

## A COMPARISON OF DUALITY AND ENERGY A POSTERIORI ESTIMATES FOR $L_\infty(0, T; L_2(\Omega))$ IN PARABOLIC PROBLEMS

OMAR LAKKIS, CHARALAMBOS MAKRIDAKIS, AND TRISTAN PRYER

ABSTRACT. We use the elliptic reconstruction technique in combination with a duality approach to prove a posteriori error estimates for fully discrete backward Euler scheme for linear parabolic equations. As an application, we combine our result with the residual based estimators from the a posteriori estimation for elliptic problems to derive space-error indicators and thus a fully practical version of the estimators bounding the error in the  $L_\infty(0, T; L_2(\Omega))$  norm. These estimators, which are of optimal order, extend those introduced by Eriksson and Johnson in 1991 by taking into account the error induced by the mesh changes and allowing for a more flexible use of the elliptic estimators. For comparison with previous results we derive also an energy-based a posteriori estimate for the  $L_\infty(0, T; L_2(\Omega))$ -error which simplifies a previous one given by Lakkis and Makridakis in 2006. We then compare both estimators (duality vs. energy) in practical situations and draw conclusions.

### 1. INTRODUCTION

A posteriori error estimators and their use to derive adaptive mesh refinement algorithms to solve time-dependent problems constitute the object of current research. The problem is appealing for the theoretician as a test ground for novel analytical techniques as well as for those practitioners who are interested in minimizing the amount of computational time in order to obtain satisfactory accuracy in the computer simulations of time-dependent PDE's. Both the theoretical and practical aspects of a posteriori-based adaptive numerical methods for evolution partial differential equations has benefited immensely from the surge in the production of dedicated papers in the last 20 years, although fundamental questions such as convergence of adaptive algorithm remain open.

In this paper, we address the problem of a posteriori error estimation for the time-dependent model problem

$$(1.1) \quad \partial_t u(\mathbf{x}, t) + \mathcal{A}u(\mathbf{x}, t) = f(\mathbf{x}, t)$$

for  $\mathbf{x} \in \Omega \subseteq \mathbb{R}^d$  and  $0 \leq t \leq T$ , where  $\mathcal{A}$  is an elliptic operator to be described in further detail in §2. This simple model of *linear parabolic* PDE has been used as a test-bed for adaptive mesh and timestep refinement methods for

---

Received by the editor September 14, 2012 and, in revised form, July 6, 2013 and November 16, 2013.

2010 *Mathematics Subject Classification*. Primary: 65N30, 65M15, 80M10; Secondary: 65M50, 93C40.

This work was partially supported by the E.U. RTN *Hyke* HPRN-CT-2002-00282 and the Marie Curie Fellowship Foundation. The first author wishes to thank the Hausdorff Institute for Mathematics, Bonn.

The second author was supported at Sussex by a EPSRC D. Phil. postgraduate research fellowship.

evolution equations by [CF04] and [KMSS12] to cite two examples. In both these papers the authors focus on *energy estimates* using the a posteriori estimator for the  $L_2(0, T; H_0^1(\Omega))$ -error as a target for an adaptive optimization procedure known as “indicator equidistribution” in the original spirit of [BR78]. From the discussion in [KMSS12], it turns out that the  $L_2(0, T)$  accumulation in time is problematic when designing adaptive algorithms, due to the loss of causality (which is built-in in timestepping methods) when trying to equidistribute “in the future”. This problem is apparent as well in [CF04] in that the algorithm therein proposed does not guarantee to reach the final time  $T$  under certain tolerance. In [KMSS12], this difficulty is overcome by establishing a priori estimates that provide a priori criteria, based on the data at hand, that will guarantee the termination of the proposed adaptive algorithm. Another way of circumventing the problem of  $L_2(0, T)$  accumulation is to look for error estimators in other functional spaces, say the  $L_\infty(0, T; L_2(\Omega))$ . While the  $L_\infty(0, T; L_2(\Omega))$  norm appears as a “companion norm” to  $L_2(0, T; H_0^1(\Omega))$  in *energy-based a posteriori estimates* [Pic98, Ver03, BBM05], the estimators turn out to be suboptimal in the spatial error because the  $H_0^1(\Omega)$  terms must have lower order. Sharp error estimates in the  $L_\infty(0, T; L_2(\Omega))$ -norm were first derived by [EJ91], using a duality argument. In [LM06], we showed that sharp  $L_\infty(0, T; L_2(\Omega))$ -norm a posteriori error estimates were possible with *energy arguments* if one uses the *elliptic reconstruction* introduced by [MN03]. The elliptic reconstruction has since then been used as an analytical tool, in combination with energy or other techniques, such as heat-kernel or semigroup techniques, to deal with time, in order to establish estimates in various norms for linear and nonlinear problems; see, e.g., [DLM09, DM10, EM09, GL10, GLV11, BM11].

Prior to the introduction of the elliptic reconstruction,  $L_\infty(0, T; L_2(\Omega))$ -error estimates could be derived by using the *duality technique*, at the cost of assuming restrictive assumptions on the domain (e.g., convexity) as well as on the mesh [EJ91]. Our chief goal in this paper is twofold:

1. We explore the possibility of using the *elliptic reconstruction* technique in conjunction with the *duality* technique as introduced by Eriksson and Johnson [EJ91];
2. We compare *duality estimates* with *energy estimates* for the  $L_\infty(0, T; L_2(\Omega))$ -norm.

To our knowledge, both these objectives have not been treated previously in the literature. The *duality* technique provides an important alternative to energy techniques and is widely used for the derivation of a priori and a posteriori error estimates both for elliptic and parabolic problems. Since being first considered, it has been developed in many different directions, including its use in *implicit and goal oriented a posteriori error estimates*.

The *elliptic reconstruction* has been used in combination with energy estimates, where one mimics the energy estimates for the parabolic equation in order to derive error estimates from a PDE where the error, or part thereof, is the “unknown”. In this paper, the use of the elliptic reconstruction is crucial in providing two simple abstract results for both the energy-based and the duality-based estimates in the  $L_\infty(0, T; L_2(\Omega))$ -norm. From a technical analysis viewpoint, we show that the elliptic reconstruction technique can be completely decoupled from energy considerations (or any other method used to deal with time integration and timestepping, for that matter). This is not obvious; indeed, in many works

a posteriori analysis, the elliptic part is entangled with the parabolic part and there is not a clear cut difference between elliptic and parabolic effects. As noted in recent work on a posteriori analysis for time-dependent problems (see, e.g., [AMN06, BBM05, BV04, dFN02, Pic98]) understanding the splitting between the elliptic, stationary, and parabolic, time-dependent, errors, as well as the part of the error where these effects are coupled, is important in designing adaptive methods and avoiding repetition.

An important by-product of our approach is that the mesh-change in time is considered as part of the proofs of our theorems. Indeed, unlike former derivations a posteriori error estimates via duality (mainly [EJ91]), we do not impose on the mesh any assumptions that are susceptible of violation in a practical implementation of the scheme, such as the no-refinement assumptions.

From a more practical side, we give an application of our theory, by comparing in a series of benchmarks where elliptic  $L_2(\Omega)$  residual-based estimators are used [AO00, LN03]. We emphasize, however, that our results are not limited to the use of residual-based estimators and that other estimators which work for the  $L_2(\Omega)$  norms in elliptic problems could be used; see e.g., [LP12].

Our main results in this paper are duality-based estimates, Theorem 4.5 and Corollary 4.6, an energy-based estimate, Theorem 6.4 and a computer experiment designed to compare in practice both estimators. From a theoretical perspective, Theorem 4.5 and Corollary 4.6, generalize the duality estimates of [EJ91], both by providing a wider choice of elliptic error estimators (other than residual being possible) and removing unrealistic assumptions on the meshes. A direct application of the duality-estimate Theorem 4.5 provides finer estimates with respect to time accumulation. This is especially helpful in situations where the error (on a time-invariant mesh) decreases with time and for long-time integration. Finally, energy-estimate Theorem 6.4 is, to the best of our knowledge, a novel result, which simplifies, by using the Poincaré–Friedrichs inequality, special cases from [LM06] and provides the basis for our comparison. The numerical results show that the estimators behave roughly the same, with a slight edge for the energy-based ones when it comes to time accumulation and long-time integration. This is a confirmation of the theoretical observation that the “tails” of the coefficients for the time-accumulation are much heavier for the duality estimators (see Figure 1). The energy estimator benefits from an exponential decay in these coefficients which also provides a faster way of computing them and a more economical storage. In practice, testing the estimators with residual-based elliptic ones, we found that energy estimators are slightly better suited for long-time integration (due to the better accumulation weights), while the duality-based ones can be more precise on short-times where error measures in  $L_\infty(0, T; L_2(\Omega))$  is important. Numerics show also that the logarithmic accumulation weights found in the duality analysis are in fact too pessimistic, and can be safely ignored, while the energy exponential weights are effective. On the other hand, the energy estimator’s “constants” in the exponential weights depend on the Poincaré–Friedrichs constant, which may in turn be pessimistic in some cases.

In this paper we focus on proving upper error-estimator bounds only. Lower bounds are important (although not essential to prove convergence as previously thought, as some recent results in adaptive schemes for elliptic equations show, e.g., [CKNS08, Remark 4.2]). In principle, partial lower bounds, excluding data

oscillation terms and global time-effects, in the spirit of [Ver03, eq. (1.6)] are possible for  $L_2$ -norms. Indeed, the elliptic technique traced in [Ver96] and recently recalled in [DS11], should be applicable to energy estimates. However, lower bounds, because of their local nature, are not relevant when comparing between duality and energy analysis methods (which is the main goal in this paper), and we therefore leave their derivation as an interesting open problem.

The rest of this article is organized as follows: In §2 we recall the main tools related to the elliptic reconstruction. In §3 we analyze the spatially semidiscrete scheme using a duality approach. In §4 we extend the §3 to the fully discrete scheme and in §5 give the proof of those results. In §6 we state and prove the estimates based on the energy approach. Finally, in §7 we summarize our computer experiments from which we drew the main practical conclusions of this research.

## 2. THE DISCRETE SCHEME AND THE ELLIPTIC RECONSTRUCTION

In this section we introduce the numerical schemes that we study, some basic tools including the definition of the elliptic reconstruction.

**2.1. Basic set-up.** We introduce next the PDE whose discretization is the object of this paper. Let  $\Omega$  be a bounded domain of the Euclidean space  $\mathbb{R}^d$ , for some fixed positive integer *space dimension*  $d$  and a *final time*  $T \in \mathbb{R}^+$ . We shall assume throughout this paper's discussion that the domain  $\Omega$  is a convex polytope, noticing that all the results can be extended to certain non-convex domains, like domains with reentrant corners in  $d = 2$ , following ideas of [LN03] regarding the elliptic a posteriori  $L_2(\Omega)$ -error estimates.

Given a Lebesgue measurable set  $D \subset \mathbb{R}^d$ , we define

$$(2.1) \quad \langle \phi, \psi \rangle_D := \int_D \phi(\mathbf{x})\psi(\mathbf{x})\mu(d\mathbf{x}),$$

$$(2.2) \quad \|\phi\|_D := \|\phi\|_{L_2(D)} := \langle \phi, \phi \rangle_D^{1/2},$$

$$(2.3) \quad |\phi|_{k,D} := \|\mathbf{D}^k \phi\|_D, \text{ for } k \in \mathbb{Z}^+,$$

$$(2.4) \quad \|\phi\|_{k,D} := \left( \|\phi\|_D^2 + \sum_{j=1}^k |\phi|_{j,D}^2 \right)^{1/2}, \text{ for } k \in \mathbb{Z}^+,$$

where  $\mu(d\mathbf{x})$  denotes either the Lebesgue measure element  $d\mathbf{x}$ , when  $D$ 's such measure is positive, or the  $(d-1)$ -dimensional (Hausdorff) measure  $ds\mathbf{x}$ , when  $D$  has zero Lebesgue measure. In many instances, in order to compress notation and when there is no danger of engendering confusion, we may drop altogether the ‘‘differential’’ symbol from integrals. This convention applies also to integrals in time.

We will use the standard [Eva98] function spaces  $L_2(D)$ ,  $H^k(D)$ ,  $H_0^k(D)$  and denote by  $H^{-1}(D)$  the dual space of  $H_0^1(D)$  with the corresponding pairing written as  $\langle \cdot | \cdot \rangle_D$ . We omit the subscript  $D$  whenever  $D = \Omega$ . We denote the Poincaré–Friedrichs constant associated with  $\Omega$  by  $C_{PF}$  and we take the seminorm  $|\cdot|_1$  to be the norm of  $H_0^1(\Omega)$ . We use the usual duality identification

$$(2.5) \quad H_0^1(\Omega) \subset L_2(\Omega) \sim L_2(\Omega)' \subset H^{-1}(\Omega)$$

and the dual norm

$$(2.6) \quad \|\psi\|_{-1} := \sup_{0 \neq \phi \in \mathbf{H}_0^1(\Omega)} \frac{\langle \psi | \phi \rangle}{|\phi|_1} \left( = \sup_{0 \neq \phi \in \mathbf{H}_0^1(\Omega)} \frac{\langle \psi, \phi \rangle}{|\phi|_1}, \text{ if } \psi \in L_2(\Omega) \right).$$

Let  $a$  be the elliptic bilinear form defined on  $\mathbf{H}_0^1(\Omega)$  by

$$(2.7) \quad a(v, \psi) := \langle \mathbf{A} \nabla v, \nabla \psi \rangle \quad \forall v, \psi \in \mathbf{H}_0^1(\Omega),$$

where “ $\nabla$ ” denotes the spatial gradient, and the matrix-valued function

$$\mathbf{A} \in L_\infty(\Omega)^{d \times d}$$

is such that

$$(2.8) \quad a(\psi, \phi) \leq \beta |\psi|_1 |\phi|_1 \quad \forall \phi, \psi \in \mathbf{H}_0^1(\Omega),$$

$$(2.9) \quad a(\phi, \phi) \geq \alpha |\phi|_1^2 \quad \forall \phi \in \mathbf{H}_0^1(\Omega),$$

with  $\alpha, \beta \in \mathbb{R}^+$ . We also use the *energy norm*  $|\cdot|_a$  defined as

$$(2.10) \quad |\phi|_a := a(\phi, \phi)^{1/2} \quad \forall \phi \in \mathbf{H}_0^1(\Omega).$$

It is equivalent to the norm  $|\cdot|_1$  on the space  $\mathbf{H}_0^1(\Omega)$ , in view of (2.8) and (2.9). In particular, we will often use the following inequality:

$$(2.11) \quad |\phi|_1 \leq \alpha^{-1/2} |\phi|_a \quad \forall \phi \in \mathbf{H}_0^1(\Omega).$$

Let  $u \in L_\infty(0, T; \mathbf{H}_0^1(\Omega))$ , with  $\partial_t u \in L_2(0, T; \mathbf{H}^{-1}(\Omega))$ , be the unique solution of the linear parabolic problem

$$(2.12) \quad \begin{aligned} \langle \partial_t u | \phi \rangle + a(u, \phi) &= \langle f, \phi \rangle \quad \forall \phi \in \mathbf{H}_0^1(\Omega), \\ \text{and } u(0) &= g, \end{aligned}$$

where  $f \in L_2(\Omega \times (0, T))$  and  $g \in \mathbf{H}_0^1(\Omega)$ . Whenever not stated explicitly, we assume that the data  $f, g, \mathbf{A}$  and the solution  $u$  of the above problem are sufficiently regular for all the norms involved to make sense.

In order to discretize the time variable in (2.12), we introduce the partition  $0 = t_0 < t_1 < \dots < t_N = T$  of  $[0, T]$ . Let  $I_n := (t_{n-1}, t_n]$  and we denote by  $\tau_n := t_n - t_{n-1}$  the timesteps. We will consistently use the following “superscript convention”: whenever a function depends on time, e.g.,  $f(\mathbf{x}, t)$ , and the time is fixed to be  $t = t_n$ ,  $n = 0, \dots, N$  we denote it by  $f^n(\mathbf{x})$ . Moreover, we often drop the space dependence explicitly, e.g., we write  $f(t)$  and  $f^n$  in reference to the previous sentence.

We use a conforming fixed polynomial degree FEM to discretize the space variable. Let  $(\mathcal{T}_n)_{n=0, \dots, N}$  be a family of conforming triangulations of the domain  $\Omega$  [BS07, Cia78]. These triangulations are allowed to change at each timestep, as long as they stay *compatible*. By *compatible meshes* we mean that they all descend, as nested refinements by bisection from the same macro-triangulation. This is an extremely mild requirement automatically implemented by many refinement methods [Kos94, for details on compatible meshes; SS05; LM06].

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

$$(2.13) \quad h_n(\mathbf{x}) = \text{diam}(K), \text{ where } K \in \mathcal{T}_n \text{ and } \mathbf{x} \in K,$$

for all  $\mathbf{x} \in \Omega$ . We also denote by  $\mathcal{S}_n$  the set of *internal sides* of  $\mathcal{T}_n$ , these are edges in  $d = 2$ —or faces in  $d = 3$ —that are contained in the interior of  $\Omega$ ; the *interior*

*mesh of edges*  $\Sigma_n$  is then defined as the union of all internal sides  $\bigcup_{E \in \mathcal{T}_n} E$ . We associate with these triangulations the *finite element spaces*

$$(2.14) \quad \mathbb{V}^n := \{ \phi \in H_0^1(\Omega) : \forall K \in \mathcal{T}_n : \upharpoonright \phi|_K \in \mathbb{P}^\ell \},$$

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  $\mathcal{T}_{n-1}$  and  $\mathcal{T}_n$ , we define  $\hat{h}_n := \max(h_n, h_{n-1})$  [LM06, Appendix]. We will also use the sets  $\hat{\Sigma}_n := \Sigma_n \cap \Sigma_{n-1}$  and  $\tilde{\Sigma}_n := \Sigma_n \cup \Sigma_{n-1}$ . Since the definitions in some parts of this section are independent of the time discretization and could be applied to any finite element space, in this section we use two generic  $H_0^1(\Omega)$ -conforming Lagrange finite element spaces  $\mathbb{V}$  and  $\mathbb{W}$ . Whenever  $\mathbb{V}$  or  $\mathbb{W}$  coincides with one of the  $\mathbb{V}^n$  introduced, we replace all indexes  $\mathbb{V}$  by  $n$ .

**2.2. Definition (fully discrete scheme).** We consider the following fully discrete scheme of problem (2.12) associated with the finite element spaces  $\mathbb{V}^n$ :

$$(2.15) \quad \begin{aligned} U^0 &:= I^0 u(0), \text{ and} \\ \tau_n^{-1} \langle U^n - U^{n-1}, \Phi_n \rangle + a(U^n, \Phi_n) &= \langle \tilde{f}^n, \Phi_n \rangle \quad \forall \Phi_n \in \mathbb{V}^n, \text{ for } n = 1, \dots, N. \end{aligned}$$

Here the operator  $I^0$  is some suitable interpolation or projection operator from  $H_0^1(\Omega)$  or  $L_2(\Omega)$ , onto  $\mathbb{V}^n$ , and  $\tilde{f}^n$  equals either the value of  $f$  at  $t_n$ ,  $f^n := f(\cdot, t_n)$  or its time-average on  $I_n$ ,  $\int_{t_{n-1}}^{t_n} f(\cdot, t) dt / (t_n - t_{n-1})$ . This scheme is the standard backward (or implicit) Euler–Galerkin finite element scheme [Tho06]. Below we shall use a continuous piecewise linear extension in time of the sequence  $(U^n)$  which we denote by  $U(t)$  for  $t \in [0, T]$  (see §2.9 for the precise definition).

**2.3. A posteriori estimates and reconstruction operators.** The elliptic reconstruction, as described by [MN03] consists in associating with  $U : [0, T] \rightarrow \mathbb{V}$  an auxiliary function  $\omega : [0, T] \rightarrow H_0^1(\Omega)$ , in such a way that when the *total error*

$$(2.16) \quad e := U - u$$

is decomposed as

$$(2.17) \quad e = \rho - \epsilon,$$

$$(2.18) \quad \epsilon := \omega - U, \quad \rho := \omega - u,$$

then the following properties are satisfied:

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

Therefore in order to successfully apply this idea we must select a suitable reconstructed function  $\omega$ . In our case, this choice is dictated by the elliptic operator at hand; the precise definition is given in §2.5. In addition, the effect of mesh modification will reflect in the right-hand side of the equation for  $\rho$ . As a result of our choice for  $\omega$  we are able to derive optimal order estimators for the error in  $L_\infty(0, T; L_2(\Omega))$ , as well as in  $L_\infty(0, T; H_0^1(\Omega))$  and  $H^1(0, T; L_2(\Omega))$ . In addition, our choosing  $\omega$  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  $\epsilon$  which will be referred to as the *elliptic reconstruction error* whereas the time approximation error information is conveyed by  $\rho$ , a fact that motivates the name *main parabolic error* for this term. This “splitting” of the error is already apparent in the spatially discrete case [MN03].

**2.4. Definition (representation of the elliptic operator, discrete elliptic operator, projections).** Suppose a function  $W \in \mathbb{W}$ , the bilinear form can be then represented as

$$(2.19) \quad a(W, \phi) = \sum_{K \in \mathcal{T}} \langle -\operatorname{div}[\mathbf{A}\nabla W], \phi \rangle_K + \sum_{E \in \mathcal{S}} \langle J[W], \phi \rangle_E \quad \forall \phi \in H_0^1(\Omega),$$

where  $J[W]$  is the *spatial jump of the field  $\mathbf{A}\nabla W$  across an element side  $E \in \mathcal{S}$*  defined as

$$(2.20) \quad \begin{aligned} \uparrow J[W]E(\mathbf{x}) &= [[\mathbf{A}\nabla W]]_E(\mathbf{x}) \\ &:= \lim_{\epsilon \rightarrow 0} [\mathbf{A}(\mathbf{x})\nabla W(\mathbf{x} + \epsilon \boldsymbol{\nu}_E(\mathbf{x})) - \mathbf{A}(\mathbf{x})\nabla W(\mathbf{x} - \epsilon \boldsymbol{\nu}_E(\mathbf{x}))] \cdot \boldsymbol{\nu}_E(\mathbf{x}), \end{aligned}$$

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

Since we use the representation (2.19) 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,  $W \in \mathbb{W}$  (or more generally for any Lipschitz continuous function  $w$  that is  $C^2(\operatorname{int}(K))$ , for each  $K \in \mathcal{T}$ ), denote by  $\mathcal{A}_{\text{el}}W$  the *regular part* of the distribution  $-\operatorname{div}[\mathbf{A}\nabla W]$ , which is defined as a piecewise continuous function such that

$$(2.21) \quad \langle \mathcal{A}_{\text{el}}W, \phi \rangle = \sum_{K \in \mathcal{T}} \int_K -\operatorname{div}[\mathbf{A}(\mathbf{x})\nabla W(\mathbf{x})]\phi(\mathbf{x}) \, d\mathbf{x} \quad \forall \phi \in H_0^1(\Omega).$$

The operator  $\mathcal{A}_{\text{el}}$  is sometime referred to, in the finite element community, as the *elementwise elliptic operator*, as it can be viewed as the result of the application of  $-\operatorname{div}[\mathbf{A}\nabla \cdot]$  only on the interior of each element  $K \in \mathcal{T}$ . Although this is a misnomer (as the operator itself does not depend in any way on the finite element space) this observation justifies our subscript in the notation. We shall write the representation (2.19) in the shorter form

$$(2.22) \quad a(W, \phi) = \langle \mathcal{A}_{\text{el}}W, \phi \rangle + \langle J[W], \phi \rangle_{\Sigma} \quad \forall \phi \in H_0^1(\Omega),$$

where  $\Sigma = \bigcup_{E \in \mathcal{S}} E$ .

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  $\mathbb{V}$  is the operator  $A^{\mathbb{V}} : \mathbb{V} \rightarrow \mathbb{V}$  defined by

$$(2.23) \quad \langle A^{\mathbb{V}}W, \Phi \rangle = a(W, \Phi) \quad \forall \Phi \in \mathbb{V},$$

for  $W \in \mathbb{V}$ .

The  $L_2(\Omega)$ -*projection operator* is defined as the operator  $P_0^{\mathbb{V}} : L_2(\Omega) \rightarrow \mathbb{V}$  such that

$$(2.24) \quad \langle P_0^{\mathbb{V}}v, \Phi \rangle = \langle v, \Phi \rangle \quad \forall \Phi \in \mathbb{V},$$

for  $v \in L_2(\Omega)$ ; and the *elliptic projection operator*  $P_1^{\mathbb{V}} : H_0^1(\Omega) \rightarrow \mathbb{V}$  is defined by

$$(2.25) \quad a(P_1^{\mathbb{V}}v, \Phi) = a(v, \Phi) \quad \forall \Phi \in \mathbb{V}.$$

**2.5. Definition (elliptic reconstruction).** We define the *elliptic reconstruction operator* associated with the bilinear form  $a$  and a given finite element space  $\mathbb{V}$  to be the unique linear operator  $\mathcal{R}^{\mathbb{V}} : \mathbb{V} \rightarrow \mathbf{H}_0^1(\Omega)$  such that

$$(2.26) \quad a(\mathcal{R}^{\mathbb{V}}W, \phi) = \langle A^{\mathbb{V}}W, \phi \rangle \quad \forall \phi \in \mathbf{H}_0^1(\Omega),$$

for each given  $W \in \mathbb{V}$  where  $A^{\mathbb{V}}$  is the discrete elliptic operator defined in (2.23). The  $\mathbf{H}_0^1(\Omega)$  function  $\mathcal{R}^{\mathbb{V}}W$  is referred to as the *elliptic reconstruction* of  $W$ .

Note that the domain of the reconstruction operator  $\mathcal{R}^{\mathbb{V}}$ , as well as that of the discrete elliptic operator  $A^{\mathbb{V}}$ , may be extended (constructively) to all of  $\mathbf{H}_0^1(\Omega)$ , but since  $\mathcal{R}^{\mathbb{V}}$  will be used effectively on the finite element space and we generally consider its restriction to  $\mathbb{V}$ . It is also worth noting that elliptic reconstruction operator  $\mathcal{R}^{\mathbb{V}}$ , is a right, but not left, inverse of the well-known elliptic (or Ritz) projection discussed in [Whe73] (see also [Tho06]).

**2.6. Remark (Galerkin orthogonality).** A crucial property of the elliptic reconstruction operator  $\mathcal{R}^{\mathbb{V}}$  is that for  $v \in \mathbf{H}_0^1(\Omega)$ ,  $v - \mathcal{R}^{\mathbb{V}}v$  is  $a(\cdot, \cdot)$ -orthogonal to  $\mathbb{V}^n$ , i.e.,

$$(2.27) \quad a(v - \mathcal{R}^{\mathbb{V}}v, \Phi) = 0 \quad \forall \Phi \in \mathbb{V}^n.$$

This is known as the *Galerkin orthogonality* of the error in the finite element literature and is the crucial property that allows to obtain a priori and a posteriori error estimates.

**2.7. Definition (elliptic a posteriori error estimator functional).** Given a normed functional space  $\mathcal{V}$  containing  $\mathbf{H}_0^1(\Omega)$ , (e.g.,  $\mathcal{V} = \mathbf{L}_2(\Omega)$  or  $\mathbf{H}_0^1(\Omega)$ ) and a generic finite dimensional subspace  $\mathbb{V}$ , we call *estimator functional* associated with the bilinear form  $a$ , defined in (2.7), the space  $\mathbb{V}$  in the the norm of  $\mathcal{V}$ , a functional of the form

$$(2.28) \quad \mathcal{E}[\cdot, \mathbb{V}, \mathcal{V}] : \mathbb{V} \rightarrow \mathbb{R}$$

such that for each  $V \in \mathbb{V}$  we have

$$(2.29) \quad \|V - \mathcal{R}^{\mathbb{V}}V\|_{\mathcal{V}} \leq \mathcal{E}[V, \mathbb{V}, \mathcal{V}].$$

Thanks to many different techniques [AO00, Bra01, Ver96], it is well known that there exist many such functionals. One of the simplest examples is given by the residual-based estimator functional, justified next by Lemma 2.8, which we will use in this work, but we note that our approach can be easily adapted to accommodate other estimators.

**2.8. Lemma (residual-based a posteriori error estimates).** *Let  $\mathbb{V}$  be a finite element space on a triangulation  $\mathcal{T}$  with edge set  $\Sigma$  of the polygonal domain  $\Omega$  as defined in §2.1. For any  $V \in \mathbb{V}$  we have*

$$(2.30) \quad \|\mathcal{R}^{\mathbb{V}}V - V\|_1 \leq \frac{C_1}{\alpha} \|(\mathcal{A}_{\text{el}}V - A^nV)h_n\| + \frac{C_2}{\alpha} \|J[V]h_n^{1/2}\|_{\Sigma_n},$$

and, if furthermore  $\Omega$  is convex, then

$$(2.31) \quad \|\mathcal{R}^{\mathbb{V}}V - V\| \leq C_3 \|(\mathcal{A}_{\text{el}}V - A^nV)h_n^2\| + C_4 \|J[V]h_n^{3/2}\|_{\Sigma_n},$$

for the  $\alpha$  given by (2.11) and some ( $V$ -independent) constants  $C_j$ ,  $j = 1, 2, 3, 4$ .



**2.9. 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.,

$$(2.32) \quad U(t) := l_{n-1}(t)U^{n-1} + l_n(t)U^n, \text{ for } t \in I_n \text{ and } n = 1, \dots, N;$$

where the functions  $l_n$  are the hat (linear Lagrange basis) functions defined by

$$(2.33) \quad l_n(t) := \frac{t - t_{n-1}}{\tau_n} \mathbb{1}_{I_n}(t) - \frac{t - t_{n+1}}{\tau_{n+1}} \mathbb{1}_{I_{n+1}}(t), \text{ for } t \in [0, T] \text{ and } n = 0, \dots, N,$$

$\mathbb{1}_X$  denoting the characteristic function of the set  $X$ . This extension is for a posteriori analysis purposes, it is not needed for computations.

In the sequel will use the following shorthand

$$(2.34) \quad \omega^\vee = \mathcal{R}^\vee U \quad (\text{and thus } \omega^n = \mathcal{R}^n U^n),$$

to denote the elliptic reconstruction of the (semi-)discrete solution  $U$  and  $U^n$ .

The *time-dependent elliptic reconstruction* of  $U$  is the function

$$(2.35) \quad \omega(t) := l_{n-1}(t)\mathcal{R}^{n-1}U^{n-1} + l_n(t)\mathcal{R}^n U^n, \text{ for } t \in I_n \text{ and } n = 1, \dots, N,$$

which results in a Lipschitz continuous function of time.

We introduce next time-discrete derivative (i.e., difference) operators:

(a) *Discrete (backward) time derivative*

$$(2.36) \quad \partial U^n := \frac{U^n - U^{n-1}}{\tau_n}.$$

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

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

$$(2.37) \quad \partial^2 U^n := \frac{\partial U^{n+1} - \partial U^n}{\tau_n}.$$

(c) *Averaged ( $L_2(\Omega)$ -projected) discrete time derivative*

$$(2.38) \quad \bar{\partial} U^n := P_0^n \partial U^n = \frac{U^n - P_0^n U^{n-1}}{\tau_n} \quad \forall n = 1, \dots, N.$$

This last definition stems from  $\partial U^n$  not necessarily belonging to  $\mathbb{V}^n$  (e.g., when  $\mathbb{V}^{n-1} \not\subseteq \mathbb{V}^n$ ), whereas  $\bar{\partial} U^n \in \mathbb{V}^n$  is always satisfied.

**2.10. Remark (pointwise form).** The discrete elliptic operators  $A^n$  can be employed to write the fully discrete scheme (2.15) in the following *pointwise form*:

$$(2.39) \quad \bar{\partial} U^n(\mathbf{x}) + A^n U^n(\mathbf{x}) = P_0^n \tilde{f}^n(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega.$$

Indeed, in view of  $\bar{\partial} U^n + A^n U^n - P_0^n \tilde{f}^n \in \mathbb{V}^n$ , (2.15), and (2.23), we have

$$(2.40) \quad \begin{aligned} \langle A^n U^n + \bar{\partial} U^n - P_0^n \tilde{f}^n, \phi \rangle &= \langle A^n U^n + \bar{\partial} U^n - P_0^n \tilde{f}^n, P_0^n \phi \rangle \\ &= a(U^n, P_0^n \phi) + \langle \tau_n^{-1}(U^n - U^{n-1}) - \tilde{f}^n, P_0^n \phi \rangle = 0, \end{aligned}$$

for any  $\phi \in \mathbb{H}_0^1(\Omega)$ . Therefore the function  $\bar{\partial} U^n + A^n U^n - P_0^n \tilde{f}^n$  vanishes.

**2.11. Error equation.** Let us consider the (*full*) *error*, the *elliptic reconstruction error* and the *parabolic error* which are defined, respectively, as follows:

$$(2.41) \quad e = U - u,$$

$$(2.42) \quad \epsilon = \omega - U,$$

$$(2.43) \quad \rho = \omega - u.$$

We have the following decomposition of the error:

$$(2.44) \quad e = \rho - \epsilon.$$

We can also readily derive the following error relation for the parabolic error in terms of the reconstruction error and the reconstruction itself [LM06]:

$$(2.45) \quad \begin{aligned} \langle \partial_t \rho(t), \phi \rangle + a(\rho(t), \phi) &= \langle \partial_t \epsilon(t), \phi \rangle + a(\omega(t) - \omega^n, \phi) \\ &+ \tau_n^{-1} \langle P_0^n U^{n-1} - U^{n-1}, \phi \rangle + \langle P_0^n \tilde{f}^n - f(t), \phi \rangle \end{aligned}$$

for all  $\phi \in H_0^1(\Omega)$ ,  $t \in I_n$  and  $n = 1, \dots, N$ .

### 3. DUALITY-RECONSTRUCTION A POSTERIORI ERROR ESTIMATES

In this section we synthetically describe how the combination of the elliptic reconstruction and the parabolic duality techniques provide a posteriori error estimates. To keep the discussion as simple as possible, we study first the *spatially semidiscrete scheme*. This simplification allows us to expose our main ideas, which we employ later for the fully discrete case in §4.

**3.1. Notational warning.** Since we will be dealing with the space semidiscrete scheme only, we will use the same symbols introduced for the fully discrete scheme in §2, albeit in their semidiscrete analog by dropping the index  $n$ . *The notation now introduced is valid only in this section.* In particular time-dependent functions, such as  $U$ ,  $\omega$ ,  $e$ ,  $\epsilon$  and  $\rho$ , to be introduced next, should not be confused with their fully-discrete analogs introduced earlier in §2 and valid outside this section.

**3.2. Notation, spatially semidiscrete scheme and the error relation.** Let  $\mathbb{V}$  be a given (time-invariant) finite element space, as defined in §2.1, consider the function  $U : [0, T] \rightarrow \mathbb{V}$  which satisfies the following *semidiscrete Galerkin finite element scheme* associated with the PDE (2.12):

$$(3.1) \quad \begin{aligned} U(0) &:= Iu(0), \text{ and} \\ \langle \partial_t U(t), \Phi \rangle + a(U(t), \Phi) &= \langle f(t), \Phi \rangle \quad \forall \Phi \in \mathbb{V}, t \in [0, T], \end{aligned}$$

where the operator  $I$  is a suitable interpolation or projection operator from  $H^1(\Omega)$ , or  $L_2(\Omega)$ , onto  $\mathbb{V}$ .

We define the (*full*) *error* at time  $t$  to be  $e(t) := U(t) - u(t)$  and the *semidiscrete elliptic reconstruction* to be  $\omega(t) := \mathcal{R}^{\mathbb{V}}U(t)$ , where  $\mathcal{R}^{\mathbb{V}}$  is the elliptic reconstruction operator associated with the space  $\mathbb{V}$ , defined in Definition 2.5. In analogy with the fully discrete notation in §2.9, we define the *semidiscrete elliptic reconstruction error*  $\epsilon := \omega - U$  and the *semidiscrete parabolic error*  $\rho := \omega - u = \epsilon + e$ , keeping in mind the warning §3.1

We observe that while in the simplified semidiscrete setting one assumes the discrete space  $\mathbb{V}$  to be invariant in time, in the fully discrete setting (cf. §4) we will take into account the possibility of the discrete space to change, with respect to the

timestep. For instance, in an adaptive mesh refinement scheme the space change derives from the mesh's modification from one time to the next.

Correspondingly to the fully discrete case (2.45), we may write the following *semidiscrete the parabolic-elliptic error relation*:

$$(3.2) \quad \langle \partial_t \rho(t), \phi \rangle + a(\rho(t), \phi) = \langle \partial_t \epsilon(t), \phi \rangle + \langle P_0^\vee f(t) - f(t), \phi \rangle, \quad \forall \phi \in H_0^1(\Omega), t \in (0, T].$$

**3.3. The dual solution.** The *parabolic dual solution* was introduced by [EJ91] as a tool for a posteriori error estimation. We will use it now to derive such error estimates from (3.2).

For each  $s \leq T$ , consider the *dual solution* to be the function

$$(3.3) \quad z(x, t; s) = z_s(x, t), \text{ for } x \in \Omega \text{ and } 0 \leq t \leq s,$$

which satisfies  $z_s \in L_2(0, T; H_0^1(\Omega))$ ,  $\partial_t z_s \in L_2(0, T; H^{-1}(\Omega))$ , and solves the following backward parabolic *dual problem*:

$$(3.4) \quad \begin{aligned} -\langle \partial_t z_s(t), \phi \rangle + a(\phi, z_s(t)) &= 0, \quad \forall \phi \in H_0^1(\Omega), t \in [0, s], \\ z_s(x, s) &= \rho(x, s) \quad \forall x \in \Omega \end{aligned}$$

for each  $s \in [0, T]$ . Notice that  $\phi$  can be taken to be time dependent, with the appropriate differentiability properties.

The dual solution enjoys stability properties which we will use in the sequel. An immediate property is the usual energy identity

$$(3.5) \quad \|z_s(t)\|^2 + 2 \int_t^s |z_s|_a^2 = \|\rho(s)\|^2 \quad \forall t \in [0, s].$$

A more intricate stability property of  $z_s$  is given by the following result.

**3.4. Lemma (Strong stability estimate [EJ91, Lem. 4.2]).** *For each  $s \in [0, T]$ ,*

$$(3.6) \quad \left\{ \int_0^s \|\partial_t z_s(t)\|^2 (s-t) dt, \int_0^s \|-\operatorname{div}[\mathbf{A}\nabla z_s](t)\|^2 (s-t) dt \right\} \leq \frac{1}{4} \|\rho(s)\|^2.$$

**3.5. A posteriori error analysis via parabolic duality.** Integrating in (3.4) by parts in time implies that

$$(3.7) \quad \begin{aligned} \langle \rho(s), \phi(s) \rangle &= \langle z_s(s), \phi(s) \rangle \\ &= \langle z_s(0), \phi(0) \rangle + \int_0^s \langle \partial_t \phi(t), z_s(t) \rangle + a(\phi(t), z_s(t)) dt, \end{aligned}$$

for all  $\phi \in L_2(0, T; H_0^1(\Omega))$  such that  $\partial_t \phi \in L_2(0, T; H^{-1}(\Omega))$ .

Take  $\phi = \rho$ , use (3.2) and assume  $P_0 \tilde{f} - f = 0$  momentarily—in the proof of Theorem 4.5 we shall remove this assumption—we obtain

$$(3.8) \quad \|\rho(s)\|^2 = \langle \rho(0), z_s(0) \rangle + \int_0^s \langle \partial_t \epsilon(t), z_s(t) \rangle dt.$$

The first term on the right-hand side, is easily estimated, with Lemma 3.4 in mind, as follows:

$$(3.9) \quad \langle \rho(0), z_s(0) \rangle \leq \|\rho(0)\| \sup_{[0, s]} \|z_s\|.$$

As for the second term on the right-hand side of (3.8) we have a choice of two different ways for estimating it.

(a) A direct estimate yields

$$(3.10) \quad \int_0^s \langle \partial_t \epsilon, z_s \rangle \leq \sup_{[0,s]} \|z_s\| \int_0^s \|\partial_t \epsilon\|.$$

Notice that the term  $\partial_t \epsilon$  can be estimated via elliptic a posteriori error estimates because it is the difference between  $\partial_t U$  and its reconstruction  $\mathcal{R}^\vee \partial_t U = \partial_t \mathcal{R}^\vee U$ . Nonetheless, a term involving  $\partial_t \epsilon$  is less desirable than one involving only  $\epsilon$ .

(b) A less direct estimate, that would avoid the appearance of time derivatives in the indicator, is obtained by first integrating by parts in time:

$$(3.11) \quad \int_0^s \langle \partial_t \epsilon, z_s \rangle = \langle \epsilon(s), z_s(s) \rangle - \langle \epsilon(0), z_s(0) \rangle - \int_0^s \langle \epsilon(t), \partial_t z_s(t) \rangle dt.$$

The last integral can be then bounded as follows:

$$(3.12) \quad \begin{aligned} \int_0^s \epsilon(t) \partial_t z_s(t) dt &\leq \int_0^s \frac{\|\epsilon(t)\|}{\sqrt{s-t}} \|\partial_t z_s(t)\| \sqrt{s-t} dt \\ &\leq \left( \int_0^s \frac{\|\epsilon(t)\|^2}{s-t} dt \right)^{1/2} \left( \int_0^s \|\partial_t z_s(t)\|^2 (s-t) dt \right)^{1/2}. \end{aligned}$$

Unfortunately this bound turns out not to be useful, as it stands, due to the weight in the first integral on the last right-hand side. Namely, for this term to be finite it is necessary for  $\epsilon$ , the error between the discrete solution and its reconstruction, to vanish at  $s$ . Heuristically this can be interpreted as the mesh having to become infinitely fine as time gets closer to  $s$ : an unrealistic option.

To circumvent this difficulty, without totally sacrificing  $\|\epsilon\|$  to  $\|\partial_t \epsilon\|$ , we compromise between approach (a) and (b) by following through from (3.8) as follows: fix  $r \in (0, s)$  (think of it as a close point to  $s$ ), split the integral and integrate by parts in time

$$(3.13) \quad \begin{aligned} \|\rho(s)\|^2 &= \langle z_s(0), \rho(0) \rangle + \left[ \int_0^r + \int_r^s \right] \langle \partial_t \epsilon, z_s \rangle \\ &= \langle z_s(0), \rho(0) - \epsilon(0) \rangle + \langle z_s(r), \epsilon(r) \rangle - \int_0^r \langle \epsilon, \partial_t z_s \rangle + \int_r^s \langle \partial_t \epsilon, z_s \rangle \\ &\leq \sup_{[0,s]} \|z_s\| \left( \|e(0)\| + \|\epsilon(r)\| + \int_r^s \|\partial_t \epsilon\| \right) \\ &\quad + \left( \int_0^r \|\partial_t z_s(t)\|^2 (s-t) dt \right)^{1/2} \left( \int_0^r \frac{\|\epsilon(t)\|^2}{s-t} dt \right)^{1/2}. \end{aligned}$$

The stability estimates (3.5) and (3.6) imply that

$$(3.14) \quad \|\rho(s)\| \leq \|e(0)\| + \|\epsilon(r)\| + \int_r^s \|\partial_t \epsilon\| + \frac{1}{2} \left( \int_0^r \frac{\|\epsilon(t)\|^2}{s-t} dt \right)^{1/2}.$$

This discussion's outcome can be summarized into the following result.

**3.6. Theorem (Semidiscrete duality-reconstruction a posteriori error estimate).** *Suppose that  $f(t) \in \mathbb{V}$ , for  $t \in [0, T]$ , and that there exists an a posteriori*

elliptic  $L_2(\Omega)$ -error estimator functional  $\mathcal{E}[\cdot, \mathbb{V}, L_2(\Omega)]$ , as defined in §2.7, then the error occurring in the semidiscrete scheme (3.1) obeys the a posteriori bound

$$(3.15) \quad \sup_{t \in [0, s]} \|U(t) - u(t)\| \leq \|U(0) - u(0)\| + L(s, r) \sup_{[0, s]} \mathcal{E}[U, \mathbb{V}, L_2(\Omega)] \\ + (s - r) \sup_{[r, s]} \mathcal{E}[\partial_t U, \mathbb{V}, L_2(\Omega)]$$

for each  $s, r$ ,  $0 \leq r < s \leq T$ , and with

$$(3.16) \quad L(s, r) := 2 + \frac{1}{2} \sqrt{\log \frac{s}{s-r}}.$$

**Proof.** Fix  $r$  and  $s$  and use (3.14) to get

$$(3.17) \quad \|e(s)\| \leq \|e(0)\| + \|\epsilon(r)\| + \|\epsilon(s)\| + \int_r^s \|\partial_t \epsilon\| + \frac{1}{2} \left( \int_0^r \frac{\|\epsilon(t)\|^2}{s-t} dt \right)^{1/2}.$$

Basic manipulations and the use of the estimator functional  $\mathcal{E}[\cdot, \mathbb{V}, L_2(\Omega)]$  leads to

$$(3.18) \quad \|\epsilon(r)\| + \|\epsilon(s)\| + \frac{1}{2} \left( \int_0^r \frac{\|\epsilon(t)\|^2}{s-t} dt \right)^{1/2} \\ \leq \left( 2 + \frac{1}{2} \left( \int_0^r \frac{dt}{s-t} \right)^{1/2} \right) \sup_{0 \leq t \leq s} \|\epsilon(t)\| \\ \leq \left( 2 + \frac{1}{2} \sqrt{\log \frac{s}{s-r}} \right) \sup_{0 \leq t \leq s} \mathcal{E}[U(t), \mathbb{V}, L_2(\Omega)]$$

and

$$(3.19) \quad \int_r^s \|\partial_t \epsilon\| \leq (s-r) \sup_{r \leq t \leq s} \mathcal{E}[\partial_t U(t), \mathbb{V}, L_2(\Omega)].$$

The result follows by using (3.18) and (3.19) in (3.17).  $\square$

**3.7. Corollary (Semidiscrete duality-residual a posteriori estimates).** *If  $\Omega$  is a convex domain in  $\mathbb{R}^d$  and  $f(t) \in L_2(\Omega)$  for each  $t \in [0, T]$ , then the following a posteriori error estimate holds:*

$$(3.20) \quad \sup_{[0, s]} \|U - u\| \leq \|U(0) - u(0)\| \\ + L(s, r) \sup_{[0, r]} (C_3 \|h^2(\mathcal{A}_{e1} - A^{\mathbb{V}})U\| \\ + C_5 \left\| h^{3/2} J[U] \right\|_{\Sigma} + C_7 \left\| h^2(P_0 \tilde{f} - f) \right\|) \\ + (s-r) \sup_{[r, s]} (C_3 \|h^2(\mathcal{A}_{e1} - A^{\mathbb{V}})\partial_t U\| \\ + C_5 \left\| h^{3/2} J[\partial_t U] \right\|_{\Sigma} + \frac{1}{2\sqrt{\alpha}} \left\| h(P_0 \tilde{f} - f) \right\|).$$

**Proof.** From Theorem 3.6 and Lemma 2.8, the result follows when  $f(t) \in \mathbb{V}$  for all  $t \in (0, T)$ .  $\square$

## 4. ESTIMATES FOR THE FULLY DISCRETE SCHEME

Bearing in mind the techniques of the last section, we now turn our attention to the analysis of the fully discrete scheme (2.15). For convenience, we now switch notation slightly and use the symbol  $U$  (even without the superscript  $n$  sometimes) for the fully discrete solution and its piecewise linear interpolation. We introduce first some extra “discrete-time” notation to be used in this section.

**4.1. Definition (duality time-accumulation coefficients).** In developing the error bounds via duality, we shall need the following (logarithmic) *time accumulation coefficients*:

$$(4.1) \quad \begin{aligned} b_n &:= \begin{cases} \frac{1}{4} \log \left( \frac{T-t_{n-1}}{T-t_n} \right), & \text{for } n = 1, \dots, N-1, \\ \frac{1}{8}, & \text{for } n = N, \end{cases} \\ a_n &:= \int_0^{t_{N-1}} \frac{l_n(t) dt}{T-t} = \begin{cases} 1 - \lambda \left( -\frac{\tau_n}{T} \right), & \text{for } n = 0, \\ \lambda \left( \frac{\tau_n}{T-t_n} \right) - \lambda \left( -\frac{\tau_{n+1}}{T-t_n} \right), & \text{for } n = 1, \dots, N-2, \\ \lambda \left( \frac{\tau_{N-1}}{\tau_N} \right) - 1, & \text{for } n = N-1, \end{cases} \end{aligned}$$

where

$$(4.2) \quad \lambda(x) := \begin{cases} (1 + 1/x) \log(1 + x) & \text{for } |x| \in (0, 1), \\ 1 & \text{for } x = 0 \end{cases}$$

is an increasing function of  $x$ . We observe that the functions  $\lambda(x) - 1$ ,  $1 - \lambda(-x)$  and  $\lambda(x) - \lambda(-y)$  are positive for  $(x, y) \in (0, 1)^2$ , a fact that makes the coefficients  $a_n$  be positive. These coefficients can be appreciated graphically in Figure 1.

**4.2. Lemma (duality time-accumulation coefficients properties).** *The coefficients  $a_n$  and  $b_n$ , defined in §4.1 for  $n = 0, \dots, N$ , satisfy the following:*

$$(4.3) \quad \sum_{n=0}^{N-1} a_n = \log \frac{T}{\tau_N}, \quad \int_{t_{n-1}}^{t_n} \frac{l_{n-1}(t)^2}{T-t} dt \leq b_n \quad \text{and} \quad \sum_{n=1}^N b_n = \frac{1}{4} \left( \frac{1}{2} + \log \frac{T}{\tau_N} \right).$$

**Proof.** The results follow from the definitions and basic calculus.  $\square$

**4.3. Definition (error indicators).** We suppose an a posteriori elliptic error estimator functional  $\mathcal{E}[\cdot, \cdot, \cdot]$ , as defined in §2.7, is available and we introduce the following (time-local)  $\mathcal{E}$ -based *spatial error indicators*:

$$(4.4) \quad \varepsilon_n := \mathcal{E}[U^n, \mathbb{V}^n, \mathbf{L}_2(\Omega)],$$

$$(4.5) \quad \eta_n := \tau_n \mathcal{E}[\partial U^n, \mathbb{V}^n \cap \mathbb{V}^{n-1}, \mathbf{L}_2(\Omega)],$$

and the *time error indicator*

$$(4.6) \quad \theta_n := \|A^{n-1}U^{n-1} - A^nU^n\| = \begin{cases} \frac{1}{2} \left\| P_0^1 \tilde{f}^1 - \bar{\partial}U^1 - A^0U^0 \right\| & \text{for } n = 1, \\ \frac{1}{2} \left\| \partial \left( P_0^n \tilde{f}^n - \bar{\partial}U^n \right) \right\| \tau_n & \text{for } n = 2, \dots, N. \end{cases}$$

or, in some cases, the alternative version

$$(4.7) \quad \theta_n := \|U^{n-1} - U^n\| + \eta_n.$$

In the numerical experiments we only use definition (4.6) of  $\theta_n$ .

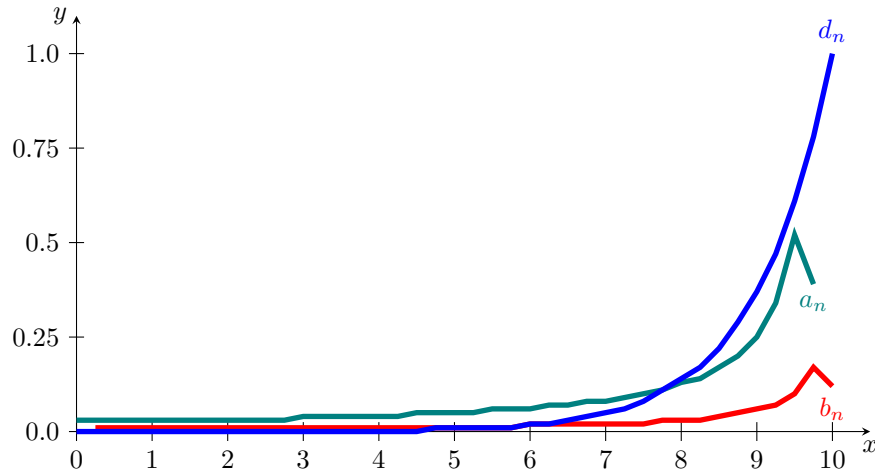


FIGURE 1. An example of the time accumulation coefficients  $(a_n)_n$ ,  $(b_n)_n$  and  $(d_n)_n$  defined in (4.1) and (6.10), respectively. This is the situation for a uniform timestep  $1/4$  over the interval  $[0, 10]$ . All coefficients exhibit a backward decaying “tail” (cf. Theorem 4.5). Noting how this tail is much heavier for  $(a_n)_n$  and  $(b_n)_n$  than for  $(d_n)_n$  it follows that the energy estimator “forgets” much faster than the duality one.

We also introduce the *data approximation error indicator*

$$(4.8) \quad \beta_n := \int_{t_{n-1}}^{t_n} \|f^n - f(t)\| dt,$$

the associated *global data approximation indicator*

$$(4.9) \quad \tilde{\beta}_N := \begin{cases} \sum_{n=1}^N \beta_n & \text{if } \tilde{f}^n = f(t_n), \\ \beta_N + 2 \left( \sum_{n=1}^{N-1} b_n \beta_n^2 \right)^{1/2} & \text{if } \tilde{f}^n = \int_{t_{n-1}}^{t_n} f(t) dt / \tau_n, \end{cases}$$

and the *mesh change error indicator function*

$$(4.10) \quad \gamma_n := \frac{P_0^n U^{n-1} - U^{n-1}}{\tau_n} + P_0^n \tilde{f}^n - \tilde{f}^n = (P_0^n - I)(\tau_n^{-1} U^{n-1} + \tilde{f}^n).$$

**4.4. Remark (smooth data approximation).** For  $f$  smooth enough, we can redefine the indicator  $\beta_n$  in relation (4.8) by the right-hand side of the following inequality:

$$(4.11) \quad \int_{t_{n-1}}^{t_n} \|f^n - f(t)\| dt \leq \|\partial_t f\|_{L_1(I_n, L_2(\Omega))} \tau_n.$$

**4.5. Theorem (general duality a posteriori parabolic-error estimate).** *Let  $u$  be the exact solution of (2.12),  $(U^n)_{n=0, \dots, N}$  the (corresponding fully discrete) solution of (2.15) and  $\omega^n = \mathcal{R}^n U^n$  the elliptic reconstruction of  $U^n$ , for  $n =$*

$0, \dots, N$  as defined by (2.26). Then, with reference to Definition 4.3, the following a posteriori error estimate holds:

$$(4.12) \quad \begin{aligned} \|\omega^N - u(T)\| \leq & \|U^0 - u(0)\| + \left( \sum_{n=0}^{N-1} a_n \varepsilon_n^2 \right)^{1/2} + \eta_N + \left( \sum_{n=1}^N b_n \theta_n^2 \right)^{1/2} \\ & + C_{5.13} \sqrt{\frac{\tau_N}{2}} \|\gamma_N h_N\| + C_{5.13} \left( \sum_{n=1}^{N-1} b_n \|\gamma_n h_n^2\|^2 \right)^{1/2} + \tilde{\beta}_N, \end{aligned}$$

with  $C_{5.13}$  defined in (5.13). The proof of this result is the object of §5. We state and application of this result, which we will also prove later in §5.5.

**4.6. Corollary (Duality a posteriori full error estimates).** *With the same notation as in Theorem 4.5 and supposing that  $\tilde{f}^n = \int_{t_{n-1}}^{t_n} f(t) dt / \tau_n$  we have*

$$(4.13) \quad \begin{aligned} \|U^N - u(T)\| \leq & \|U^0 - u(0)\| + \sqrt{\frac{\tau_N}{2}} \|\gamma_N h_N\| + \sum_{n=1}^N \tau_n \|\partial_t f\|_{L_1(I_n; L_2(\Omega))} \\ & + \sqrt{1 + \log \frac{T}{\tau_N}} \left( \max_{n=0, \dots, N} \varepsilon_n + 2 \max_{n=1, \dots, N-1} \|\gamma_n h_n^2\| + \frac{1}{2} \max_{n=1, \dots, N} \theta_n \right). \end{aligned}$$

**4.7. Remark (comparison between Theorem 4.5 and Corollary 4.6).** Corollary 4.6 has a simpler estimate than Theorem 4.5 in that it involves less terms and does not require as much memory. Notice, however, that from an error bound viewpoint, the theorem's tighter bound may be more effective as the time accumulation is not as strict as in the corollary. This is especially true in problems, typical in the parabolic setting, where the initial error may be very big and gets damped with time.

## 5. PROOF OF THEOREM 4.5

As with the semidiscrete case that we dealt with in §3 to prove Theorem 3.6, our starting point to prove (4.12) is the fully discrete analog of (3.8), which is readily obtained from (2.45) and (3.4):

$$(5.1) \quad \begin{aligned} \|\rho(T)\|^2 = & \langle \rho(0), z_T(0) \rangle + \int_0^T \langle \partial_t \epsilon(t), z_T(t) \rangle dt \\ & + \sum_{n=1}^N \int_{t_{n-1}}^{t_n} a(\omega(t) - \omega^n, z_T(t)) + \langle \gamma_n, z_T(t) \rangle + \langle \tilde{f}^n - f(t), z_T(t) \rangle dt. \end{aligned}$$

We recall that  $\rho$  and  $\epsilon$  are defined in §3.2, the functions  $\gamma_n$  are defined, for each  $n = 1, \dots, N$ , by (4.10) and  $z_T$  is the solution of the dual problem (3.4) with  $s = T$ .



**5.1. Space error estimate.** The first two terms are estimated, similarly to (3.13), as follows:

$$(5.2) \quad \langle \rho(0), z_T(0) \rangle + \int_0^T \langle \partial_t \epsilon(t), z_T(t) \rangle dt \\ \leq \|\rho(T)\| \left( \|\epsilon(0)\| + \|\epsilon(t_{N-1})\| + \int_{t_{N-1}}^T \|\partial_t \epsilon\| + \frac{1}{2} \left( \int_0^{t_{N-1}} \frac{\|\epsilon(t)\|^2}{T-t} dt \right)^{1/2} \right).$$

To proceed we observe that

$$(5.3) \quad \int_{t_{N-1}}^T \|\partial_t \epsilon\| = \|\epsilon^N - \epsilon^{N-1}\| \leq \eta_N,$$

and, by convexity and affinity of  $l_n$ , this implies

$$(5.4) \quad \int_0^{t_{N-1}} \frac{\|\epsilon(t)\|^2}{T-t} dt = \int_0^{t_{N-1}} \frac{\left\| \sum_{n=0}^{N-1} \epsilon^n l_n(t) \right\|^2}{T-t} dt \\ \leq \sum_{n=0}^{N-1} \|\epsilon^n\|^2 \int_0^{t_{N-1}} \frac{l_n(t)}{T-t} dt = \sum_{n=0}^{N-1} a_n \epsilon_n^2.$$

Thus we obtain

$$(5.5) \quad \langle \rho(0), z_T(0) \rangle + \int_0^T \langle \partial_t \epsilon(t), z_T(t) \rangle dt \\ \leq \|\rho(T)\| \left( \|\epsilon(0)\| + \left( \sum_{n=0}^{N-1} a_n \epsilon_n^2 \right)^{1/2} + \eta_N \right).$$

**5.2. Time error estimate.** The third term in (5.1), which accounts mainly for the time error, can be bounded as follows:

$$(5.6) \quad \sum_{n=1}^N \int_{t_{n-1}}^{t_n} a(\omega(t) - \omega^n, z_T(t)) dt \\ \leq \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \|\omega^{n-1} - \omega^n\| l_{n-1}(t) \|\operatorname{div} [A \nabla z_T](t)\| dt \\ \leq \frac{1}{2} \|\rho(T)\| \left( \sum_{n=1}^N \|\omega^{n-1} - \omega^n\|^2 \int_{t_{n-1}}^{t_n} \frac{l_{n-1}(t)^2}{T-t} dt \right)^{1/2} \\ \leq \frac{1}{2} \|\rho(T)\| \left( \frac{1}{2} \|\omega^N - \omega^{N-1}\|^2 + \sum_{n=1}^{N-1} \|\omega^n - \omega^{n-1}\|^2 \log \left( \frac{T-t_{n-1}}{T-t_n} \right) \right)^{1/2}.$$

Notice that in the time integral on the last interval  $(t_{N-1}, T]$  how the numerator  $l_{N-1}(t) = O(T-t)$  compensates for the singularity of  $1/(T-t)$ . The terms  $\|\omega^{n-1} - \omega^n\|$  appearing in this estimate still need to be estimated, as there is no explicit knowledge of the reconstructed functions  $\omega^n = \mathcal{R}^n U^n$ . These terms can be dealt with in two different ways.

(a) One way to estimate these terms is given by

$$(5.7) \quad \begin{aligned} \|\omega^{n-1} - \omega^n\| &\leq \|U^{n-1} - U^n\| + \|\omega^{n-1} - \omega^n - U^{n-1} + U^n\| \\ &= \|U^{n-1} - U^n\| + \tau_n \|\partial_t \epsilon^n\| \leq \|U^{n-1} - U^n\| + \eta_n, \end{aligned}$$

for all  $t \in I_n$ . Thus we obtain the estimate

$$(5.8) \quad \begin{aligned} \sum_{n=1}^N \int_{t_{n-1}}^{t_n} a(\omega(t) - \omega^n, z_T(t)) \, dt &\leq \|\rho(T)\| \left( \sum_{n=1}^N b_n (\|U^{n-1} - U^n\| + \eta_n)^2 \right)^{1/2} \\ &= \|\rho(T)\| \left( \sum_{n=1}^N b_n \theta_n^2 \right)^{1/2}, \end{aligned}$$

by using  $\theta_n$ 's alternative definition (4.7).

(b) Another way to estimate  $\|\omega^{n-1} - \omega^n\|$  consists in using again the definition of elliptic reconstruction Cauchy–Bunyakovskii–Schwarz inequality and the Poincaré–Friedrichs inequality as follows:

$$(5.9) \quad \begin{aligned} \|\omega^{n-1} - \omega^n\|^2 &\leq C_{\text{PF}} a(\omega^{n-1} - \omega^n, \omega^{n-1} - \omega^n) \\ &= C_{\text{PF}} \langle A^{n-1}U^{n-1} - A^nU^n, \omega^{n-1} - \omega^n \rangle \\ &\leq C_{\text{PF}} \|A^{n-1}U^{n-1} - A^nU^n\| \|\omega^{n-1} - \omega^n\| \\ &\leq C_{\text{PF}}^2 \|A^{n-1}U^{n-1} - A^nU^n\|^2. \end{aligned}$$

Using the definition of  $\theta_n$  in (4.6) yields

$$(5.10) \quad \begin{aligned} \sum_{n=1}^N \int_{t_{n-1}}^{t_n} a(\omega(t) - \omega^n, z_T(t)) \, dt \\ \leq \|\rho(T)\| \left( \sum_{n=1}^N b_n C_{\text{PF}}^2 \|A^{n-1}U^{n-1} - A^nU^n\|^2 \right)^{1/2} = \|\rho(T)\| \left( \sum_{n=1}^N b_n \hat{\theta}_n^2 \right)^{1/2}. \end{aligned}$$

Here  $\hat{\theta}_n := C_{\text{PF}}\theta_n$ .

**5.3. Mesh-change error estimates.** To bound the third term in (5.1), we use  $\gamma_n$ 's  $L_2(\Omega)$  orthogonality and the orthogonal projector  $\Pi^n : L_2(\Omega) \rightarrow \tilde{\mathbb{V}}^n$  as follows:

$$(5.11) \quad \begin{aligned} \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \langle \gamma_n, z_T(t) \rangle \, dt &= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \langle \gamma_n, z_T(t) - \Pi^n z_T(t) \rangle \, dt \\ &\leq C_{2,\text{proj}} \sum_{n=1}^{N-1} \int_{t_{n-1}}^{t_n} \|\gamma_n h_n^2\| |z_T(t)|_2 \, dt + C_{1,\text{proj}} \int_{t_{N-1}}^T \|\gamma_N h_N\| |z_T(t)|_1 \, dt \\ &\leq \frac{C_{\text{proj}} C_{\text{reg}}}{2} \|\rho(T)\| \left( \sqrt{2\tau_N} \|\gamma_N h_N\| + \left( \sum_{n=1}^{N-1} \|\gamma_n h_n^2\|^2 \log \left( \frac{T - t_{n-1}}{T - t_n} \right) \right)^{1/2} \right) \\ &= \|\rho(T)\| C_{5.13} \left( \sqrt{\frac{\tau_N}{2}} \|\gamma_N h_N\| + \left( \sum_{n=1}^{N-1} b_n \|\gamma_n h_n^2\|^2 \right)^{1/2} \right). \end{aligned}$$

Here we have used the fact that  $\Omega$  is convex in order to apply the regularity estimate

$$(5.12) \quad |z_T(t)|_i \leq C_{\text{reg}} \|\text{div} [\mathbf{A}\nabla z_T](t)\|, \text{ for } i = 1, 2,$$

and then apply the strong stability estimate (3.6), and

$$(5.13) \quad C_{5.13} := C_{\text{proj}} C_{\text{reg}} \text{ and } C_{\text{proj}} := \max_{i=1,2} C_{i,\text{proj}}.$$

**5.4. Data-approximation error estimates.** The fourth term in (5.1) can be bounded in two different ways depending on which definition for  $\tilde{f}^n$  appearing in the fully discrete scheme (2.15) is chosen.

(a) If  $\tilde{f}^n = f^n$ , then we can proceed as follows:

$$(5.14) \quad \begin{aligned} \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \langle \tilde{f}^n - f(t), z_T(t) \rangle dt &\leq \sum_{n=1}^N \max_{I_n} \|z_T\| \int_{t_{n-1}}^{t_n} \|\tilde{f}^n - f(t)\| dt \\ &\leq \|\rho(T)\| \sum_{n=1}^N \beta_n. \end{aligned}$$

(b) If instead of  $\tilde{f}^n = f(t_n)$  we have  $\tilde{f}^n = \int_{t_{n-1}}^{t_n} f(t) dt / \tau_n$ , which is the  $L_2(\Omega)$ -projection of  $f$  onto constants in time, then we can exploit the orthogonality and write, for each  $n = 1, \dots, N-1$ ,

$$(5.15) \quad \begin{aligned} \int_{t_{n-1}}^{t_n} \langle \tilde{f}^n - f(t), z_T(t) \rangle dt &= \int_{t_{n-1}}^{t_n} \langle \tilde{f}^n - f(t), z_T(t) - z_T(t_{n-1}) \rangle dt \\ &\leq \max_{t \in I_n} \|z_T(t) - z_T(t_{n-1})\| \int_{t_{n-1}}^{t_n} \|\tilde{f}^n - f(t)\| dt. \end{aligned}$$

By noticing that

$$(5.16) \quad \begin{aligned} \max_{t \in I_n} \|z_T(t) - z_T(t_{n-1})\| &= \max_{t \in I_n} \left\| \int_{t_{n-1}}^t \partial_s z_T(s) ds \right\| \\ &\leq \max_{t \in I_n} \int_{t_{n-1}}^t \|\partial_s z_T(s)\| ds \leq \int_{t_{n-1}}^{t_n} \|\partial_t z_T\| \\ &\leq \log \left( \frac{T - t_{n-1}}{T - t_n} \right)^{1/2} \left( \int_{t_{n-1}}^{t_n} \|\partial_t z_T(t)\|^2 (T - t) dt \right)^{1/2} \\ &= 2b_n^{1/2} \left( \int_{t_{n-1}}^{t_n} \|\partial_t z_T(t)\|^2 (T - t) dt \right)^{1/2}. \end{aligned}$$

Summing up, applying Cauchy–Bunyakovskii–Schwarz inequality, and using the strong stability estimate (3.6) we obtain

$$(5.17) \quad \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \langle \tilde{f}^n - f(t), z_T(t) \rangle dt \leq \|\rho(T)\| \left( \beta_N + 2 \left( \sum_{n=1}^{N-1} b_n \beta_n^2 \right)^{1/2} \right).$$

Employing estimates (5.5), (5.8) (or (5.10)), (5.11) and (5.14) (or (5.17)) into the relation (5.1) we obtain the result of Theorem 4.5.  $\square$

**5.5. Proof of Corollary 4.6.** Referring to the notation introduced in §4.3, the indicator  $\eta_n$  defined by (4.5) and appearing in (4.12) can be substituted by the more “practical” one:  $\varepsilon_n + \varepsilon_{n-1}$ . To see this let us first revisit estimate (5.3) and recall definition (4.4) to write

$$(5.18) \quad \|\partial_t \epsilon^N\| \tau_n = \|\epsilon^N - \epsilon^{N-1}\| \leq \|\epsilon^N\| + \|\epsilon^{N-1}\| \leq \varepsilon_N + \varepsilon_{N-1}.$$

It follows that

$$(5.19) \quad \begin{aligned} \|\omega^N - u(T)\| &\leq \|U^0 - u(0)\| + \left( \sum_{n=0}^{N-1} a_n \varepsilon_n^2 \right)^{1/2} + \varepsilon_{N-1} + \varepsilon_N \\ &\quad + \left( \sum_{n=1}^N b_n \theta_n^2 \right)^{1/2} + \sqrt{\frac{T_N}{2}} \|\gamma_N h_N\| + \left( \sum_{n=1}^{N-1} b_n \|\gamma_n h_n^2\|^2 \right)^{1/2} + \tilde{\beta}_N. \end{aligned}$$

To close use the splitting  $e = \rho - \epsilon$  to obtain

$$(5.20) \quad \begin{aligned} \|\epsilon^N\| &\leq \|e^0\| + \left( \sum_{n=0}^{N-1} a_n \varepsilon_n^2 \right)^{1/2} + \varepsilon_{N-1} + \varepsilon_N \\ &\quad + \left( \sum_{n=1}^N b_n \theta_n^2 \right)^{1/2} + \sqrt{\frac{T_N}{2}} \|\gamma_N h_N\| \\ &\quad + \left( \sum_{n=1}^{N-1} b_n \|\gamma_n h_n^2\|^2 \right)^{1/2} + \beta_N + 2 \left( \sum_{n=1}^{N-1} b_n \beta_n^2 \right)^{1/2}. \end{aligned}$$

We notice that the former estimate implies the more traditional one [EJ91],

$$(5.21) \quad \begin{aligned} \int_0^{t_{N-1}} \frac{\|\epsilon(t)\|^2}{T-t} dt &\leq \max_{n=0, \dots, N-1} \|\epsilon^n\|^2 \left( 1 + \sum_{n=1}^{N-1} a_n \right) \\ &= (1 + 4L(T, t_{N-1})^2) \max_{n=0, \dots, N-1} \|\epsilon^n\|^2, \end{aligned}$$

where  $L(T, t_{N-1})$  is the logarithmic factor defined in (3.16).

Also, here the indicator can be simplified if we relax the bound by taking the maximum norm in time as follows:

$$(5.22) \quad \begin{aligned} \sum_{n=1}^N \int_{t_{n-1}}^{t_n} a(\omega(t) - \omega^n, z_T(t)) dt \\ \leq \|\rho(T)\| \max_{n=1, \dots, N} \theta_n \sqrt{\left( \frac{1}{8} + L(T, t_{N-1})^2 \right)}. \end{aligned}$$

As with the space and time estimates, this estimate can be simplified, with some loss of sharpness, as follows:

$$(5.23) \quad \begin{aligned} \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \langle \gamma_n, z_T(t) \rangle dt \\ \leq \|\rho(T)\| \left( \sqrt{\frac{T_N}{2}} \|\gamma_N h_N\| + L(T, t_{N-1}) \max_{n=1, \dots, N-1} \|\gamma_n h_n^2\| \right). \end{aligned}$$

Like earlier estimates, this estimate can be further simplified, by taking the maximum and slightly relaxing it, into

$$(5.24) \quad \sum_{n=1}^N \int_{t_{n-1}}^{t_n} \langle \tilde{f}^n - f(t), z_T(t) \rangle dt \leq \|\rho(T)\| \left( \beta_N + 2L(T, t_{N-1}) \max_{n=1, \dots, N-1} \beta_n \right). \quad \square$$

### 6. THE ENERGY-RECONSTRUCTIVE APPROACH

In a previous paper [LM06], we analyzed the combination of classical energy methods for parabolic equations with the elliptic reconstruction to obtain a posteriori  $L_\infty(L_2(\Omega))$ -error estimates. In this section we give a similar analysis that yields tighter bounds with respect to time accumulation, that will be compared with the ones derived by duality.

**6.1. Theorem (Semidiscrete energy-reconstruction a posteriori error estimate).** *Let the notation and conditions of Theorem 3.6 hold then the following a posteriori bound is true:*

$$(6.1) \quad \sup_{[0,s]} \|U - u\| \leq \|U(0) - u(0)\| + \mathcal{E}[U(0), \tilde{V}, L_2(\Omega)] + \sup_{[0,s]} \mathcal{E}[U, \tilde{V}, L_2(\Omega)] + \int_0^s \exp\left(\frac{\alpha}{C_{PF}^2}(s-t)\right) \mathcal{E}[\partial_t U, \tilde{V}, L_2(\Omega)] dt.$$

**Proof.** Testing equation (3.2) with  $\rho$  and noting  $f(t) \in \tilde{V}$  yields

$$(6.2) \quad \frac{1}{2} d_t \|\rho\|^2 + |\rho|_a^2 = \langle \partial_t \epsilon, \rho \rangle.$$

In view of the Poincaré–Friedrichs inequality and the equivalence of the energy norm to the  $H^1(\Omega)$  seminorm (2.11),

$$(6.3) \quad \frac{1}{2} d_t \|\rho\|^2 + |\rho|_a^2 \geq \frac{1}{2} d_t \|\rho\|^2 + \frac{\alpha}{C_{PF}^2} \|\rho\|^2,$$

and hence

$$(6.4) \quad \frac{1}{2} d_t \|\rho\|^2 + \frac{\alpha}{C_{PF}^2} \|\rho\|^2 \leq \langle \partial_t \epsilon, \rho \rangle \leq \|\rho\| \|\partial_t \epsilon\|.$$

Dividing through by  $\|\rho\|$  gives

$$(6.5) \quad d_t \|\rho\| + \frac{\alpha}{C_{PF}^2} \|\rho\| \leq \|\partial_t \epsilon\|.$$

Using the integrating factor  $\exp(\alpha t/C_{PF}^2)$  we conclude that

$$(6.6) \quad \|\rho(t)\| \leq \|\rho(0)\| + \int_0^t \exp\left(\frac{\alpha}{C_{PF}^2}(s-t)\right) \|\partial_t \epsilon(s)\| ds. \quad \square$$

**6.2. Corollary (Semidiscrete energy-residual a posteriori estimates).** *Let the assumptions of Corollary 3.7 hold, then the following a posteriori bound holds*

$$(6.7) \quad \begin{aligned} \sup_{[0,s]} \|U - u\| &\leq \|U(0) - u(0)\| + C_3 \|h^2[\mathcal{A}_{\text{el}} - A^{\text{V}}] U(0)\| + C_5 \left\| h^{3/2} J[U(0)] \right\|_{\Sigma} \\ &\quad + \sup_{[0,s]} \left( C_3 \|h^2[\mathcal{A}_{\text{el}} - A^{\text{V}}] U\| + C_5 \left\| h^{3/2} J[U] \right\|_{\Sigma} \right) \\ &\quad + \int_0^s \exp\left(\frac{\alpha}{C_{\text{PF}}^2}(s-t)\right) \sup_{[0,s]} \left( C_3 \|h^2[\mathcal{A}_{\text{el}} - A^{\text{V}}] \partial_t U\| \right. \\ &\quad \left. + C_5 \left\| h^{3/2} J[\partial_t U] \right\|_{\Sigma} + \left\| P_0 \tilde{f} - f \right\| \right) dt. \end{aligned}$$

**Proof.** Removing the assumption  $f(t) \in \tilde{\mathcal{V}}$  from the proof of Theorem 6.1 gives

$$(6.8) \quad \|\rho(t)\| \leq \|\rho(0)\| + \int_0^t \exp\left(\frac{\alpha}{C_{\text{PF}}^2}(s-t)\right) \left( \|\partial_t \epsilon(s)\| + \left\| (P_0 \tilde{f} - f)(s) \right\| \right) ds$$

as an analog of (6.6). The splitting  $e(t) = \rho(t) - \epsilon(t)$  together with the error estimates from Lemma 2.8 yield the desired result.  $\square$

**6.3. Definition (energy time accumulation coefficients).** The *energy time-accumulation function* is defined as

$$(6.9) \quad d(t, s) := \exp(a(t-s)), \quad 0 \leq t < s \leq T, \quad a := \alpha/C_{\text{PF}}^2$$

where  $C_{\text{PF}}$  is the Poincaré–Friedrichs constant and  $\alpha$  is the coercivity constant (2.8); we denote  $d(t, T) =: d(t)$ . The *energy time-accumulation coefficients* are defined, for  $0 \leq n < m \leq N$ , by

$$(6.10) \quad d_n^m := \int_{t_{n-1}}^{t_n} d(t, t_m) dt = \frac{1}{a} \exp(a(t_n - t_m)) (1 - \exp(-a\tau_n)).$$

When  $m = N$  we drop it and simply write  $d_n$  instead of  $d_n^m$ . Note the useful recursive relation

$$(6.11) \quad d_n^{m+1} = d_n^m \exp(-a\tau_{m+1}).$$

This will imply, for many of the estimators are of the following form, say for the time estimator, for  $m \geq 0$ , that

$$(6.12) \quad \begin{aligned} \Theta_{m+1} &:= \sum_{n=1}^{m+1} d_n^{m+1} \theta_n = \exp(-a\tau_{m+1}) \sum_{n=1}^m d_n^m \theta_n + d_{m+1}^{m+1} \theta_{m+1} \\ &= \exp(-a\tau_{m+1}) \Theta_m + \frac{1 - \exp(-a\tau_{m+1})}{a} \theta_{m+1} \end{aligned}$$

with  $\Theta_0 := 0$ .

**6.4. Theorem (general energy a posteriori parabolic-error estimate).** *Making use of the same notation as in Theorem 4.5 the following a posteriori estimate holds:*

(6.13)

$$\begin{aligned} \max_{t_n \in [0, T]} \|U^n - u(t_n)\| &\leq \|\rho(0)\| + \max_{n \in [0: N]} \varepsilon_n + 2 \sum_{n=1}^N (\eta_n \tau_n^{-1} + \beta_n + \gamma_n + \theta_n) d_n \\ &=: \mathcal{E}_0 + \mathcal{E}_\infty(N) + \mathcal{E}_1(N). \end{aligned}$$

*Proof.* We utilize the arguments of [LM06] under the approach described in Theorem 6.1. The starting point for this estimate is the parabolic error identity (2.45) tested with  $\rho$  as follows:

$$\begin{aligned} \frac{1}{2} d_t \|\rho\|^2 + \|\rho\|_a^2 &= \langle \partial_t \epsilon, \rho \rangle + a(\omega - \omega^n, \rho) + \tau_n^{-1} \langle P_0^n U^{n-1} - U^{n-1}, \rho \rangle \\ &\quad + \langle P_0^n \tilde{f}^n - f, \rho \rangle \\ &=: \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4. \end{aligned} \tag{6.14}$$

Analogously to the semidiscrete we make use of a Poincaré–Friedrichs inequality and the coercivity of  $a$  to absorb the energy norm into the  $L_2(\Omega)$ -norm as follows:

$$\frac{1}{2} d_t \|\rho\|^2 + |\rho|_a^2 \geq \frac{1}{2} d_t \|\rho\|^2 + \frac{\alpha}{C_{\text{PF}}^2} \|\rho\|^2, \tag{6.15}$$

giving

$$\frac{1}{2} d_t \|\rho\|^2 + \frac{\alpha}{C_{\text{PF}}^2} \|\rho\|^2 \leq |\mathcal{I}_1| + |\mathcal{I}_2| + |\mathcal{I}_3| + |\mathcal{I}_4|. \tag{6.16}$$

Solving the differential equation with an integrating factor approach and integrating from 0 to  $T$  we see that

$$\frac{1}{2} \|\rho(T)\|^2 - \frac{1}{2} \|\rho(0)\|^2 \leq \int_0^T \exp\left(\frac{\alpha}{C_{\text{PF}}^2}(s - T)\right) (|\mathcal{I}_1| + |\mathcal{I}_2| + |\mathcal{I}_3| + |\mathcal{I}_4|) ds. \tag{6.17}$$

Denote  $t_* \in [0, T]$  to be the time such that

$$\|\rho(t_*)\| = \max_{t \in [0, T]} \|\rho(t)\|. \tag{6.18}$$

Now we see that

$$\frac{1}{2} \|\rho(t_*)\|^2 - \frac{1}{2} \|\rho(0)\|^2 \leq \int_0^{t_*} \exp\left(\frac{\alpha}{C_{\text{PF}}^2}(s - t_*)\right) (|\mathcal{I}_1| + |\mathcal{I}_2| + |\mathcal{I}_3| + |\mathcal{I}_4|) ds. \tag{6.19}$$

It then follows that

$$\begin{aligned} \frac{1}{2} \|\rho(t_*)\|^2 - \frac{1}{2} \|\rho(0)\|^2 &\leq \int_0^T \exp\left(\frac{\alpha}{C_{\text{PF}}^2}(s - T)\right) (|\mathcal{I}_1| + |\mathcal{I}_2| + |\mathcal{I}_3| + |\mathcal{I}_4|) ds \\ &\leq \sum_{n=1}^N \int_{t_{n-1}}^{t_n} d(s) (|\mathcal{I}_1| + |\mathcal{I}_2| + |\mathcal{I}_3| + |\mathcal{I}_4|) ds. \end{aligned} \tag{6.20}$$

The terms  $\mathcal{I}_1$ ,  $\mathcal{I}_3$ ,  $\mathcal{I}_4$  are all dealt with in a similar way, by using Cauchy–Bunyakovskii–Schwarz inequality and a maximum argument. For example,

$$\begin{aligned}
\int_{t_{n-1}}^{t_n} d(s) |\mathcal{I}_1| \, ds &\leq \int_{t_{n-1}}^{t_n} d(s) |\langle \partial_t \epsilon(s), \rho(s) \rangle| \, ds \\
&\leq \int_{t_{n-1}}^{t_n} d(s) \|\partial_t \epsilon(s)\| \|\rho(s)\| \, ds \\
(6.21) \quad &\leq \|\rho(t_*)\| \int_{t_{n-1}}^{t_n} d(s) \|\partial_t \epsilon(s)\| \, ds \\
&\leq \|\rho(t_*)\| \int_{t_{n-1}}^{t_n} d(s) \, ds \|\epsilon^n - \epsilon^{n-1}\| \tau_n^{-1} \\
&\leq \|\rho(t_*)\| \int_{t_{n-1}}^{t_n} d(s) \, ds \eta_n \tau_n^{-1}.
\end{aligned}$$

The term  $\mathcal{I}_2$  that will eventually yield a time error indicator requires a little more care:

$$\begin{aligned}
\int_{t_{n-1}}^{t_n} d(s) |\mathcal{I}_2| \, ds &= \int_{t_{n-1}}^{t_n} d(s) |a(\omega(s) - \omega^n, \rho(s))| \, ds \\
&= \int_{t_{n-1}}^{t_n} d(s) |a(l_{n-1}(s) \mathcal{R}^{n-1} U^{n-1} + l_n(s) \mathcal{R}^n U^n - \mathcal{R}^n U^n, \rho(s))| \, ds \\
(6.22) \quad &= \int_{t_{n-1}}^{t_n} d(s) l_{n-1}(s) |a(\mathcal{R}^{n-1} U^{n-1} - \mathcal{R}^n U^n, \rho(s))| \, ds \\
&= \int_{t_{n-1}}^{t_n} d(s) l_{n-1}(s) |\langle A^{n-1} U^{n-1} - A^n U^n, \rho(s) \rangle| \, ds \\
&\leq \|\rho_*\| \int_{t_{n-1}}^{t_n} d(s) \|A^{n-1} U^{n-1} - A^n U^n\| \, ds \\
&\leq \|\rho_*\| \int_{t_{n-1}}^{t_n} d(s) \theta_n \, ds.
\end{aligned}$$

Combining the results together we see that

$$(6.23) \quad \|\rho(t_*)\|^2 \leq \|\rho(0)\|^2 + 2 \|\rho(t_*)\| \sum_{n=1}^N \int_{t_{n-1}}^{t_n} d(s) (\eta_n \tau_n^{-1} + \theta_n + \beta_n + \gamma_n) \, ds.$$

Making use of the  $L_2(\Omega)$  simplification rule [LM06, §3.8] it follows that

$$(6.24) \quad \|\rho(t_*)\| \leq \|\rho(0)\| + 2 \sum_{n=1}^N \int_{t_{n-1}}^{t_n} d(s) (\eta_n \tau_n^{-1} + \theta_n + \beta_n + \gamma_n) \, ds,$$

which yields the desired result.  $\square$



## 7. NUMERICAL COMPARISON OF DUALITY AND ENERGY, VIA SPATIAL RESIDUALS

We close the paper with a sample application of the “abstract” a posteriori estimates derived in §§6–4.

In particular, we summarise next numerical experiments to test the asymptotic behaviour of the estimators given in Theorem 4.5, Corollary 4.6 and Theorem 6.4. The C code used for these computational experiments is an extension of the library ALBERTA [SS05]. The code is freely available by emailing one of the authors. To make the effects of numerical quadrature negligible we choose the quadrature formula such that it is exact on polynomials of degree 17 and less.

**7.1. Residuals.** Since our aim is to compare the numerical performance of the duality and the energy based estimates, which differ mostly in their time-accumulation and time-estimation aspects, we use the same type of spatial indicators given by the residual estimators function introduced in Lemma 2.8 and refer to [LM06] or [LP12] for more details.

The residuals constitute the building blocks of the a posteriori estimators used in our computer experiments. We associate with equations (2.12) and (2.39) two residual functions: the *inner residual* is defined as

$$(7.1) \quad \begin{aligned} R^0 &:= \mathcal{A}_{\text{el}}U^0 - A^0U^0, \\ R^n &:= \mathcal{A}_{\text{el}}U^n - A^nU^n = \mathcal{A}_{\text{el}}U^n - P_0^n \tilde{f}^n + \bar{\partial}U^n, \text{ for } n = 1, \dots, N, \end{aligned}$$

and the *jump residual* which is defined as

$$(7.2) \quad J^n := J[U^n] = \llbracket \nabla U^n \rrbracket.$$

With definition §2.4 in mind, the inner residual terms can be written explicitly as

$$(7.3) \quad \langle R^n, \phi \rangle = \sum_{K \in \mathcal{T}_n} \left\langle -\operatorname{div}[\mathbf{A}\nabla U^n] - P_0^n \tilde{f}^n + \frac{U^n - P_0^n U^{n-1}}{\tau_n}, \phi \right\rangle_K.$$

We can now introduce, for  $n = 0, \dots, N$ , the *elliptic reconstruction error indicators*

$$(7.4) \quad \varepsilon_n := C_{6,2} \|h_n^2 R^n\| + C_{10,2} \|h_n^{3/2} J^n\|_{\Sigma_n},$$

and, for  $n = 1, \dots, N$ , the *space error indicator*

$$(7.5) \quad \eta_n := C_{6,2} \|\hat{h}_n^2 \partial R^n\| + C_{10,2} \|\hat{h}_n^{3/2} \partial J^n\|_{\hat{\Sigma}_n} + C_{14,2} \|\hat{h}_n^{3/2} \partial J^n\|_{\hat{\Sigma}_n \setminus \hat{\Sigma}_n}.$$

**7.2. The benchmark problem.** We take  $\mathbf{A} = -\mathbf{I}$  such that the parabolic problem (2.12) coincides with the heat equation. We tune data functions  $f$  and  $u_0$  of this parabolic problem so that its exact solution  $u$  is given by

$$(7.6) \quad u(\mathbf{x}, t) = \sin(\pi t) \exp(-10|\mathbf{x}|^2).$$

We fix  $d = 2$  and take  $\Omega = (-1, 1) \times (-1, 1)$ .

**7.3. Definition (experimental order of convergence).** Given two sequences  $a(i)$  and  $h(i)$ ,  $i = 0, \dots, N$  we define the experimental order of convergence (EOC) to be:

$$(7.7) \quad \text{EOC}(a, h; i) = \frac{\ln(a(i+1)/a(i))}{\ln(h(i+1)/h(i))}.$$

**7.4. Definition (effectivity index and its inverse).** The main tool for deciding the quality of an estimator is the effectivity index (EI) which is the ratio of the error and the estimator, i.e., using the estimators from the duality-based Theorem 4.5 at time  $t_m$ , for some  $m = 1, \dots, N$ ,

$$(7.8) \quad \text{EI}(t_m) = \frac{(\sum_{n=1}^m b_n \theta_n^2)^{1/2} + (\sum_{n=0}^{m-1} a_n \epsilon_n^2)^{1/2} + \eta_m}{\|U - u\|_{L_\infty(0, t_m; L_2(\Omega))}},$$

using the results of Corollary 4.6,

$$(7.9) \quad \text{EI}(t_m) = \frac{\max_{n \in [0:m]} \theta_n + \max_{n \in [0:m-1]} \epsilon_n + \eta_m}{\|U - u\|_{L_\infty(0, t_m; L_2(\Omega))}},$$

and for the estimator associated with the energy-estimator Theorem 6.4,

$$(7.10) \quad \text{EI}(t_m) = \frac{\sum_{n=1}^m \int_{t_{n-1}}^{t_n} d(s)(\theta_n + \eta_n) + \max_{n \in [0:m]} \epsilon_n}{\|U - u\|_{L_\infty(0, t_m; L_2(\Omega))}}.$$

Since it is much easier to visualise we will be computing the *inverse effectivity index*,  $1/\text{EI}(t_m)$ .

**7.5. Comparing duality estimates with energy estimates.** The second main objective of this research was to compare the numerical results associated with the duality-based estimator of Theorem 4.5 and Corollary 4.6 to those of the energy-based estimator of Theorem 6.4.

Based on the theory developed in the paper it is not clear which estimator is “better”. On one hand, due to the better time-accumulation weights  $d_n^m$  (exponentially “forgetful”) in Theorem 6.4, as compared to the  $a_n$  and  $b_n$  (logarithmically accumulative) in Theorem 4.5, the energy-based estimators have the advantage of being more localized in time. On the other hand, the energy-based estimators depend on the Poincaré–Friedrichs constant, which measures the worse-case scenario and might drive up the effectivity index.

Extensive numerical experimentation lead us to the following “practical” conclusions: energy-based estimators perform slightly better than duality-based ones only over long-time integration, but, used with the residual elliptic estimators, they yield much higher effectivity indexes, because of the presence of the  $\eta_n$  tail.

We illustrate this with a benchmark problem (7.6) that we have chosen as to emphasise long-time behavior of the estimators. The initial condition is zero, the boundary values are not exactly zero but negligible, hence little interpolation error is committed here; however, some care must be taken dealing with these small numbers. The model problem (7.6) is approximated on a stationary mesh in time, hence  $\tilde{V}^{n-1} = \tilde{V}^n$  for all  $n = 1, \dots, N$ . Our results are valid for any polynomial order for the spatial finite elements, but we report results only for  $\mathbb{P}^1$  elements. In order to emphasise the time-estimator, we take  $\tau \propto h^2$  in all these experiments.

All estimators appearing in Theorems 4.5, 4.6 and 6.4 are computed except for  $\gamma_n$  and  $\beta_n$ : the mesh-change estimator  $\gamma_n = 0$  in our tests here, because the triangulation  $\mathcal{T}^n$  is in fact constant with respect to  $n$ ; for examples with  $\gamma_n \neq 0$  we refer to [LP12]. We do not track the data error indicator  $\beta_n$  either, since it can be shown to be of higher order in our case, due to regular nature of  $f$  and does not influence the difference between the two types of error estimators.

Comparing the results in Figure 2 (duality, Theorem 4.5), Figure 3 (max-duality, Corollary 4.6) and Figure 4 (energy, Theorem 6.4), we see that the last two offer a more “stable” long-time behavior, while the duality estimator yields better effectivity indexes for short times. Our interpretation for this is that it is due to a combination of Poincaré–Friedrichs constant and the use of residual-based elliptic estimators (which are known to have high effectivity indexes). A comparison using other elliptic estimators, such as gradient-recovery ones [LP12], is likely to yield better effectivity indexes and might be worth considering.

Otherwise, our conclusions are that there is no clear-cut advantage to any of the estimators, and in the linear case, where both estimators can be shown to be reliable, it is a matter of personal choice which one to use.

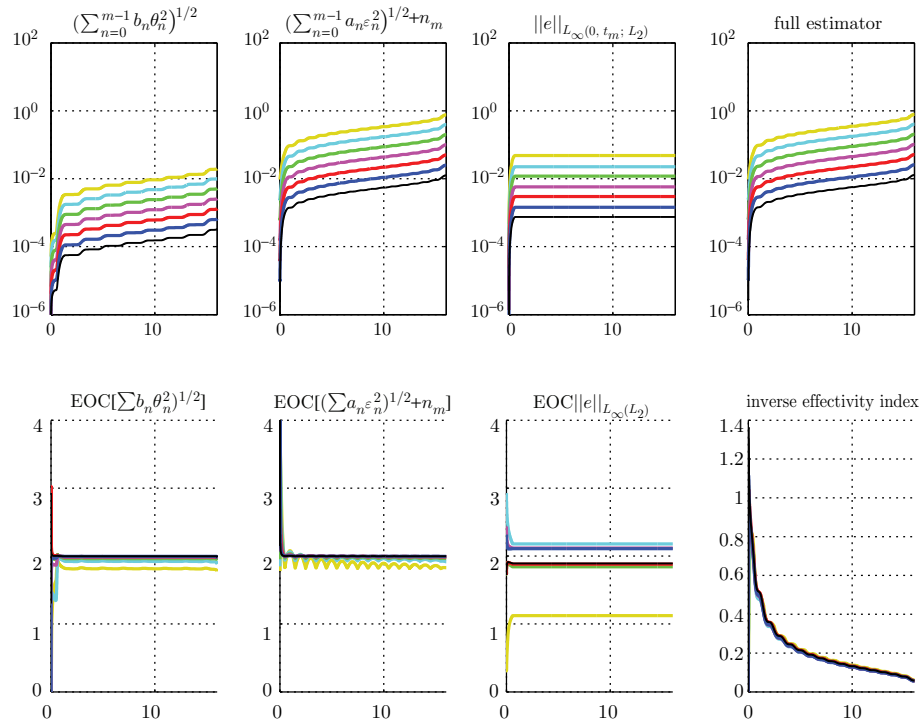


FIGURE 2. Experimental convergence of the exact error and the duality estimators of §4 with exact solution given by (7.6) with long time integration,  $T = 16$ . We use  $\mathbb{P}^1$  elements on uniform meshes and timestep coupled by  $\tau(i) = 0.05h(i)^2$  and  $h(i) = 2^{-i}$ ,  $i = 5, \dots, 11$ . The coupling is such that time and space estimators both have the same order of convergence as the exact error. We plot all quantities as functions of (PDE) time. Rates for each error/indicator can be read from the experimental order of convergence (EOC) of the associated part of the estimator together with its value on a logarithmic scale. The colour/grey-scale is such that black/dark is finest and yellow/light is coarsest. Since the benchmark problem is chosen to avoid any initial error the usual (straight) effectivity index is singular at time 0, so we employ the *inverse effectivity index*. The overall conclusions are summarized in §7.5. Here, we look at the time- $L_2$  behavior of the duality estimators appearing Theorem 4.5.

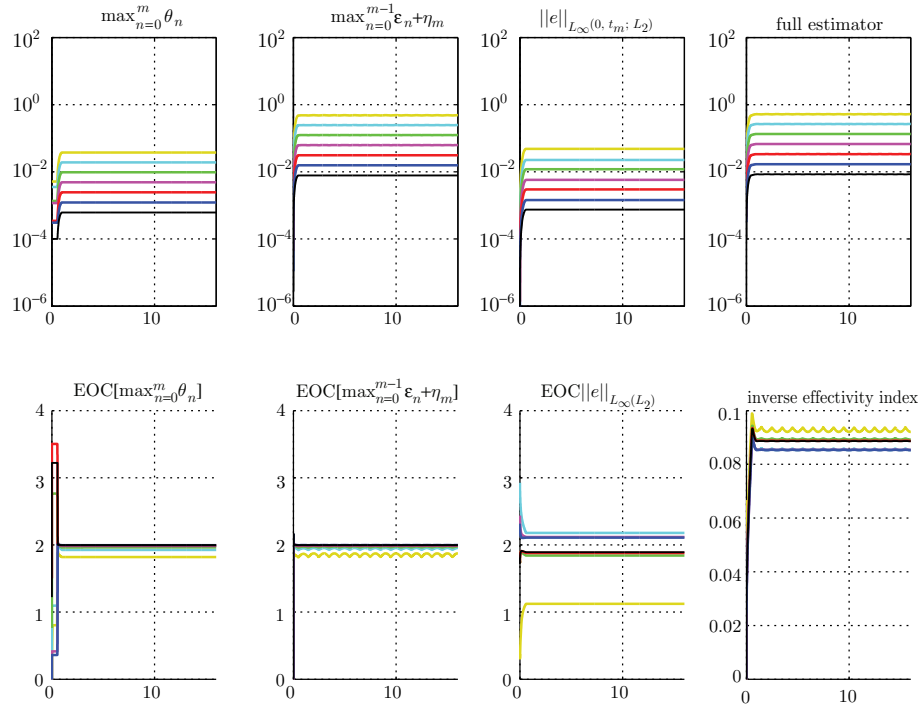


FIGURE 3. Experimental convergence of the exact error and the duality estimators of §4 with exact solution given by (7.6) with long time integration,  $T = 16$ . We use  $\mathbb{P}^1$  elements on uniform meshes and timestep coupled by  $\tau(i) = 0.05h(i)^2$  and  $h(i) = 2^{-i}$ ,  $i = 5, \dots, 11$ . The coupling is such that time and space estimators both have the same order of convergence as the exact error. We plot all quantities as functions of (PDE) time. Rates for each error/indicator can be read from the experimental order of convergence (EOC) of the associated part of the estimator together with its value on a logarithmic scale. The colour/grey-scale is such that black/dark is finest and yellow/light is coarsest. Since the benchmark problem is chosen to avoid any initial error the usual (straight) effectivity index is singular at time 0, so we employ the *inverse effectivity index*. The overall conclusions are summarized in §7.5. Here we study the time- $L_\infty$  estimator from Corollary 4.6. Compared to Figure 2, we see that replacing the  $L_2$  by  $L_\infty$  leads only to a deterioration of the effectivity index over short times, but improves the qualitative behavior for long integration times. This behaviour vindicates [EJ91]’s choice of “considering log-factors to be constants in most practical situations”, but also indicates that the log-factor may be too pessimistic.

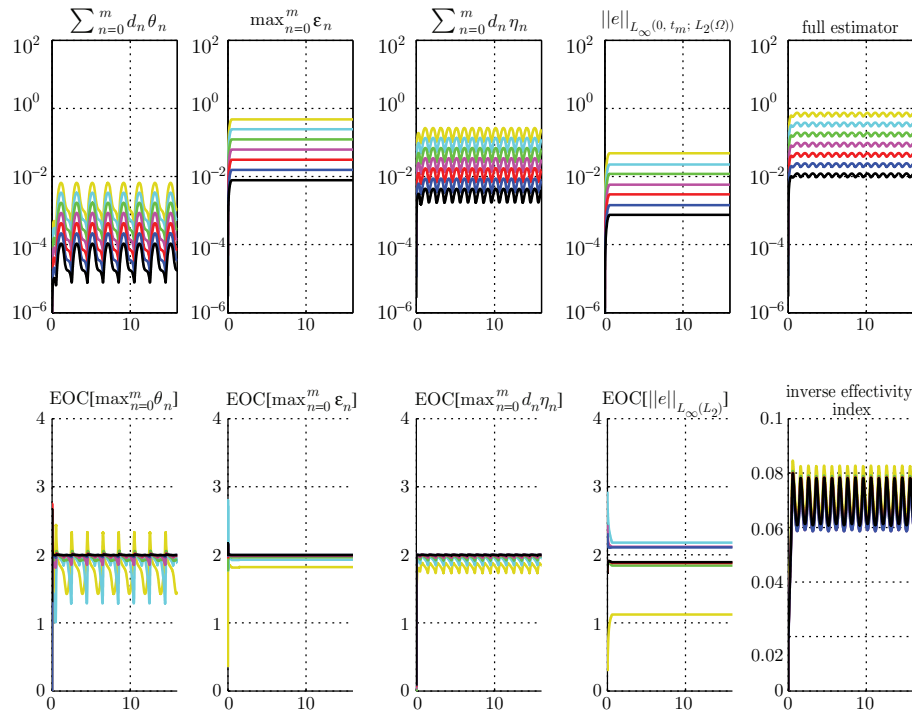


FIGURE 4. Convergence for of the energy-based estimator from Theorem 6.4. We use  $\mathbb{P}^1$  elements on uniform meshes and timestep coupled by  $\tau(i) = 0.05h(i)^2$  and  $h(i) = 2^{-i}$ ,  $i = 5, \dots, 11$ . The coupling is such that time and space estimators both have the same order of convergence as the exact error. We plot all quantities as functions of (PDE) time. Rates for each error/indicator can be read from the experimental order of convergence (EOC) of the associated part of the estimator together with its value on a logarithmic scale. The colour/grey-scale is such that black/dark is finest and yellow/light is coarsest. Since the benchmark problem is chosen to avoid any initial error the usual (straight) effectivity index is singular at time 0, so we employ the *inverse effectivity index*. The overall conclusions are summarized in §7.5. Problem (7.6) with time-oscillation factor long time integration  $T = 16$ , to be compared with Figure 2 and 3. Convergence rates, as in the duality-based estimators, optimal but the effectivity index is more oscillatory but more stable over long times with respect to time in this case, as expected given the better accumulation of the estimator in time (cf. Figure 1). The overall conclusion is that the energy estimator matches qualitatively the maximum-duality estimator of Corollary 4.6.

## ACKNOWLEDGMENTS

Part of this research originates from O.L.'s work with C.M. at FORTH in Crete in the framework of a Marie Curie fellowship and the HYKE RTN. T.P. was funded mostly by an EPSRC postgraduate research fellowship during this research. O.L. thanks Christoph Ortner and Sören Bartels for the interesting discussions about elliptic reconstruction and the energy approach at the Hausdorff Institute for Mathematics, Bonn, which we also thank for its kind generosity and outstanding hospitality.

## REFERENCES

- [AMN06] Georgios Akrivis, Charalambos Makridakis, and Ricardo H. Nochetto, *A posteriori error estimates for the Crank-Nicolson method for parabolic equations*, Math. Comp. **75** (2006), no. 254, 511–531, DOI 10.1090/S0025-5718-05-01800-4. MR2196979 (2007a:65114)
- [AO00] Mark Ainsworth and J. Tinsley Oden, *A Posteriori Error Estimation in Finite Element Analysis*, Wiley-Interscience [John Wiley & Sons], New York, 2000. MR1885308
- [BBM05] A. Bergam, C. Bernardi, and Z. Mghazli, *A posteriori analysis of the finite element discretization of some parabolic equations*, Math. Comp. **74** (2005), no. 251, 1117–1138 (electronic), DOI 10.1090/S0025-5718-04-01697-7. MR2136996 (2007c:65072)
- [BM11] Sören Bartels and Rüdiger Müller, *Quasi-optimal and robust a posteriori error estimates in  $L^\infty(L^2)$  for the approximation of Allen-Cahn equations past singularities*, Math. Comp. **80** (2011), no. 274, 761–780, DOI 10.1090/S0025-5718-2010-02444-5. MR2772095 (2012c:65152)
- [BR78] Ivo Babuška and Werner C. Rheinboldt, *Error estimates for adaptive finite element computations*, SIAM J. Numer. Anal. **15** (1978), no. 4, 736–754. MR58:3400
- [Bra01] Dietrich Braess, *Finite Elements*, 2nd ed., Cambridge University Press, Cambridge, 2001. Theory, fast solvers, and applications in solid mechanics; Translated from the 1992 German edition by Larry L. Schumaker. MR1827293 (2001k:65002)
- [BS07] Susanne C. Brenner and L. Ridgway Scott, *The Mathematical Theory of Finite Element Methods*, third ed., Texts in Applied Mathematics, vol. 15, Springer-Verlag, New York, 2007.
- [BV04] Christine Bernardi and Rüdiger Verfürth, *A posteriori error analysis of the fully discretized time-dependent Stokes equations*, M2AN Math. Model. Numer. Anal. **38** (2004), no. 3, 437–455, DOI 10.1051/m2an:2004021. MR2075754 (2005g:65131)
- [CF04] Zhiming Chen and Jia Feng, *An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems*, Math. Comp. **73** (2004), no. 247, 1167–1193 (electronic), DOI 10.1090/S0025-5718-04-01634-5. MR2047083 (2005e:65131)
- [Cia78] Philippe G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4. MR0520174 (58 #25001)
- [CKNS08] J. Manuel Cascon, Christian Kreuzer, Ricardo H. Nochetto, and Kunibert G. Siebert, *Quasi-optimal convergence rate for an adaptive finite element method*, SIAM J. Numer. Anal. **46** (2008), no. 5, 2524–2550, DOI 10.1137/07069047X. MR2421046 (2009h:65174)
- [dFN02] Javier de Frutos and Julia Novo, *Postprocessing the linear finite element method*, SIAM J. Numer. Anal. **40** (2002), no. 3, 805–819 (electronic), DOI 10.1137/S0036142900375438. MR1949394 (2004b:65148)
- [DLM09] Alan Demlow, Omar Lakkis, and Charalambos Makridakis, *A posteriori error estimates in the maximum norm for parabolic problems*, SIAM J. Numer. Anal. **47** (2009), no. 3, 2157–2176, DOI 10.1137/070708792. MR2519598 (2010e:65142)
- [DM10] Alan Demlow and Charalambos Makridakis, *Sharply local pointwise a posteriori error estimates for parabolic problems*, Math. Comp. **79** (2010), no. 271, 1233–1262, DOI 10.1090/S0025-5718-10-02346-X. MR2629992 (2011d:65280)

- [DS11] Alan Demlow and Rob Stevenson, *Convergence and quasi-optimality of an adaptive finite element method for controlling  $L_2$  errors*, Numer. Math. **117** (2011), no. 2, 185–218, DOI 10.1007/s00211-010-0349-9. MR2754849 (2011m:65278)
- [EJ91] Kenneth Eriksson and Claes Johnson, *Adaptive finite element methods for parabolic problems. I. A linear model problem*, SIAM J. Numer. Anal. **28** (1991), no. 1, 43–77, DOI 10.1137/0728003. MR1083324 (91m:65274)
- [EM09] Alexandre Ern and Sébastien Meunier, *A posteriori error analysis of Euler-Galerkin approximations to coupled elliptic-parabolic problems*, M2AN Math. Model. Numer. Anal. **43** (2009), no. 2, 353–375, DOI 10.1051/m2an:2008048. MR2512500 (2010d:65248)
- [Eva98] Lawrence C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 1998. MR1625845 (99e:35001)
- [GL10] Emmanuil H. Georgoulis and Omar Lakkis, *A posteriori error bounds for discontinuous galerkin methods for quasilinear parabolic problems*, Numerical Mathematics and Advanced Applications 2009 (Gunilla Kreiss, Per Lötstedt, Axel Målqvist, and Maya Neytcheva, eds.), Springer Berlin Heidelberg, 2010, pp. 351–358 (English).
- [GLV11] Emmanuil H. Georgoulis, Omar Lakkis, and Juha M. Virtanen, *A posteriori error control for discontinuous Galerkin methods for parabolic problems*, SIAM J. Numer. Anal. **49** (2011), no. 2, 427–458, DOI 10.1137/080722461. MR2784879 (2012d:65205)
- [KMSS12] Christian Kreuzer, Christian A. Möller, Alfred Schmidt, and Kunibert G. Siebert, *Design and convergence analysis for an adaptive discretization of the heat equation*, IMA J. Numer. Anal. **32** (2012), no. 4, 1375–1403, DOI 10.1093/imanum/drr026. MR2991832
- [Kos94] Igor Kossaczky, *A recursive approach to local mesh refinement in two and three dimensions*, J. Comput. Appl. Math. **55** (1994), no. 3, 275–288, DOI 10.1016/0377-0427(94)90034-5. MR1329875 (95m:65207)
- [LM06] Omar Lakkis and Charalambos Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp. **75** (2006), no. 256, 1627–1658, DOI 10.1090/S0025-5718-06-01858-8. MR2240628 (2007e:65122)
- [LN03] Xiaohai Liao and Ricardo H. Nochetto, *Local a posteriori error estimates and adaptive control of pollution effects*, Numer. Methods Partial Differential Equations **19** (2003), no. 4, 421–442, DOI 10.1002/num.10053. MR1980188 (2004c:65130)
- [LP12] Omar Lakkis and Tristan Pryer, *Gradient recovery in adaptive finite-element methods for parabolic problems*, IMA J. Numer. Anal. **32** (2012), no. 1, 246–278, DOI 10.1093/imanum/drr019. MR2875251
- [MN03] Charalambos Makridakis and Ricardo H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal. **41** (2003), no. 4, 1585–1594, DOI 10.1137/S0036142902406314. MR2034895 (2004k:65157)
- [Pic98] Marco Picasso, *Adaptive finite elements for a linear parabolic problem*, Comput. Methods Appl. Mech. Engrg. **167** (1998), no. 3-4, 223–237, DOI 10.1016/S0045-7825(98)00121-2. MR1673951 (2000b:65188)
- [SS05] Alfred Schmidt and Kunibert G. Siebert, *Design of adaptive Finite Element Software*, Lecture Notes in Computational Science and Engineering, vol. 42, Springer-Verlag, Berlin, 2005. The finite element toolbox ALBERTA; With 1 CD-ROM (Unix/Linux). MR2127659 (2005i:65003)
- [Tho06] Vidar Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer Series in Computational Mathematics, vol. 25, Springer-Verlag, Berlin, 2006. MR2249024 (2007b:65003)
- [Ver96] Rüdiger Verfürth, *A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner, Chichester-Stuttgart, 1996.
- [Ver03] R. Verfürth, *A posteriori error estimates for finite element discretizations of the heat equation*, Calcolo **40** (2003), no. 3, 195–212, DOI 10.1007/s10092-003-0073-2. MR2025602 (2005f:65131)
- [Whe73] Mary Fanett Wheeler, *A priori  $L_2$  error estimates for Galerkin approximations to parabolic partial differential equations*, SIAM J. Numer. Anal. **10** (1973), 723–759. MR0351124 (50 #3613)



DEPARTMENT OF MATHEMATICS, UNIVERSITY OF SUSSEX, BRIGHTON, GB-BN1 9QH, ENGLAND  
UNITED KINGDOM

*URL:* <http://www.maths.sussex.ac.uk/Staff/OL>

*E-mail address:* [o.lakkis@sussex.ac.uk](mailto: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 HERAK-  
LION, GREECE

*E-mail address:* [makr@tem.uoc.gr](mailto:makr@tem.uoc.gr)

SCHOOL OF MATHEMATICS, STATISTICS & ACTUARIAL SCIENCE, UNIVERSITY OF KENT, CAN-  
TERBURY, GB-CT2 7NF, ENGLAND UNITED KINGDOM

*E-mail address:* [T.Pryer@kent.ac.uk](mailto:T.Pryer@kent.ac.uk)