Adaptive Optimal Control of Diffusion-Convection-Reaction Equations

Bülent Karasözen, METU

Joint work with
Hamdullah Yücel (METU) & Mathias Heinkenschloss (Rice University)

March 23, 2012

Outline

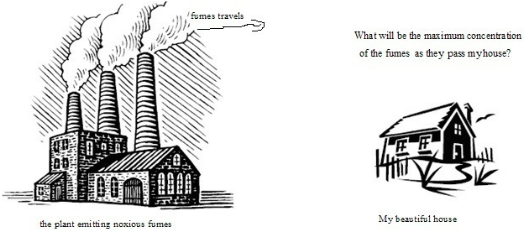
Outline

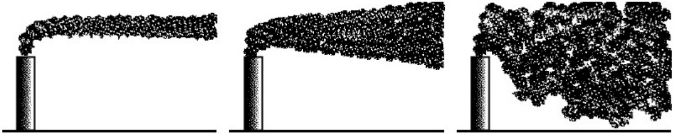
Outline

Outline

Outline

Air Pollution





(a) stable (b) neutral (c) instable

Figure: Illustrative example of the effects of air stability on a pollutants plume emitted by a chimney.

L. Dede', and A. Quarteroni, Optimal control and numerical adaptivity for advection diffusion-equations, *2005.*

Water Pollution

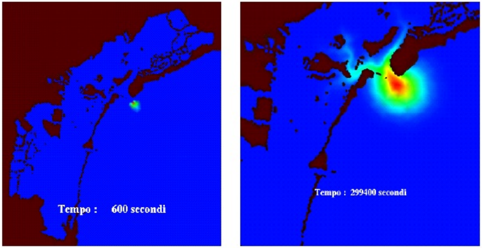


Figure 1.10: Concentration of a pollutant released in front of the Venice Lagoon at two different time steps.

A. Quarteroni, L. Bonaventura, L. Ded'e, E. Miglio, A. Quaini, M. Restelli, G. Rozza, and F. Saleri, Modellistica matematica in problemi ambientali *, 2006.*

Cooling of Steel Profiles

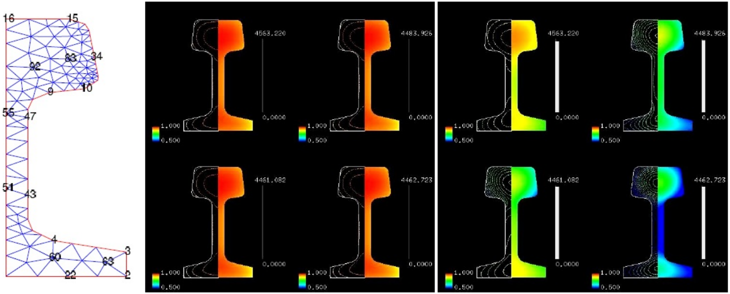


Fig. 1 The domain Ω is a half cross section of a rail profile. Different cost functionals (e.g. different output matrices) produce different final temperatures according to experimental observation.

J. Saak, and P. Benner, Efficient numerical solution of the LQR-problem for the heat equation*, 2004.*

Optimal Control Problem

- $\Omega \in \mathbb{R}^d (d = 2,3)$ with $\Gamma = \partial\Omega$ is bounded, open, and convex
- The linear-quadratic optimal control problem

$$\text{minimize } J(y,u) := \frac{1}{2}\int_\Omega (y(x) - y_d(x))^2 dx + \frac{\omega}{2}\int_\Omega u(x)^2 dx$$

subject to

$$\begin{aligned} -\varepsilon\Delta y(x) + \beta(x)\cdot\nabla y(x) + r(x)y(x) &= f(x) + u(x), \quad && x \in \Omega, \\ y(x) &= g_D(x), && x \in \Gamma, \end{aligned}$$

- source function $f \in L^2(\Omega)$, desired state $y_d \in L^2(\Omega)$, convection term $\beta(x)$, reaction term $r(x)$, diffusion term $0 < \varepsilon \ll 1$ and the regularization parameter $0 < \omega \leq 1$
- y : the state and u : the control

Optimal Control Problem

- $\Omega \in \mathbb{R}^d (d = 2, 3)$ with $\Gamma = \partial\Omega$ is bounded, open, and convex
- The linear-quadratic optimal control problem

$$\text{minimize } J(y,u) := \frac{1}{2}\int_\Omega (y(x) - y_d(x))^2 dx + \frac{\omega}{2}\int_\Omega u(x)^2 dx$$

subject to

$$\begin{aligned} -\varepsilon\Delta y(x) + \beta(x)\cdot\nabla y(x) + r(x)y(x) &= f(x) + u(x), & x \in \Omega, \\ y(x) &= g_D(x), & x \in \Gamma, \end{aligned}$$

- source function $f \in L^2(\Omega)$, desired state $y_d \in L^2(\Omega)$, convection term $\beta(x)$, reaction term $r(x)$, diffusion term $0 < \varepsilon \ll 1$ and the regularization parameter $0 < \omega \leq 1$
- y : the state and u : the control

Optimal Control Problem

- $\Omega \in \mathbb{R}^d (d = 2, 3)$ with $\Gamma = \partial\Omega$ is bounded, open, and convex
- The linear-quadratic optimal control problem

$$\text{minimize} \;\; J(y,u) := \frac{1}{2}\int_\Omega (y(x) - y_d(x))^2 dx + \frac{\omega}{2}\int_\Omega u(x)^2 dx$$

subject to

$$\begin{aligned} -\varepsilon\Delta y(x) + \beta(x)\cdot\nabla y(x) + r(x)y(x) &= f(x) + u(x), \quad x \in \Omega, \\ y(x) &= g_D(x), \quad\quad\quad x \in \Gamma, \end{aligned}$$

- source function $f \in L^2(\Omega)$, desired state $y_d \in L^2(\Omega)$, convection term $\beta(x)$, reaction term $r(x)$, diffusion term $0 < \varepsilon \ll 1$ and the regularization parameter $0 < \omega \leq 1$
- y : the state and u : the control

Optimal Control Problem

- $\Omega \in \mathbb{R}^d (d = 2,3)$ with $\Gamma = \partial\Omega$ is bounded, open, and convex
- The linear-quadratic optimal control problem

$$\text{minimize } J(y,u) := \frac{1}{2}\int_\Omega (y(x) - y_d(x))^2 dx + \frac{\omega}{2}\int_\Omega u(x)^2 dx$$

subject to

$$\begin{aligned} -\varepsilon\Delta y(x) + \beta(x)\cdot\nabla y(x) + r(x)y(x) &= f(x) + u(x), \quad && x \in \Omega, \\ y(x) &= g_D(x), && x \in \Gamma, \end{aligned}$$

- source function $f \in L^2(\Omega)$, desired state $y_d \in L^2(\Omega)$, convection term $\beta(x)$, reaction term $r(x)$, diffusion term $0 < \varepsilon \ll 1$ and the regularization parameter $0 < \omega \leq 1$
- y : the state and u : the control

Weak formulation

- State space $Y = \{y \in H^1(\Omega) : \; y = g_D \text{ on } \Gamma\}$,
- Control space $U = L^2(\Omega)$,
- Space of the test functions $V = \{\upsilon \in H^1(\Omega) : \; \upsilon = 0 \text{ on } \Gamma\}$

Weak form of the state equation

$$a(y,\upsilon) + b(u,\upsilon) = (f,\upsilon), \qquad \forall \upsilon \in V,$$

$$\begin{aligned} a(y,\upsilon) &= \int_\Omega (\varepsilon \nabla y \cdot \nabla \upsilon + \beta \cdot \nabla y \upsilon + r y \upsilon)\, dx, \\ b(u,\upsilon) &= -\int_\Omega u\upsilon \, dx, \qquad (f,\upsilon) = \int_\Omega f \upsilon \, dx. \end{aligned}$$

Weak formulation

- State space $Y = \{y \in H^1(\Omega) : \; y = g_D \text{ on } \Gamma\}$,
- Control space $U = L^2(\Omega)$,
- Space of the test functions $V = \{\upsilon \in H^1(\Omega) : \; \upsilon = 0 \text{ on } \Gamma\}$

Weak form of the state equation

$$a(y,\upsilon) + b(u,\upsilon) = (f,\upsilon), \qquad \forall \upsilon \in V,$$

$$\begin{aligned} a(y,\upsilon) &= \int_\Omega (\varepsilon \nabla y \cdot \nabla \upsilon + \beta \cdot \nabla y \upsilon + r y \upsilon)\, dx, \\ b(u,\upsilon) &= -\int_\Omega u\upsilon\, dx, \qquad (f,\upsilon) = \int_\Omega f\upsilon\, dx. \end{aligned}$$

Weak formulation

- State space $Y = \{y \in H^1(\Omega): \ y = g_D$ on $\Gamma\}$,
- Control space $U = L^2(\Omega)$,
- Space of the test functions $V = \{\upsilon \in H^1(\Omega): \ \upsilon = 0$ on $\Gamma\}$

Weak form of the state equation

$$a(y,\upsilon) + b(u,\upsilon) = (f,\upsilon), \qquad \forall \upsilon \in V,$$

$$\begin{aligned} a(y,\upsilon) &= \int_\Omega (\varepsilon \nabla y \cdot \nabla \upsilon + \beta \cdot \nabla y \upsilon + ry\upsilon)\, dx, \\ b(u,\upsilon) &= -\int_\Omega u\upsilon\, dx, \qquad (f,\upsilon) = \int_\Omega f\upsilon\, dx. \end{aligned}$$

Optimality Conditions

- Optimal control problem in variational form:

$$\text{minimize } J(y,u) := \frac{1}{2}\|y-y_d\|_\Omega^2 + \frac{\omega}{2}\|u\|_\Omega^2$$

$$\text{s.t. } a(y,\upsilon)+b(u,\upsilon) = (f,\upsilon), \ \forall (y,u,\upsilon) \in Y \times U \times V.$$

- Lagrangian

$$L(y,u,p) = \frac{1}{2}\|y-y_d\|_{L^2(\Omega)}^2 + \frac{\omega}{2}\|u\|_{L^2(\Omega)}^2 + a(y,p) + b(u,p) - (f,p).$$

- First order optimality conditions: $\nabla L(y,u,p) = 0$

$$\begin{aligned} a(\psi,p) &= -(y-y_d,\psi), && \forall \psi \in V, \\ b(w,p)+\omega(u,w) &= 0, && \forall w \in U, \\ a(y,\upsilon)+b(u,\upsilon) &= (f,\upsilon), && \forall \upsilon \in V. \end{aligned}$$

Optimality Conditions

- Optimal control problem in variational form:

$$\text{minimize } J(y,u) := \frac{1}{2}\|y - y_d\|_\Omega^2 + \frac{\omega}{2}\|u\|_\Omega^2$$

$$\text{s.t. } a(y,\upsilon) + b(u,\upsilon) = (f,\upsilon), \; \forall(y,u,\upsilon) \in Y \times U \times V.$$

- Lagrangian

$$L(y,u,p) = \frac{1}{2}\|y - y_d\|_{L^2(\Omega)}^2 + \frac{\omega}{2}\|u\|_{L^2(\Omega)}^2 + a(y,p) + b(u,p) - (f,p).$$

- First order optimality conditions: $\nabla L(y,u,p) = 0$

$$\begin{aligned} a(\psi,p) &= -(y - y_d,\psi), && \forall \psi \in V, \\ b(w,p) + \omega(u,w) &= 0, && \forall w \in U, \\ a(y,\upsilon) + b(u,\upsilon) &= (f,\upsilon), && \forall \upsilon \in V. \end{aligned}$$

Optimality Conditions

- Optimal control problem in variational form:

$$\text{minimize } J(y,u) := \frac{1}{2}\|y-y_d\|_\Omega^2 + \frac{\omega}{2}\|u\|_\Omega^2$$

$$\text{s.t. } a(y,\upsilon)+b(u,\upsilon) = (f,\upsilon), \ \forall (y,u,\upsilon) \in Y \times U \times V.$$

- Lagrangian

$$L(y,u,p) = \frac{1}{2}\|y-y_d\|_{L^2(\Omega)}^2 + \frac{\omega}{2}\|u\|_{L^2(\Omega)}^2 + a(y,p) + b(u,p) - (f,p).$$

- First order optimality conditions: $\nabla L(y,u,p) = 0$

$$\begin{aligned} a(\psi,p) &= -(y-y_d,\psi), && \forall \psi \in V, \\ b(w,p)+\omega(u,w) &= 0, && \forall w \in U, \\ a(y,\upsilon)+b(u,\upsilon) &= (f,\upsilon), && \forall \upsilon \in V. \end{aligned}$$

Optimality Conditions ...

$$\begin{aligned} a(\psi,p) &= -(y-y_d,\psi), & \forall \psi \in V, \\ b(w,p)+\omega(u,w) &= 0, & \forall w \in U, \\ a(y,\upsilon)+b(u,\upsilon) &= (f,\upsilon), & \forall \upsilon \in V. \end{aligned}$$

- adjoint equation

$$\begin{aligned} -\varepsilon\nabla p(x) - \beta(x)\cdot\nabla p(x) + (r(x)-\nabla\cdot\beta(x))p(x) &= -(y(x)-y_d(x)), \\ p(x) &= 0, \end{aligned}$$

- gradient equation

$$p(x) = \omega u(x).$$

- state equation

$$\begin{aligned} -\varepsilon\Delta y(x) + \beta(x)\cdot\nabla y(x) + r(x)y(x) &= f(x)+u(x), \\ y(x) &= g_D(x), \end{aligned}$$

Optimality Conditions ...

$$\begin{aligned} a(\psi,p) &= -(y-y_d,\psi), & \forall\psi\in V, \\ b(w,p)+\omega(u,w) &= 0, & \forall w\in U, \\ a(y,\upsilon)+b(u,\upsilon) &= (f,\upsilon), & \forall\upsilon\in V. \end{aligned}$$

- adjoint equation

$$\begin{aligned} -\varepsilon\nabla p(x)-\beta(x)\cdot\nabla p(x)+(r(x)-\nabla\cdot\beta(x))p(x) &= -(y(x)-y_d(x)), \\ p(x) &= 0, \end{aligned}$$

- gradient equation

$$p(x)=\omega u(x).$$

- state equation

$$\begin{aligned} -\varepsilon\Delta y(x)+\beta(x)\cdot\nabla y(x)+r(x)y(x) &= f(x)+u(x), \\ y(x) &= g_D(x), \end{aligned}$$

Optimality Conditions ...

$$
\begin{aligned}
a(\psi,p) &= -(y-y_d,\psi), && \forall \psi \in V,\\
b(w,p)+\omega(u,w) &= 0, && \forall w \in U,\\
a(y,\upsilon)+b(u,\upsilon) &= (f,\upsilon), && \forall \upsilon \in V.
\end{aligned}
$$

- adjoint equation

$$
\begin{aligned}
-\varepsilon\nabla p(x) - \beta(x)\cdot\nabla p(x) + (r(x)-\nabla\cdot\beta(x))p(x) &= -(y(x)-y_d(x)),\\
p(x) &= 0,
\end{aligned}
$$

- gradient equation

$$
p(x) = \omega u(x).
$$

- state equation

$$
\begin{aligned}
-\varepsilon\Delta y(x) + \beta(x)\cdot\nabla y(x) + r(x)y(x) &= f(x)+u(x),\\
y(x) &= g_D(x),
\end{aligned}
$$

Existence & uniqueness

Assumptions:

- $f, y_D \in L^2(\Omega), g_D \in H^{3/2}(\Gamma)$,
- $0 < \varepsilon, \beta(x) \in W^{1,\infty}(\Omega)^2, 0 < \omega$ and $r \in L^\infty(\Omega)$,
- $r(x) - \frac{1}{2}\nabla \cdot \beta(x) \geq r_0 \geq 0$,
- $|| - \nabla \cdot \beta(x) + r(x)||_{L^\infty(\Omega)} \leq c_* r_0$.

$(y,u) \in Y \times U$ solve the optimal control problem if and only if $(y,u,p) \in Y \times U \times Y$ is unique solution for the following optimality system [Lions, Tröltzsch]:

$$\begin{aligned} a(\psi,p) + (y,\psi) &= (y_d,\psi), & \forall \psi \in Y, \\ b(w,p) + \omega(u,w) &= 0, & \forall w \in U, \\ a(y,\upsilon) + b(u,\upsilon) &= (f,\upsilon), & \forall \upsilon \in Y. \end{aligned}$$

Existence & uniqueness

Assumptions:

- $f, y_D \in L^2(\Omega), g_D \in H^{3/2}(\Gamma)$,
- $0 < \varepsilon, \beta(x) \in W^{1,\infty}(\Omega)^2, 0 < \omega$ and $r \in L^\infty(\Omega)$,
- $r(x) - \frac{1}{2}\nabla \cdot \beta(x) \geq r_0 \geq 0$,
- $|| - \nabla \cdot \beta(x) + r(x)||_{L^\infty(\Omega)} \leq c_* r_0$.

$(y,u) \in Y \times U$ solve the optimal control problem if and only if $(y,u,p) \in Y \times U \times Y$ is unique solution for the following optimality system [Lions, Tröltzsch]:

$$
\begin{aligned}
a(\psi,p) + (y,\psi) &= (y_d,\psi), && \forall \psi \in Y, \\
b(w,p) + \omega(u,w) &= 0, && \forall w \in U, \\
a(y,\upsilon) + b(u,\upsilon) &= (f,\upsilon), && \forall \upsilon \in Y.
\end{aligned}
$$

Existence & uniqueness

Assumptions:

- $f, y_D \in L^2(\Omega), g_D \in H^{3/2}(\Gamma)$,
- $0 < \varepsilon, \beta(x) \in W^{1,\infty}(\Omega)^2, 0 < \omega$ and $r \in L^\infty(\Omega)$,
- $r(x) - \frac{1}{2}\nabla \cdot \beta(x) \geq r_0 \geq 0$,
- $||-\nabla \cdot \beta(x) + r(x)||_{L^\infty(\Omega)} \leq c_* r_0$.

$(y,u) \in Y \times U$ solve the optimal control problem if and only if $(y,u,p) \in Y \times U \times Y$ is unique solution for the following optimality system [Lions, Tröltzsch]:

$$
\begin{aligned}
a(\psi,p) + (y,\psi) &= (y_d,\psi), && \forall \psi \in Y, \\
b(w,p) + \omega(u,w) &= 0, && \forall w \in U, \\
a(y,\upsilon) + b(u,\upsilon) &= (f,\upsilon), && \forall \upsilon \in Y.
\end{aligned}
$$

Existence & uniqueness

Assumptions:

- $f, y_D \in L^2(\Omega), g_D \in H^{3/2}(\Gamma)$,
- $0 < \varepsilon, \beta(x) \in W^{1,\infty}(\Omega)^2, 0 < \omega$ and $r \in L^\infty(\Omega)$,
- $r(x) - \frac{1}{2}\nabla \cdot \beta(x) \geq r_0 \geq 0$,
- $|| - \nabla \cdot \beta(x) + r(x)||_{L^\infty(\Omega)} \leq c_* r_0$.

$(y,u) \in Y \times U$ solve the optimal control problem if and only if $(y,u,p) \in Y \times U \times Y$ is unique solution for the following optimality system [Lions, Tröltzsch]:

$$\begin{aligned} a(\psi,p) + (y,\psi) &= (y_d,\psi), && \forall \psi \in Y, \\ b(w,p) + \omega(u,w) &= 0, && \forall w \in U, \\ a(y,\upsilon) + b(u,\upsilon) &= (f,\upsilon), && \forall \upsilon \in Y. \end{aligned}$$

Solution of the Optimal Control Problem

- Discretize then optimize,
- Optimize then discretize.

Solution of the Optimal Control Problem

- Discretize then optimize,
- Optimize then discretize.

Solution of the Optimal Control Problem

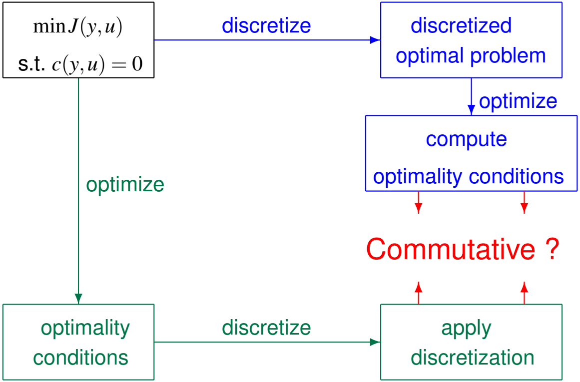
- Discretize then optimize,
- Optimize then discretize.

Solution of the Optimal Control Problem

- Discretize then optimize,
- Optimize then discretize.

Solution of the Optimal Control Problem

- Discretize then optimize,
- Optimize then discretize.



Numerical Methods

	Complex geometries	Higher-order accuracy and hp-adaptivity	Local mass Conservation
FDM	$\times$	$\checkmark$	$\checkmark$
FVM	$\checkmark$	$\times$	$\checkmark$
FEM	$\checkmark$	$\checkmark$	$\times$
DG	$\checkmark$	$\checkmark$	$\checkmark$

- Locally-higher order/flexible element as in FEM
- Local preservation of mass, energy as in FVM

Discontinuous Galerkin Finite Element Method

Numerical Methods

	Complex geometries	Higher-order accuracy and hp-adaptivity	Local mass Conservation
FDM	$\times$	$\checkmark$	$\checkmark$
FVM	$\checkmark$	$\times$	$\checkmark$
FEM	$\checkmark$	$\checkmark$	$\times$
DG	$\checkmark$	$\checkmark$	$\checkmark$

- Locally-higher order/flexible element as in FEM
- Local preservation of mass, energy as in FVM

Discontinuous Galerkin Finite Element Method

Numerical Methods

	Complex geometries	Higher-order accuracy and hp-adaptivity	Local mass Conservation
FDM	$\times$	$\checkmark$	$\checkmark$
FVM	$\checkmark$	$\times$	$\checkmark$
FEM	$\checkmark$	$\checkmark$	$\times$
DG	$\checkmark$	$\checkmark$	$\checkmark$

- Locally-higher order/flexible element as in FEM
- Local preservation of mass, energy as in FVM

Discontinuous Galerkin Finite Element Method

Numerical Methods

	Complex geometries	Higher-order accuracy and hp-adaptivity	Local mass Conservation
FDM	×	✓	✓
FVM	✓	×	✓
FEM	✓	✓	×
DG	✓	✓	✓

- Locally-higher order/flexible element as in FEM
- Local preservation of mass, energy as in FVM

Discontinuous Galerkin Finite Element Method

Numerical Methods

	Complex geometries	Higher-order accuracy and hp-adaptivity	Local mass Conservation
FDM	$\times$	$\checkmark$	$\checkmark$
FVM	$\checkmark$	$\times$	$\checkmark$
FEM	$\checkmark$	$\checkmark$	$\times$
DG	$\checkmark$	$\checkmark$	$\checkmark$

- Locally-higher order/flexible element as in FEM
- Local preservation of mass, energy as in FVM

Discontinuous Galerkin Finite Element Method

Discontinuous Galerkin Methods

DG is a class of FEMs which use discontinuous functions as the solution (and the test functions)

- Pros:
  - Flexibility for approximation order and complex meshes
  - Local conservation of physical quantities such as mass, momentum, and energy
  - Increase of the robustness and accuracy
  - Facilitation of parallelization
- Cons:
  - Large number of degrees of freedom
  - Ill-conditioning and denser global matrix with increasing approximation order

Discontinuous Galerkin Methods

DG is a class of FEMs which use discontinuous functions as the solution (and the test functions)

- Pros:
  - Flexibility for approximation order and complex meshes
  - Local conservation of physical quantities such as mass, momentum, and energy
  - Increase of the robustness and accuracy
  - Facilitation of parallelization
- Cons:
  - Large number of degrees of freedom
  - Ill-conditioning and denser global matrix with increasing approximation order

Discontinuous Galerkin Methods

DG is a class of FEMs which use discontinuous functions as the solution (and the test functions)

- Pros:
  - Flexibility for approximation order and complex meshes
  - Local conservation of physical quantities such as mass, momentum, and energy
  - Increase of the robustness and accuracy
  - Facilitation of parallelization
- Cons:
  - Large number of degrees of freedom
  - Ill-conditioning and denser global matrix with increasing approximation order

Discontinuous Galerkin Methods

DG is a class of FEMs which use discontinuous functions as the solution (and the test functions)

- Pros:
  - Flexibility for approximation order and complex meshes
  - Local conservation of physical quantities such as mass, momentum, and energy
  - Increase of the robustness and accuracy
  - Facilitation of parallelization
- Cons:
  - Large number of degrees of freedom
  - Ill-conditioning and denser global matrix with increasing approximation order

Discontinuous Galerkin Methods

DG is a class of FEMs which use discontinuous functions as the solution (and the test functions)

- Pros:
  - Flexibility for approximation order and complex meshes
  - Local conservation of physical quantities such as mass, momentum, and energy
  - Increase of the robustness and accuracy
  - Facilitation of parallelization
- Cons:
  - Large number of degrees of freedom
  - Ill-conditioning and denser global matrix with increasing approximation order

Discontinuous Galerkin Methods

DG is a class of FEMs which use discontinuous functions as the solution (and the test functions)

- Pros:
  - Flexibility for approximation order and complex meshes
  - Local conservation of physical quantities such as mass, momentum, and energy
  - Increase of the robustness and accuracy
  - Facilitation of parallelization
- Cons:
  - Large number of degrees of freedom
  - Ill-conditioning and denser global matrix with increasing approximation order

Discontinuous Galerkin Methods

DG is a class of FEMs which use discontinuous functions as the solution (and the test functions)

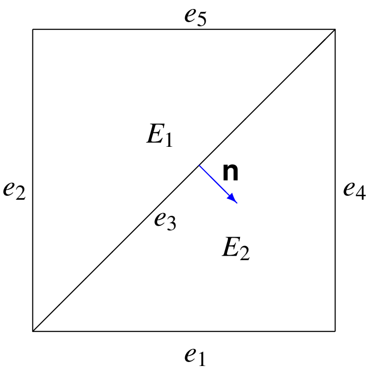
- Pros:
  - Flexibility for approximation order and complex meshes
  - Local conservation of physical quantities such as mass, momentum, and energy
  - Increase of the robustness and accuracy
  - Facilitation of parallelization
- Cons:
  - Large number of degrees of freedom
  - Ill-conditioning and denser global matrix with increasing approximation order

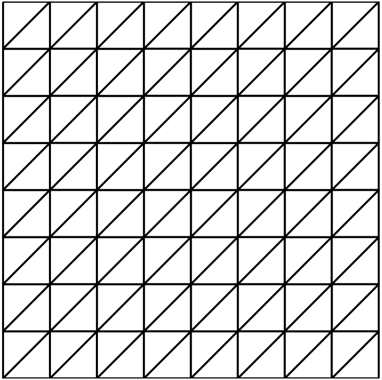
Discontinuous Galerkin Methods

DG is a class of FEMs which use discontinuous functions as the solution (and the test functions)

- Pros:
  - Flexibility for approximation order and complex meshes
  - Local conservation of physical quantities such as mass, momentum, and energy
  - Increase of the robustness and accuracy
  - Facilitation of parallelization
- Cons:
  - Large number of degrees of freedom
  - Ill-conditioning and denser global matrix with increasing approximation order

Mesh





DG Discretization

- ξ_h: partition of a domain with the conformity and shape regularity
- Γ_h: set of all edges and the interior edges and boundary edges are denoted by Γ_h^0 and Γ_h^∂, respectively
- An element and an edge are denoted by E and e, respectively
- $|E|$: the area of triangle E and $|e|$ denote the length of edge e
- The boundary edges are decomposed into the **inflow** and **outflow** edges;

$$\begin{aligned} \Gamma_h^- &= \{x \in \partial\Omega : \beta(x) \cdot \mathbf{n} < 0\}, \\ \Gamma_h^+ &= \{x \in \partial\Omega : \beta(x) \cdot \mathbf{n} \geq 0\}. \end{aligned}$$

DG Discretization

- ξ_h: partition of a domain with the conformity and shape regularity
- Γ_h: set of all edges and the interior edges and boundary edges are denoted by Γ_h^0 and Γ_h^∂, respectively
- An element and an edge are denoted by E and e, respectively
- $|E|$: the area of triangle E and $|e|$ denote the length of edge e
- The boundary edges are decomposed into the **inflow** and **outflow** edges;

$$\begin{aligned} \Gamma_h^- &= \{x \in \partial\Omega : \beta(x) \cdot \mathbf{n} < 0\}, \\ \Gamma_h^+ &= \{x \in \partial\Omega : \beta(x) \cdot \mathbf{n} \geq 0\}. \end{aligned}$$

DG Discretization

- ξ_h: partition of a domain with the conformity and shape regularity
- Γ_h: set of all edges and the interior edges and boundary edges are denoted by Γ_h^0 and Γ_h^∂, respectively
- An element and an edge are denoted by E and e, respectively
- $|E|$: the area of triangle E and $|e|$ denote the length of edge e
- The boundary edges are decomposed into the **inflow** and **outflow** edges;

$$\begin{aligned}\Gamma_h^- &= \{x \in \partial\Omega : \beta(x)\cdot\mathbf{n} < 0\},\\ \Gamma_h^+ &= \{x \in \partial\Omega : \beta(x)\cdot\mathbf{n} \geq 0\}.\end{aligned}$$

DG Discretization

- ξ_h: partition of a domain with the conformity and shape regularity
- Γ_h: set of all edges and the interior edges and boundary edges are denoted by Γ_h^0 and Γ_h^∂, respectively
- An element and an edge are denoted by E and e, respectively
- $|E|$: the area of triangle E and $|e|$ denote the length of edge e
- The boundary edges are decomposed into the **inflow** and **outflow** edges;

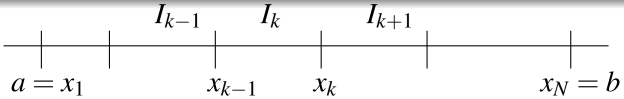
$$\Gamma_h^- = \{x \in \partial\Omega : \beta(x) \cdot \mathbf{n} < 0\},$$
$$\Gamma_h^+ = \{x \in \partial\Omega : \beta(x) \cdot \mathbf{n} \geq 0\}.$$

DG Discretization

- ξ_h: partition of a domain with the conformity and shape regularity
- Γ_h: set of all edges and the interior edges and boundary edges are denoted by Γ_h^0 and Γ_h^∂, respectively
- An element and an edge are denoted by E and e, respectively
- $|E|$: the area of triangle E and $|e|$ denote the length of edge e
- The boundary edges are decomposed into the **inflow** and **outflow** edges;

$$\begin{aligned} \Gamma_h^- &= \{x \in \partial\Omega : \beta(x) \cdot \mathbf{n} < 0\}, \\ \Gamma_h^+ &= \{x \in \partial\Omega : \beta(x) \cdot \mathbf{n} \geq 0\}. \end{aligned}$$

DG Discretization in 1 dimension



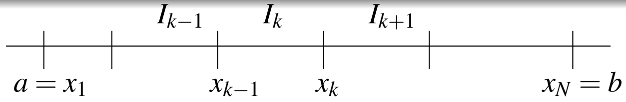
(continuous) FEM : $$v(x) = \sum_{i=1}^{N_{nodes}} v_i \varphi_i(x)$$

DGFEM : $$v(x) = \sum_{m=1}^{N_{el}} \sum_{j=1}^{N_{loc}} v_j^m \varphi_m^j(x)$$

N_{nodes}: number of nodes
N_{el} : number of elements
$N_{loc} = \frac{(k+1)(k+2)}{2}$ local dimension with approximation order k

- The jump operator $[v]_{x_k} = v|_{I_k}(x_k) - v_{I_{k+1}}(x_k)$
- The average operator $\{v\}_{x_k} = \frac{1}{2}(v_{I_k}(x_k) + v_{I_{k+1}}(x_k))$

DG Discretization in 1 dimension



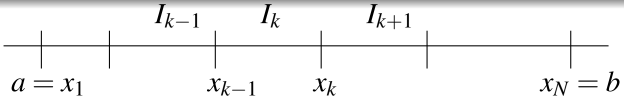
(continuous) FEM : $$v(x) = \sum_{i=1}^{N_{nodes}} v_i \varphi_i(x)$$

DGFEM : $$v(x) = \sum_{m=1}^{N_{el}} \sum_{j=1}^{N_{loc}} v_j^m \varphi_m^j(x)$$

N_{nodes}: number of nodes
N_{el} : number of elements
$N_{loc} = \frac{(k+1)(k+2)}{2}$ local dimension with approximation order k

- The jump operator $[v]_{x_k} = v|_{I_k}(x_k) - v_{I_{k+1}}(x_k)$
- The average operator $\{v\}_{x_k} = \frac{1}{2}(v_{I_k}(x_k) + v_{I_{k+1}}(x_k))$

DG Discretization in 1 dimension



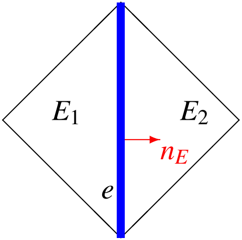
(continuous) FEM : $$v(x) = \sum_{i=1}^{N_{nodes}} v_i \varphi_i(x)$$

DGFEM : $$v(x) = \sum_{m=1}^{N_{el}} \sum_{j=1}^{N_{loc}} v_j^m \varphi_m^j(x)$$

N_{nodes}: number of nodes
N_{el} : number of elements
$N_{loc} = \frac{(k+1)(k+2)}{2}$ local dimension with approximation order k

- The jump operator $[v]_{x_k} = v|_{I_k}(x_k) - v_{I_{k+1}}(x_k)$
- The average operator $\{v\}_{x_k} = \frac{1}{2}(v_{I_k}(x_k) + v_{I_{k+1}}(x_k))$

DG Discretization in 2 dimensions

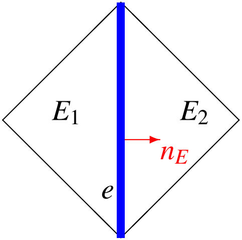


- The diffusion term is discretized by using
  - the jump operator $[v] = (v|_{E_1^e} - v|_{E_2^e})$
  - the average operator $\{v\} = \frac{1}{2}(v|_{E_1^e} + v|_{E_2^e})$

- The convection term is discretized by upwind discretization

$$y^+ = \begin{cases} y|_{E^1}, & \text{if } \beta \cdot n_e < 0, \\ y|_{E^2}, & \text{if } \beta \cdot n_e \geq 0, \end{cases} \qquad y^- = \begin{cases} y|_{E^2}, & \text{if } \beta \cdot n_e < 0, \\ y|_{E^1}, & \text{if } \beta \cdot n_e \geq 0. \end{cases}$$

DG Discretization in 2 dimensions



- The diffusion term is discretized by using
  - the jump operator $[\upsilon] = (\upsilon|_{E_1^e} - \upsilon|_{E_2^e})$
  - the average operator $\{\upsilon\} = \frac{1}{2}(\upsilon|_{E_1^e} + \upsilon|_{E_2^e})$

- The convection term is discretized by upwind discretization

$$y^+ = \begin{cases} y|_{E^1}, & \text{if } \beta \cdot n_e < 0, \\ y|_{E^2}, & \text{if } \beta \cdot n_e \geq 0, \end{cases} \qquad y^- = \begin{cases} y|_{E^2}, & \text{if } \beta \cdot n_e < 0, \\ y|_{E^1}, & \text{if } \beta \cdot n_e \geq 0. \end{cases}$$

Discretize then Optimize

DGFEM spaces on ξ_h

- $V_h = Y_h = \{y_h \in L^2(\Omega) \mid y|_E \in \mathbb{P}_n(E), \ \forall E \in \xi_h\}$,
- $U_h = \{u_h \in L^2(\Omega) \mid u|_E \in \mathbb{P}_m(E), \ \forall E \in \xi_h\}$.

Lagrangian of the discretized optimal control problem:

$$L_h(y_h,u_h,p_h) = \frac{1}{2}\sum_{E\in\xi_h}\|y_h - y_d\|_E^2 + \frac{\omega}{2}\sum_{E\in\xi_h}\|u_h\|_E^2 + a_h^s(y_h,p_h) + b_h(u_h,p_h) - l_h^s(p_h),$$

Optimality system of the discretized optimal control problem:

- discretized state equation

$$a_h^s(y_h,\upsilon_h) + b_h(u_h,\upsilon_h) = l_h^s(\upsilon_h), \qquad \forall \upsilon_h \in V_h$$

- discrete adjoint equation

$$a_h^s(\psi_h,p_h) = -(y_h - y_d,\psi_h), \qquad \forall \psi_h \in V_h$$

- discrete gradient equation

$$b_h(w_h,p_h) + \omega(u_h,w_h) = 0, \qquad \forall w_h \in U_h$$

Discretize then Optimize

DGFEM spaces on ξ_h

- $V_h = Y_h = \{y_h \in L^2(\Omega) \mid y|_E \in \mathbb{P}_n(E), \ \forall E \in \xi_h\}$,
- $U_h = \{u_h \in L^2(\Omega) \mid u|_E \in \mathbb{P}_m(E), \ \forall E \in \xi_h\}$.

Lagrangian of the discretized optimal control problem:

$$L_h(y_h,u_h,p_h) = \frac{1}{2}\sum_{E\in\xi_h} \|y_h - y_d\|_E^2 + \frac{\omega}{2}\sum_{E\in\xi_h} \|u_h\|_E^2 + a_h^s(y_h,p_h) + b_h(u_h,p_h) - l_h^s(p_h),$$

Optimality system of the discretized optimal control problem:

- discretized state equation

$$a_h^s(y_h, v_h) + b_h(u_h, v_h) = l_h^s(v_h), \qquad \forall v_h \in V_h$$

- discrete adjoint equation

$$a_h^s(\psi_h, p_h) = -(y_h - y_d, \psi_h), \qquad \forall \psi_h \in V_h$$

- discrete gradient equation

$$b_h(w_h, p_h) + \omega(u_h, w_h) = 0, \qquad \forall w_h \in U_h$$

Discretize then Optimize

DGFEM spaces on ξ_h

- $V_h = Y_h = \{y_h \in L^2(\Omega) \mid y|_E \in \mathbb{P}_n(E),\ \forall E \in \xi_h\}$,
- $U_h = \{u_h \in L^2(\Omega) \mid u|_E \in \mathbb{P}_m(E),\ \forall E \in \xi_h\}$.

Lagrangian of the discretized optimal control problem:

$$L_h(y_h,u_h,p_h) = \frac{1}{2}\sum_{E\in\xi_h}\|y_h - y_d\|_E^2 + \frac{\omega}{2}\sum_{E\in\xi_h}\|u_h\|_E^2 + a_h^s(y_h,p_h) + b_h(u_h,p_h) - l_h^s(p_h),$$

Optimality system of the discretized optimal control problem:

- discretized state equation

$$a_h^s(y_h, v_h) + b_h(u_h, v_h) = l_h^s(v_h), \qquad \forall v_h \in V_h$$

- discrete adjoint equation

$$a_h^s(\psi_h, p_h) = -(y_h - y_d, \psi_h), \qquad \forall \psi_h \in V_h$$

- discrete gradient equation

$$b_h(w_h, p_h) + \omega(u_h, w_h) = 0, \qquad \forall w_h \in U_h$$

Discretize then Optimize

DGFEM spaces on ξ_h

- $V_h = Y_h = \{y_h \in L^2(\Omega) \mid y|_E \in \mathbb{P}_n(E), \ \forall E \in \xi_h\}$,
- $U_h = \{u_h \in L^2(\Omega) \mid u|_E \in \mathbb{P}_m(E), \ \forall E \in \xi_h\}$.

Lagrangian of the discretized optimal control problem:

$$L_h(y_h, u_h, p_h) = \frac{1}{2} \sum_{E \in \xi_h} \|y_h - y_d\|_E^2 + \frac{\omega}{2} \sum_{E \in \xi_h} \|u_h\|_E^2 + a_h^s(y_h, p_h) + b_h(u_h, p_h) - l_h^s(p_h),$$

Optimality system of the discretized optimal control problem:

- discretized state equation

$$a_h^s(y_h, v_h) + b_h(u_h, v_h) = l_h^s(v_h), \qquad \forall v_h \in V_h$$

- discrete adjoint equation

$$a_h^s(\psi_h, p_h) = -(y_h - y_d, \psi_h), \qquad \forall \psi_h \in V_h$$

- discrete gradient equation

$$b_h(w_h, p_h) + \omega(u_h, w_h) = 0, \qquad \forall w_h \in U_h$$

Discretize then Optimize

DGFEM spaces on ξ_h

- $V_h = Y_h = \{y_h \in L^2(\Omega) \mid y|_E \in \mathbb{P}_n(E),\ \forall E \in \xi_h\}$,
- $U_h = \{u_h \in L^2(\Omega) \mid u|_E \in \mathbb{P}_m(E),\ \forall E \in \xi_h\}$.

Lagrangian of the discretized optimal control problem:

$$L_h(y_h,u_h,p_h) = \frac{1}{2}\sum_{E\in\xi_h} \|y_h - y_d\|_E^2 + \frac{\omega}{2}\sum_{E\in\xi_h} \|u_h\|_E^2 + a_h^s(y_h,p_h) + b_h(u_h,p_h) - l_h^s(p_h),$$

Optimality system of the discretized optimal control problem:

- discretized state equation

$$a_h^s(y_h,v_h) + b_h(u_h,v_h) = l_h^s(v_h), \qquad \forall v_h \in V_h$$

- discrete adjoint equation

$$a_h^s(\psi_h,p_h) = -(y_h - y_d, \psi_h), \qquad \forall \psi_h \in V_h$$

- discrete gradient equation

$$b_h(w_h,p_h) + \omega(u_h,w_h) = 0, \qquad \forall w_h \in U_h$$

Discretize then Optimize

DGFEM spaces on ξ_h

- $V_h = Y_h = \{y_h \in L^2(\Omega) \mid y|_E \in \mathbb{P}_n(E), \ \forall E \in \xi_h\}$,
- $U_h = \{u_h \in L^2(\Omega) \mid u|_E \in \mathbb{P}_m(E), \ \forall E \in \xi_h\}$.

Lagrangian of the discretized optimal control problem:

$$L_h(y_h,u_h,p_h) = \frac{1}{2}\sum_{E\in\xi_h} \|y_h - y_d\|_E^2 + \frac{\omega}{2}\sum_{E\in\xi_h} \|u_h\|_E^2 + a_h^s(y_h,p_h) + b_h(u_h,p_h) - l_h^s(p_h),$$

Optimality system of the discretized optimal control problem:

- discretized state equation

$$a_h^s(y_h,\upsilon_h) + b_h(u_h,\upsilon_h) = l_h^s(\upsilon_h), \qquad \forall \upsilon_h \in V_h$$

- discrete adjoint equation

$$a_h^s(\psi_h,p_h) = -(y_h - y_d, \psi_h), \qquad \forall \psi_h \in V_h$$

- discrete gradient equation

$$b_h(w_h,p_h) + \omega(u_h,w_h) = 0, \qquad \forall w_h \in U_h$$

Discretize then Optimize

DGFEM spaces on ξ_h

- $V_h = Y_h = \{y_h \in L^2(\Omega) \mid y|_E \in \mathbb{P}_n(E), \ \forall E \in \xi_h\}$,
- $U_h = \{u_h \in L^2(\Omega) \mid u|_E \in \mathbb{P}_m(E), \ \forall E \in \xi_h\}$.

Lagrangian of the discretized optimal control problem:

$$L_h(y_h,u_h,p_h) = \frac{1}{2}\sum_{E\in\xi_h} \|y_h - y_d\|_E^2 + \frac{\omega}{2}\sum_{E\in\xi_h} \|u_h\|_E^2 + a_h^s(y_h,p_h) + b_h(u_h,p_h) - l_h^s(p_h),$$

Optimality system of the discretized optimal control problem:

- discretized state equation

$$a_h^s(y_h, \upsilon_h) + b_h(u_h, \upsilon_h) = l_h^s(\upsilon_h), \qquad \forall \upsilon_h \in V_h$$

- discrete adjoint equation

$$a_h^s(\psi_h, p_h) = -(y_h - y_d, \psi_h), \qquad \forall \psi_h \in V_h$$

- discrete gradient equation

$$b_h(w_h, p_h) + \omega(u_h, w_h) = 0, \qquad \forall w_h \in U_h$$

Discretize then Optimize

DGFEM spaces on ξ_h

- $V_h = Y_h = \{y_h \in L^2(\Omega) \mid y|_E \in \mathbb{P}_n(E), \ \forall E \in \xi_h\}$,
- $U_h = \{u_h \in L^2(\Omega) \mid u|_E \in \mathbb{P}_m(E), \ \forall E \in \xi_h\}$.

Lagrangian of the discretized optimal control problem:

$$L_h(y_h,u_h,p_h) = \frac{1}{2}\sum_{E\in\xi_h} \|y_h - y_d\|_E^2 + \frac{\omega}{2}\sum_{E\in\xi_h} \|u_h\|_E^2 + a_h^s(y_h,p_h) + b_h(u_h,p_h) - l_h^s(p_h),$$

Optimality system of the discretized optimal control problem:

- discretized state equation

$$a_h^s(y_h,\upsilon_h) + b_h(u_h,\upsilon_h) = l_h^s(\upsilon_h), \qquad \forall \upsilon_h \in V_h$$

- discrete adjoint equation

$$a_h^s(\psi_h,p_h) = -(y_h - y_d, \psi_h), \qquad \forall \psi_h \in V_h$$

- discrete gradient equation

$$b_h(w_h,p_h) + \omega(u_h,w_h) = 0, \qquad \forall w_h \in U_h$$

Interior Penalty Galerkin Methods

$$
\begin{aligned}
a_h^s(y_h, \upsilon_h) \quad = \quad & \sum_{E\in\xi_h} (\varepsilon \nabla y_h, \nabla \upsilon_h)_E \\
+ \quad & \kappa \sum_{e\in\Gamma_h} (\{\varepsilon \nabla \upsilon_h \cdot n_e\}, [y_h])_e - \sum_{e\in\Gamma_h} (\{\varepsilon \nabla y_h \cdot n_e\}, [\upsilon_h])_e \\
+ \quad & \sum_{e\in\Gamma_h} \frac{\sigma\varepsilon}{h_e^{\beta_0}} ([y_h], [\upsilon_h])_e + \sum_{E\in\xi_h} (\beta \cdot \nabla y_h + r y_h, \upsilon_h)_E \\
+ \quad & \sum_{e\in\Gamma_h^0} (y_h^+ - y_h^-, |n\cdot\beta| \upsilon_h^+)_e + \sum_{e\in\Gamma_h^-} (y_h^+, \upsilon_h^+ |n\cdot\beta|)_e,
\end{aligned}
$$

with σ penalty parameter and β_0 superpenalization parameter.

- if $\kappa = -1$, **SIPG**, i.e., symmetric interior penalty Galerkin,
- if $\kappa = 1$, **NIPG**, i.e., nonsymmetric interior penalty Galerkin,
- if $\kappa = 0$, **IIPG**, i.e., incomplete interior penalty Galerkin.

D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *2002.*

Interior Penalty Galerkin Methods

$$
\begin{aligned}
a_h^s(y_h, \upsilon_h) \quad = \quad & \sum_{E\in\xi_h} (\varepsilon\nabla y_h, \nabla \upsilon_h)_E \\
+ \quad & \kappa \sum_{e\in\Gamma_h} (\{\varepsilon\nabla \upsilon_h \cdot n_e\}, [y_h])_e - \sum_{e\in\Gamma_h} (\{\varepsilon\nabla y_h \cdot n_e\}, [\upsilon_h])_e \\
+ \quad & \sum_{e\in\Gamma_h} \frac{\sigma\varepsilon}{h_e^{\beta_0}} ([y_h], [\upsilon_h])_e + \sum_{E\in\xi_h} (\beta \cdot \nabla y_h + r y_h, \upsilon_h)_E \\
+ \quad & \sum_{e\in\Gamma_h^0} (y_h^+ - y_h^-, |n\cdot\beta| \upsilon_h^+)_e + \sum_{e\in\Gamma_h^-} (y_h^+, \upsilon_h^+ |n\cdot\beta|)_e,
\end{aligned}
$$

with σ penalty parameter and β_0 superpenalization parameter.

- if $\kappa = -1$, **SIPG**, i.e., symmetric interior penalty Galerkin,
- if $\kappa = 1$, **NIPG**, i.e., nonsymmetric interior penalty Galerkin,
- if $\kappa = 0$, **IIPG**, i.e., incomplete interior penalty Galerkin.

D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *2002.*

Interior Penalty Galerkin Methods

$$
\begin{aligned}
a_h^s(y_h, \upsilon_h) &= \sum_{E \in \xi_h} (\varepsilon \nabla y_h, \nabla \upsilon_h)_E \\
&+ \kappa \sum_{e \in \Gamma_h} (\{\varepsilon \nabla \upsilon_h \cdot n_e\}, [y_h])_e - \sum_{e \in \Gamma_h} (\{\varepsilon \nabla y_h \cdot n_e\}, [\upsilon_h])_e \\
&+ \sum_{e \in \Gamma_h} \frac{\sigma \varepsilon}{h_e^{\beta_0}} ([y_h], [\upsilon_h])_e + \sum_{E \in \xi_h} (\beta \cdot \nabla y_h + r y_h, \upsilon_h)_E \\
&+ \sum_{e \in \Gamma_h^0} (y_h^+ - y_h^-, |n \cdot \beta| \upsilon_h^+)_e + \sum_{e \in \Gamma_h^-} (y_h^+, \upsilon_h^+ |n \cdot \beta|)_e,
\end{aligned}
$$

with σ penalty parameter and β_0 superpenalization parameter.

- if $\kappa = -1$, **SIPG**, i.e., symmetric interior penalty Galerkin,
- if $\kappa = 1$, **NIPG**, i.e., nonsymmetric interior penalty Galerkin,
- if $\kappa = 0$, **IIPG**, i.e., incomplete interior penalty Galerkin.

D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *2002.*

Interior Penalty Galerkin Methods

$$
\begin{aligned}
a_h^s(y_h, \upsilon_h) \quad = \quad & \sum_{E \in \xi_h} (\varepsilon \nabla y_h, \nabla \upsilon_h)_E \\
+ \quad & \kappa \sum_{e \in \Gamma_h} (\{\varepsilon \nabla \upsilon_h \cdot n_e\}, [y_h])_e - \sum_{e \in \Gamma_h} (\{\varepsilon \nabla y_h \cdot n_e\}, [\upsilon_h])_e \\
+ \quad & \sum_{e \in \Gamma_h} \frac{\sigma \varepsilon}{h_e^{\beta_0}} ([y_h], [\upsilon_h])_e + \sum_{E \in \xi_h} (\beta \cdot \nabla y_h + r y_h, \upsilon_h)_E \\
+ \quad & \sum_{e \in \Gamma_h^0} (y_h^+ - y_h^-, |n \cdot \beta| \upsilon_h^+)_e + \sum_{e \in \Gamma_h^-} (y_h^+, \upsilon_h^+ |n \cdot \beta|)_e,
\end{aligned}
$$

with σ penalty parameter and β_0 superpenalization parameter.

- if $\kappa = -1$, **SIPG**, i.e., symmetric interior penalty Galerkin,
- if $\kappa = 1$, **NIPG**, i.e., nonsymmetric interior penalty Galerkin,
- if $\kappa = 0$, **IIPG**, i.e., incomplete interior penalty Galerkin.

D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *2002.*

Interior Penalty Galerkin Methods

$$
\begin{aligned}
a_h^s(y_h, \upsilon_h) &= \sum_{E\in\xi_h} (\varepsilon \nabla y_h, \nabla \upsilon_h)_E \\
&+ \kappa \sum_{e\in\Gamma_h} (\{\varepsilon \nabla \upsilon_h \cdot n_e\}, [y_h])_e - \sum_{e\in\Gamma_h} (\{\varepsilon \nabla y_h \cdot n_e\}, [\upsilon_h])_e \\
&+ \sum_{e\in\Gamma_h} \frac{\sigma \varepsilon}{h_e^{\beta_0}} ([y_h], [\upsilon_h])_e + \sum_{E\in\xi_h} (\beta \cdot \nabla y_h + r y_h, \upsilon_h)_E \\
&+ \sum_{e\in\Gamma_h^0} (y_h^+ - y_h^-, |n\cdot\beta| \upsilon_h^+)_e + \sum_{e\in\Gamma_h^-} (y_h^+, \upsilon_h^+ |n\cdot\beta|)_e,
\end{aligned}
$$

with σ penalty parameter and β_0 superpenalization parameter.

- if $\kappa = -1$, **SIPG**, i.e., symmetric interior penalty Galerkin,
- if $\kappa = 1$, **NIPG**, i.e., nonsymmetric interior penalty Galerkin,
- if $\kappa = 0$, **IIPG**, i.e., incomplete interior penalty Galerkin.

D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *2002.*

Interior Penalty Galerkin Methods

$$
\begin{aligned}
a_h^s(y_h, \upsilon_h) \quad = \quad & \sum_{E\in\xi_h} (\varepsilon \nabla y_h, \nabla \upsilon_h)_E \\
+ \quad & \kappa \sum_{e\in\Gamma_h} (\{\varepsilon \nabla \upsilon_h \cdot n_e\}, [y_h])_e - \sum_{e\in\Gamma_h} (\{\varepsilon \nabla y_h \cdot n_e\}, [\upsilon_h])_e \\
+ \quad & \sum_{e\in\Gamma_h} \frac{\sigma\varepsilon}{h_e^{\beta_0}} ([y_h], [\upsilon_h])_e + \sum_{E\in\xi_h} (\beta \cdot \nabla y_h + r y_h, \upsilon_h)_E \\
+ \quad & \sum_{e\in\Gamma_h^0} (y_h^+ - y_h^-, |n \cdot \beta| \upsilon_h^+)_e + \sum_{e\in\Gamma_h^-} (y_h^+, \upsilon_h^+ |n \cdot \beta|)_e,
\end{aligned}
$$

with σ penalty parameter and β_0 superpenalization parameter.

- if $\kappa = -1$, **SIPG**, i.e., symmetric interior penalty Galerkin,
- if $\kappa = 1$, **NIPG**, i.e., nonsymmetric interior penalty Galerkin,
- if $\kappa = 0$, **IIPG**, i.e., incomplete interior penalty Galerkin.

D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *2002.*

Interior Penalty Galerkin Methods

$$
\begin{aligned}
a_h^s(y_h, \upsilon_h) &= \sum_{E\in\xi_h} (\varepsilon\nabla y_h, \nabla \upsilon_h)_E \\
&+ \kappa \sum_{e\in\Gamma_h} (\{\varepsilon\nabla \upsilon_h \cdot n_e\}, [y_h])_e - \sum_{e\in\Gamma_h} (\{\varepsilon\nabla y_h \cdot n_e\}, [\upsilon_h])_e \\
&+ \sum_{e\in\Gamma_h} \frac{\sigma\varepsilon}{h_e^{\beta_0}} ([y_h], [\upsilon_h])_e + \sum_{E\in\xi_h} (\beta \cdot \nabla y_h + r y_h, \upsilon_h)_E \\
&+ \sum_{e\in\Gamma_h^0} (y_h^+ - y_h^-, |n\cdot\beta| \upsilon_h^+)_e + \sum_{e\in\Gamma_h^-} (y_h^+, \upsilon_h^+ |n\cdot\beta|)_e,
\end{aligned}
$$

with σ penalty parameter and β_0 superpenalization parameter.

- if $\kappa = -1$, **SIPG**, i.e., symmetric interior penalty Galerkin,
- if $\kappa = 1$, **NIPG**, i.e., nonsymmetric interior penalty Galerkin,
- if $\kappa = 0$, **IIPG**, i.e., incomplete interior penalty Galerkin.

D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *2002.*

Interior Penalty Galerkin Methods

$$
\begin{aligned}
a_h^s(y_h, \upsilon_h) \quad = \quad & \sum_{E \in \xi_h} (\varepsilon \nabla y_h, \nabla \upsilon_h)_E \\
+ \quad & \kappa \sum_{e \in \Gamma_h} (\{\varepsilon \nabla \upsilon_h \cdot n_e\}, [y_h])_e - \sum_{e \in \Gamma_h} (\{\varepsilon \nabla y_h \cdot n_e\}, [\upsilon_h])_e \\
+ \quad & \sum_{e \in \Gamma_h} \frac{\sigma \varepsilon}{h_e^{\beta_0}} ([y_h], [\upsilon_h])_e + \sum_{E \in \xi_h} (\beta \cdot \nabla y_h + r y_h, \upsilon_h)_E \\
+ \quad & \sum_{e \in \Gamma_h^0} (y_h^+ - y_h^-, |n \cdot \beta| \upsilon_h^+)_e + \sum_{e \in \Gamma_h^-} (y_h^+, \upsilon_h^+ |n \cdot \beta|)_e ,
\end{aligned}
$$

with σ penalty parameter and β_0 superpenalization parameter.

- if $\kappa = -1$, **SIPG**, i.e., symmetric interior penalty Galerkin,
- if $\kappa = 1$, **NIPG**, i.e., nonsymmetric interior penalty Galerkin,
- if $\kappa = 0$, **IIPG**, i.e., incomplete interior penalty Galerkin.

D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini, Unified analysis of discontinuous Galerkin methods for elliptic problems, *2002.*

Discretize-Optimize System

- $b_h(u_h, v_h) = -\sum_{E\in\xi_h} (u_h, v_h)_E$
- and the linear right-hand side

$$
\begin{aligned}
l_h^s(v_h) &= \sum_{E\in\xi_h} (f, v_h)_E + \sum_{e\in\Gamma_h^\partial} \frac{\sigma\varepsilon}{h_e^{\beta_0}} (g_D, [v_h])_e \\
&+ \kappa \sum_{e\in\Gamma_h^\partial} (\varepsilon g_D, \{\nabla v_h \cdot n_e\})_e + \sum_{e\in\Gamma_h^-} (g_D, v_h^+ |n\cdot\beta|)_e.
\end{aligned}
$$

Discretize-Optimize System

- $b_h(u_h, v_h) = -\sum_{E\in\xi_h}(u_h, v_h)_E$
- and the linear right-hand side

$$
\begin{aligned}
l_h^s(v_h) \quad &= \quad \sum_{E\in\xi_h}(f, v_h)_E + \sum_{e\in\Gamma_h^\partial}\frac{\sigma\varepsilon}{h_e^{\beta_0}}(g_D, [v_h])_e \\
&+ \quad \kappa \sum_{e\in\Gamma_h^\partial}(\varepsilon g_D, \{\nabla v_h \cdot n_e\})_e + \sum_{e\in\Gamma_h^-}(g_D, v_h^+|n\cdot\beta|)_e.
\end{aligned}
$$

Optimize-Discretize

The discretized state, adjoint and gradient equations:

$$
\begin{aligned}
a_h^s(y_h, \upsilon_h) + b_h(u_h, \upsilon_h) &= l_h^s(\upsilon_h), && \forall v_h \in Y_h, \\
a_h^a(p_h, \psi_h) + (y_h, \psi_h) &= (y_d, \psi_h), && \forall \psi_h \in \Lambda_h, \\
b_h(w_h, p_h) + \omega(u_h, w_h) &= 0, && \forall w_h \in U_h,
\end{aligned}
$$

where

$$
\begin{aligned}
a_h^a(p_h, \psi_h) = & \sum_{E \in \xi_h} (\varepsilon \nabla p_h, \nabla \psi_h)_E \\
& + \kappa \sum_{e \in \Gamma_h} (\{\varepsilon \nabla \psi_h \cdot n_e\}, [p_h])_e - \sum_{e \in \Gamma_h} (\{\varepsilon \nabla p_h \cdot n_e\}, [\psi_h])_e \\
& + \sum_{e \in \Gamma_h} \frac{\sigma \varepsilon}{h_e^{\beta_0}} ([p_h], [\psi_h])_e + \sum_{E \in \xi_h} (-\beta \cdot \nabla p_h + (r - \nabla \cdot \beta) p_h, \psi_h)_E \\
& + \sum_{e \in \Gamma_h^0} (p_h^+ - p_h^-, |n \cdot \beta| \psi_h^+)_e + \sum_{e \in \Gamma_h^+} (p_h^+, \psi_h^+ |n \cdot \beta|)_e .
\end{aligned}
$$

Optimize-Discretize

The discretized state, adjoint and gradient equations:

$$
\begin{aligned}
a_h^s(y_h, \upsilon_h) + b_h(u_h, \upsilon_h) &= l_h^s(\upsilon_h), && \forall v_h \in Y_h, \\
a_h^a(p_h, \psi_h) + (y_h, \psi_h) &= (y_d, \psi_h), && \forall \psi_h \in \Lambda_h, \\
b_h(w_h, p_h) + \omega(u_h, w_h) &= 0, && \forall w_h \in U_h,
\end{aligned}
$$

where

$$
\begin{aligned}
a_h^a(p_h, \psi_h) &= \sum_{E \in \xi_h} (\varepsilon \nabla p_h, \nabla \psi_h)_E \\
&+ \kappa \sum_{e \in \Gamma_h} (\{\varepsilon \nabla \psi_h \cdot n_e\}, [p_h])_e - \sum_{e \in \Gamma_h} (\{\varepsilon \nabla p_h \cdot n_e\}, [\psi_h])_e \\
&+ \sum_{e \in \Gamma_h} \frac{\sigma \varepsilon}{h_e^{\beta_0}} ([p_h], [\psi_h])_e + \sum_{E \in \xi_h} (-\beta \cdot \nabla p_h + (r - \nabla \cdot \beta) p_h, \psi_h)_E \\
&+ \sum_{e \in \Gamma_h^0} (p_h^+ - p_h^-, |n \cdot \beta| \psi_h^+)_e + \sum_{e \in \Gamma_h^+} (p_h^+, \psi_h^+ |n \cdot \beta|)_e .
\end{aligned}
$$

Optimize-Discretize

The discretized state, adjoint and gradient equations:

$$\begin{aligned} a_h^s(y_h, \upsilon_h) + b_h(u_h, \upsilon_h) &= l_h^s(\upsilon_h), & \forall v_h \in Y_h, \\ a_h^a(p_h, \psi_h) + (y_h, \psi_h) &= (y_d, \psi_h), & \forall \psi_h \in \Lambda_h, \\ b_h(w_h, p_h) + \omega(u_h, w_h) &= 0, & \forall w_h \in U_h, \end{aligned}$$

where

$$\begin{aligned} a_h^a(p_h, \psi_h) &= \sum_{E \in \xi_h} (\varepsilon \nabla p_h, \nabla \psi_h)_E \\ &+ \kappa \sum_{e \in \Gamma_h} (\{\varepsilon \nabla \psi_h \cdot n_e\}, [p_h])_e - \sum_{e \in \Gamma_h} (\{\varepsilon \nabla p_h \cdot n_e\}, [\psi_h])_e \\ &+ \sum_{e \in \Gamma_h} \frac{\sigma \varepsilon}{h_e^{\beta_0}} ([p_h], [\psi_h])_e + \sum_{E \in \xi_h} (-\beta \cdot \nabla p_h + (r - \nabla \cdot \beta) p_h, \psi_h)_E \\ &+ \sum_{e \in \Gamma_h^0} (p_h^+ - p_h^-, |n \cdot \beta| \psi_h^+)_e + \sum_{e \in \Gamma_h^+} (p_h^+, \psi_h^+ |n \cdot \beta|)_e. \end{aligned}$$

DO versus OD

Discretize-then-Optimize

$$\begin{pmatrix} \mathbb{M} & 0 & \mathbb{A}_s^T \\ 0 & \omega\mathbb{Q} & \mathbb{B}^T \\ \mathbb{A}_s & \mathbb{B} & 0 \end{pmatrix} \begin{pmatrix} \vec{y} \\ \vec{u} \\ \vec{p} \end{pmatrix} = \begin{pmatrix} \vec{b} \\ 0 \\ \vec{f} \end{pmatrix}.$$

Optimize-then-Discretize

$$\begin{pmatrix} \mathbb{M} & 0 & \mathbb{A}_a \\ 0 & \omega\mathbb{Q} & \mathbb{B} \\ \mathbb{A}_s & \mathbb{B} & 0 \end{pmatrix} \begin{pmatrix} \vec{y} \\ \vec{u} \\ \vec{p} \end{pmatrix} = \begin{pmatrix} \vec{b} \\ 0 \\ \vec{f} \end{pmatrix}.$$

DO versus OD

Discretize-then-Optimize

$$\begin{pmatrix} \mathbb{M} & 0 & \mathbb{A}_s^T \\ 0 & \omega\mathbb{Q} & \mathbb{B}^T \\ \mathbb{A}_s & \mathbb{B} & 0 \end{pmatrix} \begin{pmatrix} \vec{y} \\ \vec{u} \\ \vec{p} \end{pmatrix} = \begin{pmatrix} \vec{b} \\ 0 \\ \vec{f} \end{pmatrix}.$$

Optimize-then-Discretize

$$\begin{pmatrix} \mathbb{M} & 0 & \mathbb{A}_a \\ 0 & \omega\mathbb{Q} & \mathbb{B} \\ \mathbb{A}_s & \mathbb{B} & 0 \end{pmatrix} \begin{pmatrix} \vec{y} \\ \vec{u} \\ \vec{p} \end{pmatrix} = \begin{pmatrix} \vec{b} \\ 0 \\ \vec{f} \end{pmatrix}.$$

Main Result

Theorem

The discretize-then-optimize and the optimize-then-discretize lead the same scheme for symmetric DG methods, i.e., SIPG, but not for nonsymmetric DG methods, i.e., NIPG, IIPG.

H. Yücel, M. Heikenschloss, and B. Karasözen, Distributed Optimal Control of Diffusion-Convection-Reaction Equations Using Discontinuous Galerkin Methods, *to appear in the Proceedings of ENUMATH 2011 Conference, Leicester, England, 5-9 September 2011*

Main Result

Theorem

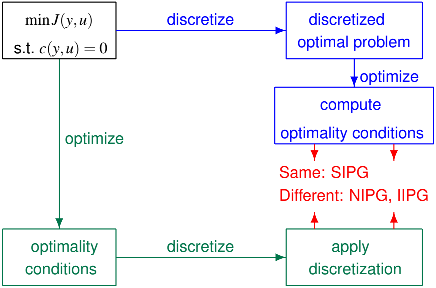
The discretize-then-optimize and the optimize-then-discretize lead the same scheme for symmetric DG methods, i.e., SIPG, but not for nonsymmetric DG methods, i.e., NIPG, IIPG.

H. Yücel, M. Heikenschloss, and B. Karasözen, Distributed Optimal Control of Diffusion-Convection-Reaction Equations Using Discontinuous Galerkin Methods, *to appear in the Proceedings of ENUMATH 2011 Conference, Leicester, England, 5-9 September 2011*

Main Result

Theorem

The discretize-then-optimize and the optimize-then-discretize lead the same scheme for symmetric DG methods, i.e., SIPG, but not for nonsymmetric DG methods, i.e., NIPG, IIPG.



H. Yücel, M. Heikenschloss, and B. Karasözen, Distributed Optimal Control of Diffusion-Convection-Reaction Equations Using Discontinuous Galerkin Methods*, to appear in the Proceedings of ENUMATH 2011 Conference, Leicester, England, 5-9 September 2011*

Comparison of DO and OD with Boundary Layer

Example

[Collis, Heinkenschloss, 2002] Let
$\Omega = [0,1]^2, \varepsilon = 10^{-2}, \theta = 45^o, \beta = (\cos\theta, \sin\theta), r = 0$ and $\omega = 1$.
The exact solutions:

$$y_{ex}(x_1,x_2) = \eta(x_1)\eta(x_2), \quad p_{ex}(x_1,x_2) = \mu(x_1)\mu(x_2),$$

where

$$\eta(z) = z - \frac{\exp((z-1)/\varepsilon) - \exp(-1/\varepsilon)}{1 - \exp(-1/\varepsilon)},$$

$$\mu(z) = 1 - z - \frac{\exp(-z/\varepsilon) - \exp(-1/\varepsilon)}{1 - \exp(-1/\varepsilon}.$$

- For SIPG and IIPG, $\sigma = 3k(k+1) \ \forall e \in \Gamma_h^0$ and $\sigma = 6k(k+1) \ \forall e \in \Gamma_h^\partial$
- For NIPG, $\sigma = 1, \forall e \in \Gamma_h$
- NIPG with standart penalization denoted by NIPG1, i.e., $\beta_0 = 1$
- NIPG with superpenalization denoted by NIPG3, i.e., $\beta_0 = 3$

Comparison of DO and OD with Boundary Layer

Example

[Collis, Heinkenschloss, 2002] Let
$\Omega = [0,1]^2, \varepsilon = 10^{-2}, \theta = 45^o, \beta = (\cos\theta, \sin\theta), r = 0$ and $\omega = 1$.
The exact solutions:

$$y_{ex}(x_1,x_2) = \eta(x_1)\eta(x_2), \quad p_{ex}(x_1,x_2) = \mu(x_1)\mu(x_2),$$

where

$$\eta(z) = z - \frac{\exp((z-1)/\varepsilon) - \exp(-1/\varepsilon)}{1-\exp(-1/\varepsilon)},$$

$$\mu(z) = 1 - z - \frac{\exp(-z/\varepsilon) - \exp(-1/\varepsilon)}{1-\exp(-1/\varepsilon}.$$

- For SIPG and IIPG, $\sigma = 3k(k+1) \ \forall e \in \Gamma_h^0$ and $\sigma = 6k(k+1) \ \forall e \in \Gamma_h^\partial$
- For NIPG, $\sigma = 1, \forall e \in \Gamma_h$
- NIPG with standart penalization denoted by NIPG1, i.e., $\beta_0 = 1$
- NIPG with superpenalization denoted by NIPG3, i.e., $\beta_0 = 3$

Comparison of DO and OD with Boundary Layer

Example

[Collis, Heinkenschloss, 2002] Let
$\Omega = [0,1]^2, \varepsilon = 10^{-2}, \theta = 45^o, \beta = (\cos\theta, \sin\theta), r = 0$ and $\omega = 1$.
The exact solutions:

$$y_{ex}(x_1,x_2) = \eta(x_1)\eta(x_2), \quad p_{ex}(x_1,x_2) = \mu(x_1)\mu(x_2),$$

where

$$\eta(z) = z - \frac{\exp((z-1)/\varepsilon) - \exp(-1/\varepsilon)}{1-\exp(-1/\varepsilon)},$$

$$\mu(z) = 1 - z - \frac{\exp(-z/\varepsilon) - \exp(-1/\varepsilon)}{1-\exp(-1/\varepsilon}.$$

- For SIPG and IIPG, $\sigma = 3k(k+1) \ \forall e \in \Gamma_h^0$ and $\sigma = 6k(k+1) \ \forall e \in \Gamma_h^\partial$
- For NIPG, $\sigma = 1, \forall e \in \Gamma_h$
- NIPG with standart penalization denoted by NIPG1, i.e., $\beta_0 = 1$
- NIPG with superpenalization denoted by NIPG3, i.e., $\beta_0 = 3$

Comparison of DO and OD with Boundary Layer

Example

[Collis, Heinkenschloss, 2002] Let
$\Omega = [0,1]^2, \varepsilon = 10^{-2}, \theta = 45^o, \beta = (\cos\theta, \sin\theta), r = 0$ and $\omega = 1$.
The exact solutions:

$$y_{ex}(x_1,x_2) = \eta(x_1)\eta(x_2), \quad p_{ex}(x_1,x_2) = \mu(x_1)\mu(x_2),$$

where

$$\begin{aligned} \eta(z) &= z - \frac{\exp((z-1)/\varepsilon) - \exp(-1/\varepsilon)}{1-\exp(-1/\varepsilon)}, \\ \mu(z) &= 1 - z - \frac{\exp(-z/\varepsilon) - \exp(-1/\varepsilon)}{1-\exp(-1/\varepsilon}. \end{aligned}$$

- For SIPG and IIPG, $\sigma = 3k(k+1) \ \forall e \in \Gamma_h^0$ and $\sigma = 6k(k+1) \ \forall e \in \Gamma_h^\partial$
- For NIPG, $\sigma = 1, \forall e \in \Gamma_h$
- NIPG with standart penalization denoted by NIPG1, i.e., $\beta_0 = 1$
- NIPG with superpenalization denoted by NIPG3, i.e., $\beta_0 = 3$

Comparison of DO and OD with Boundary Layer

Example

[Collis, Heinkenschloss, 2002] Let
$\Omega = [0,1]^2, \varepsilon = 10^{-2}, \theta = 45^o, \beta = (\cos\theta, \sin\theta), r = 0$ and $\omega = 1$.
The exact solutions:

$$y_{ex}(x_1,x_2) = \eta(x_1)\eta(x_2), \quad p_{ex}(x_1,x_2) = \mu(x_1)\mu(x_2),$$

where

$$\eta(z) = z - \frac{\exp((z-1)/\varepsilon) - \exp(-1/\varepsilon)}{1-\exp(-1/\varepsilon)},$$

$$\mu(z) = 1 - z - \frac{\exp(-z/\varepsilon) - \exp(-1/\varepsilon)}{1-\exp(-1/\varepsilon}.$$

- For SIPG and IIPG, $\sigma = 3k(k+1) \ \forall e \in \Gamma_h^0$ and $\sigma = 6k(k+1) \ \forall e \in \Gamma_h^\partial$
- For NIPG, $\sigma = 1, \forall e \in \Gamma_h$
- NIPG with standart penalization denoted by NIPG1, i.e., $\beta_0 = 1$
- NIPG with superpenalization denoted by NIPG3, i.e., $\beta_0 = 3$

SIPG Method

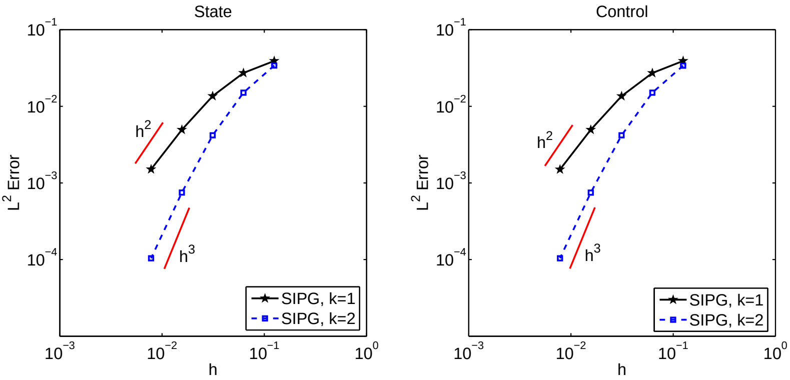


Figure: L_2 error for SIPG with $\varepsilon = 10^{-2}$.

NIPG1-NIPG3 Methods

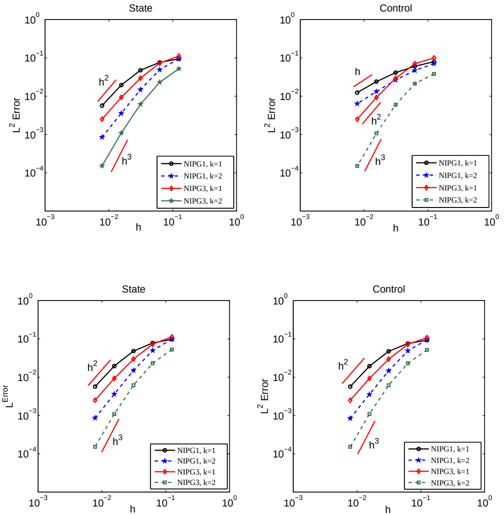


Figure: L_2 error for NIPG1 and NIPG3 with $\varepsilon = 10^{-2}$: *discretize-then-optimize* (upper), *optimize-then-discretize* (lower)

Motivation

- Convection dominate problems cause boundary and/or interior layers
- Need more elements to obtain more accurate solution
- Instead of refine all region, place only more grid-points where the solution is less regular,i.e., refine the discretization near the layers

Motivation

- Convection dominate problems cause boundary and/or interior layers
- Need more elements to obtain more accurate solution
- Instead of refine all region, place only more grid-points where the solution is less regular,i.e., refine the discretization near the layers

Motivation

- Convection dominate problems cause boundary and/or interior layers
- Need more elements to obtain more accurate solution
- Instead of refine all region, place only more grid-points where the solution is less regular,i.e., refine the discretization near the layers

A Priori Error Estimates

$$\|u - u_h\| \leq CE(u,h)$$

Contains the unknown solution u

Insufficient since it provides information about the asymptotic error behavior

A Posteriori Error Estimates

- It has the following form

$$\|u - u_h\| \leq \underbrace{C\underbrace{E(u_h, h, data_h)}_{\text{Error Indicator}}}_{\text{Error Estimator}} + \underbrace{\|data - data_h\|}_{\text{data oscillations}}$$

- Extracted from the computed numerical solution and from the given data of the problem
- Global upper bounds are sufficient to obtain a numerical solution with an accuracy a prescribed tolerance
- Local lower bounds are necessary to ensure that the grid is correctly refined

A Posteriori Error Estimates

- It has the following form

$$\|u - u_h\| \leq \underbrace{C\underbrace{E(u_h, h, data_h)}_{\text{Error Indicator}}}_{\text{Error Estimator}} + \underbrace{\|data - data_h\|}_{\text{data oscillations}}$$

- Extracted from the computed numerical solution and from the given data of the problem
- Global upper bounds are sufficient to obtain a numerical solution with an accuracy a prescribed tolerance
- Local lower bounds are necessary to ensure that the grid is correctly refined

A Posteriori Error Estimates

- It has the following form

$$\|u - u_h\| \leq \underbrace{C \underbrace{E(u_h, h, data_h)}_{\text{Error Indicator}}}_{\text{Error Estimator}} + \underbrace{\|data - data_h\|}_{\text{data oscillations}}$$

- Extracted from the computed numerical solution and from the given data of the problem
- Global upper bounds are sufficient to obtain a numerical solution with an accuracy a prescribed tolerance
- Local lower bounds are necessary to ensure that the grid is correctly refined

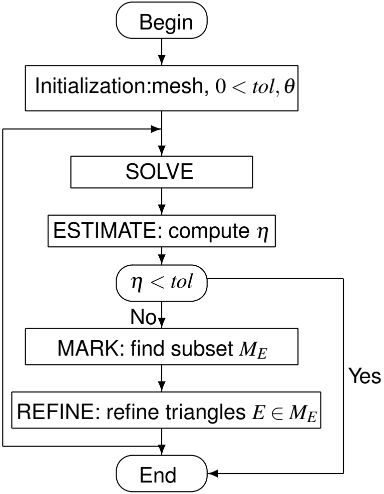
A Posteriori Error Estimates

- It has the following form

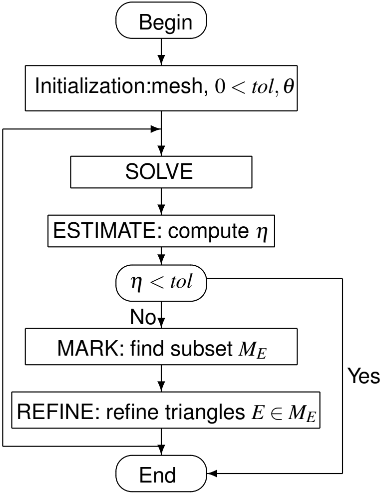
$$\|u-u_h\| \leq \underbrace{C\underbrace{E(u_h,h,data_h)}_{\text{Error Indicator}}}_{\text{Error Estimator}} + \underbrace{\|data-data_h\|}_{\text{data oscillations}}$$

- Extracted from the computed numerical solution and from the given data of the problem
- Global upper bounds are sufficient to obtain a numerical solution with an accuracy a prescribed tolerance
- Local lower bounds are necessary to ensure that the grid is correctly refined

Adaptive Strategy



- SOLVE stands for solution of the optimal control problem using DG discretization on ξ_h mesh



- **SOLVE** stands for solution of the optimal control problem using DG discretization on ξ_h mesh

Estimator

- **ESTIMATE:** [Schötzau and Zhu, 2009]

$$\eta^y = \left(\sum_{E \in \xi_h} (\eta_E^y)^2 \right)^{\frac{1}{2}}, \ \eta^p = \left(\sum_{E \in \xi_h} (\eta_E^p)^2 \right)^{\frac{1}{2}}, \ \eta^u = \left(\sum_{E \in \xi_h} (\eta_E^u)^2 \right)^{\frac{1}{2}}.$$

where

$$\begin{aligned} (\eta_E^y)^2 &= \left[(\eta_{E_R}^y)^2 + (\eta_{e_D}^y)^2 + (\eta_{e_J}^y)^2 \right], \\ (\eta_E^p)^2 &= \left[(\eta_{E_R}^p)^2 + (\eta_{e_D}^p)^2 + (\eta_{e_J}^y)^2 \right], \\ (\eta_E^u)^2 &= \left[(\eta_{E_R}^u)^2 \right]. \end{aligned}$$

η_E: the element residual

$$\begin{aligned} \eta_{E_R}^y &= \rho_E \| f_h + u_h + \varepsilon \Delta y_h - \beta_h \cdot \nabla y_h - r_h y_h \|_{L^2(E)}, & E \in \xi_h, \\ \eta_{E_R}^p &= \rho_E \| -(y_h - (y_d)_h + \varepsilon \Delta p_h + \beta_h \cdot \nabla p_h - r_h p_h \|_{L^2(E)}, & E \in \xi_h, \\ \eta_{E_R}^u &= \| \omega u_h - p_h \|_{L^2(E)}, & E \in \xi_h. \end{aligned}$$

Estimator

- **ESTIMATE:** [Schötzau and Zhu, 2009]

$$\eta^y = \left(\sum_{E \in \xi_h} (\eta_E^y)^2 \right)^{\frac{1}{2}}, \ \eta^p = \left(\sum_{E \in \xi_h} (\eta_E^p)^2 \right)^{\frac{1}{2}}, \ \eta^u = \left(\sum_{E \in \xi_h} (\eta_E^u)^2 \right)^{\frac{1}{2}}.$$

where

$$\begin{aligned} (\eta_E^y)^2 &= \left[(\eta_{E_R}^y)^2 + (\eta_{e_D}^y)^2 + (\eta_{e_J}^y)^2 \right], \\ (\eta_E^p)^2 &= \left[(\eta_{E_R}^p)^2 + (\eta_{e_D}^p)^2 + (\eta_{e_J}^y)^2 \right], \\ (\eta_E^u)^2 &= \left[(\eta_{E_R}^u)^2 \right]. \end{aligned}$$

η_E: the element residual

$$\begin{aligned} \eta_{E_R}^y &= \rho_E \| f_h + u_h + \varepsilon \Delta y_h - \beta_h \cdot \nabla y_h - r_h y_h \|_{L^2(E)}, & E \in \xi_h, \\ \eta_{E_R}^p &= \rho_E \| -(y_h - (y_d)_h + \varepsilon \Delta p_h + \beta_h \cdot \nabla p_h - r_h p_h \|_{L^2(E)}, & E \in \xi_h, \\ \eta_{E_R}^u &= \| \omega u_h - p_h \|_{L^2(E)}, & E \in \xi_h. \end{aligned}$$

Estimator

- **ESTIMATE:** [Schötzau and Zhu, 2009]

$$\eta^y = \left(\sum_{E\in\xi_h} (\eta_E^y)^2 \right)^{\frac{1}{2}}, \ \eta^p = \left(\sum_{E\in\xi_h} (\eta_E^p)^2 \right)^{\frac{1}{2}}, \ \eta^u = \left(\sum_{E\in\xi_h} (\eta_E^u)^2 \right)^{\frac{1}{2}}.$$

where

$$\begin{aligned} (\eta_E^y)^2 &= \left[(\eta_{E_R}^y)^2 + (\eta_{e_D}^y)^2 + (\eta_{e_J}^y)^2 \right], \\ (\eta_E^p)^2 &= \left[(\eta_{E_R}^p)^2 + (\eta_{e_D}^p)^2 + (\eta_{e_J}^y)^2 \right], \\ (\eta_E^u)^2 &= \left[(\eta_{E_R}^u)^2 \right]. \end{aligned}$$

η_E: the element residual

$$\begin{aligned} \eta_{E_R}^y &= \rho_E \| f_h + u_h + \varepsilon \Delta y_h - \beta_h \cdot \nabla y_h - r_h y_h \|_{L^2(E)}, && E \in \xi_h, \\ \eta_{E_R}^p &= \rho_E \| -(y_h - (y_d)_h + \varepsilon \Delta p_h + \beta_h \cdot \nabla p_h - r_h p_h \|_{L^2(E)}, && E \in \xi_h, \\ \eta_{E_R}^u &= \| \omega u_h - p_h \|_{L^2(E)}, && E \in \xi_h. \end{aligned}$$

Estimator

- **ESTIMATE:** [Schötzau and Zhu, 2009]

$$\eta^y = \left(\sum_{E \in \xi_h} (\eta_E^y)^2 \right)^{\frac{1}{2}}, \ \eta^p = \left(\sum_{E \in \xi_h} (\eta_E^p)^2 \right)^{\frac{1}{2}}, \ \eta^u = \left(\sum_{E \in \xi_h} (\eta_E^u)^2 \right)^{\frac{1}{2}}.$$

where

$$\begin{aligned} (\eta_E^y)^2 &= \left[(\eta_{E_R}^y)^2 + (\eta_{e_D}^y)^2 + (\eta_{e_J}^y)^2 \right], \\ (\eta_E^p)^2 &= \left[(\eta_{E_R}^p)^2 + (\eta_{e_D}^p)^2 + (\eta_{e_J}^y)^2 \right], \\ (\eta_E^u)^2 &= \left[(\eta_{E_R}^u)^2 \right]. \end{aligned}$$

η_E: the element residual

$$\begin{aligned} \eta_{E_R}^y &= \rho_E \| f_h + u_h + \varepsilon \Delta y_h - \beta_h \cdot \nabla y_h - r_h y_h \|_{L^2(E)}, && E \in \xi_h, \\ \eta_{E_R}^p &= \rho_E \| -(y_h - (y_d)_h + \varepsilon \Delta p_h + \beta_h \cdot \nabla p_h - r_h p_h \|_{L^2(E)}, && E \in \xi_h, \\ \eta_{E_R}^u &= \| \omega u_h - p_h \|_{L^2(E)}, && E \in \xi_h. \end{aligned}$$

Edge part of Estimator

The edge residuals denoted by η_{e_D} and η_{e_J} coming from the jump in the numerical solutions

$$
\begin{aligned}
(\eta^y_{e_D})^2 &= \frac{1}{2}\sum_{\Gamma_h^0} \varepsilon^{-\frac{1}{2}}\rho_e \|[\varepsilon\nabla y_h]\|_e^2, \\
(\eta^y_{e_J})^2 &= \frac{1}{2}\sum_{\Gamma_h^0}(\frac{\sigma\varepsilon}{h_e} + r_0 h_e + \frac{h_e}{\varepsilon})\|[y_h]\|_e^2 + \sum_{\Gamma_h^\partial}(\frac{\sigma\varepsilon}{h_e} + r_0 h_e + \frac{h_e}{\varepsilon})\|[g_D - y_h]\| \\
(\eta^p_{e_D})^2 &= \frac{1}{2}\sum_{\Gamma_h^0} \varepsilon^{-\frac{1}{2}}\rho_e \|[\varepsilon\nabla p_h]\|_e^2, \\
(\eta^p_{e_J})^2 &= \frac{1}{2}\sum_{\Gamma_h^0}(\frac{\sigma\varepsilon}{h_e} + r_0 h_e + \frac{h_e}{\varepsilon})\|[p_h]\|_e^2 + \sum_{\Gamma_h^\partial}(\frac{\sigma\varepsilon}{h_e} + r_0 h_e + \frac{h_e}{\varepsilon})\|[p_h]\|_e^2.
\end{aligned}
$$

with

$$\rho_E = \min\{h_E\varepsilon^{-\frac{1}{2}}, r_0^{-\frac{1}{2}}\},\ \rho_e = \min\{h_e\varepsilon^{-\frac{1}{2}}, r_0^{-\frac{1}{2}}\}.$$

Data Approximation Errors

Data approximation terms:

$$
\begin{aligned}
(\theta_E^y)^2 &= \rho_E^2(\|f-f_h\|^2_{L^2(E)} + \|(\beta-\beta_h)\cdot\nabla y_h\|^2_{L^2(E)} + \|(r-r_h)y_h\|^2_{L^2(E)}),\\
(\theta_E^p)^2 &= \rho_E^2(\|(y_d)_h - y_d\|^2_{L^2(E)} + \|(\beta-\beta_h)\cdot\nabla p_h\|^2_{L^2(E)}\\
&+ \|(r-\nabla\cdot\beta)-(r_h-\nabla\cdot\beta_h)p_h\|^2_{L^2(E)}).
\end{aligned}
$$

The data approximation errors:

$$
\theta^y = \left(\sum_{E\in\xi_h}(\theta_E^y)^2\right)^{\frac{1}{2}}, \qquad \theta^p = \left(\sum_{E\in\xi_h}(\theta_E^p)^2\right)^{\frac{1}{2}}.
$$

Marking

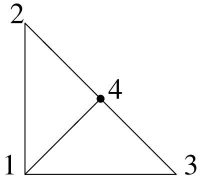
- **Marking Strategy**

  For a given universal constant θ, we choose subsets $M_E \subset \xi_h$ such that the following bulk criterion [Dörfler, 1996] is satisfied:

$$\sum_{E\in\xi_h} (\eta_E)^2 \leq \theta \sum_{E\in M_E} (\eta_E)^2.$$

Refinement

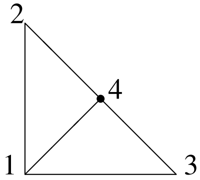
- In **Refinement** step, the marked elements are refined by longest edge bisection,



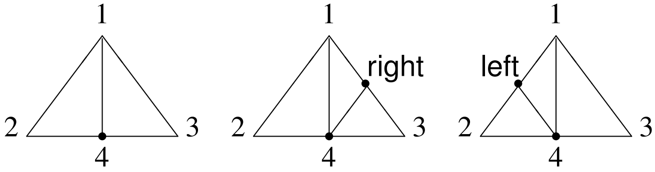
- whereas the elements of the marked edges are refined by bisection

Refinement

- In **Refinement** step, the marked elements are refined by longest edge bisection,



- whereas the elements of the marked edges are refined by bisection



A Posteriori Error Analysis

- Energy Norm

$$|||y|||^2 = \sum_{E\in\xi_h} (\|\varepsilon\nabla y\|^2_{L^2(E)} + r_0\|y\|^2_{L^2(E)}) + \sum_{e\in\Gamma_h} \frac{\sigma\varepsilon}{h_e}\|[y]\|^2_{L^2(e)}$$

- The semi-norm $|\cdot|_A$ with convective term [Verfürth,2005]

$$|y|_A^2 = |\beta y|_*^2 + \sum_{e\in\Gamma} (r_0 h_e + \frac{h_e}{\varepsilon})\|[y]\|^2_{L^2(e)},$$

where for $q \in L^2(\Omega)^2$,

$$|q|_* = \sup_{v\in H_0^1(\Omega)\setminus\{0\}} \frac{\int_\Omega q\cdot\nabla v dx}{|||v|||}.$$

- Assume that $r_0 > 0$ and $g_D = 0$.

A Posteriori Error Analysis

- Energy Norm

$$|||y|||^2 = \sum_{E\in\xi_h} (\|\varepsilon\nabla y\|^2_{L^2(E)} + r_0\|y\|^2_{L^2(E)}) + \sum_{e\in\Gamma_h} \frac{\sigma\varepsilon}{h_e}\|[y]\|^2_{L^2(e)}$$

- The semi-norm $|\cdot|_A$ with convective term [Verfürth,2005]

$$|y|^2_A \quad = \quad |\beta y|^2_* + \sum_{e\in\Gamma}(r_0 h_e + \frac{h_e}{\varepsilon})\|[y]\|^2_{L^2(e)},$$

where for $q \in L^2(\Omega)^2$,

$$|q|_* = \sup_{\upsilon\in H^1_0(\Omega)\backslash\{0\}} \frac{\int_\Omega q\cdot\nabla\upsilon dx}{|||\upsilon|||}.$$

- Assume that $r_0 > 0$ and $g_D = 0$.

A Posteriori Error Analysis

- Energy Norm

$$|||y|||^2 = \sum_{E\in\xi_h} (\|\varepsilon\nabla y\|^2_{L^2(E)} + r_0\|y\|^2_{L^2(E)}) + \sum_{e\in\Gamma_h} \frac{\sigma\varepsilon}{h_e}\|[y]\|^2_{L^2(e)}$$

- The semi-norm $|\cdot|_A$ with convective term [Verfürth,2005]

$$|y|^2_A \quad = \quad |\beta y|^2_* + \sum_{e\in\Gamma}(r_0 h_e + \frac{h_e}{\varepsilon})\|[y]\|^2_{L^2(e)},$$

where for $q \in L^2(\Omega)^2$,

$$|q|_* = \sup_{\upsilon\in H^1_0(\Omega)\setminus\{0\}} \frac{\int_\Omega q\cdot\nabla\upsilon dx}{|||\upsilon|||}.$$

- Assume that $r_0 > 0$ and $g_D = 0$.

A Posteriori Error Analysis

- Energy Norm

$$|||y|||^2 = \sum_{E\in\xi_h} (\|\varepsilon\nabla y\|^2_{L^2(E)} + r_0\|y\|^2_{L^2(E)}) + \sum_{e\in\Gamma_h} \frac{\sigma\varepsilon}{h_e}\|[y]\|^2_{L^2(e)}$$

- The semi-norm $|\cdot|_A$ with convective term [Verfürth,2005]

$$|y|^2_A \quad = \quad |\beta y|^2_* + \sum_{e\in\Gamma}(r_0 h_e + \frac{h_e}{\varepsilon})\|[y]\|^2_{L^2(e)},$$

where for $q \in L^2(\Omega)^2$,

$$|q|_* = \sup_{\upsilon\in H^1_0(\Omega)\setminus\{0\}} \frac{\int_\Omega q\cdot\nabla\upsilon dx}{|||\upsilon|||}.$$

- Assume that $r_0 > 0$ and $g_D = 0$.

Reliability

- Connection between the control and the adjoint

$$\|u-u_h\|^2_{L^2(\Omega)} \lesssim \|p_h - p[u_h]\|^2_{L^2(\Omega)} + (\eta^u)^2,$$

where $p[u_h]$ satisfies the following equation:

$$\begin{aligned} a(y[u_h],w)-(u_h,w) &= (f,w), \qquad \forall w \in V, \\ a(w,p[u_h])+(y[u_h],w) &= (y_d,w), \qquad \forall w \in V. \end{aligned}$$

- Connection between the adjoint and the state It holds

$$|||p[u_h]-p_h||| + |p[u_h]-p_h|_A \lesssim \eta^p + \theta^p + \|y_h - y[u_h]\|_{L^2(\Omega)}.$$

Reliability

- Connection between the control and the adjoint
$$\|u-u_h\|^2_{L^2(\Omega)} \lesssim \|p_h - p[u_h]\|^2_{L^2(\Omega)} + (\eta^u)^2,$$
where $p[u_h]$ satisfies the following equation:
$$\begin{aligned} a(y[u_h],w) - (u_h,w) &= (f,w), \qquad \forall w \in V, \\ a(w,p[u_h]) + (y[u_h],w) &= (y_d,w), \qquad \forall w \in V. \end{aligned}$$
- Connection between the adjoint and the state It holds
$$|||p[u_h]-p_h||| + |p[u_h]-p_h|_A \lesssim \eta^p + \theta^p + \|y_h - y[u_h]\|_{L^2(\Omega)}.$$

Reliability

- Connection between the control and the adjoint
$$\|u-u_h\|^2_{L^2(\Omega)} \lesssim \|p_h - p[u_h]\|^2_{L^2(\Omega)} + (\eta^u)^2,$$

  where $p[u_h]$ satisfies the following equation:
$$\begin{aligned} a(y[u_h],w) - (u_h,w) &= (f,w), \qquad \forall w \in V, \\ a(w,p[u_h]) + (y[u_h],w) &= (y_d,w), \qquad \forall w \in V. \end{aligned}$$

- Connection between the adjoint and the state It holds
$$|||p[u_h]-p_h||| + |p[u_h]-p_h|_A \lesssim \eta^p + \theta^p + \|y_h - y[u_h]\|_{L^2(\Omega)}.$$

Reliability

- Upper bound for state

$$|||y[u_h]-y_h|||+|y[u_h]-y_h|_A \lesssim \eta^y+\theta^y$$

- Reliability of the estimator

$$\begin{aligned}\|u-u_h\|_{L^2(\Omega)}+|||y-y_h|||+|y-y_h|_A &+ |||p-p_h|||+|p-p_h|_A \\ &\lesssim \eta^u+\eta^y+\theta^y+\eta^p+\theta^p\end{aligned}$$

Reliability

- Upper bound for state

$$|||y[u_h]-y_h|||+|y[u_h]-y_h|_A \lesssim \eta^y+\theta^y$$

- Reliability of the estimator

$$\begin{aligned}\|u-u_h\|_{L^2(\Omega)}+|||y-y_h|||+|y-y_h|_A \quad &+ \quad |||p-p_h|||+|p-p_h|_A \\ &\lesssim \quad \eta^u+\eta^y+\theta^y+\eta^p+\theta^p\end{aligned}$$

Efficiency

- Bounds for the estimator of the state and the adjoint

$$
\begin{aligned}
\eta^y &\lesssim |||y-y_h||| + |y-y_h|_A + \theta^y + \|u-u_h\|_{L^2(\Omega)} \\
\eta^p &\lesssim |||p-p_h||| + |p-p_h|_A + \theta^p + \|y-y_h\|_{L^2(\Omega)}
\end{aligned}
$$

  hold.
- Efficiency of the estimator

$$
\begin{aligned}
\eta^y + \eta^p + \eta^u &\lesssim \|u-u_h\|_{L^2(\Omega)} + |||y-y_h||| + |y-y_h|_A \\
&+ |||p-p_h||| + |p-p_h|_A + \theta^y + \theta^p .
\end{aligned}
$$

Efficiency

- Bounds for the estimator of the state and the adjoint

$$\eta^y \lesssim |||y - y_h||| + |y - y_h|_A + \theta^y + \|u - u_h\|_{L^2(\Omega)}$$
$$\eta^p \lesssim |||p - p_h||| + |p - p_h|_A + \theta^p + \|y - y_h\|_{L^2(\Omega)}$$

hold.

- Efficiency of the estimator

$$\eta^y + \eta^p + \eta^u \lesssim \|u - u_h\|_{L^2(\Omega)} + |||y - y_h||| + |y - y_h|_A + |||p - p_h||| + |p - p_h|_A + \theta^y + \theta^p.$$

Efficiency

- Bounds for the estimator of the state and the adjoint

$$\begin{aligned}\eta^y &\lesssim |||y-y_h||| + |y-y_h|_A + \theta^y + \|u-u_h\|_{L^2(\Omega)}\\ \eta^p &\lesssim |||p-p_h||| + |p-p_h|_A + \theta^p + \|y-y_h\|_{L^2(\Omega)}\end{aligned}$$

  hold.
- Efficiency of the estimator

$$\begin{aligned}\eta^y + \eta^p + \eta^u &\lesssim \|u-u_h\|_{L^2(\Omega)} + |||y-y_h||| + |y-y_h|_A\\ &+ |||p-p_h||| + |p-p_h|_A + \theta^y + \theta^p.\end{aligned}$$

Boundary Layer

Example (Collis, Heinkenschloss, 2002)

Let
$\Omega = [0,1]^2, \varepsilon = 10^{-3}, \theta = 45^o, \beta = (\cos\theta, \sin\theta), r = 0$ and $\omega = 1$.
The exact solutions:

$$y_{ex}(x_1,x_2) = \eta(x_1)\eta(x_2), \quad p_{ex}(x_1,x_2) = \mu(x_1)\mu(x_2),$$

$$\begin{aligned} \eta(z) &= z - \frac{\exp((z-1)/\varepsilon) - \exp(-1/\varepsilon)}{1-\exp(-1/\varepsilon)}, \\ \mu(z) &= 1 - z - \frac{\exp(-z/\varepsilon) - \exp(-1/\varepsilon)}{1-\exp(-1/\varepsilon}. \end{aligned}$$

H. Yücel, M. Heikenschloss, and B. Karasözen, An Adaptive discontinuous Galerkin method for convection dominated distributed optimal control problems*, Applied Numerical Mathematics, 2012. Submitted.*

Uniform Refinement

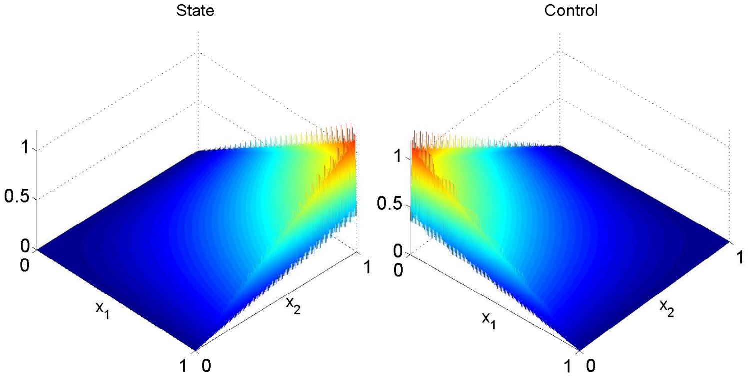


Figure: Uniformly refined mesh (16641 nodes) for $\varepsilon = 10^{-3}$.

Adaptive Refinement

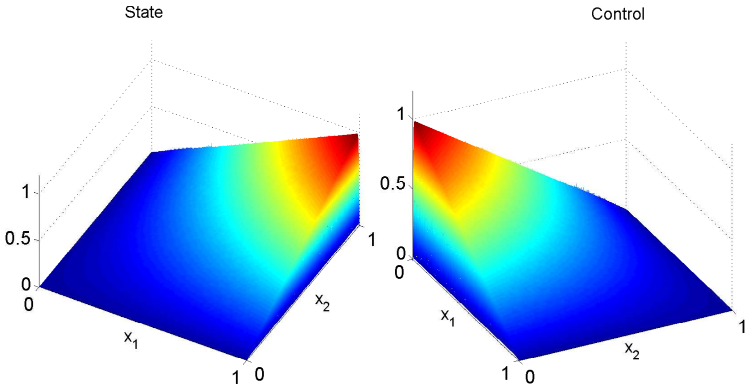
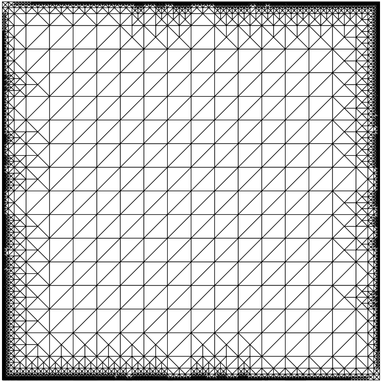


Figure: Adaptively refined mesh (15032 nodes) for $\varepsilon = 10^{-3}$.

Adaptive Mesh

[level,nodes]=[14,15032]



Global Errors

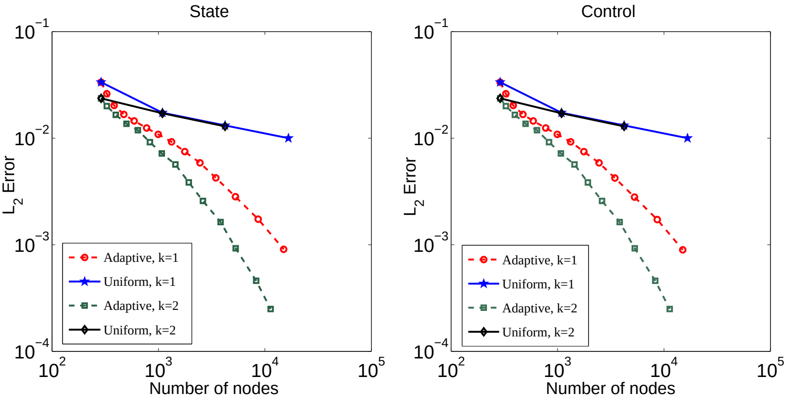


Figure: Errors in L^2 norm using linear and quadratic elements for $\varepsilon = 10^{-3}$.

Interior Layer

Example (Heinkenschloss and Leykekhman, 2008)

$$\Omega = [0,1]^2, \quad \varepsilon = 10^{-7}, \quad \beta = (1,2), \quad r = 0 \text{ and } \omega = 10^{-2}.$$

Exact solution:

$$\begin{aligned} y_{ex}(x_1, y_1) &= (1 - x_1)^3 \arctan\left(\frac{x_2 - 0.5}{\varepsilon}\right), \\ p_{ex}(x_1, x_2) &= x_1(1 - x_1)x_2(1 - x_2). \end{aligned}$$

Exact Solutions

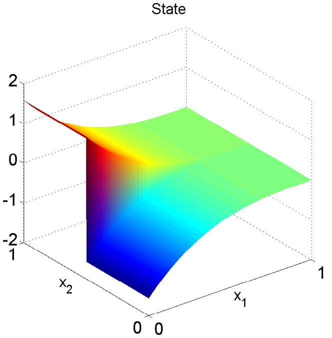


Figure: Surfaces of the exact state (left) and the exact control (right) for $\varepsilon = 10^{-7}$.

Errors

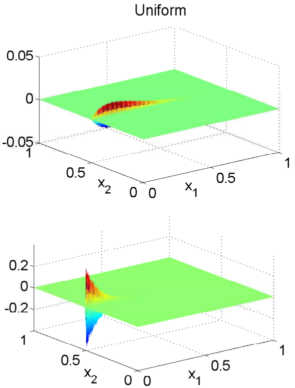


Figure: Error on uniformly refined mesh (16641 nodes) and adaptively refined mesh (9252 nodes) using linear elements for $\varepsilon = 10^{-7}$: state (top row), control (bottom row).

Adaptive Mesh

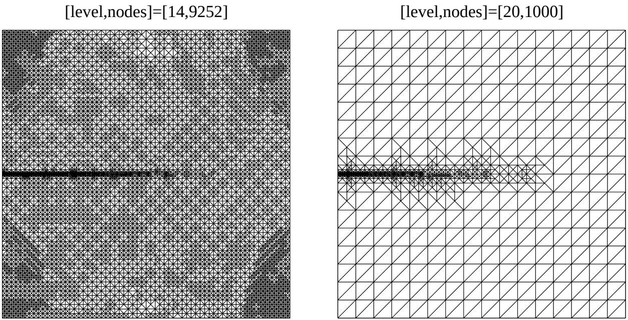


Figure: Adaptively refined meshes with linear elements (left,9252 nodes) and quadratic elements (right, 1000 nodes) for $\varepsilon = 10^{-7}$.

Global Errors

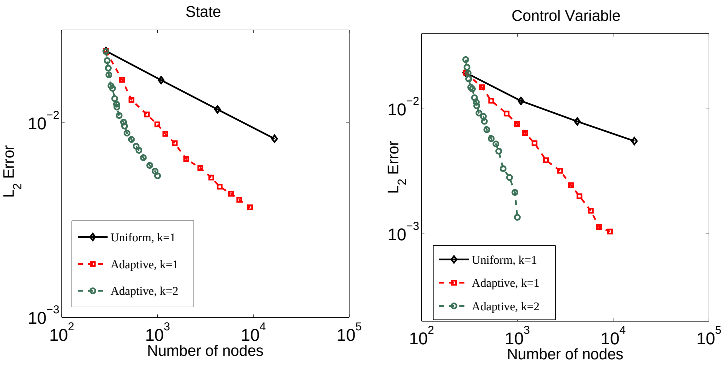


Figure: Errors in L_2 norm using linear and quadratic elements for $\varepsilon = 10^{-7}$.

Control Constrained Optimal Control Problem

$$\min_{u \in U_{ad} \subset U} J(y,u) := \frac{1}{2}\int_\Omega (y(x) - y_d(x))^2 dx + \frac{\omega}{2}\int_\Omega u(x)^2 dx$$

subject to

$$\begin{aligned} -\varepsilon \Delta y(x) + \beta(x)\cdot\nabla y(x) + r(x)y(x) &= f(x) + u(x), \quad x \in \Omega, \\ y(x) &= g_D(x), \quad x \in \Gamma, \end{aligned}$$

where a closed convex set $U_{ad} \subset U = L^2(\Omega)$

$$U_{ad} = \{u \in U : \ u_a \le u \le u_b, \ \text{a.e in } \Omega\},$$

with constants u_a, u_b.

Optimality Conditions

The Lagrange multipliers: $\lambda_a, \lambda_b \in L^2(\Omega)$

$$
\begin{aligned}
-\varepsilon\Delta y + \beta\cdot\nabla y + ry &= f + u, && x \in \Omega,\\
y &= g_D, && x \in \Gamma,\\
-\varepsilon\Delta p - \beta\cdot\nabla p + (r - \nabla\cdot\beta)p &= -(y - y_d), && x \in \Omega,\\
p &= 0, && x \in \Gamma,\\
\omega u - p - \lambda_a + \lambda_b &= 0, && \text{a.e in } \Omega,\\
\lambda_a \geq 0, \quad u_a - u &\leq 0, \quad \lambda_a(u - u_a) = 0 && \text{a.e. in } \Omega,\\
\lambda_b \geq 0, \quad u - u_b &\leq 0, \quad \lambda_b(u_b - u) = 0 && \text{a.e. in } \Omega.
\end{aligned}
$$

Primal Dual Active Set Strategy (PDAS) with Semi-Smooth Newton Method

- Solution operators S, S^* and $\lambda = \lambda_a - \lambda_b$, the complementary conditions [Bergounioux, Ito and Kunish, 1999]:

$$\begin{aligned} -S^*(Su - y_d) + \omega u + \lambda &= 0, \\ \lambda - \min\{0, \lambda - c(u_a - u)\} - \max\{0, \lambda + c(u - u_b)\} &= 0. \end{aligned}$$

- Taking $c = \omega$,

$$\begin{aligned} F(u) := -S^*(Su - y_d) + \omega u &+ \min\{0, S^*(Su - y_d) - \omega u_a\} \\ &+ \max\{0, S^*(Su - y_d) - \omega u_b\} = 0. \end{aligned}$$

- The Newton derivative of $F(u)$

$$G(u) = -S^*S + \omega + (\chi_{A^-(u)} + \chi_{A^+(u)})S^*S = -\chi_{I(u)}S^*S + \omega$$

$$\chi_{A(u)} = \begin{cases} 1, & \text{if } x \in A(u) \\ 0, & \text{otherwise.} \end{cases}$$

Primal Dual Active Set Strategy (PDAS) with Semi-Smooth Newton Method

- Solution operators S, S^* and $\lambda = \lambda_a - \lambda_b$, the complementary conditions [Bergounioux, Ito and Kunish, 1999]:

$$\begin{aligned} -S^*(Su - y_d) + \omega u + \lambda &= 0, \\ \lambda - \min\{0, \lambda - c(u_a - u)\} - \max\{0, \lambda + c(u - u_b)\} &= 0. \end{aligned}$$

- Taking $c = \omega$,

$$\begin{aligned} F(u) := -S^*(Su - y_d) + \omega u \quad &+ \quad \min\{0, S^*(Su - y_d) - \omega u_a\} \\ &+ \quad \max\{0, S^*(Su - y_d) - \omega u_b\} = 0. \end{aligned}$$

- The Newton derivative of $F(u)$

$$G(u) = -S^*S + \omega + (\chi_{A^-(u)} + \chi_{A^+(u)})S^*S = -\chi_{I(u)}S^*S + \omega$$

$$\chi_{A(u)} = \begin{cases} 1, & \text{if } x \in A(u) \\ 0, & \text{otherwise.} \end{cases}$$

Primal Dual Active Set Strategy (PDAS) with Semi-Smooth Newton Method

- Solution operators S, S^* and $\lambda = \lambda_a - \lambda_b$, the complementary conditions [Bergounioux, Ito and Kunish, 1999]:

$$\begin{aligned} -S^*(Su - y_d) + \omega u + \lambda &= 0, \\ \lambda - \min\{0, \lambda - c(u_a - u)\} - \max\{0, \lambda + c(u - u_b)\} &= 0. \end{aligned}$$

- Taking $c = \omega$,

$$\begin{aligned} F(u) := -S^*(Su - y_d) + \omega u &+ \min\{0, S^*(Su - y_d) - \omega u_a\} \\ &+ \max\{0, S^*(Su - y_d) - \omega u_b\} = 0. \end{aligned}$$

- The Newton derivative of $F(u)$

$$G(u) = -S^*S + \omega + (\chi_{A^-(u)} + \chi_{A^+(u)})S^*S = -\chi_{I(u)}S^*S + \omega$$

$$\chi_{A(u)} = \begin{cases} 1, & \text{if } x \in A(u) \\ 0, & \text{otherwise.} \end{cases}$$

Primal Dual Active Set Strategy (PDAS) with Semi-Smooth Newton Method

- Solution operators S, S^* and $\lambda = \lambda_a - \lambda_b$, the complementary conditions [Bergounioux, Ito and Kunish, 1999]:

$$\begin{aligned} -S^*(Su - y_d) + \omega u + \lambda &= 0, \\ \lambda - \min\{0, \lambda - c(u_a - u)\} - \max\{0, \lambda + c(u - u_b)\} &= 0. \end{aligned}$$

- Taking $c = \omega$,

$$\begin{aligned} F(u) := -S^*(Su - y_d) + \omega u &+ \min\{0, S^*(Su - y_d) - \omega u_a\} \\ &+ \max\{0, S^*(Su - y_d) - \omega u_b\} = 0. \end{aligned}$$

- The Newton derivative of $F(u)$

$$G(u) = -S^*S + \omega + (\chi_{A^-(u)} + \chi_{A^+(u)})S^*S = -\chi_{I(u)}S^*S + \omega$$

$$\chi_{A(u)} = \begin{cases} 1, & \text{if } x \in A(u) \\ 0, & \text{otherwise.} \end{cases}$$

Primal Dual Active Set Strategy (PDAS) with Semi-Smooth Newton Method

- Solution operators S, S^* and $\lambda = \lambda_a - \lambda_b$, the complementary conditions [Bergounioux, Ito and Kunish, 1999]:

$$\begin{aligned} -S^*(Su - y_d) + \omega u + \lambda &= 0, \\ \lambda - \min\{0, \lambda - c(u_a - u)\} - \max\{0, \lambda + c(u - u_b)\} &= 0. \end{aligned}$$

- Taking $c = \omega$,

$$\begin{aligned} F(u) := -S^*(Su - y_d) + \omega u &+ \min\{0, S^*(Su - y_d) - \omega u_a\} \\ &+ \max\{0, S^*(Su - y_d) - \omega u_b\} = 0. \end{aligned}$$

- The Newton derivative of $F(u)$

$$G(u) = -S^*S + \omega + (\chi_{A^-(u)} + \chi_{A^+(u)})S^*S = -\chi_{I(u)}S^*S + \omega$$

$$\chi_{A(u)} = \begin{cases} 1, & \text{if } x \in A(u) \\ 0, & \text{otherwise.} \end{cases}$$

Optimality System

The active sets

$$
\begin{aligned}
A^-(u) &= \{x \in \Omega :\ S^*(Su - y_d) - \omega u_a < 0\},\\
A^+(u) &= \{x \in \Omega :\ S^*(Su - y_d) - \omega u_b > 0\},
\end{aligned}
$$

The inactive set $I(u) = \Omega \backslash (A^+(u) \cup A^-(u))$.

- Newton's method,

$$
\omega u_{n+1} - \chi_{I_n} S^*(Su_{n+1} - y_d) = \chi_{A_n^-} \omega u_a + \chi_{A_n^+} \omega u_b.
$$

- DG discretized optimality system:

$$
\begin{pmatrix} \mathbb{M} & 0 & \mathbb{A}_a \\ 0 & \omega \mathbb{Q} & \text{diag}(\chi_I)\mathbb{B} \\ \mathbb{A}_s & \mathbb{B} & 0 \end{pmatrix}
\begin{pmatrix} \vec{y} \\ \vec{u} \\ \vec{p} \end{pmatrix}
=
\begin{pmatrix} \vec{b} \\ \omega \mathbb{Q}(\chi_{A^-} u_a + \chi_{A^+} u_b) \\ \vec{f} \end{pmatrix}.
$$

Optimality System

The active sets

$$A^-(u) = \{x \in \Omega : S^*(Su - y_d) - \omega u_a < 0\},$$
$$A^+(u) = \{x \in \Omega : S^*(Su - y_d) - \omega u_b > 0\},$$

The inactive set $I(u) = \Omega \backslash (A^+(u) \cup A^-(u))$.

- Newton's method,

$$\omega u_{n+1} - \chi_{I_n} S^*(Su_{n+1} - y_d) = \chi_{A_n^-} \omega u_a + \chi_{A_n^+} \omega u_b.$$

- DG discretized optimality system:

$$\begin{pmatrix} \mathbb{M} & 0 & \mathbb{A}_a \\ 0 & \omega\mathbb{Q} & \text{diag}(\chi_I)\mathbb{B} \\ \mathbb{A}_s & \mathbb{B} & 0 \end{pmatrix} \begin{pmatrix} \vec{y} \\ \vec{u} \\ \vec{p} \end{pmatrix} = \begin{pmatrix} \vec{b} \\ \omega\mathbb{Q}(\chi_{A^-} u_a + \chi_{A^+} u_b) \\ \vec{f} \end{pmatrix}.$$

Optimality System

The active sets

$$\begin{aligned} A^-(u) &= \{x \in \Omega : \; S^*(Su - y_d) - \omega u_a < 0\}, \\ A^+(u) &= \{x \in \Omega : \; S^*(Su - y_d) - \omega u_b > 0\}, \end{aligned}$$

The inactive set $I(u) = \Omega \backslash (A^+(u) \cup A^-(u))$.

- Newton's method,

$$\omega u_{n+1} - \chi_{I_n} S^*(Su_{n+1} - y_d) = \chi_{A_n^-} \omega u_a + \chi_{A_n^+} \omega u_b.$$

- DG discretized optimality system:

$$\begin{pmatrix} \mathbb{M} & 0 & \mathbb{A}_a \\ 0 & \omega\mathbb{Q} & \text{diag}(\chi_I)\mathbb{B} \\ \mathbb{A}_s & \mathbb{B} & 0 \end{pmatrix} \begin{pmatrix} \vec{y} \\ \vec{u} \\ \vec{p} \end{pmatrix} = \begin{pmatrix} \vec{b} \\ \omega\mathbb{Q}(\chi_{A^-} u_a + \chi_{A^+} u_b) \\ \vec{f} \end{pmatrix}.$$

Full discretization and variational discretization

- $\|u - u_h\|_{L^2(\Omega)} = \mathscr{O}(h^{3/2})$ by the fully discrete approaches
  - in [Becker and Vexler, 2007] with local projection based stabilization
  - in [Yan and Zhou, 2009] with edge stabilization
- $\|u - u_h\|_{L^2(\Omega)} = \mathscr{O}(h^2)$ by variational discretization, i.e., the control is not discretized, in [Hinze, Yan and Zhou, 2009]
- $\|u - u_h\|_{L^2(\Omega)} = \mathscr{O}(h^2)$ by the **fully discrete approaches using DG** in [Yücel, Heinkenschloss and Karasözen, 2012]

H. Yücel, M. Heikenschloss, and B. Karasözen, A posteriori error estimates of constrained optimal control problem governed by convection diffusion equations using symmetric interior penalty Galerkin method*, Institute of Applied Mathematics, Middle East Technical University, 2012, Preprint*

Full discretization and variational discretization

- $\|u-u_h\|_{L^2(\Omega)} = \mathcal{O}(h^{3/2})$ by the fully discrete approaches
  - in [Becker and Vexler, 2007] with local projection based stabilization
  - in [Yan and Zhou, 2009] with edge stabilization
- $\|u-u_h\|_{L^2(\Omega)} = \mathcal{O}(h^{2})$ by variational discretization, i.e., the control is not discretized, in [Hinze, Yan and Zhou, 2009]
- $\|u-u_h\|_{L^2(\Omega)} = \mathcal{O}(h^{2})$ by the **fully discrete approaches using DG** in [Yücel, Heinkenschloss and Karasözen, 2012]

H. Yücel, M. Heikenschloss, and B. Karasözen, A posteriori error estimates of constrained optimal control problem governed by convection diffusion equations using symmetric interior penalty Galerkin method, *Institute of Applied Mathematics, Middle East Technical University, 2012, Preprint*

Full discretization and variational discretization

- $\|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h^{3/2})$ by the fully discrete approaches
  - in [Becker and Vexler, 2007] with local projection based stabilization
  - in [Yan and Zhou, 2009] with edge stabilization
- $\|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h^2)$ by variational discretization, i.e., the control is not discretized, in [Hinze, Yan and Zhou, 2009]
- $\|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h^2)$ by the **fully discrete approaches using DG** in [Yücel, Heinkenschloss and Karasözen, 2012]

H. Yücel, M. Heikenschloss, and B. Karasözen, A posteriori error estimates of constrained optimal control problem governed by convection diffusion equations using symmetric interior penalty Galerkin method, *Institute of Applied Mathematics, Middle East Technical University, 2012, Preprint*

Full discretization and variational discretization

- $\|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h^{3/2})$ by the fully discrete approaches
  - in [Becker and Vexler, 2007] with local projection based stabilization
  - in [Yan and Zhou, 2009] with edge stabilization
- $\|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h^2)$ by variational discretization, i.e., the control is not discretized, in [Hinze, Yan and Zhou, 2009]
- $\|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h^2)$ by the **fully discrete approaches using DG** in [Yücel, Heinkenschloss and Karasözen, 2012]

H. Yücel, M. Heikenschloss, and B. Karasözen, A posteriori error estimates of constrained optimal control problem governed by convection diffusion equations using symmetric interior penalty Galerkin method, *Institute of Applied Mathematics, Middle East Technical University, 2012, Preprint*

Full discretization and variational discretization

- $\|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h^{3/2})$ by the fully discrete approaches
  - in [Becker and Vexler, 2007] with local projection based stabilization
  - in [Yan and Zhou, 2009] with edge stabilization
- $\|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h^2)$ by variational discretization, i.e., the control is not discretized, in [Hinze, Yan and Zhou, 2009]
- $\|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h^2)$ by the **fully discrete approaches using DG** in [Yücel, Heinkenschloss and Karasözen, 2012]

H. Yücel, M. Heikenschloss, and B. Karasözen, A posteriori error estimates of constrained optimal control problem governed by convection diffusion equations using symmetric interior penalty Galerkin method*, Institute of Applied Mathematics, Middle East Technical University, 2012, Preprint*

Numerical Results

Example (Hinze, Yan and Zhou, 2009)

Let

$$\Omega = [0,1]^2, \varepsilon = 10^{-3}, \beta = (2,3)^T \text{ and } r = 2.$$

The admissible set $U_{ad} = \{\upsilon \in U : \ \upsilon \geq 0\}$. Exact state, adjoint and controls

$$\begin{aligned} y(x_1,x_2) &= 100(1-x_1)^2 x_1^2 x_2 (1-2x_2)(1-x_2), \\ p(x_1,x_2) &= 50(1-x_1)^2 x_1^2 x_2 (1-2x_2)(1-x_2), \\ u(x_1,x_2) &= \max\{0, -\frac{1}{\omega} p(x_1,x_2)\}. \end{aligned}$$

Nodes	$\|y-y_h\|_{L^2}$	order	$\|p-p_h\|_{L^2}$	order	$\|u-u_h\|_{L^2}$	order
25	4.68e-2	-	2.82e-2	-	1.70e-1	-
81	1.24e-2	1.92	6.10e-3	1.90	4.84e-2	1.82
289	3.10e-3	2.00	1.54e-3	1.99	1.20e-2	2.02
1089	7.62e-4	2.02	3.80e-4	2.02	2.86e-3	2.06
4225	1.87e-4	2.02	9.38e-5	2.02	6.92e-4	2.05

Table: Convergence results on uniform meshes

Circular and Straight Interior Layer Example

Example (Hinze, Yan and Zhou, 2009)

$$\Omega = [0,1]^2, \quad \beta = (2,3)^T, \quad r = 1 \text{ and } \omega = 0.1.$$

Exact state

$$y(x_1,x_2) \quad = \quad \frac{2}{\pi}\arctan\left(\frac{1}{\sqrt{\varepsilon}}\left[-\frac{1}{2}x_1 + x_2 - \frac{1}{4}\right]\right),$$

Straight interior layer with the corresponding adjoint

$$\begin{aligned} p(x_1,x_2) &= 16x_1(1-x_1)x_2(1-x_2) \\ &\times \left(\frac{1}{2} + \frac{1}{\pi}\arctan\left[\frac{2}{\sqrt{\varepsilon}}\left(\frac{1}{16} - \left(x_1 - \frac{1}{2}\right)^2 - \left(x_2 - \frac{1}{2}\right)^2\right)\right]\right), \end{aligned}$$

Circular interior layer. Optimal control

$$u(x_1,x_2) = \max\{-5, \min\{-1, -\frac{1}{\omega}p(x_1,x_2)\}\}.$$

Uniform Refinement

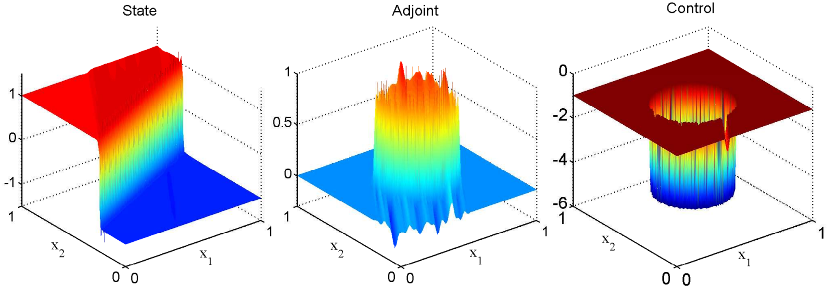


Figure: Uniform mesh (4225 nodes) for $\varepsilon = 10^{-6}$.

Adaptive Refinement

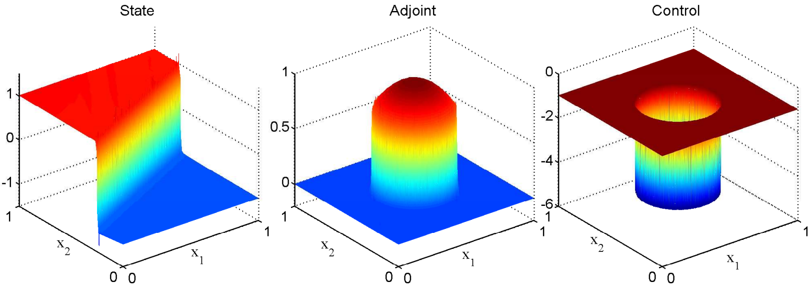


Figure: Adaptively refined mesh (4135 nodes) for $\varepsilon = 10^{-6}$.

Adaptive Mesh

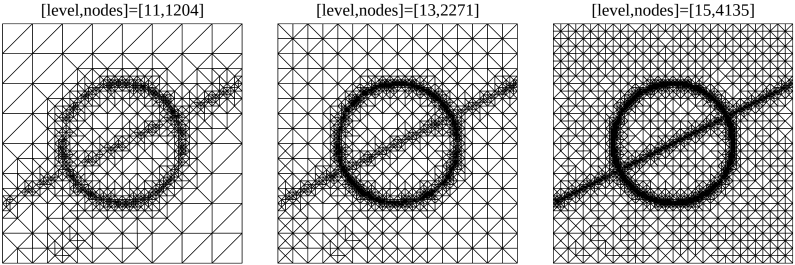


Figure: Adaptively refined meshes at various refinement levels for $\varepsilon = 10^{-6}$.

Global Errors

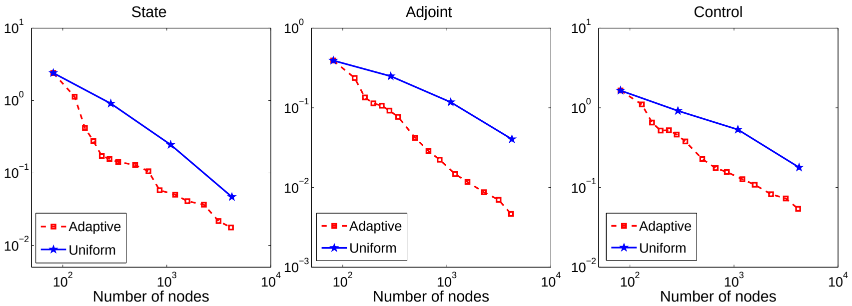


Figure: Errors in L_2 norm for $\varepsilon = 10^{-6}$.

Example with Control desired

$$\underset{u \in U_{ad} \subset U}{\text{minimize}} \; J(y,u) := \frac{1}{2}\int_\Omega (y(x) - y_d(x))^2 dx + \frac{\omega}{2}\int_\Omega (u(x) - u_d(x))^2 dx$$

Example (Yan, Zhou, 2009)

$$\Omega = [0,1]^2, \quad \varepsilon = 10^{-4}, \quad \beta = (1,0), \quad r = 1 \text{ and } \omega = 1.$$

Exact solutions

$$\begin{aligned} y(x_1,x_2) &= 4e^{(-((x_1-1/2)^2+3(x_2-0.5)^2)/\sqrt{\varepsilon})} \sin(\pi x_1)\sin(\pi x_2), \\ p(x_1,x_2) &= e^{(-((x_1-1/2)^2+3(x_2-0.5)^2)/\sqrt{\varepsilon})} \sin(\pi x_1)\sin(\pi x_2), \\ u(x_1,x_2) &= \max\{0, 2\cos(\pi x_1)\cos(\pi x_2) - 1\}. \end{aligned}$$

Uniform Refinement

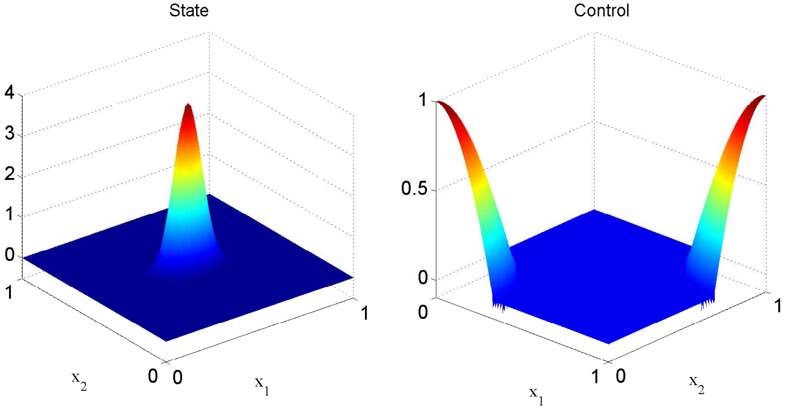


Figure: Uniform mesh (4225 nodes).

Adaptive Refinement

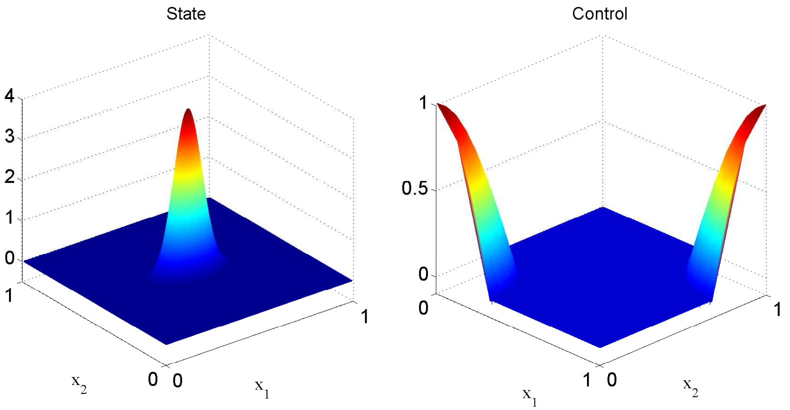


Figure: Adaptively refined mesh (2867 nodes).

Adaptive Mesh

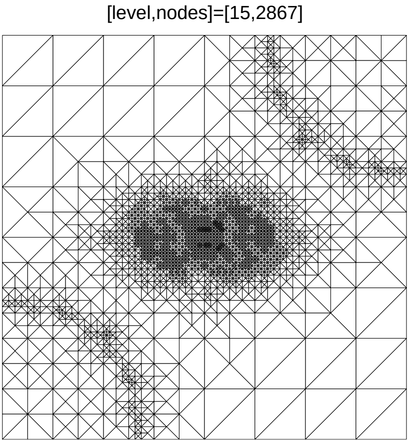


Figure: Adaptively mesh

Global Errors

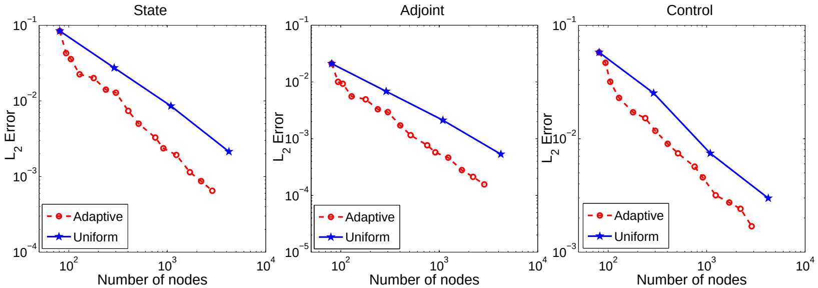


Figure: L_2 errors in the state, the adjoint and the control.

Conclusions

- State and adjoints are polluted with errors around the boundary and interior layers using adaptive FEM with SUPG stabilization.
- For the adaptive SIPG method, meshes are only refined in regions where states or adjoints exhibit layers.
- Optimal convergence orders are obtained for the control constrained problems.

Conclusions

- State and adjoints are polluted with errors around the boundary and interior layers using adaptive FEM with SUPG stabilization.
- For the adaptive SIPG method, meshes are only refined in regions where states or adjoints exhibit layers.
- Optimal convergence orders are obtained for the control constrained problems.

Conclusions

- State and adjoints are polluted with errors around the boundary and interior layers using adaptive FEM with SUPG stabilization.
- For the adaptive SIPG method, meshes are only refined in regions where states or adjoints exhibit layers.
- Optimal convergence orders are obtained for the control constrained problems.

Future Work

- Comparison of different error estimators and convergence analysis
- hp-adaptivity
- Nonconforming meshes
- Goal-oriented error estimates for optimal control problems using DG discretization
- Boundary control problems
- Time dependent optimal control problems
- Nonlinear problems

Future Work

- Comparison of different error estimators and convergence analysis
- hp-adaptivity
- Nonconforming meshes
- Goal-oriented error estimates for optimal control problems using DG discretization
- Boundary control problems
- Time dependent optimal control problems
- Nonlinear problems

Future Work

- Comparison of different error estimators and convergence analysis
- hp-adaptivity
- Nonconforming meshes
- Goal-oriented error estimates for optimal control problems using DG discretization
- Boundary control problems
- Time dependent optimal control problems
- Nonlinear problems

Future Work

- Comparison of different error estimators and convergence analysis
- hp-adaptivity
- Nonconforming meshes
- Goal-oriented error estimates for optimal control problems using DG discretization
- Boundary control problems
- Time dependent optimal control problems
- Nonlinear problems

Future Work

- Comparison of different error estimators and convergence analysis
- hp-adaptivity
- Nonconforming meshes
- Goal-oriented error estimates for optimal control problems using DG discretization
- Boundary control problems
- Time dependent optimal control problems
- Nonlinear problems

Future Work

- Comparison of different error estimators and convergence analysis
- hp-adaptivity
- Nonconforming meshes
- Goal-oriented error estimates for optimal control problems using DG discretization
- Boundary control problems
- Time dependent optimal control problems
- Nonlinear problems

Future Work

- Comparison of different error estimators and convergence analysis
- hp-adaptivity
- Nonconforming meshes
- Goal-oriented error estimates for optimal control problems using DG discretization
- Boundary control problems
- Time dependent optimal control problems
- Nonlinear problems

THANK YOU !