

MARTIN WEISER

TOBIAS GÄNZLER

ANTON SCHIELA

A Control Reduced Primal Interior Point Method for PDE Constrained Optimization¹

¹Supported by the DFG Research Center "Mathematics for key technologies" (Matheon) in Berlin.

A Control Reduced Primal Interior Point Method for PDE Constrained Optimization[†]

Martin Weiser Tobias Gänzler Anton Schiela

September 15, 2004

Abstract

A primal interior point method for control constrained optimal control problems with PDE constraints is considered. Pointwise elimination of the control leads to a homotopy in the remaining state and dual variables, which is addressed by a short step pathfollowing method. The algorithm is applied to the continuous, infinite dimensional problem, where discretization is performed only in the innermost loop when solving linear equations. The a priori elimination of the least regular control permits to obtain the required accuracy with comparatively coarse meshes. Convergence of the method and discretization errors are studied, and the method is illustrated at two numerical examples.

AMS MSC 2000: 90C51, 65N15, 49M15

Keywords: interior point methods in function space, optimal control, finite elements, discretization error

1 Introduction

Optimal control problems with PDE constraints are not only very important in practical applications, but also very expensive to solve numerically. Independently of where discretization is located in an algorithm, at the outermost loop as in direct methods or at the innermost loop as in function space oriented methods, the resulting finite dimensional equations are quite large. This is particularly true if highly local features of the solution need to be represented accurately, which requires fine meshes and thus leads to large finite dimensional subproblems.

In control constrained optimization problems such local features arise at the boundaries of active sets, where the control exhibits kinks or even jumps in bang-bang control problems. Unfortunately, the boundaries of active sets are in general not grid-aligned. This leads to error estimates of $\mathcal{O}(h)$ for piecewise constant and $\mathcal{O}(h^{3/2})$ for piecewise linear control discretizations [1, 3, 11]. For the faithful representation of an approximate solution up to a requested accuracy, massive mesh

[†]Supported by the DFG Research Center "Mathematics for key technologies" (Matheon) in Berlin.

refinement along the boundaries of active sets is necessary — see e.g. [14] and Fig. 5. In certain situations, the need for refinement can be alleviated by a special postprocessing procedure [8], which results in an approximation order of $\mathcal{O}(h^2)$.

As a different approach to alleviate the need for mesh refinement, HINZE [7] suggests to eliminate the control u analytically from the optimality system by computing it from the dual variable λ . For a certain class of optimal control problems, this is a simple pointwise calculation. The resulting variables that have to be discretized, the state y and dual variable λ , are comparatively smooth across boundaries of active sets. For this reason, a rather coarse mesh is sufficient, and a control error of order $\mathcal{O}(h^2)$ is obtained. Due to the elimination of u , however, the resulting equation system is no longer smooth.

Both primal and primal-dual adaptive function space oriented interior-point methods for optimal control problems have been developed, analyzed and applied in [13, 12, 14, 10]. They rely on approximating the control in a finite dimensional space and hence need to refine the mesh substantially, which impedes their computational efficiency. The abovementioned idea of eliminating the control can also be applied to primal interior point methods, which is addressed in this paper. Compared to the interior point methods above, the necessity of mesh refinement is decreased to a similar amount as reported in [7]. As is characteristic for interior point methods, instead of a nonsmooth equation system, a homotopy of smooth equation systems results.

The remainder of the paper is organized as follows: The elimination of the control from the optimality system and its interior point regularization is described in Section 2, where also convergence of the central path is stated. Section 3 is devoted to linear convergence of a short step pathfollowing method, whereas discretization error estimates are given in Section 4. Finally, numerical examples illustrate the method in Section 5.

2 Elimination of controls

For ease of presentation we restrict the discussion to the simple elliptic optimal control problem

$$\min_{y \in H_0^1, u \in L^2} \frac{1}{2} \|y - y_d\|_{L^2}^2 + \frac{\alpha}{2} \|u\|_{L^2}^2 \quad \text{subject to} \quad Ly = u, \quad -1 \leq u \leq 1 \quad (1)$$

on some domain $\Omega \subset \mathbb{R}^d$. $y_d \in L^2$ is the desired state, and $\alpha > 0$ is a fixed regularization parameter. $Ly = -\operatorname{div}(a(x)\nabla y) + b(x)y$ with symmetric $a(x) \in \mathbb{R}^{2 \times 2}$ uniformly positive definite and $b(x) \in \mathbb{R}$ uniformly positive is a second order elliptic differential operator. With $S : H^{-1} \rightarrow H_0^1$ we denote the symmetric positive definite solution operator for the state equation.

We do emphasize, however, that the following discretization concept can be directly extended to more complex and nonlinear control constrained problems as

long as (i) the control constraints are defined pointwise and (ii) the derivative of the objective function w.r.t. the control is an invertible Nemyckii operator of u .

For problem (1) the first order necessary conditions state the existence of Lagrange multipliers $\lambda \in H_0^1$ and $\underline{\eta}, \bar{\eta} \in L^2$, such that

$$\begin{aligned} y - y_d + L\lambda &= 0 \\ \alpha u - \lambda - \underline{\eta} + \bar{\eta} &= 0 \end{aligned} \tag{2}$$

$$Ly - u = 0 \tag{3}$$

$$\langle \underline{\eta}, u + 1 \rangle = \langle \bar{\eta}, 1 - u \rangle = 0 \tag{4}$$

$$u + 1, 1 - u, \underline{\eta}, \bar{\eta} \geq 0. \tag{5}$$

The approach of [7] is to use (2), (4), and (5) in order to eliminate

$$u = u(\lambda) = \max(-1, \min(\lambda/\alpha, 1)), \tag{6}$$

which results in the nonsmooth system

$$\begin{aligned} y - y_d + L\lambda &= 0 \\ Ly - u(\lambda) &= 0. \end{aligned}$$

Notice that the expensive to discretize control u does no longer appear as iteration variable, and actually need not to be discretized. Equivalently, the nonsmooth system may be reformulated as

$$u = u(S(y_d - Su)),$$

where the iteration variable u can be discretized implicitly in terms of discrete approximate solutions of $S(y_d - Su)$.

On the other hand, primal-dual interior point methods substitute (4) by the regularized equations

$$\underline{\eta}(u + 1) = \bar{\eta}(1 - u) = \mu \tag{7}$$

for $\mu > 0$ and thus define the central path $\mu \mapsto (y, u, \lambda, \underline{\eta}, \bar{\eta})$. This approach has been analyzed and justified in [12, 14]. Using (7) to eliminate the Lagrange multipliers $\underline{\eta}$ and $\bar{\eta}$ results in the primal interior point method given by

$$\begin{aligned} y - y_d + L\lambda &= 0 \\ \alpha u - \lambda - \frac{\mu}{u + 1} + \frac{\mu}{1 - u} &= 0 \end{aligned} \tag{8}$$

$$Ly - u = 0$$

$$u + 1, 1 - u \geq 0. \tag{9}$$

These are just the first order necessary conditions for the logarithmic barrier reformulation of (1),

$$\min \frac{1}{2} \|y - y_d\|_{L^2}^2 + \frac{\alpha}{2} \|u\|_{L^2}^2 + \mu \|\log(u + 1) + \log(1 - u)\|_{L^1} \quad \text{subject to} \quad Ly = u.$$

Existence and convergence of the central path defined by primal interior point methods for control constrained problems has been established in [10], along with convergence of a function space oriented pathfollowing method.

Again, we can use (8) and (9) in order to eliminate $u = u(\lambda)$. Instead of the L^2 -projection (6) we need to solve a scalar cubic equation in every point in space. As before, we obtain the equation system

$$\begin{aligned} y - y_d + L\lambda &= 0 \\ Ly - u(\lambda; \mu) &= 0, \end{aligned}$$

which is, however, smooth. Its unique solvability is a consequence of the following Lemma.

Lemma 2.1. *There is exactly one $u(\lambda) \in]-1, 1[$ that satisfies (6). Moreover, $u(\lambda)$ is twice continuously differentiable with respect to λ , and*

$$0 \leq u'(\lambda) \leq \frac{1}{\alpha + 2\mu} \quad \text{and} \quad |u''(\lambda)| \leq \min\left(\frac{1}{\sqrt{\mu\alpha^3}}, \frac{8}{\mu^2}\right).$$

Proof. The left hand side of (6) as a function of $u \in]-1, 1[$ is monotonically increasing and tends to $\pm\infty$ for $u \rightarrow \pm 1$. Therefore, equation (8) has exactly one root $u(\lambda)$, which is well defined.

As for the derivatives, we apply the implicit function theorem to (6) and obtain

$$\left(\alpha + \frac{\mu}{(u(\lambda) + 1)^2} + \frac{\mu}{(1 - u(\lambda))^2}\right) u'(\lambda) - 1 = 0,$$

which gives the first order bound. A second application of the implicit function theorem gives

$$\begin{aligned} \left(-\frac{\mu}{(u(\lambda) + 1)^3} + \frac{\mu}{(1 - u(\lambda))^3}\right) u'(\lambda)^2 \\ + \left(\alpha + \frac{\mu}{(u(\lambda) + 1)^2} + \frac{\mu}{(1 - u(\lambda))^2}\right) u''(\lambda) = 0 \end{aligned}$$

which yields

$$u''(\lambda) = \mu \frac{(u(\lambda) + 1)^{-3} - (1 - u(\lambda))^{-3}}{\left(\alpha + \frac{\mu}{(u(\lambda) + 1)^2} + \frac{\mu}{(1 - u(\lambda))^2}\right)^3}.$$

For the case $u(\lambda) \leq 0$ we infer

$$\begin{aligned} |u''(\lambda)| &\leq \mu \frac{(u(\lambda) + 1)^{-3}}{\left(\alpha + \frac{\mu}{(u(\lambda) + 1)^2} + \frac{\mu}{(1 - u(\lambda))^2}\right)^3} \\ &\leq \mu \left(\alpha(u(\lambda) + 1) + \frac{\mu}{u(\lambda) + 1}\right)^{-3} \\ &\leq \frac{\mu}{\max(2\sqrt{\alpha\mu}, \mu/2)^3} = \min\left(\frac{1}{\sqrt{\mu\alpha^3}}, \frac{8}{\mu^2}\right). \end{aligned}$$

Due to symmetry, this bound is also valid for $u(\lambda) > 0$. \square

Remark 2.2. In actual computation, a numerically stable evaluation of $u(\lambda)$ is of utmost importance for certain ranges of α , λ , and μ .

We define $v = (y, \lambda)$ and the homotopy

$$F(v; \mu) = \begin{bmatrix} y - y_d + L\lambda \\ Ly - u(\lambda; \mu) \end{bmatrix} = 0, \quad (10)$$

which in turn defines the central path $v(\mu)$.

Theorem 2.3. *For each $\mu > 0$ there is a corresponding unique $v(\mu) \in H_0^1 \times H_0^1$ satisfying (10). $v(\mu)$ is a continuously differentiable path with $v'(\mu) \leq \frac{c}{\sqrt{\mu}}$ for some generic constant c independent of μ . Moreover, $v(\mu)$ converges to the solution $v(0)$ of (1) at a rate of*

$$\|v(\mu) - v(0)\|_{H^1} \leq c\sqrt{\mu}. \quad (11)$$

Proof. This result is a direct consequence of [10]. Alternatively, existence of central path solutions for $\mu > 0$ can be shown directly by applying Schauder's fixed point theorem to $u = u(S(Su - y_d))$. Continuous differentiability of the path results from the implicit function theorem. The necessary invertibility of $\partial_v F(v; \mu)$ is shown in Lemma 3.2. Integrating the slope of the central path yields the estimate (11). \square

3 A pathfollowing method

This section is devoted to the analysis of a pathfollowing method for solving (10), which employs an inexact Newton corrector. Since the exact Newton correction Δv^k defined by the Newton equation

$$\partial_v F(v^k; \mu^k) \Delta v^k = -F(v^k; \mu^k)$$

is numerically unavailable due to discretization and iteration errors, we resort to inexact Newton methods, where an inner residual r^k remains when computing the inexact Newton correction δv^k by

$$\partial_v F(v^k; \mu^k) \delta v^k = -F(v^k; \mu^k) + r^k. \quad (12)$$

Algorithm 3.1.

select $\mu^0 > 0$, $\delta > 0$, $0 < \sigma < 1$, and v^0 with $\|v^0 - v(\mu^0)\| \leq \rho\sqrt{\mu^0}$

For $k = 0, \dots$

$$\mu^{k+1} = \sigma\mu^k$$

solve (12) for δv^k with a relative tolerance of $\|\delta v^k - \Delta v^k\| \leq \delta\|\Delta v^k\|$

$$v^{k+1} = v^k + \delta v^k$$

In the remainder of the section we show that for suitable choice of δ , σ , and ρ this algorithm is well defined and computes iterates that converge to the solution $v(0)$. First we derive a Lipschitz constant for the derivative of F .

Lemma 3.2. For $\alpha > 0$ there is a constant $\omega < \infty$ such that

$$\|\partial_v F(v; \mu)^{-1}(\partial_v F(v; \mu) - \partial_v F(\hat{v}, \mu))(v - \hat{v})\| \leq \frac{\omega}{\sqrt{\mu\alpha^3}} \|v - \hat{v}\|^2.$$

Proof. First we note that

$$\partial_v F(v; \mu) = \begin{bmatrix} I & L \\ L & -u'(\lambda) \end{bmatrix},$$

such that by Lemma 2.1

$$\begin{aligned} \|(\partial_v F(v; \mu) - \partial_v F(\hat{v}, \mu))(v - \hat{v})\|_{L^2} &= \|(u'(\hat{\lambda}) - u'(\lambda))(\lambda - \hat{\lambda})\|_{L^2} \\ &\leq \frac{1}{\sqrt{\mu\alpha^3}} \|(\lambda - \hat{\lambda})^2\|_{L^2} \\ &= \frac{1}{\sqrt{\mu\alpha^3}} \|\lambda - \hat{\lambda}\|_{L^4}^2 \\ &\leq \frac{1}{\sqrt{\mu\alpha^3}} \|\lambda - \hat{\lambda}\|_{H^1}^2. \end{aligned}$$

Next we note that $\partial_v F(v; \mu) : (H_0^1)^2 \rightarrow (H^{-1})^2$ satisfies the assumptions of the saddle point Lemma given by [2]. In particular, $u'(\lambda)$ is a positive semidefinite and bounded Nemyckii operator. From this we obtain the existence of a constant $\omega = \omega((\alpha + 2\mu)^{-1})$ such that

$$\begin{aligned} \|\partial_v F(v; \mu)^{-1}(\partial_v F(v; \mu) - \partial_v F(\hat{v}, \mu))(v - \hat{v})\|_{H^1} \\ \leq \omega \|(\partial_v F(v; \mu) - \partial_v F(\hat{v}, \mu))(v - \hat{v})\|_{L^2} \\ \leq \frac{\omega}{\sqrt{\mu\alpha^3}} \|\lambda - \hat{\lambda}\|_{H^1}^2. \end{aligned}$$

□

Now we are ready to formulate the main result of this section.

Theorem 3.3. Assume that

$$\rho \leq \frac{\sqrt{\alpha^3}}{4\omega}, \quad \sigma \geq \frac{1}{2}, \quad \frac{1-\sigma}{\sqrt{\sigma}} \leq \frac{\rho}{2c_0} \left(1 + \frac{2\omega c_0}{\sqrt{\alpha^3}}\right)^{-1}, \quad \delta \leq \frac{\sqrt{\sigma}}{4}, \quad (13)$$

and $\|v^0 - v(\mu^0)\|_{H^1} \leq \rho\sqrt{\mu^0}$. Then the iterates defined by Algorithm 3.1 are all well defined and converge linearly towards the limit point $v(0)$. More precisely,

$$\|v^k - v(\mu^k)\|_{H^1} \leq \rho\sqrt{\mu^k} \quad \text{and} \quad \|v^k - v(0)\|_{H^1} \leq c\sqrt{\mu^k} \quad (14)$$

with some generic constant c .

Proof. By induction, assume that (14) holds for some k . Using the estimate for the derivative of the central path given by Theorem 2.3 we derive

$$\begin{aligned} \|v^k - v(\mu^{k+1})\|_{H^1} &\leq \|v^k - v(\mu^k)\|_{H^1} + \|v(\mu^k) - v(\mu^{k+1})\|_{H^1} \\ &\leq \rho\sqrt{\mu^k} + \frac{c}{\sqrt{\sigma\mu^k}}(1-\sigma)\mu^k \\ &= \left(\rho + c\frac{1-\sigma}{\sqrt{\sigma}}\right)\sqrt{\mu^k}. \end{aligned}$$

By the Newton-Mysovskii theorem, one exact Newton step for $\mu = \mu^{k+1}$ yields

$$\begin{aligned} \|v^k + \Delta v^k - v(\mu^{k+1})\|_{H^1} &\leq \frac{\omega}{2\sqrt{\sigma\mu^k\alpha^3}}\|v^k - v(\mu^{k+1})\|_{H^1}^2 \\ &\leq \frac{\omega}{2\sqrt{\sigma\mu^k\alpha^3}}\left(\rho + c\frac{1-\sigma}{\sqrt{\sigma}}\right)^2\mu^k \\ &\leq \frac{\omega}{2\sigma\sqrt{\alpha^3}}\left(\rho + c\frac{1-\sigma}{\sqrt{\sigma}}\right)^2\sqrt{\mu^{k+1}}. \end{aligned}$$

The length of the exact Newton correction is bounded by

$$\begin{aligned} \|\Delta v^k\|_{H^1} &\leq \|v(\mu^{k+1}) - v^k - \Delta v^k\|_{H^1} + \|v(\mu^{k+1}) - v^k\|_{H^1} \\ &\leq \frac{\omega}{2\sigma\sqrt{\alpha^3}}\left(\rho + c\frac{1-\sigma}{\sqrt{\sigma}}\right)^2\sqrt{\mu^{k+1}} + \frac{\rho}{\sqrt{\sigma}}\sqrt{\mu^{k+1}}. \end{aligned}$$

Using the assumptions (13), we can now estimate the error of the next iterate v^{k+1} obtained by an inexact Newton correction as

$$\begin{aligned} \|v^{k+1} - v(\mu^{k+1})\|_{H^1} &\leq \left[\frac{\omega}{\sigma\sqrt{\alpha^3}}\left(\rho + c\frac{1-\sigma}{\sqrt{\sigma}}\right)^2 + \delta\frac{\rho}{\sqrt{\sigma}}\right]\sqrt{\mu^{k+1}} \\ &\leq \left[\frac{\omega}{\sigma\sqrt{\alpha^3}}\left(\rho^2 + 2\rho c\frac{1-\sigma}{\sqrt{\sigma}} + c^2\left(\frac{1-\sigma}{\sqrt{\sigma}}\right)^2\right) + \frac{\rho}{4}\right]\sqrt{\mu^{k+1}} \\ &\leq \left[2\left(\frac{\rho}{8} + c\frac{1-\sigma}{2\sqrt{\sigma}} + \frac{\omega}{\sqrt{\alpha^3}}c^2\frac{1-\sigma}{\sqrt{\sigma}}\right) + \frac{\rho}{4}\right]\sqrt{\mu^{k+1}} \\ &\leq \left[c\left(1 + \frac{2\omega}{\sqrt{\alpha^3}}c\right)\frac{1-\sigma}{\sqrt{\sigma}} + \frac{\rho}{2}\right]\sqrt{\mu^{k+1}} \\ &\leq \rho\sqrt{\mu^{k+1}}, \end{aligned}$$

which completes the induction. \square

4 Finite element discretization

The advantage of eliminating the control pointwise is apparent when it comes to discretize the variables. Remember that Algorithm 3.1 requires the solution of (12) up to a relative (discretization) error of $\|\delta v^k - \Delta v^k\| \leq \delta\|\Delta v^k\|$.

Throughout the section we assume $\mu > 0$ to be fixed, and hence omit it from notation.

For simplicity we assume the domain Ω to be convex with boundary of class C^∞ , such that we need not care for limited regularity of PDE solutions due to boundary effects.

Lemma 4.1. *Assume that $y_d \in H^1$. Then there exists a constant $c < \infty$ independent of μ and α , such that the central path solutions (y, λ) and $u(\lambda; \mu)$ satisfy the following regularity conditions:*

$$\begin{aligned} \|y\|_{H^2} \leq c, \quad \|y\|_{H^3} \leq c \left(1 + \frac{1}{\alpha}\right), \quad \|\lambda\|_{H^3} \leq c, \\ \|u(\lambda)\|_{L^2} \leq c, \quad \|u(\lambda)\|_{H^1} \leq c \left(1 + \frac{1}{\alpha}\right), \quad \|u(\lambda)\|_{H^2} \leq c \left(1 + \frac{1}{\alpha} + \frac{1}{\sqrt{\alpha^3 \mu}}\right). \end{aligned}$$

Proof. We denote by c a generic constant which is independent of μ and α . Since $|u| \leq 1$ by construction, $\|u(\lambda)\|_{L^2} \leq \sqrt{|\Omega|}$ is immediately clear. By standard regularity results for elliptic PDEs we have immediately $\|y\|_{H^2} \leq c\|u\|_{L^2} \leq c$ and $\|\lambda\|_{H^3} \leq c\|y - y_d\|_{H^1} \leq c$. By the Sobolev lemma we obtain $\lambda \in C^1$ and with $\nabla u = u'(\lambda)\nabla\lambda$ also $\|u\|_{H^1} \leq \frac{1}{\alpha}\|\lambda\|_{H^1} \leq c(1 + \frac{1}{\alpha})$. This in turn implies $\|y\|_{H^3} \leq c(1 + \frac{1}{\alpha})$. Concerning the H^2 -estimate for u we compute $\nabla^2 u = u''(\lambda)(\nabla\lambda, \nabla\lambda) + u'(\lambda)\nabla^2\lambda$ and conclude $\|u\|_{H^2} \leq c\|u''\|_\infty\|\lambda\|_{C^1}^2 + \|u'\|_\infty\|\lambda\|_{H^2}$, which yields the assertion. \square

We consider finite element discretizations on a sequence of uniformly shape-regular and quasi-uniform triangulations \mathcal{T}_h of Ω with maximal element size h . On \mathcal{T}_h we define for $p = 1, 2$ the ansatz space $V_h^p = \{\phi \in (H_0^1(\Omega))^2 \mid \forall T \in \mathcal{T}_h : \phi|_T \in \mathbb{P}_p^2\}$ of piecewise linear and piecewise quadratic functions, respectively. Since the accuracy of the numerical integration used for assembling $\langle u, \phi \rangle$ turns out to be crucial, it is considered separately from the finite element discretization. We interpret the numerical integration as exact integration of a projection $P_h u$ of u . The discrete approximation $F_h : V_h^p \rightarrow (V_h^p)^*$ is then defined by

$$\langle F_h(vh), (\phi_1, \phi_2) \rangle = \langle y_h - y_d + L\lambda_h, \phi_1 \rangle + \langle Ly_h - P_h u(\lambda_h), \phi_2 \rangle \quad \forall (\phi_1, \phi_2) \in V_h^p.$$

With $S_h : H^{-1} \rightarrow V_h^p, z \mapsto \zeta$, we denote the solution operator of the discretized problem

$$\langle L\zeta - z, \phi \rangle = 0 \quad \forall \phi \in V_h^p$$

with exact integration.

In the following subsection, we derive discretization error estimates based on certain accuracy assumptions imposed on P_h , whereas the next subsection is devoted to the question how to realize an integration scheme that satisfies these assumptions.

4.1 Error estimates

First we collect some immediate results of standard finite element theory.

Lemma 4.2. *The following estimates hold for $p = 1, 2$:*

$$\|Sz\|_{H^m} \leq c\|z\|_{H^{m-2}} \quad \text{for } m = 1, 2, 3 \quad (15)$$

$$\|(S_h - S)z\|_{H^1} \leq ch^p\|z\|_{H^{p-1}} \quad (16)$$

$$\|(S_h - S)z\|_{L^2} \leq ch^{p+1}\|z\|_{H^{p-1}} \quad (17)$$

$$\|(S_h S_h - SS)z\|_{L^2} \leq ch^{p+1}\|z\|_{H^1} \quad (18)$$

$$\|S_h z\|_{H^2, \mathcal{T}} \leq c\|z\|_{H^1} \quad (19)$$

Proof. The regularity result (15) holds for sufficiently smooth $\partial\Omega$ (see e.g. [5]). For (16) and (17) we refer to [6, Thms. 8.4.11, 8.5.1 and §8.6]. As for (18) we estimate

$$\begin{aligned} \|(S_h S_h - SS)z\|_{L^2} &\leq \|S_h(S_h - S)z\|_{L^2} + \|(S_h - S)Sz\|_{L^2} \\ &\leq c\|(S_h - S)z\|_{L^2} + ch^{p+1}\|Sz\|_{H^{p-1}} \\ &\leq ch^{p+1}\|z\|_{H^{p-1}} + ch^{p+1}\|z\|_{L^2} \\ &\leq ch^{p+1}\|z\|_{H^1}. \end{aligned}$$

Concerning (19), we only need to consider the case $p = 2$, and we first notice that $Sz \in H^3$. Thus the Lagrange interpolate $I_h Sz$ of Sz corresponding to the quadratic finite elements satisfies the error estimates:

$$\begin{aligned} \|I_h Sz - Sz\|_{H^2, \mathcal{T}} &\leq hc\|z\|_{H^1} \\ \|I_h Sz - Sz\|_{H^1} &\leq h^2c\|z\|_{H^1} \end{aligned}$$

and by (16) we obtain

$$\|I_h Sz - S_h z\|_{H^1} \leq \|I_h Sz - Sz\|_{H^1} + \|S_h z - Sz\|_{H^1} \leq h^2c\|z\|_{H^1}.$$

Now we may use an inverse inequality $\|v_h\|_{H^2, \mathcal{T}} \leq \frac{c}{h}\|v_h\|_{H^1}$ to obtain

$$\begin{aligned} \|S_h z\|_{H^2, \mathcal{T}} &\leq \|I_h Sz - S_h z\|_{H^2, \mathcal{T}} + \|I_h Sz - Sz\|_{H^2, \mathcal{T}} + \|Sz\|_{H^2, \mathcal{T}} \\ &\leq c(1+h)\|z\|_{H^1}. \end{aligned}$$

□

Lemma 4.3. *Let c denote some generic constant independent of μ , u , and v . If the linear projector P_h satisfies*

$$\|P_h u - u\|_{H^{-1}} \leq ch^p \sqrt{\mu} \|u\|_{H^2, \mathcal{T}}, \quad (20)$$

then there is a discrete central path solution $v_h \in V_h^p$ with associated control $u_h = u(\lambda_h)$, such that for both linear ($p = 1$) and quadratic ($p = 2$) finite elements the error estimates

$$\begin{aligned} \|v_h - v(\mu)\|_{H^1} &\leq ch^p \\ \|u_h - u(\mu)\|_{L^2} &\leq ch^p \end{aligned}$$

hold.

Proof. (i) *Existence of v_h .* The set $M = \{u \in L^2 : |u| \leq 1 \text{ a.e.}\}$ is nonempty, closed, convex, and bounded, and the mapping

$$T : L^2 \rightarrow L^2, u \mapsto u(\lambda_h(y_h(P_h u) - y_d)) = u(-S_h(S_h P_h u - y_d))$$

is a compact operator that maps M into itself. By the Schauder fixed-point theorem (cf. [15, Theorem 2.A]) T has a fixed point $u_h = u(-S_h(S_h P_h u - y_d))$ and corresponding finite element solutions $v_h = (y_h, \lambda_h)$, such that $F_h(v_h) = 0$.

(ii) *L^2 error estimates for u_h .* First we note that

$$\lambda(u) = \alpha u - \frac{\mu}{u+1} + \frac{\mu}{1-u}$$

obtained from (8) is monotonically increasing with $\lambda'(u) \geq \alpha + \mu$, and holds for both the exact central path solution as well as for any discrete solution $\lambda_h = -S_h(S_h P_h u - y_d)$. We adapt the proof given in [7] to the present setting and estimate

$$\begin{aligned} (\alpha + \mu)\|u - u_h\|_{L^2}^2 &\leq \langle \lambda(u) - \lambda(u_h), u - u_h \rangle \\ &= \langle -S(Su - y_d) + S_h(S_h P_h u_h - y_d), u - u_h \rangle \\ &= \langle -SSu + Sy_d + S_h S_h P_h u_h - S_h y_d, u - u_h \rangle \\ &= \langle (S_h S_h - SS)u, u - u_h \rangle + \underbrace{\langle S_h S_h(u_h - u), u - u_h \rangle}_{\leq 0} \\ &\quad + \langle S_h S_h(P_h u_h - u_h), u - u_h \rangle + \langle (S - S_h)y_d, u - u_h \rangle. \end{aligned}$$

Dividing by $\|u - u_h\|_{L^2}$ we obtain

$$\begin{aligned} (\alpha + \mu)\|u - u_h\|_{L^2} &\leq \|(S_h S_h - SS)u\|_{L^2} + \|S_h S_h(P_h - I)u_h\|_{L^2} + \|(S - S_h)y_d\|_{L^2} \\ &\leq ch^{p+1}\|u\|_{H^1} + ch^p\sqrt{\mu}\|u_h\|_{H^2, \mathcal{T}} + ch^{p+1}\|y_d\|_{H^1} \\ &\leq ch^p. \end{aligned}$$

(iii) *H^1 error estimates for y_h and λ_h .* Equipped with the estimate $\|u - u_h\|_{L^2} \leq ch^p$ we obtain by standard error estimates for finite element solutions

$$\begin{aligned} \|y - y_h\|_{H^1} &\leq \|Su - Su_h\|_{H^1} + \|Su_h - S_h u_h\|_{H^1} + \|S_h u_h - S_h P_h u_h\|_{H^1} \\ &\leq c\|u - u_h\|_{L^2} + ch^p\|u_h\|_{H^{p-1}} + ch^p\sqrt{\mu}\|u_h\|_{H^2, \mathcal{T}} \\ &\leq ch^p + ch^p\sqrt{\mu}\frac{c}{\sqrt{\alpha^3\mu}}\|\lambda_h\|_{H^2, \mathcal{T}} \\ &\leq ch^p \end{aligned}$$

and

$$\begin{aligned} \|\lambda - \lambda_h\|_{H^1} &\leq \|S(y - y_h)\|_{H^1} + \|(S_h - S)y_h\|_{H^1} \\ &\leq ch^p + ch^p\|y_h\|_{H^{p-1}} \\ &\leq ch^p. \end{aligned}$$

□

Corollary 4.4. *If (20) holds, Algorithm 3.1 computes discrete central path solutions $v_h \in V_h^p$ with associated control u_h with distance to the solution $v(0)$ of (1) bounded by*

$$\begin{aligned}\|v_h - v(0)\|_{H^1} &\leq ch^p \\ \|u_h - u(0)\|_{L^2} &\leq ch^p.\end{aligned}$$

Proof. For $\mu > 0$, Theorem 3.3 requires discrete solutions v_h with $\|v_h - v(\mu)\|_{H^1} \leq \rho\sqrt{\mu}$, which is satisfied for

$$ch^p \leq \rho\sqrt{\mu} \tag{21}$$

by Lemma 4.3. Using the coarsest possible discretization, i.e. choosing $h(\mu)$ or equivalently $\mu(h)$ such that equality holds in (21), we obtain

$$\|v_h - v(0)\|_{H^1} \leq \|v_h - v(\mu)\|_{H^1} + \|v(\mu) - v(0)\|_{H^1} \leq ch^p$$

by Theorem 2.3. □

Remark 4.5. Higher order results $\|u_h - u(0)\|_{L^2} \leq ch^{3/2}$ for linear and $\|u_h - u(0)\|_{L^2} \leq ch^3$ for quadratic finite elements can be obtained by using more accurate numerical integration and superconvergence properties of the central path, if the solution $v(0)$ exhibits certain regularity properties.

4.2 Numerical integration schemes

We will now analyse a simple possibility to construct an interpolation operator P_h that satisfies (20). Since we only have regularity in the norm $\|\cdot\|_{H^2, \mathcal{T}}$, we define $P_h = P_{h, \tilde{h}}$ as a piecewise linear interpolation on triangles with grid-size $\tilde{h} \leq h$, that are aligned to the finite-element grid. By standard interpolation theory we obtain the error estimate

$$\|P_h u - u\|_{L^2} \leq c\tilde{h}^2 \|u\|_{H^2, \mathcal{T}},$$

thus we have to choose

$$\tilde{h}^2 := ch^p \sqrt{\mu}.$$

Considering (21) this leads us to the choice

$$\tilde{h} := ch^p.$$

Thus, in the case of linear finite elements we may use a fixed subdivision of the finite element grid to interpolate u_h , but in the case of quadratic finite elements we have to introduce finer and finer subgrids for assembling u_h sufficiently accurate.

Remark 4.6. In actual computation with quadratic finite elements one will choose the accuracy of the integration adaptively. If the boundary $\partial\Omega_A$ of the active set Ω_A of u is not too complex, the additional computational effort for such an adaptive integration will be bounded by a fixed factor, since sharp bends in u_h are to be expected in the vicinity of $\partial\Omega_A$ only. Moreover, the contribution of this region to the overall error will be small, due to its shrinking size.

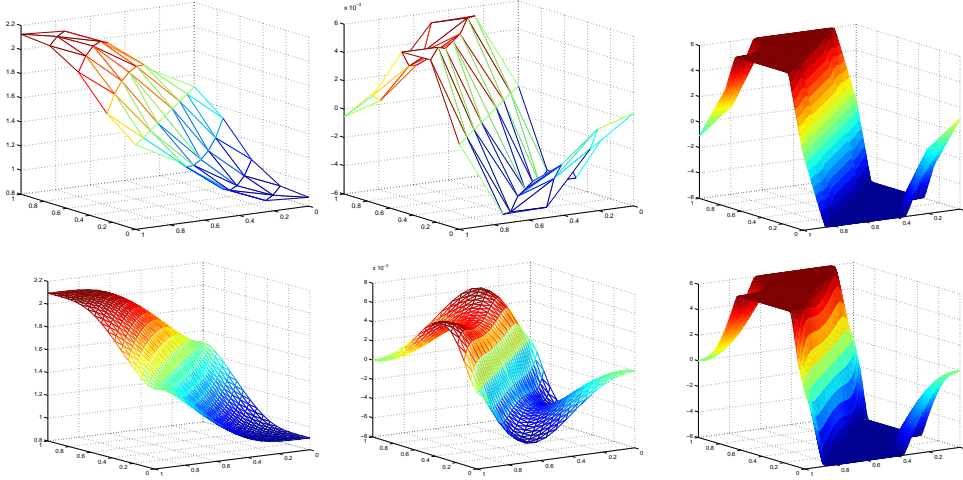


Figure 1: Discrete solution at $\mu = 2^{-18}$. From left to right: state y , multiplier λ , and control u . Top row: $h = 1/4$. Bottom row: $h = 1/16$. In order to display the accuracy gain, the control is interpolated on a significantly finer grid than y and λ were computed on.

5 Numerical examples

This section is devoted to demonstrate the method at some illustrative examples.

Example 1. As a completely artificial example we choose

$$\min_{y \in H^1, u \in L^2} \frac{1}{2} \|y - y_d\|_{L^2}^2 + \frac{\alpha}{2} \|u\|_{L^2}^2$$

subject to $-\Delta y = u$ in Ω , $\partial_n y = 0$ on $\partial\Omega$, $-6 \leq u \leq 6$

with $\Omega = [0, 1]^2$, $\alpha = 5e - 4$, and

$$y_d = \begin{cases} 1, & x_1 + x_2 < 1 \\ 2, & \text{otherwise} \end{cases}.$$

In order to examine the numerical convergence properties with respect to the mesh size h we employ uniform criss-cross triangulations with $h = 2^{-n}$.

Discrete central path solutions for different mesh sizes are shown in Figure 1, with a zoom into the control given in Figure 2. The structure of the control and in particular shape and location of the developing kinks are surprisingly well represented already on very coarse meshes.

This translates into the maximal order of convergence that can be expected for linear and quadratic finite elements, respectively. Discretization errors for state y and multiplier λ , both in $\|\cdot\|_{H^1}$ and $\|\cdot\|_{L^2}$, are given in Figure 3. Note that the quadratic and cubic convergence order of λ in L^2 for linear and quadratic elements, respectively, translates directly into the same order of convergence for u in L^2 .

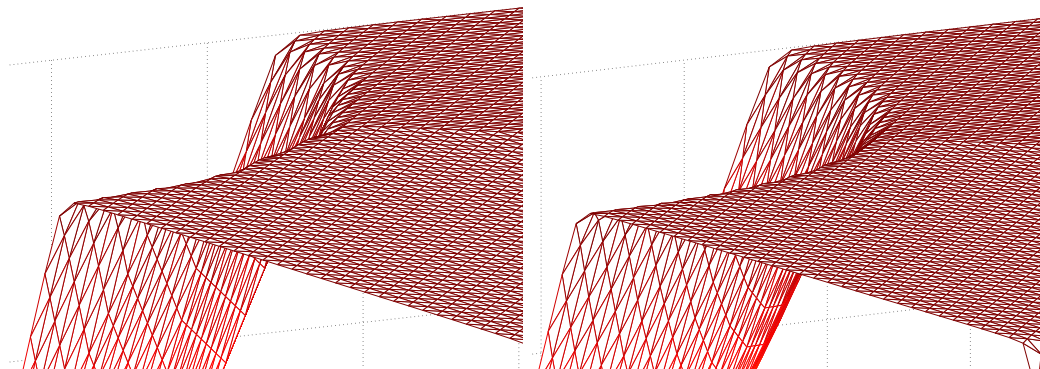


Figure 2: Zoom of control for $\mu = 2^{-18}$. Left: $h = 1/4$. Right: $h = 1/16$.

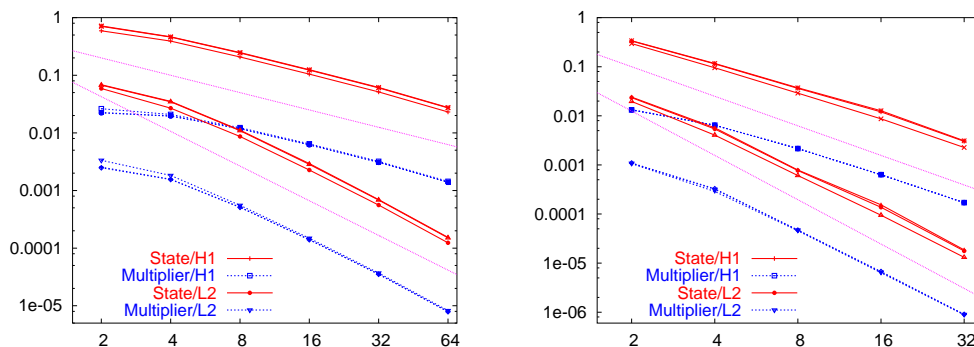


Figure 3: Convergence rates for linear (left) and quadratic (right) finite elements at different values of $\mu = 2^{-8}, 2^{-13}, 2^{-15}$.

Example 2. This problem taken from [14] is a drastically simplified benchmark problem for applicator development in regional hyperthermia treatment, a cancer therapy that aims at heating deeply seated tumors by microwave radiation in order to make it more susceptible to an accompanying radio or chemo therapy [4]. The governing PDE is the stationary bio-heat-transfer equation [9]

$$\begin{aligned} -\nabla(\kappa\nabla y) + (y - 37)w &= u && \text{in } \Omega \\ \partial_n y &= 0 && \text{on } \partial\Omega \end{aligned}$$

for the temperature y on the relevant part Ω of the human body. The control u , assumed to be freely adjustable within the bounds $0 \leq u \leq u_{\max}$, is the energy absorption of the tissue and is directly related to the amplitude of the time harmonic electric field generated by the microwave generator. The thermal effect of perfusion w with arterial blood of 37°C from different regions of the body is accounted for by

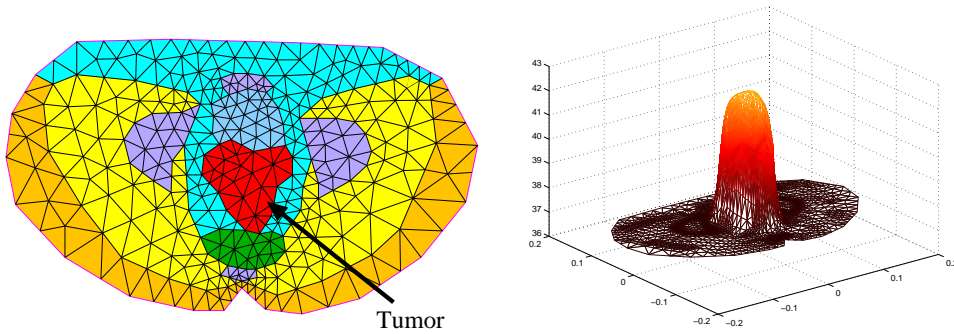


Figure 4: Cross section Ω of the pelvic region with different tissue types (left) and optimal state (right).

the Helmholtz term. We aim at a therapeutical temperature

$$y_d = \begin{cases} 45 & \text{in } \Omega_t \\ 37 & \text{in } \Omega \setminus \Omega_t \end{cases}$$

that affects only the tumor tissue $\Omega_t \subset \Omega$ (see Fig. 4). For this example, the regularization parameter α has been set to 10^{-12} . As can be expected, the optimal control just deposits almost all the energy into the tumor region and almost nothing outside. The very small value of α leads to a very thin band of steep increase in the control, which is, however, *not* aligned to the coarse grid.

This problem has been solved numerically by a primal-dual function space oriented interior point method as described in [13, 14]. The discretized control needs to represent the ‘discontinuity’ of the solution with sufficient accuracy, such that massive grid refinement occurs along the boundaries of the active sets (see Fig. 5, left). The finest computational grid contained about 40000 triangles, most of them concentrated along the control ‘discontinuity’.

We also solved this problem with the control reduced primal interior point method, however on a much coarser grid with about 2500 triangles. Then the control was computed via (8) on a uniformly refined grid (see Fig. 5, right), which yielded a comparable result with much less computational effort.

Conclusion

A novel discretization scheme for primal interior point methods applied to PDE constrained optimization problems has been presented. Pointwise elimination of the control, which is the least regular variable, enables high accuracy with comparatively coarse meshes.

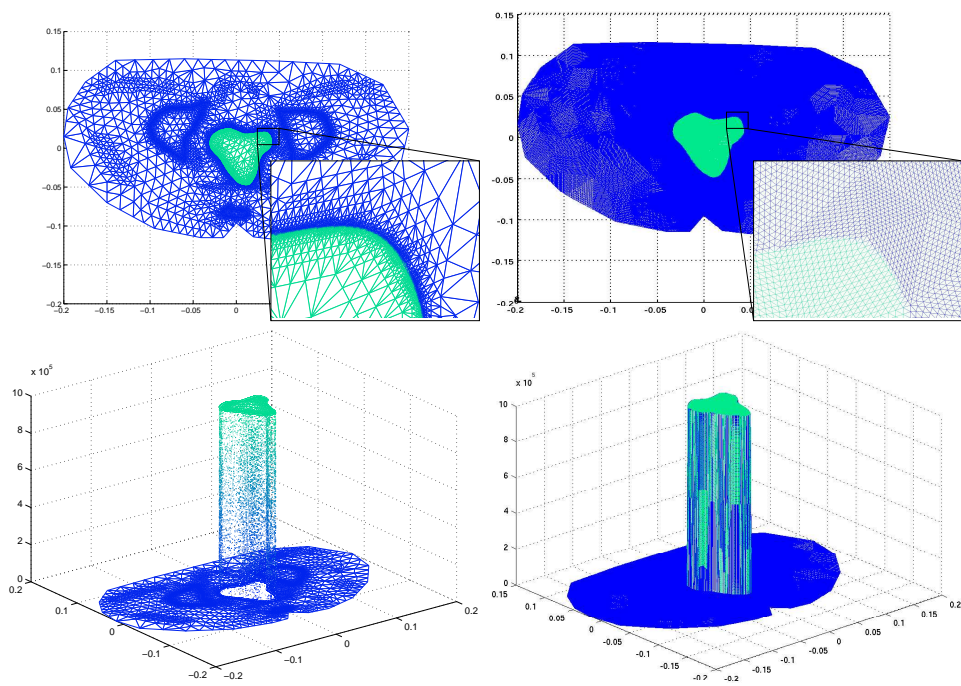


Figure 5: Mesh refinement and discrete control for the primal dual approach with piecewise constant control discretization (left) and for the control reduced primal approach (right) computed on a coarse grid.

References

- [1] N. Arada, E. Casas, and F. Tröltzsch. Error estimates for the numerical approximation of a semilinear elliptic control problem. *Comput. Optim. Appl.*, 23:201–229, 2002.
- [2] D. Braess and C. Blömer. A multigrid method for a parameter dependent problem in solid mechanics. *Numer. Math.*, 57:747–761, 1990.
- [3] E. Casas and F. Tröltzsch. Error estimates for linear-quadratic elliptic control problems. In V. Barbu et al., editor, *Analysis and Optimization of Differential Systems*, pages 89–100. Kluwer, 2003.
- [4] P. Deuffhard, M. Seebaß, D. Stalling, R. Beck, and H.-C. Hege. Hyperthermia Treatment Planning in Clinical Cancer Therapy: Modelling, Simulation and Visualization. In A. Sydow, editor, *Proc. of the 15th IMACS World Congress 1997 on Scientific Computation: Modelling and Applied Mathematics*, volume 3, pages 9–17. Wissenschaft und Technik Verlag, 1997.
- [5] D. Gilbarg and N.S. Trudinger. *Elliptic Partial Differential Equations of Second Order*. Springer, 1977.
- [6] W. Hackbusch. *Theorie und Numerik elliptischer Differentialgleichungen*. Teubner, 1986.
- [7] M. Hinze. A generalized discretization concept for optimal control problems with control constraints. *Comput. Optim. Appl.*, 2004 (to appear).
- [8] C. Meyer and A. Rösch. Superconvergence properties of optimal control problems. *SIAM J. Control Optim.*, to appear.
- [9] H.H. Pennes. Analysis of tissue and arterial blood temperatures in the resting human forearm. *J. Appl. Phys.*, 1:93–122, 1948.
- [10] U. Prüfert, F. Tröltzsch, and M. Weiser. The convergence of an interior point method for an elliptic control problem with mixed control-state constraints. in preparation.
- [11] A. Rösch. Error estimates for linear-quadratic control problems with control constraints. Preprint 0740-02, TU Berlin, 2002.
- [12] M. Weiser. Interior point methods in function space. ZIB Report 03-35, Zuse Institute Berlin, 2003.
- [13] M. Weiser and P. Deuffhard. The central path towards the numerical solution of optimal control problems. ZIB Report 01-12, Zuse Institute Berlin, 2001.
- [14] M. Weiser and A. Schiela. Function space interior point methods for PDE constrained optimization. ZIB Report 04-27, Zuse Institute Berlin, 2004.

- [15] E. Zeidler. *Nonlinear Functional Analysis and its Applications*, volume I. Springer, New York, 1986.