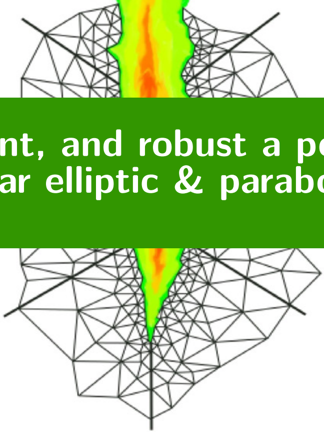
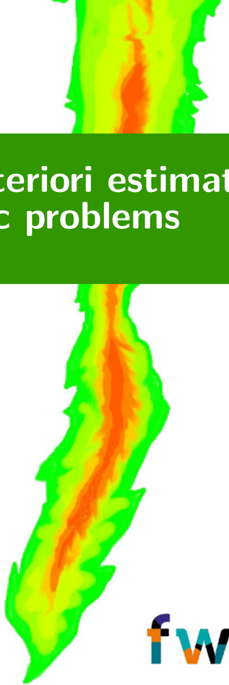




Reliable, efficient, and robust a posteriori estimates for nonlinear elliptic & parabolic problems

Koondanibha (Koondi) Mitra

joint work with Martin Vohralík

CEA-EdF-Inria summer school

June 2022

UHASSELT

fwo

0 Outline

1 Outline

1 Introduction

Nonlinear elliptic problems

For $d \in \mathbb{N}$, let $\Omega \subset \mathbb{R}^d$ be an open and bounded polytope. Let $u \in H_0^1(\Omega)$ solve the **nonlinear** elliptic operator equation: for $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$,

$$\langle \mathcal{R}(u), \varphi \rangle = 0 \quad \forall \varphi \in H_0^1(\Omega).$$

1 Introduction

Nonlinear elliptic problems

For $d \in \mathbb{N}$, let $\Omega \subset \mathbb{R}^d$ be an open and bounded polytope. Let $u \in H_0^1(\Omega)$ solve the **nonlinear** elliptic operator equation: for $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$,

$$\langle \mathcal{R}(u), \varphi \rangle = 0 \quad \forall \varphi \in H_0^1(\Omega).$$

Assumption 1 $\mathcal{R}$ is monotone & Lipschitz*

For a numerical approximation $u_\ell \in H_0^1(\Omega)$, and constants $\lambda_\mathrm{M} > \lambda_\mathrm{m} > 0$,

$$\lambda_\mathrm{m} \operatorname{dist}(u_\ell, u) \leq \sup_{\varphi \in H_0^1(\Omega)} \frac{\langle \mathcal{R}(u_\ell) - \mathcal{R}(u), \varphi \rangle}{\|\nabla \varphi\|} \leq \lambda_\mathrm{M} \operatorname{dist}(u_\ell, u).$$

1 Introduction

Nonlinear elliptic problems

For $d \in \mathbb{N}$, let $\Omega \subset \mathbb{R}^d$ be an open and bounded polytope. Let $u \in H_0^1(\Omega)$ solve the **nonlinear** elliptic operator equation: for $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$,

$$\langle \mathcal{R}(u), \varphi \rangle = 0 \quad \forall \varphi \in H_0^1(\Omega).$$

Assumption 1 $\mathcal{R}$ is monotone & Lipschitz*

For a numerical approximation $u_\ell \in H_0^1(\Omega)$, and constants $\lambda_{\mathrm{M}} > \lambda_{\mathrm{m}} > 0$,

$$\lambda_{\mathrm{m}} \operatorname{dist}(u_\ell, u) \leq \sup_{\varphi \in H_0^1(\Omega)} \frac{\langle \mathcal{R}(u_\ell), \varphi \rangle}{\|\nabla \varphi\|} \leq \lambda_{\mathrm{M}} \operatorname{dist}(u_\ell, u).$$

Nonlinear elliptic problems

For $d \in \mathbb{N}$, let $\Omega \subset \mathbb{R}^d$ be an open and bounded polytope. Let $u \in H_0^1(\Omega)$ solve the **nonlinear** elliptic operator equation: for $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$,

$$\langle \mathcal{R}(u), \varphi \rangle = 0 \quad \forall \varphi \in H_0^1(\Omega).$$

Assumption 1 $\mathcal{R}$ is monotone & Lipschitz*

For a numerical approximation $u_\ell \in H_0^1(\Omega)$, and constants $\lambda_{\mathrm{M}} > \lambda_{\mathrm{m}} > 0$,

$$\lambda_{\mathrm{m}} \operatorname{dist}(u_\ell, u) \leq \underbrace{\sup_{\varphi \in H_0^1(\Omega)} \frac{\langle \mathcal{R}(u_\ell), \varphi \rangle}{\|\nabla \varphi\|}}_{\|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)}} \leq \lambda_{\mathrm{M}} \operatorname{dist}(u_\ell, u).$$

Nonlinear elliptic problems

For $d \in \mathbb{N}$, let $\Omega \subset \mathbb{R}^d$ be an open and bounded polytope. Let $u \in H_0^1(\Omega)$ solve the **nonlinear** elliptic operator equation: for $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$,

$$\langle \mathcal{R}(u), \varphi \rangle = 0 \quad \forall \varphi \in H_0^1(\Omega).$$

Assumption 1 $\mathcal{R}$ is monotone & Lipschitz*

For a numerical approximation $u_\ell \in H_0^1(\Omega)$, and constants $\lambda_{\mathrm{M}} > \lambda_{\mathrm{m}} > 0$,

$$\lambda_{\mathrm{m}} \,\mathrm{dist}(u_\ell, u) \leq \|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)} \leq \lambda_{\mathrm{M}} \,\mathrm{dist}(u_\ell, u).$$

Nonlinear elliptic problems

For $d \in \mathbb{N}$, let $\Omega \subset \mathbb{R}^d$ be an open and bounded polytope. Let $u \in H_0^1(\Omega)$ solve the **nonlinear** elliptic operator equation: for $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$,

$$\langle \mathcal{R}(u), \varphi \rangle = 0 \quad \forall \varphi \in H_0^1(\Omega).$$

Assumption 1 $\mathcal{R}$ is monotone & Lipschitz*

For a numerical approximation $u_\ell \in H_0^1(\Omega)$, and constants $\lambda_\mathrm{M} > \lambda_\mathrm{m} > 0$,

$$\lambda_\mathrm{m} \operatorname{dist}(u_\ell, u) \leq \|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)} \leq \lambda_\mathrm{M} \operatorname{dist}(u_\ell, u).$$

The strength of the nonlinearity is measued by $\lambda_\mathrm{M}/\lambda_\mathrm{m}$

Nonlinear elliptic problems

For $d \in \mathbb{N}$, let $\Omega \subset \mathbb{R}^d$ be an open and bounded polytope. Let $u \in H_0^1(\Omega)$ solve the **nonlinear** elliptic operator equation: for $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$,

$$\langle \mathcal{R}(u), \varphi \rangle = 0 \quad \forall\, \varphi \in H_0^1(\Omega).$$

Assumption 1 $\mathcal{R}$ is monotone & Lipschitz*

For a numerical approximation $u_\ell \in H_0^1(\Omega)$, and constants $\lambda_{\mathrm{M}} > \lambda_{\mathrm{m}} > 0$,

$$\lambda_{\mathrm{m}} \operatorname{dist}(u_\ell, u) \leq \|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)} \leq \lambda_{\mathrm{M}} \operatorname{dist}(u_\ell, u).$$

The strength of the nonlinearity is measued by $\lambda_{\mathrm{M}}/\lambda_{\mathrm{m}}$

Then the estimate [Chaillou & Suri (2006), Kim (2007), Houston *et al* (2008), Garau *et al* (2011),...],

$$\lambda_{\mathrm{m}} \operatorname{dist}(u_\ell, u) \leq \eta(u_\ell) \leq C\lambda_{\mathrm{M}} \operatorname{dist}(u_\ell, u)$$

is not robust with respect to $\lambda_{\mathrm{M}}/\lambda_{\mathrm{m}}$

1 Dual norm of the residual estimate

Reliable, and **locally efficient** a posteriori error estimates **robust** with respect to the strength of the **nonlinearity** $\lambda_{\mathrm{M}}/\lambda_{\mathrm{m}}$

$$\|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)} \leq \eta(u_\ell) \leq C\|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)}$$

[Chaillou & Suri (2006), El Alaoui *et al* (2011), Ern & Vohralík (2013), Blechta *et al* (2018)]

1 Dual norm of the residual estimate

Reliable, and **locally efficient** a posteriori error estimates **robust** with respect to the strength of the **nonlinearity** $\lambda_{\mathrm{M}}/\lambda_{\mathrm{m}}$

$$\|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)} \leq \eta(u_\ell) \leq C\|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)}$$

[Chaillou & Suri (2006), El Alaoui *et al* (2011), Ern & Vohralík (2013), Blechta *et al* (2018)]

- The dual norm of the residual might be too weak an error measure

1 A linear example

Consider the diffusion eq: $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}\nabla u, \nabla \varphi) = 0.$

Let $\lambda_{\mathrm{m}} \, |\boldsymbol{y}|^2 \leq \boldsymbol{y}^{\mathrm{T}} \mathcal{D} \boldsymbol{y} \leq \lambda_{\mathrm{M}} \, |\boldsymbol{y}|^2$, for all $\boldsymbol{y} \in \mathbb{R}^d$.

1 A linear example

Consider the diffusion eq: $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}\nabla u, \nabla \varphi) = 0.$

Let $\lambda_{\rm m} |\boldsymbol{y}|^2 \leq \boldsymbol{y}^{\rm T}\mathcal{D}\boldsymbol{y} \leq \lambda_{\rm M} |\boldsymbol{y}|^2$, for all $\boldsymbol{y} \in \mathbb{R}^d$. If $u_\ell \in V_\ell \subset H_0^1(\Omega)$ is the FE solution of the problem then Cea's lemma gives

$$\|\nabla(u - u_\ell)\| \leq \tfrac{\lambda_{\rm M}}{\lambda_{\rm m}} \|\nabla(u - \varphi_\ell)\| \quad \forall \varphi_\ell \in V_\ell.$$

1 A linear example

Consider the diffusion eq: $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}\nabla u, \nabla \varphi) = 0.$

Let $\lambda_{\mathrm{m}} |\boldsymbol{y}|^2 \leq \boldsymbol{y}^{\mathrm{T}} \mathcal{D} \boldsymbol{y} \leq \lambda_{\mathrm{M}} |\boldsymbol{y}|^2$, for all $\boldsymbol{y} \in \mathbb{R}^d$. If $u_\ell \in V_\ell \subset H_0^1(\Omega)$ is the FE solution of the problem then Cea's lemma gives

$$\|\nabla(u - u_\ell)\| \leq \tfrac{\lambda_{\mathrm{M}}}{\lambda_{\mathrm{m}}} \|\nabla(u - \varphi_\ell)\| \quad \forall \varphi_\ell \in V_\ell.$$

In this case $\|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)}$ can be estimated robustly, but might be too weak an error measure.

$$\lambda_{\mathrm{m}} \|\nabla(u - u_\ell)\| \leq \|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)} \leq \lambda_{\mathrm{M}} \|\nabla(u - u_\ell)\|.$$

1 A linear example

Consider the diffusion eq: $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}\nabla u, \nabla \varphi) = 0.$

Let $\lambda_{\mathrm{m}} |\boldsymbol{y}|^2 \le \boldsymbol{y}^{\mathrm{T}} \mathcal{D} \boldsymbol{y} \le \lambda_{\mathrm{M}} |\boldsymbol{y}|^2$, for all $\boldsymbol{y} \in \mathbb{R}^d$. If $u_\ell \in V_\ell \subset H_0^1(\Omega)$ is the FE solution of the problem then Cea's lemma gives

$$\|\nabla(u - u_\ell)\| \le \tfrac{\lambda_{\mathrm{M}}}{\lambda_{\mathrm{m}}} \|\nabla(u - \varphi_\ell)\| \quad \forall\, \varphi_\ell \in V_\ell.$$

In this case $\|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)}$ can be estimated robustly, but might be too weak an error measure.
However, defining the energy norm $|\!|\!|\varphi|\!|\!|_{1,\mathcal{D}} = \|\mathcal{D}^{\frac{1}{2}} \nabla \varphi\|$ one has

$$|\!|\!|u - u_\ell|\!|\!|_{1,\mathcal{D}} \le |\!|\!|u - \varphi_\ell|\!|\!|_{1,\mathcal{D}}, \quad \forall\, \varphi_\ell \in V_\ell.$$

1 A linear example

Consider the diffusion eq: $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}\nabla u, \nabla \varphi) = 0.$

Let $\lambda_{\mathrm{m}} |\boldsymbol{y}|^2 \leq \boldsymbol{y}^{\mathrm{T}} \mathcal{D} \boldsymbol{y} \leq \lambda_{\mathrm{M}} |\boldsymbol{y}|^2$, for all $\boldsymbol{y} \in \mathbb{R}^d$. If $u_\ell \in V_\ell \subset H_0^1(\Omega)$ is the FE solution of the problem then Cea's lemma gives

$$\|\nabla(u - u_\ell)\| \leq \tfrac{\lambda_{\mathrm{M}}}{\lambda_{\mathrm{m}}} \|\nabla(u - \varphi_\ell)\| \quad \forall \varphi_\ell \in V_\ell.$$

In this case $\|\mathcal{R}(u_\ell)\|_{H^{-1}(\Omega)}$ can be estimated robustly, but might be too weak an error measure.
However, defining the energy norm $|\!|\!|\varphi|\!|\!|_{1,\mathcal{D}} = \|\mathcal{D}^{\frac{1}{2}} \nabla \varphi\|$ one has

$$|\!|\!|u - u_\ell|\!|\!|_{1,\mathcal{D}} \leq |\!|\!|u - \varphi_\ell|\!|\!|_{1,\mathcal{D}}, \quad \forall \varphi_\ell \in V_\ell.$$

This motivates rather the error measure

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

which also results in robust estimates

1 Moving to the nonlinear case

Example (nonlinear diffusion): $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u)\nabla u, \nabla \varphi) = 0.$

1 Moving to the nonlinear case

Example (nonlinear diffusion): $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u)\nabla u, \nabla \varphi) = 0.$

Then $|||\mathcal{R}(\cdot)|||_{-1,\mathcal{D}(u)}$ cannot be defined since $u \in H_0^1(\Omega)$ is unknown.

1 Moving to the nonlinear case

Example (nonlinear diffusion): $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u)\nabla u, \nabla \varphi) = 0.$

Linearization iterations

We generally solve nonlinear equations by linearization iterations, i.e., by finding a sequence $\{u_\ell^i\}_{i\in\mathbb{N}} \subset V_\ell \subset H_0^1(\Omega)$.

Example (nonlinear diffusion): $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u)\nabla u, \nabla \varphi) = 0.$

Linearization iterations

We generally solve nonlinear equations by linearization iterations, i.e., by finding a sequence $\{u_\ell^i\}_{i\in\mathbb{N}} \subset V_\ell \subset H_0^1(\Omega)$.

Example (Fixed point iteration) For each $i \in \mathbb{N}$ and $u_\ell^i \in V_\ell$, let $u_\ell^{i+1} \in V_\ell$ solve $(\mathcal{D}(u_\ell^i)\nabla u_\ell^{i+1}, \nabla \varphi_\ell) = (f, \varphi_\ell)$ for all $\varphi_\ell \in V_\ell$.

Example (nonlinear diffusion): $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u)\nabla u, \nabla\varphi) = 0$.

Linearization iterations

We generally solve nonlinear equations by linearization iterations, i.e., by finding a sequence $\{u_\ell^i\}_{i\in\mathbb{N}} \subset V_\ell \subset H_0^1(\Omega)$.

Example (Fixed point iteration) For each $i \in \mathbb{N}$ and $u_\ell^i \in V_\ell$, let $u_\ell^{i+1} \in V_\ell$ solve $(\mathcal{D}(u_\ell^i)\nabla u_\ell^{i+1}, \nabla\varphi_\ell) = (f, \varphi_\ell)$ for all $\varphi_\ell \in V_\ell$.
This is the FE approximation of $u_{\langle\ell\rangle}^{i+1} \in H_0^1(\Omega)$ solving the linear problem

$$\langle \mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_{\langle\ell\rangle}^{i+1}), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u_\ell^i)\nabla u_{\langle\ell\rangle}^{i+1}, \nabla\varphi) = 0 \qquad \forall \varphi \in H_0^1(\Omega).$$

Example (nonlinear diffusion): $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u)\nabla u, \nabla \varphi) = 0.$

Linearization iterations

We generally solve nonlinear equations by linearization iterations, i.e., by finding a sequence $\{u_\ell^i\}_{i\in\mathbb{N}} \subset V_\ell \subset H_0^1(\Omega)$.

Example (Fixed point iteration) For each $i \in \mathbb{N}$ and $u_\ell^i \in V_\ell$, let $u_\ell^{i+1} \in V_\ell$ solve $(\mathcal{D}(u_\ell^i)\nabla u_\ell^{i+1}, \nabla \varphi_\ell) = (f, \varphi_\ell)$ for all $\varphi_\ell \in V_\ell$.
This is the FE approximation of $u_{\langle\ell\rangle}^{i+1} \in H_0^1(\Omega)$ solving the linear problem

$$\langle \mathcal{R}_{\rm lin}^{u_\ell^i}(u_{\langle\ell\rangle}^{i+1}), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u_\ell^i)\nabla u_{\langle\ell\rangle}^{i+1}, \nabla\varphi) = 0 \qquad \forall \varphi \in H_0^1(\Omega).$$

Then defining the **iteration-dependent energy norm**

$$\begin{cases} |||\varphi|||_{1,u_\ell^i} := \|\mathcal{D}(u_\ell^i)^{\frac{1}{2}}\nabla\varphi\| & \text{for } \varphi \in H_0^1(\Omega), \\ |||\varsigma|||_{-1,u_\ell^i} = \sup_{\varphi\in H_0^1(\Omega)} \langle \varsigma, \varphi\rangle / |||\varphi|||_{1,u_\ell^i} & \text{for } \varsigma \in H^{-1}(\Omega), \end{cases}$$

1 Moving to the nonlinear case

Example (nonlinear diffusion): $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u)\nabla u, \nabla \varphi) = 0.$

Linearization iterations

We generally solve nonlinear equations by linearization iterations, i.e., by finding a sequence $\{u_\ell^i\}_{i\in\mathbb{N}} \subset V_\ell \subset H_0^1(\Omega)$.

Example (Fixed point iteration) For each $i \in \mathbb{N}$ and $u_\ell^i \in V_\ell$, let $u_\ell^{i+1} \in V_\ell$ solve $(\mathcal{D}(u_\ell^i)\nabla u_\ell^{i+1}, \nabla \varphi_\ell) = (f, \varphi_\ell)$ for all $\varphi_\ell \in V_\ell$.
This is the FE approximation of $u_{\langle \ell \rangle}^{i+1} \in H_0^1(\Omega)$ solving the linear problem

$$\langle \mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_{\langle\ell\rangle}^{i+1}), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u_\ell^i)\nabla u_{\langle\ell\rangle}^{i+1}, \nabla\varphi) = 0 \qquad \forall \varphi \in H_0^1(\Omega).$$

Then defining the **iteration-dependent energy norm**

$$\begin{cases} |||\varphi|||_{1,u_\ell^i} := \|\mathcal{D}(u_\ell^i)^{\frac{1}{2}}\nabla\varphi\| & \text{for } \varphi \in H_0^1(\Omega), \\ |||\varsigma|||_{-1,u_\ell^i} = \sup_{\varphi\in H_0^1(\Omega)} \langle \varsigma, \varphi\rangle / |||\varphi|||_{1,u_\ell^i} & \text{for } \varsigma \in H^{-1}(\Omega), \end{cases}$$

we have (under conditions) robust estimates of

$$\left|\left|\left| \mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_\ell^{i+1}) \right|\right|\right|_{-1,u_\ell^i} = \left|\left|\left| u_{\langle\ell\rangle}^{i+1} - u_\ell^{i+1} \right|\right|\right|_{1,u_\ell^i}$$

1 Moving to the nonlinear case

Example (nonlinear diffusion): $\langle \mathcal{R}(u), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u)\nabla u, \nabla \varphi) = 0.$

Linearization iterations

We generally solve nonlinear equations by linearization iterations, i.e., by finding a sequence $\{u_\ell^i\}_{i\in\mathbb{N}} \subset V_\ell \subset H_0^1(\Omega)$.

Example (Fixed point iteration) For each $i \in \mathbb{N}$ and $u_\ell^i \in V_\ell$, let $u_\ell^{i+1} \in V_\ell$ solve $(\mathcal{D}(u_\ell^i)\nabla u_\ell^{i+1}, \nabla \varphi_\ell) = (f, \varphi_\ell)$ for all $\varphi_\ell \in V_\ell$.
This is the FE approximation of $u_{\langle \ell \rangle}^{i+1} \in H_0^1(\Omega)$ solving the linear problem

$$\langle \mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_{\langle \ell \rangle}^{i+1}), \varphi \rangle := (f, \varphi) - (\mathcal{D}(u_\ell^i)\nabla u_{\langle \ell \rangle}^{i+1}, \nabla \varphi) = 0 \qquad \forall \varphi \in H_0^1(\Omega).$$

Noting that

$$\langle \mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_\ell^{i+1}), \varphi \rangle := -(\mathcal{D}(u_\ell^i)\nabla(u_\ell^{i+1} - u_\ell^i), \nabla \varphi) + \langle \mathcal{R}(u_\ell^i), \varphi \rangle$$

can we provide a robust estimate for $|||\mathcal{R}(u_\ell^i)|||_{-1,u_\ell^i}$?

2 Outline

2 An orthogonal decomposition result

Theorem 1 Decomposition of the total error

Under Assumption 1, provided that the linearization iterations $\{u_\ell^i\}_{i\in\mathbb{N}} \subset V_\ell$ are generated by FE approximations of $u_{\langle\ell\rangle}^i \in H_0^1(\Omega)$ solving

$$\langle \mathcal{R}_{\rm lin}^{u_\ell^i}(u_{\langle\ell\rangle}^{i+1}), \varphi\rangle := -\mathfrak{L}(u_\ell^i; u_{\langle\ell\rangle}^{i+1} - u_\ell^i, \varphi) + \langle \mathcal{R}(u_\ell^i), \varphi\rangle = 0 \quad \forall\, \varphi \in H_0^1(\Omega)$$

and $i \geq 0$, for a symmetric, bounded, coercive, bilinear form $\mathfrak{L}(u_\ell^i, \cdot, \cdot)$,

2 An orthogonal decomposition result

Theorem 1 Decomposition of the total error

Under Assumption 1, provided that the linearization iterations $\{u_\ell^i\}_{i\in\mathbb{N}} \subset V_\ell$ are generated by FE approximations of $u^i_{\langle\ell\rangle} \in H_0^1(\Omega)$ solving

$$\langle \mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u^{i+1}_{\langle\ell\rangle}), \varphi\rangle := -\mathfrak{L}(u_\ell^i; u^{i+1}_{\langle\ell\rangle} - u_\ell^i, \varphi) + \langle \mathcal{R}(u_\ell^i), \varphi\rangle = 0 \quad \forall\, \varphi \in H_0^1(\Omega)$$

and $i \geq 0$, for a symmetric, bounded, coercive, bilinear form $\mathfrak{L}(u_\ell^i, \cdot, \cdot)$,

Remark We would consider $\mathfrak{L} : H_0^1(\Omega) \times H_0^1(\Omega) \mapsto \mathbb{R}$ corresponding to **linear reaction-diffusion** problems, i.e,

$$\mathfrak{L}(u_\ell^i; v, w) := (\underbrace{L(\boldsymbol{x}, u_\ell^i)}_{\text{known reaction coeff.}}\, v, w) + (\underbrace{\mathfrak{a}(\boldsymbol{x}, u_\ell^i)}_{\text{known diffusion coeff.}} \nabla v, \nabla w).$$

2 An orthogonal decomposition result

Theorem 1 Decomposition of the total error

Under Assumption 1, provided that the linearization iterations $\{u_\ell^i\}_{i\in\mathbb{N}} \subset V_\ell$ are generated by FE approximations of $u_{\langle\ell\rangle}^i \in H_0^1(\Omega)$ solving

$$\langle \mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_{\langle\ell\rangle}^{i+1}), \varphi\rangle := -\mathfrak{L}(u_\ell^i; u_{\langle\ell\rangle}^{i+1} - u_\ell^i, \varphi) + \langle \mathcal{R}(u_\ell^i), \varphi\rangle = 0 \quad \forall\, \varphi \in H_0^1(\Omega)$$

and $i \geq 0$, for a symmetric, bounded, coercive, bilinear form $\mathfrak{L}(u_\ell^i, \cdot, \cdot)$, and

$$|||\varphi|||_{1,u_\ell^i} := \mathfrak{L}(u_\ell^i; \varphi, \varphi)^{\frac{1}{2}}, \quad |||\varsigma|||_{-1,u_\ell^i} := \sup_{\varphi\in H_0^1(\Omega)} \frac{\langle \varsigma, \varphi\rangle}{|||\varphi|||_{1,u_\ell^i}},$$

2 An orthogonal decomposition result

Theorem 1 Decomposition of the total error

Under Assumption 1, provided that the linearization iterations $\{u_\ell^i\}_{i\geq 0} \subset V_\ell$ are generated by FE approximations of $u_{\langle\ell\rangle}^i \in H_0^1(\Omega)$ solving

$$\langle \mathcal{R}_{\rm lin}^{u_\ell^i}(u_{\langle\ell\rangle}^{i+1}), \varphi\rangle := -\mathfrak{L}(u_\ell^i; u_{\langle\ell\rangle}^{i+1} - u_\ell^i, \varphi) + \langle \mathcal{R}(u_\ell^i), \varphi\rangle = 0 \quad \forall\, \varphi \in H_0^1(\Omega)$$

and $i \geq 0$, for a symmetric, bounded, coercive, bilinear form $\mathfrak{L}(u_\ell^i, \cdot, \cdot)$, and

$$|||\varphi|||_{1,u_\ell^i} := \mathfrak{L}(u_\ell^i; \varphi, \varphi)^{\frac{1}{2}}, \quad |||\varsigma|||_{-1,u_\ell^i} := \sup_{\varphi\in H_0^1(\Omega)} \frac{\langle \varsigma, \varphi\rangle}{|||\varphi|||_{1,u_\ell^i}},$$

we have

$$\underbrace{|||\mathcal{R}(u_\ell^i)|||_{-1,u_\ell^i}^2}_{\text{total error}} = \underbrace{\left|\left|\left|\mathcal{R}_{\rm lin}^{u_\ell^i}(u_\ell^{i+1})\right|\right|\right|_{-1,u_\ell^i}^2}_{\text{discretization error of the linerization step}} + \underbrace{|||u_\ell^{i+1} - u_\ell^i|||_{1,u_\ell^i}^2}_{\text{linearization error}}.$$

Theorem 1 Decomposition of the total error

Under Assumption 1, provided that the linearization iterations $\{u_\ell^i\}_{i\geq 0} \subset V_\ell$ are generated by FE approximations of $u_{\langle\ell\rangle}^i \in H_0^1(\Omega)$ solving

$$\langle \mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_{\langle\ell\rangle}^{i+1}), \varphi\rangle := -\mathfrak{L}(u_\ell^i; u_{\langle\ell\rangle}^{i+1} - u_\ell^i, \varphi) + \langle \mathcal{R}(u_\ell^i), \varphi\rangle = 0 \quad \forall\, \varphi \in H_0^1(\Omega)$$

and $i \geq 0$, for a symmetric, bounded, coercive, bilinear form $\mathfrak{L}(u_\ell^i, \cdot, \cdot)$, and

$$|\!|\!|\varphi|\!|\!|_{1,u_\ell^i} := \mathfrak{L}(u_\ell^i; \varphi, \varphi)^{\frac{1}{2}}, \quad |\!|\!|\varsigma|\!|\!|_{-1,u_\ell^i} := \sup_{\varphi \in H_0^1(\Omega)} \frac{\langle \varsigma, \varphi\rangle}{|\!|\!|\varphi|\!|\!|_{1,u_\ell^i}},$$

we have

$$\underbrace{\underbrace{|\!|\!|\mathcal{R}(u_\ell^i)|\!|\!|_{-1,u_\ell^i}^2}_{\text{total error}}}_{|\!|\!|u_\ell^i - u_{\langle\ell\rangle}^{i+1}|\!|\!|_{1,u_\ell^i}^2} = \underbrace{\underbrace{\left|\!\left|\!\left|\mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_\ell^{i+1})\right|\!\right|\!\right|_{-1,u_\ell^i}^2}_{\text{discretization error of the linerization step}}}_{|\!|\!|u_{\langle\ell\rangle}^{i+1} - u_\ell^{i+1}|\!|\!|_{1,u_\ell^i}^2} + \underbrace{|\!|\!|u_\ell^{i+1} - u_\ell^i|\!|\!|_{1,u_\ell^i}^2}_{\text{linearization error}}.$$

Proof: Since $u_\ell^{i+1} - u_\ell^i \in V_\ell$,

$$\left|\left|\left|\mathcal{R}(u_\ell^i)\right|\right|\right|^2_{-1,u_\ell^i} = \left|\left|\left|u_\ell^i - u_{\langle\ell\rangle}^{i+1}\right|\right|\right|^2_{1,u_\ell^i} = \left|\left|\left|(u_\ell^i - u_\ell^{i+1}) + (u_\ell^{i+1} - u_{\langle\ell\rangle}^{i+1})\right|\right|\right|^2_{1,u_\ell^i}$$

$$= \left|\left|\left|u_{\langle\ell\rangle}^{i+1} - u_\ell^{i+1}\right|\right|\right|^2_{1,u_\ell^i} + \left|\left|\left|u_\ell^{i+1} - u_\ell^i\right|\right|\right|^2_{1,u_\ell^i} + 2\underbrace{\mathfrak{L}(u_\ell^i; u_{\langle\ell\rangle}^{i+1} - u_\ell^{i+1}, u_\ell^{i+1} - u_\ell^i)}_{=0\text{, due to Galerkin orthogonality}}$$

$$= \left|\left|\left|\mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_\ell^{i+1})\right|\right|\right|^2_{-1,u_\ell^i} + \left|\left|\left|u_\ell^{i+1} - u_\ell^i\right|\right|\right|^2_{1,u_\ell^i}.$$

Proof: Since $u_\ell^{i+1} - u_\ell^i \in V_\ell$,

$$\left|\!\left|\!\left|\mathcal{R}(u_\ell^i)\right|\!\right|\!\right|_{-1,u_\ell^i}^2 = \left|\!\left|\!\left|u_\ell^i - u_{\langle\ell\rangle}^{i+1}\right|\!\right|\!\right|_{1,u_\ell^i}^2 = \left|\!\left|\!\left|(u_\ell^i - u_\ell^{i+1}) + (u_\ell^{i+1} - u_{\langle\ell\rangle}^{i+1})\right|\!\right|\!\right|_{1,u_\ell^i}^2$$

$$= \left|\!\left|\!\left|u_{\langle\ell\rangle}^{i+1} - u_\ell^{i+1}\right|\!\right|\!\right|_{1,u_\ell^i}^2 + \left|\!\left|\!\left|u_\ell^{i+1} - u_\ell^i\right|\!\right|\!\right|_{1,u_\ell^i}^2 + 2\underbrace{\mathfrak{L}(u_\ell^i; u_{\langle\ell\rangle}^{i+1} - u_\ell^{i+1}, u_\ell^{i+1} - u_\ell^i)}_{=0,\text{ due to Galerkin orthogonality}}$$

$$= \left|\!\left|\!\left|\mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_\ell^{i+1})\right|\!\right|\!\right|_{-1,u_\ell^i}^2 + \left|\!\left|\!\left|u_\ell^{i+1} - u_\ell^i\right|\!\right|\!\right|_{1,u_\ell^i}^2.$$

- The linerization error is computed directly, we define

$$\eta_{\mathrm{lin},\Omega}^i := \left|\!\left|\!\left|u_\ell^{i+1} - u_\ell^i\right|\!\right|\!\right|_{1,u_\ell^i}.$$

Proof: Since $u_\ell^{i+1} - u_\ell^i \in V_\ell$,

$$\left|\left|\left|\mathcal{R}(u_\ell^i)\right|\right|\right|^2_{-1,u_\ell^i} = \left|\left|\left| u_\ell^i - u_{\langle\ell\rangle}^{i+1} \right|\right|\right|^2_{1,u_\ell^i} = \left|\left|\left|(u_\ell^i - u_\ell^{i+1}) + (u_\ell^{i+1} - u_{\langle\ell\rangle}^{i+1})\right|\right|\right|^2_{1,u_\ell^i}$$

$$= \left|\left|\left| u_{\langle\ell\rangle}^{i+1} - u_\ell^{i+1} \right|\right|\right|^2_{1,u_\ell^i} + \left|\left|\left| u_\ell^{i+1} - u_\ell^i \right|\right|\right|^2_{1,u_\ell^i} + 2\underbrace{\mathfrak{L}(u_\ell^i; u_{\langle\ell\rangle}^{i+1} - u_\ell^{i+1}, u_\ell^{i+1} - u_\ell^i)}_{=0,\text{ due to Galerkin orthogonality}}$$

$$= \left|\left|\left|\mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_\ell^{i+1})\right|\right|\right|^2_{-1,u_\ell^i} + \left|\left|\left| u_\ell^{i+1} - u_\ell^i \right|\right|\right|^2_{1,u_\ell^i}.$$

- The linerization error is computed directly, we define

$$\eta^i_{\mathrm{lin},\Omega} := \left|\left|\left| u_\ell^{i+1} - u_\ell^i \right|\right|\right|_{1,u_\ell^i}.$$

- For estimating $\left|\left|\left|\mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_\ell^{i+1})\right|\right|\right|_{-1,u_\ell^i}$ we introduce $\eta^i_{\mathrm{disc},\Omega}$, following the analysis on robust estimates of **singularly perturbed reaction -diffusion problems** in [Verfürth (1998)], [Ainsworth & Vejchodský (2011, 2014)], [Smears & Vohralík (2020)]

UHASSELT fwo

Theorem 2 Reliable, efficient, and robust a posteriori estimates

Global reliability

$$|||\mathcal{R}(u_\ell^i)|||_{-1,u_\ell^i}^2 \leq [\eta_\Omega^i]^2 := \sum_{K\in\mathcal{T}_\ell}([\eta_{\mathrm{disc},K}^i]^2 + [\eta_{\mathrm{lin},K}^i]^2).$$

Theorem 2 Reliable, efficient, and robust a posteriori estimates

Global reliability

$$|||\mathcal{R}(u_\ell^i)|||^2_{-1,u_\ell^i} \leq [\eta_\Omega^i]^2 := \sum_{K\in\mathcal{T}_\ell} ([\eta^i_{\mathrm{disc},K}]^2 + [\eta^i_{\mathrm{lin},K}]^2).$$

Global efficiency

$$[\eta_\Omega^i]^2 \lesssim |||\mathcal{R}(u_\ell^i)|||^2_{-1,u_\ell^i} + \text{ (data oscillation terms)}.$$

Theorem 2 Reliable, efficient, and robust a posteriori estimates

Global reliability

$$|||\mathcal{R}(u_\ell^i)|||^2_{-1,u_\ell^i} \le [\eta_\Omega^i]^2 := \sum_{K\in\mathcal{T}_\ell} \left([\eta^i_{\mathrm{disc},K}]^2 + [\eta^i_{\mathrm{lin},K}]^2\right).$$

Global efficiency

$$[\eta_\Omega^i]^2 \lesssim |||\mathcal{R}(u_\ell^i)|||^2_{-1,u_\ell^i} + \text{ (data oscillation terms)}.$$

Local efficiency

For $\omega \subset \Omega$, there exists a neighbourhood $\mathfrak{T}_\omega \subseteq \Omega$ such that

$$[\eta_\omega^i]^2 \lesssim |||\mathcal{R}(u_\ell^{i+1})|||^2_{-1,u_\ell^i,\mathfrak{T}_\omega} + [\eta^i_{\mathrm{lin},\mathfrak{T}_\omega}]^2 + \text{(data oscillation terms)}.$$

UHASSELT fwo

3 Class of problems

Class 1: gradient-dependent diffusivity problems

For all $\varphi \in H_0^1(\Omega)$, $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$ is defined as

$$\langle \mathcal{R}(u), \varphi \rangle := \langle f(\boldsymbol{x}, u), \varphi \rangle - (\boldsymbol{\sigma}(\boldsymbol{x}, \nabla u), \nabla \varphi)$$

3 Class of problems

Class 1: gradient-dependent diffusivity problems

For all $\varphi \in H_0^1(\Omega)$, $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$ is defined as

$$\langle \mathcal{R}(u), \varphi \rangle := \langle f(\boldsymbol{x}, u), \varphi \rangle - (\boldsymbol{\sigma}(\boldsymbol{x}, \nabla u), \nabla \varphi)$$

Assumption 1 is satisfied if $f(\boldsymbol{x}, \cdot)$, $\boldsymbol{\sigma}(\boldsymbol{x}, \cdot)$ are monotone and Lipschitz

$$(\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{y}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{z})) \cdot (\boldsymbol{y} - \boldsymbol{z}) \geq \lambda_{\mathrm{m}} |\boldsymbol{y} - \boldsymbol{z}|^2 \quad \text{for } \boldsymbol{x} \in \Omega \text{ and } \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^d,$$

$$|\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{y}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{z})| \leq \lambda_{\mathrm{M}} |\boldsymbol{y} - \boldsymbol{z}| \quad \text{for } \boldsymbol{x} \in \Omega \text{ and } \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^d.$$

with

$$\mathrm{dist}(u, v) = \|\nabla (u - v)\|$$

3 Class of problems

Class 1: gradient-dependent diffusivity problems

For all $\varphi \in H_0^1(\Omega)$, $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$ is defined as

$$\langle \mathcal{R}(u), \varphi \rangle := \langle f(\boldsymbol{x}, u), \varphi \rangle - (\boldsymbol{\sigma}(\boldsymbol{x}, \nabla u), \nabla \varphi)$$

Assumption 1 is satisfied if $f(\boldsymbol{x}, \cdot)$, $\boldsymbol{\sigma}(\boldsymbol{x}, \cdot)$ are monotone and Lipschitz

$$(\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{y}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{z})) \cdot (\boldsymbol{y} - \boldsymbol{z}) \geq \lambda_{\mathrm{m}} |\boldsymbol{y} - \boldsymbol{z}|^2 \quad \text{for } \boldsymbol{x} \in \Omega \text{ and } \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^d,$$

$$|\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{y}) - \boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{z})| \leq \lambda_{\mathrm{M}} |\boldsymbol{y} - \boldsymbol{z}| \quad \text{for } \boldsymbol{x} \in \Omega \text{ and } \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^d.$$

with

$$\mathrm{dist}(u, v) = \|\nabla(u - v)\|$$

Example (Mean curvature flow) For $a(\cdot)$ satisfying ellipticity condition and $b(\cdot) > 0$: $\boldsymbol{\sigma}(\boldsymbol{x}, \boldsymbol{y}) = a(\boldsymbol{x}) + \frac{b(\boldsymbol{x})\boldsymbol{y}}{(1+|\boldsymbol{y}|^2)^{\frac{1}{2}}}$

Linearization operator

Considering the linearization operator

$$\mathfrak{L}(u_\ell^i; v, w) := (L(\boldsymbol{x}, u_\ell^i)\, v, w) + (\mathfrak{a}(\boldsymbol{x}, u_\ell^i)\nabla v, \nabla w),$$

the coefficient functions for commonly used linearization schemes are

Scheme	$L(\boldsymbol{x}, v)$	$\mathfrak{a}(\boldsymbol{x}, v)/\tau$
Kačanov (fixed point)	$\partial_\xi f(\boldsymbol{x}, v)$	$A(\boldsymbol{x}, \lvert\nabla v\rvert)$
Zarantonello	0	Λ (constant) > 0

3 Class of problems

Class 2: gradient-independent diffusivity problems

For all $\varphi \in H_0^1(\Omega)$, $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$ is defined as

$$\langle \mathcal{R}(u), \varphi \rangle := \langle f(\boldsymbol{x}, u), \varphi \rangle - \tau(\bar{\mathbf{K}}(\boldsymbol{x})(\mathcal{D}(\boldsymbol{x}, u)\nabla u + \boldsymbol{q}(\boldsymbol{x}, u)), \nabla \varphi)$$

Class 2: gradient-independent diffusivity problems

For all $\varphi \in H_0^1(\Omega)$, $\mathcal{R} : H_0^1(\Omega) \to H^{-1}(\Omega)$ is defined as

$$\langle \mathcal{R}(u), \varphi \rangle := \langle f(\boldsymbol{x}, u), \varphi \rangle - \tau(\bar{\mathbf{K}}(\boldsymbol{x})(\mathcal{D}(\boldsymbol{x}, u)\nabla u + \boldsymbol{q}(\boldsymbol{x}, u)), \nabla\varphi)$$

Assumption 1 is satisfied if $\tau > 0$ is small and

- $\mathcal{D} : \Omega \times \mathbb{R} \to \mathbb{R}^+$ is bounded and Lipschitz
- $\bar{\mathbf{K}} : \Omega \to \mathbb{R}^{d\times d}$ is symmetric positive definite
- $f : \Omega \times \mathbb{R} \to \mathbb{R}$ is monotone and Lipschitz upto the boundary
- $\boldsymbol{q} : \Omega \times \mathbb{R} \to \mathbb{R}^d$ is bounded and satisfies a Lipschitz condition*

with

$$\mathrm{dist}(u, v) = \left\| \bar{\mathbf{K}}^{\frac{1}{2}} \nabla \int_u^v \mathcal{D} \right\|$$

3 Class of problems

Class 2: gradient-independent diffusivity problems

For all $\varphi \in H^1_0(\Omega)$, $\mathcal{R} : H^1_0(\Omega) \to H^{-1}(\Omega)$ is defined as

$$\langle \mathcal{R}(u), \varphi \rangle := \langle f(\boldsymbol{x}, u), \varphi \rangle - \tau(\bar{\mathbf{K}}(\boldsymbol{x})(\mathcal{D}(\boldsymbol{x}, u)\nabla u + \boldsymbol{q}(\boldsymbol{x}, u)), \nabla \varphi)$$

Semilinear equations $-\Delta u = f(\boldsymbol{x}, u)$
Such equations pop up in quantum mechanics (special solutions to nonlinear Klein–Gordon equations), gravitation influences on stars, membrane buckling problems...

Time-discrete nonlinear advection-reaction-diffusion equations

with time-step $\tau > 0$, the following evolutions equations reduce to this case

poro-Fischer equations: $\partial_t u = \Delta u^m + \lambda\, u\,(1 - u)$

the Richards equation: $\partial_t S(u) = \nabla \cdot [\bar{\mathbf{K}}(\boldsymbol{x})\kappa(S(u))(\nabla u + \boldsymbol{g})] + f(\boldsymbol{x}, u)$

biofilm equations: $\partial_t u_k = \mu_k \Delta \Phi_k(u_k) + f_k((u_k)_{k=1}^n)$

Abstract linearization

Considering the linearization operator

$$\mathfrak{L}(u_\ell^i; v, w) := (L(\boldsymbol{x}, u_\ell^i)\, v, w) + (\mathfrak{a}(\boldsymbol{x}, u_\ell^i)\nabla v, \nabla w),$$

the coefficient functions for commonly used linearization schemes are

Scheme	$L(\boldsymbol{x}, v)$	$\mathfrak{a}(\boldsymbol{x}, v)/\tau$
Picard (fixed point)	$\partial_\xi f(\boldsymbol{x}, v)$	$\bar{\mathbf{K}}(\boldsymbol{x})\,\mathcal{D}(\boldsymbol{x}, v)$
Jäger–Kačur	$\max_{\xi\in\mathbb{R}} \left(\frac{f(\boldsymbol{x},\xi)-f(\boldsymbol{x},v)}{\xi-v}\right)$	$\bar{\mathbf{K}}(\boldsymbol{x})\,\mathcal{D}(\boldsymbol{x}, v)$
L-scheme	L (constant) $\geq \frac{1}{2}\sup \partial_\xi f$	$\bar{\mathbf{K}}(\boldsymbol{x})\,\mathcal{D}(\boldsymbol{x}, v)$
M-scheme	$\partial_\xi f(\boldsymbol{x}, v) + M\tau$ (constant)	$\bar{\mathbf{K}}(\boldsymbol{x})\,\mathcal{D}(\boldsymbol{x}, v)$

Abstract linearization

Considering the linearization operator

$$\mathfrak{L}(u_\ell^i; v, w) := (L(\boldsymbol{x}, u_\ell^i)\, v, w) + (\mathfrak{a}(\boldsymbol{x}, u_\ell^i)\nabla v, \nabla w),$$

the coefficient functions for commonly used linearization schemes are

Scheme	$L(\boldsymbol{x}, v)$	$\mathfrak{a}(\boldsymbol{x}, v)/\tau$
Picard (fixed point)	$\partial_\xi f(\boldsymbol{x}, v)$	$\bar{\mathbf{K}}(\boldsymbol{x})\, \mathcal{D}(\boldsymbol{x}, v)$
Jäger–Kačur	$\max_{\xi\in\mathbb{R}} \left(\frac{f(\boldsymbol{x},\xi)-f(\boldsymbol{x},v)}{\xi-v}\right)$	$\bar{\mathbf{K}}(\boldsymbol{x})\, \mathcal{D}(\boldsymbol{x}, v)$
L-scheme	L (constant) $\geq \frac{1}{2}\sup \partial_\xi f$	$\bar{\mathbf{K}}(\boldsymbol{x})\, \mathcal{D}(\boldsymbol{x}, v)$
M-scheme	$\partial_\xi f(\boldsymbol{x}, v) + M\tau$ (constant)	$\bar{\mathbf{K}}(\boldsymbol{x})\, \mathcal{D}(\boldsymbol{x}, v)$

- Newton scheme leads to a non-symmetric $\mathfrak{L}$ and is treated separately

Effectivity indices

Global effectivity index: $\text{Eff. Ind.} := \eta_{\Omega}^{i} / |||\mathcal{R}(u_\ell^i)|||_{-1,u_\ell^i}$

Local effectivity index: $(\text{Eff. Ind.})_K := \eta_K^i / |||\mathcal{R}(u_\ell^i)|||_{-1,u_\ell^i,K}, \qquad K \in \mathcal{T}_\ell,$

4 Gradient-independent diffusivity case: the Richards equation

For $\Omega = (0,1) \times (0,1)$ we study

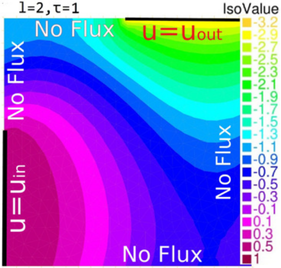
$$\langle \mathcal{R}(u_\ell), \varphi \rangle = (S(\bar{u}) - S(u_\ell), \varphi)$$
$$-\tau(\bar{\mathbf{K}}\kappa(S(u_\ell))[\nabla u_\ell - \boldsymbol{g}], \nabla\varphi)$$

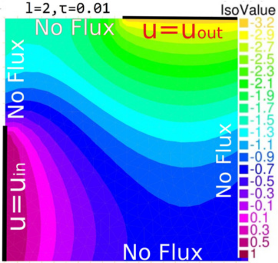
where the van Genuchten parametrization for S, κ is used:

$$\begin{cases} S(\xi) := \left(1 + (2-\xi)^{\frac{1}{1-\lambda}}\right)^{-\lambda}, \\ \kappa(s) := \sqrt{s}\left(1 - (1 - s^{\frac{1}{\lambda}})^{\lambda}\right)^2, \end{cases}$$

with $\lambda = 0.5$, $u_\ell^0 = 0$,

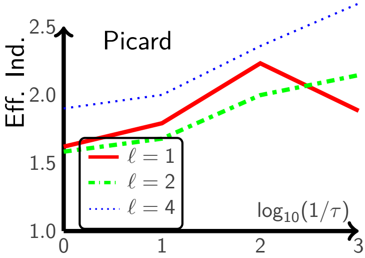
$$\bar{\mathbf{K}} = \begin{bmatrix} 1 & 0.2 \\ 0.2 & 1 \end{bmatrix}, \text{ and } \boldsymbol{g} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

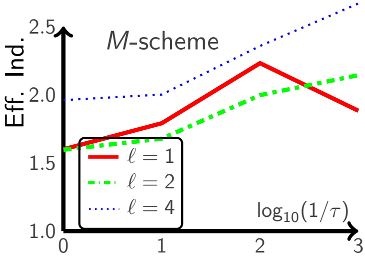


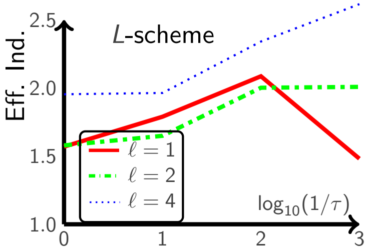


UHASSELT fwo

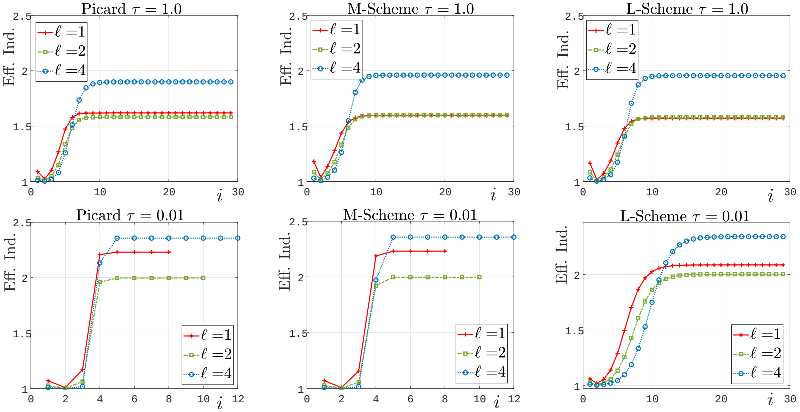
4 Robustness with respect to λ_M/λ_m represented by $1/\tau$



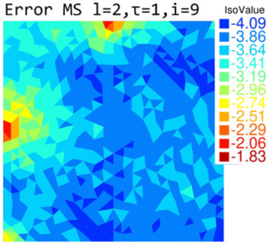




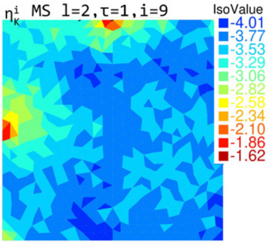
4 Global effectivity



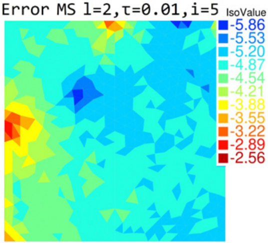
4 Distribution of error vs. estimates

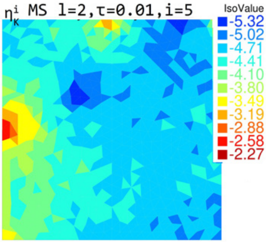


Error

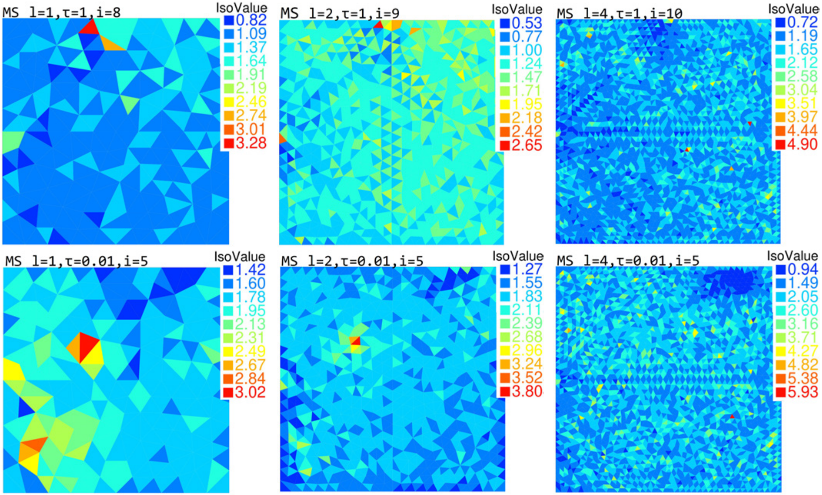


Estimate

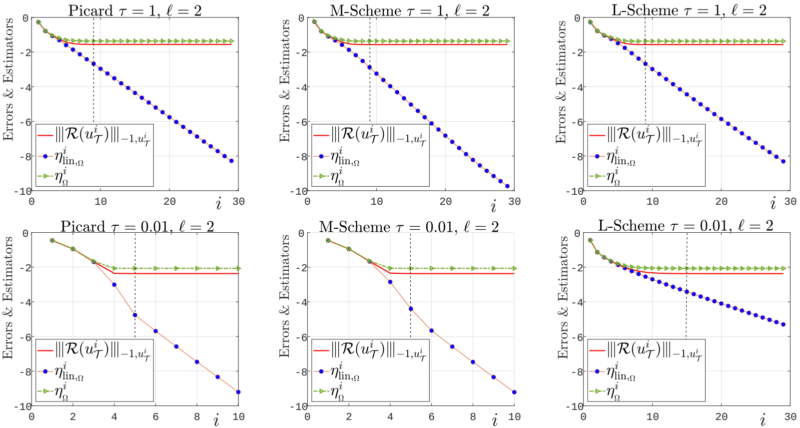




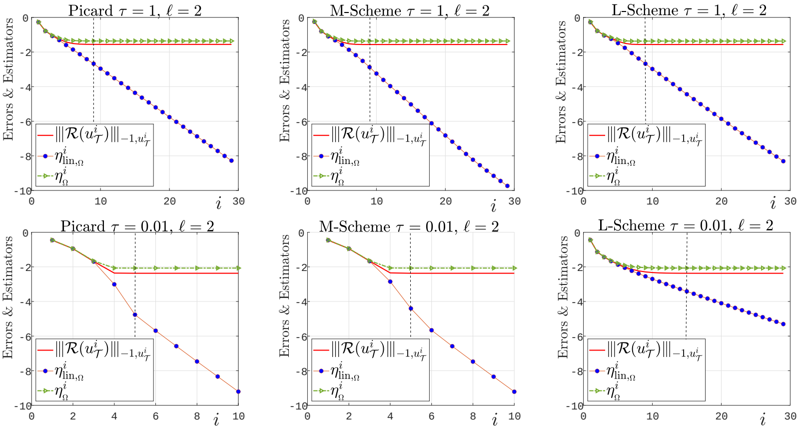
4 Local effectivity



4 Error with linearization iterations



4 Error with linearization iterations



Adaptive iteration stopping criteria:

$$\eta^i_{\mathrm{lin},\Omega} \leq 0.05\,[\eta^i_\Omega].$$

UHASSELT fwo

4 Gradient independent diffusivity case

We consider in Ω the equation

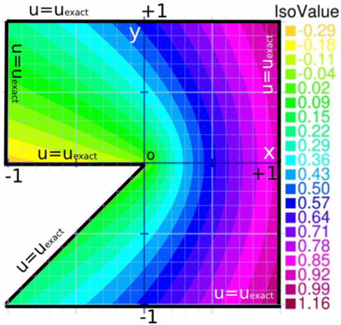
$$\varepsilon u - \nabla \cdot [A(|\nabla u|)\nabla u] = f$$

where

$$A(\mathbf{y}) = 2 + \frac{\mathbf{y}}{(1 + |\mathbf{y}|^2)},$$

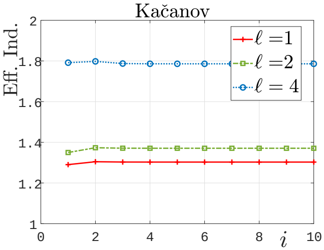
$\varepsilon = 10^{-2}$, and a singular $f \in H^{-1}(\Omega)$ is chosen such that the solution becomes

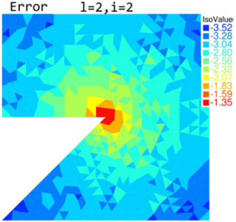
$$u_{\text{exact}} = r^{\frac{4}{7}} cos\left(\tfrac{4}{7}\theta\right).$$

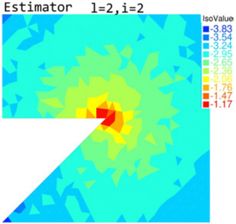


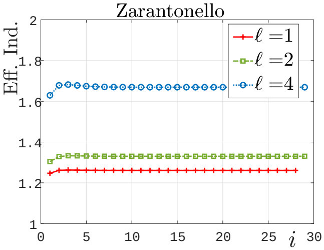
Ω

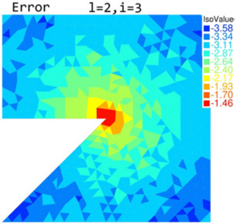
4 Global effectivity and distribution of error

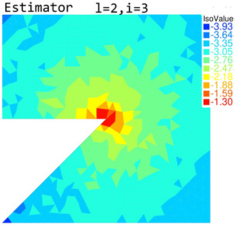




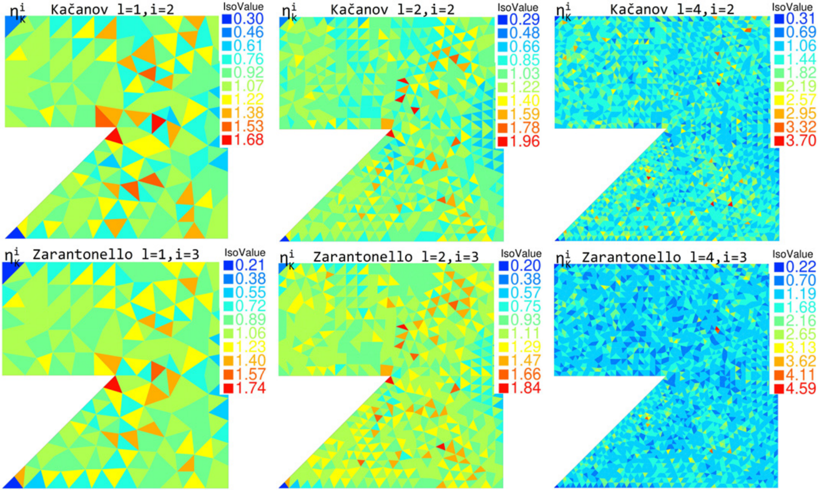




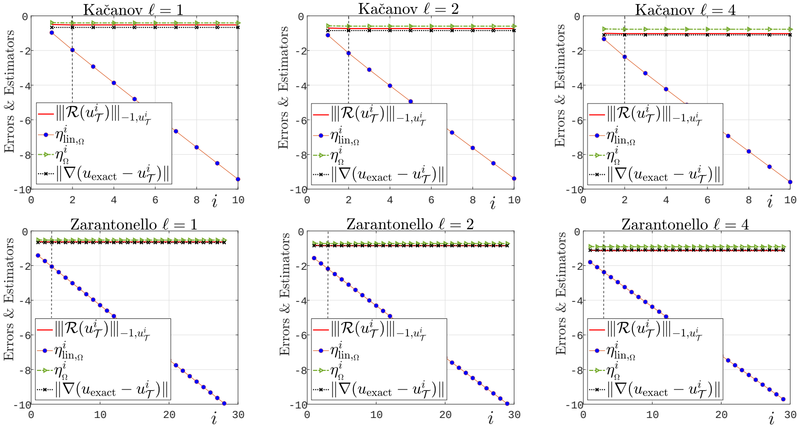




4 Local effectivity



4 Error with linearization iterations



4 The Newton scheme

For the Newton scheme, the linearization operator

$$\mathfrak{L}(u_\ell^i; v, w) := (L(\boldsymbol{x}, u_\ell^i)\, v, w) + (\mathfrak{a}(\boldsymbol{x}, u_\ell^i)\nabla v, \nabla w) + (\boldsymbol{w}(\boldsymbol{x}, u_\ell^i) v, \nabla w),$$

is non-symmetric.

4 The Newton scheme

For the Newton scheme, the linearization operator

$$\mathfrak{L}(u_\ell^i; v, w) := (L(\boldsymbol{x}, u_\ell^i)\, v, w) + (\mathfrak{a}(\boldsymbol{x}, u_\ell^i)\nabla v, \nabla w) + (\boldsymbol{w}(\boldsymbol{x}, u_\ell^i) v, \nabla w),$$

is non-symmetric. However, if for some $C_N \in [0, 2)$ we have

$$\boldsymbol{w}(\boldsymbol{x}, u_\ell^i)\, \mathfrak{a}^{-1}(\boldsymbol{x}, u_\ell^i)\, \boldsymbol{w}(\boldsymbol{x}, u_\ell^i) \leq C_N^2\, L(\boldsymbol{x}, u_\ell^i), \quad \forall\, \boldsymbol{x} \in \Omega, \text{ and } i \in \mathbb{N},$$

4 The Newton scheme

For the Newton scheme, the linearization operator

$$\mathfrak{L}(u_\ell^i; v, w) := (L(\boldsymbol{x}, u_\ell^i)\, v, w) + (\mathfrak{a}(\boldsymbol{x}, u_\ell^i)\nabla v, \nabla w) + (\boldsymbol{w}(\boldsymbol{x}, u_\ell^i) v, \nabla w),$$

is non-symmetric. However, if for some $C_N \in [0, 2)$ we have

$$\boldsymbol{w}(\boldsymbol{x}, u_\ell^i)\, \mathfrak{a}^{-1}(\boldsymbol{x}, u_\ell^i)\, \boldsymbol{w}(\boldsymbol{x}, u_\ell^i) \leq C_N^2\, L(\boldsymbol{x}, u_\ell^i), \quad \forall\, \boldsymbol{x} \in \Omega, \text{ and } i \in \mathbb{N},$$

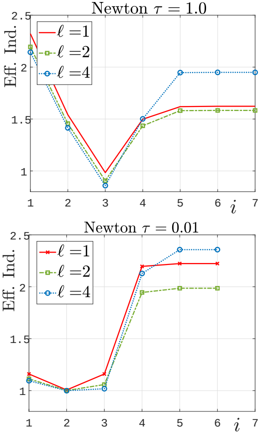
then,

$$C_{\mathrm{m}}(C_N) \left[\left|\!\left|\!\left| \mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_\ell^{i+1}) \right|\!\right|\!\right|_{-1,u_\ell^i}^2 + \left|\!\left|\!\left| u_\ell^{i+1} - u_\ell^i \right|\!\right|\!\right|_{1,u_\ell^i}^2 \right] \leq \left|\!\left|\!\left| \mathcal{R}(u_\ell^i) \right|\!\right|\!\right|_{-1,u_\ell^i}^2$$

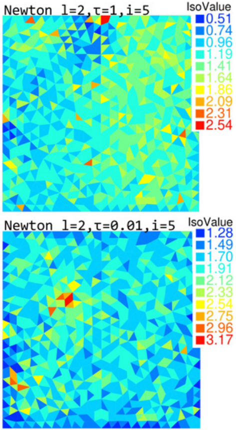
$$\leq C_{\mathrm{M}}(C_N) \left[\left|\!\left|\!\left| \mathcal{R}_{\mathrm{lin}}^{u_\ell^i}(u_\ell^{i+1}) \right|\!\right|\!\right|_{-1,u_\ell^i}^2 + \left|\!\left|\!\left| u_\ell^{i+1} - u_\ell^i \right|\!\right|\!\right|_{1,u_\ell^i}^2 \right]$$

with $C_{\mathrm{m}}(C_N), C_{\mathrm{M}}(C_N) \to 1$ if $C_N \searrow 0$.

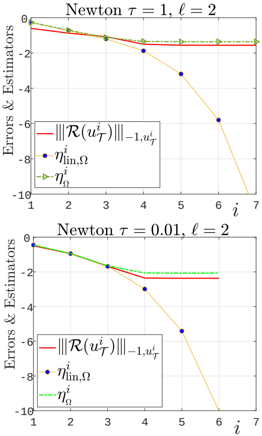
For gradient independent diffusivity case, we have



Global Effectivity



Local Effectivity



Error with iterations

5 Outline

5 Nonlinear advection-reaction-diffusion equation

Richards equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}} \kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

p is pressure, $s := S(p)$ is saturation

5 Nonlinear advection-reaction-diffusion equation

Richards equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}}\kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

p is pressure, $s := S(p)$ is saturation

- Obtained from combining *mass balance*

$$\partial_t s + \nabla \cdot \boldsymbol{\sigma} = f(s, \boldsymbol{x}, t),$$

5 Nonlinear advection-reaction-diffusion equation

Richards equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}}\kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

p is pressure, $s := S(p)$ is saturation

- Obtained from combining *mass balance*

$$\partial_t s + \nabla \cdot \boldsymbol{\sigma} = f(s, \boldsymbol{x}, t),$$

- the *Darcy Law*

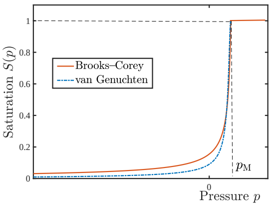
$$\boldsymbol{\sigma} = -\bar{\mathbf{K}}\kappa(s)(\nabla p + \boldsymbol{g}),$$

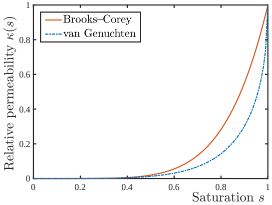
Richards equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}} \kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

Function properties

- $S \in \mathrm{Lip}(\mathbb{R})$ is increasing in $(-\infty, p_{\mathrm{M}})$, $S(-\infty) = 0$ and
$$S'(p) = 0,\ S(p) = 1 \text{ for all } p > p_{\mathrm{M}}.$$
- $\kappa \in C^1([0,1])$ is increasing with $\kappa(0) \geq 0$ and $\kappa(1) = 1$.





5 Nonlinear advection-reaction-diffusion equation

Richards equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}}\kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

Function properties

- $S \in \mathrm{Lip}(\mathbb{R})$ is increasing in $(-\infty, p_{\mathrm{M}})$, $S(-\infty) = 0$ and
$$S'(p) = 0,\ S(p) = 1 \text{ for all } p > p_{\mathrm{M}}.$$
- $\kappa \in C^1([0,1])$ is increasing with $\kappa(0) \geq 0$ and $\kappa(1) = 1$.
- $\bar{\mathbf{K}} : \Omega \mapsto \mathbb{R}^{d\times d}$ is piece-wise constant in Ω, bounded, symmetric positive definite, and satisfies the ellipticity condition,
$$K_{\mathrm{m}}|\boldsymbol{\zeta}|^2 \leq \boldsymbol{\zeta}^{\mathrm{T}}\bar{\mathbf{K}}\boldsymbol{\zeta} \leq K_{\mathrm{M}}|\boldsymbol{\zeta}|^2, \quad \forall \zeta \in \mathbb{R}^d/\{0\}.$$

Richards equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}} \kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

Function properties

- $S \in \mathrm{Lip}(\mathbb{R})$ is increasing in $(-\infty, p_{\mathrm{M}})$, $S(-\infty) = 0$ and
$$S'(p) = 0,\ S(p) = 1 \text{ for all } p > p_{\mathrm{M}}.$$
- $\kappa \in C^1([0,1])$ is increasing with $\kappa(0) \geq 0$ and $\kappa(1) = 1$.
- $\bar{\mathbf{K}} : \Omega \mapsto \mathbb{R}^{d\times d}$ is piece-wise constant in Ω, bounded, symmetric positive definite, and satisfies the ellipticity condition,
$$K_{\mathrm{m}}|\zeta|^2 \leq \zeta^{\mathrm{T}} \bar{\mathbf{K}} \zeta \leq K_{\mathrm{M}}|\zeta|^2, \quad \forall \zeta \in \mathbb{R}^d/\{0\}.$$
- $f \in C^1([0,1] \times \Omega \times \mathbb{R})$.

5 Nonlinear advection-reaction-diffusion equation

Richards Equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}} \kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

Main Challenges

1. Nonlinearity 
2. Degeneracy 

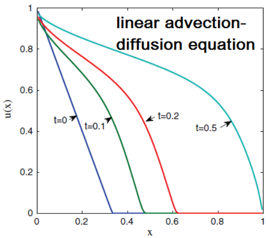
5 Nonlinear advection-reaction-diffusion equation

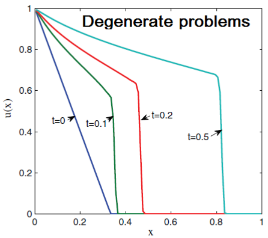
Richards Equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}}\kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

Main Challenges

1. Nonlinearity
2. Degeneracy
   > Parabolic–Hyperbolic: at $s = 0$ if $\kappa(0) = 0$
   $\partial_t s = f$ or $\partial_t s + \nabla \cdot \boldsymbol{F}(s) = f$ for multiphase problems





5 Nonlinear advection-reaction-diffusion equation

Richards Equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}} \kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

Main Challenges

1. Nonlinearity 
2. Degeneracy 
   > Parabolic–Hyperbolic: at $s = 0$ if $\kappa(0) = 0$
   $\partial_t s = f$ or $\partial_t s + \nabla \cdot \boldsymbol{F}(s) = f$ for multiphase problems

   > Parabolic–Elliptic: at $s = 1$ since $S'(p) = 0$ for $p > p_{\mathrm{M}}$
   $\nabla \cdot [\bar{\mathbf{K}} \kappa(1)(\nabla p + \boldsymbol{g})] + f(1, \boldsymbol{x}, t) = 0$

5 Nonlinear advection-reaction-diffusion equation

Richards Equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}} \kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

Main Challenges

1. Nonlinearity
2. Degeneracy
   - Parabolic–Hyperbolic: at $s = 0$ if $\kappa(0) = 0$
     $\partial_t s = f$ or $\partial_t s + \nabla \cdot \boldsymbol{F}(s) = f$ for multiphase problems
   - Parabolic–Elliptic: at $s = 1$ since $S'(p) = 0$ for $p > p_{\mathrm{M}}$
     $\nabla \cdot [\bar{\mathbf{K}} \kappa(1)(\nabla p + \boldsymbol{g})] + f(1, \boldsymbol{x}, t) = 0$
3. Solutions lack regularity

5 Nonlinear advection-reaction-diffusion equation

Richards Equation: modelling flow of water through soil

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}}\kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

Main Challenges

1. Nonlinearity
2. Degeneracy
   - Parabolic–Hyperbolic: at $s = 0$ if $\kappa(0) = 0$
     $\partial_t s = f$ or $\partial_t s + \nabla \cdot \boldsymbol{F}(s) = f$ for multiphase problems
   - Parabolic–Elliptic: at $s = 1$ since $S'(p) = 0$ for $p > p_{\mathrm{M}}$
     $\nabla \cdot [\bar{\mathbf{K}}\kappa(1)(\nabla p + \boldsymbol{g})] + f(1, \boldsymbol{x}, t) = 0$
3. Solutions lack regularity

Literature: [Dolejší *et al* (2013)][Bernardi *et al* (2014)][Cancès *et al* (2014)] [Verfürth (2004)]; [Di Pietro *et al* (2015)]; [Ohlberger (2001)]

Pressure formulation

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}} \kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

5 Alternative formulations

Pressure formulation

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}} \kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

The Kirchhoff transform and some definitions

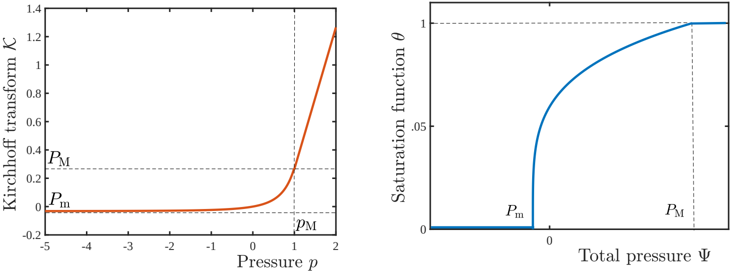
$$\mathcal{K}(p) = \int_0^p \kappa(S(\varrho)) \, \mathrm{d}\varrho, \qquad \theta = S \circ \mathcal{K}^{-1}$$

Pressure formulation

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}} \kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

The Kirchhoff transform and some definitions

$$\mathcal{K}(p) = \int_0^p \kappa(S(\varrho))\, \mathrm{d}\varrho, \qquad \theta = S \circ \mathcal{K}^{-1}$$



5 Alternative formulations

Pressure formulation

$$\partial_t S(p) = \nabla \cdot [\bar{\mathbf{K}} \kappa(S(p))(\nabla p + \boldsymbol{g})] + f(S(p), \boldsymbol{x}, t)$$

The Kirchhoff transform and some definitions

$$\mathcal{K}(p) = \int_0^p \kappa(S(\varrho))\, \mathrm{d}\varrho, \qquad \theta = S \circ \mathcal{K}^{-1}$$

Total pressure formulation

For $\Psi = \mathcal{K}(p)$,

$$\partial_t \theta(\Psi) = \nabla \cdot [\bar{\mathbf{K}}(\nabla \Psi + \mathcal{K}(\theta(\Psi))\boldsymbol{g})] + f(\theta(\Psi), \boldsymbol{x}, t)$$

5 Well-posedness

Weak total pressure formulation

For the initial condition s_0 bounded in $(0,1]$ a.e., find $\Psi \in L^2(0,T;H_0^1(\Omega))$, $s=\theta(\Psi) \in H^1(0,T;H^{-1}(\Omega))$, $s(0)=s_0$ satisying $\forall\,\varphi \in L^2(0,T;H_0^1(\Omega))$,

$$\int_0^T [\langle \partial_t s, \varphi\rangle + (\bar{\mathbf{K}}[\nabla\Psi + \kappa(s)\boldsymbol{g}], \nabla\varphi)] = \int_0^T (f(s,\boldsymbol{x},t), \varphi)$$

5 Well-posedness

Weak total pressure formulation

For the initial condition s_0 bounded in $(0,1]$ a.e., find $\Psi \in L^2(0,T;H^1_0(\Omega))$, $s = \theta(\Psi) \in H^1(0,T;H^{-1}(\Omega))$, $s(0) = s_0$ satisying $\forall\,\varphi \in L^2(0,T;H^1_0(\Omega))$,

$$\int_0^T [\langle \partial_t s, \varphi\rangle + (\bar{\mathbf{K}}[\nabla\Psi + \kappa(s)\boldsymbol{g}], \nabla\varphi)] = \int_0^T (f(s,\boldsymbol{x},t),\varphi)$$

Theorem [Alt & Luckhaus (1983)][Otto (1991)]

There exists a unique weak solution Ψ for the total pressure formulations.

5 Maximum principle

To avoid the parabolic-hyperbolic degeneracy we need $s \geq S_{\mathrm{m}} > 0$

To avoid the parabolic-hyperbolic degeneracy we need $s \geq S_{\mathrm{m}} > 0$

Proposition

If s_0 is bounded in $[\varepsilon, 1]$ for some $\varepsilon > 0$, then there exists *saturation lower-bound function* $S_{\mathrm{m}} : [0, T] \to (0, 1]$ such that for almost all $(\boldsymbol{x}, t) \in \Omega \times [0, T]$,

$$s(\boldsymbol{x}, t) = S(p(\boldsymbol{x}, t)) \geq S_{\mathrm{m}}(t) > 0.$$

To avoid the parabolic-hyperbolic degeneracy we need $s \geq S_{\mathrm{m}} > 0$

Proposition

If s_0 is bounded in $[\varepsilon, 1]$ for some $\varepsilon > 0$, then there exists *saturation lower-bound function* $S_{\mathrm{m}} : [0, T] \to (0, 1]$ such that for almost all $(\boldsymbol{x}, t) \in \Omega \times [0, T]$,

$$s(\boldsymbol{x}, t) = S(p(\boldsymbol{x}, t)) \geq S_{\mathrm{m}}(t) > 0.$$

Computing S_{m}

For example, under minor restrictions

$$S_{\mathrm{m}}(t) = \min_{\boldsymbol{x}\in\Omega}\{s_0(\boldsymbol{x})\} + \int_0^t \min_{\boldsymbol{x}\in\Omega,\varrho>0}\{f(S_{\mathrm{m}}(\varrho), \boldsymbol{x}, \varrho)\}\,\mathrm{d}\varrho$$

is a saturation lower-bound function.

Residual

For $\Psi_{h\tau} \in L^2(0,T;H_0^1(\Omega))$, $s_{h\tau} = \theta(\Psi_{h\tau}) \in H^1(0,T;H^{-1}(\Omega))$ the residual $\mathcal{R}(\Psi_{h\tau}) \in L^2(0,T;H^{-1}(\Omega))$ is

$$\int_0^T \langle \mathcal{R}(\Psi_{h\tau}), \varphi \rangle = \int_0^T [(f(s_{h\tau}, \boldsymbol{x}, t), \varphi) - \langle \partial_t s_{h\tau}, \varphi \rangle - (\bar{\mathbf{K}}[\nabla \Psi_{h\tau} + \kappa(s_{h\tau})\boldsymbol{g}], \nabla \varphi)]$$

The $H^{\pm 1}_{\bar{\mathbf{K}}}$ norm

For $\omega \subseteq \Omega$, the following equivalent norms of $H^{\pm 1}(\omega)$ are defined

$$\|\varrho\|_{H^1_{\bar{\mathbf{K}}}(\omega)} := \|\bar{\mathbf{K}}^{\frac{1}{2}} \nabla \varrho\|_{L^2(\omega)},$$

$$\|\varrho\|_{H^{-1}_{\bar{\mathbf{K}}}(\omega)} := \sup_{\varphi \in H^1_0(\omega)} \frac{\langle \varrho, \varphi \rangle_{H^{-1}(\omega), H^1_0(\omega),}}{\|\varphi\|_{H^1_{\bar{\mathbf{K}}}(\omega)}}.$$

5 Norms

The $H^{\pm 1}_{\bar{\mathbf{K}}}$ norm

For $\omega \subseteq \Omega$, the following equivalent norms of $H^{\pm 1}(\omega)$ are defined

$$\|\varrho\|_{H^1_{\bar{\mathbf{K}}}(\omega)} := \|\bar{\mathbf{K}}^{\frac{1}{2}} \nabla \varrho\|_{L^2(\omega)},$$

$$\|\varrho\|_{H^{-1}_{\bar{\mathbf{K}}}(\omega)} := \sup_{\varphi \in H^1_0(\omega)} \frac{\langle \varrho, \varphi \rangle_{H^{-1}(\omega), H^1_0(\omega),}}{\|\varphi\|_{H^1_{\bar{\mathbf{K}}}(\omega)}}.$$

The equivalent norm of $L^2([0,T])$

For $\alpha : \mathbb{R}^+ \to [0, \infty)$, a time-smoothened equivalent of $L^2([0,T])$-norm is

$$\mathcal{J}_\alpha(\varrho) := \left[\exp\left(-\int_0^T \alpha \right) \int_0^T \left(\varrho^2(t) + \alpha(t) \exp\left(\int_t^T \alpha \right) \int_0^t \varrho^2 \right) \mathrm{d}t \right]^{\frac{1}{2}}.$$

5 Error measure

- The error measure $\|\mathcal{R}(\Psi_{h\tau})\|_{L^2(0,T;H^{-1}_{\bar{\kappa}}(\Omega))}$ might again be too weak

5 Error measure

- The error measure $\|\mathcal{R}(\Psi_{h\tau})\|_{L^2(0,T;H_{\bar{\kappa}}^{-1}(\Omega))}$ might again be too weak

The error metric

For $\omega \subseteq \Omega$, interval $I \subseteq [0, T]$, and $\alpha : \mathbb{R}^+ \to [0, \infty)$, we choose

$$\begin{aligned} \mathrm{dist}_{\omega,I}^{\alpha}(\Psi_1, \Psi_2) := \|\Psi_1 - \Psi_2\|_{L^2(I,H_{\bar{\kappa}}^1(\omega))} \\ + \|\alpha(\theta(\Psi_1) - \theta(\Psi_2))\|_{L^2(\omega\times I)} \\ + \|\partial_t(\theta(\Psi_1) - \theta(\Psi_2))\|_{L^2(I;H_{\bar{\kappa}}^{-1}(\omega))}. \end{aligned}$$

*In the linear case, $\alpha = 0$

5 Lower bound on error by residual

Theorem 3 (a)

For a time-interval $I \in [0, T]$, $\omega \subseteq \Omega$, and arbitrary $\Psi_{h\tau} \in L^2(0, T; H^1_0(\Omega))$ with $\theta(\Psi_{h\tau}) \in H^1(0, T; H^{-1}(\Omega))$ we have

$$\|\mathcal{R}(\Psi_{h\tau})\|_{L^2(I;H^{-1}_{\bar{\mathsf{K}}}(\omega))} \leq \mathrm{dist}^{\alpha}_{\omega,I}(\Psi, \Psi_{h\tau}).$$

5 Lower bound on error by residual

Theorem 3 (a)

For a time-interval $I \in [0, T]$, $\omega \subseteq \Omega$, and arbitrary $\Psi_{h\tau} \in L^2(0, T; H^1_0(\Omega))$ with $\theta(\Psi_{h\tau}) \in H^1(0, T; H^{-1}(\Omega))$ we have

$$\|\mathcal{R}(\Psi_{h\tau})\|_{L^2(I;H^{-1}_{\bar{\mathsf{K}}}(\omega))} \leq \mathrm{dist}^{\alpha}_{\omega,I}(\Psi, \Psi_{h\tau}).$$

proof: Use triangle inequality for the norm $\|\cdot\|_{L^2(I;H^{-1}_{\bar{\mathsf{K}}}(\omega))}$

5 Upper bound on error by residual

Additional quantities

- For $C^{\infty}_{h\tau}(t) := \|\bar{\mathbf{K}}^{\frac{1}{2}} \nabla s_{h\tau}(t)\|^2_{L^\infty(\Omega)}$, assume that $\int_0^T C^{\infty}_{h\tau}(t)\,\mathrm{d}t < \infty$.

UHASSELT fwo

Additional quantities

- For $C^{\infty}_{h\tau}(t) := \|\bar{\mathbf{K}}^{\frac{1}{2}}\nabla s_{h\tau}(t)\|^2_{L^\infty(\Omega)}$, assume that $\int_0^T C^{\infty}_{h\tau}(t)\,\mathrm{d}t < \infty$.
- Parabolic–hyperbolic degeneracy
  Assume that $s(t) \geq S_{\mathrm{m}}(t) > 0$ a.e. in Ω for $t > 0$.

Additional quantities

- For $C^{\infty}_{h\tau}(t) := \|\bar{\mathbf{K}}^{\frac{1}{2}}\nabla s_{h\tau}(t)\|^2_{L^\infty(\Omega)}$, assume that $\int_0^T C^{\infty}_{h\tau}(t)\,\mathrm{d}t < \infty$.

- **Parabolic–hyperbolic degeneracy**
  Assume that $s(t) \geq S_{\mathrm{m}}(t) > 0$ a.e. in Ω for $t > 0$.

- **Parabolic–elliptic degeneracy**
  For $\Omega^{\mathrm{deg}} \supseteq \{s = 1\} \cup \{s_{h\tau} = 1\}$ we define

$$\eta^{\mathrm{deg}}(t) := \sqrt{\tfrac{2}{D(1)}}\left[\|[\Psi_{h\tau}(t) - P_{\mathrm{M}}]_+\|^2_{H^1_{\bar{\mathbf{K}}}(\Omega)} + \|[f(1,\boldsymbol{x},t)]_+\|_{H^{-1}_{\bar{\mathbf{K}}}(\Omega^{\mathrm{deg}}(t))} + \|(\bar{\mathbf{K}}^{\frac{1}{2}} - \frac{\bar{\mathbf{K}}^{-\frac{1}{2}}}{|\Omega^{\mathrm{deg}}(t)|}\int_{\Omega^{\mathrm{deg}}(t)}\bar{\mathbf{K}})\boldsymbol{g}\|_{\Omega^{\mathrm{deg}}(t)})^2\right]^{\frac{1}{2}}$$

Theorem 3 (b)

Estimate in the $L^2(\Omega \times [0,T])$ **norm:**

$$\mathcal{J}_{\mathfrak{C}_1}(\underline{\lambda}_1 \|s - s_{h\tau}\|)^2$$

$$\leq \|s_0 - s_{h\tau}(0)\|^2_{H^{-1}_{\bar{\kappa}}(\Omega)} + \mathcal{J}_{\mathfrak{C}_1}(\bar{\lambda}_1 \|\mathcal{R}(\Psi_{h\tau})\|_{H^{-1}_{\bar{\kappa}}(\Omega)})^2,$$

5 Global reliability

Theorem 3 (b)

Estimate in the $L^2(\Omega \times [0, T])$ **norm:**

$$\mathcal{J}_{\mathfrak{C}_1}(\underline{\lambda}_1 \|s - s_{h\tau}\|)^2$$

$$\leq \|s_0 - s_{h\tau}(0)\|^2_{H^{-1}_{\bar{\kappa}}(\Omega)} + \mathcal{J}_{\mathfrak{C}_1}(\bar{\lambda}_1 \|\mathcal{R}(\Psi_{h\tau})\|_{H^{-1}_{\bar{\kappa}}(\Omega)})^2,$$

Estimate in the $L^2(0, T; H^1(\Omega))$ **norm:**

$$\tfrac{1}{2}\mathcal{J}_{\mathfrak{C}_2}(\underline{\lambda}_2 \|\Psi - \Psi_{h\tau}\|_{H^1_{\bar{\kappa}}(\Omega)})^2$$

$$\leq \|s_0 - s_{h\tau}(0)\|^2 + \mathcal{J}_{\mathfrak{C}_2}\left(\eta^{\mathrm{deg}}\right)^2 + 4\,\mathcal{J}_{\mathfrak{C}_2}\left(\bar{\lambda}_2 \|\mathcal{R}(\Psi_{h\tau})\|_{H^{-1}_{\bar{\kappa}}(\Omega)}\right)^2.$$

5 Global reliability

Theorem 3 (b)

Estimate in the $L^2(\Omega \times [0, T])$ **norm:**

$$\mathcal{J}_{\mathfrak{C}_1}(\underline{\lambda}_1 \|s - s_{h\tau}\|)^2$$

$$\leq \|s_0 - s_{h\tau}(0)\|^2_{H^{-1}_{\bar{\kappa}}(\Omega)} + \mathcal{J}_{\mathfrak{C}_1}(\bar{\lambda}_1 \|\mathcal{R}(\Psi_{h\tau})\|_{H^{-1}_{\bar{\kappa}}(\Omega)})^2,$$

Estimate in the $L^2(0, T; H^1(\Omega))$ **norm:**

$$\tfrac{1}{2}\mathcal{J}_{\mathfrak{C}_2}(\underline{\lambda}_2 \|\Psi - \Psi_{h\tau}\|_{H^1_{\bar{\kappa}}(\Omega)})^2$$

$$\leq \|s_0 - s_{h\tau}(0)\|^2 + \mathcal{J}_{\mathfrak{C}_2}\left(\eta^{\mathrm{deg}}\right)^2 + 4\,\mathcal{J}_{\mathfrak{C}_2}\left(\bar{\lambda}_2 \|\mathcal{R}(\Psi_{h\tau})\|_{H^{-1}_{\bar{\kappa}}(\Omega)}\right)^2.$$

*Similar expression holds for the $\|\partial_t(\theta(\Psi_1) - \theta(\Psi_2))\|_{L^2(I;H^{-1}_{\bar{\kappa}}(\Omega))}$ error component

Theorem 3 (b)

Estimate in the $L^2(\Omega \times [0,T])$ **norm:**

$$\mathcal{J}_{\mathfrak{C}_1}(\underline{\lambda}_1 \|s - s_{h\tau}\|)^2$$

$$\leq \|s_0 - s_{h\tau}(0)\|^2_{H^{-1}_{\bar{\kappa}}(\Omega)} + \mathcal{J}_{\mathfrak{C}_1}(\bar{\lambda}_1 \|\mathcal{R}(\Psi_{h\tau})\|_{H^{-1}_{\bar{\kappa}}(\Omega)})^2,$$

Estimate in the $L^2(0,T;H^1(\Omega))$ **norm:**

$$\tfrac{1}{2}\mathcal{J}_{\mathfrak{C}_2}(\underline{\lambda}_2 \|\Psi - \Psi_{h\tau}\|_{H^1_{\bar{\kappa}}(\Omega)})^2$$

$$\leq \|s_0 - s_{h\tau}(0)\|^2 + \mathcal{J}_{\mathfrak{C}_2}\left(\eta^{\rm deg}\right)^2 + 4\,\mathcal{J}_{\mathfrak{C}_2}\left(\bar{\lambda}_2 \|\mathcal{R}(\Psi_{h\tau})\|_{H^{-1}_{\bar{\kappa}}(\Omega)}\right)^2.$$

*Similar expression holds for the $\|\partial_t(\theta(\Psi_1) - \theta(\Psi_2))\|_{L^2(I;H^{-1}_{\bar{\kappa}}(\Omega))}$ error component

**In the linear case $\mathfrak{C}_1 = \mathfrak{C}_2 = 0$

5 Finite element solution

- Let $\{t_0 := 0, t_1, \ldots, t_N := T\}$ be the time-discretization, with $\tau_n := t_n - t_{n-1}$, and $I_n := (t_{n-1}, t_n]$

5 Finite element solution

- Let $\{t_0 := 0, t_1, \dots, t_N := T\}$ be the time-discretization, with $\tau_n := t_n - t_{n-1}$, and $I_n := (t_{n-1}, t_n]$
- Let $\{\mathcal{T}_n\}_{n=1}^N$ be a sequence of triangulations and $\{V_{n,h}\}_{n=1}^N$ the corresponding finite element spaces

5 Finite element solution

- Let $\{t_0 := 0, t_1, \ldots, t_N := T\}$ be the time-discretization, with $\tau_n := t_n - t_{n-1}$, and $I_n := (t_{n-1}, t_n]$
- Let $\{\mathcal{T}_n\}_{n=1}^N$ be a sequence of triangulations and $\{V_{n,h}\}_{n=1}^N$ the corresponding finite element spaces
- Let $\{p_{n,h} \in V_{n,h}\}_{n=1}^N$ be the sequence of finite elements solutions for backward Euler time discretization of the pressure formulation, i.e.,

$$\left(\frac{1}{\tau_n}(S(p_{n,h}) - S(p_{n-1,h})), \nabla\varphi_h\right) + (\bar{\mathbf{K}}\kappa(S(p_{n,h}))[\nabla p_{n,h} + \boldsymbol{g}], \varphi_h)$$

$$= (f(S(p_{n,h}), \boldsymbol{x}, t_n), \varphi_h), \qquad \forall \varphi_h \in V_{n,h}.$$

5 Finite element solution

- Let $\{t_0 := 0, t_1, \ldots, t_N := T\}$ be the time-discretization, with $\tau_n := t_n - t_{n-1}$, and $I_n := (t_{n-1}, t_n]$
- Let $\{\mathcal{T}_n\}_{n=1}^N$ be a sequence of triangulations and $\{V_{n,h}\}_{n=1}^N$ the corresponding finite element spaces
- Let $\{p_{n,h} \in V_{n,h}\}_{n=1}^N$ be the sequence of finite elements solutions for backward Euler time discretization of the pressure formulation, i.e.,

$$\left(\frac{1}{\tau_n}(S(p_{n,h}) - S(p_{n-1,h})), \nabla\varphi_h\right) + (\bar{\mathbf{K}}\kappa(S(p_{n,h}))[\nabla p_{n,h} + \boldsymbol{g}], \varphi_h)$$

$$= (f(S(p_{n,h}), \boldsymbol{x}, t_n), \varphi_h), \qquad \forall\varphi_h \in V_{n,h}.$$

- Define the time-discrete total pressure and saturation as

$$\Psi_{n,h} := \mathcal{K}(p_{n,h}) \text{ and } s_{n,h} := S(p_{n,h}) = \theta(\Psi_{n,h}).$$

- Define the time-discrete total pressure and saturation as

$$\Psi_{n,h} := \mathcal{K}(p_{n,h}) \text{ and } s_{n,h} := S(p_{n,h}) = \theta(\Psi_{n,h}).$$

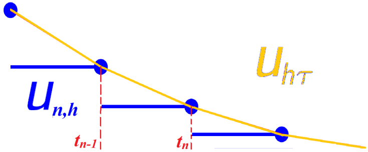
5 Finite element solution

- Define the time-discrete total pressure and saturation as

$$\Psi_{n,h} := \mathcal{K}(p_{n,h}) \text{ and } s_{n,h} := S(p_{n,h}) = \theta(\Psi_{n,h}).$$

- Let $\Psi_{h\tau} \in C(0,T;H^1(\Omega))$ with $s_{h\tau} \in W^{1,\infty}(0,T;L^2(\Omega))$ be their time-continuous interpolations, i.e., they satisfy

$$\Psi_{h\tau}(t_n) = \Psi_{n,h}, \quad s_{h\tau}(t_n) = s_{n,h}.$$



5 Finite element solution

- Define the time-discrete total pressure and saturation as

$$\Psi_{n,h} := \mathcal{K}(p_{n,h}) \text{ and } s_{n,h} := S(p_{n,h}) = \theta(\Psi_{n,h}).$$

- Let $\Psi_{h\tau} \in C(0,T;H^1(\Omega))$ with $s_{h\tau} \in W^{1,\infty}(0,T;L^2(\Omega))$ be their time-continuous interpolations, i.e., they satisfy

$$\Psi_{h\tau}(t_n) = \Psi_{n,h}, \quad s_{h\tau}(t_n) = s_{n,h}.$$

- We introduce the a posteriori estimator $\eta_\Omega : [0,T] \to [0,\infty)$,

$$\|\mathcal{R}(\Psi_{h\tau}(t))\|_{H^{-1}_{\bar{\kappa}}(\Omega)} \leq \eta_\Omega(t), \quad \forall\, t \in [0,T].$$

5 Global reliability

Theorem 4 (a)

Estimate in the $L^2(\Omega \times [0, T])$ **norm:**

$$[\mathcal{E}_{L^2}]^2 := \mathcal{J}_{\mathfrak{E}_1}(\underline{\lambda}_1 \|s - s_{h\tau}\|)^2$$

$$\leq \|s_0 - s_{h\tau}(0)\|^2_{H^{-1}_{\bar{\kappa}}(\Omega)} + \mathcal{J}_{\mathfrak{E}_1}(\bar{\lambda}_1 \eta_{\Omega})^2 =: [\eta_{L^2}]^2,$$

Theorem 4 (a)

Estimate in the $L^2(\Omega \times [0,T])$ **norm:**

$$[\mathcal{E}_{L^2}]^2 := \mathcal{J}_{\mathfrak{C}_1}(\underline{\lambda}_1 \|s - s_{h\tau}\|)^2$$

$$\leq \|s_0 - s_{h\tau}(0)\|^2_{H^{-1}_{\bar{\kappa}}(\Omega)} + \mathcal{J}_{\mathfrak{C}_1}(\bar{\lambda}_1 \eta_\Omega)^2 =: [\eta_{L^2}]^2,$$

Estimate in the $L^2(0,T;H^1(\Omega))$ **norm:**

$$[\mathcal{E}_{H^1}]^2 := \tfrac{1}{2}\mathcal{J}_{\mathfrak{C}_2}(\underline{\lambda}_2 \|\Psi - \Psi_{h\tau}\|_{H^1_{\bar{\kappa}}(\Omega)})^2$$

$$\leq \|s_0 - s_{h\tau}(0)\|^2 + \mathcal{J}_{\mathfrak{C}_2}\left(\eta^{\mathrm{deg}}\right)^2 + 4\,\mathcal{J}_{\mathfrak{C}_2}\left(\bar{\lambda}_2 \eta_\Omega\right)^2 =: [\eta_{H^1}]^2.$$

Theorem 4 (Local lower bounds)

For $n \in \{1, \dots, N\}$, $\omega \subseteq \Omega$ and some $\mathfrak{T}_\omega \subset \Omega$ such that $\omega \subseteq \Omega$,

$$\int_{I_n} \left([\eta_\omega]^2 + \|\Psi_{h\tau} - \Psi_{n,h}\|^2_{H^1_{\bar{\mathbf{K}}}(\omega)} \right)$$

$$\lesssim \mathrm{dist}^{\alpha}_{\mathfrak{T}_\omega, I_n}(\Psi, \Psi_{h\tau})^2 + \left(\substack{\text{Data oscillation, quadrature, \&} \\ \text{temporal discretization estimator}} \right).$$

Theorem 4 (Local lower bounds)

For $n \in \{1, \dots, N\}$, $\omega \subseteq \Omega$ and some $\mathfrak{T}_\omega \subset \Omega$ such that $\omega \subseteq \Omega$,

$$\int_{I_n} \left([\eta_\omega]^2 + \|\Psi_{h\tau} - \Psi_{n,h}\|^2_{H^1_{\bar{\kappa}}(\omega)} \right)$$

$$\lesssim \mathrm{dist}^{\alpha}_{\mathfrak{T}_\omega, I_n}(\Psi, \Psi_{h\tau})^2 + \begin{pmatrix} \text{Data oscillation, quadrature, \&} \\ \text{temporal discretization estimator} \end{pmatrix}.$$

Similar estimate holds for the estimator

$$[\eta^n_{\mathrm{LB}}]^2 := \int_{I_n} ([\eta_\Omega]^2 + \|\Psi_{h\tau} - \Psi_{n,h}\|^2_{H^1_{\bar{\kappa}}(\Omega)})$$

and the global-in-space error $\mathrm{dist}^{\alpha}_{\Omega, I_n}(\Psi_{h\tau}, \Psi)$.

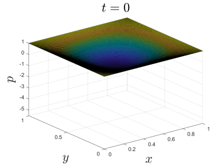
5 Numerical results: non-degenerate case

Solution

$p_{\text{exact}}(x, y, t) = 2 - e^{16(1+t^2)xy(1-x)(1-y)}$ in $(0, 1)^2$

$k(s) = s^3$, $S(p) = \frac{1}{(2-p)^{\frac{1}{3}}}$ (Brooks-Corey type)

$\bar{\boldsymbol{K}} = \mathbb{I}$, $\boldsymbol{g} = -\boldsymbol{e}_x$, $f((x, y), t)$ set accordingly



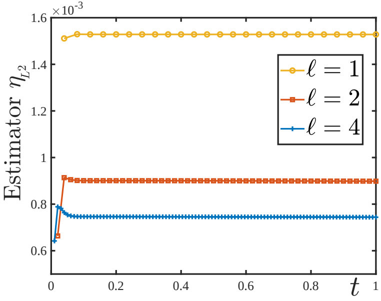
5 Reliability (upper bound) estimates

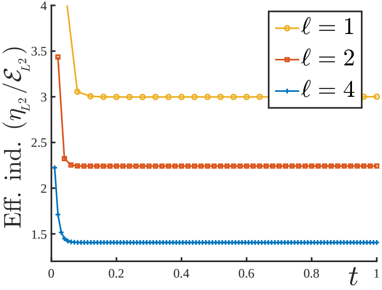
Effectivity

$$\text{Effectivity index} := \text{upper bound}/\text{error}$$

Effectivity

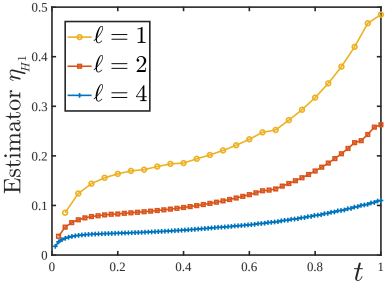
$$\text{Effectivity index} := \text{upper bound}/\text{error} = \eta_{L^2}/\mathcal{E}_{L^2}$$

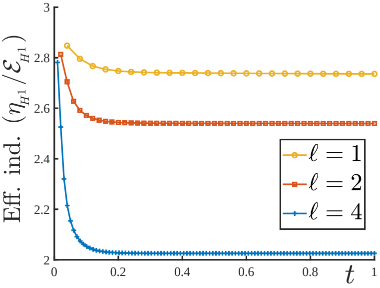




Effectivity

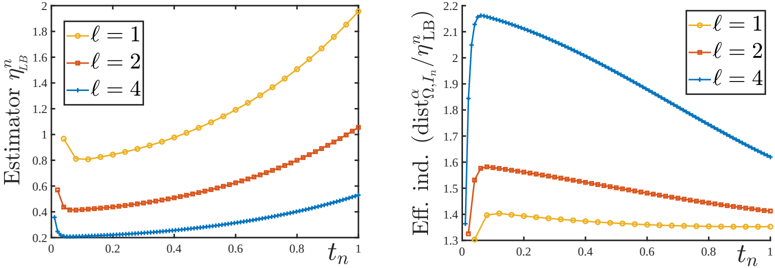
$$\text{Effectivity index} := \text{upper bound}/\text{error} = \eta_{H^1}/\mathcal{E}_{H^1}$$



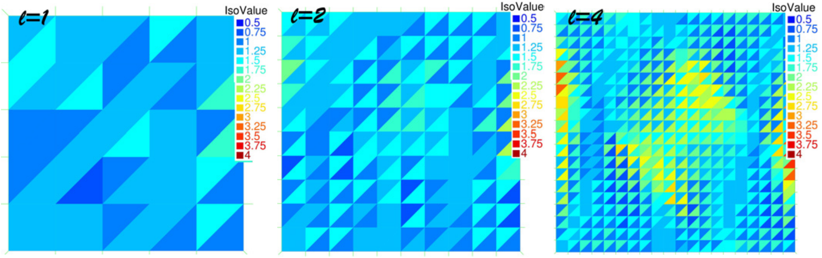


Effectivity

$$\text{Effectivity index} := \text{error/lower bound} = \mathrm{dist}^{\alpha}_{\Omega, I_n}(\Psi, \Psi_{h\tau})/\eta^{n}_{\mathrm{LB}},$$



5 Local efficiency



5 Numerical results: degenerate case

Solution

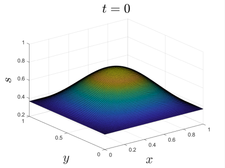
$$\Psi_{\mathrm{exact}}(x, y, t) = 12(1 + t^2)xy(1 - x)(1 - y)$$

$$\theta(\Psi) = \begin{cases} \exp(\Psi - 1) & \text{if } \Psi < 1 \\ 1 & \text{if } \Psi \geq 1 \end{cases}$$

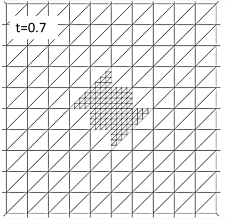
$$k(s) = \begin{cases} s & \text{if } s < 1 \\ 1 & \text{if } s \geq 1 \end{cases}$$

$\bar{\boldsymbol{K}} = \mathbb{I}$, $\boldsymbol{g} = -\boldsymbol{e}_x$

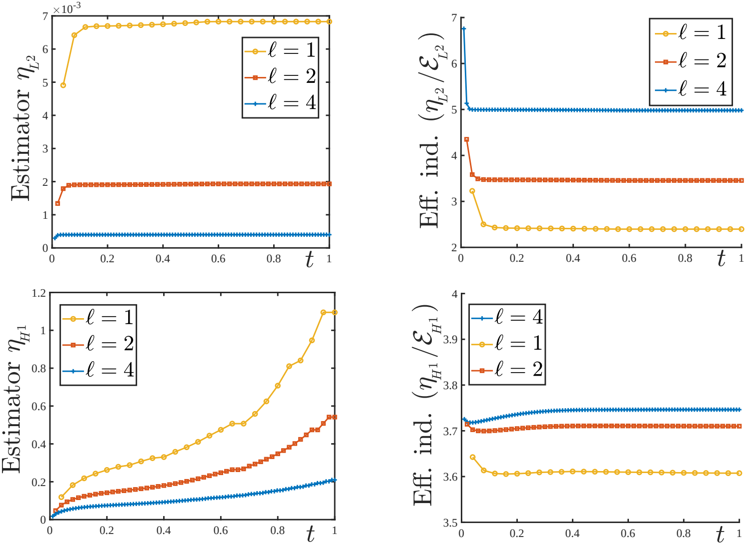
$f(x, y, t)$ set accordingly



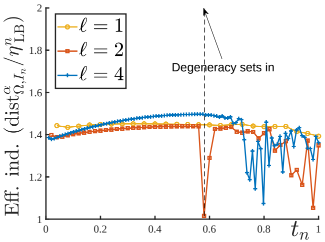
Degenerate domains

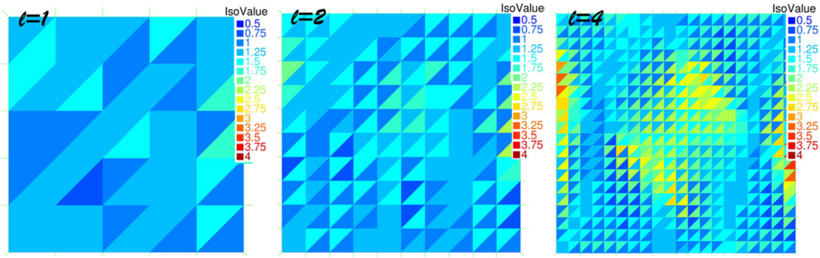


5 Reliability estimates

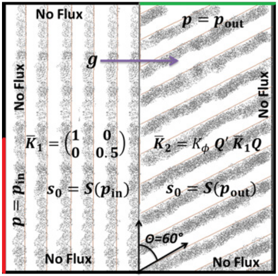


5 Efficiency

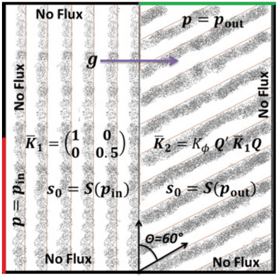


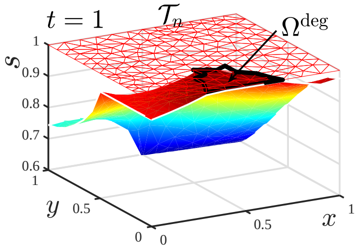


5 Numerical results: realistic case

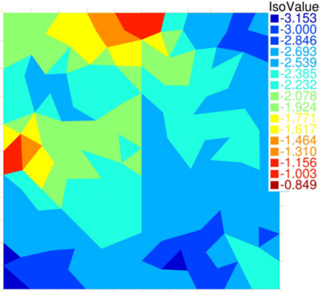


5 Numerical results: realistic case

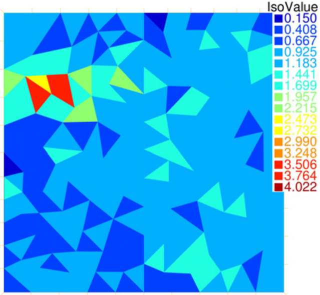




5 Local efficiency



Estimate $\log_{10}([\eta^{\mathrm{F}}_{n,h,K}])$



Effectivity index

K. Mitra, & M. Vohralík. A posteriori error estimates for the Richards equation. arXiv preprint arXiv:2108.12507

5 Thank you for your time