Electronic Transactions on Numerical Analysis.
Volume 60, pp. 99–122, 2024.

ISSN 1068–9613.
DOI: 10.1553/etna_vol60s99

ETNA
Kent State University and
Johann Radon Institute (RICAM)

A REVIEW OF MAXIMUM-NORM A POSTERIORI ERROR BOUNDS FOR TIME-SEMIDISCRETISATIONS OF PARABOLIC EQUATIONS*

TORSTEN LINSS†, NATALIA KOPTEVA‡, GORAN RADOJEV§, AND MARTIN OSSADNIK†

Abstract. A posteriori error estimates in the maximum norm are studied for various time-semidiscretisations applied to a class of linear parabolic equations. We summarise results from the literature and present some new improved error bounds. Crucial ingredients are certain bounds in the L_1-norm for the Green's function associated with the parabolic operator and its derivatives.

Key words. parabolic problems, maximum-norm a posteriori error estimates, backward Euler, Crank–Nicolson, extrapolation, discontinuous Galerkin–Radau, backward differentiation formulae, Green's function

AMS subject classifications. 65M15, 65M60

1. Introduction. Consider the linear parabolic equation:

$$\mathcal{K}u := \partial_t u + \mathcal{L}u = f\,, \qquad \text{in} \quad Q := \Omega \times (0,T], \tag{1.1a}$$

with a second-order linear elliptic operator $\mathcal{L}$ in a spatial domain $\Omega \subset \mathbb{R}^n$ with Lipschitz boundary and some function $f \in C\big(0,T;L_\infty(\Omega)\big)$, subject to the initial condition

$$u(x,0) = u^0(x)\,, \quad \text{for} \quad x \in \bar{\Omega}, \tag{1.1b}$$

and the Dirichlet boundary condition

$$u(x,t) = 0\,, \qquad \text{for} \quad (x,t) \in \partial\Omega \times [0,T]. \tag{1.1c}$$

The initial datum u^0 is assumed to be compatible with the boundary conditions, i.e., $u^0|_{\partial\Omega} = 0$.

Following [7] and [5], the authors of the present study have published a number of results on residual-type a posteriori error estimates in the maximum norm for parabolic equations utilising and merging various approaches and considering various classes of temporal discretisation [5, 6, 11, 12, 15, 17]. In this survey, we review these results in a unified manner. Revisiting those results and their proofs, we are able to present some improvements — namely sharper error bounds — for the implicit Euler method, the Crank–Nicolson method, and the dG(1)-method. These improvements are made possible by using local bounds for the Green's function rather than global stability results. Details will be highlighted in the course of the paper. We also present some new results (most notably Theorems 4.5 and 5.2). Furthermore, numerical results are given to compare the various approaches.

The general idea is to set up a parabolic PDE for the error with the residual on the right-hand side. Then the error (at final time T) can be represented by means of the Green's function associated with $\mathcal{K}$ and that residual. To this end, the time-discrete approximations need to be extended to a function defined on $[0,T]$. Furthermore, bounds in the L_1-norm for the Green's function and its time-derivatives are required; see Section 2 for details.

*Received December 23, 2022. Accepted November 29, 2023. Published online on February 29, 2024. Recommended by Christian Vergara.

†Fakultät für Mathematik und Informatik, FernUniversität in Hagen, Universitätsstraße 11, 58095 Hagen, Germany ({torsten.linss, martin.ossadnik}@fernuni-hagen.de).

‡Department of Mathematics and Statistics, University of Limerick, Limerick, V94 T9PX, Ireland (natalia.kopteva@ul.ie).

§Department of Mathematics and Computer Science, Faculty of Sciences, University of Novi Sad, Trg Dositeja Obradovića 4, 21000 Novi Sad, Serbia, (goran.radojev@dmi.uns.ac.rs).

In the present paper we study semidiscretisations in time only. However, these are essential building blocks in deriving error estimates for full (space-and-time) discretisations. Using elliptic reconstructions, a concept introduced by Makridakis and Nochetto [16], they can be combined with error estimators for discretisations of elliptic problems to give error bounds for parabolic problems; see, e.g., [2, 14] in the context of L_2-norm and H^1-norm error estimation.

The paper is organised as follows. In Section 2 we specify our general assumptions for the a posteriori error analysis, in particular we stipulate the validity of certain bounds for the Green's function of the parabolic problem. Thereafter, we present result for various discretisations:

- the simple first-order implicit Euler method (Section 4),
- the second-order Crank–Nicolson method (Section 5),
- an extrapolated Euler method of second order (Section 6),
- the third-order discontinuous Galerkin method with polynomials of degree 1 (dG(1) for short), which is equivalent to the Runge–Kutta–Radau-IIA method (Section 7), and finally
- the backward-differentiation formula of order 2 (Section 8).

We complement the theoretical finds with results of numerical experiments. The test problem is introduced in Section 3.

Notation: For functions $w\colon \Omega\times[0,T]\to\mathbb{R}$ we shall use the shortend notation $w(t) := w(\cdot,t)$, which for each time t is a function mapping from Ω to $\mathbb{R}$.

2. The Green's function. In this section we consider the Green's function $\mathcal{G}$ associated with the operator $\mathcal{K}$ in (1.1). It will be used to express the error of a numerical approximation in terms of its residual in the differential equation. For definitions and properties of fundamental solutions and Green's functions of parabolic operators, we refer the reader to the survey by Friedman [9], in particular Chapter 1.

For fixed $x\in\Omega$, the Green's function associated with $\mathcal{K}$ and x solves

$$\begin{aligned} \partial_t\mathcal{G} + \mathcal{L}^*\mathcal{G} &= 0, && \text{in } \Omega\times\mathbb{R}^+,\\ \mathcal{G}\big|_{\partial\Omega} &= 0, \\ \mathcal{G}(0) &= \delta_x = \delta(\cdot - x)\,, \end{aligned}$$

with δ denoting the Dirac δ-distribution. Let $\langle\cdot,\cdot\rangle$ denote both the duality pairing on $H^{-1}(\Omega)\times H_0^1(\Omega)$ and the $L_2(\Omega)$-scalar product. Then for all $w\in W^{1,1}\big([0,T],H_0^1(\Omega)\big)$ and $t\in(0,T]$, we have

$$w(x,t) = \langle\mathcal{G}(t),w(0)\rangle + \int_0^t \big\langle \mathcal{G}(t-s),\big(\mathcal{K}w\big)(s)\big\rangle\,\mathrm{d}s. \tag{2.1}$$

We will make frequent use of this representation of a function w in terms of $\mathcal{K}w$.

Throughout the paper we shall assume there exist non-negative constants κ_0, κ_1, κ_2, κ_1', κ_2', and γ such that (with formally setting $\kappa_0' = 0$)

$$\|\partial_t^p\mathcal{G}(t)\|_{1,\Omega} \le \left(\frac{\kappa_p}{t^p} + \kappa_p'\right)\,\mathrm{e}^{-\gamma t} =: \varphi_p(t), \qquad \begin{array}{l}\text{for all } x\in\Omega,\ t\in[0,T],\\ \text{and } p=0,1,2.\end{array} \tag{2.2}$$

Here $\|\cdot\|_{p,\Omega}$, $p\in[1,\infty]$, denotes the standard norm in $L_p(\Omega)$.

2.1. Problems that satisfy (2.2). A number of problems that satisfies these assumptions are gathered in [13, §2.1]. Let us mention some of those.

(i). For the heat equation

$$u_t - \Delta u = \varphi(x,t) \qquad \text{in } \Omega \subset \mathbb{R}^n, n \in \mathbb{N},$$

we have (2.2) with $\kappa_0 = 1$, $\kappa_p' = 0$, and $\kappa_p = p!\, 18^{p-1} 3^n\, 2^{-(n/2+1)}$.

(ii). For the singularly perturbed reaction-diffusion equation

$$u_t - \varepsilon^2 u_{xx} + r(x)u = \varphi(x,t) \qquad \text{in } \Omega = (0,1),$$

with $0 < \varepsilon \ll 1$, $r \in C^{0,1}[0,1]$, $r \geq \varrho^2$, $\varrho > 0$, the bounds of (2.2) hold true with $\gamma = \varrho$, $\kappa_0 = 1$, $\kappa_1 = \sqrt{2/(\pi \mathrm{e})}$, and $\kappa_1' = \kappa_0 \|r\|_{\infty,[0,1]} + \mathcal{O}(\varepsilon)$, $\varepsilon \to 0$; see [11, §2]. Furthermore,

$$\kappa_2 = \frac{4}{\sqrt{\pi}} \int_{\mathbb{R}} \left|4s^4 - 12s^2 + 3\right| \mathrm{e}^{-s^2} \mathrm{d}s + c_0 \approx 0.70015 + c_0,$$

with an arbitrary $c_0 > 0$, and $\kappa_2' = \|r\|^2_{\infty,[0,1]} \left(1 + \kappa_1^2 c_0^{-1}\right) + \mathcal{O}(\varepsilon)$.

(iii). For the reaction-diffusion equation

$$u_t - \varepsilon^2 \Delta u + r(x)u = \varphi(x,t) \qquad \text{in } \Omega \subset \mathbb{R}^n, n > 1,$$

with $\varepsilon \in (0,1]$, $r \in C^{0,1}(\bar{\Omega})$, $r \geq \varrho^2 \geq 0$, one has the estimate (2.2) with $\gamma^2 = \varrho^2/2$, $\kappa_0 = 1$, $\kappa_p = p!\, 18^{p-1} 3^n\, 2^{-(n/2+1)}$, and $\kappa_p' = 0$; see [12, §12] for $p = 1$. Then the bound for $p = 2$ can be obtained employing [4, Corollary 5].

REMARK 2.1. In the context of our investigations, the bounds above are exemplary. The constants κ_p and γ reappear in our a posteriori error bounds later. This means that any improvement in these bounds, i.e., smaller κ's or larger γ, will yield sharper error bounds.

2.2. Auxiliary calculations. The rest of this section is rather technical as we will precompute some coefficients that feature in our error bounds later. They appear after Hölder's inequality and (2.2) have been applied to integrals involving (derivatives of) the Green's function. The calculations are elementary, albeit tedious. The results may be verified using MAPLE. Those integrals are of the form

$$\int_{t_{j-1}}^{t_j} \xi(s)\varphi_p(T-s)\mathrm{d}s, \qquad \begin{array}{l} \text{with } \ 0 \leq t_{j-1} < t_j \leq T, \ \ p = 0,1,2, \\ \text{and a function } \ \xi \in C^0[t_{j-1}, t_j]. \end{array}$$

In a first step, these are bounded as follows

$$\left| \int_{t_{j-1}}^{t_j} \xi(s)\varphi_p(T-s)\mathrm{d}s \right| \leq \mathrm{e}^{-\gamma(T-t_j)} \int_{t_{j-1}}^{t_j} |\xi(s)| \left(\frac{\kappa_p}{(T-s)^p} + \kappa_p' \right) \mathrm{d}s\,. \tag{2.3}$$

Sharper bounds are obtained for particular p and for functions ξ that are polynomials in s (with $\tau_j := t_j - t_{j-1}$):

$$\left| \int_{t_{j-1}}^{t_j} \xi(s)\varphi_0(T-s)\mathrm{d}s \right| \leq \kappa_0 \mathrm{e}^{-\gamma(T-t_j)} \int_{t_{j-1}}^{t_j} |\xi(s)|\, \mathrm{d}s\,, \tag{2.4}$$

$$0 \leq \int_{t_{j-1}}^{t_j} \varphi_1(T-s)\mathrm{d}s = \mathrm{e}^{-\gamma(T-t_j)} \vartheta_j, \quad \vartheta_j := \left\{ \kappa_1 \ln\left(1 + \frac{\tau_j}{T-t_j}\right) + \kappa_1' \tau_j \right\} \tag{2.5}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

and

$$
\begin{aligned}
&0 \le \int_{t_{j-1}}^{t_j} (t_j - s)\, \varphi_1(T-s)\mathrm{d}s = \mathrm{e}^{-\gamma(T-t_j)}\varrho_j, \\
&\varrho_j := \left\{ \kappa_1 \left[\tau_j - (T-t_j) \ln\left(1 + \frac{\tau_j}{T-t_j}\right)\right] + \kappa_1' \frac{\tau_j^2}{2} \right\}.
\end{aligned} \tag{2.6}
$$

Another example that appears frequently is, for $k = 0, 1, \dots,$

$$
\left| \int_{t_{j-1}}^{t_j} \left(t_j - s\right)^k \left(s - t_{j-1}\right) \partial_t \mathcal{G}(T-s) \mathrm{d}s \right| \le \mathrm{e}^{-\gamma(T-t_j)} \Phi_{k,j}\,, \tag{2.7}
$$

where

$$
\Phi_{k,j} := \kappa_1 \mu_{k,j} + \kappa_1' \frac{\tau_j^{k+2}}{(k+1)(k+2)} \quad \text{and} \quad \mu_{k,j} := \int_{t_{j-1}}^{t_j} \frac{\left(t_j - s\right)^k \left(s - t_{j-1}\right)}{T-s} \mathrm{d}s\,.
$$

The last integral can be computed recursively:

$$
\begin{aligned}
\mu_{0,j} &= -\tau_j + (T - t_{j-1}) \ln\left(1 + \frac{\tau_j}{T-t_j}\right), \\
\mu_{k,j} &= \frac{\tau_j^{k+1}}{k(k+1)} + (t_j - T)\, \mu_{k-1,j}, \qquad k = 1, 2, \dots
\end{aligned}
$$

However, when t_j is close to 0, destructive cancellation occurs because the two summands are of like magnitude but of different sign. In this case an alternative is to compute $\mu_{k,j}$ using a suitable truncation of the series expansion

$$
\mu_{k,j} = \tau_j^{k+1} \sum_{\ell=1}^{\infty} \frac{(-1)^{\ell+1}}{(\ell+k)(\ell+k+1)} \left(\frac{\tau_j}{T-t_j}\right)^{\ell}.
$$

In our numerical experiments we used the first 5 terms of this expansion.

Furthermore, for $k > 0$,

$$
\int_{t_{j-1}}^{t_j} \left(t_j - s\right)^k \left(s - t_{j-1}\right) \partial_t \mathcal{G}(T-s) \mathrm{d}s = \int_{t_{j-1}}^{t_j} \frac{\mathrm{d}}{\mathrm{d}s} \left[\left(t_j - s\right)^k \left(s - t_{j-1}\right) \right] \mathcal{G}(T-s) \mathrm{d}s.
$$

Application of (2.2), gives the alternative bound

$$
\begin{aligned}
&\left| \int_{t_{j-1}}^{t_j} \left(t_j - s\right)^k \left(s - t_{j-1}\right) \partial_t \mathcal{G}(T-s) \mathrm{d}s \right| \le \mathrm{e}^{-\gamma(T-t_j)} \Phi_{k,j}^*, \\
&\Phi_{k,j}^* := \kappa_0 \int_{t_{j-1}}^{t_j} \left| \frac{\mathrm{d}}{\mathrm{d}s} \left[\left(t_j - s\right)^k \left(s - t_{j-1}\right) \right] \right| \mathrm{d}s\,.
\end{aligned} \tag{2.8}
$$

Combining (2.7) and (2.8) gives

$$
\left| \int_{t_{j-1}}^{t_j} \left(t_j - s\right)^k \left(s - t_{j-1}\right) \partial_t \mathcal{G}(T-s) \mathrm{d}s \right| \le \mathrm{e}^{-\gamma(T-t_j)} \min\left\{ \Phi_{k,j}, \Phi_{k,j}^* \right\} =: \Psi_{k,j}\,. \tag{2.9}
$$

3. Test problem. Throughout the paper we shall give numerical results for the linear reaction-diffusion equation

$$\partial_t u - u_{xx} + (5x+6)u = \mathrm{e}^{-4t} - \cos\big(\pi(x+t)^3\big) \quad \text{in} \quad (-1,1)\times(0,1], \tag{3.1a}$$

subject to the initial condition

$$u(x,0) = u^0(x) = \sin\frac{\pi(1+x)}{2} \qquad \text{for} \quad x\in[-1,1], \tag{3.1b}$$

and the Dirichlet boundary condition

$$u(x,t) = 0 \qquad\qquad \text{for} \quad (x,t)\in\{-1,1\}\times[0,1]. \tag{3.1c}$$

The Green's function for this problem satisfies [4, Corollary 5]

$$\|\mathcal{G}(t)\|_{1,\Omega} \le \mathrm{e}^{-t/2}, \qquad \|\partial_t^p \mathcal{G}(t)\|_{1,\Omega} \le \frac{3}{2^{3/2}}\frac{p!18^{p-1}}{t^p}\,\mathrm{e}^{-t/2}, \;\; p\in\{1,2\}.$$

The elliptic problems obtained after semi-discretisation in time are solved using a spectral collocation method with polynomials of degree 31. This allows to solve those problems almost to machine accuracy. We are interested in the errors and error estimates at final time T. A reference solution is computed using dG(2) in time. This is a method of order 5; cf. [8, 10].

4. The implicit Euler method. We consider the first-order backward Euler discretisations in time applied to problem (1.1). Let an arbitrary mesh in time be given by

$$\mathcal{T}_M := \{t_j\}_{j=0}^M, \qquad 0 = t_0 < t_1 < \cdots < t_M = T.$$

For $j=1,\ldots,M$, we set

$$I_j := \big(t_{j-1},t_j\big)\,, \quad \tau_j := t_j - t_{j-1}, \qquad \text{and} \quad \tau := \max_{j=1,\ldots,M}\tau_j\,.$$

Furthermore, for $\varsigma\in[0,1]$, let $t_{j-\varsigma} := t_j - \varsigma\tau_j$ and $v^{j-\varsigma} := v(t_{j-\varsigma})$.

We discretise the abstract parabolic problem (1.1) in time on the mesh $\mathcal{T}_M$ using the first-order backward Euler method as follows. We associate an approximate solution $U^j\in H_0^1(\Omega)$ with the time level t_j and require it to satisfy

$$\delta_t U^j + \mathcal{L}U^j = f^j \qquad \text{in } \Omega, \qquad j=1,\ldots,M; \quad U^0 = u^0, \tag{4.1}$$

where

$$\delta_t U^j := \frac{U^j - U^{j-1}}{\tau_j} \qquad \text{and} \quad f^j := f(\cdot,t_j).$$

§4.1. The central idea is to extend U^j to a piecewise linear function $\hat{U}$ that is defined on all of the interval $[0,T]$ and then invoke (2.1) with $w = u-\hat{U}$. To this end, for any function v defined on $\mathcal{T}_M$, $t_j\mapsto v^j$, we denote by $\hat{v}$ its piecewise linear interpolant, i.e.,

$$\begin{aligned}\hat{v}(s) := v^j - \big(t_j - s\big)\delta_t v^j &= v^{j-1} + \big(s-t_{j-1}\big)\delta_t v^j \\ &= v^{j-1/2} + \big(s - t_{j-1/2}\big)\delta_t v^j\,, \qquad s\in\bar{I}_j,\;\; j=1,\ldots,M.\end{aligned}$$

Note that $\partial_t \hat{v}(s) = \delta_t v^j$, for $s \in I_j$, $j = 1, \ldots, M$. Hence,

$$\partial_t \hat{U} = \delta_t U^j = f^j - \mathcal{L}U^j \qquad \text{in } \ I_j\,, \ \text{ by (4.1).}$$

Recalling (1.1), the residual of $\hat{U}$ in the differential equation admits the representation

$$\begin{aligned} \left(\mathcal{K}\big(u - \hat{U}\big)\right)(s) &= f(s) - \partial_t \hat{U}(s) - \mathcal{L}\left(U^j - (t_j - s)\delta_t U^j\right) \\ &= f(s) - f^j + (t_j - s)\, \delta_t \left(\mathcal{L}U\right)^j, \qquad s \in I_j. \end{aligned} \tag{4.2}$$

Invoking (2.1), we obtain for the error at final time $T = t_M$

$$\begin{aligned} u(x,T) - U^M(x) &= \big(u - \hat{U}\big)(x,T) \\ &= \sum_{j=1}^{M} \left\{ \int_{I_j} \left\langle \mathcal{G}(T-s), f(s) - f^j \right\rangle \mathrm{d}s + \int_{I_j} (t_j - s) \left\langle \mathcal{G}(T-s), \delta_t \left(\mathcal{L}U\right)^j \right\rangle \mathrm{d}s \right\} \\ &= \sum_{j=1}^{M} \left\{ \int_{I_j} \left\langle \mathcal{G}(T-s), f(s) - f^j \right\rangle \mathrm{d}s + \int_{I_j} (t_j - s) \left\langle \partial_t \mathcal{G}(T-s), \delta_t U^j \right\rangle \mathrm{d}s \right\}, \end{aligned} \tag{4.3}$$

because $(\partial_t + \mathcal{L}^*)\,\mathcal{G} = 0$. Using the Hölder inequality and (2.2), we obtain two bounds:

$$\begin{aligned} \left\| u(T) - U^M \right\|_{\infty,\Omega} \le \sum_{j=1}^{M} \Bigg\{ &\int_{I_j} \varphi_0(T-s) \left\| f(s) - f^j \right\|_{\infty,\Omega} \mathrm{d}s \\ &+ \int_{I_j} (t_j - s)\, \varphi_0(T-s) \mathrm{d}s \left\| \delta_t \left(\mathcal{L}U\right)^j \right\|_{\infty,\Omega} \Bigg\} \end{aligned} \tag{4.4a}$$

and

$$\begin{aligned} \left\| u(T) - U^M \right\|_{\infty,\Omega} \le \sum_{j=1}^{M} \Bigg\{ &\int_{I_j} \varphi_0(T-s) \left\| f(s) - f^j \right\|_{\infty,\Omega} \mathrm{d}s \\ &+ \int_{I_j} \varphi_1(T-s) (t_j - s) \mathrm{d}s \left\| \delta_t U^j \right\|_{\infty,\Omega} \Bigg\}. \end{aligned} \tag{4.4b}$$

Upon noting that the φ_i, $i = 0, 1$, are non-increasing, we obtain the following theorems. The first resembles the result given in [5, §4.3, Theorem 4.2] while the second was derived in [12, §4, Theorem 4.1]. A version of the latter is also given in [7, §1, Theorem 1.3] but without providing a proof and without fixing the constants.

THEOREM 4.1. *The maximum-norm error of the backward Euler time discretisation* (4.1) *satisfies the a posteriori bound*

$$\left\| u(T) - U^M \right\|_{\infty,\Omega} \le \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T - t_j)} \left(\eta_{\bar{f}}^j + \eta_{\delta \mathcal{L} U}^j \right)$$

with

$$\eta_{\bar{f}}^j := \kappa_0 \int_{I_j} \left\| f(s) - f^j \right\|_{\infty,\Omega} \mathrm{d}s \qquad \textit{and} \qquad \eta_{\delta \mathcal{L} U}^j := \frac{\kappa_0 \tau_j^2}{2} \left\| \delta_t \left(\mathcal{L}U\right)^j \right\|_{\infty,\Omega}.$$

TABLE 4.1
Error estimators of Theorems 4.1–4.3 for the Euler method applied to the test problem (3.1). Simpson's rule is used to estimate the $\eta_{\bar{f}}^j$.

M	err	Theorem 4.1 est	Theorem 4.1 eff	Theorem 4.2 est	Theorem 4.2 eff	Theorem 4.3 est	Theorem 4.3 eff
256	$1.045 \cdot 10^{-4}$	$7.333 \cdot 10^{-2}$	1/702	$3.970 \cdot 10^{-2}$	1/380	$3.872 \cdot 10^{-2}$	1/370
512	$5.175 \cdot 10^{-5}$	$3.664 \cdot 10^{-2}$	1/708	$2.003 \cdot 10^{-2}$	1/387	$1.934 \cdot 10^{-2}$	1/374
1024	$2.575 \cdot 10^{-5}$	$1.831 \cdot 10^{-2}$	1/711	$1.011 \cdot 10^{-2}$	1/393	$9.662 \cdot 10^{-3}$	1/375
2048	$1.284 \cdot 10^{-5}$	$9.155 \cdot 10^{-3}$	1/713	$5.106 \cdot 10^{-3}$	1/398	$4.829 \cdot 10^{-3}$	1/376
4096	$6.412 \cdot 10^{-6}$	$4.577 \cdot 10^{-3}$	1/714	$2.578 \cdot 10^{-3}$	1/402	$2.414 \cdot 10^{-3}$	1/377
8192	$3.204 \cdot 10^{-6}$	$2.288 \cdot 10^{-3}$	1/714	$1.302 \cdot 10^{-3}$	1/406	$1.207 \cdot 10^{-3}$	1/377
16384	$1.601 \cdot 10^{-6}$	$1.144 \cdot 10^{-3}$	1/715	$6.576 \cdot 10^{-4}$	1/411	$6.035 \cdot 10^{-4}$	1/377
32768	$8.006 \cdot 10^{-7}$	$5.721 \cdot 10^{-4}$	1/715	$3.320 \cdot 10^{-4}$	1/415	$3.017 \cdot 10^{-4}$	1/377
65536	$4.002 \cdot 10^{-7}$	$2.860 \cdot 10^{-4}$	1/715	$1.676 \cdot 10^{-4}$	1/419	$1.509 \cdot 10^{-4}$	1/377

THEOREM 4.2. *The maximum-norm error of the backward Euler time discretisation (4.1) satisfies the a posteriori bound*

$$\left\|u(T) - U^M\right\|_{\infty,\Omega} \le \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T-t_j)} \left(\eta_{\bar{f}}^j + \eta_{\delta U}^j\right),$$

with $\eta_{\bar{f}}^j$ as in Theorem 4.1, $\eta_{\delta U}^j := \varrho_j \left\|\delta_t U^j\right\|_{\infty,\Omega}$, and ϱ_j from (2.6).

The derivation of Theorem 4.2 in [12] uses a different, global argument employing a piecewise constant and discontinuous interpolant of the U^j. In doing so, it passed unnoticed that these bounds can be combined by locally taking, for each $j = 1, \dots, M$, the smaller of the two bounds in (4.4). We arrive at the following novel result.

THEOREM 4.3. *The maximum-norm error of the backward Euler time discretisation (4.1) satisfies the a posteriori bound*

$$\left\|u(T) - U^M\right\|_{\infty,\Omega} \le \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T-t_j)} \left(\eta_{\bar{f}}^j + \eta_{\min}^j\right), \quad \textit{with} \quad \eta_{\min}^j := \min\left\{\eta_{\delta U}^j, \eta_{\delta \mathcal{L} U}^j\right\}$$

and the notation from Theorems 4.1 and 4.2.

REMARK 4.4. The integral defining $\eta_{\bar{f}}^j$ can (in general) not been evaluated exactly but needs to be approximated. Possible options are

$$\int_{I_j} \left\|f(s) - f^j\right\|_{\infty,\Omega} \mathrm{d}s \approx \frac{\tau_j}{2} \left\|f^{j-1} - f^j\right\|_{\infty,\Omega} \qquad \text{trapezoidal rule,}$$

$$\int_{I_j} \left\|f(s) - f^j\right\|_{\infty,\Omega} \mathrm{d}s \approx \frac{\tau_j}{6} \left\{\left\|f^{j-1} - f^j\right\|_{\infty,\Omega} + 4\left\|f^{j-1/2} - f^j\right\|_{\infty,\Omega}\right\} \quad \text{Simpson's rule.}$$

Application of quadrature introduces additional error terms. They are associated with oscillations of the RHS f and of higher order in τ. In the language of a posteriori error for elliptic equations they are referred to as "higher-order terms" and are typically ignored [3].

Numerical results. Table 4.1 displays the results of our test computations for (3.1). The first column contains the number of mesh intervals used on the spatial domain $[0, 1]$. To avoid special effects from uniform meshes, we have chosen the mesh sizes to satisfy $\tau_j = 2\tau_{j-1}$ for $j = 2, 4, 6, \dots, M$. The second column of the table displays the actual errors of the backward

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Euler semidisretisation (4.1). We observe convergence of order 1—each time the number of mesh intervals is doubled, the error is divided by (approximately) two.

Columns 3 and 4 contain the a posteriori error bounds provided by Theorem 4.1 and its efficiency, i.e., the actual error divided by the error estimator. There is a strong correlation between the two. However, the errors are overestimated by a factor of about 700.

In columns 5 and 6 we have the corresponding numbers for Theorem 4.2. It gives sharper bounds than Theorem 4.1, but the efficiency is slightly deteriorating with the logarithm of the mesh size. (Our test problem somewhat favours Theorem 4.2. There are other equations where Theorem 4.1 gives sharper bounds.)

Finally, in the last two columns of Table 4.1 we present our results for Theorem 4.3. It gives sharper bounds than both Theorems 4.1 and 4.2, which had to be expected from its derivation. Moreover, we do not witness any deterioration of the efficiency with a refinement of the mesh. Since the error bound of Theorem 4.3 contains the minimum of two terms, $\eta^j_{\delta U}$ and $\eta^j_{\delta \mathcal{L} U}$, it is interesting to study when which term is active. We will do this in a broader context later.

§4.2. The preceding error bounds all contain a piecewise constant approximation of the RHS f of the PDE. Now we shall involve its piecewise linear interpolation $\hat{f}$. To this end we use $\hat{f}(t) = f^j - (t_j - t)\,\delta_t f^j$ and rewrite the residuum in (4.2) as

$$\Big(\mathcal{K}\big(u - \hat{U}\big)\Big)(t) = \big(f - \hat{f}\big)(t) + \big(t_j - t\big)\,\delta_t\,(\mathcal{L}U - f)^j \qquad t \in I_j\,.$$

In view of (4.1) we set $\delta_t U^0 := f^0 - \mathcal{L}U^0$, introduce

$$\delta_t^2 v^j := \frac{\delta_t v^j - \delta_t v^{j-1}}{\tau_j}\,, \qquad j = 1, \ldots, M,$$

and obtain

$$\Big(\mathcal{K}\big(u - \hat{U}\big)\Big)(t) = \big(f - \hat{f}\big)(t) - \big(t_j - t\big)\,\delta_t^2 U^j\,, \qquad t \in I_j,\;\; j = 1, \ldots, M.$$

Proceeding as before, we get the following theorem:

THEOREM 4.5. *The maximum-norm error of the backward Euler time discretisation* (4.1) *satisfies the a posteriori bound*

$$\big\|u(T) - U^M\big\|_{\infty,\Omega} \le \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T-t_j)}\left(\eta^j_{\hat{f}} + \eta^j_{\delta^2 U}\right)$$

with

$$\eta^j_{\hat{f}} := \kappa_0 \int_{I_j} \Big\|\big(f - \hat{f}\big)(s)\Big\|_{\infty,\Omega}\,\mathrm{d}s\,, \qquad \eta^j_{\delta^2 U} := \frac{\kappa_0 \tau_j^2}{2}\,\big\|\delta_t^2 U^j\big\|_{\infty,\Omega}\,.$$

REMARK 4.6. Again, the integrals composing $\eta_{\hat{f}}$ need to be approximated. This time the trapezoidal rule would always give zero. One possibility is Simpson's rule which gives

$$\int_{I_j} \Big\|\big(f - \hat{f}\big)(s)\Big\|_{\infty,\Omega}\,\mathrm{d}s \approx \frac{2\tau_j}{3}\,\Big\|\big(\hat{f} - f\big)^{j-1/2}\Big\|_{\infty,\Omega} = \frac{\tau_j}{3}\,\Big\|f^j - 2f^{j-1/2} + f^{j-1}\Big\|_{\infty,\Omega}\,.$$

Taking minima locally for each time level $j = 1, \ldots, M$, Theorem 4.3 and Theorem 4.5 can be combined to give the following sharpened result.

TABLE 4.2
Error estimators of Theorems 4.5 and 4.7 for the Euler method applied to (3.1). Simpson's rule is used to approximate $\eta^j_{\hat{f}}$, $\eta^j_{\bar{f}}$.

M	err	Theorem 4.5 est	Theorem 4.5 eff	Theorem 4.7 est	Theorem 4.7 eff
256	$1.045 \cdot 10^{-4}$	$9.900 \cdot 10^{-3}$	1/95	$6.498 \cdot 10^{-3}$	1/62
512	$5.175 \cdot 10^{-5}$	$4.796 \cdot 10^{-3}$	1/93	$3.079 \cdot 10^{-3}$	1/59
1024	$2.575 \cdot 10^{-5}$	$2.360 \cdot 10^{-3}$	1/92	$1.498 \cdot 10^{-3}$	1/58
2048	$1.284 \cdot 10^{-5}$	$1.171 \cdot 10^{-3}$	1/91	$7.389 \cdot 10^{-4}$	1/58
4096	$6.412 \cdot 10^{-6}$	$5.834 \cdot 10^{-4}$	1/91	$3.670 \cdot 10^{-4}$	1/57
8192	$3.204 \cdot 10^{-6}$	$2.911 \cdot 10^{-4}$	1/91	$1.829 \cdot 10^{-4}$	1/57
16384	$1.601 \cdot 10^{-6}$	$1.454 \cdot 10^{-4}$	1/91	$9.130 \cdot 10^{-5}$	1/57
32768	$8.006 \cdot 10^{-7}$	$7.269 \cdot 10^{-5}$	1/91	$4.561 \cdot 10^{-5}$	1/57
65536	$4.002 \cdot 10^{-7}$	$3.634 \cdot 10^{-5}$	1/91	$2.280 \cdot 10^{-5}$	1/57

THEOREM 4.7. *The maximum-norm error of the backward Euler time discretisation (4.1) satisfies the a posteriori bound*

$$\left\|u(T) - U^M\right\|_{\infty,\Omega} \leq \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T-t_j)} \min\left\{\eta^j_{\bar{f}} + \eta^j_{\min}, \eta^j_{\hat{f}} + \eta^j_{\delta^2 U}\right\},$$

with the notation from Theorems 4.1–4.5.

Numerical results and discussions. Table 4.2 contains our results for Theorems 4.5 and 4.7. Both give sharper bounds than Theorems 4.1–4.3. This was expected for Theorem 4.7.

How do the various components of the error estimators behave? Figure 4.1 depicts plots of the four terms $\eta_{\bar{f}}$, $\eta_{\delta\mathcal{L}U}$, $\eta_{\delta U}$, and $\eta_{\delta^2 U}$. We have chosen a uniform mesh as otherwise there would be oscillations because the components are correlated with powers of the local mesh step size. Also the term $\eta_{\hat{f}}$ is omitted because it is of higher order and close to zero. For the same reason graphs of $\eta_{\delta^2 U}$ and $\eta_{\delta^2 U} + \eta_{\hat{f}}$ would be virtually indistinguishable.

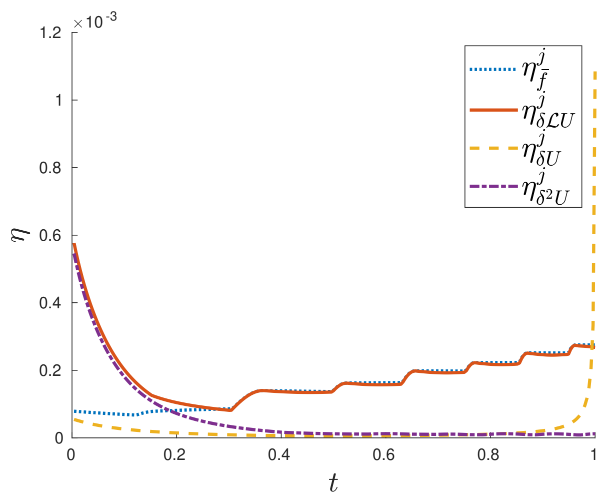

FIG. 4.1. *The various parts of the error estimators in Theorems 4.1–4.7, uniform time stepping, $M = 256$ steps.*

First, we notice that $\eta_{\delta\mathcal{L}U}$ and $\eta_{\delta^2 U}$ attain large values near initial time. Second, $\eta_{\delta U}$ becomes large towards the final time. This can be explained by the behaviour of the ϱ_j in Theorem 4.2. At final time $t_M = T$, we have $\varrho^M = \tau_M$. But further back in time, they become second order: $\varrho^j \sim \tau_j^2$.

Theorems 4.1 and 4.2 differ in the use of $\eta_{\delta\mathcal{L}U}$ (solid red line) and $\eta_{\delta U}$ (dashed yellow line). On most of the domain we have $\eta^j_{\delta\mathcal{L}U} > \eta^j_{\delta U}$, only for the last few steps the relation is reversed. This illustrates how Theorem 4.3 takes advantage by picking the minimum of the two at each time step.

Finally, one notices that for times $t \geq 0.3$ the terms $\eta_{\bar{f}}$ and $\eta_{\delta\mathcal{L}U}$ take very similar values. This suggests that in deriving Theorem 4.1 a triangle inequality might have been applied inadequately. To illustrate this we look at the two representations of the residuum used above:

$$\underbrace{f(s) - f^j}_{\to\, \eta^j_{\bar{f}}} + \underbrace{\big(t_j - s\big)\, \delta_t\, (\mathcal{L}U)^j}_{\to\, \eta^j_{\delta\mathcal{L}U}} = \underbrace{\big(f - \hat{f}\big)(s)}_{\to\, \eta^j_{\hat{f}}} + \underbrace{\big(t_j - s\big)\, \delta_t \big(\mathcal{L}U - f\big)^j}_{\to\, \eta^j_{\delta^2 U}} .$$

Generically, the term $\eta^j_{\hat{f}}$ is of order 3 (in τ_j), while the other three terms are of order 2 only. Therefore, asymptotically we have

$$\eta^j_{\delta^2 U} \leq \big(1 + \mathcal{O}(\tau_j)\big) \left(\eta^j_{\bar{f}} + \eta^j_{\delta\mathcal{L}U}\right) \qquad (\tau_j \to 0)\, .$$

Thus, in general Theorem 4.5 will give sharper bounds than Theorem 4.1. In practice Theorem 4.7 should be given preference as it gives the sharpest error bound.

§4.3. Concluding our study of the backward-Euler scheme, we like to review an idea presented in [13]. The primary intention of the authors was to eliminate the logarithmic dependence on the time step size observed in Theorem 4.2.

Let

$$W^j := \frac{1}{2}\left[\tau_j \delta_t U^j - \tau_M \delta_t U^M\right], \qquad j = 1, \ldots, M.$$

The expectation in [13] was that as j approaches M, the W^j behave similar to $T - t_j$ and therefore compensate for the term $T - s$ in the denominator of the bound φ_1 for $\mathcal{G}_t$. Then,

$$\big(t_j - s\big)\, \delta_t U^j = \frac{\tau_M}{2} \delta_t U^M + W^j + \big(t_{j-1/2} - s\big)\, \delta_t U^j\,, \quad s \in \big(t_{j-1}, t_j\big], \;\; j = 1, \ldots, M.$$

Define

$$\omega(s) := \frac{\big(t_j - s\big)\big(s - t_{j-1}\big)}{2}, \qquad s \in \bar{I}_j, \;\; j = 1, \ldots, M,$$

and note that

$$t_{j-1/2} - s = \omega'(s), \qquad s \in I_j\, .$$

Fix $J \in \{1,\dots,M\}$. Integration by parts for the interval $[t_{J-1}, t_{M-1}]$ applied to the second term on the RHS of (4.3) gives

$$\begin{aligned}
&u(x,T) - U^M(x) \\
&\quad = \sum_{j=1}^{M} \int_{I_j} \left\langle \mathcal{G}(T-s), f(s) - f^j \right\rangle \mathrm{d}s \\
&\qquad + \sum_{j\in\{1,\dots,J,M\}} \int_{I_j} \left(t_j - s\right) \left\langle \partial_t \mathcal{G}(T-s), \delta_t U^j \right\rangle \mathrm{d}s \\
&\qquad - \sum_{j=J}^{M-1} \left\{ \int_{I_j} \omega(s) \left\langle \partial_t^2 \mathcal{G}(T-s), \delta_t U^j \right\rangle \mathrm{d}s - \int_{I_j} \left\langle \partial_t \mathcal{G}(T-s), W^j \right\rangle \mathrm{d}s \right\} \\
&\qquad - \frac{\tau_M}{2} \left\langle \mathcal{G}(T-t_{M-1}) - \mathcal{G}(T-t_{J-1}), \delta_t U^M \right\rangle .
\end{aligned}$$

The first and second integral are estimated as in the derivation of Theorem 4.3. To the third and fourth integral we apply (2.3). The last one is bounded using Hölder's inequality again and (2.2).

THEOREM 4.8. *For any $J \in \{1,\dots,M\}$, the maximum-norm error of the backward Euler time discretisation* (4.1) *satisfies the a posteriori bound*

$$\begin{aligned}
\left\| u(T) - U^M \right\|_{\infty,\Omega} \le \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T-t_j)} \eta_{\bar f}^j + \sum_{j\in\{1,\dots,J,M\}} \mathrm{e}^{-\gamma(T-t_j)} \min\left\{ \eta_{\delta U}^j, \eta_{\delta \mathcal{L} U}^j \right\} \\
+ \sum_{j=J}^{M-1} \mathrm{e}^{-\gamma(T-t_j)} \left\{ \eta_{\delta U,*}^j + \eta_W^j \right\} \\
+ \frac{\kappa_0 \tau_M}{2} \left(\mathrm{e}^{-\gamma(T-t_{M-1})} + \mathrm{e}^{-\gamma(T-t_{J-1})} \right) \left\| \delta_t U^M \right\|_{\infty,\Omega} .
\end{aligned}$$

with $\eta_{\bar f}^j$ and $\eta_{\delta\mathcal{L}U}^j$ from Theorem 4.1 and $\eta_{\delta U}^j$ from Theorem 4.2 and the new terms

$$\eta_{\delta U,*}^j := \left(\kappa_2 \mu_j^* + \frac{\kappa_2' \tau_j^3}{6} \right) \left\| \delta_t U^j \right\|_{\infty,\Omega}, \quad \eta_W^j := \vartheta_j \left\| W^j \right\|_{\infty,\Omega}, \quad \mu_j^* := \int_{I_j} \frac{\omega(s)}{(T-s)^2} \mathrm{d}s,$$

and ϑ_j defined in (2.5).

REMARK 4.9. In [13] the result is derived for $J = 1$ and with only $\eta_{\delta U}^j$ in the second sum instead of $\min\left\{\eta_{\delta U}^j, \eta_{\delta\mathcal{L}U}^j\right\}$.

The drawback of this approach is that in order to compute the W^j, one has to know U^M and U^{M-1}. Hence, one either has to perform two runs for $j = J,\dots,M$, the first to determine $\delta_t U^M$ and the second to compute the W^j, or one needs to store the approximations at those time levels.

Numerical results. Table 4.3 displays our numerical results for Theorem 4.8. We witness a slight improvement over the error bounds of Theorems 4.3 but not over Theorem 4.7.

5. The Crank–Nicolson method. We discretise the abstract parabolic problem (1.1) in time on the mesh $\mathcal{T}_M$ using the second-order Crank–Nicolson method as follows. We associate an approximate solution $U^j \in H_0^1(\Omega)$ with the time level t_j and require it to satisfy

$$\delta_t U^j + \mathcal{L}\hat{U}^{j-1/2} = \hat{f}^{j-1/2} \qquad \text{in } \Omega, \quad j = 1,\dots,M; \quad U^0 = u^0, \tag{5.1}$$

TABLE 4.3
Error estimator of Theorem 4.8, $J = 1$, for the Euler method applied to the test problem (3.1). Simpson's rule is used again to estimate the $\eta_{\bar{f}}^{j}$.

M	err	Theorem 4.8 est	Theorem 4.8 eff
256	$1.045 \cdot 10^{-4}$	$3.596 \cdot 10^{-2}$	1/344
512	$5.175 \cdot 10^{-5}$	$1.778 \cdot 10^{-2}$	1/344
1024	$2.575 \cdot 10^{-5}$	$8.833 \cdot 10^{-3}$	1/343
2048	$1.284 \cdot 10^{-5}$	$4.396 \cdot 10^{-3}$	1/342
4096	$6.412 \cdot 10^{-6}$	$2.190 \cdot 10^{-3}$	1/342
8192	$3.204 \cdot 10^{-6}$	$1.092 \cdot 10^{-3}$	1/341
16384	$1.601 \cdot 10^{-6}$	$5.446 \cdot 10^{-4}$	1/340
32768	$8.006 \cdot 10^{-7}$	$2.716 \cdot 10^{-4}$	1/339
65536	$4.002 \cdot 10^{-7}$	$1.355 \cdot 10^{-4}$	1/338

i.e.,

$$\frac{U^j - U^{j-1}}{\tau_j} + \frac{\mathcal{L}U^j + \mathcal{L}U^{j-1}}{2} = \frac{f^j + f^{j-1}}{2} \qquad \text{in } \Omega, \quad j = 1, \ldots, M; \quad U^0 = u^0 .$$

§5.1. We extend the U^j to a globally defined function using piecewise linear interpolation:

$$\hat{U}(s) = U^j - \left(t_j - s\right)\delta_t U^j = \hat{U}^{j-1/2} + \left(s - t_{j-1/2}\right)\delta_t U^j , \qquad s \in \hat{I}_j , \ \ j = 1, \ldots, M.$$

The residuum of $\hat{U}$ in the PDE admits the representation

$$\Big(\mathcal{K}\big(u - \hat{U}\big)\Big)(s) = f(s) - \partial_t \hat{U}(s) - \mathcal{L}\left(\hat{U}^{j-1/2} + \left(s - t_{j-1/2}\right)\delta_t U^j\right), \qquad s \in I_j .$$

Let $\psi^j := \left(\mathcal{L}U - f\right)^j$. Then by (5.1), we have

$$\partial_t \hat{U}(s) = \delta_t U^j = \hat{f}^{j-1/2} - \mathcal{L}\hat{U}^{j-1/2} = -\hat{\psi}^{j-1/2}$$

for $s \in I_j$. This gives

$$\begin{aligned}\Big(\mathcal{K}\big(u - \hat{U}\big)\Big)(s) &= f(s) - \hat{f}^{j-1/2} + \left(t_{j-1/2} - s\right)\delta_t\left(\mathcal{L}U\right)^j \\ &= f(s) - \hat{f}(s) + \left(t_{j-1/2} - s\right)\delta_t \psi^j , \qquad s \in I_j .\end{aligned}$$

We substitute into (2.1) and obtain

$$\begin{aligned} u(x,T) - U^M(x) = \sum_{j=1}^{M} \bigg\{ &\int_{I_j} \Big\langle \mathcal{G}(T-s), \big(f - \hat{f}\big)(s) \Big\rangle \,\mathrm{d}s \\ &+ \int_{I_j} \left(t_{j-1/2} - s\right) \left\langle \mathcal{G}(T-s), \delta_t \psi^j \right\rangle \mathrm{d}s \bigg\}. \end{aligned} \tag{5.2}$$

To the first integral we apply (2.4). When bounding the second one, note that $\left(t_{j-1/2} - s\right) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}s}\left(t_j - s\right)\left(s - t_{j-1}\right)$. Therefore, we can avail of (2.9) for $k = 1$. We arrive at the following theorem which is a slight modification of the result given in [12, §5, Theorem 5.1].

THEOREM 5.1. *The maximum-norm error of the Crank–Nicolson method* (5.1) *satisfies the a posteriori error bound*

$$\left\|u(T) - U^M\right\|_{\infty,\Omega} \le \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T-t_j)} \left(\eta_{\hat f}^j + \eta_{\delta\psi}^j\right)$$

with $\eta_{\hat f}^j$ *as in Theorem* 4.5,

$$\eta_{\delta\psi}^j := \frac{\Psi_{1,j}}{2}\left\|\delta_t \psi^j\right\|_{\infty,\Omega}, \qquad \psi^j := \left(\mathcal{L}U - f\right)^j,$$

and $\Psi_{1,j}$ *from* (2.9).

§5.2. When studying the backward Euler semidiscretisation, the use of a higher order interpolant of the RHS f turned out to be useful. This time, we define a piecewise quadratic interpolant $\tilde f$ by

$$\tilde f(s) := \hat f(s) + \beta_j \omega(s), \quad s \in \bar I_j, \quad \text{with} \quad \beta_j := -4\frac{f^j - 2f^{j-1/2} + f^{j-1}}{\tau_j^2} \approx -\left(f''\right)^{j-1/2}.$$

It interpolates f at the mesh points of $\mathcal{T}_M$ and at the midpoint of its mesh intervals. Let $\mathcal{L}^{-1}\beta^j := q^j \in H_0^1(\Omega)$ be the unique solution of $\mathcal{L}q^j = \beta^j$. Then,

$$\begin{aligned}\int_{I_j} \omega(s)\left\langle \mathcal{G}(T-s), \beta^j\right\rangle \mathrm{d}s &= \int_{I_j} \omega(s)\left\langle \mathcal{G}(T-s), \mathcal{L}q^j\right\rangle \mathrm{d}s \\ &= -\int_{I_j} \omega(s)\left\langle \partial_t\mathcal{G}(T-s), q^j\right\rangle \mathrm{d}s = -\int_{I_j} \omega'(s)\left\langle \mathcal{G}(T-s), q^j\right\rangle \mathrm{d}s\end{aligned}$$

because $\mathcal{L}^*\mathcal{G} = -\partial_t\mathcal{G}$ and by integration by parts. Then, from (5.2),

$$\begin{aligned}u(x,T) - U^M(x) = \sum_{j=1}^{M}\Bigg\{ &\int_{I_j} \left\langle \mathcal{G}(T-s), (f-\tilde f)(s)\right\rangle \mathrm{d}s \\ &+ \int_{I_j} \omega'(s)\left\langle \mathcal{G}(T-s), \delta_t\psi^j - q^j\right\rangle \mathrm{d}s\Bigg\}.\end{aligned}$$

Using the Hölder inequality, (2.4) and (2.9), we obtain our next result.

THEOREM 5.2. *The maximum-norm error of the Crank–Nicolson method* (5.1) *satisfies the a posteriori bound*

$$\left\|u(T) - U^M\right\|_{\infty,\Omega} \le \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T-t_j)} \left(\eta_{\tilde f}^j + \eta_{\delta\psi q}^j\right)$$

with $q^j \in H_0^1(\Omega)$ *solving* $\mathcal{L}q^j = \beta^j$,

$$\eta_{\tilde f}^j := \kappa_0 \int_{I_j} \left\|\left(f - \tilde f\right)(s)\right\|_{\infty,\Omega} \mathrm{d}s, \qquad \eta_{\delta\psi q}^j := \frac{\Psi_{1,j}}{2}\left\|\delta_t\psi^j - q^j\right\|_{\infty,\Omega},$$

$\psi^j := \left(\mathcal{L}U - f\right)^j$, *and* $\Psi_{1,j}$ *from* (2.9).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 5.1
Error estimators of Theorems 5.1–5.5 for the Crank–Nicolson method applied to test problem (3.1); Simpson's rule used to estimate $\eta_{\hat{f}}$.

M	err	Theorem 5.1 est	Theorem 5.1 eff	Theorem 5.2 est	Theorem 5.2 eff	Theorem 5.5 est	Theorem 5.5 eff
256	$4.284 \cdot 10^{-7}$	$7.221 \cdot 10^{-4}$	1/1686	$1.102 \cdot 10^{-4}$	1/257	$1.102 \cdot 10^{-4}$	1/257
512	$1.048 \cdot 10^{-7}$	$1.806 \cdot 10^{-4}$	1/1724	$2.134 \cdot 10^{-5}$	1/204	$2.033 \cdot 10^{-5}$	1/194
1024	$2.592 \cdot 10^{-8}$	$4.529 \cdot 10^{-5}$	1/1747	$4.744 \cdot 10^{-6}$	1/183	$4.253 \cdot 10^{-6}$	1/164
2048	$6.445 \cdot 10^{-9}$	$1.145 \cdot 10^{-5}$	1/1776	$1.245 \cdot 10^{-6}$	1/193	$1.087 \cdot 10^{-6}$	1/169
4096	$1.607 \cdot 10^{-9}$	$2.875 \cdot 10^{-6}$	1/1789	$3.239 \cdot 10^{-7}$	1/202	$2.750 \cdot 10^{-7}$	1/171
8192	$4.011 \cdot 10^{-10}$	$7.221 \cdot 10^{-7}$	1/1800	$8.631 \cdot 10^{-8}$	1/215	$7.192 \cdot 10^{-8}$	1/179
16384	$1.002 \cdot 10^{-10}$	$1.814 \cdot 10^{-7}$	1/1810	$2.318 \cdot 10^{-8}$	1/231	$1.904 \cdot 10^{-8}$	1/190
32768	$2.503 \cdot 10^{-11}$	$4.557 \cdot 10^{-8}$	1/1821	$6.222 \cdot 10^{-9}$	1/249	$5.053 \cdot 10^{-9}$	1/202
65536	$6.224 \cdot 10^{-12}$	$1.145 \cdot 10^{-8}$	1/1839	$1.666 \cdot 10^{-9}$	1/268	$1.340 \cdot 10^{-9}$	1/215

REMARK 5.3. The integral defining $\eta^j_{\tilde{f}}$ can (in general) not be evaluated exactly but needs to be approximated. For example, Simpson's rule can be applied on the two subintervals $[t_{j-1}, t_{j-1/2}]$ and $[t_{j-1/2}, t_j]$ to give

$$\int_{I_j} \left\| (f - \tilde{f})(s) \right\|_{\infty,\Omega} \mathrm{d}s \approx \frac{\tau_j}{3} \left\{ \left\| (f - \tilde{f})^{j-3/4} \right\|_{\infty,\Omega} + \left\| (f - \tilde{f})^{j-1/4} \right\|_{\infty,\Omega} \right\}.$$

REMARK 5.4. The above choice of a piecewise quadratic interpolation of f corresponds to a piecewise quadratic reconstruction $\tilde{U}(s) = \hat{U}(s) + \mathcal{L}^{-1}\beta^j\omega(s)$ of the approximations U^j.

In [1] the authors also used a special piecewise quadratic reconstruction of the U^j in an a posteriori error analysis but in the context of error estimation in L_2-type norms.

Again, taking minima locally for each time level $j = 1, \ldots, M$, the bounds of the previous two theorems can be combined to give the sharpened result:

THEOREM 5.5. *The maximum-norm error of the Crank–Nicolson method* (5.1) *satisfies the a posteriori bound*

$$\left\| u(T) - U^M \right\|_{\infty,\Omega} \leq \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T-t_j)} \min\left\{ \eta^j_{\hat{f}} + \eta^j_{\delta\psi}, \eta^j_{\tilde{f}} + \eta^j_{\delta\psi q} \right\},$$

with $\eta^j_{\hat{f}}$ as in Theorem 4.5, $\eta^j_{\delta\psi}$ in Theorem 5.1, and $\eta^j_{\tilde{f}}$ and $\eta^j_{\delta\psi q}$ from Theorem 5.2.

Numerical results. Numerical results for the Crank–Nicolson method are given in Table 5.1. For our test problem, the estimator of Theorem 5.1 overestimates the errors by a factor of almost 2000. In contrast, Theorems 5.2 and 5.5 yield sharper error bounds. Of course with Theorem 5.5 giving the best. However, for all three, the efficiency slightly deteriorates as the mesh is refined.

§5.3. Concluding our study of the Crank–Nicolson method, we review an idea presented in [13]. Let

$$W^j_\psi := \frac{1}{12}\left[\tau_j^2 \delta_t \psi^j - \tau_M^2 \delta_t \psi^M\right] \qquad \text{and} \qquad \tilde{\omega}(s) := \omega(s) - \frac{\tau_j^2}{12}, \;\; s \in I_j, \;\; j = 1, \ldots, M.$$

The expectation in [13] was that for $j \to M$, the W_ψ^j behave similar to $T - t_j$ and therefore compensate for the term $T - s$ in the denominator of the bound φ_1 for $\mathcal{G}_t$; see (2.2). Then,

$$\omega(s)\delta_t\psi^j = \frac{\tau_M^2}{12}\delta_t\psi^M + W_\psi^j + \tilde{\omega}(s)\,\delta_t\psi^j\,, \qquad s \in \bar{I}_j,\ \ j = 1,\ldots,M.$$

Define

$$\pi(s) := \int_{t_{j-1}}^{s} \tilde{\omega}(\sigma)\,\mathrm{d}\sigma = \frac{1}{6}\big(t_j - s\big)\big(t_{j-1/2} - s\big)\big(t_{j-1} - s\big)\,, \qquad s \in \bar{I}_j,\ \ j = 1,\ldots,M.$$

Fix $J \in \{1,\ldots,M\}$. Integration by parts applied to parts of the RHS of (5.2) gives

$$\begin{aligned}
&u(x,T) - U^M(x)\\
&\quad = \sum_{j=1}^{M} \int_{I_j} \Big\langle \mathcal{G}(T-s), \big(f - \hat{f}\big)(s)\Big\rangle\,\mathrm{d}s\\
&\qquad\quad + \sum_{j\in\{1,\ldots,J,M\}} \int_{I_j} \omega(s)\,\big\langle \partial_t\mathcal{G}(T-s), \delta_t\psi^j\big\rangle\,\mathrm{d}s\\
&\qquad\quad + \sum_{j=J}^{M-1} \left\{ \int_{I_j} \pi(s)\,\big\langle \partial_t^2\mathcal{G}(T-s), \delta_t\psi^j\big\rangle\,\mathrm{d}s + \int_{I_j} \Big\langle \partial_t\mathcal{G}(T-s), W_\psi^j\Big\rangle\,\mathrm{d}s \right\}\\
&\qquad\quad + \frac{\tau_M^2}{12}\,\big\langle \mathcal{G}(T-t_{M-1}) - \mathcal{G}(T-t_{J-1}), \delta_t\psi^M\big\rangle\,.
\end{aligned}$$

We employ our standard machinery and arrive at the next theorem:

THEOREM 5.6. *For any $J \in \{1,\ldots,M\}$ the maximum-norm error of the Crank–Nicolson method (5.1) satisfies the a posteriori bound*

$$\begin{aligned}
\big\|u(T) - U^M\big\|_{\infty,\Omega} \le \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T-t_j)}\eta_{\hat{f}}^j &+ \sum_{j\in\{1,\ldots,J,M\}} \mathrm{e}^{-\gamma(T-t_j)}\eta_{\delta\psi}^j\\
&+ \sum_{j=J}^{M-1} \mathrm{e}^{-\gamma(T-t_j)} \left\{\eta_{\delta\psi,*}^j + \eta_{W_\psi}^j\right\}\\
&+ \frac{\kappa_0\tau_M^2}{12}\Big(\mathrm{e}^{-\gamma(T-t_{M-1})} + \mathrm{e}^{-\gamma(T-t_{J-1})}\Big)\,\big\|\delta_t\psi^M\big\|_{\infty,\Omega}\,,
\end{aligned}$$

with $\eta_{\hat{f}}^j$ and $\eta_{\delta\psi}^j$ from Theorems 4.5 and 5.1 and the new terms

$$\eta_{\delta\psi,*}^j := \left(\kappa_2\sigma_j^* + \frac{\kappa_2'\tau_j^4}{144}\right)\big\|\delta_t\psi^j\big\|_{\infty,\Omega}\,,\quad \eta_{W_\psi}^j := \vartheta_j\,\Big\|W_\psi^j\Big\|_{\infty,\Omega}\,,\quad \sigma_j^* := \int_{I_j} \frac{|\pi(s)|}{(T-s)^2}\mathrm{d}s.$$

Note, that Remark 4.9 holds accordingly. Numerical results are presented in Table 5.2. They are very similar to those of Theorem 5.1 but worse than Theorems 5.2 and 5.5. Hence, those should be preferred.

6. Extrapolated Euler method. This extrapolation method combines two approximations by the backward Euler-method on the mesh $\mathcal{T}_M$ and on a mesh that is twice as fine. They are defined by

TABLE 5.2
Error estimator of Theorem 5.6, $J = 1$, for the Crank–Nicolson method applied to the test problem (3.1); Simpson's rule to estimate the $\eta^j_{\bar{f}}$.

M	err	Theorem 5.6 est	Theorem 5.6 eff
256	$4.284 \cdot 10^{-7}$	$7.432 \cdot 10^{-4}$	1/1735
512	$1.048 \cdot 10^{-7}$	$1.858 \cdot 10^{-4}$	1/1773
1024	$2.592 \cdot 10^{-8}$	$4.649 \cdot 10^{-5}$	1/1794
2048	$6.445 \cdot 10^{-9}$	$1.163 \cdot 10^{-5}$	1/1805
4096	$1.607 \cdot 10^{-9}$	$2.910 \cdot 10^{-6}$	1/1811
8192	$4.011 \cdot 10^{-10}$	$7.283 \cdot 10^{-7}$	1/1815
16384	$1.002 \cdot 10^{-10}$	$1.822 \cdot 10^{-7}$	1/1818
32768	$2.503 \cdot 10^{-11}$	$4.560 \cdot 10^{-8}$	1/1822
65536	$6.225 \cdot 10^{-12}$	$1.141 \cdot 10^{-8}$	1/1833

One-step Euler: $V^0 = u^0$,

$$\delta_t V^j + \mathcal{L}V^j = f^j, \qquad j = 1, 2, \ldots, M, \tag{6.1a}$$

Two-step Euler: $W^0 = u^0$,

$$\begin{aligned} \frac{W^{j-1/2} - W^{j-1}}{\tau_j/2} + \mathcal{L}W^{j-1/2} &= f^{j-1/2}, \\ \frac{W^j - W^{j-1/2}}{\tau_j/2} + \mathcal{L}W^j &= f^j, \end{aligned} \qquad j = 1, \ldots, M. \tag{6.1b}$$

Extrapolation:

$$U^j := 2W^j - V^j, \qquad j = 1, \ldots, M. \tag{6.1c}$$

We follow [15] and consider a piecewise linear reconstruction $\hat{U}$ of the approximations U^j, $j = 0, 1, \ldots, M$. First, adding the two equations in (6.1b) and subtracting (6.1a) yields

$$\partial_t \hat{U} = \delta_t U^j = 2\delta_t W^j - \delta_t V^j = f^{j-1/2} - \mathcal{L}\left(W^{j-1/2} + W^j - V^j\right) .$$

This implies for the residuum

$$\begin{aligned} \left(\mathcal{K}(u - \hat{U})\right)(s) &= f(s) - \partial_t \hat{U}(s) - \mathcal{L}\hat{U}(s) \\ &= f(s) - f^{j-1/2} + \mathcal{L}\left(W^{j-1/2} - W^j\right) + \mathcal{L}\left(U^j - \hat{U}(s)\right) . \end{aligned} \tag{6.2}$$

Next,

$$U^j - \hat{U}(s) = -(s - t_j)\delta_t U^j = -(s - t_{j-1/2})\delta_t U^j + \frac{\tau_j}{2}\delta_t U^j ,$$

which implies

$$\mathcal{L}\left(U^j - \hat{U}(s)\right) = -(s - t_{j-1/2})\mathcal{L}\delta_t U^j + \frac{1}{2}\mathcal{L}\left(U^j - U^{j-1}\right) .$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

This is substituted into (6.2) to give

$$\left(\mathcal{K}(u-\hat{U})\right)(s) = \big(f-\hat{f}\big)(s) + \hat{f}^{j-1/2} - f^{j-1/2} + (t_{j-1/2}-s)\,\delta_t\big(\mathcal{L}U-f\big)^j + \mathcal{L}\left(W^{j-1/2} - W^{j-1} - \frac{V^j - V^{j-1}}{2}\right).$$

Setting

$$\begin{aligned} Z^j &:= W^{j-1/2} - W^{j-1} - \frac{V^j - V^{j-1}}{2}, \\ F(s) &:= f(s) - f^{j-1/2}, \qquad s \in (t_{j-1}, t_j),\ \ j = 1,\dots,M, \end{aligned} \tag{6.3}$$

and

$$\psi^j := \big(\mathcal{L}U - f\big)^j, \qquad j = 0,\dots,M,$$

the residuum takes the form

$$\left(\mathcal{K}(u-\hat{U})\right)(s) = \big(F-\hat{F}\big)(s) + (t_{j-1/2}-s)\,\delta_t\psi^j + \mathcal{L}Z^j.$$

Then (2.1) yields

$$u(x,T) - U^M(x) = \sum_{j=1}^{M} \left\{ \int_{I_j} \left\langle \mathcal{G}(T-s), \big(F-\hat{F}\big)(s) + \mathcal{L}Z^j \right\rangle \mathrm{d}s + \int_{I_j} \big(t_{j-1/2}-s\big) \left\langle \mathcal{G}(T-s), \delta_t\psi^j \right\rangle \mathrm{d}s \right\}.$$

Using (2.4) and (2.9), we obtain

$$\left| \int_{I_j} \left\langle \mathcal{G}(T-s), \big(F-\hat{F}\big)(s) \right\rangle \mathrm{d}s \right| \le \kappa_0 \mathrm{e}^{-\gamma(T-t_j)} \int_{I_j} \left\| \big(F-\hat{F}\big)(s) \right\|_{\infty,\Omega} \mathrm{d}s,$$

$$\left| \int_{I_j} \big(t_{j-1/2}-s\big) \left\langle \mathcal{G}(T-s), \delta_t\psi^j \right\rangle \mathrm{d}s \right| \le \frac{\Psi_{1,j}}{2} \mathrm{e}^{-\gamma(T-t_j)} \left\| \delta_t\psi^j \right\|_{\infty,\Omega},$$

and

$$\left| \int_{I_j} \left\langle \mathcal{G}(T-s), \mathcal{L}Z^j \right\rangle \mathrm{d}s \right| \le \kappa_0 \tau_j \mathrm{e}^{-\gamma(T-t_j)} \left\| \mathcal{L}Z^j \right\|_{\infty,\Omega}. \tag{6.4}$$

Furthermore,

$$\int_{I_j} \left\langle \mathcal{G}(T-s), \mathcal{L}Z^j \right\rangle \mathrm{d}s = \int_{I_j} \left\langle \mathcal{L}^*\mathcal{G}(T-s), Z^j \right\rangle \mathrm{d}s = -\int_{I_j} \left\langle \partial_t\mathcal{G}(T-s), Z^j \right\rangle \mathrm{d}s$$

gives an alternative bound for (6.4):

$$\left| \int_{I_j} \left\langle \mathcal{G}(T-s), \mathcal{L}Z^j \right\rangle \mathrm{d}s \right| \le \int_{I_j} \varphi_1(T-s)\mathrm{d}s\, \mathrm{e}^{-\gamma(T-t_j)} \left\| Z^j(s) \right\|_{\infty,\Omega}.$$

TABLE 6.1

Error estimators of Theorem 6.1 for the extrapolated Euler method applied to test problem (3.1); Simpson's rule used to estimate $\eta_{\hat{F}}$.

M	err	est	eff
256	$8.780 \cdot 10^{-7}$	$4.302 \cdot 10^{-4}$	1/490
512	$2.214 \cdot 10^{-7}$	$1.058 \cdot 10^{-4}$	1/478
1024	$5.536 \cdot 10^{-8}$	$2.644 \cdot 10^{-5}$	1/478
2048	$1.382 \cdot 10^{-8}$	$6.742 \cdot 10^{-6}$	1/488
4096	$3.448 \cdot 10^{-9}$	$1.699 \cdot 10^{-6}$	1/493
8192	$8.611 \cdot 10^{-10}$	$4.281 \cdot 10^{-7}$	1/497
16384	$2.151 \cdot 10^{-10}$	$1.079 \cdot 10^{-7}$	1/502
32768	$5.369 \cdot 10^{-11}$	$2.721 \cdot 10^{-8}$	1/507
65536	$1.330 \cdot 10^{-11}$	$6.859 \cdot 10^{-9}$	1/516

We arrive at the following theorem.

THEOREM 6.1. *The maximum-norm error of the extrapolated Euler method (6.1) satisfies the a posteriori error bound*

$$\left\|u(T) - U^M\right\|_{\infty,\Omega} \le \eta_{\mathrm{eE}}^M := \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T-t_j)} \left(\eta_{\hat{F}}^j + \eta_{\delta\psi}^j + \eta_Z^j\right),$$

with the Z^j defined in (6.3),

$$\eta_{\hat{F}}^j := \kappa_0 \int_{I_j} \left\|\left(F - \hat{F}\right)(s)\right\|_{\infty,\Omega} \mathrm{d}s\,, \qquad \eta_{\delta\psi}^j := \frac{\Psi_{1,j}}{2} \left\|\delta_t \psi^j\right\|_{\infty,\Omega},$$

$$\eta_Z^j := \min\left\{\kappa_0 \tau_j \left\|\mathcal{L}Z^j\right\|_{\infty,\Omega}, \vartheta_j \left\|Z^j\right\|_{\infty,\Omega}\right\}.$$

REMARK 6.2. The integrals composing $\eta_{\hat{F}}$ need to be approximated. One possibility is Simpson's rule which gives

$$\int_{I_j} \left\|\left(F - \hat{F}\right)(s)\right\|_{\infty,\Omega} \mathrm{d}s \approx \eta_{\hat{F},\mathrm{simp}}^j := \frac{\tau_j}{6} \left\|f^j - 2f^{j-1/2} + f^{j-1}\right\|_{\infty,\Omega}$$
$$\approx \frac{\tau_j^3}{24} \left\|\partial_t^2 f(t_{j-1/2})\right\|_{\infty,\Omega}.$$

REMARK 6.3. Theorem 6.1 can be used to establish an asymptotically exact error estimator for the underlying backward-Euler discretisation:

$$u(t_j) - V^j = u(t_j) - V^j + U^j - U^j = 2\,(W - V)^j + u(t_j) - U^j\,, \qquad j = 0, \dots, M.$$

Application of the triangle inequality gives

$$\left\|u(T) - V^M\right\|_{\infty,\Omega} \le 2\left\|W^M - V^M\right\|_{\infty,\Omega} + \eta_{\mathrm{eE}}^M.$$

Similarly,

$$\left\|u(T) - W^M\right\|_{\infty,\Omega} \le \left\|W^M - V^M\right\|_{\infty,\Omega} + \eta_{\mathrm{eE}}^M.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 6.2
Asymptotically exact error estimation for the backward Euler method according to Remark 6.3.

M	$\left\|u(T) - W^M\right\|_{\infty,\Omega}$	est	eff	$\left\|W^M - V^M\right\|_{\infty,\Omega}$	η_{eE}^M
256	$5.203 \cdot 10^{-5}$	$4.827 \cdot 10^{-4}$	1/9	$5.249 \cdot 10^{-5}$	$4.302 \cdot 10^{-4}$
512	$2.582 \cdot 10^{-5}$	$1.317 \cdot 10^{-4}$	1/5	$2.593 \cdot 10^{-5}$	$1.058 \cdot 10^{-4}$
1024	$1.286 \cdot 10^{-5}$	$3.933 \cdot 10^{-5}$	1/3	$1.289 \cdot 10^{-5}$	$2.644 \cdot 10^{-5}$
2048	$6.417 \cdot 10^{-6}$	$1.317 \cdot 10^{-5}$	1/2	$6.424 \cdot 10^{-6}$	$6.742 \cdot 10^{-6}$
4096	$3.205 \cdot 10^{-6}$	$4.905 \cdot 10^{-6}$	1/2	$3.207 \cdot 10^{-6}$	$1.699 \cdot 10^{-6}$
8192	$1.602 \cdot 10^{-6}$	$2.030 \cdot 10^{-6}$	1/1	$1.602 \cdot 10^{-6}$	$4.281 \cdot 10^{-7}$
16384	$8.006 \cdot 10^{-7}$	$9.087 \cdot 10^{-7}$	1/1	$8.007 \cdot 10^{-7}$	$1.079 \cdot 10^{-7}$
32768	$4.003 \cdot 10^{-7}$	$4.275 \cdot 10^{-7}$	1/1	$4.003 \cdot 10^{-7}$	$2.721 \cdot 10^{-8}$
65536	$2.001 \cdot 10^{-7}$	$2.070 \cdot 10^{-7}$	1/1	$2.001 \cdot 10^{-7}$	$6.859 \cdot 10^{-9}$

Numerical results. Numerical results for the extrapolated Euler method are given in Table 6.1. They are clear illustrations for the bounds given in Theorem 6.1. The efficiency is around 500 but slowly decreasing (with $\ln M$) as the mesh is refined.

Table 6.2 illustrates Remark 6.3. Using extrapolation, an asymptotically exact error estimator for the underlying Euler method is obtained. This kind of error control for initial-value problems is well established (see, e.g., [10, II.4]): A higher-order method is used to estimate the error of a lower-order method. However, this approach does not guarantee upper bounds for the discretisation error because the error of the higher-order method is not controlled. Additional bounds like Theorem 6.1 cure this defect.

7. Discontinuous Galerkin method, dG(1). Given $U^0 = u^0$, we seek approximations $U^{j-2/3}, U^j \in H_0^1(\Omega)$, of $u(t_{j-2/3})$ and $u(t_j)$ as solutions of

$$\begin{aligned} U^{j-2/3} - U^{j-1} + \frac{\tau_j}{12}\left(5\mathcal{L}U^{j-2/3} - \mathcal{L}U^j\right) &= \frac{\tau_j}{12}\left(5f^{j-2/3} - f^j\right), \\ U^j - U^{j-1} + \frac{\tau_j}{4}\left(3\mathcal{L}U^{j-2/3} + \mathcal{L}U^j\right) &= \frac{\tau_j}{4}\left(3f^{j-2/3} + f^j\right), \end{aligned} \quad j = 1, \dots, M. \tag{7.1a}$$

Let $\psi := f - \mathcal{L}U$. Then (7.1) can be rewritten as

$$\begin{aligned} U^{j-2/3} - U^{j-1} &= \frac{\tau_j}{12}\left(5\psi^{j-2/3} - \psi^j\right), \\ U^j - U^{j-1} &= \frac{\tau_j}{4}\left(3\psi^{j-2/3} + \psi^j\right), \end{aligned} \qquad j = 1, \dots, M. \tag{7.2}$$

We summarise the analysis from [12, §6] and set

$$\zeta(s) := 3(s-1)(s-1/3) \qquad \text{and} \qquad Z(s) := \int_0^s \zeta(\sigma)\,\mathrm{d}\sigma = s(s-1)^2,$$

and note that $\zeta'(s) = 6(s - 2/3)$.

Given a function v, we define a piecewise linear (possibly discontinuous) interpolant $\bar{v}$ by

$$\bar{v}(t) := v^j - \frac{3}{2}\frac{t_j - t}{\tau_j}\left(v^j - v^{j-2/3}\right), \qquad t \in (t_{j-1}, t_j],$$

and a continuous piecewise quadratic interpolant $\breve{v}$ by

$$\breve{v}(t) := v^j - \frac{3}{2}\frac{t_j - t}{\tau_j}\left(v^j - v^{j-2/3}\right) + \frac{v^j - 3v^{j-2/3} + 2v^{j-1}}{2}\zeta\left(\frac{t - t_{j-1}}{\tau_j}\right), \quad t \in \bar{I}_j,$$

Then, by (7.2)

$$\breve{U}'(t) = \frac{3\big(U^j - U^{j-2/3}\big)}{2\tau_j} + 3\frac{U^j - 3U^{j-2/3} + 2U^{j-1}}{\tau_j}\frac{t - t_{j-2/3}}{\tau_j} = \bar{\psi}(t).$$

This yields for the residuum

$$\begin{aligned}\mathcal{K}\big(u - \breve{U}\big)(t) &= f(t) - \big(\breve{U}' + \mathcal{L}\breve{U}\big)(t) = \big(f - \breve{f}\big)(t) - \breve{U}'(t) + \breve{\psi}(t) \\ &= \big(f - \breve{f}\big)(t) + \big(\breve{\psi} - \bar{\psi}\big)(t) \\ &= \big(f - \breve{f}\big)(t) + \frac{\psi^j - 3\psi^{j-2/3} + 2\psi^{j-1}}{2}\zeta\left(\frac{t - t_{j-1}}{\tau_j}\right), \qquad t \in \bar{I}_j.\end{aligned}$$

Set

$$\chi^j := \frac{\psi^j - 3\psi^{j-2/3} + 2\psi^{j-1}}{2\tau_j^2}, \qquad j = 1, \dots, M. \tag{7.3}$$

Then the residuum can be rewritten into

$$\begin{aligned}\Big(\mathcal{K}\big(u - \breve{U}\big)\Big)(t) &= \big(f - \breve{f}\big)(t) + 3\chi^j\big(t - t_j\big)\big(t - t_{j-2/3}\big) \\ &= \big(f - \breve{f}\big)(t) + \chi^j \frac{\mathrm{d}}{\mathrm{d}t}\left[\big(t - t_j\big)^2\big(t - t_{j-1}\big)\right], \qquad t \in \bar{I}_j,\end{aligned}$$

where we have used integration by parts. Next, we multiply by the Green's function and integrate over $(0, T)$ to obtain the following a posteriori error bound:

THEOREM 7.1. *The error of the dG(1) method* (7.1) *satisfies*

$$\left\|u(T) - U^M\right\|_{\infty,\Omega} \le \sum_{j=1}^{M} \mathrm{e}^{-\gamma(T - t_j)}\left(\eta_{\breve{f}}^j + \eta_\chi^j\right),$$

with χ^j from (7.3),

$$\eta_{\breve{f}}^j := \kappa_0 \int_{I_j} \left\|\big(f - \breve{f}\big)(s)\right\|_{\infty,\Omega} \mathrm{d}s \qquad \textit{and} \qquad \eta_\chi^j := \Psi_{2,j}\left\|\chi^j\right\|_{\infty,\Omega}.$$

This result is a slight improvement over Theorem 6.1 in [12] as it employs local bounds for the Green's function rather then a global argument. An a posteriori error bound for the dG(1)-method is also given in [7, §1, Theorem 1.3] but without a proof and without fixing the constants. Furthermore, a remark in [7] suggests this bound is only second-order time accurate, while Theorem 7.1 provides a bound of order 3.

REMARK 7.2. Again, the integral defining $\eta_{\breve{f}}^j$ needs to be approximated. Simpson's rule can be applied to give

$$\begin{aligned}\int_{I_j} \left\|\big(f - \breve{f}\big)(s)\right\|_{\infty,\Omega} \mathrm{d}s &\approx \frac{2\tau_j}{3}\left\|\big(\breve{f} - f\big)\big(t_{j-1/2}\big)\right\|_{\infty,\Omega} \\ &= \frac{2\tau_j}{3}\left\|\frac{f^j + 9f^{j-2/3} - 2f^{j-1}}{8} - f^{j-1/2}\right\|_{\infty,\Omega} =: f_{\breve{f},\mathrm{simp}}^j.\end{aligned}$$

Numerical results. Numerical results for the dG(1)-method are presented in Table 7.1. The results are in agreement with Theorem 7.1. Again, looking at $M = 2^{10}, \dots, 2^{14}$, we witness a slight deterioration (with $\ln M$) when the mesh is refined. For larger M we are operating close to machine accuracy and the results get erratic.

TABLE 7.1
Error estimator of Theorem 7.1 for dG(1) applied to the test problem (3.1).

M	err	est	eff
256	$6.799 \cdot 10^{-8}$	$5.739 \cdot 10^{-6}$	1/84
512	$9.859 \cdot 10^{-9}$	$7.270 \cdot 10^{-7}$	1/74
1024	$1.296 \cdot 10^{-9}$	$9.225 \cdot 10^{-8}$	1/71
2048	$1.631 \cdot 10^{-10}$	$1.170 \cdot 10^{-8}$	1/72
4096	$2.032 \cdot 10^{-11}$	$1.481 \cdot 10^{-9}$	1/73
8192	$2.531 \cdot 10^{-12}$	$1.872 \cdot 10^{-10}$	1/74
16384	$3.169 \cdot 10^{-13}$	$2.364 \cdot 10^{-11}$	1/75
32768	$4.178 \cdot 10^{-14}$	$2.984 \cdot 10^{-12}$	1/71
65536	$2.645 \cdot 10^{-14}$	$3.771 \cdot 10^{-13}$	1/145

8. BDF-2.

The backward differentiation formulae (BDF-k) are a family of multistep methods for the approximation of initial-(boundary) value problems and commonly used for stiff problems. Here we restrict ourselves to the simplest BDF-2 version; higher-order BDF-methods are studied in [17], too.

Given $U^0 = u^0$, we seek approximations $U^j \in H_0^1(\Omega)$ of $u(t_j)$ as solutions of

$$\delta_t U^1 + \mathcal{L}U^1 = f^1 \tag{8.1a}$$

$$D_t U^j + \mathcal{L}U^j = f^j, \qquad j = 2, 3, \dots, M, \tag{8.1b}$$

where

$$D_t v^n := \delta_t v^n + \tau_n \delta_t^2 v^n, \qquad \delta_t^2 v^n := \frac{\delta_t v^n - \delta_t v^{n-1}}{\tau_n + \tau_{n-1}} \qquad \text{and} \qquad \delta_t v^n := \frac{v^n - v^{n-1}}{\tau_n}.$$

Again, we extend the U^j to a piecewise linear function $\hat{U}$ defined on $[0, T]$.

On the first interval, the discretisation (8.1a) consists of a single step of the implicit Euler method (4.1). In view of our discussions following Theorem 4.7, we use the argument that led to Theorem 4.3.

For $s \in (t_{j-1}, t_j)$, $j = 2, 3, \dots, M$, the residuum satisfies

$$\begin{aligned}\mathcal{K}\big(u - \hat{U}\big)(s) &= f(s) - \partial_t \hat{U}(s) - \mathcal{L}\hat{U}(s) \\ &= \big(f - \hat{f}\big)(s) - \delta_t U^j + \big(f - \mathcal{L}U\big)^j + \frac{\big(f - \mathcal{L}U\big)^j - \big(f - \mathcal{L}U\big)^{j-1}}{\tau_j}(s - t_j)\,.\end{aligned}$$

By (8.1b) we have

$$\big(f - \mathcal{L}U\big)^j = \begin{cases} \delta_t U^j, & j = 1, \\ \delta_t U^j + \tau_j \delta_t^2 U^j, & j = 2, \dots, M. \end{cases}$$

Thus,

$$\begin{aligned}\mathcal{K}\big(u - \hat{U}\big)(s) = \big(f - \hat{f}\big)(s) &+ 2\left(s - t_{j-1/2}\right)\delta_t^2 U^j \\ &+ (s - t_j)\frac{\tau_{j-1}}{\tau_j}\left(\delta_t^2 U^j - \delta_t^2 U^{j-1}\right), \end{aligned} \qquad \begin{cases} s \in (t_{j-1}, t_j), \\ j = 3, \dots, M, \end{cases}$$

and

$$\mathcal{K}\big(u - \hat{U}\big)(s) = \big(f - \hat{f}\big)(s) + 2\left(s - t_{j-1/2}\right)\delta_t^2 U^j + (s - t_j)\frac{\tau_{j-1}}{\tau_j}\delta_t^2 U^j\,, \quad s \in (t_1, t_2)$$

TABLE 8.1
Error estimator of Theorem 8.1 for BDF-2 applied to the test problem (3.1).

M	err	est	eff
256	$1.1943 \cdot 10^{-6}$	$1.4205 \cdot 10^{-3}$	1/1189
512	$2.9418 \cdot 10^{-7}$	$4.0910 \cdot 10^{-4}$	1/1391
1024	$7.2959 \cdot 10^{-8}$	$1.0916 \cdot 10^{-4}$	1/1496
2048	$1.8164 \cdot 10^{-8}$	$1.7075 \cdot 10^{-5}$	1/940
4096	$4.5315 \cdot 10^{-9}$	$4.3742 \cdot 10^{-6}$	1/965
8192	$1.1317 \cdot 10^{-9}$	$1.1241 \cdot 10^{-6}$	1/993
16384	$2.8277 \cdot 10^{-10}$	$2.9030 \cdot 10^{-7}$	1/1027
32768	$7.0673 \cdot 10^{-11}$	$7.4963 \cdot 10^{-8}$	1/1061
65536	$1.7642 \cdot 10^{-11}$	$1.9242 \cdot 10^{-8}$	1/1091

Multiplying with $\mathcal{G}(T-s)$ and using both (2.4) and (2.9), we obtain the following result:

THEOREM 8.1. *The maximum-norm error of the BDF-2 discretisation (8.1) satisfies the a posteriori bound*

$$\begin{aligned}\left\|u(T)-U^M\right\|_{\infty,\Omega} \le\ & \mathrm{e}^{-\gamma(T-t_1)}\left(\eta^1_{\bar f}+\min\left\{\eta^1_{\delta U},\eta^1_{\delta\mathcal{L}U}\right\}\right)\\ &+\mathrm{e}^{-\gamma(T-t_2)}\left(\eta^2_{\hat f}+\left(\Psi_{1,2}+\kappa_0\frac{\tau_1\tau_2}{2}\right)\left\|\delta_t^2U^2\right\|_{\infty,\Omega}\right)\\ &+\sum_{j=3}^{M}\mathrm{e}^{-\gamma(T-t_j)}\Bigg(\eta^j_{\hat f}+\Psi_{1,j}\left\|\delta_t^2U^j\right\|_{\infty,\Omega}\\ &\qquad+\kappa_0\frac{\tau_{j-1}\tau_j}{2}\left\|\delta_t^2U^j-\delta_t^2U^{j-1}\right\|_{\infty,\Omega}\Bigg),\end{aligned}$$

with $\eta^j_{\bar f}$ and $\eta^j_{\delta\mathcal{L}U}$ from Theorem 4.1 and $\eta^j_{\delta U}$ from Theorem 4.2.

REMARK 8.2. The term $\delta_t^2U^j-\delta_t^2U^{j-1}$ is a difference quotient of order 3. For a BDF-k method the technique developed in [17] involves difference quotients of order $2k-1$. Also note that in the above analysis we had to consider the first two time steps separately. For the BDF-k method different arguments will be required for the first $2(k-1)$ steps.

Numerical results. Numerical results for the BDF-2 method are given in Table 8.1. There is a jump in the efficiency when going from $M=2^{10}$ to $M=2^{11}$ that we do not have an explanation for. Apart from this, a slight deterioration (with $\ln M$) is observed again when the mesh is refined.

9. Summary and open questions. In this paper we have reexamined (and improved) a posteriori error bounds for semidiscretisations of parabolic PDEs. In particular we have considered

- the backward Euler method,
- the Crank–Nicolson method,
- the extrapolated Euler method,
- the discontinuous Galerkin method with polynomial degree 1, dG(1),
- the BDF-2 method.

Numerical experiments have been conducted for those methods. They showed that the error are overestimated by a factor ranging from 50 to 1000. A natural question that arises is: *Can these estimates be improved to give sharper error bounds.* Ideally, one likes the efficiency of the estimators to be close to 1.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

In the framework of a posteriori error estimation and adaptivity for elliptic equations, lower bounds for the error turned out to be useful; see [3]. Typically, error estimation for elliptic problems is presented in Sobolev spaces. In contrast, we are considering the errors in L_∞, a Banach space with less structure! We are not aware of any lower a posteriori error bounds in L_∞. To the best of our knowledge this is still an important open problem. Also, the design and convergence in L_∞ of adaptive strategies remains open.

But there are further questions that may be of interest:

- Richardson extrapolation: *Is there an elegant way to derive error bounds for extrapolation of arbitrary order in a common framework?*
- Discontinuous Galerkin: The technique derived in [12, §6] for the dG(r) method gives a posteriori bounds with time accuracy of order $r+2$, while the method converges with order $2r+1$. Thus for $r \geq 2$ there is a discrepancy, and the efficiency of the estimator decays with the number of time steps (to the power of $r-1$). *Is there an alternative analysis that gives efficient a posteriori estimators for the dG(r) methods?*
- The backward differentiation formulae (BDF-k): As noted in Remark 8.2, the estimators derived in [17] involve difference quotients of order $2k-1$ while $k+1$ seems to be the natural order. Further complications arise from the necessity to have k starting values. Again: *Is there an elegant way to derive error bounds for BDF methods of arbitrary order in a common framework?*
- Continuous Galerkin: Except for the special case of the Crank–Nicolson method no results are available yet.
- Finally, error estimation for operator splitting methods and ADI methods might be interesting as these are particularly efficient methods for parabolic equations.

REFERENCES

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

[2] E. Bänsch, F. Karakatsani, and Ch. Makridakis, *A posteriori error control for fully discrete Crank–Nicolson schemes*, SIAM J. Numer. Anal., 50 (2012), pp. 2845–2872.

[3] C. Carstensen, M. Feischl, M. Page, and D. Praetorius, *Axioms of adaptivity*, Comput. Math. Appl., 67 (2014), pp. 1195–1253.

[4] E. B. Davies, *Non-Gaussian aspects of heat kernel behaviour*, J. London Math. Soc. (2), 55 (1997), pp. 105–125.

[5] A. Demlow, O. Lakkis, and Ch. Makridakis, *A posteriori error estimates in the maximum norm for parabolic problems*, SIAM J. Numer. Anal., 47 (2009), pp. 2157–2176.

[6] A. Demlow and Ch. Makridakis, *Sharply local pointwise a posteriori error estimates for parabolic problems*, Math. Comp., 79 (2010), pp. 1233–1262.

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

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

[9] A. Friedman, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, Englewood Cliffs, 1964.

[10] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations. I*, 2nd ed., Springer, Berlin, 1993.

[11] N. Kopteva and T. Linss, *Maximum norm a posteriori error estimation for a time-dependent reaction-diffusion problem*, Comput. Methods Appl. Math., 12 (2012), pp. 189–205.

[12] ———, *Maximum norm a posteriori error estimation for parabolic problems using elliptic reconstructions*, SIAM J. Numer. Anal., 51 (2013), pp. 1494–1524.

[13] ———, *Improved maximum-norm a posteriori error estimates for linear and semilinear parabolic equations*, Adv. Comput. Math., 43 (2017), pp. 999–1022.

[14] O. Lakkis and Ch. Makridakis, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp., 75 (2006), pp. 1627–1658.

[15] T. LINSS AND G. RADOJEV, *Maximum-norm a posteriori error bounds for an extrapolated Euler/finite element discretisation of parabolic equations*, Preprint on arXiv, 2022. https://arxiv.org/abs/2208.08153

[16] CH. MAKRIDAKIS AND R. H. NOCHETTO, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 1585–1594.

[17] M. OSSADNIK, *A posteriori Fehlerschätzer für BDF-Verfahren*, Master Thesis, FernUniversität in Hagen, Hagen 10/2022. In German.