Numer. Math. (2010) 115:475–509
DOI 10.1007/s00211-009-0285-8

Numerische Mathematik

A posteriori error estimation for *hp*-version time-stepping methods for parabolic partial differential equations

Dominik Schötzau · Thomas P. Wihler

Received: 16 January 2008 / Revised: 10 November 2008 / Published online: 13 January 2010

Abstract The aim of this paper is to develop an *hp*-version a posteriori error analysis for the time discretization of parabolic problems by the continuous Galerkin (cG) and the discontinuous Galerkin (dG) time-stepping methods, respectively. The resulting error estimators are fully explicit with respect to the local time-steps and approximation orders. Their performance within an *hp*-adaptive refinement procedure is illustrated with a series of numerical experiments.

Mathematics Subject Classification (2000) 65M60 · 65J10

1 Introduction

Galerkin time-stepping methods for initial-value problems were introduced, for example, in [13,14,17,22]. These methods are based on variational formulations of the initial-value problems and provide piecewise polynomial approximations in time. The approximation can be chosen to be either continuous or discontinuous at the time discretization points, thereby giving rise to the so-called continuous Galerkin (cG) and discontinuous Galerkin (dG) time-stepping methods, respectively. For both approaches, the discrete Galerkin formulations decouple into local problems on each

D. Schötzau was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC).

D. Schötzau
Department of Mathematics, University of British Columbia, 1984 Mathematics Road, Vancouver, BC V6T 1Z2, Canada
e-mail: schoetzau@math.ubc.ca

T. P. Wihler (✉)
Mathematics Institute, Universität Bern, Sidlerstrasse 5, 3012 Bern, Switzerland
e-mail: wihler@math.unibe.ch

time-step, and the discretizations can therefore be understood as implicit one-step schemes. The cG and dG time-stepping methods have been analyzed for ordinary differential equations (ODEs), e.g., in [3,5,10,11,16]. The application of cG and dG approaches to the time discretization of parabolic partial differential equations (PDEs) has been studied in [6–9,15,28] and the references therein.

The variational character of the cG and dG methods allows for arbitrary variation in the size of the time-steps and the local approximation orders. Therefore, they can be extended naturally to hp-version Galerkin schemes; see, e.g., [23,31] for hp-version cG and dG discretizations of initial-value problems for ODEs. The main feature of these hp-methods is their ability to approximate smooth solutions—with possible local singularities—at high algebraic or even exponential rates of converge. In particular, exponential convergence can be achieved in the numerical approximation of problems with start-up singularities; see [24,25,30] for linear parabolic PDEs and [4] for Volterra integro-differential equations with weakly singular kernels. Furthermore, it has been proved in [20,21] that the combination of hp-version time-stepping with suitable wavelet spatial discretizations results in log-linear complexity solution algorithms for nonlocal evolution processes involving pseudo-differential operators. In addition, the discretization of high-dimensional parabolic problems, using sparse grids in space, has been analyzed in [29].

In order to obtain a posteriori error estimates for the cG and dG methods several techniques have been proposed in the literature. For example, a posteriori error estimation based on duality techniques can be found in, e.g., [3,10,16] and the references therein. An alternative approach, applied to dG methods, has been recently presented in [19]. Here, the analysis is based on an appropriate reconstruction of the dG solution which allows the dG variational formulation to be rewritten in strong form, and subsequently, enables the application of natural energy arguments. In this paper, we shall use the h-version approach presented in [19], and extend it to the hp-version cG and dG methods. Specifically, we present a posteriori error estimators for both the hp-version cG and dG schemes which are fully explicit with respect to the size of the time-steps and polynomial degrees. The main novelties of the present analysis are: (1) a complete hp-characterization of the reconstruction operator from [19] and its difference (measured in a suitable norm) to the dG solution, and (2) the application of a similar approach to the hp-version cG scheme. The resulting error estimators are tested within an hp-adaptive algorithm based on local Sobolev regularity estimation, as proposed in [12] for elliptic problems. Our numerical experiments indicate that exponential rates of convergence can be achieved for smooth problems with start-up singularities (as induced, for example, by incompatible initial data).

Let us discuss some notation that will be used throughout the paper: The inner product and associated norm of a Hilbert space V are denoted by $(\cdot,\cdot)_V$ and $\|\cdot\|_V$, respectively. Furthermore, $V^\star$ signifies the dual space of V. Additionally, the duality pairing in $V^\star \times V$ is denoted by $\langle\cdot,\cdot\rangle_{V^\star\times V}$ and the dual norm by $\|\cdot\|_{V^\star}$. For an interval $I=(a,b)$, the space $C^0(\overline{I};V)$ consists of all functions $u:\overline{I}\to V$ that are continuous on $\overline{I}$ with values in V. It is endowed with the standard maximum norm

$$\|u\|_{C^0(\overline{I};V)} = \max_{t\in\overline{I}} \|u(t)\|_V.$$

Moreover, $L^2(I;V)$ signifies the space of (classes of) measurable functions $u: I \to V$ so that $\|u(t)\|_V$ is square-integrable over I (with respect to the Lebesgue measure on I). We notice that $L^2(I;V)$ is a Hilbert space with the inner product

$$(u, v)_{L^2(I;V)} = \int_I (u(t), v(t))_V \, \mathrm{d}t.$$

The norm induced by $(\cdot, \cdot)_{L^2(I;V)}$ is

$$\|u\|_{L^2(I;V)} = \left(\int_I \|u(t)\|_V^2 \, \mathrm{d}t \right)^{\frac{1}{2}}.$$

Moreover, we recall that $L^2(I;V)^\star = L^2(I;V^\star)$. For $u \in L^2(I;V^\star)$ and $v \in L^2(I;V)$, we then have the duality pairing

$$\langle u, v \rangle_{L^2(I;V^\star)\times L^2(I;V)} = \int_I \langle u(t), v(t) \rangle_{V^\star \times V} \, \mathrm{d}t.$$

Let us now introduce the linear parabolic problems considered in this work. This shall be done within an abstract Hilbert space setting. More precisely, let $X \hookrightarrow H$ be two Hilbert spaces with continuous embedding. We further suppose that X is dense in H. Upon identification of H with its dual space $H^\star$, the following Gelfand triple of Hilbert spaces is obtained: $X \hookrightarrow H \cong H^\star \hookrightarrow X^\star$. For $T > 0$ and given data

$$u_0 \in H, \quad g \in L^2((0,T); X^\star),$$

we consider the parabolic problem

$$\begin{aligned} u'(t) + \mathsf{A}u(t) &= g(t), \quad t \in (0,T), \\ u(0) &= u_0. \end{aligned} \tag{1}$$

Here, $\mathsf{A} : X \to X^\star$ is a linear elliptic operator (in space) defined by

$$\langle \mathsf{A}u, v \rangle_{X^\star \times X} = a(u, v) \quad \forall u, v \in X,$$

for a bilinear form $a : X \times X \to \mathbb{R}$ that is assumed to be continuous and coercive. More precisely, there are two constants $\alpha, \beta > 0$ such that

$$|a(u, v)| \leq \alpha \|u\|_X \|v\|_X \qquad \forall u, v \in X, \tag{2}$$

$$a(u, u) \geq \beta \|u\|_X^2 \qquad \forall u \in X. \tag{3}$$

The standard weak formulation of the parabolic problem (1) is to find $u(t)$ such that $u(0) = u_0$ and

$$\langle u'(t), v\rangle_{X^\star \times X} + a(u(t), v) = \langle g(t), v\rangle_{X^\star \times X}$$

for all $v \in X$ and $t \in (0, T)$. Under the above assumptions, this variational problem has a unique weak solution u satisfying

$$u \in L^2((0, T); X) \cap C^0([0, T]; H), \quad u' \in L^2((0, T); X^\star).$$

Additionally, there holds the stability estimate

$$\begin{aligned} &\|u\|_{C^0([0,T];H)} + \|u\|_{L^2((0,T);X)} + \|u'\|_{L^2((0,T);X^\star)} \\ &\leq C \left(\|g\|_{L^2((0,T);X^\star)} + \|u_0\|_H\right), \end{aligned}$$

for a constant $C > 0$ that only depends on α and β in (2) and (3); see [18, Theorem 4.1]. Finally, we denote by $\widehat{u}_0 \in X$ a generic approximation of $u_0 \in H$.

Remark 1 We remark that, if a satisfies a Gårding inequality of the form

$$a(u, u) \geq \beta\|u\|_X^2 - \gamma\|u\|_H^2 \quad \forall u \in X,$$

then the corresponding parabolic problem can be cast into our setting by applying the substitution $u = e^{\gamma t}\widetilde{u}$.

The outline of the rest of the paper is as follows: in Sect. 2, we introduce the *hp*-version time-stepping methods for the parabolic model problem (1). In Sect. 3, we shall present our *hp*-version a posteriori error estimates which constitute the main results of this work. Furthermore, Sect. 4 contains the proof of an *hp*-version approximation property that is crucial in the analysis of dG methods. Additionally, we display a series of numerical results in Sect. 5. Finally, we present some concluding remarks in Sect. 6.

2 Time-stepping methods

In this section, we shall recall and discuss the *hp*-version formulations of the continuous and discontinuous Galerkin time-stepping methods from [23,24] and [31], respectively, for the time (semi-) discretization of the parabolic problem (1).

We consider a partition $\mathcal{M} = \{I_m\}_{m=1}^M$ of the time interval $(0, T)$ into M open subintervals $I_m = (t_{m-1}, t_m)$, $m = 1, 2, \ldots, M$, obtained from a set of nodes

$$0 = t_0 < t_1 < t_2 < \ldots < t_{M-1} < t_M = T.$$

We set $k_m = t_m - t_{m-1}$ and refer to I_m as the m^{th} time-step. Furthermore, to each time-step I_m we assign a polynomial degree $r_m \geq 0$ and store these numbers in the vector $\boldsymbol{r} = \{r_m\}_{m=1}^M$. In the sequel, for an integer ℓ, we write $\boldsymbol{r} \pm \ell$ to denote the degree vector $\{r_m \pm \ell\}_{m=1}^M$. Additionally, we use the notation $\mathcal{P}^r(I; X)$ to denote the space of all polynomials of degree at most $r \in \mathbb{N}_0$ on I with coefficients in X, i.e.,

$$\mathcal{P}^r(I; X) = \left\{ p \in C^0(\overline{I}; X) \,:\, p(t) = \sum_{i=0}^{r} x_i t^i,\ x_i \in X \right\}.$$

Then, for a partition $\mathcal{M}$ and a degree vector $\boldsymbol{r}$, we introduce the discontinuous Galerkin space

$$V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X) = \{ U \in L^2((0,T); X) : U|_{I_m} \in \mathcal{P}^{r_m}(I_m; X),\ 1 \le m \le M \},$$

and the continuous Galerkin space

$$\begin{aligned} V_{\mathrm{cG}}^{\boldsymbol{r}}(\mathcal{M}; X) &= V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X) \cap C^0([0,T]; X) \\ &= \{ U \in C^0([0,T]; X) : U|_{I_m} \in \mathcal{P}^{r_m}(I_m; X),\ 1 \le m \le M \}. \end{aligned}$$

Furthermore, for a piecewise continuous function U, we define the one-sided limits of U in X or H at the node t_m by

$$U_m^+ = \lim_{s \downarrow 0} U(t_m + s), \quad U_m^- = \lim_{s \downarrow 0} U(t_m - s).$$

The jump of U at t_m, $1 \le m \le M-1$, is defined by $[\![U]\!]_m = U_m^+ - U_m^-$. Note that, if U is continuous at t_m, we have $[\![U]\!]_m = 0$.

2.1 Continuous Galerkin time-stepping

For given partition $\mathcal{M}$ and a degree vector $\boldsymbol{r}$, the *hp*-version continuous Galerkin (cG) method for the approximation of (1) is to find $U_{\mathrm{cG}} \in V_{\mathrm{cG}}^{\boldsymbol{r}+1}(\mathcal{M}; X)$ such that $U_{\mathrm{cG}}(0) = \widehat{u}_0 \in X$ and

$$B_{\mathrm{cG}}(U_{\mathrm{cG}}, V) = F_{\mathrm{cG}}(V) \quad \forall V \in V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X). \tag{4}$$

Here, for $U \in V_{\mathrm{cG}}^{\boldsymbol{r}+1}(\mathcal{M}; X)$, $V \in V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X)$, the continuous Galerkin forms are given by

$$B_{\mathrm{cG}}(U, V) = \sum_{m=1}^{M} \int_{I_m} \left\{ (U', V)_H + a(U, V) \right\} \mathrm{d}t,$$

$$F_{\mathrm{cG}}(V) = \sum_{m=1}^{M} \int_{I_m} \langle g(t), V \rangle_{X^\star \times X} \,\mathrm{d}t.$$

It is well-known that the cG method in (4) is consistent and uniquely solvable; see, e.g., [13,14,31]. Furthermore, since the test functions are discontinuous, the variational problem (4) decouples into local problems on each time-step, giving rise to an implicit one-step time marching scheme. Indeed, suppose that the approximate solution U_{cG} is given on the time-steps $I_1, \ldots, I_{m-1}$, then $U_{\mathrm{cG}}|_{I_m} \in \mathcal{P}^{r_m+1}(I_m; X)$ on I_m is found by solving

$$\int_{I_m} \left\{ (U'_{\mathrm{cG}}, V)_H + a(U_{\mathrm{cG}}, V) \right\} \mathrm{d}t = \int_{I_m} \langle g(t), V \rangle_{X^\star \times X} \, \mathrm{d}t$$

for all $V \in \mathcal{P}^{r_m}(I_m; X)$, with the additional condition that

$$U^+_{\mathrm{cG},m-1} = U^-_{\mathrm{cG},m-1}.$$

Here, we use the convention that $U^-_{\mathrm{cG},0} = \widehat{u}_0 \in X$.

2.2 Discontinuous Galerkin time-stepping

Given a partition $\mathcal{M}$ and a degree vector $\boldsymbol{r}$, the *hp*-discontinuous Galerkin (dG) method for the approximation of (1) reads: Find $U_{\mathrm{dG}} \in V^{\boldsymbol{r}}_{\mathrm{dG}}(\mathcal{M}; X)$ such that

$$B_{\mathrm{dG}}(U_{\mathrm{dG}}, V) = F_{\mathrm{dG}}(V) \quad \forall V \in V^{\boldsymbol{r}}_{\mathrm{dG}}(\mathcal{M}; X). \tag{5}$$

Here, for $U, V \in V^{\boldsymbol{r}}_{\mathrm{dG}}(\mathcal{M}; X)$, the discontinuous Galerkin forms are given by

$$\begin{aligned} B_{\mathrm{dG}}(U, V) &= \sum_{m=1}^{M} \int_{I_m} \left\{ (U', V)_H + a(U, V) \right\} \mathrm{d}t \\ &\quad + \sum_{m=2}^{M} ([\![U]\!]_{m-1}, V^+_{m-1})_H + (U^+_0, V^+_0)_H, \\ F_{\mathrm{dG}}(V) &= \sum_{m=1}^{M} \int_{I_m} \langle g(t), V \rangle_{X^\star \times X} \, \mathrm{d}t + (\widehat{u}_0, V^+_0)_H. \end{aligned}$$

The dG method in (5) is consistent and has a unique solution; see [28, Chapter 12] or [23,24]. Similarly to the cG method, it can also be interpreted as an implicit one-step time-stepping scheme. Suppose again that the approximate solution U_{dG} is given on the time-steps $I_1, \ldots, I_{m-1}$. Then $U_{\mathrm{dG}}|_{I_m} \in \mathcal{P}^{r_m}(I_m; X)$ on I_m is found by solving

$$\int_{I_m} \left\{(U_{\mathrm{dG}}', V)_H + a(U_{\mathrm{dG}}, V)\right\} \mathrm{d}t + ([\![U_{\mathrm{dG}}]\!]_{m-1}, V_{m-1}^{+})_H = \int_{I_m} \langle g(t), V\rangle_{X^\star \times X} \,\mathrm{d}t$$

for all $V \in \mathcal{P}^{r_m}(I_m; X)$, where we let $[\![U_{\mathrm{dG}}]\!]_0 = \widehat{u}_0 \in X$.

3 A posteriori error estimation

In this section, we shall develop *hp*-version a posteriori error estimates for the time-stepping schemes in (4) and (5). In order to prepare the tools for the ensuing analysis, we shall first recall two families of specialized polynomial bases and prove some auxiliary results related to L^2-projections. Then, in Sects. 3.3 and 3.4, the cG and dG methods shall be analyzed, respectively.

3.1 Polynomial bases

We shall express functions in the space $V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X)$ in terms of stepwise Legendre series. To this end, let $\widehat{K}_i(t)$ denote the standard Legendre polynomial of degree $i \geq 0$ on the unit interval $\widehat{I} = (-1, 1)$, normalized such that $\widehat{K}_i(1) = 1$; cf., e.g., [1]. These polynomials satisfy the orthogonality relation

$$\int_{\widehat{I}} \widehat{K}_i(t)\widehat{K}_j(t)\,\mathrm{d}t = 2\gamma_i \delta_{i,j} \quad i, j \geq 0, \tag{6}$$

with $\gamma_i = \frac{1}{2i+1}$ and $\delta_{i,j}$ denoting the Kronecker symbol. Furthermore, there holds $\widehat{K}_i(-1) = (-1)^i$.

The reference interval $\widehat{I}$ can be mapped onto $I_m = (t_{m-1}, t_m)$, $1 \leq m \leq M$, by the use of the affine mapping

$$F_m : \widehat{I} \to I_m, \quad \hat{t} \mapsto t = F_m(\hat{t}) = \frac{k_m}{2}\hat{t} + \frac{t_{m-1} + t_m}{2}. \tag{7}$$

On I_m, the mapped Legendre polynomials are then defined by

$$K_i^m(t) = \widehat{K}_i(F_m^{-1}(t)), \quad t \in I_m.$$

A simple change of variables shows that

$$\int_{I_m} K_i^m(t)K_j^m(t)\,\mathrm{d}t = k_m \gamma_i \delta_{i,j} \quad i, j \geq 0. \tag{8}$$

Let now U be a piecewise polynomial in $V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X)$. On each time-step I_m, it is a polynomial of degree r_m. Hence, it can be expanded in the form

$$U|_{I_m}(t) = \sum_{i=0}^{r_m} u_i^m K_i^m(t), \quad t \in I_m,$$

with coefficients $u_i^m \in X$. From the orthogonality property (8) of the Legendre polynomials, it follows that

$$\|U\|_{L^2(I_m;X)}^2 = k_m \sum_{i=0}^{r_m} \gamma_i \|u_m^i\|_X^2.$$

We shall also consider the integrated Legendre polynomials. On the unit interval $\widehat{I} = (-1, 1)$ we define

$$\widehat{Q}_i(t) = \gamma_{i-1}\left(\widehat{K}_i(t) - \widehat{K}_{i-2}(t)\right), \quad i \geq 0, \tag{9}$$

where we set $\widehat{K}_{-1} = \widehat{K}_{-2} = 0$ and $\gamma_{-1} = 1$. The function $\widehat{Q}_i(t)$ is a polynomial of degree i and the set $\{\widehat{Q}_i(t)\}_{i=0}^r$ forms a basis of the polynomial space $\mathcal{P}^r(\widehat{I}; \mathbb{R})$. Note that $\widehat{Q}_0(t) = 1$, $\widehat{Q}_1(t) = t$, and $\widehat{Q}_i(\pm 1) = 0$ for $i \geq 2$. Moreover, well-known properties of the Legendre polynomials, see, e.g., [27, Section A.4], and the orthogonality conditions in (6) readily imply the following result.

Lemma 1 *For any $i \geq 2$, there holds*

$$\widehat{Q}_i(t) = \int_{-1}^{t} \widehat{K}_{i-1}(\tau)\, \mathrm{d}\tau.$$

Furthermore, for $i, j \geq 0$, the following relations hold:

$$\int_{\widehat{I}} \widehat{Q}_i(t)\widehat{Q}_j(t)\, \mathrm{d}t = \begin{cases} 2\gamma_{i-1}^2\gamma_i & \text{if } i = j = 0, 1, \\ 2\gamma_{i-1}^2(\gamma_i + \gamma_{i-2}) & \text{if } i = j \geq 2, \\ -2\gamma_{i-1}\gamma_{i-3}\gamma_{i-2} & \text{if } i = j + 2, \\ -2\gamma_{i-1}\gamma_{i+1}\gamma_i & \text{if } j = i + 2, \\ 0 & \text{otherwise}. \end{cases}$$

3.2 L^2-Projection

Next, we define the L^2-projection for functions in $L^2(I_m; X^\star)$ and $L^2((0, T); X^\star)$. We begin by showing the following auxiliary result.

Lemma 2 *If* $p \in \mathcal{P}^{r_m}(I_m; X^\star)$ *satisfies*

$$\langle p, q \rangle_{L^2(I_m;X^\star)\times L^2(I_m;X)} = 0 \quad \forall q \in \mathcal{P}^{r_m}(I_m; X),$$

then p *is the zero polynomial.*

Proof We expand p in the Legendre basis and have

$$p(t) = \sum_{j=0}^{r_m} p_j^\star K_j^m(t),$$

with coefficients $p_j^\star \in X^\star$. For $0 \le i \le r_m$ and $v \in X$, consider the polynomial $q(t) = vK_i^m(t) \in \mathcal{P}^{r_m}(I_m; X)$. Then the orthogonality (8) of the Legendre polynomials yields

$$\begin{aligned} 0 &= \int_{I_m} \langle p(t), q(t) \rangle_{X^\star \times X}\, \mathrm{d}t = \sum_{j=0}^{r_m} \langle p_j^\star, v \rangle_{X^\star\times X} \int_{I_m} K_j^m(t) K_i^m(t)\, \mathrm{d}t \\ &= k_m \gamma_i \langle p_i^\star, v \rangle_{X^\star \times X}. \end{aligned}$$

Since $0 \le i \le r_m$ and $v \in X$ is arbitrary, we conclude that

$$p_i^\star = 0 \quad \forall 0 \le i \le r_m.$$

This completes the proof. □

For a function $u \in L^2(I_m; X^\star)$, we define its L^2-projection $\Pi_m^{\star,r_m} u$ to the space $\mathcal{P}^{r_m}(I_m; X^\star)$ by requiring that

$$\langle \Pi_m^{\star,r_m} u, V \rangle_{L^2(I_m;X^\star)\times L^2(I_m;X)} = \langle u, V \rangle_{L^2(I_m;X^\star)\times L^2(I_m;X)} \tag{10}$$

for all $V \in \mathcal{P}^{r_m}(I_m; X)$. In view of Lemma 2, the L^2-projection is uniquely defined. In fact, the following result holds true.

Lemma 3 *For* $u \in L^2(I_m; X^\star)$*, we have*

$$(\Pi_m^{\star,r_m} u)(t) = \sum_{i=0}^{r_m} u_i^\star K_i^m(t),$$

with $u_i^\star \in X^\star$ *given by*

$$\langle u_i^\star, v \rangle_{X^\star \times X} = \frac{1}{k_m \gamma_i} \langle u, vK_i^m \rangle_{L^2(I_m;X^\star)\times L^2(I_m;X)}, \quad v \in X.$$

Proof We first remark that

$$|\langle u_i^\star, v\rangle_{X^\star\times X}| \le \frac{1}{k_m\gamma_i}\|u\|_{L^2(I_m;X^\star)}\|vK_i^m\|_{L^2(I_m;X)}.$$

Therefore, since

$$\|vK_i^m\|^2_{L^2(I_m;X)} = \|v\|_X^2\int\limits_{I_m} K_i^m(t)^2\,\mathrm{d}t = k_m\gamma_i\|v\|_X^2,$$

the functional $u_i^\star$ belongs to $X^\star$.

Let now

$$U(t) = \sum_{i=0}^{r_m} u_i^\star K_i^m(t),$$

and let $V \in \mathcal{P}^{r_m}(I_m; X)$ be given by

$$V(t) = \sum_{j=0}^{r_m} v_j K_j^m(t),$$

with coefficients $v_j \in X$. From the orthogonality relation (8) and the definition of $u_i^\star$, we then obtain

$$\begin{aligned}\langle U, V\rangle_{L^2(I_m;X^\star)\times L^2(I_m;X)} &= \sum_{i,j=0}^{r_m}\langle u_i^\star, v_j\rangle_{X^\star\times X}\int\limits_{I_m} K_i^m(t)K_j^m(t)\,\mathrm{d}t\\ &= \sum_{i=0}^{r_m} k_m\gamma_i\langle u_i^\star, v_i\rangle_{X^\star\times X}\\ &= \sum_{i=0}^{r_m}\langle u, v_iK_i^m\rangle_{L^2(I_m;X^\star)\times L^2(I_m;X)}\\ &= \langle u, V\rangle_{L^2(I_m;X^\star)\times L^2(I_m;X)}.\end{aligned}$$

Hence, U is the L^2-projection. This completes the proof. □

Next, let us consider the L^2-projection of functions in $L^2(I_m; X)$ that are of the form

$$u(t) = \sum_{j=0}^{R} u_j K_j^m(t), \tag{11}$$

with $R \in \mathbb{N}$, i.e., only finitely many coefficients $u_j \in X$ in the above sum differ from zero.

Lemma 4 *Let $u \in L^2(I_m; X)$ be given by the series (11). Then we have*

$$\Pi_m^{\star,r_m}(\mathsf{A}u) = \mathsf{A}(\Pi_m^{r_m} u),$$

where $\Pi_m^{r_m}$ signifies the L^2-projection from $L^2(I_m; H)$ to $\mathcal{P}^{r_m}(I_m, H)$ with respect to the inner product $(\cdot, \cdot)_{L^2(I_m;H)}$.

Proof We first note that

$$\left(\Pi_m^{r_m} u\right)(t) = \sum_{i=0}^{r_m} u_i K_i^m(t) \in L^2(I_m; X). \tag{12}$$

Hence,

$$\mathsf{A}(\Pi_m^{r_m} u) = \sum_{i=0}^{r_m} \mathsf{A}u_i K_i^m(t). \tag{13}$$

On the other hand, we have

$$\mathsf{A}u = \sum_{j=0}^{R} \mathsf{A}u_j K_j^m(t).$$

Applying Lemma 3, we obtain

$$\Pi_m^{\star,r_m}(\mathsf{A}u) = \sum_{i=0}^{r_m} y_i^\star K_i^m(t), \tag{14}$$

where, for any $v \in X$, there holds

$$\begin{aligned}
\langle y_i^\star, v\rangle_{X^\star \times X} &= \frac{1}{k_m \gamma_i} \langle \mathsf{A}u, v K_i^m\rangle_{L^2(I_m;X^\star)\times L^2(I_m;X)} \\
&= \frac{1}{k_m \gamma_i} \int_{I_m} \langle \mathsf{A}u, v K_i^m\rangle_{X^\star\times X}\, \mathrm{d}t \\
&= \frac{1}{k_m \gamma_i} \sum_{j=0}^{R} \langle \mathsf{A}u_j, v\rangle_{X^\star\times X} \int_{I_m} K_j^m(t) K_i^m(t)\, \mathrm{d}t,
\end{aligned}$$

$i = 0, 1, \ldots, r_m$. Thus, using the orthogonality property (8), yields

$$\langle y_i^\star, v\rangle_{X^\star\times X} = \langle \mathsf{A}u_i, v\rangle_{X^\star\times X} \quad \forall v \in X.$$

Therefore, by Lemma 2, we have that $y_i^\star = \mathsf{A}u_i$, $i = 0, 1, \ldots, r_m$. Comparing (13) and (14) completes the proof. □

Finally, we define the following global L^2-projections elementwise as

$$\Pi^{\star,\boldsymbol{r}}|_{I_m} = \Pi_m^{\star,r_m}, \quad \Pi^{\boldsymbol{r}}|_{I_m} = \Pi_m^{r_m}, \quad 1 \leq m \leq M. \tag{15}$$

3.3 Error estimation for the cG time-stepping method

Let u be the solution of (1) and $U_{\mathrm{cG}} \in V_{\mathrm{cG}}^{\boldsymbol{r}+1}(\mathcal{M}; X)$ the cG approximation from (4). On each time-step I_m, the function U_{cG} is a polynomial of degree $r_m + 1$. Hence, it can be expanded in the form

$$U_{\mathrm{cG}}(t) = \sum_{i=0}^{r_m+1} u_i^m K_i^m(t), \quad t \in I_m, \tag{16}$$

with coefficients $u_i^m \in X$.

We now define the error measure

$$\mathcal{E}_{\mathrm{cG}} = \max\left\{ E_{1,\mathrm{cG}}, E_{2,\mathrm{cG}}, E_{3,\mathrm{cG}} \right\}, \tag{17}$$

where

$$\begin{aligned} E_{1,\mathrm{cG}} &= \|u - U_{\mathrm{cG}}\|_{C^0([0,T];H)}, \\ E_{2,\mathrm{cG}} &= \sqrt{\frac{\alpha}{2}} \|u - U_{\mathrm{cG}}\|_{L^2((0,T);X)}, \\ E_{3,\mathrm{cG}} &= \sqrt{\frac{\alpha}{2}} \|u - \Pi^{\boldsymbol{r}} U_{\mathrm{cG}}\|_{L^2((0,T);X)}. \end{aligned}$$

Next, we introduce the elemental error indicator

$$\eta_{\mathrm{cG},m}^2 = \frac{k_m}{2r_m + 3} \|u_{r_m+1}^m\|_X^2, \quad m = 1, 2, \ldots, M, \tag{18}$$

where $u_{r_m+1}^m$ is the Legendre coefficient of order $r_m + 1$ in the expansion (16). We further set

$$\eta_{\mathrm{cG}}^2 = \sum_{m=1}^{M} \eta_{\mathrm{cG},m}^2. \tag{19}$$

The following estimate is our main result for the continuous Galerkin time-stepping method. It shows that the error indicator η_{cG} gives rise to a reliable and efficient a posteriori error estimate, up to data approximation terms.

Theorem 1 *Let η_{cG} be defined in (19). Then the error measure (17) satisfies the upper bound*

$$\mathcal{E}_{\mathrm{cG}}^2 \leq \frac{\beta^2}{\alpha} \eta_{\mathrm{cG}}^2 + \frac{2}{\alpha} \|g - \Pi^{\star,\boldsymbol{r}} g\|_{L^2((0,T);X^\star)}^2 + \|u_0 - \widehat{u}_0\|_H^2,$$

and the lower bound

$$\frac{\alpha}{8}\eta_{\mathrm{cG}}^2 \le \mathcal{E}_{\mathrm{cG}}^2.$$

Remark 2 The estimates in Theorem 1 are fully explicit with respect to both the step sizes and the local polynomial degrees, and the constants occurring in the error bounds are explicitly given in dependence on the continuity and coercivity constants α, β of the bilinear form $a(\cdot,\cdot)$; cf. (2) and (3).

The proof of Theorem 1 will follow along the lines of [19, Theorem 3.3 and Theorem 3.8]. Before carrying it out in detail, we adapt [19, Lemma 3.2] to our setting. This leads to the following auxiliary result.

Lemma 5 *There holds*

$$a(U_1 - u, u - U_2) \le -\tfrac{\alpha}{2}\max\left\{\|u-U_1\|_X^2, \|u-U_2\|_X^2\right\} + \tfrac{\beta^2}{2\alpha}\|U_1-U_2\|_X^2$$

for any $u, U_1, U_2 \in X$.

Proof Using the coercivity and continuity of the elliptic form $a(\cdot,\cdot)$ in (2) and (3), we obtain

$$\begin{aligned} a(U_1-u, u-U_2) &= a(U_1-u, u-U_1) + a(U_1-u, U_1-U_2) \\ &\le -\alpha\|u-U_1\|_X^2 + \beta\|u-U_1\|_X\|U_1-U_2\|_X. \end{aligned}$$

The inequality $|ab| \le \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2$, with $\varepsilon = \alpha/\beta$, then gives

$$a(U_1-u, u-U_2) \le -\frac{\alpha}{2}\|u-U_1\|_X^2 + \frac{\beta^2}{2\alpha}\|U_1-U_2\|_X^2. \tag{20}$$

Similarly, there holds

$$\begin{aligned} a(U_1-u, u-U_2) &= a(U_2-u, u-U_2) + a(U_1-U_2, u-U_2) \\ &\le -\frac{\alpha}{2}\|u-U_2\|_X^2 + \frac{\beta^2}{2\alpha}\|U_1-U_2\|_X^2. \end{aligned} \tag{21}$$

The inequalities (20) and (21) now imply the desired bound. □

We are now ready to show Theorem 1.

Proof of Theorem 1 The cG solution $U_{\mathrm{cG}} \in V_{\mathrm{cG}}^{\boldsymbol{r}+1}(\mathcal{M}; X)$ satisfies

$$\int_0^T \left\{(U_{\mathrm{cG}}', V)_H + \langle \mathsf{A}U_{\mathrm{cG}}, V\rangle_{X^\star\times X}\right\} \mathrm{d}t = \int_0^T \langle g, V\rangle_{X^\star\times X}\,\mathrm{d}t$$

for all $V \in V^{r}_{\mathrm{dG}}(\mathcal{M};X)$. We note that $U'_{\mathrm{cG}} \in V^{r}_{\mathrm{dG}}(\mathcal{M};X)$. Furthermore, we have $(U'_{\mathrm{cG}}, V)_H = \langle U'_{\mathrm{cG}}, V\rangle_{X^\star\times X}$ for all $V \in V^{r}_{\mathrm{dG}}(\mathcal{M};X)$. Using the definition of the L^2-projection $\Pi^{\star,r}$, see (10) and (15), and Lemma 2, there holds:

$$U'_{\mathrm{cG}} + \Pi^{\star,r}\mathsf{A}U_{\mathrm{cG}} = \Pi^{\star,r}g \quad \text{in } V^{r}_{\mathrm{dG}}(\mathcal{M};X^\star), \quad U_{\mathrm{cG}}(0) = \widehat{u}_0.$$

Furthermore, applying Lemma 4 leads to

$$U'_{\mathrm{cG}} + \mathsf{A}\Pi^{r}U_{\mathrm{cG}} = \Pi^{\star,r}g, \quad U_{\mathrm{cG}}(0) = \widehat{u}_0.$$

We conclude that

$$\begin{aligned}(u-U_{\mathrm{cG}})' + \mathsf{A}(u-\Pi^{r}U_{\mathrm{cG}}) = g - \Pi^{\star,r}g \quad \text{in } L^2((0,T);X^\star),\\ (u-U_{\mathrm{cG}})(0) = u_0 - \widehat{u}_0.\end{aligned}$$

Testing the above equation with $(u - U_{\mathrm{cG}})$ gives

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u-U_{\mathrm{cG}}\|_H^2 - a(\Pi^{r}U_{\mathrm{cG}}-u, u-U_{\mathrm{cG}}) = \langle g-\Pi^{\star,r}g, u-U_{\mathrm{cG}}\rangle_{X^\star\times X}.$$

Here, we have used the fact that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u-U_{\mathrm{cG}}\|_H^2 = \langle (u-U_{\mathrm{cG}})', u-U_{\mathrm{cG}}\rangle_{X^\star\times X}.$$

Moreover, from Lemma 5, it follows that

$$-a(\Pi^{r}U_{\mathrm{cG}}-u, u-U_{\mathrm{cG}}) \geq \frac{\alpha}{2}M(t)^2 - \frac{\beta^2}{2\alpha}\|U_{\mathrm{cG}}-\Pi^{r}U_{\mathrm{cG}}\|_X^2,$$

where

$$M(t)^2 = \max\left\{\|(u-U_{\mathrm{cG}})(t)\|_X^2, \|(u-\Pi^{r}U_{\mathrm{cG}})(t)\|_X^2\right\}.$$

Therefore,

$$\begin{aligned}&\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u-U_{\mathrm{cG}}\|_H^2 + \frac{\alpha}{2}M(t)^2\\ &\qquad\leq \frac{\beta^2}{2\alpha}\|U_{\mathrm{cG}}-\Pi^{r}U_{\mathrm{cG}}\|_X^2 + |\langle g-\Pi^{\star,r}g, u-U_{\mathrm{cG}}\rangle_{X^\star\times X}|. \qquad (22)\end{aligned}$$

The inequality $|ab| \leq \frac{\varepsilon}{2}a^2 + \frac{1}{2\varepsilon}b^2$, with $\varepsilon = 2/\alpha$, then shows that

$$|\langle g - \Pi^{\star,r} g, u - U_{\mathrm{cG}} \rangle_{X^\star \times X}| \leq \|g - \Pi^{\star,r} g\|_{X^\star} M(t) \\ \leq \frac{1}{\alpha}\|g - \Pi^{\star,r} g\|_{X^\star}^2 + \frac{\alpha}{4} M(t)^2. \tag{23}$$

We now combine the inequalities in (22) and (23), subtract the term $\frac{\alpha}{4}M(t)^2$ on both sides, and multiply the resulting inequality by 2. We obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u - U_{\mathrm{cG}}\|_H^2 + \frac{\alpha}{2}M(t)^2 \leq \frac{\beta^2}{\alpha}\|U_{\mathrm{cG}} - \Pi^r U_{\mathrm{cG}}\|_X^2 + \frac{2}{\alpha}\|g - \Pi^{\star,r} g\|_{X^\star}^2.$$

We recall that $u - U_{\mathrm{cG}} \in C^0([0, T], H)$. Hence, integrating the above estimate over $(0, t)$ for $0 \leq t \leq T$ and observing the initial condition lead to

$$\|(u - U_{\mathrm{cG}})(t)\|_H^2 + \frac{\alpha}{2}\|M\|_{L^2((0,t);X)}^2 \leq \frac{\beta^2}{\alpha}\|U_{\mathrm{cG}} - \Pi^r U_{\mathrm{cG}}\|_{L^2((0,T);X)}^2 \\ + \frac{2}{\alpha}\|g - \Pi^{\star,r} g\|_{L^2((0,T);X^\star)}^2 \\ + \|u_0 - \widehat{u}_0\|_H^2.$$

Since this holds for any $0 \leq t \leq T$, we obtain

$$\mathcal{E}_{\mathrm{cG}}^2 \leq \frac{\beta^2}{\alpha}\|U_{\mathrm{cG}} - \Pi^r U_{\mathrm{cG}}\|_{L^2((0,T);X)}^2 + \frac{2}{\alpha}\|g - \Pi^{\star,r} g\|_{L^2((0,T);X^\star)}^2 + \|u_0 - \widehat{u}_0\|_H^2. \tag{24}$$

Furthermore, using the triangle inequality, we immediately obtain the lower bound

$$\|U_{\mathrm{cG}} - \Pi^r U_{\mathrm{cG}}\|_{L^2((0,T);X)}^2 \\ \leq \frac{4}{\alpha}\left(\frac{\alpha}{2}\|u - U_{\mathrm{cG}}\|_{L^2((0,T);X)}^2 + \frac{\alpha}{2}\|u - \Pi^r U_{\mathrm{cG}}\|_{L^2((0,T);X)}^2\right) \\ \leq \frac{8}{\alpha}\mathcal{E}_{\mathrm{cG}}^2. \tag{25}$$

Then, from the solution representation in (16) and from (12), it follows that

$$\Pi^r U_{\mathrm{cG}}(t) = \sum_{i=0}^{r_m} u_i^m K_i^m(t), \quad t \in I_m,$$

Hence, we have

$$U_{\mathrm{cG}}(t) - \Pi^r U_{\mathrm{cG}}(t) = u_{r_m+1}^m K_{r_m+1}^m(t), \quad t \in I_m,$$

and

$$\|U_{\mathrm{cG}} - \Pi^{r} U_{\mathrm{cG}}\|^2_{L^2(I_m;X)} = k_m \gamma_{r_m+1} \|u^m_{r_m+1}\|^2_X. \tag{26}$$

The proof of Theorem 1 now follows from (24) to (26). □

3.4 Error estimation for the dG time-stepping method

We shall now derive an *hp*-version a posteriori error estimate for the dG method (5). For this purpose, consider a dG function $U \in V^{r}_{\mathrm{dG}}(\mathcal{M}; X)$. Then, following [19, Section 2.1], we define a reconstruction $\widehat{U} = \mathsf{R}(U) \in V^{r+1}_{\mathrm{cG}}(\mathcal{M}; X)$ of U by requiring that

$$\int_{I_m} (\widehat{U}', V)_H \,\mathrm{d}t = \int_{I_m} (U', V)_H \,\mathrm{d}t + ([\![U]\!]_{m-1}, V)_H, \tag{27}$$

for all $V \in \mathcal{P}^{r_m}(I_m; X)$, and $\widehat{U}^+_{m-1} = U^-_{m-1}$. On the first element, we take $\widehat{U}^+_0 = \widehat{u}_0$. From [19, Lemma 2.1], it follows that $\widehat{U}$ is well-defined and continuous on $[0, T]$. We note that the operator R is only required for the purpose of the analysis and does not need to be computed in practice.

A second main result of this paper is the complete *hp*-version characterization of the difference $U - \widehat{U}$ appearing in the a posteriori error analysis of the dG method (5). More precisely, we will prove the following identities.

Theorem 2 *Let $U \in V^{r}_{\mathrm{dG}}(\mathcal{M}; X)$. Then, for $1 \le m \le M$, we have*

$$\|U - \widehat{U}\|^2_{L^2(I_m;X)} = k_m \frac{r_m + 1}{(2r_m + 1)(2r_m + 3)} \|[\![U]\!]_{m-1}\|^2_X.$$

Consequently, there holds

$$\|U - \widehat{U}\|^2_{L^2((0,T);X)} = \sum_{m=1}^{M} k_m \frac{r_m + 1}{(2r_m + 1)(2r_m + 3)} \|[\![U]\!]_{m-1}\|^2_X.$$

The proof of this result will be given in Sect. 4.

Let now u be the solution of the parabolic equation (1) and $U_{\mathrm{dG}} \in V^{r}_{\mathrm{dG}}(\mathcal{M}; X)$ the discontinuous Galerkin approximation from (5). We define the error measure

$$\mathcal{E}_{\mathrm{dG}} = \max\left\{ E_{1,\mathrm{dG}}, E_{2,\mathrm{dG}}, E_{3,\mathrm{dG}} \right\}, \tag{28}$$

where

$$E_{1,\mathrm{dG}} = \sqrt{\frac{\alpha}{2}}\|u - U_{\mathrm{dG}}\|_{L^2((0,T);X)},$$
$$E_{2,\mathrm{dG}} = \|u - \widehat{U}_{\mathrm{dG}}\|_{C^0([0,T];H)},$$
$$E_{3,\mathrm{dG}} = \sqrt{\frac{\alpha}{2}}\|u - \widehat{U}_{\mathrm{dG}}\|_{L^2((0,T);X)}.$$

Applying the ideas in [19], and proceeding similarly to the analysis of the cG method, the following *hp*-version a posteriori error estimates are obtained.

Proposition 1 *The error measure* (28) *satisfies the upper bound*

$$\mathcal{E}_{\mathrm{dG}}^2 \le \frac{\beta^2}{\alpha}\|U_{\mathrm{dG}} - \widehat{U}_{\mathrm{dG}}\|_{L^2((0,T);X)}^2 + \frac{2}{\alpha}\|g - \Pi^{\star,\boldsymbol{r}} g\|_{L^2((0,T);X^\star)}^2 + \|u_0 - \widehat{u}_0\|_H^2,$$

and the lower bound

$$\frac{\alpha}{8}\|U_{\mathrm{dG}} - \widehat{U}_{\mathrm{dG}}\|_{L^2((0,T);X)}^2 \le \mathcal{E}_{\mathrm{dG}}^2.$$

Proof Recalling the definition of the dG scheme and of the reconstruction $\widehat{U}_{\mathrm{dG}} = \mathsf{R}(U_{\mathrm{dG}})$ of U_{dG}, there holds

$$\int_0^T \left\{(\widehat{U}_{\mathrm{dG}}', V)_H + \langle \mathsf{A}U_{\mathrm{dG}}, V\rangle_{X^\star\times X}\right\} \mathrm{d}t = \int_0^T \langle g, V\rangle_{X^\star\times X}\,\mathrm{d}t$$

for all $V \in V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X)$. Then, since $\widehat{U}_{\mathrm{dG}}', \mathsf{A}U_{\mathrm{dG}} \in V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X^\star)$, it follows from the definitions (10) and (15) of the L^2-projection that

$$(u - \widehat{U}_{\mathrm{dG}})' + \mathsf{A}(u - U_{\mathrm{dG}}) = g - \Pi^{\star,\boldsymbol{r}} g \quad \text{in } L^2((0,T); X^\star),$$
$$(u - \widehat{U}_{\mathrm{dG}})(0) = u_0 - \widehat{u}_0.$$

Testing this equation with $u - \widehat{U}_{\mathrm{dG}}$ gives

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|u - \widehat{U}_{\mathrm{dG}}\|_H^2 - a(U_{\mathrm{dG}} - u, u - \widehat{U}_{\mathrm{dG}}) = \langle g - \Pi^{\star,\boldsymbol{r}} g, u - \widehat{U}_{\mathrm{dG}}\rangle_{X^\star\times X}.$$

Upon setting

$$M(t)^2 = \max\left\{\|(u - U_{\mathrm{dG}})(t)\|_X^2, \|(u - \widehat{U}_{\mathrm{dG}})(t)\|_X^2\right\},$$

the bound in Lemma 5 ensures that

$$-a(U_{\mathrm{dG}} - u, u - \widehat{U}_{\mathrm{dG}}) \ge \frac{\alpha}{2}M(t)^2 - \frac{\beta^2}{2\alpha}\|U_{\mathrm{dG}} - \widehat{U}_{\mathrm{dG}}\|_X^2.$$

Furthermore, similarly to (23), it holds that

$$|\langle g - \Pi^{\star,r} g, u - \widehat{U}_{\mathrm{dG}} \rangle_{X^\star \times X}| \leq \frac{1}{\alpha} \| g - \Pi^{\star,r} g \|_{X^\star}^2 + \frac{\alpha}{4} M(t)^2.$$

Hence, as in the proof of the a posteriori error estimate for the cG method, see (22) and (23), it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t} \| u - \widehat{U}_{\mathrm{dG}} \|_H^2 + \frac{\alpha}{2} M(t)^2 \leq \frac{\beta^2}{\alpha} \| U_{\mathrm{dG}} - \widehat{U}_{\mathrm{dG}} \|_X^2 + \frac{2}{\alpha} \| g - \Pi^{\star,r} g \|_{X^\star}^2.$$

Therefore, integrating from 0 to t leads to

$$\begin{aligned} \|(u - \widehat{U}_{\mathrm{dG}})(t)\|_H^2 + \frac{\alpha}{2} \|M\|_{L^2((0,t);X)}^2 \leq \frac{\beta^2}{\alpha} & \| U_{\mathrm{dG}} - \widehat{U}_{\mathrm{dG}} \|_{L^2((0,T);X)}^2 \\ &+ \frac{2}{\alpha} \| g - \Pi^{\star,r} g \|_{L^2((0,T);X^\star)}^2 \\ &+ \| u_0 - \widehat{u}_0 \|_H^2. \end{aligned}$$

This readily implies the lower bound.

The upper bound follows from the triangle inequality

$$\begin{aligned} \| U_{\mathrm{dG}} - \widehat{U}_{\mathrm{dG}} \|_{L^2((0,T);X)}^2 &\leq \frac{4}{\alpha} \left(\frac{\alpha}{2} \| u - U_{\mathrm{dG}} \|_{L^2((0,T);X)}^2 + \frac{\alpha}{2} \| u - \widehat{U}_{\mathrm{dG}} \|_{L^2((0,T);X)}^2 \right) \\ &\leq \frac{8}{\alpha} \mathcal{E}_{\mathrm{dG}}^2, \end{aligned}$$

which completes the proof. □

For $m = 1, 2, \dots M$, we now introduce the elemental error indicator

$$\eta_{\mathrm{dG},m}^2 = k_m \frac{r_m + 1}{(2r_m + 1)(2r_m + 3)} \| [\![U_{\mathrm{dG}}]\!]_{m-1} \|_X^2, \tag{29}$$

and set

$$\eta_{\mathrm{dG}}^2 = \sum_{m=1}^{M} \eta_{\mathrm{dG},m}^2. \tag{30}$$

The combination of Theorem 2 and Proposition 1 yields the following *hp*-version a posteriori error estimate for the dG time-stepping scheme (5).

Theorem 3 *Let η_{dG} be defined in (30). Then the error measure (28) satisfies the upper bound*

$$\mathcal{E}_{\mathrm{dG}}^2 \leq \frac{\beta^2}{\alpha} \eta_{\mathrm{dG}}^2 + \frac{2}{\alpha} \| g - \Pi^{\star,r} g \|_{L^2((0,T);X^\star)}^2 + \| u_0 - \widehat{u}_0 \|_H^2,$$

and the lower bound

$$\frac{\alpha}{8}\eta_{\mathrm{dG}}^2 \leq \mathcal{E}_{\mathrm{dG}}^2.$$

We note that, as for the cG method, the above a posteriori error estimates are fully explicit with respect to the time-steps, the local polynomial degrees, and the stability constants α and β corresponding to the bilinear form $a(\cdot,\cdot)$; cf. (2) and (3).

4 Proof of Theorem 2

In this section, we will present the proof of Theorem 2. In Sect. 4.1, we will derive a representation formula for the difference $U - \widehat{U}$ in terms of a lifting operator. Section 4.2 focuses on the properties of (more general) polynomial lifting operators. Finally, in Sect. 4.3 we will complete the proof of Theorem 2.

4.1 A representation formula

We shall first derive a representation formula for the reconstruction error $U - \widehat{U}$ occurring in Theorem 2.

Let $U \in V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X)$ be a dG function. As in the unifying framework proposed in [2] for the analysis of discontinuous Galerkin methods for elliptic problems, we rewrite the jumps of U in terms of lifting operators. Specifically, for $1 \leq m \leq M$, we define the lifting

$$\mathsf{L}_m : V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X) \to \mathcal{P}^{r_m}(I_m; X)$$

by requiring that

$$\int_{I_m} (\mathsf{L}_m(U), V)_H \,\mathrm{d}t = \left([\![U]\!]_{m-1}, V_{m-1}^+\right)_H \quad \forall V \in \mathcal{P}^{r_m}(I_m; X), \tag{31}$$

where we use $[\![U]\!]_0 = U_0^+ - \widehat{u}_0$ on the first element.

The following identity holds. It will be derived from a more general result and will be proved at the end of Sect. 4.2.

Proposition 2 *The lifting $\mathsf{L}_m(U)$ from (31) is well-defined and we have*

$$\|\mathsf{L}_m(U)\|_{L^2(I_m;X)}^2 = \frac{(r_m+1)^2}{k_m}\|[\![U]\!]_{m-1}\|_X^2$$

for any $U \in V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X)$.

In order to relate the reconstruction $\widehat{U} = \mathsf{R}(U)$ of a dG function U defined in (27) to the lifting $\mathsf{L}_m(U)$, we notice that, for $U, V \in V_{\mathrm{dG}}^{\boldsymbol{r}}(\mathcal{M}; X)$, the discontinuous Galerkin form B_{dG} can be expressed as

$$B_{\mathrm{dG}}(U, V) = \sum_{m=1}^{M} \int_{I_m} \left\{(U' + \mathsf{L}_m(U), V)_H + a(U, V)\right\} \mathrm{d}t.$$

More importantly, the defining properties of the reconstruction $\widehat{U}$ in (27) can be written as

$$\int_{I_m} (\widehat{U}', V)_H \,\mathrm{d}t = \int_{I_m} (U' + \mathsf{L}_m(U), V)_H \,\mathrm{d}t \quad \forall V \in \mathcal{P}^{r_m}(I_m; X),$$

and

$$\widehat{U}^{+}_{m-1} = U^{-}_{m-1}, \tag{32}$$

with $U_0^- = \widehat{u}_0$. Hence, since $\widehat{U}' \in \mathcal{P}^{r_m}(I_m; X)$, we have

$$\widehat{U}' = U' + \mathsf{L}_m(U) \quad \text{on } I_m.$$

Integrating this equation over (t_{m-1}, t) for $t_{m-1} \le t \le t_m$, we obtain

$$\widehat{U}(t) - \widehat{U}^{+}_{m-1} = U(t) - U^{+}_{m-1} + \int_{t_{m-1}}^{t} \mathsf{L}_m(U)\,\mathrm{d}\tau.$$

Then, applying (32), we obtain the representation formula

$$U(t) - \widehat{U}(t) = [\![U]\!]_{m-1} - \int_{t_{m-1}}^{t} \mathsf{L}_m(U)\,\mathrm{d}\tau, \quad t_{m-1} \le t \le t_m, \tag{33}$$

for $1 \le m \le M$, where $[\![U]\!]_0 = U_0^+ - \widehat{u}_0$ on the first element.

4.2 Lifting operators

In this section, we introduce and analyze generic lifting operators. In addition, we will prove the stability result in Proposition 2.

We shall first consider lifting operators on the unit interval $\widehat{I} = (-1, 1)$. Let $r \ge 0$ and $z \in X$ be fixed. We denote by $\widehat{L}^r_z$ the polynomial in $\mathcal{P}^r(\widehat{I}; X)$ that satisfies

$$\int_{\widehat{I}} (\widehat{L}^r_z(t), \widehat{V}(t))_H \,\mathrm{d}t = (z, \widehat{V}(-1))_H \quad \forall \widehat{V} \in \mathcal{P}^r(\widehat{I}; X). \tag{34}$$

Lemma 6 *The lifting* $\widehat{L}_z^r$ *is well-defined and can be represented in the form*

$$\widehat{L}_z^r(t) = \frac{z}{2}\sum_{i=0}^{r}(-1)^i(2i+1)\widehat{K}_i(t).$$

Furthermore, the identity

$$\|\widehat{L}_z^r\|^2_{L^2(\widehat{I};X)} = \frac{\|z\|_X^2}{2}(r+1)^2$$

holds.

Proof By expanding $\widehat{L}_z^r(t)$ in Legendre polynomials, we have

$$\widehat{L}_z^r(t) = \sum_{i=0}^{r} a_i\widehat{K}_i(t),$$

with coefficients $a_i \in X$ to be determined. In view of the orthogonality property (6) of the Legendre polynomials, we see that

$$a_i = \frac{1}{2\gamma_i}\int_{\widehat{I}} \widehat{L}_z^r(t)\widehat{K}_i(t)\,\mathrm{d}t.$$

Therefore,

$$(a_i, v)_H = \frac{1}{2\gamma_i}\int_{\widehat{I}} (\widehat{L}_z^r(t), v\widehat{K}_i(t))_H\,\mathrm{d}t \quad \forall v \in X.$$

Testing (34) with $\widehat{V}(t) = v\widehat{K}_i(t)$, $v \in X$, and noting that $\widehat{K}_i(-1) = (-1)^i$, yields that

$$\begin{aligned}(a_i, v)_H &= \frac{1}{2\gamma_i}\int_{\widehat{I}} (\widehat{L}_z^r(t), v\widehat{K}_i(t))_H\,\mathrm{d}t = \frac{1}{2\gamma_i}(z, v\widehat{K}_i(-1))_H \\ &= \frac{(-1)^i}{2\gamma_i}(z, v)_H.\end{aligned}$$

Since this holds for any $v \in X$ and X is dense in H, we obtain

$$a_i = \frac{z}{2}\frac{1}{\gamma_i}(-1)^i = \frac{z}{2}(2i+1)(-1)^i.$$

This shows the first part of the lemma.

Moreover, using this representation of $\widehat{L}_z^r(t)$ and the orthogonality properties of the Legendre polynomials, we conclude that

$$\|\widehat{L}_z^r\|^2_{L^2(\widehat{I};X)} = \frac{\|z\|_X^2}{4}\sum_{i=0}^{r} 2\gamma_i(2i+1)^2 = \frac{\|z\|_X^2}{2}\sum_{i=0}^{r}(2i+1)$$
$$= \frac{\|z\|_X^2}{2}(r+1)^2.$$

This proves the second claim. □

Let us now consider the interval $I_m = (t_{m-1}, t_m)$. For $r_m \geq 0$ and $z \in X$, we define the lifting $L_{m,z}^{r_m} \in \mathcal{P}^{r_m}(I_m; X)$ by

$$\int_{I_m} (L_{m,z}^{r_m}(t), V(t))_H \, \mathrm{d}t = (z, V(t_{m-1}))_H \quad \forall V \in \mathcal{P}^{r_m}(I_m; X).$$

Lemma 7 *We have*

$$L_{m,z}^{r_m}(F_m(\hat{t})) = \frac{2}{k_m}\widehat{L}_z^{r_m}(\hat{t}), \quad \hat{t} \in \widehat{I},$$

and

$$\|L_{m,z}^{r_m}\|^2_{L^2(I_m;X)} = \frac{2}{k_m}\|\widehat{L}_z^{r_m}\|^2_{L^2(\widehat{I};X)},$$

where F_m is the element mapping in (7).

Proof Let $\widehat{V} \in \mathcal{P}^{r_m}(\widehat{I}; X)$. Define $V \in \mathcal{P}^{r_m}(I_m; X)$ by $V(F_m(\hat{t})) = \widehat{V}(\hat{t})$. Then, the definition of the lifting operators and a change of variables lead to

$$\int_{\widehat{I}} (\widehat{L}_z^{r_m}(\hat{t}), \widehat{V}(\hat{t}))_H \, \mathrm{d}\hat{t} = (z, \widehat{V}(-1))_H = (z, V(t_{m-1}))_H$$
$$= \int_{I_m} (L_{m,z}^{r_m}(t), V(t))_H \, \mathrm{d}t$$
$$= \frac{k_m}{2}\int_{\widehat{I}} (L_{m,z}^{r_m}(F_m(\hat{t})), \widehat{V}(\hat{t}))_H \, \mathrm{d}\hat{t}.$$

Since this holds for any $\widehat{V} \in \mathcal{P}^{r_m}(\widehat{I}; X)$, the first claim follows.

Moreover, this result and the change of variables $t = F_m(\hat{t})$ show that

$$\int_{I_m} \|L^{r_m}_{m,z}(t)\|^2_X \, \mathrm{d}t = \frac{k_m}{2} \int_{\widehat{I}} \|L^{r_m}_{m,z}(F_m(\hat{t}))\|^2_X \, \mathrm{d}\hat{t} = \frac{2}{k_m} \int_{\widehat{I}} \|\widehat{L}^{r_m}_z(\hat{t})\|^2_X \, \mathrm{d}\hat{t}.$$

This completes the proof. □

We are now ready to prove Proposition 2.

Proof of Proposition 2 Consider a function $U \in V^{\boldsymbol{r}}_{\mathrm{dG}}(\mathcal{M}; X)$, a time-step $I_m = (t_{m-1}, t_m)$ and a polynomial degree $r_m \geq 0$. Taking $z = [\![U]\!]_{m-1}$, it follows directly that $\mathsf{L}_m(U) = L^{r_m}_{m,z} = L^{r_m}_{m,[\![U]\!]_{m-1}}$. The claim in Proposition 2 now follows by combining Lemma 6 and Lemma 7. More precisely, we have

$$\|\mathsf{L}_m(U)\|^2_{L^2(I_m;X)} = \|L^{r_m}_{m,z}\|^2_{L^2(I_m;X)} = \tfrac{2}{k_m} \|\widehat{L}^{r_m}_z\|^2_{L^2(\widehat{I};X)} = \tfrac{2}{k_m} \tfrac{(r_m+1)^2}{2} \|[\![U]\!]_{m-1}\|^2_X,$$

which is the identity in Proposition 2. □

4.3 Reconstruction error

In this subsection, we present an *hp*-version analysis for the difference $U - \widehat{U}$ in (33) and prove Theorem 2.

Again, we first consider the unit interval $\widehat{I} = (-1, 1)$. As before, let $r \geq 0$ and $z \in X$ be fixed. Then, motivated by (33), we define

$$\widehat{E}^r_z(t) = -\int_{-1}^{t} \widehat{L}^r_z(\tau) \, \mathrm{d}\tau + z, \quad t \in \widehat{I}, \tag{35}$$

with $\widehat{L}^r_z$ defined in (34).

Lemma 8 *For $r \geq 0$, we have* $\widehat{E}^r_z(t) = \frac{z}{2} \left(\sum_{i=0}^{r+1} b_i \widehat{Q}_i(t) \right)$ *with*

$$b_i = \begin{cases} 1 & \text{if } i = 0, \\ (-1)^i (2i - 1) & \text{if } i \geq 1. \end{cases}$$

Here, $\widehat{Q}_i$ is the i^{th} integrated Legendre polynomial from Lemma 1.

Proof Integrating the representation of $\widehat{L}_z^r$ in Lemma 6 yields

$$\int_{-1}^{t} \widehat{L}_z^r(\tau)\,\mathrm{d}\tau = \frac{z}{2}\left(\sum_{i=0}^{r}(-1)^i(2i+1)\int_{-1}^{t}\widehat{K}_i(\tau)\,\mathrm{d}\tau\right)$$
$$= \frac{z}{2}\left(\sum_{i=1}^{r+1}(-1)^{i-1}(2i-1)\int_{-1}^{t}\widehat{K}_{i-1}(\tau)\,\mathrm{d}\tau\right).$$

Using the first identity in Lemma 1 and the fact that $\widehat{K}_0(t)=1$, we conclude that

$$\int_{-1}^{t} \widehat{L}_z^r(\tau)\,\mathrm{d}\tau = \frac{z}{2}\left(1+t+\sum_{i=2}^{r+1}(-1)^{i-1}(2i-1)\widehat{Q}_i(t)\right).$$

Then,

$$\widehat{E}_z^r(t) = \frac{z}{2}\left(1-t+\sum_{i=2}^{r+1}(-1)^{i}(2i-1)\widehat{Q}_i(t)\right).$$

Since $\widehat{Q}_0(t)=1$ and $\widehat{Q}_1(t)=t$, the claim follows. □

Let us consider the lowest-order cases where $r=0$ and $r=1$, respectively. It can be readily seen that

$$\begin{aligned}\widehat{E}_z^0(t) &= \frac{z}{2}(1-t),\\ \widehat{E}_z^1(t) &= \frac{z}{4}(-1-2t+3t^2).\end{aligned} \tag{36}$$

We then obtain

$$\begin{aligned}\|\widehat{E}_z^0\|^2_{L^2(\widehat{I};X)} &= \frac{\|z\|_X^2}{4}\int_{\widehat{I}}(1-t)^2\,\mathrm{d}t = \frac{2}{3}\|z\|_X^2,\\ \|\widehat{E}_z^1\|^2_{L^2(\widehat{I};X)} &= \frac{\|z\|_X^2}{16}\int_{\widehat{I}}(1-2t+3t^2)^2\,\mathrm{d}t = \frac{4}{15}\|z\|_X^2.\end{aligned} \tag{37}$$

For general $r \geq 0$, the following result holds.

Lemma 9 *For $r \geq 0$, we have*

$$\|\widehat{E}_z^r\|^2_{L^2(\widehat{I};X)} = \frac{2r+2}{(2r+1)(2r+3)}\|z\|_X^2.$$

Proof We start from the representation in Lemma 8, and note that

$$\|\widehat{E}_z^r\|^2_{L^2(\widehat{I};X)} = \frac{\|z\|_X^2}{4}\left(\sum_{i,j=0}^{r+1} b_i b_j \int_{\widehat{I}} \widehat{Q}_i(t)\widehat{Q}_j(t)\,\mathrm{d}t\right), \quad r \geq 0.$$

To simplify notation, let us define

$$T_r = \sum_{i,j=0}^{r+1} b_i b_j M_{i,j}, \quad M_{i,j} = \int_{\widehat{I}} \widehat{Q}_i(t)\widehat{Q}_j(t)\,\mathrm{d}t.$$

Hence, we need to show that

$$T_r = \frac{8r+8}{(2r+1)(2r+3)}. \tag{38}$$

We prove (38) by induction. From (37) we see that

$$T_0 = \frac{8}{3}, \quad T_1 = \frac{16}{15},$$

and hence, relation (38) holds true for $r = 0$ and $r = 1$. Suppose now that it holds for $r \geq 1$. We need to prove that

$$T_{r+1} = \frac{8r+16}{(2r+3)(2r+5)}. \tag{39}$$

To that end, we note that

$$T_{r+1} = \sum_{i,j=0}^{r+2} b_i b_j M_{i,j} = \sum_{i,j=0}^{r+1} b_i b_j M_{i,j} + \sum_{i=0}^{r+1} b_i b_{r+2} M_{i,r+2} + \sum_{j=0}^{r+1} b_{r+2} b_j M_{r+2,j} + b_{r+2}^2 M_{r+2,r+2}.$$

Since $M_{i,j} = M_{j,i}$, we obtain

$$T_{r+1} = T_r + 2b_{r+2}\sum_{i=0}^{r+1} b_i M_{i,r+2} + b_{r+2}^2 M_{r+2,r+2}.$$

From Lemma 1 we conclude that

$$2b_{r+2}\sum_{i=0}^{r+1} b_i M_{i,r+2} = -4b_{r+2}b_r\gamma_{r-1}\gamma_{r+1}\gamma_r,$$

and

$$b_{r+2}^2 M_{r+2,r+2} = 2b_{r+2}^2 \gamma_{r+1}^2 (\gamma_{r+2} + \gamma_r).$$

Therefore, it follows that

$$T_{r+1} = T_r - 4b_{r+2} b_r \gamma_{r-1} \gamma_{r+1} \gamma_r + 2b_{r+2}^2 \gamma_{r+1}^2 (\gamma_{r+2} + \gamma_r).$$

Using the induction assumption that (38) holds, and the definition of γ_i, we obtain

$$\begin{aligned} T_{r+1} = {} & \frac{8r+8}{(2r+1)(2r+3)} - \frac{4b_{r+2} b_r}{(2r-1)(2r+1)(2r+3)} \\ & + \frac{4b_{r+2}^2}{(2r+1)(2r+3)(2r+5)}. \end{aligned}$$

Note that, from Lemma 8, we have $b_r = (-1)^r (2r-1)$. Using this identity and elementary manipulations yields (39). □

Next, we scale the results above to an interval $I_m = (t_{m-1}, t_m)$ of length k_m and a polynomial degree $r_m \geq 0$. Similarly to (35), we define the function $E_{m,z}^{r_m}$ on I_m by

$$E_{m,z}^{r_m}(t) = - \int_{t_{m-1}}^{t} L_{m,z}^{r_m}(\tau)\, \mathrm{d}\tau + z, \quad t \in I_m,$$

with $L_{m,z}^{r_m}$ defined above. Recalling that F_m is the elemental mapping in (7), we have the following identity.

Lemma 10 *There holds*

$$E_{m,z}^{r_m}(F_m(\hat{t})) = \widehat{E}_z^{r_m}(\hat{t}), \quad \hat{t} \in \widehat{I}.$$

Moreover,

$$\| E_{m,z}^{r_m} \|_{L^2(I_m;X)}^2 = \frac{k_m}{2} \| \widehat{E}_z^{r_m} \|_{L^2(\widehat{I};X)}^2.$$

Proof Changing variables and the relation in Lemma 7 yield

$$\begin{aligned} E^{r_m}_{m,z}(F_m(\hat{t})) &= -\int_{F_m(-1)}^{F_m(\hat{t})} L^{r_m}_{m,z}(\tau)\,\mathrm{d}\tau + z \\ &= -\frac{k_m}{2}\int_{-1}^{\hat{t}} L^{r_m}_{m,z}(F_m(\hat{\tau}))\,\mathrm{d}\hat{\tau} + z \\ &= -\int_{-1}^{\hat{t}} \widehat{L}^{r_m}_{z}(\hat{\tau})\,\mathrm{d}\hat{\tau} + z = \widehat{E}^{r_m}_{z}(\hat{t}). \end{aligned}$$

This shows the first claim.

Furthermore, using this identity, we have that

$$\begin{aligned} \int_{I_m} \|E^{r_m}_{m,z}(t)\|^2_X\,\mathrm{d}t &= \frac{k_m}{2}\int_{\widehat{I}} \|E^{r_m}_{m,z}(F_m(\hat{t}))\|^2_X\,\mathrm{d}\hat{t} \\ &= \frac{k_m}{2}\int_{\widehat{I}} \|\widehat{E}^{r_m}_{z}(\hat{t}))\|^2_X\,\mathrm{d}\hat{t}. \end{aligned}$$

This completes the proof. □

Let us now turn to the proof of Theorem 2.

Proof of Theorem 2 Let $U \in V^{\boldsymbol{r}}_{\mathrm{dG}}(\mathcal{M}; X)$. Consider the interval $I_m = (t_{m-1}, t_m)$ and the polynomial degree $r_m \geq 0$. Setting $z = [\![U]\!]_{m-1}$, we conclude from equations (31), (33) and the definition of $E^{r_m}_{m,z}$ that

$$U(t) - \widehat{U}(t) = E^{r_m}_{m,z}(t) = E^{r_m}_{m,[\![U]\!]_{m-1}}(t), \quad t \in I_m.$$

Hence, by using Lemma 9 and Lemma 10, we obtain

$$\begin{aligned} \|U - \widehat{U}\|^2_{L^2(I_m;X)} = \|E^{r_m}_{m,z}\|^2_{L^2(I_m;X)} &= \frac{k_m}{2}\|\widehat{E}^{r_m}_{z}\|^2_{L^2(\widehat{I};X)} \\ &= \frac{k_m}{2}\frac{2r_m+2}{(2r_m+1)(2r_m+3)}\|[\![U]\!]_{m-1}\|^2_X. \end{aligned}$$

This shows Theorem 2. □

5 Numerical experiments

In this section we shall illustrate the performance of the error estimators η_{cG} and η_{dG} from Theorem 1 and Theorem 3 for the cG and dG schemes, respectively, within an

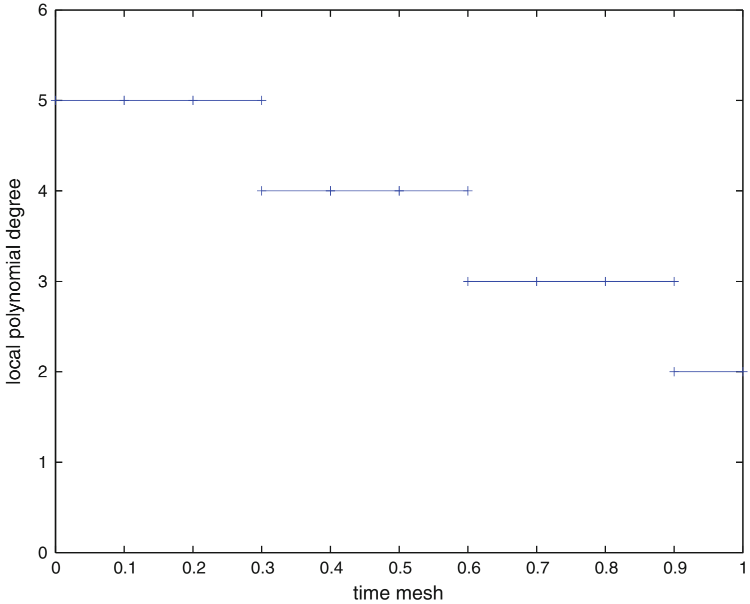


Fig. 1 Example 1: Adaptively refined time mesh and polynomial degrees for the dG method

hp-adaptive refinement algorithm. Specifically, given the numerical solution on the entire interval $(0, T)$, we use the local error estimators $\eta_{\mathrm{cG},m}$ and $\eta_{\mathrm{dG},m}$ in (18) and (29), respectively, to identify those time-steps on which large errors occur. More precisely, an element I_m is marked for refinement if $\eta_{\mathrm{cG},m} > \theta \max_i \eta_{\mathrm{cG},i}$, respectively $\eta_{\mathrm{dG},m} > \theta \max_i \eta_{\mathrm{dG},i}$. Here, θ is a parameter which we set to be 0.5 in all of our numerical experiments. Subsequently, for each of the marked elements a decision is made whether h- or p-refinement is applied. To that end, we use a local smoothness estimation technique as presented in, e.g., [12]. Starting from a coarse initial time partition and a low-order polynomial degree vector, this procedure is repeated until a given tolerance τ is met.

Our goal is to show that, for the examples considered, the error estimators for the cG and dG approximations over the whole interval $(0, T)$ decay at the same (asymptotic) rate as the actual error measures in (17) and (28), respectively, and that exponential convergence is achieved even for solutions with start-up singularities. For detailed implementational techniques for *hp*-version dG time-stepping methods we refer to [24,29]. We discretize our model problems using a high-order finite element method in space so that the spatial errors are dominated by the errors resulting from the cG and dG time discretizations.

For our numerical experiments, we shall consider the homogeneous heat equation in one space dimension,

$$u_t(x,t) - u_{xx}(x,t) = 0, \quad (x,t) \in (0,1) \times (0,1), \tag{40}$$

subject to Dirichlet boundary conditions

$$u(x,t) = 0 \quad (x,t) \in \{0,1\} \times (0,1),$$

and the initial condition

$$u(x,0) = u_0(x) \quad x \in (0,1).$$

Here, the Hilbert spaces from Sect. 1 are the standard Sobolev spaces of order zero and one: $H = L^2(0,1)$ and $X = H_0^1(0,1)$. Furthermore, the elliptic operator is $\mathsf{A} = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}$ (in the weak sense). Hence, $\alpha = \beta = 1$ in (2) and (3), respectively.

We will investigate the above problem numerically for different choices of u_0. Specifically, we shall look at

$$\begin{aligned} &\text{Example 1}: \quad u_0^{(1)}(x) = \sin(\pi x),\\ &\text{Example 2}: \quad u_0^{(2)}(x) = x(1-x),\\ &\text{Example 3}: \quad u_0^{(3)}(x) = 1. \end{aligned}$$

We denote by $u^{(i)}$ the solution corresponding to the initial data $u_0^{(i)}$, $i = 1,2,3$. The solution $u^{(1)}$ is given by $u^{(1)}(x,t) = e^{-\pi^2 t}\sin(\pi x)$; it is arbitrarily smooth in x and t. The solutions $u^{(2)}$ and $u^{(3)}$ can be readily expressed in terms of Fourier series; they both have start-up singularities at $t = 0$. More precisely, it can be seen

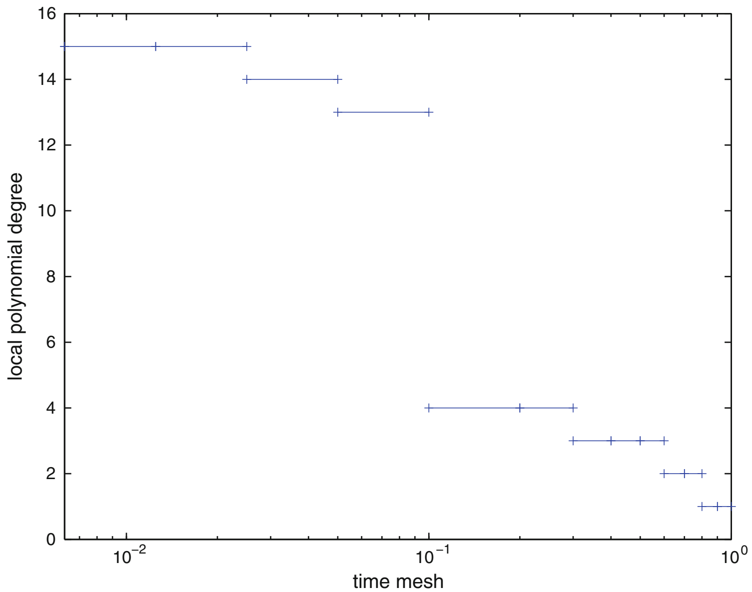


Fig. 2 Example 2: Adaptively refined time mesh and polynomial degrees for the dG method

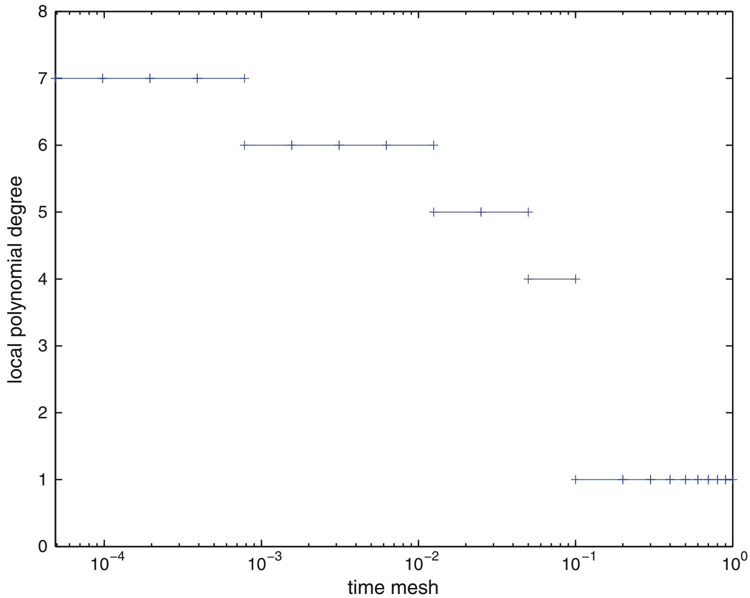


Fig. 3 Example 3: Adaptively refined time mesh and polynomial degrees for the dG method

(see [24, p. 868]) that

$$u^{(2)} \in H^{\frac{5}{4}-\varepsilon}\left((0,1); H_0^1(0,1)\right), \quad u^{(3)} \in H^{\frac{1}{4}-\varepsilon}\left((0,1); H_0^1(0,1)\right),$$

for any $\varepsilon > 0$. Here, $H^s((0,1); H_0^1(0,1))$ denotes the Sobolev space of order s of functions on $(0,1)$ with values in $H_0^1(0,1)$.

In Figs. 1, 2 and 3 we plot the time meshes and local polynomial degrees resulting from the hp-adaptive dG time discretizations of Examples 1–3. The error tolerance is chosen to be $\tau = 10^{-6}$ for Example 1, $\tau = 10^{-5}$ for Example 2, and $\tau = 10^{-1}$ for Example 3. Furthermore, in Examples 1 and 2, we use 4 elements in space and a uniform spatial polynomial degree of 10. For Example 3, geometric mesh refinement in space was applied (with a theoretically optimal factor of 0.17) to appropriately resolve the incompatibility of the initial datum at $x = 0$ and $x = 1$; cf. [26]. The horizontal axis represents the time partition, and on the vertical axis, the local polynomial degrees are displayed. As expected, the smooth solution in Example 1 is approximated on a uniform time mesh, and higher polynomial degrees are used. For Examples 2 and 3, due to the singular behavior of the solution at the start, the time partition is geometrically refined at $t = 0$ (note that we have used a logarithmic scale on the horizontal axis). This is in correspondence with the a priori error analysis in [24]. Furthermore, we notice that the polynomial degrees tend to decrease away from $t = 0$, which is due to the fact that the right-hand side of (40) is zero, and hence, the solution is flattened out quickly. Similar results (particularly for Examples 1 and 2) are obtained for the cG method.

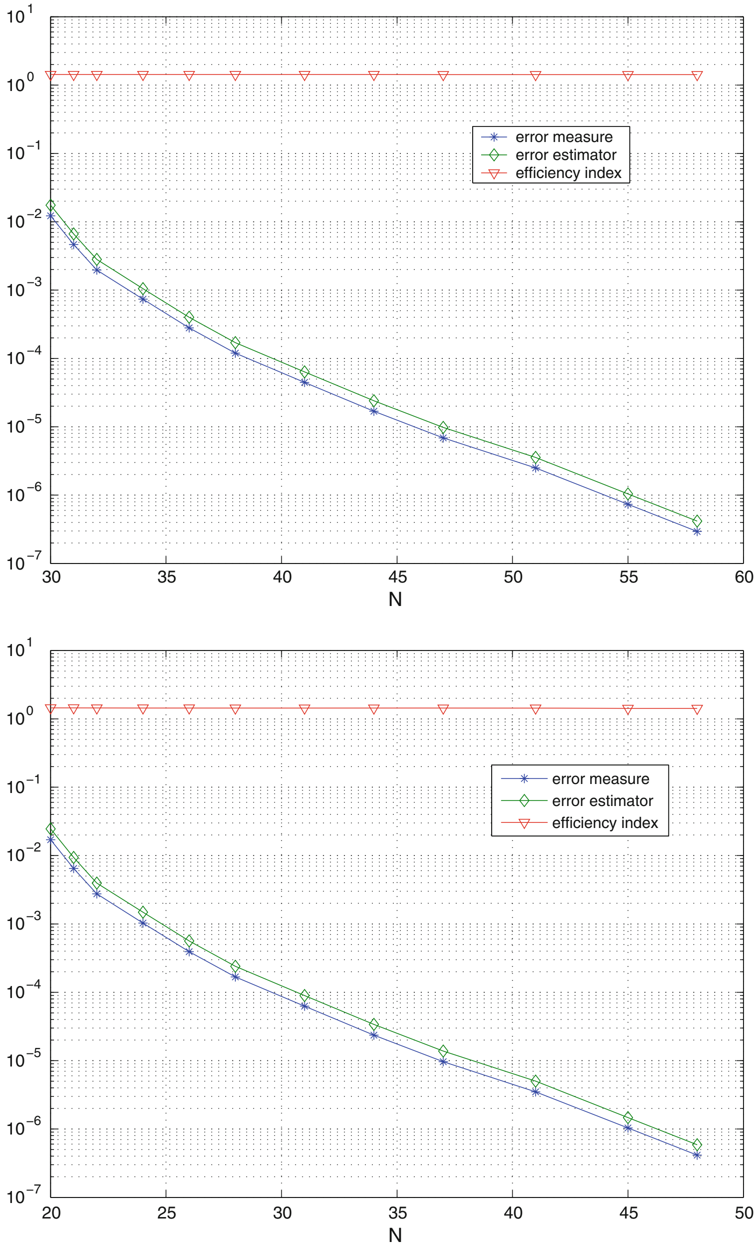
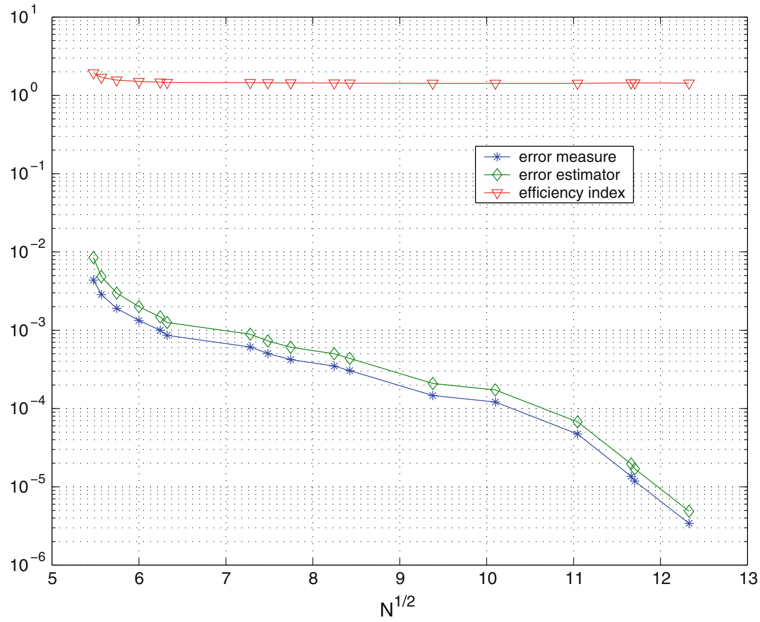


Fig. 4 Example 1: Performance of cG (*top*) and dG (*bottom*)

Figures 4, 5 and 6 display the behavior of the errors (i.e., the error measures in (17) and (28)) and the error estimators η_{cG} and η_{dG} of the cG and dG time discretizations for Examples 1–3. Here, the numerical solutions are compared to the known exact solutions, and integrals are computed using Gaussian quadrature of sufficiently high



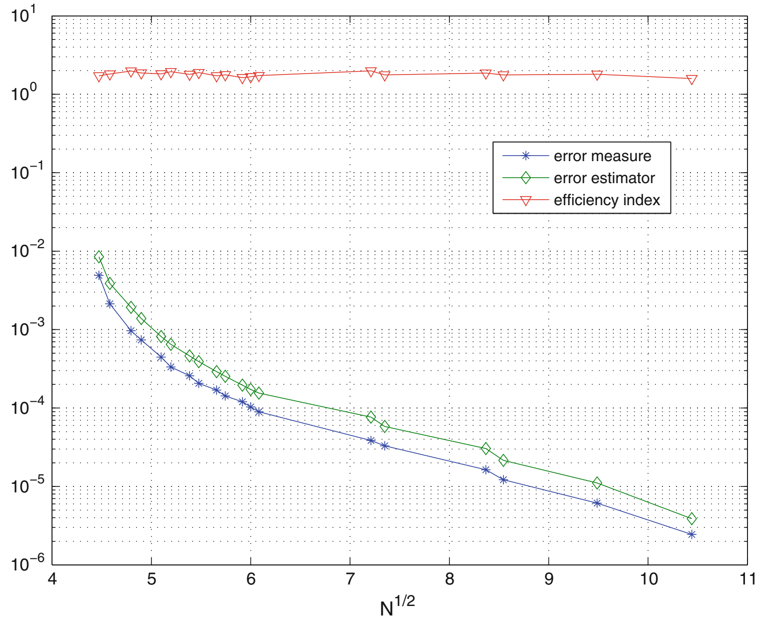


Fig. 5 Example 2: Performance of cG (*top*) and dG (*bottom*)

order. In addition, the efficiency indices, i.e., the ratio between the error estimators and the actual errors, are shown. All graphs are presented in a semi-logarithmic coordinate system. The horizontal coordinate axes correspond to the number of degrees

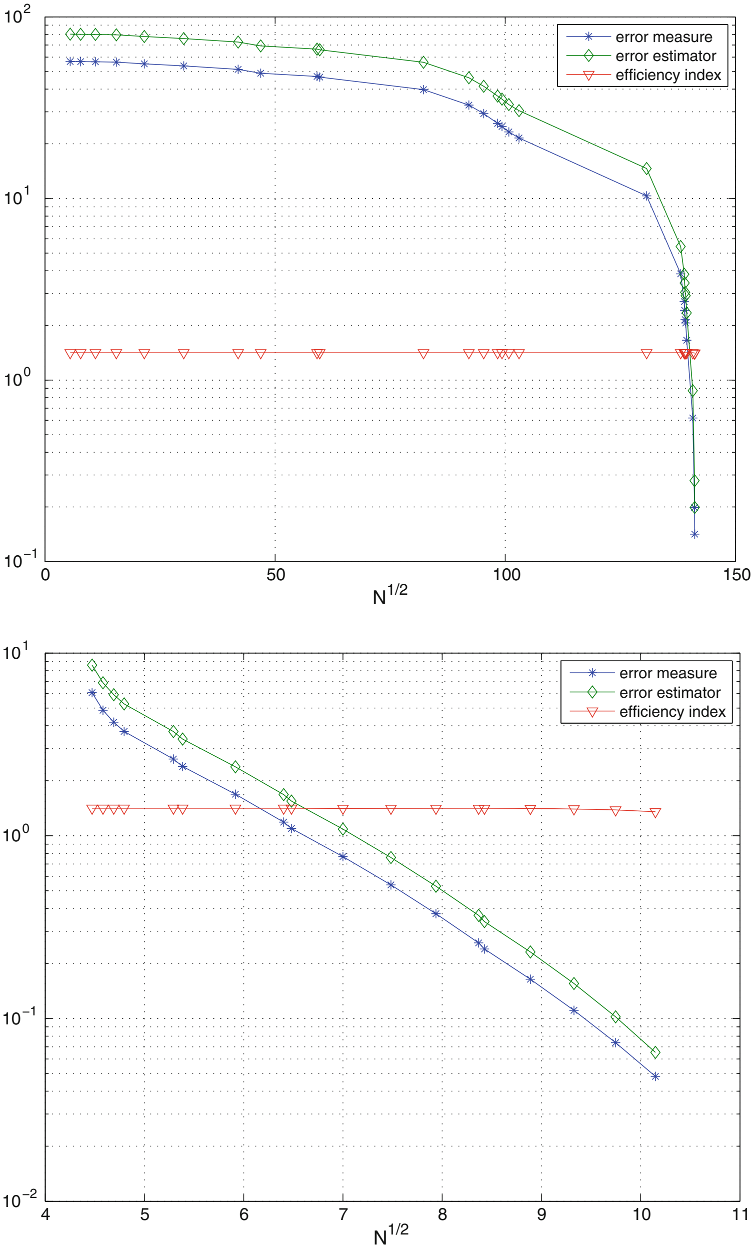


Fig. 6 Example 3: Performance of cG (*top*) and dG (*bottom*)

of freedom N in the time discretization (more precisely, N for Example 1, and $N^{\frac{1}{2}}$ for Examples 2 and 3; cf. [24]). We see that the errors decay exponentially. Moreover, the efficiency indices are consistently between 1 and 2, thereby indicating the sharpness and asymptotic correctness of the estimators for the given examples. The slow initial decay of the errors for the cG method in Example 3 may be explained by the fact that the cG time discretization is less dissipative than the dG scheme, and therefore, large errors at $\{x = 0, 1\} \times \{t = 0\}$ caused by the incompatibility of the initial conditions might be smoothed out at a comparatively low rate during the initial refinements.

6 Concluding remarks

In this paper, we have presented an hp-version a posteriori error analysis for the continuous and discontinuous Galerkin time-stepping schemes for linear parabolic PDEs. The resulting error estimators are reliable and efficient, and all constants are given fully explicitly in terms of the step sizes, the local polynomial degrees, and the continuity and coercivity constants of the spatial operator. One of the important components in finding the a posteriori error estimates is to rewrite the variational formulations of the cG and dG schemes in a strong form as previously presented in [19] for the (h-version) dG time-stepping method. Furthermore, a careful investigation of the estimators' dependence on the local polynomial degrees using suitable polynomial bases was required. Future work includes the numerical analysis of discretizations of parabolic PDEs that are fully hp-adaptive in time and space, as well as extensions to nonlinear parabolic PDEs.

References

1. Abramowitz, M., Stegun, I.A. (eds.): Handbook of Mathematical Functions, 9th edn. Dover, New York (1972)
2. Arnold, D.N., Brezzi, F., Cockburn, B., Marini, L.D.: Unified analysis of discontinuous Galerkin methods for elliptic problems. SIAM J. Numer. Anal. **39**, 1749–1779 (2001)
3. Bangerth, W., Rannacher, R.: Adaptive finite element methods for differential equations. Lectures in Mathematics, ETH Zürich. Birkhäuser, Basel (2003)
4. Brunner, H., Schötzau, D.: hp-discontinuous Galerkin time-stepping for Volterra integrodifferential equations. SIAM J. Numer. Anal. **44**, 224–245 (2006)
5. Delfour, M., Hager, W., Trochu, F.: Discontinuous Galerkin methods for ordinary differential equations. Math. Comput. **36**, 455–473 (1981)
6. Eriksson, K., Johnson, C.: Adaptive finite element methods for parabolic problems I: a linear model problem. SIAM J. Numer. Anal. **28**, 43–77 (1991)
7. Eriksson, K., Johnson, C.: 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**, 706–740 (1995)
8. Eriksson, K., Johnson, C.: Adaptive finite element methods for parabolic problems IV: a nonlinear model problem. SIAM J. Numer. Anal. **32**, 1729–1749 (1995)
9. Eriksson, K., Johnson, C.: Adaptive finite element methods for parabolic problems V: long-time integration. SIAM J. Numer. Anal. **32**, 1750–1763 (1995)
10. Estep, D.: A posteriori error bounds, global error control for approximation of ordinary differential equations. SIAM J. Numer. Anal. **32**, 1–48 (1995)
11. Estep, D., French, D.: Global error control for the continuous Galerkin finite element method for ordinary differential equations. RAIRO Modél. Math. Anal. Numér. **28**, 815–852 (1994)
12. Houston, P., Süli, E.: A note on the design of hp-adaptive finite element methods for elliptic partial differential equations. Comput. Methods Appl. Mech. Eng. **194**, 229–243 (2005)
13. Hulme, B.L.: Discrete Galerkin and related one-step methods for ordinary differential equations. Math. Comput. **26**, 881–891 (1972)

14. Hulme, B.L.: One-step piecewise Galerkin methods for initial value problems. Math. Comput. **26**, 415–425 (1972)
15. Jamet, P.: Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain. SIAM J. Numer. Anal. **15**, 912–928 (1978)
16. Johnson, C.: Error estimates and adaptive time-step control for a class of one-step methods for stiff ordinary differential equations. SIAM J. Numer. Anal. **25**, 908–926 (1988)
17. Lesaint, P., Raviart, P.A.: On a finite element method for solving the neutron transport equation. In: de Boor, C. (ed.) Mathematical Aspects of Finite Elements in Partial Differential Equations, pp. 89–145. Academic Press, New York (1974)
18. Lions, J.L., Magenes, E.: Non-Homogeneous Boundary Value Problems and Applications, vol. 1. Springer, New York (1972)
19. Makridakis, C., Nochetto, R.H.: A posteriori error analysis for higher order dissipative methods for parabolic problems. Numer. Math. **104**, 489–514 (2006)
20. Matache, A.-M., Schwab, C., Wihler, T.P.: Fast numerical solution of parabolic integrodifferential equations with applications in finance. SIAM J. Sci. Comput. **27**, 369–393 (2005)
21. Matache, A.-M., Schwab, C., Wihler, T.P.: Linear complexity solution of parabolic integro-differential equations. Numer. Math. **104**, 69–102 (2006)
22. Reed, W.H., Hill, T.R.: Triangular mesh methods for the neutron transport equation. Report LA-UR-73-479, Los Alamos Scientific Laboratory (1973)
23. Schötzau, D., Schwab, C.: An *hp* a-priori error analysis of the DG time-stepping method for initial value problems. Calcolo **37**, 207–232 (2000)
24. Schötzau, D., Schwab, C.: Time discretization of parabolic problems by the *hp*-version of the discontinuous Galerkin finite element method. SIAM J. Numer. Anal. **38**, 837–875 (2000)
25. Schötzau, D., Schwab, C.: *hp*-discontinuous Galerkin time-stepping for parabolic problems. C. R. Acad. Sci. Paris I **333**, 1121–1126 (2001)
26. Schwab, C.: *p*- and *hp*-Finite Element Methods. Oxford University Press, Oxford (1998)
27. Szabó, B., Babuška, I.: Finite Element Analysis. Wiley, New York (1991)
28. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems, 2nd edn. Springer, Berlin (1997)
29. von Petersdorff, T., Schwab, C.: Numerical solution of parabolic equations in high dimensions. Math. Model. Anal. Numer. **38**, 93–127 (2004)
30. Werder, T., Gerdes, K., Schötzau, D., Schwab, C.: *hp*-discontinuous Galerkin time-stepping for parabolic problems. Comput. Methods Appl. Mech. Eng. **190**, 6685–6708 (2001)
31. Wihler, T.P.: An a-priori error analysis of the *hp*-version of the continuous Galerkin FEM for nonlinear initial value problems. J. Sci. Comput. **25**, 523–549 (2005)