Convex relaxation of hybrid discrete–continuous penalties

Christian Clason[1] Karl Kunisch[2]

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

Inverse Problems – from Theory to Application
Bristol, August 28, 2014

Motivation: discrete optimization

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 lower-semicontinuous
- not coercive $\rightsquigarrow$ no regularization

Binary penalties

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

- $\mathcal{F}(u)$ 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

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

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 parameters (e.g., tissue types)
- $\beta > 0$ 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
- $\beta > 0$ large penalizes free arc $u_1u_2 \neq 0$
- $\alpha > 0$ penalizes magnitude of active u_i

Convex analysis approach

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

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

Primal-dual optimality system

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

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

Convex analysis approach

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

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

Primal-dual optimality system

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

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

Convex analysis approach

UNIVERSITÄT DUISBURG ESSEN Open-Minded

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

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

Primal-dual optimality(?) system

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

- $\mathcal{G}^*$ 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

UNIVERSITÄT DUISBURG ESSEN
Open-Minded

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

$\rightsquigarrow$ reformulation: consider for $\gamma > 0$

Proximal mapping

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

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

Numerical solution

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

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

Approach:

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

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

Formulation

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

- $u_1 < \cdots < u_d$ given parameter values $(d > 2)$
- $y^\delta \in Y$ noisy data, Y Hilbert space
- $K : Y \to Y^* \equiv Y$ linear, bounded
- assumption: $K^* : Y \to V \hookrightarrow L^2(\Omega)$
- e.g., $K = A^{-1}$ for elliptic operator $A : V \to V^*$

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 $\bar{v}$,

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

Fenchel conjugate

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

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

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

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

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

Fenchel conjugate

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

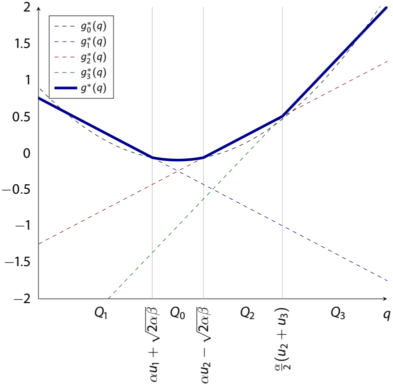
$$g^*(q) = \begin{cases} qu_i - \frac{\alpha}{2}u_i^2 & q \in \overline{Q}_i, \quad 1 \leq i \leq d \\ \frac{1}{2\alpha}q^2 - \beta & q \in \overline{Q}_0 \end{cases}$$

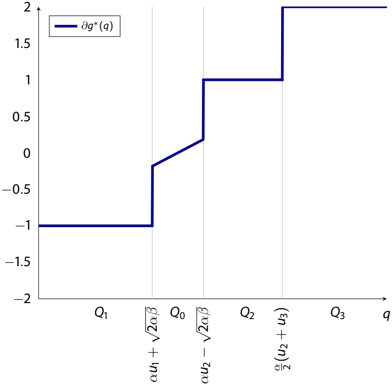
continuous, piecewise differentiable:

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

(no explicit dependence on β!)

Fenchel conjugate: sketch





Optimality system

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

- 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

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

- 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: $Q_0 = \emptyset$, free arc

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

- singular arc corresponds to set-valued subdifferential:

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

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

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

(Sub)optimality

- 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

Moreau–Yosida regularization

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

$$\begin{aligned} Q_i^\gamma &= \Big\{ q : |q - (\alpha+\gamma)u_i| < \sqrt{2\alpha\beta} \wedge \\ &\qquad \tfrac{\alpha}{2}\left(u_{i-1} + \left(1 + \tfrac{2\gamma}{\alpha}\right) u_i\right) < q < \tfrac{\alpha}{2}\left(\left(1 + \tfrac{2\gamma}{\alpha}\right) u_i + u_{i+1}\right) \Big\} \\ Q_0^\gamma &= \Big\{ q : |q - (\alpha+\gamma)u_j| > \sqrt{2\alpha\beta} \;\wedge\; (\alpha+\gamma)u_1 < q < (\alpha+\gamma)u_d \Big\} \\ Q_{i,i+1}^\gamma &= \Big\{ q : \tfrac{\alpha}{2}\left(\left(1 + \tfrac{2\gamma}{\alpha}\right) u_i + u_{i+1}\right) \le q \le \tfrac{\alpha}{2}\left(u_i + \left(1 + \tfrac{2\gamma}{\alpha}\right) u_{i+1}\right) \Big\} \\ Q_{i0}^\gamma &= \Big\{ q : \sqrt{2\alpha\beta} \le q - (\alpha+\gamma)u_i \le \left(1 + \tfrac{\gamma}{\alpha}\right)\sqrt{2\alpha\beta} \Big\} \end{aligned}$$

Regularized optimality system

UNIVERSITÄT DUISBURG ESSEN Open-Minded

$$\begin{cases} p_\gamma = K^*(y^\delta - Ku_\gamma) \\ u_\gamma = \partial\mathcal{G}_\gamma^*(p_\gamma) \end{cases}$$

- $\partial\mathcal{G}_\gamma^*$ maximal monotone $\rightsquigarrow$ unique solution (u_γ, p_γ)
- $(u_\gamma, p_\gamma) \rightharpoonup (\bar{u}, \bar{p})$ as $\gamma \to 0$
- $\partial\mathcal{G}_\gamma^*$ Lipschitz continuous, piecewise C^1, norm gap $V \hookrightarrow L^2(\Omega)$
- $\rightsquigarrow$ semismooth Newton method, continuation in $\gamma \to 0$
- globalization: backtracking line search based on residual norm
- only number of sets Q_i depends on d $\rightsquigarrow$ linear complexity

Numerical examples

- $\Omega = [0,1]^2, \quad K = (-\Delta)^{-1}$

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

- $d = 3, \quad u_1 = 0, \quad u_2 = 0.1, \quad u_3 \in \{0.2, 0.12\}$

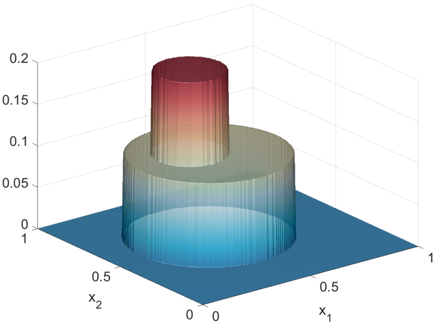
- $y^\delta \in \mathcal{N}\left(Ku^\dagger, \delta \|Ku^\dagger\|_\infty\right)$

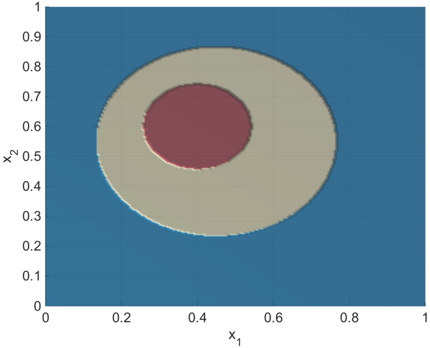
- finite element discretization: uniform grid, 256×256 nodes

- $\alpha = 5 \cdot 10^{-5}$, $\beta = 10^{-1}$ (no free arc)

- terminate at $\gamma < 10^{-12}$

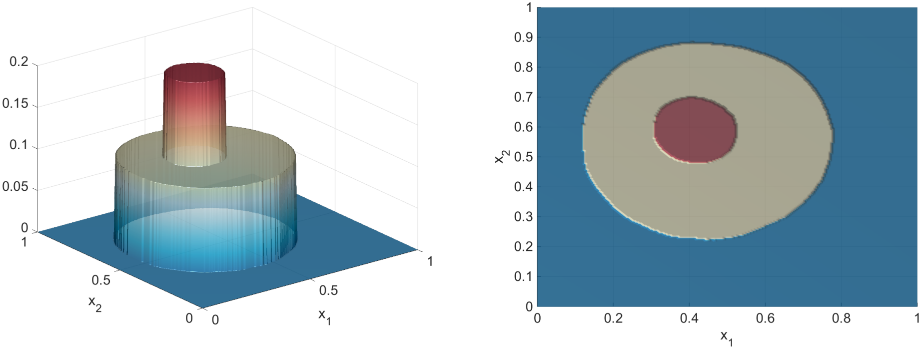
Numerical example: $u_3 = 0.2$





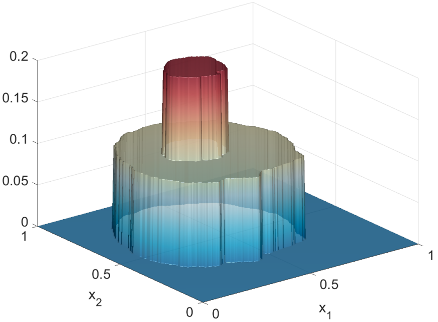
exact parameter $u^{\dagger}$

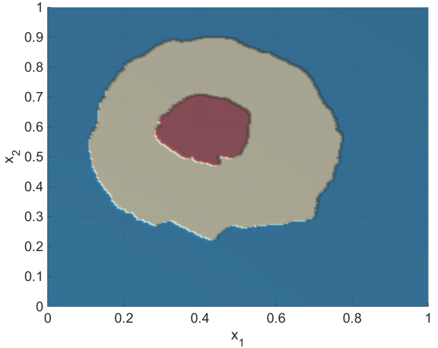
Numerical example: $u_3 = 0.2$



reconstruction u^δ, $\quad \delta = 0.1$

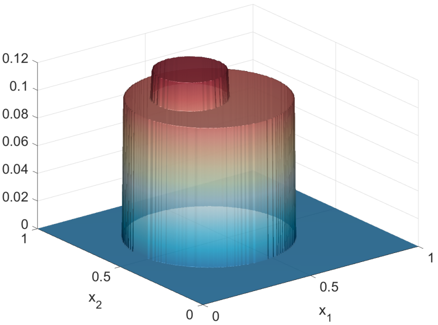
Numerical example: $u_3 = 0.2$

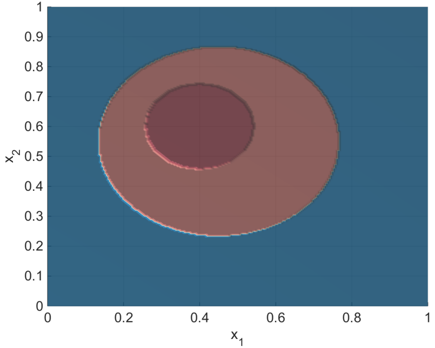




reconstruction u^δ, $\delta = 0.5$

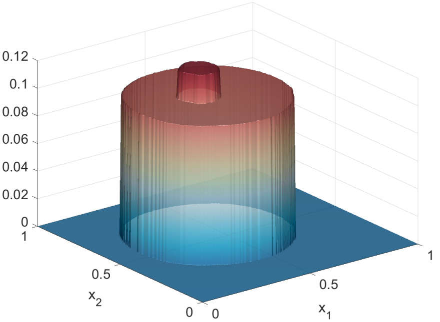
Numerical example: $u_3 = 0.12$

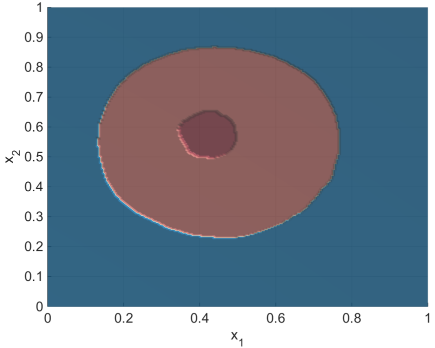




exact parameter $u^\dagger$

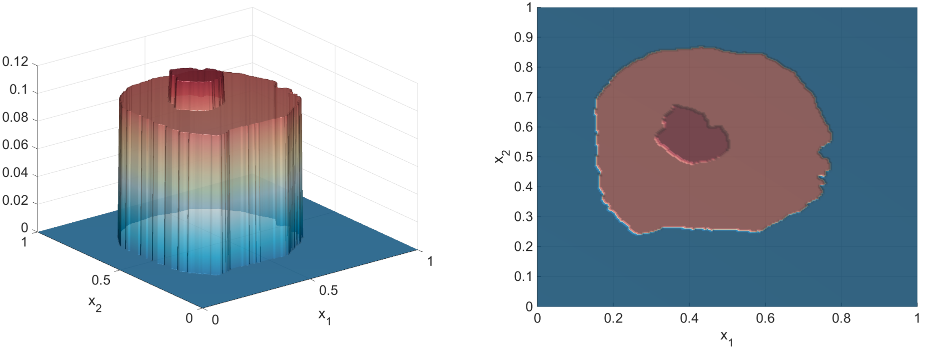
Numerical example: $u_3 = 0.12$





reconstruction u^δ, $\quad \delta = 0.1$

Numerical example: $u_3 = 0.12$



reconstruction u^δ, $\delta = 0.5$

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:

- regularization properties, parameter choice
- nonlinear inverse problems
- other hybrid discrete–continuous problems

Preprint, MATLAB/Python codes:

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