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

NATALIA KOPTEVA† AND TORSTEN LINSS‡

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.

1. Introduction.

Consider a semilinear parabolic equation in the form

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

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

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

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

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

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 analyze the third-order discontinuous Galerkin method dG(1) with Radau quadrature, and generalize our results for higher-order discontinuous Galerkin methods.

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

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

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

$$\text{(i)} \quad \varepsilon = 1, \;\; \gamma \ge 0; \qquad\qquad \text{(ii)} \quad \varepsilon \ll 1, \;\; \gamma > 0.$$

Note that regime (ii) yields a singularly perturbed reaction-diffusion equation, whose solutions may exhibit sharp layer phenomena. So it is important in this regime that

*This publication has emanated from research conducted with the financial support of Science Foundation Ireland under the Research Frontiers Programme 2008; Grant 08/RFP/MTH1536.

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

‡Fakultät für Mathematik und Informatik, FernUniversität in Hagen, Lützowstr. 125, 58095 Hagen, Germany (torsten.linss@fernuni-hagen.de).

a posteriori error estimates are robust in the sense that any dependence on the small perturbation parameter ε should be shown explicitly [19, 23].

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 τ_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 [20, 17, 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 $\partial_t - \triangle$ 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 are given either for regular linear problems [8, 3, 6, 7], or are not robust in the sense that they involve negative powers of ε [3]. For a more detailed comparison of our results with various earlier a posteriori error estimates, we refer the reader to Remarks 5.4, 9.6, 9.11 and 11.8 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. Sections 3–6 are devoted to semidiscrete methods, while in Sections 8–11 we deal with fully discrete methods as summarized in the table:

	Summary of results	backward Euler	Crank-Nicolson	dG(1)	dG(r)
Semidiscretizations	§3	§4	§5	§6	§6.3
Full discretizations	§8	§9	§10	§11	Remark 11.6

Note that elliptic reconstructions for all discretizations are defined in Section 8.2. Furthermore, in Section 7 we cite some elliptic a posteriori error estimates, used in the analysis of fully discrete methods, while the final Section 12 gives a proof of certain Green's function bounds deferred from Section 2.

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 ε 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 $L_p(\Omega)$, $1 \le p \le \infty$, with the norm $\|\cdot\|_{p,\Omega}$, while $\langle \phi, \psi \rangle = \int_\Omega \phi(x)\psi(x)\,\mathrm{d}x$ denotes the inner product in $L_2(\Omega)$.

Distributions and left-continuity convention. Certain functions will be understood as distributions [12], 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 [11, Chap. 1 and §7 of Chap. 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 $a(x,t) := \int_0^1 \partial_z f(x,t,w+z[v-w])\,\mathrm{d}z$. Hence, the difference $v-w$ is represented as

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

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

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

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

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

Here $\delta(\cdot)$ is the Dirac δ-distribution in $\mathbb{R}^n$ [12], 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 κ_0, $\kappa_1 > 0$ and $\kappa_2 \geq 0$ such that the Green's function $\mathcal{G}$ of* (2.2), (1.2) *satisfies*

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

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

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 ε-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 $\gamma = 0$.*

Proof. We defer the proof to Section 12. □

REMARK 2.3. *The constants κ_0 and κ_1 given by Lemma 2.2 are reasonably sharp. E.g., 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 yields $\|\mathcal{G}(x,t;\cdot,s)\|_{1,\Omega} = 1$ and $\|\partial_s\mathcal{G}(x,t;\cdot,s)\|_{1,\Omega} \leq \big(\sqrt{\tfrac{2}{\pi\,\mathrm{e}}}(t-s)^{-1} + \gamma^2\big)\,\mathrm{e}^{-\gamma^2(t-s)}$ so Condition 2.1 is satisfied with $\kappa_0 = 1$ (as in Lemma 2.2), $\kappa_1 = \sqrt{\tfrac{2}{\pi\,\mathrm{e}}} \approx 0.48$, $\kappa_2 = 1$, while Lemma 2.2 gives $\kappa_1 = \frac{3}{2^{3/2}} \approx 1.06$.*

The above 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

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

Lemma 2.4. *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 $v(\cdot,t)$, $w(\cdot,t) \in H^1_0(\Omega)\cap C(\bar\Omega)$ for $t\in[0,T]$, and*

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

where the function μ 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 $1\le m\le M$, and $\|\vartheta(\cdot,s)\|_{\infty,\Omega}$ is integrable w.r.t. s in $(0,t_m)$ (possibly, in the sense of distributions). Then

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

where $\ell_m=\ell_m(\gamma):=\int_{\tau_m}^{t_m} s^{-1}\mathrm{e}^{-\frac12\gamma^2 s}\,\mathrm{d}s \le \ln(t_m/\tau_m)$.

Remark 2.5. *The term $\partial_t\mu$ in the right-hand side of (2.4) is understood in the sense of distributions. A typical μ 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 to avoid ambiguity when integrating by parts. It may help the reader to consider an equivalent interpretation of such evaluations. For some small positive λ, one can replace t_j^+ by $t_j+\lambda$, and μ by μ_λ 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 $\lambda\to 0^+$.*

Similarly, in all calculations involving Γ, one can initially replace it by a regular function Γ_λ obtained using a regular approximation δ_λ of δ in (2.2b), and then let $\lambda\to0^+$. With regard to the regularity of Γ, 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 2.6. *One can easily check that if $\gamma=0$, then $\ell_m=\ln(t_m/\tau_m)$. Otherwise, if $\gamma>0$, one has $\ell_m(\gamma)=E_1(\frac12\gamma^2\tau_m)-E_1(\frac12\gamma^2 t_m)$, where $E_1(t)=\int_t^\infty s^{-1}\mathrm{e}^{-s}\,\mathrm{d}s$; so $\ell_m(\gamma)\le|\ln(\frac12\gamma^2\tau_m)|$ provided that $\frac12\gamma^2\tau_m\le 0.67$ (this is easily checked by finding the only root ≈ 0.67 of the equation $E_1(s)=|\ln s|$ on $(0,1)$). Note also that $\ell_1=0$ for any $\gamma\ge0$.*

Proof of Lemma 2.4. 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

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

Here, in view of (2.4), the integral on the right-hand side involves μ and ϑ, so can be represented as a sum $J_\mu+J_\vartheta$ of the corresponding integrals, which we consider separately. We use the notation $\int^{b^+}:=\lim_{\beta\to0^+}\int^{b+\beta}$ and so split J_μ as

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

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

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

Consequently, we arrive at

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

where the last term represents J_ϑ. Finally, Condition 2.1 implies that

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

so we get the desired result. ∎

The following version of Lemma 2.4 involves certain approximations Γ_h^j of $\Gamma(\cdot, t_j)$.

LEMMA 2.4*. *Under conditions of Lemma 2.4, 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* $j = 1, \ldots, m-1$, *and* $\sum_{j=1}^{m-1} \tau_j\, \|\mathcal{H}_j^{-2}\{\Gamma(\cdot,t_j) - \Gamma_h^j\}\|_{1,\Omega} \le \kappa_3\,\ell(\tau,t)$ *for some positive weight functions* $\{\mathcal{H}_j\}$ *and some constant* κ_3, *then the statement of Lemma 2.4 remains valid, only with an additional term* $\kappa_3\,\ell(\tau,t)\,\max_{j=1,\ldots,m-1}\big\{\tau_j^{-1}\|\mathcal{H}_j^2\,\vartheta^j\|_{\infty,\Omega}\big\}$ *in the final line of (2.5).*

Proof. Imitate the proof of Lemma 2.4, and note that now we have (2.6) with an additional term $\sum_{j=1}^{m-1}\langle \vartheta^j,\, \Gamma(\cdot,t_j)\rangle = \sum_{j=1}^{m-1}\langle \vartheta^j,\, \Gamma(\cdot,t_j) - \Gamma_h^j\rangle$. ∎

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

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

The quantities that appear in this estimate are specified by Theorems 4.1, 5.1 and 6.4 below, and can be summarized as follows:

	p	χ^{j+1}	ϑ	C_1	C_2
backward Euler	1	$U^{j+1} - U^j$	$\tilde\psi - \tilde\psi^j$ on $(t_{j-1}, t_j]$	1	2
Crank-Nicolson	2	$\tau_{j+1}(\psi^{j+1} - \psi^j)$	$\tilde\psi - I_{1,t}\tilde\psi$	$\frac{1}{8}$	$\frac{5}{8}$
dG(1)-Radau	3	$3\tau_{j+1}(2\psi^j - 3\psi^{j+1/3} + \psi^{j+1})$	$\tilde\psi - I_{2,t}\tilde\psi$	$\frac{2}{81}$	$\frac{1}{18}$

Here for the evaluation of χ^{j+1} and ϑ we use

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

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.

To be more precise, on $(t_j, t_{j+1}]$ we use the constant interpolation for the backward Euler method with $\tilde U = U^{j+1}$, the linear interpolation $I_{1,t}$ for the Crank-Nicolson method with the interpolation nodes $\{t_j, t_{j+1}\}$, and the quadratic interpolation $I_{2,t}$ for the dG(1) method with the interpolation nodes $\{t_j, t_{j+1/3}, t_{j+1}\}$.

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 χ^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 (pth-order estimates). *Remark 3.1 and the definitions of ϑ 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

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

where

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

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

Theorem 4.1. *Let u solve the 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 ϑ defined by*

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

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

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

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

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

(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

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

Here we also used the observation that, by (4.4), the regular function ϑ 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 2.4 with $\mu := \tilde U - I_{1,t}U$ and ϑ 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

$$\mu = \tfrac{t_j - t}{\tau_j}\,(U^j - U^{j-1}) = \tfrac{t_j - t}{\tau_j}\,\chi^j \qquad \Longrightarrow \qquad |\mu| \le \left|\chi^j\right|, \quad \tau_j\,|\partial_t \mu| = \left|\chi^j\right|.$$

This completes the proof. □

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* κ_0*,* κ_1*,* κ_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

$$\delta_t U^j + \tfrac12\big(\mathcal{L}^{j-1}U^{j-1} + \mathcal{L}^j U^j\big) + \tfrac12(f^{j-1} + f^j) = 0 \quad \text{in } \Omega, \quad j = 1, \ldots, M, \tag{5.1a}$$

where we again let

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

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

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

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 U^j solve the corresponding semidiscrete problem (5.1). Then for $m = 1, \ldots, M$, one has* (3.1) *with* $\chi^j = \tau_j\left(\psi^j - \psi^{j-1}\right)$ *using* $\psi^j = \mathcal{L}^j U^j + f^j$*,* $C_1 = \frac18$*,* $C_2 = \frac58$*, and ϑ defined by*

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

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. Consider $t \in [t_{j-1}, t_j]$. In view of (5.2), for any function $w = w(t)$ with the notation $w^j := w(t_j)$ and $\delta_t w^j := (w^j - w^{j-1})/\tau_j$, one has $I_{1,t}w(t) = \frac{1}{2}(w^{j-1} + w^j) + (t - t_{j-1/2})\,\delta_t w^j$. So using (5.3) combined with $\tilde\psi(\cdot, t_j) = \psi^j$, and then this property, for $t \in [t_{j-1}, t_j]$, one easily gets

$$\begin{aligned}\tilde\psi - \tfrac{1}{2}\big(\psi^{j-1} + \psi^j\big) &= \big\{I_{1,t}\tilde\psi - \tfrac{1}{2}\big(\psi^{j-1} + \psi^j\big)\big\} + \vartheta\\ &= (t - t_{j-1/2})\,\delta_t\psi^j + \vartheta\\ &= (t - t_{j-1/2})\,\tau_j^{-2}\chi^j + \vartheta.\end{aligned}$$

Note also that $\tilde U(\cdot,t) = I_{1,t}U(\cdot,t)$ implies $\partial_t\tilde U = \delta_t U^j = \frac{1}{2}(\psi^{j-1} + \psi^j)$ for $t \in (t_{j-1}, t_j]$ (where we also invoked (5.1a)). Combining these two observations, one deduces that

$$\partial_t\tilde U + \tilde\psi = (t - t_{j-1/2})\,\tau_j^{-2}\chi^j + \vartheta \quad \text{for } t \in (t_{j-1}, t_j] \tag{5.4}$$

(here the left-hand side is a regular function). As $\mathcal{M}\tilde U = \partial_t\tilde U + \tilde\psi$ and $\mathcal{M}u = 0$, so (5.4) yields

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

where $\mu = \mu(x,t)$ is a continuous function defined by

$$\mu(\cdot,t) := -\tfrac{1}{2}(t_j - t)(t - t_{j-1})\cdot\tau_j^{-2}\chi^j \qquad \text{for } t \in [t_{j-1}, t_j]. \tag{5.6}$$

This is easily checked by using the relation $\frac{d}{dt}[-\frac{1}{2}(t_j - t)(t - t_{j-1})] = t - t_{j-1/2}$ to evaluate $\partial_t\mu$.

Now the desired bound for $U^m - u(\cdot,t_m) = [\tilde U - u](\cdot,t_m)$ is obtained by an application of Lemma 2.4 to the equation (5.5) with μ defined by (5.6), and ϑ 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

$$|\mu| \le \tfrac{1}{8}\,|\chi^j| \quad \text{and} \quad \tau_j\,|\partial_t\mu| \le \tfrac{1}{2}\,|\chi^j|,$$

with $\frac{1}{8} + \frac{1}{2} = \frac{5}{8}$. This completes the proof. ∎

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 κ_0, κ_1, κ_2 from Lemma 2.2.*

Remark 5.3. *If the term $\frac{1}{2}(f^{j-1} + f^j)$ in the Crank-Nicolson discretization (5.1) is replaced by $f_h^j := \tau_j^{-1}\int_{t_{j-1}}^{t_j} f(\cdot,t,\tilde U)\,dt$, then the proof of Theorem 5.1 remains applicable with the only modification that the right-hand side in (5.5) involves, for $t \in (t_{j-1}, t_j]$, an additional term $-\bar\vartheta_f := \frac{1}{2}(f^{j-1} + f^j) - f_h^j$. Consequently, the statement of Theorem 5.1 remains valid with ϑ in the final line of (3.1) replaced by $\vartheta - \bar\vartheta_f$, where, as one can easily deduce, $\bar\vartheta_f = \tau_j^{-1}\int_{t_{j-1}}^{t_j}\vartheta_f\,dt$ is the average value on $[t_{j-1}, t_j]$ of the component $\vartheta_f := f(\cdot,t,\tilde U) - I_{1,t}f(\cdot,t,\tilde U)$ of ϑ.*

Remark 5.4. *The a posteriori error estimates given by Theorem 5.1 and Remark 5.3 resemble (but are not identical with) 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(1) (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 third-order discontinuous Galerkin method dG(1) (described, e.g., in [9, 25]) as follows.

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 an auxiliary approximate solution $y^j \in H_0^1(\Omega)\cap C(\bar\Omega)$ associated with the time level t_j^+ and an approximate solution $U^{j+1} \in H_0^1(\Omega)\cap C(\bar\Omega)$ associated with the time level t_{j+1} to satisfy

$$\big\langle y^j - U^j, v\big\rangle + \int_{t_j^+}^{t_{j+1}} \big\langle \partial_t U + \psi,\, v\phi_j + w\phi_{j+1}\big\rangle\, \mathrm{d}t = 0 \qquad \forall\, v, w \in H_0^1(\Omega), \tag{6.1}$$

where $\{\phi_j\}_{j=0}^M$ are the standard piecewise-linear hat-functions in time, with each ϕ_j having support $[t_{j-1}, t_{j+1}]$, so $\phi_j(t) = \frac{t_{j+1}-t}{\tau_{j+1}}$ and $\phi_{j+1}(t) = \frac{t-t_j}{\tau_{j+1}}$ for $t \in [t_j, t_{j+1}]$,

$$U := y^j\,\phi_j + U^{j+1}\phi_{j+1}, \qquad \psi := \mathcal{L}(t)\,U + f(\cdot, t, U) \qquad \text{for } t \in (t_j, t_{j+1}]. \tag{6.2}$$

As $v, w \in H_0^1(\Omega)$ are arbitrary, the above relation (6.1) is equivalent to the system for y^j, $U^{j+1} \in H_0^1(\Omega)\cap C(\bar\Omega)$:

$$y^j - U^j + \int_{t_j^+}^{t_{j+1}} \big(\partial_t U + \psi\big)\,\phi_j\ \mathrm{d}t = 0, \qquad \int_{t_j^+}^{t_{j+1}} \big(\partial_t U + \psi\big)\,\phi_{j+1}\ \mathrm{d}t = 0.$$

Next, we take the sum and difference of these two equations. As $\phi_j + \phi_{j+1} = 1$ and $\int_{t_j^+}^{t_{j+1}} \partial_t U\ \mathrm{d}t = U^{j+1} - y^j$, while $\int_{t_j^+}^{t_{j+1}} \partial_t U\,(\phi_j - \phi_{j+1})\ \mathrm{d}t = 0$, so

$$U^{j+1} - U^j + \int_{t_j^+}^{t_{j+1}} \psi\ \mathrm{d}t = 0, \tag{6.3a}$$

$$y^j - U^j + \int_{t_j^+}^{t_{j+1}} \psi\,(\phi_j - \phi_{j+1})\ \mathrm{d}t = 0. \tag{6.3b}$$

Note that the system (6.3) is equivalent to (6.1), so the semidiscrete dG(1) method can be equivalently defined by (6.3) combined with (6.2).

Furthermore, applying the Radau quadrature to both integrals in (6.3), one gets the semidiscrete dG(1) method with quadrature:

$$U^{j+1} - U^j + \tfrac14\,\tau_{j+1}\,\big(3\psi^{j+1/3} + \psi^{j+1}\big) = 0, \tag{6.4a}$$

$$y^j - U^j + \tfrac14\,\tau_{j+1}\,\big(\psi^{j+1/3} - \psi^{j+1}\big) = 0, \tag{6.4b}$$

where

$$\psi^{j+1/3} := \psi(\cdot, t_{j+1/3}), \qquad \psi^{j+1} := \psi(\cdot, t_{j+1}), \tag{6.4c}$$

are computed using $U(\cdot, t_{j+1/3}) = \frac23 y^j + \frac13 U^{j+1}$ and $U(\cdot, t_{j+1}) = U^{j+1}$, and also $\mathcal{L}(t_{j+1/3})$ and $\mathcal{L}(t_{j+1})$, respectively, by virtue of (6.2). Note that one can equivalently rewrite (6.4) as a system for U^{j+1} and $U^{j+1/3} := U(\cdot, t_{j+1/3})$ (rather than y^j).

REMARK 6.1. *The semidiscrete dG(1) method with quadrature, defined by (6.4) combined with (6.2) is, in fact, an implicit two-stage Runge-Kutta method of order three with the Butcher tableau:*

$$\begin{array}{c|cc} \frac13 & \frac{5}{12} & -\frac{1}{12} \\ 1 & \frac34 & \frac14 \\ \hline & \frac34 & \frac14 \end{array}$$

Remark 6.2. *The functions y^j and U^{j+1} obtained using the semidiscrete dG(1) method (6.3), as well as its version (6.4) with quadrature, respectively give a second-order approximation at the time level t_j and a third-order approximation at the time level t_{j+1}. Consequently, the more accurate approximate solutions $\{U^j\}_{j=1}^M$ are to be used, while the approximate solutions $\{y^j\}_{j=0}^{M-1}$ are auxiliary and may be discarded.*

6.1. A posteriori error estimate in the semidiscrete case. To formulate an a posteriori error estimate for the semidiscrete dG(1), we need to introduce some notation. Note that the computed solution interpolant U of (6.2) is only second-order accurate (as y^j is only second-order accurate, and also since U is a piecewise-linear interpolant). So we introduce a continuous *piecewise-quadratic* interpolant $\tilde{U}$ of the computed solution as follows:

$$\tilde{U} := U^j\phi_j + U^{j+1}\phi_{j+1} + \nu, \qquad \nu := 3\,\phi_j\,\phi_{j+1}\,\{y^j - U^j\} \qquad \text{on } [t_j, t_{j+1}]. \tag{6.5}$$

Remark 6.3. *The truncation error analysis shows that $U^j - y^j \approx \frac{1}{6}\tau_{j+1}^2\partial_t^2 u(\cdot, t_j)$, while U^j and U^{j+1} are third-order approximations of u at time levels t_j and t_{j+1}. So the piecewise-quadratic interpolant $\tilde{U}$ gives a third-order approximation to the exact solution u in the entire domain $\bar{Q}$.*

Furthermore, a comparison of (6.5) with (6.2) shows that

$$\tilde{U}(\cdot, t_{j+1/3}) = U(\cdot, t_{j+1/3}) = \tfrac{2}{3}\,y^j + \tfrac{1}{3}\,U^{j+1}, \quad \tilde{U}(\cdot, t_{j+1}) = U(\cdot, t_{j+1}) = U^{j+1}, \tag{6.6}$$

so $\frac{2}{3}\,y^j + \frac{1}{3}\,U^{j+1}$ gives a third-order approximation to $u(\cdot, t_{j+1/3})$.

Next, similarly to ψ of (6.2) , we define

$$\tilde{\psi} := \mathcal{L}(t)\,\tilde{U} + f(\cdot, t, \tilde{U}) \qquad \text{for } t \in [t_j, t_{j+1}]. \tag{6.7}$$

for which, (6.6) implies that

$$\tilde{\psi}(\cdot, t_{j+1/3}) = \psi^{j+1/3}, \qquad \tilde{\psi}(\cdot, t_{j+1}) = \psi^{j+1}.$$

Consequently, the quadratic interpolant $I_{2,t}\tilde{\psi}$ of $\tilde{\psi}$ in time on $[t_j, t_{j+1}]$, using the interpolation nodes t_j, $t_{j+1/3}$ and t_{j+1}, allows two equivalent representations:

$$\begin{aligned} I_{2,t}\tilde{\psi} &= \psi^j\,\phi_j + \psi^{j+1}\,\phi_{j+1} - \tfrac{1}{2}\tau_{j+1}^{-1}\,\phi_j\,\phi_{j+1}\,\chi^{j+1} \\ &= \psi^{j+1} - \tfrac{3}{2}\{\psi^{j+1} - \psi^{j+1/3}\}\,\phi_j - \tfrac{1}{2}\,\tau_{j+1}^{-2}(t - t_{j+1/3})\,\phi_{j+1}\,\chi^{j+1}, \end{aligned} \tag{6.8}$$

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

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

For the semidiscrete method with quadrature, we give the following result (a simplified version of this result for a t-independent elliptic operator $\mathcal{L}$ will be given in Section 6.2).

Theorem 6.4. *Let u solve the problem (1.1) with the parabolic operator $\mathcal{M}$ satisfying (1.2) and Condition 2.1, and U^j solve the corresponding semidiscrete problem (6.4), (6.2). Then for $m = 1, \ldots, M$, one has (3.1) with $\vartheta = \tilde{\psi} - I_{2,t}\tilde{\psi}$, $C_1 = \frac{2}{81}$, $C_2 = \frac{1}{18}$, and the notation (6.5)–(6.9).*

Proof. In view of (6.4a) combined with $\partial_t(U^j\phi_j+U^{j+1}\phi_{j+1})=(U^{j+1}-U^j)/\tau_{j+1}$, for $\tilde U$ of (6.5) one gets

$$\partial_t\tilde U=\partial_t\nu-\tfrac14\big(3\psi^{j+1/3}+\psi^{j+1}\big)\quad\text{for }\ t\in(t_j,t_{j+1}]. \tag{6.10}$$

Next, note that for any linear function $w=w(t)$

$$w(t)-w(t_{j+1/2})=(t-t_{j+1/2})\,w'=\partial_t[-\tfrac12\tau_{j+1}^2\,\phi_j\phi_{j+1}\,w'].$$

So for the function $\psi^{j+1}-\frac32\{\psi^{j+1}-\psi^{j+1/3}\}\,\phi_j$, linear for $t\in[t_j,t_{j+1}]$, one gets

$$\begin{aligned}\big(\psi^{j+1}-\tfrac32\{\psi^{j+1}-\psi^{j+1/3}\}\,\phi_j\big)-\tfrac14\big(3\psi^{j+1/3}+\psi^{j+1}\big)&\\ =\partial_t\big[-\tfrac34\tau_{j+1}\,\phi_j\phi_{j+1}\,\{\psi^{j+1}-\psi^{j+1/3}\}\big]&=-\partial_t\nu,\end{aligned}$$

where we also used (6.4b) and the definition of ν in (6.5). Also, a calculation yields

$$\tfrac12\,\tau_{j+1}^{-2}(t-t_{j+1/3})\,\phi_{j+1}\,\chi^{j+1}=\partial_t\big[-\tfrac16\,\phi_j^2\,\phi_{j+1}\,\chi^{j+1}\big].$$

So, taking the difference of the above two relations and recalling (6.8), one gets

$$I_{2,t}\tilde\psi-\tfrac14\big(3\psi^{j+1/3}+\psi^{j+1}\big)=\partial_t\mu-\partial_t\nu,\qquad \mu:=\tfrac16\,\phi_j^2\,\phi_{j+1}\,\chi^{j+1} \tag{6.11}$$

for $t\in(t_j,t_{j+1}]$. For this function μ, a calculation shows that

$$|\mu|\le\tfrac{2}{81}|\chi^{j+1}|,\quad \tau_{j+1}\,|\partial_t\mu|\le\tfrac{1}{18}|\chi^{j+1}|,\quad\text{for }\ t\in(t_j,t_{j+1}]. \tag{6.12}$$

Now, combining (6.10) with (6.11), we arrive at

$$\partial_t\tilde U=\partial_t\mu-I_{2,t}\tilde\psi\qquad\text{for }\ (x,t)\in Q.$$

As $\mathcal M\tilde U=\partial_t\tilde U+\tilde\psi$ and $\mathcal Mu=0$, while $\vartheta=\tilde\psi-I_{2,t}\tilde\psi$, so

$$\mathcal M\tilde U-\mathcal Mu=\partial_t\mu+\vartheta\qquad\text{in }\ Q.$$

The desired bound for $U^m-u(\cdot,t_m)=[\tilde U-u](\cdot,t_m)$ is then obtained by an application of Lemma 2.4 with μ of (6.11), using (6.12) combined with $\mu(\cdot,t_{m-1}^+)=0$ and the observation that $[\tilde U-u-\mu](\cdot,0)=U^0-\varphi=0$. This completes the proof. ∎

REMARK 6.5 (Computability). *The computation of the right-hand side in the estimate (3.1) involves computing χ^{j+1} of (6.9) for $j<m$. Note that the terms $\psi^{j+1/3}$ and ψ^{j+1}, which appear in (6.9), can be explicitly represented using (6.4); to be more precise, one gets*

$$\psi^{j+1/3}=-\frac{U^{j+1}-U^j}{\tau_{j+1}}-\frac{y^j-U^j}{\tau_{j+1}},\quad \psi^{j+1}=-\frac{U^{j+1}-U^j}{\tau_{j+1}}+3\,\frac{y^j-U^j}{\tau_{j+1}}. \tag{6.13}$$

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. Note that in this case, the semidiscrete method (6.4), (6.2) can be rewritten as

$$U^{j+1}-U^j+\tfrac12\,\tau_{j+1}\,\mathring{\mathcal L}\big(y^j+U^{j+1}\big)+\tfrac14\,\tau_{j+1}\,\big(3f^{j+1/3}+f^{j+1}\big)=0, \tag{6.14a}$$

$$y^j-U^j+\tfrac16\,\tau_{j+1}\,\mathring{\mathcal L}\big(y^j-U^{j+1}\big)+\tfrac14\,\tau_{j+1}\,\big(f^{j+1/3}-f^{j+1}\big)=0, \tag{6.14b}$$

where we use the notation

$$f^{j+1/3} := f(\cdot, t_{j+1/3}, \tfrac{2}{3}\, y^j + \tfrac{1}{3}\, U^{j+1}), \qquad f^{j+1} := f(\cdot, t_j, U^{j+1}). \tag{6.15}$$

Here we also used the observations that

$$\psi^{j+1/3} = \mathring{\mathcal{L}}[\tfrac{2}{3}\, y^j + \tfrac{1}{3}\, U^{j+1}] + f^{j+1/3}, \quad \psi^{j+1} = \mathring{\mathcal{L}} U^{j+1} + f^{j+1}.$$

Now consider the quantities that appear in the estimator given by Theorem 6.4. A similar calculation for χ^{j+1} of (6.9) yields

$$\chi^{j+1} = 6\tau_{j+1}\, \mathring{\mathcal{L}}[U^j - y^j] + 3\tau_{j+1}\, [2f^j - 3f^{j+1/3} + f^{j+1}]. \tag{6.16}$$

The estimator also involves $\vartheta = \tilde{\psi} - I_{2,t}\tilde{\psi}$. By (6.7), $\tilde{\psi} = \mathring{\mathcal{L}}\, \tilde{U} + f(\cdot, t, \tilde{U})$, where $I_{2,t}[\mathring{\mathcal{L}}\, \tilde{U}] = \mathring{\mathcal{L}}\, [I_{2,t}\tilde{U}] = \mathring{\mathcal{L}}\, \tilde{U}$, so

$$\vartheta = f(\cdot, t, \tilde{U}) - I_{2,t}[f(\cdot, t, \tilde{U})]. \tag{6.17}$$

Note that this quantity 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 6.6. *Let the elliptic operator $\mathcal{L}(t) = \mathring{\mathcal{L}}$ be independent of the variable t; then the statement of Theorem 6.4 remains valid with the simplifications (6.16), (6.17) and the notation (6.15).*

Finally, recall that in the model problem (1.3) the elliptic operator $\mathcal{L} = -\varepsilon^2 \triangle$ is t-independent, so we apply Corollary 6.6 to this problem.

COROLLARY 6.7. *Under assumption (1.2), the a posteriori error estimate of Theorem 6.4 applies to the model problem (1.3) with the constants κ_0, κ_1, κ_2 from Lemma 2.2, the simplifications (6.16), (6.17) and the notation (6.15).*

6.3. Generalization for higher-order discontinuous Galerkin methods. In this section we generalize our results for the dG(1) method to higher-order discontinuous Galerkin methods dG(r) for $r \geq 1$ with Radau quadrature.

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

$$\zeta_{r+1}(0) = 1, \qquad \zeta_{r+1}(\alpha_k) = 0 \quad \text{for } k = 0, \ldots, r, \qquad C_\zeta := \tfrac{d^{r+1}}{ds^{r+1}} \zeta_{r+1}(s). \tag{6.18}$$

Let us introduce two interpolants on $(t_j, t_{j+1}]$: $\hat{I}_{r,t}\phi \in \Pi_r$ with $\big(\hat{I}_{r,t}\phi\big)(t_{j+\alpha}) = \phi(t_{j+\alpha})$ for $\alpha \in \mathcal{A}_R$, and $I_{r+1,t}\phi \in \Pi_{r+1}$ with $\big(I_{r+1,t}\phi\big)(t_{j+\alpha}) = \phi(t_{j+\alpha})$ for $\alpha \in \mathcal{A}$. (e.g., for $r = 1$, the interpolant $I_{r+1,t}$ coincides with $I_{2,t}$ in (6.8)).

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 $k = 0, \ldots, r$, respectively associated with the time levels $t_{j+\alpha_k}$, to satisfy

$$\big[U(\cdot, t_j^+) - U^j\big]\, \varphi_k(0) + \int_{t_j^+}^{t_{j+1}} \big[\partial_t U + \hat{I}_{r,t}\psi\big]\, \varphi_k\big(\tfrac{t - t_j}{\tau_{j+1}}\big)\, dt = 0 \quad \text{for } k = 0, \ldots, r, \tag{6.19a}$$

$$\text{where} \quad U := \sum_{k=0}^{r} U^{j+\alpha_k}\, \varphi_k\big(\tfrac{t - t_j}{\tau_{j+1}}\big), \quad \psi := \mathcal{L}(t)\, U + f(\cdot, t, U) \quad \text{for } t \in (t_j, t_{j+1}]. \tag{6.19b}$$

Note that (6.19) (equivalent to (6.4) for $r = 1$) 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 ψ, then we get the dG(r) method without quadrature (in fact, equivalent to (6.3) for $r = 1$).

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 $\tilde{\psi}$:

$$\tilde{U} := U - \big[U(\cdot, t_j^+) - U^j\big]\, \zeta_{r+1}\big(\tfrac{t-t_j}{\tau_{j+1}}\big), \qquad \tilde{\psi} = \mathcal{L}(t)\, \tilde{U} + f(\cdot, t, \tilde{U}). \tag{6.20}$$

Note that $I_{r+1,t}\tilde{\psi}$ allows a representation (similar to (6.8), (6.9)):

$$I_{r+1,t}\tilde{\psi} = \hat{I}_{r,t}\psi + \chi^{j+1}\, \tau_{j+1}^{-1}\, C_\zeta^{-1}\, \zeta_{r+1}\big(\tfrac{t-t_j}{\tau_{j+1}}\big), \tag{6.21}$$

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

$$\chi^{j+1} := \tau_{j+1}\, C_\zeta \Big[\psi^j - \hat{I}_{r,t}\psi(\cdot, t_j^+)\Big] = \tau_{j+1}\, C_\zeta \Big[\psi^j - \sum_{k=0}^{r} \psi^{j+\alpha_k}\, \varphi_k(0)\Big]. \tag{6.22}$$

THEOREM 6.8. *Let u solve the problem (1.1) with the parabolic operator $\mathcal{M}$ satisfying (1.2) and Condition 2.1, and U^j solve the corresponding semidiscrete problem (6.19). Then for $m = 1, \ldots, M$, one has (3.1) with $\vartheta = \tilde{\psi} - I_{r+1,t}\tilde{\psi}$, the constants $C_1 = C_\zeta^{-1} \max_{s\in[0,1]} |\int_0^s \zeta_{r+1}(\sigma)\, d\sigma|$ and $C_2 = C_\zeta^{-1} \max_{s\in[0,1]} |\zeta_{r+1}(s)|$, and the notation (6.18), (6.20), (6.22).*

Proof. First, note that (6.19a) is equivalent to

$$\int_{t_j}^{t_{j+1}} \big[\partial_t\tilde{U} + \hat{I}_{r,t}\psi\big]\, \varphi_k\big(\tfrac{t-t_j}{\tau_{j+1}}\big)\, dt = 0 \qquad \text{for } k = 0, \ldots, r. \tag{6.23}$$

This is easily checked by getting $\partial_t[\tilde{U} - U]$ from the first relation in (6.20), and then noting that $\int_{t_j}^{t_{j+1}} \partial_t\big\{-\zeta_{r+1}\big(\tfrac{t-t_j}{\tau_{j+1}}\big)\big\} \cdot \varphi_k\big(\tfrac{t-t_j}{\tau_{j+1}}\big)\, dt = \varphi_k(0)$ (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 (6.23) 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

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

It should be noted that by virtue of (6.21), the function μ is continuous in time (this follows from ζ_{r+1} vanishing at the Radau points). Furthermore, μ satisfies the bounds (6.12) with $\frac{2}{81}$ and $\frac{1}{18}$ respectively replaced by C_1 and C_2. Hence, the desired bound for $U^m - u(\cdot, t_m) = [\tilde{U} - u](\cdot, t_m)$ is then obtained by an application of Lemma 2.4. ∎

REMARK 6.9. *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 6.8 gives an a posteriori error estimate of order $r+2$.*

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

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

and its discretizations in the form

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

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

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

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

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

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 either for the elliptic problem or the parabolic problem.

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 $\Omega \subset \mathbb{R}^n$:

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

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

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

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 7.1. *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 $\langle \mathcal{L}_h \varphi, w_h \rangle = \varepsilon^2 \langle \nabla\varphi, \nabla w_h \rangle$ and $\langle \mathcal{P}_h q, w_h \rangle = \langle q, w_h \rangle_h$, 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 7.2. *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 ϕ_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* $v_h \in \mathring{V}_h$ *by applying the normalized stiffness matrix to the column vector of nodal values* $\{v_{h,i}\}$*.*

We now cite elliptic estimators of type (7.3) for particular cases of (7.4) and (7.5).

7.2. 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 $\bar\Omega$ made of elements T, we let V_h be the space of continuous piecewise polynomial finite element functions of degree $l \geq 1$, and $\mathring{V}_h := V_h \cap H_0^1(\Omega)$. 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 [22, Theorem 4.2], an a posteriori error estimate of type (7.3) is given with $\eta = \eta_0$ defined by

$$\eta_0\big(V_h, v_h, g(\cdot, v_h)\big) := \Big[c_0 \max_{T\in\mathcal{T}_h}\Big\{h_T^2\big\|(\triangle v_h - g(\cdot, v_h))\big\|_{\infty,T} + h_T\big\|[\![\partial_n v_h]\!]\big\|_{\infty,\partial T\setminus\partial\Omega}\Big\} + c_1\,\big\|\nu^{q}_{n/2,T}\big\|_{l_{n/2}} + c_2\,\big\|h_T\,\nu^{q-1}_{n,T}\big\|_{l_n}\Big] \times |\ln h_{\min}|^2, \tag{7.6}$$

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 inter-element side, $\|\cdot\|_{l_p}$ is the l_p norm, and the quantity

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

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

7.3. Elliptic model problem: singularly-perturbed regime in one dimension. 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 $0 = x_0 < x_1 < \cdots < x_N = 1$ 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., [19]).

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

$$\langle\varphi, w_h\rangle_h := \langle I_h\varphi, w_h\rangle, \qquad \text{(quadrature)} \tag{7.7a}$$

$$\langle\varphi, w_h\rangle_h := \langle I_h[\varphi w_h], 1\rangle. \qquad \text{(lumped-mass quadrature)} \tag{7.7b}$$

REMARK 7.3. *To illustrate Remarks 7.1 and 7.2, note that the described two discretizations using either (7.7a) or (7.7b) are of type (7.2). In particular, for (7.7a), we get* $\mathcal{L}_h := -\varepsilon^2[\partial_x^2]_h$ *and* $\mathcal{P}_h := I_h$*. Here the operator* $[\partial_x^2]_h : H_0^1(\Omega) \to \mathring{V}_h +$

$\varepsilon^{-2} I_h[g(\cdot,0)]$ 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 (7.7a) may be represented as

$$-\varepsilon^2\,[\partial_x^2]_h v_h + I_h[g(\cdot, v_h)] = 0. \tag{7.8a}$$

By contrast, (7.7b) 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 $\delta_x^2 v_{h,i} := \frac{2}{h_i+h_{i+1}}\big[\frac{1}{h_{i+1}}(v_{h,i+1} - v_{h,i}) - \frac{1}{h_i}(v_{h,i} - v_{h,i-1})\big]$ 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 (7.7b) as

$$-\varepsilon^2\, I_h[\delta_x^2 v_h] + I_h[g(\cdot, v_h)] = 0, \tag{7.8b}$$

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

We cite a posteriori error bounds [13, 18, 19] of type (7.3) with $\eta := \eta_\varepsilon\big(V_h, g(\cdot, v_h)\big)$ for (7.7a) and $\eta := \eta_{\varepsilon;\,\text{l.m.}}\big(V_h, g(\cdot, v_h)\big)$ for (7.7b), respectively, defined by

$$\eta_\varepsilon\big(V_h, g_*\big) := \max_{i=1,\dots,N}\Big\{\tfrac{h_i^2}{4\varepsilon^2}\,\|I_h g_*\|_{\infty,(x_{i-1},x_i)}\Big\} + \gamma^{-2}\|g_* - I_h g_*\|_{\infty,(0,1)}, \tag{7.9a}$$

$$\eta_{\varepsilon;\,\text{l.m.}}\big(V_h, g_*\big) := \eta_\varepsilon + \max_{i=1,\dots,N}\Big\{\tfrac{h_i^2}{6\gamma\,\varepsilon}\,\|\partial_x(I_h g_*)\|_{\infty,(x_{i-1},x_i)}\Big\}, \tag{7.9b}$$

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

Remark 7.4. *The error estimators (7.9a) and (7.9b) are **robust** although they involve negative powers of the small parameter ε. Indeed, an inspection of representations (7.8a) and (7.8b) 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 $\eta := \max_{i=1,\dots,N}\big\{\frac{h_i^2}{4\varepsilon^2}\,\|g_*\|_{\infty,(x_{i-1},x_i)}\big\}$. 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 (7.8a), 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 ε 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 with (7.9a), only I_h in η_ε 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).

Remark 7.5. *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 [14, 4].*

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 semidsicrete 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 $u_h^{j+1} \in \mathring{V}_h^{j+1} := V_h^{j+1} \cap H_0^1(\Omega)$ 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

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

Note two particular cases of interest for the auxiliary computed solution $\hat u_h^j$:

$$\text{Case }\mathbf{A}:\quad \hat u_h^j := I_*^{j+1} u_h^j, \quad I_*^{j+1} : \mathring{V}_h^j \to \mathring{V}_h^{j+1} \quad \Rightarrow \quad \mathcal{P}_h^{j+1} \hat u_h^j = \hat u_h^j, \tag{8.2a}$$

$$\text{Case }\mathbf{B}:\quad \hat u_h^j := u_h^j \quad \Rightarrow \quad \hat u_h^j \in \mathring{V}_h^j, \;\; u_h^{j+1} \in \mathring{V}_h^{j+1}. \tag{8.2b}$$

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 frequently assumed that $I_*^{j+1} w_h = w_h$ for all $w_h \in \mathring{V}_h^{j+1}$. To define I_*^{j+1}, one may employ, e.g., the standard Lagrange interpolation or the L_2 projection. Note that if $V_h^j = V_h^{j+1}$, then Cases A and B are identical.

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

$$\begin{aligned} \big\|u_h^m - u(\cdot,t_m)\big\|_{\infty,\Omega} \le\; & \kappa_0\, \mathrm{e}^{-\gamma^2 t_m} \big\|u_h^0 - \varphi\big\|_{\infty,\Omega} \\ & + (\kappa_1\, \ell_m + \kappa_2) \max_{j=1,\ldots,m-1} \Big\{ C_1 \,\big\|\chi_h^j\big\|_{\infty,\Omega} + C_1^*\, \eta^j \Big\} \\ & + C_2\, \kappa_0\, \big\|\chi_h^m\big\|_{\infty,\Omega} + (C_2^* \kappa_0 + 1)\, \eta^m \\ & + \kappa_0 \sum_{j=1}^{m} \int_{t_{j-1}}^{t_j} \mathrm{e}^{-\gamma^2 (t_m - s)} \big\|\vartheta_h(\cdot,s)\big\|_{\infty,\Omega}\, \mathrm{d}s \\ & + \kappa_0 \sum_{j=1}^{m-1} \mathrm{e}^{-\gamma^2 (t_m - t_j)} \big\|\hat u_h^j - u_h^j\big\|_{\infty,\Omega}\,. \end{aligned} \tag{8.3}$$

The quantities that appear in this estimate are specified by Theorems 9.3, 10.4 and 11.5 below, and can be summarized as follows:

	p	χ_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	$\tau_{j+1}(\psi_h^{j+1} - \hat\psi_h^j)$	$\frac{1}{8}$	$\frac{5}{8}$	2	4	$\{0^+, 1\}$
dG(1)-Radau	3	$3\tau_{j+1}(2\hat\psi_h^j - 3\psi_h^{j+1/3} + \psi_h^{j+1})$	$\frac{2}{81}$	$\frac{1}{18}$	$\frac{5}{3}$	10	$\{0^+, \frac{1}{3}, 1\}$

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

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

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

where α takes values from $\mathcal{A}\backslash\{0^+\}$, for which the computed solution $u_h^{j+\alpha}$ is available from the definition of the method, while for $\alpha = 0^+$ we have $\hat u_h^j$ 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 η^j in (8.3) is related to the error due to the spatial discretization used; it is defined using the elliptic estimator η from (7.3) by

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

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

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

where

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

The quantity ϑ_h is similar to ϑ 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 $j = 1, \ldots, M$, 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 9.4, which applies to the backward Euler, and also to the Crank-Nicolson and dG(1) methods.*

8.1. Computability of χ_h^j and η^j in the a posteriori error estimate (8.3). For the ***backward Euler*** method we shall use (see Remark 9.1 on ψ_h^{j+1})

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

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

For the ***Crank-Nicolson*** and ***dG*(1)** methods, the computability of χ_h^j and η^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 ψ_h^{j+1} (by means of (10.3) assuming that $\hat{\psi}_h^j$ is available), while for the dG(1) method, $\psi_h^{j+1/3}$ and ψ_h^{j+1} are explicitly computable (by means of (11.2a)). So, if $\hat{\psi}_h^j$ is available, one can indeed explicitly compute χ_h^{j+1} and η^{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, $\hat{u}_h^j \in \mathring{V}_h^{j+1}$ 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 7.1 and 7.2; see also Remark 7.3. (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 ψ_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 7.2, $\mathcal{P}_h^{j+1} = I_h^{j+1}$ is some interpolation operator onto V_h^{j+1}, while $\mathcal{L}_h(t_j^+)\,\hat{u}_h^j$ is easily computable for any $\hat{u}_h^j \in \mathring{V}_h^{j+1}$ 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) Recall that $\hat{u}_h^j$ is obtained from u_h^j by means of some linear interpolation I_*^{j+1} in (8.2a). One possible choice is $\hat{u}_h^j \in \mathring{V}_h^{j+1}$ such that

$$\varepsilon^2 \big\langle \nabla \hat{u}_h^j, \nabla w_h \big\rangle + \big\langle f(\cdot, t_j, \hat{u}_h^j),\, w_h \big\rangle_h = \big\langle \psi_h^j,\, w_h \big\rangle_h \qquad \forall w_h \in \mathring{V}_h^{j+1}.$$

Note that then the computation of $\hat{u}_h^j$ is more expensive than when it is obtained by an explicit computation (also, $\mathring{V}_h^j \subset \mathring{V}_h^{j+1}$ does not imply $\hat{u}_h^j = u_h^j$). But this choice implies that $\hat{\psi}_h^j = \mathcal{P}_h^{j+1}\psi_h^j$, where ψ_h^j has already been computed. If, furthermore, a quadrature of type $\langle \varphi, w_h \rangle_h := \langle I_h^{j+1}\varphi, w_h \rangle$ is used for all $\varphi \in C(\bar{\Omega})$ and $w_h \in \mathring{V}_h^{j+1}$, then $\mathcal{P}_h^{j+1}$ is identical with the interpolation I_h^{j+1} (see Remark 7.1(ii)), so the computation of $\hat{\psi}_h^j$ becomes explicit.

(iv) 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 7.1 implies that, roughly speaking, $\mathcal{L}_h(t_j^+)v_h$ for any $v_h \in \mathring{V}_h^{j+1}$ 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 $\big\|\hat{\psi}_h^j - w_h\big\|_{\infty,\Omega} \le \|M^{-1}\|_\infty \cdot \big\|M(\hat{\psi}_h^j - w_h)\big\|_{\infty,\Omega}$, 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 $V_h^j \ne V_h^{j+1}$, for the Crank-Nicolson method, $\hat{\psi}_h^j$ is not given by the right-hand side in (8.4a), so ψ_h^{j+1} and $\hat{\psi}_h^j$ are computed by means of (10.2), using the above items* (ii) *or* (iv) *in the computation of ψ_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 11.4.*

8.2. Elliptic Reconstruction. Definition of ϑ_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 [20, 17, 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 $\alpha \in \mathcal{A}\backslash\{0^+\}$ 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

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

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

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

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 η^{j+1} of (8.5), one gets

$$\|\hat{R}^j - \hat{u}_h^j\|_{\infty,\Omega} \le \eta^{j+1} \text{ if } 0^+ \in \mathcal{A}, \quad \|R^{j+\alpha} - u_h^{j+\alpha}\|_{\infty,\Omega} \le \eta^{j+1} \text{ for } \alpha \in \mathcal{A}\backslash\{0^+\}. \tag{8.9}$$

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

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

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

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

8.3. Estimation of $\boldsymbol{\vartheta_h}$. We now briefly discuss possible approaches to the estimation of ϑ_h in the case of a ***t-independent $\mathcal{L}$***, which includes the model problem (1.3) (for a general $\mathcal{L}$, see Remark 8.6 below). Then ϑ_h of (8.10) simplifies to

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

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. E.g., for $t \in (t_{j-1},t_j]$, we have*

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

where we used (8.9), and η^j is computed using (8.5a). In fact, if $|\partial_t\partial_z f| \le C$, 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 η^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 $|\partial_z f| \le C_f$ for some constant C_f, using*

$$\big|\vartheta_{f,\tilde{R}} - \vartheta_{f,\tilde{u}_h}\big| \le \big|f(\cdot,t,\tilde{u}_h) - f(\cdot,t,\tilde{R})\big| + \big|I^*_{p,t}[f(\cdot,t,\tilde{u}_h) - f(\cdot,t,\tilde{R})\big|,$$

one easily gets a very crude bound $\|[\vartheta_{f,R} - \vartheta_{f,\tilde{u}_h}](\cdot,t)\|_{\infty,\Omega} \le C_* C_f\, \eta^{j+1}$ *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* $\|[\vartheta_{f,\tilde{R}} - \vartheta_{f,I\tilde{u}_h}](\cdot,t)\|_{\infty,\Omega} \le C\,\tau_{j+1}\,\eta^{j+1}$ *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* η^{j+1} *that already appear in (8.3).*

Remark 8.6 (Case $\mathcal{L} = \mathcal{L}(t)$). *One can use similar strategies to estimate* ϑ_h *even if* $\mathcal{L} = \mathcal{L}(t)$*. For example, for the backward Euler method,* $\vartheta_h = \vartheta_{\mathcal{L},\tilde{R}} + \vartheta_{f,\tilde{R}}$ *with* $\vartheta_{f,\tilde{R}}$ *of (8.12), and an additional component* $\vartheta_{\mathcal{L},\tilde{R}} := [\mathcal{L}(t) - \mathcal{L}^j]\, R^j$ *for* $t \in (t_{j-1}, t_j]$*. Similarly to* $\vartheta_{f,\tilde{R}}$*, one may expect* $\vartheta_{\mathcal{L},\tilde{R}} \approx \vartheta_{\mathcal{L},\tilde{u}_h}$*, where for some operators* $\mathcal{L}(t)$ *the quantity* $\vartheta_{\mathcal{L},\tilde{u}_h}$ *can be computed using sampling, while to estimate* $\vartheta_{\mathcal{L},\tilde{R}} - \vartheta_{\mathcal{L},\tilde{u}_h} = [\mathcal{L}(t) - \mathcal{L}^j]\,(R^j - u_h^j)$ *one would need an appropriate elliptic estimator for certain spatial derivatives of* $R^j - u_h^j$*.*

9. Fully discrete backward Euler method. To fully discretize the abstract parabolic problem (1.1), we now apply a spatial discretization of type (7.2) to the semidiscrete problem (4.1) as follows. We associate a finite-element space $V_h^j \subset C(\bar{\Omega})$ and a computed solution $u_h^j \in \mathring{V}_h^j := V_h^j \cap H_0^1(\Omega)$ with the time level t_j and require, for $j = 1, \dots, M$, that

$$\mathcal{L}_h^j\, u_h^j + \mathcal{P}_h^j\, [f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j] = 0, \tag{9.1a}$$

with $\mathcal{L}_h^j := \mathcal{L}_h(t_j)$ and $\mathcal{P}_h^j$ subject to (8.1), and some $u_h^0 \approx \varphi$ in $\mathring{V}_h^0$. Note that $\mathcal{L}_h^j\, u_h^j \in \mathring{V}_h^j - I_h^j[f(\cdot, t_j, 0)]$, while as both u_h^j and $\delta_t^* u_h^j$ vanish on $\partial\Omega$, so $\mathring{V}_h^j - I_h^j[f(\cdot, t_j, 0)]$ coincides with $\mathring{V}_h^j - I_h^j[f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j]$, so the definition (9.1) is consistent.

The term $\delta_t^* u_h^j$ in (9.1a) approximates $\partial_t u$ and is defined by

$$\delta_t^* u_h^j := \frac{u_h^j - \hat{u}_h^{j-1}}{\tau_j}, \quad \text{where } \hat{u}_h^0 := u_h^0. \tag{9.1b}$$

The operator δ_t^* is identical with δ_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).

9.1. A posteriori error estimate using a piecewise-constant elliptic reconstruction. To estimate the error of the fully discrete backward Euler method (9.1), 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 (9.1a),* $\psi_h^j = -\delta_t^* u_h^j$ *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.* ψ_h^j *satisfies (8.4b). In Case A of (8.2a), this relation simplifies using* $\mathcal{P}_h^j\, \psi_h^j = \psi_h^j$*.*

Remark 9.2. *Remark 9.1 implies that* R^j *satisfies (8.9) with* $\mathcal{A} := \{1\}$*.*

We now give an a posteriori error estimate for the fully discrete method (9.1).

Theorem 9.3. *Let* u *solve the problem (1.1), (1.2) with the parabolic operator* $\mathcal{M}$ *satisfying Condition 2.1,* u_h^j *solve the discrete problem (9.1). Then for* $m = 1, \dots, M$*, one has (8.3) with* η^j *and* χ_h^j *defined by (8.5a), (8.6),* $C_1 = 1$*,* $C_2 = 2$*,* $C_1^* = C_2^* = 1$*, and a regular function* ϑ_h *defined, for* $t \in (t_{j-1}, t_j]$*,* $j = 1, \dots, M$*, by*

$$\vartheta_h(\cdot, t) = \tilde{\psi}_R(\cdot, t) - \tilde{\psi}_R(\cdot, t_j), \quad \tilde{\psi}_R(\cdot, t) = \mathcal{L}(t) R^j + f(\cdot, t, R^j) \ \text{ for } t \in (t_{j-1}, t_j]. \tag{9.2}$$

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

THEOREM 9.3*. *The statement of Theorem 9.3 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* $\mathrm{e}^{-\gamma^2(t_m-t_j)}$ *replaced by* $\mathrm{e}^{-\gamma^2(t_m-t_{j+1})}$*.*

We first give a proof of Theorem 9.3*, and then generalize it to prove Therorem 9.3.

Proof of Theorem 9.3.* In view of Remark 9.2, $\|R^j - u_h^j\|_{\infty,\Omega} \le \eta^j$, 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:

$$I_{1,t}u_h(\cdot, t) := \tfrac{t_j - t}{\tau_j}\, u_h^{j-1} + \tfrac{t - t_{j-1}}{\tau_j}\, u_h^j \quad \text{for } t \in [t_{j-1}, t_j], \quad j = 1, \ldots, M. \tag{9.3}$$

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

$$\tilde{R}(\cdot, t) := R^j \quad \text{for } t \in (t_{j-1}, t_j], \quad j = 1 \ldots, M; \qquad \tilde{R}(\cdot, 0) := R^1, \tag{9.4}$$

(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 (9.2) are identical with those in (8.10) (using $p = 0$), 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 (9.2) yields $\tilde{\psi}_R = \vartheta_h - \delta_t^* u_h^j$, so

$$\partial_t(I_{1,t}u_h) + \tilde{\psi}_R = \vartheta_h + \vartheta_* \qquad \text{in } Q, \tag{9.5}$$

where ϑ_* is a regular function defined by

$$\vartheta_*(\cdot, t) := \partial_t(I_{1,t}u_h) - \delta_t^* u_h^j \quad \text{for } t \in (t_{j-1}, t_j]. \tag{9.6}$$

As $\mathcal{M}\tilde{R} = \partial_t \tilde{R} + \tilde{\psi}_R$ and $\mathcal{M}u = 0$, so (9.5) yields

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

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 2.4 with $\mu := \tilde{R} - I_{1,t}u_h$ and $\vartheta := \vartheta_h + \vartheta_*$, using the following three observations. First, note that

$$[\tilde{R} - u - \mu](\cdot, 0) = R^1 - \varphi - (R^1 - u_h^0) = u_h^0 - \varphi. \tag{9.7}$$

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,

$$|\mu| \le |R^j - u_h^j| + |u_h^j - u_h^{j-1}| \quad \text{and} \quad \tau_j\, |\partial_t \mu| = |u_h^j - u_h^{j-1}|, \tag{9.8}$$

where $\|R^j - u_h^j\|_{\infty,\Omega} \le \eta^j$. Finally, (9.6) combined with (9.1b), (9.3) 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,

$$\int_{t_j}^{t_{j+1}} \mathrm{e}^{-\gamma^2(t_m - s)}\, \big\|\vartheta_*(\cdot, s)\big\|_{\infty,\Omega}\, \mathrm{d}s \le \mathrm{e}^{-\gamma^2(t_m - t_{j+1})}\big\|\hat{u}_h^j - u_h^j\big\|_{\infty,\Omega}, \tag{9.9}$$

where $\hat{u}_h^0 - u_h^0 = 0$. The three observations (9.7), (9.8), (9.9) yield the required bound for $\|R^m - u(\cdot, t_m)\|_{\infty,\Omega}$. □

Proof of Theorem 9.3. We imitate the proof of Theorem 9.3*, only $I_{1,t}u_h$ of (9.3) is replaced everywhere by the piecewise-continuous interpolant

$$I^*_{1,t}u_h(\cdot,t) := \tfrac{t_j-t}{\tau_j}\,\hat{u}_h^{j-1} + \tfrac{t-t_{j-1}}{\tau_j}\,u_h^j \quad \text{for } t\in(t_{j-1},t_j],\quad j=1,\ldots,M, \tag{9.10}$$

with $I^*_{1,t}u_h(\cdot,0) := \hat{u}_h^0 = u_h^0$. Furthermore, ϑ_* is defined not by (9.6), but by

$$\vartheta_*(\cdot,t) := \partial_t(I^*_{1,t}u_h) - \delta_t^* u_h^j = [\hat{u}_h^{j-1} - u_h^{j-1}]\,\delta(t-t^+_{j-1}) \quad \text{for } t\in(t_{j-1},t_j], \tag{9.11}$$

where $\delta(\cdot)$ is the one-dimensional Dirac δ-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 (9.9) we use

$$\int_0^{t_m} \mathrm{e}^{-\gamma^2(t_m-s)}\,\|\vartheta_*(\cdot,s)\|_{\infty,\Omega}\,\mathrm{d}s \le \sum_{j=1}^{m-1} \mathrm{e}^{-\gamma^2(t_m-t_j)}\|\hat{u}_h^j - u_h^j\|_{\infty,\Omega}\,. \tag{9.12}$$

The required bound for $R^m - u(\cdot,t_m) = [\tilde{R}-u](\cdot,t_m)$ is again obtained by an application of Lemma 2.4 only with $\mu := \tilde{R} - I^*_{1,t}u_h$, for which we have a version of (9.8) with u_h^{j-1} replaced by $\hat{u}_h^{j-1}$. □

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

$$\text{Find } u_h^j \in \mathring{V}_h^j:\quad \varepsilon^2\langle \nabla u_h^j, \nabla w_h\rangle + \big\langle f(\cdot,t_j,u_h^j) + \delta_t^* u_h^j,\, w_h\big\rangle_h = 0 \quad \forall\, w_h\in \mathring{V}_h^j, \tag{9.13}$$

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 $\langle\cdot,\cdot\rangle$, and $\delta_t^* u_h^j$ is defined by (9.1b), (8.2).

Note that the full discretization (9.13) is of type (9.1). 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 7.1 and 7.2 only using V_h^j instead of V_h.

REMARK 9.4 (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*

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

with $\mathcal{H}_j$ representing the local mesh size associated with $\mathring{V}_h^j$. This version of (8.3) is easily obtained by an application of Lemma 2.4 using ϑ_* from (9.11) 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 $\mathring{V}_h^j$. Then $\langle u_h^{j-1} - \hat{u}_h^{j-1}, \Gamma_h^j\rangle = 0$ for any $\Gamma_h^j\in\mathring{V}_h^j$. So choosing $\Gamma_h^j = I_h^j\Gamma(\cdot,t_j)$, it suffices to show that $\|\mathcal{H}_j^{-2}\{\Gamma(\cdot,t_j)-\Gamma_h^j\}\|_{1,\Omega} \le \kappa_3\,\frac{C}{t_m-t_j}$. The desired result follows if one has $\|\Gamma(\cdot,t_j)\|_{W_1^2(\Omega)} \le \kappa_3\,\frac{C}{t_m-t_j}$. The latter bound is crucial in this argument; it involves the spatial derivatives of Γ and can be obtained from [7, (2.2)] if $\mathcal{L} = -\triangle + 1$ in a smooth domain and $f=f(x,t)$, with an unspecified $\kappa_3 = \mathcal{O}(1)$; and from [15, (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), $\triangle\Gamma = \varepsilon^{-2}[-\partial_s + a]\Gamma$; see also [15]). 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).

9.2.1. 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 u_h^j solve the discrete problem (9.13) with V_h^j and $\langle \cdot, \cdot \rangle_h$ defined, for each time level t_j, as in §7.2. 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 $\mathring{V}_h^j := V_h^j \cap H_0^1(\Omega)$. 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 §7.2.

Corollary 9.5. *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 9.3 and 9.3* are valid with ϑ_h simplified to (8.12) and estimated as described in Remark 8.4, and*

$$\eta^j := \eta_0\big(V_h^j, u_h^j, f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j\big) \qquad \textit{for } j = 1, \ldots, M,$$

where η_0 is defined in (7.6).

Remark 9.6. *The backward Euler method for a linear version of (1.3) with $\varepsilon = 1$ was considered in [8, 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 $\mathring{V}_h^j$. The a posteriori error estimate of Corollary 9.5 resembles (but is not identical with) the ones of [8, (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 [8, (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 9.11 and 11.8 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 $\tau_j |\partial_t^2 u + \ldots|$, which seems less suitable for a first-order method in time.

The mesh-coarsening terms in [8, 7] are similar to (9.14).

9.2.2. 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 (9.13) with V_h^j and $\langle \cdot, \cdot \rangle_h$ defined, for each time level t_j, as in §7.3. Thus V_h^j is the space of continuous ***piecewise-linear*** finite element functions on an arbitrary nonuniform mesh $\{x_i^j\}_{i=1}^N$ with $0 = x_0^j < x_1^j < \cdots < x_N^j = 1$ under absolutely no mesh regularity assumptions. We consider the two choices (7.7a) and (7.7b) of $\langle \cdot, \cdot \rangle_h$, using the piecewise-linear interpolant $I_h := I_h^j$ onto V_h^j.

Corollary 9.7. *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 9.3 and 9.3* are valid with ϑ_h simplified to (8.12) and estimated as described in Remark 8.4, and*

$$\eta^j := \begin{cases} \eta_\varepsilon\big(V_h^j, f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j\big) & \textit{for (7.7a)}, \\ \eta_{\varepsilon;\,\mathrm{l.m.}}\big(V_h^j, f(\cdot, t_j, u_h^j) + \delta_t^* u_h^j\big) & \textit{for (7.7b)}, \end{cases}$$

for $j = 1, \ldots, M$, where η_ε and $\eta_{\varepsilon;\,\mathrm{l.m.}}$ are defined in (7.9), with I_h replaced by I_h^j.

Remark 9.8. *The a posteriori error estimators of Corollary 9.7 are **robust**. Indeed, the only terms in (8.3) that involve the small parameter ε are η^j. As the argument of Remark 7.4 applies to η^j with v replaced by $u(\cdot, t_j)$, so $\varepsilon^{-2} h_i^2 |I_h g_*|$ approximates $h_i^2 |\partial_x^2 u(\cdot, t_j)|$, while the term $\varepsilon^{-1} h_i^2 |\partial_x (I_h g_*)|$ approximates $\varepsilon\, |\partial_x^3 u(\cdot, t_j)|$, which has similar magnitude to $h_i^2 |\partial_x^2 u(\cdot, t_j)|$ in the layer regions.*

Furthermore, the numerical results in [15, §4 with Remark 3.2] show that at least on a fixed layer-adapted mesh, our estimator is quite efficient independently of ε.

REMARK 9.9. *Consider the term* $\|g_* - I_h^j g_*\|_{\infty,\Omega}$ *in the error estimators of Corollary 9.7 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 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 9.10. *Using Remark 7.5, one can extend the error bound of Corollary 9.7 to the singularly perturbed equation of (1.3) in two and three dimensions.*

REMARK 9.11. *The backward Euler method for equation (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* ε *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. We also refer the reader to a recent paper [15], where we obtain a similar to Corollary 9.7, 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.*

10. Fully discrete Crank-Nicolson method. We now describe a full discretization of Crank-Nicolson type for the abstract parabolic problem (1.1). To this end, we apply a spatial discretization of type (7.2) to the semidiscrete problem (5.1) as follows. A finite-element space $V_h^j \subset C(\bar{\Omega})$ and a computed solution $u_h^j \in \mathring{V}_h^j := V_h^j \cap H_0^1(\Omega)$ are associated with the time level t_j, while a computed solution $\hat{u}_h^{j-1} \in H_0^1(\Omega)$ is associated with the time level t_{j-1}^+ (this is indicated by the hat notation). We require, for $j = 1, \ldots, M$, that

$$\mathcal{P}_h^j[\delta_t^* u_h^j] + \tfrac{1}{2}\big(\hat{\mathcal{L}}_h^{j-1} \hat{u}_h^{j-1} + \mathcal{L}_h^j u_h^j\big) + \tfrac{1}{2}\mathcal{P}_h^j[f(\cdot, t_{j-1}, \hat{u}_h^{j-1}) + f(\cdot, t_j, u_h^j)] = 0 \quad (10.1a)$$

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), and some $u_h^0 \approx \varphi$ in $\mathring{V}_h^0$. Note that $\hat{\mathcal{L}}_h^{j-1} \hat{u}_h^{j-1} \in \mathring{V}_h^j - I_h[f(\cdot, t_{j-1}, 0)]$ and $\mathcal{L}_h^j u_h^j \in \mathring{V}_h^j - I_h[f(\cdot, t_j, 0)]$, while for any $w \in H_0^1(\Omega)$ one has $\mathring{V}_h^j - I_h[f(\cdot, t_k, 0)] = \mathring{V}_h^j - I_h[f(\cdot, t_k, w)]$. As we also have $\delta_t^* u_h^j \in H_0^1(\Omega)$, so the definition (10.1a) is consistent.

The term $\delta_t^* u_h^j$ in (10.1a) approximates $\partial_t u$ and is identical with (9.1b):

$$\delta_t^* u_h^j := \frac{u_h^j - \hat{u}_h^{j-1}}{\tau_j}, \quad \text{where } \hat{u}_h^0 := u_h^0. \quad (10.1b)$$

The operator δ_t^* is identical with δ_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).

10.1. A posteriori error estimate using piecewise-linear elliptic reconstructions. To estimate the error of the fully discrete Crank-Nicolson method (10.1), set $\mathcal{A} := \{0^+, 1\}$ and recall the elliptic reconstructions $\hat{R}^{j-1}$ and R^j defined for $j = 1, \ldots, M$ 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 ψ_h^j that we now define by

$$\hat{\psi}_h^{j-1} := -\psi_h^j - 2\delta_t^* u_h^j, \qquad \psi_h^j := \mathcal{L}_h(t_j)\, u_h^j + \mathcal{P}_h^j[f(\cdot, t_j, u_h^j)]. \quad (10.2)$$

Note that the first relation here yields

$$\delta_t^* u_h^j + \tfrac{1}{2}\big(\hat\psi_h^{j-1} + \psi_h^j\big) = 0. \tag{10.3}$$

REMARK 10.1. *The definition of ψ_h^j in (10.2) implies $\psi_h^j \in \mathring{V}_h^j$ so $\mathcal{P}_h^j\psi_h^j = \psi_h^j$ so ψ_h^j satisfies (8.4b). Next, $\hat\psi_h^{j-1}$ of (10.2) 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 (10.1a), 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).*

REMARK 10.2. *Remark 10.1 implies that $\hat R^{j-1}$ and R^j satisfy (8.9).*

REMARK 10.3. *Theorem 10.4 and further results of this section remain valid for any pair $\hat\psi_h^{j-1}$, ψ_h^j that satisfy (8.4) and (10.3). For example, alternatively to the definition of ψ_h^j in (10.2), 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 ψ_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

$$I_{1,t}^* w(t) = \tfrac{t_j - t}{\tau_j}\, w(t_{j-1}^+) + \tfrac{t - t_{j-1}}{\tau_j}\, w(t_j) \quad \text{for } t \in (t_{j-1}, t_j], \quad j = 1, \dots, M. \tag{10.4}$$

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

$$\tilde R(\cdot, t) = \tfrac{t_j - t}{\tau_j}\, \hat R^{j-1} + \tfrac{t - t_{j-1}}{\tau_j}\, R^j \ \text{ for } t \in (t_{j-1}, t_j],\ \ j = 1, \dots, M,\ \ \tilde R(\cdot, 0) = \hat R^0. \tag{10.5}$$

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

THEOREM 10.4. *Let u solve the problem (1.1), (1.2) with a parabolic operator $\mathcal{M}$ satisfying Condition 2.1, u_h^j solve the discrete problem (10.1). Then for $m = 1, \dots, M$, one has (8.3) with $\chi_h^j = \tau_j\,\big(\psi_h^j - \hat\psi_h^{j-1}\big)$ using ψ^j and $\hat\psi_h^{j-1}$ from (10.2), η^j from (8.5) with $\mathcal{A} = \{0^+, 1\}$, $C_1 = \frac{1}{8}$, $C_2 = \frac{5}{8}$, $C_1^* = 2$, $C_2^* = 4$, and ϑ_h defined by*

$$\vartheta_h := \tilde\psi_R - I_{1,t}^*\tilde\psi_R, \qquad \tilde\psi_R := \mathcal{L}(t)\tilde R + f(\cdot, t, \tilde R) \tag{10.6}$$

for $t \in [0, T]$, with $I_{1,t}^*$ and $\tilde R$ from (10.4) and (10.5).

Proof. As Remark 10.2 gives $\|R^j - u_h^j\|_{\infty,\Omega} \le \eta^j$, 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 $C_2^*\kappa_0 = 4\kappa_0$, and then apply the triangle inequality. So we consider $\tilde R - u$ only.

We partially imitate the proof of Theorem 5.1. Let $t \in (t_{j-1}, t_j]$. In view of (10.4), for any left-continuous function $w = w(t)$ with the notation $w^j := w(t_j)$, $\hat w^{j-1} := w(t_{j-1}^+)$, $\delta_t^* w^j := (w^j - \hat w^{j-1})/\tau_j$, one has $I_{1,t}^* w(t) = \frac{1}{2}(\hat w^{j-1} + w^j) + (t - t_{j-1/2})\,\delta_t^* w^j$. Note that with our definition of $\tilde R$, the functions in (10.6) are identical with those in (8.10) (using $p = 1$), so we also enjoy the observation (8.11), which can be rewritten as $\tilde\psi_R(\cdot, t_{j-1}^+) = \hat\psi_h^{j-1}$ and $\tilde\psi_R(\cdot, t_j) = \psi_h^j$. Now, for $t \in [t_{j-1}, t_j]$, one easily gets

$$\begin{aligned}
\tilde\psi_R - \tfrac{1}{2}\big(\hat\psi_h^{j-1} + \psi_h^j\big) &= \big\{I_{1,t}\tilde\psi_R - \tfrac{1}{2}\big(\hat\psi_h^{j-1} + \psi_h^j\big)\big\} + \vartheta_h \\
&= (t - t_{j-1/2})\,\delta_t^*\psi_h^j + \vartheta_h \\
&= (t - t_{j-1/2})\,\tau_j^{-2}\chi_h^j + \vartheta_h.
\end{aligned}$$

Combining this with (10.3), one deduces that

$$\begin{aligned}\delta_t^* u_h^j + \tilde{\psi}_R &= (t - t_{j-1/2})\,\tau_j^{-2}\chi_h^j + \vartheta_h \\ &= \partial_t \mu_h + \vartheta_h \qquad\qquad \text{for } t \in (t_{j-1}, t_j],\end{aligned} \tag{10.7}$$

where $\mu_h = \mu_h(x,t)$ is a continuous function defined by

$$\mu_h(\cdot,t) := -\tfrac{1}{2}(t_j - t)(t - t_{j-1}) \cdot \tau_j^{-2}\chi_h^j \qquad \text{for } t \in [t_{j-1}, t_j]. \tag{10.8}$$

This is easily checked by using the relation $\frac{d}{dt}[-\frac{1}{2}(t_j - t)(t - t_{j-1})] = t - t_{j-1/2}$ to evaluate $\partial_t \mu_h$.

Next, we invoke $I_{1,t}^* u_h$ defined by (9.10), for which we have (9.11) and (9.12). As $\mathcal{M}\tilde{R} = \partial_t \tilde{R} + \tilde{\psi}_R$ and $\mathcal{M}u = 0$, so (10.7) yields

$$\mathcal{M}\tilde{R} - \mathcal{M}u = \partial_t(\tilde{R} - I_{1,t}^* u_h) + \partial_t \mu_h + [\vartheta_h + \vartheta_*] \quad \text{in } Q, \tag{10.9}$$

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 $C_2^*\kappa_0 = 4\kappa_0$, is obtained by an application of Lemma 2.4 to the equation (10.9) with $\mu := \mu_R + \mu_h := (\tilde{R} - I_{1,t}^* u_h) + \mu_h$ and $\vartheta := \vartheta_h + \vartheta_*$, using (9.12) 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

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

where we used $\mu_R = \tilde{R} - I_{1,t}^* u_h = I_{1,t}^*(R - u_h)$ combined with Remark 10.2, and (10.8). Note also that $\frac{1}{8} + \frac{1}{2} = \frac{5}{8}$ and $2\eta^j + 2\eta^j = 4\eta^j$. This completes the proof. □

10.2. Applications to the model problem. 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 $u_h^j \in \mathring{V}_h^j$ such that

$$\varepsilon^2 \big\langle \tfrac{1}{2}\nabla(\hat{u}_h^{j-1} + u_h^j), \nabla w_h \big\rangle + \big\langle \tfrac{1}{2}[f(\cdot, t_{j-1}, \hat{u}_h^{j-1}) + f(\cdot, t_j, u_h^j)] + \delta_t^* u_h^j,\, w_h \big\rangle_h = 0, \tag{10.10}$$

$\forall\, w_h \in \mathring{V}_h^j$, 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 $\langle\cdot,\cdot\rangle$, and $\delta_t^* u_h^j$ is defined by (10.1b), (8.2).

Note that the full discretization (10.10) is of type (10.1) with $\hat{\mathcal{L}}_h^{j-1} := \mathcal{L}_h^j$. 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 7.1 and 7.2 only using V_h^j instead of V_h.

10.2.1. Model problem (1.3): regular regime. Let u solve problem (1.3) with $\varepsilon = 1$, $\gamma \ge 0$, posed in a bounded polyhedral spatial domain $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, and u_h^j solve the discrete problem (10.10) with V_h^j and $\langle\cdot,\cdot\rangle_h$ defined, for each time level t_j, as in §9.2.1.

COROLLARY 10.5. *Let the above numerical method be applied to problem (1.3) with $\varepsilon = 1$, $\gamma \ge 0$. Then the a posteriori error estimate of Theorem 10.4 is valid with ϑ_h simplified to (8.12), where $p = 1$, and (8.5b) using $\eta := \eta_0$ with η_0 from (7.6).*

10.2.2. 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 (10.10) with V_h^j and $\langle\cdot,\cdot\rangle_h$ defined, for each time level t_j, as in §9.2.2. We consider the two choices (7.7a) and (7.7b) of $\langle\cdot,\cdot\rangle_h$, using nodal piecewise-linear interpolation $I_h := I_h^j$ onto V_h^j.

COROLLARY 10.6. *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 10.4 is valid with ϑ_h simplified to (8.12), where $p = 1$. The definition (8.5b) of η^{j+1} uses $\eta := \eta_\varepsilon$ for (7.7a), and $\eta := \eta_{\varepsilon;\mathrm{l.m.}}$ for (7.7b), while η_ε and $\eta_{\varepsilon;\mathrm{l.m.}}$ are defined in (7.9a) and (7.9b), respectively, in which I_h is now replaced by I_h^j.*

11. Fully discrete discontinuous Galerkin method dG(1). We now describe a full discretization of dG(1) type for the abstract parabolic problem (1.1). To this end, we apply a spatial discretization of type (7.2) to the semidiscrete problem (6.2), (6.4) as follows. A finite-element space $V_h^{j+1} \subset C(\bar\Omega)$ and a computed solution $u_h^{j+1} \in \mathring{V}_h^{j+1} := V_h^{j+1} \cap H_0^1(\Omega)$ are associated with the time level t_{j+1}, while auxiliary computed solutions $\hat u_h^j \in H_0^1(\Omega)$, $\hat y_h^j \in \mathring{V}_h^{j+1}$ are associated with the time level t_j^+ (this is indicated by the hat notation; note possible choices (8.2) for $\hat u_h^j$).

Imitating (6.2) (and using $\hat y_h^j$ and u_h^{j+1} as discrete analogues of U^{j+1} and y^j), let

$$u_h := \hat y_h^j\,\phi_j + u_h^{j+1}\phi_{j+1}, \quad \Psi := \mathcal{L}_h(t)\,u_h + \mathcal{P}_h^{j+1}\big[f(\cdot,t,u_h)\big] \quad \text{for } t \in (t_j, t_{j+1}], \tag{11.1a}$$

with $\mathcal{L}_h(t)$ and $\mathcal{P}_h^{j+1}$ subject to (8.1). Note that u_h vanishes on $\partial\Omega$, so $I_h[f(\cdot,t_k,0)] = I_h[f(\cdot,t_k,u_h)]$ on $\partial\Omega$, so $\Psi \in \mathring{V}_h^{j+1}$ for $t \in (t_j, t_{j+1}]$.

Now, imitating (6.4), we require, for $j = 0,\dots,M-1$, that

$$\mathcal{P}_h^{j+1}\big[u_h^{j+1} - \hat u_h^j\big] + \tfrac14\,\tau_{j+1}\big(3\Psi(\cdot,t_{j+1/3}) + \Psi(\cdot,t_{j+1})\big) = 0, \tag{11.1b}$$

$$\mathcal{P}_h^{j+1}\big[\hat y_h^j - \hat u_h^j\big] + \tfrac14\,\tau_{j+1}\big(\Psi(\cdot,t_{j+1/3}) - \Psi(\cdot,t_{j+1})\big) = 0, \tag{11.1c}$$

where $\Psi(\cdot,t_{j+1/3})$ and $\Psi(\cdot,t_{j+1})$ are computed using $u_h(\cdot,t_{j+1/3}) = \frac23\hat y_h^j + \frac13 u_h^{j+1}$ and $u_h(\cdot,t_j) = u_h^{j+1}$, respectively, as well as $\mathcal{L}_h(t_{j+1/3})$ and $\mathcal{L}_h(t_{j+1})$, by virtue of (11.1a).

11.1. A posteriori error estimate using piecewise-quadratic elliptic reconstructions. To estimate the error of the fully discrete dG(1) method (11.1), we partially imitate the arguments of Section 6.1 for the related semidiscrete method. First, set $\mathcal{A} := \{0^+, \frac13, 1\}$ 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 ψ_h^{j+1} that we now define by

$$\psi_h^{j+1/3} := -\frac{u_h^{j+1} - \hat u_h^j}{\tau_{j+1}} - \frac{\hat y_h^j - \hat u_h^j}{\tau_{j+1}}, \quad \psi_h^{j+1} := -\frac{u_h^{j+1} - \hat u_h^j}{\tau_{j+1}} + 3\,\frac{\hat y_h^j - \hat u_h^j}{\tau_{j+1}}, \tag{11.2a}$$

$$\hat\psi_h^j := \mathcal{L}_h(t_j^+)\,\hat u_h^j + \mathcal{P}_h^{j+1}[f(\cdot,t_j,\hat u_h^j)]. \tag{11.2b}$$

Note that (11.2a) is a discrete version of (6.13). Furthermore, (11.2a) implies a version of (11.1):

$$u_h^{j+1} - \hat u_h^j + \tfrac14\,\tau_{j+1}\big(3\psi_h^{j+1/3} + \psi_h^{j+1}\big) = 0, \tag{11.3a}$$

$$\hat y_h^j - \hat u_h^j + \tfrac14\,\tau_{j+1}\big(\psi_h^{j+1/3} - \psi_h^{j+1}\big) = 0. \tag{11.3b}$$

In fact, if $\hat{u}_h^j \in \mathring{V}_h^{j+1}$ (Case A of (8.2a)), then (11.3) and (11.1) are equivalent (and one has $\Psi(\cdot,t_{j+\alpha}) = \psi_h^{j+\alpha}$ for $\alpha = \frac{1}{3}$, 1).

REMARK 11.1. *A comparison of (11.3) and (11.1) implies* $\Psi(\cdot,t_{j+\alpha}) = \mathcal{P}_h^{j+1}\psi_h^{j+\alpha}$ *for* $\alpha = \frac{1}{3}$, 1*. Recalling the definition of* Ψ *in (11.1a), one concludes that* $\psi_h^{j+1/3}$ *and* ψ_h^{j+1} *satisfy (8.4b). Next,* $\hat{\psi}_h^j$ *of (11.2b) is in* $\mathring{V}_h^{j+1}$*, so* $\hat{\psi}_h^j$ *satisfies (8.4a).*

REMARK 11.2. *Remark 11.1 implies that* $\hat{R}^j$*,* $R^{j+1/3}$ *and* R^{j+1} *satisfy (8.9) with* $\mathcal{A} := \{0^+, \frac{1}{3}, 1\}$ *and* η^{j+1} *of (8.5b) using* $\hat{u}_h^j$*,* $u_h^{j+1/3} = \frac{2}{3}\hat{y}_h^j + \frac{1}{3}u_h^{j+1}$ *and* u_h^{j+1}*.*

REMARK 11.3. *Theorem 11.5 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 (11.3). For example, one can replace* $\mathcal{P}_h^{j+1}[f(\cdot,t_j,\hat{u}_h^j)]$ *in (11.2b) 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 11.4 (Case B). *In case (8.2b) with* $\hat{u}_h^j := u_h^j$*, it is more natural to replace (11.2b) 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* $\eta\big(V_h^{j+1}, \hat{u}_h^j, \hat{g}^j(\cdot,\hat{u}_h^j)\big)$ *in the definition (8.5b) of* η^{j+1} *by* $\eta\big(V_h^j, u_h^j, g^j(\cdot,u_h^j)\big)$*. Consequently, Theorem 11.5 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

$$I^*_{2,t}w(0) := w(0), \quad I^*_{2,t}w(t) := w(t_j^+)\,\phi_j + w(t_{j+1})\,\phi_{j+1} - \tfrac{1}{2}\,\phi_j\,\phi_{j+1}\,W^{j+1},$$
$$\text{where} \quad W^{j+1} := 3\,[2w(t_j^+) - 3w(t_{j+1/3}) + w(t_{j+1})], \tag{11.4}$$

for $t \in (t_j, t_{j+1}]$, $j = 0, \ldots, M-1$. 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

$$\tilde{R}(\cdot,0) := \hat{R}^0 = R^0, \qquad \tilde{R}(\cdot,t) := I^*_{2,t}\{\hat{R}^j, R^{j+1/3}, R^{j+1}\} \quad \text{for } t \in (0,T]. \tag{11.5}$$

Finally, similarly to (6.5), define a piecewise-quadratic computed solution in time by $\tilde{u}_h(\cdot,0) := \hat{u}_h^0 = u_h^0$ and

$$\tilde{u}_h := \hat{u}_h^j\phi_j + u_h^{j+1}\phi_{j+1} + \nu_h, \qquad \nu_h := 3\,\phi_j\,\phi_{j+1}\,\{\hat{y}_h^j - \hat{u}_h^j\} \qquad \text{for } t \in (t_j, t_{j+1}]. \tag{11.6}$$

Note that, similarly to (6.6),

$$\tilde{u}_h(\cdot,t_{j+1/3}) = u_h(\cdot,t_{j+1/3}) = \tfrac{2}{3}\,\hat{y}_h^j + \tfrac{1}{3}\,u_h^{j+1}, \;\; \tilde{u}_h(\cdot,t_{j+1}) = u_h(\cdot,t_{j+1}) = u_h^{j+1}. \tag{11.7}$$

Consequently, $\tilde{u}_h$ can be obtained by applying the same interpolation $I^*_{2,t}$ to $\hat{u}_h^j$, $\frac{2}{3}\,\hat{y}_h^j + \frac{1}{3}\,u_h^{j+1}$ and u_h^{j+1} associated with the time levels t_j^+, $t_{j+1/3}$ and t_{j+1}.

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

THEOREM 11.5. *Let* u *solve the problem (1.1), (1.2) satisfying Condition 2.1,* u_h^j *solve the discrete problem (11.1) with any* $\hat{u}_h^j \in H_0^1(\Omega)$*. Then one has (8.3) with*

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

using $\hat\psi^j$, $\psi_h^{j+1/3}$ *and* ψ_h^{j+1} *of (11.2),* η^j *from (8.5) with* $\mathcal{A}=\{0^+,\frac13,1\}$ *using* $\hat u_h^j$, $u_h^{j+1/3}=\frac23\hat y_h^j+\frac13 u_h^{j+1}$ *and* u_h^{j+1}*, the constants* $C_1=\frac{2}{81}$, $C_2=\frac{1}{18}$, $C_1^*=\frac53$, $C_2^*=10$, *and* ϑ_h *defined by*

$$\vartheta_h := \tilde\psi_R - I^*_{2,t}\tilde\psi_R, \qquad \tilde\psi_R := \mathcal{L}(t)\tilde R + f(\cdot,t,\tilde R) \tag{11.9}$$

for $t\in[0,T]$*, with* $I^*_{2,t}$ *and* $\tilde R$ *from (11.4) and (11.5).*

Proof. As Remark 11.2 gives $\|R^j-u_h^j\|_{\infty,\Omega}\le\eta^j$, 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 $C_2^*\kappa_0=10\kappa_0$, and then apply the triangle inequality. So we consider $\tilde R-u$ only.

We partially imitate the proof of Theorem 6.4. Let $t\in(t_j,t_{j+1}]$. It is convenient to treat the left-continuous function $\tilde u_h$ of (11.6) as being discontinuous at t_j^+ rather than at t_j. Now, combining (11.3a) with $\partial_t(\hat u_h^j\phi_j+u_h^{j+1}\phi_{j+1})=(u_h^{j+1}-\hat u_h^j)/\tau_{j+1}$, for $\tilde u_h$ one gets

$$\partial_t\tilde u_h = \partial_t\nu_h - \tfrac14\bigl(3\psi_h^{j+1/3}+\psi_h^{j+1}\bigr)+\vartheta_* \qquad \text{for } t\in(t_j,t_{j+1}]. \tag{11.10}$$

Here the discontinuity of $\tilde u_h$ at t_j^+ yielded the term

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

which is identical with ϑ_* of (9.11) and so satisfies (9.12).

Next, note that with our definition of $\tilde R$, the functions in (11.9) are identical with those in (8.10) (using $p=2$), so we also enjoy the observation (8.11), which can be rewritten as $\tilde\psi_R(\cdot,t_j^+)=\hat\psi_h^j$ and $\tilde\psi_R(\cdot,t_{j+\alpha})=\psi_h^{j+\alpha}$ for $\alpha=\frac13$, 1. Now, a comparison of χ_h^j in (11.8) with the definition (11.4) 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, $I^*_{2,t}\tilde\psi_R$ allows the representation (compare with (6.8), (6.9)):

$$I^*_{2,t}\tilde\psi_R = \psi_h^{j+1} - \tfrac32\{\psi_h^{j+1}-\psi_h^{j+1/3}\}\,\phi_j - \tfrac12\,\tau_{j+1}^{-2}(t-t_{j+1/3})\,\phi_{j+1}\,\chi_h^{j+1}. \tag{11.12}$$

So imitating the derivation of (6.11), from (11.12) for $t\in(t_j,t_{j+1}]$ one gets

$$I^*_{2,t}\tilde\psi_R - \tfrac14\bigl(3\psi_h^{j+1/3}+\psi_h^{j+1}\bigr) = \partial_t[\mu_h-\nu_h], \qquad \mu_h := \tfrac16\,\phi_j^2\,\phi_{j+1}\,\chi_h^{j+1}. \tag{11.13}$$

For this function μ_h, similarly to (6.12), a calculation shows that

$$|\mu_h|\le\tfrac{2}{81}|\chi_h^{j+1}|, \quad \tau_{j+1}\,|\partial_t\mu_h|\le\tfrac{1}{18}|\chi_h^{j+1}| \quad \text{for } t\in(t_j,t_{j+1}]. \tag{11.14}$$

Now, combining (11.10) with (11.13), we arrive at

$$\partial_t\tilde u_h = \partial_t\mu_h - I^*_{2,t}\tilde\psi_R + \vartheta_* \qquad \text{for } (x,t)\in Q.$$

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

$$\mathcal{M}\tilde R-\mathcal{M}u = \partial_t[\tilde R-\tilde u_h+\mu_h]+\vartheta_h+\vartheta_* \qquad \text{for } (x,t)\in Q.$$

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 2.4

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 μ_h, we recall (11.14) 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 11.2, $|[\tilde{R}-\tilde{u}_h](\cdot,t_{m-1}^+)| \le \eta^m$, and a calculation yields

$$|\tilde{R}-\tilde{u}_h| \le \tfrac{5}{3}\eta^{j+1}, \qquad \tau_{j+1}|\partial_t(\tilde{R}-\tilde{u}_h)| \le 9\eta^{j+1} \qquad \text{for } t \in (t_j,t_{j+1}]. \tag{11.15}$$

Finally, for ϑ_*, we invoke (11.11). Combining these observations in the application of Lemma 2.4 completes the proof. □

REMARK 11.6 (dG(r) for $r > 1$). *The results of the section, including the a posteriori error estimate of Theorem 11.5 can be generalized to a fully discrete dG(r) method with Radau quadrature for $r > 1$ (in lines with the analysis of §6.3 for a semidiscrete dG(r) method). In fact, then the error estimate (8.3) will involve χ_h^{j+1} defined by (6.22) with $\psi^{j+\alpha}$ replaced by $\psi_h^{j+\alpha}$, and the same constants C_1 and C_2 as in Theorem 6.8.*

11.2. Application to a general t-independent operator $\mathcal{L}$. Let the coefficients of the elliptic operator $\mathcal{L}$ be independent of the variable t; as in Section 6.2, we highlight this case by using the special notation $\mathring{\mathcal{L}} := \mathcal{L}$ for this operator. Note that its discrete counterpart $\mathcal{L}_h^{j+1}(t) : H_0^1(\Omega) \to \mathring{V}_h^{j+1} - I_h^{j+1}[f(\cdot,t,0)]$ of (8.1) remains dependent on t (and it is not linear), so it is convenient to also use its linear t-independent version $\mathring{\mathcal{L}}_h^{j+1}$ and the related version $\mathring{\mathcal{P}}^{j+1}$ of $\mathcal{P}^{j+1}$ defined by

$$\begin{aligned} &\mathring{\mathcal{L}}_h^{j+1} : H_0^1(\Omega) \to \mathring{V}_h^{j+1}, \qquad \mathring{\mathcal{P}}_h^{j+1}\mathcal{L}_h(t) = \mathring{\mathcal{L}}_h^{j+1} \ \text{ for } t \in (t_j,t_{j+1}],\\ &\mathring{\mathcal{P}}_h^{j+1} : C(\bar{\Omega}) \to \mathring{V}_h^{j+1}, \qquad \mathring{\mathcal{P}}_h^{j+1}v = \mathcal{P}_h^{j+1}v \quad \forall\, v \in H_0^1(\Omega)\cup C(\bar{\Omega}). \end{aligned} \tag{11.16}$$

Note that the function Ψ of (11.1a) is in $\mathring{V}_h^{j+1}$, so it can now be represented for $t \in (t_j,t_{j+1}]$ as $\Psi = \mathring{\mathcal{L}}_h^{j+1}\, u_h + \mathring{\mathcal{P}}_h^{j+1}\big[f(\cdot,t,u_h)\big]$. Consequently, the fully discrete method (11.1) can be rewritten (similarly to (6.14)) as

$$\begin{aligned} \mathring{\mathcal{P}}_h^{j+1}\big(u_h^{j+1}-\hat{u}_h^j\big) + \tfrac12\,\tau_{j+1}\,\mathring{\mathcal{L}}_h^{j+1}\big(\hat{y}_h^j + u_h^{j+1}\big) + \tfrac14\,\tau_{j+1}\,\mathring{\mathcal{P}}_h^{j+1}\big(3f_h^{j+1/3}+f_h^{j+1}\big) &= 0,\\ \mathring{\mathcal{P}}_h^{j+1}\big(\hat{y}_h^j-\hat{u}_h^j\big) + \tfrac16\,\tau_{j+1}\,\mathring{\mathcal{L}}_h^{j+1}\big(\hat{y}_h^j - u_h^{j+1}\big) + \tfrac14\,\tau_{j+1}\,\mathring{\mathcal{P}}_h^{j+1}\big(f_h^{j+1/3}-f_h^{j+1}\big) &= 0, \end{aligned} \tag{11.17a}$$

with the notation

$$f_h^{j+1/3} := f(\cdot,t_{j+1/3},\tfrac23\,\hat{y}_h^j + \tfrac13\,u_h^{j+1}), \qquad f_h^{j+1} := f(\cdot,t_j,u_h^{j+1}). \tag{11.17b}$$

Furthermore, we have a version of (6.16) for χ_h^{j+1} of (11.8) in the case $\hat{u}_h^j \in \mathring{V}_h^{j+1}$:

$$\chi_h^{j+1} = 6\tau_{j+1}\,\mathring{\mathcal{L}}_h^{j+1}[\hat{u}_h^j - \hat{y}_h^j] + 3\tau_{j+1}\,\mathring{\mathcal{P}}_h^{j+1}[2f(\cdot,t_j,\hat{u}_h^j) - 3f_h^{j+1/3} + f_h^{j+1}]. \tag{11.18}$$

11.3. Application to the model problem (1.3). Consider a fully discrete discontinuous Galerkin method dG(1) for (1.3), obtained by applying the spatial discretization (7.5) to the semidiscrete problem (6.4), (6.2): Find $\hat{y}_h^j$, $u_h^{j+1} \in \mathring{V}_h^{j+1}$ such that

$$\begin{aligned} \Big\langle \frac{u_h^{j+1}-\hat{u}_h^j}{\tau_{j+1}},\, w_h\Big\rangle_h + \varepsilon^2\Big\langle \tfrac12\nabla(\hat{y}_h^j + u_h^{j+1}),\nabla w_h\Big\rangle + \tfrac14\Big\langle 3f_h^{j+1/3}+f_h^{j+1},\, w_h\Big\rangle_h &= 0,\\ \Big\langle \frac{\hat{y}_h^j-\hat{u}_h^j}{\tau_{j+1}},\, w_h\Big\rangle_h + \varepsilon^2\Big\langle \tfrac16\nabla(\hat{y}_h^j - u_h^{j+1}),\nabla w_h\Big\rangle + \tfrac14\Big\langle f_h^{j+1/3}-f_h^{j+1},\, w_h\Big\rangle_h &= 0, \end{aligned} \tag{11.19}$$

$\forall\, w_h \in \mathring{V}_h^{j+1}$, with the notation (11.17b) (compare with (11.17a)). 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 that the full discretization (11.19) is of type (11.1). 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 7.1 and 7.2 only using V_h^{j+1} instead of V_h.

Note also that the elliptic operator $\mathcal{L} = -\varepsilon^2\triangle$ in (1.3) is t-independent, so all the observations of Section 11.2 apply to this problem.

11.3.1. 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 (10.10) with V_h^j and $\langle\cdot,\cdot\rangle_h$ defined, for each time level t_j, as in §9.2.1.

Corollary 11.7. *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 11.5 is valid with ϑ_h simplified to (8.12), where $p = 2$, and (8.5b) using $\eta := \eta_0$ with η_0 from (7.6).*

Remark 11.8. *A discontinuous Galerkin method dG(1) for a linear version of (1.3) with $\varepsilon = 1$ was considered in [8, 3]. In this particular case, $f = f(x,t)$ implies that (11.18) can be rewritten as $\chi_h^{j+1} = 6\tau_{j+1}\,\mathring{\mathcal{L}}_h^{j+1}[\hat{u}_h^j - \hat{y}_h^j] + \tau_{j+1}^3\partial_t^2 f(x,t')$ for some intermediate $t' \in [t_j, t_{j+1}]$. With this simplification, the a posteriori error estimate of Corollary 11.7 resembles (but is not identical with) the one of [8, (1.14)] in that it involves terms of type $\tau_{j+1}\|\mathring{\mathcal{L}}_h^{j+1}[\hat{u}_h^j - \hat{y}_h^j]\|_{\infty,\Omega}$ and $\tau_{j+1}^3\|\partial_t^2 f\|_{\infty,\Omega}$. (Note also that [8, (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 $\|\hat{u}_h^j - \hat{y}_h^j\|_{\infty,\Omega} = \mathcal{O}(\tau_{j+1}^2)$. We also note the paper [21], which gives a posteriori estimates for discontinuous Galerkin time discretizations in other norms.*

11.3.2. 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 (10.10) with V_h^j and $\langle\cdot,\cdot\rangle_h$ defined, for each time level t_j, as in §9.2.2. We consider the two choices (7.7a) and (7.7b) of $\langle\cdot,\cdot\rangle_h$, using nodal piecewise-linear interpolation $I_h := I_h^j$ onto V_h^j.

Corollary 11.9. *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 11.5 is valid ϑ_h simplified to (8.12), where $p = 2$. The definition (8.5b) of η^{j+1} uses $\eta := \eta_\varepsilon$ for (7.7a), and $\eta := \eta_{\varepsilon;\,\mathrm{l.m.}}$ for (7.7b), while η_ε and $\eta_{\varepsilon;\,\mathrm{l.m.}}$ are defined in (7.9a) and (7.9b), respectively, 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 Ω and the Green's function $\hat{\mathcal{G}}$ for the related problem $\hat{\mathcal{M}}\hat{u} := \partial_t\hat{u} - \triangle\hat{u} + f(x/\varepsilon, t, \hat{u}) = 0$ 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 κ_0, κ_1 and κ_2 independent of $|\Omega|$, so throughout the proof we set $\mathcal{L}^* = -\triangle$ in (2.2a).

(i) We start by proving the first bound in Condition 2.1. The Green's function $\bar{\mathcal{G}}$ associated with $\bar{\mathcal{M}} := \partial_t - \triangle + \gamma^2$ 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 [24, §III.3], [10, §2.3.1]. So one gets

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

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

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

As $\int_{\mathbb{R}^n} \psi(\zeta)\,\mathrm{d}\zeta = 1$, we immediately get $\|\bar{\mathcal{G}}(x,t;\cdot,s)\|_{1,\Omega} \le 1$, 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 $0 \le \mathcal{G} \le \bar{\mathcal{G}}$, an application of [5, Corollary 5] with $\beta = 2$, $\gamma = 1$, $c_1 = \frac{1}{4}$, $c_2 = \frac{4}{9}c_1$ and $\alpha(t) = \frac{\mathrm{e}^{-\gamma^2 t}}{(4\pi t)^{n/2}}$ yields $|\partial_s \mathcal{G}(x,t;\xi,s)| \le 18 c_1 c_2\,(t-s)^{-1}\,\alpha(\frac{1}{2}[t-s])\,\mathrm{e}^{-(c_2/c_1)|\zeta|^2}$, where ζ is chosen as in part (i) of this proof. Now an observation similar to (12.2) leads to the estimate $\|\partial_s \mathcal{G}(x,t;\cdot,s)\|_{1,\Omega} \le \kappa_1 (t-s)^{-1}\mathrm{e}^{-\frac{1}{2}\gamma^2(t-s)}$, 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 $\gamma^2 \le a(\xi,s) \le \bar{\gamma}^2$. 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 $\partial_t - \triangle + \gamma^2$ in the domain Ω so $\hat{\Gamma}(\xi,s)$ satisfies a version of (2.2) with a replaced by γ^2. Comparing this problem with the problem (2.2) for Γ and noting that $\mathcal{L} = \mathcal{L}^* = \triangle$, 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

$$[-\partial_s - \triangle + \gamma^2\,]\,v(\xi,s) = F(\xi,s) \quad \text{for} \quad (\xi,s) \in \Omega \times [0,t), \tag{12.3a}$$

$$v(\xi,t) = 0 \quad \text{for} \quad \xi \in \Omega, \tag{12.3b}$$

$$v(\xi,s) = 0 \quad \text{for} \quad (\xi,s) \in \partial\Omega \times [0,t], \tag{12.3c}$$

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

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

Note that in part (ii) we have shown that $\hat{\Gamma}$ satisfies the second bound in Condition 2.1 with $\kappa_2 = 0$. So 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). □

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 $[\partial_t - \triangle + \gamma^2]\,v = F$ in Q and vanish for $t = 0$ and $x \in \partial\Omega$, where $0 \le F(x,t) \le g(x - x_0, t)$ with g from (12.1) and some $x_0 \in \Omega$. Then $\int_0^T \|\partial_t v(\cdot,t)\|_{1,\Omega}\,\mathrm{d}t \le \hat{\kappa}_2$, 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 $F(x,t) \le g(x,t)$. Recall that $\bar{\mathcal{M}} g = 0$ with $\bar{\mathcal{M}} = \partial_t - \triangle + \gamma^2$; this implies that $\bar{M}[tg] = g$, so an application of the maximum principle yields

$$0 \le v(x,t) \le t\,g(x,t). \tag{12.4}$$

(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 $\varrho := t^{\frac{1}{3}}\,\mathrm{e}^{\frac{1}{2}\gamma^2 t}$ so $\varrho' = (\frac{1}{3}t^{-1} + \frac{1}{2}\gamma^2)\,\varrho$. Note that

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

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

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

(so $\hat{\kappa}_3^2 \le 3T^{1/3}$ for $\gamma \ge 0$, and $\int_0^\infty t^{-\frac{2}{3}}\,\mathrm{e}^{-t}\,\mathrm{d}t \approx 2.7$ 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 $\bar{\mathcal{M}}w = \varrho\,F + \varrho'\,v \le \varrho\,g + \varrho'\,v$ and so apply an a priori estimate [16, (6.6) of Chapter III]:

$$\|\partial_t w\|_{2,\Omega\times[0,T]} \le \|\bar{\mathcal{M}}w\|_{2,\Omega\times[0,T]} \tag{12.6}$$

(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 $\varrho'\,v \le (\frac{1}{3} + \frac{1}{2}\gamma^2\,t)\,\varrho\,g$ (which follows from (12.4)), one gets

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

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

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

(this integral is convergent as $\frac{n}{2} - \frac{2}{3} < 1$ for $n \le 3$). 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 non-overlapping subdomains $\Omega_0 := \{|x| < 2\}$ and $\Omega_j := \{2^j < |x| < 2^{j+1}\}$ for $j = 1,\ldots$; furthermore let $\Omega_0' := \Omega$ and $\Omega_j' := \{2^{j-1} < |x| < 2^{j+2}\} \supset \Omega_j$. Note that

$$|\Omega_j|^{\frac{1}{2}} \le c_n\,2^{\frac{1}{2}nj}. \tag{12.8}$$

Now we partially imitate the proof in part (i). First, note that one has the bound (12.5) with Ω replaced by Ω_j for $j = 0, 1, \ldots$. So for $j = 0$, using the results of part (i), one immediately gets

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

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

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

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

with the constant $\bar{C}$ independent of Ω and T (this estimate is obtained similarly to [16, (6.6), (6.11) of Chapter III]). Here $\bar{\mathcal{M}}w$ is estimated as in part (i), while $w = \varrho\,v \le t\varrho\,g$ by (12.4). This yields a local version of (12.7):

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

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

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

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

$$\|\partial_t v\|_{1,(\Omega\cap\Omega_j)\times[0,T]} \le 2\,\hat\kappa_3\,c_n \begin{cases} \hat\kappa_4 & \text{for } j=0,\\ \sqrt{c_n''}\,2^{-\frac12 nj} & \text{for } j\ge 1.\end{cases}$$

This immediately yields the desired bound with $\hat\kappa_2 := 2\,\hat\kappa_3\,c_n[\hat\kappa_4 + \sqrt{c_n''}\,(2^{\frac12 n}-1)^{-1}]$ independent of $|\Omega|$. □

REFERENCES

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

[2] S. Bartels and R. Müller, *Quasi-optimal and robust a posteriori error estimates in L^∞ (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] 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.

[9] 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.

[10] L. C. Evans, *Partial differential equations*, American Mathematical Society, Providence, 1998.

[11] A. Friedman, *Partial differential equations of parabolic type*, Prentice-Hall, Englewood Cliffs, N.J., 1964.

[12] D. H. Griffel, *Applied functional analysis*, Dover Publications, Mineola, NY, 2002.

[13] 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.

[14] 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.

[15] 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.

[16] O. A. Ladyzhenskaya, V. A. Solonnikov, and N. N. Ural'tseva, *Linear and quasi-linear equations of parabolic type*, American Mathematical Society, 1968.

[17] 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.

[18] 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.

[19] T. Linss, *Layer-adapted meshes for reaction-convection-diffusion problems*, vol. 1985 of Lecture Notes in Mathematics, Springer, Berlin, 2010.

[20] 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.

[21] 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.

[22] 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.

[23] H.-G. Roos, M. Stynes and L. Tobiska, *Robust Numerical Methods for Singularly Perturbed Differential Equations*, Springer, Berlin, 2008.

[24] A. N. Tikhonov and A. A. Samarskiĭ, *Equations of mathematical physics*, Dover Publications Inc., New York, 1990. Translated from Russian by A. R. M. Robson and P. Basu.

[25] V. Thomée, *Galerkin finite element methods for parabolic problems*, Springer, Berlin, 2006.