

SPACE AND TIME RECONSTRUCTIONS IN A POSTERIORI ANALYSIS OF EVOLUTION PROBLEMS *

CHARALAMBOS MAKRIDAKIS¹

Dedicated to Michel Crouzeix

Abstract. We review some recent results on a posteriori error analysis of evolution equations based on appropriate reconstruction operators. These operators are used to derive reformulations of the discretization methods as perturbed evolution pdes with source terms of a posteriori type. As a result energy or other standard techniques yield optimal a posteriori error estimates not known before.

1. INTRODUCTION

In this paper we discuss recent results related to a posteriori error control of evolution problems

$$u'(t) + A(u(t)) = 0,$$

with given initial value u^0 . Let U be an approximation to u obtained by a numerical scheme. For a given norm $\|\cdot\|$ we would like to show an estimate of the form

$$\|u - U\| \leq \eta(U)$$

where

- the estimator $\eta(U)$ is a computable quantity which depends on the approximate solution U and the data of the problem;
- $\eta(U)$ decreases with optimal order with respect to the mesh discretization parameters requiring the lowest possible regularity permitted by our problem;

in addition it will be desirable that the constants involved in the estimator $\eta(U)$ be explicit and/or easily computable. Such estimates are basic tools for adaptive mesh refinement/redistribution

* *Partially supported by the program Pythagoras of EPEAEK II funded by the European Commission and the Greek Ministry of National Education.*

¹ Department of Applied Mathematics, University of Crete, 71409 Heraklion-Crete, Greece and Institute of Applied and Computational Mathematics, FORTH, 71110 Heraklion-Crete, Greece. <http://www.tem.uoc.gr/~makr>, email: makr@tem.uoc.gr

methods for PDEs and have been the object of intense study in recent years. In addition a posteriori estimates provide a new point of view in the theoretical investigation of the behavior of schemes. Especially they open new directions for successful theoretical backup of adaptive methods for singular nonlinear problems.

In contrast to the elliptic case, [2, 28], available results for a posteriori error control of evolution problems, even in the linear case, are limited. In the late 80's and the beginning of 90's several results for the discontinuous Galerkin time discretization method based on duality arguments were proved mainly by Errikson, Johnson, Thomée, see [15], [27] and their references. Space-discrete and fully-discrete schemes for parabolic equations (discretized mainly by backward Euler in time) were considered in several papers lately, [26], [1, 5], [7], [6] and their references. The analysis is based on energy methods and yields a posteriori bounds of optimal order in $L^2(H^1)$ or other norms. Nonlinear problems were considered in [29], [25], [24] and [6].

Error representation formulas and reconstruction. In most of the above studies the analysis is based on error representation formulas obtained by a *direct* comparison of

$$u \text{ and } U.$$

The Galerkin orthogonality is then a main ingredient in the proofs. Our approach is based on the comparison with an intermediate function \hat{U} which is called *reconstruction*. In particular our aim is to find an appropriate *reconstruction* \hat{U} which is defined through the approximation U . \hat{U} is a “continuous object” obtained by applying an appropriate operator to U . Then we have to estimate

$$u - \hat{U} \text{ and } \hat{U} - U.$$

The use of an intermediate auxiliary function, a *projection* into the discrete space, is very natural, and is the typical case in the a priori error estimates of evolution PDEs, cf. e.g., [27]. The typical case in the a priori analysis is to use appropriate projections and to form the corresponding discrete error equations. Then the a priori control of the total error is reduced on proving discrete stability bounds for these equations. The situation in the a posteriori analysis is similar with the roles of “continuous” and “discrete” sort of reversed.

In the next section we describe the basic ideas which when implemented appropriately will lead to optimal a posteriori bounds. The first paper where ideas of this type were applied successfully in a posteriori error analysis was the work of Nochetto, Savaré and Verdi [24] for backward Euler time discrete approximations of certain classes of nonlinear parabolic problems.

2. APPROPRIATE RECONSTRUCTIONS

The function \hat{U} is important to satisfy certain properties that will lead to optimal a posteriori bounds. It will be useful to use the compact notation $B(u) = 0$ for the evolution problem at hand and $B_h(U) = 0$ for the given approximation scheme. I.e., the exact and approximate solutions of the evolution PDE satisfy

Exact solution u of : $B(u) = 0$, $u \in S$ (evolution PDE),

Discrete approximation U : $B_h(U) = 0$, $U \in S_h$ (time, space or fully discrete scheme).

Then we seek a

$$\begin{aligned} & \text{reconstruction } \hat{U} \text{ of the discrete approximation } U \\ U & \mapsto \hat{U} \in S, \end{aligned}$$

such that at one hand

- $U - \hat{U}$ is *computable* or can be bounded by *a posteriori (computable)* quantities of optimal order,

and in addition \hat{U} should satisfy three basic properties

- B is applicable to \hat{U} , i.e., $R = B(\hat{U})$ is *well defined* and *not singular*
- $R = B(\hat{U})$ is *computable* or can be bounded by *a posteriori (computable)* quantities
- The *reconstruction residual* $R = B(\hat{U})$ is a perturbation of the PDE of optimal order.

The above properties are not trivial to obtain and heavily depend on the problem and the discretization at hand. In the sequel we will see various choices of reconstructions for space, time or fully-discrete schemes.

Let us emphasize that for

$$B(u(t)) = u'(t) + A(u(t)),$$

the above properties imply that

$$\hat{U}'(t) + A(\hat{U}(t)) = R(t), \quad (2.1)$$

where $R(t)$ is an *a posteriori*-type term of optimal order.

2.1. Why is a comparison with a *reconstruction* useful?

The pointwise equation (2.1) is a useful tool but at this point it is not clear why one should prefer to work with it rather than the residual equation of the original approximation U :

$$B(U) = R_U. \quad (2.2)$$

One can see that even in the case of linear evolution problems (2.2) is not appropriate for mainly one of the following reasons:

- $B(U)$ does not make sense usually – or gives rise to singular right hand sides
- In evolution problems the residual $R_U = B(U)$ is *not of optimal* order
- Usually U does not satisfy structural conditions that u satisfies. Then the reconstruction \hat{U} can be defined to cure this problem by satisfying these structural conditions pointwise. A typical case is the divergence free condition in Stokes equations.

2.2. Deriving *a posteriori* estimates using the reconstruction

Linear Problems. In the linear case the derivation of the *a posteriori* estimates follows by direct stability PDE bounds to the Error Equation. In fact by the properties of the function \hat{U}

- *Error Equation for the linear case:* $B(\hat{e}) = B(u - \hat{U}) = -R$,

i.e., the error $\hat{e} = u - \hat{U}$ satisfies the same PDE as u but with the additional *a posteriori* source term R . Then PDE stability arguments imply

$$\|u - \hat{U}\|_A \leq C_{\text{pde}} \|R\|_B.$$

Then we obtain the *a posteriori* estimate

$$\|u - U\|_A \leq \|U - \hat{U}\|_A + C_{\text{pde}}\|R\|_B.$$

An *important remark* here is that *the stability step above is at our disposal and can be done by any PDE stability method suitable for our purposes.*

Nonlinear Problems. In the nonlinear case the derivation of the *a posteriori* estimates is, of course, more involved. Then the properties of the function \hat{U} lead to the nonlinear relation

- *Error Equation for the nonlinear case:* $B(u) - B(\hat{U}) = -R,$

i.e., \hat{U} satisfies a perturbation of the original PDE with source term $R = -B(\hat{U})$. In that case perturbation PDE estimates are needed to guarantee

$$\|u - \hat{U}\|_A \leq C_{\text{pde}}\|R\|_B. \quad (2.3)$$

The above step is very delicate. Its completion requires certain assumptions/restrictions on the perturbation R and on the closeness of \hat{U} to u , and the right choice of the PDE stability method used. On the other hand the fact that \hat{U} satisfies (2.1) is a very useful starting point for the derivation of meaningful *a posteriori* bounds for nonlinear evolution problems. A very attractive scenario would be to show *conditional* estimates in the spirit of [10], [21], [23]. I.e., to complete the stability step above under solely *a posteriori*, and thus in principle verifiable, assumptions on \hat{U} and therefore on U . By this we mean that we can show that (2.3) holds provided that

$$\Theta(U, \hat{U}) \leq \beta.$$

Here β is a fixed known number and Θ is an *a posteriori* functional, i.e. it depends on U, \hat{U} and the data but not on the unknown solution u . The final *a posteriori* estimate will have the same form:

$$\|u - U\|_A \leq \|U - \hat{U}\|_A + C_{\text{pde}}\|R\|_B.$$

One should note that *nonlinear PDEs are special* and therefore to obtain meaningful *a posteriori* results under realistic assumptions one has to take full advantage of the particular stability and structural properties of the equation under study.

3. TIME RECONSTRUCTIONS

Next we will see time reconstructions for first or second order time discrete schemes and then for higher order schemes. In this note we consider only linear parabolic problems,

$$\begin{cases} u'(t) + Au(t) = f(t), & 0 < t < T, \\ u(0) = u^0, \end{cases} \quad (3.1)$$

with a given $f : [0, T] \rightarrow H$. Here A is a positive definite, selfadjoint, linear operator on a Hilbert space $(H, \langle \cdot, \cdot \rangle)$ with domain $D(A)$ dense in H , and $u^0 \in H$. We let $V := D(A^{1/2})$ and denote the norms in H and in V by $|\cdot|$ and $\|\cdot\|$, with $\|v\| := |A^{1/2}v| = \langle Av, v \rangle^{1/2}$, respectively. Let V^* be the dual of V ($V \subset H \subset V^*$). We still denote by $\langle \cdot, \cdot \rangle$ the duality pairing between V^* and V , and by $\|\cdot\|_*$ the dual norm on V^* .

For *a posteriori* estimates for nonlinear problems proved with appropriate reconstructions we refer to [3, 4, 9, 23, 24]. Conditional estimates in the spirit of [10, 21] were obtained in [23] for the

minimal surface operator by energy methods, and in [9] for general quasilinear parabolic problems in Banach spaces by using semigroup-type techniques.

3.1. Two simple examples: Backward Euler and Crank-Nicolson

Backward Euler. Let $0 = t_0 < t_1 < \dots < t_N = T$ be a partition of $[0, T]$, $I_n := (t_n, t_{n+1}]$, and $k_n := t_{n+1} - t_n$. The Backward Euler method for (3.1) is

$$\frac{1}{k_n}(U^{n+1} - U^n) + AU^{n+1} = f_k^{n+1}. \quad (3.2)$$

$f_k^{n+1} = f(t_{n+1})$. The alternative discretization of f , $f_k^{n+1} = \frac{1}{k_n} \int_{I_n} f(s) ds$, gives rise to the piecewise constant Discontinuous Galerkin method (dG0). One can consider approximations to be piecewise constant in time. I.e. define U as the piecewise constant function and the projection $\Pi_0 f$ of f :

$$U|_{I_n} \in \mathbb{P}_0(I_n), \quad U|_{I_n} = U^{n+1}, \quad \Pi_0 f = f_k^{n+1}.$$

In this case the equation for the residual $R_U = B(U)$ is singular since U is discontinuous in time. The *Backward Euler Reconstruction* \hat{U} is then defined as the natural continuous piecewise linear interpolant at the time nodes. Then at each I_n , $\hat{U}' = \frac{1}{k_n}(U^{n+1} - U^n)$, and the scheme can be written as

$$\hat{U}'(t) + A\hat{U}(t) = \Pi_0 f + A[\hat{U}(t) - U(t)], \quad t \in I_n. \quad (3.3)$$

Then

- $\hat{U}(t) - U(t) = \hat{U}(t) - U^{n+1} = \ell_0^n(t)(U^n - U^{n+1})$

where $\hat{U}(t) = \ell_0^n(t)U^n + \ell_1^n(t)U^{n+1}$. It is clear now that (3.3) is of the form (2.1) with $R = \Pi_0 f + \ell_0^n(t)A[U^n - U^{n+1}]$, the *a posteriori* right hand side. Pointwise equations of the form (2.1), (3.3) were first used in a posteriori analysis in the important paper by Nochetto, Savaré, and Verdi, [24], for time discrete approximations by Backward Euler of linear and several nonlinear problems.

Error equation and estimate. Let $\hat{e} = u - \hat{U}(t)$; then the error equation in the linear case reads

$$\hat{e}'(t) + A\hat{e}(t) = (f - \Pi_0 f) - A[\hat{U}(t) - U(t)], \quad t \in I_n. \quad (3.4)$$

By employing energy arguments we conclude to the a posteriori estimate

$$\max_{0 \leq t \leq T} |\hat{e}|^2 + \int_0^T \|\hat{e}\|^2 dt \leq \alpha \left(\sum_{n=0}^{N-1} k_n \|A^{1/2}(U^{n+1} - U^n)\|^2 + \int_0^T \|f - \Pi_0 f\|_*^2 \right).$$

Semigroup approach : estimates via Duhamel's principle. We shall use Duhamel's principle in the above error equation. Let $E_A(t)$ be the solution operator of the homogeneous equation $u'(t) + Au(t) = 0$, $u(0) = w$, i.e., $u(t) = E_A(t)w$. It is well known that the family of operators $E_A(t)$ has several nice properties, in particular it is a semigroup of contractions on H with generator the operator A . Duhamel's principle states in the case where $f = 0$

$$\hat{e}(t) = \int_0^t E_A(t-s)[A[U(t) - \hat{U}(t)]] ds.$$

Then, using the smoothing property, $|E_A(t) A^\ell w| \leq C_A \frac{1}{t^{\ell-m}} |A^m w|$, $\ell \geq m \geq 0$, and adopting in the Backward Euler case arguments of [3] we conclude

$$|\hat{e}(t^{n+1})| \leq C(1 + \ln(\frac{t^{n+1}}{k_n})) \sup_{0 \leq j \leq n} |U^{j+1} - U^j|$$

which is an alternative a posteriori bound. Note that a similar estimate can be obtained by duality arguments, [27], [14], but here we use the strong stability of the *forward* problem.

Crank-Nicolson. The other popular low order method has different structure. Usually, Crank-Nicolson is associated directly to a continuous piecewise linear function U , but now the problem is that $R_U = B(U)$ is suboptimal. To see this and to motivate the definition of the appropriate reconstruction, consider the Crank-Nicolson-Galerkin method which is the lowest order method of *Continuous Galerkin time discretization schemes*. To this end let \mathcal{V}_q be the space of *continuous* functions $g : (0, T] \rightarrow D(A)$ which are polynomials of degree q in I_n , $g|_{I_n}(t) = \sum_{j=0}^q t^j w_j$, $w_j \in D(A)$. We denote by $\mathcal{V}_q(I_n)$ the space consisting of the restrictions of the elements of \mathcal{V}_q to I_n . The spaces \mathcal{H}_q and $\mathcal{H}_q(I_n)$ are defined analogously by requiring $w_j \in H$. The time discrete *continuous Galerkin* approximation to u is defined as follows: We seek $U \in \mathcal{V}_q$ such that

$$\int_{I_n} [\langle U', v \rangle + \langle F(t, U), v \rangle] dt = \int_{I_n} [\langle U', v \rangle + \langle AU, v \rangle - \langle f(t), v \rangle] dt = 0 \quad \forall v \in \mathcal{V}_{q-1}(I_n) \quad (3.5)$$

for $n = 0, \dots, N-1$, with $F(t, v) = Av - f(t)$, and $U(0) = u(0)$. Since the Crank-Nicolson-Galerkin (CN-G) approximation U corresponds to $q = 1$ it satisfies the following pointwise equation

$$U'(t) + \mathcal{P}_0 F(t, U(t)) = 0 \quad \forall t \in I_n$$

with \mathcal{P}_r denoting the L^2 orthogonal projection operator onto $\mathcal{V}_r(I_n)$.

Failure of the energy method. Let $R_U(t) = U'(t) + F(t, U(t)) = B(U(t))$, $R(t) = F(t, U(t)) - \mathcal{P}_0 F(t, U(t))$. Clearly, this pointwise equation is not appropriate, since even in the case of the trivial ode, $F(t, v) = F(t)$, $R(t) = F(t) - \mathcal{P}_0 F(t)$, i.e., $R(t)$ can only be of *first order* although Crank-Nicolson-Galerkin is second order approximation to u . Thus energy methods applied directly to the pointwise equation of U will not work.

The CN-G reconstruction. We introduce, [3], $\hat{U} \in \mathcal{V}_2(I_n)$, $\hat{U}(t) := U(t^n) - \int_{t^n}^t \mathcal{P}_1 F(s, U(s)) ds$, $\forall t \in I_n$. Then, \hat{U} satisfies the following pointwise equation

$$\hat{U}'(t) + AU = \mathcal{P}_1 f(t) \quad \forall t \in I_n,$$

and several desirable properties that allow an analysis of optimal a posteriori error bounds by energy methods, cf. [3] for details.

The *Crank-Nicolson method* can be written as variation of the CN-G method

$$U'(t) + AU^{n+\frac{1}{2}} = f(t^{n+\frac{1}{2}}), \quad n = 1, \dots, N,$$

with $U^0 := u^0$, and $v^{n+\frac{1}{2}} := \frac{1}{2}(v^{n+1} + v^n)$, $n = 0, \dots, N-1$. Here as before U is the continuous piecewise linear function such that $U(t_n) = U^n$, and thus $U'(t) = \frac{1}{k_n}(U^{n+1} - U^n)$. The above piecewise equation leads to suboptimal energy bounds. The definition of *Crank-Nicolson reconstruction* \hat{U} of U , is motivated by this of CN-G reconstruction. We note however that a more involved choice for approximating the contribution of f is necessary. To this end, [3], we let $\varphi : I_n \rightarrow H$ be the

linear interpolant of f at the nodes t^n and $t^{n+\frac{1}{2}}$, Define a piecewise quadratic polynomial Φ by $\Phi(t) = \int_{t^n}^t \varphi(t) dt = (t-t^n)f(t^{n+\frac{1}{2}}) - \frac{1}{k_n}(t-t^n)(t^{n+1}-t)[f(t^{n+\frac{1}{2}}) - f(t^n)]$. Then the *Crank-Nicolson reconstruction* \hat{U} of U is defined

$$\hat{U}(t) := U^n - \int_{t^n}^t AU(s) ds + \Phi(t).$$

Error Control. To simplify the presentation, we assume next that $f = 0$ and we refer to [3] for detailed results including the verification of the optimality of the estimates. When $f = 0$ CN-G and Crank-Nicolson coincide since the mid-point rule is exact for linear polynomials. The difference $U - \hat{U}$ is given explicitly,

$$U(t) - \hat{U}(t) = -\frac{(t-t^n)(t^{n+1}-t)}{2k_n}A(U^{n+1} - U^n). \quad (3.6)$$

For $e = u - U$ and $\hat{e} = u - \hat{U}$, energy estimates will imply for an explicit constant β ,

$$|\hat{e}(t^{m+1})|^2 + \int_0^{t^{m+1}} \|\hat{e}(s)\|^2 ds \leq \beta \sum_{n=0}^m k_n^5 |A^{3/2} \bar{\partial} U^n|^2, \quad \text{where } \bar{\partial} U^n = k_n^{-1}(U^{n+1} - U^n).$$

In addition the following *Lower Bound* for $L^2(V)$ holds

$$\frac{\beta}{24} \sum_{n=0}^{m-1} k_n^5 |A^{3/2} \bar{\partial} U^n|^2 \leq \int_0^t \left(\|e(s)\|^2 + \frac{1}{2} \|\hat{e}(s)\|^2 \right) ds \leq \frac{\beta}{2} \sum_{n=0}^{m-1} k_n^5 |A^{3/2} \bar{\partial} U^n|^2.$$

Finally, arguments based on Duhamel's principle provide the alternative estimate at the time nodes,

$$|\hat{e}(t^{n+1})| \leq C \left(1 + \ln \frac{t^{n+1}}{k_n} \right) \max_{0 \leq j \leq n} k_j |A(U^{j+1} - U^j)|.$$

3.2. Higher order Galerkin and Runge-Kutta methods

We are now interested in the *continuous* Galerkin (cG) and *Runge-Kutta collocation* (RK-C) time-stepping schemes. RK-C schemes consist the most important class of implicit Runge-Kutta methods. Following [4] we cast these methods in a unified form with the aid of a *projection operator* Π_ℓ onto $\mathcal{H}_\ell(I_n)$, $n = 1, \dots, N$: Seek $U \in \mathcal{V}_q$, such that

$$U'(t) + \Pi_{q-1}AU(t) = \Pi_{q-1}f(t) \quad \forall t \in I_n, \quad n = 1, \dots, N \quad U(0) = u^0. \quad (3.7)$$

Since all terms in this equation belong to $\mathcal{V}_{q-1}(I_n)$, (3.7) admits the Galerkin formulation

$$\int_{I_n} [\langle U', v \rangle + \langle \Pi_{q-1}AU, v \rangle] dt = \int_{I_n} \langle \Pi_{q-1}f, v \rangle dt \quad \forall v \in \mathcal{V}_{q-1}(I_n), \quad (3.8)$$

The *continuous* Galerkin method corresponds to the choice $\Pi_{q-1} := \mathcal{P}_{q-1}$, with \mathcal{P}_ℓ denoting the L^2 orthogonal projection operator onto $\mathcal{H}_\ell(I_n)$. The *Runge-Kutta collocation* methods are described by this formulation. In fact, [4], all RK-C methods can be obtained by choosing $\Pi_{q-1} := \mathcal{I}_{q-1}$, with \mathcal{I}_{q-1} denoting the interpolation operator by elements of $\mathcal{V}_{q-1}(I_n)$ at the nodes $t^n + \tau_i k_n$, $i = 1, \dots, q$, $n = 0, \dots, N-1$, with appropriate $0 \leq \tau_1 < \dots < \tau_q \leq 1$.

Reconstruction. It is easy to observe that $R_U = B(U)$ is suboptimal for these methods. The higher order *reconstruction* \hat{U} is required to belong in $\mathcal{H}_{q+1}(I_n)$. Its definition is based on appropriate projection operators $\hat{\Pi}_q$ onto $\mathcal{H}_q(I_n)$, $n = 1, \dots, N$. Then, [4],

$$\hat{U}(t) := U(t^{n-1}) - \int_{t^{n-1}}^t [AU(s) - \hat{\Pi}_q f(s)] ds \quad \forall t \in I_n. \quad (3.9)$$

A fundamental property which $\hat{\Pi}_q$ satisfy by construction is that for some $r \geq 0$

$$\int_{I_n} \langle (\hat{\Pi}_q - \Pi_{q-1})w(s), v(s) \rangle ds = 0 \quad \forall w \in C(I_n; H), \quad v \in \mathcal{H}_r(I_n). \quad (3.10)$$

In [4] the various properties of $\hat{\Pi}_q$, Π_{q-1} are studied in detail. Based then on the pointwise equation

$$\hat{U}'(t) + AU(t) = \hat{\Pi}_q f(t) \quad \forall t \in I_n, \quad (3.11)$$

and the approximation properties of $\hat{\Pi}_q$ energy a posteriori bounds are shown.

Nodal superconvergence. We next examine the phenomenon of *superconvergence* at the time nodes $\{t^n\}_{n=1}^N$ from a posteriori point of view. Superconvergence is attained if the interpolatory quadrature formula with nodes τ_1, \dots, τ_q integrates elements of \mathbb{P}_{q+r} exactly. Then, if the methods of this section are applied to a non-stiff ODE the error at the nodes will be $\mathcal{O}(k^{q+1+r})$. We let $p := q + 1 + r$ be the superconvergence order at the nodes. The *superconvergence order* for RK-C is the *classical order* of the method in the standard terminology of RK methods. Since the seminal work of Crouzeix [8] it is known that this order is limited, even for the linear problem (3.1), by requiring nontrivial conditions of the form

$$f \in D(A^\rho), \quad U^0 \in D(A^{\rho+1}) \quad (3.12)$$

which may fail to be fulfilled in applications. This lack of superconvergence at the nodes is usually called *order reduction* in the literature. A corresponding a posteriori analysis is carried out in [4]. The main estimate can be stated roughly as follows: Let the super-order of a q -stage RK-C method satisfy $p \in \{q + 2, \dots, 2q\}$. Then the following a posteriori error estimate is valid at the nodes $\{t^n\}_{n=1}^N$ for linear equations

$$|\hat{e}(t^n)| \leq C_I \left(\left(\ln \frac{t^n}{k_n} \right)^{1/2} + 1 \right) (\mathcal{E}_1 + \mathcal{E}_2),$$

with

$$\begin{aligned} \mathcal{E}_1 &= \max_{1 \leq m \leq n-1} \left(k_m^{p-q-1} |A^{p-q-2}(\rho_{\hat{U}} + \rho_f)|_{L^\infty(I_m)}, k_n^{p-q-2} \int_{I_n} |A^{p-q-2}(\rho_{\hat{U}} + \rho_f)| dt \right), \\ \mathcal{E}_2 &= \sum_{j=0}^{p-q-2} \max_{1 \leq m \leq n-1} \left(k_m^j |A^{j-1}(f - \hat{I}_{p-j-1}f)|_{L^\infty(I_m)}, k_n^{j-1} \int_{I_n} |A^{j-1}(f - \hat{I}_{p-j-1}f)| dt \right). \end{aligned}$$

Here $\rho_{\hat{U}} := A(U - \hat{U})$, $\rho_f = f - \hat{\Pi}_q f$ and \hat{I}_ℓ are appropriate polynomial interpolation operators of degree ℓ . The order of these estimators is p , provided the solution is smooth. In addition, note that we avoid time derivatives of f in the final estimate by introducing the intermediate terms $k_m^j |A^{j-1}(f - \hat{I}_{p-j-1}f)|_{L^\infty(I_m)}$, cf. [4] for details.

3.3. DG methods

Next we briefly discuss the discontinuous Galerkin time discrete scheme and the corresponding reconstruction. We denote \mathcal{V}_q^d the space of possibly discontinuous piecewise polynomials in time of degree $\leq q$, i.e., the discontinuous version of \mathcal{V}_q . The *discontinuous Galerkin method* dG(q) of order $q \geq 1$ is defined as follows [27]: given $U_0 := u_0$, seek $U \in \mathcal{V}_q^d$ such that

$$\int_{I_n} \left(\langle U', v \rangle + \langle F(U), v \rangle \right) dt + \langle U_n^+ - U_n, v_n^+ \rangle = 0, \quad \forall v \in \mathcal{V}_q^d(I_n), \quad 0 \leq n \leq N-1, \quad (3.13)$$

where $v_n := v(t_n)$, $v_n^+ := \lim_{s \downarrow 0} v(t_n + s)$. A detailed a posteriori analysis for the application of this method in various linear and nonlinear problems is presented in [23]. The main tool is the corresponding reconstruction introduced in [23]: \hat{U} is required to belong in \mathcal{V}_{q+1} , i.e., it will be a globally continuous function in time such that

$$\hat{U}_n^+ = U_n \quad \text{and} \quad \int_{I_n} \langle \hat{U}', v \rangle dt = \int_{I_n} \langle U', v \rangle dt + \langle U_n^+ - U_n, v_n^+ \rangle, \quad \forall v \in \mathcal{V}_q^d(I_n). \quad (3.14)$$

In [23] it is proved that \hat{U} is well defined, is continuous, and it coincides with U at the $q+1$ Radau points of each I_n . Therefore the following representation holds

$$(\hat{U} - U)|_{I_n}(t) = \hat{\ell}_{n,0}(t)(U_n - U_n^+), \quad \forall t \in I_n, \quad (3.15)$$

where $\hat{\ell}_{n,0}$ is the Lagrange polynomial of degree $q+1$ which equals 0 at the $q+1$ points of I_n and 1 at t_n . All the estimators in the linear case have very simple form and depend only on the jumps $U_n - U_n^+$, cf. [23] for detailed results.

4. SPACE AND SPACE-TIME RECONSTRUCTIONS

4.1. Space discrete finite element approximations

To define the space discrete formulation we specify the choice of A , being a positive definite, selfadjoint, second order elliptic operator over a open polygonal domain Ω equipped with zero Dirichlet boundary conditions. We will use then the notation $H := L^2(\Omega)$, $V := H_0^1(\Omega)$ $V^* := H^{-1}(\Omega)$ be the dual of V . The corresponding norms are denoted by $\|v\|_V$, $\|\cdot\|_H = \|\cdot\|_{L^2}$. Let $a(\cdot, \cdot)$ be the bilinear form associated to A . A weak formulation of (3.1) is then

$$\langle u_t(t), v \rangle + a(u(t), v) = \langle f, v \rangle \quad \text{for all } v \in V, \text{ a.e. } t \in [0, T],$$

with u^0 given. Let $V_h = \{\chi \in H_0^1(\Omega) : \chi|_K \in \mathbb{P}_s(K) \quad \forall K \in \mathcal{T}_h\}$, where $\mathbb{P}_s(K)$ is the space of polynomials of degree $\leq s$ over K , be a standard finite element space, see e.g. [20, 22] for details of the finite element framework used. The spatially-discrete *finite element approximation* $u_h : [0, T] \rightarrow V_h$ is defined by

$$\begin{aligned} \langle u_{h,t}, \chi \rangle + a(u_h, \chi) &= \langle f, \chi \rangle \quad \text{for all } \chi \in V_h, \text{ a.e. } t \in [0, T], \\ u_h(\cdot, 0) &= u_h^0 \in V_h. \end{aligned}$$

Direct a posteriori error estimation. Let us consider the linear functional $R_{u_h}(t) \in V^*$, defined by

$$\langle R_{u_h}, v \rangle = \langle u_{h,t}, v \rangle + a(u_h, v) - \langle f, v \rangle. \quad (4.1)$$

Then, denoting the error to be estimated by $e = u - u_h$, the following energy error identity follows

$$\frac{1}{2}\|e(T)\|_H^2 + \int_0^T a(e(t), e(t))dt = \frac{1}{2}\|e(0)\|_H^2 + \int_0^T \langle R_{u_h}(t), e \rangle dt. \quad (4.2)$$

It is clear that $\|R_{u_h}(t)\|_{V^*}$ dominates the error. On the other hand, Galerkin orthogonality, implies $\langle R_{u_h}, v \rangle = \langle u_{h,t}, v - I_h v \rangle + a(u_h, v - I_h v) - \langle f, v - I_h v \rangle$. Therefore, due to the presence of $a(u_h, e - I_h e)$, the resulting a posteriori estimator for $\|R_{u_h}(t)\|_{V^*}$ will be of optimal order in $V = H_0^1$, as it corresponds to an estimate of $a(e, e)$, but *suboptimal* in $H = L^2$. This means that (4.2) cannot lead to optimal estimator in $L^\infty(0, T; H)$ due to the lack of optimality of $\|R_{u_h}(t)\|_{V^*}$. This issue was resolved in [22] for semidiscrete schemes and in [20] for fully discrete schemes, cf. Section 4.5.

4.2. The elliptic reconstruction

Let $A_h : V_h \rightarrow V_h$ be the following discrete version of A :

$$\langle A_h v, \chi \rangle = a(v, \chi) \quad \text{for all } \chi \in V_h. \quad (4.3)$$

The *elliptic reconstruction* \hat{U} of u_h , [22], is defined through an operator $\mathcal{R} : V_h \rightarrow V$. For fixed $t \in [0, T]$, $\hat{U} = \mathcal{R}u_h \in H_0^1(\Omega)$ is defined to be the solution of the elliptic problem

$$a(\hat{U}(t), v) = \langle g_h(t), v \rangle \quad \text{for all } v \in H_0^1(\Omega), t \in [0, T], \quad (4.4)$$

where

$$g_h := A_h u_h - f_h + f, \quad f_h = Pf \text{ the } L^2\text{-projection of } f \text{ into } V_h. \quad (4.5)$$

The following two properties of \hat{U} are fundamental:

- a:** \hat{U} satisfies : $A\hat{U} = A_h u_h - f_h + f$
- b:** u_h is the finite element solution of the elliptic problem whose exact solution is the elliptic reconstruction \hat{U} .

In fact the last statement follows from the observation

$$a(\hat{U}, \varphi) = a(u_h, \varphi) - \langle f_h - f, \varphi \rangle = a(u_h, \varphi) \quad \text{for all } \varphi \in V_h. \quad (4.6)$$

4.3. A posteriori error estimates

By the definitions of u_h and \hat{U} , we have

$$u_{h,t} + A\hat{U} = f,$$

whence \hat{U} satisfies the following pointwise equation

$$\hat{U}_t + A\hat{U} = f + (\hat{U} - u_h)_t.$$

Thus with the notation of Section 2, $B(\hat{U}) = R = (\hat{U} - u_h)_t$. We will see that \hat{U} is an appropriate reconstruction in the sense of Section 2. The main difference with the time reconstructions considered in Section 3 is that $\hat{U} - u_h$ is not any more computable, but rather can be estimated by computable estimators of optimal order.

The error equation for $\hat{e} = u - \hat{U}$, $\hat{e}_t + A\hat{e} = (u_h - \hat{U})_t$ implies, [22],

$$\|\hat{e}(t)\|_H^2 + \int_0^t \|\hat{e}(s)\|_V^2 ds \leq \|u(0) - \hat{U}(0)\|_H^2 + \int_0^t \|u_{h,t} - \mathcal{R}u_{h,t}\|_{V^*}^2 ds \quad (4.7)$$

where we used the fact $(\mathcal{R}u_h)_t = \mathcal{R}u_{h,t}$. Therefore, as in Section 2.2,

$$\|u - u_h(t)\|_H \leq \|u_h(t) - \mathcal{R}u_h(t)\|_H + \|u(0) - \hat{U}(0)\|_H + \left(\int_0^t \|u_{h,t} - \mathcal{R}u_{h,t}\|_{V^*}^2 ds \right)^{1/2} \quad (4.8)$$

Elliptic reconstruction vs elliptic projection. It is interesting to observe that the elliptic reconstruction is an “*a posteriori dual*” to Wheeler’s elliptic projection $P_1 : V \rightarrow V_h$ defined by $a(P_1 v, \chi) = a(P_1 v, \chi) \forall \chi \in V_h$. Essentially the property **(b)** above can be restated as

$$\hat{U} = \mathcal{R}u_h \quad \text{implies} \quad P_1 \hat{U} = u_h.$$

Furthermore (4.7) is indeed an *a posteriori dual* to the classical a priori estimate for semidiscrete linear parabolic problems, [27],

$$\|(u_h - P_1 u)(t)\|_H^2 + \int_0^t \|u_h - P_1 u\|_V^2 ds \leq \|u_h(0) - P_1 u(0)\|_H^2 + \int_0^t \|u_t - P_1 u_t\|_{V^*}^2 ds. \quad (4.9)$$

4.4. A posteriori control of $\|u_h - \hat{U}\|$ via elliptic a posteriori estimators

The last piece that remains to close the estimate (4.8), as Section 2 suggests, is to estimate a posteriori the terms $\|u_h(t) - \mathcal{R}u_h(t)\|_H$ and $\|u_{h,t} - \mathcal{R}u_{h,t}\|_{V^*}$. We emphasize again that we do not intend to compute the difference $u_h - \mathcal{R}u_h$ but rather to estimate it. To this end, we assume that we have at our disposal *elliptic a posteriori estimators*: Let w and w_h be the solution of an elliptic problem and its finite element approximation:

$$\begin{aligned} w \in V : \quad a(w, v) &= \langle g, v \rangle \quad \text{for all } v \in V, \\ w_h \in V_h : \quad a(w_h, \chi) &= \langle g, \chi \rangle \quad \text{for all } \chi \in V_h. \end{aligned}$$

We assume that there exists an a posteriori estimator function $\mathcal{E} = \mathcal{E}(w_h, g; X)$, which depends on w_h, g and the space $X = H, V$, or V^* such that

$$\|w - w_h\|_X \leq \mathcal{E}(w_h, g; X). \quad (4.10)$$

There are many different choices for optimal order estimators $\mathcal{E}(w_h, g; X)$ provided by the a posteriori theory of elliptic problems, [2, 28]. Then the property **(b)** of the elliptic reconstruction implies,

$$\|u_h(t) - \mathcal{R}u_h(t)\|_H \leq \mathcal{E}(u_h, g_h; H) \quad \text{and} \quad \|u_{h,t} - \mathcal{R}u_{h,t}\|_{V^*} \leq \mathcal{E}(u_{h,t}, g_{h,t}; V^*). \quad (4.11)$$

These bounds are optimal and therefore the final a posteriori estimate for $u - u_h$ in $L^\infty(L^2)$ is optimal. In addition it parallels the classical a priori results, [27]. We refer to [22] for more details and to particular application of the above estimators when residual-type elliptic estimators are considered. In that case the final estimate is expressed in terms of inner residuals $r|_K = u_{h,t} - \Delta u_h - f$, $K \in \mathcal{T}_h$, and their time derivatives.

4.5. Fully discrete schemes

Next we consider a fully-discrete scheme. We use the notation for backward Euler time discretization of Section 2. We would like to allow the finite element spaces to be variable with n . In fact we associate to each time node t_n a finite element space V_h^n which is defined in the same manner as V_h . Then the fully discrete scheme is defined by requiring $U^j \in V_h^j$ to satisfy

$$\frac{1}{k_n} \langle (U^{n+1} - U^n), \phi \rangle + a(U^{n+1}, \phi) = \langle f^{n+1}, \phi \rangle, \quad \forall \phi \in V_h^{n+1}, \quad (4.12)$$

where U^0 is a given approximation of u^0 on an initial space V_h^0 .

Space-time reconstruction. The space-time reconstruction is now defined by combining the corresponding ideas in Sections 3.1 and 4.2. In fact we define the *space-time reconstruction* \hat{U}_{ST} of U as the function, [20],

$$\hat{U}_{ST}(t) = \ell_0^n(t) \mathcal{R}^n U^n + \ell_1^n(t) \mathcal{R}^{n+1} U^{n+1} \quad \text{for } t \in I_n. \quad (4.13)$$

Here $\mathcal{R}^j : V_h^j \rightarrow V$, are the elliptic reconstruction operators defined by (4.4), (4.5), with V_h^j, A_h^j, U^j replacing V_h, A_h, u_h respectively. Notice that \hat{U} is a Lipschitz continuous function of time.

The main parabolic error equation. It is interesting to write the equation for $\hat{e} = u - \hat{U} = u - \hat{U}_{ST}$. To this end, it will be useful to introduce the time reconstruction only: $\hat{U}_T(t) = \ell_0^n(t) U^n + \ell_1^n(t) U^{n+1}$. Then for each n , and for each $v \in H_0^1$, there holds, [20],

$$\begin{aligned} \langle \partial_t \hat{e}, v \rangle + a(\hat{e}, v) &= \langle \partial_t (\hat{U}_T - \hat{U}_{ST}), v \rangle - a(\hat{U}_{ST} - \hat{U}_{ST}(t^{n+1}), v) \\ &\quad - \langle P^{n+1} f^{n+1} - f, v \rangle - k_n^{-1} \langle P^{n+1} U^n - U^n, v \rangle \quad \text{on } I_n. \end{aligned} \quad (4.14)$$

Some remarks on the structure of this error equation follow. The last term of (4.14) reflects the effect of the mesh change with n . The third term in the right hand side has this form, partly since in [20] we use a modification of the definition of the elliptic reconstruction without the term $-f_h + f$ in (4.5) (of course then all the crucial properties of \mathcal{R}^j are valid). The first term in the rhs of (4.14) is a term that reflects the spatial error only. The similarity with $-R = -(\hat{U} - u_h)_t = -\mathcal{R}u_{h,t} - u_{h,t}$ the rhs of the semidiscrete error equation is entirely evident in the case we do not change the finite element spaces with time. Then, obviously, $\mathcal{R}^j = \mathcal{R}$ and $\partial_t(\hat{U}_T - \hat{U}_{ST}) = (I - \mathcal{R})\partial_t \hat{U}_T$. Finally, the second term in the rhs of (4.14) is a term that reflects the time discretization error only—compare it with (3.4). Of course this error equation is not that easy to handle as the corresponding equations we had when we discretize only in time or in space. Nevertheless, the analysis in [20] provides meaningful estimators that are easy to compute, are of optimal order in both $L^\infty(L^2(\Omega))$ and $L^2(H_0^1(\Omega))$, and take into account the mesh change (coarsening errors) in a natural way. In addition optimal a posteriori bounds in the *strong* norm $L^\infty(H_0^1(\Omega))$, are shown—the standard approach requires severe *restrictions in the variation of the mesh from n to $n+1$* , cf. [20] for details.

4.6. Remarks-extensions

Operators that are related or are similar to reconstructions have been used in different contexts, e.g., in [16, 17, 30]. In [30] high-order polynomials in time that interpolate the values of a time discretization scheme were used as a tool to computationally estimate the error. Similar to the elliptic reconstruction are the operators defined in [17], as a tool in the a priori analysis of Navier-Stokes equations, and in [16], to post-process the semidiscrete finite element approximation at a

fixed time T^* . The ideas in [16] were useful in estimating the error of the p -version of the finite element method in [11] and in [12] for nonlinear semidiscrete schemes.

It is interesting to note the successful application of the ideas of Section 2 in the case of time dependent Stokes equations, [18, 19]. The problem with Stokes (and therefore with Navier-Stokes) is twofold. At one hand the spatially discrete approximations are rarely divergence-free functions, and therefore the direct approach will give rise to a Stokes problem with a non-homogeneous divergence condition. This technical difficulty creates severe restrictions in the analysis. An appropriately defined *Stokes reconstruction* addresses this issue, since simply this reconstruction is always *divergence free* and thus standard PDE estimates are applicable, [19]. On the other hand, time discretization of Stokes and Navier-Stokes is not an easy task. Backward Euler is too dissipative, so Crank-Nicolson and especially the θ -*fractional step* methods are more appropriate. The θ -*fractional step* is a particular method, not close to Galerkin type methods, and still appropriate reconstructions lead to meaningful and optimal a posteriori bounds, [18].

Of course there are several open questions associated in particular to fully discrete schemes. E.g., it will be interesting to have optimal a posteriori estimators for general high-order fully discrete schemes which allow for mesh change with time. The error control of important nonlinear PDEs via conditional estimates is probably a hard task, but leads to very interesting problems.

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)
2. M. Ainsworth and J. T. Oden, *A posteriori error estimation in finite element analysis*. Wiley-Interscience, New York, 2000.
3. G. Akrivis, Ch. Makridakis and R. H. Nochetto, *A posteriori error estimates for the Crank-Nicolson method for parabolic equations*, Math. Comp. **75** (2006) 511-531.
4. G. Akrivis, Ch. Makridakis and R. H. Nochetto, *Optimal a posteriori error estimates for continuous Galerkin and Runge-Kutta methods*, Preprint 2006.
5. Babuška, M. Feistauer, and P. Šolin, *On one approach to a posteriori error estimates for evolution problems solved by the method of lines*. Numer. Math., **89** (2001) 225-256.
6. A. Bergam, C. Bernardi and Z. Mghazli, *A posteriori analysis of the finite element discretization of some parabolic equations*, Math. Comp. **74** (2005) 1117-1138.
7. Z. Chen and J. Feng, *An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems*. Math. Comp., **73** (2004) 1167-1193
8. M. Crouzeix, *Sur l'approximation des équations différentielles opérationnelles linéaires par des méthodes de Runge-Kutta*, Thèse, Université de Paris VI, 1975.
9. E. Cuesta and Ch. Makridakis, *A posteriori error estimates and maximal regularity for approximations of fully nonlinear parabolic equations in Banach spaces*, Preprint 2006.
10. F. Fierro and A. Veiser, *On the a posteriori error analysis for equations of prescribed mean curvature* Math. Comp., **72**(2003), 1611-1634.
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**(2002), 4893-4904.
12. J. de Frutos and J. Novo, *Element-wise a posteriori estimates based on hierarchical bases for non-linear parabolic problems*. Internat. J. Numer. Methods Engrg. **63** (2005), 1146-1173.
13. K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems. I. A linear model problem*, SIAM J. Numer. Anal. **28** (1991) 43-77.
14. K. Eriksson, C. Johnson and S. Larsson, *Adaptive finite element methods for parabolic problems. VI. Analytic semigroups*, SIAM J. Numer. Anal. **35** (1998) 1315-1325.
15. K. Eriksson, C. Johnson and A. Logg, *Adaptive computational methods for parabolic problems*, Encyclopedia of Computational Mechanics, Edited by E. Stein, R. de Borst and T.J.R. Houghes, 1-44, J. Wiley & Sons, Ltd., New York, 2004.

16. B. Garca-Archilla and E. S. Titi, *Postprocessing the Galerkin method: the finite-element case*. SIAM J. Numer. Anal., **37** (2000), 470-499.
17. 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**(1986), 750-777.
18. F. Karakatsani, *A posteriori error control for the linearized Navier-Stokes equations*. PhD Thesis, University of Crete 2006.
19. F. Karakatsani and Ch. Makridakis, *A posteriori estimates for approximations of the time-dependent Stokes equations*. Advance Electronic Access published on December 2006, IMA J. Numer. Anal. (2006).
20. O. Lakkis and Ch. Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp. **75** (2006), 1627-1658.
21. O. Lakkis and R. H. Nochetto, *A posteriori error analysis for the mean curvature flow of graphs*, SIAM J. Numer. Anal. **42** (2005) 1875-1898.
22. Ch. Makridakis and R. H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal. **41** (2003) 1585-1594.
23. Ch. Makridakis and R. H. Nochetto, *A posteriori error analysis for higher order dissipative methods for evolution problems*. Numer. Math. **104** (2006) 489-514.
24. 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** (2000) 525-589.
25. R. H. Nochetto, A. Schmidt, and C. Verdi, *A posteriori error estimation and adaptivity for degenerate parabolic problems*, Math. Comp. **69** (2000) 1-24.
26. M. Picasso, *Adaptive finite elements for a linear parabolic problem*. Comput. Methods Appl. Mech. Engrg., **167**(1998), 223-237.
27. V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*. 2nd ed., Springer-Verlag, Berlin, 2006.
28. R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Wiley-Teubner, Chichester-Stuttgart, 1996.
29. R. Verfürth, *A posteriori error estimates for nonlinear problems. -error estimates for finite element discretizations of parabolic equations*. Math. Comp., **67**(1998), 1335-1360.
30. P. E. Zadunaisky, *On the estimation of errors propagated in the numerical integration of ordinary differential equations*, Numer. Math. **27** (1976) 21-39.