Multi-bang control of elliptic equations

Christian Clason[1] Karl Kunisch[2]

[1] Faculty of Mathematics, Universität Duisburg-Essen
[2] Institute of Mathematics and scientific Computing, Karl-Franzens-Universität Graz

SIAM Conference on Optimization
San Diego, May 22, 2014

Motivation

Optimal control of elliptic PDE $Ay = u$

Bang-bang control:

- control constraints $u_1 \leq u(x) \leq u_2$, no control costs
- $\rightsquigarrow$ control $u(x) \in \{u_1, u_2\}$ almost everywhere

Multi-bang control

- control $u(x) \in \{u_1, \ldots, u_d\}$ almost everywhere
- motivation: control by discrete voltages, velocities, …
- hybrid discrete–continuous problem, combinatorial
- solve using continuous relaxation $\rightsquigarrow$ linear complexity in d

Formulation

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

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

- $u_1 < \cdots < u_d$, $d \geq 2$, desired control states
- $|t|_0 := \begin{cases} 0 & \text{if } t = 0 \\ 1 & \text{if } t \neq 0 \end{cases}$ binary penalty
- $\rightsquigarrow$ non-smooth, non-convex, not lower-semicontinuous
- $A : V \to V^*$ isomorphism for Hilbert space $V \hookrightarrow L^2(\Omega) \hookrightarrow V^*$

Approach

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

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

Necessary optimality conditions

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

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

Approach

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

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

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

Necessary optimality conditions

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

- $\mathcal{G}^*(p) = \sup_u \langle u, p \rangle - \mathcal{G}(u)$ Fenchel conjugate

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
- $\rightsquigarrow$ well-defined, unique solution $\bar{u}$ ($= \arg\min \mathcal{F}(u) + \mathcal{G}^{**}(u)$)
- but: $\bar{u}$ in general not minimizer of $J \rightsquigarrow$ sub-optimal

Optimality system

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

Here:

$$\mathcal{F} : L^2(\Omega) \to \mathbb{R}, \quad u \mapsto \frac{1}{2}\|A^{-1}u - z\|_{L^2}^2$$

$$\mathcal{G} : L^2(\Omega) \to \overline{\mathbb{R}}, \quad u \mapsto \int_\Omega \left(\frac{\alpha}{2} u(x)^2 + \beta \prod_{i=1}^{d} |u(x) - u_i|_0 \right) dx + \delta_U(u)$$

- δ_U indicator function of

$$U := \left\{ u \in L^2(\Omega) : u_1 \leq u(x) \leq u_d \quad \text{a.e.} \right\}$$

- $\mathcal{G}$ defined pointwise $\rightsquigarrow$ compute Fenchel conjugate, subdifferential pointwise

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 $\bar{v}$,

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

Fenchel conjugate

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

$$P_1 := \left\{ q : q - \alpha u_1 < \sqrt{2\alpha\beta} \ \wedge \ q < \frac{\alpha}{2}(u_1 + u_2) \right\}$$

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

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

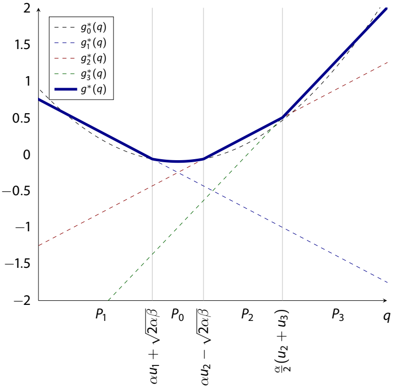
$$P_0 := \left\{ q : |q - \alpha u_j| > \sqrt{2\alpha\beta} \text{ for all } j \ \wedge \ \alpha u_1 < q < \alpha u_d \right\}$$

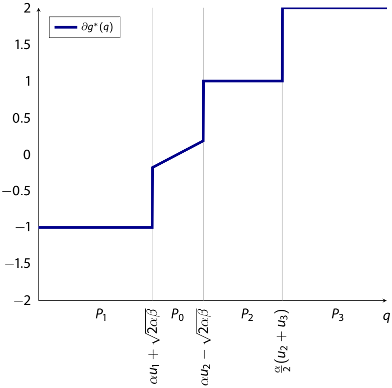
Fenchel conjugate

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

$$\partial g^*(q) = \begin{cases} \{u_i\} & q \in P_i,\ 1 \leq i < d \\ \{\frac{1}{\alpha}q\} & q \in P_0 \\ [u_i, u_{i+1}] & q \in \overline{P}_i \cap \overline{P}_{i+1},\ 1 \leq i < d \\ \left[\min\{u_i, \frac{1}{\alpha}q\}, \max\{u_i, \frac{1}{\alpha}q\}\right] & q \in \overline{P}_i \cap \overline{P}_0,\ 1 \leq i \leq d \end{cases}$$

Fenchel conjugate: sketch





Optimality system

$$
\begin{aligned}
-\bar{p} &= A^{-*}(A^{-1}\bar{u} - z)\\
\bar{u} &\in \partial\mathcal{G}^*(\bar{p})\\
&= \begin{cases}
\{u_i\} & \bar{p}(x) \in P_i\\
\{\frac{1}{\alpha}\bar{p}(x)\} & \bar{p}(x) \in P_0\\
[u_i, u_{i+1}] & \bar{p}(x) \in \overline{P}_i \cap \overline{P}_{i+1}\\
\left[\min(u_i, \frac{1}{\alpha}\bar{p}(x)), \max\{u_i, \frac{1}{\alpha}\bar{p}(x)\}\right] & \bar{p}(x) \in \overline{P}_i \cap \overline{P}_0
\end{cases}
\end{aligned}
$$

- necessary conditions for $\min_u \mathcal{F}(u) + \mathcal{G}^{**}(u)$ (convex, l.s.c.)
- $\rightsquigarrow$ unique solution $(\bar{u}, \bar{p})$

Relation to bang-bang control

- for $d = 2$ and β sufficiently large :

$$\bar{u}(x) \in \partial g^*(\bar{p}(x)) = \begin{cases} \{u_1\} & \text{if } \bar{p}(x) < \frac{\alpha}{2}(u_1 + u_d) \\ \{u_d\} & \text{if } \bar{p}(x) > \frac{\alpha}{2}(u_1 + u_d) \\ [u_1, u_d] & \text{if } \bar{p}(x) = \frac{\alpha}{2}(u_1 + u_d) \end{cases}$$

- bang-bang control

$$\bar{u}(x) \in \partial g^*(\bar{p}(x)) = \begin{cases} \{u_1\} & \text{if } \bar{p}(x) < 0 \\ \{u_d\} & \text{if } \bar{p}(x) > 0 \\ [u_1, u_d] & \text{if } \bar{p}(x) = 0 \end{cases}$$

- same for multi-bang with any $d \geq 2$ and $\alpha = 0$

  $\rightsquigarrow \alpha > 0$ necessary for multi-bang control

Structure of solution

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

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

Generalized multi-bang principle

- if β sufficiently large: $P_0 = \emptyset$, free arc

$$\mathcal{F} \subset \{\bar{p}(x) \in P_0\} = \emptyset$$

- singular arc corresponds to set-valued subdifferential:

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

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

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

(Sub)optimality

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

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

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

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

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

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

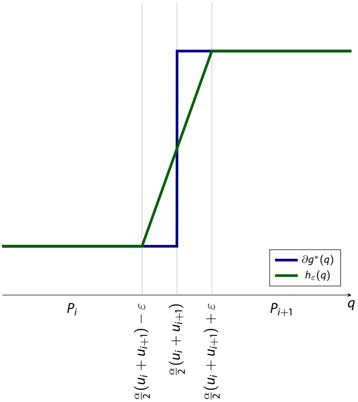
- but: $\bar{u}$ true multi-bang $\rightsquigarrow$ $|\mathcal{S}| = 0$ $\rightsquigarrow$ $\bar{u}$ optimal

Numerical solution of optimality system

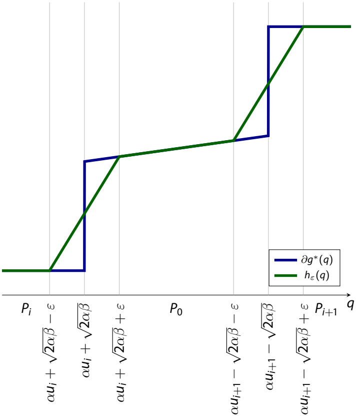
$$
\begin{aligned}
A\bar{y} &= \bar{u} \\
A^*\bar{p} &= z - \bar{y} \\
\bar{u} &\in \begin{cases}
\{u_i\} & \bar{p}(x) \in P_i \\
\{\frac{1}{\alpha}\bar{p}(x)\} & \bar{p}(x) \in P_0 \\
[u_i, u_{i+1}] & \bar{p}(x) \in \overline{P}_i \cap \overline{P}_{i+1} \\
\left[\min(u_i, \frac{1}{\alpha}\bar{p}(x)), \max\{u_i, \frac{1}{\alpha}\bar{p}(x)\}\right] & \bar{p}(x) \in \overline{P}_i \cap \overline{P}_0
\end{cases}
\end{aligned}
$$

- set-valued, not differentiable
- $\rightsquigarrow$ replace set-valued ∂g^* by piecewise linear h_ε, let $\varepsilon \to 0$

Regularization: sketch



(a) $q \in \overline{P}_i \cap \overline{P}_{i+1}$ (no free arc)



(b) $q \in \overline{P}_i \cap \overline{P}_0$ (free arc)

Regularized system

$$\begin{cases} Ay_\varepsilon = u_\varepsilon \\ A^* p_\varepsilon = z - y_\varepsilon \\ u_\varepsilon = H_\varepsilon(p_\varepsilon) \end{cases}$$

- H_ε maximal monotone $\rightsquigarrow$ unique solution $(u_\varepsilon, p_\varepsilon, y_\varepsilon)$
- convergence $(u_\varepsilon, p_\varepsilon, y_\varepsilon) \rightharpoonup (\bar{u}, \bar{p}, \bar{y})$ as $\varepsilon \to 0$
- h_ε Lipschitz continuous, norm gap $\rightsquigarrow$ H_ε semismooth
- only number of sets P_i depends on d $\rightsquigarrow$ linear complexity

Numerical examples

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

- $\Omega = [0,1]^2, \quad A = -\Delta$

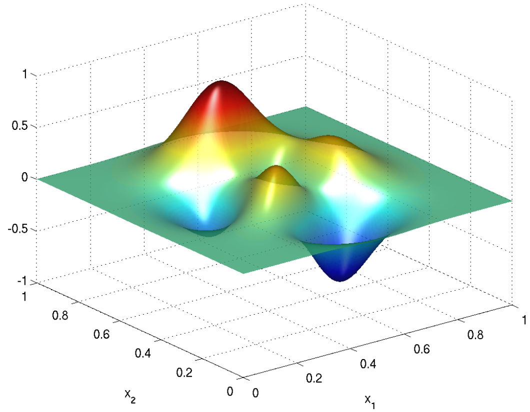
- finite element discretization: uniform grid, 256×256 nodes
- state, adjoint: piecewise linear
- control: eliminated (variational discretization)

- $d = 5, \quad (u_1, \dots, u_5) = (-2, 1, 0, 1, 2)$

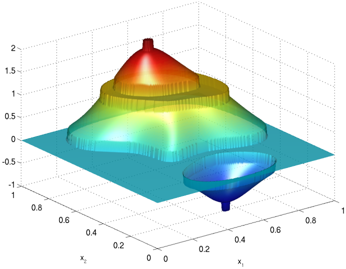
- $\varepsilon = 0$: regularized active sets empty, true multi-bang

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

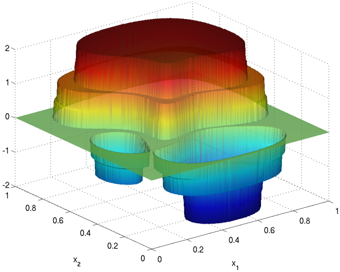
Numerical examples: target



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

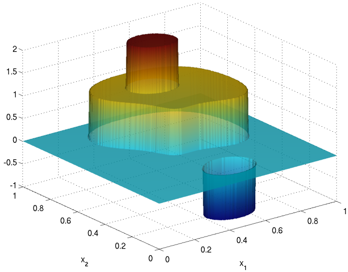


(a) $\alpha = 5 \cdot 10^{-3}$ ($\varepsilon = 0$)

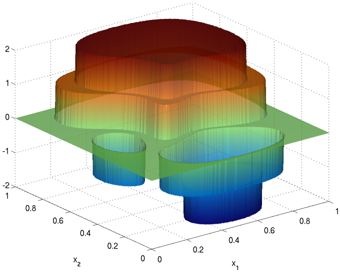


(b) $\alpha = 10^{-3}$ ($\varepsilon \approx 10^{-8}$)

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

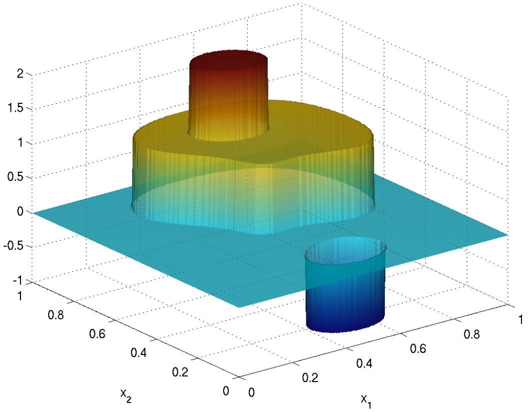


(c) $\alpha = 5 \cdot 10^{-3}$ ($\varepsilon = 0$)



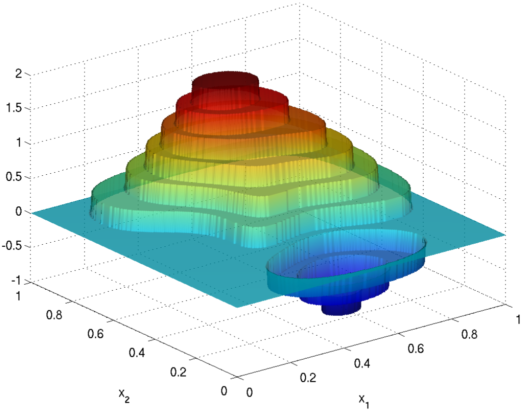
(d) $\alpha = 10^{-3}$ ($\varepsilon \approx 10^{-7}$)

Controls: effect of d



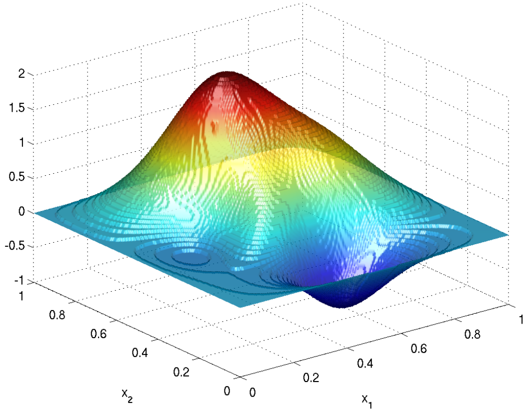
(a) $d = 5$ $(\varepsilon = 0)$

Controls: effect of d

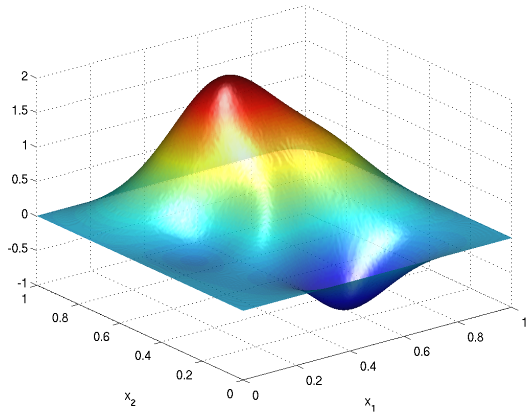


(b) $d = 15$ ($\varepsilon = 0$)

Controls: effect of d



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



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

Conclusion

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

(Non-)convex relaxation of discrete control problem:

- well-posed primal-dual optimality system
- controls optimal under reasonable assumptions
- linear complexity in number of desired states
- $\rightsquigarrow$ efficient numerical solution (superlinear convergence)

Outlook:

- inverse problems (identification of known tissue types)
- nonlinear control-to-state mapping
- other hybrid discrete–continuous problems (e.g., switching)

Preprint, MATLAB/Python codes:

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