Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs

Alexandre Ern and Martin Vohralík

INRIA Paris-Rocquencourt

Schnelle Löser für Partielle Differentialgleichungen
Oberwolfach, May 11–17, 2014
Inexact iterative linearization

System of nonlinear algebraic equations
Nonlinear operator $A : \mathbb{R}^N \to \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.
$$A(U) = F$$

Algorithm (Inexact iterative linearization)

1. Choose initial vector $U^0$. Set $k := 1$.
2. $U^{k-1} \Rightarrow$ matrix $A^{k-1}$ and vector $F^{k-1}$: find $U^k$ s.t.
   $$A^{k-1}U^k \approx F^{k-1}.$$ 
3. 1. Set $U^{k,0} := U^{k-1}$ and $i := 1$.
   2. Do 1 algebraic solver step $\Rightarrow$ $U^{k,i}$ s.t. ($R^{k,i}$ algebraic res.)
      $$A^{k-1}U^{k,i} = F^{k-1} - R^{k,i}.$$ 
   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 $\mathcal{A}^{k-1}$ and vector $F^{k-1}$: find $U^k$ s.t.
$$\mathcal{A}^{k-1} U^k \approx F^{k-1}.$$
3. Set $U^{k,0} := U^{k-1}$ and $i := 1$.
   1. Do 1 algebraic solver step $\Rightarrow U^{k,i}$ s.t. ($R^{k,i}$ algebraic res.)
   $$\mathcal{A}^{k-1} U^{k,i} = F^{k-1} - R^{k,i}.$$  
   2. 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 $A : \mathbb{R}^N \rightarrow \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.

$A(U) = F$

Algorithm (Inexact iterative linearization)

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

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

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

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

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 $A : \mathbb{R}^N \rightarrow \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.
$$A(U) = F$$

Algorithm (Inexact iterative linearization)

1. Choose initial vector $U^0$. Set $k := 1$.
2. $U^{k-1} \Rightarrow$ matrix $A^{k-1}$ and vector $F^{k-1}$: find $U^k$ s.t.
$$A^{k-1}U^k \approx F^{k-1}.$$
3. 1. Set $U^{k:0} := U^{k-1}$ and $i := 1$.
   2. Do 1 algebraic solver step $\Rightarrow U^{k:i}$ s.t. $(R^{k:i}$ algebraic res.)
$$A^{k-1}U^{k:i} = F^{k-1} - R^{k:i}.$$  
   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 $A : \mathbb{R}^N \rightarrow \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.
$$A(U) = F$$

Algorithm (Inexact iterative linearization)

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

System of nonlinear algebraic equations
Nonlinear operator $A : \mathbb{R}^N \rightarrow \mathbb{R}^N$, vector $F \in \mathbb{R}^N$: find $U \in \mathbb{R}^N$ s.t.

$A(U) = F$

Algorithm (Inexact iterative linearization)

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

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

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

4. Convergence? $\text{OK} \Rightarrow \text{finish}$. $\text{KO} \Rightarrow k := k + 1$, back to 2.
System of nonlinear algebraic equations
Nonlinear operator \( A : \mathbb{R}^N \rightarrow \mathbb{R}^N \), vector \( F \in \mathbb{R}^N \): find \( U \in \mathbb{R}^N \) s.t. \( A(U) = F \)

Algorithm (Inexact iterative linearization)

1. Choose initial vector \( U^0 \). Set \( k := 1 \).
2. \( U^{k-1} \Rightarrow \) matrix \( A^{k-1} \) and vector \( F^{k-1} \): find \( U^k \) s.t. \( A^{k-1} U^k \approx F^{k-1} \).
3. 1. Set \( U^{k,0} := U^{k-1} \) and \( i := 1 \).
   2. Do 1 algebraic solver step \( \Rightarrow U^{k,i} \) s.t. (\( R^{k,i} \) algebraic res.) \( A^{k-1} U^{k,i} = F^{k-1} - R^{k,i} \).
   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 $A(U^{k,i}) = F$

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 $A(U_{k,i}) = F$

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 $A(U^{k,i}) = F$

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

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 $\|u - u^{k,i}_h\|$ 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 $A(U^{k,i}) = F$

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

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 $\|u - u^{k,i}_h\|$ 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 \( A(U^{k,i}) = F \)

Numerical method
- underlying numerical method: the vector \( U^{k,i} \) is associated with a (piecewise polynomial) approximation \( u^{k,i}_h \)

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 \( \| u - u^{k,i}_h \| \) on Newton step \( k \) and algebraic solver step \( i \), how is it distributed?
Outline

1. Bibliography
2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results
3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results
4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results
5. Conclusions and future directions
Outline

1. Bibliography

2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results

3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results

4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results

5. Conclusions and future directions
Previous results

Inexact Newton method

- Eisenstat and Walker (1990’s) (conception, convergence, a priori error estimates)
- Moret (1989) (discrete a posteriori error estimates)

Adaptive inexact Newton method

- Bank and Rose (1982), combination with multigrid
- Hackbusch and Reusken (1989), damping and multigrid
- Deuflhard (1990’s, 2004 book), adaptivity

Stopping criteria for algebraic solvers

- engineering literature, since 1950’s
- Becker, Johnson, and Rannacher (1995), multigrid stopping criterion
- Arioli (2000’s), comparison of the algebraic and discretization errors
Previous results

**Inexact Newton method**
- Eisenstat and Walker (1990’s) (conception, convergence, a priori error estimates)
- Moret (1989) (discrete a posteriori error estimates)

**Adaptive inexact Newton method**
- Bank and Rose (1982), combination with multigrid
- Hackbusch and Reusken (1989), damping and multigrid
- Deuflhard (1990’s, 2004 book), adaptivity

**Stopping criteria for algebraic solvers**
- engineering literature, since 1950’s
- Becker, Johnson, and Rannacher (1995), multigrid stopping criterion
- Arioli (2000’s), comparison of the algebraic and discretization errors
Previous results

Inexact Newton method
- Eisenstat and Walker (1990’s) (conception, convergence, a priori error estimates)
- Moret (1989) (discrete a posteriori error estimates)

Adaptive inexact Newton method
- Bank and Rose (1982), combination with multigrid
- Hackbusch and Reusken (1989), damping and multigrid
- Deuflhard (1990’s, 2004 book), adaptivity

Stopping criteria for algebraic solvers
- engineering literature, since 1950’s
- Becker, Johnson, and Rannacher (1995), multigrid stopping criterion
- Arioli (2000’s), comparison of the algebraic and discretization errors
Previous results

A posteriori error estimates for numerical discretizations of nonlinear problems

- Ladevèze (since 1990’s), guaranteed upper bound
- Han (1994), general framework
- Verfürth (1994), residual estimates
- Carstensen and Klose (2003), guaranteed estimates
- Chaillou and Suri (2006, 2007), distinguishing discretization and linearization errors
- Kim (2007), guaranteed estimates, locally conservative methods
Outline

1. Bibliography
2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results
3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results
4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results
5. Conclusions and future directions
Outline

1. Bibliography

2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results

3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results

4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results

5. Conclusions and future directions
Theorem (A guaranteed a posteriori error estimate)

Let \( u \in H^1_0(\Omega) \) be the weak solution,

\( u_h \in H^1(\mathcal{T}_h) := \{ v \in L^2(\Omega), v|_K \in H^1(K) \; \forall K \in \mathcal{T}_h \} \) be arbitrary,

\( s_h \in H^1_0(\Omega) \) and \( \sigma_h \in H(\text{div}, \Omega) \) with \( (\nabla \cdot \sigma_h, 1)_K = (f, 1)_K \) for all \( K \in \mathcal{T}_h \) be arbitrary.

Then
\[
\| \nabla (u - u_h) \|^2 \leq \sum_{K \in \mathcal{T}_h} \left( \| \nabla u_h + \sigma_h \|_K + \frac{h_K}{\pi} \| f - \nabla \cdot \sigma_h \|_K \right)^2 + \sum_{K \in \mathcal{T}_h} \| \nabla (u_h - s_h) \|^2_K.
\]

Proof (Spirit of Prager–Synge (1947)).

- define \( s \in H^1_0(\Omega) \) by
  \[
  (\nabla s, \nabla v) = (\nabla u_h, \nabla v) \quad \forall v \in H^1_0(\Omega)
  \]
- develop (Pythagoras)
  \[
  \| \nabla (u - u_h) \|^2 = \| \nabla (u - s) \|^2 + \| \nabla (s - u_h) \|^2
  \]
Theorem (A guaranteed a posteriori error estimate)

Let \( u \in H^1_0(\Omega) \) be the weak solution,

\[ u_h \in H^1(\mathcal{T}_h) := \{ v \in L^2(\Omega), v|_K \in H^1(K) \ \forall K \in \mathcal{T}_h \} \]

be arbitrary,

\[ s_h \in H^1_0(\Omega) \]

and \( \sigma_h \in H(\text{div}, \Omega) \) with \((\nabla \cdot \sigma_h, 1)_K = (f, 1)_K\) for all \( K \in \mathcal{T}_h \)

be arbitrary.

Then

\[
\| \nabla (u - u_h) \|^2 \leq \sum_{K \in \mathcal{T}_h} \left( \| \nabla u_h + \sigma_h \|_K + \frac{h_K}{\pi} \| f - \nabla \cdot \sigma_h \|_K \right)^2 + \sum_{K \in \mathcal{T}_h} \| \nabla (u_h - s_h) \|^2_K.
\]

Proof (Spirit of Prager–Synge (1947)).

Define \( s \in H^1_0(\Omega) \) by

\[
(\nabla s, \nabla v) = (\nabla u_h, \nabla v) \quad \forall v \in H^1_0(\Omega)
\]

develop (Pythagoras)

\[
\| \nabla (u - u_h) \|^2 = \| \nabla (u - s) \|^2 + \| \nabla (s - u_h) \|^2
\]
A posteriori error estimate, $-\Delta u = f$ in $\Omega$, $u = 0$ on $\partial\Omega$

**Theorem (A guaranteed a posteriori error estimate)**

- Let $u \in H^1_0(\Omega)$ be the weak solution,
- $u_h \in H^1(T_h) := \{v \in L^2(\Omega), v|_K \in H^1(K) \forall K \in T_h\}$ be arbitrary,
- $s_h \in H^1_0(\Omega)$ and $\sigma_h \in H(\text{div}, \Omega)$ with $(\nabla \cdot \sigma_h, 1)_K = (f, 1)_K$ for all $K \in T_h$ be arbitrary.

Then

$$\|\nabla (u - u_h)\|^2 \leq \sum_{K \in T_h} \left( \|\nabla u_h + \sigma_h\|_K + \frac{h_K}{\pi} \|f - \nabla \cdot \sigma_h\|_K \right)^2 + \sum_{K \in T_h} \|\nabla (u_h - s_h)\|^2_K.$$

**Proof (Spirit of Prager–Synge (1947)).**

- define $s \in H^1_0(\Omega)$ by
  $$(\nabla s, \nabla v) = (\nabla u_h, \nabla v) \quad \forall v \in H^1_0(\Omega)$$
- develop (Pythagoras)
  $$\|\nabla (u - u_h)\|^2 = \|\nabla (u - s)\|^2 + \|\nabla (s - u_h)\|^2$$
Proof (continuation).

- **projection:**
  \[
  \|\nabla(u - u_h)\|^2 = \sup_{\varphi \in H^1_0(\Omega); \|\nabla \varphi\| = 1} (\nabla(u - u_h), \nabla \varphi)^2 + \min_{v \in H^1_0(\Omega)} \|\nabla(v - u_h)\|^2
  \]
  dual norm of the residual
  distance of \(u_h\) to \(H^1_0(\Omega)\)

- **minimization upper bound:**
  \[
  \min_{v \in H^1_0(\Omega)} \|\nabla(v - u_h)\| \leq \|\nabla(u_h - s_h)\|
  \]

- **weak solution definition, equilibrated flux:**
  \[
  (\nabla(u - u_h), \nabla \varphi) = (f, \varphi) - (\nabla u_h, \nabla \varphi) = (f - \nabla \cdot \sigma_h, \varphi) - (\nabla u_h + \sigma_h, \nabla \varphi)
  \]

- **Cauchy–Schwarz and Poincaré inequalities:**
  \[
  -(\nabla u_h + \sigma_h, \nabla \varphi) \leq \sum_{K \in T_h} \|\nabla u_h + \sigma_h\|_K \|\nabla \varphi\|_K,
  \]
  \[
  (f - \nabla \cdot \sigma_h, \varphi) = \sum_{K \in T_h} (f - \nabla \cdot \sigma_h, \varphi - \varphi_K)_K \leq \sum_{K \in T_h} \frac{h_K}{\pi} \|f - \nabla \cdot \sigma_h\|_K \|\nabla \varphi\|_K
  \]
Potential and flux reconstruction

Ideally

$$\sigma_h := \arg \min_{v_h \in V_h, \nabla \cdot v_h = \Pi_{Q_h} f} \| \nabla u_h + v_h \|$$

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

- too expensive

Partition of unity

$$\sigma^a_h := \arg \min_{v_h \in V^a_h, \nabla \cdot v_h = ?} \| \psi_a \nabla u_h + v_h \|_{\omega_a}$$

$$s^a_h := \arg \min_{v_h \in V^a_h} \| \nabla (\psi_a u_h - v_h) \|_{\omega_a}$$

- $$\sigma_h := \sum_{a \in V_h} \sigma^a_h$$,  $$s_h := \sum_{a \in V_h} s^a_h$$
- local minimizations
Potential and flux reconstruction

Ideally

\[ \sigma_h := \arg \min_{v_h \in V_h, \nabla \cdot v_h = \Pi Q_h f} \| \nabla u_h + v_h \| \]

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

- too expensive

Partition of unity

\[ \sigma^a_h := \arg \min_{v_h \in V^a_h, \nabla \cdot v_h = ?} \| \psi^a \nabla u_h + v_h \|_{\omega^a} \]

\[ s^a_h := \arg \min_{v_h \in V^a_h} \| \nabla (\psi^a u_h - v_h) \|_{\omega^a} \]

- \[ \sigma_h := \sum_{a \in V_h} \sigma^a_h, \quad s_h := \sum_{a \in V_h} s^a_h \]

- local minimizations
Assumption A (Galerkin orthogonality)

There holds

\[(\nabla u_h, \nabla \psi_a)_{\omega_a} = (f, \psi_a)_{\omega_a} \quad \forall a \in V_h^{\text{int}}.\]

Definition (Construction of \(\sigma_h\), Destuynder and Métivet (1999) & Braess and Schöberl (2008))

Let Assumption A be satisfied. For each \(a \in V_h\), prescribe \(\varsigma^a_h \in V^a_h\) and \(\bar{r}^a_h \in Q^a_h\) by solving the local MFE problem

\[
(\varsigma^a_h, v_h)_{\omega_a} - (\bar{r}^a_h, \nabla \cdot v_h)_{\omega_a} = -((\psi_a \nabla u_h, v_h)_{\omega_a} \quad \forall v_h \in V^a_h,
\]

\[
(\nabla \cdot \varsigma^a_h, q_h)_{\omega_a} = (\psi_a f - \nabla \psi_a \cdot \nabla u_h, q_h)_{\omega_a} \quad \forall q_h \in Q^a_h,
\]

with \(V^a_h \times Q^a_h\) mixed finite element spaces (hom. Neumann BC for \(a \in V_h^{\text{int}}\), hom. Dirichlet BC on \(\partial \omega_a \cap \partial \Omega\) for \(a \in V_h^{\text{ext}}\)). Set

\[\sigma_h := \sum_{a \in V_h} \varsigma^a_h.\]
Practical flux reconstruction

**Assumption A (Galerkin orthogonality)**

There holds

\[
(\nabla u_h, \nabla \psi_a)_{\omega_a} = (f, \psi_a)_{\omega_a} \quad \forall a \in \mathcal{V}_h^{\text{int}}.
\]

**Definition (Construction of } \sigma_h \text{, Destuynder and Métivet (1999) & Braess and Schöberl (2008))**

Let Assumption A be satisfied. For each \( a \in \mathcal{V}_h \), prescribe \( \varsigma^a_h \in \mathcal{V}_h^a \) and \( \bar{r}^a_h \in \mathcal{Q}_h^a \) by solving the local MFE problem

\[
(\varsigma^a_h, v_h)_{\omega_a} - (\bar{r}^a_h, \nabla \cdot v_h)_{\omega_a} = -(\psi_a \nabla u_h, v_h)_{\omega_a} \quad \forall v_h \in \mathcal{V}_h^a,
\]

\[
(\nabla \cdot \varsigma^a_h, q_h)_{\omega_a} = (\psi_a f - \nabla \psi_a \cdot \nabla u_h, q_h)_{\omega_a} \quad \forall q_h \in \mathcal{Q}_h^a,
\]

with \( \mathcal{V}_h^a \times \mathcal{Q}_h^a \) mixed finite element spaces (hom. Neumann BC for \( a \in \mathcal{V}_h^{\text{int}} \), hom. Dirichlet BC on \( \partial \omega_a \cap \partial \Omega \) for \( a \in \mathcal{V}_h^{\text{ext}} \)). Set

\[
\sigma_h := \sum_{a \in \mathcal{V}_h} \varsigma^a_h.
\]
Assumption A (Galerkin orthogonality)

There holds

\[(\nabla u_h, \nabla \psi_a)_{\omega_a} = (f, \psi_a)_{\omega_a} \quad \forall a \in \mathcal{V}_h^{\text{int}}.\]

Definition (Construction of $\sigma_h$, Destuynder and Métivet (1999) & Braess and Schöberl (2008))

Let Assumption A be satisfied. For each $a \in \mathcal{V}_h$, prescribe $\varsigma_a^h \in \mathcal{V}_a^h$ and $\bar{r}_a^h \in \mathcal{Q}_a^h$ by solving the local MFE problem

\[
\begin{align*}
(\varsigma_a^h, v_h)_{\omega_a} - (\bar{r}_a^h, \nabla \cdot v_h)_{\omega_a} &= - (\psi_a \nabla u_h, v_h)_{\omega_a} \quad \forall v_h \in \mathcal{V}_a^h, \\
(\nabla \cdot \varsigma_a^h, q_h)_{\omega_a} &= (\psi_a f - \nabla \psi_a \cdot \nabla u_h, q_h)_{\omega_a} \quad \forall q_h \in \mathcal{Q}_a^h,
\end{align*}
\]

with $\mathcal{V}_a^h \times \mathcal{Q}_a^h$ mixed finite element spaces (hom. Neumann BC for $a \in \mathcal{V}_h^{\text{int}}$, hom. Dirichlet BC on $\partial \omega_a \cap \partial \Omega$ for $a \in \mathcal{V}_h^{\text{ext}}$). Set

$$\sigma_h := \sum_{a \in \mathcal{V}_h} \varsigma_a^h.$$
Definition (Construction of $s_h$)

For each $a \in \mathcal{V}_h$, prescribe $s_h^a \in V_h^a$ and $\bar{r}_h^a \in Q_h^a$ by solving the local MFE problem

$$
(s_h^a, v_h)_{\omega_a} - (\bar{r}_h^a, \nabla \cdot v_h)_{\omega_a} = -(R \frac{\pi}{2} \nabla (\psi_a u_h), v_h)_{\omega_a} \quad \forall v_h \in V_h^a,
$$

$$
(\nabla \cdot s_h^a, q_h)_{\omega_a} = (0, q_h)_{\omega_a} \quad \forall q_h \in Q_h^a,
$$

with $V_h^a \times Q_h^a$ mixed finite element spaces (hom. Neumann BC for all $a \in \mathcal{V}_h$). Set

$$
-R \frac{\pi}{2} \nabla s_h^a = s_h^a, \\
S_h^a = 0 \text{ on } \partial \omega_a, \\
S_h := \sum_{a \in \mathcal{V}_h} S_h^a.
$$

Remark

- The same problems, only RHS/BC different.

Alexandre Ern and Martin Vohralík
Adaptive inexact Newton methods
Practical potential reconstruction \((d = 2)\)

**Definition (Construction of \(s_h\))**

For each \(a \in V_h\), prescribe \(\varsigma^a_h \in V^a_h\) and \(\bar{r}^a_h \in Q^a_h\) by solving the local MFE problem

\[
(\varsigma^a_h, v_h)_{\omega_a} - (\bar{r}^a_h, \nabla \cdot v_h)_{\omega_a} = -(R \frac{\pi}{2} \nabla (\psi^a u_h), v_h)_{\omega_a} \quad \forall v_h \in V^a_h,
\]

\[
(\nabla \cdot \varsigma^a_h, q_h)_{\omega_a} = (0, q_h)_{\omega_a} \quad \forall q_h \in Q^a_h,
\]

with \(V^a_h \times Q^a_h\) mixed finite element spaces (hom. Neumann BC for all \(a \in V_h\)). Set

\[-R \frac{\pi}{2} \nabla s^a_h = \varsigma^a_h,\]

\[s^a_h = 0 \text{ on } \partial \omega_a,\]

\[s_h := \sum_{a \in V_h} s^a_h.\]

**Remark**

- The same problems, only RHS/BC different.

Alexandre Ern and Martin Vohralík

Adaptive inexact Newton methods
Practical potential reconstruction \((d = 2)\)

**Definition (Construction of \(s_h\))**

For each \(a \in \mathcal{V}_h\), prescribe \(s_h^a \in \mathcal{V}_h^a\) and \(r_h^a \in \mathcal{Q}_h^a\) by solving the local MFE problem

\[
(s_h^a, v_h)_{\omega_a} - (r_h^a, \nabla \cdot v_h)_{\omega_a} = -(R \frac{\pi}{2} \nabla (\psi_a u_h), v_h)_{\omega_a} \quad \forall v_h \in \mathcal{V}_h^a,
\]

\[
(\nabla \cdot s_h^a, q_h)_{\omega_a} = (0, q_h)_{\omega_a} \quad \forall q_h \in \mathcal{Q}_h^a,
\]

with \(\mathcal{V}_h^a \times \mathcal{Q}_h^a\) mixed finite element spaces (hom. Neumann BC for all \(a \in \mathcal{V}_h\)). Set

\[
-R \frac{\pi}{2} \nabla s_h^a = s_h^a,
\]

\(s_h^a = 0\) on \(\partial \omega_a\),

\(s_h := \sum_{a \in \mathcal{V}_h} s_h^a\).

**Remark**

- The same problems, only RHS/BC different.
Outline

1. Bibliography

2. **Laplace equation**
   - A guaranteed a posteriori error estimate
   - **Polynomial-degree-robust local efficiency**
   - Application and numerical results

3. **Quasi-linear elliptic problems**
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results

4. **Two-phase flow in porous media**
   - A guaranteed a posteriori error estimate
   - Applications and numerical results

5. **Conclusions and future directions**
Theorem (Continuous efficiency, Carstensen & Funken (1999), Braess, Pillwein, and Schöberl (2009))

Let $u$ be the weak solution and let $u_h \in H^1(\mathcal{T}_h)$ be arbitrary. Let $a \in \mathcal{V}_h$ and let $r_a \in H^1_*(\omega_a)$ solve

$$(\nabla r_a, \nabla v)_{\omega_a} = - (\psi_a \nabla u_h, \nabla v)_{\omega_a} + (\psi_a f - \nabla \psi_a \cdot \nabla u_h, v)_{\omega_a} \quad \forall v \in H^1_*(\omega_a)$$

with $H^1_*(\omega_a) := \{ v \in H^1(\omega_a); (v, 1)_{\omega_a} = 0 / v = 0$ on $\partial \omega_a \cap \partial \Omega \}$. Then there exists a constant $C_{\text{cont,PF}} > 0$ only depending on the shape-regularity parameter $\kappa_T$ such that

$$\| \nabla r_a \|_{\omega_a} \leq C_{\text{cont,PF}} \| \nabla (u - u_h) \|_{\omega_a}.$$
Potential reconstruction \((d = 2)\)

**Assumption B (Weak continuity)**

There holds

\[
\langle [u_h], 1 \rangle_e = 0 \quad \forall e \in \mathcal{E}_h.
\]

**Theorem (Continuous efficiency)**

Let \(u\) be the weak solution and let \(u_h \in H^1(\mathcal{T}_h)\) satisfying Assumption B be arbitrary. Let \(a \in \mathcal{V}_h\) and let \(r_a \in H^1_*(\omega_a)\) solve

\[
(\nabla r_a, \nabla v)_{\omega_a} = -\left(\frac{R}{2}, \nabla (\psi a u_h), \nabla v\right)_{\omega_a} + (0, v)_{\omega_a} \quad \forall v \in H^1_*(\omega_a)
\]

with \(H^1_*(\omega_a) := \{ v \in H^1(\omega_a); (v, 1)_{\omega_a} = 0 \}\). Then there exists a constant \(C_{\text{cont, bPF}} > 0\) only depending on the shape-regularity parameter \(\kappa_{\mathcal{T}}\) such that

\[
\|\nabla r_a\|_{\omega_a} \leq C_{\text{cont, bPF}} \|\nabla (u - u_h)\|_{\omega_a}.
\]
Potential reconstruction ($d = 2$)

### Assumption B (Weak continuity)

There holds

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

### Theorem (Continuous efficiency)

Let $u$ be the weak solution and let $u_h \in H^1(\mathcal{T}_h)$ satisfying Assumption B be arbitrary. Let $a \in \mathcal{V}_h$ and let $r_a \in H^1_*(\omega_a)$ solve

$$(\nabla r_a, \nabla v)_{\omega_a} = -(R \frac{\pi}{2} \nabla (\psi_a u_h), \nabla v)_{\omega_a} + (0, v)_{\omega_a} \quad \forall v \in H^1_*(\omega_a)$$

with $H^1_*(\omega_a) := \{ v \in H^1(\omega_a); (v, 1)_{\omega_a} = 0 \}$. Then there exists a constant $C_{\text{cont,bPF}} > 0$ only depending on the shape-regularity parameter $\kappa_T$ such that

$$\| \nabla r_a \|_{\omega_a} \leq C_{\text{cont,bPF}} \| \nabla (u - u_h) \|_{\omega_a}.$$
Theorem (MFE stability, Braess, Pillwein, and Schöberl (2009))

Let $u$ be the weak solution and let $u_h$ and $f$ be piecewise polynomial. Consider corresponding polynomial degree MFE reconstructions. Then there exists a constant $C_{st} > 0$ only depending on the shape-regularity parameter $\kappa T$ such that

$$
\|\varsigma_h^a + \tau_h^a\|_{\omega_a} \leq C_{st} \|\nabla r_a\|_{\omega_a},
$$

with $\tau_h^a = \psi_a \nabla u_h$ for the flux reconstruction and $\tau_h^a = R_{\frac{\pi}{2}} \nabla (\psi_a u_h)$ for the potential reconstruction.
Theorem (Polynomial-degree-robust efficiency)

Let \( u \) be the weak solution and let \( u_h \) and \( f \) be piecewise polynomial. Let \( u_h \) satisfy Assumptions A and B. Then, for corresponding polynomial degree MFE reconstructions,

\[
\| \nabla u_h + \sigma_h \|_K \leq C_{st} C_{cont,PF} \sum_{a \in V_K} \| \nabla (u - u_h) \|_{\omega_a},
\]

\[
\| \nabla (u_h - s_h) \|_K \leq C_{st} C_{cont,PF} \sum_{a \in V_K} \| \nabla (u - u_h) \|_{\omega_a}.
\]

Remarks

- \( C_{st} \) can be bounded by solving the local Neumann problems by a conforming FEs
- maximal overestimation factor guaranteed
Theorem (Polynomial-degree-robust efficiency)

Let \( u \) be the weak solution and let \( u_h \) and \( f \) be piecewise polynomial. Let \( u_h \) satisfy Assumptions A and B. Then, for corresponding polynomial degree MFE reconstructions,

\[
\| \nabla u_h + \sigma_h \|_K \leq C_{\text{st}} C_{\text{cont,PF}} \sum_{a \in V_K} \| \nabla (u - u_h) \|_{\omega_a},
\]

\[
\| \nabla (u_h - s_h) \|_K \leq C_{\text{st}} C_{\text{cont,bPF}} \sum_{a \in V_K} \| \nabla (u - u_h) \|_{\omega_a}.
\]

Remarks

- \( C_{\text{st}} \) can be bounded by solving the local Neumann problems by a conforming FEs
- maximal overestimation factor guaranteed
Outline

1. Bibliography
2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results
3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results
4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results
5. Conclusions and future directions
Conforming finite elements

Find $u_h \in V_h$ such that

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

- $V_h := \mathbb{P}_p(T_h) \cap H^1_0(\Omega), \ p \geq 1$
- Assumption A: take $v_h = \psi_a$
- $V_h \subset H^1_0(\Omega): s_h := u_h$, no need for Assumption B
Conforming finite elements

Find $u_h \in V_h$ such that

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

- $V_h := \mathbb{P}_p(T_h) \cap H^1_0(\Omega), \ p \geq 1$
- **Assumption A**: take $v_h = \psi_a$
- $V_h \subset H^1_0(\Omega)$: $s_h := u_h$, no need for **Assumption B**
Nonconforming finite elements

Find \( u_h \in V_h \) such that

\[
(\nabla u_h, \nabla v_h) = (f, v_h) \quad \forall v_h \in V_h.
\]

- \( V_h := \mathbb{P}_p(T_h), \ p \geq 1, \ v_h \in V_h \) satisfy

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

- Assumption A: take \( v_h = \psi_a \)
- Assumption B: building requirement for the space \( V_h \)
Nonconforming finite elements

Find $u_h \in V_h$ such that

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

$V_h := \mathbb{P}_p(T_h)$, $p \geq 1$, $v_h \in V_h$ satisfy

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

- **Assumption A**: take $v_h = \psi_a$
- **Assumption B**: building requirement for the space $V_h$
Discontinuous Galerkin finite elements

Find \( u_h \in V_h \) such that

\[
\sum_{K \in T_h} (\nabla u_h, \nabla v_h)_K - \sum_{e \in \mathcal{E}_h} \left\{ \langle \{\nabla u_h\} \cdot n_e, [v_h] \rangle_e + \theta \langle \{\nabla v_h\} \cdot n_e, [u_h] \rangle_e \right\} \\
+ \sum_{e \in \mathcal{E}_h} \langle \alpha h_e^{-1} [u_h], [v_h] \rangle_e = (f, v_h) \quad \forall v_h \in V_h
\]

- \( V_h := \mathbb{P}_p(T_h), p \geq 1 \)
- Assumption A: take \( v_h = \psi_a \) for \( \theta = 0 \), otherwise consider the discrete gradient
- Assumption B not satisfied, but an easy adjustment by including the jump terms in the norm
Discontinuous Galerkin finite elements

Find $u_h \in V_h$ such that

$$\sum_{K \in \mathcal{T}_h} (\nabla u_h, \nabla v_h)_K - \sum_{e \in \mathcal{E}_h} \left\{ \langle \{\nabla u_h\} \cdot n_e, [v_h] \rangle_e + \theta \langle \{\nabla v_h\} \cdot n_e, [u_h] \rangle_e \right\}$$

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

- $V_h := \mathbb{P}_p(\mathcal{T}_h), \ p \geq 1$
- **Assumption A**: take $v_h = \psi_a$ for $\theta = 0$, otherwise consider the discrete gradient
- **Assumption B** not satisfied, but an easy adjustment by including the jump terms in the norm
Mixed finite elements

Find a couple \((\sigma_h, \tilde{u}_h) \in V_h \times Q_h\) such that

\[
\begin{align*}
(\sigma_h, v_h) - (\tilde{u}_h, \nabla \cdot v_h) &= 0 & \forall v_h \in V_h, \\
(\nabla \cdot \sigma_h, q_h) &= (f, q_h) & \forall q_h \in Q_h.
\end{align*}
\]

- postprocessed solution \(u_h \in V_h, V_h := \mathbb{P}_p(T_h), p \geq 1,\)
  \(v_h \in V_h\) satisfy
  \[
  \langle [v_h], q_h \rangle_e = 0 & \forall q_h \in \mathbb{P}_{p'}(e), \forall e \in E_h
  \]

- Assumption A: no need for flux reconstruction, \(\sigma_h\) comes from the discretization

- Assumption B satisfied, building requirement for the space \(V_h\)
Mixed finite elements

Find a couple \((\sigma_h, \bar{u}_h) \in V_h \times Q_h\) such that

\[
(\sigma_h, v_h) - (\bar{u}_h, \nabla \cdot v_h) = 0 \quad \forall v_h \in V_h,
\]
\[
(\nabla \cdot \sigma_h, q_h) = (f, q_h) \quad \forall q_h \in Q_h.
\]

• postprocessed solution \(u_h \in V_h, V_h := \mathbb{P}_p(T_h), p \geq 1, v_h \in V_h\) satisfy

\[
\langle [v_h], q_h \rangle_e = 0 \quad \forall q_h \in \mathbb{P}_{p'}(e), \forall e \in \mathcal{E}_h
\]

• Assumption A: no need for flux reconstruction, \(\sigma_h\) comes from the discretization

• Assumption B satisfied, building requirement for the space \(V_h\)
Model problem

\[-\Delta u = f \quad \text{in} \quad \Omega := ]0, 1[ \times ]0, 1[,\]
\[u = u_D \quad \text{on} \quad \partial \Omega\]

Exact solution

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

Discretization
incomplete interior penalty discontinuous Galerkin method
Model problem

\[-\Delta u = f \quad \text{in } \Omega := ]0, 1[ \times ]0, 1[ ,
\]
\[u = u_D \quad \text{on } \partial \Omega\]

Exact solution

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

Discretization

incomplete interior penalty discontinuous Galerkin method
Numerics: discontinuous Galerkin

Model problem

$$-\Delta u = f \quad \text{in } \Omega := ]0,1[ \times ]0,1[,$$

$$u = u_D \quad \text{on } \partial \Omega$$

Exact solution

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

$$c_1 = -e^{-\alpha}, \quad c_2 = -1 - c_1, \quad \alpha = 10$$

Discretization

incomplete interior penalty discontinuous Galerkin method
### Biblio Laplace Nonlinear Laplace Two-phase Conclusions

Estimate Efficiency Application and numerical results

<table>
<thead>
<tr>
<th>$h$</th>
<th>$p$</th>
<th>$| u - u_h |$</th>
<th>$| \nabla (u - u_h) |$</th>
<th>$| \nabla u_h + \sigma_h |$</th>
<th>$\frac{h K}{n}$</th>
<th>$| f - \nabla \cdot \sigma_h |$</th>
<th>$| \nabla (u_h - s_h) |$</th>
<th>$\eta$</th>
<th>$l_{\text{eff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3E-01</td>
<td>1</td>
<td>1.39E-02</td>
<td>4.98E-01</td>
<td>5.32E-01</td>
<td>4.34E-02</td>
<td>3.52E-02</td>
<td>5.71E-01</td>
<td>1.15</td>
<td></td>
</tr>
<tr>
<td>6.3E-02</td>
<td>(EOC)</td>
<td>(1.85)</td>
<td>(0.90)</td>
<td>(0.93)</td>
<td>(2.79)</td>
<td>(0.85)</td>
<td>(1.00)</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td>3.1E-02</td>
<td>(EOC)</td>
<td>(1.96)</td>
<td>(0.97)</td>
<td>(0.98)</td>
<td>(2.94)</td>
<td>(0.96)</td>
<td>(1.00)</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>1.6E-02</td>
<td>(EOC)</td>
<td>(1.99)</td>
<td>(0.99)</td>
<td>(0.99)</td>
<td>(1.01)</td>
<td>(1.00)</td>
<td>1.04</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$h$</th>
<th>$p$</th>
<th>$| u - u_h |$</th>
<th>$| \nabla (u - u_h) |$</th>
<th>$| \nabla u_h + \sigma_h |$</th>
<th>$\frac{h K}{n}$</th>
<th>$| f - \nabla \cdot \sigma_h |$</th>
<th>$| \nabla (u_h - s_h) |$</th>
<th>$\eta$</th>
<th>$l_{\text{eff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3E-01</td>
<td>2</td>
<td>2.08E-03</td>
<td>9.38E-02</td>
<td>9.48E-02</td>
<td>8.83E-03</td>
<td>1.27E-02</td>
<td>1.03E-01</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>6.3E-02</td>
<td>(EOC)</td>
<td>(2.37)</td>
<td>(1.84)</td>
<td>(1.85)</td>
<td>(3.79)</td>
<td>(1.65)</td>
<td>(1.93)</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>3.1E-02</td>
<td>(EOC)</td>
<td>(2.28)</td>
<td>(1.95)</td>
<td>(1.96)</td>
<td>(3.95)</td>
<td>(1.77)</td>
<td>(1.97)</td>
<td>1.02</td>
<td></td>
</tr>
<tr>
<td>1.6E-02</td>
<td>(EOC)</td>
<td>(2.16)</td>
<td>(1.98)</td>
<td>(1.99)</td>
<td>(1.88)</td>
<td>(1.99)</td>
<td>1.02</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$h$</th>
<th>$p$</th>
<th>$| u - u_h |$</th>
<th>$| \nabla (u - u_h) |$</th>
<th>$| \nabla u_h + \sigma_h |$</th>
<th>$\frac{h K}{n}$</th>
<th>$| f - \nabla \cdot \sigma_h |$</th>
<th>$| \nabla (u_h - s_h) |$</th>
<th>$\eta$</th>
<th>$l_{\text{eff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3E-01</td>
<td>3</td>
<td>2.29E-04</td>
<td>1.55E-02</td>
<td>1.51E-02</td>
<td>1.38E-03</td>
<td>2.36E-03</td>
<td>1.66E-02</td>
<td>1.07</td>
<td></td>
</tr>
<tr>
<td>6.3E-02</td>
<td>(EOC)</td>
<td>(3.80)</td>
<td>(2.81)</td>
<td>(2.81)</td>
<td>(4.79)</td>
<td>(2.62)</td>
<td>(2.89)</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>3.1E-02</td>
<td>(EOC)</td>
<td>(3.94)</td>
<td>(2.95)</td>
<td>(2.94)</td>
<td>(4.94)</td>
<td>(2.87)</td>
<td>(2.97)</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>1.6E-02</td>
<td>(EOC)</td>
<td>(3.99)</td>
<td>(2.99)</td>
<td>(2.98)</td>
<td>(4.99)</td>
<td>(2.99)</td>
<td>(2.99)</td>
<td>1.01</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$h$</th>
<th>$p$</th>
<th>$| u - u_h |$</th>
<th>$| \nabla (u - u_h) |$</th>
<th>$| \nabla u_h + \sigma_h |$</th>
<th>$\frac{h K}{n}$</th>
<th>$| f - \nabla \cdot \sigma_h |$</th>
<th>$| \nabla (u_h - s_h) |$</th>
<th>$\eta$</th>
<th>$l_{\text{eff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3E-01</td>
<td>4</td>
<td>2.64E-05</td>
<td>2.28E-03</td>
<td>2.17E-03</td>
<td>1.69E-04</td>
<td>3.46E-04</td>
<td>2.37E-03</td>
<td>1.04</td>
<td></td>
</tr>
<tr>
<td>6.3E-02</td>
<td>(EOC)</td>
<td>(4.61)</td>
<td>(3.80)</td>
<td>(3.79)</td>
<td>(5.79)</td>
<td>(3.41)</td>
<td>(3.86)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>3.1E-02</td>
<td>(EOC)</td>
<td>(4.52)</td>
<td>(3.95)</td>
<td>(3.94)</td>
<td>(5.94)</td>
<td>(3.73)</td>
<td>(3.95)</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>1.6E-02</td>
<td>(EOC)</td>
<td>(4.52)</td>
<td>(3.95)</td>
<td>(4.52)</td>
<td>(5.94)</td>
<td>(3.73)</td>
<td>(3.95)</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$h$</th>
<th>$p$</th>
<th>$| u - u_h |$</th>
<th>$| \nabla (u - u_h) |$</th>
<th>$| \nabla u_h + \sigma_h |$</th>
<th>$\frac{h K}{n}$</th>
<th>$| f - \nabla \cdot \sigma_h |$</th>
<th>$| \nabla (u_h - s_h) |$</th>
<th>$\eta$</th>
<th>$l_{\text{eff}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3E-01</td>
<td>5</td>
<td>2.69E-06</td>
<td>2.78E-04</td>
<td>2.60E-04</td>
<td>1.69E-05</td>
<td>4.39E-05</td>
<td>2.81E-04</td>
<td>1.01</td>
<td></td>
</tr>
<tr>
<td>6.3E-02</td>
<td>(EOC)</td>
<td>(5.77)</td>
<td>(4.80)</td>
<td>(4.78)</td>
<td>(6.78)</td>
<td>(4.53)</td>
<td>(4.84)</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>3.1E-02</td>
<td>(EOC)</td>
<td>(6.02)</td>
<td>(4.95)</td>
<td>(4.93)</td>
<td>(6.44)</td>
<td>(4.88)</td>
<td>(4.94)</td>
<td>0.99</td>
<td></td>
</tr>
</tbody>
</table>
Outline

1. Bibliography

2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results

3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results

4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results

5. Conclusions and future directions
Quasi-linear elliptic problem

\[-\nabla \cdot \sigma(u, \nabla u) = f \quad \text{in } \Omega,\]
\[u = 0 \quad \text{on } \partial \Omega\]

- quasi-linear diffusion problem
  \[\sigma(v, \xi) = A(v)\xi \quad \forall (v, \xi) \in \mathbb{R} \times \mathbb{R}^d\]

- Leray–Lions problem
  \[\sigma(v, \xi) = A(\xi)\xi \quad \forall \xi \in \mathbb{R}^d\]

- \(p > 1, \ q := \frac{p}{p-1}, \ f \in L^q(\Omega)\)

Example
\(p\)-Laplacian: Leray–Lions setting with \(A(\xi) = |\xi|^{p-2}\)

Nonlinear operator \(A : V := \mathcal{W}^{1,p}_0(\Omega) \to V'\)

\[\langle A(u), v \rangle_{V', V} := (\sigma(u, \nabla u), \nabla v)\]

Weak formulation
Find \(u \in V\) such that
\[A(u) = f \text{ in } V'\]
Quasi-linear elliptic problem

\[ -\nabla \cdot \sigma(u, \nabla u) = f \quad \text{in } \Omega, \]
\[ u = 0 \quad \text{on } \partial \Omega \]

- quasi-linear diffusion problem

\[ \sigma(v, \xi) = A(v)\xi \quad \forall (v, \xi) \in \mathbb{R} \times \mathbb{R}^d \]

- Leray–Lions problem

\[ \sigma(v, \xi) = A(\xi)\xi \quad \forall \xi \in \mathbb{R}^d \]

- \( p > 1, \; q := \frac{p}{p-1}, \; f \in L^q(\Omega) \)

Example

\( p \)-Laplacian: Leray–Lions setting with \( A(\xi) = |\xi|^{p-2}I \)

Nonlinear operator \( A : V := W_0^{1,p}(\Omega) \rightarrow V' \)

\[ \langle A(u), v \rangle_{V', V} := \langle \sigma(u, \nabla u), \nabla v \rangle \]

Weak formulation

Find \( u \in V \) such that \( A(u) = f \) in \( V' \)
Quasi-linear elliptic problem

\[-\nabla \cdot \sigma(u, \nabla u) = f \quad \text{in } \Omega, \]
\[u = 0 \quad \text{on } \partial \Omega\]

- quasi-linear diffusion problem

\[\sigma(v, \xi) = A(v)\xi \quad \forall (v, \xi) \in \mathbb{R} \times \mathbb{R}^d\]

- Leray–Lions problem

\[\sigma(v, \xi) = A(\xi)\xi \quad \forall \xi \in \mathbb{R}^d\]

\[p > 1, \quad q := \frac{p}{p-1}, \quad f \in L^q(\Omega)\]

Example

\- $p$-Laplacian: Leray–Lions setting with $A(\xi) = |\xi|^{p-2}\xi$

Nonlinear operator $A : V := W^{1,p}_0(\Omega) \to V'$

\[\langle A(u), v \rangle_{V',V} := (\sigma(u, \nabla u), \nabla v)\]

Weak formulation

Find $u \in V$ such that $A(u) = f$ in $V'$.
Quasi-linear elliptic problem

Quasi-linear elliptic problem

\[-\nabla \cdot \sigma(u, \nabla u) = f \quad \text{in } \Omega,\]
\[u = 0 \quad \text{on } \partial \Omega\]

- quasi-linear diffusion problem

\[\sigma(v, \xi) = A(v)\xi \quad \forall (v, \xi) \in \mathbb{R} \times \mathbb{R}^d\]

- Leray–Lions problem

\[\sigma(v, \xi) = A(\xi)\xi \quad \forall \xi \in \mathbb{R}^d\]

- \(p > 1, \quad q := \frac{p}{p-1}, \quad f \in L^q(\Omega)\)

Example

\(p\)-Laplacian: Leray–Lions setting with \(A(\xi) = |\xi|^{p-2}I\)

Nonlinear operator \(A : V := W_0^{1,p}(\Omega) \rightarrow V'\)

\[\langle A(u), v \rangle_{V', V} := (\sigma(u, \nabla u), \nabla v)\]

Weak formulation

Find \(u \in V\) such that \(A(u) = f \text{ in } V'\)
Quasi-linear elliptic problem

Quasi-linear elliptic problem

\[-\nabla \cdot \sigma(u, \nabla u) = f \quad \text{in } \Omega,
\]
\[u = 0 \quad \text{on } \partial \Omega\]

- quasi-linear diffusion problem

\[\sigma(v, \xi) = A(v)\xi \quad \forall (v, \xi) \in \mathbb{R} \times \mathbb{R}^d\]

- Leray–Lions problem

\[\sigma(v, \xi) = A(\xi)\xi \quad \forall \xi \in \mathbb{R}^d\]

\[p > 1, \quad q := \frac{p}{p-1}, \quad f \in L^q(\Omega)\]

Example

$p$-Laplacian: Leray–Lions setting with $A(\xi) = |\xi|^{p-2}I$

Nonlinear operator $A : V := \mathcal{W}^{1,p}_0(\Omega) \rightarrow V'$

\[\langle A(u), v \rangle_{V', V} := (\sigma(u, \nabla u), \nabla v)\]

Weak formulation
Find $u \in V$ such that

\[A(u) = f \quad \text{in } V'\]
Quasi-linear elliptic problem

Quasi-linear elliptic problem

\[-\nabla \cdot \sigma(u, \nabla u) = f \quad \text{in } \Omega,\]
\[u = 0 \quad \text{on } \partial \Omega\]

- quasi-linear diffusion problem

\[\sigma(v, \xi) = A(v) \xi \quad \forall (v, \xi) \in \mathbb{R} \times \mathbb{R}^d\]

- Leray–Lions problem

\[\sigma(v, \xi) = A(\xi) \xi \quad \forall \xi \in \mathbb{R}^d\]

- \(p > 1, \ q := \frac{p}{p-1}, \ f \in L^q(\Omega)\)

Example

\(p\)-Laplacian: Leray–Lions setting with \(A(\xi) = |\xi|^{p-2}I\)

Nonlinear operator \(A : V := W_0^{1,p}(\Omega) \to V'\)

\[\langle A(u), v \rangle_{V', V} := (\sigma(u, \nabla u), \nabla v)\]

Weak formulation

Find \(u \in V\) such that \(A(u) = f\) in \(V'\)
Quasi-linear elliptic problem

Quasi-linear elliptic problem

\[-\nabla \cdot \sigma(u, \nabla u) = f \quad \text{in } \Omega,\]
\[u = 0 \quad \text{on } \partial \Omega\]

- quasi-linear diffusion problem

\[\sigma(v, \xi) = A(v)\xi \quad \forall (v, \xi) \in \mathbb{R} \times \mathbb{R}^d\]

- Leray–Lions problem

\[\sigma(v, \xi) = A(\xi)\xi \quad \forall \xi \in \mathbb{R}^d\]

- \(p > 1, \; q := \frac{p}{p-1}, \; f \in L^q(\Omega)\)

Example

\(p\)-Laplacian: Leray–Lions setting with \(A(\xi) = |\xi|^{p-2}\)

Nonlinear operator \(A : V := \mathcal{W}_{0}^{1,p}(\Omega) \rightarrow V'\)

\[\langle A(u), v \rangle_{V', V} := (\sigma(u, \nabla u), \nabla v)\]

Weak formulation

Find \(u \in V\) such that \(A(u) = f \; \text{in} \; V'\)
Approximate solution and error measure

Approximate solution

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

Error measure

$$
\mathcal{J}_u(u_h^{k,i}) := \sup_{\varphi \in V; \|\nabla \varphi\|_p = 1} (\sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i}), \nabla \varphi) + \mathcal{J}_{u,NC}(u_h^{k,i})
$$

$$
\mathcal{J}_{u,NC}(u_h^{k,i}) := \left\{ \sum_{K \in \mathcal{T}_h} \sum_{e \in \mathcal{E}_K} h_e^{1-q} \|u - u_h^{k,i}\|_{q,e}^q \right\}^{1/q}
$$

- dual norm of the residual + nonconformity
- there holds $\mathcal{J}_u(u_h^{k,i}) = 0$ if and only if $u = u_h^{k,i}$
- link: strong difference of the fluxes + nonconformity

$$
\mathcal{J}_u(u_h^{k,i}) \leq \mathcal{J}_u^{up}(u_h^{k,i}) := \|\sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i})\|_q + \mathcal{J}_{u,NC}(u_h^{k,i})
$$
Approximate solution and error measure

Approximate solution

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

Error measure

\[
\mathcal{J}_u(u_h^{k,i}) := \sup_{\varphi \in V; \|\nabla \varphi\|_p = 1} (\sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i}), \nabla \varphi) + \mathcal{J}_{u,\text{NC}}(u_h^{k,i})
\]

\[
\mathcal{J}_{u,\text{NC}}(u_h^{k,i}) := \left\{ \sum_{K \in \mathcal{T}_h} \sum_{e \in \mathcal{E}_K} h_e^{1-q} \|u - u_h^{k,i}\|_q, e \right\}^{1/q}
\]

- Dual norm of the residual + nonconformity
- there holds \( \mathcal{J}_u(u_h^{k,i}) = 0 \) if and only if \( u = u_h^{k,i} \)
- link: strong difference of the fluxes + nonconformity

\[
\mathcal{J}_u(u_h^{k,i}) \leq \mathcal{J}_{u,\text{up}}(u_h^{k,i}) := \|\sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i})\|_q + \mathcal{J}_{u,\text{NC}}(u_h^{k,i})
\]
Approximate solution and error measure

Approximate solution

- \( u_{h}^{k,i} \in V(\mathcal{T}_h) \not\subset V \), \( u_{h}^{k,i} \) not necessarily in \( V \)
- \( V(\mathcal{T}_h) := \{ v \in L^p(\Omega), \ v|_K \in W^{1,p}(K) \ \forall K \in \mathcal{T}_h \} \)

Error measure

\[
J_u(u_{h}^{k,i}) := \sup_{\varphi \in V; \|\nabla \varphi\|_p=1} (\sigma(u,\nabla u) - \sigma(u_{h}^{k,i},\nabla u_{h}^{k,i}), \nabla \varphi) + J_{u,NC}(u_{h}^{k,i})
\]

\[
J_{u,NC}(u_{h}^{k,i}) := \left\{ \sum_{K \in \mathcal{T}_h} \sum_{e \in \mathcal{E}_K} h_e^{1-q} \|\| u - u_{h}^{k,i} \|\|_q, e \right\}^{1/q}
\]

- dual norm of the residual + nonconformity
- there holds \( J_u(u_{h}^{k,i}) = 0 \) if and only if \( u = u_{h}^{k,i} \)
- link: strong difference of the fluxes + nonconformity

\[
J_u(u_{h}^{k,i}) \leq J_{u,up}(u_{h}^{k,i}) := \|\sigma(u,\nabla u) - \sigma(u_{h}^{k,i},\nabla u_{h}^{k,i})\|_q + J_{u,NC}(u_{h}^{k,i})
\]
Approximate solution and error measure

Approximate solution

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

Error measure

\[
\mathcal{J}_u(u_h^{k,i}) := \sup_{\varphi \in V; \|\nabla \varphi\|_p = 1} (\sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i}), \nabla \varphi) + \mathcal{J}_{u,NC}(u_h^{k,i})
\]

\[
\mathcal{J}_{u,NC}(u_h^{k,i}) := \left\{ \sum_{K \in \mathcal{T}_h} \sum_{e \in \mathcal{E}_K} h_e^{1-q} \| [u - u_h^{k,i}] \|_{q,e} \right\}^{1/q}
\]

- dual norm of the residual + nonconformity
- there holds \( \mathcal{J}_u(u_h^{k,i}) = 0 \) if and only if \( u = u_h^{k,i} \)
- link: strong difference of the fluxes + nonconformity

\[
\mathcal{J}_u(u_h^{k,i}) \leq \mathcal{J}^{up}_u(u_h^{k,i}) := \| \sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i}) \|_q + \mathcal{J}_{u,NC}(u_h^{k,i})
\]
Approximate solution and error measure

Approximate solution

- \( u_h^{k,i} \in V(\mathcal{T}_h) \not\subset V \), \( u_h^{k,i} \) not necessarily in \( V \)
- \( V(\mathcal{T}_h) := \{ v \in L^p(\Omega), \ v|_K \in W^{1,p}(K) \ \forall K \in \mathcal{T}_h \} \)

Error measure

\[
\mathcal{J}_u(u_h^{k,i}) := \sup_{\varphi \in V; \|\nabla \varphi\|_p = 1} (\sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i}), \nabla \varphi) + \mathcal{J}_{u,NC}(u_h^{k,i})
\]

\[
\mathcal{J}_{u,NC}(u_h^{k,i}) := \left\{ \sum_{K \in \mathcal{T}_h} \sum_{e \in \mathcal{E}_K} h_e^{1-q} \| [u - u_h^{k,i}] \|_{q,e} \right\}^{1/q}
\]

- dual norm of the residual + nonconformity
- there holds \( \mathcal{J}_u(u_h^{k,i}) = 0 \) if and only if \( u = u_h^{k,i} \)
- link: strong difference of the fluxes + nonconformity

\[
\mathcal{J}_u(u_h^{k,i}) \leq \mathcal{J}_u^{up}(u_h^{k,i}) := \| \sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i}) \|_q + \mathcal{J}_{u,NC}(u_h^{k,i})
\]
Approximate solution

- \( u_h^{k,i} \in V(\mathcal{T}_h) \not\subset V \), \( u_h^{k,i} \) not necessarily in \( V \)
- \( V(\mathcal{T}_h) := \{ v \in L^p(\Omega), v|_K \in W^{1,p}(K) \ \forall K \in \mathcal{T}_h \} \)

Error measure

\[
\mathcal{J}_u(u_h^{k,i}) := \sup_{\varphi \in V; \|\nabla \varphi\|_p=1} (\sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i}), \nabla \varphi) + \mathcal{J}_{u,NC}(u_h^{k,i})
\]

\[
\mathcal{J}_{u,NC}(u_h^{k,i}) := \left\{ \sum_{K \in \mathcal{T}_h} \sum_{e \in \mathcal{E}_K} h_e^{1-q} \|u - u_h^{k,i}\|_{q,e} \right\}^{1/q}
\]

- dual norm of the residual + nonconformity
- there holds \( \mathcal{J}_u(u_h^{k,i}) = 0 \) if and only if \( u = u_h^{k,i} \)
- link: strong difference of the fluxes + nonconformity

\[
\mathcal{J}_u(u_h^{k,i}) \leq \mathcal{J}_u^{up}(u_h^{k,i}) := \|\sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i})\|_q + \mathcal{J}_{u,NC}(u_h^{k,i})
\]
Outline

1. Bibliography
2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results
3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results
4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results
5. Conclusions and future directions

Alexandre Ern and Martin Vohralík

Adaptive inexact Newton methods
Assumption A (Total quasi-equilibrated flux reconstruction)

There exists a flux reconstruction \( \sigma_{h}^{k,i} \in H^q(\text{div},\Omega) \) and an algebraic remainder \( \rho_{h}^{k,i} \in L^q(\Omega) \) such that

\[
\nabla \cdot \sigma_{h}^{k,i} = f_h - \rho_{h}^{k,i},
\]

with the data approximation \( f_h \) s.t. \( (f_h,1)_K = (f,1)_K \ \forall K \in T_h \).

Theorem (A guaranteed a posteriori error estimate)

Let

- \( u \in V \) be the weak solution,
- \( u_{h}^{k,i} \in V(T_h) \) be arbitrary,
- Assumption A hold.

Then there holds

\[
J_u(u_{h}^{k,i}) \leq \eta_{h}^{k,i},
\]

where \( \eta_{h}^{k,i} \) is fully computable from \( u_{h}^{k,i}, \sigma_{h}^{k,i}, \) and \( \rho_{h}^{k,i} \).
**A posteriori error estimate**

**Assumption A (Total quasi-equilibrated flux reconstruction)**

There exists a flux reconstruction \( \sigma_{h}^{k,i} \in H^{q}(\text{div}, \Omega) \) and an algebraic remainder \( \rho_{h}^{k,i} \in L^{q}(\Omega) \) such that

\[
\nabla \cdot \sigma_{h}^{k,i} = f_{h} - \rho_{h}^{k,i},
\]

with the data approximation \( f_{h} \) s.t. \( (f_{h}, 1)_{K} = (f, 1)_{K} \) \( \forall K \in T_{h} \).

**Theorem (A guaranteed a posteriori error estimate)**

Let

- \( u \in V \) be the weak solution,
- \( u_{h}^{k,i} \in V(T_{h}) \) be arbitrary,
- **Assumption A** hold.

Then there holds

\[
J_{u}(u_{h}^{k,i}) \leq \eta_{h}^{k,i},
\]

where \( \eta_{h}^{k,i} \) is fully computable from \( u_{h}^{k,i}, \sigma_{h}^{k,i}, \) and \( \rho_{h}^{k,i} \).
A posteriori error estimate

Assumption A (Total quasi-equilibrated flux reconstruction)

There exists a flux reconstruction \( \sigma_h^{k,i} \in H^q(\text{div}, \Omega) \) and an algebraic remainder \( \rho_h^{k,i} \in L^q(\Omega) \) such that

\[
\nabla \cdot \sigma_h^{k,i} = f - \rho_h^{k,i},
\]

with the data approximation \( f_h \) s.t. \( (f_h, 1)_K = (f, 1)_K \) \( \forall K \in T_h \).

Theorem (A guaranteed a posteriori error estimate)

Let

- \( u \in V \) be the weak solution,
- \( u_h^{k,i} \in V(T_h) \) be arbitrary,
- Assumption A hold.

Then there holds

\[
\mathcal{J}_u(u_h^{k,i}) \leq \eta_h^{k,i},
\]

where \( \eta_h^{k,i} \) is fully computable from \( u_h^{k,i} \), \( \sigma_h^{k,i} \), and \( \rho_h^{k,i} \).
Distinguishing error components

Assumption B (Discretization, linearization, and algebraic errors)

There exist fluxes $d_{h}^{k,i}, l_{h}^{k,i}, a_{h}^{k,i} \in [L^q(\Omega)]^d$ such that

(i) $d_{h}^{k,i} + l_{h}^{k,i} + a_{h}^{k,i} = \sigma_{h}^{k,i}$;

(ii) as the linear solver converges, $\|a_{h}^{k,i}\|_q \to 0$;

(iii) as the nonlinear solver converges, $\|l_{h}^{k,i}\|_q \to 0$.

Comments

- $d_{h}^{k,i}$: discretization flux reconstruction
- $l_{h}^{k,i}$: linearization error flux reconstruction
- $a_{h}^{k,i}$: algebraic error flux reconstruction

Alexandre Ern and Martin Vohralík
Adaptive inexact Newton methods
Assumption B (Discretization, linearization, and algebraic errors)

There exist fluxes $d_{h}^{k,i}, l_{h}^{k,i}, a_{h}^{k,i} \in [L^q(\Omega)]^d$ such that

(i) $d_{h}^{k,i} + l_{h}^{k,i} + a_{h}^{k,i} = \sigma_{h}^{k,i}$;

(ii) as the linear solver converges, $\|a_{h}^{k,i}\|_q \to 0$;

(iii) as the nonlinear solver converges, $\|l_{h}^{k,i}\|_q \to 0$.

Comments

- $d_{h}^{k,i}$: discretization flux reconstruction
- $l_{h}^{k,i}$: linearization error flux reconstruction
- $a_{h}^{k,i}$: algebraic error flux reconstruction
Theorem (Estimate distinguishing different error components)

Let

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

Then there holds

\[
J_u(u_h^{k,i}) \leq \eta_{k,i}^{k,i} := \eta_{\text{disc}}^{k,i} + \eta_{\text{lin}}^{k,i} + \eta_{\text{alg}}^{k,i} + \eta_{\text{rem}}^{k,i} + \eta_{\text{quad}}^{k,i} + \eta_{\text{osc}}^{k,i}.
\]
Theorem (Estimate distinguishing different error components)

Let
- $u \in V$ be the weak solution,
- $u_{h}^{k,i} \in V(T_{h})$ be arbitrary,
- Assumptions A and B hold.

Then there holds

$$J_{u}(u_{h}^{k,i}) \leq \eta^{k,i} := \eta_{\text{disc}}^{k,i} + \eta_{\text{lin}}^{k,i} + \eta_{\text{alg}}^{k,i} + \eta_{\text{rem}}^{k,i} + \eta_{\text{quad}}^{k,i} + \eta_{\text{osc}}^{k,i}.$$
Estimators

- **discretization estimator**
  \[
  \eta_{\text{disc},K}^{k,i} := 2^{1/p} \left( \| \sigma_h^{k,i} + d_h^{k,i} \|_{q,K} + \left\{ \sum_{e \in \mathcal{E}_K} h_e^{1-q} \left\| \left[ u_h^{k,i}, \nabla u_h^{k,i} \right] \right\|_{q,e} \right\}^{1/q} \right)
  \]

- **linearization estimator**
  \[
  \eta_{\text{lin},K}^{k,i} := \| l_h^{k,i} \|_{q,K}
  \]

- **algebraic estimator**
  \[
  \eta_{\text{alg},K}^{k,i} := \| a_h^{k,i} \|_{q,K}
  \]

- **algebraic remainder estimator**
  \[
  \eta_{\text{rem},K}^{k,i} := h_\Omega \| \rho_h^{k,i} \|_{q,K}
  \]

- **quadrature estimator**
  \[
  \eta_{\text{quad},K}^{k,i} := \| \sigma(u_h^{k,i}, \nabla u_h^{k,i}) - \sigma_h^{k,i} \|_{q,K}
  \]

- **data oscillation estimator**
  \[
  \eta_{\text{osc},K}^{k,i} := C_{p,h} h_K \| f - f_h \|_{q,K}
  \]

- \[
  \eta^{k,i} := \left\{ \sum_{K \in \mathcal{T}_h} (\eta_{\cdot,K}^{k,i})^q \right\}^{1/q}
  \]
Outline

1. Bibliography
2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results
3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - **Stopping criteria and efficiency**
   - Application and numerical results
4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results
5. Conclusions and future directions
Global stopping criteria

- stop whenever:

\[ \eta_{\text{rem}}^{k,i} \leq \gamma_{\text{rem}} \max \{ \eta_{\text{disc}}^{k,i}, \eta_{\text{lin}}^{k,i}, \eta_{\text{alg}}^{k,i} \}, \]

\[ \eta_{\text{alg}}^{k,i} \leq \gamma_{\text{alg}} \max \{ \eta_{\text{disc}}^{k,i}, \eta_{\text{lin}}^{k,i} \}, \]

\[ \eta_{\text{lin}}^{k,i} \leq \gamma_{\text{lin}} \eta_{\text{disc}}^{k,i}. \]

- \( \gamma_{\text{rem}}, \gamma_{\text{alg}}, \gamma_{\text{lin}} \approx 0.1 \)

Local stopping criteria

- stop whenever:

\[ \eta_{\text{rem},K}^{k,i} \leq \gamma_{\text{rem},K} \max \{ \eta_{\text{disc},K}^{k,i}, \eta_{\text{lin},K}^{k,i}, \eta_{\text{alg},K}^{k,i} \} \quad \forall K \in \mathcal{T}_h, \]

\[ \eta_{\text{alg},K}^{k,i} \leq \gamma_{\text{alg},K} \max \{ \eta_{\text{disc},K}^{k,i}, \eta_{\text{lin},K}^{k,i} \} \quad \forall K \in \mathcal{T}_h, \]

\[ \eta_{\text{lin},K}^{k,i} \leq \gamma_{\text{lin},K} \eta_{\text{disc},K}^{k,i} \quad \forall K \in \mathcal{T}_h. \]

- \( \gamma_{\text{rem},K}, \gamma_{\text{alg},K}, \gamma_{\text{lin},K} \approx 0.1 \)
Stopping criteria

Global stopping criteria

- stop whenever:
  \[ \eta_{\text{rem}}^{k,i} \leq \gamma_{\text{rem}} \max\{ \eta_{\text{disc}}^{k,i}, \eta_{\text{lin}}^{k,i}, \eta_{\text{alg}}^{k,i} \} , \]
  \[ \eta_{\text{alg}}^{k,i} \leq \gamma_{\text{alg}} \max\{ \eta_{\text{disc}}^{k,i}, \eta_{\text{lin}}^{k,i} \} , \]
  \[ \eta_{\text{lin}}^{k,i} \leq \gamma_{\text{lin}} \eta_{\text{disc}}^{k,i} \]
  \[ \gamma_{\text{rem}}, \gamma_{\text{alg}}, \gamma_{\text{lin}} \approx 0.1 \]

Local stopping criteria

- stop whenever:
  \[ \eta_{\text{rem},K}^{k,i} \leq \gamma_{\text{rem},K} \max\{ \eta_{\text{disc},K}^{k,i}, \eta_{\text{lin},K}^{k,i}, \eta_{\text{alg},K}^{k,i} \} \quad \forall K \in \mathcal{T}_h, \]
  \[ \eta_{\text{alg},K}^{k,i} \leq \gamma_{\text{alg},K} \max\{ \eta_{\text{disc},K}^{k,i}, \eta_{\text{lin},K}^{k,i} \} \quad \forall K \in \mathcal{T}_h, \]
  \[ \eta_{\text{lin},K}^{k,i} \leq \gamma_{\text{lin}} \eta_{\text{disc},K}^{k,i} \quad \forall K \in \mathcal{T}_h \]
  \[ \gamma_{\text{rem},K}, \gamma_{\text{alg},K}, \gamma_{\text{lin},K} \approx 0.1 \]
Assumption C (Approximation property)

For all $K \in T_h$, there holds

$$
\| \overline{\sigma}^{k,i}_h + d^{k,i}_h \|_{q,K} \lesssim \eta^{k,i}_{\#;\xi_K} + \eta^{k,i}_{osc;\xi_K},
$$

where

$$
\eta^{k,i}_{\#;\xi_K} := \left\{ \sum_{K' \in \xi_K} h^{q}_{K'} \| f_h + \nabla \cdot \overline{\sigma}^{k,i}_h \|_{q,K'} + \sum_{e \in \mathcal{C}_{K}^{\text{int}}} h_e \| [\overline{\sigma}^{k,i}_h \cdot n_e] \|_{q,e} \right\}^{1/q}
$$

$$
+ \sum_{e \in \mathcal{E}_K} h^{1-q}_{e} \| [u^{k,i}_h] \|_{q,e} \right\}^{1/q}.
$$

Alexandre Ern and Martin Vohralík

Adaptive inexact Newton methods
Theorem (Global efficiency)

Let the mesh $\mathcal{T}_h$ be shape-regular and let the global stopping criteria hold. Recall that $J_u(u_h^{k,i}) \leq \eta^{k,i}$. Then, under Assumption C,

$$\eta^{k,i} \lesssim J_u(u_h^{k,i}) + \eta^{k,i}_{\text{quad}} + \eta^{k,i}_{\text{osc}},$$

where $\lesssim$ means up to a constant independent of $\sigma$ and $q$.

- **robustness** with respect to the nonlinearity thanks to the choice of the dual norm as error measure.
Global efficiency

Theorem (Global efficiency)

Let the mesh $\mathcal{T}_h$ be shape-regular and let the global stopping criteria hold. Recall that $J_u(u_h^{k,i}) \leq \eta_{k,i}^k$. Then, under Assumption C,

$$\eta_{k,i}^k \leq J_u(u_h^{k,i}) + \eta_{quad}^k + \eta_{osc}^k,$$

where $\leq$ means up to a constant independent of $\sigma$ and $q$.

- **robustness** with respect to the nonlinearity thanks to the choice of the dual norm as error measure

Alexandre Ern and Martin Vohralík

Adaptive inexact Newton methods
Global efficiency

Theorem (Global efficiency)

Let the mesh $\mathcal{T}_h$ be shape-regular and let the global stopping criteria hold. Recall that $J_u(u_h^{k,i}) \leq \eta^{k,i}$. Then, under Assumption C,

$$\eta^{k,i} \lesssim J_u(u_h^{k,i}) + \eta^{k,i}_{\text{quad}} + \eta^{k,i}_{\text{osc}},$$

where $\lesssim$ means up to a constant independent of $\sigma$ and $q$.

- robustness with respect to the nonlinearity thanks to the choice of the dual norm as error measure
Theorem (Global efficiency)

Let the mesh $\mathcal{T}_h$ be shape-regular and let the global stopping criteria hold. Recall that $\mathcal{J}_u(u_h^{k,i}) \leq \eta^{k,i}$. Then, under Assumption C,

$$\eta^{k,i} \lesssim \mathcal{J}_u(u_h^{k,i}) + \eta_{\text{quad}}^{k,i} + \eta_{\text{osc}}^{k,i},$$

where $\lesssim$ means up to a constant independent of $\sigma$ and $q$.

- robustness with respect to the nonlinearity thanks to the choice of the dual norm as error measure
Theorem (Local efficiency)

Let the mesh $\mathcal{T}_h$ be shape-regular and let the local stopping criteria hold. Then, under Assumption C,

$$\eta_{k,i}^{\text{disc},K} + \eta_{k,i}^{\text{lin},K} + \eta_{k,i}^{\text{alg},K} + \eta_{k,i}^{\text{rem},K} \lesssim \mathcal{J}_{u,\bar{\Sigma}_K}^\text{up}(u_h^{k,i}) + \eta_{k,i}^{\text{quad},\bar{\Sigma}_K} + \eta_{k,i}^{\text{osc},\bar{\Sigma}_K}$$

for all $K \in \mathcal{T}_h$.

- **robustness** and **local efficiency** for an upper bound on the dual norm
Theorem (Local efficiency)

Let the mesh $\mathcal{T}_h$ be shape-regular and let the local stopping criteria hold. Then, under Assumption C,

$$
\eta_{\text{disc},K}^k,i + \eta_{\text{lin},K}^k,i + \eta_{\text{alg},K}^k,i + \eta_{\text{rem},K}^k,i \\
\lesssim \mathcal{J}_{u,\xi_K}^\text{up}(u_h^k,i) + \eta_{\text{quad},\xi_K}^k,i + \eta_{\text{osc},\xi_K}^k,i
$$

for all $K \in \mathcal{T}_h$.

- robustness and **local efficiency** for an upper bound on the dual norm.
Local efficiency

Theorem (Local efficiency)

Let the mesh $T_h$ be shape-regular and let the local stopping criteria hold. Then, under Assumption C,

$$\eta_{k,i}^{\text{disc},K} + \eta_{k,i}^{\text{lin},K} + \eta_{k,i}^{\text{alg},K} + \eta_{k,i}^{\text{rem},K} \lesssim \mathcal{J}_{u,\mathcal{K}}^{\text{up}}(u_h^{k,i}) + \eta_{k,i}^{\text{quad},\mathcal{K}} + \eta_{k,i}^{\text{osc},\mathcal{K}}$$

for all $K \in T_h$.

- robustness and local efficiency for an upper bound on the dual norm
Outline

1. Bibliography

2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results

3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results

4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results

5. Conclusions and future directions
### Construction of $a_h^{k,i}$ and $\rho_h^{k,i}$

- On linearization step $k$ and algebraic step $i$, we have

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

- Do $\nu$ additional steps of the algebraic solver, yielding

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

- Construct the function $\rho_h^{k,i}$ from the algebraic residual vector $R^{k,i+\nu}$ (lifting into appropriate discrete space).

- Suppose we can obtain discretization and linearization flux reconstructions $d_h^{k,i}, l_h^{k,i}$ on each algebraic step. Then set

$$a_h^{k,i} := (d_h^{k,i+\nu} + l_h^{k,i+\nu}) - (d_h^{k,i} + l_h^{k,i}).$$

- Independent of the algebraic solver.
Construction of $a_h^{k,i}$ and $\rho_h^{k,i}$

- On linearization step $k$ and algebraic step $i$, we have
  \[ A^{k-1}U^{k,i} = F^{k-1} - R^{k,i}. \]
- Do $\nu$ additional steps of the algebraic solver, yielding
  \[ A^{k-1}U^{k,i+\nu} = F^{k-1} - R^{k,i+\nu}. \]
- Construct the function $\rho_h^{k,i}$ from the algebraic residual vector $R^{k,i+\nu}$ (lifting into appropriate discrete space).
- Suppose we can obtain discretization and linearization flux reconstructions $d_h^{k,i}$, $l_h^{k,i}$ on each algebraic step. Then set
  \[ a_h^{k,i} := (d_h^{k,i+\nu} + l_h^{k,i+\nu}) - (d_h^{k,i} + l_h^{k,i}). \]
- Independent of the algebraic solver.
Construction of $a_{h,i}^{k}$ and $\rho_{h,i}^{k}$

On linearization step $k$ and algebraic step $i$, we have

$$\mathbb{A}^{k-1} U_{k,i} = F^{k-1} - R_{k,i}.$$ 

Do $\nu$ additional steps of the algebraic solver, yielding

$$\mathbb{A}^{k-1} U_{k,i+\nu} = F^{k-1} - R_{k,i+\nu}.$$ 

Construct the function $\rho_{h,i}^{k}$ from the algebraic residual vector $R_{k,i+\nu}$ (lifting into appropriate discrete space).

Suppose we can obtain discretization and linearization flux reconstructions $d_{h,i}^{k}$, $l_{h,i}^{k}$ on each algebraic step. Then set

$$a_{h,i}^{k} := (d_{h,i+\nu}^{k} + l_{h,i+\nu}^{k}) - (d_{h,i}^{k} + l_{h,i}^{k}).$$

Independent of the algebraic solver.
Algebraic error flux reconstruction and remainder

Construction of $a_h^{k,i}$ and $\rho_h^{k,i}$

On linearization step $k$ and algebraic step $i$, we have

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

Do $\nu$ additional steps of the algebraic solver, yielding

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

Construct the function $\rho_h^{k,i}$ from the algebraic residual vector $R_h^{k,i+\nu}$ (lifting into appropriate discrete space).

Suppose we can obtain discretization and linearization flux reconstructions $d_h^{k,i}$, $l_h^{k,i}$ on each algebraic step. Then set

$$a_h^{k,i} := (d_h^{k,i+\nu} + l_h^{k,i+\nu}) - (d_h^{k,i} + l_h^{k,i}).$$

Independent of the algebraic solver.
Construction of $a_h^{k,i}$ and $\rho_h^{k,i}$

- On linearization step $k$ and algebraic step $i$, we have
  \[ \mathbb{A}^{k-1} U^{k,i} = F^{k-1} - R^{k,i}. \]

- Do $\nu$ additional steps of the algebraic solver, yielding
  \[ \mathbb{A}^{k-1} U^{k,i+\nu} = F^{k-1} - R^{k,i+\nu}. \]

- Construct the function $\rho_h^{k,i}$ from the algebraic residual vector $R_h^{k,i+\nu}$ (lifting into appropriate discrete space).

- Suppose we can obtain discretization and linearization flux reconstructions $d_h^{k,i}$, $l_h^{k,i}$ on each algebraic step. Then set
  \[ a_h^{k,i} := (d_h^{k,i+\nu} + l_h^{k,i+\nu}) - (d_h^{k,i} + l_h^{k,i}). \]

- Independent of the algebraic solver.
Construction of $a_h^{k,i}$ and $\rho_h^{k,i}$

- On linearization step $k$ and algebraic step $i$, we have
  \[ \mathbb{A}^{k-1} U^{k,i} = F^{k-1} - R^{k,i}. \]

- Do $\nu$ additional steps of the algebraic solver, yielding
  \[ \mathbb{A}^{k-1} U^{k,i+\nu} = F^{k-1} - R^{k,i+\nu}. \]

- Construct the function $\rho_h^{k,i}$ from the algebraic residual vector $R_h^{k,i+\nu}$ (lifting into appropriate discrete space).

- Suppose we can obtain discretization and linearization flux reconstructions $d_h^{k,i}, l_h^{k,i}$ on each algebraic step. Then set
  \[ a_h^{k,i} := (d_h^{k,i+\nu} + l_h^{k,i+\nu}) - (d_h^{k,i} + l_h^{k,i}). \]

- Independent of the algebraic solver.
Nonconforming finite elements for the $p$-Laplacian

Discretization
Find $u_h \in V_h$ such that

$$(\sigma(\nabla u_h), \nabla v_h) = (f_h, v_h) \quad \forall v_h \in V_h.$$  

- $\sigma(\nabla u_h) = |\nabla u_h|^{p-2}\nabla u_h$
- $V_h$ the Crouzeix–Raviart space
- $f_h := \Pi_0 f$
- leads to the system of nonlinear algebraic equations

$$\mathcal{A}(U) = F$$
Nonconforming finite elements for the \( p \)-Laplacian

**Discretization**
Find \( u_h \in V_h \) such that

\[
(\sigma(\nabla u_h), \nabla v_h) = (f_h, v_h) \quad \forall v_h \in V_h.
\]

- \( \sigma(\nabla u_h) = |\nabla u_h|^{p-2}\nabla u_h \)
- \( V_h \) the Crouzeix–Raviart space
- \( f_h := \Pi_0 f \)
- leads to the system of nonlinear algebraic equations

\[
\mathcal{A}(U) = F
\]
Linearization

Find $u^k_h \in V_h$ such that

$$(\sigma^{k-1}(\nabla u^k_h), \nabla \psi_e) = (f_h, \psi_e) \quad \forall e \in \mathcal{E}^\text{int}_h.$$ 

- $u^0_h \in V_h$ yields the initial vector $U^0$
- fixed-point linearization

$$\sigma^{k-1}(\xi) := |\nabla u^{k-1}_h|^{p-2} \xi$$

- Newton linearization

$$\sigma^{k-1}(\xi) := |\nabla u^{k-1}_h|^{p-2} \xi + (p - 2)|\nabla u^{k-1}_h|^{p-4}$$

$$(\nabla u^{k-1}_h \otimes \nabla u^{k-1}_h)(\xi - \nabla u^{k-1}_h)$$

- leads to the system of linear algebraic equations

$$\mathbf{A}^{k-1} U^k = F^{k-1}$$

Alexandre Ern and Martin Vohralík
Adaptive inexact Newton methods
Linearization

Find $u_h^k \in V_h$ such that

$$(\sigma^{k-1}(\nabla u_h^k), \nabla \psi_e) = (f_h, \psi_e) \quad \forall e \in \mathcal{E}_h^{\text{int}}.$$ 

- $u_h^0 \in V_h$ yields the initial vector $U^0$
- fixed-point linearization
  $$\sigma^{k-1}(\xi) := |\nabla u_h^{k-1}|^{p-2} \xi$$
- Newton linearization
  $$\sigma^{k-1}(\xi) := |\nabla u_h^{k-1}|^{p-2} \xi + (p - 2)|\nabla u_h^{k-1}|^{p-4} \left(\nabla u_h^{k-1} \otimes \nabla u_h^{k-1}\right)(\xi - \nabla u_h^{k-1})$$

leads to the system of linear algebraic equations

$$A^{k-1}U^k = F^{k-1}$$
Algebraic solution

Find $u_h^{k,i} \in V_h$ such that

$$
(\sigma^{k-1}(\nabla u_h^{k,i}), \nabla \psi_e) = (f_h, \psi_e) - R_e^{k,i} \quad \forall e \in \mathcal{E}_h^{\text{int}}.
$$

- algebraic residual vector $R^{k,i} = \{R_e^{k,i}\}_{e \in \mathcal{E}_h^{\text{int}}}$
- discrete system

$$
A^{k-1} U^k = F^{k-1} - R^{k,i}
$$
Algebraic solution

Find \( u_h^{k,i} \in V_h \) such that

\[
(\sigma^{k-1}(\nabla u_h^{k,i}), \nabla \psi_e) = (f_h, \psi_e) - R_e^{k,i} \quad \forall e \in \mathcal{E}_h^{\text{int}}.
\]

- algebraic residual vector \( R^{k,i} = \{R_e^{k,i}\}_{e \in \mathcal{E}_h^{\text{int}}} \)
- discrete system

\[
\Delta^{k-1} U^k = F^{k-1} - R^{k,i}
\]
Flux reconstructions

Definition (Construction of \( (d_h^{k,i} + l_h^{k,i}) \))

For all \( K \in \mathcal{T}_h \),
\[
(d_h^{k,i} + l_h^{k,i})|_K := -\sigma^{k-1}(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d} (x - x_K) - \sum_{e \in \mathcal{E}_K} \frac{R_e^{k,i}}{d|D_e|} (x - x_K)|_{K_e},
\]
where, \( R_e^{k,i} = (f_h, \psi_e) - (\sigma^{k-1}(\nabla u_h^{k,i}), \nabla \psi_e) \quad \forall e \in \mathcal{E}_h^{\text{int}}. \)

Definition (Construction of \( d_h^{k,i} \))

For all \( K \in \mathcal{T}_h \),
\[
d_h^{k,i}|_K := -\sigma(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d} (x - x_K) - \sum_{e \in \mathcal{E}_K} \frac{R_e^{k,i}}{d|D_e|} (x - x_K)|_{K_e},
\]
where \( R_e^{k,i} := (f_h, \psi_e) - (\sigma(\nabla u_h^{k,i}), \nabla \psi_e) \quad \forall e \in \mathcal{E}_h^{\text{int}}. \)

Definition (Construction of \( \sigma_h^{k,i} \))

Set \( \sigma_h^{k,i} := \sigma(\nabla u_h^{k,i}) \). Consequently, \( \eta^{k,i}_{\text{quad},K} = 0 \) for all \( K \in \mathcal{T}_h \).
Flux reconstructions

Definition (Construction of \( (d_h^{k,i} + l_h^{k,i}) \))

For all \( K \in T_h \),

\[
(d_h^{k,i} + l_h^{k,i})|_K := -\sigma^{k-1}(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d}(x - x_K) - \sum_{e \in E_K} \frac{R_e^{k,i}}{d|D_e|}(x - x_K)|_{Ke},
\]

where,

\[
R_e^{k,i} = (f_h, \psi_e) - (\sigma^{k-1}(\nabla u_h^{k,i}), \nabla \psi_e) \quad \forall e \in E_h^{int}.
\]

Definition (Construction of \( d_h^{k,i} \))

For all \( K \in T_h \),

\[
d_h^{k,i}|_K := -\sigma(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d}(x - x_K) - \sum_{e \in E_K} \frac{R_e^{k,i}}{d|D_e|}(x - x_K)|_{Ke},
\]

where,

\[
R_e^{k,i} := (f_h, \psi_e) - (\sigma(\nabla u_h^{k,i}), \nabla \psi_e) \quad \forall e \in E_h^{int}.
\]

Definition (Construction of \( \sigma_h^{k,i} \))

Set \( \sigma_h^{k,i} := \sigma(\nabla u_h^{k,i}) \). Consequently, \( \eta_{quad,K}^{k,i} = 0 \) for all \( K \in T_h \).
**Flux reconstructions**

**Definition (Construction of \((d_h^{k,i} + l_h^{k,i})\))**

For all \(K \in T_h\),

\[
(d_h^{k,i} + l_h^{k,i})|_K := -\sigma^{-1}(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d}(x - x_K) - \sum_{e \in E_K} \frac{R_e^{k,i}}{d|D_e|}(x - x_K)|_{K_e},
\]

where, \(R_e^{k,i} = (f_h, \psi_e) - (\sigma^{-1}(\nabla u_h^{k,i}), \nabla \psi_e)\) \(\forall e \in E_h^{\text{int}}\).

**Definition (Construction of \(d_h^{k,i}\))**

For all \(K \in T_h\),

\[
d_h^{k,i}|_K := -\sigma(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d}(x - x_K) - \sum_{e \in E_K} \frac{\tilde{R}_e^{k,i}}{d|D_e|}(x - x_K)|_{K_e},
\]

where \(\tilde{R}_e^{k,i} := (f_h, \psi_e) - (\sigma(\nabla u_h^{k,i}), \nabla \psi_e)\) \(\forall e \in E_h^{\text{int}}\).

**Definition (Construction of \(\sigma_h^{k,i}\))**

Set \(\sigma_h^{k,i} := \sigma(\nabla u_h^{k,i})\). Consequently, \(\eta_{\text{quad},K}^{k,i} = 0\) for all \(K \in T_h\).
Verification of the assumptions – upper bound

Lemma (Assumptions A and B)

Assumptions A and B hold.

Comments

- \( \|a_h^{k,i}\|_{q,K} \to 0 \) as the linear solver converges by definition.
- \( \|l_h^{k,i}\|_{q,K} \to 0 \) as the nonlinear solver converges by the construction of \( l_h^{k,i} \).
- Both \( (d_h^{k,i} + l_h^{k,i}) \) and \( d_h^{k,i} \) belong to \( \text{RTN}_0(S_h) \) ⇒ \( a_h^{k,i} \in \text{RTN}_0(S_h) \) and \( \sigma_h^{k,i} \in \text{RTN}_0(S_h) \).
Verification of the assumptions – upper bound

Lemma (Assumptions A and B)

Assumptions A and B hold.

Comments

- $\|a_{h}^{k,i}\|_{q,K} \to 0$ as the linear solver converges by definition.
- $\|l_{h}^{k,i}\|_{q,K} \to 0$ as the nonlinear solver converges by the construction of $l_{h}^{k,i}$.
- Both $(d_{h}^{k,i} + l_{h}^{k,i})$ and $d_{h}^{k,i}$ belong to $\text{RTN}_0(S_h)$.$\Rightarrow a_{h}^{k,i} \in \text{RTN}_0(S_h)$ and $\sigma_{h}^{k,i} \in \text{RTN}_0(S_h)$. 

Alexandre Ern and Martin Vohralík

Adaptive inexact Newton methods
Verification of the assumptions – efficiency

**Lemma (Assumption C)**

**Assumption C holds.**

**Comments**

- $d_h^{k,i}$ close to $\sigma(\nabla u_h^{k,i})$
- approximation properties of Raviart–Thomas–Nédélec spaces
Verification of the assumptions – efficiency

Lemma (Assumption C)

**Assumption C holds.**

Comments

- $d_h^{k,i}$ close to $\sigma(\nabla u_h^{k,i})$
- approximation properties of Raviart–Thomas–Nédélec spaces
Summary

Discretization methods

- conforming finite elements
- nonconforming finite elements
- discontinuous Galerkin
- various finite volumes
- mixed finite elements

Linearizations

- fixed point
- Newton

Linear solvers

- independent of the linear solver

... all Assumptions A to C verified
Summary

Discretization methods
- conforming finite elements
- nonconforming finite elements
- discontinuous Galerkin
- various finite volumes
- mixed finite elements

Linearizations
- fixed point
- Newton

Linear solvers
- independent of the linear solver

... all Assumptions A to C verified
Summary

Discretization methods

- conforming finite elements
- nonconforming finite elements
- discontinuous Galerkin
- various finite volumes
- mixed finite elements

Linearizations

- fixed point
- Newton

Linear solvers

- independent of the linear solver

... all Assumptions A to C verified
Summary

Discretization methods

- conforming finite elements
- nonconforming finite elements
- discontinuous Galerkin
- various finite volumes
- mixed finite elements

Linearizations

- fixed point
- Newton

Linear solvers

- independent of the linear solver

... all Assumptions A to C verified
Numerical experiment I

Model problem

- $p$-Laplacian

$$\nabla \cdot (|\nabla u|^{p-2} \nabla u) = f \quad \text{in } \Omega,$$

$$u = u_D \quad \text{on } \partial \Omega$$

- weak solution (used to impose the Dirichlet BC)

$$u(x, y) = -\frac{p-1}{p} \left( (x - \frac{1}{2})^2 + (y - \frac{1}{2})^2 \right)^{\frac{p}{2(p-1)}} + \frac{p-1}{p} \left( \frac{1}{2} \right)^{\frac{p}{p-1}}$$

- tested values $p = 1.5$ and 10
- nonconforming finite elements
Analytical and approximate solutions

Case \( p = 1.5 \)

Case \( p = 10 \)
Error and estimators as a function of CG iterations, $p = 10$, 6th level mesh, 6th Newton step.

Newton
inexact Newton
ad. inexact Newton
Error and estimators as a function of Newton iterations, $\rho = 10$, 6th level mesh

- Newton
- inexact Newton
- ad. inexact Newton
Error and estimators, $p = 10$

- Newton
- inexact Newton
- ad. inexact Newton

Alexandre Ern and Martin Vohralík
Adaptive inexact Newton methods
Effectivity indices, $p = 10$

Newton

inexact Newton

ad. inexact Newton
Error distribution, $p = 10$

Estimated error distribution

Exact error distribution
Newton and algebraic iterations, $p = 10$

Newton it. / refinement  alg. it. / Newton step  alg. it. / refinement

Alexandre Ern and Martin Vohralík  Adaptive inexact Newton methods
Error and estimators as a function of CG iterations, $p = 1.5$, 6th level mesh, 1st Newton step.
Error and estimators as a function of Newton iterations, $\rho = 1.5$, 6th level mesh

Newton inexact Newton ad. inexact Newton

Alexandre Ern and Martin Vohralík
Adaptive inexact Newton methods
Error and estimators, $p = 1.5$

Newton  
inexact Newton  
ad. inexact Newton
Effectivity indices, $p = 1.5$

Newton

inexact Newton

ad. inexact Newton
Newton and algebraic iterations, $p = 1.5$

- Newton it. / refinement
- alg. it. / Newton step
- alg. it. / refinement
Numerical experiment II

Model problem

- \( p \)-Laplacian

\[
\nabla \cdot (|\nabla u|^{p-2} \nabla u) = f \quad \text{in } \Omega,
\]

\[
u = u_D \quad \text{on } \partial\Omega
\]

- weak solution (used to impose the Dirichlet BC)

\[
u(r, \theta) = r^{7/8} \sin(\theta^{7/8})
\]

- \( p = 4 \), L-shape domain, singularity in the origin (Carstensen and Klose (2003))

- nonconforming finite elements
Error distribution on an adaptively refined mesh

Estimated error distribution

Exact error distribution

Alexandre Ern and Martin Vohralík
Adaptive inexact Newton methods
Estimated and actual errors and the effectivity index

Estimated and actual errors

Effectivity index
Energy error and overall performance

Energy error

Overall performance

Alexandre Ern and Martin Vohralík

Adaptive inexact Newton methods
Outline

1. Bibliography

2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results

3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results

4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results

5. Conclusions and future directions
Two-phase flow in porous media

Two-phase flow in porous media

\[ \partial_t (\phi s_\alpha) + \nabla \cdot u_\alpha = q_\alpha, \quad \alpha \in \{n, w\}, \]
\[ -\lambda_\alpha (s_w) K (\nabla p_\alpha + \rho_\alpha g \nabla z) = u_\alpha, \quad \alpha \in \{n, w\}, \]
\[ s_n + s_w = 1, \]
\[ p_n - p_w = p_c(s_w) \]

Mathematical issues

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

Alexandre Ern and Martin Vohralík
Adaptive inexact Newton methods
Two-phase flow in porous media

\[
\begin{align*}
\partial_t (\phi s_\alpha) + \nabla \cdot \mathbf{u}_\alpha &= q_\alpha, & \alpha \in \{n, w\}, \\
-\lambda_\alpha (s_w) K (\nabla p_\alpha + \rho_\alpha g \nabla z) &= \mathbf{u}_\alpha, & \alpha \in \{n, w\}, \\
\quad s_n + s_w &= 1, \\
\quad p_n - p_w &= p_c (s_w)
\end{align*}
\]

Mathematical issues

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

1. Bibliography
2. Laplace equation
   - A guaranteed a posteriori error estimate
   - Polynomial-degree-robust local efficiency
   - Application and numerical results
3. Quasi-linear elliptic problems
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Application and numerical results
4. Two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Applications and numerical results
5. Conclusions and future directions
Theorem (Link energy-type error – dual norm of the residual)

Let \((s_w, p_w)\) be the weak solution. Let \((s_{w,h_T}, p_{w,h_T})\) be a vertex-centered finite volume / backward Euler approximation. Then

\[
\|s_w - s_{w,h_T}\|_{L^2((0,T);H^{-1}(\Omega))} + \|q(s_w) - q(s_{w,h_T})\|_{L^2(\Omega \times (0,T))} \\
+ \|p(s_w, p_w) - p(s_{w,h_T}, p_{w,h_T})\|_{L^2((0,T);H^1_0(\Omega))} \leq C \left\{ \sum_{n=1}^{N} \|\|s_w - s_{w,h_T}, p_w - p_{w,h_T}\|_{L^2_{I_n}}^2 \right\}^{\frac{1}{2}}
\]
Theorem (Link energy-type error – dual norm of the residual)

Let \((s_w, p_w)\) be the weak solution. Let \((s_w, h_T, p_w, h_T)\) be a vertex-centered finite volume / backward Euler approximation. Then

\[
\|s_w - s_w, h_T\|_{L^2((0,T);H^{-1}(\Omega))} + \|q(s_w) - q(s_w, h_T)\|_{L^2(\Omega \times (0,T))} + \|p(s_w, p_w) - p(s_w, h_T, p_w, h_T)\|_{L^2((0,T);H^1_0(\Omega))} \\
\leq C \left\{ \sum_{n=1}^{N} \|(s_w - s_w, h_T, p_w - p_w, h_T)\|_{I_n}^2 \right\}^{\frac{1}{2}}
\]
Link energy-type error – dual norm of the residual

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

Let \((s_w, p_w)\) be the weak solution. Let \((s_{w,h\tau}, p_{w,h\tau})\) be a vertex-centered finite volume / backward Euler approximation. Then

\[
\|s_w - s_{w,h\tau}\|_{L^2((0,T);H^{-1}(\Omega))} + \|q(s_w) - q(s_{w,h\tau})\|_{L^2(\Omega \times (0,T))} + \|p(s_w, p_w) - p(s_{w,h\tau}, p_{w,h\tau})\|_{L^2((0,T);H^1_0(\Omega))} \\
\leq C \left\{ \sum_{n=1}^{N} \|s_w - s_{w,h\tau}, p_w - p_{w,h\tau}\|_{l_n} \right\}^{\frac{1}{2}}
\]
Distinguishing the error components

**Theorem (Distinguishing the error components)**

Let
- \( n \) be the **time** step,
- \( k \) be the **linearization** step,
- \( i \) be the **algebraic solver** step,

with the approximations \((s_{w,h_T}^{n,k,i}, p_{w,h_T}^{n,k,i})\). Then

\[
\|\|\| (s_w - s_{w,h_T}^{n,k,i}, p_w - p_{w,h_T}^{n,k,i}) \|\|_n \leq \eta_{sp}^{n,k,i} + \eta_{tm}^{n,k,i} + \eta_{lin}^{n,k,i} + \eta_{alg}^{n,k,i}.
\]

**Error components**

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

Alexandre Ern and Martin Vohralík

Adaptive inexact Newton methods
Distinguishing the error components

**Theorem (Distinguishing the error components)**

Let

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

with the approximations \((s^n_{w,h_T}, p^n_{w,h_T})\). Then

\[
\left\| (s_w - s^n_{w,h_T}, p_w - p^n_{w,h_T}) \right\|_n \leq \eta^{n,k,i}_{sp} + \eta^{n,k,i}_{tm} + \eta^{n,k,i}_{lin} + \eta^{n,k,i}_{alg}.
\]

**Error components**

- \( \eta^{n,k,i}_{sp} \): spatial discretization
- \( \eta^{n,k,i}_{tm} \): temporal discretization
- \( \eta^{n,k,i}_{lin} \): linearization
- \( \eta^{n,k,i}_{alg} \): algebraic solver
Estimators and stopping criteria

Estimators in function of GMRes iterations

Estimators in function of iterative coupling iterations

Adaptive inexact Newton methods
GMRes relative residual/iterative coupling iterations

GMRes relative residual

Iterative coupling iterations

Alexandre Ern and Martin Vohralík

Adaptive inexact Newton methods
GMRes iterations

Per time and iterative coupling step

Cumulated number of GMRes iterations

Cumulated

Alexandre Ern and Martin Vohralík

Adaptive inexact Newton methods
Fully adaptive computation
Outline

1 Bibliography

2 Laplace equation
   • A guaranteed a posteriori error estimate
   • Polynomial-degree-robust local efficiency
   • Application and numerical results

3 Quasi-linear elliptic problems
   • A guaranteed a posteriori error estimate
   • Stopping criteria and efficiency
   • Application and numerical results

4 Two-phase flow in porous media
   • A guaranteed a posteriori error estimate
   • Applications and numerical results

5 Conclusions and future directions
Conclusions

Entire adaptivity

- only a necessary number of algebraic/linearization solver iterations
- "online decisions": algebraic step / linearization step / space mesh refinement / time step modification
- important computational savings
- guaranteed and robust a posteriori error estimates

Future directions

- other coupled nonlinear systems
- convergence and optimality
Conclusions

Entire adaptivity

- only a necessary number of algebraic/linearization solver iterations
- “online decisions”: algebraic step / linearization step / space mesh refinement / time step modification
- important computational savings
- guaranteed and robust a posteriori error estimates

Future directions

- other coupled nonlinear systems
- convergence and optimality


Thank you for your attention!