

## MAXIMUM NORM A POSTERIORI ERROR ESTIMATION FOR PARABOLIC PROBLEMS USING ELLIPTIC RECONSTRUCTIONS\*

NATALIA KOPTEVA<sup>†</sup> AND TORSTEN LINSS<sup>‡</sup>

**Abstract.** A semilinear second-order parabolic equation is considered in a regular and a singularly perturbed regime. For this equation, we give computable a posteriori error estimates in the maximum norm. Semidiscrete and fully discrete versions of the backward Euler, Crank–Nicolson, and discontinuous Galerkin dG( $r$ ) methods are addressed. For their full discretizations, we employ elliptic reconstructions that are, respectively, piecewise-constant, piecewise-linear, and piecewise-quadratic for  $r = 1$  in time. We also use certain bounds for the Green’s function of the parabolic operator.

**Key words.** a posteriori error estimate, maximum norm, singular perturbation, elliptic reconstruction, backward Euler, Crank–Nicolson, discontinuous Galerkin, parabolic equation, reaction-diffusion

**AMS subject classifications.** 65M15, 65M60

**DOI.** 10.1137/110830563

**1. Introduction.** Consider a semilinear parabolic equation in the form

$$(1.1a) \quad \mathcal{M}u := \partial_t u + \mathcal{L}u + f(x, t, u) = 0 \quad \text{for } (x, t) \in Q := \Omega \times (0, T]$$

with a second-order linear elliptic operator  $\mathcal{L} = \mathcal{L}(t)$  in a spatial domain  $\Omega \subset \mathbb{R}^n$  with Lipschitz boundary, subject to the initial and Dirichlet boundary conditions

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

We assume that  $f$  is continuous in  $\bar{\Omega} \times [0, T] \times \mathbb{R}$ , is differentiable in the third argument, and, for some nonnegative constants  $\gamma$  and  $\bar{\gamma}$ , satisfies

$$(1.2) \quad 0 \leq \gamma^2 \leq \partial_z f(x, t, z) \leq \bar{\gamma}^2 \quad \text{for } (x, t, z) \in \bar{\Omega} \times [0, T] \times \mathbb{R}.$$

The purpose of this paper is to obtain computable a posteriori error estimates for fully discrete methods applied to problem (1.1). We consider the first-order backward Euler and the second-order Crank–Nicolson discretizations in time. Furthermore, we study the discontinuous Galerkin method dG( $r$ ),  $r \geq 1$ , with Radau quadrature.

These results are applied to the model equation with  $\mathcal{L} := -\varepsilon^2 \Delta = -\varepsilon^2 \sum_{i=1}^n \partial_{x_i}^2$ :

$$(1.3) \quad \mathcal{M}u := \partial_t u - \varepsilon^2 \Delta u + f(x, t, u) = 0$$

posed in a bounded polyhedral spatial domain  $\Omega \subset \mathbb{R}^n$  with  $n = 1, 2, 3$ . This equation will be considered in two regimes:

$$(i) \quad \varepsilon = 1, \quad \gamma \geq 0; \quad (ii) \quad \varepsilon \ll 1, \quad \gamma > 0.$$

\*Received by the editors April 13, 2011; accepted for publication (in revised form) March 20, 2013; published electronically May 22, 2013. This research was conducted with the financial support of Science Foundation Ireland under Research Frontiers Programme 2008 grant 08/RFP/MTH1536. <http://www.siam.org/journals/sinum/51-3/83056.html>

<sup>†</sup>Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland (natalia.kopteva@ul.ie).

<sup>‡</sup>Fakultät für Mathematik und Informatik, FernUniversität in Hagen, 58095 Hagen, Germany (torsten.linss@fernuni-hagen.de).

Note that regime (ii) yields a singularly perturbed reaction-diffusion equation, whose solutions may exhibit sharp layer phenomena. It is important in this regime that a posteriori error estimates are robust in the sense that any dependence on the small perturbation parameter  $\varepsilon$  should be shown explicitly [21, 25].

We will give error estimates in the *maximum norm*, which is sufficiently strong to capture sharp layers and singularities that may occur, in particular, if problem (1.1) is of singularly perturbed type. Our estimates will be of *interpolation type* in the sense that they will include certain terms that may be interpreted as approximating  $\tau_j^p |\partial_t^p u|$ , where  $p$  and  $\tau_j$  are the discretization order and local step size in time, respectively.

We employ the *elliptic reconstruction* technique, which was introduced in the recent papers [22, 19, 6] as a counterpart of the Ritz-projection in the a posteriori error estimation for parabolic problems. We also use certain bounds for the *Green’s function* of the continuous parabolic operator in a manner similar to [6], only for a more general semilinear parabolic operator of (1.3) (compared to  $\partial_t - \Delta$  in [6]).

One distinctive feature of our analysis in this paper (compared, e.g., to [1, 6]) is that we use computed solutions and elliptic reconstructions that are piecewise-polynomial of degree  $p - 1$  in time, where  $p$  is the time discretization order. In particular, they are *piecewise-constant* in time when dealing with the first-order backward Euler method and *piecewise-linear* and *-quadratic*, respectively, when dealing with the second-order Crank–Nicolson method, and the third-order dG(1) method. Consequently, we allow the residuals of computed solutions, as well as other functions, to be understood as *distributions*; this inclusion plays a crucial role in our analysis.

Note that earlier pointwise/maximum norm a posteriori error estimates for parabolic equations either are given for regular linear problems [9, 3, 6, 7] or are not robust in the sense that they involve negative powers of  $\varepsilon$  [3]. For a more detailed comparison of our results with various earlier a posteriori error estimates, we refer the reader to Remarks 5.3, 9.5, 9.9, and 11.7 below.

The paper is organized as follows. In section 2, we introduce the Green’s function and obtain a certain stability lemma, which is the key ingredient of our a posteriori error analysis. The contents of sections 3–6 and 8–11 are summarized in the table below, while section 7 looks into elliptic a posteriori error estimators.

	Summary of results	Backward Euler	Crank–Nicolson	dG( $r$ )
Semidiscretizations	section 3	section 4	section 5	section 6
Full discretizations	section 8	section 9	section 10	section 11

*Notation.* Throughout the paper,  $C$ , as well as  $c$ , denotes a generic positive constant that may take different values in different formulas but is *independent of the diffusion coefficient  $\varepsilon$  and any mesh sizes*. We use  $|x|$  for the Euclidean norm of  $x \in \mathbb{R}^n$ . The usual spaces  $C(\bar{\Omega})$  and  $H_0^1(\Omega)$  are used, as well as the spaces  $L_p(\Omega)$ ,  $1 \leq p \leq \infty$ , with the norm  $\|\cdot\|_{p,\Omega}$ , while  $\langle \phi, \psi \rangle = \int_{\Omega} \phi(x)\psi(x) dx$  denotes the inner product in  $L_2(\Omega)$ .

*Distributions and left-continuity convention.* Certain functions will be understood as distributions [13], which will in most cases be indicated. By contrast, if a certain function is Lebesgue-integrable in  $\Omega \times (0, T)$ , we shall refer to it as a regular function. Whenever we deal with a regular function, it will be understood as *left-continuous* for all  $t \in (0, T]$ . In particular, this convention will be applied to all piecewise-continuous temporal derivatives.

**2. The Green's function of the parabolic operator.** In this section we consider the Green's function  $\mathcal{G}$  associated with the operator  $\mathcal{M}$  of (1.1). Our interest in the Green's function is in that it will be used to express the error of a numerical approximation in terms of its residual.

For definitions and properties of fundamental solutions and Green's functions of parabolic operators with variable coefficients, we refer the reader to [12, Chapter 1 and section 7 of Chapter 3]. For any pair of bounded functions  $v$  and  $w$  that vanish on  $\partial\Omega$ , the standard linearization yields  $\mathcal{M}v - \mathcal{M}w = [\partial_t + \mathcal{L} + a(x, t)](v - w)$ , where  $a(x, t) := \int_0^1 \partial_z f(x, t, w + z[v - w]) dz$ . Hence, the difference  $v - w$  is represented as

$$(2.1) \quad [v - w](x, t) = \int_{\Omega} \mathcal{G}(x, t; \xi, 0) [v - w](\xi, 0) d\xi + \int_0^t \int_{\Omega} \mathcal{G}(x, t; \xi, s) [\mathcal{M}v - \mathcal{M}w](\xi, s) d\xi ds$$

with the help of the Green's function  $\mathcal{G}$  that we now define. For fixed  $(x, t) \in Q$ , the Green's function  $\mathcal{G}(x, t; \xi, s) =: \Gamma(\xi, s)$  solves the adjoint terminal-value problem

$$(2.2a) \quad [-\partial_s - \mathcal{L}^* + a(\xi, s)] \Gamma(\xi, s) = 0 \quad \text{for } (\xi, s) \in \Omega \times [0, t),$$

$$(2.2b) \quad \Gamma(\xi, t) = \delta(\xi - x) \quad \text{for } \xi \in \Omega,$$

$$(2.2c) \quad \Gamma(\xi, s) = 0 \quad \text{for } (\xi, s) \in \partial\Omega \times [0, t].$$

Here  $\delta(\cdot)$  is the Dirac  $\delta$ -distribution in  $\mathbb{R}^n$  [13], and  $\mathcal{L}^*$  is the adjoint operator to the linear operator  $\mathcal{L}$ .

The analysis in this paper will be carried out under the following condition.

CONDITION 2.1. *There are constants  $\kappa_0, \kappa_1 > 0$  and  $\kappa_2 \geq 0$  such that the Green's function  $\mathcal{G}$  of (2.2), (1.2) satisfies*

$$\|\mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} \leq \kappa_0 e^{-\gamma^2(t-s)}, \quad \int_0^{t-\tau} \|\partial_s \mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} ds \leq \kappa_1 \ell(\tau, t) + \kappa_2,$$

where  $x \in \Omega$ ,  $\tau \in (0, t]$ ,  $t \in (0, T]$ , and  $\ell(\tau, t) := \int_{\tau}^t s^{-1} e^{-\frac{1}{2}\gamma^2 s} ds \leq \ln(t/\tau)$ .

Note that our model problem satisfies this condition as follows.

LEMMA 2.2. *Let  $\varepsilon \in (0, 1]$  and  $\gamma \geq 0$ . Under assumption (1.2), the model problem (1.3) satisfies Condition 2.1 with  $\kappa_0 := 1$ ,  $\kappa_1 := \frac{3^n}{2^{n/2+1}}$ , and an  $\varepsilon$ -independent constant  $\kappa_2 \geq 0$ . If  $f(x, t, z) = a(x)z + b(x, t)$ , then  $\kappa_2 = 0$ . In general,  $\kappa_2 = (\bar{\gamma}^2 - \gamma^2) \hat{\kappa}_2$ , where  $\hat{\kappa}_2 = \hat{\kappa}_2(\gamma)$  if  $\gamma > 0$  and  $\hat{\kappa}_2 = \hat{\kappa}_2(T)$  if  $\gamma = 0$ .<sup>1</sup>*

*Proof.* We defer the proof to section 12.  $\square$

Condition 2.1 will be employed by means of the following lemma, which plays a crucial role in our analysis. The lemma is formulated in the context of an arbitrary nonuniform mesh in the time direction

$$(2.3) \quad 0 = t_0 < t_1 < t_2 < \dots < t_M = T \quad \text{with } \tau_j = t_j - t_{j-1} \quad \text{for } j = 1, \dots, M.$$

LEMMA 2.3. *Suppose the parabolic operator  $\mathcal{M}$  of (1.1) satisfies (1.2) and Condition 2.1, and  $v, w$  are bounded in  $\bar{\Omega} \times [0, T]$ . Furthermore, let  $v(\cdot, t), w(\cdot, t) \in H_0^1(\Omega) \cap C(\bar{\Omega})$  for  $t \in [0, T]$ , and*

<sup>1</sup>The constants  $\kappa_0$  and  $\kappa_1$  given by Lemma 2.2 are reasonably sharp. For example, for the constant-coefficient version  $\partial_t u - \varepsilon^2 \partial_x^2 u + \gamma^2 u = b(x, t)$  of (1.3) in the spatial domain  $\Omega := \mathbb{R}$ , a calculation [16] yields  $\|\mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} = e^{-\gamma^2(t-s)}$  and  $\|\partial_s \mathcal{G}(x, t; \cdot, s)\|_{1, \Omega} \leq (\sqrt{2(\pi\varepsilon)^{-1}}(t-s)^{-1} + \gamma^2) e^{-\gamma^2(t-s)}$ , so Condition 2.1 is satisfied with  $\kappa_0 = 1$  (as in Lemma 2.2),  $\kappa_1 = \sqrt{2(\pi\varepsilon)^{-1}} \approx 0.48$ ,  $\kappa_2 = 1$ , while Lemma 2.2 gives  $\kappa_1 = 3 \cdot 2^{-3/2} \approx 1.06$ .

$$(2.4) \quad \mathcal{M}v - \mathcal{M}w = \partial_t \mu + \vartheta \quad \text{in } Q,$$

where the function  $\mu$  is continuous and bounded on  $[t_0, t_1]$  and each  $(t_{j-1}, t_j]$ , while  $\partial_t \mu$  is continuous and bounded on  $(t_{m-1}, t_m]$  for some  $1 \leq m \leq M$ , and  $\|\vartheta(\cdot, s)\|_{\infty, \Omega}$  is integrable w.r.t.  $s$  in  $(0, t_m)$  (possibly, in the sense of distributions). Then

$$(2.5) \quad \begin{aligned} & \| [v - w](\cdot, t_m) \|_{\infty, \Omega} \\ & \leq \kappa_0 e^{-\gamma^2 t_m} \| [v - w - \mu](\cdot, 0) \|_{\infty, \Omega} + (\kappa_1 \ell_m + \kappa_2) \sup_{s \in [0, t_{m-1}]} \| \mu(\cdot, s) \|_{\infty, \Omega} \\ & \quad + \kappa_0 \lim_{s \rightarrow t_{m-1}^+} \| \mu(\cdot, s) \|_{\infty, \Omega} + \kappa_0 \tau_m \sup_{s \in (t_{m-1}, t_m]} \| \partial_s \mu(\cdot, s) \|_{\infty, \Omega} \\ & \quad + \kappa_0 \int_0^{t_m} e^{-\gamma^2(t_m-s)} \| \vartheta(\cdot, s) \|_{\infty, \Omega} ds, \end{aligned}$$

where  $\ell_m = \ell_m(\gamma) := \int_{\tau_m}^{t_m} s^{-1} e^{-\frac{1}{2}\gamma^2 s} ds \leq \ln(t_m/\tau_m)$ .

*Remark 2.4.* The term  $\partial_t \mu$  in the right-hand side of (2.4) is understood in the sense of distributions. A typical  $\mu$  is continuously differentiable in time on each  $(t_{j-1}, t_j]$  and has jumps at  $t \in \{t_j\}_{j=1}^{m-1}$ , but our left-continuity convention allows us to avoid ambiguity when integrating by parts. It may help the reader to consider an equivalent interpretation of such evaluations. For some small positive  $\lambda$ , one can replace  $t_j^+$  by  $t_j + \lambda$  and  $\mu$  by  $\mu_\lambda$  such that  $\mu_\lambda = \mu$  for  $t \in [t_{j-1} + \lambda, t_j]$ , and it is continuous and linear in time on each  $[t_j, t_j + \lambda]$ . Then one deals with a regular function  $\partial_t \mu_\lambda$ , while the final result is obtained by taking the limit as  $\lambda \rightarrow 0^+$ .

Similarly, in all calculations involving  $\Gamma$ , one can initially replace it by a regular function  $\Gamma_\lambda$  obtained using a regular approximation  $\delta_\lambda$  of  $\delta$  in (2.2b), and then let  $\lambda \rightarrow 0^+$ . With regard to the regularity of  $\Gamma$ , Condition 2.1 implies for any  $\tau \in (0, t)$  that  $\partial_s \Gamma \in L_1(\Omega \times [0, t - \tau])$ , while an inspection of the proof of Lemma 2.2 yields a stronger regularity with  $\partial_s \Gamma \in L_2(\Omega \times [0, t - \tau])$ .

*Remark 2.5.* One can easily check that if  $\gamma = 0$ , then  $\ell_m = \ln(t_m/\tau_m)$ . Otherwise, if  $\gamma > 0$ , one has  $\ell_m(\gamma) = E_1(\frac{1}{2}\gamma^2 \tau_m) - E_1(\frac{1}{2}\gamma^2 t_m)$ , where  $E_1(t) = \int_t^\infty s^{-1} e^{-s} ds$ ; so  $\ell_m(\gamma) \leq |\ln(\frac{1}{2}\gamma^2 \tau_m)|$  provided that  $\frac{1}{2}\gamma^2 \tau_m \leq 0.67$ . (This is easily checked by finding the only root  $\approx 0.67$  of the equation  $E_1(s) = |\ln s|$  on  $(0, 1)$ .) Note also that  $\ell_1 = 0$  for any  $\gamma \geq 0$ .

*Proof of Lemma 2.3.* Combining representation (2.1) with the notation  $\Gamma(\xi, s) := \mathcal{G}(x, t_m; \xi, s)$  for the Green's function of (2.2), one gets

$$[v - w](x, t_m) = \langle [v - w](\cdot, 0), \Gamma(\cdot, 0) \rangle + \int_0^{t_m} \langle [\mathcal{M}v - \mathcal{M}w](\cdot, s), \Gamma(\cdot, s) \rangle ds.$$

Here, in view of (2.4), the integral on the right-hand side involves  $\mu$  and  $\vartheta$  and so can be represented as a sum  $J_\mu + J_\vartheta$  of the corresponding integrals, which we consider separately. We use the notation  $\int^{b^+} := \lim_{\beta \rightarrow 0^+} \int^{b+\beta}$  and so split  $J_\mu$  as

$$J_\mu = J_\mu^{(1)} + J_\mu^{(2)} := \int_0^{t_{m-1}^+} \langle \partial_s \mu, \Gamma(\cdot, s) \rangle ds + \int_{t_{m-1}^+}^{t_m} \langle \partial_s \mu, \Gamma(\cdot, s) \rangle ds.$$

Here, for  $J_\mu^{(1)}$ , an integration by parts yields

$$J_\mu^{(1)} = \langle \mu(\cdot, t_{m-1}^+), \Gamma(\cdot, t_{m-1}) \rangle - \langle \mu(\cdot, 0), \Gamma(\cdot, 0) \rangle - \int_0^{t_{m-1}} \langle \mu(\cdot, s), \partial_s \Gamma(\cdot, s) \rangle ds.$$

Consequently, we arrive at

$$\begin{aligned}
 (2.6) \quad [v - w](x, t_m) &= \langle [v - w - \mu](\cdot, 0), \Gamma(\cdot, 0) \rangle - \int_0^{t_{m-1}} \langle \mu(\cdot, s), \partial_s \Gamma(\cdot, s) \rangle ds \\
 &\quad + \langle \mu(\cdot, t_{m-1}^+), \Gamma(\cdot, t_{m-1}) \rangle + \int_{t_{m-1}^+}^{t_m} \langle \partial_s \mu, \Gamma(\cdot, s) \rangle ds \\
 &\quad + \int_0^{t_m} \langle \vartheta(\cdot, s), \Gamma(\cdot, s) \rangle ds,
 \end{aligned}$$

where the last term represents  $J_\vartheta$ . Finally, Condition 2.1 implies that

$$\|\Gamma(\cdot, s)\|_{1,\Omega} \leq \kappa_0 e^{-\gamma^2(t_m-s)} \leq \kappa_0, \quad \int_0^{t_{m-1}} \|\partial_s \Gamma(\cdot, s)\|_{1,\Omega} ds \leq \kappa_1 \ell_m + \kappa_2,$$

so we get the desired result.  $\square$

The following version of Lemma 2.3 involves certain approximations  $\Gamma_h^j$  of  $\Gamma(\cdot, t_j)$ .

LEMMA 2.3\*. *Under conditions of Lemma 2.3, suppose that instead of (2.4) one has  $\mathcal{M}v - \mathcal{M}w = \partial_t \mu + \vartheta + \vartheta_*$ , where  $\vartheta_*(\cdot, t) = \sum_{j=1}^{m-1} \vartheta^j \delta(t - t_j)$  for  $t \in [0, t_m]$ . If there exist some functions  $\{\Gamma_h^j\}_{j=1}^{m-1}$  such that  $\langle \vartheta^j, \Gamma_h^j \rangle = 0$  for  $j = 1, \dots, m-1$ , and  $\sum_{j=1}^{m-1} \tau_j \|\mathcal{H}_j^{-2} \{\Gamma(\cdot, t_j) - \Gamma_h^j\}\|_{1,\Omega} \leq \kappa_3 \ell(\tau, t)$  for some positive weight functions  $\{\mathcal{H}_j\}$  and some constant  $\kappa_3$ , then the statement of Lemma 2.3 remains valid, only with an additional term  $\kappa_3 \ell(\tau, t) \max_{j=1, \dots, m-1} \{\tau_j^{-1} \|\mathcal{H}_j^2 \vartheta^j\|_{\infty,\Omega}\}$  in the final line of (2.5).*

*Proof.* Imitate the proof of Lemma 2.3, and note that now we have (2.6) with an additional term  $\sum_{j=1}^{m-1} \langle \vartheta^j, \Gamma(\cdot, t_j) \rangle = \sum_{j=1}^{m-1} \langle \vartheta^j, \Gamma(\cdot, t_j) - \Gamma_h^j \rangle$ .  $\square$

**3. Summary of results for semidiscrete methods (no spatial discretization).** In this section we describe our results for the abstract parabolic problem (1.1) discretized in time on an arbitrary nonuniform mesh (2.3) using semidiscrete backward Euler, Crank–Nicolson, and discontinuous Galerkin methods.

Let  $u$  solve problem (1.1) with the parabolic operator  $\mathcal{M}$  satisfying (1.2) and let Condition 2.1, and let  $U^j \in H_0^1(\Omega) \cap C(\bar{\Omega})$ , associated with the time level  $t_j$ , solve a corresponding semidiscrete problem with  $U^0 = \varphi$ . Then, for  $m = 1, \dots, M$ , we give a *posteriori error estimates* of the type

$$\begin{aligned}
 (3.1) \quad \|U^m - u(\cdot, t_m)\|_{\infty,\Omega} &\leq C_1(\kappa_1 \ell_m + \kappa_2) \max_{j=1, \dots, m-1} \|\chi^j\|_{\infty,\Omega} + C_2 \kappa_0 \|\chi^m\|_{\infty,\Omega} \\
 &\quad + \kappa_0 \sum_{j=1}^m \int_{t_{j-1}}^{t_j} e^{-\gamma^2(t_m-s)} \|\vartheta(\cdot, s)\|_{\infty,\Omega} ds.
 \end{aligned}$$

The quantities that appear in this estimate are specified by Theorems 4.1 and 5.1 and Corollary 6.3 below, and can be summarized as follows:

	$p$	$\chi^{j+1}$	$\vartheta$	$C_1$	$C_2$
Backward Euler	1	$U^{j+1} - U^j$	$\tilde{\psi} - \psi^j$ on $(t_{j-1}, t_j]$	1	2
Crank–Nicolson	2	$\tau_{j+1}(\psi^{j+1} - \psi^j)$	$\tilde{\psi} - I_{1,t} \tilde{\psi}$	$\frac{1}{8}$	$\frac{1}{2}$
dG(1)-Radau	3	$3\tau_{j+1}(2\psi^j - 3\psi^{j+1/3} + \psi^{j+1})$	$\tilde{\psi} - I_{2,t} \tilde{\psi}$	$\frac{2}{81}$	$\frac{1}{6}$

Here for the evaluation of  $\chi^{j+1}$  and  $\vartheta$  we use

$$(3.2) \quad \psi^{j+\alpha} := \mathcal{L}(t_{j+\alpha}) U^{j+\alpha} + f(\cdot, t_{j+\alpha}, U^{j+\alpha}), \quad \tilde{\psi} := \mathcal{L}(t) \tilde{U} + f(\cdot, t, \tilde{U}),$$

where  $\alpha \in (0, 1]$  is any value for which the approximate solution  $U^{j+\alpha}$  at time  $t_{j+\alpha} := t_j + \alpha\tau_{j+1}$  is available from the definition of the semidiscrete method. Also,  $\tilde{U}$  is a piecewise-polynomial interpolant of the computed solution of degree  $p - 1$ , while  $I_{p-1,t}\tilde{\psi}$  is a piecewise-polynomial interpolant of  $\tilde{\psi}$  of the same degree using the same interpolation points.

*Remark 3.1* (interpolation-type estimates). The quantity  $|\chi^j|$  in (3.1) approximates  $\tau_j^p |\partial_t^p u(\cdot, t_j)|$ . This immediately follows from  $\chi^j = U^j - U^{j-1}$  for the backward Euler method. For the Crank–Nicolson and dG(1) methods, note that  $\psi^{j+\alpha}$  approximates  $\mathcal{L}u + f(\cdot, t, u)$  at  $t = t_{j+\alpha}$ , so  $\chi^j$  approximates  $\tau_j^p |\partial_t^{p-1}(\mathcal{L}u + f(\cdot, t, u))|$  (in fact,  $\chi^j = \tau_j^p \partial_t^{p-1}(I_{p-1,t}\tilde{\psi})$ ), while (1.1) gives  $|\partial_t^{p-1}(\mathcal{L}u + f(\cdot, t, u))| = |\partial_t^p u|$ .

*Remark 3.2* ( $p$ th-order estimates). Remark 3.1 and the definitions of  $\vartheta$  for the backward Euler, Crank–Nicolson, and dG(1) methods imply that (3.1) gives an a posteriori error estimate of order  $p$  with  $p = 1, 2,$  and  $3$ , respectively.

**4. Semidiscrete backward Euler method (no spatial discretization).**

Consider an arbitrary nonuniform mesh (2.3) in the time direction and discretize the abstract parabolic problem (1.1) in time using the first-order backward Euler method as follows. We associate an approximate solution  $U^j \in H_0^1(\Omega) \cap C(\bar{\Omega})$  with the time level  $t_j$  and require it to satisfy

$$(4.1a) \quad \delta_t U^j + \mathcal{L}^j U^j + f^j = 0 \quad \text{in } \Omega, \quad j = 1, \dots, M, \quad U^0 = \varphi,$$

$$(4.1b) \quad \text{where } \delta_t U^j := \frac{U^j - U^{j-1}}{\tau_j}, \quad \mathcal{L}^j := \mathcal{L}(t_j), \quad \text{and } f^j := f(\cdot, t_j, U^j).$$

For this discretization, we give the following a posteriori error estimate.

**THEOREM 4.1.** *Let  $u$  solve problem (1.1) with the parabolic operator  $\mathcal{M}$  satisfying (1.2) and Condition 2.1 and  $U^j$  solve the corresponding semidiscrete problem (4.1). Then, for  $m = 1, \dots, M$ , one has (3.1) with  $\chi^j = U^j - U^{j-1}$ ,  $C_1 = 1$ ,  $C_2 = 2$ , and  $\vartheta$  defined by*

$$(4.2) \quad \vartheta(\cdot, t) = \tilde{\psi}(\cdot, t) - \tilde{\psi}(\cdot, t_j), \quad \tilde{\psi}(\cdot, t) = \mathcal{L}(t)U^j + f(\cdot, t, U^j) \quad \text{for } t \in (t_{j-1}, t_j].$$

*Proof.* Let  $I_{1,t}U$  be the standard piecewise-linear interpolant of  $U^j$  in time:

$$(4.3) \quad I_{1,t}U(\cdot, t) := \frac{t_j - t}{\tau_j} U^{j-1} + \frac{t - t_{j-1}}{\tau_j} U^j \quad \text{for } t \in [t_{j-1}, t_j], \quad j = 1, \dots, M.$$

Furthermore, we define a *piecewise-constant* interpolant  $\tilde{U}$  of  $U^j$  by

$$(4.4) \quad \tilde{U}(\cdot, t) := U^j \quad \text{for } t \in (t_{j-1}, t_j], \quad j = 1 \dots, M, \quad \tilde{U}(\cdot, 0) := U^1$$

(so  $\tilde{U}$  is continuous on  $[t_0, t_1]$ ). Note that the temporal derivative  $\partial_t \tilde{U}$  is understood as a distribution, while  $\partial_t(I_{1,t}U)$  is a regular function, equal to  $\delta_t U^j$  for  $t \in (t_{j-1}, t_j]$  (in agreement with our left-continuity convention). Consequently, (4.1a) implies that

$$(4.5) \quad \partial_t(I_{1,t}U) + \tilde{\psi} = \vartheta \quad \text{for } (x, t) \in Q.$$

Here we also used the observation that by (4.4), the regular function  $\vartheta$  of (4.2) can be rewritten as  $\vartheta = \tilde{\psi} - [\mathcal{L}^j U^j + f^j]$  for  $t \in (t_{j-1}, t_j]$ .

As  $\mathcal{M}\tilde{U} = \partial_t \tilde{U} + \tilde{\psi}$  and  $\mathcal{M}u = 0$ , so (4.5) implies that

$$\mathcal{M}\tilde{U} - \mathcal{M}u = \partial_t[\tilde{U} - I_{1,t}U] + \vartheta \quad \text{in } Q.$$

Now the desired bound for  $U^m - u(\cdot, t_m) = [\tilde{U} - u](\cdot, t_m)$  is obtained by an application of Lemma 2.3 with  $\mu := \tilde{U} - I_{1,t}U$  and  $\vartheta$  of (4.2), using the following two observations. First, we note that  $[\tilde{U} - u - \mu](\cdot, 0) = U^1 - \varphi - (U^1 - \varphi) = 0$ . Second, for  $t \in (t_{j-1}, t_j]$ , one has

$$\mu = \frac{t_i - t}{\tau_j} (U^j - U^{j-1}) = \frac{t_i - t}{\tau_j} \chi^j \implies |\mu| \leq |\chi^j|, \quad \tau_j |\partial_t \mu| = |\chi^j|.$$

This completes the proof.  $\square$

**COROLLARY 4.2.** *Under assumption (1.2), the a posteriori error estimate of Theorem 4.1 applies to the model problem (1.3) with  $\vartheta = f(\cdot, t, U^j) - f(\cdot, t_j, U^j)$  and the constants  $\kappa_0, \kappa_1, \kappa_2$  from Lemma 2.2.*

**5. Semidiscrete Crank–Nicolson method (no spatial discretization).** Consider an arbitrary nonuniform mesh (2.3) in the time direction and discretize the abstract parabolic problem (1.1) in time using the second-order Crank–Nicolson method as follows. We associate an approximate solution  $U^j \in H_0^1(\Omega) \cap C(\bar{\Omega})$  with the time level  $t_j$  and require it to satisfy

$$(5.1a) \quad \delta_t U^j + \frac{1}{2}(\mathcal{L}^{j-1}U^{j-1} + \mathcal{L}^jU^j) + \frac{1}{2}(f^{j-1} + f^j) = 0 \quad \text{in } \Omega, \quad j = 1, \dots, M,$$

where we again let

$$(5.1b) \quad U^0 = \varphi, \quad \delta_t U^j := \frac{U^j - U^{j-1}}{\tau_j}, \quad \mathcal{L}^j := \mathcal{L}(t_j), \quad \text{and} \quad f^j := f(\cdot, t_j, U^j).$$

To give an a posteriori error estimate for this discretization, we will use the standard piecewise linear interpolation  $I_{1,t}$ , which, for any continuous function  $w = w(t)$ , is defined by

$$(5.2) \quad I_{1,t}w(t) := \frac{t_j - t}{\tau_j} w(t_{j-1}) + \frac{t - t_{j-1}}{\tau_j} w(t_j) \quad \text{for } t \in [t_{j-1}, t_j], \quad j = 1, \dots, M.$$

Recall an almost identical definition (4.3) for the *piecewise-linear* interpolant  $I_{1,t}U$  of the computed solution; the latter plays a crucial role in our analysis of this section.

**THEOREM 5.1.** *Let  $u$  solve the problem (1.1) with the parabolic operator  $\mathcal{M}$  satisfying (1.2) and Condition 2.1, and let  $U^j$  solve the corresponding semidiscrete problem (5.1). Then for  $m = 1, \dots, M$ , one has (3.1) with  $\chi^j = \tau_j (\psi^j - \psi^{j-1})$  using  $\psi^j = \mathcal{L}^jU^j + f^j$ ,  $C_1 = \frac{1}{8}$ ,  $C_2 = \frac{1}{2}$ , and  $\vartheta$  defined by*

$$(5.3) \quad \vartheta = \tilde{\psi} - I_{1,t}\tilde{\psi}, \quad \tilde{\psi} = \mathcal{L}(t)\tilde{U} + f(\cdot, t, \tilde{U}), \quad \tilde{U}(\cdot, t) = I_{1,t}U(\cdot, t)$$

for  $t \in [0, T]$ , where we use  $I_{1,t}U(\cdot, t)$  of (4.3) and  $I_{1,t}$  of (5.2).

*Proof.* Let  $t \in [t_{j-1}, t_j]$ . First, note that  $\tilde{\psi} = I_{1,t}\tilde{\psi} + \vartheta = \frac{1}{2}(\psi^{j-1} + \psi^j) + \partial_t \mu + \vartheta$ , where  $\mu := \int_{t_j}^t [I_{1,t}\tilde{\psi} - \frac{1}{2}(\psi^{j-1} + \psi^j)] dt$ , so

$$(5.4) \quad \mu = \tau_j^{-1} \chi^j \int_{t_j}^t (t - t_{j-1/2}) dt = -\frac{1}{2}(t_j - t)(t - t_{j-1}) \cdot \tau_j^{-2} \chi^j \quad \text{for } t \in [t_{j-1}, t_j].$$

Next, note that  $\tilde{U}(\cdot, t) = I_{1,t}U(\cdot, t)$  implies that  $\partial_t \tilde{U} = \delta_t U^j = -\frac{1}{2}(\psi^{j-1} + \psi^j)$  for  $t \in (t_{j-1}, t_j]$  (where we also invoked (5.1a)). Combining these two observations, one deduces that  $\partial_t \tilde{U} + \tilde{\psi} = \partial_t \mu + \vartheta$ . As  $\mathcal{M}\tilde{U} = \partial_t \tilde{U} + \tilde{\psi}$  and  $\mathcal{M}u = 0$ , so

$$(5.5) \quad \mathcal{M}\tilde{U} - \mathcal{M}u = \partial_t \mu + \vartheta \quad \text{in } Q.$$

Both sides in this relation are regular functions; it is valid for  $t \in (0, T]$  as  $\mu$  of (5.4) is continuous for  $t \in [0, T]$ .

Now the desired bound for  $U^m - u(\cdot, t_m) = [\tilde{U} - u](\cdot, t_m)$  is obtained by an application of Lemma 2.3 to (5.5) with  $\mu$  given by (5.4) and  $\vartheta$  by (5.3), using the following two observations. First, note that  $[\tilde{U} - u - \mu](\cdot, 0) = U^0 - \varphi - 0 = 0$ . Second, for  $t \in (t_{j-1}, t_j]$ , one has

$$(5.6) \quad |\mu| \leq \frac{1}{8} |\chi^j| \quad \text{and} \quad \tau_j |\partial_t \mu| \leq \frac{1}{2} |\chi^j|,$$

while  $\mu(\cdot, t_{m-1}^+) = 0$ . This completes the proof.  $\square$

**COROLLARY 5.2.** *Under assumption (1.2), the a posteriori error estimate of Theorem 5.1 applies to the model problem (1.3) with  $\vartheta = f(\cdot, t, I_{1,t}U) - I_{1,t}[f(\cdot, t, I_tU)]$  and the constants  $\kappa_0, \kappa_1, \kappa_2$  from Lemma 2.2.*

*Remark 5.3.* The a posteriori error estimate given by Theorem 5.1 resembles (but is not identical to) error estimates of [1]. Our analysis of the semidiscrete Crank–Nicolson method seems more straightforward as we work with the standard piecewise linear interpolant of the computed solution, while the analysis in [1] involves a construction of a certain piecewise-quadratic polynomial of the computed solution in time. Furthermore, in section 10, we derive a posteriori error estimates for fully discrete Crank–Nicolson methods, which were not considered in [1].

**6. Semidiscrete discontinuous Galerkin method dG(r) with Radau quadrature (no spatial discretization).** Consider an arbitrary nonuniform mesh (2.3) in the time direction and discretize the abstract parabolic problem (1.1) in time using the discontinuous Galerkin method dG(r) (described, e.g., in [10, 27]) as follows.

First, introduce the Radau points  $\mathcal{A}_R := \{\alpha_k : 0 < \alpha_0 < \alpha_1 < \dots < \alpha_r = 1\}$  (e.g.,  $r = 1$  corresponds to  $\mathcal{A}_R = \{\frac{1}{3}, 1\}$ ). We shall also use the augmented set  $\mathcal{A} := \{0\} \cup \mathcal{A}_R$  of  $r + 2$  points. Next, on  $[0, 1]$  introduce the basis  $\{\phi_k(s)\}_{k=0}^r$  for polynomials of degree  $r$  with the property  $\varphi_k(\alpha_l) = \delta_{kl}$  and the polynomial  $\zeta_{r+1}$  of degree  $r + 1$  such that

$$(6.1) \quad \zeta_{r+1}(0) = 1, \quad \zeta_{r+1}(\alpha_k) = 0 \quad \text{for } k = 0, \dots, r, \quad C_\zeta := \frac{d^{r+1}}{ds^{r+1}} \zeta_{r+1}(s).$$

Also define the two interpolants on  $(t_j, t_{j+1}]$ :  $\hat{I}_{r,t}\phi \in \Pi_r$  with  $(\hat{I}_{r,t}\phi)(t_{j+\alpha}) = \phi(t_{j+\alpha})$  for  $\alpha \in \mathcal{A}_R$  and  $I_{r+1,t}\phi \in \Pi_{r+1}$  with  $(I_{r+1,t}\phi)(t_{j+\alpha}) = \phi(t_{j+\alpha})$  for  $\alpha \in \mathcal{A}$ .

Let  $U^0 := \varphi$ . Given an approximate solution  $U^j \in H_0^1(\Omega) \cap C(\bar{\Omega})$  associated with the time level  $t_j$ , we require approximate solutions  $U^{j+\alpha_k} \in H_0^1(\Omega) \cap C(\bar{\Omega})$ , for  $k = 0, \dots, r$ , respectively associated with the time levels  $t_{j+\alpha_k}$ , to satisfy

$$(6.2a) \quad [U(\cdot, t_j^+) - U^j] \varphi_k(0) + \int_{t_j^+}^{t_{j+1}} [\partial_t U + \hat{I}_{r,t}\psi] \varphi_k\left(\frac{t-t_j}{\tau_{j+1}}\right) dt = 0 \quad \text{for } k = 0, \dots, r,$$

$$(6.2b) \quad \text{where } U := \sum_{k=0}^r U^{j+\alpha_k} \varphi_k\left(\frac{t-t_j}{\tau_{j+1}}\right), \quad \psi := \mathcal{L}(t)U + f(\cdot, t, U) \quad \text{for } t \in (t_j, t_{j+1}].$$

Note that (6.2) represents the dG(r) method with Radau quadrature, exact for polynomials of degree  $2r$ , while if the term  $\hat{I}_{r,t}\psi$  is replaced by  $\psi$ , then we get the dG(r) method without quadrature.

Next, an application of  $I_{r+1,t}$  to the approximate solutions  $\{U^{j+\alpha}, \alpha \in \mathcal{A}\}$  generates  $\tilde{U}$  and the related function  $\tilde{\psi}$ :

$$(6.3) \quad \tilde{U} := U - [U(\cdot, t_j^+) - U^j] \zeta_{r+1}\left(\frac{t-t_j}{\tau_{j+1}}\right), \quad \tilde{\psi} = \mathcal{L}(t)\tilde{U} + f(\cdot, t, \tilde{U}).$$



Note that  $I_{r+1,t}\tilde{\psi}$  allows a representation

$$(6.4) \quad I_{r+1,t}\tilde{\psi} = \hat{I}_{r,t}\psi + \chi^{j+1} \tau_{j+1}^{-1} C_\zeta^{-1} \zeta_{r+1}\left(\frac{t-t_j}{\tau_{j+1}}\right),$$

where  $\chi^{j+1} := \tau_{j+1}^{r+2} \partial_t^{r+1}[I_{r+1,t}\tilde{\psi}]$ , so, with the notation  $\psi^{j+\alpha} = \psi(\cdot, t_{j+\alpha})$ , one has

$$(6.5) \quad \chi^{j+1} = \tau_{j+1} C_\zeta \left[ \psi^j - \hat{I}_{r,t}\psi(\cdot, t_j^+) \right] = \tau_{j+1} C_\zeta \left[ \psi^j - \sum_{k=0}^r \psi^{j+\alpha_k} \varphi_k(0) \right].$$

**THEOREM 6.1.** *Let  $u$  solve the problem (1.1) with the parabolic operator  $\mathcal{M}$  satisfying (1.2) and Condition 2.1, and let  $U^j$  solve the corresponding semidiscrete problem (6.2). Then for  $m = 1, \dots, M$ , one has (3.1) with  $\vartheta = \tilde{\psi} - I_{r+1,t}\tilde{\psi}$ , the constants  $C_1 = C_\zeta^{-1} \max_{s \in [0,1]} \left| \int_0^s \zeta_{r+1}(\sigma) d\sigma \right|$  and  $C_2 = C_\zeta^{-1} \max_{s \in [0,1]} |\zeta_{r+1}(s)|$ , and the notation (6.1), (6.3), (6.5).*

*Proof.* First, note that (6.2a) is equivalent to

$$(6.6) \quad \int_{t_j}^{t_{j+1}} [\partial_t \tilde{U} + \hat{I}_{r,t}\psi] \varphi_k\left(\frac{t-t_j}{\tau_{j+1}}\right) dt = 0 \quad \text{for } k = 0, \dots, r.$$

This is easily checked by getting  $\partial_t[\tilde{U} - U]$  from the first relation in (6.3) and then noting that  $\int_{t_j}^{t_{j+1}} \partial_t \left\{ -\zeta_{r+1}\left(\frac{t-t_j}{\tau_{j+1}}\right) \right\} \cdot \varphi_k\left(\frac{t-t_j}{\tau_{j+1}}\right) dt = \varphi_k(0)$ . (The latter is easily obtained using integration by parts and the fact that  $\int_{t_j}^{t_{j+1}} p(t) dt = 0$  for any polynomial  $p$  of degree  $2r$  vanishing at the Radau points.)

Next, note that (6.6) yields  $\partial_t \tilde{U} + \hat{I}_{r,t}\psi = 0$  (as this function is a polynomial of degree  $r$  on  $[t_j, t_{j+1}]$ ). Now, as  $\mathcal{M}\tilde{U} = \partial_t \tilde{U} + \tilde{\psi}$  and  $\mathcal{M}u = 0$ , while  $\tilde{\psi} = I_{r+1,t}\tilde{\psi} + \vartheta$ , so

$$\mathcal{M}\tilde{U} - \mathcal{M}u = \partial_t \mu + \vartheta, \quad \mu := \int_{t_j}^t [I_{r+1,t}\tilde{\psi} - \hat{I}_{r,t}\psi] dt \quad \text{for } t \in [t_j, t_{j+1}].$$

It should be noted that by virtue of (6.4), the function  $\mu$  is continuous in time. (This follows from  $\zeta_{r+1}$  vanishing at the Radau points.) Furthermore,  $\mu$  satisfies the bounds (5.6) with  $\frac{1}{8}$  and  $\frac{1}{2}$  respectively replaced by  $C_1$  and  $C_2$ , while  $\mu(\cdot, t_{m-1}^+) = 0$ . The desired bound for  $U^m - u(\cdot, t_m) = [\tilde{U} - u](\cdot, t_m)$  is then obtained by an application of Lemma 2.3.  $\square$

*Remark 6.2.* Similarly to Remark 3.1, the quantity  $|\chi^j|$  in (3.1) approximates  $\tau_j^{r+2} |\partial_t^{r+2} u(\cdot, t_j)|$ , so Theorem 6.1 gives an a posteriori error estimate of order  $r + 2$ .

**6.1. Particular case dG(1).** For  $r = 1$ , the Radau points  $\mathcal{A}_R = \{\frac{1}{3}, 1\}$  are used, so (6.2) is equivalent to<sup>2</sup>

$$(6.7a) \quad U^{j+1} - U^j + \frac{1}{4} \tau_{j+1} (3\psi^{j+1/3} + \psi^{j+1}) = 0,$$

$$(6.7b) \quad U^{j+1/3} - U^j + \frac{1}{12} \tau_{j+1} (5\psi^{j+1/3} - \psi^{j+1}) = 0.$$

Furthermore, a calculation using (6.1), (6.5) yields  $\zeta_2(s) = 3(s-1)(s-\frac{1}{3})$  and  $C_\zeta = 6$ , and also

$$(6.8) \quad I_{2,t}\tilde{\psi} = \psi^{j+1} - \{\psi^{j+1} - \psi^{j+1/3}\} \cdot \frac{3}{2} \left(\frac{t_{j+1}-t}{\tau_{j+1}}\right) + \chi^{j+1} \cdot \frac{1}{6} \tau_{j+1}^{-1} \zeta_2\left(\frac{t-t_j}{\tau_{j+1}}\right),$$

<sup>2</sup>This is, in fact, an implicit two-stage Runge–Kutta method of order 3. The functions  $U^{j+1/3}$  and  $U^{j+1}$  obtained from (6.7) give third-order approximations to  $u$  at the time levels  $t_{j+1/3}$  and  $t_{j+1}$ , respectively.

where  $\chi^{j+1} = \tau_{j+1}^3 \partial_t^2(I_{2,t}\tilde{\psi})$  is given by

$$(6.9) \quad \chi^{j+1} = 3\tau_{j+1} (2\psi^j - 3\psi^{j+1/3} + \psi^{j+1}).$$

Note that  $\tilde{U}$  is generated similarly to  $I_{2,t}\tilde{\psi}$  by an application of the interpolant  $I_{2,t}$  to the approximate solutions  $\{U^j, U^{j+1/3}, U^{j+1}\}$ .

**COROLLARY 6.3 (dG(1)).** *For the semidiscrete method (6.7), the statement of Theorem 6.1 is valid with the notation (6.8), (6.9) and  $C_1 = \frac{2}{81}$ ,  $C_2 = \frac{1}{6}$ .*

*Remark 6.4 (computability).* The computation of the right-hand side in the estimate (3.1) involves computing  $\chi^{j+1}$  of (6.9) for  $j < m$ . Note that the terms  $\psi^{j+1/3}$  and  $\psi^{j+1}$ , which appear in (6.9), can be explicitly represented using (6.7).

**6.2. Application to a general  $t$ -independent operator  $\mathcal{L}$  and the model problem (1.3).** Suppose that the coefficients of the linear elliptic operator  $\mathcal{L}(t)$  are independent of the variable  $t$ ; we shall highlight this case by using the special notation  $\mathring{\mathcal{L}} := \mathcal{L}$  for this operator.

Recall that the estimator of Theorem 6.1 (as well as Corollary 6.3) involves  $\vartheta = \tilde{\psi} - I_{r+1,t}\tilde{\psi}$  with  $\tilde{\psi} = \mathring{\mathcal{L}}\tilde{U} + f(\cdot, t, \tilde{U})$ . As  $I_{r+1,t}[\mathring{\mathcal{L}}\tilde{U}] = \mathring{\mathcal{L}}[I_{r+1,t}\tilde{U}] = \mathring{\mathcal{L}}\tilde{U}$ , so

$$(6.10) \quad \vartheta = f(\cdot, t, \tilde{U}) - I_{r+1,t}[f(\cdot, t, \tilde{U})].$$

Note that now  $\vartheta$  does not involve  $\mathring{\mathcal{L}}$  and can be bounded using the properties of the function  $f$ . Our findings are summarized in the following result.

**COROLLARY 6.5.** *Let the elliptic operator  $\mathcal{L}(t) = \mathring{\mathcal{L}}$  be independent of the variable  $t$ ; then the statements of Theorem 6.1 and Corollary 6.3 remain valid with the simplification (6.10).*

Finally, recall that in the model problem (1.3) the elliptic operator  $\mathcal{L} = -\varepsilon^2\Delta$  is  $t$ -independent, so we apply Corollary 6.5 to this problem.

**COROLLARY 6.6.** *Under assumption (1.2), the a posteriori error estimates of Theorem 6.1 and Corollary 6.3 apply to the model problem (1.3) with the constants  $\kappa_0, \kappa_1, \kappa_2$  from Lemma 2.2, and the simplification (6.10).*

**7. Elliptic a posteriori error estimators.** In this section, we consider a steady-state version of the abstract parabolic problem (1.1):

$$(7.1) \quad \mathcal{L}v + g(\cdot, v) = 0 \quad \text{in } \Omega, \quad v = 0 \quad \text{on } \partial\Omega,$$

and its discretizations in the form

$$(7.2a) \quad \text{Find } v_h \in \mathring{V}_h : \quad \mathcal{L}_h v_h + \mathcal{P}_h[g(\cdot, v_h)] = 0, \quad \text{where } \mathring{V}_h := V_h \cap H_0^1(\Omega).$$

Here  $V_h \subset C(\bar{\Omega})$  is some finite element space, and with some interpolation operator  $I_h : C(\bar{\Omega}) \rightarrow V_h$ , we use some operators  $\mathcal{L}_h$  and  $\mathcal{P}_h$  such that

$$(7.2b) \quad \begin{aligned} \mathcal{L}_h &: H_0^1(\Omega) \rightarrow \mathring{V}_h - I_h[g(\cdot, 0)], \\ \mathcal{P}_h v &\in \mathring{V}_h + I_h v \quad \forall v \in C(\bar{\Omega}), \quad \mathcal{P}_h v_h = v_h \quad \forall v_h \in V_h. \end{aligned}$$

Note that as any  $v_h \in \mathring{V}_h$  vanishes on  $\partial\Omega$ , so  $\mathring{V}_h - I_h[g(\cdot, 0)] = \mathring{V}_h - I_h[g(\cdot, v_h)]$ , so the definition (7.2) is consistent.

**Assumptions.** We assume, for any admissible  $g$ , that

- (i) there exist unique solutions  $v$  and  $v_h$  of problems (7.1) and (7.2), respectively;
- (ii) an a posteriori error estimate is available for these solutions in the form

$$(7.3) \quad \|v - v_h\|_{\infty, \Omega} \leq \eta(V_h, v_h, g(\cdot, v_h)).$$

Note that the availability of elliptic a posteriori error estimates, such as (7.3), enables one to employ elliptic reconstructions of computed solutions in the a posteriori error estimation of the related parabolic problems. Moreover,  $\mathcal{L}_h$  and  $\mathcal{P}_h$  are not necessarily needed to be evaluated explicitly to compute the a posteriori estimator for either the elliptic problem or the parabolic problem.

*Remark 7.1* (relation of  $g$  to  $f$ ). We shall employ the functions  $g$  of the type  $g(x, v) = f(x, t, v) + \psi(x)$  for some fixed  $t$  and some function  $\psi$  (which approximates  $\partial_t^2 u(x, t)$ ). So problem (7.1) will typically have a unique solution by virtue of (1.2).

*Remark 7.2* (uniqueness assumption). The uniqueness assumption (i) is not essential but simplifies the presentation. In fact, one can replace assumptions (i), (ii) by the following *alternative assumption*: For each solution  $v_h$  of problem (7.2), there exists a solution  $v$  of problem (7.1) such that one has (7.3).

To be more specific, the uniqueness assumption is used only in section 8.2 below to establish (8.9). Under the alternative assumption,  $\hat{u}_h^j$  and  $u_h^{j+\alpha}$  will be particular solutions of the two discrete problems in (8.8), for which one then employs appropriate particular solutions  $\hat{R}^j$  and  $R^{j+\alpha}$  of the two corresponding elliptic problems in (8.7) such that one can deduce (8.9).

**7.1. Elliptic model problem.** Many standard finite element discretizations of elliptic equations (including those with quadrature) allow a representation of type (7.2). For example, consider a steady-state elliptic version of our model problem (1.3) posed in a bounded polyhedral domain  $\Omega \subset \mathbb{R}^n$ :

$$(7.4) \quad -\varepsilon^2 \Delta v + g(x, v) = 0 \text{ for } x \in \Omega, \quad v = 0 \text{ for } x \in \partial\Omega, \quad \partial_z g(x, z) \geq \gamma^2 \geq 0.$$

With a finite element space  $V_h \subset C(\bar{\Omega})$  and  $\mathring{V}_h := V_h \cap H_0^1(\Omega)$ , a standard Galerkin finite element method for this problem can be described by

$$(7.5) \quad \text{Find } v_h \in \mathring{V}_h : \quad \varepsilon^2 \langle \nabla v_h, \nabla w_h \rangle + \langle g(\cdot, v_h), w_h \rangle_h = 0 \quad \forall w_h \in \mathring{V}_h,$$

where  $\langle \cdot, \cdot \rangle_h$  is either exactly the inner product  $\langle \cdot, \cdot \rangle$  in  $L_2(\Omega)$  or some quadrature formula for  $\langle \cdot, \cdot \rangle$ .

*Remark 7.3.* The discretization (7.5) is of type (7.2) provided that the Gram matrix  $\langle \phi_i, \phi_j \rangle_h$  of the basis  $\{\phi_i\}$  in  $\mathring{V}_h$  is invertible. Then let  $\langle \mathcal{L}_h \varphi, w_h \rangle_h = \varepsilon^2 \langle \nabla \varphi, \nabla w_h \rangle$  and  $\langle \mathcal{P}_h q, w_h \rangle_h = \langle q, w_h \rangle_h$ , subject to (7.2b), for all  $\varphi \in H_0^1(\Omega)$ ,  $q \in C(\bar{\Omega})$ ,  $w_h \in \mathring{V}_h$ .

Suppose, for example, that  $\langle q_h, w_h \rangle_h = \langle q_h, w_h \rangle$  for all  $q_h, w_h \in V_h$ . Then  $\langle \mathcal{L}_h \varphi, w_h \rangle_h = \varepsilon^2 \langle \nabla \varphi, \nabla w_h \rangle$  and  $\langle \mathcal{P}_h q, w_h \rangle_h = \langle q, w_h \rangle_h$ , subject to (7.2b), for all  $\varphi \in H_0^1(\Omega)$ ,  $q \in C(\bar{\Omega})$  and  $w_h \in \mathring{V}_h$ . In particular,

- (i) if  $\langle \cdot, \cdot \rangle_h := \langle \cdot, \cdot \rangle$  (i.e., no quadrature is used), then  $\mathcal{P}_h$  is the  $L_2$  projection;
- (ii) if a quadrature of type  $\langle q, w_h \rangle_h := \langle I_h q, w_h \rangle$  is used, where  $I_h$  is some interpolation operator onto  $V_h$ , then  $\mathcal{P}_h := I_h$ .

*Remark 7.4.* Suppose that one employs a quadrature of lumped-mass type defined by  $\langle q, \phi_i \rangle_h = \langle I_h(q\phi_i), 1 \rangle = q_i \langle \phi_i, 1 \rangle$  for all basis functions  $\phi_i$  of  $V_h$ , where  $q \in C(\bar{\Omega})$  and  $\sum q_i \phi_i = I_h q$ . Then again  $\mathcal{P}_h := I_h$ , but  $\mathcal{L}_h v_h := \sum a_i \phi_i$  with  $a_i := \varepsilon^2 \frac{\langle \nabla v_h, \nabla \phi_i \rangle}{\langle \phi_i, 1 \rangle}$  for interior mesh nodes, and  $a_i := -[g(\cdot, 0)]_i$  for boundary mesh nodes. Consequently,  $\mathcal{L}_h v_h$  is easily computable for any  $v_h \in V_h$  by applying the normalized stiffness matrix to the column vector of nodal values  $\{v_{h,i}\}$ .

We cite elliptic estimators of type (7.3) for particular cases of (7.4) and (7.5) in Appendix A (for both  $\varepsilon = 1$  and  $\varepsilon \ll 1$ ).

**8. Summary of results for fully discrete methods. Computability.** In this section we describe our results for full discretizations of the abstract parabolic problem (1.1) satisfying (1.2) and Condition 2.1. To fully discretize this problem, we apply a spatial discretization of type (7.2) to the semidiscrete backward Euler, Crank–Nicolson, and discontinuous Galerkin methods as follows.

A finite element space  $V_h^{j+1} \subset C(\bar{\Omega})$  and a computed solution  $u_h^{j+1} \in \hat{V}_h^{j+1} := V_h^{j+1} \cap H_0^1(\Omega)$  are associated with the time level  $t_{j+1}$ , while an auxiliary computed solution  $\hat{u}_h^j \in H_0^1(\Omega)$  is associated with the time level  $t_j^+$ . (This is indicated by the hat notation, typically, either  $\hat{u}_h^j \in V_h^j$  or  $\hat{u}_h^j \in V_h^{j+1}$ .) A full discretization is then obtained from a semidiscretization using operators  $\mathcal{L}_h(t)$  and  $\mathcal{P}_h^{j+1}$ , for which, in agreement with (7.2b), with some interpolation operator  $I_h^{j+1}$  onto  $V_h^{j+1}$ , we assume that

$$(8.1) \quad \begin{aligned} \mathcal{L}_h(t) : H_0^1(\Omega) &\rightarrow \hat{V}_h^{j+1} - I_h^{j+1}[f(\cdot, t, 0)] \quad \text{for } t \in (t_j, t_{j+1}], \\ \mathcal{P}_h^{j+1}v &\in \hat{V}_h^{j+1} + I_h^{j+1}v \quad \forall v \in C(\bar{\Omega}), \quad \mathcal{P}_h^{j+1}v_h = v_h \quad \forall v_h \in V_h^{j+1}. \end{aligned}$$

Note two particular cases of interest for the auxiliary computed solution  $\hat{u}_h^j$ :

$$(8.2a) \text{ Case A: } \hat{u}_h^j := I_*^{j+1}u_h^j, \quad I_*^{j+1} : \hat{V}_h^j \rightarrow \hat{V}_h^{j+1} \quad \Rightarrow \quad \mathcal{P}_h^{j+1}\hat{u}_h^j = \hat{u}_h^j;$$

$$(8.2b) \text{ Case B: } \hat{u}_h^j := u_h^j \quad \Rightarrow \quad \hat{u}_h^j \in \hat{V}_h^j, \quad u_h^{j+1} \in \hat{V}_h^{j+1}.$$

Here, in Case A,  $\hat{u}_h^j$  is obtained by applying some linear interpolation operator  $I_*^{j+1}$  to  $u_h^j$ , for which it is assumed that  $I_*^{j+1}w_h = w_h$  for all  $w_h \in \hat{V}_h^{j+1}$ . To define  $I_*^{j+1}$ , one may employ, e.g., the standard Lagrange interpolation or the  $L_2$  projection. Note that if  $V_h^j \subseteq V_h^{j+1}$ , then Cases A and B are identical.

For  $m = 1, \dots, M$ , we give *a posteriori error estimates* of the type

$$(8.3) \quad \begin{aligned} \|u_h^m - u(\cdot, t_m)\|_{\infty, \Omega} &\leq \kappa_0 e^{-\gamma^2 t_m} \|u_h^0 - \varphi\|_{\infty, \Omega} \\ &\quad + (\kappa_1 \ell_m + \kappa_2) \max_{j=1, \dots, m-1} \{C_1 \|\chi_h^j\|_{\infty, \Omega} + C_1^* \eta^j\} \\ &\quad + C_2 \kappa_0 \|\chi_h^m\|_{\infty, \Omega} + (C_2^* \kappa_0 + 1) \eta^m \\ &\quad + \kappa_0 \sum_{j=1}^m \int_{t_{j-1}}^{t_j} e^{-\gamma^2(t_m-s)} \|\vartheta_h(\cdot, s)\|_{\infty, \Omega} ds \\ &\quad + \kappa_0 \sum_{j=1}^{m-1} e^{-\gamma^2(t_m-t_j)} \|\hat{u}_h^j - u_h^j\|_{\infty, \Omega}. \end{aligned}$$

The quantities that appear in this estimate are specified by Theorems 9.2, 10.3, and 11.4 below and can be summarized as follows:

	$p$	$\chi_h^{j+1}$	$C_1$	$C_2$	$C_1^*$	$C_2^*$	$\mathcal{A}$
Backward Euler	1	$u_h^{j+1} - \hat{u}_h^j$	1	2	1	1	{1}
Crank–Nicolson	2	$\tau_{j+1}(\psi_h^{j+1} - \hat{\psi}_h^j)$	$\frac{1}{8}$	$\frac{1}{2}$	2	3	{0 <sup>+</sup> , 1}
dG(1)-Radau	3	$3\tau_{j+1}(2\hat{\psi}_h^j - 3\psi_h^{j+1/3} + \psi_h^{j+1})$	$\frac{2}{81}$	$\frac{1}{6}$	$\frac{5}{3}$	10	{0 <sup>+</sup> , $\frac{1}{3}$ , 1}

Here for the evaluation of  $\chi_h^{j+1}$  we use  $\hat{\psi}_h^j$  and  $\psi_h^{j+\alpha}$  that satisfy (similarly to (3.2))

$$(8.4a) \quad \mathcal{P}_h^{j+1} \hat{\psi}_h^j = \mathcal{L}_h(t_j^+) \hat{u}_h^j + \mathcal{P}_h^{j+1}[f(\cdot, t_j, \hat{u}_h^j)],$$

$$(8.4b) \quad \mathcal{P}_h^{j+1} \psi_h^{j+\alpha} = \mathcal{L}_h(t_{j+\alpha}) u_h^{j+\alpha} + \mathcal{P}_h^{j+1}[f(\cdot, t_{j+\alpha}, u_h^{j+\alpha})],$$

where  $\alpha$  takes values from  $\mathcal{A} \setminus \{0^+\}$ , for which the computed solution  $u_h^{j+\alpha}$  is available from the definition of the method, while for  $\alpha = 0^+$  we use  $\hat{u}_h^j$  and  $\hat{\psi}_h^j$ . Note that in Case A of (8.2a), relations (8.4) simplify using  $\hat{\psi}_h^j = \mathcal{P}_h^{j+1} \hat{\psi}_h^j$  and  $\psi_h^{j+\alpha} = \mathcal{P}_h^{j+1} \psi_h^{j+\alpha}$ .

The quantity  $\eta^j$  in (8.3) is related to the error due to the spatial discretization used; it is defined using the elliptic estimator  $\eta$  from (7.3) by

$$(8.5a) \quad \eta^{j+1} := \eta(V_h^{j+1}, u_h^{j+1}, g^{j+1}(\cdot, u_h^{j+1})) \quad \text{if } \mathcal{A} = \{1\},$$

i.e., for the backward Euler method, and, otherwise, by

$$(8.5b) \quad \eta^{j+1} := \max \left\{ \eta(V_h^{j+1}, \hat{u}_h^j, \hat{g}^j(\cdot, \hat{u}_h^j)), \max_{\alpha \in \mathcal{A} \setminus \{0^+\}} \eta(V_h^{j+1}, u_h^{j+\alpha}, g^{j+\alpha}(\cdot, u_h^{j+\alpha})) \right\},$$

where

$$(8.5c) \quad \hat{g}^j(\cdot, v) := f(\cdot, t_j, v) - \hat{\psi}_h^j, \quad g^{j+\alpha}(\cdot, v) := f(\cdot, t_{j+\alpha}, v) - \psi_h^{j+\alpha}.$$

The quantity  $\vartheta_h$  is similar to  $\vartheta$  in (3.1) but involves the so-called elliptic reconstruction of the computed solution, so we defer the definition and estimation of this quantity to sections 8.2 and 8.3. The constants  $C_1$  and  $C_2$  in (8.3) are the same as in the estimate (3.1) for the corresponding semidiscrete method.

*Remark 8.1* (interpolation-type estimates). Similarly to Remarks 3.1 and 3.2 for the semidiscrete methods, the quantity  $|\chi_h^j|$  in (8.3) approximates  $\tau_j^p |\partial_t^p u(\cdot, t_j)|$ ; consequently, (8.3) gives an a posteriori error estimate of order  $p$  with  $p = 1, 2$ , and 3 for the backward Euler, Crank–Nicolson, and dG(1) methods, respectively.

*Remark 8.2.* The final term in the error estimate (8.3) vanishes when one has  $\hat{u}_h^j = u_h^j$  for all  $j = 1, \dots, M$ , i.e., in Case B of (8.2), and also in Case A if the mesh is not coarsened. Note also that in some cases the final term in (8.3) can be improved to (9.14); see Remark 9.3, which applies to the backward Euler as well as to the Crank–Nicolson and dG( $r$ ) methods.

### 8.1. Computability of $\chi_h^j$ and $\eta^j$ in the a posteriori error estimate (8.3).

For the *backward Euler* method we shall use (see Remark 9.1 on  $\psi_h^{j+1}$ )

$$(8.6) \quad \chi_h^{j+1} = u_h^{j+1} - \hat{u}_h^j, \quad \psi_h^{j+1} = -\frac{u_h^{j+1} - \hat{u}_h^j}{\tau_{j+1}}, \quad g^{j+1}(\cdot, u_h^{j+1}) = f(\cdot, t_{j+1}, u_h^{j+1}) - \psi_h^{j+1},$$

where the relation for  $g^{j+1}$  agrees with (8.5c). As  $u_h^{j+1}$  and  $\hat{u}_h^j$  are available during the computation process, so  $\chi_h^{j+1}$  and  $\eta^{j+1}$  of (8.5a) are easily explicitly computable.

For the *Crank–Nicolson* and *dG(1)* methods, the computability of  $\chi_h^j$  and  $\eta^j$  of (8.5), being somewhat less straightforward, reduces to the availability of  $\hat{\psi}_h^j$ . Indeed, for the Crank–Nicolson method, one can explicitly represent  $\psi_h^{j+1}$  (by means of (10.4) assuming that  $\hat{\psi}_h^j$  is available), while for the dG(1) method,  $\psi_h^{j+1/3}$  and  $\psi_h^{j+1}$  are explicitly computable (by means of (11.3a)). So, if  $\hat{\psi}_h^j$  is available, one can indeed explicitly compute  $\chi_h^{j+1}$  and  $\eta^{j+1}$ .

We now briefly discuss possible approaches to the computation of  $\hat{\psi}_h^j$  when applied to the *model problem* (1.3) in *Case A* of (8.2). In this case,  $\hat{u}_h^j \in \tilde{V}_h^{j+1}$  and (8.4a) simplifies to  $\hat{\psi}_h^j = \mathcal{L}_h(t_j^+) \hat{u}_h^j + \mathcal{P}_h^{j+1}[f(\cdot, t_j, \hat{u}_h^j)]$ , so it may help the reader to recall Remarks 7.3 and 7.4; see also Remark A.1. (For Case B, we give Remark 8.3 below.)

(i) Suppose  $V_h^j = V_h^{j+1}$ . Then, by (8.4) combined with  $\psi_h^{j+1} = \mathcal{P}_h^{j+1}\psi_h^{j+1}$ , one enjoys  $\hat{\psi}_h^j = \psi_h^j$ , where  $\psi_h^j$  has already been computed.

(ii) Suppose that  $\mathcal{P}_h^{j+1}$  is associated with a *lumped-mass* quadrature  $\langle q, \phi_i \rangle_h$ . Then, as described in Remark 7.4,  $\mathcal{P}_h^{j+1} = I_h^{j+1}$  is some interpolation operator onto  $V_h^{j+1}$ , while  $\mathcal{L}_h(t_j^+) \hat{u}_h^j$  is easily computable for any  $\hat{u}_h^j \in \hat{V}_h^{j+1}$  by applying the normalized stiffness matrix to the column vector of nodal values  $\{\hat{u}_{h,i}^j\}$ . Consequently, the computation of  $\hat{\psi}_h^j$  using the right-hand side in (8.4a) involves only explicit computations.

(iii) In the general case, the computation of  $\hat{\psi}_h^j$  by means of the right-hand side in (8.4a) involves an application of  $\mathcal{L}_h(t_j^+)$  and  $\mathcal{P}_h^{j+1}$ . Note that Remark 7.3 implies that, roughly speaking,  $\mathcal{L}_h(t_j^+)v_h$  for any  $v_h \in \hat{V}_h^{j+1}$  can be obtained by an application of  $M_{j+1}^{-1}K_{j+1}$  to the column vector of nodal values  $\{v_{h,i}\}$ , where  $M_{j+1}$  is the mass matrix and  $K_{j+1}$  is the stiffness matrix associated with the time level  $t_{j+1}$ . Such computations may be expensive.

Note also that, in some cases, an inversion of the mass matrix may be entirely avoided as follows. Suppose  $\hat{\psi}_h^j - w_h$  is involved in the estimator with some function  $w_h$ , and an inversion of  $M := M_{j+1}$  is required to compute  $\hat{\psi}_h^j$ . Then one can instead use the bound  $\|\hat{\psi}_h^j - w_h\|_{\infty, \Omega} \leq \|M^{-1}\|_{\infty} \cdot \|M(\hat{\psi}_h^j - w_h)\|_{\infty, \Omega}$ , where  $\|M^{-1}\|_{\infty}$  denotes the associated matrix norm (which may be bounded a priori). As  $M\hat{\psi}_h^j$  is explicitly computable (using an application of the normalized stiffness matrix to the column vector of nodal values associated with  $\hat{u}_h^j$ ), all the computations become explicit.

*Remark 8.3 (Case B).* In case (8.2b) with  $V_h^j \not\subseteq V_h^{j+1}$ , for the Crank–Nicolson method,  $\hat{\psi}_h^j$  is not given by the right-hand side in (8.4a), so  $\psi_h^{j+1}$  and  $\hat{\psi}_h^j$  are computed by means of (10.3), using the above items (ii) or (iii) in the computation of  $\psi_h^{j+1}$ . For the dG(1) method in this case, one can use  $\hat{\psi}_h^j = \psi_h^j$  by virtue of Remark 11.3.

**8.2. Elliptic reconstruction. Definition of  $\vartheta_h$ .** In our error analysis for fully discrete methods, we employ the *elliptic reconstruction* of the computed solution, which was introduced in the recent papers [22, 19, 6] as a counterpart of the Ritz-projection in the a posteriori error estimation for parabolic problems.

We associate elliptic reconstructions  $\hat{R}^j$  with the time level  $t_j^+$  and  $R^{j+\alpha}$  for  $\alpha \in \mathcal{A} \setminus \{0^+\}$  with the time level  $t_{j+\alpha}$ . They are defined, using  $\hat{g}^j$  and  $g^{j+\alpha}$  of (8.5c), as the unique solutions in  $H_0^1(\Omega) \cap C(\bar{\Omega})$  of the elliptic problems

$$(8.7) \quad \mathcal{L}(t_j) \hat{R}^j + \hat{g}^j(x, \hat{R}^j) = 0, \quad \mathcal{L}(t_{j+\alpha}) R^{j+\alpha} + g^{j+\alpha}(x, R^{j+\alpha}) = 0.$$

Note that (8.7) describes two versions of the elliptic problem (7.1) with  $\mathcal{L} := \mathcal{L}(t_j)$ ,  $g := \hat{g}^j$ , and with  $\mathcal{L} := \mathcal{L}(t_{j+\alpha})$ ,  $g := g^{j+\alpha}$ , and exact solutions  $\hat{R}^j$  and  $R^{j+\alpha}$ , respectively. Furthermore, the numerical method (7.2), using the finite element space  $V_h^{j+1}$ , applied to these two problems yields

$$(8.8) \quad \mathcal{L}_h(t_j^+) \hat{R}_h^j + \mathcal{P}_h^{j+1} [\hat{g}^j(x, \hat{R}_h^j)] = 0, \quad \mathcal{L}_h(t_{j+\alpha}) R_h^{j+\alpha} + \mathcal{P}_h^{j+1} [g^{j+\alpha}(x, R_h^{j+\alpha})] = 0.$$

We have assumed that solutions of these two discrete problems are unique. Thus,  $\hat{R}_h^j = \hat{u}_h^j$  and  $R_h^{j+\alpha} = u_h^{j+\alpha}$ . This is easily checked by combining (8.8) with the definitions of  $\hat{g}^j$  and  $g^{j+\alpha}$  in (8.5c) and then using (8.4). Consequently, applying the elliptic a posteriori error estimate (7.3) to the exact solutions  $\hat{R}^j$  and  $R^{j+\alpha}$  and the corresponding computed solutions  $\hat{u}_h^j$  and  $u_h^{j+\alpha}$ , and recalling  $\eta^{j+1}$  of (8.5), one gets

(8.9)

$$\|\hat{R}^j - \hat{u}_h^j\|_{\infty, \Omega} \leq \eta^{j+1} \text{ if } 0^+ \in \mathcal{A}, \quad \|R^{j+\alpha} - u_h^{j+\alpha}\|_{\infty, \Omega} \leq \eta^{j+1} \text{ for } \alpha \in \mathcal{A} \setminus \{0^+\}.$$

Next, similarly to  $\tilde{U}$ ,  $\tilde{\psi}$ , and  $\vartheta$  of section 3, we define a piecewise-polynomial  $\tilde{R}$  and then  $\tilde{\psi}_R$  and  $\vartheta_h$  by

$$(8.10) \quad \tilde{R} := I_{p-1,t}^* R, \quad \tilde{\psi}_R := \mathcal{L}(t) \tilde{R} + f(\cdot, t, \tilde{R}), \quad \vartheta_h := \tilde{\psi}_R - I_{p-1,t}^* \tilde{\psi}_R.$$

Here  $I_{p-1,t}$  is a piecewise-polynomial interpolation operator of degree  $p-1$  using the interpolation points  $\{t_{j+\alpha}, \alpha \in \mathcal{A}\}$  on each  $(t_j, t_{j+1}]$ . (The difference between  $I_{p-1,t}^*$  and  $I_{p-1,t}$  is that now we use the interpolation point  $t_j^+$  rather than  $t_j$ , while  $I_{0,t}^* = I_{0,t}$ .)

Note that by virtue of (8.7), (8.5c), the definition of  $\tilde{\psi}_R$  in (8.10) implies that

$$(8.11) \quad \tilde{\psi}_R(\cdot, t_j^+) = \hat{\psi}_h^j \text{ if } 0^+ \in \mathcal{A}, \quad \tilde{\psi}_R(\cdot, t_{j+\alpha}) = \psi_h^{j+\alpha} \text{ for } \alpha \in \mathcal{A} \setminus \{0^+\}.$$

**8.3. Estimation of  $\vartheta_h$ .** We now briefly discuss possible approaches to the estimation of  $\vartheta_h$  in the case of a  $t$ -independent  $\mathcal{L}$ , which includes the model problem (1.3). Then  $\vartheta_h$  of (8.10) simplifies to

$$(8.12) \quad \vartheta_h = \vartheta_{f, \tilde{R}} := f(\cdot, t, \tilde{R}) - I_{p-1,t}^*[f(\cdot, t, \tilde{R})].$$

*Remark 8.4* (backward Euler). For the backward Euler method,  $\tilde{R} = R^j$  so (8.12) simplifies to  $\vartheta_h = \vartheta_{f, \tilde{R}} = f(\cdot, t, R^j) - f(\cdot, t_j, R^j)$  for  $t \in (t_{j-1}, t_j]$ . As  $\vartheta_{f, \tilde{R}}$  involves the elliptic reconstruction  $R^j$ , which is unavailable during the computation process, instead one can use  $\vartheta_{f, \tilde{u}_h}$  (where  $\tilde{u}_h = u_h^j$ ), which can be estimated by sampling (it suffices to use a few values of  $t$  on each interval  $(t_{j-1}, t_j]$ ). Note that the discrepancy of  $\vartheta_{f, \tilde{R}}$  from  $\vartheta_{f, \tilde{u}_h}$  can be easily estimated. For example, for  $t \in (t_{j-1}, t_j]$ , we have

$$\begin{aligned} \|[\vartheta_{f, \tilde{R}} - \vartheta_{f, \tilde{u}_h}](\cdot, t)\|_{\infty, \Omega} &\leq \eta^j \sup_{(t_{j-1}, t_j] \times \mathbb{R}} \|\partial_z f(\cdot, t, z) - \partial_z f(\cdot, t_j, z)\|_{\infty, \Omega} \\ &\leq \tau_j \eta^j \sup_{(t_{j-1}, t_j] \times \mathbb{R}} \|\partial_t \partial_z f(\cdot, t, z)\|_{\infty, \Omega}, \end{aligned}$$

where we used (8.9), and  $\eta^j$  is computed using (8.5a). In fact, if  $|\partial_t \partial_z f| \leq C$ , then the discrepancy  $\|[\vartheta_{f, \tilde{R}} - \vartheta_{f, \tilde{u}_h}](\cdot, t)\|_{\infty, \Omega}$  between  $\vartheta_{f, \tilde{R}}$  and  $\vartheta_{f, \tilde{u}_h}$  becomes  $\mathcal{O}(\tau_j \eta^j)$ , i.e., negligible compared with the terms  $\eta^j$  that explicitly appear in (8.3).

*Remark 8.5* (Crank–Nicolson and dG(1)). In general, for the estimation of  $\vartheta_{f, \tilde{R}}$  in (8.12), one can use  $\vartheta_{f, \tilde{u}_h}$  with  $\tilde{u}_h := I_{p-1,t}^* u_h$ , which can be estimated by sampling, as one expects  $\vartheta_{f, \tilde{R}} \approx \vartheta_{f, \tilde{u}_h}$ . For example, if  $|\partial_z f| \leq C_f$  for some constant  $C_f$ , using

$$|\vartheta_{f, \tilde{R}} - \vartheta_{f, \tilde{u}_h}| \leq |f(\cdot, t, \tilde{u}_h) - f(\cdot, t, \tilde{R})| + |I_{p,t}^*[f(\cdot, t, \tilde{u}_h) - f(\cdot, t, \tilde{R})]|,$$

one easily gets a very crude bound  $\|[\vartheta_{f, \tilde{R}} - \vartheta_{f, \tilde{u}_h}](\cdot, t)\|_{\infty, \Omega} \leq C_* C_f \eta^{j+1}$  for  $t \in (t_j, t_{j+1}]$  with  $C_* = 2$  for the Crank–Nicolson method and  $C_* = \frac{10}{3}$  for the dG(1) method. Furthermore, in some special cases (e.g., if  $f$  is linear in the third argument) one can, in fact, get a sharper bound of type  $\|[\vartheta_{f, \tilde{R}} - \vartheta_{f, I \tilde{u}_h}](\cdot, t)\|_{\infty, \Omega} \leq C \tau_{j+1} \eta^{j+1}$  for  $t \in (t_j, t_{j+1}]$  with some constant  $C$ . Then the discrepancy between  $\vartheta_{f, \tilde{R}}$  and  $\vartheta_{f, \tilde{u}_h}$  becomes negligible compared with the terms  $\eta^{j+1}$  that already appear in (8.3).

**9. Fully discrete backward Euler method.** Consider a fully discrete backward Euler method for the model problem (1.3), obtained by applying the spatial discretization (7.5) to a version of the semidiscrete backward Euler method (4.1):

$$(9.1) \text{ Find } u_h^j \in \hat{V}_h^j : \quad \varepsilon^2 \langle \nabla u_h^j, \nabla w_h \rangle + \langle f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j, w_h \rangle_h = 0 \quad \forall w_h \in \hat{V}_h^j$$

for  $j = 1, \dots, M$ , where  $\langle \cdot, \cdot \rangle_h$  is either exactly the inner product  $\langle \cdot, \cdot \rangle$  in  $L_2(\Omega)$  or some quadrature formula for  $\langle \cdot, \cdot \rangle$ .

The term  $\delta_t^* u_h^j$  approximates  $\partial_t u$  and is defined by

$$(9.2a) \quad \delta_t^* u_h^j := \frac{u_h^j - \hat{u}_h^{j-1}}{\tau_j}, \quad \text{where } \hat{u}_h^0 := u_h^0.$$

The operator  $\delta_t^*$  is identical to  $\delta_t$  of (4.1b) for  $j = 1$ , while for  $j > 1$  it involves the intermediate computed solution  $\hat{u}_h^{j-1} \in H_0^1(\Omega)$  that we associate with the time level  $t_{j-1}^+$ , for which we note possible choices (8.2).

Note that the full discretization (9.1) can be represented as

$$(9.2b) \quad \mathcal{L}_h^j u_h^j + \mathcal{P}_h^j [f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j] = 0$$

with  $\mathcal{L}_h^j := \mathcal{L}_h(t_j)$  and  $\mathcal{P}_h^j$  subject to (8.1). For some particular cases of  $\langle \cdot, \cdot \rangle_h$ , the operators  $\mathcal{L}_h^j$  and  $\mathcal{P}_h^j$  are defined as in Remarks 7.3 and 7.4, only using  $V_h^j$  instead of  $V_h$ . Furthermore, (9.2b) can be considered a full discretization of the abstract parabolic problem (1.1) obtained by applying a spatial discretization of type (7.2) to the semidiscrete problem (4.1).

Note that by virtue of (8.1),  $\mathcal{L}_h^j u_h^j \in \hat{V}_h^j - I_h^j[f(\cdot, t_j, 0)]$ , while as both  $u_h^j$  and  $\delta_t^* u_h^j$  vanish on  $\partial\Omega$ , so  $\hat{V}_h^j - I_h^j[f(\cdot, t_j, 0)]$  coincides with  $\hat{V}_h^j - I_h^j[f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j]$ , so the definition (9.2) is consistent.

**9.1. A posteriori error estimate using a piecewise-constant elliptic reconstruction.** To estimate the error of the fully discrete backward Euler method (9.2), set  $\mathcal{A} := \{1\}$  (i.e., always use  $j + \alpha = j + 1$ ) and recall the elliptic reconstructions  $R^j$  defined for  $j = 1, \dots, M$  by (8.7). This definition involves  $g^j$ , which in its turn involves  $\psi_h^j = -\delta_t^* u_h^j$ , both defined in (8.6).

*Remark 9.1.* By (9.2b),  $\psi_h^j = -\delta_t^* u_h^j$  implies  $\mathcal{P}_h^j \psi_h^j = \mathcal{L}_h^j u_h^j + \mathcal{P}_h^j [f(\cdot, t_j, u_h^j)]$ , i.e.,  $\psi_h^j$  satisfies (8.4b). (In Case A of (8.2), this relation simplifies using  $\mathcal{P}_h^j \psi_h^j = \psi_h^j$ .) Consequently,  $R^j$  satisfies (8.9) with  $\mathcal{A} := \{1\}$ .

We now give an a posteriori error estimate for the fully discrete method (9.2).

**THEOREM 9.2.** *Let  $u$  solve the problem (1.1), (1.2) with the parabolic operator  $\mathcal{M}$  that satisfies Condition 2.1 and  $u_h^j$  solve the discrete problem (9.2). Then for  $m = 1, \dots, M$ , one has (8.3) with  $\eta^j$  and  $\chi_h^j$  defined by (8.5a), (8.6),  $C_1 = 1$ ,  $C_2 = 2$ ,  $C_1^* = C_2^* = 1$ , and a regular function  $\vartheta_h$  defined, for  $t \in (t_{j-1}, t_j]$ ,  $j = 1, \dots, M$ , by*

$$(9.3) \quad \vartheta_h(\cdot, t) = \tilde{\psi}_R(\cdot, t) - \tilde{\psi}_R(\cdot, t_j), \quad \tilde{\psi}_R(\cdot, t) = \mathcal{L}(t)R^j + f(\cdot, t, R^j) \text{ for } t \in (t_{j-1}, t_j].$$

Here  $R^j$  is the elliptic reconstruction defined by (8.7), (8.6) using  $\mathcal{A} := \{1\}$ .

**THEOREM 9.2\*.** *The statement of Theorem 9.2 is valid with the terms  $\|\chi_h^j\|_{\infty, \Omega}$  and  $\|\chi_h^m\|_{\infty, \Omega}$  in (8.3) respectively replaced by  $\|u_h^j - u_h^{j-1}\|_{\infty, \Omega}$  and  $\|u_h^m - u_h^{m-1}\|_{\infty, \Omega}$  and also  $e^{-\gamma^2(t_m - t_j)}$  replaced by  $e^{-\gamma^2(t_m - t_{j+1})}$ .*



We first give a proof of Theorem 9.2\* and then generalize it to prove Theorem 9.2.

*Proof of Theorem 9.2\*.* In view of Remark 9.1,  $\|R^j - u_h^j\|_{\infty, \Omega} \leq \eta^j$ , so to get the desired bound of type (8.3) for  $u_h^m - u(\cdot, t_m)$ , it suffices to obtain a bound of type (8.3) for  $R^m - u(\cdot, t_m)$  only with  $(C_2^* \kappa_0 + 1)$  replaced by  $C_2^* \kappa_0 = \kappa_0$ , and then apply the triangle inequality. So we focus on estimating  $R^m - u(\cdot, t_m)$ .

We partially imitate the proof of Theorem 4.1. Let  $I_{1,t}u_h$  be a standard piecewise-linear interpolant of  $u_h^j$  in time:

$$(9.4) \quad I_{1,t}u_h(\cdot, t) := \frac{t_j - t}{\tau_j} u_h^{j-1} + \frac{t - t_{j-1}}{\tau_j} u_h^j \quad \text{for } t \in [t_{j-1}, t_j], \quad j = 1, \dots, M.$$

Furthermore, we define a *piecewise-constant* interpolant  $\tilde{R}$  of  $R^j$  in time by

$$(9.5) \quad \tilde{R}(\cdot, t) := R^j \quad \text{for } t \in (t_{j-1}, t_j], \quad j = 1 \dots, M; \quad \tilde{R}(\cdot, 0) := R^1$$

(so  $\tilde{R}$  is continuous on  $[t_0, t_1]$ ; compare with  $\tilde{U}$  of (4.4)). The temporal derivative  $\partial_t \tilde{R}$  is understood in the sense of distributions, while  $\partial_t(I_{1,t}u_h)$  is a regular function.

Note that with our definition of  $\tilde{R}$ , the functions in (9.3) are identical with those in (8.10) (using  $p = 1$ ), so we also enjoy the observation (8.11), which can be rewritten as  $\tilde{\psi}_R(\cdot, t_j) = \psi_h^j = -\delta_t^* u_h^j$ . Combining this with (9.3) yields  $\psi_R = \vartheta_h - \delta_t^* u_h^j$ , so

$$(9.6) \quad \partial_t(I_{1,t}u_h) + \tilde{\psi}_R = \vartheta_h + \vartheta_* \quad \text{in } Q,$$

where  $\vartheta_*$  is a regular function defined by

$$(9.7) \quad \vartheta_*(\cdot, t) := \partial_t(I_{1,t}u_h) - \delta_t^* u_h^j \quad \text{for } t \in (t_{j-1}, t_j].$$

As  $\mathcal{M}\tilde{R} = \partial_t \tilde{R} + \tilde{\psi}_R$  and  $\mathcal{M}u = 0$ , so (9.6) yields

$$\mathcal{M}\tilde{R} - \mathcal{M}u = \partial_t[\tilde{R} - I_{1,t}u_h] + [\vartheta_h + \vartheta_*] \quad \text{in } Q.$$

Now the desired bound of type (8.3) for  $R^m - u(\cdot, t_m) = [\tilde{R} - u](\cdot, t_m)$  only with  $(C_2^* \kappa_0 + 1)$  replaced by  $C_2^* \kappa_0 = \kappa_0$  is obtained by an application of Lemma 2.3 with  $\mu := \tilde{R} - I_{1,t}u_h$  and  $\vartheta := \vartheta_h + \vartheta_*$ , using the following three observations. First, note that

$$(9.8) \quad [\tilde{R} - u - \mu](\cdot, 0) = R^1 - \varphi - (R^1 - u_h^0) = u_h^0 - \varphi.$$

Next, for  $t \in (t_{j-1}, t_j]$ , we have  $\mu = R^j - u_h^j + \frac{t_j - t}{\tau_j} (u_h^j - u_h^{j-1})$ . Thus,

$$(9.9) \quad |\mu| \leq |R^j - u_h^j| + |u_h^j - u_h^{j-1}| \quad \text{and} \quad \tau_j |\partial_t \mu| = |u_h^j - u_h^{j-1}|,$$

where  $\|R^j - u_h^j\|_{\infty, \Omega} \leq \eta^j$ . Finally, (9.7) combined with (9.2a), (9.4) implies that  $\vartheta_*(\cdot, t) = \frac{1}{\tau_j} (\hat{u}_h^{j-1} - u_h^{j-1})$  for  $t \in (t_{j-1}, t_j]$ . Therefore,

$$(9.10) \quad \int_{t_j}^{t_{j+1}} e^{-\gamma^2(t_m - s)} \|\vartheta_*(\cdot, s)\|_{\infty, \Omega} ds \leq e^{-\gamma^2(t_m - t_{j+1})} \|\hat{u}_h^j - u_h^j\|_{\infty, \Omega},$$

where  $\hat{u}_h^0 - u_h^0 = 0$ . The three observations (9.8), (9.9), (9.10) yield the required bound for  $\|R^m - u(\cdot, t_m)\|_{\infty, \Omega}$ .  $\square$

*Proof of Theorem 9.2.* We imitate the proof of Theorem 9.2\*, except  $I_{1,t}u_h$  of (9.4) is replaced everywhere by the piecewise-continuous interpolant

$$(9.11) \quad I_{1,t}^*u_h(\cdot, t) := \frac{t_j - t}{\tau_j} \hat{u}_h^{j-1} + \frac{t - t_{j-1}}{\tau_j} u_h^j \quad \text{for } t \in (t_{j-1}, t_j], \quad j = 1, \dots, M,$$

with  $I_{1,t}^* u_h(\cdot, 0) := \hat{u}_h^0 = u_h^0$ . Furthermore,  $\vartheta_*$  is defined not by (9.7), but by

$$(9.12) \quad \vartheta_*(\cdot, t) := \partial_t(I_{1,t}^* u_h) - \delta_t^* u_h^j = [\hat{u}_h^{j-1} - u_h^{j-1}] \delta(t - t_{j-1}^+) \quad \text{for } t \in (t_{j-1}, t_j],$$

where  $\delta(\cdot)$  is the one-dimensional Dirac  $\delta$ -distribution. (Note that  $\hat{u}_h^0 = u_h^0$  and the right-continuity convention at  $t = 0$  imply that  $\vartheta_* = 0$  on  $[0, t_1]$ .) So instead of (9.10) we use

$$(9.13) \quad \int_0^{t_m} e^{-\gamma^2(t_m-s)} \|\vartheta_*(\cdot, s)\|_{\infty, \Omega} ds \leq \sum_{j=1}^{m-1} e^{-\gamma^2(t_m-t_j)} \|\hat{u}_h^j - u_h^j\|_{\infty, \Omega}.$$

The required bound for  $R^m - u(\cdot, t_m) = [\tilde{R} - u](\cdot, t_m)$  is again obtained by an application of Lemma 2.3 only with  $\mu := \tilde{R} - I_{1,t}^* u_h$ , for which we have a version of (9.9) with  $u_h^{j-1}$  replaced by  $\hat{u}_h^{j-1}$ .  $\square$

*Remark 9.3* (improved mesh-coarsening term). In some cases the coarsening term that appears in the final line of (8.3) can be improved to the form

$$(9.14) \quad \kappa_3 \ell(\tau, t) \max_{j=1, \dots, m-1} \{ \tau_j^{-1} \|\mathcal{H}_j^2 \vartheta^j\|_{\infty, \Omega} \}, \quad \text{where } \vartheta^j = u_h^{j-1} - \hat{u}_h^{j-1},$$

with  $\mathcal{H}_j$  representing the local mesh size associated with  $\hat{V}_h^j$ . This version of (8.3) is easily obtained by an application of Lemma 2.3\* using  $\vartheta_*$  from (9.12) provided one has a version of Lemma 2.2 for spatial derivatives of the Green’s function. Indeed, let  $\hat{u}_h^{j-1}$  be the  $L_2$  projection of  $u_h^{j-1}$  onto  $\hat{V}_h^j$ . Then  $\langle u_h^{j-1} - \hat{u}_h^{j-1}, \Gamma_h^j \rangle = 0$  for any  $\Gamma_h^j \in \hat{V}_h^j$ . So choosing  $\Gamma_h^j = I_h^j \Gamma(\cdot, t_j)$ , it suffices to show that  $\|\mathcal{H}_j^{-2} \{\Gamma(\cdot, t_j) - \Gamma_h^j\}\|_{1, \Omega} \leq \kappa_3 \frac{C}{t_m - t_j}$ . The desired result follows if one has  $\|\Gamma(\cdot, t_j)\|_{W_1^2(\Omega)} \leq \kappa_3 \frac{C}{t_m - t_j}$ . The latter bound is crucial in this argument; it involves the spatial derivatives of  $\Gamma$  and can be obtained from [7, (2.2)] if  $\mathcal{L} = -\Delta + 1$  in a smooth domain and  $f = f(x, t)$ , with an unspecified  $\kappa_3 = \mathcal{O}(1)$ , and from [16, (2.18b)] if  $\mathcal{L} = -\varepsilon^2 \partial_{x_1}^2 u + a(x_1)$ , with  $\kappa_3 = \mathcal{O}(\varepsilon^{-2})$ .

It is important to note that  $\kappa_3 = \mathcal{O}(\varepsilon^{-2})$  (as, by (2.2a),  $\Delta \Gamma = \varepsilon^{-2}[-\partial_s + a]\Gamma$ ; see also [16]). So in the singularly perturbed regime  $\varepsilon \ll 1$ , the mesh-coarsening term (9.14) may be considerably larger than the original final term in (8.3). Whether the latter is sharp is still an open question. (See [17] for preliminary numerical results.)

Note also that unless the mesh is coarsened a finite number of times, the choice of appropriate strategies for mesh coarsening/updating remains a very delicate issue even in the regular regime; see the counterexample in [8, section 4].

**9.2. Model problem (1.3): Regular regime.** Let  $u$  solve the problem (1.3) with  $\varepsilon = 1, \gamma \geq 0$ , posed in a bounded polyhedral spatial domain  $\Omega \subset \mathbb{R}^n, n = 2, 3$ , and let  $u_h^j$  solve the discrete problem (9.1) with  $V_h^j$  and  $\langle \cdot, \cdot \rangle_h$  defined, for each time level  $t_j$ , as in section A.1. To be more specific, we let  $\mathcal{T}_h^j$  be a conforming and shape-regular triangulation of  $\bar{\Omega}$  made of elements  $T, V_h^j$  be the space of continuous piecewise polynomial finite element functions of degree  $l \geq 1$ , and  $\hat{V}_h^j := V_h^j \cap H_0^1(\Omega)$ . We then employ a quadrature formula  $\langle \varphi, w \rangle_h := \sum_{T \in \mathcal{T}_h^j} Q_T(\varphi w)$ , as described in section A.1.

**COROLLARY 9.4.** *Let the above numerical method be applied to problem (1.3) with  $\varepsilon = 1, \gamma \geq 0$ . Then the a posteriori error estimates of Theorems 9.2 and 9.2\* are valid with  $\vartheta_h$  simplified to (8.12) and estimated as described in Remark 8.4, and*

$$\eta^j := \eta_0 (V_h^j, u_h^j, f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j) \quad \text{for } j = 1, \dots, M,$$

where  $\eta_0$  is defined in (A.1).

*Remark 9.5.* The backward Euler method for a linear version of (1.3) with  $\varepsilon = 1$  was considered in [9, 3, 6, 7], in Case B of (8.2), equivalent to Case A with  $\hat{u}_h^{j-1}$  being the  $L_2$  projection of  $u_h^{j-1}$  onto  $\hat{V}_h^j$ . The a posteriori error estimate of Corollary 9.4 resembles (but is not identical to) the ones of [9, (1.13)] and [3] in that it involves terms such as  $|u_h^j - u_h^{j-1}|$  that may be interpreted as approximating  $\tau_j |\partial_t u|$ . Note also that [9, (1.13)] is given without proof and does not appear to be proved elsewhere. The proofs in [3] invoke bounds of temporal and spatial derivatives of a generalized parabolic Green's function in the  $L_1(\Omega)$  norm and appear fairly complicated compared to our approach. (We also discuss [3] in Remarks 9.9 and 11.7 below.)

By contrast, the a posteriori error estimates of [6, 7] include terms (denoted by  $\tau_j |g^j - g^{j-1}|$  in [6]) that may be interpreted as approximating the quantity  $\tau_j |\partial_t^2 u + \dots|$ , which seems less suitable for a first-order method in time.

The mesh-coarsening terms in [9, 7] are similar to (9.14).

**9.3. Model problem (1.3): Singularly perturbed regime in one dimension.** Now, consider  $\varepsilon \ll 1$ . Let  $u$  solve (1.3) with  $\varepsilon \in (0, 1]$ ,  $\gamma > 0$ , posed in the domain  $\Omega := (0, 1)$ . Let  $u_h$  solve the discrete problem (9.1) with  $V_h^j$  and  $\langle \cdot, \cdot \rangle_h$  defined, for each time level  $t_j$ , as described in section A.2. Thus we use the space  $V_h^j$  of continuous piecewise-linear finite element functions on an arbitrary nonuniform mesh  $\{x_i^j\}_{i=1}^N$  with  $0 = x_0^j < x_1^j < \dots < x_N^j = 1$  under absolutely no mesh regularity assumptions. Two choices (A.2a) and (A.2b) of  $\langle \cdot, \cdot \rangle_h$  are discussed in section A.2; both should now use the piecewise-linear interpolant  $I_h := I_h^j$  onto  $V_h^j$ .

**COROLLARY 9.6.**<sup>3</sup> *Let the above numerical method be applied to problem (1.3) with  $\varepsilon \in (0, 1]$ ,  $\gamma > 0$ ,  $\Omega := (0, 1)$ . Then the a posteriori error estimates of Theorems 9.2 and 9.2\* are valid with  $\vartheta_h$  simplified to (8.12) and estimated as described in Remark 8.4, and*

$$\eta^j := \eta_\varepsilon(V_h^j, f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j) \quad \text{for } j = 1, \dots, M,$$

where  $\eta_\varepsilon$  is defined in (A.4) with  $I_h$  replaced by  $I_h^j$ .

We also refer the reader to a recent paper [16], where we obtain a similar but slightly sharper result by using a more intricate direct analysis that invokes sharp bounds of the spatial derivatives of the parabolic Green's function.

*Remark 9.7.* The a posteriori error estimators of Corollary 9.6 are *robust*. Indeed, the only terms in (8.3) that involve the small parameter  $\varepsilon$  are the spatial estimators  $\eta^j$ , whose robustness can be discussed similarly to the steady-state case; see Remark A.2 below. In fact, this remark applies to  $\eta^j$  with  $g_* = f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j$  and  $v$  replaced by  $u(\cdot, t_j)$ . Thus  $\eta^j$  involves  $\varepsilon^{-2} h_i^2 |I_h g_*|$ , which approximates  $h_i^2 |\partial_x^2 u(\cdot, t_j)|$ , and also  $\varepsilon^{-1} h_i^2 |\partial_x(I_h g_*)|$ , which approximates  $\varepsilon |\partial_x^3 u(\cdot, t_j)|$  and has similar magnitude to  $h_i^2 |\partial_x^2 u(\cdot, t_j)|$  in the layer regions.

Furthermore, the numerical results in [16, section 4 with Remark 3.2] show that at least on a fixed layer-adapted mesh, our estimator is quite efficient independently of  $\varepsilon$ .

*Remark 9.8.* Consider the ingredient  $\|g_* - I_h^j g_*\|_{\infty, \Omega}$  in the spatial estimator  $\eta^j$  of Corollary 9.6 for Cases A and B of (8.2). In Case A, one has  $\delta_t^* u_h^j - I_h^j[\delta_t^* u_h^j] = 0$ , hence  $\|g_* - I_h^j g_*\|_{\infty, \Omega}$  simplifies to  $\|f(\cdot, t_j, u_h^j) - I_h^j f(\cdot, t_j, u_h^j)\|_{\infty, \Omega}$ . In Case B, the final

<sup>3</sup>By plugging the elliptic estimators of [15, 4] into (8.3), one can extend this corollary to (1.3) in two and three dimensions.

term in (8.3) vanishes. However,  $g_* - I_h^j g_*$  again involves  $f(\cdot, t_j, u_h^j) - I_h^j f(\cdot, t_j, u_h^j)$  and, furthermore,  $\delta_t^* u_h^j - I_h^j [\delta_t^* u_h^j] = -\frac{1}{\tau_j} (u_h^{j-1} - I_h^j [u_h^{j-1}])$ .

Interestingly, Case A and Case B with  $I_*^j := I_h^j$  are identical but, in view of the above, yield different error estimators. Note that one seems to get a sharper estimator when this method is interpreted as Case A with  $I_*^j := I_h^j$ .

*Remark 9.9.* The backward Euler method for (1.3) with  $\varepsilon \ll 1$  is a particular case of a singularly perturbed convection-reaction-diffusion equation considered in [3]; however, the a posteriori estimate for this equation in [3] is not robust as, e.g., it involves the term  $\varepsilon^{-1} \max_j \|u_h^j - u_h^{j-1}\|_{\infty, \Omega}$  (rather than  $\max_j \|u_h^j - \hat{u}_h^{j-1}\|_{\infty, \Omega}$ , which appears in our estimator).

Similarly, the a posteriori error estimates [2] for a singularly perturbed Allen–Cahn equation (given in the weaker  $L_\infty(L_2)$  norm) involve negative powers of  $\varepsilon$  in various terms. Note that the analysis in [2] invokes elliptic reconstructions for a semilinear parabolic equation, but in contrast to our definition (8.7), they are defined as solutions to linear Laplace equations. It should also be noted that the results of this paper do not directly apply to the Allen–Cahn equation because of the monotonicity assumption (1.2), which is essential in our estimation of the Green’s function. If (1.2) is replaced with  $|\partial_z f(x, t, z)| \leq \bar{\gamma}^2$ , then in some cases, our results can be extended (see [16, Remark 2.2]), only the error estimate will involve additional factors of type  $e^{\bar{\gamma}^2 t_m}$  (which, however, become unbounded if long-term computations are required).

**10. Fully discrete Crank–Nicolson method.** Consider a fully discrete Crank–Nicolson method for (1.3), obtained by applying the spatial discretization (7.5) to the semidiscrete problem (5.1): Find  $u_h^j \in \hat{V}_h^j$  such that

$$(10.1) \quad \varepsilon^2 \langle \frac{1}{2} \nabla (\hat{u}_h^{j-1} + u_h^j), \nabla w_h \rangle + \langle \frac{1}{2} [f(\cdot, t_{j-1}, \hat{u}_h^{j-1}) + f(\cdot, t_j, u_h^j)] + \delta_t^* u_h^j, w_h \rangle_h = 0$$

for all  $w_h \in \hat{V}_h^j$ ,  $j = 1, \dots, M$ , where  $\langle \cdot, \cdot \rangle_h$  is either exactly the inner product  $\langle \cdot, \cdot \rangle$  in  $L_2(\Omega)$  or some quadrature formula for  $\langle \cdot, \cdot \rangle$ . Here a computed solution  $u_h^j \in \hat{V}_h^j := V_h^j \cap H_0^1(\Omega)$  and an auxiliary computed solution  $\hat{u}_h^{j-1} \in H_0^1(\Omega)$  are respectively associated with the time levels  $t_j$  and  $t_{j-1}^+$  (the latter is reflected in the hat notation).

The term  $\delta_t^* u_h^j$  approximates  $\partial_t u$  and is identical with (9.2a):

$$(10.2a) \quad \delta_t^* u_h^j := \frac{u_h^j - \hat{u}_h^{j-1}}{\tau_j}, \quad \text{where } \hat{u}_h^0 := u_h^0.$$

The operator  $\delta_t^*$  is identical with  $\delta_t$  of (5.1b) for  $j = 1$ , while for  $j > 1$  it involves  $\hat{u}_h^{j-1} \in H_0^1(\Omega)$ , for which we note possible choices (8.2).

Note that the full discretization (10.1) can be represented as

$$(10.2b) \quad \mathcal{P}_h^j [\delta_t^* u_h^j] + \frac{1}{2} (\hat{\mathcal{L}}_h^{j-1} \hat{u}_h^{j-1} + \mathcal{L}_h^j u_h^j) + \frac{1}{2} \mathcal{P}_h^j [f(\cdot, t_{j-1}, \hat{u}_h^{j-1}) + f(\cdot, t_j, u_h^j)] = 0$$

with  $\hat{\mathcal{L}}_h^{j-1} := \mathcal{L}_h(t_{j-1}^+)$ ,  $\mathcal{L}_h^j := \mathcal{L}_h(t_j)$  and  $\mathcal{P}_h^j$  subject to (8.1). For some particular cases of  $\langle \cdot, \cdot \rangle_h$ , the operators  $\mathcal{L}_h^j$  and  $\mathcal{P}_h^j$  are defined as in Remarks 7.3 and 7.4 only using  $V_h^j$  instead of  $V_h$ . Furthermore, (10.2) can be considered a full discretization for the abstract parabolic problem (1.1) obtained by applying a spatial discretization of type (7.2) to the semidiscrete problem (5.1).

Note that by virtue of (8.1),  $\hat{\mathcal{L}}_h^{j-1} \hat{u}_h^{j-1} \in \hat{V}_h^j - I_h[f(\cdot, t_{j-1}, 0)]$  and  $\mathcal{L}_h^j u_h^j \in \hat{V}_h^j - I_h[f(\cdot, t_j, 0)]$ , while for any  $w \in H_0^1(\Omega)$  one has  $\hat{V}_h^j - I_h[f(\cdot, t_k, 0)] = \hat{V}_h^j - I_h[f(\cdot, t_k, w)]$ . As we also have  $\delta_t^* u_h^j \in H_0^1(\Omega)$ , the definition (10.2b) is consistent.

**10.1. A posteriori error estimate using piecewise-linear elliptic reconstructions.** To estimate the error of the fully discrete Crank–Nicolson method (10.2), set  $\mathcal{A} := \{0^+, 1\}$  and recall the elliptic reconstructions  $\hat{R}^{j-1}$  and  $R^j$  defined for  $j = 1, \dots, M$  by (8.7). These definitions involve  $\hat{g}^{j-1}$  and  $g^j$ , defined in (8.5c), which in their turn involve  $\hat{\psi}_h^{j-1}$  and  $\psi_h^j$  that we now define by

$$(10.3) \quad \hat{\psi}_h^{j-1} := -\psi_h^j - 2\delta_t^* u_h^j, \quad \psi_h^j := \mathcal{L}_h(t_j) u_h^j + \mathcal{P}_h^j[f(\cdot, t_j, u_h^j)].$$

Note that the first relation here yields

$$(10.4) \quad \delta_t^* u_h^j + \frac{1}{2}(\hat{\psi}_h^{j-1} + \psi_h^j) = 0.$$

*Remark 10.1.* The definition of  $\psi_h^j$  in (10.3) implies  $\psi_h^j \in \mathring{V}_h^j$  so  $\mathcal{P}_h^j \psi_h^j = \psi_h^j$  so  $\psi_h^j$  satisfies (8.4b). Next,  $\hat{\psi}_h^{j-1}$  of (10.3) satisfies  $\mathcal{P}_h^j \hat{\psi}_h^{j-1} = -\psi_h^j - 2\mathcal{P}_h^j[\delta_t^* u_h^j]$  for any  $\hat{u}^{j-1} \in H_0^1(\Omega)$ , which, in view of (10.2b), yields  $\mathcal{P}_h^j \hat{\psi}_h^{j-1} = \mathcal{L}_h(t_{j-1}^+) \hat{u}_h^{j-1} + \mathcal{P}_h^j[f(\cdot, t_{j-1}, \hat{u}_h^{j-1})]$ , i.e.,  $\hat{\psi}_h^{j-1}$  satisfies (8.4a). Therefore  $\hat{R}^{j-1}$  and  $R^j$  satisfy (8.9).

*Remark 10.2.* Theorem 10.3 and further results of this section remain valid for any pair  $\hat{\psi}_h^{j-1}, \psi_h^j$  that satisfy (8.4) and (10.4). For example, alternatively to the definition of  $\psi_h^j$  in (10.3), one can use  $\psi_h^j := \mathcal{L}_h(t_j) u_h^j + f(\cdot, t_j, u_h^j)$ , but this modification does not seem to improve the computability of  $\psi_h^j$ .

To formulate our a posteriori error estimate for  $u_h - u$ , we generalize the piecewise-linear interpolation  $I_{1,t}$  of (5.2) to any *left-continuous* function  $w = w(t)$  by setting

$$(10.5) \quad I_{1,t}^* w(t) = \frac{t_j - t}{\tau_j} w(t_{j-1}^+) + \frac{t - t_{j-1}}{\tau_j} w(t_j) \quad \text{for } t \in (t_{j-1}, t_j], \quad j = 1, \dots, M.$$

In a similar manner, we apply the *piecewise-linear* interpolation  $I_{1,t}^*$  to the elliptic reconstructions  $\hat{R}^{j-1}$  and  $R^j$  associated with the time levels  $t_{j-1}^+$  and  $t_j$  and define

$$(10.6) \quad \tilde{R}(\cdot, t) = \frac{t_j - t}{\tau_j} \hat{R}^{j-1} + \frac{t - t_{j-1}}{\tau_j} R^j \quad \text{for } t \in (t_{j-1}, t_j], \quad j = 1, \dots, M, \quad \tilde{R}(\cdot, 0) = \hat{R}^0.$$

Note that we impose that both  $\tilde{R}$  and  $I_{1,t}^* w$  are right-continuous at  $t = 0$ .

**THEOREM 10.3.** *Let  $u$  solve problem (1.1), (1.2) with a parabolic operator  $\mathcal{M}$  satisfying Condition 2.1, and let  $u_h^j$  solve the discrete problem (10.2). Then for  $m = 1, \dots, M$ , one has (8.3) with  $\chi_h^j = \tau_j (\psi_h^j - \hat{\psi}_h^{j-1})$  using  $\psi_h^j$  and  $\hat{\psi}_h^{j-1}$  from (10.3),  $\eta^j$  from (8.5) with  $\mathcal{A} = \{0^+, 1\}$ ,  $C_1 = \frac{1}{8}$ ,  $C_2 = \frac{1}{2}$ ,  $C_1^* = 2$ ,  $C_2^* = 3$ , and  $\vartheta_h$  defined by*

$$(10.7) \quad \vartheta_h := \tilde{\psi}_R - I_{1,t}^* \tilde{\psi}_R, \quad \tilde{\psi}_R := \mathcal{L}(t) \tilde{R} + f(\cdot, t, \tilde{R})$$

for  $t \in [0, T]$  with  $I_{1,t}^*$  and  $\tilde{R}$  from (10.5) and (10.6).

*Proof.* As Remark 10.1 gives  $\|R^j - u_h^j\|_{\infty, \Omega} \leq \eta^j$ , so to get the desired bound (8.3) for  $u_h^m - u(\cdot, t_m)$ , it suffices to obtain a bound of type (8.3) for  $R^m - u(\cdot, t_m) = [\tilde{R} - u](\cdot, t_m)$ , with  $(C_2^* \kappa_0 + 1)$  replaced by  $C_2^* \kappa_0 = 3\kappa_0$ , and then apply the triangle inequality. So we consider  $\tilde{R} - u$  only.

We partially imitate the proof of Theorem 5.1. For  $t \in (t_{j-1}, t_j]$ , one has  $\tilde{\psi}_R = I_{1,t}^* \tilde{\psi}_R + \vartheta_h = \frac{1}{2}(\hat{\psi}_h^{j-1} + \psi_h^j) + \partial_t \mu_h + \vartheta_h$ , where  $\mu_h := \int_{t_j}^t [I_{1,t}^* \tilde{\psi}_R - \frac{1}{2}(\hat{\psi}_h^{j-1} + \psi_h^j)] dt$ . Note that with our definition of  $\tilde{R}$ , the functions in (10.7) are identical with those in (8.10) (using  $p = 2$ ), so we also enjoy the observation (8.11), which can be rewritten as  $\tilde{\psi}_R(\cdot, t_{j-1}^+) = \hat{\psi}_h^{j-1}$  and  $\tilde{\psi}_R(\cdot, t_j) = \psi_h^j$ . Consequently, we get a version of (5.4):

$$(10.8) \quad \mu_h = \tau_j^{-1} \chi_h^j \int_{t_j}^t (t - t_{j-1/2}) dt = -\frac{1}{2}(t_j - t)(t - t_{j-1}) \cdot \tau_j^{-2} \chi_h^j \quad \text{for } t \in [t_{j-1}, t_j].$$

Combining our findings with (10.4) yields

$$(10.9) \quad \delta_t^* u_h^j + \tilde{\psi}_R = \partial_t \mu_h + \vartheta_h \quad \text{for } t \in (t_{j-1}, t_j].$$

Next, we invoke  $I_{1,t}^* u_h$  defined by (9.11), for which we have (9.12) and (9.13). As  $\mathcal{M}\tilde{R} = \partial_t \tilde{R} + \tilde{\psi}_R$  and  $\mathcal{M}u = 0$ , so (10.9) implies that

$$(10.10) \quad \mathcal{M}\tilde{R} - \mathcal{M}u = \partial_t(\tilde{R} - I_{1,t}^* u_h) + \partial_t \mu_h + [\vartheta_h + \vartheta_*] \quad \text{in } Q.$$

Note that here  $\partial_t \mu_h$  is a regular function as  $\mu_h$  of (10.8) is continuous for  $t \in [0, T]$ .

Now the desired bound of type (8.3) for  $R^m - u(\cdot, t_m) = [\tilde{R} - u](\cdot, t_m)$ , only with  $(C_2^* \kappa_0 + 1)$  replaced by  $C_2^* \kappa_0 = 3\kappa_0$ , is obtained by an application of Lemma 2.3 to (10.10) with  $\mu := \mu_R + \mu_h := (\tilde{R} - I_{1,t}^* u_h) + \mu_h$  and  $\vartheta := \vartheta_h + \vartheta_*$ , using (9.13) and the following two observations. First,  $[\tilde{R} - u - \mu](\cdot, 0) = \hat{R}^0 - \varphi - [(\hat{R}^0 - u_h^0) + 0] = u_h^0 - \varphi$ . Second, for  $t \in (t_{j-1}, t_j]$ , we have

$$\begin{aligned} |\mu_R| &\leq |\hat{R}^{j-1} - \hat{u}_h^{j-1}| + |R^j - u_h^j| \leq 2\eta^j, & |\mu_h| &\leq \frac{1}{8} |\chi_h^j|, \\ \tau_j |\partial_t \mu_R| &\leq |\hat{R}^{j-1} - \hat{u}_h^{j-1}| + |R^j - u_h^j| \leq 2\eta^j, & \tau_j |\partial_t \mu_h| &\leq \frac{1}{2} |\chi_h^j|, \end{aligned}$$

where we used  $\mu_R = \tilde{R} - I_{1,t}^* u_h = I_{1,t}^*(R - u_h)$  combined with Remark 10.1, and also (10.8). Finally note that  $|\mu(\cdot, t_{m-1}^+)| = |\mu_R(\cdot, t_{m-1}^+)| \leq \eta^m$  and so  $\eta^m$  is multiplied by  $1 + 2 = 3 = C_2^*$ . This completes the proof.  $\square$

**10.2. Model problem (1.3): Regular regime.** Let  $u$  solve problem (1.3) with  $\varepsilon = 1$ ,  $\gamma \geq 0$ , posed in a bounded polyhedral spatial domain  $\Omega \subset \mathbb{R}^n$ ,  $n = 2, 3$ , and  $u_h^j$  solve the discrete problem (10.1) with  $V_h^j$  and  $\langle \cdot, \cdot \rangle_h$  defined, for each time level  $t_j$ , as in section 9.2.

**COROLLARY 10.4.** *Let the above numerical method be applied to problem (1.3) with  $\varepsilon = 1$ ,  $\gamma \geq 0$ . Then the a posteriori error estimate of Theorem 10.3 is valid with  $\vartheta_h$  simplified to (8.12), where  $p = 2$ . The definition (8.5b) of  $\eta^{j+1}$  uses  $\eta := \eta_0$ , where  $\eta_0$  is from (A.1).*

**10.3. Model problem (1.3): Singularly perturbed regime in one dimension.** Now consider the regime of  $\varepsilon \ll 1$ . Let  $u$  solve the problem (1.3) with  $\varepsilon \in (0, 1]$ ,  $\gamma > 0$ , posed in the domain  $\Omega := (0, 1)$ , and  $u_h$  solve the discrete problem (10.1) with  $V_h^j$  and  $\langle \cdot, \cdot \rangle_h$  defined, for each time level  $t_j$ , as in section 9.3.

**COROLLARY 10.5.** *Let the above numerical method be applied to problem (1.3) with  $\varepsilon \in (0, 1]$ ,  $\gamma > 0$ ,  $\Omega := (0, 1)$ . Then the a posteriori error estimate of Theorem 10.3 is valid with  $\vartheta_h$  simplified to (8.12) for  $p = 2$ . The definition (8.5b) of  $\eta^{j+1}$  uses  $\eta := \eta_\varepsilon$  with  $\eta_\varepsilon$  defined in (A.4), in which  $I_h$  is now replaced by  $I_h^j$ .*

**11. Fully discrete discontinuous Galerkin method dG(1).** To simplify the presentation, in this section we mainly focus on dG( $r$ ) with  $r = 1$ ; for  $r > 1$  see Remark 11.5. Consider a fully discrete discontinuous Galerkin method dG(1) for (1.3), obtained by applying the spatial discretization (7.5) to the semidiscrete problem (6.7): For  $j = 0, \dots, M - 1$ , find  $u_h^{j+1/3}, u_h^{j+1} \in \hat{V}_h^{j+1}$  such that

$$(11.1) \quad \begin{aligned} \left\langle \frac{u_h^{j+1} - \hat{u}_h^j}{\tau_{j+1}}, w_h \right\rangle_h + \varepsilon^2 \left\langle \frac{1}{4} \nabla (3u_h^{j+1/3} + u_h^{j+1}), \nabla w_h \right\rangle + \frac{1}{4} \left\langle 3f_h^{j+1/3} + f_h^{j+1}, w_h \right\rangle_h &= 0, \\ \left\langle \frac{u_h^{j+1/3} - \hat{u}_h^j}{\tau_{j+1}}, w_h \right\rangle_h + \varepsilon^2 \left\langle \frac{1}{12} \nabla (u_h^{j+1/3} - u_h^{j+1}), \nabla w_h \right\rangle + \frac{1}{12} \left\langle 5f_h^{j+1/3} - f_h^{j+1}, w_h \right\rangle_h &= 0 \end{aligned}$$

for all  $w_h \in \mathring{V}_h^{j+1}$ , with the notation  $f_h^{j+\alpha} := f(\cdot, t_{j+\alpha}, u_h^{j+\alpha})$ . Here  $\langle \cdot, \cdot \rangle_h$  is either exactly the inner product  $\langle \cdot, \cdot \rangle$  in  $L_2(\Omega)$  or some quadrature formula for  $\langle \cdot, \cdot \rangle$ . Note possible choices (8.2) for  $\hat{u}_h^j$ .

Note that the full discretization (11.1) can be represented as

$$(11.2a) \quad \mathcal{P}_h^{j+1}[u_h^{j+1} - \hat{u}_h^j] + \frac{1}{4} \tau_{j+1} (3\Psi^{j+1/3} + \Psi^{j+1}) = 0,$$

$$(11.2b) \quad \mathcal{P}_h^{j+1}[u_h^{j+1/3} - \hat{u}_h^j] + \frac{1}{12} \tau_{j+1} (5\Psi^{j+1/3} - \Psi^{j+1}) = 0,$$

where

$$(11.2c) \quad \Psi^{j+\alpha} := \mathcal{L}_h(t_{j+\alpha}) u_h^{j+\alpha} + \mathcal{P}_h^{j+1}[f(\cdot, t_{j+\alpha}, u_h^{j+\alpha})] \quad \text{for } \alpha \in \{\frac{1}{3}, 1\}$$

with  $\mathcal{L}_h(t)$  and  $\mathcal{P}_h^{j+1}$  subject to (8.1). For some particular cases of  $\langle \cdot, \cdot \rangle_h$ , the operators  $\mathcal{L}_h(t)$  and  $\mathcal{P}_h^{j+1}$  are defined as in Remarks 7.3 and 7.4 except using  $V_h^{j+1}$  instead of  $V_h$ . Furthermore, (11.2) gives a full discretization of dG(1) type for the abstract parabolic problem (1.1) obtained by applying a spatial discretization of type (7.2) to the semidiscrete problem (7.5).

Note that  $u_h^{j+\alpha}$  vanishes on  $\partial\Omega$ , so  $I_h[f(\cdot, t_{j+\alpha}, 0)] = I_h[f(\cdot, t_{j+\alpha}, u_h^{j+\alpha})]$  on  $\partial\Omega$ , so  $\Psi^{j+\alpha} \in \mathring{V}_h^{j+1}$  for  $\alpha \in \{\frac{1}{3}, 1\}$ , and hence (11.2) is consistent.

**11.1. A posteriori error estimate using piecewise-quadratic elliptic reconstructions.** To estimate the error of the fully discrete dG(1) method (11.2), we partially imitate the arguments of section 6 for the related semidiscrete method. First, set  $\mathcal{A} := \{0^+, \frac{1}{3}, 1\}$  and recall the elliptic reconstructions  $\hat{R}^j$ ,  $R^{j+1/3}$ , and  $R^{j+1}$  defined by (8.7). These definitions involve  $\hat{g}^j$ ,  $g^{j+1/3}$ , and  $g^{j+1}$ , defined in (8.5c), which in their turn involve  $\hat{\psi}_h^j$ ,  $\psi_h^{j+1/3}$ , and  $\psi_h^{j+1}$  that we now define by

(11.3a)

$$\psi_h^{j+1/3} := -\frac{u_h^{j+1} - \hat{u}_h^j}{2\tau_{j+1}} - 3\frac{u_h^{j+1/3} - \hat{u}_h^j}{2\tau_{j+1}}, \quad \psi_h^{j+1} := -5\frac{u_h^{j+1} - \hat{u}_h^j}{2\tau_{j+1}} + 9\frac{u_h^{j+1/3} - \hat{u}_h^j}{2\tau_{j+1}},$$

$$(11.3b) \quad \hat{\psi}_h^j := \mathcal{L}_h(t_j^+) \hat{u}_h^j + \mathcal{P}_h^{j+1}[f(\cdot, t_j, \hat{u}_h^j)].$$

Note that (11.3a) implies a version of (11.2):

$$(11.4a) \quad u_h^{j+1} - \hat{u}_h^j + \frac{1}{4} \tau_{j+1} (3\psi_h^{j+1/3} + \psi_h^{j+1}) = 0,$$

$$(11.4b) \quad u_h^{j+1/3} - \hat{u}_h^j + \frac{1}{12} \tau_{j+1} (5\psi_h^{j+1/3} - \psi_h^{j+1}) = 0.$$

In fact, if  $\hat{u}_h^j \in \mathring{V}_h^{j+1}$  (Case A of (8.2a)), then (11.4) and (11.2) are equivalent (and one has  $\Psi^{j+\alpha} = \psi_h^{j+\alpha}$  for  $\alpha \in \{\frac{1}{3}, 1\}$ ).

*Remark 11.1.* A comparison of (11.4) and (11.2) implies  $\Psi^{j+\alpha} = \mathcal{P}_h^{j+1} \psi_h^{j+\alpha}$  for  $\alpha \in \{\frac{1}{3}, 1\}$ . So, by virtue of (11.2c), one concludes that  $\psi_h^{j+1/3}$  and  $\psi_h^{j+1}$  satisfy (8.4b). Next,  $\hat{\psi}_h^j$  of (11.3b) is in  $\mathring{V}_h^{j+1}$ , so  $\hat{\psi}_h^j$  satisfies (8.4a). Consequently,  $\hat{R}^j$ ,  $R^{j+1/3}$  and  $R^{j+1}$  satisfy (8.9) with  $\mathcal{A} := \{0^+, \frac{1}{3}, 1\}$  and  $\eta^{j+1}$  of (8.5b).

*Remark 11.2.* Theorem 11.4 and further results of this section remain valid for any triple  $\hat{\psi}_h^j, \psi_h^{j+1/3}, \psi_h^{j+1}$  that satisfy (8.4) and (11.4). For example, one can replace  $\mathcal{P}_h^{j+1}[f(\cdot, t_j, \hat{u}_h^j)]$  in (11.3b) by  $f(\cdot, t_j, \hat{u}_h^j)$ , but this modification does not seem to improve the computability of  $\hat{\psi}_h^j$ .

*Remark 11.3 (Case B).* In case (8.2b) with  $\hat{u}_h^j := u_h^j$ , it is more natural to replace (11.3b) by  $\hat{\psi}_h^j := \psi_h^j$  (and this makes  $\hat{\psi}_h^j$  easily explicitly computable). Then (8.4a) is

no longer true, but we still enjoy (8.9) provided that we replace  $\eta(V_h^{j+1}, \hat{u}_h^j, \hat{g}^j(\cdot, \hat{u}_h^j))$  in the definition (8.5b) of  $\eta^{j+1}$  by  $\eta(V_h^j, u_h^j, g^j(\cdot, u_h^j))$ . Consequently, Theorem 11.4 and further results of this section remain valid for these modifications.

To formulate our a posteriori error estimate for  $u_h - u$ , we generalize the *piecewise-quadratic interpolation*  $I_{2,t}$  of (6.8) to any *left-continuous* function  $w = w(t)$  by using the interpolation nodes  $t_j^+$ ,  $t_{j+1/3}$  and  $t_{j+1}$ , so

$$(11.5) \quad \begin{aligned} I_{2,t}^* w(0) &:= w^0, \quad I_{2,t}^* w(t) := \hat{I}_{1,t} w(t) + 3 [2\hat{w}^j - 3w^{j+1/3} + w^{j+1}] \cdot \frac{1}{6} \tau_{j+1}^{-1} \zeta_2\left(\frac{t-t_j}{\tau_{j+1}}\right), \\ \text{where } \hat{I}_{1,t} w(t) &:= w^{j+1} - \{w^{j+1} - w^{j+1/3}\} \cdot \frac{3}{2} \left(\frac{t_{j+1}-t}{\tau_{j+1}}\right) \end{aligned}$$

for  $t \in (t_j, t_{j+1}]$ ,  $j = 0, \dots, M - 1$ , with the notation  $\hat{w}^j := w(t_j^+)$  and  $w^{j+\alpha} := w(t_{j+\alpha})$ . By applying  $I_{2,t}^*$  to the elliptic reconstructions  $\hat{R}^j$ ,  $R^{j+1/3}$ , and  $R^{j+1}$  associated with the time levels  $t_j^+$ ,  $t_{j+1/3}$ , and  $t_{j+1}$ , we now define

$$(11.6) \quad \tilde{R}(\cdot, 0) := \hat{R}^0 = R^0, \quad \tilde{R}(\cdot, t) := I_{2,t}^* \{\hat{R}^j, R^{j+1/3}, R^{j+1}\} \quad \text{for } t \in (0, T].$$

Similarly, define a piecewise-quadratic computed solution in time by

$$(11.7) \quad \tilde{u}_h(\cdot, 0) := \hat{u}_h^0 = u_h^0, \quad \tilde{u}_h(\cdot, t) := I_{2,t}^* \{\hat{u}_h^j, u_h^{j+1/3}, u_h^{j+1}\} \quad \text{for } t \in (0, T].$$

We are now prepared to formulate our main result for the dG(1) method.

**THEOREM 11.4.** *Let  $u$  solve problem (1.1), (1.2) satisfying Condition 2.1,  $u_h^j$  solve the discrete problem (11.2) with any  $\hat{u}_h^j \in H_0^1(\Omega)$ . Then one has (8.3) with*

$$(11.8) \quad \chi_h^{j+1} := 3\tau_{j+1} [2\hat{\psi}_h^j - 3\psi_h^{j+1/3} + \psi_h^{j+1}]$$

using  $\hat{\psi}^j$ ,  $\psi_h^{j+1/3}$ , and  $\psi_h^{j+1}$  of (11.3),  $\eta^j$  from (8.5) with  $\mathcal{A} = \{0^+, \frac{1}{3}, 1\}$ , the constants  $C_1 = \frac{2}{81}$ ,  $C_2 = \frac{1}{6}$ ,  $C_1^* = \frac{5}{3}$ ,  $C_2^* = 10$ , and  $\vartheta_h$  defined by

$$(11.9) \quad \vartheta_h := \tilde{\psi}_R - I_{2,t}^* \tilde{\psi}_R, \quad \tilde{\psi}_R := \mathcal{L}(t)\tilde{R} + f(\cdot, t, \tilde{R})$$

for  $t \in [0, T]$ , with  $I_{2,t}^*$  and  $\tilde{R}$  from (11.5) and (11.6).

*Proof.* As Remark 11.1 gives  $\|R^j - u_h^j\|_{\infty, \Omega} \leq \eta^j$ , to get the desired bound (8.3) for  $u_h^m - u(\cdot, t_m)$  it suffices to obtain a bound of type (8.3) for  $R^m - u(\cdot, t_m) = [\tilde{R} - u](\cdot, t_m)$ , with  $(C_2^* \kappa_0 + 1)$  replaced by  $C_2^* \kappa_0 = 10\kappa_0$ , and then apply the triangle inequality. So we consider  $\tilde{R} - u$  only.

We partially imitate the proof of Theorem 6.1. On each  $(t_j, t_{j+1}]$ , the function  $\tilde{u}_h$  is quadratic in time and satisfies  $\partial_t \tilde{u}_h + \hat{I}_{1,t} \psi_h = 0$  (where  $\hat{I}_{1,t}$  is specified in (11.5)). This relation is a version of  $\partial_t \tilde{U} + \hat{I}_{r,t} \psi_h = 0$  used in the proof of Theorem 6.1 and can be obtained similarly. Alternatively, it can be checked by a direct calculation using (11.7), (11.5), and (11.3a).

It is convenient to treat the left-continuous function  $\tilde{u}_h$  of (11.7) as being discontinuous at  $t_j^+$  rather than at  $t_j$ . Now, letting  $t \in (0, T]$ , one gets

$$(11.10) \quad \partial_t \tilde{u}_h + \hat{I}_{1,t} \psi_h = \vartheta_* \quad \text{in } Q.$$

Here the discontinuity of  $\tilde{u}_h$  at  $t_j^+$  yielded the term

$$(11.11) \quad \vartheta_*(\cdot, t) := [\hat{u}_h^j - u_h^j] \delta(t - t_j^+) \quad \text{for } t \in (t_j, t_{j+1}],$$

which is identical to  $\vartheta_*$  of (9.12) and so satisfies (9.13).



Next, note that with our definition of  $\tilde{R}$ , the functions in (11.9) are identical with those in (8.10) (using  $p = 3$ ), so we also enjoy the observation (8.11), which can be rewritten as  $\tilde{\psi}_R(\cdot, t_j^+) = \hat{\psi}_h^j$  and  $\tilde{\psi}_R(\cdot, t_{j+\alpha}) = \psi_h^{j+\alpha}$  for  $\alpha \in \{\frac{1}{3}, 1\}$ . Furthermore, a comparison of  $\chi_h^j$  in (11.8) with the definition (11.5) of  $I_{2,t}^*$  implies that  $\chi_h^{j+1} = \tau_{j+1}^3 \partial_t^2(I_{2,t}^* \tilde{\psi}_R)$  for  $t \in (t_j, t_{j+1}]$ . Consequently,  $I_{2,t}^* \tilde{\psi}_R = \hat{I}_{1,t} \psi_h + \chi_h^{j+1} \cdot \frac{1}{6} \tau_{j+1}^{-1} \zeta_2(\frac{t-t_j}{\tau_{j+1}})$ . Combining this with (11.10) yields

$$\partial_t \tilde{u}_h + I_{2,t}^* \tilde{\psi}_R = \partial_t \mu_h + \vartheta_*, \quad \text{where} \quad \mu_h := \chi_h^{j+1} \cdot \frac{1}{6} \tau_{j+1}^{-1} \int_{t_j}^t \zeta_2(\frac{t-t_j}{\tau_{j+1}}).$$

For this function  $\mu_h$  (similar to  $\mu$  in section 6.1), a calculation yields

$$(11.12) \quad |\mu_h| \leq \frac{2}{81} |\chi_h^{j+1}|, \quad \tau_{j+1} |\partial_t \mu_h| \leq \frac{1}{6} |\chi_h^{j+1}| \quad \text{for } t \in (t_j, t_{j+1}].$$

As  $\mathcal{M}\tilde{R} = \partial_t \tilde{R} + \tilde{\psi}_R$  and  $\mathcal{M}u = 0$ , while  $\tilde{\psi}_R = I_{2,t}^* \tilde{\psi}_R + \vartheta_h$ , so

$$\mathcal{M}\tilde{R} - \mathcal{M}u = \partial_t [\tilde{R} - \tilde{u}_h + \mu_h] + \vartheta_h + \vartheta_* \quad \text{in } Q.$$

Now the desired bound of type (8.3) for  $R^m - u(\cdot, t_m) = [\tilde{R} - u](\cdot, t_m)$ , only with  $(C_2^* \kappa_0 + 1)$  replaced by  $C_2^* \kappa_0 = 10\kappa_0$ , is obtained by an application of Lemma 2.3 with  $\mu := (\tilde{R} - \tilde{u}_h) + \mu_h$  and  $\vartheta := \vartheta_h + \vartheta_*$ , for which we make a few observations. First, note that  $[\tilde{R} - u - \mu](\cdot, 0) = R^0 - \varphi - (R^0 - u_h^0) = u_h^0 - \varphi$ . For  $\mu_h$ , we recall (11.12) and also note that  $\mu_h(\cdot, t_{m-1}^+) = 0$ . For the piecewise-quadratic function  $\tilde{R} - \tilde{u}_h$ , by virtue of Remark 11.1  $|\tilde{R} - \tilde{u}_h(\cdot, t_{m-1}^+)| \leq \eta^m$ , and a calculation yields

$$(11.13) \quad |\tilde{R} - \tilde{u}_h| \leq \frac{5}{3} \eta^{j+1}, \quad \tau_{j+1} |\partial_t (\tilde{R} - \tilde{u}_h)| \leq 9 \eta^{j+1} \quad \text{for } t \in (t_j, t_{j+1}],$$

so  $\eta^m$  is multiplied by  $1 + 9 = 10 = C_2^*$ . Finally, for  $\vartheta_*$ , we invoke (11.11). Combining these observations in the application of Lemma 2.3 completes the proof.  $\square$

*Remark 11.5* (dG( $r$ ) for  $r > 1$ ). The results of the section, including the a posteriori error estimate of Theorem 11.4, can be generalized to a fully discrete dG( $r$ ) method with Radau quadrature for  $r > 1$  (in lines with the analysis of section 6 for a semidiscrete dG( $r$ ) method). In fact, then the error estimate (8.3) will involve  $\chi_h^{j+1}$  defined by (6.5) with  $\psi^{j+\alpha}$  replaced by  $\psi_h^{j+\alpha}$  and the same constants  $C_1$  and  $C_2$  as in Theorem 6.1.

**11.2. Model problem (1.3): Regular regime.** Let  $u$  solve problem (1.3) with  $\varepsilon = 1$ ,  $\gamma \geq 0$ , posed in a bounded polyhedral spatial domain  $\Omega \subset \mathbb{R}^n$ ,  $n = 2, 3$ , and let  $w_h^j$  solve the discrete problem (10.1) with  $V_h^j$  and  $\langle \cdot, \cdot \rangle_h$  defined as in section 9.2.

**COROLLARY 11.6.** *Let the above numerical method be applied to problem (1.3) with  $\varepsilon = 1$ ,  $\gamma \geq 0$ . Then the a posteriori error estimate of Theorem 11.4 is valid with  $\vartheta_h$  simplified to (8.12), where  $p = 3$ , and (8.5b) using  $\eta := \eta_0$ , where  $\eta_0$  from (A.1).*

*Remark 11.7.* A discontinuous Galerkin method dG(1) for a linear version of (1.3) with  $\varepsilon = 1$  was considered in [9, 3]. In this particular case,  $f = f(x, t)$  implies that (11.8) can be rewritten as  $\chi_h^{j+1} = 6\tau_{j+1} \hat{\mathcal{L}}_h^{j+1} [\hat{u}_h^j - \hat{I}_{1,t} u_h(t_j^+)] + \tau_{j+1}^3 \partial_t^2 f(x, t')$  for some intermediate  $t' \in [t_j, t_{j+1}]$ . Here we use the piecewise-linear Radau interpolant  $\hat{I}_{1,t}$  described in (11.5), and the discrete operator  $\hat{\mathcal{L}}_h^{j+1}$  is similar to  $\mathcal{L}_h^{j+1}$  only  $\hat{\mathcal{L}}_h^{j+1} : H_0^1(\Omega) \rightarrow \hat{V}_h^{j+1}$ . With this simplification, the a posteriori error estimate of Corollary 11.6 resembles (but is not identical to) the one of [9, (1.14)] in that it involves terms of type  $\tau_{j+1} \|\hat{\mathcal{L}}_h^{j+1} [\hat{u}_h^j - \hat{I}_{1,t} u_h(t_j^+)]\|_{\infty, \Omega}$  and  $\tau_{j+1}^3 \|\partial_t^2 f\|_{\infty, \Omega}$ . (Note also that [9, (1.14)] is given without proof and does not appear to be proved elsewhere.)

The a posteriori estimate in [3] is of the lower order 2 in time as it involves the terms of type  $\|\hat{u}_h^j - \hat{I}_{1,t}u_h(t_j^+)\|_{\infty,\Omega} = \mathcal{O}(\tau_{j+1}^2)$ . We also note the paper [23], which gives a posteriori estimates for discontinuous Galerkin time discretizations in other norms.

**11.3. Model problem (1.3): Singularly perturbed regime in one dimension.** Now consider the regime of  $\varepsilon \ll 1$ . Let  $u$  solve the problem (1.3) with  $\varepsilon \in (0, 1]$ ,  $\gamma > 0$ , posed in the domain  $\Omega := (0, 1)$ , and let  $u_h$  solve the discrete problem (10.1) with  $V_h^j$  and  $\langle \cdot, \cdot \rangle_h$  defined, for each time level  $t_j$ , as in section 9.3. We consider the two choices (A.2) of  $\langle \cdot, \cdot \rangle_h$ , using nodal piecewise-linear interpolation  $I_h := I_h^j$  onto  $V_h^j$ .

**COROLLARY 11.8.** *Let the above numerical method be applied to problem (1.3) with  $\varepsilon \in (0, 1]$ ,  $\gamma > 0$ ,  $\Omega := (0, 1)$ . Then the a posteriori error estimate of Theorem 11.4 is valid  $\vartheta_h$  simplified to (8.12), where  $p = 3$ . The definition (8.5b) of  $\eta^{j+1}$  uses  $\eta := \eta_\varepsilon$  with  $\eta_\varepsilon$  defined in (A.4), in which  $I_h$  is now replaced by  $I_h^j$ .*

**12. Proof of Lemma 2.2.** First, note that the Green’s function  $\mathcal{G}$  associated with our problem (1.3) in the spatial domain  $\Omega$  and the Green’s function  $\hat{\mathcal{G}}$  for the related problem  $\hat{\mathcal{M}}\hat{u} := \partial_t\hat{u} - \Delta\hat{u} + f(x/\varepsilon, t, \hat{u}) = 0$  in the spatial domain  $\hat{\Omega} := \Omega/\varepsilon$  satisfy  $\|\partial_s^k \mathcal{G}(x, t; \cdot, s)\|_{1,\Omega} = \|\partial_s^k \hat{\mathcal{G}}(x/\varepsilon, t; \cdot, s)\|_{1,\hat{\Omega}}$  for  $k = 0, 1$ . Consequently, it suffices to prove Condition 2.1 for the case of  $\varepsilon = 1$  with  $\kappa_0, \kappa_1$ , and  $\kappa_2$  independent of  $|\Omega|$ , so throughout the proof we set  $\mathcal{L}^* = -\Delta$  in (2.2a).

(i) We start by proving the first bound in Condition 2.1. The Green’s function  $\bar{\mathcal{G}}$  associated with  $\bar{\mathcal{M}} := \partial_t - \Delta + \gamma^2$  in the domain  $\bar{\Omega} := \mathbb{R}^n$  can be easily obtained from the fundamental solution of the heat equation. (The latter can be found, e.g., in [26, section III.3], [11, section 2.3.1].) One gets

$$(12.1) \quad \bar{\mathcal{G}}(x, t; \xi, s) = g(x - \xi, t - s), \quad \text{where} \quad g(x, t) := \frac{e^{-\gamma^2 t}}{(4\pi t)^{n/2}} \exp\left(-\frac{|x|^2}{4t}\right).$$

Next, note that by (1.2), the coefficient  $a$  in (2.2a) satisfies  $a \geq \gamma^2$ , so an application of the maximum principle to problem (2.2) yields  $0 \leq \mathcal{G} \leq \bar{\mathcal{G}}$ . Finally, note that

$$(12.2) \quad \bar{\mathcal{G}}(x, t; \xi, s) \, d\xi = e^{-\gamma^2(t-s)} \psi(\zeta) \, d\zeta, \quad \text{where} \quad \psi(\zeta) := \frac{e^{-|\zeta|^2}}{\pi^{n/2}}, \quad \zeta := \frac{\xi - x}{2\sqrt{t-s}}.$$

As  $\int_{\mathbb{R}^n} \psi(\zeta) \, d\zeta = 1$ , we immediately get  $\|\bar{\mathcal{G}}(x, t; \cdot, s)\|_{1,\Omega} \leq 1$ , which yields the first bound in Condition 2.1 with  $\kappa_0 = 1$ .

(ii) Next, we prove the second bound in Condition 2.1 in the linear case of  $f(x, t, z) = a(x)z + b(x, t)$  with  $\kappa_2 = 0$ . In this case, the differential operator in (2.2) does not involve  $s$ , so one can invoke [5, Corollary 5]. (In using this result, we imitate the proof of [6, Lemma 2.1].) In view of the above bound  $0 \leq \mathcal{G} \leq \bar{\mathcal{G}}$ , an application of [5, Corollary 5] with  $\beta = 2$ ,  $\gamma = 1$ ,  $c_1 = \frac{1}{4}$ ,  $c_2 = \frac{4}{9}c_1$ , and  $\alpha(t) = \frac{e^{-\gamma^2 t}}{(4\pi t)^{n/2}}$  yields  $|\partial_s \mathcal{G}(x, t; \xi, s)| \leq 18c_1c_2(t-s)^{-1} \alpha(\frac{1}{2}[t-s]) e^{-(c_2/c_1)|\zeta|^2}$ , where  $\zeta$  is chosen as in part (i) of this proof. Now an observation similar to (12.2) leads to the estimate  $\|\partial_s \mathcal{G}(x, t; \cdot, s)\|_{1,\Omega} \leq \kappa_1(t-s)^{-1} e^{-\frac{1}{2}\gamma^2(t-s)}$ , which immediately implies the second bound in Condition 2.1 with  $\kappa_2 = 0$ .

(iii) It remains to establish the second bound in Condition 2.1 in the general case of  $f(x, t, z)$  satisfying (1.2), which implies that for the coefficient  $a$  in (2.2a) one has  $\gamma^2 \leq a(\xi, s) \leq \bar{\gamma}^2$ . For any fixed  $(x, t) \in \Omega \times (0, T]$ , consider the Green’s function  $\hat{\mathcal{G}}(x, t; \xi, s) =: \hat{\Gamma}(\xi, s)$  associated with the operator  $\partial_t - \Delta + \gamma^2$  in the domain  $\Omega$  so

$\hat{\Gamma}(\xi, s)$  satisfies a version of (2.2) with  $a$  replaced by  $\gamma^2$ . Comparing this problem with the problem (2.2) for  $\Gamma$  and noting that  $\mathcal{L} = \mathcal{L}^* = \Delta$ , we find that for any fixed  $(x, t)$ , the function  $v(\xi, s) := \hat{\Gamma}(\xi, s) - \Gamma(\xi, s)$  solves the terminal-value problem

$$\begin{aligned} (12.3a) \quad & [-\partial_s - \Delta + \gamma^2] v(\xi, s) = F(\xi, s) \quad \text{for } (\xi, s) \in \Omega \times [0, t], \\ (12.3b) \quad & v(\xi, t) = 0 \quad \text{for } \xi \in \Omega, \\ (12.3c) \quad & v(\xi, s) = 0 \quad \text{for } (\xi, s) \in \partial\Omega \times [0, t], \end{aligned}$$

where  $F(\xi, s) := [a(\xi, s) - \gamma^2] \Gamma(\xi, s)$  so, using  $\Gamma \leq \bar{G}$  and (12.1),

$$(12.3d) \quad 0 \leq F(\xi, s) \leq (\bar{\gamma}^2 - \gamma^2) g(x - \xi, t - s).$$

Note that in part (ii) we have shown that  $\hat{\Gamma}$  satisfies the second bound in Condition 2.1 with  $\kappa_2 = 0$ . It remains to show that  $v$  satisfies the second bound in Condition 2.1 with  $\kappa_1 = 0$  and  $\kappa_2 = (\bar{\gamma}^2 - \gamma^2)\hat{\kappa}_2$ . This latter bound is immediately obtained by an application of Lemma 12.1 below to the terminal-value problem (12.3).  $\square$

The next lemma is applied to the terminal-value problem (12.3), but it is convenient to formulate it in the context of an initial-value problem.

LEMMA 12.1. *Let  $v$  satisfy  $[\partial_t - \Delta + \gamma^2]v = F$  in  $Q$  and vanish for  $t = 0$  and  $x \in \partial\Omega$ , where  $0 \leq F(x, t) \leq g(x - x_0, t)$  with  $g$  from (12.1) and some  $x_0 \in \Omega$ . Then  $\int_0^T \|\partial_t v(\cdot, t)\|_{1,\Omega} dt \leq \hat{\kappa}_2$ , where  $\hat{\kappa}_2$  is independent of  $|\Omega|$ , and  $\hat{\kappa}_2 = \hat{\kappa}_2(\gamma)$  if  $\gamma > 0$ , while  $\hat{\kappa}_2 = \hat{\kappa}_2(T)$  if  $\gamma = 0$ .*

*Proof.* Without loss of generality, assume that  $x_0 = 0 \in \Omega$  so  $F(x, t) \leq g(x, t)$ . Recall that  $\bar{\mathcal{M}}g = 0$  with  $\bar{\mathcal{M}} = \partial_t - \Delta + \gamma^2$ ; this implies that  $\bar{M}[tg] = g$ , so an application of the maximum principle yields

$$(12.4) \quad 0 \leq v(x, t) \leq tg(x, t).$$

(i) First we establish the desired estimate with  $\hat{\kappa}_2$  that depends on  $|\Omega|$ . Let  $w(x, t) := \varrho(t)v$  with the weight  $\varrho := t^{\frac{1}{3}} e^{\frac{1}{2}\gamma^2 t}$  so  $\varrho' = (\frac{1}{3}t^{-1} + \frac{1}{2}\gamma^2)\varrho$ . Note that

$$(12.5) \quad \begin{aligned} \|\partial_t v\|_{1,\Omega \times [0,T]} &\leq \|\varrho^{-1}\|_{2,\Omega \times [0,T]} \|\varrho \partial_t v\|_{2,\Omega \times [0,T]} \\ &\leq \hat{\kappa}_3 |\Omega|^{\frac{1}{2}} \left( \|\partial_t w\|_{2,\Omega \times [0,T]} + \|\varrho' v\|_{2,\Omega \times [0,T]} \right), \end{aligned}$$

where we used  $\varrho \partial_t v = \partial_t w - \varrho' v$  and

$$\|\varrho^{-1}\|_{2,\Omega \times [0,T]}^2 = |\Omega| \int_0^T t^{-\frac{2}{3}} e^{-\gamma^2 t} dt =: |\Omega| \hat{\kappa}_3^2$$

(so  $\hat{\kappa}_3^2 \leq 3T^{1/3}$  for  $\gamma \geq 0$ , and  $\int_0^\infty t^{-\frac{2}{3}} e^{-t} dt \approx 2.7$  implies  $\hat{\kappa}_3^2 \lesssim 2.7\gamma^{-2/3}$  for  $\gamma > 0$ ). To estimate  $\partial_t w$  in (12.5), we note that  $\bar{\mathcal{M}}w = \varrho F + \varrho' v \leq \varrho g + \varrho' v$  and so apply an a priori estimate [18, (6.6) of Chapter III]:

$$(12.6) \quad \|\partial_t w\|_{2,\Omega \times [0,T]} \leq \|\bar{\mathcal{M}}w\|_{2,\Omega \times [0,T]}.$$

(In fact, the cited estimate is given for a slightly different differential operator, but the argument also applies to  $\bar{\mathcal{M}}$ .) In view of  $\varrho' v \leq (\frac{1}{3} + \frac{1}{2}\gamma^2 t)\varrho g$  (which follows from (12.4)), one gets

$$(12.7) \quad \|\partial_t v\|_{1,\Omega \times [0,T]} \leq 2\hat{\kappa}_3 |\Omega|^{\frac{1}{2}} \|\hat{\varrho} g\|_{2,\Omega \times [0,T]}, \quad \text{where } \hat{\varrho} := \left(\frac{4}{3} + \frac{1}{2}\gamma^2 t\right)\varrho.$$

Finally, a calculation using  $\zeta := \frac{x}{\sqrt{2t}}$  and  $\psi(\zeta)$  from (12.2) yields

$$\|\hat{\varrho}g\|_{2,\Omega \times [0,T]}^2 \leq \int_0^T \frac{\hat{\varrho}^2(t) e^{-2\gamma^2 t}}{(8\pi t)^{n/2}} \int_{\mathbb{R}^n} \psi(\zeta) \, d\zeta \, dt = \int_0^T \frac{(\frac{4}{3} + \frac{1}{2}\gamma^2 t)^2 t^{2/3} e^{-\gamma^2 t}}{(8\pi t)^{n/2}} \, dt =: \hat{\kappa}_4^2.$$

(This integral is convergent as  $\frac{n}{2} - \frac{2}{3} < 1$  for  $n \leq 3$ .) Combining this with (12.7), we arrive at the desired bound with  $\hat{\kappa}_2 := 2 \hat{\kappa}_3 \hat{\kappa}_4 |\Omega|^{\frac{1}{2}}$ .

(ii) Now we shall show the desired result with  $\hat{\kappa}_2$  independent of  $|\Omega|$  (which requires a more subtle estimation). Divide  $\mathbb{R}^n$  into the nonoverlapping subdomains  $\Omega_0 := \{|x| < 2\}$  and  $\Omega_j := \{2^j < |x| < 2^{j+1}\}$  for  $j = 1, \dots$ ; furthermore let  $\Omega'_0 := \Omega_0$  and  $\Omega'_j := \{2^{j-1} < |x| < 2^{j+2}\} \supset \Omega_j$ . Note that

$$(12.8) \quad |\Omega_j|^{\frac{1}{2}} \leq c_n 2^{\frac{1}{2}nj}.$$

Now we partially imitate the proof in part (i). First, note that one has the bound (12.5) with  $\Omega$  replaced by  $\Omega \cap \Omega_j$  for  $j = 0, 1, \dots$ . Hence for  $j = 0$ , using the results of part (i), one immediately gets

$$(12.9) \quad \|\partial_t v\|_{1,(\Omega \cap \Omega_0) \times [0,T]} \leq 2 \hat{\kappa}_3 \hat{\kappa}_4 |\Omega_0|^{\frac{1}{2}}.$$

(Compare with  $\hat{\kappa}_2$  from part (i).)

For  $j \geq 1$ , we combine the local version of (12.5) with a local version of the global estimate (12.6) from

$$\|\partial_t w\|_{2,(\Omega \cap \Omega_j) \times [0,T]} \leq \bar{C} \left\{ \|\bar{\mathcal{M}}w\|_{2,(\Omega \cap \Omega'_j) \times [0,T]} + \|w\|_{2,(\Omega \cap \Omega'_j) \times [0,T]} \right\}$$

with the constant  $\bar{C}$  independent of  $\Omega$  and  $T$ . (This estimate is obtained similarly to [18, (6.6), (6.11) of Chapter III].) Here  $\bar{\mathcal{M}}w$  is estimated as in part (i), while  $w = \varrho v \leq t\varrho g$  by (12.4). This yields a local version of (12.7):

$$(12.10) \quad \|\partial_t v\|_{1,(\Omega \cap \Omega_j) \times [0,T]} \leq 2 \hat{\kappa}_3 |\Omega_j|^{\frac{1}{2}} \bar{C} \|(\hat{\varrho} + t\varrho)g\|_{2,\Omega'_j \times [0,T]} \quad \text{for } j \geq 1.$$

Next, we use  $\zeta := \frac{x}{\sqrt{2t}}$  and  $\psi(\zeta)$  from (12.2) and also the observation that as  $j \geq 1$  so  $(\exp(-\frac{|x|^2}{4t}))^2 \leq e^{-\frac{4j^2-2}{t}} e^{-|\zeta|^2} \leq c'_n (\frac{t}{4j})^n e^{-|\zeta|^2}$ . So for  $j \geq 1$  a calculation shows that

$$\|(\hat{\varrho} + t\varrho)g\|_{2,\Omega'_j \times [0,T]}^2 \leq c'_n 4^{-jn} \int_0^T \frac{(\hat{\varrho} + t\varrho)^2 e^{-2\gamma^2 t} t^n}{(4\pi t)^{n/2}} \int_{\mathbb{R}^n} \psi(\zeta) \, d\zeta \, dt = c''_n 4^{-jn}.$$

Combining this with (12.10) and then with (12.9) and (12.8), we arrive at

$$\|\partial_t v\|_{1,(\Omega \cap \Omega_j) \times [0,T]} \leq 2 \hat{\kappa}_3 c_n \begin{cases} \hat{\kappa}_4 & \text{for } j = 0, \\ \sqrt{c''_n} 2^{-\frac{1}{2}nj} & \text{for } j \geq 1. \end{cases}$$

This immediately yields the desired bound with  $\hat{\kappa}_2 := 2 \hat{\kappa}_3 c_n [\hat{\kappa}_4 + \sqrt{c''_n} (2^{\frac{1}{2}n} - 1)^{-1}]$  independent of  $|\Omega|$ .  $\square$

**Appendix A. Elliptic estimators.** We now cite error estimators of type (7.3) for particular cases of the elliptic model problem (7.4) and its discretizations (7.5).

**A.1. Elliptic model problem: Regular regime.** We first consider the steady-state version (7.4) of our model problem (1.3) in the regular regime of  $\varepsilon := 1$ .

Let  $v$  solve the problem (7.4) with  $\varepsilon = 1$ ,  $\gamma \geq 0$ , posed in a bounded polyhedral domain  $\Omega \subset \mathbb{R}^n$ ,  $n = 2, 3$ , and  $v_h$  solve the discrete problem (7.5) with  $V_h$  and  $\langle \cdot, \cdot \rangle_h$  defined as follows. Given a conforming and shape-regular triangulation  $\mathcal{T}_h$  of  $\Omega$  made of elements  $T$ , we let  $V_h$  be the space of continuous piecewise polynomial finite element functions of degree  $l \geq 1$ , and  $\tilde{V}_h := V_h \cap H_0^1(\Omega)$ . We employ  $\langle \varphi, w \rangle_h := \sum_{T \in \mathcal{T}_h} Q_T(\varphi w)$ , where  $Q_T$  is a quadrature formula for the integral over  $T$  with positive weights, and quadrature points contained in  $T$ , such that  $Q_T$  is exact for the polynomials of degree  $q$  with  $q \geq \max\{2l - 2, 1\}$ .

In [24, Theorem 4.2], an a posteriori error estimate of type (7.3) is given with  $\eta = \eta_0$  defined by

$$(A.1) \quad \eta_0(V_h, v_h, g(\cdot, v_h)) := \left[ c_0 \max_{T \in \mathcal{T}_h} \left\{ h_T^2 \|(\Delta v_h - g(\cdot, v_h))\|_{\infty, T} + h_T \|[\![\partial_n v_h]\!] \|_{\infty, \partial T \setminus \partial \Omega} \right\} \right. \\ \left. + c_1 \|\nu_{n/2, T}^q\|_{l_{n/2}} + c_2 \|h_T \nu_{n, T}^{q-1}\|_{l_n} \right] \times |\ln h_{\min}|^2,$$

where  $h_{\min}$  is the smallest mesh size,  $h_T$  is the diameter of  $T$ ,  $[\![\partial_n v_h]\!]$  is the jump of the normal derivatives across an interelement side,  $\|\cdot\|_{l_p}$  is the  $l_p$  norm, and the quantity

$$\nu_{n', T}^{q'} := |T|^{1/n'} \|g(\cdot, v_h) - I_{h, q'}[g(\cdot, v_h)]\|_{\infty, T}$$

is defined using the Lagrange interpolation operator  $I_{h, q'}$  onto the space of piecewise polynomials of degree  $\leq q'$ .<sup>4</sup>

**A.2. Elliptic model problem: Singularly perturbed regime in one dimension.**<sup>5</sup> We now consider the steady-state version (7.4) of our model problem (1.3) in the singularly perturbed regime of  $\varepsilon \ll 1$ .

Let  $v$  solve the problem (7.4) with  $\varepsilon \in (0, 1]$  and  $\gamma > 0$ , posed in the domain  $\Omega := (0, 1)$ , and  $v_h$  solve the discrete problem (7.5) using the space  $V_h$  of continuous piecewise-linear finite element functions on an arbitrary nonuniform mesh  $\{x_i\}_{i=1}^N$  with  $0 = x_0 < x_1 < \dots < x_N = 1$  and  $h_i := x_i - x_{i-1}$ . Note that here we make absolutely no mesh regularity assumptions (as solutions of our problem typically exhibit sharp layers, so a suitable mesh is expected to be highly nonuniform; see, e.g., [21]).

Consider two choices of  $\langle \cdot, \cdot \rangle_h$ , which are defined using the standard piecewise-linear nodal interpolation operator  $I_h$ :

$$(A.2a) \quad \langle \varphi, w_h \rangle_h := \langle I_h \varphi, w_h \rangle \quad (\text{quadrature}),$$

$$(A.2b) \quad \langle \varphi, w_h \rangle_h := \langle I_h[\varphi w_h], 1 \rangle \quad (\text{lumped-mass quadrature}).$$

*Remark A.1.* To illustrate Remarks 7.3 and 7.4, note that the described two discretizations using either (A.2a) or (A.2b) are of type (7.2). In particular, for (A.2a), we get  $\mathcal{L}_h := -\varepsilon^2 [\partial_x^2]_h$  and  $\mathcal{P}_h := I_h$ . Here the operator  $[\partial_x^2]_h : H_0^1(\Omega) \rightarrow$

<sup>4</sup>It is noted in [6] that if the domain  $\Omega$  has cracks, it is not entirely clear whether (A.1) of [24] still holds. We refer the reader to [6, Remark 2.4] for a further discussion of (A.1) in this case and related literature.

<sup>5</sup>Similar elliptic estimators for two- and three-dimensional steady-state versions of (1.3) in the singularly perturbed regime  $\varepsilon \ll 1$  are given in [15, 4].

$\mathring{V}_h + \varepsilon^{-2} I_h[g(\cdot, 0)]$  is defined by  $\langle -[\partial_x^2]_h \varphi, w_h \rangle = \langle \varphi', w'_h \rangle$  for all  $\varphi \in H_0^1(\Omega)$ ,  $w_h \in \mathring{V}_h$ . Consequently, the discrete problem using (A.2a) may be represented as

$$(A.3a) \quad -\varepsilon^2 [\partial_x^2]_h v_h + I_h[g(\cdot, v_h)] = 0.$$

By contrast, (A.2b) can be rewritten as a difference scheme:  $-\varepsilon^2 \delta_x^2 v_{h,i} + g(x_i, v_{h,i}) = 0$ , for  $i = 1, \dots, N - 1$ , where  $\delta_x^2 v_{h,i} := \frac{2}{h_i + h_{i+1}} [\frac{1}{h_{i+1}}(v_{h,i+1} - v_{h,i}) - \frac{1}{h_i}(v_{h,i} - v_{h,i-1})]$  is the standard finite-difference operator. Letting  $\delta_x^2 v_{h,i} := \varepsilon^{-2} g(x_i, v_{h,i})$  for  $i = 0, N$  and applying the linear interpolation  $I_h$  to  $\{\delta_x^2 v_{h,i}\}_{i=0}^N$ , we can represent the discrete problem using (A.2b) as

$$(A.3b) \quad -\varepsilon^2 I_h[\delta_x^2 v_h] + I_h[g(\cdot, v_h)] = 0,$$

where the values  $\delta_x^2 v_{h,i}$  are easily explicitly computable.

We cite a posteriori estimates [14, 20, 21] of type (7.3) with  $\eta := \eta_\varepsilon(V_h, g(\cdot, v_h))$ , where  $\eta_\varepsilon = \eta_{\varepsilon, (A.2a)}$  for (A.2a) and  $\eta_\varepsilon = \eta_{\varepsilon, (A.2b)}$  for (A.2b) are given by

$$(A.4a) \quad \eta_{\varepsilon, (A.2a)}(V_h, g_*) := \max_{i=1, \dots, N} \left\{ \frac{h_i^2}{4\varepsilon^2} \|I_h g_*\|_{\infty, (x_{i-1}, x_i)} \right\} + \gamma^{-2} \|g_* - I_h g_*\|_{\infty, (0,1)},$$

$$(A.4b) \quad \eta_{\varepsilon, (A.2b)}(V_h, g_*) := \eta_{\varepsilon, (A.2a)} + \max_{i=1, \dots, N} \left\{ \frac{h_i^2}{6\gamma\varepsilon} \|\partial_x(I_h g_*)\|_{\infty, (x_{i-1}, x_i)} \right\},$$

where  $g_* := g(\cdot, v_h)$ .

*Remark A.2.* The error estimators (A.4a) and (A.4b) are *robust* although they involve negative powers of the small parameter  $\varepsilon$ . Indeed, an inspection of representations (A.3a) and (A.3b) for the two considered numerical methods shows that  $\varepsilon^{-2} h_i^2 |I_h g_*| = \varepsilon^{-2} h_i^2 |I_h [g(\cdot, v_h)]|$  becomes  $h_i^2 |[\partial_x^2]_h v_h|$  or  $h_i^2 |\delta_x^2 v_h|$ , so it approximates  $h_i^2 |\partial_x^2 v|$ , where  $v$  is the exact solution of our equation  $-\varepsilon^2 \partial_x^2 v + g(\cdot, v) = 0$ . Similarly, the term  $\varepsilon^{-1} h_i^2 |\partial_x(I_h g_*)|$  approximates  $\varepsilon |\partial_x^3 v|$ , which has similar magnitude to  $h_i^2 |\partial_x^2 v|$  in the layer regions.

By contrast, if  $\langle \cdot, \cdot \rangle_h := \langle \cdot, \cdot \rangle$  (i.e. no quadrature is used), then one can obtain a simpler-looking error estimate of type (7.3) with  $\eta := \max_{i=1, \dots, N} \left\{ \frac{h_i^2}{4\varepsilon^2} \|g_*\|_{\infty, (x_{i-1}, x_i)} \right\}$ .

However, this estimate is not robust. To see this, split  $g_* = P_h g_* + (g_* - P_h g_*)$  using the standard  $L_2$  projection  $P_h$ . Then, instead of (A.3a), we have the representation  $-\varepsilon^2 [\partial_x^2]_h v_h + P_h [g_*] = 0$  for our numerical method. The component  $\varepsilon^{-2} h_i^2 |P_h g_*|$  approximates  $h_i^2 |\partial_x^2 v|$  so it yields a robust part of the estimator. But the other component  $\varepsilon^{-2} h_i^2 |g_* - P_h g_*|$  may be as large as  $\mathcal{O}(\varepsilon^{-2} h_i^4)$ , which may become quite large if  $\varepsilon$  is small compared to the local mesh size. For this numerical method one can, in fact, obtain a robust error estimator, which is almost identical to (A.4a), only  $I_h$  in  $\eta_\varepsilon$  should be replaced by  $P_h$  (but this latter estimator is less practical, as it requires the  $L_2$  projection  $P_h g_*$  to be explicitly computed).

REFERENCES

[1] G. AKRIVIS, C. 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] S. BARTELS AND R. MÜLLER, *Quasi-optimal and robust a posteriori error estimates in  $L^\infty(L^2)$  for the approximation of Allen-Cahn equations past singularities*, Math. Comp., 80 (2011), pp. 761–780.  
 [3] M. BOMAN, *On A Posteriori Error Analysis in the Maximum Norm*, Ph.D. thesis, Chalmers University of Technology and Göthenburg University, Göthenburg, Sweden, 2000.

- [4] N. M. CHADHA AND N. KOPTEVA, *Maximum norm a posteriori error estimate for a 3d singularly perturbed semilinear reaction-diffusion problem*, Adv. Comput. Math., 35 (2011), pp. 33–55.
- [5] E. B. DAVIES, *Non-Gaussian aspects of heat kernel behaviour*, J. London Math. Soc. (2), 55 (1997), pp. 105–125.
- [6] A. DEMLOW, O. LAKKIS, AND C. MAKRIDAKIS, *A posteriori error estimates in the maximum norm for parabolic problems*, SIAM J. Numer. Anal., 47 (2009), pp. 2157–2176.
- [7] A. DEMLOW AND C. MAKRIDAKIS, *Sharply local pointwise a posteriori error estimates for parabolic problems*, Math. Comp., 79 (2010), pp. 1233–1262.
- [8] T. DUPONT, *Mesh modification for evolution equations*, Math. Comp., 39 (1982), pp. 85–107.
- [9] 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.
- [10] 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.
- [11] L. C. EVANS, *Partial Differential Equations*, AMS, Providence, RI, 1998.
- [12] A. FRIEDMAN, *Partial Differential Equations of Parabolic Type*, Prentice-Hall, Englewood Cliffs, NJ, 1964.
- [13] D. H. GRIFFEL, *Applied Functional Analysis*, Dover Publications, Mineola, NY, 2002.
- [14] N. KOPTEVA, *Maximum norm a posteriori error estimates for a 1D singularly perturbed semilinear reaction-diffusion problem*, IMA J. Numer. Anal., 27 (2007), pp. 576–592.
- [15] N. KOPTEVA, *Maximum norm a posteriori error estimate for a 2d singularly perturbed reaction-diffusion problem*, SIAM J. Numer. Anal., 46 (2008), pp. 1602–1618.
- [16] 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.
- [17] N. KOPTEVA AND T. LINSS, *Numerical Study of Maximum Norm a Posteriori Error Estimates for Singularly Perturbed Parabolic Problems*, Lecture Notes in Comput. Sci., Springer, Berlin, to appear.
- [18] O. A. LADYZHENSKAYA, V. A. SOLONNIKOV, AND N. N. URAL'TSEVA, *Linear and Quasi-Linear Equations of Parabolic Type*, AMS, Providence, RI, 1968.
- [19] O. LAKKIS AND C. MAKRIDAKIS, *Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems*, Math. Comp., 75 (2006), pp. 1627–1658.
- [20] T. LINSS, *Maximum-norm error analysis of a non-monotone FEM for a singularly perturbed reaction-diffusion problem*, BIT Numer. Math., 47 (2007), pp. 379–391.
- [21] T. LINSS, *Layer-Adapted Meshes for Reaction-Convection-Diffusion Problems*, Lecture Notes in Math. 1985, Springer, Berlin, 2010.
- [22] C. MAKRIDAKIS AND R. H. NOCHETTO, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, SIAM J. Numer. Anal., 41 (2003), pp. 1585–1594.
- [23] C. MAKRIDAKIS AND R. H. NOCHETTO, *A posteriori error analysis for higher order dissipative methods for evolution problems*, Numer. Math., 104 (2006), pp. 489–514.
- [24] R. H. NOCHETTO, A. SCHMIDT, K. G. SIEBERT, AND A. VEESER, *Pointwise a posteriori error estimates for monotone semi-linear equations*, Numer. Math., 104 (2006), pp. 515–538.
- [25] H.-G. ROOS, M. STYNES AND L. TÖBISKA, *Robust Numerical Methods for Singularly Perturbed Differential Equations*, Springer, Berlin, 2008.
- [26] A. N. TIKHONOV AND A. A. SAMARSKII, *Equations of Mathematical Physics*, Dover Publications, New York, 1990.
- [27] V. THOMÉE, *Galerkin Finite Element Methods for Parabolic Problems*, Springer, Berlin, 2006.