The current landscape of energy a posteriori error estimators

Martin Vohralík

INRIA Paris

28th Chemnitz FEM symposium, September 29, 2015

Outline

Outline

What is an a posteriori error estimate

A posteriori error estimate

- Let u be a weak solution of a PDE.
- Let u_h be its approximate numerical solution.
- A priori error estimate: $\|\nabla(u - u_h)\| \leq C(u)h^k$. Dependent on u, not computable. Useful in theoretical assessment of convergence.
- A posteriori error estimate: $\|\nabla(u - u_h)\| \leq C\eta(u_h)$. Only uses u_h, computable. Great in practical calculation.

Usual form

- Element indicators $\eta_K(u_h)$, $K \in \mathcal{T}_h$.
- Can be used to determine mesh elements with large error.
- We can then refine these elements: mesh adaptivity.

What is an a posteriori error estimate

A posteriori error estimate

- Let u be a weak solution of a PDE.
- Let u_h be its approximate numerical solution.
- A priori error estimate: $\|\nabla(u - u_h)\| \leq C(u)h^k$. Dependent on u, not computable. Useful in theoretical assessment of convergence.
- A posteriori error estimate: $\|\nabla(u - u_h)\| \leq C\eta(u_h)$. Only uses u_h, computable. Great in practical calculation.

Usual form

- Element indicators $\eta_K(u_h)$, $K \in \mathcal{T}_h$.
- Can be used to determine mesh elements with large error.
- We can then refine these elements: mesh adaptivity.

What is an a posteriori error estimate

A posteriori error estimate

- Let u be a weak solution of a PDE.
- Let u_h be its approximate numerical solution.
- A priori error estimate: $\|\nabla(u - u_h)\| \leq C(u)h^k$. Dependent on u, not computable. Useful in theoretical assessment of convergence.
- A posteriori error estimate: $\|\nabla(u - u_h)\| \leq C\eta(u_h)$. Only uses u_h, computable. Great in practical calculation.

Usual form

- Element indicators $\eta_K(u_h)$, $K \in \mathcal{T}_h$.
- Can be used to determine mesh elements with large error.
- We can then refine these elements: mesh adaptivity.

What is an a posteriori error estimate

A posteriori error estimate

- Let u be a weak solution of a PDE.
- Let u_h be its approximate numerical solution.
- A priori error estimate: $\|\nabla(u - u_h)\| \leq C(u)h^k$. Dependent on u, not computable. Useful in theoretical assessment of convergence.
- A posteriori error estimate: $\|\nabla(u - u_h)\| \leq C\eta(u_h)$. Only uses u_h, computable. Great in practical calculation.

Usual form

- Element indicators $\eta_K(u_h)$, $K \in \mathcal{T}_h$.
- Can be used to determine mesh elements with large error.
- We can then refine these elements: mesh adaptivity.

What is an a posteriori error estimate

A posteriori error estimate

- Let u be a weak solution of a PDE.
- Let u_h be its approximate numerical solution.
- A priori error estimate: $\|\nabla(u - u_h)\| \leq C(u)h^k$. Dependent on u, not computable. Useful in theoretical assessment of convergence.
- A posteriori error estimate: $\|\nabla(u - u_h)\| \leq C\eta(u_h)$. Only uses u_h, computable. Great in practical calculation.

Usual form

- Element indicators $\eta_K(u_h)$, $K \in \mathcal{T}_h$.
- Can be used to determine mesh elements with large error.
- We can then refine these elements: mesh adaptivity.

Outline

1. Introduction and the current landscape
   - Optimal properties
   - Various energy a posteriori error estimators
2. Equilibrated fluxes in the Laplace case
   - Guaranteed upper bound
   - Polynomial-degree-robust local efficiency
   - Applications
   - Numerical results
3. Adaptive inexact Newton method
   - Guaranteed a posteriori error estimate
   - Stopping criteria, efficiency, and robustness
   - Numerical results
4. Full adaptivity for unsteady nonlinear problems
5. Guaranteed bounds for Laplace eigen-values and -vectors
6. References and bibliography

Model problem

Model problem

$$
\begin{aligned}
-\Delta u &= f \qquad \text{in } \Omega, \\
u &= 0 \qquad \text{on } \partial\Omega
\end{aligned}
$$

- $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ polygon/polyhedron
- $f \in L^2(\Omega)$

Weak formulation

Find $u \in H_0^1(\Omega)$ such that

$$(\nabla u, \nabla v) = (f, v) \qquad \forall v \in H_0^1(\Omega)$$

Finite element solution

Find $u_h \in V_h := \mathbb{P}_p(\mathcal{T}_h) \cap H_0^1(\Omega)$, $p \geq 1$, such that

$$(\nabla u_h, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_h$$

Model problem

Model problem

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

- $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ polygon/polyhedron
- $f \in L^2(\Omega)$

Weak formulation

Find $u \in H_0^1(\Omega)$ such that

$$(\nabla u, \nabla v) = (f, v) \qquad \forall v \in H_0^1(\Omega)$$

Finite element solution

Find $u_h \in V_h := \mathbb{P}_p(\mathcal{T}_h) \cap H_0^1(\Omega)$, $p \geq 1$, such that

$$(\nabla u_h, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_h$$

Model problem

Model problem

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

- $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ polygon/polyhedron
- $f \in L^2(\Omega)$

Weak formulation

Find $u \in H_0^1(\Omega)$ such that

$$(\nabla u, \nabla v) = (f, v) \qquad \forall v \in H_0^1(\Omega)$$

Finite element solution

Find $u_h \in V_h := \mathbb{P}_p(\mathcal{T}_h) \cap H_0^1(\Omega)$, $p \geq 1$, such that

$$(\nabla u_h, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_h$$

What an a posteriori error estimate should fulfill

Optimal estimate

- **guaranteed upper bound**:

$$\|\nabla(u-u_h)\| \leq \left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}$$

- **local efficiency**:

$$\eta_K(u_h) \leq C_{\mathrm{eff}}\|\nabla(u-u_h)\|_{\omega_K} \qquad \forall K \in \mathcal{T}_h$$

- **asymptotic exactness**:

$$\frac{\left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}}{\|\nabla(u-u_h)\|} \searrow 1$$

- **robustness**: C_{eff} independent of the parameters (size and shape of Ω, regularity of u, refinement of $\mathcal{T}_h$, polynomial degree of u_h) (final time, singular perturbation . . .)
- **small evaluation cost** of $\eta_K(u_h)$
- **error components identification**: discretization, algebraic (linearization, time stepping . . .)

What an a posteriori error estimate should fulfill

Optimal estimate

- **guaranteed upper bound**:

$$\|\nabla(u-u_h)\| \leq \left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}$$

- **local efficiency**:

$$\eta_K(u_h) \leq C_{\mathrm{eff}}\|\nabla(u-u_h)\|_{\omega_K} \qquad \forall K\in\mathcal{T}_h$$

- **asymptotic exactness**:

$$\frac{\left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}}{\|\nabla(u-u_h)\|} \searrow 1$$

- **robustness**: C_{eff} independent of the parameters (size and shape of Ω, regularity of u, refinement of $\mathcal{T}_h$, polynomial degree of u_h) (final time, singular perturbation ...)
- **small evaluation cost** of $\eta_K(u_h)$
- **error components identification**: discretization, algebraic (linearization, time stepping ...)

What an a posteriori error estimate should fulfill

Optimal estimate

- **guaranteed upper bound**:

$$\|\nabla(u-u_h)\| \leq \left\{ \sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2 \right\}^{\frac{1}{2}}$$

- **local efficiency**:

$$\eta_K(u_h) \leq C_{\rm eff}\|\nabla(u-u_h)\|_{\omega_K} \qquad \forall K \in \mathcal{T}_h$$

- **asymptotic exactness**:

$$\frac{\left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}}{\|\nabla(u-u_h)\|} \searrow 1$$

- **robustness**: $C_{\rm eff}$ independent of the parameters (size and shape of Ω, regularity of u, refinement of $\mathcal{T}_h$, polynomial degree of u_h) (final time, singular perturbation . . .)
- **small evaluation cost** of $\eta_K(u_h)$
- **error components identification**: discretization, algebraic (linearization, time stepping . . .)

What an a posteriori error estimate should fulfill

Optimal estimate

- **guaranteed upper bound**:

$$\|\nabla(u-u_h)\| \le \left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}$$

- **local efficiency**:

$$\eta_K(u_h) \le C_{\mathrm{eff}}\|\nabla(u-u_h)\|_{\omega_K} \qquad \forall K\in\mathcal{T}_h$$

- **asymptotic exactness**:

$$\frac{\left\{\sum_{K\in\mathcal{T}_h}\eta_K(u_h)^2\right\}^{\frac{1}{2}}}{\|\nabla(u-u_h)\|} \searrow 1$$

- **robustness**: C_{eff} independent of the parameters (size and shape of Ω, regularity of u, refinement of $\mathcal{T}_h$, polynomial degree of u_h) (final time, singular perturbation ...)
- **small evaluation cost** of $\eta_K(u_h)$
- **error components identification**: discretization, algebraic (linearization, time stepping ...)

What an a posteriori error estimate should fulfill

Optimal estimate

- **guaranteed upper bound**:

$$\|\nabla(u-u_h)\| \le \left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}$$

- **local efficiency**:

$$\eta_K(u_h) \le C_{\rm eff}\|\nabla(u-u_h)\|_{\omega_K} \qquad \forall K \in \mathcal{T}_h$$

- **asymptotic exactness**:

$$\frac{\left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}}{\|\nabla(u-u_h)\|} \searrow 1$$

- **robustness**: $C_{\rm eff}$ independent of the parameters (size and shape of Ω, regularity of u, refinement of $\mathcal{T}_h$, polynomial degree of u_h) (final time, singular perturbation . . .)
- **small evaluation cost** of $\eta_K(u_h)$
- **error components identification**: discretization, algebraic (linearization, time stepping . . .)

What an a posteriori error estimate should fulfill

Optimal estimate

- **guaranteed upper bound**:

$$\|\nabla(u - u_h)\| \leq \left\{ \sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2 \right\}^{\frac{1}{2}}$$

- **local efficiency**:

$$\eta_K(u_h) \leq C_{\text{eff}} \|\nabla(u - u_h)\|_{\omega_K} \qquad \forall K \in \mathcal{T}_h$$

- **asymptotic exactness**:

$$\frac{\left\{ \sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2 \right\}^{\frac{1}{2}}}{\|\nabla(u - u_h)\|} \searrow 1$$

- **robustness**: C_{eff} independent of the parameters (size and shape of Ω, regularity of u, refinement of $\mathcal{T}_h$, polynomial degree of u_h) (final time, singular perturbation ...)
- **small evaluation cost** of $\eta_K(u_h)$
- **error components identification**: discretization, algebraic (linearization, time stepping ...)

What an a posteriori error estimate should fulfill

Optimal estimate

- **guaranteed upper bound**:

$$\|\nabla(u-u_h)\| \le \left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}$$

- **local efficiency**:

$$\eta_K(u_h) \le C_{\mathrm{eff}}\|\nabla(u-u_h)\|_{\omega_K} \qquad \forall K \in \mathcal{T}_h$$

- **asymptotic exactness**:

$$\frac{\left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}}{\|\nabla(u-u_h)\|} \searrow 1$$

- **robustness**: C_{eff} independent of the parameters (size and shape of Ω, regularity of u, refinement of $\mathcal{T}_h$, polynomial degree of u_h) (final time, singular perturbation ...)
- **small evaluation cost** of $\eta_K(u_h)$
- **error components identification**: discretization, algebraic (linearization, time stepping ...)

What an a posteriori error estimate should fulfill

Optimal estimate

- **guaranteed upper bound**:

$$\|\nabla(u-u_h)\| \leq \left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}$$

- **local efficiency**:

$$\eta_K(u_h) \leq C_{\mathrm{eff}}\|\nabla(u-u_h)\|_{\omega_K} \qquad \forall K \in \mathcal{T}_h$$

- **asymptotic exactness**:

$$\frac{\left\{\sum_{K\in\mathcal{T}_h} \eta_K(u_h)^2\right\}^{\frac{1}{2}}}{\|\nabla(u-u_h)\|} \searrow 1$$

- **robustness**: C_{eff} independent of the parameters (size and shape of Ω, regularity of u, refinement of $\mathcal{T}_h$, polynomial degree of u_h) (final time, singular perturbation . . .)
- **small evaluation cost** of $\eta_K(u_h)$
- **error components identification**: discretization, algebraic (linearization, time stepping . . .)

Outline

1. Introduction and the current landscape
   - Optimal properties
   - Various energy a posteriori error estimators
2. Equilibrated fluxes in the Laplace case
   - Guaranteed upper bound
   - Polynomial-degree-robust local efficiency
   - Applications
   - Numerical results
3. Adaptive inexact Newton method
   - Guaranteed a posteriori error estimate
   - Stopping criteria, efficiency, and robustness
   - Numerical results
4. Full adaptivity for unsteady nonlinear problems
5. Guaranteed bounds for Laplace eigen-values and -vectors
6. References and bibliography

Residual-based estimates

Theorem (Residual-based estimate, Verfürth (1989))

Let $u \in H_0^1(\Omega)$ be the weak solution and let $u_h \in V_h$ be its FE approximation. Then there exists $C_{\rm up} > 0$ only depending on d (space dimension) and $\kappa_{\mathcal{T}_h}$ (shape-regularity of $\mathcal{T}_h$) such that

$$\|\nabla(u-u_h)\| \leq C_{\rm up} \left\{ \sum_{K \in \mathcal{T}_h} (\underbrace{h_K \|f + \Delta u_h\|_K + h_K^{\frac{1}{2}} \|[\![\nabla u_h]\!] \cdot \mathbf{n}_K\|_{\partial K}}_{\eta_{{\rm res},K}})^2 \right\}^{\frac{1}{2}}.$$

For $C_{\rm low}$ additionally depending p (polynomial degree),

$$\eta_{{\rm res},K} \leq C_{\rm low}(\|\nabla(u-u_h)\|_{\omega_K} + \|f - \Pi_p f\|_{\omega_K}) \quad \forall K \in \mathcal{T}_h.$$

Properties

- ⊕ rigorous
- ⊕ cheap (explicit)
- ⊖ what is $C_{\rm up}$?
- ⊖ $C_{\rm low}$ depends on p

Residual-based estimates

Theorem (Residual-based estimate, Verfürth (1989))

Let $u \in H_0^1(\Omega)$ be the weak solution and let $u_h \in V_h$ be its FE approximation. Then there exists $C_{\rm up} > 0$ only depending on d (space dimension) and $\kappa_{\mathcal{T}_h}$ (shape-regularity of $\mathcal{T}_h$) such that

$$\|\nabla(u-u_h)\| \leq C_{\rm up}\left\{\sum_{K\in\mathcal{T}_h}\underbrace{(h_K\|f+\Delta u_h\|_K + h_K^{\frac{1}{2}}\|[\![\nabla u_h]\!]\cdot\mathbf{n}_K\|_{\partial K})}_{\eta_{{\rm res},K}}^2\right\}^{\frac{1}{2}}.$$

For $C_{\rm low}$ additionally depending p (polynomial degree),

$$\eta_{{\rm res},K} \leq C_{\rm low}(\|\nabla(u-u_h)\|_{\omega_K} + \|f-\Pi_p f\|_{\omega_K}) \quad \forall K \in \mathcal{T}_h.$$

Properties

- ⊕ rigorous
- ⊕ cheap (explicit)
- ⊖ what is $C_{\rm up}$?
- ⊖ $C_{\rm low}$ depends on p

Residual-based estimates

Theorem (Residual-based estimate, Verfürth (1989))

Let $u \in H_0^1(\Omega)$ be the weak solution and let $u_h \in V_h$ be its FE approximation. Then there exists $C_{\rm up} > 0$ only depending on d (space dimension) and $\kappa_{\mathcal{T}_h}$ (shape-regularity of $\mathcal{T}_h$) such that

$$\|\nabla(u-u_h)\| \leq C_{\rm up} \left\{ \sum_{K \in \mathcal{T}_h} \underbrace{(h_K \|f + \Delta u_h\|_K + h_K^{\frac{1}{2}} \|[\![\nabla u_h]\!] \cdot \mathbf{n}_K\|_{\partial K})}_{\eta_{{\rm res},K}}{}^2 \right\}^{\frac{1}{2}}.$$

For $C_{\rm low}$ additionally depending p (polynomial degree),

$$\eta_{{\rm res},K} \leq C_{\rm low} (\|\nabla(u-u_h)\|_{\omega_K} + \|f - \Pi_p f\|_{\omega_K}) \quad \forall K \in \mathcal{T}_h.$$

Properties

- ⊕ rigorous
- ⊕ cheap (explicit)
- ⊖ what is $C_{\rm up}$?
- ⊖ $C_{\rm low}$ depends on p

Averaging estimates

Theorem (Averaging estimate, Zienkiewicz and Zhu (1987))

Let $u \in H_0^1(\Omega)$ and let $u_h \in V_h$ be the FE approximation. Then

$$\|\nabla(u - u_h)\| \lesssim \|\nabla u_h + \boldsymbol{\sigma}_h\|,$$

where $\boldsymbol{\sigma}_h \in [\mathbb{P}_{p'}(\mathcal{T}_h) \cap H^1(\Omega)]^d$ is obtained by local averaging (smoothing) of $-\nabla u_h$.

Properties

- ⊕ cheap (explicit)
- ⊕ often asymptotically exact
- problem-independent
- ⊖ upper bound (may be severely violated on coarse meshes), efficiency (under certain conditions, however, equivalence with residual-based estimates can be shown)

Averaging estimates

Theorem (Averaging estimate, Zienkiewicz and Zhu (1987))

Let $u \in H_0^1(\Omega)$ and let $u_h \in V_h$ be the FE approximation. Then

$$\|\nabla(u - u_h)\| \lesssim \|\nabla u_h + \boldsymbol{\sigma}_h\|,$$

where $\boldsymbol{\sigma}_h \in [\mathbb{P}_{p'}(\mathcal{T}_h) \cap H^1(\Omega)]^d$ is obtained by local averaging (smoothing) of $-\nabla u_h$.

Properties

- ⊕ cheap (explicit)
- ⊕ often asymptotically exact
- problem-independent
- ⊖ upper bound (may be severely violated on coarse meshes), efficiency (under certain conditions, however, equivalence with residual-based estimates can be shown)

Functional estimates

Theorem (Functional estimate, $\approx$ Prager and Synge (1947), Hlaváček, Haslinger, Nečas, and Lovíšek (1979), Repin (1997)**)**

Let $u \in H_0^1(\Omega)$ be the weak solution and let $u_h \in H_0^1(\Omega)$ be arbitrary. Let $\boldsymbol{\sigma}_h \in \mathbf{H}(\mathrm{div}, \Omega)$ be arbitrary. Then

$$\|\nabla(u - u_h)\| \leq \|\nabla u_h + \boldsymbol{\sigma}_h\| + h_\Omega \|f - \nabla\cdot\boldsymbol{\sigma}_h\|.$$

Properties

- ⊕ general (no requirement on u_h)
- ⊕ guaranteed upper bound
- ⊖ h_Ω may be large
- ⊖ no local efficiency
- ⊖ potentially expensive (global solve for σ_h)

Functional estimates

Theorem (Functional estimate, ≈ Prager and Synge (1947), Hlaváček, Haslinger, Nečas, and Lovíšek (1979), Repin (1997))

Let $u \in H_0^1(\Omega)$ be the weak solution and let $u_h \in H_0^1(\Omega)$ be arbitrary. Let $\boldsymbol{\sigma}_h \in \mathbf{H}(\mathrm{div}, \Omega)$ be arbitrary. Then

$$\|\nabla(u - u_h)\| \leq \|\nabla u_h + \boldsymbol{\sigma}_h\| + h_\Omega \|f - \nabla\cdot\boldsymbol{\sigma}_h\|.$$

Properties

- ⊕ general (no requirement on u_h)
- ⊕ guaranteed upper bound
- ⊖ h_Ω may be large
- ⊖ no local efficiency
- ⊖ potentially expensive (global solve for $\boldsymbol{\sigma}_h$)

The current landscape

Estimator	Guar. bound	Loc. eff.	Asymp. exact.	Rob. data	Rob. p	Cost	Maximal overest.	Err. comps.
Residual	✗	✓	✗	✓	✗	explicit	✗	✗
Averaging	✗	✗	✓	✗	✗	explicit	✗	✓
Hierarchical	✗	✓	✓	✓	✓	global	✗	✗
Functional	✓	✗	✓	✗	✗	global	✗	✓
Equil. res.	✗	✓	✓	✓	✗	implicit	✗	✓
Geometric	✗	✗	✓	✗	✗	explicit	✗	✓
Equil. flux	✓	✓	✓	✓	✓	implicit	✓	✓

Outline

Outline

Equilibrated flux a posteriori estimate

Properties of the weak solution

- $u \in H^1_0(\Omega)$ (constraint)
- $\boldsymbol{\sigma} := -\nabla u$ (constitutive relation)
- $\nabla\cdot\boldsymbol{\sigma} = f$ (equilibrium)
- $\boldsymbol{\sigma} \in \mathbf{H}(\mathrm{div}, \Omega)$ (constraint)

Theorem (Equilibrated flux estimate, ≈ Prager and Synge (1947)**)**

- *Let $u \in H^1_0(\Omega)$ be the weak solution;*
- *$u_h \in H^1(\mathcal{T}_h)$ (piecewise $H^1(K)$) be arbitrary;*
- *$s_h \in H^1_0(\Omega)$, $\boldsymbol{\sigma}_h \in \mathbf{H}(\mathrm{div}, \Omega)$, $(\nabla\cdot\boldsymbol{\sigma}_h, 1)_K = (f, 1)_K \ \forall K \in \mathcal{T}_h$.*

Then

$$\|\nabla(u-u_h)\|^2 \le \sum_{K\in\mathcal{T}_h}\Bigg(\underbrace{\|\nabla u_h + \boldsymbol{\sigma}_h\|_K}_{\text{constitutive relation}} + \underbrace{\frac{h_K}{\pi}\|f-\nabla\cdot\boldsymbol{\sigma}_h\|_K}_{\text{equilibrium}}\Bigg)^2 + \sum_{K\in\mathcal{T}_h}\underbrace{\|\nabla(u_h-s_h)\|_K^2}_{\text{constraint}}.$$

Equilibrated flux a posteriori estimate

Properties of the weak solution

- $u \in H_0^1(\Omega)$ (constraint)
- $\boldsymbol{\sigma} := -\nabla u$ (constitutive relation)
- $\nabla \cdot \boldsymbol{\sigma} = f$ (equilibrium)
- $\boldsymbol{\sigma} \in \mathbf{H}(\mathrm{div}, \Omega)$ (constraint)

Theorem (Equilibrated flux estimate, ≈ Prager and Synge (1947))

- *Let $u \in H_0^1(\Omega)$ be the weak solution;*
- *$u_h \in H^1(\mathcal{T}_h)$ (piecewise $H^1(K)$) be arbitrary;*
- $s_h \in H_0^1(\Omega)$, $\boldsymbol{\sigma}_h \in \mathbf{H}(\mathrm{div}, \Omega)$, $(\nabla \cdot \boldsymbol{\sigma}_h, 1)_K = (f, 1)_K \ \forall K \in \mathcal{T}_h$.

Then

$$\|\nabla(u - u_h)\|^2 \leq \sum_{K \in \mathcal{T}_h} \Big(\underbrace{\|\nabla u_h + \boldsymbol{\sigma}_h\|_K}_{\text{constitutive relation}} + \underbrace{\frac{h_K}{\pi}\|f - \nabla \cdot \boldsymbol{\sigma}_h\|_K}_{\text{equilibrium}} \Big)^2 + \sum_{K \in \mathcal{T}_h} \underbrace{\|\nabla(u_h - s_h)\|_K^2}_{\text{constraint}}.$$

Equilibrated flux a posteriori estimate

Properties of the weak solution

- $u \in H_0^1(\Omega)$ (constraint)
- $\boldsymbol{\sigma} := -\nabla u$ (constitutive relation)
- $\nabla \cdot \boldsymbol{\sigma} = f$ (equilibrium)
- $\boldsymbol{\sigma} \in \mathbf{H}(\mathrm{div}, \Omega)$ (constraint)

Theorem (Equilibrated flux estimate, ≈ Prager and Synge (1947))

- *Let $u \in H_0^1(\Omega)$ be the weak solution;*
- *$u_h \in H^1(\mathcal{T}_h)$ (piecewise $H^1(K)$) be arbitrary;*
- $s_h \in H_0^1(\Omega)$, $\boldsymbol{\sigma}_h \in \mathbf{H}(\mathrm{div}, \Omega)$, $(\nabla \cdot \boldsymbol{\sigma}_h, 1)_K = (f, 1)_K \ \forall K \in \mathcal{T}_h$.

Then

$$\|\nabla(u - u_h)\|^2 \leq \sum_{K \in \mathcal{T}_h} \Big(\underbrace{\|\nabla u_h + \boldsymbol{\sigma}_h\|_K}_{\text{constitutive relation}} + \underbrace{\frac{h_K}{\pi} \|f - \nabla \cdot \boldsymbol{\sigma}_h\|_K}_{\text{equilibrium}} \Big)^2 + \sum_{K \in \mathcal{T}_h} \underbrace{\|\nabla(u_h - s_h)\|_K^2}_{\text{constraint}}.$$

Global potential and flux reconstructions

Ideally

$$\boldsymbol{\sigma}_h := \arg \min_{\mathbf{v}_h \in \mathbf{V}_h,\, \nabla\cdot\mathbf{v}_h = \Pi_{Q_h} f} \|\nabla u_h + \mathbf{v}_h\|$$

$$s_h := \arg \min_{v_h \in V_h} \|\nabla(u_h - v_h)\|$$

- $\mathbf{V}_h \subset \mathbf{H}(\mathrm{div}, \Omega)$, $Q_h \subset L^2(\Omega)$, $V_h \subset H_0^1(\Omega)$
- too expensive, **global** minimization problems (the hypercircle method)

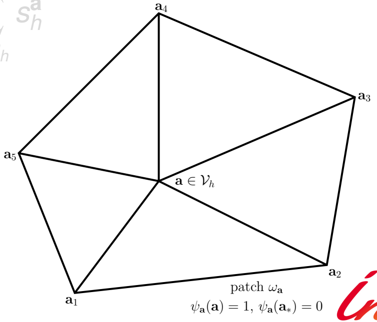
Local potential and flux reconstructions

Partition of unity localization

$$\boldsymbol{\sigma}_h^{\mathbf{a}} := \arg \min_{\mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},\ \nabla\cdot\mathbf{v}_h = ?} \|\psi_{\mathbf{a}} \nabla u_h + \mathbf{v}_h\|_{\omega_{\mathbf{a}}}$$

$$s_h^{\mathbf{a}} := \arg \min_{v_h \in V_h^{\mathbf{a}}} \|\nabla(\psi_{\mathbf{a}} u_h - v_h)\|_{\omega_{\mathbf{a}}}$$

- cut-off by hat basis functions $\psi_{\mathbf{a}} \in \mathbb{P}_1(\mathcal{T}_h) \cap H^1(\Omega)$
- $\boldsymbol{\sigma}_h := \sum_{\mathbf{a}\in\mathcal{V}_h} \boldsymbol{\sigma}_h^{\mathbf{a}}, \quad s_h := \sum_{\mathbf{a}\in\mathcal{V}_h} s_h^{\mathbf{a}}$
- local minimizations



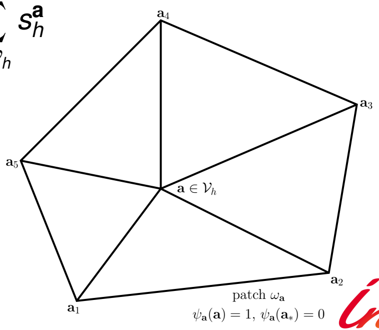
Local potential and flux reconstructions

Partition of unity localization

$$\boldsymbol{\sigma}_h^{\mathbf{a}} := \arg \min_{\mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},\, \nabla\cdot\mathbf{v}_h = ?} \|\psi_{\mathbf{a}} \nabla u_h + \mathbf{v}_h\|_{\omega_{\mathbf{a}}}$$

$$s_h^{\mathbf{a}} := \arg \min_{v_h \in V_h^{\mathbf{a}}} \|\nabla(\psi_{\mathbf{a}} u_h - v_h)\|_{\omega_{\mathbf{a}}}$$

- cut-off by hat basis functions $\psi_{\mathbf{a}} \in \mathbb{P}_1(\mathcal{T}_h) \cap H^1(\Omega)$
- $\boldsymbol{\sigma}_h := \sum_{\mathbf{a}\in\mathcal{V}_h} \boldsymbol{\sigma}_h^{\mathbf{a}}, \quad s_h := \sum_{\mathbf{a}\in\mathcal{V}_h} s_h^{\mathbf{a}}$
- local minimizations



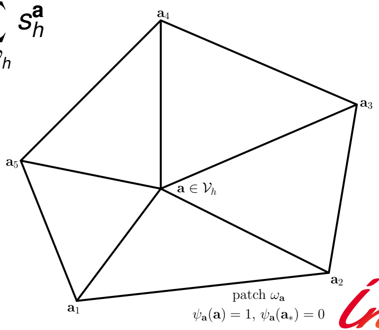
Local potential and flux reconstructions

Partition of unity localization

$$\boldsymbol{\sigma}_h^{\mathbf{a}} := \arg \min_{\mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},\, \nabla\cdot\mathbf{v}_h = ?} \|\psi_{\mathbf{a}} \nabla u_h + \mathbf{v}_h\|_{\omega_{\mathbf{a}}}$$

$$s_h^{\mathbf{a}} := \arg \min_{v_h \in V_h^{\mathbf{a}}} \|\nabla(\psi_{\mathbf{a}} u_h - v_h)\|_{\omega_{\mathbf{a}}}$$

- cut-off by hat basis functions $\psi_{\mathbf{a}} \in \mathbb{P}_1(\mathcal{T}_h) \cap H^1(\Omega)$
- $\boldsymbol{\sigma}_h := \sum_{\mathbf{a}\in\mathcal{V}_h} \boldsymbol{\sigma}_h^{\mathbf{a}}, \quad s_h := \sum_{\mathbf{a}\in\mathcal{V}_h} s_h^{\mathbf{a}}$
- **local** minimizations



Local equilibrated flux reconstruction

Assumption A (Galerkin orthogonality wrt hat functions)

There holds $u_h \in H^1(\mathcal{T}_h)$ *and*

$$(\nabla u_h, \nabla \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = (f, \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} \qquad \forall \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}.$$

$\mathbf{V}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$: MFE spaces (hom. Neumann BC for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$, homogeneous Dirichlet BC on $\partial\omega_{\mathbf{a}} \cap \partial\Omega$ for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}$)

Definition (Constr. of $\boldsymbol{\sigma}_h$, Destuynder and Métivet (1999) & Braess and Schöberl (2008))

Let Assumption A be satisfied. For each $\mathbf{a} \in \mathcal{V}_h$, prescribe $\boldsymbol{\sigma}_h^{\mathbf{a}} \in \mathbf{V}_h^{\mathbf{a}}$ and $\bar{r}_h^{\mathbf{a}} \in Q_h^{\mathbf{a}}$ by solving the **local mixed FE problem**

$$\boldsymbol{\sigma}_h^{\mathbf{a}} := \arg \min_{\mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},\, \nabla\cdot\mathbf{v}_h = \Pi_{Q_h^{\mathbf{a}}}(\psi_{\mathbf{a}} f - \nabla\psi_{\mathbf{a}}\cdot\nabla u_h)} \|\psi_{\mathbf{a}} \nabla u_h + \mathbf{v}_h\|_{\omega_{\mathbf{a}}}$$

$$\Updownarrow$$

$$\begin{aligned}
(\boldsymbol{\sigma}_h^{\mathbf{a}}, \mathbf{v}_h)_{\omega_{\mathbf{a}}} - (\bar{r}_h^{\mathbf{a}}, \nabla\cdot\mathbf{v}_h)_{\omega_{\mathbf{a}}} &= -(\psi_{\mathbf{a}} \nabla u_h, \mathbf{v}_h)_{\omega_{\mathbf{a}}} \qquad && \forall \mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},\\
(\nabla\cdot\boldsymbol{\sigma}_h^{\mathbf{a}}, q_h)_{\omega_{\mathbf{a}}} &= (\psi_{\mathbf{a}} f - \nabla\psi_{\mathbf{a}}\cdot\nabla u_h, q_h)_{\omega_{\mathbf{a}}} && \forall q_h \in Q_h^{\mathbf{a}}.
\end{aligned}$$

Local equilibrated flux reconstruction

Assumption A (Galerkin orthogonality wrt hat functions)

There holds $u_h \in H^1(\mathcal{T}_h)$ and

$$(\nabla u_h, \nabla \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = (f, \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} \qquad \forall \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}.$$

$\mathbf{V}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$: MFE spaces (hom. Neumann BC for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$, homogeneous Dirichlet BC on $\partial\omega_{\mathbf{a}} \cap \partial\Omega$ for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}$)

Definition (Constr. of $\boldsymbol{\sigma}_h$, Destuynder and Métivet (1999) & Braess and Schöberl (2008))

Let Assumption A be satisfied. For each $\mathbf{a} \in \mathcal{V}_h$, prescribe $\boldsymbol{\sigma}_h^{\mathbf{a}} \in \mathbf{V}_h^{\mathbf{a}}$ and $\bar{r}_h^{\mathbf{a}} \in Q_h^{\mathbf{a}}$ by solving the **local mixed FE problem**

$$\boldsymbol{\sigma}_h^{\mathbf{a}} := \arg \min_{\mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},\, \nabla\cdot\mathbf{v}_h = \Pi_{Q_h^{\mathbf{a}}}(\psi_{\mathbf{a}} f - \nabla\psi_{\mathbf{a}}\cdot\nabla u_h)} \|\psi_{\mathbf{a}}\nabla u_h + \mathbf{v}_h\|_{\omega_{\mathbf{a}}}$$

$$\Updownarrow$$

$$(\boldsymbol{\sigma}_h^{\mathbf{a}}, \mathbf{v}_h)_{\omega_{\mathbf{a}}} - (\bar{r}_h^{\mathbf{a}}, \nabla\cdot\mathbf{v}_h)_{\omega_{\mathbf{a}}} = -(\psi_{\mathbf{a}}\nabla u_h, \mathbf{v}_h)_{\omega_{\mathbf{a}}} \qquad \forall \mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},$$
$$(\nabla\cdot\boldsymbol{\sigma}_h^{\mathbf{a}}, q_h)_{\omega_{\mathbf{a}}} = (\psi_{\mathbf{a}} f - \nabla\psi_{\mathbf{a}}\cdot\nabla u_h, q_h)_{\omega_{\mathbf{a}}} \qquad \forall q_h \in Q_h^{\mathbf{a}}.$$

ematics

Local equilibrated flux reconstruction

Assumption A (Galerkin orthogonality wrt hat functions)

There holds $u_h \in H^1(\mathcal{T}_h)$ *and*

$$(\nabla u_h, \nabla \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = (f, \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} \qquad \forall \mathbf{a} \in \mathcal{V}_h^{\rm int}.$$

$\mathbf{V}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$: MFE spaces (hom. Neumann BC for $\mathbf{a} \in \mathcal{V}_h^{\rm int}$, homogeneous Dirichlet BC on $\partial\omega_{\mathbf{a}} \cap \partial\Omega$ for $\mathbf{a} \in \mathcal{V}_h^{\rm ext}$)

Definition (Constr. of $\boldsymbol{\sigma}_h$, Destuynder and Métivet (1999) & Braess and Schöberl (2008))

Let Assumption A be satisfied. For each $\mathbf{a} \in \mathcal{V}_h$, prescribe $\boldsymbol{\sigma}_h^{\mathbf{a}} \in \mathbf{V}_h^{\mathbf{a}}$ and $\bar{r}_h^{\mathbf{a}} \in Q_h^{\mathbf{a}}$ by solving the **local mixed FE problem**

$$\boldsymbol{\sigma}_h^{\mathbf{a}} := \arg \min_{\mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},\, \nabla\cdot\mathbf{v}_h = \Pi_{Q_h^{\mathbf{a}}}(\psi_{\mathbf{a}} f - \nabla\psi_{\mathbf{a}}\cdot\nabla u_h)} \|\psi_{\mathbf{a}} \nabla u_h + \mathbf{v}_h\|_{\omega_{\mathbf{a}}}$$

$$\Updownarrow$$

$$(\boldsymbol{\sigma}_h^{\mathbf{a}}, \mathbf{v}_h)_{\omega_{\mathbf{a}}} - (\bar{r}_h^{\mathbf{a}}, \nabla\cdot\mathbf{v}_h)_{\omega_{\mathbf{a}}} = -(\psi_{\mathbf{a}} \nabla u_h, \mathbf{v}_h)_{\omega_{\mathbf{a}}} \qquad \forall \mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},$$

$$(\nabla\cdot\boldsymbol{\sigma}_h^{\mathbf{a}}, q_h)_{\omega_{\mathbf{a}}} = (\psi_{\mathbf{a}} f - \nabla\psi_{\mathbf{a}}\cdot\nabla u_h, q_h)_{\omega_{\mathbf{a}}} \qquad \forall q_h \in Q_h^{\mathbf{a}}.$$

Local equilibrated flux reconstruction

Assumption A (Galerkin orthogonality wrt hat functions)

There holds $u_h \in H^1(\mathcal{T}_h)$ *and*

$$(\nabla u_h, \nabla \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = (f, \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} \qquad \forall \mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}.$$

$\mathbf{V}_h^{\mathbf{a}} \times Q_h^{\mathbf{a}}$: MFE spaces (hom. Neumann BC for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$, homogeneous Dirichlet BC on $\partial\omega_{\mathbf{a}} \cap \partial\Omega$ for $\mathbf{a} \in \mathcal{V}_h^{\mathrm{ext}}$)

Definition (Constr. of $\boldsymbol{\sigma}_h$, Destuynder and Métivet (1999) & Braess and Schöberl (2008))

Let Assumption A be satisfied. For each $\mathbf{a} \in \mathcal{V}_h$, prescribe $\boldsymbol{\sigma}_h^{\mathbf{a}} \in \mathbf{V}_h^{\mathbf{a}}$ and $\bar{r}_h^{\mathbf{a}} \in Q_h^{\mathbf{a}}$ by solving the **local mixed FE problem**

$$\boldsymbol{\sigma}_h^{\mathbf{a}} := \arg \min_{\mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},\, \nabla\cdot\mathbf{v}_h = \Pi_{Q_h^{\mathbf{a}}}(\psi_{\mathbf{a}} f - \nabla\psi_{\mathbf{a}}\cdot\nabla u_h)} \|\psi_{\mathbf{a}} \nabla u_h + \mathbf{v}_h\|_{\omega_{\mathbf{a}}}$$

$$\Updownarrow$$

$$(\boldsymbol{\sigma}_h^{\mathbf{a}}, \mathbf{v}_h)_{\omega_{\mathbf{a}}} - (\bar{r}_h^{\mathbf{a}}, \nabla\cdot\mathbf{v}_h)_{\omega_{\mathbf{a}}} = -(\psi_{\mathbf{a}} \nabla u_h, \mathbf{v}_h)_{\omega_{\mathbf{a}}} \qquad \forall \mathbf{v}_h \in \mathbf{V}_h^{\mathbf{a}},$$

$$(\nabla\cdot\boldsymbol{\sigma}_h^{\mathbf{a}}, q_h)_{\omega_{\mathbf{a}}} = (\psi_{\mathbf{a}} f - \nabla\psi_{\mathbf{a}}\cdot\nabla u_h, q_h)_{\omega_{\mathbf{a}}} \qquad \forall q_h \in Q_h^{\mathbf{a}}.$$

Local potential reconstruction

$V_h^{\mathbf{a}}$: FE space (hom. Dirichlet BC on $\partial\omega_{\mathbf{a}}$ for all $\mathbf{a} \in \mathcal{V}_h$)

Definition (Construction of s_h, ≈ Carstensen and Merdon (2013))

Let $u_h \in H^1(\mathcal{T}_h)$. For each $\mathbf{a} \in \mathcal{V}_h$, prescribe $s_h^{\mathbf{a}} \in V_h^{\mathbf{a}}$ by solving the **local conforming finite element problem**

$$s_h^{\mathbf{a}} := \arg \min_{v_h \in V_h^{\mathbf{a}}} \|\nabla(\psi_{\mathbf{a}} u_h - v_h)\|_{\omega_{\mathbf{a}}}$$

$$\Updownarrow$$

$$(\nabla s_h^{\mathbf{a}}, \nabla v_h)_{\omega_{\mathbf{a}}} = (\nabla(\psi_{\mathbf{a}} u_h), \nabla v_h)_{\omega_{\mathbf{a}}} \qquad \forall v_h \in V_h^{\mathbf{a}}.$$

Local potential reconstruction

$V_h^{\mathbf{a}}$: FE space (hom. Dirichlet BC on $\partial\omega_{\mathbf{a}}$ for all $\mathbf{a} \in \mathcal{V}_h$)

Definition (Construction of s_h, ≈ Carstensen and Merdon (2013))

Let $u_h \in H^1(\mathcal{T}_h)$. For each $\mathbf{a} \in \mathcal{V}_h$, prescribe $s_h^{\mathbf{a}} \in V_h^{\mathbf{a}}$ by solving the **local conforming finite element problem**

$$s_h^{\mathbf{a}} := \arg\min_{v_h \in V_h^{\mathbf{a}}} \|\nabla(\psi_{\mathbf{a}} u_h - v_h)\|_{\omega_{\mathbf{a}}}$$

$$\Updownarrow$$

$$(\nabla s_h^{\mathbf{a}}, \nabla v_h)_{\omega_{\mathbf{a}}} = (\nabla(\psi_{\mathbf{a}} u_h), \nabla v_h)_{\omega_{\mathbf{a}}} \qquad \forall v_h \in V_h^{\mathbf{a}}.$$

Local potential reconstruction

$V_h^{\mathbf{a}}$: FE space (hom. Dirichlet BC on $\partial\omega_{\mathbf{a}}$ for all $\mathbf{a} \in \mathcal{V}_h$)

Definition (Construction of s_h, ≈ Carstensen and Merdon (2013))

Let $u_h \in H^1(\mathcal{T}_h)$. For each $\mathbf{a} \in \mathcal{V}_h$, prescribe $s_h^{\mathbf{a}} \in V_h^{\mathbf{a}}$ by solving the **local conforming finite element problem**

$$s_h^{\mathbf{a}} := \arg \min_{v_h \in V_h^{\mathbf{a}}} \|\nabla(\psi_{\mathbf{a}} u_h - v_h)\|_{\omega_{\mathbf{a}}}$$

$$\Updownarrow$$

$$(\nabla s_h^{\mathbf{a}}, \nabla v_h)_{\omega_{\mathbf{a}}} = (\nabla(\psi_{\mathbf{a}} u_h), \nabla v_h)_{\omega_{\mathbf{a}}} \qquad \forall v_h \in V_h^{\mathbf{a}}.$$

Outline

Assumptions

Assumption B (Weak continuity)

There holds $\langle [\![u_h]\!], 1 \rangle_e = 0 \qquad \forall e \in \mathcal{E}_h.$

Assumption C (Piecewise polynomials, data, and meshes)

The approximation u_h and the datum f are piecewise polynomial. The **degrees** *of the MFE reconstructions σ_h and s_h are chosen correspondingly. The meshes $\mathcal{T}_h$ are shape-regular.*

Assumptions

Assumption B (Weak continuity)

There holds $$\langle [\![u_h]\!], 1 \rangle_e = 0 \qquad \forall e \in \mathcal{E}_h.$$

Assumption C (Piecewise polynomials, data, and meshes)

*The approximation u_h and the datum f are piecewise polynomial. The **degrees** of the MFE reconstructions $\boldsymbol{\sigma}_h$ and s_h are chosen correspondingly. The meshes $\mathcal{T}_h$ are shape-regular.*

Polynomial-degree-robust efficiency

Theorem (Polynomial-degree-robust efficiency)

Let u be the weak solution. Under Assumptions A, B, and C,

$$\|\nabla u_h + \boldsymbol{\sigma}_h\|_K \leq C_{\rm st} C_{\rm cont,PF} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\nabla(u-u_h)\|_{\omega_{\mathbf{a}}},$$

$$\|\nabla(u_h - s_h)\|_K \leq C_{\rm st} C_{\rm cont,bPF} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\nabla(u-u_h)\|_{\omega_{\mathbf{a}}}.$$

Remarks

- $C_{\rm cont,PF}$: $\approx 1 + 2/\pi$ on convex patches $\omega_{\mathbf{a}}$
- $C_{\rm st}$ can be bounded by solving the local Neumann problems by conforming FEs: find $r_h^{\mathbf{a}} \in V_h^{\mathbf{a}} \subset H^1_*(\omega_{\mathbf{a}})$ s.t.

$$(\nabla r_h^{\mathbf{a}}, \nabla v_h)_{\omega_{\mathbf{a}}} = -(\psi_{\mathbf{a}} \nabla u_h, \nabla v_h)_{\omega_{\mathbf{a}}} + (\psi_{\mathbf{a}} f - \nabla\psi_{\mathbf{a}}\cdot\nabla u_h, v_h)_{\omega_{\mathbf{a}}} \ \forall v_h \in V_h^{\mathbf{a}};$$

  then the first $C_{\rm st} \leq \|\psi_{\mathbf{a}}\nabla u_h + \boldsymbol{\sigma}_h^{\mathbf{a}}\|_{\omega_{\mathbf{a}}} / \|\nabla r_h^{\mathbf{a}}\|_{\omega_{\mathbf{a}}}$
- $\Rightarrow$ maximal overestimation factor guaranteed

Polynomial-degree-robust efficiency

Theorem (Polynomial-degree-robust efficiency)

Let u be the weak solution. Under Assumptions A, B, and C,

$$\|\nabla u_h + \boldsymbol{\sigma}_h\|_K \leq C_{\rm st} C_{\rm cont,PF} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\nabla(u-u_h)\|_{\omega_{\mathbf{a}}},$$

$$\|\nabla(u_h - s_h)\|_K \leq C_{\rm st} C_{\rm cont,bPF} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\nabla(u-u_h)\|_{\omega_{\mathbf{a}}}.$$

Remarks

- $C_{\rm cont,PF}$: $\approx 1 + 2/\pi$ on convex patches $\omega_{\mathbf{a}}$
- $C_{\rm st}$ can be bounded by solving the local Neumann problems by conforming FEs: find $r_h^{\mathbf{a}} \in V_h^{\mathbf{a}} \subset H^1_*(\omega_{\mathbf{a}})$ s.t.

$$(\nabla r_h^{\mathbf{a}}, \nabla v_h)_{\omega_{\mathbf{a}}} = -(\psi_{\mathbf{a}} \nabla u_h, \nabla v_h)_{\omega_{\mathbf{a}}} + (\psi_{\mathbf{a}} f - \nabla\psi_{\mathbf{a}} \cdot \nabla u_h, v_h)_{\omega_{\mathbf{a}}} \ \forall v_h \in V_h^{\mathbf{a}};$$

  then the first $C_{\rm st} \leq \|\psi_{\mathbf{a}} \nabla u_h + \boldsymbol{\sigma}_h^{\mathbf{a}}\|_{\omega_{\mathbf{a}}} / \|\nabla r_h^{\mathbf{a}}\|_{\omega_{\mathbf{a}}}$
- $\Rightarrow$ maximal overestimation factor guaranteed

Outline

Conforming finite elements

Conforming finite elements

Find $u_h \in V_h$ such that

$$(\nabla u_h, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_h.$$

- $V_h := \mathbb{P}_p(\mathcal{T}_h) \cap H_0^1(\Omega)$, $p \geq 1$
- Assumption A: take $v_h = \psi_{\mathbf{a}}$
- $V_h \subset H_0^1(\Omega)$: $s_h := u_h$, no need for Assumption B

Conforming finite elements

Conforming finite elements

Find $u_h \in V_h$ such that

$$(\nabla u_h, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_h.$$

- $V_h := \mathbb{P}_p(\mathcal{T}_h) \cap H_0^1(\Omega)$, $p \geq 1$
- Assumption A: take $v_h = \psi_{\mathbf{a}}$
- $V_h \subset H_0^1(\Omega)$: $s_h := u_h$, no need for Assumption B

Nonconforming finite elements

Nonconforming finite elements
Find $u_h \in V_h$ such that

$$(\nabla u_h, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_h.$$

- $V_h := \mathbb{P}_p(\mathcal{T}_h)$, $p \geq 1$, $v_h \in V_h$ satisfy

$$\langle [\![v_h]\!], q_h \rangle_e = 0 \qquad \forall q_h \in \mathbb{P}_{p-1}(e), \forall e \in \mathcal{E}_h$$

- Assumption A: take $v_h = \psi_{\mathbf{a}}$
- Assumption B: building requirement for the space V_h

Nonconforming finite elements

Nonconforming finite elements

Find $u_h \in V_h$ such that

$$(\nabla u_h, \nabla v_h) = (f, v_h) \qquad \forall v_h \in V_h.$$

- $V_h := \mathbb{P}_p(\mathcal{T}_h)$, $p \geq 1$, $v_h \in V_h$ satisfy

$$\langle [\![v_h]\!], q_h \rangle_e = 0 \qquad \forall q_h \in \mathbb{P}_{p-1}(e), \, \forall e \in \mathcal{E}_h$$

- Assumption A: take $v_h = \psi_{\mathbf{a}}$
- Assumption B: building requirement for the space V_h

Discontinuous Galerkin finite elements

Discontinuous Galerkin finite elements

Find $u_h \in V_h$ such that

$$\sum_{K\in\mathcal{T}_h}(\nabla u_h, \nabla v_h)_K - \sum_{e\in\mathcal{E}_h}\{\langle\{\!\{\nabla u_h\}\!\}\cdot\mathbf{n}_e, [\![v_h]\!]\rangle_e + \theta\langle\{\!\{\nabla v_h\}\!\}\cdot\mathbf{n}_e, [\![u_h]\!]\rangle_e\}$$

$$+ \sum_{e\in\mathcal{E}_h}\langle\alpha h_e^{-1}[\![u_h]\!], [\![v_h]\!]\rangle_e = (f, v_h) \qquad \forall v_h \in V_h.$$

- $V_h := \mathbb{P}_p(\mathcal{T}_h)$, $p \geq 1$
- Assumption A: take $v_h = \psi_{\mathbf{a}}$ for $\theta = 0$, otherwise estimates for the discrete gradient $\mathfrak{G}(u_h) := \nabla u_h - \theta \sum_{e\in\mathcal{E}_h} \mathfrak{l}_e([\![u_h]\!])$
- NIPG, IIPG: broken Poincaré–Friedrichs inequality & jumps
- symmetric version:

$$\|\mathfrak{G}(u_h) + \sigma_h\|_K \leq C_{\rm st} C_{\rm cont,PF} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\mathfrak{G}(u - u_h)\|_{\omega_{\mathbf{a}}},$$

$$\|\mathfrak{G}(u_h - s_h)\|_K \leq C_{\rm st} C_{\rm cont,P} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\mathfrak{G}(u - u_h)\|_{\omega_{\mathbf{a}}}$$

Discontinuous Galerkin finite elements

Discontinuous Galerkin finite elements

Find $u_h \in V_h$ such that

$$\sum_{K\in\mathcal{T}_h}(\nabla u_h, \nabla v_h)_K - \sum_{e\in\mathcal{E}_h}\{\langle\{\!\{\nabla u_h\}\!\}\cdot\mathbf{n}_e, [\![v_h]\!]\rangle_e + \theta\langle\{\!\{\nabla v_h\}\!\}\cdot\mathbf{n}_e, [\![u_h]\!]\rangle_e\}$$

$$+\sum_{e\in\mathcal{E}_h}\langle\alpha h_e^{-1}[\![u_h]\!], [\![v_h]\!]\rangle_e = (f, v_h) \qquad \forall v_h \in V_h.$$

- $V_h := \mathbb{P}_p(\mathcal{T}_h)$, $p \geq 1$
- Assumption A: take $v_h = \psi_{\mathbf{a}}$ for $\theta = 0$, otherwise estimates for the discrete gradient $\mathfrak{G}(u_h) := \nabla u_h - \theta\sum_{e\in\mathcal{E}_h} \mathfrak{l}_e([\![u_h]\!])$
- NIPG, IIPG: broken Poincaré–Friedrichs inequality & jumps
- symmetric version:

$$\|\mathfrak{G}(u_h) + \sigma_h\|_K \leq C_{\rm st} C_{\rm cont,PF} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\mathfrak{G}(u - u_h)\|_{\omega_{\mathbf{a}}},$$

$$\|\mathfrak{G}(u_h - s_h)\|_K \leq C_{\rm st} C_{\rm cont,P} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\mathfrak{G}(u - u_h)\|_{\omega_{\mathbf{a}}}$$

Discontinuous Galerkin finite elements

Discontinuous Galerkin finite elements

Find $u_h \in V_h$ such that

$$\sum_{K\in\mathcal{T}_h}(\nabla u_h, \nabla v_h)_K - \sum_{e\in\mathcal{E}_h}\{\langle\{\!\{\nabla u_h\}\!\}\cdot\mathbf{n}_e, [\![v_h]\!]\rangle_e + \theta\langle\{\!\{\nabla v_h\}\!\}\cdot\mathbf{n}_e, [\![u_h]\!]\rangle_e\}$$

$$+ \sum_{e\in\mathcal{E}_h}\langle\alpha h_e^{-1}[\![u_h]\!], [\![v_h]\!]\rangle_e = (f, v_h) \qquad \forall v_h \in V_h.$$

- $V_h := \mathbb{P}_p(\mathcal{T}_h)$, $p \geq 1$
- Assumption A: take $v_h = \psi_{\mathbf{a}}$ for $\theta = 0$, otherwise estimates for the discrete gradient $\mathfrak{G}(u_h) := \nabla u_h - \theta \sum_{e\in\mathcal{E}_h} \mathfrak{l}_e([\![u_h]\!])$
- NIPG, IIPG: broken Poincaré–Friedrichs inequality & jumps
- symmetric version:

$$\|\mathfrak{G}(u_h) + \sigma_h\|_K \leq C_{\rm st} C_{\rm cont,PF} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\mathfrak{G}(u - u_h)\|_{\omega_{\mathbf{a}}},$$

$$\|\mathfrak{G}(u_h - s_h)\|_K \leq C_{\rm st} C_{\rm cont,P} \sum_{\mathbf{a}\in\mathcal{V}_K} \|\mathfrak{G}(u - u_h)\|_{\omega_{\mathbf{a}}}$$

Discontinuous Galerkin finite elements

Discontinuous Galerkin finite elements

Find $u_h \in V_h$ such that

$$\sum_{K\in\mathcal{T}_h}(\nabla u_h, \nabla v_h)_K - \sum_{e\in\mathcal{E}_h}\{\langle\{\!\{\nabla u_h\}\!\}\cdot\mathbf{n}_e, [\![v_h]\!]\rangle_e + \theta\langle\{\!\{\nabla v_h\}\!\}\cdot\mathbf{n}_e, [\![u_h]\!]\rangle_e\}$$

$$+\sum_{e\in\mathcal{E}_h}\langle\alpha h_e^{-1}[\![u_h]\!], [\![v_h]\!]\rangle_e = (f, v_h) \qquad \forall v_h \in V_h.$$

- $V_h := \mathbb{P}_p(\mathcal{T}_h)$, $p \geq 1$
- Assumption A: take $v_h = \psi_{\mathbf{a}}$ for $\theta = 0$, otherwise estimates for the discrete gradient $\mathfrak{G}(u_h) := \nabla u_h - \theta\sum_{e\in\mathcal{E}_h}\mathfrak{l}_e([\![u_h]\!])$
- NIPG, IIPG: broken Poincaré–Friedrichs inequality & jumps
- symmetric version:

$$\|\mathfrak{G}(u_h) + \sigma_h\|_K \leq C_{\rm st}C_{\rm cont,PF}\sum_{\mathbf{a}\in\mathcal{V}_K}\|\mathfrak{G}(u-u_h)\|_{\omega_{\mathbf{a}}},$$

$$\|\mathfrak{G}(u_h - s_h)\|_K \leq C_{\rm st}C_{\rm cont,P}\sum_{\mathbf{a}\in\mathcal{V}_K}\|\mathfrak{G}(u-u_h)\|_{\omega_{\mathbf{a}}}$$

Discontinuous Galerkin finite elements

Discontinuous Galerkin finite elements

Find $u_h \in V_h$ such that

$$\sum_{K\in\mathcal{T}_h}(\nabla u_h,\nabla v_h)_K - \sum_{e\in\mathcal{E}_h}\{\langle\{\!\{\nabla u_h\}\!\}\cdot\mathbf{n}_e,[\![v_h]\!]\rangle_e+\theta\langle\{\!\{\nabla v_h\}\!\}\cdot\mathbf{n}_e,[\![u_h]\!]\rangle_e\}$$

$$+\sum_{e\in\mathcal{E}_h}\langle\alpha h_e^{-1}[\![u_h]\!],[\![v_h]\!]\rangle_e=(f,v_h)\qquad\forall v_h\in V_h.$$

- $V_h := \mathbb{P}_p(\mathcal{T}_h)$, $p \geq 1$
- Assumption A: take $v_h = \psi_{\mathbf{a}}$ for $\theta = 0$, otherwise estimates for the discrete gradient $\mathfrak{G}(u_h) := \nabla u_h - \theta\sum_{e\in\mathcal{E}_h}\mathfrak{l}_e([\![u_h]\!])$
- NIPG, IIPG: broken Poincaré–Friedrichs inequality & jumps
- symmetric version:

$$\|\mathfrak{G}(u_h)+\boldsymbol{\sigma}_h\|_K \leq C_{\rm st}C_{\rm cont,PF}\sum_{\mathbf{a}\in\mathcal{V}_K}\|\mathfrak{G}(u-u_h)\|_{\omega_{\mathbf{a}}},$$

$$\|\mathfrak{G}(u_h-s_h)\|_K \leq C_{\rm st}C_{\rm cont,P}\sum_{\mathbf{a}\in\mathcal{V}_K}\|\mathfrak{G}(u-u_h)\|_{\omega_{\mathbf{a}}}$$

Outline

Numerics: smooth test case

Model problem

$$-\Delta u = f \quad \text{in } \Omega :=]0,1[^2$$
$$u = u_{\mathrm{D}} \quad \text{on } \partial\Omega$$

Exact solution

$$u(\mathbf{x}) = (c_1 + c_2(1 - x_1) + e^{-\alpha x_1})(c_1 + c_2(1 - x_2) + e^{-\alpha x_2})$$
$$c_1 = -e^{-\alpha}, \quad c_2 = -1 - c_1, \quad \alpha = 10$$

Discretization

- incomplete interior penalty discontinuous Galerkin method
- unstructured nested triangular grids
- uniform refinement

Numerics: smooth test case

Model problem

$$-\Delta u = f \quad \text{in } \Omega :=]0,1[^2$$
$$u = u_{\mathrm{D}} \quad \text{on } \partial\Omega$$

Exact solution

$$u(\mathbf{x}) = (c_1 + c_2(1 - x_1) + e^{-\alpha x_1})(c_1 + c_2(1 - x_2) + e^{-\alpha x_2})$$
$$c_1 = -e^{-\alpha}, \quad c_2 = -1 - c_1, \quad \alpha = 10$$

Discretization

- incomplete interior penalty discontinuous Galerkin method
- unstructured nested triangular grids
- uniform refinement

Numerics: smooth test case

Model problem

$$-\Delta u = f \quad \text{in } \Omega :=]0,1[^2$$
$$u = u_{\mathrm{D}} \quad \text{on } \partial\Omega$$

Exact solution

$$u(\mathbf{x}) = (c_1 + c_2(1-x_1) + e^{-\alpha x_1})(c_1 + c_2(1-x_2) + e^{-\alpha x_2})$$
$$c_1 = -e^{-\alpha}, \quad c_2 = -1 - c_1, \quad \alpha = 10$$

Discretization

- incomplete interior penalty discontinuous Galerkin method
- unstructured nested triangular grids
- uniform refinement

Estimates, errors, and effectivity indices

h	p	$\|\nabla(u-u_h)\|$	$\|u-u_h\|_{DG}$	$\|\nabla u_h+\boldsymbol{\sigma}_h\|$	$\|\nabla(u_h-s_h)\|$	η_{osc}	η	η_{DG}	I^{eff}	I^{eff}_{DG}
$h_0/1$	1	1.21E+00	1.22E+00	1.24E+00	1.07E-01	5.56E-02	1.30E+00	1.31E+00	1.07	1.07
$h_0/2$		6.18E-01	6.22E-01	6.38E-01	5.09E-02	7.02E-03	6.47E-01	6.50E-01	1.05	1.05
		(0.97)	(0.97)	(0.96)	(1.07)	(2.99)	(1.01)	(1.01)		
$h_0/4$		3.12E-01	3.13E-01	3.22E-01	2.43E-02	8.80E-04	3.24E-01	3.25E-01	1.04	1.04
		(0.99)	(0.99)	(0.99)	(1.07)	(3.00)	(1.00)	(1.00)		
$h_0/8$		1.56E-01	1.57E-01	1.61E-01	1.18E-02	1.10E-04	1.62E-01	1.63E-01	1.04	1.04
		(1.00)	(1.00)	(1.00)	(1.05)	(3.00)	(1.00)	(1.00)		
$h_0/1$	2	1.50E-01	1.53E-01	1.49E-01	2.76E-02	5.10E-03	1.56E-01	1.59E-01	1.04	1.04
$h_0/2$		3.85E-02	3.92E-02	3.83E-02	7.99E-03	3.22E-04	3.94E-02	4.01E-02	1.03	1.02
		(1.96)	(1.96)	(1.96)	(1.79)	(3.98)	(1.98)	(1.98)		
$h_0/4$		9.70E-03	9.88E-03	9.68E-03	2.12E-03	2.02E-05	9.93E-03	1.01E-02	1.02	1.02
		(1.99)	(1.99)	(1.98)	(1.92)	(4.00)	(1.99)	(1.99)		
$h_0/8$		2.43E-03	2.48E-03	2.43E-03	5.42E-04	1.26E-06	2.49E-03	2.54E-03	1.02	1.02
		(1.99)	(1.99)	(1.99)	(1.96)	(4.00)	(1.99)	(1.99)		
$h_0/1$	3	1.32E-02	1.34E-02	1.29E-02	2.52E-03	3.58E-04	1.35E-02	1.37E-02	1.03	1.03
$h_0/2$		1.67E-03	1.69E-03	1.65E-03	3.13E-04	1.13E-05	1.70E-03	1.71E-03	1.01	1.01
		(2.98)	(2.98)	(2.97)	(3.01)	(4.99)	(3.00)	(3.00)		
$h_0/4$		2.11E-04	2.13E-04	2.09E-04	3.83E-05	3.53E-07	2.12E-04	2.15E-04	1.01	1.01
		(2.99)	(2.99)	(2.99)	(3.03)	(5.00)	(3.00)	(3.00)		
$h_0/8$		2.64E-05	2.67E-05	2.61E-05	4.69E-06	1.10E-08	2.66E-05	2.69E-05	1.01	1.01
		(3.00)	(3.00)	(3.00)	(3.03)	(5.00)	(3.00)	(3.00)		
$h_0/1$	4	9.36E-04	9.54E-04	9.05E-04	2.41E-04	2.12E-05	9.57E-04	9.74E-04	1.02	1.02
$h_0/2$		5.93E-05	6.05E-05	5.77E-05	1.68E-05	3.36E-07	6.04E-05	6.16E-05	1.02	1.02
		(3.98)	(3.98)	(3.97)	(3.84)	(5.98)	(3.99)	(3.98)		
$h_0/4$		3.72E-06	3.80E-06	3.63E-06	1.10E-06	5.31E-09	3.80E-06	3.87E-06	1.02	1.02
		(3.99)	(3.99)	(3.99)	(3.94)	(5.98)	(3.99)	(3.99)		
$h_0/8$		2.33E-07	2.38E-07	2.27E-07	7.02E-08	8.30E-11	2.38E-07	2.43E-07	1.02	1.02
		(4.00)	(4.00)	(4.00)	(3.97)	(6.00)	(4.00)	(3.99)		
$h_0/1$	5	5.41E-05	5.50E-05	5.22E-05	1.38E-05	1.06E-06	5.50E-05	5.58E-05	1.02	1.02
$h_0/2$		1.70E-06	1.72E-06	1.65E-06	4.39E-07	9.35E-09	1.72E-06	1.74E-06	1.01	1.01
		(4.99)	(5.00)	(4.98)	(4.98)	(6.82)	(5.00)	(5.00)		
$h_0/4$		5.32E-08	5.39E-08	5.19E-08	1.40E-08	7.67E-11	5.38E-08	5.45E-08	1.01	1.01
		(5.00)	(5.00)	(4.99)	(4.97)	(6.93)	(5.00)	(5.00)		
$h_0/8$		1.66E-09	1.69E-09	1.62E-09	4.41E-10	5.99E-13	1.68E-09	1.70E-09	1.01	1.01
		(5.00)	(5.00)	(5.00)	(4.99)	(7.00)	(5.00)	(5.00)		

Numerics: singular test case & *hp*-adaptivity

Model problem

$$\begin{aligned} -\Delta u &= 0 \quad \text{in } \Omega := \Omega :=]-1,1[^2 \backslash [0,1]^2, \\ u &= u_{\mathrm{D}} \quad \text{on } \partial\Omega \end{aligned}$$

Exact solution

$$u(r,\phi) = r^{2/3} \sin(2\phi/3)$$

Discretization

- incomplete interior penalty discontinuous Galerkin method
- unstructured non-nested triangular grids
- *hp*-adaptive refinement

Numerics: singular test case & *hp*-adaptivity

Model problem

$$\begin{aligned} -\Delta u &= 0 \quad \text{in } \Omega := \Omega :=]-1,1[^2 \backslash [0,1]^2, \\ u &= u_{\mathrm{D}} \quad \text{on } \partial\Omega \end{aligned}$$

Exact solution

$$u(r,\phi) = r^{2/3} \sin(2\phi/3)$$

Discretization

- incomplete interior penalty discontinuous Galerkin method
- unstructured non-nested triangular grids
- *hp*-adaptive refinement

Numerics: singular test case & *hp*-adaptivity

Model problem

$$\begin{aligned} -\Delta u &= 0 \quad \text{in } \Omega := \Omega :=]-1,1[^2 \backslash [0,1]^2, \\ u &= u_{\mathrm{D}} \quad \text{on } \partial\Omega \end{aligned}$$

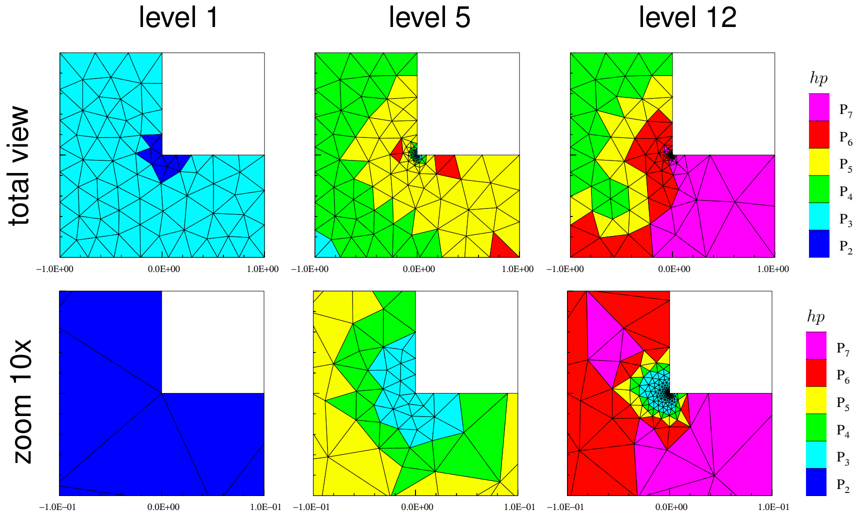
Exact solution

$$u(r,\phi) = r^{2/3} \sin(2\phi/3)$$

Discretization

- incomplete interior penalty discontinuous Galerkin method
- unstructured non-nested triangular grids
- *hp*-adaptive refinement

hp-refinement grids



Estimates, errors, and effectivity indices

lev	$\|\mathcal{T}_h\|$	DoF	$\|\nabla(u-u_h)\|$	$\|\nabla u_h + \boldsymbol{\sigma}_h\|$	$\eta_{\rm osc}$	$\|\nabla(u_h - s_h)\|$	$\eta_{\rm BC}$	η	$I^{\rm eff}$
0	114	684	6.22E-02	6.63E-02	1.89E-15	4.48E-02	3.81E-02	1.05E-01	1.69
1	122	1180	4.28E-02	4.27E-02	1.18E-14	3.08E-02	2.92E-02	7.29E-02	1.70
2	139	1919	3.28E-02	3.37E-02	8.21E-14	2.09E-02	2.12E-02	5.36E-02	1.64
3	165	2573	2.32E-02	2.30E-02	3.88E-13	1.50E-02	1.03E-02	3.41E-02	1.47
4	174	2858	1.02E-02	1.01E-02	4.48E-13	8.22E-03	9.19E-03	1.99E-02	1.96
5	199	3351	6.27E-03	6.21E-03	1.12E-12	4.81E-03	6.18E-03	1.25E-02	2.00
6	237	3926	4.21E-03	4.23E-03	1.98E-12	3.15E-03	3.29E-03	7.66E-03	1.82
7	285	4537	2.84E-03	2.91E-03	7.47E-12	2.13E-03	2.42E-03	5.33E-03	1.88
8	338	5257	2.04E-03	2.19E-03	4.63E-11	1.45E-03	1.32E-03	3.51E-03	1.72
9	372	5658	1.21E-03	1.23E-03	1.11E-11	9.07E-04	9.99E-04	2.26E-03	1.87
10	426	6500	7.70E-04	7.69E-04	5.69E-11	5.55E-04	6.95E-04	1.46E-03	1.89
11	453	7010	4.95E-04	5.04E-04	9.77E-11	3.97E-04	4.74E-04	9.91E-04	2.00
12	469	7308	3.41E-04	3.47E-04	1.13E-10	2.55E-04	2.88E-04	6.40E-04	1.88
13	463	7286	2.42E-04	2.42E-04	1.39E-10	1.73E-04	1.94E-04	4.37E-04	1.81
14	458	7215	1.69E-04	1.69E-04	1.23E-10	1.19E-04	1.53E-04	3.17E-04	1.88
15	440	6955	1.29E-04	1.31E-04	1.45E-10	9.21E-05	9.10E-05	2.24E-04	1.73
16	435	7035	9.71E-05	9.91E-05	1.39E-10	6.89E-05	7.63E-05	1.74E-04	1.79
17	434	7167	8.52E-05	8.97E-05	1.41E-10	5.76E-05	5.47E-05	1.42E-04	1.67
18	419	6960	7.51E-05	7.97E-05	1.44E-10	5.00E-05	4.15E-05	1.21E-04	1.60
19	410	6838	6.06E-05	6.35E-05	1.47E-10	3.87E-05	3.65E-05	9.69E-05	1.60

Outline

Inexact iterative linearization

System of nonlinear algebraic equations

Nonlinear operator $\mathcal{A} : \mathbb{R}^N \to \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.

$$\mathcal{A}(U) = F$$

Algorithm (Inexact iterative linearization)

1. *Choose initial vector* U^0. *Set* $k := 1$.
2. $U^{k-1} \Rightarrow$ *matrix* $\mathbb{A}^{k-1}$ *and vector* F^{k-1}: *find* U^k *s.t.*

$$\mathbb{A}^{k-1} U^k \approx F^{k-1}.$$

3.
   1. *Set* $U^{k,0} := U^{k-1}$ *and* $i := 1$.
   2. *Do an algebraic solver step* $\Rightarrow U^{k,i}$ *s.t.* ($R^{k,i}$ *algebraic res.*)

$$\mathbb{A}^{k-1} U^{k,i} = F^{k-1} - R^{k,i}.$$

   3. *Convergence? OK* $\Rightarrow U^k := U^{k,i}$. *KO* $\Rightarrow i := i + 1$, *back to 3.2.*
4. *Convergence? OK* $\Rightarrow$ *finish. KO* $\Rightarrow k := k + 1$, *back to 2.*

Inexact iterative linearization

System of nonlinear algebraic equations

Nonlinear operator $\mathcal{A} : \mathbb{R}^N \to \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.

$$\mathcal{A}(U) = F$$

Algorithm (Inexact iterative linearization)

1. *Choose initial vector U^0. Set $k := 1$.*
2. *$U^{k-1} \Rightarrow$ matrix $\mathbb{A}^{k-1}$ and vector F^{k-1}: find U^k s.t.*
$$\mathbb{A}^{k-1} U^k \approx F^{k-1}.$$
3.
   1. *Set $U^{k,0} := U^{k-1}$ and $i := 1$.*
   2. *Do an algebraic solver step $\Rightarrow U^{k,i}$ s.t. ($R^{k,i}$ algebraic res.)*
   $$\mathbb{A}^{k-1} U^{k,i} = F^{k-1} - R^{k,i}.$$
   3. *Convergence? OK $\Rightarrow U^k := U^{k,i}$. KO $\Rightarrow i := i + 1$, back to 3.2.*
4. *Convergence? OK $\Rightarrow$ finish. KO $\Rightarrow k := k + 1$, back to 2.*

Inexact iterative linearization

System of nonlinear algebraic equations

Nonlinear operator $\mathcal{A}:\mathbb{R}^N \to \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.

$$\mathcal{A}(U) = F$$

Algorithm (Inexact iterative linearization)

1. *Choose initial vector U^0. Set $k := 1$.*
2. *$U^{k-1} \Rightarrow$ matrix $\mathbb{A}^{k-1}$ and vector F^{k-1}: find U^k s.t.*

$$\mathbb{A}^{k-1} U^k \approx F^{k-1}.$$

3.
   1. *Set $U^{k,0} := U^{k-1}$ and $i := 1$.*
   2. *Do an algebraic solver step $\Rightarrow U^{k,i}$ s.t. ($R^{k,i}$ algebraic res.)*

      $$\mathbb{A}^{k-1} U^{k,i} = F^{k-1} - R^{k,i}.$$

   3. *Convergence? OK $\Rightarrow U^k := U^{k,i}$. KO $\Rightarrow i := i+1$, back to 3.2.*
4. *Convergence? OK $\Rightarrow$ finish. KO $\Rightarrow k := k+1$, back to 2.*

Inexact iterative linearization

System of nonlinear algebraic equations

Nonlinear operator $\mathcal{A} : \mathbb{R}^N \to \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.

$$\mathcal{A}(U) = F$$

Algorithm (Inexact iterative linearization)

1. *Choose initial vector* U^0. *Set* $k := 1$.
2. $U^{k-1} \Rightarrow$ *matrix* $\mathbb{A}^{k-1}$ *and vector* F^{k-1}: *find* U^k *s.t.*
$$\mathbb{A}^{k-1} U^k \approx F^{k-1}.$$
3.
   1. *Set* $U^{k,0} := U^{k-1}$ *and* $i := 1$.
   2. *Do an algebraic solver step* $\Rightarrow U^{k,i}$ *s.t.* ($R^{k,i}$ *algebraic res.*)
   $$\mathbb{A}^{k-1} U^{k,i} = F^{k-1} - R^{k,i}.$$
   3. *Convergence? OK* $\Rightarrow U^k := U^{k,i}$. *KO* $\Rightarrow i := i + 1$, *back to 3.2.*
4. *Convergence? OK* $\Rightarrow$ *finish. KO* $\Rightarrow k := k + 1$, *back to 2.*

Inexact iterative linearization

System of nonlinear algebraic equations

Nonlinear operator $\mathcal{A} : \mathbb{R}^N \to \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.

$$\mathcal{A}(U) = F$$

Algorithm (Inexact iterative linearization)

1. *Choose initial vector U^0. Set $k := 1$.*
2. *$U^{k-1} \Rightarrow$ matrix $\mathbb{A}^{k-1}$ and vector F^{k-1}: find U^k s.t.*
   $$\mathbb{A}^{k-1} U^k \approx F^{k-1}.$$
3.
   1. *Set $U^{k,0} := U^{k-1}$ and $i := 1$.*
   2. *Do an algebraic solver step $\Rightarrow U^{k,i}$ s.t. ($R^{k,i}$ algebraic res.)*
      $$\mathbb{A}^{k-1} U^{k,i} = F^{k-1} - R^{k,i}.$$
   3. *Convergence? OK $\Rightarrow U^k := U^{k,i}$. KO $\Rightarrow i := i + 1$, back to 3.2.*
4. *Convergence? OK $\Rightarrow$ finish. KO $\Rightarrow k := k + 1$, back to 2.*

Inexact iterative linearization

System of nonlinear algebraic equations

Nonlinear operator $\mathcal{A}:\mathbb{R}^N \to \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.

$$\mathcal{A}(U) = F$$

Algorithm (Inexact iterative linearization)

1. *Choose initial vector U^0. Set $k := 1$.*
2. *$U^{k-1} \Rightarrow$ matrix $\mathbb{A}^{k-1}$ and vector F^{k-1}: find U^k s.t.*

$$\mathbb{A}^{k-1}U^k \approx F^{k-1}.$$

3.
   1. *Set $U^{k,0} := U^{k-1}$ and $i := 1$.*
   2. *Do an algebraic solver step $\Rightarrow$ $U^{k,i}$ s.t. ($R^{k,i}$ algebraic res.)*

   $$\mathbb{A}^{k-1}U^{k,i} = F^{k-1} - R^{k,i}.$$

   3. *Convergence? OK $\Rightarrow$ $U^k := U^{k,i}$. KO $\Rightarrow$ $i := i+1$, back to 3.2.*
4. *Convergence? OK $\Rightarrow$ finish. KO $\Rightarrow$ $k := k+1$, back to 2.*

Inexact iterative linearization

System of nonlinear algebraic equations

Nonlinear operator $\mathcal{A}:\mathbb{R}^N \to \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.

$$\mathcal{A}(U) = F$$

Algorithm (Inexact iterative linearization)

1. *Choose initial vector* U^0. *Set* $k := 1$.
2. $U^{k-1} \Rightarrow$ *matrix* $\mathbb{A}^{k-1}$ *and vector* F^{k-1}: *find* U^k *s.t.*

$$\mathbb{A}^{k-1} U^k \approx F^{k-1}.$$

3.
   1. *Set* $U^{k,0} := U^{k-1}$ *and* $i := 1$.
   2. *Do an algebraic solver step* $\Rightarrow U^{k,i}$ *s.t.* ($R^{k,i}$ *algebraic res.)*

   $$\mathbb{A}^{k-1} U^{k,i} = F^{k-1} - R^{k,i}.$$

   3. *Convergence? OK* $\Rightarrow U^k := U^{k,i}$. *KO* $\Rightarrow i := i+1$, *back to 3.2.*
4. *Convergence? OK* $\Rightarrow$ *finish. KO* $\Rightarrow k := k+1$, *back to 2.*

Context and questions

Approximate solution

- approximate solution $U^{k,i}$ does not solve $\mathcal{A}(U^{k,i}) = F$

Numerical method

- underlying numerical method: the vector $U^{k,i}$ is associated with a (piecewise polynomial) approximation $u_h^{k,i}$

Partial differential equation

- underlying PDE, u its weak solution: $A(u) = f$

Question (Stopping criteria)

- *What is a good stopping criterion for the linear solver?*
- *What is a good stopping criterion for the nonlinear solver?*

Question (Error)

- *How big is the error $\|u - u_h^{k,i}\|$ on Newton step k and algebraic solver step i, how is it distributed?*

Context and questions

Approximate solution

- approximate solution $U^{k,i}$ does not solve $\mathcal{A}(U^{k,i}) = F$

Numerical method

- underlying numerical method: the vector $U^{k,i}$ is associated with a (piecewise polynomial) approximation $u_h^{k,i}$

Partial differential equation

- underlying PDE, u its weak solution: $A(u) = f$

Question (Stopping criteria)

- *What is a good stopping criterion for the linear solver?*
- *What is a good stopping criterion for the nonlinear solver?*

Question (Error)

- *How big is the error $\|u - u_h^{k,i}\|$ on Newton step k and algebraic solver step i, how is it distributed?*

Context and questions

Approximate solution

- approximate solution $U^{k,i}$ does not solve $\mathcal{A}(U^{k,i}) = F$

Numerical method

- underlying numerical method: the vector $U^{k,i}$ is associated with a (piecewise polynomial) approximation $u_h^{k,i}$

Partial differential equation

- underlying PDE, u its weak solution: $A(u) = f$

Question (Stopping criteria)

- *What is a good stopping criterion for the linear solver?*
- *What is a good stopping criterion for the nonlinear solver?*

Question (Error)

- *How big is the error $\|u - u_h^{k,i}\|$ on Newton step k and algebraic solver step i, how is it distributed?*

ematics

Context and questions

Approximate solution

- approximate solution $U^{k,i}$ does not solve $\mathcal{A}(U^{k,i}) = F$

Numerical method

- underlying numerical method: the vector $U^{k,i}$ is associated with a (piecewise polynomial) approximation $u_h^{k,i}$

Partial differential equation

- underlying PDE, u its weak solution: $A(u) = f$

Question (Stopping criteria)

- *What is a good stopping criterion for the linear solver?*
- *What is a good stopping criterion for the nonlinear solver?*

Question (Error)

- *How big is the error $\|u - u_h^{k,i}\|$ on Newton step k and algebraic solver step i, how is it distributed?*

Context and questions

Approximate solution

- approximate solution $U^{k,i}$ does not solve $\mathcal{A}(U^{k,i}) = F$

Numerical method

- underlying numerical method: the vector $U^{k,i}$ is associated with a (piecewise polynomial) approximation $u_h^{k,i}$

Partial differential equation

- underlying PDE, u its weak solution: $A(u) = f$

Question (Stopping criteria)

- *What is a good stopping criterion for the linear solver?*
- *What is a good stopping criterion for the nonlinear solver?*

Question (Error)

- *How big is the error $\|u - u_h^{k,i}\|$ on Newton step k and algebraic solver step i, how is it distributed?*

Model steady problem, discretization

Quasi-linear elliptic problem

$$-\nabla\cdot\overline{\boldsymbol{\sigma}}(u,\nabla u) = f \qquad \text{in } \Omega,$$
$$u = 0 \qquad \text{on } \partial\Omega$$

- $p > 1$, $q := \frac{p}{p-1}$, $f \in L^q(\Omega)$
- example: p-Laplacian with $\overline{\boldsymbol{\sigma}}(u,\nabla u) = |\nabla u|^{p-2}\nabla u$
- f piecewise polynomial for simplicity
- weak solution: $u \in V := W_0^{1,p}(\Omega)$ such that

$$(\overline{\boldsymbol{\sigma}}(u,\nabla u),\nabla v) = (f,v) \qquad \forall v \in V$$

Numerical approximation

- simplicial mesh $\mathcal{T}_h$, linearization step k, algebraic step i
- $u_h^{k,i} \in V(\mathcal{T}_h) := \{v \in L^p(\Omega),\, v|_K \in W^{1,p}(K) \quad \forall K \in \mathcal{T}_h\} \not\subset V$

Model steady problem, discretization

Quasi-linear elliptic problem

$$\begin{aligned} -\nabla\cdot\overline{\boldsymbol{\sigma}}(u,\nabla u) &= f \qquad \text{in } \Omega,\\ u &= 0 \qquad \text{on } \partial\Omega \end{aligned}$$

- $p > 1$, $q := \frac{p}{p-1}$, $f \in L^q(\Omega)$
- example: p-Laplacian with $\overline{\boldsymbol{\sigma}}(u,\nabla u) = |\nabla u|^{p-2}\nabla u$
- f piecewise polynomial for simplicity
- weak solution: $u \in V := W_0^{1,p}(\Omega)$ such that

$$(\overline{\boldsymbol{\sigma}}(u,\nabla u),\nabla v) = (f,v) \qquad \forall v \in V$$

Numerical approximation

- simplicial mesh $\mathcal{T}_h$, linearization step k, algebraic step i
- $u_h^{k,i} \in V(\mathcal{T}_h) := \{v \in L^p(\Omega),\ v|_K \in W^{1,p}(K) \quad \forall K \in \mathcal{T}_h\} \not\subset V$

Outline

Abstract assumptions

Assumption A (Total flux reconstruction)

There exists $\boldsymbol{\sigma}_h^{k,i} \in \mathbf{H}^q(\mathrm{div}, \Omega)$ *and* $\rho_h^{k,i} \in L^q(\Omega)$ *such that*

$$\nabla \cdot \boldsymbol{\sigma}_h^{k,i} = f - \underbrace{\rho_h^{k,i}}_{\text{algebraic remainder}}.$$

Assumption B (Discretization, linearization, and alg. fluxes)

There exist fluxes $\mathbf{d}_h^{k,i}, \mathbf{l}_h^{k,i}, \mathbf{a}_h^{k,i} \in [L^q(\Omega)]^d$ *such that*

(i) $\boldsymbol{\sigma}_h^{k,i} = \mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i} + \mathbf{a}_h^{k,i}$;

(ii) *as the linear solver converges,* $\|\mathbf{a}_h^{k,i}\|_q \to 0$;

(iii) *as the nonlinear solver converges,* $\|\mathbf{l}_h^{k,i}\|_q \to 0$.

Abstract assumptions

Assumption A (Total flux reconstruction)

There exists $\boldsymbol{\sigma}_h^{k,i} \in \mathbf{H}^q(\mathrm{div}, \Omega)$ *and* $\rho_h^{k,i} \in L^q(\Omega)$ *such that*

$$\nabla \cdot \boldsymbol{\sigma}_h^{k,i} = f - \underbrace{\rho_h^{k,i}}_{\text{algebraic remainder}} .$$

Assumption B (Discretization, linearization, and alg. fluxes)

There exist fluxes $\mathbf{d}_h^{k,i}, \mathbf{l}_h^{k,i}, \mathbf{a}_h^{k,i} \in [L^q(\Omega)]^d$ *such that*

(i) $\boldsymbol{\sigma}_h^{k,i} = \mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i} + \mathbf{a}_h^{k,i}$;

(ii) *as the linear solver converges,* $\|\mathbf{a}_h^{k,i}\|_q \to 0$;

(iii) *as the nonlinear solver converges,* $\|\mathbf{l}_h^{k,i}\|_q \to 0$.

Abstract assumptions

Assumption A (Total flux reconstruction)

There exists $\boldsymbol{\sigma}_h^{k,i} \in \mathbf{H}^q(\mathrm{div}, \Omega)$ *and* $\rho_h^{k,i} \in L^q(\Omega)$ *such that*

$$\nabla \cdot \boldsymbol{\sigma}_h^{k,i} = f - \underbrace{\rho_h^{k,i}}_{\text{algebraic remainder}}.$$

Assumption B (Discretization, linearization, and alg. fluxes)

There exist fluxes $\mathbf{d}_h^{k,i}, \mathbf{l}_h^{k,i}, \mathbf{a}_h^{k,i} \in [L^q(\Omega)]^d$ *such that*

(i) $\boldsymbol{\sigma}_h^{k,i} = \mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i} + \mathbf{a}_h^{k,i}$;

(ii) *as the linear solver converges,* $\|\mathbf{a}_h^{k,i}\|_q \to 0$;

(iii) *as the nonlinear solver converges,* $\|\mathbf{l}_h^{k,i}\|_q \to 0$.

Estimate distinguishing error components

Theorem (Estimate distinguishing different error components)

Let

- *$u \in V$ be the weak solution,*
- *$u_h^{k,i} \in V(\mathcal{T}_h)$ be arbitrary,*
- *Assumptions A and B hold.*

Then there holds

$$\underbrace{\mathcal{J}_u(u_h^{k,i})}_{\text{dual norm of the residual + NC}} \leq \eta_{\text{disc}}^{k,i} + \underbrace{\eta_{\text{lin}}^{k,i}}_{\|\mathbf{l}_h^{k,i}\|_q} + \underbrace{\eta_{\text{alg}}^{k,i}}_{\|\mathbf{a}_h^{k,i}\|_q} + \underbrace{\eta_{\text{rem}}^{k,i}}_{h_\Omega \|\rho_h^{k,i}\|_{q,K}} + \eta_{\text{quad}}^{k,i} + \eta_{\text{osc}}.$$

Outline

Stopping criteria and efficiency

Global stopping criteria $\gamma_{\rm rem}, \gamma_{\rm alg}, \gamma_{\rm lin} \approx 0.1$

$$\eta_{\rm rem}^{k,i} \leq \gamma_{\rm rem} \max\{\eta_{\rm disc}^{k,i}, \eta_{\rm lin}^{k,i}, \eta_{\rm alg}^{k,i}\},$$
$$\eta_{\rm alg}^{k,i} \leq \gamma_{\rm alg} \max\{\eta_{\rm disc}^{k,i}, \eta_{\rm lin}^{k,i}\},$$
$$\eta_{\rm lin}^{k,i} \leq \gamma_{\rm lin} \eta_{\rm disc}^{k,i}$$

Theorem (Global efficiency)

Under the global stopping criteria and the usual assumptions,

$$\eta_{\rm disc}^{k,i} + \eta_{\rm lin}^{k,i} + \eta_{\rm alg}^{k,i} + \eta_{\rm rem}^{k,i} \leq C(\mathcal{J}_u(u_h^{k,i}) + \eta_{\rm quad}^{k,i} + \eta_{\rm osc}),$$

where C is independent of $\overline{\sigma}$ and q.

- local (elementwise) stopping criteria ⇒ **local** efficiency
- **robustness** with respect to the **nonlinearity** thanks to the choice of $\mathcal{J}_u$ as error measure

Stopping criteria and efficiency

Global stopping criteria $\gamma_{\rm rem}, \gamma_{\rm alg}, \gamma_{\rm lin} \approx 0.1$

$$\begin{aligned} \eta_{\rm rem}^{k,i} &\le \gamma_{\rm rem} \max\{\eta_{\rm disc}^{k,i}, \eta_{\rm lin}^{k,i}, \eta_{\rm alg}^{k,i}\}, \\ \eta_{\rm alg}^{k,i} &\le \gamma_{\rm alg} \max\{\eta_{\rm disc}^{k,i}, \eta_{\rm lin}^{k,i}\}, \\ \eta_{\rm lin}^{k,i} &\le \gamma_{\rm lin} \eta_{\rm disc}^{k,i} \end{aligned}$$

Theorem (Global efficiency)

Under the global stopping criteria and the usual assumptions,

$$\eta_{\rm disc}^{k,i} + \eta_{\rm lin}^{k,i} + \eta_{\rm alg}^{k,i} + \eta_{\rm rem}^{k,i} \le C(\mathcal{J}_u(u_h^{k,i}) + \eta_{\rm quad}^{k,i} + \eta_{\rm osc}),$$

where C is independent of $\overline{\sigma}$ and q.

- local (elementwise) stopping criteria $\Rightarrow$ **local** efficiency
- **robustness** with respect to the **nonlinearity** thanks to the choice of $\mathcal{J}_u$ as error measure

Stopping criteria and efficiency

Global stopping criteria $\gamma_{\text{rem}}, \gamma_{\text{alg}}, \gamma_{\text{lin}} \approx 0.1$

$$\begin{aligned} \eta_{\text{rem}}^{k,i} &\le \gamma_{\text{rem}} \max\{\eta_{\text{disc}}^{k,i}, \eta_{\text{lin}}^{k,i}, \eta_{\text{alg}}^{k,i}\}, \\ \eta_{\text{alg}}^{k,i} &\le \gamma_{\text{alg}} \max\{\eta_{\text{disc}}^{k,i}, \eta_{\text{lin}}^{k,i}\}, \\ \eta_{\text{lin}}^{k,i} &\le \gamma_{\text{lin}} \eta_{\text{disc}}^{k,i} \end{aligned}$$

Theorem (Global efficiency)

Under the global stopping criteria and the usual assumptions,

$$\eta_{\text{disc}}^{k,i} + \eta_{\text{lin}}^{k,i} + \eta_{\text{alg}}^{k,i} + \eta_{\text{rem}}^{k,i} \le C(\mathcal{J}_u(u_h^{k,i}) + \eta_{\text{quad}}^{k,i} + \eta_{\text{osc}}),$$

where C is independent of $\overline{\sigma}$ and q.

- local (elementwise) stopping criteria ⇒ **local** efficiency
- **robustness** with respect to the **nonlinearity** thanks to the choice of $\mathcal{J}_u$ as error measure

Stopping criteria and efficiency

Global stopping criteria $\gamma_{\rm rem}, \gamma_{\rm alg}, \gamma_{\rm lin} \approx 0.1$

$$\begin{aligned} \eta_{\rm rem}^{k,i} &\leq \gamma_{\rm rem} \max\{\eta_{\rm disc}^{k,i}, \eta_{\rm lin}^{k,i}, \eta_{\rm alg}^{k,i}\}, \\ \eta_{\rm alg}^{k,i} &\leq \gamma_{\rm alg} \max\{\eta_{\rm disc}^{k,i}, \eta_{\rm lin}^{k,i}\}, \\ \eta_{\rm lin}^{k,i} &\leq \gamma_{\rm lin} \eta_{\rm disc}^{k,i} \end{aligned}$$

Theorem (Global efficiency)

Under the global stopping criteria and the usual assumptions,

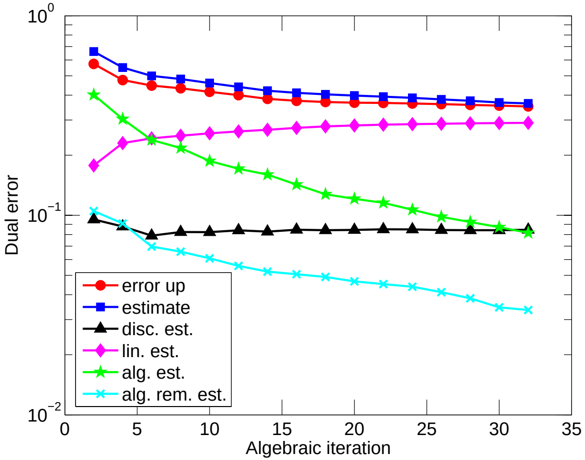
$$\eta_{\rm disc}^{k,i} + \eta_{\rm lin}^{k,i} + \eta_{\rm alg}^{k,i} + \eta_{\rm rem}^{k,i} \leq C(\mathcal{J}_u(u_h^{k,i}) + \eta_{\rm quad}^{k,i} + \eta_{\rm osc}),$$

where C is independent of $\overline{\sigma}$ and q.

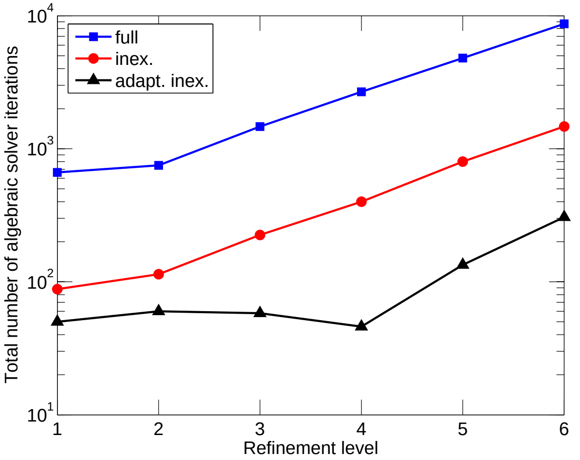
- local (elementwise) stopping criteria ⇒ **local** efficiency
- **robustness** with respect to the **nonlinearity** thanks to the choice of $\mathcal{J}_u$ as error measure

Outline

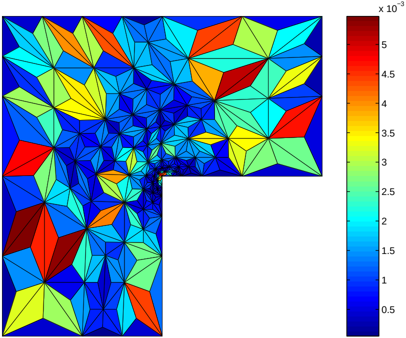
Error and estimators as a function of CG iterations, regular 10-Laplacian, 6th level mesh, 6th Newton step.



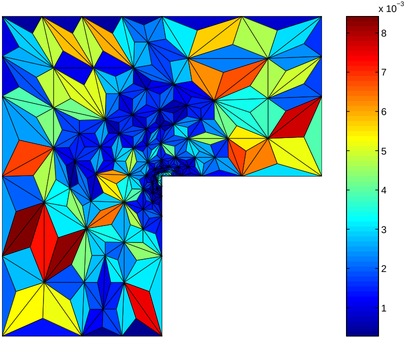
Overall algebraic solver iterations, regular 10-Laplacian



Full adaptivity, singular 4-Laplacian



Estimated error distribution



Exact error distribution

Outline

Two-phase flow in porous media

Two-phase flow in porous media

$$
\begin{aligned}
\partial_t(\phi s_\alpha) + \nabla\cdot\mathbf{u}_\alpha &= q_\alpha, && \alpha \in \{\mathrm{o}, \mathrm{w}\},\\
-\lambda_\alpha(s_\mathrm{w})\underline{\mathbf{K}}(\nabla p_\alpha + \rho_\alpha g \nabla z) &= \mathbf{u}_\alpha, && \alpha \in \{\mathrm{o}, \mathrm{w}\},\\
s_\mathrm{o} + s_\mathrm{w} &= 1,\\
p_\mathrm{o} - p_\mathrm{w} &= p_\mathrm{c}(s_\mathrm{w})
\end{aligned}
$$

+ boundary & initial conditions

Mathematical issues

- coupled system
- unsteady, nonlinear
- elliptic–degenerate parabolic type
- dominant advection

Two-phase flow in porous media

Two-phase flow in porous media

$$\begin{aligned} \partial_t(\phi s_\alpha) + \nabla\cdot\mathbf{u}_\alpha &= q_\alpha, & \alpha \in \{\mathrm{o}, \mathrm{w}\}, \\ -\lambda_\alpha(s_\mathrm{w})\underline{\mathbf{K}}(\nabla p_\alpha + \rho_\alpha g \nabla z) &= \mathbf{u}_\alpha, & \alpha \in \{\mathrm{o}, \mathrm{w}\}, \\ s_\mathrm{o} + s_\mathrm{w} &= 1, \\ p_\mathrm{o} - p_\mathrm{w} &= p_\mathrm{c}(s_\mathrm{w}) \end{aligned}$$

+ boundary & initial conditions

Mathematical issues

- coupled system
- unsteady, nonlinear
- elliptic–degenerate parabolic type
- dominant advection

Global and complementary pressures

Global pressure

$$\mathfrak{p}(s_{\mathrm{w}}, p_{\mathrm{w}}) := p_{\mathrm{w}} + \int_0^{s_{\mathrm{w}}} \frac{\lambda_{\mathrm{o}}(a)}{\lambda_{\mathrm{w}}(a) + \lambda_{\mathrm{o}}(a)} p_{\mathrm{c}}'(a) \mathrm{d}a$$

Complementary pressure

$$\mathfrak{q}(s_{\mathrm{w}}) := -\int_0^{s_{\mathrm{w}}} \frac{\lambda_{\mathrm{w}}(a)\lambda_{\mathrm{o}}(a)}{\lambda_{\mathrm{w}}(a) + \lambda_{\mathrm{o}}(a)} p_{\mathrm{c}}'(a) \mathrm{d}a$$

Comments

- necessary for the correct definition of the weak solution
- equivalent Darcy velocities expressions

$$\mathbf{u}_{\mathrm{w}}(s_{\mathrm{w}}, p_{\mathrm{w}}) := -\mathbf{K}\big(\lambda_{\mathrm{w}}(s_{\mathrm{w}})\nabla\mathfrak{p}(s_{\mathrm{w}}, p_{\mathrm{w}}) + \nabla\mathfrak{q}(s_{\mathrm{w}}) + \lambda_{\mathrm{w}}(s_{\mathrm{w}})\rho_{\mathrm{w}} g \nabla z\big),$$

$$\mathbf{u}_{\mathrm{o}}(s_{\mathrm{w}}, p_{\mathrm{w}}) := -\mathbf{K}\big(\lambda_{\mathrm{o}}(s_{\mathrm{w}})\nabla\mathfrak{p}(s_{\mathrm{w}}, p_{\mathrm{w}}) - \nabla\mathfrak{q}(s_{\mathrm{w}}) + \lambda_{\mathrm{o}}(s_{\mathrm{w}})\rho_{\mathrm{o}} g \nabla z\big)$$

Global and complementary pressures

Global pressure

$$\mathfrak{p}(s_{\rm w}, p_{\rm w}) := p_{\rm w} + \int_0^{s_{\rm w}} \frac{\lambda_{\rm o}(a)}{\lambda_{\rm w}(a) + \lambda_{\rm o}(a)} p_{\rm c}'(a) {\rm d}a$$

Complementary pressure

$$\mathfrak{q}(s_{\rm w}) := -\int_0^{s_{\rm w}} \frac{\lambda_{\rm w}(a)\lambda_{\rm o}(a)}{\lambda_{\rm w}(a) + \lambda_{\rm o}(a)} p_{\rm c}'(a) {\rm d}a$$

Comments

- necessary for the correct definition of the weak solution
- equivalent Darcy velocities expressions

$$\mathbf{u}_{\rm w}(s_{\rm w}, p_{\rm w}) := -\mathbf{K}\big(\lambda_{\rm w}(s_{\rm w})\nabla\mathfrak{p}(s_{\rm w}, p_{\rm w}) + \nabla\mathfrak{q}(s_{\rm w}) + \lambda_{\rm w}(s_{\rm w})\rho_{\rm w} g \nabla z\big),$$

$$\mathbf{u}_{\rm o}(s_{\rm w}, p_{\rm w}) := -\mathbf{K}\big(\lambda_{\rm o}(s_{\rm w})\nabla\mathfrak{p}(s_{\rm w}, p_{\rm w}) - \nabla\mathfrak{q}(s_{\rm w}) + \lambda_{\rm o}(s_{\rm w})\rho_{\rm o} g \nabla z\big)$$

Global and complementary pressures

Global pressure

$$\mathfrak{p}(s_{\rm w}, p_{\rm w}) := p_{\rm w} + \int_0^{s_{\rm w}} \frac{\lambda_{\rm o}(a)}{\lambda_{\rm w}(a) + \lambda_{\rm o}(a)} p_{\rm c}'(a) {\rm d}a$$

Complementary pressure

$$\mathfrak{q}(s_{\rm w}) := -\int_0^{s_{\rm w}} \frac{\lambda_{\rm w}(a)\lambda_{\rm o}(a)}{\lambda_{\rm w}(a) + \lambda_{\rm o}(a)} p_{\rm c}'(a) {\rm d}a$$

Comments

- necessary for the correct definition of the weak solution
- equivalent Darcy velocities expressions

$$\mathbf{u}_{\rm w}(s_{\rm w}, p_{\rm w}) := -\mathbf{K}\big(\lambda_{\rm w}(s_{\rm w})\nabla\mathfrak{p}(s_{\rm w}, p_{\rm w}) + \nabla\mathfrak{q}(s_{\rm w}) + \lambda_{\rm w}(s_{\rm w})\rho_{\rm w} g\nabla z\big),$$

$$\mathbf{u}_{\rm o}(s_{\rm w}, p_{\rm w}) := -\mathbf{K}\big(\lambda_{\rm o}(s_{\rm w})\nabla\mathfrak{p}(s_{\rm w}, p_{\rm w}) - \nabla\mathfrak{q}(s_{\rm w}) + \lambda_{\rm o}(s_{\rm w})\rho_{\rm o} g\nabla z\big)$$

Weak formulation

Energy space

$$X := L^2((0,T); H^1_{\rm D}(\Omega))$$

Definition (Weak solution (Arbogast 1992, Chen 2001))

Find $(s_{\rm w}, p_{\rm w})$ such that, with $s_{\rm o} := 1 - s_{\rm w}$,

$$s_{\rm w} \in C([0,T]; L^2(\Omega)),\ s_{\rm w}(\cdot, 0) = s^0_{\rm w},$$
$$\partial_t s_{\rm w} \in L^2((0,T); (H^1_{\rm D}(\Omega))'),$$
$$\mathfrak{p}(s_{\rm w}, p_{\rm w}) \in X,$$
$$\mathfrak{q}(s_{\rm w}) \in X,$$
$$\int_0^T \{\langle \partial_t(\phi s_\alpha), \varphi\rangle - (\mathbf{u}_\alpha(s_{\rm w}, p_{\rm w}), \nabla\varphi) - (q_\alpha, \varphi)\}{\rm d}t = 0$$
$$\forall \varphi \in X,\ \alpha \in \{{\rm o}, {\rm w}\}.$$

Weak formulation

Energy space

$$X := L^2((0,T); H^1_{\mathrm{D}}(\Omega))$$

Definition (Weak solution (Arbogast 1992, Chen 2001))

Find $(s_{\mathrm{w}}, p_{\mathrm{w}})$ such that, with $s_{\mathrm{o}} := 1 - s_{\mathrm{w}}$,

$$s_{\mathrm{w}} \in C([0,T]; L^2(\Omega)), \; s_{\mathrm{w}}(\cdot,0) = s^0_{\mathrm{w}},$$

$$\partial_t s_{\mathrm{w}} \in L^2((0,T); (H^1_{\mathrm{D}}(\Omega))'),$$

$$\mathfrak{p}(s_{\mathrm{w}}, p_{\mathrm{w}}) \in X,$$

$$\mathfrak{q}(s_{\mathrm{w}}) \in X,$$

$$\int_0^T \{\langle \partial_t(\phi s_\alpha), \varphi\rangle - (\mathbf{u}_\alpha(s_{\mathrm{w}}, p_{\mathrm{w}}), \nabla\varphi) - (q_\alpha, \varphi)\}\mathrm{d}t = 0$$

$$\forall \varphi \in X, \; \alpha \in \{\mathrm{o}, \mathrm{w}\}.$$

Link energy-type error – dual norm of the residual

Dual norm of the residual on the time interval I_n

$$\mathcal{J}^n_{s_{\rm w},p_{\rm w}}(s_{{\rm w},h\tau},p_{{\rm w},h\tau}) := \left\{ \sum_{\alpha\in\{{\rm o},{\rm w}\}} \left\{ \sup_{\varphi\in X_n,\, \|\varphi\|_{X_n}=1} \int_{I_n} \{\langle \partial_t(\phi s_\alpha) - \partial_t(\phi s_{\alpha,h\tau}), \varphi\rangle - (\mathbf{u}_\alpha(s_{\rm w},p_{\rm w}) - \mathbf{u}_\alpha(s_{{\rm w},h\tau},p_{{\rm w},h\tau}), \nabla\varphi)\}{\rm d}t \right\}^2 \right\}^{\frac{1}{2}}$$

Theorem (Link energy-type error – dual norm of the residual)

Let $(s_{\rm w}, p_{\rm w})$ be the weak solution. Let $(s_{{\rm w},h\tau}, p_{{\rm w},h\tau})$ be arbitrary such that $\mathfrak{p}(s_{{\rm w},h\tau}, p_{{\rm w},h\tau}) \in X$ and $\mathfrak{q}(s_{{\rm w},h\tau}) \in X$ (and satisfying the initial and boundary conditions for simplicity). Then

$$\begin{aligned}
&\|s_{\rm w} - s_{{\rm w},h\tau}\|_{L^2((0,T);H^{-1}(\Omega))} + \|\mathfrak{q}(s_{\rm w}) - \mathfrak{q}(s_{{\rm w},h\tau})\|_{L^2(\Omega\times(0,T))} \\
&+ \|\mathfrak{p}(s_{\rm w},p_{\rm w}) - \mathfrak{p}(s_{{\rm w},h\tau},p_{{\rm w},h\tau})\|_{L^2((0,T);H^1_0(\Omega))} \\
&\le C\left\{\sum_{n=1}^N \mathcal{J}^n_{s_{\rm w},p_{\rm w}}(s_{{\rm w},h\tau},p_{{\rm w},h\tau})^2\right\}^{\frac{1}{2}}.
\end{aligned}$$

Link energy-type error – dual norm of the residual

Dual norm of the residual on the time interval I_n

$$\mathcal{J}^n_{s_{\rm w},p_{\rm w}}(s_{{\rm w},h\tau},p_{{\rm w},h\tau}) := \left\{ \sum_{\alpha\in\{{\rm o},{\rm w}\}} \left\{ \sup_{\varphi\in X_n,\, \|\varphi\|_{X_n}=1} \int_{I_n} \{\langle \partial_t(\phi s_\alpha) - \partial_t(\phi s_{\alpha,h\tau}), \varphi\rangle - (\mathbf{u}_\alpha(s_{\rm w},p_{\rm w}) - \mathbf{u}_\alpha(s_{{\rm w},h\tau},p_{{\rm w},h\tau}), \nabla\varphi)\}{\rm d}t \right\}^2 \right\}^{\frac{1}{2}}$$

Theorem (Link energy-type error – dual norm of the residual)

Let $(s_{\rm w},p_{\rm w})$ be the weak solution. Let $(s_{{\rm w},h\tau},p_{{\rm w},h\tau})$ be arbitrary such that $\mathfrak{p}(s_{{\rm w},h\tau},p_{{\rm w},h\tau}) \in X$ and $\mathfrak{q}(s_{{\rm w},h\tau}) \in X$ (and satisfying the initial and boundary conditions for simplicity). Then

$$\|s_{\rm w} - s_{{\rm w},h\tau}\|_{L^2((0,T);H^{-1}(\Omega))} + \|\mathfrak{q}(s_{\rm w}) - \mathfrak{q}(s_{{\rm w},h\tau})\|_{L^2(\Omega\times(0,T))} + \|\mathfrak{p}(s_{\rm w},p_{\rm w}) - \mathfrak{p}(s_{{\rm w},h\tau},p_{{\rm w},h\tau})\|_{L^2((0,T);H^1_0(\Omega))} \le C\left\{\sum_{n=1}^N \mathcal{J}^n_{s_{\rm w},p_{\rm w}}(s_{{\rm w},h\tau},p_{{\rm w},h\tau})^2\right\}^{\frac{1}{2}}.$$

Link energy-type error – dual norm of the residual

Dual norm of the residual on the time interval I_n

$$\mathcal{J}^n_{s_{\rm w},p_{\rm w}}(s_{{\rm w},h\tau},p_{{\rm w},h\tau}) := \left\{ \sum_{\alpha\in\{{\rm o},{\rm w}\}} \left\{ \sup_{\varphi\in X_n,\,\|\varphi\|_{X_n}=1} \int_{I_n} \{\langle \partial_t(\phi s_\alpha) - \partial_t(\phi s_{\alpha,h\tau}), \varphi\rangle - (\mathbf{u}_\alpha(s_{\rm w},p_{\rm w}) - \mathbf{u}_\alpha(s_{{\rm w},h\tau},p_{{\rm w},h\tau}), \nabla\varphi)\}{\rm d}t \right\}^2 \right\}^{\frac{1}{2}}$$

Theorem (Link energy-type error – dual norm of the residual)

Let $(s_{\rm w},p_{\rm w})$ be the weak solution. Let $(s_{{\rm w},h\tau},p_{{\rm w},h\tau})$ be arbitrary such that $\mathfrak{p}(s_{{\rm w},h\tau},p_{{\rm w},h\tau})\in X$ and $\mathfrak{q}(s_{{\rm w},h\tau})\in X$ (and satisfying the initial and boundary conditions for simplicity). Then

$$\|s_{\rm w}-s_{{\rm w},h\tau}\|_{L^2((0,T);H^{-1}(\Omega))} + \|\mathfrak{q}(s_{\rm w})-\mathfrak{q}(s_{{\rm w},h\tau})\|_{L^2(\Omega\times(0,T))} + \|\mathfrak{p}(s_{\rm w},p_{\rm w})-\mathfrak{p}(s_{{\rm w},h\tau},p_{{\rm w},h\tau})\|_{L^2((0,T);H^1_0(\Omega))} \le C\left\{\sum_{n=1}^N \mathcal{J}^n_{s_{\rm w},p_{\rm w}}(s_{{\rm w},h\tau},p_{{\rm w},h\tau})^2\right\}^{\frac{1}{2}}.$$

Link energy-type error – dual norm of the residual

Dual norm of the residual on the time interval I_n

$$\mathcal{J}^n_{s_{\rm w},p_{\rm w}}(s_{{\rm w},h\tau},p_{{\rm w},h\tau}) := \left\{ \sum_{\alpha\in\{{\rm o},{\rm w}\}} \left\{ \sup_{\varphi\in X_n,\, \|\varphi\|_{X_n}=1} \int_{I_n} \{\langle \partial_t(\phi s_\alpha) - \partial_t(\phi s_{\alpha,h\tau}), \varphi\rangle - (\mathbf{u}_\alpha(s_{\rm w},p_{\rm w}) - \mathbf{u}_\alpha(s_{{\rm w},h\tau},p_{{\rm w},h\tau}), \nabla\varphi)\}\,{\rm d}t \right\}^2 \right\}^{\frac{1}{2}}$$

Theorem (Link energy-type error – dual norm of the residual)

Let $(s_{\rm w},p_{\rm w})$ be the weak solution. Let $(s_{{\rm w},h\tau},p_{{\rm w},h\tau})$ be arbitrary such that $\mathfrak{p}(s_{{\rm w},h\tau},p_{{\rm w},h\tau}) \in X$ and $\mathfrak{q}(s_{{\rm w},h\tau}) \in X$ (and satisfying the initial and boundary conditions for simplicity). Then

$$\|s_{\rm w} - s_{{\rm w},h\tau}\|_{L^2((0,T);H^{-1}(\Omega))} + \|\mathfrak{q}(s_{\rm w}) - \mathfrak{q}(s_{{\rm w},h\tau})\|_{L^2(\Omega\times(0,T))} + \|\mathfrak{p}(s_{\rm w},p_{\rm w}) - \mathfrak{p}(s_{{\rm w},h\tau},p_{{\rm w},h\tau})\|_{L^2((0,T);H^1_0(\Omega))} \le C\left\{\sum_{n=1}^N \mathcal{J}^n_{s_{\rm w},p_{\rm w}}(s_{{\rm w},h\tau},p_{{\rm w},h\tau})^2\right\}^{\frac{1}{2}}.$$

Distinguishing the error components

Theorem (Distinguishing the error components)

Consider a vertex-centered finite volume / backward Euler approximation and Newton linearization. Let

- n *be the* *time* *step,*
- k *be the* *linearization* *step,*
- i *be the* *algebraic solver* *step,*

with the approximations $(s_{\mathrm{w},h\mathcal{T}}^{n,k,i}, p_{\mathrm{w},h\mathcal{T}}^{n,k,i})$*. Then*

$$\mathcal{J}_{s_\mathrm{w},p_\mathrm{w}}^{n}(s_{\mathrm{w},h\mathcal{T}}^{n,k,i}, p_{\mathrm{w},h\mathcal{T}}^{n,k,i}) \le \eta_{\mathrm{sp}}^{n,k,i} + \eta_{\mathrm{tm}}^{n,k,i} + \eta_{\mathrm{lin}}^{n,k,i} + \eta_{\mathrm{alg}}^{n,k,i}.$$

Error components

- $\eta_{\mathrm{sp}}^{n,k,i}$: spatial discretization
- $\eta_{\mathrm{tm}}^{n,k,i}$: temporal discretization
- $\eta_{\mathrm{lin}}^{n,k,i}$: linearization
- $\eta_{\mathrm{alg}}^{n,k,i}$: algebraic solver

Distinguishing the error components

Theorem (Distinguishing the error components)

Consider a vertex-centered finite volume / backward Euler approximation and Newton linearization. Let

- n *be the time step,*
- k *be the linearization step,*
- i *be the algebraic solver step,*

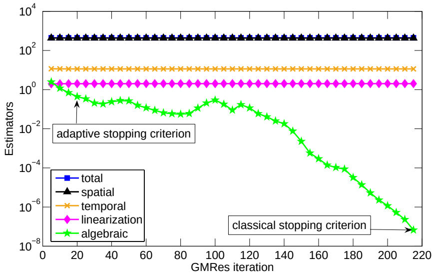
with the approximations $(s_{\mathrm{w},h\tau}^{n,k,i}, p_{\mathrm{w},h\tau}^{n,k,i})$. *Then*

$$\mathcal{J}_{s_\mathrm{w},p_\mathrm{w}}^{n}(s_{\mathrm{w},h\tau}^{n,k,i}, p_{\mathrm{w},h\tau}^{n,k,i}) \leq \eta_{\mathrm{sp}}^{n,k,i} + \eta_{\mathrm{tm}}^{n,k,i} + \eta_{\mathrm{lin}}^{n,k,i} + \eta_{\mathrm{alg}}^{n,k,i}.$$

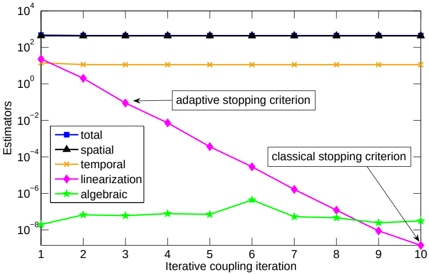
Error components

- $\eta_{\mathrm{sp}}^{n,k,i}$: spatial discretization
- $\eta_{\mathrm{tm}}^{n,k,i}$: temporal discretization
- $\eta_{\mathrm{lin}}^{n,k,i}$: linearization
- $\eta_{\mathrm{alg}}^{n,k,i}$: algebraic solver

Estimators and stopping criteria

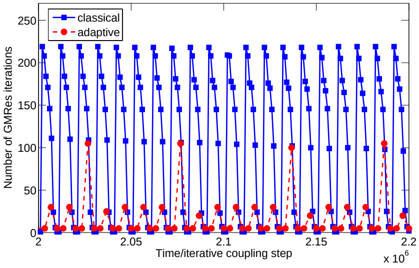


Estimators in function of GMRes iterations

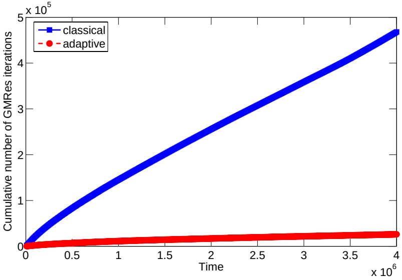


Estimators in function of iterative coupling iterations

GMRes iterations

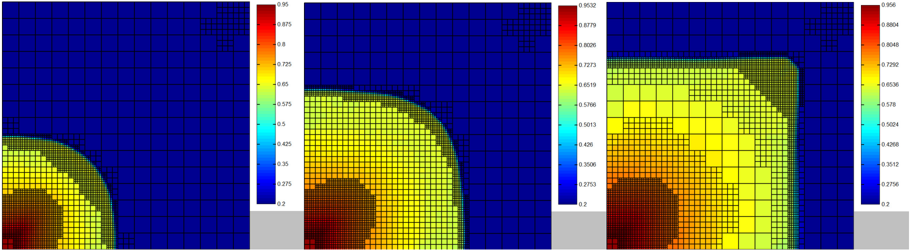


Per time and iterative coupling step



Cumulated

Space/time/nonlinear solver/linear solver adaptivity



Fully adaptive computation

Full adaptivity

Full adaptivity

- only a **necessary number** of **algebraic/linearization solver iterations**
- adaptive **regularization**, model adaptation, adaptive choice of the scheme parameters
- **"online decisions"**: algebraic step / linearization step / space mesh refinement / time step modification
- important computational savings

Outline

Setting

Laplace eigenvalue problem

Find eigenvector & eigenvalue pair (u, λ) such that $\|u\| = 1$ and

$$\begin{aligned} -\Delta u &= \lambda u \qquad \text{in } \Omega \subset \mathbb{R}^d, d = 2, 3, \\ u &= 0 \qquad \text{on } \partial\Omega. \end{aligned}$$

Full problem, weak formulation

Find $(u_k, \lambda_k) \in V := H_0^1(\Omega) \times \mathbb{R}^+$ with $\|u_k\| = 1$, $k \geq 1$, s. t.

$$(\nabla u_k, \nabla v) = \lambda_k (u_k, v) \qquad \forall v \in V$$

- $0 < \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_k \to \infty$

Setting

Laplace eigenvalue problem

Find eigenvector & eigenvalue pair (u, λ) such that $\|u\| = 1$ and

$$\begin{aligned} -\Delta u &= \lambda u \qquad \text{in } \Omega \subset \mathbb{R}^d, d = 2, 3, \\ u &= 0 \qquad \text{on } \partial\Omega. \end{aligned}$$

Full problem, weak formulation

Find $(u_k, \lambda_k) \in V := H_0^1(\Omega) \times \mathbb{R}^+$ with $\|u_k\| = 1$, $k \geq 1$, s. t.

$$(\nabla u_k, \nabla v) = \lambda_k (u_k, v) \qquad \forall v \in V$$

- $0 < \lambda_1 < \lambda_2 \leq \ldots \leq \lambda_k \to \infty$

Approximate solution and assumptions

Assumption A (Conforming variational solution)

There holds

- $(u_h, \lambda_h) \in V \times \mathbb{R}^+$
- $\|u_h\| = 1$
- $(u_h, 1) > 0$
- $\|\nabla u_h\|^2 = \lambda_h$

Assumption B (Galerkin orthogonality of the residual to $\psi_{\mathbf{a}}$)

There holds, for all $\mathbf{a} \in \mathcal{V}_h^{\mathrm{int}}$,

$$\lambda_h(u_h, \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} - (\nabla u_h, \nabla \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = 0.$$

Assumption C (Piecewise polynomial form)

There holds $u_h \in \mathbb{P}_p(\mathcal{T}_h)$, $p \geq 1$, *and the mixed finite element spaces* $\mathbf{V}_h \times Q_h$ *are of degree* $p+1$.

Approximate solution and assumptions

Assumption A (Conforming variational solution)

There holds

- $(u_h, \lambda_h) \in V \times \mathbb{R}^+$
- $\|u_h\| = 1$
- $(u_h, 1) > 0$
- $\|\nabla u_h\|^2 = \lambda_h$

Assumption B (Galerkin orthogonality of the residual to $\psi_{\mathbf{a}}$)

There holds, for all $\mathbf{a} \in \mathcal{V}_h^{\rm int}$,

$$\lambda_h(u_h, \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} - (\nabla u_h, \nabla \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = 0.$$

Assumption C (Piecewise polynomial form)

There holds $u_h \in \mathbb{P}_p(\mathcal{T}_h)$, $p \geq 1$, *and the mixed finite element spaces* $\mathbf{V}_h \times Q_h$ *are of degree* $p + 1$.

ematics

Approximate solution and assumptions

Assumption A (Conforming variational solution)

There holds

- $(u_h, \lambda_h) \in V \times \mathbb{R}^+$
- $\|u_h\| = 1$
- $(u_h, 1) > 0$
- $\|\nabla u_h\|^2 = \lambda_h$

Assumption B (Galerkin orthogonality of the residual to $\psi_{\mathbf{a}}$)

There holds, for all $\mathbf{a} \in \mathcal{V}_h^{\text{int}}$,

$$\lambda_h (u_h, \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} - (\nabla u_h, \nabla \psi_{\mathbf{a}})_{\omega_{\mathbf{a}}} = 0.$$

Assumption C (Piecewise polynomial form)

There holds $u_h \in \mathbb{P}_p(\mathcal{T}_h)$, $p \geq 1$, *and the mixed finite element spaces* $\mathbf{V}_h \times Q_h$ *are of degree* $p+1$.

Guaranteed bounds for the first eigenvalue

Theorem (Eigenvalue bounds)

Let $0 < \underline{\lambda_2} \leq \lambda_2$ and $0 < \underline{\lambda_1} \leq \lambda_1$. Let Assumptions A and B hold and let $\lambda_h < \underline{\lambda_2}$. Let $\boldsymbol{\sigma}_h$ be an equilibrated flux and let

$$\underbrace{\beta_h}_{\searrow 0} := \frac{1}{\sqrt{\underline{\lambda_1}}}\left(1 - \frac{\lambda_h}{\underline{\lambda_2}}\right)^{-1} \|\nabla u_h + \boldsymbol{\sigma}_h\| < 1,$$

$$\underbrace{\alpha_h^2}_{\searrow 0} := 2\left(1 - \sqrt{1 - \beta_h^2}\right) \leq |\Omega|^{-1}(u_h, 1)^2.$$

Then

$$\lambda_1 \geq \lambda_h - \overbrace{\left(1 - \frac{\lambda_h}{\underline{\lambda_2}}\right)^{-2}\left(1 - \frac{\alpha_h^2}{4}\right)^{-1} \|\nabla u_h + \boldsymbol{\sigma}_h\|^2}^{\eta^2},$$

$$\lambda_1 \leq \lambda_h - \tilde{\eta}^2.$$

Guaranteed bounds for the first eigenvalue

Theorem (Eigenvalue bounds)

Let $0 < \underline{\lambda_2} \le \lambda_2$ and $0 < \underline{\lambda_1} \le \lambda_1$. Let Assumptions A and B hold and let $\lambda_h < \underline{\lambda_2}$. Let $\boldsymbol{\sigma}_h$ be an equilibrated flux and let

$$\underbrace{\beta_h}_{\searrow 0} := \frac{1}{\sqrt{\underline{\lambda_1}}}\left(1 - \frac{\lambda_h}{\underline{\lambda_2}}\right)^{-1} \|\nabla u_h + \boldsymbol{\sigma}_h\| < 1,$$

$$\underbrace{\alpha_h^2}_{\searrow 0} := 2\left(1 - \sqrt{1 - \beta_h^2}\right) \le |\Omega|^{-1}(u_h, 1)^2.$$

Then

$$\lambda_1 \ge \lambda_h - \overbrace{\left(1 - \frac{\lambda_h}{\underline{\lambda_2}}\right)^{-2}\left(1 - \frac{\alpha_h^2}{4}\right)^{-1} \|\nabla u_h + \boldsymbol{\sigma}_h\|^2}^{\eta^2},$$

$$\lambda_1 \le \lambda_h - \tilde{\eta}^2.$$

Guaranteed bounds for the first eigenvalue

Theorem (Eigenvalue bounds)

Let $0 < \underline{\lambda_2} \leq \lambda_2$ *and* $0 < \underline{\lambda_1} \leq \lambda_1$. *Let Assumptions A and B hold and let* $\lambda_h < \underline{\lambda_2}$. *Let* $\boldsymbol{\sigma}_h$ *be an equilibrated flux and let*

$$\underbrace{\beta_h}_{\searrow 0} := \frac{1}{\sqrt{\underline{\lambda_1}}} \left(1 - \frac{\lambda_h}{\underline{\lambda_2}}\right)^{-1} \|\nabla u_h + \boldsymbol{\sigma}_h\| < 1,$$

$$\underbrace{\alpha_h^2}_{\searrow 0} := 2\left(1 - \sqrt{1 - \beta_h^2}\right) \leq |\Omega|^{-1}(u_h, 1)^2.$$

Then

$$\lambda_1 \geq \lambda_h - \overbrace{\left(1 - \frac{\lambda_h}{\underline{\lambda_2}}\right)^{-2} \left(1 - \frac{\alpha_h^2}{4}\right)^{-1} \|\nabla u_h + \boldsymbol{\sigma}_h\|^2}^{\eta^2},$$

$$\lambda_1 \leq \lambda_h - \tilde{\eta}^2.$$

Guaranteed bounds for the first eigenvector

Theorem (Eigenvector bounds)

Let the assumptions of the eigenvalue theorem be verified. Then

$$\|\nabla(u_1 - u_h)\| \leq \eta.$$

Moreover, under Assumption C,

$$\eta \leq (d+1)C_{\mathrm{cont,PF}}C_{\mathrm{st}} \underbrace{\left(\frac{\|\nabla(u_1-u_h)\|^2}{\lambda_1}+1\right)^{\frac{1}{2}}}_{\searrow 1}$$

$$\underbrace{\left(1-\frac{\lambda_h}{\lambda_2}\right)^{-1}}_{\searrow\left(1-\frac{\lambda_1}{\lambda_2}\right)^{-1}} \underbrace{\left(1-\frac{\alpha_h^2}{4}\right)^{-\frac{1}{2}}}_{\searrow 1} \|\nabla(u_1-u_h)\|.$$

Guaranteed bounds for the first eigenvector

Theorem (Eigenvector bounds)

Let the assumptions of the eigenvalue theorem be verified. Then

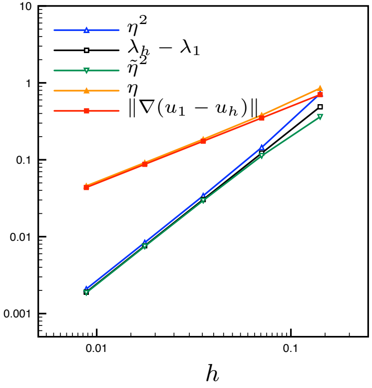
$$\|\nabla(u_1 - u_h)\| \leq \eta.$$

Moreover, under Assumption C,

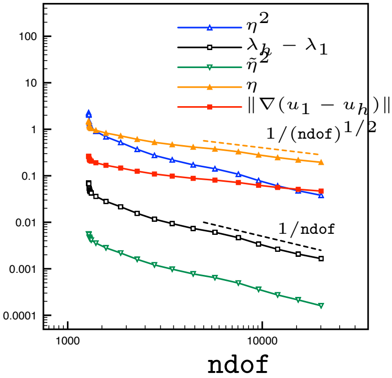
$$\eta \leq (d+1) C_{\mathrm{cont,PF}} C_{\mathrm{st}} \underbrace{\left(\frac{\|\nabla(u_1 - u_h)\|^2}{\lambda_1} + 1 \right)^{\frac{1}{2}}}_{\searrow 1}$$

$$\underbrace{\left(1 - \frac{\lambda_h}{\underline{\lambda_2}}\right)^{-1}}_{\searrow \left(1 - \frac{\lambda_1}{\underline{\lambda_2}}\right)^{-1}} \underbrace{\left(1 - \frac{\alpha_h^2}{4}\right)^{-\frac{1}{2}}}_{\searrow 1} \|\nabla(u_1 - u_h)\|.$$

Guaranteed eigenvalue and eigenvector bounds



Square, uniform refinement



L-shape, adaptive refinement

Outline

Previous results (Laplace)

Global flux reconstructions

- Prager and Synge (1947):

$$\|\nabla u + \boldsymbol{\sigma}_h\|^2 + \|\nabla(u - u_h)\|^2 = \|\nabla u_h + \boldsymbol{\sigma}_h\|^2$$

for **any** $u_h \in H_0^1(\Omega)$ and **any** $\boldsymbol{\sigma}_h \in \mathbf{H}(\mathrm{div}, \Omega)$ s.t. $\nabla \cdot \boldsymbol{\sigma}_h = f$

- Hlaváček, Haslinger, Nečas, and Lovíšek (1979), Repin (1997), . . . : global construction of $\boldsymbol{\sigma}_h$: unprecise/costly

Local flux reconstructions

- Ladevèze and Leguillon (1983), equilibrated face fluxes
- Destuynder and Métivet (1999), discrete flux σ_h
- Luce and Wohlmuth (2004), local efficiency proof
- Kim (2007), Ainsworth (2007), Ern, Nicaise, and Vohralík (2007), discontinuous Galerkin method elementwise prescription
- Braess and Schöberl (2008), Vohralík (2008), Ern and Vohralík (2009), **local Neumann MFE problems**

Previous results (Laplace)

Global flux reconstructions

- Prager and Synge (1947):

$$\|\nabla u + \boldsymbol{\sigma}_h\|^2 + \|\nabla(u - u_h)\|^2 = \|\nabla u_h + \boldsymbol{\sigma}_h\|^2$$

for **any** $u_h \in H^1_0(\Omega)$ and **any** $\boldsymbol{\sigma}_h \in \mathbf{H}(\mathrm{div}, \Omega)$ s.t. $\nabla\cdot\boldsymbol{\sigma}_h = f$

- Hlaváček, Haslinger, Nečas, and Lovíšek (1979), Repin (1997), . . . : global construction of $\boldsymbol{\sigma}_h$: unprecise/costly

Local flux reconstructions

- Ladevèze and Leguillon (1983), equilibrated face fluxes
- Destuynder and Métivet (1999), discrete flux $\boldsymbol{\sigma}_h$
- Luce and Wohlmuth (2004), local efficiency proof
- Kim (2007), Ainsworth (2007), Ern, Nicaise, and Vohralík (2007), discontinuous Galerkin method elementwise prescription
- Braess and Schöberl (2008), Vohralík (2008), Ern and Vohralík (2009), **local Neumann MFE problems**

Previous results (nonlinear and stopping criteria)

A posteriori error estimates for numerical discretizations of nonlinear problems

- Ladevèze (since 1990's), guaranteed upper bound
- Verfürth (1994), residual estimates
- Chaillou and Suri (2006, 2007), distinguishing discretization and linearization errors

Stopping criteria for algebraic solvers

- Becker, Johnson, and Rannacher (1995), multigrid stopping criterion
- Arioli (2000's), comparison of the algebraic and discretization errors

Previous results (nonlinear and stopping criteria)

A posteriori error estimates for numerical discretizations of nonlinear problems

- Ladevèze (since 1990's), guaranteed upper bound
- Verfürth (1994), residual estimates
- Chaillou and Suri (2006, 2007), distinguishing discretization and linearization errors

Stopping criteria for algebraic solvers

- Becker, Johnson, and Rannacher (1995), multigrid stopping criterion
- Arioli (2000's), comparison of the algebraic and discretization errors

Previous results (unsteady nonlinear)

Nonlinear unsteady problems

- Eriksson and Johnson (1995), $L^\infty(0, T; L^2(\Omega))$ estimates exploiting stability of the adjoint problem
- Gallimard, Ladevèze, Pelle (1997), const. rel. estimates
- Verfürth (1998), framework for energy control, efficiency
- Ohlberger (2001), non-energy estimates, hyperbolic limit
- Akrivis, Makridakis, and Nochetto (2006), higher-order temporal discretizations

Degenerate parabolic problems

- Nochetto, Schmidt, Verdi (2000), Stefan problem

Two-phase flows

- Chen and Ewing (2003), mesh adaptivity

Previous results (unsteady nonlinear)

Nonlinear unsteady problems

- Eriksson and Johnson (1995), $L^\infty(0, T; L^2(\Omega))$ estimates exploiting stability of the adjoint problem
- Gallimard, Ladevèze, Pelle (1997), const. rel. estimates
- Verfürth (1998), framework for energy control, efficiency
- Ohlberger (2001), non-energy estimates, hyperbolic limit
- Akrivis, Makridakis, and Nochetto (2006), higher-order temporal discretizations

Degenerate parabolic problems

- Nochetto, Schmidt, Verdi (2000), Stefan problem

Two-phase flows

- Chen and Ewing (2003), mesh adaptivity

Previous results (unsteady nonlinear)

Nonlinear unsteady problems

- Eriksson and Johnson (1995), $L^\infty(0, T; L^2(\Omega))$ estimates exploiting stability of the adjoint problem
- Gallimard, Ladevèze, Pelle (1997), const. rel. estimates
- Verfürth (1998), framework for energy control, efficiency
- Ohlberger (2001), non-energy estimates, hyperbolic limit
- Akrivis, Makridakis, and Nochetto (2006), higher-order temporal discretizations

Degenerate parabolic problems

- Nochetto, Schmidt, Verdi (2000), Stefan problem

Two-phase flows

- Chen and Ewing (2003), mesh adaptivity

Previous results (eigenproblems)

Eigenvalues

- Carstensen and Gedicke (2014): guaranteed bound, lowest-order nonconforming FEs
- Hu, Huang, Lin (2014): nonconforming FEs (saturation assumption may be necessary)
- Šebestová and Vejchodský (2014), Kuznetsov and Repin (2013): guaranteed bounds (condition on applicability, suboptimal convergence speed)
- Liu and Oishi (2013): guaranteed bound, lowest-order conforming FEs (auxiliary eigenvalue problem)

Eigenvectors

- Rannacher, Westenberger, Wollner (2010), Grubišić and Ovall (2009), Durán, Padra, Rodríguez (2003), Heuveline and Rannacher (2002), Larson (2000), Maday and Patera (2000), Verfürth (1994) (uncomputable higher-order terms)

Bibliography

Bibliography

- ERN A., VOHRALÍK M., Polynomial-degree-robust a posteriori estimates in a unified setting for conforming, nonconforming, discontinuous Galerkin, and mixed discretizations, *SIAM J. Numer. Anal.*, **53** (2015), 1058–1081.
- CANCÈS E., DUSSON G., MADAY Y., STAMM B., VOHRALÍK M., Guaranteed and robust a posteriori bounds for Laplace eigenvalues and eigenvectors: conforming approximations, HAL Preprint 01194364.
- ERN A., VOHRALÍK M., Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs, *SIAM J. Sci. Comput.* **35** (2013), A1761–A1791.
- CANCÈS C., POP I. S., VOHRALÍK M., An a posteriori error estimate for vertex-centered finite volume discretizations of immiscible incompressible two-phase flow, *Math. Comp.* **83** (2014), 153–188.

Thank you for your attention!