Nested dual–residual a posteriori error estimators for advection-diffusion-reaction problems *

Stefano Micheletti[♯] and Simona Perotto[♯]

30th August 2005

[♯] MOX– Modellistica e Calcolo Scientifico
Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano
via Bonardi 9, 20133 Milano, Italy
{stefano.micheletti,simona.perotto}@mate.polimi.it

Keywords: Goal oriented a posteriori error analysis, residual based error estimators, mesh adaption, advection-diffusion-reaction problems, finite element method.

AMS Subject Classification: 65N15, 65N30, 65N50

Abstract

In this work we introduce a fully computable dual-based a posteriori error estimator for standard scalar advection-diffusion-reaction problems. In particular, such an estimator does not depend on neither the primal nor the dual exact solution, but only on the corresponding Galerkin finite element approximations. This new approach merges the main advantages of the dual-based and of the residual-based error analysis, being devised as a residual-based estimator "nested" in a dual-based one. This allows us to explicitly approximate suitable functionals of the solution, in the spirit of a classical goal-oriented analysis, at the same cost as a dual-based strategy, the solution of two differential problems being involved. The related issue of optimal mesh adaptivity is also addressed. Several two-dimensional numerical test cases validate the proposed theory as well as the employed adaptive procedure.

*This work has been supported by the Project COFIN 2003 "Metodi Numerici per Applicazioni Avanzate in Meccanica dei Fluidodinamica ed Elettromagnetismo".

1 Introduction and motivations

Nowadays mesh adaptivity finds a large following in many engineering applications where the numerical solution of partial differential equations is required. The basic idea is to reduce the computational cost involved in the numerical approximation of the phenomenon at hand, by suitably distributing the mesh elements to better capture the nonsmooth behavior and the directional features (shocks, boundary and internal layers, singularities, etc.) of the exact solution u. We may distinguish the different approaches employed for mesh adaptivity in *heuristic* and *theoretically based* adaption techniques. In the first class, the gradient or the Hessian of the numerical solution is typically employed to drive the mesh adaption procedure. For instance, this is the case of the well-known Zienkiewicz-Zhu error estimator [33, 34, 35]. The second group consists of more theoretically sound techniques, moving from suitable a priori or a posteriori estimators for the discretization error. In this framework the specific literature is broad (see, e.g., [1, 5, 6, 8, 13, 19, 28, 32]). Among the approaches based on a posteriori error estimates, the most known are the so-called *residual-based* and *dual-based* estimators (as reviewing works see, for example, [1, 32] for the first ones and [8, 19] for the second ones). The residual-based analysis has the advantage of providing fully computable and easily implementable estimators. However, only the energy norm of the discretization error can be controlled in such a case. On the contrary, the dual-based estimators are more flexible, since linear or nonlinear functionals $J(u)$ of the solution can now be controlled, via the resolution of a suitable adjoint problem. These functionals are typically associated with quantities significant in engineering applications, known as goal or target quantities: meaningful examples are the lift and drag around bodies in external flows or mean and local values in computational fluid dynamics, the torsion moment, the point-wise stress or the surface tension in structural mechanics, the total output electrical current in a semiconductor device in microelectronics. Nevertheless, these estimators suffer the limitation to be not thoroughly computable. In more detail, let us suppose to have at our disposal a discrete solution u_h to u, for example a Galerkin finite element solution, so as to approximate the goal quantity $J(u)$ via the value $J(u_h)$. In this case, the well-known Dual Weighted Residual (DWR) theory (see [7, 8]) provides us with an error estimator η for the functional error $J(u) - J(u_h)$ of the form

$$\eta = \sum_{K \in \mathcal{T}_h} \rho_K(u_h)\, \omega_K(z),$$

or, more in general,

$$\eta = \frac{1}{2} \sum_{K \in \mathcal{T}_h} \rho_K(u_h)\, \omega_K(z) + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \rho_K(z_h)\, \omega_K(u), \tag{1}$$

$\rho_K(u_h), \rho_K(z_h)$ and $\omega_K(u), \omega_K(z)$ being suitable residuals (depending on the computable numerical solutions u_h, z_h) and weights (associated with the exact

primal and dual solutions u and z), respectively. Thus, while the elemental residuals $\rho_K(u_h)$ and $\rho_K(z_h)$ can be computed explicitly, the weights $\omega_K(u)$ and $\omega_K(z)$ have to be suitably approximated, at the price of increasing the overall computational burden. Different strategies have been pursued in the literature to replace the unknown solutions u and z with sufficiently "fine" approximations [8, 19], the error introduced in this last step remaining however not precisely assessed.

Aim of this work is the derivation of a fully computable dual-based a posteriori error estimator, merging the main advantages of both the dual-based and the residual-based analysis, at the same cost of the most expensive approach, i.e. the dual-based one. We name this new estimator as nested dual-residual, being devised, as shown later, as a residual-based estimator "nested" in a dual-based one.

Let us briefly explain the idea behind such an approach. Assume that the primal and the dual problems have the following structure

$$\begin{cases} \mathcal{L}u = f, \\ + b.c.\,, \end{cases} \qquad \begin{cases} \mathcal{L}^* z = g, \\ + b.c.\,, \end{cases}$$

respectively, where $\mathcal{L}$ and $\mathcal{L}^*$ represent the primal operator and its adjoint form, f and g denote the primal and the dual source term, and $b.c.$ stands for some prescribed boundary conditions. The goal quantity can thus be written as the linear functional

$$J(u) = \int_\Omega g\, u\, d\mathbf{x}.$$

To begin, we derive the functional error representation

$$J(u - u_h) = \theta \int_\Omega f\,(z - z_h)\, d\mathbf{x} + (1-\theta) \int_\Omega g\,(u - u_h)\, d\mathbf{x}, \tag{2}$$

with $0 \leq \theta \leq 1$ an arbitrary constant, and u_h and z_h suitable Galerkin approximations to u and z. The right-hand side of (2) provides us with a quantity which, though exact, is not computable, due to the presence of the primal and of the dual discretization errors. To overcome such a limitation, we first exploit the Cauchy-Schwarz inequality to get a bound for $|J(u-u_h)|$ involving the L^2-norm (or the H^1-norm) of both $(u-u_h)$ and $(z-z_h)$. Then, these errors are bounded in turn, simply by using standard residual-based a posteriori estimators. The resulting quantity is thus thoroughly computable.

The considered numerical test cases confirm the good quality of this new estimator. Moreover the same idea has been generalized to the model adaption framework, yielding, also in this case, satisfactory results [11].

The outline of the paper is as follows. In § 2, we introduce the basic idea of the nested dual-residual analysis on the standard Poisson problem. In § 3

we generalize this framework to a standard scalar advection-diffusion-reaction equation. The iterative procedure used to drive the mesh adaptivity is discussed in § 4. The numerical validation is carried out in § 5 for both the Poisson and the advection-diffusion problems. Some conclusions and perspectives are drawn in § 6, while in the Appendices A. 1. and A. 2., self-contained results on the a posteriori analysis and some issues about optimal mesh generation are addressed.

2 The reference problem

To detail the idea behind the a posteriori error analysis proposed in this paper, let us move from a model problem, namely the standard Poisson problem. In § 3 we generalize this analysis to the more significant advection-diffusion-reaction framework. In such a case the issue of the stabilization makes the presentation of our new approach less immediate. Moreover, while for the Poisson problem two optimal (with respect to the convergence rate) nested a posteriori error estimators can be derived, according to the regularity of the solution, only an optimal error estimate has been rigorously proved for the advection-diffusion-reaction problem.

Thus let us introduce, as primal problem, the standard Poisson equation provided with homogeneous Dirichlet boundary conditions

$$\begin{cases} -\Delta u = f & \text{in} \quad \Omega, \\ u = 0 & \text{on} \quad \partial\Omega, \end{cases} \tag{3}$$

where Ω is a domain in $\mathbb{R}^d$ (with $d = 2, 3$) with Lipschitz continuous boundary $\partial\Omega$, and $f \in L^2(\Omega)$ is the given source term. In the sequel Ω always coincides with a polygonal domain.

Throughout, we adopt the standard notation $H^s(\Omega)$ to denote the Sobolev spaces of functions with Lebesgue measurable derivatives [22]. In particular, notice that the space $H^0(\Omega)$ coincides with the space $L^2(\Omega)$ of square-integrable functions, while $H_0^1(\Omega)$ stands for the closure in $H^1(\Omega)$ of the space $\mathscr{D}(\Omega)$ of infinitely differentiable functions with a compact support in Ω. Moreover, from the Poincaré-Friedrichs inequality, it follows that the H^1-seminorm is a norm on $H_0^1(\Omega)$, equivalent to the norm of $H^1(\Omega)$.

In view of a finite element approximation of problem (3), let us introduce the corresponding weak form which reads as: find $u \in V \equiv H_0^1(\Omega)$ such that

$$\int_\Omega \nabla u \cdot \nabla v \, d\mathbf{x} = \int_\Omega f \, v \, d\mathbf{x}, \qquad \forall v \in V, \tag{4}$$

with $\mathbf{x} = (x_i)^T \in \Omega$, for $i = 1, \ldots, d$.

Let $\{\mathcal{T}_h\}_h$ be a family of conforming triangulations of $\overline{\Omega}$ into triangles K of diameter $h_K \leq h$, where $h = \max_{K \in \mathcal{T}_h} h_K$. Each element K can be viewed

as the image, via an invertible affine mapping T_K, of the reference triangle $\widehat{K}$, usually chosen as the right triangle $(0,0),(1,0),(0,1)$ or as the equilateral one $(-1/2,0),(1/2,0),(0,\sqrt{3}/2)$. Now let $V_h \subset V$ denote the finite element space of continuous affine functions [9]. Then the discrete form associated with (4) is given by: find $u_h \in V_h$ such that

$$\int_\Omega \nabla u_h \cdot \nabla v_h \, d\mathbf{x} = \int_\Omega f \, v_h \, d\mathbf{x}, \qquad \forall v_h \in V_h. \tag{5}$$

Simply by subtracting (5) from (4), for the choice $v = v_h$, we get the so-called Galerkin orthogonality property

$$\int_\Omega \nabla e_h \cdot \nabla v_h \, d\mathbf{x} = 0, \qquad \forall v_h \in V_h, \tag{6}$$

$e_h = u - u_h$ being the discretization error associated with the primal problem (3).

As we are interested in a goal-oriented analysis [7, 8, 18, 19, 24, 28, 29, 31], let us introduce the linear goal functional $J : V \to \mathbb{R}$, identifying the quantity we aim to control. Such a functional is related to the source term of the dual problem associated with (3), whose weak form reads as: find $z \in V$ such that

$$\int_\Omega \nabla \varphi \cdot \nabla z \, d\mathbf{x} = J(\varphi), \qquad \forall \varphi \in V. \tag{7}$$

According to the theory in [23], let us assume that the representation

$$J(\varphi) = \int_\Omega g \, \varphi \, d\mathbf{x}, \tag{8}$$

holds for the functional J, g being a suitable function in $L^2(\Omega)$. Equality (8) allows us to provide the differential form of the dual problem (7), given by

$$\begin{cases} -\Delta z = g & \text{in} \quad \Omega, \\ z = 0 & \text{on} \quad \partial\Omega, \end{cases} \tag{9}$$

i.e., to rewrite the weak form (7) as

$$\int_\Omega \nabla \varphi \cdot \nabla z \, d\mathbf{x} = \int_\Omega g \, \varphi \, d\mathbf{x}, \qquad \forall \varphi \in V. \tag{10}$$

Remark 2.1 *The choice* (8) *for the output functional J makes the choice of the space V for the adjoint problem* (7) *consistent with the theory in* [18].

To discretize the dual problem (7), let us exploit the same scheme adopted for the primal problem. Notice that this choice simplifies the numerical validation, a single computational mesh being involved. Thus the discrete form of the dual problem reads as: find $z_h \in V_h \subset V$, such that

$$\int_\Omega \nabla\varphi_h \cdot \nabla z_h \, d\mathbf{x} = J(\varphi_h) \left(= \int_\Omega g\, \varphi_h \, d\mathbf{x} \right), \qquad \forall \varphi_h \in V_h. \tag{11}$$

Remark 2.2 *Moving from the goal functional representation* (8)*, both the weak and the discrete forms,* (10) *and* (11) *can be expressed in terms of the same function g. Alternatively, some suitable finite dimensional approximation of g may be employed in* (11)*.*

Analogously to (6), by subtracting (11) from (7), for the choice $\varphi = \varphi_h$, we get the Galerkin orthogonality property referred to the dual problem, given by

$$\int_\Omega \nabla\varphi_h \cdot \nabla e_h^* \, d\mathbf{x} = 0, \qquad \forall \varphi_h \in V_h, \tag{12}$$

where $e_h^* = z - z_h$ stands for the discretization error associated with the dual problem (9). In the sequel, we refer to properties (6) and (12) as to the *primal* and the *dual Galerkin orthogonality*, respectively.

2.1 Preliminaries 1: interpolation error estimates

In view of the a posteriori analysis below, let us first recall some classical results related to the finite element interpolation error theory.

In more detail, let $\Pi_h : C^0(\overline{\Omega}) \to V_h$ and $I_h : L^2(\Omega) \to V_h$ be the standard Lagrange and Clément affine interpolants, respectively, Π_K and I_K denoting the corresponding restrictions to the general element $K \in \mathcal{T}_h$ [9, 10]. Moreover, the definition of the Clément interpolation operator leads us to introduce the patch $\Delta_K = \{\bigcup_{T\in\mathcal{T}_h} T \ : T \cap K \neq \emptyset\}$ of all the elements sharing (at least) a vertex with K. In the sequel we assume the cardinality of any patch Δ_K to be uniformly bounded, independently of the geometry of the mesh, i.e., there exists a positive integer N such that, for any $K \in \mathcal{T}_h$,

$$\mathrm{card}(\Delta_K) \leq N. \tag{13}$$

Notice that with this requirement we are essentially avoiding too thin triangles in the computational mesh $\mathcal{T}_h$.

The following local interpolation error estimates represent a significant ingredient for the a posteriori analysis of both § 2.3 and § 3.2.

Lemma 2.1 *Let $v \in H^2(K)$, for any $K \in \mathcal{T}_h$, with ∂K denoting the boundary of the element K. Then there exist two positive constants $\widetilde{C}_i = \widetilde{C}_i(\widehat{K}, d)$, with $i = 1, 2$, such that*

$$\begin{aligned} \|v - \Pi_K(v)\|_{L^2(K)} &\leq \widetilde{C}_1 \, h_K^2 \, |v|_{H^2(K)}, \\ \|v - \Pi_K(v)\|_{L^2(\partial K)} &\leq \widetilde{C}_2 \, h_K^{3/2} \, |v|_{H^2(K)}. \end{aligned} \tag{14}$$

Concerning the Clément interpolant, the following inequalities can be proved.

Lemma 2.2 *Let $v \in H^1(\Omega)$. Then there exist three positive constants $C_j = C_j(\widehat{K}, d, N)$, with $j = 1, 2, 3$, such that*

$$\begin{aligned} \|v - I_K(v)\|_{L^2(K)} &\leq C_1 \, h_K \, |v|_{H^1(\Delta_K)}, \\ \|v - I_K(v)\|_{L^2(\partial K)} &\leq C_2 \, h_K^{1/2} \, \|v\|_{H^1(\Delta_K)}, \\ \|\nabla(v - I_K(v))\|_{L^2(K)} &\leq C_3 \, |v|_{H^1(\Delta_K)}, \end{aligned} \tag{15}$$

N being defined according to (13).

Remark 2.3 *Results similar to the ones in Lemma* 2.2 *hold also for Clément-like interpolants, whose general expression is given by*

$$\widetilde{I}_h(v) = \sum_{\mathbf{a}_j \in \mathcal{N}_h} (P_j v)(\mathbf{a}_j) \, \psi_j,$$

with $v \in H^1(\Omega)$, where $\mathcal{N}_h = \{\mathbf{a}_j\}$ denotes the set of all the mesh nodes, ψ_j is the general nodal basis function, while with P_j we mean suitable local L^2-projection operators over a proper patch of elements $\Delta_{\mathbf{a}_j}$ associated with the node $\mathbf{a}_j$. The choice of such a patch identifies the different Clément-like interpolation operators [2, 10, 30].

2.2 Preliminaries 2: residual-based a posteriori analysis

The second essential ingredient for the analysis in § 2.3 is represented by some standard results of the residual-based a posteriori error analysis for the Poisson problem (3) [13, 32].
Let us begin by recalling the a posteriori estimate with respect to the H^1-norm.

Lemma 2.3 *Let u and u_h be the solution of the weak Poisson problem* (4) *and of the corresponding discrete formulation* (5)*, respectively, and let $e_h = u - u_h$ denote the associated discretization error. Then there exists a positive constant*

$C^{**} = C^{**}(\widehat{K}, d, N, C_P)$, *with* C_P *the Poincaré constant and* N *defined according to* (13)*, such that the following error estimate holds:*

$$\|e_h\|_{H^1(\Omega)} \leq C^{**} \left(\sum_{K \in \mathcal{T}_h} [\, \rho_K(u_h) \,]^2 \right)^{1/2}, \tag{16}$$

where

$$\rho_K(u_h) = h_K \, \|r_K(u_h)\|_{L^2(K)} + \frac{1}{2} \, h_K^{1/2} \left\| \left[\frac{\partial u_h}{\partial n} \right] \right\|_{L^2(\partial K)} \tag{17}$$

is the elemental primal residual term, with $r_K(u_h) = f + \Delta u_h$ *the internal residual associated with the primal problem* (3)*, and where* $[\partial u_h / \partial n]$ *stands for the jump of the normal derivative of* u_h *over the boundary* ∂K *of the element* K*.*

For the sake of completeness, we provide the proof of estimate (16) in Appendix A. 1. Due to the self-adjoint nature of the Laplace operator, a result similar to (16) can be stated moving from the dual problem (9).

Lemma 2.4 *Let* z *and* z_h *be the solution of the weak Poisson problem* (10) *and of the corresponding discrete formulation* (11)*, respectively, and let* $e_h^* = z - z_h$ *denote the associated discretization error. Then there exists a positive constant* $C^{**} = C^{**}(\widehat{K}, d, N, C_P)$*, with* C_P *the Poincaré constant and* N *defined according to* (13)*, such that the following error estimate holds:*

$$\|e_h^*\|_{H^1(\Omega)} \leq C^{**} \left(\sum_{K \in \mathcal{T}_h} [\, \rho_K^*(z_h) \,]^2 \right)^{1/2} \tag{18}$$

where

$$\rho_K^*(z_h) = h_K \, \|r_K^*(z_h)\|_{L^2(K)} + \frac{1}{2} \, h_K^{1/2} \left\| \left[\frac{\partial z_h}{\partial n} \right] \right\|_{L^2(\partial K)} \tag{19}$$

is the elemental dual residual term, with $r_K^*(z_h) = g + \Delta z_h$ *the internal residual associated with the dual problem* (9)*, and where* $[\partial z_h / \partial n]$ *stands for the jump of the normal derivative of* z_h *over the boundary* ∂K *of the element* K*.*

Via the well-known Aubin-Nitsche trick (see [4, 26]), a residual-based a posteriori error control can be derived also with respect to the L^2-norm.

Lemma 2.5 *Let* Ω *coincide with a convex polygonal domain. Let* u *and* u_h *be the solution of the weak Poisson problem* (4) *and of the corresponding discrete formulation* (5)*, respectively, and let* $e_h = u - u_h$ *denote the associated discretization error. Then there exists a positive constant* $C_{**} = C_{**}(\widehat{K}, d)$*, such that the following error estimate holds:*

$$\|e_h\|_{L^2(\Omega)} \leq C_{**} \left(\sum_{K \in \mathcal{T}_h} [\, h_K \, \rho_K(u_h) \,]^2 \right)^{1/2}, \tag{20}$$

where the elemental primal residual term $\rho_K(u_h)$ is defined according to (17).

As also Lemma 2.5 refers to the well-established a posteriori analysis for the Poisson problem, we confine the corresponding proof into Appendix A. 1. A similar result can be stated in the dual framework:

Lemma 2.6 *Let Ω coincide with a convex polygonal domain. Let z and z_h be the solution of the weak Poisson problem* (10) *and of the corresponding discrete formulation* (11)*, respectively, and let $e_h^* = z - z_h$ denote the associated discretization error. Then there exists a positive constant $C_{**} = C_{**}(\widehat{K}, d)$, such that the following error estimate holds:*

$$\|e_h^*\|_{L^2(\Omega)} \leq C_{**} \left(\sum_{K \in \mathcal{T}_h} [\, h_K \, \rho_K^*(z_h) \,]^2 \right)^{1/2}, \tag{21}$$

where the elemental dual residual term $\rho_K^*(z_h)$ is defined according to (19).

We have now all the ingredients necessary for the *nested a posteriori analysis* of the Poisson problem.

2.3 The a posteriori analysis

Let us recall that our actual goal is to approximate the goal quantity $J(u)$ with the approximate value $J(u_h)$. At the same time, we would like to provide a *fully computable* (a posteriori) estimate for the error $J(e_h) = J(u) - J(u_h)$, so as to guarantee that the value $J(u_h)$ differs from the exact one $J(u)$ within a prescribed tolerance τ. As stated in § 1, this aim is pursued by introducing a new kind of a posteriori error estimator, identified as *nested dual-residual*.
A first example of such an estimator is provided by the result below.

Proposition 2.1 *Let u and u_h be the solution of the weak Poisson problem* (4) *and of the corresponding discrete formulation* (5)*, respectively, $e_h = u - u_h$ denoting the associated discretization error. Let $J : V \to \mathbb{R}$ be the goal linear functional we are interested in. Then there exists a positive constant $C^{**} = C^{**}(\widehat{K}, d, N, C_P)$, with C_P the Poincaré constant and N defined according to* (13)*, such that*

$$\begin{aligned} |J(e_h)| \leq C^{**} \quad & \Bigg\{ \quad \theta \, \|f\|_{L^2(\Omega)} \Bigg(\sum_{K \in \mathcal{T}_h} [\, \rho_K^*(z_h) \,]^2 \Bigg)^{1/2} \\ & + \quad (1-\theta) \, \|g\|_{L^2(\Omega)} \left(\sum_{K \in \mathcal{T}_h} [\, \rho_K(u_h) \,]^2 \right)^{1/2} \Bigg\}, \end{aligned} \tag{22}$$

where the elemental primal and dual residual terms $\rho_K(u_h)$ and $\rho_K^*(z_h)$ are defined according to (17) *and* (19)*, respectively, while $\theta \in \mathbb{R}$ is a suitable constant, with $0 \leq \theta \leq 1$.*

Proof. Let us suitably rewrite equation (7), after identifying the generic function $\varphi \in V$ with the primal discretization error e_h:

$$J(e_h) = \int_\Omega \nabla e_h \cdot \nabla z \, d\mathbf{x} = \theta \underbrace{\int_\Omega \nabla e_h \cdot \nabla z \, d\mathbf{x}}_{\text{(A)}} + (1-\theta) \underbrace{\int_\Omega \nabla e_h \cdot \nabla z \, d\mathbf{x}}_{\text{(B)}}, \tag{23}$$

with θ to be properly chosen. We analyze separately the two terms (A) and (B) in (23), starting from (A). By exploiting the primal Galerkin orthogonality (6), we get

$$\begin{aligned} \text{(A)} &= \int_\Omega \nabla e_h \cdot \nabla z \, d\mathbf{x} = \int_\Omega \nabla e_h \cdot \nabla (z - v_h) \, d\mathbf{x} \\ &= \int_\Omega \nabla u \cdot \nabla (z - v_h) \, d\mathbf{x} - \int_\Omega \nabla u_h \cdot \nabla (z - v_h) \, d\mathbf{x}, \end{aligned} \tag{24}$$

with $v_h \in V_h$. Now, by choosing $v_h = z_h$ and thanks to the dual Galerkin orthogonality (12), with $\varphi_h = u_h$, we have that the last term in (24) is identically equal to zero. We remark that the choice made above for the functions v_h and φ_h is coherent with the employment of the same discrete space V_h, namely of the same mesh $\mathcal{T}_h$, for both the primal and the dual problems. Thus, via (4), we obtain

$$\text{(A)} = \int_\Omega \nabla u \cdot \nabla e_h^* \, d\mathbf{x} = \int_\Omega f \, e_h^* \, d\mathbf{x}. \tag{25}$$

Now let us consider the term (B). The weak form (10) immediately yields

$$\text{(B)} = \int_\Omega \nabla e_h \cdot \nabla z \, d\mathbf{x} = \int_\Omega g \, e_h \, d\mathbf{x}. \tag{26}$$

Notice that the approach followed to rewrite terms (A) and (B) is different: no Galerkin orthogonality is exploited in the second case. By inserting relations (25) and (26) into (23), we derive

$$J(e_h) = \theta \int_\Omega f \, e_h^* \, d\mathbf{x} + (1-\theta) \int_\Omega g \, e_h \, d\mathbf{x}. \tag{27}$$

The relation above provides us with an exact representation of the error $J(e_h)$. However it is fully useless as depending on the (unknown) exact primal and dual solutions u and z. The idea is to make this expression helpful from a computational viewpoint, by suitably rewriting the right-hand side of (27) in terms of (discrete) computable quantities only. Thanks to the Cauchy-Schwarz inequality, we immediately get

$$|J(e_h)| \leq \theta \, \|f\|_{L^2(\Omega)} \underbrace{\|e_h^*\|_{L^2(\Omega)}}_{\text{(C)}} + (1-\theta) \, \|g\|_{L^2(\Omega)} \underbrace{\|e_h\|_{L^2(\Omega)}}_{\text{(D)}}. \tag{28}$$

The terms (C) and (D) can now be explicitly computed by means of proper *residual based* a posteriori error estimators. In more detail, the (standard) regularity requirements made in the primal and dual weak formulations (4) and (7), respectively ($u, z \in H_0^1(\Omega)$), lead us to resort to residual based a posteriori error estimates for the norms $\|e_h^*\|_{H^1(\Omega)}$ and $\|e_h\|_{H^1(\Omega)}$ (see Remark 2.5 for more details), as it trivially follows that

$$|J(e_h)| \leq \theta \, \|f\|_{L^2(\Omega)} \underbrace{\|e_h^*\|_{H^1(\Omega)}}_{(\widetilde{\mathrm{C}})} + (1-\theta) \, \|g\|_{L^2(\Omega)} \underbrace{\|e_h\|_{H^1(\Omega)}}_{(\widetilde{\mathrm{D}})}. \tag{29}$$

Thanks to inequalities (18) and (16), we get result (22), where C^{**} is defined as in Lemmas 2.3 and 2.4. □

Remark 2.4 *A heuristic recipe to suitably choose a value for the constant θ in* (22) *is suggested in the sequel* (*see* § 4.2).

Up to the constant C^{**}, the right-hand side of inequality (22) provides us with an explicitly evaluable estimator for the functional error $J(e_h)$, given by

$$\eta^1_{D,nested} = \theta \, \|f\|_{L^2(\Omega)} \left(\sum_{K \in \mathcal{T}_h} [\, \rho_K^*(z_h) \,]^2 \right)^{1/2} + (1-\theta) \, \|g\|_{L^2(\Omega)} \left(\sum_{K \in \mathcal{T}_h} [\, \rho_K(u_h) \,]^2 \right)^{1/2}. \tag{30}$$

As well as for the estimation (1), a balance between the contribution of the primal and of the dual problem characterizes such an estimator, the primal and the dual residual terms being weighted by the data (the forcing terms) of the dual and of the primal problem, respectively. Moreover, notice that the explicitly computable nature of estimator (30) is obtained without demanding additional data or the resolution of extra discrete or differential problems.

An alternative *nested dual-residual* a posteriori estimator for the error $J(e_h)$ can be derived by directly bounding the L^2-norms $\|e_h^*\|_{L^2(\Omega)}$ and $\|e_h\|_{L^2(\Omega)}$ in (28). The corresponding result is stated in the following

Proposition 2.2 *Let Ω coincide with a convex polygonal domain. Let u and u_h be the solution of the weak Poisson problem* (4) *and of the corresponding discrete formulation* (5)*, respectively, $e_h = u - u_h$ denoting the associated discretization error. Let $J : V \to \mathbb{R}$ be the goal linear functional we are interested in. Then there exists a positive constant $C_{**} = C_{**}(\widehat{K}, d)$, such that*

$$\begin{aligned} |J(e_h)| \;\; \leq \;\; & C_{**} \left\{ \theta \, \|f\|_{L^2(\Omega)} \left(\sum_{K \in \mathcal{T}_h} [\, h_K \, \rho_K^*(z_h) \,]^2 \right)^{1/2} \right. \\ + \;\; & \left. (1-\theta) \, \|g\|_{L^2(\Omega)} \left(\sum_{K \in \mathcal{T}_h} [\, h_K \, \rho_K(u_h) \,]^2 \right)^{1/2} \right\}, \end{aligned} \tag{31}$$

where the primal and the dual residual terms $\rho_K(u_h)$ and $\rho_K^*(z_h)$ are defined as in (17) *and* (19)*, respectively, while $\theta \in \mathbb{R}$ is a suitable constant, with $0 \leq \theta \leq 1$.*

Proof. We exactly follow the proof of Proposition 2.1 till inequality (28). Then it suffices to estimate the terms (C) and (D) via the residual based error estimates (21) and (20), respectively, the constant C_{**} being defined as in Lemmas 2.5 and 2.6. □

Moving from inequality (31), we have at our disposal another fully computable estimator for the functional error $J(e_h)$, represented, up to the constant C_{**}, by

$$\begin{aligned} \eta^2_{D,nested} &= \theta \, \|f\|_{L^2(\Omega)} \left(\sum_{K \in \mathcal{T}_h} [\, h_K \, \rho^*_K(z_h) \,]^2 \right)^{1/2} \\ &+ (1-\theta) \, \|g\|_{L^2(\Omega)} \left(\sum_{K \in \mathcal{T}_h} [\, h_K \, \rho_K(u_h) \,]^2 \right)^{1/2}. \end{aligned} \tag{32}$$

A balance between the primal and the dual problems characterizes this estimator as well as $\eta^1_{D,nested}$.

Remark 2.5 *Both the choices* (30) *and* (32) *for the error estimator of the functional error $J(e_h)$ have advantages and drawbacks. Essentially, while with the estimator* (32) *we gain an order with respect to the convergence rate, we are implicitly requiring, via the Aubin-Nitsche trick, more regularity on the solutions u and z of the weak problems* (4) *and* (7) *(an H^2-regularity instead of the standard H^1, sufficient to derive $\eta^1_{D,nested}$). On the other hand, moving from* (28) *to* (29)*, we are loosing some information. Thus estimator $\eta^1_{D,nested}$ is, in some sense, defective.*

Remark 2.6 *The dual Galerkin orthogonality* (12) *plays a two-fold role in the analysis above. It is essential to suitably rewrite the term* (A)*, when passing from* (24) *to* (25)*, while, at the same time, it is used to derive both the residual-based estimates in Lemmas* 2.4 *and* 2.6 (*see Appendix* A. 1.).

3 Generalization to an advection-diffusion-reaction problem

Let us generalize the a posteriori analysis in the previous section to the standard scalar advection-diffusion-reaction problem, completed with homogeneous Dirichlet boundary conditions:

$$\begin{cases} -\mu \Delta u + \boldsymbol{\beta} \cdot \nabla u + \alpha \, u = f & \text{in } \Omega \,, \\ u = 0 & \text{on } \partial\Omega \,, \end{cases} \tag{33}$$

where Ω and $\partial\Omega$ are defined as in (3), while the source $f \in L^2(\Omega)$, the diffusivity $\mu \in \mathbb{R}^+$, the advective field $\boldsymbol{\beta} \in (W^{1,\infty}(\Omega))^2$, with $\nabla \cdot \boldsymbol{\beta} = 0$, and the reaction coefficient $\alpha \in L^\infty(\Omega)$, with $\alpha \geq 0$ a.e. in Ω, are given data. What makes

this analysis more complex compared with the one of the previous section is essentially the employment of a stabilized scheme to discretize the problem under investigation.

The weak form associated with (33) is: find $u \in V \equiv H_0^1(\Omega)$ such that, for any $v \in V$,

$$\int_\Omega \mu \nabla u \cdot \nabla v \, d\mathbf{x} + \int_\Omega (\boldsymbol{\beta} \cdot \nabla u + \alpha \, u) v \, d\mathbf{x} = \int_\Omega f \, v \, d\mathbf{x} \, . \tag{34}$$

Remark 3.1 *The requests made on the data of problem* (33) *guarantee both the continuity and the coercivity of the bilinear form identified by the left-hand side of* (34)*. In particular, the coercivity property is exploited in the a posteriori analysis of* § 3.1*, the corresponding coercivity constant being denoted with* c_α*.*

In view of an advection-dominated problem, let us discretize the weak form (34) by projecting onto the space $V_h \subset V$ of continuous piecewise affine finite elements and by stabilizing through the SUPG method [15]. Notice that, due to the choice made for the finite elements, all the standard stabilized techniques actually do coincide with each other. The discrete form thus reads as: find $u_h \in V_h$ such that, for any $v_h \in V_h$,

$$\begin{aligned} & \int_\Omega \mu \nabla u_h \cdot \nabla v_h \, d\mathbf{x} + \int_\Omega (\boldsymbol{\beta} \cdot \nabla u_h + \alpha \, u_h) v_h \, d\mathbf{x} \\ + \quad & \sum_{K \in \mathcal{T}_h} \tau_K \int_K (-\mu \Delta u_h + \boldsymbol{\beta} \cdot \nabla u_h + \alpha \, u_h) \, (\boldsymbol{\beta} \cdot \nabla v_h) \, d\mathbf{x} \\ = \quad & \int_\Omega f \, v_h \, d\mathbf{x} + \sum_{K \in \mathcal{T}_h} \tau_K \int_K f (\boldsymbol{\beta} \cdot \nabla v_h) \, d\mathbf{x} \, , \end{aligned} \tag{35}$$

τ_K denoting the element-wise stabilization parameters defined according to [15]. To shorten the notations, we introduce in the sequel the stabilized bilinear and linear forms $A_\tau : V \times V \to \mathbb{R}$ and $F_\tau : V \to \mathbb{R}$, respectively, given by

$$\begin{aligned} A_\tau(u, v) \quad = \quad & \int_\Omega \mu \nabla u \cdot \nabla v \, d\mathbf{x} + \int_\Omega (\boldsymbol{\beta} \cdot \nabla u + \alpha \, u) v \, d\mathbf{x} \\ & + \sum_{K \in \mathcal{T}_h} \tau_K \int_K (-\mu \Delta u + \boldsymbol{\beta} \cdot \nabla u + \alpha \, u) \, (\boldsymbol{\beta} \cdot \nabla v) \, d\mathbf{x} \, , \\ F_\tau(v) \quad = \quad & \int_\Omega f \, v \, d\mathbf{x} + \sum_{K \in \mathcal{T}_h} \tau_K \int_K f (\boldsymbol{\beta} \cdot \nabla v) \, d\mathbf{x} \, , \end{aligned} \tag{36}$$

with u and v functions smooth enough. Thus problem (35) can be rewritten as: find $u_h \in V_h$ such that, for any $v_h \in V_h$,

$$A_\tau(u_h, v_h) = F_\tau(v_h) \, .$$

Notice that the exact solution u to (34) satisfies $A_\tau(u,v) = F_\tau(v)$, for any $v \in V$, provided that u has extra regularity, namely $u \in V$ with $\Delta u|_K \in L^2(K)$, for any $K \in \mathcal{T}_h$. For example, $u \in H^2(\Omega)$ fits this framework. Whenever this extra regularity is guaranteed, a standard Galerkin orthogonality property (the analogous of (6)) holds, that is we have

$$A_\tau(e_h, v_h) = 0\,, \qquad \forall v_h \in V_h\,, \tag{37}$$

with $e_h = u - u_h$ the discretization error associated with the primal problem (33). In general, if no extra regularity is demanded, u satisfies the equality

$$A_0(u,v) = F_0(v)\,, \qquad \forall v \in V\,, \tag{38}$$

A_0 and F_0 being the non-stabilized bilinear and linear form, respectively, obtained simply by choosing in (36) $\tau_K = 0$, for any $K \in \mathcal{T}_h$. In such a case relation (37) is replaced by the (weaker) Galerkin orthogonality property

$$\begin{aligned} A_0(e_h,\, v_h) &= \int_\Omega \mu \nabla e_h \cdot \nabla v_h \, d\mathbf{x} + \int_\Omega (\boldsymbol{\beta} \cdot \nabla e_h + \alpha\, e_h) v_h \, d\mathbf{x} \\ &= \sum_{K \in \mathcal{T}_h} \tau_K \int_K (-\mu \Delta u_h + \boldsymbol{\beta} \cdot \nabla u_h + \alpha\, u_h - f)\, (\boldsymbol{\beta} \cdot \nabla v_h)\, d\mathbf{x}\,, \end{aligned} \tag{39}$$

obtained simply by subtracting (35) from (34), tested against $v = v_h$, for any $v_h \in V_h$. The extra regularity asked for guaranteeing relation (37) is difficult to obtain in practice. For instance, when mixed boundary conditions complete the advection-diffusion-reaction equation in (33), no H^2-regularity can be assured on u.

Remark 3.2 *Relation* (37) *tacitly implies the stabilization of the weak form* (34)*. This idea may sound a little bit odd, if one is used to associate the concept of stabilization with the discrete setting. However, notice that the extra regularity demanded to have* (37)*, automatically guarantees $A_\tau(u,v) = A_0(u,v)$, for any $u, v \in V$. On the other hand, the stabilization of the weak form is a rather widespread approach adopted for the a posteriori analysis of an advective (or, more in general, a hyperbolic) problem, with the aim of deriving estimates optimal with respect to the convergence rate* (*see, e.g.,* [8, 19, 21])*.*

In view of the a posteriori error analysis of § 3.2, let us introduce the dual problem associated with the weak form (34), given by: find $z \in V$ such that

$$A_0^*(z, \varphi) = J(\varphi) \qquad \forall \varphi \in V, \tag{40}$$

with $J : V \to \mathbb{R}$ the linear functional defining the target quantity we are interested in, and where $A_0^* : V \times V \to \mathbb{R}$ denotes the adjoint form to A_0, such

that $A_0^*(u,v) = A_0(v,u)$, for any u and $v \in V$. As in § 2, let us assume that the quantity $J(\varphi)$ can be represented in the form (8), for a proper choice of the function $g \in L^2(\Omega)$. The differential form of the dual problem (40) can thus be stated as

$$\begin{cases} -\mu\Delta z - \boldsymbol{\beta}\cdot\nabla z + \alpha\, z = g & \text{in } \Omega\,, \\ z = 0 & \text{on } \partial\Omega\,, \end{cases} \tag{41}$$

as well as the weak form (40) rewritten as: find $z \in V$ such that

$$\int_\Omega \mu\nabla z\cdot\nabla\varphi\, d\mathbf{x} - \int_\Omega (\boldsymbol{\beta}\cdot\nabla z - \alpha\, z)\varphi\, d\mathbf{x} = \int_\Omega g\,\varphi\, d\mathbf{x} \qquad \forall\varphi \in V. \tag{42}$$

The analogue of Remark 2.1 still holds in the advective-diffusive-reactive framework. Notice the reverse direction of the dual advective field with respect to the primal one. On the other hand, if the primal problem is advection dominated, the dual problem preserves such a feature. This justifies the employment of the stabilized SUPG method also to discretize problem (42). We get: find $z_h \in V_h$, such that, for any $\varphi_h \in V_h$,

$$\begin{aligned} & \int_\Omega \mu\nabla z_h\cdot\nabla\varphi_h\, d\mathbf{x} - \int_\Omega (\boldsymbol{\beta}\cdot\nabla z_h - \alpha\, z_h)\varphi_h\, d\mathbf{x} \\ + \quad & \sum_{K\in\mathcal{T}_h} \tau_K \int_K (-\mu\Delta z_h - \boldsymbol{\beta}\cdot\nabla z_h + \alpha\, z_h)\,(-\boldsymbol{\beta}\cdot\nabla\varphi_h)\, d\mathbf{x} \\ = \quad & \int_\Omega g\,\varphi_h\, d\mathbf{x} + \sum_{K\in\mathcal{T}_h} \tau_K \int_K g(-\boldsymbol{\beta}\cdot\nabla\varphi_h)\, d\mathbf{x}\,. \end{aligned} \tag{43}$$

Without requiring extra regularity on the dual solution z, a property similar to (39) can be stated in the dual framework. Simply by subtracting (43) from (42), with $\varphi = \varphi_h$, for any $\varphi_h \in V_h$, it follows that

$$\begin{aligned} A_0^*(e_h^*,\,\varphi_h) &= \int_\Omega \mu\nabla e_h^*\cdot\nabla\varphi_h\, d\mathbf{x} - \int_\Omega (\boldsymbol{\beta}\cdot\nabla e_h^* - \alpha\, e_h^*)\varphi_h\, d\mathbf{x} \\ &= \sum_{K\in\mathcal{T}_h} \tau_K \int_K (-\mu\Delta z_h - \boldsymbol{\beta}\cdot\nabla z_h + \alpha\, z_h - g)\,(-\boldsymbol{\beta}\cdot\nabla\varphi_h)\, d\mathbf{x}\,, \end{aligned} \tag{44}$$

with $e_h^* = z - z_h$ the discretization error associated with the dual problem (41). Likewise to relation (37), we can state a standard Galerkin orthogonality property ($A_\tau^*(e_h^*,\varphi_h) = 0$, for any $\varphi_h \in V_h$) also for the dual problem, provided that extra regularity is demanded on the solution z of (42). For the same reasons furnished in the primal framework, we decide to exploit the (weaker) Galerkin orthogonality relations (39) and (44) for the primal and the dual problem, respectively, thus referring to them as to the *primal* and the *dual Galerkin orthogonality*.

Remark 3.3 *The choice of the same discrete space V_h for both the primal and the dual problems leads to a possible weak point of our analysis, i.e., the employment of the same computational grid for two problems with advective fields of opposite direction. However this is not so an unusual approach in the literature (see, for instance,* [8, 19]*). From a computational viewpoint, a reasonable choice is to move from an initial uniform mesh $\mathcal{T}_h$, taking into account the directionalities neither of the primal nor of the dual problem. This is the approach followed in* § 5*.*

3.1 Preliminaries: residual-based a posteriori analysis

In the same fashion as for the Poisson problem, basic ingredients for the a posteriori analysis in § 3.2 are the interpolation error estimates (14) and (15) together with suitable a posteriori residual-based error estimators for the primal and the dual problems (33) and (41), respectively. Since the derivation of these last estimates is less immediate with respect to the corresponding analysis of § 2.2, we provide the corresponding complete proofs just below, without confining them in the Appendix A. 1.

Lemma 3.1 *Let u and u_h be the solution of the advection-diffusion-reaction problem* (34) *and of the corresponding discrete formulation* (35)*, respectively, and let $e_h = u - u_h$ denote the associated discretization error. Then there exists a positive constant $\widetilde{C}^{**} = \widetilde{C}^{**}(\widehat{K}, d, N, c_\alpha)$, with c_α the coercivity constant associated with the bilinear form $A_0(\cdot,\cdot)$ and N defined according to* (13)*, such that the following error estimate holds:*

$$\|e_h\|_{H^1(\Omega)} \leq \widetilde{C}^{**} \left(\sum_{K \in \mathcal{T}_h} [\, \widetilde{\rho}_K(u_h)\,]^2 \right)^{1/2}, \tag{45}$$

where

$$\widetilde{\rho}_K(u_h) = \left(h_K + \tau_K \, \|\boldsymbol{\beta}\|_{L^\infty(K)}\right) \|\widetilde{r}_K(u_h)\|_{L^2(K)} + \frac{1}{2}\, h_K^{1/2}\, \mu \left\| \left[\frac{\partial u_h}{\partial n} \right] \right\|_{L^2(\partial K)} \tag{46}$$

defines the elemental primal residual term, with $\widetilde{r}_K(u_h) = f + \mu\, \Delta u_h - \boldsymbol{\beta} \cdot \nabla u_h - \alpha\, u_h$ the internal residual associated with the primal problem (33)*, and where $[\partial u_h / \partial n]$ stands for the jump of the normal derivative of u_h over the boundary ∂K of the element K.*

Proof. Let us move from the equalities below:

$$
\begin{aligned}
A_0(e_h, v) &= \int_\Omega \mu \nabla e_h \cdot \nabla v \, d\mathbf{x} + \int_\Omega (\boldsymbol{\beta} \cdot \nabla e_h + \alpha \, e_h) v \, d\mathbf{x} \\
&= \int_\Omega \mu \nabla (u - u_h) \cdot \nabla (v - v_h) \, d\mathbf{x} + \int_\Omega \big(\boldsymbol{\beta} \cdot \nabla (u - u_h) + \alpha \, (u - u_h)\big)(v - v_h) \, d\mathbf{x} \\
&+ \sum_{K \in \mathcal{T}_h} \tau_K \int_K (-\mu \Delta u_h + \boldsymbol{\beta} \cdot \nabla u_h + \alpha \, u_h - f) \, (\boldsymbol{\beta} \cdot \nabla v_h) \, d\mathbf{x} \\
&= \int_\Omega f \, (v - v_h) \, d\mathbf{x} - \int_\Omega \mu \nabla u_h \cdot \nabla (v - v_h) \, d\mathbf{x} - \int_\Omega (\boldsymbol{\beta} \cdot \nabla u_h + \alpha \, u_h)(v - v_h) \, d\mathbf{x} \\
&+ \sum_{K \in \mathcal{T}_h} \tau_K \int_K (-\mu \Delta u_h + \boldsymbol{\beta} \cdot \nabla u_h + \alpha \, u_h - f) \, (\boldsymbol{\beta} \cdot \nabla v_h) \, d\mathbf{x} \\
&= \sum_{K \in \mathcal{T}_h} \Big\{ \int_K (f + \mu \, \Delta u_h - \boldsymbol{\beta} \cdot \nabla u_h - \alpha \, u_h) \, (v - v_h) \, d\mathbf{x} \\
&- \tau_K \int_K (f + \mu \, \Delta u_h - \boldsymbol{\beta} \cdot \nabla u_h - \alpha \, u_h) \, (\boldsymbol{\beta} \cdot \nabla v_h) \, d\mathbf{x} - \frac{1}{2} \, \mu \int_{\partial K} \Big[\frac{\partial u_h}{\partial n} \Big] \, (v - v_h) \, d\gamma \Big\},
\end{aligned}
\tag{47}
$$

with $v \in V$, and where we have essentially exploited the primal Galerkin orthogonality (39) together with the weak form (34). Via the Cauchy-Schwarz inequality and thanks to the interpolation error estimates $(15)_1$ and $(15)_2$, we have

$$
\begin{aligned}
|A_0(e_h, v)| &\leq \sum_{K \in \mathcal{T}_h} \Big\{ \|\widetilde{r}_K(u_h)\|_{L^2(K)} \, \|v - v_h\|_{L^2(K)} \\
&+ \frac{1}{2} \, \mu \, \Big\| \Big[\frac{\partial u_h}{\partial n} \Big] \Big\|_{L^2(\partial K)} \, \|v - v_h\|_{L^2(\partial K)} + \tau_K \, \|\widetilde{r}_K(u_h)\|_{L^2(K)} \, \|\boldsymbol{\beta} \cdot \nabla v_h\|_{L^2(K)} \Big\} \\
&\leq \sum_{K \in \mathcal{T}_h} \Big\{ \|\widetilde{r}_K(u_h)\|_{L^2(K)} \, C_1 \, h_K \, |v|_{H^1(\Delta_K)} + \frac{1}{2} \, \mu \, \Big\| \Big[\frac{\partial u_h}{\partial n} \Big] \Big\|_{L^2(\partial K)} \, C_2 \, h_K^{1/2} \, \|v\|_{H^1(\Delta_K)} \\
&+ \tau_K \, \|\widetilde{r}_K(u_h)\|_{L^2(K)} \, \|\boldsymbol{\beta} \cdot \nabla I_h(v)\|_{L^2(K)} \Big\},
\end{aligned}
$$

after choosing $v_h = I_h(v)$ and introducing the definition of the internal residual $\widetilde{r}_K(u_h)$. Let us consider separately the norm $\|\boldsymbol{\beta} \cdot \nabla I_h(v)\|_{L^2(K)}$:

$$
\begin{aligned}
&\|\boldsymbol{\beta} \cdot \nabla I_h(v)\|_{L^2(K)} \leq \|\boldsymbol{\beta}\|_{L^\infty(K)} \Big\{ \|\nabla (I_h(v) - v)\|_{L^2(K)} + \|\nabla v\|_{L^2(K)} \Big\} \\
&\leq \|\boldsymbol{\beta}\|_{L^\infty(K)} \Big\{ C_3 \, |v|_{H^1(\Delta_K)} + \|\nabla v\|_{L^2(K)} \Big\} \leq (C_3 + 1) \, \|\boldsymbol{\beta}\|_{L^\infty(K)} \, \|v\|_{H^1(\Delta_K)},
\end{aligned}
\tag{48}
$$

estimate $(15)_3$ having been used. Thus we get

$$
|A_0(e_h, v)| \leq C \sum_{K \in \mathcal{T}_h} \widetilde{\rho}_K(u_h) \, \|v\|_{H^1(\Delta_K)} \leq C \Big(\sum_{K \in \mathcal{T}_h} [\widetilde{\rho}_K(u_h)]^2 \Big)^{1/2} \Big(\sum_{K \in \mathcal{T}_h} \|v\|^2_{H^1(\Delta_K)} \Big)^{1/2},
$$

with $C = \max(C_1, C_2, C_3 + 1)$, $\widetilde{\rho}_K(u_h)$ defined as in (46), and where the discrete Cauchy-Schwarz inequality has been exploited. Now, as

$$
\Big(\sum_{K \in \mathcal{T}_h} \|v\|^2_{H^1(\Delta_K)} \Big)^{1/2} \leq \sqrt{N} \, \|v\|_{H^1(\Omega)}, \tag{49}
$$

with N defined according to (13), we can state that

$$|A_0(e_h, v)| \leq \widetilde{C}^* \Big(\sum_{K \in \mathcal{T}_h} [\widetilde{\rho}_K(u_h)]^2 \Big)^{1/2} \|v\|_{H^1(\Omega)},$$

with $\widetilde{C}^* = C\sqrt{N}$. Taking $v = e_h$, we obtain

$$|A_0(e_h, e_h)| \leq \widetilde{C}^* \Big(\sum_{K \in \mathcal{T}_h} [\widetilde{\rho}_K(u_h)]^2 \Big)^{1/2} \|e_h\|_{H^1(\Omega)}.$$

Thanks to the coercivity of the bilinear form $A_0(\cdot, \cdot)$, we have

$$c_\alpha \, \|e_h\|^2_{H^1(\Omega)} \leq |A_0(e_h, e_h)| \leq \widetilde{C}^* \Big(\sum_{K \in \mathcal{T}_h} [\widetilde{\rho}_K(u_h)]^2 \Big)^{1/2} \|e_h\|_{H^1(\Omega)}$$

c_α being the coercivity constant, i.e., the desired result (45) after simplifying $\|e_h\|_{H^1(\Omega)}$ and by choosing $\widetilde{C}^{**} = \widetilde{C}^*/c_\alpha$. □

Remark 3.4 *For both the convective and the diffusive dominant regimes, the factor* $\big(h_K + \tau_K \, \|\boldsymbol{\beta}\|_{L^\infty(K)}\big)$ *multiplying in* (46) *the* L^2*-norm of the internal residual* $\widetilde{r}_K(u_h)$ *is an* $\mathcal{O}(h_K)$*,* τ_K *being an* $\mathcal{O}(h_K)$ *and an* $\mathcal{O}(h_K^2)$ *in the convective dominant and in the diffusive dominant case, respectively.*

A residual based a posteriori error estimator with respect to the H^1-norm is required also for the dual problem (41). A proof exactly equal to the one of Lemma 3.1 leads to the following result:

Lemma 3.2 *Let* z *and* z_h *be the solution of the advection-diffusion-reaction problem* (42) *and of the corresponding discrete formulation* (43)*, respectively, and let* $e_h^* = z - z_h$ *denote the associated discretization error. Then there exists a positive constant* $\widetilde{C}^{**} = \widetilde{C}^{**}(\widehat{K}, d, N, c_\alpha)$*, with* c_α *the coercivity constant associated with the bilinear form* $A_0(\cdot, \cdot)$ *and* N *defined according to* (13)*, such that the following error estimate holds:*

$$\|e_h^*\|_{H^1(\Omega)} \leq \widetilde{C}^{**} \left(\sum_{K \in \mathcal{T}_h} [\, \widetilde{\rho}_K^*(z_h) \,]^2 \right)^{1/2}, \tag{50}$$

where

$$\widetilde{\rho}_K^*(z_h) = \big(h_K + \tau_K \, \|\boldsymbol{\beta}\|_{L^\infty(K)}\big) \, \|\widetilde{r}_K^*(z_h)\|_{L^2(K)} + \frac{1}{2} \, h_K^{1/2} \, \mu \, \Big\| \Big[\frac{\partial z_h}{\partial n} \Big] \Big\|_{L^2(\partial K)} \tag{51}$$

defines the elemental dual residual term, with $\widetilde{r}_K^*(z_h) = g + \mu \, \Delta z_h + \boldsymbol{\beta} \cdot \nabla z_h - \alpha \, z_h$ *the internal residual associated with the dual problem* (41)*, and where* $[\partial z_h / \partial n]$ *stands for the jump of the normal derivative of* z_h *over the boundary* ∂K *of the element* K*.*

We are now in a position to supply the *nested a posteriori analysis* for the advection-diffusion-reaction problem (33).

3.2 The a posteriori analysis

The approach pursued in the sequel to estimate the goal quantity $J(u)$ is different from that of § 2.3. Following the *error correction* strategy reviewed in [19], the value $J(u)$ is no longer approximated by $J(u_h)$ but via a new estimator $\widetilde{J}_h = \widetilde{J}_h(u_h, z_h)$, obtained by adding a suitable corrective term $J_h^C(u_h, z_h)$ to the old estimator $J(u_h)$, so that the quantity $\widetilde{J}_h$ represents a better approximation for $J(u)$ with respect to $J(u_h)$. The good property of such an approach is that the corrective term $J_h^C(u_h, z_h)$ is directly yielded by the a posteriori analysis itself.

In more detail, let us introduce the new estimator

$$\widetilde{J}_h = \widetilde{J}_h(u_h, z_h) = J(u_h) + J_h^C(u_h, z_h), \tag{52}$$

with corrective term $J_h^C(u_h, z_h)$ given by

$$J_h^C(u_h, z_h) = \theta \Big[\sum_{K \in \mathcal{T}_h} \tau_K \int_K \widetilde{r}_K(u_h)\,(-\boldsymbol{\beta}\cdot\nabla z_h)\,d\mathbf{x} - \sum_{K \in \mathcal{T}_h} \tau_K \int_K \widetilde{r}_K^*(z_h)\,(\boldsymbol{\beta}\cdot\nabla u_h)\,d\mathbf{x} \Big], \tag{53}$$

where $\theta \in \mathbb{R}$ is a suitable constant, with $0 \le \theta \le 1$.

With the aim of estimating, within a prescribed tolerance, the goal quantity $J(u)$ via the new error estimator $\widetilde{J}_h$, the following result can be stated:

Proposition 3.1 *Let u and u_h be the solution of the advection-diffusion-reaction problem* (34) *and of the corresponding discrete formulation* (35)*, respectively. Let $J : V \to \mathbb{R}$ be the goal linear functional we are interested in and let $\widetilde{J}_h$ denote the estimator of $J(u)$ defined by* (52)*. Then there exists a positive constant $\widetilde{C}^{**} = \widetilde{C}^{**}(\widehat{K}, d, N, c_\alpha)$, with c_α the coercivity constant associated with the bilinear form $A_0(\cdot,\cdot)$ and N defined according to* (13)*, such that the following error estimate holds:*

$$\begin{aligned} |J(u) - \widetilde{J}_h(u_h, z_h)| \;\le\; & \widetilde{C}^{**} \Big\{ \theta\, \|f\|_{L^2(\Omega)} \Big(\sum_{K \in \mathcal{T}_h} [\,\widetilde{\rho}_K^*(z_h)\,]^2 \Big)^{1/2} \\ & + \; (1-\theta)\, \|g\|_{L^2(\Omega)} \Big(\sum_{K \in \mathcal{T}_h} [\,\widetilde{\rho}_K(u_h)\,]^2 \Big)^{1/2} \Big\}, \end{aligned} \tag{54}$$

where $\widetilde{\rho}_K(u_h)$ and $\widetilde{\rho}_K^*(z_h)$ are defined by (46) *and* (51)*, respectively, and $\theta \in \mathbb{R}$ is a suitable constant, with $0 \le \theta \le 1$.*

Proof. First let us properly rewrite the value $J(e_h)$. Thanks to (40) and to the standard relation between the bilinear form $A_0(\cdot,\cdot)$ and the corresponding adjoint form $A_0^*(\cdot,\cdot)$, we get

$$J(e_h) = A_0^*(z, e_h) = \theta\, A_0^*(z, e_h) + (1-\theta)\, A_0^*(z, e_h) = \theta\, \underbrace{A_0(e_h, z)}_{(\widetilde{\mathrm{A}})} + (1-\theta)\, \underbrace{A_0^*(z, e_h)}_{(\widetilde{\mathrm{B}})}, \tag{55}$$

with $0 \leq \theta \leq 1$, constant. The terms $(\widetilde{\mathrm{A}})$ and $(\widetilde{\mathrm{B}})$ are analyzed following two different approaches. Let us move from the term $(\widetilde{\mathrm{A}})$. Thanks to the primal Galerkin orthogonality (39) and by exploiting the weak form (38) and the definition of the internal residual $\widetilde{r}_K(u_h)$ associated with the primal problem (33), we have

$$\begin{aligned}(\widetilde{\mathrm{A}}) &= A_0(e_h, z - v_h) + \sum_{K \in \mathcal{T}_h} \tau_K \int_K (-\mu \Delta u_h + \boldsymbol{\beta} \cdot \nabla u_h + \alpha\, u_h - f)\,(\boldsymbol{\beta} \cdot \nabla v_h)\, d\mathbf{x} \\ &= (f, z - v_h) - A_0(u_h, z - v_h) - \sum_{K \in \mathcal{T}_h} \tau_K \int_K \widetilde{r}_K(u_h)\,(\boldsymbol{\beta} \cdot \nabla v_h)\, d\mathbf{x}.\end{aligned}$$

Now, after choosing $v_h = z_h$, and by using the dual Galerkin orthogonality (44), the definition of the internal residual $\widetilde{r}^*_K(z_h)$ associated with the dual problem (41) and the relation between the bilinear form $A_0(\cdot,\cdot)$ and its adjoint $A_0^*(\cdot,\cdot)$, we derive

$$\begin{aligned}(\widetilde{\mathrm{A}}) &= (f, e_h^*) - A_0^*(e_h^*, u_h) + \sum_{K \in \mathcal{T}_h} \tau_K \int_K \widetilde{r}_K(u_h)\,(-\boldsymbol{\beta} \cdot \nabla z_h)\, d\mathbf{x} \\ &= (f, e_h^*) + \sum_{K \in \mathcal{T}_h} \tau_K \int_K (-\mu \Delta z_h - \boldsymbol{\beta} \cdot \nabla z_h + \alpha\, z_h - g)\,(\boldsymbol{\beta} \cdot \nabla u_h)\, d\mathbf{x} \\ &+ \sum_{K \in \mathcal{T}_h} \tau_K \int_K \widetilde{r}_K(u_h)\,(-\boldsymbol{\beta} \cdot \nabla z_h)\, d\mathbf{x} \\ &= (f, e_h^*) - \sum_{K \in \mathcal{T}_h} \tau_K \int_K \widetilde{r}^*_K(z_h)\,(\boldsymbol{\beta} \cdot \nabla u_h)\, d\mathbf{x} + \sum_{K \in \mathcal{T}_h} \tau_K \int_K \widetilde{r}_K(u_h)\,(-\boldsymbol{\beta} \cdot \nabla z_h)\, d\mathbf{x}.\end{aligned} \tag{56}$$

On the term $(\widetilde{\mathrm{B}})$ we simply exploit relation (42), thus writing

$$(\widetilde{\mathrm{B}}) = (g, e_h). \tag{57}$$

Coming back to equality (55), thanks to relations (56) and (57) and to definition (53), we obtain

$$\begin{aligned}J(e_h) &= \theta \Big[(f, e_h^*) - \sum_{K \in \mathcal{T}_h} \tau_K \int_K \widetilde{r}^*_K(z_h)\,(\boldsymbol{\beta} \cdot \nabla u_h)\, d\mathbf{x} \\ &+ \sum_{K \in \mathcal{T}_h} \tau_K \int_K \widetilde{r}_K(u_h)\,(-\boldsymbol{\beta} \cdot \nabla z_h)\, d\mathbf{x}\Big] + (1 - \theta)\,(g, e_h) \\ &= \theta\,(f, e_h^*) + J_h^C(u_h, z_h) + (1 - \theta)\,(g, e_h).\end{aligned}$$

Via the Cauchy-Schwarz inequality, we get

$$\begin{aligned}|J(u) - \widetilde{J}_h(u_h, z_h)| &\leq \theta\, \|f\|_{L^2(\Omega)}\, \|e_h^*\|_{L^2(\Omega)} + (1 - \theta)\, \|g\|_{L^2(\Omega)}\, \|e_h\|_{L^2(\Omega)} \\ &\leq \theta\, \|f\|_{L^2(\Omega)}\, \|e_h^*\|_{H^1(\Omega)} + (1 - \theta)\, \|g\|_{L^2(\Omega)}\, \|e_h\|_{H^1(\Omega)},\end{aligned}$$

that is result (54) after exploiting the residual based error estimators (50) and (45). □

Remark 3.5 *The corrective term $J_h^C(u_h, z_h)$ defined in* (53) *differs with respect to the error correction introduced in* [19]*, essentially as it depends on both the primal and the dual residuals $\widetilde{r}_K(u_h)$ and $\widetilde{r}_K^*(z_h)$ rather than only on the primal one.*

Up to the constant $\widetilde{C}^{**}$, the right-hand side of inequality (54) provides us with an explicitly evaluable estimator for the functional error $J(u) - \widetilde{J}_h(u_h, z_h)$, given by

$$\begin{aligned} \eta^1_{ADR,nested} &= \theta \, \|f\|_{L^2(\Omega)} \Big(\sum_{K \in \mathcal{T}_h} [\, \widetilde{\rho}^*_K(z_h) \,]^2 \Big)^{1/2} \\ &+ (1-\theta) \, \|g\|_{L^2(\Omega)} \Big(\sum_{K \in \mathcal{T}_h} [\, \widetilde{\rho}_K(u_h) \,]^2 \Big)^{1/2}. \end{aligned} \tag{58}$$

Analogously to the estimators $\eta^p_{D,nested}$, with $p = 1, 2$, defined in (30) and (32), a balance between the contribution of the primal and of the dual problem characterizes $\eta^1_{ADR,nested}$. Moreover the quantity now estimated is not the standard one $J(e_h)$, even if this value can be easily recovered via relation (52).

Remark 3.6 *Inspired by the error estimators $\eta^p_{D,nested}$, with $p = 1, 2$, derived for the Poisson equation, one would expect that the alternative nested error estimator*

$$\begin{aligned} \eta_{ADR,nested} &= \theta \, \|f\|_{L^2(\Omega)} \Big(\sum_{K \in \mathcal{T}_h} [\, h_K \, \widetilde{\rho}^*_K(z_h) \,]^2 \Big)^{1/2} \\ &+ (1-\theta) \, \|g\|_{L^2(\Omega)} \Big(\sum_{K \in \mathcal{T}_h} [\, h_K \, \widetilde{\rho}_K(u_h) \,]^2 \Big)^{1/2} \end{aligned} \tag{59}$$

existed in the advective-diffusive-reactive framework. However, this is not true as, to our knowledge, it has never been proved for problem (33)*, discretized as in* (35)*, a residual-based a posteriori error estimator, with respect to the L^2-norm, of the form*

$$\|e_h\|_{L^2(\Omega)} \leq C \Big(\sum_{K \in \mathcal{T}_h} [\, h_K \, \widetilde{\rho}_K(u_h) \,]^2 \Big)^{1/2},$$

$\widetilde{\rho}_K(u_h)$ being defined according to (46)*. On the other hand, the nonoptimal (with respect to the convergence rate) estimate*

$$\|e_h\|_{L^2(\Omega)} \leq C \Big(\sum_{K \in \mathcal{T}_h} [\, \breve{\rho}_K(u_h) \,]^2 \Big)^{1/2},$$

can be easily derived, with

$$\breve{\rho}_K(u_h) = \big(h_K^2 + (h_K + 1) \, \tau_K \, \|\boldsymbol{\beta}\|_{L^\infty(K)} \big) \, \|\widetilde{r}_K(u_h)\|_{L^2(K)} + \frac{1}{2} \, h_K^{3/2} \, \mu \, \Big\| \Big[\frac{\partial u_h}{\partial n} \Big] \Big\|_{L^2(\partial K)}.$$

This leads to the nonoptimal nested a posteriori estimator for the goal value $J(u) - \widetilde{J}_h(u_h, z_h)$

$$\begin{aligned}\eta^2_{ADR,nested} &= \theta \, \|f\|_{L^2(\Omega)} \Big(\sum_{K \in \mathcal{T}_h} [\, \breve{\rho}^*_K(z_h) \,]^2 \Big)^{1/2} \\ &+ (1-\theta) \, \|g\|_{L^2(\Omega)} \Big(\sum_{K \in \mathcal{T}_h} [\, \breve{\rho}_K(u_h) \,]^2 \Big)^{1/2},\end{aligned}$$

with

$$\breve{\rho}^*_K(z_h) = \left(h_K^2 + (h_K+1)\,\tau_K \, \|\boldsymbol{\beta}\|_{L^\infty(K)}\right) \|\widetilde{r}^*_K(z_h)\|_{L^2(K)} + \frac{1}{2}\, h_K^{3/2}\, \mu \left\| \left[\frac{\partial z_h}{\partial n} \right] \right\|_{L^2(\partial K)}.$$

Estimator $\eta^2_{ADR,nested}$ *preserves the desired optimality only for a diffusion dominant problem, the term* $\left(h_K^2 + (h_K+1)\,\tau_K \, \|\boldsymbol{\beta}\|_{L^\infty(K)}\right)$ *being an* $\mathcal{O}(h_K^2)$*. Thus the derivation of the optimal estimator* (59) *represents an open problem in the advection dominant case.*

Remark 3.7 *The theory developed above for homogeneous Dirichlet boundary conditions can be easily carried over to the nonhomogeneous case by using the standard approach based on the extension of the boundary data into* Ω*. In this case, the four estimators* $\eta^i_{D,nested}$ *and* $\eta^i_{ADR,nested}$*, with* $i = 1, 2$*, can be defined as well.*
More interesting is the case when the primal problem is completed with mixed boundary conditions, i.e., of Dirichlet type on Γ_D *and of Neumann type on* Γ_N*,* Γ_D *and* Γ_N *being two disjoint subsets of* $\partial\Omega$*, such that* $\overline{\Gamma}_D \cup \overline{\Gamma}_N = \partial\Omega$*. In this case, the functional* J *may assume the more general form*

$$J(u) = \int_\Omega g\, u \, d\mathbf{x} + \int_{\Gamma_N} h\, u \, d\gamma,$$

whereby it is possible to control the value of the solution u *also on* Γ_N *via a suitable weigh function* h*. As a consequence, the dual problem enjoys mixed boundary conditions too, on the same partitions* Γ_D *and* Γ_N*, function* h *representing the Neumann data for the dual solution* z [18]*. While the estimators* $\eta^1_{D,nested}$ *and* $\eta^1_{ADR,nested}$ *can be extended to this more general case, it is no longer possible to build the estimators* $\eta^2_{D,nested}$ *and* $\eta^2_{ADR,nested}$*, due to the lack of the elliptic regularity in the presence of mixed boundary conditions.*

4 How to get the new mesh

In this section we discuss the iterative procedure used to drive the mesh adaptivity. For the sake of presentation, we describe this procedure for the Poisson

problem (3) and by choosing a two-dimensional framework. Straightforward modifications are required in the case of the advection-diffusion-reaction problem (33) and of higher spatial dimensions. We first note that the estimators (30) and (32) can be re-written, in compact form, as

$$\eta^p_{D,nested} = \underbrace{\theta \, \|f\|_{L^2(\Omega)} \Bigg(\sum_{K \in \mathcal{T}h} \left[h_K^p \Big(\|r_K^*(z_h)\|_{L^2(K)} + \frac{1}{2} \, h_K^{-1/2} \Big\| \Big[\frac{\partial z_h}{\partial n} \Big] \Big\|_{L^2(\partial K)} \Big) \right]^2 \Bigg)^{1/2}}_{\text{(I)}}$$
$$+ \underbrace{(1-\theta) \, \|g\|_{L^2(\Omega)} \Bigg(\sum_{K \in \mathcal{T}h} \left[h_K^p \Big(\|r_K(u_h)\|_{L^2(K)} + \frac{1}{2} \, h_K^{-1/2} \Big\| \Big[\frac{\partial u_h}{\partial n} \Big] \Big\|_{L^2(\partial K)} \Big) \right]^2 \Bigg)^{1/2}}_{\text{(II)}}, \tag{60}$$

where the internal residuals $r_K(u_h)$, $r_K^*(z_h)$, and the jumps $[\partial u_h/\partial n]$, $[\partial z_h/\partial n]$ are defined as in Lemmas 2.3 and 2.4, while p is set to 1 or 2, according to whether an H^1- or an L^2- residual-based estimate has been adopted. To use this estimator in a *predictive* fashion, we resort to an *isotropic metric based* strategy, discussed in § 4.1, while providing some comments on the choice of θ in § 4.2.

4.1 Metric based mesh adaptivity

The metric is represented by a tensor field $\mathcal{M} : \Omega \to \mathbb{R}^{2\times 2}$, such that all the elements K of a given mesh $\mathcal{T}_h$ are unit equilateral triangles with respect to such a tensor, i.e.,

$$\|e\|_{\mathcal{M}} = \int_0^{|e|} \sqrt{\mathbf{t}^T \mathcal{M}(s) \mathbf{t}} \, d\gamma = 1,$$

with $e \in \partial K$ any edge of K, $|e|$ its length, and $\mathbf{t} \in e$ the unit tangent vector. In the case of an isotropic metric, it holds $\mathcal{M}(\mathbf{x}) = h(\mathbf{x})^{-2}\mathcal{I}$, where $h = h(\mathbf{x})$ is the actual mesh size at point $\mathbf{x}$ and $\mathcal{I}$ is the unit tensor. Thus the elements are approximately equilateral also with respect to the Euclidean metric. We provide the following

Definition 4.1 (Matching Condition) *A mesh $\mathcal{T}_h$ satisfies the matching condition with respect to a given metric field $\mathcal{M}$ if $\|e\|_{\mathcal{M}} = 1$, for any edge $e \in \partial K$, and for any $K \in \mathcal{T}_h$.*

For practical purposes, the metric is approximated by a piecewise constant function over a given triangulation $\mathcal{T}_h$, i.e., $\mathcal{M}\big|_K = \mathcal{M}_K \in \mathbb{R}^{2\times 2}$, and, in the isotropic case, $\mathcal{M}_K = h_K^{-2}\mathcal{I}$, for any $K \in \mathcal{T}_h$, so that the triangles can be thoroughly characterized by their diameter h_K only. More general approaches consider anisotropic meshes, where the elements are no longer restricted to be

equilateral but may become arbitrarily stretched along some directions (see, e.g., [3, 14, 16, 17]).

Concerning the adaptive algorithm used in the numerical test cases, we have adopted the following

`Adaptive Procedure`

i) `set` $k = 0$ `and build the ansatz mesh` $\mathcal{T}_h^{(k)}$;

ii) `compute the solutions` u_h `and` z_h `on` $\mathcal{T}_h^{(k)}$;

iii) `moving from the estimator` $\eta^p_{D,nested}$, `derive the mesh size function` $h = h(\mathbf{x})$ `and in turn the metric` $\mathcal{M}$;

iv) `build the new mesh` $\mathcal{T}_h^{(k+1)}$ `satisfying the matching condition` (4.1) `with respect to` $\mathcal{M}$. `If any stopping criterion is met then return, else set` $k \leftarrow k+1$ `and goto` ii).

In more detail, the adaptive procedure usually starts from a uniform and coarse ansatz mesh $\mathcal{T}_h^{(0)}$. Concerning step `iv)`, two alternative approaches are commonly employed for building each new mesh, either local modification techniques (see, e.g., [12]) or complete remeshing of the domain (see, for example, [17]). In the following we choose this second one.
Finally, about step `iii)`, for an assigned value of θ, once the numerical approximations u_h and z_h have been computed on a given mesh $\mathcal{T}_h^{(k)}$:

1. from each of the terms (I) and (II) in (60), via an *equidistribution criterion*, we obtain a prediction of the new element size, say $h_{K,z}^{(k+1)}$ and $h_{K,u}^{(k+1)}$, where the subscripts K, z and K, u refer to the local mesh sizes derived from the dual solution z_h (term (I)) and from the primal solution u_h (term (II)), respectively;

2. the mesh size $h_K^{(k+1)}$ for the new mesh at step $(k+1)$ is obtained by the linear combination

$$h_K^{(k+1)} = \theta\, h_{K,z}^{(k+1)} + (1-\theta)\, h_{K,u}^{(k+1)}. \tag{61}$$

Remark 4.1 *Alternatively to* (61)*, a metric intersection approach can be used, i.e.,*

$$h_K^{(k+1)} = \min\left\{ h_{K,z}^{(k+1)},\ h_{K,u}^{(k+1)} \right\}.$$

This choice generally yields an excessive mesh refinement compared with the procedure (61)*, as shown in the numerical results of* § 5.

Let us detail point 1., focusing on the term (I), analogous arguments holding for the term (II). In particular we show how to get a metric out of this term. The equidistribution criterion implies that, for each $K \in \mathcal{T}_h$,

$$\eta_K^p = \|f\|_{L^2(\Omega)}\, h_{K,z}^p \left(\|r_K^*(z_h)\|_{L^2(K)} + \frac{1}{2}\, h_K^{-1/2} \left\| \left[\frac{\partial z_h}{\partial n} \right] \right\|_{L^2(\partial K)} \right) = \tau, \qquad (62)$$

τ being a local tolerance. To extract quantitative information from the constraint above, we first scale the term in brackets by the square root of the area of the element K. The rationale for this is that, e.g., $\|r_K^*(z_h)\|_{L^2(K)} = \mathcal{O}(|K|^{1/2})$, at least for a sufficiently small element, the other term $h_K^{-1/2}\|[\partial z_h/\partial n]\|_{L^2(\partial K)}$ behaving similarly. Then we recall that, for an isotropic element, it holds $|K| = C_K h_K^2$, where $C_K = \sqrt{3}/4$. After multiplying and dividing by $|K|^{1/2}$ in (62), we obtain

$$C_K^{1/2}\|f\|_{L^2(\Omega)}\, |K|^{-1/2}\, h_{K,z}^{p+1} \left(\|r_K^*(z_h)\|_{L^2(K)} + \frac{1}{2}\, h_K^{-1/2} \left\| \left[\frac{\partial z_h}{\partial n} \right] \right\|_{L^2(\partial K)} \right) = \tau.$$

This relation, which holds on a given mesh, say on the mesh $\mathcal{T}_h^{(k)}$ at step k, can be used in a predictive fashion by evaluating the term in brackets scaled by the factor $|K|^{-1/2}$ on the current mesh $\mathcal{T}_h^{(k)}$, while regarding the term $h_{K,z}^{p+1}$ as linked to the unknown mesh size $h_{K,z}^{(k+1)}$ of the next mesh $\mathcal{T}_h^{(k+1)}$. Thus we have

$$h_{K,z}^{(k+1)} = \tau^{\frac{1}{p+1}} \left[C_K^{1/2}\|f\|_{L^2(\Omega)}|K|^{-1/2} \left(\|r_K^*(z_h)\|_{L^2(K)} + \frac{1}{2}\, h_K^{-1/2} \left\| \left[\frac{\partial z_h}{\partial n} \right] \right\|_{L^2(\partial K)} \right) \right]^{-\frac{1}{p+1}}. \qquad (63)$$

Analogously, by repeating the same arguments for the term (II), we obtain

$$h_{K,u}^{(k+1)} = \tau^{\frac{1}{p+1}} \left[C_K^{1/2}\|g\|_{L^2(\Omega)}|K|^{-1/2} \left(\|r_K(u_h)\|_{L^2(K)} + \frac{1}{2}\, h_K^{-1/2} \left\| \left[\frac{\partial u_h}{\partial n} \right] \right\|_{L^2(\partial K)} \right) \right]^{-\frac{1}{p+1}}. \qquad (64)$$

Then the actual new mesh size is obtained by the linear combination (61).

We postpone to Appendix A. 2. two further techniques for mesh generation in the framework of constrained optimization. They can be used alternatively to the `Adaptive Procedure` i)-iv) described above.

4.2 The choice of θ

We comment on the choice of the parameter θ, so far being any constant $0 \leq \theta \leq 1$. Since θ weights the dual residual term while $1-\theta$ is related to the primal one, we would like θ to be close to the value 1 when the dual problem plays a major role, and, vice-versa, we expect θ to be nearly equal to 0 when the primal problem is more relevant. A first attempt in this direction leads us to define

$$\theta = \frac{\|g\|_{H^1(\Omega)}}{\|g\|_{H^1(\Omega)} + \|f\|_{H^1(\Omega)}}. \qquad (65)$$

The idea behind this choice is that the H^1-norm measures the "badness" or "roughness" of a function.

For example, we expect the dual problem to be more relevant as far as the computation of the functional J is concerned, when $\|g\|_{H^1(\Omega)} \gg \|f\|_{H^1(\Omega)}$, and in this case (65) correctly yields $\theta \simeq 1$. In more detail, let us assume that the source term f is smooth (the solution u being smooth as well), while the function g is rough, e.g., localized in a small portion of the domain, say Ω_g, the dual solution z being in turn rough too. In such a case, it is reasonable that the evaluation of the functional $J(u)$ in (8) is mostly influenced by the behavior of g in Ω_g. Thus, applying the mesh adaptive procedure, we expect the mesh size $h_{K,u}$ to be larger than $h_{K,z}$ in Ω_g, so that, in this region, the optimal mesh should be dictated by the finer "dual" mesh. This is the case for the metric intersection procedure, as well as for the linear combination one, as $\theta \simeq 1$. On the other hand, following the approach (61), we expect the mesh size to depend, also outside of Ω_g, on the "dual" mesh only, independently of the behavior of the primal solution u. This is not the case of the approach based on the metric intersection, the primal solution u being now not completely ignored. A finer grid will be typically yielded in this second case. Of course, the roles played by the primal and dual problems swap in the opposite case of $\|f\|_{H^1(\Omega)} \gg \|g\|_{H^1(\Omega)}$.

As a second example, let us consider the case where f and g are functions characterized by a similar roughness but localized in different areas of the domain Ω, say Ω_f and Ω_g, respectively, with $\overline{\Omega}_f \cap \overline{\Omega}_g = \emptyset$. In this case, (65) yields $\theta \simeq 0.5$, so that the primal and dual meshes are weighted alike. With reference to the representation (8) of the goal quantity $J(u)$, we can argue that now both the regions Ω_g, due to g, and Ω_f, via u, influence the value $J(u)$, a balancing between the two mesh sizes being thus reasonable.

As a final comment, we point out that the recipe (65) for θ is not to be considered as the "panacea", even if providing satisfactory results, as shown in the numerical tests presented below.

5 Numerical test cases

In this section we collect the results of some 2D test problems for both the Poisson and the advection-diffusion-reaction problems. We refer to these problems as Pi and $ADRi$, respectively, where the integer i counts the progressive number of the test case. The further specification $H1$ or $L2$ is used, essentially in the captions, to indicate whether the H^1- or the L^2- residual-based analysis has been used. In particular, we deal with nontrivial numerical test cases, in the sense that the data, f and g, for the primal and dual problems may both be very nonsmooth functions. This is the more interesting and realistic scenario in view of the applications. For example, a very peaked and narrow function may reasonably approximate a Dirac distribution, so that we are able, via J, of controlling point-wise values. The case where all the data are smooth has been

checked a priori but is not reported in the sequel, as the results are satisfactory and show that the nested dual-residual estimators work well in practice.
Finally, for all the test cases below, the numerical code is fully based on the software `FreeFem++` [20].

5.1 The Poisson problem

Test case P1 With reference to (3), the domain Ω is chosen as the unit square $(0,1)^2$ while the solution u is identified with

$$u(x_1,x_2) = 50\,\exp\big(-10\big[(x_1-0.5)^2+(x_2-0.5)^2\big]\big)\,x_1(1-x_1)x_2(1-x_2),$$

i.e., with a Gaussian centered at (0.5,0.5) modulated by a fourth-order bubble function. This choice corresponds to the source term

$$\begin{aligned} f(x_1,x_2) = \;-\;& 100\,\exp\big(-10\big[(x_1-0.5)^2+(x_2-0.5)^2\big]\big) \\ & \big[x_2(x_2-1)+x_1(x_1-1)\big]\big[200\,x_1x_2(x_1x_2-x_1-x_2+1)-9\big]. \end{aligned}$$

The solution to the dual problem is selected coinciding with the gaussian function

$$z(x_1,x_2) = 10\,\exp\big(-10^4\big[(x_1-0.5)^2+(x_2-0.5)^2\big]\big),$$

obtained by setting

$$\begin{aligned} g(x_1,x_2) \;=\;& 4\cdot 10^5\big(1-10^4\big[(x_1-0.5)^2+(x_2-0.5)^2\big]\big) \\ & \exp\big(-10^4\big[(x_1-0.5)^2+(x_2-0.5)^2\big]\big). \end{aligned}$$

According to (8), this is equivalent to controlling the goal value $J(u)=\int_\Omega g\,u\,d\mathbf{x} = 5.483363104887761\cdot 10^{-1}$ via a highly nonsmooth term g, localized in a region Ω_g around the center of the domain. Notice that both u and z are "confined" inside Ω_g but z is rougher than u.

We employ the a posteriori error estimator $\eta^1_{D,nested}$ defined in (30), following the iterative procedure of § 4.1, and by choosing $\tau=1$ as local tolerance. Table 1 gathers the main information about such a procedure. In particular, from left to right, we show the iteration counter k, the cardinality $\#\mathcal{T}_h^{(k)}$ of the mesh $\mathcal{T}_h^{(k)}$, the actual error $|J(u)-J(u_h)|$ on the functional, the value of the estimator $\eta^1_{D,nested}$, the value of the estimator

$$\eta_{\mathrm{BR}} = \sum_{K\in\mathcal{T}_h} h_K^{-1/2}\rho_K(u_h)\left\|\left[\frac{\partial z_h}{\partial n}\right]\right\|_{L^2(\partial K)} \tag{66}$$

proposed in [8], $\rho_K(u_h)$ and $[\partial z_h/\partial n]$ being defined as in Lemmas 2.3 and 2.4, respectively, and the quantity θ defined in (65), computed via a high-order quadrature rule. The estimator (66), introduced for comparison purposes, is the one, among those proposed in [8], with a structure more similar to $\eta^1_{D,nested}$. The sequence of all the quantities in the table show a very fast saturation trend.

Table 1: Test case P1-H1: main quantities characterizing the adaptive procedure.

k	$\#\mathcal{T}_h^{(k)}$	$\lvert J(u) - J(u_h)\rvert$	$\eta^1_{D,nested}$	η_{BR}	θ
0	800	$4.20821 \cdot 10^{+1}$	$7.03368 \cdot 10^{+3}$	$1.38184 \cdot 10^{+4}$	0.999519
1	5486	$5.96337 \cdot 10^{-3}$	$1.16667 \cdot 10^{+3}$	$4.54094 \cdot 10^{+4}$	0.999532
2	15460	$6.25207 \cdot 10^{-3}$	$1.34155 \cdot 10^{+2}$	$4.78638 \cdot 10^{+4}$	0.999556
3	25304	$7.08442 \cdot 10^{-3}$	$9.47073 \cdot 10^{+1}$	$4.74209 \cdot 10^{+4}$	0.999557
4	25670	$6.73411 \cdot 10^{-3}$	$9.34529 \cdot 10^{+1}$	$4.74321 \cdot 10^{+4}$	0.999557
5	25742	$6.96019 \cdot 10^{-3}$	$9.30096 \cdot 10^{+1}$	$4.74585 \cdot 10^{+4}$	0.999556
6	25816	$6.98817 \cdot 10^{-3}$	$9.26988 \cdot 10^{+1}$	$4.74819 \cdot 10^{+4}$	0.999556
7	25838	$7.10147 \cdot 10^{-3}$	$9.24959 \cdot 10^{+1}$	$4.74714 \cdot 10^{+4}$	0.999556
8	25872	$6.91328 \cdot 10^{-3}$	$9.23436 \cdot 10^{+1}$	$4.74750 \cdot 10^{+4}$	0.999556
9	25884	$7.02271 \cdot 10^{-3}$	$9.22537 \cdot 10^{+1}$	$4.74929 \cdot 10^{+4}$	0.999556

In particular, the stagnation of the values of the cardinalities proves the efficiency of the adopted iterative procedure. On comparing the values of the third and of the fourth columns in Table 1, we observe that the estimator $\eta^1_{D,nested}$ over-estimates the true error by a large factor. Nevertheless it saturates fast, thus guaranteeing that the effectivity index $\eta^1_{D,nested}/|J(u) - J(u_h)|$ tends to a constant value already after few iterations. In this case, a suitable scaling of $\eta^1_{D,nested}$ suffices to provide a correct estimate of the actual error. Notice that the estimator η_{BR} is even more over-estimating and shows an increasing trend. Finally, as expected, the values of θ are very close to one, the "badness" of the dual problem being greater than the one of the primal one.
Figure 1 displays the sequence of the first three meshes, obtained starting from the initial uniform mesh on the left. The successive meshes are not shown as

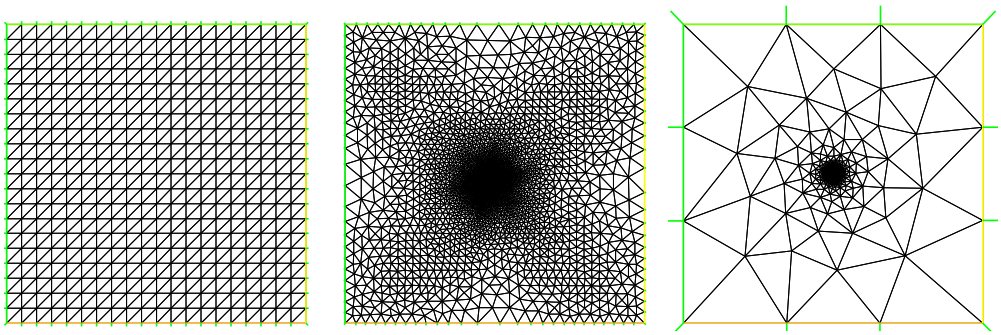

Figure 1: Test case P1-H1. Sequence of adapted meshes: initial (on the left), first adapted (center), and second adapted (on the right).

they actually differ slightly from the second adapted one. The elements of the

adapted meshes correctly cluster around the center of the domain, in correspondence with the domain Ω_g. Figure 2 shows the numerical solutions to the primal (left) and the dual (right) problems, computed on the ninth adapted mesh. The different roughness of the two solutions is clearly highlighted.

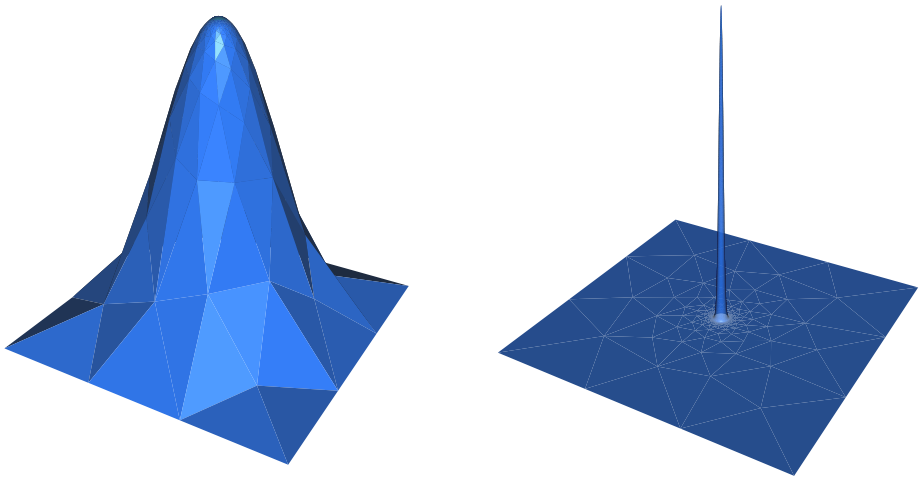

Figure 2: Test case P1-H1: primal (left) and dual (right) solutions on the ninth adapted grid.

The same test case has been solved employing the error estimator $\eta^2_{D,nested}$ defined in (32). The local tolerance is now chosen as $\tau = 10^{-3}$. Table 2 collects the main quantities related to such a procedure, while Figure 3 shows the sequence of the first three meshes. Conclusions similar to those of Table 1 and Figure 1 can be drawn also in this case.

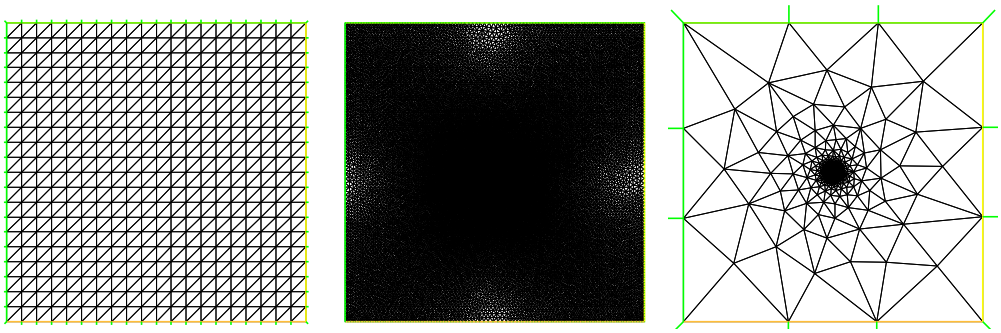

Figure 3: Test case P1-L2. Sequence of adapted meshes: initial (on the left), first adapted (center), and second adapted (on the right).

Table 2: Test case P1-L2: main quantities characterizing the adaptive procedure.

k	$\#\mathcal{T}_h^{(k)}$	$\lvert J(u)-J(u_h)\rvert$	$\eta^2_{D,nested}$	η_{BR}	θ
0	800	$4.2082\cdot 10^{+1}$	$4.97356\cdot 10^{+2}$	$1.38184\cdot 10^{+4}$	0.999519
1	43486	$1.5095\cdot 10^{-2}$	$5.53604\cdot 10^{+0}$	$4.52972\cdot 10^{+4}$	0.999533
2	16452	$2.1530\cdot 10^{-3}$	$3.30031\cdot 10^{+0}$	$4.77086\cdot 10^{+4}$	0.999557
3	18800	$1.6822\cdot 10^{-3}$	$3.70198\cdot 10^{+0}$	$4.75192\cdot 10^{+4}$	0.999554
4	18968	$1.5797\cdot 10^{-3}$	$3.73464\cdot 10^{+0}$	$4.75221\cdot 10^{+4}$	0.999555
5	19024	$1.6322\cdot 10^{-3}$	$3.75909\cdot 10^{+0}$	$4.75642\cdot 10^{+4}$	0.999555
6	19060	$1.6329\cdot 10^{-3}$	$3.75670\cdot 10^{+0}$	$4.75787\cdot 10^{+4}$	0.999555
7	19102	$1.6979\cdot 10^{-3}$	$3.73328\cdot 10^{+0}$	$4.76068\cdot 10^{+4}$	0.999556
8	19134	$1.6767\cdot 10^{-3}$	$3.71768\cdot 10^{+0}$	$4.76207\cdot 10^{+4}$	0.999556
9	19180	$1.6020\cdot 10^{-3}$	$3.71709\cdot 10^{+0}$	$4.76450\cdot 10^{+4}$	0.999556

Test case P2 As in the previous test case, and with reference to the Poisson problem (3), the domain Ω is identified with the unit square $(0,1)^2$ while the solution u is picked as

$$u(x_1,x_2) = 100\,\exp\big(-50\big[(x_1-0.25)^2+(x_2-0.5)^2\big]\big)x_1(1-x_1)x_2(1-x_2),$$

i.e., as a gaussian centered at (0.25,0.5) and modulated by a fourth-degree bubble function. On the other hand, the solution z to the dual problem (9) is chosen so as to have the same structure as u but centered at (0.75,0.5), being given by

$$z(x_1,x_2) = 100\,\exp\big(-50\big[(x_1-0.75)^2+(x_2-0.5)^2\big]\big)x_1(1-x_1)x_2(1-x_2).$$

In more detail, we choose the data for this test case to validate our theory on the interesting case when f and g have equal roughness but localized in two different sub-domains Ω_f and Ω_g of Ω, respectively.
We aim to control the value $J(u)=\int_\Omega g\,u\,d\mathbf{x} = -1.130514581510517$, by means of the error estimator $\eta^1_{D,nested}$ in (30). In particular, we are interested in comparing the results provided by the adaptive procedure based on the linear combination (61) and by the metric intersection based approach (see Remark 4.1), at equal number of mesh elements in the final adapted grid.

Let us start with the adaptive procedure exploiting (61), by choosing $\tau = 0.54$ as the local tolerance. The corresponding main information are summarized in Table 3, the entries preserving the same meaning as in Table 1 and 2. Remarks analogous to those made for the test case *P1* hold. Notice also that, as expected from § 4.2, the values of θ are now very close to 0.5. Moreover, in this case, the estimator η_{BR} is slightly more over-estimating than $\eta^1_{D,nested}$.
Figure 4 displays the initial uniform mesh together with the first two adapted ones. The grid elements essentially thicken in correspondence with the intersection of the supports of the two gaussians. The numerical solutions to the primal

Table 3: Test case P2-H1: main quantities characterizing the adaptive procedure based on (61).

k	$\#\mathcal{T}_h^{(k)}$	$\lvert J(u) - J(u_h)\rvert$	$\eta^1_{D,nested}$	η_{BR}	θ
0	800	$3.36581 \cdot 10^{-2}$	$1.84141 \cdot 10^{+3}$	$1.97887 \cdot 10^{+3}$	0.500000
1	9057	$5.54020 \cdot 10^{-4}$	$1.62847 \cdot 10^{+3}$	$1.80163 \cdot 10^{+3}$	0.497560
2	8218	$8.85588 \cdot 10^{-4}$	$1.77222 \cdot 10^{+3}$	$1.79804 \cdot 10^{+3}$	0.501040
3	8249	$9.12444 \cdot 10^{-4}$	$1.75479 \cdot 10^{+3}$	$1.80104 \cdot 10^{+3}$	0.501011
4	8246	$1.01235 \cdot 10^{-3}$	$1.76347 \cdot 10^{+3}$	$1.79955 \cdot 10^{+3}$	0.501657
5	8257	$8.01808 \cdot 10^{-4}$	$1.68548 \cdot 10^{+3}$	$1.80251 \cdot 10^{+3}$	0.501319
6	8247	$6.63675 \cdot 10^{-4}$	$1.69885 \cdot 10^{+3}$	$1.80322 \cdot 10^{+3}$	0.501185
7	8242	$7.03683 \cdot 10^{-4}$	$1.71205 \cdot 10^{+3}$	$1.80465 \cdot 10^{+3}$	0.500959
8	8242	$7.81432 \cdot 10^{-4}$	$1.74411 \cdot 10^{+3}$	$1.80208 \cdot 10^{+3}$	0.500738
9	8241	$7.69929 \cdot 10^{-4}$	$1.72267 \cdot 10^{+3}$	$1.80093 \cdot 10^{+3}$	0.500632

(left) and the dual (right) problems, computed on the ninth adapted mesh, are provided in Figure 5.

Let us move to the metric intersection based adaptive procedure cited in Remark 4.1. The local tolerance τ is chosen equal to 10 to guarantee, approximately, the same number of triangles in the final adapted mesh. Table 4 is the counterpart of Table 3 for this second approach. Notice that the true error is

Table 4: Test case P2-H1: main quantities characterizing the metric intersection based adaptive procedure.

k	$\#\mathcal{T}_h^{(k)}$	$\lvert J(u) - J(u_h)\rvert$	$\eta^1_{D,nested}$	η_{BR}	θ
0	800	$3.36581 \cdot 10^{-2}$	$1.84141 \cdot 10^{+3}$	$1.97887 \cdot 10^{+3}$	0.500000
1	9342	$3.22127 \cdot 10^{-3}$	$2.81231 \cdot 10^{+2}$	$1.83311 \cdot 10^{+3}$	0.499997
2	8339	$3.15491 \cdot 10^{-3}$	$2.81407 \cdot 10^{+2}$	$1.84629 \cdot 10^{+3}$	0.499992
3	8280	$3.16755 \cdot 10^{-3}$	$2.79663 \cdot 10^{+2}$	$1.84844 \cdot 10^{+3}$	0.499990
4	8264	$3.16807 \cdot 10^{-3}$	$2.78542 \cdot 10^{+2}$	$1.84759 \cdot 10^{+3}$	0.499988
5	8264	$3.16085 \cdot 10^{-3}$	$2.77892 \cdot 10^{+2}$	$1.85035 \cdot 10^{+3}$	0.499990
6	8250	$3.17709 \cdot 10^{-3}$	$2.77580 \cdot 10^{+2}$	$1.85085 \cdot 10^{+3}$	0.499992
7	8243	$3.17312 \cdot 10^{-3}$	$2.77180 \cdot 10^{+2}$	$1.85070 \cdot 10^{+3}$	0.499991
8	8230	$3.18087 \cdot 10^{-3}$	$2.77242 \cdot 10^{+2}$	$1.85112 \cdot 10^{+3}$	0.499991
9	8235	$3.17312 \cdot 10^{-3}$	$2.77015 \cdot 10^{+2}$	$1.85151 \cdot 10^{+3}$	0.499993

larger than the corresponding one in Table 3. On the other hand, the estimator $\eta^1_{D,nested}$, though still over-estimating, turns out to be more accurate when using the metric intersection based procedure. Finally, the values predicted by the

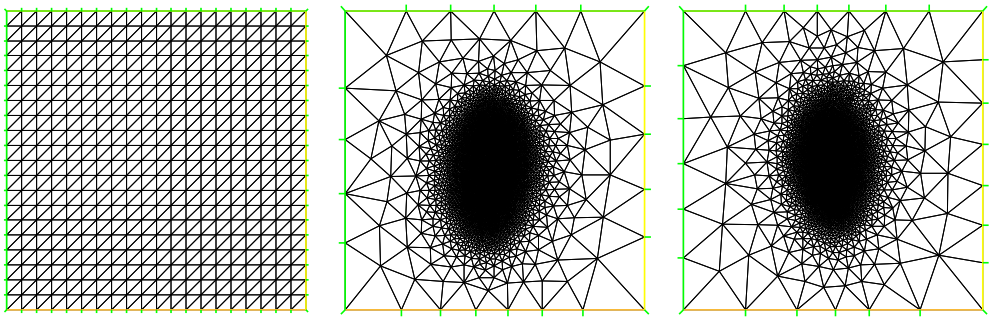

Figure 4: Test case P2-H1. Sequence of adapted meshes for the procedure based on (61): initial (on the left), first adapted (center), and second adapted (on the right).

estimator η_{BR} are even larger than in Table 3.

The initial mesh and the first two adapted ones are shown in Figure 6. A comparison between these grids and the corresponding ones in Figure 4 highlights the different elements distribution: now the essential features of both the primal and dual solutions are captured. Notice that, despite the apparently more reasonable distribution obtained with the metric intersection based procedure, the actual error is smaller using the adaptive technique exploiting the linear combination (61). Finally, in Figure 7 we provide the approximate primal (left) and dual (right) solutions computed on the ninth adapted grid. The different distribution of the mesh elements justifies the higher smoothness of the functions in Figure 7, compared with the corresponding ones in Figure 5. However, this can be deceptive in view of a reduction of the actual error.

5.2 The advection-diffusion-reaction problem

Test case ADR1 Let us move from the reference problem (33), by setting the reactive term zero. We take $\Omega = (0,1)^2 \backslash [0.5,1]^2$, namely an L-shaped domain, $\mu = 10^{-4}$, $\boldsymbol{\beta} = [x_2, -x_1]^T$, while the source term f is chosen such that the exact primal solution is

$$u(x_1, x_2) = x_1(1-x_1)(x_1-0.5)x_2(1-x_2)(x_2-0.5).$$

On the other hand, the forcing term for the dual problem is identified with the function $g = -\Delta w$, where

$$w(x_1, x_2) = 10^5 \exp\big(-10^4\left[(x_1-0.25)^2 + (x_2-0.25)^2\right]\big)$$

is a Gaussian function centered at $(0.25,\ 0.25)$, i.e.,

$$\begin{aligned} g(x_1, x_2) = \ - \ & 4 \cdot 10^9 \exp\big(-10^4\left[(x_1-0.25)^2 + (x_2-0.25)^2\right]\big) \\ & \left[1249 + 10^4 x_1^2 - 5000\, x_1 + 10^4 x_2^2 - 5000\, x_2\right]. \end{aligned}$$

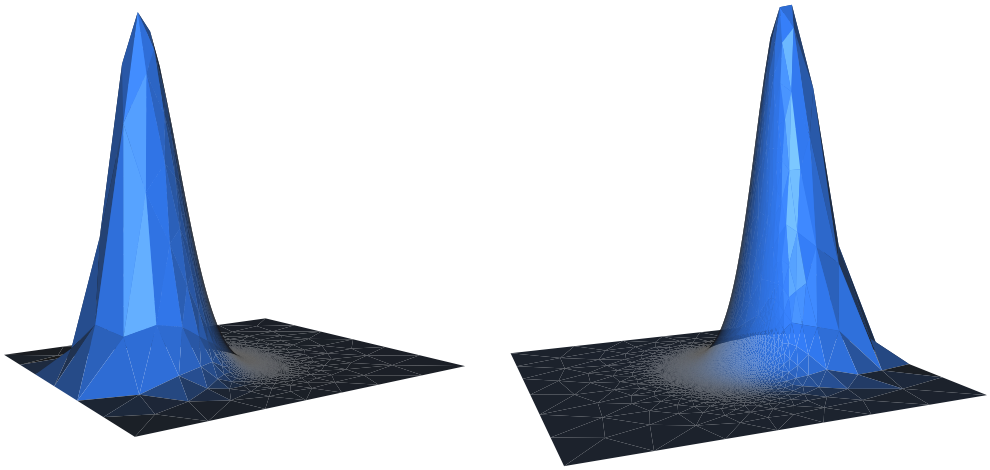

Figure 5: Test case P2-H1. Adaptive procedure based on (61): primal (left) and dual (right) solutions on the ninth adapted grid.

Due to the reverse direction of the dual convective field and to the localized source term g, we expect the dual solution z to be localized around the circular streamlines coming out, in counterclockwise direction, from a neighborhood of the point (0.25,0.25). The goal value we are interested in is, according to the representation (8), $J(u) = \int_\Omega gu\, d\mathbf{x} = 4.414330377375358$. With this aim, we exploit the error estimator $\eta^2_{ADR,nested}$ defined in Remark 3.6, employing the adaptive procedure based on the linear combination (61), with local tolerance $\tau = 10^{-2}$.

Table 5 collects the main quantities characterizing this procedure. The entries preserve the same meaning as in § 5.1, except for the new quantity in fourth column, providing the functional error with respect to the corrected value $\widetilde{J}_h$ defined in (52). No comparison with the estimators defined in [8] is carried out in such a case. All the quantities in the table exhibit a stagnation trend, as in § 5.1. On comparing the values of the third and of the fourth columns, we remark that the error with respect to the corrected functional $\widetilde{J}_h$ is smaller than the standard one $|J(u) - J(u_h)|$, even if not appreciably. This underlines that the corrective term $J^C_h(u_h, z_h)$ in (53) does not play an important role for this test case. Concerning the error estimator $\eta^2_{ADR,nested}$, it under-estimates the true error by about an order of magnitude. Nevertheless, the corresponding fast saturation trend assures that the effectivity index $\eta^2_{ADR,nested}/|J(u)-J(u_h)|$ tends quickly to a constant value. Finally, notice that, as the term g is extremely rough, the values of θ are, in practice, identically equal to one.

Figure 8 displays the initial mesh (left) and the first two adapted ones (center and right), the following grids being nearly identical to the second adapted one. As expected, the mesh elements cluster along the circular streamlines with origin

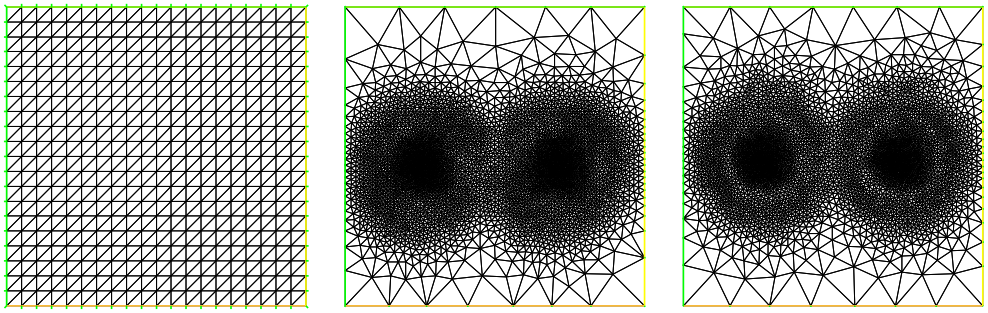

Figure 6: Test case P2-H1. Sequence of adapted meshes for the metric intersection based adaptive procedure: initial (on the left), first adapted (center), and second adapted (on the right).

about the point (0.5,0.5). The solutions to the primal (left) and dual (right) problems, evaluated on the ninth adapted grid, are represented in Figure 9. Notice that, while the main features of the dual solution are very well captured, the accuracy characterizing the approximate primal solution is, undoubtedly, poorer.

For the sake of comparison, we gather in Table 6 the results obtained on a uniform mesh with, approximately, the same number of elements as the ninth adapted grid. Notice that both the functional errors $|J(u) - J(u_h)|$ and $|J(u) - \widetilde{J}_h|$, as well as the estimate furnished by $\eta^2_{ADR,nested}$, are larger than the corresponding values in Table 5. We show in Figure 10 the solutions to the primal (left) and dual (right) problems on the uniform mesh. It is not surprising that the approximate dual solution is less accurate compared with the one in Figure 9, a uniform mesh being not able to detect any significant feature of the considered solution.

Finally, we compare, at equal local tolerance ($\tau = 10^{-2}$), the results in Table 5 and Figure 8 with those yielded by $\eta^2_{ADR,nested}$ when the metric intersection based adaptive procedure is adopted. Table 7 is the counterpart of Table 5: if the values of the functional errors $|J(u) - J(u_h)|$ and $|J(u) - \widetilde{J}_h|$ and of the error estimator $\eta^2_{ADR,nested}$ are similar, the number of the mesh elements is about three times as large when the metric intersection based adaptive algorithm is employed. Also in this case, a non-significant reduction of the error is provided by the corrected functional $\widetilde{J}_h$.

The three grids corresponding to the ones in Figure 8 are gathered in Figure 11: the large number of mesh elements is confirmed by the widespread over-refinement of the adapted grids.

Test case ADR2 We consider the stationary version of a test case in [27], obtained following the procedure below. The full time dependent solution, say

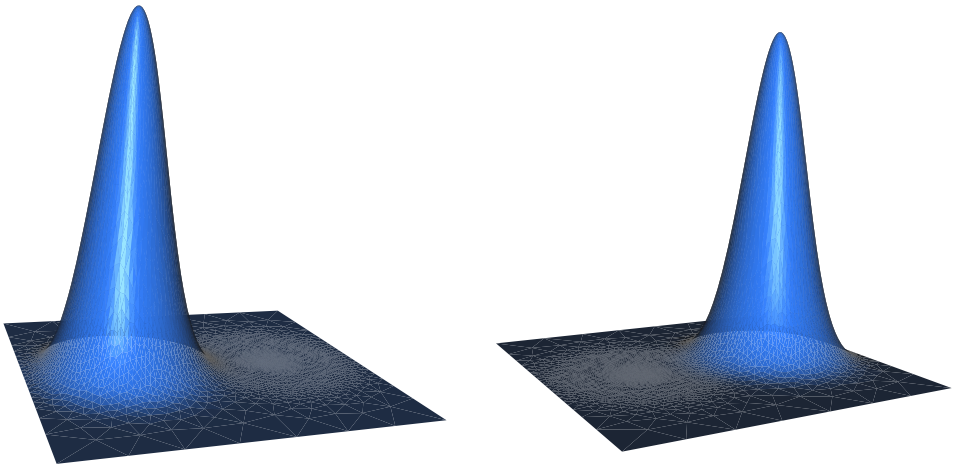

Figure 7: Test case P2-H1. Metric intersection based adaptive procedure: primal (left) and dual (right) solutions on the ninth adapted grid.

u_{NT}, is

$$u_{\mathrm{NT}}(x_1, x_2, t) = \frac{1}{4t+1} \exp\left(- \frac{(x_1 - \beta_1 t - x_0)^2}{k_1(4t+1)} - \frac{(x_2 - \beta_2 t - y_0)^2}{k_2(4t+1)} \right)$$

and solves the equation

$$\frac{\partial u_{\mathrm{NT}}}{\partial t} - k_1 \frac{\partial^2 u_{\mathrm{NT}}}{\partial x_1^2} - k_2 \frac{\partial^2 u_{\mathrm{NT}}}{\partial x_2^2} + \beta_1 \frac{\partial u_{\mathrm{NT}}}{\partial x_1} + \beta_2 \frac{\partial u_{\mathrm{NT}}}{\partial x_2} = 0,$$

on the domain $\Omega = (0, 2)^2$, the initial condition being given by $u_{\mathrm{NT}}(x_1, x_2, 0) = \exp(-(x_1 - x_0)^2/k_1 - (x_2 - y_0)^2/k_2)$, i.e., a gaussian centered at (x_0, y_0). This

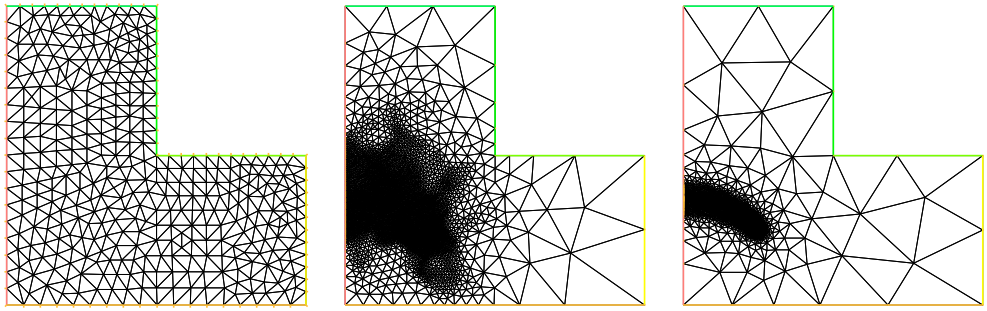

Figure 8: Test case ADR1-L2. Sequence of adapted meshes: initial (on the left), first adapted (center), and second adapted (on the right).

Table 5: Test case ADR1-L2: main quantities characterizing the adaptive procedure.

k	$\#\mathcal{T}_h^{(k)}$	$\lvert J(u)-J(u_h)\rvert$	$\lvert J(u)-\widetilde{J}_h\rvert$	$\eta^2_{ADR,nested}$	θ
0	724	$3.61179\cdot 10^{+2}$	$2.47725\cdot 10^{+2}$	$1.62522\cdot 10^{+2}$	1
1	7962	$2.03342\cdot 10^{+0}$	$2.19336\cdot 10^{+0}$	$2.26353\cdot 10^{+0}$	1
2	6436	$2.09627\cdot 10^{+0}$	$1.85854\cdot 10^{+0}$	$3.42932\cdot 10^{-1}$	1
3	6210	$2.11235\cdot 10^{+0}$	$1.82676\cdot 10^{+0}$	$3.09424\cdot 10^{-1}$	1
4	6314	$2.08382\cdot 10^{+0}$	$1.91844\cdot 10^{+0}$	$3.01776\cdot 10^{-1}$	1
5	6325	$2.07586\cdot 10^{+0}$	$1.91160\cdot 10^{+0}$	$2.77325\cdot 10^{-1}$	1
6	6225	$2.08385\cdot 10^{+0}$	$2.02936\cdot 10^{+0}$	$2.95370\cdot 10^{-1}$	1
7	6322	$2.08841\cdot 10^{+0}$	$2.04413\cdot 10^{+0}$	$2.88449\cdot 10^{-1}$	1
8	6318	$2.08101\cdot 10^{+0}$	$2.01486\cdot 10^{+0}$	$2.86172\cdot 10^{-1}$	1
9	6378	$2.07900\cdot 10^{+0}$	$2.01578\cdot 10^{+0}$	$2.68347\cdot 10^{-1}$	1

Table 6: Test case ADR1-L2: main quantities characterizing the uniform mesh case.

$\#\mathcal{T}_h$	$\lvert J(u)-J(u_h)\rvert$	$\lvert J(u)-\widetilde{J}_h\rvert$	$\eta^2_{ADR,nested}$	θ
6594	$2.46529\cdot 10^{+2}$	$2.45826\cdot 10^{+2}$	$4.07048\cdot 10^{+1}$	1

function is then advected along the 45° diagonal with speed $\boldsymbol{\beta}=[\beta_1,\beta_2]^T$, while it diffuses with coefficients k_1 and k_2 in the x_1 and x_2 direction, respectively, the time interval being $0<t<1.25$. To set our problem, we assume $\beta_1=\beta_2=0.8$, $k_1=k_2=\mu=10^{-2}$, $x_0=y_0=0.5$, and we fix $t=\bar{t}=0.625$, corresponding to the time when the top of $u_{\rm NT}$ reaches the center $(1,1)$ of the domain. Then we identify the solution of problem (33) with $u(x_1,x_2)=u_{\rm NT}(x_1,x_2,\bar{t})$, the right-hand side being computed as

$$f(x_1,x_2) \;=\; -\frac{\partial u_{\rm NT}}{\partial t}\bigg|_{t=\bar{t}} = -\frac{16}{343}\exp\Big(-\frac{200}{7}\big[(x_1-1)^2+(x_2-1)^2\big]\Big)$$
$$\big(200\big[(x_1-0.3)^2+(x_2-0.3)^2\big]-203\big).$$

On the other hand, the source term g for the dual problem (41) is chosen coinciding with the characteristic function of the square $S=(0.5,1.5)^2$, so that we aim at controlling the mean value of u over S, the goal value $J(u)$ being represented by $\int_\Omega gu\,d\mathbf{x}=\int_S u=3.140605942473758\cdot 10^{-2}$.
The adaptive procedure is first driven by the error estimator $\eta^2_{ADR,nested}$ introduced in Remark 3.6, employing the linear combination strategy (61), and with a local tolerance $\tau=10^{-7}$. The main quantities related to this process are collected in Table 8, the meaning of the corresponding entries being preserved with

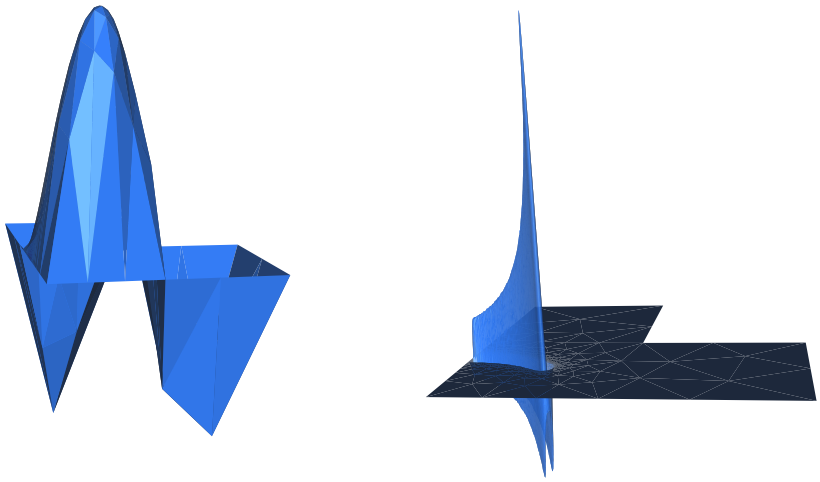

Figure 9: Test case ADR1-L2: primal (left) and dual (right) solutions on the ninth adapted grid.

respect to the previous test case. All the values in the table exhibit a stagnation trend. In particular, we remark that the corrected functional $\widetilde{J}_h$ now provides a reasonably better approximation for the goal quantity $J(u)$, the error $|J(u)-\widetilde{J}_h|$ being about a factor five lower than the standard one $|J(u) - J(u_h)|$. On the other hand, the values yielded by the error estimator $\eta^2_{ADR,nested}$ are of about three orders of magnitude larger compared with the errors $|J(u) - J(u_h)|$ and $|J(u) - \widetilde{J}_h|$, thus over-estimating the considered goal quantity. However, as in the previous test cases, a suitable scaling of $\eta^2_{ADR,nested}$ guarantees a correct prediction for the actual error. Finally, the parameter θ tends to the value 0.82. If, a priori, it is not easy to establish which solution, between u and z, is the roughest, this trend seems to suggest, a posteriori, the greater roughness of the dual solution (in agreement with Figure 13 too).
Figure 12 displays the initial mesh (left) together with the first two adapted ones (center and right). Notice how the elements highlight the boundary of the square S, as a consequence of the choice for g, and the 45° diagonal inside of S, the main transport phenomena, due to both the primal and dual problems, being drawn along this direction. The adapted meshes successive to the second one are not included in Figure 12, being, actually, slightly different.
Figure 13 shows the solutions to the primal (left) and dual (right) problems, computed of the ninth adapted grid. While the primal solution is mostly localized at the center of the domain, the dual solution has a pyramidal structure inside S, because of the source term g and of the downward advection along the 45° diagonal.

The same test case is now solved using the error estimator $\eta^1_{ADR,nested}$ defined in (58), using again the linear combination technique (61), and with a local

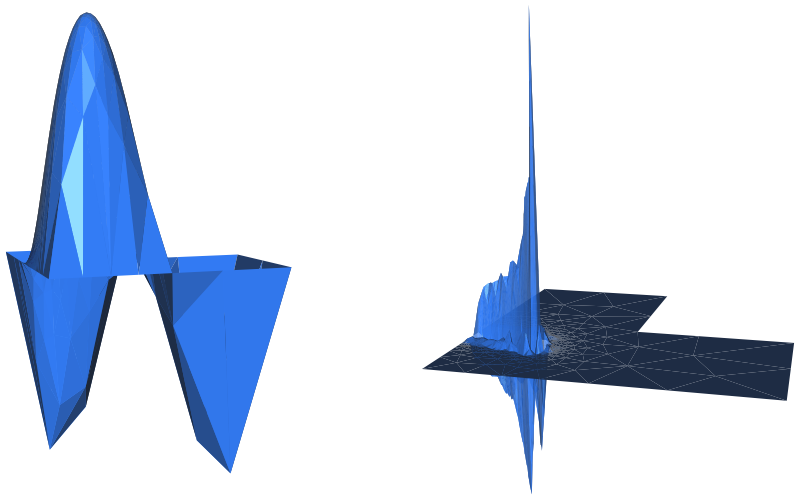

Figure 10: Test case ADR1-L2: primal (left) and dual (right) solutions on a uniform mesh.

tolerance τ now set to 10^{-5}. The corresponding results are summarized in Table 9. Comments analogous to those made for Table 8 hold. In particular, we underline the good quality of the corrected functional $\widetilde{J}_h$, the error $|J(u) - \widetilde{J}_h|$ being now one order of magnitude smaller than the standard one $|J(u) - J(u_h)|$.

Also the mesh elements distribution characterizing the initial mesh and the first two adapted grids in Figure 14 is very similar to the one in Figure 12.

For comparison purposes, we have also validated the adaptive procedure based on the metric intersection approach, still moving from the estimator $\eta^1_{ADR,nested}$, and at equal local tolerance $\tau = 10^{-5}$. However, after computing the second adapted grid, the program runs out of memory, because of the large number of mesh elements (see Table 10 for the few data collected).

Finally, Figure 15 shows the initial mesh together with the first two adapted ones: the elements gather at the boundary of the square S, at the center of the domain, along three diagonals and, mostly, along the lower part of the inflow boundaries stemming from the bottom-left corner of the domain. These zones take into account simultaneously the effects of both the primal and dual problems.

6 Conclusions and future developments

An account of facts In this paper, we have introduced a fully computable dual-based a posteriori error estimator in the framework of a standard elliptic second-order advection-diffusion-reaction problem. However, more complex linear differential problems can be easily taken into account. In contrast with the standard dual-based approach, this estimator does not depend any more on the

Table 7: Test case ADR1-L2: main quantities characterizing the metric intersection based adaptive procedure.

k	$\#\mathcal{T}_h^{(k)}$	$\lvert J(u) - J(u_h)\rvert$	$\lvert J(u) - \widetilde{J}_h\rvert$	$\eta^2_{ADR,nested}$	θ
0	724	$3.61179 \cdot 10^{+2}$	$2.47725 \cdot 10^{+2}$	$1.62522 \cdot 10^{+2}$	1
1	24683	$1.92316 \cdot 10^{+0}$	$2.05020 \cdot 10^{+0}$	$2.33526 \cdot 10^{+0}$	1
2	18850	$2.11551 \cdot 10^{+0}$	$1.85814 \cdot 10^{+0}$	$3.12239 \cdot 10^{-1}$	1
3	19041	$2.09626 \cdot 10^{+0}$	$1.81105 \cdot 10^{+0}$	$2.84783 \cdot 10^{-1}$	1
4	18840	$2.10890 \cdot 10^{+0}$	$1.94381 \cdot 10^{+0}$	$2.82432 \cdot 10^{-1}$	1
5	18810	$2.10232 \cdot 10^{+0}$	$1.81713 \cdot 10^{+0}$	$2.81132 \cdot 10^{-1}$	1
6	18924	$2.09329 \cdot 10^{+0}$	$1.88339 \cdot 10^{+0}$	$2.68156 \cdot 10^{-1}$	1
7	18676	$2.09209 \cdot 10^{+0}$	$1.96082 \cdot 10^{+0}$	$2.81861 \cdot 10^{-1}$	1
8	18577	$2.10036 \cdot 10^{+0}$	$2.01276 \cdot 10^{+0}$	$2.82031 \cdot 10^{-1}$	1
9	18608	$2.08934 \cdot 10^{+0}$	$2.02379 \cdot 10^{+0}$	$2.80160 \cdot 10^{-1}$	1

exact primal and/or dual solutions, as only their Galerkin finite element approximations now play a role. This new estimator joins the main advantages of the dual-based and of the residual-based a posteriori error analysis, being obtained by "nesting" a residual-based estimator into a dual-based one. Thus, according to a classical goal-oriented strategy, suitable functionals of the solution can be, now explicitly, evaluated, at the same cost as the standard dual-based approach, the solution of two differential problems being required. The proposed two-dimensional numerical test cases have validated positively the estimator and the associated adaptive procedure. The mesh adaptivity process, founded on an isotropic metric-based remeshing of the domain, proves to be very effective in all cases, as all the quantities under control show a saturation trend just after few (three/four) iterations. On the other hand, the estimator generally overestimates the goal quantity. Nevertheless, as also the effectivity index of the nested estimator saturates fast, a suitable scaling suffices to guarantee a correct estimate of the quantity of interest.

A look towards the 3D case One possible future development of the issues addressed in this paper consists of extending the nested a posteriori error analysis of § 2 and § 3, together with the corresponding adaptive procedure, to the 3D case. Here we report only on a very preliminary result. The reference problem, denoted as *P3D-L2*, is the Poisson problem (3) defined on the unit cube domain $\Omega = (0,1)^3$, with source term

$$\begin{aligned} f(x_1,x_2,x_3) &= \sqrt{246960}\,\big\{ -6\,x_1^3x_2^3x_3^3(1-x_1)(1-x_2)(1-x_3) \\ &\quad (x_1^2x_2^2 + x_2^2x_3^2 + x_3^2x_1^2) + 6\,x_1^2x_2^2x_3^2\big[x_1x_2(1-x_1)(1-x_2) \\ &+ x_2x_3(1-x_2)(1-x_3) + x_3x_1(1-x_3)(1-x_1)\big]\big\} \end{aligned}$$

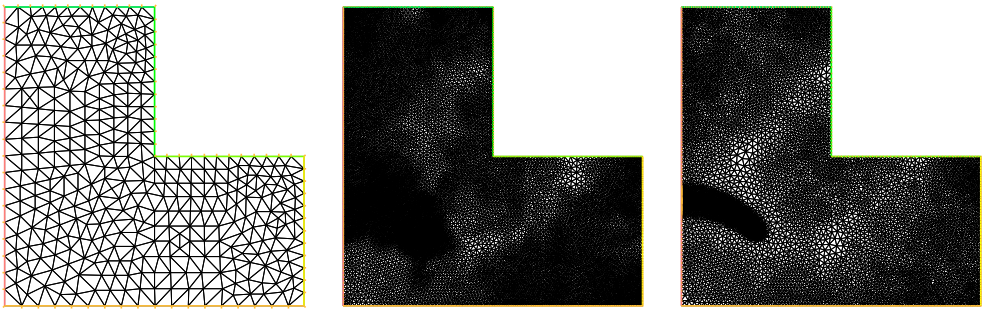

Figure 11: Test case ADR1-L2. Sequence of adapted meshes: initial (left), first adapted (center), and second adapted (right), for the metric intersection based adaptive procedure.

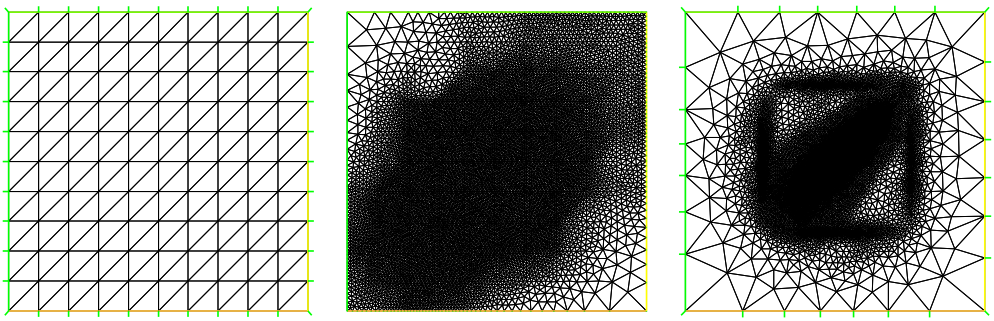

Figure 12: Test case ADR2-L2. Sequence of adapted meshes: initial (left), first adapted (center), and second adapted (right).

so that the exact solution is given by

$$u(x_1, x_2, x_3) = \sqrt{246960}\; x_1^3\, x_2^3\, x_3^3 (1 - x_1)(1 - x_2)(1 - x_3).$$

The solution to the dual problem is chosen coinciding with the primal solution u, so that it holds $z = u$ and $g = f$. This data allow us to control the energy norm of the primal solution since, from the weak form (4) of the primal problem and from the definition (8) of the functional J, it follows

$$J(u) = \int_\Omega g\, u\, d\mathbf{x} = \int_\Omega f\, u\, d\mathbf{x} = \int_\Omega |\nabla u|^2\, d\mathbf{x}.$$

In this case we obtain $J(u) = 1$. Moreover, since the exploited adaptive procedure uses the same computational mesh for both the primal and dual problems, it follows that $u_h = z_h$, that is only one problem, say the primal one, needs

Table 8: Test case ADR2-L2: main quantities characterizing the adaptive procedure.

k	$\#\mathcal{T}_h^{(k)}$	$\lvert J(u) - J(u_h)\rvert$	$\lvert J(u) - \widetilde{J}_h\rvert$	$\eta^2_{ADR,nested}$	θ
0	200	$2.85778 \cdot 10^{-4}$	$2.55089 \cdot 10^{-4}$	$4.17099 \cdot 10^{-2}$	0.587253
1	18067	$5.66531 \cdot 10^{-6}$	$4.15210 \cdot 10^{-6}$	$1.26022 \cdot 10^{-3}$	0.791472
2	15576	$8.17383 \cdot 10^{-5}$	$1.24872 \cdot 10^{-5}$	$6.00322 \cdot 10^{-2}$	0.816543
3	15822	$7.97920 \cdot 10^{-5}$	$1.59511 \cdot 10^{-5}$	$3.84169 \cdot 10^{-2}$	0.819435
4	15728	$8.14190 \cdot 10^{-5}$	$2.22639 \cdot 10^{-5}$	$5.00908 \cdot 10^{-2}$	0.820639
5	15734	$8.12009 \cdot 10^{-5}$	$1.56315 \cdot 10^{-5}$	$5.04778 \cdot 10^{-2}$	0.821605
6	15692	$8.16338 \cdot 10^{-5}$	$1.10712 \cdot 10^{-5}$	$5.28990 \cdot 10^{-2}$	0.822241
7	15740	$8.18816 \cdot 10^{-5}$	$1.24143 \cdot 10^{-5}$	$5.25231 \cdot 10^{-2}$	0.821633
8	15672	$8.16624 \cdot 10^{-5}$	$1.56772 \cdot 10^{-5}$	$4.76920 \cdot 10^{-2}$	0.820837
9	15654	$8.03826 \cdot 10^{-5}$	$1.60333 \cdot 10^{-5}$	$5.09971 \cdot 10^{-2}$	0.820209

to be solved at each iteration. The equivalence between u_h and z_h also implies that the parameter θ in (65) is always 0.5, while the local mesh sizes $h_{K,u}^{(k+1)}$ and $h_{K,z}^{(k+1)}$ computed in (63) and (64), respectively, do coincide, for any $k \geq 0$. This suggests that the two adaptive procedures, the one based on the linear combination (61) and that addressed in Remark 4.1, relying on the metric intersection, yield the same results.
The 3D counterpart of the error estimator $\eta^2_{D,nested}$ defined in (32) has been adopted to drive the adaptive procedure, the local tolerance being set as $\tau = 10^{-2}$. Table 11 collects some quantities related to this adaptive process. From left to right the table shows the iteration counter k, the cardinality $\#\mathcal{T}_h^{(k)}$ of the k-th mesh, the actual functional error $|J(u) - J(u_h)|$, and the value of the estimator $\eta^2_{D,nested}$. Notice that the actual error is over-estimated by a factor of about 30. Finally, Figure 16 displays, in the top row, the initial uniform mesh (left), consisting of $6 \cdot 32^3$ tetrahedra, together with the first adapted grid from two opposite viewpoints (center and right). In the bottom row, the mesh size distribution computed on the initial grid (left) and on the first adapted one (right) is shown. The elements are more refined in correspondence with the corner of coordinates (1,1,1), where the solutions displays the largest curvature.
A more thorough investigation will be the subject of a paper in preparation [25].

Appendix A. 1.

Proof of Lemma 2.3. Let us move from the following chain of equalities

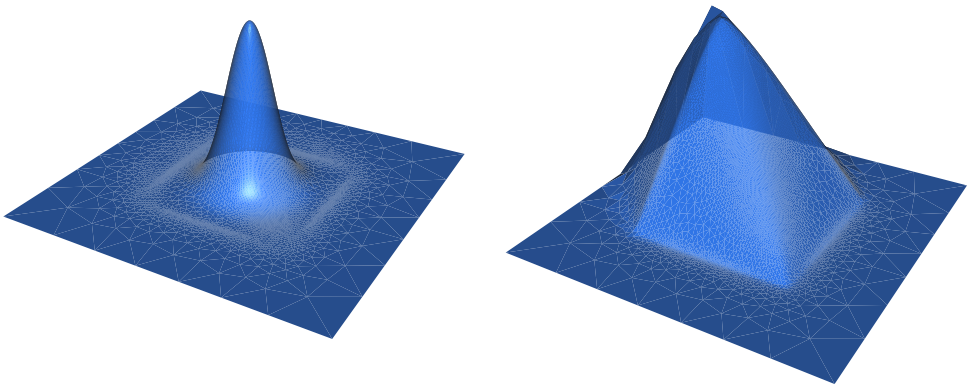

Figure 13: Test case ADR2-L2: primal (left) and dual (right) solutions on the ninth adapted grid.

$$
\begin{aligned}
\int_\Omega \nabla(u-u_h)\cdot\nabla v\,d\mathbf{x} &= \int_\Omega \nabla(u-u_h)\cdot\nabla(v-v_h)\,d\mathbf{x} \\
&= \int_\Omega f\,(v-v_h)\,d\mathbf{x} - \int_\Omega \nabla u_h\cdot\nabla(v-v_h)\,d\mathbf{x} \\
&= \sum_{K\in\mathcal{T}_h}\Big\{\int_K (f+\Delta u_h)\,(v-v_h)\,d\mathbf{x} - \frac{1}{2}\int_{\partial K}\Big[\frac{\partial u_h}{\partial n}\Big]\,(v-v_h)\,d\gamma\Big\},
\end{aligned}
\tag{67}
$$

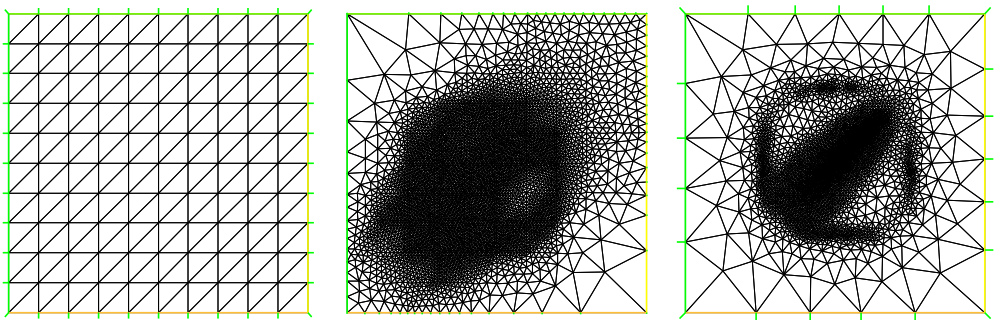

Figure 14: Test case ADR2-H1. Sequence of adapted meshes: initial (left), first adapted (center), and second adapted (right).

Table 9: Test case ADR2-H1: main quantities characterizing the adaptive procedure.

k	$\#\mathcal{T}_h^{(k)}$	$\lvert J(u)-J(u_h)\rvert$	$\lvert J(u)-\widetilde{J}_h\rvert$	$\eta^1_{ADR,nested}$	θ
0	200	$2.85778\cdot 10^{-4}$	$2.55089\cdot 10^{-4}$	$1.47467\cdot 10^{-1}$	0.587253
1	10056	$1.68397\cdot 10^{-5}$	$5.69512\cdot 10^{-7}$	$7.83522\cdot 10^{-2}$	0.766018
2	6786	$1.49120\cdot 10^{-4}$	$2.70413\cdot 10^{-5}$	$1.85350\cdot 10^{-1}$	0.760064
3	6888	$1.51209\cdot 10^{-4}$	$2.80378\cdot 10^{-5}$	$1.73055\cdot 10^{-1}$	0.763476
4	6948	$1.52768\cdot 10^{-4}$	$3.31795\cdot 10^{-5}$	$1.74590\cdot 10^{-1}$	0.763804
5	6912	$1.53964\cdot 10^{-4}$	$3.49269\cdot 10^{-5}$	$1.73883\cdot 10^{-1}$	0.760653
6	6948	$1.52886\cdot 10^{-4}$	$3.09147\cdot 10^{-5}$	$1.72371\cdot 10^{-1}$	0.763004
7	6924	$1.55948\cdot 10^{-4}$	$2.78691\cdot 10^{-5}$	$1.72781\cdot 10^{-1}$	0.762574
8	6900	$1.52257\cdot 10^{-4}$	$2.50379\cdot 10^{-5}$	$1.73106\cdot 10^{-1}$	0.762149
9	6908	$1.49186\cdot 10^{-4}$	$1.78508\cdot 10^{-5}$	$1.73456\cdot 10^{-1}$	0.761349

Table 10: Test case ADR2-H1: main quantities characterizing the metric intersection based adaptive procedure.

k	$\#\mathcal{T}_h^{(k)}$	$\lvert J(u)-J(u_h)\rvert$	$\lvert J(u)-\widetilde{J}_h\rvert$	$\eta^1_{ADR,nested}$	θ
0	200	$2.85778\cdot 10^{-4}$	$2.55089\cdot 10^{-4}$	$1.47467\cdot 10^{-1}$	0.587253
1	20185	$1.23210\cdot 10^{-7}$	$9.02067\cdot 10^{-6}$	$3.35886\cdot 10^{-2}$	0.797567
2	116955	—	—	—	—

with $v\in V$, $v_h\in V_h$, and where the primal Galerkin orthogonality (6) together with the weak form (4) have been essentially exploited. Then thanks to the Cauchy-Schwarz inequality and to the interpolation estimates $(15)_1$ and $(15)_2$, we get

$$
\begin{aligned}
&\Big|\int_\Omega \nabla(u-u_h)\cdot\nabla v\,d\mathbf{x}\Big| \\
&\le \sum_{K\in\mathcal{T}_h}\Big\{\,\|f+\Delta u_h\|_{L^2(K)}\,\|v-v_h\|_{L^2(K)}+\frac{1}{2}\Big\|\Big[\frac{\partial u_h}{\partial n}\Big]\Big\|_{L^2(\partial K)}\,\|v-v_h\|_{L^2(\partial K)}\Big\} \\
&\le \sum_{K\in\mathcal{T}_h}\Big\{\,\|f+\Delta u_h\|_{L^2(K)}\,C_1\,h_K\,|v|_{H^1(\Delta_K)}+\frac{1}{2}\Big\|\Big[\frac{\partial u_h}{\partial n}\Big]\Big\|_{L^2(\partial K)}\,C_2\,h_K^{1/2}\,\|v\|_{H^1(\Delta_K)}\Big\} \\
&\le C\sum_{K\in\mathcal{T}_h}\rho_K(u_h)\,\|v\|_{H^1(\Delta_K)}\le C\Big(\sum_{K\in\mathcal{T}_h}[\rho_K(u_h)]^2\Big)^{1/2}\Big(\sum_{K\in\mathcal{T}_h}\|v\|^2_{H^1(\Delta_K)}\Big)^{1/2},
\end{aligned}
$$

with $C=\max(C_1,C_2)$, $\rho_K(u_h)$ defined as in (17), and where the choice $v_h=I_h(v)$ is done. Moreover the last quantity is yielded via the discrete Cauchy-

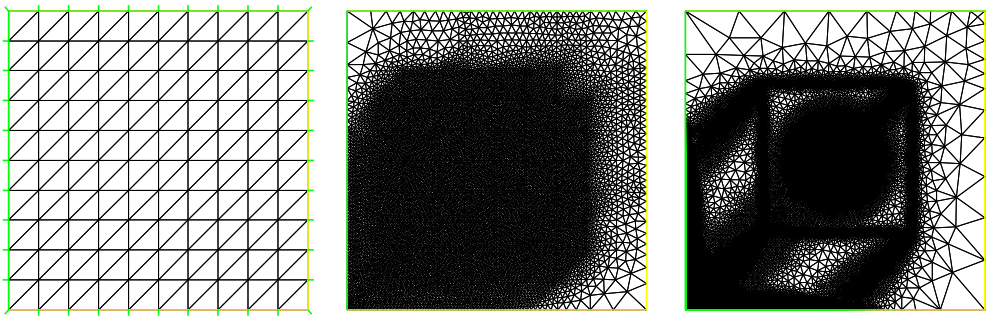

Figure 15: Test case ADR2-H1. Sequence of adapted meshes for the metric intersection based adaptive procedure: initial (left), first adapted (center), and second adapted (right).

Table 11: Test case P3D-L2: main quantities characterizing the adaptive procedure.

k	$\#\mathcal{T}_h^{(k)}$	$\lvert J(u) - J(u_h)\rvert$	$\eta^2_{D,nested}$
0	196608	$1.30770 \cdot 10^{-2}$	$4.19245 \cdot 10^{-1}$
1	138212	$7.99201 \cdot 10^{-3}$	$2.35281 \cdot 10^{-1}$
2	141088	$7.61104 \cdot 10^{-3}$	$2.39810 \cdot 10^{-1}$

Schwarz inequality. Now, thanks to relation (49), we can state that

$$\Big|\int_\Omega \nabla(u-u_h)\cdot\nabla v\, d\mathbf{x}\Big| \le C^* \Big(\sum_{K\in\mathcal{T}_h} [\rho_K(u_h)]^2\Big)^{1/2} \|v\|_{H^1(\Omega)}, \tag{68}$$

where $C^* = C\sqrt{N}$. Then let us choose in (68) $v = e_h \in V$. Thanks to the Poincaré-Friedrichs inequality, we obtain

$$\Big(\frac{1}{1+C_P^2}\Big)\|e_h\|^2_{H^1(\Omega)} \le |||e_h|||^2 \le C^* \Big(\sum_{K\in\mathcal{T}_h} [\rho_K(u_h)]^2\Big)^{1/2} \|e_h\|_{H^1(\Omega)}$$

with $|||\cdot||| \equiv |\cdot|_{H^1(\Omega)}$ the energy norm and C_P the Poincaré constant. After simplifying $\|e_h\|_{H^1(\Omega)}$ and by identifying $C^{**} = C^*(1+C_P^2)$, we recognize the desired result (16). □

Lemma 2.4 can be proved in an identical way.

Proof of Lemma 2.5. By resorting to the Aubin-Nitsche trick, let us introduce the auxiliary problem:

$$\begin{cases} -\Delta\xi = e_h & \text{in } \Omega, \\ \xi = 0 & \text{on } \partial\Omega. \end{cases} \tag{69}$$

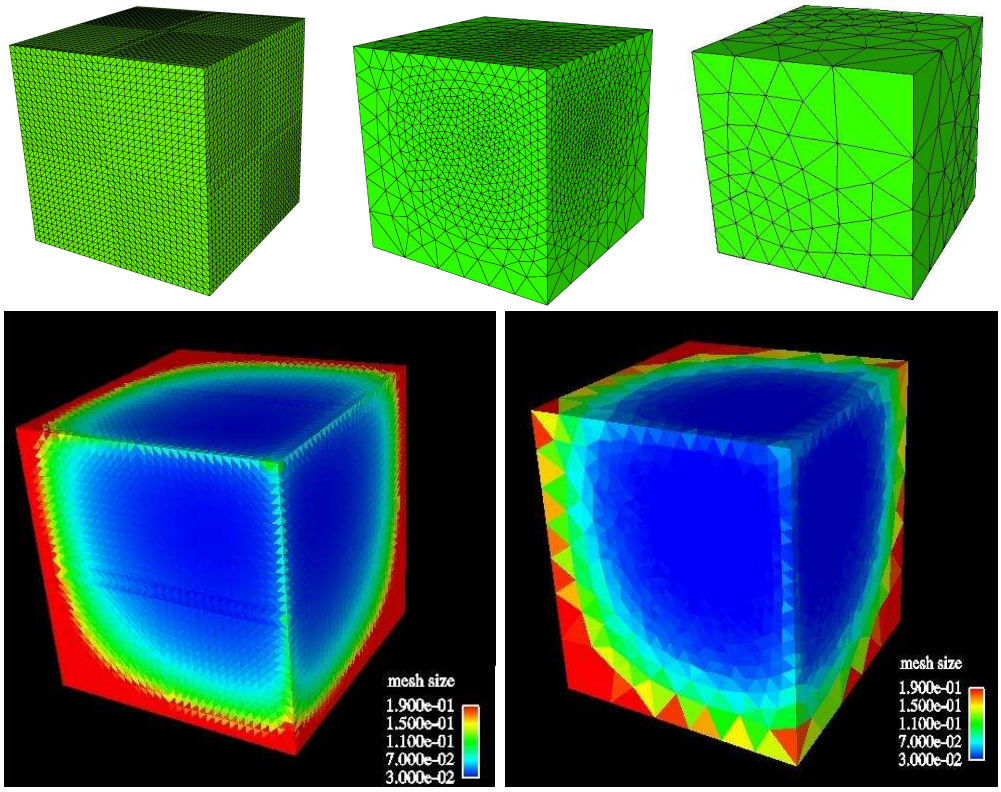

Figure 16: Test case P3D-L2. Top row: initial mesh (on the left) and first adapted grid (center and right) from two different viewpoints. Bottom row: mesh size distribution on the initial grid (left) and on the first adapted one (right).

As Ω is identified with a convex polygonal domain and since $e_h \in L^2(\Omega)$, an *elliptic regularity* property can be stated for problem (69), namely we have that $\xi \in H^2(\Omega)$ and there exists a constant $C_{AN} > 0$, such that

$$\|\xi\|_{H^2(\Omega)} \leq C_{AN} \|e_h\|_{L^2(\Omega)}. \tag{70}$$

Now let us multiply the differential equation in (69) by e_h, while integrating over Ω. We get

$$\begin{aligned}
\|e_h\|^2_{L^2(\Omega)} &= -\int_\Omega \Delta\xi\, e_h\, d\mathbf{x} = \int_\Omega \nabla\xi\cdot\nabla e_h\, d\mathbf{x} = \int_\Omega \nabla(\xi-\xi_h)\cdot\nabla e_h\, d\mathbf{x} \\
&= \int_\Omega f\,(\xi-\xi_h)\, d\mathbf{x} - \int_\Omega \nabla(\xi-\xi_h)\cdot\nabla u_h\, d\mathbf{x} \\
&= \sum_{K\in\mathcal{T}_h}\Big\{\int_K (f+\Delta u_h)\,(\xi-\xi_h)\, d\mathbf{x} - \frac{1}{2}\int_{\partial K}\Big[\frac{\partial u_h}{\partial n}\Big]\,(\xi-\xi_h)\, d\gamma\Big\},
\end{aligned} \tag{71}$$

with $\xi_h \in V_h$, and where integration by parts, together with the primal Galerkin orthogonality (6) and the weak form (4), have been suitably employed. Let us choose $\xi_h = \Pi_h(\xi)$. Notice that the H^2-regularity, guaranteed to the function ξ by the elliptic regularity property, allows us to uniquely define the interpolant $\Pi_h(\xi)$ either in the 2D case or in the 3D one. Via the Cauchy-Schwarz inequality and by exploiting the interpolation estimates (14), we obtain

$$\begin{aligned}
&\|e_h\|^2_{L^2(\Omega)} \\
&\le \sum_{K\in\mathcal{T}_h}\Big\{\|f+\Delta u_h\|_{L^2(K)}\,\|\xi-\Pi_K(\xi)\|_{L^2(K)} + \frac{1}{2}\,\Big\|\Big[\frac{\partial u_h}{\partial n}\Big]\Big\|_{L^2(\partial K)}\,\|\xi-\Pi_K(\xi)\|_{L^2(\partial K)}\Big\} \\
&\le \sum_{K\in\mathcal{T}_h}\Big\{\|f+\Delta u_h\|_{L^2(K)}\,\widetilde{C}_1\,h_K^2\,|\xi|_{H^2(K)} + \frac{1}{2}\,\Big\|\Big[\frac{\partial u_h}{\partial n}\Big]\Big\|_{L^2(\partial K)}\,\widetilde{C}_2\,h_K^{3/2}\,|\xi|_{H^2(K)}\Big\} \\
&\le \widetilde{C}\sum_{K\in\mathcal{T}_h} h_K\,\rho_K(u_h)\,|\xi|_{H^2(K)} \le \widetilde{C}\Big(\sum_{K\in\mathcal{T}_h}[\,h_K\,\rho_K(u_h)]^2\Big)^{1/2}\Big(\sum_{K\in\mathcal{T}_h}|\xi|^2_{H^2(K)}\Big)^{1/2},
\end{aligned}$$

with $\widetilde{C} = \max(\widetilde{C}_1, \widetilde{C}_2)$, $\rho_K(u_h)$ defined as in (17), and where the discrete Cauchy-Schwarz inequality has been used to get the last term. Finally, thanks to relation (70), we derive that

$$\begin{aligned}
\|e_h\|^2_{L^2(\Omega)} &\le \widetilde{C}\Big(\sum_{K\in\mathcal{T}_h}[\,h_K\,\rho_K(u_h)]^2\Big)^{1/2}|\xi|_{H^2(\Omega)} \\
&\le \widetilde{C}\Big(\sum_{K\in\mathcal{T}_h}[\,h_K\,\rho_K(u_h)]^2\Big)^{1/2}\|\xi\|_{H^2(\Omega)} \\
&\le C_{**}\Big(\sum_{K\in\mathcal{T}_h}[\,h_K\,\rho_K(u_h)]^2\Big)^{1/2}\|e_h\|_{L^2(\Omega)},
\end{aligned}$$

namely the desired estimate (20), with $C_{**} = \widetilde{C}\,C_{AN}$. □

We remark that the auxiliary problem (69) is a mere trick to get the desired estimate for the L^2-norm $\|e_h\|_{L^2(\Omega)}$, the corresponding solution ξ being never computed.

Lemma 2.6 can be proved in a similar way, moving from another suitable auxiliary problem.

Appendix A. 2.

For optimal mesh generation, the following two criteria are widely used in the literature (see, e.g., [8, 19]): *i*) for a given number of mesh elements, maximize the accuracy of the approximation; *ii*) for a given accuracy, minimize the number of mesh elements. Let us detail both these procedures moving from the estimator corresponding to the term (I) in (60), analogous conclusions holding for the estimator associated with the term (II).

Let us move from the local error estimator defined in (62)

$$\eta_K = \|f\|_{L^2(\Omega)}\, h_K^p \left(\|r_K^*(z_h)\|_{L^2(K)} + \frac{1}{2}\, h_K^{-1/2} \|R_K(z_h)\|_{L^2(\partial K)} \right)$$

where, to simplify the notation throughout this section, η_K^p is denoted by η_K, $h_{K,z}^p$ by h_K^p, and $R_K(z_h)$ stands for the jump $[\partial z_h/\partial n]$ of the normal derivative of z_h across ∂K. Moreover, we recall that it holds $|K| = C_K h_K^2$, with $C_K = \sqrt{3}/4$. The issue is to compute, over the domain Ω, the optimal mesh size distribution represented by the piecewise constant function $h = h(\mathbf{x})$, with $h(\mathbf{x})\big|_K = h_K$.

Criterion i) The optimal mesh size distribution is given by the solution of the following minimization problem:

$$\begin{gathered} \textit{find } h = h(\mathbf{x}) \textit{ such that} \\ \varepsilon^2 = \sum_{K \in \mathcal{T}_h} \eta_K^2 \textit{ is minimized, subject to the constraint } \left[\int_\Omega \frac{1}{C_K} \frac{1}{h^2}\, d\mathbf{x} \right] = N, \end{gathered} \tag{72}$$

N being the desired number of elements, and $[\,\cdot\,]$ denoting the integer part.

In the following, we compute the solution to this minimization problem, by first re-writing the objective function in a continuous form. With this aim, let us remark that

$$\begin{aligned} \varepsilon^2 &= \sum_{K \in \mathcal{T}_h} \eta_K^2 = \|f\|_{L^2(\Omega)}^2 \sum_{K \in \mathcal{T}_h} h_K^{2p} \Big(\|r_K^*(z_h)\|_{L^2(K)} + \frac{1}{2}\, h_K^{-1/2} \|R_K(z_h)\|_{L^2(\partial K)} \Big)^2 \\ &\leq 2\, \|f\|_{L^2(\Omega)}^2 \sum_{K \in \mathcal{T}_h} h_K^{2p} \Big(\|r_K^*(z_h)\|_{L^2(K)}^2 + \frac{1}{2}\, h_K^{-1} \|R_K(z_h)\|_{L^2(\partial K)}^2 \Big) \\ &= 2\, \|f\|_{L^2(\Omega)}^2 \sum_{K \in \mathcal{T}_h} \int_K h_K^{2p} \Big\{ \big(r_K^*(z_h)\big)^2 + \frac{1}{2}\, h_K^{-1}\, |K|^{-1} \Big(\sum_{e \in \partial K} \big(R_e(z_h)\big)^2\, |e| \Big) \Big\}\, d\mathbf{x}, \end{aligned}$$

where $R_e(z_h) = [\partial z_h/\partial n]\big|_e$, that is coincides with the restriction of the piecewise constant jump term on the edge e of K, for any $e \cap \Omega \neq \emptyset$, while $|e|$ stands for

the length of the edge e. Let us define the weighted average

$$\left(\overline{R}_K(z_h)\right)^2 = \frac{\displaystyle\sum_{e\in\partial K} \left(R_e(z_h)\right)^2 |e|}{\displaystyle\sum_{e\in\partial K} |e|} = \frac{1}{|\partial K|} \sum_{e\in\partial K} \left(R_e(z_h)\right)^2 |e|,$$

with $|\partial K| = 3\,h_K$ the measure of the boundary of K, since all the triangles are supposed to be equilateral (see § 4.1). Thus the objective function can be bounded as

$$\begin{aligned} \varepsilon^2 &\leq 2\,\|f\|^2_{L^2(\Omega)} \sum_{K\in\mathcal{T}_h} \int_K h_K^{2p} \Big\{ \left(r_K^*(z_h)\right)^2 + \frac{1}{2}\, h_K^{-1}\, |K|^{-1}\, |\partial K|\, \left(\overline{R}_K(z_h)\right)^2 \Big\}\, d\mathbf{x} \\ &= 2\,\|f\|^2_{L^2(\Omega)} \sum_{K\in\mathcal{T}_h} \int_K h_K^{2p} \Big\{ \left(r_K^*(z_h)\right)^2 + \frac{1}{2}\, \widetilde{C}_K\, \left(h_K^{-1}\, \overline{R}_K(z_h)\right)^2 \Big\}\, d\mathbf{x}, \end{aligned}$$

where the geometrical relations among $h_K, |K|$ and $|\partial K|$ are exploited, with $\widetilde{C}_K = 3/C_K = 4\sqrt{3}$. We notice that the term $h_K^{-1}\,\overline{R}_K(z_h)$ can be considered as a difference quotient for the second order derivative (across the normal direction to the edge) of the exact dual solution z computed on z_h, so that the above bound for ε^2 can be re-written in a continuous form as

$$\eta(h) = 2\,\|f\|^2_{L^2(\Omega)} \int_\Omega h^{2p} \Big(\left(r^*\right)^2 + \frac{1}{2}\, \widetilde{C}_K\, \left(D^2 z_h\right)^2 \Big)\, d\mathbf{x} = 2\,\|f\|^2_{L^2(\Omega)} \int_\Omega h^{2p}\, \rho^2\, d\mathbf{x}, \tag{73}$$

where $r^* = r^*(\mathbf{x})$ and $D^2 z_h = D^2 z_h(\mathbf{x})$ are such that $r^*\big|_K = r_K^*(z_h)$, $D^2 z_h\big|_K = h_K^{-1}\,\overline{R}_K(z_h)$, respectively, while $\rho = \rho(\mathbf{x})$ is defined via the relation $\rho^2 = \left(r^*\right)^2 + \frac{1}{2}\,\widetilde{C}_K\,\left(D^2 z_h\right)^2$ and can be considered independent of h.

Going back to the minimization problem (72), we are thus led to minimize $\eta(h)$ such that the constraint in (72) is guaranteed.

Proposition 6.1 *The constrained minimization problem* (72) *admits as possible solution the mesh size distribution* $h = h(\mathbf{x})$ *such that*

$$\begin{aligned} h\big|_K = h_K &= \left(\frac{1}{C_K\, N}\right)^{\frac{1}{2}} \left(\sum_{K\in\mathcal{T}_h} \int_K \Big(\left(r_K^*(z_h)\right)^2 + \frac{1}{2}\, \widetilde{C}_K\, \left(h_K^{-1}\, \overline{R}_K(z_h)\right)^2 \Big)^{\frac{1}{p+1}}\, d\mathbf{x} \right)^{\frac{1}{2}} \\ &\quad \left(\left(r_K^*(z_h)\right)^2 + \tfrac{1}{2}\, \widetilde{C}_K\, \left(h_K^{-1}\, \overline{R}_K(z_h)\right)^2 \right)^{-\frac{1}{2(p+1)}}. \end{aligned} \tag{74}$$

Proof. The stationarity of $\eta(h)$ yields

$$\int_\Omega h^{2p-1}\,\rho^2\,\delta h\;d\mathbf{x}=0,$$

for any admissible variation δh such that

$$\delta N=-\frac{2}{C_K}\int_\Omega h^{-3}\,\delta h\,d\mathbf{x}=0.$$

A solution of this constrained minimization problem is

$$h=h(\mathbf{x})=C_N\,\rho(\mathbf{x})^{-\frac{1}{p+1}},$$

for a suitable constant C_N to be determined by the satisfaction of the constraint in (72). This yields

$$\frac{1}{C_K\,C_N^2}\int_\Omega \rho(\mathbf{x})^{\frac{2}{p+1}}\,d\mathbf{x}=N,$$

from which we obtain

$$C_N=\left(\frac{1}{C_K\,N}\right)^{\frac{1}{2}}\left(\int_\Omega \rho(\mathbf{x})^{\frac{2}{p+1}}\,d\mathbf{x}\right)^{\frac{1}{2}}.$$

Going back to the elementwise notation, the optimal mesh size can thus be written as

$$\begin{aligned}
h\big|_K = \quad & h_K=\left(\frac{1}{C_K\,N}\right)^{\frac{1}{2}}\left(\int_\Omega \rho(\mathbf{x})^{\frac{2}{p+1}}\,d\mathbf{x}\right)^{\frac{1}{2}}\rho\big|_K^{-\frac{1}{p+1}}\\
= \quad & \left(\frac{1}{C_K\,N}\right)^{\frac{1}{2}}\left(\int_\Omega\left(\left(r^*\right)^2+\frac{1}{2}\,\widetilde{C}_K\left(D^2z_h\right)^2\right)^{\frac{1}{p+1}}d\mathbf{x}\right)^{\frac{1}{2}}\\
& \left(\left(r^*\right)^2+\frac{1}{2}\,\widetilde{C}_K\left(D^2z_h\right)^2\right)\Big|_K\Bigg)^{-\frac{1}{2(p+1)}}\\
= \quad & \left(\frac{1}{C_K\,N}\right)^{\frac{1}{2}}\left(\sum_{K\in\mathcal{T}_h}\int_K\left(\left(r_K^*(z_h)\right)^2+\frac{1}{2}\,\widetilde{C}_K\left(h_K^{-1}\,\overline{R}_K(z_h)\right)^2\right)^{\frac{1}{p+1}}d\mathbf{x}\right)^{\frac{1}{2}}\\
& \left(\left(r_K^*(z_h)\right)^2+\frac{1}{2}\,\widetilde{C}_K\left(h_K^{-1}\,\overline{R}_K(z_h)\right)^2\right)^{-\frac{1}{2(p+1)}}.
\end{aligned}$$

This relation identifies the desired mesh size distribution. □

Criterion ii) In this case the optimization problem reads as:

$$\textit{find } h=h(\mathbf{x}) \textit{ such that}$$
$$N=\left[\int_\Omega \frac{1}{C_K}\frac{1}{h^2}\,d\mathbf{x}\right] \textit{ is minimized, subject to the constraint } \sum_{K\in\mathcal{T}_h}\eta_K^2=\varepsilon^2, \quad (75)$$

ε being the desired tolerance.

The same notation and arguments as in the previous case hold, provided that the quantity defining the objective function and the constraint are exchanged.

Proposition 6.2 *The constrained minimization problem* (75) *admits as possible solution the mesh size distribution $h = h(\mathbf{x})$ such that*

$$h\big|_K = h_K = \left(\frac{\varepsilon}{2^{\frac{1}{2}}\,\|f\|_{L^2(\Omega)}}\right)^{\frac{1}{p}} \left(\left(r_K^*(z_h)\right)^2 + \frac{1}{2}\,\widetilde{C}_K\left(h_K^{-1}\,\overline{R}_K(z_h)\right)^2\right)^{-\frac{1}{2(p+1)}} \left(\sum_{K\in\mathcal{T}_h}\int_K \left(\left(r_K^*(z_h)\right)^2 + \frac{1}{2}\,\widetilde{C}_K\left(h_K^{-1}\,\overline{R}_K(z_h)\right)^2\right)^{\frac{1}{p+1}} d\mathbf{x}\right)^{-\frac{1}{2p}}. \tag{76}$$

Proof. The stationarity of the number of elements N yields

$$\int_\Omega h^{-3}\,\delta h\,d\mathbf{x} = 0$$

for any admissible variation δh such that

$$\delta\varepsilon^2 = 4\,p\,\|f\|^2_{L^2(\Omega)}\int_\Omega h^{2p-1}\,\rho^2\,\delta h\,d\mathbf{x} = 0,$$

after replacing $\displaystyle\sum_{K\in\mathcal{T}_h}\eta_K^2$ by the continuous quantity $\eta(h)$ defined in (73). A solution to this constrained minimization problem is

$$h = h(\mathbf{x}) = C_\varepsilon\,\rho(\mathbf{x})^{-\frac{1}{p+1}},$$

for a suitable constant C_ε to be determined by satisfying the constraint in (75). We obtain

$$2\,\|f\|^2_{L^2(\Omega)}\,C_\varepsilon^{2p}\int_\Omega \rho(\mathbf{x})^{\frac{2}{p+1}}\,d\mathbf{x} = \varepsilon^2,$$

from which we get

$$C_\varepsilon = \left(\frac{\varepsilon}{2^{\frac{1}{2}}\,\|f\|_{L^2(\Omega)}}\right)^{\frac{1}{p}}\left(\int_\Omega \rho(\mathbf{x})^{\frac{2}{p+1}}\,d\mathbf{x}\right)^{-\frac{1}{2p}}.$$

Using the elementwise notation, the optimal mesh size can thus be written as

$$
\begin{aligned}
h\big|_K = \qquad & h_K = \left(\frac{\varepsilon}{2^{\frac{1}{2}}\,\|f\|_{L^2(\Omega)}}\right)^{\frac{1}{p}} \left(\int\limits_{\Omega} \rho(\mathbf{x})^{\frac{2}{p+1}}\, d\mathbf{x}\right)^{-\frac{1}{2p}} \rho\big|_K^{-\frac{1}{p+1}} \\
= \qquad & \left(\frac{\varepsilon}{2^{\frac{1}{2}}\,\|f\|_{L^2(\Omega)}}\right)^{\frac{1}{p}} \left(\int\limits_{\Omega} \left(\left(r^*\right)^2 + \frac{1}{2}\,\widetilde{C}_K\left(D^2 z_h\right)^2\right)^{\frac{1}{p+1}} d\mathbf{x}\right)^{-\frac{1}{2p}} \\
& \left(\left(\left(r^*\right)^2 + \frac{1}{2}\,\widetilde{C}_K\left(D^2 z_h\right)^2\right)\Big|_K\right)^{-\frac{1}{2(p+1)}} \\
= \qquad & \left(\frac{\varepsilon}{2^{\frac{1}{2}}\,\|f\|_{L^2(\Omega)}}\right)^{\frac{1}{p}} \left(\sum_{K\in\mathcal{T}_h} \int\limits_{K} \left(\left(r_K^*(z_h)\right)^2 + \frac{1}{2}\,\widetilde{C}_K\left(h_K^{-1}\,\overline{R}_K(z_h)\right)^2\right)^{\frac{1}{p+1}} d\mathbf{x}\right)^{-\frac{1}{2p}} \\
& \left(\left(r_K^*(z_h)\right)^2 + \frac{1}{2}\,\widetilde{C}_K\left(h_K^{-1}\,\overline{R}_K(z_h)\right)^2\right)^{-\frac{1}{2(p+1)}}.
\end{aligned}
$$

The desired mesh size distribution is thus identified. □

Remark 6.1 *A procedure related to criterion ii), and known as the optimized mesh strategy, is considered, for instance, in* [8] (*see Section* 5.2) *and* [19] (*see pp.* 197–199). *In such a case, the optimal mesh size distribution is computed via a different approach* (*a Lagrangian technique*) *but coincides with* (76).

Remark 6.2 *Except for the different constants, the dependence on the residual terms of the optimal mesh size distributions* (74) *and* (76) *is the same as the one obtained in* § 4.1, *moving from the equidistribution criterion. This seems to be a feature of the isotropic mesh adaption and may not hold in the case of anisotropic mesh adaptivity.*

Remark 6.3 *The analysis of Propositions* 6.1 *and* 6.2 *can be extended, in a straightforward way, to the 3D case, though the corresponding implementation is more demanding* ([25]).

Acknowledgments. We gratefully acknowledge Dr. M. Prosi for carrying out the 3D numerical computations by the finite element library `LifeV` (`www.lifev.org`), and the automatic 3D tetrahedral mesh generator `Netgen` (`www.hpfem.jku.at/netgen`).

References

[1] M. Ainsworth and J.T. Oden, *A posteriori error estimation in finite element analysis*, John Wiley & Sons, Inc., New-York, 2000.

[2] T. Apel, *Interpolation of non-smooth functions on anisotropic finite element meshes*, M2AN Math. Model. Numer. Anal., 33 (1999), pp. 1149–1185.

[3] T. Apel, *Anisotropic finite elements: local estimates and applications*, Book Series: Advances in Numerical Mathematics, Teubner, Stuttgart, 1999.

[4] J.P. Aubin, *Behavior of the error of the approximate solutions of boundary value problems for linear elliptic operators by Galerkin's and finite difference methods*, Ann. Scuola Norm. Sup. Pisa, 21 (1967), pp. 599–637.

[5] I. Babuška and W. Rheinboldt, *A posteriori error estimates for the finite element method*, Int. J. Numer. Methods Eng., 12 (1978), pp. 1597-1615.

[6] R.E. Bank and A. Weiser, *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comp., 44 (1985), 283-301.

[7] R. Becker and R. Rannacher, *A feed-back approach to error control in finite element methods: basic analysis and examples*, East-West J. Numer. Math., 4 (1996), pp. 237–264.

[8] R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta Numerica, 10 (2001), pp. 1–102.

[9] Ph. Ciarlet, *The finite element method for elliptic problems*, North-Holland Publishing Company, Amsterdam, 1978.

[10] Ph. Clément, *Approximation by finite element functions using local regularization*, RAIRO Anal. Numér., 2 (1975), pp. 77–84.

[11] L. Dedè and S. Perotto, in preparation.

[12] V. Dolejší, *Anisotropic mesh adaptation for finite volume and finite element methods on triangular meshes*, Comput. Vis. Sci., 1 (1998), pp. 165–178.

[13] K. Eriksson, D. Estep, P. Hansbo and C. Johnson, *Introduction to adaptive methods for differential equations*, Acta Numerica, 4 (1995), 105-158.

[14] L. Formaggia, S. Micheletti and S. Perotto, *Anisotropic mesh adaptation in Computational Fluid Dynamics: application to the advection-diffusion-reaction and the Stokes problems*, Appl. Numer. Math., 51 (2004), pp. 511–533.

[15] L.P. Franca, S.L. Frey and T.J.R. Hughes, *Stabilized finite element methods. I. Application to the advective-diffusive model*, Comput. Methods Appl. Mech. Engrg., 95 (1992), pp. 253–276.

[16] P.J. Frey and P.L. George, *Mesh generation. Application to finite elements*, Hermes Science, Paris, 2000

[17] P.L. George and H. Borouchaki, *Delaunay triangulation and meshing-application to finite element*, Editions Hermes, Paris, 1998.

[18] M.B. Giles and N.A. Pierce, *Adjoint equations in CFD: duality, boundary conditions and solution behaviour*, AIAA Paper, 97-1850 (1997).

[19] M.B. Giles and E. Süli, *Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality*, Acta Numerica, 11 (2002), pp. 145-236.

[20] F. Hecht, O. Pironneau and K. Ohtsuka, *FreeFem++ manual – version 1.45-7*, `http://www.ann.jussieu.fr/~hecht/freefem++.htm`, 2005.

[21] P. Houston, R. Rannacher and E. Süli, *A posteriori error analysis for stabilised finite element approximations of transport problems*, Comput. Methods Appl. Mech. Engrg., 190 (2000), pp. 1483-1508.

[22] J.L. Lions and E. Magenes, *Non-Homogeneous boundary value problem and application*, Volume I, Springer-Verlag, Berlin, 1972.

[23] L. Machiels, Y. Maday and A.T. Patera, *Output bounds for reduced-order approximations of elliptic partial differential equations*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 3413–3426.

[24] Y. Maday and A.T. Patera, *Numerical analysis of a posteriori finite element bounds for linear functional outputs*, M3AS Math. Models Methods Appl. Sci., 10 (2000), pp. 785–799.

[25] S. Micheletti, S. Perotto and M. Prosi, in preparation.

[26] J.A. Nitsche, *Ein kriterium für die quasi-optimalitat des Ritzchen Verfahrens*, Numer. Math., 11 (1968), pp. 346–348.

[27] B.J. Noye and H.H. Tan, *Finite difference methods for solving the two dimensional advection-diffusion equation*, Int. J. Numer. Meth. Fl., 9 (1989), pp. 75–98.

[28] J.T. Oden and S. Prudhomme, *Goal-oriented error estimation and adaptivity for the finite element method*, Comput. Math. Appl., 41 (2001), pp. 735–756.

[29] 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 (1997), pp. 289–312.

[30] L.R. Scott and S. Zhang, *Finite element interpolation of non-smooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.

[31] D.A. Venditti and D.L. Darmofal, *Adjoint error estimation and grid adaptation for functional outputs: application to quasi-one-dimensional flow*, J. Comput. Phys., 164 (2000), pp. 204–227.

[32] R. Verfürth, *A review of a posteriori error estimation and adaptive mesh-refinement techniques*, Wiley - Teubner, New York, 1996.

[33] O.C. Zienkiewicz and J.Z. Zhu, *A simple error estimator and adaptive procedure for practical engineering analysis*, Int. J. Numer. Methods Eng., 24 (1987), pp. 337–357.

[34] O.C. Zienkiewicz and J.Z. Zhu, *The superconvergent patch recovery and a posteriori error estimates, Part 1: the recovery technique*, Int. J. Numer. Methods Eng., 33 (1992), pp. 1331–1364.

[35] O.C. Zienkiewicz and J.Z. Zhu, *The superconvergent patch recovery and a posteriori error estimates, Part 2: error estimates and adaptivity*, Int. J. Numer. Methods Eng., 33 (1992), pp. 1365–1382.