

## A convergent time–space adaptive dG(s) finite element method for parabolic problems motivated by equal error distribution

FERNANDO D. GASPOZ\* AND KUNIBERT SIEBERT

*Institut für Angewandte Analysis und Numerische Simulation, Fachbereich Mathematik,  
Universität Stuttgart, Pfaffenwaldring 57, D-70569 Stuttgart, Germany*

\*Corresponding author: fernando.gaspoz@ians.uni-stuttgart.de  
kg.siebert@ians.uni-stuttgart.de

CHRISTIAN KREUZER

*Lehrstuhl LSX, Fakultät für Mathematik, Technische Universität Dortmund,  
Vogelpothsweg 87, D-44227 Dortmund, Germany  
Christian.Kreuzer@math.tu-dortmund.de*

AND

DANIEL A. ZIEGLER

*Institut für Angewandte und Numerische Mathematik, Karlsruher Institut für Technologie,  
Englerstrasse 2, D-76131 Karlsruhe, Germany  
daniel.ziegler@kit.edu*

[Received on 21 October 2016; revised on 13 September 2017]

We shall develop a fully discrete space-time adaptive method for linear parabolic problems based on new, reliable and efficient *a posteriori* analysis for higher order dG(s) finite element discretisations. The adaptive strategy is motivated by the principle of equally distributing the *a posteriori* indicators in time and the convergence of the method is guaranteed by the uniform energy estimate from Kreuzer *et al.* (2012, Design and convergence analysis for an adaptive discretization of the heat equation. *IMA J. Numer. Anal.*, **32**, 1375–1403) under minimal assumptions on the regularity of the data.

*Keywords:* adaptive finite elements; parabolic problems; convergence.

### 1. Introduction

Let  $\Omega$  be a bounded polyhedral domain in  $\mathbb{R}^d$ ,  $d \in \mathbb{N}$ . We consider the linear parabolic partial differential equation

$$\begin{aligned} \partial_t u + \mathcal{L}u &= f && \text{in } \Omega \times (0, T), \\ u &= 0 && \text{on } \partial\Omega \times (0, T), \\ u(\cdot, 0) &= u_0 && \text{in } \Omega. \end{aligned} \tag{1.1}$$

Hereafter,  $\mathcal{L}u = -\operatorname{div}\mathbf{A}\nabla u + cu$  is a second-order elliptic operator with respect to space and  $\partial_t u = \frac{\partial u}{\partial t}$  denotes the partial derivative with respect to time. In the simplest setting  $\mathcal{L} = -\Delta$ , whence (1.1) is the heat equation. Precise assumptions on data are provided in Section 2.1.

The objective of this paper is the design and a detailed convergence analysis of an efficient adaptive finite element method for solving (1.1) numerically. To this end we resort to adaptive finite elements in space combined with a discontinuous Galerkin dG( $s$ ) time-stepping scheme in Section 2.2, compared with Thomée (2006). The conforming finite element spaces are continuous piecewise polynomials of fixed degree over a simplicial triangulation of the domain  $\Omega$ . In each single we reduce or enlarge the local time-step size and refine and coarsen the underlying triangulation.

The adaptive decisions are based on *a posteriori* error indicators. Numerous such estimators for various error notions are available in the literature. In fact, Eriksson & Johnson (1991, 1995) have analysed adaptive dG( $s$ ) methods for parabolic problems in their studies providing *a priori* and *a posteriori* bounds for the  $L^\infty$  ( $L^2$ ) error using duality techniques. Lakkis & Makridakis (2006) proved *a posteriori* bounds based on the elliptic reconstruction technique, which was introduced in the study by Makridakis & Nochetto (2003) in the semidiscrete context. The  $L^2$  ( $H^1$ ), respectively  $L^2$  ( $H^1$ )  $\cap$   $H^1$  ( $H^{-1}$ ), error bounds in the studies by Picasso (1998) and Verfürth (2003) are based on energy techniques and have been used with a dG(0) time-stepping scheme in the adaptive methods and convergence analysis presented in the studies by Chen & Feng (2004) and Kreuzer et al. (2012). For our purpose we generalise the residual-based estimator (Verfürth, 2003) to higher-order dG( $s$ ) schemes in Section 3. The estimator is built from five indicators: an indicator for the initial error, indicators for the temporal and spatial errors, a coarsening error indicator and an indicator controlling the so-called consistency error. It is important to notice that besides the first indicator all other indicators accumulate in  $L^2$  in time. The adaptation of the time-step size uses information of the indicators for the temporal error and the consistency error. The adaptation of the spatial triangulation is based on refinement by bisection using information from the indicators for the spatial error and for the coarsening error. Very recently, an independently developed guaranteed *a posteriori* estimator for higher-order dG( $s$ ) schemes was provided in the study by Ern et al. (2017) using equilibrated flux-based bounds for the spatial error.

By now the convergence and optimality of adaptive methods for stationary inf–sup stable, respectively coercive, problems are well established (Dörfler, 1996; Morin et al., 2000, 2002; Binev et al., 2004; Chen & Feng, 2004; Mekchay & Nochetto, 2005; Stevenson, 2007; Cascon et al., 2008; Diening & Kreuzer, 2008; Morin et al., 2008; Kreuzer & Siebert, 2011; Siebert, 2011; Diening et al., 2016); compare also with the overview article by Nochetto et al. (2009). The essential design principle motivating the adaptive strategies in most of the above methods is the equal distribution of the error. The importance of this principle is highlighted by the near characterisations of nonlinear approximation classes with the help of a thresholding algorithm in the studies by Binev et al. (2002) and Gaspoz & Morin (2014).

In contrast to the situation for above-mentioned problems the convergence analysis of adaptive approximation of time-dependent problems is still in its infancy. In the study by Schwab & Stevenson (2009) optimal computational complexity of an adaptive wavelet method for parabolic problems is proved using a symmetric and coercive discretisation based on a least squares formulation. To the best of our knowledge there exist only two results (Chen & Feng, 2004; Kreuzer et al., 2012) concerned with a rigorous convergence analysis of time-space adaptive finite element methods. In the study by Chen & Feng (2004) it is proved for the dG(0) time-stepping scheme that each single terminates and that the error of the computed approximation is below a prescribed tolerance when the final time is reached. This, however, is not guaranteed and thus theoretically the adaptively generated sequence of time instances  $\{t_n\}_{n \geq 0}$  may not be finite and such that  $t_n \rightarrow t_* < T$  as  $n \rightarrow \infty$ . This drawback has been overcome in the study by Kreuzer et al. (2012) with the help of an *a priori* computed minimal time-step

size in terms of the right-hand side  $f$  and the discrete initial value  $U_0$ . However, the design of neither of the two methods heeds the principle of equally distributing the error. Let us shed some light on this fact with the help of the initial value problem

$$\partial_t u + u = f \quad \text{in } (0, T) \quad \text{and} \quad u(0) = u_0.$$

Let  $0 = t_0 < t_1 < \dots < t_N = T$  be some partition of  $(0, T)$ . Using the dG(0) time-stepping scheme we obtain  $\{U_n\}_{n=0}^N$  such that

$$\frac{U_n - U_{n-1}}{\tau_n} + U_n = f_n := \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} f \, dt, \quad n = 1, \dots, N, \quad U_0 = u_0,$$

where  $\tau_n = t_n - t_{n-1}$ . Let  $\mathcal{U}$  be the piecewise affine interpolation  $\mathcal{U}$  of the nodal values  $\{U_n\}_{n=0}^N$ . Then we have with Young's inequality that

$$\begin{aligned} \int_0^T \frac{1}{2} \partial_t |u - \mathcal{U}|^2 + |u - \mathcal{U}|^2 \, dt &= \sum_{n=1}^N \int_{t_{n-1}}^{t_n} (f - f_n)(u - \mathcal{U}) + (U_n - \mathcal{U})(u - \mathcal{U}) \, dt \\ &\leqslant \sum_{n=1}^N \int_{t_{n-1}}^{t_n} |f - f_n|^2 + |U_n - \mathcal{U}|^2 + \frac{1}{2} |u - \mathcal{U}|^2 \, dt. \end{aligned}$$

A simple computation reveals  $\int_{t_{n-1}}^{t_n} |U_n - \mathcal{U}|^2 \, dt = \frac{1}{3} \tau_n |U_n - U_{n-1}|^2$ . This term and  $\int_{t_{n-1}}^{t_n} |f - f_n|^2 \, dt$  are the so-called time and consistency *a posteriori* indicators. In order to illustrate the basic differences in the design of the adaptive schemes we shall concentrate on the time indicator. In the studies by [Chen & Feng \(2004\)](#) and [Kreuzer et al. \(2012\)](#) the partition is constructed such that

$$|U_n - U_{n-1}|^2 \leq \frac{\text{TOL}^2}{T}, \quad \text{which implies} \quad \sum_{n=1}^N \tau_n |U_n - U_{n-1}|^2 \leqslant \sum_{n=1}^N \tau_n \frac{\text{TOL}^2}{T} = \text{TOL}^2,$$

i.e., the accumulated indicator is below the prescribed tolerance  $\text{TOL}$ . We call this the  $L^\infty$ -strategy and remark that it does not aim to equally distribute the local indicators. In contrast to this we shall use the  $L^2$ -strategy

$$\tau_n |U_n - U_{n-1}|^2 \leq \text{TOL}^2.$$

Thanks to the uniform energy bound

$$\sum_{n=1}^N |U_n - U_{n-1}|^2 \leqslant \int_0^T |f|^2 \, dt + |U_0|^2 \tag{1.2}$$

(see Corollary 3.11 below) we conclude then that

$$\begin{aligned} \sum_{n=1}^N \tau_n |U_n - U_{n-1}|^2 &= \sum_{\tau_n \leq \delta}^N \tau_n |U_n - U_{n-1}|^2 + \sum_{\tau_n > \delta}^N \tau_n |U_n - U_{n-1}|^2 \\ &\leq \delta \left( \int_0^T |f|^2 dt + |U_0|^2 \right) + \frac{T}{\delta} \text{tol}^2 = T^{1/2} \left( \int_0^T |f|^2 dt + |U_0|^2 \right)^{1/2} \text{tol}, \end{aligned}$$

where  $\delta = \text{tol} T^{1/2} / \left( \int_0^T |f|^2 dt + |U_0|^2 \right)^{1/2}$ . Taking  $\text{tol} = \text{TOL}^2 / \delta$  guarantees that the error is below the prescribed tolerance  $\text{TOL}$ .

These arguments directly generalise to semidiscretisations of (1.1) in time. In the case of a full space-time discretisation of (1.1) additional indicators regarding the space discretisation are involved. We shall roughly explain the principle adaptive strategies of algorithm TAFEM for (1.1). In a preprocessing step an  $L^2$ -tolerance for handling the consistency error is computed. Then in each we first control the consistency error with an  $L^2$ -strategy since this does not require solving expensive linear systems. Second, we control the time indicator using an  $L^2$ -strategy and (1.2). Unfortunately, for the spatial indicator a control similar to (1.2) is not available; therefore, we enforce in a third step that the spatial indicators are bounded by the time or the consistency indicator. In the case where these indicators are equally distributed in time and no massive undershooting appears, this likely results also in an equal distribution of the spatial indicators. In order to handle the other case we include the  $L^\infty$ -strategy from the studies by Chen & Feng (2004) and Kreuzer et al. (2012) for the spatial indicators as a backup strategy. The detailed algorithm TAFEM is presented in Section 4 and its convergence analysis is given in Section 5.

The advantage of our new approach over the algorithms in Chen & Feng (2004) and Kreuzer et al. (2012) is twofold. First, from the fact that TAFEM aims for an equal distribution of the error, we expect an improved performance. Second, we use an  $L^2$ -strategy for the consistency error, which requires only  $L^2$ -regularity of  $f$  in time instead of the  $H^1$ -regularity needed for the  $L^\infty$ -strategy in the studies by Chen & Feng (2004) and Kreuzer et al. (2012). This makes the proposed method suitable for problems where the existing approaches may even fail completely. We conclude the paper in Section 6 with comments on the implementation in DUNE (Blatt et al., 2016) and some numerical experiments. The experiments confirm the expectations and show that the performance of our algorithm TAFEM is more than competitive.

## 2. The continuous and discrete problems

In this section we state the weak formulation of the continuous problem together with the assumptions on data. Then the discretisation by adaptive finite elements in space combined with the dG(s) scheme in time is introduced.

### 2.1 The weak formulation

For  $d \in \mathbb{N}$ , let  $\Omega \subset \mathbb{R}^d$  be a bounded, polyhedral domain that is meshed by some conforming simplicial mesh  $\mathcal{G}_{\text{init}}$ . We denote by  $H^1(\Omega)$  the Sobolev space of square integrable functions  $L^2(\Omega)$  whose first derivatives are in  $L^2(\Omega)$  and we let  $\mathbb{V} := H_0^1(\Omega)$  be the space of functions in  $H^1(\Omega)$  with vanishing

trace on  $\partial\Omega$ . For any measurable set  $\omega$  and  $k \in \mathbb{N}$  we denote by  $\|\cdot\|_\omega$  the  $L^2(\omega; \mathbb{R}^k)$ -norm, whence  $\|v\|_{H^1(\Omega)}^2 = \|v\|_\Omega^2 + \|\nabla v\|_\Omega^2$ .

We suppose that the data of (1.1) have the following properties:  $\mathbf{A}: \Omega \rightarrow \mathbb{R}^{d \times d}$  is piecewise Lipschitz over  $\mathcal{G}_{\text{init}}$  and is symmetric positive definite with eigenvalues  $0 < a_* \leq a^* < \infty$ , i.e.,

$$a_* |\xi|^2 \leq \mathbf{A}(x) \xi \cdot \xi \leq a^* |\xi|^2 \quad \text{for all } \xi \in \mathbb{R}^d, x \in \Omega; \quad (2.1)$$

$c \in L^\infty(\Omega)$  is non-negative, i.e.,  $c \geq 0$  in  $\Omega$ ;  $f \in L^2((0, T); L^2(\Omega)) = L^2(\Omega \times (0, T))$  and  $u_0 \in L^2(\Omega)$ .

We next turn to the weak formulation of (1.1); compare with Evans (2010, Chap. 7). We let  $\mathcal{B}: \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}$  be the symmetric bilinear form associated to the weak form of the elliptic operator  $\mathcal{L}$ , i.e.,

$$\mathcal{B}[w, v] := \int_{\Omega} \mathbf{A} \nabla v \cdot \nabla w + c v w \, dV \quad \text{for all } v, w \in \mathbb{V}.$$

Recalling the Poincaré–Friedrichs inequality  $\|v\|_\Omega \leq C_F \|\nabla v\|_\Omega$  for all  $v \in \mathbb{V}$  with  $C_F = C_F(d, \Omega)$  (Gilbarg & Trudinger, 2001, p. 158) we deduce from (2.1) that  $\mathcal{B}$  is a scalar product on  $\mathbb{V}$  with induced norm

$$\|v\|^2 := \mathcal{B}[v, v] = \int_{\Omega} \mathbf{A} \nabla v \cdot \nabla v + c v^2 \, dV \quad \text{for all } v \in H_0^1(\Omega).$$

This *energy norm* is equivalent to the  $H^1$ -norm  $\|\cdot\|_{H^1(\Omega)}$  and we shall use the energy norm in the subsequent analysis. We denote the restriction of the energy norm to some subset  $\omega \subset \Omega$  by  $\|\cdot\|_\omega$  and let  $\mathbb{V}^* := H^{-1}(\Omega)$  be the dual space of  $H_0^1(\Omega)$  equipped with the operator norm  $\|g\|_* := \sup_{v \in \mathbb{V}} \frac{\langle g, v \rangle}{\|v\|_\Omega}$ , where  $\langle \cdot, \cdot \rangle$  denotes the usual duality bracket in  $H^{-1}(\Omega) \times H_0^1(\Omega)$ .

The weak solution space

$$\mathbb{W}(0, T) := \left\{ u \in L^2(0, T; \mathbb{V}) \mid \partial_t u \in L^2(0, T; \mathbb{V}^*) \right\}$$

is a Banach space endowed with the norm

$$\|v\|_{\mathbb{W}(0, T)}^2 = \int_0^T \|\partial_t v\|_*^2 + \|v\|_\Omega^2 \, dt + \|v(T)\|_\Omega^2, \quad v \in \mathbb{W}(0, T).$$

Moreover, it is continuously embedded into  $C^0([0, T]; L^2(\Omega))$ ; see, e.g., Evans (2010, Chap. 5).

After this preparation we are in a position to state the weak formulation of (1.1): a function  $u \in \mathbb{W}(0, T)$  is a weak solution to (1.1) if it satisfies

$$\langle \partial_t u(t), v \rangle + \mathcal{B}[u(t), v] = \langle f(t), v \rangle_\Omega \quad \text{for all } v \in \mathbb{V}, \text{ a.e. } t \in (0, T), \quad (2.2a)$$

$$u(0) = u_0. \quad (2.2b)$$

Hereafter,  $\langle \cdot, \cdot \rangle_\Omega$  denotes the  $L^2(\Omega)$  scalar product. Since the operator  $\mathcal{L}$  is elliptic, problem (2.2) admits for any  $f \in L^2(0, T; L^2(\Omega))$  and  $u_0 \in L^2(\Omega)$  a unique weak solution; compare, e.g., with Evans (2010, Chap. 7).

## 2.2 The discrete problem

For the discretisation of (2.2) we use adaptive finite elements in space and a dG( $s$ ) scheme with adaptive time-step size control.

*Adaptive grids and time steps.* For the adaptive space discretisation we restrict ourselves to simplicial grids and local refinement by bisection; compare with the studies by Bänsch (1991); Kossaczky (1994); Maubach (1995); Traxler (1997), as well as Schmidt & Siebert (2005) and Nochetto *et al.* (2008) and the references therein. To be more precise, refinement is based on the initial conforming triangulation  $\mathcal{G}_{\text{init}}$  of  $\Omega$  and a procedure `REFINE` with the following properties: given a conforming triangulation  $\mathcal{G}$  and a subset  $\mathcal{M} \subset \mathcal{G}$  of *marked elements*, then

$$\text{REFINE}(\mathcal{G}, \mathcal{M})$$

outputs a conforming refinement  $\mathcal{G}_+$  of  $\mathcal{G}$  such that all elements in  $\mathcal{M}$  are bisected at least once. In general, additional elements are refined in order to ensure conformity. The input  $\mathcal{G}$  can be either  $\mathcal{G}_{\text{init}}$  or the output of a previous application of `REFINE`. We denote by  $\mathbb{G}$  the class of all conforming triangulations that can be produced from  $\mathcal{G}_{\text{init}}$  by finitely many applications of `REFINE`. For  $\mathcal{G} \in \mathbb{G}$  we call  $\mathcal{G}_+ \in \mathbb{G}$  a *refinement* of  $\mathcal{G}$  if  $\mathcal{G}_+$  is produced from  $\mathcal{G}$  by a finite number of applications of `REFINE` and we denote this by  $\mathcal{G} \leqslant \mathcal{G}_+$  or  $\mathcal{G}_+ \geqslant \mathcal{G}$ . Conversely, we call any  $\mathcal{G}_- \in \mathbb{G}$  with  $\mathcal{G}_- \leqslant \mathcal{G}$  a *coarsening* of  $\mathcal{G}$ .

Throughout the discussion we deal only with conforming grids; this means that whenever we refer to some triangulations  $\mathcal{G}$ ,  $\mathcal{G}_+$  and  $\mathcal{G}_-$  we tacitly assume  $\mathcal{G}, \mathcal{G}_+, \mathcal{G}_- \in \mathbb{G}$ . One key property of the refinement by bisection is uniform shape regularity for any  $\mathcal{G} \in \mathbb{G}$ . This means that all constants depending on the shape regularity are uniformly bounded depending on  $\mathcal{G}_{\text{init}}$ .

For the discretisation in time we let  $0 = t_0 < t_1 < \dots < t_N = T$  be a partition of  $(0, T)$  into half-open subintervals  $I_n = (t_{n-1}, t_n]$  with corresponding local time-step sizes  $\tau_n := |I_n| = t_n - t_{n-1}$ ,  $n = 1, \dots, N$ .

*Space-time discretisation.* For the spatial discretisation we use Lagrange finite elements. For any  $\mathcal{G} \in \mathbb{G}$  the finite element space  $\mathbb{V}(\mathcal{G})$  consists of all continuous, piecewise polynomials of fixed degree  $\ell \geq 1$  over  $\mathcal{G}$  that vanish on  $\partial\Omega$ . This gives a conforming discretisation of  $\mathbb{V}$ , i.e.,  $\mathbb{V}(\mathcal{G}) \subset \mathbb{V}$ . Moreover, Lagrange finite elements give nested spaces, i.e.,  $\mathbb{V}(\mathcal{G}) \subset \mathbb{V}(\mathcal{G}_+)$  whenever  $\mathcal{G} \leqslant \mathcal{G}_+$ .

We denote by  $\mathcal{G}_0$  the triangulation at  $t_0 = 0$  and for  $n \geq 1$ , we denote by  $\mathcal{G}_n$  the grid in  $I_n$  and let  $\mathbb{V}_n = \mathbb{V}(\mathcal{G}_n)$ ,  $n = 0, \dots, N$  be the corresponding finite element spaces. For  $\mathcal{G} \in \mathbb{G}$  we denote by  $\prod_{\mathcal{G}}: L^2(\Omega) \rightarrow \mathbb{V}(\mathcal{G})$  the  $L^2$ -projection onto  $\mathbb{V}(\mathcal{G})$  and set  $\prod_n := \prod_{\mathcal{G}_n}$ .

On each time interval the discrete approximation is polynomial in time over the corresponding spatial finite element space. Let  $s \in \mathbb{N}_0$  for any real vector space  $\mathbb{U}$  and interval  $I \subset \mathbb{R}$ ; we denote by

$$\mathbb{P}_s(I; \mathbb{U}) := \left\{ t \mapsto \sum_{i=0}^s t^i V_i : V_i \in \mathbb{U} \right\}$$

the space of all polynomials with degree less than or equal to  $s$  over  $\mathbb{U}$ . We write  $\mathbb{P}_s(\mathbb{U}) := \mathbb{P}_s(\mathbb{R}, \mathbb{U})$  and  $\mathbb{P}_s := \mathbb{P}_s(\mathbb{R})$ .

Furthermore, for an interval  $I \subset (0, T)$  we let

$$f_I \in \mathbb{P}_s(I; L^2(\Omega))$$

be the best approximation of  $f|_I$  in  $L^2(I; L^2(\Omega))$ . In particular we use  $f_n := f_{I_n}$  as a time discretisation of  $f$  on  $I_n$ . For  $s = 0$ ,  $f_I = \frac{1}{|I|} \int_I f \, dt$  is the mean value of  $f$  on  $I$ .

In the following we introduce the so-called discontinuous Galerkin time-stepping scheme dG( $s$ ) of degree  $s$ , where dG(0) is the well-known implicit Euler scheme. To this end we denote for  $n \geq 1$  the actual grid on  $I_n$  by  $\mathcal{G}_n$  and let  $\mathbb{V}_n = \mathbb{V}(\mathcal{G}_n)$  be the corresponding finite element space. We start with a suitable initial refinement  $\mathcal{G}_0$  of  $\mathcal{G}_{\text{init}}$  and an approximation  $U_0 = \Pi_0 u_0 = \Pi_{\mathcal{G}_0} u_0 \in \mathbb{V}_0$  of the initial value  $u_0$ . Note that in principle any suitable interpolation operator can be used instead of  $\Pi_0$ . We then inductively compute for  $n > 0$  a solution  $U|_{I_n} \in \mathbb{P}_s(\mathbb{V}_n)$  to the problem

$$\int_{I_n} \langle \partial_t U|_{I_n}, V \rangle_{\Omega} + \mathcal{B}[U|_{I_n}, V] dt + \langle [U]_{n-1}, V(t_{n-1}) \rangle_{\Omega} = \int_{I_n} \langle f_n, V \rangle_{\Omega} \quad (2.3)$$

for all  $V \in \mathbb{P}_s(\mathbb{V}_n)$ . Thereby,  $f_n := f|_{I_n}$  and  $[U]_{n-1}$  denotes the jump

$$[U]_{n-1} := U_{n-1}^+ - U_{n-1}^-$$

of  $U$  across  $t_{n-1}$ , where we used  $U_{n-1}^+ := \lim_{t \downarrow t_{n-1}} U|_{I_n}(t)$ ,  $U_n^- := U|_{I_n}(t_n)$ ,  $n = 1, \dots, N$  and  $U_0^- := U_0$ . Note that with this definition we have  $U_{n-1}^- = U(t_{n-1})$ . The solution  $U$  is uniquely defined (Thomée, 2006) and we will see below that (2.3) is equivalent to an  $(s+1)$ -dimensional second-order elliptic system. Note that  $U$  is allowed to be discontinuous across the nodal points  $t_0, \dots, t_N$  and hence in general  $U \notin \mathbb{W}(0, T)$ .

In order to construct from  $U$  a conforming function we shall now define the Radau reconstruction, which was first introduced and analysed by Makridakis & Nochetto (2006) and is motivated by the close relation of the dG( $s$ ) to Runge–Kutta Radau IIA collocation methods. In fact, let  $c_1 < \dots < c_{s+1}$  and  $b_1, \dots, b_{s+1}$  be the abscissae and weights of the RadauIIA quadrature formula, which is exact of degree  $2s$ , i.e.,

$$\sum_{j=1}^{s+1} b_j P(c_j) = \int_0^1 P(t) dt \quad \text{for all } P \in \mathbb{P}[2s]. \quad (2.4)$$

We define  $\mathcal{U} \in \mathbb{W}$ ,  $\mathcal{U}|_{I_n} \in \mathbb{P}_{s+1}(I_n; \mathbb{V})$  as the piecewise interpolation of  $U$  at the local RadauIIA points  $t_n^j := t_{n-1} + c_j \tau_n$ , i.e.,

$$\mathcal{U}\left(t_n^j\right) = U|_{I_n}\left(t_n^j\right) \in \mathbb{V}_n, \quad j = 1, \dots, s+1. \quad (2.5a)$$

The continuous embedding of  $\mathbb{W}$  in  $C^0([0, T]; L^2(\Omega))$  additionally enforces

$$\mathcal{U}(t_{n-1}) = U_{n-1}^- \in \mathbb{V}_{n-1}, \quad (2.5b)$$

which uniquely defines  $\mathcal{U} \in \mathbb{W}$  since  $0 =: c_0 < c_1 < \dots < c_s$  are pairwise disjoint. For convenience, we shall adopt the representation of  $\mathcal{U}$  with Legendre polynomials from Davis & Rabinowitz (1984). We observe from Davis & Rabinowitz (1984) that the  $t_n^j, j = 1, \dots, s$  are the zeros of

$$\frac{(-1)^s}{2} (L_s^n - L_{s+1}^n) \in \mathbb{P}_{s+1},$$

where for  $r \in \mathbb{N}_0$ ,  $L_r^n$  denotes the affine transformation of the  $r$ th Legendre polynomial  $L_r$  on  $(-1, 1)$  to the interval  $I_n$ , i.e.,

$$L_r^n(t) := L_r\left(2\frac{t - t_{n-1}}{\tau_n} - 1\right).$$

Therefore, we have that the  $\{L_r^n\}_{r \in \mathbb{N}_0}$  are  $L^2$ -orthogonal on  $I_n$  and that  $L_r^n(t_n) = 1$ ,  $L_r^n(t_{n-1}) = (-1)^r$ , as well as  $\int_{I_n} |L_r^n|^2 dt = \frac{\tau_n}{2r+1}$ . These properties readily imply the representation

$$\mathcal{U}|_{I_n} = U|_{I_n} + \frac{(-1)^s}{2} (L_{s+1}^n - L_s^n) (U_{n-1}^+ - U_{n-1}^-); \quad (2.6)$$

compare with the studies by [Ern et al. \(2017\)](#) and [Smears \(2017\)](#). Using integration by parts with respect to time, (2.4) and (2.5), we observe that (2.3) is equivalent to

$$\int_{I_n} \langle \partial_t \mathcal{U}, V \rangle + \mathcal{B}[U, V] dt = \int_{I_n} \langle f_n, V \rangle_{\Omega} dt \quad (2.7)$$

for all  $n = 1, \dots, n$  and  $V \in \mathbb{P}_s(\mathbb{V}_n)$ .

We emphasise that  $\mathcal{U}(t)$  is a finite element function, since for  $t \in I_n$ , we have  $\mathcal{U}(t) \in \mathbb{V}(\mathcal{G}_{n-1} \oplus \mathcal{G}_n) \subset \mathbb{V}$ , where  $\mathcal{G}_{n-1} \oplus \mathcal{G}_n \in \mathbb{G}$  is the smallest common refinement of  $\mathcal{G}_{n-1}$  and  $\mathcal{G}_n$ , which we call *overlay*. Continuity of  $\mathcal{U}$  in time in combination with  $\mathcal{U}(t) \in \mathbb{V}$  for all  $t \in I$  then implies  $\mathcal{U} \in \mathbb{W}$ .

**REMARK 2.1** For  $s = 0$  we see from (2.3) that in each time step  $n \in \mathbb{N}$  we need to solve for partial differential operators of the form  $-\Delta + \mu$ , with  $\mu = \frac{1}{\tau_n}$ , in order to compute  $U_n$ . Unfortunately, for  $s > 0$ , though still coercive, (2.3) becomes an  $(s+1)$ -dimensional coupled nonsymmetric system. Recently, in the study by [Smears \(2017\)](#) a preconditioned conjugate gradient (PCG) method for a symmetrisation of (2.3) is proposed, which is fully robust with respect to the discretisation parameters  $s$  and  $\tau$ , provided a solver for the operator  $-\Delta + \mu$ ,  $\mu \geq 0$  is available.

### 3. A posteriori error estimation

One basic ingredient of adaptive methods is *a posteriori* error indicator building up a reliable upper bound for the error in terms of the discrete solution and given data. The dG(0) method corresponds to the implicit Euler scheme and residual-based estimators for the heat equation can be found in the study by [Verfürth \(2003\)](#). In this section we generalise this result and prove reliable and efficient residual-based estimators for dG( $s$ ) schemes (2.3), with arbitrary  $s \in \mathbb{N}_0$ .

Some arguments in this section are straightforward generalisations of those in the study by [Verfürth \(2003\)](#) and we only sketch their proofs; others are based on new ideas and therefore we shall prove them in detail.

#### 3.1 Equivalence of error and residual

In order to prove residual-based error estimators one first has to relate the error to the residual. To this end we note that (2.2) can be taken as an operator equation in  $L^2(0, T; \mathbb{V}^*) \times L^2(\Omega)$ . Its residual  $\text{Res}(\mathcal{U})$

in  $\mathcal{U} \in \mathbb{W}(0, T)$  is given by

$$\begin{aligned}\langle \text{Res}(\mathcal{U}), v \rangle &= \langle \partial_t(u - \mathcal{U}), v \rangle + \mathcal{B}[u - \mathcal{U}, v] \\ &= \langle f - \partial_t \mathcal{U}, v \rangle - \mathcal{B}[\mathcal{U}, v] \quad \text{for all } v \in \mathbb{V}.\end{aligned}\tag{3.1}$$

From the study by Tantardini & Veeser (2016) we have the following identity between the residual and the error.

**PROPOSITION 3.1** (Abstract error bound). Let  $u \in \mathbb{W}$  be the solution of (2.2) and let  $\mathcal{U} \in \mathbb{W}$  be the approximation defined in (2.5) for time instances  $t_0 = 0, \dots, t_N = T$  and time-step sizes  $\tau_n := t_n - t_{n-1}$ ,  $n = 1, \dots, N$ . Then it holds for  $0 \leq k \leq N$  that

$$\|u - \mathcal{U}\|_{\mathbb{W}(0, T)}^2 = \|u_0 - U_0\|_{\Omega}^2 + \|\text{Res}(\mathcal{U})\|_{L^2(0, T; \mathbb{V}^*)}^2\tag{3.2a}$$

and

$$\|\text{Res}(\mathcal{U})\|_{L^2(I_n, \mathbb{V}^*)}^2 \leq \left\{ 2\|\partial_t(u - \mathcal{U})\|_{L^2(I_n; \mathbb{V}^*)}^2 + \|u - \mathcal{U}\|_{L^2(I_n; \mathbb{V})}^2 \right\}.\tag{3.2b}$$

The rest of this section concentrates on proving computable upper and lower bounds for the error. We note that the initial error  $\|u_0 - U_0\|_{\Omega}$  in (3.2) is already *a posteriori* computable, whence it remains to estimate the dual norm of the residual. However, there is another issue of separating the influence of the temporal and the spatial discretisations on the error. In particular, defining the temporal residual  $\text{Res}_\tau(\mathcal{U}) \in L^2(0, T; \mathbb{V}^*)$  as

$$\langle \text{Res}_\tau(\mathcal{U}), v \rangle := \mathcal{B}[U - \mathcal{U}, v]\tag{3.3}$$

and the spatial residual  $\text{Res}_h(\mathcal{U}) \in L^2(0, T; \mathbb{V}^*)$  as

$$\langle \text{Res}_h(\mathcal{U}), v \rangle := \langle f_n - \partial_t \mathcal{U}, v \rangle - \mathcal{B}[U, v] \quad \text{on } I_n,\tag{3.4}$$

we obtain

$$\text{Res}(\mathcal{U}) = \text{Res}_\tau(\mathcal{U}) + \text{Res}_h(\mathcal{U}) + f - f_n \quad \text{on } I_n.\tag{3.6}$$

In what follows we use this decomposition to prove separated time and space error indicators, which build up a reliable and efficient bound for the error.

### 3.2 Temporal residual

Recalling (2.6), we have

$$\mathcal{U}(t) - U(t) = \frac{(-1)^s}{2} (L_{s+1}^n - L_s^n) (U_{n-1}^+ - U_{n-1}^-) \quad \text{for } t \in I_n, n = 1, \dots, N,$$

and thanks to (2.5) and (2.4), we obtain

$$\begin{aligned} \int_{I_n} \|\text{Res}_\tau(\mathcal{U})\|_{\mathbb{V}^*}^2 dt &= \int_{I_n} \|\|U - \mathcal{U}\|\|_\Omega^2 dt = \|\|U_{n-1}^+ - U_{n-1}^-\|\|_\Omega^2 \int_{I_n} \frac{1}{4} |L_{s+1}^n - L_s^n|^2 dt \\ &= \tau_n C_\tau \|\|U_{n-1}^+ - U_{n-1}^-\|\|_\Omega^2, \end{aligned} \quad (3.6)$$

where  $C_\tau = C_\tau(s) := \frac{1}{4} \left( \frac{1}{2s+3} + \frac{1}{2s+1} \right)$ ; see also the study by Ern *et al.* (2017).

### 3.3 The spatial residual

In this section we estimate the spatial residual.

**LEMMA 3.2** Let  $U$  be the approximation of (2.3) to the solution  $u$  of (2.2) and let  $\mathcal{U}$  be its interpolation defined by (2.5). Then there exists a constant  $C_{\mathbb{G}} > 0$  such that

$$\int_{I_n} \|\text{Res}_h(\mathcal{U})\|_{\mathbb{V}^*}^2 dt \leq C_{\mathbb{G}} \sum_{E \in \mathcal{G}_n} \int_{I_n} h_E^2 \|\partial_t \mathcal{U} + \mathcal{L}U - f_n\|_E^2 + h_E \|J(U)\|_{\partial E}^2 dt$$

for all  $1 \leq n \leq N$ . Thereby, for  $V \in \mathbb{V}_n$  we denote by  $J(V)|_S$  for an interior side  $S$  the jump of the normal flux  $\mathbf{A} \nabla V \cdot \vec{n}$  across  $S$  and for boundary sides  $S$  we set  $J(V)|_S \equiv 0$ . The mesh size of an element  $E \in \mathcal{G}$  is given by  $h_E := |E|^{1/d}$ .

*Proof.* Recalling (3.4), we first observe that  $\|\text{Res}_h(\mathcal{U})\|_{\mathbb{V}^*}^2 \in \mathbb{P}_{2s}$ , whence by (2.4) we have

$$\int_{I_n} \|\text{Res}_h(\mathcal{U})\|_{\mathbb{V}^*}^2 dt = \tau_n \sum_{j=1} b_j \left\| \text{Res}_h(\mathcal{U}) \left( t_n^j \right) \right\|_{\mathbb{V}^*}^2.$$

Therefore, it suffices to estimate  $\|\text{Res}_h(\mathcal{U})\|_{\mathbb{V}^*}^2$  at the abscissae of the RadauIIA quadrature formula. For arbitrary  $V_j \in \mathbb{V}_n, j = 1, \dots, s+1$  choose  $V \in \mathbb{P}_s(\mathbb{V}_n)$  in (2.7) such that  $V(t + c_i \tau_n) = V_j \delta_{ij}$ ,  $1 \leq i \leq s+1$ . Then again exploiting (2.4) yields the Galerkin orthogonality

$$\left\langle \text{Res}_h(\mathcal{U}) \left( t_n^j \right), V_j \right\rangle = 0, \quad j = 1, \dots, s+1.$$

Since  $V_j \in \mathbb{V}_n$  was arbitrary we have for any  $v \in \mathbb{V}$  that

$$\left\langle \text{Res}_h(\mathcal{U}) \left( t_n^j \right), v \right\rangle = \left\langle \text{Res}_h(\mathcal{U}) \left( t_n^j \right), v - V \right\rangle \quad \text{for all } V \in \mathbb{V}_n.$$

Using integration by parts with respect to the space variable, the Cauchy–Schwarz inequality, and the scaled trace inequality and choosing  $V$  as a suitable interpolation of  $v$ , we arrive at

$$\left\| \text{Res}_h(\mathcal{U}) \left( t_n^j \right) \right\|_{\mathbb{V}^*}^2 \leq C_{\mathbb{G}} \sum_{E \in \mathcal{G}_n} \left\{ h_E^2 \left\| (\partial_t \mathcal{U} + \mathcal{L}U - f_n) \left( t_n^j \right) \right\|_E^2 + h_E \|J(U) \left( t_n^j \right)\|_{\partial E}^2 \right\}.$$

The right-hand side is a pointwise evaluation of a polynomial of degree  $2s$  and thus the claimed upper bound follows from (2.4).  $\square$

The following result shows that the spatial indicators are locally efficient in time as well.

LEMMA 3.3 Under the conditions of Lemma 3.2 we have

$$\sum_{E \in \mathcal{G}_n} \int_{I_n} h_E^2 \|\partial_t \mathcal{U} + \mathcal{L}U - f_n\|_E^2 + h_E \|J(U)\|_{\partial E}^2 dt \leq C \left\{ \int_{I_n} \|\text{Res}_h(\mathcal{U})\|_{V^*}^2 + \text{osc}_{\mathcal{G}_n}^2(f_n, \mathcal{U}) dt \right\},$$

where

$$\text{osc}_{\mathcal{G}_n}^2(f_n, \mathcal{U}) := \sum_{E \in \mathcal{G}_n} h_E^2 \|\partial_t \mathcal{U} + \mathcal{L}U - f_n - R_E\|_E^2 + h_E \|J(U) - J_E\|_{\partial E}^2$$

with, at time  $t \in I_n$ , pointwise  $L^2(\Omega)$ -best approximations  $R_E(t) \in \mathbb{P}_s[2\ell-2](E)$ , respectively  $J_E(t)|_S \in \mathbb{P}_s[2\ell-1](S)$ , for each side  $S \subset \partial E$ . The constant  $C > 0$  depends solely on the shape regularity of  $\mathbb{G}$ .

*Proof.* With the same arguments as in the proof of Lemma 3.2, for each  $1 \leq j \leq s+1$  it suffices to prove that

$$\begin{aligned} & C_{\mathbb{G}} \sum_{E \in \mathcal{G}_n} h_E^2 \left\| (\partial_t \mathcal{U} + \mathcal{L}U - f_n) \left( t_n^j \right) \right\|_E^2 + h_E \left\| J(U) \left( t_n^j \right) \right\|_{\partial E}^2 \\ & \leq C \left\{ \left\| \text{Res}_h(\mathcal{U}) \left( t_n^j \right) \right\|_{V^*}^2 + \text{osc}_{\mathcal{G}_n}^2(f_n, U) \left( t_n^j \right) \right\}. \end{aligned}$$

This, however, follows with standard techniques used in *a posteriori* estimation of elliptic second-order problems; see, e.g., Verfürth (2013) and Mekchay & Nochetto (2005) and compare with the case of the implicit Euler scheme  $s = 0$  in the study by Verfürth (2003).  $\square$

### 3.4 Estimation of the error

By means of the decomposition of the residual (3.5) we can combine the above results to obtain a reliable and efficient error estimator for (1.1). To this end we introduce the following error indicators for the sake of brevity of presentation: for  $\mathcal{G} \in \mathbb{G}$  and  $v \in \mathbb{V}$  the estimator for the initial value is given by

$$\mathcal{E}_0^2(v, \mathcal{G}) := \|v - \mathcal{I}_{\mathcal{G}} v\|_{\Omega}^2. \quad (3.7a)$$

For  $f \in L^2(0, T; L^2(\Omega))$ ,  $t_* \in (0, T)$  and  $I = (t_*, t_* + \tau] \subset (t_*, T]$ , the so-called consistency error which is inherited by the decomposition of the residual (3.5) is defined by

$$\mathcal{E}_f^2(f, t_*, \tau) := 3 \frac{C_F^2}{a_*} \int_I \|f - f_I\|_{\Omega}^2 dt. \quad (3.7b)$$

For  $v^-, v^+ \in \mathbb{V}$ ,  $\mathcal{G} \in \mathbb{G}$ ,  $V \in \mathbb{P}_s(\mathbb{V}(G))$ ,  $E \in \mathcal{G}$  and  $g \in \mathbb{P}_s(L^2(\Omega))$  the indicator

$$\mathcal{E}_{ct}^2(v^+, v^-, \tau) := \tau 3 C_{\tau} \|v^- - v^+\|_{\Omega}^2 \quad (3.7c)$$

is motivated by (3.6) and Lemma 3.2 suggests defining the spatial indicators by

$$\begin{aligned} \mathcal{E}_{\mathcal{G}}^2(V, v^-, t_*, \tau, g, \mathcal{G}, E) &:= 3C_{\mathbb{G}} \int_I h_E^2 \|\partial_t \mathcal{V} + \mathcal{L}V - g\|_E^2 + h_E \|J(V)\|_{\partial E}^2 dt \\ &= 3C_{\mathbb{G}} \tau \sum_{j=1}^{s+1} b_j \left\{ h_E^2 \|(\partial_t \mathcal{V} + \mathcal{L}V - g)(t_* + c_j \tau)\|_E^2 + h_E \|J(V)(t_* + c_j \tau)\|_{\partial E}^2 \right\}. \end{aligned} \quad (3.7d)$$

Here we have used, analogously to (2.6), that

$$\mathcal{V}(t) := V + \frac{(-1)^s}{2} (L_{s+1}^I - L_s^I) (V(t) - v^-), \quad (3.8)$$

where for  $r \in \mathbb{N}_0$ ,  $L_r^I$  denotes the affine transformation of the Legendre polynomial  $L_r$  to  $I$ .

**PROPOSITION 3.4** (Upper bound). Let  $u \in W$  be the solution of (2.2) and let  $\mathcal{U} \in W$  be the approximation defined in (2.5) for time instances  $t_0 = 0, \dots, t_N = T$  and time-step sizes  $\tau_n := t_n - t_{n-1}$ ,  $n = 1, \dots, N$ . Then we have the estimate

$$\begin{aligned} \|u - \mathcal{U}\|_{W(0,T)}^2 &\leq \mathcal{E}_0^2(u_0, \mathcal{G}_0) + \sum_{n=1}^N \left\{ \mathcal{E}_{c\tau}^2(U_{n-1}^+, U_{n-1}^-, \tau_n) + \mathcal{E}_{\mathcal{G}}^2(U, U_{n-1}^-, t_n, \tau_n, f_n, \mathcal{G}_n) \right. \\ &\quad \left. + 3\|f - f_n\|_{L^2(I_n; \mathbb{V}^*)}^2 \right\}. \end{aligned}$$

In particular, this implies

$$\|u - \mathcal{U}\|_{W(0,T)}^2 \leq \mathcal{E}_0^2(u_0, \mathcal{G}_0) + \sum_{n=1}^N \left\{ \mathcal{E}_{c\tau}^2(U_{n-1}^+, U_{n-1}^-, \tau_n) + \mathcal{E}_{\mathcal{G}}^2(U, U_{n-1}^-, t_n, \tau_n, f_n, \mathcal{G}_n) + \mathcal{E}_f^2(f, t_{n-1}, \tau_n) \right\}.$$

*Proof.* By the decomposition of the residual (3.5) and the triangle inequality we estimate on each interval  $I_n$ ,  $n = 1, \dots, N$ ,

$$\|\text{Res}(\mathcal{U})\|_{L^2(I_n; \mathbb{V}^*)}^2 \leq 3\|\text{Res}_\tau(\mathcal{U})\|_{L^2(I_n; \mathbb{V}^*)}^2 + 3\|\text{Res}_h(\mathcal{U})\|_{L^2(I_n; \mathbb{V}^*)}^2 + 3\|f - f_n\|_{L^2(I_n; \mathbb{V}^*)}^2.$$

Now the first assertion follows with Proposition 3.1, (3.6) and Lemma 3.2. The second bound follows from the Friedrichs inequality and the ellipticity (2.1) of  $\mathcal{B}$ . Indeed, we have

$$\|f - f_n\|_{L^2(I_n; \mathbb{V}^*)}^2 \lesssim \frac{C_F^2}{a_*} \|f - f_n\|_{L^2(I_n; L^2(\Omega))}^2. \quad (3.9)$$

□

PROPOSITION 3.5 (Lower bound). Supposing the conditions of Proposition 3.4 we have

$$\begin{aligned} & \mathcal{E}_{ct}^2(U_{n-1}^+, U_{n-1}^-, \tau_n) + \mathcal{E}_{\mathcal{G}}^2(U, U_{n-1}^-, t_n, \tau_n, f_n, \mathcal{G}_n) \\ & \leq C \left\{ \|\partial_t(u - \mathcal{U})\|_{L^2(I_n; \mathbb{V}^*)}^2 + \|u - \mathcal{U}\|_{L^2(I_n; \mathbb{V})}^2 + \int_{I_n} \text{osc}_{\mathcal{G}_n}^2(f_n, \mathcal{U}) dt + \|f - f_n\|_{L^2(I_n; \mathbb{V}^*)}^2 \right\}, \end{aligned}$$

where the constant  $C$  depends solely on the shape regularity of  $\mathbb{G}$  and on  $\ell$ .

*Proof.* We first consider the spatial indicators. By Lemma 3.3 there exists  $C > 0$  such that

$$\mathcal{E}_{\mathcal{G}}^2(U, U_{n-1}^-, t_n, \tau_n, f_n, \mathcal{G}_n) \leq C \|\text{Res}_h(\mathcal{U})\|_{L^2(I_n; \mathbb{V}^*)}^2 + C \int_{I_n} \text{osc}_{\mathcal{G}_n}^2(f_n, \mathcal{U}) dt.$$

The first term on the right-hand side can be further estimated using the decomposition of the residual, the triangle inequality and (3.6) to obtain

$$\|\text{Res}_h(\mathcal{U})\|_{L^2(I_n; \mathbb{V}^*)} \leq \|\text{Res}(\mathcal{U})\|_{L^2(I_n; \mathbb{V}^*)} + \|f - f_n\|_{L^2(I_n; \mathbb{V}^*)} + \mathcal{E}_{ct}^2(U_{n-1}^-, U_{n-1}^+, \tau_n). \quad (3.10)$$

The time/coarsening indicator  $\mathcal{E}_{ct}$  can be bounded as in the study by Ern *et al.* (2017). In order to keep the paper self-contained we shall reproduce the proof here. Recalling the properties of the transformed Legendre functions we conclude with (2.6) that

$$\begin{aligned} \frac{1}{4} \frac{\tau_n}{2s+3} \|U_{n-1}^+ - U_{n-1}^-\|_{\Omega}^2 &= \frac{1}{4} \|U_{n-1}^+ - U_{n-1}^-\|_{\Omega}^2 \int_{I_n} (L_{s+1}^n)^2 dt \\ &= \|U_{n-1}^+ - U_{n-1}^-\|_{\Omega}^2 \int_{I_n} \frac{(-1)^s}{2} (L_{s+1}^n - L_s^n) \frac{(-1)^s}{2} L_{s+1}^n dt \\ &= \int_{I_n} \left\langle \nabla(\mathcal{U} - U), \nabla(U_{n-1}^+ - U_{n-1}^-) \frac{(-1)^s}{2} L_{s+1}^n \right\rangle_{\Omega} dt \\ &= \int_{I_n} \left\langle \text{Res}(\mathcal{U}), (U_{n-1}^- - U_{n-1}^+) \frac{(-1)^s}{2} L_{s+1}^n \right\rangle dt \\ &\quad + \int_{I_n} \left\langle f - f_n, (U_{n-1}^- - U_{n-1}^+) \frac{(-1)^s}{2} L_{s+1}^n \right\rangle_{\Omega} dt. \end{aligned} \quad (3.11)$$

Here we have used in the last step that  $L_{s+1}^n$  is  $L^2$ -orthogonal on  $I_n$  onto  $\mathbb{P}_s$  and thus we have  $\int_{I_n} \langle \partial_t \mathcal{U}, L_{s+1}^n \rangle dt = \int_{I_n} \langle f_n, L_{s+1}^n \rangle_{\Omega} dt = \int_{I_n} \langle \nabla U, L_{s+1}^n \rangle_{\Omega} dt = 0$ . The assertion follows then from a Cauchy–Schwarz inequality. Indeed, recalling (3.7c) and observing that

$$\frac{\frac{1}{4} \frac{1}{2s+3}}{C_{\tau}} = \frac{2s+1}{4s+4} \in \left[ \frac{1}{4}, \frac{1}{2} \right),$$

we have that the involved constants are robust with respect to  $s$ .  $\square$

**REMARK 3.6** (Consistency indicator  $\mathcal{E}_f$ ). The upper and lower bounds in Propositions 3.4 and 3.5 in principle involve the consistency terms  $\|f - f_n\|_{L^2(I_n; \mathbb{V}^*)}^2$ ,  $n = 1, \dots, N$ , which are not computable in general. Though these terms can be bounded by  $\mathcal{E}_f^2(f, t_{n-1}, \tau_n)$  (compare with (3.11)), this still involves the best approximation  $f_n \in \mathbb{P}_s(I_n; L^2(\Omega))$  of  $f|_{I_n}$  in  $L^2(I_n; L^2(\Omega))$ . In a practical algorithm  $f_n$  may therefore be replaced by some quasi-best approximation in  $\mathbb{P}_s(I_n; L^2(\Omega))$  like a Clément-type interpolation in time; see the study by Clément (1975).

**REMARK 3.7** (Temporal residual). We emphasise that the techniques from the study by Ern et al. (2017) for bounding the indicator  $\mathcal{E}_{ct}$  in the proof of Proposition 3.5 are different from the one in the study by Verfürth (2003) for the dG(0) scheme and it is noteworthy that this improves the constant even in this case. In fact, it follows from the proof of Proposition 3.5 that

$$\frac{1}{8} \mathcal{E}_f^2(U_{n-1}^+, U_{n-1}^-, \tau_n)^2 \leq \| \text{Res}(\mathcal{U}) \|_{L^2(I_n; \mathbb{V}^*)}^2 + \| f - f_n \|_{L^2(I_n; \mathbb{V}^*)}^2;$$

compare also with the study by Ern et al. (2017).

**REMARK 3.8** (Time-space efficiency). We note that the lower bound in Proposition 3.5 is local in time but, however, global in space. In fact, though the estimate (3.12) of the indicators  $\mathcal{E}_G$  in principle is local also in space, we have that the estimate of  $\mathcal{E}_{ct}$  in (3.11) follows from a global relation in space. Ern et al. (2017) presented an estimator, which is locally efficient in space and time for the modified error norm

$$\|u - U\|_{\mathcal{E}_Y} := \left( \|u - \mathcal{U}\|_{W(0,T)}^2 + \sum_{n=1}^N \mathcal{E}_{ct}^2(U_{n-1}^+, U_{n-1}^-, \tau_n)^2 \right)^{1/2}, \quad (3.12)$$

which can be interpreted as a dG( $s$ ) norm controlling the temporal jumps of the discrete solution. With respect to this norm the local space and time efficiency of the indicator  $\mathcal{E}_{ct}$  with respect to  $\|u - U\|_{\mathcal{E}_Y}$  is then immediate. Moreover, by resorting to equilibrated flux estimators instead of the residual-based ones in (3.7d), the *a posteriori* bounds in the study by Ern et al. (2017) are also robust with respect to the polynomial degree  $\ell$  in space.

**REMARK 3.9** (Elliptic problem). In case of the implicit Euler scheme dG(0) it is well known that in each  $1 \leq n \leq N$ ,  $U|_{I_n} \in \mathbb{P}_s[0](\mathbb{V}_n) = \mathbb{V}_n$  is the Ritz approximation to a coercive elliptic problem. Moreover, the spatial estimators (3.7d) are the standard residual-based estimators for this elliptic problem. This observation transfers to the dG( $s$ ) scheme for  $s \geq 1$ . To see this we observe that (after transformation to the unit interval) (2.3) is a Galerkin approximation to the solution  $u_\tau \in \mathbb{P}_s(\mathbb{V})$  of a problem of the kind

$$\int_0^1 \frac{1}{\tau} \langle \partial_t u_\tau, v \rangle_\Omega + \mathcal{B}[u_\tau, v] dt + \frac{1}{\tau} \langle u_\tau(0), v(0) \rangle_\Omega = \int_0^1 \langle \bar{f}, v \rangle_\Omega dt + \frac{1}{\tau} \langle v^-, v(0) \rangle_\Omega \quad (3.13)$$

for all  $v \in \mathbb{P}_s(\mathbb{V})$  and some data  $\bar{f} \in \mathbb{P}_s(L^2(\Omega))$ ,  $v^- \in L^2(\Omega)$  and  $\tau > 0$ . The mappings  $v \mapsto v(0)$  and  $v \mapsto \partial_t v$  are linear and continuous on  $\mathbb{P}_s(\mathbb{V})$ , whence this equation can be taken as a vector-valued linear

variational problem of second order on  $\mathbb{V}^{s+1}$ . Testing with  $v = u_\tau$  proves coercivity:

$$\int_0^1 \frac{1}{\tau} \langle \partial_t u_\tau, u_\tau \rangle_{\Omega} + \mathcal{B}[u_\tau, u_\tau] dt + \frac{1}{\tau} \langle u_\tau(0), u_\tau(0) \rangle_{\Omega} = \frac{1}{2\tau} \|u_\tau(0)\|_{\Omega}^2 + \frac{1}{2\tau} \|u_\tau(1)\|_{\Omega}^2 + \int_0^1 \|u_\tau\|_{\Omega}^2 dt.$$

Obviously, its residual in  $V \in \mathbb{P}_s(\mathbb{V})$  is given by

$$\langle \text{Res}_h(\mathcal{V}), v \rangle = \langle \bar{f} - \partial_t \mathcal{V}, v \rangle - \mathcal{B}[V, v], \quad v \in \mathbb{P}_s(\mathbb{V}),$$

where  $\mathcal{V} \in \mathbb{P}_s[s+1](\mathbb{V})$  is such that  $\mathcal{V}(c_j) = V(c_j)$ ,  $j = 1, \dots, s$  and  $\mathcal{V}(0) = v^-$ ; compare with (2.5). Thanks to Lemmas 3.2 and 3.3, for  $V \in \mathbb{P}_s(\mathbb{V}(\mathcal{G}))$ ,  $\mathcal{G} \in \mathbb{G}$ , the standard residual-based estimator for this problem is given by  $\mathcal{E}_{\mathcal{G}}^2(V, v^-, \tau, 0, \bar{f}, \mathcal{G})$ .

*Energy estimation.* We shall now generalise the energy estimate from the study by Kreuzer *et al.* (2012) to higher-order dG(s) schemes. Similar estimates have been obtained in the studies by Eriksson & Johnson (1991, 1995), although they use different techniques involving restrictions on the time step and mesh sizes, as well as constants resulting from inverse inequalities.

**PROPOSITION 3.10** (Uniform global energy estimate). Assume  $N \in \mathbb{N} \cup \{\infty\}$  and arbitrary time instances  $0 = t_0 < \dots < t_N \leq T$  with time-step sizes  $\tau_1, \dots, \tau_N > 0$ . Let  $U_0 = \Pi_0 u_0$  and for  $1 \leq n \leq N$  let  $U|_{I_n} \in \mathbb{P}_s(\mathbb{V}_n)$  be the discrete solutions to (2.3) and let  $\mathcal{U} \in W$  as defined in (2.5). Then for any  $m = 1, \dots, N$  we have

$$\sum_{n=1}^m \int_{I_n} \|\partial_t \mathcal{U}\|_{\Omega}^2 dt + \|U_{n-1}^+ - \Pi_n U_{n-1}^-\|_{\Omega}^2 + \|U_n^-\|_{\Omega}^2 - \|\Pi_n U_{n-1}^-\|_{\Omega}^2 \leq \sum_{n=1}^m \int_{I_n} \|f_n\|_{\Omega}^2 dt.$$

*Proof.* We choose  $V := \Pi_n \partial_t \mathcal{U}|_{I_n} \in \mathbb{P}_s(\mathbb{V}_n)$  as a test function in (2.7) obtaining

$$\int_{I_n} \|\Pi_n \partial_t \mathcal{U}\|_{\Omega}^2 + \mathcal{B}[U, \Pi_n \partial_t \mathcal{U}] dt = \int_{I_n} \langle f_n, \Pi_n \partial_t \mathcal{U} \rangle_{\Omega} dt. \quad (3.14)$$

In order to analyse the second term on the left-hand side we first observe that  $\Pi_n \partial_t \mathcal{U}|_{I_n} = \partial_t \Pi_n \mathcal{U}|_{I_n} \in \mathbb{P}_s(\mathbb{V}_n)$ . Recalling (2.5b) and that  $\mathcal{B} : \mathbb{V} \times \mathbb{V} \rightarrow \mathbb{R}$  is constant in time we obtain, integrating by parts, that

$$\int_{I_n} \mathcal{B}[U, \Pi_n \partial_t \mathcal{U}] dt = - \int_{I_n} \mathcal{B}[\partial_t U, \Pi_n \mathcal{U}] dt + \|U_n^-\|_{\Omega}^2 - \mathcal{B}[U_{n-1}^+, \Pi_n U_{n-1}^-].$$

Since  $\mathcal{B}[\partial_t U, \Pi_n \mathcal{U}]|_{I_n} \in \mathbb{P}_s[2s]$  we can apply (2.4) and conclude with (2.5a) that

$$\begin{aligned} \int_{I_n} \mathcal{B}[U, \Pi_n \partial_t \mathcal{U}] dt &= - \int_{I_n} \mathcal{B}[\partial_t U, U] dt + \|U_n^-\|_{\Omega}^2 - \mathcal{B}[U_{n-1}^+, \Pi_n U_{n-1}^-] \\ &= \frac{1}{2} \|U_{n-1}^+ - \Pi_n U_{n-1}^-\|_{\Omega}^2 - \frac{1}{2} \|\Pi_n U_{n-1}^-\|_{\Omega}^2 + \frac{1}{2} \|U_n^-\|_{\Omega}^2, \end{aligned}$$

where we used that  $\mathcal{B}[\partial_t U|_{I_n}, U|_{I_n}] = \frac{1}{2} \partial_t \|U|_{I_n}\|_{\Omega}^2$ . Inserting this in (3.14) yields

$$\int_{I_n} \|\Pi_n \partial_t \mathcal{U}\|_{\Omega}^2 dt + \frac{1}{2} \|U_{n-1}^+ - \Pi_n U_{n-1}^-\|_{\Omega}^2 - \frac{1}{2} \|\Pi_n U_{n-1}^-\|_{\Omega}^2 + \frac{1}{2} \|U_n^-\|_{\Omega}^2 = \int_{I_n} \langle f_n, \Pi_n \partial_t \mathcal{U} \rangle_{\Omega} dt.$$

Estimating the right-hand side with the help of the Cauchy–Schwarz and the Young inequalities proves the assertion.  $\square$

**COROLLARY 3.11** Under the conditions of Proposition 3.10 assume that

$$\|U_{n-1}^-\|_{\Omega}^2 - \|\Pi_n U_{n-1}^-\|_{\Omega}^2 + \frac{1}{2} \int_{I_n} \|\Pi_n \partial_t \mathcal{U}\|_{\Omega}^2 dt \geq 0 \quad \text{for } n = 1, \dots, N. \quad (3.15)$$

Then we have the estimate

$$\sum_{n=1}^m \frac{1}{2} \int_{I_n} \|\Pi_n \partial_t \mathcal{U}\|_{\Omega}^2 dt + \|U_{n-1}^+ - \Pi_n U_{n-1}^-\|_{\Omega}^2 \leq \|f\|_{\Omega \times (0, t_m)}^2 + \|U_0\|_{\Omega}^2 - \|U_m^-\|_{\Omega}^2.$$

In particular, the series  $\sum_{n=1}^N \|U_{n-1}^+ - \Pi_n U_{n-1}^-\|_{\Omega}^2$  is uniformly bounded irrespective of the sequence of time-step sizes used.

*Proof.* Summing the non-negative terms in (3.15) yields

$$0 \leq \sum_{n=1}^m \|U_{n-1}^-\|_{\Omega}^2 - \|\Pi_n U_{n-1}^-\|_{\Omega}^2 + \frac{1}{2} \int_{I_n} \|\Pi_n \partial_t \mathcal{U}\|_{\Omega}^2 dt,$$

which is equivalent to

$$\|U_m^-\|_{\Omega}^2 - \|U_0\|_{\Omega}^2 \leq \sum_{n=1}^m \|U_{n-1}^+\|_{\Omega}^2 - \|\Pi_n U_{n-1}^-\|_{\Omega}^2 + \frac{1}{2} \int_{I_n} \|\Pi_n \partial_t \mathcal{U}\|_{\Omega}^2 dt.$$

Using this in the estimate of Proposition 3.10 yields the desired estimate.  $\square$

Having a closer look at the indicator  $\mathcal{E}_{ct}$  we note that since we allow for coarsening, it is not a pure temporal error indicator. Coarsening may cause the loss of information and too little information may lead to wrong decisions within the adaptive method. For this reason we use the triangle inequality to split

$$\mathcal{E}_{ct}^2(v^-, v^+, \tau, \mathcal{G}) \leq \mathcal{E}_c^2(v^-, \tau, \mathcal{G}) + \mathcal{E}_{\tau}^2(v^+, v^-, \tau, \mathcal{G}) \quad (3.16a)$$

into a measure

$$\mathcal{E}_c^2(v^-, \tau, \mathcal{G}) := \sum_{E \in \mathcal{G}} \mathcal{E}_c^2(v^-, \tau, \mathcal{G}, E) := 6C_{\tau} \sum_{E \in \mathcal{G}} \tau \|\Pi_{\mathcal{G}} v^- - v^-\|_E^2 \quad (3.16b)$$

for the coarsening error and

$$\mathcal{E}_\tau^2(v^+, v^-, \tau, \mathcal{G}) := 6C_\tau \tau \|\|v^+ - \Pi_{\mathcal{G}} v^-\|\|_\Omega^2, \quad (3.16c)$$

which serves as an indicator for the temporal error. This allows us to control the coarsening error separately.

Assuming that (3.15) holds, Corollary 3.11 provides control of the sum of the time error indicators  $\mathcal{E}_\tau^2(U_{n-1}^+, U_{n-1}^-, \tau_n, \mathcal{G}_n) = 6C_\tau \tau \|\|U_{n-1}^+ - \Pi_n U_{n-1}^-\|\|_\Omega^2$ . Assumption (3.15) would trivially be satisfied for the Ritz projection  $R_n U_{n-1}^-$  of  $U_{n-1}^-$  into  $\mathbb{V}_n$ , since  $\|\|R_n U_{n-1}^-\|\|_\Omega \leq \|\|U_{n-1}^-\|\|_\Omega$ . The  $L^2$ -projection  $\Pi_n U_{n-1}^-$ , however, does not satisfy this monotonicity property in general and therefore coarsening may lead to an increase of energy. The algorithm presented below ensures that (3.15) is fulfilled at the end of every time step. To this end, using the notation (3.8) we define for  $V \in \mathbb{P}_s(\mathbb{V}(\mathcal{G}))$ ,  $v^- \in \mathbb{V}$ ,  $t_\star \in (0, T)$ ,  $I = (t_\star, t_\star + \tau] \subset (t_\star, T]$ ,  $\mathcal{G} \in \mathbb{G}$  and  $E \in \mathcal{G}$  the indicators

$$\mathcal{E}_*^2(V, v^-, t_\star, \tau, \mathcal{G}, E) := \|\| \Pi_{\mathcal{G}} v^-\|\|_E^2 - \|\|v^-\|\|_E^2 - \frac{1}{2} \int_I \|\| \Pi_{\mathcal{G}} \partial_t V \|\|_E^2 dt,$$

as well as the convenient notation  $\mathcal{E}_*^2(V, v^-, t_\star, \tau, \mathcal{G}) := \sum_{E \in \mathcal{G}} \mathcal{E}_*^2(V, v^-, t_\star, \tau, \mathcal{G}, E)$ . Condition (3.15) is then equivalent to  $\mathcal{E}_*^2(U, U_{n-1}^-, t_{n-1}, \tau_n, \mathcal{G}_n) \leq 0$ ,  $n = 1, \dots, N$ . Note that the term  $-\int_{I_n} \|\| \Pi_n \partial_t U \|\|_E^2$  may compensate for  $\|\| \Pi_n U_{n-1}^-\|\|_E^2 > \|\|U_{n-1}^-\|\|_E^2$ .

## 4. The adaptive algorithm TAFEM

Based on the observations in the previous section and a new concept for marking we shall next describe the adaptive algorithm TAFEM in this section. In contrast to the algorithms presented in the studies by Kreuzer *et al.* (2012) and Chen & Feng (2004), TAFEM is based on a different marking philosophy. In fact, they mark according to the same indicators, (3.7b)–(3.7d) and (3.16a), but in contrast to the studies by Kreuzer *et al.* (2012) and Chen & Feng (2004), TAFEM uses an  $L^2$ -instead of an  $L^\infty$ -strategy. Philosophically, this aims at an  $L^2$  rather than an  $L^\infty$  equal distribution of the error in time; compare also with the introductory Section 1.

We follow a bottom-up approach, i.e., we first state basic properties on some rudimentary modules that are treated as black-box routines, then describe three core modules in detail and finally combine these procedures in the adaptive algorithm TAFEM .

### 4.1 Black-box modules

As in the study by Kreuzer *et al.* (2012) we use standard modules `ADAPT_INIT`, `COARSEN`, `MARK_REFINE` and `SOLVE` as black-box routines, while adding the module `ENLARGE` as a black box as well. In particular, we use the subroutine `MARK_REFINE` in an object-oriented fashion, i.e., the functionality of `MARK_REFINE` changes according to its arguments. We next state the basic properties of these routines.

**ASSUMPTION 4.1** (Properties of modules). We suppose that all rudimentary modules terminate with an output having the following properties:

- (1) For a given initial datum  $u_0 \in L^2(\Omega)$  and tolerance  $\text{TOL}_0 > 0$  the output

$$(U_0, \mathcal{G}_0) = \text{ADAPT\_INIT}(u_0, \mathcal{G}_{\text{init}}, \text{TOL}_0)$$

is a refinement  $\mathcal{G}_0 \geq \mathcal{G}_{\text{init}}$  and an approximation  $U_0 \in \mathbb{V}(\mathcal{G}_0)$  to  $u_0$  such that  $\mathcal{E}_0^2(u_0, \mathcal{G}_0) \leq \text{TOL}_0^2$ .

- (2) For given  $g \in L^2(\Omega), \bar{f} \in \mathbb{P}_s(L^2(\Omega)), t_* \in (0, T), I = (t_*, t_* + \tau] \subset (t_*, T]$  and  $\mathcal{G} \in \mathbb{G}$ , the output

$$U_I = \text{SOLVE}(g, \bar{f}, t, \tau, \mathcal{G})$$

is the solution  $U_I \in \mathbb{P}_s(I, \mathbb{V}(\mathcal{G}))$  to the discrete elliptic problem

$$\int_I \langle \partial_t U_I, V \rangle_{\Omega} + \mathcal{B}[U_I, V] dt + \langle U_I(t), V(t) \rangle_{\Omega} = \langle g, V \rangle_{\Omega} + \int_I \langle \bar{f}, V \rangle_{\Omega} dt$$

for all  $V \in \mathbb{P}_s(\mathbb{V}(\mathcal{G}))$ ; compare with (2.3). Hereby, we assume exact integration and linear algebra.

- (3) For a given grid  $\mathcal{G} \in \mathbb{G}$  and a discrete function  $V \in \mathbb{V}(\mathcal{G})$ , the output

$$\mathcal{G}_* = \text{COARSEN}(V, \mathcal{G})$$

satisfies  $\mathcal{G}_* \leq \mathcal{G}$ .

- (4) For a given function  $f \in L^2(\Omega \times (0, T))$ , an initial time  $t \in (0, T)$  and an initial step size  $\tau \in (0, T - t]$ , the output

$$\tau_* = \text{ENLARGE}(f, t, \tau, \text{tol}_f)$$

satisfies  $\tau_* \geq \tau$  or  $\tau = T - t$ . The argument  $\text{tol}_f$  can be used to additionally restrict the enlargement; compare with Remark 4.3 below.

- (5) For a given grid  $\mathcal{G}$  and a set of indicators  $\{\mathcal{E}_E\}_{E \in \mathcal{G}}$ , the output

$$\mathcal{G}_* = \text{MARK\_REFINE}(\{\mathcal{E}_E\}_{E \in \mathcal{G}}, \mathcal{G}) \in \mathbb{G}$$

is a conforming refinement of  $\mathcal{G}$ , where at least one element in the subset  $\text{argmax}\{\mathcal{E}_E : E \in \mathcal{G}\} \subset \mathcal{G}$  has been refined.

- (6) For given grids  $\mathcal{G}, \mathcal{G}_{\text{old}} \in \mathbb{G}$  and a set of indicators  $\{\mathcal{E}_E\}_{E \in \mathcal{G}}$ , the output

$$\mathcal{G}_* = \text{MARK\_REFINE}(\{\mathcal{E}_E\}_{E \in \mathcal{G}}, \mathcal{G}, \mathcal{G}_{\text{old}}) \in \mathbb{G}$$

is a conforming refinement of  $\mathcal{G}$ , where at least one element of the set  $\{E \in \mathcal{G} : h_{\mathcal{G}|E} > h_{\mathcal{G}_{\text{old}}|E}\}$  of coarsened elements (with respect to  $\mathcal{G}_{\text{old}}$ ) is refined.

For a more detailed description of these modules see the study by Kreuzer *et al.* (2012, Section 3.3.1).

#### 4.2 The core modules

The first core module **CONSISTENCY** controls the consistency error  $\mathcal{E}_f$ . Recalling its definition in (3.7b) we see that the consistency error is solely influenced by the time-step size and can be computed without solving expensive discrete systems. Therefore, **CONSISTENCY** is used in the initialisation of each time step to adjust the time-step size such that the local consistency indicator  $\mathcal{E}_f^2(f, t, \tau)$  is below a local tolerance  $\text{tol}_f$ . It is important to notice that this module can be chosen to follow the classic *thresholding* algorithm, which ensures quasi-optimal order of convergence in terms of the degrees of freedom (DoFs); see Remark 4.3.

---

**Algorithm 1** Module **CONSISTENCY** (Parameter  $\kappa_1 \in (0, 1)$ )

---

```

CONSISTENCY( $f, t, \tau, \text{tol}_f$ )
1:  $\tau = \text{ENLARGE}(f, t, \tau)$ 
2: compute  $\mathcal{E}_f^2(f, t, \tau)$ 
3: while  $\mathcal{E}_f^2(f, t, \tau) > \text{tol}_f^2$  do ★ reduce  $\tau$ 
4:    $\tau = \kappa_1 \tau$ 
5:   compute  $\mathcal{E}_f^2(f, t, \tau)$ 
6: end while
7:  $\bar{f} = f_{[t, t+\tau]}$ 
8: return  $\bar{f}, \tau$ 

```

---

We start with termination of the module **CONSISTENCY**.

**LEMMA 4.2** (Termination of **CONSISTENCY**). Assume  $f \in L^2((0, T); L^2(\Omega))$ . Then for any  $t \in (0, T)$  and  $\tau^{\text{in}} \in (0, T - t]$ ,

$$(\bar{f}, \tau) = \text{CONSISTENCY}\left(f, t, \tau^{\text{in}}, \text{tol}_f\right)$$

terminates and

$$\mathcal{E}_f^2(f, t, \tau) \leq \text{tol}_f^2. \quad (4.1)$$

*Proof.* The proof is straightforward since  $\mathcal{E}_f^2(f, t, \tau)$  is monotone nonincreasing and  $\mathcal{E}_f^2(f, t, \tau) \rightarrow 0$  when  $\tau \rightarrow 0$ .  $\square$

**REMARK 4.3** We note that the enlargement step in line 2 of **CONSISTENCY** is not relevant for Lemma 4.2, i.e., it could for instance be implemented by Algorithm 2.

In this case the module **CONSISTENCY** becomes a classical *thresholding* algorithm; compare, e.g., with the study by [Binev et al. \(2002\)](#). However, a too aggressive enlargement of the time-step size may possibly need to be reduced later because of the time estimator  $\mathcal{E}_\tau$  (e.g., if  $\mathcal{E}_f \equiv 0$ ), which requires the solution of expensive linear systems. Therefore, we decided to introduce the abstract enlargement routine **ENLARGE**, which can be adjusted individually to the characteristics of the considered problem; compare also with Section 6.2.3, where we choose  $\kappa_2 > 1$ , i.e.,  $\min\{\kappa_2 \tau, T - t\} = \text{ENLARGE}(f, t, \tau, \text{tol}_f)$  for the rough initial data experiment.

**Algorithm 2** ENLARGE (Parameter  $\kappa_2$ )

---

```

ENLARGE( $f, t, \tau, \text{tol}_f$ )
1: compute  $\mathcal{E}_f^2(f, t, \tau)$ 
2: while  $\mathcal{E}_f^2(f, t, \tau) < \text{tol}_f^2$  and  $\tau < T - t$  do ★ enlarge  $\tau$ 
3:    $\tau = \min\{\kappa_2\tau, T - t\}$ 
4:   compute  $\mathcal{E}_f^2(f, t, \tau)$ 
5: end while
6: return  $\tau$ 

```

---

Obviously, a local control of the form (4.1) does not guarantee that the global consistency error is below some prescribed tolerance  $\text{TOL}_f$ . For this reason we first precompute some local tolerance  $\text{tol}_f$  from the global tolerance  $\text{TOL}_f$  by the following module TOLFIND.

**Algorithm 3** TOLFIND (Parameter  $\tilde{\tau}_0$ )

---

```

TOLFIND( $f, T, \text{TOL}_f$ )
1: initialise  $N_f$  and set  $\text{tol}_f = \text{TOL}_f, \tilde{\tau}_0 = 0$ 
2: loop forever
3:    $\epsilon = n = 0$ 
4:   do
5:      $n = n + 1$ 
6:      $(f_n, \tilde{\tau}_n) = \text{CONSISTENCY}(f, \tilde{\tau}_{n-1}, \tilde{\tau}_{n-1}, \text{tol}_f)$ 
7:      $\epsilon = \epsilon + \mathcal{E}_f^2(f, \tilde{\tau}_{n-1}, \tilde{\tau}_n)$ 
8:     while  $\tilde{\tau}_n = \tilde{\tau}_{n-1} + \tilde{\tau}_n < T$ 
9:      $N_f = n$ 
10:    if  $\epsilon > \frac{1}{2}\text{TOL}_f^2$  then ★ std. exit
11:       $\text{tol}_f^2 = \frac{1}{2}\text{tol}_f^2$ 
12:    else
13:      break
14:    end if
15:  end loop forever
16:   $\text{tol}_f^2 = \min\left\{\text{tol}_f^2, \frac{\text{TOL}_f^2}{2N_f}\right\}$ 
17: return  $\text{tol}_f$ 

```

---

The next result states that if all local consistency indicators are below the threshold  $\text{tol}_f$  then the accumulation of the consistency indicators stays indeed below the prescribed global tolerance  $\text{TOL}_f$ .

LEMMA 4.4 (Termination of TOLFIND). Assume  $f \in L^2((0, T); L^2(\Omega))$ . Then for any  $\text{TOL}_f > 0$  we have that

$$\text{tol}_f = \text{TOLFIND}(f, T, \text{TOL}_f) > 0$$

terminates. Moreover, let  $0 = t_0 < t_1 < \dots < t_N = T$  be arbitrary with  $\tau_n = t_n - t_{n-1}$ ,  $n = 1, \dots, N$ , then

$$\mathcal{E}_f^2(f, t_{n-1}, \tau_n) \leq \text{tol}_f^2, \quad n = 1, \dots, N \quad \Rightarrow \quad \sum_{n=1}^N \mathcal{E}_f^2(f, t_{n-1}, \tau_n) \leq \text{TOL}_f^2. \quad (4.2)$$

*Proof.* The proof is divided into three steps.

- (1) We show that the process from lines 4 to 8 terminates. To this end we recall the parameter  $\kappa_1 \in (0, 1)$  from  $\text{CONSISTENCY}(f, \tilde{t}_{n-1}, \tilde{\tau}_{n-1}, \text{tol}_f)$ . We argue by contradiction and assume that an infinite monotone sequence  $\{\tilde{t}_n\}_{n \geq 0} \subset [0, T]$  is constructed by TOLFIND with  $\lim_{n \rightarrow \infty} \tilde{t}_n = t^* \in (0, T]$ . Defining the intervals  $I_n^* = (\tilde{t}_n, \tilde{t}_n^*]$  where  $\tilde{t}_n^* = \min \left\{ \kappa_1^{-1}(t^* - \tilde{t}_n) + \tilde{t}_n, T \right\}$  we have that there exists  $n_0 \in \mathbb{N}$  such that

$$\|f\|_{\Omega \times (\tilde{t}_n, \tilde{t}_n^*)}^2 \leq \text{tol}_f^2$$

for all  $n \geq n_0$  since  $\tilde{t}_n \rightarrow t^*$ . Let  $m > n_0$  such that the condition in line 3 of the module  $\text{CONSISTENCY}(f, \tilde{t}_{m-1}, \tilde{\tau}_{m-1}, \text{tol}_f)$  is activated at least once (the existence of such  $m$  is a direct consequence of  $\tilde{t}_n \rightarrow t^*$ ), then

$$\begin{aligned} \text{tol}_f^2 &< \mathcal{E}_f^2(f, \tilde{t}_{m-1}, \kappa_1^{-1}\tilde{\tau}_m) \\ &\leq \|f\|_{\Omega \times (\tilde{t}_{m-1}, \tilde{t}_{m-1} + \kappa_1^{-1}\tilde{\tau}_m)}^2 \\ &\leq \|f\|_{\Omega \times (\tilde{t}_{m-1}, \tilde{t}_{m-1}^*)}^2 \leq \text{tol}_f^2. \end{aligned}$$

This contradiction implies that the sequence  $\{\tilde{t}_n\}_{n \geq 0}$  is finite.

- (2) We next check that the condition of line 10 is violated after finitely many steps. Since the span of characteristics of dyadic intervals is dense in  $L^2(0, T)$  we can choose  $M > 0$  such that the squared consistency error on the grid of  $2^M$  uniform intervals is below  $\frac{1}{4}\text{TOL}_f^2$ . We split the intervals generated in  $\text{TOLFIND}(f, T, \text{tol}_f)$  into

$$\mathbb{I}_{\text{in}} := \left\{ n : (\tilde{t}_{n-1}, \tilde{t}_n] \subset T(m2^{-M}, (m+1)2^{-M}] \text{ for some } m \in \{0, \dots, 2^M - 1\} \right\}$$

and  $\mathbb{I}_{\text{out}} := \{1, \dots, N_f\} \setminus \mathbb{I}_{\text{in}}$  according to whether or not they are included in one of the dyadic intervals. Therefore, we have, with the monotonicity of the consistency error and  $\#\mathbb{I}_{\text{out}} \leq 2^M$ ,

that

$$\epsilon = \sum_{n \in \mathbb{I}_{\text{in}}} \mathcal{E}_f^2(f, \tilde{t}_{n-1}, \tilde{\tau}_n) + \sum_{n \in \mathbb{I}_{\text{out}}} \mathcal{E}_f^2(f, \tilde{t}_{n-1}, \tilde{\tau}_n) \leq \frac{1}{4} \text{TOL}_f^2 + 2^M \text{tol}_f^2.$$

Taking  $\text{tol}_f^2 < 2^{-(M+2)} \text{TOL}_f^2$  we see that the condition of line 10 is violated, which proves the assertion.

- (3) Combining the above steps we conclude that TOLFIND terminates and it remains to prove (4.2). To this end we proceed similarly as in (2) and let

$$\mathbb{I}_{\text{in}} := \{n : (t_{n-1}, t_n] \subset (\tilde{t}_{m-1}, \tilde{t}_m] \text{ for some } m \in \{1, \dots, N_f\}\}$$

and  $\mathbb{I}_{\text{out}} := \{1, \dots, N\} \setminus \mathbb{I}_{\text{in}}$ . By monotonicity we have  $\sum_{n \in \mathbb{I}_{\text{in}}} \mathcal{E}_f^2(f, t_{n-1}, \tau_n) \leq \sum_{n=1}^{N_f} \mathcal{E}_f^2(f, \tilde{t}_{n-1}, \tilde{\tau}_n) \leq \text{TOL}_f^2/2$  and thus the assertion follows from

$$\begin{aligned} \sum_{n=1}^N \mathcal{E}_f^2(f, t_{n-1}, \tau_n) &= \sum_{n \in \mathbb{I}_{\text{in}}} \mathcal{E}_f^2(f, t_{n-1}, \tau_n) + \sum_{n \in \mathbb{I}_{\text{out}}} \mathcal{E}_f^2(f, t_{n-1}, \tau_n) \\ &\leq \frac{\text{TOL}_f^2}{2} + N_f \text{tol}_f^2 = \frac{\text{TOL}_f^2}{2} + N_f \frac{\text{TOL}_f^2}{2N_f} \leq \text{TOL}_f^2. \end{aligned}$$

□

**REMARK 4.5** (Estimation of  $\text{tol}_f$  under regularity assumptions). Supposing the regularity assumption  $f \in H^s((0, T); L^2(\Omega))$ ,  $s \in (0, 1]$  the following idea may be used as an alternative for the estimation of  $\text{tol}_f$  with TOLFIND.

Let  $\delta > 0$ . Then using Lemma 4.2 together with Poincaré's inequality in  $H^s$  and the fact that there are at most  $\frac{T}{\delta}$  disjoint intervals of length  $\delta$  in  $(0, T]$ , we obtain

$$\begin{aligned} \sum_{n=1}^N \mathcal{E}_f^2(f, t_{n-1}, \tau_n) &= \sum_{\tau_n > \delta} \mathcal{E}_f^2(f, t_{n-1}, \tau_n) + \sum_{\tau_n \leq \delta} \mathcal{E}_f^2(f, t_{n-1}, \tau_n) \\ &\leq \frac{T}{\delta} \text{tol}_f^2 + \sum_{\tau_n \leq \delta} \tau_n^{2s} \|f\|_{H^s(t_{n-1}, t_n; L^2(\Omega))}^2 \\ &= \frac{T}{\delta} \text{tol}_f^2 + \delta^{2s} \|f\|_{H^s(0, T; L^2(\Omega))}^2. \end{aligned}$$

By choosing  $\delta = \left( \frac{T \text{tol}_f}{\|f\|_{H^s(0, T; L^2(\Omega))}} \right)^{\frac{2}{2s+1}}$  the previous estimate turns into

$$\sum_{n=1}^N \mathcal{E}_f^2(f, t_{n-1}, \tau_n) \leq 2T^{\frac{2s}{2s+1}} \|f\|_{H^s(0, T; L^2(\Omega))}^{\frac{2}{2s+1}} \text{tol}_f^{\frac{4s}{2s+1}}.$$

In other words if *a priori* knowledge of the regularity of the right-hand side is available then TOLFIND can be replaced by the somewhat simpler term

$$\text{tol}_f^2 = 2^{-\frac{2s+1}{2s}} T^{-1} \|f\|_{H^s(0,T;L^2(\Omega))}^{-\frac{1}{s}} \text{TOL}_f^{\frac{2s+1}{s}}.$$

---

**Algorithm 4** Module ST\\_ADAPTATION (Parameter  $\kappa \in (0, 1)$ )

---

```

ST_ADIAPTION( $U_t^-$ ,  $f$ ,  $t$ ,  $\tau$ ,  $\mathcal{G}$ ,  $\mathcal{G}_{\text{old}}$ ,  $\text{tol}_{\mathcal{G}\tau}$ )
1: compute  $\mathcal{E}_f^2(f, t, \tau)$ 
2: loop forever
3:    $I = [t, t + \tau]$ 
4:    $\bar{f} = f_I$ 
5:    $U_I = \text{SOLVE}(U_t^-, \bar{f}, t, \tau, \mathcal{G})$ 
6:    $U_t^+ = \lim_{s \searrow t} U_I(s)$ 
7:   compute  $\left\{ \mathcal{E}_{\mathcal{G}}^2(U_I, U_t^-, t, \tau, \bar{f}, \mathcal{G}, E) \right\}_{E \in \mathcal{G}}$ ,  $\left\{ \mathcal{E}_*^2(U_t^+, U_t^-, \tau, \mathcal{G}, E) \right\}_{E \in \mathcal{G}}$ ,  $\mathcal{E}_{\tau}^2(U_t^+, U_t^-, \tau, \mathcal{G})$ 
    and  $\left\{ \mathcal{E}_c^2(U_t^-, \tau, \mathcal{G}, E) \right\}_{E \in \mathcal{G}}$ 
8:   if  $\mathcal{E}_c^2(U_t^-, \tau, \mathcal{G}) > \mathcal{E}_{\tau}^2(U_t^+, U_t^-, \tau, \mathcal{G}) + \mathcal{E}_f^2(f, t, \tau) + \tau \text{tol}_{\mathcal{G}\tau}$  then
9:      $\mathcal{G} = \text{MARK\_REFINE}\left(\left\{ \mathcal{E}_c^2(U_t^-, \tau, \mathcal{G}, E) \right\}_{E \in \mathcal{G}}, \mathcal{G}\right)$  A
10:    else if  $\mathcal{E}_*^2(U_t^+, U_t^-, \tau, \mathcal{G}) > 0$  then
11:       $\mathcal{G} = \text{MARK\_REFINE}\left(\left\{ \mathcal{E}_*^2(U_t^+, U_t^-, \tau, \mathcal{G}, E) \right\}_{E \in \mathcal{G}}, \mathcal{G}, \mathcal{G}_{\text{old}}\right)$  B
12:    else if  $\mathcal{E}_{\tau}^2(U_t^+, U_t^-, \tau, \mathcal{G}) > \text{tol}_{\mathcal{G}\tau}^2$  then
13:       $\tau = \kappa \tau$  C
14:      compute  $\mathcal{E}_f^2(f, t, \tau)$ 
15:    else if  $\mathcal{E}_{\tau}^2(U_I, U_t^-, t, \tau, \bar{f}, \mathcal{G}) > \mathcal{E}_{\tau}^2(U_t^+, U_t^-, \tau, \mathcal{G}) + \mathcal{E}_f^2(f, t, \tau) + \tau \text{tol}_{\mathcal{G}\tau}$  then
16:       $\mathcal{G} = \text{MARK\_REFINE}\left(\left\{ \mathcal{E}_{\tau}^2(U_I, U_t^-, t, \tau, \bar{f}, \mathcal{G}, E) \right\}_{E \in \mathcal{G}}, \mathcal{G}\right)$  D
17:    else
18:      break ★ exit
19:    end if
20:  end loop forever
21:  return  $U_I, \tau, \bar{f}, \mathcal{G}$ 

```

---

We turn to the module ST\\_ADAPTATION, listed in Algorithm 4, which handles a single time step. The module adapts the grid and the time-step size according to the indicators involving the discrete solution of the current time step, namely the space indicator  $\mathcal{E}_{\mathcal{G}}$  and the separated coarsening and time indicators  $\mathcal{E}_c$  and  $\mathcal{E}_{\tau}$ . The routine requires, right at the start of each iteration, the computation of the discrete solution on the actual grid and with the current time-step size; see line 5. Note that in ST\\_ADAPTATION only refinements are performed (both in space and in time). Recalling the discussion in the introductory Section 1, we aim to use a thresholding algorithm for the indicators  $\mathcal{E}_{\tau}$  in order to

equally distribute the time error. To this end we need to guarantee  $\mathcal{E}_* \leq 0$  in order to control the global time error with the help of the uniform energy estimate from Corollary 3.11. Since there is no similar control available for either the space or the coarsening errors, we relate the corresponding indicators to the time or the consistency indicator, i.e., to adapt the spatial triangulation until

$$\mathcal{E}_c^2, \mathcal{E}_{\mathcal{G}}^2 \leq \mathcal{E}_{\tau}^2 + \mathcal{E}_f^2. \quad (4.3)$$

Here we have invoked the consistency indicator  $\mathcal{E}_f$  on the right-hand side although it is controlled by CONSISTENCY outside ST\_ADAPTATION—note that  $\mathcal{E}_f$  does not depend on the discrete solution. In fact, from the uniform energy estimate, Corollary 3.11, we have that  $\mathcal{E}_{\tau}$  vanishes faster than  $\mathcal{E}_f$  by one order, when no additional regularity of  $f$  is assumed. Consequently, in this case the time-step size is dictated by  $\mathcal{E}_f$ , which may lead to  $\mathcal{E}_{\tau} \ll \text{tol}_{\mathcal{G}\tau}$ . Thanks to Lemma 4.4 we expect that (4.3) leads to an equal distribution of the errors in time in most cases. However, the case  $\max\{\mathcal{E}_{\tau}, \mathcal{E}_f\} \ll \min\{\text{tol}_{\mathcal{G}\tau}, \text{tol}_f\}$  cannot be avoided theoretically; hence, we have supplemented (4.3) with the safeguard  $L^\infty$  marking tolerance  $\tau \text{tol}_{\mathcal{G}\tau}$ ; compare with lines 8 and 15 of ST\_ADAPTATION. In particular, in order to not waste tolerances we have implemented the safeguard additively thanks to

$$\max\{\mathcal{E}_{\tau}^2 + \mathcal{E}_f^2, \tau \text{tol}_{\mathcal{G}\tau}\} \leq \mathcal{E}_{\tau}^2 + \mathcal{E}_f^2 + \tau \text{tol}_{\mathcal{G}\tau} \leq 2 \max\{\mathcal{E}_{\tau}^2 + \mathcal{E}_f^2, \tau \text{tol}_{\mathcal{G}\tau}\}.$$

Note that in the above discussion we have concentrated on an equal distribution in time and have tacitly assumed that in each time step the local space indicators are optimally distributed, which is motivated by the optimal convergence analysis for elliptic problems; compare, e.g., with the studies by Cascon *et al.* (2008); Diening *et al.* (2016) and Stevenson (2007).

**REMARK 4.6** We note that the **if** conditions in lines 15 and 8 of ST\_ADAPTATION may involve additional parameters. For instance, line 8 may be replaced by

15 : **else if**  $\mathcal{E}_c^2(U_t^-, \tau, \mathcal{G}) \geq \gamma_c \mathcal{E}_{\tau}^2(U_t^+, U_t^-, \tau, \mathcal{G}) + \rho_c \mathcal{E}_f^2(f, t, \tau) + \sigma_c \tau \text{tol}_{\mathcal{G}\tau}$  **then**

with  $\gamma_c, \rho_c, \sigma_c > 0$  and similar for the space indicator  $\mathcal{E}_{\mathcal{G}}$  in line 15 with constants  $\gamma_{\mathcal{G}}, \rho_{\mathcal{G}}, \sigma_{\mathcal{G}} > 0$ . This requires some modifications of TAFEM, which would make the presentation more technical. For the sake of clarity of the presentation we decided to skip these customisation possibilities; compare also with Remark 5.3.

#### 4.3 The main module TAFEM

We are now in a position to formulate TAFEM in Algorithm 5 below.

In the initialisation phase the given tolerance  $\text{TOL} > 0$  is split into tolerances  $\text{TOL}_0, \text{TOL}_f, \text{TOL}_{\mathcal{G}\tau} > 0$ . Next ADAPT\_INIT provides a sufficiently good approximation  $U_0$  of the initial datum  $u_0$ . The constant  $C_T$  is computed in order to determine the right scaling of the tolerances in the module ST\_ADAPTATION. In particular, the last term  $2T$  in  $C_T$  accounts for the  $L^\infty$  safeguard marking, whereas the other terms result from (4.3) in combination with the estimate for  $\mathcal{E}_{\tau}$  from Corollary 3.11; compare also with the proof of Theorem 5.2. Then the time-step iteration is entered, where each single time step consists of the following main steps. We first initialise the time-step size by CONSISTENCY and then conduct one coarsening step with COARSEN. The adaptation of the grid and time-step size with respect to the indicators for the spatial, temporal and coarsening error is done by ST\_ADAPTATION.

**Algorithm 5** TAFEM

---

```

1: initialise  $\mathcal{G}_{\text{init}}$ ,  $\tau_0$  and set  $t_0 = 0, n = 0$ 
2: split tolerance  $\text{TOL} > 0$  such that  $\text{TOL}_0^2 + 3 \text{TOL}_f^2 + \text{TOL}_{\mathcal{G}\tau}^2 = \text{TOL}^2$ 
3:  $\text{tol}_f = \text{TOLFIND}(f, T, \text{TOL}_f)$ 
4:  $(U_0^-, \mathcal{G}_0) = \text{ADAPT\_INIT}(u_0, \mathcal{G}_{\text{init}}, \text{TOL}_0)$ 
5: compute  $C_T := 6\sqrt{6C_\tau T} \left( \|f\|_{\Omega \times (0,T)}^2 + \|U_0^-\|^2 \right)^{\frac{1}{2}} + 2T$ 
6: do
7:    $n = n + 1$ 
8:    $\tau_n = \min \{\tau_{n-1}, T - t\}$ 
9:    $\tau_n = \text{CONSISTENCY}(f, t_{n-1}, \tau_{n-1}, \text{tol}_f)$ 
10:   $\mathcal{G}_n = \text{COARSEN}(U_{n-1}^-, \mathcal{G}_{n-1})$ 
11:   $(U|_{I_n}, \tau_n, f_n, \mathcal{G}_n) = \text{ST\_ADAPTATION}(U_{n-1}^-, t_n, \tau_n, f, \mathcal{G}_n, \mathcal{G}_{n-1}, \text{TOL}_{\mathcal{G}\tau}^2/C_T)$ 
12:   $U_n^- = U|_{I_n}(t_{n-1} + \tau_n)$ 
13: while  $t_n = t_{n-1} + \tau_n < T$ 

```

---

**5. Convergence**

In this section we first prove that the core modules and TAFEM terminate and then verify that the estimators and thus the error are below the given tolerance. Throughout the section we suppose that the black-box modules satisfy Assumption 4.1.

Before turning to the main module ST\_ADAPTATION, as an auxiliary result, we shall consider convergence of the adaptive finite element method for stationary elliptic problems of the kind (3.13), which have to be solved in each time step.

**Algorithm 6** AFEM

---

```

AFEM( $v^-, \bar{f}, t, \tau, \mathcal{G}^0$ )
1: set  $k = 0$ 
2: loop forever
3:    $U_\tau^k = \text{SOLVE}(v^-, \bar{f}, 0, \tau, \mathcal{G}^k)$ 
4:   compute  $\left\{ \mathcal{E}_{\mathcal{G}}^2(U_\tau^k, v^-, 0, \tau, \bar{f}, \mathcal{G}, E) \right\}_{E \in \mathcal{G}}$ 
5:    $\mathcal{G}^{k+1} = \text{MARK\_REFINE}\left(\left\{ \mathcal{E}_{\mathcal{G}}^2(U_\tau^k, v^-, 0, \tau, \bar{f}, \mathcal{G}^k, E) \right\}_{E \in \mathcal{G}}, \mathcal{G}^k \right)$ 
6:    $k = k + 1$ 
7: end loop forever

```

---

**PROPOSITION 4.7** (Convergence for the elliptic problem). Suppose that  $v^- \in L^2(\Omega)$ ,  $\bar{f} \in \mathbb{P}_s(L^2(\Omega))$  and  $\tau > 0$ . Then starting from any grid  $\mathcal{G}^0 \in \mathbb{G}$  we have for the sequence  $\{\mathcal{G}^k, U_\tau^k\}_{k \geq 0} \subset \mathbb{G} \times \mathbb{P}_s(\mathbb{V})$

generated by AFEM( $v^-$ ,  $\bar{f}$ ,  $t$ ,  $\tau$ ,  $\mathcal{G}^0$ ) that

$$\mathcal{E}_{\mathcal{G}}^2 \left( U_{\tau}^k, v^-, \tau, t, \bar{f}, \mathcal{G}^k \right) \rightarrow 0 \quad \text{as } k \rightarrow \infty.$$

*Proof.* Recalling Remark 3.9 we have that  $\mathcal{E}_{\mathcal{G}}^2(U_{\tau}^k, v^-, \tau, t, \bar{f}, \mathcal{G}^k)$  are the standard residual-based *a posteriori* error estimators for the coercive problem (3.13). From Lemmas 3.2 and 3.3 and Assumption 4.1 on MARK\_REFINE we have that the conditions of Siebert (2011) are satisfied. This yields the assertion.  $\square$

LEMMA 4.8 (Termination of ST\_ADAPTATION). For any  $t \in (0, T)$ ,  $\tau^{\text{in}} \in (0, T - t]$ ,  $\mathcal{G}, \mathcal{G}_{\text{old}} \in \mathbb{G}$  and  $U_t^- \in \mathbb{V}(\mathcal{G}_{\text{old}})$  we have that

$$(U_I, \tau, \bar{f}, \mathcal{G}) = \text{ST\_ADAPTATION}\left(U_t^-, f, t, \tau^{\text{in}}, \mathcal{G}_{\text{in}}, \mathcal{G}_{\text{old}}, \text{tol}_{\mathcal{G}\tau}\right)$$

terminates. Moreover, we have  $\mathcal{G} \geq \mathcal{G}_0$ ,  $\mathcal{E}_{*}^2(U_t^+, U_t^-, \tau, \mathcal{G}) \leq 0$ ,

$$\tau^{\text{in}} \geq \tau \geq \min \left\{ \tau^{\text{in}}, \frac{\kappa \text{tol}_{\mathcal{G}\tau}^2}{6 \left( \|f\|_{\Omega \times (t, t+\tau)}^2 + \|U_t^-\|_{\Omega}^2 \right)} \right\}$$

and the indicators satisfy the tolerances

$$\begin{aligned} \mathcal{E}_{\tau}^2(U_t^+, U_t^-, \tau, \mathcal{G}) &\leq \text{tol}_{\mathcal{G}\tau}^2, \\ \mathcal{E}_{\mathcal{G}}^2(U_I, U_t^-, t, \tau, \bar{f}, \mathcal{G}) &\leq \mathcal{E}_{\tau}^2(U_t^+, U_t^-, \tau, \mathcal{G}) + \mathcal{E}_{\bar{f}}^2(f, t, \tau) + \tau \text{tol}_{\mathcal{G}\tau}, \\ \mathcal{E}_{\mathcal{c}}^2(U_t^-, \tau, \mathcal{G}) &\leq \mathcal{E}_{\tau}^2(U_t^+, U_t^-, \tau, \mathcal{G}) + \mathcal{E}_{\bar{f}}^2(f, t, \tau) + \tau \text{tol}_{\mathcal{G}\tau}, \end{aligned}$$

where  $U_t^+ = \lim_{s \searrow t} U_{(t, t+\tau]}(s)$ .

*Proof.* In each iteration of the loop in ST\_ADAPTATION at first, a discrete solution  $U_I$  is computed on the current grid  $\mathcal{G}$  with the actual time-step size  $\tau$ . Then either the time-step size is reduced or the actual grid is refined. More precisely, exactly one of the statements labelled **A**, ..., **D** in Algorithm 4 is executed, any of them terminating by Assumption 4.1. Whenever one of these statements is executed the corresponding indicator is positive.

In statement **A** the grid is refined due to the coarsening indicator  $\mathcal{E}_{\mathcal{c}}$ . Thanks to Assumption 4.1 (5), after a finite number of executions of **A**, a grid  $\mathcal{G}$  is obtained with  $\mathcal{G}_{\text{old}} \leq \mathcal{G}$  and thus  $\mathcal{E}_{\mathcal{c}}^2(U_t^-, \mathcal{G}) = 0$ , i.e., statement **A** is not entered anymore. This happens irrespective of refinements in other statements.

In statement **B** the grid is refined with respect to the indicators  $\mathcal{E}_{*}$  controlling the energy gain due to coarsening. Therefore, it follows from the same reasoning as for statement **A** that statement **B** is also executed at most until the coarsening is fully removed after finitely many refinement steps.

Assume that statement **C** is executed infinitely many times. Then according to the above considerations, we have that, after a finite number of iterations, statement **B** is not executed anymore,

i.e., (3.15) is always satisfied according to line 10. Consequently, the condition in line 12 and Corollary 3.11 imply

$$\frac{1}{\tau} \leq \frac{1}{\tau} 6 \tau \|U_t^+ - \Pi_{\mathcal{G}} U_t^-\|_{\Omega}^2 \frac{1}{\text{tol}_{\mathcal{G}\tau}^2} \leq \frac{1}{\text{tol}_{\mathcal{G}\tau}^2} 6 \left( \|f\|_{\Omega \times (t, t+\tau)}^2 + \|U_t^-\|_{\Omega}^2 \right),$$

which contradicts the assumption that statement  $\boxed{\mathbf{C}}$  is executed infinitely many times. The same argument also proves the asserted lower bound on the finally accepted time-step size.

Assuming that ST\\_ADAPTATION does not terminate we infer from the fact that all other statements are conducted only finitely many times, that statement  $\boxed{\mathbf{D}}$  has to be executed infinitely many times. In other words the loop reduces to the adaptive iteration AFEM with fixed data  $U_t^-, \bar{f}, t$  and  $\tau$ . Therefore, Proposition 4.7 contradicts the condition in line 15.

In summary, we deduce that ST\\_ADAPTATION terminates and the iteration is abandoned in line 18. This proves the assertion.  $\square$

We next address the termination of the main module TAFEM .

**PROPOSITION 5.1** (Termination of TAFEM). The adaptive algorithm TAFEM terminates for any initial time-step size  $\tau_0 > 0$  and produces a finite number of time instances  $0 = t_0 < \dots < t_N = T$ .

Moreover, we have  $\mathcal{E}_0^2(u_0, \mathcal{G}_0) \leq \text{tol}_0^2$  and that the consistency error complies with (4.2). For all  $n = 1, \dots, N$  we have that the estimates in Lemma 4.8 are satisfied with  $t = t_{n-1}$ ,  $\tau = \tau_n$ ,  $U_I = U|_{I_n}$ ,  $U_t^\pm = U_{n-1}^\pm$ ,  $\mathcal{G} = \mathcal{G}_n$  and  $\mathcal{G}_{\text{old}} = \mathcal{G}_{n-1}$ .

*Proof.* Each loop starts with setting the time-step size such that  $\tau_n \leq T - t_n$ ,  $n \in \mathbb{N}$ . Thanks to Assumption 4.1 for the black-box modules, Lemma 4.2 for CONSISTENCY and Lemma 4.8 for ST\\_ADAPTATION, all modules of TAFEM terminate and in each time step the asserted properties are satisfied.

Since we have  $\mathcal{E}_*^2(U_{n-1}^+, U_{n-1}^-, \tau_n, \mathcal{G}_n) \leq 0$  for all  $n$  we may conclude  $\|U_{n-1}^-\|_{\Omega} \leq \|f\|_{\Omega \times (0, T)}^2 + \|U_0\|_{\Omega}$  from Lemma 3.11 and thus it follows with Lemma 4.8 that

$$\tau_n^{\text{in}} \geq \tau_n \geq \min \left\{ \tau_n^{\text{in}}, \frac{\kappa \text{tol}_{\mathcal{G}\tau}^2}{12 \left( \|f\|_{\Omega \times (0, T)}^2 + \|U_0\|_{\Omega}^2 \right)} \right\},$$

where  $\tau_n^{\text{in}} = \text{CONSISTENCY}(f, t_{n-1}, \tau_{n-1}, \text{tol}_f)$ . Assuming that the final time is not reached implies  $\tau_n \rightarrow 0$  as  $n \rightarrow \infty$  and therefore there exists  $n_0 \in \mathbb{N}$  such that  $\tau_n = \tau_n^{\text{in}}$  for all  $n \geq n_0$ . Now the contradiction follows as in step (1) of the proof of Lemma 4.4.  $\square$

Collecting the results derived above allows us to prove the main result.

**THEOREM 5.2** (Convergence into tolerance). Algorithm TAFEM computes for any prescribed tolerance  $\text{TOL} > 0$  and initial time-step size  $\tau_0 > 0$  a partition  $0 < t_0 < \dots < t_N = T$  with associated meshes  $\{\mathcal{G}_n\}_{n=0,\dots,N}$ , such that we have for the corresponding approximation  $\mathcal{U} \in \mathbf{W}$  from (2.7) that

$$\|u - \mathcal{U}\|_{\mathbf{W}(0, T)} \leq \text{TOL}.$$

*Proof.* Thanks to Proposition 5.1 we have that TAFEM terminates and it remains to prove the error bound. For the sake of brevity of the presentation we shall use the abbreviations

$$\begin{aligned}\mathcal{E}_\tau^2(n) &:= \mathcal{E}_\tau^2(U_{n-1}^+, U_{n-1}^-, \tau_n, \mathcal{G}_n), & \mathcal{E}_f^2(n) &:= \mathcal{E}_f^2(f, t_{n-1}, \tau_n), \\ \mathcal{E}_{\mathcal{G}}^2(n) &:= \mathcal{E}_{\mathcal{G}}^2(U, U_{n-1}^-, t_{n-1}, \tau_n, f_n, \mathcal{G}_n) \quad \text{and} \quad \mathcal{E}_c^2(n) &:= \mathcal{E}_c^2(U_{n-1}^-, \tau_n, \mathcal{G}_n).\end{aligned}$$

The initial error satisfies  $\mathcal{E}_0^2(u_0, \mathcal{G}_0) \leq \text{TOL}_0^2$  by Assumption 4.1. Thanks to the choice of the precomputed local tolerance  $\text{tol}_f$  we know from Lemma 4.4 that the consistency error is bounded by  $\text{TOL}_f$ , i.e., we have (4.2).

When finalizing a time step, we also have from Lemma 5.1 that

$$\mathcal{E}_\tau^2(n) \leq \text{tol}_{\mathcal{G}\tau}^2 \quad \text{and} \quad \mathcal{E}_{\mathcal{G}}^2(n), \mathcal{E}_c^2(n) \leq \mathcal{E}_\tau^2(n) + \mathcal{E}_f^2(n) + \tau_n \text{tol}_{\mathcal{G}\tau},$$

with  $\text{tol}_{\mathcal{G}\tau} = \text{TOL}_{\mathcal{G}\tau}^2 / C_T$ . Combining this with (3.16a) and (4.2) we conclude that

$$\begin{aligned}\sum_{n=1}^N \mathcal{E}_{\mathcal{G}}^2(n) + \mathcal{E}_{c\tau}^2(U_{n-1}^+, U_{n-1}^-, \tau_n) &\leq \sum_{n=1}^N \mathcal{E}_{\mathcal{G}}^2(n) + \mathcal{E}_c^2(n) + \mathcal{E}_\tau^2(n) \\ &\leq \sum_{n=1}^N 2\tau_n \text{tol}_{\mathcal{G}\tau} + 2\mathcal{E}_f^2(n) + 3\mathcal{E}_\tau^2(n) \\ &\leq 2T \text{tol}_{\mathcal{G}\tau} + 2\text{TOL}_f^2 + 3 \sum_{n=1}^N \mathcal{E}_\tau^2(n).\end{aligned}$$

Using Corollary 3.11 for the last term we get for any  $\delta > 0$  that

$$\begin{aligned}\sum_{n=1}^N \mathcal{E}_\tau^2(n) &= \sum_{\tau_n > \delta} \mathcal{E}_\tau^2(n) + \sum_{\tau_n \leq \delta} \mathcal{E}_\tau^2(n) \\ &\leq \frac{T}{\delta} \text{tol}_{\mathcal{G}\tau}^2 + \delta \sum_{n=1}^N 6C_\tau \|U_{n-1}^+ - \Pi_{\mathcal{G}_n} U_{n-1}^-\|_{\Omega}^2 \\ &\leq \frac{T}{\delta} \text{tol}_{\mathcal{G}\tau}^2 + \delta 6C_\tau \left( \|f\|_{\Omega \times (0,T)}^2 + \|U_0\|_{\Omega}^2 \right)\end{aligned}$$

and by choosing

$$\delta = \left( \frac{T}{6C_\tau \left( \|f\|_{\Omega \times (0,T)}^2 + \|U_0\|_{\Omega}^2 \right)} \right)^{\frac{1}{2}} \text{tol}_{\mathcal{G}\tau},$$

we obtain

$$\sum_{n=1}^N \mathcal{E}_\tau^2(n) \leq 2 \left( 6C_\tau T \left( \|f\|_{\Omega \times (0,T)}^2 + \|U_0\|_{\Omega}^2 \right) \right)^{\frac{1}{2}} \text{tol}_{\mathcal{G}\tau}.$$

Inserting this into the above estimate yields

$$\begin{aligned} \sum_{n=1}^N \mathcal{E}_\mathcal{G}^2(n) + \mathcal{E}_{c\tau}^2(U_{n-1}^+, U_{n-1}^-, \tau_n) &\leq \underbrace{\left( 6\sqrt{6C_\tau T} \left( \|f\|_{\Omega \times (0,T)}^2 + \|U_0\|_{\Omega}^2 \right)^{\frac{1}{2}} + 2T \right)}_{=C_T} \text{tol}_{\mathcal{G}\tau} + 2 \text{TOL}_f^2 \\ &\leq \text{TOL}_{\mathcal{G}\tau}^2 + 2 \text{TOL}_f^2. \end{aligned}$$

Collecting the bounds for the indicators  $\mathcal{E}_0$ ,  $\mathcal{E}_\mathcal{G}$ ,  $\mathcal{E}_{c\tau}$ , and  $\mathcal{E}_f$ , recalling the splitting

$$\text{TOL}_0^2 + 3 \text{TOL}_f^2 + \text{TOL}_{\mathcal{G}\tau}^2 = \text{TOL}^2$$

and taking into account the upper bound of Proposition 3.4 proves the assertion.  $\square$

**REMARK 5.3** In order to guarantee the main result (Theorem 5.2) also for the modifications of Remark 4.6, line 5 in TAFEM must be changed to

$$5: \text{compute } C_T := (1 + \gamma_c + \gamma_g) 2 \sqrt{6C_\tau T} \left( \|f\|_{\Omega \times (0,T)}^2 + \|U_0^-\|_{\Omega}^2 \right)^{\frac{1}{2}} + (\sigma_c + \sigma_g) T.$$

Moreover, the splitting of the tolerances in line 2 must be changed to

$$2: \text{split tolerance } \text{TOL} > 0 \text{ such that } \text{TOL}_0^2 + (1 + \rho_g + \rho_c) \text{TOL}_f^2 + \text{TOL}_{\mathcal{G}\tau}^2 = \text{TOL}^2.$$

## 6. Numerical aspects and experiments

We conclude the article by illustrating some practical aspects of the implementation with three numerical experiments. We compare the presented algorithm TAFEM with the algorithm ASTFEM introduced in the study by Kreuzer *et al.* (2012).

### 6.1 The implementation

The experiments are implemented in DUNE (Blatt *et al.*, 2016) using the DUNE-ACFEM (<https://www.dune-project.org/modules/dune-acfem/>) module. The computations utilise linear finite elements in space and the dG(0) time-stepping scheme. All simulations were performed on an AMD® Opteron™7274 Processor with 128 GB RAM.

Both algorithms TAFEM and ASTFEM start from exactly the same initial mesh  $\mathcal{G}_{\text{init}}$ . The initial values are interpolated on the mesh and local refinements are performed in order to comply with the initial tolerance. On the resulting meshes the needed constants are computed (the minimal time-step size  $\tau^*$  for ASTFEM and  $\text{tol}_f$  from TOLFIND for TAFEM ).

In order to control  $\mathcal{E}_c$  and  $\mathcal{E}_*$  the algorithms need to handle two meshes and corresponding finite element spaces at every new time step. This is realised by exploiting the tree structure of the refinements of macroelements as in the study by Kreuzer *et al.* (2012). At every new time step the time-step size is adjusted by the module CONSISTENCY and all elements of the mesh are marked to be coarsened up to two times and then adapted again if necessary. The mentioned estimators are computed only

up to constants and used for the adaptive refinement progress. The spatial marking relies on the equidistribution strategy, which marks every element with an estimator bigger than the arithmetic mean.

The following remark lists the tolerance splitting used by ASTFEM.

**REMARK 6.1** In the study by Kreuzer *et al.* (2012), ASTFEM uses the tolerance splitting

$$\text{TOL}^2 = \text{TOL}_0^2 + T\widetilde{\text{TOL}}_f^2 + T\widetilde{\text{TOL}}_{\mathcal{G}\tau}^2 + \widetilde{\text{TOL}}_*^2.$$

Thereby,  $\text{TOL}_*^2$  is used to compute a minimal safeguard step size  $\tau_*$ . The method computes then an approximation  $\mathcal{U} \in \mathbf{W}$  to (2.2) such that

$$\mathcal{E}_0^2(u_0, \mathcal{G}_0) \leq \text{TOL}_0^2, \quad \sum_{n=1}^N \left\{ \mathcal{E}_f^2(f, t_{n-1}, \tau_n) \right\} \leq T\widetilde{\text{TOL}}_f^2$$

and

$$\sum_{n=1}^N \left\{ \mathcal{E}_0^2(U_{n-1}^+, U_{n-1}^-, \tau_n) + \mathcal{E}_0^2(U, U_{n-1}^-, t_n, \tau_n, f_n, \mathcal{G}_n) \right\} \leq T\widetilde{\text{TOL}}_{\mathcal{G}\tau}^2 + \widetilde{\text{TOL}}_*^2.$$

This motivates the relations

$$T\widetilde{\text{TOL}}_f^2 = 3\text{TOL}_f^2 \quad \text{and} \quad T\widetilde{\text{TOL}}_{\mathcal{G}\tau}^2 + \widetilde{\text{TOL}}_*^2 = \text{TOL}_{\mathcal{G}\tau}^2$$

in the examples below.

For the simulations presented below we have used the following comparable splittings for the two methods ASTFEM and TAFEM relative to the total tolerance  $\text{TOL}$ :

- $\text{TOL}_0^2 = \frac{1}{10}\text{TOL}^2$ ,
- $\text{TOL}_f^2 = T\widetilde{\text{TOL}}_f^2 = \frac{4}{10}\text{TOL}^2$ ,
- $\text{TOL}_{\mathcal{G}\tau}^2 = T\widetilde{\text{TOL}}_{\mathcal{G}\tau}^2 + \widetilde{\text{TOL}}_*^2 = \frac{6}{10}\text{TOL}^2$ ,
- $\widetilde{\text{TOL}}_*^2 = \frac{1}{100}\text{TOL}^2$ .

## 6.2 The experiments

In this section we introduce the three numerical experiments in detail and discuss the numerical results.

**6.2.1 Singularity in time.** This numerical experiment is constructed on the spatial domain  $\Omega = (0, 1)^2 \subset \mathbb{R}^2$  over the time interval  $(0, T) = (0, 2)$  with homogeneous Dirichlet boundary conditions and homogeneous initial data. The right-hand side  $f$  is chosen such that the exact solution is given by

$$u(x, t) = |t - \bar{t}|^\alpha \sin(\pi(x^2 - x)t) \sin(\pi(y^2 - y)t)$$

with parameters  $\bar{t} = \frac{\pi}{3}$  and  $\alpha = 0.7$ . The graph of  $u$  has a singularity in time at  $t = \frac{\pi}{3}$ . Hence, the right-hand side contains the term  $\text{sgn}(t - \bar{t})\alpha|t - \bar{t}|^{\alpha-1}$ . A direct calculation shows that this term is  $L^2$ -integrable but is not in  $H^1$ . This particular example shows one main advantage of TAFEM. In fact, in contrast to ASTFEM, TAFEM does not require the right-hand side  $f$  to have a temporal derivative in  $L^2$  in order to control the consistency error  $\mathcal{E}_f$ . As in the example of Section 6.2.2 the threshold-like

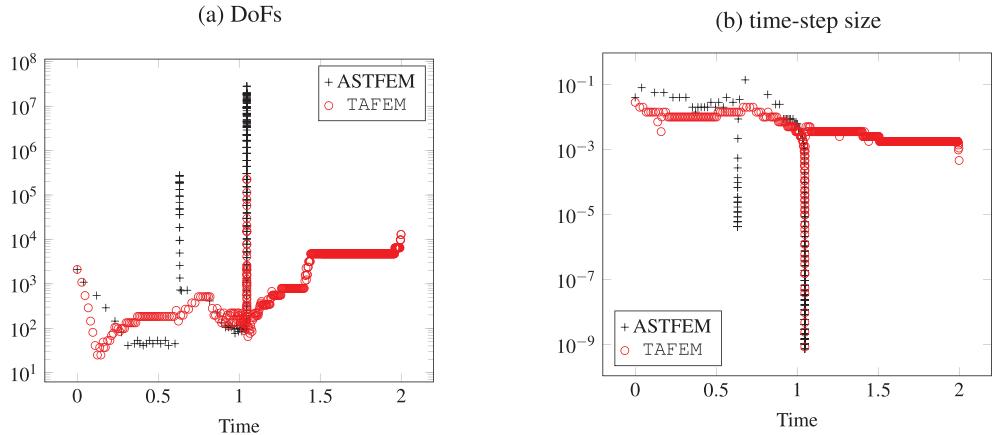
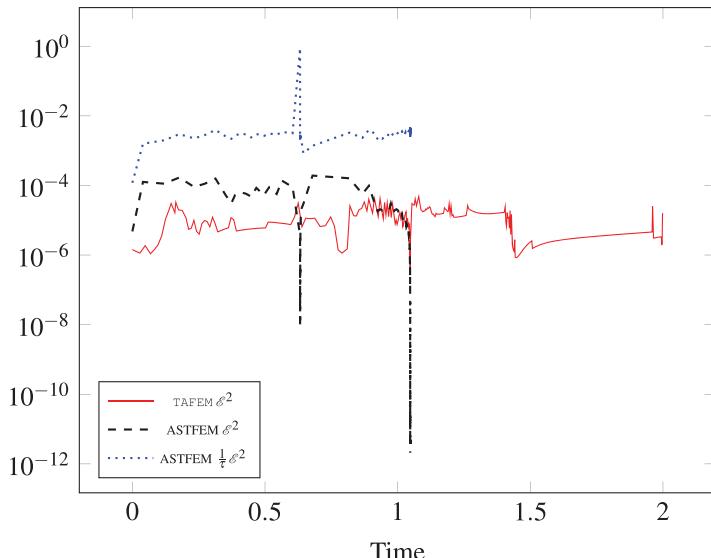


FIG. 1. DoFs and time-step sizes for the singularity in time problem.

FIG. 2. The local error estimators  $\mathcal{E}_\tau^2 + \mathcal{E}_G^2 + \mathcal{E}_c^2 + \mathcal{E}_f^2$  for TAFEM and ASTFEM, as well as the sum of local  $L^\infty$  indicators  $\frac{1}{\tau}(\mathcal{E}_\tau^2 + \mathcal{E}_G^2 + \mathcal{E}_c^2 + \mathcal{E}_f^2)$  used by ASTFEM for the singularity in time problem.

version of ENLARGE from Remark 4.3 is used for the computations, with enlargement constant  $\kappa_2 = 4$ . However, test runs with the fixed enlargement, i.e.,  $\min\{\kappa_2\tau, T-t\} = \text{ENLARGE}(f, t, \tau, \text{tol}_f)$ , produce nearly the same results.

ASTFEM was killed after time step 321 in which 27 513 668 DoFs are used as well as a time-step size of  $7.3172e-10$ . As can be observed from Fig. 1, ASTFEM massively refines in time and space. It was killed before reaching the singularity at  $\tilde{t} = \frac{\pi}{3}$ , thereby accumulating a total number of 1070 889 132 DoFs. The reason for this behaviour lies in the  $L^\infty$  marking. Indeed, Fig. 2 shows

TABLE 1 Time steps and DoFs of ASTFEM and TAFEM for the singularity in time problem

Time	Time step ASTFEM	DoFs ASTFEM	Time step TAFEM	DoFs TAFEM
1.0	0.00613614	97	0.00353591	221
1.02	0.00433893	85	0.00250028	221
1.03	0.0030681	97	0.0025003	105
1.04	0.00153406	97	0.00125015	181
1.042	0.00108475	97	0.00125016	128
1.044	0.00108475	97	0.000883996	128
1.047	3.3899e-05	713	0.0001105	478
1.0471	1.19852e-05	2073	3.90677e-05	1136
1.0472	9.36409e-08	27513668	4.76905e-09	224914
1.0473	non	non	5.5251e-05	759

that ASTFEM equally distributes the  $L^\infty$  indicators, thereby leading to very small local errors, which cause the strong spatial refinement. Note that the minimal step size  $\tau_*$  in ASTFEM applies only when temporal refinement is performed due to the time indicator  $\mathcal{E}_\tau$ , i.e., time-step sizes below the threshold  $\tau_*$  can be chosen when required by the consistency estimator  $\mathcal{E}_f$ , which is the case close to the singularity. Consequently, the behaviour of ASTFEM cannot essentially be improved by a different choice of  $TOL_*$ . In contrast, the local estimators in TAFEM appear to be quite equally distributed. It uses slightly larger time steps and far fewer DoFs close to the singularity; compare with the table of Table 1. It completely outperforms ASTFEM and reaches the final time with a total of 2 481 808 DoFs in 575 time steps.

**6.2.2 Jumping singularity.** Inspired by Morin *et al.* (2000, example 5.3) we construct an experiment where the solution has a strong spatial singularity that changes its position in time. In the domain  $\Omega \times (0, 4]$ , with  $\Omega = (0, 3) \times (0, 3)$ , we define the elliptic operator  $\mathcal{L}u = -\operatorname{div} \mathbf{A}\nabla u$ , where

$$\mathbf{A}(t, x) = \begin{cases} a_1 \mathbb{I} & \text{if } (x - x_i)(y - y_i) \geq 0, \\ a_2 \mathbb{I} & \text{if } (x - x_i)(y - y_i) < 0, \end{cases}$$

with  $a_1 = 161.4476387975881$ ,  $a_2 = 1$ ,  $i = \lceil t \rceil$ ,  $(x_1, y_1) = (1, 2)$ ,  $(x_2, y_2) = (1, 1)$ ,  $(x_3, y_3) = (2, 1)$  and  $(x_4, y_4) = (2, 2)$ . This operator will ‘move’ the singularity through the points  $x_i$ . Let  $u$  be the function

$$u(x, t) = \sum_{i=1}^4 s_i(t) r_i^\gamma \mu(\theta_i),$$

where

$$s_i(t) = \begin{cases} (t - (i-1))^2(t - i)^2 & \text{if } i-1 \leq t \leq i, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\mu(\theta) = \begin{cases} \cos((\frac{\pi}{2} - \sigma)\gamma) \cos((\theta - \frac{\pi}{2} + \rho)\gamma) & \text{if } 0 \leq \theta < \frac{1}{2}\pi, \\ \cos(\rho\gamma) \cos((\theta - \pi + \sigma)\gamma) & \text{if } \frac{1}{2}\pi \leq \theta < \pi, \\ \cos(\sigma\gamma) \cos((\theta - \pi - \rho)\gamma) & \text{if } \pi \leq \theta < \frac{3}{2}\pi, \\ \cos((\frac{\pi}{2} - \rho)\gamma) \cos((\theta - \frac{3\pi}{2} - \sigma)\gamma) & \text{if } \frac{3}{2}\pi \leq \theta < 2\pi, \end{cases}$$

with  $\gamma = 0.1$ ,  $\rho = \frac{\pi}{4}$ ,  $\sigma = -14.92256510455152$ ,  $x - x_i = r_i \cos(\theta_i)$  and  $y - y_i = r_i \sin(\theta_i)$ . It is easy to check that  $u$  satisfies

$$\partial_t u(x, t) + \mathcal{L}u(x, t) = \sum_{i=1}^4 r_i^\gamma \mu(\theta_i) \partial_t s_i(t).$$

Based on the ideas presented in Remark 6.1 we compare TAFEM and ASTFEM with the same tolerance  $\text{TOL} = 0.007$ .

ASTFEM makes excessive use of the nonstandard exit, i.e., the time-step sizes equal minimal time-step size  $\tau_* = 0.0123477$  for 231 of a total of 302 time steps, and uses a total of 893 771 DoFs. The  $L^2 - H^1$  error is 0.0546689, the  $L^2 - L^2$  error is 0.0355061 and the total computation time was 628.398 seconds.

TAFEM uses a total of 508 352 DoFs in 228 time steps. The  $L^2(0, 4, H^1(\Omega))$  error is 0.0563979, the  $L^2(0, 4, L^2(\Omega))$  error is 0.0367746 and the total computation time was 646.126 seconds (including TOLFIND). By looking at Fig. 3 we see that TAFEM makes more use of the spatial and temporal adaptivity and achieves a similar result with slightly less effort. The adaptive meshes generated by TAFEM are displayed in Fig. 4. We see that the spatial adaptivity captures the position of the singularity by local refinement and coarsens the region when the singularity has passed by.

The advantages of TAFEM come fully into their own in the presence of singularities in time (see Section 6.2.1). For regular (in time) problems, TAFEM is expected to perform similarly to ASTFEM up to the disadvantage that, at the beginning, the module TOLFIND needs several adaptive iterations over the time span, whereas the computation for the minimal time-step size in ASTFEM iterates only once over the time. This is reflected in the comparable computing times for the jumping singularity problem.

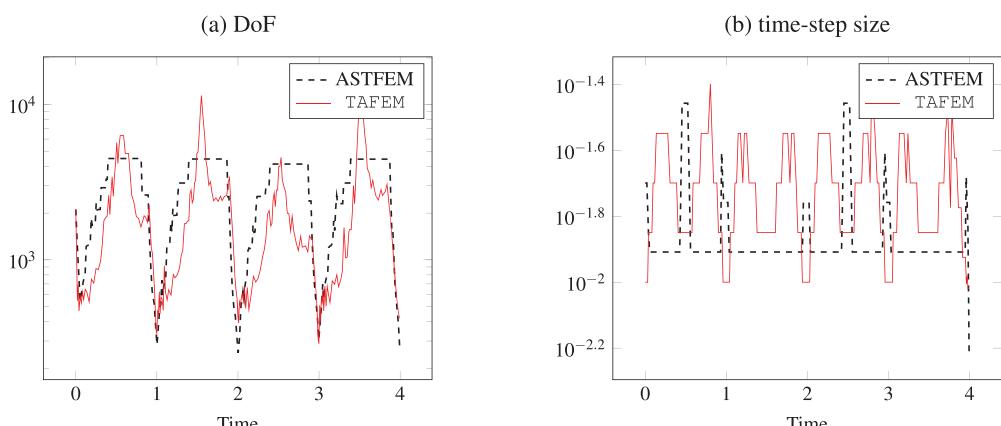


FIG. 3. DoFs and time-step sizes for the jumping singularity problem.

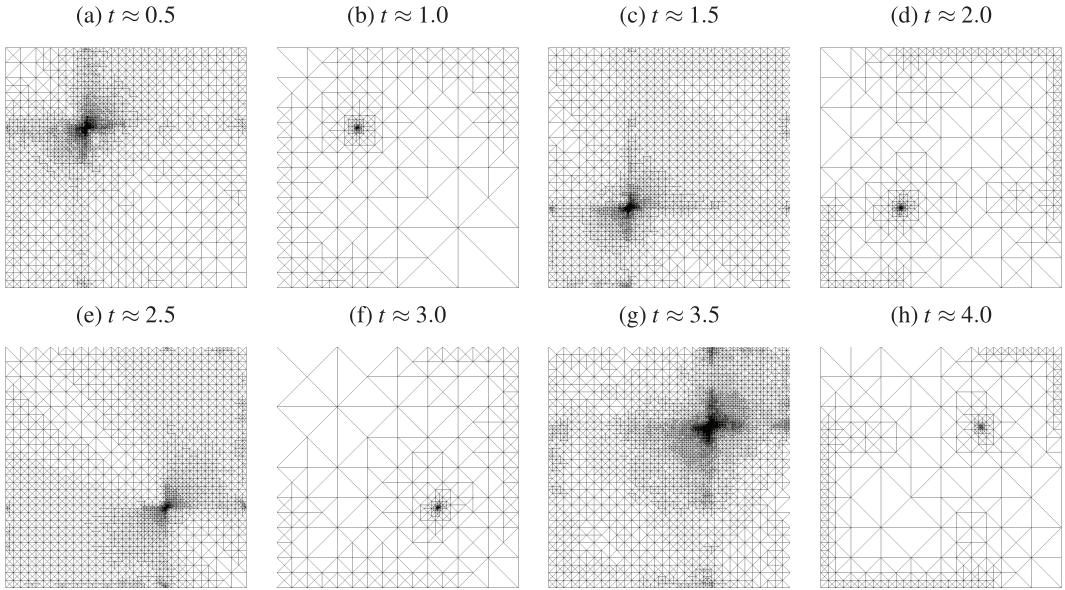
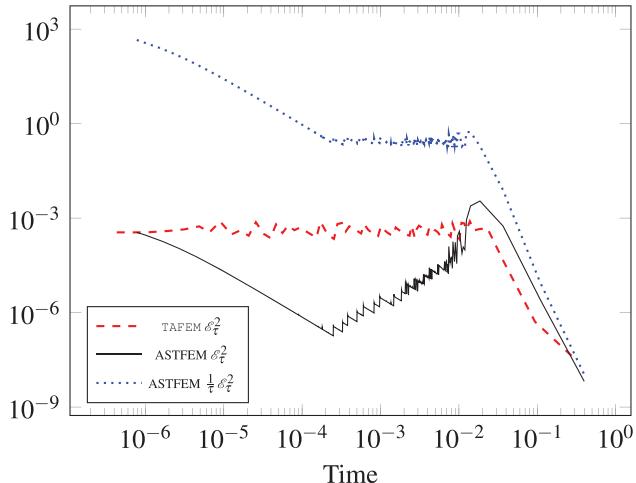


FIG. 4. Adaptive grids for the jumping singularity problem.

FIG. 5. The local time indicator  $\mathcal{E}_t^2$  for TAFEM and ASTFEM, as well as the local  $L^\infty$  indicators  $\frac{1}{\tau}\mathcal{E}_t^2$  used by ASTFEM for the rough initial data problem.

**6.2.3 Rough initial data.** We conclude with an example inspired by Kreuzer *et al.* (2012, numerical experiment 5.3.2) the numerical experiment 5.3.2 with homogeneous Dirichlet boundary conditions and homogeneous right-hand side  $f \equiv 0$ . As initial data we choose a checkerboard pattern over  $\Omega = (0, 1)^2$  where  $u_0 \equiv -1$  on  $\Omega_1 = \left(\frac{1}{3}, \frac{2}{3}\right) \times \left(0, \frac{1}{3}\right) \cup \left(\frac{2}{3}, 1\right) \times \left(0, \frac{1}{3}\right)$  and  $u_0 \equiv 1$  on

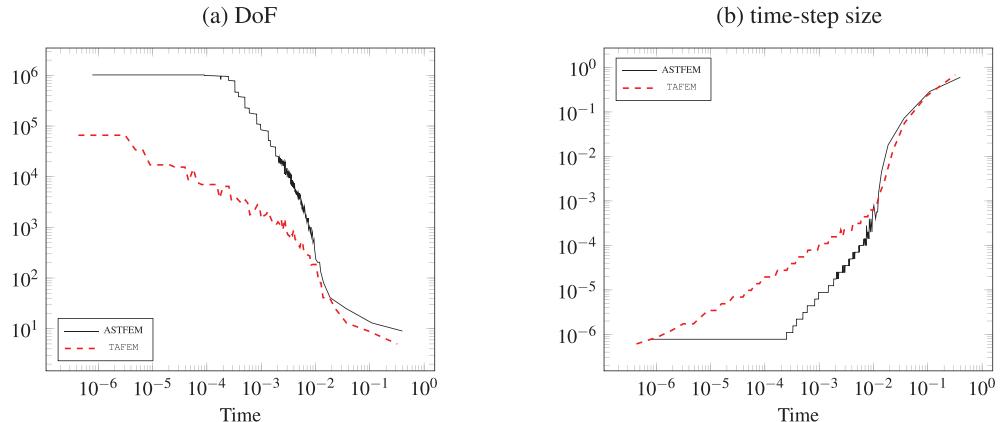


FIG. 6. DoFs and time-step sizes for the rough initial data problem.

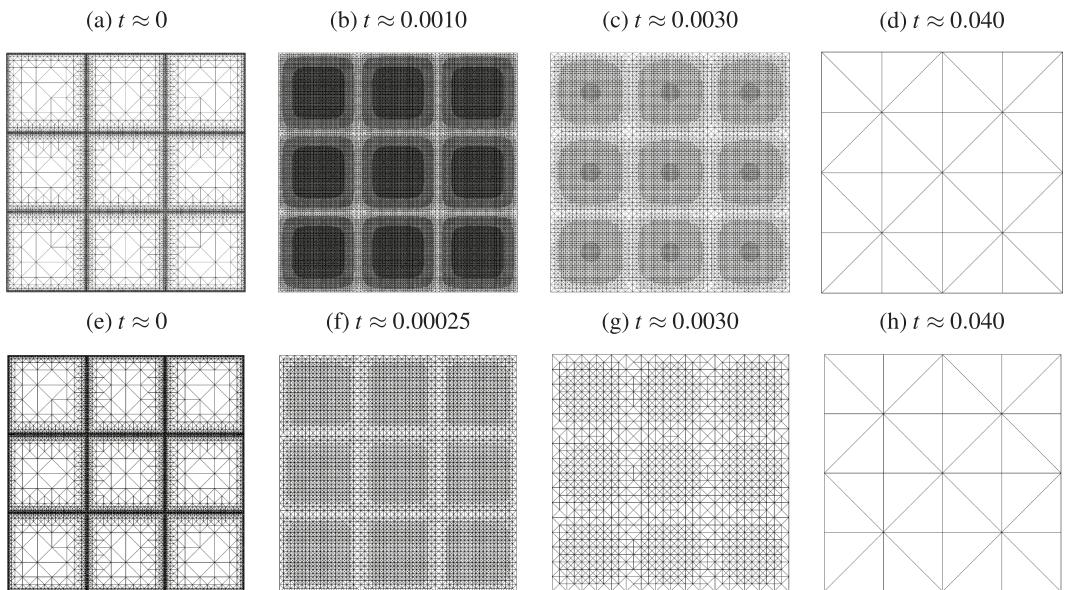


FIG. 7. Adapted meshes generated with ASTFEM (a–d) and TAFEM (e–h) for the rough initial data problem.

$\Omega \setminus \Omega_1$  and  $u_0 \equiv 0$  on  $\partial\Omega$ . Starting with an initial mesh with only 5 DoFs, the approximation of  $u_0$  uses Lagrange interpolation and refines the mesh until  $\|U_0 - u_0\|_{\Omega}^2 \leq \text{TOL}_0^2 = 10^{-2}$  is fulfilled. For this example we choose  $\min\{\kappa_2 \tau, T - t\} = \text{ENLARGE}(f, t, \tau, \text{tol}_f)$ . In fact, since  $f \equiv 0$ , using the threshold-like version of ENLARGE from Remark 4.3 would yield  $T - t = \text{ENLARGE}(f, t, \tau, \text{tol}_f)$  and then in ST\_ADAPTATION the time step must be reduced according to  $\mathcal{E}_\tau$ , again requiring the solution of an expensive linear system in each iteration. This slows down TAFEM by a factor of about 10.

Starting ASTFEM and TAFEM with a tolerance of  $\text{TOL} = 10^{-1}$  and running to the final time  $T = 1$  we get the following results: ASTFEM needs 811 time steps, a total of 436 199 377 DoFs, with an estimated total error of 0.0230905 and a total computation time of 144.114 seconds. ASTFEM makes use of the nonstandard exit for the first 270 time steps, with minimal time-step size of  $\tau_* = 7.77573e-7$ ; the small size of the time steps at the beginning is also accompanied by extreme spatial refinements contributing to the large total number of DoFs. This is due to the  $L^\infty$ -strategy that aims to equally distribute the time-indicators  $\frac{1}{\tau}\mathcal{E}_\tau^2$  rather than  $\mathcal{E}_\tau^2$ . In order to highlight this effect close to the initial time we have used a log scale for the time in Figs. 5 and 6. TAFEM needs only 81 time steps and a total of 672 159 DoFs, resulting in an estimated total error of 0.0574711. It is about 725 times faster with a total computation time of 198.682 seconds (including TOLFIND). TAFEM refines the mesh initially and then almost steadily coarsens in time and space (see Figs. 6 and 7 (e–h)). Figure 5 shows that the time indicators  $\mathcal{E}_\tau^2$  are nearly equally distributed. Both algorithms reduce the spatial resolution once the singular behaviour of the solution is reduced; see Figs. 6 and 7.

## Funding

Deutsche Forschungsgemeinschaft (DFG) project SI 814/7-1.

## REFERENCES

- BÄNSCH, E. (1991) Local mesh refinement in 2 and 3 dimensions. *IMPACT Comput. Sci. Eng.*, **3**, 181–191.
- BINEV, P., DAHMEN, W. & DEVORE, R. A. (2004) Adaptive finite element methods with convergence rates. *Numer. Math.*, **97**, 219–268.
- BINEV, P., DAHMEN, W., DEVORE, R. A. & PETRUSHEV, P. (2002) Approximation classes for adaptive methods. *Serdica Math. J.*, **28**, 391–416.
- BLATT, M., BURCHARDY, A., DEDNER, A., ENGWER, C., FAHLLE, J., FLEMISCH, B., GERSBACHER, C., GRÄSER, C., GRUBER, F., GRÜNINGER, C., KEMPF, D., KLÖFKORN, R., MALKMUS, T., MÜTHING, S., NOLTE, M., PIATKOWSKI, M. & SANDER, O. (2016) The distributed and unified numerics environment (DUNE) *Arch. Numerical Software*, **100**, 13–29.
- CASCON, J. M., KREUZER, C., NOCHETTO, R. H. & SIEBERT, K. G. (2008) Quasi-optimal convergence rate for an adaptive finite element method. *SIAM J. Numer. Anal.*, **46**, 2524–2550.
- CHEN, Z. & FENG, J. (2004) An adaptive finite element algorithm with reliable and efficient error control for linear parabolic problems. *Math. Comp.*, **73**, 1167–1193.
- CLÉMENT, P. (1975) Approximation by finite element functions using local regularization. *Rev. Française Automat. Informat. Recherche Opérationnelle Sér. RAIRO Analyse Numérique*, **9**, 77–84.
- DAVIS, P. J. & RABINOWITZ, P. (1984) *Methods of Numerical Integration*, 2nd edn. Computer Science and Applied Mathematics. Orlando, FL: Academic Press.
- DIENING, L. & KREUZER, C. (2008) Convergence of an adaptive finite element method for the  $p$ -Laplacian equation. *SIAM J. Numer. Anal.*, **46**, 614–638.
- DIENING, L., KREUZER, C. & STEVENSON, R. (2016) Instance optimality of the adaptive maximum strategy. *Found. Comut. Math.*, **16**, 33–68.
- DÖRFLER, W. (1996) A convergent adaptive algorithm for Poisson's equation. *SIAM J. Numer. Anal.*, **33**, 1106–1124.
- ERIKSSON, K. & JOHNSON, C. (1991) Adaptive finite element methods for parabolic problems I: a linear model problem. *SIAM J. Numer. Anal.*, **28**, 43–77.
- ERIKSSON, K. & JOHNSON, C. (1995) 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.
- ERN, A., SMEARS, I. & VOHRALÍK, M. (2017) Guaranteed, locally space-time efficient, and polynomial-degree robust a posteriori error estimates for high-order discretizations of parabolic problems. *SIAM J. Numer. Anal.*, **55**, 2811–2834.

- EVANS, L. C. (2010) *Partial Differential Equations*, 2nd edn. Graduate Studies in Mathematics, vol. 19. Providence, RI: American Mathematical Society.
- GASPOZ, F. D. & MORIN, P. (2014) Approximation classes for adaptive higher order finite element approximation. *Math. Comp.*, **83**, 2127–2160.
- GILBARG, D. & TRUDINGER, N. S. (2001) *Elliptic Partial Differential Equations of Second Order*. Classics in Mathematics. Berlin: Springer.
- KOSSACZKÝ, I. (1994) A recursive approach to local mesh refinement in two and three dimensions. *J. Comput. Appl. Math.*, **55**, 275–288.
- KREUZER, C., MÖLLER, C. A., SCHMIDT, A. & SIEBERT, K. G. (2012) Design and convergence analysis for an adaptive discretization of the heat equation. *IMA J. Numer. Anal.*, **32**, 1375–1403.
- KREUZER, C. & SIEBERT, K. G. (2011) Decay rates of adaptive finite elements with Dörfler marking. *Numer. Math.*, **117**, 679–716.
- LAKKIS, O. & MAKRIDAKIS, C. (2006) Elliptic reconstruction and a posteriori error estimates for fully discrete linear parabolic problems. *Math. Comp.*, **75**, 1627–1658.
- MAKRIDAKIS, C. & NOCHETTO, R. H. (2003) Elliptic reconstruction and a posteriori error estimates for parabolic problems. *SIAM J. Numer. Anal.*, **41**, 1585–1594.
- MAKRIDAKIS, C. & NOCHETTO, R. H. (2006) A posteriori error analysis for higher order dissipative methods for evolution problems. *Numer. Math.*, **104**, 489–514.
- MAUBACH, J. M. (1995) Local bisection refinement for  $n$ -simplicial grids generated by reflection. *SIAM J. Sci. Comput.*, **16**, 210–227.
- MEKCHAY, K. & NOCHETTO, R. H. (2005) Convergence of adaptive finite element methods for general second order linear elliptic PDEs. *SIAM J. Numer. Anal.*, **43**, 1803–1827.
- MORIN, P., NOCHETTO, R. H. & SIEBERT, K. G. (2000) Data oscillation and convergence of adaptive FEM. *SIAM J. Numer. Anal.*, **38**, 466–488.
- MORIN, P., NOCHETTO, R. H. & SIEBERT, K. G. (2002) Convergence of adaptive finite element methods. *SIAM Rev.*, **44**, 631–658.
- MORIN, P., SIEBERT, K. G. & VEESER, A. (2008) A basic convergence result for conforming adaptive finite elements. *Math. Models Methods Appl. Sci.*, **18**, 707–737.
- NOCHETTO, R. H., SIEBERT, K. G. & VEESER, A. (2009) Theory of adaptive finite element methods: an introduction. *Multiscale, Nonlinear and Adaptive Approximation*. Berlin: Springer, pp. 409–542.
- PICASSO, M. (1998) Adaptive finite elements for a linear parabolic problem. *Comput. Methods Appl. Mech. Eng.*, **167**, 223–237.
- SCHMIDT, A. & SIEBERT, K. G. (2005) *Design of Adaptive Finite Element Software: The Finite Element Toolbox ALBERTA*. Lecture Notes in Computational Science and Engineering, vol. 42. Berlin: Springer.
- SCHWAB, C. & STEVENSON, R. (2009) Space-time adaptive wavelet methods for parabolic evolution problems. *Math. Comp.*, **78**, 1293–1318.
- SIEBERT, K. G. (2011) A convergence proof for adaptive finite elements without lower bound. *IMA J. Numer. Anal.*, **31**, 947–970.
- SMEARS, I. (2017) Robust and efficient preconditioners for the discontinuous Galerkin time-stepping method. *IMA J. Numer. Anal.*, **37**, 1961–1985.
- STEVENSON, R. (2007) Optimality of a standard adaptive finite element method. *Found. Comput. Math.*, **7**, 245–269.
- TANTARDINI, F. & VEESER, A. (2016) The  $L_2$ -projection and quasi-optimality of Galerkin methods for parabolic equations. *SIAM J. Numer. Anal.*, **54**, 317–340.
- THOMÉE, V. (2006) *Galerkin Finite Element Methods for Parabolic Problems*, 2nd edn. Springer Series in Computational Mathematics, vol. 25. Berlin: Springer.
- TRAXLER, C. T. (1997) An algorithm for adaptive mesh refinement in  $n$  dimensions. *Computing*, **59**, 115–137.
- VERFÜRTH, R. (2003) A posteriori error estimates for finite element discretizations of the heat equation. *Calcolo*, **40**, 195–212.
- VERFÜRTH, R. (2013) *A Posteriori Error Estimation Techniques for Finite Element Methods*. Numerical Mathematics and Scientific Computation. Oxford: Oxford University Press.