Adaptive finite element methods for PDE-constrained optimization problems

Boris Vexler

joint work with: Dominik Meidner, Olaf Benedix, Winnifried Wollner, Michael Schmich

Fakultät für Mathematik
Technische Universität München

June 4, 2009

Outline

Outline

Motivation

→ Applications from engineering and natural sciences

→ mathematical models with partial differential equations (PDEs)

→ Goals:

- Numerical simulation of the process
- Calibration of the mathematical model
- Optimization / Optimal control of the underlying process

Optimization problem (infinite dimensional)

$$\text{Minimize} \quad J(q,u), \qquad q \in Q_{\text{ad}}, \; u \in V_{\text{ad}}$$

subject to

$$A(q,u) = 0$$

Discretized optimization problem

$$\text{Minimize} \quad J(q_h,u_h), \qquad q_h \in Q_h \cap Q_{\text{ad}}\,, \; u_h \in V_h \cap V_{\text{ad}}$$

subject to

$$A_h(q_h,u_h) = 0$$

- $\dim V_h, \dim Q_h < \infty$
- Discretization error between (q,u) and (q_h,u_h)
- Numerical effort depends on $\dim V_h$ and $\dim Q_h$

Motivation

Optimization problem (infinite dimensional)

$$\text{Minimize} \quad J(q,u), \qquad q \in Q_{\text{ad}}, \; u \in V_{\text{ad}}$$

subject to

$$A(q,u) = 0$$

Discretized optimization problem

$$\text{Minimize} \quad J(q_h,u_h), \qquad q_h \in Q_h \cap Q_{\text{ad}}, \; u_h \in V_h \cap V_{\text{ad}}$$

subject to

$$A_h(q_h,u_h) = 0$$

- $\dim V_h, \dim Q_h < \infty$
- Discretization error between (q,u) and (q_h,u_h)
- Numerical effort depends on $\dim V_h$ and $\dim Q_h$

Optimization problem (infinite dimensional)

$$\text{Minimize} \quad J(q, u), \qquad q \in Q_{\text{ad}}, \ u \in V_{\text{ad}}$$

subject to

$$A(q, u) = 0$$

Discretized optimization problem

$$\text{Minimize} \quad J(q_h, u_h), \qquad q_h \in Q_h \cap Q_{\text{ad}}, \ u_h \in V_h \cap V_{\text{ad}}$$

subject to

$$A_h(q_h, u_h) = 0$$

- $\dim V_h, \dim Q_h < \infty$
- Discretization error between (q, u) and (q_h, u_h)
- Numerical effort depends on $\dim V_h$ and $\dim Q_h$

Optimization problem (infinite dimensional)

$$\text{Minimize} \quad J(q,u), \qquad q \in Q_{\text{ad}}, \; u \in V_{\text{ad}}$$

subject to

$$A(q,u) = 0$$

Discretized optimization problem

$$\text{Minimize} \quad J(q_h,u_h), \qquad q_h \in Q_h \cap Q_{\text{ad}}\,, \; u_h \in V_h \cap V_{\text{ad}}$$

subject to

$$A_h(q_h,u_h) = 0$$

- $\dim V_h, \dim Q_h < \infty$
- Discretization error between (q,u) and (q_h,u_h)
- Numerical effort depends on $\dim V_h$ and $\dim Q_h$

Motivation

→ Goal:

- "Discretization error" $\leq$ tolerance
- Numerical effort → min!

→ Questions:

- What kind of error is important?

→ Quantity of interest: $I(q, u)$

- How to measure/estimate this error?

→ A posteriori error estimation: $I(q, u) - I(q_h, u_h) \approx \eta_h$

- How to construct an appropriate discretization?

→ Adaptive choice of discretizations: locally refined meshes

Motivation

→ Goal:

- "Discretization error" $\leq$ tolerance
- Numerical effort $\rightarrow$ min!

→ Questions:

- What kind of error is important?

→ Quantity of interest: $I(q, u)$

- How to measure/estimate this error?

→ A posteriori error estimation: $I(q, u) - I(q_h, u_h) \approx \eta_h$

- How to construct an appropriate discretization?

→ Adaptive choice of discretizations: locally refined meshes

Motivation

→ Goal:

- "Discretization error" $\leq$ tolerance
- Numerical effort → min!

→ Questions:

- What kind of error is important?

→ Quantity of interest: $I(q, u)$

- How to measure/estimate this error?

→ A posteriori error estimation: $I(q, u) - I(q_h, u_h) \approx \eta_h$

- How to construct an appropriate discretization?

→ Adaptive choice of discretizations: locally refined meshes

Motivation

→ Goal:

- "Discretization error" $\leq$ tolerance
- Numerical effort $\rightarrow$ min!

→ Questions:

- What kind of error is important?

→ Quantity of interest: $I(q, u)$

- How to measure/estimate this error?

→ A posteriori error estimation: $I(q, u) - I(q_h, u_h) \approx \eta_h$

- How to construct an appropriate discretization?

→ Adaptive choice of discretizations: locally refined meshes

Motivation

→ Goal:

- "Discretization error" $\leq$ tolerance
- Numerical effort → min!

→ Questions:

- What kind of error is important?

→ Quantity of interest: $I(q, u)$

- How to measure/estimate this error?

→ A posteriori error estimation: $I(q, u) - I(q_h, u_h) \approx \eta_h$

- How to construct an appropriate discretization?

→ Adaptive choice of discretizations: locally refined meshes

Motivation

→ Goal:

- "Discretization error" $\leq$ tolerance
- Numerical effort → min!

→ Questions:

- What kind of error is important?

→ Quantity of interest: $I(q, u)$

- How to measure/estimate this error?

→ A posteriori error estimation: $I(q, u) - I(q_h, u_h) \approx \eta_h$

- How to construct an appropriate discretization?

→ Adaptive choice of discretizations: locally refined meshes

Motivation

→ Goal:

- "Discretization error" $\leq$ tolerance
- Numerical effort $\rightarrow$ min!

→ Questions:

- What kind of error is important?

→ Quantity of interest: $I(q, u)$

- How to measure/estimate this error?

→ A posteriori error estimation: $I(q, u) - I(q_h, u_h) \approx \eta_h$

- How to construct an appropriate discretization?

→ Adaptive choice of discretizations: locally refined meshes

Motivation

$\rightarrow$ Goal:

- "Discretization error" $\leq$ tolerance
- Numerical effort $\rightarrow$ min!

$\rightarrow$ Questions:

- What kind of error is important?

$\rightarrow$ Quantity of interest: $I(q, u)$

- How to measure/estimate this error?

$\rightarrow$ A posteriori error estimation: $I(q, u) - I(q_h, u_h) \approx \eta_h$

- How to construct an appropriate discretization?

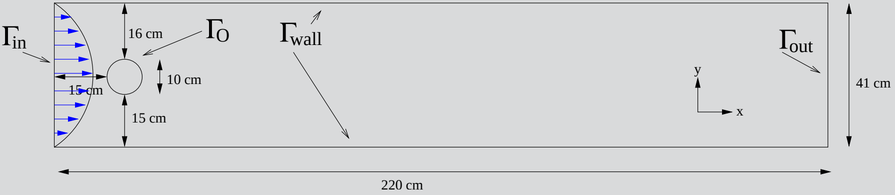
$\rightarrow$ Adaptive choice of discretizations: locally refined meshes

Literature

- Becker, Kapp, Rannacher
- Liu et al.
- Hoppe, Kieweg, Hintermüller
- Vexler, Benedix, Wollner
- Günther, Hinze, Schiela
- Meidner, Vexler
- ...

Outline

Flow around a cylinder



Optimization problem

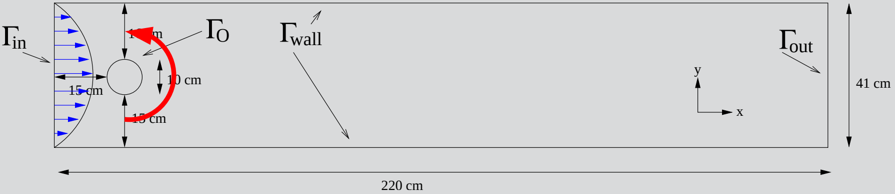
Minimize $\quad J(u) = c_0 \int_{\Gamma_O} n \cdot \sigma(u) \cdot d \, ds, \quad \sigma(u) = \frac{\nu}{2}(\nabla v + \nabla v^T) - pI, \quad u = (v, p)$

$$-\nu \Delta v + v \cdot \nabla v + \nabla p = 0, \quad \nabla \cdot v = 0 \text{ in } \Omega$$

$$v = v_{in} \text{ on } \Gamma_{in}, \quad v = 0 \text{ on } \Gamma_{wall}, \quad \nu \partial_n v - p \cdot n = 0 \text{ on } \Gamma_{out}$$

$$v \cdot n = 0 \text{ on } \Gamma_O, \quad v \cdot \tau = q \text{ on } \Gamma_O$$

Flow around a cylinder



Optimization problem

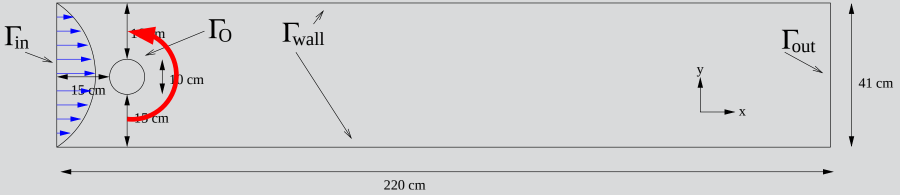
Minimize $J(u) = c_0 \int_{\Gamma_O} n \cdot \sigma(u) \cdot d \, ds, \quad \sigma(u) = \frac{\nu}{2}(\nabla v + \nabla v^T) - pI, \quad u = (v, p)$

$$-\nu \Delta v + v \cdot \nabla v + \nabla p = 0, \quad \nabla \cdot v = 0 \text{ in } \Omega$$

$$v = v_{in} \text{ on } \Gamma_{in}, \quad v = 0 \text{ on } \Gamma_{wall}, \quad \nu \partial_n v - p \cdot n = 0 \text{ on } \Gamma_{out}$$

$$v \cdot n = 0 \text{ on } \Gamma_O, \quad v \cdot \tau = q \text{ on } \Gamma_O$$

Flow around a cylinder



Optimization problem

Minimize $J(u) = c_0 \int_{\Gamma_O} n \cdot \sigma(u) \cdot d \, ds, \quad \sigma(u) = \frac{\nu}{2}(\nabla v + \nabla v^T) - pI, \quad u = (v, p)$

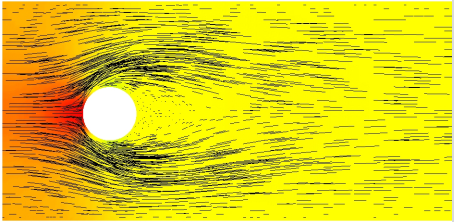
$$-\nu \Delta v + v \cdot \nabla v + \nabla p = 0, \quad \nabla \cdot v = 0 \text{ in } \Omega$$

$$v = v_{in} \text{ on } \Gamma_{in}, \quad v = 0 \text{ on } \Gamma_{wall}, \quad \nu \partial_n v - p \cdot n = 0 \text{ on } \Gamma_{out}$$

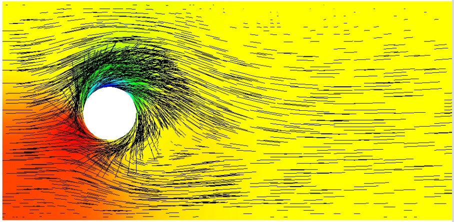
$$v \cdot n = 0 \text{ on } \Gamma_O, \qquad v \cdot \tau = q \text{ on } \Gamma_O$$

Example 1: Drag minimization

uncontrolled flow, $J(u) \approx 5.58$



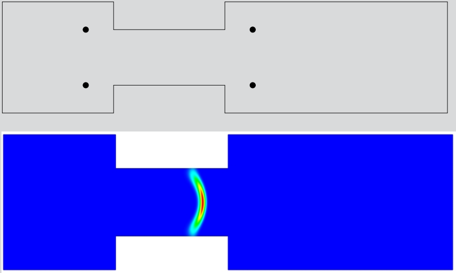
controlled flow, $J(u) \approx 5.04$



- Quantity of interest $I(q,u) = J(u)$ (QI = Cost functional)
- Discretization of the state variable $u = (v, p)$ in space
- Error estimation $J(u) - J(u_h)$

Example 2: Parameter Identification

Configuration



Mathematical model

$$\partial_t \theta - \Delta\theta = \omega \qquad \text{in } (0, T) \times \Omega$$

$$\partial_t y - \frac{1}{L}\Delta y = -\omega \qquad \text{in } (0, T) \times \Omega$$

$$\omega = \frac{\beta^2}{2\gamma}\, y \exp\left(\frac{\beta(\theta - 1)}{1 + \alpha(\theta - 1)}\right)$$

+ Initial and bounday conditions

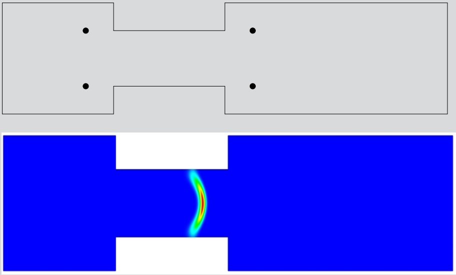
- The state $u = (\theta, y)$ consists of temperature and concentration of the fuel
- The parameter $q = \alpha$ has to be estimated

State equation

$$A(q, u) = 0$$

Example 2: Parameter Identification

Configuration



Mathematical model

$$\partial_t \theta - \Delta\theta = \omega \qquad \text{in } (0,T)\times\Omega$$

$$\partial_t y - \frac{1}{L}\Delta y = -\omega \qquad \text{in } (0,T)\times\Omega$$

$$\omega = \frac{\beta^2}{2\gamma}\, y \exp\left(\frac{\beta(\theta-1)}{1+\alpha(\theta-1)}\right)$$

+ Initial and bounday conditions

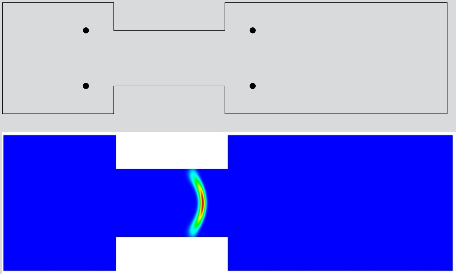
- The state $u = (\theta, y)$ consists of temperature and concentration of the fuel
- The parameter $q = \alpha$ has to be estimated

State equation

$$A(q,u) = 0$$

Example 2: Parameter Identification

Configuration



Mathematical model

$$\partial_t \theta - \Delta\theta = \omega \qquad \text{in } (0,T)\times\Omega$$

$$\partial_t y - \frac{1}{L}\Delta y = -\omega \qquad \text{in } (0,T)\times\Omega$$

$$\omega = \frac{\beta^2}{2\gamma}\, y \exp\left(\frac{\beta(\theta-1)}{1+\alpha(\theta-1)}\right)$$

+ Initial and bounday conditions

- The state $u = (\theta, y)$ consists of temperature and concentration of the fuel
- The parameter $q = \alpha$ has to be estimated

State equation

$$A(q,u) = 0$$

Example 2: Parameter Identification

- Comparison of an experiment and the simulation
- Observation $\overline{C}$ vs. $C(u)$

Optimization problem

$$\text{Minimize} \quad J(q,u) = \frac{1}{2}\|C(u) - \overline{C}\|_Z^2, \quad q \in Q, \; u \in V$$

subject to

$$A(q,u) = 0$$

- Quantity of interest: $I(q,u) = \alpha$
- Discretization of the state variable in space and time

Example 2: Parameter Identification

- Comparison of an experiment and the simulation
- Observation $\overline{C}$ vs. $C(u)$

Optimization problem

$$\text{Minimize} \quad J(q,u) = \frac{1}{2}\|C(u) - \overline{C}\|_Z^2, \qquad q \in Q, \; u \in V$$

subject to

$$A(q,u) = 0$$

- Quantity of interest: $I(q,u) = \alpha$
- Discretization of the state variable in space and time

Example 2: Parameter Identification

- Comparison of an experiment and the simulation
- Observation $\overline{C}$ vs. $C(u)$

Optimization problem

$$\text{Minimize} \quad J(q,u) = \frac{1}{2}\|C(u) - \overline{C}\|_Z^2, \qquad q \in Q,\; u \in V$$

subject to

$$A(q,u) = 0$$

- Quantity of interest: $I(q,u) = \alpha$
- Discretization of the state variable in space and time

Example 2: Parameter Identification

- Comparison of an experiment and the simulation
- Observation $\overline{C}$ vs. $C(u)$

Optimization problem

$$\text{Minimize} \quad J(q,u) = \frac{1}{2}\|C(u) - \overline{C}\|_Z^2, \qquad q \in Q,\ u \in V$$

subject to

$$A(q,u) = 0$$

- Quantity of interest: $I(q,u) = \alpha$
- Discretization of the state variable in space and time

Example 2: Parameter Identification

- Comparison of an experiment and the simulation
- Observation $\overline{C}$ vs. $C(u)$

Optimization problem

$$\text{Minimize} \quad J(q,u) = \frac{1}{2}\|C(u) - \overline{C}\|_Z^2, \quad q \in Q,\ u \in V$$

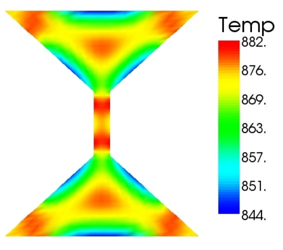
subject to

$$A(q,u) = 0$$

- Quantity of interest: $I(q,u) = \alpha$
- Discretization of the state variable in space and time

Example 3: Glass cooling

Configuration



Mathematical model

$$
\begin{aligned}
\partial_t \theta - k\Delta\theta &= \tfrac{1}{3\kappa}\Delta\rho && \text{in } (0,T)\times\Omega \\
-\tfrac{\varepsilon^2}{3\kappa}\Delta\rho + \kappa\rho &= 4\kappa\pi a\theta^4 && \text{in } (0,T)\times\Omega \\
\tfrac{1}{\varepsilon k}\theta + n\cdot\nabla\theta &= \tfrac{1}{\varepsilon k} q && \text{auf } (0,T)\times\partial\Omega \\
\tfrac{3\kappa}{2\varepsilon}\rho + n\cdot\nabla\rho &= \tfrac{3\kappa}{2\varepsilon}4\pi a q^4 && \text{auf } (0,T)\times\partial\Omega
\end{aligned}
$$

- The state $u = (\theta, \rho)$ consist of the temperature and the radiation transfer
- Die control q is the ambient temperature

State equation

$$A(q, u) = 0$$

Optimization problem

Minimize $J(q,u) = \int_0^T \int_\Omega (\theta - \theta_d)^2 \, dx \, dt + \frac{\alpha}{2} \int_0^T q^2 \, dt, \quad q \in Q_{\text{ad}}, \; u \in V_{\text{ad}}$

subject to

$$A(q,u) = 0$$

- Control constraints

$$q \in Q_{\text{ad}} = \left\{ q \in L^2(0,T) \;\middle|\; q_a \leq q(t) \leq q_b \text{ a. e. in } (0,T) \right\}$$

- State constraints

$$\theta_a \leq \theta(t,x) \leq \theta_b \quad \text{and/or} \quad |\nabla\theta(t,x)| \leq c_b$$

- Discretization of the state variable in space and time
- Discretization of the control variable in time

Example 3: Glass cooling

Optimization problem

Minimize $J(q,u) = \int_0^T \int_\Omega (\theta - \theta_d)^2 \, dx \, dt + \frac{\alpha}{2} \int_0^T q^2 \, dt, \quad q \in Q_{\text{ad}}, \; u \in V_{\text{ad}}$

subject to

$$A(q,u) = 0$$

- Control constraints

$$q \in Q_{\text{ad}} = \left\{ q \in L^2(0,T) \;\middle|\; q_a \le q(t) \le q_b \text{ a.e. in } (0,T) \right\}$$

- State constraints

$$\theta_a \le \theta(t,x) \le \theta_b \quad \text{and/or} \quad |\nabla\theta(t,x)| \le c_b$$

- Discretization of the state variable in space and time
- Discretization of the control variable in time

Optimization problem

Minimize $J(q,u) = \int_0^T \int_\Omega (\theta - \theta_d)^2 \, dx \, dt + \frac{\alpha}{2} \int_0^T q^2 \, dt, \quad q \in Q_{\text{ad}}, \; u \in V_{\text{ad}}$

subject to

$$A(q,u) = 0$$

- Control constraints

$$q \in Q_{\text{ad}} = \{\, q \in L^2(0,T) \;|\; q_a \le q(t) \le q_b \;\text{ a.e. in } (0,T) \,\}$$

- State constraints

$$\theta_a \le \theta(t,x) \le \theta_b \quad \text{and/or} \quad |\nabla\theta(t,x)| \le c_b$$

- Discretization of the state variable in space and time
- Discretization of the control variable in time

Optimization problem

Minimize $J(q,u) = \int_0^T \int_\Omega (\theta - \theta_d)^2 \, dx \, dt + \frac{\alpha}{2} \int_0^T q^2 \, dt, \quad q \in Q_{\text{ad}}, \; u \in V_{\text{ad}}$

subject to

$$A(q,u) = 0$$

- Control constraints

$$q \in Q_{\text{ad}} = \left\{ q \in L^2(0,T) \;\middle|\; q_a \le q(t) \le q_b \text{ a.e. in } (0,T) \right\}$$

- State constraints

$$\theta_a \le \theta(t,x) \le \theta_b \quad \text{and/or} \quad |\nabla \theta(t,x)| \le c_b$$

- Discretization of the state variable in space and time
- Discretization of the control variable in time

Optimization problem

Minimize $J(q,u) = \int_0^T \int_\Omega (\theta - \theta_d)^2 \, dx \, dt + \frac{\alpha}{2} \int_0^T q^2 \, dt, \quad q \in Q_{\text{ad}}, \; u \in V_{\text{ad}}$

subject to

$$A(q,u) = 0$$

- Control constraints

$$q \in Q_{\text{ad}} = \left\{ \, q \in L^2(0,T) \; \middle| \; q_a \le q(t) \le q_b \;\text{ a.e. in } (0,T) \, \right\}$$

- State constraints

$$\theta_a \le \theta(t,x) \le \theta_b \quad \text{and/or} \quad |\nabla\theta(t,x)| \le c_b$$

- Discretization of the state variable in space and time
- Discretization of the control variable in time

Outline

Degrees of freedom for the choice of discretization

- Temporal discretization
  $\rightarrow$ Choice of the time steps
- Spatial discretization
  $\rightarrow$ Choice of spatial meshes for each time step
- Control discretization
  $\rightarrow$ Choice of the discretization of the control space

Goal: Error estimators

$$I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_l$$

Degrees of freedom for the choice of discretization

- Temporal discretization
  → Choice of the time steps
- Spatial discretization
  → Choice of spatial meshes for each time step
- Control discretization
  → Choice of the discretization of the control space

Goal: Error estimators

$$I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_I$$

Degrees of freedom for the choice of discretization

- Temporal discretization
  $\rightarrow$ Choice of the time steps
- Spatial discretization
  $\rightarrow$ Choice of spatial meshes for each time step
- Control discretization
  $\rightarrow$ Choice of the discretization of the control space

Goal: Error estimators

$$I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_l$$

Degrees of freedom for the choice of discretization

- Temporal discretization
  $\rightarrow$ Choice of the time steps
- Spatial discretization
  $\rightarrow$ Choice of spatial meshes for each time step
- Control discretization
  $\rightarrow$ Choice of the discretization of the control space

Goal: Error estimators

$$I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_l$$

Adaptive algorithm

1. Choose an initial discretization $\mathcal{T}_{\sigma_0}$, set $n = 0$
2. Compute (q_σ, u_σ)
3. Evaluate $\eta_{k_n}, \eta_{h_n}, \eta_{l_n}$
4. if $\eta_{k_n} + \eta_{h_n} + \eta_{l_n} \leq TOL$ break;
5. else
   - $\eta_k \gg \eta_h,\ \eta_k \gg \eta_l \ \Rightarrow$ Refine time discr.
   - $\eta_k \approx \eta_h \gg \eta_l \ \Rightarrow$ Refine time & space discr.
   - $\eta_k \approx \eta_h \approx \eta_l \Rightarrow$ Refine time, space & control discr.
   - ...
6. $n = n + 1$, go to 2

$\rightarrow$ Refinement is based on local information from η_k, η_h, η_l

Adaptive algorithm

1. Choose an initial discretization $\mathcal{T}_{\sigma_0}$, set $n = 0$
2. Compute (q_σ, u_σ)
3. Evaluate $\eta_{k_n}, \eta_{h_n}, \eta_{l_n}$
4. if $\eta_{k_n} + \eta_{h_n} + \eta_{l_n} \leq TOL$ break;
5. else
   - $\eta_k \gg \eta_h,\ \eta_k \gg \eta_l \ \Rightarrow$ Refine time discr.
   - $\eta_k \approx \eta_h \gg \eta_l \ \Rightarrow$ Refine time & space discr.
   - $\eta_k \approx \eta_h \approx \eta_l \Rightarrow$ Refine time, space & control discr.
   - ...
6. $n = n + 1$, go to 2

$\rightarrow$ Refinement is based on local information from η_k, η_h, η_l

Outline

→ State equation: $u_t + A(q,u) = f, \; u(0) = u_0$

Weak formulation

$$u \in X = W(0,T;V,V^*) \;:\; a(q,u)(\phi) = f(\phi) \quad \forall \phi \in X$$

with

$$a(q,u)(\phi) = \int_0^T \Big(\langle \partial_t u(t), \phi(t) \rangle_{V^*,V} + a_s(q,u(t))(\phi(t)) \Big) dt + (u(0) - u_0, \phi(0))$$

Optimization problem

$$\text{Minimize} \quad J(q,u), \qquad q \in Q_{\text{ad}}, \; u \in X$$

$$\text{s.t.} \quad a(q,u)(\phi) = f(\phi) \quad \forall \phi \in X$$

Parabolic problems

→ State equation: $u_t + A(q,u) = f,\; u(0) = u_0$

Weak formulation

$$u \in X = W(0,T;V,V^*) \;:\; a(q,u)(\phi) = f(\phi) \quad \forall \phi \in X$$

with

$$a(q,u)(\phi) = \int_0^T \Big(\langle \partial_t u(t), \phi(t) \rangle_{V^*,V} + a_s(q,u(t))(\phi(t)) \Big) dt + (u(0) - u_0, \phi(0))$$

Optimization problem

$$\text{Minimize} \quad J(q,u), \qquad q \in Q_{\text{ad}},\; u \in X$$

$$\text{s.t.} \quad a(q,u)(\phi) = f(\phi) \quad \forall \phi \in X$$

→ State equation: $u_t + A(q,u) = f,\ u(0) = u_0$

Weak formulation

$$u \in X = W(0,T;V,V^*)\ :\ a(q,u)(\phi) = f(\phi) \quad \forall \phi \in X$$

with

$$a(q,u)(\phi) = \int_0^T \Big(\langle \partial_t u(t), \phi(t) \rangle_{V^*,V} + a_s(q,u(t))(\phi(t)) \Big) dt + (u(0) - u_0, \phi(0))$$

Optimization problem

$$\text{Minimize} \quad J(q,u), \qquad q \in Q_{\text{ad}},\ u \in X$$

$$\text{s.t.} \quad a(q,u)(\phi) = f(\phi) \quad \forall \phi \in X$$

Optimality conditions

- Lagrange functional

$$\mathcal{L}(q,u,z) = J(q,u) + f(z) - a(q,u)(z)$$

- Necessary optimality conditions:

$$\mathcal{L}'(q,u,z)(\delta q, \delta u, \delta z) = 0 \quad \forall \delta q \in Q,\ \forall \delta u, \delta z \in X$$

Optimality system

$$\begin{aligned} a(q,u)(v) &= f(v) && \forall v \in X && \text{(state equation)} \\ a'_u(q,u)(v,z) &= J'_u(q,u)(v) && \forall v \in X && \text{(adjoint equation)} \\ a'_q(q,u)(\delta q,z) &= J'_q(q,u)(\delta q) && \forall \delta q \in Q && \text{(gradient equation)} \end{aligned}$$

$$\rightarrow \quad J'_q(q,u)(\delta q - q) - a'_q(q,u)(\delta q - q, z) \geq 0 \quad \forall \delta q \in Q_{ad}$$

- Lagrange functional

$$\mathcal{L}(q,u,z) = J(q,u) + f(z) - a(q,u)(z)$$

- Necessary optimality conditions:

$$\mathcal{L}'(q,u,z)(\delta q, \delta u, \delta z) = 0 \quad \forall \delta q \in Q,\ \forall \delta u, \delta z \in X$$

Optimality system

$$\begin{aligned} a(q,u)(v) &= f(v) & \forall v \in X & \quad \text{(state equation)} \\ a'_u(q,u)(v,z) &= J'_u(q,u)(v) & \forall v \in X & \quad \text{(adjoint equation)} \\ a'_q(q,u)(\delta q, z) &= J'_q(q,u)(\delta q) & \forall \delta q \in Q & \quad \text{(gradient equation)} \end{aligned}$$

$\rightarrow \quad J'_q(q,u)(\delta q - q) - a'_q(q,u)(\delta q - q, z) \geq 0 \quad \forall \delta q \in Q_{ad}$

Optimality conditions

- Lagrange functional

$$\mathcal{L}(q,u,z) = J(q,u) + f(z) - a(q,u)(z)$$

- Necessary optimality conditions:

$$\mathcal{L}'(q,u,z)(\delta q, \delta u, \delta z) = 0 \quad \forall \delta q \in Q,\ \forall \delta u, \delta z \in X$$

Optimality system

$$\begin{aligned}
a(q,u)(v) &= f(v) && \forall v \in X && \text{(state equation)}\\
a'_u(q,u)(v,z) &= J'_u(q,u)(v) && \forall v \in X && \text{(adjoint equation)}\\
a'_q(q,u)(\delta q,z) &= J'_q(q,u)(\delta q) && \forall \delta q \in Q && \text{(gradient equation)}
\end{aligned}$$

$$\to \quad J'_q(q,u)(\delta q - q) - a'_q(q,u)(\delta q - q, z) \geq 0 \quad \forall \delta q \in Q_{\text{ad}}$$

Optimality conditions

- Lagrange functional

$$\mathcal{L}(q,u,z) = J(q,u) + f(z) - a(q,u)(z)$$

- Necessary optimality conditions:

$$\mathcal{L}'(q,u,z)(\delta q, \delta u, \delta z) = 0 \quad \forall \delta q \in Q,\ \forall \delta u, \delta z \in X$$

Optimality system

$$\begin{aligned}
a(q,u)(v) &= f(v) && \forall v \in X && \text{(state equation)}\\
a'_u(q,u)(v,z) &= J'_u(q,u)(v) && \forall v \in X && \text{(adjoint equation)}\\
a'_q(q,u)(\delta q,z) &= J'_q(q,u)(\delta q) && \forall \delta q \in Q && \text{(gradient equation)}
\end{aligned}$$

$$\rightarrow \quad J'_q(q,u)(\delta q - q) - a'_q(q,u)(\delta q - q, z) \geq 0 \quad \forall \delta q \in Q_{\text{ad}}$$

Outline

Optimize-then-Discretize versus Discretize-then-Optimize

→ Optimize-then-Discretize

- Build up the optimality system on the continuous level
- Discretize the state, the adjoint and the gradient equations

→ Appropriate and stable discretizations of all involved equations

→ Discretize-then-Optimize

- Discretize the state equation
- Build up the optimality system for the discrete optimization problem

→ Preserving problem structure (symmetry of the optimality system)

⟶ For a pure Galerkin discretization:

Optimize-then-Discretize = Discretize-then-Optimize

→ Exact discrete derivatives

Space-time finite elements

- Use finite elements for spatial and temporal discretization
  - $\rightarrow$ Optimize-then-Discretize = Discretize-then-Optimize
  - $\rightarrow$ Systematic a priori error analysis
  - $\rightarrow$ Systematic a posteriori error estimation
  - $\rightarrow$ Dynamic meshes

- Temporal discretization: Discontinuous Galerkin methods
  - Time partitioning: $\bar{I} = \{0\} \cup I_1 \cup I_2 \ldots \cup I_M$ with

$$I_m = (t_{m-1}, t_m] \text{ and } 0 = t_0 < t_1 < \ldots < t_{M-1} < t_M = T$$

  - Semidiscretized state space

$$X_k^r = \{v_k \in L^2(I, V) \mid v_k\big|_{I_m} \in P_r(I_m, V) \text{ and } v_k(0) \in H\}$$

- Use finite elements for spatial and temporal discretization
  - → Optimize-then-Discretize = Discretize-then-Optimize
  - → Systematic a priori error analysis
  - → Systematic a posteriori error estimation
  - → Dynamic meshes

- Temporal discretization: Discontinuous Galerkin methods
  - Time partitioning: $\bar{I} = \{0\} \cup I_1 \cup I_2 \ldots \cup I_M$ with

$$I_m = (t_{m-1}, t_m] \text{ and } 0 = t_0 < t_1 < \ldots < t_{M-1} < t_M = T$$

  - Semidiscretized state space

$$X_k^r = \{v_k \in L^2(I, V) \mid v_k\big|_{I_m} \in P_r(I_m, V) \text{ and } v_k(0) \in H\}$$

Semi-discretization in time

$$u_k \in X_k^r \ : \ a_k(q_k, u_k)(\phi_k) = f(\phi_k) \quad \forall \phi_k \in X_k^r$$

with

$$a_k(q_k, u_k)(\phi_k) := a(q_k, u_k)(\phi) + \sum_{m=1}^{M-1} ([u_k]_m, \phi_m^+)$$

Spatial discretization

$$u_{kh} \in X_{kh}^{r,s} \ : \ a_k(q_{kh}, u_{kh})(\phi_{kh}) = f(\phi_{kh}) \quad \forall \phi_{kh} \in X_{kh}^{r,s}$$

with the discrete state space:

$$X_{kh}^{r,s} = \{ v_{kh} \in X_k^r \ : \ v_{kh}\big|_{I_m} \in P_r(I_m, V_{h,m}^s) \text{ and } v_k(0) \in V_{h,0}^s \}$$

Semi-discretization in time

$$u_k \in X_k^r \;:\; a_k(q_k, u_k)(\phi_k) = f(\phi_k) \quad \forall \phi_k \in X_k^r$$

with

$$a_k(q_k, u_k)(\phi_k) := a(q_k, u_k)(\phi) + \sum_{m=1}^{M-1} ([u_k]_m, \phi_m^+)$$

Spatial discretization

$$u_{kh} \in X_{kh}^{r,s} \;:\; a_k(q_{kh}, u_{kh})(\phi_{kh}) = f(\phi_{kh}) \quad \forall \phi_{kh} \in X_{kh}^{r,s}$$

with the discrete state space:

$$X_{kh}^{r,s} = \{v_{kh} \in X_k^r \;:\; v_{kh}\big|_{I_m} \in P_r(I_m, V_{h,m}^s) \text{ and } v_k(0) \in V_{h,0}^s\}$$

→ Reinterpret Galerkin methods as time stepping schemes

- dG(0) $\Longrightarrow$ (variant of) implicit Euler scheme
- dG(r) $\Longrightarrow$ A-stable scheme of order $r+1$

Discrete optimization problem

$$\text{Minimize} \quad J(q_\sigma, u_\sigma), \qquad q_\sigma \in Q_l, u_\sigma \in X_{kh}^{r,s}$$

subject to

$$a_k(q_\sigma, u_\sigma)(\phi_{kh}) = f(\phi_{kh}) \quad \forall \phi_{kh} \in X_{kh}^{r,s}$$

Outline

A posteriori error estimation

→ Goal: $I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_I$

→ Lagrange functional: $\mathcal{L}(q,u,z) = J(q,u) + f(z) - a_k(q,u)(z)$

$$\begin{aligned}
\xi &= (q,u,z) &&\in Q \times X \times X && \mathcal{L}'(\xi)(\delta\xi) = 0\\
\xi_k &= (q_k,u_k,z_k) &&\in Q \times X_k \times X_k && \mathcal{L}'(\xi_k)(\delta\xi_k) = 0\\
\xi_{kh} &= (q_{kh},u_{kh},z_{kh}) &&\in Q \times X_{kh} \times X_{kh} && \mathcal{L}'(\xi_{kh})(\delta\xi_{kh}) = 0\\
\xi_\sigma &= (q_\sigma,u_\sigma,z_\sigma) &&\in Q_I \times X_{kh} \times X_{kh} && \mathcal{L}'(\xi_\sigma)(\delta\xi_\sigma) = 0
\end{aligned}$$

Error spliting

$$\begin{aligned}
I(q,u) - I(q_\sigma,u_\sigma) = I(q,u) - I(q_k,u_k)&\\
+ I(q_k,u_k) - I(q_{kh},u_{kh})&\\
+ I(q_{kh},u_{kh}) - I(q_\sigma,u_\sigma)&
\end{aligned}$$

→ Goal: $I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_l$

→ Lagrange functional: $\mathcal{L}(q,u,z) = J(q,u) + f(z) - a_k(q,u)(z)$

$\xi = (q,u,z)$	$\in$	$Q \times X \times X$	$\mathcal{L}'(\xi)(\delta\xi)$	$= 0$
$\xi_k = (q_k, u_k, z_k)$	$\in$	$Q \times X_k \times X_k$	$\mathcal{L}'(\xi_k)(\delta\xi_k)$	$= 0$
$\xi_{kh} = (q_{kh}, u_{kh}, z_{kh})$	$\in$	$Q \times X_{kh} \times X_{kh}$	$\mathcal{L}'(\xi_{kh})(\delta\xi_{kh})$	$= 0$
$\xi_\sigma = (q_\sigma, u_\sigma, z_\sigma)$	$\in$	$Q_l \times X_{kh} \times X_{kh}$	$\mathcal{L}'(\xi_\sigma)(\delta\xi_\sigma)$	$= 0$

Error spliting

$$\begin{aligned} I(q,u) - I(q_\sigma,u_\sigma) = I(q,u) - I(q_k,u_k) \\ + I(q_k,u_k) - I(q_{kh},u_{kh}) \\ + I(q_{kh},u_{kh}) - I(q_\sigma,u_\sigma) \end{aligned}$$

→ Goal: $I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_l$

→ Lagrange functional: $\mathcal{L}(q,u,z) = J(q,u) + f(z) - a_k(q,u)(z)$

$\xi = (q,u,z)$	$\in$	$Q \times X \times X$	$\mathcal{L}'(\xi)(\delta\xi)$	$= 0$
$\xi_k = (q_k,u_k,z_k)$	$\in$	$Q \times X_k \times X_k$	$\mathcal{L}'(\xi_k)(\delta\xi_k)$	$= 0$
$\xi_{kh} = (q_{kh},u_{kh},z_{kh})$	$\in$	$Q \times X_{kh} \times X_{kh}$	$\mathcal{L}'(\xi_{kh})(\delta\xi_{kh})$	$= 0$
$\xi_\sigma = (q_\sigma,u_\sigma,z_\sigma)$	$\in$	$Q_l \times X_{kh} \times X_{kh}$	$\mathcal{L}'(\xi_\sigma)(\delta\xi_\sigma)$	$= 0$

Error spliting

$$\begin{aligned} I(q,u) - I(q_\sigma,u_\sigma) = I(q,u) - I(q_k,u_k) & \\ + I(q_k,u_k) - I(q_{kh},u_{kh}) & \\ + I(q_{kh},u_{kh}) - I(q_\sigma,u_\sigma) & \end{aligned}$$

→ Goal: $I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_l$

Error estimation w.r.t the cost functional

$$J(q,u) - J(q_k,u_k) = \frac{1}{2}\rho_u(x_\sigma)(u - \tilde{u}_k) + \frac{1}{2}\rho_z(x_\sigma)(z - \tilde{z}_k) + R_k$$

$$J(q_k,u_k) - J(q_{kh},u_{kh}) = \frac{1}{2}\rho_u(x_\sigma)(u_k - \tilde{u}_{kh}) + \frac{1}{2}\rho_z(x_\sigma)(z_k - \tilde{z}_{kh}) + R_{kh}$$

$$J(q_{kh},u_{kh}) - J(q_\sigma,u_\sigma) = \frac{1}{2}\rho_q(x_\sigma)(q_{kh} - \tilde{q}_\sigma) + R_\sigma$$

- Residuals:

$$\rho_u(x_\sigma)(\phi) = J'_u(q_\sigma,u_\sigma)(\phi) - a'_{k,u}(q_\sigma,u_\sigma)(\phi, z_\sigma)$$
$$\rho_z(x_\sigma)(\phi) = f(\phi) - a_k(q_\sigma,u_\sigma)(\phi)$$
$$\rho_q(x_\sigma)(\phi) = J'_q(q_\sigma,u_\sigma)(\phi) - a'_{k,q}(q_\sigma,u_\sigma)(\phi, z_\sigma)$$

- Remainder terms: $R_k, R_{kh}, R_\sigma \to \mathcal{O}(\|\text{error}\|^3)$

→ Goal: $I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_l$

Error estimation w.r.t the cost functional

$$J(q,u) - J(q_k, u_k) = \frac{1}{2}\rho_u(x_\sigma)(u - \tilde{u}_k) + \frac{1}{2}\rho_z(x_\sigma)(z - \tilde{z}_k) + R_k$$
$$J(q_k,u_k) - J(q_{kh}, u_{kh}) = \frac{1}{2}\rho_u(x_\sigma)(u_k - \tilde{u}_{kh}) + \frac{1}{2}\rho_z(x_\sigma)(z_k - \tilde{z}_{kh}) + R_{kh}$$
$$J(q_{kh},u_{kh}) - J(q_\sigma, u_\sigma) = \frac{1}{2}\rho_q(x_\sigma)(q_{kh} - \tilde{q}_\sigma) + R_\sigma$$

- Residuals:

$$\rho_u(x_\sigma)(\phi) = J'_u(q_\sigma, u_\sigma)(\phi) - a'_{k,u}(q_\sigma, u_\sigma)(\phi, z_\sigma)$$
$$\rho_z(x_\sigma)(\phi) = f(\phi) - a_k(q_\sigma, u_\sigma)(\phi)$$
$$\rho_q(x_\sigma)(\phi) = J'_q(q_\sigma, u_\sigma)(\phi) - a'_{k,q}(q_\sigma, u_\sigma)(\phi, z_\sigma)$$

- Remainder terms: $R_k, R_{kh}, R_\sigma \to \mathcal{O}(\|\text{error}\|^3)$

→ Goal: $I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_l$

Error estimation w.r.t the cost functional

$$J(q,u) - J(q_k, u_k) = \frac{1}{2}\rho_u(x_\sigma)(u - \tilde{u}_k) + \frac{1}{2}\rho_z(x_\sigma)(z - \tilde{z}_k) + R_k$$

$$J(q_k,u_k) - J(q_{kh}, u_{kh}) = \frac{1}{2}\rho_u(x_\sigma)(u_k - \tilde{u}_{kh}) + \frac{1}{2}\rho_z(x_\sigma)(z_k - \tilde{z}_{kh}) + R_{kh}$$

$$J(q_{kh},u_{kh}) - J(q_\sigma, u_\sigma) = \frac{1}{2}\rho_q(x_\sigma)(q_{kh} - \tilde{q}_\sigma) + R_\sigma$$

- Residuals:

$$\rho_u(x_\sigma)(\phi) = J'_u(q_\sigma, u_\sigma)(\phi) - a'_{k,u}(q_\sigma, u_\sigma)(\phi, z_\sigma)$$
$$\rho_z(x_\sigma)(\phi) = f(\phi) - a_k(q_\sigma, u_\sigma)(\phi)$$
$$\rho_q(x_\sigma)(\phi) = J'_q(q_\sigma, u_\sigma)(\phi) - a'_{k,q}(q_\sigma, u_\sigma)(\phi, z_\sigma)$$

- Remainder terms: $R_k, R_{kh}, R_\sigma \to \mathcal{O}(\|\text{error}\|^3)$

→ Goal: $I(q,u) - I(q_\sigma, u_\sigma) \approx \eta_k + \eta_h + \eta_l$

Error estimation w.r.t the cost functional

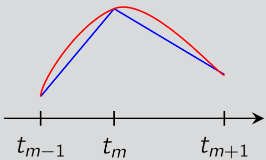
$$J(q,u) - J(q_k,u_k) = \frac{1}{2}\rho_u(x_\sigma)(u - \tilde{u}_k) + \frac{1}{2}\rho_z(x_\sigma)(z - \tilde{z}_k) + R_k$$
$$J(q_k,u_k) - J(q_{kh},u_{kh}) = \frac{1}{2}\rho_u(x_\sigma)(u_k - \tilde{u}_{kh}) + \frac{1}{2}\rho_z(x_\sigma)(z_k - \tilde{z}_{kh}) + R_{kh}$$
$$J(q_{kh},u_{kh}) - J(q_\sigma,u_\sigma) = \frac{1}{2}\rho_q(x_\sigma)(q_{kh} - \tilde{q}_\sigma) + R_\sigma$$

- Residuals:

$$\rho_u(x_\sigma)(\phi) = J'_u(q_\sigma,u_\sigma)(\phi) - a'_{k,u}(q_\sigma,u_\sigma)(\phi, z_\sigma)$$
$$\rho_z(x_\sigma)(\phi) = f(\phi) - a_k(q_\sigma,u_\sigma)(\phi)$$
$$\rho_q(x_\sigma)(\phi) = J'_q(q_\sigma,u_\sigma)(\phi) - a'_{k,q}(q_\sigma,u_\sigma)(\phi, z_\sigma)$$

- Remainder terms: $R_k, R_{kh}, R_\sigma \to \mathcal{O}(\|\text{error}\|^3)$

$u - \tilde{u}_k \approx i_{2k}^{(2)} u_\sigma - u_\sigma$

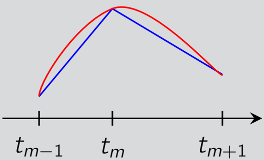


$z - \tilde{z}_k \approx i_{2k}^{(1)} z_\sigma - z_\sigma$

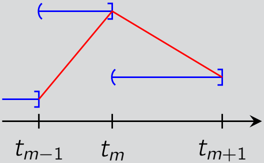
$u_k - \tilde{u}_{kh} \approx i_{2h}^{(2)} u_\sigma - u_\sigma$

Approximation of weights

$u - \tilde{u}_k \approx i_{2k}^{(2)} u_\sigma - u_\sigma$

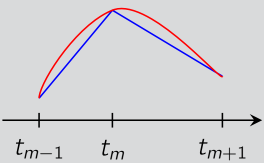


$z - \tilde{z}_k \approx i_{2k}^{(1)} z_\sigma - z_\sigma$

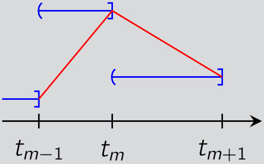


$u_k - \tilde{u}_{kh} \approx i_{2h}^{(2)} u_\sigma - u_\sigma$

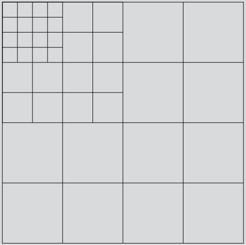
$u - \tilde{u}_k \approx i_{2k}^{(2)} u_\sigma - u_\sigma$



$z - \tilde{z}_k \approx i_{2k}^{(1)} z_\sigma - z_\sigma$



$u_k - \tilde{u}_{kh} \approx i_{2h}^{(2)} u_\sigma - u_\sigma$



Numerical effort

- Error estimation w.r.t the cost functional
  - Evaluation of residuals
  - Approximation of weights

- Error estimation w.r.t a quantity of interest
  - Solution of a linear problem ($\sim$ Newton step)

$$\chi = (p, v, y) \; : \; \mathcal{L}''(q, u, z)\chi = -I'_q(q, u) - I'_u(q, u)$$

  - Evaluation of residuals
  - Approximation of weights

Numerical effort

- Error estimation w.r.t the cost functional
  - Evaluation of residuals
  - Approximation of weights

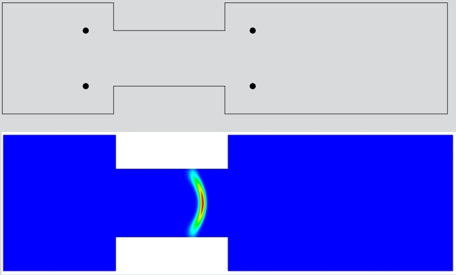
- Error estimation w.r.t a quantity of interest
  - Solution of a linear problem ($\sim$ Newton step)

$$\chi = (p, v, y) \ : \ \mathcal{L}''(q, u, z)\chi = -I'_q(q, u) - I'_u(q, u)$$

  - Evaluation of residuals
  - Approximation of weights

Outline

Configuration



Mathematical model

$$\partial_t \theta - \Delta\theta = \omega \qquad \text{in } (0,T)\times\Omega$$

$$\partial_t y - \frac{1}{L}\Delta y = -\omega \qquad \text{in } (0,T)\times\Omega$$

$$\omega = \frac{\beta^2}{2\gamma}\, y \exp\left(\frac{\beta(\theta-1)}{1+\alpha(\theta-1)}\right)$$

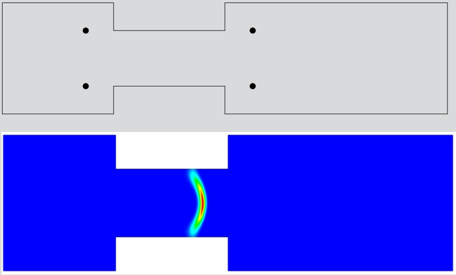
+ Initial and boundary conditions

Goal:

Estimation of Arrhenius parameters α using measurements $\theta(T,\xi_i)$ and $Y(T,\xi_i)$

$\rightarrow$ Quantity of interest: $I(\alpha) = \alpha$

Configuration



Mathematical model

$$\partial_t \theta - \Delta\theta = \omega \quad \text{in } (0, T) \times \Omega$$

$$\partial_t y - \frac{1}{L}\Delta y = -\omega \quad \text{in } (0, T) \times \Omega$$

$$\omega = \frac{\beta^2}{2\gamma}\, y \exp\left(\frac{\beta(\theta - 1)}{1 + \alpha(\theta - 1)}\right)$$

+ Initial and boundary conditions

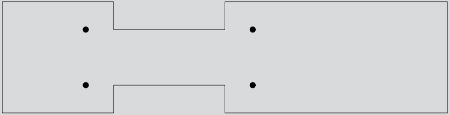
Goal:

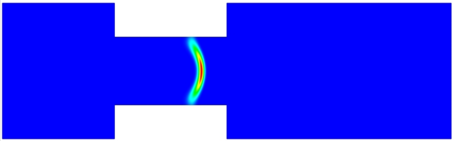
Estimation of Arrhenius parameters α using measurements $\theta(T, \xi_i)$ and $Y(T, \xi_i)$

$\rightarrow$ Quantity of interest: $I(\alpha) = \alpha$

Numerical example

Configuration





Mathematical model

$$\partial_t \theta - \Delta\theta = \omega \qquad \text{in } (0, T) \times \Omega$$

$$\partial_t y - \frac{1}{L}\Delta y = -\omega \qquad \text{in } (0, T) \times \Omega$$

$$\omega = \frac{\beta^2}{2\gamma}\, y \exp\left(\frac{\beta(\theta - 1)}{1 + \alpha(\theta - 1)}\right)$$

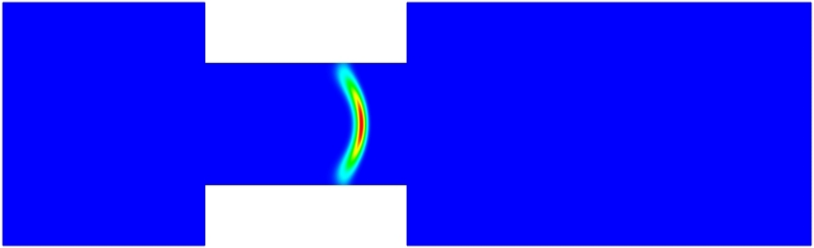
+ Initial and boundary conditions

Goal:

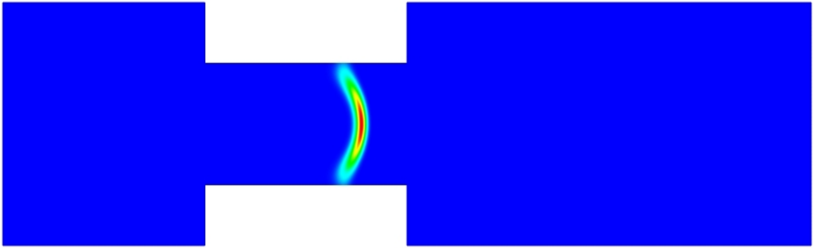
Estimation of Arrhenius parameters α using measurements $\theta(T, \xi_i)$ and $Y(T, \xi_i)$

$\rightarrow$ Quantity of interest: $I(\alpha) = \alpha$

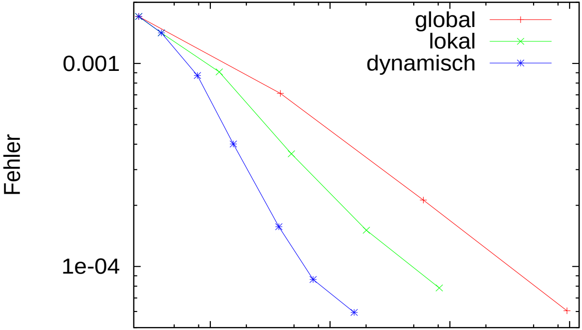
Simulation:



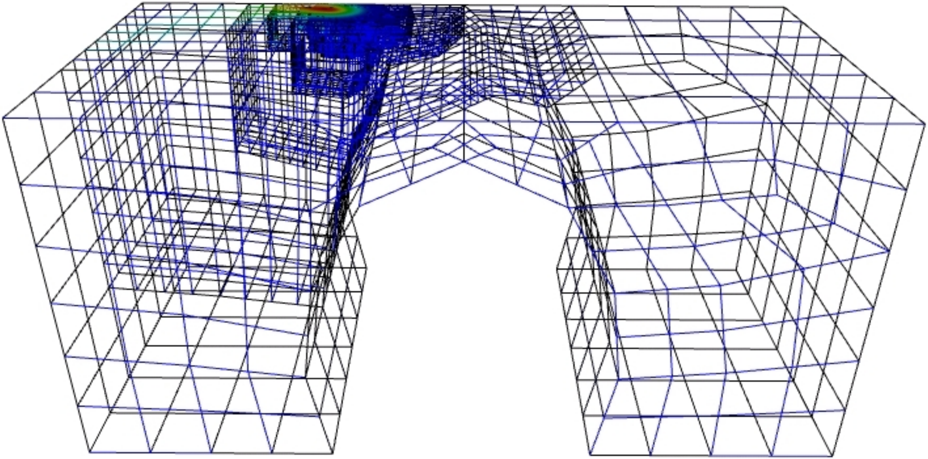
Optimization (Parameter identification):



M	$N_{\max}$	η_k	η_h	$\eta_k + \eta_h$	$J(u) - J(u_{kh})$	I_{eff}
256	985	$-7.8 \cdot 10^{-04}$	$4.1 \cdot 10^{-05}$	$-7.4 \cdot 10^{-04}$	$-1.7 \cdot 10^{-03}$	2.29
396	985	$-7.4 \cdot 10^{-04}$	$2.0 \cdot 10^{-04}$	$-5.4 \cdot 10^{-04}$	$-1.4 \cdot 10^{-03}$	2.63
616	1427	$-2.5 \cdot 10^{-04}$	$-3.3 \cdot 10^{-04}$	$-5.8 \cdot 10^{-04}$	$-8.7 \cdot 10^{-04}$	1.51
872	2309	$-1.0 \cdot 10^{-04}$	$-1.4 \cdot 10^{-04}$	$-2.4 \cdot 10^{-04}$	$-4.0 \cdot 10^{-04}$	1.64
1370	3927	$-5.0 \cdot 10^{-05}$	$-6.8 \cdot 10^{-05}$	$-1.2 \cdot 10^{-04}$	$-1.6 \cdot 10^{-04}$	1.33
1528	6927	$-4.6 \cdot 10^{-05}$	$-2.8 \cdot 10^{-05}$	$-7.4 \cdot 10^{-05}$	$-8.6 \cdot 10^{-05}$	1.17
1772	14683	$-4.0 \cdot 10^{-05}$	$-1.1 \cdot 10^{-06}$	$-5.2 \cdot 10^{-05}$	$-5.9 \cdot 10^{-05}$	1.15



3D – laser surface hardening of steel



Conclusions

→ Space-time finite elements methods

- optimize-then-discretize = discretize-then-optimize
- exact derivatives on the discrete level

→ Error estimation w.r.t. to a given quantity of interest

- separation of time, space, control error
- dynamic meshes

→ Extensions: problems with inequality constraints

- control constraints
- state constraints