A POSTERIORI ERROR ESTIMATES
FOR THE CRANK–NICOLSON METHOD
FOR PARABOLIC EQUATIONS

GEORGIOS AKRIVIS, CHARALAMBOS MAKRIDAKIS, AND RICARDO H. NOCHETTO

ABSTRACT. We derive optimal order a posteriori error estimates for time discretizations by both the Crank–Nicolson and the Crank–Nicolson–Galerkin methods for linear and nonlinear parabolic equations. We examine both smooth and rough initial data. Our basic tool for deriving a posteriori estimates are second-order Crank–Nicolson reconstructions of the piecewise linear approximate solutions. These functions satisfy two fundamental properties: (i) they are explicitly computable and thus their difference to the numerical solution is controlled a posteriori, and (ii) they lead to optimal order residuals as well as to appropriate pointwise representations of the error equation of the same form as the underlying evolution equation. The resulting estimators are shown to be of optimal order by deriving upper and lower bounds for them depending only on the discretization parameters and the data of our problem. As a consequence we provide alternative proofs for known a priori rates of convergence for the Crank–Nicolson method.

1. INTRODUCTION

In this paper we derive a posteriori error estimates for time discretizations by Crank–Nicolson type methods for parabolic partial differential equations (p.d.e.'s). The Crank–Nicolson scheme is one of the most popular time-stepping methods; however, optimal order a posteriori estimates for it have not yet been derived. Most of the (many) contributions in the last years devoted to a posteriori error control for time dependent equations concern the discretization in time of linear or nonlinear equations with dissipative character by the backward Euler method or by higher order discontinuous Galerkin methods; cf., e.g., [12], [7], [8], [20], [9], [19] and [13]. Let u and U be the exact and the numerical solution of a given problem. In a posteriori error analysis

$$\|u - U\| \le \eta(U),$$

we seek computable *estimators* $\eta(U)$ depending on the approximate solution U and the data of the problem such that (i) $\eta(U)$ decreases with optimal order for the lowest possible regularity permitted by our problem, and (ii) the constants involved in the estimator $\eta(U)$ are explicit and easily computable.

In this paper we derive optimal order estimators of various types for the Crank–Nicolson and the Crank–Nicolson–Galerkin time-stepping methods applied to evolution problems of the form: Find $u : [0, T] \to D(A)$ satisfying

$$(1.1) \qquad \begin{cases} u'(t) + Au(t) = B(t, u(t)), & 0 < t < T, \\ u(0) = u^0, \end{cases}$$

with $A : D(A) \to H$ a positive definite, selfadjoint, linear operator on a Hilbert space $(H, (\cdot, \cdot))$ with domain $D(A)$ dense in H, $B(t, \cdot) : D(A) \to H, t \in [0, T]$, a (possibly) nonlinear operator, and $u^0 \in H$. The structural assumption (6.1) on $B(t, \cdot)$ implies that problem (1.1) is parabolic.

A main novel feature of our approach is the *Crank–Nicolson reconstruction* $\hat{U}$ of the numerical approximation U. This function satisfies two fundamental properties: (i) it is explicitly computable and thus its difference to the numerical solution is controlled a posteriori, and (ii) it leads to an appropriate pointwise representation of the error equation, of the same form as the original evolution equation. Then by employing techniques developed for the underlying p.d.e. we conclude the final estimates. Of course, depending on the stability methods that are used, we obtain different estimators. The resulting estimators are shown to be of optimal order by deriving upper bounds for them, depending only on the discretization parameters and the data of our problem. As a consequence we provide alternative proofs for *a priori* estimates, depending only on the data and corresponding known rates of convergence for the Crank–Nicolson method.

The above idea is related to earlier work on a posteriori analysis of time or space discrete approximations of evolution problems [19, 18, 17]. It provides the means to show optimal order error estimates with energy as well as with other stability techniques. An alternative approach for a posteriori analysis of time dependent problems, based on the direct comparison of u and U via parabolic duality, was considered in [12], [7], [20], [9] for p.d.e.'s and in [11], [10] for ordinary differential equations (o.d.e.'s). In particular Estep and French [10] considered the continuous Galerkin method for o.d.e's. Its lowest order representative corresponds to a variant of the Crank–Nicolson method—the Crank–Nicolson–Galerkin method—considered also in this paper. A posteriori bounds with energy techniques for Crank–Nicolson methods for the linear Schrödinger equation were proved by Dörfler [6] and for the heat equation by Verfürth [22]; the upper bounds in [6], [22] are of suboptimal order.

Most of this paper is devoted to linear parabolic equations, namely $B(t, u(t)) = f(t)$ for a given forcing function f. The general nonlinear problem (1.1) is only briefly discussed in the last section. The paper is organized as follows. We start in Section 2 by introducing the necessary notation, the Crank–Nicolson and the Crank–Nicolson–Galerkin (CNG) methods for the linear problem (2.1). We then observe that the direct use of standard piecewise linear interpolation at the approximate nodal values (see (2.3)), would lead to suboptimal estimates as in [6] and [22]. The Crank–Nicolson and Crank–Nicolson–Galerkin reconstructions $\hat{U}$ are, instead,

continuous piecewise quadratic functions which are defined in (2.9) and (2.22), respectively. In Section 3 we estimate $\hat{U} - U$. Section 4 is devoted to the a posteriori error analysis for linear equations. Error estimates are obtained by using energy techniques, as well as Duhamel's principle. Both estimators lead to second order convergence rates. Note the interesting similarity of the estimator obtained by Duhamel's principle to those established in the literature by parabolic duality. In Section 5 we discuss the form of estimators in the case of nonsmooth initial data. In Section 6 we finally conclude with the case of nonlinear equations.

2. Crank–Nicolson methods for linear equations

Most of this paper focuses on the case of a linear equation,

(2.1)
$$\begin{cases} u'(t) + Au(t) = f(t), & 0 < t < T, \\ u(0) = u^0, \end{cases}$$

with $f : [0, T] \to H$. Let $0 = t^0 < t^1 < \cdots < t^N = T$ be a partition of $[0, T]$, $I_n := (t^{n-1}, t^n]$, and $k_n := t^n - t^{n-1}$.

2.1. **The Crank–Nicolson method.** For given $\{v^n\}_{n=0}^N$ we will use the notation

$$\bar{\partial} v^n := \frac{v^n - v^{n-1}}{k_n}, \quad v^{n-\frac{1}{2}} := \frac{1}{2}(v^n + v^{n-1}), \quad n = 1, \ldots, N.$$

The Crank–Nicolson nodal approximations $U^m \in D(A)$ to the values $u^m := u(t^m)$ of the solution u of (2.1) are defined by

(2.2)
$$\bar{\partial} U^n + AU^{n-\frac{1}{2}} = f(t^{n-\frac{1}{2}}), \quad n = 1, \ldots, N,$$

with $U^0 := u^0$. Since the error $u^m - U^m$ is of second order, to obtain a second–order approximation $U(t)$ to $u(t)$, for all $t \in [0, T]$, we define the Crank–Nicolson approximation $U : [0, T] \to D(A)$ to u by linearly interpolating between the nodal values U^{n-1} and U^n,

(2.3)
$$U(t) = U^{n-\frac{1}{2}} + (t - t^{n-\frac{1}{2}})\bar{\partial} U^n, \quad t \in I_n.$$

Let $R(t) \in H$,

(2.4)
$$R(t) := U'(t) + AU(t) - f(t), \quad t \in I_n,$$

denote the *residual* of U, i.e., the amount by which the approximate solution U misses being an exact solution of (2.1). Now

$$U'(t) + AU(t) = \bar{\partial} U^n + AU^{n-\frac{1}{2}} + (t - t^{n-\frac{1}{2}})A\bar{\partial} U^n, \quad t \in I_n,$$

whence, in view of (2.2),

$$U'(t) + AU(t) = f(t^{n-\frac{1}{2}}) + (t - t^{n-\frac{1}{2}})A\bar{\partial} U^n, \quad t \in I_n.$$

Therefore, the residual can also be written in the form

(2.5)
$$R(t) = (t - t^{n-\frac{1}{2}})A\bar{\partial} U^n + [f(t^{n-\frac{1}{2}}) - f(t)], \quad t \in I_n.$$

Obviously, $R(t)$ is an a posteriori quantity of first order, even in the case of a scalar o.d.e. $u'(t) = f(t)$, although the Crank–Nicolson scheme yields second-order accuracy. Since the error $e := u - U$ satisfies $e' + Ae = -R$, applying energy techniques to this error equation, as in [6], [22], leads inevitably to suboptimal bounds.

2.2. **Crank–Nicolson reconstruction.** To recover the optimal order we introduce a *Crank–Nicolson reconstruction* $\hat{U}$ of U, namely a continuous piecewise quadratic polynomial in time $\hat{U} : [0, T] \to H$ defined as follows. First, we let $\varphi : I_n \to H$ be the linear interpolant of f at the nodes t^{n-1} and $t^{n-\frac{1}{2}}$,

$$(2.6) \qquad \varphi(t) := f(t^{n-\frac{1}{2}}) + \frac{2}{k_n}(t - t^{n-\frac{1}{2}})[f(t^{n-\frac{1}{2}}) - f(t^{n-1})], \quad t \in I_n,$$

and define a piecewise quadratic polynomial Φ by $\Phi(t) := \int_{t^{n-1}}^t \varphi(s)ds, t \in I_n$, i.e.,

$$(2.7) \qquad \Phi(t) = (t - t^{n-1})f(t^{n-\frac{1}{2}}) - \frac{1}{k_n}(t - t^{n-1})(t^n - t)[f(t^{n-\frac{1}{2}}) - f(t^{n-1})].$$

As will become evident in the sequel, an important property of Φ is that

$$(2.8) \qquad \Phi(t^{n-1}) = 0, \quad \Phi(t^n) = k_n f(t^{n-\frac{1}{2}}) = \int_{I_n} f(t^{n-\frac{1}{2}})dt.$$

We now introduce the *Crank–Nicolson reconstruction* $\hat{U}$ of U by

$$(2.9) \qquad \hat{U}(t) := U^{n-1} - \int_{t^{n-1}}^t AU(s)\,ds + \Phi(t) \quad \forall t \in I_n.$$

We can interpret this formula as the result of formally replacing the constants $U^{n-\frac{1}{2}}$ and $f(t^{n-\frac{1}{2}})$ in (2.2) by their piecewise linear counterparts U and φ, and next integrating $-AU + \varphi$ from t^{n-1} to t. Consequently

$$\hat{U}'(t) + AU(t) = \varphi(t) \quad \forall t \in I_n.$$

Evaluating the integral in (2.9) by the trapezoidal rule, we obtain

$$(2.10) \qquad \hat{U}(t) = U^{n-1} - \frac{1}{2}(t - t^{n-1})A[U(t) + U^{n-1}] + \Phi(t) \quad \forall t \in I_n,$$

which can also be written as

$$\hat{U}(t) = U^{n-1} - A[(t - t^{n-1})U^{n-1} + \frac{1}{2}(t - t^{n-1})^2 \bar{\partial}U^n] + \Phi(t) \quad \forall t \in I_n.$$

Obviously $\hat{U}(t^{n-1}) = U^{n-1}$. Furthermore, in view of (2.8) and (2.2), we have

$$\hat{U}(t^n) = U^{n-1} - k_n A U^{n-\frac{1}{2}} + \Phi(t^n)$$
$$= U^{n-1} + k_n[-AU^{n-\frac{1}{2}} + f(t^{n-\frac{1}{2}})] = U^{n-1} + k_n \bar{\partial}U^n = U^n.$$

Thus, $\hat{U}$ and U coincide at the nodes $t^0, \dots, t^N$; in particular, $\hat{U} : [0, T] \to H$ is continuous.

Remark 2.1 (Choice of φ). Let $\tilde{t} \in I_n$. Since $f(t) = f(\tilde{t}) + f'(\tilde{t})(t - \tilde{t}) + O(k_n^2)$, $t \in I_n$, it is easily seen that the only affine functions φ satisfying

$$\sup_{t \in I_n} |f(t) - \varphi(t)| = O(k_n^2)$$

are the ones of the form

$$\varphi(t) = f(\tilde{t}) + \big[f'(\tilde{t}) + O(k_n)\big](t - \tilde{t}) + O(k_n^2).$$

Obviously

$$\int_{I_n} \varphi(t)dt = k_n f(\tilde{t}) + \frac{1}{2}\big[f'(\tilde{t}) + O(k_n)\big]\big\{(t^n - \tilde{t})^2 - (t^{n-1} - \tilde{t})^2\big\} + O(k_n^3);$$

therefore, for $f'(\tilde{t}) \neq 0$, the requirement $\int_{I_n} \varphi(t)dt = k_n f(\tilde{t})$ leads to $\tilde{t} = t^{n-\frac{1}{2}}$ and

$$\varphi(t) = f(t^{n-\frac{1}{2}}) + \left[f'(t^{n-\frac{1}{2}}) + O(k_n)\right](t - t^{n-\frac{1}{2}}).$$

These comments demonstrate the special features of the midpoint method among the one-stage Runge–Kutta methods.

Furthermore, our choice (2.6) is motivated by the fact that for all affine functions φ on I_n we have $\int_{I_n} \varphi(s)ds = k_n\varphi(t^{n-\frac{1}{2}})$, whence the requirement $\int_{I_n} \varphi(s)ds = k_n f(t^{n-\frac{1}{2}})$ (see (2.8)), is satisfied if and only if φ interpolates f at $t^{n-\frac{1}{2}}$. Now, to ensure that $\varphi(t)$ is a second order approximation to $f(t)$, we let φ interpolate f at an additional point $t^{n,\star} \in [t^{n-1}, t^n]$; of course, in the case $t^{n,\star} = t^{n-\frac{1}{2}}$, φ is the affine Taylor polynomial of f around $t^{n-\frac{1}{2}}$. In the sequel we let $t^{n,\star} := t^{n-1}$.   □

In view of (2.9) and (2.6), we have

$$(2.11) \quad \hat{U}'(t) + AU(t) = f(t^{n-\frac{1}{2}}) + \frac{2}{k_n}(t - t^{n-\frac{1}{2}})[f(t^{n-\frac{1}{2}}) - f(t^{n-1})], \quad t \in I_n;$$

therefore, the *residual* $\hat{R}(t)$ of $\hat{U}$,

$$(2.12) \qquad\qquad \hat{R}(t) := \hat{U}'(t) + A\hat{U}(t) - f(t), \quad t \in I_n,$$

can be written in the form

$$(2.13) \quad \begin{aligned} \hat{R}(t) &= A[\hat{U}(t) - U(t)] \\ &+ \left\{ f(t^{n-\frac{1}{2}}) + \frac{2}{k_n}(t - t^{n-\frac{1}{2}})[f(t^{n-\frac{1}{2}}) - f(t^{n-1})] - f(t) \right\}, \quad t \in I_n. \end{aligned}$$

We will see later that the a posteriori quantity $\hat{R}(t)$ is of second order; compare with (2.5).

2.3. The Crank–Nicolson–Galerkin method.

Next we consider the discretization of (2.1) by the Crank–Nicolson–Galerkin method. The Crank–Nicolson–Galerkin approximation to u is defined as follows: We seek $U : [0, T] \to D(A)$, continuous and piecewise linear in time, which interpolates the values $\{U^n\}_{n=0}^N$ given by

$$(2.14) \qquad\qquad \bar{\partial}U^n + AU^{n-\frac{1}{2}} = \frac{1}{k_n} \int_{I_n} f(t)dt, \quad n = 1, \dots, N,$$

with $U^0 = u^0$. This function U can be expressed in terms of its nodal values,

$$(2.15) \qquad\qquad U(t) = U^{n-\frac{1}{2}} + (t - t^{n-\frac{1}{2}})\bar{\partial}U^n, \quad t \in I_n.$$

For $t \in I_n$, $U'(t) = \bar{\partial}U^n$, and (2.14) takes the form

$$(2.16) \qquad\qquad U'(t) + AU^{n-\frac{1}{2}} = \frac{1}{k_n} \int_{I_n} f(t)dt, \quad n = 1, \dots, N.$$

Now, $\frac{1}{k_n} \int_{I_n} \psi(t)dt$ is the L^2 orthogonal projection of a function ψ onto the space of constant functions on I_n, and $\int_{I_n} U(t)dt = k_n U^{n-\frac{1}{2}}$; therefore (2.16) yields the pointwise equation for the Crank–Nicolson–Galerkin approximation

$$(2.17) \qquad\qquad U'(t) + P_0 AU(t) = P_0 f(t) \quad \forall t \in I_n,$$

with P_0 denoting the L^2 orthogonal projection operator onto the space of constant functions in I_n. Equivalently, as it is customary, this method can be seen as a finite element in time method, [4],

$$(2.18) \qquad \int_{I_n} \left[(U', v) + (AU, v)\right] dt = \int_{I_n} (f, v) dt \quad \forall v \in D(A).$$

For a priori results for general Continuous Galerkin methods for various evolution p.d.e.'s cf. [3, 4, 14, 15].

Let $R(t)$,

$$(2.19) \qquad R(t) := U'(t) + AU(t) - f(t),$$

denote the *residual* of U. In view of (2.17), the residual can also be written in the form

$$(2.20) \qquad R(t) = A[U(t) - P_0U(t)] - [f(t) - P_0f(t)].$$

However, this residual is not appropriate for our purposes, since, even in the case of an o.d.e. $u'(t) = f(t)$, $R(t)$ can only be of first order, although our approximations are piecewise polynomials of degree one.

2.4. Crank–Nicolson–Galerkin reconstruction.

To recover the optimal order $O(k_n^2)$, we introduce the *Crank–Nicolson–Galerkin reconstruction* $\hat{U}$ of the approximate solution U, namely, the continuous and piecewise quadratic function $\hat{U} : [0, T] \to H$ defined by

$$(2.21) \qquad \hat{U}(t) := U^{n-1} - \int_{t^{n-1}}^{t} [AU(s) - P_1f(s)] \, ds \quad \forall t \in I_n.$$

Hence

$$(2.22) \quad \hat{U}(t) = U^{n-1} - \frac{1}{2}(t - t^{n-1})A[U(t) + U^{n-1}] + \int_{t^{n-1}}^{t} P_1f(s) ds, \quad t \in I_n,$$

with P_1 denoting the L^2 orthogonal projection operator onto the space of linear polynomials in I_n; that $\hat{U}(t)$ is continuous, namely, $\hat{U}(t^n) = U^n$, is a consequence of $\int_{I_n} P_1f = \int_{I_n} f$. Obviously, $\hat{U}$ satisfies the following pointwise equation:

$$(2.23) \qquad \hat{U}'(t) + AU(t) = P_1f(t) \quad \forall t \in I_n;$$

compare with (2.17). In view of (2.23), the residual $\hat{R}(t)$,

$$(2.24) \qquad \hat{R}(t) := \hat{U}'(t) + A\hat{U}(t) - f(t),$$

of $\hat{U}$ can also be written as

$$(2.25) \qquad \hat{R}(t) = A[\hat{U}(t) - U(t)] + [P_1f(t) - f(t)], \quad t \in I_n.$$

$\hat{R}(t)$ is an a posteriori quantity and, as we will see in Section 3, it is of second order at least in some cases.

3. Estimation of $\hat{U} - U$

In this section we will estimate $\hat{U} - U$ for both the Crank–Nicolson and the Crank–Nicolson–Galerkin methods; also we will derive representations of $\hat{U} - U$ that will be useful in the sequel.

We let $V := D(A^{1/2})$ and denote the norms in H and in V by $|\cdot|$ and $\|\cdot\|$, $\|v\| := |A^{1/2}v| = (Av, v)^{1/2}$, respectively. We identify H with its dual, and let $V^\star$ be the dual of V ($V \subset H \subset V^\star$). We still denote by $(\cdot, \cdot)$ the duality pairing between $V^\star$ and V, and by $\|\cdot\|_\star$ the dual norm on $V^\star$, $\|v\|_\star := |A^{-1/2}v| = (v, A^{-1}v)^{1/2}$.

3.1. The Crank–Nicolson method. From (2.10) we obtain

$$\hat{U}(t) - U(t) = U^{n-1} - U(t) - \frac{1}{2}(t - t^{n-1})A[U(t) + U^{n-1}] + \Phi(t)$$

$$= -(t - t^{n-1})\bar{\partial}U^n - \frac{1}{2}(t - t^{n-1})A[U(t) + U^{n-1}] + \Phi(t).$$

Therefore, in view of (2.2),

$$\hat{U}(t) - U(t) = (t - t^{n-1})\big[AU^{n-\frac{1}{2}} - f(t^{n-\frac{1}{2}})\big] - \frac{1}{2}(t - t^{n-1})A[U(t) + U^{n-1}] + \Phi(t)$$

$$= -\frac{1}{2}(t - t^{n-1})A[U(t) - U^n] + \Phi(t) - (t - t^{n-1})f(t^{n-\frac{1}{2}}),$$

whence, using (2.7), for $t \in I_n$,

$$(3.1) \qquad \hat{U}(t) - U(t) = (t - t^{n-1})(t^n - t)\Big(\frac{1}{2}A\bar{\partial}U^n - \frac{1}{k_n}[f(t^{n-\frac{1}{2}}) - f(t^{n-1})]\Big),$$

from which we immediately see that $\max_{t \in I_n} |\hat{U}(t) - U(t)| = O(k_n^2)$.

3.2. The Crank–Nicolson–Galerkin method. Subtracting (2.15) from (2.22), and utilizing (2.14), for $t \in I_n$ we obtain

$$(3.2) \qquad \begin{aligned} \hat{U}(t) - U(t) = &\frac{1}{2}(t - t^{n-1})(t^n - t)A\bar{\partial}U^n \\ &- \frac{t - t^{n-1}}{k_n}\int_{I_n} f(s)ds + \int_{t^{n-1}}^t P_1 f(s)ds. \end{aligned}$$

Now, it is easily seen that

$$(3.3) \qquad (P_1 f)(t) = \frac{1}{k_n}\int_{I_n} f(s)ds + \frac{12}{k_n^3}(t - t^{n-\frac{1}{2}})\int_{I_n} f(s)(s - t^{n-\frac{1}{2}})ds,$$

and (3.2) can be rewritten in the form

$$(3.4) \qquad \hat{U}(t) - U(t) = (t - t^{n-1})(t^n - t)\Big(\frac{1}{2}A\bar{\partial}U^n - \frac{6}{k_n^3}\int_{I_n} f(s)(s - t^{n-\frac{1}{2}})ds\Big).$$

Therefore, U and $\hat{U}$ coincide at the endpoints of I_n, and, consequently, at all nodes $t^0, \ldots, t^N$. From (3.4) we immediately see that $\max_{t \in I_n} |\hat{U}(t) - U(t)| = O(k_n^2)$.

Let us write both (3.1) and (3.4) in the form

$$(3.5) \qquad \hat{U}(t) - U(t) = \frac{1}{2}(t - t^{n-1})(t^n - t)\big(A\bar{\partial}U^n - \rho_{f,n}\big);$$

here $\rho_{f,n} = \rho_{f,n}^{\mathrm{CN}}$ for the Crank–Nicolson method and $\rho_{f,n} = \rho_{f,n}^{\mathrm{CNG}}$ for the Crank–Nicolson–Galerkin method, respectively, with

$$(3.6) \qquad \rho_{f,n}^{\mathrm{CN}} := \frac{2}{k_n}\big[f(t^{n-\frac{1}{2}}) - f(t^{n-1})\big]$$

and

$$(3.7) \quad \rho_{f,n}^{\mathrm{CNG}} := \frac{12}{k_n^3} \int_{I_n} f(s)(s - t^{n-\frac{1}{2}}) ds = \frac{12}{k_n^3} \int_{I_n} \left(f(s) - f(t^{n-\frac{1}{2}})\right)(s - t^{n-\frac{1}{2}}) ds.$$

Consequently, both $\rho_{f,n}^{\mathrm{CN}}$ and $\rho_{f,n}^{\mathrm{CNG}}$ depend on the first derivative of f.

4. Smooth data error estimates

Let the errors e and $\hat{e}$ be defined by $e := u - U$ and $\hat{e} := u - \hat{U}$. Subtracting (2.11) or (2.23), respectively, from the differential equation in (2.1), we obtain

$$(4.1) \qquad\qquad\qquad \hat{e}'(t) + Ae(t) = R_f(t),$$

with $R_f = R_f^{\mathrm{CN}}$ for the Crank–Nicolson method and $R_f = R_f^{\mathrm{CNG}}$ for the Crank–Nicolson–Galerkin method, respectively, defined by

$$(4.2) \quad R_f^{\mathrm{CN}}(t) := f(t) - \left\{ f(t^{n-\frac{1}{2}}) + \frac{2}{k_n}(t - t^{n-\frac{1}{2}})[f(t^{n-\frac{1}{2}}) - f(t^{n-1})] \right\}, \quad t \in I_n,$$

and

$$(4.3) \qquad\qquad R_f^{\mathrm{CNG}}(t) := f(t) - P_1 f(t), \quad t \in I_n.$$

We make the following further regularity assumption on $\hat{U}$, defined in (2.9) and (2.21):

$$\hat{U}(t) \in V, \quad \forall t \in [0, T].$$

4.1. **Energy estimates.** Taking in (4.1) the inner product with $\hat{e}(t)$, we obtain

$$(4.4) \qquad\qquad \frac{1}{2} \frac{d}{dt} |\hat{e}(t)|^2 + \left(Ae(t), \hat{e}(t)\right) = \left(R_f(t), \hat{e}(t)\right).$$

Now,

$$\left(Ae(t), \hat{e}(t)\right) = \frac{1}{2} \left(\|e(t)\|^2 + \|\hat{e}(t)\|^2 - \|\hat{e}(t) - e(t)\|^2 \right)$$

and

$$\left(R_f(t), \hat{e}(t)\right) \le \|R_f(t)\|_\star^2 + \frac{1}{4} \|\hat{e}(t)\|^2 ;$$

therefore, (4.4) yields

$$(4.5) \qquad \frac{d}{dt} |\hat{e}(t)|^2 + \|e(t)\|^2 + \frac{1}{2} \|\hat{e}(t)\|^2 \le \|\hat{U}(t) - U(t)\|^2 + 2\|R_f(t)\|_\star^2 .$$

We recall that $\|v\| = |A^{1/2}v|$ and $\|v\|_\star = |A^{-1/2}v|$.

4.1.1. *Upper bound.* Since $\hat{e}(0) = 0$, integration of (4.5) from 0 to $t \le T$ yields

$$(4.6) \qquad \begin{aligned} |\hat{e}(t)|^2 &+ \int_0^t \left(\|e(s)\|^2 + \frac{1}{2} \|\hat{e}(s)\|^2 \right) ds \\ &\le \int_0^t \|\hat{U}(s) - U(s)\|^2 \, ds + 2 \int_0^t \|R_f(s)\|_\star^2 \, ds . \end{aligned}$$

From (4.6) we easily conclude that

$$(4.7) \qquad \begin{aligned} \max_{0 \le \tau \le t} &\left\{ |\hat{e}(\tau)|^2 + \int_0^\tau \left(\|e(s)\|^2 + \frac{1}{2} \|\hat{e}(s)\|^2 \right) ds \right\} \\ &\le \int_0^t \|\hat{U}(s) - U(s)\|^2 \, ds + 2 \int_0^t \|R_f(s)\|_\star^2 \, ds . \end{aligned}$$

Next, let β be given by

$$(4.8) \qquad \beta := \int_0^1 t^2 (1-t)^2 dt = \frac{1}{30} \, ;$$

then, obviously,

$$\int_{I_n} (t - t^{n-1})^2 (t^n - t)^2 dt = \beta \, k_n^5 \, .$$

With this notation, in view of (3.5), we have

$$(4.9) \qquad \int_0^{t^m} \|\hat{U}(t) - U(t)\|^2 \, dt \leq \frac{\beta}{2} \sum_{n=1}^m k_n^5 \left(|A^{3/2} \bar{\partial} U^n|^2 + \|\rho_{f,n}\|^2 \right).$$

4.1.2. *Lower bound.* Obviously,

$$\|\hat{U}(s) - U(s)\| \leq \|e(s)\| + \|\hat{e}(s)\|$$

and thus

$$(4.10) \qquad \|\hat{U}(s) - U(s)\|^2 \leq 3 \left(\|e(s)\|^2 + \frac{1}{2} \|\hat{e}(s)\|^2 \right).$$

In particular, combining the upper and lower bounds, we have

$$(4.11) \qquad \begin{aligned} \frac{1}{3} \int_0^t \|\hat{U}(s) - U(s)\|^2 \, ds &\leq \int_0^t \left(\|e(s)\|^2 + \frac{1}{2} \|\hat{e}(s)\|^2 \right) ds \\ &\leq \int_0^t \|\hat{U}(s) - U(s)\|^2 \, ds + 2 \int_0^t \|R_f(s)\|_\star^2 \, ds \, . \end{aligned}$$

Invoking (4.6), this shows that $|\hat{e}(t)|^2$ is dominated by $\int_0^t \left(\|e(s)\|^2 + \frac{1}{2} \|\hat{e}(s)\|^2 \right) ds$, the energy norm error, plus data oscillation $\int_0^t \|R_f(s)\|_\star^2 ds$. We will next estimate the lower bounds above in terms of U^n and data in analogy to the upper bound in (4.9). To this end we first note that (3.5) yields

$$\|\hat{U}(t) - U(t)\|^2 \geq \frac{1}{4} (t - t^{n-1})^2 (t^n - t)^2 \left(\frac{1}{2} |A^{3/2} \bar{\partial} U^n|^2 - \|\rho_{f,n}\|^2 \right),$$

whence, in view of (4.8), we have

$$(4.12) \qquad \int_0^{t^m} \|\hat{U}(s) - U(s)\|^2 \, ds \geq \frac{\beta}{8} \sum_{n=1}^m k_n^5 |A^{3/2} \bar{\partial} U^n|^2 - \frac{\beta}{4} \sum_{n=1}^m k_n^5 \|\rho_{f,n}\|^2 \, .$$

Therefore, (4.11), (4.9) and (4.12) imply

$$(4.13) \qquad \begin{aligned} \frac{\beta}{24} \sum_{n=1}^m & k_n^5 |A^{3/2} \bar{\partial} U^n|^2 - \frac{\beta}{12} \sum_{n=1}^m k_n^5 \|\rho_{f,n}\|^2 \\ &\leq \int_0^{t^m} \left(\|e(s)\|^2 + \frac{1}{2} \|\hat{e}(s)\|^2 \right) ds \\ &\leq \frac{\beta}{2} \sum_{n=1}^m k_n^5 \left(|A^{3/2} \bar{\partial} U^n|^2 + \|\rho_{f,n}\|^2 \right) + 2 \int_0^{t^m} \|R_f(s)\|_\star^2 \, ds \, . \end{aligned}$$

Note that error bounds of this type are customary in the a posteriori analysis of elliptic problems in which data oscillation appear with different signs in the upper

and lower bounds. On the other hand, if f is constant, then the lower and upper bounds are exactly the same up to a constant:

(4.14)
$$\frac{\beta}{24} \sum_{n=1}^{m} k_n^5 \, |A^{3/2} \bar{\partial} U^n|^2 \leq \int_0^{t^m} \left(\|e(s)\|^2 + \frac{1}{2} \|\hat{e}(s)\|^2 \right) ds$$
$$\leq \frac{\beta}{2} \sum_{n=1}^{m} k_n^5 \, |A^{3/2} \bar{\partial} U^n|^2 \, .$$

Let us also note that, in the case of f constant, in view of (3.5), the estimate (4.11), for $t = t^m$, can be written in the form of (4.14) with the lower bound multiplied by 2 and the upper bound by $1/2$.

Remark 4.1 (Optimality of the lower bound). The pointwise lower bound

(4.15)
$$|(\hat{U} - U)(s)| \leq |e(s)| + |\hat{e}(s)|, \quad s \in [0, T],$$

cannot be expected to be of exactly second order for all s, since $\hat{U} - U$ vanishes at the nodes $t^0, \ldots, t^N$. However, we can conclude from (4.15) the following lower bound in the $\| \cdot \|_{L^\infty(H)}$-norm, with $\|v\|_{L^\infty(H)} := \max_{t \in [0,T]} |v(t)|$:

(4.16)
$$\|\hat{U} - U\|_{L^\infty(H)} \leq \|e\|_{L^\infty(H)} + \|\hat{e}\|_{L^\infty(H)}.$$

For the trivial case that the exact solution u is an affine function, and so is f, then both the Crank–Nicolson approximation U and the Crank–Nicolson reconstruction $\hat{U}$ coincide with u; thus (4.16) is an equality. Next consider the case of u nonaffine. In view of (3.1), we have

$$|(\hat{U} - U)(t^{n-\frac{1}{2}})| = \frac{k_n^2}{4} \left| \frac{1}{2} A \bar{\partial} U^n - \frac{1}{k_n} \left[f(t^{n-\frac{1}{2}}) - f(t^{n-1}) \right] \right|.$$

Now, for smooth data, we have, as $k_n \to 0$,

$$\frac{1}{2} A \bar{\partial} U^n - \frac{1}{k_n} \left[f(t^{n-\frac{1}{2}}) - f(t^{n-1}) \right] \to \frac{1}{2} \left[A u'(t^{n-1}) - f'(t^{n-1}) \right] = -\frac{1}{2} u''(t^{n-1}).$$

If $u''(t^{n-1}) \neq 0$, we then have $\|\hat{U} - U\|_{L^\infty(H;I_n)} \geq c k_n^2$ with a positive constant c. This is the generic situation.

That the lower bound in the $L^2(V)$-norm is of the same form as the upper bound, in the case of f constant, can be seen from (4.14). Otherwise, let us note that, in view of (3.1) and (4.8),

$$\|\hat{U} - U\|_{L^2(V)}^2 = \beta \sum_{n=1}^{N} k_n^5 \left\| \frac{1}{2} A \bar{\partial} U^n - \frac{1}{k_n} \left[f(t^{n-\frac{1}{2}}) - f(t^{n-1}) \right] \right\|^2$$

and, assuming that the partition is quasi-uniform and letting $k := \max_n k_n$,

$$\|\hat{U} - U\|_{L^2(V)}^2 \geq c \beta k^4 \sum_{n=1}^{N} k_n \left\| \frac{1}{2} A \bar{\partial} U^n - \frac{1}{k_n} \left[f(t^{n-\frac{1}{2}}) - f(t^{n-1}) \right] \right\|^2.$$

Now, as $k \to 0$,

$$\sum_{n=1}^{N} k_n \left\| \frac{1}{2} A \bar{\partial} U^n - \frac{1}{k_n} \left[f(t^{n-\frac{1}{2}}) - f(t^{n-1}) \right] \right\|^2 \to \left\| \frac{1}{2} u'' \right\|_{L^2(V)}^2,$$

whence, if u is not affine,

(4.17)
$$\|\hat{U} - U\|_{L^2(V)} \geq C k^2. \quad \square$$

Remark 4.2 (Alternative estimate in $L^\infty(H)$). Combining (4.10) and (4.6) we can replace the factor 1 on the right-hand side of (4.7) by $\frac{2}{3}$, if we only want to estimate the first term on the left-hand side of (4.7). Indeed, in view of (4.10), from (4.6) we obtain

$$|\hat{e}(t)|^2 + \frac{1}{3}\int_0^t \|\hat{U}(s) - U(s)\|^2\,ds$$

$$\leq |\hat{e}(t)|^2 + \int_0^t \left(\|e(s)\|^2 + \frac{1}{2}\|\hat{e}(s)\|^2\right) ds$$

$$\leq \int_0^t \|\hat{U}(s) - U(s)\|^2\,ds + 2\int_0^t \|R_f(s)\|_\star^2\,ds\,,$$

i.e.,

$$|\hat{e}(t)|^2 \leq \frac{2}{3}\int_0^t \|\hat{U}(s) - U(s)\|^2\,ds + 2\int_0^t \|R_f(s)\|_\star^2\,ds\,,$$

and thus

$$(4.18) \qquad \max_{0\leq\tau\leq t}|\hat{e}(\tau)|^2 \leq \frac{2}{3}\int_0^t \|\hat{U}(s) - U(s)\|^2\,ds + 2\int_0^t \|R_f(s)\|_\star^2\,ds\,. \qquad \square$$

The following *stability* estimate for the Crank–Nicolson scheme will be useful in the convergence proof:

Lemma 4.1 (Stability). *Let $\{U^n\}_{n=0}^N$ be the Crank–Nicolson approximations for* (2.1),

$$(4.19) \qquad\qquad \bar{\partial}U^n + AU^{n-\frac{1}{2}} = \bar{f}^n\,,$$

where either $\bar{f}^n = f(t^{n-\frac{1}{2}})$ or $\bar{f}^n = \frac{1}{k_n}\int_{I_n} f(s)ds$. Then the following estimate holds for $m \leq N$:

$$(4.20) \qquad \sum_{n=1}^m k_n|A^{3/2}\bar{\partial}U^n|^2 + |A^2U^m|^2 \leq |A^2U^0|^2 + \sum_{n=1}^m k_n|A^{3/2}\bar{f}^n|^2\,.$$

Proof. We apply A to the scheme

$$A\bar{\partial}U^n + A^2U^{n-\frac{1}{2}} = A\bar{f}^n\,,$$

and take the inner product with $2k_nA^2(U^n - U^{n-1})$ to obtain

$$2k_n|A^{3/2}\bar{\partial}U^n|^2 + |A^2U^n|^2 - |A^2U^{n-1}|^2 = 2k_n\,(A\bar{f}^n, A^2\bar{\partial}U^n)\,.$$

Summing here from $n = 1$ to m, we get

$$\sum_{n=1}^m 2k_n|A^{3/2}\bar{\partial}U^n|^2 + |A^2U^m|^2 = |A^2U^0|^2 + 2\sum_{n=1}^m k_n(A\bar{f}^n, A^2\bar{\partial}U^n),$$

whence the Cauchy–Schwarz and the arithmetic-geometric mean inequalities yield

$$\sum_{n=1}^m 2k_n|A^{3/2}\bar{\partial}U^n|^2 + |A^2U^m|^2 \leq |A^2U^0|^2$$

$$+ \sum_{n=1}^m k_n|A^{3/2}\bar{f}^n|^2 + \sum_{n=1}^m k_n|A^{3/2}\bar{\partial}U^n|^2,$$

and the proof is complete. $\qquad\qquad \square$

From (4.6), (4.9), (4.20) and (3.5) we conclude the following Theorem. We emphasize that the optimal order *a priori* error estimate (4.23), depending only on the data (see below), follows from our a posteriori estimate (4.6). This shows, in particular, that the a posteriori estimate is of optimal (second) order.

Theorem 4.1 (Error estimates). *Let $\{U^n\}_{n=0}^N$ be either the Crank–Nicolson approximations or the Crank–Nicolson–Galerkin approximations to the solution of problem* (2.1), $e = u - U$ *and* $\hat{e} = u - \hat{U}$. *The following a posteriori estimate is valid for $m \leq N$:*

$$(4.21) \quad |\hat{e}(t^m)|^2 + \int_0^{t^m} \left(\|e(s)\|^2 + \frac{1}{2}\|\hat{e}(s)\|^2 \right) ds \leq \frac{\beta}{2} \sum_{n=1}^m k_n^5 |A^{3/2}\bar{\partial}U^n|^2 + \mathcal{E}[f],$$

with β given by (4.8) *and*

$$(4.22) \qquad \mathcal{E}[f] := 2 \int_0^{t^m} \|R_f(s)\|_\star^2 ds + \frac{\beta}{2} \sum_{n=1}^m k_n^5 \|\rho_{f,n}\|^2.$$

Here R_f and $\rho_{f,n}$ are given by (4.2) *and* (3.6), *respectively, for the Crank–Nicolson method, and by* (4.3) *and* (3.7) *for the Crank–Nicolson–Galerkin method. Furthermore, if $U^0 \in D(A^2)$ and $f(t) \in D(A^{3/2})$, then the following a priori estimate holds for $m \leq N$:*

$$(4.23) \quad \begin{aligned} |\hat{e}(t^m)|^2 &+ \int_0^{t^m} \left(\|e(s)\|^2 + \frac{1}{2}\|\hat{e}(s)\|^2 \right) ds \\ &\leq \frac{\beta}{2} \max_n k_n^4 \left(|A^2 U^0|^2 + \sum_{n=1}^m k_n |A^{3/2}\bar{f}^n|^2 \right) + \mathcal{E}[f], \end{aligned}$$

with $\bar{f}^n := f(t^{n-\frac{1}{2}})$ for the Crank–Nicolson method, and $\bar{f}^n := \frac{1}{k_n}\int_{I_n} f(s)ds$ for the Crank–Nicolson–Galerkin method.

Remark 4.3 (Equivalent upper bound for CNG). In the case of the Crank–Nicolson–Galerkin method, it is easily seen that (4.23) yields

$$(4.24) \quad \begin{aligned} |\hat{e}(t^m)|^2 &+ \int_0^{t^m} \left(\|e(s)\|^2 + \frac{1}{2}\|\hat{e}(s)\|^2 \right) ds \\ &\leq \frac{\beta}{2} \max_n k_n^4 \left(|A^2 U^0|^2 + \int_0^{t^m} |A^{3/2}f(s)|^2 ds \right) + \mathcal{E}[f]. \qquad \square \end{aligned}$$

Remark 4.4 (Estimate for $\mathcal{E}[f]$). As a by-product of interpolation theory, we realize that the following optimal order estimate is valid for the error $\mathcal{E}[f]$ in the forcing term f:

$$\mathcal{E}[f] \leq C \sum_{n=1}^m k_n^2 \int_{I_n} \left(\|f''(s)\|_\star^2 + \|f'(s)\|^2 \right) ds. \qquad \square$$

4.2. Estimates via Duhamel's principle. We first rewrite the relation (4.1) in the form

$$(4.25) \qquad \hat{e}'(t) + A\hat{e}(t) = R_{\hat{U}}(t) + R_f(t)$$

with

$$(4.26) \qquad R_{\hat{U}}(t) := A[U(t) - \hat{U}(t)].$$

We will use Duhamel's principle in (4.25). Let $E_A(t)$ be the solution operator of the homogeneous equation

$$(4.27) \qquad v'(t) + Av(t) = 0, \quad v(0) = w,$$

i.e., $v(t) = E_A(t)w$. 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 generated by the operator A. Duhamel's principle applied to (4.25) yields

$$(4.28) \qquad \hat{e}(t) = \int_0^t E_A(t-s)\big[R_{\hat{U}}(s) + R_f(s)\big]\, ds.$$

In the sequel we will use the smoothing property (cf., e.g., Crouzeix [5], Thomée [21])

$$(4.29) \qquad |A^\ell E_A(t)w| \le C_A \frac{1}{t^{\ell-m}}|A^m w|, \quad \ell \ge m \ge 0.$$

In addition, note that A and E_A commute, i.e., $A E_A(t)w = E_A(t)\, Aw$. In particular, (4.29) can also be written in the form

$$(4.30) \qquad |E_A(t)\, A^\ell w| \le C_A \frac{1}{t^{\ell-m}}|A^m w|, \quad \ell \ge m \ge 0,$$

whence $|E_A(t)w| \le C_A t^{-m}|A^{-m}w|$. Next, using (4.28) for $t = t^n$, we have

$$
\begin{aligned}
|\hat{e}(t^n)| &\le \int_{I_n} \big|E_A(t^n - s)\big[R_{\hat{U}}(s) + R_f(s)\big]\big|\, ds \\
&\quad + \int_0^{t^{n-1}} \big|E_A(t^n - s)\big[R_{\hat{U}}(s) + R_f(s)\big]\big|\, ds \\
&\le C_A \int_{I_n} \frac{1}{t^n - s}\big|A^{-1}R_{\hat{U}}(s)\big|\, ds + C_A \int_{I_n} \big|R_f(s)\big|\, ds \\
&\quad + C_A \int_0^{t^{n-1}} \frac{1}{t^n - s}\big|A^{-1}\big[R_{\hat{U}}(s) + R_f(s)\big]\big|\, ds
\end{aligned}
$$

and thus

$$(4.31)
\begin{aligned}
|\hat{e}(t^n)| &\le C_A \int_{I_n} \frac{1}{t^n - s}\big|A^{-1}R_{\hat{U}}(s)\big|\, ds + C_A \int_{I_n} \big|R_f(s)\big|\, ds \\
&\quad + C_A \sup_{s \in [0, t^{n-1}]} \big|A^{-1}\big[R_{\hat{U}}(s) + R_f(s)\big]\big| \int_0^{t^{n-1}} \frac{1}{t^n - s}\, ds.
\end{aligned}
$$

We now recall (4.26) and (3.5), namely,

$$(4.32) \qquad A^{-1}R_{\hat{U}}(s) = U(s) - \hat{U}(s) = \frac{1}{2}(t - t^{n-1})(t^n - t)\big(A\bar{\partial}U^n - \rho_{f,n}\big),$$

to obtain a bound for the first two terms on the right-hand side of (4.31),

$$(4.33) \qquad \int_{I_n} \frac{1}{t^n - s}\big|A^{-1}R_{\hat{U}}(s)\big|\, ds + \int_{I_n} \big|R_f(s)\big|\, ds \le \frac{k_n^2}{4}\big|A\bar{\partial}U^n\big| + \mathcal{E}_1[I_n; f],$$

with

$$(4.34) \qquad \mathcal{E}_1[I_n; f] := \frac{k_n^2}{4}\big|\rho_{f,n}\big| + k_n \max_{s \in I_n}\big|R_f(s)\big|.$$

In addition, again using (4.32), we have

$$
(4.35) \quad \sup_{s \in [0, t^{n-1}]} \left| A^{-1} \left[R_{\hat{U}}(s) + R_f(s) \right] \right| \int_0^{t^{n-1}} \frac{1}{t^n - s} ds
$$
$$
\leq \frac{1}{8} \ln(\frac{t^n}{k_n}) \max_{1 \leq j \leq n-1} \left(k_j^2 |A \bar{\partial} U^j| \right) + \ln(\frac{t^n}{k_n}) \, \mathcal{E}_2[\,[0, t^{n-1}]\,; f]
$$

with

$$
(4.36) \quad \mathcal{E}_2[\,[0, t^{n-1}]\,; f] := \frac{1}{8} \max_{1 \leq j \leq n-1} \left(k_j^2 |\rho_{f,j}| \right) + \max_{1 \leq j \leq n-1} \sup_{s \in I_j} |A^{-1} R_f(s)| .
$$

We have thus proved

Theorem 4.2 (Error estimates). *Let $\{U^n\}_{n=0}^N$ be either the Crank–Nicolson or the Crank–Nicolson–Galerkin approximations to the solution of problem (2.1). Then, with the notation of Theorem 4.1, the following a posteriori estimate is valid for $n \leq N$:*

$$
(4.37) \quad |\hat{e}(t^n)| \leq \frac{1}{8} C_A \left(2 k_n^2 |A \bar{\partial} U^n| + \ln(\frac{t^n}{k_n}) \max_{0 \leq j \leq n-1} k_j^2 |A \bar{\partial} U^j| \right)
$$
$$
+ C_A \left(\mathcal{E}_1[I_n; f] + \ln(\frac{t^n}{k_n}) \, \mathcal{E}_2[\,[0, t^{n-1}]\,; f] \right),
$$

with C_A the constant of (4.29) for $\ell = 1$, $m = 0$, and the terms involving f are defined in (4.34) and (4.36). Furthermore, if $U^0 \in D(A^2), f(t) \in D(A^{3/2})$, and $k := \max_{0 \leq j \leq n} k_j \big(2 + \ln(\frac{t^n}{k_n}) \big)^{1/2}$, the following a priori estimate holds for $n \leq N$:

$$
(4.38) \quad |\hat{e}(t^n)| \leq \frac{1}{8} \, C_A \, k^2 \Big\{ \Big(|A^2 U^0|^2 + \sum_{j=1}^n k_j |A^{3/2} \bar{f}^j|^2 \Big)^{1/2} + \max_{1 \leq j \leq n} |A \bar{f}^j| \Big\}
$$
$$
+ C_A \left(\mathcal{E}_1[I_n; f] + \ln(\frac{t^n}{k_n}) \, \mathcal{E}_2[\,[0, t^{n-1}]\,; f] \right).
$$

Proof. It only remains to prove (4.38). It immediately follows from (4.19) that

$$
|A \bar{\partial} U^n| \leq |A^2 U^{n-\frac{1}{2}}| + |A \bar{f}^n| \leq \max \left(|A^2 U^{n+1}|, |A^2 U^n| \right) + |A \bar{f}^n|,
$$

and (4.38) thus results from (4.37) in light of (4.20). $\qquad \square$

Remark 4.5 (Alternative bound for CNG). In the case of the Crank–Nicolson–Galerkin method, it is easily seen that (4.38) yields

$$
(4.39) \quad |\hat{e}(t^n)| \leq \frac{1}{8} C_A k^2 \Big\{ \Big(|A^2 U^0|^2 + \int_0^{t^n} |A^{3/2} f(s)|^2 ds \Big)^{1/2} + \max_{0 \leq s \leq t^n} |A f(s)| \Big\}
$$
$$
+ C_A \left(\mathcal{E}_1[I_n; f] + \ln(\frac{t^n}{k_n}) \, \mathcal{E}_2[\,[0, t^n]\,; f] \right). \qquad \square
$$

5. ESTIMATES FOR INITIAL DATA WITH REDUCED SMOOTHNESS

In this section our objective is the derivation of a posteriori error estimates in the case of initial data with reduced smoothness.

Since the initial value problem (2.1) can be split into one with homogeneous initial data and one with homogeneous equation, and we are mainly concerned with

the effect of the initial data, we focus on the homogeneous initial value problem

(5.1)
$$\begin{cases} u'(t) + Au(t) = 0, & 0 < t < T, \\ u(0) = u^0. \end{cases}$$

A typical a priori nonsmooth data estimate reads (see [21]),

(5.2)
$$|u(t^n) - U^n| \le C \left(\frac{k}{t^n}\right)^s |U^0|,$$

s being the order of the method at hand and k the constant time step, $t^n = nk$. It is well known that estimates of this type are available for strongly $A(0)$-stable schemes, such as the Runge–Kutta–Radau IIA methods, and, in particular, the backward Euler method. Similar estimates are not available for $A(0)$-stable schemes. A way of securing such estimates for $A(0)$-stable schemes is to start the time-stepping procedure with a few steps of a strongly $A(0)$-stable scheme of order $s - 1$ and then proceed with the main scheme. For instance, for the Crank–Nicolson method it suffices to perform the first two steps with the backward Euler scheme and subsequently proceed with the Crank–Nicolson method. The use of the Euler scheme as a starting procedure for the Crank–Nicolson method in order to establish a priori error estimates for nonsmooth initial data was suggested by Luskin and Rannacher (see [21], Theorems 7.4 and 7.5). In the a posteriori error analysis we have to derive estimates with reduced regularity requirements on the initial data, but in a form that allows efficient monitoring of the time steps. To this end we will establish estimates that are the a posteriori analogs of Theorems 7.4 and 7.5 of [21] for the following modification of the scheme: We let $U^0 := u^0$ and define the approximations U^m to $u^m := u(t^m)$, $m = 1, \ldots, N$, by

(5.3i)
$$\bar{\partial} U^n + AU^n = 0, \quad n = 1, 2,$$

(5.3ii)
$$\bar{\partial} U^n + AU^{n-\frac{1}{2}} = 0, \quad n = 3, \ldots, N.$$

Note first that, even for $U^0 \in H$, due to the fact that U^1 and U^2 are backward Euler approximations, we have $U^1 \in D(A)$ and $U^2 \in D(A^2)$; then, obviously, $U^n \in D(A^2)$ for $n \ge 2$. We now proceed to the definition of the reconstruction U.

5.1. **Reconstruction.** Given the nodal approximations $U^0, \ldots, U^N$, we define the approximation $U(t)$ to $u(t)$, $t \in [0, T]$, in the intervals I_1 and I_2 by interpolating by piecewise constants at their right ends and at the other subintervals by linearly interpolating between the nodal values, i.e.,

(5.4i)
$$U(t) = U^n, \quad t \in I_n, \ n = 1, 2,$$

(5.4ii)
$$U(t) = U^n + (t - t^n)\bar{\partial} U^n, \quad t \in I_n, \ n = 3, \ldots, N-1$$

(cf. (2.3)). Furthermore, we let the reconstruction $\hat{U}$ of U be given by

(5.5)
$$\hat{U}(t) := U^{n-1} - \int_{t^{n-1}}^t AU(s)\, ds, \quad t \in I_n,$$

i.e.,

(5.6i)
$$\hat{U}(t) = U^{n-1} - (t - t^{n-1})AU^n, \quad t \in I_n, \ n = 1, 2,$$

(5.6ii)
$$\hat{U}(t) = U^{n-1} - \frac{1}{2}(t - t^{n-1})A[U(t) + U^{n-1}], \quad t \in I_n, \ 3 \le n \le N.$$

We observe that $(5.6i)$ coincides with the continuous piecewise linear reconstruction of [19], whereas $(5.6ii)$ agrees with the continuous piecewise quadratic reconstruction (2.10) for $f = 0$. In view of $(5.3i)$, we obtain

$(5.7i)$ $\qquad\qquad U(t) - \hat{U}(t) = -(t^n - t)AU^n, \quad t \in I_n,\ n = 1, 2.$

Furthermore, in view of $(5.3ii)$,

$(5.7ii)$ $\qquad U(t) - \hat{U}(t) = -\frac{1}{2}(t - t^{n-1})(t^n - t)A\bar{\partial}U^n, \quad t \in I_n,\ 3 \le n \le N$

(see (3.1)). Note that, even for $U^0 \in H\backslash D(A)$, $\hat{U}$ is well defined; this would not be the case if U^1 and U^2 were Crank–Nicolson approximations because then $U^n \notin D(A)$ for any n.

It immediately follows from (5.5) that

(5.8) $\qquad\qquad\qquad\qquad \hat{e}'(t) + Ae(t) = 0, \quad 0 < t < T;$

compare with (4.1) for $f = 0$. Next, we will briefly discuss some of the estimators we obtain by applying the energy method or Duhamel's principle to the above error equation.

5.2. **Estimate I: Energy method.** Taking in (5.8) the inner product with $\hat{e}(t)$, and recalling that

$$\big(Ae(t), \hat{e}(t)\big) = \frac{1}{2}\Big(\|e(t)\|^2 + \|\hat{e}(t)\|^2 - \|\hat{e}(t) - e(t)\|^2\Big),$$

we obtain

(5.9) $\quad |\hat{e}(t)|^2 + \int_0^t \Big(\|\hat{e}(s)\|^2 + \|e(s)\|^2\Big)ds \le \int_0^t \|\hat{U}(s) - U(s)\|^2\,ds, \quad 0 \le t \le T$

(cf. (4.6)). Now, in view of $(5.7i)$,

$$\|\hat{U}(t) - U(t)\|^2 = (t^n - t)^2 |A^{3/2}U^n|^2, \quad n = 1, 2;$$

therefore

(5.10)
$$\int_0^{t^2} \|\hat{U}(s) - U(s)\|^2\,ds = \frac{1}{3}\Big(k_1^3|A^{3/2}U^1|^2 + k_2^3|A^{3/2}U^2|^2\Big)$$
$$= \frac{1}{3}\Big(k_1^3|A^{1/2}\bar{\partial}U^1|^2 + k_2^3|A^{1/2}\bar{\partial}U^2|^2\Big).$$

Furthermore,

(5.11) $\qquad\qquad \int_{t^2}^{t^m} \|\hat{U}(s) - U(s)\|^2\,ds \le \frac{\beta}{2}\sum_{n=2}^{m-1} k_n^5|A^{3/2}\bar{\partial}U^n|^2,$

with β as in (4.8). We thus deduce the *upper* a posteriori estimate

(5.12)
$$|\hat{e}(t)|^2 + \int_0^t \Big(\|\hat{e}(s)\|^2 + \|e(s)\|^2\Big)ds \le \frac{1}{3}\Big(k_1^3|A^{1/2}\bar{\partial}U^1|^2 + k_2^3|A^{1/2}\bar{\partial}U^2|^2\Big)$$
$$+ \frac{\beta}{2}\sum_{n=2}^{m-1} k_n^5|A^{3/2}\bar{\partial}U^n|^2;$$

compare with (4.21) for $f = 0$. Proceeding as in Subsection 4.1.2 we also get a sharp *lower* bound. Note that the above estimate holds, provided that $U^0 \in D(A^{1/2})$. Further reasonable error control based on this estimate requires us to balance the terms $k_1|A^{1/2}\bar{\partial}U^1|$, $k_2|A^{1/2}\bar{\partial}U^2|$ and $k_n^2|A^{3/2}\bar{\partial}U^n|$.

5.3. **Estimate II: Duhamel principle.** We next modify arguments of Sections 4.2 to obtain the final result

(5.13)
$$
\begin{aligned}
|\hat{e}(t^n)| \leq \frac{1}{8} C_A \Big\{ &2k_n^2 |A\bar{\partial}U^n| \\
&+ \ln(\frac{t^n}{k_n}) \max \Big(k_1 |\bar{\partial}U^1|, k_2 |\bar{\partial}U^2|, \max_{2 \leq j \leq n-1} \big(k_j^2 |A\bar{\partial}U^j| \big) \Big) \Big\},
\end{aligned}
$$

which could be compared with (4.37) for $f = 0$. Note that the above estimate holds, provided that $U^0 \in H$. Further reasonable error control based on this estimate requires us to balance the terms $k_1 |\bar{\partial}U^1|, k_2 |\bar{\partial}U^2|$ and $k_n^2 |A\bar{\partial}U^n|$.

6. Error estimates for nonlinear equations

In this section we consider the discretization of (1.1). We assume that $B(t, \cdot)$ can be extended to an operator from V into $V^\star$. A natural condition for (1.1) to be locally of parabolic type is the following local one-sided Lipschitz condition:

(6.1)
$$
\big(B(t,v) - B(t,w), v - w \big) \leq \lambda \|v - w\|^2 + \mu |v - w|^2 \quad \forall v, w \in T_u
$$

in a tube T_u, $T_u := \{v \in V : \min_t \|u(t) - v\| \leq 1\}$, around the solution u, uniformly in t, with a constant λ less than one and a constant μ. With $F(t,v) := Av - B(t,v)$, it is easily seen that (6.1) can be written in the form of a Gårding-type inequality,

(6.2)
$$
(F(t,v) - F(t,w), v - w) \geq (1-\lambda)\|v - w\|^2 - \mu |v - w|^2 \quad \forall v, w \in T_u .
$$

Furthermore, in order to ensure that an appropriate residual is of the correct order, we will make use of the following local Lipschitz condition for $B(t, \cdot)$:

(6.3)
$$
\|B(t,v) - B(t,w)\|_\star \leq L \|v - w\| \quad \forall v, w \in T_u
$$

with a constant L, *not necessarily less than one.* Here the tube T_u is defined in terms of the norm of V for concreteness. The analysis may be modified to yield a posteriori error estimates under analogous conditions for v and w belonging to tubes defined in terms of other norms, not necessarily the same for both arguments.

In the sequel the estimates are proved *under the assumption* that $\hat{U}(t), U(t) \in T_u$, for all $t \in [0, T]$. This assumption can, in some cases, be verified a posteriori under *conditional assumptions* on $\hat{U}$ and U. Thus the final result will hold, pending on a condition that $\hat{U}$ and U may or may not satisfy. However, the validity of this condition can be computationally verified. The derivation of these bounds requires the use of fine properties of the specific underlying p.d.e., as was done in [16, 18], and therefore goes beyond the scope of the present paper.

We refer to [3] for existence and local uniqueness results for the continuous Galerkin approximations, in particular for the Crank–Nicolson–Galerkin approximations, as well as for a priori error estimates. Concrete examples of parabolic equations satisfying (6.1) and (6.3) in appropriate tubes are given in [1] and [2].

6.1. **The Crank–Nicolson–Galerkin method.** We recall that this method for (1.1) consists of seeking a function $U : [0, T] \to V$, continuous and piecewise linear, such that $U(0) = u(0)$ and

(6.4)
$$
\bar{\partial}U^n + AU^{n-\frac{1}{2}} = \frac{1}{k_n} \int_{I_n} B(s, U(s)) ds,
$$

where $U^n := U(t^n), U^{n-\frac{1}{2}} := \frac{1}{2}(U^n + U^{n-1}) = U(t^{n-\frac{1}{2}})$. The Crank–Nicolson–Galerkin approximate solution U can be expressed in terms of its nodal values U^{n-1} and U^n,

$$(6.5) \qquad U(t) = U^{n-\frac{1}{2}} + (t - t^{n-\frac{1}{2}})\bar{\partial}U^n, \quad t \in I_n.$$

In view of (2.22), we let the *Crank–Nicolson–Galerkin reconstruction* $\hat{U}$ of U be

$$(6.6) \qquad \hat{U}(t) := U^{n-1} - A \int_{t^{n-1}}^t U(s)\, ds + \int_{t^{n-1}}^t P_1 B(s, U(s)) ds,$$

i.e.,

$$(6.7)\ \hat{U}(t) = U^{n-1} - \frac{1}{2}(t - t^{n-1})A[U(t) + U^{n-1}] + \int_{t^{n-1}}^t P_1 B(s, U(s)) ds, \quad t \in I_n.$$

It immediately follows from (6.6) that

$$(6.8) \qquad \hat{U}'(t) + AU(t) = P_1 B(t, U(t)).$$

Using (6.4), it easily follows from (6.5) and (6.7) that

$$(6.9) \qquad \begin{aligned} U(t) - \hat{U}(t) = &-\frac{1}{2}(t - t^{n-1})(t^n - t)A\bar{\partial}U^n \\ &+ \frac{t - t^{n-1}}{k_n} \int_{I_n} B(s, U(s)) ds - \int_{t^{n-1}}^t P_1 B(s, U(s)) ds; \end{aligned}$$

therefore, again using (6.4), we have

$$(6.10) \qquad \begin{aligned} U(t) - \hat{U}(t) = &\frac{1}{2}(t - t^{n-1})(t^n - t)\Big(A^2 U^{n-\frac{1}{2}} - \frac{1}{k_n} \int_{I_n} AB(s, U(s)) ds\Big) \\ &+ \frac{t - t^{n-1}}{k_n} \int_{I_n} B(s, U(s)) ds - \int_{t^{n-1}}^t P_1 B(s, U(s)) ds, \end{aligned}$$

$t \in I_n$. Hence, in view of (3.3),

$$(6.10') \qquad \begin{aligned} U(t) - \hat{U}(t) = &\frac{1}{2}(t - t^{n-1})(t^n - t)\Big(A^2 U^{n-\frac{1}{2}} - \frac{1}{k_n} \int_{I_n} AB(s, U(s)) ds\Big) \\ &+ \frac{6}{k_n^3}(t - t^{n-1})(t^n - t) \int_{I_n} B(s, U(s))(s - t^{n-\frac{1}{2}}) ds, \end{aligned}$$

$t \in I_n$, from which we immediately see that $\max_{t \in I_n} |U(t) - \hat{U}(t)| = O(k_n^2)$.

6.2. The Crank–Nicolson method. The Crank–Nicolson approximations $U^m \in D(A)$ to the nodal values $u^m := u(t^m)$ of the solution u of (1.1) are defined by

$$(6.11) \qquad \bar{\partial}U^n + AU^{n-\frac{1}{2}} = B(t^{n-\frac{1}{2}}, U^{n-\frac{1}{2}}), \quad n = 1, \dots, N,$$

with $U^0 := u(0)$. As before, we define the Crank–Nicolson approximation U to u by linearly interpolating between the nodal values U^{n-1} and U^n,

$$(6.12) \qquad U(t) = U^{n-\frac{1}{2}} + (t - t^{n-\frac{1}{2}})\bar{\partial}U^n, \quad t \in I_n.$$

Let $b : I_n \to H$ be the linear interpolant of $B(\cdot, U(\cdot))$ at the nodes t^{n-1} and $t^{n-\frac{1}{2}}$,

$$(6.13) \qquad \begin{aligned} b(t) = &B(t^{n-\frac{1}{2}}, U^{n-\frac{1}{2}}) \\ &+ \frac{2}{k_n}(t - t^{n-\frac{1}{2}})[B(t^{n-\frac{1}{2}}, U^{n-\frac{1}{2}}) - B(t^{n-1}, U^{n-1})], \quad t \in I_n. \end{aligned}$$

Inspired by (2.9), we define the *Crank–Nicolson reconstruction* $\hat{U}$ of U by

$$(6.14) \qquad \hat{U}(t) := U^{n-1} - A \int_{t^{n-1}}^{t} U(s)\, ds + \int_{t^{n-1}}^{t} b(s)\, ds, \quad t \in I_n,$$

i.e.,

$$(6.15) \qquad \begin{aligned} \hat{U}(t) =& U^{n-1} - \frac{1}{2}(t - t^{n-1})A[U(t) + U^{n-1}] + (t - t^{n-1})B(t^{n-\frac{1}{2}}, U^{n-\frac{1}{2}}) \\ &+ \frac{1}{k_n}(t - t^{n-1})(t^n - t)[B(t^{n-\frac{1}{2}}, U^{n-\frac{1}{2}}) - B(t^{n-1}, U^{n-1})], \end{aligned}$$

$t \in I_n$. Note that (6.15) reduces to (2.10) in the case B is independent of u.

It immediately follows from (6.14) that

$$(6.16) \qquad \hat{U}'(t) + AU(t) = b(t), \quad t \in I_n.$$

Furthermore, it is easily seen that, for $t \in I_n$,

$$(6.17) \qquad \begin{aligned} &U(t) - \hat{U}(t) \\ &= -\frac{1}{2}(t - t^{n-1})(t^n - t)\Big\{ A\bar{\partial}U^n - \frac{2}{k_n}\big[B(t^{n-\frac{1}{2}}, U^{n-\frac{1}{2}}) - B(t^{n-1}, U^{n-1})\big] \Big\}. \end{aligned}$$

6.3. Error estimates. We now derive a posteriori error estimates for both the Crank–Nicolson–Galerkin and the Crank–Nicolson method. Let $e := u - U$ and $\hat{e} := u - \hat{U}$. The following estimates hold *under the assumption* that $\hat{U}(t), U(t) \in T_u$ for all $t \in [0, T]$.

Crank–Nicolson–Galerkin method. Subtracting (6.8) from the differential equation in (1.1), we obtain

$$(6.18) \qquad \hat{e}'(t) + Ae(t) = B(t, u(t)) - P_1 B(t, U(t)),$$

which we rewrite in the form

$$(6.19) \qquad \hat{e}'(t) + Ae(t) = B(t, u(t)) - B(t, U(t)) + R_U(t),$$

with

$$(6.20) \qquad R_U(t) = B(t, U(t)) - P_1 B(t, U(t)).$$

Now,

$$\big(B(t, u(t)) - B(t, U(t)), \hat{e}(t)\big) = \big(B(t, u(t)) - B(t, U(t)), e(t)\big) \\ + \big(B(t, u(t)) - B(t, U(t)), U(t) - \hat{U}(t)\big),$$

and, in view of (6.1) and (6.3), elementary calculations yield

$$(6.21) \quad \big(B(t, u(t)) - B(t, U(t)), \hat{e}(t)\big) \le \lambda\|\hat{e}(t)\|^2 + \mu|e(t)|^2 + L\|e(t)\|\,\|(\hat{U} - U)(t)\|.$$

Similarly,

$$\big(B(t, u(t)) - B(t, U(t)), \hat{e}(t)\big) = \big(B(t, u(t)) - B(t, \hat{U}(t)), \hat{e}(t)\big) \\ + \big(B(t, \hat{U}(t)) - B(t, U(t)), \hat{e}(t)\big)$$

and

$$(6.22) \quad \big(B(t, u(t)) - B(t, U(t)), \hat{e}(t)\big) \le \lambda\|\hat{e}(t)\|^2 + \mu|\hat{e}(t)|^2 + L\|\hat{e}(t)\|\,\|(\hat{U} - U)(t)\|.$$

Summing (6.21) and (6.22), we obtain

$$(6.23) \qquad \begin{aligned} 2\big(B(t, u(t)) - B(t, U(t)), \hat{e}(t)\big) \le& \lambda\big(\|\hat{e}(t)\|^2 + \|e(t)\|^2\big) \\ &+ \mu\big(|\hat{e}(t)|^2 + |e(t)|^2\big) + L\big(\|\hat{e}(t)\| + \|e(t)\|\big)\|(\hat{U} - U)(t)\|. \end{aligned}$$

Now,

$$|e(t)|^2 \leq \big(|\hat{e}(t)| + |(\hat{U} - U)(t)|\big)^2 \leq 2|\hat{e}(t)|^2 + 2|(\hat{U} - U)(t)|^2$$

and

$$L\big(\|\hat{e}(t)\| + \|e(t)\|\big)\|(\hat{U} - U)(t)\|$$

$$\leq \frac{\varepsilon}{2}\big(\|\hat{e}(t)\| + \|e(t)\|\big)^2 + \frac{L^2}{2\varepsilon}\|(\hat{U} - U)(t)\|^2$$

$$\leq \varepsilon\big(\|\hat{e}(t)\|^2 + \|e(t)\|^2\big) + \frac{L^2}{2\varepsilon}\|(\hat{U} - U)(t)\|^2.$$

Consequently, from (6.23), we obtain

$$(6.24) \quad \begin{aligned} 2\big(B(t, u(t)) - B(t, U(t)), \hat{e}(t)\big) &\leq \lambda\big(\|\hat{e}(t)\|^2 + \|e(t)\|^2\big) + 3\mu|e(t)|^2 \\ &+ 2\mu|(\hat{U} - U)(t)|^2 + \varepsilon\big(\|\hat{e}(t)\|^2 + \|e(t)\|^2\big) + \frac{L^2}{2\varepsilon}\|(\hat{U} - U)(t)\|^2 \end{aligned}$$

for any positive ε.

Taking in (6.19) the inner product with $\hat{e}(t)$, we obtain

$$(6.25) \quad \begin{aligned} \frac{d}{dt}|\hat{e}(t)|^2 + \|e(t)\|^2 + \|\hat{e}(t)\|^2 &= \|\hat{U}(t) - U(t)\|^2 \\ &+ 2\big(B(t, u(t)) - B(t, U(t)), \hat{e}(t)\big) + 2\big(R_U(t), \hat{e}(t)\big), \end{aligned}$$

whence, in view of (6.24),

$$(6.26) \quad \begin{aligned} \frac{d}{dt}|\hat{e}(t)|^2 + (1 - \lambda - 2\varepsilon)\big(\|\hat{e}(t)\|^2 + \|e(t)\|^2\big) &\leq 3\mu|\hat{e}(t)|^2 \\ &+ \big(1 + \frac{L^2}{2\varepsilon}\big)\|(\hat{U} - U)(t)\|^2 + 2\mu|(\hat{U} - U)(t)|^2 + \frac{1}{\varepsilon}\|R_U(t)\|_\star^2. \end{aligned}$$

We thus easily obtain the desired a posteriori error estimate via Gronwall's lemma

$$(6.27) \quad \begin{aligned} |\hat{e}(t)|^2 + (1 - \lambda - 2\varepsilon)\int_0^t e^{3\mu(t-s)}\Big(\|\hat{e}(s)\|^2 + \|e(s)\|^2\Big)ds \\ \leq \int_0^t e^{3\mu(t-s)}\Big[\big(1 + \frac{L^2}{2\varepsilon}\big)\|(\hat{U} - U)(s)\|^2 \\ + 2\mu|(\hat{U} - U)(s)|^2 + \frac{1}{\varepsilon}\|R_U(s)\|_\star^2\Big]ds. \end{aligned}$$

Crank–Nicolson method. Subtracting (6.16) from the differential equation in (1.1), we obtain

$$\hat{e}'(t) + Ae(t) = B(t, u(t)) - b(t).$$

Therefore, (6.27) is valid for the Crank–Nicolson method as well, this time with

$$R_U(t) := B(t, U(t)) - b(t).$$

References

1. G. Akrivis, M. Crouzeix, and Ch. Makridakis, *Implicit-explicit multistep finite element methods for nonlinear parabolic problems*, Math. Comp. **67** (1998) 457–477. MR1458216 (98g:65088)
2. G. Akrivis, M. Crouzeix, and Ch. Makridakis, *Implicit-explicit multistep methods for quasilinear parabolic equations*, Numer. Math. **82** (1999) 521–541. MR1701828 (2000e:65075)
3. G. Akrivis and Ch. Makridakis, *Galerkin time-stepping methods for nonlinear parabolic equations*, M^{2}AN Math. Mod. Numer. Anal. **38** (2004) 261–289. MR2069147 (2005f:65124)
4. A. K. Aziz and P. Monk, *Continuous finite elements in space and time for the heat equation*, Math. Comp. **52** (1989) 255–274. MR0983310 (90d:65189)

5. M. Crouzeix, *Parabolic Evolution Problems*. Unpublished manuscript, 2003.
6. W. Dörfler, *A time- and space-adaptive algorithm for the linear time-dependent Schrödinger equation*, Numer. Math. **73** (1996) 419–448. MR1393174 (97g:65183)
7. 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. MR1083324 (91m:65274)
8. K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems. IV. Nonlinear problems*, SIAM J. Numer. Anal. **32** (1995) 1729–1749. MR1360457 (96i:65081)
9. 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. MR1620144 (99d:65281)
10. D. Estep and D. French, *Global error control for the continuous Galerkin finite element method for ordinary differential equations*, RAIRO Math. Mod. Numer. Anal. **28** (1994) 815–852. MR1309416 (95k:65079)
11. C. Johnson, *Error estimates and adaptive time-step control for a class of one-step methods for stiff ordinary differential equations*, SIAM J. Numer. Anal. **25** (1988) 908–926. MR0954791 (90a:65160)
12. C. Johnson, Y.-Y. Nie, and V. Thomée, *An a posteriori error estimate and adaptive timestep control for a backward Euler discretization of a parabolic problem*, SIAM J. Numer. Anal. **27** (1990) 277–291. MR1043607 (91g:65199)
13. C. Johnson and A. Szepessy, *Adaptive finite element methods for conservation laws based on a posteriori error estimates*, Comm. Pure Appl. Math. **48** (1995) 199–234. MR1322810 (97b:76084)
14. O. Karakashian and Ch. Makridakis, *A space-time finite element method for the nonlinear Schrödinger equation: the continuous Galerkin method*, SIAM J. Numer. Anal. **36** (1999) 1779–1807. MR1712169 (2000h:65139)
15. O. Karakashian and Ch. Makridakis, *Convergence of a continuous Galerkin method with mesh modification for nonlinear wave equations*, Math. Comp. **74** (2005) 85–102. MR2085403 (2005g:65147)
16. 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. MR2139228
17. Ch. Makridakis and R. H. Nochetto, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal. **41** (2003) 1585–1594. MR2034895 (2004k:65157)
18. Ch. Makridakis and R. H. Nochetto, *A posteriori error analysis for higher order dissipative methods for evolution problems*. (Submitted for publication).
19. 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. MR1737503 (2000k:65142)
20. R. H. Nochetto, A. Schmidt, and C. Verdi, *A posteriori error estimation and adaptivity for degenerate parabolic problems*, Math. Comp. **69** (2000) 1–24. MR1648399 (2000i:65136)
21. V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*. Springer-Verlag, Berlin, 1997. MR1479170 (98m:65007)
22. R. Verfürth, *A posteriori error estimates for finite element discretizations of the heat equation*, Calcolo **40** (2003) 195–212. MR2025602 (2005f:65131)

COMPUTER SCIENCE DEPARTMENT, UNIVERSITY OF IOANNINA, 451 10 IOANNINA, GREECE
E-mail address: `akrivis@cs.uoi.gr`

DEPARTMENT OF APPLIED MATHEMATICS, UNIVERSITY OF CRETE, 71409 HERAKLION-CRETE, GREECE – AND – INSTITUTE OF APPLIED AND COMPUTATIONAL MATHEMATICS, FORTH, 71110 HERAKLION-CRETE, GREECE
URL: `http://www.tem.uoc.gr/~makr`
E-mail address: `makr@math.uoc.gr`
E-mail address: `makr@tem.uoc.gr`

DEPARTMENT OF MATHEMATICS AND INSTITUTE FOR PHYSICAL SCIENCE AND TECHNOLOGY, UNIVERSITY OF MARYLAND, COLLEGE PARK, MARYLAND 20742
URL: `http://www.math.umd.edu/~rhn`
E-mail address: `rhn@math.umd.edu`