

Bounds of functional outputs for parabolic problems. Part I: Exact bounds of the Discontinuous Galerkin time discretization^{*}

Núria Parés, Pedro Díez and Antonio Huerta

Laboratori de Càlcul Numèric (LaCàN)
Universitat Politècnica de Catalunya
e-mail: nuria.pares,pedro.diez,antonio.huerta@upc.edu

Abstract

Classical implicit residual type error estimators require using an underlying spatial finer mesh to compute bounds for some quantity of interest. Consequently, the bounds obtained are only guaranteed asymptotically that is with respect to the reference solution computed with the fine mesh. Exact bounds, that is bounds guaranteed with respect to the exact solution, are needed to properly certify the accuracy of the results, especially if the meshes are coarse. The paper introduces a procedure to compute strict upper and lower bounds of the error in linear functional outputs of parabolic problems. In this first part, the bounds account for the error associated with the spatial discretization. The error coming from the time marching scheme is therefore assumed to be negligible in front of the spatial error. The time discretization is performed using the discontinuous Galerkin method, both for the primal and adjoint problems. In the error estimation procedure, equilibrated fluxes at interelement edges are calculated using hybridization techniques.

1 Introduction

Simulation-Based Engineering requires accurate and reliable numerical tools. In practice, the engineering design is based on bounds of quantities of interest, expressed as functional outputs of the solution of some boundary value problem. These bounds are supportive of important and critical decisions. Hence, the accuracy of the numerical results has to be certified without any shadow of uncertainty.

Much work has been devoted to develop goal-oriented adaptive strategies and the related error assessment techniques for steady elliptic problems [10, 9, 2, 4]. The standard approach consists on combining upper and lower bounds for the energy norm of the original problem (primal) and of an auxiliary problem (adjoint) associated with the selected output. The bounds of the quantity of interest are readily obtained operating with the energy estimates. These error estimators require using an underlying spatial finer mesh to compute the energy bounds. Thus,

[†]Submitted to Computer Methods in Applied Mechanics and Engineering

the bounds obtained are only guaranteed asymptotically, that is with respect to the reference solution computed with the fine mesh.

Further research has been carried out to obtain exact bounds, that is bounds guaranteed with respect to the exact solution, independently of any underlying reference mesh. The motivation to develop these numerical tools is to certify the accuracy of the solutions of boundary value and/or evolution problems. These techniques, also provide certificates which may be used to check the correctness of the bounds using a simple algorithm independent of the original code used to compute the bounds and the certificates [13, 14, 11, 17].

In the context of transient parabolic problems, the list of related references is much shorter, see for instance [1, 3, 15, 8, 5], and the possibility of producing exact bounds has not been explored in this framework. This paper and the forthcoming second part [12] provide a methodology to obtain computable strict bounds for quantities of interest in the context of parabolic problems. The strategy presented here uses ideas from [8, 14, 6].

Part of the ingredients used in this work are taken from previous references. The series of papers [1, 3, 15] presents the construction of a posteriori error estimates for an appropriate space-time energy-like measure. No piece of information on specific quantity of interest is provided. On the other hand, the strategy presented in [8] allows obtaining asymptotic bounds for the functional outputs, that is with respect to a reference solution and assuming that the error introduced by the time marching scheme is negligible.

The methodology presented here, also assumes that the error introduced by the time marching scheme is negligible. The contribution of this work is to remove the necessity of using an underlying reference mesh. Exact bounds of the spatial error are obtained instead of asymptotic bounds. In fact, the Discontinuous Galerkin method is used for the time integration and therefore the presented methodology produces bounds for linear functional outputs of the Discontinuous Galerkin discretizations of the transient convection-diffusion-reaction equation. Also an h -adaptive strategy is developed based on the information obtained from the bounds.

In the forthcoming paper [12], also the assumption of neglecting the error introduced in the time integration is removed. The computed bounds account also for the error arising from the time discretization. Thus, the obtained bounds are used in an adaptive procedure accounting both for the error in space and time and refining both h and Δt . Both in this first part and in the companion paper polynomial fields are provided which certify the computed bounds.

2 Problem statement

2.1 Model problem

Let us consider the transient convection-reaction-diffusion equation on a domain $\Omega \subset \mathbb{R}^{n_{sd}}$ polygonal for $n_{sd} = 2$, polyhedral for $n_{sd} = 3$, where n_{sd} is the number of spatial dimensions:

$$\begin{aligned} \dot{u} - \nu \Delta u + \boldsymbol{\alpha} \cdot \nabla u + \sigma u &= f & \text{in } \Omega & \text{ for } t > 0, \\ u &= 0 & \text{on } \partial\Omega & \text{ for } t > 0, \\ u &= u_0 & \text{in } \Omega & \text{ for } t = 0. \end{aligned} \tag{1}$$

Here $u = u(\boldsymbol{x}, t)$, \dot{u} denotes $\partial u / \partial t$ and the time interval of interest is $I =]0, T]$. For the sake of simplicity, the presentation concerns only Dirichlet homogeneous boundary conditions, but the methodology is general and it is also applicable to other type of boundary conditions.

In order to introduce the weak variational form of the previous problem it is useful to redefine $u = u(\boldsymbol{x}, t)$ as a time-dependent function $u(t)$ such that, for each $t \in I$, $u(t) \in$

$\mathcal{H}_0^1(\Omega) := \mathcal{V}$, where $\mathcal{H}_0^1(\Omega)$ denotes the standard Sobolev space of functions vanishing on $\partial\Omega$. Schematically:

$$\begin{aligned} u : I &\longrightarrow \mathcal{V} & \text{and} & & u(t) : \Omega &\longrightarrow \mathbb{R} \\ t &\longmapsto u(t) & & & \mathbf{x} &\longmapsto u(t)(\mathbf{x}) =: u(\mathbf{x}, t). \end{aligned}$$

In this case, \dot{u} denotes du/dt . Also, let \mathcal{V}' be the dual space of \mathcal{V} , $\mathcal{V}' = (\mathcal{H}_0^1)' = \mathcal{H}^{-1}(\Omega)$. Then, the weak solution of (1) belongs to the space

$$\mathcal{W} := \{v \in \mathcal{L}^2(I; \mathcal{V}) \text{ such that } \dot{v} \in \mathcal{L}^2(I; \mathcal{V}')\},$$

where $\mathcal{L}^2(I; \mathcal{V})$ (resp. $\mathcal{L}^2(I; \mathcal{V}')$) denotes the Bochner space associated to \mathcal{V} of square-integrable functions from I into \mathcal{V} (resp. \mathcal{V}')

$$\mathcal{L}^2(I; \mathcal{V}) := \{v : I \rightarrow \mathcal{V}, v \text{ is measurable and } \int_0^T \|v(t)\|_{\mathcal{V}}^2 dt < +\infty\},$$

$\|\cdot\|_{\mathcal{V}}$ being the norm associated with \mathcal{V} , $\|\cdot\|_{\mathcal{V}} = \|\cdot\|_1$.

Remark 1 *In fact, if $v \in \mathcal{W}$, then v is almost everywhere equal to a function continuous from $[0, T]$ into $\mathcal{L}^2(\Omega)$ (see [Temam, Ch. III, §1]). That is, given $v \in \mathcal{W}$, for any $t \in I$ $v(t) \in \mathcal{V}$, and moreover v is a.e. a continuous function with respect to time.*

Finally, the following weak (both in space and time) variational form of the convection-reaction-diffusion equation given by (1) is considered: find $u \in \mathcal{W}$ such that

$$A(u, v) = L(v) \quad \forall v \in \mathcal{L}^2(I; \mathcal{V}), \quad (2)$$

for

$$A(w, v) := \int_0^T [\langle \dot{w}, v \rangle + a(t; w, v)] dt + (w^+(0), v^+(0)),$$

and

$$L(v) := \int_0^T \ell(t; v) dt + (u_0, v^+(0)),$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between \mathcal{V}' and \mathcal{V} , (\cdot, \cdot) denotes the $\mathcal{L}^2(\Omega)$ inner product and $v^\pm(t)$ denote the directional limits of the function v

$$v^\pm(t) := \lim_{s \rightarrow 0^\pm} v(t + s). \quad (3)$$

Here, $u_0 \in \mathcal{V}$ and the bilinear form $a(t; \cdot, \cdot)$ and the linear functional $\ell(t; \cdot)$ are

$$a(t; w, v) := \int_\Omega [\nu(t) \nabla w \cdot \nabla v + \boldsymbol{\alpha}(t) \cdot \nabla w v + \sigma(t) w v] d\Omega,$$

and

$$\ell(t; v) := \langle f(t), v \rangle = \int_\Omega f(t) v d\Omega,$$

where $f \in \mathcal{L}^2(I; \mathcal{V}')$ and for each $t \in I$, $\nu(t) \in \mathcal{L}^\infty(\Omega)$ is a strictly positive real coefficient, $\sigma(t) \in \mathcal{L}^\infty(\Omega)$ is a nonnegative real coefficient and $\boldsymbol{\alpha}(t) \in \mathcal{H}(\text{div}; \Omega)$ is a prescribed vector field which is assumed for simplicity to be incompressible, $\nabla \cdot \boldsymbol{\alpha}(t) = 0$. Moreover ν , σ and $\boldsymbol{\alpha}$ are assumed to be sufficiently smooth in time.

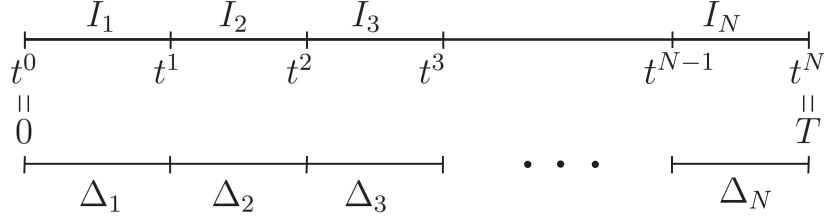


Figure 1: Partition of I into time slabs $I_n =]t^{n-1}, t^n[$ of length $\Delta t_n = t^n - t^{n-1}$.

2.2 Discontinuous Galerkin discretization

Let $0 = t^0 < t^1 < \dots < t^n < \dots < t^N = T$ be a partition of the time interval I into time slabs $I_n =]t^{n-1}, t^n[$, and let $\Delta t_n := t^n - t^{n-1}$ be the length of the time slab I_n (see figure 1).

The discontinuous Galerkin method in time, also denoted as dG(q), seeks a semi-discrete approximation of u, u_τ , which is allowed to be discontinuous at $t^n, n = 1, \dots, N-1$, but which is forced to be polynomial in time (of degree q) inside the time slabs I_n . That is, u_τ belongs to the space

$$\widehat{\mathcal{W}}^\tau := \{v \in \mathcal{L}^2(I; \mathcal{V}), v|_{I_n} \in \mathbb{P}^q(I_n; \mathcal{V}), n = 1, \dots, N\},$$

where, denoting by $N_{nj}(\cdot), j = 0, \dots, q$ the one dimensional Lagrangian shape functions of degree q in the interval I_n ,

$$\mathbb{P}^q(I_n; \mathcal{V}) := \{v : I_n \rightarrow \mathcal{V}, v(t) = \sum_{j=0}^q v_j N_{nj}(t), v_j \in \mathcal{V}\}. \quad (4)$$

The functions $v \in \widehat{\mathcal{W}}^\tau$ may be discontinuous at every $t^n, n = 1, \dots, N-1$ and the jump at $t = t^n$ is defined as $\llbracket v \rrbracket_n := v^+(t^n) - v^-(t^n)$, see (3). Figure 2 shows the time-dependency of the functions in $\mathcal{W}, \widehat{\mathcal{W}}^\tau$ and also in the space $\widehat{\mathcal{W}}$ which is the space where the error $e_\tau := u - u_\tau$ belongs. It is worth noting that both $\mathcal{W} \subset \widehat{\mathcal{W}}$ and $\widehat{\mathcal{W}}^\tau \subset \widehat{\mathcal{W}}$ but $\widehat{\mathcal{W}}^\tau \not\subset \mathcal{W}$.

With these definitions, the discontinuous Galerkin, dG(q), approximation of u is the solution of the weak problem: find $u_\tau \in \widehat{\mathcal{W}}^\tau$ such that

$$A_\tau(u_\tau, v) = L(v) \quad \forall v \in \widehat{\mathcal{W}}^\tau, \quad (5a)$$

where

$$A_\tau(w, v) := A(w, v) + \sum_{n=1}^{N-1} (\llbracket w \rrbracket_n, v^+(t^n)). \quad (5b)$$

It is worth noting that in equations (5), the bilinear form $A(\cdot, \cdot)$ and the linear functional $L(\cdot)$ have been generalized to accept discontinuous-in-time functions in its arguments. Thus, for $w, v \in \widehat{\mathcal{W}}^\tau$,

$$A(w, v) = \sum_{i=1}^N \int_{I_n} [\langle \dot{w}, v \rangle + a(t; w, v)] dt + (w^+(0), v^+(0)),$$

and

$$L(v) = \sum_{i=1}^N \int_{I_n} \ell(t; v) dt + (u_0, v^+(0)),$$

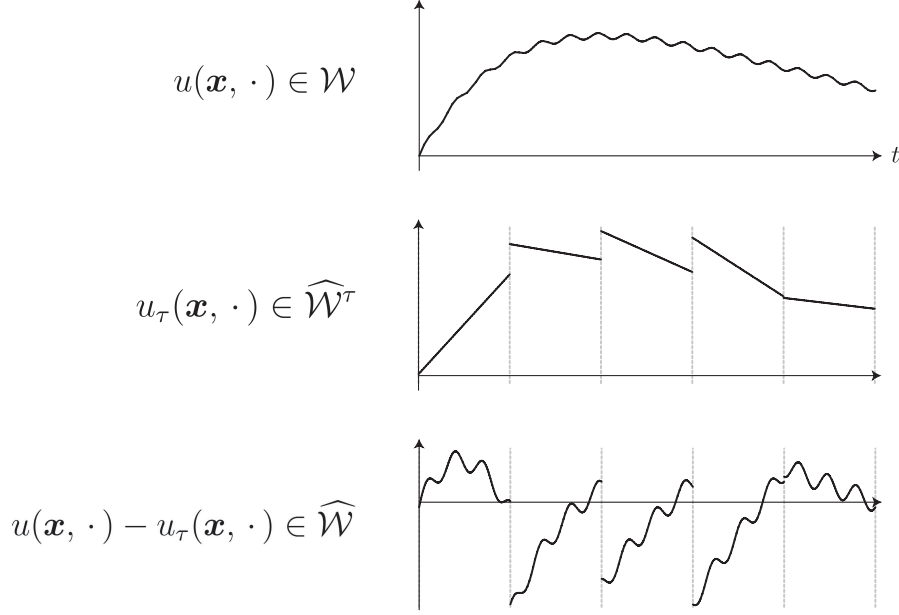


Figure 2: Representation of the time-dependency of the interpolation spaces \mathcal{W} (continuous in time), $\widehat{\mathcal{W}}$ (discontinuous at $t^n, n = 1, \dots, N - 1$) and $\widehat{\mathcal{W}}^\tau$ (piecewise polynomial of degree q in time inside each time slab and discontinuous at $t^n, n = 1, \dots, N - 1$).

where the time differential operator is applied to the space $\widehat{\mathcal{W}}^\tau$ piecewise (inside each time slab I_n) and not in the sense of distributions.

Moreover, equation (5a) decouples into N local-in-time problems posed over the time slabs: for $n = 1, 2, \dots, N$ find $u_\tau|_{I_n} \in \mathbb{P}^q(I_n; \mathcal{V})$ such that

$$\begin{aligned} \int_{I_n} \left[\langle \dot{u}_\tau, v \rangle + a(t; u_\tau, v) \right] dt + (u_\tau^+(t^{n-1}), v^+(t^{n-1})) \\ = \int_{I_n} \ell(t; v) dt + (u_\tau^-(t^{n-1}), v^+(t^{n-1})) \quad \forall v \in \mathbb{P}^q(I_n; \mathcal{V}), \end{aligned} \quad (6)$$

where for $n = 1$ the initial condition of (1) is used, that is $u_\tau^-(0) = u_0$.

However, these local-in-time problems can not be solved exactly since they are posed over an infinite dimensional space. Therefore, in order to find a computable approximation for u , it is also necessary to introduce a finite element mesh of the domain Ω and its associated finite element interpolation space $\mathcal{V}^h \subset \mathcal{V}$. Then, the weak problem for u_τ posed by equation (5a) is approximated by: find $u_{\tau,h} \in \widehat{\mathcal{W}}^{\tau,h}$ such that

$$A_\tau(u_{\tau,h}, v) = L(v) \quad \forall v \in \widehat{\mathcal{W}}^{\tau,h}, \quad (7)$$

where

$$\widehat{\mathcal{W}}^{\tau,h} := \{v \in \mathcal{L}^2(I; \mathcal{V}), v|_{I_n} \in \mathbb{P}^q(I_n; \mathcal{V}^h), n = 1, \dots, N\}.$$

As in equation (6), the previous problem decouples into N problems posed over the time slabs. However, now, the local-in-time problems are computable since $\mathbb{P}^q(I_n; \mathcal{V}^h)$ is a finite-dimensional space of dimension $(q + 1)$ times the dimension of \mathcal{V}^h . To be precise, $u_{\tau,h}$ is computed recursively starting from I_1 and going forward in time (from $n = 1$ to N): in each

time slab I_n , $u_{\tau,h}|_{I_n} \in \mathbb{P}^q(I_n; \mathcal{V}^h)$ is the solution of

$$\begin{aligned} & \int_{I_n} \left[\langle \dot{u}_{\tau,h}, v \rangle + a(t; u_{\tau,h}, v) \right] dt + (u_{\tau,h}^+(t^{n-1}), v^+(t^{n-1})) \\ & = \int_{I_n} \ell(t; v) dt + (u_{\tau,h}^-(t^{n-1}), v^+(t^{n-1})) \quad \forall v \in \mathbb{P}^q(I_n; \mathcal{V}^h), \end{aligned} \quad (8)$$

where for the first slab, I_1 , $u_{\tau,h}^-(0) = u_0$, and for the rest, the solution $u_{\tau,h}^-(t^{n-1})$ computed in the previous slab I_{n-1} is taken as an initial condition for I_n .

The approximation $u_{\tau,h}$ is also denoted as the cG(p)dG(q) approximation of u . It is a standard continuous Galerkin finite element approximation of degree p in space (where p denotes the degree of the complete polynomials used in the interpolation of \mathcal{V}^h) and it is a discontinuous Galerkin approximation of degree q in time.

Remark 2 Note that if $w \in \mathcal{W}$, since $[[w]]_n = 0 \quad \forall n = 1, \dots, N-1$, then $A_\tau(w, v) = A(w, v) \quad \forall v \in \mathcal{L}^2(I; \mathcal{V})$. In particular, the weak problem for u , equation (2), is equivalent to: find $u \in \mathcal{W}$ such that

$$A_\tau(u, v) = L(v) \quad \forall v \in \mathcal{L}^2(I; \mathcal{V}). \quad (9)$$

2.3 Error equations and error decomposition

A posteriori error estimation techniques aim at assessing the accuracy of the approximations of the exact solution. That is, the goal is to evaluate and measure the error $e_{\tau,h} := u - u_{\tau,h}$ either in a specific norm, or in a quantity of interest.

Since $u \in \mathcal{W} \subset \mathcal{L}^2(I; \mathcal{V})$, $u_{\tau,h} \in \widehat{\mathcal{W}}^\tau \subset \mathcal{L}^2(I; \mathcal{V})$ and neither $\widehat{\mathcal{W}}^\tau \subset \mathcal{W}$ nor $\mathcal{W} \subset \widehat{\mathcal{W}}^\tau$, the error belongs to $\mathcal{L}^2(I; \mathcal{V})$. In fact, a smaller space containing both the exact solution u and its cG(p)dG(q) approximation $u_{\tau,h}$ is the space $\widehat{\mathcal{W}}$, where

$$\widehat{\mathcal{W}} := \{v \in \mathcal{L}^2(I; \mathcal{V}) \text{ such that } v|_{I_n} \in \mathcal{L}^2(I_n; \mathcal{V}) \text{ and } \dot{v}|_{I_n} \in \mathcal{L}^2(I_n; \mathcal{V}')\}.$$

That is, $\widehat{\mathcal{W}}$ may be obtained from \mathcal{W} allowing time discontinuities at each time stage t^n , $n = 1, \dots, N-1$ (see figure 2). Therefore, $e_{\tau,h} \in \widehat{\mathcal{W}}$ is continuous inside the time slabs I_n and may present discontinuities at $t = t^n$.

An error equation may be recovered using remark 2, replacing u by $u_{\tau,h} + e_{\tau,h}$ in equation (9) and using the linearity of the first argument of $A_\tau(\cdot, \cdot)$

$$A_\tau(e_{\tau,h}, v) = L(v) - A_\tau(u_{\tau,h}, v) =: R^P(v) \quad \forall v \in \mathcal{L}^2(I; \mathcal{V}),$$

where $R^P(\cdot)$ stands for a weak residual associated to the approximation $u_{\tau,h}$. Moreover, the Galerkin orthogonality condition of the residual,

$$A_\tau(e_{\tau,h}, v) = R^P(v) = 0 \quad \forall v \in \widehat{\mathcal{W}}^{\tau,h}, \quad (10)$$

is verified directly from equation (7).

It will be useful in the following to decompose the error $e_{\tau,h}$ into

$$e_{\tau,h} = u - u_{\tau,h} = (u - u_\tau) + (u_\tau - u_{\tau,h}) = e_\tau + e_h, \quad (11)$$

where $e_\tau := u - u_\tau \in \widehat{\mathcal{W}}$ is the error introduced by the time-discretization of equation (2) using the dG(q) method, and where $e_h := u_\tau - u_{\tau,h} \in \widehat{\mathcal{W}}^\tau$ is the error induced by the space-discretization of equation (5a) using the finite element method, cG(p).

Substituting $u_\tau = u_{\tau,h} + e_h$ into equation (5a), the following weak equation for the spatial error is derived: find $e_h \in \widehat{\mathcal{W}}^\tau$ such that

$$A_\tau(e_h, v) = R^P(v) \quad \forall v \in \widehat{\mathcal{W}}^\tau. \quad (12)$$

Note that thanks to the orthogonality of the residual $R^P(\cdot)$, equation (10), the spatial error e_h is orthogonal to $\widehat{\mathcal{W}}^{\tau,h}$, that is:

$$A_\tau(e_h, v) = 0 \quad \forall v \in \widehat{\mathcal{W}}^{\tau,h}. \quad (13)$$

Remark 3 From now on, in order to simplify the notation $u_{\tau,h}$ will be denoted as u_h . Similarly, the space $\widehat{\mathcal{W}}^{\tau,h}$ will be denoted by $\widehat{\mathcal{W}}^h$. Figure 3 summarizes the notation of the approximations of u and its associated errors.

$$\begin{array}{ccccc}
 \mathcal{W} & \xrightarrow{\text{dG}(q)} & \widehat{\mathcal{W}}^\tau & \supset & \widehat{\mathcal{W}}^h = \widehat{\mathcal{W}}^{\tau,h} \\
 u & & u_\tau & & u_h = u_{\tau,h} \\
 \underbrace{\hspace{10em}} & & \underbrace{\hspace{10em}} & & \\
 e_\tau = u - u_\tau \in \widehat{\mathcal{W}} & & e_h = u_\tau - u_h \in \widehat{\mathcal{W}}^\tau & & \\
 \underbrace{\hspace{10em}} & & \underbrace{\hspace{10em}} & & \\
 e_{\tau,h} = u - u_h \in \widehat{\mathcal{W}} & & & &
 \end{array}$$

Figure 3: Approximations of u and its associated errors.

3 Outputs of interest and adjoint problem

Attention is usually centered in providing upper and lower bounds for quantities of interest depending on the exact solution u . Here, the quantities of interest are restricted to depend linearly on u (this restriction may be relaxed in some problems see [17]) and to be of the form

$$L^\mathcal{O}(u) := \int_0^T \ell^\mathcal{O}(t; u) dt + (u_T^\mathcal{O}, u^-(T)), \quad (14)$$

where $u_T^\mathcal{O} \in \mathcal{V}$ and the linear functional $\ell^\mathcal{O}(t; \cdot)$ can be written as

$$\ell^\mathcal{O}(t; v) := \langle f^\mathcal{O}(t), v \rangle = \int_\Omega f^\mathcal{O}(t)v d\Omega,$$

for $f^\mathcal{O} \in \mathcal{L}^2(I; \mathcal{V}')$. That is, the quantity of interest may depend on the solution at the final time and also it may account for the behavior of the solution along the complete time evolution.

The linearity of the functional $L^\mathcal{O}(\cdot)$ reduces the problem of finding bounds for the output $L^\mathcal{O}(u)$, to the problem of finding bounds for the error committed by approximating $L^\mathcal{O}(u)$ by $L^\mathcal{O}(u_h)$. That is, defining $s := L^\mathcal{O}(u) - L^\mathcal{O}(u_h) = L^\mathcal{O}(e_{\tau,h})$, the obtention of bounds for s , namely

$$s^{lb} \leq s \leq s^{ub},$$

yields the bounds for $L^\mathcal{O}(u)$,

$$L^\mathcal{O}(u_h) + s^{lb} \leq L^\mathcal{O}(u) \leq L^\mathcal{O}(u_h) + s^{ub}.$$

Moreover, the decomposition of the error $e_{\tau,h}$ given in equation (11), yields the natural decomposition of the error in the output s

$$s = L^\mathcal{O}(e_{\tau,h}) = L^\mathcal{O}(e_\tau) + L^\mathcal{O}(e_h) = s_\tau + s_h,$$

where s_τ is the contribution of the time discretization error, and s_h is the contribution of the space discretization error.

Similarly to [8, 3], in this work it is assumed that either the time discretization error has no effect on the error in the output, $s_\tau = L^\mathcal{O}(e_\tau) = 0$, or that at least, the influence of the time discretization to s is negligible in front of the space discretization influence, $s_\tau \ll s_h$, and therefore $s \approx s_h$. The assumption $s_\tau = 0$ holds, for instance, for model problems where the exact solution u is piecewise polynomial in time. In this case, $u = u_\tau$ and therefore since $e_\tau = 0$, $s_\tau = L^\mathcal{O}(e_\tau) = 0$. For more general solutions, it is assumed that the time discretization is enough fine to ensure that the discontinuous Galerkin discretization is much more accurate than the spatial finite element space discretization. Or equivalently that the error introduced by the time marching scheme can be neglected. That is, it is assumed that $e_\tau \ll e_h$ or at least that $s_\tau \ll s_h$. This assumption is removed in the forthcoming paper [12], where the computed bounds account also for the error arising from the time discretization.

The goal of obtaining bounds for the total error in the output s , is modified to obtain bounds for the spatial contribution s_h . Moreover, the obtention of bounds for s_h , namely

$$s_h^{lb} \leq s_h \leq s_h^{ub},$$

yields the bounds for $L^\mathcal{O}(u_\tau)$,

$$L^\mathcal{O}(u_h) + s_h^{lb} \leq L^\mathcal{O}(u_\tau) \leq L^\mathcal{O}(u_h) + s_h^{ub}.$$

In order to derive upper and lower bounds for s_h , an adjoint problem with respect to the selected output is introduced: find $\psi_\tau \in \widehat{\mathcal{W}}^\tau$ such that

$$A_\tau(v, \psi_\tau) = L^\mathcal{O}(v) \quad \forall v \in \widehat{\mathcal{W}}^\tau, \quad (15)$$

together with its corresponding finite element approximation, $\psi_h \in \widehat{\mathcal{W}}^h$

$$A_\tau(v, \psi_h) = L^\mathcal{O}(v) \quad \forall v \in \widehat{\mathcal{W}}^h. \quad (16)$$

Following the same strategy that decomposes the problem defined by (7) into the one posed by (8), the previous problem decouples into N local-in-time problems posed over the time slabs (see appendix A). However now, the adjoint approximation ψ_h is computed recursively starting from I_N and going backward in time (from $n = N$ to $n = 1$): in each time slab I_n , $\psi_h|_{I_n} \in \mathbb{P}^q(I_n; \mathcal{V}^h)$ is the solution of

$$\begin{aligned} & \int_{I_n} \left[-\langle \dot{\psi}_h, v \rangle + a(t; v, \psi_h) \right] dt + (\psi_h^-(t^n), v^-(t^n)) \\ & = \int_{I_n} \ell^\mathcal{O}(t; v) dt + (\psi_h^+(t^n), v^-(t^n)) \quad \forall v \in \mathbb{P}^q(I_n; \mathcal{V}^h), \end{aligned} \quad (17)$$

where for the last slab I_N , $\psi_h^+(T) = u_T^\mathcal{O}$, and for the rest, the solution $\psi_h^+(t^n)$ computed in the previous slab I_{n+1} is taken as a final condition for I_n .

The error in the adjoint solution associated with the approximation ψ_h is $\varepsilon_h := \psi_\tau - \psi_h \in \widehat{\mathcal{W}}^\tau$, and it is such that

$$A_\tau(v, \varepsilon_h) = L^\mathcal{O}(v) - A_\tau(v, \psi_h) =: R^D(v) \quad \forall v \in \widehat{\mathcal{W}}^\tau, \quad (18)$$

where $R^D(\cdot)$ is the weak adjoint residual associated with ψ_h . In particular, for $v \in \widehat{\mathcal{W}}^h$, from equation (16) it follows that

$$A_\tau(v, \varepsilon_h) = R^D(v) = 0 \quad \forall v \in \widehat{\mathcal{W}}^h. \quad (19)$$

4 Output bounds

This section details the obtention of bounds for the error in the output s_h . The bounds, which are computed from the approximations u_h and ψ_h of the primal and adjoint problems (5a) and (15) respectively, are strict regardless of the underlying time partition or finite element meshes used to compute the approximations u_h and ψ_h . Moreover, the bounds are also bounds with respect to the exact error $s = L^\mathcal{O}(e_{\tau,h})$ in the cases where the error in the time discretization does not affect the value of the output, that is, for $s_\tau = L^\mathcal{O}(e_\tau) = 0$. In particular, the bounds are strict for s in the cases where the exact solution u is piecewise polynomial in time, that is, if $u \in \widehat{\mathcal{W}}^\tau$ or equivalently, if $e_\tau = 0$.

The key ingredient to our bound procedure is to relate the error in the quantity of interest, s_h , to the error in the approximations u_h and ψ_h measured in some norm. In this work the norm considered in [8] is selected, but other choices could be used [1, 3]. In order to define the norm in which the error will be measured, the inner product associated to the symmetric counterpart of the bilinear form $a(t; \cdot, \cdot)$ is introduced,

$$a^s(t; w, v) := \frac{1}{2} (a(t; w, v) + a(t; v, w)).$$

For a general convection-reaction-diffusion model problem, the bilinear form $a^s(t; \cdot, \cdot)$ may be rewritten in the form

$$a^s(t; w, v) = \int_\Omega [\nu \nabla w \cdot \nabla v + (\sigma - \nabla \cdot \boldsymbol{\alpha}) w v] \, d\Omega + \frac{1}{2} \int_{\partial\Omega \setminus \Gamma_D} \boldsymbol{\alpha} \cdot \mathbf{n} w v \, d\Gamma,$$

where Γ_D corresponds to the portion of $\partial\Omega$ where Dirichlet boundary conditions are imposed. Recall that for simplicity here $\Gamma_D = \partial\Omega$ and $\nabla \cdot \boldsymbol{\alpha} = 0$ thus,

$$a^s(t; w, v) = \int_\Omega [\nu \nabla w \cdot \nabla v + \sigma w v] \, d\Omega.$$

This inner product induces the norm denoted by $\|\cdot\|$, $\|v\|^2 := a^s(t; v, v) = a(t; v, v)$, and allows to define the space-time norm

$$\| \|v\| \|^2 := \int_0^T \|v\|^2 \, dt,$$

associated to the inner product $\int_0^T a^s(t; \cdot, \cdot) \, dt$. Moreover, the bilinear form $A_\tau(\cdot, \cdot)$ and the space-time norm $\| \|\cdot\| \|$ are related by the following Lemma proved in appendix B. Its proof is trivial given the definition of $A_\tau(\cdot, \cdot)$.

Lemma 1 For any $v \in \widehat{\mathcal{W}}$, $A_\tau(v, v) \geq \|v\|^2$.

The following result shows that bounding s_h reduces to determine upper bounds for errors measured in the space-time norm $\|\cdot\|$.

Theorem 1 Let e_h^s and $\varepsilon_h^s \in \widehat{\mathcal{W}}^\tau$ be such that: for any $v \in \widehat{\mathcal{W}}^\tau$

$$\int_0^T a^s(t; e_h^s, v) dt = R^P(v) \quad \text{and} \quad \int_0^T a^s(t; \varepsilon_h^s, v) dt = R^D(v). \quad (20)$$

Then,

$$-\frac{1}{4} \left\| \left\| \kappa e_h^s - \frac{1}{\kappa} \varepsilon_h^s \right\| \right\|^2 \leq s_h \leq \frac{1}{4} \left\| \left\| \kappa e_h^s + \frac{1}{\kappa} \varepsilon_h^s \right\| \right\|^2,$$

for any nonzero parameter $\kappa \in \mathbb{R}$.

Proof. Combining equation (13) — the primal error e_h is orthogonal to the space $\widehat{\mathcal{W}}^h$ — and equation (18) — definition of the adjoint residual — induces the following representation for the error in the output

$$s_h = L^\mathcal{O}(e_h) = L^\mathcal{O}(e_h) - A_\tau(e_h, \psi_h) = R^D(e_h).$$

Also, taking $v = e_h \in \widehat{\mathcal{W}}^\tau \subset \widehat{\mathcal{W}}$ in equation (12) and using the relation between the bilinear form $A_\tau(\cdot, \cdot)$ and $\|\cdot\|$ given in Lemma 1 it follows that

$$R^P(e_h) = A_\tau(e_h, e_h) \geq \|e_h\|^2.$$

The proof now follows from a simple algebraic manipulation. Indeed, let κ be a nonzero real parameter and consider the obvious inequality

$$\left\| \left\| \frac{1}{2}(\kappa e_h^s \pm \frac{1}{\kappa} \varepsilon_h^s) - \kappa e_h \right\| \right\|^2 \geq 0. \quad (21)$$

Expanding the norm yields

$$\begin{aligned} & \left\| \left\| \frac{1}{2}(\kappa e_h^s \pm \frac{1}{\kappa} \varepsilon_h^s) - \kappa e_h \right\| \right\|^2 \\ &= \frac{1}{4} \left\| \left\| \kappa e_h^s \pm \frac{1}{\kappa} \varepsilon_h^s \right\| \right\|^2 + \kappa^2 \|e_h\|^2 - \kappa \int_0^T a^s(t; \kappa e_h^s \pm \frac{1}{\kappa} \varepsilon_h^s, e_h) dt. \end{aligned} \quad (22)$$

Moreover, using $v = e_h$ in the definition of the primal and dual errors e_h^s and ε_h^s , i.e. equations (20), the last term in the r.h.s. of the previous equality can be rewritten as

$$\begin{aligned} \kappa \int_0^T a^s(t; \kappa e_h^s \pm \frac{1}{\kappa} \varepsilon_h^s, e_h) dt &= \kappa^2 \int_0^T a^s(t; e_h^s, e_h) dt \pm \int_0^T a^s(t; \varepsilon_h^s, e_h) dt \\ &= \kappa^2 R^P(e_h) \pm R^D(e_h) \geq \kappa^2 \|e_h\|^2 \pm s_h. \end{aligned}$$

Finally, joining equation (21) along with equation (22) yields to:

$$0 \leq \frac{1}{4} \left\| \left\| \kappa e_h^s \pm \frac{1}{\kappa} \varepsilon_h^s \right\| \right\|^2 \mp s_h$$

which ends the proof. \square

Theorem 1 reveals that in order to obtain bounds for s_h it is sufficient to compute upper bounds for the norm of the combination of the primal and adjoint symmetric errors e_h^s and ε_h^s , namely

$$-\frac{1}{4} \left\| \left\| \kappa e_h^s - \frac{1}{\kappa} \varepsilon_h^s \right\| \right\|_{\text{UB}}^2 \leq s_h \leq \frac{1}{4} \left\| \left\| \kappa e_h^s + \frac{1}{\kappa} \varepsilon_h^s \right\| \right\|_{\text{UB}}^2, \quad (23)$$

where the subscript UB denotes upper bound. It is worth noting that e_h^s and ε_h^s are the (non-computable) solutions of the problems given by equation (20) which are symmetric (both in space and time) versions of the residual problems posed by equations (12) and (18) respectively.

Next section is devoted to detail the obtention of computable upper bounds for the space-time norm $\|\cdot\|$ of the symmetric errors e_h^s and ε_h^s .

5 Upper bounds for the space-time norm

Consider the auxiliary function $z \in \widehat{\mathcal{W}}^\tau$ solution of

$$\int_0^T a^s(t; z, v) dt = R^*(v) \quad \forall v \in \widehat{\mathcal{W}}^\tau, \quad (24)$$

where $R^*(v) = \alpha R^P(v) + \beta R^D(v)$ for $\alpha, \beta \in \mathbb{R}$. Note that for $\alpha = 1$ and $\beta = 0$, then $R^*(v) = R^P(v)$ and problem (24) is the residual problem for e_h^s . Therefore in this case $z = e_h^s$. Analogously, the choice of $\alpha = 0$ and $\beta = 1$, produces $R^*(v) = R^D(v)$ and the residual problem for ε_h^s is recovered yielding $z = \varepsilon_h^s$. In particular, $\alpha = \kappa$ and $\beta = \pm 1/\kappa$ will be used later to obtain the required upper bounds for $\|\kappa e_h^s \pm 1/\kappa \varepsilon_h^s\|^2$.

The purpose of this section is to establish a procedure to compute upper bounds on $\|z\|^2$. It is worth noting that the model problem under consideration, equation (24), is symmetric both in space and time and that it does not contain derivatives with respect to time.

In order to come up with a computable expression for an upper bound of $\|z\|^2$ the following three steps are considered. First, it is shown that $z \in \widehat{\mathcal{W}}^\tau$ may be computed solving $q + 1$ independent steady diffusion-reaction problems in each slab I_n . Second, for each (infinitely dimensional) steady diffusion-reaction problem posed over the hole domain Ω , a domain decomposition strategy is used to decompose the global problem into n_{el} independent (infinite dimensional) local problems defined over the elements of the mesh (triangles in our case), n_{el} being the number of elements of the underlying spatial mesh. Finally, duality is exploited to transform each local steady problem (posed over an infinite dimensional space) into a feasibility discrete problem which is shown to yield upper bounds for the optimal solution.

5.1 Time decomposition

Let $\nu(t)$, $\alpha(t)$ and $\sigma(t)$ be piecewise constant-in-time functions inside each time slab, that is:

$$\nu(t)|_{I_n} = \nu_n, \quad \alpha(t)|_{I_n} = \alpha_n, \quad \sigma(t)|_{I_n} = \sigma_n,$$

for $\nu_n, \sigma_n \in \mathcal{L}^\infty(\Omega)$, $\alpha_n \in \mathcal{H}(\text{div}; \Omega)$ and $\nabla \cdot \alpha_n = 0$. Working with piecewise constant-in-time parameters has the advantage of avoiding the notational complexity introduced by more complex time dependencies. However, the methodology can be applied to parameters $\nu(t)$, $\alpha(t)$ and $\sigma(t)$ which are piecewise polynomial functions in time.

In this case, the bilinear forms $a(t; \cdot, \cdot)$ and $a^s(t; \cdot, \cdot)$ are piecewise constant-in-time inside the time slabs, that is:

$$a(t; w, v)|_{I_n} = a_n(w, v) \quad \forall w, v \in \mathcal{L}^2(I; \mathcal{V}),$$

for

$$a_n(w, v) = \int_{\Omega} \left[\nu_n \nabla w \cdot \nabla v + \alpha_n \cdot \nabla w v + \sigma_n w v \right] d\Omega,$$

and

$$a^s(t; w, v)|_{I_n} = a_n^s(w, v) = \int_{\Omega} \left[\nu_n \nabla w \cdot \nabla v + \sigma_n w v \right] d\Omega. \quad (25)$$

Equation (24) may be then rewritten as: find $z \in \widehat{\mathcal{W}}^\tau$ such that

$$\sum_{n=1}^N \int_{I_n} a_n^s(z_n, v) dt = R^*(v) \quad \forall v \in \widehat{\mathcal{W}}^\tau,$$

where $z_n := z|_{I_n} \in \mathbb{P}^q(I_n; \mathcal{V})$. In particular, noting that any function $v \in \mathbb{P}^q(I_n; \mathcal{V})$ may be seen as a function of $\widehat{\mathcal{W}}^\tau$ extending v to zero outside the time interval I_n , $z_n \in \mathbb{P}^q(I_n; \mathcal{V})$ is the solution of

$$\int_{I_n} a_n^s(z_n, v) dt = R^*(v) \quad v \in \mathbb{P}^q(I_n; \mathcal{V}). \quad (26)$$

From the definition of the space $\mathbb{P}^q(I_n; \mathcal{V})$ given in (4), it is clear that solving equation (26) for any $v(t) \in \mathbb{P}^q(I_n; \mathcal{V})$ is equivalent to solve it for any $v(t) = v N_{nj}(t)$, $v \in \mathcal{V}$, $j = 0, \dots, q$ which yields the following system of $q + 1$ equations:

$$\int_{I_n} a_n^s(z_n(t), v N_{nj}(t)) dt = R^*(v N_{nj}(t)) \quad \forall v \in \mathcal{V}, j = 0, \dots, q. \quad (27)$$

Note that $z_n(t) \in \mathbb{P}^q(I_n; \mathcal{V})$ is uniquely characterized in time by $q + 1$ independent degrees of freedom. The choice of the degrees of freedom plays an important role in avoiding having to deal with a system of $q + 1$ coupled equations. Consider first the natural degrees of freedom given by the decomposition:

$$z_n(t) = \sum_{i=0}^q z_{ni} N_{ni}(t) \quad z_{ni} \in \mathcal{V}. \quad (28)$$

Defining $t_{ni} := t^{n-1} + i\Delta_n$, $i = 0, \dots, q$, z_{ni} is the value of the function z_n at the time t_{ni} , that is $z_{ni} = z_n(t_{ni})$, see figure 4.

Then, introducing the decomposition given by equation (28) into (27) yields to the following coupled system of $q + 1$ steady problems: find $z_{ni} \in \mathcal{V}$ such that

$$\sum_{i=0}^q \int_{I_n} a_n^s(z_{ni} N_{ni}(t), v N_{nj}(t)) dt = R^*(v N_{nj}(t)) \quad \forall v \in \mathcal{V}, j = 0, \dots, q.$$

Since the functions $N_{nj}(t)$ are constant-in-space and since the bilinear form $a_n^s(\cdot, \cdot)$ is constant-in-time, the l.h.s. of the previous equation is rewritten as:

$$\int_{I_n} a_n^s \left(\sum_{i=0}^q z_{ni} N_{ni}(t), v N_{nj}(t) \right) dt = a_n^s \left(\sum_{i=0}^q c_{ji} z_{ni}, v \right),$$

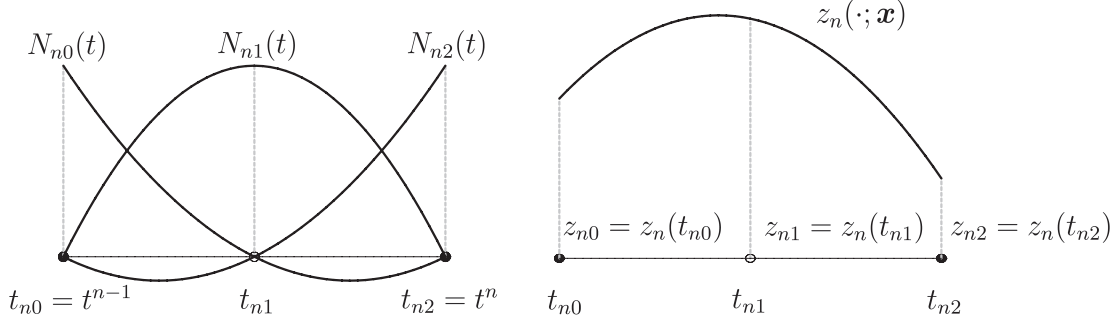


Figure 4: Representation for $q = 2$ of the time dependency of the function z_n .

where $c_{ji} := \int_{I_n} N_{nj}(t)N_{ni}(t) dt$. Therefore, choosing the degrees of freedom of z_n to be

$$Z_{nj} = \sum_{i=0}^q c_{ji}z_{ni} = \sum_{i=0}^q c_{ji}z_n(t_{ni}), \quad (29)$$

reduces the problem of computing z_n , i.e. to find the solution of equation (24), to the determination of $Z_{nj} \in \mathcal{V}, j = 0, \dots, q$ such that

$$a_n^s(Z_{nj}, v) = R^*(vN_{nj}(t)) \quad \forall v \in \mathcal{V}. \quad (30)$$

That is, problem (24) can be reduced to solve for each time slab I_n , $q + 1$ independent steady diffusion-reaction problems. Once the functions Z_{nj} are computed, it is simple to reconstruct the actual time-dependent function $z_n(t)$: first, the estimates $z_{ni} \in \mathcal{V}$ are recovered using equation (29), see remark 4, and then z_n is recovered using equation (28).

Remark 4 Consider the function-vectors

$$\mathbf{Z}_n = (Z_{n0}, Z_{n1}, \dots, Z_{nq}) \quad \text{and} \quad \mathbf{z}_n = (z_{n0}, z_{n1}, \dots, z_{nq}).$$

Then, the transformation given in equation (29) is

$$\mathbf{Z}_n = \mathbf{C} \mathbf{z}_n,$$

where $\mathbf{C} \in \mathcal{M}_{(q+1) \times (q+1)}(\mathbb{R})$. In particular, for $q = 1$ and $q = 2$, the matrix \mathbf{C} is

$$\mathbf{C}_1 = \frac{\Delta t_n}{6} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \quad \text{and} \quad \mathbf{C}_2 = \frac{\Delta t_n}{30} \begin{pmatrix} 4 & 2 & -1 \\ 2 & 16 & 2 \\ -1 & 2 & 4 \end{pmatrix}$$

respectively. Moreover, note that the inverse transformation is

$$\mathbf{z}_n = \mathbf{C}^{-1} \mathbf{Z}_n, \quad (31)$$

where the matrix \mathbf{C}^{-1} may be computed explicitly as

$$\mathbf{C}_1^{-1} = \frac{1}{\Delta t_n} \begin{pmatrix} 4 & -2 \\ -2 & 4 \end{pmatrix}, \quad \mathbf{C}_2^{-1} = \frac{1}{4\Delta t_n} \begin{pmatrix} 36 & -6 & 12 \\ -6 & 9 & -6 \\ 12 & -6 & 36 \end{pmatrix}$$

for $q = 1$ and $q = 2$ respectively.

Remark 5 The r.h.s. of equation (30) for $R^*(\cdot)$ being the primal and adjoint residuals is

$$R^P(vN_{nj}(t)) = \langle f_{nj}, v \rangle - a_n(u_h^{nj}, v), \quad (32)$$

and

$$R^D(vN_{nj}(t)) = \langle f_{nj}^{\mathcal{O}}, v \rangle - a_n(v, \psi_h^{nj}), \quad (33)$$

where

$$f_{nj} = \int_{I_n} \left[f(t)N_{nj}(t) - \dot{u}_h N_{nj}(t) \right] dt - \llbracket u_h \rrbracket_{n-1} N_{nj}(t^{n-1}) \text{ and } u_h^{nj} = \int_{I_n} u_h N_{nj}(t) dt,$$

and

$$f_{nj}^{\mathcal{O}} = \int_{I_n} \left[f^{\mathcal{O}}(t)N_{nj}(t) + \dot{\psi}_h N_{nj}(t) + \llbracket \psi_h \rrbracket_n N_{nj}(t^n) \right] dt \text{ and } \psi_h^{nj} = \int_{I_n} \psi_h N_{nj}(t) dt.$$

For a detailed computation see appendix C.

5.2 Domain decomposition

The steady diffusion-reaction problems given by equation (30) must be solved over the hole domain Ω . This section provides a brief account of the domain decomposition strategy which allows to decompose the steady global problems into solving local spatial problems for each element of the finite element mesh. This approach is standard and it is widely used in a posteriori error estimation for steady problems [7, 10, 2].

Consider a triangulation of the computational domain Ω into n_{el} triangles and denote by Ω_k a general triangle, $k = 1, \dots, n_{\text{el}}$. Equilibrated residual methods compute estimates \hat{Z}_{nj} of Z_{nj} that fulfill equation (30),

$$a_n^s(\hat{Z}_{nj}, v) = a_n^s(Z_{nj}, v) = R^*(vN_{nj}(t)) \quad \forall v \in \mathcal{V}, \quad (34)$$

which are allowed to present discontinuities across the edges of the mesh and which are not forced to verify the Dirichlet boundary conditions. That is, the estimates \hat{Z}_{nj} belong to the ‘‘broken’’ space $\hat{\mathcal{V}}$,

$$\hat{\mathcal{V}} := \{v \in \mathcal{L}^2(\Omega), v|_{\Omega_k} \in \mathcal{H}^1(\Omega_k)\}.$$

In equation (34), the bilinear form $a_n^s(\cdot, \cdot)$ and the residual $R^*(\cdot)$ have been generalized to accept ‘broken’ functions in space, that is, given $v, w \in \hat{\mathcal{V}}$,

$$a_n^s(w, v) = \sum_{k=1}^{n_{\text{el}}} \int_{\Omega_k} \left[\nu_n \nabla w \cdot \nabla v + \sigma_n w v \right] d\Omega =: \sum_{k=1}^{n_{\text{el}}} a_{nk}^s(w, v), \quad (35)$$

and $R^*(v) = \sum_{k=1}^{n_{\text{el}}} R_k^*(v)$ where $R_k^*(\cdot)$ is the restriction of the residual $R^*(\cdot)$ into the element Ω_k .

In fact, equilibrated residual methods compute estimates $\hat{Z}_{nj} \in \hat{\mathcal{V}}$ verifying

$$a_n^s(\hat{Z}_{nj}, v) = R^*(vN_{nj}(t)) + \sum_{\gamma \in \Gamma_h} \int_{\gamma} \lambda_{nj}[v] d\Gamma \quad \forall v \in \hat{\mathcal{V}}. \quad (36)$$

Here Γ_h denotes the set of all edges in the mesh, $[v]_\gamma$ is the jump of the function v along the edge γ if it is an interior edge or $[v]_{\partial\Omega} = v$ for the exterior edges, and $\lambda_{nj} \in \prod_{k=1}^{n_{el}} \mathcal{H}^{-\frac{1}{2}}(\partial\Omega_k)$ are the equilibrated fluxes added to the r.h.s. of equation (36) in order to yield equilibrated local problems in each element. The different existing equilibration techniques differ in the choice of the equilibrated fluxes λ_{nj} which may be computed with an asymptotic complexity that is linear in the number of vertices of the mesh using, for instance, the procedure proposed by Ladeveze in [7].

It is a relatively simple matter to see that the estimates \hat{Z}_{nj} computed from equation (36) verify equation (34). Indeed, for any $v \in \mathcal{V}$, that is, for any v continuous and vanishing on the boundary of the domain, $[v]_\gamma = 0 \forall \gamma \in \Gamma_h$. Therefore, taking $v \in \mathcal{V}$ in equation (36) yields

$$a_n^s(\hat{Z}_{nj}, v) = R^*(vN_{nj}(t)) + \sum_{\gamma \in \Gamma_h} \int_{\gamma} \lambda_{nj}[v] \, d\Gamma = R^*(vN_{nj}(t)),$$

as required in equation (34).

Remark 6 *It is worth noting that when the problem also includes Neumann boundary conditions, Γ_h only contains the interior and the Dirichlet boundaries, so that for $v \in \mathcal{V}$ it is still valid that $[v]_\gamma = 0 \forall \gamma \in \Gamma_h$. Therefore, also in this case the estimates computed from equation (36) verify equation (34).*

Note that for a given choice of the equilibrated fluxes λ_{nj} , the estimate $\hat{Z}_{nj} \in \hat{\mathcal{V}}$ solution of (36) can be computed solving independent problems posed over the elements of the mesh: find $\hat{Z}_{nj,k} := \hat{Z}_{nj}|_{\Omega_k} \in \mathcal{H}^1(\Omega_k)$ such that

$$a_{nk}^s(\hat{Z}_{nj,k}, v) = R_k^*(vN_{nj}(t)) + \int_{\partial\Omega_k} \tau_k \lambda_{nj} v \, d\Gamma \quad \forall v \in \mathcal{H}^1(\Omega_k), \quad (37)$$

where for a given element Ω_k and an arbitrary ordering of the elements of the mesh, τ_k is defined as

$$\tau_k(\mathbf{x}) = \begin{cases} -1 & \mathbf{x} \in \bar{\Omega}_k \cap \bar{\Omega}_l, k < l \\ +1 & \text{otherwise.} \end{cases}$$

It is worth emphasizing that the estimates $\hat{Z}_{nj,k}$ are computed for $n = 1, \dots, N$, $j = 0, \dots, q$ and $k = 1, \dots, n_{el}$, where subscript n stands for the time interval we are considering, I_n , the subscript j stands for the subtime inside the time interval, and the index k stands for the element of the mesh.

Remark 7 *It is tacitly assumed that problems given in equation (37) have at least one solution. For problems with $\sigma_n \neq 0$ the kernel of the local bilinear form $a_{nk}^s(\cdot, \cdot)$ is empty, and therefore, equation (37) has a unique solution. However, for $\sigma_n = 0$, the kernel of $a_{nk}^s(\cdot, \cdot)$ are the constant functions. In this case, the problem is solvable if and only if the following compatibility condition holds:*

$$R_k^*(N_{nj}(t)) + \int_{\partial\Omega_k} \tau_k \lambda_{nj} \, d\Gamma = 0,$$

that is, if the r.h.s. of equation (37) vanishes for $v = 1|_{\Omega_k}$. This previous condition expresses that the boundary data must be in equilibrium with the interior load so that the local problems

are solvable. This is precisely the required condition for the fluxes λ_{nj} to be equilibrated. It is worth noting that the previous conditions do not uniquely determine the equilibrated fluxes λ_{nj} , thus yielding different equilibration techniques. Moreover, the existence of the equilibrated fluxes λ_{nj} is ensured by the orthogonality of the residual $R^*(\cdot)$ with respect to the space $\widehat{\mathcal{W}}^h$, see equations (10) and (19).

5.3 Complementary energy relaxation

The local steady diffusion-reaction problems which define the estimates \hat{Z}_{nj} , equation (37), can not be solved exactly since they are posed over the infinite dimensional space $\mathcal{H}^1(\Omega_k)$. Machiels [8] propose to solve these local problems using an underlying finer mesh, yielding non-guaranteed bounds with respect to $L^{\mathcal{O}}(e_h)$. In order to circumvent this drawback, the proposed approach uses the technique introduced in [13, 14] based on the use of the complementary energy principle, yielding to strict bounds for $L^{\mathcal{O}}(e_h)$.

The key idea is that instead of computing the estimates $\hat{Z}_{nj} \in \mathcal{H}^1(\Omega_k)$ verifying equation (37), the problem is relaxed to determine a pair of dual estimates $\hat{\mathbf{P}}_{nj} \in [\mathcal{L}^2(\Omega_k)]^{\text{nsd}}$ and $\hat{R}_{nj} \in \mathcal{L}^2(\Omega_k)$ such that

$$\int_{\Omega_k} \left[\nu_n \hat{\mathbf{P}}_{nj} \cdot \nabla v + \sigma_n \hat{R}_{nj} v \right] d\Omega = a_{nk}^s(\hat{Z}_{nj}, v) \quad \forall v \in \mathcal{H}^1(\Omega_k) \quad (38)$$

$$k = 1, \dots, n_{\text{el}} \quad j = 0, \dots, q \quad n = 1, \dots, N.$$

It is worth noting that the previous restrictions *do not* determine the dual estimates $\hat{\mathbf{P}}_{nj}$ and \hat{R}_{nj} uniquely.

The essential feature of the method is that if the fields $u_0, u_T^{\mathcal{O}}, f(t)$ and $f^{\mathcal{O}}(t)$ are piecewise polynomial functions in space for each $t \in I$, it is possible to determine — amongst all the dual estimates $\hat{\mathbf{P}}_{nj} \in [\mathcal{L}^2(\Omega_k)]^{\text{nsd}}$ and $\hat{R}_{nj} \in \mathcal{L}^2(\Omega_k)$ verifying equation (38) — two *polynomial* fields $\hat{\mathbf{P}}_{nj}$ and \hat{R}_{nj} verifying equation (38). That is, for a given suitable interpolation degree r , it is possible to find $\hat{\mathbf{P}}_{nj} \in [\mathbb{P}^r(\Omega_k)]^{\text{nsd}}$ and $\hat{R}_{nj} \in \mathbb{P}^r(\Omega_k)$ verifying equation (38). Consequently, the problem of determining the dual polynomial estimates is a discrete solvable problem, see [13, 14].

5.4 Upper bound computation

The procedure to obtain bounds for the time-space norm $\|z\|^2$ can be described in the following five steps:

1. Compute the dual polynomial estimates $\hat{\mathbf{P}}_{nj} \in [\mathbb{P}^r(\Omega_k)]^{\text{nsd}}$, $\hat{R}_{nj} \in \mathbb{P}^r(\Omega_k)$ such that for all $v \in \mathcal{H}^1(\Omega_k)$

$$\int_{\Omega_k} \left[\nu_n \hat{\mathbf{P}}_{nj} \cdot \nabla v + \sigma_n \hat{R}_{nj} v \right] d\Omega = R_k^*(v N_{nj}(t)) + \int_{\partial\Omega_k} \tau_k \lambda_{nj} v d\Gamma. \quad (39)$$

2. Recover the global-in-space piecewise polynomial dual estimates $\hat{\mathbf{P}}_{nj}$ and \hat{R}_{nj} ,

$$\hat{\mathbf{P}}_{nj} = \sum_{k=1}^{n_{\text{el}}} \hat{\mathbf{P}}_{nj} \quad \text{and} \quad \hat{R}_{nj} = \sum_{k=1}^{n_{\text{el}}} \hat{R}_{nj}, \quad (40)$$

where in the previous additions the functions $\hat{\mathbf{P}}_{njk}$ and \hat{R}_{njk} are seen as discontinuous functions vanishing outside the element Ω_k .

3. Compute the dual estimates $\hat{\mathbf{p}}_{ni}$ and \hat{r}_{ni} associated to $\hat{\mathbf{P}}_{nj}$ and \hat{R}_{nj} via the change of variables given in equation (31),

$$\hat{\mathbf{P}}_{nj} = \sum_{i=0}^q c_{ji} \hat{\mathbf{p}}_{ni} \quad \text{and} \quad \hat{R}_{nj} = \sum_{i=0}^q c_{ji} \hat{r}_{ni}. \quad (41)$$

4. Recover the dual time-dependent estimates in the time slab I_n ,

$$\hat{\mathbf{p}}_n = \sum_{i=0}^q \hat{\mathbf{p}}_{ni} N_{ni}(t) \quad \text{and} \quad \hat{r}_n = \sum_{i=0}^q \hat{r}_{ni} N_{ni}(t).$$

5. Evaluate the upper bound for the space-time norm

$$\|z\|^2 \leq \sum_{i=1}^N \int_{I_n} \int_{\Omega} [\nu_n \hat{\mathbf{p}}_n \cdot \hat{\mathbf{p}}_n + \sigma_n \hat{r}_n \hat{r}_n] \, d\Omega \, dt.$$

The following result proves the previous assertion.

Theorem 2 *Let $\hat{\mathbf{P}}_{njk} \in [\mathbb{P}^r(\Omega_k)]^{\text{nsd}}$ and $\hat{R}_{njk} \in \mathbb{P}^r(\Omega_k)$ be a pair of dual estimates verifying equation (39), and consider the global-in-space estimates $\hat{\mathbf{P}}_{nj}$ and \hat{R}_{nj} obtained adding the local fields $\hat{\mathbf{P}}_{njk}$ and \hat{R}_{njk} respectively, see equation (40). Moreover, let $\hat{\mathbf{p}}_{ni}$ and \hat{r}_{ni} be the dual estimates associated to $\hat{\mathbf{P}}_{nj}$ and \hat{R}_{nj} via the change of variables given in equation (31), see equation (41). Finally, consider the time-dependent estimates*

$$\hat{\mathbf{p}}_n = \sum_{i=0}^q \hat{\mathbf{p}}_{ni} N_{ni}(t) \quad \text{and} \quad \hat{r}_n = \sum_{i=0}^q \hat{r}_{ni} N_{ni}(t). \quad (42)$$

Then, an upper bound for the space-time norm of the solution z of the problem given by (24) can be computed as

$$\|z\|^2 = \sum_{i=1}^N \int_{I_n} \|z_n\|^2 \, dt \leq \sum_{i=1}^N \int_{I_n} \int_{\Omega} [\nu_n \hat{\mathbf{p}}_n \cdot \hat{\mathbf{p}}_n + \sigma_n \hat{r}_n \hat{r}_n] \, d\Omega \, dt.$$

Proof. Using the definition of the dual estimates $\hat{\mathbf{P}}_{nj}$ and \hat{R}_{nj} , equation (40), along with equations (38) and (35), it easily follows that for any $v \in \hat{\mathcal{V}}$

$$\begin{aligned} \int_{\Omega} [\nu_n \hat{\mathbf{P}}_{nj} \cdot \nabla v + \sigma_n \hat{R}_{nj} v] \, d\Omega &= \sum_{k=1}^{n_{\text{el}}} \int_{\Omega_k} [\nu_n \hat{\mathbf{P}}_{njk} \cdot \nabla v + \sigma_n \hat{R}_{njk} v] \, d\Omega \\ &= \sum_{k=1}^{n_{\text{el}}} a_{nk}^s(\hat{Z}_{njk}, v) = a_n^s(\hat{Z}_{nj}, v). \end{aligned}$$

In particular, taking $v \in \mathcal{V} \subset \widehat{\mathcal{V}}$ in the previous equation and using equation (34) yields

$$\int_{\Omega} \left[\nu_n \widehat{\mathbf{P}}_{nj} \cdot \nabla v + \sigma_n \widehat{R}_{nj} v \right] d\Omega = a_n^s(\widehat{Z}_{nj}, v) = R^*(v N_{nj}(t)) \quad \forall v \in \mathcal{V}.$$

Now, from the definition of the estimates $\widehat{\mathbf{P}}_{nj}$ and \widehat{R}_{nj} , equation (41), and using both the definition of the constants c_{ij} and that the functions $N_{ni}(t)$ are constant-in-time functions, the r.h.s. of the previous equation may be rewritten as:

$$\begin{aligned} \int_{\Omega} \left[\nu_n \widehat{\mathbf{P}}_{nj} \cdot \nabla v + \sigma_n \widehat{R}_{nj} v \right] d\Omega &= \int_{\Omega} \left[\nu_n \left(\sum_{i=0}^q c_{ji} \widehat{\mathbf{p}}_{ni} \right) \cdot \nabla v + \sigma_n \left(\sum_{i=0}^q c_{ji} \widehat{r}_{ni} \right) v \right] d\Omega \\ &= \int_{I_n} \int_{\Omega} \left[\nu_n \left(\sum_{i=0}^q \widehat{\mathbf{p}}_{ni} N_{ni}(t) \right) \cdot \nabla (v N_{nj}(t)) + \sigma_n \left(\sum_{i=0}^q \widehat{r}_{ni} N_{ni}(t) \right) (v N_{nj}(t)) \right] d\Omega dt \\ &= \int_{I_n} \int_{\Omega} \left[\nu_n \widehat{\mathbf{p}}_n \cdot \nabla (N_{nj}(t) v) + \sigma_n \widehat{r}_n (N_{nj}(t) v) \right] d\Omega dt, \end{aligned}$$

where the definition of the estimates $\widehat{\mathbf{p}}_n$ and \widehat{r}_n , equation (42), has been used in the last equality.

Consequently adding the two previous results it follows that

$$\int_{I_n} \int_{\Omega} \left[\nu_n \widehat{\mathbf{p}}_n \cdot \nabla (v N_{nj}(t)) + \sigma_n \widehat{r}_n (v N_{nj}(t)) \right] d\Omega dt = R^*(v N_{nj}(t))$$

for all $v \in \mathcal{V}$ and $j = 0, \dots, q$. Hence, from the definition of the space $\mathbb{P}^q(I_n; \mathcal{V})$ given by equation (4), it is clear that imposing the previous $q + 1$ equations is equivalent to enforce that

$$\int_{I_n} \int_{\Omega} \left[\nu_n \widehat{\mathbf{p}}_n \cdot \nabla v + \sigma_n \widehat{r}_n v \right] d\Omega dt = R^*(v) \quad \forall v \in \mathbb{P}^q(I_n; \mathcal{V}).$$

Now, taking $v = z_n \in \mathbb{P}^q(I_n; \mathcal{V})$ both in the definition of the function z_n given by equation (26) and in the previous equation one has that

$$\int_{I_n} a_n^s(z_n, z_n) dt = R^*(z_n) = \int_{I_n} \int_{\Omega} \left[\nu_n \widehat{\mathbf{p}}_n \cdot \nabla z_n + \sigma_n \widehat{r}_n z_n \right] d\Omega dt. \quad (43)$$

At this point, the previous equality along with an elementary algebraic manipulations reveal that

$$\int_{I_n} a_n^s(z_n, z_n) dt \leq \int_{I_n} \int_{\Omega} \left[\nu_n \widehat{\mathbf{p}}_n \cdot \widehat{\mathbf{p}}_n + \sigma_n (\widehat{r}_n)^2 \right] d\Omega dt. \quad (44)$$

Indeed, the result is obtained using the obvious inequality

$$\int_{I_n} \int_{\Omega} \left[\nu_n (\widehat{\mathbf{p}}_n - \nabla z_n) \cdot (\widehat{\mathbf{p}}_n - \nabla z_n) + \sigma_n (\widehat{r}_n - z_n)^2 \right] d\Omega dt \geq 0$$

along with the algebraic manipulation

$$\begin{aligned}
& \int_{I_n} \int_{\Omega} \left[\nu_n (\hat{\mathbf{p}}_n - \nabla z_n) \cdot (\hat{\mathbf{p}}_n - \nabla z_n) + \sigma_n (\hat{r}_n - z_n)^2 \right] d\Omega dt \\
&= \int_{I_n} \int_{\Omega} \left[\nu_n \hat{\mathbf{p}}_n \cdot \hat{\mathbf{p}}_n + \sigma_n (\hat{r}_n)^2 \right] d\Omega dt + \int_{I_n} \int_{\Omega} \left[\nu_n \nabla z_n \cdot \nabla z_n + \sigma_n (z_n)^2 \right] d\Omega dt \\
&\quad - 2 \int_{I_n} \int_{\Omega} \left[\nu_n \hat{\mathbf{p}}_n \cdot \nabla z_n + \sigma_n \hat{r}_n z_n \right] d\Omega dt \\
&= \int_{I_n} \int_{\Omega} \left[\nu_n \hat{\mathbf{p}}_n \cdot \hat{\mathbf{p}}_n + \sigma_n (\hat{r}_n)^2 \right] d\Omega dt + \int_{I_n} a_n^s(z_n, z_n) dt - 2 \int_{I_n} a_n^s(z_n, z_n) dt \\
&= \int_{I_n} \int_{\Omega} \left[\nu_n \hat{\mathbf{p}}_n \cdot \hat{\mathbf{p}}_n + \sigma_n (\hat{r}_n)^2 \right] d\Omega dt - \int_{I_n} a_n^s(z_n, z_n) dt,
\end{aligned}$$

where both equation (43) and the definition of the symmetric counterpart of the bilinear form $a_n(\cdot, \cdot)$, $a_n^s(\cdot, \cdot)$, equation (25), have been used.

Finally, using the inequality given by equation (44)

$$\begin{aligned}
\|z\|^2 &= \int_0^T \|z\|^2 dt = \int_0^T a^s(t; z, z) dt = \sum_{n=1}^N \int_{I_n} a^s(t; z, z) dt \\
&= \sum_{n=1}^N \int_{I_n} a_n^s(z_n, z_n) dt \leq \sum_{n=1}^N \int_{I_n} \int_{\Omega} \left[\nu_n \hat{\mathbf{p}}_n \cdot \hat{\mathbf{p}}_n + \sigma_n (\hat{r}_n)^2 \right] d\Omega dt,
\end{aligned}$$

concluding the proof. \square

6 Bounds for the output of interest $L^{\mathcal{O}}(u_{\tau})$: computational aspects

The complete method to derive bounds for $L^{\mathcal{O}}(u_{\tau})$ can be described in the following steps:

1. Compute and store the primal solution u_h recursively starting from I_1 and going forward in time (from $n = 1$ to N), where in each time slab I_n $u_h|_{I_n}$ is the solution of (8). Following reference [16], the computation of $u_h|_{I_n}$ can be decoupled to solve $q + 1$ independent systems of dimension $\dim(\mathcal{V}^h) \times \dim(\mathcal{V}^h)$.
2. For each time slab I_n from $n = N$ to $n = 1$ do:
 - 2.1. Compute and store the dual solution $\bar{\psi}_h|_{I_n}$ solution of (17) which can also be computed solving $q + 1$ independent systems of equations of dimension $\dim(\mathcal{V}^h) \times \dim(\mathcal{V}^h)$.
 - 2.2. For each subtime slab associated to $t_{nj}, j = 0, \dots, q$ do:
 - 2.2.1. Compute and store the primal and adjoint equilibrated fluxes λ_{nj}^P and λ_{nj}^D solution of

$$R^P(vN_{nj}(t)) + \sum_{\gamma \in \Gamma_h} \int_{\gamma} \lambda_{nj}^P[v] d\Gamma = 0 \quad \forall v \in \widehat{\mathcal{V}}^h,$$

$$R^D(vN_{nj}(t)) + \sum_{\gamma \in \Gamma_h} \int_{\gamma} \lambda_{nj}^D[v] d\Gamma = 0 \quad \forall v \in \widehat{\mathcal{V}}^h,$$

where $\widehat{\mathcal{V}}^h$ is the “broken” space associated to \mathcal{V}^h . Reference [2] provides a clear and detailed construction of the equilibrated fluxes.

2.2.2. For each element of the mesh $k = 1, \dots, n_{\text{el}}$ compute the primal and adjoint dual estimates $\hat{\mathbf{P}}_{nj}^P, \hat{\mathbf{P}}_{nj}^D \in [\mathbb{P}^r(\Omega_k)]^{\text{nsd}}$, $\hat{R}_{nj}^P, \hat{R}_{nj}^D \in \mathbb{P}^r(\Omega_k)$ such that for all $v \in \mathcal{H}^1(\Omega_k)$:

$$\int_{\Omega_k} \left[\nu_n \hat{\mathbf{P}}_{nj}^P \cdot \nabla v + \sigma_n \hat{R}_{nj}^P v \right] d\Omega = R_k^P(N_{nj}(t)v) + \int_{\partial\Omega_k} \tau_k \lambda_{nj}^P v d\Gamma, \quad (45a)$$

$$\int_{\Omega_k} \left[\nu_n \hat{\mathbf{P}}_{nj}^D \cdot \nabla v + \sigma_n \hat{R}_{nj}^D v \right] d\Omega = R_k^D(N_{nj}(t)v) + \int_{\partial\Omega_k} \tau_k \lambda_{nj}^D v d\Gamma. \quad (45b)$$

See appendix D and references [13, 14] for a detailed construction of the dual estimates.

2.3 Compute the dual estimates $\hat{\mathbf{p}}_{nik}^P, \hat{\mathbf{p}}_{nik}^D$ and $\hat{r}_{nik}^P, \hat{r}_{nik}^D$ associated to $\hat{\mathbf{P}}_{nj}^P, \hat{\mathbf{P}}_{nj}^D, \hat{R}_{nj}^P, \hat{R}_{nj}^D$ via the change of variables (31),

$$\hat{\mathbf{P}}_{nj}^P = \sum_{i=0}^q c_{ji} \hat{\mathbf{p}}_{nik}^P \quad \text{and} \quad \hat{R}_{nj}^P = \sum_{i=0}^q c_{ji} \hat{r}_{nik}^P,$$

$$\hat{\mathbf{P}}_{nj}^D = \sum_{i=0}^q c_{ji} \hat{\mathbf{p}}_{nik}^D \quad \text{and} \quad \hat{R}_{nj}^D = \sum_{i=0}^q c_{ji} \hat{r}_{nik}^D.$$

2.5 Recover the dual time-dependent estimates in the time slab I_n for each element Ω_k ,

$$\hat{\mathbf{p}}_{nk}^P = \sum_{i=0}^q \hat{\mathbf{p}}_{nik}^P N_{ni}(t) \quad \text{and} \quad \hat{r}_{nk}^P = \sum_{i=0}^q \hat{r}_{nik}^P N_{ni}(t),$$

$$\hat{\mathbf{p}}_{nk}^D = \sum_{i=0}^q \hat{\mathbf{p}}_{nik}^D N_{ni}(t) \quad \text{and} \quad \hat{r}_{nk}^D = \sum_{i=0}^q \hat{r}_{nik}^D N_{ni}(t).$$

2.6 Compute and store the three scalar quantities

$$\eta_n^P := \sum_{k=1}^{n_{\text{el}}} \eta_{nk}^P = \sum_{k=1}^{n_{\text{el}}} \int_{I_n} \int_{\Omega_k} \left[\nu_n \hat{\mathbf{p}}_{nk}^P \cdot \hat{\mathbf{p}}_{nk}^P + \sigma_n \hat{r}_{nk}^P \hat{r}_{nk}^P \right] d\Omega dt,$$

$$\eta_n^D := \sum_{k=1}^{n_{\text{el}}} \eta_{nk}^D = \sum_{k=1}^{n_{\text{el}}} \int_{I_n} \int_{\Omega_k} \left[\nu_n \hat{\mathbf{p}}_{nk}^D \cdot \hat{\mathbf{p}}_{nk}^D + \sigma_n \hat{r}_{nk}^D \hat{r}_{nk}^D \right] d\Omega dt,$$

$$\eta_n^{PD} := \sum_{k=1}^{n_{\text{el}}} \eta_{nk}^{PD} = \sum_{k=1}^{n_{\text{el}}} \int_{I_n} \int_{\Omega_k} \left[\nu_n \hat{\mathbf{p}}_{nk}^P \cdot \hat{\mathbf{p}}_{nk}^D + \sigma_n \hat{r}_{nk}^P \hat{r}_{nk}^D \right] d\Omega dt,$$

and free all the other computed quantities in the time slab I_n .

3. Compute the global quantities

$$\eta^P = \left(\sum_{i=1}^N \eta_n^P \right)^{\frac{1}{2}}, \quad \eta^D = \left(\sum_{i=1}^N \eta_n^D \right)^{\frac{1}{2}} \quad \text{and} \quad \eta^{PD} = \sum_{i=1}^N \eta_n^{PD},$$

and recover the bounds for the output

$$L^{\mathcal{O}}(u_h) - \frac{1}{2}\eta^P\eta^D + \frac{1}{2}\eta^{PD} \leq L^{\mathcal{O}}(u_\tau) \leq L^{\mathcal{O}}(u_h) + \frac{1}{2}\eta^P\eta^D + \frac{1}{2}\eta^{PD}.$$

Remark 8 *The final expression for the bounds for the output $L^{\mathcal{O}}(u_\tau)$ are recovered by means of the following considerations. First theorem 1 states that in order to obtain bounds for the error in the output s_h it is sufficient to obtain upper bounds for the quantities $\left\| \kappa e_h^s \pm \frac{1}{\kappa} \varepsilon_h^s \right\|_{UB}^2$, see equation (23). In order to compute the upper bounds for the space-time norm, the procedure detailed in Section 5 is considered for $z = \kappa e_h^s \pm \frac{1}{\kappa} \varepsilon_h^s$. Then, from theorem 2, the following upper bounds are obtained:*

$$\left\| \kappa e_h^s \pm \frac{1}{\kappa} \varepsilon_h^s \right\|_{UB}^2 = \sum_{i=1}^N \int_{I_n} \int_{\Omega} \left[\nu_n \left(\kappa \hat{\mathbf{p}}_n^P \pm \frac{1}{\kappa} \hat{\mathbf{p}}_n^D \right) \cdot \left(\kappa \hat{\mathbf{p}}_n^P \pm \frac{1}{\kappa} \hat{\mathbf{p}}_n^D \right) + \sigma_n \left(\kappa \hat{r}_n^P \pm \frac{1}{\kappa} \hat{r}_n^D \right)^2 \right] d\Omega dt.$$

Finally the given expressions for the bounds are obtained taking $\kappa^2 = \eta^D / \eta^P$ and rearranging terms.

Remark 9 *It is worth noting that although an adjoint problem has to be solved backwards in time in order to recover the bounds, since the problems for the estimates are decoupled in time (one problem for each time slab I_n), there is no need to store the whole adjoint solution. In fact, once the primal solution u_h is computed and stored, in order to recover the bounds for each time slab one only has to store the three scalar quantities η_n^P , η_n^D and η_n^{PD} . Another option is to store both the primal and adjoint approximations and parallelize the computation of η_n^P , η_n^D and η_n^{PD} , since they may be computed independently for each time slab.*

7 Numerical examples

In the numerical tests presented in this section the upper and lower bounds introduced above are denoted by s^+ and s^- , that is

$$s^- := L^{\mathcal{O}}(u_h) - \frac{1}{2}\eta^P\eta^D + \frac{1}{2}\eta^{PD} \quad \text{and} \quad s^+ := L^{\mathcal{O}}(u_h) + \frac{1}{2}\eta^P\eta^D + \frac{1}{2}\eta^{PD}.$$

In the following, the bound average, $s^{\text{ave}} := (s^+ + s^-)/2$, is taken as a new approximation of the quantity of interest and the half bound gap, $\Delta := (s^+ - s^-)/2$, is seen as an error indicator. The relative counterpart of the bound gap, $\Delta_{\text{rel}} = \Delta/s^{\text{ave}}$, is also used in the presentation.

The meshes are adapted to reduce the half bound gap, Δ . In the examples a simple adaptive strategy is used based on the decomposition of Δ into local positive contributions from the elements:

$$\Delta = \sum_{k=1}^{n_{\text{el}}} \Delta_k$$

where the element contribution to the bound gap Δ_k is

$$\Delta_k := \sum_{i=1}^N \left[\frac{1}{4} \kappa^2 \eta_{nk}^P + \frac{1}{4\kappa^2} \eta_{nk}^D \right].$$

Note that this decomposition is valid because

$$\begin{aligned} \Delta &= \frac{s^+ - s^-}{2} = \frac{1}{2} \eta^P \eta^D = \frac{1}{4} \kappa^2 (\eta^P)^2 + \frac{1}{4\kappa^2} (\eta^D)^2 = \sum_{i=1}^N \left[\frac{1}{4} \kappa^2 \eta_n^P + \frac{1}{4\kappa^2} \eta_n^D \right] \\ &= \sum_{k=1}^{n_{el}} \sum_{i=1}^N \left[\frac{1}{4} \kappa^2 \eta_{nk}^P + \frac{1}{4\kappa^2} \eta_{nk}^D \right] = \sum_{k=1}^{n_{el}} \Delta_k. \end{aligned}$$

The remeshing strategy consists in subdividing, at each step of the adaptive procedure, the elements with the larger values of Δ_k .

7.1 Example 1: uniformly forced square domain

The transient pure diffusion equation ($\nu = 1$, $\sigma = 0$, $\alpha = \mathbf{0}$ in (1)) is solved in the squared domain $\Omega = [0, 1] \times [0, 1]$ and for a final time $T = 0.1$. A constant source term $f(t) = \sqrt{10}$ and homogeneous Dirichlet boundary conditions and initial condition ($u_0 = 0$ in (1)) are considered.

The quantity of interest is an average of the space-time solution

$$L^{\mathcal{O}}(u) = \int_0^T \int_{\Omega} \sqrt{10} u(x, y, t) \, d\Omega \, dt,$$

that is $f^{\mathcal{O}} = \sqrt{10}$ and $u_T^{\mathcal{O}} = 0$ in equation (14). The solution ψ of the adjoint problem is in this case such that $u(t) = \psi(T - t)$.

Two spatial discretizations are used in this test: linear and quadratic triangular elements in space, $p = 1$ and $p = 2$. The time discretization is a piecewise linear interpolation, $q = 1$, with a constant time step $\Delta t = 0.002$, that is $N = 50$. In the computation of the hybrid fluxes, the equilibrated normal fluxes along the edges of the elements are linear, both for $p = 1$ and $p = 2$. The local approximation to \hat{Z}_{njk} and its fluxes in the interior of the elements, $\hat{P}_{njk} \in [\mathbb{P}^r(\Omega_k)]^{\text{nsd}}$ and $\hat{R}_{njk} \in \mathbb{P}^r(\Omega_k)$ are fourth order polynomials, i.e. $r = 4$.

The convergence of the bounds is analyzed for a uniform mesh refinement in a series of structured meshes. The initial mesh is composed by 8 triangular elements (half squares) and in each refinement step every triangle is divided in four similar triangles. The results are displayed in tables 1 and 2 and in figure 5. The evolution of the bounds shows very approximately the optimal finite element convergence rate of $\mathcal{O}(h^p)$, see figures 5 and 6. For quadratic elements ($p = 2$) a slight drop off in the rate of convergence is observed, probably due to the use of linear equilibrated fluxes for λ_{nj}^P and λ_{nj}^D .

In order to check if the error associated with the time integration is indeed negligible, the same cases are performed doubling the number of time steps. All the figures displayed in the tables are identical up to the seventh significant digit. The error introduced by the time marching scheme is therefore negligible in front of the error introduced by the space discretization, which is affecting the third significant digit even for the finer mesh.

n_{el}	linear elements					
	$L^{\mathcal{O}}(u_h)$	s^-	s^+	s^{ave}	Δ	Δ_{rel}
8	0.017253	0.016684	0.022344	0.019514	2.8×10^{-3}	14.5%
32	0.018584	0.018035	0.023163	0.020599	2.6×10^{-3}	12.4%
128	0.019846	0.019685	0.021317	0.020501	8.2×10^{-4}	4.00%
512	0.020295	0.020253	0.020688	0.020470	2.2×10^{-4}	1.55%
2048	0.020419	0.020408	0.020519	0.020463	5.5×10^{-5}	0.26%

Table 1: Example 1: series of uniformly h -refined linear triangular meshes.

n_{el}	quadratic elements					
	$L^{\mathcal{O}}(u_h)$	s^-	s^+	s^{ave}	Δ	Δ_{rel}
8	0.01989196	0.01975772	0.02081192	0.02028482	5.27×10^{-4}	2.60%
32	0.02036561	0.02035690	0.02051254	0.02043472	7.78×10^{-5}	0.38%
128	0.02045199	0.02045133	0.02046651	0.02045892	7.59×10^{-6}	0.04%
512	0.02046074	0.02046069	0.02046196	0.02046132	6.3×10^{-7}	0.003%
2048	0.02046147	0.02046147	0.02046157	0.02046152	5.00×10^{-8}	0.0002%

Table 2: Example 1: series of uniformly h -refined quadratic triangular meshes.

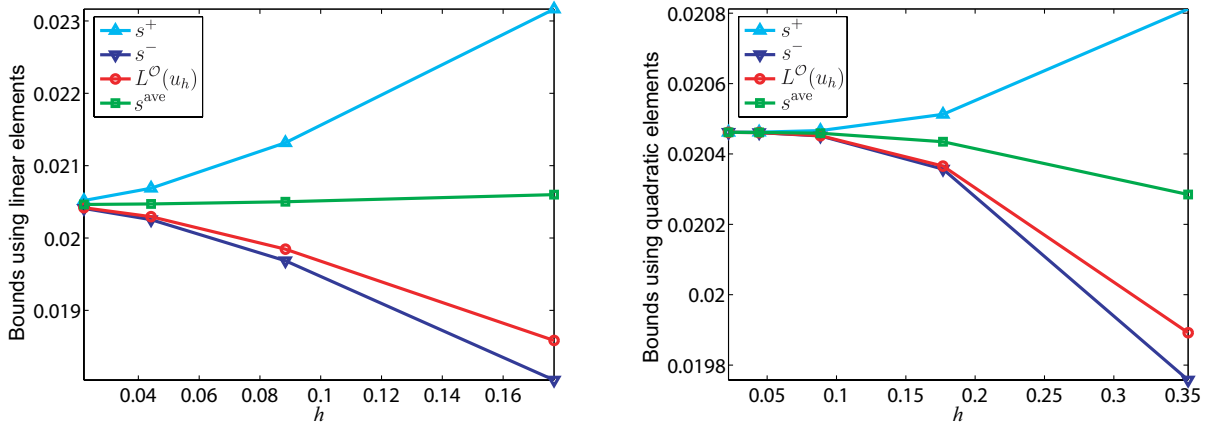


Figure 5: Example 1: convergence of bounds in uniformly h -refined meshes for linear triangles (left) and quadratic triangles (right).

7.2 Example 2: composite material

The unsteady heat equation (1) is solved in the domain described in figure 7 for a non-uniform (composite) material. The problem is purely diffusive ($\sigma = 0$ and $\alpha = \mathbf{0}$ in (1)). The thermal conductivity is smaller in the rectangular inclusions ($\nu = 0.01$) and larger for the bulk material ($\nu = 1$). The specimen is assumed to be thermally isolated on the lateral sides, that is homogeneous Neumann boundary conditions are prescribed in this part of the boundary. The temperature is set to zero on the top ($u = 0$) and a prescribed heat flux is imposed on the

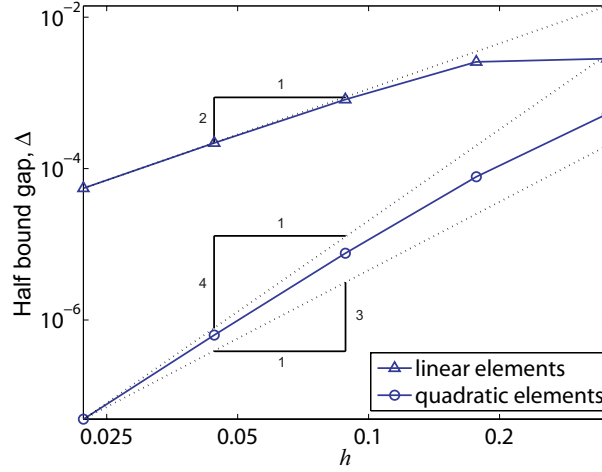


Figure 6: Example 1: convergence of the half bound gap Δ in uniformly h -refined meshes for linear and quadratic triangles.

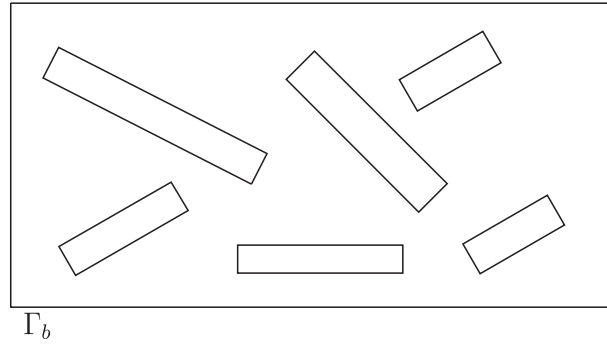


Figure 7: Example 2. Computational domain. Inside the rectangles the thermal conductivity is $\nu = 0.01$, in the rest of the domain $\nu = 1$.

bottom, Γ_b , that is $\nabla u \cdot \mathbf{n} = g_N(t)$, where

$$g_N(t) = \begin{cases} 4t(1-t) + 1 & \text{for } t \in [0, 1] \\ 4(1-t)(2-t) + 1 & \text{for } t \in (1, 2]. \end{cases}$$

The initial thermal state is assumed to be $u(0) = 0$ and the time interval is taken from $t = 0$ to $t = T = 2$.

Here, both space and time discretizations are quadratic, that is $p = 2$ and $q = 2$. The constant time step is $\Delta t_n = 0.1$, which corresponds to $N = 20$. The selected output of interest is the average of the temperature on the bottom, Γ_b :

$$L^{\mathcal{O}}(u) = \frac{1}{\text{meas } \Gamma_b} \int_0^T \int_{\Gamma_b} u(x, y, t) \, d\Gamma \, dt.$$

This quantity of interest is represented by a Neumann boundary condition for the dual problem $g_N^{\mathcal{O}}(t) = 1/\text{meas}(\Gamma_b)$ on Γ_b .

An adaptive procedure is carried out by subdividing the elements with larger valued of Δ_k (larger contributions to the bound gap). The bounds are obtained using a fourth order interpolation for the dual estimates, $r = 4$. The results along the adaptive process are shown in table 3 and in figure 8, where the representative mesh size h is defined as $h = 2\sqrt{1/n_{el}}$. At

n_{el}	$L^{\mathcal{O}}(u_h)$	s^-	s^+	s^{ave}	Δ_{rel}
454	3.19987	3.13879	3.99617	3.56748	12.01%
553	3.20497	3.17302	3.62413	3.39858	6.64%
687	3.20972	3.19091	3.45625	3.32358	3.99%
876	3.21430	3.20177	3.37144	3.28661	2.58%
1102	3.21849	3.21110	3.31225	3.26168	1.55%
1400	3.22169	3.21727	3.27380	3.24553	0.87%
1758	3.22361	3.22108	3.25134	3.23621	0.47%

Table 3: Example 2: series of adaptive h -refinement.

each remeshing step 2% of the elements are marked to be refined. The adaptive procedure is continued until the relative bound gap reaches a target value of 0.5%, that is, $\Delta_{\text{rel}} \leq 0.005$. The convergence is stated with respect to a solution computed with an overkill mesh, $L^{\mathcal{O}}(u_{\tau}) \approx s^{\text{ovk}} = 3.2248603$. Figure 8 shows that the upper and lower bounds and the average approximation reach the expected finite element convergence rate $\mathcal{O}(h^4)$. Note that in this example the lower bound estimates provide better approximation than the upper bound and the bound average.

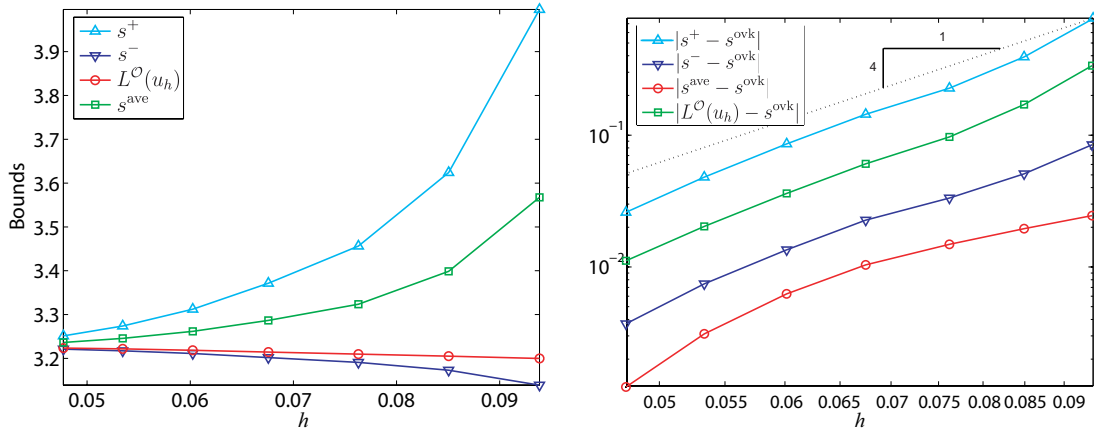


Figure 8: Example 2: Computed bounds for an adaptive h -refinement (left) and its convergence (right).

Using the information provided by the error estimates, in the initial mesh the bounds for the quantity of interest are guarantying that it ranges in an interval of 12.01%, i.e. $L^{\mathcal{O}}(u_{\tau}) = 3.56748 \pm 12.01\%$, while in the final mesh the interval is reduced to 0.47%, namely $L^{\mathcal{O}}(u_{\tau}) = 3.23621 \pm 0.47\%$.

7.3 Example 3: quasi-2D transport

This example is the transient version of a steady quasi-2D transport problem introduced in [14]. The effect of including the convective term is analyzed in this simple problem for different values of the velocity α . Equation (1) is solved in the unit square, $\Omega = [0, 1] \times [0, 1]$, for $\nu = 1$, $\sigma = 1$ and a uniform horizontal velocity field $\alpha = (\alpha, 0)$. The performance of the introduced estimates is tested for different values of α . The boundary conditions are of Dirichlet type on

the lateral sides, homogeneous on the right $u(1, y) = 0$ and set to 1 on the left $u(0, y) = 1$. The boundary condition on both the top and bottom are Neumann homogeneous. The source term is $f = 0$, and the initial condition is $u_0(x, y) = 1 - x$. Time integration is performed to $t = T = 1$ with $\Delta t_n = 0.05$ ($N = 20$). The degrees of the space and time interpolations are $p = 1$, $q = 1$ and $r = 4$.

The quantity of interest is an overall average of the solution, that is

$$L^{\mathcal{O}}(u) = \int_0^T \int_{\Omega} u(x, y, t) \, d\Omega \, dt,$$

which corresponds to $f^{\mathcal{O}} = 1$.

The error estimation strategies and the computation of bounds are performed for a series of uniformly h -refined meshes and different values of α , the results are displayed in table 4 and figure 9. For all the values of α , the rate of convergence of the bound gap is found to be equal to the expected one for the error, that is $\mathcal{O}(h^2)$. It is worth noting that the bound gap is larger as α increases. For $\alpha = 100$ the bound gap is 4 orders of magnitude larger than the for $\alpha = 0$, being the quantity of interest of the same order.

This increment in the bound gap does not correspond to the actual error increment and therefore it has to be concluded that the efficiency of the computed error bounds is deteriorated if the convection parameter is large.

n_{el}	$\alpha = 0$		$\alpha = 1$		$\alpha = 5$		$\alpha = 10$	
	s^{ave}	Δ	s^{ave}	Δ	s^{ave}	Δ	s^{ave}	Δ
32	0.466028	0.001480	0.532580	0.001795	0.738153	0.034904	0.844928	0.168573
128	0.465682	0.000346	0.532864	0.000418	0.739191	0.008002	0.849497	0.038800
512	0.465587	0.000083	0.532940	0.000100	0.739410	0.001854	0.849563	0.008828
2048	0.465562	0.000020	0.532959	0.000024	0.739465	0.000440	0.849580	0.002043
8192	0.465555	0.000005	0.532964	0.000006	0.739479	0.000107	0.849585	0.000485

Table 4: Example 3: results in a series of uniformly h -refined meshes.

7.4 Example 4: rotating transport

Again, a transient version of a steady problem analyzed in [14] is considered. The computational domain is $\Omega = [0, 1] \times [0, 1]$ and the parameters entering in (1) are $\nu = 1$, $\sigma = 10$ and $\alpha = 250(y - \frac{1}{2}, \frac{1}{2} - x)$. The boundary conditions are Dirichlet homogeneous on the whole boundary $\partial\Omega$ and the initial condition is $u_0 = 0$. A localized source term is $f = 100$ in the square $[0.7, 0.8] \times [0.7, 0.8]$ and $f = 0$ elsewhere, see figure 10. The output of interest is a local average in the square region $[0.2, 0.3] \times [0.2, 0.3]$, that is $f^{\mathcal{O}} = 1$ in $[0.2, 0.3] \times [0.2, 0.3]$ and $f^{\mathcal{O}} = 0$ elsewhere.

The parameters describing the space-time discretization are $p = 1$ and $q = 1$, and for a final time $T = 0.03$ a constant time step $\Delta t_n = 0.001$ which corresponds to $N = 30$. A series of adapted meshes is produced by subdividing at each remeshing step 2% of the elements, those with the larger contributions to the bound gap. The initial mesh of 322 elements certifies a wide interval for the quantity of interest, $L^{\mathcal{O}}(u_{\tau}) = 3.76946 \cdot 10^{-5} \pm 7.74669 \cdot 10^{-5}$, after

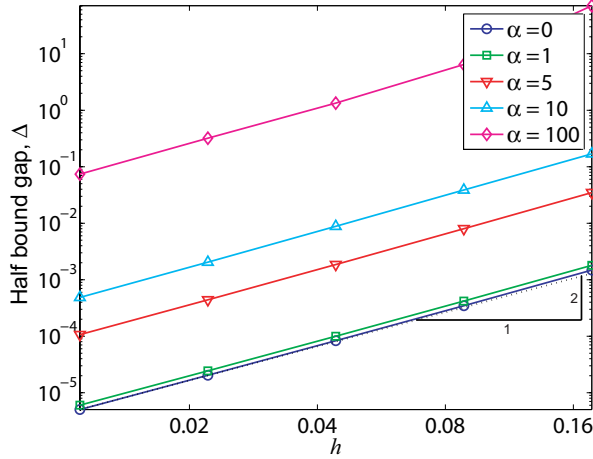


Figure 9: Example 3: convergence of the bound gap for different values of the convection parameter, $\alpha = 0, 1, 5, 10$ and 100 .

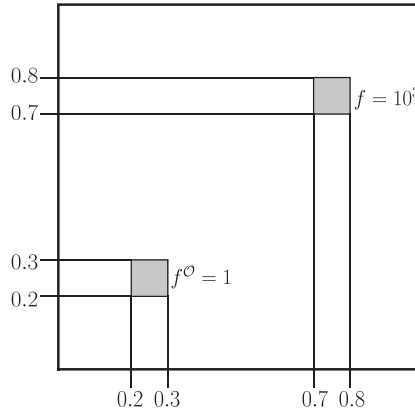


Figure 10: Example 4: Rotating transport forcing and output regions.

remeshing the bounds associated with the final mesh of 4238 elements set a much narrower interval, $L^{\mathcal{O}}(u_{\tau}) = 3.79472 \cdot 10^{-5} \pm 0.71541 \cdot 10^{-5}$, see table 5. The primal and adjoint solutions at the final computational times ($t = T$ for the primal and $t = 0$ for the adjoint) are displayed in figure 11.

n_{el}	$L^{\mathcal{O}}(u_h)$	s^-	s^+	s^{ave}	Δ	Δ_{rel}
322	$3.673 \cdot 10^{-5}$	$-3.977 \cdot 10^{-5}$	$11.516 \cdot 10^{-5}$	$3.769 \cdot 10^{-5}$	$7.747 \cdot 10^{-5}$	205.51%
4238	$3.791 \cdot 10^{-5}$	$3.079 \cdot 10^{-5}$	$4.510 \cdot 10^{-5}$	$3.795 \cdot 10^{-5}$	$0.715 \cdot 10^{-5}$	18.85%

Table 5: Example 4: bounds for the initial and final meshes of the adaptive h -refinement procedure.

The local elementary contributions Δ_k to the global bound gap are plotted in figure 12 for the initial mesh and for an intermediate mesh of the adaptive procedure. The larger values of the local contributions are precisely in the zones where either the primal or the adjoint solutions have larger gradients. Also in figure 12 the resulting final mesh is displayed.

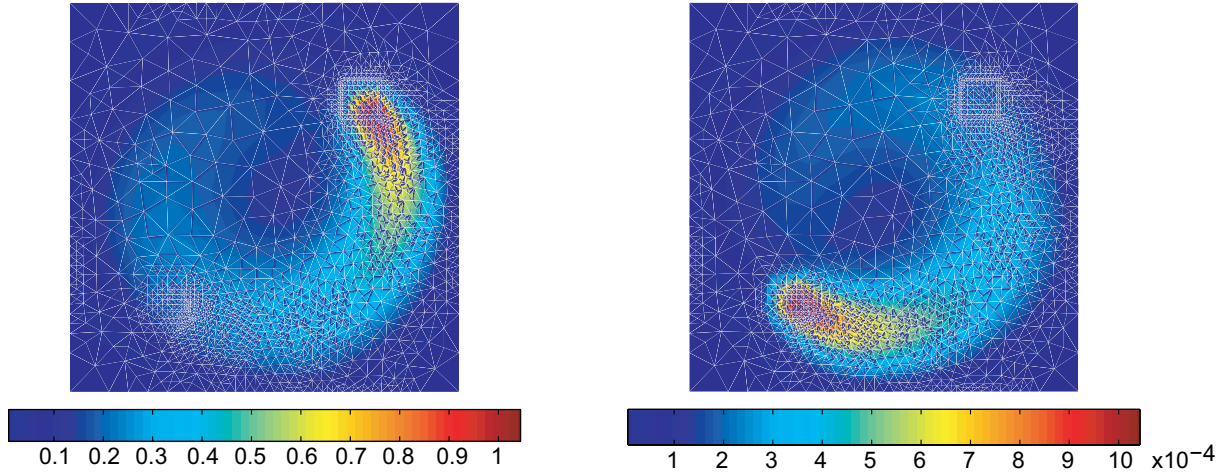


Figure 11: Example 4: Primal and adjoint solution for the final mesh at the time $t = T$ for the primal and $t = 0$ for the adjoint.

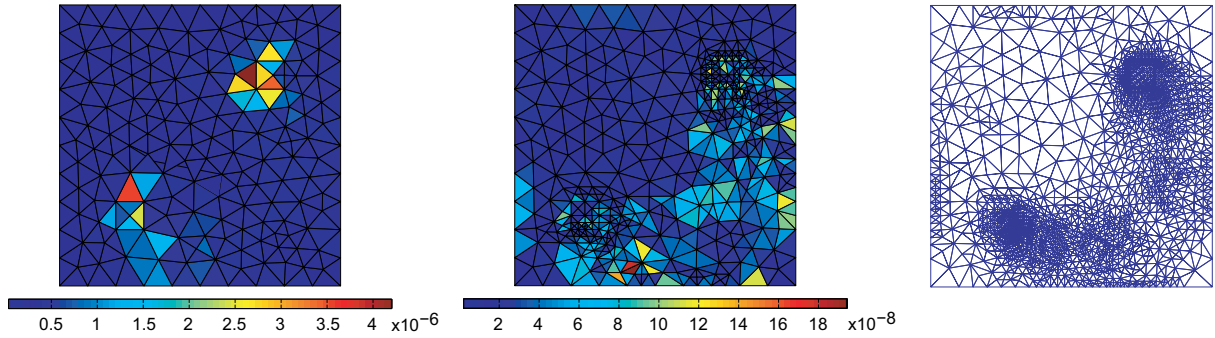


Figure 12: Example 4: Elementary contributions to bound gap in initial mesh (left) and intermediate mesh with $n_{el} = 793$ (center). Final adapted mesh (right).

7.5 Example 5: canister

The final example represents the transport of pollutant inside an active carbon filter. Simulating accurately the behavior of this kind of devices, also named canisters, is of practical importance for the automotive industry. The transient convection-reaction-diffusion equation is solved in the simplified canister geometry shown in figure 13. The diffusion is constant $\nu = 0.01$ whereas the reaction is larger in the outlet of the canister ($\sigma = 10$ in $\Omega_2 \cup \Omega^O$) and smaller in the rest of the canister ($\sigma = 0.1$ in Ω_1). Thus, Ω^O is a pollutant trap capturing all the pollutant that the actual filter (domain Ω_1) is not able to retain. The advection field, α , is a piecewise linear field (see figure 13) resulting from a finite element computation of a potential flow in the same mesh. The inlet concentration of pollutant is set to one ($u = 1$ in Γ_1) and the outlet concentration of pollutant is set to zero ($u = 0$ in Γ_2). The rest of the boundary conditions are Neumann homogeneous because the walls of the canister are considered to be impermeable.

The initial pollutant concentration state is taken as $u_0 = 1$ in the inlet boundary, Γ_1 , and zero elsewhere. In practice, the initial condition u_0 has to be interpolated in the mesh and therefore u_0 is set to 1 in the nodes of the mesh lying in Γ_1 and set to zero in the rest of the nodes. The time interval is taken from $t = 0$ to $t = T = 2$. Here, both the spatial and time discretizations

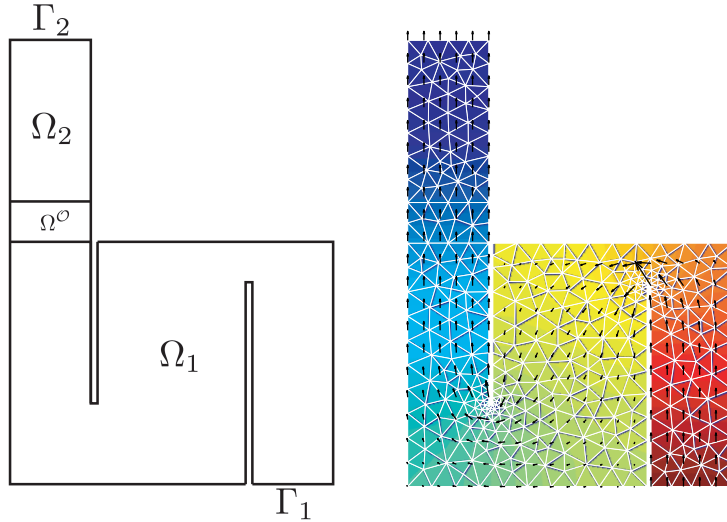


Figure 13: Example 5: Computational domain (left) and incompressible advection field α (right).

are linear, that is $p = 1$ and $q = 1$, for a constant time step $\Delta t_n = 0.08$, which corresponds to $N = 25$.

The quantity of interest, is the total pollutant captured by the trap domain Ω^O along the complete time evolution. Note that the canister is considered to work properly if this quantity is small enough. If the outcome of pollutant exceeds a threshold value, the canister breaks and the design is not admissible. This quantity is expressed in terms of the solution by

$$L^O(u) = \int_0^T \int_{\Omega^O} u(x, y, t) \, d\Omega \, dt,$$

that is $u_T^O = 0$ and $f^O = 1$ in Ω^O and zero elsewhere, in equation (14).

As in the previous examples, the adaptive procedure subdivides the elements with larger contribution to the bound gap, until the relative half bound gap reaches a target value of 2.5%. Also in this example, in each step 2% of the elements are marked to be refined. Again, the bounds are obtained using a fourth order interpolation for the dual estimates $r = 4$. The initial and final meshes are shown in figure 14, as well as the evolution of the bounds along the adaptive process. The corresponding values of the bounds are displayed in table 6. In the final mesh, the bounds for the quantity of interest guarantee that $L^O(u_\tau) = 7.606 \cdot 10^{-4} \pm 0.188 \cdot 10^{-4} = 7.606 \cdot 10^{-4} \pm 2.47\%$.

n_{el}	$L^O(u_h)$	s^-	s^+	s^{ave}	Δ	Δ_{rel}
686	$7.740 \cdot 10^{-4}$	$5.744 \cdot 10^{-4}$	$9.581 \cdot 10^{-4}$	$7.663 \cdot 10^{-4}$	$1.918 \cdot 10^{-4}$	25.03%
5786	$7.609 \cdot 10^{-4}$	$7.419 \cdot 10^{-4}$	$7.794 \cdot 10^{-4}$	$7.606 \cdot 10^{-4}$	$0.188 \cdot 10^{-4}$	2.47%

Table 6: Example 5: bounds for the initial and final meshes of the adaptive process

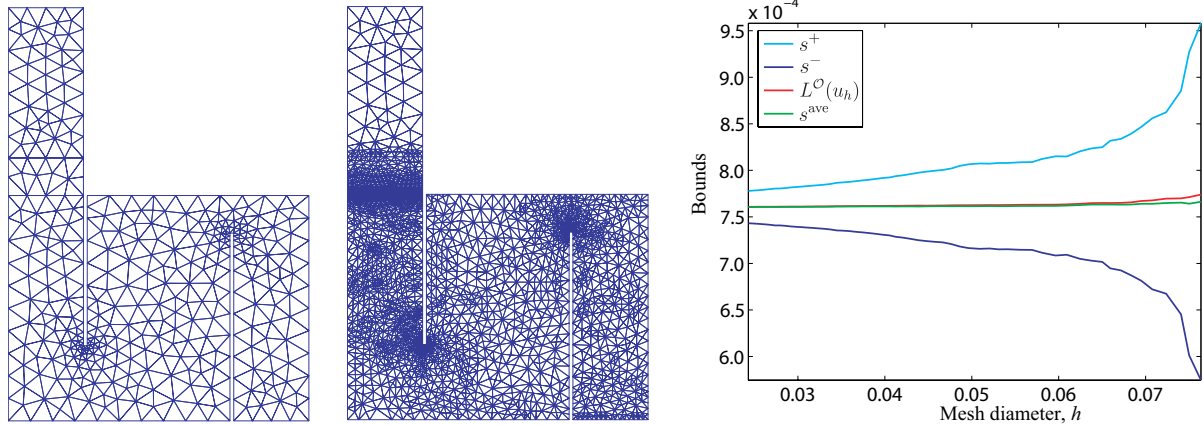


Figure 14: Example 5: initial and final meshes with 686 and 5786 elements (left and center) and convergence of bounds along the adaptive h -refinement (right).

References

- [1] S. Adjerid, J.E. Flaherty, and I. Babuška. A posteriori error estimation for the finite element method-of-lines solution of parabolic problems. *Math. Models Methods Appl. Sci.*, 9(2):261–286, 1999.
- [2] M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis*. John Wiley & Sons, Chichester, 2000.
- [3] I. Babuška and S. Ohnibus. A posteriori error estimation for the semidiscrete finite element method of parabolic differential equations. *Comput. Methods Appl. Mech. Engrg.*, 190(35-36):4691–4712, 2001.
- [4] I. Babuška and T. Strouboulis. *The finite element method and its reliability*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York, 2001.
- [5] P. Díez and G. Calderón. Goal-oriented error estimation for transient parabolic problems. *Comput. Mech.*, electronic, 2007.
- [6] A. Huerta and P. Díez. Error estimation including pollution assessment for nonlinear finite element analysis. *Comput. Methods Appl. Mech. Eng.*, 181(1–3):21–41, 2000.
- [7] P. Ladevèze and D. Leguillon. Error estimate procedure in the finite element method and applications. *SIAM J. Numer. Anal.*, 20(3):485–509, 1983.
- [8] L. Machiels. A posteriori finite element bounds for output functionals of discontinuous Galerkin discretizations of parabolic problems. *Comput. Methods Appl. Mech. Engrg.*, 190(26-27):3401–3411, 2001.
- [9] J. T. Oden and S. Prudhomme. Goal-oriented error estimation and adaptivity for the finite element method. *Comput. Math. Appl.*, 41(5-6):735–756, 2001.

- [10] M. Paraschivoiu, J. Peraire, and A. T. Patera. A posteriori finite element bounds for linear-functional outputs of elliptic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 150(1-4):289–312, 1997. Symposium on Advances in Computational Mechanics, Vol. 2 (Austin, TX, 1997).
- [11] N. Parés, J. Bonet, A. Huerta, and J. Peraire. The computation of bounds for linear-functional outputs of weak solutions to the two-dimensional elasticity equations. *Comput. Methods Appl. Mech. Engrg.*, 195(4-6):406–429, 2006.
- [12] N. Parés, P. Díez, and A. Huerta. Bounds of functional outputs for parabolic problems. Part II: Bounds of the exact solution.
- [13] A. M. Sauer-Budge, J. Bonet, A. Huerta, and J. Peraire. Computing bounds for linear functionals of exact weak solutions to Poisson’s equation. *SIAM J. Numer. Anal.*, 42(4):1610–1630 (electronic), 2004.
- [14] A. M. Sauer-Budge and J. Peraire. Computing bounds for linear functionals of exact weak solutions to the advection-diffusion-reaction equation. *SIAM J. Sci. Comput.*, 26(2):636–652 (electronic), 2004.
- [15] T. Strouboulis, I. Babuška, and D. K. Datta. Guaranteed a posteriori error estimation for fully discrete solutions of parabolic problems. *Internat. J. Numer. Methods Engrg.*, 56(9):1243–1259, 2003.
- [16] T. Werder, K. Gerdes, D. Schötzau, and C. Schwab. *hp*-discontinuous Galerkin time stepping for parabolic problems. *Comput. Methods Appl. Mech. Engrg.*, 190(49-50):6685–6708, 2001.
- [17] Z. C. Xuan, N. Parés, and J. Peraire. Computing upper and lower bounds for the *J*-integral in two-dimensional linear elasticity. *Comput. Methods Appl. Mech. Engrg.*, 195(4-6):430–443, 2006.

A Discretization of the adjoint problem

In this appendix the adjoint problem given in equation (16), namely, find $\psi_h \in \widehat{\mathcal{W}}^h$ such that

$$A_\tau(v, \psi_h) = L^\mathcal{O}(v) \quad \forall v \in \widehat{\mathcal{W}}^h, \quad (46)$$

is decomposed into N local-in-time problems posed over each time slab (equation (17)).

First, the r.h.s. of the previous equation is expanded as

$$\begin{aligned} A_\tau(v, \psi_h) &= \int_0^T \left[\langle \dot{v}, \psi_h \rangle + a(t; v, \psi_h) \right] dt + (v^+(0), \psi_h^+(0)) \\ &\quad + \sum_{n=1}^{N-1} \left[(v^+(t^n), \psi_h^+(t^n)) - (v^-(t^n), \psi_h^+(t^n)) \right]. \end{aligned}$$

Since the term containing the time derivative verifies

$$\begin{aligned} \int_0^T \langle \dot{v}, \psi_h \rangle dt &= \int_0^T -\langle v, \dot{\psi}_h \rangle dt + \sum_{n=1}^N \left[(v^-(t^n), \psi_h^-(t^n)) - (v^+(t^{n-1}), \psi_h^+(t^{n-1})) \right] \\ &= \int_0^T -\langle v, \dot{\psi}_h \rangle dt + (v^-(T), \psi_h^-(T)) - (v^+(0), \psi_h^+(0)) \\ &\quad + \sum_{n=1}^{N-1} \left[(v^-(t^n), \psi_h^-(t^n)) - (v^+(t^n), \psi_h^+(t^n)) \right], \end{aligned}$$

the term $A_\tau(v, \psi_h)$ may be rewritten as

$$\begin{aligned} A_\tau(v, \psi_h) &= \int_0^T \left[-\langle \dot{\psi}_h, v \rangle + a(t; v, \psi_h) \right] dt + \sum_{n=1}^N (v^-(t^n), \psi_h^-(t^n)) \\ &\quad - \sum_{n=1}^{N-1} (v^-(t^n), \psi_h^+(t^n)). \end{aligned} \quad (47)$$

Second, the final condition for the adjoint problem $\psi_h^+(T) = u_T^\mathcal{O}$ is used to rewrite the functional output as

$$L^\mathcal{O}(v) = \int_0^T \ell^\mathcal{O}(t; v) dt + (\psi_h^+(T), v^-(T)). \quad (48)$$

Finally, joining the expressions for the l.h.s. and r.h.s. of equation (46), equations (47) and (48) respectively, allows to rewrite the problem (46) as

$$\begin{aligned} \int_0^T \left[-\langle \dot{\psi}_h, v \rangle + a(t; v, \psi_h) \right] dt + \sum_{n=1}^N (\psi_h^-(t^n), v^-(t^n)) \\ = \int_0^T \ell^\mathcal{O}(t; v) dt + \sum_{n=1}^N (\psi_h^+(t^n), v^-(t^n)), \end{aligned} \quad (49)$$

yielding the local-in-time problems given in equation (17). It is also usual to write the adjoint problem given in equation (49) as

$$\int_0^T \left[-\langle \dot{\psi}_h, v \rangle + a(t; v, \psi_h) \right] dt - \sum_{n=1}^N (\llbracket \psi_h \rrbracket_n, v^-(t^n)) = \int_0^T \ell^\mathcal{O}(t; v) dt.$$

B Proof of Lemma 1

Lemma 1 For any $v \in \widehat{\mathcal{W}}$, $A_\tau(v, v) \geq \|v\|^2$.

Proof. For any $v \in \widehat{\mathcal{W}}$

$$A_\tau(v, v) = \int_0^T \left[\langle \dot{v}, v \rangle + a(t; v, v) \right] dt + (v^+(0), v^+(0)) + \sum_{n=1}^{N-1} (\llbracket v \rrbracket_n, v^+(t^n)).$$

Now,

$$\begin{aligned} \int_0^T \langle \dot{v}, v \rangle dt &= \sum_{n=1}^N \int_{I_n} \langle \dot{v}, v \rangle dt = \frac{1}{2} \sum_{n=1}^N (v^-(t^n), v^-(t^n)) - (v^+(t^{n-1}), v^+(t^{n-1})) \\ &= \frac{1}{2} (v_N^-, v_N^-) - \frac{1}{2} (v_0^+, v_0^+) + \frac{1}{2} \sum_{n=1}^{N-1} [(v_n^-, v_n^-) - (v_n^+, v_n^+)], \end{aligned}$$

where the notation $v_n^\pm := v^\pm(t^n)$ has been used. Therefore

$$\begin{aligned} \int_0^T \langle \dot{v}, v \rangle dt + (v^+(0), v^+(0)) + \sum_{n=1}^{N-1} (\llbracket v \rrbracket_n, v^+(t^n)) \\ &= \frac{1}{2} (v_N^-, v_N^-) + \frac{1}{2} (v_0^+, v_0^+) + \sum_{n=1}^{N-1} \left[\frac{1}{2} (v_n^-, v_n^-) - \frac{1}{2} (v_n^+, v_n^+) + (\llbracket v \rrbracket_n, v_n^+) \right] \\ &= \frac{1}{2} (v_N^-, v_N^-) + \frac{1}{2} (v_0^+, v_0^+) + \sum_{n=1}^{N-1} \left[\frac{1}{2} (v_n^-, v_n^-) + \frac{1}{2} (v_n^+, v_n^+) - (v_n^-, v_n^+) \right] \\ &= \frac{1}{2} (v_N^-, v_N^-) + \frac{1}{2} (v_0^+, v_0^+) + \frac{1}{2} \sum_{n=1}^{N-1} (\llbracket v \rrbracket_n, \llbracket v \rrbracket_n). \end{aligned}$$

Now substituting the previous expression in the definition of $A_\tau(v, v)$ yields

$$A_\tau(v, v) = \int_0^T a(t; v, v) dt + \frac{1}{2} \left((v_0^+, v_0^+) + \sum_{n=1}^{N-1} (\llbracket v \rrbracket_n, \llbracket v \rrbracket_n) + (v_N^-, v_N^-) \right).$$

Therefore,

$$A_\tau(v, v) \geq \int_0^T a(t; v, v) dt = \|v\|_0^2,$$

which concludes the proof. \square

C Proof of remark 5

This appendix rewrites the primal and adjoint residuals $R^P(vN_{nj}(t))$ and $R^D(vN_{nj}(t))$ in the form given by remark 5.

For any $v(t) \in \widehat{\mathcal{W}}^\tau$, the primal residual can be rewritten as:

$$\begin{aligned} R^P(v) &= L(v) - A_\tau(u_h, v) = L(v) - A(u_h, v) - \sum_{n=1}^{N-1} (\llbracket u_h \rrbracket_n, v^+(t^n)) \\ &= \int_0^T \left[\ell(t; v) - \langle \dot{u}_h, v \rangle - a(t; u_h, v) \right] dt + (u_0 - u_h^+(0), v^+(0)) \\ &\quad - \sum_{n=1}^{N-1} (\llbracket u_h \rrbracket_n, v^+(t^n)) \\ &= \int_0^T \left[\ell(t; v) - \langle \dot{u}_h, v \rangle - a(t; u_h, v) \right] dt - \sum_{n=0}^{N-1} (\llbracket u_h \rrbracket_n, v^+(t^n)), \end{aligned}$$

where $\llbracket u_h \rrbracket_0 := u_h^+(0) - u_0$. Now, taking $v(t) = vN_{nj}(t)$ for $v \in \mathcal{V}$, and noting that $N_{nj}(t)$ vanishes outside the slab I_n yields

$$R^P(vN_{nj}(t)) = \int_{I_n} \left[\ell(t; vN_{nj}(t)) - \langle \dot{u}_h, vN_{nj}(t) \rangle - a(t; u_h, vN_{nj}(t)) \right] dt \\ - (\llbracket u_h \rrbracket_{n-1}, vN_{nj}(t^{n-1})).$$

Finally, the expression given in remark 5 for $R^P(vN_{nj}(t))$, equation (32), is obtained rearranging terms and noting that $N_{nj}(t)$ is a constant-in-space function:

$$R^P(vN_{nj}(t)) = \int_{I_n} \left[\langle f(t), vN_{nj}(t) \rangle - \langle \dot{u}_h, vN_{nj}(t) \rangle - a_n(u_h, vN_{nj}(t)) \right] dt \\ - (\llbracket u_h \rrbracket_{n-1}, vN_{nj}(t^{n-1})) \\ = \left\langle \int_{I_n} \left[f(t)N_{nj}(t) - \dot{u}_h N_{nj}(t) \right] dt - \llbracket u_h \rrbracket_{n-1} N_{nj}(t^{n-1}), v \right\rangle \\ - a_n \left(\int_{I_n} u_h N_{nj}(t) dt, v \right).$$

In order to rewrite the adjoint residue, the following alternative representation of $A_\tau(v, \psi_h)$, which may be easily derived from equation (47) in appendix A, is used

$$A_\tau(v, \psi_h) = \int_0^T \left[-\langle \dot{\psi}_h, v \rangle + a(t; v, \psi_h) \right] dt + (\psi_h^-(T), v^-(T)) - \sum_{n=1}^{N-1} (\llbracket \psi_h \rrbracket_n, v^-(t^n)).$$

In this case, the dual residual can be rewritten as:

$$R^D(v) = L^O(v) - A_\tau(v, \psi_h) \\ = \int_0^T \left[\ell^O(t; v) + \langle \dot{\psi}_h, v \rangle - a(t; v, \psi_h) \right] dt + (u_T^O - \psi_h^-(T), v^-(T)) \\ + \sum_{n=1}^{N-1} (\llbracket \psi_h \rrbracket_n, v^-(t^n)) \\ = \int_0^T \left[\ell^O(t; v) + \langle \dot{\psi}_h, v \rangle - a(t; v, \psi_h) \right] dt + \sum_{n=1}^N (\llbracket \psi_h \rrbracket_n, v^-(t^n)),$$

where $\llbracket \psi_h \rrbracket_N := u_T^O - \psi_h^-(T)$. Finally, the expression given in remark 5 for $R^D(vN_{nj}(t))$, equation (33), is recovered taking $v(t) = vN_{nj}(t)$, rearranging terms and recalling that $N_{nj}(t)$

is a constant-in-space function:

$$\begin{aligned}
R^D(vN_{nj}(t)) &= \int_{I_n} \left[\ell^{\mathcal{O}}(t; vN_{nj}(t)) + \langle \dot{\psi}_h, vN_{nj}(t) \rangle - a(t; vN_{nj}(t), \psi_h) \right] dt \\
&\quad + ([\psi_h]_n, vN_{nj}(t^n)) \\
&= \int_{I_n} \left[\langle f^{\mathcal{O}}(t), vN_{nj}(t) \rangle + \langle \dot{\psi}_h, vN_{nj}(t) \rangle - a_n(v, \psi_h N_{nj}(t)) \right] dt \\
&\quad + ([\psi_h]_n N_{nj}(t^n), v) \\
&= \left\langle \int_{I_n} \left[f^{\mathcal{O}}(t) N_{nj}(t) + \dot{\psi}_h N_{nj}(t) + [\psi_h]_n N_{nj}(t^n) \right] dt \right\rangle \\
&\quad - a_n(v, \int_{I_n} \psi_h N_{nj}(t) dt).
\end{aligned}$$

D Computation of the dual estimates

This appendix intends to provide a sketch of the computation of the polynomial dual estimates $\hat{\mathbf{P}}_{nj}^P$, $\hat{\mathbf{P}}_{nj}^D$ and \hat{R}_{nj}^P , \hat{R}_{nj}^D solution of (45). A detailed construction can be found in [13, 14].

First the computation of $\hat{\mathbf{P}}_{nj}^P$ and \hat{R}_{nj}^P is detailed and the analogous computation is then given for the adjoint problem.

Using remark 5, equation (45a) may be rewritten as:

$$\int_{\Omega_k} \left[\nu_n \hat{\mathbf{P}}_{nj}^P \cdot \nabla v + \sigma_n \hat{R}_{nj}^P v \right] d\Omega = \langle f_{nj}, v \rangle_k - a_{nk}(u_h^{nj}, v) + \int_{\partial\Omega_k} \tau_k \lambda_{nj}^P v d\Gamma,$$

where $\langle \cdot, \cdot \rangle_k$ and $a_{nk}(\cdot, \cdot)$ are the restrictions of the bilinear forms $\langle \cdot, \cdot \rangle$ and $a_n(\cdot, \cdot)$ to the element Ω_k . Expanding now the r.h.s. of the previous equation yields:

$$\begin{aligned}
&\langle f_{nj}, v \rangle_k - a_{nk}(u_h^{nj}, v) + \int_{\partial\Omega_k} \tau_k \lambda_{nj}^P v d\Gamma \\
&= \int_{\Omega_k} \left[f_{nj} - \boldsymbol{\alpha}_n \cdot \nabla u_h^{nj} - \sigma_n u_h^{nj} \right] v d\Omega - \int_{\Omega_k} \nu_n \nabla u_h^{nj} \cdot \nabla v d\Omega + \int_{\partial\Omega_k} \tau_k \lambda_{nj}^P v d\Gamma.
\end{aligned}$$

Therefore the problem reduces to determine $\hat{\mathbf{P}}_{nj}^P$ and \hat{R}_{nj}^P such that

$$\begin{aligned}
&\int_{\Omega_k} \left[\nu_n (\hat{\mathbf{P}}_{nj}^P + \nabla u_h^{nj}) \cdot \nabla v + \sigma_n \hat{R}_{nj}^P v \right] d\Omega \\
&= \int_{\Omega_k} \left[f_{nj} - \boldsymbol{\alpha}_n \cdot \nabla u_h^{nj} - \sigma_n u_h^{nj} \right] v d\Omega + \int_{\partial\Omega_k} \tau_k \lambda_{nj}^P v d\Gamma \quad \forall v \in \mathcal{H}^1(\Omega_k).
\end{aligned}$$

Finally, denoting by $\tilde{\mathbf{P}}_{nj}^P := \hat{\mathbf{P}}_{nj}^P + \nabla u_h^{nj}$ and integrating by parts the previous equation, yields the strong equations to determine $\tilde{\mathbf{P}}_{nj}^P$ and \hat{R}_{nj}^P :

$$\begin{aligned}
\nu_n \tilde{\mathbf{P}}_{nj}^P \cdot \mathbf{n} &= \tau_k \lambda_{nj}^P, \\
-\nu_n \nabla \cdot \tilde{\mathbf{P}}_{nj}^P + \sigma_n \hat{R}_{nj}^P &= f_{nj} - \boldsymbol{\alpha}_n \cdot \nabla u_h^{nj} - \sigma_n u_h^{nj}.
\end{aligned}$$

It is worth noting that the previous equations do not uniquely determine $\tilde{\mathbf{P}}_{njk}^P$ and \hat{R}_{njk}^P . It is also clear at this point that if u_0 and $f(t)$ are polynomial functions in each element of the mesh, then it is possible to choose a suitable interpolation degree r so that the dual estimates may be sought in $[\mathbb{P}^r(\Omega_k)]^{\text{nsd}}$ and $\mathbb{P}^r(\Omega_k)$ respectively.

The same reasoning yields to the following strong equations for the adjoint dual estimates:

$$\begin{aligned}\nu_n \tilde{\mathbf{P}}_{njk}^D \cdot \mathbf{n} &= \tau_k \lambda_{nj}^D, \\ -\nu_n \nabla \cdot \tilde{\mathbf{P}}_{njk}^D + \sigma_n \hat{R}_{njk}^D &= f_{nj}^{\mathcal{O}} - \sigma_n \psi_h^{nj}.\end{aligned}$$

where $\hat{\mathbf{P}}_{njk}^D := \tilde{\mathbf{P}}_{njk}^D - \nabla \psi_h^{nj} - 1/\nu_n \boldsymbol{\alpha}_n \psi_h^{nj}$.