A posteriori error estimates robust with respect to nonlinearities and final time

Martin Vohralík

in collaboration with André Harnist, Koondanibha Mitra, and Ari Rappaport

Inria Paris & Ecole des Ponts

Meknès, October 17, 2023

Outline

Modelling flow of water and air through soil

The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\boldsymbol{K}\kappa(S(u))(\nabla u + \boldsymbol{g})] &= f \quad \text{in } \Omega \times (0, T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Modelling flow of water and air through soil

The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\boldsymbol{K}\kappa(S(u))(\nabla u + \boldsymbol{g})] &= f \quad \text{in } \Omega \times (0, T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Setting

- u: pressure
- $s = S(u)$: saturation
- $\Omega \subset \mathbb{R}^d$, $1 \leq d \leq 3$, open bounded polytope with Lipschitz boundary $\partial\Omega$
- T: final time
- diffusion tensor $\boldsymbol{K}$, source term f, gravity $\boldsymbol{g}$, initial saturation s_0
- **nonlinear (degenerate) functions** S and κ

Modelling flow of water and air through soil

The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\boldsymbol{K}\kappa(S(u))(\nabla u + \boldsymbol{g})] &= f \quad \text{in } \Omega \times (0, T),\\ u &= 0 \quad \text{on } \partial\Omega \times (0, T),\\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Setting

- u: pressure
- $s = S(u)$: saturation
- $\Omega \subset \mathbb{R}^d$, $1 \le d \le 3$, open bounded polytope with Lipschitz boundary $\partial\Omega$
- T: final time
- diffusion tensor $\boldsymbol{K}$, source term f, gravity $\boldsymbol{g}$, initial saturation s_0
- **nonlinear (degenerate) functions** S and κ

Modelling flow of water and air through soil

The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\boldsymbol{K}\kappa(S(u))(\nabla u + \boldsymbol{g})] &= f \quad \text{in } \Omega \times (0, T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Setting

- u: pressure
- $s = S(u)$: saturation
- $\Omega \subset \mathbb{R}^d$, $1 \leq d \leq 3$, open bounded polytope with Lipschitz boundary $\partial\Omega$
- T: final time
- diffusion tensor $\boldsymbol{K}$, source term f, gravity $\boldsymbol{g}$, initial saturation s_0
- **nonlinear (degenerate) functions** S and κ

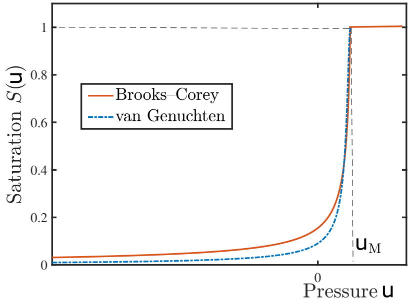
Modelling flow of water and air through soil

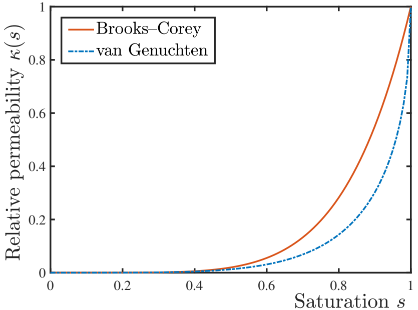
The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\mathbf{K}\kappa(S(u))(\nabla u + \mathbf{g})] &= f && \text{in } \Omega \times (0,T), \\ u &= 0 && \text{on } \partial\Omega \times (0,T), \\ (S(u))(0) &= s_0 && \text{in } \Omega. \end{aligned}$$

Nonlinear (degenerate) functions S and κ





Modelling flow of water and air through soil

The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\boldsymbol{K}\kappa(S(u))(\nabla u + \boldsymbol{g})] &= f \quad \text{in } \Omega \times (0, T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Degeneracies

- parabolic–hyperbolic: $\kappa(0) = 0$ leads to

$$\partial_t S(u) = f$$

Modelling flow of water and air through soil

The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\mathbf{K}\kappa(S(u))(\nabla u + \mathbf{g})] &= f \quad \text{in } \Omega \times (0, T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Degeneracies

- parabolic–hyperbolic: $\kappa(0) = 0$ leads to

$$\partial_t S(u) = f$$

- parabolic–elliptic: $S'(u) = 0$ for $u > u_{\mathrm{M}}$ leads to

$$-\nabla\cdot[\mathbf{K}\kappa(S(u))(\nabla u + \mathbf{g})] = f$$

A posteriori error estimates

Purpose

- provide sharp **computable bounds** on the unknown error between the unavailable exact solution and its numerical approximation
- predict the **error localization** (in space and in time)
- **adapt** the regularization parameters, linear solver, nonlinear solver, space mesh, time mesh ...

A posteriori error estimates

Purpose

- provide sharp **computable bounds** on the unknown error between the unavailable exact solution and its numerical approximation
- predict the **error localization** (in space and in time)
- **adapt** the regularization parameters, linear solver, nonlinear solver, space mesh, time mesh . . .

A posteriori error estimates

Purpose

- provide sharp **computable bounds** on the unknown error between the unavailable exact solution and its numerical approximation
- predict the **error localization** (in space and in time)
- **adapt** the regularization parameters, linear solver, nonlinear solver, space mesh, time mesh . . .

Goals

Nonlinear problems

a posteriori error estimates

$$|||u - u_\ell||| \leq \eta(u_\ell)$$

Goals

Nonlinear problems

Guaranteed a posteriori error estimates

$$|||u - u_\ell||| \leq \eta(u_\ell)$$

Goals

Nonlinear problems

Guaranteed a posteriori error estimates **efficient**

$$|||u - u_\ell||| \leq \eta(u_\ell) \leq C_{\text{eff}} |||u - u_\ell|||,$$

Goals

Nonlinear problems

Guaranteed a posteriori error estimates **efficient** and **robust** with respect to the **strength of nonlinearities**.

$$|||u - u_\ell||| \leq \eta(u_\ell) \leq C_{\mathrm{eff}} |||u - u_\ell|||, \quad C_{\mathrm{eff}} \text{ independent of nonlinearities}$$

Goals

Nonlinear problems

Guaranteed a posteriori error estimates **efficient** and **robust** with respect to the **strength of nonlinearities**.

$$|||u - u_\ell||| \leq \left\{ \sum_{K \in \mathcal{T}_\ell} \eta_K(u_\ell)^2 \right\}^{1/2} \leq C_{\text{eff}} |||u - u_\ell|||,$$

Goals

Nonlinear problems

Guaranteed a posteriori error estimates **locally efficient** and **robust** with respect to the **strength of nonlinearities**.

$$\eta_K(u_\ell) \leq C_{\text{eff}} |||u - u_\ell|||_{\omega_K}, \qquad \text{for all } K \in \mathcal{T}_\ell$$

Goals

Nonlinear problems

Guaranteed a posteriori error estimates **locally efficient** and **robust** with respect to the **strength of nonlinearities**.

$$\eta_K(u_\ell) \leq C_{\text{eff}} |||u - u_\ell|||_{\omega_K}, \qquad \text{for all } K \in \mathcal{T}_\ell$$

Unsteady problems

Guaranteed a posteriori error estimates

$$\int_0^T |||u - u_\ell|||^2 \leq \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_\ell^n} \eta_K^n(u_\ell)^2$$

Goals

Nonlinear problems

Guaranteed a posteriori error estimates **locally efficient** and **robust** with respect to the **strength of nonlinearities**.

$$\eta_K(u_\ell) \leq C_{\text{eff}} |||u - u_\ell|||_{\omega_K}, \qquad \text{for all } K \in \mathcal{T}_\ell$$

Unsteady problems

Guaranteed a posteriori error estimates **efficient**

$$\int_0^T |||u - u_\ell|||^2 \leq \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_\ell^n} \eta_K^n(u_\ell)^2 \leq C_{\text{eff}}^2 \int_0^T |||u - u_\ell|||^2,$$

Goals

Nonlinear problems

Guaranteed a posteriori error estimates **locally efficient** and **robust** with respect to the **strength of nonlinearities**.

$$\eta_K(u_\ell) \le C_{\mathrm{eff}} |||u - u_\ell|||_{\omega_K}, \qquad \text{for all } K \in \mathcal{T}_\ell$$

Unsteady problems

Guaranteed a posteriori error estimates **efficient** and **robust** with respect to the **final time**.

$$\int_0^T |||u-u_\ell|||^2 \le \sum_{n=1}^{N} \sum_{K\in\mathcal{T}_\ell^n} \eta_K^n(u_\ell)^2 \le C_{\mathrm{eff}}^2 \int_0^T |||u-u_\ell|||^2, \quad C_{\mathrm{eff}} \text{ independent of } T$$

Goals

Nonlinear problems

Guaranteed a posteriori error estimates **locally efficient** and **robust** with respect to the **strength of nonlinearities**.

$$\eta_K(u_\ell) \le C_{\mathrm{eff}} |||u - u_\ell|||_{\omega_K}, \qquad \text{for all } K \in \mathcal{T}_\ell$$

Unsteady problems

Guaranteed a posteriori error estimates **locally space-time efficient** and **robust** with respect to the **final time**.

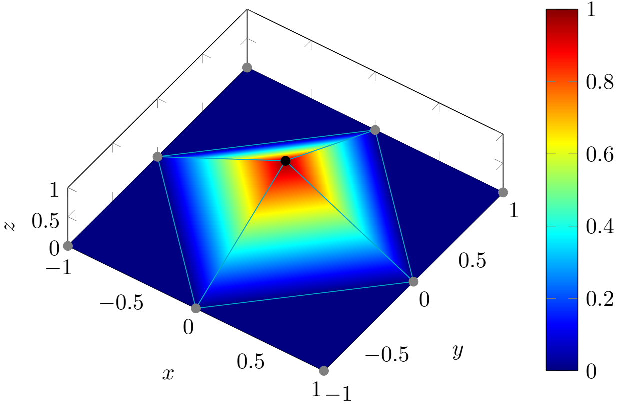
$$\eta_K^n(u_\ell)^2 \le C_{\mathrm{eff}}^2 \int_{t^{n-1}}^{t^n} |||u - u_\ell|||_{\omega_K}^2, \quad \text{for all } n \text{ and } K \in \mathcal{T}_\ell^n$$

Outline

Outline

Partition of unity

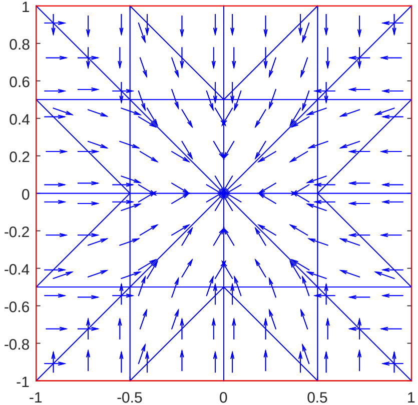
$$\sum_{\boldsymbol{a}\in\mathcal{V}_\ell} \psi^{\boldsymbol{a}} = 1$$



Hat basis function $\psi^{\boldsymbol{a}}$

Equilibrated flux reconstruction

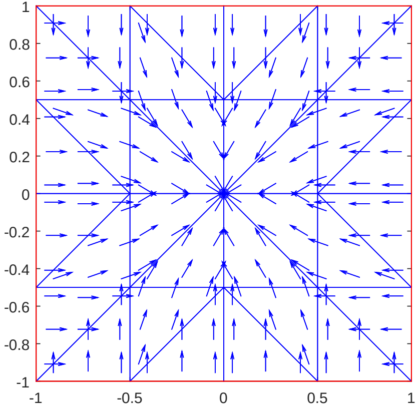
Destuynder and Métivet (1998), Braess & Schöberl (2008), Ern & Vohralík (2013)



Flux $\boldsymbol{\iota}_\ell \notin \boldsymbol{H}(\mathrm{div})$ (e.g. FE flux $-\nabla u_\ell$)

Equilibrated flux reconstruction

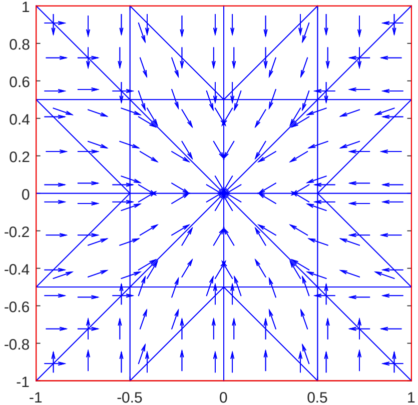
Destuynder and Métivet (1998), Braess & Schöberl (2008), Ern & Vohralík (2013)



Flux $\boldsymbol{\iota}_\ell \notin \boldsymbol{H}(\mathrm{div})$, $\nabla \cdot \boldsymbol{\iota}_\ell \neq f$

Equilibrated flux reconstruction

Destuynder and Métivet (1998), Braess & Schöberl (2008), Ern & Vohralík (2013)

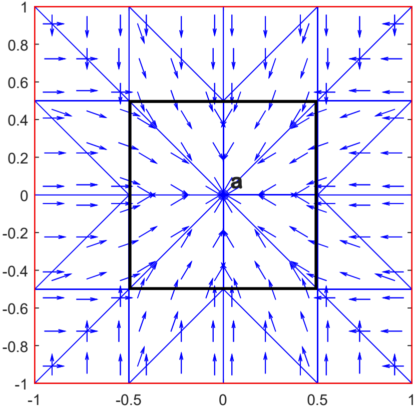


Flux $\boldsymbol{\iota}_\ell \notin \mathbf{H}(\mathrm{div})$, $\nabla\cdot\boldsymbol{\iota}_\ell \neq f$

$$\underbrace{\boldsymbol{\iota}_\ell \in \mathcal{RT}_p(\mathcal{T}_\ell), f \in \mathcal{P}_p(\mathcal{T}_\ell)}$$

Equilibrated flux reconstruction

Destuynder and Métivet (1998), Braess & Schöberl (2008), Ern & Vohralík (2013)

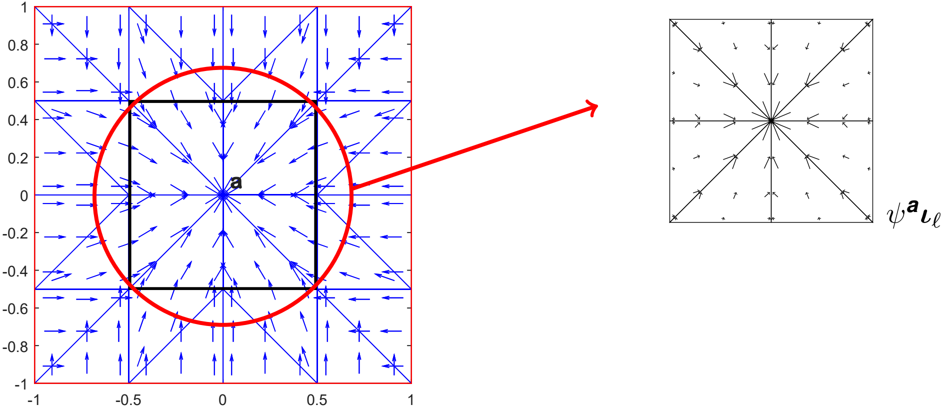


Flux $\boldsymbol{\iota}_\ell \notin \boldsymbol{H}(\mathrm{div})$, $\nabla \cdot \boldsymbol{\iota}_\ell \neq f$

$$\underbrace{\boldsymbol{\iota}_\ell \in \mathcal{RT}_p(\mathcal{T}_\ell), f \in \mathcal{P}_p(\mathcal{T}_\ell)}_{(f, \psi^{\boldsymbol{a}})_{\omega_{\boldsymbol{a}}} + (\boldsymbol{\iota}_\ell, \nabla \psi^{\boldsymbol{a}})_{\omega_{\boldsymbol{a}}} = 0 \ \forall \boldsymbol{a} \in \mathcal{V}_\ell^{\mathrm{int}}}$$

Equilibrated flux reconstruction

Destuynder and Métivet (1998), Braess & Schöberl (2008), Ern & Vohralík (2013)

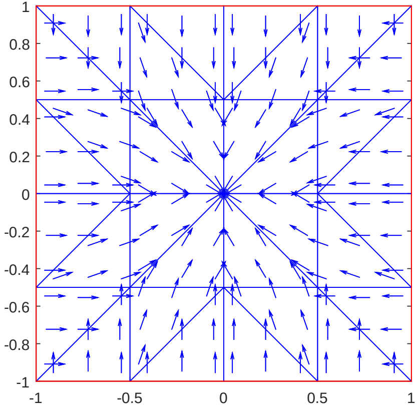


Flux $\boldsymbol{\iota}_\ell \notin \boldsymbol{H}(\mathrm{div})$, $\nabla\cdot\boldsymbol{\iota}_\ell \neq f$

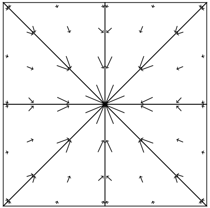
$$\underbrace{\boldsymbol{\iota}_\ell \in \mathcal{RT}_p(\mathcal{T}_\ell),\, f \in \mathcal{P}_p(\mathcal{T}_\ell)}$$

Equilibrated flux reconstruction

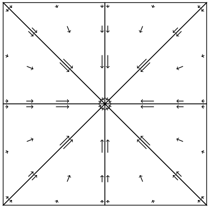
Destuynder and Métivet (1998), Braess & Schöberl (2008), Ern & Vohralík (2013)



Flux $\boldsymbol{\iota}_\ell \notin \mathbf{H}(\mathrm{div})$, $\nabla \cdot \boldsymbol{\iota}_\ell \neq f$



$\psi^{\mathbf{a}} \boldsymbol{\iota}_\ell$

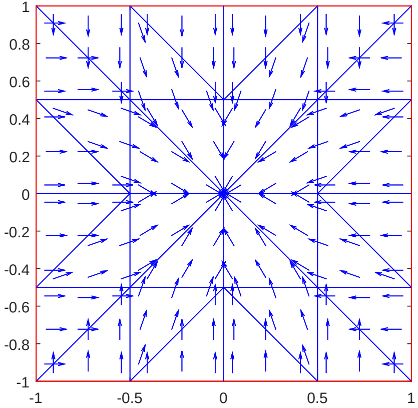


$\boldsymbol{\sigma}_\ell^{\mathbf{a}}$

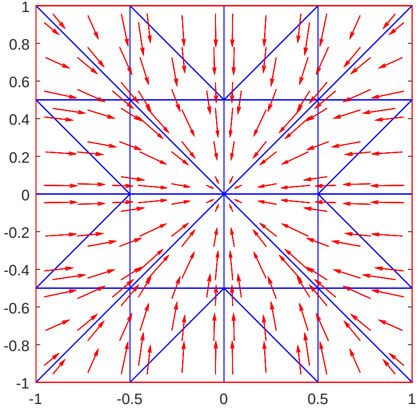
$$\underbrace{\boldsymbol{\iota}_\ell \in \mathcal{RT}_p(\mathcal{T}_\ell),\, f \in \mathcal{P}_p(\mathcal{T}_\ell)} \qquad \boldsymbol{\sigma}_\ell^{\mathbf{a}} := \arg \min_{\substack{\mathbf{v}_\ell \in \mathcal{RT}_{p+1}(\mathcal{T}_{\mathbf{a}}) \cap \mathbf{H}_0(\mathrm{div},\omega_{\mathbf{a}}) \\ \nabla \cdot \mathbf{v}_\ell = f\psi^{\mathbf{a}} + \boldsymbol{\iota}_\ell \cdot \nabla \psi^{\mathbf{a}}}} \|\psi^{\mathbf{a}} \boldsymbol{\iota}_\ell - \mathbf{v}_\ell\|^2_{\omega_{\mathbf{a}}}$$

Equilibrated flux reconstruction

Destuynder and Métivet (1998), Braess & Schöberl (2008), Ern & Vohralík (2013)



Flux $\boldsymbol{\iota}_\ell \notin \boldsymbol{H}(\mathrm{div})$, $\nabla\cdot\boldsymbol{\iota}_\ell \neq f$

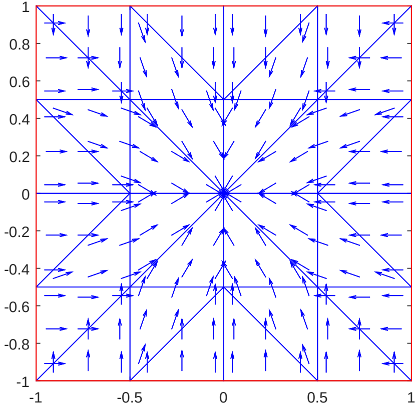


Equilibrated flux $\boldsymbol{\sigma}_\ell \in \boldsymbol{H}(\mathrm{div})$, $\nabla\cdot\boldsymbol{\sigma}_\ell = f$

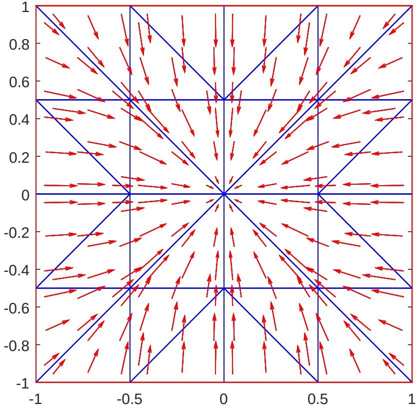
$$\underbrace{\boldsymbol{\iota}_\ell \in \mathcal{RT}_p(\mathcal{T}_\ell), f \in \mathcal{P}_p(\mathcal{T}_\ell)} \to \boldsymbol{\sigma}_\ell := \sum_{\boldsymbol{a}\in\mathcal{V}_\ell} \boldsymbol{\sigma}_\ell^{\boldsymbol{a}} \in \mathcal{RT}_{p+1}(\mathcal{T}_\ell) \cap \boldsymbol{H}(\mathrm{div}), \nabla\cdot\boldsymbol{\sigma}_\ell = f$$

Equilibrated flux reconstruction

Destuynder and Métivet (1998), Braess & Schöberl (2008), Ern & Vohralík (2013)



Flux $\iota_\ell \notin \boldsymbol{H}(\mathrm{div})$, $\nabla\cdot\iota_\ell \neq f$



Equilibrated flux $\sigma_\ell \in \boldsymbol{H}(\mathrm{div})$, $\nabla\cdot\sigma_\ell = f$

Equilibrated flux reconstruction

Use

- **a posteriori error estimates**
  - comparison of the original & reconstructed flux $\|\nabla u_\ell + \sigma_\ell\|$: discretization error
  - error component fluxes: linearization and algebraic errors
- recovery of **mass conservative fluxes**
  - local on patches of mesh elements from FE-type approximations
  - local on elements from FV- & DG-type approximations
  - inexact nonlinear solvers (still local)
  - inexact linear solvers (price of one MG iteration)

Equilibrated flux reconstruction

Use

- **a posteriori error estimates**
  - comparison of the original & reconstructed flux $\|\nabla u_\ell + \boldsymbol{\sigma}_\ell\|$: discretization error
  - error component fluxes: linearization and algebraic errors
- recovery of **mass conservative fluxes**
  - local on patches of mesh elements from FE-type approximations
  - local on elements from FV- & DG-type approximations
  - inexact nonlinear solvers (still local)
  - inexact linear solvers (price of one MG iteration)

Equilibrated flux reconstruction

Use

- **a posteriori error estimates**
  - comparison of the original & reconstructed flux $\|\nabla u_\ell + \boldsymbol{\sigma}_\ell\|$: discretization error
  - error component fluxes: linearization and algebraic errors
- recovery of **mass conservative fluxes**
  - local on patches of mesh elements from FE-type approximations
  - local on elements from FV- & DG-type approximations
  - inexact nonlinear solvers (still local)
  - inexact linear solvers (price of one MG iteration)

Equilibrated flux reconstruction

Use

- **a posteriori error estimates**
  - comparison of the original & reconstructed flux $\|\nabla u_\ell + \boldsymbol{\sigma}_\ell\|$: discretization error
  - error component fluxes: linearization and algebraic errors
- recovery of **mass conservative fluxes**
  - local on patches of mesh elements from FE-type approximations
  - local on elements from FV- & DG-type approximations
  - inexact nonlinear solvers (still local)
  - inexact linear solvers (price of one MG iteration)

Equilibrated flux reconstruction

Use

- **a posteriori error estimates**
  - comparison of the original & reconstructed flux $\|\nabla u_\ell + \boldsymbol{\sigma}_\ell\|$: discretization error
  - error component fluxes: linearization and algebraic errors
- recovery of **mass conservative fluxes**
  - local on patches of mesh elements from FE-type approximations
  - local on elements from FV- & DG-type approximations
  - inexact nonlinear solvers (still local)
  - inexact linear solvers (price of one MG iteration)

Equilibrated flux reconstruction

Use

- **a posteriori error estimates**
  - comparison of the original & reconstructed flux $\|\nabla u_\ell + \boldsymbol{\sigma}_\ell\|$: discretization error
  - error component fluxes: linearization and algebraic errors
- recovery of **mass conservative fluxes**
  - local on patches of mesh elements from FE-type approximations
  - local on elements from FV- & DG-type approximations
  - inexact nonlinear solvers (still local)
  - inexact linear solvers (price of one MG iteration)

Equilibrated flux reconstruction

Use

- **a posteriori error estimates**
  - comparison of the original & reconstructed flux $\|\nabla u_\ell + \boldsymbol{\sigma}_\ell\|$: discretization error
  - error component fluxes: linearization and algebraic errors
- recovery of **mass conservative fluxes**
  - local on patches of mesh elements from FE-type approximations
  - local on elements from FV- & DG-type approximations
  - inexact nonlinear solvers (still local)
  - inexact linear solvers (price of one MG iteration)

Outline

Outline

How **large** is the error? (steady linear Darcy, known solution)

$h_\ell \approx 1/\lvert\mathcal{T}_\ell\rvert^{\frac{1}{2}}$	p	relative error estimate $\frac{\eta(u_\ell)}{\lVert\nabla u_\ell\rVert}$	relative error $\frac{\lVert\nabla(u-u_\ell)\rVert}{\lVert\nabla u_\ell\rVert}$	effectivity index $\frac{\eta(u_\ell)}{\lVert\nabla(u-u_\ell)\rVert}$
h_0	1	28%	24%	1.17
$\approx h_0/2$		14%	13%	
$\approx h_0/4$		7.0%	6.6%	
$\approx h_0/8$		3.3%	3.1%	
$\approx h_0/2$	2	9.5×10^{-1}%	9.2×10^{-1}%	
$\approx h_0/4$	3	5.9×10^{-3}%	5.9×10^{-3}%	
$\approx h_0/8$	4	5.9×10^{-5}%	5.8×10^{-5}%	

A. Ern, M. Vohralík, SIAM Journal on Numerical Analysis (2015)
V. Dolejší, A. Ern, M. Vohralík, SIAM Journal on Scientific Computing (2016)

How **large** is the error? (steady linear Darcy, known solution)

$h_\ell \approx 1/\lvert\mathcal{T}_\ell\rvert^{\frac{1}{2}}$	p	relative error estimate $\frac{\eta(u_\ell)}{\lVert\nabla u_\ell\rVert}$	relative error $\frac{\lVert\nabla(u-u_\ell)\rVert}{\lVert\nabla u_\ell\rVert}$	effectivity index $\frac{\eta(u_\ell)}{\lVert\nabla(u-u_\ell)\rVert}$
h_0	1	28%	24%	1.17

How **large** is the error? (steady linear Darcy, known solution)

$h_\ell \approx 1/\lvert\mathcal{T}_\ell\rvert^{\frac{1}{2}}$	p	relative error estimate $\frac{\eta(u_\ell)}{\lVert\nabla u_\ell\rVert}$	relative error $\frac{\lVert\nabla(u-u_\ell)\rVert}{\lVert\nabla u_\ell\rVert}$	effectivity index $\frac{\eta(u_\ell)}{\lVert\nabla(u-u_\ell)\rVert}$
h_0	1	28%	24%	1.17
$\approx h_0/2$		14%	13%	1.09
$\approx h_0/4$		7.0%	6.6%	[illegible]
$\approx h_0/8$		3.3%	3.1%	[illegible]
$\approx h_0/2$	2	9.5×10^{-1}%	9.2×10^{-1}%	[illegible]
$\approx h_0/4$	3	5.9×10^{-3}%	5.9×10^{-3}%	[illegible]
$\approx h_0/8$	4	5.9×10^{-6}%	5.8×10^{-6}%	[illegible]

A. Ern, M. Vohralík, SIAM Journal on Numerical Analysis (2015)
V. Dolejší, A. Ern, M. Vohralík, SIAM Journal on Scientific Computing (2016)

How **large** is the error? (steady linear Darcy, known solution)

$h_\ell \approx 1/\lvert\mathcal{T}_\ell\rvert^{\frac{1}{2}}$	p	relative error estimate $\frac{\eta(u_\ell)}{\lVert\nabla u_\ell\rVert}$	relative error $\frac{\lVert\nabla(u-u_\ell)\rVert}{\lVert\nabla u_\ell\rVert}$	effectivity index $\frac{\eta(u_\ell)}{\lVert\nabla(u-u_\ell)\rVert}$
h_0	1	28%	24%	**1.17**
$\approx h_0/2$		14%	13%	**1.09**
$\approx h_0/4$		7.0%	6.6%	**1.06**
$\approx h_0/8$		3.3%	3.1%	**1.04**
$\approx h_0/2$	2	9.5×10^{-1}%	9.2×10^{-1}%	**1.04**
$\approx h_0/4$	3	5.9×10^{-3}%	5.9×10^{-3}%	**1.01**
$\approx h_0/8$	4	5.9×10^{-6}%	5.8×10^{-6}%	**1.01**

A. Ern, M. Vohralík, SIAM Journal on Numerical Analysis (2015)
V. Dolejší, A. Ern, M. Vohralík, SIAM Journal on Scientific Computing (2016)

How **large** is the error? (steady linear Darcy, known solution)

$h_\ell \approx 1/\vert\mathcal{T}_\ell\vert^{\frac{1}{2}}$	p	relative error estimate $\frac{\eta(u_\ell)}{\Vert\nabla u_\ell\Vert}$	relative error $\frac{\Vert\nabla(u-u_\ell)\Vert}{\Vert\nabla u_\ell\Vert}$	effectivity index $\frac{\eta(u_\ell)}{\Vert\nabla(u-u_\ell)\Vert}$
h_0	1	28%	24%	**1.17**
$\approx h_0/2$		14%	13%	**1.09**
$\approx h_0/4$		7.0%	6.6%	**1.06**
$\approx h_0/8$		3.3%	3.1%	**1.04**
$\approx h_0/2$	2	9.5×10^{-1}%	9.2×10^{-1}%	**1.04**
$\approx h_0/4$	3	5.9×10^{-3}%	5.9×10^{-3}%	**1.01**
$\approx h_0/8$	4	5.9×10^{-6}%	5.8×10^{-6}%	**1.01**

A. Ern, M. Vohralík, SIAM Journal on Numerical Analysis (2015)
V. Dolejší, A. Ern, M. Vohralík, SIAM Journal on Scientific Computing (2016)

How **large** is the error? (steady linear Darcy, known solution)

$h_\ell \approx 1/\lvert\mathcal{T}_\ell\rvert^{\frac{1}{2}}$	p	relative error estimate $\frac{\eta(u_\ell)}{\lVert\nabla u_\ell\rVert}$	relative error $\frac{\lVert\nabla(u-u_\ell)\rVert}{\lVert\nabla u_\ell\rVert}$	effectivity index $\frac{\eta(u_\ell)}{\lVert\nabla(u-u_\ell)\rVert}$
h_0	1	28%	24%	**1.17**
$\approx h_0/2$		14%	13%	**1.09**
$\approx h_0/4$		7.0%	6.6%	**1.06**
$\approx h_0/8$		3.3%	3.1%	**1.04**
$\approx h_0/2$	2	9.5×10^{-1}%	9.2×10^{-1}%	**1.04**
$\approx h_0/4$	3	5.9×10^{-3}%	5.9×10^{-3}%	**1.01**
$\approx h_0/8$	4	5.9×10^{-6}%	5.8×10^{-6}%	**1.01**

A. Ern, M. Vohralík, SIAM Journal on Numerical Analysis (2015)
V. Dolejší, A. Ern, M. Vohralík, SIAM Journal on Scientific Computing (2016)

How **large** is the error? (steady linear Darcy, known solution)

$h_\ell \approx 1/\lvert\mathcal{T}_\ell\rvert^{\frac{1}{2}}$	p	relative error estimate $\frac{\eta(u_\ell)}{\lVert\nabla u_\ell\rVert}$	relative error $\frac{\lVert\nabla(u-u_\ell)\rVert}{\lVert\nabla u_\ell\rVert}$	effectivity index $\frac{\eta(u_\ell)}{\lVert\nabla(u-u_\ell)\rVert}$
h_0	1	28%	24%	**1.17**
$\approx h_0/2$		14%	13%	**1.09**
$\approx h_0/4$		7.0%	6.6%	**1.06**
$\approx h_0/8$		3.3%	3.1%	**1.04**
$\approx h_0/2$	2	9.5×10^{-1}%	9.2×10^{-1}%	**1.04**
$\approx h_0/4$	3	5.9×10^{-3}%	5.9×10^{-3}%	**1.01**
$\approx h_0/8$	4	5.9×10^{-6}%	5.8×10^{-6}%	**1.01**

A. Ern, M. Vohralík, SIAM Journal on Numerical Analysis (2015)
V. Dolejší, A. Ern, M. Vohralík, SIAM Journal on Scientific Computing (2016)

How **large** is the error? (steady linear Darcy, known solution)

$h_\ell \approx 1/\lvert\mathcal{T}_\ell\rvert^{\frac{1}{2}}$	p	relative error estimate $\frac{\eta(u_\ell)}{\lVert\nabla u_\ell\rVert}$	relative error $\frac{\lVert\nabla(u-u_\ell)\rVert}{\lVert\nabla u_\ell\rVert}$	effectivity index $\frac{\eta(u_\ell)}{\lVert\nabla(u-u_\ell)\rVert}$
h_0	1	28%	24%	**1.17**
$\approx h_0/2$		14%	13%	**1.09**
$\approx h_0/4$		7.0%	6.6%	**1.06**
$\approx h_0/8$		3.3%	3.1%	**1.04**
$\approx h_0/2$	2	9.5×10^{-1}%	9.2×10^{-1}%	**1.04**
$\approx h_0/4$	3	5.9×10^{-3}%	5.9×10^{-3}%	**1.01**
$\approx h_0/8$	4	5.9×10^{-6}%	5.8×10^{-6}%	**1.01**

A. Ern, M. Vohralík, SIAM Journal on Numerical Analysis (2015)
V. Dolejší, A. Ern, M. Vohralík, SIAM Journal on Scientific Computing (2016)

How **large** is the error? (steady linear Darcy, known solution)

$h_\ell \approx 1/\lvert\mathcal{T}_\ell\rvert^{\frac{1}{2}}$	p	relative error estimate $\frac{\eta(u_\ell)}{\|\nabla u_\ell\|}$	relative error $\frac{\|\nabla(u-u_\ell)\|}{\|\nabla u_\ell\|}$	effectivity index $\frac{\eta(u_\ell)}{\|\nabla(u-u_\ell)\|}$
h_0	1	28%	24%	**1.17**
$\approx h_0/2$		14%	13%	**1.09**
$\approx h_0/4$		7.0%	6.6%	**1.06**
$\approx h_0/8$		3.3%	3.1%	**1.04**
$\approx h_0/2$	2	$9.5 \times 10^{-1}\%$	$9.2 \times 10^{-1}\%$	**1.04**
$\approx h_0/4$	3	$5.9 \times 10^{-3}\%$	$5.9 \times 10^{-3}\%$	**1.01**
$\approx h_0/8$	4	$5.9 \times 10^{-6}\%$	$5.8 \times 10^{-6}\%$	**1.01**

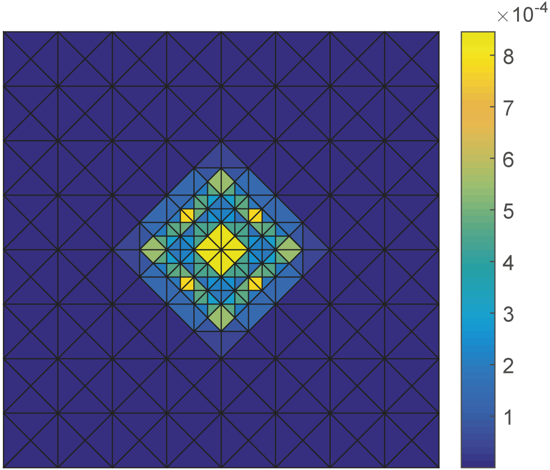
A. Ern, M. Vohralík, SIAM Journal on Numerical Analysis (2015)
V. Dolejší, A. Ern, M. Vohralík, SIAM Journal on Scientific Computing (2016)

How **large** is the error? (steady linear Darcy, known solution)

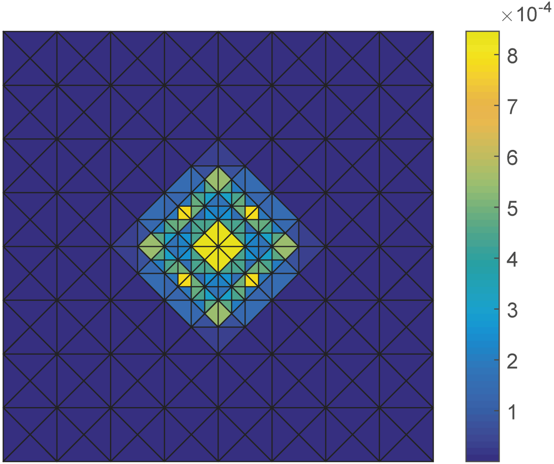
$h_\ell \approx 1/\lvert\mathcal{T}_\ell\rvert^{\frac{1}{2}}$	p	relative error estimate $\frac{\eta(u_\ell)}{\lVert\nabla u_\ell\rVert}$	relative error $\frac{\lVert\nabla(u-u_\ell)\rVert}{\lVert\nabla u_\ell\rVert}$	effectivity index $\frac{\eta(u_\ell)}{\lVert\nabla(u-u_\ell)\rVert}$
h_0	1	28%	24%	**1.17**
$\approx h_0/2$		14%	13%	**1.09**
$\approx h_0/4$		7.0%	6.6%	**1.06**
$\approx h_0/8$		3.3%	3.1%	**1.04**
$\approx h_0/2$	2	9.5×10^{-1}%	9.2×10^{-1}%	**1.04**
$\approx h_0/4$	3	5.9×10^{-3}%	5.9×10^{-3}%	**1.01**
$\approx h_0/8$	4	5.9×10^{-6}%	5.8×10^{-6}%	**1.01**

A. Ern, M. Vohralík, SIAM Journal on Numerical Analysis (2015)
V. Dolejší, A. Ern, M. Vohralík, SIAM Journal on Scientific Computing (2016)

Where (in space) is the error **localized**? (steady linear Darcy)



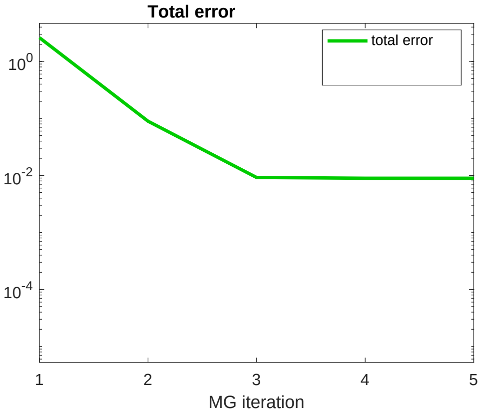
Estimated local error $\eta_K(u_\ell) = \|\nabla u_\ell + \boldsymbol{\sigma}_\ell\|_K$



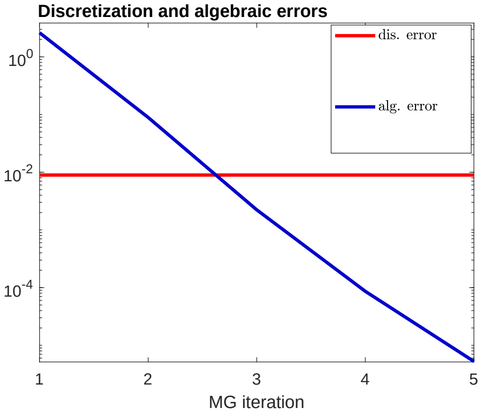
Exact local error $\|\nabla(u - u_\ell)\|_K$

P. Daniel, A. Ern, I. Smears, M. Vohralík, Computers & Mathematics with Applications (2018)

How **large** is the total error and its components? ($\mathbb{A}_\ell \mathsf{U}_\ell^i \neq \mathsf{F}_\ell$)



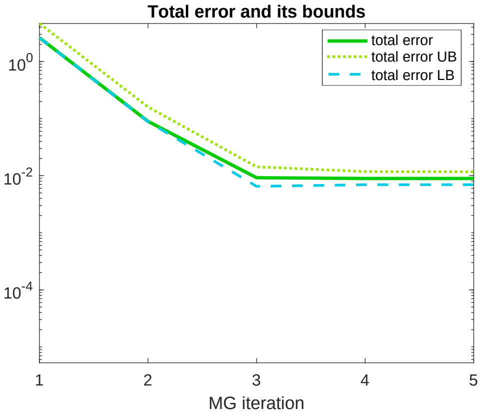
Total error



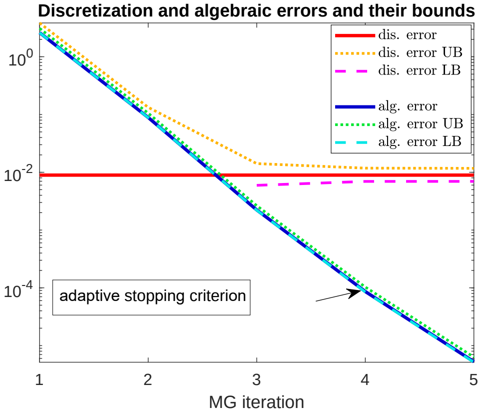
Error components

J. Papež, U. Rüde, M. Vohralík, B. Wohlmuth, Computer Methods in Applied Mechanics and Engineering (2020)

How **large** is the total error and its components? ($\mathbb{A}_\ell \mathsf{U}_\ell^i \neq \mathsf{F}_\ell$)



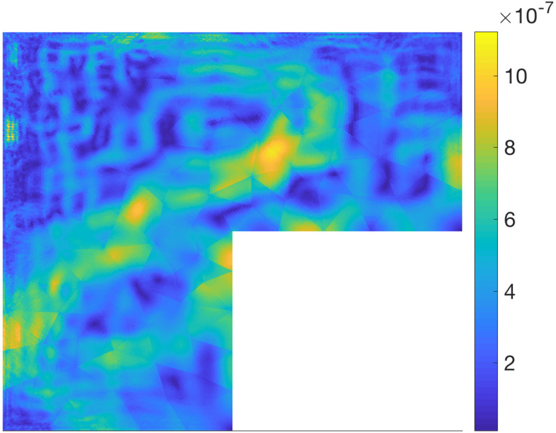
Total error



Error components and stopping criteria

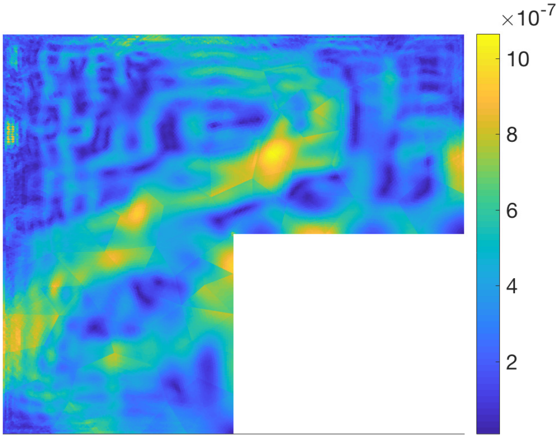
J. Papež, U. Rüde, M. Vohralík, B. Wohlmuth, Computer Methods in Applied Mechanics and Engineering (2020)

Where (in space) is the algebraic error localized? ($\mathbb{A}_\ell U_\ell^i \neq F_\ell$)



Estimated local algebraic errors

$\eta_{\mathrm{alg},K}(u_\ell^i) = \|\boldsymbol{\sigma}_{\mathrm{alg},\ell}^i\|_K$

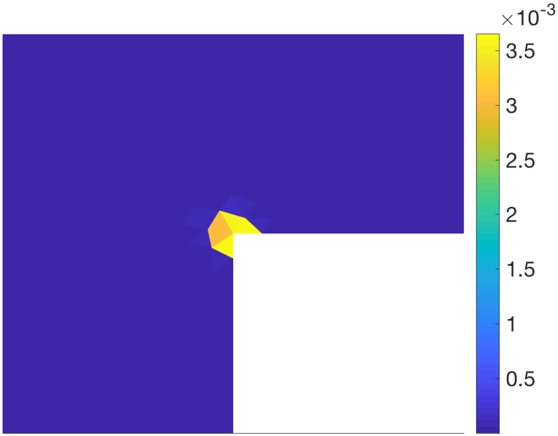


Exact local algebraic errors

$\|\nabla(u_\ell - u_\ell^i)\|_K$

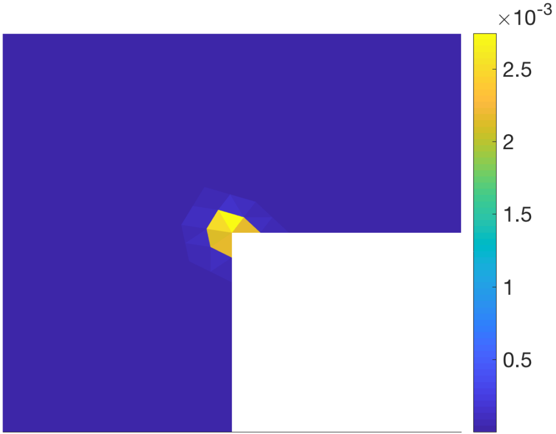
J. Papež, U. Rüde, M. Vohralík, B. Wohlmuth, Computer Methods in Applied Mechanics and Engineering (2020)

Where (in space) is the total error **localized**? ($\mathbb{A}_\ell \mathsf{U}_\ell^i \neq \mathsf{F}_\ell$)



Estimated local total errors

$\eta_K(u_\ell^i) = \|\nabla u_\ell + \boldsymbol{\sigma}_\ell^i\|_K$



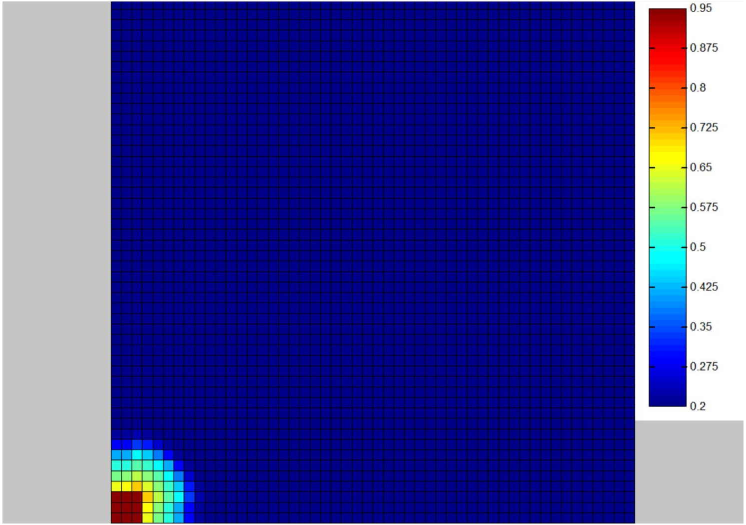
Exact local total errors

$\|\nabla(u - u_\ell^i)\|_K$

J. Papež, U. Rüde, M. Vohralík, B. Wohlmuth, Computer Methods in Applied Mechanics and Engineering (2020)

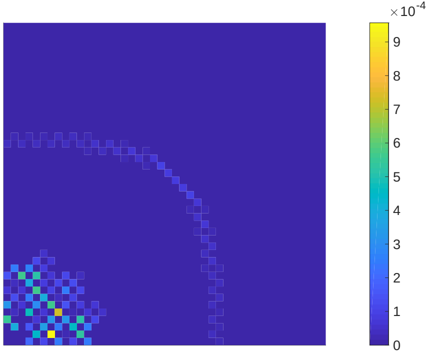
Outline

Two-phase flow, water saturation

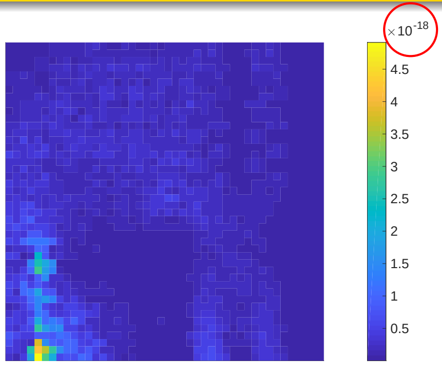


M. Vohralík, M. Wheeler, Computational Geosciences (2013)

Recovering mass balance: two-phase flow (inexact solver, water)



original mass balance misfit (m^2s^{-1})



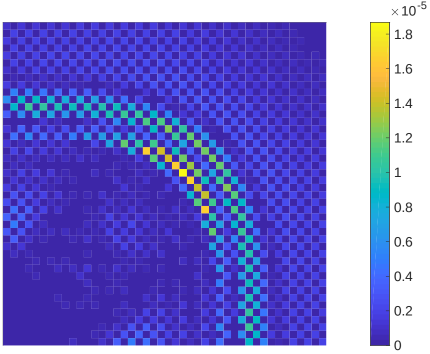
corrected mass balance misfit (m^2s^{-1})

Setting

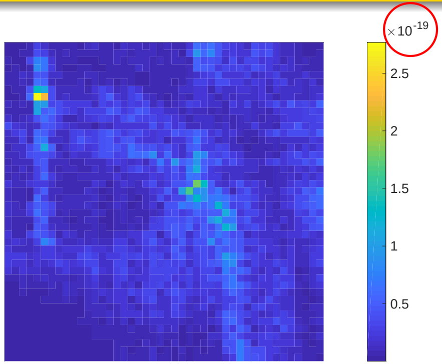
- fully implicit discretization of a two-phase oil–water flow
- cell-centered finite volumes on a square mesh
- time step 260, 1st Newton linearization, GMRes iteration 195

J. Papež, U. Rüde, M. Vohralík, B. Wohlmuth, Computer Methods in Applied Mechanics and Engineering (2020)

Recovering mass balance: two-phase flow (inexact solver, oil)



original mass balance misfit $(\mathrm{m}^2\mathrm{s}^{-1})$



corrected mass balance misfit $(\mathrm{m}^2\mathrm{s}^{-1})$

Setting

- fully implicit discretization of a two-phase oil–water flow
- cell-centered finite volumes on a square mesh
- time step 260, 1st Newton linearization, GMRes iteration 195

J. Papež, U. Rüde, M. Vohralík, B. Wohlmuth, Computer Methods in Applied Mechanics and Engineering (2020)

Outline

Outline

A model steady nonlinear problem

Nonlinear elliptic problem

Find $u : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\nabla\cdot(a(|\nabla u|)\nabla u) &= f \quad \text{in} \quad \Omega, \\ u &= 0 \quad \text{on} \quad \partial\Omega. \end{aligned}$$

- $\Omega \subset \mathbb{R}^d$, $1 \leq d \leq 3$, open bounded polytope with Lipschitz boundary $\partial\Omega$
- f piecewise polynomial for simplicity

A model steady nonlinear problem

Nonlinear elliptic problem

Find $u : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\nabla\cdot(a(|\nabla u|)\nabla u) &= f \quad \text{in} \quad \Omega, \\ u &= 0 \quad \text{on} \quad \partial\Omega. \end{aligned}$$

- $\Omega \subset \mathbb{R}^d$, $1 \leq d \leq 3$, open bounded polytope with Lipschitz boundary $\partial\Omega$
- f piecewise polynomial for simplicity

Assumption (Nonlinear function a)

Function $a : [0, \infty) \to (0, \infty)$, *for all* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$,

$$\begin{aligned} |a(|\boldsymbol{x}|)\boldsymbol{x} - a(|\boldsymbol{y}|)\boldsymbol{y}| &\leq a_{\rm c}|\boldsymbol{x} - \boldsymbol{y}| && \textit{(Lipschitz continuity)}, \\ (a(|\boldsymbol{x}|)\boldsymbol{x} - a(|\boldsymbol{y}|)\boldsymbol{y}) \cdot (\boldsymbol{x} - \boldsymbol{y}) &\geq a_{\rm m}|\boldsymbol{x} - \boldsymbol{y}|^2 && \textit{(strong monotonicity)}. \end{aligned}$$

A model steady nonlinear problem

Nonlinear elliptic problem

Find $u : \Omega \to \mathbb{R}$ such that

$$
\begin{aligned}
-\nabla\cdot(a(|\nabla u|)\nabla u) &= f \quad \text{in} \quad \Omega,\\
u &= 0 \quad \text{on} \quad \partial\Omega.
\end{aligned}
$$

- $\Omega \subset \mathbb{R}^d$, $1 \le d \le 3$, open bounded polytope with Lipschitz boundary $\partial\Omega$
- f piecewise polynomial for simplicity

Assumption (Nonlinear function a)

Function $a : [0,\infty) \to (0,\infty)$, *for all* $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^d$,

$$
\begin{aligned}
|a(|\boldsymbol{x}|)\boldsymbol{x} - a(|\boldsymbol{y}|)\boldsymbol{y}| &\le a_{\mathrm{c}}|\boldsymbol{x}-\boldsymbol{y}| \qquad &\text{(Lipschitz continuity)},\\
(a(|\boldsymbol{x}|)\boldsymbol{x} - a(|\boldsymbol{y}|)\boldsymbol{y})\cdot(\boldsymbol{x}-\boldsymbol{y}) &\ge a_{\mathrm{m}}|\boldsymbol{x}-\boldsymbol{y}|^2 \qquad &\text{(strong monotonicity)}.
\end{aligned}
$$

- $a_{\mathrm{m}} \le a(r) \le a_{\mathrm{c}}$, $a_{\mathrm{m}} \le (a(r)r)' \le a_{\mathrm{c}}$

Example of the nonlinear function a

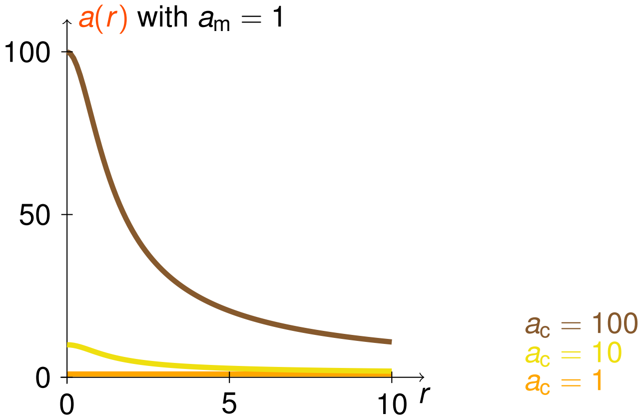
Example (Mean curvature nonlinearity)

$$a(r) := a_{\mathrm{m}} + \frac{a_{\mathrm{c}} - a_{\mathrm{m}}}{\sqrt{1 + r^2}}.$$

Example of the nonlinear function a

Example (Mean curvature nonlinearity)

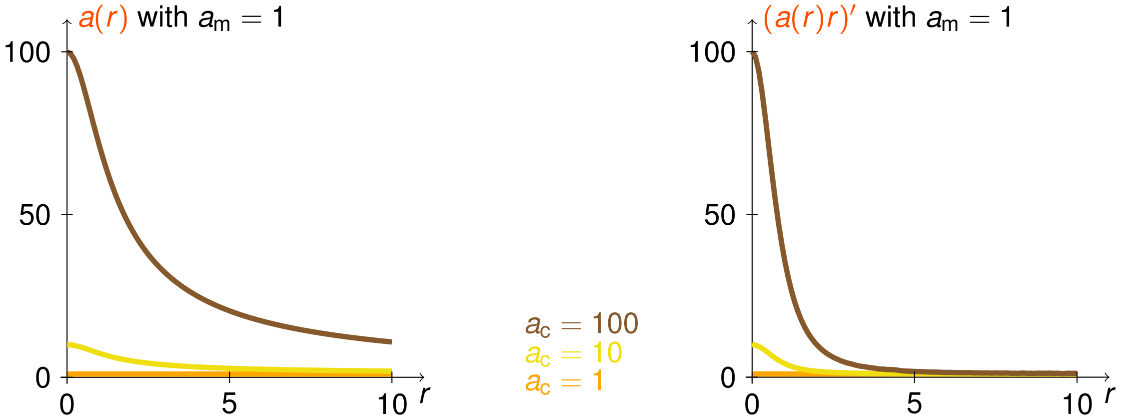
$$a(r) := a_m + \frac{a_c - a_m}{\sqrt{1 + r^2}}.$$



Example of the nonlinear function a

Example (Mean curvature nonlinearity)

$$a(r) := a_{\mathrm{m}} + \frac{a_{\mathrm{c}} - a_{\mathrm{m}}}{\sqrt{1 + r^2}}.$$



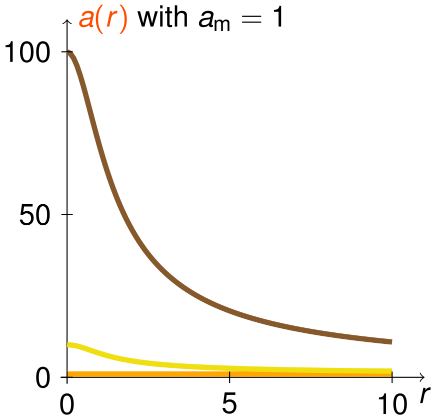
Example of the nonlinear function a

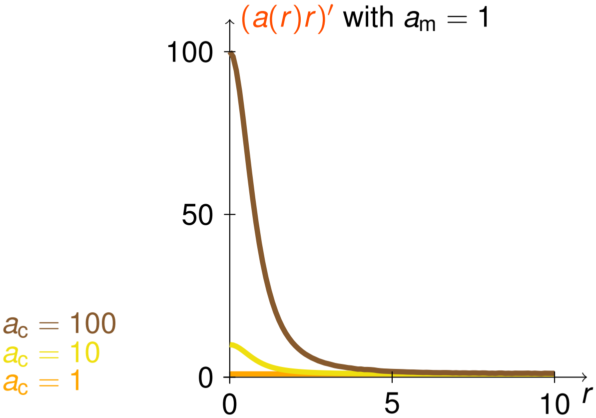
Example (Mean curvature nonlinearity)

$$a(r) := a_{\mathrm{m}} + \frac{a_{\mathrm{c}} - a_{\mathrm{m}}}{\sqrt{1+r^2}}.$$

Strength of the nonlinearity

$$\frac{a_{\mathrm{c}}}{a_{\mathrm{m}}} = \frac{\text{Lipschitz continuity}}{\text{strong monotonicity}}$$





Weak solution

Definition (Weak solution)

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

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

Energy

Definition (Energy functional)

$\mathcal{J} : H_0^1(\Omega) \to \mathbb{R}$

$$\mathcal{J}(v) := \int_\Omega \phi(|\nabla v|) - (f, v), \quad v \in H_0^1(\Omega),$$

with function $\phi : [0, \infty) \to [0, \infty)$ such that, for all $r \in [0, \infty)$,

$$\phi(r) := \int_0^r a(s) s \, \mathrm{d}s.$$

Equivalently

$$u = \arg \min_{v \in H_0^1(\Omega)} \mathcal{J}(v)$$

Energy

Definition (Energy functional)

$\mathcal{J} : H_0^1(\Omega) \to \mathbb{R}$

$$\mathcal{J}(v) := \int_\Omega \phi(|\nabla v|) - (f, v), \quad v \in H_0^1(\Omega),$$

with function $\phi : [0, \infty) \to [0, \infty)$ such that, for all $r \in [0, \infty)$,

$$\phi(r) := \int_0^r a(s) s \, \mathrm{d}s.$$

Equivalently

$$u = \arg \min_{v \in H_0^1(\Omega)} \mathcal{J}(v)$$

Finite element approximation

Definition (Finite element approximation)

$u_\ell \in V_\ell^p$ such that

$$(a(|\nabla u_\ell|)\nabla u_\ell, \nabla v_\ell) = (f, v_\ell) \qquad \forall v_\ell \in V_\ell^p.$$

- $\mathcal{T}_\ell$ simplicial mesh of Ω
- $p \geq 1$ polynomial degree
- $V_\ell^p := \mathcal{P}_p(\mathcal{T}_\ell) \cap H_0^1(\Omega)$
- conforming finite elements

Equivalently

$$u_\ell = \arg\min_{v_\ell \in V_\ell^p} \mathcal{J}(v_\ell)$$

Finite element approximation

Definition (Finite element approximation)

$u_\ell \in V_\ell^p$ such that

$$(a(|\nabla u_\ell|)\nabla u_\ell, \nabla v_\ell) = (f, v_\ell) \qquad \forall v_\ell \in V_\ell^p.$$

- $\mathcal{T}_\ell$ simplicial mesh of Ω
- $p \geq 1$ polynomial degree
- $V_\ell^p := \mathcal{P}_p(\mathcal{T}_\ell) \cap H_0^1(\Omega)$
- conforming finite elements

Equivalently

$$u_\ell = \arg \min_{v_\ell \in V_\ell^p} \mathcal{J}(v_\ell)$$

Finite element approximation

Definition (Finite element approximation)

$u_\ell \in V_\ell^p$ such that

$$(a(|\nabla u_\ell|)\nabla u_\ell, \nabla v_\ell) = (f, v_\ell) \qquad \forall v_\ell \in V_\ell^p.$$

- $\mathcal{T}_\ell$ simplicial mesh of Ω
- $p \geq 1$ polynomial degree
- $V_\ell^p := \mathcal{P}_p(\mathcal{T}_\ell) \cap H_0^1(\Omega)$
- conforming finite elements

Equivalently

$$u_\ell = \arg \min_{v_\ell \in V_\ell^p} \mathcal{J}(v_\ell)$$

Energy difference

Energy difference

$$\boxed{\mathcal{J}(u_\ell) - \mathcal{J}(u)}$$

- $\mathcal{J}(u_\ell) - \mathcal{J}(u) \geq 0$, $\mathcal{J}(u_\ell) - \mathcal{J}(u) = 0$ if and only if $u_\ell = u$
- physically-based error measure

Previous results

Energy difference (not robust wrt $\frac{a_c}{a_m}$)

$$\mathcal{J}(u_\ell) - \mathcal{J}(u) \leq \eta(u_\ell)^2 \leq C_{\text{eff}}^2 \frac{a_c^2}{a_m^2} \left(\mathcal{J}(u_\ell) - \mathcal{J}(u) \right)$$

- Zeidler (1992), Han (1994), Repin (1997), Ladevèze & Moës (1997), Diening & Kreuzer (2008), Bartels & Milicevic (2020), ...

Sobolev norm

$$a_m \|\nabla(u_\ell - u)\| \leq \eta(u_\ell) \leq C_{\text{eff}} a_c \|\nabla(u_\ell - u)\|$$

Dual norm of the residual

$$|||\mathcal{R}(u_\ell)|||_{-1} \leq \eta(u_\ell) \leq C_{\text{eff}} |||\mathcal{R}(u_\ell)|||_{-1}$$

Previous results

Energy difference (**not robust** wrt $\frac{a_c}{a_m}$)

$$\mathcal{J}(u_\ell) - \mathcal{J}(u) \le \eta(u_\ell)^2 \le C_{\rm eff}^2 \frac{a_c^2}{a_m^2} (\mathcal{J}(u_\ell) - \mathcal{J}(u))$$

- Zeidler (1992), Han (1994), Repin (1997), Ladevèze & Moës (1997), Diening & Kreuzer (2008), Bartels & Milicevic (2020), ...

Sobolev norm

$$a_m \|\nabla(u_\ell - u)\| \le \eta(u_\ell) \le C_{\rm eff}\, a_c \|\nabla(u_\ell - u)\|$$

Dual norm of the residual

$$|||\mathcal{R}(u_\ell)|||_{-1} \le \eta(u_\ell) \le C_{\rm eff} |||\mathcal{R}(u_\ell)|||_{-1}$$

Previous results

Energy difference (**not robust** wrt $\frac{a_c}{a_m}$)

$$\mathcal{J}(u_\ell) - \mathcal{J}(u) \leq \eta(u_\ell)^2 \leq C_{\rm eff}^2 \frac{a_c^2}{a_m^2} (\mathcal{J}(u_\ell) - \mathcal{J}(u))$$

- Zeidler (1992), Han (1994), Repin (1997), Ladevèze & Moës (1997), Diening & Kreuzer (2008), Bartels & Milicevic (2020), ...

Sobolev norm (**not robust** wrt $\frac{a_c}{a_m}$)

$$a_m \|\nabla(u_\ell - u)\| \leq \eta(u_\ell) \leq C_{\rm eff} a_c \|\nabla(u_\ell - u)\|$$

- Pousin & Rappaz (1994), Verfürth (1994), Kim (2007), Houston, Süli & Wihler (2008), Garau, Morin, & Zuppa (2011), Gantner, Haberl, Praetorius, & Stiftner (2018), Heid & Wihler (2020), ...

Dual norm of the residual

$$|||\mathcal{R}(u_\ell)|||_{-1} \leq \eta(u_\ell) \leq C_{\rm eff} |||\mathcal{R}(u_\ell)|||_{-1}$$

Previous results

Energy difference (**not robust** wrt $\frac{a_c}{a_m}$)

$$\mathcal{J}(u_\ell) - \mathcal{J}(u) \leq \eta(u_\ell)^2 \leq C_{\text{eff}}^2 \frac{a_c^2}{a_m^2} (\mathcal{J}(u_\ell) - \mathcal{J}(u))$$

- Zeidler (1992), Han (1994), Repin (1997), Ladevèze & Moës (1997), Diening & Kreuzer (2008), Bartels & Milicevic (2020), ...

Sobolev norm (**not robust** wrt $\frac{a_c}{a_m}$)

$$a_m \|\nabla(u_\ell - u)\| \leq \eta(u_\ell) \leq C_{\text{eff}} a_c \|\nabla(u_\ell - u)\|$$

- Pousin & Rappaz (1994), Verfürth (1994), Kim (2007), Houston, Süli, & Wihler (2008), Garau, Morin, & Zuppa (2011), Gantner, Haberl, Praetorius, & Stiftner (2018), Heid & Wihler (2020), ...

Dual norm of the residual

$$|||\mathcal{R}(u_\ell)|||_{-1} \leq \eta(u_\ell) \leq C_{\text{eff}} |||\mathcal{R}(u_\ell)|||_{-1}$$

Previous results

Energy difference (**not robust** wrt $\frac{a_c}{a_m}$)

$$\mathcal{J}(u_\ell) - \mathcal{J}(u) \leq \eta(u_\ell)^2 \leq C_{\text{eff}}^2 \frac{a_c^2}{a_m^2} (\mathcal{J}(u_\ell) - \mathcal{J}(u))$$

- Zeidler (1992), Han (1994), Repin (1997), Ladevèze & Moës (1997), Diening & Kreuzer (2008), Bartels & Milicevic (2020), ...

Sobolev norm (**not robust** wrt $\frac{a_c}{a_m}$)

$$a_m \|\nabla(u_\ell - u)\| \leq \eta(u_\ell) \leq C_{\text{eff}} a_c \|\nabla(u_\ell - u)\|$$

- Pousin & Rappaz (1994), Verfürth (1994), Kim (2007), Houston, Süli, & Wihler (2008), Garau, Morin, & Zuppa (2011), Gantner, Haberl, Praetorius, & Stiftner (2018), Heid & Wihler (2020), ...

Dual norm of the residual

$$|||\mathcal{R}(u_\ell)|||_{-1} \leq \eta(u_\ell) \leq C_{\text{eff}} |||\mathcal{R}(u_\ell)|||_{-1}$$

Previous results

Energy difference (**not robust** wrt $\frac{a_c}{a_m}$)

$$\mathcal{J}(u_\ell) - \mathcal{J}(u) \leq \eta(u_\ell)^2 \leq C_{\text{eff}}^2 \frac{a_c^2}{a_m^2} (\mathcal{J}(u_\ell) - \mathcal{J}(u))$$

- Zeidler (1992), Han (1994), Repin (1997), Ladevèze & Moës (1997), Diening & Kreuzer (2008), Bartels & Milicevic (2020), ...

Sobolev norm (**not robust** wrt $\frac{a_c}{a_m}$)

$$a_m \|\nabla(u_\ell - u)\| \leq \eta(u_\ell) \leq C_{\text{eff}} a_c \|\nabla(u_\ell - u)\|$$

- Pousin & Rappaz (1994), Verfürth (1994), Kim (2007), Houston, Süli, & Wihler (2008), Garau, Morin, & Zuppa (2011), Gantner, Haberl, Praetorius, & Stiftner (2018), Heid & Wihler (2020), ...

Dual norm of the residual (robust wrt $\frac{a_c}{a_m}$), "bypasses" the nonlinearity

$$|||\mathcal{R}(u_\ell)|||_{-1} \leq \eta(u_\ell) \leq C_{\text{eff}} |||\mathcal{R}(u_\ell)|||_{-1}$$

Previous results

Energy difference (**not robust** wrt $\frac{a_c}{a_m}$)

$$\mathcal{J}(u_\ell) - \mathcal{J}(u) \leq \eta(u_\ell)^2 \leq C_{\text{eff}}^2 \frac{a_c^2}{a_m^2} (\mathcal{J}(u_\ell) - \mathcal{J}(u))$$

- Zeidler (1992), Han (1994), Repin (1997), Ladevèze & Moës (1997), Diening & Kreuzer (2008), Bartels & Milicevic (2020), ...

Sobolev norm (**not robust** wrt $\frac{a_c}{a_m}$)

$$a_m \|\nabla(u_\ell - u)\| \leq \eta(u_\ell) \leq C_{\text{eff}} a_c \|\nabla(u_\ell - u)\|$$

- Pousin & Rappaz (1994), Verfürth (1994), Kim (2007), Houston, Süli, & Wihler (2008), Garau, Morin, & Zuppa (2011), Gantner, Haberl, Praetorius, & Stiftner (2018), Heid & Wihler (2020), ...

Dual norm of the residual (**robust** wrt $\frac{a_c}{a_m}$), "bypasses" the nonlinearity

$$|||\mathcal{R}(u_\ell)|||_{-1} \leq \eta(u_\ell) \leq C_{\text{eff}} |||\mathcal{R}(u_\ell)|||_{-1}$$

- El Alaoui, Ern, & Vohralík (2011), Blechta, Málek, & Vohralík (2020), ...

Previous results

Energy difference (**not robust** wrt $\frac{a_c}{a_m}$)

$$\mathcal{J}(u_\ell) - \mathcal{J}(u) \leq \eta(u_\ell)^2 \leq C_{\text{eff}}^2 \frac{a_c^2}{a_m^2} (\mathcal{J}(u_\ell) - \mathcal{J}(u))$$

- Zeidler (1992), Han (1994), Repin (1997), Ladevèze & Moës (1997), Diening & Kreuzer (2008), Bartels & Milicevic (2020), ...

Sobolev norm (**not robust** wrt $\frac{a_c}{a_m}$)

$$a_m \|\nabla(u_\ell - u)\| \leq \eta(u_\ell) \leq C_{\text{eff}} a_c \|\nabla(u_\ell - u)\|$$

- Pousin & Rappaz (1994), Verfürth (1994), Kim (2007), Houston, Süli, & Wihler (2008), Garau, Morin, & Zuppa (2011), Gantner, Haberl, Praetorius, & Stiftner (2018), Heid & Wihler (2020), ...

Dual norm of the residual (**robust** wrt $\frac{a_c}{a_m}$), "bypasses" the nonlinearity

$$|||\mathcal{R}(u_\ell)|||_{-1} \leq \eta(u_\ell) \leq C_{\text{eff}} |||\mathcal{R}(u_\ell)|||_{-1}$$

- El Alaoui, Ern, & Vohralík (2011), Blechta, Málek, & Vohralík (2020), ...

Previous results

Energy difference (**not robust** wrt $\frac{a_c}{a_m}$)

$$\mathcal{J}(u_\ell) - \mathcal{J}(u) \leq \eta(u_\ell)^2 \leq C_{\text{eff}}^2 \frac{a_c^2}{a_m^2} (\mathcal{J}(u_\ell) - \mathcal{J}(u))$$

- Zeidler (1992), Han (1994), Repin (1997), Ladevèze & Moës (1997), Diening & Kreuzer (2008), Bartels & Milicevic (2020), ...

Sobolev norm (**not robust** wrt $\frac{a_c}{a_m}$)

$$a_m \|\nabla(u_\ell - u)\| \leq \eta(u_\ell) \leq C_{\text{eff}} a_c \|\nabla(u_\ell - u)\|$$

- Pousin & Rappaz (1994), Verfürth (1994), Kim (2007), Houston, Süli, & Wihler (2008), Garau, Morin, & Zuppa (2011), Gantner, Haberl, Praetorius, & Stiftner (2018), Heid & Wihler (2020), ...

Dual norm of the residual (**robust** wrt $\frac{a_c}{a_m}$), "bypasses" the nonlinearity

$$|||\mathcal{R}(u_\ell)|||_{-1} \leq \eta(u_\ell) \leq C_{\text{eff}} |||\mathcal{R}(u_\ell)|||_{-1}$$

- El Alaoui, Ern, & Vohralík (2011), Blechta, Málek, & Vohralík (2020), ...

Previous results

Energy difference (**not robust** wrt $\frac{a_c}{a_m}$)

$$\mathcal{J}(u_\ell) - \mathcal{J}(u) \leq \eta(u_\ell)^2 \leq C_{\text{eff}}^2 \frac{a_c^2}{a_m^2} (\mathcal{J}(u_\ell) - \mathcal{J}(u))$$

- Zeidler (1992), Han (1994), Repin (1997), Ladevèze & Moës (1997), Diening & Kreuzer (2008), Bartels & Milicevic (2020), ...

Sobolev norm (**not robust** wrt $\frac{a_c}{a_m}$)

$$a_m \|\nabla(u_\ell - u)\| \leq \eta(u_\ell) \leq C_{\text{eff}} a_c \|\nabla(u_\ell - u)\|$$

- Pousin & Rappaz (1994), Verfürth (1994), Kim (2007), Houston, Süli, & Wihler (2008), Garau, Morin, & Zuppa (2011), Gantner, Haberl, Praetorius, & Stiftner (2018), Heid & Wihler (2020), ...

Dual norm of the residual (**robust** wrt $\frac{a_c}{a_m}$), "bypasses" the nonlinearity

$$|||\mathcal{R}(u_\ell)|||_{-1} \leq \eta(u_\ell) \leq C_{\text{eff}} |||\mathcal{R}(u_\ell)|||_{-1}$$

- El Alaoui, Ern, & Vohralík (2011), Blechta, Málek, & Vohralík (2020), ...

Iterative linearization

Need to solve a nonlinear system

$$\mathcal{A}_\ell(\mathsf{U}_\ell) = \mathsf{F}_\ell$$

Iterative linearization

Need to solve a nonlinear system

$$\mathcal{A}_\ell(\mathsf{U}_\ell) = \mathsf{F}_\ell$$

Definition (Linearized finite element approximation)

$u_\ell^k \in V_\ell^p$ such that

$$(\boldsymbol{A}_\ell^{k-1} \nabla u_\ell^k, \nabla v_\ell) = (f, v_\ell) + (\boldsymbol{b}_\ell^{k-1}, \nabla v_\ell) \qquad \forall v_\ell \in V_\ell^p.$$

Iterative linearization

Need to solve a nonlinear system

$$\mathcal{A}_\ell(\mathsf{U}_\ell) = \mathsf{F}_\ell$$

Definition (Linearized finite element approximation)

$u_\ell^k \in V_\ell^p$ such that

$$(\boldsymbol{A}_\ell^{k-1}\nabla u_\ell^k, \nabla v_\ell) = (f, v_\ell) + (\boldsymbol{b}_\ell^{k-1}, \nabla v_\ell) \qquad \forall v_\ell \in V_\ell^p.$$

- $u_\ell^0 \in V_\ell^p$ a given initial guess
- iterative linearization index $k \geq 1$
- **linearization**: $\boldsymbol{A}_\ell^{k-1}\colon \Omega \to \mathbb{R}^{d\times d}$ matrix, $\boldsymbol{b}_\ell^{k-1}\colon \Omega \to \mathbb{R}^d$ vector constructed from u_ℓ^{k-1}

Examples

Example (Picard (fixed-point))

$$\boldsymbol{A}_\ell^{k-1} = a(|\nabla u_\ell^{k-1}|)\boldsymbol{I}_d, \quad \boldsymbol{b}_\ell^{k-1} = \boldsymbol{0}.$$

Example (Zarantonello)

$$\boldsymbol{A}_\ell^{k-1} = \gamma \boldsymbol{I}_d, \quad \boldsymbol{b}_\ell^{k-1} = (\gamma - a(|\nabla u_\ell^{k-1}|))\nabla u_\ell^{k-1},$$

with $\gamma \geq \frac{a_c^2}{a_m}$ *a constant parameter.*

Example (Newton)

$$\boldsymbol{A}_\ell^{k-1} = a(|\nabla u_\ell^{k-1}|)\boldsymbol{I}_d + \frac{a'(|\nabla u_\ell^{k-1}|)}{|\nabla u_\ell^{k-1}|}\nabla u_\ell^{k-1} \otimes \nabla u_\ell^{k-1},$$

$$\boldsymbol{b}_\ell^{k-1} = a'(|\nabla u_\ell^{k-1}|)|\nabla u_\ell^{k-1}|\nabla u_\ell^{k-1}.$$

Examples

Example (Picard (fixed-point))

$$\boldsymbol{A}_\ell^{k-1} = a(|\nabla u_\ell^{k-1}|)\boldsymbol{I}_d, \quad \boldsymbol{b}_\ell^{k-1} = \boldsymbol{0}.$$

Example (Zarantonello)

$$\boldsymbol{A}_\ell^{k-1} = \gamma \boldsymbol{I}_d, \quad \boldsymbol{b}_\ell^{k-1} = (\gamma - a(|\nabla u_\ell^{k-1}|))\nabla u_\ell^{k-1},$$

with $\gamma \geq \frac{a_c^2}{a_m}$ *a constant parameter.*

Example (Newton)

$$\boldsymbol{A}_\ell^{k-1} = a(|\nabla u_\ell^{k-1}|)\boldsymbol{I}_d + \frac{a'(|\nabla u_\ell^{k-1}|)}{|\nabla u_\ell^{k-1}|}\nabla u_\ell^{k-1} \otimes \nabla u_\ell^{k-1},$$

$$\boldsymbol{b}_\ell^{k-1} = a'(|\nabla u_\ell^{k-1}|)|\nabla u_\ell^{k-1}|\nabla u_\ell^{k-1}.$$

Examples

Example (Picard (fixed-point))

$$\boldsymbol{A}_\ell^{k-1} = a(|\nabla u_\ell^{k-1}|)\boldsymbol{I}_d, \quad \boldsymbol{b}_\ell^{k-1} = \boldsymbol{0}.$$

Example (Zarantonello)

$$\boldsymbol{A}_\ell^{k-1} = \gamma \boldsymbol{I}_d, \quad \boldsymbol{b}_\ell^{k-1} = (\gamma - a(|\nabla u_\ell^{k-1}|))\nabla u_\ell^{k-1},$$

with $\gamma \geq \frac{a_c^2}{a_m}$ *a constant parameter.*

Example (Newton)

$$\boldsymbol{A}_\ell^{k-1} = a(|\nabla u_\ell^{k-1}|)\boldsymbol{I}_d + \frac{a'(|\nabla u_\ell^{k-1}|)}{|\nabla u_\ell^{k-1}|}\nabla u_\ell^{k-1} \otimes \nabla u_\ell^{k-1},$$

$$\boldsymbol{b}_\ell^{k-1} = a'(|\nabla u_\ell^{k-1}|)|\nabla u_\ell^{k-1}|\nabla u_\ell^{k-1}.$$

Main idea

Observation

None of the known approaches employs **in the analysis**, to define norms, the **iterative linearization**, i.e., **how** do we solve the nonlinear system $\mathcal{A}_\ell(\mathsf{U}_\ell) = \mathsf{F}_\ell$.

Main idea

Observation

None of the known approaches employs **in the analysis**, to define norms, the **iterative linearization**, i.e., **how** do we solve the nonlinear system $\mathcal{A}_\ell(\mathsf{U}_\ell) = \mathsf{F}_\ell$.

Definition (Linearized energy functional)

$\mathcal{J}_\ell^{k-1} : H_0^1(\Omega) \to \mathbb{R}$

$$\mathcal{J}_\ell^{k-1}(v) := \frac{1}{2}\left\|(\boldsymbol{A}_\ell^{k-1})^{\frac{1}{2}}\nabla v\right\|^2 - (f, v) - (\boldsymbol{b}_\ell^{k-1}, \nabla v), \quad v \in H_0^1(\Omega).$$

Main idea

Observation

None of the known approaches employs **in the analysis**, to define norms, the **iterative linearization**, i.e., **how** do we solve the nonlinear system $\mathcal{A}_\ell(\mathsf{U}_\ell) = \mathsf{F}_\ell$.

Definition (Linearized energy functional)

$\mathcal{J}_\ell^{k-1} : H_0^1(\Omega) \to \mathbb{R}$

$$\mathcal{J}_\ell^{k-1}(v) := \frac{1}{2} \left\| (\boldsymbol{A}_\ell^{k-1})^{\frac{1}{2}} \nabla v \right\|^2 - (f, v) - (\boldsymbol{b}_\ell^{k-1}, \nabla v), \quad v \in H_0^1(\Omega).$$

Equivalently

$$u_\ell^k := \arg \min_{v_\ell \in V_\ell^p} \mathcal{J}_\ell^{k-1}(v_\ell)$$

Outline

A posteriori error estimates for an augmented energy difference

Theorem (A posteriori estimate of augmented energy)

For all linearization steps $k \geq 1$, $$\mathcal{E}_\ell^k \leq \eta_\ell^k.$$

A posteriori error estimates for an augmented energy difference

Theorem (A posteriori estimate of augmented energy)

For all linearization steps $k \geq 1$, $\mathcal{E}_\ell^k \leq \eta_\ell^k$.

Moreover, for k *satisfying a stopping criterion, there holds*

$$\eta_\ell^k \leq C_{\text{eff}}(d, \kappa_{\mathcal{T}}) C_\ell^k \mathcal{E}_\ell^k + \text{quadrature error terms},$$

A posteriori error estimates for an augmented energy difference

Theorem (A posteriori estimate of augmented energy)

For all linearization steps $k \geq 1$, $\mathcal{E}_\ell^k \leq \eta_\ell^k$.

Moreover, for k satisfying a stopping criterion, there holds

$$\eta_\ell^k \leq C_{\text{eff}}(d, \kappa_{\mathcal{T}}) C_\ell^k \mathcal{E}_\ell^k + \textit{quadrature error terms},$$

where

$$C_\ell^k \begin{cases} = 1 & \textit{Zarantonello} \\ & \end{cases}$$

A posteriori error estimates for an augmented energy difference

Theorem (A posteriori estimate of augmented energy)

For all linearization steps $k \geq 1$, $\mathcal{E}_\ell^k \leq \eta_\ell^k$.

Moreover, for k satisfying a stopping criterion, there holds

$$\eta_\ell^k \leq C_{\text{eff}}(d, \kappa_{\mathcal{T}}) C_\ell^k \mathcal{E}_\ell^k + \textit{quadrature error terms},$$

where

$$C_\ell^k \begin{cases} = 1 & \textit{Zarantonello} \\ & \end{cases}$$

✓ $C_\ell^k = 1$ for Zarantonello $\Longrightarrow$ **robustness** wrt the **strength of nonlinearities**

A posteriori error estimates for an augmented energy difference

Theorem (A posteriori estimate of augmented energy)

For all linearization steps $k \geq 1$, $\quad \mathcal{E}_\ell^k \leq \eta_\ell^k.$

Moreover, for k satisfying a stopping criterion, there holds

$$\eta_\ell^k \leq C_{\text{eff}}(d, \kappa_{\mathcal{T}}) C_\ell^k \mathcal{E}_\ell^k + \textit{quadrature error terms},$$

where

$$C_\ell^k := \max_{\mathbf{a} \in \mathcal{V}_\ell} \left(\frac{\textit{max. eig. } \mathbf{A}_\ell^{k-1}|_{\omega_\ell^{\mathbf{a}}}}{\textit{min. eig. } \mathbf{A}_\ell^{k-1}|_{\omega_\ell^{\mathbf{a}}}} \right) \begin{cases} = 1 & \textit{Zarantonello} \end{cases}$$

✓ $C_\ell^k = 1$ for Zarantonello $\Longrightarrow$ **robustness** wrt the **strength of nonlinearities**

A posteriori error estimates for an augmented energy difference

Theorem (A posteriori estimate of augmented energy)

For all linearization steps $k \geq 1$, $\mathcal{E}_\ell^k \leq \eta_\ell^k$.

Moreover, for k satisfying a stopping criterion, there holds

$$\eta_\ell^k \leq C_{\text{eff}}(d, \kappa_{\mathcal{T}}) C_\ell^k \mathcal{E}_\ell^k + \textit{quadrature error terms},$$

where

$$C_\ell^k := \max_{\boldsymbol{a} \in \mathcal{V}_\ell} \left(\frac{\textit{max. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_\ell^{\boldsymbol{a}}}}{\textit{min. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_\ell^{\boldsymbol{a}}}} \right) \begin{cases} = 1 & \textit{Zarantonello} \end{cases}$$

- ✓ $C_\ell^k = 1$ for Zarantonello $\Longrightarrow$ **robustness** wrt the **strength of nonlinearities**
- ✓ C_ℓ^k given by **local conditioning** of the linearization matrix $\boldsymbol{A}_\ell^{k-1}$:

A posteriori error estimates for an augmented energy difference

Theorem (A posteriori estimate of augmented energy)

For all linearization steps $k \geq 1$, $\mathcal{E}_\ell^k \leq \eta_\ell^k.$

Moreover, for k satisfying a stopping criterion, there holds

$$\eta_\ell^k \leq C_{\rm eff}(d, \kappa_{\mathcal{T}}) C_\ell^k \mathcal{E}_\ell^k + \text{quadrature error terms},$$

where

$$C_\ell^k := \max_{\boldsymbol{a} \in \mathcal{V}_\ell} \left(\frac{\text{max. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_\ell^{\boldsymbol{a}}}}{\text{min. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_\ell^{\boldsymbol{a}}}} \right) \begin{cases} = 1 & \text{Zarantonello} \\ \leq \dfrac{\text{max. eig. } \boldsymbol{A}_\ell^{k-1}|_\Omega}{\text{min. eig. } \boldsymbol{A}_\ell^{k-1}|_\Omega} \leq \dfrac{a_{\rm c}}{a_{\rm m}} & \text{in general.} \end{cases}$$

- ✓ $C_\ell^k = 1$ for Zarantonello $\Longrightarrow$ **robustness** wrt the **strength of nonlinearities**
- ✓ C_ℓ^k given by **local conditioning** of the linearization matrix $\boldsymbol{A}_\ell^{k-1}$:

A posteriori error estimates for an augmented energy difference

Theorem (A posteriori estimate of augmented energy)

For all linearization steps $k \geq 1$, $\mathcal{E}_\ell^k \leq \eta_\ell^k$.
Moreover, for k satisfying a stopping criterion, there holds

$$\eta_\ell^k \leq C_{\text{eff}}(d, \kappa_{\mathcal{T}}) C_\ell^k \mathcal{E}_\ell^k + \text{quadrature error terms},$$

where

$$C_\ell^k := \max_{\boldsymbol{a} \in \mathcal{V}_\ell} \left(\frac{\text{max. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_\ell^{\boldsymbol{a}}}}{\text{min. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_\ell^{\boldsymbol{a}}}} \right) \begin{cases} = 1 & \text{Zarantonello} \\ \leq \dfrac{\text{max. eig. } \boldsymbol{A}_\ell^{k-1}|_\Omega}{\text{min. eig. } \boldsymbol{A}_\ell^{k-1}|_\Omega} \leq \dfrac{a_c}{a_m} & \text{in general.} \end{cases}$$

- ✓ $C_\ell^k = 1$ for Zarantonello $\Longrightarrow$ **robustness** wrt the **strength of nonlinearities**
- ✓ C_ℓ^k given by **local conditioning** of the linearization matrix $\boldsymbol{A}_\ell^{k-1}$: typically **much better** than a_c/a_m,

A posteriori error estimates for an augmented energy difference

Theorem (A posteriori estimate of augmented energy)

For all linearization steps $k \geq 1$, $\mathcal{E}_\ell^k \leq \eta_\ell^k$.

Moreover, for k satisfying a stopping criterion, there holds

$$\eta_\ell^k \leq C_{\mathrm{eff}}(d, \kappa_{\mathcal{T}}) C_\ell^k \mathcal{E}_\ell^k + \textit{quadrature error terms},$$

where

$$C_\ell^k := \max_{\mathbf{a} \in \mathcal{V}_\ell} \left(\frac{\textit{max. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_\ell^{\mathbf{a}}}}{\textit{min. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_\ell^{\mathbf{a}}}} \right) \begin{cases} = 1 & \textit{Zarantonello} \\ \leq \dfrac{\textit{max. eig. } \boldsymbol{A}_\ell^{k-1}|_\Omega}{\textit{min. eig. } \boldsymbol{A}_\ell^{k-1}|_\Omega} \leq \dfrac{a_{\mathrm{c}}}{a_{\mathrm{m}}} & \textit{in general.} \end{cases}$$

- ✓ $C_\ell^k = 1$ for Zarantonello $\Longrightarrow$ **robustness** wrt the **strength of nonlinearities**
- ✓ C_ℓ^k given by **local conditioning** of the linearization matrix $\boldsymbol{A}_\ell^{k-1}$: typically **much better** than $a_{\mathrm{c}}/a_{\mathrm{m}}$, **improves** with **mesh refinement**

A posteriori error estimates for an augmented energy difference

Theorem (A posteriori estimate of augmented energy)

For all linearization steps $k \geq 1$, $\mathcal{E}_\ell^k \leq \eta_\ell^k$.
Moreover, for k satisfying a stopping criterion, there holds

$$\eta_\ell^k \leq C_{\mathrm{eff}}(d, \kappa_{\mathcal{T}}) C_\ell^k \mathcal{E}_\ell^k + \text{quadrature error terms},$$

where

$$C_\ell^k := \max_{\boldsymbol{a} \in \mathcal{V}_\ell} \left(\frac{\text{max. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_\ell^{\boldsymbol{a}}}}{\text{min. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_\ell^{\boldsymbol{a}}}} \right) \begin{cases} = 1 & \text{Zarantonello} \\ \leq \dfrac{\text{max. eig. } \boldsymbol{A}_\ell^{k-1}|_\Omega}{\text{min. eig. } \boldsymbol{A}_\ell^{k-1}|_\Omega} \leq \dfrac{a_{\mathrm{c}}}{a_{\mathrm{m}}} & \text{in general.} \end{cases}$$

- ✓ $C_\ell^k = 1$ for Zarantonello $\Longrightarrow$ **robustness** wrt the **strength of nonlinearities**
- ✓ C_ℓ^k given by **local conditioning** of the linearization matrix $\boldsymbol{A}_\ell^{k-1}$: typically **much better** than $a_{\mathrm{c}}/a_{\mathrm{m}}$, **improves** with **mesh refinement**
- ✓ C_ℓ^k **computable**: we can affirm **robustness *a posteriori***, for the given case

A posteriori error estimates for an augmented energy difference

Augmented energy difference

$$\mathcal{E}_\ell^k = \frac{1}{2}\text{energy difference} + \lambda_\ell^k \times \frac{1}{2}(\text{linearized energy difference})$$

A posteriori error estimates for an augmented energy difference

Augmented energy difference

$$\mathcal{E}_\ell^k = \frac{1}{2}\text{energy difference} + \lambda_\ell^k \times \frac{1}{2}(\text{linearized energy difference})$$

$$\mathcal{E}_\ell^k := \frac{1}{2}\underbrace{(\mathcal{J}(u_\ell^k) - \mathcal{J}(u))}_{\text{energy difference}}$$

A posteriori error estimates for an augmented energy difference

Augmented energy difference

$$\mathcal{E}_\ell^k = \frac{1}{2}\text{energy difference} + \lambda_\ell^k \times \frac{1}{2}(\text{linearized energy difference})$$

$$\mathcal{E}_\ell^k := \frac{1}{2}\underbrace{(\mathcal{J}(u_\ell^k) - \mathcal{J}(u))}_{\text{energy difference}} + \lambda_\ell^k \frac{1}{2}\underbrace{\left(\mathcal{J}_\ell^{k-1}(u_\ell^k) - \mathcal{J}_\ell^{k-1}(u_{\langle\ell\rangle}^k)\right)}_{\text{linearized en. diff.}}$$

A posteriori error estimates for an augmented energy difference

Augmented energy difference

$$\mathcal{E}_\ell^k = \frac{1}{2}\text{energy difference} + \lambda_\ell^k \times \frac{1}{2}(\text{linearized energy difference})$$

$$\mathcal{E}_\ell^k := \frac{1}{2}\underbrace{(\mathcal{J}(u_\ell^k) - \mathcal{J}(u))}_{\text{energy difference}} + \lambda_\ell^k \frac{1}{2}\underbrace{\left(\mathcal{J}_\ell^{k-1}(u_\ell^k) - \mathcal{J}_\ell^{k-1}(u_{\langle\ell\rangle}^k)\right)}_{\text{linearized en. diff.}}$$

$$\eta_\ell^k := \frac{1}{2}\underbrace{(\mathcal{J}(u_\ell^k) - \mathcal{J}^*(\boldsymbol{\sigma}_\ell^k))}_{\text{en. diff. estimate}}$$

A posteriori error estimates for an augmented energy difference

Augmented energy difference

$$\mathcal{E}_\ell^k = \frac{1}{2}\text{energy difference} + \lambda_\ell^k \times \frac{1}{2}(\text{linearized energy difference})$$

$$\mathcal{E}_\ell^k := \frac{1}{2}\underbrace{(\mathcal{J}(u_\ell^k) - \mathcal{J}(u))}_{\text{energy difference}} + \lambda_\ell^k \frac{1}{2}\Big(\underbrace{\mathcal{J}_\ell^{k-1}(u_\ell^k) - \mathcal{J}_\ell^{k-1}(u_{\langle\ell\rangle}^k)}_{\text{linearized en. diff.}}\Big)$$

$$\eta_\ell^k := \frac{1}{2}\underbrace{(\mathcal{J}(u_\ell^k) - \mathcal{J}^*(\boldsymbol{\sigma}_\ell^k))}_{\text{en. diff. estimate}} + \lambda_\ell^k \frac{1}{2}\Big(\underbrace{\mathcal{J}_\ell^{k-1}(u_\ell^k) - \mathcal{J}_\ell^{*,k-1}(\boldsymbol{\sigma}_\ell^k)}_{\text{linearized en. diff. estimate}}\Big)$$

A posteriori error estimates for an augmented energy difference

Augmented energy difference

$$\boxed{\mathcal{E}_\ell^k = \frac{1}{2}\text{energy difference} + \lambda_\ell^k \times \frac{1}{2}(\text{linearized energy difference})}$$

$$\mathcal{E}_\ell^k := \frac{1}{2}\underbrace{(\mathcal{J}(u_\ell^k) - \mathcal{J}(u))}_{\text{energy difference}} + \lambda_\ell^k \frac{1}{2}\Big(\underbrace{\mathcal{J}_\ell^{k-1}(u_\ell^k) - \mathcal{J}_\ell^{k-1}(u_{\langle\ell\rangle}^k)}_{\text{linearized en. diff.}}\Big)$$

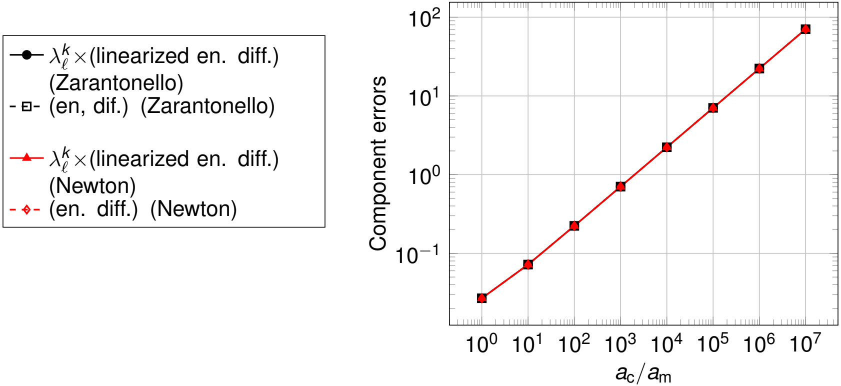
$$\eta_\ell^k := \frac{1}{2}\underbrace{(\mathcal{J}(u_\ell^k) - \mathcal{J}^*(\boldsymbol{\sigma}_\ell^k))}_{\text{en. diff. estimate}} + \lambda_\ell^k \frac{1}{2}\Big(\underbrace{\mathcal{J}_\ell^{k-1}(u_\ell^k) - \mathcal{J}_\ell^{*,k-1}(\boldsymbol{\sigma}_\ell^k)}_{\text{linearized en. diff. estimate}}\Big)$$

- λ_ℓ^k computable weight to make the two components comparable

A posteriori error estimates for an augmented energy difference

Augmented energy difference

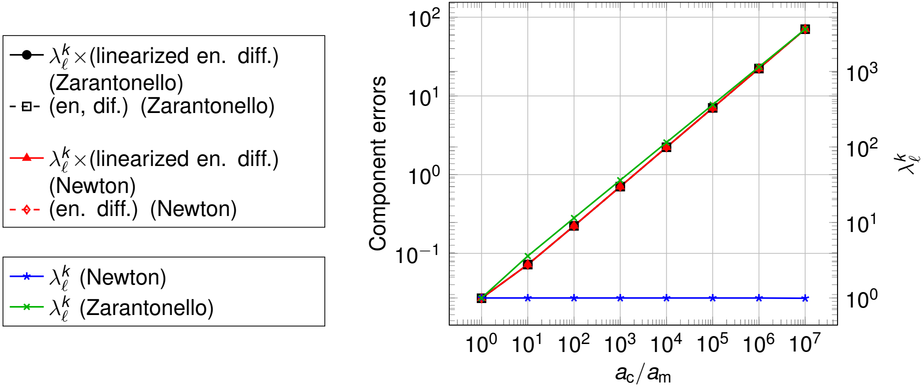
$$\mathcal{E}_\ell^k = \frac{1}{2}\text{energy difference} + \lambda_\ell^k \times \frac{1}{2}(\text{linearized energy difference})$$



A posteriori error estimates for an augmented energy difference

Augmented energy difference

$$\mathcal{E}_\ell^k = \frac{1}{2}\text{energy difference} + \lambda_\ell^k \times \frac{1}{2}(\text{linearized energy difference})$$



A posteriori error estimates for an augmented energy difference

Augmented energy difference

$$\mathcal{E}_\ell^k = \frac{1}{2}\text{energy difference} + \lambda_\ell^k \times \frac{1}{2}(\text{linearized energy difference})$$

Practically

$$\mathcal{E}_\ell^k = \mathcal{J}(u_\ell^k) - \mathcal{J}(u) \text{ at convergence}$$

Outline

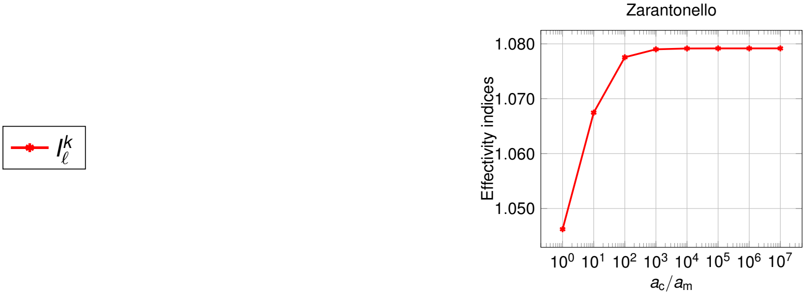
Smooth solution

Setting

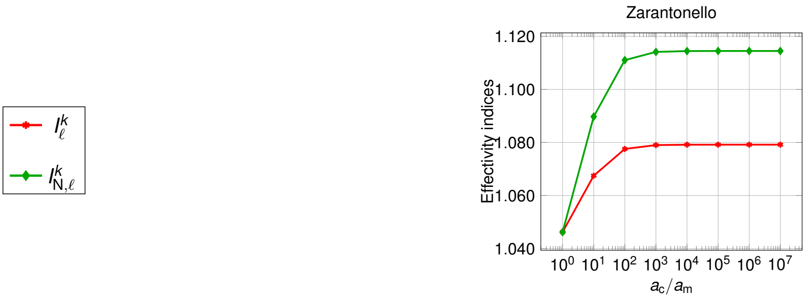
- unit square $\Omega = (0,1)^2$
- known smooth solution $u(x,y) := 10\,x(x-1)y(y-1)$
- $p = 1$
- effectivity indices

$$\underbrace{I_\ell^k := \left(\frac{\eta_\ell^k}{\mathcal{E}_\ell^k}\right)^{\frac{1}{2}}}_{\text{total}}, \quad \underbrace{I_{\mathrm{N},\ell}^k := \left(\frac{\mathcal{J}(u_\ell^k) - \mathcal{J}^*(\boldsymbol{\sigma}_\ell^k)}{\mathcal{J}(u_\ell^k) - \mathcal{J}(u)}\right)^{\frac{1}{2}}}_{\text{energy difference}}$$

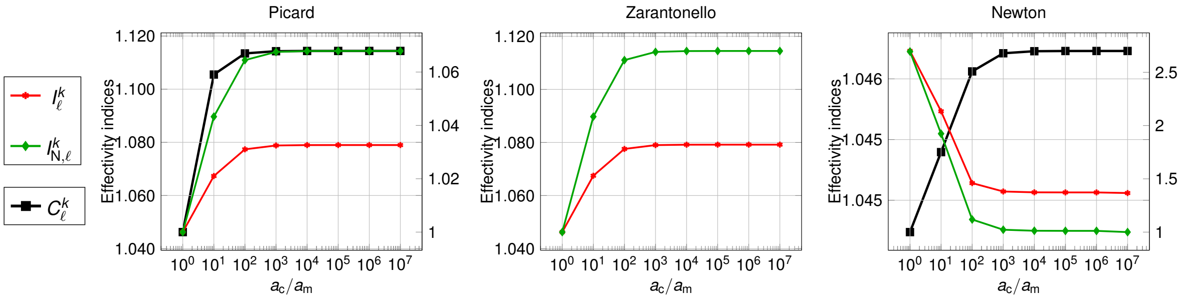
How **large** is the error? **Robustness** wrt the nonlinearities $(a(r) = a_{\mathrm{m}} + \frac{a_{\mathrm{c}} - a_{\mathrm{m}}}{\sqrt{1+r^2}})$



How **large** is the error? **Robustness** wrt the nonlinearities $\left(a(r) = a_{\mathrm{m}} + \frac{a_{\mathrm{c}} - a_{\mathrm{m}}}{\sqrt{1+r^2}}\right)$



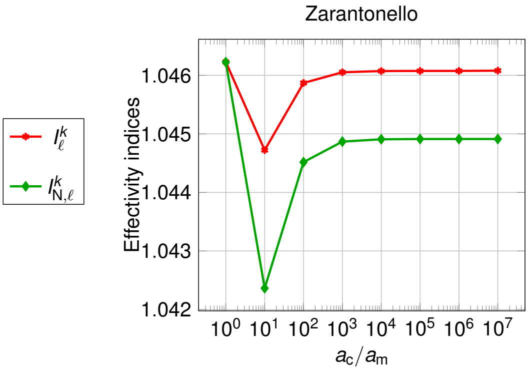
How **large** is the error? **Robustness** wrt the nonlinearities $\left(a(r) = a_{\mathrm{m}} + \frac{a_{\mathrm{c}} - a_{\mathrm{m}}}{\sqrt{1+r^2}}\right)$



A. Harnist, K. Mitra, A. Rappaport, M. Vohralík, preprint (2023)

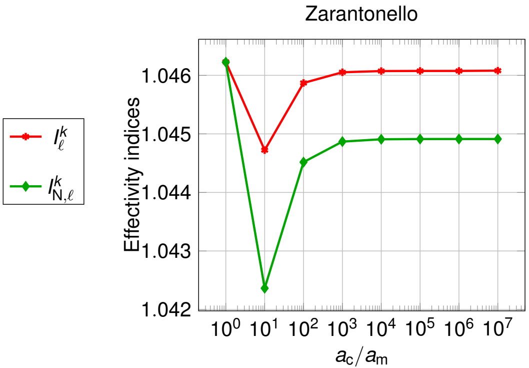
How **large** is the error? **Robustness** wrt the nonlinearities

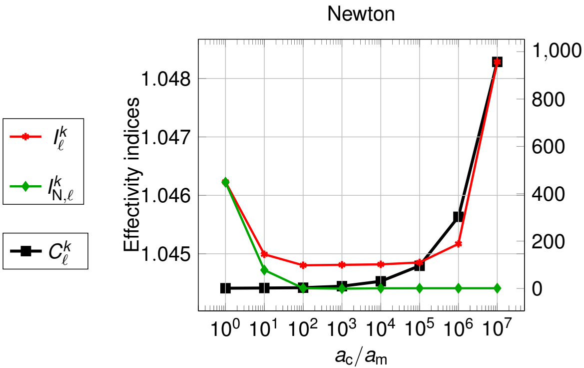
$$\left(a(r) = a_m + (a_c - a_m)\frac{1-e^{-\frac{3}{2}r^2}}{1+2e^{-\frac{3}{2}}}\right)$$



How **large** is the error? **Robustness** wrt the nonlinearities

$(a(r) = a_{\rm m} + (a_{\rm c} - a_{\rm m})\frac{1-e^{-\frac{3}{2}r^2}}{1+2e^{-\frac{3}{2}}}$, robustness only for Zarantonello)





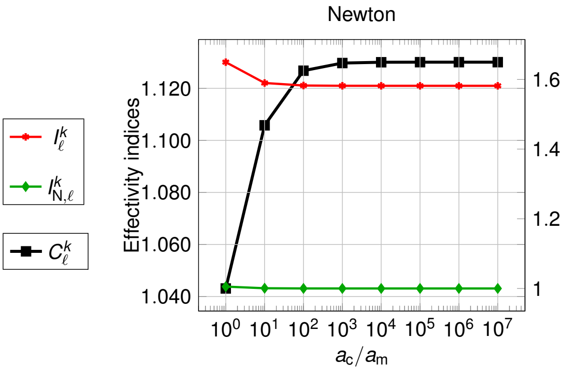
A. Harnist, K. Mitra, A. Rappaport, M. Vohralík, preprint (2023)

Singular solution

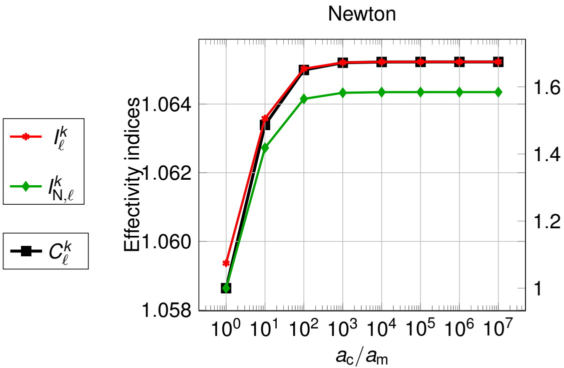
Setting

- L-shaped domain $\Omega = (-1,1)^2 \setminus ([0,1) \times (-1,0])$
- known singular solution $u(\rho,\theta) = \rho^{\frac{2}{3}} \sin(\frac{2}{3}\theta)$
- $a(r) = a_{\mathrm{m}} + (a_{\mathrm{c}} - a_{\mathrm{m}})\frac{1-e^{-\frac{3}{2}r^2}}{1+2e^{-\frac{3}{2}}}$
- $p = 1$
- uniform or adaptive mesh refinement

How **large** is the error? **Robustness** wrt the nonlinearities



Uniform mesh refinement



Adaptive mesh refinement

A. Harnist, K. Mitra, A. Rappaport, M. Vohralík, preprint (2023)

Outline

Observation

Observation

Not all nonlinear problems admit an energy minimization structure.

A model steady nonlinear problem

Nonlinear elliptic problem

Find $u : \Omega \to \mathbb{R}$ such that

$$-\nabla\cdot\big(\tau \boldsymbol{K}(\boldsymbol{x})(\underbrace{\mathcal{D}(\boldsymbol{x},u)}_{\text{diffusion}}\nabla u + \underbrace{\boldsymbol{q}(\boldsymbol{x},u)}_{\text{advection}})\big) + \underbrace{f(\boldsymbol{x},u)}_{\text{reaction}} = 0 \quad \text{in} \quad \Omega,$$

$$u = 0 \quad \text{on} \quad \partial\Omega.$$

- $\tau > 0$ a parameter (time step size in transient problems: applies to Richards on each time step)

Assumption (Nonlinear functions $\mathcal{D}$, $\boldsymbol{q}$, and f)

$|\mathcal{D}(\boldsymbol{x}_1,u_1) - \mathcal{D}(\boldsymbol{x}_2,u_2)| \le \mathcal{D}_\mathrm{M}(|\boldsymbol{x}_1 - \boldsymbol{x}_2| + |u_1 - u_2|) \quad \forall \boldsymbol{x}_1,\boldsymbol{x}_2 \in \Omega$ *and* $u_1,u_2 \in \mathbb{R}$,

$0 \le f(\boldsymbol{x},u_2) - f(\boldsymbol{x},u_1) \le f_\mathrm{M}(u_2 - u_1) \quad \forall \boldsymbol{x} \in \Omega$ *and* $u_1,u_2 \in \mathbb{R}, u_2 \ge u_1$,

$\boldsymbol{q}$ *is "small" wrt* $\boldsymbol{K}\mathcal{D}$.

A model steady nonlinear problem

Nonlinear elliptic problem

Find $u : \Omega \to \mathbb{R}$ such that

$$-\nabla\cdot\big(\tau \boldsymbol{K}(\boldsymbol{x})(\underbrace{\mathcal{D}(\boldsymbol{x},u)}_{\text{diffusion}}\nabla u + \underbrace{\boldsymbol{q}(\boldsymbol{x},u)}_{\text{advection}})\big) + \underbrace{f(\boldsymbol{x},u)}_{\text{reaction}} = 0 \quad \text{in} \quad \Omega,$$

$$u = 0 \quad \text{on} \quad \partial\Omega.$$

- $\tau > 0$ a parameter (time step size in transient problems: applies to Richards on each time step)

Assumption (Nonlinear functions $\mathcal{D}$, $\boldsymbol{q}$, and f)

$|\mathcal{D}(\boldsymbol{x}_1,u_1) - \mathcal{D}(\boldsymbol{x}_2,u_2)| \leq \mathcal{D}_{\mathrm{M}}(|\boldsymbol{x}_1 - \boldsymbol{x}_2| + |u_1 - u_2|) \quad \forall \boldsymbol{x}_1,\boldsymbol{x}_2 \in \Omega$ *and* $u_1,u_2 \in \mathbb{R}$,

$0 \leq f(\boldsymbol{x},u_2) - f(\boldsymbol{x},u_1) \leq f_{\mathrm{M}}(u_2 - u_1) \quad \forall \boldsymbol{x} \in \Omega$ *and* $u_1,u_2 \in \mathbb{R}, u_2 \geq u_1$,

$\boldsymbol{q}$ *is "small" wrt* $\boldsymbol{K}\mathcal{D}$.

Finite element discretization and iterative linearization

Definition (Linearized finite element approximation)

$u_\ell^k \in V_\ell^p$ such that

$$((u_\ell^k - u_\ell^{k-1}, v_\ell))_{u_\ell^{k-1}} = -\langle \underbrace{\mathcal{R}(u_\ell^{k-1})}_{\text{residual}}, v_\ell \rangle \qquad \forall v_\ell \in V_\ell^p.$$

- covers most linearization schemes: Picard (fixed-point), L & M-schemes, ...
- linearization: **reaction–diffusion scalar product**

$$((w, v))_{u_\ell^{k-1}} := (\underbrace{L_\ell^{k-1}}_{\text{reaction coef.}} w, v) + (\underbrace{A_\ell^{k-1}}_{\text{diffusion coef.}} \nabla w, \nabla v), \quad w, v \in H_0^1(\Omega)$$

Iteration-dependent norm

- $$|||v|||^2_{1,u_\ell^{k-1}} := ((v, v))_{u_\ell^{k-1}} = \|(L_\ell^{k-1})^{1/2} v\|^2 + \|(A_\ell^{k-1})^{1/2} \nabla v\|^2, \quad v \in H_0^1(\Omega)$$
- induced by the linearization scalar product

Finite element discretization and iterative linearization

Definition (Linearized finite element approximation)

$u_\ell^k \in V_\ell^p$ such that

$$((u_\ell^k - u_\ell^{k-1}, v_\ell))_{u_\ell^{k-1}} = -\underbrace{\langle \mathcal{R}(u_\ell^{k-1}), v_\ell \rangle}_{\text{residual}} \qquad \forall v_\ell \in V_\ell^p.$$

- covers most linearization schemes: Picard (fixed-point), L & M-schemes, ...
- linearization: **reaction–diffusion scalar product**

$$((w, v))_{u_\ell^{k-1}} := (\underbrace{L_\ell^{k-1}}_{\text{reaction coef. } =0 \text{ if } f=f(x)} w, v) + (\underbrace{\mathbf{A}_\ell^{k-1}}_{\text{diffusion coef.}} \nabla w, \nabla v), \quad w, v \in H_0^1(\Omega)$$

Iteration-dependent norm

- $|||v|||^2_{1,u_\ell^{k-1}} := ((v, v))_{u_\ell^{k-1}} = \|(L_\ell^{k-1})^{1/2} v\|^2 + \|(\mathbf{A}_\ell^{k-1})^{1/2} \nabla v\|^2, \quad v \in H_0^1(\Omega)$
- induced by the linearization scalar product

Finite element discretization and iterative linearization

Definition (Linearized finite element approximation)

$u_\ell^k \in V_\ell^p$ such that

$$((u_\ell^k - u_\ell^{k-1}, v_\ell))_{u_\ell^{k-1}} = -\langle \underbrace{\mathcal{R}(u_\ell^{k-1})}_{\text{residual}}, v_\ell \rangle \qquad \forall v_\ell \in V_\ell^p.$$

- covers most linearization schemes: Picard (fixed-point), L & M-schemes, ...
- linearization: **reaction–diffusion scalar product**

$$((w, v))_{u_\ell^{k-1}} := (\underbrace{L_\ell^{k-1}}_{\text{reaction coef. } =0 \text{ if } f=f(\boldsymbol{x})} w, v) + (\underbrace{\boldsymbol{A}_\ell^{k-1}}_{\text{diffusion coef. } =\tau \boldsymbol{K}(\boldsymbol{x})\mathcal{D}(\boldsymbol{x}, u_\ell^{k-1})} \nabla w, \nabla v), \quad w, v \in H_0^1(\Omega)$$

Iteration-dependent norm

- $$|||v|||_{1,u_\ell^{k-1}}^2 := ((v, v))_{u_\ell^{k-1}} = \|(L_\ell^{k-1})^{1/2} v\|^2 + \|(\boldsymbol{A}_\ell^{k-1})^{1/2} \nabla v\|^2, \quad v \in H_0^1(\Omega)$$
- induced by the linearization scalar product

Finite element discretization and iterative linearization

Definition (Linearized finite element approximation)

$u_\ell^k \in V_\ell^p$ such that

$$((u_\ell^k - u_\ell^{k-1}, v_\ell))_{u_\ell^{k-1}} = -\langle \underbrace{\mathcal{R}(u_\ell^{k-1})}_{\text{residual}}, v_\ell \rangle \qquad \forall v_\ell \in V_\ell^p.$$

- covers most linearization schemes: Picard (fixed-point), L & M-schemes, ...
- linearization: **reaction–diffusion scalar product**

$$((w, v))_{u_\ell^{k-1}} := (\underbrace{L_\ell^{k-1}}_{\text{reaction coef. } =0 \text{ if } f=f(\boldsymbol{x})} w, v) + (\underbrace{\boldsymbol{A}_\ell^{k-1}}_{\text{diffusion coef. } =\tau \boldsymbol{K}(\boldsymbol{x})\mathcal{D}(\boldsymbol{x}, u_\ell^{k-1})} \nabla w, \nabla v), \quad w, v \in H_0^1(\Omega)$$

Iteration-dependent norm

- $$|||v|||_{1,u_\ell^{k-1}}^2 := ((v, v))_{u_\ell^{k-1}} = \|(L_\ell^{k-1})^{1/2} v\|^2 + \|(\boldsymbol{A}_\ell^{k-1})^{1/2} \nabla v\|^2, \quad v \in H_0^1(\Omega)$$
- induced by the linearization scalar product

Finite element discretization and iterative linearization

Definition (Linearized finite element approximation)

$u_\ell^k \in V_\ell^p$ such that

$$((u_\ell^k - u_\ell^{k-1}, v_\ell))_{u_\ell^{k-1}} = -\langle \underbrace{\mathcal{R}(u_\ell^{k-1})}_{\text{residual}}, v_\ell \rangle \qquad \forall v_\ell \in V_\ell^p.$$

- covers most linearization schemes: Picard (fixed-point), L & M-schemes, ...
- linearization: **reaction–diffusion scalar product**

$$((w, v))_{u_\ell^{k-1}} := (\underbrace{L_\ell^{k-1}}_{\text{reaction coef. } =0 \text{ if } f=f(\boldsymbol{x})} w, v) + (\underbrace{\boldsymbol{A}_\ell^{k-1}}_{\text{diffusion coef. } =\tau \boldsymbol{K}(\boldsymbol{x})\mathcal{D}(\boldsymbol{x}, u_\ell^{k-1})} \nabla w, \nabla v), \quad w, v \in H_0^1(\Omega)$$

Iteration-dependent norm

- $$|||v|||^2_{1,u_\ell^{k-1}} := ((v, v))_{u_\ell^{k-1}} = \|(L_\ell^{k-1})^{1/2} v\|^2 + \|(\boldsymbol{A}_\ell^{k-1})^{1/2} \nabla v\|^2, \quad v \in H_0^1(\Omega)$$
- induced by the linearization scalar product

Finite element discretization and iterative linearization

Definition (Linearized finite element approximation)

$u_\ell^k \in V_\ell^p$ such that

$$((u_\ell^k - u_\ell^{k-1}, v_\ell))_{u_\ell^{k-1}} = -\langle \underbrace{\mathcal{R}(u_\ell^{k-1})}_{\text{residual}}, v_\ell \rangle \qquad \forall v_\ell \in V_\ell^p.$$

- covers most linearization schemes: Picard (fixed-point), L & M-schemes, ...
- linearization: **reaction–diffusion scalar product**

$$((w, v))_{u_\ell^{k-1}} := (\underbrace{L_\ell^{k-1}}_{\text{reaction coef. } =0 \text{ if } f=f(\boldsymbol{x})} w, v) + (\underbrace{\boldsymbol{A}_\ell^{k-1}}_{\text{diffusion coef. } =\tau \boldsymbol{K}(\boldsymbol{x})\mathcal{D}(\boldsymbol{x}, u_\ell^{k-1})} \nabla w, \nabla v), \quad w, v \in H_0^1(\Omega)$$

Iteration-dependent norm

- $$|||v|||_{1,u_\ell^{k-1}}^2 := ((v, v))_{u_\ell^{k-1}} = \|(L_\ell^{k-1})^{1/2} v\|^2 + \|(\boldsymbol{A}_\ell^{k-1})^{1/2} \nabla v\|^2, \quad v \in H_0^1(\Omega)$$
- induced by the linearization scalar product

An orthogonal decomposition of the total residual/error

Theorem (Orthogonal decomposition of the total error into linearization and discretization components)

For all linearization steps $k \geq 1$, there holds

$$\underbrace{|||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1}}^2}_{\substack{\text{total residual/error} \\ |||u_\ell^{k-1}-u_{\langle\ell\rangle}^k|||_{1,u_\ell^{k-1}}}} = \underbrace{|||u_\ell^{k-1} - u_\ell^k|||_{1,u_\ell^{k-1}}^2}_{\substack{\text{linearization} \\ \text{error}}} + \underbrace{|||\mathcal{R}_{\mathrm{disc}}^{u_\ell^{k-1}}(u_\ell^k)|||_{-1,u_\ell^{k-1}}^2}_{\substack{\text{discretization residual/error} \\ |||u_\ell^k-u_{\langle\ell\rangle}^k|||_{1,u_\ell^{k-1}}}}.$$

- **orthogonal decomposition**
- **error components**
- $u_{\langle\ell\rangle}^k \in H_0^1(\Omega)$ such that

$$((u_{\langle\ell\rangle}^k - u_\ell^{k-1}, v))_{u_\ell^{k-1}} = \underbrace{-\langle\mathcal{R}(u_\ell^{k-1}), v\rangle}_{\text{residual}} \qquad \forall v \in H_0^1(\Omega)$$

An orthogonal decomposition of the total residual/error

Theorem (Orthogonal decomposition of the total error into linearization and discretization components)

For all linearization steps $k \geq 1$, there holds

$$\underbrace{|||\mathcal{R}(u_\ell^{k-1})|||^2_{-1,u_\ell^{k-1}}}_{\substack{\text{total residual/error} \\ |||u_\ell^{k-1} - u_{\langle \ell \rangle}^k|||_{1,u_\ell^{k-1}}}} = \underbrace{|||u_\ell^{k-1} - u_\ell^k|||^2_{1,u_\ell^{k-1}}}_{\substack{\text{linearization} \\ \text{error}}} + \underbrace{|||\mathcal{R}_{\text{disc}}^{u_\ell^{k-1}}(u_\ell^k)|||^2_{-1,u_\ell^{k-1}}}_{\substack{\text{discretization residual/error} \\ |||u_\ell^k - u_{\langle \ell \rangle}^k|||_{1,u_\ell^{k-1}}}}.$$

- **orthogonal decomposition**
- **error components**
- $u_{\langle \ell \rangle}^k \in H_0^1(\Omega)$ such that

$$((u_{\langle \ell \rangle}^k - u_\ell^{k-1}, v))_{u_\ell^{k-1}} = -\underbrace{\langle \mathcal{R}(u_\ell^{k-1}), v \rangle}_{\text{residual}} \qquad \forall v \in H_0^1(\Omega)$$

An orthogonal decomposition of the total residual/error

Theorem (Orthogonal decomposition of the total error into linearization and discretization components)

For all linearization steps $k \geq 1$, there holds

$$\underbrace{|||\mathcal{R}(u_\ell^{k-1})|||^2_{-1,u_\ell^{k-1}}}_{\substack{\text{total residual/error} \\ |||u_\ell^{k-1}-u_{\langle\ell\rangle}^k|||_{1,u_\ell^{k-1}}}} = \underbrace{|||u_\ell^{k-1} - u_\ell^k|||^2_{1,u_\ell^{k-1}}}_{\substack{\text{linearization} \\ \text{error}}} + \underbrace{|||\mathcal{R}_{\mathrm{disc}}^{u_\ell^{k-1}}(u_\ell^k)|||^2_{-1,u_\ell^{k-1}}}_{\substack{\text{discretization residual/error} \\ |||u_\ell^k-u_{\langle\ell\rangle}^k|||_{1,u_\ell^{k-1}}}}.$$

- **orthogonal decomposition**
- **error components**
- $u_{\langle\ell\rangle}^k \in H_0^1(\Omega)$ such that

$$((u_{\langle\ell\rangle}^k - u_\ell^{k-1}, v))_{u_\ell^{k-1}} = -\langle \underbrace{\mathcal{R}(u_\ell^{k-1})}_{\text{residual}}, v\rangle \qquad \forall v \in H_0^1(\Omega)$$

Outline

A posteriori error estimates for an iteration-dependent norm

Theorem (A posteriori estimate of iteration-dependent norm)

For all linearization steps $k \geq 1$,

$$|||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1}} \leq \eta(u_\ell^k).$$

A posteriori error estimates for an iteration-dependent norm

Theorem (A posteriori estimate of iteration-dependent norm)

For all linearization steps $k \geq 1$,

$$|||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1}} \leq \eta(u_\ell^k).$$

Moreover, for all linearization steps $k \geq 1$, there holds

$$\eta(u_\ell^k) \leq C_{\text{eff}}(d, \kappa_{\mathcal{T}}\quad) C_\ell^k |||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1}} \quad + \text{quadrature error terms},$$

A posteriori error estimates for an iteration-dependent norm

Theorem (A posteriori estimate of iteration-dependent norm)

For all linearization steps $k \geq 1$,

$$|||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1}} \leq \eta(u_\ell^k).$$

Moreover, for all linearization steps $k \geq 1$ and **for each element** *$K \in \mathcal{T}_\ell$, there holds*

$$\eta_K(u_\ell^k) \leq C_{\text{eff}}(d, \kappa_{\mathcal{T}})C_K^k |||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1},\omega_K} + \textit{quadrature error terms},$$

A posteriori error estimates for an iteration-dependent norm

Theorem (A posteriori estimate of iteration-dependent norm)

For all linearization steps $k \geq 1$,

$$|||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1}} \leq \eta(u_\ell^k).$$

Moreover, for all linearization steps $k \geq 1$ and **for each element** *$K \in \mathcal{T}_\ell$, there holds*

$$\eta_K(u_\ell^k) \leq C_{\mathrm{eff}}(d, \kappa_{\mathcal{T}}, p) C_K^k |||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1},\omega_K} + \textit{quadrature error terms},$$

where

$$C_K^k := \left(\frac{\textit{max. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_K}}{\textit{min. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_K}} \right)^{1/2} + \left(\frac{\textit{max. } L_\ell^{k-1}|_{\omega_K}}{\textit{min. } L_\ell^{k-1}|_{\omega_K}} \right)^{1/2} \quad \textit{if react. dom.}$$

A posteriori error estimates for an iteration-dependent norm

Theorem (A posteriori estimate of iteration-dependent norm)

For all linearization steps $k \geq 1$,

$$|||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1}} \leq \eta(u_\ell^k).$$

Moreover, for all linearization steps $k \geq 1$ and **for each element** *$K \in \mathcal{T}_\ell$, there holds*

$$\eta_K(u_\ell^k) \leq C_{\mathrm{eff}}(d, \kappa_{\mathcal{T}}, p) C_K^k |||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1},\omega_K} + \text{quadrature error terms},$$

where

$$C_K^k := \left(\frac{\text{max. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_K}}{\text{min. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_K}}\right)^{1/2} + \left(\frac{\text{max. } L_\ell^{k-1}|_{\omega_K}}{\text{min. } L_\ell^{k-1}|_{\omega_K}}\right)^{1/2} \quad \text{if react. dom.}$$

✓ C_K^k given by **local conditioning** of the linearization matrix $\boldsymbol{A}_\ell^{k-1}$: typically **much better** than global conditioning (= worst-case scenario)

A posteriori error estimates for an iteration-dependent norm

Theorem (A posteriori estimate of iteration-dependent norm)

For all linearization steps $k \geq 1$,

$$|||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1}} \leq \eta(u_\ell^k).$$

Moreover, for all linearization steps $k \geq 1$ and **for each element** *$K \in \mathcal{T}_\ell$, there holds*

$$\eta_K(u_\ell^k) \leq C_{\text{eff}}(d, \kappa_{\mathcal{T}}, p) C_K^k |||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1},\omega_K} + \text{quadrature error terms},$$

where

$$C_K^k := \left(\frac{\text{max. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_K}}{\text{min. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_K}}\right)^{1/2} + \left(\frac{\text{max. } L_\ell^{k-1}|_{\omega_K}}{\text{min. } L_\ell^{k-1}|_{\omega_K}}\right)^{1/2} \quad \text{if react. dom.}$$

- ✓ C_K^k given by **local conditioning** of the linearization matrix $\boldsymbol{A}_\ell^{k-1}$: typically **much better** than global conditioning (= worst-case scenario)
- ✓ C_K^k **computable**: we can affirm **robustness *a posteriori***, for the given case

A posteriori error estimates for an iteration-dependent norm

Theorem (A posteriori estimate of iteration-dependent norm)

For all linearization steps $k \geq 1$,

$$|||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1}} \leq \eta(u_\ell^k).$$

Moreover, for all linearization steps $k \geq 1$ and **for each element** *$K \in \mathcal{T}_\ell$, there holds*

$$\eta_K(u_\ell^k) \leq C_{\mathrm{eff}}(d, \kappa_{\mathcal{T}}, p) C_K^k |||\mathcal{R}(u_\ell^{k-1})|||_{-1,u_\ell^{k-1},\omega_K} + \textit{quadrature error terms},$$

where

$$C_K^k := \left(\frac{\textit{max. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_K}}{\textit{min. eig. } \boldsymbol{A}_\ell^{k-1}|_{\omega_K}}\right)^{1/2} + \left(\frac{\textit{max. } L_\ell^{k-1}|_{\omega_K}}{\textit{min. } L_\ell^{k-1}|_{\omega_K}}\right)^{1/2} \quad \textit{if react. dom.}$$

- ✓ C_K^k given by **local conditioning** of the linearization matrix $\boldsymbol{A}_\ell^{k-1}$: typically **much better** than global conditioning (= worst-case scenario)
- ✓ C_K^k **computable**: we can affirm **robustness *a posteriori***, for the given case
- ✓ **local efficiency**

Outline

One time step of the Richards equation

Setting

- unit square $\Omega = (0,1)^2$
- realistic data

$$f(\boldsymbol{x},u) = S(u) - S(u_\ell^{n-1}(\boldsymbol{x})), \quad \mathcal{D}(\boldsymbol{x},u) = \kappa(S(u)), \quad \boldsymbol{q}(\boldsymbol{x},u) = -\kappa(S(u))\,\boldsymbol{g},$$

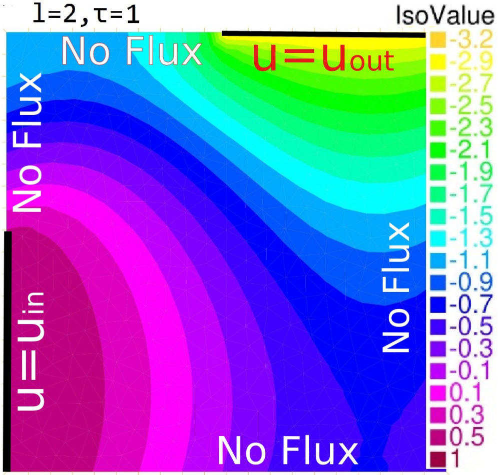
$$\boldsymbol{K} = \begin{bmatrix} 1 & 0.2 \\ 0.2 & 1 \end{bmatrix}, \quad \boldsymbol{g} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

- **van Genuchten saturation** and **permeability** laws

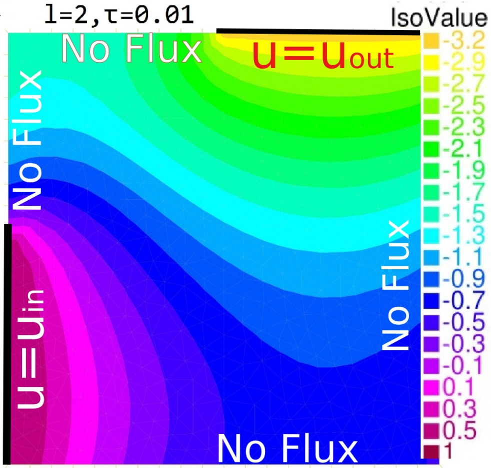
$$S(u) := \left(1 + (2-u)^{\frac{1}{1-\lambda}}\right)^{-\lambda}, \quad \kappa(s) := \sqrt{s}\left(1 - (1 - s^{\frac{1}{\lambda}})^{\lambda}\right)^2, \quad \lambda = 0.5$$

- time step length $\tau \in [10^{-3}, 1]$

One time step of the Richards equation: saturation u

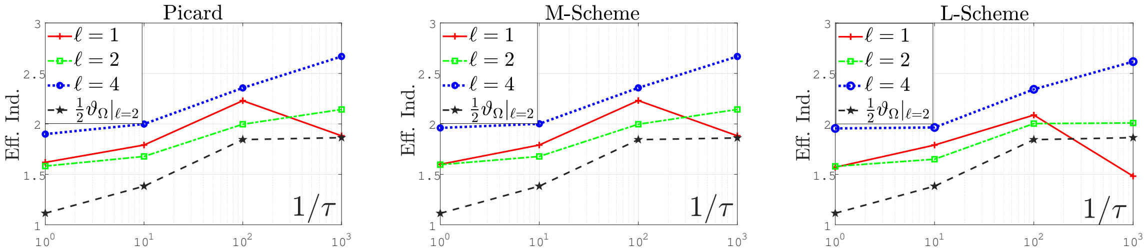


Time step length $\tau = 1$



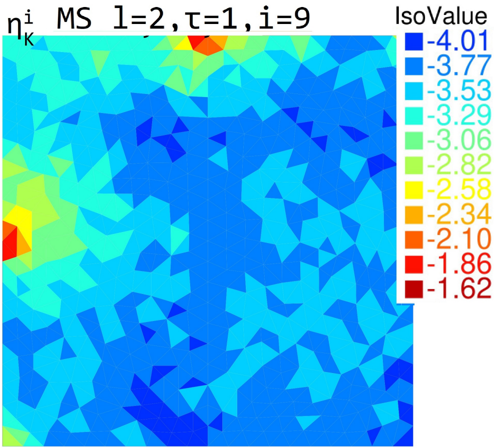
Time step length $\tau = 0.01$

How **large** is the error? **Robustness** wrt the nonlinearities

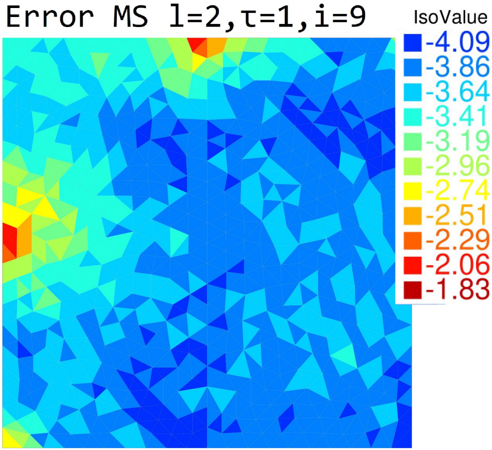


K. Mitra, M. Vohralík, preprint (2023)

Where is the error **localized**?

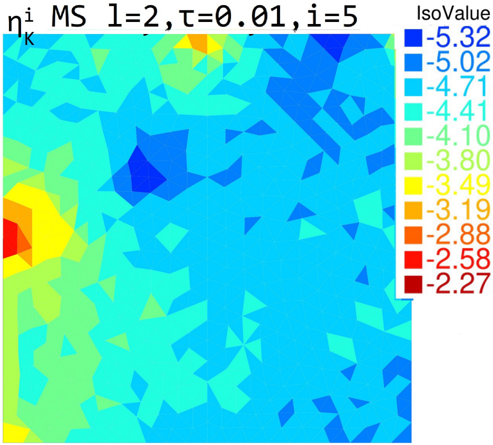


Estimated local error, $\tau = 1$

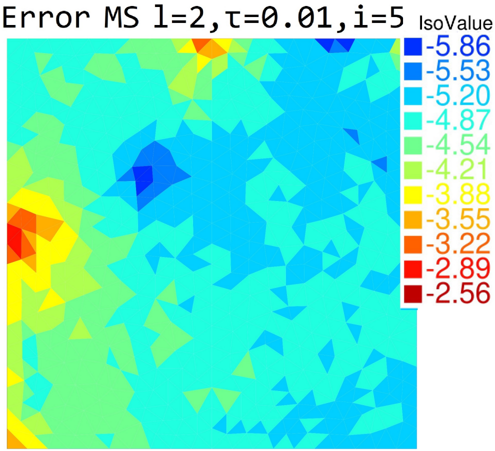


Exact local error, $\tau = 1$

Where is the error **localized**?



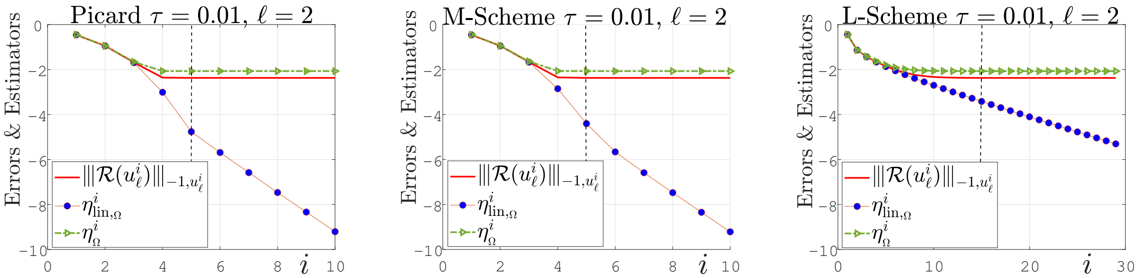
Estimated local error, $\tau = 0.01$



Exact local error, $\tau = 0.01$

K. Mitra, M. Vohralík, preprint (2023)

Error components and adaptivity via **stopping criteria**



Time step length $\tau = 0.01$

K. Mitra, M. Vohralík, preprint (2023)

Outline

A model unsteady linear problem

The unsteady linear Darcy (heat) equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t u - \Delta u &= f && \text{in } \Omega \times (0, T), \\ u &= 0 && \text{on } \partial\Omega \times (0, T), \\ u(0) &= u_0 && \text{in } \Omega. \end{aligned}$$

- T: final time
- f and u_0 piecewise polynomial for simplicity

Spaces and norms

$$\begin{aligned} X &:= L^2(0, T; H_0^1(\Omega)), \\ \|v\|_X^2 &:= \int_0^T \|\nabla v\|^2 \,\mathrm{d}t, \\ Y &:= L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)), \\ \|v\|_Y^2 &:= \int_0^T \|\partial_t v\|_{H^{-1}(\Omega)}^2 + \|\nabla v\|^2 \,\mathrm{d}t + \|v(T)\|^2 \end{aligned}$$

A model unsteady linear problem

The unsteady linear Darcy (heat) equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t u - \Delta u &= f \quad \text{in } \Omega \times (0, T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\ u(0) &= u_0 \quad \text{in } \Omega. \end{aligned}$$

- T: final time
- f and u_0 piecewise polynomial for simplicity

Spaces and norms

$$\begin{aligned} X &:= L^2(0, T; H_0^1(\Omega)), \\ \|v\|_X^2 &:= \int_0^T \|\nabla v\|^2 \, \mathrm{d}t, \\ Y &:= L^2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)), \\ \|v\|_Y^2 &:= \int_0^T \|\partial_t v\|_{H^{-1}(\Omega)}^2 + \|\nabla v\|^2 \, \mathrm{d}t + \|v(T)\|^2 \end{aligned}$$

Weak solution

Definition (Weak solution)

$u \in Y$ with $u(0) = u_0$ such that

$$\int_0^T \langle \partial_t u, v \rangle + (\nabla u, \nabla v) \, \mathrm{d}t = \int_0^T (f, v) \, \mathrm{d}t \qquad \forall v \in X.$$

Nonsymmetry

Trial space Y, test space X.

Weak solution

Definition (Weak solution)

$u \in Y$ with $u(0) = u_0$ such that

$$\int_0^T \langle \partial_t u, v \rangle + (\nabla u, \nabla v) \, \mathrm{d}t = \int_0^T (f, v) \, \mathrm{d}t \qquad \forall v \in X.$$

Nonsymmetry

Trial space Y, test space X.

Previous results

- Picasso / Verfürth (1998), work with the energy norm of X:
  - ✓ upper bound $\|u - u_\ell\|_X^2 \leq C^2 \sum_{n=1}^N \sum_{K \in \mathcal{T}_\ell^n} \eta_K^n(u_\ell)^2$
  - ✗ constrained lower bound (number of mesh elements $|\mathcal{T}_\ell^n|$ and time step τ strongly linked)
- Verfürth (2003) (cf. also Bergam, Bernardi, and Mghazli (2005)), work with the Y norm:
  - ✓ upper bound $\|u - \mathcal{I}u_\ell\|_Y^2 \leq C^2 \sum_{n=1}^N \sum_{K \in \mathcal{T}_\ell^n} \eta_K^n(u_\ell)^2$
  - ✓ efficiency $\sum_{K \in \mathcal{T}_\ell^n} \eta_K^n(u_\ell)^2 \leq C^2 \|u - \mathcal{I}u_\ell\|_{Y(I_n)}^2$
  - ✓ robustness with respect to the final time T, no link $|\mathcal{T}_\ell^n| \leftrightarrow \tau$
  - ✗ efficiency local in time but global in space
  - ✗ restrictions on mesh coarsening between time steps

Previous results

- Picasso / Verfürth (1998), work with the energy norm of X:
  - ✓ upper bound $\|u - u_\ell\|_X^2 \leq C^2 \sum_{n=1}^N \sum_{K\in\mathcal{T}_\ell^n} \eta_K^n(u_\ell)^2$
  - ✗ constrained lower bound (number of mesh elements $|\mathcal{T}_\ell^n|$ and time step τ strongly linked)
- Verfürth (2003) (cf. also Bergam, Bernardi, and Mghazli (2005)), work with the Y norm:
  - ✓ upper bound $\|u - \mathcal{I}u_\ell\|_Y^2 \leq C^2 \sum_{n=1}^N \sum_{K\in\mathcal{T}_\ell^n} \eta_K^n(u_\ell)^2$
  - ✓ efficiency $\sum_{K\in\mathcal{T}_\ell^n} \eta_K^n(u_\ell)^2 \leq C^2 \|u - \mathcal{I}u_\ell\|_{Y(I_n)}^2$
  - ✓ robustness with respect to the final time T, no link $|\mathcal{T}_\ell^n| \leftrightarrow \tau$
  - ✗ efficiency local in time but global in space
  - ✗ restrictions on mesh coarsening between time steps

Guaranteed upper bound

Augmented Y norm

$$\|u - u_\ell\|_{\mathcal{E}_Y}^2 := \|u - \mathcal{I}u_\ell\|_Y^2 + \underbrace{\|u_\ell - \mathcal{I}u_\ell\|_X^2}_{\text{known, computable, measures time jumps}}$$

Guaranteed upper bound

Augmented Y norm

$$\|u - u_\ell\|_{\mathcal{E}_Y}^2 := \|u - \mathcal{I}u_\ell\|_Y^2 + \underbrace{\|u_\ell - \mathcal{I}u_\ell\|_X^2}_{\text{known, computable, measures time jumps}}$$

Theorem (Guaranteed and locally space–time efficient estimate)

There holds

$$\|u - u_\ell\|_{\mathcal{E}_Y}^2 \leq \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_\ell^n} \int_{I_n} \|\boldsymbol{\sigma}_\ell + \nabla \mathcal{I}u_\ell\|_K^2 + \|\nabla(u_\ell - \mathcal{I}u_\ell)\|_K^2 \,\mathrm{d}t.$$

Guaranteed upper bound

Augmented Y norm

$$\|u-u_\ell\|_{\mathcal{E}_Y}^2 := \|u-\mathcal{I}u_\ell\|_Y^2 + \underbrace{\|u_\ell-\mathcal{I}u_\ell\|_X^2}_{\text{known, computable, measures time jumps}}$$

Theorem (Guaranteed and locally space–time efficient estimate)

There holds

$$\|u-u_\ell\|_{\mathcal{E}_Y}^2 \leq \sum_{n=1}^{N}\sum_{K\in\mathcal{T}_\ell^n}\int_{I_n}\|\boldsymbol{\sigma}_\ell+\nabla\mathcal{I}u_\ell\|_K^2+\|\nabla(u_\ell-\mathcal{I}u_\ell)\|_K^2\,\mathrm{d}t.$$

Moreover, for each time-step interval I_n and for each element $K\in\mathcal{T}_\ell^n$, there holds

$$\int_{I_n}\|\boldsymbol{\sigma}_\ell+\nabla\mathcal{I}u_\ell\|_K^2+\|\nabla(u_\ell-\mathcal{I}u_\ell)\|_K^2\mathrm{d}t \leq C_{\mathrm{eff}}^2\sum_{\mathbf{a}\in\mathcal{V}_K}|u-u_\ell|_{\mathcal{E}_Y^{\mathbf{a},n}}^2.$$

Guaranteed upper bound

Augmented Y norm

$$\|u - u_\ell\|^2_{\mathcal{E}_Y} := \|u - \mathcal{I}u_\ell\|^2_Y + \underbrace{\|u_\ell - \mathcal{I}u_\ell\|^2_X}_{\text{known, computable, measures time jumps}}$$

Theorem (Guaranteed and locally space–time efficient estimate)

There holds

$$\|u - u_\ell\|^2_{\mathcal{E}_Y} \leq \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_\ell^n} \int_{I_n} \|\boldsymbol{\sigma}_\ell + \nabla \mathcal{I}u_\ell\|^2_K + \|\nabla(u_\ell - \mathcal{I}u_\ell)\|^2_K \,\mathrm{d}t.$$

Moreover, for each time-step interval I_n and for each element $K \in \mathcal{T}_\ell^n$, there holds

$$\int_{I_n} \|\boldsymbol{\sigma}_\ell + \nabla \mathcal{I}u_\ell\|^2_K + \|\nabla(u_\ell - \mathcal{I}u_\ell)\|^2_K \mathrm{d}t \leq C^2_{\mathrm{eff}} \sum_{\mathbf{a} \in \mathcal{V}_K} |u - u_\ell|^2_{\mathcal{E}_Y^{\mathbf{a},n}}.$$

Comments

✓ C_{eff} only depends on mesh shape regularity $\kappa_{\mathcal{T}}$ and space dimension d ⇒ **robustness** w.r.t the final time T

Guaranteed upper bound

Augmented Y norm

$$\|u-u_\ell\|^2_{\mathcal{E}_Y} := \|u-\mathcal{I}u_\ell\|^2_Y + \underbrace{\|u_\ell-\mathcal{I}u_\ell\|^2_X}_{\text{known, computable, measures time jumps}}$$

Theorem (Guaranteed and locally space–time efficient estimate)

There holds

$$\|u-u_\ell\|^2_{\mathcal{E}_Y} \le \sum_{n=1}^{N} \sum_{K\in\mathcal{T}_\ell^n} \int_{I_n} \|\boldsymbol{\sigma}_\ell + \nabla\mathcal{I}u_\ell\|^2_K + \|\nabla(u_\ell-\mathcal{I}u_\ell)\|^2_K \,\mathrm{d}t.$$

Moreover, for each time-step interval I_n and for each element $K\in\mathcal{T}_\ell^n$, there holds

$$\int_{I_n} \|\boldsymbol{\sigma}_\ell + \nabla\mathcal{I}u_\ell\|^2_K + \|\nabla(u_\ell-\mathcal{I}u_\ell)\|^2_K \mathrm{d}t \le C^2_{\mathrm{eff}} \sum_{\mathbf{a}\in\mathcal{V}_K} |u-u_\ell|^2_{\mathcal{E}_Y^{\mathbf{a},n}}.$$

Comments

✓ C_{eff} only depends on mesh shape regularity $\kappa_{\mathcal{T}}$ and space dimension d $\Rightarrow$ **robustness** w.r.t the final time T and the polynomial degrees p and q

Guaranteed upper bound

Augmented Y norm

$$\|u - u_\ell\|^2_{\mathcal{E}_Y} := \|u - \mathcal{I}u_\ell\|^2_Y + \underbrace{\|u_\ell - \mathcal{I}u_\ell\|^2_X}_{\text{known, computable, measures time jumps}}$$

Theorem (Guaranteed and locally space–time efficient estimate)

There holds

$$\|u - u_\ell\|^2_{\mathcal{E}_Y} \leq \sum_{n=1}^{N} \sum_{K \in \mathcal{T}_\ell^n} \int_{I_n} \|\boldsymbol{\sigma}_\ell + \nabla \mathcal{I}u_\ell\|^2_K + \|\nabla(u_\ell - \mathcal{I}u_\ell)\|^2_K \,\mathrm{d}t.$$

Moreover, for each time-step interval I_n and for each element $K \in \mathcal{T}_\ell^n$, there holds

$$\int_{I_n} \|\boldsymbol{\sigma}_\ell + \nabla \mathcal{I}u_\ell\|^2_K + \|\nabla(u_\ell - \mathcal{I}u_\ell)\|^2_K \mathrm{d}t \leq C_{\mathrm{eff}}^2 \sum_{\boldsymbol{a} \in \mathcal{V}_K} |u - u_\ell|^2_{\mathcal{E}_Y^{\boldsymbol{a},n}}.$$

Comments

- ✓ C_{eff} only depends on mesh shape regularity $\kappa_{\mathcal{T}}$ and space dimension d $\Rightarrow$ **robustness** w.r.t the final time T and the polynomial degrees p and q
- ✓ **local** in **space** and in **time** efficiency

Outline

Modelling flow of water and air through soil

The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\mathbf{K}\kappa(S(u))(\nabla u + \mathbf{g})] &= f \quad \text{in } \Omega \times (0, T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Modelling flow of water and air through soil

The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\boldsymbol{K}\kappa(S(u))(\nabla u + \boldsymbol{g})] &= f \quad \text{in } \Omega \times (0, T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Setting

- u: pressure
- $s = S(u)$: saturation
- $\Omega \subset \mathbb{R}^d$, $1 \leq d \leq 3$, open bounded polytope with Lipschitz boundary $\partial\Omega$
- T: final time
- diffusion tensor $\boldsymbol{K}$, source term f, gravity $\boldsymbol{g}$, initial saturation s_0
- **nonlinear (degenerate) functions** S and κ

Modelling flow of water and air through soil

The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\boldsymbol{K}\kappa(S(u))(\nabla u + \boldsymbol{g})] &= f \quad \text{in } \Omega \times (0, T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Setting

- u: pressure
- $s = S(u)$: saturation
- $\Omega \subset \mathbb{R}^d$, $1 \le d \le 3$, open bounded polytope with Lipschitz boundary $\partial\Omega$
- T: final time
- diffusion tensor $\boldsymbol{K}$, source term f, gravity $\boldsymbol{g}$, initial saturation s_0
- **nonlinear (degenerate) functions** S and κ

Modelling flow of water and air through soil

The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\mathbf{K}\kappa(S(u))(\nabla u + \mathbf{g})] &= f \quad \text{in } \Omega \times (0, T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Setting

- u: pressure
- $s = S(u)$: saturation
- $\Omega \subset \mathbb{R}^d$, $1 \leq d \leq 3$, open bounded polytope with Lipschitz boundary $\partial\Omega$
- T: final time
- diffusion tensor $\mathbf{K}$, source term f, gravity $\mathbf{g}$, initial saturation s_0
- **nonlinear (degenerate) functions** S and κ

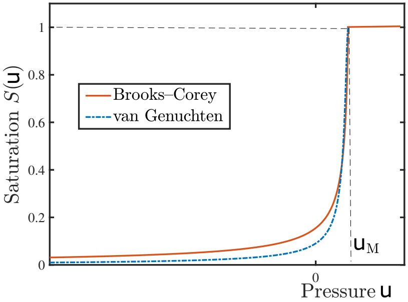
Modelling flow of water and air through soil

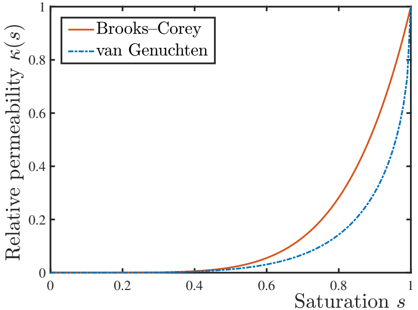
The Richards equation

Find $u : \Omega \times (0, T) \to \mathbb{R}$ such that

$$\begin{aligned} \partial_t S(u) - \nabla\cdot[\mathbf{K}\kappa(S(u))(\nabla u + \mathbf{g})] &= f \quad \text{in } \Omega \times (0,T), \\ u &= 0 \quad \text{on } \partial\Omega \times (0,T), \\ (S(u))(0) &= s_0 \quad \text{in } \Omega. \end{aligned}$$

Nonlinear (degenerate) functions S and κ





A posteriori error estimates

Nonsymmetry, nonlinearity, degeneracy

- Use all the tools from the above cases.
- Treatment of time-dependent nonlinearity: **sharp Gronwall lemma** not neglecting the integral terms.
- **Avoids** the appearance of the usual factor e^T, but gives rise to **integrated norms**.

A posteriori error estimates

Nonsymmetry, nonlinearity, degeneracy

- Use all the tools from the above cases.
- Treatment of time-dependent nonlinearity: **sharp Gronwall lemma** not neglecting the integral terms.
- **Avoids** the appearance of the usual factor e^T, but gives rise to **integrated norms**.

✓ **local** in **space** and in **time** efficiency

A posteriori error estimates

Nonsymmetry, nonlinearity, degeneracy

- Use all the tools from the above cases.
- Treatment of time-dependent nonlinearity: **sharp Gronwall lemma** not neglecting the integral terms.
- **Avoids** the appearance of the usual factor e^T, but gives rise to **integrated norms**.

✓ **local** in **space** and in **time** efficiency

✓ **robustness** w.r.t the final time T

A posteriori error estimates

Nonsymmetry, nonlinearity, degeneracy

- Use all the tools from the above cases.
- Treatment of time-dependent nonlinearity: **sharp Gronwall lemma** not neglecting the integral terms.
- **Avoids** the appearance of the usual factor e^T, but gives rise to **integrated norms**.

✓ **local** in **space** and in **time** efficiency

✓ **robustness** w.r.t the final time T

✗ **heuristic** estimators for the **treatment of degeneracy**

A posteriori error estimates

Nonsymmetry, nonlinearity, degeneracy

- Use all the tools from the above cases.
- Treatment of time-dependent nonlinearity: **sharp Gronwall lemma** not neglecting the integral terms.
- **Avoids** the appearance of the usual factor e^T, but gives rise to **integrated norms**.

✓ **local** in **space** and in **time** efficiency

✓ **robustness** w.r.t the final time T

✗ **heuristic** estimators for the **treatment of degeneracy**

✗ **norm change** between efficiency and reliability

A posteriori error estimates

Nonsymmetry, nonlinearity, degeneracy

- Use all the tools from the above cases.
- Treatment of time-dependent nonlinearity: **sharp Gronwall lemma** not neglecting the integral terms.
- **Avoids** the appearance of the usual factor e^T, but gives rise to **integrated norms**.

✓ **local** in **space** and in **time** efficiency

✓ **robustness** w.r.t the final time T

✗ **heuristic** estimators for the **treatment of degeneracy**

✗ **norm change** between efficiency and reliability

✗ **no robustness** wrt the **strength of nonlinearities**

A posteriori error estimates

Nonsymmetry, nonlinearity, degeneracy

- Use all the tools from the above cases.
- Treatment of time-dependent nonlinearity: **sharp Gronwall lemma** not neglecting the integral terms.
- **Avoids** the appearance of the usual factor e^T, but gives rise to **integrated norms**.

✓ **local** in **space** and in **time** efficiency

✓ **robustness** w.r.t the final time T

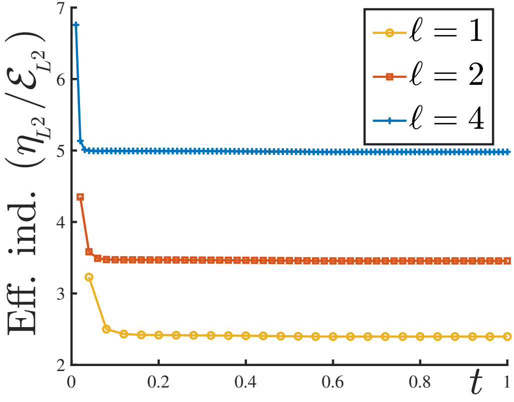
✗ **heuristic** estimators for the **treatment of degeneracy**

✗ **norm change** between efficiency and reliability

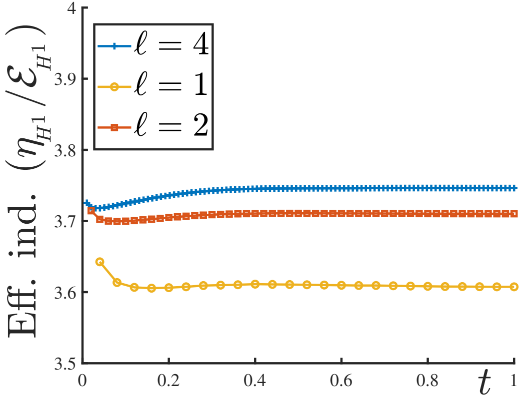
✗ **no robustness** wrt the **strength of nonlinearities**

- Details in K. Mitra, M. Vohralík, preprint (2022)

How **large** is the error? **Robustness** wrt the final time (known sol.)



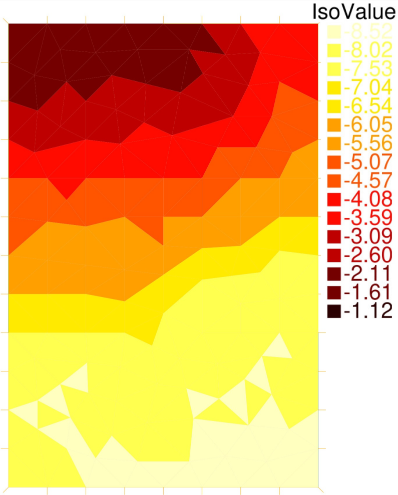
Integrated L^2 effectivity indices



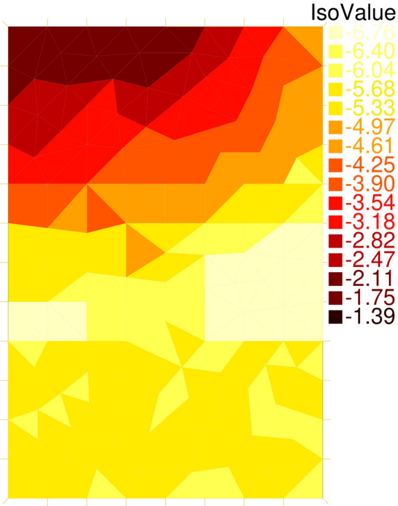
Integrated H^1 effectivity indices

K. Mitra, M. Vohralík, preprint (2022)

Where (in space and time) is the error **localized**? (benchmark case)



Estimated local error



Exact local error

K. Mitra, M. Vohralík, preprint (2022)

Realistic case

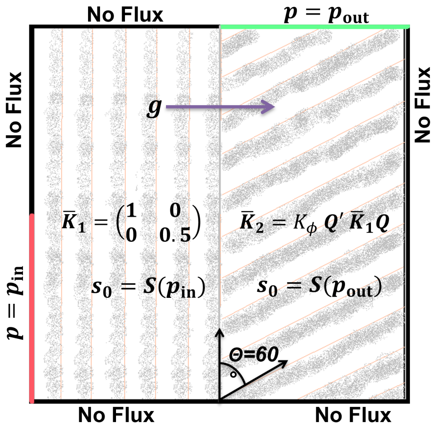
Setting

- unit square $\Omega = (0,1)^2$
- $T = 1$
- $f(\boldsymbol{x}, u) = 0$, heterogeneous and anisotropic $\boldsymbol{K}$, $\boldsymbol{g} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$
- **Brooks–Corey**-type **saturation** and **permeability** laws

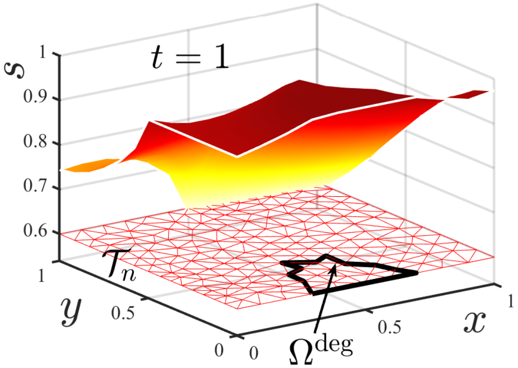
$$S(u) := \begin{cases} \frac{1}{(2-u)^{\frac{1}{3}}} & \text{if } u < 1, \\ 1 & \text{if } u \geq 1 \end{cases}, \quad \kappa(s) := s^3$$

- $(h, \tau) = (h_0, \tau_0)/\ell$ with $\ell \in \{1, 2, 4\}$, $h_0 = 0.2$, and $\tau_0 = 0.04$

Realistic case

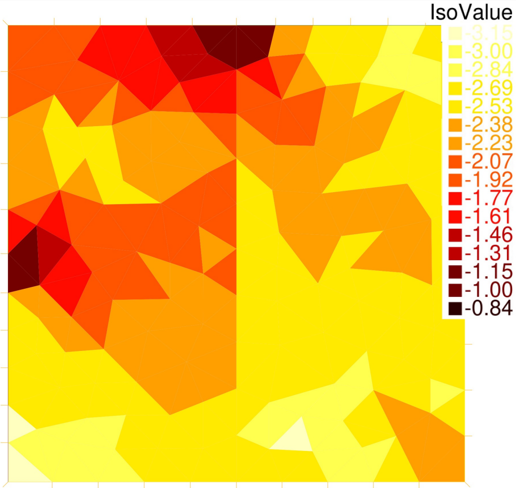


Setting

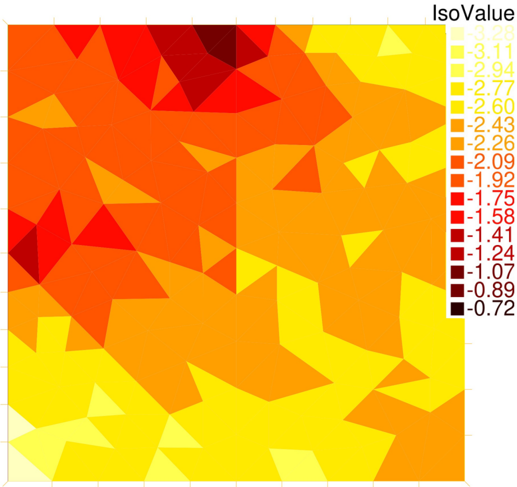


Numerical saturation for $\ell = 2$ at $t = 1$

Where (in space and time) is the error **localized**? (realistic test case)



Estimated local error



Exact local error

K. Mitra, M. Vohralík, preprint (2022)

Outline

Conclusions

Conclusions

- a posteriori **certification** of the **error** for nonlinear and unsteady problems
- **robustness** with respect to the **strength of nonlinearities** and **final time** for model cases (nonlinear or unsteady)
- **localization** of the error in **space** and in **time**
- **theory** and **sound numerical performance** for the Richards equation

Conclusions

Conclusions

- a posteriori **certification** of the **error** for nonlinear and unsteady problems
- **robustness** with respect to the **strength of nonlinearities** and **final time** for model cases (nonlinear or unsteady)
- **localization** of the error in **space** and in **time**
- **theory** and **sound numerical performance** for the Richards equation

HARNIST A., MITRA K., RAPPAPORT A., VOHRALÍK M. Robust energy a posteriori estimates for nonlinear elliptic problems. HAL Preprint 04033438, 2023.

MITRA K., VOHRALÍK M. Guaranteed, locally efficient, and robust a posteriori estimates for nonlinear elliptic problems in iteration-dependent norms. An orthogonal decomposition result based on iterative linearization. HAL Preprint 04156711, 2023.

ERN A., SMEARS, I., VOHRALÍK M. Guaranteed, locally space-time efficient, and polynomial-degree robust a posteriori error estimates for high-order discretizations of parabolic problems, *SIAM J. Numer. Anal.* **55** (2017), 2811–2834.

MITRA K., VOHRALÍK M. A posteriori error estimates for the Richards equation. HAL Preprint 03328944, 2022.

Conclusions

Conclusions

- a posteriori **certification** of the **error** for nonlinear and unsteady problems
- **robustness** with respect to the **strength of nonlinearities** and **final time** for model cases (nonlinear or unsteady)
- **localization** of the error in **space** and in **time**
- **theory** and **sound numerical performance** for the Richards equation

HARNIST A., MITRA K., RAPPAPORT A., VOHRALÍK M. Robust energy a posteriori estimates for nonlinear elliptic problems. HAL Preprint 04033438, 2023.

MITRA K., VOHRALÍK M. Guaranteed, locally efficient, and robust a posteriori estimates for nonlinear elliptic problems in iteration-dependent norms. An orthogonal decomposition result based on iterative linearization. HAL Preprint 04156711, 2023.

ERN A., SMEARS, I., VOHRALÍK M. Guaranteed, locally space-time efficient, and polynomial-degree robust a posteriori error estimates for high-order discretizations of parabolic problems, *SIAM J. Numer. Anal.* **55** (2017), 2811–2834.

MITRA K., VOHRALÍK M. A posteriori error estimates for the Richards equation. HAL Preprint 03328944, 2022.

Thank you for your attention!

Outline

Sobolev space and error

Sobolev space

$$H_0^1(\Omega)$$

Sobolev norm error

$$\boxed{\|\nabla(u_\ell - u)\|}$$

Residual and its dual norm

Definition (Residual)

$\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$; for $w \in H_0^1(\Omega)$, $\mathcal{R}(w) \in H^{-1}(\Omega)$ is given by

$$\langle \mathcal{R}(w), v \rangle := (a(|\nabla w|)\nabla w, \nabla v) - (f, v), \quad v \in H_0^1(\Omega).$$

Definition (Dual norm of the finite element residual)

$$|||\mathcal{R}(u_\ell) - \mathcal{R}(u)|||_{-1} = \boxed{|||\mathcal{R}(u_\ell)|||_{-1}} := \sup_{v \in H_0^1(\Omega)} \frac{\langle \mathcal{R}(u_\ell), v \rangle}{|||v|||}.$$

- $|||\mathcal{R}(u_\ell)|||_{-1} \geq 0$, $|||\mathcal{R}(u_\ell)|||_{-1} = 0$ if and only if $u_\ell = u$
- subordinate to the choice of the norm $|||\cdot|||$ on the Sobolev space $H_0^1(\Omega)$
- the most straightforward choice: $|||v||| := \|\nabla v\|$
- mathematically-based error measure

Residual and its dual norm

Definition (Residual)

$\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$; for $w \in H_0^1(\Omega)$, $\mathcal{R}(w) \in H^{-1}(\Omega)$ is given by

$$\langle \mathcal{R}(w), v \rangle := (a(|\nabla w|)\nabla w, \nabla v) - (f, v), \quad v \in H_0^1(\Omega).$$

Definition (Dual norm of the finite element residual)

$$|||\mathcal{R}(u_\ell) - \mathcal{R}(u)|||_{-1} = \boxed{|||\mathcal{R}(u_\ell)|||_{-1}} := \sup_{v \in H_0^1(\Omega)} \frac{\langle \mathcal{R}(u_\ell), v \rangle}{|||v|||}.$$

- $|||\mathcal{R}(u_\ell)|||_{-1} \geq 0$, $|||\mathcal{R}(u_\ell)|||_{-1} = 0$ if and only if $u_\ell = u$
- subordinate to the choice of the norm $||| \cdot |||$ on the Sobolev space $H_0^1(\Omega)$
- the most straightforward choice: $|||v||| := \|\nabla v\|$
- mathematically-based error measure

Outline

Fenchel conjugate, dual energy, flux equilibration

Definition (Fenchel conjugate)

$$\phi^*(\cdot, s) := \sup_{r \in [0,\infty)} (sr - \phi(\cdot, r)).$$

Definition (Dual energy)

$$\mathcal{J}^*(\boldsymbol{v}) := -\int_\Omega \phi^*(\cdot, |\boldsymbol{v}|), \quad \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega).$$

Definition (Flux equilibration)

$$\boldsymbol{\sigma}_\ell^{\boldsymbol{a},k} := \arg \min_{\substack{\boldsymbol{v}_\ell \in \mathcal{RT}_{p+1}(\mathcal{T}_{\boldsymbol{a}}) \cap \boldsymbol{H}_0(\mathrm{div},\omega_{\boldsymbol{a}}) \\ \nabla \cdot \boldsymbol{v}_\ell = \Pi_{\ell,p}(\psi^{\boldsymbol{a}} f - \nabla \psi^{\boldsymbol{a}} \cdot (\boldsymbol{A}_\ell^{k-1} \nabla u_\ell^k - \boldsymbol{b}_\ell^{k-1}))}} \|(\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}}(\psi^{\boldsymbol{a}} \Pi_{\ell,p-1}^{RTN}(\boldsymbol{A}_\ell^{k-1} \nabla u_\ell^k - \boldsymbol{b}_\ell^{k-1}) + \boldsymbol{v}_\ell)\|_{\omega_{\boldsymbol{a}}}^2.$$

Fenchel conjugate, dual energy, flux equilibration

Definition (Fenchel conjugate)

$$\phi^*(\cdot, s) := \sup_{r \in [0,\infty)} (sr - \phi(\cdot, r)).$$

Definition (Dual energy)

$$\mathcal{J}^*(\mathbf{v}) := -\int_\Omega \phi^*(\cdot, |\mathbf{v}|), \quad \mathbf{v} \in \mathbf{H}(\mathrm{div}, \Omega).$$

Definition (Flux equilibration)

$$\boldsymbol{\sigma}_\ell^{\mathbf{a},k} := \arg \min_{\substack{\mathbf{v}_\ell \in \mathcal{RT}_{p+1}(\mathcal{T}_{\mathbf{a}}) \cap \mathbf{H}_0(\mathrm{div}, \omega_{\mathbf{a}}) \\ \nabla \cdot \mathbf{v}_\ell = \Pi_{\ell,p}(\psi^{\mathbf{a}} f - \nabla \psi^{\mathbf{a}} \cdot (\mathbf{A}_\ell^{k-1} \nabla u_\ell^k - \mathbf{b}_\ell^{k-1}))}} \|(\mathbf{A}_\ell^{k-1})^{-\frac{1}{2}} (\psi^{\mathbf{a}} \Pi_{\ell,p-1}^{\mathbf{RTN}} (\mathbf{A}_\ell^{k-1} \nabla u_\ell^k - \mathbf{b}_\ell^{k-1}) + \mathbf{v}_\ell)\|_{\omega_{\mathbf{a}}}^2.$$

Fenchel conjugate, dual energy, flux equilibration

Definition (Fenchel conjugate)

$$\phi^*(\cdot, s) := \sup_{r \in [0,\infty)} (sr - \phi(\cdot, r)).$$

Definition (Dual energy)

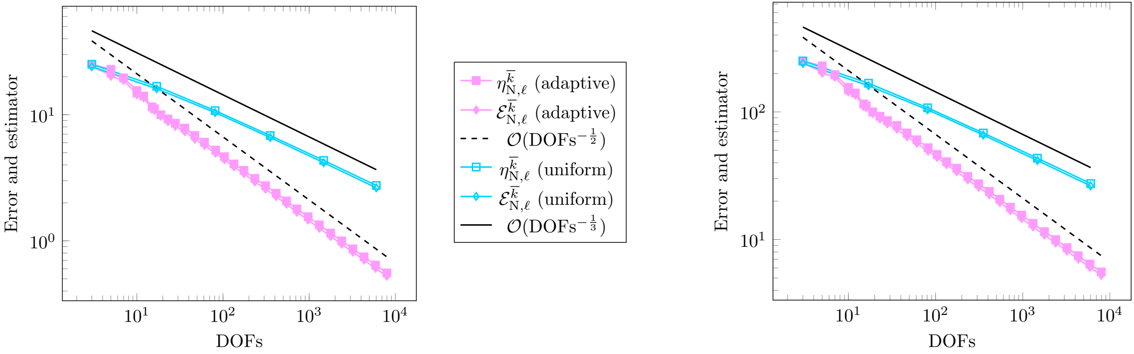
$$\mathcal{J}^*(\boldsymbol{v}) := -\int_\Omega \phi^*(\cdot, |\boldsymbol{v}|), \qquad \boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}, \Omega).$$

Definition (Flux equilibration)

$$\boldsymbol{\sigma}_\ell^{\boldsymbol{a},k} := \arg \min_{\substack{\boldsymbol{v}_\ell \in \mathcal{RT}_{p+1}(\mathcal{T}_{\boldsymbol{a}}) \cap \boldsymbol{H}_0(\mathrm{div}, \omega_{\boldsymbol{a}}) \\ \nabla \cdot \boldsymbol{v}_\ell = \Pi_{\ell,p}(\psi^{\boldsymbol{a}} f - \nabla \psi^{\boldsymbol{a}} \cdot (\boldsymbol{A}_\ell^{k-1} \nabla u_\ell^k - \boldsymbol{b}_\ell^{k-1}))}} \|(\boldsymbol{A}_\ell^{k-1})^{-\frac{1}{2}} (\psi^{\boldsymbol{a}} \boldsymbol{\Pi}_{\ell,p-1}^{\boldsymbol{RTN}} (\boldsymbol{A}_\ell^{k-1} \nabla u_\ell^k - \boldsymbol{b}_\ell^{k-1}) + \boldsymbol{v}_\ell)\|_{\omega_{\boldsymbol{a}}}^2.$$

Outline

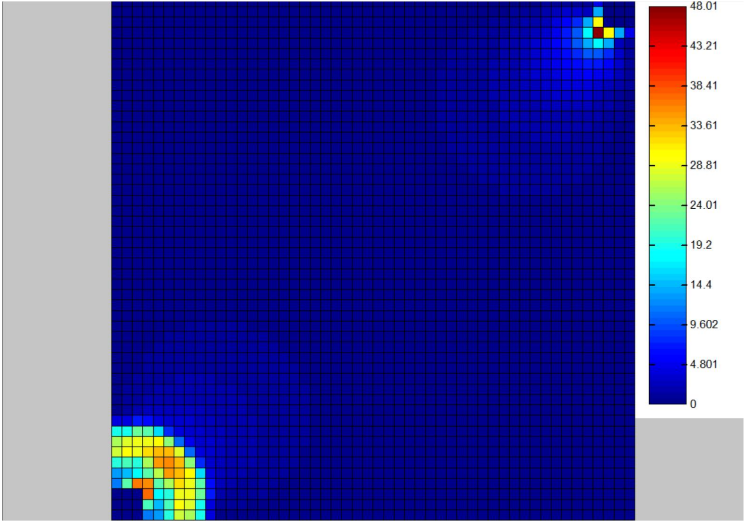
Decreasing the error efficiently: optimal decay rate wrt DoFs



$$\frac{a_{\mathrm{c}}}{a_{\mathrm{m}}} = 10^3 \qquad\qquad \frac{a_{\mathrm{c}}}{a_{\mathrm{m}}} = 10^6$$

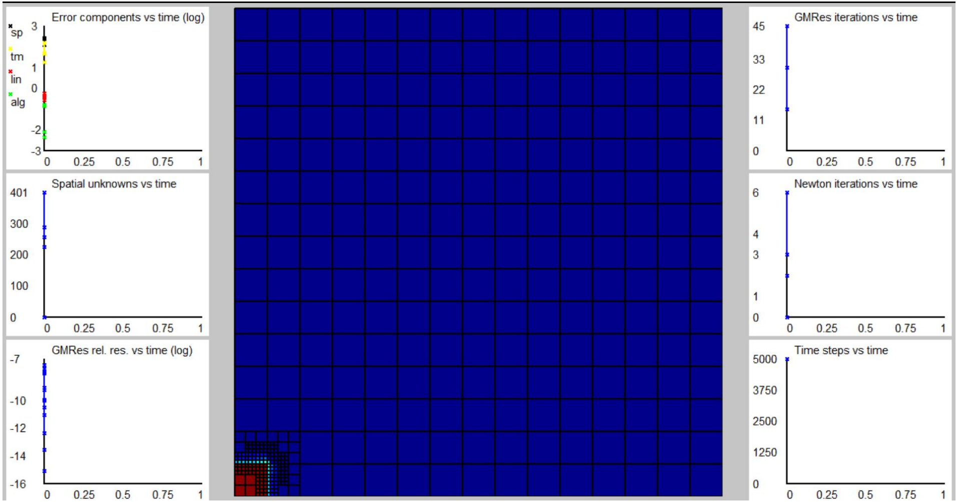
Outline

Where (in space and time) is the error **localized**? (two-phase flow)



M. Vohralík, M. Wheeler, Computational Geosciences (2013)

All error components (two-phase flow)



M. Vohralík, M. Wheeler, Computational Geosciences (2013)