Convex regularization of discrete-valued inverse problems

Christian Clason

Faculty of Mathematics, Universität Duisburg-Essen

joint work with Thi Bich Tram Do, Florian Kruse, Karl Kunisch

Joint Meeting of the KMS and DMV
Seoul, Korea, October 4, 2018

Discrete-valued inverse problem

$$\min_{u \in U} \frac{1}{2}\|F(u) - y^\delta\|_Y^2 + \frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2$$

- $F : L^2(\Omega) \to Y$ forward mapping, $y^\delta \in Y$ noisy data
- a priori information: $u^\dagger \in U$ discrete set

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

- $u_1, \dots, u_d$ given voltages, velocities, materials, ...
  (assumed here: ranking by magnitude possible!)
- goal: include discrete a priori information in regularization

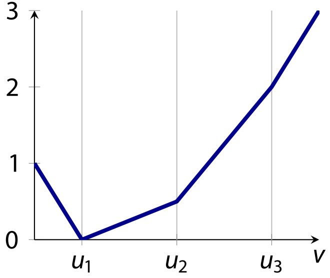
Multi-bang penalty

Approach:

- promote $u(x) \in \{u_1, \ldots, u_d\}$ by convex pointwise penalty

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

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



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

Multi-bang regularization

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

- $K : L^2(\Omega) \to Y$ (linear) 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 $(d > 2)$
- $\mathcal{G}$ multi-bang penalty

Multi-bang penalty

Penalty: pointwise

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

Subdifferential

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

Multi-bang regularization

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

- $\mathcal{G}$ multi-bang 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. $\delta \to 0, \alpha \to 0, \delta\alpha^{-2} \to 0$ implies $u_\alpha^\delta \rightharpoonup u^\dagger$

(standard arguments, e.g. [Burger/Osher 04, Ito/Jin 14])

Multi-bang regularization

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

- standard source condition: $p^\dagger := K^* w \in \partial\mathcal{G}(u^\dagger)$ for $w \in Y$,

  1. a priori choice $\alpha(\delta) \sim \delta$
  2. a posteriori choice $\|Ku^\delta_{\alpha(\delta)} - y^\delta\|_Y \leqslant \tau\delta, \quad \tau > 1$

  $\rightsquigarrow$ convergence rate

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

  in Bregman distance

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

Multi-bang regularization

Pointwise definition of Bregman distance, $\partial \mathcal{G}$:

- $u^\dagger(x) = u_i$ 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

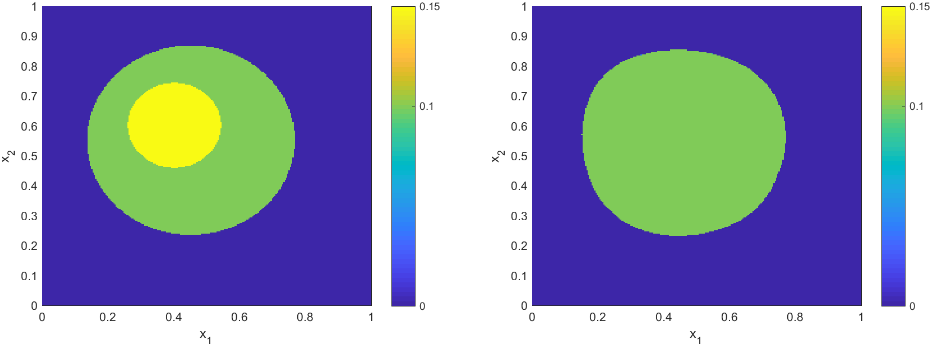
$$D_g^{p^\dagger(x)}(u(x), u^\dagger(x)) = 0 \qquad \text{for any } u(x) \in [u_i, u_{i+1}]$$

- $\rightsquigarrow$ $u^\delta_{\alpha(\delta)} \to u^\dagger$ pointwise a.e. iff $u^\dagger(x) \in \{u_1, \dots, u_d\}$ a.e.
- (convergence not uniform $\rightsquigarrow$ no pointwise rates)

Example: linear inverse problem

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

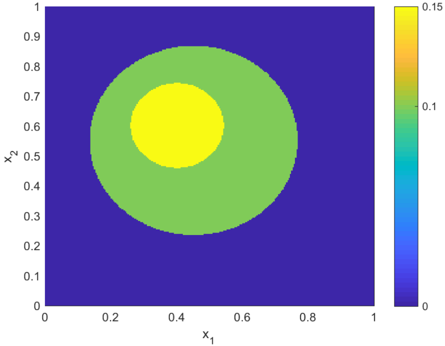
Numerical example: $u_3 = 0.15$



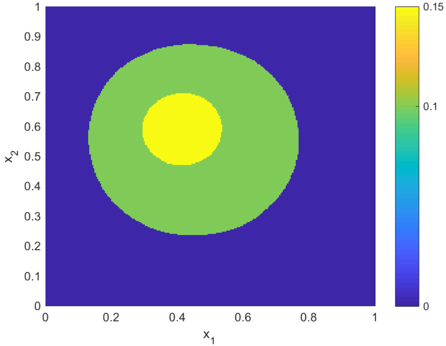
(a) $u^\dagger$

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

Numerical example: $u_3 = 0.15$

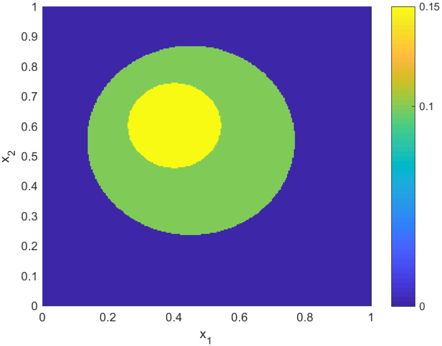


(c) $u^\dagger$

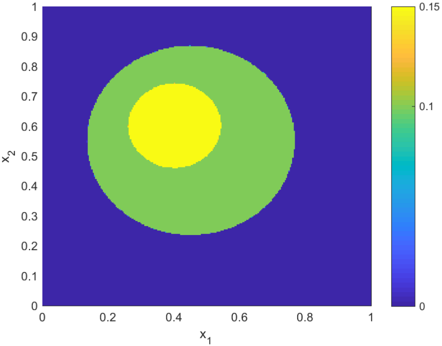


(d) u_α^δ, $\delta \approx 2.37 \cdot 10^{-2}$

Numerical example: $u_3 = 0.15$

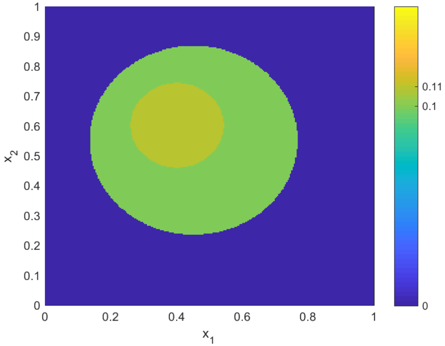


(e) $u^\dagger$

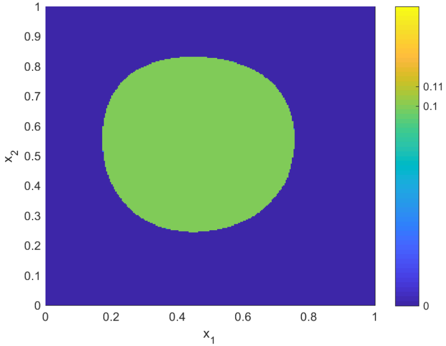


(f) u_α^δ, $\delta \approx 3.69 \cdot 10^{-4}$

Numerical example: $u_3 = 0.11$

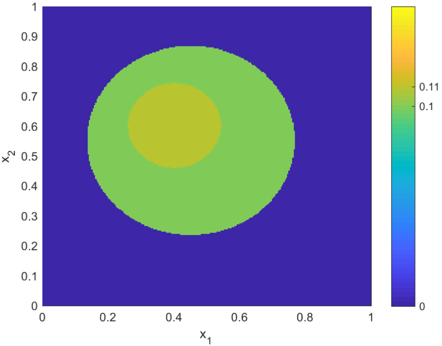


(a) $u^\dagger$

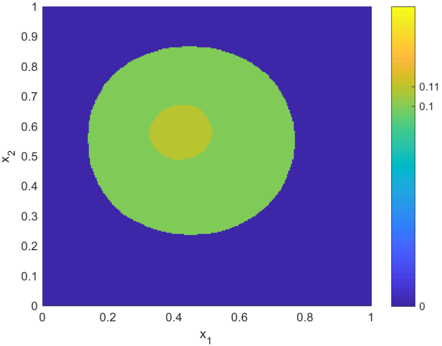


(b) u_α^δ, $\delta \approx 1.68 \cdot 10^{-1}$

Numerical example: $u_3 = 0.11$

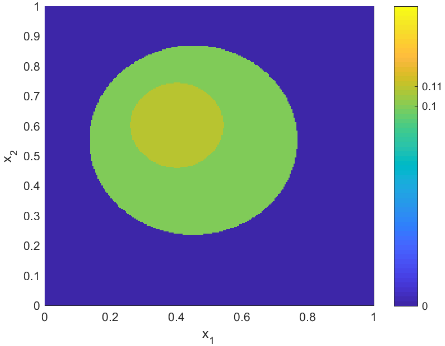


(c) $u^\dagger$

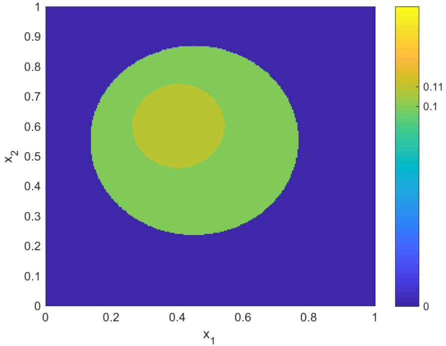


(d) u_α^δ, $\delta \approx 2.17 \cdot 10^{-2}$

Numerical example: $u_3 = 0.11$

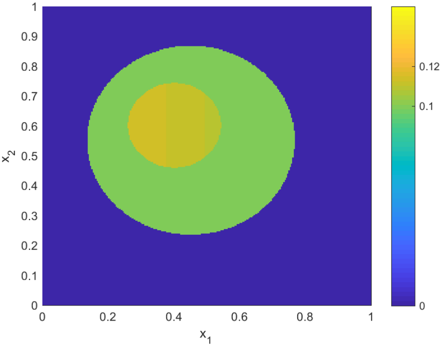


(e) $u^\dagger$

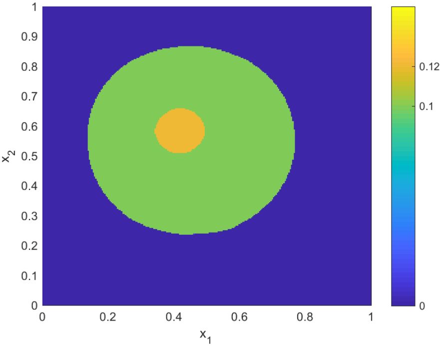


(f) u_α^δ, $\delta \approx 3.29 \cdot 10^{-4}$

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

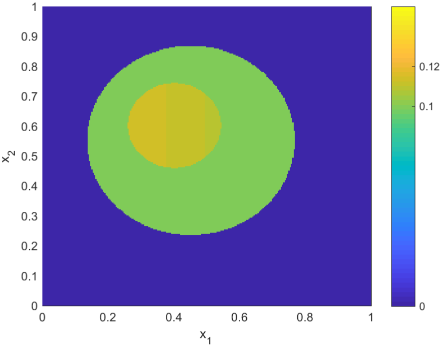


(a) $u^\dagger$

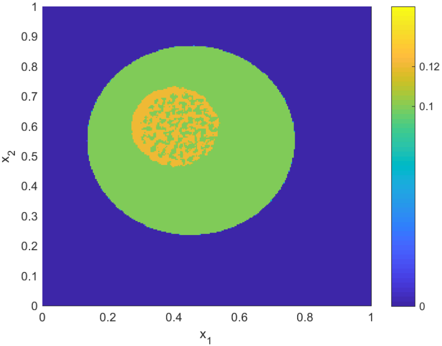


(b) u_α^δ, $\delta \approx 2.11 \cdot 10^{-2}$

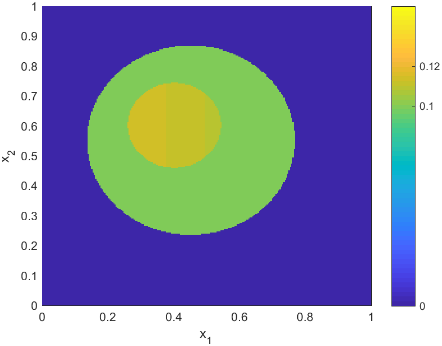
Numerical example: $u_3(x) = 0.12(1 - x_1)$



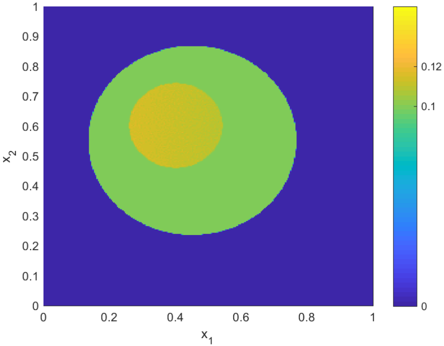
(c) $u^\dagger$



(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 EIT

- $F : u \mapsto y$ solving

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

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

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

- remedies: higher regularity of y or u by

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

TV regularization

Difficulty:

- existence requires box constraints $\rightsquigarrow$ use penalty

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

(here: G multi-bang penalty with 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^∞
- $\rightsquigarrow$ no explicit characterization of minimizers
- $\rightsquigarrow$ replace box constraints by ($C^{1,1}$) projection of $u \in L^1(\Omega)$

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

TV regularization: existence

$$\begin{cases} \min\limits_{u \in BV(\Omega)} \frac{1}{2}\|y - z\|^2_{L^2(\Omega)} + \alpha\, G(u) + \beta\, TV(u) \\ \text{s.t.} \quad -\nabla \cdot (\Phi_\varepsilon(u) \nabla y) = f \text{ in } \Omega \\ \qquad\qquad y = 0 \text{ on } \partial\Omega \end{cases}$$

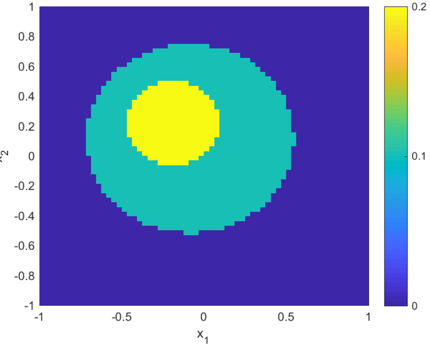
- existence of optimal $\bar{u} \in BV(\Omega) \cap L^\infty(\Omega)$ for $\varepsilon \geqslant 0$
- tracking term Fréchet differentiable in $\Phi_\varepsilon(u) \in L^\infty$ for $\varepsilon > 0$
- regularity of state, adjoint $\rightsquigarrow$ derivative in $L^r(\Omega)$, $r > 1$ (instead of $L^\infty(\Omega)^*$)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

TV regularization: optimality conditions

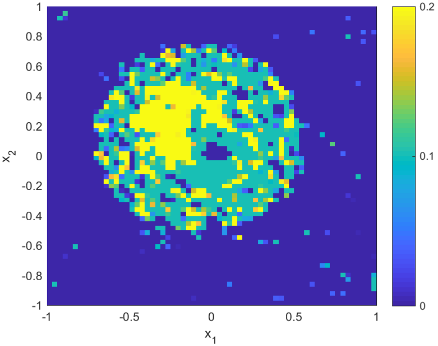
$$\begin{cases} 0 = F'(\Phi(\bar{u}))\Phi'_\varepsilon(\bar{u}) + \alpha\bar{q} + \beta\bar{\xi} \\ \bar{u} \in \partial G^*(\bar{q}) \\ \bar{\xi} \in \partial TV(\bar{u}) \end{cases}$$

- $F'(\Phi_\varepsilon(\bar{u})) = (\nabla\bar{y} \cdot \nabla\bar{p}) \in L^r(\Omega)$ (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)

Numerical example: total variation

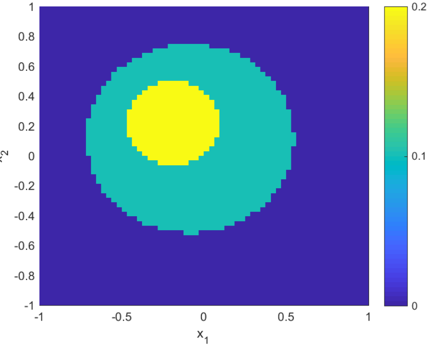


(a) $u^\dagger$

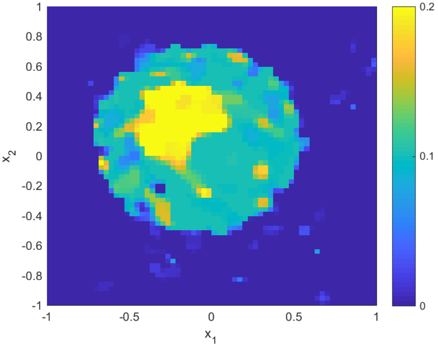


(b) $\alpha = 5 \cdot 10^{-4}, \beta = 0$

Numerical example: total variation

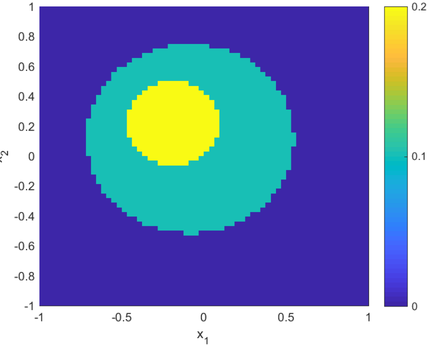


(c) $u^\dagger$

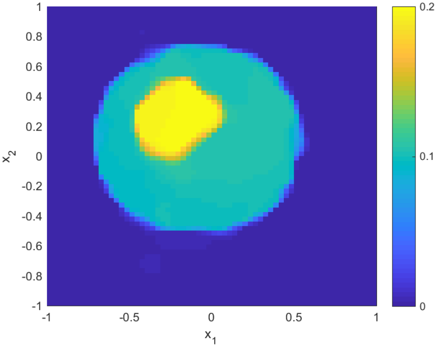


(d) $\alpha = 5 \cdot 10^{-4}, \beta = 10^{-6}$

Numerical example: total variation



(e) $u^\dagger$



(f) $\alpha = 5 \cdot 10^{-4}, \beta = 10^{-5}$

Conclusion

Convex relaxation of discrete regularization:

- well-posed regularization method
- pointwise convergence under general assumptions
- strong structural regularization
- efficient numerical solution (superlinear convergence)

Outlook:

- (heuristic) parameter choice
- nonlinear inverse problems: EIT
- vector-valued multibang
- other hybrid discrete–continuous problems

Preprint, MATLAB/Python codes:

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