Adaptive High-Order Methods for Elliptic Problems: Convergence and Optimality

Claudio Canuto

Department of Mathematical Sciences
Politecnico di Torino, Italy

Joint work with

Ricardo H. Nochetto, University of Maryland, U.S.A.

Rob Stevenson, Korteweg-de Vries Institute for Mathematics, The Netherlands

Marco Verani, Politecnico di Milano, Italy

Foundations of Computational Mathematics
Barcelona, July 14 2017

Outline

Outline

Adaptive approximation of elliptic problems: the state of the art

- **Adaptivity for finite-order methods** [wavelets, h-type finite elements]: well-understood in terms of algorithms and theory (convergence, optimality)

  [Dörfler 1996, Morin, Nochetto and Siebert 2000, Binev, Dahmen and DeVore 2004, Stevenson 2007, Cascon, Kreuzer, Nochetto and Siebert 2008]

Adaptive approximation of elliptic problems: the state of the art

- **Adaptivity for finite-order methods** [wavelets, h-type finite elements]: well-understood in terms of algorithms and theory (convergence, optimality)

  [Dörfler 1996, Morin, Nochetto and Siebert 2000, Binev, Dahmen and DeVore 2004, Stevenson 2007, Cascon, Kreuzer, Nochetto and Siebert 2008]

- **Adaptivity for high-order methods** [spectral, hp-type finite elements]: heuristic algorithms, partial theory
  - **A posteriori error analysis**:

    [Gui and Babuška 1986, Oden, Demkowicz et al '89, Bernardi '96, Ainsworth and Senior '98, Schmidt and Siebert '00, Melenk and Wohlmuth '01, Heuvelin and Rannacher '03, Houston and Süli '05, Eibner and Melenk '07, Braess, Pillwein and Schöberl '08, Ern and Vohralík '14, ...]
  - **Convergence and optimality**:

    [Scherer 1982, Schmidt and Siebert 2000, Dörfler and Heuveline 2007, Bürg and Dörfler 2011, Bank, Parsania, and Sauter 2014, our work (2012 $\rightarrow$)]

Challenges for high-order adaptivity

- A suitable combination of 'h-refinement' and 'p-enrichment' may yield a fast (e.g., exponential) decay of the approximation error, even for functions with poor global smoothness.
  - For instance, the function $u(x) = x^\alpha$ with $\alpha < 1$ on $I = [0,1]$ can be approximated with an error of the form

    $$\text{approximation error} \sim C\,\mathrm{e}^{-\beta\sqrt{N}} \qquad N = \#\text{degrees of freedom}$$

    on a graded mesh geometrically refined towards the origin, with polynomial degrees linearly growing away from the origin.
    [DeVore-Scherer '79, Babuška-Guo '86].

Challenges for high-order adaptivity

- A suitable combination of 'h-refinement' and 'p-enrichment' may yield a fast (e.g., exponential) decay of the approximation error, even for functions with poor global smoothness.
  - For instance, the function $u(x) = x^\alpha$ with $\alpha < 1$ on $I = [0,1]$ can be approximated with an error of the form

    $$\text{approximation error} \sim C\,\mathrm{e}^{-\beta\sqrt{N}} \qquad N = \#\text{degrees of freedom}$$

    on a graded mesh geometrically refined towards the origin, with polynomial degrees linearly growing away from the origin.
    [DeVore-Scherer '79, Babuška-Guo '86].

- Need of dealing with approximation classes of functions for which the (best) approximation error decays faster than algebraically (e.g., exponentially).

Challenges for high-order adaptivity

- A suitable combination of 'h-refinement' and 'p-enrichment' may yield a fast (e.g., exponential) decay of the approximation error, even for functions with poor global smoothness.
  - For instance, the function $u(x) = x^\alpha$ with $\alpha < 1$ on $I = [0, 1]$ can be approximated with an error of the form

    $$\text{approximation error} \sim C\,\mathrm{e}^{-\beta\sqrt{N}} \qquad N = \#\text{degrees of freedom}$$

    on a graded mesh geometrically refined towards the origin, with polynomial degrees linearly growing away from the origin.
    [DeVore-Scherer '79, Babuška-Guo '86].

- Need of dealing with approximation classes of functions for which the (best) approximation error decays faster than algebraically (e.g., exponentially).

- The choice between 'h-refinement' and 'p-enrichment' is quite delicate.

  In an iterative adaptive algorithm, one of the two choices may appear preferable in an earlier stage, but eventually it may reveal itself short-sighted and non-optimal.

  One should incorporate the possibility of stepping back, and correcting early errors in the adaptive strategy.

Approximation classes

- **Best N-term approximation error:** Given $v \in V$, define

$$\sigma_N(v) = \inf_{\substack{V_N \subset V \\ \text{dim } V_N = N}} \inf_{w \in V_N} \|v - w\|_V .$$

- **Decay vs N identifies an approximation class:**

$$\sigma_N(v) \lesssim \phi(N) \qquad \text{with} \quad \phi \to 0 \quad \text{as} \quad N \to \infty.$$

Approximation classes

- **Best N-term approximation error:** Given $v \in V$, define

$$\sigma_N(v) = \inf_{\substack{V_N \subset V \\ \dim V_N = N}} \inf_{w \in V_N} \|v - w\|_V .$$

- **Decay vs N identifies an approximation class:**

$$\sigma_N(v) \lesssim \phi(N) \qquad \text{with} \quad \phi \to 0 \quad \text{as} \quad N \to \infty.$$

- **Algebraic class (finite-order methods):**

$$v \in \mathcal{A}_B^s \qquad \text{iff} \qquad |v|_{\mathcal{A}_B^s} := \sup_N \ \sigma_N(v) N^{s/d} < \infty.$$

- **Exponential class (infinite-order methods):**

$$v \in \mathcal{A}_G^{\eta,t} \qquad \text{iff} \qquad |v|_{\mathcal{A}_G^{\eta,t}} := \sup_N \ \sigma_N(v) \mathrm{e}^{\eta N^\tau} < \infty.$$

Complexity

Question: What is the cost involved in reducing the best approximation error $E(v_k) = \|v - v_k\|_V$ for a given function v by a fixed factor $\rho < 1$?

Complexity

Question: What is the cost involved in reducing the best approximation error $E(v_k) = \|v - v_k\|_V$ for a given function v by a fixed factor $\rho < 1$?

- **Algebraic decay:** Let $E(v_k)$ decay algebraically

$$E(v_k) = AN_k^{-s}$$

  in terms of degrees of freedom N_k. Then, a simple calculation yields

$$N_{k+1} = \rho^{-\frac{1}{s}} N_k$$

  The new number of degrees of freedom N_{k+1} is proportional to the current one N_k. This is what the h-theory predicts.

Complexity

Question: What is the cost involved in reducing the best approximation error $E(v_k) = \|v - v_k\|_V$ for a given function v by a fixed factor $\rho < 1$?

- **Algebraic decay:** Let $E(v_k)$ decay algebraically

$$E(v_k) = AN_k^{-s}$$

in terms of degrees of freedom N_k. Then, a simple calculation yields

$$N_{k+1} = \rho^{-\frac{1}{s}} N_k$$

The new number of degrees of freedom N_{k+1} is proportional to the current one N_k. This is what the h-theory predicts.

- **Exponential decay:** Let $E(v_k)$ decay exponentially

$$E(v_k) = Ae^{-\eta N_k}.$$

Then, a simple calculation reveals that

$$N_{k+1} - N_k = -\eta^{-1} \log \rho$$

and the number of degrees of freedom must only grow by an additive constant. This property is very delicate to prove!

Outline

Fourier methods

- **Periodic elliptic problem in** $\Omega = (0, 2\pi)^d$

$$-\nabla \cdot (\nu \nabla u) + \sigma u = f \ \text{ in } \Omega \,, \qquad\qquad u \ \ (2\pi)^d\text{-periodic},$$

formulated variationally in $V = H^1_{\text{per}}(\Omega)$ as

$$u \in V \quad : \quad a(u, v) = \langle f, v \rangle \qquad \forall v \in V,$$

and assumed to be continuous and coercive in V.

Fourier methods

- **Periodic elliptic problem in $\Omega = (0, 2\pi)^d$**

$$-\nabla \cdot (\nu \nabla u) + \sigma u = f \quad \text{in } \Omega \,, \qquad\qquad u \quad (2\pi)^d\text{-periodic,}$$

formulated variationally in $V = H^1_{\text{per}}(\Omega)$ as

$$u \in V \quad : \quad a(u, v) = \langle f, v \rangle \qquad \forall v \in V,$$

and assumed to be continuous and coercive in V.

- **Fourier basis** $\{\phi_k : k \in \mathbb{Z}^d\}$, normalized in V

$$v = \sum_k \hat{v}_k \phi_k \;, \qquad \text{with} \quad \|v\|_V^2 = \sum_k |\hat{v}_k|^2 \;.$$

Fourier methods

- **Periodic elliptic problem in $\Omega = (0, 2\pi)^d$**

$$-\nabla \cdot (\nu \nabla u) + \sigma u = f \quad \text{in } \Omega\,, \qquad u \quad (2\pi)^d\text{-periodic,}$$

formulated variationally in $V = H^1_{\text{per}}(\Omega)$ as

$$u \in V \quad : \quad a(u, v) = \langle f, v \rangle \qquad \forall v \in V,$$

and assumed to be continuous and coercive in V.

- **Fourier basis** $\{\phi_k : k \in \mathbb{Z}^d\}$, normalized in V

$$v = \sum_k \hat{v}_k \phi_k\,, \qquad \text{with} \quad \|v\|_V^2 = \sum_k |\hat{v}_k|^2\,.$$

- **Finite dimensional subspaces:** For arbitrary finite $\Lambda \subset \mathbb{Z}^d$, define

$$V_\Lambda = \text{span}\,\{\phi_k \;:\; k \in \Lambda\}.$$

and the orthogonal projection $P_\Lambda : V \to V_\Lambda$.

Galerkin approximation and residual

- **Galerkin projection**

$$u_\Lambda \in V_\Lambda \quad : \quad a(u_\Lambda, v_\Lambda) = \langle f, v_\Lambda \rangle \qquad \forall v_\Lambda \in V_\Lambda \ .$$

Galerkin approximation and residual

- **Galerkin projection**

$$u_\Lambda \in V_\Lambda \quad : \quad a(u_\Lambda, v_\Lambda) = \langle f, v_\Lambda \rangle \qquad \forall v_\Lambda \in V_\Lambda \ .$$

- **Residual** $r_\Lambda = r(u_\Lambda) \in V'$ defined by

$$\langle r_\Lambda, v \rangle = \langle f, v_\Lambda \rangle - a(u_\Lambda, v) \qquad \forall v \in V.$$

It satisfies

$$\|r_\Lambda\|^2_{V'} = \sum_{k \notin \Lambda} |\hat{r}_k|^2, \qquad \hat{r}_k = \langle r_\Lambda, \phi_k \rangle.$$

Galerkin approximation and residual

- **Galerkin projection**

$$u_\Lambda \in V_\Lambda \quad : \quad a(u_\Lambda, v_\Lambda) = \langle f, v_\Lambda \rangle \qquad \forall v_\Lambda \in V_\Lambda \ .$$

- **Residual** $r_\Lambda = r(u_\Lambda) \in V'$ defined by

$$\langle r_\Lambda, v \rangle = \langle f, v_\Lambda \rangle - a(u_\Lambda, v) \qquad \forall v \in V.$$

  It satisfies

$$\|r_\Lambda\|_{V'}^2 = \sum_{k \notin \Lambda} |\hat{r}_k|^2, \qquad \hat{r}_k = \langle r_\Lambda, \phi_k \rangle.$$

- **(Ideal) efficient and reliable error estimator**

$$\frac{1}{\alpha^*} \|r(u_\Lambda)\|_{V'} \leq \|u - u_\Lambda\|_V \leq \frac{1}{\alpha_*} \|r(u_\Lambda)\|_{V'} \ ,$$

Dörfler marking

- **Active basis updating:** Fix any $\theta \in (0,1)$. Given Λ, $u_\Lambda \in V_\Lambda$ and $r_\Lambda \in V'$, select

$$\Lambda_{\text{new}} = \Lambda \cup \partial\Lambda$$

by the condition

$$\|P_{\partial\Lambda} r_\Lambda\|_{V'} \geq \theta \|r_\Lambda\|_{V'}\ , \qquad \text{i.e.,} \qquad \sum_{k \in \partial\Lambda} |\hat{r}_k|^2 \ \geq\ \theta^2 \sum_{k \in \mathbb{Z}^d} |\hat{r}_k|^2\ .$$

Dörfler marking

- **Active basis updating:** Fix any $\theta \in (0,1)$. Given Λ, $u_\Lambda \in V_\Lambda$ and $r_\Lambda \in V'$, select
$$\Lambda_{\text{new}} = \Lambda \cup \partial\Lambda$$
by the condition
$$\|P_{\partial\Lambda} r_\Lambda\|_{V'} \geq \theta \|r_\Lambda\|_{V'} \,, \qquad \text{i.e.,} \qquad \sum_{k\in\partial\Lambda} |\hat{r}_k|^2 \;\geq\; \theta^2 \sum_{k\in\mathbb{Z}^d} |\hat{r}_k|^2 \,.$$

- **Minimality:** $\partial\Lambda$ may be chosen of minimal cardinality by a greedy approach, bases on the decreasing rearrangement of the moduli of the Fourier coefficients of r_Λ.

- **Feasibility:** Exploiting some more information on data, a feasible version exists, which requires exploring only a finite number of Fourier coefficients of r_Λ.

An ideal adaptive algorithm

Algorithm ADFOUR(θ, tol)

Set $r_0 := f$, $\Lambda_0 := \emptyset$, $n = -1$

do

$n \leftarrow n + 1$

$\partial\Lambda_n :=$ **DÖRFLER**(r_n, θ)

$\Lambda_{n+1} := \Lambda_n \cup \partial\Lambda_n$

$u_{n+1} :=$ **GAL**(Λ_{n+1})

$r_{n+1} :=$ **RES**(u_{n+1})

while $\|r_{n+1}\|_{V'} > tol$

An ideal adaptive algorithm

Algorithm ADFOUR(θ, tol)

Set $r_0 := f$, $\Lambda_0 := \emptyset$, $n = -1$
do
  $n \leftarrow n + 1$
  $\partial\Lambda_n :=$ **DÖRFLER**(r_n, θ)
  $\Lambda_{n+1} := \Lambda_n \cup \partial\Lambda_n$
  $u_{n+1} :=$ **GAL**(Λ_{n+1})
  $r_{n+1} :=$ **RES**(u_{n+1})
while $\|r_{n+1}\|_{V'} > \mathit{tol}$

Theorem (contraction property of ADFOUR). Let $\theta \in (0,1)$ and let $\{\Lambda_n,\, u_n\}_{n\geq 0}$ be the sequence generated by the adaptive algorithm above. Then,

$$\|u - u_{n+1}\| \leq \underbrace{\sqrt{1 - \frac{\alpha_*}{\alpha^*}\theta^2}}_{\rho(\theta)<1} \; \|u - u_n\|$$

where $\|v\| = \sqrt{a(v,v)}$.

A more aggressive version

If the coefficients of the equation are analytic, the Galerkin matrix is "quasi sparse". Exploiting this property, one can slightly enrich the active set produced by Dörfler's marking, and push the contraction constant towards 0 .

A more aggressive version

If the coefficients of the equation are analytic, the Galerkin matrix is "quasi sparse". Exploiting this property, one can slightly enrich the active set produced by Dörfler's marking, and push the contraction constant towards 0 .

Algorithm A-ADFOUR$(\theta,\ tol)$

Set $r_0 := f$, $\Lambda_0 := \emptyset$, $n = -1$
do
$n \leftarrow n+1$
$\widetilde{\partial\Lambda}_n :=$ **DÖRFLER**(r_n, θ)
$\partial\Lambda_n :=$ **ENRICH**$(\widetilde{\partial\Lambda}_n, \theta)$
$\Lambda_{n+1} := \Lambda_n \cup \partial\Lambda_n$
$u_{n+1} :=$ **GAL**(Λ_{n+1})
$r_{n+1} :=$ **RES**(u_{n+1})
while $\|r_{n+1}\|_{V'} > tol$

Theorem (contraction property of A-ADFOUR). Let $\theta \in (0,1)$ and let $\{\Lambda_n,\ u_n\}_{n\geq 0}$ be the sequence generated by **A-ADFOUR**. Then,

$$\|u - u_{n+1}\| \leq 2\sqrt{\frac{\alpha_*}{\alpha^*}}\ \sqrt{1-\theta^2}\ \|u - u_n\| \ .$$

Optimality issues

- **Target cardinality growth:** If the solution u belongs to some exponential class $\mathcal{A}_G^{\eta,t}$, one should expect

$$\#\Lambda_n \leq \left(\frac{1}{\eta} \log \frac{|u|_{\mathcal{A}_G^{\eta,t}}}{\|u - u_n\|} \right)^{1/\tau} + C \,, \qquad n = 0, 1, 2, \ldots$$

Optimality issues

- **Target cardinality growth:** If the solution u belongs to some exponential class $\mathcal{A}_G^{\eta,t}$, one should expect

$$\#\Lambda_n \leq \left(\frac{1}{\eta} \log \frac{|u|_{\mathcal{A}_G^{\eta,t}}}{\|u - u_n\|} \right)^{1/\tau} + C\,, \qquad n = 0, 1, 2, \ldots$$

- **Residual obstruction:** Dörfler marking is based on the current residual r_Λ. In general, this belongs to a worse approximation class than the solution.

$$v \in \mathcal{A}_G^{\eta,t} \qquad \Rightarrow \qquad r(v) \in \mathcal{A}_G^{\bar{\eta},\bar{t}} \text{ for some } \bar{\eta} \leq \eta,\ \bar{\tau} \leq \tau.$$

Optimality issues

- **Target cardinality growth:** If the solution u belongs to some exponential class $\mathcal{A}_G^{\eta,t}$, one should expect

$$\#\Lambda_n \le \left(\frac{1}{\eta} \log \frac{|u|_{\mathcal{A}_G^{\eta,t}}}{\|u-u_n\|} \right)^{1/\tau} + C\,, \qquad n = 0,1,2,\dots$$

- **Residual obstruction:** Dörfler marking is based on the current residual r_Λ. In general, this belongs to a worse approximation class than the solution.

$$v \in \mathcal{A}_G^{\eta,t} \qquad \Rightarrow \qquad r(v) \in \mathcal{A}_G^{\bar\eta,\bar t} \text{ for some } \bar\eta \le \eta,\ \bar\tau \le \tau.$$

- **Estimate on cardinality growth:**

$$\#\Lambda_n \le \left(\frac{1}{\bar\eta} \log \frac{|u|_{\mathcal{A}_G^{\eta,t}}}{\|u-u_n\|} \right)^{1/\bar\tau} + C\,, \qquad n = 0,1,2,\dots$$

- **Remedies...**

Remedy I: incorporating a coarsening step

- $\Lambda := \mathbf{COARSE}(w, \varepsilon)$
  Given $u \in \mathcal{A}_G^{\eta,\tau}$ and a function $w \in V$, which is known to satisfy
  $$\|u - w\| \le \varepsilon \ ,$$
  the output Λ is a set of minimal cardinality such that
  $$\|w - P_\Lambda w\| \le 2\varepsilon \ ,$$
  and
  $$\#\Lambda \le \left(\frac{1}{\eta} \log \frac{|u|_{\mathcal{A}_G^{\eta,\tau}}}{\varepsilon} \right)^{1/\tau} + 1.$$

Remedy I: incorporating a coarsening step

- $\Lambda := \mathbf{COARSE}(w, \varepsilon)$
  Given $u \in \mathcal{A}_G^{\eta,\tau}$ and a function $w \in V$, which is known to satisfy

$$\|u - w\| \le \varepsilon \ ,$$

the output Λ is a set of minimal cardinality such that

$$\|w - P_\Lambda w\| \le 2\varepsilon \ ,$$

and

$$\#\Lambda \le \left(\frac{1}{\eta} \log \frac{|u|_{\mathcal{A}_G^{\eta,\tau}}}{\varepsilon}\right)^{1/\tau} + 1.$$

Algorithm AC-ADFOUR$(\theta,\ tol)$

Set $r_0 := f$, $\Lambda_0 := \emptyset$, $n = -1$

do

$n \leftarrow n + 1$

$\widetilde{\partial\Lambda}_n :=$ **DÖRFLER**(r_n, θ)

$\widehat{\partial\Lambda}_n :=$ **ENRICH**$(\widetilde{\partial\Lambda}_n, \theta)$

$\widehat{\Lambda}_{n+1} := \Lambda_n \cup \widehat{\partial\Lambda}_n$

$\hat{u}_{n+1} :=$ **GAL**$(\widetilde{\Lambda}_{n+1})$

$\Lambda_{n+1} :=$ **COARSE**$(\hat{u}_{n+1}, \varepsilon_n)$

$u_{n+1} :=$ **GAL**(Λ_{n+1})

$r_{n+1} :=$ **RES**(u_{n+1})

while $\|r_{n+1}\|_{V'} > tol$

- The final sets Λ_n have quasi-optimal cardinality
  (while the intermediate sets have only suboptimal cardinality)

Remedy II: applying a super-aggressive Dörfler marking

- **Dynamic choice of Dörfler parameter:**

$$\theta \quad \to \quad \theta_n \qquad \text{such that} \quad \sqrt{1-\theta_n^2} \;\simeq\; \|r_n\| \; .$$

Remedy II: applying a super-aggressive Dörfler marking

- **Dynamic choice of Dörfler parameter:**

$$\theta \quad \to \quad \theta_n \qquad \text{such that} \quad \sqrt{1-\theta_n^2} \ \simeq \ \|r_n\| \, .$$

This yields:

- **quadratic convergence of the algorithm:**

$$\|u - u_{n+1}\| \lesssim \|u - u_n\|^2 \; ;$$

Remedy II: applying a super-aggressive Dörfler marking

- **Dynamic choice of Dörfler parameter:**

$$\theta \quad \to \quad \theta_n \qquad \text{such that} \quad \sqrt{1-\theta_n^2} \simeq \|r_n\| \, .$$

This yields:

- **quadratic convergence of the algorithm:**

$$\|u-u_{n+1}\| \lesssim \|u-u_n\|^2 \; ;$$

- **linear computational cost:** the computational cost of the algorithm scales linearly with $\#\Lambda_n$.

Remedy II: applying a super-aggressive Dörfler marking

- **Dynamic choice of Dörfler parameter:**

$$\theta \quad \to \quad \theta_n \qquad \text{such that} \quad \sqrt{1-\theta_n^2} \simeq \|r_n\| \, .$$

This yields:

- **quadratic convergence of the algorithm:**

$$\|u - u_{n+1}\| \lesssim \|u - u_n\|^2 \, ;$$

- **linear computational cost:** the computational cost of the algorithm scales linearly with $\#\Lambda_n$.

- **cardinalities $\#\partial\Lambda_k$ grow at a geometric rate**, showing that $\#\Lambda_n \simeq \#\partial\Lambda_n$, i.e., past iterations do not significantly influence the current cardinality;

Remedy II: applying a super-aggressive Dörfler marking

- **Dynamic choice of Dörfler parameter:**

$$\theta \quad \to \quad \theta_n \qquad \text{such that} \quad \sqrt{1-\theta_n^2} \simeq \|r_n\| \; .$$

This yields:

- **quadratic convergence of the algorithm:**

$$\|u - u_{n+1}\| \lesssim \|u - u_n\|^2 \; ;$$

- **linear computational cost:** the computational cost of the algorithm scales linearly with $\#\Lambda_n$.
- **cardinalities $\#\partial\Lambda_k$ grow at a geometric rate**, showing that $\#\Lambda_n \simeq \#\partial\Lambda_n$, i.e., past iterations do not significantly influence the current cardinality;
- **no need of coarsening.**

Outline

Abstract Framework for hp-Adaptivity

- **Operator equation:** Consider a, possibly, parametric operator equation (eg., a PDE) in a domain $\Omega \subset \mathbb{R}^n$

$$A_\lambda u = g.$$

  - The forcing g and the parameter λ (representing, e.g., the coefficients of the operator) are taken from some spaces G and Λ of functions on Ω.
  - For short, we will write $f = (g, \lambda) \in F = G \times \Lambda$.
  - We assume there exists a unique solution $u = u(f) \in V$, a space of functions on Ω. We assume, for simplicity, that V and F are Hilbert spaces over $\mathbb{R}$.

Abstract Framework for hp-Adaptivity

- **Operator equation:** Consider a, possibly, parametric operator equation (eg., a PDE) in a domain $\Omega \subset \mathbb{R}^n$
$$A_\lambda u = g.$$
  - The forcing g and the parameter λ (representing, e.g., the coefficients of the operator) are taken from some spaces G and Λ of functions on Ω.
  - For short, we will write $f = (g, \lambda) \in F = G \times \Lambda$.
  - We assume there exists a unique solution $u = u(f) \in V$, a space of functions on Ω. We assume, for simplicity, that V and F are Hilbert spaces over $\mathbb{R}$.

- **Binary trees:** From an initial partition of Ω, we generate an infinite binary master tree $\mathfrak{K}$ by recursively halving each element K in two children K' and K''.

- h**-partitions:** A finite subtree of $\mathfrak{K}$ defines an essentially disjoint h-partition $\mathcal{K}$ of Ω, by collecting all the leaves of the subtree. The set of all h-partitions is denoted by $\mathbb{K}$.

- hp-**partitions:** A hp-element is a pair $D = (K, d) \in \mathfrak{K} \times \mathbb{N}$, i.e., a geometric element K together with a dimension d.

  Given a h-partition $\mathcal{K}$, an associated hp-partition of Ω is a collection

$$\mathcal{D} = \{D = (K_D, d_D) : K_D \in \mathcal{K}\}.$$

  The set of all hp-partitions is denoted by $\mathbb{D}$.

Local Spaces for hp-Adaptivity

- **Local spaces:** For all $K \in \mathfrak{K}$, let V_K and F_K be (infinite dimensional) spaces of functions on K, such that for any $\mathcal{K} \in \mathbb{K}$ we have

$$V \subseteq \prod_{K \in \mathcal{K}} V_K, \qquad F \subseteq \prod_{K \in \mathcal{K}} F_K.$$

Local Spaces for hp-Adaptivity

- **Local spaces:** For all $K \in \mathfrak{K}$, let V_K and F_K be (infinite dimensional) spaces of functions on K, such that for any $\mathcal{K} \in \mathbb{K}$ we have

$$V \subseteq \prod_{K \in \mathcal{K}} V_K, \qquad F \subseteq \prod_{K \in \mathcal{K}} F_K.$$

- **Discrete local spaces:** For all hp-elements $D = (K, d) \in \mathfrak{K} \times \mathbb{N}$, given $Z \in \{V, F\}$ we let $Z_{K,d} \subset Z_K$ be finite dimensional spaces of functions on K such that

$$Z_{K,d} \subseteq Z_{K,d+1}, \qquad Z_{K,d} \subset Z_{K',d} \times Z_{K'',d}.$$

  We write $Z_D = Z_{K,d}$ and observe that any Z_D will be a polynomial space of dimension $\approx d$.

Local Spaces for hp-Adaptivity

- **Local spaces:** For all $K \in \mathfrak{K}$, let V_K and F_K be (infinite dimensional) spaces of functions on K, such that for any $\mathcal{K} \in \mathbb{K}$ we have

$$V \subseteq \prod_{K \in \mathcal{K}} V_K, \qquad F \subseteq \prod_{K \in \mathcal{K}} F_K.$$

- **Discrete local spaces:** For all hp-elements $D = (K, d) \in \mathfrak{K} \times \mathbb{N}$, given $Z \in \{V, F\}$ we let $Z_{K,d} \subset Z_K$ be finite dimensional spaces of functions on K such that

$$Z_{K,d} \subseteq Z_{K,d+1}, \qquad Z_{K,d} \subset Z_{K',d} \times Z_{K'',d}.$$

  We write $Z_D = Z_{K,d}$ and observe that any Z_D will be a polynomial space of dimension $\approx d$.

- **Example:** When K is an n-simplex, $V_{K,d}$ may be chosen as $\mathbb{P}_p(K)$, where the associated polynomial degree $p = p(d)$ can be defined as the largest value in $\mathbb{N}$ such that $\dim \mathbb{P}_{p-1}(K) = \binom{n+p-1}{p-1} \leq d$.
  - This definition normalizes the starting value $p(1) = 1$ for all $n \in \mathbb{N}$.
  - Only for $n = 1$, it holds that $p(d) = d$ for all $d \in \mathbb{N}$.

Local Error Functional and Monotonicity

- **Local error functional:** This is a quantity

$$e_D = e_D(v, f) \geq 0,$$

defined for all $(v, f) \in V \times F$, which measures the (squared) distance between $(v|_{K_D}, f|_{K_D})$ and its local approximation (v_D, f_D).

Local Error Functional and Monotonicity

- **Local error functional:** This is a quantity

$$e_D = e_D(v, f) \geq 0,$$

defined for all $(v, f) \in V \times F$, which measures the (squared) distance between $(v|_{K_D}, f|_{K_D})$ and its local approximation (v_D, f_D).

- **Local monotonicity:** We assume that $e_D = e_D(v, f)$ is non-increasing under both 'h-refinements' and 'p-enrichments', in the sense that

  - h-**refinement**

    $$e_{D'} + e_{D''} \leq e_D \quad \text{if } K_{D'}, K_{D''} \text{ are children of } K_D \text{ and } d_{D'} = d_{D''} = d_D;$$

  - p-**enrichment**

    $$e_{D'} \leq e_D \quad \text{if } K_{D'} = K_D \text{ and } d_{D'} \geq d_D.$$

Global Error Functional and Monotonicity

- **Global error functional:** For an hp-partition $\mathcal{D} = \{D = (K_D, d_D)\}$ of Ω, the *global error functional*

$$\mathrm{E}_{\mathcal{D}}(v,f) := \sum_{D\in\mathcal{D}} e_D(v,f),$$

  measures the (squared) distance between (v,f) and its projection onto $V_{\mathcal{D}} \times F_{\mathcal{D}}$, where $Z_{\mathcal{D}} = \big(Z_D\big)_{D\in\mathcal{D}}$.

- **Global monotonicity:**

$$\mathrm{E}_{\widetilde{\mathcal{D}}}(v,f) \le \mathrm{E}_{\mathcal{D}}(v,f) \qquad \text{if} \quad \mathcal{D} \le \widetilde{\mathcal{D}}.$$

Global Error Functional and Monotonicity

- **Global error functional:** For an hp-partition $\mathcal{D} = \{D = (K_D, d_D)\}$ of Ω, the *global error functional*

$$\mathrm{E}_{\mathcal{D}}(v, f) := \sum_{D \in \mathcal{D}} e_D(v, f),$$

measures the (squared) distance between (v, f) and its projection onto $V_{\mathcal{D}} \times F_{\mathcal{D}}$, where $Z_{\mathcal{D}} = \big(Z_D\big)_{D \in \mathcal{D}}$.

- **Global monotonicity:**

$$\mathrm{E}_{\widetilde{\mathcal{D}}}(v, f) \leq \mathrm{E}_{\mathcal{D}}(v, f) \qquad \text{if} \quad \mathcal{D} \leq \widetilde{\mathcal{D}}.$$

- **Notation:**
  - The cardinality of $\mathcal{D}$ is defined as $\#\mathcal{D} := \sum_{D \in \mathcal{D}} d_D$ (d_D local dimension).
  - The set $\mathbb{D}$ of all hp-partitions contains the subset $\mathbb{D}^c$ of the 'conforming' partitions. We assume that for any $\mathcal{D} \in \mathbb{D}$ there exists a conforming partition $\mathcal{C}(\mathcal{D})$ such that

$$\mathcal{D} \leq \mathcal{C}(\mathcal{D}).$$

Example of Choice of Error Functional e_D

- Consider the model elliptic problem in Ω

$$-\Delta u = f, \qquad u = 0 \quad \text{in} \quad \partial\Omega,$$

- Define as a local error functional

$$\mathsf{e}_D(v,f) := |v - P^1_{p_D} v|^2_{H^1(K_D)} + \frac{1}{\kappa} \|p_D^{-1} h_D (f - P^0_{p_D-1} f)\|^2_{L^2(K_D)} \qquad \forall D \in \mathcal{D},$$

where

  - P^1_p, P^0_p resp. are orthogonal projectors on $\mathbb{P}_p(K_D)$ in the inner products of $L^2(K_D)$, $H^1_0(K_D)$, resp.
  - κ is a parameter to be chosen later on.

Outline

- **Goal:** Given two functions $(v, f) \in V \times F$ and a target accuracy $\varepsilon > 0$, find a "near optimal" hp-partition $\mathcal{D}$ such that

$$E_{\mathcal{D}}(v, f) \leq \varepsilon.$$

- The task will be realized in two stages...

h-Adaptive Tree Approximation

- **Admissible binary tree:** Given $\mathcal{K} \in \mathbb{K}$, an admissible binary tree $\mathcal{T}$ is the set of all $K \in \mathcal{K}$ and their ancestors. We note that $\mathcal{T} \subset \mathfrak{K}$ is finite and denote by $\mathcal{L}(\mathcal{T})$ the leaves of $\mathcal{T}$, i.e. elements without successors.

- **Local h-error functional:** This is a subadditive quantity e_K

$$e_{K'} + e_{K''} \leq e_K \quad \forall K \in \mathfrak{K},$$

  where K' and K'' denote the children of K. Given a function $v \in L^2(\Omega)$, e_K is simply the square of the best L^2-error in K.

- **Global h-error functional:** $\quad \mathrm{E}_{\mathcal{K}} = \sum_{K \in \mathcal{K}} e_K \quad \forall \mathcal{K} \in \mathbb{K}.$

- **Best h-approximation:** Given $N \in \mathbb{N}$, let

$$\sigma_N := \inf_{\#\mathcal{K} \leq N} \mathrm{E}_{\mathcal{K}}\,.$$

  For functions in $L^2(\Omega)$ this gives the best L^2-error but computing a tree that realizes the $\min$ has exponential complexity.

Near-Best h-Adaptive Tree Approximation (Binev-DeVore)

- **Modified local error functional:** $\tilde{e}_K$ for all $K \in \mathcal{K}$
  - $\tilde{e}_K := e_K$ if K is a root;
  - $\frac{1}{\tilde{e}_K} := \frac{1}{e_K} + \frac{1}{\tilde{e}_{K^*}}$ where K^* is the parent of K and $e_K \neq 0$; else $\tilde{e}_K = 0$.

Near-Best h-Adaptive Tree Approximation (Binev-DeVore)

- **Modified local error functional:** $\tilde{e}_K$ for all $K \in \mathcal{K}$
  - $\tilde{e}_K := e_K$ if K is a root;
  - $\frac{1}{\tilde{e}_K} := \frac{1}{e_K} + \frac{1}{\tilde{e}_{K^*}}$ where K^* is the parent of K and $e_K \neq 0$; else $\tilde{e}_K = 0$.

- **Greedy algorithm on $\{\tilde{e}_K\}_{K\in\mathcal{K}}$:** Given a tree $\mathcal{K}_N$, with $\#\mathcal{K}_N = N$, construct $\mathcal{K}_{N+1}$ by bisecting the leaf $K \in \mathcal{L}(\mathcal{K})$ with largest $\tilde{e}_K$.

Near-Best h-Adaptive Tree Approximation (Binev-DeVore)

- **Modified local error functional:** $\tilde{e}_K$ for all $K \in \mathcal{K}$
  - $\tilde{e}_K := e_K$ if K is a root;
  - $\frac{1}{\tilde{e}_K} := \frac{1}{e_K} + \frac{1}{\tilde{e}_{K^*}}$ where K^* is the parent of K and $e_K \neq 0$; else $\tilde{e}_K = 0$.
- **Greedy algorithm on** $\{\tilde{e}_K\}_{K\in\mathcal{K}}$**:** Given a tree $\mathcal{K}_N$, with $\#\mathcal{K}_N = N$, construct $\mathcal{K}_{N+1}$ by bisecting the leaf $K \in \mathcal{L}(\mathcal{K})$ with largest $\tilde{e}_K$.
- **Instance optimality:** The sequence of trees $\{\mathcal{K}_N\}$ given by the greedy algorithm on $\{\tilde{e}_K\}_{K\in\mathcal{K}}$ provides a near-best h-adaptive approximation in the sense
$$\mathrm{E}_{\mathcal{K}_N} \leq \frac{N}{N-n+1}\sigma_n$$
for any integer $n \leq N$. The complexity for obtaining $\mathcal{K}_N$ is $\mathcal{O}(N)$.

Near-Best h-Adaptive Tree Approximation (Binev-DeVore)

- **Modified local error functional:** $\tilde{e}_K$ for all $K \in \mathcal{K}$
  - $\tilde{e}_K := e_K$ if K is a root;
  - $\frac{1}{\tilde{e}_K} := \frac{1}{e_K} + \frac{1}{\tilde{e}_{K^*}}$ where K^* is the parent of K and $e_K \neq 0$; else $\tilde{e}_K = 0$.

- **Greedy algorithm on $\{\tilde{e}_K\}_{K \in \mathcal{K}}$:** Given a tree $\mathcal{K}_N$, with $\#\mathcal{K}_N = N$, construct $\mathcal{K}_{N+1}$ by bisecting the leaf $K \in \mathcal{L}(\mathcal{K})$ with largest $\tilde{e}_K$.

- **Instance optimality:** The sequence of trees $\{\mathcal{K}_N\}$ given by the greedy algorithm on $\{\tilde{e}_K\}_{K \in \mathcal{K}}$ provides a near-best h-adaptive approximation in the sense

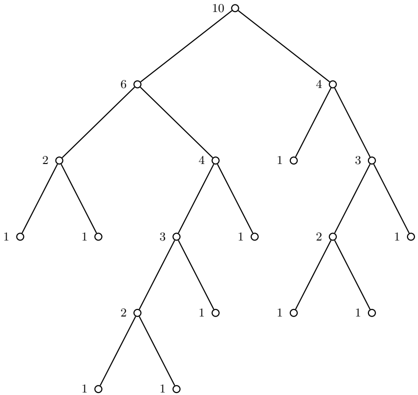
$$\mathrm{E}_{\mathcal{K}_N} \leq \frac{N}{N-n+1} \sigma_n$$

for any integer $n \leq N$. The complexity for obtaining $\mathcal{K}_N$ is $\mathcal{O}(N)$.

- **Interpretation:** Given N let $n = \lceil \frac{N}{2} \rceil$. Then $N - n + 1 \geq N/2$ and

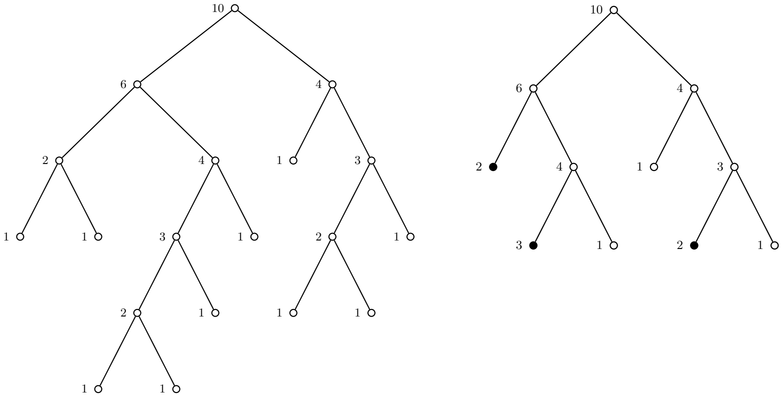
$$\mathrm{E}_{\mathcal{K}_N} \leq 2\sigma_{\lceil \frac{N}{2} \rceil}.$$

hp-adaptivity: the Ghost h-Tree...



Ghost h-tree $\mathcal{T}$ (left) with 10 leaves ($\#\mathcal{L}(\mathcal{T}) = 10$); the root K of $\mathcal{T}$ thus has an admissible dimension (polynomial degree for $n = 1$) $d(K, \mathcal{T}) = 10$.

hp-adaptivity: the Ghost h-Tree... and Subordinate hp-Tree (Binev)



Ghost h-tree $\mathcal{T}$ (left) with 10 leaves ($\#\mathcal{L}(\mathcal{T}) = 10$); the root K of $\mathcal{T}$ thus has an admissible dimension (polynomial degree for $n = 1$) $d(K, \mathcal{T}) = 10$.
The subordinate hp-tree $\mathcal{P}$ (right) results from $\mathcal{T}$ upon trimming 3 subtrees and raising the polynomial degrees of the interior nodes of $\mathcal{T}$, now leaves of $\mathcal{P}$, to $d(K, \mathcal{T}) = 2, 3, 2$ respectively.

Adaptive Strategy for hp-Refinements: hp-NEARBEST(Binev)

- **Ghost h-tree $\mathcal{T}$:** This is the previous h-tree associated with $v \in L^2(\Omega)$.
- **Admissible dimension:** Given $K \in \mathcal{T}$, the dimension $d(K, \mathcal{T})$ is

$$d(K, \mathcal{T}) = \#\mathcal{L}(\mathcal{T}(K)),$$

where $\mathcal{T}(K)$ is the subtree of $\mathcal{T}$ emanating from K. This quantity depends on both T and the underlying tree $\mathcal{T}$.

Adaptive Strategy for hp-Refinements: hp-NEARBEST(Binev)

- **Ghost h-tree $\mathcal{T}$:** This is the previous h-tree associated with $v \in L^2(\Omega)$.
- **Admissible dimension:** Given $K \in \mathcal{T}$, the dimension $d(K, \mathcal{T})$ is

$$d(K, \mathcal{T}) = \#\mathcal{L}(\mathcal{T}(K)),$$

  where $\mathcal{T}(K)$ is the subtree of $\mathcal{T}$ emanating from K. This quantity depends on both T and the underlying tree $\mathcal{T}$.
- **Local hp-error functionals:** Let $e_{K,d}$ be the local error functional on $K \in \mathcal{T}$ with polynomial dimension d. The modified local hp-error functional $e_K(\mathcal{T})$ reads
  - $e_K(\mathcal{T}) := e_{K,1}$ provided $K \in \mathcal{L}(\mathcal{T})$ is a leaf;
  - $e_K(\mathcal{T}) := \min\left\{e_{K'}(\mathcal{T}) + e_{K''}(\mathcal{T}), e_{K,d(K,\mathcal{T})}\right\}$ otherwise.
- **Subordinate hp-tree $\mathcal{P}$:** This tree is obtained from the h-tree upon eliminating the subtree $\mathcal{T}(K)$ whenever increasing the polynomial dimension in K from 1 to $d(K, \mathcal{T})$ reduces the error, i.e.

$$e_K(\mathcal{T}) = e_{K,d(K,\mathcal{T})}.$$

Instance Optimality of hp-NEARBEST

- **Theorem (Binev):** The subordinate hp-tree $\mathcal{P}_N$ with cardinality

$$\#\mathcal{P}_N = \sum_{K \in \mathcal{L}(\mathcal{P}_N)} d(K, \mathcal{T}_N) = \#\mathcal{L}(\mathcal{T}_N) = N$$

gives a hp partition $\mathcal{D}_N$ with $\#\mathcal{D}_N = N$ and near-best hp-approximation over $\mathcal{D}_N$ in the sense that the global error functional satisfies

$$\mathrm{E}_{\mathcal{D}_N}(v, f) \leq \frac{2N}{N - n + 1} \sigma_n(v, f) \quad \forall\, N \geq n,$$

where σ_n is the best hp-error for (v, f) with n total degrees of freedom.

Instance Optimality of hp-NEARBEST

- **Theorem (Binev):** The subordinate hp-tree $\mathcal{P}_N$ with cardinality

$$\#\mathcal{P}_N = \sum_{K \in \mathcal{L}(\mathcal{P}_N)} d(K, \mathcal{T}_N) = \#\mathcal{L}(\mathcal{T}_N) = N$$

gives a hp partition $\mathcal{D}_N$ with $\#\mathcal{D}_N = N$ and near-best hp-approximation over $\mathcal{D}_N$ in the sense that the global error functional satisfies

$$\mathrm{E}_{\mathcal{D}_N}(v, f) \leq \frac{2N}{N - n + 1} \sigma_n(v, f) \quad \forall\, N \geq n,$$

where σ_n is the best hp-error for (v, f) with n total degrees of freedom.

  - The cost for constructing $\mathcal{D}_N$ is bounded by $\mathcal{O}\big(\sum_{K \in \mathcal{T}_N} d(K, \mathcal{T}_N)\big)$, and varies from $\mathcal{O}(N \log N)$ for well balanced trees to $\mathcal{O}(N^2)$ for highly unbalanced trees.

Instance Optimality of hp-NEARBEST

- **Theorem (Binev):** The subordinate hp-tree $\mathcal{P}_N$ with cardinality

$$\#\mathcal{P}_N = \sum_{K\in\mathcal{L}(\mathcal{P}_N)} d(K,\mathcal{T}_N) = \#\mathcal{L}(\mathcal{T}_N) = N$$

gives a hp partition $\mathcal{D}_N$ with $\#\mathcal{D}_N = N$ and near-best hp-approximation over $\mathcal{D}_N$ in the sense that the global error functional satisfies

$$\mathrm{E}_{\mathcal{D}_N}(v,f) \leq \frac{2N}{N-n+1}\sigma_n(v,f) \quad \forall\, N \geq n,$$

where σ_n is the best hp-error for (v,f) with n total degrees of freedom.

  - The cost for constructing $\mathcal{D}_N$ is bounded by $\mathcal{O}\big(\sum_{K\in\mathcal{T}_N} d(K,\mathcal{T}_N)\big)$, and varies from $\mathcal{O}(N\log N)$ for well balanced trees to $\mathcal{O}(N^2)$ for highly unbalanced trees.

- **Interpretation:** Choosing $B>1$, $n=\frac{N}{B}$ and $b=\frac{1}{2}(1-\frac{1}{B})<1$ implies
  - $\mathrm{E}_{\mathcal{D}_N}(v,f)\leq\varepsilon$
  - $\#\mathcal{D}_N \leq B\#\mathcal{D}$ for all $\mathcal{D}\in\mathbb{D}$ such that $\mathrm{E}_{\mathcal{D}}(v,f)\leq b\varepsilon$.

Outline

Basic Modules

We assume availability of the following routines, which realize the two fundamental steps of the algorithm.

- $[\mathcal{D}, f_{\mathcal{D}}] :=$ **hp-NEARBEST**(ε, v, f)

  The routine **hp-NEARBEST** takes as input $\varepsilon > 0$ and $(v, f) \in V \times F$, and outputs $\mathcal{D} \in \mathbb{D}$ as well as $f_{\mathcal{D}}$, such that

  (i) $\mathrm{E}_{\mathcal{D}}(v, f)^{\frac{1}{2}} \leq \varepsilon$;

  (ii) $\#\mathcal{D} \leq B\#\widehat{\mathcal{D}}$ for any $\widehat{\mathcal{D}} \in \mathbb{D}$ with $\mathrm{E}_{\widehat{\mathcal{D}}}(v, f)^{\frac{1}{2}} \leq b\varepsilon$, for some constants $0 < b \leq 1 \leq B$.

  This routine may be implemented via Binev's algorithm.

- $[\bar{\mathcal{D}}, \bar{u}] :=$ **PDE**$(\varepsilon, \mathcal{D}, f_{\mathcal{D}})$

  The routine **PDE** takes as input $\varepsilon > 0$, $\mathcal{D} \in \mathbb{D}^c$, and data $f_{\mathcal{D}} \in F_{\mathcal{D}}$. It outputs $\bar{\mathcal{D}} \in \mathbb{D}^c$ with $\mathcal{D} \leq \bar{\mathcal{D}}$ and $\bar{u} \in V^c_{\bar{\mathcal{D}}}$ such that $\|u(f_{\mathcal{D}}) - \bar{u}\|_V \leq \varepsilon$.

Assumptions on Global Error Functional

We assume the existence of constants $C_1, C_2 > 0$ with

$$C_1 C_2 < b,$$

such that the following properties hold:

- **Continuity of the solution upon data:**

$$\|u(f) - u(f_{\mathcal{D}})\|_V \le C_1 \inf_{w \in V} \mathrm{E}_{\mathcal{D}}(w, f)^{\frac{1}{2}} \qquad \forall \mathcal{D} \in \mathbb{D}, \forall\, f \in F,$$

- **Lipschitz continuity of $\mathrm{E}_{\mathcal{D}}$ upon state:**

$$\sup_{f \in F} |\, \mathrm{E}_{\mathcal{D}}(w, f)^{\frac{1}{2}} - \mathrm{E}_{\mathcal{D}}(v, f)^{\frac{1}{2}}\,| \le C_2 \|w - v\|_V \qquad \forall \mathcal{D} \in \mathbb{D}, \forall v, w \in V.$$

Verifying the Assumptions on Global Error Functional

Error and Oscillation: Let

$$\mathsf{e}_D(v,f) := |v - P^1_{p_D} v|^2_{H^1(K_D)} + \frac{1}{\kappa}\operatorname{osc}_D(f)^2 \qquad \forall D \in \mathcal{D},$$

with

$$\operatorname{osc}_D(f)^2 = \|p_D^{-1} h_D (f - P^0_{p_D - 1} f)\|^2_{L^2(K_D)}$$

- **Continuity of the solution upon data:**

$$\|u(f) - u(f_{\mathcal{D}})\|_{H^1(\Omega)} \le C \operatorname{osc}_{\mathcal{D}}(f) \le \underbrace{C\kappa^{\frac{1}{2}}}_{=C_1} \mathrm{E}_{\mathcal{D}}(w,f)^{\frac{1}{2}} \quad \forall w \in V.$$

- **Lipschitz continuity of $\mathrm{E}_{\mathcal{D}}$ upon state:**

$$\left| \mathrm{E}_{\mathcal{D}}(w,f)^{\frac{1}{2}} - \mathrm{E}_{\mathcal{D}}(v,f)^{\frac{1}{2}} \right| \le \|w - v\|_{H^1(\Omega)} \quad \Rightarrow \quad C_2 = 1.$$

Verifying the Assumptions on Global Error Functional

Error and Oscillation: Let

$$\mathsf{e}_D(v,f) := |v - P^1_{p_D} v|^2_{H^1(K_D)} + \frac{1}{\kappa}\mathrm{osc}_D(f)^2 \qquad \forall D \in \mathcal{D},$$

with

$$\mathrm{osc}_D(f)^2 = \|p_D^{-1} h_D (f - P^0_{p_D-1} f)\|^2_{L^2(K_D)}$$

- **Continuity of the solution upon data:**

$$\|u(f) - u(f_{\mathcal{D}})\|_{H^1(\Omega)} \le C\,\mathrm{osc}_{\mathcal{D}}(f) \le \underbrace{C\kappa^{\frac{1}{2}}}_{=C_1} \mathrm{E}_{\mathcal{D}}(w,f)^{\frac{1}{2}} \quad \forall w \in V.$$

- **Lipschitz continuity of $\mathrm{E}_{\mathcal{D}}$ upon state:**

$$\left|\,\mathrm{E}_{\mathcal{D}}(w,f)^{\frac{1}{2}} - \mathrm{E}_{\mathcal{D}}(v,f)^{\frac{1}{2}}\,\right| \le \|w - v\|_{H^1(\Omega)} \quad \Rightarrow \quad C_2 = 1.$$

- **Bound on constants:** $C_1 C_2 = C\kappa^{\frac{1}{2}} C_2 < b$ for κ sufficiently small.

Basic hp-AFEM

- **hp-AFEM:**

  hp-AFEM$(\bar{u}_0, f, \varepsilon_0)$

  % Input: $(\bar{u}_0, f) \in V \times F$, $\varepsilon_0 > 0$ with $\|u(f) - \bar{u}_0\|_V \leq \varepsilon_0$.

  % Parameters: $\mu \in (0,1)$ such that $C_1 C_2 < b(1-\mu)$, and $\omega \in (\frac{C_2}{b}, \frac{1-\mu}{C_1})$.

  for $i = 1, 2, \ldots$ do

  $[\mathcal{D}_i, f_{\mathcal{D}_i}]$:=**hp-NEARBEST**$(\omega\varepsilon_{i-1}, \bar{u}_{i-1}, f)$

  $[\bar{\mathcal{D}}_i, \bar{u}_i]$:= **PDE**$(\mu\varepsilon_{i-1}, \mathcal{C}(\mathcal{D}_i), f_{\mathcal{D}_i})$

  $\varepsilon_i := (\mu + C_1\omega)\varepsilon_{i-1}$

  end do

- **Error Reduction:** Note that $\varepsilon_i = (\mu + C_1\omega)^i \varepsilon_0$ where $\mu + C_1\omega < 1$.

Basic hp-AFEM

- **hp-AFEM:**

hp-AFEM$(\bar{u}_0, f, \varepsilon_0)$

% Input: $(\bar{u}_0, f) \in V \times F$, $\varepsilon_0 > 0$ with $\|u(f) - \bar{u}_0\|_V \leq \varepsilon_0$.

% Parameters: $\mu \in (0,1)$ such that $C_1 C_2 < b(1-\mu)$, and $\omega \in (\frac{C_2}{b}, \frac{1-\mu}{C_1})$.

```
for i = 1, 2, ... do
    [D_i, f_{D_i}] := hp-NEARBEST(ω ε_{i-1}, ū_{i-1}, f)
    [D̄_i, ū_i] := PDE(μ ε_{i-1}, C(D_i), f_{D_i})
    ε_i := (μ + C_1 ω) ε_{i-1}
end do
```

- **Error Reduction:** Note that $\varepsilon_i = (\mu + C_1\omega)^i \varepsilon_0$ where $\mu + C_1\omega < 1$.

- **Coarsening:** Tolerance for $\bar{u}_i$ within **PDE** is $\tau_i := \mu\varepsilon_{i-1}$ and the subsequent input tolerance of **hp-NEARBEST** is $\omega\varepsilon_i = \omega(\mu + C_1\omega)\varepsilon_{i-1} > \omega\tau_i$. Since in our applications $C_2 = 1$ and $\omega > C_2/b \geq 1$, we see that $\omega\varepsilon_i > \tau_i = \mu\varepsilon_{i-1}$.

Convergence and Instance Optimality

Theorem. Let the previous assumptions on the global error functional $E_{\mathcal{D}}$ be satisfied. Then, for the sequences $(\bar{u}_i)$, $(\mathcal{D}_i)$ produced in **hp-AFEM**, it holds that

$$\|u - \bar{u}_i\|_V \leq \varepsilon_i \quad \forall i \geq 0, \qquad \mathrm{E}_{\mathcal{D}_i}(u,f)^{\frac{1}{2}} \leq \frac{\omega + C_2}{\mu + C_1\omega}\,\varepsilon_i \quad \forall i \geq 1,$$

and

$$\#\mathcal{D}_i \leq B\#\mathcal{D} \quad \text{for any } \mathcal{D} \in \mathbb{D} \ \text{ with } \mathrm{E}_{\mathcal{D}}(u,f)^{\frac{1}{2}} \leq \frac{b\omega - C_2}{\mu + C_1\omega}\,\varepsilon_i\,,$$

where $u = u(f)$.

Outline

A practical hp-adaptive algorithm

- **Recall PDE:**
  - $[\bar{\mathcal{D}}, \bar{u}] :=$ **PDE**$(\varepsilon, \mathcal{D}, f_{\mathcal{D}})$

    The routine **PDE** takes as input $\varepsilon > 0$, $\mathcal{D} \in \mathbb{D}^c$, and data $f_{\mathcal{D}} \in F_{\mathcal{D}}$. It outputs $\bar{\mathcal{D}} \in \mathbb{D}^c$ with $\mathcal{D} \leq \bar{\mathcal{D}}$ and $\bar{u} \in V^c_{\bar{\mathcal{D}}}$ such that $\|u(f_{\mathcal{D}}) - \bar{u}\|_V \leq \varepsilon$.

A practical hp-adaptive algorithm

- **Recall PDE:**
  - $[\bar{\mathcal{D}}, \bar{u}] :=$ **PDE**$(\varepsilon, \mathcal{D}, f_{\mathcal{D}})$

    The routine **PDE** takes as input $\varepsilon > 0$, $\mathcal{D} \in \mathbb{D}^c$, and data $f_{\mathcal{D}} \in F_{\mathcal{D}}$. It outputs $\bar{\mathcal{D}} \in \mathbb{D}^c$ with $\mathcal{D} \leq \bar{\mathcal{D}}$ and $\bar{u} \in V^c_{\bar{\mathcal{D}}}$ such that $\|u(f_{\mathcal{D}}) - \bar{u}\|_V \leq \varepsilon$.

- **Error reduction:** For efficiency, **PDE** should exploit the work already carried out within $\mathbf{hp}$-**AFEM**.
  Precisely, for any desired error reduction factor $\varrho \in (0,1)$, it should give

$$\|u(f_{\mathcal{D}}) - \bar{u}\|_V \leq \varrho \inf_{v \in V^c_{\mathcal{D}}} \|u(f_{\mathcal{D}}) - v\|_V .$$

A practical hp-adaptive algorithm

- **Recall PDE:**
  - $[\bar{\mathcal{D}}, \bar{u}] :=$ **PDE**$(\varepsilon, \mathcal{D}, f_{\mathcal{D}})$

    The routine **PDE** takes as input $\varepsilon > 0$, $\mathcal{D} \in \mathbb{D}^c$, and data $f_{\mathcal{D}} \in F_{\mathcal{D}}$. It outputs $\bar{\mathcal{D}} \in \mathbb{D}^c$ with $\mathcal{D} \leq \bar{\mathcal{D}}$ and $\bar{u} \in V^c_{\bar{\mathcal{D}}}$ such that $\|u(f_{\mathcal{D}}) - \bar{u}\|_V \leq \varepsilon$.

- **Error reduction:** For efficiency, **PDE** should exploit the work already carried out within **hp-AFEM**.
  Precisely, for any desired error reduction factor $\varrho \in (0,1)$, it should give

$$\|u(f_{\mathcal{D}}) - \bar{u}\|_V \leq \varrho \inf_{v \in V^c_{\mathcal{D}}} \|u(f_{\mathcal{D}}) - v\|_V .$$

- Indeed, at each call of **PDE** within **hp-AFEM**, we will be already guaranteed to have

$$\inf_{v \in V^c_{\mathcal{D}}} \|u(f_{\mathcal{D}}) - v\|_V \leq C\varepsilon,$$

  whence a suitable choice of ϱ will yield

$$\|u(f_{\mathcal{D}}) - \bar{u}\|_V \leq \varepsilon.$$

A practical hp-adaptive algorithm

- **Recall PDE:**
  - $[\bar{\mathcal{D}}, \bar{u}] :=$ **PDE**$(\varepsilon, \mathcal{D}, f_{\mathcal{D}})$

    The routine **PDE** takes as input $\varepsilon > 0$, $\mathcal{D} \in \mathbb{D}^c$, and data $f_{\mathcal{D}} \in F_{\mathcal{D}}$. It outputs $\bar{\mathcal{D}} \in \mathbb{D}^c$ with $\mathcal{D} \leq \bar{\mathcal{D}}$ and $\bar{u} \in V^c_{\bar{\mathcal{D}}}$ such that $\|u(f_{\mathcal{D}}) - \bar{u}\|_V \leq \varepsilon$.

- **Error reduction:** For efficiency, **PDE** should exploit the work already carried out within **hp-AFEM**.
  Precisely, for any desired error reduction factor $\varrho \in (0,1)$, it should give
  $$\|u(f_{\mathcal{D}}) - \bar{u}\|_V \leq \varrho \inf_{v \in V^c_{\mathcal{D}}} \|u(f_{\mathcal{D}}) - v\|_V .$$

- Indeed, at each call of **PDE** within **hp-AFEM**, we will be already guaranteed to have
  $$\inf_{v \in V^c_{\mathcal{D}}} \|u(f_{\mathcal{D}}) - v\|_V \leq C\varepsilon,$$
  whence a suitable choice of ϱ will yield
  $$\|u(f_{\mathcal{D}}) - \bar{u}\|_V \leq \varepsilon.$$

- **Remark:** The input data $f_{\mathcal{D}}$ is piecewise polynomial on the input partition $\mathcal{D}$, hence no data oscillation appears.

- **PDE:** This module may be implemented by the usual loop

SOLVE $\rightarrow$ **ESTIMATE** $\rightarrow$ **MARK** $\rightarrow$ **GROW**

where

GROW may be either an h-refinement or a p-enrichment.

- **PDE:** This module may be implemented by the usual loop

$$\textbf{SOLVE} \rightarrow \textbf{ESTIMATE} \rightarrow \textbf{MARK} \rightarrow \textbf{GROW}$$

where

GROW may be either an h-refinement or a p-enrichment.

- **Features:**
  - **ESTIMATE**: should be based on a reliable and efficient *a posteriori* error indicator, with constants independent of the current hp-partition $\mathcal{D}$ (in particular, "p-robust").

- **PDE:** This module may be implemented by the usual loop

  SOLVE $\rightarrow$ **ESTIMATE** $\rightarrow$ **MARK** $\rightarrow$ **GROW**

  where

  GROW may be either an h-refinement or a p-enrichment.

- **Features:**

  - **ESTIMATE**: should be based on a reliable and efficient *a posteriori* error indicator, with constants independent of the current hp-partition $\mathcal{D}$ (in particular, "p-robust").

  - **GROW**: should yield a new hp partition $\mathcal{D}_{\text{new}}$ with $\#\mathcal{D}_{\text{new}} \lesssim \#\mathcal{D}$ and guaranteed *saturation property*

    $$\|u - u_{\mathcal{D}}\|_V \lesssim \|u_{\mathcal{D}_{\text{new}}} - u_{\mathcal{D}}\|_V$$

    (all constants independent of $\mathcal{D}$).

- **PDE:** This module may be implemented by the usual loop

  SOLVE $\rightarrow$ **ESTIMATE** $\rightarrow$ **MARK** $\rightarrow$ **GROW**

  where

  GROW may be either an h-refinement or a p-enrichment.

- **Features:**
  - **ESTIMATE**: should be based on a reliable and efficient *a posteriori* error indicator, with constants independent of the current hp-partition $\mathcal{D}$ (in particular, "p-robust").
  - **GROW**: should yield a new hp partition $\mathcal{D}_{\text{new}}$ with $\#\mathcal{D}_{\text{new}} \lesssim \#\mathcal{D}$ and guaranteed *saturation property*

    $$\|u - u_{\mathcal{D}}\|_V \lesssim \|u_{\mathcal{D}_{\text{new}}} - u_{\mathcal{D}}\|_V$$

    (all constants independent of $\mathcal{D}$).
  - **MARK**: might be skipped. Indeed, since the task of producing a near-best hp-partition is assigned to **hp-NEARBEST**, in principle even a *uniform* refinement/enrichment is allowed.

Applications to elliptic self-adjoint problems

We detail three possible realizations of **PDE**, based on:

- a residual estimator, in dimension 1,
- a residual estimator, in dimension 2
- an equilibrated flux estimator, in dimension 2.

Applications to elliptic self-adjoint problems

We detail three possible realizations of **PDE**, based on:

- a residual estimator, in dimension 1,
- a residual estimator, in dimension 2
- an equilibrated flux estimator, in dimension 2.

- **In a 1D domain**, for the general elliptic self-adjoint problem

$$u \in H_0^1(\Omega) \ : \quad -(\mu u_x)_x + \sigma u = f + g_x \quad \text{in} \quad H^{-1}(\Omega),$$

  - the residual-based error estimator $\eta_{\mathcal{D}}(u_{\mathcal{D}}, f_{\mathcal{D}})$ defined by

$$\eta^2_{\mathcal{D}}(u_{\mathcal{D}}, f_{\mathcal{D}}) = \sum_{D \in \mathcal{D}} \|r_D\|^2_{H^{-1}(K_D)}$$

    is p-robust and easily computable element-wise;
  - the saturation property is guaranteed by raising the polynomial degree from p_D to some $\hat{p}_D \leq 2p_D + 3$ in each marked element.
  - Thus, **hp-AFEM** is fully optimal.

Residual-based estimators

- **In a 2D domain (a polygon)**, for the model problem

$$u \in H_0^1(\Omega) \;\; : \quad -\Delta u = f \quad \text{in } \Omega,$$

  - the Melenk-Wohlmuth residual-based error estimator $\eta_{\mathcal{D}}(u_{\mathcal{D}}, f_{\mathcal{D}})$ defined element-wise by

$$\eta_D^2(u_{\mathcal{D}}, f_{\mathcal{D}}) := \frac{|K_D|}{p_D^2} \|f_{\mathcal{D}} + \Delta u_{\mathcal{D}}\|_{L^2(K_D)}^2 + \sum_{\{e \in \mathcal{E}(\mathcal{D}) : e \subset \partial K_D \cap \Omega\}} \frac{|e|}{2p_{e,\mathcal{D}}} \|[\![\nabla u_{\mathcal{D}} \cdot \mathbf{n}_e]\!]\|_{L^2(e)}^2 .$$

    induces a factor $\simeq \|p_{\mathcal{D}}\|_\infty^{-2-2\varepsilon}$ in the efficiency estimate.

    Consequently, the number M of iterations in each call of **PDE** for reducing the error by a factor ϱ scales like $M \approx \log \varrho^{-1} \, \|p_{\mathcal{D}}\|_\infty^{2+\varepsilon}$, leading to a convergence analysis that is not p-robust.

  - The saturation property is guaranteed if each marked element is replaced by its four grandchildren, while preserving the polynomial degree.

Equilibrated Flux Estimators

- ***p*-robust convergence:** This can be achieved for **hp-AFEM** in 2D upon resorting to equilibrated flux estimators.

- **Equilibrated flux estimator:** We introduce the following standard notation:

  - Given a partition $\mathcal{D}$ made of triangles K, with vertices $a \in \mathcal{A}_{\mathcal{D}}$, denote by ω_a the star (or patch) of elements containing a.

  - For any such vertex, define the local energy space

$$H^1_*(\omega_a) := \begin{cases} \{v \in H^1(\omega_a) \colon \langle v, 1 \rangle_{\omega_a} = 0\} & a \in \mathcal{A}_{\mathcal{D}}^{\text{int}}, \\ \{v \in H^1(\omega_a) \colon v = 0 \text{ on } \partial\omega_a \cap \partial\Omega\} & a \in \mathcal{A}_{\mathcal{D}}^{\text{bdry}}. \end{cases}$$

  - Define the global and local residuals for the Galerkin solution $u_{\mathcal{D}} \in V_{\mathcal{D}}$

$$r(v) := \langle f, v \rangle_\Omega - \langle \nabla u_{\mathcal{D}}, \nabla v \rangle_\Omega, \qquad r_a(v) = r(\phi_a v).$$

    where ϕ_a is the piecewise linear hat function centered at a.

p-Robust A Posteriori Estimates

- **Upper and lower bounds:**

$$\|\nabla(u-u_{\mathcal{D}})\|_\Omega^2 \leq 3 \sum_{a\in\mathcal{A}_{\mathcal{D}}} \|r_a\|^2_{H^1_*(\omega_a)'}, \quad \|r_a\|_{H^1_*(\omega_a)'} \lesssim \|\nabla(u-u_{\mathcal{D}})\|_{\omega_a} \quad \forall a \in \mathcal{A}_{\mathcal{D}}.$$

p-Robust A Posteriori Estimates

- **Upper and lower bounds:**

$$\|\nabla(u-u_{\mathcal{D}})\|_{\Omega}^2 \le 3 \sum_{a\in\mathcal{A}_{\mathcal{D}}} \|r_a\|^2_{H^1_*(\omega_a)'}, \quad \|r_a\|_{H^1_*(\omega_a)'} \lesssim \|\nabla(u-u_{\mathcal{D}})\|_{\omega_a} \quad \forall a \in \mathcal{A}_{\mathcal{D}}.$$

- **p-robust equivalence:**

$$\|r_a\|_{H^1_*(\omega_a)'} \simeq \|\boldsymbol{\sigma}_a\|_{\omega_a}$$

where $\boldsymbol{\sigma}_a \in \mathcal{RT}(\mathcal{D}_a)$ is a suitable equilibrated flux for $u_{\mathcal{D}}$ (i.e., it satisfies $\langle \nabla\cdot\boldsymbol{\sigma}_a, 1\rangle_T = \langle f, 1\rangle_T$ for all $T \subset \omega_a$),
[Braess, Pillwein and Schöberl (2009)]

p-Robust A Posteriori Estimates

- **Upper and lower bounds:**

$$\|\nabla(u-u_{\mathcal{D}})\|_{\Omega}^2 \leq 3 \sum_{a\in\mathcal{A}_{\mathcal{D}}} \|r_a\|^2_{H^1_*(\omega_a)'}, \qquad \|r_a\|_{H^1_*(\omega_a)'} \lesssim \|\nabla(u-u_{\mathcal{D}})\|_{\omega_a} \quad \forall a \in \mathcal{A}_{\mathcal{D}}.$$

- **p-robust equivalence:**

$$\|r_a\|_{H^1_*(\omega_a)'} \simeq \|\boldsymbol{\sigma}_a\|_{\omega_a}$$

where $\boldsymbol{\sigma}_a \in \mathcal{RT}(\mathcal{D}_a)$ is a suitable equilibrated flux for $u_{\mathcal{D}}$ (i.e., it satisfies $\langle \nabla\cdot\boldsymbol{\sigma}_a, 1\rangle_T = \langle f, 1\rangle_T$ for all $T \subset \omega_a$),
[Braess, Pillwein and Schöberl (2009)]

- **Computability.** A particular equilibrated flux $\boldsymbol{\sigma}_a$ can be efficiently computed by solving a discrete saddle-point problem in the space $\mathcal{RT}(\mathcal{D}_a)$.
[Ern and Vohralík (2015)]

- **Upper and lower bound for discrete functions:** If $\widetilde{\mathcal{D}} \geq \mathcal{D}$ yields $V_{\mathcal{D}} \subset V_{\widetilde{\mathcal{D}}}$, then

$$\|\nabla(u_{\widetilde{\mathcal{D}}} - u_{\mathcal{D}})\|_{\Omega}^2 \leq 3 \sum_{a \in \mathcal{A}_{\mathcal{D}}} \|r_a\|^2_{(H^1_*(\omega_a) \cap V_{\widetilde{\mathcal{D}}}(\omega_a))'},$$

and

$$\|r_a\|_{(H^1_*(\omega_a) \cap V_{\widetilde{\mathcal{D}}}(\omega_a))'} \lesssim \|\nabla(u_{\widetilde{\mathcal{D}}} - u_{\mathcal{D}})\|_{\omega_a} \qquad \forall a \in \mathcal{A}_{\mathcal{D}}.$$

- **Upper and lower bound for discrete functions:** If $\widetilde{\mathcal{D}} \geq \mathcal{D}$ yields $V_{\mathcal{D}} \subset V_{\widetilde{\mathcal{D}}}$, then

$$\|\nabla(u_{\widetilde{\mathcal{D}}} - u_{\mathcal{D}})\|_{\Omega}^2 \leq 3 \sum_{a \in \mathcal{A}_{\mathcal{D}}} \|r_a\|^2_{(H^1_*(\omega_a) \cap V_{\widetilde{\mathcal{D}}}(\omega_a))'},$$

and

$$\|r_a\|_{(H^1_*(\omega_a) \cap V_{\widetilde{\mathcal{D}}}(\omega_a))'} \lesssim \|\nabla(u_{\widetilde{\mathcal{D}}} - u_{\mathcal{D}})\|_{\omega_a} \qquad \forall a \in \mathcal{A}_{\mathcal{D}}.$$

- **Marking:** Suppose we apply a *star-based Dörfler marking*, and that for any marked star we can find a local space $V_{\widetilde{\mathcal{D}}}(\omega_a) \supset V_{\mathcal{D}}(\omega_a)$ for which it holds

$$\|r_a\|_{H^1_*(\omega_a)'} \lesssim \|r_a\|_{(H^1_*(\omega_a) \cap V_{\widetilde{\mathcal{D}}}(\omega_a))'} \qquad \text{uniformly in } p.$$

- **Upper and lower bound for discrete functions:** If $\widetilde{\mathcal{D}} \geq \mathcal{D}$ yields $V_{\mathcal{D}} \subset V_{\widetilde{\mathcal{D}}}$, then

$$\|\nabla(u_{\widetilde{\mathcal{D}}} - u_{\mathcal{D}})\|_{\Omega}^2 \leq 3 \sum_{a \in \mathcal{A}_{\mathcal{D}}} \|r_a\|^2_{(H^1_*(\omega_a) \cap V_{\widetilde{\mathcal{D}}}(\omega_a))'},$$

  and

$$\|r_a\|_{(H^1_*(\omega_a) \cap V_{\widetilde{\mathcal{D}}}(\omega_a))'} \lesssim \|\nabla(u_{\widetilde{\mathcal{D}}} - u_{\mathcal{D}})\|_{\omega_a} \qquad \forall a \in \mathcal{A}_{\mathcal{D}}.$$

- **Marking:** Suppose we apply a *star-based Dörfler marking*, and that for any marked star we can find a local space $V_{\widetilde{\mathcal{D}}}(\omega_a) \supset V_{\mathcal{D}}(\omega_a)$ for which it holds

$$\|r_a\|_{H^1_*(\omega_a)'} \lesssim \|r_a\|_{(H^1_*(\omega_a) \cap V_{\widetilde{\mathcal{D}}}(\omega_a))'} \qquad \text{uniformly in } p.$$

- **Saturation property:** Then, we immediately obtain the p-robust saturation property

$$\|\nabla(u - u_{\mathcal{D}})\|_{\Omega}^2 \lesssim \|\nabla(u_{\widetilde{\mathcal{D}}} - u_{\mathcal{D}})\|_{\Omega}^2$$

- **Contraction property:** This implies the following bound with $\varrho < 1$

$$\|\nabla(u - u_{\widetilde{\mathcal{D}}})\|_{\Omega} \leq \varrho \|\nabla(u - u_{\mathcal{D}})\|_{\Omega}$$

Checking the saturation condition

- **Reduction to a reference domain.** The problem of verifying

$$\|r_a\|_{H^1_*(\omega_a)'} \lesssim \|r_a\|_{(H^1_*(\omega_a)\cap V_{\widetilde{\mathcal{D}}}(\omega_a))'}$$

for a suitable $V_{\widetilde{\mathcal{D}}}(\omega_a) \supset V_{\mathcal{D}}(\omega_a)$ can be reduced to the problem of establishing, in a reference domain, norm equivalences between the exact and the Galerkin solution of certain elliptic problems with polynomial data p, assuming that the Galerkin solution is a polynomial of suitable degree $q > p$.

Checking the saturation condition

- **Reduction to a reference domain.** The problem of verifying

$$\|r_a\|_{H^1_*(\omega_a)'} \lesssim \|r_a\|_{(H^1_*(\omega_a)\cap V_{\widetilde{\mathcal{D}}}(\omega_a))'}$$

for a suitable $V_{\widetilde{\mathcal{D}}}(\omega_a) \supset V_{\mathcal{D}}(\omega_a)$ can be reduced to the problem of establishing, in a reference domain, norm equivalences between the exact and the Galerkin solution of certain elliptic problems with polynomial data p, assuming that the Galerkin solution is a polynomial of suitable degree $q > p$.

- **A prototypal problem is as follows**:

Let $\hat{E}$ be a reference triangle or square. For any given $g \in \mathbb{P}_p(\hat{E})$, let $u = u(g) \in \hat{V} := H^1_0(\hat{E})$ be the solution of

$$\int_{\hat{E}} \nabla u \cdot \nabla v = \int_{\hat{E}} g\, v \qquad \forall v \in H^1_0(\hat{E}),$$

and let $u_q = u_q(g) \in \hat{V}_q := H^1_0(\hat{E}) \cap \mathbb{P}_q(\hat{E})$ be the solution of

$$\int_{\hat{E}} \nabla u_q \cdot \nabla v = \int_{\hat{E}} g\, v \qquad \forall v \in H^1_0(\hat{E}) \cap \mathbb{P}_q(\hat{E}).$$

- **Prototypal problem (cont'd)** One seeks a function

$$q = q(p) > p$$

and a constant $C > 0$ independent of g and p such that

$$\|\nabla u\|_{0,\hat{E}} \leq C \, \|\nabla u_{q(p)}\|_{0,\hat{E}}.$$

- **Prototypal problem (cont'd)** One seeks a function

$$q = q(p) > p$$

and a constant $C > 0$ independent of g and p such that

$$\|\nabla u\|_{0,\hat{E}} \leq C \, \|\nabla u_{q(p)}\|_{0,\hat{E}}.$$

- On the reference square $\hat{R} = (0,1)^2$, the result is proven to be true with $q(p) = p + c$ for a suitable constant $c > 0$.

- On the reference simplex $\hat{T}$, there is clear numerical evidence of a similar result (and the proof is under construction).

Outline

Conclusions

- We have considered several adaptive spectral methods, with guaranteed linear or quadratic convergence; we have discussed their optimality properties in terms of cardinality of activated degrees of feedom.
- We have introduced an abstract framework for hp-adaptivity.
- We have presented an algorithm for hp-adaptive approximation, with instance optimality.
- We have considered a general, convergent and nearly-optimal hp-adaptive finite element method, and we have discussed several specific realizations.
- Various extension are waiting:
  - Discontinuous Galerkin (underway)
  - Stokes system
  - anisotropic adaptivity
  - non-symmetric operators
  - ...

Some references

- C. Canuto, R.H. Nochetto and M. Verani, *Adaptive Fourier-Galerkin Methods*, Math. Comp. 83 (2014), 1645–1687.
- C. Canuto, R.H. Nochetto and M. Verani, *Contraction and optimality properties of adaptive Legendre-Galerkin methods: the 1-dimensional case*, Comput. & Math. with Appl. 67 (2014), no.4, 752–770.
- C. Canuto, V. Simoncini and M. Verani, *On the decay of the inverse of matrices that are sum of Kronecker products*, Linear Algebra Appl., 452 (2014), 21–39
- C. Canuto, V. Simoncini and M. Verani, *Adaptive Legendre-Galerkin methods: the multidimensional case*, J. Sci. Comput. 63 (2015), 769–798
- C. Canuto, R.H. Nochetto, R. Stevenson, M. Verani, *Adaptive spectral Galerkin methods with dynamic marking*, SIAM J. Numer. Anal. 54 (2016), 3193–3213
- C. Canuto, R.H. Nochetto, R. Stevenson, M. Verani, *Convergence and optimality of hp-AFEM*, Numer. Math. 135 (2017), 1073-1119
- C. Canuto, R.H. Nochetto, R. Stevenson, M. Verani, *On p-robust saturation for hp-AFEM*, Comput. & Math. with Appl. 73 (2017), 2004–2022