
Computational Optimization and Applications, 31, 309–333, 2005

A Multigrid Scheme for Elliptic Constrained Optimal Control Problems*

A. BORZÌ alfio.borzi@uni-graz.at
K. KUNISCH karl.kunisch@uni-graz.at
Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstr. 36, A-8010 Graz, Austria

Received April 15, 2003; Revised February 20, 2004; Accepted August 18, 2004

Abstract. A multigrid scheme for the solution of constrained optimal control problems discretized by finite differences is presented. This scheme is based on a new relaxation procedure that satisfies the given constraints pointwise on the computational grid. In applications, the cases of distributed and boundary control problems with box constraints are considered. The efficient and robust computational performance of the present multigrid scheme allows to investigate bang-bang control problems.

Keywords: constrained optimal control problems, finite differences, multigrid methods

AMS Subject Classification: 49J20, 65N06, 65N12, 65N55

1. Introduction

There is large theoretical and experimental evidence that multigrid methods [13, 21] solve elliptic problems with optimal computational complexity. This is also true for multigrid methods applied to unconstrained optimal control problems [2–5, 9, 12, 20]. Furthermore, it was demonstrated that multigrid methods are robust with respect to changes of the weight of the cost.

In case of optimal control problems with constraints on the control much less is known regarding the implementation and the analysis of multigrid schemes. By reformulating the presence of constraints as a nonlinear equation relating the control to the adjoint variable, it is possible to solve constrained optimal control problems using nonlinear multigrid methods [5]. The resulting multigrid scheme shows typical multigrid efficiency for sufficiently large values of the weight of the cost of the control. For small values of the weight the convergence of the multigrid iterations deteriorates, showing a lack of robustness of this approach.

In this paper we propose a different technique where the constraints are enforced at each grid point in the smoothing procedure. This procedure appears to be robust with respect to changes of the value ν of the weight of the cost and, in particular, it allows the choice $\nu = 0$. The present approach applies to distributed as well as to boundary optimal control problems. In all cases, for moderate values of the weight, the resulting multigrid iteration

*Supported in part by the SFB 03 "Optimization and Control".

shows typical multigrid convergence factors which are mesh independent. For very small values of ν down to $\nu = 0$ new features can arise. In the case of bang-bang-type controls, multigrid efficiency is retained. If chattering phenomena take place, the computational performance of the multigrid scheme worsens. This typically arises in the case of local attainability of the objective function.

These facts make our multigrid algorithm a useful tool for investigating bang-bang type control phenomena for elliptic problems, which has not received much attention in the literature on optimal control problems apart from some non recent contribution as in [10, 11] where some theoretical and numerical results on the bang-bang principle for parabolic boundary control problems are given.

In the following section we introduce and analyze our model problem with distributed control. In particular, we consider the case where the control acts only in a part of the whole domain.

In Section 3, the accuracy of the finite difference approximations to the solution of the optimal control problem is considered. While finite element approximations are well investigated [1, 17–19], much less results are available for finite difference methods. Thus, before presenting our multigrid approach, we discuss convergence of finite differences deriving some sub-optimal estimates. On the other hand, results of numerical experiments demonstrate that finite differences provide optimal accuracy as it is proved in [19] in the context of finite elements.

In Section 4, the multigrid scheme is formulated in the framework of nonlinear multigrid methods. This approach is appropriate in order to implement the inequality constraint at all levels of the multigrid process. We introduce a new relaxation scheme obtained by solving at each grid point the optimality system under the constraints imposed on the control function. Numerical experiments follow to demonstrate the ability of our algorithm to solve constrained optimal control problems also in the limit case of bang-bang control.

In Section 6, the extension of our algorithm to solve boundary control problems is presented and validated with numerical experiments. A section of conclusion completes the exposition of our work.

2. Constrained optimal control problems

We consider the optimal control problem

$$\begin{cases} \min_{u \in U_{ad}} J(y, u), \\ -\Delta y = Bu + g \quad \text{in } \Omega, \\ y = 0 \quad \text{on } \partial\Omega, \end{cases} \tag{1}$$

where Ω is a open bounded set in $\mathbf{R}^2$, with boundary $\partial\Omega$ and $u \in U_{ad} \subset L^2(\Omega)$. We assume that Ω is convex or that $\partial\Omega$ is $C^{1,1}$ smooth. The cost functional J is of the tracking type and is given by

$$J(y, u) = \frac{1}{2}\|y - z\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|Bu\|^2_{L^2(\Omega)}, \tag{2}$$

where $g \in L^2(\Omega)$, $z \in L^2(\Omega)$ is the objective function, and $\nu \geq 0$ is the weight of the cost of the control. The set of admissible controls is the closed convex subset of $L^2(\Omega)$ given by

$$U_{ad} = \{u \in L^2(\omega) \mid \underline{u}(\mathbf{x}) \leq u(\mathbf{x}) \leq \bar{u}(\mathbf{x}) \text{ a.e. in } \omega \subset \Omega\}, \tag{3}$$

where $\underline{u}$ and $\bar{u}$ are elements of $L^\infty(\Omega)$ and ω is a subset of Ω. The extension operator $B : L^2(\omega) \to L^2(\Omega)$ is defined as follows

$$Bu = \begin{cases} u & \text{in } \omega, \\ 0 & \text{in } \Omega \setminus \omega. \end{cases}$$

Existence of a unique solution to (1) and its characterization are well known. For completeness we give a short derivation. Let $y(u)$ denote the solution of the equality constraints in (1) as a function of u. The mapping $u \to y(u)$ from $L^2(\Omega)$ to $H_0^1(\Omega) \cap H^2(\Omega)$ is affine and continuous. Let us denote its first derivative in the direction δu by $y'(\delta u)$. It is characterized as the solution to

$$\begin{aligned} -\Delta y'(\delta u) &= B\,\delta u \quad \text{in } \Omega, \\ y'(\delta u) &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{4}$$

The second derivative of $u \to y(u)$ is zero.

Define $\hat{J}(u) = J(y(u), u)$. The mapping $u \to \hat{J}(u)$ is twice Frechet differentiable and its second derivative is given by

$$\hat{J}''(u)(\delta u, \delta u) = ||y'(\delta u)||^2_{L^2(\Omega)} + \nu ||B\delta u||^2_{L^2(\Omega)}.$$

It follows that $u \to \hat{J}(u)$ is uniformly convex if $\nu > 0$. This implies existence of a unique solution u^* to (1). For $\nu = 0$ the y-coordinate of the solution is unique. Since $u \to y(u)$ is injective, the solution u^* is again unique. The solution to (1) is characterized by $\hat{J}'(u)(u^*, \delta u) \geq 0$ and consequently

$$\hat{J}'(u^*, v - u^*) = (y(u^*) - z, y'(v - u^*))_{L^2(\Omega)} + \nu(Bu^*, B(v - u^*))_{L^2(\Omega)} \geq 0,$$

for all $v \in U_{ad}$.

Introduce $p^* \in H^2(\Omega) \cap H_0^1(\Omega)$ as the unique solution to

$$\begin{aligned} -\Delta p^* &= -(y^* - z) \quad \text{in } \Omega, \\ p^* &= 0 \quad \text{on } \partial\Omega, \end{aligned} \tag{5}$$

where $y^* = y(u^*)$. Then by (4) and (5) we have

$$\begin{aligned} \hat{J}'(u^*, v - u^*) &= -(p^*, B(v - u^*))_{L^2(\Omega)} + \nu(Bu^*, B(v - u^*))_{L^2(\Omega)} \\ &\geq 0 \text{ for all } v \in U_{ad}, \end{aligned}$$

which constitutes the necessary and sufficient optimality condition for the given optimal control problem.

Summarizing, the solution to (1) is characterized by the following optimality system

$$\begin{aligned} -\Delta y &= Bu + g && \text{in } \Omega, \\ y &= 0 && \text{on } \partial\Omega, \\ -\Delta p &= -(y - z) && \text{in } \Omega, \\ p &= 0 && \text{on } \partial\Omega, \\ (\nu u - B^* p,\ v - u) &\geq 0 && \text{for all } v \in U_{ad}, \end{aligned} \tag{6}$$

where B^* denotes the adjoint of B.

Notice that the last equation in (6) giving the optimality condition is equivalent to (see [16, 17])

$$u = \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu} p(u)\right\}\right\} \text{ in } \omega, \quad \nu > 0. \tag{7}$$

For the analysis that follows we need to define the operator $Q : L^2(\omega) \to L^2(\omega)$ given by

$$Q(\cdot) = \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu} p(\cdot)\right\}\right\},$$

when $\nu > 0$. Equation (7) can thus be written as $u = Q(u)$. Further, we introduce the operator $K : L^2(\omega) \to L^2(\omega)$ by $K(u) = u - Q(u)$.

Recall that for $\phi \in H^1(\Omega)$ we have $\max\{0, \phi\} \in H^1(\Omega)$ as well. Hence from (7) we obtain the following regularity result.

Lemma 1. If $\nu > 0$ and $\underline{u}, \bar{u} \in H^1(\Omega)$ then $u \in H^1(\Omega)$.

This regularity result will be used for the analysis of finite difference approximations of the optimality system that we discuss in the next section.

We complete this section discussing the solution to (1) with $\nu = 0$. It is simple to argue that the unique solution u satisfies

$$\begin{aligned} -\Delta y &= Bu + g \quad \text{in } \Omega, \\ y &= 0 \quad \text{on } \partial\Omega, \\ -\Delta p &= -(y - z) \quad \text{in } \Omega, \\ p &= 0 \quad \text{on } \partial\Omega, \\ p &= \min\{0, p + u - \underline{u}\} + \max\{0, p + u - \bar{u}\} \quad \text{in } \Omega. \end{aligned} \tag{8}$$

We are interested in solutions for which the inactive set $\mathcal{I} = \{\mathbf{x} : \underline{u} < u < \bar{u}\}$ is small.

For the construction of test examples the following lemma is useful.

Lemma 2. If $z(\mathbf{x}) \neq y(\mathbf{x})$ a.e. in Ω, with y the optimal state, then the inactive set $\mathcal{I}$ contains no interior points.

Proof: Proceeding by contradiction assume that $\mathcal{J}$ is an open set in $\mathcal{I}$. Then $p = 0$ on $\mathcal{I}$ by the last equation in (8) and hence $z(\mathbf{x}) = y(\mathbf{x})$ on $\mathcal{J}$ which contradicts our assumption. □

To verify the non-attainability assumption of Lemma 2 the following result can be used [8].

Lemma 3. Let $f \in L^\infty(\Omega)$, $\phi \in L^\infty(\partial\Omega)$, and let w be any function in $C^2(\Omega)$ satisfying

$$\begin{aligned} -\Delta w &\geq |f|_\infty \quad \text{in } \Omega, \\ w &\geq |\phi|_\infty \quad \text{on } \partial\Omega. \end{aligned} \tag{9}$$

Then $|v(\mathbf{x})| \leq w(\mathbf{x})$ for all $\mathbf{x} \in \Omega$ and v satisfying $-\Delta v = f$ in Ω, $v = \phi$ on $\partial\Omega$.

Lemma 3 allows to construct tracking functions z which are not attainable by any $u \in U_{ad}$. For example, if Ω is the unit square, we set $u_m = \max\{|\underline{u}|, |\bar{u}|\}$ and define

$$w(x_1, x_2) = \frac{u_m}{4}[x_1(1 - x_1) + x_2(1 - x_2)].$$

Then if $|z(\mathbf{x})| > \max|w| = u_m/8$ a.e. in Ω, we have

$$|y(u)(\mathbf{x})| < |z(\mathbf{x})| \quad \text{for all } \mathbf{x} \in \Omega \quad \text{and} \quad u \in U_{ad}.$$

Consequently the inactive set $\mathcal{I}$ cannot contain interior points.

Sufficient conditions for the more restrictive requirement $meas(\mathcal{I}) = 0$, i.e. that the control is bang-bang, appear to be an interesting open problem.

The following example is defined based on the estimate given above. It is computed using finite difference discretization and the multigrid scheme discussed in later sections. Assume $\underline{u} = -8$ and $\bar{u} = 8$, then $u_m = 8$. Take

$$z(x_1, x_2) = 1.1 \operatorname{sign}\{\sin(4\pi x_1)\sin(4\pi x_2)\}.$$

Clearly $|z|_\infty > \max|w| = 1$. The numerical solution of (1) for this setting is depicted in figure 1.

3. Finite difference optimality system

In this section accuracy of finite difference approximations to (6) is investigated in the framework given in [14]. It is assumed throughout that $\nu > 0$.

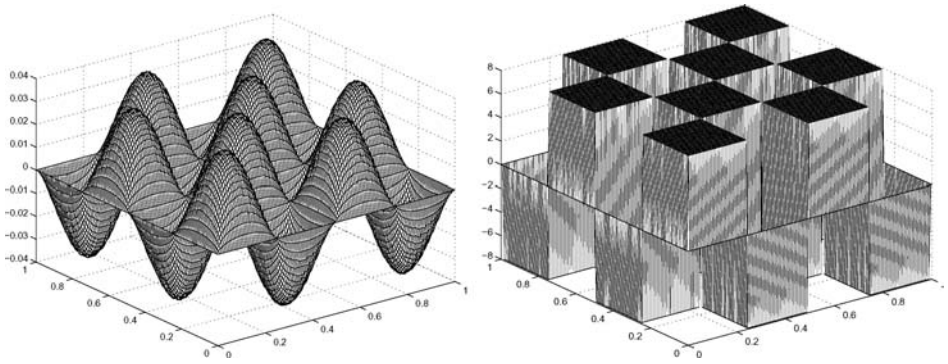

Figure 1. Numerical solution with $z = 1.1\, sign\{\sin(4\pi x_1)\sin(4\pi x_2)\}$ and $\nu = 0$. The state (left) and the control (right); 513×513 mesh.

Consider a sequence of grids $\{\Omega_h\}_{h>0}$ defined by

$$\Omega_h = \{\mathbf{x} \in \mathbf{R}^2 : x_i = s_i\, h, \quad s_i \in \mathbb{Z}\} \cap \Omega.$$

We assume that Ω is a rectangular domain and that the values of h are chosen such that the boundaries of Ω coincide with grid lines.

Define $\omega_h = \Omega_h \cap \omega$ and introduce a binary grid operator which is the discrete analog of the operator B defined above:

$$B_h u_h = \begin{cases} u_h & \text{in } \omega_h, \\ 0 & \text{in } \Omega_h \backslash \omega_h. \end{cases}$$

The negative Laplacian with homogeneous Dirichlet boundary conditions is approximated by the five-point stencil and is denoted by $-\Delta_h$.

For grid functions v_h and w_h defined on Ω_h we introduce the discrete L^2-scalar product

$$(v_h, w_h)_{L_h^2} = h^2 \sum_{\mathbf{x} \in \Omega_h} v_h(\mathbf{x})\, w_h(\mathbf{x}),$$

with associated norm $|v_h|_0 = (v_h, v_h)_{L_h^2}^{1/2}$. We require as well the discrete H^1-product given by

$$|v_h|_1 = \left(|v_h|_0^2 + \sum_{i=1}^{2} |\partial_i^- v_h|_0^2 \right)^{1/2},$$

where ∂_i^- denotes the backward difference quotient in the x_i direction and v_h is extended by 0 on grid points outside of Ω; see [14]. The spaces L_h^2 and H_h^1 consist of the sets of grid

functions v_h endowed with $|v_h|_0$, respectively $|v_h|_1$, as norm. For the definition of H_h^2 we refer to [14], as well.

Functions in $L^2(\Omega)$, $H^1(\Omega)$, and $H^2(\Omega)$ are approximated by grid functions defined through their mean values with respect to elementary cells $[x_1-\frac{h}{2}, x_1+\frac{h}{2}]\times[x_2-\frac{h}{2}, x_2+\frac{h}{2}]$. This gives rise to the following restriction operators; see [14] p. 232 for more details:

The restriction $R_h : H^2(\Omega)\cap H_0^1(\Omega) \to H_h^2$ where

$$(R_h v)(x, y) = \frac{1}{h^2}\int_{-h/2}^{h/2}\int_{-h/2}^{h/2} v(x+\xi, y+\eta)\, d\xi\, d\eta.$$

In the following, this operator is also used as mapping $R_h : H^1 \to H_h^1$.

For L^2 functions, we have the restriction operator $\tilde{R}_h : L^2(\Omega) \to L_h^2$ where

$$(\tilde{R}_h v)(x, y) = \frac{1}{h^2}\int_{-h/2}^{h/2}\int_{-h/2}^{h/2}\int_{-h/2}^{h/2}\int_{-h/2}^{h/2} v(x+\xi+\xi', y+\eta+\eta')\, d\xi\, d\xi'\, d\eta\, d\eta'.$$

The following properties can be proved using the Newton-Leibnitz formula.

$$|\tilde{R}_h v - R_h v|_0 \le ch\,|v|_{H^1(\Omega)} \quad \text{for all } v \in H^1(\Omega). \tag{10}$$

$$|\tilde{R}_h v - R_h v|_0 \le ch^2\,|v|_{H^2(\Omega)} \quad \text{for all } v \in H^2(\Omega). \tag{11}$$

Here and below, c denotes a positive constant which does not depend on h.

Now consider the discrete optimal control problem

$$\begin{cases} \min \dfrac{1}{2}|y_h - \tilde{R}_h z|_0^2 + \dfrac{\nu}{2}|u_h|_0^2, \\ -\Delta_h y_h = B_h u_h + \tilde{R}_h g, \end{cases} \tag{12}$$

where $u_h \in U_{adh} = U_{ad} \cap L_h^2$.

Let u_h^* denote the unique solution to (12) and set $y_h^* = y_h(u_h^*)$. The optimality system related to (12) is found to be

$$\begin{aligned} -\Delta_h y_h^* &= B_h u_h^* + \tilde{R}_h g, \\ -\Delta_h p_h^* &= -(y_h^* - \tilde{R}_h z), \\ (\nu u_h^* - B_h^* p_h^*)\cdot(v_h - u_h^*) &\ge 0 \quad \text{for all } v_h \in U_{adh}. \end{aligned} \tag{13}$$

The optimality condition can be reformulated as

$$u_h^* - \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}p_h(u_h^*)\right\}\right\} = 0 \text{ in } \omega_h, \tag{14}$$

which is equivalent to $u_h^* - Q_h(u_h^*) = 0$ where

$$Q_h(\cdot) = \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p_h(\cdot)\right\}\right\}.$$

The discrete counterpart of K is the operator $K_h(u_h) = u_h - Q_h(u_h)$.

To investigate the accuracy of the solution to (13) we use the approach of [17] extended to the present finite difference framework. In the remainder of this section we require $\underline{u},\ \bar{u} \in C^{0,1}(\Omega)$.

From Lemma 1.1 of [17] we have that

$$|u_h^* - \tilde{R}_h u^*|_0 \le \gamma\, |K_h(u_h^*) - K_h(\tilde{R}_h u^*)|_0,$$

where γ is a positive constant independent of the mesh size. The proof in [17] uses a finite element setting which can readily be adapted to the present finite difference framework. Due to the fact that $|u_h^*|_0$, $|\tilde{R}_h u^*|_0$, $|y_h(u_h^*)|_0$ and $|y_h(\tilde{R}_h u^*)|_0$ are uniformly bounded with respect to h, conditions (A.4)–(A.6) in [17] are satisfied.

Since $K_h(u_h^*) = 0$ and $K(u^*) = 0$ it follows that $K_h(u_h^*) = R_h K(u^*)$. Consequently we obtain

$$\begin{aligned} |u_h^* - \tilde{R}_h u^*|_0 &\le \gamma\, |K_h(u_h^*) - K_h(\tilde{R}_h u^*)|_0 = \gamma\, |R_h K(u^*) - K_h(\tilde{R}_h u^*)|_0 \\ &\le \gamma\, (|\tilde{R}_h u^* - R_h u^*|_0 + |R_h Q(u^*) - Q_h(\tilde{R}_h u^*)|_0). \end{aligned}$$

Notice that the solution u_h^* does not appear in the right-hand side of this inequality. From Lemma 1 and (10) we have $|\tilde{R}_h u^* - R_h u^*|_0 \le ch$.

To estimate the second term on the right-hand side we use the fact that $p(u^*) \in C^{0,1}(\Omega)$; see, e.g., [15] pg. 237, and consider the following chain of inequalities

$$\begin{aligned} &|R_h Q(u^*) - Q_h(\tilde{R}_h u^*)|_0 \\ &\quad = \left| R_h \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p(u^*)\right\}\right\} - \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p_h(\tilde{R}_h u^*)\right\}\right\}\right|_0 \\ &\quad \le \left| R_h \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p(u^*)\right\}\right\} - \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p(u^*)\right\}\right\}\right|_0 \\ &\qquad + \left| \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p(u^*)\right\}\right\} - \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, R_h p(u^*)\right\}\right\}\right|_0 \\ &\qquad + \left| \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, R_h p(u^*)\right\}\right\} - \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p_h(\tilde{R}_h u^*)\right\}\right\}\right|_0. \\ &\quad \le \left| R_h \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p(u^*)\right\}\right\} - \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p(u^*)\right\}\right\}\right|_0 \\ &\qquad + \frac{1}{\nu}\, |p(u^*) - R_h p(u^*)|_0 + \frac{1}{\nu}\, |R_h p(u^*) - p_h(\tilde{R}_h u^*)|_0. \end{aligned}$$

Since $|p_h(\tilde{R}_h u^*) - R_h p(u^*)|_0 \le ch^2$ and because of the approximation property of R_h (see [14]) the last two terms in the above inequality are of order h^2. The final estimate

$$\begin{aligned} &\left| R_h \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p(u^*)\right\}\right\} - \max\left\{\underline{u}, \min\left\{\bar{u}, \frac{1}{\nu}\, p(u^*)\right\}\right\}\right|_0 \\ &\quad \le ch\, |p(u^*)|_{C^{0,1}(\Omega)}, \end{aligned}$$

follows from [14], pg. 233, where we use the fact that $\max\{\underline{u}, \min\{\bar{u}, \frac{1}{\nu} p(u^*)\}\} \in C^{0,1}(\Omega)$. It follows that $|u_h^* - \tilde{R}_h u^*|_0 \leq ch$. Therefore, using (10), we have

$$|u_h^* - R_h u^*|_0 \leq ch. \tag{15}$$

Now we consider the state equation. Denote by $\bar{y}_h$ the solution to

$$-\Delta_h \bar{y}_h = \tilde{R}_h B u^* + \tilde{R}_h g.$$

We have that $|\bar{y}_h - R_h y^*|_0 \leq ch^2$. Furthermore we have $|y_h^* - \bar{y}_h|_1 \leq c\,|u_h^* - \tilde{R}_h u^*|_0$ and $|y_h^* - \bar{y}_h|_0 \leq c\,|y_h^* - \bar{y}_h|_1$. Using these estimates and (15) in $|y_h^* - R_h y^*|_0 \leq |y_h^* - \bar{y}_h|_0 + |\bar{y}_h - R_h y^*|_0$ we obtain

$$|y_h^* - R_h y^*|_0 \leq ch. \tag{16}$$

From this estimate and (13) the following estimate results

$$|p_h^* - R_h p^*|_0 \leq ch. \tag{17}$$

Estimates (15), (16), and (17) are sub-optimal in the sense that $H^2(\Omega)$-regularity of y^* and p^* would suggest $\mathcal{O}(h^2)$ convergence estimates. Such results are impeded by the lack of the estimate $|y_h^* - \bar{y}_h|_0 \leq ch\,|y_h^* - \bar{y}_h|_1$. For finite element approximations this is a consequence of the Aubin-Nitsche duality argument. If the estimate were to hold, we obtain h^2 in (16) and (17). Moreover, our numerical experiments show that the estimate (15) for the control is pessimistic. In fact, we observe

$$|u_h^* - R_h u^*|_0 \leq ch^{3/2}, \tag{18}$$

and $\mathcal{O}(h^2)$ convergence for the state and the adjoint variables. In the recent publication [19] these convergence rates could be verified for finite element approximations.

To illustrate that these convergence rates are obtained by our algorithm, consider the following exact solution to (6) with

$$g(x_1, x_2) = -u + 2\pi^2 \sin(\pi x_1) \sin(\pi x_2), \tag{19}$$

$$z(x_1, x_2) = -\Delta p + y, \tag{20}$$

where

$$y(x_1, x_2) = \sin(\pi x_1) \sin(\pi x_2), \tag{21}$$

$$p(x_1, x_2) = \sin(8\pi x_1) \sin(8\pi x_2), \tag{22}$$

$$u(x_1, x_2) = \max\{-1,\ \min\{1,\ p/\nu\}\}. \tag{23}$$

Note that the control is active for each $\nu < 1$. Results of experiments with this test case are reported in Table 1. One can see that the solution errors for y and p reduce as a factor of

Table 1. Accuracy results; $\nu = 10^{-8}$.

Mesh	$\|y - y_h\|_0$	$\|p - p_h\|_0$	$\|u - u_h\|_0$
129×129	$1.63\,10^{-7}$	$1.60\,10^{-3}$	$1.54\,10^{-3}$
257×257	$4.20\,10^{-8}$	$4.03\,10^{-4}$	$5.26\,10^{-4}$
513×513	$1.11\,10^{-8}$	$1.00\,10^{-4}$	$1.80\,10^{-4}$
1025×1025	$2.93\,10^{-9}$	$2.51\,10^{-5}$	$6.26\,10^{-5}$

four by halving the mesh size, showing second-order convergence. On the other hand, the error for the control scales as $h^{3/2}$ with $c \approx 2$ in (18).

4. The multigrid method

The purpose of this paper is to present multigrid algorithms for optimality systems, arising from constrained optimal control problems, that have multigrid efficiency and are robust with respect to a large range of values of the cost of the control.

For the presentation, let us index the operators and variables defined on the grid with mesh size $h = h_k = h_0/2^k$, $k = 1, \ldots, L$, with the index k.

For multigrid purposes, in order to formulate (13) on all grids we need to have a multigrid full approximation storage (FAS) [6] representation of the problem. That is, a representation where the solution and not the error is computed on all grids. In fact, in order to impose the constraints the variables u_h and p_h must be available at all levels. In the remainder of this section we describe our algorithm formulated in the FAS framework.

Recall (13) given by

$$-\Delta_h y_h - B_h u_h = g_h, \tag{24}$$

$$-\Delta_h p_h + y_h = z_h, \tag{25}$$

$$(\nu u_h - B_h^* p_h) \cdot (v_h - u_h) \geq 0 \quad \text{for all } v_h \in U_{ad\,h}. \tag{26}$$

In general, an initial approximation to the solution of this problem will differ from the exact solution because of errors involving high-frequency as well as low-frequency components. In order to solve for all frequency components of the error, the multigrid strategy combines two complementary schemes. The high-frequency components of the error are reduced by smoothing iterations while the low-frequency error components are effectively reduced by a coarse-grid correction method.

On the grid of level k, the smoothing procedure is denoted by S_k, and S_k^m is the smoothing operator applied m times on the pair $\mathbf{w}_h = (y_h, p_h)$. This operator is defined later in this section. To correct for the smooth component of the error, a coarse grid correction is defined. First, a coarse grid problem is constructed on the grid with mesh size $H = h_{k-1}$. That is,

$$-\Delta_H y_H - B_H u_H = I_h^H g_h + \tau(y)_h^H, \tag{27}$$

$$-\Delta_H p_H + y_H = I_h^H z_h + \tau(p)_h^H,$$

$$(\nu u_H - B_H^* p_H) \cdot (v_H - u_H) \geq 0 \quad \text{for all } v_H \in U_{ad\,H}, \tag{28}$$

where $I_h^H : L_h^2 \to L_H^2$ denotes a restriction operator, and $\tau(y)_h^H$ and $\tau(y)_h^H$ are fine-to-coarse defect corrections defined by

$$\tau(y)_h^H = -\Delta_H \hat{I}_h^H y_h - B_H \hat{I}_h^H u_h - I_h^H(-\Delta_h y_h - B_h u_h), \tag{29}$$

$$\tau(p)_h^H = -\Delta_H \hat{I}_h^H p_h + \hat{I}_h^H y_h - I_h^H(-\Delta_h p_h + y_h), \tag{30}$$

with $\hat{I}_h^H : L_h^2 \to L_H^2$ a restriction operator not necessarily equal to I_h^H. We choose $\hat{I}_h^H$ to be straight injection. Once the coarse grid problem is solved, the coarse grid correction follows

$$y_h^{new} = y_h + I_H^h\big(y_H - \hat{I}_h^H y_h\big), \tag{31}$$

$$p_h^{new} = p_h + I_H^h\big(p_H - \hat{I}_h^H p_h\big), \tag{32}$$

where $I_H^h : L_H^2 \to L_h^2$ represents an interpolation operator. If the high frequency components of the error on the finer grid are indeed well damped, then the grid Ω_H should provide enough resolution for the error of $\mathbf{w}_h$ and hence $\mathbf{w}_H - \hat{I}_h^H \mathbf{w}_h$ should be a good approximation to this error. This idea of transferring the problem to be solved to a coarser grid can be applied along the set of nested meshes. One starts at level L with a zero approximation and applies the smoothing iteration m_1 times. Then the problem is transferred to a coarser grid and so on. Once the coarsest grid is reached, one solves the coarsest problem to convergence by applying, as we do, a few steps of the smoothing iteration. The solution obtained on each grid is then used to correct the approximation on the next finer grid. The coarse grid correction followed by m_2 post-smoothing steps is applied from one grid to the next, down to the finest grid with level L. This entire process represents one FAS-(m_1, m_2) multigrid cycle.

The application of N FAS cycles is denoted by N-FAS. One can choose a starting grid with a level number $K < L$ which is coarser than the finest grid where the solution is desired. In this case one applies N-FAS on level K and then the solution is interpolated on the next finer grid. The interpolation provides a first approximation for the N-FAS on this finer level and so on until the finest grid is reached. The combination of the nested iteration technique and the FAS scheme is called the full multigrid (FMG) method.

Denote the optimality system by $A(\mathbf{w}_h) = \mathbf{f}_h$. Then the multigrid FAS-(m_1, m_2)-cycle algorithm, expressed in terms of the (nonlinear) multigrid iteration operator B_K to solve this problem in recursive form is given as follows:

Multigrid FAS-(m_1, m_2)-Cycle

Set $B_1(\mathbf{w}^{(0)}) \approx A_1^{-1}$ (e.g., iterating with S_1). For $k = 2, \ldots, K$ define B_k in terms of B_{k-1} as follows. Let $\mathbf{q}^0 = 0$.

1. Set the starting approximation $\mathbf{w}_k^{(0)}$.
2. Pre-smoothing. Define $\mathbf{w}_k^{(l)}$ for $l = 1, \ldots, m_1$, by

$$\mathbf{w}_k^{(l)} = S_k\big(\mathbf{w}_k^{(l-1)}, \mathbf{f}_k\big).$$

3. Coarse grid correction. Set $\mathbf{w}_k^{(m_1+1)} = \mathbf{w}_k^{(m_1)} + I_{k-1}^k(\mathbf{q}^m - \hat{I}_k^{k-1}\mathbf{w}_k^{(m_1)})$ where $\mathbf{q}^i$ for $i = 1, \ldots, m$ is defined by

$$\mathbf{q}^i = \mathbf{q}^{i-1} + B_{k-1}\big(\hat{I}_k^{k-1}\mathbf{w}_k^{(m_1)}\big)\big[I_k^{k-1}\big(\mathbf{f}_k - A_k\big(\mathbf{w}_k^{(m_1)}\big)\big) + A_{k-1}\big(\hat{I}_k^{k-1}\mathbf{w}_k^{(m_1)}\big) - A_{k-1}\mathbf{q}^{i-1}\big].$$

4. Post-smoothing. Define $\mathbf{w}_k^{(l)}$ for $l = m_1 + 2, \ldots, m_1 + m_2 + 1$, by

$$\mathbf{w}_k^{(l)} = S_k\big(\mathbf{w}_k^{(l-1)}, \mathbf{f}_k\big).$$

5. Set $B_k(\mathbf{w}_k^{(0)})\,\mathbf{f}_k = \mathbf{w}_k^{(m_1+m_2+1)}$.

Notice that we can perform m two-grid iterations at each working level. For $m = 1$ we have a V-cycle and for $m = 2$ we have a W-cycle; m is called the cycle index [21]. Our numerical experience shows that in case of constrained control problems the use of W-cycles results in a robust multigrid iteration.

Our main contribution to the formulation of multigrid schemes for constrained optimal control problems is a new relaxation scheme. In the present context, the smoothing iteration must reduce the high-frequency components of the error and must preserve the inequality constraint. In case of unconstrained optimal control problems of the type considered here, the Fourier analysis presented in [4] proves that collective Gauss-Seidel iteration has the smoothing property required in the multigrid method. In order to present our modification of this iteration that takes account of the presence of constraints consider (24) and (25) at $\mathbf{x} \in \Omega_h$, where $\mathbf{x} = (ih, jh)$ and i, j index the grid points, e.g., lexicographically. We have

$$-(y_{i-1\,j} + y_{i+1\,j} + y_{i\,j-1} + y_{i\,j+1}) + 4y_{ij} - h^2 B_h u_{ij} = h^2 g_{ij} + h^2 f_{ij}^{(y)}, \tag{33}$$

$$-(p_{i-1\,j} + p_{i+1\,j} + p_{i\,j-1} + p_{i\,j+1}) + 4p_{ij} + h^2 y_{ij} = h^2 z_{ij} + h^2 f_{ij}^{(p)}, \tag{34}$$

$$(\nu u_{ij} - B_h^* p_{ij}) \cdot (v_{ij} - u_{ij}) \geq 0 \quad \text{for all} v_h \in U_{adh}, \tag{35}$$

where $f^{(y)}$ and $f^{(p)}$ have been introduced to take into account the presence of defect corrections in (27) and (28).

A Gauss-Seidel step at $\mathbf{x}$ consists in updating the values y_{ij} and p_{ij} such that the resulting residuals of the two equations at that point are zero. The neighboring variables are considered constant during this process. Therefore, define the two constants

$$C_y = (y_{i-1\,j} + y_{i+1\,j} + y_{i\,j-1} + y_{i\,j+1}) + h^2 g_{ij} + h^2 f_{ij}^{(y)},$$

and

$$C_p = (p_{i-1\,j} + p_{i+1\,j} + p_{i\,j-1} + p_{i\,j+1}) + h^2 f_{ij}^{(p)}.$$

Replacing these two constants in (33) and (34), we obtain y_{ij} and p_{ij} as functions of u_{ij} as follows

$$y_{ij} = (C_y + h^2 B_h u_{ij})/4, \tag{36}$$

and

$$p_{ij} = (4 C_p - h^2 C_y + 4 h^2 z_{ij} - h^4 B_h u_{ij})/16. \tag{37}$$

Now to obtain the u_{ij} update, replace the expression for p_{ij} in the inequality constraint and define the auxiliary variable

$$\tilde{u}_{ij} = \frac{1}{16\nu + h^4} B_h^* (4 C_p - h^2 C_y + 4 h^2 z_{ij}). \tag{38}$$

Then, the new value for u_{ij} resulting from our Gauss-Seidel step is given by

$$u_{ij} = \begin{cases} \bar{u}_{ij} & \text{if} \quad \tilde{u}_{ij} \geq \bar{u}_{ij} \\ \tilde{u}_{ij} & \text{if} \quad \underline{u}_{ij} < \tilde{u}_{ij} < \bar{u}_{ij} \\ \underline{u}_{ij} & \text{if} \quad \tilde{u}_{ij} \leq \underline{u}_{ij} \end{cases} \tag{39}$$

for all $\mathbf{x} = (ih, jh) \in \omega_h$, $u_{ij} = 0$ otherwise. Possible ill-conditioning due to the division by $(16\nu + h^4)$ in (38) is confined by the constraints expressed in (39). With the new value of u_{ij} given, new values for y_{ij} and p_{ij} are obtained. This completes the Gauss-Seidel step.

The collective Gauss-Seidel step defined by (36), (37), (38), and (39) satisfies the inequality constraint. In fact, consider any grid point such that $\tilde{u} > \bar{u}$; then from (39) we have $u = \bar{u}$. Therefore $(v - u) \leq 0$ for any $v \in U_{adh}$. On the other hand we have

$$\begin{aligned} \nu u - p &= \nu u - (4 C_p - h^2 C_y + 4 h^2 z_{ij} - h^4 B_h u_{ij})/16 \\ &= [(16\nu + h^4)u - (4 C_p - h^2 C_y + 4 h^2 z_{ij})]/16 \\ &< [(16\nu + h^4)\tilde{u} - (4 C_p - h^2 C_y + 4 h^2 z_{ij})]/16 = 0. \end{aligned}$$

Therefore $(\nu u - p) \cdot (v - u) \geq 0$ for all $v \in U_{adh}$. Similarly one proves that if $\tilde{u} < \underline{u}$, then the choice $u = \underline{u}$ satisfies the inequality constraint. The case $\underline{u} \leq \tilde{u} \leq \bar{u}$ is obvious. Similarly in case $\nu = 0$, it can be proved that the Gauss-Seidel iteration defined above satisfies (8). Because of (39) we can consider the present iteration belongs to the class of projected Gauss-Seidel schemes [7].

We complete this section describing the prolongation and restriction operators used in our algorithm. Between two grids $\bar{\Omega}_k$ and $\bar{\Omega}_{k-1}$, corresponding to mesh sizes h_k and h_{k-1}, we apply the bilinear prolongation operator, $I_{k-1}^k : L_{k-1}^2 \to L_k^2$, given in stencil form by

$$I_{k-1}^k = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \tag{40}$$

For restricting the residuals we use the half-weighting restriction operator, $I_k^{k-1} : L_k^2 \to L_{k-1}^2$, given in stencil form by

$$I_k^{k-1} = \frac{1}{8}\begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \tag{41}$$

We conclude this section with remarks on the complexity of our multigrid algorithm. Regarding memory complexity notice that we have followed a standard implementation of the FAS structure. Therefore we have $\sigma \approx 1.5$ where σ is the grid complexity defined as the total number of unknowns on all levels divided by the number of unknowns on the finest level. Further, mesh independent convergence factors imply optimal computational complexity.

5. Numerical experiments

In this section we present a numerical investigation of the computational performance of the proposed multigrid method to solve constrained optimal control problems. We report the values of the tracking functional $|y_h - z_h|_0$ depending on the value of the cost of the control and the residual of the primal and of the adjoint equations which are denoted by $r_h(y)$ and $r_h(p)$, respectively.

We also report values of the convergence factor defined as the "asymptotic" value of the ratio between the discrete L^2-norm of the residuals resulting from two successive multigrid cycles on a given mesh [13]. That is,

$$\rho(y) = \lim_N \frac{\left|r_h(y)^{(N+1)}\right|_0}{\left|r_h(y)^{(N)}\right|_0} \quad \text{and} \quad \rho(p) = \lim_N \frac{\left|r_h(p)^{(N+1)}\right|_0}{\left|r_h(p)^{(N)}\right|_0}. \tag{42}$$

For the coarsest grid we have $h_1 = 1/8$ and we use up to eleven levels. For $L = 11$ we have a 8193×8193 mesh. In all experiments we use $m_1 = m_2 = 2$ smoothing steps and the FMG-W-cycle version of the algorithm with initial level $K = 3$. Starting with $k = K$ three W-cycles are performed at this level. Then the solution is (cubic) interpolated to the next finer working level and the same number of W-cycles are performed. This process is continued until the finest level is reached where $k = L$ and at most ten W-cycles are carried out. The stopping criterion is $|r_h(y)|_0 + |r_h(p)|_0 < 10^{-10}$. Recall that with this setting the typical multigrid convergence factor for the Poisson problem in a square domain is $\rho \approx 0.08$.

We investigate the convergence behavior of the multigrid solver depending on the mesh size, on the value of ν, and on the presence of constraints.

Four cases are considered. In the first case ω coincides with the entire domain and no constraints are given. Results for this case are reported to allow comparison with the forthcoming cases. In the second case we also have $\omega = \Omega$ and severe constraints are prescribed. In the third and fourth case we investigate the multigrid solution process when

$\omega \subset \Omega$, without and with constraints, respectively. In the experiments that follow we choose $g = 0$. A first series of experiments is performed with the following objective function

$$z(x_1, x_2) = \sin(2\pi x_1) \sin(\pi x_2).$$

Case 1. $\omega = \Omega$, no constraints.

This case has been investigated by the authors in [4]. In this reference stability of the finite difference optimality system and optimal-order error estimates in the discrete L^2-norm and in the discrete H^1-norm were proved. Sharp convergence factor estimates of the two grid method for the optimality system were obtained by means of local Fourier analysis. A multigrid convergence theory was provided which guarantees convergence of the multigrid process towards solutions of the optimality system. The smoothing iteration considered in [4] was the collective Gauss-Seidel iteration that coincides with the smoothing iteration considered here when no constraints are imposed. The collective Gauss-Seidel iteration is also a special case of the Newton-Gauss-Seidel iteration presented in [3] to solve singular optimal control problems. In all cases optimal convergence factors are obtained as reported in Table 2. The state and control solutions for the present case are depicted in figure 2.

Notice that, in the unconstrained case the convergence behavior of the multigrid solution process is independent of the value of ν and of the mesh size. Note also that for the present example z is attainable, i.e. there exists u such that $y(u) = z$. Similar computational performance is obtained for non-attainable z.

Case 2. $\omega = \Omega$, constraints: $\underline{u} = -30$ and $\bar{u} = 30$.

For this case, the constraints are active in large portions of the domain for all three choices of $\nu = \{10^{-4}, 10^{-6}, 10^{-8}\}$ considered here. For $\nu = 10^{-6}$ this can be seen in figure 2. From the results of numerical experiments reported in Table 3 we observe that for $\nu = 10^{-4}$ the multigrid convergence behavior is similar to that observed in the unconstrained case. Reducing the value of ν results in steeper gradients of the adjoint and control variables,

Table 2. Case 1: Results of experiments.

Mesh	$\rho(y)$, $\rho(p)$	$\lvert y - z\rvert_0$	$\lvert r(y)\rvert_0$, $\lvert r(p)\rvert_0$
		$\nu = 10^{-6}$	
129×129	0.03, 0.07	$1.21\,10^{-3}$	$4.6\,10^{-10}$, $1.3\,10^{-13}$
257×257	0.03, 0.02	$1.21\,10^{-3}$	$1.4\,10^{-10}$, $1.2\,10^{-14}$
513×513	0.03, 0.01	$1.21\,10^{-3}$	$1.7\,10^{-10}$, $8.9\,10^{-15}$
		$\nu = 10^{-8}$	
129×129	0.03, 0.05	$1.21\,10^{-5}$	$1.2\,10^{-10}$, $1.6\,10^{-14}$
257×257	0.03, 0.07	$1.21\,10^{-5}$	$8.9\,10^{-11}$, $3.6\,10^{-15}$
513×513	0.03, 0.07	$1.21\,10^{-5}$	$1.7\,10^{-10}$, $2.7\,10^{-15}$

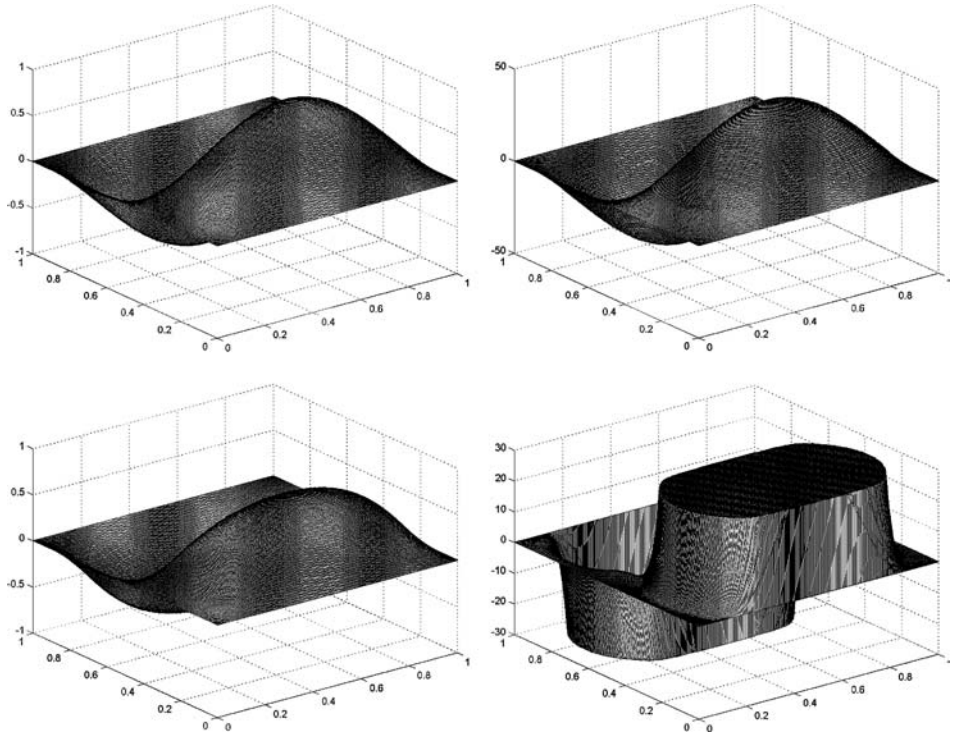

Figure 2. Numerical solutions for Case 1 (top) and Case 2 (bottom) the state (left) and the control (right); $\nu = 10^{-6}$ and 513×513 mesh.

particularly close to the boundary where p and u are required to be zero. Furthermore, decreasing ν results in an increasingly more complex switching structure of the control between upper and lower bounds; see figure 3. The results for $\nu = 10^{-6}$ in Table 3 suggest that once the mesh size is sufficiently fine to resolve completely the switching structure the typical multigrid convergence rate is obtained. They further indicate that the multigrid convergence factor depends only weakly on the mesh size provided the problem is sufficiently well resolved on the mesh.

The ability of the multigrid scheme in solving constrained control problems with very small value of ν allows to investigate the occurrence of bang-bang control for the present class of problems. In particular, with the choice of z given above we can observe fast switching of the control function in the x_2 direction as depicted in figure 3. In this figure we give plots of the control function for $x_1 = 3/4$ and $x_2 \in [0, 1]$ for the following choices of $\nu \in \{10^{-8}, 10^{-10}, 10^{-12}, 0\}$. We can see that as the value of ν is reduced the number of switching points increases.

The solution obtained for $\nu = 0$ is interesting. In this case, by further refining the mesh size additional switching points can be seen closer to the boundary while the existing switching points obtained at the previous coarser grids are retained; see figure 4.

Table 3. Case 2: Results of experiments.

Mesh	$\rho(y)$, $\rho(p)$	$\|y - z\|_0$	$\|r(y)\|_0$, $\|r(p)\|_0$
		$\nu = 10^{-4}$	
129×129	0.04, 0.04	$1.11\,10^{-1}$	$3.1\,10^{-10}$, $1.2\,10^{-13}$
257×257	0.03, 0.04	$1.11\,10^{-1}$	$6.8\,10^{-10}$, $7.1\,10^{-14}$
513×513	0.03, 0.04	$1.11\,10^{-1}$	$4.9\,10^{-10}$, $1.5\,10^{-13}$
1025×1025	0.03, 0.03	$1.11\,10^{-1}$	$3.2\,10^{-10}$, $7.2\,10^{-13}$
		$\nu = 10^{-6}$	
129×129	0.56, 0.56	$5.30\,10^{-2}$	$1.3\,10^{-6}$, $2.2\,10^{-10}$
257×257	0.52, 0.51	$5.30\,10^{-2}$	$1.5\,10^{-7}$, $1.3\,10^{-11}$
513×513	0.03, 0.03	$5.30\,10^{-2}$	$3.5\,10^{-10}$, $5.3\,10^{-14}$
1025×1025	0.03, 0.03	$5.30\,10^{-2}$	$2.2\,10^{-10}$, $2.2\,10^{-13}$
		$\nu = 10^{-8}$	
129×129	0.63, 0.63	$5.28\,10^{-2}$	$1.6\,10^{-3}$, $8.3\,10^{-8}$
257×257	0.54, 0.54	$5.28\,10^{-2}$	$2.4\,10^{-6}$, $7.4\,10^{-11}$
513×513	0.64, 0.60	$5.28\,10^{-2}$	$2.5\,10^{-7}$, $3.7\,10^{-12}$
1025×1025	0.68, 0.66	$5.28\,10^{-2}$	$2.7\,10^{-7}$, $2.1\,10^{-12}$
2049×2049	0.74, 0.71	$5.28\,10^{-2}$	$7.8\,10^{-7}$, $3.5\,10^{-12}$
4097×4097	0.76, 0.70	$5.28\,10^{-2}$	$7.4\,10^{-8}$, $2.9\,10^{-12}$

We complete the discussion of Case 2. considering another desired state given by

$$z_1(x_1, x_2) = \sin(4\pi x_1)\sin(2\pi x_2).$$

The difference between this objective function and the previous one is that the gradient of z_1 is larger close to the boundary. For the choice $\nu = 0$ the constraints are everywhere active, i.e. differently from the previous case with desired state z the control is bang-bang. Moreover no fast switching of the control occurs. In figure 5 the optimal control and the corresponding state for $\nu = 0$ are depicted. The numerical results in Table 4 document the convergence factors.

Table 4. Results of experiments with z_1 and $\nu = 0$.

Mesh	$\rho(y)$, $\rho(p)$	$\|y - z\|_0$	$\|r(y)\|_0$, $\|r(p)\|_0$
513×513	0.12, 0.13	$3.70\,10^{-1}$	$2.9\,10^{-8}$, $1.3\,10^{-13}$
1025×1025	0.12, 0.13	$3.70\,10^{-1}$	$2.5\,10^{-8}$, $4.2\,10^{-13}$
2049×2049	0.12, 0.16	$3.70\,10^{-1}$	$1.9\,10^{-8}$, $1.6\,10^{-12}$

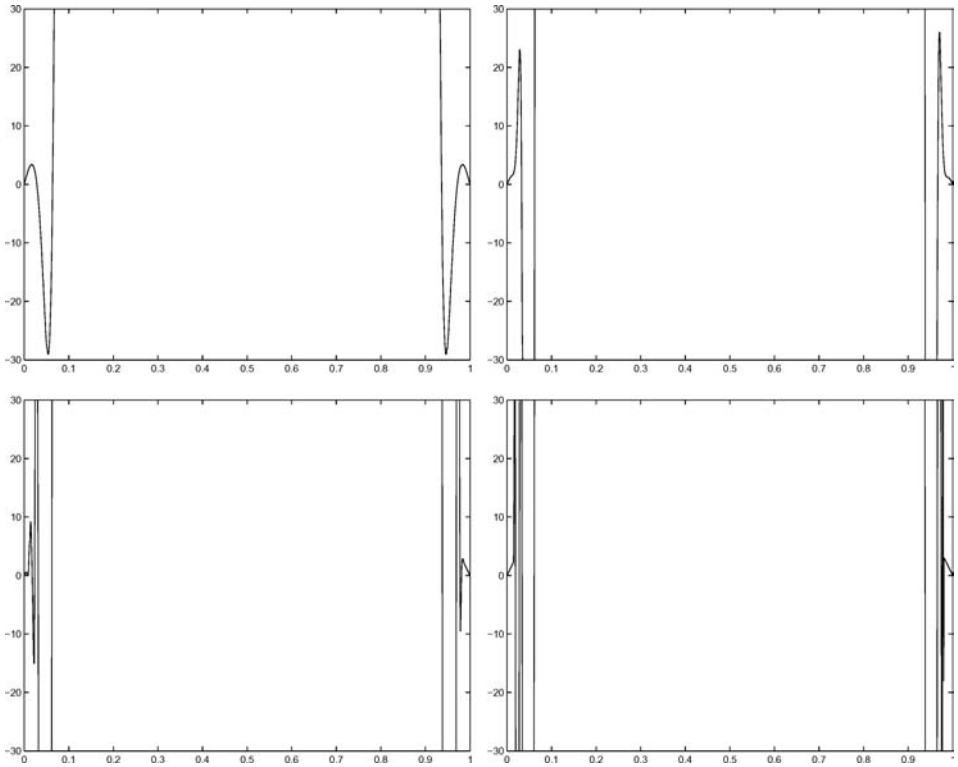

Figure 3. The control function for $x_1 = 3/4$ and $x_2 \in [0, 1]$ obtained with $\nu = 10^{-8}$ (top left), $\nu = 10^{-10}$ (top right), $\nu = 10^{-12}$ (bottom left), and $\nu = 0$ (bottom right); 2049×2049 mesh.

Case 3. $\omega \subset \Omega$, no constraints. Here ω is given by

$$\omega = \{\mathbf{x} \in \Omega : (x_1 - 1/2)^2 + (x_2 - 1/2)^2 < \sqrt{(7/160)}\},$$

and z is the desired state.

In this case we have $u = p/\nu$ in ω and otherwise we set $u = 0$. Results of experiments with the present setting are reported in Table 5 where typical multigrid convergence behavior can be observed.

Case 4. $\omega \subset \Omega$ as in Case 3, and constraints given by $\underline{u} = -30$ and $\bar{u} = 30$.

In Table 6, optimal multigrid convergence factors for the present case with moderate values of ν are reported. For ν sufficiently small the convergence factors worsen (while

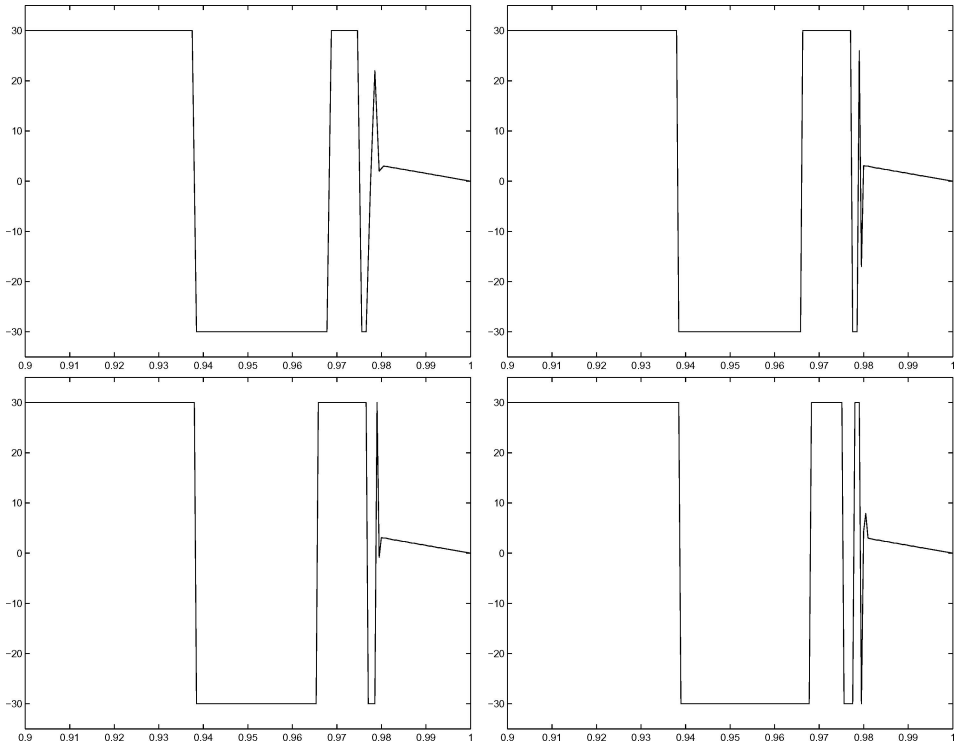

Figure 4. Switching of the control function for $x_1 = 3/4$ and $x_2 \in [0.9, 1]$ (notice the scaling) obtained with $\nu = 0$ on increasingly finer meshes: 1025×1025 (top left), 2049×2049 (top right), 4097×4097 (bottom left), and 8193×8193 (bottom right).

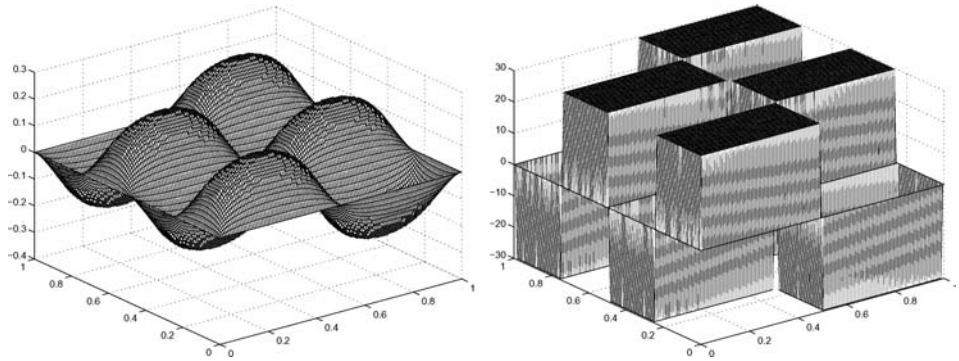

Figure 5. Numerical solutions with z_1 and $\nu = 0$. The state (left) and the control (right); 257×257 mesh.

Table 5. Case 3: Results of experiments.

Mesh	$\rho(y)$, $\rho(p)$	$\|y - z\|_0$	$\|r(y)\|_0$, $\|r(p)\|_0$
		$\nu = 10^{-6}$	
129×129	0.10, 0.10	$1.85\,10^{-1}$	$1.1\,10^{-9}$, $1.0\,10^{-13}$
257×257	0.05, 0.05	$1.84\,10^{-1}$	$3.3\,10^{-10}$, $5.0\,10^{-14}$
513×513	0.05, 0.05	$1.84\,10^{-1}$	$2.5\,10^{-10}$, $1.1\,10^{-13}$
1025×1025	0.05, 0.05	$1.84\,10^{-1}$	$3.3\,10^{-10}$, $5.5\,10^{-13}$
		$\nu = 10^{-8}$	
129×129	0.37, 0.37	$1.58\,10^{-1}$	$5.6\,10^{-3}$, $2.5\,10^{-7}$
257×257	0.16, 0.15	$1.57\,10^{-1}$	$4.2\,10^{-7}$, $7.8\,10^{-12}$
513×513	0.05, 0.05	$1.57\,10^{-1}$	$1.7\,10^{-10}$, $1.0\,10^{-13}$
1025×1025	0.05, 0.05	$1.57\,10^{-1}$	$1.9\,10^{-10}$, $4.4\,10^{-13}$

Table 6. Case 4: Results of experiments.

Mesh	$\rho(y)$, $\rho(p)$	$\|y - z\|_0$	$\|r(y)\|_0$, $\|r(p)\|_0$
		$\nu = 10^{-4}$	
129×129	0.05, 0.05	$4.24\,10^{-1}$	$4.4\,10^{-10}$, $5.0\,10^{-11}$
257×257	0.05, 0.05	$4.23\,10^{-1}$	$3.2\,10^{-10}$, $1.4\,10^{-13}$
513×513	0.05, 0.05	$4.23\,10^{-1}$	$2.3\,10^{-10}$, $5.6\,10^{-13}$
1025×1025	0.05, 0.05	$4.23\,10^{-1}$	$2.7\,10^{-10}$, $2.6\,10^{-12}$
		$\nu = 10^{-6}$	
129×129	0.61, 0.61	$4.23\,10^{-1}$	$6.9\,10^{-6}$, $1.1\,10^{-9}$
257×257	0.67, 0.67	$4.22\,10^{-1}$	$3.0\,10^{-6}$, $2.7\,10^{-10}$
513×513	0.71, 0.70	$4.22\,10^{-1}$	$1.1\,10^{-6}$, $5.7\,10^{-11}$
1025×1025	0.72, 0.70	$4.22\,10^{-1}$	$3.7\,10^{-7}$, $1.0\,10^{-11}$
		$\nu = 10^{-8}$	
129×129	0.57, 0.57	$4.23\,10^{-1}$	$7.8\,10^{-5}$, $1.3\,10^{-8}$
257×257	0.58, 0.57	$4.22\,10^{-1}$	$1.2\,10^{-5}$, $6.9\,10^{-10}$
513×513	0.64, 0.63	$4.22\,10^{-1}$	$5.5\,10^{-6}$, $1.2\,10^{-10}$
1025×1025	0.70, 0.68	$4.22\,10^{-1}$	$2.5\,10^{-6}$, $2.7\,10^{-11}$

remaining approximately mesh independent) taking values typical of multigrid applied to complementarity problems as in [7].

For comparison we report results obtained with the choice $\nu = 0$ and z_1 as objective function. Results of numerical experiments for this case are reported in Table 7. In figure 6, bang-bang control in ω can be seen.

Table 7. Results of experiments with z_1, $\omega \subset \Omega$, and $\nu = 0$.

Mesh	$\rho(y)$, $\rho(p)$	$\|y - z\|_0$	$\|r(y)\|_0$, $\|r(p)\|_0$
513×513	0.27, 0.23	$4.87\,10^{-1}$	$6.6\,10^{-8}$, $2.4\,10^{-13}$
1025×1025	0.22, 0.21	$4.87\,10^{-1}$	$1.6\,10^{-7}$, $6.3\,10^{-13}$
2049×2049	0.20, 0.21	$4.87\,10^{-1}$	$9.0\,10^{-9}$, $2.5\,10^{-12}$

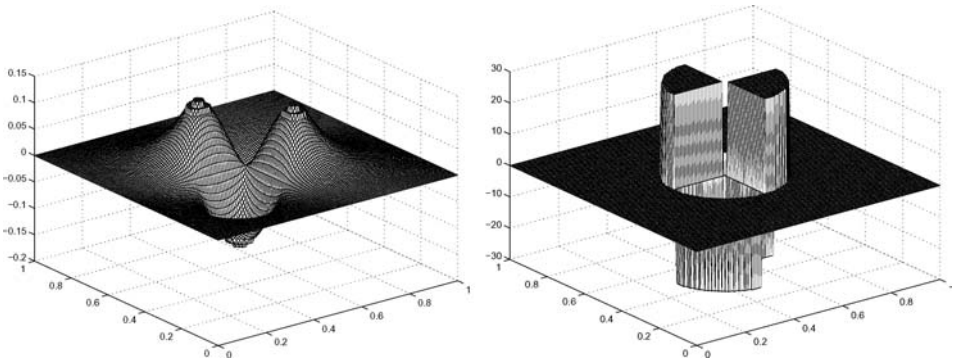

Figure 6. Numerical solutions with z_1 and $\omega \subset \Omega$ and $\nu = 0$. The state (left) and the control (right); 257×257 mesh.

6. Constrained boundary optimal control problems

We now turn to boundary optimal control problems with constraints. We focus on the following optimal control problem. Minimize

$$J(y, u) = \frac{1}{2}||y - z||^2_{L^2(\Omega)} + \frac{\nu}{2}||u||^2_{L^2(\partial\Omega)}, \tag{43}$$

subject to $u \in U_{ad} \subset L^2(\partial\Omega)$ and

$$\begin{aligned} -\Delta y + y &= g \quad \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= u \quad \text{on } \partial\Omega, \end{aligned} \tag{44}$$

where Ω is a open bounded set of $\mathbf{R}^2$, $g \in L_2(\Omega)$, $z \in L_2(\Omega)$ is the objective function, and $\nu \geq 0$. The set of admissible controls is given in this case by

$$U_{ad} = \{u \in L^2(\partial\Omega) \,|\, \underline{u}(\mathbf{x}) \leq u(\mathbf{x}) \leq \bar{u}(\mathbf{x}) \text{ a.e. in } \partial\Omega\}, \tag{45}$$

where $\underline{u}$ and $\bar{u}$ are functions of $L^\infty(\partial\Omega)$.

For the existence of a unique solution to (43)–(45) we refer to [16]. The solution is characterized by the following optimality system

$$\begin{aligned} -\Delta y + y &= g \quad \text{in } \Omega, \\ \frac{\partial y}{\partial n} &= u \quad \text{on } \partial\Omega, \\ -\Delta p + p &= -(y - z) \quad \text{in } \Omega, \\ \frac{\partial p}{\partial n} &= 0 \quad \text{on } \partial\Omega, \\ (\nu u - p, v - u) &\geq 0 \quad \text{for all } v \in U_{ad}. \end{aligned} \tag{46}$$

After discretization the optimal control problem becomes

$$\begin{cases} \min \dfrac{1}{2}|y_h - \tilde{R}_h z|_0^2 + \dfrac{\nu}{2}|u_h|_0^2, \\ -\Delta_h y_h = \tilde{R}_h g, \\ \partial_h^n y_h = u_h \qquad\qquad . \end{cases} \tag{47}$$

Here, ∂_h^n denotes the second-order centered difference quotient with orientation normal to the boundary. The optimality system related to (47) is found to be

$$\begin{aligned} -\Delta_h y_h + y_h &= g_h, \\ \partial_h^n y_h &= u_h, \\ -\Delta_h p_h + p_h &= -(y_h - z_h), \\ \partial_h^n p_h &= 0, \\ (\nu u_h - p_h) \cdot (v_h - u_h) &\geq 0 \quad \text{for all } v_h \in u_{ad_h}, \end{aligned} \tag{48}$$

where $g_h = \tilde{R}_h g$ and $z_h = \tilde{R}_h z$.

Notice that to solve (48) we need to realize the control on the boundary. For this purpose we eliminate the Neumann boundary conditions including them in the stencil of the differential operator considered at the boundary. We discuss this approach explicitly for one lateral boundary of $\Omega = (0, 1) \times (0, 1)$.

Let $\mathbf{x} = (ih,\ jh)$ be a boundary grid point on the side $x = 0$.

$$\begin{aligned} -(y_{i-1\,j} + y_{i+1\,j} + y_{i\,j-1} + y_{i\,j+1}) + (4 + h^2) y_{ij} &= h^2 g_{ij} + h^2 f_{ij}^{(y)}, \\ y_{i-1\,j} - y_{i+1\,j} &= 2h u_{ij}, \\ -(p_{i-1\,j} + p_{i+1\,j} + p_{i\,j-1} + p_{i\,j+1}) + (4 + h^2) p_{ij} + h^2 y_{ij} &= h^2 z_{ij} + h^2 f_{ij}^{(p)}, \\ p_{i-1\,j} - p_{i+1\,j} &= 0. \end{aligned}$$

Summing up the minus Laplacian stencil with the normal derivative the (ghost) variables outside of Ω are eliminated. We have

$$\begin{aligned} -(2y_{i+1\,j} + y_{i\,j-1} + y_{i\,j+1}) + (4 + h^2) y_{ij} - 2h u_{ij} &= h^2 g_{ij} + h^2 f_{ij}^{(y)}, \\ -(2p_{i+1\,j} + p_{i\,j-1} + p_{i\,j+1}) + (4 + h^2) p_{ij} + h^2 y_{ij} &= h^2 z_{ij} + h^2 f_{ij}^{(p)}. \end{aligned}$$

On the corners we consider the minus Laplacian stencil with the normal derivatives in both directions. The equations obtained in this way have the same structure as (33) and (34) and can be solved by the FAS multigrid method described in Section 4.

The application of the collective Gauss-Seidel iteration follows along the same lines as described in the previous sections. In the interior of the computational domain the collective Gauss-Seidel iteration reduces in this case to the single Gauss-Seidel iteration for the state equation while the residuals of the state equation and of the adjoint equation both enter in the relaxation of the adjoint variable.

To describe the Gauss-Seidel iteration on the boundary $x = 0$ in more detail, define

$$C_y = (2y_{i+1\,j} + y_{i\,j-1} + y_{i\,j+1}) + h^2 g_{ij} + h^2 f_{ij}^{(y)},$$
$$C_p = (2p_{i+1\,j} + p_{i\,j-1} + p_{i\,j+1}) + h^2 f_{ij}^{(p)}.$$

We obtain y_{ij} and p_{ij} as functions of u_{ij} as follows

$$y_{ij} = (C_y + 2hu_{ij})/(4 + h^2),$$

and

$$p_{ij} = ((4 + h^2)\, C_p - h^2\, C_y + (4 + h^2)\, h^2\, z_{ij} - 2h^3 u_{ij})/(4 + h^2)^2.$$

To obtain the u_{ij} update, replace the expression for p_{ij} in the inequality constraint and define

$$\tilde{u}_{ij} = \frac{1}{(4 + h^2)^2\, \nu + 2h^3}((4 + h^2)\, C_p - h^2\, C_y - (4 + h^2)\, h^2\, z_{ij}).$$

Then, the new value for u_{ij} resulting from Gauss-Seidel step on the boundary is given by

$$u_{ij} = \begin{cases} \bar{u}_{ij} & \text{if} \quad \tilde{u}_{ij} \geq \bar{u}_{ij} \\ \tilde{u}_{ij} & \text{if} \quad \underline{u}_{ij} < \tilde{u}_{ij} < \bar{u}_{ij} \\ \underline{u}_{ij} & \text{if} \quad \tilde{u}_{ij} \leq \underline{u}_{ij} \end{cases} \tag{49}$$

for all $\mathbf{x} = (ih,\, jh)$ on $\partial\Omega_h$. With the new value of u_{ij} given, the values for y_{ij} and p_{ij} are updated.

The bilinear prolongation operator described above applies also to the boundary variables. For restricting the residuals we use the full-weighting restriction operator, $I_k^{k-1} : L_k^2 \to L_{k-1}^2$, given in stencil form by

$$I_k^{k-1} = \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \tag{50}$$

This choice is necessary to guarantee at the boundary the right scaling for the coarse-grid problem formulation; see the discussion in [21]. Clearly, on the boundary the restriction

Table 8. Results of experiments, boundary control problems; 1025×1025 mesh.

ν	$\rho(y)$, $\rho(p)$	$\|y - z\|_0$	$\|r(y)\|_0$, $\|r(p)\|_0$
10^{-6}	0.05, 0.05	$8.09\,10^{-2}$	$1.7\,10^{-10}$, $2.9\,10^{-13}$
10^{-8}	0.14, 0.12	$8.09\,10^{-2}$	$3.7\,10^{-8}$,$2.9\,10^{-13}$
10^{-10}	0.28, 0.28	$8.09\,10^{-2}$	$4.7\,10^{-5}$,$9.9\,10^{-11}$
0	0.25, 0.26	$8.09\,10^{-2}$	$3.5\,10^{-5}$,$4.8\,10^{-11}$

operator is mirrored. For example on the left vertical boundary, we have $I_k^{k-1} = \frac{1}{16}\left[\begin{smallmatrix} 0 & 2 & 2 \\ 0 & 4 & 4 \\ 0 & 2 & 2 \end{smallmatrix}\right]$ and in the left-low corner $I_k^{k-1} = \frac{1}{16}\left[\begin{smallmatrix} 0 & 4 & 4 \\ 0 & 4 & 4 \\ 0 & 0 & 0 \end{smallmatrix}\right]$.

To numerically validate the present algorithm for solving boundary optimal control problems, consider the desired state given by

$$z(x_1, x_2) = \left(x_1^2 - x_2^2\right) \sin(\pi x_1) \sin(\pi x_2),$$

and $g = 0$. We choose constraints given by $\underline{u} = -1$ and $\bar{u} = 1$ which are active in part of the boundary for $\nu \leq 10^{-6}$. The multigrid setting is the same as in previous sections. Results for this case are reported in Table 8.

7. Conclusions

We presented a robust multigrid method for a class of optimality systems arising from elliptic constrained optimal control problems. The multigrid method was developed based on a collective Gauss-Seidel scheme that satisfies the given constraints pointwise. In applications we considered distributed and boundary control problems. With the present algorithm it was possible to investigate bang-bang control solutions.

Acknowledgment

We thank E. Sachs and F. Tröltzsch for interesting discussions on the bang-bang principle, and the Referees for helpful comments.

References

1. N. Arada, E. Casas, and F. Tröltzsch, "Error estimates for a semilinear elliptic control problem," Comp. Optim. Appl., vol. 23 pp. 201–229, 2002.
2. E. Arian and S. Ta'asan, "Smoothers for optimization problems," in Seventh Copper Mountain Conference on Multigrid Methods, Vol. CP3339, NASA Conference Publication, NASA, N. Duane Melson, T.A. Manteuffel, S.F. McCormick, and C.C. Douglas (Eds.), Hampton: VA, 1995, pp. 15–30.
3. A. Borzì and K. Kunisch, "The numerical solution of the steady state solid fuel ignition model and its optimal control," SIAM J. Sci. Comp., vol. 22 no. 1, pp. 263–284, 2000.
4. A. Borzì, K. Kunisch, and D.Y. Kwak, "Accuracy and convergence properties of the finite difference multigrid solution of an optimal control optimality system," SIAM J. Control Opt., vol. 41, pp. 1477–1497, 2003.

5. A. Borzì, "Multigrid methods for parabolic distributed optimal control problems," J. Comp. Appl. Math, vol. 157 pp. 365–382, 2003.
6. A. Brandt, "Multi-level adaptive solutions to boundary-value problems," Mathematics of Computation, vol. 31 pp. 333–390, 1977.
7. A. Brandt and C.W. Cryer, "Multi-grid algorithms for the solution of linear complementarity problems arising from free boundary problems," SIAM J. Sci. Stat. Comp., vol. 4, pp. 655–684, 1983.
8. R. Courant and D. Hilbert, Methods of Mathematical Physics, Interscience Publishing: New York, 1966.
9. Th. Dreyer, B. Maar, and V. Schulz, "Multigrid Optimization in Applications," J. Comput. Appl. Math., vol. 120, pp. 67–84, 2000.
10. K. Eppler and F. Tröltzsch, "On switching points of optimal controls for coercive parabolic boundary control problems," Optimization, vol. 17, pp. 93–101, 1986.
11. K. Glashoff and E. Sachs, "On theoretical and numerical aspects of the bang-bang principle," Numer. Math., vol. 29, pp. 93–113, 1977.
12. W. Hackbusch, "Fast solution of elliptic control problems," Journal of Optimization Theory and Application, vol. 31, pp. 565–581, 1980.
13. W. Hackbusch, Multi-grid Methods and Applications, Springer-Verlag: New York, 1985.
14. W. Hackbusch, Elliptic Differential Equations, Springer-Verlag: New York, 1992.
15. O. Ladyzhenskaya and N. Ural'tseva, Linear and Quasilinear Elliptic Equations, Academic Press: New York, 1968.
16. J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer: Berlin, 1971.
17. K. Malanowski, "Convergence of approximations vs. regularity of solutions for convex, control-constrained optimal-control problems," Appl. Math. Optim., vol. 8, pp. 69–95, 1981.
18. C. Meyer and A. Rösch, "Superconvergence properties of optimal control problems," to appear in SIAM J. Control Opt.
19. A. Rösch, "Error estimates for linear-quadratic control problems with control constraints," submitted to Optimization Methods and Software.
20. V. Schulz and G. Wittum, "Multigrid optimization methods for stationary parameter identification problems in groundwater flow," Multigrid Methods V, Lecture Notes in Computational Science and Engineering 3, in W. Hackbusch and G. Wittum (Eds.), Springer, 1998, pp. 276–288.
21. U. Trottenberg, C. Oosterlee, and A. Schüller, "Multigrid," Academic Press: London, 2001.