On Nonlinear Optimal Control Problems with an L^1 Norm

Eduardo Casas   Roland Herzog   Gerd Wachsmuth

University of Cantabria   Numerical Mathematics   CHEMNITZ UNIVERSITY OF TECHNOLOGY

Workshop on Inverse Problems and Optimal Control for PDEs

Warwick, May 23–27, 2011

Overview

1. Introduction and Problem Setting
2. 1st- and 2nd-Order Optimality Conditions
3. Finite Element Error Estimates and Examples
4. Extension: Directional Sparsity
(joint with Georg Stadler, ICES, Texas)

Problem Setting for this Talk

Control problem

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)} \\ \text{such that} \quad & u_a \le u \le u_b \qquad (u_a < 0 < u_b) \\ \text{and} \quad & y \text{ solves the PDE} \end{aligned}$$

Semilinear partial differential equation

$$\begin{aligned} -\Delta y + a(\cdot, y) &= u \quad \text{in } \Omega \\ y &= 0 \quad \text{on } \Gamma \end{aligned}$$

Why Consider $\|u\|_{L^1(\Omega)}$?

- The L^1-norm
$$\|u\|_{L^1(\Omega)} = \int_\Omega |u(x)| \, \mathrm{d}x$$
is often a natural measure of the true control cost.
- It also has the effect of promoting sparse controls.

[Vossen, Maurer (2006); Stadler (2009); Clason, Kunisch (2011)]

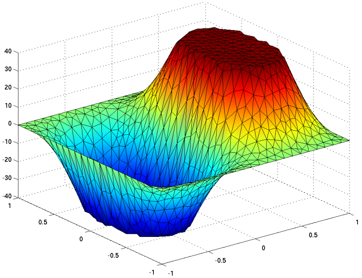
Why Consider $\|u\|_{L^1(\Omega)}$?

- The L^1-norm
$$\|u\|_{L^1(\Omega)} = \int_\Omega |u(x)| \, \mathrm{d}x$$
is often a natural measure of the true control cost.
- It also has the effect of promoting sparse controls.
- Applications in control:
  - actuator placement
  - on/off control structure desired
  - true measure of control cost
- Other applications using the 1-norm:
  - compressed sensing
  - TV-based image restoration

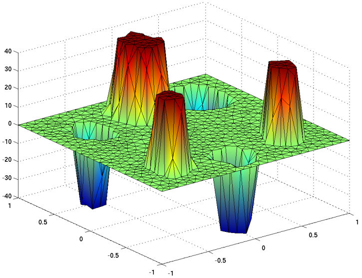
[Vossen, Maurer (2006); Stadler (2009); Clason, Kunisch (2011)]

A First Glance at Sparsity

$\mu = 0$



$\mu > 0$

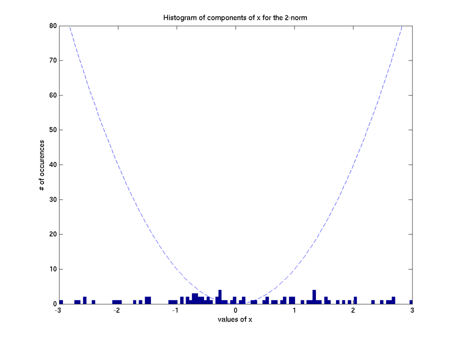


A First Glance at Sparsity

Smooth minimization problem

minimize $\frac{1}{2}\|x\|_2^2$ s.t. $Ax = b$

Histogram (solution components' sizes)



A First Glance at Sparsity

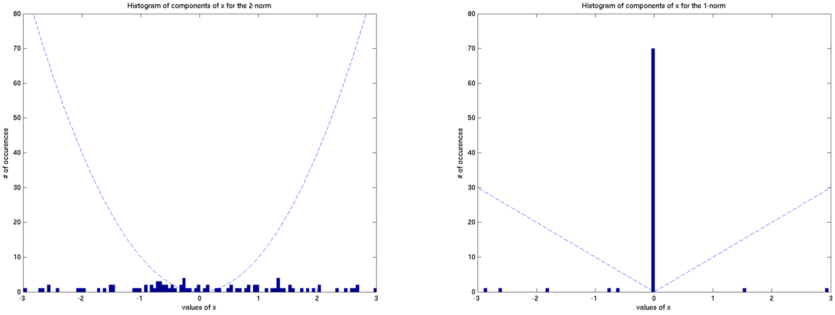
Smooth minimization problem

minimize $\frac{1}{2}\|x\|_2^2$ s.t. $Ax = b$

Convex minimization problem

minimize $\|x\|_1$ s.t. $Ax = b$

Histogram (solution components' sizes)



A First Glance at Sparsity

Smooth minimization problem

$$\text{minimize} \quad \frac{1}{2}\|x\|_2^2 \quad \text{s.t.} \quad Ax = b$$

$$x + A^\top p = 0$$
$$Ax - b = 0$$

Convex minimization problem

$$\text{minimize} \quad \|x\|_1 \quad \text{s.t.} \quad Ax = b$$

$$\lambda + A^\top p = 0, \quad \lambda \in \partial\|x\|_1$$
$$Ax - b = 0$$

$$\begin{aligned} \lambda_i &= -1 && \text{if } x_i < 0 \\ \lambda_i &= +1 && \text{if } x_i > 0 \\ \lambda_i &\in [-1, 1] && \text{if } x_i = 0 \end{aligned}$$

A First Glance at Sparsity

Smooth minimization problem

minimize $\frac{1}{2}\|x\|_2^2$ s.t. $Ax = b$

$$x + A^\top p = 0$$
$$Ax - b = 0$$

Convex minimization problem

minimize $\|x\|_1$ s.t. $Ax = b$

$$\lambda + A^\top p = 0, \quad \lambda \in \partial\|x\|_1$$
$$Ax - b = 0$$

$$\begin{aligned} \lambda_i &= -1 && \text{if } x_i < 0 \\ \lambda_i &= +1 && \text{if } x_i > 0 \\ \lambda_i &\in [-1, 1] && \text{if } x_i = 0 \end{aligned}$$

$$x_i = \max\{0, x_i + c\,(\lambda_i - 1)\} + \min\{0, x_i + c\,(\lambda_i + 1)\}$$

Problem Setting for this Talk

Control problem

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)} \\ \text{such that} \quad & u_a \le u \le u_b \qquad (u_a < 0 < u_b) \\ \text{and} \quad & y \text{ solves the PDE} \end{aligned}$$

Semilinear partial differential equation

$$\begin{aligned} -\Delta y + a(\cdot, y) &= u \quad \text{in } \Omega \\ y &= 0 \quad \text{on } \Gamma \end{aligned}$$

Basic Assumptions Concerning the PDE

Semilinear partial differential equation

$$\begin{aligned} -\Delta y + a(\cdot, y) &= u \quad \text{in } \Omega \\ y &= 0 \quad \text{on } \Gamma \end{aligned}$$

Assumptions

- $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$, with $C^{1,1}$-boundary or convex, polygonal set
- a is Carathéodory-function, monotone, C^2 w.r.t. y

Properties

- For $u \in L^p(\Omega)$, $n/2 < p \le 2$ the solution $y = G(u) \in W^{2,p}(\Omega)$
- $G : L^p(\Omega) \to W^{2,p}(\Omega)$ is C^2, derivatives by linearization

Basic Assumptions Concerning the PDE

Semilinear partial differential equation

$$\begin{aligned} -\operatorname{div}(A\,\nabla y) + a(\cdot, y) &= u \quad \text{in } \Omega \\ y &= 0 \quad \text{on } \Gamma \end{aligned}$$

Assumptions

- $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$, with $C^{1,1}$-boundary or convex, polygonal set
- a is Carathéodory-function, monotone, C^2 w.r.t. y
- $\xi^\top A(x)\,\xi \geq \underline{a}\,\|\xi\|^2$ for all $\xi \in \mathbb{R}^n$, $\underline{a} > 0$

Properties

- For $u \in L^p(\Omega)$, $n/2 < p \leq 2$ the solution $y = G(u) \in W^{2,p}(\Omega)$
- $G : L^p(\Omega) \to W^{2,p}(\Omega)$ is C^2, derivatives by linearization

Overview

1 Introduction and Problem Setting

2 1st- and 2nd-Order Optimality Conditions

3 Finite Element Error Estimates and Examples

4 Extension: Directional Sparsity
(joint with Georg Stadler, ICES, Texas)

Problem Setting

Control problem

$$\text{Minimize} \quad \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu \, \|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

$$\text{and} \quad y \text{ solves the PDE}$$

Semilinear partial differential equation

$$-\operatorname{div}(A\,\nabla y) + a(\cdot, y) = u \quad \text{in } \Omega$$

$$y = 0 \quad \text{on } \Gamma$$

Problem Setting

Control problem

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

Problem Setting

Control problem

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

Properties

- ▢ is differentiable w.r.t. $u \in L^2(\Omega)$

Problem Setting

Control problem

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|_{L^2(\Omega)}^2 + \frac{\nu}{2}\|u\|_{L^2(\Omega)}^2 + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

Properties

- ▢ is differentiable w.r.t. $u \in L^2(\Omega)$

Problem Setting

Control problem

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

Properties

- ▢ is differentiable w.r.t. $u \in L^2(\Omega)$

Problem Setting

Control problem

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

Properties

- (orange box) is differentiable w.r.t. $u \in L^2(\Omega)$
- (blue box) is convex w.r.t. u

Problem Setting

Control problem

$$\text{Minimize} \quad \underbrace{\frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}}_{\text{f(u)}} + \underbrace{\mu\,\|u\|_{L^1(\Omega)}}_{\text{j(u)}}$$

$$\text{such that} \quad u_a \le u \le u_b$$

Properties

- (orange box) is differentiable w.r.t. $u \in L^2(\Omega)$
- (blue box) is convex w.r.t. u

Sums of Differentiable and Convex Function

Definition of a generalized subdifferential

Let f be differentiable and j convex, $J = f + j$. The generalized subdifferential $\partial J(x)$ is defined as

$$\partial J(x) = \nabla f(x) + \partial j(x)$$

- This coincides with known generalized derivatives (e.g. Fréchet, Clarke) on this class of functions.
- This ensures the uniqueness, i.e. ∂J does not depend on the splitting of J into f and j.

Sums of Differentiable and Convex Function

Definition of a generalized subdifferential

Let f be differentiable and j convex, $J = f + j$. The generalized subdifferential $\partial J(x)$ is defined as

$$\partial J(x) = \nabla f(x) + \partial j(x)$$

Necessary optimality condition of first order

$$0 \in \partial J(x) = \nabla f(x) + \partial j(x)$$

First-Order Necessary Condition

$$f(u) = \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}, \qquad j(u) = \|u\|_{L^1(\Omega)}$$

First-Order Necessary Condition

$$f(u) = \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}, \qquad j(u) = \|u\|_{L^1(\Omega)}$$

$$\nabla f(\overline{u}) = \underbrace{G'(\overline{u})^\star(\overline{y} - y_d)}_{\text{adjoint state } \overline{p}} + \nu\,\overline{u}, \quad \text{where} \quad \overline{y} = G(\overline{u})$$

First-Order Necessary Condition

$$f(u) = \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}, \qquad j(u) = \|u\|_{L^1(\Omega)}$$

$$\nabla f(\overline{u}) = \underbrace{G'(\overline{u})^\star(\overline{y} - y_d)}_{\text{adjoint state } \overline{p}} + \nu\,\overline{u}, \quad \text{where} \quad \overline{y} = G(\overline{u})$$

First-order necessary optimality conditions

$$0 \in \nabla f(\overline{u}) + \mu\,\partial j(\overline{u})$$

$$\Leftrightarrow \quad 0 = \nabla f(\overline{u}) + \mu\,\overline{\lambda}, \quad \overline{\lambda} \in \partial j(\overline{u})$$

First-Order Necessary Condition

$$f(u) = \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}, \qquad j(u) = \|u\|_{L^1(\Omega)}$$

$$\nabla f(\overline{u}) = \underbrace{G'(\overline{u})^\star(\overline{y} - y_d)}_{\text{adjoint state } \overline{p}} + \nu\,\overline{u}, \quad \text{where} \quad \overline{y} = G(\overline{u})$$

First-order necessary optimality conditions

$$0 \in \nabla f(\overline{u}) + \mu\,\partial j(\overline{u})$$

$$\Leftrightarrow \quad 0 = \nabla f(\overline{u}) + \mu\,\overline{\lambda}, \quad \overline{\lambda} \in \partial j(\overline{u})$$

... with convex control constraints: $U_{\text{ad}} = \{u \in L^2(\Omega) : u_a \le u \le u_b\}$

$$0 \le \langle \nabla f(\overline{u}) + \mu\,\overline{\lambda},\ u - \overline{u}\rangle_{L^2(\Omega)} \quad \text{for all } u \in U_{\text{ad}}, \quad \overline{\lambda} \in \partial j(\overline{u})$$

First-Order Necessary Condition

$$f(u) = \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)}, \qquad j(u) = \|u\|_{L^1(\Omega)}$$

$$\nabla f(\overline{u}) = \underbrace{G'(\overline{u})^\star(\overline{y} - y_d)}_{\text{adjoint state } \overline{p}} + \nu\,\overline{u}, \quad \text{where} \quad \overline{y} = G(\overline{u})$$

First-order necessary optimality conditions

$$0 \in \nabla f(\overline{u}) + \mu\,\partial j(\overline{u})$$

$$\Leftrightarrow \quad 0 = \nabla f(\overline{u}) + \mu\,\overline{\lambda}, \quad \overline{\lambda} \in \partial j(\overline{u})$$

... with convex control constraints: $U_{\text{ad}} = \{u \in L^2(\Omega) : u_a \le u \le u_b\}$

$$0 \le \langle \nabla f(\overline{u}) + \mu\,\overline{\lambda},\ u - \overline{u}\rangle_{L^2(\Omega)} \quad \text{for all } u \in U_{\text{ad}}, \quad \overline{\lambda} \in \partial j(\overline{u})$$

CHEMNITZ UNIVERSITY OF TECHNOLOGY

First-Order Necessary Condition

Theorem

Let $\overline{u}$ be a local min. with state $\overline{y} = G(\overline{u})$. Then there exist an adjoint state $\overline{p} = G'(\overline{u})^\star(\overline{y} - y_d)$ and a subgradient $\overline{\lambda} \in \partial j(\overline{u}) = \partial \|\overline{u}\|_{L^1(\Omega)}$ s.t.

$$\langle \overline{p} + \nu\, \overline{u} + \mu\, \overline{\lambda},\, u - \overline{u} \rangle_{L^2(\Omega)} \geq 0 \quad \text{for all } u \in U_{\text{ad}}.$$

First-Order Necessary Condition

Theorem

Let $\overline{u}$ be a local min. with state $\overline{y} = G(\overline{u})$. Then there exist an adjoint state $\overline{p} = G'(\overline{u})^\star(\overline{y} - y_d)$ and a subgradient $\overline{\lambda} \in \partial j(\overline{u}) = \partial \|\overline{u}\|_{L^1(\Omega)}$ s.t.

$$\langle \overline{p} + \nu\, \overline{u} + \mu\, \overline{\lambda},\, u - \overline{u} \rangle_{L^2(\Omega)} \geq 0 \quad \text{for all } u \in U_{\text{ad}}.$$

Subgradient of the L^1 norm

$$\overline{\lambda}(x) \begin{cases} = +1 & \text{where } \overline{u}(x) > 0 \\ \in [-1, 1] & \text{where } \overline{u}(x) = 0 \\ = -1 & \text{where } \overline{u}(x) < 0 \end{cases}$$

First-Order Necessary Condition

Theorem

Let $\overline{u}$ be a local min. with state $\overline{y} = G(\overline{u})$. Then there exist an adjoint state $\overline{p} = G'(\overline{u})^\star(\overline{y} - y_d)$ and a subgradient $\overline{\lambda} \in \partial j(\overline{u}) = \partial \|\overline{u}\|_{L^1(\Omega)}$ s.t.

$$\langle \overline{p} + \nu\,\overline{u} + \mu\,\overline{\lambda},\, u - \overline{u} \rangle_{L^2(\Omega)} \geq 0 \quad \text{for all } u \in U_{\text{ad}}.$$

Adjoint equation

$$-\operatorname{div}(A^\top \nabla p) + \frac{\partial a}{\partial y}(\cdot, \overline{y})\,\overline{p} = \overline{y} - y_d \quad \text{in } \Omega$$

$$\overline{p} = 0 \quad \text{on } \Gamma$$

First-Order Necessary Condition

Theorem

Let $\overline{u}$ be a local min. with state $\overline{y} = G(\overline{u})$. Then there exist an adjoint state $\overline{p} = G'(\overline{u})^\star(\overline{y} - y_d)$ and a subgradient $\overline{\lambda} \in \partial j(\overline{u}) = \partial \|\overline{u}\|_{L^1(\Omega)}$ s.t.

$$\langle \overline{p} + \nu\,\overline{u} + \mu\,\overline{\lambda},\, u - \overline{u} \rangle_{L^2(\Omega)} \ge 0 \quad \text{for all } u \in U_{\text{ad}}.$$

Corollary: projection formulas

$$\overline{u}(x) = \operatorname{proj}_{[u_a,u_b]}\Big(-\frac{1}{\nu}\big(\overline{p}(x) + \mu\,\overline{\lambda}(x)\big)\Big)$$

$$\overline{\lambda}(x) = \operatorname{proj}_{[-1,+1]}\Big(-\frac{1}{\mu}\overline{p}(x)\Big)$$

$$\overline{u}(x) = 0 \quad \Longleftrightarrow \quad |\overline{p}(x)| \le \mu$$

First-Order Necessary Condition

Theorem

Let $\overline{u}$ be a local min. with state $\overline{y} = G(\overline{u})$. Then there exist an adjoint state $\overline{p} = G'(\overline{u})^\star(\overline{y} - y_d)$ and a subgradient $\overline{\lambda} \in \partial j(\overline{u}) = \partial \|\overline{u}\|_{L^1(\Omega)}$ s.t.

$$\langle \overline{p} + \nu\, \overline{u} + \mu\, \overline{\lambda},\, u - \overline{u} \rangle_{L^2(\Omega)} \ge 0 \quad \text{for all } u \in U_{\text{ad}}.$$

Corollary: projection formulas

$$\overline{u}(x) = \operatorname{proj}_{[u_a, u_b]} \Big(-\frac{1}{\nu}\big(\overline{p}(x) + \mu\, \overline{\lambda}(x)\big)\Big)$$

$$\overline{\lambda}(x) = \operatorname{proj}_{[-1,+1]} \Big(-\frac{1}{\mu}\overline{p}(x)\Big)$$

$$\overline{u}(x) = 0 \quad \Longleftrightarrow \quad |\overline{p}(x)| \le \mu$$

It follows that $\overline{u}, \overline{\lambda} \in C^{0,1}(\overline{\Omega}) = W^{1,\infty}(\Omega)$.
Moreover, $\overline{\lambda} \in \partial \|\overline{u}\|_{L^1(\Omega)}$ is unique.

Second-Order Optimality Conditions

Critical cone at stationary point $\overline{u}$ with associated $\overline{\lambda} \in \partial j(\overline{u})$

$$\mathcal{C}^+_{\overline{u}} := \left\{ v \in L^2(\Omega) : f'(\overline{u})\, v + \mu \, \langle \overline{\lambda},\, v \rangle = 0 \right\}$$

Second-Order Optimality Conditions

Critical cone at stationary point $\overline{u}$ with associated $\overline{\lambda} \in \partial j(\overline{u})$

$$\mathcal{C}_{\overline{u}}^{+} := \left\{ v \in L^2(\Omega) : f'(\overline{u})\, v + \mu \, \langle \overline{\lambda}, \, v \rangle = 0 \right\}$$

$$\langle f''(\overline{u})\, v, \, v \rangle > 0 \quad \text{for all } v \in \mathcal{C}_{\overline{u}}^{+} \setminus \{0\} \quad \Rightarrow \quad \overline{u} \text{ is locally optimal}$$

CHEMNITZ UNIVERSITY OF TECHNOLOGY

Second-Order Optimality Conditions

Critical cone at stationary point $\overline{u}$ with associated $\overline{\lambda} \in \partial j(\overline{u})$

$$\mathcal{C}_{\overline{u}}^{+} := \left\{ v \in L^2(\Omega) : f'(\overline{u})\, v + \mu \, \langle \overline{\lambda}, v \rangle = 0 \right\}$$

$$\langle f''(\overline{u})\, v,\ v \rangle > 0 \quad \text{for all } v \in \mathcal{C}_{\overline{u}}^{+} \setminus \{0\} \quad \Rightarrow \quad \overline{u} \text{ is locally optimal}$$

$$\langle f''(\overline{u})\, v,\ v \rangle \geq 0 \quad \text{for all } v \in \mathcal{C}_{\overline{u}}^{+} \quad \not\Leftarrow \quad \overline{u} \text{ is locally optimal}$$

Second-Order Optimality Conditions

Critical cone at stationary point $\overline{u}$ with associated $\overline{\lambda} \in \partial j(\overline{u})$

$$\mathcal{C}_{\overline{u}}^{+} := \left\{ v \in L^2(\Omega) : f'(\overline{u})\, v + \mu \, \langle \overline{\lambda},\, v \rangle = 0 \right\}$$ too large

$$\langle f''(\overline{u})\, v,\ v \rangle > 0 \quad \text{for all } v \in \mathcal{C}_{\overline{u}}^{+} \setminus \{0\} \quad \Rightarrow \quad \overline{u} \text{ is locally optimal}$$

$$\langle f''(\overline{u})\, v,\ v \rangle \geq 0 \quad \text{for all } v \in \mathcal{C}_{\overline{u}}^{+} \quad \not\Leftarrow \quad \overline{u} \text{ is locally optimal}$$

Second-Order Optimality Conditions

Critical cone at stationary point $\overline{u}$ with associated $\overline{\lambda} \in \partial j(\overline{u})$

$$\mathcal{C}_{\overline{u}}^{+} := \big\{ v \in L^2(\Omega) : f'(\overline{u})\, v + \mu \, \langle \overline{\lambda}, \, v \rangle = 0 \big\} \qquad \text{too large}$$

$$\mathcal{C}_{\overline{u}} := \big\{ v \in L^2(\Omega) : f'(\overline{u})\, v + \mu \, j'(\overline{u}; v) = 0 \big\} \qquad \text{correct}$$

$$\langle f''(\overline{u})\, v, \, v \rangle > 0 \quad \text{for all } v \in \mathcal{C}_{\overline{u}} \setminus \{0\} \quad \Rightarrow \quad \overline{u} \text{ is locally optimal}$$

$$\langle f''(\overline{u})\, v, \, v \rangle \geq 0 \quad \text{for all } v \in \mathcal{C}_{\overline{u}} \quad \Leftarrow \quad \overline{u} \text{ is locally optimal}$$

CHEMNITZ UNIVERSITY OF TECHNOLOGY

Second-Order Optimality Conditions

Critical cone at stationary point $\overline{u}$ with associated $\overline{\lambda} \in \partial j(\overline{u})$

$$\mathcal{C}_{\overline{u}}^{+} := \{ v \in L^2(\Omega) : f'(\overline{u})\, v + \mu \, \langle \overline{\lambda}, \, v \rangle = 0 \} \qquad \text{too large}$$

$$\mathcal{C}_{\overline{u}} := \{ v \in L^2(\Omega) : f'(\overline{u})\, v + \mu \, j'(\overline{u}; v) = 0 \} \qquad \text{correct}$$

$$\langle f''(\overline{u})\, v, \, v \rangle > 0 \quad \text{for all } v \in \mathcal{C}_{\overline{u}} \setminus \{0\} \quad \Rightarrow \quad \overline{u} \text{ is locally optimal}$$

$$\langle f''(\overline{u})\, v, \, v \rangle \geq 0 \quad \text{for all } v \in \mathcal{C}_{\overline{u}} \quad \Leftarrow \quad \overline{u} \text{ is locally optimal}$$

. . . with control constraints

$$\mathcal{C}_{\overline{u}} := \left\{ v \in L^2(\Omega) : \begin{array}{l} f'(\overline{u})\, v + \mu \, j'(\overline{u}; v) = 0 \\ \left.\begin{array}{l} v \geq 0 \text{ where } \overline{u} = u_a \\ v \leq 0 \text{ where } \overline{u} = u_b \end{array}\right\} \; v \in \mathcal{T}_{U_{ad}}(\overline{u}) \end{array} \right\}$$

CHEMNITZ UNIVERSITY OF TECHNOLOGY

Second-Order Sufficient Conditions

Critical cone (closed, convex)

$$C_{\overline{u}} := \{ v \in \mathcal{T}_{U_{\text{ad}}}(\overline{u}) : f'(\overline{u})\, v + \mu\, j'(\overline{u}; v) = 0 \}$$

Theorem

Let $\overline{u} \in U_{\text{ad}}$ and $\overline{\lambda} \in \partial j(\overline{u})$ satisfy the first order necessary condition. Assume $\langle f''(\overline{u})\, v,\, v\rangle > 0$ holds for all $v \in C_{\overline{u}} \setminus \{0\}$. Then there exist $\delta > 0$, $\varepsilon > 0$ such that

$$J(\overline{u}) + \frac{\delta}{2} \| u - \overline{u} \|^2_{L^2(\Omega)} \le J(u) \quad \text{for all } u \in U_{\text{ad}} \cap B^{L^2}_{\varepsilon}(\overline{u}).$$

Second-Order Sufficient Conditions

Critical cone (closed, convex)

$$C_{\overline{u}} := \{v \in \mathcal{T}_{U_{\text{ad}}}(\overline{u}) : f'(\overline{u})\, v + \mu\, j'(\overline{u}; v) = 0\}$$

Theorem

Let $\overline{u} \in U_{\text{ad}}$ and $\overline{\lambda} \in \partial j(\overline{u})$ satisfy the first order necessary condition. Assume $\langle f''(\overline{u})\, v,\ v\rangle > 0$ holds for all $v \in C_{\overline{u}} \setminus \{0\}$. Then there exist $\delta > 0$, $\varepsilon > 0$ such that

$$J(\overline{u}) + \frac{\delta}{2}\|u - \overline{u}\|^2_{L^2(\Omega)} \le J(u) \quad \text{for all } u \in U_{\text{ad}} \cap B_\varepsilon^{L^2}(\overline{u}).$$

Corollary

There exist $\tau > 0$, $\delta_2 > 0$ such that $\langle f''(\overline{u})\, v,\ v\rangle \ge \delta_2\, \|v\|^2_{L^2(\Omega)}$ for all

$$v \in C_{\overline{u}}^{\tau} = \{v \in \mathcal{T}_{U_{\text{ad}}}(\overline{u}) : f'(\overline{u})\, v + \mu\, j'(\overline{u}; v) \le \tau\, \|v\|_{L^2(\Omega)}\}$$

Overview

1. Introduction and Problem Setting
2. 1st- and 2nd-Order Optimality Conditions
3. Finite Element Error Estimates and Examples
4. Extension: Directional Sparsity (joint with Georg Stadler, ICES, Texas)

Finite Element Approximation

- Regular triangulation $\{\mathcal{T}_h\}$ of Ω, $\Omega_h = \cup_{T\in\mathcal{T}_h} T$.
- Discrete space of (adjoint) states (piecewise linear):

$$Y_h = \{y_h \in C(\overline{\Omega}) : y_{h|T} \in \mathcal{P}_1 \text{ for all } T \in \mathcal{T}_h, \text{ and } y_h = 0 \text{ on } \overline{\Omega} \setminus \Omega_h\}$$

- Discrete PDE:

$$\int_{\Omega_h} \nabla z_h^\top A \nabla y_h + a(\cdot, y_h)\,\mathrm{d}x = \int_{\Omega_h} u\, z_h\,\mathrm{d}x \quad \text{for all } z_h \in Y_h$$

- Discrete space of controls (piecewise constant):

$$U_h = \{u_h \in L^2(\Omega_h) : u_{h|T} \equiv \mathrm{const} \text{ for all } T \in \mathcal{T}_h\}$$

Discrete Problem

Discrete optimization problem

$$\begin{aligned} \text{Minimize} \quad & \frac{1}{2}\|G_h(u_h) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u_h\|^2_{L^2(\Omega)} + \mu\,\|u_h\|_{L^1(\Omega)} \\ \text{such that} \quad & u_a \le u_h \le u_b \\ \text{and} \quad & u_h \in U_h \end{aligned}$$

Convergence of Minimizers

Theorem (approximation of global minima)

For every $h > 0$ let $\overline{u}_h$ be a global solution of the discrete problem. Then the sequence $\{\overline{u}_h\}_{h>0}$ is bounded in $L^\infty(\Omega)$ and there exist subsequences, denoted in the same way, converging to a point $\overline{u}$ in the weak* $L^\infty(\Omega)$ topology. Any of these limit points is a global solution of the continuous problem. Moreover, we have

$$\lim_{h\to 0} \left\{ \|\overline{u} - \overline{u}_h\|_{L^\infty(\Omega_h)} \right\} = 0 \quad \text{and} \quad \lim_{h\to 0} J_h(\overline{u}_h) = J(\overline{u}).$$

CHEMNITZ UNIVERSITY OF TECHNOLOGY

Convergence of Minimizers

Theorem (approximation of global minima)

For every $h > 0$ let $\overline{u}_h$ be a global solution of the discrete problem. Then the sequence $\{\overline{u}_h\}_{h>0}$ is bounded in $L^\infty(\Omega)$ and there exist subsequences, denoted in the same way, converging to a point $\overline{u}$ in the weak* $L^\infty(\Omega)$ topology. Any of these limit points is a global solution of the continuous problem. Moreover, we have

$$\lim_{h\to 0} \left\{ \|\overline{u} - \overline{u}_h\|_{L^\infty(\Omega_h)} \right\} = 0 \quad \text{and} \quad \lim_{h\to 0} J_h(\overline{u}_h) = J(\overline{u}).$$

Theorem (approximation of strict local minima)

Let $\overline{u}$ be a strict local minimum of the continuous problem, then there exists a sequence $\{\overline{u}_h\}_{h>0}$ of local minima of the discrete problems which converge towards $\overline{u}$.

Error Estimates

Theorem (piecewise constant discretization)

Let $\overline{u}$ be a solution of the continuous problem and $\{\overline{u}_h\}$ a sequence of solutions of the discrete problems converging towards $\overline{u}$. Moreover, assume that the second-order sufficient condition is satisfied.
Then there exists $C > 0$ such that

$$\|\overline{u} - \overline{u}_h\|_{L^\infty(\Omega_h)} + \|\overline{y} - \overline{y}_h\|_{L^\infty(\Omega_h)} + \|\overline{p} - \overline{p}_h\|_{L^\infty(\Omega_h)} + \|\overline{\lambda} - \overline{\lambda}_h\|_{L^\infty(\Omega_h)} \leq C\,h.$$

Error Estimates

Theorem (piecewise constant discretization)

Let $\overline{u}$ be a solution of the continuous problem and $\{\overline{u}_h\}$ a sequence of solutions of the discrete problems converging towards $\overline{u}$. Moreover, assume that the second-order sufficient condition is satisfied.
Then there exists $C > 0$ such that

$$\|\overline{u}-\overline{u}_h\|_{L^\infty(\Omega_h)}+\|\overline{y}-\overline{y}_h\|_{L^\infty(\Omega_h)}+\|\overline{p}-\overline{p}_h\|_{L^\infty(\Omega_h)}+\|\overline{\lambda}-\overline{\lambda}_h\|_{L^\infty(\Omega_h)} \le C\,h.$$

Idea of the proof

Extend $\overline{u}_h$ to $\Omega \setminus \Omega_h$ by $\overline{u}$. We obtain by optimality

$$f'(\overline{u})(\overline{u}_h - \overline{u}) + \mu \int_\Omega \overline{\lambda}\ (\overline{u}_h - \overline{u}\ )\,\mathrm{d}x \ge 0$$

$$f_h'(\overline{u}_h)(u_h - \overline{u}_h) + \mu \int_\Omega \overline{\lambda}_h(u_h - \overline{u}_h)\,\mathrm{d}x \ge 0 \quad \text{for all } u_h \in U_h \cap U_{\mathrm{ad}}$$

Error Estimates

Theorem (piecewise constant discretization)

Let $\overline{u}$ be a solution of the continuous problem and $\{\overline{u}_h\}$ a sequence of solutions of the discrete problems converging towards $\overline{u}$. Moreover, assume that the second-order sufficient condition is satisfied.
Then there exists $C > 0$ such that

$$\|\overline{u}-\overline{u}_h\|_{L^\infty(\Omega_h)}+\|\overline{y}-\overline{y}_h\|_{L^\infty(\Omega_h)}+\|\overline{p}-\overline{p}_h\|_{L^\infty(\Omega_h)}+\|\overline{\lambda}-\overline{\lambda}_h\|_{L^\infty(\Omega_h)} \le C\, h.$$

Idea of the proof

$$\le \big[f'(\overline{u}_h) - f'(\overline{u})\big](\overline{u}_h - \overline{u}) \le \dots$$

Error Estimates

Theorem (piecewise constant discretization)

Let $\overline{u}$ be a solution of the continuous problem and $\{\overline{u}_h\}$ a sequence of solutions of the discrete problems converging towards $\overline{u}$. Moreover, assume that the second-order sufficient condition is satisfied.

Then there exists $C > 0$ such that

$$\|\overline{u}-\overline{u}_h\|_{L^\infty(\Omega_h)} + \|\overline{y}-\overline{y}_h\|_{L^\infty(\Omega_h)} + \|\overline{p}-\overline{p}_h\|_{L^\infty(\Omega_h)} + \|\overline{\lambda}-\overline{\lambda}_h\|_{L^\infty(\Omega_h)} \le C\, h.$$

Idea of the proof

$$\frac{\delta}{2}\|\overline{u}_h - \overline{u}\|^2_{L^2(\Omega)} \le \big[f'(\overline{u}_h) - f'(\overline{u})\big](\overline{u}_h - \overline{u}) \le \dots$$

since $\overline{u}_h - \overline{u} \in C^\tau_{\overline{u}}$ and SSC hold

CHEMNITZ UNIVERSITY OF TECHNOLOGY

Error Estimates

Theorem (piecewise constant discretization)

Let $\overline{u}$ be a solution of the continuous problem and $\{\overline{u}_h\}$ a sequence of solutions of the discrete problems converging towards $\overline{u}$. Moreover, assume that the second-order sufficient condition is satisfied.
Then there exists $C > 0$ such that

$$\|\overline{u}-\overline{u}_h\|_{L^\infty(\Omega_h)}+\|\overline{y}-\overline{y}_h\|_{L^\infty(\Omega_h)}+\|\overline{p}-\overline{p}_h\|_{L^\infty(\Omega_h)}+\|\overline{\lambda}-\overline{\lambda}_h\|_{L^\infty(\Omega_h)} \le C\,h.$$

Theorem (variational discretization, Hinze (2005))

Let $\overline{u}$ be a solution of the continuous problem and $\{\overline{u}_h\}$ a sequence of solutions of the variational discretized probem, converging towards $\overline{u}$. Moreover, assume that the second-order sufficient condition is satisfied.
Then there is $C > 0$, such that

$$\|\overline{u}-\overline{u}_h\|_{L^2(\Omega_h)}+\|\overline{y}-\overline{y}_h\|_{L^2(\Omega_h)}+\|\overline{p}-\overline{p}_h\|_{L^2(\Omega_h)}+\|\overline{\lambda}-\overline{\lambda}_h\|_{L^2(\Omega_h)} \le C\,h^2.$$

Test Problem

Control problem

$$\text{Minimize} \quad \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + 10^{-3}\,\|u\|^2_{L^2(\Omega)} + 3\cdot 10^{-2}\,\|u\|_{L^1(\Omega)}$$

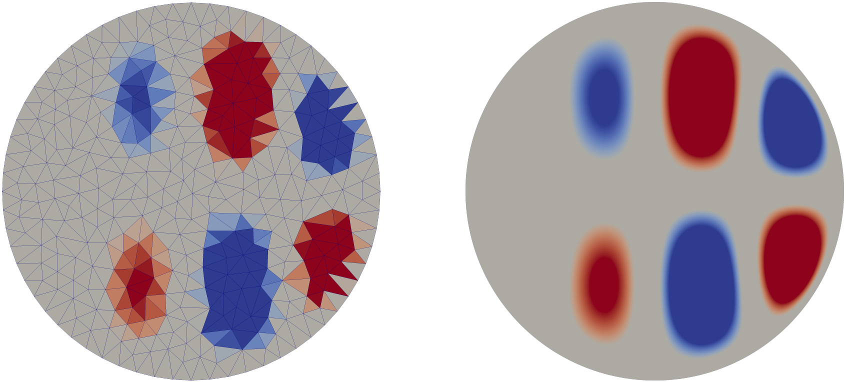
$$\text{such that} \quad u_a \le u \le u_b$$

- $y_d(x_1, x_2) = 2\,\sin(2\,\pi\,x_1)\,\sin(\pi\,x_2)\,\exp(x_1)$
- PDE:

$$\begin{aligned} -\Delta y + y^3 &= u \quad \text{in } \Omega \\ y &= 0 \quad \text{on } \Gamma \end{aligned}$$

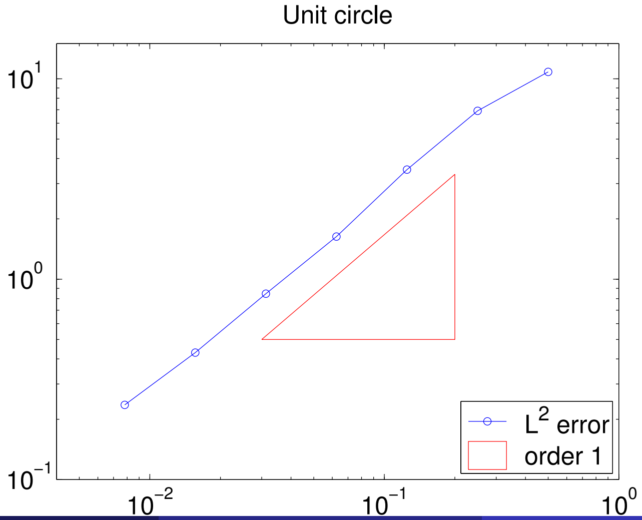
CHEMNITZ UNIVERSITY OF TECHNOLOGY

Solutions for $h = 2^{-3}$ and $h = 2^{-8}$



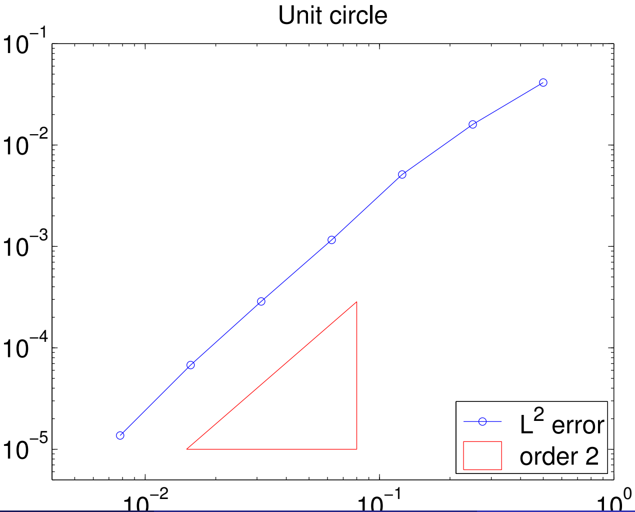
Convergence (Full Discretization)

Error in the control:



Convergence (Variational Discretization)

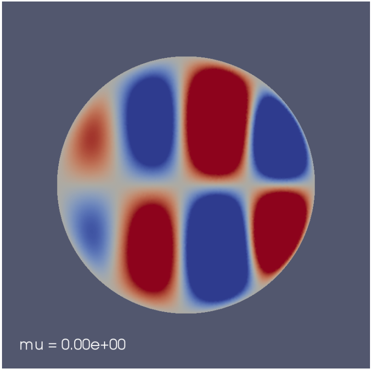
Error in the adjoint:



Influence of Parameter μ

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

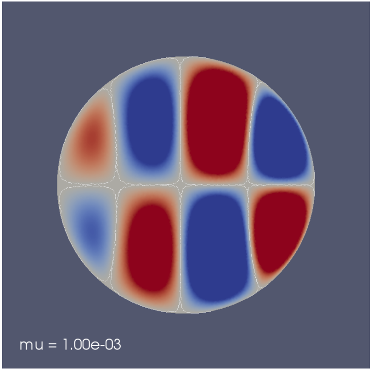


$\mu = 0.00$

Influence of Parameter μ

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu \, \|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

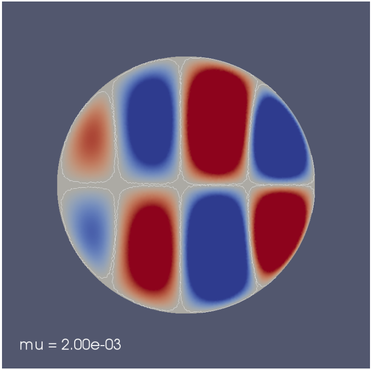


$\mu = 1.00\text{E–}03$

Influence of Parameter μ

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

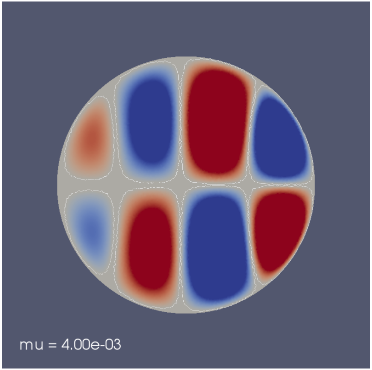


$\mu = 2.00\text{E–}03$

Influence of Parameter μ

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

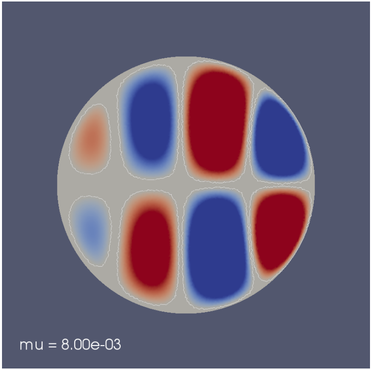


$\mu = 4.00\text{E–}03$

Influence of Parameter μ

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

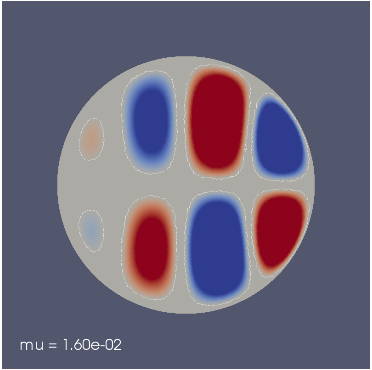


$\mu = 8.00\text{E–}03$

Influence of Parameter μ

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

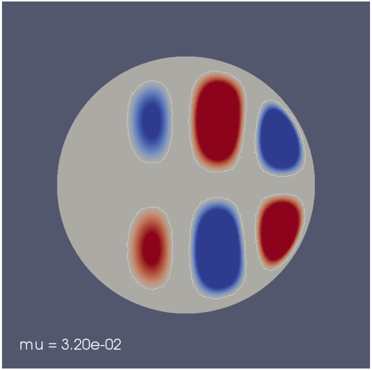


$\mu = 1.60\text{E–}02$

Influence of Parameter μ

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

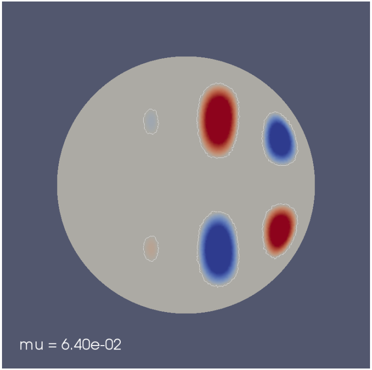


$\mu = 3.20\text{E–}02$

Influence of Parameter μ

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$

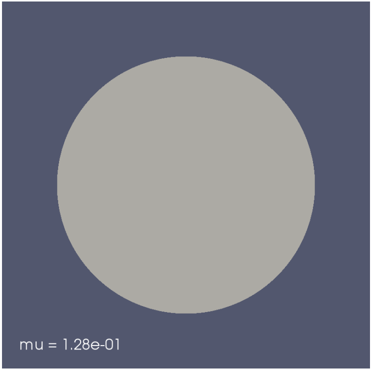


$\mu = 6.40\text{E}–02$

Influence of Parameter μ

$$\text{Minimize} \quad \frac{1}{2}\|G(u) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} + \mu\,\|u\|_{L^1(\Omega)}$$

$$\text{such that} \quad u_a \le u \le u_b \qquad (u_a < 0 < u_b)$$



$\mu = 1.28\text{E}{-}01$

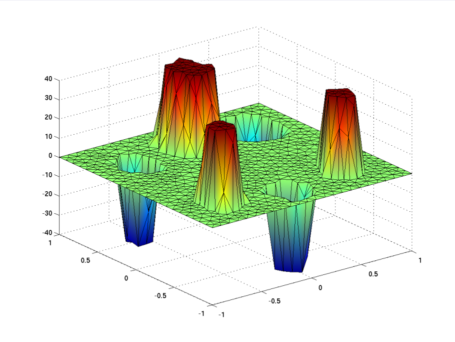
CHEMNITZ UNIVERSITY OF TECHNOLOGY

Overview

1. Introduction and Problem Setting
2. 1st- and 2nd-Order Optimality Conditions
3. Finite Element Error Estimates and Examples
4. Extension: Directional Sparsity
(joint with Georg Stadler, ICES, Texas)

Can we do Better Than Just Sparse?

Sparsity

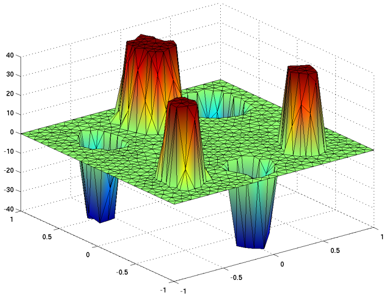


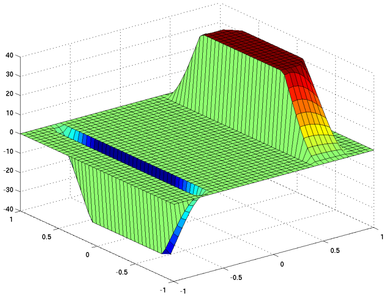
Objective function

$$\frac{1}{2}\|y - y_d\|_{L^2}^2 + \beta\,\|u\|_{L^1}$$

Can we do Better Than Just Sparse?

Sparsity vs. directional sparsity





Objective function

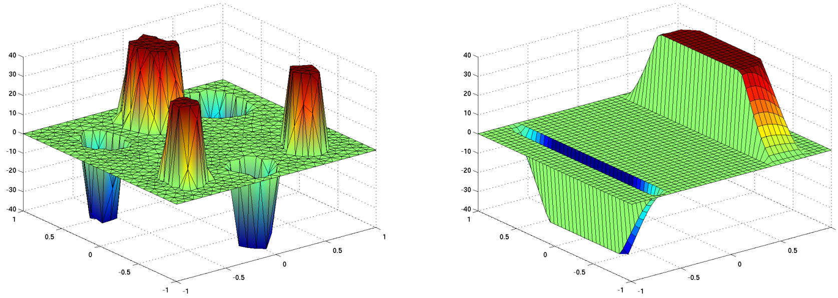
$$\frac{1}{2}\|y - y_d\|_{L^2}^2 + \beta \, \|u\|_{L^1}$$

Objective function

$$\frac{1}{2}\|y - y_d\|_{L^2}^2 + \beta \, \|u\|_{L^1(L^2)}$$

Can we do Better Than Just Sparse?

Sparsity vs. directional sparsity

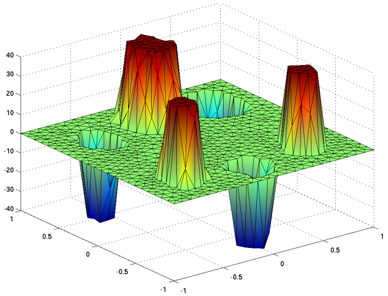


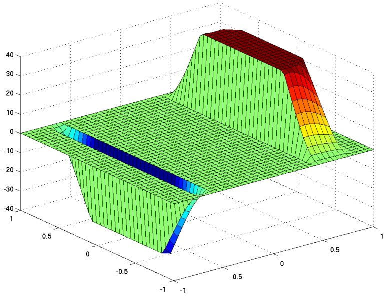
Properties

- no structural assumptions made

Can we do Better Than Just Sparse?

Sparsity vs. directional sparsity





Properties

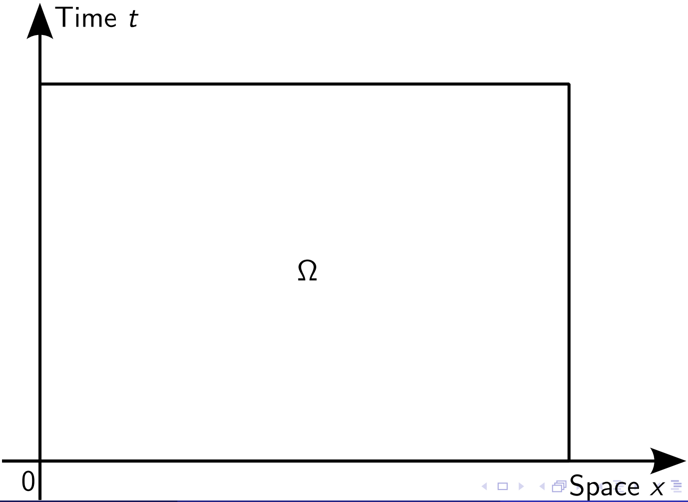
- no structural assumptions made

Properties

- exploits known or desired group sparsity structure

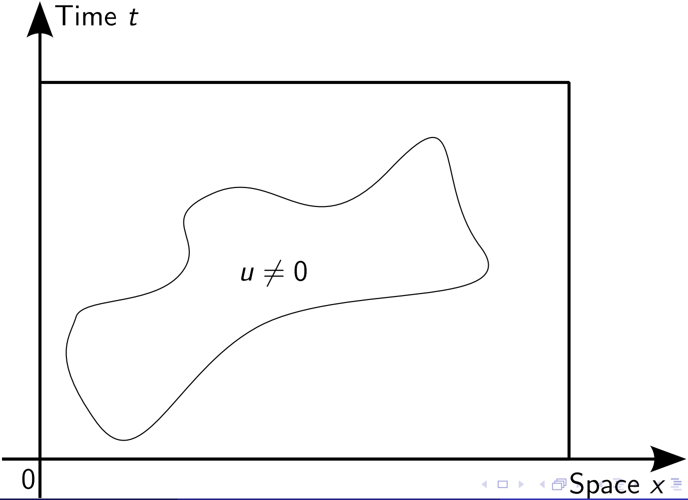
Directional Sparsity with Parabolic PDEs

Placement of actuators for a parabolic problem



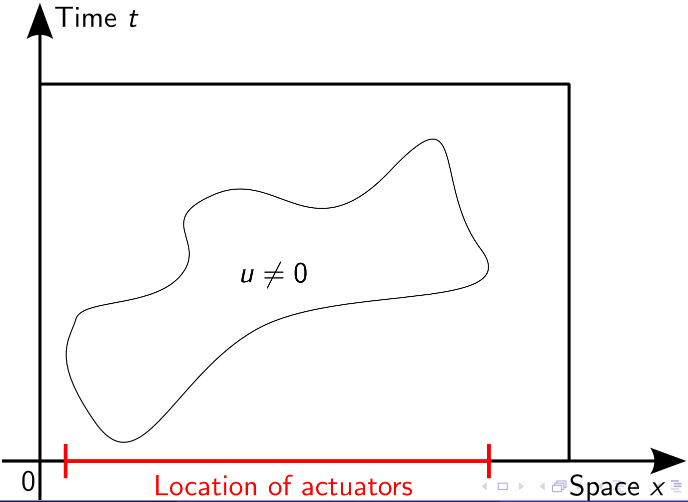
Directional Sparsity with Parabolic PDEs

With Sparsity functional



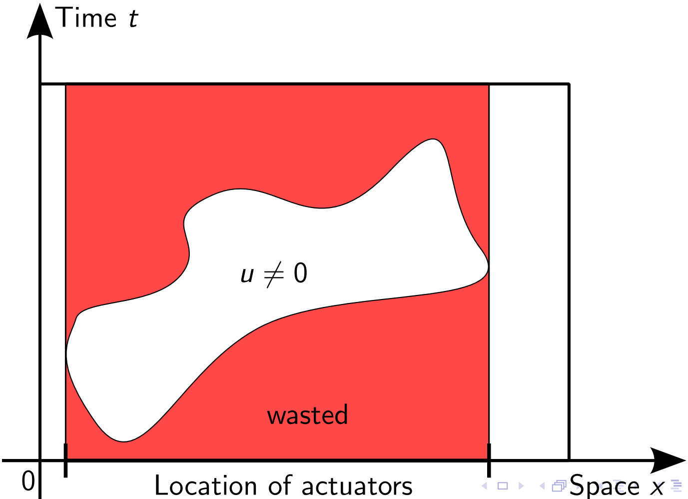
Directional Sparsity with Parabolic PDEs

With Sparsity functional



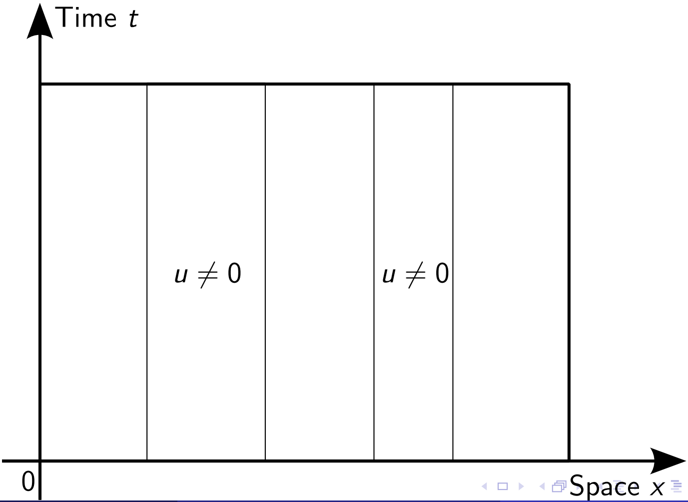
Directional Sparsity with Parabolic PDEs

With Sparsity functional



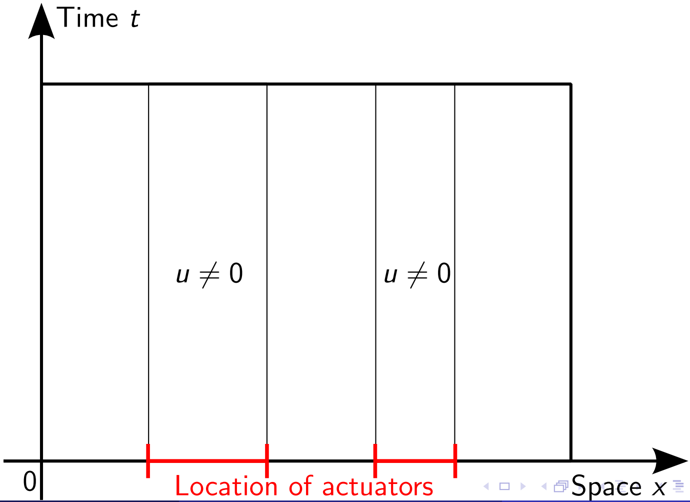
Directional Sparsity with Parabolic PDEs

With Directional Sparsity functional



Directional Sparsity with Parabolic PDEs

With Directional Sparsity functional

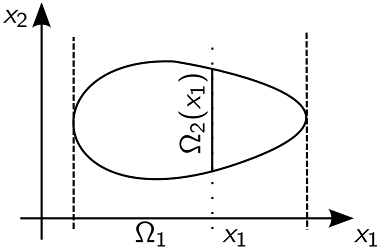


Directional Sparsity: Basic Definition

Problem formulation

$$\min \quad \frac{1}{2}\|\mathcal{S}u - y_d\|_H^2 + \frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2 + \beta\,\|u\|_{L^1(L^2)}$$

$$\text{s.t.} \quad u_a \le u \le u_b \quad \text{a.e. in } \Omega$$

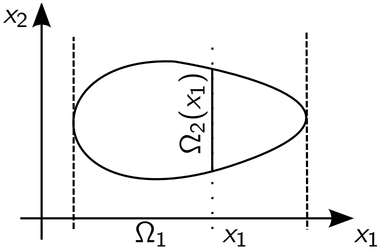


Directional Sparsity: Basic Definition

Problem formulation

$$\min \quad \frac{1}{2}\|\mathcal{S}u - y_d\|_H^2 + \frac{\alpha}{2}\|u\|_{L^2(\Omega)}^2 + \beta \int_{\Omega_1} \left(\int_{\Omega_2(x_1)} u(x_1,x_2)^2 \, \mathrm{d}x_2 \right)^{1/2} \mathrm{d}x_1$$

$$\text{s.t.} \quad u_a \le u \le u_b \quad \text{a.e. in } \Omega$$



Related Approaches

Joint sparsity in image restoration

$$\Psi(u) = \sum_{\lambda \in \Lambda} \omega_\lambda \, |\vec{u}_\lambda|_q^p, \quad q = 2, p = 1$$

- $\Omega_1 \mathrel{\widehat{=}} \Lambda$
- $\Omega_2 = \{1, 2, \ldots, \# \text{ of channels}\}$
- with dx_2 = counting measure

[Fornasier, Ramlau, Teschke (2008)]

TV-based image restoration

$$\Psi(u) = \int_{\Omega_1} |\nabla u| \, dx_1, \quad \Omega_2 = \{1, 2, \ldots, N\} \text{ for N-D images}$$

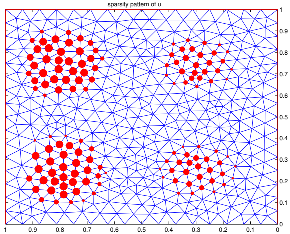
Parabolic Example with Spatial Sparsity

Parabolic example

$$\text{minimize} \quad \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \beta\,\|u\|_{L^1(L^2)}$$

$$\text{s.t.} \quad \begin{cases} y_t - \frac{1}{10}\Delta y = u & \text{in } \Omega = \Omega_1 \times (0, T) \\ y = 0 & \text{on } \Gamma \times (0, T) \\ y(\cdot, 0) = 0 & \text{in } \Omega_1 \end{cases}$$

$$\text{and} \quad u_a \le u \le u_b \quad \text{a.e. in } \Omega$$



$n = 2$ sparse directions (space)

$N - n = 1$ non-sparse direction (time)

Summary

- use of $\|u\|_{L^1}$ induces sparse solutions
- it is often an appropriate measure of control cost
- applications in actuator placement problems
- presented 1st- and new 2nd-order optimality conditions
- used them to derive FE error estimates
- extension to directional sparsity concept

References I

E. Casas, R. Herzog, and G. Wachsmuth.
Optimality conditions and error analysis of semilinear elliptic control problems with L^1 cost functional.
Technical report, 2010.

C. Clason and K. Kunisch.
A duality-based approach to elliptic control problems in non-reflexive Banach spaces.
ESAIM: Control, Optimisation, and Calculus of Variations, in print.
doi: 10.1051/cocv/2010003.

M. Fornasier, R. Ramlau, and G. Teschke.
The application of joint sparsity and total variation minimization algorithms in a real-life art restoration problem.
Advances in Computational Mathematics, 31(1–3):301–329, 2009.
URL http://dx.doi.org/10.1007/s10444-008-9103-6.

M. Hinze.
A variational discretization concept in control constrained optimization: The linear-quadratic case.
Computational Optimization and Applications, 30(1):45–61, 2005.

References II

G. Stadler.
Elliptic optimal control problems with L^1-control cost and applications for the placement of control devices.
Computational Optimization and Applications, 44(2):159–181, 2009.
URL `http://dx.doi.org/10.1007/s10589-007-9150-9`.

G. Vossen and H. Maurer.
On L^1-minimization in optimal control and applications to robotics.
Optimal Control Applications and Methods, 27(6):301–321, 2006.