A POSTERIORI ERROR ANALYSIS FOR HIGHER ORDER DISSIPATIVE METHODS FOR EVOLUTION PROBLEMS

CHARALAMBOS MAKRIDAKIS AND RICARDO H. NOCHETTO

ABSTRACT. We prove a posteriori error estimates for time discretizations by the discontinuous Galerkin method dG(q) and the corresponding implicit Runge-Kutta-Radau method IRK-R(q) of arbitrary order $q \geq 0$ for both linear and nonlinear evolution problems of the form $u' + \mathfrak{F}(u) = f$, with γ^2-angle bounded operator $\mathfrak{F}$. The key ingredient is a novel higher order reconstruction $\widehat{U}$ of the discrete solution U, which restores continuity and leads to the differential equation $\widehat{U}' + \Pi\mathfrak{F}(U) = F$ for a suitable interpolation operator Π and piecewise polynomial approximation F of f. We discuss applications to linear PDE, such as the convection-diffusion equation ($\gamma \geq \frac{1}{2}$) and the wave equation (formally $\gamma = \infty$), and nonlinear PDE corresponding to subgradient operators ($\gamma = 1$), such as the p-Laplacian, as well as Lipschitz operators ($\gamma \geq \frac{1}{2}$). We also derive conditional a posteriori error estimates for the time-dependent minimal surface problem.

1. Introduction

In this paper we study the time discretization, via the discontinuous Galerkin method dG(q) and the corresponding implicit Runge-Kutta-Radau IIA method IRK-R(q) of any order $q \geq 0$, of the *dissipative* initial value problem

$$u' + \mathfrak{F}(u) = f, \qquad u(0) = u_0, \tag{1.1}$$

governed by a γ^2-angle bounded operator $\mathfrak{F}$ in a Hilbert space $(\mathcal{H}, \langle\cdot,\cdot\rangle, |\cdot|)$. We present a rigorous energy a posteriori error analysis based on the novel idea of *reconstruction*. Our results can be viewed as extensions of the optimal error estimates of Nochetto, Savaré and Verdi for the implicit Euler method ($q = 0$) [34, 35] to the higher order dG(q) and IRK-R(q) methods; see also [31, 32].

We would like first to place our contributions in context. A posteriori error analysis for time dependent partial differential equations (PDE) has received a great deal of attention in the last few years, particularly so for low order methods. The existing estimators can be classified as either *analytical* or *computational*, and they have been derived by two different, though somewhat related, techniques. They are the *energy* and *duality* methods, which are briefly described below. The former

Date: March 14, 2006.

2000 *Mathematics Subject Classification.* 65M15, 65M50.

Key words and phrases. A posteriori error analysis, discontinuous Galerkin, Runge-Kutta-Radau, reconstruction, γ^2-angle bounded operator, energy method.

consists of testing the error representation formula with the error (or an integral or derivative of it) and employing stability properties of the PDE. This technique has been used by Nochetto, Savaré and Verdi [35] to study the backward Euler method for a class of dissipative PDE closely related to those in this paper; see [31, 32] and also [4, 8, 37, 43, 44] for analogous results for fully discrete schemes. In contrast, the duality method, initiated by Johnson, relies on the stability properties of a linear backward dual problem [3, 11, 12, 10, 13, 16, 17, 26]. This theory is sharp for linear PDE [3, 11, 10, 13, 26]. For nonlinear PDE an analysis based on duality has to overcome a subtle issue: the strong stability of the *linearized* dual problem. The nonlinear PDE must be linearized around a convex combination of the *exact and discrete* solution to ensure an exact error representation formula [17, Section 2.3]. This leads to a posteriori bounds provided one is able to show that the linearized dual problem is stable with a stability factor independent of the exact solution. This question, being problem dependent, remains open in general; we refer to [36] for a successfull application to nonlinear (degenerate) parabolic problems, which possess rather weak stability properties. Treating higher order methods, and extending their applicability to nonlinear problems under realistic and computationally verifiable assumptions, requires new ideas. In this paper we bridge the gap for the class of γ^2-angle bounded operators upon deriving new error representation formulas for the discontinuous Galerkin method dG(q) and the corresponding implicit Runge-Kutta-Radau IIA method IRK-R(q) of any order $q \geq 0$. These formulas can be used then to obtain error estimates by either energy or duality methods; here we use the former because it handles easily the nonlinearities under study. Since these error estimates hinge upon analysis of the underlying PDE, they are *analytical*.

In contrast, recent activity has centered around the actual computation of the dual solution, see e.g., [3, 14, 17, 22]. For nonlinear PDE, this entails replacing the unknown exact solution u by the computed discrete solution in the coefficients of the linearized dual problem. In addition, the choice of terminal data for the dual problem requires special care except when estimating a linear functional of the solution [17, Section 42]; this is for instance the case of [3, 21, 22]. However, if the goal is to compute a *norm* of the error, as in this paper, then the choice of data is unclear because it depends on the error itself. To deal with this matter, a probabilistic approach is proposed in [17, Section 4.2], whereas computational evidence is provided in [14] that the strong stability factors are of moderate size and somewhat insensitive to the terminal condition as well as time. Once this crucial issue has been resolved, this approach gives a *computational* upper error estimate.

If u denotes the exact solution of (1.1), U an approximation of u and $\|\cdot\|$ a norm, not necessarily $|\cdot|$, the aim of this paper is to obtain *a posteriori* error estimates of the form

$$\|u - U\| \leq \eta(U, f, u_0)$$

where the estimator $\eta(U, f, u_0)$ exhibits the following properties:

- $\eta(U, f, u_0)$ is a computable quantity which solely depends on U, f and u_0;
- $\eta(U, f, u_0)$ is of optimal order and entails minimal regularity;
- $\eta(U, f, u_0)$ utilizes explicit and computable constants for linear problems.

To this end, we restrict ourselves to the class of γ^2-angle bounded operators. The operator $\mathfrak{F} : D(\mathfrak{F}) \to \mathcal{H}$ is γ^2-*angle bounded* provided

$$\langle \mathfrak{F}(v) - \mathfrak{F}(w), w - z\rangle \le \gamma^2 \langle \mathfrak{F}(v) - \mathfrak{F}(z), v - z\rangle \quad \forall\, v, w, z \in D(\mathfrak{F}) \tag{1.2}$$

with $\gamma \ge \frac{1}{2}$ [6, 7, 34, 35]. This class of operators is a natural extension of linear *sectorial* operators [6]; this connection is explored in Sections 3 and 4. Upon taking $w = z$, the structural condition (1.2) implies that $\mathfrak{F}$ is *monotone*

$$0 \le \langle \mathfrak{F}(v) - \mathfrak{F}(z), v - z\rangle \quad \forall\, v, z \in D(\mathfrak{F}), \tag{1.3}$$

but it is indeed much stronger; we formally write $\gamma = \infty$ in (1.2) to indicate (1.3). We develop most of the theory under the assumption (1.2), which yields optimal order-regularity error estimates. We also consider monotone operators for which optimal order comes at the expense of extra regularity. In addition to (1.2), we assume that $\mathfrak{F}$ satisfies *coercivity* conditions with respect to a norm $\|\cdot\|$ on $D(\mathfrak{F})$, namely

$$\langle \mathfrak{F}(v) - \mathfrak{F}(w), v - w\rangle \ge \|v - w\|^2 \quad \forall\, v, w \in D(\mathfrak{F}). \tag{1.4}$$

This is the case of linear problems with smoothing effect ($\mathfrak{F}$ sectorial) in Section 3, and several nonlinear counterparts in Sections 4 and 5. Other notions of dissipativity weaker than (1.4) have been studied in [18, 40] and [17] from the dynamical system and invariant regions point of view, respectively. The emphasis has then been on qualitative properties of the numerical scheme, particularly long time behavior, rather than rigorous error analysis. Since the class of $gamma^2$-angle bounded operators is intimately related to optimal regularity [7], and this concept is equivalent to optimal approximability, the class of operators (1.2) might perhaps be the largest one admiting optimal order-regularity a posteriori and a priori error estimates [34, 35]. Monotonicity alone is certainly not enough [38]; see also Section 3.4.

The key novel ingredient of our approach to a posteriori error analysis is a higher order *reconstruction* $\widehat{U}$, of degree $q + 1$, which yields the differential equation

$$\widehat{U}' + \Pi\mathfrak{F}(U) = F. \tag{1.5}$$

Here $\widehat{U}$ is a suitable *continuous* interpolant of the discontinuous discrete solution U, constructed elementwise, Π is an operator into the space of discontinuous piecewise polynomials $\mathcal{V}_k(q)$ of degree $\le q$, and F is an approximation of f within $\mathcal{V}_k(q)$. Expression (1.5) extends to $q > 0$ the pointwise representations of [33, 34, 35] for $q = 0$. In these works $\widehat{U}$ is the natural piecewise linear interpolant of the piecewise constant backward Euler approximation U. Rewriting (1.5) in the form

$$\widehat{U}' + \mathfrak{F}(U) = F + \mathfrak{F}(U) - \Pi\mathfrak{F}(U) =: \mathfrak{R} \tag{1.6}$$

reveals the fundamental principle behind our a posteriori error analysis: the *residual* $\mathfrak{R}$ measures the amount by which the pair $(\widehat{U}, U)$ misses to be a solution of (1.1). Therefore, *stability* of the continuous problem (1.1) dictates error estimates in terms of $\mathfrak{R}$. Regarding $(\widehat{U}, U)$ as a *relaxed* solution is a natural concept developed in [33] in Banach spaces for $q = 0$. Higher order reconstruction is also crucial for conservative schemes such as the Crank-Nicolson method [1]. Theories in both [1, 33] differ from that herein.

It is interesting to note the relation between our approach and the technique of Zadunaisky [45] for error control of ODEs. The idea in [45] is to construct a high degree polynomial by interpolating the approximate values on several consecutive time intervals, along with a perturbed ODE satisfied by such polynomial. Applying the same numerical scheme to this auxiliary ODE gives a heuristic estimate of the error, which sometimes can be made rigorous provided the continuous and discrete solutions, u and U, of (1.1) are sufficiently close. We refer to Skeel [39], who describes variations of [45] such as the popular defect-correction methods, and mentions justifications based on asymptotics. Since our focus is the derivation of optimal a posteriori estimates, the reconstruction $\widehat{U}$ and corresponding perturbed ODE (1.5) are dictated by the following fundamental principles: optimal approximability, global continuity, and elementwise construction. Our estimates do not require fine partitions on $[0, T]$ and do not need $\widehat{U}$ explicitly if $\mathfrak{F}$ is linear.

We next recall the two time discretizations dG(q) and IRK-R(q) we are interested in. Let $0 = t_0 < t_1 < \cdots < t_N = T$ be a partition $\mathfrak{P}$ of $[0, T]$, $I_n := (t_n, t_{n+1}]$, and $k_n := t_{n+1} - t_n$ be the variable time-step. We denote by $\mathbb{P}(q)$ the space of polynomials of degree $\le q$, and by $\mathcal{V}_k(q)$ the space of discontinuous piecewise polynomials of degree $\le q$ over $\mathfrak{P}$ with values in $D(\mathfrak{F})$: hence $g \in \mathcal{V}_k(q)$ reads

$$g|_{I_n}(t) = \sum_{j=0}^{q} t^j w_j \qquad (w_j \in D(\mathfrak{F}) \quad 0 \le j \le q).$$

The *discontinuous Galerkin method* dG(q) of order $q \ge 0$ is defined as follows [3, 11, 12, 10, 13, 14, 15, 16, 17, 18, 21, 22, 23, 24, 25, 26, 29, 40, 41, 42]: given $U_0 := u_0$, seek $U \in \mathcal{V}_k(q)$ such that

$$\int_{I_n} \Big(\langle U', v \rangle + \langle \mathfrak{F}(U), v \rangle \Big)\, dt + \langle U_n^+ - U_n, v_n^+ \rangle = \int_{I_n} \langle f, v \rangle dt\,, \quad \forall v \in \mathcal{V}_k(q) \tag{1.7}$$

for $0 \le n \le N-1$; hereafter $v_n := v(t_n)$, $v_n^+ := \lim_{s \downarrow 0} v(t_n + s)$. We consider also the corresponding Galerkin method with quadrature at the Radau points: given $V_0 := u_0$, find $V \in \mathcal{V}_k(q)$ such that

$$\int_{I_n} \Big(\langle V', v \rangle + \langle \mathcal{I}\,\mathfrak{F}(V), v \rangle \Big)\, dt + \langle V_n^+ - V_n, v_n^+ \rangle = \int_{I_n} \langle \mathcal{I} f, v \rangle dt\,, \quad \forall v \in \mathcal{V}_k(q) \tag{1.8}$$

for $0 \leq n \leq N-1$; hereafter $\mathcal{I}$ is the interpolation operator onto $\mathcal{V}_k(q)$ at the Radau points of each I_n. We show in Section 2 that $V(t^{n+1})$ coincides with V^{n+1}, the solution of the *Implicit Runge-Kutta Radau IIA method* with $q+1$ intermediate stages [9, 20]; we will thus refer to (1.8) as IRK-R(q) for short. As in [27], writing the solution U of dG(q) in terms of Radau polynomials turns out to be extremely useful. This is what establishes the connection between dG(q) and IRK-R(q), leads to (1.8) and thus to (1.5) (see Section 2).

The paper is organized as follows. We discuss in Section 2 the *reconstruction* of either U or V above, along with the crucial pointwise representation (1.5), after showing the relation between (1.8) and IRK-R(q). In Section 3 we study *linear* operators $\mathfrak{F}$, for which $\Pi\mathfrak{F} = \mathfrak{F}$. We first examine sectorial operators in Subsection 3.1 and apply our results to convection-diffusion problems in Subsection 3.2. The estimators are of optimal order-regularity, possess the extremely simple form $\sum_{n=0}^{N-1} k_n \|U_n^+ - U_n\|^2$ just involving jumps, and have absolute and explicit stability constants as in [34, 35]. Compared with [13], our estimators provide additional control from above and below of the full energy norm at all times. We analyze monotone operators in Subsection 3.3 and apply our results to the wave equation in Subsection 3.4. The estimates are of optimal order but require higher regularity, as expected for hyperbolic problems [25]. We deal with nonlinear γ^2-*angle bounded* operators (1.2) in Section 4. We consider subgradient operators in Subsection 4.1, such as the p-Laplacian, and Lipschitz operators in Subsection 4.2. We finally derive a *conditional* a posteriori error estimate in Section 5 for the minimal surface operator, for which the condition is also a posteriori and thus verifiable. Conditional estimates are somehow natural but rare in the literature. We refer to Fierro and Veeser [19] for (elliptic) problems of prescribed mean curvature and Lakkis and Nochetto [28] for the (parabolic) mean curvature flow of graphs, both for polynomial degree $q = 1$.

2. Reconstruction

In this section we derive the representation formula (1.6) for both dG(q) and IRK-R(q), namely,

$$\widehat{U}' + \mathfrak{F}(U) = F + \mathfrak{F}(U) - \Pi\mathfrak{F}(U), \tag{2.1}$$

where $F := \Pi f$, $\Pi = \mathcal{P}$ is the L^2-projection onto $\mathcal{V}_k(q)$ for dG(q) and $\Pi = \mathcal{I}$ is the Lagrange interpolation operator at the Radau points for IRK-R(q). These two methods are indeed closely related [14, 18, 27, 29].

2.1. **Reconstruction Operator.** Let $\{\tau_j\}_{j=1}^{q+1}$ be the Radau points in $[0,1]$. Then $0 < \tau_1 < \cdots < \tau_{q+1} = 1$ and for appropriate weights $\{w_j\}_{j=1}^{q+1}$ the Radau integration rule on $[0,1]$

$$\int_0^1 g(\tau)d\tau \cong \sum_{j=1}^{q+1} w_j g(\tau_j) \tag{2.2}$$

is exact for all polynomials of degree $\le 2q$. Let $\{\ell_j\}_{j=1}^{q+1} \subset \mathbb{P}(q)$ and $\{\widehat{\ell}_j\}_{j=0}^{q+1} \subset \mathbb{P}(q+1)$ be the Lagrange polynomials associated with either $\{\tau_j\}_{j=1}^{q+1}$ or $\{\tau_j\}_{j=0}^{q+1}$ with $\tau_0 = 0$. The corresponding Radau points in $\bar{I}_n$ are denoted by $t_{n,j}$, the Lagrange polynomials by $\ell_{n,j}$, $\widehat{\ell}_{n,j}$, and they satisfy

$$\begin{aligned} &t_{n,j} = t_n + \tau_j k_n \quad j = 0, \dots, q+1 \qquad (t_{n,0} = t_n,\ t_{n,q+1} = t_{n+1}) \\ &\ell_{n,j}(t) = \ell_j(\tau),\ \widehat{\ell}_{n,j}(t) = \widehat{\ell}_j(\tau), \qquad t = t_n + \tau k_n. \end{aligned} \tag{2.3}$$

The quadrature (2.2) induces a similar formula in I_n with nodes $\{t_{n,j}\}_{j=1}^{q+1}$ and weights $w_{n,j} = k_n w_j$. Moreover, let the interpolation operator $\mathcal{I} : C(0,T;D(\mathfrak{F})) \to \mathcal{V}_k(q)$ be

$$\mathcal{I}\ v|_{I_n}(t) := \sum_{j=1}^{q+1} \ell_{n,j}(t) v(t_{n,j}). \tag{2.4}$$

Consequently, if $V(t)$ is a polynomial in $\mathcal{V}_k(q)$, then $V(t) = \mathcal{I}\,V(t)$ for all t.

The *reconstruction operator* $\widehat{\mathcal{I}} : \mathcal{V}_k(q) \to \mathcal{V}_k(q+1)$ is now defined as follows: $\widehat{U} := \widehat{\mathcal{I}}\,U \in \mathcal{V}_k(q+1)$ satisfies in I_n

$$\begin{aligned} &\widehat{U}_n^+ = U_n\,. \\ &\int_{I_n} \langle \widehat{U}', v\rangle dt = \int_{I_n} \langle U', v\rangle dt + \langle U_n^+ - U_n, v_n^+\rangle, \quad \forall v \in \mathcal{V}_k(q)\,. \end{aligned} \tag{2.5}$$

In the sequel, we show that $\widehat{U}$ is well defined and exhibits some useful properties.

Lemma 2.1 (Reconstruction)**.** *The function $\widehat{U}$ is uniquely defined by (2.5), is globally continuous, and satisfies*

$$\widehat{U}(t_{n,j}) = U(t_{n,j}), \quad j = 0, \dots, q+1 \qquad (U(t_{n,0}) = U_n).$$

Proof. Integrating (2.5) by parts we get

$$\begin{aligned} -\int_{I_n} \langle \widehat{U}, v'\rangle dt + \langle \widehat{U}_{n+1}^-, v_{n+1}^-\rangle - \langle \widehat{U}_n^+, v_n^+\rangle \\ = -\int_{I_n} \langle U, v'\rangle dt + \langle U_{n+1}^-, v_{n+1}^-\rangle - \langle U_n, v_n^+\rangle. \end{aligned} \tag{2.6}$$

Since $\widehat{U}_n^+ = U_n$, selecting v constant in time we get that $\widehat{U}_{n+1}^- = U_{n+1}^- = U_{n+1}$. Since $t_{n,q+1} = t_{n+1}$ and $\widehat{U}_{n,q+1} = U_{n,q+1}$, using the exactness of the Radau integration rule (2.2) in I_n, (2.6) can be written as

$$\sum_{i=1}^{q} w_{n,i} \langle \widehat{U}_{n,i}, v'_{n,i}\rangle = \sum_{i=1}^{q} w_{n,i} \langle U_{n,i}, v'_{n,i}\rangle \quad (w_{n,i} = k_n w_i)\,.$$

Since v is arbitrary in $\mathcal{V}_k(q)$, we obtain $\widehat{U}_{n,j} = U_{n,j}$ for $1 \le j \le q$. This completes the proof. $\square$

A consequence of the fact that $\widehat{U}$ interpolates U at the Radau points is the following crucial properties for the estimates to follow.

Lemma 2.2 (Properties of $\widehat{U}$). *The following representation of $\widehat{U}-U$ is valid*

$$(\widehat{U}-U)|_{I_n}(t)=\widehat{\ell}_{n,0}(t)(U_n-U_n^+), \qquad \forall\, t\in I_n. \tag{2.7}$$

In addition, if

$$\alpha_p := \Big(\int_0^1 \big|\widehat{\ell}_0(\tau)\big|^p d\tau\Big)^{1/p}, \qquad \forall\, 1\le p\le\infty, \tag{2.8}$$

then for any semi-norm $\|\cdot\|$ in $\mathcal{H}$

$$\Big(\int_{I_n}\|\widehat{U}-U\|^p dt\Big)^{1/p}=\alpha_p\, k_n^{1/p}\|U_n^+-U_n\|, \qquad \forall\, 1\le p\le\infty. \tag{2.9}$$

Proof. Note that $\widehat{U}-U$ in I_n is a polynomial of degree $q+1$ which, in view of Lemma 2.1, vanishes at the Radau points $\{t_{n,j}\}_{j=1}^{q+1}$. Since $(\widehat{U}-U)(t_{n,0})=U_n-U_n^+$, we readily deduce (2.7). The identity

$$\int_{I_n}\big|\widehat{\ell}_{n,0}(t)\big|^p dt = k_n\int_0^1\big|\widehat{\ell}_0(\tau)\big|^p d\tau = k_n\alpha_p^p,$$

obviously implies (2.9). The proof is thus complete. □

2.2. **Discontinuous Galerkin Methods.** In view of (2.5), we can rewrite (1.7) as

$$\int_{I_n}\Big(\langle \widehat{U}', v\rangle\ +\langle \mathfrak{F}(U), v\rangle\Big)dt=\int_{I_n}\langle f, v\rangle dt\,, \quad \forall v\in\mathcal{V}_k(q)\,. \tag{2.10}$$

If $\mathcal{P}$ is the L^2-projection onto $\mathcal{V}_k(q)$, then (2.10) readily implies

$$\widehat{U}'+\mathcal{P}\,\mathfrak{F}(U)=F\,, \tag{2.11}$$

with $F:=\mathcal{P}\,f$. For piecewise constant solutions U, that is $q=0$, we have $\mathcal{P}\,\mathfrak{F}(U)=\mathfrak{F}(U)$. An expression similar to (2.11) was first used in [35] for subgradient and γ^2-angle bounded operators, and later extended in [33] to accretive operators in Banach spaces for $q=0$ and in [1] to the Crank-Nicolson method in Hilbert spaces.

2.3. **Runge-Kutta-Radau Methods.** We now consider the Implicit Runge-Kutta Radau IIA method with $q+1$ intermediate stages $\{V_{n,j}\}_{j=1}^{q+1}$ [9, 20]. It is known that the coefficients of this RK method are [20]

$$a_{i,j}=\int_0^{\tau_i}\ell_j(\tau)d\tau, \quad b_i=\int_0^1\ell_i(\tau)d\tau\ (=a_{q+1,i}), \qquad \forall\, 1\le i,j\le q+1, \tag{2.12}$$

and that the following implicit relation for $\{V_{n,j}\}_{j=1}^{q+1}$ holds

$$V_{n,i}-V_n+k_n\sum_{j=1}^{q+1}a_{i,j}\big(\mathfrak{F}(V_{n,j})-f_{n,j}\big)=0, \qquad \forall\, 1\le i\le q+1, \tag{2.13}$$

where $f_{n,j} := f(t_{n,j})$ and $V_{n+1} := V_{n,q+1}$. It is instructive to see the connection between (2.13) and its Galerkin counterpart (1.8), by using the interpolant $\widehat{V} = \widehat{\mathcal{I}}\, V$. In fact, next lemma links two well known facts: The connection of collocation and RK Radau methods, [20], and the connection of (1.8) to the RK Radau methods.

Lemma 2.3 (Equivalence between (1.8) and (2.13)). *Formulations* (1.8) *and* (2.13) *of IRK-R(q) are equivalent and, in fact, they are a collocation method for $\widehat{V}$ in each interval I_n with starting value $\widehat{V}_{n,0} = V_n$, namely*

$$\widehat{V}'_{n,i} + \mathfrak{F}(\widehat{V}_{n,i}) = f_{n,i}, \quad \forall\ 1 \le i \le q+1. \tag{2.14}$$

Proof. The exactness of the Radau quadrature for polynomials of degree $\le 2q$ on I_n implies that

$$\langle \mathcal{I}\,\mathfrak{F}(V) - \mathcal{I}\, f, v\rangle = \sum_{j=1}^{q+1} w_{n,j} \langle \mathfrak{F}(V_{n,j}) - f_{n,j}, v(t_{n,j})\rangle, \quad \forall v \in \mathcal{V}_k(q). \tag{2.15}$$

Therefore, (1.8) yields the expression

$$\int_{I_n} \langle V', v\rangle dt + \langle V_n^+ - V_n, v_n^+\rangle + \sum_{j=1}^{q+1} w_{n,j} \langle \mathfrak{F}(V_{n,j}) - f_{n,j}, v(t_{n,j})\rangle \ = 0\,, \quad \forall v \in \mathcal{V}_k(q),$$

or, with the help of (2.5), the simpler expression

$$\int_{I_n} \langle \widehat{V}', v\rangle dt + \sum_{j=1}^{q+1} w_{n,j} \langle \mathfrak{F}(V_{n,j}) - f_{n,j}, v(t_{n,j})\rangle = 0, \quad \forall v \in \mathcal{V}_k(q). \tag{2.16}$$

Since $\widehat{V}' \in \mathcal{V}_k(q)$ on I_n, taking $v = \ell_{n,i}$ for $1 \le i \le q+1$, and making use again of the Radau quadrature, we end up with (2.14). Consequently

$$\widehat{V}'(t) = \sum_{j=1}^{q+1} \big(f_{n,j} - \mathfrak{F}(V_{n,j})\big) \ell_{n,j}(t), \qquad \forall\ t \in I_n,$$

whence

$$\begin{aligned} V_{n,i} - V_n &= \int_{t_n}^{t_{n,i}} \widehat{V}'(t) dt \\ &= \sum_{j=1}^{q+1} \big(f_{n,j} - \mathfrak{F}(V_{n,j})\big) \int_{t_n}^{t_{n,i}} \ell_{n,j}(t) dt = k_n \sum_{j=1}^{q+1} a_{i,j} \big(f_{n,j} - \mathfrak{F}(V_{n,j})\big), \end{aligned}$$

which is (2.13). This completes the proof. ∎

Expression (2.16) also reads

$$\int_{I_n} \Big(\langle \widehat{V}', v\rangle + \langle \mathcal{I}\,\mathfrak{F}(V), v\rangle\Big) dt = \int_{I_n} \langle \mathcal{I}\, f, v\rangle dt, \quad \forall v \in \mathcal{V}_k(q),$$

or equivalently

$$\widehat{V}' + \mathcal{I}\,\mathfrak{F}(V) = F\,, \tag{2.17}$$

with $F := \mathcal{I}\,f$. A comparison of (2.17) with (2.11) leads to the interesting conclusion that the pointwise representations of dG(q) and IRK-R(q) differ only in the form of the operator acting on $\mathfrak{F}$ and f. This will be instrumental below.

3. A Posteriori Error Estimates for Linear Operators

In this section we assume that $\mathfrak{F} : D(\mathfrak{F}) \to \mathcal{H}$ is *linear*, whence $\Pi\mathfrak{F}(U) = \mathfrak{F}(U)$ and $F = \Pi f$ for either dG(q) or IRK-R(q), and (2.1) becomes

$$\widehat{U}' + \mathfrak{F}(U) = F, \tag{3.1}$$

as in [33, 34, 35]. In view of (3.1), we now examine both methods at once but distinguish between sectorial and monotone operators.

3.1. Sectorial Operators.

Let $\mathfrak{F} : D(\mathfrak{F}) \to \mathcal{H}$ be linear and monotone (see (1.3)). We define the *energy* semi-norm associated with $\mathfrak{F}$ by

$$\|v\| := \langle \mathfrak{F}(v), v\rangle^{\frac{1}{2}}, \qquad \forall\, v \in D(\mathfrak{F}), \tag{3.2}$$

and $\mathcal{V} := \{v \in \mathcal{H} : \|v\| < \infty\}$. In addition, we assume that $\mathfrak{F}$ satisfies the *strong sector* condition

$$|\langle \mathfrak{F}(v), w\rangle|^2 \le 4\gamma^2 \|v\|^2\, \|w\|^2, \qquad \forall\, v, w \in D(\mathfrak{F}). \tag{3.3}$$

This implies that $\mathfrak{F}$ is continuous and $\|\mathfrak{F}(v)\|_\star := \sup_{w\in D(\mathfrak{F})} \frac{\langle \mathfrak{F}(v), w\rangle}{\|w\|}$ satisfies

$$\frac{1}{4\gamma^2}\|\mathfrak{F}(v)\|_\star^2 \le \|v\|^2 \le \|\mathfrak{F}(v)\|_\star^2, \qquad \forall\, v \in D(\mathfrak{F}); \tag{3.4}$$

hence $\gamma \ge \frac{1}{2}$. Condition (3.3) is equivalent to the following inequality for the skew-symmetric part of the operator [6, Proposition 1]

$$\big|\langle \mathfrak{F}(v), w\rangle - \langle \mathfrak{F}(w), v\rangle\big| \le 2\mu \|v\| \|w\|, \qquad \forall\, v, w \in D(\mathfrak{F}), \tag{3.5}$$

with $\gamma^2 = (\mu^2 + 1)/4$; note that $\mu = 0$ and thus $\gamma = \frac{1}{2}$ if $\mathfrak{F}$ is symmetric.

Lemma 3.1 (Linear γ^2-Angle Bounded Operators)**.** *The strong sector condition* (3.3) *is equivalent to the γ^2-angle bounded condition* (1.2)*, namely,*

$$\langle \mathfrak{F}(v - w), w - z\rangle \le \gamma^2 \langle \mathfrak{F}(v - z), v - z\rangle \quad \forall\, v, w, z \in D(\mathfrak{F}). \tag{3.6}$$

Proof. We simply set $\tilde{v} = v - z$ and $\tilde{w} = w - z$ in (3.6) to get the equivalent formulation (we omit the tildes)

$$\langle \mathfrak{F}(v), w\rangle \le \gamma^2 \langle \mathfrak{F}(v), v\rangle + \langle \mathfrak{F}(w), w\rangle, \quad \forall\, v, w \in D(\mathfrak{F}). \tag{3.7}$$

Then replace v by λv with $\lambda \in \mathbb{R}$, and argue with the resulting quadratic inequality in λ to realize that (3.3) and (3.7) are equivalent. $\square$

Lemma 3.2 (Coercivity). *If $\mathfrak{F}$ satisfies* (3.3), *then for all $v, w, z \in D(\mathfrak{F})$*

$$\langle \mathfrak{F}(v-w), w-z\rangle \le 2\gamma^2 \langle \mathfrak{F}(v-z), v-z\rangle - \frac{1}{2}\max\big(\|v-w\|^2, \|z-w\|^2\big). \tag{3.8}$$

Proof. Elementary calculations based on (3.4) yield

$$\begin{aligned}\langle \mathfrak{F}(v-w), w-z\rangle &= \langle \mathfrak{F}(v-w), w-v\rangle + \langle \mathfrak{F}(v-w), v-z\rangle \\ &\le -\|v-w\|^2 + 2\gamma\|v-w\|\|v-z\| \\ &\le -\frac{1}{2}\|v-w\|^2 + 2\gamma^2\langle \mathfrak{F}(v-z), v-z\rangle.\end{aligned}$$

On the other hand, a symmetric argument implies

$$\begin{aligned}\langle \mathfrak{F}(v-w), w-z\rangle &= \langle \mathfrak{F}(z-w), w-z\rangle + \langle \mathfrak{F}(v-z), w-z\rangle \\ &\le -\frac{1}{2}\|w-z\|^2 + 2\gamma^2\langle \mathfrak{F}(v-z), v-z\rangle.\end{aligned}$$

Combining these two inequalities, we easily obtain (3.8). □

We are now ready to prove both upper and lower a posteriori error bounds. To this end, we first need to introduce the error measure $\mathfrak{E}$:

$$\mathfrak{E} := \left\{\max\Big(\max_{0\le t\le T} |(u-\widehat{U})(t)|^2, \frac{1}{2}\int_0^T \|u-\widehat{U}\|^2 dt, \frac{1}{2}\int_0^T \|u-U\|^2 dt\Big)\right\}^{1/2}. \tag{3.9}$$

Theorem 3.3 (Upper Bound). *If $u_0 \in \mathcal{V}$, then the following estimate is valid for sectorial operators $\mathfrak{F}$ and for both dG(q) and IRK-R(q) for any $q \ge 0$:*

$$\mathfrak{E}^2 \le 4\gamma^2\alpha_2^2 \sum_{n=0}^{N-1} k_n \|U_n^+ - U_n\|^2 + 2\int_0^T \|f-F\|_\star^2 dt. \tag{3.10}$$

Proof. Subtract (3.1) from (1.1) to obtain the error equation

$$\frac{d}{dt}(u-\widehat{U}) + \mathfrak{F}(u-U) = f - F.$$

We next multiply this equation by $u-\widehat{U}$ to see that

$$\frac{1}{2}\frac{d}{dt}|u-\widehat{U}|^2 - \langle \mathfrak{F}(U-u), u-\widehat{U}\rangle = \langle f-F, u-\widehat{U}\rangle,$$

whence, in view of (3.8), we deduce with $M = \max\big(\|u-U\|, \|u-\widehat{U}\|\big)$

$$\frac{1}{2}\frac{d}{dt}|u-\widehat{U}|^2 + \frac{1}{2}M^2 \le 2\gamma^2\|U-\widehat{U}\|^2 + M\|f-F\|_\star \le 2\gamma^2\|U-\widehat{U}\|^2 + \frac{M^2}{4} + \|f-F\|_\star^2.$$

The asserted estimate (3.10) follows from (2.9) after integration in time. □

Remark 3.4 (Energy Dissipation)**.** A striking property of (3.10) is that, except for data oscillation, the energy dissipation (or jump discontinuity) $\|U_n^+ - U_n\|$ is what controls the error. This estimate for dG(q) as well as for IRK-R(q) extends the estimates of Nochetto, Savaré and Verdi for the implicit Euler scheme ($q = 0$) to higher order ($q > 0$) without changing their structure [34, 35]. Similar estimates were obtained by Eriksson, Johnson and Larsson via duality, but without the coercivity components of $\mathfrak{E}$ [13].

Remark 3.5 (Stiff ODE)**.** This theory applies to *stiff* ODE systems and yields a posteriori estimates which are dimension independent. The nature of these estimates is different though from those in [16, 24] in that our results incorporate energy terms and the estimators accumulate in time in the L^2 norm instead of the L^∞ norm.

Remark 3.6 (Smooth Data A Priori Error Estimates)**.** We assert that the error estimates in Theorem 3.3 are of optimal order-regularity provided the initial data and forcing term are smooth. To see this, we consider $f = 0$ and recall the a priori estimate of [42, Theorem 12.1] extended to sectorial operators,

$$\int_0^T \|u - U\|^2 \, dt \le Ck^{2(q+1)} \int_0^T \|\partial_t^{q+1} u\|^2 \, dt, \tag{3.11}$$

where $k = \max_n k_n$ is the largest step-size. Applying inverse estimates, we get

$$\begin{aligned} k_n \|U_n^+ - U_n\|^2 &\le 2k_n \|U_n^+ - u(t_n)\|^2 + 2k_n \|U_n - u(t_n)\|^2 \\ &\le C \int_{t_n}^{t_{n+1}} \|U - \mathcal{I}\, u\|^2 dt + C \frac{k_n}{k_{n-1}} \int_{t_{n-1}}^{t_n} \|U - \mathcal{I}\, u\|^2 dt, \end{aligned}$$

whence, invoking (3.11) and interpolation theory for $\int_0^T \|u - \mathcal{I}\, u\|^2 dt$, we deduce

$$\sum_{n=0}^{N-1} k_n \|U_n^+ - U_n\|^2 \le Ck^{2(q+1)} \int_0^T \|\partial_t^{q+1} u\|^2 \, dt,$$

provided $k_n \le Ck_{n-1}$; the latter is a reasonable constraint between consecutive time-steps. Compared with the estimates of Eriksson, Johnson, and Larsson [11, 12, 10, 13], which require the regularity $\max_{0 \le t \le T} |\partial_t^{q+1} u|$, we observe that both

$$\max_{0 \le t \le T} |\partial_t^{q+1} u| \quad \text{and} \quad \Big(\int_0^T \|\partial_t^{q+1} u\|^2 \, dt \Big)^{1/2}$$

are bounded by the same constant depending on data. Therefore, in the linear case considered here, their control require the same regularity on the data (u_0, f) of problem (1.1).

Remark 3.7 (Comparison with Duality)**.** We now show the striking agreement between the stability constant γ^2 in Theorem 3.3 and the corresponding one of Eriksson, Johnson and Larsson [13] for analytic semigroups based on duality arguments for dG(q). The a posteriori error estimate shown in [13] has the form

$$\max_{1\le n\le N}|u(t_n)-U(t_n)| \le \; C_I C_S L_N \max_{0\le n\le N-1}\Big(|U_n^+-U_n| + \max_{t\in I_n} k_n|f-Pf|\Big), \tag{3.12}$$

where C_I is an interpolation constant, L_N grows logarithmicaly with respect to k_N and C_S is the stability constant of an homogeneous *backward dual* problem. A sharp bound for C_S can be found by a simple energy argument [42]. Upon changing variables $t\to T-t$, we consider the corresponding homogeneous forward problem

$$w_t+\mathfrak{F}^\star(w)=0, \quad w(0)=w^0, \tag{3.13}$$

where $\mathfrak{F}^\star$ is the adjoint of $\mathfrak{F}$. Then C_S is the constant of the *strong* stability estimate

$$|\mathfrak{F}^\star(w(t))| \le \frac{C_S}{t}\,|w^0|\,. \tag{3.14}$$

We assert that $C_S\approx\gamma$ for sectorial operators. To see this, we deal with $v=tw$ and the equation that it satisfies. Since $v_t+\mathfrak{F}(v)=w$, then

$$v_{tt}+\mathfrak{F}^\star(v_t)=w_t,$$

whence

$$\langle v_{tt},v_t\rangle+\langle\mathfrak{F}^\star(v_t),v_t\rangle=\langle w_t,v_t\rangle=-\langle\mathfrak{F}^\star(w),v_t\rangle\le 2\gamma\|w\|\|v_t\|\le\gamma^2\|w\|^2+\|v_t\|^2\,.$$

Since $\langle\mathfrak{F}^\star(v_t),v_t\rangle=\|v_t\|^2$, $v_t(0)=w(0)$, and $\int_0^t\|w\|^2\,ds\le\frac12|w^0|^2$, we thus obtain

$$|v_t|^2\le 2\gamma^2\int_0^t\|w\|^2\,dt+|w^0|^2\le(\gamma^2+1)|w^0|^2\,.$$

Moreover, the fact that $|w(t)|\le|w^0|$, in conjunction with $v_t=w+tw_t$, finally yields

$$t|\mathfrak{F}^\star(w)|=t|w_t|\le|w|+|v_t|\le\big((\gamma^2+1)^{1/2}+1\big)\,|w^0|.$$

Therefore (3.14) holds with

$$C_S=(\gamma^2+1)^{1/2}+1\,.$$

This shows that although the approach of this paper is based on simple energy arguments, it gives a posteriori error bounds that for linear sectorial operators compare remarkably well with the estimates based on duality techniques [13].

Theorem 3.8 (Lower Bound). *If $u_0\in\mathcal{V}$, then the following estimate is valid for sectorial operators $\mathfrak{F}$ and for both dG(q) and IRK-R(q) for any $q\ge0$:*

$$\alpha_2^2\sum_{n=0}^{N-1}k_n\|U_n^+-U_n\|^2\le 8\mathfrak{E}^2. \tag{3.15}$$

Proof. This is a trivial consequence of (2.9) because

$$\alpha_2^2 \sum_{n=0}^{N-1} k_n \|U_n^+ - U_n\|^2 = \int_0^T \|\widehat{U} - U\|^2 dt \\ \le 2 \int_0^T \Big(\|u - \widehat{U}\|^2 + \|u - U\|^2\Big) dt \le 8\mathfrak{E}^2. \qquad \square$$

Remark 3.9 (Dominant Term)**.** A simple by-product of (3.10) and the above proof is the following upper bound

$$\max_{0\le t\le T} |(u-\widehat{U})(t)|^2 \le 8\gamma^2 \max\Big(\int_0^T \|u-\widehat{U}\|^2, \int_0^T \|u-U\|^2\Big) + 2\int_0^T \|f-F\|_\star^2.$$

This shows that, up to data oscillation, the energy error $L^2(\mathcal{V})$ dominates the $L^\infty(\mathcal{H})$ error.

3.2. **Application: Convection-Diffusion Equation.** Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$ with any $d \ge 1$. Consider the initial boundary value problem

$$\begin{aligned} u_t - \epsilon^2\Delta u + \boldsymbol{b}\cdot\nabla u + cu &= f, && \text{in } \Omega\times[0,T],\\ u(\cdot,0) &= u_0, && \text{in } \Omega,\\ u &= 0, && \text{on } \partial\Omega\times[0,T], \end{aligned} \tag{3.16}$$

with $\mathcal{H} := L^2(\Omega)$ and norm $|\cdot|$. The coefficients $\boldsymbol{b}\in W^{1,\infty}(\Omega), c \in L^\infty(\Omega)$ satisfy

$$|\boldsymbol{b}(x)| \le b_0, \quad d(x) := -\tfrac12 \operatorname{div}\boldsymbol{b}(x) + c(x) \ge d_0^2 \ge 0, \quad \text{a.e. } x\in\Omega, \tag{3.17}$$

and $\epsilon^2 > 0$. The underlying elliptic operator $\mathfrak{F}(v) := -\epsilon^2\Delta v + \boldsymbol{b}\cdot\nabla v + cv$ has a domain $D(\mathfrak{F}) := \{v \in H_0^1(\Omega) : \Delta v \in \mathcal{H}\}$, and induces the energy norm in $\mathcal{V} := H_0^1(\Omega)$

$$\|v\|^2 = \langle \mathfrak{F}(v), v\rangle = \int_\Omega (\epsilon^2|\nabla v|^2 + d|v|^2)dx \qquad \forall\, v\in\mathcal{V}. \tag{3.18}$$

Let $p_D > 0$ be the constant of the Poincaré inequality $p_D|w| \le |\nabla w|$ for all $w \in \mathcal{V}$. The following result is well known [35, Lemma 5.1], but we prove it for completeness.

Lemma 3.10 ($\mathfrak{F}$ is Sectorial)**.** *The operator $\mathfrak{F}$ satisfies* (3.3) *with constant γ given by*

$$\gamma^2 = \frac14\Big(1 + \frac{b_0^2}{\epsilon^2(d_0^2+\epsilon^2p_D^2)}\Big). \tag{3.19}$$

Proof. This proof hinges on (3.5). Since

$$\|v\|^2 \ge \epsilon^2|\nabla v|^2, \qquad \|v\|^2 \ge \big(\epsilon^2p_D^2 + d_0^2\big)|v|^2 \qquad \forall\, v\in\mathcal{V},$$

the skew-symmetric part of $\mathfrak{F}$ satisfies for all $v, w \in \mathcal{V}$

$$\big|\langle\mathfrak{F}(v), w\rangle - \langle\mathfrak{F}(w), v\rangle\big| \le b_0|\nabla v||w| + b_0|\nabla w||v| \le \frac{2b_0}{\epsilon\big(\epsilon^2p_D^2+d_0^2\big)^{\frac12}}\|v\|\|w\|.$$

Consequently $\mu = b_0\epsilon^{-1}\big(\epsilon^2 p_D^2 + d_0^2\big)^{-\frac{1}{2}}$ and the expression for γ follows from $\gamma^2 = (1+\mu^2)/4$; see [6, Proposition 1]. This concludes the proof. $\square$

The following a posteriori error estimates is a simple consequence of Theorems 3.3 and 3.8.

Corollary 3.11 (Error Estimates for Convection-Diffusion Equations). *If $u_0 \in \mathcal{V}$, then the following estimates are valid for the convection-diffusion problem* (3.16), (3.17) *with γ given by* (3.19) *and for both dG(q) and IRK-R(q) for any $q \geq 0$:*

$$\frac{1}{8}\alpha_2^2 \sum_{n=0}^{N-1} k_n \|U_n^+ - U_n\|^2 \leq \mathfrak{E}^2 \leq 4\gamma^2\alpha_2^2 \sum_{n=0}^{N-1} k_n \|U_n^+ - U_n\|^2 + 2\int_0^T \|f-F\|_\star^2 dt.$$

3.3. **Monotone Operators.** We consider a monotone linear operator $\mathfrak{F}$, cf. (1.3). In the linear case (1.3) reduces to $\langle \mathfrak{F}(v), v\rangle \geq 0$ for all $v \in D(\mathfrak{F})$. This assumption is insufficient to guarantee optimal a priori error estimates [38]. The same happens with the a posteriori error analysis.

Theorem 3.12 (Error Estimates for Monotone Operators). *If $u_0 \in D(\mathfrak{F})$, then*

$$\max_{0\leq t\leq T} |u - \widehat{U}| \leq \alpha_1 \sum_{n=0}^{N-1} k_n |\mathfrak{F}(U_n^+ - U_n)| + \int_0^T |f - F| \; dt.$$

Proof. We repeat the argument of Theorem 3.3 except that we can no longer exploit coercivity. Since

$$\frac{1}{2}\frac{d}{dt}|u-\widehat{U}|^2 + \langle \mathfrak{F}(u-\widehat{U}), u-\widehat{U}\rangle = \langle \mathfrak{F}(U-\widehat{U}), u-\widehat{U}\rangle + \langle f-F, u-\widehat{U}\rangle$$

we deduce

$$\frac{d}{dt}|u-\widehat{U}|^2 \leq 2|u-\widehat{U}|\Big(|\mathfrak{F}(U-\widehat{U})| + |f-F|\Big).$$

We now invoke the inequalities

$$\frac{d}{dt}a(t)^2 \leq 2a(t)b(t) \quad \Rightarrow \quad \frac{d}{dt}a(t) \leq b(t) \quad \Rightarrow \quad \max_{0\leq t\leq T} a(t) \leq a(0) + \int_0^T b(t)dt.$$

The assertion finally follows from Lemma 2.2. $\square$

3.4. **Application: Wave Equation.** Let Ω be a bounded Lipschitz domain in $\mathbb{R}^d$ for any $d \geq 1$. We consider the initial boundary value problem:

$$\begin{aligned} &v_{tt} - \Delta v = g, && \text{in } \Omega\times(0,T),\\ &v(0) = v_0, v_t(0) = v_1, && \text{in } \Omega,\\ &v|_{\partial\Omega} = 0\,, && \text{on } \partial\Omega\times(0,T), \end{aligned} \tag{3.20}$$

with $D(-\Delta) = H^2(\Omega) \cap H_0^1(\Omega)$. To write (3.20) in the form $u_t + \mathfrak{F}(u) = f$, we set $u = (u_1, u_2)$ and reduce the order as follows:

$$u = \begin{bmatrix} v \\ v_t \end{bmatrix}, \quad \mathfrak{F}(u) = -\begin{bmatrix} u_2 \\ \Delta u_1 \end{bmatrix}, \quad f = \begin{bmatrix} 0 \\ g \end{bmatrix}.$$

Let $U = (U_1, U_2)$ be either the dG(q) or IRK-R(q) approximation of u, and let $G = \Pi g$. The next issue is to state the functional setting. We start with $\mathcal{H} := H_0^1(\Omega) \times L^2(\Omega)$ with scalar product $\langle v, w\rangle = \langle \nabla v_1, \nabla w_1\rangle + \langle v_2, w_2\rangle$ and corresponding norm $|\cdot|$; hence $\langle \mathfrak{F}(v), v\rangle = 0$ for all $v \in D(\mathfrak{F}) = D(-\Delta) \times L^2(\Omega)$. The error is then

$$\mathfrak{E} := \max_{0\le t\le T} \Big(\|\nabla(u_1 - \widehat{U}_1)\|^2_{L^2(\Omega)} + \|u_2 - \widehat{U}_2\|^2_{L^2(\Omega)}\Big)^{\frac{1}{2}}.$$

Corollary 3.13 (Energy Norm Estimate). *If $v_0 \in D(-\Delta)$ and $v_1 \in H_0^1(\Omega)$, then*

$$\mathfrak{E} \le \alpha_1 \sum_{n=0}^{N-1} k_n \Big(\|\Delta(U^+_{1,n} - U_{1,n})\|^2_{L^2(\Omega)} + \|\nabla(U^+_{2,n} - U_{2,n})\|^2_{L^2(\Omega)}\Big)^{\frac{1}{2}} + \int_0^T \|g - G\|_{L^2(\Omega)}.$$

This estimate provides an optimal order error bound at the expense of additional regularity. In order to reduce the regularity demands, we seek an alternative choice of $\mathcal{H}$ which leads to an estimate in a weaker norm but also with lower data regularity requirements.

Let $T := (-\Delta)^{-1}$ be the restriction to $L^2(\Omega)$ of the inverse Laplacian with zero Dirichlet condition. Let $\mathcal{H} := L^2(\Omega) \times H^{-1}(\Omega)$ with scalar product, [2]

$$\langle v, w\rangle := \langle v_1, w_1\rangle + \langle T v_2, w_2\rangle, \quad \forall\ v, w \in \mathcal{H};$$

hence $\langle \mathfrak{F}(v), v\rangle = -\langle v_2, v_1\rangle + \langle T(-\Delta)v_1, v_2\rangle = 0$ for all $w \in D(\mathfrak{F})$. The error is

$$\mathfrak{E} := \max_{0\le t\le T} \Big(\|u_1 - \widehat{U}_1\|^2_{L^2(\Omega)} + \|u_2 - \widehat{U}_2\|^2_{H^{-1}(\Omega)}\Big)^{\frac{1}{2}}.$$

Corollary 3.14 (Weak Estimate). *If $v_0 \in H_0^1(\Omega)$ and $v_1 \in L^2(\Omega)$, then*

$$\mathfrak{E} \le \alpha_1 \sum_{n=0}^{N-1} k_n \Big(\|\nabla(U^+_{1,n} - U_{1,n})\|^2_{L^2(\Omega)} + \|U^+_{2,n} - U_{2,n}\|^2_{L^2(\Omega)}\Big)^{\frac{1}{2}} + \int_0^T \|g - G\|_{H^{-1}(\Omega)} dt.$$

Remark 3.15 (Hyperbolic vs. Parabolic Character). A posteriori error estimates for the fully discrete discontinuous Galerkin method were proved in [25] for $q = 1$ via duality. In contrast to the parabolic case, the estimators were expressed in terms of discrete-time L^1 norms. This is due to the fact that *strong* stability estimates of the form (3.14) are not valid for the wave equation. Our estimators in this case are also expressed in terms of discrete-time L^1 norms. Compared to the parabolic case, the increased regularity required in the estimators appears also in [25]. This is expected for problems of non-parabolic character as the a priori results for the Schrödinger equation show [27]; see also [2]. Compare also with the a priori and a posteriori results for first order hyperbolic problems derived in [21, 22, 41].

4. A Posteriori Error Estimates for Nonlinear Operators

In this section we consider the nonlinear case. The notion of linear γ-angle bounded operators (3.6) extends naturally to nonlinear $\mathfrak{F}$ [6, 34, 35].

Definition 4.1 (Nonlinear γ^2-Angle Bounded Operators). *The (possibly) nonlinear operator $\mathfrak{F}$ is called γ^2-angle bounded if it satisfies for some $\gamma > 0$*

$$\langle \mathfrak{F}(v) - \mathfrak{F}(w), w - z\rangle \le \gamma^2 \langle \mathfrak{F}(v) - \mathfrak{F}(z), v - z\rangle \quad \forall\, v, w, z \in D(\mathfrak{F}). \tag{4.1}$$

We derive our results under the assumption that an amount of *coercivity* is inherited by (4.1). To this end we introduce the nonnegative quantity for $\eta \ge \gamma$:

$$\sigma_\eta(v,w,z) := \eta^2\langle \mathfrak{F}(v) - \mathfrak{F}(z), v - z\rangle - \langle \mathfrak{F}(v) - \mathfrak{F}(w), w - z\rangle \quad \forall\, v, w, z \in D(\mathfrak{F}), \tag{4.2}$$

and assume the following coercivity condition.

Definition 4.2 ($p-$Coercivity). *Let $D(\mathfrak{F})$ be equiped with a lower semicontinuous (l.s.c) seminorm $\|\cdot\|$. The operator $\mathfrak{F}$ is called* p-coercive *if for some $p \ge 2$ and $\eta \ge \gamma$ there exists $\delta > 0$, depending on η, such that*

$$\sigma_\eta(v,w,z) \ge \frac{\delta}{p} \max\Big(\|v - w\|^p, \|w - z\|^p\Big), \quad \forall\, v, w, z \in D(\mathfrak{F}). \tag{4.3}$$

This notion of coercivity is a natural extension of the linear case (3.8) in Lemma 3.2, where $p = 2$, $\eta = \sqrt{2}\,\gamma$, and $\delta = 1$; likewise, we set $\mathcal{V} := \{v \in \mathcal{H} : \|v\| < \infty\}$. Note that even in the linear case we need $\eta > \gamma$ to gain the above coercivity. Examples of (4.3) are given in Subsections 4.1 and 4.2.

We are now ready to prove a posteriori error bounds in the nonlinear case similar to Theorem 3.3. By analogy with (3.9), we introduce the following error measure

$$\mathfrak{E} := \left\{ \max\Big(\max_{0\le t\le T} |(u - \widehat{U})(t)|^2, \frac{\delta}{p}\int_0^T \|u - \widehat{U}\|^p dt, \frac{\delta}{p}\int_0^T \|u - U\|^p dt \Big) \right\}^{1/2}.$$

Theorem 4.3 (Error Bound for Nonlinear Operators). *Let $\mathfrak{F}$ be γ^2-angle bounded and* p-coercive *with respect to the seminorm $\|\cdot\|$ for $\eta \ge \gamma$. We denote by $\|\cdot\|_\star$ the dual of the seminorm $\|\cdot\|$. If $p^\star = p/(p-1)$, then the following estimate is valid for both dG(q) and IRK-R(q) for any $q \ge 0$:*

$$\begin{aligned} \mathfrak{E}^2 \le\ & 2\eta^2 \int_0^T \langle \mathfrak{F}(\widehat{U}) - \mathfrak{F}(U), \widehat{U} - U\rangle dt \\ & + \frac{2}{p^\star}\Big(\frac{2}{\delta}\Big)^{p^\star/p} \int_0^T \Big(\|\mathfrak{F}(U) - \Pi\mathfrak{F}(U)\|_\star + \|f - F\|_\star \Big)^{p^\star} dt. \end{aligned} \tag{4.4}$$

Proof. Subtract (2.1) from (1.1), to obtain the error equation

$$\frac{d}{dt}(u - \widehat{U}) + \mathfrak{F}(u) - \mathfrak{F}(U) = (\Pi\mathfrak{F}(U) - \mathfrak{F}(U)) + (f - F) \qquad \forall\, 0 \le t \le T.$$

Testing with $u-\widehat{U}$ and using (4.3), we see that

$$\frac{1}{2}\frac{d}{dt}|u-\widehat{U}|^2+\frac{\delta}{p}\max\Big(\|u-U\|^p\,,\;\|u-\widehat{U}\|^p\Big)$$
$$\leq \eta^2\langle\mathfrak{F}(\widehat{U})-\mathfrak{F}(U),\widehat{U}-U\rangle+\langle(\Pi\mathfrak{F}(U)-\mathfrak{F}(U))+(f-F),u-\widehat{U}\rangle\,.$$

Next, we use Young's inequality, $ab\leq\frac{\varepsilon}{p}a^p+\frac{\varepsilon^{-p^\star/p}}{p^\star}b^{p^\star}$, with $\varepsilon=\frac{\delta}{2}$ to arrive at

$$|\langle(\Pi\mathfrak{F}(U)-\mathfrak{F}(U))+(f-F),u-\widehat{U}\rangle|\leq\frac{\delta}{2p}\|u-\widehat{U}\|^p$$
$$+\frac{1}{p^\star}\Big(\frac{\delta}{2}\Big)^{-p^\star/p}\Big(\|\mathfrak{F}(U)-\Pi\mathfrak{F}(U)\|_\star+\|f-F\|_\star\Big)^{p^\star},$$

whence (4.4) follows immediately upon integration. □

Remark 4.4 (Nonlinearity)**.** In contrast with Theorem 3.3, we notice two nonlinear effects. First, we can no longer express $\langle\mathfrak{F}(\widehat{U})-\mathfrak{F}(U),\widehat{U}-U\rangle$ in terms of the jump residual $\|U_n^+-U_n\|^2$. Secondly, the new estimator $\|\mathfrak{F}(U)-\Pi\mathfrak{F}(U)\|_\star$ accounts for the approximation of $\mathfrak{F}(U)\notin\mathcal{V}_k(q)$ by piecewise polynomials of degree $\leq q$.

4.1. **Application: Subgradient Operators.** A subclass of γ^2-angle bounded operators are the *subgradient* operators, which are characterized by the existence of a proper lower semicontinuous convex function

$$\phi:\mathcal{H}\to(-\infty,+\infty],\quad D(\phi)=\{v\in\mathcal{H}:\phi(v)<\infty\},$$

such that $\mathfrak{F}=\partial\phi$ is the subgradient of ϕ. This means that $\mathfrak{F}$ and ϕ satisfy

$$\phi(w)-\phi(v)-\langle\mathfrak{F}(v),w-v\rangle\geq 0,\quad\forall\ v\in D(\mathfrak{F}),w\in D(\phi).\tag{4.5}$$

Moreover, the following well-known characterization has been used in [34, 35] to derive a posteriori error estimates for the implicit Euler method.

Lemma 4.5 (Subgradient Operators are 1-Angle Bounded)**.** *All subgradient operators $\mathfrak{F}$ are 1-angle bounded. If, in addition, ϕ is Frechet differentiable and $\mathfrak{F}$ satisfies*

$$\|v-w\|^p\leq\langle\mathfrak{F}(v)-\mathfrak{F}(w),v-w\rangle,\quad\forall\ v,w\in D(\mathfrak{F})\tag{4.6}$$

then $\mathfrak{F}$ is p-coercive with $\delta=1$; moreover, it holds

$$\sigma_1(v,w,z)\geq\frac{1}{p}\Big(\|v-w\|^p\;+\;\|w-z\|^p\Big),\quad\forall\ v,w,z\in D(\mathfrak{F}).\tag{4.7}$$

Proof. Since

$$\langle\mathfrak{F}(v)-\mathfrak{F}(w),w-z\rangle=\langle\mathfrak{F}(v)-\mathfrak{F}(z),v-z\rangle+\langle\mathfrak{F}(v),w-v\rangle+\langle\mathfrak{F}(w),z-w\rangle+\langle\mathfrak{F}(z),v-z\rangle,$$

in view of (4.5), we first see that $\mathfrak{F}$ is 1-angle bounded, i.e.

$$\langle\mathfrak{F}(v)-\mathfrak{F}(w),w-z\rangle\leq\langle\mathfrak{F}(v)-\mathfrak{F}(z),v-z\rangle.$$

On the other hand, using the mean value theorem in (4.5) and (4.6), we deduce

$$\begin{aligned}\phi(w) - \phi(v) - \langle \mathfrak{F}(v), w - v\rangle &= \int_0^1 \langle \mathfrak{F}(w_s) - \mathfrak{F}(v), w_s - v\rangle \frac{ds}{s} \\ &\geq \int_0^1 \|w_s - v\|^p \frac{ds}{s} = \|w - v\|^p \int_0^1 s^{p-1} ds = \frac{1}{p}\|w - v\|^p,\end{aligned}$$

where $w_s = sw + (1-s)v$. This implies (4.7) and concludes the proof. □

In view of (4.7), we now define the error measure to be

$$\mathfrak{E} := \left\{ \max\Big(\max_{0\leq t\leq T} |(u - \widehat{U})(t)|^2, \frac{1}{p}\int_0^T \|u - \widehat{U}\|^p dt + \frac{2}{p}\int_0^T \|u - U\|^p dt\Big) \right\}^{1/2}.$$

Corollary 4.6 (Error Estimates for Subgradient Operators)**.** *Let $\mathfrak{F} = \nabla\phi$ be a subgradient operator with ϕ Frechet differentiable and satisfying* (4.6)*. Then the following error estimate is valid for both dG(q) and IRK-R(q) for all $q \geq 0$:*

$$\mathfrak{E}^2 \leq 2\int_0^T \Big(\|\mathfrak{F}(\widehat{U}) - \mathfrak{F}(U)\|_\star^{p^\star} + \frac{2^{p^\star/p}}{p^\star}\big(\|\mathfrak{F}(U) - \Pi\mathfrak{F}(U)\|_\star + \|f - F\|_\star\big)^{p^\star}\Big) dt.$$

Proof. Since $\eta = \delta = 1$ from Lemma 4.5, it suffices to show

$$\langle \mathfrak{F}(\widehat{U}) - \mathfrak{F}(U), \widehat{U} - U\rangle \leq \|\mathfrak{F}(\widehat{U}) - \mathfrak{F}(U)\|_\star^{p^\star}.$$

This follows from $\|\widehat{U} - U\|^p \leq \|\mathfrak{F}(\widehat{U}) - \mathfrak{F}(U)\|_\star^{p^\star}$ which, in light of (4.6), results from

$$\|\widehat{U} - U\|^p \leq \langle \mathfrak{F}(\widehat{U}) - \mathfrak{F}(U), \widehat{U} - U\rangle \leq \frac{1}{p^\star}\|\mathfrak{F}(\widehat{U}) - \mathfrak{F}(U)\|_\star^{p^\star} + \frac{1}{p}\|\widehat{U} - U\|^p.$$

Consequently, inserting this bound into (4.4) we obtain the asserted estimate. □

Remark 4.7 ($p-$Laplacian Operator)**.** Given $p \geq 2$, let $\mathfrak{F}(v) := -\operatorname{div}(|\nabla v|^{p-2}\nabla v)$ be the p-Laplacian operator [5, 30, 35]. This operator is the subgradient of $\phi(v) := \frac{1}{p}|\nabla v\|_{L^p(\Omega)}^p$ in $\mathcal{H} = L^2(\Omega)$, and has the following coercivity property [35]

$$\langle \mathfrak{F}(u) - \mathfrak{F}(v), u - v\rangle \geq \Lambda_p \|\nabla(u - v)\|^p, \tag{4.8}$$

for a suitable constant $\Lambda_p > 0$. Hence $\mathfrak{F}$ is p-coercive with respect to the norm $\|v\| = \Lambda_p^{1/p}\|\nabla v\|_{L^p(\Omega)}$, and Corollary 4.6 applies with dual norm $\|v\|_\star := \Lambda_p^{-1/p}\|v\|_{W^{-1}_{p^\star}(\Omega)}$.

Remark 4.8 (Porous Medium Operator)**.** Given $p \geq 2$, let $\mathfrak{F}(v) := -\Delta(|v|^{p-2}v)$ be the porous medium operator. This prototype of *degenerate* operator is the subgradient of $\phi(v) := \frac{1}{p}\|v\|_{L^p(\Omega)}$ in $\mathcal{H} = H^{-1}(\Omega)$, and is p-coercive in $L^p(\Omega)$ [35], i.e.

$$\langle \mathfrak{F}(v) - \mathfrak{F}(w), v - w\rangle \geq \lambda_p \|v - w\|_{L^p(\Omega)}^p,$$

for a suitable constant $\lambda_p > 0$. Corollary 4.6 applies again.

4.2. **Application: Lipschitz Operators.** We consider now a subclass of nonlinear operators which extend the class of linear sectorial operators of §3.1. We assume that $\mathfrak{F}$ satisfies (4.6) with $p = 2$, namely,

$$\|v - w\|^2 \le \langle \mathfrak{F}(v) - \mathfrak{F}(w), v - w\rangle \quad \forall\ v, w \in D(\mathfrak{F}), \tag{4.9}$$

as well as the following Lipschitz condition for some $\gamma > 0$

$$\|\mathfrak{F}(v) - \mathfrak{F}(w)\|_\star \le\ 2\gamma\|v - w\| \quad \forall\ v, w \in D(\mathfrak{F}); \tag{4.10}$$

compare with (3.3). Combining (4.9) and (4.10) gives $\gamma \ge \frac{1}{2}$. The following lemma extends Lemma 3.2, and is proved in [35, Lemma 4.3]. We present its proof here for completeness.

Lemma 4.9 (Angle Boundedness and Coercivity). *If $\mathfrak{F}$ satisfies* (4.9) *and* (4.10)*, then for all $v, w, z \in D(\mathfrak{F})$*

$$\langle \mathfrak{F}(v) - \mathfrak{F}(w), w - z\rangle \le 2\gamma^2\|v - z\|^2 - \frac{1}{2}\max\Big(\|v - w\|^2, \|z - w\|^2\Big). \tag{4.11}$$

Proof. Proceeding as in Lemma 3.2, we obtain

$$\begin{aligned}\langle \mathfrak{F}(v) - \mathfrak{F}(w), w - z\rangle &= \langle \mathfrak{F}(v) - \mathfrak{F}(w), w - v\rangle + \langle \mathfrak{F}(v) - \mathfrak{F}(w), v - z\rangle\\ &\le -\|v - w\|^2 + 2\gamma\|v - w\|\|v - z\| \le -\frac{1}{2}\|v - w\|^2 + 2\gamma^2\|v - z\|^2,\end{aligned}$$

as well as

$$\begin{aligned}\langle \mathfrak{F}(v) - \mathfrak{F}(w), w - z\rangle &= \langle \mathfrak{F}(z) - \mathfrak{F}(w), w - z\rangle + \langle \mathfrak{F}(v) - \mathfrak{F}(z), w - z\rangle\\ &\le -\|z - w\|^2 + 2\gamma\|v - z\|\|w - z\| \le -\frac{1}{2}\|w - z\|^2 + 2\gamma^2\|v - z\|^2.\end{aligned}$$

Combining these inequalities, we deduce the estimate (4.11). □

Note that (4.11) implies (4.3) with $p = 2$, $\eta = \sqrt{2}\,\gamma$, $\delta = 1$. Therefore Theorem 4.3 is applicable with an error measure $\mathfrak{E}$ of the form:

$$\mathfrak{E} := \left\{\max\Big(\max_{0\le t\le T}|(u - \widehat{U})(t)|^2, \frac{1}{2}\int_0^T \|u - \widehat{U}\|^2 dt, \frac{1}{2}\int_0^T \|u - U\|^2 dt\Big)\right\}^{1/2}.$$

Corollary 4.10 (Error Estimates for Lipschitz Operators). *If $u_0 \in \mathcal{V}$ and* (4.9) *and* (4.10) *hold, then the following lower and upper bounds are valid for both $dG(q)$ and IRK-$R(q)$ for all $q \ge 0$:*

$$\begin{aligned}\frac{1}{8}\alpha_2^2\sum_{n=0}^{N-1} k_n\|U_n^+ - U_n\|^2 \le \mathfrak{E}^2 \le\ & 4\gamma^2\alpha_2^2\sum_{n=0}^{N-1} k_n\|U_n^+ - U_n\|^2\\ &+ 2\int_0^T \Big(\|\mathfrak{F}(U) - \Pi\mathfrak{F}(U)\|_\star + \|f - F\|_\star\Big)^2 dt.\end{aligned} \tag{4.12}$$

Proof. We note that (4.11) with $\eta = \sqrt{2}\gamma$ yields (4.2) and (4.3) with $\delta = 1, p = 2$. Then Theorem 4.3 gives the upper bound. To derive the lower bound, we proceed as in Theorem 3.8, whose proof does not rely on the structure of $\mathfrak{F}$. $\square$

5. Conditional Estimates: Minimal Surface Equation

For quasilinear operators $\mathfrak{F}$ the theory of §4 does not always apply. It may, in particular, be difficult to find a suitable Sobolev setting. This is the case of the *minimal surface* operator over a domain Ω of $\mathbb{R}^d$

$$\mathfrak{F}(u) := \operatorname{div} \frac{\nabla u}{Q(u)}, \qquad Q(u) := \sqrt{1 + |\nabla u|^2}, \tag{5.1}$$

which is better studied in terms of geometric quantities. One such quantity is

$$N(u) = \frac{(-\nabla u\,,\, 1)}{Q(u)}, \tag{5.2}$$

the unit normal to the graph $\Gamma_u := \{(x, u(x)) : x \in \Omega\}$. The following equality has been first observed by Fierro and Veeser [19]:

$$\langle \mathfrak{F}(u) - \mathfrak{F}(v), u - v\rangle = \langle \frac{\nabla u}{Q(u)} - \frac{\nabla v}{Q(v)}\,,\, \nabla(u - v)\rangle = \int_\Omega |N(u) - N(v)|^2\, \frac{Q(u) + Q(v)}{2}.$$

This can be viewed as a geometric notion of coercivity in that we have a weighted L^2-estimate for the normals to the graphs Γ_u, Γ_v measured on Ω or, equivalently, an L^2-estimate measured on an 'average' of the graphs. For this to make sense we need $u, v \in W^1_1(\Omega)$.

Let U be either the dG(q) or IRK-R(q) approximation to the time-dependent prescribed mean curvature equation:

$$\partial_t u - \operatorname{div} \frac{\nabla u}{Q(u)} = f \quad \text{in } \Omega \times (0, T), \tag{5.3}$$

subject to an initial and lateral boundary condition $u = u_0$. We assume that (5.3) admits a smooth solution in $W^1_\infty(\Omega \times (0, T))$. We denote by

$$\mathfrak{J}(U) := U_n^+ - U_n, \qquad \mathfrak{N}(U) := \frac{\nabla U}{Q(U)} - \Pi\, \frac{\nabla U}{Q(U)}; \tag{5.4}$$

the piecewise constant jump $\mathfrak{J}(U)$ is a measure of numerical dissipation and $\mathfrak{N}(U)$ is a measure of the nonlinearity of $\mathfrak{F}$, which vanishes for $q = 0$ as in [34, 35, 28]. The latter is used next to quantify proximity to the exact solution u. We first set $\mathcal{H} = L^2(\Omega)$, and define the notion of coercivity

$$\rho(v, w; z) := \int_\Omega |N(v) - N(w)|^2\, Q(z)\; dx\,,$$

along with that of error for $0 \le \lambda \le 1$

$$\mathfrak{E} := \max \left\{ \max_{0 \le t \le T} |u - \widehat{U}|^2; \int_0^T \Big(\frac{1}{4} \rho(U, u; U) + (1 - \lambda) \rho(U, u; u) \Big) dt \right\}^{\frac{1}{2}}.$$

Proposition 5.1 (Conditional Estimate). *If the solution U of $dG(q)$ or IRK-R(q) with $q \ge 0$ verifies the a posteriori condition*

$$\lambda = 2 \|\mathfrak{N}(U) Q(U)^2\|_{L^\infty(\Omega)} \le 1, \tag{5.5}$$

and α_∞ is defined as in (2.8)*, then the error $\mathfrak{E}$ satisfies*

$$\mathfrak{E} \le \left\{ 5\alpha_\infty^2 \int_0^T \int_\Omega \frac{|\nabla \mathfrak{J}(U)|^2}{Q(U)} \; dxdt + 9 \int_0^T \int_\Omega |\mathfrak{N}(U)|^2 Q(U)^3 \; dxdt \right\}^{\frac{1}{2}} + \int_0^T |F - f| dt.$$

Remark 5.2 (Coercivity). It is worth noticing that the second term in the definition of $\mathfrak{E}$ is not a norm but a geometric quantity without homogeneity. Therefore, the technique of Lemma 4.5 does not apply to yield coercivity. However, the minimal surface operator is 1-angle bounded from Lemma 4.5, which leads to an optimal a posteriori error estimate for polynomial degree $q = 0$ as in [35]. For $q > 0$ we need control of the nonlinear term $\mathfrak{N}(U)$, which vanishes otherwise, and this is achieved via the coercivity term.

Remark 5.3 (Conditional Estimates). We point out that the condition (5.5) is a posteriori, and thus verifiable. It is conceivable that, for a sufficiently fine partition of $[0, T]$, (5.5) would be valid. Conditional estimates are somehow natural for nonlinear equations but rather unusual in a posteriori error analysis. We refer to Fierro and Veeser [19] for (elliptic) problems of prescribed mean curvature and Lakkis and Nochetto [28] for the (parabolic) mean curvature flow of graphs, both for $q = 1$.

Proof of Proposition 5.1. Subtracting (1.1) from (2.1). we obtain the error equation

$$\langle (\widehat{U} - u)', v \rangle \; + \langle \mathfrak{F}(U) - \mathfrak{F}(u), v \rangle \; = \langle \mathfrak{F}(U) - \Pi \mathfrak{F}(U), v \rangle + \langle F - f, v \rangle \, .$$

We now choose $v = \widehat{U} - u$ which, in view of (2.7), reads $v = U - u + \ell \mathfrak{J}(U)$ where $\ell(t) = -\widehat{\ell}_{n,0}(t)$ for $t \in I_n$. Since

$$\langle \mathfrak{F}(U) - \mathfrak{F}(u), U - u \rangle = \frac{1}{2} \Big(\rho(U, u; u) + \rho(U, u; U) \Big) \, ,$$

we have

$$\begin{aligned} \frac{d}{dt} |u - \widehat{U}|^2 + \rho(U, u; u) + \rho(U, u; U) \; &\le \; 2|\ell| \; |\langle \mathfrak{F}(U) - \mathfrak{F}(u), \mathfrak{J}(U) \rangle| \\ &\quad + 2\langle \mathfrak{N}(U), \nabla(\widehat{U} - u) \rangle + 2\langle F - f, \widehat{U} - u \rangle \, . \end{aligned} \tag{5.6}$$

Then, since $|\frac{\nabla u}{Q(u)} - \frac{\nabla U}{Q(U)}| \le |N(u) - N(U)|$, it is easily seen that

$$2|\ell| \; |\langle \mathfrak{F}(U) - \mathfrak{F}(u), \mathfrak{J}(U) \rangle| \le \; \frac{1}{4} \rho(U, u; U) + 4\alpha_\infty^2 \int_\Omega \frac{|\nabla \mathfrak{J}(U)|^2}{Q(U)} dx \, .$$

It remains to estimate the last two terms in the right hand side of (5.6). Both terms require finding a bound for $|\nabla(U-u)|$ in terms of geometric quantities we have control of. To this end, we proceed as in [19]. If we define $\mathbf{p},\mathbf{q}\in\mathbb{R}^{d+1}$ to be

$$\mathbf{p}=(\nabla u,1),\quad \mathbf{q}=(\nabla U,1)$$

for any $(x,t)\in\Omega\times(0,T)$, and denote with $|\mathbf{p}|$, $|\mathbf{q}|$ their ℓ^2-norms, then

$$\begin{aligned}|\mathbf{p}-\mathbf{q}| &= \Big||\mathbf{p}|\,\frac{\mathbf{p}}{|\mathbf{p}|}-|\mathbf{q}|\,\frac{\mathbf{q}}{|\mathbf{q}|}\Big| \\ &\le \Big|\frac{\mathbf{p}}{|\mathbf{p}|}-\frac{\mathbf{q}}{|\mathbf{q}|}\Big|\,|\mathbf{q}| + \Big||\mathbf{p}|\,\frac{\mathbf{p}}{|\mathbf{p}|}-|\mathbf{q}|\,\frac{\mathbf{p}}{|\mathbf{p}|}\Big| = \Big|\frac{\mathbf{p}}{|\mathbf{p}|}-\frac{\mathbf{q}}{|\mathbf{q}|}\Big|\,|\mathbf{q}| + \Big||\mathbf{p}|-|\mathbf{q}|\Big|,\end{aligned}$$

whence

$$|\nabla(U-u)|\le|N(u)-N(U)|\,Q(U)\,+\big|Q(u)-Q(U)\big|. \tag{5.7}$$

In addition, for $a,b\in\mathbb{R}$, we have

$$\frac{|b-a|}{b^2}\le\frac{|b-a|}{ab}+\frac{|b-a|}{b}\Big|\frac{1}{a}-\frac{1}{b}\Big|=\Big|\frac{1}{a}-\frac{1}{b}\Big|+\Big|\frac{1}{a}-\frac{1}{b}\Big|^2a.$$

Hence, taking $a=Q(u),b=Q(U)$ and using $|\frac{1}{a}-\frac{1}{b}|\le|N(u)-N(U)|$, we get

$$|Q(u)-Q(U)|\le Q(U)^2|N(u)-N(U)|+Q(u)Q(U)^2|N(u)-N(U)|^2. \tag{5.8}$$

Combining (5.7) and (5.8), and making use of $Q(U)\ge1$, we arrive at

$$|\nabla(U-u)|\le 2|N(u)-N(U)|Q(U)^2+|N(u)-N(U)|^2Q(u)Q(U)^2.$$

Consequently, if $\lambda=2\|\mathfrak{N}(U)Q(U)^2\|_{L^\infty(\Omega)}$, then we get

$$|\langle\mathfrak{N}(U),\nabla(U-u)\rangle|\le\frac{1}{4}\rho(U,u;U)+4\int_\Omega\mathfrak{N}(U)^2Q(U)^3\,dx+\frac{\lambda}{2}\rho(U,u;u),$$

whence, using $Q(U)\ge1$,

$$\begin{aligned}2|\langle\mathfrak{N}(U),\nabla(\widehat{U}-u)\rangle| &\le \frac{1}{2}\,\rho(U,u;U)+\lambda\,\rho(U,u;u)\\ &\quad+9\int_\Omega|\mathfrak{N}(U)|^2Q(U)^3\,+\alpha_\infty^2\int_\Omega\frac{|\nabla\mathfrak{I}(U)|^2}{Q(U)}.\end{aligned}$$

Inserting the above estimates back into (5.6), we find

$$\begin{aligned}&\frac{d}{dt}|u-\widehat{U}|^2+(1-\lambda)\rho(U,u;u)+\frac{1}{4}\rho(U,u;U)\\ &\qquad\le 9\int_\Omega|\mathfrak{N}(U)|^2Q(U)^3\,+5\alpha_\infty^2\int_\Omega\frac{|\nabla\mathfrak{I}(U)|^2}{Q(U)}+2|F-f|\;|\widehat{U}-u|.\end{aligned}$$

Since $\lambda\le1$ in view of (5.5), we thus have an expression of the form

$$\frac{d}{dt}a^2(t)+b^2(t)\le c^2(t)+2d(t)a(t).$$

The following Gronwall-like inequality [35, Lemma 3.7]

$$\max\Big\{\max_{0\le t\le T} a(t), \Big(\int_0^T b^2(t)dt\Big)^{\frac{1}{2}}\Big\} \le \Big(\int_0^T c^2(t)dt\Big)^{\frac{1}{2}} + \int_0^T d(t)dt$$

yields the asserted estimate and completes the proof. □

Acknowledgement: We are grateful to the referees for their valuable comments and suggestions, and for bringing up Zadunaisky's work [45].

References

[1] G. Akrivis, Ch. G. Makridakis, and R.H. Nochetto. A posteriori error estimates for the Crank-Nicolson method for parabolic equations. *Math. Comp.*, (to appear).

[2] G. A. Baker and J. H. Bramble. Semidiscrete and single step fully discrete approximations for second order hyperbolic equations. *RAIRO Anal. Numér.*, 13(2):75–100, 1979.

[3] W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations.* Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.

[4] A. Bergam, C. Bernardi, and Z. Mghazli. A posteriori analysis of the finite element discretization of some parabolic equations. *Math. Comp.*, 74(251):1117–1138 (electronic), 2005.

[5] H. Brézis. Problèmes unilatéraux. *J. Math. Pures Appl. (9)*, 51:1–168, 1972.

[6] H. Brézis and F. E. Browder. Nonlinear integral equations and systems of Hammerstein type. *Advances in Math.*, 18(2):115–147, 1975.

[7] L. A. Caffarelli. The regularity of monotone maps of finite compression. *Comm. Pure Appl. Math.*, 50(6):563–591, 1997.

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

[9] K. Dekker and J. G. Verwer. *Stability of Runge-Kutta methods for stiff nonlinear differential equations.* North-Holland Publishing Co., Amsterdam, 1984.

[10] K Eriksson, D Estep, P Hansbo, and C Johnson. Introduction to adaptive methods for differential equations. *Acta Numerica*, pages 105–158, 1995.

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

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

[13] K. Eriksson, C. Johnson, and S. Larsson. Adaptive finite element methods for parabolic problems. VI. Analytic semigroups. *SIAM J. Numer. Anal.*, 35(4):1315–1325 (electronic), 1998.

[14] K. Eriksson, C. Johnson, and A. Logg. Adaptive computational methods for parabolic problems. In *Encyclopedia of Computational Mechanics*, Edited by E. Stein, R. de Borst and T.J.R. Houghes, pages 1–44. J. Wiley & Sons, Ltd., New York, 2004.

[15] K. Eriksson, C. Johnson, and V. Thomée. Time discretization of parabolic problems by the discontinuous Galerkin method. *RAIRO Modél. Math. Anal. Numér.*, 19(4):611–643, 1985.

[16] D. J. Estep. A posteriori error bounds and global error control for approximation of ordinary differential equations. *SIAM J. Numer. Anal.*, 32(1):1–48, 1995.

[17] D. J. Estep, M. G. Larson, and R. D. Williams. Estimating the error of numerical solutions of systems of reaction-diffusion equations. *Mem. Amer. Math. Soc.*, 146(696):viii+109, 2000.

[18] D. J. Estep and A. M. Stuart. The dynamical behavior of the discontinuous Galerkin method and related difference schemes. *Math. Comp.*, 71(239):1075–1103 (electronic), 2002.

[19] F. Fierro and A. Veeser. On the a posteriori error analysis for equations of prescribed mean curvature. *Math. Comp.*, 72:1611–1634, 2003.

[20] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations. II Stiff and Differential-Algebraic Problems.* Springer-Verlag, Berlin, 1991.

[21] P. Houston, B. Senior, and E. Süli. *hp*-discontinuous Galerkin finite element methods for hyperbolic problems: error analysis and adaptivity. *Internat. J. Numer. Methods Fluids*, 40(1-2):153–169, 2002. ICFD Conference on Numerical Methods for Fluid Dynamics (Oxford, 2001).

[22] P. Houston and E. Süli. *hp*-adaptive discontinuous Galerkin finite element methods for first-order hyperbolic problems. *SIAM J. Sci. Comput.*, 23(4):1226–1252 (electronic), 2001.

[23] P. Jamet. Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain. *SIAM J. Numer. Anal.*, 15(5):912–928, 1978.

[24] 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(4):908–926, 1988.

[25] C. Johnson. Discontinuous Galerkin finite element methods for second order hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, 107(1-2):117–129, 1993.

[26] 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(2):277–291, 1990.

[27] O. Karakashian and Ch. Makridakis. A space-time finite element method for the nonlinear Schrödinger equation: the discontinuous Galerkin method. *Math. Comp.*, 67(222):479–499, 1998.

[28] O. Lakkis and R. H. Nochetto. A posteriori error analysis for the mean curvature flow of graphs. *SIAM J. Numer. Anal.*, 2004.

[29] P. Lasaint and P.-A. Raviart. On a finite element method for solving the neutron transport equation, in mathematical aspects of finite elements in partial differential equations. *Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Academic Press*, pages 89–123. Publication No. 33, 1974.

[30] J.-L. Lions. *Quelques Méthodes de Résolution des Problèmes aux Limites Non Linéaires.* Dunod, 1969.

[31] G. Lippold. Error estimates for the implicit Euler approximation of an evolution inequality. *Nonlinear Anal.*, 15(11):1077–1089, 1990.

[32] G. Lippold. Error estimates and step-size control for the approximate solution of a first order evolution equation. *RAIRO Modél. Math. Anal. Numér.*, 25(1):111–128, 1991.

[33] R. H. Nochetto and G. Savaré. Nonlinear evolution equations governed by accretive operators in banach spaces: error control and applications. *Math. Models Methods Appl. Sci.*, (to appear).

[34] R. H. Nochetto, G Savaré, and C. Verdi. Error control for nonlinear evolution equations. *C.R. Acad. Sci. Paris, Serie I*, 326:1437–1442, 1998.

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

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

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

[38] J. Rulla. Error analysis for implicit approximations to solutions to cauchy problems. *SIAM J. Numer. Anal.*, 33(1):68–87, 1996.

[39] R. D. Skeel. Thirteen ways to estimate global error. *Numer. Math.*, 48(1):1–20, 1986.

[40] A. M. Stuart and A. R. Humphries. The essential stability of local error control for dynamical systems. *SIAM J. Numer. Anal.*, 32(6):1940–1971, 1995.

[41] E. Süli and P. Houston. Finite element methods for hyperbolic problems: a posteriori error analysis and adaptivity. In *The state of the art in numerical analysis (York, 1996)*, volume 63 of *Inst. Math. Appl. Conf. Ser. New Ser.*, pages 441–471. Oxford Univ. Press, New York, 1997.

[42] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*. Springer-Verlag, Berlin, 1997.

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

[44] R. Verfürth. A posteriori error estimates for finite element discretizations of the heat equation. *Calcolo*, 40:195–212, 2003.

[45] P. E. Zadunaisky. On the estimation of errors propagated in the numerical integration of ordinary differential equations. *Numer. Math.*, 27(1):21–39, 1976/77.

Department of Applied Mathematics, University of Crete, 71409 Heraklion-Crete, Greece and Institute of Applied and Computational Mathematics, FORTH, 71110 Heraklion-Crete, Greece. Partially supported by the European Union RTN-network HYKE, HPRN-CT-2002-00282, and the EU Marie Curie Development Host Site, HPMD-CT-2001-00121.

URL: http://www.tem.uoc.gr/~makr
E-mail address: makr@tem.uoc.gr

Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, MD 20742, USA. Partially supported by NSF Grants DMS-9971450 and DMS-0204670 and the General Research Board of the University of Maryland.

URL: http://www.math.umd.edu/~rhn
E-mail address: rhn@math.umd.edu