Analysis of Adaptive Finite Elements for Control Constrained Optimal Control Problems

Kunibert G. Siebert

Fakutät für Mathematik

UNIVERSITÄT DUISBURG ESSEN

joint with Kristina Kohls and Arnd Rösch

Partially supported by DFG SPP 1253

A posteriori error estimates and mesh adaptivity for evolutionary and nonlinear problems, Paris, July 2010

Outline

Outline

Distributed Control Constrained Optimal Control Problem

Control Constrained Optimal Control Problem (P)

Solve the constrained minimization problem

$$\min_{u \in \mathbb{U}^{\text{ad}}} \psi(y) + \frac{\alpha}{2}\|u\|_{\mathbb{U}}^2 \quad \text{subject to} \quad \mathcal{B}[y,\, w] = \langle u,\, w\rangle_{\mathbb{U}} \quad \forall w \in \mathbb{Y},$$

where

- $\mathcal{B}$ is a bilinear form that arises from the weak formulation of a PDE over the Hilbert space $\mathbb{Y}$;
- $\mathbb{U}^{\text{ad}} \subset \mathbb{U}$ is a set of admissible controls;
- ψ is some objective functional for the state y;
- $\alpha > 0$ is a cost parameter for the control.

We next state assumptions that give existence and uniqueness of a solution (u, y) to (P) ...

Variational Problems

State Space and Bilinear Form

Let $\mathbb{Y}$ be an L^2-based Hilbert-space for the state and $\mathcal{B}\colon \mathbb{Y}\times\mathbb{Y}\to\mathbb{R}$ be a continuous bilinear form satisfying the inf-sup condition

$$\inf_{\substack{v\in\mathbb{Y}\\ \|v\|_{\mathbb{Y}}=1}} \sup_{\substack{w\in\mathbb{Y}\\ \|w\|_{\mathbb{Y}}=1}} \mathcal{B}[v,\,w] = \inf_{\substack{w\in\mathbb{Y}\\ \|w\|_{\mathbb{Y}}=1}} \sup_{\substack{v\in\mathbb{Y}\\ \|v\|_{\mathbb{Y}}=1}} \mathcal{B}[v,\,w] = \gamma > 0.$$

Theorem (Nečas (1962))

The inf-sup condition is equivalent to solvability of the variational problems

$$\hat{y}\in\mathbb{Y}: \qquad \mathcal{B}[\hat{y},\,w] = \langle f,\,w\rangle \qquad \forall\, w\in\mathbb{Y}$$

and

$$\hat{p}\in\mathbb{Y}: \qquad \mathcal{B}[v,\,\hat{p}] = \langle g,\,v\rangle \qquad \forall\, v\in\mathbb{Y}$$

for any $f,g\in\mathbb{Y}^*$.

Variational Problems

State Space and Bilinear Form

Let $\mathbb{Y}$ be an L^2-based Hilbert-space for the state and $\mathcal{B}\colon \mathbb{Y} \times \mathbb{Y} \to \mathbb{R}$ be a continuous bilinear form satisfying the inf-sup condition

$$\inf_{\substack{v\in\mathbb{Y}\\ \|v\|_{\mathbb{Y}}=1}} \sup_{\substack{w\in\mathbb{Y}\\ \|w\|_{\mathbb{Y}}=1}} \mathcal{B}[v,\, w] = \inf_{\substack{w\in\mathbb{Y}\\ \|w\|_{\mathbb{Y}}=1}} \sup_{\substack{v\in\mathbb{Y}\\ \|v\|_{\mathbb{Y}}=1}} \mathcal{B}[v,\, w] = \gamma > 0.$$

Theorem (Nečas (1962))

The inf-sup condition is equivalent to solvability of the variational problems

$$\hat{y} \in \mathbb{Y}: \qquad \mathcal{B}[\hat{y},\, w] = \langle f,\, w\rangle \qquad \forall\, w \in \mathbb{Y}$$

and

$$\hat{p} \in \mathbb{Y}: \qquad \mathcal{B}[v,\, \hat{p}] = \langle g,\, v\rangle \qquad \forall\, v \in \mathbb{Y}$$

for any $f, g \in \mathbb{Y}^*$.

Examples I

Example (Poisson Problem with Dirichlet Boundary Condition)

The variational formulation of

$$-\Delta y = u \quad \text{in } \Omega, \qquad y = 0 \quad \text{on } \partial\Omega$$

utilizes on $\mathbb{Y} = H_0^1(\Omega)$ the coercive and continuous bilinear form

$$\mathcal{B}[v,\, w] = \int_\Omega \nabla v \cdot \nabla w \, dV.$$

Example (Laplace Equation with Neumann Boundary Condition)

The variational formulation of

$$-\Delta y = 0 \quad \text{in } \Omega, \qquad \frac{\partial y}{\partial n} = u \quad \text{on } \partial\Omega$$

utilizes on $\mathbb{Y} = H^1(\Omega)/\mathbb{R}$ the coercive and continuous bilinear form

$$\mathcal{B}[v,\, w] = \int_\Omega \nabla v \cdot \nabla w \, dV.$$

Examples I

Example (Poisson Problem with Dirichlet Boundary Condition)

The variational formulation of

$$-\Delta y = u \quad \text{in } \Omega, \qquad y = 0 \quad \text{on } \partial\Omega$$

utilizes on $\mathbb{Y} = H_0^1(\Omega)$ the coercive and continuous bilinear form

$$\mathcal{B}[v,\, w] = \int_\Omega \nabla v \cdot \nabla w \, dV.$$

Example (Laplace Equation with Neumann Boundary Condition)

The variational formulation of

$$-\Delta y = 0 \quad \text{in } \Omega, \qquad \frac{\partial y}{\partial n} = u \quad \text{on } \partial\Omega$$

utilizes on $\mathbb{Y} = H^1(\Omega)/\mathbb{R}$ the coercive and continuous bilinear form

$$\mathcal{B}[v,\, w] = \int_\Omega \nabla v \cdot \nabla w \, dV.$$

Examples II

Example (The Stokes Problem with No-Slip Boundary Condition)

The variational formulation of

$$\begin{aligned} -\Delta \boldsymbol{v} + \nabla p &= \boldsymbol{u} && \text{in } \Omega, \\ \nabla \cdot \boldsymbol{v} &= 0 && \text{in } \Omega, \\ \boldsymbol{v} &= \boldsymbol{0} && \text{on } \partial\Omega \end{aligned}$$

utilizes with $\mathbb{Y} = H_0^1(\Omega; \mathbb{R}^d) \times L_0^2(\Omega)$ the bilinear form

$$\mathcal{B}[(\boldsymbol{v}, p),\, (\boldsymbol{w}, q)] := \int_\Omega \nabla \boldsymbol{w} : \nabla \boldsymbol{v}\, dV - \int_\Omega \nabla \cdot \boldsymbol{w}\, p\, dV - \int_\Omega q\, \nabla \cdot \boldsymbol{v}\, dV.$$

The continuous and symmetric bilinear form $\mathcal{B}$ satisfies an inf-sup condition.

Examples III

Example (The Stokes Problem with Slip Boundary Condition)

The variational form of

$$
\begin{aligned}
-\Delta \boldsymbol{v} + \nabla p &= \boldsymbol{f} && \text{in } \Omega,\\
\nabla \cdot \boldsymbol{v} &= 0 && \text{in } \Omega,\\
\boldsymbol{v} \cdot \boldsymbol{n} &= 0 && \text{on } \partial\Omega,\\
(2D(\boldsymbol{v}) - p)\boldsymbol{n} \cdot \boldsymbol{\tau_i} &= \boldsymbol{u} \cdot \boldsymbol{\tau_i}, \quad i = 1, \dots, d-1 && \text{on } \partial\Omega,
\end{aligned}
$$

with the symmetric gradient $D(\boldsymbol{v}) = \frac{1}{2}\left(\nabla \boldsymbol{v} + \nabla \boldsymbol{v}^T\right)$ is formulated in

$$\mathbb{Y} = \{\boldsymbol{w} \in H^1(\Omega; \mathbb{R}^d)/R \mid \boldsymbol{w} \cdot \boldsymbol{n} = 0 \text{ on } \partial\Omega\} \times L_0^2(\Omega),$$

where R is the set of rigid body motions of Ω. The corresponding bilinear form

$$\mathcal{B}[(\boldsymbol{v}, p),\, (\boldsymbol{w}, q)] := 2\int_\Omega D(\boldsymbol{w}) : D(\boldsymbol{v})\, dV - \int_\Omega \nabla \cdot \boldsymbol{w}\, p\, dV - \int_\Omega q\, \nabla \cdot \boldsymbol{v}\, dV$$

is continuous and symmetric and satisfies an inf-sup condition.

Solution Operators

Solution Operators

Denote by $S, S^* : \mathbb{Y}^* \to \mathbb{Y}$ the solution operator of the primal problem, i. e.,

$$\mathcal{B}[Sf,\, w] = \langle f,\, w\rangle \qquad \forall\, w \in \mathbb{Y}, \tag{LP}$$

respectively adjoint problem, i. e.,

$$\mathcal{B}[v,\, S^*g] = \langle g,\, v\rangle \qquad \forall\, v \in \mathbb{Y}. \tag{LP*}$$

Lemma (Invertibility and Boundedness)

The solution operators $S, S^* : \mathbb{Y}^* \to \mathbb{Y}$ are one-to-one and there holds

$$\|S\|_{L(\mathbb{Y}^*,\mathbb{Y})} = \|S^*\|_{L(\mathbb{Y}^*,\mathbb{Y})} = \gamma^{-1}.$$

Solution Operators

Solution Operators

Denote by $S, S^* \colon \mathbb{Y}^* \to \mathbb{Y}$ the solution operator of the primal problem, i. e.,

$$\mathcal{B}[Sf,\, w] = \langle f,\, w\rangle \qquad \forall\, w \in \mathbb{Y}, \tag{LP}$$

respectively adjoint problem, i. e.,

$$\mathcal{B}[v,\, S^*g] = \langle g,\, v\rangle \qquad \forall\, v \in \mathbb{Y}. \tag{LP*}$$

Lemma (Invertibility and Boundedness)

The solution operators $S, S^* \colon \mathbb{Y}^* \to \mathbb{Y}$ are one-to-one and there holds

$$\|S\|_{L(\mathbb{Y}^*,\mathbb{Y})} = \|S^*\|_{L(\mathbb{Y}^*,\mathbb{Y})} = \gamma^{-1}.$$

Control Space and Tichonov Functional

Control Space

Let $\mathbb{U}$ be another L^2-based Hilbert space such that

$$\mathbb{Y} \hookrightarrow \mathbb{U} \hookrightarrow \mathbb{Y}^*$$

in the sense that $v \in \mathbb{Y}$ implies $v \in \mathbb{U}$ with $\|v\|_{\mathbb{U}} \leq C\|v\|_{\mathbb{Y}}$ and $u \in \mathbb{U}$ implies $u \in \mathbb{Y}^*$ by

$$\langle u, v\rangle = \langle u, v\rangle_{\mathbb{Y}^* \times \mathbb{Y}} := \langle u, v\rangle_{\mathbb{U}} \qquad \forall\, v \in \mathbb{Y}.$$

Let $\emptyset \neq \mathbb{U}^{\mathsf{ad}} \subset \mathbb{U}$ be a convex set of admissible controls.

Tichonov Functional

We assume that the Tichonov functional $\psi\colon \mathbb{Y} \to \mathbb{R}$ for the state is quadratic and convex.

Control Space and Tichonov Functional

Control Space

Let $\mathbb{U}$ be another L^2-based Hilbert space such that

$$\mathbb{Y} \hookrightarrow \mathbb{U} \hookrightarrow \mathbb{Y}^*$$

in the sense that $v \in \mathbb{Y}$ implies $v \in \mathbb{U}$ with $\|v\|_{\mathbb{U}} \leq C\|v\|_{\mathbb{Y}}$ and $u \in \mathbb{U}$ implies $u \in \mathbb{Y}^*$ by

$$\langle u,\, v\rangle = \langle u,\, v\rangle_{\mathbb{Y}^* \times \mathbb{Y}} := \langle u,\, v\rangle_{\mathbb{U}} \qquad \forall\, v \in \mathbb{Y}.$$

Let $\emptyset \neq \mathbb{U}^{\mathsf{ad}} \subset \mathbb{U}$ be a convex set of admissible controls.

Tichonov Functional

We assume that the Tichonov functional $\psi\colon \mathbb{Y} \to \mathbb{R}$ for the state is quadratic and convex.

Examples of Control Spaces

Example (Distributed Control)

Typical example for distributed control is

$$\mathbb{U} = L^2(\Omega; \mathbb{R}^n)$$

for suitable $n \in \mathbb{N}$. For all the above examples we have

$$\mathbb{Y} \subset \mathbb{U} \subset \mathbb{Y}^*.$$

Example (Boundary Control)

Typical example for boundary control is

$$\mathbb{U} = L^2(\partial\Omega; \mathbb{R}^n)$$

for suitable $n \in \mathbb{N}$. For the above problems we have

$$\mathbb{Y} \hookrightarrow \mathbb{U} \hookrightarrow \mathbb{Y}^*$$

in the sense of traces.

Examples of Control Spaces

Example (Distributed Control)

Typical example for distributed control is

$$\mathbb{U} = L^2(\Omega; \mathbb{R}^n)$$

for suitable $n \in \mathbb{N}$. For all the above examples we have

$$\mathbb{Y} \subset \mathbb{U} \subset \mathbb{Y}^*.$$

Example (Boundary Control)

Typical example for boundary control is

$$\mathbb{U} = L^2(\partial\Omega; \mathbb{R}^n)$$

for suitable $n \in \mathbb{N}$. For the above problems we have

$$\mathbb{Y} \hookrightarrow \mathbb{U} \hookrightarrow \mathbb{Y}^*$$

in the sense of traces.

Examples of Admissible Controls

Box Constraints

For $a, b \in \mathbb{U}$ with $a \leq b$ let

$$\mathbb{U}^{\text{ad}} := \{u \in \mathbb{U} \mid a \leq u \leq b \text{ in } \Omega\}.$$

For $n \geq 1$ the constraint $a \leq u \leq b$ has to be read component-wise.

Norm Constraint

For $n \geq 2$ we may also consider for $r > 0$

$$\mathbb{U}^{\text{ad}} := \{u \in \mathbb{U} \mid |u|_2 \leq r \text{ in } \Omega\}.$$

Remark

A locally acting control in a subset $\Omega_0 \subset \Omega$ is included by

$$a, b \equiv 0 \qquad \text{in } \Omega \setminus \Omega_0,$$

or allowing for $r = r(x)$ and $r \equiv 0$ in $\Omega \setminus \Omega_0$.

Examples of Admissible Controls

Box Constraints

For $a, b \in \mathbb{U}$ with $a \le b$ let

$$\mathbb{U}^{\text{ad}} := \{u \in \mathbb{U} \mid a \le u \le b \text{ in } \Omega\}.$$

For $n \ge 1$ the constraint $a \le u \le b$ has to be read component-wise.

Norm Constraint

For $n \ge 2$ we may also consider for $r > 0$

$$\mathbb{U}^{\text{ad}} := \{u \in \mathbb{U} \mid |u|_2 \le r \text{ in } \Omega\}.$$

Remark

A locally acting control in a subset $\Omega_0 \subset \Omega$ is included by

$$a, b \equiv 0 \qquad \text{in } \Omega \setminus \Omega_0,$$

or allowing for $r = r(x)$ and $r \equiv 0$ in $\Omega \setminus \Omega_0$.

Examples of Admissible Controls

Box Constraints

For $a, b \in \mathbb{U}$ with $a \le b$ let

$$\mathbb{U}^{\text{ad}} := \{u \in \mathbb{U} \mid a \le u \le b \text{ in } \Omega\}.$$

For $n \ge 1$ the constraint $a \le u \le b$ has to be read component-wise.

Norm Constraint

For $n \ge 2$ we may also consider for $r > 0$

$$\mathbb{U}^{\text{ad}} := \{u \in \mathbb{U} \mid |u|_2 \le r \text{ in } \Omega\}.$$

Remark

A locally acting control in a subset $\Omega_0 \subset \Omega$ is included by

$$a, b \equiv 0 \qquad \text{in } \Omega \setminus \Omega_0,$$

or allowing for $r = r(x)$ and $r \equiv 0$ in $\Omega \setminus \Omega_0$.

Examples of Tichonov Functionals

Example (Desired L^2 Shape)

For $\Omega_d \subset \Omega$ and $z_d \in L^2(\Omega_d)$ set $\psi(z) := \frac{1}{2}\|z - z_d\|^2_{L^2(\Omega_d)}$.

Here, we have $\langle \psi'(z), v\rangle = \int_{\Omega_d} (z - z_d)\, v\, dV$.

Example (Desired Trace)

For $z_d \in L^2(\partial\Omega)$ set $\psi(z) := \frac{1}{2}\|z - z_d\|^2_{L^2(\partial\Omega)}$.

It holds $\langle \psi'(z), v\rangle = \int_{\partial\Omega} (z - z_d)\, v\, dA$.

Example (Desired H^1 Shape)

For $\Omega_d \subset \Omega$ and $z_d \in H^1(\Omega_d)$ set $\psi(z) := \frac{1}{2}\|\nabla(z - z_d)\|^2_{L^2(\Omega_d)}$.

Here, $\langle \psi'(z), v\rangle = \int_{\Omega_d} \nabla(z - z_d)\, \nabla v\, dV$.

Examples of Tichonov Functionals

Example (Desired L^2 Shape)

For $\Omega_d \subset \Omega$ and $z_d \in L^2(\Omega_d)$ set $\psi(z) := \frac{1}{2}\|z - z_d\|^2_{L^2(\Omega_d)}$.

Here, we have $\langle \psi'(z), v\rangle = \int_{\Omega_d} (z - z_d)\, v \, dV$.

Example (Desired Trace)

For $z_d \in L^2(\partial\Omega)$ set $\psi(z) := \frac{1}{2}\|z - z_d\|^2_{L^2(\partial\Omega)}$.

It holds $\langle \psi'(z), v\rangle = \int_{\partial\Omega} (z - z_d)\, v \, dA$.

Example (Desired H^1 Shape)

For $\Omega_d \subset \Omega$ and $z_d \in H^1(\Omega_d)$ set $\psi(z) := \frac{1}{2}\|\nabla(z - z_d)\|^2_{L^2(\Omega_d)}$.

Here, $\langle \psi'(z), v\rangle = \int_{\Omega_d} \nabla(z - z_d)\, \nabla v \, dV$.

Examples of Tichonov Functionals

Example (Desired L^2 Shape)

For $\Omega_d \subset \Omega$ and $z_d \in L^2(\Omega_d)$ set $\psi(z) := \frac{1}{2}\|z - z_d\|^2_{L^2(\Omega_d)}$.

Here, we have $\langle \psi'(z), v\rangle = \int_{\Omega_d} (z - z_d)\, v \, dV$.

Example (Desired Trace)

For $z_d \in L^2(\partial\Omega)$ set $\psi(z) := \frac{1}{2}\|z - z_d\|^2_{L^2(\partial\Omega)}$.

It holds $\langle \psi'(z), v\rangle = \int_{\partial\Omega} (z - z_d)\, v \, dA$.

Example (Desired H^1 Shape)

For $\Omega_d \subset \Omega$ and $z_d \in H^1(\Omega_d)$ set $\psi(z) := \frac{1}{2}\|\nabla(z - z_d)\|^2_{L^2(\Omega_d)}$.

Here, $\langle \psi'(z), v\rangle = \int_{\Omega_d} \nabla(z - z_d)\, \nabla v \, dV$.

Distributed Control Constrained Optimal Control Problem

Control Constrained Optimal Control Problem (P)

Solve the constrained minimization problem

$$\min_{u \in \mathbb{U}^{\text{ad}}} \psi(y) + \frac{\alpha}{2}\|u\|_{\mathbb{U}}^2 \quad \text{subject to} \quad \mathcal{B}[y,\, w] = \langle u,\, w\rangle_{\mathbb{U}} \quad \forall w \in \mathbb{Y}.$$

or, equivalently,

$$\min_{u \in \mathbb{U}^{\text{ad}}} \frac{1}{2}\psi(Su) + \frac{\alpha}{2}\|u\|_{\mathbb{U}}^2.$$

Theorem (Existence and Uniqueness)

There exists a unique pair $(\hat{y}, \hat{u}) \in \mathbb{Y} \times \mathbb{U}^{ad}$ *solving (P).*

Compare for instance with [Lions] or [Tröltzsch].

First Order Optimality Systems

First Order Optimality System

A solution $(\hat{y}, \hat{u}) \in \mathbb{Y} \times \mathbb{U}^{\text{ad}}$ is characterized by

$$\begin{aligned} &\text{state equation:} & \mathcal{B}[\hat{y},\, w] &= \langle \hat{u},\, w \rangle_{\mathbb{U}} & \forall\, w \in \mathbb{Y},\\ &\text{adjoint equation:} & \mathcal{B}[v,\, \hat{p}] &= \langle \psi'(\hat{y}),\, v \rangle & \forall\, v \in \mathbb{Y},\\ &\text{gradient equation:} & \langle \alpha\hat{u} + \hat{p},\, \hat{u} - u \rangle_{\mathbb{U}} &\le 0 & \forall\, u \in \mathbb{U}^{\text{ad}}, \end{aligned}$$

where $\hat{p} \in \mathbb{Y}$ is the adjoint state.

The gradient equation characterizes $\hat{u} = \Pi(\hat{p})$ as the best approximation of $-\frac{1}{\alpha}\hat{p}$ in $\mathbb{U}^{\text{ad}}$.

Reduced First Order Optimality System

The pair $(\hat{y}, \hat{p})$ is a solution to the reduced first order system

$$\begin{aligned} \mathcal{B}[\hat{y},\, w] &= \langle \Pi(\hat{p}),\, w \rangle_{\mathbb{U}} & \forall\, w \in \mathbb{Y},\\ \mathcal{B}[v,\, \hat{p}] &= \langle \psi'(\hat{y}),\, v \rangle & \forall\, v \in \mathbb{Y}. \end{aligned}$$

First Order Optimality Systems

First Order Optimality System

A solution $(\hat{y}, \hat{u}) \in \mathbb{Y} \times \mathbb{U}^{\text{ad}}$ is characterized by

$$\begin{aligned}
&\text{state equation:} && \mathcal{B}[\hat{y},\, w] = \langle \hat{u},\, w\rangle_{\mathbb{U}} && \forall\, w \in \mathbb{Y},\\
&\text{adjoint equation:} && \mathcal{B}[v,\, \hat{p}] = \langle \psi'(\hat{y}),\, v\rangle && \forall\, v \in \mathbb{Y},\\
&\text{gradient equation:} && \langle \alpha\hat{u} + \hat{p},\, \hat{u} - u\rangle_{\mathbb{U}} \le 0 && \forall\, u \in \mathbb{U}^{\text{ad}},
\end{aligned}$$

where $\hat{p} \in \mathbb{Y}$ is the adjoint state.

The gradient equation characterizes $\hat{u} = \Pi(\hat{p})$ as the best approximation of $-\frac{1}{\alpha}\hat{p}$ in $\mathbb{U}^{\text{ad}}$.

Reduced First Order Optimality System

The pair $(\hat{y}, \hat{p})$ is a solution to the reduced first order system

$$\begin{aligned}
\mathcal{B}[\hat{y},\, w] &= \langle \Pi(\hat{p}),\, w\rangle_{\mathbb{U}} && \forall\, w \in \mathbb{Y},\\
\mathcal{B}[v,\, \hat{p}] &= \langle \psi'(\hat{y}),\, v\rangle && \forall\, v \in \mathbb{Y}.
\end{aligned}$$

Outline

Conforming Discretization with Adaptive Finite Elements

Adaptive Finite Element Discretization

- Let $\mathcal{T}_0$ be an initial, conforming triangulation of Ω and let $\mathbb{T}$ be the set of all conforming refinements of $\mathcal{T}_0$ created by bisection.
- Let $\mathbb{Y}(\mathcal{T}) \subset \mathbb{Y}$ and $\mathbb{U}(\mathcal{T}) \subset \mathbb{U}$ be finite element spaces over $\mathcal{T}$.

$\mathcal{T}_0$ $\qquad$ $\mathcal{T} \in \mathbb{T}$

Nesting of Spaces

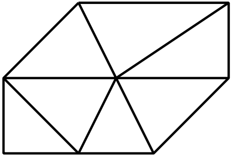
Let $\mathcal{T}_* \subset \mathbb{T}$ be a refinement of $\mathcal{T} \in \mathbb{T}$. If $\mathbb{Y}(\mathcal{T})$ and $\mathbb{U}(\mathcal{T})$ are built from piecewise polynomials then the finite element spaces are nested, i. e.,

$$\mathbb{Y}(\mathcal{T}) \subset \mathbb{Y}(\mathcal{T}_*) \qquad \text{and} \qquad \mathbb{U}(\mathcal{T}) \subset \mathbb{U}(\mathcal{T}_*).$$

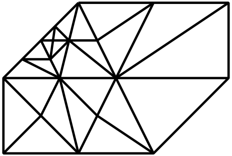
This property is important when analyzing a sequence of discrete solutions.

Adaptive Finite Element Discretization

- Let $\mathcal{T}_0$ be an initial, conforming triangulation of Ω and let $\mathbb{T}$ be the set of all conforming refinements of $\mathcal{T}_0$ created by bisection.
- Let $\mathbb{Y}(\mathcal{T}) \subset \mathbb{Y}$ and $\mathbb{U}(\mathcal{T}) \subset \mathbb{U}$ be finite element spaces over $\mathcal{T}$.



$\mathcal{T}_0$



$\mathcal{T} \in \mathbb{T}$

Nesting of Spaces

Let $\mathcal{T}_* \subset \mathbb{T}$ be a refinement of $\mathcal{T} \in \mathbb{T}$. If $\mathbb{Y}(\mathcal{T})$ and $\mathbb{U}(\mathcal{T})$ are built from piecewise polynomials then the finite element spaces are nested, i. e.,

$$\mathbb{Y}(\mathcal{T}) \subset \mathbb{Y}(\mathcal{T}_*) \qquad \text{and} \qquad \mathbb{U}(\mathcal{T}) \subset \mathbb{U}(\mathcal{T}_*).$$

This property is important when analyzing a sequence of discrete solutions.

The Discrete Inf-Sup Condition

Discrete Inf-Sup Condition

We assume that there holds

$$\inf_{\substack{V\in\mathbb{Y}(\mathcal{T})\\ \|V\|_{\mathbb{Y}}=1}} \sup_{\substack{W\in\mathbb{Y}(\mathcal{T})\\ \|W\|_{\mathbb{Y}}=1}} \mathcal{B}[V,\,W] = \inf_{\substack{W\in\mathbb{Y}(\mathcal{T})\\ \|W\|_{\mathbb{Y}}=1}} \sup_{\substack{V\in\mathbb{Y}(\mathcal{T})\\ \|V\|_{\mathbb{Y}}=1}} \mathcal{B}[V,\,W] = \gamma(\mathcal{T}) > 0.$$

This is equivalent to solvability for any $F, G \in \mathbb{Y}(\mathcal{T})^*$ of the discrete problems

$$\begin{aligned} \mathcal{B}[\hat{Y},\,W] &= \langle F,\,W\rangle \qquad \forall\, W \in \mathbb{Y}(\mathcal{T}),\\ \mathcal{B}[V,\,\hat{P}] &= \langle G,\,V\rangle \qquad \forall\, V \in \mathbb{Y}(\mathcal{T}). \end{aligned}$$

Lemma (Invertibility and Boundedness)

The discrete solution operators $S_{\mathcal{T}}, S_{\mathcal{T}}^*\colon \mathbb{Y}^*(\mathcal{T}) \to \mathbb{Y}(\mathcal{T})$ are one-to-one and there holds

$$\|S_{\mathcal{T}}\|_{L(\mathbb{Y}^*(\mathcal{T}),\mathbb{Y}(\mathcal{T}))} = \|S_{\mathcal{T}}^*\|_{L(\mathbb{Y}^*(\mathcal{T}),\mathbb{Y}(\mathcal{T}))} = \gamma(\mathcal{T})^{-1}.$$

Stable discretization, i. e., $\gamma(\mathcal{T}) \geq \underline{\gamma} > 0$, are important for uniform bounds.

The Discrete Inf-Sup Condition

Discrete Inf-Sup Condition

We assume that there holds

$$\inf_{\substack{V\in\mathbb{Y}(\mathcal{T})\\ \|V\|_{\mathbb{Y}}=1}} \sup_{\substack{W\in\mathbb{Y}(\mathcal{T})\\ \|W\|_{\mathbb{Y}}=1}} \mathcal{B}[V,\,W] = \inf_{\substack{W\in\mathbb{Y}(\mathcal{T})\\ \|W\|_{\mathbb{Y}}=1}} \sup_{\substack{V\in\mathbb{Y}(\mathcal{T})\\ \|V\|_{\mathbb{Y}}=1}} \mathcal{B}[V,\,W] = \gamma(\mathcal{T}) > 0.$$

This is equivalent to solvability for any $F, G \in \mathbb{Y}(\mathcal{T})^*$ of the discrete problems

$$\begin{aligned} \mathcal{B}[\hat{Y},\,W] &= \langle F,\,W\rangle \qquad \forall\, W \in \mathbb{Y}(\mathcal{T}),\\ \mathcal{B}[V,\,\hat{P}] &= \langle G,\,V\rangle \qquad \forall\, V \in \mathbb{Y}(\mathcal{T}). \end{aligned}$$

Lemma (Invertibility and Boundedness)

The discrete solution operators $S_{\mathcal{T}}, S_{\mathcal{T}}^*\colon \mathbb{Y}^*(\mathcal{T}) \to \mathbb{Y}(\mathcal{T})$ are one-to-one and there holds

$$\|S_{\mathcal{T}}\|_{L(\mathbb{Y}^*(\mathcal{T}),\mathbb{Y}(\mathcal{T}))} = \|S_{\mathcal{T}}^*\|_{L(\mathbb{Y}^*(\mathcal{T}),\mathbb{Y}(\mathcal{T}))} = \gamma(\mathcal{T})^{-1}.$$

Stable discretization, i. e., $\gamma(\mathcal{T}) \geq \underline{\gamma} > 0$, are important for uniform bounds.

Discretization I: Discretized Control

Discretized First Order Optimality System

Let $(\hat{Y}, \hat{P}, \hat{U}) \in \mathbb{Y}(\mathcal{T}) \times \mathbb{Y}(\mathcal{T}) \times \mathbb{U}^{\mathsf{ad}}(\mathcal{T})$ be the solution to

$$\begin{aligned} \mathcal{B}[\hat{Y}, W] &= \langle \hat{U}, W \rangle_{\mathbb{U}} && \forall\, W \in \mathbb{Y}(\mathcal{T}), \\ \mathcal{B}[V, \hat{P}] &= \langle \psi'(\hat{Y}), V \rangle && \forall\, V \in \mathbb{Y}(\mathcal{T}), \\ \langle \alpha \hat{U} + \hat{P}, \hat{U} - U \rangle_{\mathbb{U}} &\le 0 && \forall\, U \in \mathbb{U}^{\mathsf{ad}}(\mathcal{T}), \end{aligned}$$

where $\mathbb{U}^{\mathsf{ad}}(\mathcal{T}) := \mathbb{U}^{\mathsf{ad}} \cap \mathbb{U}(\mathcal{T})$ is the set of discrete admissible controls.

Remarks

1. $\mathbb{U}^{\mathsf{ad}}(\mathcal{T})$ is non-empty if, for instance, $a, b \in \mathbb{U}(\mathcal{T})$ in case of box constraints.
2. The problem can be solved by SSN methods; compare for instance with [Hintermüller, Ito, Kunisch].
3. The discrete control $\hat{U} = \Pi_{\mathcal{T}}(\hat{P})$ is the best approximation of $-\frac{1}{\alpha}\hat{P}$ in $\mathbb{U}^{\mathsf{ad}}(\mathcal{T})$.

Discretization I: Discretized Control

Discretized First Order Optimality System

Let $(\hat{Y}, \hat{P}, \hat{U}) \in \mathbb{Y}(\mathcal{T}) \times \mathbb{Y}(\mathcal{T}) \times \mathbb{U}^{\mathsf{ad}}(\mathcal{T})$ be the solution to

$$\begin{aligned} \mathcal{B}[\hat{Y}, W] &= \langle \hat{U}, W \rangle_{\mathbb{U}} && \forall\, W \in \mathbb{Y}(\mathcal{T}), \\ \mathcal{B}[V, \hat{P}] &= \langle \psi'(\hat{Y}), V \rangle && \forall\, V \in \mathbb{Y}(\mathcal{T}), \\ \langle \alpha \hat{U} + \hat{P}, \hat{U} - U \rangle_{\mathbb{U}} &\le 0 && \forall\, U \in \mathbb{U}^{\mathsf{ad}}(\mathcal{T}), \end{aligned}$$

where $\mathbb{U}^{\mathsf{ad}}(\mathcal{T}) := \mathbb{U}^{\mathsf{ad}} \cap \mathbb{U}(\mathcal{T})$ is the set of discrete admissible controls.

Remarks

1. $\mathbb{U}^{\mathsf{ad}}(\mathcal{T})$ is non-empty if, for instance, $a, b \in \mathbb{U}(\mathcal{T})$ in case of box constraints.
2. The problem can be solved by SSN methods; compare for instance with [Hintermüller, Ito, Kunisch].
3. The discrete control $\hat{U} = \Pi_{\mathcal{T}}(\hat{P})$ is the best approximation of $-\frac{1}{\alpha}\hat{P}$ in $\mathbb{U}^{\mathsf{ad}}(\mathcal{T})$.

Discrete Problem II: Non-Discretized Control

Discretized Reduced First Order Optimality System

Determine $(\hat{Y}, \hat{P}) \in \mathbb{Y}(\mathcal{T}) \times \mathbb{Y}(\mathcal{T})$ as the solution to the discretized reduced first order system

$$\begin{aligned} \mathcal{B}[\hat{Y}, W] &= \langle \Pi(\hat{P}), W \rangle_{\mathbb{U}} \qquad && \forall\, W \in \mathbb{Y}(\mathcal{T}), \\ \mathcal{B}[V, \hat{P}] &= \langle \psi'(\hat{Y}), V \rangle && \forall\, V \in \mathbb{Y}(\mathcal{T}). \end{aligned}$$

Remarks

1. This approach avoids to discretize the control space $\mathbb{U}$.
2. The discrete problem can be solved assuming that $\Pi(V)$ can be computed for discrete functions [Hinze].
3. The corresponding control $\hat{U} = \Pi(\hat{P})$ is the best approximation of $-\frac{1}{\alpha}\hat{P}$ in $\mathbb{U}^{\mathrm{ad}}$ and is in general not a discrete function.

In this talk we focus on non-discretized control. The results of are also valid for discretized control (up to modifications) provided that

$$\Pi_{\mathcal{T}} = \Pi_{|\mathbb{U}(\mathcal{T})}.$$

Discrete Problem II: Non-Discretized Control

Discretized Reduced First Order Optimality System

Determine $(\hat{Y}, \hat{P}) \in \mathbb{Y}(\mathcal{T}) \times \mathbb{Y}(\mathcal{T})$ as the solution to the discretized reduced first order system

$$\mathcal{B}[\hat{Y}, W] = \langle \Pi(\hat{P}), W \rangle_{\mathbb{U}} \qquad \forall\, W \in \mathbb{Y}(\mathcal{T}),$$
$$\mathcal{B}[V, \hat{P}] = \langle \psi'(\hat{Y}), V \rangle \qquad \forall\, V \in \mathbb{Y}(\mathcal{T}).$$

Remarks

1. This approach avoids to discretize the control space $\mathbb{U}$.
2. The discrete problem can be solved assuming that $\Pi(V)$ can be computed for discrete functions [Hinze].
3. The corresponding control $\hat{U} = \Pi(\hat{P})$ is the best approximation of $-\frac{1}{\alpha}\hat{P}$ in $\mathbb{U}^{\text{ad}}$ and is in general not a discrete function.

In this talk we focus on non-discretized control. The results of are also valid for discretized control (up to modifications) provided that

$$\Pi_{\mathcal{T}} = \Pi_{|\mathbb{U}(\mathcal{T})}.$$

Discrete Problem II: Non-Discretized Control

Discretized Reduced First Order Optimality System

Determine $(\hat{Y}, \hat{P}) \in \mathbb{Y}(\mathcal{T}) \times \mathbb{Y}(\mathcal{T})$ as the solution to the discretized reduced first order system

$$\begin{aligned} \mathcal{B}[\hat{Y}, W] &= \langle \Pi(\hat{P}), W \rangle_{\mathbb{U}} \qquad && \forall\, W \in \mathbb{Y}(\mathcal{T}), \\ \mathcal{B}[V, \hat{P}] &= \langle \psi'(\hat{Y}), V \rangle && \forall\, V \in \mathbb{Y}(\mathcal{T}). \end{aligned}$$

Remarks

1. This approach avoids to discretize the control space $\mathbb{U}$.
2. The discrete problem can be solved assuming that $\Pi(V)$ can be computed for discrete functions [Hinze].
3. The corresponding control $\hat{U} = \Pi(\hat{P})$ is the best approximation of $-\frac{1}{\alpha}\hat{P}$ in $\mathbb{U}^{\text{ad}}$ and is in general not a discrete function.

In this talk we focus on non-discretized control. The results of are also valid for discretized control (up to modifications) provided that

$$\Pi_{\mathcal{T}} = \Pi_{|\mathbb{U}(\mathcal{T})}.$$

The Adaptive Iteration

Start with a conforming triangulation $\mathcal{T}_0$ of Ω and iterate

SOLVE $\longrightarrow$ ESTIMATE $\longrightarrow$ MARK $\longrightarrow$ REFINE

Modules of the Adaptive Iteration

- SOLVE: solve the discretized reduced first order optimality system in $\mathbb{Y}(\mathcal{T})$;
- ESTIMATE: compute an a posteriori error estimator build from error indicators $\{\mathcal{E}_{\mathcal{T}}(T)\}_{T\in\mathcal{T}}$;
- MARK: collect in $\mathcal{M}$ elements of $\mathcal{T}$ subject to refinement;
- REFINE: refine at least all elements $\mathcal{M}$ and output a conforming refinement $\mathcal{T}_*$ of $\mathcal{T}$.

State of the Art in the A Posteriori Error Analysis

Linear PDEs:

Starting in the late 1970s, an exhaustive analysis was developed for different kinds of estimators, various problem classes, and norms yielding reliable and efficient estimators.

Optimal Control

Much less results. Liu and Yan started the a posteriori error analysis at the beginning of this century for the residual estimator.

Then contributions by [Liu, Yan], [Hintermüller, Hoppe], [Hinze, Yan, Zhou], [Li, Liu, Yan], [Yan, Zhou], ...
Dual Weighted Residual Indicators by [Becker, Rannacher], [Günther, Hinze], [Vexler, Wollner], ...

Main Drawback of the Analysis for Optimal Control

Well-known results for linear problems are not directly utilized.

State of the Art in the A Posteriori Error Analysis

Linear PDEs:

Starting in the late 1970s, an exhaustive analysis was developed for different kinds of estimators, various problem classes, and norms yielding reliable and efficient estimators.

Optimal Control

Much less results. Liu and Yan started the a posteriori error analysis at the beginning of this century for the residual estimator.

Then contributions by [Liu, Yan], [Hintermüller, Hoppe], [Hinze, Yan, Zhou], [Li, Liu, Yan], [Yan, Zhou], . . .
Dual Weighted Residual Indicators by [Becker, Rannacher], [Günther, Hinze], [Vexler, Wollner], . . .

Main Drawback of the Analysis for Optimal Control

Well-known results for linear problems are not directly utilized.

State of the Art in the A Posteriori Error Analysis

Linear PDEs:

Starting in the late 1970s, an exhaustive analysis was developed for different kinds of estimators, various problem classes, and norms yielding reliable and efficient estimators.

Optimal Control

Much less results. Liu and Yan started the a posteriori error analysis at the beginning of this century for the residual estimator.

Then contributions by [Liu, Yan], [Hintermüller, Hoppe], [Hinze, Yan, Zhou], [Li, Liu, Yan], [Yan, Zhou], . . .
Dual Weighted Residual Indicators by [Becker, Rannacher], [Günther, Hinze], [Vexler, Wollner], . . .

Main Drawback of the Analysis for Optimal Control

Well-known results for linear problems are not directly utilized.

State of the Art in the Convergence and Optimality Analysis

Linear PDEs:

Starting in the mid 1990s, the convergence analysis for linear elliptic PDEs is more or less settled: [Dörfler] [Morin, Nochetto, S.] [Morin, S., Veeser]

Optimal decay rates in terms of degrees of freedom can be shown for energy minimization: [Binev, Dahmen, DeVore] [Stevenson] [Cascón, Kreuzer, Nochetto, S.] [Kreuzer, Diening] [Kreuzer, S.]

Optimal Control:

The only preliminary convergence results are due to [Gaevskaya, Hoppe, Iliash, Kieweg] [Hintermüller, Hoppe].
Optimal error decay is completely open.

Main Drawback of the Analysis for Optimal Control

1. Plain convergence needs an assumption on smallness of the maximal mesh size.
2. Error reduction property needs an additional assumption on the closeness of true and discrete active sets.

State of the Art in the Convergence and Optimality Analysis

Linear PDEs:

Starting in the mid 1990s, the convergence analysis for linear elliptic PDEs is more or less settled: [Dörfler] [Morin, Nochetto, S.] [Morin, S., Veeser]

Optimal decay rates in terms of degrees of freedom can be shown for energy minimization: [Binev, Dahmen, DeVore] [Stevenson] [Cascón, Kreuzer, Nochetto, S.] [Kreuzer, Diening] [Kreuzer, S.]

Optimal Control:

The only preliminary convergence results are due to [Gaevskaya, Hoppe, Iliash, Kieweg] [Hintermüller, Hoppe].
Optimal error decay is completely open.

Main Drawback of the Analysis for Optimal Control

1. Plain convergence needs an assumption on smallness of the maximal mesh size.
2. Error reduction property needs an additional assumption on the closeness of true and discrete active sets.

State of the Art in the Convergence and Optimality Analysis

Linear PDEs:

Starting in the mid 1990s, the convergence analysis for linear elliptic PDEs is more or less settled: [Dörfler] [Morin, Nochetto, S.] [Morin, S., Veeser]

Optimal decay rates in terms of degrees of freedom can be shown for energy minimization: [Binev, Dahmen, DeVore] [Stevenson] [Cascón, Kreuzer, Nochetto, S.] [Kreuzer, Diening] [Kreuzer, S.]

Optimal Control:

The only preliminary convergence results are due to [Gaevskaya, Hoppe, Iliash, Kieweg] [Hintermüller, Hoppe].
Optimal error decay is completely open.

Main Drawback of the Analysis for Optimal Control

1. Plain convergence needs an assumption on smallness of the maximal mesh size.
2. Error reduction property needs an additional assumption on the closeness of true and discrete active sets.

Outline

Continuous and Discrete Problem

The continuous problem reads: solve for $(\hat{y}, \hat{p}, \hat{u}) \in \mathbb{Y} \times \mathbb{Y} \times \mathbb{U}^{\mathsf{ad}}$

$$
\begin{aligned}
\mathcal{B}[\hat{y},\, w] &= \langle \hat{u},\, w\rangle_{\mathbb{U}} && \forall\, w \in \mathbb{Y},\\
\mathcal{B}[v,\, \hat{p}] &= \langle \psi'(\hat{y}),\, v\rangle && \forall\, v \in \mathbb{Y},\\
\langle \alpha\hat{u} + \hat{p},\, \hat{u} - u\rangle_{\mathbb{U}} &\le 0 && \forall\, u \in \mathbb{U}^{\mathsf{ad}}.
\end{aligned}
$$

The discrete problem reads: solve for $(\hat{Y}, \hat{P}, \hat{U}) \in \mathbb{Y}(\mathcal{T}) \times \mathbb{Y}(\mathcal{T}) \times \mathbb{U}^{\mathsf{ad}}$

$$
\begin{aligned}
\mathcal{B}[\hat{Y},\, W] &= \langle \hat{U},\, W\rangle_{\mathbb{U}} && \forall\, W \in \mathbb{Y}(\mathcal{T}),\\
\mathcal{B}[V,\, \hat{P}] &= \langle \psi'(\hat{Y}),\, V\rangle && \forall\, V \in \mathbb{Y}(\mathcal{T}),\\
\langle \alpha\hat{U} + \hat{P},\, \hat{U} - u\rangle_{\mathbb{U}} &\le 0 && \forall\, u \in \mathbb{U}^{\mathsf{ad}}.
\end{aligned}
$$

Main Problem in the A Posteriori Error Analysis

The discrete state $\hat{Y}$ is not the Galerkin approximation to the continuous state $\hat{y} = S(\hat{u})$ but it is the Galerkin approximation to $\bar{y} = S(\hat{U})$.

Likewise, $\hat{P}$ is the Galerkin approximation to to $\bar{p} = S^*(\psi'(\hat{Y}))$ rather than to the continuous adjoint state $\hat{p} = S^*(\psi'(\hat{y}))$.

Continuous and Discrete Problem

The continuous problem reads: solve for $(\hat{y}, \hat{p}, \hat{u}) \in \mathbb{Y} \times \mathbb{Y} \times \mathbb{U}^{\text{ad}}$

$$\begin{aligned} \mathcal{B}[\hat{y},\, w] &= \langle \hat{u},\, w\rangle_{\mathbb{U}} && \forall\, w \in \mathbb{Y}, \\ \mathcal{B}[v,\, \hat{p}] &= \langle \psi'(\hat{y}),\, v\rangle && \forall\, v \in \mathbb{Y}, \\ \langle \alpha\hat{u} + \hat{p},\, \hat{u} - u\rangle_{\mathbb{U}} &\le 0 && \forall\, u \in \mathbb{U}^{\text{ad}}. \end{aligned}$$

The discrete problem reads: solve for $(\hat{Y}, \hat{P}, \hat{U}) \in \mathbb{Y}(\mathcal{T}) \times \mathbb{Y}(\mathcal{T}) \times \mathbb{U}^{\text{ad}}$

$$\begin{aligned} \mathcal{B}[\hat{Y},\, W] &= \langle \hat{U},\, W\rangle_{\mathbb{U}} && \forall\, W \in \mathbb{Y}(\mathcal{T}), \\ \mathcal{B}[V,\, \hat{P}] &= \langle \psi'(\hat{Y}),\, V\rangle && \forall\, V \in \mathbb{Y}(\mathcal{T}), \\ \langle \alpha\hat{U} + \hat{P},\, \hat{U} - u\rangle_{\mathbb{U}} &\le 0 && \forall\, u \in \mathbb{U}^{\text{ad}}. \end{aligned}$$

Main Problem in the A Posteriori Error Analysis

The discrete state $\hat{Y}$ is not the Galerkin approximation to the continuous state $\hat{y} = S(\hat{u})$ but it is the Galerkin approximation to $\bar{y} = S(\hat{U})$.

Likewise, $\hat{P}$ is the Galerkin approximation to to $\bar{p} = S^*(\psi'(\hat{Y}))$ rather than to the continuous adjoint state $\hat{p} = S^*(\psi'(\hat{y}))$.

Estimate for State and Adjoint State

Lemma

For $\hat{y} = S(\hat{u})$, $\hat{p} = S^*(\psi'(\hat{y}))$, $\bar{y} = S(\hat{U})$, *and* $\bar{p} = S^*(\psi'(\hat{Y}))$ *holds*

$$\|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}},$$
$$\|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}},$$

Proof. Use the continuous primal problem to derive

$$\|\bar{y} - \hat{y}\|_{\mathbb{Y}} = \|S(\hat{U} - \hat{u})\|_{\mathbb{Y}} \le \gamma^{-1}\|\hat{U} - \hat{u}\|_{\mathbb{Y}^*} \le \gamma^{-1}C\|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

Therefore,

$$\|\hat{Y} - \hat{y}\|_{\mathbb{Y}} \le \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\bar{y} - \hat{y}\|_{Y} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

Similarly, using the continuous adjoint problem we estimate

$$\|\hat{P} - \hat{p}\|_{\mathbb{Y}} \lesssim \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

For the second estimate simply exchange $\hat{y}$ with $\bar{y}$ and $\hat{p}$ with $\bar{p}$. □

Estimate for State and Adjoint State

Lemma

For $\hat{y} = S(\hat{u})$, $\hat{p} = S^*(\psi'(\hat{y}))$, $\bar{y} = S(\hat{U})$, *and* $\bar{p} = S^*(\psi'(\hat{Y}))$ *holds*

$$\|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}},$$
$$\|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}},$$

Proof. Use the continuous primal problem to derive

$$\|\bar{y} - \hat{y}\|_{\mathbb{Y}} = \|S(\hat{U} - \hat{u})\|_{\mathbb{Y}} \le \gamma^{-1}\|\hat{U} - \hat{u}\|_{\mathbb{Y}^*} \le \gamma^{-1}C\|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

Therefore,

$$\|\hat{Y} - \hat{y}\|_{\mathbb{Y}} \le \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\bar{y} - \hat{y}\|_{Y} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

Similarly, using the continuous adjoint problem we estimate

$$\|\hat{P} - \hat{p}\|_{\mathbb{Y}} \lesssim \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

For the second estimate simply exchange $\hat{y}$ with $\bar{y}$ and $\hat{p}$ with $\bar{p}$. □

Estimate for State and Adjoint State

Lemma

For $\hat{y} = S(\hat{u})$, $\hat{p} = S^*(\psi'(\hat{y}))$, $\bar{y} = S(\hat{U})$, *and* $\bar{p} = S^*(\psi'(\hat{Y}))$ *holds*

$$\|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}},$$

$$\|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}},$$

Proof. Use the continuous primal problem to derive

$$\|\bar{y} - \hat{y}\|_{\mathbb{Y}} = \|S(\hat{U} - \hat{u})\|_{\mathbb{Y}} \le \gamma^{-1}\|\hat{U} - \hat{u}\|_{\mathbb{Y}^*} \le \gamma^{-1}C\|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

Therefore,

$$\|\hat{Y} - \hat{y}\|_{\mathbb{Y}} \le \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\bar{y} - \hat{y}\|_{Y} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

Similarly, using the continuous adjoint problem we estimate

$$\|\hat{P} - \hat{p}\|_{\mathbb{Y}} \lesssim \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

For the second estimate simply exchange $\hat{y}$ with $\bar{y}$ and $\hat{p}$ with $\bar{p}$. □

Estimate for State and Adjoint State

Lemma

For $\hat{y} = S(\hat{u})$, $\hat{p} = S^*(\psi'(\hat{y}))$, $\bar{y} = S(\hat{U})$, *and* $\bar{p} = S^*(\psi'(\hat{Y}))$ *holds*

$$\|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}},$$

$$\|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} + \|\hat{P} - \hat{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}},$$

Proof. Use the continuous primal problem to derive

$$\|\bar{y} - \hat{y}\|_{\mathbb{Y}} = \|S(\hat{U} - \hat{u})\|_{\mathbb{Y}} \leq \gamma^{-1}\|\hat{U} - \hat{u}\|_{\mathbb{Y}^*} \leq \gamma^{-1}C\|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

Therefore,

$$\|\hat{Y} - \hat{y}\|_{\mathbb{Y}} \leq \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\bar{y} - \hat{y}\|_{Y} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

Similarly, using the continuous adjoint problem we estimate

$$\|\hat{P} - \hat{p}\|_{\mathbb{Y}} \lesssim \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{Y} - \hat{y}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{U} - \hat{u}\|_{\mathbb{U}}.$$

For the second estimate simply exchange $\hat{y}$ with $\bar{y}$ and $\hat{p}$ with $\bar{p}$. □

Estimate for the Control

Lemma

There holds

$$\|\hat{U} - \hat{u}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}}.$$

Main Idea of the Proof. First estimate

$$\begin{aligned}\alpha\|\hat{U} - \hat{u}\|_{\mathbb{U}}^2 &= \langle \alpha\hat{U} + \hat{P}, \hat{U} - \hat{u}\rangle_{\mathbb{U}} + \langle \alpha\hat{u} + \hat{p}, \hat{u} - \hat{U}\rangle_{\mathbb{U}} + \langle \hat{p} - \hat{P}, \hat{U} - \hat{u}\rangle_{\mathbb{U}} \\ &\le \langle \hat{p} - \hat{P}, \hat{U} - \hat{u}\rangle_{\mathbb{U}},\end{aligned}$$

by the discrete and continuous gradient equations. Then estimate

$$\langle \hat{p} - \hat{P}, \hat{U} - \hat{u}\rangle_{\mathbb{U}} \lesssim \left(\|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{Y} - \bar{y}\|_{\mathbb{Y}}\right)\|\hat{U} - \hat{u}\|_{\mathbb{U}}$$

using the definitions of $\bar{y}$, $\bar{p}$, and the convexity of ψ. □

Estimate for the Control

Lemma

There holds

$$\|\hat{U} - \hat{u}\|_{\mathbb{Y}} \lesssim \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}}.$$

Main Idea of the Proof. First estimate

$$\begin{aligned}\alpha\|\hat{U} - \hat{u}\|_{\mathbb{U}}^2 &= \langle \alpha\hat{U} + \hat{P}, \hat{U} - \hat{u}\rangle_{\mathbb{U}} + \langle \alpha\hat{u} + \hat{p}, \hat{u} - \hat{U}\rangle_{\mathbb{U}} + \langle \hat{p} - \hat{P}, \hat{U} - \hat{u}\rangle_{\mathbb{U}} \\ &\le \langle \hat{p} - \hat{P}, \hat{U} - \hat{u}\rangle_{\mathbb{U}},\end{aligned}$$

by the discrete and continuous gradient equations. Then estimate

$$\langle \hat{p} - \hat{P}, \hat{U} - \hat{u}\rangle_{\mathbb{U}} \lesssim \left(\|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{Y} - \bar{y}\|_{\mathbb{Y}}\right)\|\hat{U} - \hat{u}\|_{\mathbb{U}}$$

using the definitions of $\bar{y}$, $\bar{p}$, and the convexity of ψ. □

Equivalence of Errors

Recall: The functions $\bar{y}$ and $\bar{p}$ are solutions of the linear problems

$$\bar{y} \in \mathbb{Y}: \qquad \mathcal{B}[\bar{y},\, w] = \langle \hat{U},\, w\rangle_{\mathbb{U}} \qquad \forall\, w \in \mathbb{Y}, \tag{LP}$$

$$\bar{p} \in \mathbb{Y}: \qquad \mathcal{B}[v,\, \bar{p}] = \langle \psi'(\hat{Y}),\, v\rangle \qquad \forall\, v \in \mathbb{Y}, \tag{LP*}$$

and $\hat{Y}, \hat{P} \in \mathbb{Y}(\mathcal{T})$ are their respective Galerkin approximations.

Proposition (Equivalence of Errors [Kohls, Rösch, S. '10])

The error of the optimal control problem is equivalent to the error in the corresponding linear problems, i. e.,

$$\|\hat{Y} - \hat{y},\, \hat{P} - \hat{p},\, \hat{U} - \hat{u}\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} \approx \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}}.$$

A posteriori error estimation for the optimal control problem reduces to error estimates for the linear sub-problems.

Equivalence of Errors

Recall: The functions $\bar{y}$ and $\bar{p}$ are solutions of the linear problems

$$\bar{y} \in \mathbb{Y}: \qquad \mathcal{B}[\bar{y}, w] = \langle \hat{U}, w \rangle_{\mathbb{U}} \qquad \forall\, w \in \mathbb{Y}, \qquad \text{(LP)}$$

$$\bar{p} \in \mathbb{Y}: \qquad \mathcal{B}[v, \bar{p}] = \langle \psi'(\hat{Y}), v \rangle \qquad \forall\, v \in \mathbb{Y}, \qquad \text{(LP}^*\text{)}$$

and $\hat{Y}, \hat{P} \in \mathbb{Y}(\mathcal{T})$ are their respective Galerkin approximations.

Proposition (Equivalence of Errors [Kohls, Rösch, S. '10])

The error of the optimal control problem is equivalent to the error in the corresponding linear problems, i. e.,

$$\|\hat{Y} - \hat{y},\, \hat{P} - \hat{p},\, \hat{U} - \hat{u}\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} \approx \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}}.$$

A posteriori error estimation for the optimal control problem reduces to error estimates for the linear sub-problems.

Equivalence of Errors

Recall: The functions $\bar{y}$ and $\bar{p}$ are solutions of the linear problems

$$\bar{y} \in \mathbb{Y}: \qquad \mathcal{B}[\bar{y},\, w] = \langle \hat{U},\, w\rangle_{\mathbb{U}} \qquad \forall\, w \in \mathbb{Y}, \qquad \text{(LP)}$$

$$\bar{p} \in \mathbb{Y}: \qquad \mathcal{B}[v,\, \bar{p}] = \langle \psi'(\hat{Y}),\, v\rangle \qquad \forall\, v \in \mathbb{Y}, \qquad \text{(LP}^*\text{)}$$

and $\hat{Y}, \hat{P} \in \mathbb{Y}(\mathcal{T})$ are their respective Galerkin approximations.

Proposition (Equivalence of Errors [Kohls, Rösch, S. '10])

The error of the optimal control problem is equivalent to the error in the corresponding linear problems, i. e.,

$$\|\hat{Y} - \hat{y},\, \hat{P} - \hat{p},\, \hat{U} - \hat{u}\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} \approx \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}}.$$

A posteriori error estimation for the optimal control problem reduces to error estimates for the linear sub-problems.

Remarks

Remark (A Priori or A Posteriori Character)

1. The result has an a priori flavor. But it is, philosophically speaking, a real a posteriori result in that only the continuous inf-sup constant enters but not the discrete one.

2. It cannot be directly used for an a priori estimate since $\bar{y}$ and $\bar{p}$ depend on $\mathcal{T}$, i. e., $\bar{y} = \bar{y}(\mathcal{T})$ and $\bar{p} = \bar{p}(\mathcal{T})$.

   A priori estimates for (LP) and (LP*) only transfer to optimal control problems having uniform bounds on higher derivatives of $\bar{y}$ and $\bar{p}$.

Remark (Discretized Control)

In case of discretized control $\hat{U} = \Pi_{\mathcal{T}}(\hat{P}) \in \mathbb{U}^{\mathrm{ad}}(\mathcal{T})$ we have the equivalence

$$\|\hat{Y} - \hat{y},\, \hat{P} - \hat{p},\, \hat{U} - \hat{u}\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} \approx \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{U} - \Pi(\hat{P})\|_{\mathbb{U}}.$$

Remarks

Remark (A Priori or A Posteriori Character)

1. The result has an a priori flavor. But it is, philosophically speaking, a real a posteriori result in that only the continuous inf-sup constant enters but not the discrete one.

2. It cannot be directly used for an a priori estimate since $\bar{y}$ and $\bar{p}$ depend on $\mathcal{T}$, i. e., $\bar{y} = \bar{y}(\mathcal{T})$ and $\bar{p} = \bar{p}(\mathcal{T})$.

   A priori estimates for (LP) and (LP*) only transfer to optimal control problems having uniform bounds on higher derivatives of $\bar{y}$ and $\bar{p}$.

Remark (Discretized Control)

In case of discretized control $\hat{U} = \Pi_{\mathcal{T}}(\hat{P}) \in \mathbb{U}^{\mathsf{ad}}(\mathcal{T})$ we have the equivalence

$$\|\hat{Y} - \hat{y},\, \hat{P} - \hat{p},\, \hat{U} - \hat{u}\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} \approx \|\hat{Y} - \bar{y}\|_{\mathbb{Y}} + \|\hat{P} - \bar{p}\|_{\mathbb{Y}} + \|\hat{U} - \Pi(\hat{P})\|_{\mathbb{U}}.$$

A Posteriori Error Bound for the Optimal Control Problem

Theorem (Reliability and Efficiency [Kohls, Rösch, S. '10])

Let $(\hat{y}, \hat{p}, \hat{u})$ *be the continuous solution and* $(\hat{Y}, \hat{P}, \hat{U})$ *be the discrete one. Let* $\mathcal{E}_{\mathcal{T}}(\hat{Y}, \hat{U}; \mathcal{T})$ *and* $\mathcal{E}^*_{\mathcal{T}}(\hat{P}, \psi'(\hat{Y}); \mathcal{T})$ *be reliable and efficient estimators for* (LP) *respectively* (LP*).

Then the sum of $\mathcal{E}_{\mathcal{T}}$ *and* $\mathcal{E}^*_{\mathcal{T}}$ *is a reliable and efficient estimator for the optimal control problem, i. e.,*

$$\|\hat{Y} - \hat{y},\, \hat{P} - \hat{p},\, \hat{U} - \hat{u}\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} \lesssim \mathcal{E}_{\mathcal{T}}(\hat{Y}, \hat{U}; \mathcal{T}) + \mathcal{E}^*_{\mathcal{T}}(\hat{P}, \psi'(\hat{Y}); \mathcal{T}) \lesssim \|\hat{Y} - \hat{y},\, \hat{P} - \hat{p},\, \hat{U} - \hat{u}\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} + \widetilde{\operatorname{osc}}_{\mathcal{T}}.$$

Remark (Approximability of the Control)

1. There is no explicit contribution in the indicators that measures $\|\hat{U} - \hat{u}\|_{\mathbb{U}}$.
2. However, in the lower bound there pops up an oscillation term $\|h_{\mathcal{T}}(P_{\mathcal{T}}\hat{u} - \hat{u})\|_{\mathbb{U}}$ that measures how good the control can be approximated on $\mathcal{T}$.

A Posteriori Error Bound for the Optimal Control Problem

Theorem (Reliability and Efficiency [Kohls, Rösch, S. '10])

Let $(\hat y, \hat p, \hat u)$ be the continuous solution and $(\hat Y, \hat P, \hat U)$ be the discrete one. Let $\mathcal{E}_{\mathcal{T}}(\hat Y, \hat U; \mathcal{T})$ and $\mathcal{E}^*_{\mathcal{T}}(\hat P, \psi'(\hat Y); \mathcal{T})$ be reliable and efficient estimators for (LP) *respectively* (LP*).

Then the sum of $\mathcal{E}_{\mathcal{T}}$ and $\mathcal{E}^*_{\mathcal{T}}$ is a reliable and efficient estimator for the optimal control problem, i. e.,

$$\|\hat Y - \hat y,\, \hat P - \hat p,\, \hat U - \hat u\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} \lesssim \mathcal{E}_{\mathcal{T}}(\hat Y, \hat U; \mathcal{T}) + \mathcal{E}^*_{\mathcal{T}}(\hat P, \psi'(\hat Y); \mathcal{T}) \lesssim \|\hat Y - \hat y,\, \hat P - \hat p,\, \hat U - \hat u\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} + \widetilde{\operatorname{osc}}_{\mathcal{T}}.$$

Remark (Approximability of the Control)

1. There is no explicit contribution in the indicators that measures $\|\hat U - \hat u\|_{\mathbb{U}}$.
2. However, in the lower bound there pops up an oscillation term $\|h_{\mathcal{T}}(P_{\mathcal{T}}\hat u - \hat u)\|_{\mathbb{U}}$ that measures how good the control can be approximated on $\mathcal{T}$.

Select any Estimator of Your Taste

Estimators for Linear Problems

1. Estimate the negative norm of the residual by a stronger but computable norm.
   - Leads to residual estimators, which are, for instance, built form scaled L^2 norms.
2. Evaluate the residual on an enriched but finite dimensional space.
   - Leads to hierarchical estimators.
3. Compute approximations to the Riesz representation of the residual in suitable local spaces.
   - Leads to local problems on enriched spaces.
4. Construct a suitable quantity that has better approximation properties than the discrete solution.
   - Leads to averaging techniques (ZZ-estimator).
5. Construct a suitable function in the setting of Prager-Synge.
   - Leads to equilibrated residual estimators.
6. ...

Example: Primal Problem with Distributed Control

The Poisson problem with distributed control reads

$$-\Delta \bar{y} = \hat{U} \quad \text{in } \Omega, \qquad \bar{y} = 0 \quad \text{on } \partial\Omega.$$

Residual Estimator

The local indicators are given by

$$\mathcal{E}_{\mathcal{T}}(\hat{Y}, \hat{U}; T)^2 = h_T^2 \| - \Delta \hat{Y} - \hat{U} \|_{L^2(T)}^2 + h_T \| [\![\nabla \hat{Y}]\!] \|_{L^2(\partial T \cap \Omega)}^2.$$

The estimator is given by

$$\mathcal{E}_{\mathcal{T}}(\hat{Y}, \hat{U}; \mathcal{T})^2 = \sum_{T \in \mathcal{T}} \mathcal{E}_{\mathcal{T}}(\hat{Y}, \hat{U}; T)^2.$$

Example: Primal Problem with Boundary Control

The Laplace equation with boundary control reads

$$-\Delta\bar{y} = 0 \quad \text{in } \Omega, \qquad \partial_{\boldsymbol{n}}\bar{y} = \hat{U} \quad \text{on } \partial\Omega.$$

Residual Estimator

The local indicators are given by

$$\mathcal{E}_{\mathcal{T}}(\hat{Y},\hat{U};T)^2 = h_T^2\|\Delta\hat{Y}\|^2_{L^2(T)} + h_T\|\,[\![\nabla\hat{Y}]\!]\,\|^2_{L^2(\partial T\cap\Omega)}, \\ + h_T\|\partial_{\boldsymbol{n}}\hat{Y} - \hat{U}\|^2_{L^2(\partial T\cap\partial\Omega)}$$

Example: Adjoint Problem for Desired L^2 State

The Poisson problem with distributed control reads

$$-\Delta \bar{p} = \hat{Y} - z_d \quad \text{in } \Omega, \qquad \bar{p} = 0 \quad \text{on } \partial\Omega.$$

Residual Estimator

The local indicators are given by

$$\mathcal{E}^*_{\mathcal{T}}(\hat{P}, \psi'(\hat{Y}); T)^2 = h_T^2 \| -\Delta \hat{P} - (\hat{Y} - z_d) \|^2_{L^2(T)} + h_T \| [\![\nabla \hat{P}]\!] \|^2_{L^2(\partial T \cap \Omega)}.$$

Example: Adjoint Problem for Desired H^1 State

The Poisson problem with distributed control reads

$$-\Delta \bar{p} = -\Delta(\hat{Y} - z_d) \quad \text{in } \Omega, \qquad \bar{p} = 0 \quad \text{on } \partial\Omega.$$

Residual Estimator

Assuming $z_d \in H^2(\Omega)$ the local indicators are given by

$$\mathcal{E}^*_{\mathcal{T}}(\hat{P}, \psi'(\hat{Y}); T)^2 = h_T^2 \|\Delta(\hat{P} - \hat{Y} + z_d)\|^2_{L^2(T)} + h_T \| [\![\nabla(\hat{P} - \hat{Y})]\!] \|^2_{L^2(\partial T \cap \Omega)}.$$

Remark

In case $z_d \notin H^2(\Omega)$ or z_d not piecewise H^2 over $\mathcal{T}$, the residual estimator cannot be efficient. In this case it is better to use one of the other estimators.

Example: Adjoint Problem for Desired H^1 State

The Poisson problem with distributed control reads

$$-\Delta\bar{p} = -\Delta(\hat{Y} - z_d) \quad \text{in } \Omega, \qquad \bar{p} = 0 \quad \text{on } \partial\Omega.$$

Residual Estimator

Assuming $z_d \in H^2(\Omega)$ the local indicators are given by

$$\mathcal{E}_{\mathcal{T}}^*(\hat{P}, \psi'(\hat{Y}); T)^2 = h_T^2 \|\Delta(\hat{P} - \hat{Y} + z_d)\|_{L^2(T)}^2 + h_T \| [\![\nabla(\hat{P} - \hat{Y})]\!] \|_{L^2(\partial T \cap \Omega)}^2.$$

Remark

In case $z_d \notin H^2(\Omega)$ or z_d not piecewise H^2 over $\mathcal{T}$, the residual estimator cannot be efficient. In this case it is better to use one of the other estimators.

Remarks

Remark (A Posteriori Error Analysis)

1. The approach simplifies the a posteriori error analysis drastically. Having reliable and efficient estimators for the linear problems (LP) and (LP*) the sum gives a reliable and efficient estimator for the optimal control problem.

2. Although there exists an exhaustive literature about a posteriori error analysis for linear problems, optimal control problems may not be included depending on the type of control and desired state.

   This may require a new error analysis but only for the linear problems.

3. For the equivalences of the errors we only used the discrete gradient equation but neither the discrete state nor the discrete adjoint problem.

   This indicates that variational crimes such as inexact solution, stabilized discretizations (SUPG for advection-diffusion equations), etc. are included in this approach.

Remarks

Remark (A Posteriori Error Analysis)

1. The approach simplifies the a posteriori error analysis drastically. Having reliable and efficient estimators for the linear problems (LP) and (LP^*) the sum gives a reliable and efficient estimator for the optimal control problem.

2. Although there exists an exhaustive literature about a posteriori error analysis for linear problems, optimal control problems may not be included depending on the type of control and desired state.

   This may require a new error analysis but only for the linear problems.

3. For the equivalences of the errors we only used the discrete gradient equation but neither the discrete state nor the discrete adjoint problem.

   This indicates that variational crimes such as inexact solution, stabilized discretizations (SUPG for advection-diffusion equations), etc. are included in this approach.

Remarks

Remark (A Posteriori Error Analysis)

1. The approach simplifies the a posteriori error analysis drastically. Having reliable and efficient estimators for the linear problems (LP) and (LP*) the sum gives a reliable and efficient estimator for the optimal control problem.

2. Although there exists an exhaustive literature about a posteriori error analysis for linear problems, optimal control problems may not be included depending on the type of control and desired state.

   This may require a new error analysis but only for the linear problems.

3. For the equivalences of the errors we only used the discrete gradient equation but neither the discrete state nor the discrete adjoint problem.

   This indicates that variational crimes such as inexact solution, stabilized discretizations (SUPG for advection-diffusion equations), etc. are included in this approach.

Outline

Convergence Analysis

Let $\{(\hat{Y}_k, \hat{P}_k, \hat{U}_k)\}_{k\geq 0}$ be the sequence of discrete solutions generated by

SOLVE $\longrightarrow$ ESTIMATE $\longrightarrow$ MARK $\longrightarrow$ REFINE

and let $(\hat{y}, \hat{p}, \hat{u})$ be the true solution.

Convergence?

Does there hold:

$$\lim_{k\to\infty} \|\hat{Y}_k - \hat{y},\ \hat{P}_k - \hat{p},\ \hat{U}_k - \hat{u}\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} = 0?$$

In proving this, we follow the ideas of [Morin, S. Veeser] and [S.] for convergence of adaptive finite elements for inf-sup stable discretizations, i. e.,

$$\inf_{\substack{V\in\mathbb{Y}(\mathcal{T})\\ \|V\|_{\mathbb{Y}}=1}} \sup_{\substack{W\in\mathbb{Y}(\mathcal{T})\\ \|W\|_{\mathbb{Y}}=1}} \mathcal{B}[V, W] = \inf_{\substack{W\in\mathbb{Y}(\mathcal{T})\\ \|W\|_{\mathbb{Y}}=1}} \sup_{\substack{V\in\mathbb{Y}(\mathcal{T})\\ \|V\|_{\mathbb{Y}}=1}} \mathcal{B}[V, W] = \gamma(\mathcal{T}) \geq \underline{\gamma} > 0.$$

of linear PDEs.

Convergence Analysis

Let $\{(\hat{Y}_k, \hat{P}_k, \hat{U}_k)\}_{k\geq 0}$ be the sequence of discrete solutions generated by

SOLVE $\longrightarrow$ ESTIMATE $\longrightarrow$ MARK $\longrightarrow$ REFINE

and let $(\hat{y}, \hat{p}, \hat{u})$ be the true solution.

Convergence?

Does there hold:

$$\lim_{k\to\infty} \|\hat{Y}_k - \hat{y},\, \hat{P}_k - \hat{p},\, \hat{U}_k - \hat{u}\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} = 0?$$

In proving this, we follow the ideas of [Morin, S. Veeser] and [S.] for convergence of adaptive finite elements for inf-sup stable discretizations, i. e.,

$$\inf_{\substack{V\in\mathbb{Y}(\mathcal{T})\\ \|V\|_{\mathbb{Y}}=1}} \sup_{\substack{W\in\mathbb{Y}(\mathcal{T})\\ \|W\|_{\mathbb{Y}}=1}} \mathcal{B}[V,\, W] = \inf_{\substack{W\in\mathbb{Y}(\mathcal{T})\\ \|W\|_{\mathbb{Y}}=1}} \sup_{\substack{V\in\mathbb{Y}(\mathcal{T})\\ \|V\|_{\mathbb{Y}}=1}} \mathcal{B}[V,\, W] = \gamma(\mathcal{T}) \geq \underline{\gamma} > 0.$$

of linear PDEs.

Basic Steps of the Convergence Proof [Morin, S., Veeser]

Let $\mathbb{V}$ be a Hilbert-space, $\hat{v}$ the true solution, and $\{V_k\}_{k\geq 0}$ the adaptively generated sequence of discrete solutions in $\{\mathbb{V}(\mathcal{T}_k)\}_{k\geq 0}$.

Basic Ingredients

1. Properties of refinement yield uniform convergence of the mesh-size functions $\{h_k\}_{k\geq 0}$:
$$\|h_k - h_\infty\|_{\infty;\Omega} \to 0.$$

2. Nesting of spaces and a uniform bound on the discrete inf-sup constant yield convergence of the discrete solutions, i. e.,
$$\|V_k - v_\infty\|_{\mathbb{V}} \to 0 \qquad \text{for some } v_\infty \in \mathbb{V}_\infty = \overline{\bigcup_{k\geq 0} \mathbb{V}(\mathcal{T}_k)}^{\|\cdot\|_{\mathbb{V}}}.$$

3. Properties of the indicators and properties of marking then yield
$$\max_{T\in\mathcal{T}_k}\{\mathcal{E}_k(V_k;T) \mid T \in \mathcal{T}_k\} \to 0.$$

4. Pointwise convergence of the indicators can be converted into integral convergence, i. e.,
$$\|V_k - v\|_{\mathbb{V}} \lesssim \mathcal{E}_k(V_k;\mathcal{T}_k) \to 0.$$

Basic Steps of the Convergence Proof [Morin, S., Veeser]

Let $\mathbb{V}$ be a Hilbert-space, $\hat{v}$ the true solution, and $\{V_k\}_{k\ge 0}$ the adaptively generated sequence of discrete solutions in $\{\mathbb{V}(\mathcal{T}_k)\}_{k\ge 0}$.

Basic Ingredients

1. Properties of refinement yield uniform convergence of the mesh-size functions $\{h_k\}_{k\ge 0}$:
$$\|h_k - h_\infty\|_{\infty;\Omega} \to 0.$$
2. Nesting of spaces and a uniform bound on the discrete inf-sup constant yield convergence of the discrete solutions, i. e.,
$$\|V_k - v_\infty\|_{\mathbb{V}} \to 0 \qquad \text{for some } v_\infty \in \mathbb{V}_\infty = \overline{\bigcup_{k\ge 0} \mathbb{V}(\mathcal{T}_k)}^{\|\cdot\|_{\mathbb{V}}}.$$
3. Properties of the indicators and properties of marking then yield
$$\max_{T\in\mathcal{T}_k} \{\mathcal{E}_k(V_k; T) \mid T \in \mathcal{T}_k\} \to 0.$$
4. Pointwise convergence of the indicators can be converted into integral convergence, i. e.,
$$\|V_k - v\|_{\mathbb{V}} \lesssim \mathcal{E}_k(V_k; \mathcal{T}_k) \to 0.$$

Basic Steps of the Convergence Proof [Morin, S., Veeser]

Let $\mathbb{V}$ be a Hilbert-space, $\hat{v}$ the true solution, and $\{V_k\}_{k\geq 0}$ the adaptively generated sequence of discrete solutions in $\{\mathbb{V}(\mathcal{T}_k)\}_{k\geq 0}$.

Basic Ingredients

1. Properties of refinement yield uniform convergence of the mesh-size functions $\{h_k\}_{k\geq 0}$:
$$\|h_k - h_\infty\|_{\infty;\Omega} \to 0.$$
2. Nesting of spaces and a uniform bound on the discrete inf-sup constant yield convergence of the discrete solutions, i. e.,
$$\|V_k - v_\infty\|_{\mathbb{V}} \to 0 \qquad \text{for some } v_\infty \in \mathbb{V}_\infty = \overline{\bigcup_{k\geq 0} \mathbb{V}(\mathcal{T}_k)}^{\|\cdot\|_{\mathbb{V}}}.$$
3. Properties of the indicators and properties of marking then yield
$$\max_{T\in\mathcal{T}_k}\{\mathcal{E}_k(V_k;T) \mid T \in \mathcal{T}_k\} \to 0.$$
4. Pointwise convergence of the indicators can be converted into integral convergence, i. e.,
$$\|V_k - v\|_{\mathbb{V}} \lesssim \mathcal{E}_k(V_k;\mathcal{T}_k) \to 0.$$

Basic Steps of the Convergence Proof [Morin, S., Veeser]

Let $\mathbb{V}$ be a Hilbert-space, $\hat{v}$ the true solution, and $\{V_k\}_{k\ge 0}$ the adaptively generated sequence of discrete solutions in $\{\mathbb{V}(\mathcal{T}_k)\}_{k\ge 0}$.

Basic Ingredients

1. Properties of refinement yield uniform convergence of the mesh-size functions $\{h_k\}_{k\ge 0}$:
$$\|h_k - h_\infty\|_{\infty;\Omega} \to 0.$$
2. Nesting of spaces and a uniform bound on the discrete inf-sup constant yield convergence of the discrete solutions, i. e.,
$$\|V_k - v_\infty\|_{\mathbb{V}} \to 0 \qquad \text{for some } v_\infty \in \mathbb{V}_\infty = \overline{\bigcup_{k\ge 0} \mathbb{V}(\mathcal{T}_k)}^{\|\cdot\|_{\mathbb{V}}}.$$
3. Properties of the indicators and properties of marking then yield
$$\max_{T\in\mathcal{T}_k} \{\mathcal{E}_k(V_k;T) \mid T \in \mathcal{T}_k\} \to 0.$$
4. Pointwise convergence of the indicators can be converted into integral convergence, i. e.,
$$\|V_k - v\|_{\mathbb{V}} \lesssim \mathcal{E}_k(V_k;\mathcal{T}_k) \to 0.$$

Convergence for Optimal Control Problems

For optimal control problems Step 2 - Step 4 have to be adapted in the convergence analysis.

Convergence of the Galerkin approximations is the tough part. Here, one has to deal with the nonlinearity of the problem.

Generalization of Step 3 and Step 4 is straight forward.

Convergence of Discrete Solutions

Proposition

Let $(\hat{y}_\infty, \hat{p}_\infty, \hat{u}_\infty) \in \mathbb{Y}_\infty \times \mathbb{Y}_\infty \times \mathbb{U}$ be the solution of the optimal control problem in

$$\mathbb{Y}_\infty = \overline{\bigcup_{k \geq 0} \mathbb{Y}(\mathcal{T}_k)}^{\|\cdot\|_{\mathbb{Y}}}$$

Then there holds

$$\lim_{k \to \infty} \|\hat{Y}_k - \hat{y}_\infty,\, \hat{P}_k - \hat{p}_\infty,\, \hat{U}_k - \hat{u}_\infty\|_{\mathbb{Y} \times \mathbb{Y} \times \mathbb{U}} = 0.$$

Main Ideas of the Proof. 1 Let $S_k, S_k^*, S_\infty, S_\infty^*$ be the solution operators in $\mathbb{Y}(\mathcal{T}_k)$ and $\mathbb{Y}_\infty$, respectively. The linear theory implies

$$\|S_k\|, \|S_k^*\|, \|S_\infty\|, \|S_\infty^*\| \leq \underline{\gamma}^{-1}$$

and pointwise convergence $S_k \to S_\infty$ and $S_k^* \to S_\infty^*$.

2 Show next via the gradient equation

$$\|\hat{U}_k - \hat{u}_\infty\|_{\mathbb{U}} \to 0.$$

3 Conclude form this convergence

$$\|\hat{Y}_k - \hat{y}_\infty\|_{\mathbb{Y}} \to 0 \quad \text{and} \quad \|\hat{P}_k - \hat{p}_\infty\|_{\mathbb{Y}} \to 0.$$

□

Convergence of Discrete Solutions

Proposition

Let $(\hat{y}_\infty, \hat{p}_\infty, \hat{u}_\infty) \in \mathbb{Y}_\infty \times \mathbb{Y}_\infty \times \mathbb{U}$ be the solution of the optimal control problem in

$$\mathbb{Y}_\infty = \overline{\bigcup_{k\geq 0} \mathbb{Y}(\mathcal{T}_k)}^{\|\cdot\|_{\mathbb{Y}}}$$

Then there holds

$$\lim_{k\to\infty} \|\hat{Y}_k - \hat{y}_\infty,\ \hat{P}_k - \hat{p}_\infty,\ \hat{U}_k - \hat{u}_\infty\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} = 0.$$

Main Ideas of the Proof. [1] Let $S_k, S_k^*, S_\infty, S_\infty^*$ be the solution operators in $\mathbb{Y}(\mathcal{T}_k)$ and $\mathbb{Y}_\infty$, respectively. The linear theory implies

$$\|S_k\|, \|S_k^*\|, \|S_\infty\|, \|S_\infty^*\| \leq \underline{\gamma}^{-1}$$

and pointwise convergence $S_k \to S_\infty$ and $S_k^* \to S_\infty^*$.

[2] Show next via the gradient equation

$$\|\hat{U}_k - \hat{u}_\infty\|_{\mathbb{U}} \to 0.$$

[3] Conclude form this convergence

$$\|\hat{Y}_k - \hat{y}_\infty\|_{\mathbb{Y}} \to 0 \qquad \text{and} \qquad \|\hat{P}_k - \hat{p}_\infty\|_{\mathbb{Y}} \to 0.$$

□

Convergence

Theorem (Convergence [Kohls, Rösch, S. '10])

Let the discrete spaces be uniformly inf-sup stable and nested. Suppose reasonable (and standard!) properties of the estimators $\mathcal{E}_{\mathcal{T}}, \mathcal{E}^*_{\mathcal{T}}$ and of marking.
Then there holds

$$\lim_{k\to\infty} \|\hat{Y}_k - \hat{y},\, \hat{P}_k - \hat{p},\, \hat{U}_k - \hat{u}\|_{\mathbb{Y}\times\mathbb{Y}\times\mathbb{U}} = 0$$

as well as

$$\lim_{k\to\infty} \mathcal{E}_k(\hat{Y}_k, \hat{U}_k; \mathcal{T}_k) + \mathcal{E}^*_k(\hat{P}_k, \psi'(\hat{Y}_k); \mathcal{T}_k) = 0.$$

Outline

Concluding Remarks

For control constrained optimal control problems discretized via the reduced first order optimality system we have presented a

1. a framework for the a posteriori error analysis for control constrained optimal control problems based on results from the linear theory;
2. a general convergence result for adaptive methods based on assumptions for the linear sub-problems.

Ideas transfer to a discretization of the full first order optimality system, i. e., including a discretization of the control.

Perspectives: The challange ahead are state constrained or mixed state-control constrained optimal control problems.

Thank you for your interest!