Semi-Smooth Newton Methods and their Applications: Part 2

K. Kunisch

Institute for Mathematics and Scientific Computing
University of Graz, Austria
Radon Institute, Austrian Academy of Sciences, Linz, Austria

Woudschoten, October 2017

How to choose γ: Path Following

$$(P_\gamma) \quad \min \tfrac{1}{2}|Au-z|^2 + \tfrac{\beta}{2}|u|^2 + \tfrac{1}{2\gamma}\int_\Omega |(\bar\lambda + \gamma(y-\psi))^+|^2 dx$$

$$\mathcal{P} = \{(y_\gamma, u_\gamma, p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2 \times L^2 \times \mathcal{W}^*\}$$

$(P_{\gamma=0})$ unconstrained, $(P_{\gamma=\infty})$ constrained.

Theorem

$\mathcal{P}$ is globally Lipschitz continuous, and $\gamma \to (p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2$ is locally Lipschitz continuous.

Remark: $\bar\lambda >>$ can guarantee feasibility of y_γ.
For the obstacle problem, choose $\bar\lambda = \max(0, f + \Delta\psi)$.

(H) $\quad S_\gamma^0 := \{x \in \Omega : \ y_\gamma - \psi = 0\}, \quad meas\, S_\gamma^0 = 0$

How to choose γ: Path Following

$$(P_\gamma) \quad \min \frac{1}{2}|Au - z|^2 + \frac{\beta}{2}|u|^2 + \frac{1}{2\gamma}\int_\Omega |(\bar{\lambda} + \gamma(y - \psi))^+|^2 dx$$

$$\mathcal{P} = \{(y_\gamma, u_\gamma, p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2 \times L^2 \times \mathcal{W}^*\}$$

$(P_{\gamma=0})$ unconstrained, $\quad (P_{\gamma=\infty})$ constrained.

Theorem

$\mathcal{P}$ is globally Lipschitz continuous, and $\gamma \to (p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2$ is locally Lipschitz continuous.

Remark: $\bar{\lambda} >>$ can guarantee feasibility of y_γ.
For the obstacle problem, choose $\bar{\lambda} = \max(0, f + \Delta\psi)$.

$(H) \quad S_\gamma^0 := \{x \in \Omega : \; y_\gamma - \psi = 0\}, \quad meas\, S_\gamma^0 = 0$

How to choose γ: Path Following

$$(P_\gamma) \quad \min \tfrac{1}{2}\,|Au - z|^2 + \tfrac{\beta}{2}\,|u|^2 + \tfrac{1}{2\gamma}\int_\Omega |(\bar{\lambda} + \gamma(y - \psi))^+|^2 dx$$

$$\mathcal{P} = \{(y_\gamma, u_\gamma, p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2 \times L^2 \times \mathcal{W}^*\}$$

$(P_{\gamma=0})$ unconstrained, $(P_{\gamma=\infty})$ constrained.

Theorem

$\mathcal{P}$ is globally Lipschitz continuous, and $\gamma \to (p_\gamma, \lambda_\gamma) \in \mathcal{W} \times L^2$ is locally Lipschitz continuous.

Remark: $\bar{\lambda} >>$ can guarantee feasibility of $\underline{y}_\gamma$.
For the obstacle problem, choose $\bar{\lambda} = \max(0, f + \Delta\psi)$.

$$(H) \quad S_\gamma^0 := \{x \in \Omega : \; y_\gamma - \psi = 0\}, \quad meas\; S_\gamma^0 = 0$$

Recall

$$(P_\gamma) \quad \begin{cases} \min \frac{1}{2}\,|y-z|^2 + \frac{\beta}{2}\,|u|^2 + \frac{1}{2\gamma}\int_\Omega |(\bar{\lambda}+\gamma(y-\psi))^+|^2 dx \\ -\Delta y = u, \ \text{ in } \ \Omega, \quad y = 0 \ \text{ on } \partial\Omega. \end{cases}$$

$$(OS_\gamma) \begin{cases} -\Delta y = u \ \text{ in } \Omega, & y = 0 \ \text{ on } \ \partial\Omega \\ -\Delta p + \lambda = -(y-z) \ \text{ in } \Omega, & p = 0 \ \text{ on } \ \partial\Omega \\ \beta u = p \\ \lambda = \max(0, \bar{\lambda} + \gamma(y-\psi)) \end{cases}$$

for the next slides: set $\bar{\lambda} = 0$

Sensitivities for Path Following

Theorem

$\gamma \to (y_\gamma, u_\gamma, p_\gamma) \in \mathcal{W} \times L^2 \times L^2_{weak}$ is differentiable and

$$(OS_\gamma) \quad -\Delta \dot{y}_\gamma = \dot{u}_\gamma, \quad -\Delta \dot{p}_\gamma + (y_\gamma - \psi + \gamma \dot{y}_\gamma)\mathcal{X}_{S_\gamma} = -\dot{y}_\gamma, \quad \beta \dot{u}_\gamma = \dot{p}_\gamma$$

where $S_\gamma = \{x : y_\gamma - \psi > 0\}$.

$$V(\gamma) = \min \; J(y_\gamma, u_\gamma) + \frac{\gamma}{2} \int_\Omega |(y_\gamma - \psi)^+|^2$$

Theorem

$$\dot{V}(\gamma) = \frac{1}{2} \int_\Omega |(y_\gamma - \psi)^+|^2 \,, \qquad \ddot{V}(\gamma) = \int_\Omega (y_\gamma - \psi)^+ \dot{y}$$

Corollary

$$\dot{V}(\gamma) > 0 \,, \quad \ddot{V}(\gamma) < 0 \,, \quad V(0) \triangleq (P_{unconstr.}) \,, \quad V(\infty) \triangleq (P)$$

Sensitivities for Path Following

Theorem

$\gamma \to (y_\gamma, u_\gamma, p_\gamma) \in \mathcal{W} \times L^2 \times L^2_{weak}$ *is differentiable and*

$$(OS_\gamma) \quad -\Delta \dot{y}_\gamma = \dot{u}_\gamma, \quad -\Delta \dot{p}_\gamma + (y_\gamma - \psi + \gamma \dot{y}_\gamma)\mathcal{X}_{S_\gamma} = -\dot{y}_\gamma, \quad \beta \dot{u}_\gamma = \dot{p}_\gamma$$

where $S_\gamma = \{x : y_\gamma - \psi > 0\}$.

$$V(\gamma) = \min\ J(y_\gamma, u_\gamma) + \frac{\gamma}{2} \int_\Omega |(y_\gamma - \psi)^+|^2$$

Theorem

$$\dot{V}(\gamma) = \frac{1}{2} \int_\Omega |(y_\gamma - \psi)^+|^2 \,, \qquad \ddot{V}(\gamma) = \int_\Omega (y_\gamma - \psi)^+ \dot{y}$$

Corollary

$$\dot{V}(\gamma) > 0 \,, \quad \ddot{V}(\gamma) < 0 \,, \quad V(0) \triangleq (P_{unconstr.}) \,, \quad V(\infty) \triangleq (P)$$

Model Function.

$$m(\gamma) = C_1 - \frac{C_2}{E+\gamma} \,.$$

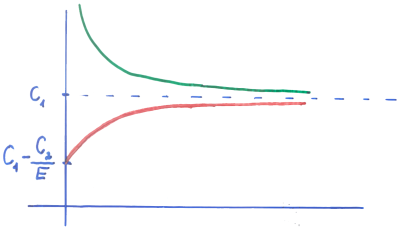
recall: $\dot{m}(\gamma) \sim \dot{V}(\gamma) = \frac{1}{2}\int_\Omega |(y_\gamma - \psi)^+|^2$

test (OS_γ) with $(y_\gamma - \psi)^+$ and "approximate" ∞ dimensional quantities by constants

$$(E+\gamma)\ddot{m}(\gamma) + 2\,\dot{m}(\gamma) = 0$$

$\rightarrow$ model function

$\dot{m} \geq 0,\ \ddot{m} \leq 0,\ \gamma^2 \dot{m}(\gamma)$ bounded for $\gamma \rightarrow \infty$.



Path-following Algorithms

Model parameters

$$m(0) = V(0)\,, \quad m(\gamma) = V(\gamma)\,, \quad \dot{m}(\gamma) = \dot{V}(\gamma) = \frac{1}{2}\int_\Omega |(y_\gamma - \psi)^+|^2$$

determine $E > 0\,,\; C_1 > 0\,,\; C_2 > 0.$

Update Strategy

$$|V^* - V(\gamma_{k+1})| \le \tau_k |V^* - V(\gamma_k)|$$

$$|C_{1,k} - m_k(\gamma_{k+1})| \le \tau_k |C_{1,k} - V(\gamma_k)| =: \beta_k$$

$$\gamma_{k+1} = \frac{C_{2,k}}{\beta_k} - E_k.$$

Theorem (exact path following)

$$\lim_{k\to\infty} (y_{\gamma_k}, u_{\gamma_k}, \lambda_{\gamma_k}) \to (y^*, u^*, \lambda^*).$$

Inexact Path-Following

$$\mathcal{N}(\gamma) = \{(y,\lambda) : |(r_\gamma^1(y,\lambda),\ r_\gamma^2(y,\lambda))|_{\mathbb{R}^2} < \frac{\tau}{\sqrt{\gamma}}\}$$

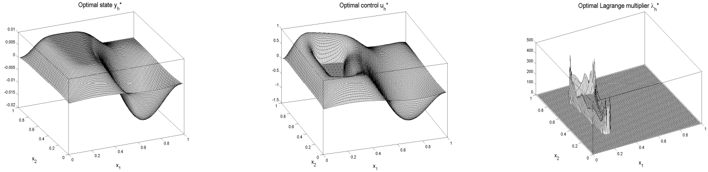
$$r_\gamma^1(y,\lambda) = |\Delta y + \frac{1}{\beta}\Delta^{-1}(\lambda + y - z)|_{H^{-1}}\ ,\ r_\gamma^2(y,\lambda) = |\lambda - \max(0, \bar{\lambda} + \gamma(y-\psi))|_{L^2}.$$

$$\gamma_{k+1} \geq \max\left(\gamma_k \max(\tau_1, \frac{\rho_{k+1}^F}{\rho_{k+1}^C}),\ \frac{1}{\max\left(\rho_{k+1}^F, \rho_{k+1}^C\right)^q}\right),$$

where $q \geq 1\,, \quad \tau_1 > 1$

$$\rho_{k+1}^F := \int_\Omega (y_{k+1} - \psi)^+ dx\ ,\ \rho_{k+1}^C := \int_{\mathcal{I}_{k+1}} (y_{k+1} - \psi)^+\, dx + \int_{\mathcal{A}_{k+1}} (y_{k+1} - \psi)^- dx.$$

Inexact pathfollowing



Optimal state (left), optimal control (middle), and optimal multiplier (right) for problem 1 with $h = 1/128$.

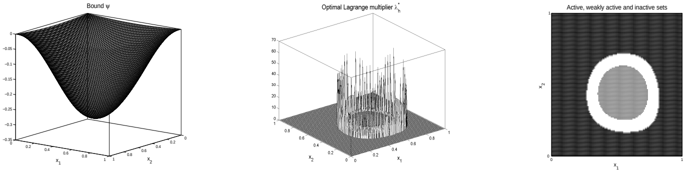
Mesh size h	1/16	1/32	1/64	1/128	1/256
PDAS	14	27	54	113	226
PDIP	12	14	15	19	19
IPF	11	15	14	13	15

Comparison of iteration numbers for different mesh sizes and methods.

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/128	1/256	total
PDAS	3	4	4	5	6	6	6	34
PDIP	3	2	4	4	5	6	7	31
IPF	4	3	3	4	5	5	5	29

Comparison of iteration numbers for different mesh sizes and methods based on nested iteration.

Problem with lack of strict complementarity



bound ψ (left), optimal multiplier (middle), active/inactive sets (right), for $h = 1/128$.

Mesh size h	1/4	1/8	1/16	1/32	1/64	1/128	1/256	total
PDAS	2	4	5	9	10	21	40	91
PDIP	3	2	3	3	6	12	11	40
IPF	7	2	4	4	6	8	15	46

Table: Comparison of iteration numbers for different mesh sizes and methods based on nested iteration.

BUT: CPU-time for PDIP 20 percent higher than for IPF

Some references

- Primal-dual Strategy for Constrained Optimal Control Problems M. BERGOUNIOUX, K. ITO and K. KUNISCH, SIAM J. Control and Optimization, 37(1999), 1176-1194 doi: 10.1137/S0363012997328609
- The primal-dual active set strategy as a semi-smooth Newton method M. HINTERMÜLLER, K. ITO and K. KUNISCH, SIAM Journal on Optimization, 13(2002), 865-888 doi: 10.1137/S1052623401383558
- The primal-dual active set method for nonlinear optimal control problems with bilateral constraints K. ITO and K. KUNISCH, SIAM J. on Control and Optimization, 43(2004), 357-376 doi: 10.1137/S0363012902411015
- Generalized Newton Methods for the 2D-Signorini Contact Problem with Friction K. KUNISCH, G. STADLER, ESAIM: M2AN, 39(2005), 827-854 doi: 10.1051/m2an:2005036
- Feasible and Non-Interior Path-Following in Constrained Minimization with Low Multiplier Regularity M. HINTERMÜLLER and K. KUNISCH, SIAM J. Control and Optim., 45(2006), 1198-1221 doi: 10.1137/050637480

Convex relaxation: motivation

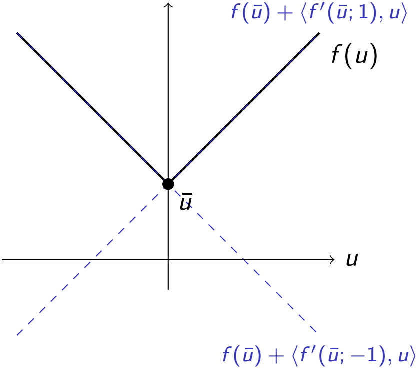
$f : \mathbb{R} \to \mathbb{R}$ not differentiable, convex:

- directional derivative:

$$f'(u; h) = \lim_{t \to 0^+} \frac{f(u + th) - f(u)}{t}$$

- but: for all h,

$$f'(\bar{u}; h) \neq 0$$



Convex relaxation: motivation

$f : \mathbb{R} \to \mathbb{R}$ not differentiable, convex:

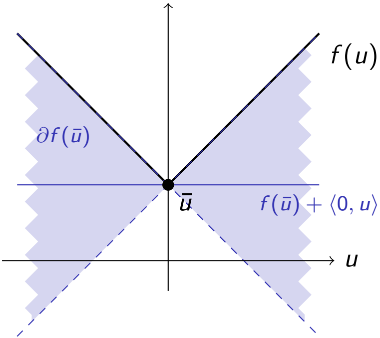
- subdifferential:

  $$\partial f(u) = \{u^* : \langle u^*, h\rangle \leq f'(u; h)\}$$

- geometrically: $\partial f(u)$ set of tangent slopes

- $f(\bar{u}) = \min_u f(u) \Rightarrow 0 \in \partial f(\bar{u})$

- calculus for ∂f



Fenchel duality

$f : V \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ convex, V Banach space, V^* dual space

- subdifferential

$$\partial f(\bar{v}) = \{v^* \in V^* : \langle v^*, v - \bar{v}\rangle_{V^*,V} \leq f(v) - f(\bar{v}) \quad \text{for all } v \in V\}$$

- Fenchel conjugate (always convex)

$$f^* : V^* \to \overline{\mathbb{R}}, \qquad f^*(v^*) = \sup_{v \in V} \langle v^*, v\rangle_{V^*,V} - f(v) \tag{1}$$

- "convex inverse function theorem":

$$v^* \in \partial f(v) \quad \Leftrightarrow \quad v \in \partial f^*(v^*) \tag{2}$$

Fenchel duality: application

$$\mathcal{F}(\bar{u}) + \mathcal{G}(\bar{u}) = \min_u \mathcal{F}(u) + \mathcal{G}(u) \tag{3}$$

1. Fermat principle: $0 \in \partial\left(\mathcal{F}(\bar{u}) + \mathcal{G}(\bar{u})\right)$
2. sum rule: $0 \in \partial\mathcal{F}(\bar{u}) + \partial\mathcal{G}(\bar{u})$, i.e., there is $\bar{p} \in V^*$ with

$$\begin{cases} -\bar{p} \in \partial\mathcal{F}(\bar{u}) \\ \quad \bar{p} \in \partial\mathcal{G}(\bar{u}) \end{cases} \tag{4}$$

3. Fenchel duality:

$$\begin{cases} -\bar{p} \in \partial\mathcal{F}(\bar{u}) \\ \quad \bar{u} \in \partial\mathcal{G}^*(\bar{p}) \end{cases} \tag{5}$$

Regularization

$\mathcal{G}$ non-smooth $\rightsquigarrow$ subdifferential $\partial\mathcal{G}^*$ set-valued $\rightsquigarrow$ regularize

$L^2(\Omega)$ setting

Proximal mapping

$$\mathrm{prox}_{\gamma\mathcal{G}^*}(p) = \arg\min_w \mathcal{G}^*(w) + \frac{1}{2\gamma}\|w - p\|^2$$

- single-valued, Lipschitz continuous
- coincides with resolvent $(\mathrm{Id} + \gamma\partial\mathcal{G}^*)^{-1}(p)$
- (also required for primal-dual first-order methods)

Regularization

Proximal mapping

$$\mathrm{prox}_{\gamma\mathcal{G}^*}(p) = \arg\min_w \mathcal{G}^*(w) + \frac{1}{2\gamma}\|w - p\|^2$$

Complementarity formulation of $u \in \partial\mathcal{G}^*(p)$

$$u = \frac{1}{\gamma}\left((p + \gamma u) - \mathrm{prox}_{\gamma\mathcal{G}^*}(p + \gamma u)\right)$$

- equivalent for every $\gamma > 0$
- single-valued, Lipschitz continuous, implicit

Regularization

Proximal mapping

$$\mathrm{prox}_{\gamma\mathcal{G}^*}(p) = \arg\min_w \mathcal{G}^*(w) + \frac{1}{2\gamma}\|w - p\|^2$$

Moreau–Yosida regularization of $u \in \partial\mathcal{G}^*(p)$

$$u_\gamma = (\partial\mathcal{G}^*)_\gamma(p) = \partial\mathcal{G}^*(I + \gamma\partial\mathcal{G}^*)^{-1}(p) = \frac{1}{\gamma}\left(p - \mathrm{prox}_{\gamma\mathcal{G}^*}(p)\right)$$

- $(\partial\mathcal{G}^*)_\gamma = \partial\left(\mathcal{G} + \frac{\gamma}{2}\|\cdot\|^2\right)^* \to \partial\mathcal{G}^*$ as $\gamma \to 0$
- single-valued, Lipschitz continuous, explicit
  $\rightsquigarrow$ nonsmooth operator equation, Newton method

SMOOTH + CONVEX

$$\begin{cases} \min_{u \in L^2(\Omega)} \dfrac{1}{2}\|y - z\|_{L^2}^2 + \alpha\, \mathcal{G}(u) \\ \quad \text{s.t.}\;\; Ay = u, \end{cases} \tag{6}$$

- $\mathcal{G} : L^2(\Omega) \to \mathbb{R}$ convex
- $z \in L^2(\Omega)$ target (or noisy data)
- $A : V \to V^*$ isomorphism for Hilbert space $V \hookrightarrow L^2(\Omega) \hookrightarrow V^*$

  (e.g., elliptic differential operator with boundary conditions)
- $\rightsquigarrow \mathcal{F}(u) = \frac{1}{2}\|A^{-1}u - z\|_{L^2}^2$ smooth

Optimality system

- let $(\bar{u}, \bar{p}) \in L^2(\Omega) \times L^2(\Omega)$ be a solution
- $S : u \mapsto y$ control-to-state mapping, $\quad S^*$ adjoint

$$\bar{p} = \tfrac{1}{\alpha} S^*(z - S\bar{u}), \qquad \bar{u} \in \partial\mathcal{G}^*(\bar{p}) \tag{7}$$

Regularized optimality system

$$\begin{cases} p_\gamma = \frac{1}{\alpha} S^*(z - Su_\gamma) \\ u_\gamma = (\partial \mathcal{G}^*)_\gamma(p_\gamma) \end{cases} \tag{8}$$

- optimality conditions for

$$\begin{cases} \min_{u \in L^2(\Omega)} \dfrac{1}{2}\|y - z\|_{L^2}^2 + \alpha\, \mathcal{G}(u) + \frac{\gamma}{2}\|u\|^2 \\ \quad \text{s. t.} \quad Ay = u, \end{cases} \tag{9}$$

- $\rightsquigarrow$ unique solution, (u_γ, p_γ) and $(u_\gamma, p_\gamma) \rightharpoonup (\bar{u}, \bar{p})$ as $\gamma \to 0$
- $\mathcal{G} : L^2(\Omega) \to \mathbb{R}, \quad \mathcal{G}(u) = \int_\Omega (g(u(x))\, dx$
- ∂g_γ^* Lipschitz continuous, piecewise C^1, norm gap $V \hookrightarrow L^2(\Omega)$
- $\rightsquigarrow$ semismooth Newton method

Regularized optimality system

$$\begin{cases} A^* p_\gamma = \frac{1}{\alpha}(z - y_\gamma) \\ \quad A y_\gamma = (\partial \mathcal{G}^*)_\gamma(p_\gamma) \end{cases} \tag{8}$$

- optimality conditions for

$$\begin{cases} \min_{u \in L^2(\Omega)} \dfrac{1}{2}\|y - z\|_{L^2}^2 + \alpha\, \mathcal{G}(u) + \frac{\gamma}{2}\|u\|^2 \\ \quad \text{s. t.} \;\; Ay = u, \end{cases} \tag{9}$$

- $\rightsquigarrow$ unique solution, (u_γ, p_γ) and $(u_\gamma, p_\gamma) \rightharpoonup (\bar{u}, \bar{p})$ as $\gamma \to 0$
- $\mathcal{G} : L^2(\Omega) \to \mathbb{R}, \quad \mathcal{G}(u) = \int_\Omega (g(u(x))\, dx$
- ∂g_γ^* Lipschitz continuous, piecewise C^1, norm gap $V \hookrightarrow L^2(\Omega)$
- $\rightsquigarrow$ semismooth Newton method
- introduce $y_\gamma = S u_\gamma$, eliminate $u_\gamma = \mathcal{G}_\gamma^*(p_\gamma)$

Semismooth Newton method

$$\begin{pmatrix} \frac{1}{\alpha}\,\mathrm{Id} & A^* \\ A & -D_N\mathcal{G}_\gamma^*(p) \end{pmatrix} \begin{pmatrix} \delta y \\ \delta p \end{pmatrix} = -\begin{pmatrix} A^*p + \frac{1}{\alpha}(y-z) \\ Ay - \mathcal{G}_\gamma^*(p) \end{pmatrix} \tag{10}$$

- symmetric, but: local convergence
- $\rightsquigarrow$ continuation in $\gamma \to 0$
- $\rightsquigarrow$ pathfollowing, ((or backtracking with line search based on residual norm))

Switching Control

$$\begin{cases} \displaystyle\min_{u\in L^2(0,T;\mathbb{R}^N)} \frac{1}{2}\|y-y^d\|^2_{L^2(0,T;L^2(\omega_{\mathrm{obs}}))} + \frac{\alpha}{2}\int_0^T |u(t)|_2^2\,dt, \\ \text{s.t.} \quad Ly = Bu, \quad y(0) = y_0, \end{cases}$$

where

$$(Bu)(t,x) = \sum_{i=1}^{N} \chi_{\omega_i}(x) u_i(t),$$

promote switching

$$\beta \int_0^T \sum_{\substack{i,j=1\\ i<j}}^{N} |u_i(t)u_j(t)|\,dt$$

for $\beta = \alpha$,

$$\begin{cases} \displaystyle\min_{u\in L^2(0,T;\mathbb{R}^N)} \frac{1}{2}\|y-y^d\|^2_{L^2(0,T;L^2(\omega_{\mathrm{obs}}))} + \frac{\alpha}{2}\int_0^T |u(t)|_1^2\,dt, \\ \text{s.t.} \quad Ly = Bu, \quad y(0) = y_0, \end{cases} \tag{P}$$

Switching Control, Abstract Setting

$$\min_u \mathcal{F}(u) + \mathcal{G}(u),$$

with $\mathcal{F} : L^2(0, T; \mathbb{R}^N) \to \mathbb{R}$ and $\mathcal{G} : L^2(0, T; \mathbb{R}^N) \to \mathbb{R}$ given by

$$\mathcal{F}(u) = \frac{1}{2}\|Su - y^d\|^2_{L^2(0,T;L^2(\omega_{\mathrm{obs}}))}, \qquad \mathcal{G}(u) = \frac{\alpha}{2}\int_0^T |u(t)|_1^2 \, dt,$$

Proposition

The control $\bar{u} \in L^2(0, T; \mathbb{R}^N)$ is a minimizer for (P) *if and only if there exists a $\bar{p} \in L^2(0, T; \mathbb{R}^N)$ such that*

$$\begin{cases} -\bar{p} = \mathcal{F}'(\bar{u}), \\ \quad \bar{u} \in \partial\mathcal{G}^*(\bar{p}), \end{cases} \tag{OS}$$

holds.

Switching Control: Regularized

$$\begin{cases} -p_\gamma = \mathcal{F}'(u_\gamma), \\ \quad u_\gamma = (\partial \mathcal{G}^*)_\gamma(p_\gamma). \end{cases} \tag{OS_γ}$$

$$\mathcal{G}(u) = \int_0^T g(u(t))\,dt,$$

$$g : \mathbb{R}^N \to \mathbb{R}, \qquad g(v) = \frac{\alpha}{2}|v|_1^2.$$

$$g^* : \mathbb{R}^N \to \mathbb{R}, \qquad g^*(q) = \frac{1}{2\alpha}|q|_\infty^2 = \max_{1\le i\le N} \frac{1}{2\alpha} q_i^2.$$

$$[\partial \mathcal{G}_\gamma^*(p)](t) = \partial g_\gamma^*(p(t)) = \frac{1}{\gamma}\left(p(t) - \mathrm{prox}_{\gamma g^*}(p(t))\right) \quad \text{for a.e. } t \in (0, T).$$

$$[\mathrm{prox}_{\gamma g^*}(v)]_j = \begin{cases} sign(v_j)\frac{\alpha}{d\alpha+\gamma}\sum_{i=1}^d |v_i| & \text{if } j \le d, \\ v_j & \text{if } j > d. \end{cases} \tag{11}$$

where $v \in \mathbb{R}^N$ be sorted by decreasing magnitude and d be the smallest index for which

$$|v_{d+1}| < \frac{\alpha}{d\alpha + \gamma}\sum_{i=1}^d |v_i|$$

Switching Control: Regularized

$$\begin{cases} -p_\gamma = \mathcal{F}'(u_\gamma), \\ \quad u_\gamma = (\partial \mathcal{G}^*)_\gamma(p_\gamma). \end{cases} \tag{OS$_\gamma$}$$

$$\mathcal{G}(u) = \int_0^T g(u(t))\,dt,$$

$$g : \mathbb{R}^N \to \mathbb{R}, \qquad g(v) = \frac{\alpha}{2}|v|_1^2.$$

$$g^* : \mathbb{R}^N \to \mathbb{R}, \qquad g^*(q) = \frac{1}{2\alpha}|q|_\infty^2 = \max_{1\le i\le N} \frac{1}{2\alpha} q_i^2.$$

$$[\partial g_\gamma^*(q)]_j = \begin{cases} \frac{1}{\alpha+\gamma} q_j & \text{if } |q_j| = \max_i |q_i| \text{ and } |q_j| > \left(1 + \frac{\gamma}{\alpha}\right)|q_i|, i \neq j, \\ 0 & \text{if } |q_j| < \max_i |q_i|. \end{cases}$$

Switching Control: Newton Step

$$\begin{cases} y_\gamma = S(\partial\mathcal{G}_\gamma)^*(p_\gamma), \\ p_\gamma = -S^*(y_\gamma - y^d), \end{cases}$$

$$[D_N h_\gamma(q)]_{ji} = \begin{cases} \frac{(d-1)\alpha+\gamma}{\gamma(d\alpha+\gamma)} & \text{if } j = i \leq d, \\ -\frac{\alpha}{\gamma(d\alpha+\gamma)} sign(q_j q_i) & \text{if } j \leq d,\ i \leq d,\ i \neq j, \\ 0 & \text{if } j > d \text{ or } i > d, \end{cases}$$

$$\begin{cases} \delta y - S_0 D_N H_\gamma(p^k)\delta p = -y^k + SH_\gamma(p^k), \\ \qquad\qquad \delta p + S^*\delta y = -p^k - S^*(y^k - y^d), \end{cases}$$

$$\delta p + S^* S_0 D_N H_\gamma(p^k)\delta p = -\left(p^k + S^*(SH_\gamma(p^k) - y^d)\right)$$

S_0 homogenous initial and boundary conditions

Switching Control: Numerical Example

$$y_t - \Delta y = \sum_{i=1}^{N} \chi_{\omega_i}(x) u_i(t) \quad \text{on } \Omega_T,$$

$$y^d = \sum_{i=1}^{N} \cos(i+t) \sin^2\left(2\pi \frac{t}{T}\right) |x - x_i|^2.$$

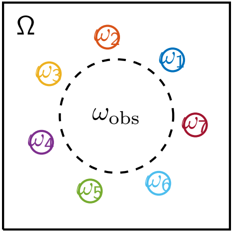


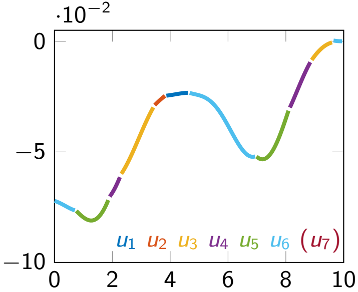
Figure: Problem setting for $N = 7$ control components

Discretization:

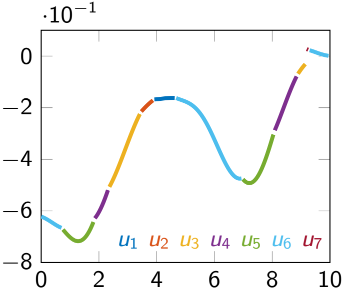
- linear FE in space, cG(1) Petrov-Galerkin method in time
- 200 temporal degrees of freedom per control component

Switching Control: Numerical Example

Optimal controls for $N = 7$ and different α



(a) $\alpha = 10^{-1}$ ($\bar{\gamma} = 10^{-12}$, $\tau_1 = 200$)



(b) $\alpha = 10^{-2}$ ($\bar{\gamma} = 10^{-12}$, $\tau_1 = 200$)

Switching Control: Numerical Example

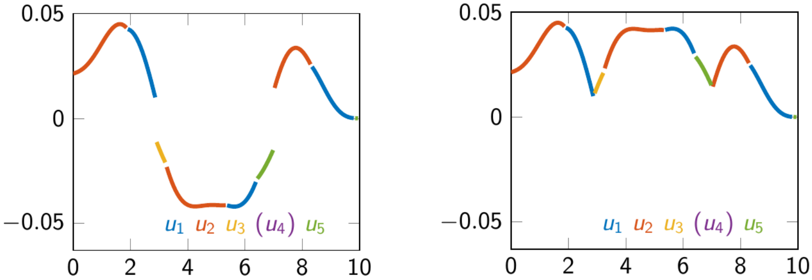
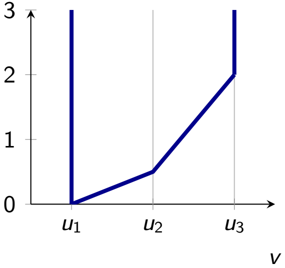
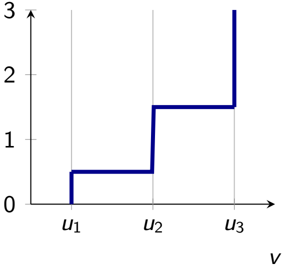


Figure: $N = 5$, $\alpha = 10^{-1}$: optimal controls (left) and their absolute value (right)

Multi-bang penalty: sketch



g ($u_1 = 0$, $u_2 = 1$, $u_3 = 2$)



∂g ($u_1 = 0$, $u_2 = 1$, $u_3 = 2$)

Multi-bang penalty

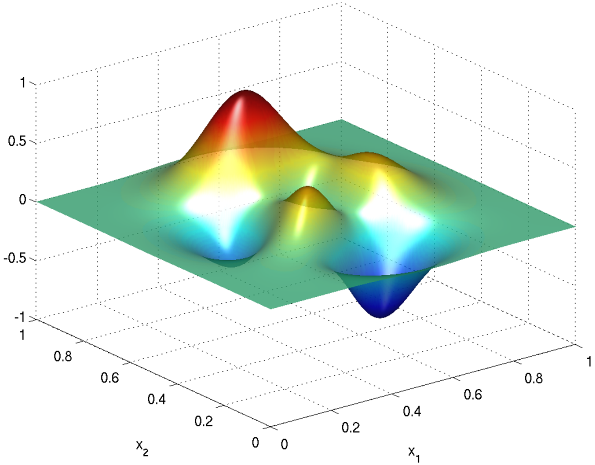
$$\begin{cases} \min\limits_{u,y} \dfrac{1}{2}\|y-z\|_{L^2}^2 + \dfrac{\alpha}{2}\|u\|_{L^2}^2 + \beta \displaystyle\int_\Omega \prod_{i=1}^d |u(x)-u_i|^0 \, dx \\ \text{s. t. } \; Ay = u, \quad u_1 \leq u(x) \leq u_d \; \text{ for almost every } x \in \Omega \end{cases}$$

- U discrete set, $U = \{u \in L^2(\Omega) : u(x) \in \{u_1, \dots, u_d\} \text{ a.e.}\}$
- $u_1, \dots, u_d$ given voltages, velocities, materials, ...
  (assumed here: ranking by magnitude possible!)
- motivation: topology optimization, medical imaging

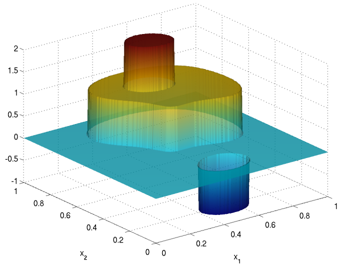
Numerical example

- $\Omega = [0,1]^2, \quad A = -\Delta$
- finite element discretization: uniform grid, 256×256 nodes
- state, adjoint: piecewise linear
- parameter: eliminated (variational discretization)
- $d = 5, \quad (u_1, \ldots, u_5) = (-2, 1, 0, 1, 2)$

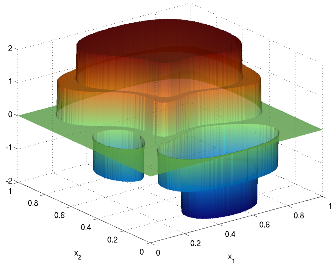
Numerical examples: desired state



Multi-bang controls



$\alpha = 5 \cdot 10^{-3}$ ($\gamma = 0$)



$\alpha = 10^{-3}$ ($\gamma \approx 10^{-7}$)

Time-optimal Control

$$\begin{cases} \min_{\tau \geq 0} \int_0^\tau dt \\ \text{subject to} \\ \frac{d}{dt}x(t) = Ax(t) + Bu(t), \\ |u(t)|_{\ell^\infty} \leq 1,\ x(0) = x_0,\ x(\tau) = x_1, \end{cases} \qquad (P)$$

Hamiltonian $\qquad H(x, u, p_0, p) = p_0 + p^T(Ax + Bu),$

$$\begin{cases} \dot{x} = Ax + Bu,\ x(0) = x_0,\ x(\tau) = x_1, \\ -\dot{p} = A^T p, \\ u = \operatorname{argmin}_{|v|_{\ell^\infty} \leq 1} H(x, v, p_0, p),\ \text{a.e. in } (0, \tau), \\ p_0 + p(t)^T(Ax(t) + Bu(t)) = 0,\ p_0 \geq 0, \end{cases}$$

$$p(t) = \exp\left(A^T(\tau - t)\right) q, \qquad q \neq 0$$

Time-optimal Control: Optimality System

$$\begin{cases} \dot{x} = Ax + Bu, \; x(0) = x_0, \; x(\tau) = x_1, \\ -\dot{p} = A^T p, \\ u = \operatorname{argmin}_{|v|_{\ell^\infty} \le 1} H(x, v, p_0, p), \; \text{a.e. in } (0, \tau), \\ p_0 + p(t)^T (Ax(t) + Bu(t)) = 0, \; p_0 \ge 0, \end{cases} \qquad (tOS)$$

$$u_i = -\sigma(b_i^T p) = -\sigma\left(b_i^T \exp\left(-A^T(\tau - t)\right) q\right),$$

$$\sigma(s) \; \in \; \begin{cases} -1 & \text{if } \; s < 0 \\ [-1, \, 1] & \text{if } \; s = 0 \\ 1 & \text{if } \; s > 0. \end{cases}$$

lack of smoothness

Time-optimal Control: Regularization

$$\begin{cases} \min_{\tau \geq 0} \int_0^\tau (1 + \frac{\varepsilon}{2} |u(t)|^2)\, dt \\ \text{subject to} \\ \frac{d}{dt}x(t) = Ax(t) + Bu(t), \\ |u(t)|_{\ell^\infty} \leq 1,\ x(0) = x_0,\ x(\tau) = x_1. \end{cases} \qquad (P_\varepsilon)$$

Behavior $\varepsilon \to 0$ is well-understood.

$$\text{There exists } i^* \text{ such that } (A, b_{i^*}) \text{ is controllable.} \qquad \text{(H1)}$$

Time-optimal Control: Regularized Optimality System

Theorem

Assume that (H1) holds and let $(x_\varepsilon, u_\varepsilon, \tau_\varepsilon)$ be a solution of (P_ε). If there exist $\eta > 0$ and an interval $I_{i^*} \subset (0,1)$ such that

$$|\,(\hat{u}_\varepsilon)_{i^*}(t)|_{\ell^\infty} \leq 1 - \eta \;\; \text{for a.e. } t \in I_{i^*}, \tag{H2}$$

then there exists an adjoint state p_ε such that

$$\begin{cases} \dot{x}_\varepsilon = Ax_\varepsilon + Bu_\varepsilon,\; x_\varepsilon(0) = x_0,\; x_\varepsilon(\tau_\varepsilon) = x_1 \\ -\dot{p}_\varepsilon = A^T p_\varepsilon \\ u_\varepsilon = -\sigma_\varepsilon(B^T p_\varepsilon) \\ 1 + \frac{\epsilon}{2}|u_\varepsilon(\tau_\varepsilon)|^2_{\mathbb{R}^m} + p_\varepsilon(\tau_\varepsilon)^T(Ax_\varepsilon(\tau_\varepsilon) + Bu_\varepsilon(\tau_\varepsilon)) = 0. \end{cases} \tag{11}$$

where

$$\sigma_\varepsilon(s) \;\in\; \begin{cases} -1 & \text{if } s \leq -\varepsilon \\ \frac{s}{\varepsilon} & \text{if } |s| < \varepsilon \\ 1 & \text{if } s \geq \varepsilon. \end{cases}$$

Time-optimal Control: Time Transformation

Under the transformation $t \to \frac{t}{\tau}$ the first order necessary optimality condition for (P_ε)

$$\begin{cases} \dot{x} = \tau(Ax + Bu), \; x(0) = x_0, \; x(1) = x_1 \\ -\dot{p} = \tau A^T p \\ u = -\sigma_\varepsilon(B^T p) \\ 1 + \frac{\epsilon}{2}|u(1)|^2 + p(1)^T(Ax(1) + Bu(1)) = 0, \end{cases} \tag{12}$$

Time-optimal Control: Semi-Smooth Newton Method

$$F(x,\, p,\, u,\, \tau) = \begin{pmatrix} \dot{x} - \tau A x - \tau B u \\ -\dot{p} - \tau A^T p \\ u + \sigma_\varepsilon(B^T p) \\ x(1) - x_1 \\ 1 + \frac{\varepsilon}{2}|u(1)|^2 + p(1)^T\,(Ax(1) + Bu(1)) \end{pmatrix}. \qquad (13)$$

$$F : D_F \to L^2(0,1;\mathbb{R}^n) \times L^2(0,1;\mathbb{R}^n) \times U \times \mathbb{R}^n \times \mathbb{R}$$

where

$$D_F = W^{1,2}(0,1;\mathbb{R}^n) \times \mathcal{U}_{p_\varepsilon} \times U \times \mathbb{R},$$

$$\mathcal{U}_{p_\varepsilon} \subset W^{1,2}(0,1;\mathbb{R}^n), \quad U = \{u \in L^2(0,\,1;\mathbb{R}^m) : u|I_{i^*} \in W^{1,2}(I_{i^*};\mathbb{R}^m)\}$$

Time-optimal Control: Convergence

$$|b_i^T p_\varepsilon(1)| \neq \varepsilon, \text{ for all } i = 1, \ldots, m. \tag{H3}$$

$$\text{The scalar-valued Schur complement for } \tau \text{ is nontrivial.} \tag{H4}$$

Theorem

If (H1)–(H4) *hold and* $(x_\varepsilon, u_\varepsilon, \tau_\varepsilon)$ *denotes a solution to* (P_ε) *with associated adjoint* p_ε*, then the semi-smooth Newton algorithm converges superlinearly, provided that the initialization is sufficiently close to* $(x_\varepsilon, p_\varepsilon, u_\varepsilon, \tau_\varepsilon)$.

Canonical Example

c	1	10	50	100	200
No. of iterations	5	46	4	4	3
Final Time	11.1088	10.6092	10.6034	10.6033	10.6031

Table 2

Some references

- A convex penalty for switching control of partial differential equations C. CLASON, A. RUND, K. KUNISCH and R. BARNARD. Systems & Control Letters
- A convex analysis approach to optimal controls with switching structure for partial differential equations C. CLASON, K. ITO and K. KUNISCH. ESAIM: Control, Optimisation and Calculus of Variations
- A convex analysis approach to multi-material topology optimization C. CLASON and K. KUNISCH. ESAIM: M2AN 50 (2016), 1917-1936
- Semi-smooth Newton Methods for Time-Optimal Control for a Class of ODEs K. ITO and K. KUNISCH. SIAM Journal on Control and Optimization, 48 (2010), 3997