MOX–Report No. 2/2008

An Adaptive Gradient-DWR Finite Element Algorithm for an Optimal Control Constrained Problem

Stefano Berrone, Marco Verani

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 29 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

An Adaptive Gradient-DWR Finite Element Algorithm for an Optimal Control Constrained Problem

Stefano Berrone[a] and Marco Verani[b]

11th January 2008

[a] Dipartimento di Matematica, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy, (sberrone@calvino.polito.it).
[b] MOX - Modelling and Scientific Computing - Dipartimento di Matematica "F. Brioschi", Politecnico di Milano, Milano, Italy (marco.verani@polimi.it)

Abstract

We present an adaptive finite element algorithm for the numerical approximation of distributed control constrained problems governed by second order elliptic PDEs. The algorithm is based on a suitable co-operation between a gradient type descent numerical scheme and the dual weighted residual (DWR) method. We assess the efficiency of the algorithm on several test problems and compare its performances with the ones of the residual based adaptive algorithm introduced in [17].

Keywords: control constrained problems, a posteriori estimates, finite elements, duality, gradient method.

AMS subject classification: 65N30, 65N50, 35J25, 49J20, 65K10

1 Introduction

The numerical solution of infinite-dimensional constrained or unconstrained minimization problems requires suitable approximation steps, which replace the infinite-dimensional feasible set by a finite-dimensional one, where the approximate minimum is looked for. The error of such an approximate minimum depends on the choice of the discrete feasible set. In general there are two ways to make the error decrease:

- to enlarge the discrete feasible set, following an *a priori* approach, in such a way that *any* element of the infinite-dimensional feasible set can be approximated, up to a certain order of accuracy, by elements belonging to the new discrete feasible set;
- to enlarge the discrete feasible sets with the help of information extracted from data and previous approximated minima, by means of suitable *a posteriori* error indicators.

Among the methods sharing the second approach, adaptive finite element methods for the solution of constrained or not-constrained optimal control problem governed by partial differential equations (see e.g. [20]) have recently received great attention; see e.g. [4, 3, 7, 25, 15, 16, 17, 19], where residual based *a posteriori* error estimator [1, 24, 2] or dual weighted residual (DWR) type estimator [5, 6] have been used. In this paper we consider the following optimal control constrained problem: find

$$u^* \;=\; \operatorname*{argmin}_{u\in\mathcal{U}_{ad}} J(y,u), \tag{1}$$

where $\mathcal{U}_{ad}$ is the set of admissible controls and y is the unique solution to the elliptic problem $\mathcal{A}y = f + u$. We propose an adaptive iterative algorithm, where, at each iteration, the strategy of enlarging the feasible set aims at performing one (or both) of the two tasks: (a) a more accurate approximated evaluation of the infinite-dimensional functional J; (b) a more accurate minimization. The choice between the two tasks is done automatically, by means of a suitable dynamical criterion, and it allows to avoid dangerous situations like investing the computational effort in the minimization of a perturbed (and possibly very different) functional.

The outline of the paper is the following: in Section 2 we will state the optimal control problem and we will recall the `primal-dual active set (pdas)` algorithm, that will be important in the construction of our adaptive method; in Section 3 we will introduce the finite element formulation of the optimal control problem; in Section 4 we will present an error estimate based on a suitable separation of the sources of the approximation error; in Section 5 we will introduce and discuss our adaptive algorithm, named ∇`-DWR Algorithm`; in Section 6 we will show several numerical experiments to assess the numerical behavior of the ∇`-DWR Algorithm`.

2 Problem Formulation

Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain with boundary $\Gamma = \partial\Omega$. We consider the following distributed optimal control problem:

Let

$$J(y,u) := \frac{1}{2}\int_\Omega (y - z_d)^2 dx + \frac{\alpha}{2}\int_\Omega (u - u_d)^2 dx, \tag{2}$$

with

$$-\Delta y = f + u \quad \text{in } \Omega, \quad y = 0 \quad \text{on } \Gamma = \partial\Omega, \tag{3}$$

where $f \in L^2(\Omega)$, $z_d, u_d \in L^\infty(\Omega)$, $\alpha > 0$ and

$$\mathcal{U}_{ad} = \{u \in L^2(\Omega) : \quad u(x) \le b(x) \text{ a.e. in } \Omega\} \subset L^2(\Omega), \tag{4}$$

with $b \in L^\infty(\Omega)$. Find

$$u^* \;=\; \operatorname*{argmin}_{u\in\mathcal{U}_{ad}} J(y,u). \tag{5}$$

It is well known (see e.g. [20]) that (2)-(5) admits a unique solution $(y^*, u^*) \in (H^2(\Omega) \cap H_0^1(\Omega)) \times L^2(\Omega)$ characterized by the following optimality system:

$$\begin{cases} -\Delta y^* = f + u^* \quad \text{in } \Omega, \quad y^* = 0 \quad \text{on } \Gamma, \\ -\Delta p^* = z_d - y^* \quad \text{in } \Omega, \quad p^* = 0 \quad \text{on } \Gamma, \\ (\alpha(u^* - u_d) - p^*, v - u^*) \geq 0, \quad \text{for all } v \in \mathcal{U}_{ad}, \end{cases} \tag{6}$$

where $p^* \in H^2(\Omega) \cap H_0^1(\Omega)$ and $(\cdot, \cdot)$ denotes the $L^2(\Omega)$ inner product. An equivalent formulation of the optimality system (6) has been stated in [18] as follows:

$$\begin{cases} -\Delta y^* = f + u^* \quad \text{in } \Omega, \quad y^* = 0 \quad \text{on } \Gamma, \\ -\Delta p^* = z_d - y^* \quad \text{in } \Omega, \quad p^* = 0 \quad \text{on } \Gamma, \\ u^* = u_d + \frac{p^* - \lambda^*}{\alpha}, \\ \lambda^* = c \max(0, u^* + \frac{\lambda^*}{c} - b), \end{cases} \tag{7}$$

for every $c > 0$. This formulation is essential to motivate the `primal-dual active set (pdas)` algorithm introduced in [9] for the approximate solution of the optimal control problem (2)-(5). For the ease of the reader we now briefly recall this algorithm, here stated at the continuous level, which will be used in the construction of the adaptive strategy discussed in this paper.

Let us introduce the active and inactive sets which are respectively defined as follows:

$$\mathcal{A}(u) = \{x \in \Omega : \quad u = b \text{ a.e.}\} \quad \text{and} \quad \mathcal{I}(u) = \{x \in \Omega : \quad u < b \text{ a.e.}\}. \tag{8}$$

The `primal-dual active set algorithm` reads as follows:

1. Initial Guess: y_0, u_0, λ_0
2. Build the active and inactive sets as:

$$\mathcal{A}_n = \{x \in \Omega : \quad u_{n-1} + \frac{\lambda_{n-1}}{c} > b \text{ a.e.}\}, \quad \mathcal{I}_n = \{x \in \Omega : \quad u_{n-1} + \frac{\lambda_{n-1}}{c} \leq b \text{ a.e.}\}$$

3. `if` $\mathcal{A}_n = \mathcal{A}_{n-1}$ then `stop`
4. `else` find (y_n, p_n) such that

$$-\Delta y_n = f + \begin{cases} b & \text{in } \mathcal{A}_n, \\ u_d + \frac{p_n}{\alpha} & \text{in } \mathcal{I}_n, \end{cases}$$

$$-\Delta p_n = z_d - y_n \quad \text{in } \Omega,$$

and set

$$u_n = \begin{cases} b & \text{in } \mathcal{A}_n, \\ u_d + \frac{p_n}{\alpha} & \text{in } \mathcal{I}_n. \end{cases}$$

5. set $\lambda_n = p_n - \alpha(u_n - u_d)$, update $n = n + 1$ and `goto` 2.

3 Finite Element Discretization

As a preliminary step towards the construction of our adaptive algorithm, we introduce the finite element approximation of the optimal control problem (2)-(5). Let us first consider the weak formulation of the optimality system (6): find $(y^*, u^*) \in H_0^1(\Omega) \times L^2(\Omega)$ such that

$$a(y^*, \phi) = (f + u^*, \phi) \qquad \forall \phi \in H_0^1(\Omega), \tag{9}$$

$$a(p^*, \phi) = -(y^* - z_d, \phi) \qquad \forall \phi \in H_0^1(\Omega), \tag{10}$$

$$u^* = u_d + \frac{1}{\alpha}(p^* - \lambda^*), \tag{11}$$

$$(\lambda^*, v - u^*) \leq 0 \qquad \forall v \in \mathcal{U}_{ad} \subset L^2(\Omega), \tag{12}$$

where $a(\cdot, \cdot) : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ is the bilinear form defined as

$$a(\phi_1, \phi_2) = \int_\Omega \nabla \phi_1 \cdot \nabla \phi_2 dx \qquad \forall \phi_1, \phi_2 \in H_0^1(\Omega).$$

It can be seen that the inequality (12) is equivalent to the following complementarity conditions:

$$\lambda^* \geq 0, \quad b - u^* \geq 0, \quad (\lambda^*, b - u^*) = 0 \quad \text{in } \Omega, \tag{13}$$

or, by using (11), to:

$$\alpha(u^* - u_d) - p^* \leq 0, \quad b - u^* \geq 0, \quad (\alpha(u^* - u_d) - p^*, b - u^*) = 0 \quad \text{in } \Omega. \tag{14}$$

Conditions (14) and (13) can be equivalently stated as

$$\lambda^*(x) \geq 0 \quad \text{a.e. in } \Omega, \tag{15}$$

$$\lambda^*(x) = 0 \quad \text{a.e. in } \mathcal{I}(u^*), \tag{16}$$

$$\lambda^*(x) = p^* - \alpha(b - u_d) \quad \text{a.e. in } \mathcal{A}(u^*). \tag{17}$$

Let us now introduce the finite element discretization of (2)-(5). Let $\mathcal{T}$ a shape-regular triangulation of the domain Ω, we denote by h_T and $|T|$ the diameter and the area of an element $T \in \mathcal{T}$. Let

$$V^h = \{v_h \in C^0(\Omega) | \quad v_h|_T \in \mathbb{P}^1(T), \ T \in \mathcal{T}\} \tag{18}$$

be the continuous piecewise linear finite element space, and

$$W^h = \{w_h \in L^2(\Omega) | \quad w_h|_T \in \mathbb{P}^0(T), T \in \mathcal{T}\} \tag{19}$$

be the (discontinuous) piecewise constant finite element space. Let $V_0^h := V^h \cap H_0^1(\Omega)$. Then a possible finite element approximation of (2)-(5) is as follows:

Let

$$J(y_h, u_h) := \frac{1}{2} \int_\Omega (y_h - z_d)^2 dx + \frac{\alpha}{2} \int_\Omega (u_h - u_d)^2 dx, \tag{20}$$

with

$$a(y_h, \phi_h) = (f + u_h, \phi_h) \qquad \forall \phi_h \in V_0^h. \tag{21}$$

Find

$$u_h^* \;=\; \underset{u \in \mathcal{U}_{ad,h}}{\operatorname{argmin}} J(y_h, u_h), \tag{22}$$

where

$$\mathcal{U}_{ad,h} = \{w_h \in W^h | \quad w_h|_T \leq b_h|_T,\ T \in \mathcal{T}\}, \tag{23}$$

with $b_h|_T := |T|^{-1} \int_T b(x)dx,\ T \in \mathcal{T}$.

In the rest of the paper we will work under the following simplifying assumption:

Assumption 3.1. *Let $b \in L^\infty(\Omega)$ such that $b \in W^h$, i.e. $b|_T = b_h|_T$, $\forall T \in \mathcal{T}$.*

The discretized control problem (20)-(22) has a unique solution $(y_h^*, u_h^*) \in V_0^h \times W^h$ which is characterized by the following optimality conditions:

$$a(y_h^*, \phi_h) = (f + u_h^*, \phi_h) \qquad \forall \phi_h \in V_0^h, \tag{24}$$

$$a(p_h^*, \phi_h) = -(y_h^* - z_d, \phi_h) \qquad \forall \phi_h \in V_0^h, \tag{25}$$

$$u_h^* = \Pi_0 u_d + \frac{1}{\alpha}(\Pi_0 p_h^* - \lambda_h^*), \tag{26}$$

$$(\lambda_h^*, v_h - u_h^*) \leq 0 \qquad \forall v_h \in \mathcal{U}_{ad,h} \subset L^2(\Omega), \tag{27}$$

where Π_0 is the projecting operator onto the piecewise constant space W^h; on each triangle $T \in \mathcal{T}$ it is defined as follows:

$$(\Pi_0 v_h)_{|_T} = |T|^{-1} \int_T v_h(x)dx. \tag{28}$$

As in the continuous case, the inequality (27) can be stated as the complementarity problem

$$\lambda_h^* \geq 0, \quad b - u_h^* \geq 0, \quad (\lambda_h^*, b - u_h^*) = 0, \tag{29}$$

which can be equivalently stated as

$$\lambda_h^*(x)|_T \geq 0 \quad T \in \mathcal{T}, \tag{30}$$

$$\lambda_h^*(x)|_T = 0 \quad T \in \mathcal{I}(u_h^*), \tag{31}$$

$$\lambda_h^*(x)|_T = (\Pi_0 p_h^*)|_T - \alpha(b - \Pi_0 u_d)|_T \quad T \in \mathcal{A}(u_h^*). \tag{32}$$

4 Split of the Error and DWR type Error Estimate

The crucial step towards the construction of our adaptive algorithm is to provide an *a posteriori* error estimate for the quantity $|J(y^*, u^*) - J(y_h^*, u_h^*)|$, being (y^*, u^*) the solution

to the optimal problem (2)-(5) and (y_h^*, u_h^*) the solution to the Galerkin problem (24)-(27). Having in mind to apply the DWR method [5, 6, 4] to the control problem (2)-(5), we introduce the Lagrangian functional $\mathcal{L} : H_0^1(\Omega) \times H_0^1(\Omega) \times L^2(\Omega) \to \mathbb{R}$ defined as

$$\mathcal{L}(y, p, u) = J(y, u) + a(y, p) - (f + u, p). \tag{33}$$

Let $v^* = (y^*, p^*, u^*)$ and $v_h = (y_h^*, p_h^*, u_h^*)$ denote respectively the continuous and discrete optimal triplets. Then the optimality conditions (9)-(10) and (24)-(25) yield

$$J(y^*, u^*) - J(y_h^*, u_h^*) \quad = \quad \mathcal{L}(y^*, p^*, u^*) - \mathcal{L}(y_h^*, p_h^*, u_h^*). \tag{34}$$

Now it is standard [5, 6, 25] to obtain the following equalities:

$$\begin{aligned}
\mathcal{L}(y^*, p^*, u^*) - \mathcal{L}(y_h^*, p_h^*, u_h^*) &= \frac{1}{2}[\nabla\mathcal{L}(v^*)(v^* - v_h) + \nabla\mathcal{L}(v_h)(v^* - v_h)] \qquad (35)\\
&= \frac{1}{2}[a(y^* - y_h^*, p_h^*) - (y_h^* - z_d, y^* - y_h^*)] \\
&+ \frac{1}{2}[a(y_h^*, p^* - p_h^*) - (f + u_h^*, p^* - p_h^*)] \\
&+ \frac{1}{2}[(\alpha(u^* - u_d) - p^*, u^* - u_h^*) + (\alpha(u_h^* - u_d) - p_h^*, u^* - u_h^*)].
\end{aligned}$$

Using (12) and bearing in mind that $\lambda^* = p^* - \alpha(u^* - u_d)$, give the following inequality

$$(\alpha(u^* - u_d) - p^*, u^* - u_h^*) = (\lambda^*, u_h^* - u^*) \leq 0, \tag{36}$$

being $\mathcal{U}_{ad,h} \subset \mathcal{U}_{ad}$ thanks to Assumption 3.1. Hence combining (34) and (35) yields

$$\begin{aligned}
J(y^*, u^*) - J(y_h^*, u_h^*) &\leq \frac{1}{2}[a(p_h^*, y^* - y_h^*) - (y_h^* - z_d, y^* - y_h^*)] \\
&+ \frac{1}{2}[a(y_h^*, p^* - p_h^*) - (f + u_h^*, p^* - p_h^*)] + (\alpha(u_h^* - u_d) - p_h^*, u^* - u_h^*) \qquad (37)
\end{aligned}$$

Unfortunately, as already pointed out by [25] in Remark 4.3, the inequality (37) does not provide an estimate for the absolute value $|J(y^*, u^*) - J(y_h^*, u_h^*)|$; to overcome this difficulty and to derive a computable DWR based *a posteriori* error estimate to be used in an adaptive strategy, various strategies have been recently proposed (see e.g. [25, 16]). In this paper we propose a different approach, somehow related to what has been presented in [14]. Let us now describe our idea.

For every control finite element approximation $u_h \in W^h$, we define the intermediate primal and dual solutions $(y(u_h), p(u_h)) \in H_0^1(\Omega) \times H_0^1(\Omega)$ as follows:

$$\begin{aligned}
a(y(u_h), v) &= (f + u_h, v)dx \quad \forall v \in H_0^1(\Omega), \qquad (38)\\
a(p(u_h), v) &= -(y(u_h) - z_d, v) \quad \forall v \in H_0^1(\Omega). \qquad (39)
\end{aligned}$$

We want to build a computable error estimator for the following quantity

$$|J(y^*, u^*) - J(y_h(u_h), u_h)|, \tag{40}$$

where $y_h(u_h) \in V_0^h$ is the Galerkin solution to (38). By using the triangle inequality we obtain the following *split of the error*:

$$\begin{aligned} |J(y^*,u^*) - J(y_h(u_h),u_h)| &\leq |J(y^*,u^*) - J(y(u_h),u_h)| \\ &+ |J(y(u_h),u_h) - J(y_h(u_h),u_h)|, \end{aligned} \tag{41}$$

which will be crucial for the design of our adaptive strategy. The right hand side of (41) identifies two sources of error:

- the approximation of the optimal control u^*;
- the approximation of $y(u_h)$.

Let now us estimate each of the two terms appearing on the right hand side of (41). Whenever u_h is taken as an approximation to u^*, the first term $|J(y^*,u^*) - J(y(u_h),u_h)|$ measures the distance, in terms of the value of functional J, between the optimal solution (y^*,u^*) and the exact intermediate solution $(y_h(u_h),u_h)$. This term can be approximated by using the truncated first order Taylor expansion:

$$\begin{aligned} J(y^*,u^*) - J(y(u_h),u_h) &\simeq D_u J(y(u_h),u_h)(u^* - u_h) \\ &= (\alpha(u_h - u_d) - p(u_h), u^* - u_h), \end{aligned} \tag{42}$$

where $D_u J(y(u_h),u_h)(u^* - u_h)$ denotes the Gateaux derivative with respect to u of the functional J in the direction $u^* - u_h$, evaluated at $(y(u_h),u_h)$, see e.g. [20]. Approximating $p(u_h)$ with its Galerkin discretization $p_h(u_h)$ yields

$$J(y^*,u^*) - J(y(u_h),u_h) \simeq (\alpha(u_h - u_d) - p_h(u_h), u^* - u_h). \tag{43}$$

The right-hand side of (43) is still non-computable as it involves the unknown function u^*; hence it cannot be used directly to drive an adaptive strategy. We will discuss in Section 5 the construction of a computable version of (43) and its usage in the design of our adaptive algorithm.

The second term $J(y(u_h),u_h) - J(y_h(u_h),u_h)$ measures the error caused by the approximation of the functional value $J(y(u_h),u_h)$ and it can be estimated by the standard tools of the DWR method [5, 6]. Let $I_h : H_0^1(\Omega) \to V_0^h$ be a suitable interpolation operator such that the following error estimates hold [12]:

$$\|y(u_h) - I_h y(u_h)\|_{0,T} + h_T^{1/2}\|y(u_h) - I_h y(u_h)\|_{0,\partial T} \leq Ch\|\nabla y(u_h)\|_{0,\mathcal{N}(T)} \tag{44}$$

$$\|p(u_h) - I_h p(u_h)\|_{0,T} + h_T^{1/2}\|p(u_h) - I_h p(u_h)\|_{0,\partial T} \leq Ch\|\nabla p(u_h)\|_{0,\mathcal{N}(T)}, \tag{45}$$

where C is a constant not depending on h and $\mathcal{N}(T) = \{T' : \ T \cap T' \neq \emptyset\}$. Then, by using (38)-(39) and the Lagrangian functional (33), it holds

$$\begin{aligned} |J(y(u_h),u_h) - J(y_h(u_h),u_h)| &\leq |\sum_{T\in\mathcal{T}} \{(R^p, p(u_h) - I_h p(u_h))_T + (r^p, p(u_h) - I_h p(u_h))_{\partial T}\}| \\ &+ |\sum_{T\in\mathcal{T}} \{(R^d, y(u_h) - I_h y(u_h))_T + (r^d, y(u_h) - I_h y(u_h))_{\partial T}\}| \\ &\leq \sum_{T\in\mathcal{T}} \rho_T^p \omega_T^d + \sum_{T\in\mathcal{T}} \rho_T^d \omega_T^p, \end{aligned} \tag{46}$$

with

$$R^p = \triangle y_h(u_h) + f, \qquad r^p = \begin{cases} \frac{1}{2}[\nu_E \cdot \nabla y_h(u_h)] & \text{if } E \subset \partial T \setminus \partial\Omega, \\ 0 & \text{if } E \subset \partial\Omega, \end{cases} \tag{47}$$

$$R^d = \triangle p_h(u_h) - (y_h(u_h) - z_d), \qquad r^d = \begin{cases} \frac{1}{2}[\nu_E \cdot \nabla p_h(u_h)] & \text{if } E \subset \partial T \setminus \partial\Omega, \\ 0 & \text{if } E \subset \partial\Omega, \end{cases} \tag{48}$$

where $[\nu_E \cdot \nabla \cdot]$ denotes the jump of the normal flux across the inter-element side E, and

$$\begin{aligned} \rho_T^p &:= \|R^p\|_{0,T} + h_T^{-1/2}\|r^p\|_{0,\partial T} \\ \rho_T^d &:= \|R^d\|_{0,T} + h_T^{-1/2}\|r^d\|_{0,\partial T} \\ \omega_T^p &:= \|y(u_h) - I_h y(u_h)\|_{0,T} + h_T^{1/2}\|y(u_h) - I_h y(u_h)\|_{0,\partial T} \quad \forall T \in \mathcal{T} \\ \omega_T^d &:= \|p(u_h) - I_h p(u_h)\|_{0,T} + h_T^{1/2}\|p(u_h) - I_h p(u_h)\|_{0,\partial T} \quad \forall T \in \mathcal{T} \end{aligned}$$

Applying the interpolation error estimates (44)-(45) to (46) yields the following error estimate:

$$|J(y(u_h), u_h) - J(y_h(u_h), u_h)| \leq C \sum_{T \in \mathcal{T}} h_T \rho_T^p \|\nabla p(u_h)\|_{0,\mathcal{N}(T)} + \sum_{T \in \mathcal{T}} h_T \rho_T^d \|\nabla y(u_h)\|_{0,\mathcal{N}(T)}. \tag{49}$$

The right hand-side still involves the unknown quantities $\nabla y(u_h), \nabla p(u_h)$. By resorting to the so-called Zienkiewicz-Zhu patch recovery technique [26, 27], we build two computable quantities $G_h(\nabla y_h(u_h))$ and $G_h(\nabla p_h(u_h))$, which are respectively approximation to $\nabla y(u_h)$ and $\nabla p(u_h)$. They enable us to define the error indicator $\mathcal{E}_{2,T}$ on the element T as

$$\mathcal{E}_{2,T} := h_T \rho_T^p \|G(\nabla y_h(u_h))\|_{0,\mathcal{N}(T)} + h_T \rho_T^d \|G(\nabla y_h(u_h))\|_{0,\mathcal{N}(T)}. \tag{50}$$

Combining (43) and (49) yields the following heuristic estimate:

$$|J(y^*, u^*) - J(y_h(u_h), u_h)| \simeq \mathcal{E}_1 + \mathcal{E}_2, \tag{51}$$

where

$$\mathcal{E}_1 := |(\alpha(u_h - u_d) - p_h(u_h), u^* - u_h)|, \tag{52}$$

$$\mathcal{E}_2 := C \sum_{T \in \mathcal{T}} \mathcal{E}_{2,T}. \tag{53}$$

Remark 4.1. While the quantity $\mathcal{E}_2$ is computable, the term $\mathcal{E}_1$, as it stands, is not computable, because it involves the unknown function u^*. In the next section we will propose a computable version of $\mathcal{E}_1$ and an adaptive algorithm based on it for the approximate solution of problem (2)-(5).

5 The Adaptive Algorithm

In this section we describe our adaptive algorithm for the approximate solution of the optimal control problem (2)-(5). In particular our adaptive scheme aims at approximating $u^* = \text{argmin}_{\mathcal{U}_{ad,h}} J(y,u)$ and it is based upon an iteration of the following main steps:

$$\texttt{minimize} \rightarrow \texttt{estimate} \rightarrow \texttt{select} \rightarrow \texttt{include},$$

where

1. `minimize` applies the primal-dual active set algorithm providing us with an approximation $u_h \in W^h$ to u^*, that minimizes the functional J among all the functions belonging to the actual piecewise constant finite element space W^h.

2. `estimate` is based on a computable version of (51) that allows to check the fulfillment of a suitable stopping test; if this is met then the algorithm stops; otherwise the `select` and the `include` procedures are performed.

3. `select` is based on $\mathcal{E}_2$ and on a localized computable version of $\mathcal{E}_1$ and it aims at selecting new basis functions to be added to the actual finite element spaces V_0^h and W^h in order to perform the following tasks:

   - enriching the set of descent directions with new feasible and effective directions, mainly governed by $\mathcal{E}_1$
   - improving the accuracy of the approximate computation of the functional J, mainly governed by $\mathcal{E}_2$.

4. `include` adds the basis functions selected in the previous step to the actual finite element spaces V_0^h and W^h.

5.1 Haar wavelets on the triangles

As an intermediate step towards the construction of a computable version of $\mathcal{E}_1$, we need to introduce the Haar wavelets on the triangles (for more detail on wavelets see [13],[21]). It is easy to see that the finite element space W^h is spanned by the Haar scaling basis functions

$$\phi_T(x) = \begin{cases} 1 & x \in T, \\ 0 & \text{otherwise}, \end{cases} \tag{54}$$

for every $T \in \mathcal{T}$. The quantity $u^* - u_h$ appearing in the definition (52) of $\mathcal{E}_1$ represents the correction to be added to the current approximation u_h in order to get u^* and it belongs to the orthogonal complement $W_\perp^h$ of W^h to $L^2(\Omega)$. A basis for $W_\perp^h$ can be obtained by considering the Haar wavelet basis functions $\psi_{j,k,T}$, defined over each triangle $T \in \mathcal{T}$. Now we briefly explain their construction (see also Figure 1). Let us first subdivide each triangle $T \in \mathcal{T}$ in four similar triangles $T_{1,k}$, $k = 0,..,3$, where $|T_{1,k}| = |T|/4$. Then it is easy to see that every function $f_T(x)$ which is defined on the triangle T and is constant on each sub-triangle $T_{1,k}$, $k = 0,..,3$ can be written as a unique linear combination of

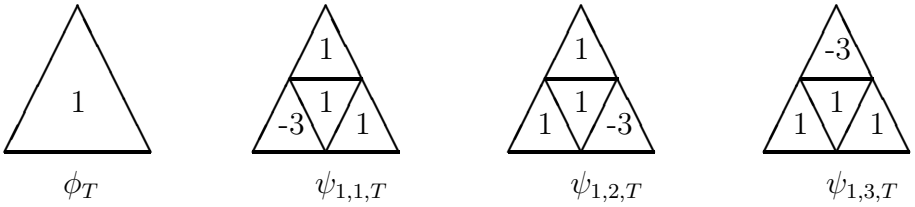

Figure 1: Haar scaling and wavelet functions

the constant function ϕ_T and three piecewise constant functions $\psi_{1,k,T}$, $k = 1, 2, 3$; i.e. $f_T = \alpha_T\phi_T + \beta_{1,T}\psi_{1,1,T} + \beta_{2,T}\psi_{1,2,T} + \beta_{3,T}\psi_{1,3,T}$. This building process can be iterated to define the Haar wavelet functions $\psi_{j,k,T}$ of level j and position k, all with support contained in the triangle T and the following orthogonality property holding for every $T \in \mathcal{T}$:

$$\int_T \phi_T \psi_{j,k,T} = 0 \quad \forall j, k \in \mathbb{N}.$$

5.2 Approximation of $\mathcal{E}_1$

We now focus on the crucial step of building a computable version of the error indicator $\mathcal{E}_1$. Let $u_h \in W^h$ and $p_h(u_h)$ be the Galerkin approximation to (39). We define the following quantities:

$$\delta J_h = \alpha(u_h - u_d) - p_h(u_h), \qquad \delta J_{h,T} = \alpha(u_h - u_d)|_T - p_h(u_h)|_T \quad \forall T \in \mathcal{T}. \tag{55}$$

In order to catch the bulk of $\mathcal{E}_1$ we compute the following indicators:

$$\mathcal{E}_{1,\psi_{1,k,T}} := (\delta J_h, \psi_{1,k,T})_T, \quad k = 1, .., 3, \tag{56}$$

$$\mathcal{E}_{1,T} := \sum_{k=1}^{3} |\mathcal{E}_{1,\psi_{1,k,T}}|. \tag{57}$$

Motivated by the fact that $\mathcal{E}_1$ is an approximation to the Gateaux derivative $D_u J(y(u_h), u_h)(u^* - u_h)$, with $u^* - u_h \in W^h_\perp$, we take the quantity $\mathcal{E}_{1,\psi_{1,k,T}}$ as an approximation to the derivative $D_u J(y(u_h), u_h)(\psi_{1,k,T})$ of J in the test direction $\psi_{1,k,T} \in W^h_\perp$ and we consider the quantity $\mathcal{E}_{1,T}$ as an indicator of the magnitude of $\mathcal{E}_1|_T$, at the resolution scale $j = 1$. The combined information about the size of $\mathcal{E}_{1,T}$ and the sign of $\delta J_{h,T}$ (see Remark 5.3) will be used in our adaptive algorithm to decide whether to add or not new directions to the set of feasible descent directions.

Remark 5.1. It is worth noticing that the idea of getting the bulk of a quantity of interest, like $\mathcal{E}_1$ in our case, by testing it with functions belonging to the complement of the actual space of discretization, has already been successfully exploited in the context of adaptive wavelet methods for the solution of elliptic PDEs (see e.g. [10]), and in the context of adaptive finite element methods for the solution of the obstacle problem [23]. The same idea has also been explored to design an adaptive finite element algorithm for the solution of shape optimization problems [22].

5.3 The adaptive algorithm

Our method iteratively solves the discrete problem (24)-(27) on the given triangulation $\mathcal{T}^{(k)}$, by means of the `primal dual active set (pdas)` algorithm and adaptively builds the new triangulation $\mathcal{T}^{(k+1)}$, by means of the error indicators $\mathcal{E}_{i,T},\ i = 1, 2$. The discrete active and inactive sets built by the primal dual active set algorithm are defined as suitable subsets of the triangulation $\mathcal{T}^{(k)}$; in particular at each iteration k the primal dual active set algorithm builds the two sets as follows:

$$\mathcal{I}^{(k)} := \{T \in \mathcal{T}^{(k)} \ : \ u_h^{(k)} + \frac{\lambda_h^{(k)}}{c} \leq b \ \text{ in } T\},$$
$$\mathcal{A}^{(k)} := \{T \in \mathcal{T}^{(k)} \ : \ T \notin \mathcal{I}^{(k)}\}.$$

Let $J_h^{(k)} := J(y_h^{(k)}, u_h^{(k)})$ denote the value of the functional at the iteration k, then we introduce the following quantity that will play an important rôle in the algorithm:

$$\mathrm{d}J/\mathrm{d}N^{(k)} := |J_h^{(k)} - J_h^{(k-1)}|/(N_{dofs}^{(k)} - N_{dofs}^{(k-1)}),$$

where $N_{dofs}^{(k)}$ is the number of degrees of freedom of the discrete problem on $\mathcal{T}^{(k)}$. The size of $\mathrm{d}J/\mathrm{d}N^{(k)}$ drives the selection of the refinement strategy among the following three ones:

- If $\mathrm{d}J/\mathrm{d}N^{(k)}$ is "large"$(\mathrm{d}J/\mathrm{d}N^{(k)} \geq \mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}})$, then refine according to the gradient-type error indicator $\mathcal{E}_1$;
- If $\mathrm{d}J/\mathrm{d}N^{(k)}$ is "small" $(\mathrm{d}J/\mathrm{d}N^{(k)} \leq \mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}})$, then refine according to the DWR-type error indicator $\mathcal{E}_2$;
- If $\mathrm{d}J/\mathrm{d}N^{(k)}$ is "of medium size" $(\mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}} < \mathrm{dJ/dN^{(k)}} < \mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}})$, then refine according to $\mathcal{E}_1$ and $\mathcal{E}_2$;

where $\mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}} > \mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}} > 0$ are given universal parameters.

Let $\mathrm{TOL}^{\mathrm{stop}} > 0$ be a user defined stop tolerance, then our adaptive algorithm, named ∇`-DWR Algorithm`, reads as follows:

Algorithm (∇`-DWR Algorithm`)

`Set` $k = 0$

`initialize`

`while` $\left(\mathrm{d}J/\mathrm{d}N^{(k)} > \mathrm{TOL}^{\mathrm{stop}}$ `or` $\mathrm{d}J/\mathrm{d}N^{(k-1)} > 10\,\mathrm{TOL}^{\mathrm{stop}}\right)$

{

$k = k + 1$

`Apply the pdas algorithm on` $\mathcal{T}^{(k)}$ `and compute` $J_h^{(k)}$

`if` $\left(\mathrm{d}J/\mathrm{d}N^{(k)} > \mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}}\right)$

`compute` $\mathcal{E}_{1,T} \ \forall T \in \mathcal{T}^{(k)}$

`if`$\big(T \in \mathcal{A}^{(k)}$ and $\delta J_{h,T} < 0\big)$

`set` $\mathcal{E}_{1,T} = 0$

`if` $\left(\mathrm{d}J/\mathrm{d}N^{(k)} < \mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}}\right)$

`compute` $\mathcal{E}_{2,T} \ \forall T \in \mathcal{T}^{(k)}$

$\mathcal{M}_T^{(k)}$ `= mark(`$\mathcal{T}^{(k)}$`)`

$\mathcal{T}^{(k+1)}$ `= refine(`$\mathcal{M}_T^{(k)}$`)`

}

Remark 5.2. The stopping criterion aims at predicting when the adaptive minimization starts to produce negligible effects: this is measured looking at the variation of the functional J_h with respect to the number of degrees of freedom. It is important to notice that spurious oscillations of the functional value coupled with the simpler stopping criterion $\{\mathrm{d}J/\mathrm{d}N^{(k)} < \mathrm{TOL}^{\mathrm{stop}}\}$ may led to an early stop of the algorithm. For that reason the safeguarded version $\{\mathrm{d}J/\mathrm{d}N^{(k)} < \mathrm{TOL}^{\mathrm{stop}}$ or $\mathrm{d}J/\mathrm{d}N^{(k-1)} < 10\,\mathrm{TOL}^{\mathrm{stop}}\}$ has been implemented.

Remark 5.3. Condition `if` $\big(T \in \mathcal{A}^{(k)}$ and $\delta J_{h,T} < 0\big)$ `set` $\mathcal{E}_{1,T} = 0$ aims at not refining an element T, when it belongs to the active set $\mathcal{A}^{(k)}$ and the inequality $\delta J_{h,T} < 0$ is satisfied. Indeed these two conditions guarantee, through (32), that T belongs to the exact active set $\mathcal{A}(u_h)$. In the actual implementation of the algorithm, we substitute the condition $\delta J_{h,T} < 0$, which is tricky in its evaluation at every point of T, with the more feasible condition $\Pi_1(\delta J_{h,T})(x_i) < 0$ for every vertices x_i of the triangle T, where Π_1 is the projector on the set of linear functions. We remark that if u_d is a piecewise linear function on $\mathcal{T}$, then the fulfillment of the new condition still implies, through (32), $T \in \mathcal{A}(u_h)$.

Remark 5.4. The `mark` and `refine` procedures have to perform clever balancing between the two terms $\mathcal{E}_{1,T}$ and $\mathcal{E}_{2,T}$ in order to avoid dangerous combinations, like a poor resolution of the functional J coupled with a rich set of descent directions. For instance this latter situation may lead to a very accurate minimization of a "perturbed" (and in principle different) functional.

The `mark` procedure is based on a bulk criterion which uses the elemental quantities $\mathcal{E}_{1,T}$ and $\mathcal{E}_{2,T}$, two tolerances $\mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}} > \mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}}$ and other two universal constants

$\Theta_{\delta J_h}$ and Θ_{DWR} to build a set $\mathcal{M}_T$ of triangles to be refined.

Algorithm (`mark`)

`set` $\mathcal{M}_T = \emptyset$

`if` $\Big(\mathrm{d}J/\mathrm{d}N^{(k)} > \mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}}\Big)$

`while` $\left(\Theta_{\delta J_h} \sum_{T\in\mathcal{T}} \mathcal{E}_{1,T} \le \sum_{T\in\mathcal{M}_T} \mathcal{E}_{1,T}\right)$ {

`let` $T_{\max}$ `the triangle that maximize` $\mathcal{E}_{1,T}$ `in` $\mathcal{T}\backslash\mathcal{M}_T$

`set` $\mathcal{M}_T = \mathcal{M}_T \cup T_{\max}$}

`if` $\big(\mathrm{d}J/\mathrm{d}N^{(k)} < \mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}}\big)$

`while` $\left(\Theta_{\mathrm{DWR}} \sum_{T\in\mathcal{T}} \mathcal{E}_{2,T} \le \sum_{T\in\mathcal{M}_T} \mathcal{E}_{2,T}\right)$ {

`let` $T_{\max}$ `the triangle that maximize` $\mathcal{E}_{2,T}$ `in` $\mathcal{T}\backslash\mathcal{M}_T$

`set` $\mathcal{M}_T = \mathcal{M}_T \cup T_{\max}$}

6 Numerical results

In this section we present a documentation of the numerical performances of the ∇-`DWR Algorithm`. In particular we consider three examples: two taken from the literature [9, 8] and the third derived from [19]. In the following we denote by N_{dofs} the number of degrees of freedom used by the `pdas` algorithm, by $\#\mathcal{I}^{(k)}$ the number of triangles contained in the discrete inactive region $\mathcal{I}^{(k)}$, by $\#\mathcal{A}^{(k)}$ the number of triangles contained in the discrete active region $\mathcal{A}^{(k)}$, by $\mathrm{d}J/\mathrm{d}N^{(k)}$ the ratio $|J_k^{(k)} - J_k^{(k-1)}|/(N_{dofs}^{(k)} - N_{dofs}^{(k-1)})$, by $\mathcal{E}_{1,\mathcal{M}_T}$ the sum $\sum_{T\in\mathcal{M}_T} \mathcal{E}_{1,T}$ and by $\mathcal{E}_{2,\mathcal{M}_T}$ the sum $\sum_{T\in\mathcal{M}_T} \mathcal{E}_{2,T}$, where $\mathcal{M}_T$ is the set of marked triangles for refinement. In order to asses the robustness of the algorithm we perform each numerical test for several values of the parameters $\Theta_{\delta J_h}$ and Θ_{DWR}. For each example we provide the plot of the adapted meshes, the history of the functional value with respect to N_{dofs} and some characteristic quantities. We also consider the influence of a modified marking step introduced in [17]: after performing the marking step described in Section 5, we enforce an extra-refinement of the elements belonging to $\mathcal{I}^{(k)}$ and $\mathcal{A}^{(k)}$ having an edge in $\mathcal{I}^{(k)} \cap \mathcal{A}^{(k)}$. The results of such extra-refinements, labeled with RefForce = 1, are displayed in some pictures; more detailed results can be found in [11]. Finally the numerical results of the ∇-`DWR Algorithm` are compared with the ones obtained by the `Residual-based Algorithm`: a slight modification of the adaptive algorithm presented in [17] and briefly recalled in Section 8.

For our computations we always set $c = 0.1$ in the `pdas` algorithm.

Table 1: Example 1, $\Theta_{\delta J_h} = 0.6$, $\Theta_{\mathrm{DWR}} = 0.6$, $\mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}} = 1.0\mathrm{E}{-8}$, $\mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}} = 5.0\mathrm{E}{-10}$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-11$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\mathcal{E}_{1,\mathcal{M}_T}$	$\mathcal{E}_{2,\mathcal{M}_T}$
802	113	143	4.192961E-02	0.0	8.1238E-04	0.0000E+00
1242	253	159	4.192891E-02	1.5751E-09	5.5672E-04	6.3780E-03
2620	565	324	4.191468E-02	1.0327E-08	3.9268E-04	0.0000E+00
4232	1042	387	4.191371E-02	6.0086E-10	2.6551E-04	3.6840E-03
9992	2167	1206	4.190961E-02	7.1255E-10	1.8127E-04	2.4434E-03
19258	4512	1927	4.190882E-02	8.4618E-11	0.0000E+00	1.6708E-03
41654	9013	4932	4.190837E-02	2.0311E-11	0.0000E+00	1.1651E-03
82284	17971	9525	4.190819E-02	4.3663E-12	0.0000E+00	8.1084E-04

Table 2: Example 1, $\Theta_{1,2,3,4} = 0.6$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-11$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\eta^2_{y,p,T,E}$	$\mu^2 + osc^2$
802	113	143	4.192961E-02	0.0	3.1180E-03	4.3970E-05
2382	378	415	4.192057E-02	5.7172E-09	1.3364E-03	1.8571E-05
3278	586	507	4.191296E-02	8.4903E-09	7.3766E-04	4.7958E-06
5078	933	784	4.191088E-02	1.1607E-09	4.2835E-04	2.0272E-06
9402	1758	1405	4.190943E-02	3.3533E-10	2.3088E-04	6.6809E-07
25150	5213	3323	4.190866E-02	4.8635E-11	1.0556E-04	2.2669E-07
38436	7284	5620	4.190835E-02	2.3545E-11	5.6781E-05	5.1092E-08
66792	12675	9730	4.190820E-02	5.0649E-12	3.2278E-05	1.8370E-08

6.1 Example 1

The first test problem is taken from [9]:
$\Omega = (0,1) \times (0,1)$, $\alpha = 0.01$, $u_d(x_1, x_2) = 0$, $b(x_1, x_2) = 0$, $f(x_1, x_2) = 0$,

$$z_d(x_1, x_2) = \sin(2\pi x_1)\sin(2\pi x_2)\frac{\exp(2x_1)}{6}.$$

In Tables 1 and 3 we report the evolution of the main quantities in the adaptive process of our algorithm. When a zero appears in the columns labeled with $\mathcal{E}_{1,\mathcal{M}_T}$ or $\mathcal{E}_{2,\mathcal{M}_T}$ means that the corresponding marking step of Algorithm `mark` is not applied to adapt the mesh; for example if a zero appears in the column labeled with $\mathcal{E}_{1,\mathcal{M}_T}$, then the refinement is performed only according to the DWR-type error indicator $\mathcal{E}_2$.

In Tables 2 and 4 we report the results obtained by applying the residual based adaptive algorithm introduced in [17]. The labels of the last two columns are defined as

Table 3: Example 1, $\Theta_{\delta J_h} = 0.4$, $\Theta_{\mathrm{DWR}} = 0.4$, $\mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}} = 1.0\mathrm{E}{-8}$, $\mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}} = 5.0\mathrm{E}{-10}$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-11$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\mathcal{E}_{1,\mathcal{M}_T}$	$\mathcal{E}_{2,\mathcal{M}_T}$
802	113	143	4.192961E-02	0.0	8.1238E-04	0.0000E+00
1062	201	151	4.193105E-02	5.5731E-09	6.3624E-04	6.9939E-03
1484	298	195	4.192084E-02	2.4214E-08	5.0402E-04	0.0000E+00
2012	449	221	4.191666E-02	7.9003E-09	4.0811E-04	5.2807E-03
3424	808	354	4.191424E-02	1.7191E-09	3.1855E-04	3.9531E-03
5370	1185	616	4.191095E-02	1.6868E-09	2.5336E-04	3.0804E-03
9666	2044	1203	4.190952E-02	3.3372E-10	0.0000E+00	2.3975E-03
14792	3444	1531	4.190912E-02	7.8790E-11	0.0000E+00	1.9022E-03
23322	4800	3001	4.190861E-02	5.9719E-11	0.0000E+00	1.5077E-03
39010	8348	4718	4.190839E-02	1.3999E-11	0.0000E+00	1.2015E-03
60236	13860	6301	4.190828E-02	5.2016E-12	0.0000E+00	9.5602E-04

Table 4: Example 1, $\Theta_{1,2,3,4} = 0.4$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-11$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\eta^2_{y,p,T,E}$	$\mu^2 + osc^2$
802	113	143	4.192961E-02	0.0	3.1180E-03	4.3970E-05
2150	336	376	4.192269E-02	5.1285E-09	1.7901E-03	2.6336E-05
8056	1316	1379	4.191553E-02	1.2134E-09	7.1830E-04	1.2143E-05
8502	1397	1439	4.191213E-02	7.6095E-09	4.8014E-04	3.5931E-06
9346	1542	1582	4.191073E-02	1.6647E-09	3.4474E-04	1.3933E-06
11262	1948	1818	4.190986E-02	4.5085E-10	2.3069E-04	5.5512E-07
14508	2789	2054	4.190894E-02	2.8483E-10	1.6003E-04	2.2680E-07
33202	7081	4152	4.190854E-02	2.1346E-11	9.0362E-05	1.0644E-07
39190	8002	5193	4.190838E-02	2.6144E-11	6.2812E-05	5.6212E-08
49332	9946	6594	4.190826E-02	1.1735E-11	4.4322E-05	2.8711E-08
103900	20473	14646	4.190818E-02	1.4508E-12	2.7545E-05	1.6572E-08

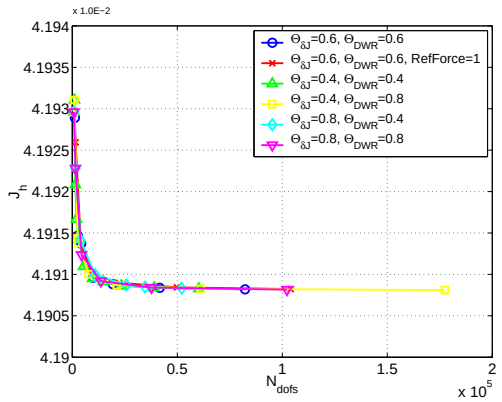

Figure 2: Example 1: convergence history of the functional J_h, for different values of $\Theta_{\delta J_h}$ and Θ_{DWR}

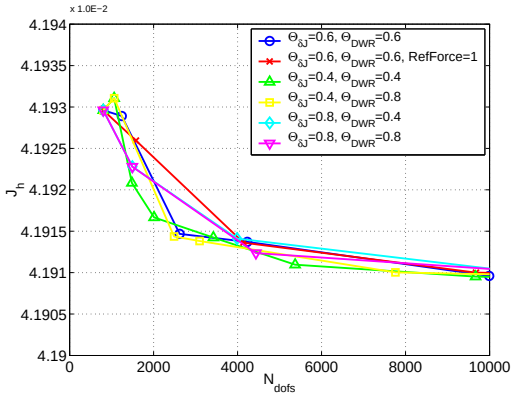

Figure 3: Example 1: convergence history of the functional J_h, for different values of $\Theta_{\delta J_h}$ and Θ_{DWR} - detail of the first iterations

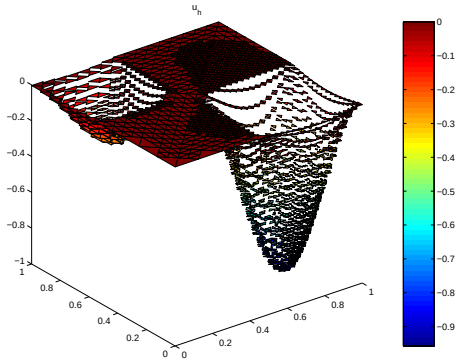

Figure 4: Example 1: discrete control function u_h at iteration $k = 7$ of Table 3

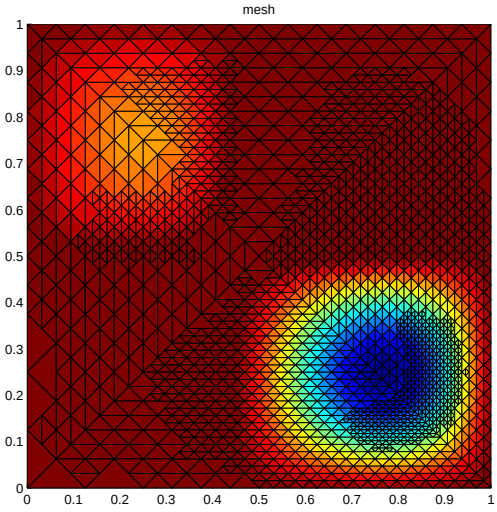

Figure 5: Example 1: adapted mesh at iteration $k = 7$ of Table 3

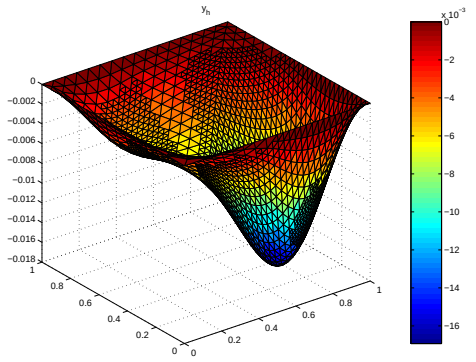

Figure 6: Example 1: state function y_h at iteration $k = 7$ of Table 3

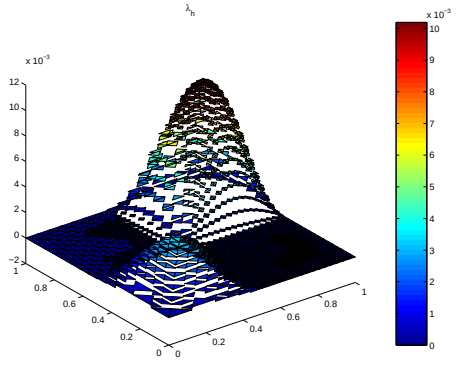

Figure 7: Example 1: co-control function λ_h at iteration $k = 7$ of Table 3

follows:

$$\begin{aligned} \eta^2_{y,p,T,E} &= \sum_{E\in\mathcal{E}} \eta^2_{y,E} + \eta^2_{p,E} + \sum_{T\in\mathcal{T}} \eta^2_{y,T} + \left(\eta^{(1)}_{p,T}\right)^2 + \left(\eta^{(2)}_{p,T}\right)^2, \\ \mu^2 + osc^2 &= \sum_{T\in\mathcal{T}} \mu^2_T(u_d) + \mu^2_T(b) + osc^2_T(y_d) + osc^2_T(f), \end{aligned}$$

where we refer to Section 8 for more details.

From Tables 1-2 we conclude that the two algorithms produce comparable results: the `Residual-based Algorithm` satisfies the stopping criterion with a slightly smaller number of degrees of freedoms, but the value of the functional J_h is slightly larger with respect to the one obtained by the ∇`-DWR Algorithm`. In Table 3 we report the results obtained by using the ∇`-DWR Algorithm` with different values of $\Theta_{\delta J_h}$ and Θ_{DWR}, which influence the growth of the number of degrees of freedom employed by the algorithm; a similar conclusion, but now in ∇`-DWR Algorithm`'s favor, can be drawn from Tables 3-4.

In Figure 2 we display the history of convergence of the ∇`-DWR Algorithm`, for different values of $\Theta_{\delta J_h}$ and Θ_{DWR} (see Figure 3 for a zoom of the first iterations or [11] for the corresponding tables).

In Figure 4 we plot the discrete control u_h obtained at the seventh adaptive step with $\Theta_{\delta J_h} = \Theta_{\mathrm{DWR}} = 0.4$, in Figure 5 we report the corresponding adapted mesh, while in Figures 6-7 we show the functions y_h and λ_h, respectively.

From all the Tables concerning the ∇`-DWR Algorithm` and from Figures 2-3 it is evident the strong relative reduction of the functional value when the mesh is adapted according to the indicator $\mathcal{E}_1$, which is an approximation of the gradient of the functional J_h.

Remark 6.1. The histories of convergence reported in Figures 2-3 and in the corresponding Tables show the ability of ∇`-DWR Algorithm` to capture the essential features of the optimal control variable u^* in order to make the functional decrease, also when the absolute variation of the functional value (and this is typical of inverse problems) is

Table 5: Example 2, $\Theta_{\delta J_h} = 0.6$, $\Theta_{\mathrm{DWR}} = 0.6$, $\mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}} = 1.0\mathrm{E}{-8}$, $\mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}} = 5.0\mathrm{E}{-10}$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-11$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\mathcal{E}_{1,\mathcal{M}_T}$	$\mathcal{E}_{2,\mathcal{M}_T}$
802	222	34	9.593094E-02	0.0	2.4756E-03	0.0000E+00
1702	517	51	9.593092E-02	2.8475E-11	0.0000E+00	3.5537E-02
2780	782	143	9.588493E-02	4.2661E-08	1.3466E-03	0.0000E+00
5526	1682	161	9.586335E-02	7.8585E-09	9.4288E-04	2.1082E-02
10644	3071	476	9.585014E-02	2.5802E-09	6.4825E-04	1.3612E-02
25938	7530	1198	9.584305E-02	4.6364E-10	0.0000E+00	9.1163E-03
49282	14124	2332	9.583968E-02	1.4466E-10	0.0000E+00	6.3567E-03
106164	29886	5674	9.583869E-02	1.7342E-11	0.0000E+00	4.4203E-03
212284	60930	10018	9.583799E-02	6.6153E-12	0.0000E+00	3.0900E-03

restricted to the fourth or fifth digit. This effective behavior is strongly related to the use of the gradient method combined with the DWR error estimate. The same kind of conclusions can be drawn from the results of Example 2 and Example 3, where the latter is particular meaningful, as the exact value of the optimal control is known.

6.2 Example 2

The second test problem is taken from [8]:
$\Omega = (0,1) \times (0,1)$, $\alpha = 0.01$, $u_d(x_1, x_2) = 0$, $b(x_1, x_2) = 1$, $f(x_1, x_2) = 0$,

$$z_d(x_1, x_2) = \begin{cases} 200 x_1 x_2 (x_1 - \frac{1}{2})^2 (1 - x_2), & \text{if } 0 < x_1 \leq 1/2, \\ 200 (x_1 - 1) x_2 (x_1 - \frac{1}{2})^2 (1 - x_2), & \text{if } 1/2 < x_1 \leq 1. \end{cases}$$

In Tables 5-8, we report the main adaptive quantities as in the previous example. From Tables 5-6 we conclude again that the two algorithms produce comparable results; the ∇`-DWR Algorithm` satisfies the stopping criterion with a slightly smaller number of degrees of freedoms, although the value of the discrete functional J_h is slightly larger with respect to the `Residual-based Algorithm`. A similar comparison, with similar conclusions, can be applied to Tables 7-8.

In Figure 8 we display the history of convergence of the ∇`-DWR Algorithm`, for different values of $\Theta_{\delta J_h}$ and Θ_{DWR} (see Figure 9 for a zoom of the first iterations or [11] for the corresponding tables).

In Figures 10-13 we plot the adapted functions u_h, y_h, λ_h and the corresponding mesh obtained at the ninth adaptive step for $\Theta_{\delta J_h} = \Theta_{\mathrm{DWR}} = 0.4$.

Table 6: Example 2, $\Theta_{1,2,3,4} = 0.6$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-11$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\eta^2_{y,p,T,E}$	$\mu^2 + osc^2$
802	222	34	9.593094E-02	0.0	1.6402E-02	5.6518E-05
2280	616	138	9.588860E-02	2.8649E-08	6.8084E-03	2.4081E-05
3602	1053	145	9.585918E-02	2.2258E-08	3.6631E-03	8.5921E-06
6274	1555	558	9.585136E-02	2.9258E-09	2.1744E-03	3.0231E-06
10846	3077	578	9.584453E-02	1.4932E-09	1.1529E-03	1.0922E-06
21264	5356	1824	9.584125E-02	3.1454E-10	6.5632E-04	4.2832E-07
34442	9288	2317	9.583963E-02	1.2289E-10	3.6541E-04	1.2760E-07
71052	18918	4933	9.583854E-02	2.9808E-11	1.9541E-04	3.1471E-08
276418	72969	19723	9.583779E-02	3.6540E-12	5.6099E-05	7.7475E-09

Table 7: Example 2, $\Theta_{\delta J_h} = 0.4$, $\Theta_{\mathrm{DWR}} = 0.4$, $\mathrm{TOL}^{\mathrm{start}}_{\mathrm{DWR}} = 1.0\mathrm{E}{-8}$, $\mathrm{TOL}^{\mathrm{stop}}_{\delta \mathrm{J_h}} = 5.0\mathrm{E}{-10}$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-11$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\mathcal{E}_{1,\mathcal{M}_T}$	$\mathcal{E}_{2,\mathcal{M}_T}$
802	222	34	9.593094E-02	0.0	2.4756E-03	0.0000E+00
1346	407	41	9.594140E-02	1.9218E-08	1.9525E-03	0.0000E+00
1986	600	70	9.592415E-02	2.6952E-08	1.5462E-03	0.0000E+00
2806	837	88	9.588509E-02	4.7635E-08	1.2352E-03	0.0000E+00
4832	1512	100	9.587344E-02	5.7502E-09	9.8955E-04	2.3628E-02
7752	2324	278	9.585864E-02	5.0671E-09	7.8091E-04	1.6801E-02
11550	3254	578	9.584749E-02	2.9371E-09	6.2151E-04	1.3069E-02
21256	6190	945	9.584416E-02	3.4290E-10	0.0000E+00	1.0105E-02
31492	8850	1714	9.584255E-02	1.5699E-10	0.0000E+00	8.0230E-03
48864	13977	2317	9.583965E-02	1.6717E-10	0.0000E+00	6.3838E-03
81630	23450	3941	9.583911E-02	1.6278E-11	0.0000E+00	5.0770E-03
123594	34563	6808	9.583857E-02	1.2952E-11	0.0000E+00	4.0412E-03
191908	54987	9091	9.583802E-02	7.9996E-12	0.0000E+00	3.2249E-03

Table 8: Example 2, $\Theta_{1,2,3,4} = 0.4$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-11$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\eta^2_{y,p,T,E}$	$\mu^2 + osc^2$
802	222	34	9.593094E-02	0.0	1.6402E-02	5.6518E-05
1642	416	131	9.590818E-02	2.7096E-08	8.9544E-03	3.5329E-05
5674	1385	530	9.588058E-02	6.8446E-09	3.9673E-03	2.1019E-05
6174	1548	532	9.586827E-02	2.4628E-08	2.6919E-03	1.0111E-05
7588	1975	576	9.585211E-02	1.1432E-08	1.8193E-03	2.1292E-06
10194	2855	578	9.584721E-02	1.8801E-09	1.2347E-03	8.6504E-07
16390	4580	932	9.584229E-02	7.9427E-10	8.3560E-04	4.2248E-07
55292	15178	3492	9.584007E-02	5.6955E-11	3.9839E-04	1.3011E-07
218258	59355	13942	9.583839E-02	1.0291E-11	1.4919E-04	5.6534E-08
221888	60523	13947	9.583804E-02	9.6623E-11	1.0292E-04	2.3603E-08
238626	65055	14923	9.583787E-02	1.0338E-11	7.3661E-05	1.1608E-08
278318	74685	18385	9.583775E-02	2.9094E-12	5.0356E-05	2.6525E-09

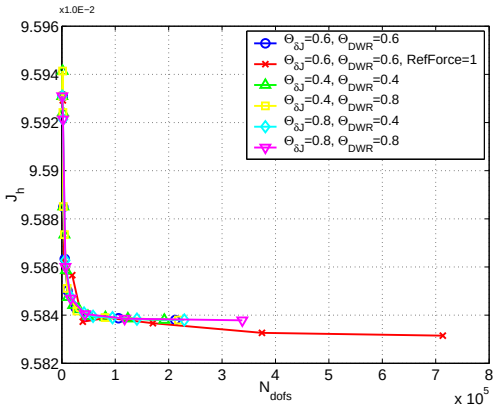

Figure 8: Example 2: convergence history of the functional J_h, for different values of $\Theta_{\delta J_h}$ and Θ_{DWR}

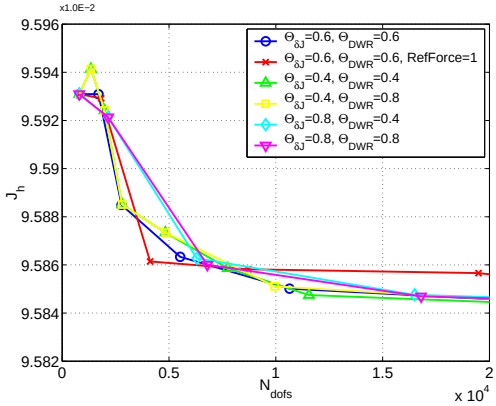

Figure 9: Example 2: convergence history of the functional J_h, for different values of $\Theta_{\delta J_h}$ and Θ_{DWR} - detail of the first iterations

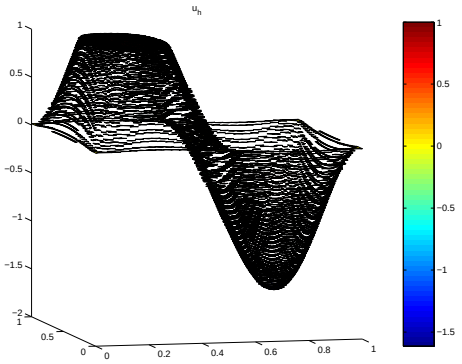

Figure 10: Example 2: discrete control function u_h at iteration $k = 9$ of Table 7

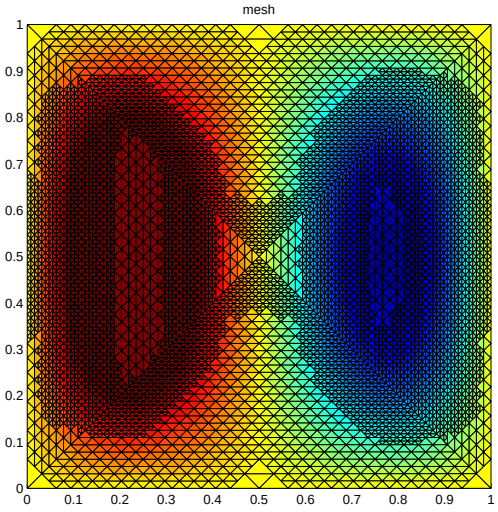

Figure 11: Example 2: adapted mesh at iteration $k = 9$ of Table 7

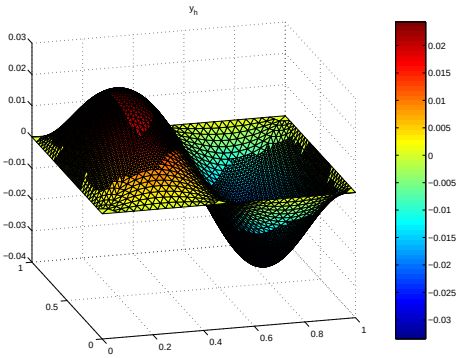

Figure 12: Example 2: state function y_h at iteration $k = 9$ of Table 7

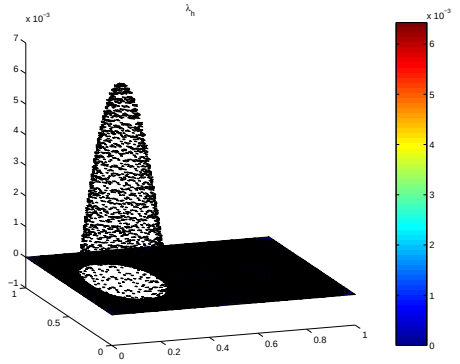

Figure 13: Example 2: co-control function λ_h at iteration $k = 9$ of Table 7

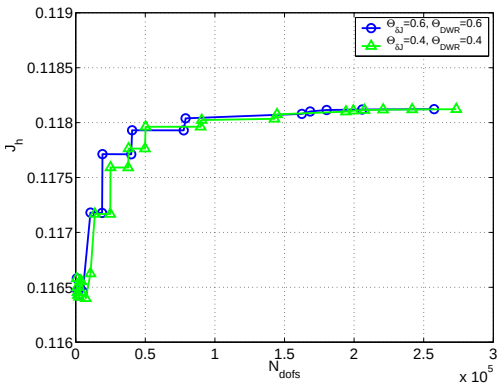

Figure 14: Example 3: convergence history of the functional J_h, for different values of $\Theta_{\delta J_h}$ and Θ_{DWR}

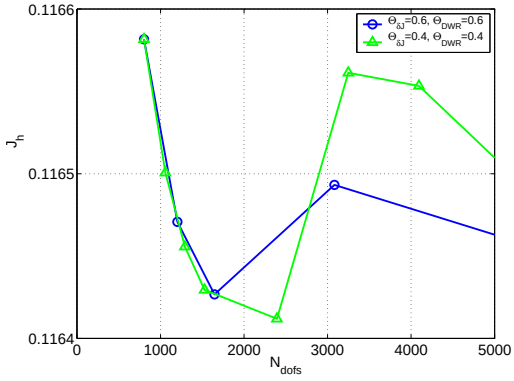

Figure 15: Example 3: convergence history of the functional J_h, for different values of $\Theta_{\delta J_h}$ and Θ_{DWR} - detail of the first iterations

6.3 Example 3

The third test problem is:
$\Omega = (0,1) \times (0,1)$, $\alpha = 1$, $b(x_1, x_2) = 0$, $z_d(x_1, x_2) = 0$,

$$u_d(x_1, x_2) = \sin(\frac{\pi}{2} x_1) + \sin(\frac{\pi}{2} x_2) - 1,$$

moreover we define:

$$\begin{aligned} p(x_1, x_2) &= -A \exp(-a(x_1 - v_1)^2) x_1 (1.0 - x_1) \exp(-a(x_2 - v_2)^2) x_2 (1.0 - x_2), \\ y(x_1, x_2) &= z_d(x_1, x_2) + \triangle p(x_1, x_2), \\ f(x_1, x_2) &= -\triangle y(x_1, x_2) - u(x_1, x_2), \end{aligned}$$

with $A = 5.0E-2$, $a = 1000$, $v_1 = v_2 = 0.85$. As boundary conditions for the primal and dual problems we choose the exact values of y and p on $\partial\Omega$. The explicit exact solution is known and is

$$u(x_1, x_2) = \min\{u_d(x_1, x_2) + p(x_1, x_2)/\alpha, 0\}.$$

This test problem is derived by modifying the one presented in [19]; in particular the primal solution y exhibits a singular behaviour inside the contact region.

In all the tables we report the error $\|u - u_h\|_0$ for the control function u and the error $|y - y_h|_1$ for the state function y, where $|\cdot|_1$ denotes the H^1 semi-norm.

Observing Tables 9-10 we conclude that the ∇-`DWR Algorithm` satisfies the stopping criterion with much less degrees of freedoms than the `Residual-based Algorithm`, although the value of the functional J_h is slightly larger. The final values of the errors $\|u - u_h\|_0$ and $|y - y_h|_1$ are comparable. Similar conclusions can be drawn from Tables 11-12.

Table 9: Example 3, $\Theta_{\delta J_h} = 0.6$, $\Theta_{\mathrm{DWR}} = 0.6$, $\mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}} = 1.0\mathrm{E}-7$, $\mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}} = 5.0\mathrm{E}-9$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-10$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\mathcal{E}_{1,\mathcal{M}_T}$	$\mathcal{E}_{2,\mathcal{M}_T}$	$\|u-u_h\|_0$	$\|y-y_h\|_1$
802	70	186	1.165817E-01	0.0	3.45E-02	0.00E+00	4.006E-04	5.286E+01
1202	199	192	1.164708E-01	2.77E-07	2.36E-02	0.00E+00	1.793E-04	5.286E+01
1646	302	236	1.164268E-01	9.92E-08	1.53E-02	2.17E-01	9.830E-05	5.286E+01
3082	764	257	1.164932E-01	4.63E-08	1.07E-02	2.78E-01	4.707E-05	3.149E+01
5100	1312	384	1.164614E-01	1.57E-08	7.21E-03	2.32E-01	2.358E-05	2.668E+01
10610	3079	462	1.171805E-01	1.30E-07	5.01E-03	0.00E+00	1.169E-05	1.190E+01
18986	5670	664	1.171774E-01	3.63E-10	0.00E+00	1.68E-01	5.714E-06	1.190E+01
19250	5670	760	1.177129E-01	2.03E-06	3.45E-03	0.00E+00	5.716E-06	6.431E+00
39918	12534	790	1.177115E-01	6.94E-11	0.00E+00	1.36E-01	2.849E-06	6.431E+00
40552	12542	998	1.179301E-01	3.45E-07	2.41E-03	0.00E+00	2.846E-06	2.696E+00
77624	24391	1512	1.179293E-01	1.99E-11	0.00E+00	9.07E-02	1.377E-06	2.696E+00
78966	24391	1977	1.180388E-01	8.16E-08	1.67E-03	6.39E-02	1.377E-06	1.317E+00
162452	51177	3043	1.180795E-01	4.87E-10	0.00E+00	4.35E-02	6.907E-07	5.657E-01
168444	51181	5082	1.181016E-01	3.69E-09	0.00E+00	3.03E-02	6.909E-07	3.232E-01
180274	51177	9070	1.181150E-01	1.14E-09	0.00E+00	2.06E-02	6.907E-07	1.336E-01
205818	51177	17641	1.181191E-01	1.57E-10	0.00E+00	1.43E-02	6.907E-07	7.713E-02
257500	51177	34999	1.181224E-01	6.50E-11	0.00E+00	9.89E-03	6.907E-07	3.382E-02

Table 10: Example 3, $\Theta_{1,2,3,4} = 0.6$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-10$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\eta^2_{y,p,T,E}$	$\mu^2 + osc^2$	$\|u-u_h\|_0$	$\|y-y_h\|_1$
802	70	186	1.165817E-01	0.0	1.89E+02	9.08E+03	4.006E-04	5.286E+01
1762	236	350	1.165363E-01	4.72E-08	5.35E+02	1.69E+03	1.302E-04	3.155E+01
2732	268	636	1.165697E-01	3.44E-08	4.77E+02	4.43E+02	1.022E-04	2.576E+01
6136	858	1192	1.172744E-01	2.07E-07	2.26E+02	5.35E+01	3.887E-05	9.200E+00
10172	1038	2362	1.175920E-01	7.87E-08	1.23E+02	1.12E+01	2.686E-05	5.150E+00
19088	2910	3451	1.177928E-01	2.25E-08	7.39E+01	4.78E+00	1.264E-05	3.197E+00
38696	4120	8847	1.180459E-01	1.29E-08	2.31E+01	8.51E-01	6.912E-06	8.689E-01
61294	6609	13807	1.180847E-01	1.72E-09	1.24E+01	2.13E-01	5.271E-06	4.482E-01
135642	15318	30094	1.181085E-01	3.20E-10	4.99E+00	6.36E-02	1.979E-06	1.949E-01
201932	17080	50337	1.181159E-01	1.12E-10	2.77E+00	1.52E-02	1.604E-06	9.920E-02
442174	56810	90815	1.181187E-01	1.16E-11	1.60E+00	5.14E-03	5.754E-07	5.842E-02

Table 11: Example 3, $\Theta_{\delta J_h} = 0.4$, $\Theta_{\mathrm{DWR}} = 0.4$, $\mathrm{TOL}_{\mathrm{DWR}}^{\mathrm{start}} = 1.0\mathrm{E}{-7}$, $\mathrm{TOL}_{\delta \mathrm{J_h}}^{\mathrm{stop}} = 5.0\mathrm{E}{-9}$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-10$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\mathcal{E}_{1,\mathcal{M}_T}$	$\mathcal{E}_{2,\mathcal{M}_T}$	$\|u-u_h\|_0$	$\|y-y_h\|_1$
802	70	186	1.165817E-01	0.0	3.45E-02	0.00E+00	4.006E-04	5.286E+01
1058	157	186	1.165007E-01	3.16E-07	2.75E-02	0.00E+00	2.386E-04	5.286E+01
1292	223	198	1.164557E-01	1.92E-07	2.11E-02	0.00E+00	1.513E-04	5.286E+01
1528	270	229	1.164295E-01	1.11E-07	1.62E-02	0.00E+00	1.033E-04	5.286E+01
2394	554	236	1.164119E-01	2.03E-08	1.28E-02	2.16E-01	6.862E-05	5.286E+01
3248	815	260	1.165613E-01	1.75E-07	1.02E-02	0.00E+00	4.222E-05	3.154E+01
4094	1040	317	1.165533E-01	9.40E-09	7.98E-03	2.98E-01	2.657E-05	3.154E+01
7272	2050	372	1.164002E-01	4.82E-08	6.27E-03	2.40E-01	1.813E-05	2.967E+01
10624	3094	450	1.166251E-01	6.71E-08	4.98E-03	1.59E-01	1.162E-05	1.741E+01
13762	4009	582	1.171687E-01	1.73E-07	3.95E-03	0.00E+00	7.191E-06	1.172E+01
24826	7614	670	1.171675E-01	1.08E-10	0.00E+00	1.65E-01	4.815E-06	1.172E+01
25002	7616	731	1.175918E-01	2.41E-06	3.14E-03	0.00E+00	4.815E-06	7.133E+00
37720	11832	757	1.175910E-01	6.75E-11	0.00E+00	1.39E-01	3.101E-06	7.133E+00
37980	11840	842	1.177642E-01	6.66E-07	2.51E-03	0.00E+00	3.101E-06	4.806E+00
49750	15577	1026	1.177636E-01	4.87E-11	0.00E+00	1.14E-01	1.915E-06	4.806E+00
50230	15565	1200	1.179626E-01	4.15E-07	2.00E-03	0.00E+00	1.914E-06	2.454E+00
89782	28429	1530	1.179623E-01	8.32E-12	0.00E+00	9.12E-02	1.260E-06	2.454E+00
90554	28435	1797	1.180233E-01	7.90E-08	1.59E-03	7.10E-02	1.260E-06	1.707E+00
142782	45391	2268	1.180344E-01	2.13E-10	0.00E+00	5.57E-02	8.215E-07	1.065E+00
144638	45391	2892	1.180755E-01	2.21E-08	1.27E-03	4.36E-02	8.215E-07	5.816E-01
194158	60400	4416	1.180999E-01	4.94E-10	0.00E+00	3.48E-02	5.102E-07	4.223E-01
199496	60404	6203	1.181083E-01	1.57E-09	0.00E+00	2.72E-02	5.103E-07	2.452E-01
207440	60396	8875	1.181143E-01	7.55E-10	0.00E+00	2.12E-02	5.101E-07	1.381E-01
220896	60396	13417	1.181176E-01	2.45E-10	0.00E+00	1.68E-02	5.101E-07	1.067E-01
241706	60396	20380	1.181201E-01	1.20E-10	0.00E+00	1.32E-02	5.101E-07	6.691E-02
273416	60396	31006	1.181220E-01	5.78E-11	0.00E+00	1.05E-02	5.101E-07	3.710E-02

Table 12: Example 3, $\Theta_{1,2,3,4} = 0.4$, $\mathrm{TOL}^{\mathrm{stop}} = 1.0\mathrm{E}-10$

N_{dofs}	$\#\mathcal{I}^{(k)}$	$\#\mathcal{A}^{(k)}$	J_h	$\mathrm{d}J/\mathrm{d}N^{(k)}$	$\eta^2_{y,p,T,E}$	$\mu^2 + osc^2$	$\|u-u_h\|_0$	$\|y-y_h\|_1$
802	70	186	1.165817E-01	0.0	1.89E+02	9.08E+03	4.006E-04	5.286E+01
1390	211	249	1.165537E-01	4.75E-08	5.35E+02	1.69E+03	1.634E-04	3.155E+01
1960	242	413	1.164547E-01	1.74E-07	5.06E+02	7.44E+02	1.237E-04	2.967E+01
2664	260	629	1.175206E-01	1.51E-06	2.51E+02	2.18E+02	1.068E-04	2.032E+01
3682	392	827	1.173007E-01	2.16E-07	2.14E+02	5.34E+01	8.713E-05	1.020E+01
6078	854	1178	1.174618E-01	6.72E-08	1.65E+02	2.23E+01	3.889E-05	8.788E+00
14770	2828	2144	1.177903E-01	3.78E-08	7.02E+01	6.31E+00	1.839E-05	3.690E+00
23776	4362	3676	1.179745E-01	2.04E-08	3.16E+01	1.62E+00	1.312E-05	1.750E+00
26522	4437	4483	1.180549E-01	2.93E-08	2.28E+01	8.61E-01	1.201E-05	1.054E+00
36316	5086	7119	1.180685E-01	1.39E-09	1.51E+01	3.69E-01	7.854E-06	6.564E-01
50248	5452	11472	1.181035E-01	2.51E-09	8.63E+00	1.54E-01	6.372E-06	3.301E-01
61372	5664	14929	1.181052E-01	1.47E-10	6.24E+00	7.51E-02	5.773E-06	2.357E-01
95518	12110	19853	1.181098E-01	1.35E-10	4.32E+00	2.94E-02	3.000E-06	1.652E-01
143886	15102	33169	1.181163E-01	1.35E-10	2.84E+00	1.66E-02	2.037E-06	1.004E-01
248506	22194	61630	1.181197E-01	3.27E-11	1.73E+00	7.70E-03	1.606E-06	6.309E-02

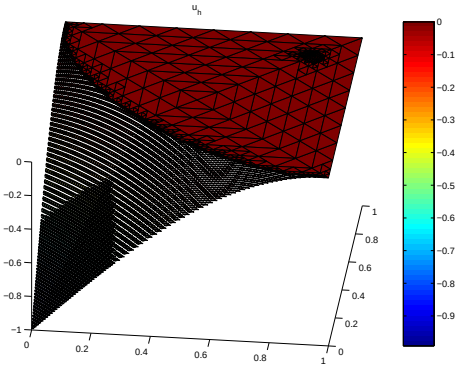

Figure 16: Example 3: discrete control function u_h at iteration $k = 12$ of Table 11

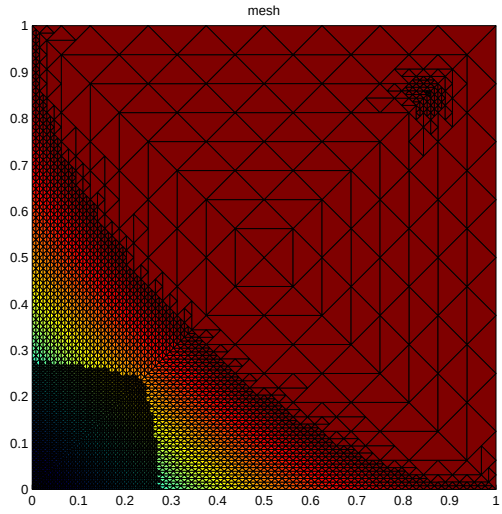

Figure 17: Example 3: adapted mesh at iteration $k = 12$ of Table 11

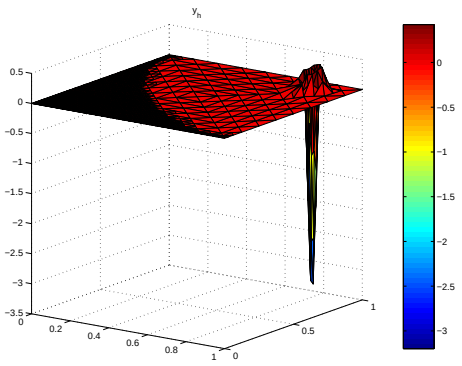

Figure 18: Example 3: state function y_h at iteration $k = 12$ of Table 11

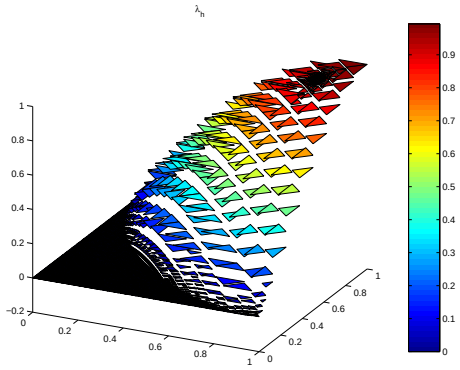

Figure 19: Example 3: co-control function λ_h at iteration $k = 12$ of Table 11

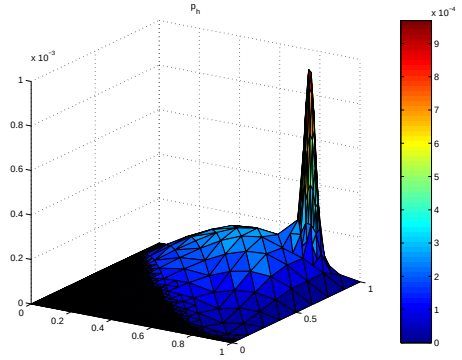

Figure 20: Example 3: co-state function p_h at iteration $k = 12$ of Table 11

In Figures 14 and 15 we depict the behaviour of the functional J_h during the adaptive process of the ∇-`DWR Algorithm`.

All the Tables concerning the ∇-`DWR Algorithm` show that each refinement based on $\mathcal{E}_1$ yields a strong reduction of the error $\|u - u_h\|_0$ and each refinement based on $\mathcal{E}_2$ yields a strong reduction of the error $|y - y_h|_1$.

In Figures 16-20 we plot the adapted mesh after 12 iterations and the corresponding discrete solutions y_h, p_h, u_h, λ_h obtained for $\Theta_{\delta J_h} = \Theta_{\mathrm{DWR}} = 0.4$.

Remark 6.2. It is important to comment on the histories of convergence depicted in Figures 14-15 and detailed in the corresponding Tables. At first sight, they may seem to be in contrast with the basic feature of the ∇-`DWR Algorithm` (i.e. being a descent-type method). However, a more careful analysis shows that the value of the functional

(a) decreases when the refinement is based only on $\mathcal{E}_1$ (the size of the gradient);

(b) increases when the refinement is based also on $\mathcal{E}_2$ (the DWR method).

This latter behavior is not surprising due to the highly singular behavior of the exact solution y^* and to the aim of the DWR method at improving the accuracy in the computation of the functional. Indeed, whenever $\mathcal{E}_2$ is used, the algorithm corrects the value of the functional J_h, which, due to the singularity of y^*, can vary substantially from one iteration to the other.

7 Conclusions

In this paper we presented a gradient-type adaptive finite element algorithm based on the DWR method for the approximate solution of an optimal control constrained problem governed by an elliptic PDE. The algorithm is based on the principle of separating the sources of the error in the adaptive process and to automatically choose the source to work on, in order to reduce the corresponding error. All the numerical experiments assess the effectiveness of this adaptive mechanism. In particular they show that:

- the refinement driven by $\mathcal{E}_1$ aims at reducing the functional value, by mainly working on the approximation of the control function u;
- the refinement driven by he DWR-based quantity $\mathcal{E}_2$ aims at keeping the discrete functional J_h close, up to a certain precision, to the continuous functional J, by mainly working on the approximation of the state variable y;
- the two refinement processes can co-operate and they yield the combined effect of minimizing the discrete functional, while increasing its accuracy.

Finally the overall behavior of our scheme shows that it is competitive with the well-established algorithm introduced in [17].

8 Appendix: Residual based adaptive algorithm

For the ease of the reader we briefly recall the essential ideas of the adaptive finite element algorithm introduced in [17], which is based on a residual type error estimator consisting of easily computable element and edge residuals with respect to the finite element approximations y_h and p_h of the state (or primal) function y and the co-state (or dual) function p, as well as of data oscillations.

Let us define the element residuals $\eta_{y,T}, \eta^{(i)}_{p,T}$, $i = 1, 2$ and the edge residuals $\eta_{y,E}, \eta_{p,E}$ as follows:

$$\begin{aligned}
\eta_{y,T} &:= h_T \|f + u_h\|_{0,T}, \\
\eta^{(1)}_{p,T} &:= h_T \|z_d - y_h\|_{0,T}, \\
\eta^{(2)}_{p,T} &:= h_T \|\Pi_0 p_h - p_h\|_{0,T}, \\
\eta_{y,E} &:= h_E^{1/2} \|\nu_E \cdot [\nabla y_h]\|_{0,E}, \\
\eta_{p,E} &:= h_E^{1/2} \|\nu_E \cdot [\nabla p_h]\|_{0,E},
\end{aligned}$$

where $E = T_1 \cap T_2$, $T_i \in \mathcal{T}$, $i = 1, 2$ and ν_E is the exterior unit normal vector on E directed towards T_2, whereas $[\nabla y_h]$ and $[\nabla p_h]$ denote the jumps of $\nabla y_h, \nabla p_h$ across E.

Moreover, let us define the low order data oscillations:

$$\begin{aligned}
\mu_h(u_d) &:= \|u_d - \Pi_0 u_d\|_{0,T}, \\
\mu_h(b) &:= \|b - b_h\|_{0,T},
\end{aligned}$$

as well as the data oscillations:

$$\begin{aligned}
osc_T(y_d) &:= h_T \|z_d - \Pi_0 z_d\|_{0,T}, \\
osc_T(f) &:= h_T \|f - f_h\|_{0,T}.
\end{aligned}$$

In what follows we drop for simplicity the iteration index k, when there will be no ambiguity.

Algorithm (`Residual-based Algorithm`)

$k = 0$

`initialize`

`apply the pdas algorithm and compute` J_h

`while` $\left(\mathrm{d}J/\mathrm{d}N^{(k)} > \mathrm{TOL}^{\mathrm{stop}} \text{ or } \mathrm{dJ/dN^{(k-1)}} > 10\,\mathrm{TOL}^{\mathrm{stop}}\right)$ {

$k = k + 1$

`apply the pdas algorithm and compute` J_h

`compute` $\eta_{y,E}$, $\eta_{p,E}$, $\forall E \in \mathcal{E}$

`compute` $\eta_{y,T}$, $\eta_{p,T}^{(1)}$, $\eta_{p,T}^{(2)}$, $\mu_T(u_d)$, $\mu_T(b)$, $osc_T(y_d)$, $osc_T(f)$, $\forall T \in \mathcal{T}$

`mark`

`refine(`$\mathcal{M}_T$`)` }

Algorithm (`mark`)

`set` $\mathcal{M}_E = \emptyset,\ \mathcal{M}_T = \emptyset$

`while` $\left(\Theta_1 \sum_{E\in\mathcal{E}} \eta_{y,E}^2 + \eta_{p,E}^2 \leq \sum_{E\in\mathcal{M}_E} \eta_{y,E}^2 + \eta_{p,E}^2 \right)$ {

`let` $E_{\max}$ `the edge that maximize` $\eta_{y,E}^2 + \eta_{p,E}^2$ `in` $\mathcal{E}\backslash\mathcal{M}_E$

`set` $\mathcal{M}_E = \mathcal{M}_E \cup E_{\max}$}

`set` $\mathcal{M}_T = \{T \in \mathcal{T} \mid \mathrm{card}(\mathcal{E}(T) \cap \mathcal{M}_E) > 2\}$

`while` $\left(\Theta_2 \sum_{T\in\mathcal{T}} \eta_{y,T}^2 + \left(\eta_{p,T}^{(1)}\right)^2 + \left(\eta_{p,T}^{(2)}\right)^2 \leq \sum_{T\in\mathcal{M}_T} \eta_{y,T}^2 + \left(\eta_{p,T}^{(1)}\right)^2 + \left(\eta_{p,T}^{(2)}\right)^2 \right)$ {

`let` $T_{\max}$ `the triangle that maximize` $\eta_{y,T}^2 + \left(\eta_{p,T}^{(1)}\right)^2 + \left(\eta_{p,T}^{(2)}\right)^2$ `in` $\mathcal{T}\backslash\mathcal{M}_T$

`set` $\mathcal{M}_T = \mathcal{M}_T \cup T_{\max}$}

`while` $\left(\Theta_3 \sum_{T\in\mathcal{T}} \mu_T^2(u_d) + \mu_T^2(b) \leq \sum_{T\in\mathcal{M}_T} \mu_T^2(u_d) + \mu_T^2(b) \right)$ {

`let` $T_{\max}$ `the triangle that maximize` $\mu_T^2(u_d) + \mu_T^2(b)$ `in` $\mathcal{T}\backslash\mathcal{M}_T$

`set` $\mathcal{M}_T = \mathcal{M}_T \cup T_{\max}$}

`while` $\left(\Theta_4 \sum_{T\in\mathcal{T}} osc_T^2(y_d) + osc_T^2(f) \leq \sum_{T\in\mathcal{M}_T} osc_T^2(y_d) + osc_T^2(f) \right)$ {

`let` $T_{\max}$ `the triangle that maximize` $\eta_{y,T}^2 + \left(\eta_{p,T}^{(1)}\right)^2 + \left(\eta_{p,T}^{(2)}\right)^2$ `in` $\mathcal{T}\backslash\mathcal{M}_T$

`set` $\mathcal{M}_T = \mathcal{M}_T \cup T_{\max}$}

Acknowledgements

Research partially supported by Italian funds MIUR-PRIN-2006 "Nuove tecniche di accoppiamento di modelli e di metodi numerici nel trattamento di problemi differenziali"

References

[1] M. Ainsworth and J. T. Oden. *A posteriori error estimation in finite element analysis.* Pure and Applied Mathematics (New York). Wiley-Interscience [John

Wiley & Sons], New York, 2000.

[2] I. Babuška and T. Strouboulis. *The finite element method and its reliability*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York, 2001.

[3] R. Becker. Estimating the control error in discretized PDE-constrained optimization. *J. Numer. Math.*, 14(3):163–185, 2006.

[4] R. Becker, H. Kapp, and R. Rannacher. Adaptive finite element methods for optimal control of partial differential equations: basic concept. *SIAM J. Control Optim.*, 39(1):113–132 (electronic), 2000.

[5] 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(4):237–264, 1996.

[6] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.*, 10:1–102, 2001.

[7] R. Becker and B. Vexler. Optimal control of the convection-diffusion equation using stabilized finite element methods. *Numer. Math.*, 106(3):349–367, 2007.

[8] M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch. A comparison of a moreau–yosida-based active set strategy and interior point methods for constrained optimal control problems. *SIAM Journal on Optimization*, 11(2):495–521, 2000.

[9] M. Bergounioux, K. Ito, and K. Kunisch. Primal-dual strategy for constrained optimal control problems. *SIAM J. Control Optim.*, 37(4):1176–1194 (electronic), 1999.

[10] S. Berrone and T. Kozubek. An adaptive WEM algorithm for solving elliptic boundary value problems in fairly general domains. *SIAM Journal of Scientific Computing*, 28(6):2114–2138, 2006.

[11] S. Berrone and M. Verani. An adaptive gradient-dwr fe algorithm for an optimal control constrained problem. Technical Report 34, Dipartimento di Matematica, Politecnico di Torino, 2007.

[12] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 2002.

[13] I. Daubechies. *Ten Lectures on Wavelets*. Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, 1992.

[14] L. Dedè and A. Quarteroni. Optimal control and numerical adaptivity for advection-diffusion equations. *M2AN Math. Model. Numer. Anal.*, 39(5):1019–1040, 2005.

[15] A. Gaevskaya, R.H.W. Hoppe, Y. Iliash, and M. Kieweg. Convergence analysis of an adaptive finite element method for distributed control problems with control constraints. In *Optimal Control for PDEs*, 2006.

[16] M. Hintermüller and R.H.W. Hoppe. Goal-oriented adaptivity in control constrained optimal control of partial differential equations. Technical report, 2007.

[17] M. Hintermüller, R.H.W. Hoppe, Y. Iliash, and M. Kieweg. An a posteriori error analysis of adaptive finite element methods for distributed elliptic control problems with control constraints. *ESAIM Control, Optimisation and Calculus of Variations*, 2006.

[18] K. Ito and K. Kunisch. Augmented Lagrangian formulation of nonsmooth convex optimization in Hilbert spaces. In *Control of partial differential equations and applications (Laredo, 1994)*, volume 174 of *Lecture Notes in Pure and Appl. Math.*, pages 107–117. Dekker, New York, 1996.

[19] R. Li, W. Liu, H. Ma, and T. Tang. Adaptive finite element approximation for distributed elliptic optimal control problems. *SIAM J. Control Optim.*, 41(5):1321–1349 (electronic), 2002.

[20] J.L. Lions. *Optimal control of systems governed by partial differential equations.* Springer-Verlag, New York, 1971.

[21] Y. Meyer. *Wavelets and operators*, volume 37 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1992.

[22] P Morin, R.H. Nochetto, and M. Verani. Adaptive finite elements for a shape optimization problem. in preparation.

[23] K.G. Siebert and A. Veeser. A unilaterally constrained quadratic minimization with adaptive finite elements. *SIAM J. Optim.*, 18(1):260–289 (electronic), 2007.

[24] R. Verfurth. *A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*. Teubner Verlag and J. Wiley, 1996.

[25] B. Vexler and W. Wollner. Adaptive finite elements for elliptic optimization problems with control constraints. *SIAM Journal on Control and Optimization*, 2007.

[26] O. C. Zienkiewicz and J. Z. Zhu. The superconvergent patch recovery and a posteriori error estimates. I. The recovery technique. *Internat. J. Numer. Methods Engrg.*, 33(7):1331–1364, 1992.

[27] O. C. Zienkiewicz and J. Z. Zhu. The superconvergent patch recovery and a posteriori error estimates. II. Error estimates and adaptivity. *Internat. J. Numer. Methods Engrg.*, 33(7):1365–1382, 1992.

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

2/2008 S. Berrone, M. Verani:
An Adaptive Gradient-DWR Finite Element Algorithm for an Optimal Control Constrained Problem.

1/2008 R. Rosso, M. Verani:
Stabilizing rôle of a curvature correction to line tension

26/2007 M. Colecchia, N. Nicolai, P. Secchi, G. Bandieramonte, A.M. Paganoni, L.M. Sangalli, L. Piva, R. Salvioni:
Penile Superficial Squamous cell Carcinoma (SCC) Submitted to CO_2 Laser Excision only: Oncologic Outcome of T1 Disease in 25 Years-Long Experience

25/2007 S. Micheletti, S. Perotto:
Space-Time Adaption for Advection-Diffusion-Reaction Problems on Anisotropic Meshes

24/2007 L.M. Sangalli, P. Secchi, S. Vantini, A. Veneziani:
A Case Study in Functional Data Analisys: Geometrical Features of the Internal Carotid Artery

23/2007 L.M. Sangalli, P. Secchi, S. Vantini, A. Veneziani:
Efficient estimation of 3-dimensional centerlines of inner carotid arteries and their curvature functions by free knot regression splines

22/2007 G.O. Roberts, L.M. Sangalli:
Latent diffusion models for event history analysis

21/2007 P. Massimi, A. Quarteroni, F. Saleri, G. Scrofani:
Modeling of Salt Tectonics

20/2007 E. Burman, A. Quarteroni, B. Stamm:
Stabilization Strategies for High Order Methods for Transport Dominated Problems

19/2007 E. Burman, A. Quarteroni, B. Stamm:
Interior Penalty Continuous and Discontinuous Finite Element Approximations of Hyperbolic Equations