A framework for robust a posteriori error control in unsteady nonlinear advection-diffusion problems

Vít Dolejší, Alexandre Ern, Martin Vohralík

To cite this version:

HAL Id: hal-00652979
https://hal.archives-ouvertes.fr/hal-00652979v2
Submitted on 5 Sep 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A framework for robust a posteriori error control in unsteady nonlinear advection-diffusion problems

Vít Dolejší† Alexandre Ern‡ Martin Vohralík§

September 5, 2012

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

1 Introduction
We consider the unsteady nonlinear problem

\[ \frac{\partial}{\partial t} u - \nabla \cdot \sigma(u, \nabla u) = f \quad \text{in } Q := \Omega \times (0, t_F), \]
\[ u = 0 \quad \text{on } \partial \Omega \times (0, t_F), \]
\[ u(\cdot, 0) = u_0 \quad \text{in } \Omega, \]

(1.1)

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 \( \sigma(u, \nabla u) \) takes the form

\[ \sigma(u, \nabla u) := K(u) \nabla u - \phi(u), \]

(1.2)

where \( K(\cdot) \) is a nonlinear, possibly degenerate, tensor-valued function associated with diffusive transport and \( \phi(\cdot) \) a nonlinear vector-valued function associated with advective transport. Precise assumptions on \( K(\cdot) \) and \( \phi(\cdot) \) are specified in Section 2.1. In what follows, \( u \) is termed the potential and \( -\sigma(u, \nabla u) \) the...
advection-diffusive flux. We assume that the problem (1.1a)–(1.1c) admits a unique weak solution $u$ and that a fully 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 et al. [5], both reliability and efficiency are proven 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 et al. [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 approaches have 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 efficient estimates for this weighted $L^2$-norm.

Our results are derived in a unified framework where the actual numerical scheme used to obtain $u_{h\tau}$ need 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 et al. [6], or Ainsworth [2].

This paper is structured as follows. We present the continuous and discrete settings and define our error measure $\mathcal{J}_h(u_{h\tau})$ in Section 2. This is a sum of the above-discussed space-time mesh-dependent dual norm $\mathcal{J}_{h, FR}(u_{h\tau})$, see (2.7), and of a weighted jump term $\mathcal{J}_{h, NC}(u_{h\tau})$ which measures possible nonconformities in space, see (2.12). We state our a posteriori error estimate, Theorem 3.3, in Section 3 yielding the error
upper bound $J_u(u_{ht}) \leq \eta_{FR} + \eta_{NC} + \eta_{H}$. Theorem 4.4 of Section 4 then establishes the error lower bound $\eta_{FR} + \eta_{NC} \lesssim J_u(u_{ht})$ 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 $J_{u,FR}(u_{ht})$. 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

\begin{align*}
X &:= L^2(0, t_F; H^1_0(\Omega)), \\
Y &:= \{ \varphi \in L^2(0, t_F; H^1_0(\Omega)); \partial_t \varphi \in L^2(Q); \varphi(\cdot, t_F) = 0 \},
\end{align*}

(2.1a)

(2.1b)

recalling that $\varphi \in L^2(0, t_F; H^1_0(\Omega))$ 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)–(1.1e) 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

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

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 $K \in L^\infty(\mathbb{R}; \mathbb{R}^{d \times d})$ 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

\[ \partial_t b(v) - \nabla \cdot (K \nabla v - \phi(v)) = f, \quad (2.3) \]

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 media 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, equation (2.3) can be recast into the form (1.1a) by setting $u := b(v)$, yielding $K(u) := K(b^{-1})(u)$ and $\phi(u) := \phi(b^{-1}(u))$. 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 $b(v) \not\in X$. 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 $\mathbf{K}(\cdot)$ is $L^2T^{-1}$ 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 other ones. For all \( 1 \leq n \leq N \), we denote by \( \mathcal{T}^{n-1,n} \) the coarsest common refinement of \( \mathcal{T}^{n-1} \) and \( \mathcal{T}^n \) and by \( \mathcal{T}^{n-1,n}_l \) the finest common coarsening of \( \mathcal{T}^{n-1} \) and \( \mathcal{T}^n \). Obviously, if the meshes are kept fixed, \( \mathcal{T}^{n-1,n} = \mathcal{T}^{n-1,n}_0 = \mathcal{T}^n = \mathcal{T}^0 \). In the general case, the mesh \( \mathcal{T}^{n-1,n} \) is needed to handle nonconforming functions in space that are continuous in time, and the mesh \( \mathcal{T}^{n-1,n}_l \) 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 \( u_{h,T} \). We simply assume that \( u_{h,T} \) 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 \( T \in \mathcal{T}^{n-1,n} \), satisfy \( \varphi|_{\mathcal{T}^n} \in L^2(I_n; H^1(T)) \). Functions in \( X_h \) can exhibit jumps across the interfaces of \( \mathcal{T}^{n-1,n} \), 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 \( T \in \mathcal{T}^{n-1,n} \) or \( T \in \mathcal{T}^{n-1,n}_l \). 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 \( \langle \cdot, \cdot \rangle_T \) and \( \langle \cdot, \cdot \rangle_{T \times I_n} \), respectively. We collect in \( F_T \) all the faces of \( T \) and in \( F_T^{\Omega} \) all those that are subsets of \( \Omega \). The set of all faces of the mesh \( \mathcal{T}^n \) is denoted by \( F^n \), a generic mesh face by \( F \), and we use a similar notation for norms and inner products as above. For an interface \( F \), \( n_F \) denotes its unit normal oriented in the sense the jump is evaluated, while, for a mesh element \( T \), \( n_T \) denotes its outward unit normal. Similarly, \( T_F \) stands for all mesh elements sharing a face with the element \( T \), whereas \( T_F^{-1} \) 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 \( C_T := 1/\pi \). Then, for all \( \varphi \in H^1(T \times I_n) \),

\[
\|\varphi - \Pi_0 \varphi\|_{T \times I_n} \leq C_T \left( h_T^2 \|\nabla \varphi\|_{T \times I_n}^2 + (\tau^n)^2 \|\partial_t \varphi\|_{T \times I_n}^2 \right)^{1/2}. \tag{2.4}
\]

**Proof.** Let \( \varphi \in H^1(T \times I_n) \) and, for all \( t \in I_n \), set \( \tilde{\varphi}(t) := |T|^{-1} \int_T \varphi(x, t) \, dx \). 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

\[
\|\varphi - \tilde{\varphi}\|_{T \times I_n}^2 \leq \int_{I_n} \left( C_T^2 h_T^2 \int_T |\nabla \varphi|^2(x, t) \, dx \right) \, dt = C_T^2 h_T^2 \|\nabla \varphi\|_{T \times I_n}^2.
\]

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

\[
\|\tilde{\varphi} - \Pi_0 \varphi\|_{T \times I_n} \leq |T|^{-1/2} (\tau^n)^{1/2} \left( \int_{I_n} \left( \int_T \partial_t \varphi(x, t) \, dx \right)^2 \, dt \right)^{1/2} \leq (\tau^n)^{1/2} \|\partial_t \varphi\|_{T \times I_n}^2.
\]

Collecting the above bounds yields (2.4).

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

$$\|\varphi\|_{Y,T \times I_n} := C_T,n \left( h_T^n \|\nabla \varphi\|_{T \times I_n}^2 + (\tau^n)^2 \|\partial_t \varphi\|_{T \times I_n}^2 \right), \quad \forall 1 \leq n \leq N, \ T \in T^{n-1,n}. \tag{2.5a}$$

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

$\|\cdot\|_Y$ 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

$$C_{T,n} := \left( (\tau^n)^{-2} T + h_T^{-2} C_{\phi,T,n} + h_T^{-2} C_{K,T,n} \right), \tag{2.6}$$

with a weight $C_{\phi,T,n}$ proportional to a length scale times an advective velocity scale and a weight $C_{K,T,n}$ 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

$$J_{u,FR}(u_{h_T}) := \sup_{\varphi \in Y, \|\varphi\|_Y = 1} \int_0^{t_F} \{ (u_{h_T} - u, \partial_t \varphi) + (\sigma(u, \nabla u) - \sigma(u_{h_T}, \nabla u_{h_T}), \nabla \varphi) \} (t) \, dt. \tag{2.7}$$

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

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

Then we infer, owing to (2.2), that

$$J_{u,FR}(u_{h_T}) = \sup_{\varphi \in Y, \|\varphi\|_Y = 1} \langle R(u_{h_T}), \varphi \rangle_{Y',Y}, \tag{2.9}$$

showing that $J_{u,FR}(u_{h_T})$ is a dual norm of the residual of the approximate solution.

The error measure $J_{u,FR}(u_{h_T})$ 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) problem: Find $\psi \in Y$ such that $(\psi, \varphi)_Y = (R(u_{h_T}), \varphi)_Y$ for all $\varphi \in Y$, where $(\cdot, \cdot)_Y$ denotes the inner product corresponding to the $\|\cdot\|_Y$-norm. Then, it is immediate that $J_{u,FR}(u_{h_T}) = \|\psi\|_Y$. However, a computable upper bound on $J_{u,FR}(u_{h_T})$ can be readily derived using the Cauchy–Schwarz inequality, leading to

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

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

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

\textbf{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 $u_{h_T}$ can be measured in the corresponding energy-type norms. For conforming approximations ($u_{h_T} \in X$), 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,

\[
J_{u,\text{NC}}(u_{h\tau}) := \left\{ \sum_{n=1}^{N} \sum_{T \in T_{n\tau}, n} \sum_{F \in F_T} C_{T,n}^{-1} h_T^{-2} C_{K,\phi,T,F,n} \left\| u - u_{h\tau} \right\|_{F \times I_n}^2 \right\}^{\frac{1}{2}},
\]  

(2.12)

where \([\cdot]\) means the jump across interfaces and the actual value at boundary faces and, for \(T \in T_{n\tau}^{n-1,n}\), \(C_{T,n} := C_{T',n}\) where \(T \subset T' \in T_{n\tau}^{n-1,n}\). The measure (2.12) is inspired by the penalty terms used in IPDG methods, cf., e.g., Arnold [4]. Owing to the flux approximation property in Assumption 4.1, the design of \(J_{u,\text{NC}}(u_{h\tau})\) is linked to the way fluxes are equilibrated. Consequently, the positive weights \(C_{K,\phi,T,F,n}\) (physical unit is \(L^2T^{-2}\)) have to be chosen accordingly; we refer to (7.7) for an example. Moreover, the summation over the common refinement \(T_{n\tau}^{n-1,n}\) is linked to continuous-in-time flux reconstruction in the Crank–Nicolson scheme, whereas \(T_{n}\) 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 \(T \in T_{n-1,n}^{n-1,n}\), and all \(F \in F_T\), \([u] = 0\) 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 \(J_{u,\text{FR}}(u_{h\tau})\), \(J_{u,\text{NC}}(u_{h\tau})\) is easily computable. Finally, note that \(J_{u,\text{NC}}(u_{h\tau}) = 0\) 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.

\[
J_u(u_{h\tau}) := J_{u,\text{FR}}(u_{h\tau}) + J_{u,\text{NC}}(u_{h\tau}).
\]  

(2.13)

There holds \(J_u(u_{h\tau}) = 0\) if and only if \(u_{h\tau} = u\). Indeed, if \(u_{h\tau} = u\), \(J_u(u_{h\tau})\) clearly equals 0. Conversely, if \(J_u(u_{h\tau}) = 0\), then \(J_{u,\text{FR}}(u_{h\tau}) = J_{u,\text{NC}}(u_{h\tau}) = 0\). Thus, \(u_{h\tau}\) is in \(X\) and, hence, since the weak solution is uniquely characterized by the property \((R(u), \varphi)_{Y',Y} = 0\) 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 \(H(\text{div}, \Omega)\) 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 \(t_{h\tau} \in L^2(0, t_F; H(\text{div}, \Omega))\), equilibrated in the sense that

\[
(f - \partial_t u_{h\tau} - \nabla t_{h\tau}, 1)_{T \times I_n} = 0 \quad \forall 1 \leq n \leq N, T \in T_{n\tau}^{n-1,n}.
\]  

(3.1)

The function \(t_{h\tau}\), which represents the advective-diffusive flux, has the opposite sign with respect to the function \(\sigma\) of (1.2), usually defined with the same sign as the gradient to yield monotonicity-type properties. Specific constructions of the flux \(t_{h\tau}\) 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], see equation (3.4) therein. 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 $T^{n-1}$ and $T^n$ only, which in particular fits the Crank–Nicolson scheme of Section 7. Local mass conservation on the elements of $T^n$ 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 $T \in \prod_{n=1}^{N} T_{n}$, we define the residual, flux, and nonconformity estimators respectively

$$
\eta^0_{R,T,n} := C_{T,n}^{-1/2} \eta_{\cdot,n} := C_{T,n}^{-1/2} \left[ f - \partial_t u_{h_T} - \nabla \cdot t_{h_T} \right]_{T \times I_n},
$$

$$
\eta^0_{F,T,n} := C_{T,n}^{-1/2} \eta_{\cdot,n} := C_{T,n}^{-1/2} \left[ \sigma(u_{h_T}, \nabla u_{h_T}) + t_{h_T} \right]_{T \times I_n},
$$

$$
\eta^0_{NC,T,n} := \left\{ \sum_{T^n \in T^{n-1}, T^n \subset T} \sum_{F^n \in F_T} C_{T,n}^{-1} h_T^2 C_{K,F,T,n} \left[ \|u_{h_T} \|_{F \times I_n}^2 \right] \right\}^{1/2},
$$

and set $\eta^0_{FR,T} := \eta^0_{F,T} + \eta^0_{NC,T}$. These estimators are local-in-space and in-time. We define their global space-time versions as $\eta^n := \left\{ \sum_{n=1}^{N} \sum_{T \in T^{n-1}} \left( \eta^0_{\cdot,T,n} \right)^2 \right\}^{1/2}$ for $\bullet \in \{F, R, FR, NC\}$. Finally, we define the initial condition estimator as

$$
\eta_C := \left\{ \sum_{n=1}^{N} \sum_{T \in T^{n-1}} \left( \eta^0_{NC,T,n} \right)^2 \right\}^{1/2}, \quad \eta^0_{IC,T} := C_{T,n}^{-1/2} \left( \tau^n \right)^{-1/2} \|u_0 - u_{h_T}(\cdot, 0)\|_T.
$$

Theorem 3.3 (Guaranteed a posteriori error estimate). Let $u \in X$ be the weak solution given by (2.2) and let $u_{h_T} \in X_h$ be arbitrary. Let Assumption 3.1 hold. Let $\eta^0_{R,T,n}, \eta^0_{F,T,n}, \eta^0_{NC,T,n}$, and $\eta_C$ be defined by (3.2)–(3.5). Then,

$$
J_n(u_{h_T}) \leq \eta_{FR} + \eta_{NC} + \eta_C.
$$

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

Remark 3.4 (Interpretation of the error estimators). The flux estimator $\eta^0_{F,T,n}$ is related to the violation of the constitutive relation (1.2), the residual estimator $\eta^0_{R,T,n}$ to the violation of the equilibrium condition (1.1a), while $\eta^0_{NC,T,n}$ and $\eta^0_{IC,T}$ 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 $t_{h_T}$ within the Crank–Nicolson setting of Section 7, $\eta^0_{F,T,n}$ and $\eta^0_{NC,T,n}$ can be viewed as spatial error estimators and $\eta^0_{R,T,n}$ as temporal error estimator.

Remark 3.6 (Initial condition estimator). Whenever the weights $C_{T,n}$ are chosen according to (2.6), there holds $C_{T,n} \geq C_{\bar{T}} \left( \tau^n \right)^{-2}$, whence we readily infer that $\eta_C \leq \|u_{h_T}(\cdot, 0) - u_0\|_\Omega$, that is, $\eta_C$ is upper-bounded by the usual $L^2(\Omega)$-norm of the approximation error on the initial condition. Moreover, for very small time steps, $\eta_C$ tends to $\|u_{h_T}(\cdot, 0) - u_0\|_\Omega$, while for very large time steps, $\eta_C$ becomes very small.

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 $T^{n-1}$; otherwise, a classical data oscillation term
has to be included in the error lower bound. We likewise assume that both \( u_{h,T} \) and \( t_{h,T} \) are piecewise space-time polynomials on \( \mathcal{T}_{n}^{1,n} \). We observe that, because of the nonlinear functions \( K(\cdot) \) and \( \phi(\cdot) \), the flux \( \mathbf{\sigma}(u_{h,T}, \nabla u_{h,T}) \) is not necessarily a piecewise polynomial, even if \( u_{h,T} \) is. Let

\[
\mathbf{\sigma}(u_{h,T}, \nabla u_{h,T})|_{T} := \mathcal{P}_{T}(\mathbf{\sigma}(u_{h,T}, \nabla u_{h,T})|_{T}), \quad \forall 1 \leq n \leq N, \quad T \in \mathcal{T}_{n}^{1,n},
\]

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

\[
\eta_{qd,T}^{n} := C_{T,n}^{-\frac{1}{2}}h_{T}^{-\frac{1}{2}}\|\mathbf{\sigma}(u_{h,T}, \nabla u_{h,T}) - \mathbf{\sigma}(u_{h,T}, \nabla u_{h,T})\|_{T \times I_{n}},
\]

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

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 \( \mathcal{T}_{n}^{1,n} \) 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 refinement of \( \mathcal{T}_{n}^{1,n} \) and \( \mathcal{T}_{n}^{1,n} \)) and \( T' \in \mathcal{T}_{n}^{1,n} \) (the common refinement of \( \mathcal{T}_{n}^{1,n} \) and \( \mathcal{T}_{n}^{1,n} \)), \( T' \) 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

Let, for all \( 1 \leq n \leq N \) and all \( T \in \mathcal{T}_{n}^{1,n} \), \( \eta_{class,T}^{n} \) be given by

\[
\begin{aligned}
&h_{T}\|f - \partial_{t}u_{h,T} + \nabla \cdot (\mathbf{\sigma}(u_{h,T}, \nabla u_{h,T}))\|_{T \times I_{n}} \\
&\quad + \left\{ \sum_{F \in \mathcal{F}_{T}^{n}} h_{F}\|\mathbf{\sigma}(u_{h,T}, \nabla u_{h,T})\|_{F \times I_{n}} \right\}^{2} + \left\{ \sum_{F \in \mathcal{F}_{T}^{n}} C_{K,F,T,F,n}\|\mathbf{u}_{h,T}\|_{F \times I_{n}}^{2} \right\}^{\frac{1}{2}}.
\end{aligned}
\]

The quantity \( \eta_{class,T}^{n} \) 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

\[
\|\mathbf{\sigma}(u_{h,T}, \nabla u_{h,T}) + t_{h,T}\|_{T \times I_{n}}^{2} \lesssim \sum_{T' \in \mathcal{T}_{n}^{1,n}, T' \subset T} (\eta_{class,T'}^{n})^{2} \quad \forall 1 \leq n \leq N, \quad T \in \mathcal{T}_{n}^{1,n}.
\]

An example on 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 \( 0 < \gamma_{qd,T}^{n} \) and require

\[
\max_{T' \in \mathcal{T}_{T}} \eta_{qd,T'}^{n} \leq \gamma_{qd,T}^{n} (\eta_{FR,T}^{n} + \eta_{NC,T}^{n}), \quad \forall 1 \leq n \leq N, \quad T \in \mathcal{T}_{n}^{1,n},
\]

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 \( \gamma_{qd,T}^{n} \) small enough, hold. Let Assumption 4.1 hold. Recall the definition (2.11) of \( \epsilon_{FR,T}^{n} \) and define \( \mathcal{J}_{a,NC,T}(u_{h,T}) \) as \( \eta_{NC,T}^{n} \) in (3.4), with \( \|u_{h,T}\| \) replaced by \( \|u - u_{h,T}\| \). Then,

\[
\eta_{FR,T}^{n} + \eta_{NC,T}^{n} \lesssim \left\{ \sum_{T' \in \mathcal{T}_{T}} (\epsilon_{FR,T'}^{n})^{2} \right\}^{\frac{1}{2}} + \mathcal{J}_{a,NC,T}(u_{h,T}).
\]

**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 \( \epsilon_{FR,T}^{n} \), whose Hilbertian sum provides an upper bound on the error measure \( \mathcal{J}_{a,FR}(u_{h,T}) \), see (2.10), augmented by the jump seminorm \( \mathcal{J}_{a,NC}(u_{h,T}) \).
4.3 Robust global efficiency

As above, in order to control quadrature errors, we introduce a coefficient $0 < \gamma_{qd}$ and require

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

Then we have the following full equivalence result between our estimators and $J_a(u_{h,t})$.

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

$$\eta_{FR} + \eta_{NC} \lesssim J_a(u_{h,t}). \quad (4.7)$$

**Remark 4.5** (Weights $C_{T,n}$). We stress that our robustness results are not influenced by the value assigned.

**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 $C_T := (h_T^2 C_{\phi,T} + h_T^2 C_{K,T})$ with $C_{\phi,T} := h_T \|\phi(u_{h,t})\|_{\infty,T}$, the steady version of the error measure $e_{FR}$ 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., 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 in two steps.

**Lemma 5.1** (Bound on $J_{a,FR}(u_{h,t})$). There holds $J_{a,FR}(u_{h,t}) \leq \eta_{FR} + \eta_{RC}$.

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

$$\int_0^{t_f} \{(f - \partial_t u_{h,t} - \nabla t_{h,t}, \varphi) + (u_{h,t}(-,0) - u_0, \partial_t \varphi) - (\sigma(u_{h,t}, \nabla u_{h,t}) + t_{h,t}, \nabla \varphi)\}(t) dt.$$

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

$$\langle R(u_{h,t}), \varphi \rangle_{Y',Y} = \sum_{n=1}^N \sum_{T \in T^{n-1,n}} \{(f - \partial_t u_{h,t} - \nabla t_{h,t}, \varphi - \Pi_0 \varphi)_{T \times I_n} + (u_{h,t}(-,0) - u_0, \partial_t \varphi)_{T \times I_n} - (\sigma(u_{h,t}, \nabla u_{h,t}) + t_{h,t}, \nabla \varphi)_{T \times I_n}\} \leq \sum_{n=1}^N \sum_{T \in T^{n-1,n}} (\eta^n_{FR,T} + (\eta^n_{FR,T})^2) \|\varphi\|_{Y,T \times I_n},$$

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

**Lemma 5.2** (Bound on $J_{a,NC}(u_{h,t})$). There holds $J_{a,NC}(u_{h,t}) = \eta_{NC}$.

**Proof.** Immediate owing to Lemma 2.4 and (3.4).
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 $T \in T^{n-1,n}$ be fixed. The first two steps of the proof are devoted to bounding the quantity $\eta_{\text{q},T}^{n}$ defined by (4.3), while staying on the common refinement $T^{n-1,n}$.

**Step 1, bound on $C_{T,n}^{-\frac{1}{2}}\|f - \partial_t u_{ht} + \nabla \sigma(u_{ht}, \nabla u_{ht})\|_{T \times I_n}$**. Let us prove that

$$C_{T,n}^{-\frac{1}{2}}\|f - \partial_t u_{ht} + \nabla \sigma(u_{ht}, \nabla u_{ht})\|_{T \times I_n} \lesssim \epsilon_{\text{FR},T}^{n} + \eta_{\text{q},T}^{n}, \quad (6.1)$$

with $\epsilon_{\text{FR},T}^{n}$ defined by (2.11). Set $v_{T,n} := (f - \partial_t u_{ht} + \nabla \sigma(u_{ht}, \nabla u_{ht}))|_{T \times I_n}$. 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 $T^{n-1}$ and $T^n$. Then, by norm equivalence in finite-dimensional spaces, there holds

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

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

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

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

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

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

Thus, using the definition of $v_{T,n}$, (6.2), (2.2), and the Green theorem, yields, with the notation $\sigma_d := \sigma_n - \sigma(u_{ht}, \nabla u_{ht})$,

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

Thus, owing to (6.4),

$$C_{T,n}^{-\frac{1}{2}}\|v_{T,n}\|_{T \times I_n} \lesssim \frac{(u_{ht} - u, \partial_t (\psi_{T,n} v_{T,n}))_{T \times I_n} + (\sigma_d, \nabla (\psi_{T,n} v_{T,n})).}{\|\psi_{T,n} v_{T,n}\|_{Y,T \times I_n}},$$

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

**Step 2, bound on $C_{T,n}^{-\frac{1}{2}}h_T^{-\frac{1}{2}}\|\nabla (u_{ht}, \nabla u_{ht})\|_{F \times I_n}$** for all $F \in \mathcal{F}_{n}^{\text{int}}$. Let $F \in \mathcal{F}_{n}^{\text{int}}$ be fixed. Recalling that $\mathcal{T}_F$ denotes the simplices in $T^{n-1,n}$ sharing the face $F$, let us prove that

$$C_{T,n}^{-\frac{1}{2}}h_T^{-\frac{1}{2}}\|\nabla (u_{ht}, \nabla u_{ht})\|_{F \times I_n} \lesssim \sum_{T \in \mathcal{T}_F} (\epsilon_{\text{FR},T}^{n} + \eta_{\text{q},T}^{n}). \quad (6.6)$$

Set $v_{F,n} := \nabla (u_{ht}, \nabla u_{ht})|_{F \times I_n}$ 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 \( T_F \) by constant values, we also infer the estimate \( \|v_{F,n}\|_{T_F \times I_n} \lesssim h_T^{\frac{1}{2}} \|v_{F,n}\|_{F \times I_n} \). Using these inequalities and proceeding as in step 1, we arrive at

\[
C_{T,n}^{-\frac{1}{2}} h_T^{-\frac{1}{2}} \|v_{F,n}\|_{F \times I_n} \lesssim C_{T,n}^{-\frac{1}{2}} \|f - \partial_t u_{hT} + \nabla \varphi(u_{hT}, \nabla u_{hT})\|_{T_T \times I_n} + (u - u_{hT}, \partial_t (\psi_{F,n}(u_{F,n}))_{T_T \times I_n} - (\sigma_d, \nabla (\psi_{F,n}(u_{F,n})))_{T_T \times I_n}.
\]

Finally, the first term on the right-hand side is bounded using (6.1) for each \( T' \in T_F \), 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 \( \mathcal{T}^{n-1,n} \) be fixed. Combining (6.1) and (6.6) yields

\[
\sum_{T' \in \mathcal{T}^{n-1,n}, T' \subset T} (C_{T,n}^{-\frac{1}{2}} h_T^{-\frac{1}{2}} \eta_{\text{cl},T,n})^2 \lesssim \sum_{T' \in T_T} (C_{FR,T,n}^n + \eta_{\text{qd},T})^2 + (\eta_{NC,T})^2.
\]

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

\[
\eta_{FR,T}^n \lesssim \sum_{T' \in T_T} (C_{FR,T,n}^n + \eta_{\text{qd},T}) + \eta_{NC,T}^n.
\]

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.

\]

**6.2 Proof of Theorem 4.4**

**Proof.** We use the notations \( v_{T,n} \) and \( v_{F,n} \) from the previous proof. It is sufficient to show that

\[
\left\{ \begin{array}{l}
\sum_{n=1}^{N} \sum_{T \in \mathcal{T}^{n-1,n}} C_{T,n}^{-1} \|v_{T,n}\|_{T_T \times I_n}^2 \lesssim \mathcal{J}_u(u_{hT}, \eta_{\text{qd}}), \\
\sum_{n=1}^{N} \sum_{T \in \mathcal{T}^{n-1,n}} \sum_{F \in \mathcal{F}^{n-1,n}} C_{T,n}^{-1} h_T^{-1} \|v_{F,n}\|_{F_T \times I_n}^2 \lesssim \mathcal{J}_u(u_{hT}, \eta_{\text{qd}}),
\end{array} \right.
\]

(6.7a)

(6.7b)

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 \( T \in \mathcal{T}^{n-1,n} \). It follows from (6.5) and the Green theorem that

\[
C_{T,n}^{-1} \|v_{T,n}\|_{T_T \times I_n}^2 \lesssim C_{T,n}^{-1} \left\{ (f - \partial_t u_{hT}, \psi_{T,n} v_{T,n})_{T_T \times I_n} - (\sigma(u_{hT}, \nabla u_{hT}), \nabla (\psi_{T,n} v_{T,n}))_{T_T \times I_n} + (\sigma(u_{hT}, \nabla u_{hT}) - \sigma(u_{hT}, \nabla u_{hT}), \nabla (\psi_{T,n} v_{T,n}))_{T_T \times I_n} \right\}.
\]

Set \( \lambda|_{T_T \times I_n} := C_{T,n}^{-1} \psi_{T,n} v_{T,n} \) and observe that \( \lambda \in Y \). Recall the notation \( R(u_{hT}) \) from (2.8), \( \mathcal{J}_{u,FR}(u_{hT}) \) 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 \( T \in \mathcal{T}^{n-1,n} \) and using the Green theorem, the Cauchy–Schwarz inequality, (6.3a), and \( \|v_{T,n}\|_{\infty,T_T \times I_n} \leq 1 \) yields

\[
\sum_{n=1}^{N} \sum_{T \in \mathcal{T}^{n-1,n}} C_{T,n}^{-1} \|v_{T,n}\|_{T_T \times I_n}^2 \lesssim \langle R(u_{hT}), \lambda \rangle_{Y', Y} + \sum_{n=1}^{N} \sum_{T \in \mathcal{T}^{n-1,n}} \eta_{\text{qd},T}^n C_{T,n}^{-1} \|v_{T,n}\|_{T_T \times I_n} \lesssim \mathcal{J}_{u,FR}(u_{hT}) \|\lambda\|_Y + \eta_{\text{qd}} \left\{ \sum_{n=1}^{N} \sum_{T \in \mathcal{T}^{n-1,n}} C_{T,n}^{-1} \|v_{T,n}\|_{T_T \times I_n}^2 \right\}^{\frac{1}{2}}.
\]

(6.8)
The definition (2.5b) of the $\|\cdot\|_Y$-norm, that of $\lambda$, the fact that we suppose that the ratio $h_T/h_{T'}$, $T \in \mathcal{T}^{n-1,n}$, $T' \subset T$, is bounded, and (6.4) lead to

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

Combining (6.8) and (6.9) proves (6.7a).

## 7 Application to interior-penalty discontinuous Galerkin 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 $t_{h_T}$ and in verifying Assumptions 3.1 and 4.1.

### 7.1 The IPDG method

For all $0 \leq n \leq N$, let $V_h^n := \mathcal{P}_p(T^n)$, $p \geq 1$, 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 $\llbracket \cdot \rrbracket$, we consider the (arithmetic) average operator $\{ \cdot \}$ (conventionally yielding the actual value at boundary faces).

The space-time approximation $u_h^n$ 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 $u_h^n$ as the $L^2$-orthogonal projection of $u_0$ onto $V_h^n$. Then, for all $1 \leq n \leq N$, we look for $u_h^n \in V_h^n$ such that

$$\begin{align*}
(\partial_t u_h^n, v_h) + \frac{1}{2} \sum_{m=1}^{n} & \left\{ \langle \sigma(u_h^n, \nabla u_h^n), \nabla v_h \rangle + \sum_{F \in \mathcal{F}^m} \alpha_{K,F}^m h_T^{-1}(\|u_h^n\|_F, \|v_h\|_F) \\
+ \sum_{F \in \mathcal{F}^m} (H_F(u_h^n, v_h))_F - \sum_{F \in \mathcal{F}^m} (\{ K(u_h^n) \nabla u_h^n \} \cdot n_F, \|v_h\|_F) \\
- \theta & \sum_{F \in \mathcal{F}^m} (\| K(u_h^n) \nabla v_h \|_F, \|v_h\|_F) - ( f^n, v_h ) \right\} = 0 \quad \forall v_h \in V_h^n,
\end{align*} \quad (7.1)$$

where $\alpha_{K,F}^m$ is the penalty coefficient, $\theta \in \{-1,0,1\}$, and $\partial_t u_h^n := (T^n)^{-1}(u_h^n - u_h^{n-1})$; 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

$$\| H_F(v_h) - \{ \phi(v_h) \} \cdot n_F \|_F \lesssim \| \phi'(v_h) \|_{\infty,T} \| v_h \|_F.$$  

An example is the numerical flux of Lax–Friedrichs type, which consists of the centered flux $\{ \phi(\cdot) \} \cdot n_F$ 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 $t_{h_T}$ continuous and piecewise affine in time, with $t_h^m := t_{h_T}(t^n)$, $0 \leq n \leq N$, in the Raviart–Thomas–Nédélec space $\mathcal{RT}_{l}(T^n) := \{ v_h \in \mathbf{H}(\text{div},\Omega); v_h|_T \in \mathcal{RT}_{l}(T) \}$ for all $T \in \mathcal{T}^n$, where $\mathcal{RT}_{l}(T) := [\mathcal{P}_l(T)]^d + x \mathcal{P}_l(T)$. In particular, $v_h \in \mathcal{RT}_{l}(T^n)$ is such that, for all $T \in \mathcal{T}^n$, $\nabla \cdot v_h \in \mathcal{P}_l(T)$, and for all $F \in \mathcal{F}^n$, $v_h|_F \in \mathcal{P}_l(F)$, with continuity across interfaces, cf. Brezzi and Fortin [7]. Following Kim [20] and [13], we set:

**Definition 7.2** (Reconstructed flux $t_{h_T}$). 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 $t_h^n \in \mathcal{RT}_{l}(T^n)$ by setting, for all $T \in \mathcal{T}^n$, all $F \in \mathcal{F}_T$, all $q_h \in \mathcal{P}_l(F)$,
To verify Assumption 3.1, By construction, we have, for 0 \leq n \leq N and all T \in \mathcal{T}^{n-1,n},

\begin{align}
(f - \partial_t u_{h_T} - \nabla \cdot u_{h_T}, v_h)_{T \times I_n} &= 0 \quad \forall v_h \in \mathcal{P}_{\min(p,l)}(T),
\end{align}

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^n \in \mathcal{L}^2(0, t_F; \mathbf{H}(\text{div}, \Omega)). Let 1 \leq n \leq N, T \in \mathcal{T}^{n-1,n}, and v_h \in \mathcal{P}_{\min(p,l)}(T) be given. Using that \partial_t u_{h_T}, f, and \mathbf{t}_h \mathbf{r}_T are affine in time on \mathcal{I}_n, that \mathcal{T} is from the common coarsening of the meshes \mathcal{T}^{n-1} and \mathcal{T}^n, and the Green theorem, we obtain

\begin{align}
(f - \partial_t u_{h_T} - \nabla \cdot u_{h_T}, v_h)_{T \times I_n} &= -\tau^n (\partial_t u_{h_T}, v_h)_T + \frac{\tau^n}{2} \sum_{m=0}^{n-1} (f^m - \nabla \cdot \mathbf{t}_h^m, v_h)_T \\
&= -\tau^n \left( \partial_t u_{h_T} - \frac{1}{2} \sum_{m=0}^{n-1} f^m, v_h \right)_T + \frac{\tau^n}{2} \sum_{m=0}^{n-1} \sum_{T' \in \mathcal{T}^m, T' \subset T} \left( \mathbf{t}_h^m \cdot \nabla v_h \right)_{T'} - \mathbf{t}_h^m \cdot \nabla v_h \right)_{\partial T'}.
\end{align}

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

\begin{align}
(\mathbf{t}_h^n \cdot \mathbf{n}_{T'}, v_h)_{\partial T'} &= -\sum_{F \in \mathcal{F}_{T'}} \mathbf{n}_{T'} \cdot \mathbf{n}_F \left( \mathbf{f} - \nabla \cdot \mathbf{t}_h^m \right)_{\mathcal{I}_n} \mathbf{n}_F + \alpha^n \sum_{m=0}^{n-1} \left( f^m + \nabla \cdot \mathbf{t}_h^m, v_h \right)_F, \\
(\mathbf{t}_h^n, \nabla v_h)_{T'} &= -\left( \mathbf{f} - \nabla \cdot \mathbf{t}_h^m \right)_{\mathcal{I}_n} \mathbf{n}_F + \alpha^n \mathbf{t}_h^m \cdot \nabla v_h \mathbf{n}_F + \left( \mathbf{t}_h^n, \nabla v_h \right)_{\partial T'}.
\end{align}

Extending \mathbf{v}_h by 0 outside \mathcal{T} so that a function in \mathcal{V}_h \mathbf{n} is obtained and using the above identities and the definition (7.1) of the scheme yields.

7.4 Verification of Assumption 4.1

To verify Assumption 4.1, we first specify the space-time piecewise polynomial \( \mathbf{\sigma}(u_h^n, \nabla u_h^n) \). Let 0 \leq n \leq N and let \mathcal{T} \in \mathbb{T}^n. We define \( \mathbf{\sigma}(u_h^n, \nabla u_h^n) \in \mathbf{RTN}_l(T) \) such that, for all \( F \in \mathcal{F}_T, \) all \( q_h \in \mathcal{P}_l(F), \) and all \( r_h \in \mathcal{P}_{l-1}(T')^d, \)

\begin{align}
(\mathbf{\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, \\
(\mathbf{\sigma}(u_h^n, \nabla u_h^n), r_h)_T &= (\sigma(u_h^n, \nabla u_h^n), r_h)_T.
\end{align}

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 \( T \in \mathcal{T}^n, \) \( \mathbf{\sigma}(u_h^n, \nabla u_h^n)|_T \) belongs to \( \mathbf{RTN}_l(T) \) (as \( \mathbf{t}_h^n \) does), but, globally, \( \mathbf{\sigma}(u_h^n, \nabla u_h^n) \notin \mathbf{RTN}_l(T^n) \) because the normal component of \( \mathbf{\sigma}(u_h^n, \nabla u_h^n) \) is in general discontinuous across interfaces. Finally, the space-time function \( \mathbf{\sigma}(u_{hT}, \nabla u_{hT}) \) is taken to be continuous and piecewise affine in time, matching the values \( \mathbf{\sigma}(u_h^n, \nabla u_h^n) \) at \( t^n \) for all 0 \leq n \leq N.
Lemma 7.5 (Flux approximation). Let $t_{h,T}$ be defined by (7.2)–(7.3) with $l \geq 0$ and let $\sigma(u_{h,T}, \nabla u_{h,T})$ be defined by (7.5)–(7.6). For $1 \leq n \leq N$ and $T \in \mathcal{T}^{-1,n}$, define the constants $C_{K, \phi, T/F,n}$ of (2.12) by

$$C_{K, \phi, T/F,n} := (\alpha_{K,F}^n)^2 h_F^{-1} + ||K(u_h^n)||^2_{\infty,T} h_F^{-1} + ||\phi'(u_h^n)||^2_{\infty,T_F} h_F^r.$$

Then, Assumption 4.1 holds.

Proof. Let $0 \leq n \leq N$ and set $v_h := \sigma(u_h^n, \nabla u_h^n) + t_h^n$. Recall that $\sigma(u_h^n, \nabla u_h^n) = K(u_h^n) \nabla u_h^n - \phi(u_h^n)$. Let $T \in \mathcal{T}^n$ and $F \in F_T$. Using (7.2), (7.5) and (7.3), (7.6) yield, respectively,

$$(v_h|_{T} n_F, q_h)_F = ((1 - w_F)\sigma(u_h^n, \nabla u_h^n) n_F + \alpha_{K,F}^n h_F^{-1} [u_h^n] + H_F(u_h^n) - \langle \phi(u_h^n)\rangle n_F, q_h)_F,$$

$$(v_h, r_h)_T \lesssim ||K(u_h^n)||_{\infty,T} ||r_h||_T \sum_{F \in F_T} h_F^{-1} ||u_h^n||_F.$$

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

$$||v_h||_F \lesssim (1 - w_F)||\sigma(u_h^n, \nabla u_h^n) n_F||_F + (\alpha_{K,F}^n h_F^{-1} + ||\phi'(u_h^n)||_{\infty,T_F}) ||u_h^n||_F,$$

$$(v_h, r_h)_T \lesssim ||K(u_h^n)||_{\infty,T} ||r_h||_T \sum_{F \in F_T} h_F^{-1} ||u_h^n||_F.$$

Using these two estimates in the classical bound $||v_h||_F^2 \lesssim \sum_{F \in F_T} h_F ||v_h n_F||_F^2 + \left(\sup_{r_h \in [\mathcal{P}_{l-1}(T)]^d} (v_h, r_h)_T\right)^2$, valid for all $v_h \in \mathcal{RTN}_l(T)$, and owing to (7.7),

$$||\sigma(u_h^n, \nabla u_h^n) + t_h^n||_F^2 \lesssim \sum_{F \in F_T} h_F ||\sigma(u_h^n, \nabla u_h^n) n_F||_F^2 + \sum_{F \in F_T} C_{K, \phi, T/F,n} ||u_h^n||_F^2. \quad (7.8)$$

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

$$||\sigma(u_{h,T}, \nabla u_{h,T}) + t_{h,T}||_{T \times J_n} \lesssim \gamma^n \sum_{m=n-1}^n ||\sigma(u_{h,m}, \nabla u_{h,m}) + t_{h,m}||_T^2$$

$$= \gamma^n \sum_{m=n-1}^n \sum_{T' \in \mathcal{T}^m} \sum_{T' \subset T} ||\sigma(u_{h,m}, \nabla u_{h,m}) + t_{h,m}||_T^2 \lesssim \sum_{T' \in \mathcal{T}^{n-1,n}, T' \subset T} (\eta_{\text{class},T'})^2,$$

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

Remark 7.6 (Choice of the reconstruction degree $l$). A typical choice for the polynomial degree in the flux reconstruction $t_{h,T}$ 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 \( F \in \mathcal{F}^n \), \( H_F(u^n_F) = \phi(u^n_F)\|F\cdot \mathbf{n}_F \) if \( \|\phi'(u^n_F)\|_F \|F\cdot \mathbf{n}_F \) otherwise, where \( v^n_F \) and \( v^n_{\mathcal{R}} \) denote traces of a function \( v^n \) on \( F \) from the direction and the opposite direction of \( \mathbf{n}_F \), respectively. The penalty parameter in (7.1) is chosen as \( c_h^0 := 10p^2\|\mathbf{K}(u^n)\|_{\infty,T_r} \); see Houston et al. [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 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 \((2p+1)\) 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 \( h_{m,p} := h_02^{-m} \), \( m,p \in \{1,2,3\} \), and \( \tau_{m,1} := \tau_02^{-1} \), \( \tau_{m,2} := 0.5\tau_02^{-1} \), \( \tau_{m,3} := 0.75\tau_02^{-1} \), \( m \in \{1,2,3\} \). We experiment on various orders of convergence in the form \( EOC := \frac{\log(E_m,p/E_{m-1,p})}{\log(h_{m,p}/h_{m-1,p})} \), \( m \in \{2,3\} \), 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 \( C_{F,T,n} \) are chosen according to (2.6) with \( C_{F,T,n} := h_{
abla} \|\phi'(u_{IR})\|_{\infty,Q} \) (here, \( h_{\nabla} \) denotes the diameter of \( \Omega \)) and \( C_{K,F,T,n} := (h_{\nabla}/h_T)\|\mathbf{K}(u_{IR})\|_{\infty,Q} \). The first two addends on the right-hand side of (2.6) are of similar size if the Courant numbers \( \tau n/\|\phi'(u_{IR})\|_{\infty,Q}/h_T \) are of order unity, as well as the ratios \( \tau T/\|\phi'(u_{IR})\|_{\infty,Q}/h_{\nabla} \) (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 \( h_T\|\phi'(u_{IR})\|_{\infty,Q}/\|\mathbf{K}(u_{IR})\|_{\infty,Q} \).

To evaluate the error measure \( J_{n,F}(u_{IR}) \), 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 \( u_{IR} \) has been computed. Recalling that \( J_{e}(u_{IR}) = J_{n,F}(u_{IR}) + J_{e,NC}(u_{IR}) \) where \( J_{e,NC}(u_{IR}) \) coincides with the nonconformity estimator \( \eta_{NC} \), we evaluate the effectivity index \( i_{e} := \eta/J_{e}(u_{IR}) \), where \( \eta := \eta_{FR} + \eta_{NC} + \eta_{e} \) to illustrate Theorems 3.3 and 4.4. Moreover, to illustrate Theorem 4.2, we evaluate the effectivity index \( i_{e,FR} := \eta/(\epsilon_{FR} + J_{e,NC}(u_{IR})) \) where \( \epsilon_{FR} \) is the locally computable upper bound on \( J_{n,F}(u_{IR}) \), see (2.10)–(2.11).

### 8.2 Linear diffusion

We consider first a model advection-diffusion problem with \( \Omega = (-1,1) \times (-1,1) \), \( \tau_T = 1 \), and linear diffusive but nonlinear advective flux. In particular, \( \mathbf{K}(u) = \varepsilon \mathbb{I} \) (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

\[
    u(x,y,t) = (1 + \exp(\vartheta(x,y,t)))^{-1}
\]

with \( \vartheta(x,y,t) := \frac{1 + 2x + 4y + 4yt^2}{4\varepsilon} \), 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 \((t_0, \tau_0) = (1/6, 0.05)\) and consider the value \( \varepsilon = 10^{-2} \). Results are reported in Table 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 

To assess robustness with respect to advection dominance and relative size of space and time steps, we consider some different choices for \( \varepsilon \) and \((t_0, \tau_0)\). In Table 2, we compare the effectivity indices \( i_{e} \) and \( i_{e,FR} \) with those of Table 1 (repeated for convenience); for \( m = p = 3 \) and \((t_0, \tau_0) = (1/6, 0.0125)\), 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 \((t_0, \tau_0)\) only mildly modify the effectivity indices.
Table 1: Example (8.1), $\varepsilon = 10^{-2}$, and $(h_0, \tau_0) = (1/6, 0.05)$: errors, estimators, and effectivity indices

| $m$ | $p$ | $J_{h,FR}(h_{m,p})$ | $\eta$ | $\eta_{h} | \eta_{FR}$ | $\eta_{h,FR}$ | $\eta_{k}$ | $\eta_{k,FR}$ | $i_{e}$ | $i_{e,FR}$ |
|-----|-----|---------------------|--------|-------------|--------------|---------|---------------|-----|---------|
| 1   | 1   | 1.56E-02            | 1.04E-04 | 0.81        | 2.21         | 1.22    | 0.34          | 1.85| 0.81    |
| 2   | 1   | 1.17E-02            | 8.63E-04 | 0.84        | 2.18         | 1.22    | 0.30          | 1.71| 0.81    |
| 3   | 1   | 1.02E-02            | 5.43E-04 | 0.84        | 2.18         | 1.22    | 0.30          | 1.25| 0.81    |
|     |     | (0.84)              | (0.84)  | (0.84)      | (0.84)       |         |               |     |         |
| 1   | 2   | 4.91E-03            | 3.38E-04 | 0.84        | 2.18         | 1.22    | 0.30          | 1.53| 0.81    |
| 2   | 2   | 1.74E-03            | 7.86E-04 | 0.84        | 2.18         | 1.22    | 0.30          | 1.75| 0.81    |
| 3   | 2   | 1.33E-05            | 3.18E-05 | 0.84        | 2.18         | 1.22    | 0.30          | 2.54| 0.97    |
|     |     | (4.71)              | (4.71)  | (4.71)      | (4.71)       |         |               |     |         |

8.3 Degenerate diffusion

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

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

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 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 $\Omega = (-6, 6) \times (-6, 6)$, $t_F = 1$, $K(u) = \kappa|u|^{\kappa-1}u$, $\kappa > 1$, and $\phi(u) = 0$. The initial and boundary conditions are chosen so as to yield the so-called Barenblatt solution

$$u(x, y, t) = \left\{ \begin{array}{ll} 
1 & \text{for } t < \tau_m, \\
\frac{1}{t + 1} \left[ 1 - \frac{1 - \kappa + \frac{1}{2} + \frac{1}{\kappa}}{4\kappa^2} \left( t + 1 \right)^{1/\kappa} \right] & \text{for } t \geq \tau_m.
\end{array} \right.$$
Table 3: Example (8.2): errors, estimators, and effectivity indices

<table>
<thead>
<tr>
<th>m</th>
<th>$p_x$</th>
<th>$J_{x,y}(u_{x,y})$</th>
<th>$\eta_p$</th>
<th>$\eta_m$</th>
<th>$\eta_{NC}$</th>
<th>$\eta_C$</th>
<th>$\eta_{q3}$</th>
<th>$\eta$</th>
<th>$t_e$</th>
<th>$t_{p,F,R}$</th>
<th>$t_e$</th>
<th>$t_{p,F,R}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.93E-03</td>
<td>1.06E-02</td>
<td>6.02E-03</td>
<td>2.73E-02</td>
<td>2.31E-02</td>
<td>2.17E-03</td>
<td>6.26E-02</td>
<td>1.76</td>
<td>0.97</td>
<td>1.55</td>
<td>1.02</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>7.39E-03</td>
<td>7.71E-03</td>
<td>5.68E-03</td>
<td>1.62E-02</td>
<td>7.71E-03</td>
<td>1.23E-03</td>
<td>3.66E-02</td>
<td>1.97</td>
<td>0.73</td>
<td>1.56</td>
<td>0.62</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4.59E-03</td>
<td>4.52E-03</td>
<td>4.95E-03</td>
<td>8.33E-03</td>
<td>1.86E-03</td>
<td>5.22E-04</td>
<td>1.89E-02</td>
<td>1.47</td>
<td>1.16</td>
<td>1.35</td>
<td>0.57</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2.62E-03</td>
<td>3.30E-03</td>
<td>5.40E-03</td>
<td>9.33E-03</td>
<td>1.09E-03</td>
<td>2.67E-04</td>
<td>8.34E-03</td>
<td>1.97</td>
<td>0.73</td>
<td>1.56</td>
<td>0.62</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4.11E-03</td>
<td>1.43E-03</td>
<td>1.93E-03</td>
<td>4.22E-03</td>
<td>1.09E-03</td>
<td>2.67E-04</td>
<td>8.34E-03</td>
<td>1.97</td>
<td>0.73</td>
<td>1.56</td>
<td>0.62</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>4.36E-04</td>
<td>6.13E-04</td>
<td>1.84E-04</td>
<td>1.51E-04</td>
<td>1.00E-04</td>
<td>3.06E-03</td>
<td>1.89E-02</td>
<td>1.47</td>
<td>1.16</td>
<td>1.35</td>
<td>0.57</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>6.48E-04</td>
<td>8.81E-04</td>
<td>1.03E-03</td>
<td>3.57E-03</td>
<td>1.89E-02</td>
<td>1.89E-02</td>
<td>6.47E-03</td>
<td>1.47</td>
<td>1.16</td>
<td>1.35</td>
<td>0.57</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1.94E-04</td>
<td>1.45E-04</td>
<td>1.21E-04</td>
<td>1.07E-04</td>
<td>6.39E-05</td>
<td>1.69E-03</td>
<td>1.93E-02</td>
<td>1.21</td>
<td>0.25</td>
<td>1.21</td>
<td>0.25</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4.26E-04</td>
<td>5.63E-04</td>
<td>6.13E-04</td>
<td>1.84E-04</td>
<td>1.51E-04</td>
<td>1.00E-04</td>
<td>3.06E-03</td>
<td>1.89E-02</td>
<td>1.47</td>
<td>1.16</td>
<td>1.35</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>m</th>
<th>$p_x$</th>
<th>$J_{x,y}(u_{x,y})$</th>
<th>$\eta_p$</th>
<th>$\eta_m$</th>
<th>$\eta_{NC}$</th>
<th>$\eta_C$</th>
<th>$\eta_{q3}$</th>
<th>$\eta$</th>
<th>$t_e$</th>
<th>$t_{p,F,R}$</th>
<th>$t_e$</th>
<th>$t_{p,F,R}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>7.90E-03</td>
<td>5.90E-03</td>
<td>1.32E-02</td>
<td>9.10E-03</td>
<td>3.23E-02</td>
<td>7.08E-05</td>
<td>5.88E-02</td>
<td>4.68</td>
<td>0.98</td>
<td>1.24</td>
<td>0.92</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>8.36E-03</td>
<td>4.64E-03</td>
<td>1.71E-02</td>
<td>8.46E-03</td>
<td>1.11E-02</td>
<td>3.99E-05</td>
<td>4.03E-02</td>
<td>2.40</td>
<td>1.46</td>
<td>3.72</td>
<td>1.62</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>8.91E-03</td>
<td>4.38E-03</td>
<td>2.18E-02</td>
<td>9.56E-03</td>
<td>3.44E-03</td>
<td>1.83E-05</td>
<td>3.87E-02</td>
<td>2.09</td>
<td>2.49</td>
<td>3.38</td>
<td>2.68</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>1.09E-03</td>
<td>1.06E-02</td>
<td>1.06E-01</td>
<td>3.12E-02</td>
<td>1.35E-02</td>
<td>1.74E-04</td>
<td>1.61E-01</td>
<td>4.99</td>
<td>4.22</td>
<td>5.13</td>
<td>2.18</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>4.02E-04</td>
<td>8.04E-03</td>
<td>8.12E-02</td>
<td>2.37E-02</td>
<td>5.16E-03</td>
<td>6.40E-05</td>
<td>1.18E-01</td>
<td>4.90</td>
<td>3.89</td>
<td>5.05</td>
<td>3.84</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>1.28E-04</td>
<td>5.22E-03</td>
<td>5.53E-02</td>
<td>1.10E-02</td>
<td>1.69E-03</td>
<td>2.23E-05</td>
<td>7.57E-02</td>
<td>4.84</td>
<td>4.26</td>
<td>4.97</td>
<td>4.30</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>6.53E-04</td>
<td>2.26E-02</td>
<td>3.27E-01</td>
<td>7.58E-02</td>
<td>8.39E-04</td>
<td>1.36E-04</td>
<td>4.33E-01</td>
<td>5.67</td>
<td>5.01</td>
<td>5.67</td>
<td>4.88</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1.78E-04</td>
<td>9.26E-03</td>
<td>1.38E-02</td>
<td>3.14E-02</td>
<td>3.14E-03</td>
<td>3.51E-05</td>
<td>1.82E-01</td>
<td>5.76</td>
<td>5.17</td>
<td>5.78</td>
<td>5.03</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>3.81E-04</td>
<td>3.41E-03</td>
<td>5.08E-02</td>
<td>1.15E-02</td>
<td>1.14E-03</td>
<td>8.89E-06</td>
<td>6.68E-02</td>
<td>5.80</td>
<td>5.21</td>
<td>5.85</td>
<td>5.10</td>
</tr>
</tbody>
</table>

Figure 1: Example (8.3), $\kappa = 4$: distribution of the element error estimate $\eta_T^p$ (left) and of the local error $e_T^p$ (right) at $t = t_F$ for $p = 2$ and $m = 2$
References


