Adaptive Methods for Control Problems with Finite-Dimensional Control Space

Saheed Akindeinde, Daniel Wachsmuth

Adaptive methods for control problems with finite-dimensional control space *

Saheed Akindeinde and Daniel Wachsmuth

Johann Radon Institute for Computational and Applied Mathematics (RICAM)
Austrian Academy of Sciences
Altenbergerstraße 69
A-4040 Linz, Austria

Abstract. We investigate adaptive methods for optimal control problems with finitely many control parameters. We analyze a-posteriori error estimates based on verification of second-order sufficient optimality conditions. Reliability and efficiency of the error estimator is shown. The estimator is used in numerical tests to guide adaptive mesh refinement.

Keywords: optimal control, numerical approximation, a-posteriori error estimates, adaptive refinement

1 Introduction

We study optimal control problems of the following type: Minimize the functional J given by

$$J(y, u) = g(y) + j(u) \tag{P}$$

over all $(y, u) \in Y \times U$ that satisfy the non-linear elliptic partial differential equation

$$E(y, u) = 0$$

and the control constraints

$$u \in U_{ad}.$$

Here, Y is a real Banach space, $U = \mathbb{R}^n$. The set $U_{ad} \subset U$ is a non-empty, convex and closed set given by $U_{ad} = \{u \in U : \ u_a \le u \le u_b\}$, where the inequalities are to be understood component-wise. Here, the cases $u_a = -\infty$ and $u_b = +\infty$ are allowed, such that problems with one-sided constraints or without control constraints are included in the analysis as well. Examples that are covered by this framework include parameter identification and optimization problems with finitely many parameters, see for instance our previous work [1].

Adaptive mesh refinement remains a valuable tool in scientific computation. The main objective of an adaptive procedure is to find a discrete solution to a problem while maintaining as few as possible numbers of unknowns with respect to a desired error estimate. As the solution and hence the error distributions on the mesh are unknown a-priori, one has to rely on a-posteriori error estimates.

* This work was funded by Austrian Science Fund (FWF) grant P21564-N18.

A-posteriori error estimates for non-linear control and identification problems can be found for instance in [2, 5, 6, 9]. However, they depend on two crucial *a-priori* assumptions: the first is that a second-order sufficient condition (SSC) has to hold at the solution of the continuous problem. With this assumption, error estimates of the type $\|\bar{u} - u_h\|_U \leq c\eta + \mathcal{R}$ can be derived, where η is a computable error indicator and $\mathcal{R}$ is a second-order remainder term. Here, the second a-priori assumption comes into play: one has to assume that $\mathcal{R}$ is small enough, in order to guarantee that mesh refinement solely based on η is meaningful. A different approach with respect to mesh refinement was followed in [13]. There the residuals in the first-order necessary optimality condition were used to derive an adaptive procedure. However, smallness of residuals does not imply smallness of errors without any further assumption. Here again, SSC as well as smallness of remainder terms is essential to draw this conclusion.

In our previous work [1], we applied a different strategy: There the sufficient optimality condition as well as smallness of remainders is checked *a-posteriori.* If both conditions are fulfilled, an error-estimator of the form

$$\|u - u_h\|_U \leq \frac{2}{\alpha}(\omega_y r_y + \omega_p r_p)$$

is available, see [1, Thm 3.22]. This error estimator is localizable if r_y and r_p are localizable error estimates for the norm of the residual in the state and adjoint equations, respectively. For earlier and related work on a special problem calls with infinite-dimensional control space, we refer to [7, 8].

In this article, we will prove a lower bound of the error estimator. For the setting $Y = H_0^1(\Omega)$, we obtain

$$\begin{aligned} r_y + r_p \leq c(\|u - u_h\|_U + \|y - y_h\|_Y + \|\nabla y - \sigma_h\|_{L^2(\Omega)} \\ + \|p - p_h\|_Y + \|\nabla p - \tau_h\|_{L^2(\Omega)} + \tilde{o}), \end{aligned}$$

where y, p and y_h, p_h are solutions of continuous and discrete state and adjoint equations, respectively, and σ_h and τ_h are approximations of ∇y_h and ∇p_h in $H(\mathrm{div})$. The term $\tilde{o}$ is a higher-order oscillation term. In addition, we have localized lower bounds for the residuals in the state and adjoint equations, respectively. This justifies the use of the a-posteriori estimator above in an adaptive mesh-refinement procedure.

1.1 The abstract framework

Let Ω be a polygonal domain in $\mathbb{R}^m$, $m = 2, 3$. The function space for the states of the optimal control problem is chosen as $Y := H_0^1(\Omega)$. Let us now specify assumptions on the abstract problem (P).

Assumption 1 *The mapping $E : Y \times U \to Y^*$, $g : Y \to \mathbb{R}$, and $j : U \to \mathbb{R}$ are twice continuously Fréchet-differentiable with locally Lipschitz continuous second-derivatives. Furthermore, we assume that the mapping E is strongly monotone with respect to the first variable at all points $u \in U_{ad}$.*

The assumptions on E are met for instance for semilinear elliptic equations with monotone nonlinearities. Under Assumption (1), the state equation $E(y,u)=0$ is uniquely solvable for each admissible control $u\in U_{ad}$, [12, Theorem 26.A, p. 557]. We remark that the differentiability assumption on E can be relaxed to accomodate a more general class of problems, see [1, Remark 1.2], e.g. differentiability of E from $(Y\cap L^\infty(\Omega))\times U$ to Y^* is sufficient.

Let us define the Lagrange functional for the abstract problem:

$$\mathcal{L}(u,y,p):=g(y)+j(u)-\langle E(y,u),\,p\rangle_{Y^*,Y}.$$

Let $(\bar y,\bar u)$ be locally optimal for (P). Then the first-order necessary optimality conditions can be expressed as $\mathcal{L}'_y(\bar y,\bar u,\bar p)=0$ and $\mathcal{L}_u(\bar y,\bar u,\bar p)(u-\bar u)\geq 0$ for all $u\in U_{ad}$, which is equivalent to

$$\begin{aligned} E_y(\bar u,\bar y)^*\bar p &= g'(\bar y)\\ \langle j'(\bar u)-E_u(\bar u,\bar y)^*\bar p,\,u-\bar u\rangle_{U^*,U} &\geq 0 \quad \forall u\in U_{ad}. \end{aligned}$$

Since the problem (P) is in general non-convex, the fulfillment of these necessary conditions does not imply optimality. In order to guarantee this, one needs additional sufficient optimality conditions of the type: There exists $\alpha>0$ such that

$$\mathcal{L}''(\bar u,\bar y,\bar p)[(z,v)^2]\geq\alpha\|v\|_U^2 \tag{1}$$

holds for all $v=u-\bar u$, $u\in U_{ad}$, and z solves the linearized equation $E_y(\bar u,\bar y)z+E_u(\bar u,\bar y)v=0$. This condition can be weakened taking strongly active inequality constraints into account, see e.g. [1, 3]. For simplicity, we chose to work with this stronger condition. The results of this article hold also under the weakened sufficient condition.

Although, the sufficient condition is of high interest, it is difficult to check numerically even when $(\bar u,\bar y,\bar p)$ are given, see e.g. [1, 7, 8]. The main difficulty here is that the function z appearing in (1) is given as solution of a partial differential equation, which cannot be solved explicitly. Any discretization of this equation introduces another error that has to be analyzed.

1.2 Discretization

In order to solve (P) numerically, we discretize the problem. Let Y_h be a finite-dimensional subspace of Y. Here and in the following, the index h denotes a discrete quantity. Then a discretization of the state equation can be obtained in the following way: A function $y_h\in Y_h$ is a solution of the discretized equation for given $u\in U_{ad}$ if and only if

$$\langle E(y_h,u),\,\phi_h\rangle_{Y^*,Y}=0 \qquad \forall\phi_h\in Y_h. \tag{2}$$

The discrete optimization problem is then given by: Minimize the functional $J(y_h,u_h)$ over all $(y_h,u_h)\in Y_h\times U_{ad}$, where y_h solves the discrete equation.

Let $(\bar{y}_h, \bar{u}_h)$ be a local solution of the discrete problem. Then it fulfills the discrete first-order necessary optimality condition, which is given as: there exists a uniquely determined discrete adjoint state $\bar{p}_h \in Y_h$ such that it holds

$$\begin{aligned} \langle E_y(\bar{y}_h, \bar{u}_h)^* \bar{p}_h,\, \phi_h \rangle_{Y^*,Y} = \langle g'(\bar{y}_h),\, \phi_h \rangle_{Y^*,Y} \quad \forall \phi_h \in Y_h \\ \langle j'(\bar{u}_h) - E_u(\bar{y}_h, \bar{u}_h)^* \bar{p}_h, u - \bar{u}_h \rangle_{U^*,U} \geq 0 \ \forall u \in U_{ad}. \end{aligned} \tag{3}$$

Throughout this work, we will assume that errors in discretizing the optimality system are controllable in the following sense. We will not make any further assumptions on the discretization, in particular, we do not assume a sufficient fine discretization.

Assumption 2 *For a fixed finite-dimensional subspace Y_h, let (u_h, y_h, p_h) be approximations of the discrete optimal control and the corresponding state and adjoint. There are positive constants r_y, r_p such that the following holds*

$$\|E(y_h, u_h)\|_{Y^*} \leq r_y, \tag{4}$$

$$\|g'(y_h) - E_y(y_h, u_h)^* p_h\|_{Y^*} \leq r_p, \tag{5}$$

$$\langle j'(u_h) - E_u(y_h, u_h)^* p_h, u - u_h \rangle_{U^*,U} \geq 0 \ \forall u \in U_{ad}. \tag{6}$$

Here, r_y and r_p are dual norms of residuals in the state and adjoint equation, respectively, and hence reflect the discretization error. We report on the computation of these residuals in Section 2.

As already mentioned, without any further assumption, smallness of the residuals in (4)–(6) does not imply smallness of the error $\|u - u_h\|_U$ in the control. In order to establish such a bound, it is essential to check that a second-order sufficient optimality condition is satisfied.

Here it is important to recognize that sufficient optimality conditions for the *discrete* problem alone are still not enough. The sufficient optimality condition for the discrete problem is given by: There exists $\alpha_h > 0$ such that

$$\mathcal{L}''(\bar{u}_h, \bar{y}_h, \bar{p}_h)[(z_h, v)^2] \geq \alpha_h \|v\|_U^2 \tag{7}$$

holds for all $v = u - \bar{u}$, $u \in U_{ad}$, and z_h solves the linearized discrete equation

$$\langle E_y(\bar{u}_h, \bar{y}_h) z_h + E_u(\bar{u}_h, \bar{y}_h) v,\ \phi_h \rangle_{Y^*,Y} = 0 \quad \forall \phi_h \in Y_h. \tag{8}$$

This condition is equivalent to positive definiteness of a certain computable matrix, see [1, Section 3.5]. Moreover we have the following estimate relating the coercivity constants α and α_h appearing in (1) and (7):

$$\alpha \geq \alpha_h - \|\mathcal{E}\|_2,$$

where $\|\mathcal{E}\|_2$ is the norm of an error matrix taking the discretization error in the linearized equation $E_y(\bar{u}, \bar{y})z + E_u(\bar{u}, \bar{y})v = 0$ into account, see [1, Section 3.5] for the details. If the computable lower bound $\alpha_h - \|\mathcal{E}\|_2$ of α is positive, then it follows that (1) is satisfied. Moreover, we have the following result:

Theorem 1 (Upper bound of the error). *Let Assumptions 1 and 2 be satisfied. Let (y_h, u_h, p_h) be a solution of the discrete optimal control problem. If $\alpha_h - \|\mathcal{E}\|_2 > 0$ holds and the residuals r_y and r_p are small enough, then there exists a local solution $\bar{u}$ of* (P) *that satisfies the error bound*

$$\|\bar{u} - u_h\|_U \leq \frac{2}{\alpha_h - \|\mathcal{E}\|_2} \left(\omega_y\, r_y + \omega_p\, r_p\right), \tag{9}$$

where the weights ω_y, ω_p depend on the discrete solution (y_h, u_h, p_h). If for different discretizations the discrete solutions $\{(y_h, u_h, p_h)\}_{h>0}$ are uniformly bounded in $Y \times U \times Y$ then the weights ω_y, ω_p are bounded as well.

For the proof, we refer to [1, Thm. 3.22]. There, precise estimates of the weights ω_y, ω_p are given. Moreover, a quantification of the smallness assumption on r_y and r_p is given, which makes this assumption verifiable *a-posteriori.*

Corollary 1. *Let the assumptions of Theorem 1 be satisfied. Let $\bar{y}$, $\bar{p}$ denote the solutions of the state and adjoint equations to $\bar{u}$, respectively. Then it holds*

$$\begin{aligned}
\|\bar{y} - y_h\|_Y &\leq \upsilon_{yu}\|\bar{u} - u_h\|_U + \delta^{-1} r_y, \\
\|\bar{p} - p_h\|_Y &\leq \upsilon_{pu}\|\bar{u} - u_h\|_U + \delta^{-1} r_p + \upsilon_{py} r_y,
\end{aligned}$$

with δ^{-1} being the global bound of $\|E_y^{-1}(y,u)\|_{\mathcal{L}(Y^*,Y)}$, and weights υ_{yu}, υ_{pu}, and υ_{py} depending on (y_h, u_h, p_h) in the same way as the weights ω_y, ω_p in Theorem 1.

Proof. The result is a consequence of Theorem 1 and [1, Lemma 3.1, 3.3].

2 Lower error bounds

In this section, we assume that the general operator E can be written as $E(y, u) = -\Delta y + d(y, u)$ with d being a superposition operator induced by a smooth function $d : \mathbb{R}^2 \to \mathbb{R}$. We remark that the subsequent analysis can be easily extended to operators in divergence form with bounded coefficients possibly depending on u [1]. We will work with a classical finite-element discretization: The discrete space Y_h is the classical space of piecewise quadratic and continuous elements (P2) on a given conforming triangulation $\mathcal{T}_h$ of Ω. The diameter of an element $T \in \mathcal{T}_h$ is denoted by h_T. We denote by $\Sigma_h \subset H(\mathrm{div})$, a conforming Raviart-Thomas (RT_1) discretization of the space $H(\mathrm{div})$.

Let us endow $Y = H_0^1(\Omega)$ with the norm $\|y\|_Y^2 := \|\nabla y\|^2_{L^2(\Omega)} + \|y\|^2_{L^2(\Omega)}$. In the sequel, let us denote the norm of the embedding $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ by I_2.

Now let us report on the computation of the residual r_y in the state equation. As required by Assumption 2, we are interested in constant-free error estimates, i.e. all constants appearing in the a-posteriori error estimate must be computable. Here, we apply the results of Vohralík [11].

Theorem 2. *Let $y_h \in Y_h \subset H_0^1(\Omega)$, $u_h \in U_{ad}$ satisfy the discrete equation* (2)*. Let $\sigma_h \in \Sigma_h \subset H(\operatorname{div})$ be given such that*

$$(\operatorname{div}\sigma_h,\, 1)_{L^2(T)} = (d(y_h, u),\, 1)_{L^2(T)} \quad \text{for all cells } T \in \mathcal{T}_h. \tag{10}$$

Let us define the cell-wise indicator $\eta_{y,T}$, $T \in \mathcal{T}_h$,

$$\eta_{y,T} := 2\|\nabla y_h - \sigma_h\|_{L^2(T)} + \pi^{-1} h_T \|d(y_h, u_h) - \operatorname{div}\sigma_h\|_{L^2(T)} \tag{11}$$

Then it holds

$$\|-\Delta y_h + d(y_h, u_h)\|^2_{H^{-1}(\Omega)} \le (1 - I_2^2)^{-1} \sum_{K\in\mathcal{T}_h} \eta^2_{y,T} =: r_y^2. \tag{12}$$

If moreover, $\mathcal{T}_h$ is shape-regular, then it holds

$$\eta_{y,T} \le C\|\nabla(\tilde{y} - y_h)\|_{L^2(T)} + c\|\nabla\tilde{y} - \sigma_h\|_{L^2(T)} + ch_T\|d(y_h, u_h) - \Pi_h d(y_h, u_h)\|_{L^2(T)}, \tag{13}$$

where $\tilde{y} := \Delta^{-1} d(y_h, u_h)$ and Π_h denotes the orthogonal L^2-projection onto Y_h. The constants C, c depend only on the spatial dimension m and the shape regularity of the triangulation.

Proof. The upper bound (12) is a consequence of [11, Thm. 6.8, 6.12] taking [11, Remark 6.3] into account for σ_h satisfying $(\operatorname{div}\sigma_h,\, 1)_{L^2(T)} = (d(y_h, u),\, 1)_{L^2(T)}$, $T \in \mathcal{T}_h$. The lower bound (13) follows from [11, Thm. 6.16], see also [10, Lemma 7.6]. Since $d(y_h, u)$ is in general not in the discrete space Y_h, we obtain the additional oscillation term $h_T\|d(y_h, u) - \pi_h d(y_h, u)\|_{L^2(T)}$ by a standard argument.

Estimates of the residual in the adjoint equation can be obtained after obvious modifications: for $\tau_h \in \Sigma_h$ satisfying

$$(\operatorname{div}\tau_h,\, 1)_{L^2(T)} = (d'(y_h, u_h)p_h - g'(y_h),\, 1)_{L^2(T)} \quad \text{for all cells } T \in \mathcal{T}_h \tag{14}$$

and with the local error indicators defined by

$$\eta_{p,T} := 2\|\nabla p_h - \tau_h\|_{L^2(T)} + \pi^{-1} h_T \|d'(y_h, u_h)p_h - g'(y_h) - \operatorname{div}\tau_h\|_{L^2(T)} \tag{15}$$

we obtain the upper bound

$$\|-\Delta p_h + d'(y_h, u_h)p_h - g'(y_h)\|^2_{H^{-1}(\Omega)} \le (1 - I_2^2)^{-1} \sum_{K\in\mathcal{T}_h} \eta^2_{p,T} =: r_p^2. \tag{16}$$

as well as the lower bound

$$\begin{aligned}\eta_{p,T} \le C\|\nabla(\tilde{p} - p_h)\|_{L^2(T)} &+ c\|\nabla\tilde{p} - \tau_h\|_{L^2(T)} \\ &+ ch_T\|(I - \Pi_h)(d'(y_h, u_h)p_h - g'(y_h))\|_{L^2(T)},\end{aligned} \tag{17}$$

where $\tilde{p} := \Delta^{-1}(d'(y_h, u_h)p_h - g'(y_h))$.

We remark that the upper bounds (12) and (16) are constant-free, making them explicitly computable. In our computations, we computed the functions

σ_h and τ_h as a minimizer of the right-hand side in (12) and (16), respectively, using Raviart-Thomas elements for discretization of $H(\mathrm{div})$. This shows that the requirements of Assumption 2 on the computability of upper bounds on the residuals can be fulfilled.

Now let us argue that under the assumptions of Theorem 1 we also obtain lower bounds for the error, which proves efficiency of the error bound.

Theorem 3. *Let the assumptions of Theorem 1 be fulfilled. Let r_y and r_p be computed according to* (12) *and* (16)*. Let $(\bar y, \bar u, \bar p)$ be the local solution of* (P) *provided by Theorem 1. Then it holds*

$$\sum_{T\in\mathcal{T}_h} r_{y,T} \le C\Big(\|\bar u - u_h\|_U + \|\bar y - y_h\|_Y + \|\nabla\bar y - \sigma_h\|_{L^2(\Omega)} + \|h_T(I-\Pi_h)d(y_h,u_h)\|_{L^2(\Omega)}\Big), \quad (18)$$

$$\sum_{T\in\mathcal{T}_h} r_{p,T} \le C\Big(\|\bar u - u_h\|_U + \|\bar y - y_h\|_Y + \|\bar p - p_h\|_Y + \|\nabla\bar p - \tau_h\|_{L^2(\Omega)} + \|h_T(I-\Pi_h)(d'(y_h,u_h)p_h - g'(y_h))\|_{L^2(\Omega)}\Big), \quad (19)$$

where $C>0$ depends only on the spatial dimension m, the shape regularity of the triangulation, and global bounds of derivatives d_y, d_u, d_{yy}, and d_{yu} of $d : Y\times U \to Y^*$ near (y_h,u_h).

Proof. Let $\tilde y$ be given by $\tilde y := \Delta^{-1} d(y_h,u_h)$. Then we can estimate

$$\begin{aligned}\|\nabla(\tilde y - y_h)\|_{L^2(\Omega)} &+ \|\nabla\tilde y - \sigma_h\|_{L^2(\Omega)}\\ &\le 2\|\nabla(\tilde y - \bar y)\|_{L^2(\Omega)} + \|\nabla(\bar y - y_h)\|_{L^2(\Omega)} + \|\nabla\bar y - \sigma_h\|_{L^2(\Omega)}.\end{aligned}$$

Using Lipschitz continuity of d, we find

$$\|\nabla(\tilde y - \bar y)\|_{L^2(\Omega)} \le C(\|\bar u - u_h\| + \|\bar y - y_h\|_Y),$$

with C depending on bounds of $\|d'\|_{\mathcal{L}(Y\times U, Y^*)}$ near (y_h,u_h). This together with (13) proves (18). The estimate (19) can be obtained analogously.

These lower bounds together with (9) and the local lower bounds (13) and (17) justify the use of the error indicators in an adaptive mesh-refinement procedure.

Remark 1. Another possibility of constant-free a-posteriori error estimators based on $H(\mathrm{div})$-functions is described in [4]. There, fluxes across edges in a dual mesh are prescribed instead of the integrals on elements as in (10) and (14). In [4] it is proven that the resulting error estimate is reliable and efficient. Moreover, the terms $\|\nabla y_h - \sigma_h\|_{L^2(\Omega)}$ and $\|\nabla p_h - \tau_h\|_{L^2(\Omega)}$ do not appear in the lower error bound when compared to (18) and (19), respectively.

3 Adaptivity

We will compare the performance of adaptive mesh refinement using different strategies to mark elements for refinement. The first one, referred to as *'verified adaptive'*, is implemented as follows: in each step the verification procedure of [1] is carried out. If it confirms that the assumptions of Theorem 1 are satisfied, then the error indicator $\omega_y r_y + \omega_p r_p$ given by (9) is used to guide the mesh-refinement. If the requirements of Theorem 1 cannot be verified, then a uniform refinement step is carried out. Here, we expect that after a small number of uniform refinement steps the requirements of Theorem 1 are confirmed a-posteriori, which coincides with the numerical experiments done in earlier work [1]. After these initial uniform refinements steps, we expect that the method proceeds with adaptive steps.

A second strategy, called *'fully adaptive'*, omits the verification step, and simply uses $\omega_y r_y + \omega_p r_p$ from (9) without checking the validity of this bound.

4 Numerical results

Let us report about the outcome of the above described adaptive methods for a selected example, taken from [1]. The functional J was chosen as

$$J(y,u) := \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\kappa}{2}\|u\|^2_{\mathbb{R}^n}.$$

The nonlinear mapping E represents a semi-linear elliptic equation given by

$$E(y,u) := -\Delta y + \sum_{k=1}^{n} u_k d_k(y) - g, \tag{20}$$

where the functions d_k are chosen as $d_1(y) = 1$, $d_j(y) = y|y|^{j-2}$ for $j = 2 \ldots n$. This example is motivated by parameter identification: given a state y_d and source term g, find the set of coefficients u such that the resulting solution y of $E(y,u) = 0$ is as close as possible to y_d.

In order to make the operator E strongly monotone, we require positivity of the coefficients u_k, i.e. we set $U_{ad} = \{u \in \mathbb{R}^n : \ u_k \geq 0 \quad \forall k = 1 \ldots n\}$. For the computations we used the following data: the source term $g = 10.0001$ and

$$\Omega = (0,1)^2,\ u_a = 0,\ u_b = 0.5,\ \kappa = 10^{-2},\ y_d(x_1,x_2) = 0.5\sin(2\pi x_1 x_2),\ n = 4.$$

Let us remark, that the function d_3 is not of class C^2 globally. Since g is non-negative, every solution y of (20) to $u \in U_{ad}$ will be non-negative. For non-negative functions y it holds $d_3(y) = y^2$, which is C^2, so the assumptions on E are satisfied. See also the discussion in [1, Section 4.3].

We employed a discretization scheme as described in Section 2. After the resulting non-linear optimization problem is solved, the error indicators according to the chosen strategy are computed. For an adaptive refinement, a subset

$\tilde{\mathcal{T}} \subset \mathcal{T}$ of elements T with large local error contributions η_T were selected for refinement that satisfies $\sum_{T\in\tilde{\mathcal{T}}} \eta_T^2 \geq \theta^2 \sum_{T\in\mathcal{T}} \eta_T^2$ with $\theta = 0.8$.

Let us report on the outcome of the different adaptive strategies as described in Section 3. For all the methods, we compare the residual norms as given in (11), i.e. with the notation of that section

$$\epsilon_{\text{residual}} := \omega_y r_y + \omega_p r_p.$$

Moreover, we employed the verification procedure of Theorem 1, see e.g. [1], and report about the upper error bound

$$\epsilon_{\text{bound}} := \frac{2}{\alpha_h - \|\mathcal{E}\|_2} \left(\omega_y \, r_y + \omega_p \, r_p\right).$$

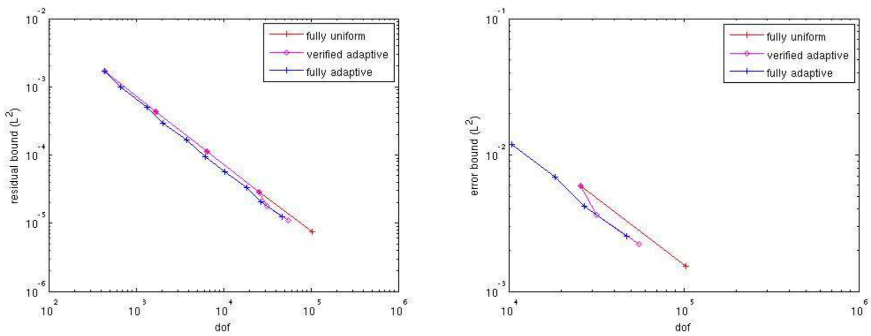

Fig. 1. (a) upper bound of residuals versus number of unknowns, (b) verified error bound versus number of unknowns

As can be expected, the assumptions of Theorem 1 are only fulfilled on a sufficiently fine discretization. This is reflected by our numerical results.

Plots of $\epsilon_{\text{residual}}$ and ϵ_{bound} versus the number of degrees of freedom can be seen in Figure 4. For reference, we provided the numerical values in Tables 1 and 2. In the tables, L refers to refinement level, where $L = 0$ is the initial mesh, which is the same for all the different adaptive methods. Moreover, dof denotes the number of degrees of freedom.

Let us comment on the observed behavior of the verified adaptive methods. The conditions of Theorem 1 are fulfilled for the first time after three uniform refinement steps. The fourth and all further refinement levels were reached by using adaptive refinement according to the error indicator based on (11). The fully adaptive scheme, which refines according to the residuals in the optimality system, obtains verified error bounds as of level 7. After the verified adaptive methods actually start adaptive refinement, they quickly reach the same ratio of error bound versus number of degrees of freedom as the full adaptive method. That means, the early (unverified) adaptive refinements of the full adaptive methods does not seem to give this method an advantage over the verified method. The same observation also applies to the residual error versus number of degrees of freedom ratio, as can be seen in Figure 4.

fully uniform			*verified adaptive*			*fully adaptive*		
L	# *dof*	ϵ_{error}	L	# *dof*	ϵ_{error}	L	# *dof*	ϵ_{error}
0	441	–	1	1681	–	5	6177	–
1	1681	–	2	6561	–	6	10341	–
2	6561	–	3	25921	$7.6226 \cdot 10^{-3}$	7	18427	$9.0724 \cdot 10^{-3}$
3	25921	$7.6226 \cdot 10^{-3}$	4	31491	$4.8745 \cdot 10^{-3}$	8	27155	$5.6376 \cdot 10^{-3}$
4	103041	$1.9183 \cdot 10^{-3}$	5	55061	$2.8728 \cdot 10^{-3}$	9	46979	$3.3167 \cdot 10^{-3}$

Table 1. Error bound estimates

fully uniform			*verified adaptive*			*fully adaptive*		
L	#*dof*	$\epsilon_{\text{residual}}$	L	#*dof*	$\epsilon_{\text{residual}}$	L	#*dof*	$\epsilon_{\text{residual}}$
1	1681	$5.4976 \cdot 10^{-4}$	1	1681	$5.4976 \cdot 10^{-4}$	2	1311	$6.6336 \cdot 10^{-4}$
2	6561	$1.4181 \cdot 10^{-4}$	2	6561	$1.4181 \cdot 10^{-4}$	5	6193	$1.2537 \cdot 10^{-4}$
3	25921	$3.6666 \cdot 10^{-5}$	3	25921	$3.6669 \cdot 10^{-5}$	7	18427	$4.3361 \cdot 10^{-5}$
			4	31491	$2.3746 \cdot 10^{-5}$	8	27155	$2.7370 \cdot 10^{-5}$
4	103041	$9.4714 \cdot 10^{-6}$	5	55061	$1.4121 \cdot 10^{-5}$	9	46979	$1.6271 \cdot 10^{-5}$

Table 2. Residual error bound estimates

References

1. S. Akindeinde and D. Wachsmuth. A-posteriori verification of optimality conditions for control problems with finite-dimensional control space. *Num. Funct. Anal. Opt.*, 33(5):473–523, 2012.
2. R. Becker. Estimating the control error in discretized pde-constrained optimization. *Journal of Numerical Mathematics*, 14:163–185, 2006.
3. E. Casas, F. Tröltzsch, and A. Unger. Second-order sufficient optimality conditions for a nonlinear elliptic control problem. *J. for Analysis and its Applications*, 15:687–707, 1996.
4. L. El Alaoui, A. Ern, and M. Vohralík. Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems. *Comput. Methods Appl. Mech. Engrg.*, 200(37-40):2782–2795, 2011.
5. A. Griesbaum, B. Kaltenbacher, and B. Vexler. Efficient computation of the Tikhonov regularization parameter by goal-oriented adaptive discretization. *Inverse Problems*, 24(2):025025, 20, 2008.

6. K. Kunisch, W. Liu, Y. Chang, N. Yan, and R. Li. Adaptive finite element approximation for a class of parameter estimation problems. *J. Comput. Math.*, 28(5):645–675, 2010.
7. A. Rösch and D. Wachsmuth. Numerical verification of optimality conditions. *SIAM J. Control Optim.*, 47(5):2557–2581, 2008.
8. A. Rösch and D. Wachsmuth. How to check numerically the sufficient optimality conditions for infinite-dimensional optimization problems. In *Control of Coupled Partial Differential Equations*, volume 158 of *Internat. Ser. Numer. Math.*, pages 297–317. Birkhäuser, Basel, 2009.
9. B. Vexler and W. Wollner. Adaptive finite elements for elliptic optimization problems with control constraints. *SIAM J. Control Optim.*, 47(1):509–534, 2008.
10. M. Vohralík. A posteriori error estimates for lowest-order mixed finite element discretizations of convection-diffusion-reaction equations. *SIAM J. Numer. Anal.*, 45(4):1570–1599, 2007.
11. M. Vohralík. Unified primal formulation-based a priori and a posteriori error analysis of mixed finite element methods. *Math. Comp.*, 79(272):2001–2032, 2010.
12. E. Zeidler. *Nonlinear functional analysis and its applications. II/B.* Springer-Verlag, New York, 1990. Nonlinear monotone operators.
13. J. C. Ziems and S. Ulbrich. Adaptive multilevel inexact SQP methods for PDE-constrained optimization. *SIAM J. Optim.*, 21(1):1–40, 2011.