A convex analysis approach to hybrid binary–continuous optimal control problems with application to switching control

Christian Clason[1] Kazufumi Ito[2] Karl Kunisch[3]

[1]Faculty of Mathematics, University of Duisburg-Essen
[2]Department of Mathematics, North Carolina State University, Raleigh
[3]Institute of Mathematics and Scientific Computing, University of Graz

10th AIMS Conference on Dynamical Systems, Differential Equations and Applications
Madrid, July 7, 2014

L^0 penalty

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

- “continuous counting measure”
- popular in sparse optimization
- binary penalty $\rightsquigarrow$ combinatorial optimization
- difficulty: non-smooth, non-convex, not lower-semicontinuous
- not a norm $\rightsquigarrow$ no regularization

Binary penalties

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

- $\mathcal{F}(u)$ tracking or discrepancy term (here: linear–quadratic)

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)
- $\beta > 0$ large penalizes *free arc* $u(x) \neq u_i$
- $\alpha > 0$ penalizes magnitude of $u(x) = u_i$

Binary penalties

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

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

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
- $\beta > 0$ large penalizes free arc $u_1 u_2 \neq 0$
- $\alpha > 0$ penalizes magnitude of active u_i

1 Overview

2 Approach

3 Switching control
- Optimality system
- Numerical solution
- Examples

Convex analysis approach

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

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

Necessary optimality conditions

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

- Fermat, sum rule for subdifferentials (under regularity condition)

Convex analysis approach

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

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

Necessary optimality conditions

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

- $\mathcal{G}^*(p) = \sup_u \langle u, p\rangle - \mathcal{G}(u)$ Fenchel conjugate
- subdifferential inversion, “inverse convex function theorem”

Convex analysis approach

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

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

Sufficient(?) optimality conditions

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

- $\mathcal{G}^*$ Fenchel conjugate: always convex, lower semi-continuous
- $\rightsquigarrow$ 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

Numerical solution

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

$\rightsquigarrow$ regularize: 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
- Hilbert space: concides with resolvent $(\mathrm{Id} + \gamma\partial\mathcal{G}^*)^{-1}(p)$

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

Proximal mapping

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

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, not semismooth

Numerical solution

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

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

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

Numerical solution

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

Approach:

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

superposition operator $\quad \partial \mathcal{G}^*_\gamma(p)(x) = \partial g^*_\gamma(p(x))$

1 Overview

2 Approach

3 Switching control
- 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

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

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

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

$$g^*(q) = \begin{cases} g_1^*(q) := \frac{1}{2\alpha} q_1^2 & \text{if } \bar{v}_2 = 0 \\ g_2^*(q) := \frac{1}{2\alpha} q_2^2 & \text{if } \bar{v}_1 = 0 \\ g_0^*(q) := \frac{1}{2\alpha}(q_1^2 + q_2^2) - \beta & \text{if } \bar{v}_1, \bar{v}_2 \neq 0 \end{cases}$$

Fenchel conjugate

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

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

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

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

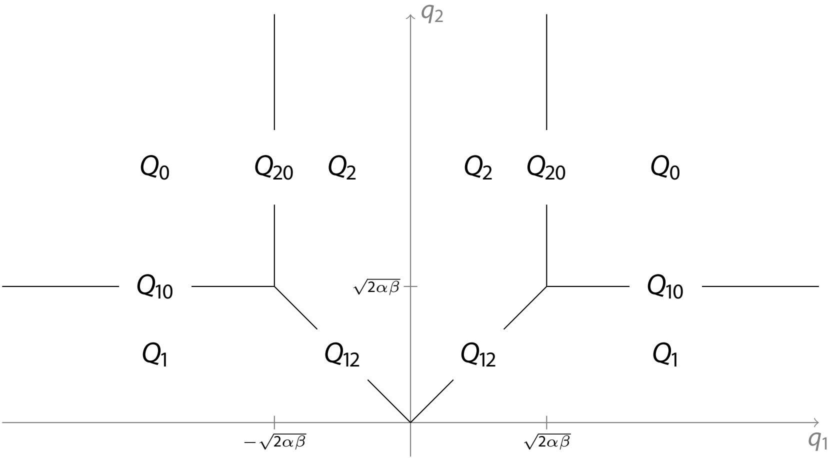
Subdifferential

$$\partial g^*(q) = \overline{\text{co}}\left(\bigcup_{\{i : g^*(q) = g_i^*(q)\}} \{(g_i^*)'(q)\} \right)$$

Six possible cases for $g^*(q) = g_i^*(q), i \in \{1, 2, 0\}$:

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

Subdifferential: sketch



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

Optimality system

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

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

Suboptimality

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

Structure of solution

- $|\partial\mathcal{I}_\beta| < |\mathcal{I}_\beta| \to 0$ as $\beta \to \infty$
- If $\bar{p}$ bounded, $|\mathcal{I}_\beta| = 0$ for β sufficiently large
- For every $\varepsilon > 0$, $|\mathcal{S}_\beta^\varepsilon| \to 0$ as $\beta \to \infty$ for

$$\mathcal{S}_\beta^\varepsilon := \mathcal{S}_\beta \cap \{x \in D : |u_{\beta,1}(x)|, |u_{\beta,2}(x)| \geq \varepsilon\}$$

- No information on remaining singular arc, but likely small since

$$0 < |\bar{p}_1(x)| = |\bar{p}_2(x)| < \alpha\varepsilon$$

Numerical solution of optimality system

Replace set-valued ∂g^* by Moreau–Yosida regularization

$$\begin{cases} p_\gamma = S^*(z - Su_\gamma) \\ u_\gamma(x) \in (\partial g^*)_\gamma(p_\gamma(x)) = \frac{1}{\gamma}\left(p_\gamma(x) - \operatorname{prox}_{\gamma g^*}(p_\gamma(x))\right) \end{cases}$$

Proximal point mapping / resolvent

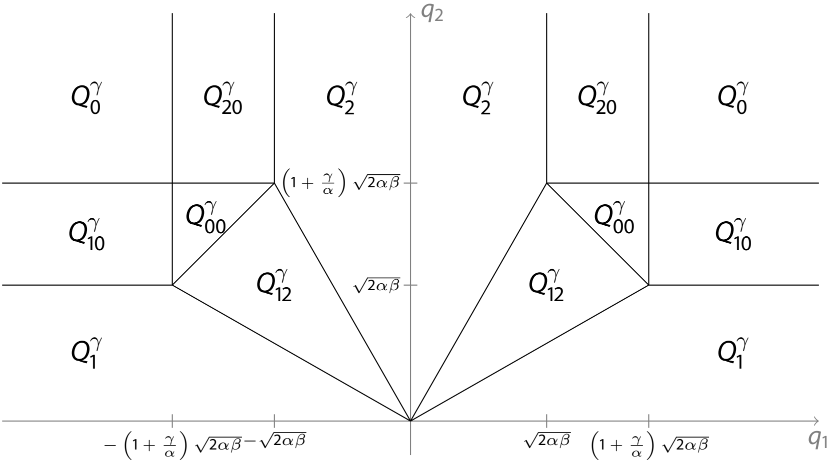
$$w := \operatorname{prox}_{\gamma g^*}(q) = (\operatorname{Id} + \gamma\partial g^*)^{-1}(q)$$

Solve for w in $v \in w + \gamma\partial g^*(w)$, case distinction

Moreau–Yosida regularization

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

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

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

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

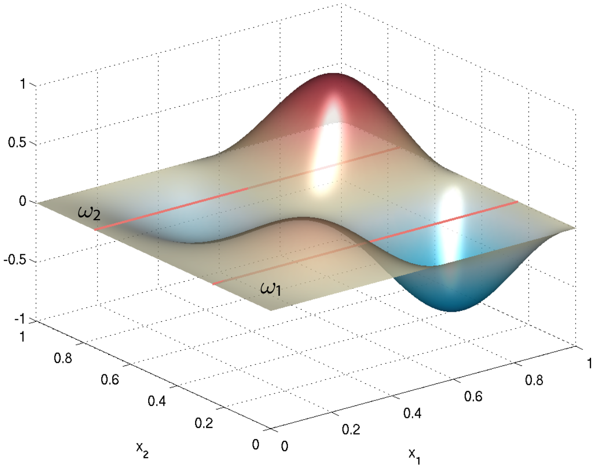
- Elliptic example: $S(u) = y$ solves

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

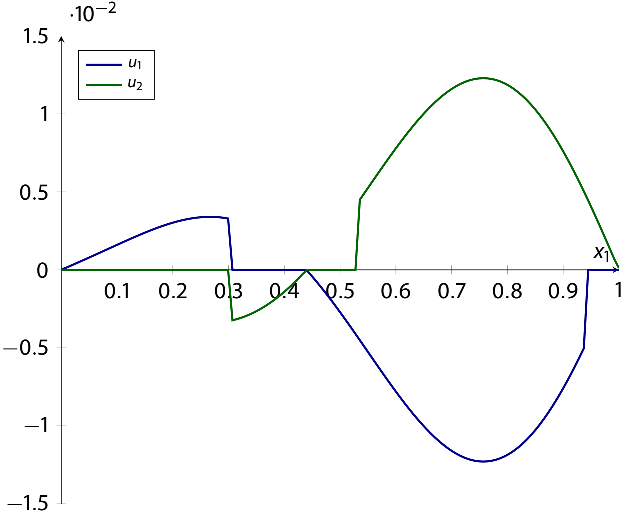
- Target

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

Numerical example: target

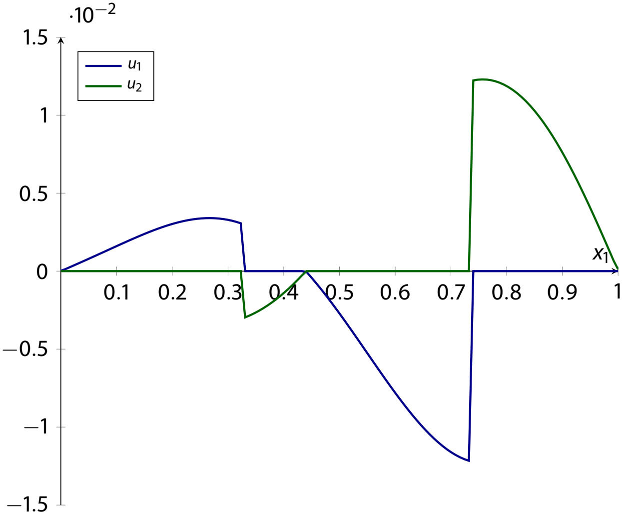


Numerical example: controls



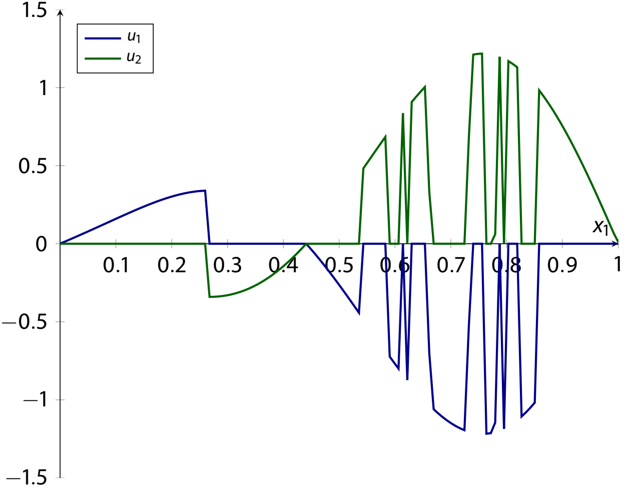
(a) $\alpha = 10^{-3}, \beta = 10^{-8}$

Numerical example: controls



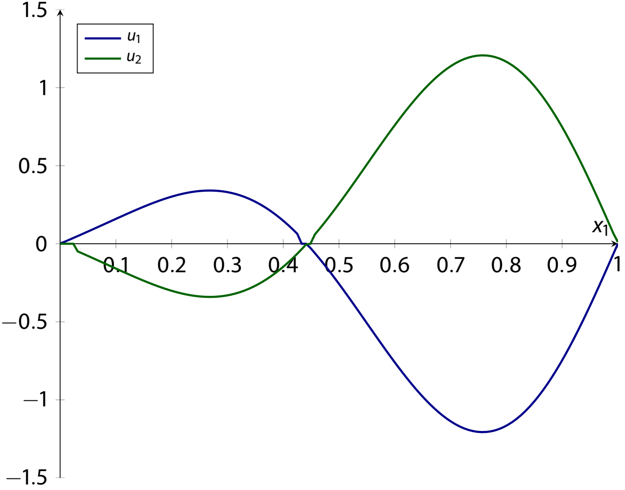
(b) $\alpha = 10^{-3}, \beta = 10^{-3}$

Numerical example: controls



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

Numerical example: controls



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

Conclusion

(Non)convex relaxation of discrete control problem:

- well-posed primal-dual optimality system
- amenable to semismooth Newton method
- efficient numerical solution of switching problems

Outlook:

- generalized switching (at most d out of m active)
- nonlinear control-to-state mapping
- other hybrid discrete–continuous problems

Preprint, MATLAB codes:

`http://www.udue.de/clason`