Department of Mathematics, University of Houston
Institut für Mathematik, Universität Augsburg

Numerical Solution of Elliptic Optimal Control Problems

Ronald H.W. Hoppe[1,2]

[1] Department of Mathematics, University of Houston
[2] Institute of Mathematics, University of Augsburg

Summer School 'Optimal Control of PDEs'

Cortona, July 12-17, 2010

Recent Books on Optimal Control of PDEs

R. Glowinski, J.L. Lions, and J. He; Exact and Approximate Controllability for Distributed Parameter Systems: A Numerical Approach. Cambridge University Press, Cambridge, 2008.

M. Hinze, R. Pinnau, and M. Ulbrich; Optimization with PDE Constraints. Springer, Berlin-Heidelberg-New York, 2008.

F. Tröltzsch; Optimal Control of Partial Differential Equations. Theory, Methods, and Applications. American Mathematical Society, Providence, 2010.

Numerical Solution of Elliptic Optimal Control Problems

- Unconstrained Elliptic Optimal Control Problems
- Control Constrained Elliptic Optimal Control Problems
- State Constrained Elliptic Optimal Control Problems

Elliptic Optimal Control Problems

Unconstrained Case

Elliptic Optimal Control Problems: Unconstrained Case

Consider the unconstrained distributed elliptic optimal control problem

$$\begin{aligned} \inf_{y,u} \; J(y,u) \; &:= \frac{1}{2} \int_\Omega |y - y^d|^2 \, dx + \frac{\alpha}{2} \int_\Omega |u|^2 \, dx, \\ (EUC)_1 \qquad -\Delta y &= u \quad \text{in } \Omega, \\ (EUC)_2 \qquad y &= 0 \quad \text{on } \Gamma, \end{aligned}$$

where $u \in L^2(\Omega)$.

Existence and Uniqueness of a Solution

We introduce the control-to-state map

$$G : L^2(\Omega) \to H_0^1(\Omega) \quad , \quad y = Gu \quad \text{solves} \quad -\Delta y = u \ .$$

Substituting $y = Gu$ allows to consider the reduced control problem

$$\inf_u J_{red}(u) := \frac{1}{2}\|Gu - y^d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{L^2(\Omega)} \ .$$

Theorem (Existence and uniqueness) The unconstrained elliptic control problem admits a unique solution $y \in H_0^1(\Omega)$.

Proof. Direct method of the calculus of variations (minimizing sequence).

Optimality Conditions for the Distributed Control Problem

There exists an adjoint state $p \in H_0^1(\Omega)$ such that the triple (y, p, u) satisfies

$$\begin{aligned} a(y, v) &= (u, v)_{L^2(\Omega)} \quad , \quad v \in H_0^1(\Omega) \ , \\ a(p, v) &= -\ (y - y^d, v)_{L^2(\Omega)} \quad , \quad v \in H_0^1(\Omega) \ , \\ p - \alpha u &= 0 \ , \end{aligned}$$

where $a(\cdot, \cdot)$ stands for the bilinear form

$$a(w, z) \ = \ \int_\Omega \nabla w \cdot \nabla z \, dx \quad , \quad w, z \in H_0^1(\Omega) \ .$$

Optimality Conditions for the FE Discretized Control Problem

There exists an adjoint state $p_h \in V_h$ such that the triple (y_h, p_h, u_h) satisfies

$$\begin{aligned} a(y_h, v_h) &= (u_h, v_h)_{L^2(\Omega)} \quad , \quad v_h \in V_h \ , \\ a(p_h, v_h) &= -\ (y_h - y^d, v_h)_{L^2(\Omega)} \quad , \quad v_h \in V_h \ , \\ p_h - \alpha u_h &= 0 \ . \end{aligned}$$

Finite Element Approximation of the Distributed Control Problem

Let $\mathcal{T}_h(\Omega)$ be a shape regular, simplicial triangulation of Ω and let

$$V_h := \{ v_h \in C(\Omega) \mid v_h|_T \in P_1(T) \ , \ T \in \mathcal{T}_h(\Omega) \ , \ v_h|_{\partial\Omega} = 0 \}$$

be the FE space of continuous, piecewise linear finite elements.

Consider the following FE Approximation of the distributed control problem

$$\begin{array}{lll} \text{Minimize} & J(y_h, u_h) \ := \ \frac{1}{2} \, \|y_h - y^d\|^2_{L^2(\Omega)} \ + \ \frac{\alpha}{2} \, \|u_h\|^2_{L^2(\Omega)} \ , \\ \text{over} & (y_h, u_h) \in V_h \times V_h \ , \\ \text{subject to} & a(y_h, v_h) \ = \ (u_h, v_h)_{L^2(\Omega)} \ , \ v_h \in V_h \ . \end{array}$$

Algebraic Formulation of the Discrete Optimality System

Let $V_\ell := \text{span}\{\varphi_\ell^{(1)}, \cdots, \varphi_\ell^{(N_\ell)}\}$, where $\varphi_\ell^{(i)}, 1 \le i \le N_\ell$, are the nodal basis functions. We refer to $M_\ell \in \mathbb{R}^{N_\ell \times N_\ell}$ and $A_\ell \in \mathbb{R}^{N_\ell \times N_\ell}$ as the associated mass matrix and stiffness matrix:

$$(M_\ell)_{ij} := (\varphi_\ell^{(i)}, \varphi_\ell^{(j)})_{L^2(\Omega)} \quad , \quad (A_\ell)_{ij} := a(\varphi_\ell^{(i)}, \varphi_\ell^{(j)}) \ , \ 1 \le i,j \le N_\ell.$$

The solution of the discrete optimality system requires the computation of $(y_\ell, p_\ell) \in \mathbb{R}^{N_\ell} \times \mathbb{R}^{N_\ell}$ as the solution of

$$\begin{pmatrix} A_\ell & -\alpha^{-1} M_\ell \\ M_\ell & A_\ell \end{pmatrix} \begin{pmatrix} y_\ell \\ p_\ell \end{pmatrix} = \begin{pmatrix} 0 \\ M_\ell y_\ell^d \end{pmatrix}.$$

Left and Right Transforming Iterations

For given $K \in \mathbb{R}^{N \times N}, d \in \mathbb{R}^N$, we want to solve the linear algebraic system

$$(*) \qquad Kx = d.$$

Let $K_L, K_R \in \mathbb{R}^{N \times N}$ be regular. Then, $(*)$ can be equivalently wriiten as

$$(**) \qquad K_L K K_R K_R^{-1} x = K_L d.$$

Assuming $\tilde{K}$ to be a suitable preconditioner for $K_L K K_R$, we consider the transforming iteration

$$K_R^{-1} x^{(\nu+1)} = K_R^{-1} x^{(\nu)} + \tilde{K}^{-1} (K_L d - K_L K x^{(\nu)}).$$

Backtransformation yields

$$x^{(\nu+1)} = x^{(\nu)} + \underbrace{K_R \tilde{K}^{-1} K_L}_{= (K_L^{-1} \tilde{K} K_R^{-1})^{-1}} (d - K x^{(\nu)}),$$

so that $\hat{K} := K_L^{-1} \tilde{K} K_R^{-1}$ is an appropriate preconditioner for the original system.

Left and Right Transforming Iterations II

We apply the idea of left/right transformations to

$$K = \begin{pmatrix} A_\ell & -\alpha^{-1}M_\ell \\ M_\ell & A_\ell \end{pmatrix} \quad , \quad d = \begin{pmatrix} 0 \\ M_\ell y_\ell^d \end{pmatrix} .$$

We choose the left and right transformations according to

$$K_L = \begin{pmatrix} \alpha^{1/2} I & 0 \\ 0 & -I \end{pmatrix} \quad , \quad K_R = \begin{pmatrix} \alpha^{-1/2} I & 0 \\ 0 & I \end{pmatrix} .$$

This leads to the transformed matrix

$$K_L K K_R = \begin{pmatrix} A_\ell & -\alpha^{-1/2} M_\ell \\ -\alpha^{-1/2} M_\ell & -A_\ell \end{pmatrix} .$$

Department of Mathematics, University of Houston
Institut für Mathematik, Universität Augsburg

Left and Right Transforming Iterations III

Denoting by S_ℓ the Schur complement $S_\ell = A_\ell + \alpha^{-1} M_\ell A_\ell^{-1} M_\ell$, we have

$$\begin{pmatrix} A_\ell & -\alpha^{-1/2} M_\ell \\ -\alpha^{-1/2} M_\ell & -A_\ell \end{pmatrix} = \begin{pmatrix} A_\ell & -\alpha^{-1/2} M_\ell \\ -\alpha^{-1/2} M_\ell & -S_\ell + \alpha^{-1} M_\ell A_\ell^{-1} M_\ell \end{pmatrix}.$$

Choosing $\hat{A}_\ell$ as a preconditioner for A_ℓ and $\hat{S}_\ell := \tau_\ell^{-1} \operatorname{diag}(A_\ell + \alpha^{-1} M_\ell \hat{A}_\ell^{-1} M_\ell)$, we obtain the symmetric Uzawa preconditioner

$$\tilde{K}_\ell = \begin{pmatrix} \hat{A}_\ell & -\alpha^{-1/2} M_\ell \\ -\alpha^{-1/2} M_\ell & -\hat{S}_\ell + \alpha^{-1} M_\ell \hat{A}_\ell^{-1} M_\ell \end{pmatrix}.$$

Backtransformation results in the following preconditioner

$$\hat{K}_\ell = K_L^{-1} \tilde{K}_\ell K_R^{-1} = \begin{pmatrix} \hat{A}_\ell & -\alpha^{-1} M_\ell \\ M_\ell & \hat{S}_\ell - \alpha^{-1} M_\ell \hat{A}_\ell^{-1} M_\ell \end{pmatrix}.$$

Multigrid Solvers for Elliptic Optimal Control Problems

A. Borzi, 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 Optimization 41, 1477-1497, 2003.

A. Borzi and V. Schulz; Multigrid methods for PDE optimization.
SIAM Rev. 51, 361-395, 2009.

J. Schöberl, R. Simon, and W. Zulehner; A robust multigrid method for elliptic optimal control problems.
Preprint, Inst. of Comput. Math., University of Linz, 2010.

Preconditioned Richardson Iteration I

Step 1: Given iterates $y_\ell^{(\nu)}, p_\ell^{(\nu)} \in \mathbb{R}^{N_\ell}$, compute the residuals

$$\begin{pmatrix} \mathrm{Res}_\ell^{(1)} \\ \mathrm{Res}_\ell^{(2)} \end{pmatrix} = \begin{pmatrix} 0 \\ M_\ell y_\ell^d \end{pmatrix} - K_\ell \begin{pmatrix} y_\ell^{(\nu)} \\ p_\ell^{(\nu)} \end{pmatrix}.$$

Step 2: Solve the linear system

$$\begin{pmatrix} \hat{A}_\ell & -\alpha^{-1} M_\ell \\ M_\ell & \hat{S}_\ell - \alpha^{-1} M_\ell \hat{A}_\ell^{-1} M_\ell \end{pmatrix} \begin{pmatrix} \Delta y_\ell^{(\nu)} \\ \Delta p_\ell^{(\nu)} \end{pmatrix} = \begin{pmatrix} \mathrm{Res}_\ell^{(1)} \\ \mathrm{Res}_\ell^{(2)} \end{pmatrix}.$$

Step 3: Compute the new iterates

$$\begin{pmatrix} y_\ell^{(\nu+1)} \\ p_\ell^{(\nu+1)} \end{pmatrix} = \begin{pmatrix} y_\ell^{(\nu)} \\ p_\ell^{(\nu)} \end{pmatrix} + \begin{pmatrix} \Delta y_\ell^{(\nu)} \\ \Delta p_\ell^{(\nu)} \end{pmatrix}.$$

Preconditioned Richardson Iteration II

Consider the linear system

$$(*) \qquad \begin{pmatrix} \hat{A}_\ell & -\alpha^{-1}M_\ell \\ M_\ell & \hat{S}_\ell - \alpha^{-1}M_\ell\hat{A}_\ell^{-1}M_\ell \end{pmatrix} \begin{pmatrix} \Delta y_\ell^{(\nu)} \\ \Delta p_\ell^{(\nu)} \end{pmatrix} = \begin{pmatrix} Res_\ell^{(1)} \\ Res_\ell^{(2)} \end{pmatrix}.$$

Elimination of $\Delta y_\ell^{(\nu)}$ results in the Schur complement system

$$\hat{S}_\ell \Delta p_\ell^{(\nu)} = Res_\ell^{(2)} - M_\ell \hat{A}_\ell^{-1} Res_\ell^{(1)}.$$

Hence, the solution of $(*)$ can be reduced to the successive solution of the three linear systems

$$\begin{aligned} \hat{A}_\ell \tilde{\Delta} y_\ell^{(\nu)} &= Res_\ell^{(1)}, \\ \hat{S}_\ell \Delta p_\ell^{(\nu)} &= Res_\ell^{(2)} - M_\ell \tilde{\Delta} y_\ell^{(\nu)}, \\ \hat{A}_\ell \Delta y_\ell^{(\nu)} &= Res_\ell^{(1)} + \alpha^{-1} M_\ell \Delta p_\ell^{(\nu)}. \end{aligned}$$

Preconditioned Richardson Iteration III

Theorem [Schöberl/Simon/Zulehner] If the preconditioners $\hat{A}_\ell$ and $\hat{S}_\ell$ are chosen according to

$$\hat{A}_\ell = \sigma_\ell^{-1} \operatorname{diag}(A_\ell) \quad , \quad \hat{S}_\ell = \tau_\ell^{-1} \operatorname{diag}(S_\ell),$$
$$\sigma_\ell \le \lambda_{\max}((\operatorname{diag}(A_\ell))^{-1} A_\ell)^{-1} \quad , \quad \tau_\ell \le \lambda_{\max}((\operatorname{diag}(S_\ell))^{-1} S_\ell)^{-1},$$

then there holds

$$\|(y_\ell^{(\nu)}, p_\ell^{(\nu)}) - (y_\ell, p_\ell)\| \le C\ (1 + \alpha^{-1/2}\ h_\ell^2)\ \eta(\nu)\ \|(y_\ell^{(0)}, p_\ell^{(0)}) - (y_\ell, p_\ell)\|,$$

where

$$\eta(\nu) \le \begin{cases} \sqrt{\frac{2}{\pi(\nu-1)}} \ , \text{ if } \nu \text{ is even} \\ \sqrt{\frac{2}{\pi\nu}} \ , \text{ if } \nu \text{ is odd} \end{cases},$$

and C is a positive constant independent of h_ℓ and α.

Multigrid Solution of the Optimality System

The previous theorem suggests to use the preconditioned iterative solver as a smoother within a multigrid solution of the optimality system.

Theorem. In case of two nested grids $\mathcal{T}_{\ell-1}(\Omega)$ and $\mathcal{T}_{\ell}(\Omega)$ let $(y_\ell^{(0)}, p_\ell^{(0)})$ be a level ℓ startiterate and let $(y_\ell^{(1)}, p_\ell^{(1)})$ be the result of a two-grid cycle with $\nu > 0$ smoothing steps. Then there holds

$$\|(y_\ell^{(1)}, p_\ell^{(1)}) - (y_\ell, p_\ell)\| \leq C \, \frac{1 + \alpha^{-1/2} h_\ell^2}{\sqrt{\nu}} \, \|(y_\ell^{(0)}, p_\ell^{(0)}) - (y_\ell, p_\ell)\|,$$

where $C > 0$ is a constant independent of h_ℓ and α.

Elliptic Optimal Control Problems

Control Constraints

Department of Mathematics, University of Houston
Institut für Mathematik, Universität Augsburg

Elliptic Optimal Control Problems: Control Constraints

Consider the control constrained distributed elliptic optimal control problem

$$\inf_{y,u} J(y,u) := \frac{1}{2}\int_\Omega |y - y^d|^2 \, dx + \frac{\alpha}{2} \int_\Omega |u|^2 \, dx,$$

$$(\mathrm{ECC})_1 \qquad -\Delta y = u \quad \text{in } \Omega,$$

$$(\mathrm{ECC})_2 \qquad y = 0 \quad \text{on } \Gamma,$$

and

$$u \in K_C := \{v \in L^2(\Omega) \mid v(x) \le \psi(x) \text{ f.a.a. } x \in \Omega\},$$

where $u \in L^2(\Omega)$.

Control Constraints: Necessary Optimality Conditions

Theorem. If $(y, u) \in H_0^1(\Omega) \times K_C$ is the optimal solution of the optimal control problem, then there exists

$$(p, \lambda) \in H_0^1(\Omega)) \times L^2(\Omega)$$

such that p satisfies the adjoint state equation

$$\begin{aligned} -\Delta p &= -(y - y^d) && \text{in } \Omega, \\ p &= 0 && \text{on } \Gamma, \end{aligned}$$

and there holds

$$\begin{aligned} p - \lambda &= \alpha u && \text{in } \Omega, \\ \lambda &\in \partial I_{K_c}(u). \end{aligned}$$

Elliptic Optimal Control Problems

Primal-Dual Active Set Strategy

Moreau-Yosida Approximation of Multivalued Maps I

Weighted Duality Mapping: Assume that V is a Banach space with dual V^* and let $h : \mathbb{R}_+ \to \mathbb{R}_+$ be a continuous and non-decreasing function such that $h(0) = 0$ and $h(t) \to \infty$ as $t \to \infty$. Then the mapping $J_h : V \to 2^{V^*}$

$$J_h(u) := \{u^* \in V^* \mid \langle u^*, u\rangle = \|u\|\|u^*\| \ , \ \|u^*\| = h(\|u\|)\}$$

is called the duality mapping with weight h.

Example: For $V = L^p(\Omega), V^* = L^q(\Omega), 1 < p, q < \infty, 1/p + 1/q = 1$, and $h(t) = t^{p-1}$ we have

$$J_h(u)(x) = \begin{cases} |u(x)|^{p-1} \operatorname{sgn}(u(x)) \ , \ u(x) \neq 0 \\ 0 \ , \ u(x) = 0 \end{cases} .$$

Department of Mathematics, University of Houston
Institut für Mathematematik, Universität Augsburg

Moreau-Yosida Approximation of Multivalued Maps II

Moreau-Yosida proximal map: Let $f : V \to \bar{\mathbb{R}}$ be a lower semi-continuous proper convex function with subdifferential ∂f. For $c > 0$, the Moreau-Yosida proximal map $P_c^{\partial f} : V \to 2^V$ is defined such that $P_c^{\partial f}(w)$, $w \in V$, is the set of minimizers of

$$\inf_{v \in V} f(v) + c j_h\left(\frac{v - w}{c}\right),$$

where $\partial j_h = J_h$.

Moreau-Yosida approximation: If J_h is single-valued, then for $c > 0$ the Moreau-Yosida approximation $(\partial f)_c$ of ∂f is given by

$$(\partial f)_c(w) := J_h(c^{-1}w - c^{-1}P_c^{\partial f}(w)).$$

Moreau-Yosida Approximation of ∂I_{K_C}

Idea: Approximate ∂I_{K_C} by its Moreau-Yosida approximation $(\partial I_{K_C})_c$.

Theorem. For any $c > 0$, we have

$$\lambda \in (\partial I_{K_C})_c,$$

if and only if there holds

$$\lambda = c\Big(u + c^{-1}\lambda - \Pi_{K_C}(u + c^{-1}\lambda)\Big) = c\ \max(0, u + c^{-1}\lambda - \psi),$$

and this is equivalent to

$$u = \Pi_{K_C}(u + c^{-1}\lambda),$$

where Π_{K_C} denotes the L^2-projection onto K_C.

Primal-Dual Active Set Strategy I

Step 1 (Initialization):

Choose $c > 0$, start-iterates $y_h^{(0)}, u_h^{(0)}, \lambda_h^{(0)}$ and set $n = 1$.

Step 2 (Specification of active/inactive sets):

Compute the active/inactive sets $\mathcal{A}_n$ and $\mathcal{I}_n$ according to

$$\mathcal{A}_n := \{1 \le i \le N \mid (u_h^{(n-1)} + c^{-1}\lambda_h^{(n-1)})_i > (\psi_h)_i\} \quad , \quad \mathcal{I}_n := \{1, \cdots, N\} \setminus \mathcal{A}_n.$$

Step 3 (Termination criterion):

If $n \ge 2$ and $\mathcal{A}_n = \mathcal{A}_{n-1}$, stop the algorithm. Otherwise, go to Step 4.

Primal-Dual Active Set Strategy II

Step 4 (Update of the state, adjoint state, and control):

Compute $y_h^{(n)}, p_h^{(n)}$ as the solution of

$$(A_h y_h^{(n)})_i = \begin{cases} (\psi_h)_i \ , & \text{if } i \in \mathcal{A}_n \\ \alpha^{-1}(M_h^{-1} p_h)_i \ , & \text{if } i \in \mathcal{I}_n \end{cases} \quad , \quad A_h p_h^{(n)} = -M_h(y_h - y_h^d),$$

and set

$$(u_h^{(n)})_i := \begin{cases} (\psi_h)_i \ , & \text{if } i \in \mathcal{A}_n \\ \alpha^{-1}(M_h^{-1} p_h^{(n)})_i \ , & \text{if } i \in \mathcal{I}_n \end{cases} .$$

Step 5 (Update of the multiplier):

Set $\lambda_h^{(n)} := p_h^{(n)} - \alpha M_h u_h^{(n)}$, $n := n + 1$, and go to Step 2.

PDAS As A Semi-Smooth Newton Method I

Newton differentiability: A function $f : S \subset X \to Y$ (X, Y Banach spaces) is called Newton differentiable at $x \in S$, if there exist a neighborhood $\mathcal{U}(x) \subset S$ and a family of mappings $g : \mathcal{U}(x) \to \mathcal{L}(X, Y)$ such that

$$\|f(x+h) - f(x) - g(x+h)(h)\|_Y = o(\|h\|_X) \quad (\text{as } \|h\|_X \to \infty).$$

$g(x)$ is said to be the generalized derivative of f at x.

Example: The max-function $\max : L^p(\Omega) \to L^q(\Omega)$ is Newton differentiable for $1 \leq q < p < \infty$. If $f : L^r(\Omega) \to L^p(\Omega)$ is Fréchet-differentiable for some $1 < r \leq \infty$, then the function

$$x \longmapsto \chi_A(x)\ f(u(x))$$

is the generalized derivative of $\max(0, f(\cdot))$. Here, χ_A denotes the characteristic function of the set A where $f(u(\cdot))$ is nonnegative.

PDAS As A Semi-Smooth Newton Method II

Setting $z_h := (y_h, u_h, \lambda_h)$, the discrete optimality conditions require the solution of the nonlinear system

$$F_h(z_h) := \begin{pmatrix} A_h y_h - M_h u_h \\ \alpha A_h u_h + M_h y_h - y_h^d + \lambda_h \\ -\lambda_h + c \max(0, u_h + c^{-1}\lambda_h - \psi_h) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

The primal-dual active set strategy (PDAS) is equivalent to the semi-smooth Newton method

$$\begin{aligned} F_h'(z_h^{(n)})\Delta z_h^{(n)} &= -F_h(z_h^{(n)}), \\ z_h^{(n+1)} &= z_h^{(n)} + \Delta z_h^{(n)}, \end{aligned}$$

where $F_h'(z_h^{(n)})$ is the generalized derivative of F_h at $z_h^{(n)}$.

Elliptic Optimal Control Problems

Interior Point Methods

Interior-Point Methods (General)

Gonzaga [1992] , Wright [1992] , Zhang/Tapia/Dennis [1992]

Conn/Gould/Toint [1996] , Heinkenschloss [2001]

Forsgren/Gill/Wright [2002] , M. Ulbrich [2002] , S. Ulbrich [2003]

Interior-Point Methods (PDE Constrained Optimization)

Ulbrich/Ulbrich/Heinkenschloss [1999] , Ulbrich/Ulbrich [2000]

H./Petrova/Schulz [2002] , H./Petrova [2004] , Weiser [2005]

Ulbrich/Ulbrich [2007] , Antil/H./Linsenmann [2007]

Interior Point Method I

Idea: Coupling of control constraints $u_h \leq \psi_h$ by, e.g., logarithmic barrier functions with barrier parameter $\beta > 0$:

$$\inf_{y_h, u_h} B_h^{(\beta)}(y_h, u_h) := J_h(y_h, u_h) + \beta \sum_{i=1}^{N} \ln((\psi_h - u_h)_i),$$

which represents a parameter-dependent family of minimization subproblems.

Coupling of the discrete state equation by a Lagrange multiplier:

$$\inf_{y_h, u_h} \sup_{p_h} \mathcal{L}_h^{(\beta)}(y_h, u_h, p_h) := B_h^{(\beta)}(y_h, u_h) + \langle p_h, A_h y_h - M_h u_h \rangle.$$

Interior Point Method II

The barrier path $\beta \longmapsto x_h(\beta) := (y_h(\beta), u_h(\beta), p_h(\beta))$ is the solution of the parameter-dependent nonlinear system:

$$F_h(x_h(\beta), \beta) = \begin{pmatrix} A_h p_h + M_h y_h - y_h^d \\ A_h y_h - M_h u_h \\ \alpha (M_h u_h)_1 + \frac{\beta}{(\psi_h - u_h)_1} \\ \cdots \\ \alpha (M_h u_h)_N + \frac{\beta}{(\psi_h - u_h)_N} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \cdot \\ 0 \end{pmatrix} .$$

Solution by path-following predictor-corrector continuation methods in the barrier parameter.

Continuation Strategies in Interior-Point Methods

(i) Static Strategies

Fiacco/McCormick [1990] , Byrd/Hribar/Nocedal [1999]

Wächter/Biegler [2005]

(ii) Dynamic Update Strategies

El-Bakry/Tapia/Tsuchiya/Zhang [1996] , Gay/Overton/Wright [1998]

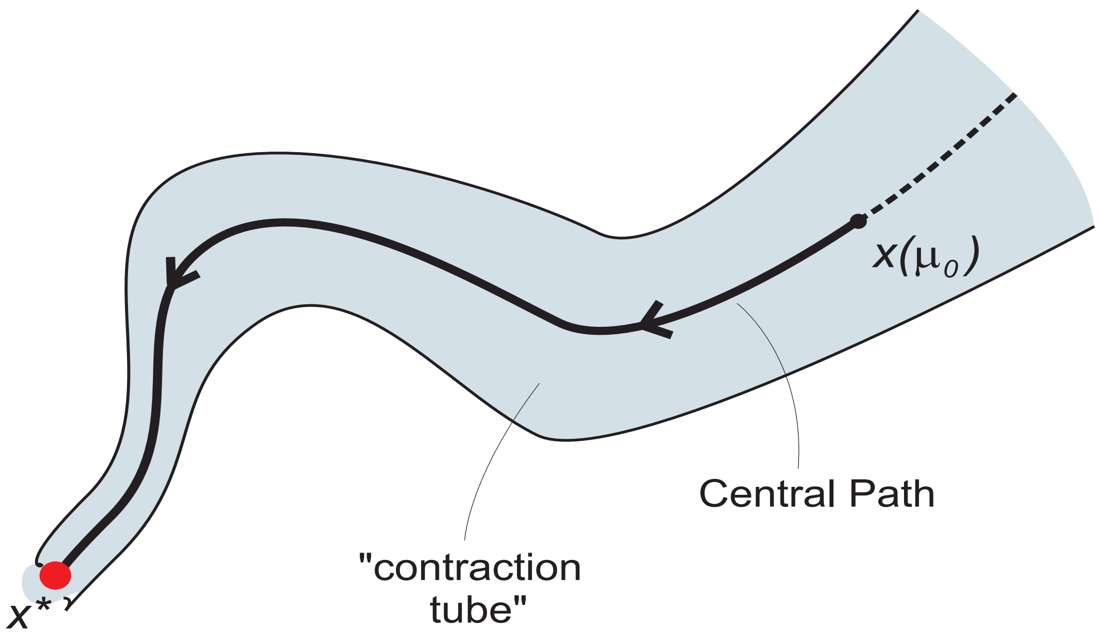
Vanderbei/Shanno [1999] , H./Petrova/Schulz [2002]

Tits/Wächter/Bakhtiari/Urban/Lawrence [2003] , H./Petrova [2004]

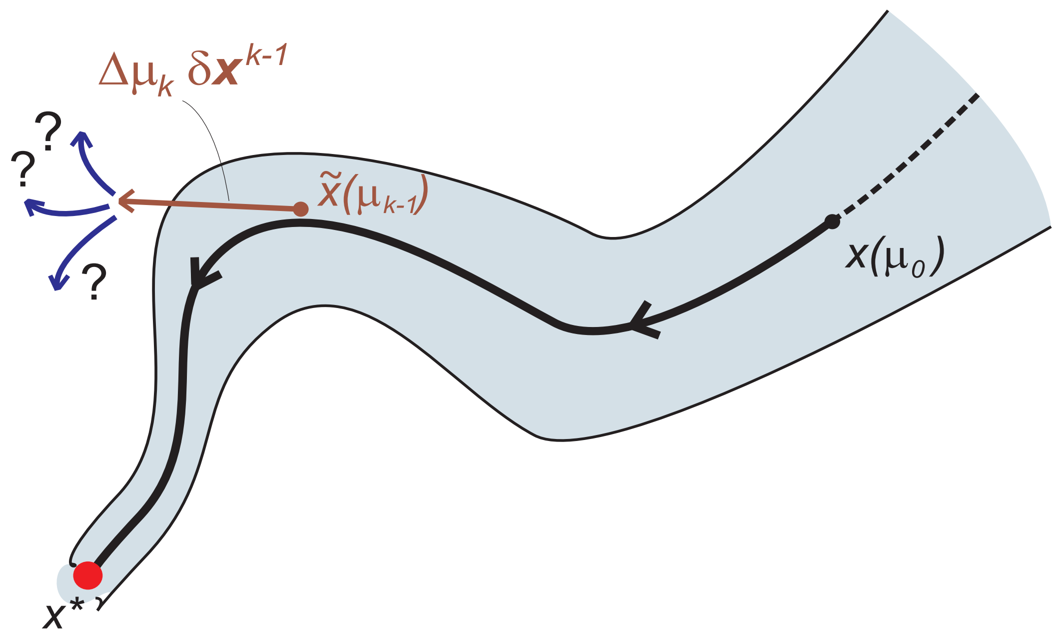
Ulbrich/Ulbrich/Vicente [2004] , H./Linsenmann/Petrova [2006]

Nocedal/Wächter/Waltz [2006] , Armand/Benoist/Orban [2007]

The Barrier Path

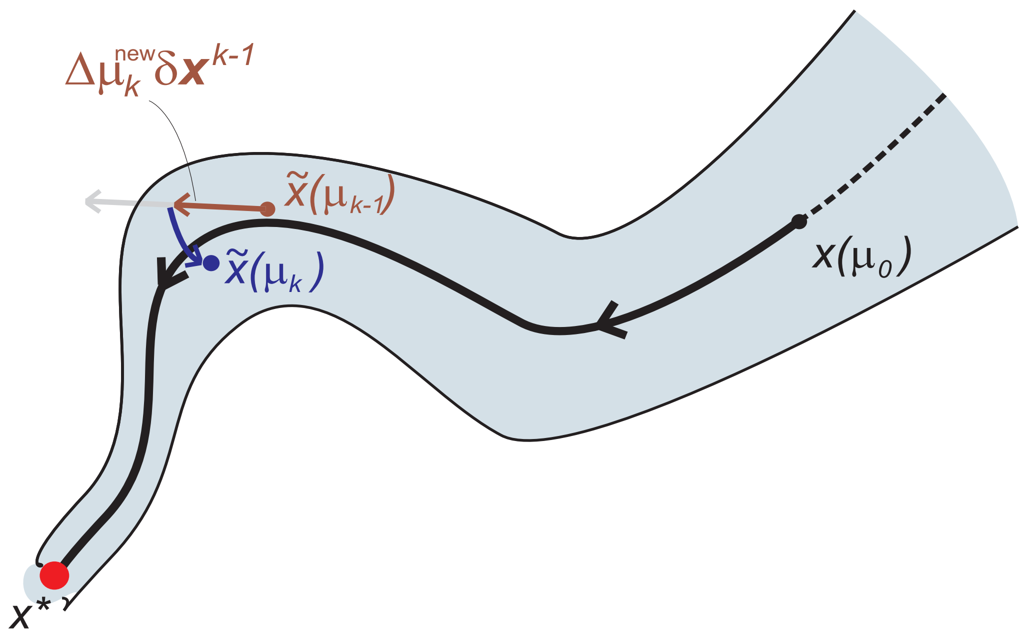


Barrier Path: The Predicted Continuation Step



$\cdots$ may leave the Kantorovich neighborhood.

Barrier Path: The Newton Correction



$\cdots$ reduces the continuation steplength adaptively.

Numerical Results: Distributed Control Problem with Control Constraints I

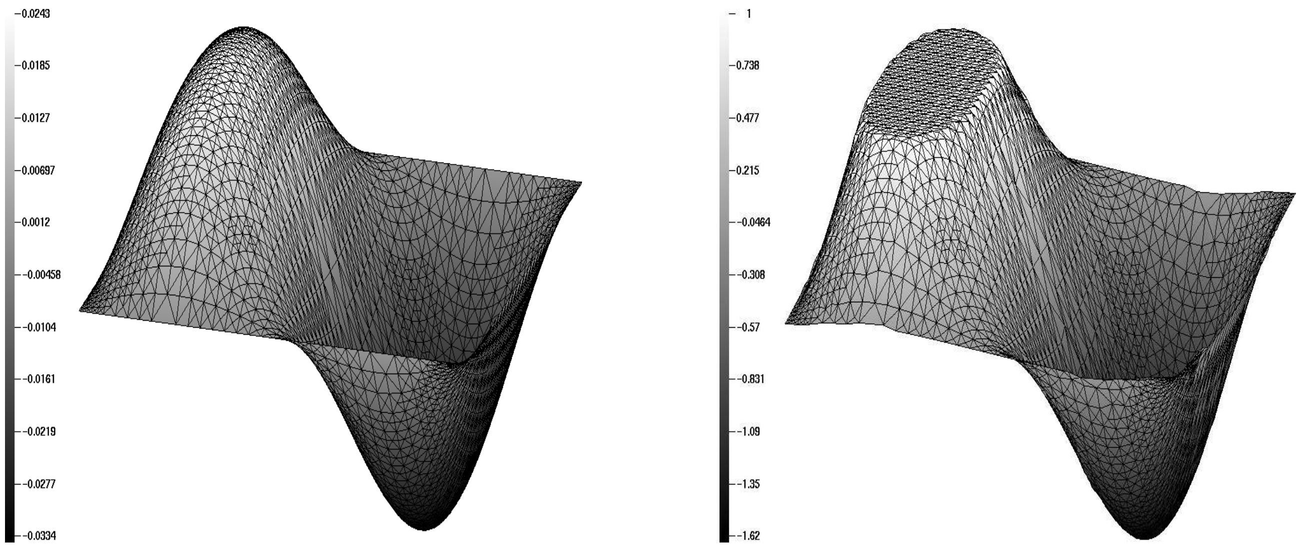
$$
\begin{array}{ll}
\text{Minimize} & J(y,u) := \frac{1}{2}\,\|y - y^d\|^2_{0,\Omega} + \frac{\alpha}{2}\,\|u\|^2_{0,\Omega} \\
\text{over} & (y,u) \in H^1_0(\Omega) \times L^2(\Omega) \\
\text{subject to} & -\Delta y = u \quad \text{in } \Omega\,, \\
 & u \in K := \{v \in L^2(\Omega) \mid v \le \psi \text{ a.e. in } \Omega\}
\end{array}
$$

Data: $\Omega := (0,1)^2$,

$$
y^d := \begin{cases} 200\; x_1\; x_2\; (x_1 - 0.5)^2\; (1 - x_2)\,, & 0 \le x_1 \le 0.5 \\ 200\; (x_1 - 1)\; (x_2\; (x_1 - 0.5)^2\; (1 - x_2)\,, & 0.5 < x_1 \le 1 \end{cases} ,
$$

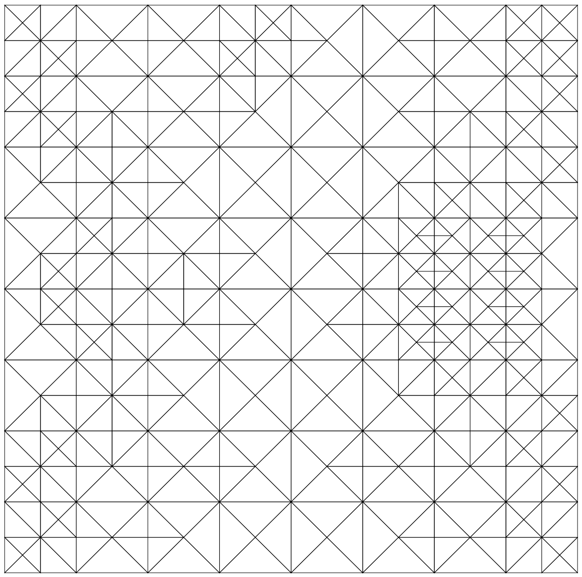
$$\alpha = 0.01 \quad , \quad \psi = 1 \, .$$

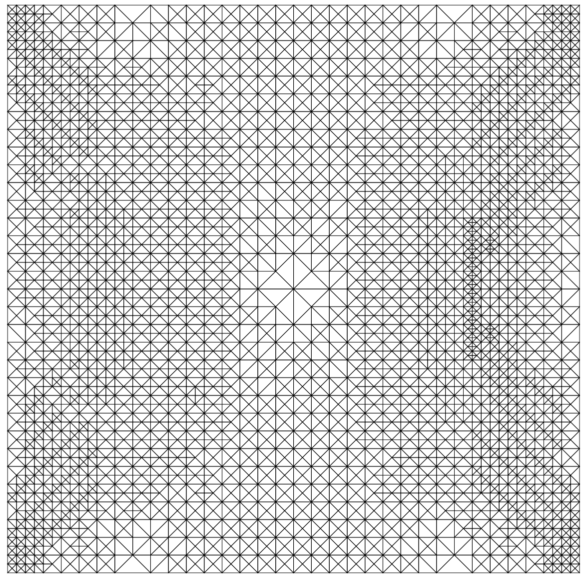
Numerical Results: Distributed Control Problem with Control Constraints I



Optimal state (left) and optimal control (right)

Numerical Results: Distributed Control Problem with Control Constraints I





Grid after 6 (left) and 10 (right) refinement steps

Numerical Results: Distributed Control Problem with Control Constraints II

$$\begin{array}{ll} \text{Minimize} & J(y,u) := \frac{1}{2}\,\|y - y^d\|^2_{0,\Omega} + \frac{\alpha}{2}\,\|u - u^d\|^2_{0,\Omega} \\ \text{over} & (y,u) \in H^1_0(\Omega) \times L^2(\Omega) \\ \text{subject to} & -\Delta y = f + u \quad \text{in } \Omega\,, \\ & u \in K := \{v \in L^2(\Omega) \mid v \le \psi \text{ a.e. in } \Omega\} \end{array}$$

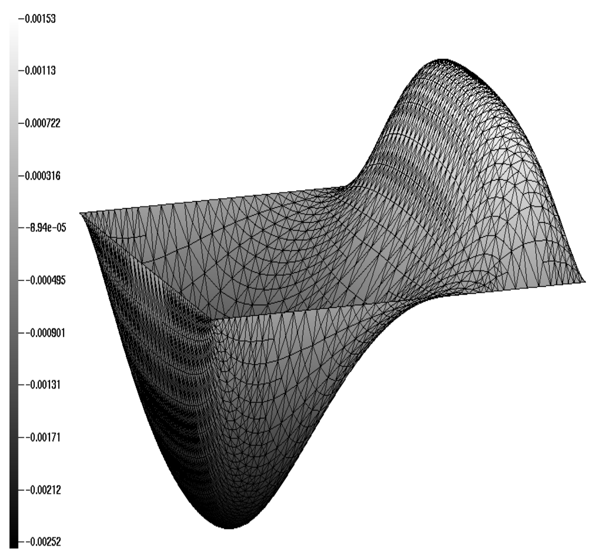
Data: $\Omega := (0,1)^2 \quad , \quad y^d := 0 \quad , \quad u^d := \hat{u} + \alpha^{-1}(\hat{\sigma} - \Delta^{-2}\hat{u})\,,$

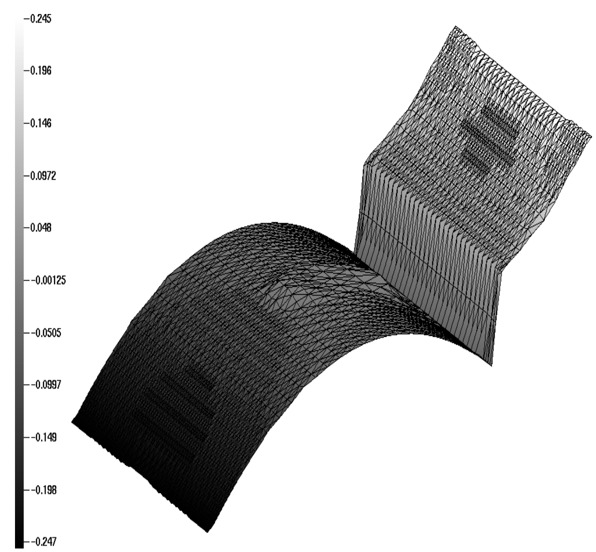
$$\psi := \begin{cases} (x_1 - 0.5)^8 & , \quad (x_1,x_2) \in \Omega_1, \\ (x_1 - 0.5)^2 & , \quad \text{otherwise} \end{cases} \quad , \quad \alpha := 0.1 \quad , \quad f := 0$$

$$\hat{u} := \begin{cases} \psi & , \quad (x_1,x_2) \in \Omega_1 \cup \Omega_2, \\ -1.01\,\psi & , \quad \text{otherwise} \end{cases} \quad , \quad \hat{\sigma} := \begin{cases} 2.25\,(x_1 - 0.75) \cdot 10^{-4} & , \quad (x_1,x_2) \in \Omega_2, \\ 0 & , \quad \text{otherwise} \end{cases} \quad ,$$

$$\Omega_1 := \{(x_1,x_2) \in \Omega \mid \Big((x_1 - 0.5)^2 + (x_2 - 0.5)^2\Big)^{1/2} \le 0.15\} \quad , \quad \Omega_2 := \{(x_1,x_2) \in \Omega \mid x_1 \ge 0.75\}\,.$$

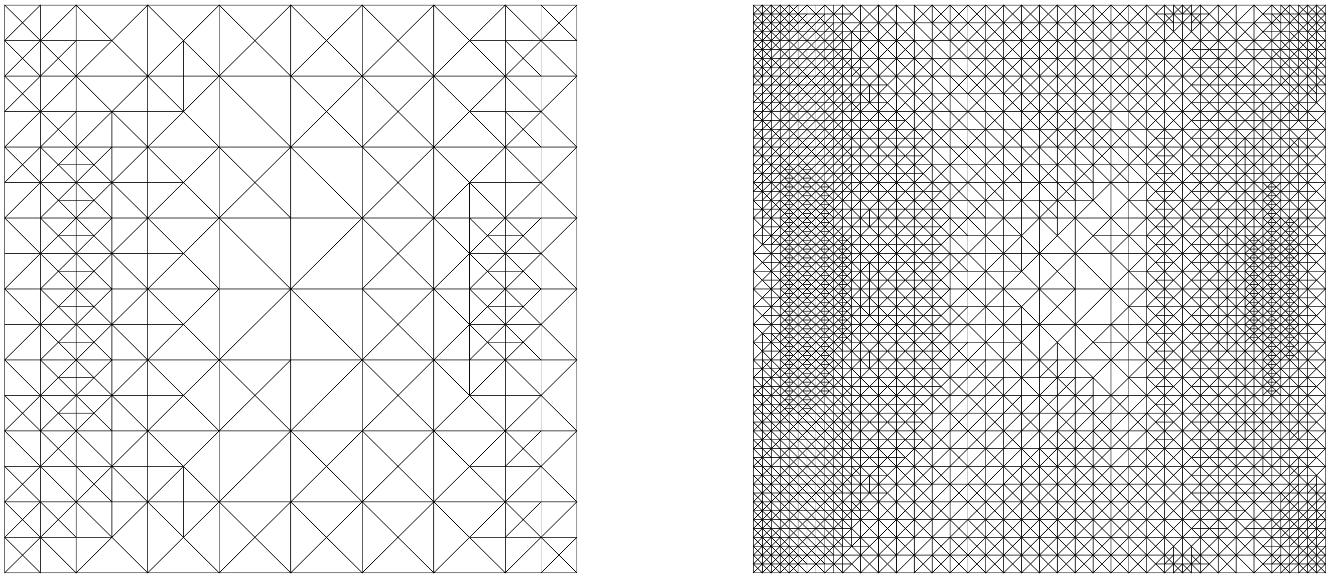
Numerical Results: Distributed Control Problem with Control Constraints II





Optimal state (left) and optimal control (right)

Numerical Results: Distributed Control Problem with Control Constraints II



Grid after 6 (left) and 10 (right) refinement steps

Elliptic Optimal Control Problems

State Constraints

Literature on State-Constrained Optimal Control Problems

M. Bergounioux, K. Ito, and K. Kunisch (1999)
M. Bergounioux, M. Haddou, M. Hintermüller, and K. Kunisch (2000)
M. Bergounioux and K. Kunisch (2002)
E. Casas (1986) E. Casas and M. Mateos (2002)
J.-P. Raymond and F. Tröltzsch (2000)
K. Deckelnick and M. Hinze (2006) M. Hintermüller and K. Kunisch (2007)
K. Kunisch and A. Rösch (2002) C. Meyer and F. Tröltzsch (2006)
C. Meyer, U. Prüfert, and F. Tröltzsch (2005)
U. Prüfert, F. Tröltzsch, and M. Weiser (2004)
H./M. Kieweg (2007) A. Günther, M. Hinze (2007) O. Benedix, B. Vexler (2008)
M. Hintermüller/H. (2008) W. Liu, W. Gong and N. Yan (2008)

Elliptic Optimal Control Problems: State Constraints

Consider the state constrained distributed elliptic optimal control problem

$$\inf_{y,u} J(y,u) := \frac{1}{2}\int_\Omega |y - y^d|^2 \, dx + \frac{\alpha}{2}\int_\Omega |u|^2 \, dx,$$

$$(ESC)_1 \qquad -\Delta y = u \quad \text{in } \Omega,$$

$$(ESC)_2 \qquad y = 0 \quad \text{on } \Gamma,$$

and

$$y \in K_S := \{y \in H_0^1(\Omega) \mid y(x) \le \psi(x) \text{ f.a.a. } x \in \Omega\},$$

where $u \in L^2(\Omega)$.

State Constraints: Necessary Optimality Conditions

Theorem. The state constrained optimal control problem admits a unique solution (y, u) with $y \in W_0^{1,r}(\Omega), r > 2$, and $u \in L^2(\Omega)$.

Theorem. Let us assume that there exists $u_0 \in L^2(\Omega)$ such that the associated solution $y_0 \in W_0^{1,r}(\Omega)$ of the state equation satisfies $y_0 \in \text{int}(K_S)$ (Slater condition).
If (y, u) is the unique solution of the state constrained optimal control problem, there exist $p \in W_0^{1,s}(\Omega), 1/r + 1/s = 1$, and $\lambda \in \mathcal{M}_+(\Omega)$ such that

$$
\begin{aligned}
(\nabla p, \nabla v)_{L^2(\Omega)} &= -(y - y^d, v)_{L^2(\Omega)} + \langle \lambda, v \rangle \quad , \quad v \in W_0^{1,r}(\Omega), \\
p &= \alpha u, \\
\langle \lambda, y - \psi \rangle &= 0.
\end{aligned}
$$

Finite Element Approximation

Let $\mathcal{T}_h(\Omega)$ be a simplicial triangulation of Ω and let

$$V_h := \{ v_h \in C(\overline{\Omega}) \mid v_h|_T \in P_1(T) , \ T \in \mathcal{T}_h(\Omega) , \ v_h|_\Gamma = 0 \}$$

be the FE space of continuous, piecewise linear functions. Let ψ_h be the V_h-interpoland of ψ. Consider the following FE Approximation of the state constrained control problem

$$\begin{aligned} &\text{Minimize} && J_h(y_h, u_h) := \frac{1}{2} \|y_h - y^d\|^2_{0,\Omega} + \frac{\alpha}{2} \|u_h\|^2_{0,\Omega} , \\ &\text{over} && (y_h, u_h) \in V_h \times V_h , \\ &\text{subject to} && (\nabla y_h, \nabla v_h)_{0,\Omega} = (u_h, v_h)_{0,\Omega} , \ v_h \in V_h , \\ & && y_h \in K_{h,S} := \{ v_h \in V_h \mid v_h(x) \le \psi_h(x) , \ x \in \overline{\Omega} \} . \end{aligned}$$

Since the constraints are point constraints associated with the nodal points, the discrete multipliers are chosen from

$$\mathcal{M}_h := \{ \lambda_h \in \mathcal{M}(\Omega) \mid \lambda_h = \sum_{a \in \mathcal{N}_h(\Omega)} \kappa_a \delta_a , \ \kappa_a \in \mathbb{R} \} .$$

Primal-Dual Active Set Strategy I

Step 1 (Initialization):

Choose $\mathbf{c} > \mathbf{0}$, start-iterates $\mathbf{y}_h^{(0)}, \mathbf{u}_h^{(0)}, \boldsymbol{\lambda}_h^{(0)}$ and set $\mathbf{n} = \mathbf{1}$.

Step 2 (Specification of active/inactive sets):

Compute the active/inactive sets $\mathcal{A}_n$ and $\mathcal{I}_n$ according to

$$\mathcal{A}_n := \{\mathbf{1} \leq \mathbf{i} \leq \mathrm{N} \mid (\mathbf{y}_h^{(n-1)} + \mathbf{c}^{-1}\boldsymbol{\lambda}_h^{(n-1)})_i > (\psi_h)_i\} \quad , \quad \mathcal{I}_n := \{\mathbf{1}, \cdots, \mathrm{N}\} \setminus \mathcal{A}_n.$$

Step 3 (Termination criterion):

If $\mathbf{n} \geq \mathbf{2}$ and $\mathcal{A}_n = \mathcal{A}_{n-1}$, stop the algorithm. Otherwise, go to Step 4.

Primal-Dual Active Set Strategy II

Step 4 (Update of the state, adjoint state, control, and multiplier):

Compute $(y_h^{(n)}, u_h^{(n)}, p_h^{(n)}, \lambda_h^{(n)})$ as the solution of

$$\begin{aligned} A_h y_h^{(n)} - M_h u_h^{(n)} &= 0, \\ A_h p_h^{(n)} - (M_h y_h^{(n)} - y_h^d) - \lambda_h^{(n)} &= 0, \\ p_h^{(n)} - \alpha M_h u_h^{(n)} &= 0, \\ (y_h^{(n)})_i &= (\psi_h)_i \quad \text{for } i \in \mathcal{A}_n, \\ (\lambda_h^{(n)})_i &= 0 \quad \text{for } i \in \mathcal{A}_n. \end{aligned}$$

Set $n := n + 1$ and go to Step 2.

Lavrentiev Regularization: Mixed Control-State Constraints

Introduce a regularization parameter $\varepsilon > 0$ and consider the mixed control-state constrained optimal control problem

$$\text{Minimize} \quad J(y^\varepsilon, u^\varepsilon) := \frac{1}{2} \|y^\varepsilon - y^d\|^2_{L^2(\Omega)} + \frac{\alpha}{2} \|u^\varepsilon\|^2_{L^2(\Omega)} ,$$

$$\text{subject to} \quad -\Delta y^\varepsilon = u^\varepsilon \text{ in } \Omega \quad , \quad y^\varepsilon = 0 \text{ on } \Gamma_D,$$

$$\varepsilon u^\varepsilon + y^\varepsilon \in K := \{v \in L^2(\Omega) \mid v(x) \leq \psi(x) \ , \ \text{f.a.a. } x \in \Omega\} \ .$$

Theorem (Optimality conditions). The optimal solution $(y^\varepsilon, u^\varepsilon) \in V \times L^2(\Omega)$ is characterized by the existence of an adjoint state $p^\varepsilon \in V$ and a multiplier $\lambda^\varepsilon \in L^2_+(\Omega)$ such that

$$(\nabla y^\varepsilon, \nabla v)_{L^2(\Omega)} = (u^\varepsilon, v)_{L^2(\Omega)} \quad , \quad v \in V \ ,$$

$$(\nabla p^\varepsilon, \nabla v)_{L^2(\Omega)} = -(y^\varepsilon - y^d, v)_{L^2(\Omega)} + (\lambda^\varepsilon, v)_{L^2(\Omega)} \quad , \quad v \in V \ ,$$

$$p^\varepsilon - \alpha u^\varepsilon + \varepsilon \lambda^\varepsilon = 0 \quad , \quad (\lambda^\varepsilon, \varepsilon u^\varepsilon + y^\varepsilon - \psi)_{L^2(\Omega)} = 0 \ .$$

Numerical Results: Distributed Control Problem with State Constraints I

$$\text{Minimize} \quad J(y,u) := \frac{1}{2}\,\|y - y^d\|^2_{0,\Omega} + \frac{\alpha}{2}\,\|u - u^d\|^2_{0,\Omega} \quad \text{over} \quad (y,u) \in H^1_0(\Omega) \times L^2(\Omega)$$

$$\text{subject to} \quad -\Delta y = u \quad \text{in } \Omega \quad , \quad y \in K := \{v \in H^1_0(\Omega) \mid v \le \psi \text{ a.e. in } \Omega\}$$

Data: $\Omega := (-2,+2)^2 \quad , \quad y^d(r) := y(r) + \Delta p(r) + \sigma(r) \quad , \quad u^d(r) := u(r) + \alpha^{-1} p(r) ,$
$\psi := 0 \quad , \quad \alpha := 0.1 ,$

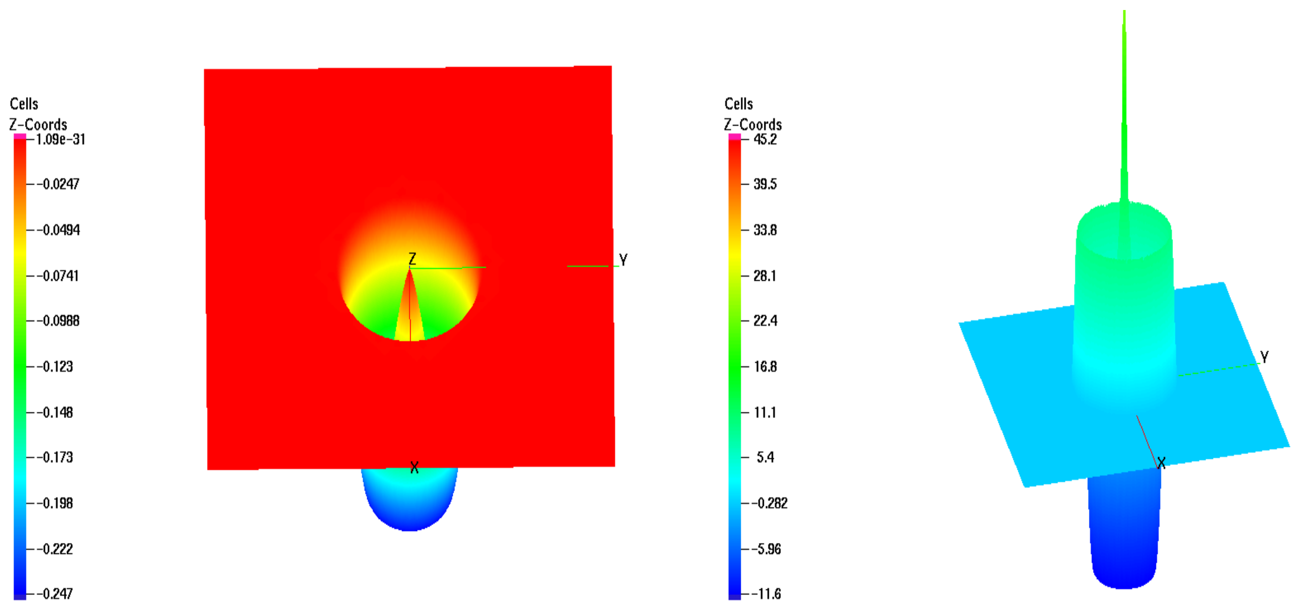
where $y(r), u(r), p(r), \sigma(r)$ is the solution of the problem:

$$y(r) := -r^{4/3} + \gamma_1(r) \ , \ u(r) = -\Delta y(r) \ , \ p(r) = \gamma_2(r) + r^4 - \frac{3}{2}r^3 + \frac{9}{16}r^2 \ , \ \sigma(r) := \begin{cases} 0.0 & , \quad r < 0.75 \\ 0.1 & , \quad \text{otherwise} \end{cases}$$

$$, \qquad \gamma_1 := \begin{cases} 1 & , \quad r < 0.25 \\ -192(r-0.25)^5 + 240(r-0.25)^4 - 80(r-0.25)^3 + 1 & , \quad 0.25 < r < 0.75 \\ 0 & , \quad \text{otherwise} \end{cases} ,$$

$$\gamma_2 := \begin{cases} 1 & , \quad r < 0.75 \\ 0 & , \quad \text{otherwise} \end{cases} .$$

Numerical Results: Distributed Control Problem with State Constraints I



Optimal state (left) and optimal control (right)

Numerical Results: Distributed Control Problem with State Constraints II

$$\text{Minimize} \quad J(y,u) := \frac{1}{2}\,\|y - y^d\|^2_{0,\Omega} + \frac{\alpha}{2}\,\|u - u^d\|^2_{0,\Omega} \quad \text{over} \;\; (y,u) \in H^1(\Omega) \times L^2(\Omega)$$

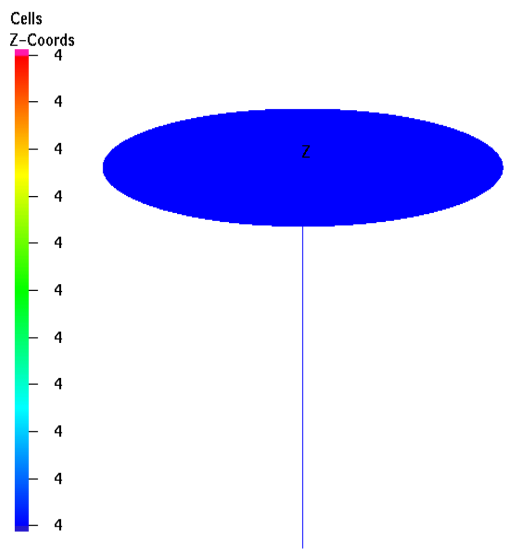
$$\text{subject to} \quad -\,\Delta y + cy = u \;\; \text{in } \Omega \;\; , \quad y \in K := \{v \in H^1(\Omega) \mid v \le \psi \text{ a.e. in } \Omega\}$$

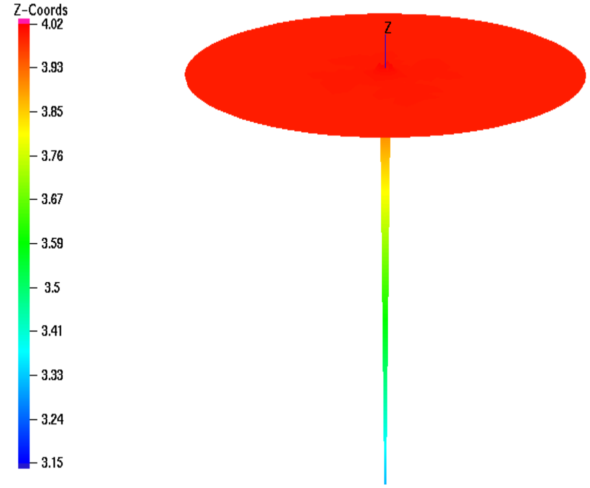
Data: $\Omega = B(0,1) := \{x = (x_1,x_2)^T \mid x_1^2 + x_2^2 < 1\}$, $y^d(r) := 4 + \frac{1}{\pi} - \frac{1}{4\pi} r^2 + \frac{1}{2\pi} \ln(r)$,
$u^d(r) := 4 + \frac{1}{4\pi} r^2 - \frac{1}{2\pi} \ln(r)$, $\psi := 4 + r$, $\alpha := 1$.

The solution $y(r), u(r), p(r), \sigma(r)$ of the problem is given by

$$y(r) \equiv 4 \quad , \quad u(r) \equiv 4 \quad , \quad p(r) = \frac{1}{4\pi} r^2 - \frac{1}{2\pi} \ln(r) \quad , \quad \sigma(r) = \delta_0 \; .$$

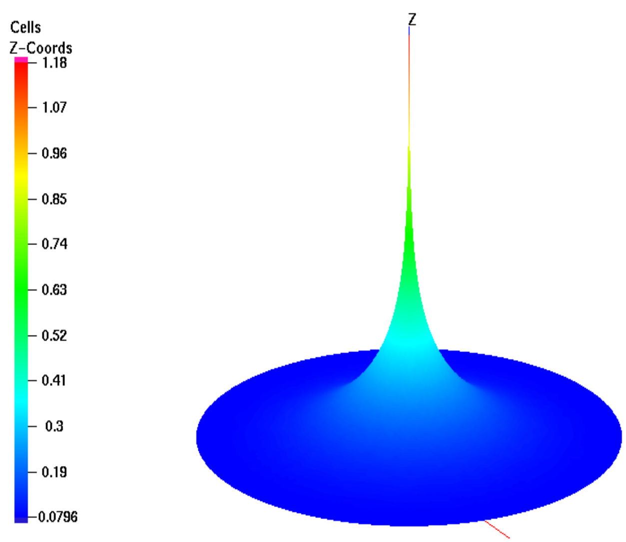
Numerical Results: Distributed Control Problem with State Constraints II

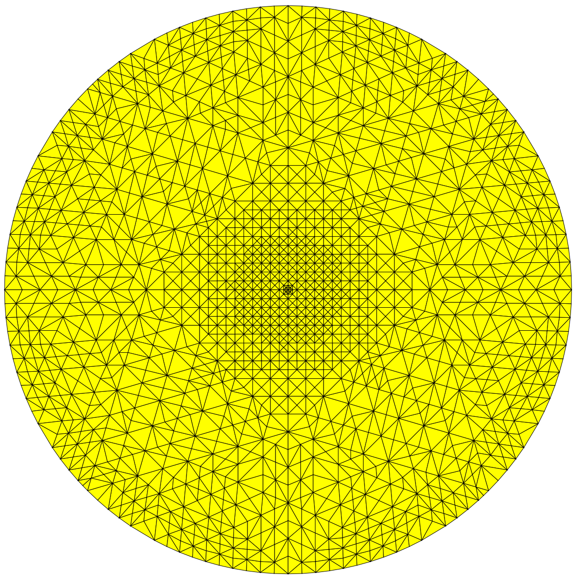




Optimal state (left) and optimal control (right)

Numerical Results: Distributed Control Problem with State Constraints II





Optimal adjoint state (left) and mesh after 16 adaptive loops (right)