NUMERICAL APPROXIMATION OF A CONTROL PROBLEM FOR ADVECTION–DIFFUSION PROCESSES

Alfio QUARTERONI*† Gianluigi ROZZA*

Luca DEDE'† Annalisa QUAINI*

21st September 2005

* École Polytechnique Fédérale de Lausanne (EPFL), FSB,
Chaire de Modelisation et Calcul Scientifique (CMCS),
Station 8, 1015, Lausanne, Switzerland.
alfio.quarteroni@epfl.ch
gianluigi.rozza@epfl.ch
annalisa.quaini@epfl.ch

† MOX– Modeling and Scientific Computing,
Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano,
via Bonardi 9, 20133, Milano, Italy.
luca.dede@mate.polimi.it

Abstract

Two different approaches are proposed to enhance the efficiency of the numerical resolution of optimal control problems governed by a linear advection–diffusion equation. In the framework of the Galerkin–Finite Element (FE) method, we adopt a novel a posteriori error estimate of the discretization error on the cost functional; this estimate is used in the course of a numerical adaptive strategy for the generation of efficient grids for the resolution of the optimal control problem. Moreover, we propose to solve the control problem by adopting a reduced basis (RB) technique, hence ensuring rapid, reliable and repeated evaluations of input–output relationship. Our numerical tests show that by this technique a substantial saving of computational costs can be achieved.

Keywords: optimal control problems; partial differential equations; finite element approximation; reduced basis techniques; advection–diffusion equations; stabilized Lagrangian; numerical adaptivity.

1 Introduction

Many physical processes, which involve diffusion and transport of scalar quantities, can be modelled by linear advection–diffusion partial differential equations. These phenomena are studied, e.g., in Environmental Sciences, to investigate the distribution forecast of pollutants in water or in atmosphere. In this context it might be of interest to regulate the source term of the advection–diffusion equation so that the solution is as near as possible to a desired one, e.g., to operate the emission rates of industrial plants to keep the concentration of pollutants near (or below) a desired level.

This problem can be conveniently accommodated in the optimal control framework for PDEs, where we consider the Lagrangian functional formulation [3], as complementary to the classical approach of J.L. Lions [6].

To avoid numerical instabilities that arise in a transport dominating regime, we propose a stabilization on the Lagrangian functional [5]. We consider two numerical approaches that allow an efficient resolution of the optimal control problem, in the context of an iterative optimization procedure. In the first case we solve the equations governing the control problem by means of the Galerkin–FE method. Grid adaptivity is driven by *a posteriori* error estimate on the cost functional, which we assume as an indicator of the whole error on the control problem [3]. Moreover, we propose a separation of the *iteration* and *discretization* error [5], for which we define a posteriori error estimate. As soon as the iteration error is brought below a desired threshold by means of the iterative optimization method, we operate the adaptive strategy to reduce the discretization error [5]. Then we solve numerically the equations governing the control problem by means of the *reduced basis* (RB) method [8], which leads to a large saving of computational costs. In fact the RB method permits a rapid, reliable and repeated evaluation of the input–output relationship [7]; in the case of the control problem the inputs are the control function for the state equation, and the observation for the adjoint one, while the outputs are respectively the state variable and the adjoint one.

At the end we report some numerical tests to validate the methods here presented, referring in particular to a pollution control problem in atmosphere.

2 Mathematical Model of the Control Problem

In this section we recall the Lagrangian functional approach for optimal control problems and the associated iterative optimization method, in a general setting [3]; then we specialize it to an advection–diffusion control problem.

2.1 The general setting

Let us consider the following control problem:

$$find \quad u \in \mathcal{U} \quad : \quad J(w,u) \quad minimum, \quad with \quad Aw = f + Bu, \tag{1}$$

where $w \in \mathcal{V}$ is the state variable, u the control function, A is an elliptic operator defined on $\mathcal{V}$ with values in $\mathcal{V}'$, B is an operator defined on $\mathcal{U}$ and valued in $\mathcal{V}'$, f is a source term, $\mathcal{V}$ and $\mathcal{U}$ are two Hilbert spaces. We write the state equation $Aw = f + Bu$ in weak form: $find\ w \in \mathcal{V}\ :\ a(w,\varphi) = (f,\varphi) + b(u,\varphi),\ \forall \varphi \in \mathcal{V}$. The associated Lagrangian functional reads:

$$\mathcal{L}(w,p,u) := J(w,u) + b(u,p) + (f,p) - a(w,p), \tag{2}$$

where $a(\cdot,\cdot)$ and $b(\cdot,\cdot)$ are the bilinear forms associated with A and B, respectively, $(\cdot,\cdot)$ is the L^2–inner product, while $p \in \mathcal{V}$ is the adjoint variable. Should there exist, the solution of the control problem (w^*, p^*, u^*) is the stationary point of $\mathcal{L}(w,p,u)$. By differentiating the Lagrangian functional, we obtain the Euler–Lagrange system governing the optimal control problem:

$$\begin{cases} \mathcal{L}_{,w}[\phi] = 0 \ \longrightarrow \ find \ \ p \in \mathcal{V} \ \ : \ \ a(\phi,p) = J_{,w}(w,u)[\phi], \ \ \forall \phi \in \mathcal{V}, \\ \mathcal{L}_{,p}[\varphi] = 0 \ \longrightarrow \ find \ \ w \in \mathcal{V} \ \ : \ \ a(w,\varphi) = (f,\varphi) + b(u,\varphi), \ \ \forall \varphi \in \mathcal{V}, \\ \mathcal{L}_{,u}[\psi] = 0 \ \longrightarrow J_{,u}(w,u)[\psi] + b(\psi,p) = 0, \ \ \forall \psi \in \mathcal{U}. \end{cases} \tag{3}$$

The first equation in (3) is the adjoint equation, the second one is the state equation, while, by the Riesz theorem, from the third one we can extract the sensitivity of the cost functional δu with respect to the control function u ($\mathcal{L}_{,u}[\psi] = (\delta u(p,u), \psi)$). The control problem can be solved by means of an iterative method [1]. At each step j we solve sequentially the state and the adjoint equation and we compute the sensitivity $\delta u(p^j, u^j)$; then we evaluate the latter in an appropriate norm, which we compare with a prescribed tolerance. If this stopping criterium is not fulfilled, we adopt an optimization iteration on the control function u, such as the *steepest–descent* method, $u^{j+1} = u^j - \tau^j \delta u(p^j, u^j)$, where τ^j is a relaxation parameter.

2.2 The case of an advection–diffusion problem

Let us consider now the specific case of a linear advection–diffusion state equation, referring to a $2D$–domain Ω:

$$\begin{cases} L(w) := -\nabla \cdot (\nu \nabla w) + \mathbf{V} \cdot \nabla w = u, \quad in\ \Omega, \\ w = 0, \quad on\ \Gamma_D, \\ \nu \frac{\partial w}{\partial n} = 0, \quad on\ \Gamma_N, \end{cases} \tag{4}$$

Γ_D and Γ_N are two disjoint portions of the domain boundary $\partial\Omega$ such that $\Gamma_D \cup \Gamma_N = \partial\Omega$, $u \in L^2(\Omega)$ is the control variable, while ν and $\mathbf{V}$ are given

functions. We assume homogeneous Dirichlet condition on the inflow boundary $\Gamma_D := \{\mathbf{x} \in \partial\Omega : \mathbf{V}(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) < 0\}$, being $\mathbf{n}(\mathbf{x})$ the unit vector directed outward, and homogeneous Neumann condition on the outflow boundary $\Gamma_N := \partial\Omega \setminus \Gamma_D$. We consider the observation on a part $D \subseteq \Omega$ of the domain, for which the control problem reads:

$$find\ u \quad : \quad J(w,u) := \frac{1}{2}\int_D (g\ w(u) - z_d)^2\ dD \quad minimum, \tag{5}$$

where $g \in C^\infty(\Omega)$ projects w in the observation space and z_d is the desired observation function. Adopting the formalism of the previous section and assuming $\mathcal{V} = H^1_{\Gamma_D} := \{v \in H^1(\Omega) : v_{|\Gamma_D} = 0\}$ and $\mathcal{U} = L^2(\Omega)$, the Lagrangian functional becomes:

$$\mathcal{L}(w,p,u) := J(w,u) + F(p;u) - a(w,p). \tag{6}$$

where:

$$a(w,\varphi) := \int_\Omega \nu\nabla w \cdot \nabla\varphi\ d\Omega + \int_\Omega \mathbf{V}\cdot\nabla w\ \varphi\ d\Omega, \tag{7}$$

$$F(\varphi;u) := \int_\Omega u\varphi\ d\Omega. \tag{8}$$

By differentiating $\mathcal{L}$ with respect to the state variable, we obtain the adjoint equation in weak form:

$$find\ p \in \mathcal{V} \quad : \quad a^{ad}(p,\phi) = F^{ad}(\phi;w), \quad \forall\phi \in \mathcal{V}, \tag{9}$$

with:

$$a^{ad}(p,\phi) := \int_\Omega \nu\nabla p \cdot \nabla\phi\ d\Omega + \int_\Omega \mathbf{V}\cdot\nabla\phi\ p\ d\Omega, \tag{10}$$

$$F^{ad}(\phi;w) = \int_D (g\ w - z_d)\ g\ \phi\ dD. \tag{11}$$

In the distributional sense this yields:

$$\begin{cases} L^{ad}(p) := -\nabla\cdot(\nu\nabla p + \mathbf{V}p) = \chi_D g\ (g\ w - z_d), \quad in\ \Omega, \\ p = 0, \quad on\ \Gamma_D, \\ \nu\frac{\partial p}{\partial n} + \mathbf{V}\cdot\mathbf{n}\ p = 0, \quad on\ \Gamma_N, \end{cases} \tag{12}$$

being χ_D the characteristic function of the subdomain D. Finally, by differentiating $\mathcal{L}$ with respect to the control function u, we have the optimal control constraint, from which we define the cost functional sensitivity: $\delta u(p) = p$.

3 Numerical Approximation and Stabilization

For the numerical resolution of both the state and adjoint equations, we adopt the Galerkin–FE method with linear elements on unstructured triangular meshes. Both equations are of advection–diffusion type with a transport term that can dominate the diffusive one; when it happens an appropriate stabilization is mandatory to avoid numerical instabilities and their propagation in the course of the optimization iterative procedure ([9]).

3.1 "Optimize–then–discretize" and "discretize–then–optimize" approaches

From a numerical point of view, the algorithm outlined in Sec.2.1 (or in Sec.2.2 for the specific advection–diffusion case) requires, at each iterative step, the approximation of the state and adjoint equations. This approximation can be based, e.g., on a suitable FE subspace $\mathcal{X}_h \subset \mathcal{V}$ and the GLS (Galerkin–Least–Squares) method [9], obtaining respectively:

$$find\ w_h \in \mathcal{X}_h \quad : \quad a(w_h, \varphi_h) + \overline{s}_h(w_h, \varphi_h) = F(\varphi_h; u_h), \quad \forall \varphi_h \in \mathcal{X}_h, \tag{13}$$

$$\overline{s}_h(w_h, \varphi_h) := \sum_{K \in \mathcal{T}_h} \delta_K \int_K R(w_h; u_h)\ L(\varphi_h)\ dK, \tag{14}$$

$$find\ p_h \in \mathcal{X}_h \quad : \quad a^{ad}(p_h, \phi_h) + \overline{s}_h^{ad}(p_h, \phi_h) = F^{ad}(\phi_h; w_h), \quad \forall \phi_h \in \mathcal{X}_h, \tag{15}$$

$$\overline{s}_h^{ad}(p_h, \phi_h) := \sum_{K \in \mathcal{T}_h} \delta_K \int_K R^{ad}(p_h; w_h)\ L^{ad}(\phi_h)\ dK, \tag{16}$$

where δ_K is a stabilization parameter, $R(w; u) := L(w) - u$, $R^{ad}(p; w) := L^{ad}(p) - G(w)$, with $G(w) := \chi_D g\ (g\ w - z_d)$. This paradigm is resumed in the slogan "optimize–then–discretize" [2, 4]. An alternative approach is "discretize–then–optimize", for which first we discretize and stabilize the state equation, e.g. still by the GLS method (Eq.(13) and (14)), then we define the discrete Lagrangian functional:

$$\mathcal{L}_h(w_h, p_h, u_h) := J(w_h, u_h) + F(p_h; u_h) - a(w_h, p_h) - \overline{s}_h(w_h, p_h), \tag{17}$$

from which, by differentiation with respect to w_h, we obtain the discrete adjoint equation (15), however with the following stabilization term:

$$\overline{s}_h^{ad}(p_h, \phi_h) = \overline{\overline{s}}_h^{ad}(p_h, \phi_h) := \sum_{K \in \mathcal{T}_h} \delta_K \int_K L(\phi_h)\ L(p_h)\ dK. \tag{18}$$

Differentiating $\mathcal{L}_h$ with respect to u_h and applying the Riesz theorem, being $u_h \in \mathcal{X}_h$, we obtain: $\delta u_h = p_h + \sum_{K \in \mathcal{T}_h} \delta_K \int_K L(p_h)\ dK$.

3.2 The stabilized Lagrangian approach

We consider a stabilization on the Lagrangian functional itself [5], for which our stabilized Lagrangian functional is:

$$\mathcal{L}_h^s(w_h, p_h, u_h) := \mathcal{L}(w_h, p_h, u_h) + S_h(w_h, p_h, u_h), \tag{19}$$

with:

$$S_h(w, p, u) := \sum_{K \in \mathcal{T}_h} \delta_K \int_K R(w; u)\ R^{ad}(p; w)\ dK. \tag{20}$$

This approach can be regarded as a particular case of the "discretize–then–optimize" one if we identify $\overline{s}_h(w_h, p_h)$ with $-S_h(w_h, p_h, u_h)$. By differentiating $\mathcal{L}_h^s$ we obtain the (stabilized) approximate state and adjoint equations (13) and (15), assuming $\overline{s}(w_h, \varphi_h) = s_h(w_h, \varphi_h; u_h)$ and $\overline{s}^{ad}(p_h, \phi_h) = s_h^{ad}(p_h, \phi_h; w_h)$, where:

$$s_h(w_h, \varphi_h; u_h) := -\sum_{K \in \mathcal{T}_h} \delta_K \int_K R(w_h; u_h)\ L^{ad}(\varphi_h)\ dK, \tag{21}$$

$$s_h^{ad}(p_h, \phi_h; w_h) := -\sum_{K \in \mathcal{T}_h} \delta_K \int_K \Big(R^{ad}(p_h; w_h)\ L(\phi_h) - R(w_h; u_h)\ G'(\phi_h) \Big)\ dK, \tag{22}$$

having set $G'(w) := \chi_D g^2 w$. Finally, the cost functional sensitivity reads:

$$\delta u_h(p_h, w_h) = p_h - \sum_{K \in \mathcal{T}_h} \delta_K\ R^{ad}(p_h; w_h). \tag{23}$$

4 A Posteriori Error Estimate

For the definition of an appropriate error estimate for the optimal control problem, we identify the error on the control problem as being the error on the cost functional, as proposed in [3]. Moreover, we propose to separate this error in two parts: the iteration and the discretization error. For the latter we define a suitable estimate according with the duality principles [3], adopted in the course of mesh adaptive strategy.

4.1 Iteration and discretization errors

At each iterative step j of the optimization procedure we consider the following error:

$$|\varepsilon^{(j)}| = |J(w^*, u^*) - J(w_h^j, u_h^j)|, \tag{24}$$

where $*$ indicates optimal variables, while w_h^j stands for the discrete variable evaluated at the step j. If we refine the mesh, according with an adaptive procedure, we certainly reduce the component of the full error $\varepsilon^{(j)}$ (Eq.(24)) related to the numerical approximation at the step j, which we call the *discretization error* $\varepsilon_D^{(j)}$. On the other hand, the part of $\varepsilon^{(j)}$ expressing the difference between the cost functional computed on continuous variables at the step j and the optimal cost functional, which we call the *iteration error* $\varepsilon_{IT}^{(j)}$, can generally increase [5]. From Eq.(24):

$$\varepsilon^{(j)} = \big(J(w^*, u^*) - J(w^j, u^j)\big) + \Big(J(w^j, u^j) - J(w_h^j, u_h^j)\Big) = \varepsilon_{IT}^{(j)} + \varepsilon_D^{(j)}; \tag{25}$$

then we will define a posteriori error estimate only for $\varepsilon_D^{(j)}$, the only part of $\varepsilon^{(j)}$ which can be reduced by mesh refinement. Since $\nabla\mathcal{L}(\mathbf{x})$ is linear in $\mathbf{x}$, the

iteration error $\varepsilon_{IT}^{(j)}$ becomes $\varepsilon_{IT}^{(j)} = \frac{1}{2}\ (\ \delta u(p^j, u^j)\ ,\ u^* - u^j\ )$, which, in the case of our advection–diffusion control problem (see Sec.2.2 and [5]), can be written as:

$$\varepsilon_{IT}^{(j)} = -\frac{1}{2}\tau\|p^j\|^2_{L^2(\Omega)} - \frac{1}{2}\tau \sum_{r=j+1}^{\infty} (\ p^j, p^r\ )_{L^2(\Omega)}. \tag{26}$$

Since the iteration error can not be correctly evaluated by means of this expression, we can assume that $|\varepsilon_{IT}^{(j)}| \approx \frac{1}{2}\tau\|p^j\|^2_{L^2(\Omega)}$, or, more simply $|\varepsilon_{IT}^{(j)}| \approx \|p^j\|^2_{L^2(\Omega)}$, which leads to the usual criterium $|\varepsilon_{IT}^{(j)}| \approx \|\delta u(p^j)\|$ (L^2–norm).

4.2 A posteriori error estimate and adaptive strategy

We define the a posteriori error estimate for the discretization error only, based on the following theorem ([5]).

Theorem 4.1 *For a linear control problem with the stabilized Lagrangian $\mathcal{L}_h^s$ (Eq.(19) and Eq.(20)), the discretization error at the j–th iteration reads:*

$$\varepsilon_D^{(j)} = \frac{1}{2}(\ \delta u(p^j, u^j), u^j - u_h^j\ ) + \frac{1}{2}\nabla\mathcal{L}_h^s(\mathbf{x}_h^j)\cdot(\mathbf{x}^j - \mathbf{x}_h^j) + \Lambda_h(\mathbf{x}_h^j), \tag{27}$$

where $\mathbf{x}_h^j := (w_h^j, p_h^j, u_h^j)$ is the Galerkin–FE approximation and $\Lambda_h(\mathbf{x}_h^j) := S_h(\mathbf{x}_h^j) + s_h(w_h^j, p_h^j; u_h^j)$, being $s_h(w_h^j, p_h^j; u_h^j)$ the stabilization term (21).

Applying (27) to our advection–diffusion control problem and highlighting the contributions on the elements of the mesh $K \in \mathcal{T}_h$ ([3]), we obtain the following estimate:

$$|\varepsilon_D^{(j)}| \le \eta_D^{(j)} := \frac{1}{2}\sum_{K\in\mathcal{T}_h}\left\{\ \left(\omega_K^p\rho_K^w + \omega_K^w\rho_K^p + \omega_K^u\rho_K^u\right) + \lambda_K\ \right\}, \tag{28}$$

where, according with the symbol definitions given in Sec.3:

$$\begin{aligned}
&\rho_K^w := \|R(w_h^j; u_h^j)\|_K + h_K^{-\frac{1}{2}}\|r(w_h^j)\|_{\partial K},\\
&\omega_K^p := \|(p^j - p_h^j) - \delta_K L^{ad}(p^j - p_h^j) + \delta_K G'(w^j - w_h^j)\|_K + h_K^{\frac{1}{2}}\|p^j - p_h^j\|_{\partial K},\\
&\rho_K^p := \|R^{ad}(p_h^j; w_h^j)\|_K + h_K^{-\frac{1}{2}}\|r^{ad}(p_h^j)\|_{\partial K},\\
&\omega_K^w := \|(w^j - w_h^j) - \delta_K L(w^j - w_h^j)\|_K + h_K^{\frac{1}{2}}\|w^j - w_h^j\|_{\partial K},\\
&\rho_K^u := \|\delta u_h(p_h^j, w_h^j) + \delta u(p^j)\|_K = \|p^j + p_h^j - \delta_K R^{ad}(p_h^j; w_h^j)\|_K,\\
&\omega_K^u := \|u^j - u_h^j\|_K,\\
&\lambda_K := 2\delta_K\|R(w_h^j; u_h^j)\|_K\ \|G(w_h^j)\|_K,\\
&r(w_h^j) := \begin{cases} -\frac{1}{2}\left[\nu\frac{\partial w_h^j}{\partial n}\right], & on\ \partial K\backslash\partial\Omega,\\ -\nu\frac{\partial w_h^j}{\partial n}, & on\ \partial K\in\Gamma_N,\end{cases}\\
&r^{ad}(p_h^j) := \begin{cases} -\frac{1}{2}\left[\nu\frac{\partial p_h^j}{\partial n} + \mathbf{V}\cdot\mathbf{n}\ p_h^j\right], & on\ \partial K\backslash\partial\Omega,\\ -\left(\nu\frac{\partial p_h^j}{\partial n} + \mathbf{V}\cdot\mathbf{n}\ p_h^j\right), & on\ \partial K\in\Gamma_N;\end{cases}
\end{aligned} \tag{29}$$

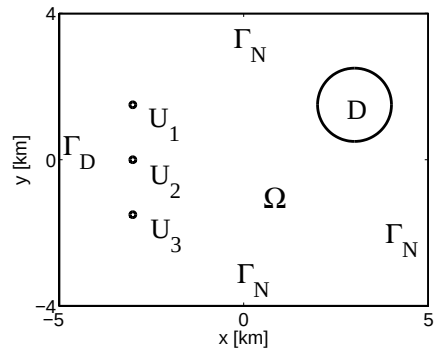


Figure 1: Domain for the pollution problem.

∂K indicates the boundary of $K \in \mathcal{T}_h$, while $[\cdot]$ stands for the jump of the embraced quantity across ∂K.

To use the estimate (28), we need to evaluate w^j, p^j and u^j. Indeed, we replace w^j and p^j by the respective quadratic reconstructions, $(w_h^j)^q$ and $(p_h^j)^q$, and u^j by $(u_h^j)^q := u_h^j - \tau(\delta u_h((p_h^j)^q, (w_h^j)^q) - \delta u_h(p_h^j, w_h^j))$, according to the steepest-descent iterative method with $\tau^j = \tau$. The following adaptive strategy is then adopted to allow an efficient generation of adapted meshes:

1. we adopt the optimization iterative method till convergence to the iteration error tolerance Tol_{IT}, assuming an initial coarse mesh;

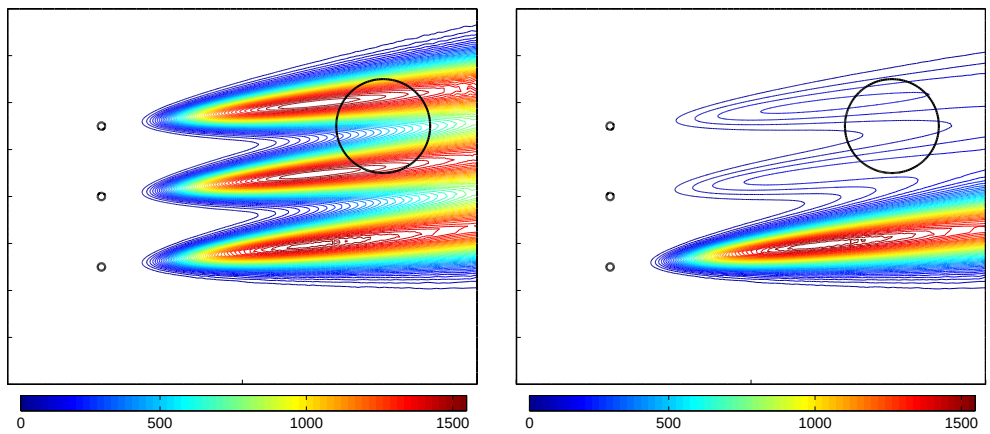
2. we adapt the mesh, balancing the error on the elements $K \in \mathcal{T}_h$, according with the error estimate $\eta_D^{(j)}$ (28), till convergence to the discretization error tolerance Tol_D;

3. we re-evaluate the variables and $\varepsilon_{IT}^{(j)}$ on the adapted mesh: if $\varepsilon_{IT}^{(j)} \geq Tol_{IT}$, we return to point 1 and we repeat the procedure, while if $\varepsilon_{IT}^{(j)}$ is inferior to Tol_{IT}, we stop.

5 A Numerical Test: Pollution Control Problem

We apply the a posteriori error estimates $\eta_D^{(j)}$ (28) for the discretization error and the strategy presented in Sec.4.2 to a numerical test, which can be regarded as a pollution control problem in atmosphere. Our goal consists in regulating the emissions of industrial chimneys to keep the pollutant concentration below a desired threshold in an observation area (a town).

To this aim we consider a simple advection–diffusion model [5, 8], which can be regarded as a *quasi*–$3D$ model: the pollutant concentration w at the emissive height H is described by the advection–diffusion equation introduced in Sec.2.2, while the concentration at soil is obtained by projection by means



(a) (b)

Figure 2: Pollutant concentration $[\mu g/m^3]$ at the ground before (a) and after (b) the regulation of the sources.

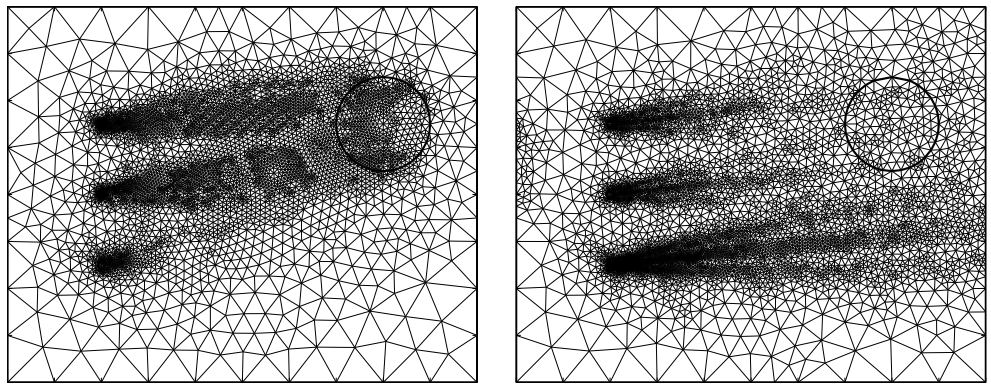
of the function $g(\mathrm{x},\mathrm{y})$ described in Sec.2.2. The values assumed by the diffusion coefficient $\nu(\mathrm{x},\mathrm{y})$ and the function $g(\mathrm{x},\mathrm{y})$ depend on the distance from the emission sources and the atmospherical stability class (stable, neutral or unstable). In particular, we consider the case of neutral atmospherical conditions and, referring to the domain reported in Fig.1, we assume $\mathbf{V} = V_{\mathrm{x}}\hat{\mathrm{x}} + V_{\mathrm{y}}\hat{\mathrm{y}}$, with $V_{\mathrm{x}} = V\cos(\frac{\pi}{30})$ and $V_{\mathrm{y}} = V\sin(\frac{\pi}{30})$, being $V = 2.5\ m/s$. Moreover we consider that the chimneys maximum rate of emission is $u_{max} = 800\ g/s$ at the emission height $H = 100\ m$, for which the pollutant concentration (we consider SO_2) is higher than the desired level $z_d = 100\ \mu g/m^3$. In Sec.2.2 we have considered the case of u distributed over all the domain Ω, while here we deal with a particular case which can be accommodated in the general case assuming $u = \sum_{i=1}^{N} u_i\chi_i$, where χ_i is the characteristic function of the chimney U_i.

In Fig.2a we report the pollutant concentration at the ground corresponding to the maximum emission rates; in Fig.2b we plot the concentration at ground at the completion of the optimization strategy; we observe that the "optimal" emission rates become $u_1 = 0.0837{\cdot}u_{max}$, $u_2 = 0.0908{\cdot}u_{max}$ and $u_3 = 1.00{\cdot}u_{max}$. In Fig.3 we report a comparison among adapted meshes; in Fig.3a that obtained by our estimator $\eta_D^{(j)}$ and in Fig.3b by the following estimators [5] (which lead to analogous results):

1. the *energy norm* indicator $(\eta_E^w)^{(j)} := \sum_{K\in\mathcal{T}_h} h_K\ \rho_K^w$;

2. the indicator $(\eta_E^{wpu})^{(j)} := \sum_{K\in\mathcal{T}_h} h_K \left\{(\rho_K^w)^2 + (\rho_K^p)^2 + (\rho_K^u)^2\right\}^{\frac{1}{2}}$.

For symbols definitions see Eq.(29); the results are compared with those obtained with a fine mesh with about 80000 elements. The adaptivity driven by the error indicator $\eta_D^{(j)}$ leads to concentrate elements in those areas that are more relevant



(a) (b)

Figure 3: Adapted meshes (about 14000 elements) obtained by $\eta_D^{(j)}$ Eq.(28) (a) and $(\eta_E^{wpu})^{(j)}$ (b) (and similarly $(\eta_E)^{(j)}$).

for the optimal control problem. This fact is underlined by comparing the errors on the cost functional and other interesting quantities for the meshes obtained with the different error indicators, but with the same number of elements. E.g., the indicator $\eta_D^{(j)}$ provides an error on the optimal cost functional J of about 20% against the 55% obtained by $(\eta_E^{w})^{(j)}$ and $(\eta_E^{wpu})^{(j)}$ with meshes with about 4000 elements, and of 6% vs. 15% with about 14000 elements. We see that the adaptivity driven by the error indicator $\eta_D^{(j)}$ permits large savings of number of mesh elements, allowing more efficient resolution of the optimal control problem.

6 A Reduced Basis Approach to Control Problems

As a second approach to improve efficiency, we consider the RB method to solve the optimal control problem, by applying the approach to the state and adjoint equations. For a review on the use of the RB method and for optimal control problems, see [7, 8, 10].

6.1 Reduced basis: abstract formulation

The RB method allows the evaluation of input–output relationships by means of a precise and efficient procedure. The goal consists in calculating a quantity (the output) $s(\mu) = l(w(\mu); \mu)$ depending on the solution of the following parametrized equation:

$$find \quad w(\mu) \in \mathcal{X} \quad : \quad a(w(\mu), v; \mu) = f(v; \mu), \quad \forall v \in \mathcal{X}, \tag{30}$$

where $\mu \in \mathcal{D}$ and $\mathcal{D}$ is a set of parameters, $\mathcal{X}$ is a Hilbert space, the form $a(\cdot, \cdot; \mu)$ is bilinear, continuous and coercive and the forms $f(\cdot; \mu)$ and $l(\cdot; \mu)$ are linear

and continuous, for all μ. Moreover, we assume that the form $a(\cdot,\cdot;\mu)$ is affine parameter dependent, that is:

$$a(w(\mu),v;\mu)=\sum_{q=1}^{Q}\sigma^q(\mu)a^q(w,v),\quad \forall w,v\in\mathcal{X},\quad \forall\mu\in\mathcal{D},\quad q=1,\dots,Q, \tag{31}$$

where $\sigma^q:\mathcal{D}\to R$ are parameter–dependent functions, while $a^q:\mathcal{X}\times\mathcal{X}\to R$ are parameter–independent forms; affine parameter dependence is required also for $f(\cdot;\mu)$ and $l(\cdot;\mu)$. To build the RB space we need to introduce a finite dimensional subspace $\mathcal{X}_h$ of $\mathcal{X}$, which we identify with a Galerkin–FE space associated with a very fine triangulation of the domain Ω. The Galerkin–FE element method consists in solving the following $\mathcal{N}$–dimensional problem $find\ w_h(\mu)\in\mathcal{X}_h\ :\ a(w_h(\mu),v;\mu)=f(v),\ \forall v\in\mathcal{X}_h$, which, if $\mathcal{N}$ is large, leads to computational expensive evaluations of the output $s_h(\mu)=l(w_h(\mu);\mu)$ for several values of the input μ. We consider a set of samples $S_N^\mu=\{\mu^i\in\mathcal{D},i=1,\dots,N\}$ and we define the N–dimensional RB space as $W_N=span\{\zeta^i,i=1,\dots,N\}$, where $\zeta^n=w_h(\mu^n)$, with $n=1,\dots,N$. The RB method consists in evaluating the output $s_N(\mu)=l(w_N(\mu);\mu)$, where $w_N(\mu)$ is given by the following problem (of dimension N):

$$find\quad w_N(\mu)\in W_N\quad :\quad a(w_N(\mu),v;\mu)=f(v),\quad \forall v\in W_N. \tag{32}$$

Then we write $w_N(\mu)$ as $w_N(\mu)=\sum_{j=1}^N w_{N_j}(\mu)\zeta^j$, being $\underline{w}_N(\mu)=\{w_{N_1}(\mu),\dots,w_{N_N}(\mu)\}^T$ the solution of the following linear system of order N:

$$\underline{A}_N(\mu)\underline{w}_N(\mu)=\underline{F}_N(\mu), \tag{33}$$

where $A_{N_{i,j}}(\mu)=a(\zeta^j,\zeta^i;\mu)$ and $F_{N_i}(\mu)=f(\zeta^i,\mu)$, with $i,j=1,\dots,N$; the output $s_N(\mu)$ is calculated as $s_N(\mu)=\underline{L}_N(\mu)^T w_N(\mu)$, where $L_{N_i}(\mu)=l(\zeta^i;\mu)$. From the affine dependence property, we can split the matrix $\underline{A}_N(\mu)$ and the vectors $\underline{F}_N(\mu)$ and $\underline{L}_N(\mu)$ into a parameter–dependent part and a parameter–independent part. In the case of matrix $\underline{A}_N(\mu)$, this means that $\underline{A}_N(\mu)=\sum_{q=1}^Q\sigma^q(\mu)\underline{A}_N^q$, where $\underline{A}_{N_{i,j}}^q=a^q(\zeta^j,\zeta^i)$ is a parameter–independent matrix. Similar expressions hold for $\underline{F}_N(\mu)$ and $\underline{L}_N(\mu)$. These decompositions allow a very convenient computational procedure composed by off–line and on–line stages. In the off–line stage we afford the larger computational costs, computing the basis ζ^n of W_N and assembling the matrices $\underline{A}_N^q$ and the vectors $\underline{F}_N^q$ and $\underline{L}_N^q$, which require N FE solutions and inner products. In the on–line stage, given μ, we assemble $\underline{A}_N(\mu)$, $\underline{F}_N(\mu)$ and $\underline{L}_N(\mu)$, we solve the system (33) and we compute $s_N(\mu)$. Let us notice that N is usually very low with respect to $\mathcal{N}$ according to the precision required on $s_N(\mu)$; this leads to large computational costs savings in case of recursive evaluation of $s_N(\mu)$ for different parameters μ.

6.2 The reduced basis method applied to control problems

The resolution of optimal control problems by an iterative method leads to a recursive resolution of the state and adjoint equations. The computational cost

of the whole procedure can therefore be quite relevant, especially if great precision is required. In this context, see Sec.2.1, the input of the state equation can be regarded as the control function u, while the output as the state variable w. Similarly, the adjoint equation input is the observation on the system, related to w, while the output is the adjoint variable p itself, which, by means of the iterative method adopted (e.g., the steepest–descent method), becomes the input for the state equation. The iterative optimization method can be seen as a recursive evaluation of an input–output relationship. Moreover, it can be interesting to perform an optimization for different values of some physical or geometrical parameters [8]: e.g., referring to the pollution control problem introduced in Sec.5, the diffusivity, the velocity of the wind, or the reciprocal distance among the chimneys. The main idea consists in parametrizing both the state and adjoint equations by the parameters of interest and the source terms (related respectively to u and w); then we solve these equations by means of the RB method in the course of the optimization iterative procedure. The strategy allows large savings of computational costs with respect to the conventional FE iterative optimization method. The RB space for the adjoint equation does not need to be coherent with that for the state equation.

6.3 Numerical tests

We report two numerical tests, which refer to the pollution control problem considered in Sec.5.
In the first case we assume that at the initial step the emissions of the chimneys are respectively the 45%, 0% and 55% of the total emission $u_{tot} = 2700\ g/s$. The optimization procedure leads to the following distribution of emissions on the three chimneys: 3.49%, 0% and 55.02% of u_{tot}. The saving in computational costs with respect to the resolution by the FE method is about the 73%, having chosen $\tau = 800$ and an error tolerance on J of 10^{-8}.
The second test considers as parameters the emissions of the chimneys and the wind velocity field, i.e. $\mu = \{u_1, u_2, u_3, V_{\mathrm{x}}, V_{\mathrm{y}}\}$, where $\mathbf{V} = V_{\mathrm{x}}\hat{\mathrm{x}} + V_{\mathrm{y}}\hat{\mathrm{y}}$, with $\sqrt{V_{\mathrm{x}}^2 + V_{\mathrm{y}}^2} = 1$. We assume $V_{\mathrm{x}} = \cos(\frac{\pi}{4})$ and $V_{\mathrm{y}} = \sin(\frac{\pi}{4})$ and we start with the following initial emissions 30%, 40% and 30% of u_{tot}. The optimization procedure provides the following optimal emissions, respectively for the three chimneys, the 30%, 38.8% and 7.3% of u_{tot}. The RB strategy, with $\tau = 800$ and the RB dimension $N = 81$, allows a 55% time saving with respect to the FE method, with a final error on J and u respectively about $3.9 \cdot 10^{-9}$ and $1.1 \cdot 10^{-6}$. The original grid is made of 1700 nodes.

7 Conclusions

We have proposed two strategies to improve the efficiency of the numerical resolution of optimal control problems governed by linear advection–diffusion equa-

tions, in the context of an iterative optimization procedure. In particular, having identified the error on the control problem as error on the cost functional, we have separated it into the iteration and discretization errors. For the latter we have proposed an a posteriori error estimate, which is adopted in a strategy of grid adaptivity. Then we have considered the RB method, applied to the equations governing the control problem itself, in order to save computational costs by adopting a reliable method. The efficiency of these approaches is proved by numerical tests, that are concerned with a pollution control problem in atmosphere.

References

[1] V. Agoshkov. *Optimal Control Methods and Adjoint Equations in Mathematical Physics Problems.* Institute of Numerical Mathematics, Russian Academy of Science, 2003.

[2] R. Becker. Mesh adaptation for stationary flow control. *J. Math. Fluid Mech.* 3:317-341, 2001.

[3] R. Becker, H. Kapp, R. Rannacher. Adaptive finite element methods for optimal control of partial differential equations: basic concepts. *SIAM J. Control Optim.* 39:113-132, 2000.

[4] S.S. Collis, M. Heinkenschloss. Analysis of the Streamline Upwind/Petrov Galerkin method applied to the solution of optimal control problems. *CAAM report* TR02-01, `http://www.caam.rice.edu`

[5] L. Dedè, A. Quarteroni. Optimal control and numerical adaptivity for advection–diffusion equations. *MOX report* 55.2005, `http://mox.polimi.it`, *M^2AN,* to appear, 2005.

[6] J.L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations.* Springer–Verlag, New York, 1971.

[7] C. Prud'homme, D. Rovas, K. Veroy, Y. Maday, A.T. Patera, G. Turinici. Reliable real–time solution of parametrized partial differential equations: reduced–basis output bound methods. *J. Fluids Eng.* 124:70-80, 2002.

[8] A. Quaini, G. Rozza. Reduced basis methods for advection–diffusion optimal control problems. In progress. 2005.

[9] A. Quarteroni, A. Valli. *Numerical Approximation of Partial Differential Equations.* Springer, Berlin and Heidelberg, 1994.

[10] G. Rozza. Optimal flow control and reduced basis techniques in shape design with applications in haemodynamics. PhD Thesis, EPFL, 2005.