Convex relaxation of (some) hybrid discrete-valued optimization problems

Christian Clason

Department of Mathematics and Scientific Computing, University of Graz

joint work with Florian Kruse, Karl Kunisch, Carla Tameling, Benedikt Wirth

4^{th} Austrian Calculus of Variations Day

University of Innsbruck, November 8, 2024

Motivation: hybrid discrete optimization

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

- $\mathcal{F}$ tracking, discrepancy term (involving PDEs)
- 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

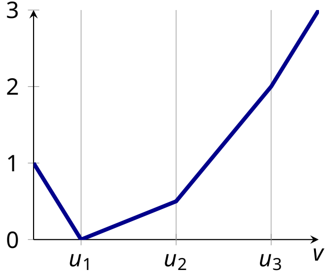
- convex relaxation: replace U by convex hull $u(x) \in [u_1, u_d]$
- works only for $d = 2$, cf. bang-bang control ($\alpha = 0$)
- $\rightsquigarrow$ promote $u(x) \in \{u_1, \dots, u_d\}$ by convex pointwise penalty

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

- 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 $\frac{1}{2}u^2 + \delta_U$
- multi-bang (generalized bang-bang) control
- $\rightsquigarrow$ non-smooth optimization in function spaces

Convex analysis: motivation

$f : \mathbb{R} \to \mathbb{R}$ differentiable:

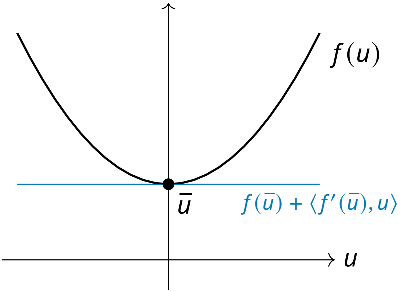
- derivative:

$$f'(u) = \lim_{h\to 0} \frac{f(u+h) - f(u)}{h}$$

- geometrically:
  $f'(u)$ tangent slope

- $f(\bar{u}) = \min_u f(u) \Rightarrow f'(\bar{u}) = 0$

- calculus for f'



Convex relaxation: motivation

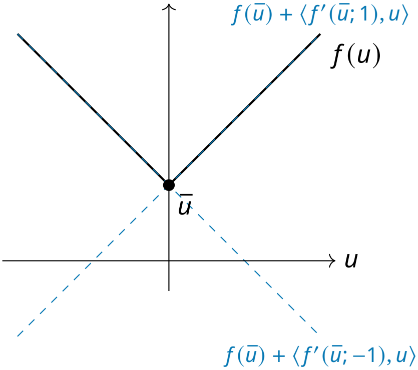
$f : \mathbb{R} \to \mathbb{R}$ not differentiable, convex:

- directional derivative:

$$f'(u;h) = \lim_{t\to 0^+} \frac{f(u+th) - f(u)}{t}$$

- but: for all h,

$$f'(\bar{u};h) \neq 0$$



Convex relaxation: motivation

$f : \mathbb{R} \to \mathbb{R}$ not differentiable, convex:

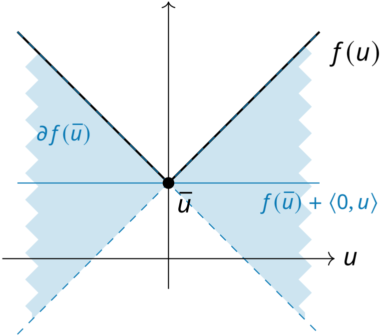
- subdifferential:

  $$\partial f(u) = \{u^* : \langle u^*, h\rangle \leq f'(u; h)\}$$

- geometrically: $\partial f(u)$ set of tangent slopes

- $f(\bar{u}) = \min_u f(u) \Rightarrow 0 \in \partial f(\bar{u})$

- calculus for ∂f under regularity conditions



Fenchel duality

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

- subdifferential

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

- Fenchel conjugate (always convex)

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

- "convex inverse function theorem" (if $\mathcal{F}$ lower semicontinuous)

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

Fenchel duality: example

- $\mathcal{G} : V \to \mathbb{R}, \quad v \mapsto \|v\|_V$:

$$\mathcal{G}^* : V^* \to \overline{\mathbb{R}}, \quad v^* \mapsto \delta_{\{\|\cdot\|_{V^*} \leq 1\}}(v^*) := \begin{cases} 0 & \text{if } \|v^*\|_{V^*} \leq 1 \\ \infty & \text{else} \end{cases}$$

- $\mathcal{G} : V \to \overline{\mathbb{R}}, \quad v \mapsto \delta_{\{\|\cdot\|_V \leq 1\}}(v)$:

$$\partial\mathcal{G}(\bar{v}) = \left\{v^* \in V^* : \langle v^*, v - \bar{v}\rangle_{V^*,V} \leq 0 \quad \text{for all} \quad \|v\|_V \leq 1\right\}$$

$\rightsquigarrow$ box-constrained optimization

Fenchel duality: application

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

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:

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

Regularization

$\mathcal{G}$ 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

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

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

Regularization

Proximal mapping

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

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

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

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

Regularization

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

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

- $\partial\mathcal{G}^*_\gamma = \partial\left(\mathcal{G} + \frac{\gamma}{2}\|\cdot\|^2\right)^* \to \partial\mathcal{G}^*$ as $\gamma \to 0$ (no smoothing of $\mathcal{G}$!)
- single-valued, Lipschitz continuous, explicit
  $\rightsquigarrow$ nonsmooth operator equation, Newton method

Regularization: example

$$\mathcal{G}^* : V^* \to \overline{\mathbb{R}}, \quad p \mapsto \delta_{\{\|\cdot\|_{V^*} \leq 1\}}(p):$$

- Proximal mapping:

$$\operatorname{prox}_{\gamma \mathcal{G}^*}(p) = \operatorname{proj}_{\{\|\cdot\|_{V^*} \leq 1\}}(p)$$

- Moreau–Yosida regularization ($V^* = L^\infty(\Omega)$):

$$\partial \mathcal{G}_\gamma^*(p) = \frac{1}{\gamma}\big(\max(0, p-1)) + \min(0, p+1)\big)$$

(max, min pointwise almost everywhere)

Generalized Newton method

Consider Banach spaces X, Y, mapping $F : X \longrightarrow Y$

Newton-type method for $F(x) = 0$

- choose $x^0 \in X$ (close to solution x^*)
- for $k = 0, 1, \dots$
  1. choose $M_k \in \mathcal{L}(X, Y)$ invertible
  2. solve for s^k:
  $$M_k s^k = -F(x^k)$$
  3. set $x^{k+1} = x^k + s^k$

Convergence of Newton method

Set $d^k = x^k - x^* \rightsquigarrow$

$$\frac{\|x^{k+1} - x^*\|_X}{\|x^k - x^*\|_X} = \frac{\|M_k^{-1}(F(x^* + d^k) - F(x^*) - M_k d^k)\|_X}{\|d^k\|_X}$$

$\rightsquigarrow$ superlinear convergence if

1. regularity condition

$$\|M_k^{-1}\|_{\mathcal{L}(Y,X)} \leq C \quad \text{for all } k$$

2. approximation condition

$$\lim_{\|d^k\|_X \to 0} \frac{\|F(x^* + d^k) - F(x^*) - M_k d^k\|_Y}{\|d^k\|_X} = 0$$

Semismooth Newton method

Goal: define Newton derivative $M_k =: D_N F(x^k)$ such that

$$x^{k+1} = x^k - D_N F(x^k)^{-1} F(x^k)$$

converges superlinearly for $F(x) = 0$ nonsmooth

- $\mathbb{R}^n$: F Lipschitz $\rightsquigarrow$ $D_N F$ from Clarke subdifferential (Rademacher)
  [Mifflin 1977, Kummer 1992, Qi/Sun 1993]
- function space: Clarke subdifferential not explicit
  $\rightsquigarrow$ define $D_N F$ via approximation condition
  [Chen/Nashed/Qi 2000, Hintermüller/Ito/Kunisch 2002]
- $f : \mathbb{R}^N \to \mathbb{R}$ semismooth $\rightsquigarrow$ superposition operator
  $F : L^p(\Omega) \to L^q(\Omega)$ semismooth for $p > q$
  [Ulbrich 2002/03/11, Schiela 2008]

Semismooth Newton method

f locally Lipschitz, piecewise C^1:

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

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

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

converges locally superlinearly

Semismooth Newton method

f locally Lipschitz, piecewise C^1:

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

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

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

converges locally superlinearly if $r > s$

Semismooth functions: example

- $f : \mathbb{R} \to \mathbb{R}, \quad t \mapsto \max(0, t)$

$$D_N f(t) \in \partial_C f(t) = \begin{cases} \{0\} & t < 0 \\ \{1\} & t > 0 \\ [0,1] & t = 0 \end{cases}$$

- $F : L^p(\Omega) \to L^q(\Omega), \quad u(x) \mapsto \max(0, u(x)), \quad p > q$

$$[D_N F(u)h](x) = \begin{cases} 0 & u(x) < 0 \\ h(x) & u(x) \geq 0 \end{cases}$$

$\rightsquigarrow$ Moreau–Yosida regularization semismooth

Numerical solution: summary

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

Approach: pointwise

1. compute subdifferential ∂g (or Fenchel conjugate g^*)
2. compute subdifferential ∂g^*
3. compute proximal mapping $\mathrm{prox}_{\gamma g^*}$
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 $\quad [\partial \mathcal{G}^*_\gamma(p)](x) = \partial g^*_\gamma(p(x))$

Formulation

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

- $u_1 < \cdots < u_d$ given parameter values $(d > 2)$
- $z \in L^2(\Omega)$ target (or noisy data)
- $A : V \to V^*$ isomorphism for Hilbert space $V \hookrightarrow L^2(\Omega) \hookrightarrow V^*$

  (e.g., elliptic differential operator with boundary conditions)
- $\rightsquigarrow \mathcal{F}(u) = \frac{1}{2}\|A^{-1}u - z\|_{L^2}^2$ smooth
- $\mathcal{G}$ multi-bang penalty (will include control constraints from now)

Multi-bang 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

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

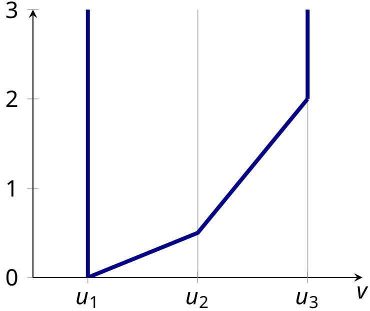
Multi-bang penalty

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

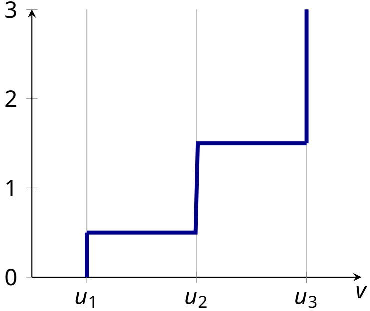
convex inverse function theorem:

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

Multi-bang penalty: sketch

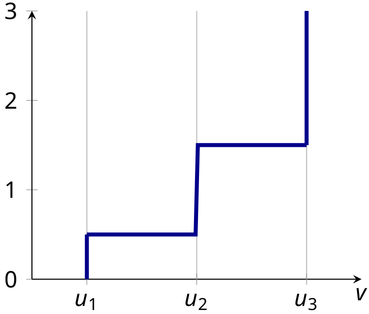


(a) g $(u_1 = 0, u_2 = 1, u_3 = 2)$

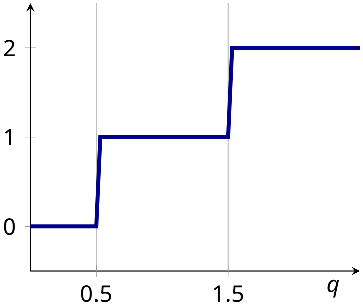


(b) ∂g $(u_1 = 0, u_2 = 1, u_3 = 2)$

Multi-bang penalty: sketch



(c) ∂g $(u_1 = 0, u_2 = 1, u_3 = 2)$



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

Optimality system

$$\bar p = \tfrac{1}{\alpha} S^*(z - S\bar u)$$

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

- $S : u \mapsto y$ control-to-state mapping, S^* adjoint
- $\rightsquigarrow$ unique solution $(\bar u, \bar p) \in L^2(\Omega) \times L^2(\Omega)$
- singular arc $\mathcal{S} = \{x : \bar u(x) \notin \{u_i\}\} \subset \{x : \bar p(x) = \tfrac{1}{2}(u_i + u_{i+1})\}$
- for suitable A, $\bar p(x)$ constant implies $[A^*\bar p](x) = [z - \bar y](x) = 0$

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

Moreau–Yosida regularization

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

case-wise inspection of subdifferential:

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

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

Regularized optimality system

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

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

Regularized optimality system

$$\begin{cases} A^* p_\gamma = \frac{1}{\alpha}(z - y_\gamma) \\ A y_\gamma = \mathcal{G}_\gamma^*(p_\gamma) \end{cases}$$

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

Semismooth Newton method

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

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

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

Numerical example

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

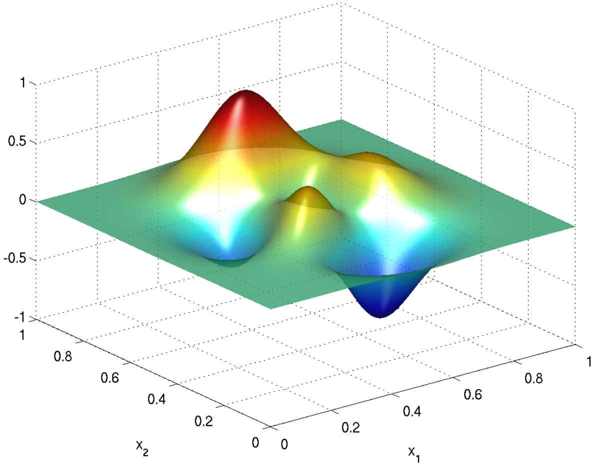
- finite element discretization: uniform grid, 256×256 nodes
- state, adjoint: piecewise linear
- parameter: eliminated (variational discretization)

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

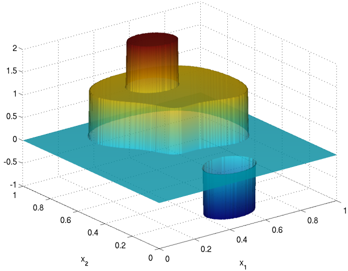
- $\gamma = 0$: regularized active sets empty, true multi-bang

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

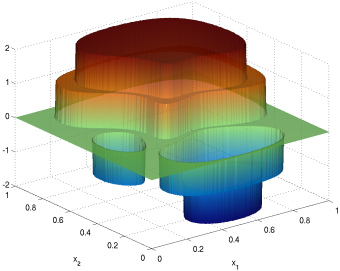
Numerical examples: desired state



Multi-bang controls

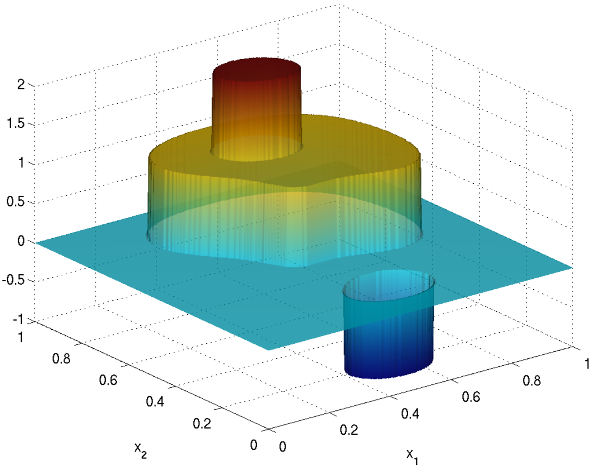


(a) $\alpha = 5 \cdot 10^{-3}$ $(\gamma = 0)$



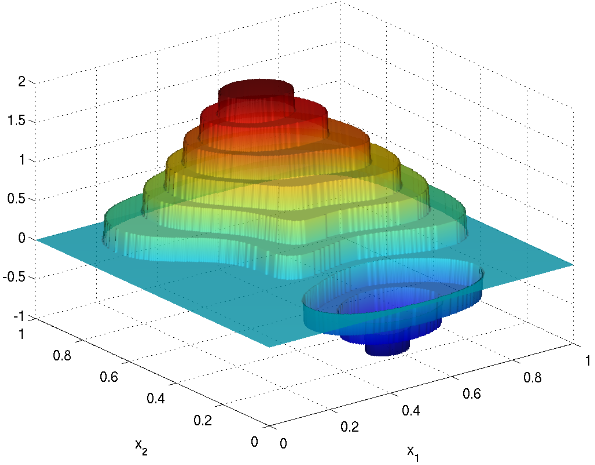
(b) $\alpha = 10^{-3}$ $(\gamma \approx 10^{-7})$

Parameters: effect of d



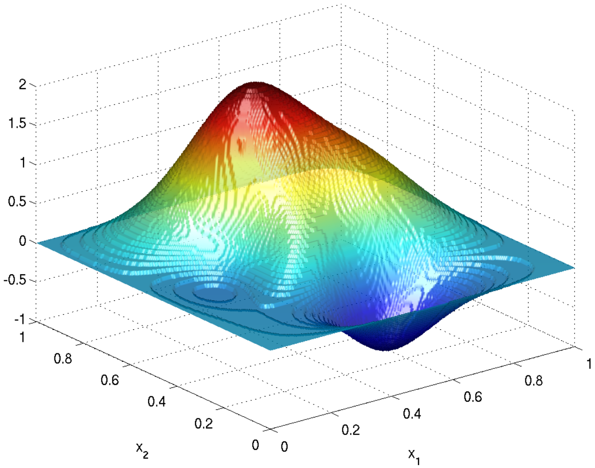
(a) $d = 5$ $(\gamma = 0)$

Parameters: effect of d



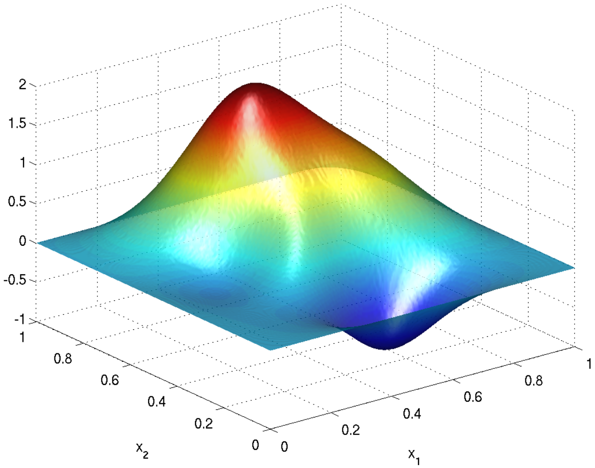
(b) $d = 15$ ($\gamma = 0$)

Parameters: effect of d



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

Parameters: effect of d



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

Vector-valued multi-bang control

Discrete vector-valued controls $u : \Omega \to U \subset \mathbb{R}^m$

Example: optimal control of Bloch equation: $\Omega = [0, T]$, $m = 2$

$$\frac{d}{dt}M(t) = M(t) \times B(t), \qquad M(0) = M_0$$

- $M(t) \in \mathbb{R}^3$ describes temporal evolution of spin ensemble
- $B(t) = (u_1(t), u_2(t), \omega)^T$ controlled time-dependent magnetic field
- ω resonance frequency (material parameter)
- applications in magnetic resonance imaging, spectroscopy
- control-to-state mapping $S : u \to M$ bilinear ($\rightsquigarrow$ chain rule, Clarke)

Vector-valued multi-bang: penalty

Here: admissible control set U of d radially distributed states, origin

$$U = \left\{ \binom{0}{0}, \begin{pmatrix} \omega_0 \cos\theta_1 \\ \omega_0 \sin\theta_1 \end{pmatrix}, \dots, \begin{pmatrix} \omega_0 \cos\theta_d \\ \omega_0 \sin\theta_d \end{pmatrix} \right\}$$

- fixed amplitude $\omega_0 > 0$
- phases $0 \leq \theta_1 < \dots < \theta_d < 2\pi$

multi-bang penalty $g = \left(\frac{1}{2}|\cdot|_2^2 + \delta_U\right)^{**}$ convex envelope

$$g^*(q) = \left(\left(\tfrac{1}{2}|\cdot|_2^2 + \delta_U\right)^{**}\right)^*(q) = \left(\tfrac{1}{2}|\cdot|_2^2 + \delta_U\right)^*(q)$$

$$= \begin{cases} 0 & \langle q, u_i\rangle \leq \frac{1}{2}\omega_0^2 \text{ for all } 1 \leq i \leq d \\ \langle q, u_i\rangle - \frac{1}{2}\omega_0^2 & \frac{\theta_{i-1}+\theta_i}{2} \leq \angle q \leq \frac{\theta_i+\theta_{i+1}}{2}, \langle q, u_i\rangle \geq \frac{1}{2}\omega_0^2 \end{cases}$$

Vector-valued multi-bang: subdifferential

Fenchel conjugate

$$g^*(q) = \begin{cases} 0 =: u_0 & q \in \overline{Q}_0 \\ \langle q, u_i \rangle - \frac{1}{2}\omega_0^2 & q \in \overline{Q}_i \end{cases}$$

Subdifferential

$$\partial g^*(q) = \begin{cases} \{u_i\} & q \in Q_i \qquad 0 \leq i \leq d \\ \operatorname{co}\{u_{i_1}, \dots, u_{i_k}\} & q \in Q_{i_1 \dots i_k} \quad 0 \leq i_1, \dots, i_k \leq d \end{cases}$$

Vector-valued multi-bang: subdifferential

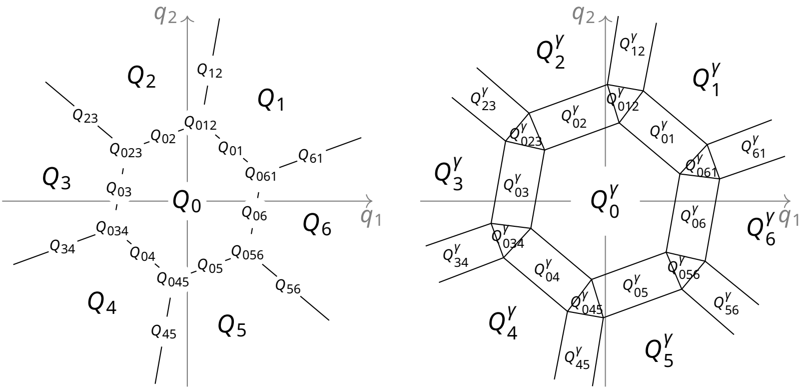
Subdifferential

$$\partial g^*(q) = \begin{cases} \{u_i\} & q \in Q_i \qquad 0 \le i \le d \\ \mathrm{co}\,\{u_{i_1}, \dots, u_{i_k}\} & q \in Q_{i_1 \dots i_k} \quad 0 \le i_1, \dots, i_k \le d \end{cases}$$

Moreau–Yosida regularization

$$(\partial g^*)_\gamma(q) = \begin{cases} u_i & q \in Q_i^\gamma \\ \left(\frac{\langle q, u_i\rangle}{\gamma\omega_0^2} - \frac{\alpha}{2\gamma}\right) u_i & q \in Q_{0,i}^\gamma \\ \frac{u_i + u_{i+1}}{2} + \frac{\langle q, u_i - u_{i+1}\rangle (u_i - u_{i+1})}{\gamma |u_i - u_{i+1}|_2^2} & q \in Q_{i,i+1}^\gamma \\ \frac{q}{\gamma} - \frac{\alpha}{\gamma}\left(\frac{\omega_0}{|u_i + u_{i+1}|_2}\right)^2 (u_i + u_{i+1}) & q \in Q_{0,i,i+1}^\gamma . \end{cases}$$

Vector-valued multi-bang: subdifferential



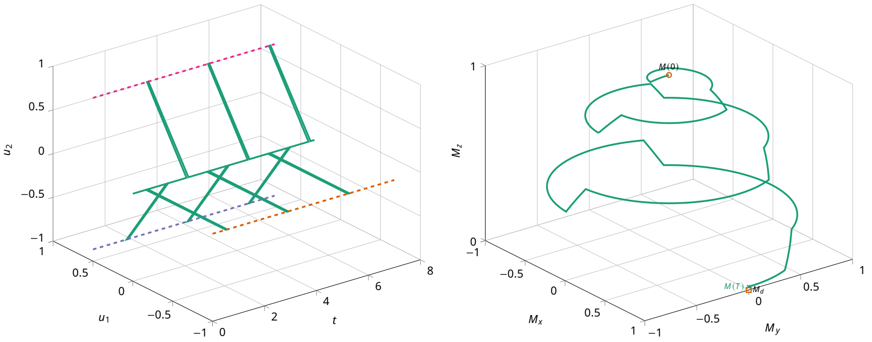
(a) subdomains for ∂g^*

(b) subdomains for $(\partial g^*)_\gamma$

Vector-valued multi-bang: examples

- goal: shift magnetization from $M_0 = (0, 0, 1)^T$ at $t = 0$
  to $M_d = (1, 0, 0)^T$ at $t = T$
- $d = 3, 6$ radially distributed admissible control states
- $n = 1, 4$ isochromats with different resonance frequencies
  1. shift all isochromats
  2. shift only one isochromat
- $\alpha = 10^{-1}$, $\omega_0 = 1$
- example motivated by [Dridi/Lapert/Salomon/Glaser/Sugny '15]
- matrix-free Krylov method for semismooth Newton step
- discretization, adjoint from [Aigner/Clason/Rund/Stollberger '16]

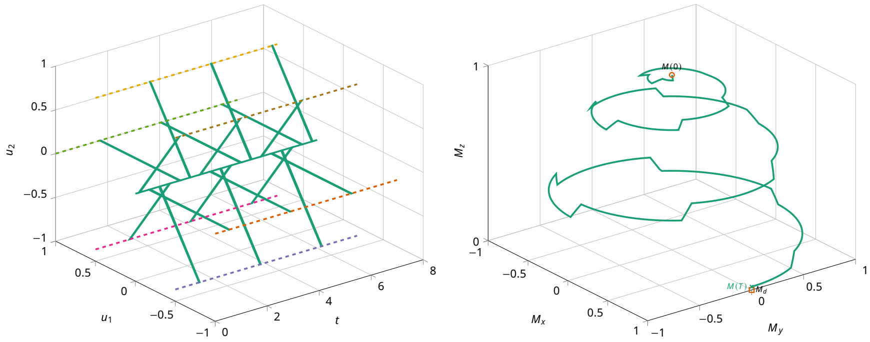
Vector-valued multi-bang: examples



(a) control $u(t)$ (b) state $M(t)$

Figure: $n = 1$ isochromat, $d = 3$ control states

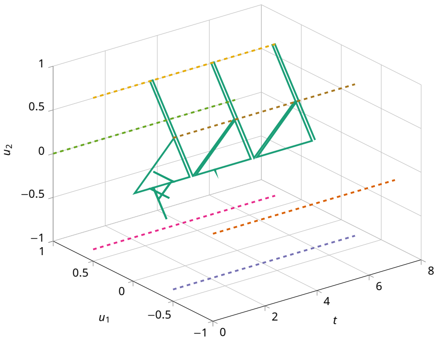
Vector-valued multi-bang: examples



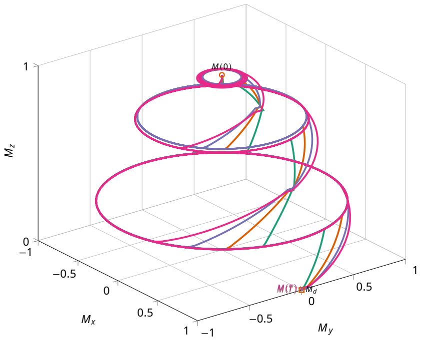
(a) control $u(t)$ (b) state $M(t)$

Figure: $n = 1$ isochromat, $d = 6$ control states

Vector-valued multi-bang: examples



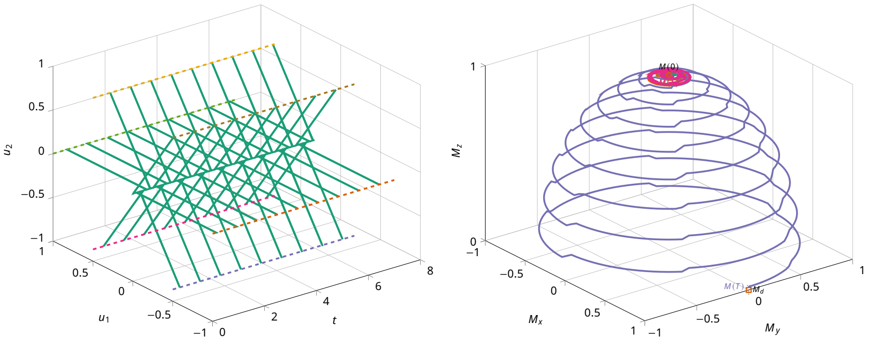
(a) control $u(t)$



(b) state $M(t)$

Figure: $n = 4$ isochromats, same target

Vector-valued multi-bang: examples



(a) control $u(t)$ (b) state $M(t)$

Figure: $J = 4$ isochromats, different targets

Vector-valued multi-bang: elasticity

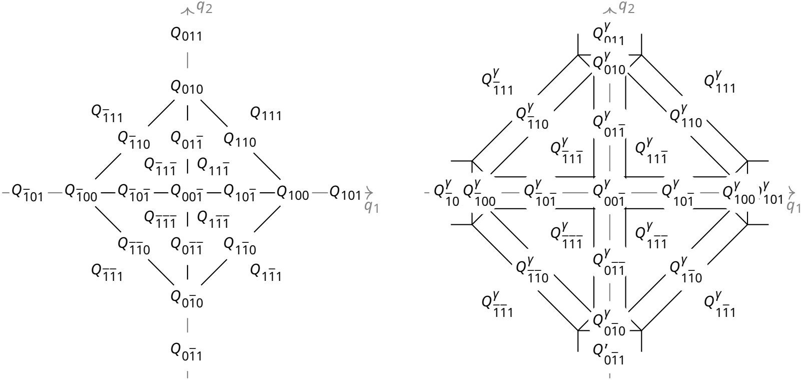
Linear elasticity: $S : u \mapsto y$ solving

$$\begin{cases} -2\mu \operatorname{div} \epsilon(y) - \lambda \operatorname{grad} \operatorname{div} y = u & \text{in } \Omega, \\ y = 0 & \text{on } \Gamma, \\ (2\mu\epsilon(y) + \lambda \operatorname{div} y)n = 0 & \text{on } \partial\Omega \setminus \Gamma \end{cases}$$

Concentric admissible set (without origin!)

$$U = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ -2 \end{pmatrix} \right\}$$

Vector-valued multi-bang: elasticity



(a) subdomains for ∂g^*

(b) subdomains for $(\partial g^*)_\gamma$

Vector-valued multi-bang: elasticity

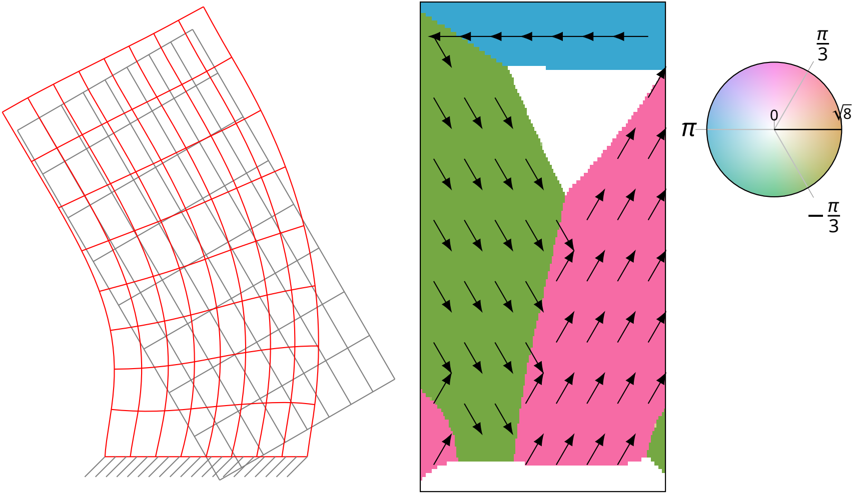


Figure: radial, $d = 3$, $\alpha = 10^{-3}$ (grey: prescribed, red: achieved deformation)

Vector-valued multi-bang: elasticity

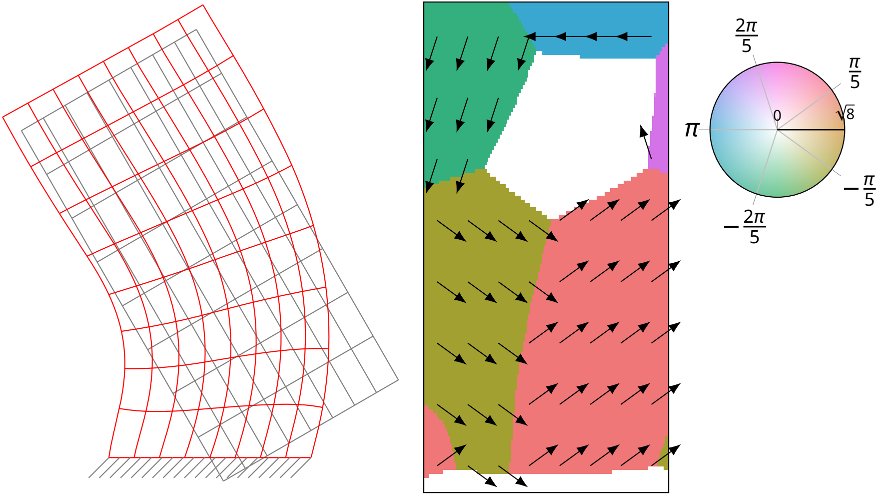


Figure: radial, $d = 5$, $\alpha = 10^{-3}$ (grey: prescribed, red: achieved deformation)

Vector-valued multi-bang: elasticity

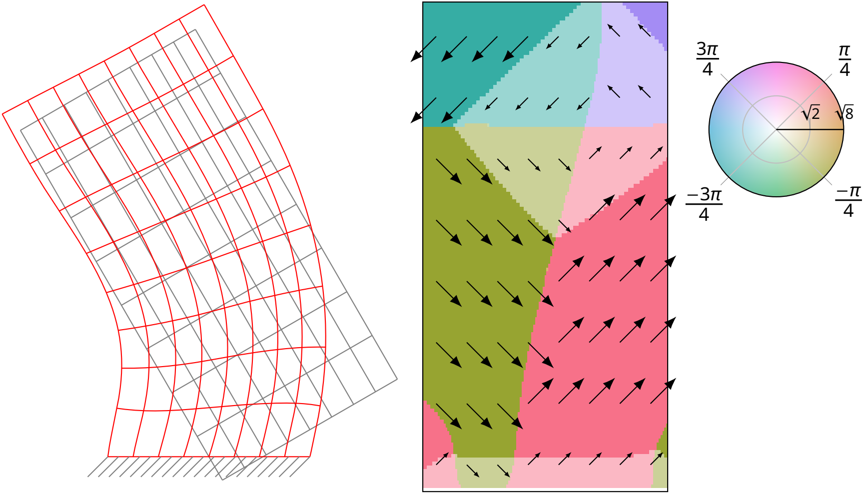


Figure: concentric, $\alpha = 10^{-3}$ (grey: prescribed, red: achieved deformation)

Vector-valued multi-bang: elasticity

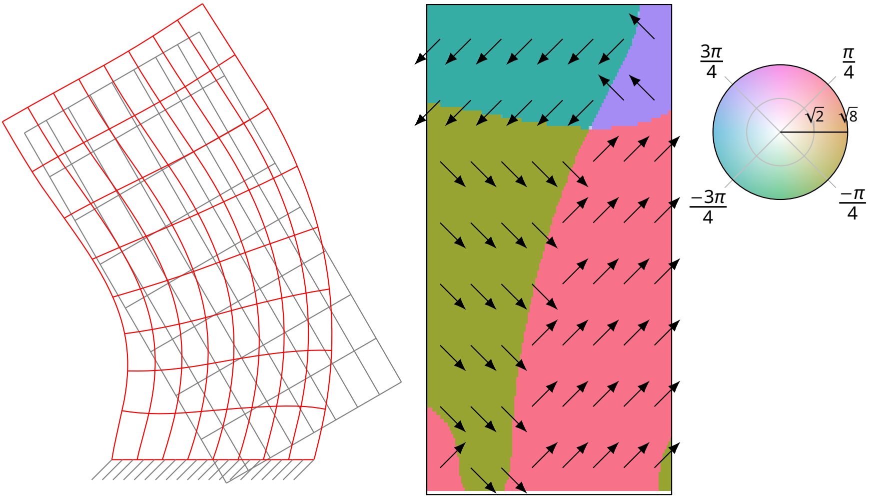
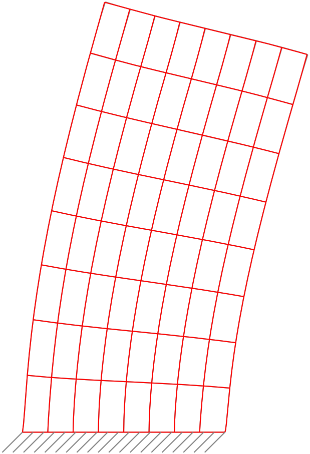


Figure: concentric, $\alpha = 10^{-5}$ (grey: prescribed, red: achieved deformation)

Vector-valued multi-bang: elasticity



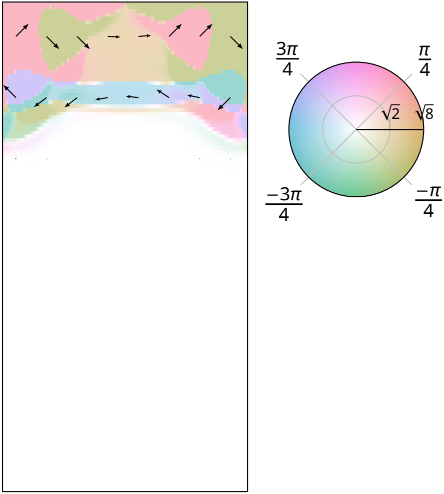


Figure: concentric, $\alpha = 10^{-5}$ (grey: prescribed, red: achieved deformation)

Vector-valued multi-bang: elasticity

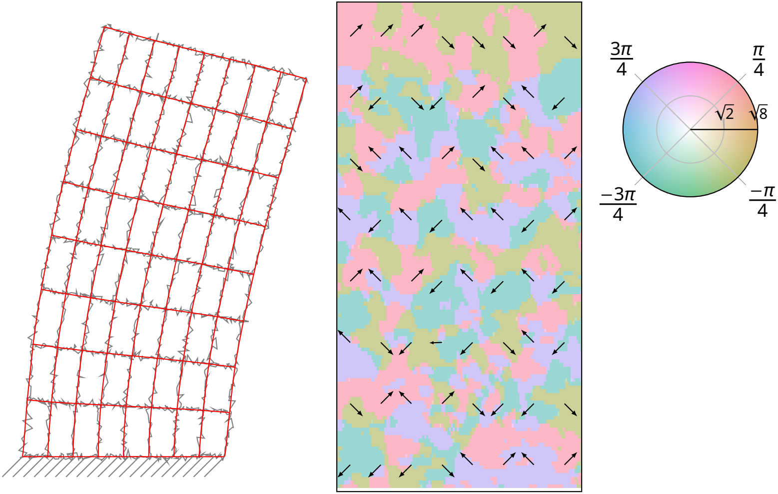


Figure: concentric, $\alpha = 10^{-5}$ (grey: prescribed, red: achieved deformation)

Vector-valued multi-bang: multimaterial transport

Multimaterial transport on graph (V, E): S graph divergence

$$Su(x) = \sum_{e \in E \text{ to } x} u(e) - \sum_{e \in E \text{ from } x} u(e)$$

$\rightsquigarrow$ tracking term penalizes material loss

$$U = \left\{ u \in \mathbb{R}^m \,\middle|\, u_i \in \{0, m_i\} \text{ or } u_i \in \{0, -m_i\} \text{ for } i = 1, \dots, m \right\}$$

$\rightsquigarrow$ all transport in same direction (same sign)

Here: multibang penalty

$$g(u) = |u|_2 + \delta_U$$

Vector-valued multi-bang: multimaterial transport

Here: multibang penalty

$$g(u) = |u|_2 + \delta_U$$

$\rightsquigarrow$ algorithmic computation of proximal mapping:

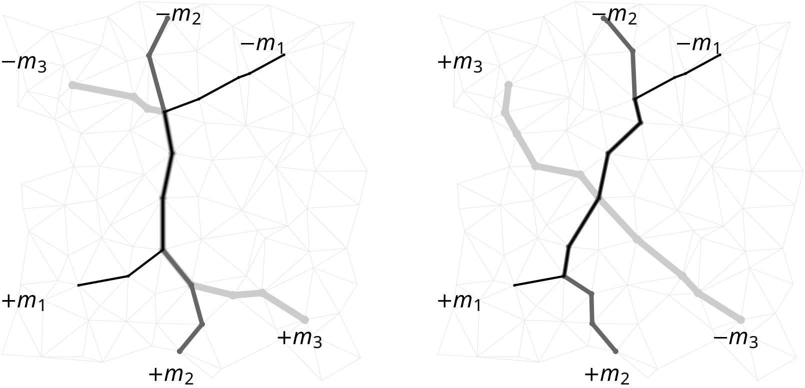
1. enumerate all faces of

$$\operatorname{epi} g^* = \{(q,t) \in \mathbb{R}^{m+1} : t \geq \langle \bar{u}_i, q\rangle - \alpha c(\bar{u}_i) \text{ for } i \in I\}$$

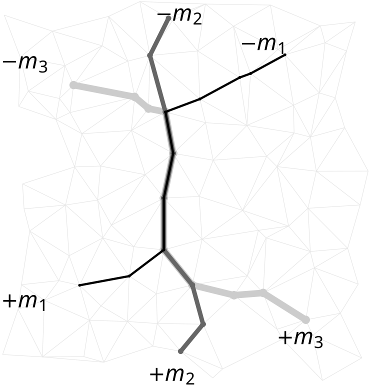
   (linear program, precompute!)

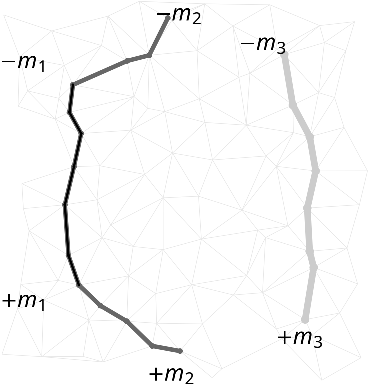
2. for each face, precompute linear system for $(\mathrm{Id} + \gamma\partial g^*)^{-1}$
3. for given q, evaluate linear system for each face and pick correct face (inequalities)

Vector-valued multi-bang: multimaterial transport

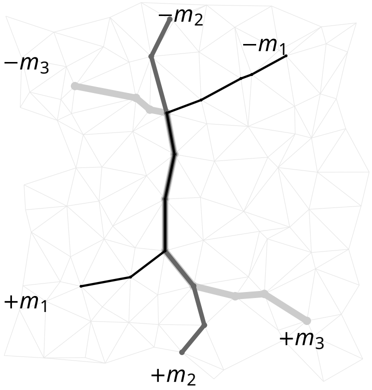


Vector-valued multi-bang: multimaterial transport





Vector-valued multi-bang: multimaterial transport



Conclusion

Discrete controls:

- can be promoted by convex penalty
- linear complexity in number of parameter values
- $\rightsquigarrow$ efficient numerical solution (superlinear convergence)
- applicable to vector-valued problems

Outlook:

- nonlinear inverse problems: seismic imaging
- combination with total variation regularization
- primal-dual proximal splitting
- other discrete–continuous problems: switching, state diagrams

Papers, Code: `https://imsc.uni-grat.at/clason/publications`
SIREV paper: `https://arxiv.org/abs/2108.10077`
Textbook: `https://arxiv.org/abs/1708.04180`