge, second edition, 2001. Theory, fast solvers, and applications in solid mechanics, Translated from the 1992 German edition by Larry L. Schumaker. MR1827293 (2001k:65002)

[8] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*. Springer-Verlag, New York, 1994. MR1278258 (95f:65001)

[9] Z. Chen and J. Feng. An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems. *Math. Comp.*, 73(247):1167–1193 (electronic), 2004. MR2047083 (2005e:65131)

[10] P. G. Ciarlet. *The finite element method for elliptic problems*. North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4. MR0520174 (58:25001)

[11] J. de Frutos and J. Novo. A posteriori error estimation with the p-version of the finite element method for nonlinear parabolic differential equations. *Comput. Methods Appl. Mech. Engrg.*, 191(43):4893–4904, 2002. MR1932022 (2003i:65080)

[12] W. Dörfler and M. Rumpf. An adaptive strategy for elliptic problems including a posteriori controlled boundary approximation. *Math. Comp.*, 67(224):1361–1382, 1998. MR1489969 (99b:65141)

[13] T. Dupont. Mesh modification for evolution equations. *Math. Comp.*, 39(159):85–107, 1982. MR0658215 (84g:65131)

[14] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.*, 28(1):43–77, 1991. MR1083324 (91m:65274)

[15] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. II. Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$. *SIAM J. Numer. Anal.*, 32(3):706–740, 1995. MR1335652 (96c:65162)

[16] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. IV. Nonlinear problems. *SIAM J. Numer. Anal.*, 32(6):1729–1749, 1995. MR1360457 (96i:65081)

[17] B. García-Archilla and E. S. Titi. Postprocessing the Galerkin method: the finite-element case. *SIAM J. Numer. Anal.*, 37(2):470–499 (electronic), 2000. MR1740770 (2001h:65112)

[18] J. G. Heywood and R. Rannacher. Finite element approximation of the nonstationary Navier-Stokes problem. II. Stability of solutions and error estimates uniform in time. *SIAM J. Numer. Anal.*, 23(4):750–777, 1986. MR0849281 (88b:65132)

[19] J. G. Heywood and R. Rannacher. Finite element approximation of the nonstationary Navier-Stokes problem. III. Smoothing property and higher order error estimates for spatial discretization. *SIAM J. Numer. Anal.*, 25(3):489–512, 1988. MR0942204 (89k:65114)

[20] O. Lakkis and C. Makridakis. A posteriori error control for parabolic problems: Duality and reconstruction methods. Preprint, Foundation for Research and Technology, Hellas, Heraklion, Greece, 2003.

[21] O. Lakkis and R. H. Nochetto. A posteriori error analysis for the mean curvature flow of graphs. *SIAM J. Numer. Anal.*, 42(5):1875–1898, 2005. MR2139228

[22] X. Liao and R. H. Nochetto. Local a posteriori error estimates and adaptive control of pollution effects. *Numer. Methods Partial Differential Equations*, 19(4):421–442, 2003. MR1980188 (2004c:65130)

[23] C. Makridakis and R. H. Nochetto. A posteriori error analysis of a class of dissipative methods for nonlinear evolution problems. *Preprint*, 2002.

[24] C. Makridakis and R. H. Nochetto. Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Numer. Anal.*, 41(4):1585–1594 (electronic), 2003. MR2034895 (2004k:65157)

[25] R. H. Nochetto, G. Savaré, and C. Verdi. A posteriori error estimates for variable time-step discretizations of nonlinear evolution equations. *Comm. Pure Appl. Math.*, 53(5):525–589, 2000. MR1737503 (2000k:65142)

[26] R. H. Nochetto, A. Schmidt, and C. Verdi. A posteriori error estimation and adaptivity for degenerate parabolic problems. *Math. Comp.*, 69(229):1–24, 2000. MR1648399 (2000i:65136)

[27] M. Picasso. Adaptive finite elements for a linear parabolic problem. *Comput. Methods Appl. Mech. Engrg.*, 167(3-4):223–237, 1998. MR1673951 (2000b:65188)

[28] A. Schmidt and K. G. Siebert. ALBERT—software for scientific computations and applications. *Acta Math. Univ. Comenian. (N.S.)*, 70(1):105–122, 2000. MR1865363

[29] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990. MR1011446 (90j:65021)

[30] R. Verfürth. *A Review of* A Posteriori *Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, Chichester-Stuttgart, 1996.

[31] R. Verfürth. A posteriori error estimates for nonlinear problems. $L^r(0,T;L^\rho(\omega))$-error estimates for finite element discretizations of parabolic equations. *Math. Comp.*, 67(224):1335–1360, 1998. MR1604371 (99b:65120)

[32] R. Verfürth. A posteriori error estimates for nonlinear problems: $L^r(0,T;W^{1,\rho}(\omega))$-error estimates for finite element discretizations of parabolic equations. *Numer. Methods Partial Differential Equations*, 14(4):487–518, 1998. MR1627578 (99g:65099)

Department of Mathematics, University of Sussex, Brighton, UK-BN1 9RF, United Kingdom

E-mail address: O.Lakkis@sussex.ac.uk

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

E-mail address: makr@tem.uoc.gr