Nonsmooth optimization of partial differential equations

Christian Clason

Faculty of Mathematics, Universität Duisburg-Essen

joint work with Karl Kunisch, Kazufumi Ito, Armin Rund, Rick Barnard

XII Conference Approximation and Optimization in the Caribbean
La Habana, June 7, 2016

Motivation: discrete optimization

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

L^0 penalty

$$\|u\|_0 := \int_\Omega |u(x)|_0 \, dx \qquad |t|_0 := \begin{cases} 0 & \text{if } t = 0 \\ 1 & \text{if } t \neq 0 \end{cases}$$

- Lebesgue measure of support of u
- popular in sparse optimization
- binary penalty $\rightsquigarrow$ combinatorial optimization
- difficulty: non-smooth, non-convex, not weakly l.s.c.
- not a norm $\rightsquigarrow$ not coercive

Binary penalties

UNIVERSITÄT DUISBURG ESSEN Open-Minded

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

- $\mathcal{F}(u)$ tracking or discrepancy term

1 $\mathcal{G}(u)$ sparsity penalty [Ito, Kunisch 2012]

$$\mathcal{G}(u) = \frac{\alpha}{2}\|u\|_{L^2}^2 + \beta\|u\|_0$$

- $\rightsquigarrow u(x) = 0$ almost everywhere
- separate penalization of support (β), magnitude (α)
- $\rightsquigarrow \alpha > 0$ necessary!

Binary penalties

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

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

2 $\mathcal{G}(u)$ multi-bang penalty [Clason, Kunisch 2013]

$$\mathcal{G}(u) = \int_\Omega \frac{\alpha}{2}|u(x)|^2 + \beta \prod_{i=1}^{d} |u(x) - u_i|_0 \, dx$$

- $\rightsquigarrow u(x) \in \{u_1, \ldots, u_d\}$ almost everywhere
- motivation: discrete control (voltages, velocities, materials)
- $\beta > 0$ large penalizes *free arc* $u(x) \neq u_i$
- $\alpha > 0$ penalizes magnitude of $u(x) = u_i$

Binary penalties

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

3 $\mathcal{G}(u)$ switching penalty, $u = (u_1, u_2)$ [Clason, Ito, Kunisch 2014]

$$\mathcal{G}(u) = \int_0^T \frac{\alpha}{2}|u(t)|_2^2 + \beta|u_1(t)u_2(t)|_0 \, dt$$

- $\rightsquigarrow u_1(t)u_2(t) = 0$ almost everywhere
- motivation: only one active control component
- $\beta > 0$ large penalizes free arc $u_1 \neq 0$ and $u_2 \neq 0$
- $\alpha > 0$ penalizes magnitude of active u_i

1 Overview

2 Approach
- Convex relaxation
- Moreau–Yosida regularization
- Semismooth Newton method

3 Multi-bang penalty
- Optimality system
- Numerical solution
- Examples

4 Switching penalty
- Optimality system
- Numerical solution
- Examples

Convex relaxation: motivation

$f : \mathbb{R} \to \mathbb{R}$ differentiable:

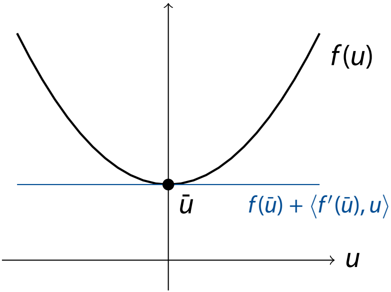
- derivative:

$$f'(u) = \lim_{h \to 0} \frac{f(u+h) - f(u)}{h}$$

- geometrically:
  $f'(u)$ tangent slope

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

- calculus for f'



Convex relaxation: motivation

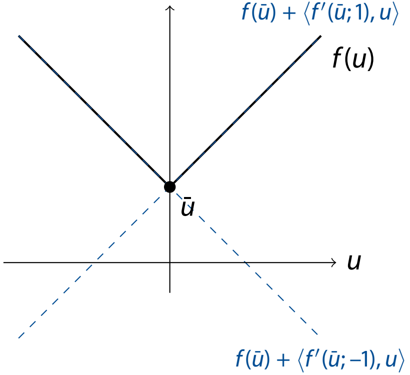
$f : \mathbb{R} \to \mathbb{R}$ not differentiable, convex:

- 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:

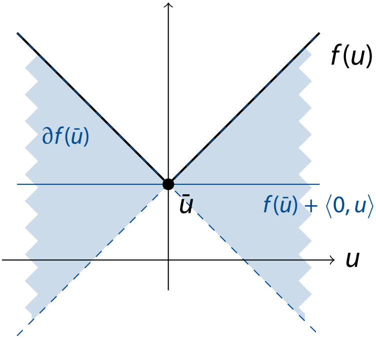
- subdifferential:

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

- 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

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

- subdifferential

$$\partial F(\bar{v}) = \left\{ v^* \in V^* : \langle v^*, v - \bar{v} \rangle_{V^*,V} \leqslant F(v) - F(\bar{v}) \quad \text{for all } v \in V \right\}$$

- Fenchel conjugate (always convex)

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

- connection:

$$v^* \in \partial F(v) \quad \Leftrightarrow \quad v \in \partial F^*(v^*)$$

"convex inverse function theorem"

Fenchel duality: application

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

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

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

3 Fenchel duality:

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

Fenchel duality: example

- $\mathcal{G} : V \to \mathbb{R}, \quad v \mapsto \|v\|_V$:

$$\mathcal{G}^* : V^* \to \overline{\mathbb{R}}, \quad v^* \mapsto \delta_{\{\|\cdot\|_{V^*} \leqslant 1\}}(v^*) := \begin{cases} 0 & \text{if } \|v^*\|_{V^*} \leqslant 1 \\ \infty & \text{else} \end{cases}$$

- $\mathcal{G} : V \to \overline{\mathbb{R}}, \quad v \mapsto \delta_{\{\|\cdot\|_V \leqslant 1\}}(v)$:

$$\partial \mathcal{G}(\bar{v}) = \left\{ v^* \in V^* : \langle v^*, v - \bar{v} \rangle_{V^*,V} \leqslant 0 \quad \text{for all} \quad \|v\|_V \leqslant 1 \right\}$$

$\rightsquigarrow$ box-constrained optimization

Convex relaxation

Consider $\mathcal{F}$ convex, $\mathcal{G}$ non-convex

$$\mathcal{J}(\bar{u}) := \mathcal{F}(\bar{u}) + \mathcal{G}(\bar{u}) = \min_u \mathcal{F}(u) + \mathcal{G}(u)$$

Optimality(?) conditions:

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

- Fenchel conjugate always convex, weakly l.s.c.
- $\rightsquigarrow$ well-defined, unique solution $\bar{u}$ (minimizes $\mathcal{F}(u) + \mathcal{G}^{**}(u)$)
- but: $\bar{u}$ in general not minimizer of $\mathcal{J}$ $\rightsquigarrow$ sub-optimal

Regularization

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

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

$u, p \in L^2$ Hilbert space $\rightsquigarrow$ consider for $\gamma > 0$

Proximal mapping

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

- single-valued, Lipschitz continuous
- coincides with resolvent $(\mathrm{Id} + \gamma\partial\mathcal{G}^*)^{-1}(p)$

Regularization

Proximal mapping

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

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

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

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

Regularization

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

Proximal mapping

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

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

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

- $\partial\mathcal{G}^*_\gamma(p) \to \partial\mathcal{G}^*(p)$ as $\gamma \to 0$
- single-valued, Lipschitz continuous, explicit
  $\rightsquigarrow$ nonsmooth operator equation, Newton method

Regularization: example

$$\mathcal{G}^* : V^* \to \overline{\mathbb{R}}, \quad p \mapsto \delta_{\{\|\cdot\|_{V^*} \leqslant 1\}}(p):$$

- Proximal mapping:

$$\mathrm{prox}_{\gamma \mathcal{G}^*}(p) = \mathrm{proj}_{\{\|\cdot\|_{V^*} \leqslant 1\}}(p)$$

- Moreau–Yosida regularization ($V^* = L^\infty(\Omega)$):

$$\partial \mathcal{G}^*_\gamma(p) = \frac{1}{\gamma}\Big(\max(0, p-1)) + \min(0, p+1)\Big)$$

(max, min pointwise almost everywhere)

Generalized Newton method

Consider Banach spaces X, Y, mapping $F : X \to Y$

Newton-type method for $F(x) = 0$

- choose $x^0 \in X$ (close to solution x^*)
- for $k = 0, 1, \ldots$
  1. choose $M_k \in \mathcal{L}(X, Y)$ invertible
  2. solve for s^k:

$$M_k s^k = -F(x^k)$$

  3. set $x^{k+1} = x^k + s^k$

Generalized Newton method

Newton-type method for $F(x) = 0$

- choose $x^0 \in X$ (close to solution x^*)
- for $k = 0, 1, \ldots$
  1. choose $M_k \in \mathcal{L}(X, Y)$ invertible
  2. solve for s^k:
  $$M_k s^k = -F(x^k)$$
  3. set $x^{k+1} = x^k + s^k$

- convergence, i.e., $\|x^k - x^*\|_X \to 0$?
- superlinear convergence, i.e., $\|x^{k+1} - x^*\|_X = o\left(\|x^k - x^*\|_X\right)$?

Convergence of Newton method

Set $d^k = x^k - x^*$ $\rightsquigarrow$

$$\|x^{k+1} - x^*\|_X = \|M_k^{-1}(F(x^* + d^k) - F(x^*) - M_k d^k)\|_X$$

$\rightsquigarrow$ superlinear convergence if

1. regularity condition

$$\|M_k^{-1}\|_{\mathcal{L}(Y,X)} \leqslant C \qquad \text{for all } k$$

2. approximation condition

$$\lim_{\|d^k\|_X \to 0} \frac{\|F(x^* + d^k) - F(x^*) - M_k d^k\|_Y}{\|d^k\|_X} = 0$$

Semismooth Newton method

Goal: define Newton derivative $M_k = D_N F$ such that

$$x^{k+1} = x^k - D_N F(x^k)^{-1} F(x^k)$$

converges superlinearly for $F(x) = 0$ nonsmooth

- $\mathbb{R}^n$: F Lipschitz $\rightsquigarrow$ $D_N F$ from Clarke subdifferential
  [Mifflin 1977, Kummer 1992, Qi/Sun 1993]
- function space: no explicit Clarke subdifferential
  $\rightsquigarrow$ define $D_N F$ via approximation condition
  [Chen/Nashed/Qi 2000, Hintermüller/Ito/Kunisch 2002]
- $f : \mathbb{R}^N \to \mathbb{R}$ semismooth $\rightsquigarrow$ superposition operator
  $F : L^p(\Omega) \to L^q(\Omega)$ semismooth for $p > q$
  [Ulbrich 2002/03/11, Schiela 2008]

Semismooth functions: example

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

- $f : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \max(0, t)$

$$D_N f(t) \in \partial_C f(t) = \begin{cases} \{0\} & t < 0 \\ \{1\} & t > 0 \\ [0,1] & t = 0 \end{cases}$$

- $F : L^p(\Omega) \to L^q(\Omega), \quad u(x) \mapsto \max(0, u(x)), \quad p > q$

$$[D_N F(u)h](x) = \begin{cases} 0 & u(x) < 0 \\ h(x) & u(x) \geqslant 0 \end{cases}$$

$\leadsto$ Moreau–Yosida regularization semismooth

Numerical solution: summary

For nonconvex $\mathcal{G} : L^2(\Omega) \to \mathbb{R}$, $\quad \mathcal{G}(u) = \int_\Omega g(u(x))\,dx$,

Approach:

1. compute Fenchel conjugate $g^*(q)$
2. compute subdifferential $\partial g^*(q)$
3. compute proximal mapping $\operatorname{prox}_{\gamma\partial g^*}(q)$
4. compute Moreau–Yosida regularization $\partial g^*_\gamma(q)$
5. $\rightsquigarrow$ semismooth Newton method, continuation in γ for superposition operator $\partial \mathcal{G}^*_\gamma(p)(x) = \partial g^*_\gamma(p(x))$

1 Overview

2 Approach
- Convex relaxation
- Moreau–Yosida regularization
- Semismooth Newton method

3 Multi-bang penalty
- Optimality system
- Numerical solution
- Examples

4 Switching penalty
- Optimality system
- Numerical solution
- Examples

Formulation

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

$$\begin{cases} \displaystyle\min_{u\in L^2(\Omega)} \frac{1}{2}\|y-z\|_{L^2}^2 + \frac{\alpha}{2}\|u\|_{L^2}^2 + \beta\int_\Omega \prod_{i=1}^d |u(x)-u_i|_0\,dx \\ \text{s.t. } Ay = u, \qquad u_1 \leqslant u(x) \leqslant u_d \end{cases}$$

- $u_1 < \cdots < u_d$ given control values $(d > 2)$
- $z \in L^2(\Omega)$ desired state
- $A : V \to V^*$ 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

Fenchel conjugate

$$g : \mathbb{R} \to \overline{\mathbb{R}}, \qquad v \mapsto \frac{\alpha}{2} v^2 + \beta \prod_{i=1}^{d} |v - u_i|_0 + \delta_{[u_1,u_d]}(v)$$

$$g^* : \mathbb{R} \to \overline{\mathbb{R}}, \qquad q \mapsto \sup_v \; q\,v - g(v)$$

Case differentiation: sup attained at $\bar{v}$,

$$g^*(q) = \begin{cases} qu_i - \frac{\alpha}{2} u_i^2 & \bar{v} = u_i, \quad 1 \leqslant i \leqslant d \\ \frac{1}{2\alpha} q^2 - \beta & \bar{v} \neq u_i, \quad 1 \leqslant i \leqslant d \end{cases}$$

Fenchel conjugate

$$g^*(q) = \begin{cases} qu_i - \frac{\alpha}{2}u_i^2 & q \in \overline{Q}_i, \quad 1 \leqslant i \leqslant d \\ \frac{1}{2\alpha}q^2 - \beta & q \in \overline{Q}_0 \end{cases}$$

$$Q_1 := \left\{q : q - \alpha u_1 < \sqrt{2\alpha\beta} \wedge q < \frac{\alpha}{2}(u_1 + u_2)\right\}$$

$$Q_i := \left\{q : |q - \alpha u_i| < \sqrt{2\alpha\beta} \wedge \frac{\alpha}{2}(u_{i-1} + u_i) < q < \frac{\alpha}{2}(u_i + u_{i+1})\right\}$$

$$Q_d := \left\{q : q - \alpha u_d > \sqrt{2\alpha\beta} \wedge \frac{\alpha}{2}(u_d + u_{d-1}) < q\right\}$$

$$Q_0 := \left\{q : |q - \alpha u_j| > \sqrt{2\alpha\beta} \text{ for all } j \wedge \alpha u_1 < q < \alpha u_d\right\}$$

Fenchel conjugate

$$g^*(q) = \begin{cases} qu_i - \frac{\alpha}{2}u_i^2 & q \in \overline{Q}_i, \quad 1 \leqslant i \leqslant d \\ \frac{1}{2\alpha}q^2 - \beta & q \in \overline{Q}_0 \end{cases}$$

continuous, piecewise differentiable:

$$\partial g^*(q) = \begin{cases} \{u_i\} & q \in Q_i,\ 1 \leqslant i < d \\ \left\{\frac{1}{\alpha}q\right\} & q \in Q_0 \\ [u_i, u_{i+1}] & q \in \overline{Q}_i \cap \overline{Q}_{i+1},\ 1 \leqslant i < d \\ \left[\min\left\{u_i, \frac{1}{\alpha}q\right\}, \max\left\{u_i, \frac{1}{\alpha}q\right\}\right] & q \in \overline{Q}_i \cap \overline{Q}_0,\ 1 \leqslant i \leqslant d \end{cases}$$

(no explicit dependence on β!)

Optimality system

$$\bar{p} = S^*(z - S\bar{u})$$
$$\bar{u} \in \partial\mathcal{G}^*(\bar{p})$$
$$= \begin{cases} \{u_i\} & \bar{p}(x) \in Q_i \\ \left\{\frac{1}{\alpha}\bar{p}(x)\right\} & \bar{p}(x) \in Q_0 \\ [u_i, u_{i+1}] & \bar{p}(x) \in \overline{Q}_i \cap \overline{Q}_{i+1} \\ \left[\min\left\{u_i, \frac{1}{\alpha}\bar{p}(x)\right\}, \max\left\{u_i, \frac{1}{\alpha}\bar{p}(x)\right\}\right] & \bar{p}(x) \in \overline{Q}_i \cap \overline{Q}_0 \end{cases}$$

- $S : u \mapsto y$ control-to-observation mapping, S^* adjoint
- necessary conditions for $\min_u \mathcal{F}(u) + \mathcal{G}^{**}(u)$ (convex, l.s.c.)
- $\rightsquigarrow$ unique solution $(\bar{u}, \bar{p}) \in L^2(\Omega) \times L^2(\Omega)$

Structure of solution

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

$$\Omega = \mathcal{A} \cup \mathcal{F} \cup \mathcal{S}$$

- multi-bang arc $\quad \mathcal{A} = \bigcup_{i=1}^{d} \{x : \bar{u}(x) = u_i\}$
- free arc $\quad \mathcal{F} = \left\{x : \bar{u}(x) = \frac{1}{\alpha}\bar{p}(x) \neq u_i\right\}$
- singular arc $\quad \mathcal{S} = \left\{x : \bar{u}(x) \notin \left\{u_i, \frac{1}{\alpha}\bar{p}(x)\right\}\right\}$

Generalized multi-bang principle

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

- if β sufficiently large: $Q_0 = \emptyset$, free arc

$$\mathcal{F} \subset \{\bar p(x) \in Q_0\} = \emptyset$$

- singular arc corresponds to set-valued subdifferential:

$$\begin{aligned}\mathcal{S} &= \left\{ \bar p(x) \in \bigcup_{i=1}^{d-1} (\overline{Q}_i \cap \overline{Q}_{i+1}) \cup \bigcup_{i=1}^{d} (\overline{Q}_i \cap \overline{Q}_0) \right\} \\ &\subset \left\{ \bar p(x) \in \left\{ \tfrac{\alpha}{2}(u_i + u_{i+1}), \alpha u_i - \sqrt{2\alpha\beta}, \alpha u_i + \sqrt{2\alpha\beta} \right\} \right\}\end{aligned}$$

- for suitable A, $\bar p(x)$ constant implies $[A^*\bar p](x) = [z - \bar y](x) = 0$,

$\rightsquigarrow$ $|\{x : \bar y(x) = z(x)\}| = 0 \;\Rightarrow\; \bar u \in \{u_1, \dots u_d\}$ a. e., true multi-bang

(Sub)optimality

- duality gap for non-convex $\mathcal{G}$:

$$\mathcal{G}(\bar{u}) + \mathcal{G}^*(\bar{p}) - \langle \bar{p}, \bar{u} \rangle \leqslant \beta|\mathcal{S}|$$

(pointwise gap of β where $\partial g^*(\bar{p}(x))$ set-valued)

- $\rightsquigarrow$ in general: $\bar{u}$ sub-optimal:

$$\mathcal{J}(\bar{u}) \leqslant \mathcal{J}(u) + \beta|\mathcal{S}| \qquad \text{for all } u$$

- but: $\bar{u}$ true multi-bang $\rightsquigarrow$ $|\mathcal{S}| = 0$ $\rightsquigarrow$ $\bar{u}$ optimal

Moreau–Yosida regularization

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

$$\partial g_\gamma^*(q) = \begin{cases} u_i & q \in Q_i^\gamma \\ \frac{1}{\alpha+\gamma} q & q \in Q_0^\gamma \\ \frac{1}{\gamma}\left(q - (\alpha u_i + \sqrt{2\alpha\beta})\right) & q \in Q_{i0}^\gamma \\ \frac{1}{\gamma}\left(q - \frac{\alpha}{2}(u_i + u_{i+1})\right) & q \in Q_{i,i+1}^\gamma \end{cases}$$

$$Q_i^\gamma = \Big\{ q : |q - (\alpha+\gamma)u_i| < \sqrt{2\alpha\beta} \wedge \frac{\alpha}{2}\left(u_{i-1} + \left(1 + \frac{2\gamma}{\alpha}\right)u_i\right) < q < \frac{\alpha}{2}\left(\left(1 + \frac{2\gamma}{\alpha}\right)u_i + u_{i+1}\right) \Big\}$$

$$Q_0^\gamma = \left\{ q : |q - (\alpha+\gamma)u_j| > \sqrt{2\alpha\beta} \wedge (\alpha+\gamma)u_1 < q < (\alpha+\gamma)u_d \right\}$$

$$Q_{i,i+1}^\gamma = \left\{ q : \frac{\alpha}{2}\left(\left(1 + \frac{2\gamma}{\alpha}\right)u_i + u_{i+1}\right) \leqslant q \leqslant \frac{\alpha}{2}\left(u_i + \left(1 + \frac{2\gamma}{\alpha}\right)u_{i+1}\right) \right\}$$

$$Q_{i0}^\gamma = \left\{ q : \sqrt{2\alpha\beta} \leqslant q - (\alpha+\gamma)u_i \leqslant \left(1 + \frac{\gamma}{\alpha}\right)\sqrt{2\alpha\beta} \right\}$$

Regularized optimality system

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

- $\partial \mathcal{G}_\gamma^*$ maximal monotone $\rightsquigarrow$ unique solution (u_γ, p_γ)
- $(u_\gamma, p_\gamma) \rightharpoonup (\bar{u}, \bar{p})$ as $\gamma \to 0$
- ∂g_γ^* Lipschitz continuous, piecewise C^1, norm gap $V \hookrightarrow L^2(\Omega)$
- $\rightsquigarrow$ semismooth Newton method

Regularized optimality system

$$\begin{cases} A^* p_\gamma = z - y_\gamma \\ A y_\gamma = \mathcal{G}_\gamma^*(p_\gamma) \end{cases}$$

- $\partial \mathcal{G}_\gamma^*$ maximal monotone $\rightsquigarrow$ unique solution (u_γ, p_γ)
- $(u_\gamma, p_\gamma) \rightharpoonup (\bar u, \bar p)$ as $\gamma \to 0$
- ∂g_γ^* Lipschitz continuous, piecewise C^1, norm gap $V \hookrightarrow L^2(\Omega)$
- $\rightsquigarrow$ semismooth Newton method
- introduce state $y_\gamma = S u_\gamma$, eliminate control $u_\gamma = \mathcal{G}_\gamma^*(p_\gamma)$

Semismooth Newton method

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

$$[D_N\mathcal{G}_\gamma^*(p)\delta p](x) = \begin{cases} \frac{1}{\alpha+\gamma}\delta p(x) & p(x) \in Q_0^\gamma \\ \frac{1}{\gamma}\delta p(x) & p(x) \in Q_{i,i+1}^\gamma \\ 0 & \text{else} \end{cases}$$

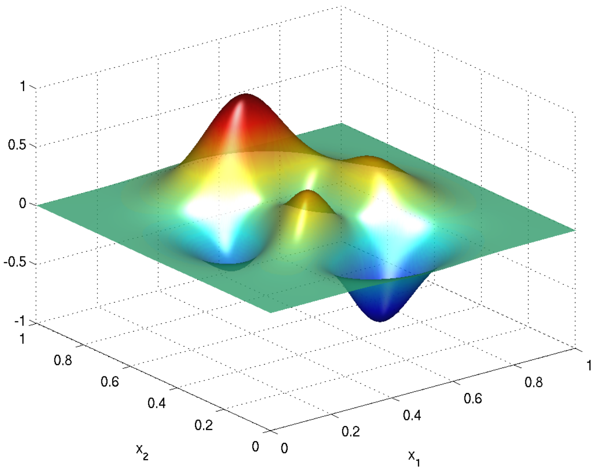
- symmetric, but: local convergence
- $\rightsquigarrow$ continuation in $\gamma \to 0$
- $\rightsquigarrow$ backtracking line search based on residual norm
- only number of sets Q_i depends on d $\rightsquigarrow$ linear complexity

Numerical examples

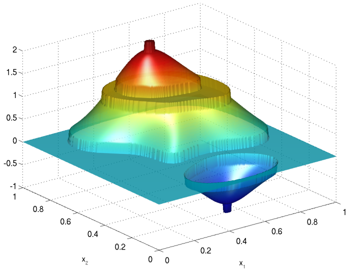
- $\Omega = [0,1]^2, \quad A = -\Delta$
- finite element discretization: uniform grid, 256×256 nodes
- state, adjoint: piecewise linear
- control: eliminated (variational discretization)
- $d = 5, \quad (u_1, \dots, u_5) = (-2, 1, 0, 1, 2)$
- $\gamma = 0$: regularized active sets empty, true multi-bang

  $\gamma > 0$: terminated with 2–21 nodes in regularized active sets

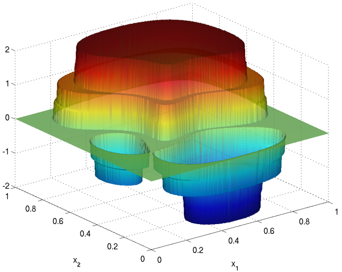
Numerical examples: desired state



Controls: $\beta = 10^{-4}$ (free arcs)

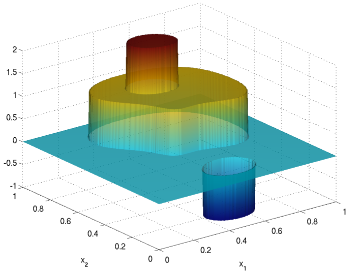


(a) $\alpha = 5 \cdot 10^{-3}$ ($\gamma = 0$)

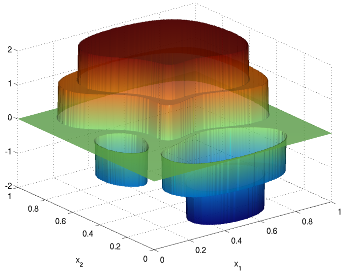


(b) $\alpha = 10^{-3}$ ($\gamma \approx 10^{-8}$)

Controls: $\beta = 10^{-3}$ (no free arcs)

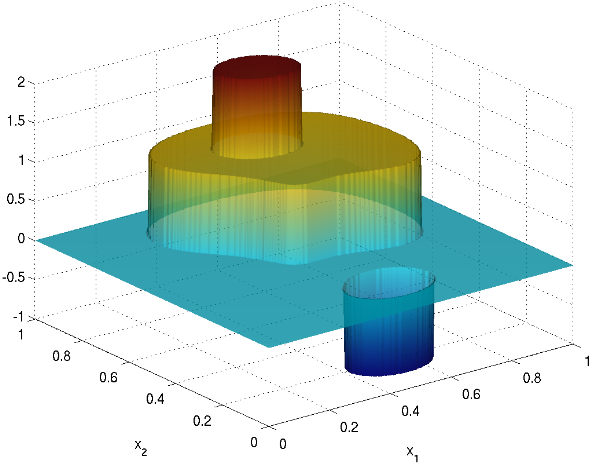


(c) $\alpha = 5 \cdot 10^{-3}$ $(\gamma = 0)$



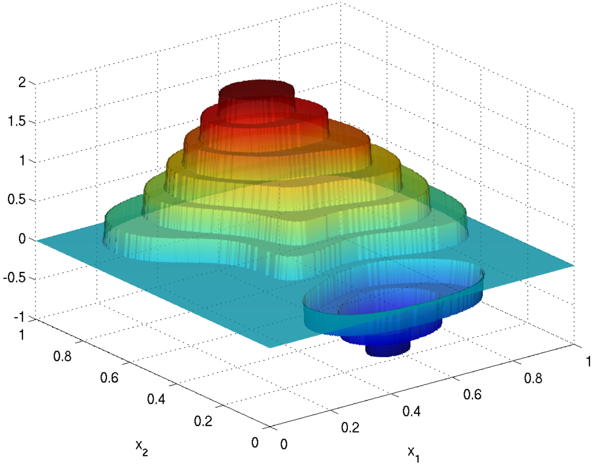
(d) $\alpha = 10^{-3}$ $(\gamma \approx 10^{-7})$

Controls: effect of d



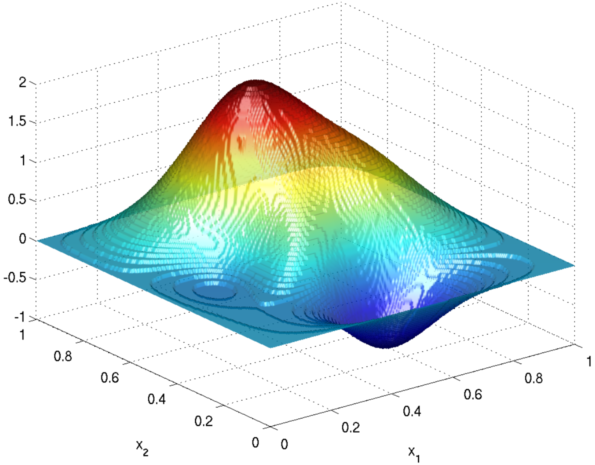
(a) $d = 5$ ($\gamma = 0$)

Controls: effect of d



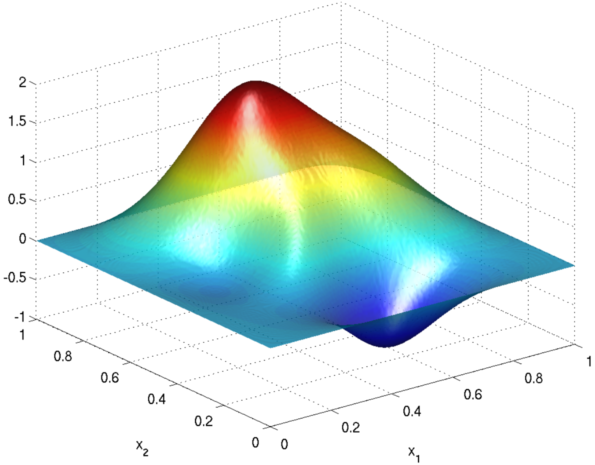
(b) $d = 15$ ($\gamma = 0$)

Controls: effect of d



(c) $d = 101$ ($\gamma \approx 10^{-9}$)

Controls: effect of d



(d) $d = 1001$ ($\gamma \approx 10^{-11}$)

Example: control in coefficient

- $S : u \mapsto y$ solving

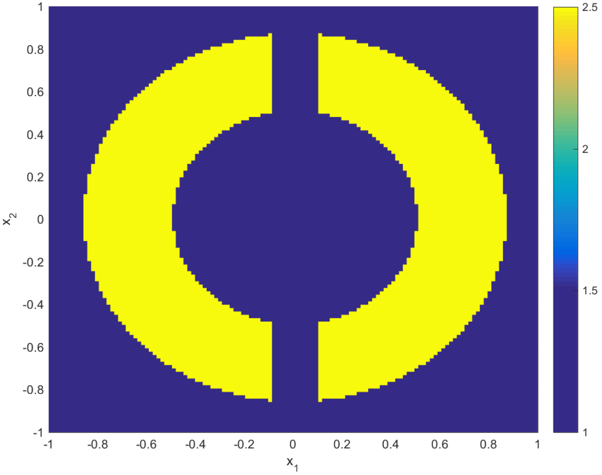
$$-\Delta y + uy = f$$

- approach applicable, but $\mathcal{F}$ nonconvex
- $\rightsquigarrow$ only local minimizers, regularity condition technical
- numerical example: $\Omega = [-1,1]^2$,

$$f(x_1, x_2) = \sin(\pi x_1)\sin(\pi x_2)$$

- fix u_{ref}, set $z = S(u_{\text{ref}})$, $(u_1, u_2, u_3, u_4) = (1, 1.5, 2, 2.5)$
- $\alpha = 10^{-6}$, $\beta \to \infty$ (formal)

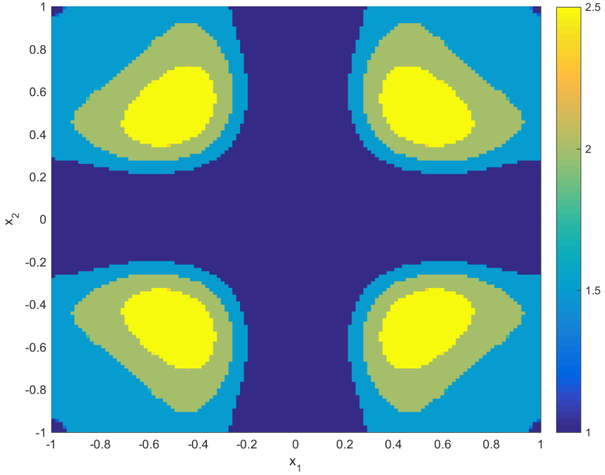
Example: reference



(a) u_{ref}

Example: optimal control ($\gamma = 10^{-12}$)

UNIVERSITÄT DUISBURG ESSEN
Open-Minded



(b) u_γ

1 Overview

2 Approach
- Convex relaxation
- Moreau–Yosida regularization
- Semismooth Newton method

3 Multi-bang penalty
- Optimality system
- Numerical solution
- Examples

4 Switching penalty
- Optimality system
- Numerical solution
- Examples

Formulation

$$\min_{u\in L^2(D;\mathbb{R}^2)} \frac{1}{2}\|Su - z\|_Y^2 + \int_D \frac{\alpha}{2}(u_1(t)^2 + u_2(t)^2) + \beta|u_1(t)u_2(t)|_0 \,dt,$$

- $S : L^2(D; \mathbb{R}^2) \to Y, \quad Y = Y^*$ Hilbert space, $\quad z \in Y$ target
- $\mathcal{F}(u) = \frac{1}{2}\|Su - z\|_Y^2$ strictly convex, smooth, coercive
- assumption: $\quad S^*(Y) \hookrightarrow L^r(D; \mathbb{R}^2)$ with $r > 2$
- e.g., $D = (0, T), \quad Y = L^2([0, T] \times \Omega), \quad S(u) = y$ solution to

$$\partial_t y - Ay = \chi_{\omega_1}(x)u_1(t) + \chi_{\omega_2}(x)u_2(t)$$

Fenchel conjugate

$$g : \mathbb{R}^2 \to \overline{\mathbb{R}}, \qquad v \mapsto \frac{\alpha}{2}\left(v_1^2 + v_2^2\right) + \beta |v_1 v_2|_0$$

$$g^* : \mathbb{R}^2 \to \overline{\mathbb{R}}, \qquad q \mapsto \sup_v \; q \cdot v - g(v)$$

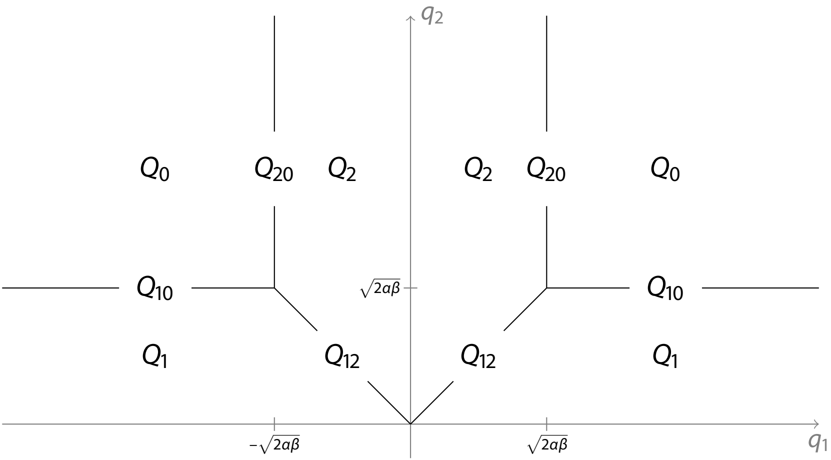
Case differentiation:

$$g^*(q) = \begin{cases} \frac{1}{2\alpha} q_1^2 & \text{if } |q_1| \geqslant |q_2| \text{ and } |q_2| \leqslant \sqrt{2\alpha\beta} \\ \frac{1}{2\alpha} q_2^2 & \text{if } |q_2| \geqslant |q_1| \text{ and } |q_1| \leqslant \sqrt{2\alpha\beta} \\ \frac{1}{2\alpha}\left(q_1^2 + q_2^2\right) - \beta & \text{if } |q_1|, |q_2| \geqslant \sqrt{2\alpha\beta} \end{cases}$$

$$g^*(q) = \begin{cases} \frac{1}{2\alpha} q_1^2 & \text{if } |q_1| \geqslant |q_2| \text{ and } |q_2| \leqslant \sqrt{2\alpha\beta} \\ \frac{1}{2\alpha} q_2^2 & \text{if } |q_2| \geqslant |q_1| \text{ and } |q_1| \leqslant \sqrt{2\alpha\beta} \\ \frac{1}{2\alpha}\left(q_1^2 + q_2^2\right) - \beta & \text{if } |q_1|, |q_2| \geqslant \sqrt{2\alpha\beta} \end{cases}$$

$$\partial g^*(q) = \begin{cases} \left(\left\{\frac{1}{\alpha}q_1\right\}, \{0\}\right) & \text{if } q \in Q_1 \\ \left(\{0\}, \left\{\frac{1}{\alpha}q_2\right\}\right) & \text{if } q \in Q_2 \\ \left(\left\{\frac{1}{\alpha}q_1\right\}, \left\{\frac{1}{\alpha}q_2\right\}\right) & \text{if } q \in Q_0 \\ \left(\left\{\frac{1}{\alpha}q_1\right\}, \left[0, \frac{1}{\alpha}q_2\right]\right) & \text{if } q \in Q_{10} \\ \left(\left[0, \frac{1}{\alpha}q_1\right], \left\{\frac{1}{\alpha}q_2\right\}\right) & \text{if } q \in Q_{20} \\ \left\{\left(\frac{t}{\alpha}q_1, \frac{1-t}{\alpha}q_2\right) : t \in [0,1]\right\} & \text{if } q \in Q_{12} \end{cases}$$

Subdifferential: domains



Optimality system

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

$$\begin{cases} -\bar{p} = S^*(S\bar{u} - z) \\ \bar{u}(x) \in \partial g^*(\bar{p}(x)) \quad \text{a.e. in } D \end{cases}$$

Structure of solution: $D = \mathcal{A} \cup \mathcal{I} \cup \mathcal{S}$,

- switching arc $\quad \mathcal{A} = \left\{x \in D : \bar{p}(x) \in Q_1 \cup Q_2 \cup \{(0,0)\}\right\}$
- free arc $\quad \mathcal{I} = \left\{x \in D : \bar{p}(x) \in Q_0 \cup Q_{10} \cup Q_{20}\right\}$, $\quad \partial\mathcal{I} = \left\{x \in D : \bar{p}(x) \in Q_{10} \cup Q_{20}\right\}$
- singular arc $\quad \mathcal{S} = \left\{x \in D : \bar{p}(x) \in Q_{12} \setminus \{(0,0)\}\right\}$

Optimality system

Structure of solution: $D = \mathcal{A} \cup \mathcal{I} \cup \mathcal{S}$,

- switching arc $\mathcal{A} = \left\{x \in D : \bar{p}(x) \in Q_1 \cup Q_2 \cup \{(0,0)\}\right\}$
- free arc $\mathcal{I} = \left\{x \in D : \bar{p}(x) \in Q_0 \cup Q_{10} \cup Q_{20}\right\}$,
  $\partial\mathcal{I} = \left\{x \in D : \bar{p}(x) \in Q_{10} \cup Q_{20}\right\}$
- singular arc $\mathcal{S} = \left\{x \in D : \bar{p}(x) \in Q_{12} \setminus \{(0,0)\}\right\}$

Suboptimality

$$\mathcal{J}(\bar{u}) \leqslant \mathcal{J}(u) + \beta\left(|\partial\mathcal{I}| + 2|\mathcal{S}|\right) \qquad \text{for all } u$$

Structure of solution

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

Suboptimality

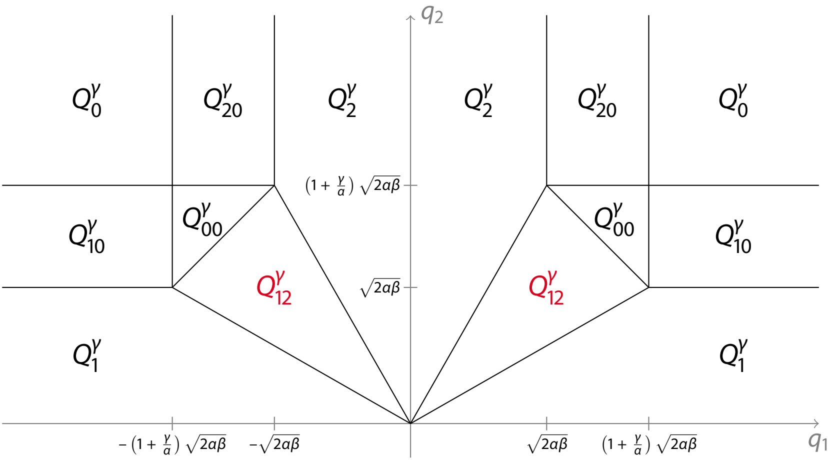
$$\mathcal{J}(\bar{u}) \leqslant \mathcal{J}(u) + \beta\left(|\partial\mathcal{I}| + 2|\mathcal{S}|\right) \qquad \text{for all } u$$

- free arc $\quad \mathcal{I}_\beta = \left\{x \in D : |\bar{p}_1(x)|, |\bar{p}_2(x)| \geqslant \sqrt{2\alpha\beta}\right\}$
- $|\partial\mathcal{I}_\beta| < |\mathcal{I}_\beta| \to 0$ as $\beta \to \infty$
- if $\bar{p}$ bounded, $|\mathcal{I}_\beta| = 0$ for β sufficiently large
- singular arc $\quad \mathcal{S}_\beta = \left\{x \in D : |\bar{p}_1(x)| = |\bar{p}_2(x)| > 0\right\}$
- $\rightsquigarrow$ switching control $\bar{u}_1(x)\bar{u}_2(x) = 0$ a. e. optimal

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

$$(\partial g^*)_\gamma(q) = \begin{cases} \left(\frac{1}{\alpha+\gamma} q_1, 0\right) & \text{if } q \in Q_1^\gamma \\ \left(0, \frac{1}{\alpha+\gamma} q_2\right) & \text{if } q \in Q_2^\gamma \\ \left(\frac{1}{\alpha+\gamma} q_1, \frac{1}{\alpha+\gamma} q_2\right) & \text{if } q \in Q_0^\gamma \\ \left(\frac{1}{\alpha+\gamma} q_1, \frac{1}{\gamma}\left(q_2 - \operatorname{sign}(q_2)\sqrt{2\alpha\beta}\right)\right) & \text{if } q \in Q_{10}^\gamma \\ \left(\frac{1}{\gamma}\left(q_1 - \operatorname{sign}(q_1)\sqrt{2\alpha\beta}\right), \frac{1}{\alpha+\gamma} q_2, \right) & \text{if } q \in Q_{20}^\gamma \\ \left(\frac{1}{\gamma}\left(q_1 - \operatorname{sign}(q_1)\sqrt{2\alpha\beta}\right), \right. & \\ \quad \left. \frac{1}{\gamma}\left(q_2 - \operatorname{sign}(q_2)\sqrt{2\alpha\beta}\right)\right) & \text{if } q \in Q_{00}^\gamma \\ \left(\frac{1}{\gamma}\left(\frac{\alpha+\gamma}{2\alpha+\gamma} q_1 - \operatorname{sign}(q_1)\frac{\alpha}{2\alpha+\gamma}|q_2|\right), \right. & \\ \quad \left. \frac{1}{\gamma}\left(\frac{\alpha+\gamma}{2\alpha+\gamma} q_2 - \operatorname{sign}(q_2)\frac{\alpha}{2\alpha+\gamma}|q_1|\right)\right) & \text{if } q \in Q_{12}^\gamma \end{cases}$$

Moreau–Yosida regularization: sketch



Moreau–Yosida regularization

$$\begin{cases} p_\gamma = S^*(z - Su_\gamma) \\ u_\gamma \in (\partial \mathcal{G}^*)_\gamma(p_\gamma) \end{cases}$$

- $(\partial \mathcal{G}^*)_\gamma$ maximal monotone $\rightsquigarrow$ unique solution (u_γ, p_γ)
- weak convergence $(u_\gamma, p_\gamma) \to (\bar{u}, \bar{p})$ as $\gamma \to 0$
- $(\partial \mathcal{G}^*)_\gamma$ Lipschitz continuous, piecewise C^1, norm gap
- $\rightsquigarrow$ semismooth Newton method, continuation in $\gamma \to 0$
- vector penalty (Q_{12}^γ): needs line search (based on residual norm)

Numerical example

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

- Domain $\Omega = [0,1]^2, \quad D = [0,1]$,

$$\omega_1 = \left\{(x_1, x_2) \in \Omega : x_2 < \tfrac{1}{4}\right\}$$
$$\omega_2 = \left\{(x_1, x_2) \in \Omega : x_2 > \tfrac{3}{4}\right\}$$

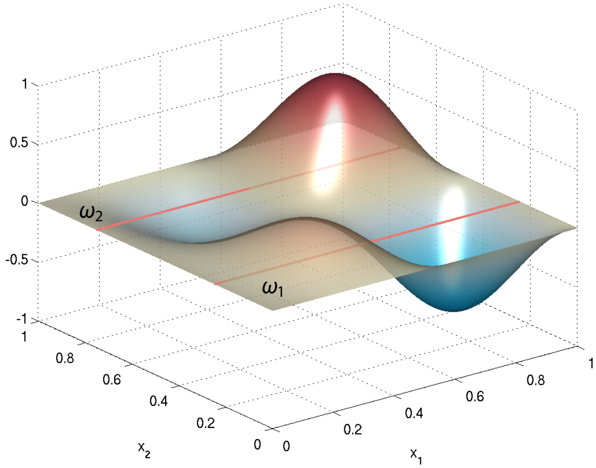
- elliptic example: $S(u) = y$ solves

$$-\Delta y = \chi_{\omega_1}(x_1, x_2)u_1(x_1) + \chi_{\omega_2}(x_1, x_2)u_2(x_1).$$

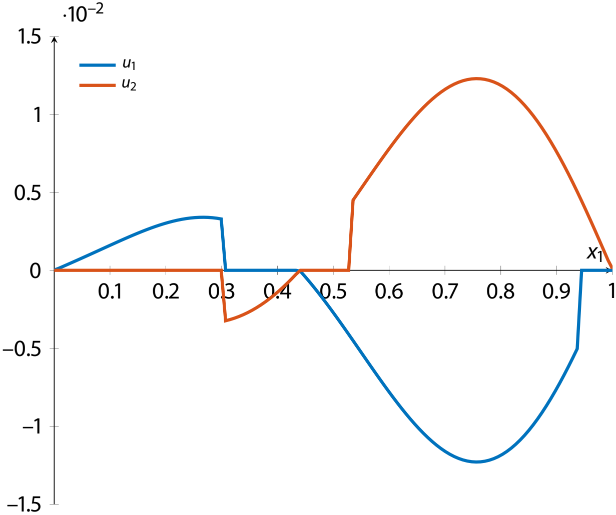
- target

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

Numerical example: target

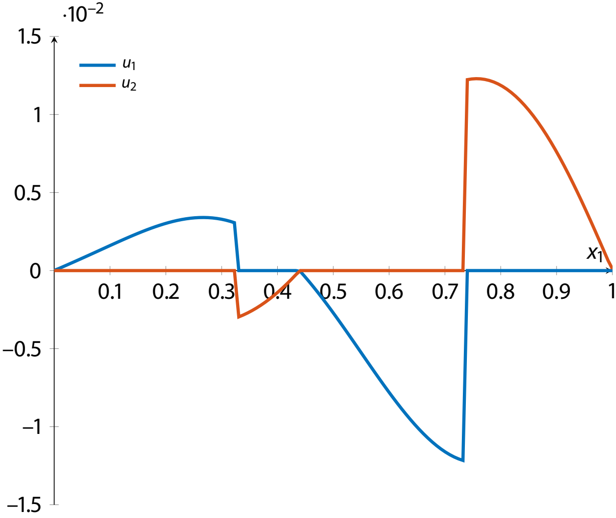


Numerical example: controls



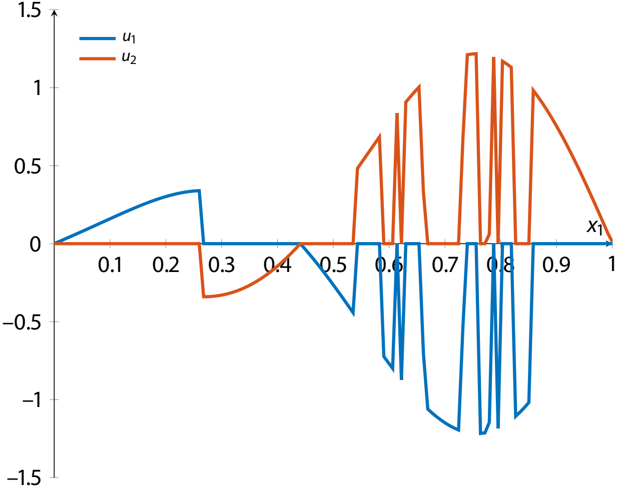
(a) $\alpha = 10^{-3}, \beta = 10^{-8}$

Numerical example: controls



(b) $\alpha = 10^{-3}, \beta = 10^{-3}$

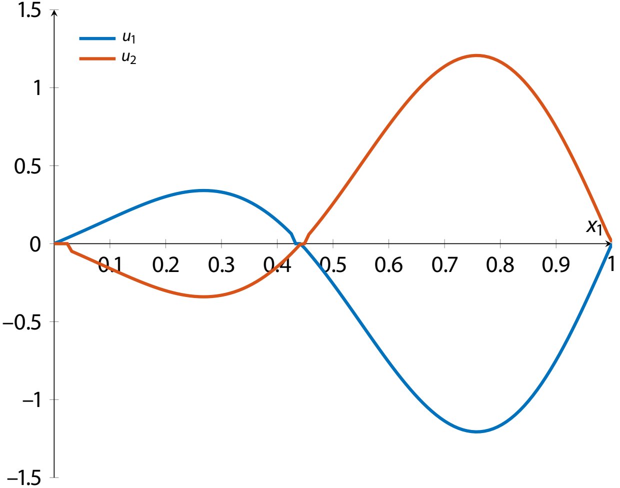
Numerical example: controls



(c) $\alpha = 10^{-5}, \beta = 10^{-3}$

Numerical example: controls

UNIVERSITÄT DUISBURG ESSEN
Open-Minded



(d) $\alpha = 10^{-5}, \beta = 10^{-8}$

Example: more controls

- $N > 2$ controls: switching penalty (convex):

$$\mathcal{G}(u) := \int_0^T \frac{\alpha}{2}|u(t)|_2^2 + \beta \sum_{i<j} |u_i(t)u_j(t)|\,dt$$

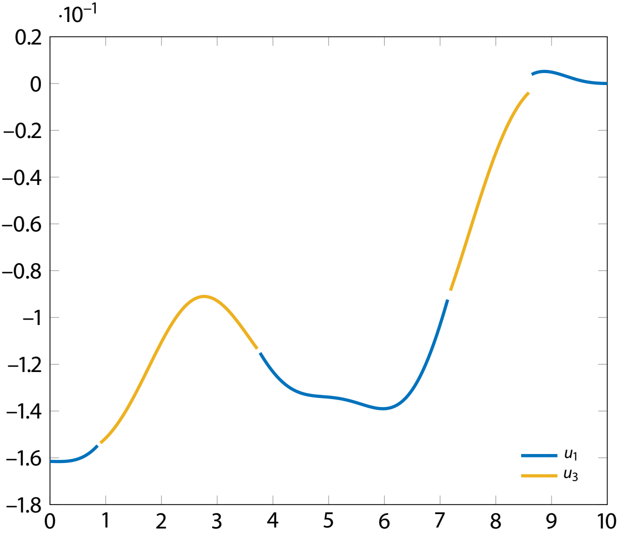
- for $\alpha = \beta$:

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

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

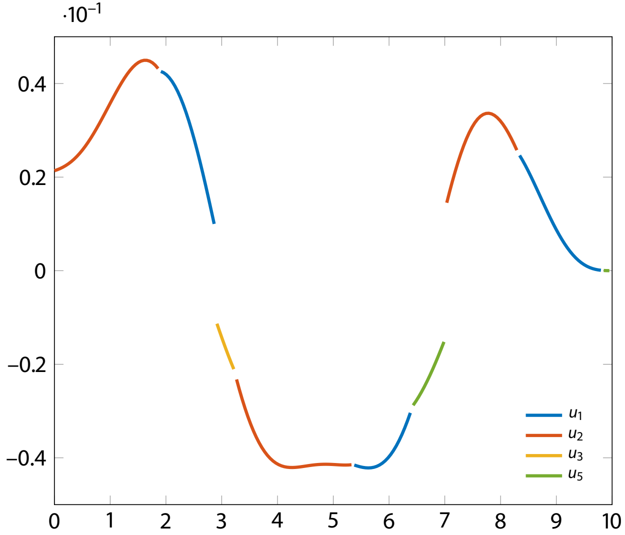
- $\rightsquigarrow$ efficient evaluation of proximal mapping by sorting
- $\rightsquigarrow$ Moreau–Yosida regularization, semismooth Newton method

Numerical example: heat equation



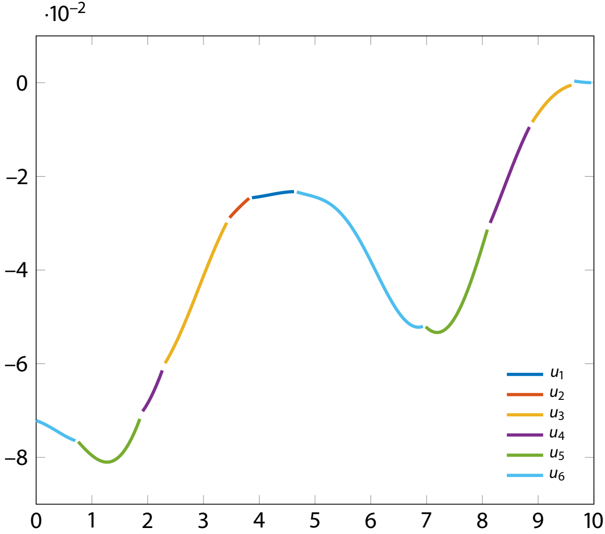
(a) $N = 3$ ($\alpha = 10^{-1}$)

Numerical example: heat equation



(b) $N = 5$ ($\alpha = 10^{-1}$)

Numerical example: heat equation



(c) $N = 7$ $(a = 10^{-1})$

Conclusion

Convex relaxation of discrete problem:

- well-posed primal-dual optimality system
- solution optimal under general assumptions
- linear complexity in number of parameter values
- $\rightsquigarrow$ efficient numerical solution (superlinear convergence)

Outlook:

- multimaterial optimization in leading coefficients
- BV regularization (done in linear case)
- other hybrid discrete–continuous problems

Preprint, MATLAB/Python codes:

http://www.uni-due.de/mathematik/agclason/clason_pub.php