Optimal interplay of adaptive mesh-refinement and iterative solvers for nonlinear elliptic PDEs

slides

Dirk Praetorius

TU WIEN ASC INSTITUTE OF ANALYSIS AND SCIENTIFIC COMPUTING

NumPDEs Workgroup on Numerics of PDEs

FWF Österreichischer Wissenschaftsfonds

What's it all about?

Nonlinear PDE: Discretize → Linearize → Solve

TU WIEN NumPDEs

1. consider: exact solution $u^\star$ to nonlinear PDE
   - strongly monotone & Lipschitz continuous energy minimization problem (scalar nonlinearity)

2. discretize: exact FEM solution $u_\ell^\star \approx u^\star$ solves nonlinear discrete system
   - lowest-order H^1-conforming FEM (continuous piecewise linears)

3. linearize: fixed point iterate $u_\ell^{k,\star} \approx u_\ell^\star$ solves linear discrete system
   - Kačanov iteration (or: Zarantonello iteration, or: damped Newton method)

4. solve: algebraic solver iterate $u_\ell^{k,j} \approx u_\ell^{k,\star}$ is computed
   - geometric multigrid (or: PCG with additive Schwarz preconditioner)

Ern, Vohralík: *SIAM J. Sci. Comput.*, 35 (2013)

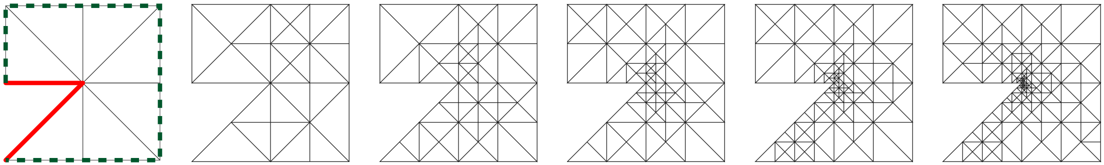
What is optimal interplay?

- overall goal: approximate PDE solution $u^\star$ by FE function u_ℓ within minimal time
- requires: resolve singularities + balance error components, i.e.,
  1. discretize: $u_\ell^\star \approx u^\star$ → control $\eta_\ell(u_\ell^\star) \gtrsim \|u^\star - u_\ell^\star\|$
  2. linearize: $u_\ell^{1,\star}, \dots, u_\ell^{K,\star} \approx u_\ell^\star$ → control $\|u_\ell^\star - u_\ell^{k,\star}\|$
  3. solve: $u_\ell^{k,1}, \dots, u_\ell^{k,J} \approx u_\ell^{k,\star}$ → control $\|u_\ell^{k,\star} - u_\ell^{k,j}\|$
- note: only the $u_\ell^{k,j}$ will be computed!
- adaptive mesh-refinement and iterative solver have optimal interplay, if
  - guaranteed convergence for any choice of parameters
  - rates = complexity for any choice of parameters
  - optimal complexity for appropriate choices of parameters

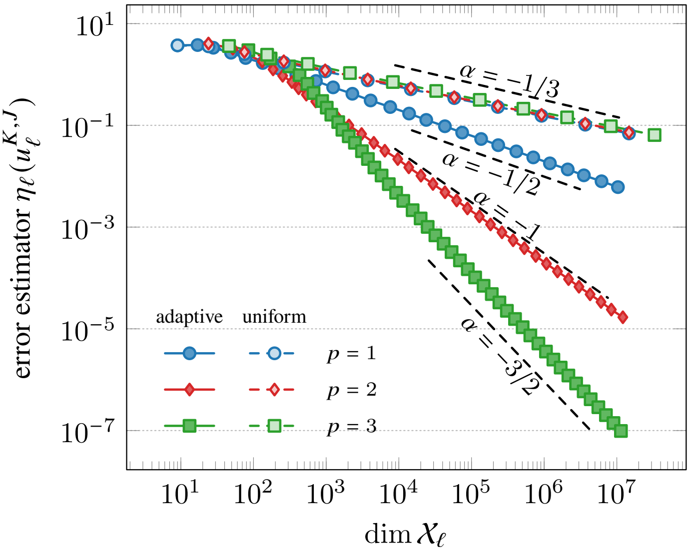
Example: AFEM for 2D model problem

$$\begin{aligned} -\mathrm{div}\big(\mu(|\nabla u^\star|)\,\nabla u^\star\big) &= 1 \quad \text{in } \Omega \\ \mu(|\nabla u^\star|)\,\nabla u^\star \cdot \boldsymbol{n} &= g \quad \text{on } \Gamma_N \\ u^\star &= 0 \quad \text{on } \Gamma_D \end{aligned}$$

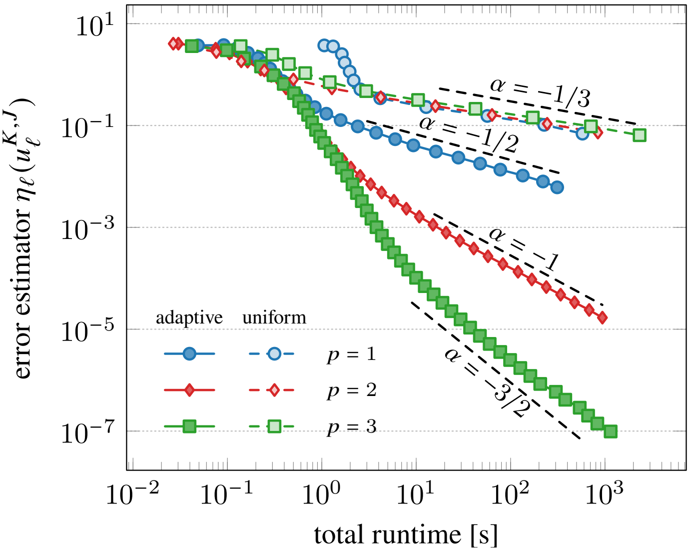
- $\mu(t) := 2 + \dfrac{1}{\sqrt{1+t^2}}$



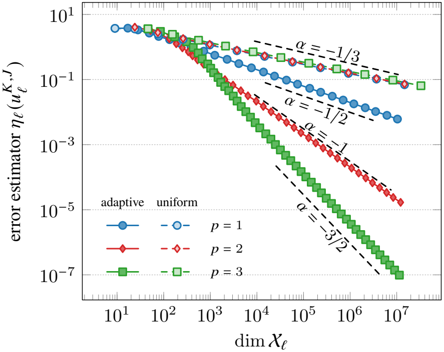
Convergence wrt. dofs

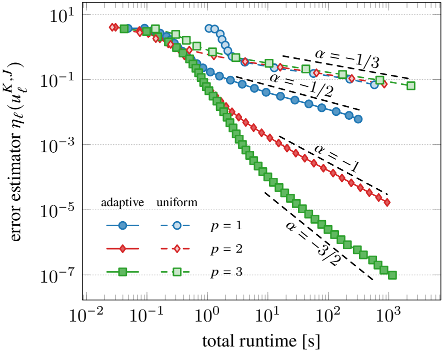


Convergence wrt. time (adaptivity pays off!)



Rates vs. complexity





- clear: $\dim \mathcal{X}_\ell \simeq \#\mathcal{T}_\ell$ for fixed p
- $\mathfrak{R}(\alpha) := \sup_{(\ell,k,j)} (\#\mathcal{T}_\ell)^\alpha \eta_\ell(u_\ell^{k,j})$
- assume: $\mathtt{work}(\mathcal{T}_{\ell'}) \simeq \#\mathcal{T}_{\ell'}$ for all steps
- $\widehat{\mathfrak{R}}(\alpha) := \sup_{(\ell,k,j)} \Big(\sum_{|\ell',k',j'|=0}^{|\ell,k,j|} \#\mathcal{T}_{\ell'} \Big)^\alpha \eta_\ell(u_\ell^{k,j})$

Rates = complexity?!

- $\mathfrak{R}(\alpha) := \sup_{(\ell,k,j)\in\mathcal{Q}} (\#\mathcal{T}_\ell)^\alpha \eta_\ell(u_\ell^{k,j}) < \infty$ describes rate wrt. $\#\mathcal{T}_\ell \simeq \dim \mathcal{X}_\ell$
- $\widehat{\mathfrak{R}}(\alpha) := \sup_{(\ell,k,j)\in\mathcal{Q}} \Big(\sum_{|\ell',k',j'|=0}^{|\ell,k,j|} \#\mathcal{T}_{\ell'} \Big)^\alpha \eta_\ell(u_\ell^{k,j}) < \infty$ describes rate wrt. total cost/runtime

Key argument = R-linear convergence $\eta_n \le C q^{n-m} \eta_m \quad \forall\, 0 \le m \le n$

$$\Longrightarrow \quad \mathfrak{R}(\alpha) := \sup_{n\in\mathbb{N}_0} (\#\mathcal{T}_n)^\alpha \eta_n \le \underbrace{\sup_{n\in\mathbb{N}_0} \Big(\sum_{m=0}^{n} \#\mathcal{T}_m \Big)^\alpha \eta_n}_{=:\,\widehat{\mathfrak{R}}(\alpha)} \le \frac{C}{(1-q^{1/\alpha})^\alpha} \mathfrak{R}(\alpha)$$

Takeaways from intro

1. AFEM analysis should address rates wrt. complexity/time instead of #dofs
2. R-linear convergence is key to relate #dofs and complexity/time (and essentially necessary)
3. needs linear cost of all modules SOLVE – ESTIMATE – MARK – REFINE
   - SOLVE is critical (beyond 1D)
   - ESTIMATE is clear (with idealized quadrature for residual estimator)
   - MARK is known (e.g., Stevenson 2007 with binning, Pfeiler–Praetorius 2020 for minimal $\mathcal{M}_\ell$)
   - REFINE is known (Binev–Dahmen–DeVore '04, Stevenson '08, Diening–Gehring–Storn '25$^+$)

Stevenson: *Found. Comput. Math.*, 7 (2007)

Pfeiler, Praetorius: *Math. Comp.*, 89 (2020)

Binev, Dahmen, DeVore: *Numer. Math.*, 97 (2004)

Stevenson: *Math. Comp.*, 77 (2008)

Diening, Gehring, Storn: *Found. Comput. Math.*, published online first (2025)

AFEM with contractive solver

Contractive solver

TU WIEN NumPDEs

- let $0 < \kappa < 1$

- consider contractive solver for discrete FE systems

$$\|u_\ell^\star - u_\ell^k\| \le \kappa \, \|u_\ell^\star - u_\ell^{k-1}\| \quad \text{for all} \quad k \in \mathbb{N}$$

- nothing but triangle inequality

$$\Longrightarrow \quad \frac{1-\kappa}{\kappa} \, \|u_\ell^\star - u_\ell^k\| \le (1-\kappa) \, \|u_\ell^\star - u_\ell^{k-1}\| \le \|u_\ell^k - u_\ell^{k-1}\| \le (1+\kappa) \, \|u_\ell^\star - u_\ell^{k-1}\|$$

Stopping criterion for contractive solver

- $$\|u^\star - u_\ell^k\| \le \|u^\star - u_\ell^\star\| + \|u_\ell^\star - u_\ell^k\| \overset{\text{reliability (A3)}}{\lesssim} \eta_\ell(u_\ell^\star) + \|u_\ell^\star - u_\ell^k\|$$
  $$\overset{\text{stability (A1)}}{\lesssim} \eta_\ell(u_\ell^k) + \|u_\ell^\star - u_\ell^k\|$$
  $$\overset{\text{solver}}{\lesssim} \eta_\ell(u_\ell^k) + \|u_\ell^k - u_\ell^{k-1}\|$$

- idea: equilibrate $\eta_\ell(u_\ell^k)$ and $\|u_\ell^k - u_\ell^{k-1}\|$

$\Longrightarrow$ stop solver for $K = k$ as soon as $\|u_\ell^K - u_\ell^{K-1}\| \le \lambda\,\eta_\ell(u_\ell^K)$

  - clear: $\|u^\star - u_\ell^K\| \lesssim \eta_\ell(u_\ell^K)$

- nested iteration: $u_{\ell+1}^0 := u_\ell^K$

$\Longrightarrow$ a-posteriori error control for all u_ℓ^k but u_0^0

AFEM with contractive solver

TU WIEN NumPDEs

Input: initial mesh $\mathcal{T}_0$, initial guess u_0^0, adaptivity parameters $0 < \theta \le 1$, $\lambda > 0$

For each $\ell = 0, 1, 2, \dots$ do

- **SOLVE & ESTIMATE:** For $k = 1, 2, 3, \dots, K =: K(\ell)$, **repeat**
  - compute u_ℓ^k and $\|u_\ell^k - u_\ell^{k-1}\|$
  - compute $\eta_\ell(T, u_\ell^k)$ for all $T \in \mathcal{T}_\ell$

  until $\|u_\ell^K - u_\ell^{K-1}\| \le \lambda\, \eta_\ell(u_\ell^K)$

- **MARK:** choose (quasi-) minimal $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ s.t. $\theta \eta_\ell(u_\ell^K)^2 \le \eta_\ell(\mathcal{M}_\ell, u_\ell^K)^2$
- **REFINE:** $\mathcal{T}_{\ell+1} := \mathrm{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$, $u_{\ell+1}^0 := u_\ell^K$,

Output: Discrete solutions u_ℓ^k, corresponding estimator $\eta_\ell(u_\ell^k) + \|u_\ell^k - u_\ell^{k-1}\|$

- note: number of solver steps $K = K(\ell)$ might vary with ℓ

Gantner, Haberl, Praetorius, Schimanko: *Math. Comp.*, 90 (2021)

Model problem & Kačanov linearization

Model problem

Strong formulation with scalar nonlinearity

$$-\mathrm{div}\,\big(\mu(|\nabla u^\star|^2)\nabla u^\star\big) = f \quad \text{in } \Omega$$
$$u^\star = 0 \quad \text{on } \partial\Omega$$

- model example: $M\,(t-s) \le \mu(t^2)t - \mu(s^2)s \le L\,(t-s) \qquad \forall\, 0 \le s \le t$

$\Longrightarrow$ PDE defines operator $\mathcal{A}\colon H^1_0(\Omega) \to H^{-1}(\Omega)$ that is

- strongly monotone $\quad M\,\|u-v\|^2 \le \langle \mathcal{A}u - \mathcal{A}v\,,\, u-v\rangle$
- Lipschitz continuous $\quad \langle \mathcal{A}u - \mathcal{A}v\,,\, w\rangle \le L\,\|u-v\|\,\|w\|$

$\Longrightarrow$ existence and uniqueness of weak solution $u^\star \in H^1_0(\Omega)$

PDE inherent energy structure

- energy functional $\mathcal{J}(v) := \frac{1}{2}\int_\Omega \int_0^{|\nabla v(x)|^2} \mu(t)\,\mathrm{d}t\,\mathrm{d}x - \langle f\,,\,v\rangle_{L^2(\Omega)}$

Lemma (e.g., Zeidler '90)

- $\frac{M}{2}\,\|u^\star - v\|^2 \le \mathcal{J}(v) - \mathcal{J}(u^\star) \le \frac{L}{2}\,\|u^\star - v\|^2 \qquad \forall\, v \in H^1_0(\Omega)$
- i.e., $u^\star$ is unique minimizer of $\mathcal{J}$ over $H^1_0(\Omega)$

- energy distance $\mathbf{d}(u,v) := \mathcal{J}(v) - \mathcal{J}(u)$ replaces energy norm $\|\cdot\|$ above

Galerkin discretization

Weak formulation

- find $u^\star \in H^1_0(\Omega)$ s.t. $\quad \langle \mu(|\nabla u^\star|^2)\nabla u^\star \,,\, \nabla v\rangle_{L^2(\Omega)} = \langle f \,,\, v\rangle_{L^2(\Omega)} \quad \forall\, v \in H^1_0(\Omega)$

Discrete formulation

- find $u_\ell^\star \in \mathcal{X}_\ell$ s.t. $\quad \langle \mu(|\nabla u_\ell^\star|^2)\nabla u_\ell^\star \,,\, \nabla v_\ell\rangle_{L^2(\Omega)} = \langle f, v_\ell\rangle_{L^2(\Omega)} \quad \forall\, v_\ell \in \mathcal{X}_\ell$

- all analytical properties transfer from $H^1_0(\Omega)$ to $\mathcal{X}_\ell$
- clear: existence and uniqueness of $u_\ell^\star$ and equivalence to energy minimization over $\mathcal{X}_\ell$

- but: corresponds to nonlinear discrete system so that $u_\ell^\star$ can hardly be computed exactly

Garau, Morin, Zuppa: *Numer. Math: Theory, Meth. Appl.*, 5 (2012)

Linearized Galerkin discretization

Discrete formulation

- find $u_\ell^\star \in \mathcal{X}_\ell$ s.t. $\langle \mu(|\nabla u_\ell^\star|^2)\nabla u_\ell^\star \, , \, \nabla v_\ell \rangle_{L^2(\Omega)} = \langle f \, , \, v_\ell \rangle_{L^2(\Omega)} \quad \forall \, v_\ell \in \mathcal{X}_\ell$

Kačanov linearization (linearized discrete formulation)

- given $u_\ell^k \in \mathcal{X}_\ell$, find $u_\ell^{k+1,\star} \in \mathcal{X}_\ell$ s.t. $\langle \mu(|\nabla u_\ell^k|^2)\nabla u_\ell^{k+1,\star} \, , \nabla v_\ell \rangle_{L^2(\Omega)} = \langle f \, , \, v_\ell \rangle_{L^2(\Omega)} \, \forall \, v_\ell \in \mathcal{X}_\ell$

- corresponds to linear discrete SPD system (i.e., Laplace-type PDE)

- recall: $M\,(t-s) \le \mu(t^2)t - \mu(s^2)s \le L\,(t-s) \quad \forall\, 0 \le s \le t$

Heid, Wihler: *Math. Comp.*, 89 (2020)

Convergence of exact Kačanov linearization

TU WIEN NumPDEs

- given $u_\ell^k \in \mathcal{X}_\ell$, find $u_\ell^{k+1,\star} \in \mathcal{X}_\ell$ s.t. $\langle \mu(|\nabla u_\ell^k|^2)\nabla u_\ell^{k+1,\star}, \nabla v_\ell \rangle_{L^2(\Omega)} = \langle f, v_\ell \rangle_{L^2(\Omega)} \; \forall v_\ell \in \mathcal{X}_\ell$

Proposition (HPW'21)

- exists $0 < q < 1$ s.t. $0 \le \mathbf{d}(u_\ell^\star, u_\ell^{k+1,\star}) \le q\,\mathbf{d}(u_\ell^\star, u_\ell^k)$

1. $\|u_\ell^{k+1,\star} - u_\ell^k\|^2 \le \frac{2}{M}\,\mathbf{d}(u_\ell^{k+1,\star}, u_\ell^k)$ — core estimate [HW'20]
2. $\|u_\ell^\star - u_\ell^k\| \le \frac{L}{M}\,\|u_\ell^{k+1,\star} - u_\ell^k\|$ — [HW'20]
3. $\mathbf{d}(u_\ell^\star, u_\ell^k) \le \frac{L}{2}\,\|u_\ell^\star - u_\ell^k\|^2$ — energy equivalence

$$\Longrightarrow \quad 0 \le \mathbf{d}(u_\ell^\star, u_\ell^{k+1,\star}) \le \Big[1 - \frac{M^3}{L^3}\Big]\,\mathbf{d}(u_\ell^\star, u_\ell^k)$$

Heid, Wihler: *Math. Comp.*, 89 (2020)

Heid, Praetorius, Wihler: *Comput. Methods Appl. Math.*, 21 (2021)

AFEM with linearization & algebraic solver

Contractive algebraic solver

Linearized discrete formulation

- given $u_\ell^{k,J} \in \mathcal{X}_\ell$, find $u_\ell^{k+1,\star} \in \mathcal{X}_\ell$ s.t. $\langle \mu(|\nabla u_\ell^{k,J}|^2)\nabla u_\ell^{k+1,\star}, \nabla v_\ell \rangle_{L^2(\Omega)} = \langle f, v_\ell \rangle_{L^2(\Omega)} \; \forall v_\ell \in \mathcal{X}_\ell$

- compute algebraic iterates $u_\ell^{k+1,j} \approx u_\ell^{k+1,\star}$ via contractive solver (for linear PDE)
- important properties:
  1. h-robustness: contraction constant $0 < \kappa < 1$ is $\mathcal{T}_\ell$-independent
  2. linear cost: cost $\mathcal{O}(\#\mathcal{T}_\ell)$ per solver step (also $\mathcal{O}(\sum_{\ell'=0}^{\ell} \#\mathcal{T}_{\ell'})$ is OK)

Chen, Nochetto, Xu: *Numer. Math.*, 120 (2012)

Wu, Zheng: *Appl. Numer. Math.*, 113 (2017)

Innerberger, Miraçi, Praetorius, Streitberger: *ESAIM Math. Model. Numer. Anal.*, 58 (2024)

AFEM algorithm: discretize → linearize → solve

Input: $\mathcal{T}_0$, $u_0^{0,0} := 0$, $0 < \theta \le 1$, $\lambda_{\rm lin} > 0$

For each $\ell = 0, 1, 2, \dots,$ do

- **SOLVE & ESTIMATE:** For $k = 1, 2, 3, \dots, K =: K(\ell)$, **repeat**
  - For $j = 1, \dots, J =: J(\ell, k)$, **repeat**
    - compute $u_\ell^{k,j}$ using multigrid solver
    - compute $\eta_\ell(T, u_\ell^{k,j})$ for all $T \in \mathcal{T}_\ell$
  - **until** $\| u_\ell^{k,j} - u_\ell^{k,j-1} \| \le$ **sufficiently small**
- **until** $\mathbf{d}(u_\ell^{K,J}, u_\ell^{K-1,J})^{1/2} \le \lambda_{\rm lin}\, \eta_\ell(u_\ell^{K,J})$
- **MARK:** choose minimal $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ such that $\theta \sum_{T \in \mathcal{T}_\ell} \eta_\ell(T, u_\ell^{K,J})^2 \le \sum_{T \in \mathcal{M}_\ell} \eta_\ell(T, u_\ell^{K,J})^2$
- **REFINE:** $\mathcal{T}_{\ell+1} := \texttt{refine}(\mathcal{T}_\ell, \mathcal{M}_\ell)$, $u_{\ell+1}^{0,0} := u_\ell^{K,J}$

Output: discrete solutions $u_\ell^{k,j}$ for all $(\ell, k, j) \in \mathcal{Q} \subset \mathbb{N}_0^3$

Termination of algebraic solver 1/4

Equilibration criterion [HPSV'21] $\mapsto$ Zarantonello linearization

- stop algebraic solver if $\quad \|u_\ell^{k,J} - u_\ell^{k,J-1}\| \le \lambda_{\rm alg}\,[\eta_\ell(u_\ell^{k,J}) + \|u_\ell^{k,J} - u_\ell^{k-1,J}\|\,]$
- stop linearization if $\quad \|u_\ell^{K,J} - u_\ell^{K-1,J}\| \le \lambda_{\rm lin}\,\eta_\ell(u_\ell^{K,J})$

$\Longrightarrow$ full R-linear convergence for sufficiently small $\lambda_{\rm lin}, \lambda_{\rm alg}$ with respect to θ

Kačanov: Given $u_\ell^k \in \mathcal{X}_\ell$, find $u_\ell^{k+1,\star} \in \mathcal{X}_\ell$ such that $\forall\, v_\ell \in \mathcal{X}_\ell$

- $\langle \mu(|\nabla u_\ell^k|^2)\nabla u_\ell^{k+1,\star}\,, \nabla v_\ell\rangle_{L^2(\Omega)} = \langle f\,,\, v_\ell\rangle_{L^2(\Omega)}$

Zarantonello: Given $u_\ell^k \in \mathcal{X}_\ell$, find $u_\ell^{k+1,\star} \in \mathcal{X}_\ell$ such that $\forall\, v_\ell \in \mathcal{X}_\ell$

- $\langle \nabla u_\ell^{k+1,\star}\,, \nabla v_\ell\rangle_{L^2(\Omega)} = \langle \nabla u_\ell^k\,, \nabla v_\ell\rangle_{L^2(\Omega)} + \delta\big[\langle f\,,\, v_\ell\rangle_{L^2(\Omega)} + \langle \mu(|\nabla u_\ell^k|^2)\nabla u_\ell^k\,, \nabla v_\ell\rangle_{L^2(\Omega)}\big]$

Ern, Vohralík: *SIAM J. Sci. Comput.*, 35 (2013)

Haberl, Praetorius, Schimanko, Vohralík: *Numer. Math.*, 147 (2021)

Termination of algebraic solver 2/4

TU WIEN NumPDEs

Equilibration criterion [HPSV'21] $\mapsto$ Zarantonello linearization

- stop algebraic solver if $\|u_\ell^{k,J} - u_\ell^{k,J-1}\| \le \lambda_{\rm alg}\,[\eta_\ell(u_\ell^{k,J}) + \|u_\ell^{k,J} - u_\ell^{k-1,J}\|\,]$
- stop linearization if $\|u_\ell^{K,J} - u_\ell^{K-1,J}\| \le \lambda_{\rm lin}\,\eta_\ell(u_\ell^{K,J})$

$\Longrightarrow$ full R-linear convergence for sufficiently small $\lambda_{\rm lin}, \lambda_{\rm alg}$ with respect to θ

Equilibration criterion [BFMPS'25] $\mapsto$ Zarantonello linearization

- stop algebraic solver if $\|u_\ell^{k,J} - u_\ell^{k,J-1}\| \le \lambda_{\rm alg}\,[\lambda_{\rm lin}\,\eta_\ell(u_\ell^{k,J}) + \|u_\ell^{k,J} - u_\ell^{k-1,J}\|\,]$
- stop linearization if $\|u_\ell^{K,J} - u_\ell^{K-1,J}\| \le \lambda_{\rm lin}\,\eta_\ell(u_\ell^{K,J})$

$\Longrightarrow$ full R-linear convergence for arbitrary $\lambda_{\rm lin}$, yet sufficiently small $\lambda_{\rm alg}$ and $\lambda_{\rm lin}\lambda_{\rm alg}$

Haberl, Praetorius, Schimanko, Vohralík: *Numer. Math.*, 147 (2021)

Bringmann, Feischl, Miraçi, Praetorius, Streitberger: *Comput. Math. Appl.*, 180 (2025)

Equilibration criterion [HPSV'21], [BFMPS'25] $\mapsto$ Zarantonello linearization

- stop algebraic solver if $\|u_\ell^{k,J} - u_\ell^{k,J-1}\| \le \lambda_{\rm alg}\,[\lambda_{\rm lin}\,\eta_\ell(u_\ell^{k,J}) + \|u_\ell^{k,J} - u_\ell^{k-1,J}\|\,]$
- stop linearization if $\|u_\ell^{K,J} - u_\ell^{K-1,J}\| \le \lambda_{\rm lin}\,\eta_\ell(u_\ell^{K,J})$

$\Longrightarrow$ full R-linear convergence for sufficiently small $\lambda_{\rm lin}, \lambda_{\rm alg}$

Energy-based criterion [MPS'25$^+$] $\mapsto$ Kačanov, Zarantonello, damped Newton

- stop algebraic solver if $\widetilde{C}_{\rm nrg}\|u_\ell^{k,J} - u_\ell^{k-1,J}\|^2 \le \mathbf{d}(u_\ell^{k,J}, u_\ell^{k-1,J})$ parameter free?!
- stop linearization if $\mathbf{d}(u_\ell^{K,J}, u_\ell^{K-1,J})^{1/2} \le \lambda_{\rm lin}\,\eta_\ell(u_\ell^{K,J})$

$\Longrightarrow$ unconditional full R-linear convergence for arbitrary $\lambda_{\rm lin}$

Haberl, Praetorius, Schimanko, Vohralík: *Numer. Math.*, 147 (2021)

Bringmann, Feischl, Miraçi, Praetorius, Streitberger: *Comput. Math. Appl.*, 180 (2025)

Miraçi, Praetorius, Streitberger: *Math. Comp.*, published online first (2025)

Termination of algebraic solver 4/4

- known: $C_{\rm nrg}\,\|u_\ell^{k,\star} - u_\ell^{k-1,J}\|^2 \le \mathbf{d}(u_\ell^{k,\star}, u_\ell^{k-1,J})$, but: $C_{\rm nrg}$ unknown

Input: $J_{\max} \in \mathbb{N}$, $C_{\min} > 0$, $0 < \varrho < 1$, $u_\ell^{k,0} := u_\ell^{k-1,J}$

- For each $j = 1, 2, 3, \dots, J =: J(\ell, k)$, **repeat**
  1. Compute $u_\ell^{k,j}$ from $u_\ell^{k,j-1}$ by one step of multigrid
  2. Compute $C_\ell^{k,j} := \mathbf{d}(u_\ell^{k,j}, u_\ell^{k-1,J}) / \|u_\ell^{k,j} - u_\ell^{k-1,J}\|^2$

  until $C_\ell^{k,j} \ge C_{\min}$ **or** $u_\ell^{k,j} = u_\ell^{k-1,J}$ **or** $\big[C_\ell^{k,j} > 0$ **and** $j > J_{\max}\big]$.
- **If** $J > J_{\max}$, **then** update $J_{\max} := j$, $C_{\min} \leftarrow \varrho C_{\min}$

Core observation from [MPS'25$^+$]

- $J_{\max}$ and $C_{\min}$ are updated only finitely often
- yields uniform energy contraction $\mathbf{d}(u_\ell^\star, u_\ell^{k,J}) \le [1 - c\, C_{\min}]\, \mathbf{d}(u_\ell^\star, u_\ell^{k-1,J})$

Main result 1: Unconditional full R-linear conv.

TU WIEN NumPDEs

Theorem (MPS'25$^+$)

- employ P1-FEM $\mathcal{X}_\ell := \mathcal{P}^1(\mathcal{T}_\ell) \cap H^1_0(\Omega)$
- arbitrary $u_0^{0,0}$, arbitrary $0 < \theta \le 1$, arbitrary $\lambda_{\mathrm{lin}} > 0$
- $\mathcal{Q} := \big\{(\ell,k,j) \in \mathbb{N}_0^3 \,:\, u_\ell^{k,j}$ is computed by adaptive algorithm$\big\}$
- step counter $|\ell,k,j| := \#\big\{(\ell',k',j') \in \mathcal{Q} \,:\, u_{\ell'}^{k',j'}$ computed earlier than $u_\ell^{k,j}\big\}$
- quasi-error $\mathsf{H}_\ell^{k,j} := \eta_\ell(u_\ell^\star) + |||u_\ell^\star - u_\ell^{k,\star}||| + |||u_\ell^{k,\star} - u_\ell^{k,j}|||$

$\Longrightarrow$ exist $C > 0$ and $0 < q < 1$ such that

$$\mathsf{H}_\ell^{k,j} \le C q^{|\ell,k,j| - |\ell',k',j'|} \mathsf{H}_{\ell'}^{k',j'} \quad \forall\, (\ell,k,j), (\ell',k',j') \in \mathcal{Q} \text{ with } |\ell',k',j'| \le |\ell,k,j|$$

- proof by tail summability $\displaystyle\sum_{|\ell',k',j'| > |\ell,k,j|} \mathsf{H}_{\ell'}^{k',j'} \le C_{\mathrm{tail}}\, \mathsf{H}_\ell^{k,j}$

Miraçi, Praetorius, Streitberger: *Math. Comp.*, published online first (2025)

Main result 2: Rates = complexity

Corollary (GHPS'21, BFMPS'25)

- quasi-error $\mathsf{H}_\ell^{k,j} := \eta_\ell(u_\ell^\star) + |||u_\ell^\star - u_\ell^{k,\star}||| + |||u_\ell^{k,\star} - u_\ell^{k,j}|||$
- assumptions for full R-linear convergence of $\mathsf{H}_\ell^{k,j}$
- $\alpha > 0$

$$\Longrightarrow \quad \sup_{(\ell,k,j)\in\mathcal{Q}} (\#\mathcal{T}_\ell)^\alpha\, \mathsf{H}_\ell^{k,j} \simeq \sup_{(\ell,k,j)\in\mathcal{Q}} \Big(\sum_{\substack{(\ell',k',j')\in\mathcal{Q} \\ |\ell',k',j'|\le|\ell,k,j|}} \#\mathcal{T}_{\ell'} \Big)^\alpha \mathsf{H}_\ell^{k,j}$$

- direct consequence of geometric series (cf. intro)

Gantner, Haberl, Praetorius, Schimanko: *Math. Comp.*, 90 (2021)

Bringmann, Feischl, Miraçi, Praetorius, Streitberger: *Comput. Math. Appl.*, 180 (2025)

Main result 3: Optimal complexity

Theorem (HPSV'21; MPS'25$^+$)

- employ P1-FEM $\mathcal{X}_\ell := \mathcal{P}^1(\mathcal{T}_\ell) \cap H^1_0(\Omega)$
- arbitrary $\alpha > 0$ with $\|u^\star\|_{\mathbb{A}_\alpha} := \sup_{N \geq \#\mathcal{T}_0} N^\alpha \Big(\min_{\#\mathcal{T}_{\rm opt} \leq N} \eta_{\rm opt}(u^\star_{\rm opt}) \Big) < \infty$
- quasi-error $\mathsf{H}_\ell^{k,j} := \eta_\ell(u_\ell^\star) + \|u_\ell^\star - u_\ell^{k,\star}\| + \|u_\ell^{k,\star} - u_\ell^{k,j}\|$
- sufficiently small $0 < \theta < 1$ and sufficiently small $\lambda > 0$ such that

$$0 < \lambda < \lambda_{\rm opt}, \qquad 0 < \Big(\frac{\theta^{1/2} + \lambda/\lambda_{\rm opt}}{1 - \lambda/\lambda_{\rm opt}} \Big)^2 =: \widetilde{\theta} < \theta_{\rm opt} := \frac{1}{1 + C_{\rm stab}^2 C_{\rm rel}^2}$$

$$\Longrightarrow \quad c_{\rm opt} \, \|u^\star\|_{\mathbb{A}_\alpha} \leq \sup_{(\ell,k,j) \in \mathcal{Q}} \Big(\sum_{\substack{(\ell',k',j') \in \mathcal{Q} \\ |\ell',k',j'| \leq |\ell,k,j|}} \#\mathcal{T}_{\ell'} \Big)^\alpha \mathsf{H}_\ell^{k,j} \leq C_{\rm opt} \, \max \big\{ \|u^\star\|_{\mathbb{A}_\alpha} , \mathsf{H}_0^{0,0} \big\}$$

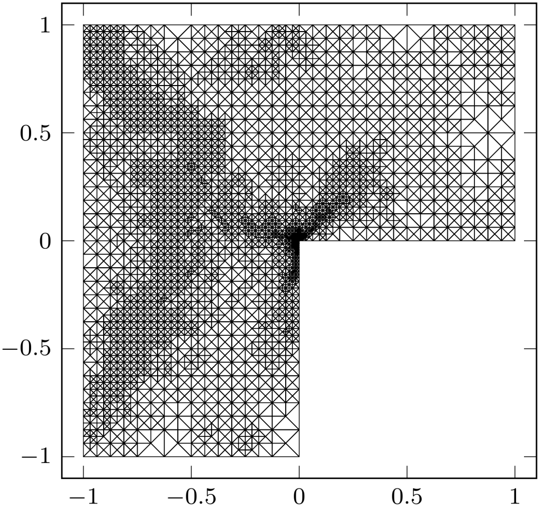
Haberl, Praetorius, Schimanko, Vohralík: *Numer. Math.*, 147 (2021)

Miraçi, Praetorius, Streitberger: *Math. Comp.*, published online first (2025)

Numerical experiment

Setting

- nonlinearity $\mu(t) = 1 + e^{-t}$
  - $M = 1 - 2\exp(-3/2)$
  - $L = 6$
- manufactured solution $u \in H_0^1(\Omega)$
  - $u(x) \simeq r^{2/3} \cos(2\varphi/3)$



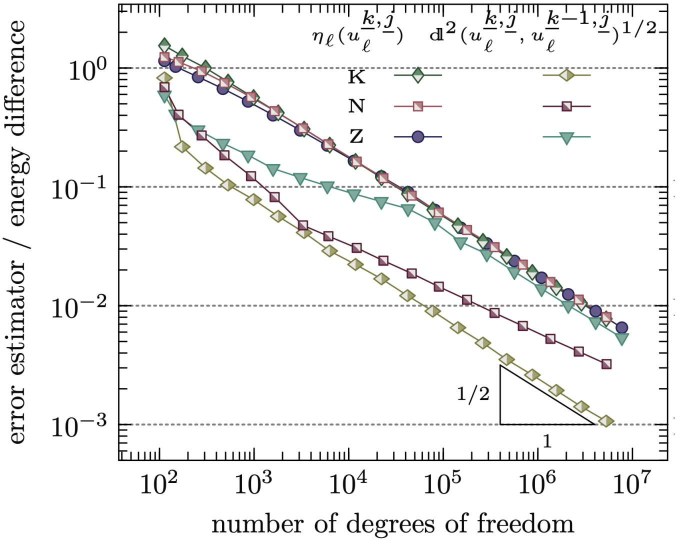
Heid, Praetorius, Wihler: *Comput. Methods Appl. Math.*, 21 (2021)

Parameter study ($p = 1$)

λ_{lin} \ θ	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	0.348	0.248	0.209	0.203	0.210	0.207	0.213	0.252	0.320
0.2	0.371	0.247	0.209	0.205	0.194	0.200	0.220	0.231	0.334
0.3	0.348	0.247	0.209	0.189	0.200	0.202	0.220	0.230	0.327
0.4	0.349	0.248	0.209	0.188	0.193	0.190	0.220	0.229	0.324
0.5	0.348	0.247	0.209	0.189	0.205	0.201	0.202	0.249	0.326
0.6	0.349	0.247	0.209	0.188	0.184	0.193	0.225	0.233	0.303
0.7	0.348	0.247	0.209	0.191	0.194	0.198	0.205	0.248	0.321
0.8	0.348	0.247	0.209	0.187	0.172	0.179	0.205	0.232	0.324
0.9	0.347	0.246	0.208	0.190	0.174	0.190	0.220	0.232	0.304

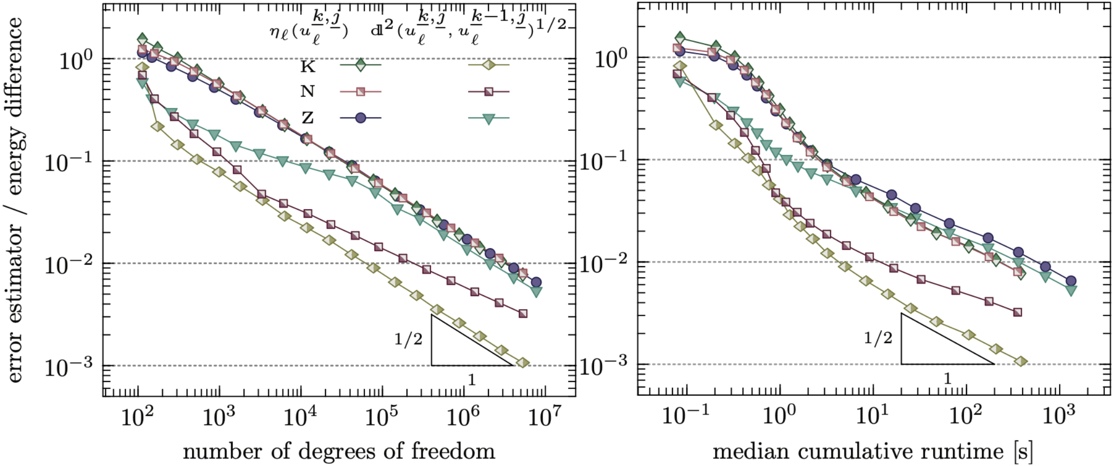
- target: $\eta_\ell(u_\ell^{K,J}) < 10^{-2}$
- show: $\eta_\ell(u_\ell^{K,J}) \times \texttt{time}(u_\ell^{K,J})$

Convergence ($p = 1, \quad \theta = 0.5, \quad \lambda = 0.7$)

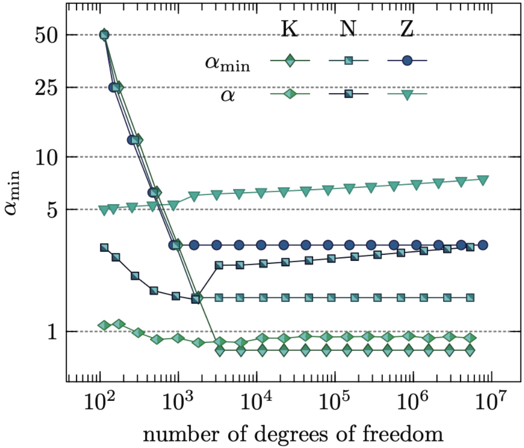


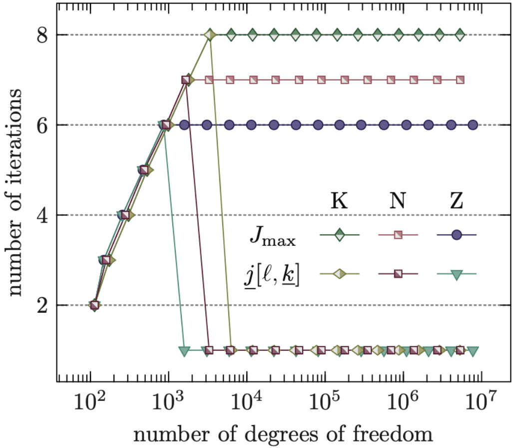
Convergence wrt. time ($p = 1, \quad \theta = 0.5, \quad \lambda = 0.7$)

TU WIEN NumPDEs

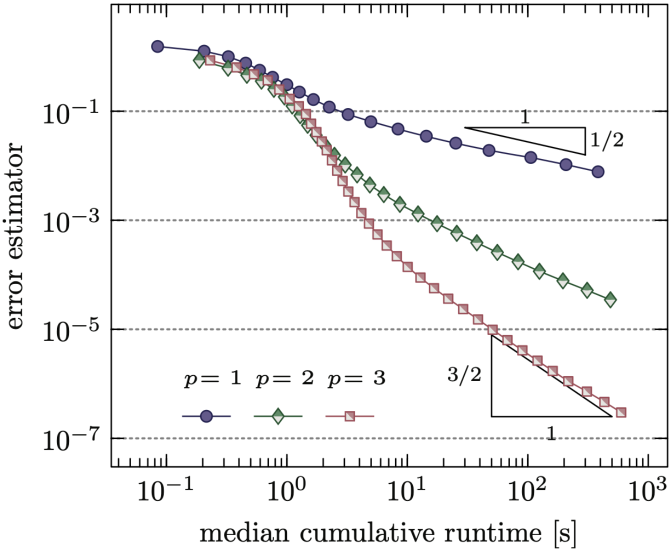


Convergence of algebra ($p = 1, \quad \theta = 0.5, \quad \lambda = 0.7$)





Convergence wrt. time ($p = 1, 2, 3$)



Conclusion

1. AFEM for nonlinear PDEs with discretize ➔ linearize ➔ solve
   - analysis only for lowest-order FEM $p = 1$
   - numerics works also for higher-order FEM $p \in \{1, 2, 3\}$
2. so far, only simple nonlinear problems
   - scalar nonlinearity ($\rightsquigarrow$ energy minimization problems)
   - strongly monotone + globally Lipschitz continuous ($\rightsquigarrow$ excludes p-Laplacian)
3. unconditional: full R-linear convergence of quasi-error $\mathrm{H}_\ell^{k,j}$ ($\Rightarrow$ rates = complexity)
4. perturbation analysis: optimal rates/complexity as for AFEM with exact solver
5. certain extensions (further linearization strategies, semilinear PDEs, goal-oriented AFEM)

Outlook

- understood: adaptive damping for Zarantonello and Newton linearization
  - open: how to deal with inexact algebra?
- understood: use linearization-adapted norms, e.g., Zarantonello
  - use $\langle \mu_\ell(|\nabla u_\ell^{k,J}|^2)\nabla u_\ell^{k+1,\star}, \nabla v_\ell\rangle_{L^2(\Omega)}$ instead of $\langle \nabla u_\ell^{k+1,\star}, \nabla v_\ell\rangle_{L^2(\Omega)}$
- open: use better estimators, e.g., equilibrated fluxes instead of standard residual estimator
  - understood: for linear problems
- open: how far can the analysis be robust w.r.t. L/M?
  - known: robust a-posteriori error estimation strategies for this problem class exist

Thank you for your attention!

TU WIEN NumPDEs

Bringmann, Feischl, Miraçi, Praetorius, Streitberger
On full linear convergence and optimal complexity of adaptive FEM with inexact solver
Comput. Math. Appl., 180 (2025)

Miraçi, Praetorius, Streitberger
Parameter-robust full linear convergence and optimal complexity of adaptive iteratively linearized FEM for nonlinear PDEs
Math. Comp., published online first (2025)

slides

Calculus

Axioms of adaptivity

- $\eta_H(v_H)$ standard residual error estimator, evaluated at $v_H \in \mathcal{X}_H$
  - $\eta_H(T, v_H)^2 := h_T^2 \|f + \operatorname{div}(\boldsymbol{A}\nabla v_H)\|_{L^2(T)}^2 + h_T \|[\![(\boldsymbol{A}\nabla v_H)\cdot \boldsymbol{n}]\!]\|_{L^2(\partial T\cap\Omega)}^2, \quad h_T := |T|^{1/d}$

$\forall \mathcal{T}_H \in \mathbb{T} \quad \forall \mathcal{T}_h \in \mathbb{T}(\mathcal{T}_H) \quad \forall \mathcal{U}_H \subseteq \mathcal{T}_H \cap \mathcal{T}_h \quad \forall v_H \in \mathcal{X}_H \quad \forall v_h \in \mathcal{X}_h$

(A1) $|\eta_h(\mathcal{U}_H, v_h) - \eta_H(\mathcal{U}_H, v_H)| \le C_{\rm stab}\, |\!|\!| v_h - v_H |\!|\!|$ stability

(A2) $\eta_h(\mathcal{T}_h \backslash \mathcal{T}_H, v_H) \le q_{\rm red}\, \eta_H(\mathcal{T}_H \backslash \mathcal{T}_h, v_H)$ (simplified) reduction

(A3) $|\!|\!| u_h^\star - u_H^\star |\!|\!| \le C_{\rm rel}\, \eta_H(\mathcal{T}_H \backslash \mathcal{T}_h, u_H^\star)$ discrete reliability

- $0 < q_{\rm red} = (1/2)^{1/(2d)} < 1$ for newest vertex bisection
- $C_{\rm rel}, C_{\rm stab} > 0$ depend only on $\boldsymbol{A}$, shape regularity, and polynomial degree

Carstensen, Feischl, Page, Praetorius: *Comput. Math. Appl.*, 67 (2014)

R-linear convergence via tail summability

TU WIEN NumPDEs

Proposition (CFPP'14; follows by elementary calculus)

For any sequence $a_\ell \in \mathbb{R}_{\ge 0}$, there holds equivalence:

1. $\exists C > 0 \, \exists 0 < q < 1 \, \forall 0 \le \ell' \le \ell : \qquad a_\ell \le C q^{\ell-\ell'} a_{\ell'}$ R-linear convergence

2. $\exists C' > 0 \, \forall \ell \ge 0 : \qquad \sum_{\ell'=\ell}^{\infty} a_{\ell'} \le C' \, a_\ell$ tail summability

- key = estimator reduction: $\eta_{\ell+1}(u_{\ell+1}^\star) \le q_\theta \, \eta_\ell(u_\ell^\star) + C_{\rm stab} \| u_{\ell+1}^\star - u_\ell^\star \|$
- remains: proof of tail summability
- requires: quasi-orthogonality (A4) provided by problem at hand

Carstensen, Feischl, Page, Praetorius: *Comput. Math. Appl.*, 67 (2014)

Tail summability criterion

Lemma (Feischl '22, BFMPS'25; follows by elementary calculus)

- $C > 0, \quad 0 < q < 1, \quad 0 < \delta \leq 1$
- $a_\ell, b_\ell \in \mathbb{R}_{\geq 0}$ sequences such that
  - $\forall \ell, n \in \mathbb{N}_0 : \quad \underbrace{a_{\ell+1} \leq q\, a_\ell + b_\ell}_{\text{estimator reduction}}, \quad \underbrace{\sum_{\ell'=\ell}^{\ell+n} b_{\ell'}^2 \leq C(n+1)^{1-\delta} a_\ell^2}_{\text{quasi-orthogonality (A4)}}$

$\Longrightarrow \quad \exists C' > 0 \, \forall\, \ell \geq 0 : \quad \sum_{\ell'=\ell}^{\infty} a_{\ell'} \leq C'\, a_\ell$ tail summability

- estimator reduction: $a_\ell := \eta_\ell(u_\ell^\star), \quad b_\ell := C_{\rm stab}\, \| u_{\ell+1}^\star - u_\ell^\star \|$
- SPD setting (or energy minimization): $\sum_{\ell'=\ell}^{\ell+n} b_{\ell'}^2 \leq C_{\rm stab}^2 \, \| u^\star - u_\ell^\star \|^2 \leq C_{\rm stab}^2 C_{\rm rel}^2 \eta_\ell(u_\ell^\star)^2$

Feischl: *Math. Comp.*, 91 (2022)

Bringmann, Feischl, Miraçi, Praetorius, Streitberger: *Comput. Math. Appl.*, 180 (2025)