

## A FRAMEWORK FOR ROBUST A POSTERIORI ERROR CONTROL IN UNSTEADY NONLINEAR ADVECTION-DIFFUSION PROBLEMS\*

VÍT DOLEJŠÍ<sup>†</sup>, ALEXANDRE ERN<sup>‡</sup>, AND MARTIN VOHRALÍK<sup>§</sup>

**Abstract.** We derive a framework for a posteriori error estimates in unsteady, nonlinear, possibly degenerate, advection-diffusion problems. Our estimators are based on a space-time equilibrated flux reconstruction and are locally computable. They are derived for the error measured in a space-time mesh-dependent dual norm stemming from the problem and meshes at hand augmented by a jump seminorm measuring possible nonconformities in space. Owing to this choice, a guaranteed and globally efficient upper bound is achieved, as well as robustness with respect to nonlinearities, advection dominance, domain size, final time, and absolute and relative size of space and time steps. Local-in-time and in-space efficiency is also shown for a localized upper bound of the error measure. In order to apply the framework to a given numerical method, two simple conditions, local space-time mass conservation and an approximation property of the reconstructed fluxes, need to be verified. We show how to do this for the interior-penalty discontinuous Galerkin method in space and the Crank–Nicolson scheme in time. Numerical experiments illustrate the theory.

**Key words.** unsteady nonlinear advection-diffusion problem, a posteriori estimate, dual norm, flux reconstruction, flux equilibration, unified framework, robustness, discontinuous Galerkin method

**AMS subject classifications.** 35K65, 65M15, 65M60, 76S05, 76R99

**DOI.** 10.1137/110859282

**1. Introduction.** We consider the unsteady nonlinear problem

$$(1.1a) \quad \partial_t u - \nabla \cdot \boldsymbol{\sigma}(u, \nabla u) = f \quad \text{in } Q := \Omega \times (0, t_F),$$

$$(1.1b) \quad u = 0 \quad \text{on } \partial\Omega \times (0, t_F),$$

$$(1.1c) \quad u(\cdot, 0) = u_0 \quad \text{in } \Omega,$$

with  $\Omega \subset \mathbb{R}^d$ ,  $d \geq 2$ , a polygonal (polyhedral) domain,  $t_F > 0$  the final time,  $f$  the source term, and  $u_0$  the initial datum. The function  $\boldsymbol{\sigma}(u, \nabla u)$  takes the form

$$(1.2) \quad \boldsymbol{\sigma}(u, \nabla u) := \underline{\mathbf{K}}(u)\nabla u - \boldsymbol{\phi}(u),$$

where  $\underline{\mathbf{K}}(\cdot)$  is a nonlinear, possibly degenerate, tensor-valued function associated with diffusive transport and  $\boldsymbol{\phi}(\cdot)$  a nonlinear vector-valued function associated with advective transport. Precise assumptions on  $\underline{\mathbf{K}}(\cdot)$  and  $\boldsymbol{\phi}(\cdot)$  are specified in section 2.1. In what follows,  $u$  is termed the *potential* and  $-\boldsymbol{\sigma}(u, \nabla u)$  the advective-diffusive *flux*. We assume that the problem (1.1a)–(1.1c) admits a unique weak solution  $u$  and that a fully

---

\*Received by the editors December 19, 2011; accepted for publication (in revised form) December 5, 2012; published electronically March 5, 2013. This work was supported by the Groupement MoMaS (PACEN/CNRS, ANDRA, BRGM, CEA, EdF, IRSN) and by the ERT project “Enhanced Oil Recovery and Geological Sequestration of CO<sub>2</sub>: Mesh Adaptivity, a Posteriori Error Control, and Other Advanced Techniques” (LJLL/IFPEN).

<http://www.siam.org/journals/sinum/51-2/85928.html>

<sup>†</sup>Charles University in Prague, Faculty of Mathematics and Physics, 186 75 Praha 8, Czech Republic (dolejsi@karlin.mff.cuni.cz). The research of the work of the first author was also supported by grant 201/08/0012 of the Czech Science Foundation.

<sup>‡</sup>Université Paris-Est, CERMICS, Ecole des Ponts, 77455 Marne-la-Vallée, France (ern@cermics.enpc.fr).

<sup>§</sup>UPMC Univ. Paris 06, UMR 7598, Laboratoire Jacques-Louis Lions, 75005 Paris, France & CNRS, UMR 7598, Laboratoire Jacques-Louis Lions, 75005 Paris, France. Current address: INRIA Paris-Rocquencourt, B.P. 105, 78153 Le Chesnay, France (martin.vohralik@inria.fr).

discrete approximate potential, say,  $u_{h\tau}$ , is available. The space discretization scheme can be nonconforming, so that  $u_{h\tau}$  exhibits jumps across the spatial mesh interfaces.

Quite a broad literature has been devoted to the a posteriori error analysis for problems of the form (1.1a)–(1.1c). One of the main issues is to simultaneously prove global upper and (local) lower bounds for the error (reliability and (local) efficiency). For the linear heat equation, reliability is achieved in the energy norm by Picasso [26] and Repin [28] and in higher-order norms by Makridakis and Nochetto [23]. In Verfürth [32] and Bergam, Bernardi, and Mghazli [5], both reliability and efficiency are proved by augmenting the energy norm by a dual norm of the time derivative. The lower bound is then local in time but global in space. A further analysis unifying various space discretization schemes has been recently given in [17]. For nonlinear parabolic problems, reliability and local and global efficiency have been derived by Verfürth [30, 31] under a restriction on the relative size of space and time steps. This restriction has been lifted by Verfürth in [33] for convex two-dimensional spatial domains, thereby achieving global efficiency; the price to pay is a solution of a linear diffusion problem by the finite element method on each time step. Robustness with respect to advection dominance or nonlinearities is not addressed in [33]. For unsteady linear advection-diffusion problems, Verfürth [34] has proved robustness with respect to advection dominance while augmenting the energy norm by a dual norm of the material derivative. A solution of a finite element reaction-diffusion problem on each time step is, again, necessary. All the above works concern the nondegenerate case; still less results are available for degenerate parabolic problems. We cite in particular Nochetto, Schmidt, and Verdi [24], where only the error upper bound is derived.

The purpose of the present work is to derive *guaranteed*, *(locally) efficient*, and *robust* a posteriori error estimates for the problem (1.1a)–(1.1c). Here, guaranteed means that the estimates represent an explicitly computable upper bound on the error and robustness means that the efficiency is independent of the main model and discretization parameters, that is, nonlinearities, advection dominance, domain size, final time, and absolute and relative size of space and time steps. While these results are of independent theoretical interest, their practical motivation is on the one hand to certify the simulation error and on the other hand to provide criteria for space-time adaptive mesh refinement in view of efficient numerical computations.

Our key idea is the introduction of a *space-time mesh-dependent dual norm* to measure the error; see (2.7). This norm includes the full space-time nonlinear advection-diffusion operator, which stands in contrast to previous work where only parts of the differential operator at hand are included (i.e., the time derivative or the material derivative as in [32, 34]). Such an approach has been used recently by Chaillou and Suri [8] and in [12] in the context of steady nonlinear diffusion problems. Moreover, evaluating the dual norm with respect to a specific mesh-dependent norm for test functions with bounded time and space derivatives in  $L^2(Q)$  (see (2.5a)–(2.5b)) also allows one to *simplify* substantially the *proof* of the *error lower bound*, whereby space-time bubble functions can be considered instead of the more usual space bubble functions at fixed times. This point also presents the practically crucial advantage that an error lower bound can be achieved using *locally computable estimators*, in contrast to [33, 34], where the solution of a global diffusion/reaction-diffusion problem on each time step is necessary to evaluate the estimators. The error measure (2.7) may seem rather weak at a first glance. However, this measure admits an easily and locally computable upper bound which consists of weighted  $L^2$ -norms of the potential error and of the error in the nonlinear advection-diffusion flux; see (2.10)–(2.11).

Our efficiency results carry over to this norm and, moreover, can then be *localized in space and in time*. The numerical experiments of section 8 actually show that our estimators generally provide reliable estimates for this weighted  $L^2$ -norm.

Our results are derived in a *unified framework* where the actual numerical scheme used to obtain  $u_{h\tau}$  needs not be specified. The error upper bound hinges on an advection-diffusion *flux reconstruction* and its local *space-time equilibration* (see Assumption 3.1), while the error lower bound requires a local *approximation property* on this flux; see Assumption 4.1. Applying the present framework to a given numerical scheme simply boils down to verifying these two assumptions. The reconstruction of the flux depends on the space discretization scheme, as discussed in [17] for the linear heat equation. Our approach is thus in line with the developments relying for linear model problems on the Prager and Synge equality [27] and pursued later by, e.g., Ladevèze [21], Luce and Wohlmuth [22], Braess, Pillwein, and Schöberl [6], or Ainsworth [2].

This paper is structured as follows. We present the continuous and discrete settings and define our error measure  $\mathcal{J}_u(u_{h\tau})$  in section 2. This is a sum of the above-discussed space-time mesh-dependent dual norm  $\mathcal{J}_{u,\text{FR}}(u_{h\tau})$  (see (2.7)) and of a weighted jump term  $\mathcal{J}_{u,\text{NC}}(u_{h\tau})$  which measures possible nonconformities of  $u_{h\tau}$  in space (see (2.12)). We state our a posteriori error estimate, Theorem 3.3, in section 3, yielding the error upper bound  $\mathcal{J}_u(u_{h\tau}) \leq \eta_{\text{FR}} + \eta_{\text{NC}} + \eta_{\text{IC}}$ . Theorem 4.4 of section 4 then establishes the error lower bound  $\eta_{\text{FR}} + \eta_{\text{NC}} \lesssim \mathcal{J}_u(u_{h\tau})$  under the assumption that the quadrature error caused by the nonlinearity of the flux  $\sigma$  is small enough. Here,  $\lesssim$  means up to a generic constant which is independent of the main model and discretization parameters as specified above. (The possible dependencies of the constant are stated in section 4.) Moreover, Theorem 4.2 provides a space-time localized version in terms of the computable error upper bound (2.10)–(2.11) on the error measure  $\mathcal{J}_{u,\text{FR}}(u_{h\tau})$ . We devote sections 5 and 6 to the proofs of these results. To illustrate the developed abstract framework, we apply it in section 7 to the interior-penalty discontinuous Galerkin (IPDG) method in space and the Crank–Nicolson scheme in time. Numerical experiments, including nonlinear degenerate advection-diffusion problems, are presented in section 8.

**2. The setting.** This section briefly describes the continuous and discrete settings and discusses the error measure.

**2.1. Continuous setting and weak solution.** We consider the function spaces

$$(2.1a) \quad X := L^2(0, t_F; H_0^1(\Omega)),$$

$$(2.1b) \quad Y := \{\varphi \in L^2(0, t_F; H_0^1(\Omega)); \partial_t \varphi \in L^2(Q); \varphi(\cdot, t_F) = 0\},$$

recalling that  $\varphi \in L^2(0, t_F; H_0^1(\Omega))$  and  $\partial_t \varphi \in L^2(0, t_F; H^{-1}(\Omega))$  classically imply  $\varphi \in C^0([0, t_F]; L^2(\Omega))$ . The weak solution  $u$  of (1.1a)–(1.1c) is sought in the space  $X$ , whereas the space  $Y$  is used as test space. Specifically, we assume that there exists a unique weak solution  $u \in X$  such that

$$(2.2) \quad \int_0^{t_F} \{(f, \varphi) + (u, \partial_t \varphi) - (\sigma(u, \nabla u), \nabla \varphi)\}(t) dt + (u_0, \varphi(\cdot, 0)) = 0 \quad \forall \varphi \in Y,$$

where  $(\cdot, \cdot)$  denotes the inner product in  $L^2(\Omega)$  or  $[L^2(\Omega)]^d$ . In order to ensure that all the terms in (2.2) are well-defined, we assume  $f \in L^2(Q)$ ,  $u_0 \in L^2(\Omega)$ , and  $\sigma(u, \nabla u) \in [L^2(Q)]^d$  which is satisfied, for  $u \in X \cap L^\infty(Q)$ , e.g., if  $\mathbf{K} \in L_{\text{loc}}^\infty(\mathbb{R}; \mathbb{R}^{d \times d})$

and  $\phi \in C^1(\mathbb{R}; \mathbb{R}^d)$ . In deriving our a posteriori error estimators in section 3, we exploit the fact that (2.2) consists, except for the contribution of the initial condition, of space-time inner products in  $L^2(Q)$ .

The present framework also covers some particular cases of degenerate parabolic equations of the form

$$(2.3) \quad \partial_t b(v) - \nabla \cdot (\underline{\mathbf{K}} \nabla v - \tilde{\phi}(v)) = f,$$

where  $b(\cdot)$  is an increasing function with locally Hölder regularity such that  $b(0) = 0$ . Problem (2.3) includes slow-diffusion-type problems, e.g., the porous medium equation for which  $b(v) = v^{1/\kappa}$ ,  $1 < \kappa < +\infty$  (so that  $b'(0) = +\infty$ ), and fast-diffusion-type problems of elliptic-parabolic form, e.g., the Richards equation for which typically  $b'(0) = 0$ . Existence, uniqueness, and regularity results for problem (2.3) can be found in the work of Alt and Luckhaus [3] and Otto [25], leading to weak solutions  $v \in X \cap L^\infty(Q)$ . Since the function  $b$  is increasing, (2.3) can be recast into the form (1.1a) by setting  $u := b(v)$ , yielding  $\underline{\mathbf{K}}(u) := \underline{\mathbf{K}}(b^{-1})'(u)$  and  $\phi(u) := \tilde{\phi}(b^{-1}(u))$ . Therefore, the present analysis can be applied to (2.3) under the assumption  $b(v) \in X$ . This assumption holds in the fast-diffusion regime but not necessarily in the slow-diffusion regime, where it is possible that  $v \in X$  but  $b(v) \notin X$ . The Stefan problem, governed by  $\partial_t u - \Delta \beta(u) = 0$ , where  $\beta(\cdot)$  is a nondecreasing Lipschitz function, is not covered by the present assumptions either, since, in this case, the weak solution  $u$  can exhibit jumps. Finally, we observe that we do not include zero-order terms in (1.1a).

*Remark 2.1* (physical units). In advection-diffusion problems, the physical unit of the components of the diffusion tensor  $\underline{\mathbf{K}}(\cdot)$  is  $L^2 T^{-1}$  and that of the components of the transport velocity  $\phi'(\cdot)$  is  $LT^{-1}$ . Here,  $L$  stands for length and  $T$  for time.

**2.2. Discrete setting and approximate solution.** We consider an increasing sequence of discrete times  $\{t^n\}_{0 \leq n \leq N}$  such that  $t^0 = 0$  and  $t^N = t_F$ . We set  $I_n := (t^{n-1}, t^n]$  and  $\tau^n := t^n - t^{n-1}$  for all  $1 \leq n \leq N$ . We consider a time sequence of matching simplicial meshes  $\{\mathcal{T}^n\}_{0 \leq n \leq N}$  of the spatial domain  $\Omega$ . The mesh  $\mathcal{T}^n$  is used to approximate the solution at the discrete time  $t^n$ . We assume that  $\mathcal{T}^n$  is obtained from  $\mathcal{T}^{n-1}$  by refining some elements and coarsening some others. For all  $1 \leq n \leq N$ , we denote by  $\overline{\mathcal{T}}^{n-1,n}$  the coarsest common refinement of  $\mathcal{T}^{n-1}$  and  $\mathcal{T}^n$  and by  $\underline{\mathcal{T}}^{n-1,n}$  the finest common coarsening of  $\mathcal{T}^{n-1}$  and  $\mathcal{T}^n$ . Obviously, if the meshes are kept fixed,  $\overline{\mathcal{T}}^{n-1,n} = \underline{\mathcal{T}}^{n-1,n} = \mathcal{T}^n = \mathcal{T}^0$ . In the general case, the mesh  $\overline{\mathcal{T}}^{n-1,n}$  is needed to handle nonconforming functions in space that are continuous in time, and the mesh  $\underline{\mathcal{T}}^{n-1,n}$  is used to facilitate the flux equilibration.

At this stage, we do not need to specify any numerical scheme with which to obtain the approximate solution  $u_{h\tau}$ . We simply assume that  $u_{h\tau}$  is in the space

$$X_h := \{\varphi \in L^2(0, t_F; H^1(\mathcal{T})); \partial_t \varphi \in L^2(Q)\}.$$

Here,  $L^2(0, t_F; H^1(\mathcal{T}))$  is spanned by those functions  $\varphi \in L^2(Q)$  that, for all  $1 \leq n \leq N$  and all  $T \in \overline{\mathcal{T}}^{n-1,n}$ , satisfy  $\varphi|_{T \times I_n} \in L^2(I_n; H^1(T))$ . Functions in  $X_h$  can exhibit jumps across the interfaces of  $\overline{\mathcal{T}}^{n-1,n}$ , but, owing to the assumption  $\partial_t \varphi \in L^2(Q)$ , they are continuous with respect to time. This latter assumption is needed to express the residual in terms of space-time inner products in  $L^2(Q)$ .

Let  $1 \leq n \leq N$  and let  $T \in \mathcal{T}^n$  or  $T \in \overline{\mathcal{T}}^{n-1,n}$  or  $T \in \underline{\mathcal{T}}^{n-1,n}$ . We use the notation  $h_T$  for the diameter of  $T$  and  $\|\cdot\|_T$  for the norm in  $L^2(T)$ . Similarly, the norm in  $L^2(T \times I_n)$  is denoted by  $\|\cdot\|_{T \times I_n}$ . The corresponding inner products are

denoted by  $(\cdot, \cdot)_T$  and  $(\cdot, \cdot)_{T \times I_n}$ , respectively. We collect in  $\mathcal{F}_T$  all the faces of  $T$  and in  $\mathcal{F}_T^{\text{int}}$  those that are subsets of  $\Omega$ . The set of all faces of the mesh  $\mathcal{T}^n$  is denoted by  $\mathcal{F}^n$ , a generic mesh face is denoted by  $F$ , and we use similar notation for norms and inner products as above. For an interface  $F$ ,  $\mathbf{n}_F$  denotes its unit normal oriented in the sense the jump is evaluated, while for a mesh element  $T$ ,  $\mathbf{n}_T$  denotes its outward unit normal. Similarly,  $\mathcal{T}_T$  stands for all mesh elements sharing a face with the element  $T$ , whereas  $\mathcal{T}_F$  denotes the mesh elements sharing the face  $F$ . Finally, we denote by  $\nabla$  the broken space gradient.

The following inequality is instrumental in deriving our error upper bound.

LEMMA 2.2 (scaled space-time Poincaré inequality). *Let  $1 \leq n \leq N$  and let  $T \in \mathcal{T}^{n-1,n}$ . Let  $\Pi_0$  denote the  $L^2(Q)$ -orthogonal projection onto constants in each space-time element  $T \times I_n$ . Set  $C_P := 1/\pi$ . Then, for all  $\varphi \in H^1(T \times I_n)$ ,*

$$(2.4) \quad \|\varphi - \Pi_0\varphi\|_{T \times I_n} \leq C_P (h_T^2 \|\nabla\varphi\|_{T \times I_n}^2 + (\tau^n)^2 \|\partial_t\varphi\|_{T \times I_n}^2)^{\frac{1}{2}}.$$

*Proof.* Let  $\varphi \in H^1(T \times I_n)$  and, for all  $t \in I_n$ , set  $\tilde{\varphi}(t) := |T|^{-1} \int_T \varphi(\mathbf{x}, t) \, d\mathbf{x}$ . Observing that  $(\varphi - \tilde{\varphi})$  and  $(\tilde{\varphi} - \Pi_0\varphi)$  are  $L^2(T \times I_n)$ -orthogonal, we infer  $\|\varphi - \Pi_0\varphi\|_{T \times I_n}^2 = \|\varphi - \tilde{\varphi}\|_{T \times I_n}^2 + \|\tilde{\varphi} - \Pi_0\varphi\|_{T \times I_n}^2$ . For the first term on the right-hand side, the usual Poincaré inequality on  $T$  (which is convex) yields

$$\|\varphi - \tilde{\varphi}\|_{T \times I_n}^2 \leq \int_{I_n} \left( C_P^2 h_T^2 \int_T |\nabla\varphi|^2(\mathbf{x}, t) \, d\mathbf{x} \right) dt = C_P^2 h_T^2 \|\nabla\varphi\|_{T \times I_n}^2.$$

For the second term, observing that  $\Pi_0\varphi = (\tau^n)^{-1} \int_{I_n} \tilde{\varphi}(t) \, dt$ , the one-dimensional Poincaré inequality on  $I_n$  and the Cauchy–Schwarz inequality yield

$$\|\tilde{\varphi} - \Pi_0\varphi\|_{T \times I_n}^2 \leq |T|^{-1} C_P^2 (\tau^n)^2 \int_{I_n} \left( \int_T \partial_t\varphi(\mathbf{x}, t) \, d\mathbf{x} \right)^2 dt \leq C_P^2 (\tau^n)^2 \|\partial_t\varphi\|_{T \times I_n}^2.$$

Collecting the above bounds yields (2.4).  $\square$

**2.3. Error measure.** The error measure described in this section combines a space-time mesh-dependent dual norm plus a nonconformity term.

**2.3.1. Space-time mesh-dependent dual norm.** Recalling definition (2.1b) of the space  $Y$ , we equip it with the norm

(2.5a)

$$\|\varphi\|_{Y, T \times I_n}^2 := C_{T,n} (h_T^2 \|\nabla\varphi\|_{T \times I_n}^2 + (\tau^n)^2 \|\partial_t\varphi\|_{T \times I_n}^2) \quad \forall 1 \leq n \leq N, T \in \mathcal{T}^{n-1,n},$$

$$(2.5b) \quad \|\varphi\|_Y^2 := \sum_{n=1}^N \sum_{T \in \mathcal{T}^{n-1,n}} \|\varphi\|_{Y, T \times I_n}^2.$$

$\|\cdot\|_Y$  is indeed a norm since functions in  $Y$  vanish on  $\Omega \times \{t_F\}$  and on  $\partial\Omega \times (0, t_F)$ . We observe that this norm depends on the space-time meshes. The positive quantities  $C_{T,n}$  are user-dependent weights, typically with physical unit  $T^{-1}$ , so that the error measure has the same unit as the classical energy norm. A possible example is

$$(2.6) \quad C_{T,n} := ((\tau^n)^{-2} t_F + h_T^{-2} C_{\phi, T, n} + h_T^{-2} C_{\mathbf{k}, T, n})$$

with a weight  $C_{\phi, T, n}$  proportional to a length scale times an advective velocity scale and a weight  $C_{\mathbf{k}, T, n}$  proportional to a diffusion scale. Since our estimators derived

in section 3 are also scaled by these weights, our robustness results are *not influenced* by the value assigned to these weights; see Remark 4.5.

The first building block of our error measure is the quantity

$$(2.7) \quad \mathcal{J}_{u,FR}(u_{h\tau}) := \sup_{\varphi \in Y, \|\varphi\|_Y=1} \int_0^{t_F} \{ (u_{h\tau} - u, \partial_t \varphi) + (\boldsymbol{\sigma}(u, \nabla u) - \boldsymbol{\sigma}(u_{h\tau}, \nabla u_{h\tau}), \nabla \varphi) \} (t) dt.$$

Let  $v \in L^2(0, t_F; H^1(\mathcal{T}))$ . Define the residual  $R(v) \in Y'$  such that, for all  $\varphi \in Y$ ,

$$(2.8) \quad \langle R(v), \varphi \rangle_{Y',Y} := \int_0^{t_F} \{ (f, \varphi) + (v, \partial_t \varphi) - (\boldsymbol{\sigma}(v, \nabla v), \nabla \varphi) \} (t) dt + (u_0, \varphi(\cdot, 0)).$$

Then we infer, owing to (2.2), that

$$(2.9) \quad \mathcal{J}_{u,FR}(u_{h\tau}) = \sup_{\varphi \in Y, \|\varphi\|_Y=1} \langle R(u_{h\tau}), \varphi \rangle_{Y',Y},$$

showing that  $\mathcal{J}_{u,FR}(u_{h\tau})$  is a *dual norm of the residual of the approximate solution*.

The error measure  $\mathcal{J}_{u,FR}(u_{h\tau})$  cannot be computed easily in practice (in the test cases where the exact solution  $u$  is available). Indeed, its evaluation requires solving the following (infinite-dimensional space-time) dual problem: Find  $\psi \in Y$  such that  $(\psi, \varphi)_Y = \langle R(u_{h\tau}), \varphi \rangle_{Y',Y}$  for all  $\varphi \in Y$ , where  $(\cdot, \cdot)_Y$  denotes the inner product corresponding to the  $\|\cdot\|_Y$ -norm. Then, it is immediate that  $\mathcal{J}_{u,FR}(u_{h\tau}) = \|\psi\|_Y$ . However, a computable upper bound on  $\mathcal{J}_{u,FR}(u_{h\tau})$  can be readily derived using the Cauchy–Schwarz inequality, leading to

$$(2.10) \quad \mathcal{J}_{u,FR}(u_{h\tau}) \leq \sup_{\varphi \in Y, \|\varphi\|_Y=1} \sum_{n=1}^N \sum_{T \in \mathcal{T}^{n-1,n}} e_{FR,T}^n \|\varphi\|_{Y,T \times I_n} \leq \left\{ \sum_{n=1}^N \sum_{T \in \mathcal{T}^{n-1,n}} (e_{FR,T}^n)^2 \right\}^{\frac{1}{2}}$$

that we denote by  $e_{FR}$  with, for all  $1 \leq n \leq N$  and all  $T \in \mathcal{T}^{n-1,n}$ ,

$$(2.11) \quad e_{FR,T}^n := C_{T,n}^{-\frac{1}{2}} \{ (\tau^n)^{-2} \|u_{h\tau} - u\|_{T \times I_n}^2 + h_T^{-2} \|\boldsymbol{\sigma}(u, \nabla u) - \boldsymbol{\sigma}(u_{h\tau}, \nabla u_{h\tau})\|_{T \times I_n}^2 \}^{\frac{1}{2}}.$$

*Remark 2.3* (dual norm of the residual and energy-type norms). Under appropriate assumptions, a functional framework can be introduced for the nonlinear differential operator in (1.1a), and the error between  $u$  and  $u_{h\tau}$  can be measured in the corresponding energy-type norms. For conforming approximations ( $u_{h\tau} \in X$ ), following Verfürth (see, e.g., [30, Proposition 2.1]), the energy error can be bounded from above and from below by a dual norm of the residual. Such equivalence results hinge on suitable a priori bounds of the linearized differential operator, where it is in particular difficult to trace the influence of the size of the nonlinearities or of the advection dominance. Here we measure the error directly by (2.7) and avoid energy-type norms and the question of their equivalence with a dual norm of the residual.

**2.3.2. Nonconformity.** As we allow for nonconformities ( $X_h$  is not a subspace of  $X$ ), we need to introduce a second building block of our error measure,

$$(2.12) \quad \mathcal{J}_{u,NC}(u_{h\tau}) := \left\{ \sum_{n=1}^N \sum_{T \in \overline{\mathcal{T}}^{n-1,n}} \sum_{F \in \mathcal{F}_T} C_{T,n}^{-1} h_T^{-2} C_{\mathbf{K},\phi,T,F,n} \|[u - u_{h\tau}]\|_{F \times I_n}^2 \right\}^{\frac{1}{2}},$$

where  $[\cdot]$  means the jump across interfaces and the actual value at boundary faces and, for  $T \in \overline{\mathcal{T}}^{n-1,n}$ ,  $C_{T,n} := C_{T',n}$ , where  $T \subset T' \in \underline{\mathcal{T}}^{n-1,n}$ . The measure (2.12) is inspired by the penalty terms used in IPDG methods (see, e.g., the work of Arnold [4]). Owing to the flux approximation property in Assumption 4.1 below, the design of  $\mathcal{J}_{u,\text{NC}}(u_{h\tau})$  is linked to the way fluxes are equilibrated. Consequently, the positive weights  $C_{\underline{\mathbf{k}},\phi,T,F,n}$  (physical unit is  $L^3T^{-2}$ ) have to be chosen accordingly; we refer to (7.7) for an example. Moreover, the summation over the common refinement  $\overline{\mathcal{T}}^{n-1,n}$  is linked to continuous-in-time flux reconstruction in the Crank–Nicolson scheme, whereas  $\mathcal{T}^n$  and piecewise constant-in-time flux reconstruction appear naturally in the backward Euler case. Furthermore, a straightforward consequence of the fact that the weak solution  $u$  is in  $X$  is the following lemma.

LEMMA 2.4 (jumps of the weak solution). *For all  $1 \leq n \leq N$ , all  $T \in \overline{\mathcal{T}}^{n-1,n}$ , and all  $F \in \mathcal{F}_T$ ,  $[[u]] = 0$  in  $L^2(F \times I_n)$ .*

Owing to Lemma 2.4,  $[[u - u_{h\tau}]]$  can be replaced by  $[[u_{h\tau}]]$  in (2.12). Thus, in contrast to  $\mathcal{J}_{u,\text{FR}}(u_{h\tau})$ ,  $\mathcal{J}_{u,\text{NC}}(u_{h\tau})$  is easily computable. Finally, note that  $\mathcal{J}_{u,\text{NC}}(u_{h\tau}) = 0$  if and only if  $u_{h\tau} \in X$ , that is, if and only if  $u_{h\tau}$  is  $X$ -conforming.

**2.3.3. Error measure.** Our error measure is the sum of (2.7) and (2.12), i.e.,

$$(2.13) \quad \mathcal{J}_u(u_{h\tau}) := \mathcal{J}_{u,\text{FR}}(u_{h\tau}) + \mathcal{J}_{u,\text{NC}}(u_{h\tau}).$$

There holds  $\mathcal{J}_u(u_{h\tau}) = 0$  if and only if  $u_{h\tau} = u$ . Indeed, if  $u_{h\tau} = u$ ,  $\mathcal{J}_u(u_{h\tau})$  clearly equals 0. Conversely, if  $\mathcal{J}_u(u_{h\tau}) = 0$ , then  $\mathcal{J}_{u,\text{FR}}(u_{h\tau}) = \mathcal{J}_{u,\text{NC}}(u_{h\tau}) = 0$ . Thus,  $u_{h\tau}$  is in  $X$  and, hence, since the weak solution is uniquely characterized by the property  $\langle R(u), \varphi \rangle_{Y',Y} = 0$  for all  $\varphi \in Y$  (see (2.2) and (2.8)), we infer  $u_{h\tau} = u$ .

**3. A posteriori error estimate.** This section collects the main results of this paper concerning the error upper bound. The approximation error is measured by (2.13), and the estimators are defined using an equilibrated flux reconstruction.

**3.1. Equilibrated flux reconstruction.** In order to proceed as generally as possible, in particular without the definition of any numerical scheme for approximating (1.1a)–(1.1c), we make the following assumption. (Recall that  $\mathbf{H}(\text{div}, \Omega)$  is spanned by vector fields in  $[L^2(\Omega)]^d$  with weak divergence in  $L^2(\Omega)$ ; cf. Brezzi and Fortin [7].)

Assumption 3.1 (space-time equilibrated flux reconstruction). There exists a flux reconstruction  $\mathbf{t}_{h\tau} \in \mathbf{L}^2(0, t_F; \mathbf{H}(\text{div}, \Omega))$ , equilibrated in the sense that

$$(3.1) \quad (f - \partial_t u_{h\tau} - \nabla \cdot \mathbf{t}_{h\tau}, 1)_{T \times I_n} = 0 \quad \forall 1 \leq n \leq N, T \in \underline{\mathcal{T}}^{n-1,n}.$$

The function  $\mathbf{t}_{h\tau}$ , which represents the advective-diffusive flux, has the opposite sign with respect to the function  $\boldsymbol{\sigma}$  of (1.2), usually defined with the same sign as the gradient to yield monotonicity-type properties. Specific constructions of the flux  $\mathbf{t}_{h\tau}$  for various spatial discretizations are discussed in [17]. We present an example in the context of IPDG methods for problem (1.1a)–(1.1c) in section 7.

Remark 3.2 (local mass conservation). Equation (3.1) expresses local mass conservation over the space-time element  $T \times I_n$ . A similar assumption has been made in [17, equation (3.4)]. In [17], however, the local mass conservation had to be satisfied over a given mesh element  $T$  for all times  $t \in I_n$ . The present assumption, being more general, allows for more flexibility. In particular, it allows us to use the scaled Poincaré inequality (2.4) in the proof of the error upper bound. This local mass conservation is supposed in (3.1) on the elements of the common coarsening of  $\mathcal{T}^{n-1}$  and

$\mathcal{T}^n$  only, which in particular fits the Crank–Nicolson scheme of section 7. Local mass conservation on the elements of  $\mathcal{T}^n$  appears naturally in the backward Euler case.

**3.2. Guaranteed a posteriori error estimate.** We are now in a position to state our main result concerning the error upper bound. For all  $1 \leq n \leq N$  and all  $T \in \underline{\mathcal{T}}^{n-1,n}$ , we define the *residual*, *flux*, and *nonconformity estimators*, respectively,

$$(3.2) \quad \eta_{R,T}^n := C_{T,n}^{-\frac{1}{2}} C_P \|f - \partial_t u_{h\tau} - \nabla \cdot \mathbf{t}_{h\tau}\|_{T \times I_n},$$

$$(3.3) \quad \eta_{F,T}^n := C_{T,n}^{-\frac{1}{2}} h_T^{-1} \|\boldsymbol{\sigma}(u_{h\tau}, \nabla u_{h\tau}) + \mathbf{t}_{h\tau}\|_{T \times I_n},$$

$$(3.4) \quad \eta_{NC,T}^n := \left\{ \sum_{T' \in \overline{\mathcal{T}}^{n-1,n}, T' \subset T} \sum_{F \in \mathcal{F}_{T'}} C_{T',n}^{-1} h_{T'}^{-2} C_{\mathbf{K},\phi,T',F,n} \|[u_{h\tau}]\|_{F \times I_n}^2 \right\}^{\frac{1}{2}},$$

and set  $\eta_{FR,T}^n := \eta_{F,T}^n + \eta_{R,T}^n$ . These estimators are local-in-space and in-time. We define their global space-time versions as  $\eta_{\bullet} := \{\sum_{n=1}^N \sum_{T \in \underline{\mathcal{T}}^{n-1,n}} (\eta_{\bullet,T}^n)^2\}^{\frac{1}{2}}$  for  $\bullet \in \{F, R, FR, NC\}$ . Finally, we define the *initial condition estimator* as

$$(3.5) \quad \eta_{IC} := \left\{ \sum_{n=1}^N \sum_{T \in \underline{\mathcal{T}}^{n-1,n}} (\eta_{IC,T}^n)^2 \right\}^{\frac{1}{2}}, \quad \eta_{IC,T}^n := C_{T,n}^{-\frac{1}{2}} (\tau^n)^{-\frac{1}{2}} \|u_0 - u_{h\tau}(\cdot, 0)\|_T.$$

**THEOREM 3.3** (guaranteed a posteriori error estimate). *Let  $u \in X$  be the weak solution satisfying (2.2) and let  $u_{h\tau} \in X_h$  be arbitrary. Let Assumption 3.1 hold. Let  $\eta_{R,T}^n, \eta_{F,T}^n, \eta_{NC,T}^n$ , and  $\eta_{IC}$  be defined by (3.2)–(3.5). Then,*

$$(3.6) \quad \mathcal{J}_u(u_{h\tau}) \leq \eta_{FR} + \eta_{NC} + \eta_{IC}.$$

The proof is given in section 5. We now present several remarks.

*Remark 3.4* (interpretation of the error estimators). The flux estimator  $\eta_{F,T}^n$  is related to the violation of the *constitutive relation* (1.2) and the residual estimator  $\eta_{R,T}^n$  to the violation of the *equilibrium condition* (1.1a), while  $\eta_{NC,T}^n$  and  $\eta_{IC,T}^n$  are related to the violation of the *constraints* ( $u \in X$  and  $u(\cdot, 0) = u_0$ ) at the discrete level.

*Remark 3.5* (spatial and temporal errors). The decomposition of our estimators into spatial and temporal error estimators depends on the time-marching scheme. For the flux reconstruction  $\mathbf{t}_{h\tau}$  within the Crank–Nicolson setting of section 7,  $\eta_{F,T}^n$  and  $\eta_{NC,T}^n$  can be viewed as spatial error estimators and  $\eta_{R,T}^n$  as a temporal error estimator.

*Remark 3.6* (initial condition estimator). Whenever the weights  $C_{T,n}$  are chosen according to (2.6), there holds  $C_{T,n} \geq t_F (\tau^n)^{-2}$ , whence we readily infer that  $\eta_{IC} \leq \|u_{h\tau}(\cdot, 0) - u_0\|_{\Omega}$ , that is,  $\eta_{IC}$  is upper-bounded by the usual  $L^2(\Omega)$ -norm of the approximation error on the initial condition. Moreover, for very small time steps,  $\eta_{IC}$  tends to  $\|u_{h\tau}(\cdot, 0) - u_0\|_{\Omega}$ , while for very large time steps,  $\eta_{IC}$  becomes very small as the error from large time steps during the simulation itself dominates.

**4. Efficiency and robustness.** This section deals with the efficiency and robustness of our estimates. For simplicity, we assume that the source term  $f$  is a piecewise space-time polynomial on  $\overline{\mathcal{T}}^{n-1,n}$ ; otherwise, a classical data oscillation term has to be included in the error lower bound. We likewise assume that both  $u_{h\tau}$  and  $\mathbf{t}_{h\tau}$  are piecewise space-time polynomials on  $\overline{\mathcal{T}}^{n-1,n}$ . We observe that, because of the nonlinear functions  $\mathbf{K}(\cdot)$  and  $\phi(\cdot)$ , the flux  $\boldsymbol{\sigma}(u_{h\tau}, \nabla u_{h\tau})$  is not necessarily a piecewise polynomial, even if  $u_{h\tau}$  is. Let

$$(4.1) \quad \overline{\boldsymbol{\sigma}}(u_{h\tau}, \nabla u_{h\tau})|_T := \mathcal{P}_T(\boldsymbol{\sigma}(u_{h\tau}, \nabla u_{h\tau})|_T) \quad \forall 1 \leq n \leq N, T \in \underline{\mathcal{T}}^{n-1,n},$$



where  $\mathcal{P}_T$  is a projection-type operator mapping onto piecewise space-time polynomials; see (7.5)–(7.6) below for an example. We define the *quadrature error estimator*

$$(4.2) \quad \eta_{\text{qd},T}^n := C_{T,n}^{-\frac{1}{2}} h_T^{-1} \|\bar{\sigma}(u_{h\tau}, \nabla u_{h\tau}) - \sigma(u_{h\tau}, \nabla u_{h\tau})\|_{T \times I_n},$$

together with its global space-time version  $\eta_{\text{qd}} := \{\sum_{n=1}^N \sum_{T \in \mathcal{T}^{n-1,n}} (\eta_{\text{qd},T}^n)^2\}^{\frac{1}{2}}$ .

Henceforth,  $A \lesssim B$  means that there exists a constant  $C$  such that  $A \leq CB$  with  $C$  only depending on space dimension, maximal polynomial degree, shape-regularity of the meshes  $\mathcal{T}^n$  for all  $1 \leq n \leq N$ , the maximal ratio  $h_T/h_{T'}$  over all  $1 \leq n \leq N$ ,  $T \in \mathcal{T}^{n-1,n}$  (the common coarsening of  $\mathcal{T}^{n-1}$  and  $\mathcal{T}^n$ ) and  $T' \in \bar{\mathcal{T}}^{n-1,n}$  (the common refinement of  $\mathcal{T}^{n-1}$  and  $\mathcal{T}^n$ ),  $T'$  a subelement of  $T$ , and the maximal ratios  $C_{T,n}/C_{T',n}$  for all  $T \in \mathcal{T}^{n-1,n}$  and  $T' \in \mathcal{T}^{n-1,n}$  sharing a face and all  $1 \leq n \leq N$ . We thus in particular suppose that both refinement and coarsening are not too abrupt.

**4.1. Approximation property of the flux reconstruction.** For all  $1 \leq n \leq N$  and all  $T \in \bar{\mathcal{T}}^{n-1,n}$ , let  $\eta_{\text{clas},T}^n$  be given by

$$(4.3) \quad \begin{aligned} & h_T \|f - \partial_t u_{h\tau} + \nabla \cdot (\bar{\sigma}(u_{h\tau}, \nabla u_{h\tau}))\|_{T \times I_n} \\ & + \left\{ \sum_{F \in \mathcal{F}_T^{\text{int}}} h_F \|\bar{\sigma}(u_{h\tau}, \nabla u_{h\tau}) \cdot \mathbf{n}_F\|_{F \times I_n}^2 \right\}^{\frac{1}{2}} + \left\{ \sum_{F \in \mathcal{F}_T} C_{\mathbf{k},\phi,T,F,n} \|\llbracket u_{h\tau} \rrbracket\|_{F \times I_n}^2 \right\}^{\frac{1}{2}}. \end{aligned}$$

The quantity  $\eta_{\text{clas},T}^n$  can be viewed as a classical residual-based a posteriori error estimator. In order to carry the analysis without specifying a numerical scheme, we make the following assumption on the reconstructed flux.

*Assumption 4.1* (flux approximation property). There holds

$$\|\bar{\sigma}(u_{h\tau}, \nabla u_{h\tau}) + \mathbf{t}_{h\tau}\|_{T \times I_n}^2 \lesssim \sum_{T' \in \bar{\mathcal{T}}^{n-1,n}, T' \subset T} (\eta_{\text{clas},T'}^n)^2 \quad \forall 1 \leq n \leq N, T \in \mathcal{T}^{n-1,n}.$$

An example of how to achieve Assumption 4.1 in the context of IPDG methods in space and the Crank–Nicolson scheme in time is presented in section 7.

**4.2. Robust local-in-space and in-time efficiency.** To control quadrature errors, we introduce coefficients  $0 < \gamma_{\text{qd},T}^n$  and require

$$(4.4) \quad \max_{T' \in \mathcal{T}_T} \eta_{\text{qd},T'}^n \leq \gamma_{\text{qd},T}^n (\eta_{\text{FR},T}^n + \eta_{\text{NC},T}^n) \quad \forall 1 \leq n \leq N, T \in \mathcal{T}^{n-1,n},$$

recalling that  $\mathcal{T}_T$  collects the mesh elements of  $\mathcal{T}^{n-1,n}$  sharing a face with  $T$ .

**THEOREM 4.2** (robust local-in-space and in-time efficiency). *Let a time step  $1 \leq n \leq N$  and a mesh element  $T \in \mathcal{T}^{n-1,n}$  be fixed. Let (4.4), with  $\gamma_{\text{qd},T}^n$  small enough, hold. Let Assumption 4.1 hold. Recall the definition (2.11) of  $e_{\text{FR},T}^n$  and define  $\mathcal{J}_{u,\text{NC},T}(u_{h\tau})$  as  $\eta_{\text{NC},T}^n$  in (3.4) with  $\llbracket u_{h\tau} \rrbracket$  replaced by  $\llbracket u - u_{h\tau} \rrbracket$ . Then,*

$$(4.5) \quad \eta_{\text{FR},T}^n + \eta_{\text{NC},T}^n \lesssim \left\{ \sum_{T' \in \mathcal{T}_T} (e_{\text{FR},T'}^n)^2 \right\}^{\frac{1}{2}} + \mathcal{J}_{u,\text{NC},T}(u_{h\tau}).$$

*Remark 4.3* (comment on Theorem 4.2). Estimate (4.5) says that our estimators represent a robust *local* lower bound for the error measures  $e_{\text{FR},T}^n$ , whose Hilbertian sum provides an upper bound on the error measure  $\mathcal{J}_{u,\text{FR}}(u_{h\tau})$  (see (2.10)), augmented by the jump seminorm  $\mathcal{J}_{u,\text{NC}}(u_{h\tau})$ .

**4.3. Robust global efficiency.** As above, in order to control quadrature errors, we introduce a coefficient  $0 < \gamma_{\text{qd}}$  and require

$$(4.6) \quad \eta_{\text{qd}} \leq \gamma_{\text{qd}}(\eta_{\text{FR}} + \eta_{\text{NC}}).$$

Then we have the following full equivalence result between our estimators and  $\mathcal{J}_u(u_{h\tau})$ .

**THEOREM 4.4** (robust global efficiency). *Let (4.6), with  $\gamma_{\text{qd}}$  small enough, be satisfied. Let Assumption 4.1 hold. Then,*

$$(4.7) \quad \eta_{\text{FR}} + \eta_{\text{NC}} \lesssim \mathcal{J}_u(u_{h\tau}).$$

*Remark 4.5* (weights  $C_{T,n}$ ). We stress that our robustness results are not influenced by the value assigned to the weights  $C_{T,n}$  of (2.5a). Indeed, multiplying them by a positive factor  $\lambda$  scales both error measures  $\mathcal{J}_{u,\text{FR}}(u_{h\tau})$  and  $\mathcal{J}_{u,\text{NC}}(u_{h\tau})$  defined by (2.7) and (2.12) by the factor  $\lambda^{-\frac{1}{2}}$ , while the error estimators  $\eta_{\text{R},T}^n, \eta_{\text{F},T}^n, \eta_{\text{NC},T}^n$ , and  $\eta_{\text{IC},T}^n$  defined by (3.2)–(3.5) are scaled by the same factor. Hence, the ratio of error measure  $\mathcal{J}_u(u_{h\tau})$  to error indicators is *independent* of the scaling factor  $\lambda$ .

*Remark 4.6* (steady case). The present approach can be applied to the steady case leading to robustness with respect to advection dominance. Moreover, choosing  $C_T := (h_T^{-2}C_{\phi,T} + h_T^{-2}C_{\mathbf{K},T})$  with  $C_{\phi,T} := h_T \|\phi'(u_{h\tau})\|_{\infty,T}$ ,  $C_{\mathbf{K},T} := \|\mathbf{K}(u_{h\tau})\|_{\infty,T}$ , the steady version of the error measure  $e_{\text{FR}}$  of (2.10) is multiplied by the square root of the Péclet number with respect to the usual energy norm, while the estimators are divided by this factor with respect to the classical estimators (without cutoff factors); see, e.g., the work of Verfürth [34] and [14]. Thus, a similarity to the approach of Sangalli [29] can be observed.

**5. Proof of the error upper bound.** The goal of this section is to prove Theorem 3.3. The proof is decomposed into two steps.

**LEMMA 5.1** (bound on  $\mathcal{J}_{u,\text{FR}}(u_{h\tau})$ ). *There holds  $\mathcal{J}_{u,\text{FR}}(u_{h\tau}) \leq \eta_{\text{FR}} + \eta_{\text{IC}}$ .*

*Proof.* Let  $\varphi \in Y$  with  $\|\varphi\|_Y = 1$  be given. Using (2.8), subtracting the quantity  $\int_0^{t_{\text{F}}} (\mathbf{t}_{h\tau}, \nabla\varphi)(t) dt + \int_0^{t_{\text{F}}} (\nabla \cdot \mathbf{t}_{h\tau}, \varphi)(t) dt$ , which is equal to zero owing to the Green theorem since  $\mathbf{t}_{h\tau} \in \mathbf{L}^2(0, t_{\text{F}}; \mathbf{H}(\text{div}, \Omega))$  by Assumption 3.1, integrating by parts in time since  $\partial_t u_{h\tau} \in L^2(Q)$  by assumption, and using that  $\varphi(\cdot, 0) = -\int_0^{t_{\text{F}}} \partial_t \varphi(t) dt$  since  $\varphi(\cdot, t_{\text{F}}) = 0$  by assumption, we infer that  $\langle R(u_{h\tau}), \varphi \rangle_{Y',Y}$  is equal to

$$\int_0^{t_{\text{F}}} \{ (f - \partial_t u_{h\tau} - \nabla \cdot \mathbf{t}_{h\tau}, \varphi) + (u_{h\tau}(\cdot, 0) - u_0, \partial_t \varphi) - (\boldsymbol{\sigma}(u_{h\tau}, \nabla u_{h\tau}) + \mathbf{t}_{h\tau}, \nabla \varphi) \}(t) dt.$$

We now employ the second part of Assumption 3.1, namely, (3.1), yielding

$$\begin{aligned} \langle R(u_{h\tau}), \varphi \rangle_{Y',Y} &= \sum_{n=1}^N \sum_{T \in \mathcal{T}^{n-1,n}} \{ (f - \partial_t u_{h\tau} - \nabla \cdot \mathbf{t}_{h\tau}, \varphi - \Pi_0 \varphi)_{T \times I_n} \\ &\quad + (u_{h\tau}(\cdot, 0) - u_0, \partial_t \varphi)_{T \times I_n} - (\boldsymbol{\sigma}(u_{h\tau}, \nabla u_{h\tau}) + \mathbf{t}_{h\tau}, \nabla \varphi)_{T \times I_n} \} \\ &\leq \sum_{n=1}^N \sum_{T \in \mathcal{T}^{n-1,n}} (\eta_{\text{R},T}^n + ((\eta_{\text{F},T}^n)^2 + (\eta_{\text{IC},T}^n)^2)^{\frac{1}{2}}) \|\varphi\|_{Y,T \times I_n}, \end{aligned}$$

where we have used the Cauchy–Schwarz inequality, the scaled space-time Poincaré inequality (2.4), and definition (2.5a). The assertion then follows from  $((\eta_{\text{F},T}^n)^2 + (\eta_{\text{IC},T}^n)^2)^{\frac{1}{2}} \leq \eta_{\text{F},T}^n + \eta_{\text{IC},T}^n$ , the Cauchy–Schwarz inequality, definition (2.5b) of the  $\|\cdot\|_Y$ -norm, the fact that  $\|\varphi\|_Y = 1$ , and relation (2.9).  $\square$

**LEMMA 5.2** (bound on  $\mathcal{J}_{u,\text{NC}}(u_{h\tau})$ ). *There holds  $\mathcal{J}_{u,\text{NC}}(u_{h\tau}) = \eta_{\text{NC}}$ .*

*Proof.* The proof is immediate owing to Lemma 2.4 and (3.4).  $\square$

**6. Proof of the error lower bound.** We present here the proofs of Theorems 4.2 and 4.4. An important ingredient is the bubble function technique (see [32, 34]) that we extend here to space-time bubbles.

**6.1. Proof of Theorem 4.2.**

*Proof.* Let a time step  $1 \leq n \leq N$  and a mesh element  $T \in \overline{\mathcal{T}}^{n-1,n}$  be fixed. The first two steps of the proof are devoted to bounding the quantity  $\eta_{\text{clas},T}^n$  defined by (4.3), while staying on the common refinement  $\overline{\mathcal{T}}^{n-1,n}$ .

*Step 1: Bound on  $C_{T,n}^{-\frac{1}{2}} \|f - \partial_t u_{h\tau} + \nabla \cdot \overline{\boldsymbol{\sigma}}(u_{h\tau}, \nabla u_{h\tau})\|_{T \times I_n}$ .* Let us prove that

$$(6.1) \quad C_{T,n}^{-\frac{1}{2}} \|f - \partial_t u_{h\tau} + \nabla \cdot \overline{\boldsymbol{\sigma}}(u_{h\tau}, \nabla u_{h\tau})\|_{T \times I_n} \lesssim e_{\text{FR},T}^n + \eta_{\text{qd},T}^n$$

with  $e_{\text{FR},T}^n$  defined by (2.11). Set  $v_{T,n} := (f - \partial_t u_{h\tau} + \nabla \cdot \overline{\boldsymbol{\sigma}}(u_{h\tau}, \nabla u_{h\tau}))|_{T \times I_n}$ . Let  $\psi_{T,n}$  be the *space-time bubble function* on  $T \times I_n$  given by the product of the barycentric coordinates on  $T$  and of the barycentric coordinates on  $I_n$ . Note that both  $\psi_{T,n}$  and  $v_{T,n}$  are (space-time) polynomials since we are on a refinement of both  $\mathcal{T}^{n-1}$  and  $\mathcal{T}^n$ . Then, by norm equivalence in finite-dimensional spaces, there holds

$$(6.2) \quad (v_{T,n}, v_{T,n})_{T \times I_n} \lesssim (v_{T,n}, \psi_{T,n} v_{T,n})_{T \times I_n}.$$

Using an inverse inequality *separately* in space and in time, we obtain

$$(6.3a) \quad h_T \|\nabla(\psi_{T,n} v_{T,n})\|_{T \times I_n} \lesssim \|\psi_{T,n} v_{T,n}\|_{T \times I_n},$$

$$(6.3b) \quad \tau^n \|\partial_t(\psi_{T,n} v_{T,n})\|_{T \times I_n} \lesssim \|\psi_{T,n} v_{T,n}\|_{T \times I_n}.$$

The norm  $\|\cdot\|_{Y,T \times I_n}$  defined by (2.5a) was precisely designed in order to use these inequalities at the present stage. Using (6.3), (2.5a), and  $\|\psi_{T,n}\|_{\infty,T,n} \leq 1$ , we infer

$$(6.4) \quad C_{T,n}^{-1} \|\psi_{T,n} v_{T,n}\|_{Y,T \times I_n}^2 = (h_T^2 \|\nabla(\psi_{T,n} v_{T,n})\|_{T \times I_n}^2 + (\tau^n)^2 \|\partial_t(\psi_{T,n} v_{T,n})\|_{T \times I_n}^2) \lesssim \|\psi_{T,n} v_{T,n}\|_{T \times I_n}^2 \leq \|v_{T,n}\|_{T \times I_n}^2.$$

Thus, using the definition of  $v_{T,n}$ , (6.2), (2.2), and the Green theorem yields, with the notation  $\boldsymbol{\sigma}_d := \boldsymbol{\sigma}(u, \nabla u) - \overline{\boldsymbol{\sigma}}(u_{h\tau}, \nabla u_{h\tau})$ ,

$$(6.5) \quad C_{T,n}^{-1} \|v_{T,n}\|_{T \times I_n}^2 \lesssim C_{T,n}^{-1} (f - \partial_t u_{h\tau} + \nabla \cdot \overline{\boldsymbol{\sigma}}(u_{h\tau}, \nabla u_{h\tau}), \psi_{T,n} v_{T,n})_{T \times I_n} = C_{T,n}^{-1} (u_{h\tau} - u, \partial_t(\psi_{T,n} v_{T,n}))_{T \times I_n} + (\boldsymbol{\sigma}_d, \nabla(\psi_{T,n} v_{T,n}))_{T \times I_n}.$$

Thus, owing to (6.4),

$$C_{T,n}^{-\frac{1}{2}} \|v_{T,n}\|_{T \times I_n} \lesssim \frac{(u_{h\tau} - u, \partial_t(\psi_{T,n} v_{T,n}))_{T \times I_n} + (\boldsymbol{\sigma}_d, \nabla(\psi_{T,n} v_{T,n}))_{T \times I_n}}{\|\psi_{T,n} v_{T,n}\|_{Y,T \times I_n}},$$

whence (6.1) follows from the Cauchy–Schwarz inequality, the definitions (2.11) of  $e_{\text{FR},T}^n$  and (4.2) of  $\eta_{\text{qd},T}^n$ , and the triangle inequality.

*Step 2: Bound on  $C_{T,n}^{-\frac{1}{2}} h_T^{-\frac{1}{2}} \|\overline{\boldsymbol{\sigma}}(u_{h\tau}, \nabla u_{h\tau})\| \cdot \mathbf{n}_F \|_{F \times I_n}$  for all  $F \in \mathcal{F}_T^{\text{int}}$ .* Let  $F \in \mathcal{F}_T^{\text{int}}$  be fixed. Recalling that  $\mathcal{T}_F$  denotes the simplices in  $\overline{\mathcal{T}}^{n-1,n}$  sharing the face  $F$ , let us prove that

$$(6.6) \quad C_{T,n}^{-\frac{1}{2}} h_T^{-\frac{1}{2}} \|\overline{\boldsymbol{\sigma}}(u_{h\tau}, \nabla u_{h\tau})\| \cdot \mathbf{n}_F \|_{F \times I_n} \lesssim \sum_{T' \in \mathcal{T}_F} (e_{\text{FR},T'}^n + \eta_{\text{qd},T'}^n).$$

Set  $v_{F,n} := \llbracket \bar{\sigma}(u_{h\tau}, \nabla u_{h\tau}) \rrbracket \cdot \mathbf{n}_F|_{F \times I_n}$ . Let  $\psi_{F,n}$  be the space-time bubble function on  $F \times I_n$  given by the product of the barycentric coordinates with vertices in  $F$  and of the barycentric coordinates on  $I_n$ . Then, by norm equivalence in finite-dimensional spaces, there holds  $(v_{F,n}, v_{F,n})_{F \times I_n} \lesssim (v_{F,n}, \psi_{F,n} v_{F,n})_{F \times I_n}$ . Using the same notation for the extension of the function  $v_{F,n}$  onto  $\mathcal{T}_F$  by constant values, we also infer the estimate  $\|v_{F,n}\|_{\mathcal{T}_F \times I_n} \lesssim h_F^{\frac{1}{2}} \|v_{F,n}\|_{F \times I_n}$ . Using these inequalities and proceeding as in Step 1, we arrive at

$$C_{T,n}^{-\frac{1}{2}} h_T^{-\frac{1}{2}} \|v_{F,n}\|_{F \times I_n} \lesssim C_{T,n}^{-\frac{1}{2}} \|f - \partial_t u_{h\tau} + \nabla \cdot \bar{\sigma}(u_{h\tau}, \nabla u_{h\tau})\|_{\mathcal{T}_F \times I_n} + \frac{(u - u_{h\tau}, \partial_t(\psi_{F,n} v_{F,n}))_{\mathcal{T}_F \times I_n} - (\sigma_d, \nabla(\psi_{F,n} v_{F,n}))_{\mathcal{T}_F \times I_n}}{\|\psi_{F,n} v_{F,n}\|_{Y, \mathcal{T}_F, n}}.$$

Finally, the first term on the right-hand side is bounded using (6.1) for each  $T' \in \mathcal{T}_F$ , while proceeding as in Step 1 for the second term yields (6.6).

*Step 3: Conclusion.* Let now  $1 \leq n \leq N$  and an element  $T$  from the common coarsening  $\underline{\mathcal{T}}^{n-1,n}$  be fixed. Combining (6.1) and (6.6) yields

$$\sum_{T' \in \overline{\mathcal{T}}^{n-1,n}, T' \subset T} (C_{T',n}^{-\frac{1}{2}} h_{T'}^{-1} \eta_{\text{clas}, T'}^n)^2 \lesssim \sum_{T' \in \mathcal{T}_T} (e_{\text{FR}, T'}^n + \eta_{\text{qd}, T'}^n)^2 + (\eta_{\text{NC}, T}^n)^2.$$

Using the triangle inequality and Assumption 4.1, we infer  $(\eta_{\text{F}, T}^n)^2 \lesssim (\eta_{\text{qd}, T}^n)^2 + C_{T,n}^{-1} h_T^{-2} \sum_{T' \in \overline{\mathcal{T}}^{n-1,n}, T' \subset T} (\eta_{\text{clas}, T'}^n)^2$ . Similarly, using in addition the inverse space inequality leads to  $(\eta_{\text{R}, T}^n)^2 \lesssim C_{T,n}^{-1} h_T^{-2} \sum_{T' \in \overline{\mathcal{T}}^{n-1,n}, T' \subset T} (\eta_{\text{clas}, T'}^n)^2$ . Combining these inequalities we obtain

$$\eta_{\text{FR}, T}^n \lesssim \sum_{T' \in \mathcal{T}_T} (e_{\text{FR}, T'}^n + \eta_{\text{qd}, T'}^n) + \eta_{\text{NC}, T}^n.$$

Finally, invoking the requirement (4.4) on quadrature errors to discard the quantities  $\eta_{\text{qd}, T'}^n$  from the right-hand side completes the proof.  $\square$

**6.2. Proof of Theorem 4.4.**

*Proof.* We use the notation  $v_{T,n}$  and  $v_{F,n}$  from the previous proof. It is sufficient to show that

$$(6.7a) \quad \left\{ \sum_{n=1}^N \sum_{T \in \overline{\mathcal{T}}^{n-1,n}} C_{T,n}^{-1} \|v_{T,n}\|_{T \times I_n}^2 \right\}^{\frac{1}{2}} \lesssim \mathcal{J}_u(u_{h\tau}) + \eta_{\text{qd}},$$

$$(6.7b) \quad \left\{ \sum_{n=1}^N \sum_{T \in \overline{\mathcal{T}}^{n-1,n}} \sum_{F \in \mathcal{F}_T^{\text{int}}} C_{T,n}^{-1} h_T^{-1} \|v_{F,n}\|_{F \times I_n}^2 \right\}^{\frac{1}{2}} \lesssim \mathcal{J}_u(u_{h\tau}) + \eta_{\text{qd}},$$

since choosing  $\gamma_{\text{qd}}$  in (4.6) small enough then yields (4.7). We only show (6.7a); (6.7b) follows by similar arguments. Let  $1 \leq n \leq N$  and let  $T \in \overline{\mathcal{T}}^{n-1,n}$ . It follows from (6.5) and the Green theorem that

$$C_{T,n}^{-1} \|v_{T,n}\|_{T \times I_n}^2 \lesssim C_{T,n}^{-1} \{ (f - \partial_t u_{h\tau}, \psi_{T,n} v_{T,n})_{T \times I_n} - (\sigma(u_{h\tau}, \nabla u_{h\tau}), \nabla(\psi_{T,n} v_{T,n}))_{T \times I_n} + (\sigma(u_{h\tau}, \nabla u_{h\tau}) - \bar{\sigma}(u_{h\tau}, \nabla u_{h\tau}), \nabla(\psi_{T,n} v_{T,n}))_{T \times I_n} \}.$$

Set  $\lambda|_{T \times I_n} := C_{T,n}^{-1} \psi_{T,n} v_{T,n}$  and observe that  $\lambda \in Y$ . Recall the notation  $R(u_{h\tau})$  from (2.8),  $\mathcal{J}_{u,\text{FR}}(u_{h\tau})$  from (2.7), and  $\eta_{\text{qd},T}^n$  from (4.2). Summing the above inequality over all  $1 \leq n \leq N$  and all  $T \in \overline{\mathcal{T}}^{n-1,n}$  and using the Green theorem, the Cauchy–Schwarz inequality, (6.3a), and  $\|\psi_{T,n}\|_{\infty,T,n} \leq 1$  yields

$$\begin{aligned}
 & \sum_{n=1}^N \sum_{T \in \overline{\mathcal{T}}^{n-1,n}} C_{T,n}^{-1} \|v_{T,n}\|_{T \times I_n}^2 \\
 & \lesssim \langle R(u_{h\tau}), \lambda \rangle_{Y',Y} + \sum_{n=1}^N \sum_{T \in \overline{\mathcal{T}}^{n-1,n}} \eta_{\text{qd},T}^n C_{T,n}^{-\frac{1}{2}} \|v_{T,n}\|_{T \times I_n} \\
 (6.8) \quad & \leq \mathcal{J}_{u,\text{FR}}(u_{h\tau}) \|\lambda\|_Y + \eta_{\text{qd}} \left\{ \sum_{n=1}^N \sum_{T \in \overline{\mathcal{T}}^{n-1,n}} C_{T,n}^{-1} \|v_{T,n}\|_{T \times I_n}^2 \right\}^{\frac{1}{2}}.
 \end{aligned}$$

The definition (2.5b) of the  $\|\cdot\|_Y$ -norm, that of  $\lambda$ , the fact that we suppose that the ratio  $h_T/h_{T'}$ ,  $T \in \underline{\mathcal{T}}^{n-1,n}$ ,  $T' \in \overline{\mathcal{T}}^{n-1,n}$ ,  $T' \subset T$ , is bounded, and (6.4) lead to

$$(6.9) \quad \|\lambda\|_Y^2 = \sum_{n=1}^N \sum_{T \in \underline{\mathcal{T}}^{n-1,n}} C_{T,n}^{-2} \|\psi_{T,n} v_{T,n}\|_{Y,T \times I_n}^2 \lesssim \sum_{n=1}^N \sum_{T \in \overline{\mathcal{T}}^{n-1,n}} C_{T,n}^{-1} \|v_{T,n}\|_{T \times I_n}^2.$$

Combining (6.8) and (6.9) proves (6.7a).  $\square$

**7. Application to IPDG methods.** We apply the framework of sections 3 and 4 to IPDG methods as an example of nonconforming space discretization with Crank–Nicolson time stepping. This consists in specifying the flux reconstruction  $\mathbf{t}_{h\tau}$  and in verifying Assumptions 3.1 and 4.1.

**7.1. The IPDG method.** For all  $0 \leq n \leq N$ , let  $V_h^n := \mathbb{P}_p(\mathcal{T}^n)$ ,  $p \geq 1$ , be spanned by piecewise polynomials on the mesh  $\mathcal{T}^n$  with total degree  $\leq p$ . For each  $F \in \mathcal{F}^n$ , along with the jump operator  $[[\cdot]]$ , we consider the (arithmetic) average operator  $\{\!\{ \cdot \}\!\}$  (conventionally yielding the actual value at boundary faces).

The space-time approximation  $u_{h\tau}$  is continuous and piecewise affine in time and is defined by its values  $u_h^n$  at  $t^n$  for all  $0 \leq n \leq N$ . We take  $u_h^0$  as the  $L^2$ -orthogonal projection of  $u_0$  onto  $V_h^0$ . Then, for all  $1 \leq n \leq N$ , we look for  $u_h^n \in V_h^n$  such that

$$\begin{aligned}
 (7.1) \quad & (\partial_t u_h^n, v_h) + \frac{1}{2} \sum_{m=n-1}^n \left\{ (\boldsymbol{\sigma}(u_h^m), \nabla u_h^m), \nabla v_h \right\} + \sum_{F \in \mathcal{F}^m} \alpha_{\mathbf{K},F}^m h_F^{-1} ([[u_h^m]], [[v_h]])_F \\
 & + \sum_{F \in \mathcal{F}^m} (H_F(u_h^m), [[v_h]])_F - \sum_{F \in \mathcal{F}^m} (\{\!\{ \mathbf{K}(u_h^m) \nabla u_h^m \}\!\}, \mathbf{n}_F, [[v_h]])_F \\
 & - \theta \sum_{F \in \mathcal{F}^m} (\{\!\{ \mathbf{K}(u_h^m) \nabla v_h \}\!\}, \mathbf{n}_F, [[u_h^m]])_F - (f^m, v_h) \left. \right\} \\
 & = 0 \quad \forall v_h \in V_h^n,
 \end{aligned}$$

where  $\alpha_{\mathbf{K},F}^n$  is the penalty coefficient,  $\theta \in \{-1, 0, 1\}$ , and  $\partial_t u_h^n := (\tau^n)^{-1} (u_h^n - u_h^{n-1})$ ; we suppose for simplicity that  $f$  is continuous and piecewise affine in time and denote

$f^n := f(t^n)$ . The advection term in (7.1) has been discretized using a numerical flux  $H_F(u_h^n)$  satisfying the following reasonable assumption.

*Assumption 7.1* (numerical flux for advection). For all  $1 \leq n \leq N$ , all mesh faces  $F \in \mathcal{F}^n$ , and all  $v_h \in V_h^n$ , there holds

$$\|H_F(v_h) - \{\!\!\{ \phi(v_h) \}\!\!\} \cdot \mathbf{n}_F\|_F \lesssim \|\phi'(v_h)\|_{\infty, \mathcal{T}_F} \| \llbracket v_h \rrbracket \|_F.$$

An example is the numerical flux of Lax–Friedrichs type, which consists of the centered flux  $\{\!\!\{ \phi(\cdot) \}\!\!\} \cdot \mathbf{n}_F$  supplemented by a stabilization term penalizing the interface jumps. In our numerical experiments, we employ a numerical upwinding flux; see section 8.1.

**7.2. Flux reconstruction.** Let  $l \geq 0$ . We construct  $\mathbf{t}_{h\tau}$  continuous and piecewise affine in time, with  $\mathbf{t}_h^n := \mathbf{t}_{h\tau}(t^n)$ ,  $0 \leq n \leq N$ , in the Raviart–Thomas–Nédélec space  $\mathbf{RTN}_l(\mathcal{T}^n) := \{ \mathbf{v}_h \in \mathbf{H}(\text{div}, \Omega); \mathbf{v}_h|_T \in \mathbf{RTN}_l(T) \text{ for all } T \in \mathcal{T}^n \}$ , where  $\mathbf{RTN}_l(T) := [\mathbb{P}_l(T)]^d + \mathbf{x}\mathbb{P}_l(T)$ . In particular,  $\mathbf{v}_h \in \mathbf{RTN}_l(\mathcal{T}^n)$  is such that for all  $T \in \mathcal{T}^n$ ,  $\nabla \cdot \mathbf{v}_h \in \mathbb{P}_l(T)$ , and for all  $F \in \mathcal{F}^n$ ,  $\mathbf{v}_h \cdot \mathbf{n}_F \in \mathbb{P}_l(F)$ , with continuity across interfaces; cf. Brezzi and Fortin [7]. Following the work of Kim [20] and [13], we set the following.

**DEFINITION 7.2** (reconstructed flux  $\mathbf{t}_{h\tau}$ ). Let  $l \geq 0$ . Set  $w_F := \frac{1}{2}$  for interfaces and  $w_F := 1$  for boundary faces. For all  $0 \leq n \leq N$ , we specify  $\mathbf{t}_h^n \in \mathbf{RTN}_l(\mathcal{T}^n)$  by setting for all  $T \in \mathcal{T}^n$ , all  $F \in \mathcal{F}_T$ , all  $q_h \in \mathbb{P}_l(F)$ , and all  $\mathbf{r}_h \in [\mathbb{P}_{l-1}(T)]^d$ ,

$$(7.2) \quad (\mathbf{t}_h^n \cdot \mathbf{n}_F, q_h)_F = (-\{\!\!\{ \mathbf{K}(u_h^n) \nabla u_h^n \}\!\!\} \cdot \mathbf{n}_F + \alpha_{\mathbf{K}, F}^n h_F^{-1} \llbracket u_h^n \rrbracket + H_F(u_h^n), q_h)_F,$$

$$(7.3) \quad (\mathbf{t}_h^n, \mathbf{r}_h)_T = (-\mathbf{K}(u_h^n) \nabla u_h^n + \phi(u_h^n), \mathbf{r}_h)_T + \theta \sum_{F \in \mathcal{F}_T} w_F (\mathbf{K}(u_h^n) \mathbf{r}_h \cdot \mathbf{n}_F, \llbracket u_h^n \rrbracket)_F.$$

*Remark 7.3* (alternative construction). Instead of prescribing directly the degrees of freedom for  $\mathbf{t}_h^n$ , it is also possible to reconstruct the flux by solving local Neumann problems by mixed finite elements; see [16]. This approach can achieve a tighter relationship between the error and the estimator but is slightly more expensive.

**7.3. Verification of Assumption 3.1.**

**LEMMA 7.4** (local conservation). Let  $u_h^n$ ,  $1 \leq n \leq N$ , solve (7.1), and let  $\mathbf{t}_{h\tau}$  be defined by (7.2)–(7.3) with  $l \geq 0$ . Then,  $\mathbf{t}_{h\tau}$  satisfies  $\mathbf{t}_{h\tau} \in \mathbf{L}^2(0, t_F; \mathbf{H}(\text{div}, \Omega))$  and, for all  $1 \leq n \leq N$  and all  $T \in \underline{\mathcal{T}}^{n-1, n}$ ,

$$(7.4) \quad (f - \partial_t u_{h\tau} - \nabla \cdot \mathbf{t}_{h\tau}, v_h)_{T \times I_n} = 0 \quad \forall v_h \in \mathbb{P}_{\min(p, l)}(T),$$

so that, in particular, the local space-time conservation property (3.1) holds.

*Proof.* By construction, we have for  $0 \leq n \leq N$ ,  $\mathbf{t}_h^n \in \mathbf{H}(\text{div}, \Omega)$ , so that  $\mathbf{t}_{h\tau} \in \mathbf{L}^2(0, t_F; \mathbf{H}(\text{div}, \Omega))$ . Let  $1 \leq n \leq N$ ,  $T \in \underline{\mathcal{T}}^{n-1, n}$ , and let  $v_h \in \mathbb{P}_{\min(p, l)}(T)$  be given. Using that  $u_{h\tau}$ ,  $f$ , and  $\mathbf{t}_{h\tau}$  are affine in time on  $I_n$ , that  $T$  is from the common coarsening of the meshes  $\mathcal{T}^{n-1}$  and  $\mathcal{T}^n$ , and the Green theorem, we obtain

$$\begin{aligned} & (f - \partial_t u_{h\tau} - \nabla \cdot \mathbf{t}_{h\tau}, v_h)_{T \times I_n} \\ &= -\tau^n (\partial_t u_{h\tau}, v_h)_T + \frac{\tau^n}{2} \sum_{m=n-1}^n (f^m - \nabla \cdot \mathbf{t}_h^m, v_h)_T \\ &= -\tau^n \left( \partial_t u_{h\tau} - \frac{1}{2} \sum_{m=n-1}^n f^m, v_h \right)_T \\ & \quad + \frac{\tau^n}{2} \sum_{m=n-1}^n \sum_{T' \in \mathcal{T}^m, T' \subset T} \{ (\mathbf{t}_h^m, \nabla v_h)_{T'} - (\mathbf{t}_h^m \cdot \mathbf{n}_{T'}, v_h)_{\partial T'} \}. \end{aligned}$$

Let  $m = n$  or  $m = n - 1$ . The facts that  $v_h|_F \in \mathbb{P}_l(F)$  for all  $F \in \mathcal{F}_{T'}$ , (7.2),  $\nabla v_h \in [\mathbb{P}_{l-1}(T')]^d$  for any  $T'$ , and (7.3) yield

$$\begin{aligned} (\mathbf{t}_h^m \cdot \mathbf{n}_{T'}, v_h)_{\partial T'} &= - \sum_{F \in \mathcal{F}_{T'}} \mathbf{n}_{T'} \cdot \mathbf{n}_F \{ (-\{\mathbf{K}(u_h^m) \nabla u_h^m\} \cdot \mathbf{n}_F \\ &\quad + \alpha_{\mathbf{K}, F}^m h_F^{-1} \llbracket u_h^m \rrbracket + H_F(u_h^m, v_h)_F \}, \\ (\mathbf{t}_h^m, \nabla v_h)_{T'} &= (-\mathbf{K}(u_h^m) \nabla u_h^m + \phi(u_h^m), \nabla v_h)_{T'} \\ &\quad + \theta \sum_{F \in \mathcal{F}_{T'}} w_F (\mathbf{K}(u_h^m) \nabla v_h \cdot \mathbf{n}_F, \llbracket u_h^m \rrbracket)_F. \end{aligned}$$

Extending  $v_h$  by 0 outside  $T$  so that a function in  $V_h^n$  is obtained and using the above identities and the definition (7.1) of the scheme yields (7.4).  $\square$

**7.4. Verification of Assumption 4.1.** To verify Assumption 4.1, we first specify the space-time piecewise polynomial  $\bar{\sigma}(u_h^n, \nabla u_h^n)$ . Let  $0 \leq n \leq N$  and let  $T \in \mathcal{T}^n$ . We define  $\bar{\sigma}(u_h^n, \nabla u_h^n)|_T \in \mathbf{RTN}_l(T)$  such that for all  $F \in \mathcal{F}_T$ , all  $q_h \in \mathbb{P}_l(F)$ , and all  $\mathbf{r}_h \in [\mathbb{P}_{l-1}(T)]^d$ ,

$$(7.5) \quad (\bar{\sigma}(u_h^n, \nabla u_h^n)|_T \cdot \mathbf{n}_F, q_h)_F = (\sigma(u_h^n, \nabla u_h^n)|_T \cdot \mathbf{n}_F, q_h)_F,$$

$$(7.6) \quad (\bar{\sigma}(u_h^n, \nabla u_h^n), \mathbf{r}_h)_T = (\sigma(u_h^n, \nabla u_h^n), \mathbf{r}_h)_T.$$

Here,  $l$  is the polynomial degree used for reconstructing the flux  $\mathbf{t}_h^n$  in section 7.2. We observe that locally in each mesh element  $T \in \mathcal{T}^n$ ,  $\bar{\sigma}(u_h^n, \nabla u_h^n)|_T$  belongs to  $\mathbf{RTN}_l(T)$  (as  $\mathbf{t}_h^n$  does), but globally,  $\bar{\sigma}(u_h^n, \nabla u_h^n) \notin \mathbf{RTN}_l(\mathcal{T}^n)$  because the normal component of  $\bar{\sigma}(u_h^n, \nabla u_h^n)$  is in general discontinuous across interfaces. Finally, the space-time function  $\bar{\sigma}(u_{h\tau}, \nabla u_{h\tau})$  is taken to be continuous and piecewise affine in time, matching the values  $\bar{\sigma}(u_h^n, \nabla u_h^n)$  at  $t^n$  for all  $0 \leq n \leq N$ .

LEMMA 7.5 (flux approximation). *Let  $\mathbf{t}_{h\tau}$  be defined by (7.2)–(7.3) with  $l \geq 0$  and let  $\bar{\sigma}(u_{h\tau}, \nabla u_{h\tau})$  be defined by (7.5)–(7.6). For  $1 \leq n \leq N$  and  $T \in \bar{\mathcal{T}}^{n-1, n}$ , define the constants  $C_{\mathbf{K}, \phi, T, F, n}$  of (2.12) by*

$$(7.7) \quad C_{\mathbf{K}, \phi, T, F, n} := (\alpha_{\mathbf{K}, F}^n)^2 h_F^{-1} + \|\mathbf{K}(u_h^n)\|_{\infty, T}^2 h_F^{-1} + \|\phi'(u_h^n)\|_{\infty, \mathcal{T}_F}^2 h_F.$$

Then, Assumption 4.1 holds.

*Proof.* Let  $0 \leq n \leq N$  and set  $\mathbf{v}_h := \bar{\sigma}(u_h^n, \nabla u_h^n) + \mathbf{t}_h^n$ . Recall that  $\sigma(u_h^n, \nabla u_h^n) = \mathbf{K}(u_h^n) \nabla u_h^n - \phi(u_h^n)$ . Let  $T \in \mathcal{T}^n$  and  $F \in \mathcal{F}_T$ . Using (7.2), (7.5) and (7.3), (7.6) yields, respectively,

$$\begin{aligned} (\mathbf{v}_h|_T \cdot \mathbf{n}_F, q_h)_F &= ((1 - w_F) \llbracket \bar{\sigma}(u_h^n, \nabla u_h^n) \rrbracket \cdot \mathbf{n}_F + \alpha_{\mathbf{K}, F}^n h_F^{-1} \llbracket u_h^n \rrbracket \\ &\quad + H_F(u_h^n) - \{\phi(u_h^n)\} \cdot \mathbf{n}_F, q_h)_F, \\ (\mathbf{v}_h, \mathbf{r}_h)_T &= \theta \sum_{F \in \mathcal{F}_T} w_F (\mathbf{K}(u_h^n) \mathbf{r}_h \cdot \mathbf{n}_F, \llbracket u_h^n \rrbracket)_F, \end{aligned}$$

for all  $q_h \in \mathbb{P}_l(F)$  and all  $\mathbf{r}_h \in [\mathbb{P}_{l-1}(T)]^d$ . Thus, using Assumption 7.1 on the numerical flux for advection and employing an inverse inequality in space, we infer

$$\begin{aligned} \|\mathbf{v}_h \cdot \mathbf{n}_F\|_F &\lesssim (1 - w_F) \|\llbracket \bar{\sigma}(u_h^n, \nabla u_h^n) \rrbracket \cdot \mathbf{n}_F\|_F + (\alpha_{\mathbf{K}, F}^n h_F^{-1} + \|\phi'(u_h^n)\|_{\infty, \mathcal{T}_F}) \|\llbracket u_h^n \rrbracket\|_F, \\ (\mathbf{v}_h, \mathbf{r}_h)_T &\lesssim \|\mathbf{K}(u_h^n)\|_{\infty, T} \|\mathbf{r}_h\|_T \sum_{F \in \mathcal{F}_T} h_F^{-\frac{1}{2}} \|\llbracket u_h^n \rrbracket\|_F. \end{aligned}$$

Using these two estimates in the classical bound  $\|\mathbf{v}_h\|_T^2 \lesssim \sum_{F \in \mathcal{F}_T} h_F \|\mathbf{v}_h \cdot \mathbf{n}_F\|_F^2 + (\sup_{\mathbf{r}_h \in [\mathbb{P}_{l-1}(T)]^d} \frac{(\mathbf{v}_h, \mathbf{r}_h)_T}{\|\mathbf{r}_h\|_T})^2$ , valid for all  $\mathbf{v}_h \in \mathbf{RTN}_l(T)$ , and owing to (7.7),

$$(7.8) \quad \|\bar{\boldsymbol{\sigma}}(u_h^n, \nabla u_h^n) + \mathbf{t}_h^n\|_T^2 \lesssim \sum_{F \in \mathcal{F}_T^{\text{int}}} h_F \|\llbracket \bar{\boldsymbol{\sigma}}(u_h^n, \nabla u_h^n) \rrbracket \cdot \mathbf{n}_F\|_F^2 + \sum_{F \in \mathcal{F}_T} C_{\mathbf{K}, \phi, T, F, n} \|[u_h^n]\|_F^2.$$

Let now  $1 \leq n \leq N$  and  $T \in \mathcal{T}^{n-1, n}$ . Using that both  $\bar{\boldsymbol{\sigma}}(u_{h\tau}, \nabla u_{h\tau})$  and  $u_{h\tau}$  are piecewise affine in time, we have (cf. [17, Lemma 6.1])

$$\begin{aligned} \|\bar{\boldsymbol{\sigma}}(u_{h\tau}, \nabla u_{h\tau}) + \mathbf{t}_{h\tau}\|_{T \times I_n}^2 &\lesssim \tau^n \sum_{m=n-1}^n \|\bar{\boldsymbol{\sigma}}(u_h^m, \nabla u_h^m) + \mathbf{t}_h^m\|_T^2 \\ &= \tau^n \sum_{m=n-1}^n \sum_{T' \in \mathcal{T}^m, T' \subset T} \|\bar{\boldsymbol{\sigma}}(u_h^m, \nabla u_h^m) + \mathbf{t}_h^m\|_{T'}^2 \\ &\lesssim \sum_{T' \in \bar{\mathcal{T}}^{n-1, n}, T' \subset T} (\eta_{\text{clas}, T'}^n)^2, \end{aligned}$$

employing (7.8) on  $T' \in \mathcal{T}^m$ ; actually, only the jump terms of  $\eta_{\text{clas}, T'}^n$  appear.  $\square$

*Remark 7.6* (choice of the reconstruction degree  $l$ ). A typical choice for the polynomial degree in the flux reconstruction  $\mathbf{t}_{h\tau}$  is  $l \in \{p - 1, p\}$ . Choosing larger values for  $l$  ensures theoretically that the quadrature errors are small enough. In our numerical experiments, the choice  $l = p$  was sufficient to deliver quadrature errors one to two orders of magnitude smaller than the leading estimators.

**8. Numerical experiments.** In this section, we present numerical experiments illustrating our a posteriori error estimates.

**8.1. Setting.** We consider the problem (1.1a)–(1.1c), where, in some cases, we replace the homogeneous Dirichlet boundary condition (1.1b) by an inhomogeneous one. The error for nonpolynomial Dirichlet boundary data is neglected, since it is generally of higher order; for a numerical study, see, e.g., Luce and Wohlmuth [22].

We employ the IPDG method (7.1) with  $\theta = 0$  and the upwind numerical flux such that for all  $1 \leq n \leq N$  and all  $F \in \mathcal{F}^n$ ,  $H_F(u_h^n)|_F := \phi(u_h^n)|_F \cdot \mathbf{n}_F$  if  $\llbracket \phi'(u_h^n) \rrbracket_F \cdot \mathbf{n}_F > 0$  and  $H_F(u_h^n)|_F := \phi(u_h^n)|_F^R \cdot \mathbf{n}_F$  otherwise, where  $v_h^n|_F^L$  and  $v_h^n|_F^R$  denote traces of a function  $v_h^n$  on  $F$  from the direction and the opposite direction of  $\mathbf{n}_F$ , respectively. The penalty parameter in (7.1) is chosen as  $\alpha_{\mathbf{K}, F}^n := 10p^2 \|\mathbf{K}(u_h^n)\|_{\infty, \mathcal{T}_F}$ ; see Houston, Robson, and Süli [18], [9], and [15] for problems with internal layers caused by locally small diffusion. We consider polynomial degrees  $p \in \{1, 2, 3\}$ . For the linearization of (7.1), we use the Newton-like method of [11], where the (approximate) construction of the Jacobian matrix is avoided via the idea of the easy-to-evaluate flux matrix. The volume integrals are evaluated with the Dunavant quadrature rule of order  $(3p + 2)$  and the face integrals with the Gauss quadrature rule with  $(2p + 2)$  nodes.

We consider square domains  $\Omega$  with uniform discretizations (consisting of right-angled triangles obtained by diagonal cuttings of squares). The time step and the meshes are kept fixed in time. We consider a family of uniformly refined space-time meshes characterized by the squares edge lengths and time steps  $(h_{m,p}, \tau_{m,p})$ ,  $m \in \{1, 2, 3\}$ ,  $p \in \{1, 2, 3\}$ . We choose  $(h_0, \tau_0)$  and then set  $h_{m,p} := h_0 2^{1-m}$ ,  $m, p \in \{1, 2, 3\}$ , and  $\tau_{m,1} := \tau_0 2^{(1-m)/2}$ ,  $\tau_{m,2} := 0.5\tau_0 2^{(1-m)}$ ,  $\tau_{m,3} := 0.2\tau_0 2^{3(1-m)/2}$ ,  $m \in \{1, 2, 3\}$ . We evaluate experimental orders of convergence in the form  $\text{EOC} :=$



TABLE 8.1

Example (8.1),  $\varepsilon = 10^{-2}$ , and  $(h_0, \tau_0) = (1/6, 0.05)$ : errors, estimators, and effectivity indices.

$m$	$p$	$\mathcal{J}_{u,FR}(u_{h\tau})$	$\eta_F$	$\eta_R$	$\eta_{NC}$	$\eta_{IC}$	$\eta_{qd}$	$\eta$	$i_e$	$i_{e,FR}$
1	1	1.50E-02	1.11E-02	2.28E-02	4.11E-02	2.94E-02	3.82E-03	1.04E-01	1.85	1.15
2	1	1.17E-02 ( 0.36)	8.30E-03 ( 0.43)	1.52E-02 ( 0.59)	2.29E-02 ( 0.84)	1.31E-02 ( 1.16)	1.92E-03 ( 0.99)	5.94E-02 ( 0.81)	1.71	1.35
3	1	1.02E-02 ( 0.20)	5.16E-03 ( 0.69)	7.78E-03 ( 0.96)	1.16E-02 ( 0.98)	2.69E-03 ( 2.29)	7.49E-04 ( 1.36)	2.72E-02 ( 1.13)	1.25	1.36
1	2	4.97E-03	3.78E-03	8.23E-03	1.23E-02	1.32E-02	9.38E-04	3.72E-02	2.15	1.01
2	2	1.74E-03 ( 1.52)	1.36E-03 ( 1.47)	2.52E-03 ( 1.71)	4.02E-03 ( 1.61)	1.76E-03 ( 2.90)	2.34E-04 ( 2.00)	9.54E-03 ( 1.96)	1.65	0.94
3	2	4.63E-04 ( 1.91)	4.00E-04 ( 1.77)	7.36E-04 ( 1.77)	1.26E-03 ( 1.67)	3.01E-04 ( 2.55)	3.97E-05 ( 2.56)	2.63E-03 ( 1.86)	1.53	1.08
1	3	1.78E-03	9.11E-04	1.69E-03	3.41E-03	3.01E-03	2.20E-04	8.88E-03	1.71	0.59
2	3	3.47E-04 ( 2.35)	1.57E-04 ( 2.54)	3.26E-04 ( 2.38)	6.06E-04 ( 2.49)	6.20E-04 ( 2.28)	2.50E-05 ( 3.14)	1.67E-03 ( 2.41)	1.75	0.73
3	3	1.33E-05 ( 4.71)	1.80E-05 ( 3.12)	3.81E-05 ( 3.10)	6.97E-05 ( 3.12)	8.88E-05 ( 2.80)	1.64E-06 ( 3.93)	2.10E-04 ( 2.99)	2.54	0.97

$\frac{\log(E_{m,p}/E_{m-1,p})}{\log(h_{m,p}/h_{m-1,p})}$ ,  $m \in \{2, 3\}$ , where  $E_{m,p}$  is either an error or an error estimator on the space-time discretization  $(h_{m,p}, \tau_{m,p})$ .

The weights  $C_{T,n}$  are chosen according to (2.6) with  $C_{\phi,T,n} := h_\Omega \|\phi'(u_{h\tau})\|_{\infty,Q}$  (here,  $h_\Omega$  denotes the diameter of  $\Omega$ ) and  $C_{\mathbf{K},T,n} := (h_\Omega/h_T) \|\mathbf{K}(u_{h\tau})\|_{\infty,Q}$ . The first two addends on the right-hand side of (2.6) are of similar size if the Courant numbers  $\tau^n \|\phi'(u_{h\tau})\|_{\infty,Q}/h_T$  are of order unity, as well as the ratios  $t_F \|\phi'(u_{h\tau})\|_{\infty,Q}/h_\Omega$  (meaning that the final time allows particles to be advected across a relevant part of the domain). Moreover, the ratio of the second to the third addend is of the order of the Péclet numbers  $h_T \|\phi'(u_{h\tau})\|_{\infty,Q} / \|\mathbf{K}(u_{h\tau})\|_{\infty,Q}$ .

To evaluate the error measure  $\mathcal{J}_{u,FR}(u_{h\tau})$ , we solve (approximately) the dual problem introduced in section 2.3.1. This problem, which is posed on the space-time domain  $Q$ , is solved using the finite difference package Fishpack [1] on a structured space-time mesh obtained by typically three successive dyadic partitions of the space-time grid on which the approximate solution  $u_{h\tau}$  has been computed. Recalling that  $\mathcal{J}_u(u_{h\tau}) = \mathcal{J}_{u,FR}(u_{h\tau}) + \mathcal{J}_{u,NC}(u_{h\tau})$ , where  $\mathcal{J}_{u,NC}(u_{h\tau})$  coincides with the non-conformity estimator  $\eta_{NC}$ , we evaluate the effectivity index  $i_e := \eta/\mathcal{J}_u(u_{h\tau})$ , where  $\eta := \eta_{FR} + \eta_{NC} + \eta_{IC}$ , to illustrate Theorems 3.3 and 4.4. Moreover, to illustrate Theorem 4.2, we evaluate the effectivity index  $i_{e,FR} := \eta/(e_{FR} + \mathcal{J}_{u,NC}(u_{h\tau}))$ , where  $e_{FR}$  is the locally computable upper bound on  $\mathcal{J}_{u,FR}(u_{h\tau})$ ; see (2.10)–(2.11).

**8.2. Linear diffusion.** We consider first a model advection-diffusion problem with  $\Omega = (-1, 1) \times (-1, 1)$ ,  $t_F = 1$ , and linear diffusive but nonlinear advective flux. In particular,  $\mathbf{K}(u) = \varepsilon \mathbb{I}$  ( $\mathbb{I}$  being the identity tensor),  $\varepsilon > 0$ , and  $\phi(u) = (u^2/2, u^2/2)^T$  in (1.2). The initial and boundary conditions are such that the exact solution is

$$(8.1) \quad u(x, y, t) = (1 + \exp(\vartheta(x, y, t)))^{-1}$$

with  $\vartheta(x, y, t) := \frac{x+y+1-t}{2\varepsilon}$ ; see [10]. This problem exhibits an inner layer moving in the diagonal direction. The steepness of the layer increases with decreasing  $\varepsilon$ .

We first employ the initial space-time step  $(h_0, \tau_0) = (1/6, 0.05)$  and consider the value  $\varepsilon = 10^{-2}$ . Results are reported in Table 8.1. The flux, residual, nonconforming, and initial condition estimators take relatively close values, while the quadrature estimator is lower. We observe that for all values of  $m$  and  $p$ , the effectivity index  $i_e$  takes stable values in a relatively narrow interval, independently of  $m$  and  $p$ . The same observation holds for the effectivity index  $i_{e,FR}$ , with, however, values  $\leq 1$ , since  $e_{FR}$  is only an upper bound on  $\mathcal{J}_{u,FR}(u_{h\tau})$ .

TABLE 8.2

Example (8.1): effectivity indices for different choices of  $\varepsilon$  and  $(h_0, \tau_0)$ .

$\varepsilon$		$10^{-2}$		$10^{-2}$		$10^{-2}$		$10^{-4}$	
$(h_0, \tau_0)$		$(1/6, 0.05)$		$(1/6, 0.2)$		$(1/6, 0.0125)$		$(1/6, 0.05)$	
$m$	$p$	$i_e$	$i_{e,FR}$	$i_e$	$i_{e,FR}$	$i_e$	$i_{e,FR}$	$i_e$	$i_{e,FR}$
1	1	1.85	1.15	2.21	1.28	3.00	0.81	1.45	0.71
2	1	1.71	1.35	2.38	1.12	2.45	1.03	1.68	1.06
3	1	1.25	1.36	2.15	0.90	1.33	1.03	1.82	1.34
1	2	2.15	1.01	3.13	1.71	3.69	0.67	1.38	0.62
2	2	1.65	0.94	2.74	1.58	2.16	0.49	1.41	0.62
3	2	1.53	1.08	2.38	1.52	1.83	0.58	1.54	0.69
1	3	1.71	0.59	2.74	1.47	3.00	0.34	1.26	0.31
2	3	1.75	0.73	2.63	1.67	3.15	0.46	1.13	0.21
3	3	2.54	0.97	2.77	1.73	—	0.69	1.03	0.15

TABLE 8.3

Example (8.2): errors, estimators, and effectivity indices.

$m$	$p$	$J_{u,FR}(u_{h,\tau})$	$\eta_F$	$\eta_R$	$\eta_{NC}$	$\eta_{IC}$	$\eta_{qd}$	$\eta$	$i_e$	$i_{e,FR}$
1	1	9.91E-03	1.00E-02	6.02E-03	2.77E-02	2.31E-02	2.17E-03	6.62E-02	1.76	0.97
2	1	7.39E-03	7.71E-03	5.68E-03	1.62E-02	7.71E-03	1.23E-03	3.66E-02	1.55	1.02
		(0.42)	(0.37)	(0.08)	(0.78)	(1.59)	(0.82)	(0.86)		
3	1	4.58E-03	4.52E-03	4.95E-03	8.33E-03	1.86E-03	5.22E-04	1.89E-02	1.47	1.16
		(0.69)	(0.77)	(0.20)	(0.96)	(2.05)	(1.23)	(0.95)		
1	2	2.62E-03	3.30E-03	5.40E-03	9.33E-03	6.27E-03	6.74E-04	2.35E-02	1.97	0.73
2	2	1.11E-03	1.43E-03	1.93E-03	4.22E-03	1.09E-03	2.67E-04	8.34E-03	1.56	0.62
		(1.23)	(1.21)	(1.48)	(1.14)	(2.52)	(1.34)	(1.50)		
3	2	4.26E-04	5.63E-04	6.13E-04	1.84E-03	1.51E-04	1.00E-04	3.06E-03	1.35	0.57
		(1.38)	(1.34)	(1.65)	(1.20)	(2.85)	(1.42)	(1.45)		
1	3	6.48E-04	8.83E-04	1.03E-03	3.57E-03	1.19E-03	2.31E-04	6.47E-03	1.53	0.36
2	3	1.94E-04	2.63E-04	1.45E-04	1.21E-03	1.07E-04	6.39E-05	1.69E-03	1.21	0.25
		(1.74)	(1.74)	(2.84)	(1.56)	(3.48)	(1.85)	(1.93)		
3	3	4.42E-05	7.58E-05	2.58E-05	4.04E-04	7.47E-06	1.67E-05	5.07E-04	1.13	0.21
		(2.13)	(1.80)	(2.49)	(1.58)	(3.84)	(1.94)	(1.74)		

To assess robustness with respect to advection dominance and relative size of space and time steps, we consider some different choices for  $\varepsilon$  and  $(h_0, \tau_0)$ . In Table 8.2, we compare the effectivity indices  $i_e$  and  $i_{e,FR}$  with those of Table 8.1 (repeated for convenience); for  $m = p = 3$  and  $(h_0, \tau_0) = (1/6, 0.0125)$ , the value of  $i_e$  is not reported because of the cost of solving accurately the dual problem. We observe that all the considered changes in  $\varepsilon$  and  $(h_0, \tau_0)$  only mildly modify the effectivity indices.

**8.3. Degenerate diffusion.** We first consider a nonlinear degenerate advection-diffusion problem from Kačur [19] with  $\Omega = (0, 1) \times (0, 1)$ ,  $t_F = 1$ ,  $\mathbf{K}(u) = 2\varepsilon u\mathbb{I}$ ,  $\varepsilon = 10^{-2}$ , and  $\phi(u) = 0.5(u^2, 0)^T$ . The initial and boundary conditions are such that the exact solution is

$$(8.2) \quad u(x, y, t) = \begin{cases} 1 - \exp\left(\frac{v(x-vt-x_0)}{2\varepsilon}\right) & \text{for } x \leq vt + x_0, \\ 0 & \text{for } x > vt + x_0, \end{cases}$$

where  $x_0 = 1/4$  is the initial position of the front. If  $u = 0$  (for  $x \geq vt + x_0$ ), the diffusive term degenerates. The initial space-time step is  $(h_0, \tau_0) = (1/8, 0.05)$ . Results are summarized in Table 8.3. We observe again that the quadrature estimator is smaller than the other estimators and that the effectivity indices  $i_e$  and  $i_{e,FR}$  take relatively stable values, independently of  $m$  and  $p$ .

Our last example is the porous medium equation with  $\Omega = (-6, 6) \times (-6, 6)$ ,  $t_F = 1$ ,  $\mathbf{K}(u) = \kappa|u|^{\kappa-1}\mathbb{I}$ ,  $\kappa > 1$ , and  $\phi(u) = \mathbf{0}$ . The initial and boundary conditions

TABLE 8.4  
 Example (8.3),  $\kappa = 2$ : errors, estimators, and effectivity indices (also for  $\kappa = 4$ ).

		$\kappa = 2$							$\kappa = 4$			
$m$	$p$	$J_{u,FR}(u_{h\tau})$	$\eta_F$	$\eta_R$	$\eta_{NC}$	$\eta_C$	$\eta_{qd}$	$\eta$	$i_e$	$i_{e,FR}$	$i_e$	$i_{e,FR}$
1	1	7.90E-03	5.90E-03	1.32E-02	9.10E-03	3.23E-02	7.08E-05	5.88E-02	3.46	0.92	4.68	0.98
2	1	8.36E-03 (-0.08)	4.64E-03 (0.35)	1.71E-02 (-0.38)	8.46E-03 (0.10)	1.11E-02 (1.54)	3.99E-05 (0.83)	4.03E-02 (0.54)	2.40	1.46	3.72	1.62
3	1	8.91E-03 (-0.09)	4.38E-03 (0.08)	2.18E-02 (-0.35)	9.56E-03 (-0.18)	3.44E-03 (1.69)	1.83E-05 (1.13)	3.87E-02 (0.06)	2.09	2.49	3.38	2.68
1	2	1.09E-03	1.06E-02	1.06E-01	3.12E-02	1.35E-02	1.74E-04	1.61E-01	4.99	3.22	5.13	3.18
2	2	4.02E-04 (1.43)	8.04E-03 (0.40)	8.12E-02 (0.39)	2.37E-02 (0.40)	5.16E-03 (1.39)	6.40E-05 (1.45)	1.18E-01 (0.45)	4.90	3.89	5.05	3.84
3	2	1.28E-04 (1.65)	5.22E-03 (0.62)	5.33E-02 (0.61)	1.55E-02 (0.61)	1.69E-03 (1.61)	2.23E-05 (1.52)	7.57E-02 (0.64)	4.84	4.26	4.97	4.30
1	3	6.53E-04	2.26E-02	3.27E-01	7.58E-02	8.39E-03	1.36E-04	4.33E-01	5.67	5.01	5.67	4.88
2	3	1.78E-04 (1.87)	9.26E-03 (1.29)	1.38E-01 (1.24)	3.13E-02 (1.27)	3.14E-03 (1.42)	3.51E-05 (1.95)	1.82E-01 (1.25)	5.76	5.17	5.78	5.03
3	3	3.83E-05 (2.22)	3.41E-03 (1.44)	5.08E-02 (1.44)	1.15E-02 (1.45)	1.14E-03 (1.46)	8.89E-06 (1.98)	6.68E-02 (1.44)	5.80	5.21	5.85	5.10

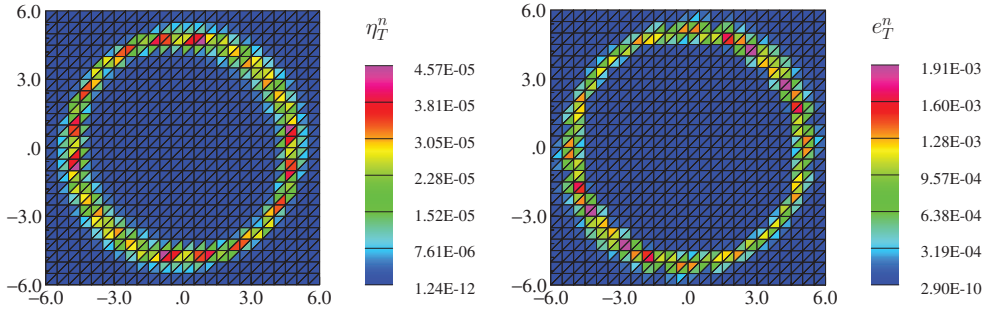


FIG. 8.1. Example (8.3),  $\kappa = 4$ : distribution of the element error estimate  $\eta_T^n$  (left) and of the local error  $e_T^n$  (right) at  $t = t_F$  for  $p = 2$  and  $m = 2$ .

are chosen so as to yield the so-called Barenblatt solution

$$(8.3) \quad u(x, y, t) = \left\{ \frac{1}{t+1} \left[ 1 - \frac{\kappa-1}{4\kappa^2} \frac{x^2 + y^2}{(t+1)^{1/\kappa}} \right]_+^{\frac{\kappa-1}{\kappa}} \right\}^{\frac{1}{\kappa}},$$

where  $[a]_+ = \max(a, 0)$ ,  $a \in \mathbb{R}$ . We consider the initial space-time steps  $(h_0, \tau_0) = (0.5, 0.02)$  and use here  $\tau_{m,p} := \tau_0 2^{(1-m)p/2}$ . We first set  $\kappa = 2$ . Results are summarized in Table 8.4. We can draw the same conclusions as above regarding the effectivity indices; we also observe that in this setting, the initial condition estimator is relatively important. To assess robustness with respect to the size of the nonlinearity, the last column of Table 8.4 reports the effectivity indices  $i_e$  and  $i_{e,FR}$  obtained for  $\kappa = 4$ ; we observe that they take close values to those obtained for  $\kappa = 2$ .

Finally, Figure 8.1 shows (for  $\kappa = 4$ ) the distribution of the local error  $e_T^n := e_{FR,T}^n + \eta_{NC,T}^n$  and the local error estimator  $\eta_T^n := \eta_{FR,T}^n + \eta_{NC,T}^n$  at  $t = t_F$  for  $m = 2$  and  $p = 2$ . We observe a close agreement in the distributions of error and estimator, thereby indicating that the latter can be used to drive an adaptive mesh procedure. A similar agreement (not shown) was observed for  $\kappa = 2$  and the other examples.

## REFERENCES

- [1] J. C. ADAMS, P. N. SWARZTRAUBER, AND R. SWEET, *FISHPACK: Efficient FORTRAN subprograms for the solution of separable elliptic partial differential equations*, <http://www2.cisl.ucar.edu/resources/legacy/fishpack> (2011).
- [2] M. AINSWORTH, *A framework for obtaining guaranteed error bounds for finite element approximations*, *J. Comput. Appl. Math.*, 234 (2010), pp. 2618–2632.
- [3] H. W. ALT AND S. LUCKHAUS, *Quasilinear elliptic-parabolic differential equations*, *Math. Z.*, 183 (1983), pp. 311–341.
- [4] D. N. ARNOLD, *An interior penalty finite element method with discontinuous elements*, *SIAM J. Numer. Anal.*, 19 (1982), pp. 742–760.
- [5] A. BERGAM, C. BERNARDI, AND Z. MGHAZLI, *A posteriori analysis of the finite element discretization of some parabolic equations*, *Math. Comp.*, 74 (2005), pp. 1117–1138.
- [6] D. BRAESS, V. PILLWEIN, AND J. SCHÖBERL, *Equilibrated residual error estimates are  $p$ -robust*, *Comput. Methods Appl. Mech. Engrg.*, 198 (2009), pp. 1189–1197.
- [7] F. BREZZI AND M. FORTIN, *Mixed and Hybrid Finite Element Methods*, Springer Ser. Comput. Math. 15, Springer-Verlag, New York, 1991.
- [8] A. L. CHAILLOU AND M. SURI, *Computable error estimators for the approximation of nonlinear problems by linearized models*, *Comput. Methods Appl. Mech. Engrg.*, 196 (2006), pp. 210–224.
- [9] V. DOLEJŠÍ, *Analysis and application of the IIPG method to quasilinear nonstationary convection-diffusion problems*, *J. Comput. Appl. Math.*, 222 (2008), pp. 251–273.
- [10] V. DOLEJŠÍ, M. FEISTAUER, V. KUČERA, AND V. SOBOTÍKOVÁ, *An optimal  $L^\infty(L^2)$ -error estimate of the discontinuous Galerkin method for a nonlinear nonstationary convection-diffusion problem*, *IMA J. Numer. Anal.*, 28 (2008), pp. 496–521.
- [11] V. DOLEJŠÍ, M. HOLÍK, AND J. HOZMAN, *Efficient solution strategy for the semi-implicit discontinuous Galerkin discretization of the Navier–Stokes equations*, *J. Comput. Phys.*, 230 (2011), pp. 4176–4200.
- [12] L. EL ALAOU, A. ERN, AND M. VOHRALÍK, *Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems*, *Comput. Methods Appl. Mech. Engrg.*, 200 (2011), pp. 2782–2795.
- [13] A. ERN, S. NICAISE, AND M. VOHRALÍK, *An accurate  $\mathbf{H}(\text{div})$  flux reconstruction for discontinuous Galerkin approximations of elliptic problems*, *C. R. Math. Acad. Sci. Paris*, 345 (2007), pp. 709–712.
- [14] A. ERN, A. F. STEPHANSEN, AND M. VOHRALÍK, *Guaranteed and robust discontinuous Galerkin a posteriori error estimates for convection-diffusion-reaction problems*, *J. Comput. Appl. Math.*, 234 (2010), pp. 114–130.
- [15] A. ERN, A. F. STEPHANSEN, AND P. ZUNINO, *A discontinuous Galerkin method with weighted averages for advection-diffusion equations with locally small and anisotropic diffusivity*, *IMA J. Numer. Anal.*, 29 (2009), pp. 235–256.
- [16] A. ERN AND M. VOHRALÍK, *Flux reconstruction and a posteriori error estimation for discontinuous Galerkin methods on general nonmatching grids*, *C. R. Math. Acad. Sci. Paris*, 347 (2009), pp. 441–444.
- [17] A. ERN AND M. VOHRALÍK, *A posteriori error estimation based on potential and flux reconstruction for the heat equation*, *SIAM J. Numer. Anal.*, 48 (2010), pp. 198–223.
- [18] P. HOUSTON, J. ROBSON, AND E. SÜLI, *Discontinuous Galerkin finite element approximation of quasilinear elliptic boundary value problems I: The scalar case*, *IMA J. Numer. Anal.*, 25 (2005), pp. 726–749.
- [19] J. KAČUR, *Solution of degenerate convection–diffusion problems by the method of characteristics*, *SIAM J. Numer. Anal.*, 39 (2001), pp. 858–879.
- [20] K. Y. KIM, *A posteriori error estimators for locally conservative methods of nonlinear elliptic problems*, *Appl. Numer. Math.*, 57 (2007), pp. 1065–1080.
- [21] P. LADEVÈZE, *Comparaison de modèles de milieux continus*, Ph.D. thesis, Université Pierre et Marie Curie, Paris 6, 1975.
- [22] R. LUCE AND B. I. WOHLMUTH, *A local a posteriori error estimator based on equilibrated fluxes*, *SIAM J. Numer. Anal.*, 42 (2004), pp. 1394–1414.
- [23] C. MAKRIDAKIS AND R. H. NOCHETTO, *Elliptic reconstruction and a posteriori error estimates for parabolic problems*, *SIAM J. Numer. Anal.*, 41 (2003), pp. 1585–1594.
- [24] R. H. NOCHETTO, A. SCHMIDT, AND C. VERDI, *A posteriori error estimation and adaptivity for degenerate parabolic problems*, *Math. Comp.*, 69 (2000), pp. 1–24.
- [25] F. OTTO,  *$L^1$ -contraction and uniqueness for quasilinear elliptic-parabolic equations*, *J. Differential Equations*, 131 (1996), pp. 20–38.

- [26] M. PICASSO, *Adaptive finite elements for a linear parabolic problem*, Comput. Methods Appl. Mech. Engrg., 167 (1998), pp. 223–237.
- [27] W. PRAGER AND J. L. SYNGE, *Approximations in elasticity based on the concept of function space*, Quart. Appl. Math., 5 (1947), pp. 241–269.
- [28] S. I. REPIN, *Estimates of deviations from exact solutions of initial-boundary value problem for the heat equation*, Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl., 13 (2002), pp. 121–133.
- [29] G. SANGALLI, *Robust a-posteriori estimator for advection-diffusion-reaction problems*, Math. Comp., 77 (2008), pp. 41–70.
- [30] R. VERFÜRTH, *A posteriori error estimates for nonlinear problems.  $L^r(0, T; L^p(\Omega))$ -error estimates for finite element discretizations of parabolic equations*, Math. Comp., 67 (1998), pp. 1335–1360.
- [31] R. VERFÜRTH, *A posteriori error estimates for nonlinear problems:  $L^r(0, T; W^{1,\rho}(\Omega))$ -error estimates for finite element discretizations of parabolic equations*, Numer. Methods Partial Differential Equations, 14 (1998), pp. 487–518.
- [32] R. VERFÜRTH, *A posteriori error estimates for finite element discretizations of the heat equation*, Calcolo, 40 (2003), pp. 195–212.
- [33] R. VERFÜRTH, *A Posteriori Error Estimates for Non-Linear Parabolic Equations*. Technical report, Ruhr-Universität Bochum, 2004.
- [34] R. VERFÜRTH, *Robust a posteriori error estimates for nonstationary convection-diffusion equations*, SIAM J. Numer. Anal., 43 (2005), pp. 1783–1802.