ASC Report No. 45/2012

A Numerical Study of Averaging Error Indicators in p-FEM

P. Dörsek, J.M. Melenk

Institute for Analysis and Scientific Computing
Vienna University of Technology — TU Wien
www.asc.tuwien.ac.at ISBN 978-3-902627-05-6

Most recent ASC Reports

44/2012 *A. Jüngel, I. Stelzer*
Existence analysis of Maxwell-Stefan systems for multicomponent mixtures

43/2012 *M. Feischl, T. Führer, and D. Praetorius*
Adaptive FEM with optimal convergence rates for a certain class of non-symmetric and possibly non-linear problems

42/2012 *L. Banăs, M. Page, D. Praetorius*
A general integrator for the Landau-Lifshitz-Gilbert equation

41/2012 *I. Rachunková, A. Spielauer, S. Staněk, E. Weinmüller*
Positive solutions of nonlinear Dirichlet BVPs in ODEs with time and space singularities

40/2012 *S. Simonov, H. Woracek*
Spectral multiplicity of selfadjoint Schrödinger operators on star-graphs with standard interface conditions

39/2012 *A. Sokolova, H. Woracek*
Congruences of Convex Algebras

38/2012 *M. Langer, H. Woracek*
Stability of the derivative of a canonical product

37/2012 *F. Bruckner, C. Vogler, B. Bergmair, T. Huber, M. Fuger, D. Suess, M. Feischl, T. Führer, M. Page, D. Praetorius*
Combining micromagnetism and magnetostatic Maxwell equations for multiscale magnetic simulations

36/2012 *M. Bukal, E. Emmrich, and A. Jüngel*
Entropy-stable and entropy-dissipative approximations of a fourth-order quantum diffusion equation

35/2012 *W. Auzinger and M. Tutz*
A review of stability and error theory for collocation methods applied to linear boundary value problems

Institute for Analysis and Scientific Computing
Vienna University of Technology
Wiedner Hauptstraße 8–10
1040 Wien, Austria

E-Mail: admin@asc.tuwien.ac.at
WWW: http://www.asc.tuwien.ac.at
FAX: +43-1-58801-10196

ISBN 978-3-902627-05-6

ASC
TU WIEN

A Numerical Study of Averaging Error Indicators in p-FEM

Philipp Dörsek and J. Markus Melenk

Abstract We consider the averaging error indicator in the context of the p-FEM. We explain how a proof of reliability and efficiency might look, and why the error indicator will behave differently than for low order methods. Using two model problems, one with nonsmooth, the other one with smooth solution, we identify appropriate spaces for the averaged fluxes in order to obtain reasonable reliability and efficiency bounds on the averaging error indicator for p-FEM. In particular, averaging over two neighbouring elements using global polynomials of the same polynomial degree as the finite element solution leads to reliability and efficiency up to a factor of order $O(p)$.

1.1 Introduction

The averaging error indicator, also called gradient recovery, superconvergent patch recovery, or Zienkiewicz-Zhu error indicator, going back originally to [17], is a widely used method for gauging errors in finite element methods and steering adaptive mesh refinements. Its main advantage is that it is very simple to compute, requiring only a local averaging of the numerical fluxes. A mathematical analysis in the low order context was performed in [16, 14, 3, 1, 6, 15, 11, 2]. In [7], the proof of reliability was reduced to the existence of approximation operators with certain additional orthogonality properties, and such approximation operators were then constructed for arbitrary, but fixed polynomial degree. It is also stated in [7, p. 991] that the

Philipp Dörsek
ETH Zürich, Rämistrasse 101, CH-8092 Zürich, Switzerland, e-mail: philipp.doersek@math.ethz.ch

J. Markus Melenk
Vienna University of Technology, Wiedner Hauptstraße 8–10, A-1040 Vienna, Austria, e-mail: melenk@tuwien.ac.at

numerical behaviour observed in an hp-adaptive strategy "suggests that those constants depend only moderately on p", where the constants referred to are the reliability and efficiency constants of the averaging error indicator.

It is therefore our aim in this paper to analyse whether the proof for reliability and efficiency in [7] can be carried over to the p-FEM. A counting argument on the degrees of freedom shows quickly that the usual good efficiency estimate (efficiency with constant 1 up to a term of higher order) cannot be expected in the high order setting at least for algebraic rates of convergence, as this would require too many degrees of freedom in the approximation space for the averaged fluxes. Hence, we perform numerical computations for two model problems, one with nonsmooth, the other with smooth solution. Our results suggest that increasing the polynomial degree by one, as is commonly done in the low order context, leads to reasonable results if the averaging is performed over four quadrilateral elements. However, in this case, we observe the p-gap, similarly as in the residual error indicator due to [8, 12], which can be removed using equilibration techniques, see [9].

1.2 The averaging error indicator

Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a bounded polygonal domain, and $f \in \mathrm{H}^{-1}(\Omega)$, where $\mathrm{H}^{-1}(\Omega) = (\mathrm{H}^1_0(\Omega))^*$ are the usual Sobolev spaces. We denote the L^2 norm by $\|u\|_0 := \left(\int_\Omega u^2 \mathrm{d}x\right)^{1/2}$ and the H^1 seminorm by $|u|_1 := \left(\int_\Omega |\nabla u|^2 \mathrm{d}x\right)^{1/2}$, where $|\cdot|$ is the Euclidean norm. Consider for simplicity the Poisson problem with homogeneous Dirichlet boundary conditions,

$$-\Delta u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega; \tag{1.1}$$

the analysis of more general boundary conditions is also possible. Defining $V := \mathrm{H}^1_0(\Omega)$, its weak formulation reads: find $u \in V$ such that

$$a(u, v) = \ell(v) \quad \text{for all } v \in V, \tag{1.2}$$

with

$$a(w, v) := \int_\Omega \nabla w \cdot \nabla v \mathrm{d}x \quad \text{and} \quad \ell(v) := \int_\Omega f v \mathrm{d}x. \tag{1.3}$$

We approximate u from the conforming hp-finite element space $V_N \subset V$, i.e., with a triangulation $\mathcal{T}_N$ of Ω into quadrilaterals and a vector $(p_{N,T})_{T \in \mathcal{T}_N}$ of polynomial degrees, we consider

$$V_N := \left\{ v \in V \colon v|_T \in \mathbb{Q}^{p_{N,T}} \right\}, \tag{1.4}$$

where $\mathbb{Q}^k$ is the usual space of tensor product polynomials of degree k in every component. Then, $u_N \in V_N$ is defined through

$$a(u_N, v_N) = \ell(v_N) \quad \text{for all } v_N \in V_N. \tag{1.5}$$

Let $\Sigma_N \subset \mathrm{H}(\nabla\cdot, \Omega) := \left\{\tau \in (\mathrm{L}^2(\Omega))^2 : \nabla \cdot \tau \in \mathrm{L}^2(\Omega)\right\}$, then the global error indicator is defined by

$$\eta_N := \inf_{\tau_N \in \Sigma_N} \|\tau_N - \nabla u_N\|_0. \tag{1.6}$$

Let $\sigma_N \in \Sigma_N$ denote the uniquely determined argument where the above infimum is attained. If Σ_N is finite-dimensional, it is clear that this quantity can be calculated by solving a system of linear equations.

Proposition 1 (Reliability). *Let $I_N : V \to V_N$ be a linear operator with*

$$|I_N v|_1 \leq C_N |v|_1 \quad \textit{for all } v \in V. \tag{1.7}$$

Assume that $\nabla u \in \mathrm{H}(\nabla\cdot, \Omega)$. Then, the error indicator η defined in (1.6) *satisfies*

$$|u - u_N|_1 \leq (1 + C_N)\eta_N + \sup_{v \in V\setminus\{0\}} \frac{\int_\Omega (f + \nabla \cdot \sigma_N)(v - I_N v)\mathrm{d}x}{|v|_1}. \tag{1.8}$$

Proof. As $\sigma_N \in \Sigma_N \subset \mathrm{H}(\nabla\cdot, \Omega)$, the Galerkin orthogonality yields

$$\begin{aligned} |u - u_N|_1 &= \sup_{v \in V\setminus\{0\}} \frac{a(u - u_N, v)}{|v|_1} = \sup_{v \in V\setminus\{0\}} \frac{a(u - u_N, v - I_N v)}{|v|_1} \qquad (1.9) \\ &= \sup_{v \in V\setminus\{0\}} \frac{\int_\Omega (f + \nabla \cdot \sigma_N)(v - I_N v)\mathrm{d}x}{|v|_1} \\ &\qquad + \frac{\int_\Omega (\sigma_N - \nabla u_N) \cdot \nabla(v - I_N v)\mathrm{d}x}{|v|_1} \\ &\leq \sup_{v \in V\setminus\{0\}} \frac{\int_\Omega (f + \nabla \cdot \sigma_N)(v - I_N v)\mathrm{d}x}{|v|_1} + (1 + C_N)\|\sigma_N - \nabla u_N\|_0. \end{aligned}$$

This proves the claimed estimate. □

Remark 1. The above result suggests to look for a linear operator $I_N : V \to V_N$ such that, ideally, its norm in V is bounded independently of N and, additionally, it has the orthogonality property

$$\int_\Omega w_N (v - I_N v)\mathrm{d}x = 0 \quad \text{for all } v \in V \text{ and } w_N \in W_N, \tag{1.10}$$

where W_N is a sufficiently large discrete space satisfying $\nabla \cdot \Sigma_N \subset W_N$. In this case, we observe

$$\int_\Omega \nabla \cdot \sigma_N (v - I_N v)\mathrm{d}x = 0 \quad \text{for all } v \in V \tag{1.11}$$

and hence

$$|u - u_N|_1 \le C\eta_N + C\gamma_N \inf_{f_N \in W_N} \|f - f_N\|_0, \tag{1.12}$$

i.e., reliability with a generic constant. Here, γ_N is defined by

$$\gamma_N := \sup_{v \in V\setminus\{0\}} \frac{\|v - I_N v\|_0}{|v|_1} \tag{1.13}$$

and usually behaves like $\gamma_N \sim h_N p_N^{-1}$ on quasi-uniform meshes and polynomial degree distributions, i.e., the last term in (1.12) is of higher order compared to $|u - u_N|_1$ if W_N is large enough.

Remark 2. If the polynomial degree is fixed and the mesh is refined, an operator I_N as required above is constructed in [7]. Their construction, however, does not generalise directly to the p-version.

In order to obtain an operator I_N for the p-version, a first step would be to let I_N be the L^2-projection operator onto $\mathbb{Q}^{p_N}$, global polynomials of degree p_N, if we assume that $W_N = \mathbb{Q}^{p_N}$ consists of global polynomials, as well. This assumption makes sense in a pure p-version context on a reasonably coarse mesh. If we ignore the issue of boundary conditions, e.g., by considering a pure Neumann problem, [10, Theorem 2.4] yields that on a quasi-uniform mesh with uniform polynomial degree,

$$\|I_N v\|_1 \le C(p_N + 1)^{1/2}\|v\|_1 \quad \text{for all } v \in \mathrm{H}^1(\Omega); \tag{1.14}$$

see also [13, Theorem 1.3] for a corresponding result for triangular and tetrahedral meshes. In this case, we obtain

$$\int_\Omega w_N(v - I_N v)\mathrm{d}x = 0 \quad \text{for all } v \in V \text{ and } w_N \in W_N. \tag{1.15}$$

Choosing $\Sigma_N := \mathbb{Q}^{(p_N+1)\times p_N} \times \mathbb{Q}^{p_N\times(p_N+1)}$, we observe $\nabla \cdot \Sigma_N \subset W_N$, and hence Proposition 1 yields

$$|u - u_N|_1 \le C(p_N + 1)^{1/2}\eta_N + Cp_N^{-1} \inf_{f_N \in W_N} \|f - f_N\|_0. \tag{1.16}$$

Proposition 2 (Efficiency). *The error indicator η defined in* (1.6) *satisfies*

$$\eta_N \le |u - u_N|_1 + \inf_{\tau_N \in \Sigma_N} \|\tau_N - \nabla u\|_0. \tag{1.17}$$

Proof. We see that

$$\|\sigma_N - \nabla u_N\|_0 \le \inf_{\tau_N \in \Sigma_N} \|\tau_N - \nabla u_N\|_0 \le |u - u_N|_1 + \inf_{\tau_N \in \Sigma_N} \|\tau_N - \nabla u\|_0, \tag{1.18}$$

from which the claim follows. □

Remark 3. In order to ensure efficiency of the error indicator, the gradient L^2 projection error

$$\xi_N := \inf_{\tau_N \in \Sigma_N} \|\tau_N - \nabla u\|_0 \tag{1.19}$$

needs to be small. In the h-version context, [7] shows that ξ_N is indeed of higher order if local averaging over edge patches is done using polynomials of degree p_N. It is unclear whether this is possible when averaging globally, see [7, Remark 4.3].

For the p-version, we cannot hope that ξ_N is of higher order: if u_N is approximated using polynomials of degree p, then, in order that ξ_N is of higher order, we need that Σ_N consists of polynomials of degree $p^{1+\alpha}$ for some $\alpha > 0$. But this is not possible if we simultaneously want to ensure existence of an operator I_N as outlined in Remark 1, as in this case $\dim \Sigma_N$ grows faster than $\dim V_N$, which is incompatible with $I_N \colon V \to V_N$ being orthogonal to $W_N \supset \nabla \cdot \Sigma_N$. However, the following argument lets us hope for efficiency, at least if the convergence is only algebraic and we are prepared to accept a p-gap. Let us restrict ourselves for ease of exposition to Ω being a square and the right-hand side being smooth; general polygonal domains can be treated in a similar fashion. Then, [4, Theorem 2.7 and 2.10] yield the sharp convergence bounds

$$c(1 + p_N)^{-4} \le |u - u_N|_1 \le C(1 + p_N)^{-4}. \tag{1.20}$$

Similarly, as the gradient of a singularity function is again a singularity function, we obtain, assuming $\mathbb{Q}^{p_N} \subset \Sigma_N$, from [4, Theorem 2.7] that

$$\inf_{\tau_N \in \Sigma_N} \|\tau_N - \nabla u\|_0 \le C(1 + p_N)^{-3}. \tag{1.21}$$

Together with (1.17), this implies

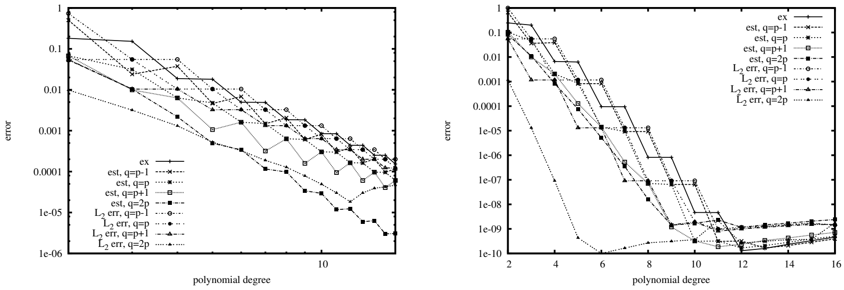
$$\eta_N \le C(1 + p_N)|u - u_N|_1. \tag{1.22}$$

1.3 Numerical examples

For our numerical computations, we consider the square domain $\Omega = (0, \pi)^2$ and solve the homogeneous Dirichlet problem

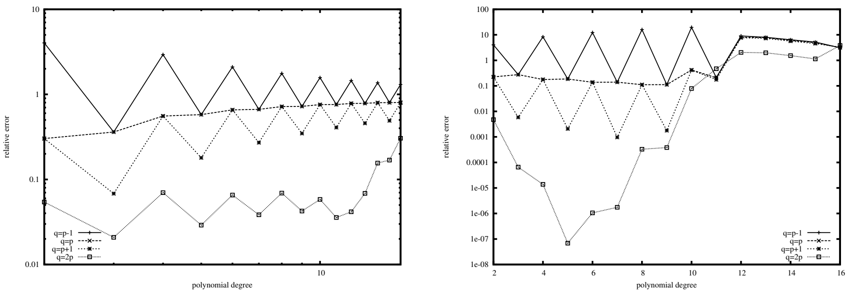
$$-\Delta u = f \quad \text{in } \Omega, \quad u = 0 \quad \text{on } \partial\Omega, \tag{1.23}$$

with the two right-hand sides $f = 1$ and $f(x, y) = 2\sin(x)\sin(y)$. In the first case, the solution is known in terms of a Fourier series, and it is in $\mathrm{H}^{3-\varepsilon}(\Omega)$ for all $\varepsilon > 0$. As the singularities are in the corners of the domain and can therefore be described using the corresponding singularity functions, the rate of convergence is known to be p^{-4}, see [5, Section 4.2], and this is confirmed



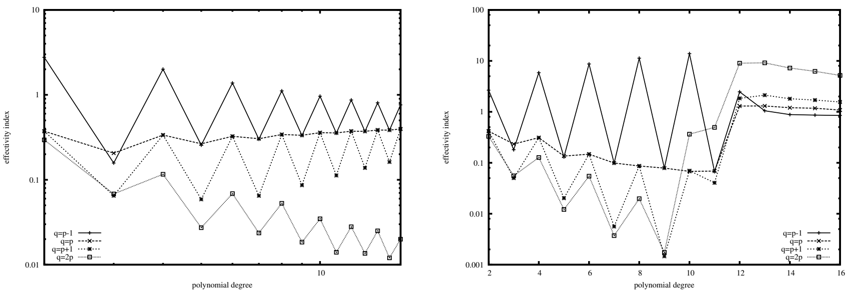
(a) Algebraic rate, two elements (b) Exponential rate, two elements

Fig. 1.1 Errors and error indicators, two elements



(a) Algebraic rate (b) Exponential rate

Fig. 1.2 Gradient L^2 projection error ξ_N relative to Galerkin error, two elements.



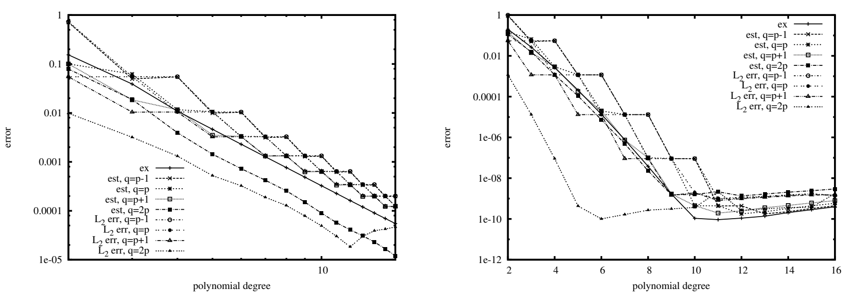
(a) Algebraic rate (b) Exponential rate

Fig. 1.3 Effectivity indices, two elements in Figures 1.1 and 1.4. In the second case, the solution $u(x,y) = \sin(x)\sin(y)$ is analytic, hence the convergence is exponential, and this is also confirmed in Figures 1.1 and 1.4.

We consider two triangulations, one with two quadrilateral elements, $\mathcal{T}_2 = \{(0,\pi/2)\times(0,\pi),(\pi/2,\pi)\times(0,\pi)\}$, and the second with four elements, $\mathcal{T}_4 = \{(0,\pi/2)\times(0,\pi/2),(\pi/2,\pi)\times(0,\pi/2),(0,\pi/2)\times(\pi/2,\pi),(\pi/2,\pi)\times(\pi/2,\pi)\}$. The finite element space is

$$V_N^{(\ell)} := \{v \in V : v|_T \in \mathbb{Q}^{p_N} \text{ for } T \in \mathcal{T}_\ell\}, \quad \ell = 2,4, \tag{1.24}$$

and the approximation space for the averaged fluxes is chosen to be global polynomials. More precisely, we set $\Sigma_N := \mathbb{Q}^{q_N+1,q_N} \times \mathbb{Q}^{q_N,q_N+1}$ with $\mathbb{Q}^{q_1,q_2}$ the space of tensor product polynomials of degree q_1 in the first and q_2 in the second component, i.e., we average the numerical flux over two or four elements using Raviart-Thomas elements. Given p_N, we consider for q_N the values $p_N - 1$, p_N, $p_N + 1$ and $2p_N$. A good choice for q_N should ensure that the effectivity indices do not decay too quickly in p, and that the gradient L^2 projection error in (1.17) is at least not more important than the error.

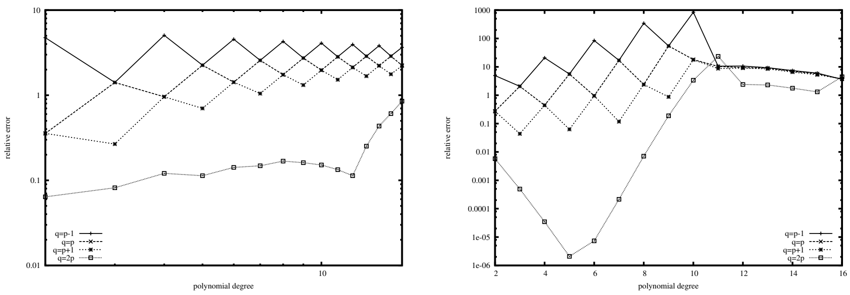


(a) Algebraic rate, four elements (b) Exponential rate, four elements

Fig. 1.4 Errors and error indicators, four elements

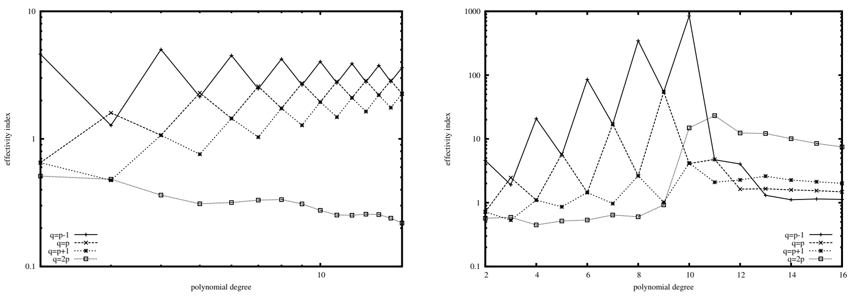
Our experiments show that the gradient L^2 projection error ξ_N of ∇u from Σ_N, see (1.19), is of higher order relative to the Galerkin error only for exponentially decaying error, and even then only for $q_N = 2p_N$. For two elements, the choices $q_N = p_N$ and $q_N = p_N + 1$ at least lead to ξ_N being not larger than the Galerkin error. When averaging over four elements, even that is only achieved using $q_N = 2p_N$.

Let us now turn to the effectivity indices. For two elements, the most reasonable choice is given by $q_N = p_N$; it leads to a reliable and efficient error indicator in the nonsmooth model problem with effectivity indices varying between 0.2 and 0.4, and only to a moderate loss of reliability (of the order



(a) Algebraic rate (b) Exponential rate

Fig. 1.5 Gradient L^2 projection error ξ_N relative to Galerkin error, four elements



(a) Algebraic rate (b) Exponential rate

Fig. 1.6 Effectivity indices, four elements

$O(p)$) in the smooth model problem. Setting $q_N = p_N + 1$ is adequate in the nonsmooth model problem, but the loss of reliability in the smooth model problem is pronounced. For four elements, both choices $q_N = p_N$ and $q_N = p_N + 1$ are reliable in both model problems, but lead to a loss of efficiency (of the order $O(p^{1.35})$ and $O(p^{0.85})$, respectively). The choice $q_N = 2p_N$, finally, leads to a slight loss in reliability in the nonsmooth model problem (of the order $O(p^{0.35})$), and is reliable and efficient in the smooth problem.

1.4 Conclusions

In contrast to low order finite elements, the use of the averaging error indicator in p-FEM leads to certain difficulties. The standard methods of proof cannot be used to obtain reliability and efficiency in the same sense as for the low order case. As explained in Remark 3, the gradient L^2 projection error present in the efficiency estimate cannot be made to be of higher order relative to the Galerkin error.

Averaging the numerical fluxes over two neighbouring quadrilaterals using Raviart-Thomas elements of degree q, reasonable results (reliability up to a factor of the order $O(p)$ and efficiency, i.e., a p-gap) in two model problems are obtained if q is set equal to the local approximation order. This choice is practically the most relevant, as this corresponds to what is known to work in h-FEM and can therefore be expected to be used in hp-FEM. When averaging over four elements, we observe the p-gap when setting $q = p$ or $q = p + 1$. In this case, however, the gradient L^2 projection error in the efficiency estimate even dominates the Galerkin error, which might be of concern theoretically. Finally, averaging over four elements and setting $q = 2p$ leads to an efficient estimator that is reliable up to $O(p^{0.35})$.

Acknowledgements The first author gratefully acknowledges support by the ETH Foundation.

References

1. Ainsworth, M., Craig, A.: A posteriori error estimators in the finite element method. Numer. Math. **60**(4), 429–463 (1992). DOI 10.1007/BF01385730. URL http://dx.doi.org/10.1007/BF01385730
2. Ainsworth, M., Oden, J.T.: A posteriori error estimation in finite element analysis. Pure and Applied Mathematics (New York). Wiley-Interscience [John Wiley & Sons], New York (2000). DOI 10.1002/9781118032824. URL http://dx.doi.org/10.1002/9781118032824
3. Ainsworth, M., Zhu, J.Z., Craig, A.W., Zienkiewicz, O.C.: Analysis of the Zienkiewicz-Zhu a posteriori error estimator in the finite element method. Internat. J. Numer. Methods Engrg. **28**(9), 2161–2174 (1989). DOI 10.1002/nme.1620280912
4. Babuška, I., Guo, B.: Optimal estimates for lower and upper bounds of approximation errors in the p-version of the finite element method in two dimensions. Numer. Math. **85**(2), 219–255 (2000). DOI 10.1007/PL00005387. URL http://dx.doi.org/10.1007/PL00005387
5. Babuška, I., Suri, M.: The p and h-p versions of the finite element method, basic principles and properties. SIAM Rev. **36**(4), 578–632 (1994). DOI 10.1137/1036141. URL http://dx.doi.org/10.1137/1036141
6. Babuška, I.M., Rodríguez, R.: The problem of the selection of an a posteriori error indicator based on smoothening techniques. Internat. J. Numer. Methods Engrg. **36**(4), 539–567 (1993). DOI 10.1002/nme.1620360402
7. Bartels, S., Carstensen, C.: Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. II. Higher order FEM. Math. Comp. **71**(239), 971–994 (electronic) (2002). DOI 10.1090/S0025-5718-02-01412-6
8. Bernardi, C.: Indicateurs d'erreur en h-N version des éléments spectraux. RAIRO Modél. Math. Anal. Numér. **30**(1), 1–38 (1996)

9. Braess, D., Pillwein, V., Schöberl, J.: Equilibrated residual error estimates are p-robust. Comput. Methods Appl. Mech. Engrg. **198**(13-14), 1189–1197 (2009). DOI 10.1016/j.cma.2008.12.010
10. Canuto, C., Quarteroni, A.: Approximation results for orthogonal polynomials in Sobolev spaces. Math. Comp. **38**(157), 67–86 (1982). DOI 10.2307/2007465. URL http://dx.doi.org/10.2307/2007465
11. Carstensen, C., Verfürth, R.: Edge residuals dominate a posteriori error estimates for low order finite element methods. SIAM J. Numer. Anal. **36**(5), 1571–1587 (1999). DOI 10.1137/S003614299732334X
12. Melenk, J.M., Wohlmuth, B.I.: On residual-based a posteriori error estimation in hp-FEM. Adv. Comput. Math. **15**(1-4), 311–331 (2002) (2001). DOI 10.1023/A:1014268310921
13. Melenk, J.M., Wurzer, T.: On the stability of the polynomial L^2-projection on triangles and tetrahedra. Tech. rep., Institute for Analysis and Scientific Computing, Vienna University of Technology (2012)
14. Rank, E., Zienkiewicz, O.: A simple error estimator in the finite element method. Commun. Appl. Numer. Methods **3**, 243–249 (1987). DOI 10.1002/cnm.1630030311
15. Rodríguez, R.: Some remarks on Zienkiewicz-Zhu estimator. Numer. Methods Partial Differential Equations **10**(5), 625–635 (1994). DOI 10.1002/num.1690100509
16. Zienkiewicz, O., Li, X.K., Nakazawa, S.: Iterative solution of mixed problems and the stress recovery procedures. Commun. Appl. Numer. Methods **1**, 3–9 (1985). DOI 10.1002/cnm.1630010103
17. Zienkiewicz, O.C., Zhu, J.Z.: A simple error estimator and adaptive procedure for practical engineering analysis. Internat. J. Numer. Methods Engrg. **24**(2), 337–357 (1987). DOI 10.1002/nme.1620240206