

Smoothers for control- and state-constrained optimal control problems

Alfio Borzi

Institut für Mathematik und Wissenschaftliches Rechnen, Karl-Franzens-Universität Graz, Heinrichstr. 36, A-8010 Graz, Austria (alfio.borzi@uni-graz.at)

Received:

Abstract. A framework for designing robust smoothing procedures for control- and state-constrained optimal control problems is presented. The focus is on minimization problems governed by elliptic partial differential equations with additional pointwise constraints on the control variable or on the state variable, respectively. The basic principle for the construction of the present smoothers is the solution of the corresponding optimality systems at grid-point level. A new approach is presented to cope with the lack of differentiability due to the presence of the constraints.

1 Introduction

We present a general framework for the design of smoothers for the multigrid solution of elliptic optimal control problems. These are constrained minimization problems governed by elliptic partial differential equations [2, 14, 19]. Specifically, we address the solution of optimal control problems by means of the corresponding optimality systems representing the first-order necessary conditions for a minimum and assume that these critical points also satisfy the second-order necessary conditions for (local) minima; for details regarding these conditions in a multigrid context see [6]. For a general and detailed description of multigrid methods see, e.g., [20].

Previous experience shows that multigrid algorithms based on present smoothing strategy can solve optimal control problems with optimal computational complexity and appear to be robust with respect to the choice of values of the optimization parameters; see [3]–[6]. Among others, representative examples demonstrating these facts can be found in [4] for the case of singular optimal control problems and in [5] for the case of linear control-constrained optimal control problems. In the latter case, the resulting multigrid algorithm allowed to investigate bang-bang control and, for the first time, to show the occurrence of chattering of control in an elliptic problem. For more details see [5].

In this paper we discuss an extension of techniques developed in [5] to the case of control-constrained semilinear elliptic optimal control problems and introduce an appropriate setting for the construction of smoothers for state-constrained optimal control problems.

The multigrid approach to optimal control problems is, apart from a few contributions, rather recent. Our framework is the one-shot multigrid strategy as first proposed in [1]. Related approaches can be found in [9] within the successive quadratic programming method, and in [11] where the optimality conditions are reformulated as fixed-point problem and solved by multigrid methods. A one-shot multigrid algorithm means solving the optimality system for the state, the adjoint, and the control variables in parallel in the multigrid process. This is in contrast to solving sequentially for the state and the adjoint equations and then updating the control variables along with the gradients provided by the optimality condition. Notice that this ‘gradient’ approach was also used in [1] to design the smoothers. This strategy has disadvantages: it cannot be applied to singular optimal control problems and, in general, it results in less robust multigrid schemes. The present approach of collectively solve for all three optimization variables overcomes these difficulties.

In the next two sections a control-constrained and a state-constrained optimal control problems are formulated. In the Sections 4 and 5 the construction of appropriate multigrid smoothers for these problems is described in detail. In Section 6, results of numerical experiments are presented. A section of conclusions completes this paper.

2 A control-constrained optimal control problem

Optimal control problems are defined for the purpose of determining the optimal way to influence equilibrium or dynamical systems towards a given task. An optimal control problem consists of a governing system, a description of the control mechanism, and a criterion defining the cost functional, that models the purpose of the control and describes the cost of its action. The formula-

tion of an optimal control problem is then to minimize the cost functional under the constraint given by the modeling equations. The necessary conditions for such a minimum result in a set of coupled equations called the optimality system. In this and in the next section we describe a control- and a state-constrained optimal control problems. For a more general and detailed discussion on optimal control problems see, e.g., [2, 14, 19].

For the purpose of illustration, consider a material plate defining a two-dimensional convex domain Ω . Let the state y of the material represents the temperature distribution which is maintained equal to zero along the boundary. Assume the presence of a given heating source f defined on Ω and thermal radiation or positive temperature feedback due to chemical reactions are represented by the term $G(y)$. This system is governed by the following equation

$$\begin{cases} \Delta y + G(y) = f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega. \end{cases} \quad (1)$$

With the setting above we may think of controlling the temperature distribution y to come close to a given target profile $z \in L^2(\Omega)$ by acting with an additional distributed source term u , the control function. Notice that the given target function needs not be as regular as the state function and it does not need satisfy the boundary conditions for the state.

An optimal control problem governed by (1) is formulated as follows

$$\begin{cases} \min_{u \in U_{ad}} J(y, u), \\ \Delta y + G(y) = u + f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega. \end{cases} \quad (2)$$

In the control-constrained case we assume that $u \in U_{ad}$, $U_{ad} \subset L^2(\Omega)$ being the convex closed set of admissible controls as follows

$$U_{ad} = \{u \in L^2(\Omega) \mid u_L(x) \leq u(x) \leq u_H(x) \text{ a.e. in } \Omega\}, \quad (3)$$

where we choose u_L and u_H elements of $L^2(\Omega)$. Other choices are possible. For a discussion on optimal control problems where the control acts only in a subdomain of Ω ; see [5].

The cost functional J is of the tracking type and is given by

$$J(y, u) = \frac{1}{2} \|y - z\|_{L^2(\Omega)}^2 + \frac{\nu}{2} \|u\|_{L^2(\Omega)}^2, \quad (4)$$

where $\nu \geq 0$ is the weight of the cost of the control.

Existence and uniqueness of solutions to (2)–(4) can be established under suitable conditions for various forms of the nonlinearity; see, e.g., [4, 10, 15]. However, for some interesting applications depending on G (e.g., [4]) multiple (local) optimal solutions may exist. In these cases the multigrid schemes discussed in this paper converge to one of the local minima; see [6].

Optimal solutions are characterized by the following optimality system

$$\begin{aligned} \Delta y + G(y) &= u + f && \text{in } \Omega, \\ y &= 0 && \text{on } \partial\Omega, \\ \Delta p + G'(y)p &= -(y - z) && \text{in } \Omega, \\ p &= 0 && \text{on } \partial\Omega, \\ (\nu u - p, v - u) &\geq 0 && \text{for all } v \in U_{ad}. \end{aligned} \quad (5)$$

We refer to the first differential equation of (5) as the state equation and to the second one as the adjoint equation. The solution to (2)–(4) with $\nu = 0$ is characterized by (5) with the inequality constraint being replaced by $p = \min\{0, p + u - u_L\} + \max\{0, p + u - u_H\}$.

The inequality in (5) gives the optimality condition. It prevents the use of a standard Newton scheme because of the lack of differentiability. In particular, it is not possible to define a smoother based on local Newton steps. We follow an alternative approach where differentiation is not required.

3 A state-constrained optimal control problem

In a state-constrained optimal control problem, bounds are given to the admissible set of values of the state variable. Similarly to the setting above we have

$$\begin{cases} \min_{u \in L^2(\Omega)} J(y, u), \\ \Delta y + G(y) = u + f & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega, \\ y_L \leq y \leq y_H & \text{a.e. in } \Omega, \end{cases}$$

where y_L and y_H are continuous functions, and $J(y, u)$ is given as in (4). Existence and uniqueness of solutions to state-constrained semilinear elliptic optimal control problems depends on the given constraints and on the nonlinearity. For a nonempty solution set, uniqueness can be proved for sufficiently regular G such that the state operator is monotone; see, e.g., [17].

The solution approach to state-constrained optimal control problems through Lagrange multipliers associated with the state constraints leads to difficulties [16]. In particular, the fact that the Lagrange multipliers associated with the state constraints are only regular Borel measures prevents us from using known approximation techniques. In fact, to the best of our knowledge, (at least in a finite difference context) no approximation methods for such a class of functions is available. The remedy is to introduce appropriate regularization; see [16] and references given there. In the following, we consider the Lavrentiev-type approach because it elegantly accommodates in our framework. The Lavrentiev-type regularization consists in approximating the pointwise state constraints $y_L(x) \leq y(x) \leq y_H(x)$ by the following

$$y_L(x) \leq y(x) - \lambda u(x) \leq y_H(x) \quad \text{a.e. in } \Omega,$$

where $\lambda > 0$ is a small parameter. As a result, the associated Lagrange multipliers can be assumed to be functions in $L^2(\Omega)$ [16]. The following regularized state-constrained optimal control problem results

$$\begin{cases} \min_{u \in L^2(\Omega)} J(y, u) := \frac{1}{2} \|y - z\|_{L^2(\Omega)}^2 + \frac{\nu}{2} \|u\|_{L^2(\Omega)}^2, \\ \Delta y + G(y) = u + f, \\ y = 0, \\ y_L \leq y - \lambda u \leq y_H. \end{cases}$$

Now, introduce the auxiliary variable $v = y - \lambda u$ and express the control function u in terms of v . The regularized

state-constrained optimal control problem becomes

$$\begin{cases} \min_{v \in L^2(\Omega)} J(y, v) := \frac{1}{2} \|y - z\|_{L^2(\Omega)}^2 + \frac{\nu}{2\lambda^2} \|y - v\|_{L^2(\Omega)}^2 & u_k^{(l)} = S_k(u_k^{(l-1)}, f_k), \quad l = m_1+2, \dots, m_1+m_2+1; \\ \Delta y + G(y) - y/\lambda + v/\lambda = f, \\ y = 0, \\ y_L \leq v \leq y_H. \end{cases}$$

Notice that after the transformation, an optimal control problem is obtained having a ‘control-constrained’ structure. The solution to this problem is characterized by the following optimality system

$$\begin{aligned} \Delta y + G(y) - y/\lambda + v/\lambda &= f & \text{in } \Omega, \\ y &= 0 & \text{on } \partial\Omega, \\ \Delta p + G'(y)p - p/\lambda + (y - z) + \gamma(y - v) &= 0 & \text{in } \Omega, \\ p &= 0 & \text{on } \partial\Omega, \\ (p/\lambda - \gamma(y - v), t - v) &\geq 0, \end{aligned} \quad (6)$$

where $\gamma = \nu/\lambda^2$ and the inequality must hold for all $t \in V_{ad}$, and V_{ad} is defined by

$$V_{ad} = \{v \in L^2(\Omega) \mid y_L(x) \leq v(x) \leq y_H(x) \text{ a.e. in } \Omega\}.$$

4 A smoother for control-constrained optimal control problems

For illustration, we focus on finite difference approximations, however the present approach can be extended to other discretization techniques. Assume that Ω is a rectangular domain and consider a sequence of grids indexed by $k = 1, \dots, L$, $\{\Omega_{h_k}\}_k$ with mesh sizes h_k such that the boundaries of Ω coincide with grid lines. Here $k = L$ denotes the finest level and h_1 is the mesh size of the coarsest grid. Correspondingly, with k we also index the operators and variables defined on $\Omega_k = \Omega_{h_k}$.

The multigrid smoothing schemes presented here are used within the full approximation storage (FAS) [7] method. This is a natural choice in the treatment of nonlinear and constrained problems since the optimization variables and not their errors are represented at all levels. We can define the multigrid FAS scheme for a generic nonlinear problem $A_k(u_k) = f_k$, where $A_k(\cdot)$ represents a nonlinear discrete operator on Ω_k , as follows.

– Multigrid FAS- (m_1, m_2) method.

- (i) If $k = 1$ solve $A_k(u_k) = f_k$ directly (e.g., repeated application of S_k).
- (ii) Pre-smoothing steps on the fine grid:

$$u_k^{(l)} = S_k(u_k^{(l-1)}, f_k), \quad l = 1, \dots, m_1;$$

- (iii) Computation of the residual: $r_k = f_k - A_k u_k^{(m_1)}$;
- (iv) Restriction of the residual: $r_{k-1} = I_{k-1}^{k-1} r_k$;
- (v) Set $u_{k-1} = \tilde{I}_{k-1}^{k-1} u_k^{(m_1)}$;
- (vi) Set $f_{k-1} = r_{k-1} + A_{k-1}(u_{k-1})$;
- (vii) Call m times FAS- (m_1, m_2) to solve

$$A_{k-1}(u_{k-1}) = f_{k-1};$$

- (viii) Coarse-grid correction:

$$u_k^{(m_1+1)} = u_k^{(m_1)} + I_{k-1}^k(u_{k-1} - \tilde{I}_{k-1}^{k-1} u_k^{(m_1)});$$

- (ix) Post-smoothing steps on the fine grid:

The variational inequality defining the optimality condition has the same form on all grids.

We choose I_{k-1}^k to be bilinear interpolation and for \tilde{I}_{k-1}^{k-1} and I_k^{k-1} we take injection and full-weighting restriction, respectively. These operators are given in stencil form as follows

$$I_{k-1}^k = \frac{1}{4} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad \text{and} \quad I_k^{k-1} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \quad (7)$$

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 [20]. Our numerical experience shows that in case of constrained control problems the use of W -cycles results in a more robust multigrid iteration.

In the following, Δ_{h_k} denotes the discretization of Laplacian by any finite differences stencil satisfying a maximum principle. Homogeneous Dirichlet boundary conditions are included in the definition of Δ_{h_k} . For a generic h_k we use h instead.

We now give a detailed derivation of a smoothing scheme for (5). Let $h = h_k$ and $x \in \Omega_h$, where $x = (ih, jh)$ and (i, j) index the grid points, e.g., lexicographically. Denote with ω_{ij} the set of grid index pairs s, t of the stencil of Δ_h centered at (i, j) . Correspondingly, denote with c_{st} , $s, t \in \omega_{ij}$, the s, t entry of the stencil of the discretized Laplacian (multiplied by h^2) centered at (i, j) . Using this notation, we can express the action of Δ_h on the function v_h in the following compact form

$$\Delta_h v_h|_{ij} = \frac{1}{h^2} \left(\sum_{s, t \in \omega_{ij}, s, t \neq i, j} c_{st} v_{st} - c_{ij} v_{ij} \right),$$

where $c_{ij} > 0$.

Consider the discretization of (5) at x , we have

$$\sum_{\substack{s, t \in \omega_{ij} \\ s, t \neq i, j}} c_{st} y_{st} - c_{ij} y_{ij} + h^2 G(y_{ij}) - h^2 u_{ij} = h^2 \tilde{f}_{ij} \quad (8)$$

$$\sum_{\substack{s, t \in \omega_{ij} \\ s, t \neq i, j}} c_{st} p_{st} - c_{ij} p_{ij} + h^2 G'(y_{ij}) p_{ij} + h^2 y_{ij} = h^2 \tilde{z}_{ij} \quad (9)$$

$$(\nu u_{ij} - p_{ij}) \cdot (v_{ij} - u_{ij}) \geq 0 \text{ for all } v \in U_{adh}, \quad (10)$$

where $U_{adh} = \{u \in L_h^2(\Omega_h) \mid u_L(\mathbf{x}) \leq u(\mathbf{x}) \leq u_H(\mathbf{x}) \text{ in } \Omega_h\}$ and \tilde{f} and \tilde{z} represent f_h and z_h including the defect corrections resulting by coarsening in nonlinear multigrid schemes [20]. We set

$$A_{ij} = \sum_{s, t \in \omega_{ij}, s, t \neq i, j} c_{st} y_{st} - h^2 \tilde{f}_{ij}$$

and

$$B_{ij} = \sum_{s, t \in \omega_{ij}, s, t \neq i, j} c_{st} p_{st} - h^2 \tilde{z}_{ij}.$$

Here A_{ij} and B_{ij} are considered constant during the update of the variables at (i, j) . Summarizing, we have the following system at (i, j) for the three scalar variables y_{ij} , p_{ij} , and u_{ij} :

$$A_{ij} - c_{ij} y_{ij} + h^2 G(y_{ij}) - h^2 u_{ij} = 0, \quad (11)$$

$$B_{ij} - c_{ij} p_{ij} + h^2 G'(y_{ij}) p_{ij} + h^2 y_{ij} = 0, \quad (12)$$

$$(\nu u_{ij} - p_{ij}) \cdot (v_{ij} - u_{ij}) \geq 0 \quad (13)$$

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

After discretization, it is required that the inequality constraint holds pointwise. This is an equivalent but more convenient discretization than expressing the inequality in terms of the discrete L^2 product.

Next, we discuss a Newton-based step which generalizes ideas in [4, 5] and avoids differentiation of (13). Consider the inverse of the Jacobian of (11)–(12) with respect to the variables y_{ij} , p_{ij} , it is given by

$$J_{ij}^{-1} = \frac{1}{\det J_{ij}} \begin{pmatrix} -c_{ij} + h^2 G'(y_{ij}) & 0 \\ -h^2(1 + G''(y_{ij}) p_{ij}) & -c_{ij} + h^2 G'(y_{ij}) \end{pmatrix}$$

where $\det J_{ij} = (-c_{ij} + h^2 G'(y_{ij}))(-c_{ij} + h^2 G'(y_{ij}))$. By computing the reduced Hessian (see, e.g., [6]) one finds that second-order necessary conditions for a minimum require $(1 + G''(y_{ij}) p_{ij}) \geq 0$.

For given u_{ij} , a local Newton update for the state and the adjoint variables \hat{y}_{ij} and \hat{p}_{ij} at (i, j) is given by

$$\begin{pmatrix} \hat{y}_{ij} \\ \hat{p}_{ij} \end{pmatrix} = \begin{pmatrix} y_{ij} \\ p_{ij} \end{pmatrix} + J_{ij}^{-1} \begin{pmatrix} r_{ij}(u_{ij}) \\ s_{ij} \end{pmatrix}, \quad (14)$$

where r_{ij} and s_{ij} denote the residuals of (11) and (12), respectively. Here $r_{ij}(u_{ij}) = -(A_{ij} - c_{ij} y_{ij} + h^2 G(y_{ij}) - h^2 u_{ij})$ depends explicitly on u_{ij} . This fact allows us to write \hat{p}_{ij} as a function of u_{ij} as follows

$$\begin{aligned} \hat{p}_{ij}(u_{ij}) &= p_{ij} \\ &- \frac{1}{\det J_{ij}} \{ h^2(1 + G''(y_{ij}) p_{ij}) r_{ij}(u_{ij}) \\ &+ (c_{ij} - h^2 G'(y_{ij})) s_{ij} \} \end{aligned} \quad (15)$$

Now in order to obtain the update for u_{ij} , replace \hat{p}_{ij} in the inequality constraint (13). Define the reduced cost functional $\hat{J}(u) = J(y(u), u)$ and recall that the optimal control solution to (2)–(4) is characterized by $\hat{J}'(u, v - u) = (\nu u - p(u), v - u) \geq 0$ for all $v \in U_{ad}$. In the absence of constraints the optimal control satisfies $\nu u - p(u) = 0$. Therefore we consider the solution to $\nu \tilde{u} - \hat{p}(\tilde{u}) = 0$ at (i, j) with $\hat{p}(u)$ given by (15). This provides the auxiliary variable (that represents the updated control without constraints)

$$\begin{aligned} \tilde{u}_{ij} &= (\nu + \frac{(1 + G''(y_{ij}) p_{ij}) h^4}{\det J_{ij}})^{-1} \times \\ &\{ p_{ij} + \frac{1}{\det J_{ij}} (h^2(1 + G''(y_{ij}) p_{ij}) \times \\ &(A_{ij} - c_{ij} y_{ij} + h^2 G(y_{ij})) \\ &+ \frac{1}{\det J_{ij}} ((c_{ij} - h^2 G'(y_{ij})) \times \\ &(B_{ij} - c_{ij} p_{ij} + h^2 G'(y_{ij}) p_{ij} + h^2 y_{ij})) \} \end{aligned}$$

In the presence of constraints, the new value for u_{ij} resulting from the relaxation step is obtained by projection of \tilde{u}_{ij} onto U_{adh} as follows

$$u_{ij} = \begin{cases} u_{Hij} & \text{if } \tilde{u}_{ij} \geq u_{Hij} \\ \tilde{u}_{ij} & \text{if } u_{Lij} < \tilde{u}_{ij} < u_{Hij} \\ u_{Lij} & \text{if } \tilde{u}_{ij} \leq u_{Lij}. \end{cases} \quad (16)$$

With u_{ij} given by (16), we use (14) to update y_{ij} and p_{ij} .

This completes the description of the smoothing step. One can verify that it satisfies the inequality constraints. In fact, assume that $\tilde{u}_{ij} > u_{Hij}$; then from (16) we have $u_{ij} = u_{Hij}$. Therefore $(v_{ij} - u_{ij}) \leq 0$ for any $v \in U_{adh}$. By inspection of (15) we see that $\hat{p}_{ij}(\tilde{u}_{ij}) < \hat{p}_{ij}(u_{ij})$. On the other hand we have

$$\nu u_{ij} - \hat{p}_{ij}(u_{ij}) < \nu \tilde{u}_{ij} - \hat{p}_{ij}(\tilde{u}_{ij}) = 0.$$

Similarly, if $\tilde{u}_{ij} < u_{Lij}$ we take $u_{ij} = u_{Lij}$ and we have that $\nu u_{ij} - \hat{p}_{ij}(u_{ij}) > 0$ together with $(v_{ij} - u_{ij}) \geq 0$ for any $v \in U_{adh}$. Therefore $(\nu u - p) \cdot (v - u) \geq 0$ for all $v \in U_{adh}$. In the case $\nu = 0$ similar arguments hold.

Because of the ‘projection’ (16), the present iteration can be considered as belonging to the class of projected Gauss-Seidel schemes [8].

5 A smoother for state-constrained optimal control problems

Next, the construction of the smoother for the regularized state-constrained optimality system given by (6) is discussed. Consider the following discretized state and adjoint equations

$$\begin{aligned} \sum_{s, t \in \omega_{ij}, s, t \neq i, j} c_{st} y_{st} - (c_{ij} + h^2/\lambda) y_{ij} + h^2 G(y_{ij}) \\ + (h^2/\lambda) v_{ij} &= h^2 \tilde{f}_{ij}, \\ \sum_{s, t \in \omega_{ij}, s, t \neq i, j} c_{st} p_{st} - (c_{ij} + h^2/\lambda) p_{ij} + h^2 G'(y_{ij}) p_{ij} \\ + (1 + \gamma) h^2 y_{ij} - \gamma h^2 v_{ij} &= h^2 \tilde{z}_{ij}. \end{aligned}$$

These two equations are written in a compact form as follows.

$$A_{ij} - \alpha_{ij} y_{ij} + h^2 G(y_{ij}) + (h^2/\lambda) v_{ij} = 0, \quad (17)$$

$$B_{ij} - \alpha_{ij} p_{ij} + h^2 G'(y_{ij}) p_{ij} + (1 + \gamma) h^2 y_{ij} - \gamma h^2 v_{ij} = 0, \quad (18)$$

where $\alpha_{ij} = c_{ij} + h^2/\lambda$. In addition, we have the inequality

$$(p_{ij}/\lambda - \gamma y_{ij} + \gamma v_{ij}) \cdot (t_{ij} - v_{ij}) \geq 0$$

for all $t \in V_{adh} = \{v \in L_h^2(\Omega_h) \mid y_L \leq v \leq y_H \text{ a.e. in } \Omega_h\}$.

Now consider the Jacobian of the system (17)–(18) with respect to y_{ij} , p_{ij} , that is,

$$J_{ij} = \begin{pmatrix} -\alpha_{ij} + h^2 G'(y_{ij}) & 0 \\ (1 + \gamma) h^2 + h^2 G''(y_{ij}) p_{ij} & -\alpha_{ij} + h^2 G'(y_{ij}) \end{pmatrix}.$$

Hence, the following local Newton update for y_{ij} and p_{ij} at (i, j) results

$$\begin{pmatrix} \hat{y}_{ij} \\ \hat{p}_{ij} \end{pmatrix} = \begin{pmatrix} y_{ij} \\ p_{ij} \end{pmatrix} + J_{ij}^{-1} \begin{pmatrix} r_{ij}(v_{ij}) \\ s_{ij}(v_{ij}) \end{pmatrix}, \quad (19)$$

where $r_{ij}(v_{ij})$ and $s_{ij}(v_{ij})$ denote the residuals of (17) and (18), respectively. Both residuals depend explicitly on the control variable v_{ij} . Therefore the update above defines the values $\hat{y}_{ij}(v_{ij})$ and $\hat{p}_{ij}(v_{ij})$ as functions of v_{ij} .

Similarly to the control-constrained case, we now denote with \tilde{v}_{ij} the solution to the linear equation

$$p_{ij}(v_{ij})/\lambda - \gamma y_{ij}(v_{ij}) + \gamma v_{ij} = 0.$$

It is given by $\tilde{v}_{ij} = N_{ij}/D_{ij}$ where

$$\begin{aligned} N_{ij} = & -(\lambda(\alpha_{ij}B_{ij} + (1 + \gamma)A_{ij}h^2 - A_{ij}\alpha_{ij}\gamma\lambda \\ & - h^2(B_{ij} - A_{ij}\gamma\lambda + ((1 + \gamma)h^2 - \alpha_{ij}\gamma\lambda)y_{ij})G'(y_{ij}) \\ & - h^4\gamma\lambda y_{ij}G'(y_{ij})^2 + h^2A_{ij}p_{ij}G''(y_{ij}) \\ & - \alpha_{ij}h^2p_{ij}y_{ij}G''(y_{ij}) + h^2G(y_{ij})(h^2(1 + \gamma) \\ & - \alpha_{ij}\gamma\lambda + h^2\gamma\lambda G'(y_{ij}) + h^2p_{ij}G''(y_{ij})))) \end{aligned}$$

and

$$\begin{aligned} D_{ij} = & h^4(1 + \gamma) - 2\alpha_{ij}h^2\gamma\lambda + \alpha_{ij}^2\gamma\lambda^2 \\ & + 2h^2\gamma\lambda(h^2 - \alpha_{ij}\lambda)G'(y_{ij}) \\ & + h^4\gamma\lambda^2G'(y_{ij})^2 + h^4p_{ij}G''(y_{ij}). \end{aligned}$$

Now recall that the update to v_{ij} must satisfy the constraints $y_L(x_{ij}) \leq v_{ij} \leq y_H(x_{ij})$. Therefore a feasible update is given by

$$v_{ij} = \begin{cases} y_{Hij} & \text{if } \tilde{v}_{ij} \geq y_{Hij} \\ \tilde{v}_{ij} & \text{if } y_{Lij} < \tilde{v}_{ij} < y_{Hij} \\ y_{Lij} & \text{if } \tilde{v}_{ij} \leq y_{Lij}. \end{cases} \quad (20)$$

Updates for the adjoint and state variables are then given by $p_{ij} = \hat{p}_{ij}(v_{ij})$ and $y_{ij} = \hat{y}_{ij}(v_{ij})$, respectively, as defined by (19).

In the linear unconstrained case, local Fourier analysis results show that the smoothing schemes (14)–(16) and (19)–(20) have typical smoothing factor values ($\mu \approx 0.5$) and this value is weakly dependent on ν . This property is especially important when solving for very small choices of ν where other methods may fail.

6 Numerical experiments

We report results of numerical experiments to discuss the computational performance of the smoothing schemes (14)–(16) and (19)–(20) by considering the convergence behavior of resulting multigrid algorithms. We use W-FAS-cycles [7] with one pre- and post-smoothing of the symmetric (i.e. a forward and a backward sweeps) version of the smoothers described above with lexicographic ordering (for details regarding coarsening of the variational inequality see [5] and [18]). W-cycles appear to be superior in the case of active constraints [5]; for the present isotropic problems ordering is not essential for convergence. In the figures we plot the log of the sum

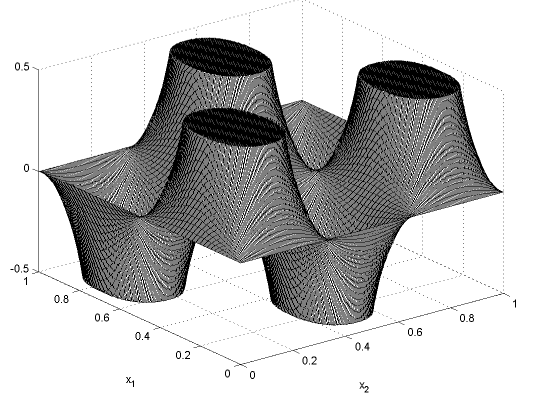


Fig. 1. Control-constrained case. The optimal control for $\nu = 10^{-7}$.

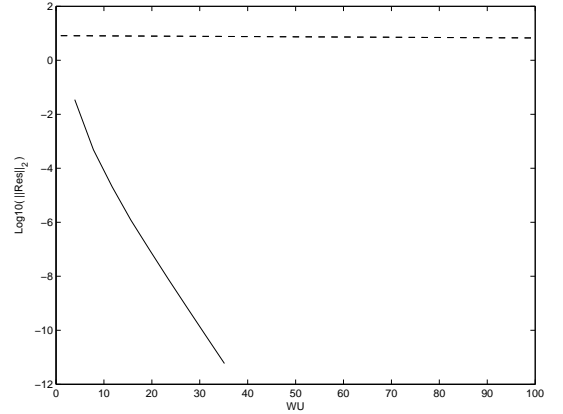


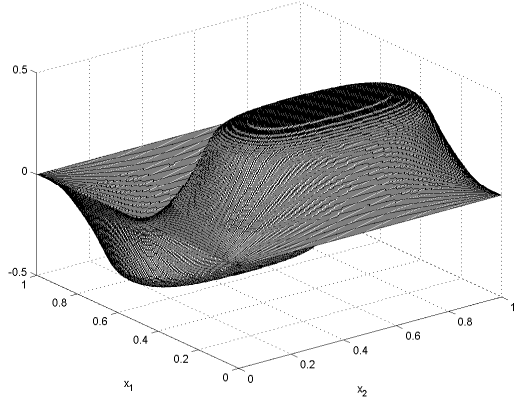
Fig. 2. Control-constrained case. Convergence history for smoothing only (dashed line) and multigrid W(1,1)-cycle; $\nu = 10^{-7}$.

of L^2 -norm of the state and adjoint equation residuals. Notice that similar results are obtained with different meshes. Let $\Omega = (0, 1)^2$ and the finest mesh is 257×257 and six levels are employed. All unknown variables are initialized to be zero and we choose $G(y) = -y^4$ (see [15]) and $f = 0$.

First we consider the following control-constrained problem. Let $z(x_1, x_2) = \sin(2\pi x_1)\sin(3\pi x_2)$ and $\nu = 10^{-7}$. With this setting and $u_L(x) = -1/2$ and $u_H(x) = 1/2$ the constraints are active in large part of the domain. This fact can be seen in Figure 6; in Figure 6 we observe typical slow convergence of the smoothing iteration and ‘textbook’ multigrid convergence of the multigrid cycle. Similar results are obtained with different choices of values of the optimization parameter.

Table 1. Convergence factors choosing $\nu = \lambda^2$.

mesh	$\lambda = 10^{-3}$	$\lambda = 10^{-4}$	$\lambda = 10^{-5}$
257×257	0.06	0.06	0.07
513×513	0.07	0.07	0.08
1025×1025	0.07	0.07	0.07

**Fig. 3.** State-constrained case. The optimal state for $\nu = 10^{-7}$ and $\lambda = 10^{-3}$.

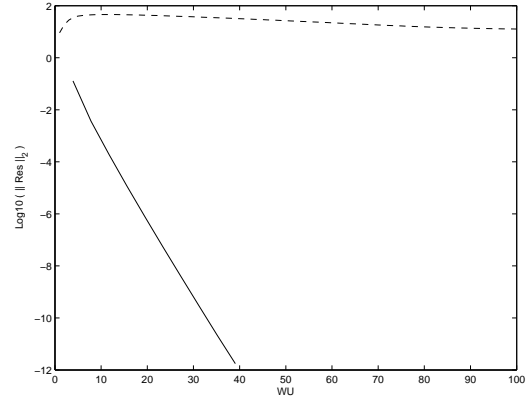
Next, consider the following state-constrained optimal control problem. Take $z(x_1, x_2) = \sin(2\pi x_1) \sin(\pi x_2)$ and $y_L(x) = -1/2$ and $y_H(x) = 1/2$. Results for the choice $\nu = 10^{-7}$ and $\lambda = 10^{-3}$ are reported in Figures 6 & 6. In Figure 6, the constrained state solution is depicted. Convergence history is reported in Figure 6. We notice an increase of the value of the sum of the L^2 -norm of the residuals during the first few iterations of the smoothing scheme. This behavior results from an increase of the residual of the state equation, while the residuals of the adjoint and control equations decrease monotonically. On the other hand, we observe typical convergence behavior of the multigrid scheme based on the proposed smoother.

The results reported in Table 1 show typical multigrid convergence factors that are mesh independent. These values are obtained choosing $\nu \approx \lambda^2$. With ν held fixed and decreasing λ the resulting convergence factors worsen.

7 Conclusion

An unified approach to the design of multigrid smoothers for constrained optimal control problems was presented. This approach defines collective smoothing steps that combine local Newton update with projection through the gradient of the reduced cost functional to overcome the lack of differentiability due to the presence of inequality constraints.

Results of numerical experiments were presented to show that the resulting multigrid schemes provide typical multigrid computational efficiency in the case of

**Fig. 4.** State-constrained case. Convergence history for smoothing only (dashed line) and multigrid W(1,1)-cycle; $\nu = 10^{-7}$ and $\lambda = 10^{-3}$.

control-constrained optimal control problems. In the state-constrained case, multigrid efficiency is obtained whenever the value of the weight of the cost of the control ν is of the same order as λ^2 . Deterioration of convergence behavior may occur when the Lavrentiev regularization parameter λ becomes too small with respect to ν .

It can be expected that improved robustness with respect to λ might be achieved when including appropriate multigrid techniques for variational inequalities as given in, e.g., [13, 18, 21].

References

1. 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.
2. A. BORZÌ, *Multigrid Methods for Optimality Systems*, habilitation thesis, University of Graz, 2003. (<http://www.uni-graz.at/imawww/borzi/publist.html>)
3. A. BORZÌ AND G. BORZÌ, *An efficient algebraic multigrid method for solving optimality systems*, Comp. Visual. Sci., 7(3/4) (2004), pp. 183–188.
4. A. BORZÌ AND K. KUNISCH, *The numerical solution of the steady state solid fuel ignition model and its optimal control*, SIAM J.Sci.Comp., 22 (2000), pp. 263–284.
5. A. BORZÌ AND K. KUNISCH, *A multigrid scheme for elliptic constrained optimal control problems*, Comput. Optim. Appl., 31 (2005), pp. 309–333.
6. A. BORZÌ AND K. KUNISCH, *A globalization strategy for the multigrid solution of elliptic optimal control problems*, Optimization Methods and Software, 21(3) (2006), pp. 445–459.
7. A. BRANDT, *Multi-level adaptive solutions to boundary-value problems*, Mathematics of Computation, 31 (1977), pp. 333–390.

8. 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., 4 (1983), pp. 655–684.
9. TH. DREYER, B. MAAR, V. SCHULZ, *Multigrid Optimization in Applications*, J. Comput. Appl. Math., 120 (2000), pp. 67–84.
10. H. GOLDBERG AND F. TRÖLTZSCH, *Second order sufficient optimality conditions for a class of non-linear parabolic boundary control problems*, SIAM J. Control Optim., 31 (1993), pp. 1007–1027.
11. W. HACKBUSCH, *Fast solution of elliptic control problems*, Journal of Optimization Theory and Application, 31 (1980), pp. 565–581.
12. W. HACKBUSCH, *Multi-grid Methods and Applications*, Springer-Verlag, New York, 1985.
13. R. KORNUBER AND R. KRAUSE, *Adaptive multigrid methods for Signorini's problem in linear elasticity*, Comp. Visual. Sci., 4 (2001), pp. 9–20.
14. J.L. LIONS, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin, 1971.
15. J.L. LIONS, *Control of Distributed Singular Systems*, Gauthier-Villars, Paris, 1985.
16. C. MEYER, A. RÖSCH AND F. TRÖLTZSCH, *Optimal control of PDEs with regularized pointwise state constraints*. TU Berlin, Preprint 2003/14. To appear in Computational Optimization and Applications.
17. C. MEYER, A. RÖSCH AND F. TRÖLTZSCH, *On an elliptic optimal control problem with pointwise mixed control-state constraints*, Recent Advances in Optimization. Proceedings of the 12th French-German-Spanish Conference on Optimization held in Avignon, September 20-24, 2004, Edited by A. Seeger. Lectures Notes in Economics and Mathematical Systems, Vol. 563, Springer-Verlag Berlin Heidelberg, 2006, pp. 187–204.
18. C. REISINGER AND G. WITTUM, *On Multigrid for anisotropic equations and variational inequalities “Pricing multi-dimensional european and american options”*, Comp. Visual. Sci., 4 (2004), pp. 189–197.
19. F. TRÖLTZSCH, *Optimale Steuerung Partieller Differentialgleichung*, Vieweg Verlag, 2005.
20. U. TROTTEBERG, C. OOSTERLEE, AND A. SCHÜLLER, *Multigrid*, Academic Press, London, 2001.
21. C.H. VENNER AND A.A. LUBRECHT, *Multilevel Methods in Lubrication*, Elsevier, Amsterdam, 1993.