

Discrete regularization of parameter identification problems

Christian Clason

Faculty of Mathematics, Universität Duisburg-Essen

joint work with Florian Kruse, Karl Kunisch

Applied Inverse Problems Conference 2017
Hangzhou, May 31, 2017

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

- \mathcal{F} discrepancy term (involving PDEs)
- U discrete set,

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

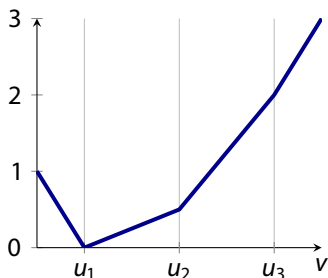
- u_1, \dots, u_d given voltages, velocities, materials, ...
(assumed here: ranking by magnitude possible!)
- motivation: topology optimization, medical imaging

- **convex relaxation**: replace U by convex hull
- 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)!

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



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

- 1 Overview
- 2 Approach
 - Convex analysis
 - Moreau–Yosida regularization
 - Semismooth Newton method
- 3 Multi-bang penalty
- 4 Nonlinear multi-bang

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

■ subdifferential

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

■ Fenchel conjugate (always convex)

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

■ “convex inverse function theorem”:

$$v^* \in \partial\mathcal{F}(v) \iff v \in \partial\mathcal{F}^*(v^*)$$

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

- 1 Fermat principle: $0 \in \partial(\mathcal{F}(\bar{u}) + \mathcal{G}(\bar{u}))$
- 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:

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

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

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

Proximal mapping

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

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

Proximal mapping

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

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

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

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

Proximal mapping

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

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

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

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

f locally Lipschitz, piecewise C^1 :

$$f(v) = 0, \quad f : \mathbb{R}^n \rightarrow \mathbb{R}$$

Newton derivative

$$D_N f(v) \delta v \in \partial_C f(v) \delta v$$

Clarke generalized gradient: convex hull of piecewise derivatives

semismooth Newton method

$$D_N f(v^k) \delta v = -f(v^k), \quad v^{k+1} = v^k + \delta v$$

converges **locally superlinearly**

f locally Lipschitz, piecewise C^1 :

$$F(u) = 0, \quad F : L^r(\Omega) \rightarrow L^s(\Omega), \quad [F(u)](x) = f(u(x))$$

Newton derivative

$$[D_N F(u)\delta u](x) \in \partial_C f(\delta u(x))\delta u(x)$$

any measurable selection of Clarke generalized gradient

semismooth Newton method

$$D_N F(u^k)\delta u = -F(u^k), \quad u^{k+1} = u^k + \delta u$$

converges **locally superlinearly** if $r > s$

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

Approach: pointwise

- 1 compute subdifferential ∂g (or Fenchel conjugate g^*)
 - 2 compute subdifferential ∂g^*
 - 3 compute proximal mapping $\text{prox}_{\gamma g^*}$
 - 4 compute Moreau–Yosida regularization ∂g_{γ}^*
 - 5 compute Newton derivative $D_N \partial g_{\gamma}^*$
- ↪ 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 analysis
 - Moreau–Yosida regularization
 - Semismooth Newton method
- 3 Multi-bang penalty**
- 4 Nonlinear multi-bang

$$\begin{cases} \min_{u \in L^2(\Omega)} \frac{1}{2} \|y - y^\delta\|_{L^2}^2 + \alpha \mathcal{G}(u) \\ \text{s. t. } Ay = u, \quad u_1 \leq u(x) \leq u_d \text{ a.e.} \end{cases}$$

- $u_1 < \dots < u_d$ given parameter values ($d > 2$)
- $y^\delta \in L^2(\Omega)$ noisy data
- $A : V \rightarrow V^*$ isomorphism for Hilbert space $V \hookrightarrow L^2(\Omega) \hookrightarrow V^*$
(e.g., elliptic differential operator with boundary conditions)
- $\rightsquigarrow \mathcal{F}(u) = \frac{1}{2} \|A^{-1}u - y^\delta\|_{L^2}^2$ smooth
- \mathcal{G} multi-bang penalty (will include control constraints from now)

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

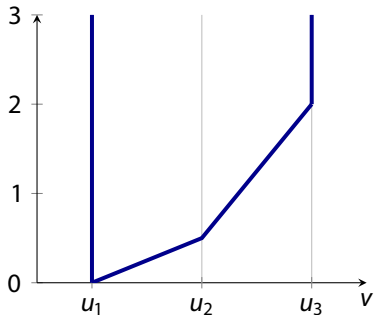
piecewise differentiable \rightsquigarrow subdifferential convex hull of derivatives

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

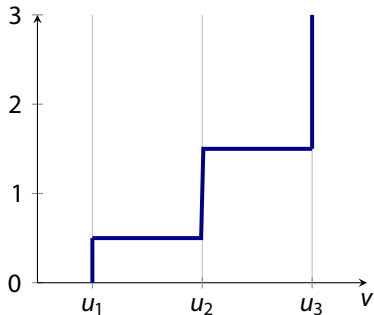
$$\partial g(v) = \begin{cases} (-\infty, \frac{1}{2}(u_1 + u_2)] & v = u_1 \\ \{\frac{1}{2}(u_i + u_{i+1})\} & v \in (u_i, u_{i+1}) \quad 1 \leq i < d \\ [\frac{1}{2}(u_{i-1} + u_i), \frac{1}{2}(u_i + u_{i+1})] & v = u_i \quad 1 < i < d \\ [\frac{1}{2}(u_{d-1} + u_d), \infty) & v = u_d \end{cases}$$

convex inverse function theorem:

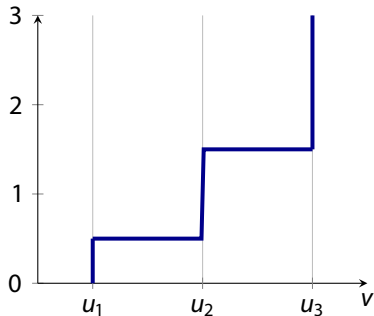
$$\partial g^*(q) \in \begin{cases} \{u_1\} & q \in (-\infty, \frac{1}{2}(u_1 + u_2)) \\ [u_i, u_{i+1}] & q = \frac{1}{2}(u_i + u_{i+1}), \quad 1 \leq i < d \\ \{u_i\} & q \in (\frac{1}{2}(u_{i-1} + u_i), \frac{1}{2}(u_i + u_{i+1})) \quad 1 < i < d, \\ \{u_d\} & q \in (\frac{1}{2}(u_{d-1} + u_d), \infty)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



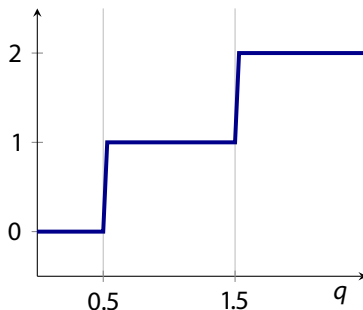
(a) $g(u_1 = 0, u_2 = 1, u_3 = 2)$



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



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



(d) $\partial g^*(u_1 = 0, u_2 = 1, u_3 = 2)$

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

- $S : u \mapsto y$ control-to-state mapping, S^* adjoint
- \rightsquigarrow unique solution $(\bar{u}, \bar{p}) \in L^2(\Omega) \times L^2(\Omega)$
- singular arc $\mathcal{S} = \{x : \bar{u}(x) \notin \{u_i\}\} \subset \{x : \bar{p}(x) = \frac{1}{2}(u_i + u_{i+1})\}$
- for suitable A , $\bar{p}(x)$ constant implies $[A^*\bar{p}](x) = [y^\delta - \bar{y}](x) = 0$
 $\rightsquigarrow |\{x : \bar{y}(x) = y^\delta(x)\}| = 0 \Rightarrow \bar{u} \in \{u_1, \dots, u_d\}$ a. e., true multi-bang

Proximal mapping $\text{prox}_{\gamma g^*}(q) = w$ iff $q \in \{w\} + \gamma \partial g^*(w)$

case-wise inspection of subdifferential:

$$\partial g_{\gamma}^*(q) = \frac{1}{\gamma} (q - \text{prox}_{\gamma g^*}(q)) = \begin{cases} u_i & q \in Q_i^{\gamma} \\ \frac{1}{\gamma} (q - \frac{1}{2}(u_i + u_{i+1})) & q \in Q_{i,i+1}^{\gamma} \end{cases}$$

$$Q_i^{\gamma} = \left(\frac{1}{2}(u_{i-1} + u_i) + \gamma u_i, \frac{1}{2}(u_i + u_{i+1}) + \gamma u_i \right)$$
$$Q_{i,i+1}^{\gamma} = \left[\frac{1}{2}(u_i + u_{i+1}) + \gamma u_i, \frac{1}{2}(u_i + u_{i+1}) + \gamma u_{i+1} \right]$$

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

- optimality conditions for $\mathcal{F}(u) + \alpha \mathcal{G}(u) + \frac{\gamma}{2} \|u\|^2$
- \rightsquigarrow unique solution (u_γ, p_γ)
- $(u_\gamma, p_\gamma) \rightarrow (\bar{u}, \bar{p})$ as $\gamma \rightarrow 0$
- $\partial \mathcal{G}_\gamma^*$ Lipschitz continuous, piecewise C^1 , norm gap $V \hookrightarrow L^2(\Omega)$
- \rightsquigarrow semismooth Newton method

$$\begin{cases} A^* p_\gamma = \frac{1}{a}(y^\delta - y_\gamma) \\ Ay_\gamma = \mathcal{G}_\gamma^*(p_\gamma) \end{cases}$$

- optimality conditions for $\mathcal{F}(u) + a \mathcal{G}(u) + \frac{\gamma}{2} \|u\|^2$
- \rightsquigarrow unique solution (u_γ, p_γ)
- $(u_\gamma, p_\gamma) \rightarrow (\bar{u}, \bar{p})$ as $\gamma \rightarrow 0$
- ∂g_γ^* Lipschitz continuous, piecewise C^1 , norm gap $V \hookrightarrow L^2(\Omega)$
- \rightsquigarrow semismooth Newton method
- introduce $y_\gamma = Su_\gamma$, eliminate $u_\gamma = \mathcal{G}_\gamma^*(p_\gamma)$

$$\begin{pmatrix} \frac{1}{a} \text{Id} & A^* \\ A & -D_N \mathcal{G}_\gamma^*(p) \end{pmatrix} \begin{pmatrix} \delta y \\ \delta p \end{pmatrix} = - \begin{pmatrix} A^* p + \frac{1}{a}(y - y^\delta) \\ Ay - \mathcal{G}_\gamma^*(p) \end{pmatrix}$$

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

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

- $\Omega = [0, 1]^2$, $A = -\Delta$
- $u^\dagger(x) = u_1 + u_2 \chi_{\{x: (x_1 - 0.45)^2 + (x_2 - 0.55)^2 < 0.1\}}(x)$
 $+ (u_3 - u_2) \chi_{\{x: (x_1 - 0.4)^2 + (x_2 - 0.6)^2 < 0.02\}}(x)$
- $d = 3$, $u_1 = 0$, $u_2 = 0.1$, $u_3 \in \{0.2, 0.12\}$
- $y^\delta = y^\dagger + \xi$, $\xi \in \mathcal{N}(y^\dagger, \delta \|y^\dagger\|_\infty)$
- finite element discretization: uniform grid, 256×256 nodes
- $\alpha = 5 \cdot 10^{-5}$
- terminate at $\gamma < 10^{-12}$

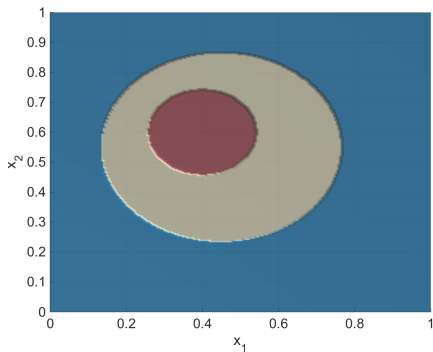
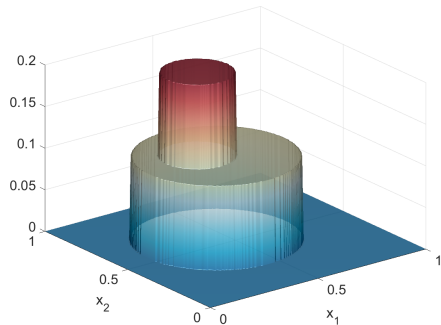


Figure: exact parameter u^\dagger

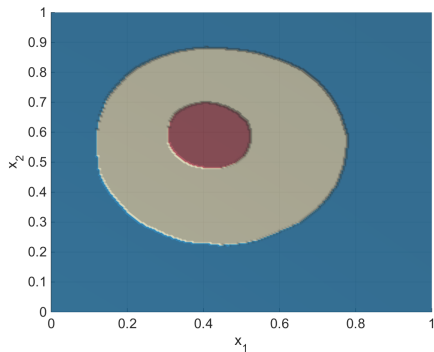
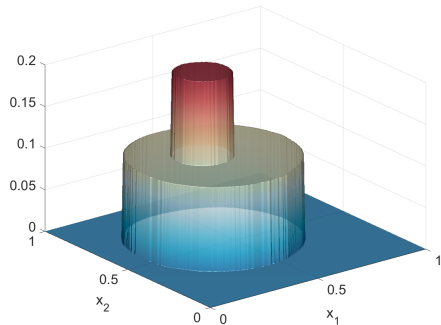


Figure: reconstruction u^δ , $\delta = 0.1$

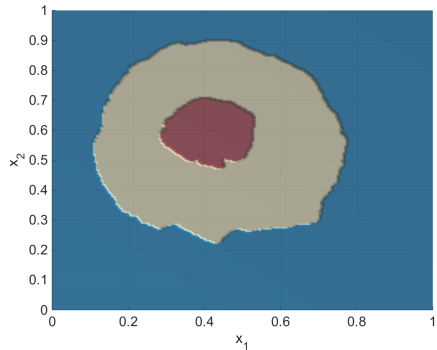
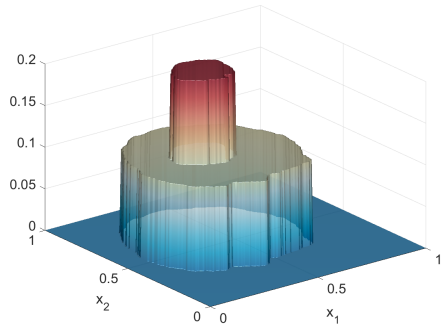


Figure: reconstruction u^δ , $\delta = 0.5$

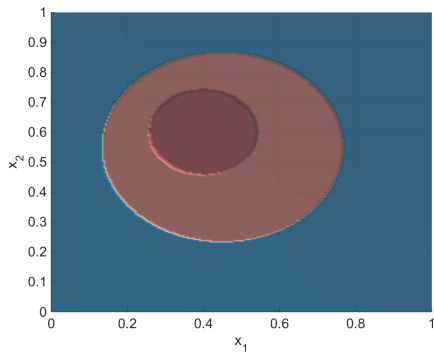
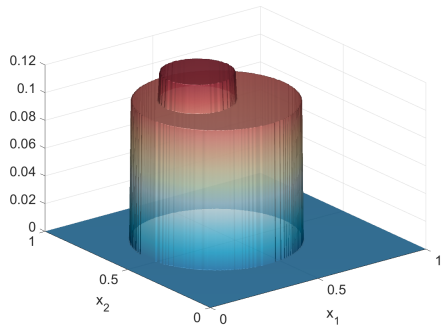


Figure: exact parameter u^\dagger

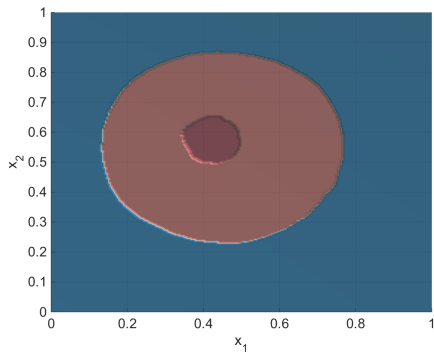
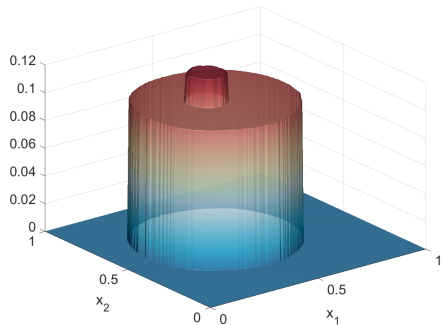


Figure: reconstruction u^δ , $\delta = 0.1$

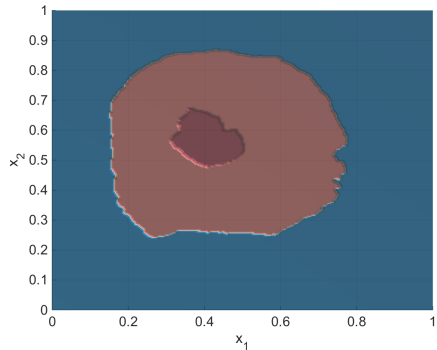
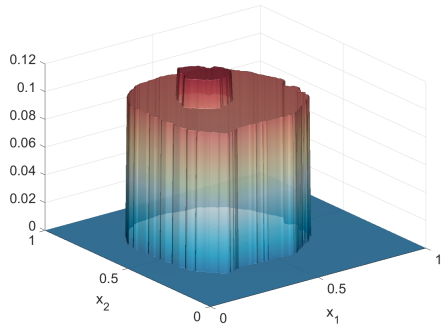


Figure: reconstruction u^δ , $\delta = 0.5$

- 1 Overview
- 2 Approach
 - Convex analysis
 - Moreau–Yosida regularization
 - Semismooth Newton method
- 3 Multi-bang penalty
- 4 **Nonlinear multi-bang**

Control-to-state mapping $S : u \mapsto y$ **nonlinear**:

- approach applicable if S

- 1 weak-to-weak continuous
- 2 twice Fréchet-differentiable

- example: $u \mapsto y$ solving $-\Delta y + uy = f$

- existence, optimality conditions

$$\begin{cases} -\bar{p} = S'(\bar{u})^*(S(\bar{u}) - y^\delta) \\ \bar{u} \in \partial \mathcal{G}^*(\bar{p}) \end{cases}$$

- semismooth Newton method (regularity condition technical)

- $S : u \mapsto y$ solving

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

- approach applicable, but \mathcal{F} nonconvex

- numerical example: $\Omega = [0, 1]^2$, $f \equiv 1$

- $u^\dagger(x) = u_1 + (u_2 - u_1) \chi_{\{x: (x_1 - 0.45)^2 + (x_2 - 0.55)^2 < 0.1\}}(x)$
 $+ (u_3 - u_2 - u_1) \chi_{\{x: (x_1 - 0.4)^2 + (x_2 - 0.6)^2 < 0.02\}}(x)$

- $y^\delta = S(u^\dagger) + \xi$

- $\alpha = 3 \cdot 10^{-5}$, $\gamma \rightarrow 10^{-12}$

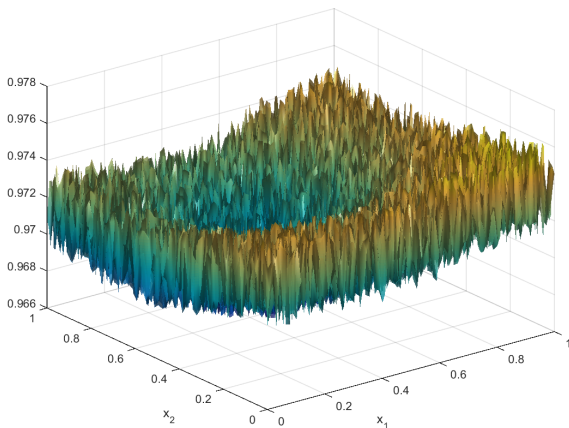


Figure: noisy data y^δ

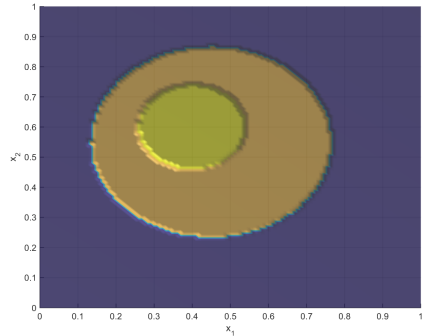
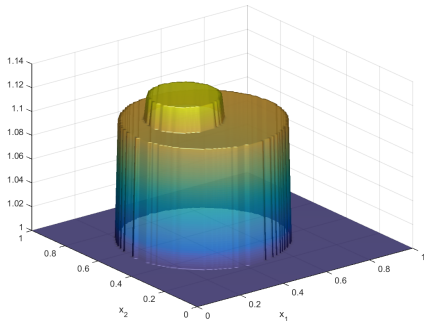


Figure: exact parameter u^\dagger

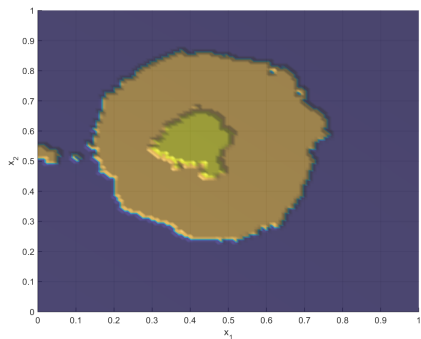
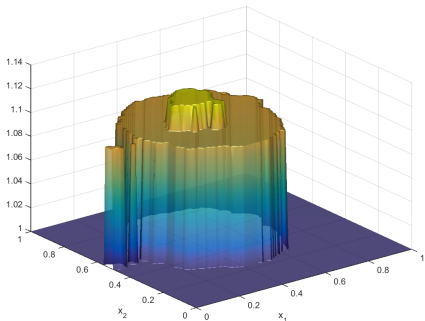


Figure: reconstruction u^δ

Goal: application to EIT

- $S : u \mapsto y$ solving

$$-\nabla \cdot (u \nabla y) = f$$

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

- 1 lack of existence of minimizer ($\bar{y} \neq S(\bar{u})$, cf. homogenization)
- 2 lack of convergence $\gamma \rightarrow 0$
- 3 lack of Newton differentiability of H_γ (no norm gap)

- **remedies:** higher regularity of y or u by

- 1 local smoothing: consider $-\nabla \cdot \left(\int_{B_\epsilon(x)} u(s) ds \nabla y \right)$
- 2 **TV regularization:** add $\|Du\|_{\mathcal{M}} \rightsquigarrow u \in BV(\Omega) \cap L^\infty(\Omega) \hookrightarrow_c L^p(\Omega)$

Difficulty:

- existence requires box constraints \rightsquigarrow use penalty

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

- **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 $u \in L^1(\Omega)$

$$[\Phi(u)](x) = \text{proj}_{[u_1, u_d]}^\varepsilon(u(x)) \quad \text{a.e. } x \in \Omega$$

$$\begin{cases} \min_{u \in BV(\Omega)} \frac{1}{2} \|y - z\|_{L^2(\Omega)}^2 + \alpha G(u) + \beta TV(u) \\ \text{s.t.} & -\nabla \cdot (\Phi(u) \nabla y) = f \text{ in } \Omega \\ & y = 0 \text{ on } \partial\Omega \end{cases}$$

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

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

- $F'(\Phi(\bar{u})) = (\nabla\bar{y} \cdot \nabla\bar{p}) \in L^r(\Omega)$ (optimal state, adjoint)
- $\bar{q} \in L^r(\Omega), r > 1 \rightsquigarrow$ pointwise **multi-bang**
- $\bar{\xi} \in L^r(\Omega), r > 1 \rightsquigarrow$ characterization via *full trace* [Bredies/Holler '12]
- \rightsquigarrow **pointwise optimality conditions**
- **semi-smooth Newton** (after discretization, regularization)

Convex relaxation of **discrete** regularization:

- **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:

- regularization properties, parameter choice
- nonlinear inverse problems: **EIT**
- combination with **BV regularization** (numerics)
- other hybrid discrete–continuous problems

Preprint, MATLAB/Python codes:

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