Multibang regularization of a coefficient inverse problem for the wave equation

Christian Clason[1] Tram Do[1]
Karl Kunisch[2] Philip Trautmann[2]

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

Computational Methods for New Directions in Inverse Problems
College Station, TX, USA, February 4, 2020

Motivation: hybrid discrete optimization

UNIVERSITÄT DUISBURG ESSEN Open-Minded

$$\min_{u \in U} \mathcal{F}(u) + \frac{\alpha}{2}\|u\|^2$$

- $\mathcal{F}$ discrepancy term (involving PDE)
- U discrete set

$$U = \left\{u \in L^p(\Omega) : u(x) \in \{u_1, \dots, u_d\} \text{ a.e.}\right\}$$

- $u_1, \dots, u_d$ given voltages, velocities, materials, …
  (assumed here: ranking by magnitude possible!)
- motivation: topology optimization, medical imaging

Motivation: penalty

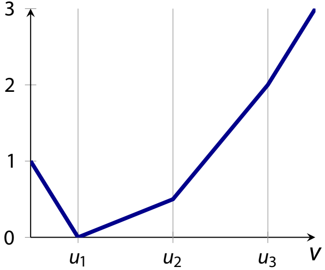
- convex relaxation: replace U by convex hull $u(x) \in [u_1, u_d]$
- works only for $d = 2$, cf. bang-bang control ($\alpha = 0$)
- $\rightsquigarrow$ promote $u(x) \in \{u_1, \dots, u_d\}$ by convex pointwise penalty

$$\mathcal{G}(u) = \int_\Omega g(u(x))\,dx$$

- generalize L^1 norm: polyhedral epigraph with vertices $u_1, \dots, u_d$
- not exact relaxation/penalization (in general)!

Motivation: penalty

- generalize L^1 norm: polyhedral epigraph with vertices $u_1, \dots, u_d$



- motivation: convex envelope of $\frac{1}{2}\|u\|^2 + \delta_U$
- multibang (generalized bang-bang) control
- $\rightsquigarrow$ non-smooth optimization in function spaces

Convex analysis approach

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

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:

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

Convex analysis approach

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

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

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

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

4. equivalent reformulation (for any $\sigma, \tau > 0$):

$$\begin{cases} \bar{p} = \operatorname{prox}_{\sigma\mathcal{F}^*}(\bar{p} + \sigma\bar{u}) \\ \bar{u} = \operatorname{prox}_{\tau\mathcal{G}}(\bar{u} - \tau\bar{p}) \end{cases}$$

Primal-dual proximal splitting

$$u^{k+1} = \mathrm{prox}_{\tau\mathcal{G}}\left(u^k - \tau K'(u^k)^* p^k\right)$$
$$\bar{u}^{k+1} = 2u^{k+1} - u^k$$
$$p^{k+1} = \mathrm{prox}_{\sigma\mathcal{F}^*}\left(p^k + \sigma K(\bar{u}^{k+1})\right)$$

- nonlinear variant of Chambolle–Pock (for $\mathcal{F}(Ku) + \mathcal{G}(u)$)
  [Valkonen '14, C./Mazurenko/Valkonen '18]
- $\tau, \sigma > 0$ step sizes
- local convergence in Hilbert space under
  1. second-order type condition on K
  2. τ, σ sufficiently small
- can be accelerated if $\mathcal{F}$ and/or $\mathcal{G}$ strongly convex

Convex analysis approach

For $\min_u \mathcal{F}(u) + \mathcal{G}(u)$, $\quad \mathcal{G}(u) = \int_\Omega g(u(x))\,dx$ convex, l.s.c.

Approach: pointwise

1. compute subdifferential ∂g (or Fenchel conjugate g^*)
2. compute conjugate subdifferential ∂g^*
3. compute proximal mapping $\text{prox}_{\gamma g}$

$\rightsquigarrow$ optimality conditions, proximal splitting methods

Multibang penalty

$$g : \mathbb{R} \to \overline{\mathbb{R}}, \qquad v \mapsto \begin{cases} \frac{1}{2}\left((u_i + u_{i+1})v - u_i u_{i+1}\right) & v \in [u_i, u_{i+1}] \\ \infty & \text{else} \end{cases}$$

piecewise differentiable $\rightsquigarrow$ subdifferential convex hull of derivatives

$$\partial g(v) = \begin{cases} \left(-\infty, \frac{1}{2}(u_1 + u_2)\right] & v = u_1 \\ \left\{\frac{1}{2}(u_i + u_{i+1})\right\} & v \in (u_i, u_{i+1}) \quad 1 \leqslant i < d \\ \left[\frac{1}{2}(u_{i-1} + u_i), \frac{1}{2}(u_i + u_{i+1})\right] & v = u_i \quad 1 < i < d \\ \left[\frac{1}{2}(u_{d-1} + u_d), \infty\right) & v = u_d \end{cases}$$

Multibang penalty

$$\partial g(v) = \begin{cases} \left(-\infty, \frac{1}{2}(u_1 + u_2)\right] & v = u_1 \\ \left\{\frac{1}{2}(u_i + u_{i+1})\right\} & v \in (u_i, u_{i+1}) \quad 1 \leqslant i < d \\ \left[\frac{1}{2}(u_{i-1} + u_i), \frac{1}{2}(u_i + u_{i+1})\right] & v = u_i \quad 1 < i < d \\ \left[\frac{1}{2}(u_{d-1} + u_d), \infty\right) & v = u_d \end{cases}$$

Fenchel duality:

$$\partial g^*(q) \in \begin{cases} \{u_1\} & q \in \left(-\infty, \frac{1}{2}(u_1 + u_2)\right) \\ [u_i, u_{i+1}] & q = \frac{1}{2}(u_i + u_{i+1}), \quad 1 \leqslant i < d \\ \{u_i\} & q \in \left(\frac{1}{2}(u_{i-1} + u_i), \frac{1}{2}(u_i + u_{i+1})\right) \quad 1 < i < d, \\ \{u_d\} & q \in \left(\frac{1}{2}(u_{d-1} + u_d), \infty\right) \end{cases}$$

Proximal point mapping

Proximal point mapping $\quad \mathrm{prox}_{\gamma g}(v) = w \;\text{ iff }\; v \in \{w\} + \gamma\partial g(w)$

case-wise inspection of subdifferential:

$$\mathrm{prox}_{\gamma g}(v) = \begin{cases} u_i & \text{if } v \in P_i^\gamma \\ v - \frac{\gamma}{2}(u_i + u_{i+1}) & \text{if } v \in P_{i,i+1}^\gamma \end{cases}$$

$$P_i^\gamma = \left[\left(1+\tfrac{\gamma}{2}\right)u_i + \tfrac{\gamma}{2}u_{i-1}, \left(1+\tfrac{\gamma}{2}\right)u_{i-1} + \tfrac{\gamma}{2}u_i\right]$$
$$P_{i,i+1}^\gamma = \left(\left(1+\tfrac{\gamma}{2}\right)u_i + \tfrac{\gamma}{2}u_{i+1}, \left(1+\tfrac{\gamma}{2}\right)u_{i+1} + \tfrac{\gamma}{2}u_i\right)$$

$\rightsquigarrow$ generalized soft thresholding

Multibang regularization

$$\min_{u \in L^2(\Omega)} \frac{1}{2} \|K(u) - y^\delta\|_Y^2 + \alpha\, \mathcal{G}(u)$$

- $K : L^2(\Omega) \to Y$ (nonlinear) forward mapping, weakly closed
- $y^\delta \in L^2(\Omega)$ noisy data with $\|y - y^\delta\|_Y \leqslant \delta$
- $u_1 < \cdots < u_d$ given parameter values ($d > 2$)
- $\mathcal{G}$ multibang penalty

Multibang regularization

$$\min_{u \in L^2(\Omega)} \frac{1}{2} \|K(u) - y^\delta\|_Y^2 + \alpha\, \mathcal{G}(u)$$

- $\mathcal{G}$ multibang penalty convex:
  1. existence of solution u_α^δ for every $\alpha > 0$
  2. $\delta \to 0$ implies $u_\alpha^\delta \rightharpoonup u_\alpha$ for every $\alpha > 0$
  3. $\delta \to 0$, $\alpha \to 0$, $\delta^2 \alpha^{-1} \to 0$ implies $u_\alpha^\delta \rightharpoonup u^\dagger$

(standard arguments, e.g. [Burger/Osher 04, Scherzer et al. 09, Kaltenbacher et al. '12, Ito/Jin 14])

Multibang regularization

$$\min_{u\in L^2(\Omega)} \frac{1}{2}\|K(u) - y^\delta\|_Y^2 + \alpha\, \mathcal{G}(u)$$

- standard source condition: $p^\dagger := K'(u^\dagger)^* w \in \partial\mathcal{G}(u^\dagger)$ for $w \in Y$,
  1. a priori choice $\alpha(\delta) = c\delta$
  2. a posteriori choice $\|K(u^\delta_{\alpha(\delta)}) - y^\delta\|_Y \leqslant \tau\delta, \tau > 1$

  $\rightsquigarrow$ convergence rate

$$d_{\mathcal{G}}^{p^\dagger}(u_\alpha^\delta, u^\dagger) \leqslant C\delta$$

  in Bregman divergence

$$d_{\mathcal{G}}^{p_1}(u_2, u_1) = \mathcal{G}(u_2) - \mathcal{G}(u_1) - \langle p_1, u_2 - u_1\rangle_X, \qquad p_1 \in \partial\mathcal{G}(u_1)$$

Pointwise definition of Bregman divergence, ∂g:

- $u^\dagger(x) = u_i$ and $p^\dagger \notin \left\{\frac{1}{2}(u_{i-1}+u_i), \frac{1}{2}(u_i+u_{i+1})\right\}$ implies

$$d_g^{p^\dagger(x)}(u^\delta_{\alpha(\delta)}(x), u^\dagger(x)) \to 0 \qquad \text{for } \delta \to 0$$

- $u^\dagger(x) \in (u_i, u_{i+1})$ implies

$$d_g^{p^\dagger(x)}(u(x), u^\dagger(x)) = 0 \qquad \text{for any } u(x) \in [u_i, u_{i+1}]$$

- $\rightsquigarrow$ $u^\delta_{\alpha(\delta)} \to u^\dagger$ pointwise a.e. iff $u^\dagger(x) \in \{u_1, \dots, u_d\}$ a.e.

- (convergence not uniform $\rightsquigarrow$ no pointwise rates)

Optimality system

$$\bar{p} = \frac{1}{\alpha} K'(\bar{u})^*(y^\delta - K(\bar{u}))$$

$$\bar{u} \in \partial \mathcal{G}^*(\bar{p}) = \begin{cases} \{u_i\} & \bar{p}(x) \in Q_i \\ [u_i, u_{i+1}] & \bar{p}(x) \in \overline{Q}_i \cap \overline{Q}_{i+1} \end{cases}$$

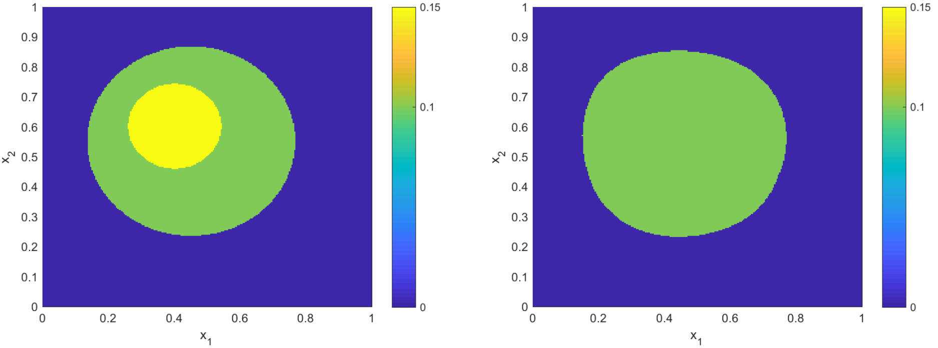
- $\rightsquigarrow$ unique solution $(\bar{u}, \bar{p}) \in L^2(\Omega) \times L^2(\Omega)$
- singular set $\mathcal{S} = \left\{x : \bar{u}(x) \neq u_i\right\} \subset \left\{x : \bar{p}(x) = \frac{1}{2}(u_i + u_{i+1})\right\}$
- for suitable K, $\bar{p}(x)$ constant implies $[y^\delta - K(\bar{u})](x) = 0$ (e.g., $K = A^{-1}$ for A pure second-order elliptic)

$\rightsquigarrow$ $\left|\{x : K(\bar{u}(x)) = y^\delta(x)\}\right| = 0 \;\Rightarrow\; \bar{u} \in \{u_1, \dots u_d\}$ a. e. (true multibang)

Example: linear inverse problem

- $\Omega = [0,1]^2, \quad K = A^{-1}, \quad A = -\Delta$
- $$u^\dagger(x) = u_1 + u_2\, \chi_{\{x:(x_1-0.45)^2+(x_2-0.55)^2<0.1\}}(x) + (u_3 - u_2)\, \chi_{\{x:(x_1-0.4)^2+(x_2-0.6)^2<0.02\}}(x)$$
- $d = 3, \quad u_1 = 0, \quad u_2 = 0.1, \quad u_3 \in \{0.15, 0.11\}$
- $y^\delta = y^\dagger + \xi, \quad \xi \in \mathcal{N}\left(y^\dagger, \delta\|y^\dagger\|_\infty\right)$
- finite element discretization: uniform grid, 256×256 nodes
- solution by (regularized) semismooth Newton method ($\gamma < 10^{-12}$)
- $\alpha = \alpha(\delta)$ by Morozov discrepancy principle

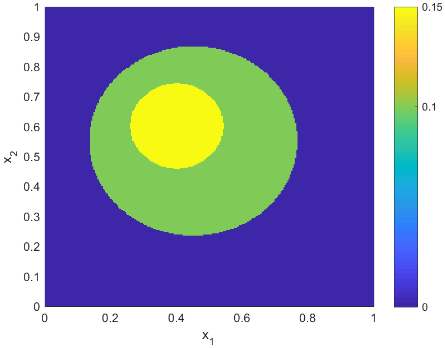
Numerical example: $u_3 = 0.15$



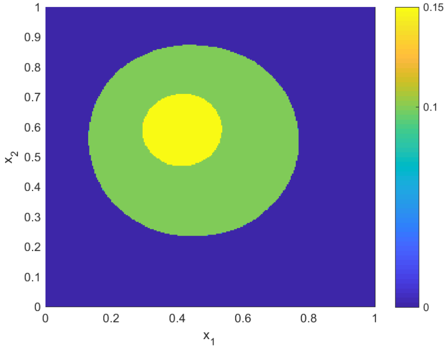
(a) $u^\dagger$

(b) u_α^δ, $\delta \approx 1.89 \cdot 10^{-1}$

Numerical example: $u_3 = 0.15$

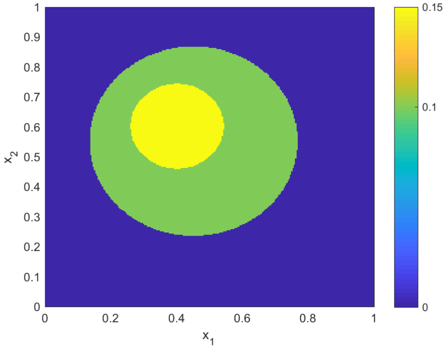


(c) $u^{\dagger}$

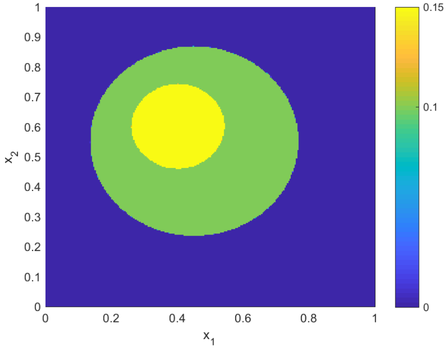


(d) u_{α}^{δ}, $\delta \approx 2.37 \cdot 10^{-2}$

Numerical example: $u_3 = 0.15$

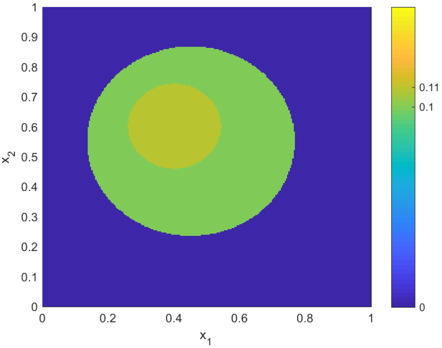


(e) $u^\dagger$

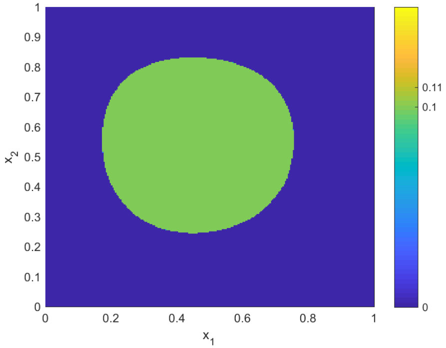


(f) u_α^δ, $\delta \approx 3.69 \cdot 10^{-4}$

Numerical example: $u_3 = 0.11$

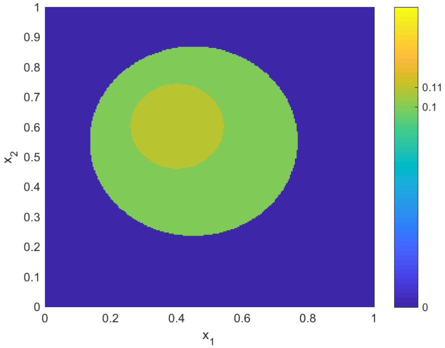


(a) $u^\dagger$

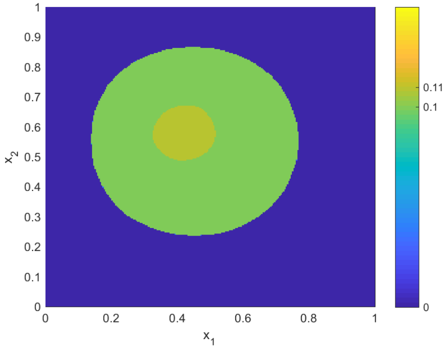


(b) u_α^δ, $\delta \approx 1.68 \cdot 10^{-1}$

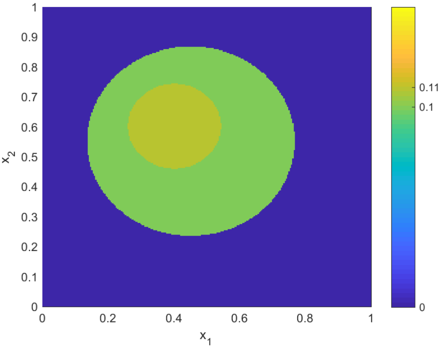
Numerical example: $u_3 = 0.11$



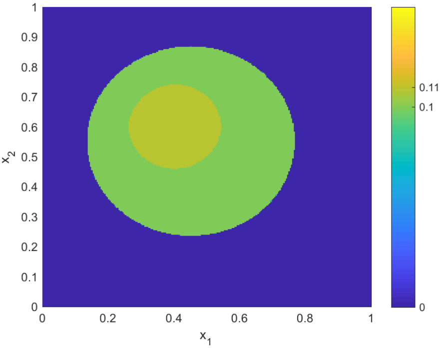
(c) $u^\dagger$



(d) u_α^δ, $\delta \approx 2.17 \cdot 10^{-2}$

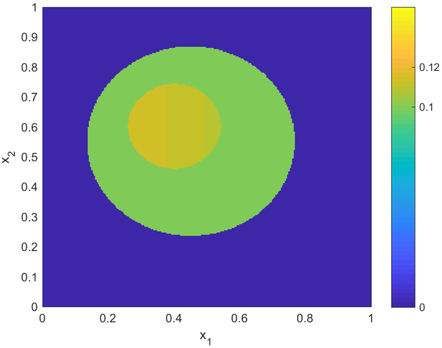


(e) $u^\dagger$

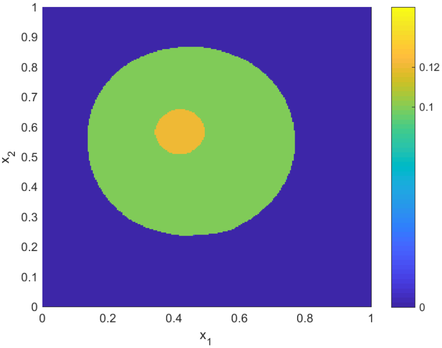


(f) u_α^δ, $\delta \approx 3.29 \cdot 10^{-4}$

Numerical example: $u_3(x) = 0.12(1 - x_1)$

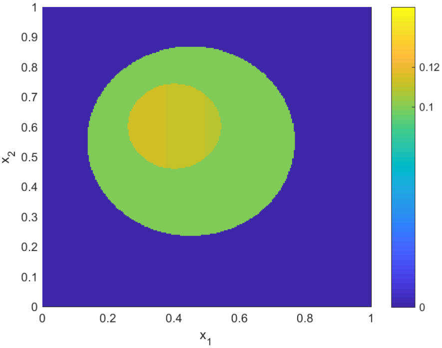


(a) $u^\dagger$

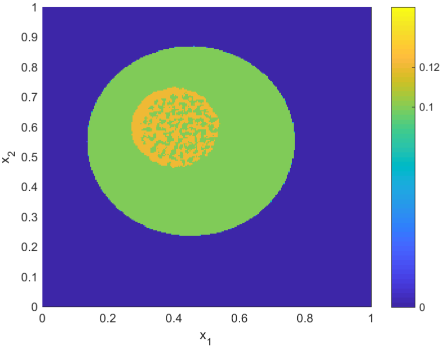


(b) u_α^δ, $\delta \approx 2.11 \cdot 10^{-2}$

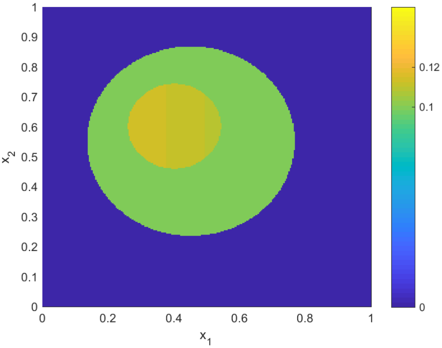
Numerical example: $u_3(x) = 0.12(1 - x_1)$



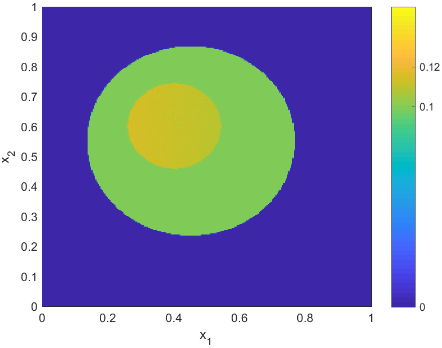
(c) $u^\dagger$



(d) u_α^δ, $\delta \approx 3.29 \cdot 10^{-4}$



(e) $u^\dagger$



(f) u_α^δ, $\delta \approx 1.29 \cdot 10^{-6}$

Nonlinear forward mapping

Goal: application to coefficient inverse problem for wave equation

- $K : u \mapsto y$ solving

$$\begin{cases} y_{tt} - \nabla \cdot (u \nabla y) = f \\ y(0) = y_0, \quad y_t(0) = 0 \end{cases}$$

- difficulty: $\bar{u} \in L^\infty(\Omega) \quad \rightsquigarrow \quad K$ not weakly-$*$ closed

  $\rightsquigarrow$ lack of existence of minimizer ($\bar{y} \neq K(\bar{u})$, cf. homogenization)

- $\rightsquigarrow$ TV regularization: add $TV(u) := \|Du\|_{\mathcal{M}}$

- $\rightsquigarrow u \in BV(\Omega) \cap L^\infty(\Omega) \hookrightarrow_c L^p(\Omega)$

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

Difficulty:

- existence requires box constraints $\rightsquigarrow$ use penalty

$$\left(G(u) + \delta_{[u_1,u_d]}(u)\right) + TV(u)$$

(here: G multibang penalty with $\operatorname{dom} G = L^1(\Omega)$)

- but: $TV(u) + \delta_{[u_1,u_d]}(u)$ not continuous on $L^p(\Omega)$, $p < \infty$
- but: multipliers $\xi \in \partial TV(u)$, $q \in \partial G(u)$ not pointwise on BV, L^∞
- $\rightsquigarrow$ replace box constraints by ($C^{1,1}$) projection of $u \in L^1(\Omega)$

$$[\Phi_\varepsilon(u)](x) = \operatorname{proj}^\varepsilon_{[u_1,u_d]}(u(x)) \qquad \text{a.e. } x \in \Omega$$

- $\rightsquigarrow$ use higher regularity of solution to wave equation

Wave equation

$$\begin{cases} \int_0^T -(\partial_t y, \partial_t v) + (u \nabla y(t), \nabla v(t))\, dt = \int_0^T (f(t), v(t))\, dt \\ y(0) = y_0 \end{cases}$$

for all $v \in W := L^2(0,T;H^1) \cap H^1(0,T;L^2)$ with $v(T) = 0$ [Lions/Magenes '72]

$\rightsquigarrow$ solution mapping $S : u \mapsto y$ on $U := \{u \in L^\infty(\Omega) : u_1 \leqslant u \leqslant u_d \text{ a.e.}\}$

- $S(u)$ uniformly bounded in $W \cap H^2(0,T;H^{-1}) := Z$
- S Lipschitz continuous from L^∞ to $L^2(0,T;L^2)$
- $S(u_n) \rightharpoonup S(u)$ in Z if $u_n \to u$ in $L^r(\Omega)$, $r \in [1,\infty]$

$$\begin{cases} \int_0^T -(\partial_t y, \partial_t v) + (u\nabla y(t), \nabla v(t))\, dt = \int_0^T (f(t), v(t))\, dt \\ y(0) = y_0 \end{cases}$$

for all $v \in W := L^2(0,T;H^1) \cap H^1(0,T;L^2)$ with $v(T) = 0$ [Lions/Magenes '72]

Assumption:

- $f \in L^2(0,T;H^1)$, $\quad y_0 \in H^2(\Omega), \partial_\nu y_0 = 0, \quad y_1 \in H^1(\Omega)$
- there is $\omega_c \subset \Omega$ with u constant on $\Omega \setminus \omega_c$, y_0 constant on ω_c

Then:

- $S(u)$ uniformly bounded in $L^\infty(0,T;W^{1,s})$ for some $s > 2$ (proof: combination of higher hyperbolic and maximal elliptic regularity [Wloka '87, Gröger '89])

TV regularization: existence

$$\begin{cases} \min_{u \in BV(\Omega)} \frac{1}{2}\|y - y^\delta\|^2_{L^2(\Omega)} + \alpha\, G(u) + \beta\, TV(u) \\ \text{s.t.} \quad y_{tt} - \nabla \cdot (\Phi_\varepsilon(u)\nabla y) = f \\ \qquad y(0) = y_0, \quad y_t(0) = y_1 \end{cases}$$

- existence of optimal $\bar{u} \in BV(\Omega) \cap U$ for $\varepsilon \geqslant 0$
- tracking term Fréchet differentiable in $\Phi_\varepsilon(u) \in L^\infty$ for $\varepsilon > 0$
- improved regularity of state $\rightsquigarrow$ derivative in $L^r(\Omega)$, $r > 1$ (instead of $L^\infty(\Omega)^*$)
- $\rightsquigarrow$ sum rule applicable, subgradients in $L^r(\Omega)$, $r > 1$

TV regularization: optimality conditions

$$\begin{cases} 0 = F'(\Phi(\bar{u}))\Phi'_\varepsilon(\bar{u}) + \alpha\bar{q} + \beta\bar{\xi} \\ \bar{u} \in \partial G^*(\bar{q}) \\ \bar{\xi} \in \partial TV(\bar{u}) \end{cases}$$

- $F'(\Phi_\varepsilon(\bar{u})) = \int_0^T \nabla\bar{y} \cdot \nabla\bar{p}\, dt \in L^r(\Omega)$, $r > 1$
  ($\bar{y}$ optimal state, $\bar{p}$ adjoint state)
- $\bar{q} \in L^r(\Omega)$, $r > 1 \rightsquigarrow$ pointwise multibang
- $\bar{\xi} \in L^r(\Omega)$, $r > 1 \rightsquigarrow$ characterization via *full trace* [Bredies/Holler '12]
- $\rightsquigarrow$ pointwise optimality conditions

Numerical solution

Approach: discretize before optimize

- consider finite element discretization of problem
  - piecewise linear in space
  - stabilized piecewise linear in time [Zlotnik '94]
  - discrete adjoint
- include projection in multi-bang penalty, eliminate Φ_ε
- apply sum rule, chain rule for $\partial TV(u_h) = -\mathrm{div}_h \partial(\|\cdot\|_1)(\nabla_h u_h)$
- $\rightsquigarrow$ apply nonlinear primal-dual proximal splitting

Primal-dual proximal splitting

$$
\begin{aligned}
u^{k+1} &= \operatorname{prox}_{\tau\mathcal{G}}\left(u^k - \tau K'(u^k)^* p^k\right)\\
\bar{u}^{k+1} &= 2u^{k+1} - u^k\\
p^{k+1} &= \operatorname{prox}_{\sigma\mathcal{F}^*}\left(p^k + \sigma K(\bar{u}^{k+1})\right)
\end{aligned}
$$

- nonlinear variant of Chambolle–Pock

  [Valkonen '14, C./Mazurenko/Valkonen '18]
- $\tau, \sigma > 0$ step sizes
- local convergence in Hilbert space under
  1. second-order type condition on K
  2. τ, σ sufficiently small
- apply to $\mathcal{F}(y,q) = \|y - y^\delta\|_2^2 + \||q|_2\|_1, \quad K(u) = (S(u), \nabla_h u)$

Primal-dual proximal splitting algorithm

$$
\begin{aligned}
u^{k+1} &= \operatorname{prox}_{\tau\alpha\mathcal{G}}\left(u^k - \tau S_h'(u^k)^*(r^k) - \tau\nabla_h^*\psi^k\right)\\
\bar{u}^{k+1} &= 2u^{k+1} - u^k\\
r^{k+1} &= \frac{1}{1+\frac{\sigma}{\nu_1}}\left(r_1^k + \sigma(S_h(\bar{u}^{k+1}) - y^d)\right)\\
q^{k+1} &= \psi^k + \sigma\nabla_h\bar{u}^{k+1}\\
\psi^{k+1} &= \frac{\beta q^{k+1}}{\max\{\beta, |q^{k+1}|_2\}}
\end{aligned}
$$

- $S_h(u)$ solution of wave equation
- $S_h'(u)^* r$ solution of wave, adjoint equation (with RHS r), integration
- proximal mappings pointwise ($\mathcal{G}$ includes projection)

Numerical example

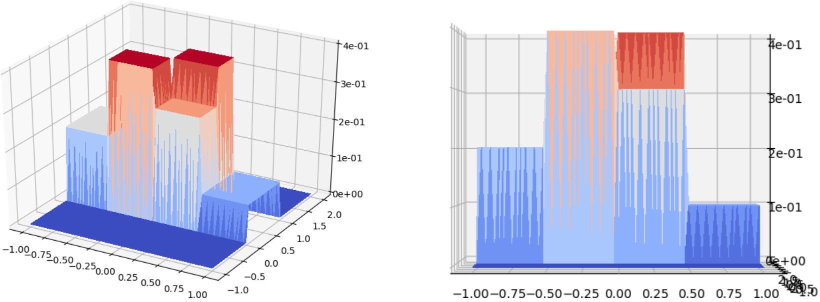
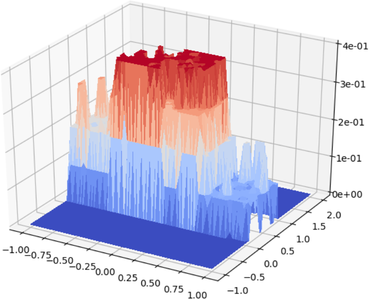


Figure: exact coefficient (front: sources; back: observation)

Numerical example



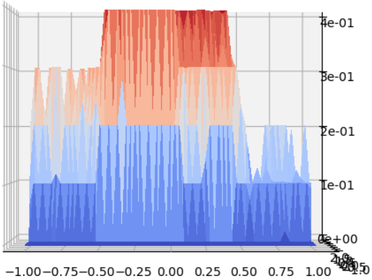


Figure: $\alpha = 10^{-5}$, $\beta = 0$, 3680 iterations

Numerical example

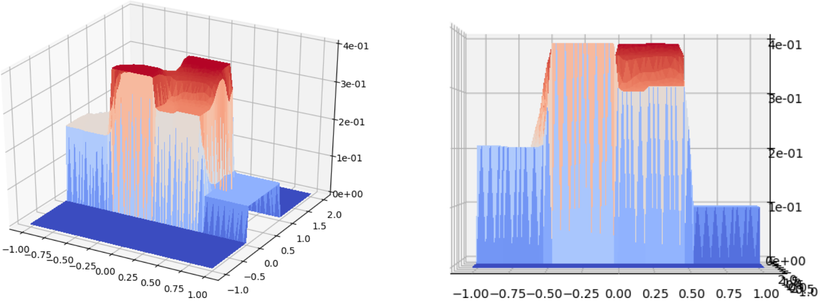


Figure: $\alpha = 0$, $\beta = 10^{-4}$, 1100 iterations

Numerical example

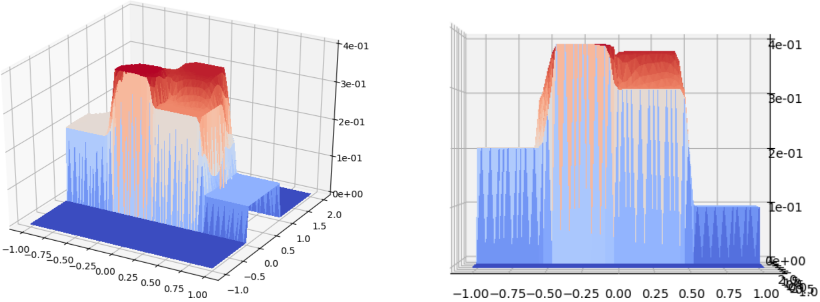


Figure: $\alpha = 10^{-5}$, $\beta = 10^{-4}$, 600 iterations

Conclusion

Multibang regularization for discrete-valued inverse problem

- well-posed convex relaxation
- combination with total variation
- applicable to wave equation

Outlook:

- (block) acceleration of proximal splitting
- boundary observation
- total generalized variation
- vector-valued coefficient

Preprints, codes:

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