On Nonlinear Optimal Control Problems with an $L^1$ Norm

Eduardo Casas
Roland Herzog
Gerd Wachsmuth

University of Cantabria
Numerical Mathematics

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)
Control problem

Minimize \[ \frac{1}{2} \| y - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2 + \mu \| u \|_{L^1(\Omega)} \]
such that \[ u_a \leq u \leq u_b \quad (u_a < 0 < u_b) \]
and \[ y \text{ solves the PDE} \]

Semilinear partial differential equation

\[ -\Delta y + a(\cdot, y) = u \quad \text{in } \Omega \]
\[ y = 0 \quad \text{on } \Gamma \]
Why Consider $\|u\|_{L^1(\Omega)}$?

- The $L^1$-norm

$$\|u\|_{L^1(\Omega)} = \int_{\Omega} |u(x)| \, dx$$

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
  \[ \|u\|_{L^1(\Omega)} = \int_{\Omega} |u(x)| \, dx \]
  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

$\mu = 0$  

$\mu > 0$
Smooth minimization problem

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

Histogram (solution components’ sizes)
A First Glance at Sparsity

Smooth minimization problem

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

Convex minimization problem

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

Histogram (solution components’ sizes)

![Histogram graphs showing component sizes of solutions for different norms.](image)
### A First Glance at Sparsity

#### Smooth minimization problem

\[
\begin{align*}
\text{minimize} & \quad \frac{1}{2} \| x \|_2^2 \\
\text{s.t.} & \quad Ax = b
\end{align*}
\]

\[
\begin{align*}
x + A^\top p &= 0 \\
Ax - b &= 0
\end{align*}
\]

#### Convex minimization problem

\[
\begin{align*}
\text{minimize} & \quad \| x \|_1 \\
\text{s.t.} & \quad Ax = b
\end{align*}
\]

\[
\begin{align*}
\lambda + A^\top p &= 0, \quad \lambda \in \partial \| x \|_1 \\
Ax - b &= 0
\end{align*}
\]

\[
\lambda_i = \begin{cases} 
-1 & \text{if } x_i < 0 \\
+1 & \text{if } x_i > 0 \\
[-1, 1] & \text{if } x_i = 0
\end{cases}
\]
### A First Glance at Sparsity

#### Smooth minimization problem

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

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

#### Convex minimization problem

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

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

\[
\lambda_i = -1 \quad \text{if } x_i < 0 \\
\lambda_i = +1 \quad \text{if } x_i > 0 \\
\lambda_i \in [-1,1] \quad \text{if } x_i = 0
\]

\[
x_i = \max\{0, x_i + c (\lambda_i - 1)\} + \min\{0, x_i + c (\lambda_i + 1)\}
\]
Control problem

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

such that \( u_a \leq u \leq u_b \) \( (u_a < 0 < u_b) \)

and \( y \) solves the PDE

Semilinear partial differential equation

\[-\Delta y + a(\cdot, y) = u \text{ in } \Omega\]

\[y = 0 \text{ on } \Gamma\]
Semilinear partial differential equation

\[- \Delta y + a(\cdot, y) = u \quad \text{in } \Omega \]
\[y = 0 \quad \text{on } \Gamma\]

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 \( u \in L^p(\Omega), n/2 < p \leq 2 \) the solution \( y = G(u) \in W^{2,p}(\Omega) \)
- \( G : L^p(\Omega) \rightarrow W^{2,p}(\Omega) \) is \( C^2 \), derivatives by linearization
Basic Assumptions Concerning the PDE

Semilinear partial differential equation

\[- \text{div}(A \nabla y) + a(\cdot, y) = u \quad \text{in } \Omega \]
\[y = 0 \quad \text{on } \Gamma\]

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. \(y\)
- \(\xi^\top A(x) \xi \geq a \|\xi\|^2\) for all \(\xi \in \mathbb{R}^n, a > 0\)

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

Minimize \[ \frac{1}{2} \| y - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2 + \mu \| u \|_{L^1(\Omega)} \]
such that \[ u_a \leq u \leq u_b \quad (u_a < 0 < u_b) \]
and \( y \) solves the PDE

Semilinear partial differential equation

\[ - \text{div}(A \nabla y) + a(\cdot, y) = u \quad \text{in } \Omega \\
\quad y = 0 \quad \text{on } \Gamma \]
Control problem

Minimize \[ \frac{1}{2} \left\| G(u) - y_d \right\|^2_{L^2(\Omega)} + \frac{\nu}{2} \left\| u \right\|^2_{L^2(\Omega)} + \mu \left\| u \right\|_{L^1(\Omega)} \]

such that \[ u_a \leq u \leq u_b \quad (u_a < 0 < u_b) \]
Control problem

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

such that \( u_a \leq u \leq u_b \quad (u_a < 0 < u_b) \)

Properties

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

Minimize

\[ \frac{1}{2} \| G(u) - y_d \|^2_{L^2(\Omega)} + \frac{\nu}{2} \| u \|^2_{L^2(\Omega)} + \mu \| u \|_{L^1(\Omega)} \]

such that

\[ u_a \leq u \leq u_b \quad (u_a < 0 < u_b) \]

Properties

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

Minimize
\[ \frac{1}{2} \| G(u) - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2 + \mu \| u \|_{L^1(\Omega)} \]

such that
\[ u_a \leq u \leq u_b \quad (u_a < 0 < u_b) \]

Properties

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

Minimize

\[ \frac{1}{2} \| G(u) - y_d \|^2_{L^2(\Omega)} + \frac{\nu}{2} \| u \|^2_{L^2(\Omega)} + \mu \| u \|_{L^1(\Omega)} \]

such that

\[ u_a \leq u \leq u_b \quad (u_a < 0 < u_b) \]

Properties

- is differentiable w.r.t. \( u \in L^2(\Omega) \)
- is convex w.r.t. \( u \)
Problem Setting

Control problem

Minimize
\[ \frac{1}{2} \| G(u) - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2 + \mu \| u \|_{L^1(\Omega)} \]

such that
\[ u_a \leq u \leq u_b \]

Properties

- is differentiable w.r.t. \( u \in L^2(\Omega) \)
- is convex w.r.t. \( u \)
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. $\partial J$ does not depend on the splitting of $J$ into $f$ and $j$. 
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

\[
    f(u) = \frac{1}{2} \| G(u) - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2, \quad j(u) = \| u \|_{L^1(\Omega)}
\]
First-Order Necessary Condition

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

\[ \nabla f(\bar{u}) = G'(\bar{u})^*(\bar{y} - y_d) + \nu \bar{u}, \quad \text{where} \quad \bar{y} = G(\bar{u}) \]

adjoint state \( \bar{p} \)
First-Order Necessary Condition

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

\[ \nabla f(\bar{u}) = G'(\bar{u})^*(\bar{y} - y_d) + \nu \bar{u}, \quad \text{where} \quad \bar{y} = G(\bar{u}) \]

adjoint state \( \bar{p} \)

First-order necessary optimality conditions

\[ 0 \in \nabla f(\bar{u}) + \mu \partial j(\bar{u}) \]

\[ \Leftrightarrow \quad 0 = \nabla f(\bar{u}) + \mu \bar{\lambda}, \quad \bar{\lambda} \in \partial j(\bar{u}) \]
First-Order Necessary Condition

\[
\begin{align*}
 f(u) &= \frac{1}{2} \| G(u) - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2, \\
 j(u) &= \| u \|_{L^1(\Omega)}, \\
 \nabla f(u) &= G'(u)^*(\bar{y} - y_d) + \nu \bar{u}, \quad \text{where} \quad \bar{y} = G(\bar{u})
\end{align*}
\]

First-order necessary optimality conditions

\[
0 \in \nabla f(u) + \mu \partial j(u)
\]

\[
\Leftrightarrow \quad 0 = \nabla f(u) + \mu \bar{\lambda}, \quad \bar{\lambda} \in \partial j(u)
\]

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

\[
0 \leq \langle \nabla f(u) + \mu \bar{\lambda}, u - \bar{u} \rangle_{L^2(\Omega)} \quad \text{for all} \quad u \in U_{\text{ad}}, \quad \bar{\lambda} \in \partial j(u)
\]
First-Order Necessary Condition

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

\[ \nabla f(\overline{u}) = G'(\overline{u})^*(\overline{y} - y_d) + \nu \overline{u}, \quad \text{where} \quad \overline{y} = G(\overline{u}) \]

adjoint state \( \overline{p} \)

First-order necessary optimality conditions

\[ 0 \in \nabla f(\overline{u}) + \mu \partial j(\overline{u}) \]

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

\[ 0 \leq \langle \nabla f(\overline{u}) + \mu \overline{\lambda}, u - \overline{u} \rangle_{L^2(\Omega)} \quad \text{for all} \quad u \in U_{\text{ad}}, \quad \overline{\lambda} \in \partial j(\overline{u}) \]
First-Order Necessary Condition

**Theorem**

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

\[
\langle \bar{p} + \nu \bar{u} + \mu \bar{\lambda}, u - \bar{u} \rangle_{L^2(\Omega)} \geq 0 \quad \text{for all } u \in U_{ad}.
\]
First-Order Necessary Condition

Theorem

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

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

Subgradient of the \( L^1 \) norm

\[
\bar{\lambda}(x) \begin{cases} 
= +1 & \text{where } \bar{u}(x) > 0 \\
\in [-1, 1] & \text{where } \bar{u}(x) = 0 \\
= -1 & \text{where } \bar{u}(x) < 0
\end{cases}
\]
First-Order Necessary Condition

**Theorem**

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

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

**Adjoint equation**

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

$$\bar{p} = 0 \quad \text{on } \Gamma$$
Theorem

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

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

Corollary: projection formulas

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

$$\overline{\lambda}(x) = \text{proj}_{[-1, 1]} \left( - \frac{1}{\mu} \overline{p}(x) \right)$$

$$\overline{u}(x) = 0 \quad \iff \quad |\overline{p}(x)| \leq \mu$$
First-Order Necessary Condition

**Theorem**

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

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

**Corollary: projection formulas**

$$\bar{u}(x) = \text{proj}_{[u_a, u_b]} \left( -\frac{1}{\nu} (\bar{p}(x) + \mu \bar{\lambda}(x)) \right)$$

$$\bar{\lambda}(x) = \text{proj}_{[-1, +1]} \left( -\frac{1}{\mu} \bar{p}(x) \right)$$

$$\bar{u}(x) = 0 \iff |\bar{p}(x)| \leq \mu$$

It follows that $\bar{u}, \bar{\lambda} \in C^{0,1}(\overline{\Omega}) = W^{1,\infty}(\Omega)$. Moreover, $\bar{\lambda} \in \partial \|\bar{u}\|_{L^1(\Omega)}$ is unique.
Critical cone at stationary point $\bar{u}$ with associated $\bar{\lambda} \in \partial j(\bar{u})$

$$C^+_u := \{ v \in L^2(\Omega) : f'(\bar{u}) v + \mu \langle \bar{\lambda}, v \rangle = 0 \}$$
Second-Order Optimality Conditions

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

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

$$\langle f''(\bar{u}) v, v \rangle > 0 \quad \text{for all } v \in C^+_{\bar{u}} \setminus \{0\} \quad \Rightarrow \quad \bar{u} \text{ is locally optimal}$$
Critical cone at stationary point $\bar{u}$ with associated $\bar{\lambda} \in \partial j(\bar{u})$

$$C_{\bar{u}}^+ := \{ v \in L^2(\Omega) : f'(\bar{u}) v + \mu \langle \bar{\lambda}, v \rangle = 0 \}$$

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

$$\langle f''(\bar{u}) v, v \rangle \geq 0 \quad \text{for all } v \in C_{\bar{u}}^+ \quad \not\Rightarrow \quad \bar{u} \text{ is locally optimal}$$
Second-Order Optimality Conditions

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

$$C_{\bar{u}}^+ := \{ v \in L^2(\Omega) : f'(\bar{u}) v + \mu \langle \bar{\lambda}, v \rangle = 0 \} \text{ too large}$$

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

$$\langle f''(\bar{u}) v, v \rangle \geq 0 \quad \text{for all } v \in C_{\bar{u}}^+ \quad \nRightarrow \quad \bar{u} \text{ is locally optimal}$$
Second-Order Optimality Conditions

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

\[ C^+_\bar{u} := \{ v \in L^2(\Omega) : f'(\bar{u}) v + \mu \langle \bar{\lambda}, v \rangle = 0 \} \quad \text{too large} \]
\[ C^-\bar{u} := \{ v \in L^2(\Omega) : f'(\bar{u}) v + \mu j'(\bar{u}; v) = 0 \} \quad \text{correct} \]

\[ \langle f''(\bar{u}) v, v \rangle > 0 \quad \text{for all } v \in C^-\bar{u} \setminus \{0\} \quad \Rightarrow \quad \bar{u} \text{ is locally optimal} \]
\[ \langle f''(\bar{u}) v, v \rangle \geq 0 \quad \text{for all } v \in C^-\bar{u} \quad \Leftrightarrow \quad \bar{u} \text{ is locally optimal} \]
Second-Order Optimality Conditions

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

\[
C^+_\bar{u} := \{ v \in L^2(\Omega) : f'(\bar{u}) v + \mu \langle \bar{\lambda}, v \rangle = 0 \} \quad \text{too large}
\]

\[
C_{\bar{u}} := \{ v \in L^2(\Omega) : f'(\bar{u}) v + \mu j'(\bar{u}; v) = 0 \} \quad \text{correct}
\]

\[
\langle f''(\bar{u}) v, v \rangle > 0 \quad \text{for all } v \in C_{\bar{u}} \setminus \{0\} \quad \Rightarrow \quad \bar{u} \text{ is locally optimal}
\]

\[
\langle f''(\bar{u}) v, v \rangle \geq 0 \quad \text{for all } v \in C_{\bar{u}} \quad \Leftarrow \quad \bar{u} \text{ is locally optimal}
\]

... with control constraints

\[
C_{\bar{u}} := \{ v \in L^2(\Omega) : f'(\bar{u}) v + \mu j'(\bar{u}; v) = 0 \}
\]

\[
\begin{align*}
&v \geq 0 \text{ where } \bar{u} = u_a \\
&v \leq 0 \text{ where } \bar{u} = u_b
\end{align*}
\]

\[
v \in T_{U_{ad}}(\bar{u})
\]
Second-Order Sufficient Conditions

Critical cone (closed, convex)

\[ C_u := \{ v \in T_{U_{\text{ad}}} (\bar{u}) : f'(\bar{u}) v + \mu j'(\bar{u}; v) = 0 \} \]

Theorem

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

\[
J(\bar{u}) + \frac{\delta}{2} \| u - \bar{u} \|_{L^2(\Omega)}^2 \leq J(u) \quad \text{for all} \quad u \in U_{\text{ad}} \cap B_{\varepsilon}^{L^2}(\bar{u}).
\]
Second-Order Sufficient Conditions

Critical cone (closed, convex)

\[ C_{\bar{u}} := \{ v \in \mathcal{T}_{U_{ad}}(\bar{u}) : f'(\bar{u}) v + \mu j'(\bar{u}; v) = 0 \} \]

Theorem

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

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

Corollary

There exist \( \tau > 0, \delta_2 > 0 \) such that \( \langle f''(\bar{u}) v, v \rangle \geq \delta_2 \| v \|_{L^2(\Omega)}^2 \) for all \( v \in C^T_{\bar{u}} = \{ v \in \mathcal{T}_{U_{ad}}(\bar{u}) : f'(\bar{u}) v + \mu j'(\bar{u}; v) \leq \tau \| v \|_{L^2(\Omega)} \} \)
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 \( \{ T_h \} \) of \( \Omega \), \( \Omega_h = \bigcup_{T \in T_h} T \).
- Discrete space of (adjoint) states (piecewise linear):
  \[
  Y_h = \{ y_h \in C(\overline{\Omega}) : y_h|_T \in P_1 \text{ for all } T \in T_h, \text{ and } y_h = 0 \text{ on } \overline{\Omega} \setminus \Omega_h \}.
  \]
- Discrete PDE:
  \[
  \int_{\Omega_h} \nabla z_h^\top A \nabla y_h + a(\cdot, y_h) \, dx = \int_{\Omega_h} u z_h \, dx \quad \text{for all } z_h \in Y_h
  \]
- Discrete space of controls (piecewise constant):
  \[
  U_h = \{ u_h \in L^2(\Omega_h) : u_h|_T \equiv \text{const} \text{ for all } T \in T_h \}.
  \]
Discrete optimization problem

Minimize \( \frac{1}{2} \| G_h(u_h) - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u_h \|_{L^2(\Omega)}^2 + \mu \| u_h \|_{L^1(\Omega)} \)

such that \( u_a \leq u_h \leq u_b \)

and \( u_h \in U_h \)
Theorem (approximation of global minima)

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

$$\lim_{h \to 0} \{ \| \bar{u} - \bar{u}_h \|_{L^\infty(\Omega_h)} \} = 0 \quad \text{and} \quad \lim_{h \to 0} J_h(\bar{u}_h) = J(\bar{u}).$$
**Theorem (approximation of global minima)**

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

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

**Theorem (approximation of strict local minima)**

Let \( \bar{u} \) be a strict local minimum of the continuous problem, then there exists a sequence \( \{ \bar{u}_h \}_{h>0} \) of local minima of the discrete problems which converge towards \( \bar{u} \).
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 \( \bar{u} \) be a solution of the continuous problem and \( \{ \bar{u}_h \} \) a sequence of solutions of the discrete problems converging towards \( \bar{u} \). Moreover, assume that the second-order sufficient condition is satisfied.

Then there exists \( C > 0 \) such that

\[
\| \bar{u} - \bar{u}_h \|_{L^\infty(\Omega_h)} + \| \bar{y} - \bar{y}_h \|_{L^\infty(\Omega_h)} + \| \bar{p} - \bar{p}_h \|_{L^\infty(\Omega_h)} + \| \bar{\lambda} - \bar{\lambda}_h \|_{L^\infty(\Omega_h)} \leq C h.
\]

**Idea of the proof**

Extend \( \bar{u}_h \) to \( \Omega \setminus \Omega_h \) by \( \bar{u} \). We obtain by optimality

\[
f'(\bar{u})(\bar{u}_h - \bar{u}) + \mu \int_{\Omega} \bar{\lambda} \ (\bar{u}_h - \bar{u}) \ dx \geq 0
\]

\[
f'_h(\bar{u}_h)(u_h - \bar{u}_h) + \mu \int_{\Omega} \bar{\lambda}_h (u_h - \bar{u}_h) \ dx \geq 0 \quad \text{for all} \ u_h \in U_h \cap U_{ad}
\]
Error Estimates

Theorem (piecewise constant discretization)

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

\[
\| \bar{u} - \bar{u}_h \|_{L^\infty(\Omega_h)} + \| \bar{y} - \bar{y}_h \|_{L^\infty(\Omega_h)} + \| \bar{p} - \bar{p}_h \|_{L^\infty(\Omega_h)} + \| \bar{\lambda} - \bar{\lambda}_h \|_{L^\infty(\Omega_h)} \leq C h.
\]

Idea of the proof

\[
\leq \left[ f'(\bar{u}_h) - f'(\bar{u}) \right] (\bar{u}_h - \bar{u}) \leq \ldots
\]
**Error Estimates**

**Theorem (piecewise constant discretization)**

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

\[
\| \bar{u} - \bar{u}_h \|_{L^\infty(\Omega_h)} + \| \bar{y} - \bar{y}_h \|_{L^\infty(\Omega_h)} + \| \bar{p} - \bar{p}_h \|_{L^\infty(\Omega_h)} + \| \bar{\lambda} - \bar{\lambda}_h \|_{L^\infty(\Omega_h)} \leq C h.
\]

**Idea of the proof**

\[
\frac{\delta}{2} \| \bar{u}_h - \bar{u} \|^2_{L^2(\Omega)} \leq \left[ f'(\bar{u}_h) - f'(\bar{u}) \right] (\bar{u}_h - \bar{u}) \leq \ldots
\]

since \( \bar{u}_h - \bar{u} \in \mathcal{C}_{\bar{u}}^T \) and SSC hold
**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

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

**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 problem, converging towards \( \overline{u} \). Moreover, assume that the second-order sufficient condition is satisfied.

Then there is \( C > 0 \), such that

\[
\| \overline{u} - \overline{u}_h \|_{L^2(\Omega_h)} + \| \overline{y} - \overline{y}_h \|_{L^2(\Omega_h)} + \| \overline{p} - \overline{p}_h \|_{L^2(\Omega_h)} + \| \overline{\lambda} - \overline{\lambda}_h \|_{L^2(\Omega_h)} \leq C h^2.
\]
Test Problem

Control problem

Minimize \[ \frac{1}{2} \| y - y_d \|_{L^2(\Omega)}^2 + 10^{-3} \| u \|_{L^2(\Omega)}^2 + 3 \cdot 10^{-2} \| u \|_{L^1(\Omega)} \]

such that \[ u_a \leq u \leq u_b \]

- \[ y_d(x_1, x_2) = 2 \sin(2 \pi x_1) \sin(\pi x_2) \exp(x_1) \]
- PDE:
  \[ -\Delta y + y^3 = u \quad \text{in } \Omega \]
  \[ y = 0 \quad \text{on } \Gamma \]
Solutions for $h = 2^{-3}$ and $h = 2^{-8}$
Convergence (Full Discretization)

Error in the control:

![Graph showing error in the control for different mesh sizes. The graph plots L^2 error against mesh size, illustrating order 1 convergence. The unit circle is represented.]
Convergence (Variational Discretization)

Error in the adjoint:

Unit circle

\( L^2 \) error

order 2

\( 10^{-2} \)

\( 10^{-1} \)

\( 10^{-3} \)

\( 10^{-4} \)

\( 10^{-5} \)

\( 10^0 \)
Influence of Parameter $\mu$

Minimize
\[
\frac{1}{2} \| G(u) - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2 + \mu \| u \|_{L^1(\Omega)}
\]
such that
\[
u_a \leq u \leq u_b \quad (u_a < 0 < u_b)
\]

$\mu = 0.00$

$\mu = 0.00e+00$
Influence of Parameter $\mu$

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

such that $u_a \leq u \leq u_b$ ($u_a < 0 < u_b$)

$\mu = 1.00E-03$
Influence of Parameter $\mu$

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

such that $u_a \leq u \leq u_b$ ($u_a < 0 < u_b$)

$\mu = 2.00 \times 10^{-3}$
Influence of Parameter $\mu$

Minimize \[ \frac{1}{2} \| G(u) - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2 + \mu \| u \|_{L^1(\Omega)} \]

such that \[ u_a \leq u \leq u_b \quad (u_a < 0 < u_b) \]

$\mu = 4.00E-03$
Influence of Parameter $\mu$

Minimize \[
\frac{1}{2} \| G(u) - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2 + \mu \| u \|_{L^1(\Omega)}
\]
such that $u_a \leq u \leq u_b$ \quad ($u_a < 0 < u_b$)

$\mu = 8.00\mathrm{E}^{-03}$
Influence of Parameter $\mu$

Minimize \[ \frac{1}{2} \| G(u) - y_d \|^2_{L^2(\Omega)} + \frac{\nu}{2} \| u \|^2_{L^2(\Omega)} + \mu \| u \|_{L^1(\Omega)} \]

such that $u_a \leq u \leq u_b$ \quad ($u_a < 0 < u_b$)

$\mu = 1.60 \text{E}-02$

$\mu u = 1.60 \text{e-02}$
Influence of Parameter $\mu$

Minimize \[
\frac{1}{2} \| G(u) - y_d \|^2_{L^2(\Omega)} + \frac{\nu}{2} \| u \|^2_{L^2(\Omega)} + \mu \| u \|_{L^1(\Omega)}
\]
such that \[ u_a \leq u \leq u_b \quad (u_a < 0 < u_b) \]

$\mu = 3.20E-02$
Influence of Parameter $\mu$

Minimize \[ \frac{1}{2} \| G(u) - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2 + \mu \| u \|_{L^1(\Omega)} \]

such that $u_a \leq u \leq u_b$ ($u_a < 0 < u_b$)

$\mu = 6.40E-02$

$\mu u = 6.40e-02$
Influence of Parameter $\mu$

Minimize \[
\frac{1}{2} \| G(u) - y_d \|_{L^2(\Omega)}^2 + \frac{\nu}{2} \| u \|_{L^2(\Omega)}^2 + \mu \| u \|_{L^1(\Omega)}
\]
such that $u_a \leq u \leq u_b$ ($u_a < 0 < u_b$)

$\mu = 1.28E-01$

$\mu u = 1.28e-01$
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?

**Objective function**

\[
\frac{1}{2} \| y - y_d \|_{L^2}^2 + \beta \| u \|_{L^1}
\]
Can we do Better Than Just Sparse?

**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?

**Properties**
- no structural assumptions made
- exploits known or desired group sparsity structure
Placement of actuators for a parabolic problem
Directional Sparsity with Parabolic PDEs

With Sparsity functional

\[ u \neq 0 \]
Directional Sparsity with Parabolic PDEs

With Sparsity functional

\[ u \neq 0 \]

Location of actuators
Directional Sparsity with Parabolic PDEs

With Sparsity functional

\[ u \neq 0 \]

Location of actuators

wasted

Time \( t \)

Space \( x \)
With Directional Sparsity functional

\[ u \neq 0 \]
With Directional Sparsity functional

\[ u \neq 0 \]

Location of actuators

Time \( t \)

Space \( x \)
Directional Sparsity: Basic Definition

Problem formulation

\[
\begin{align*}
\min & \quad \frac{1}{2} \| Su - y_d \|_H^2 + \frac{\alpha}{2} \| u \|_{L^2(\Omega)}^2 \\
& \quad + \beta \| u \|_{L^1(L^2)} \\
\text{s.t.} & \quad u_a \leq u \leq u_b \quad \text{a.e. in } \Omega
\end{align*}
\]
Directional Sparsity: Basic Definition

Problem formulation

\[ \min \frac{1}{2} \|Su - y_d\|^2_H + \frac{\alpha}{2} \|u\|^2_{L^2(\Omega)} \]
\[ + \beta \int_{\Omega_1} \left( \int_{\Omega_2(x_1)} u(x_1, x_2)^2 \, dx_2 \right)^{1/2} \, dx_1 \]
\[ \text{s.t. } u_a \leq u \leq u_b \quad \text{a.e. in } \Omega \]
Related Approaches

Joint sparsity in image restoration

\[ \Psi(u) = \sum_{\lambda \in \Lambda} \omega_{\lambda} |\tilde{u}_{\lambda}|^{p}_{q}, \quad q = 2, p = 1 \]

- \( \Omega_1 \eqqcolon \Lambda \)
- \( \Omega_2 = \{1, 2, \ldots, \# \text{ of channels}\} \)
- with \( dx_2 = \text{counting measure} \)

[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

\[
\begin{align*}
\text{minimize} \quad & \frac{1}{2} \| y - y_d \|_{L^2(\Omega)}^2 + \beta \| u \|_{L^1(L^2)} \\
\text{s.t.} \quad & \begin{cases} 
    y_t - \frac{1}{10} \Delta y = u & \text{in } \Omega = \Omega_1 \times (0, T) \\
    y = 0 & \text{on } \Gamma \times (0, T) \\
    y(\cdot, 0) = 0 & \text{in } \Omega_1 \\
\end{cases} \\
\text{and} \quad & u_a \leq u \leq u_b \quad \text{a.e. in } \Omega
\end{align*}
\]

\[ n = 2 \text{ sparse directions (space)} \]

\[ N - n = 1 \text{ non-sparse direction (time)} \]
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
E. Casas, R. Herzog, and G. Wachsmuth.
Optimality conditions and error analysis of semilinear elliptic control problems with $L^1$ cost functional.

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.
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.
G. Stadler.
Elliptic optimal control problems with $L^1$-control cost and applications for the placement of control devices.

G. Vossen and H. Maurer.
On $L^1$-minimization in optimal control and applications to robotics.