Contrôle d'erreur numérique a posteriori et critères d'arrêt pour des solveurs linéaires et non linéaires

Martin Vohralík

INRIA Paris-Rocquencourt

Séminaire de modélisation et calcul scientifique,
Rocquencourt, 15 janvier 2013

Outline

1. Introduction
2. Adaptive inexact Newton method
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Applications
   - Numerical results
3. Application to two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Application and numerical results
4. Conclusions and future directions

Outline

1. Introduction
2. Adaptive inexact Newton method
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Applications
   - Numerical results
3. Application to two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Application and numerical results
4. Conclusions and future directions

Inexact Newton method

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 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 1 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 Newton method

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 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 1 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 Newton method

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 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 1 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 Newton method

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 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 1 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 Newton method

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 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 1 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 Newton method

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 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 1 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 Newton method

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 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 1 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.*

hématiques

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?*

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?*

Previous results

Inexact Newton method

- Eisenstat and Walker (1990's)
- Moret (1989)

Stopping criteria

- engineering literature, since 1950's
- Becker, Johnson, and Rannacher (1995), multigrid st. crit.
- Arioli (2000's)

A posteriori error estimates for 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, loc. cons. methods

Previous results

Inexact Newton method

- Eisenstat and Walker (1990's)
- Moret (1989)

Stopping criteria

- engineering literature, since 1950's
- Becker, Johnson, and Rannacher (1995), multigrid st. crit.
- Arioli (2000's)

A posteriori error estimates for 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, loc. cons. methods

Previous results

Inexact Newton method

- Eisenstat and Walker (1990's)
- Moret (1989)

Stopping criteria

- engineering literature, since 1950's
- Becker, Johnson, and Rannacher (1995), multigrid st. crit.
- Arioli (2000's)

A posteriori error estimates for 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, loc. cons. methods

Outline

1. Introduction
2. Adaptive inexact Newton method
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Applications
   - Numerical results
3. Application to two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Application and numerical results
4. Conclusions and future directions

Quasi-linear elliptic problem

Quasi-linear elliptic problem

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

- quasi-linear diffusion problem

$$\boldsymbol{\sigma}(v,\boldsymbol{\xi}) = \underline{\mathbf{A}}(v)\boldsymbol{\xi} \qquad \forall (v,\boldsymbol{\xi}) \in \mathbb{R}\times\mathbb{R}^d$$

- Leray–Lions problem

$$\boldsymbol{\sigma}(v,\boldsymbol{\xi}) = \underline{\mathbf{A}}(\boldsymbol{\xi})\boldsymbol{\xi} \qquad \forall \boldsymbol{\xi} \in \mathbb{R}^d$$

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

Example

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

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

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

Weak formulation

Find $u \in V$ such that

$$A(u) = f \text{ in } V'$$

Quasi-linear elliptic problem

Quasi-linear elliptic problem

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

- quasi-linear diffusion problem

$$\boldsymbol{\sigma}(v, \xi) = \underline{\mathbf{A}}(v)\xi \qquad \forall (v, \xi) \in \mathbb{R} \times \mathbb{R}^d$$

- Leray–Lions problem

$$\boldsymbol{\sigma}(v, \xi) = \underline{\mathbf{A}}(\xi)\xi \qquad \forall \xi \in \mathbb{R}^d$$

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

Example

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

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

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

Weak formulation

Find $u \in V$ such that

$$A(u) = f \text{ in } V'$$

Quasi-linear elliptic problem

Quasi-linear elliptic problem

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

- quasi-linear diffusion problem

$$\boldsymbol{\sigma}(v,\xi) = \underline{\mathbf{A}}(v)\xi \qquad \forall (v,\xi) \in \mathbb{R}\times\mathbb{R}^d$$

- Leray–Lions problem

$$\boldsymbol{\sigma}(v,\xi) = \underline{\mathbf{A}}(\xi)\xi \qquad \forall \xi \in \mathbb{R}^d$$

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

Example

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

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

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

Weak formulation

Find $u \in V$ such that

$$A(u) = f \text{ in } V'$$

Quasi-linear elliptic problem

Quasi-linear elliptic problem

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

- quasi-linear diffusion problem

$$\boldsymbol{\sigma}(v,\xi) = \underline{\mathbf{A}}(v)\xi \qquad \forall (v,\xi) \in \mathbb{R}\times\mathbb{R}^d$$

- Leray–Lions problem

$$\boldsymbol{\sigma}(v,\xi) = \underline{\mathbf{A}}(\xi)\xi \qquad \forall \xi \in \mathbb{R}^d$$

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

Example

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

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

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

Weak formulation

Find $u \in V$ such that

$$A(u) = f \text{ in } V'$$

Quasi-linear elliptic problem

Quasi-linear elliptic problem

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

- quasi-linear diffusion problem

$$\boldsymbol{\sigma}(v,\boldsymbol{\xi}) = \underline{\mathbf{A}}(v)\boldsymbol{\xi} \qquad \forall (v,\boldsymbol{\xi}) \in \mathbb{R}\times\mathbb{R}^d$$

- Leray–Lions problem

$$\boldsymbol{\sigma}(v,\boldsymbol{\xi}) = \underline{\mathbf{A}}(\boldsymbol{\xi})\boldsymbol{\xi} \qquad \forall \boldsymbol{\xi} \in \mathbb{R}^d$$

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

Example

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

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

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

Weak formulation

Find $u \in V$ such that

$$A(u) = f \text{ in } V'$$

Quasi-linear elliptic problem

Quasi-linear elliptic problem

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

- quasi-linear diffusion problem

$$\boldsymbol{\sigma}(v,\boldsymbol{\xi}) = \underline{\mathbf{A}}(v)\boldsymbol{\xi} \qquad \forall (v,\boldsymbol{\xi}) \in \mathbb{R}\times\mathbb{R}^d$$

- Leray–Lions problem

$$\boldsymbol{\sigma}(v,\boldsymbol{\xi}) = \underline{\mathbf{A}}(\boldsymbol{\xi})\boldsymbol{\xi} \qquad \forall \boldsymbol{\xi} \in \mathbb{R}^d$$

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

Example

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

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

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

Weak formulation

Find $u \in V$ such that

$$A(u) = f \text{ in } V'$$

Quasi-linear elliptic problem

Quasi-linear elliptic problem

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

- quasi-linear diffusion problem

$$\boldsymbol{\sigma}(v,\boldsymbol{\xi}) = \underline{\mathbf{A}}(v)\boldsymbol{\xi} \qquad \forall (v,\boldsymbol{\xi}) \in \mathbb{R}\times\mathbb{R}^d$$

- Leray–Lions problem

$$\boldsymbol{\sigma}(v,\boldsymbol{\xi}) = \underline{\mathbf{A}}(\boldsymbol{\xi})\boldsymbol{\xi} \qquad \forall \boldsymbol{\xi} \in \mathbb{R}^d$$

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

Example

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

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

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

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

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

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

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

$$\mathcal{J}_u(u_h^{k,i}) \le \mathcal{J}_u^{\mathrm{up}}(u_h^{k,i}) := \|\sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i})\|_q + \mathcal{J}_{u,\mathrm{NC}}(u_h^{k,i})$$

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

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

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

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

$$\mathcal{J}_u(u_h^{k,i}) \le \mathcal{J}_u^{\mathrm{up}}(u_h^{k,i}) := \|\sigma(u, \nabla u) - \sigma(u_h^{k,i}, \nabla u_h^{k,i})\|_q + \mathcal{J}_{u,\mathrm{NC}}(u_h^{k,i})$$

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

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

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

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

$$\mathcal{J}_u(u_h^{k,i}) \le \mathcal{J}_u^{\mathrm{up}}(u_h^{k,i}) := \|\boldsymbol{\sigma}(u,\nabla u) - \boldsymbol{\sigma}(u_h^{k,i},\nabla u_h^{k,i})\|_q + \mathcal{J}_{u,\mathrm{NC}}(u_h^{k,i})$$

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

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

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

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

$$\mathcal{J}_u(u_h^{k,i}) \le \mathcal{J}_u^{\mathrm{up}}(u_h^{k,i}) := \|\boldsymbol{\sigma}(u,\nabla u) - \boldsymbol{\sigma}(u_h^{k,i},\nabla u_h^{k,i})\|_q + \mathcal{J}_{u,\mathrm{NC}}(u_h^{k,i})$$

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

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

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

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

$$\mathcal{J}_u(u_h^{k,i}) \le \mathcal{J}_u^{\mathrm{up}}(u_h^{k,i}) := \|\boldsymbol{\sigma}(u, \nabla u) - \boldsymbol{\sigma}(u_h^{k,i}, \nabla u_h^{k,i})\|_q + \mathcal{J}_{u,\mathrm{NC}}(u_h^{k,i})$$

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

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

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

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

$$\mathcal{J}_u(u_h^{k,i}) \le \mathcal{J}_u^{\mathrm{up}}(u_h^{k,i}) := \|\boldsymbol{\sigma}(u, \nabla u) - \boldsymbol{\sigma}(u_h^{k,i}, \nabla u_h^{k,i})\|_q + \mathcal{J}_{u,\mathrm{NC}}(u_h^{k,i})$$

Outline

1. Introduction
2. Adaptive inexact Newton method
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Applications
   - Numerical results
3. Application to two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Application and numerical results
4. Conclusions and future directions

A posteriori error estimate

Assumption A (Total flux reconstruction)

There exists a flux reconstruction $\mathbf{t}_h^{k,i} \in \mathbf{H}^q(\mathrm{div}, \Omega)$ *and an algebraic remainder* $\rho_h^{k,i} \in L^q(\Omega)$ *such that*

$$\nabla{\cdot}\mathbf{t}_h^{k,i} = f_h - \rho_h^{k,i},$$

with the data approximation f_h *s.t.* $(f_h, 1)_K = (f, 1)_K \;\; \forall K \in \mathcal{T}_h$.

Theorem (A guaranteed a posteriori error estimate)

Let

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

Then there holds

$$\mathcal{J}_u(u_h^{k,i}) \leq \overline{\eta}^{k,i},$$

where $\overline{\eta}^{k,i}$ *is fully computable from* $u_h^{k,i}$, $\mathbf{t}_h^{k,i}$, *and* $\rho_h^{k,i}$.

A posteriori error estimate

Assumption A (Total flux reconstruction)

There exists a flux reconstruction $\mathbf{t}_h^{k,i} \in \mathbf{H}^q(\mathrm{div}, \Omega)$ *and an algebraic remainder* $\rho_h^{k,i} \in L^q(\Omega)$ *such that*

$$\nabla \cdot \mathbf{t}_h^{k,i} = f_h - \rho_h^{k,i},$$

with the data approximation f_h *s.t.* $(f_h, 1)_K = (f, 1)_K \quad \forall K \in \mathcal{T}_h$.

Theorem (A guaranteed a posteriori error estimate)

Let

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

Then there holds

$$\mathcal{J}_u(u_h^{k,i}) \leq \overline{\eta}^{k,i},$$

where $\overline{\eta}^{k,i}$ *is fully computable from* $u_h^{k,i}$, $\mathbf{t}_h^{k,i}$, *and* $\rho_h^{k,i}$.

A posteriori error estimate

Assumption A (Total flux reconstruction)

There exists a flux reconstruction $\mathbf{t}_h^{k,i} \in \mathbf{H}^q(\mathrm{div}, \Omega)$ *and an algebraic remainder* $\rho_h^{k,i} \in L^q(\Omega)$ *such that*

$$\nabla \cdot \mathbf{t}_h^{k,i} = f_h - \rho_h^{k,i},$$

with the data approximation f_h *s.t.* $(f_h, 1)_K = (f, 1)_K \quad \forall K \in \mathcal{T}_h$.

Theorem (A guaranteed a posteriori error estimate)

Let

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

Then there holds

$$\mathcal{J}_u(u_h^{k,i}) \leq \overline{\eta}^{k,i},$$

where $\overline{\eta}^{k,i}$ *is fully computable from* $u_h^{k,i}$, $\mathbf{t}_h^{k,i}$, *and* $\rho_h^{k,i}$.

Distinguishing error components

Assumption B (Discretization, linearization, and algebraic errors)

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

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

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

Comments

- $\mathbf{d}_h^{k,i}$: *discretization flux reconstruction*
- $\mathbf{l}_h^{k,i}$: *linearization **error** flux reconstruction*
- $\mathbf{a}_h^{k,i}$: *algebraic **error** flux reconstruction*

Distinguishing error components

Assumption B (Discretization, linearization, and algebraic errors)

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

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

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

Comments

- $\mathbf{d}_h^{k,i}$: *discretization flux reconstruction*
- $\mathbf{l}_h^{k,i}$: *linearization **error** flux reconstruction*
- $\mathbf{a}_h^{k,i}$: *algebraic **error** flux reconstruction*

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

$$\mathcal{J}_u(u_h^{k,i}) \leq \eta^{k,i} := \eta_{\mathrm{disc}}^{k,i} + \eta_{\mathrm{lin}}^{k,i} + \eta_{\mathrm{alg}}^{k,i} + \eta_{\mathrm{rem}}^{k,i} + \eta_{\mathrm{quad}}^{k,i} + \eta_{\mathrm{osc}}^{k,i}.$$

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

$$\mathcal{J}_u(u_h^{k,i}) \le \eta^{k,i} := \eta_{\mathrm{disc}}^{k,i} + \eta_{\mathrm{lin}}^{k,i} + \eta_{\mathrm{alg}}^{k,i} + \eta_{\mathrm{rem}}^{k,i} + \eta_{\mathrm{quad}}^{k,i} + \eta_{\mathrm{osc}}^{k,i}.$$

Estimators

- *discretization* estimator

$$\eta_{\mathrm{disc},K}^{k,i} := 2^{1/p}\left(\|\overline{\boldsymbol{\sigma}}_h^{k,i} + \mathbf{d}_h^{k,i}\|_{q,K} + \left\{\sum_{e\in\mathcal{E}_K} h_e^{1-q}\|[\![u_h^{k,i}]\!]\|_{q,e}^q\right\}^{\frac{1}{q}}\right)$$

- *linearization* estimator

$$\eta_{\mathrm{lin},K}^{k,i} := \|\mathbf{l}_h^{k,i}\|_{q,K}$$

- *algebraic* estimator

$$\eta_{\mathrm{alg},K}^{k,i} := \|\mathbf{a}_h^{k,i}\|_{q,K}$$

- *algebraic remainder estimator*

$$\eta_{\mathrm{rem},K}^{k,i} := h_\Omega\|\rho_h^{k,i}\|_{q,K}$$

- *quadrature estimator*

$$\eta_{\mathrm{quad},K}^{k,i} := \|\boldsymbol{\sigma}(u_h^{k,i}, \nabla u_h^{k,i}) - \overline{\boldsymbol{\sigma}}_h^{k,i}\|_{q,K}$$

- *data oscillation estimator*

$$\eta_{\mathrm{osc},K}^{k,i} := C_{\mathrm{P},p} h_K\|f - f_h\|_{q,K}$$

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

Outline

1 Introduction

2 Adaptive inexact Newton method
- A guaranteed a posteriori error estimate
- Stopping criteria and efficiency
- Applications
- Numerical results

3 Application to two-phase flow in porous media
- A guaranteed a posteriori error estimate
- Application and numerical results

4 Conclusions and future directions

Stopping criteria

Global stopping criteria

- stop whenever:

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

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

Local stopping criteria

- stop whenever:

$$\eta_{\text{rem},K}^{k,i} \le \gamma_{\text{rem},K} \max\{\eta_{\text{disc},K}^{k,i}, \eta_{\text{lin},K}^{k,i}, \eta_{\text{alg},K}^{k,i}\} \qquad \forall K \in \mathcal{T}_h,$$
$$\eta_{\text{alg},K}^{k,i} \le \gamma_{\text{alg},K} \max\{\eta_{\text{disc},K}^{k,i}, \eta_{\text{lin},K}^{k,i}\} \qquad \forall K \in \mathcal{T}_h,$$
$$\eta_{\text{lin},K}^{k,i} \le \gamma_{\text{lin},K} \eta_{\text{disc},K}^{k,i} \qquad \forall K \in \mathcal{T}_h$$

- $\gamma_{\text{rem},K}, \gamma_{\text{alg},K}, \gamma_{\text{lin},K} \approx 0.1$

Stopping criteria

Global stopping criteria

- stop whenever:

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

- $\gamma_{\rm rem}, \gamma_{\rm alg}, \gamma_{\rm lin} \approx 0.1$

Local stopping criteria

- stop whenever:

$$\eta_{{\rm rem},K}^{k,i} \le \gamma_{{\rm rem},K} \max\{\eta_{{\rm disc},K}^{k,i}, \eta_{{\rm lin},K}^{k,i}, \eta_{{\rm alg},K}^{k,i}\} \qquad \forall K \in \mathcal{T}_h,$$
$$\eta_{{\rm alg},K}^{k,i} \le \gamma_{{\rm alg},K} \max\{\eta_{{\rm disc},K}^{k,i}, \eta_{{\rm lin},K}^{k,i}\} \qquad \forall K \in \mathcal{T}_h,$$
$$\eta_{{\rm lin},K}^{k,i} \le \gamma_{{\rm lin},K} \eta_{{\rm disc},K}^{k,i} \qquad \forall K \in \mathcal{T}_h$$

- $\gamma_{{\rm rem},K}, \gamma_{{\rm alg},K}, \gamma_{{\rm lin},K} \approx 0.1$

Assumption for efficiency

Assumption C (Approximation property)

For all $K \in \mathcal{T}_h$, there holds

$$\|\overline{\boldsymbol{\sigma}}_h^{k,i} + \mathbf{d}_h^{k,i}\|_{q,K} \lesssim \eta_{\sharp,\mathfrak{T}_K}^{k,i} + \eta_{\mathrm{osc},\mathfrak{T}_K}^{k,i},$$

where

$$\eta_{\sharp,\mathfrak{T}_K}^{k,i} := \left\{ \sum_{K' \in \mathfrak{T}_K} h_{K'}^q \|f_h + \nabla\cdot\overline{\boldsymbol{\sigma}}_h^{k,i}\|_{q,K'}^q + \sum_{e \in \mathfrak{E}_K^{\mathrm{int}}} h_e \|[\![\overline{\boldsymbol{\sigma}}_h^{k,i}\cdot\mathbf{n}_e]\!]\|_{q,e}^q + \sum_{e \in \mathcal{E}_K} h_e^{1-q} \|[\![u_h^{k,i}]\!]\|_{q,e}^q \right\}^{\frac{1}{q}}.$$

Global efficiency

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,

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

where $\lesssim$ means up to a constant independent of σ 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 $\mathcal{J}_u(u_h^{k,i}) \leq \eta^{k,i}$. Then, under Assumption C,

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

where $\lesssim$ means up to a constant independent of σ 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 $\mathcal{J}_u(u_h^{k,i}) \leq \eta^{k,i}$. Then, under Assumption C,

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

where $\lesssim$ means up to a constant independent of σ 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 $\mathcal{J}_u(u_h^{k,i}) \le \eta^{k,i}$. Then, under Assumption C,

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

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

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

Local efficiency

Theorem (Local efficiency)

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

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

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

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

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

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

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

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

Outline

1. Introduction
2. Adaptive inexact Newton method
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Applications
   - Numerical results
3. Application to two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Application and numerical results
4. Conclusions and future directions

Algebraic error flux reconstruction and algebraic remainder

Construction of $\mathbf{a}_h^{k,i}$ 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 ν 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 $\mathbf{d}_h^{k,i}, \mathbf{l}_h^{k,i}$ on each algebraic step. Then set

$$\mathbf{a}_h^{k,i} := (\mathbf{d}_h^{k,i+\nu} + \mathbf{l}_h^{k,i+\nu}) - (\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i}).$$

- ν chosen adaptively so that $\eta_{\mathrm{rem},K}^{k,i}$ or $\eta_{\mathrm{rem}}^{k,i}$ are small enough.
- Independent of the algebraic solver.

Algebraic error flux reconstruction and algebraic remainder

Construction of $\mathbf{a}_h^{k,i}$ 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 ν 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 $\mathbf{d}_h^{k,i}, \mathbf{l}_h^{k,i}$ on each algebraic step. Then set

$$\mathbf{a}_h^{k,i} := (\mathbf{d}_h^{k,i+\nu} + \mathbf{l}_h^{k,i+\nu}) - (\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i}).$$

- ν chosen adaptively so that $\eta_{\mathrm{rem},K}^{k,i}$ or $\eta_{\mathrm{rem}}^{k,i}$ are small enough.
- Independent of the algebraic solver.

Algebraic error flux reconstruction and algebraic remainder

Construction of $\mathbf{a}_h^{k,i}$ 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 ν 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 $\mathbf{d}_h^{k,i}, \mathbf{l}_h^{k,i}$ on each algebraic step. Then set

$$\mathbf{a}_h^{k,i} := (\mathbf{d}_h^{k,i+\nu} + \mathbf{l}_h^{k,i+\nu}) - (\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i}).$$

- ν chosen adaptively so that $\eta_{\mathrm{rem},K}^{k,i}$ or $\eta_{\mathrm{rem}}^{k,i}$ are small enough.
- Independent of the algebraic solver.

Algebraic error flux reconstruction and algebraic remainder

Construction of $\mathbf{a}_h^{k,i}$ 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 ν 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 $\mathbf{d}_h^{k,i}, \mathbf{l}_h^{k,i}$ on each algebraic step. Then set

$$\mathbf{a}_h^{k,i} := (\mathbf{d}_h^{k,i+\nu} + \mathbf{l}_h^{k,i+\nu}) - (\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i}).$$

- ν chosen adaptively so that $\eta_{\mathrm{rem},K}^{k,i}$ or $\eta_{\mathrm{rem}}^{k,i}$ are small enough.
- Independent of the algebraic solver.

Algebraic error flux reconstruction and algebraic remainder

Construction of $\mathbf{a}_h^{k,i}$ 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 ν 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 $\mathbf{d}_h^{k,i}, \mathbf{l}_h^{k,i}$ on each algebraic step. Then set
$$\mathbf{a}_h^{k,i} := (\mathbf{d}_h^{k,i+\nu} + \mathbf{l}_h^{k,i+\nu}) - (\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i}).$$
- ν chosen adaptively so that $\eta_{\mathrm{rem},K}^{k,i}$ or $\eta_{\mathrm{rem}}^{k,i}$ are small enough.
- Independent of the algebraic solver.

Algebraic error flux reconstruction and algebraic remainder

Construction of $\mathbf{a}_h^{k,i}$ 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 ν 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 $\mathbf{d}_h^{k,i}, \mathbf{l}_h^{k,i}$ on each algebraic step. Then set
$$\mathbf{a}_h^{k,i} := (\mathbf{d}_h^{k,i+\nu} + \mathbf{l}_h^{k,i+\nu}) - (\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i}).$$
- ν chosen adaptively so that $\eta_{\mathrm{rem},K}^{k,i}$ or $\eta_{\mathrm{rem}}^{k,i}$ are small enough.
- Independent of the algebraic solver.

Algebraic error flux reconstruction and algebraic remainder

Construction of $\mathbf{a}_h^{k,i}$ 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 ν 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 $\mathbf{d}_h^{k,i}, \mathbf{l}_h^{k,i}$ on each algebraic step. Then set

$$\mathbf{a}_h^{k,i} := (\mathbf{d}_h^{k,i+\nu} + \mathbf{l}_h^{k,i+\nu}) - (\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i}).$$

- ν chosen adaptively so that $\eta_{\mathrm{rem},K}^{k,i}$ or $\eta_{\mathrm{rem}}^{k,i}$ are small enough.
- Independent of the algebraic solver.

Nonconforming finite elements for the p-Laplacian

Discretization

Find $u_h \in V_h$ such that

$$(\boldsymbol{\sigma}(\nabla u_h), \nabla v_h) = (f_h, v_h) \qquad \forall v_h \in V_h.$$

- $\boldsymbol{\sigma}(\nabla u_h) = |\nabla u_h|^{p-2} \nabla u_h$
- 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

$$(\boldsymbol{\sigma}(\nabla u_h), \nabla v_h) = (f_h, v_h) \qquad \forall v_h \in V_h.$$

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

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

Linearization

Linearization

Find $u_h^k \in V_h$ such that

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

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

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

- Newton linearization

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

- leads to the system of linear algebraic equations

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

Linearization

Linearization

Find $u_h^k \in V_h$ such that

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

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

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

- Newton linearization

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

- leads to the system of linear algebraic equations

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

Algebraic solution

Algebraic solution

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

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

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

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

Algebraic solution

Algebraic solution

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

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

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

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

Flux reconstructions

Definition (Construction of $(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i})$)

For all $K \in \mathcal{T}_h$,

$$(\mathbf{d}_h^{k,i}+\mathbf{l}_h^{k,i})|_K := -\boldsymbol{\sigma}^{k-1}(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d}(\mathbf{x}-\mathbf{x}_K) - \sum_{e\in\mathcal{E}_K} \frac{R_e^{k,i}}{d|D_e|}(\mathbf{x}-\mathbf{x}_K)|_{K_e},$$

where $R_e^{k,i} := (f_h, \psi_e) - (\boldsymbol{\sigma}^{k-1}(\nabla u_h^{k,i}), \nabla\psi_e) \qquad \forall e \in \mathcal{E}_h^{\mathrm{int}}$.

Definition (Construction of $\mathbf{d}_h^{k,i}$)

For all $K \in \mathcal{T}_h$,

$$\mathbf{d}_h^{k,i}|_K := -\boldsymbol{\sigma}(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d}(\mathbf{x}-\mathbf{x}_K) - \sum_{e\in\mathcal{E}_K} \frac{\bar{R}_e^{k,i}}{d|D_e|}(\mathbf{x}-\mathbf{x}_K)|_{K_e},$$

where $\bar{R}_e^{k,i} := (f_h, \psi_e) - (\boldsymbol{\sigma}(\nabla u_h^{k,i}), \nabla\psi_e) \qquad \forall e \in \mathcal{E}_h^{\mathrm{int}}$.

Definition (Construction of $\overline{\boldsymbol{\sigma}}_h^{k,i}$)

Set $\overline{\boldsymbol{\sigma}}_h^{k,i} := \boldsymbol{\sigma}(\nabla u_h^{k,i})$. Consequently, $\eta_{\mathrm{quad},K}^{k,i} = 0$ for all $K \in \mathcal{T}_h$.

Flux reconstructions

Definition (Construction of $(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i})$)

For all $K \in \mathcal{T}_h$,

$$(\mathbf{d}_h^{k,i}+\mathbf{l}_h^{k,i})|_K := -\boldsymbol{\sigma}^{k-1}(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d}(\mathbf{x}-\mathbf{x}_K) - \sum_{e\in\mathcal{E}_K} \frac{R_e^{k,i}}{d|D_e|}(\mathbf{x}-\mathbf{x}_K)|_{K_e},$$

where $R_e^{k,i} := (f_h, \psi_e) - (\boldsymbol{\sigma}^{k-1}(\nabla u_h^{k,i}), \nabla\psi_e) \qquad \forall e \in \mathcal{E}_h^{\mathrm{int}}$.

Definition (Construction of $\mathbf{d}_h^{k,i}$)

For all $K \in \mathcal{T}_h$,

$$\mathbf{d}_h^{k,i}|_K := -\boldsymbol{\sigma}(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d}(\mathbf{x}-\mathbf{x}_K) - \sum_{e\in\mathcal{E}_K} \frac{\bar{R}_e^{k,i}}{d|D_e|}(\mathbf{x}-\mathbf{x}_K)|_{K_e},$$

where $\bar{R}_e^{k,i} := (f_h, \psi_e) - (\boldsymbol{\sigma}(\nabla u_h^{k,i}), \nabla\psi_e) \qquad \forall e \in \mathcal{E}_h^{\mathrm{int}}$.

Definition (Construction of $\overline{\boldsymbol{\sigma}}_h^{k,i}$)

Set $\overline{\boldsymbol{\sigma}}_h^{k,i} := \boldsymbol{\sigma}(\nabla u_h^{k,i})$. Consequently, $\eta_{\mathrm{quad},K}^{k,i} = 0$ for all $K \in \mathcal{T}_h$.

Flux reconstructions

Definition (Construction of $(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i})$)

For all $K \in \mathcal{T}_h$,

$$(\mathbf{d}_h^{k,i}+\mathbf{l}_h^{k,i})|_K := -\boldsymbol{\sigma}^{k-1}(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d}(\mathbf{x}-\mathbf{x}_K) - \sum_{e\in\mathcal{E}_K} \frac{R_e^{k,i}}{d|D_e|}(\mathbf{x}-\mathbf{x}_K)|_{K_e},$$

where $R_e^{k,i} := (f_h, \psi_e) - (\boldsymbol{\sigma}^{k-1}(\nabla u_h^{k,i}), \nabla\psi_e) \qquad \forall e \in \mathcal{E}_h^{\mathrm{int}}$.

Definition (Construction of $\mathbf{d}_h^{k,i}$)

For all $K \in \mathcal{T}_h$,

$$\mathbf{d}_h^{k,i}|_K := -\boldsymbol{\sigma}(\nabla u_h^{k,i})|_K + \frac{f_h|_K}{d}(\mathbf{x}-\mathbf{x}_K) - \sum_{e\in\mathcal{E}_K} \frac{\bar{R}_e^{k,i}}{d|D_e|}(\mathbf{x}-\mathbf{x}_K)|_{K_e},$$

where $\bar{R}_e^{k,i} := (f_h, \psi_e) - (\boldsymbol{\sigma}(\nabla u_h^{k,i}), \nabla\psi_e) \qquad \forall e \in \mathcal{E}_h^{\mathrm{int}}$.

Definition (Construction of $\overline{\boldsymbol{\sigma}}_h^{k,i}$)

Set $\overline{\boldsymbol{\sigma}}_h^{k,i} := \boldsymbol{\sigma}(\nabla u_h^{k,i})$. Consequently, $\eta_{\mathrm{quad},K}^{k,i} = 0$ for all $K \in \mathcal{T}_h$.

Verification of the assumptions – upper bound

Lemma (Assumptions A and B)

Assumptions A and B hold.

Comments

- $\|\mathbf{a}_h^{k,i}\|_{q,K} \to 0$ as the linear solver converges by definition.
- $\|\mathbf{l}_h^{k,i}\|_{q,K} \to 0$ as the nonlinear solver converges by the construction of $\mathbf{l}_h^{k,i}$.
- Both $(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i})$ and $\mathbf{d}_h^{k,i}$ belong to $\mathbf{RTN}_0(\mathcal{S}_h) \Rightarrow$ $\mathbf{a}_h^{k,i} \in \mathbf{RTN}_0(\mathcal{S}_h)$ and $\mathbf{t}_h^{k,i} \in \mathbf{RTN}_0(\mathcal{S}_h)$.

Verification of the assumptions – upper bound

Lemma (Assumptions A and B)

Assumptions A and B hold.

Comments

- $\|\mathbf{a}_h^{k,i}\|_{q,K} \to 0$ as the linear solver converges by definition.
- $\|\mathbf{l}_h^{k,i}\|_{q,K} \to 0$ as the nonlinear solver converges by the construction of $\mathbf{l}_h^{k,i}$.
- Both $(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i})$ and $\mathbf{d}_h^{k,i}$ belong to $\mathbf{RTN}_0(\mathcal{S}_h) \Rightarrow$ $\mathbf{a}_h^{k,i} \in \mathbf{RTN}_0(\mathcal{S}_h)$ and $\mathbf{t}_h^{k,i} \in \mathbf{RTN}_0(\mathcal{S}_h)$.

Verification of the assumptions – efficiency

Lemma (Assumption C)

Assumption C holds.

Comments

- $\mathbf{d}_h^{k,i}$ 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

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

Discontinuous Galerkin for the quasi-linear diffusion

Discretization

Find $u_h \in V_h := \mathbb{P}_m(\mathcal{T}_h)$, $m \geq 1$, such that, for all $v_h \in V_h$,

$$(\boldsymbol{\sigma}(u_h, \nabla u_h), \nabla v_h) - \sum_{e \in \mathcal{E}_h} \{ \langle \{\!\{ \boldsymbol{\sigma}(u_h, \nabla u_h) \}\!\} \cdot \mathbf{n}_e, [\![v_h]\!] \rangle_e$$

$$+ \theta \langle \{\!\{ \underline{\mathbf{A}}(u_h) \nabla v_h \}\!\} \cdot \mathbf{n}_e, [\![u_h]\!] \rangle_e \} + \sum_{e \in \mathcal{E}_h} \langle \bar{\alpha}_e h_e^{-1} [\![u_h]\!], [\![v_h]\!] \rangle_e = (f, v_h).$$

- $\theta \in \{-1, 0, 1\}$
- $\bar{\alpha}_e := \|\underline{\mathbf{A}}\|_{L^\infty(\mathbb{R})} \chi_e$, χ_e large enough
- leads to the system of nonlinear algebraic equations

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

Discontinuous Galerkin for the quasi-linear diffusion

Discretization

Find $u_h \in V_h := \mathbb{P}_m(\mathcal{T}_h)$, $m \geq 1$, such that, for all $v_h \in V_h$,

$$(\boldsymbol{\sigma}(u_h, \nabla u_h), \nabla v_h) - \sum_{e \in \mathcal{E}_h} \{ \langle \{\!\{ \boldsymbol{\sigma}(u_h, \nabla u_h) \}\!\} \cdot \mathbf{n}_e, [\![v_h]\!] \rangle_e$$

$$+ \theta \langle \{\!\{ \underline{\mathbf{A}}(u_h) \nabla v_h \}\!\} \cdot \mathbf{n}_e, [\![u_h]\!] \rangle_e \} + \sum_{e \in \mathcal{E}_h} \langle \bar{\alpha}_e h_e^{-1} [\![u_h]\!], [\![v_h]\!] \rangle_e = (f, v_h).$$

- $\theta \in \{-1, 0, 1\}$
- $\bar{\alpha}_e := \|\underline{\mathbf{A}}\|_{L^\infty(\mathbb{R})} \chi_e$, χ_e large enough
- leads to the system of nonlinear algebraic equations

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

Linearization

Linearization

Find $u_h^k \in V_h$ such that, for all $K \in \mathcal{T}_h$ and all $j \in \mathcal{C}_K := \{1, \ldots, \dim(\mathbb{P}_m(K))\}$,

$$(\boldsymbol{\sigma}^{k-1}(u_h^k, \nabla u_h^k), \nabla \psi_{K,j}) - \sum_{e \in \mathcal{E}_h} \{ \langle \{\!\{ \boldsymbol{\sigma}^{k-1}(u_h^k, \nabla u_h^k) \}\!\} \cdot \mathbf{n}_e, [\![\psi_{K,j}]\!] \rangle_e$$
$$+ \theta \langle \{\!\{ \underline{\mathbf{A}}^{k-1}(u_h^k) \nabla \psi_{K,j} \}\!\} \cdot \mathbf{n}_e, [\![u_h^k]\!] \rangle_e \} + \sum_{e \in \mathcal{E}_h} \langle \bar{\alpha}_e h_e^{-1} [\![u_h^k]\!], [\![\psi_{K,j}]\!] \rangle_e = (f, \psi_{K,j}).$$

- $u_h^0 \in V_h$ yields the initial vector U^0
- fixed-point linearization $\boldsymbol{\sigma}^{k-1}(v, \boldsymbol{\xi}) := \underline{\mathbf{A}}(u_h^{k-1}) \boldsymbol{\xi}$
- Newton linearization

$$\boldsymbol{\sigma}^{k-1}(v, \boldsymbol{\xi}) := \underline{\mathbf{A}}(u_h^{k-1}) \boldsymbol{\xi} + (v - u_h^{k-1}) \partial_v \underline{\mathbf{A}}(u_h^{k-1}) \nabla u_h^{k-1},$$
$$\underline{\mathbf{A}}^{k-1}(v) := \underline{\mathbf{A}}(u_h^{k-1}) + \partial_v \underline{\mathbf{A}}(u_h^{k-1})(v - u_h^{k-1})$$

- leads to the system of linear algebraic equations

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

Linearization

Linearization

Find $u_h^k \in V_h$ such that, for all $K \in \mathcal{T}_h$ and all $j \in \mathcal{C}_K := \{1, \ldots, \dim(\mathbb{P}_m(K))\}$,

$$(\boldsymbol{\sigma}^{k-1}(u_h^k, \nabla u_h^k), \nabla \psi_{K,j}) - \sum_{e \in \mathcal{E}_h} \{ \langle \{\!\{ \boldsymbol{\sigma}^{k-1}(u_h^k, \nabla u_h^k) \}\!\} \cdot \mathbf{n}_e, [\![\psi_{K,j}]\!] \rangle_e$$
$$+ \theta \langle \{\!\{ \underline{\mathbf{A}}^{k-1}(u_h^k) \nabla \psi_{K,j} \}\!\} \cdot \mathbf{n}_e, [\![u_h^k]\!] \rangle_e \} + \sum_{e \in \mathcal{E}_h} \langle \bar{\alpha}_e h_e^{-1} [\![u_h^k]\!], [\![\psi_{K,j}]\!] \rangle_e = (f, \psi_{K,j}).$$

- $u_h^0 \in V_h$ yields the initial vector U^0
- fixed-point linearization $\boldsymbol{\sigma}^{k-1}(v, \xi) := \underline{\mathbf{A}}(u_h^{k-1}) \xi$
- Newton linearization

$$\boldsymbol{\sigma}^{k-1}(v, \xi) := \underline{\mathbf{A}}(u_h^{k-1}) \xi + (v - u_h^{k-1}) \partial_v \underline{\mathbf{A}}(u_h^{k-1}) \nabla u_h^{k-1},$$
$$\underline{\mathbf{A}}^{k-1}(v) := \underline{\mathbf{A}}(u_h^{k-1}) + \partial_v \underline{\mathbf{A}}(u_h^{k-1})(v - u_h^{k-1})$$

- leads to the system of linear algebraic equations

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

Algebraic solution

Algebraic solution

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

$$(\boldsymbol{\sigma}^{k-1}(u_h^{k,i}, \nabla u_h^{k,i}), \nabla \psi_{K,j}) - \sum_{e \in \mathcal{E}_h} \{ \langle \{\!\{ \boldsymbol{\sigma}^{k-1}(u_h^{k,i}, \nabla u_h^{k,i}) \}\!\} \cdot \mathbf{n}_e, [\![\psi_{K,j}]\!] \rangle_e$$

$$+ \theta \langle \{\!\{ \underline{\mathbf{A}}^{k-1}(u_h^{k,i}) \nabla \psi_{K,j} \}\!\} \cdot \mathbf{n}_e, [\![u_h^{k,i}]\!] \rangle_e \} + \sum_{e \in \mathcal{E}_h} \langle \bar{\alpha}_e h_e^{-1} [\![u_h^{k,i}]\!], [\![\psi_{K,j}]\!] \rangle_e$$

$$= (f, \psi_{K,j}) - R_{K,j}^{k,i}.$$

- algebraic residual vector $R^{k,i} = \{R_{K,j}^{k,i}\}_{K \in \mathcal{T}_h, j \in \mathcal{C}_K}$
- discrete system

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

Algebraic solution

Algebraic solution

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

$$(\boldsymbol{\sigma}^{k-1}(u_h^{k,i}, \nabla u_h^{k,i}), \nabla \psi_{K,j}) - \sum_{e \in \mathcal{E}_h} \{ \langle \{\!\{ \boldsymbol{\sigma}^{k-1}(u_h^{k,i}, \nabla u_h^{k,i}) \}\!\} \cdot \mathbf{n}_e, [\![\psi_{K,j}]\!] \rangle_e$$

$$+ \theta \langle \{\!\{ \underline{\mathbf{A}}^{k-1}(u_h^{k,i}) \nabla \psi_{K,j} \}\!\} \cdot \mathbf{n}_e, [\![u_h^{k,i}]\!] \rangle_e \} + \sum_{e \in \mathcal{E}_h} \langle \bar{\alpha}_e h_e^{-1} [\![u_h^{k,i}]\!], [\![\psi_{K,j}]\!] \rangle_e$$

$$= (f, \psi_{K,j}) - R_{K,j}^{k,i}.$$

- algebraic residual vector $R^{k,i} = \{R_{K,j}^{k,i}\}_{K \in \mathcal{T}_h, j \in \mathcal{C}_K}$
- discrete system

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

Flux reconstructions

Definition (Construction of $(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i}) \in \mathbf{RTN}_l(\mathcal{T}_h)$, $l := m-1|m$)

For all $K \in \mathcal{T}_h$ and all $e \in \mathcal{E}_K$,

$$\langle(\mathbf{d}_h^{k,i}+\mathbf{l}_h^{k,i})\cdot\mathbf{n}_e,q_h\rangle_e := \langle -\{\!\{\boldsymbol{\sigma}^{k-1}(u_h^{k,i},\nabla u_h^{k,i})\}\!\}\cdot\mathbf{n}_e + \bar{\alpha}_e h_e^{-1}[\![u_h^{k,i}]\!],q_h\rangle_e,$$

$$(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i},\mathbf{r}_h)_K := -(\boldsymbol{\sigma}^{k-1}(u_h^{k,i},\nabla u_h^{k,i}),\mathbf{r}_h)_K + \theta \sum_{e\in\mathcal{E}_K} w_e\langle \underline{\mathbf{A}}^{k-1}(u_h^{k,i})\mathbf{r}_h\cdot\mathbf{n}_e,[\![u_h^{k,i}]\!]\rangle_e,$$

for all $q_h \in \mathbb{P}_l(e)$ and all $\mathbf{r}_h \in [\mathbb{P}_{l-1}(K)]^d$.

Definition (Construction of $\mathbf{d}_h^{k,i} \in \mathbf{RTN}_l(\mathcal{T}_h)$, $l := m-1$ or $l := m$)

For all $K \in \mathcal{T}_h$ and all $e \in \mathcal{E}_K$,

$$\langle\mathbf{d}_h^{k,i}\cdot\mathbf{n}_e, q_h\rangle_e := \langle -\{\!\{\boldsymbol{\sigma}(u_h^{k,i},\nabla u_h^{k,i})\}\!\}\cdot\mathbf{n}_e + \bar{\alpha}_e h_e^{-1}[\![u_h^{k,i}]\!], q_h\rangle_e,$$

$$(\mathbf{d}_h^{k,i},\mathbf{r}_h)_K := -(\boldsymbol{\sigma}(u_h^{k,i},\nabla u_h^{k,i}),\mathbf{r}_h)_K + \theta\sum_{e\in\mathcal{E}_K} w_e\langle\underline{\mathbf{A}}(u_h^{k,i})\mathbf{r}_h\cdot\mathbf{n}_e,[\![u_h^{k,i}]\!]\rangle_e,$$

for all $q_h \in \mathbb{P}_l(e)$ and all $\mathbf{r}_h \in [\mathbb{P}_{l-1}(K)]^d$.

Flux reconstructions

Definition (Construction of $(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i}) \in \mathbf{RTN}_l(\mathcal{T}_h)$, $l := m-1|m$)

For all $K \in \mathcal{T}_h$ and all $e \in \mathcal{E}_K$,

$$\langle(\mathbf{d}_h^{k,i}+\mathbf{l}_h^{k,i})\cdot\mathbf{n}_e,q_h\rangle_e := \langle-\{\!\{\boldsymbol{\sigma}^{k-1}(u_h^{k,i},\nabla u_h^{k,i})\}\!\}\cdot\mathbf{n}_e+\bar{\alpha}_e h_e^{-1}[\![u_h^{k,i}]\!],q_h\rangle_e,$$

$$(\mathbf{d}_h^{k,i}+\mathbf{l}_h^{k,i},\mathbf{r}_h)_K := -(\boldsymbol{\sigma}^{k-1}(u_h^{k,i},\nabla u_h^{k,i}),\mathbf{r}_h)_K + \theta\sum_{e\in\mathcal{E}_K} w_e\langle\underline{\mathbf{A}}^{k-1}(u_h^{k,i})\mathbf{r}_h\cdot\mathbf{n}_e,[\![u_h^{k,i}]\!]\rangle_e,$$

for all $q_h \in \mathbb{P}_l(e)$ and all $\mathbf{r}_h \in [\mathbb{P}_{l-1}(K)]^d$.

Definition (Construction of $\mathbf{d}_h^{k,i} \in \mathbf{RTN}_l(\mathcal{T}_h)$, $l := m-1$ or $l := m$)

For all $K \in \mathcal{T}_h$ and all $e \in \mathcal{E}_K$,

$$\langle\mathbf{d}_h^{k,i}\cdot\mathbf{n}_e,q_h\rangle_e := \langle-\{\!\{\boldsymbol{\sigma}(u_h^{k,i},\nabla u_h^{k,i})\}\!\}\cdot\mathbf{n}_e+\bar{\alpha}_e h_e^{-1}[\![u_h^{k,i}]\!],q_h\rangle_e,$$

$$(\mathbf{d}_h^{k,i},\mathbf{r}_h)_K := -(\boldsymbol{\sigma}(u_h^{k,i},\nabla u_h^{k,i}),\mathbf{r}_h)_K + \theta\sum_{e\in\mathcal{E}_K} w_e\langle\underline{\mathbf{A}}(u_h^{k,i})\mathbf{r}_h\cdot\mathbf{n}_e,[\![u_h^{k,i}]\!]\rangle_e,$$

for all $q_h \in \mathbb{P}_l(e)$ and all $\mathbf{r}_h \in [\mathbb{P}_{l-1}(K)]^d$.

Verification of the assumptions – upper bound

Definition (Construction of f_h, $\overline{\boldsymbol{\sigma}}_h^{k,i}$)

Set $f_h := \Pi_l f$ and $\overline{\boldsymbol{\sigma}}_h^{k,i} := \mathbf{I}_l^{\mathrm{RTN}}(\boldsymbol{\sigma}(u_h^{k,i}, \nabla u_h^{k,i}))$.

Lemma (Assumptions A and B)

Assumptions A and B hold.

Comments

- $\|\mathbf{a}_h^{k,i}\|_{q,K} \to 0$ as the linear solver converges by definition.
- $\|\mathbf{l}_h^{k,i}\|_{q,K} \to 0$ as the nonlinear solver converges by the construction of $\mathbf{l}_h^{k,i}$.
- Both $(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i})$ and $\mathbf{d}_h^{k,i}$ belong to $\mathbf{RTN}_l(\mathcal{T}_h) \Rightarrow$ $\mathbf{a}_h^{k,i} \in \mathbf{RTN}_l(\mathcal{T}_h)$ and $\mathbf{t}_h^{k,i} \in \mathbf{RTN}_l(\mathcal{T}_h)$.

Verification of the assumptions – upper bound

Definition (Construction of f_h, $\overline{\boldsymbol{\sigma}}_h^{k,i}$)

Set $f_h := \Pi_l f$ and $\overline{\boldsymbol{\sigma}}_h^{k,i} := \mathbf{I}_l^{\mathrm{RTN}}(\boldsymbol{\sigma}(u_h^{k,i}, \nabla u_h^{k,i}))$.

Lemma (Assumptions A and B)

Assumptions A and B hold.

Comments

- $\|\mathbf{a}_h^{k,i}\|_{q,K} \to 0$ as the linear solver converges by definition.
- $\|\mathbf{l}_h^{k,i}\|_{q,K} \to 0$ as the nonlinear solver converges by the construction of $\mathbf{l}_h^{k,i}$.
- Both $(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i})$ and $\mathbf{d}_h^{k,i}$ belong to $\mathbf{RTN}_l(\mathcal{T}_h) \Rightarrow$ $\mathbf{a}_h^{k,i} \in \mathbf{RTN}_l(\mathcal{T}_h)$ and $\mathbf{t}_h^{k,i} \in \mathbf{RTN}_l(\mathcal{T}_h)$.

Verification of the assumptions – upper bound

Definition (Construction of f_h, $\overline{\boldsymbol{\sigma}}_h^{k,i}$)

Set $f_h := \Pi_l f$ and $\overline{\boldsymbol{\sigma}}_h^{k,i} := \mathbf{I}_l^{\mathrm{RTN}}(\boldsymbol{\sigma}(u_h^{k,i}, \nabla u_h^{k,i}))$.

Lemma (Assumptions A and B)

Assumptions A and B hold.

Comments

- $\|\mathbf{a}_h^{k,i}\|_{q,K} \to 0$ as the linear solver converges by definition.
- $\|\mathbf{l}_h^{k,i}\|_{q,K} \to 0$ as the nonlinear solver converges by the construction of $\mathbf{l}_h^{k,i}$.
- Both $(\mathbf{d}_h^{k,i} + \mathbf{l}_h^{k,i})$ and $\mathbf{d}_h^{k,i}$ belong to $\mathbf{RTN}_l(\mathcal{T}_h) \Rightarrow$ $\mathbf{a}_h^{k,i} \in \mathbf{RTN}_l(\mathcal{T}_h)$ and $\mathbf{t}_h^{k,i} \in \mathbf{RTN}_l(\mathcal{T}_h)$.

Verification of the assumptions – efficiency

Lemma (Assumption C)

Assumption C holds.

Comments

- $\mathbf{d}_h^{k,i}$ close to $\overline{\sigma}_h^{k,i}$
- approximation properties of Raviart–Thomas–Nédélec spaces

Verification of the assumptions – efficiency

Lemma (Assumption C)

Assumption C holds.

Comments

- $\mathbf{d}_h^{k,i}$ close to $\overline{\boldsymbol{\sigma}}_h^{k,i}$
- 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

Outline

1 Introduction

2 Adaptive inexact Newton method
- A guaranteed a posteriori error estimate
- Stopping criteria and efficiency
- Applications
- Numerical results

3 Application to two-phase flow in porous media
- A guaranteed a posteriori error estimate
- Application and numerical results

4 Conclusions and future directions

Numerical experiment I

Model problem

- p-Laplacian

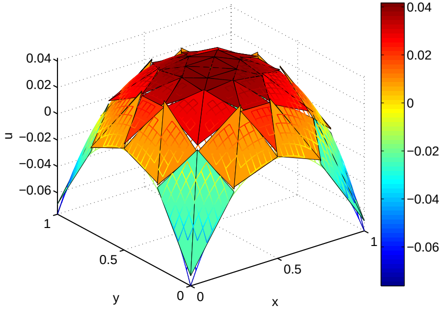
$$\nabla\cdot(|\nabla u|^{p-2}\nabla u) = f \qquad \text{in } \Omega,$$
$$u = u_0 \qquad \text{on } \partial\Omega$$

- weak solution (used to impose the Dirichlet BC)

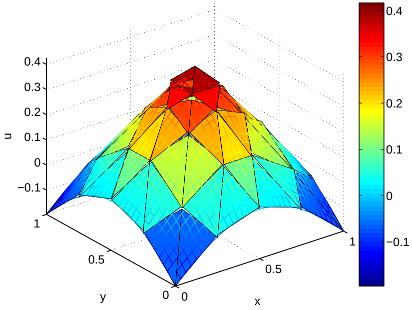
$$u(x,y) = -\frac{p-1}{p}\left((x-\tfrac{1}{2})^2 + (y-\tfrac{1}{2})^2\right)^{\frac{p}{2(p-1)}} + \frac{p-1}{p}\left(\frac{1}{2}\right)^{\frac{p}{p-1}}$$

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

Analytical and approximate solutions

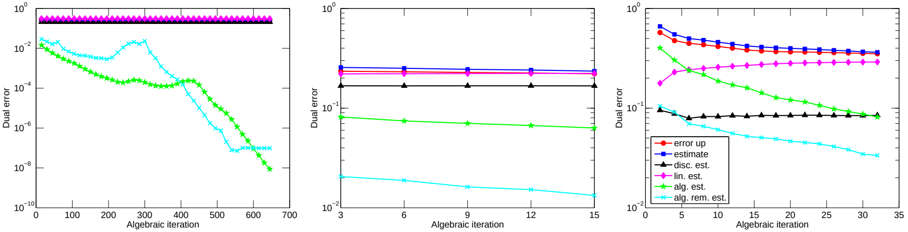


Case $p = 1.5$



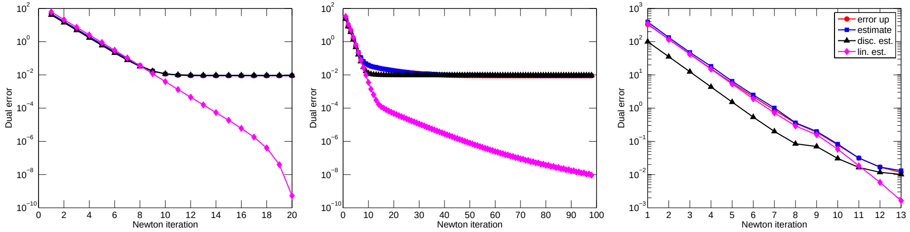
Case $p = 10$

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



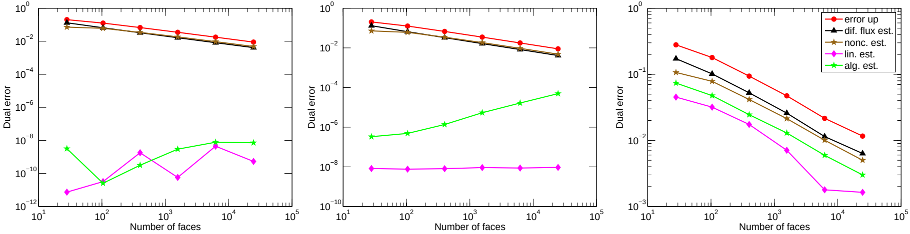
Newton inexact Newton ad. inexact Newton

Error and estimators as a function of Newton iterations, $p = 10$, 6th level mesh



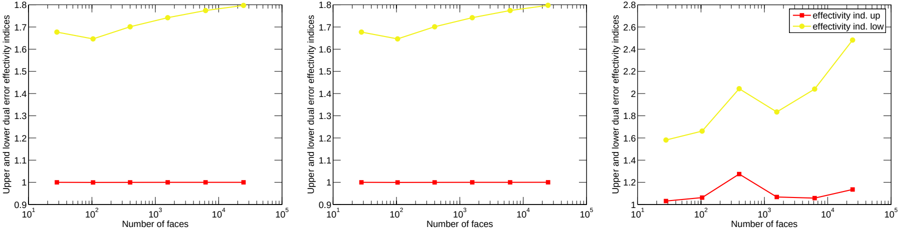
Newton inexact Newton ad. inexact Newton

Error and estimators, $p = 10$



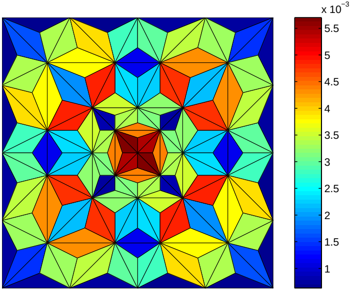
Newton inexact Newton ad. inexact Newton

Effectivity indices, $p = 10$

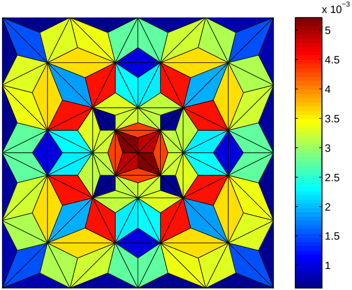


Newton inexact Newton ad. inexact Newton

Error distribution, $p = 10$

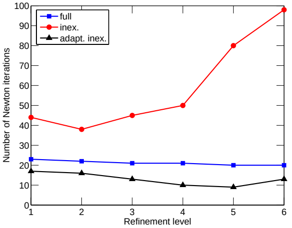


Estimated error distribution

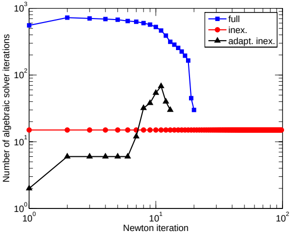


Exact error distribution

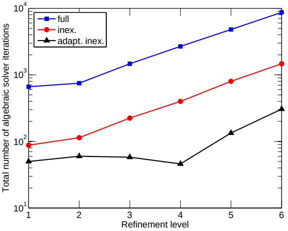
Newton and algebraic iterations, $p = 10$



Newton it. / refinement

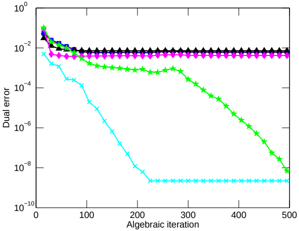


alg. it. / Newton step

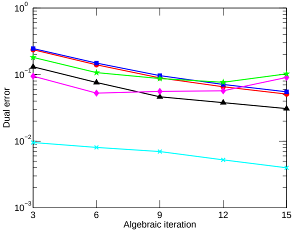


alg. it. / refinement

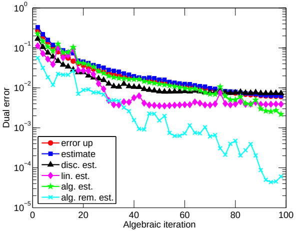
Error and estimators as a function of CG iterations, $p = 1.5$, 6th level mesh, 1st Newton step.



Newton

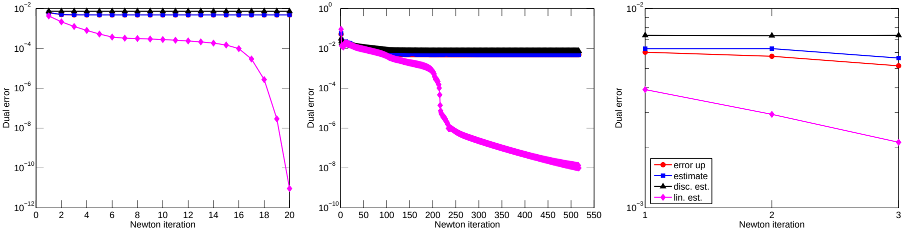


inexact Newton



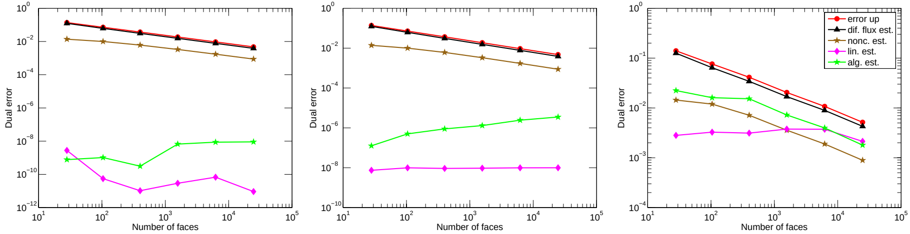
ad. inexact Newton

Error and estimators as a function of Newton iterations, $p = 1.5$, 6th level mesh



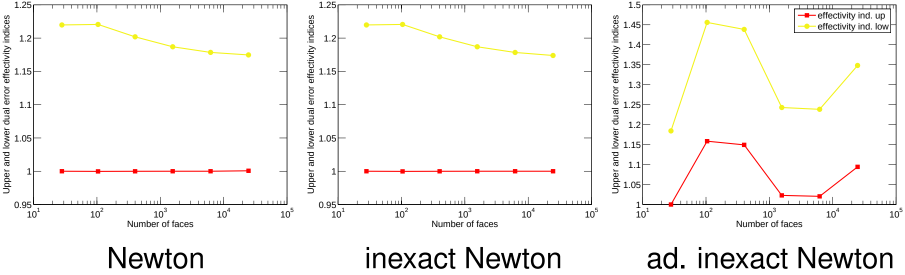
Newton inexact Newton ad. inexact Newton

Error and estimators, $p = 1.5$

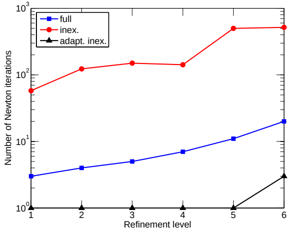


Newton inexact Newton ad. inexact Newton

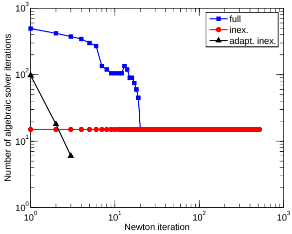
Effectivity indices, $p = 1.5$



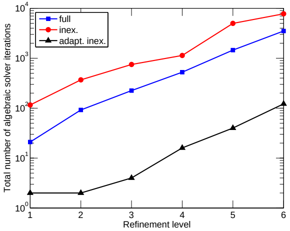
Newton and algebraic iterations, $p = 1.5$



Newton it. / refinement



alg. it. / Newton step



alg. it. / refinement

Numerical experiment II

Model problem

- *p*-Laplacian

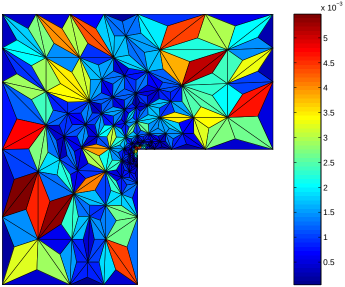
$$\begin{aligned} \nabla\cdot(|\nabla u|^{p-2}\nabla u) &= f \qquad \text{in } \Omega, \\ u &= u_0 \qquad \text{on } \partial\Omega \end{aligned}$$

- weak solution (used to impose the Dirichlet BC)

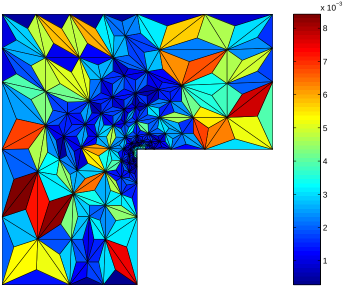
$$u(r,\theta) = r^{\frac{7}{8}} \sin(\theta \tfrac{7}{8})$$

- $p = 4$, L-shape domain, singularity in the origin (Carstensen and Klose (2003))
- nonconforming finite elements

Error distribution on an adaptively refined mesh

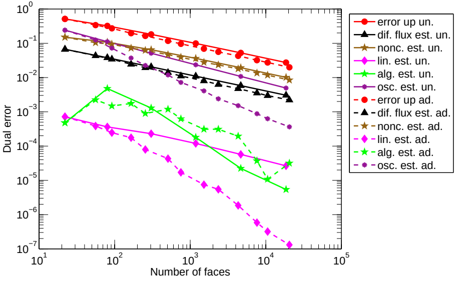


Estimated error distribution

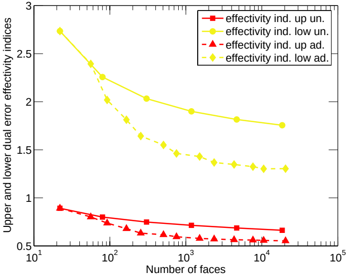


Exact error distribution

Estimated and actual errors and the effectivity index

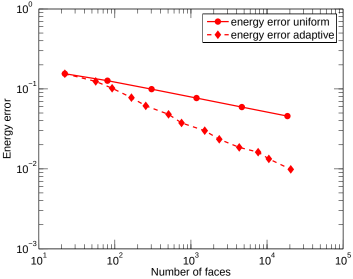


Estimated and actual errors

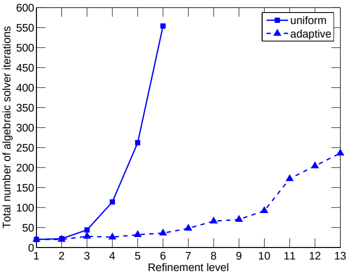


Effectivity index

Energy error and overall performance



Energy error



Overall performance

Outline

1 Introduction

2 Adaptive inexact Newton method
- A guaranteed a posteriori error estimate
- Stopping criteria and efficiency
- Applications
- Numerical results

3 Application to two-phase flow in porous media
- A guaranteed a posteriori error estimate
- Application and numerical results

4 Conclusions and future directions

Two-phase flow in porous media

Two-phase flow in porous media

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

Mathematical issues

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

Two-phase flow in porous media

Two-phase flow in porous media

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

Mathematical issues

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

Outline

1. Introduction
2. Adaptive inexact Newton method
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Applications
   - Numerical results
3. Application to two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Application and numerical results
4. Conclusions and future directions

Two-phase flow in porous media

Theorem (A posteriori error estimate distinguishing the error components)

Let

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

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

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

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

Two-phase flow in porous media

Theorem (A posteriori error estimate distinguishing the error components)

Let

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

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

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

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

Local estimators

- *spatial estimators*

$$\eta_{\mathrm{sp},K}^{n,k,i}(t) := \left\{ \sum_{\alpha\in\{\mathrm{n,w}\}} \left(\|\mathbf{d}_{\alpha,h}^{n,k,i} - \mathbf{v}_\alpha(p_{\mathrm{w},h}^{n,k,i}, s_{\mathrm{w},h}^{n,k,i})\|_K \right.\right.$$
$$+ h_K/\pi \|q_\alpha^n - \partial_t^n(\phi s_{\alpha,h\tau}^{n,k,i}) - \nabla\cdot\mathbf{u}_{\alpha,h}^{n,k,i}\|_K \Big)^2$$
$$+ \left(\|\underline{\mathbf{K}}(\lambda_\mathrm{w}(s_{\mathrm{w},h\tau}^{n,k,i}) + \lambda_\mathrm{n}(s_{\mathrm{w},h\tau}^{n,k,i}))\nabla(\mathfrak{p}(p_{\mathrm{w},h\tau}^{n,k,i}, s_{\mathrm{w},h\tau}^{n,k,i}) - \bar{\mathfrak{p}}_{h\tau}^{n,k,i})\|_K(t)\right)^2$$
$$\left.+ \left(\|\underline{\mathbf{K}}\nabla(\mathfrak{q}(s_{\mathrm{w},h\tau}^{n,k,i}) - \bar{\mathfrak{q}}_{h\tau}^{n,k,i})\|_K(t)\right)^2 \right\}^{\frac{1}{2}}$$

- *temporal estimators*

$$\eta_{\mathrm{tm},K,\alpha}^{n,k,i}(t) := \|\mathbf{v}_\alpha(p_{\mathrm{w},h\tau}^{n,k,i}, s_{\mathrm{w},h\tau}^{n,k,i})(t) - \mathbf{v}_\alpha(p_{\mathrm{w},h\tau}^{n,k,i}, s_{\mathrm{w},h\tau}^{n,k,i})(t^n)\|_K \quad \alpha\in\{\mathrm{n,w}\}$$

- *linearization estimators*

$$\eta_{\mathrm{lin},K,\alpha}^{n,k,i} := \|\mathbf{l}_{\alpha,h}^{n,k,i}\|_K \qquad \alpha\in\{\mathrm{n,w}\}$$

- *algebraic estimators*

$$\eta_{\mathrm{alg},K,\alpha}^{n,k,i} := \|\mathbf{a}_{\alpha,h}^{n,k,i}\|_K \qquad \alpha\in\{\mathrm{n,w}\}$$

Global estimators

Global estimators

$$\eta_{\mathrm{sp}}^{n,k,i} := \left\{ 3 \int_{I_n} \sum_{K \in \mathcal{T}_h^n} (\eta_{\mathrm{sp},K}^{n,k,i}(t))^2 \, \mathrm{d}t \right\}^{\frac{1}{2}},$$

$$\eta_{\mathrm{tm}}^{n,k,i} := \left\{ \sum_{\alpha \in \{\mathrm{n,w}\}} \int_{I_n} \sum_{K \in \mathcal{T}_h^n} (\eta_{\mathrm{tm},K,\alpha}^{n,k,i}(t))^2 \, \mathrm{d}t \right\}^{\frac{1}{2}},$$

$$\eta_{\mathrm{lin}}^{n,k,i} := \left\{ \sum_{\alpha \in \{\mathrm{n,w}\}} \tau^n \sum_{K \in \mathcal{T}_h^n} (\eta_{\mathrm{lin},K,\alpha}^{n,k,i})^2 \right\}^{\frac{1}{2}},$$

$$\eta_{\mathrm{alg}}^{n,k,i} := \left\{ \sum_{\alpha \in \{\mathrm{n,w}\}} \tau^n \sum_{K \in \mathcal{T}_h^n} (\eta_{\mathrm{alg},K,\alpha}^{n,k,i})^2 \right\}^{\frac{1}{2}}$$

Outline

1. Introduction
2. Adaptive inexact Newton method
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Applications
   - Numerical results
3. Application to two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Application and numerical results
4. Conclusions and future directions

Cell-centered finite volume scheme

Cell-centered finite volume scheme

For all $1 \leq n \leq N$, look for $s^n_{\mathrm{w},h}, \bar{p}^n_{\mathrm{w},h}$ such that

$$\phi \frac{s^n_{\mathrm{w},K} - s^{n-1}_{\mathrm{w},K}}{\tau^n}|K| + \sum_{e_{KK'} \in \mathcal{E}^{\mathrm{int}}_K} F_{\mathrm{w},e_{KK'}}(s^n_{\mathrm{w},h}, \bar{p}^n_{\mathrm{w},h}) = 0,$$

$$-\phi \frac{s^n_{\mathrm{w},K} - s^{n-1}_{\mathrm{w},K}}{\tau^n}|K| + \sum_{e_{KK'} \in \mathcal{E}^{\mathrm{int}}_K} F_{\mathrm{n},e_{KK'}}(s^n_{\mathrm{w},h}, \bar{p}^n_{\mathrm{w},h}) = 0,$$

where the fluxes are given by

$$F_{\mathrm{w},e_{KK'}}(s^n_{\mathrm{w},h}, \bar{p}^n_{\mathrm{w},h}) := -\frac{\lambda_{\mathrm{w}}(s^n_{\mathrm{w},K}) + \lambda_{\mathrm{w}}(s^n_{\mathrm{w},K'})}{2}|\underline{\mathbf{K}}| \frac{\bar{p}^n_{\mathrm{w},K'} - \bar{p}^n_{\mathrm{w},K}}{|\mathbf{x}_K - \mathbf{x}_{K'}|}|e_{KK'}|,$$

$$F_{\mathrm{n},e_{KK'}}(s^n_{\mathrm{w},h}, \bar{p}^n_{\mathrm{w},h}) := -\frac{\lambda_{\mathrm{n}}(s^n_{\mathrm{w},K}) + \lambda_{\mathrm{n}}(s^n_{\mathrm{w},K'})}{2}|\underline{\mathbf{K}}| \times \frac{\bar{p}^n_{\mathrm{w},K'} + \pi(s^n_{\mathrm{w},K'}) - (\bar{p}^n_{\mathrm{w},K} + \pi(s^n_{\mathrm{w},K}))}{|\mathbf{x}_K - \mathbf{x}_{K'}|}|e_{KK'}|.$$

Cell-centered finite volume scheme

Cell-centered finite volume scheme

For all $1 \leq n \leq N$, look for $s^n_{\mathrm{w},h}, \bar{p}^n_{\mathrm{w},h}$ such that

$$\phi \frac{s^n_{\mathrm{w},K} - s^{n-1}_{\mathrm{w},K}}{\tau^n}|K| + \sum_{e_{KK'} \in \mathcal{E}^{\mathrm{int}}_K} F_{\mathrm{w},e_{KK'}}(s^n_{\mathrm{w},h}, \bar{p}^n_{\mathrm{w},h}) = 0,$$

$$-\phi \frac{s^n_{\mathrm{w},K} - s^{n-1}_{\mathrm{w},K}}{\tau^n}|K| + \sum_{e_{KK'} \in \mathcal{E}^{\mathrm{int}}_K} F_{\mathrm{n},e_{KK'}}(s^n_{\mathrm{w},h}, \bar{p}^n_{\mathrm{w},h}) = 0,$$

where the fluxes are given by

$$F_{\mathrm{w},e_{KK'}}(s^n_{\mathrm{w},h}, \bar{p}^n_{\mathrm{w},h}) := -\frac{\lambda_{\mathrm{w}}(s^n_{\mathrm{w},K}) + \lambda_{\mathrm{w}}(s^n_{\mathrm{w},K'})}{2}|\underline{\mathbf{K}}| \frac{\bar{p}^n_{\mathrm{w},K'} - \bar{p}^n_{\mathrm{w},K}}{|\mathbf{x}_K - \mathbf{x}_{K'}|}|e_{KK'}|,$$

$$F_{\mathrm{n},e_{KK'}}(s^n_{\mathrm{w},h}, \bar{p}^n_{\mathrm{w},h}) := -\frac{\lambda_{\mathrm{n}}(s^n_{\mathrm{w},K}) + \lambda_{\mathrm{n}}(s^n_{\mathrm{w},K'})}{2}|\underline{\mathbf{K}}| \times \frac{\bar{p}^n_{\mathrm{w},K'} + \pi(s^n_{\mathrm{w},K'}) - (\bar{p}^n_{\mathrm{w},K} + \pi(s^n_{\mathrm{w},K}))}{|\mathbf{x}_K - \mathbf{x}_{K'}|}|e_{KK'}|.$$

Linearization and algebraic solution

Linearization step k and algebraic step i

Couple $s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}$ such that

$$\phi \frac{s_{\mathrm{w},K}^{n,k,i} - s_{\mathrm{w},K}^{n-1}}{\tau^n}|K| + \sum_{e_{KK'} \in \mathcal{E}_K^{\mathrm{int}}} F_{\mathrm{w},e_{KK'}}^{k-1}(s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}) = -R_{\mathrm{w},K}^{n,k,i},$$

$$-\phi \frac{s_{\mathrm{w},K}^{n,k,i} - s_{\mathrm{w},K}^{n-1}}{\tau^n}|K| + \sum_{e_{KK'} \in \mathcal{E}_K^{\mathrm{int}}} F_{\mathrm{n},e_{KK'}}^{k-1}(s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}) = -R_{\mathrm{n},K}^{n,k,i},$$

where the linearized fluxes are given by

$$F_{\alpha,e_{KK'}}^{k-1}(s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}) := F_{\alpha,e_{KK'}}(s_{\mathrm{w},h}^{n,k-1}, \bar{p}_{\mathrm{w},h}^{n,k-1}) + \sum_{M \in \{K,K'\}} \frac{\partial F_{\alpha,e_{KK'}}}{\partial s_{\mathrm{w},M}}(s_{\mathrm{w},h}^{n,k-1}, \bar{p}_{\mathrm{w},h}^{n,k-1}) \cdot (s_{\mathrm{w},M}^{n,k,i} - s_{\mathrm{w},M}^{n,k-1}) + \sum_{M \in \{K,K'\}} \frac{\partial F_{\alpha,e_{KK'}}}{\partial \bar{p}_{\mathrm{w},M}}(s_{\mathrm{w},h}^{n,k-1}, \bar{p}_{\mathrm{w},h}^{n,k-1}) \cdot (\bar{p}_{\mathrm{w},M}^{n,k,i} - \bar{p}_{\mathrm{w},M}^{n,k-1}).$$

Linearization and algebraic solution

Linearization step k and algebraic step i

Couple $s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}$ such that

$$\phi \frac{s_{\mathrm{w},K}^{n,k,i} - s_{\mathrm{w},K}^{n-1}}{\tau^n}|K| + \sum_{e_{KK'} \in \mathcal{E}_K^{\mathrm{int}}} F_{\mathrm{w},e_{KK'}}^{k-1}(s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}) = -R_{\mathrm{w},K}^{n,k,i},$$

$$-\phi \frac{s_{\mathrm{w},K}^{n,k,i} - s_{\mathrm{w},K}^{n-1}}{\tau^n}|K| + \sum_{e_{KK'} \in \mathcal{E}_K^{\mathrm{int}}} F_{\mathrm{n},e_{KK'}}^{k-1}(s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}) = -R_{\mathrm{n},K}^{n,k,i},$$

where the linearized fluxes are given by

$$\begin{aligned} F_{\alpha,e_{KK'}}^{k-1}(s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}) :=& F_{\alpha,e_{KK'}}(s_{\mathrm{w},h}^{n,k-1}, \bar{p}_{\mathrm{w},h}^{n,k-1}) \\ &+ \sum_{M \in \{K,K'\}} \frac{\partial F_{\alpha,e_{KK'}}}{\partial s_{\mathrm{w},M}}(s_{\mathrm{w},h}^{n,k-1}, \bar{p}_{\mathrm{w},h}^{n,k-1}) \cdot (s_{\mathrm{w},M}^{n,k,i} - s_{\mathrm{w},M}^{n,k-1}) \\ &+ \sum_{M \in \{K,K'\}} \frac{\partial F_{\alpha,e_{KK'}}}{\partial \bar{p}_{\mathrm{w},M}}(s_{\mathrm{w},h}^{n,k-1}, \bar{p}_{\mathrm{w},h}^{n,k-1}) \cdot (\bar{p}_{\mathrm{w},M}^{n,k,i} - \bar{p}_{\mathrm{w},M}^{n,k-1}). \end{aligned}$$

Fluxes reconstructions and pressure postprocessing

Fluxes reconstructions

$$(\mathbf{d}_{\alpha,h}^{n,k,i}\cdot\mathbf{n}_K, 1)_{e_{KK'}} := F_{\alpha,e_{KK'}}(s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}),$$

$$((\mathbf{d}_{\alpha,h}^{n,k,i} + \mathbf{l}_{\alpha,h}^{n,k,i})\cdot\mathbf{n}_K, 1)_{e_{KK'}} := F_{\alpha,e_{KK'}}^{k-1}(s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}),$$

$$\mathbf{a}_{\alpha,h}^{n,k,i} := \mathbf{d}_{\alpha,h}^{n,k,i+\nu} + \mathbf{l}_{\alpha,h}^{n,k,i+\nu} - (\mathbf{d}_{\alpha,h}^{n,k,i} + \mathbf{l}_{\alpha,h}^{n,k,i})$$

Phase pressures postprocessing

- Piecewise constant $\bar{p}_{\alpha,h}^{n,k,i}$ postprocessed to piecewise quadratic $p_{\alpha,h}^{n,k,i}$:

$$-\lambda_{\mathrm{w}}(s_{\mathrm{w},K}^{n,k,i})\underline{\mathbf{K}}\nabla(p_{\mathrm{w},h}^{n,k,i}|_K) = \mathbf{d}_{\mathrm{w},h}^{n,k,i}|_K,$$

$$p_{\mathrm{w},h}^{n,k,i}(\mathbf{x}_K) = \bar{p}_{\mathrm{w},K}^{n,k,i},$$

$$-\lambda_{\mathrm{n}}(s_{\mathrm{w},K}^{n,k,i})\underline{\mathbf{K}}\nabla(p_{\mathrm{n},h}^{n,k,i}|_K) = \mathbf{d}_{\mathrm{n},h}^{n,k,i}|_K,$$

$$p_{\mathrm{n},h}^{n,k,i}(\mathbf{x}_K) = \pi(s_{\mathrm{w},K}^{n,k,i}) + \bar{p}_{\mathrm{w},K}^{n,k,i}$$

Fluxes reconstructions and pressure postprocessing

Fluxes reconstructions

$$(\mathbf{d}_{\alpha,h}^{n,k,i} \cdot \mathbf{n}_K, 1)_{e_{KK'}} := F_{\alpha,e_{KK'}}(s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}),$$

$$((\mathbf{d}_{\alpha,h}^{n,k,i} + \mathbf{l}_{\alpha,h}^{n,k,i}) \cdot \mathbf{n}_K, 1)_{e_{KK'}} := F_{\alpha,e_{KK'}}^{k-1}(s_{\mathrm{w},h}^{n,k,i}, \bar{p}_{\mathrm{w},h}^{n,k,i}),$$

$$\mathbf{a}_{\alpha,h}^{n,k,i} := \mathbf{d}_{\alpha,h}^{n,k,i+\nu} + \mathbf{l}_{\alpha,h}^{n,k,i+\nu} - (\mathbf{d}_{\alpha,h}^{n,k,i} + \mathbf{l}_{\alpha,h}^{n,k,i})$$

Phase pressures postprocessing

- Piecewise constant $\bar{p}_{\alpha,h}^{n,k,i}$ postprocessed to piecewise quadratic $p_{\alpha,h}^{n,k,i}$:

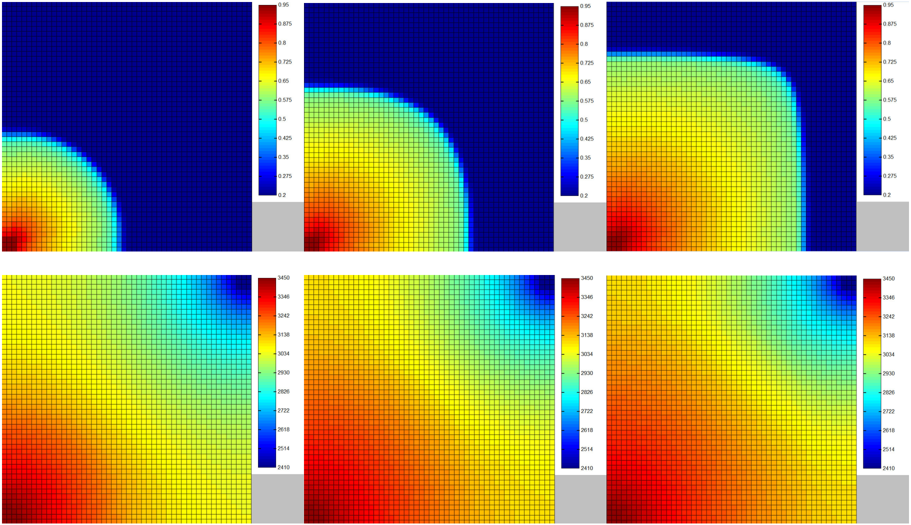
$$-\lambda_{\mathrm{w}}(s_{\mathrm{w},K}^{n,k,i})\underline{\mathbf{K}}\nabla(p_{\mathrm{w},h}^{n,k,i}|_K) = \mathbf{d}_{\mathrm{w},h}^{n,k,i}|_K,$$

$$p_{\mathrm{w},h}^{n,k,i}(\mathbf{x}_K) = \bar{p}_{\mathrm{w},K}^{n,k,i},$$

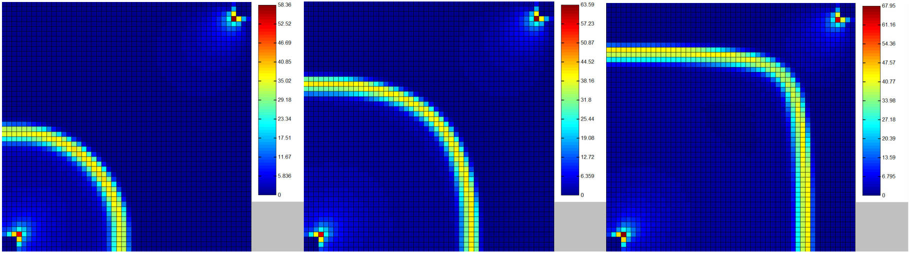
$$-\lambda_{\mathrm{n}}(s_{\mathrm{w},K}^{n,k,i})\underline{\mathbf{K}}\nabla(p_{\mathrm{n},h}^{n,k,i}|_K) = \mathbf{d}_{\mathrm{n},h}^{n,k,i}|_K,$$

$$p_{\mathrm{n},h}^{n,k,i}(\mathbf{x}_K) = \pi(s_{\mathrm{w},K}^{n,k,i}) + \bar{p}_{\mathrm{w},K}^{n,k,i}$$

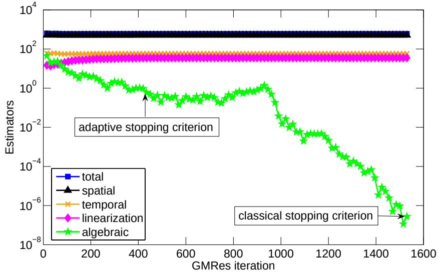
Water saturation/water pressure evolution



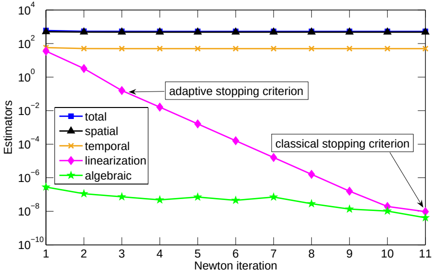
Estimators evolution



Estimators and stopping criteria

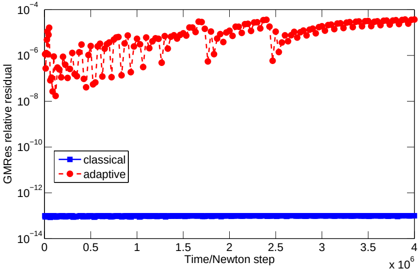


Estimators in function of GMRes iterations

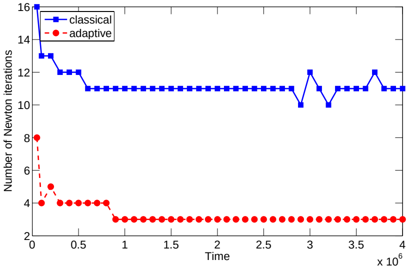


Estimators in function of Newton iterations

GMRes relative residual/Newton iterations

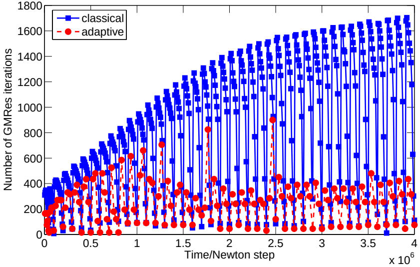


GMRes relative residual

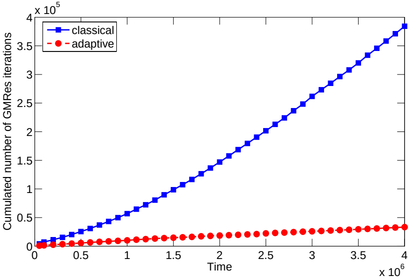


Newton iterations

GMRes iterations



Per time and Newton step



Cumulated

Outline

1. Introduction
2. Adaptive inexact Newton method
   - A guaranteed a posteriori error estimate
   - Stopping criteria and efficiency
   - Applications
   - Numerical results
3. Application to two-phase flow in porous media
   - A guaranteed a posteriori error estimate
   - Application and numerical results
4. Conclusions and future directions

Conclusions

Entire adaptivity

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

Future directions

- other coupled nonlinear systems
- convergence and optimality

Conclusions

Entire adaptivity

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

Future directions

- other coupled nonlinear systems
- convergence and optimality

Bibliography

Bibliography

- JIRÁNEK P., STRAKOŠ Z., VOHRALÍK M., A posteriori error estimates including algebraic error and stopping criteria for iterative solvers. *SIAM J.Sci.Comput.* **32** (2010),1567–1590.
- EL ALAOUI L., ERN A., VOHRALÍK M., Guaranteed and robust a posteriori error estimates and balancing discretization and linearization errors for monotone nonlinear problems. *Comput. Methods Appl. Mech. Engrg.* **200** (2011), 2782–2795.
- ERN A., VOHRALÍK M., Adaptive inexact Newton methods with a posteriori stopping criteria for nonlinear diffusion PDEs, HAL Preprint 00681422.
- VOHRALÍK M., WHEELER M. F., A posteriori error estimates, stopping criteria, and adaptivity for two-phase flows, HAL Preprint 00633594.

Merci de votre attention !