J. Numer. Math., Vol. 0, No. 0, pp. 1–18 (2003)

Prepared using jnm.sty *[Version: 02.02.2002 v1.2]*

Duality-based adaptivity in the *hp*-finite element method

V. Heuveline , R. Rannacher *

26th May 2003

Abstract — In this paper a duality-based a posteriori error analysis is developed for the conforming *hp* Galerkin finite element approximation of second-order elliptic problems. Duality arguments combined with Galerkin orthogonalty yield representations of the error in arbitrary quantities of interest. From these error estimates, criteria are derived for the simultaneous adaptation of the mesh size h and the polynomial degree p. The effectivity of this procedure is confirmed by numerical tests.

Keywords: *hp* finite element method, a posteriori error analysis, adaptivity, error control, DWR method.

1. INTRODUCTION

This paper deals with the numerical solution of second-order elliptic boundary value problems by the conforming *hp*-finite element method on tetrahedral or hexahedral meshes. In this Galerkin method the trial and test spaces consist of functions which are globally continuous and locally polynomial with possibly varying degree. Our main objective is to develop criteria for automatic local adaptation of mesh size h as well as polynomial degree p which are oriented by the goal of the computation. Conformity of the approximation is preserved by using 'hanging nodes' and 'local projection'. The *a priori* analysis of such *hp* finite element approximation is well developed (see, e.g., Schwab [11], and the literature cited therein). An *a posteriori* error analysis following the lines of traditional 'energy-error' estimation was initiated in Babuska and Dorr [2], and has recently been reconsidered in Melenk and Wohlmuth [9] where also hints to further references are given. In this paper, we develop a corresponding analysis for 'goal-oriented' error estimation and adaptivity following the DWR (Dual Weighted Residual) approach described in Becker and Rannacher [5, 6] and Bangerth and Rannacher [4].

For simplicity, we concentrate on the Poisson equation with mixed Dirichlet-Neumann boundary conditions,

$$-\Delta u = f \quad \text{in } \Omega, \qquad u = 0 \quad \text{on } \partial\Omega_D, \quad \partial_n u = g \quad \text{on } \partial\Omega_N, \tag{1.1}$$

*Institut für Angewandte Mathematik, Universität Heidelberg, INF 293, D-69120 Heidelberg, Germany.
This work has been supported by Deutsche Forschungsgemeinschaft (DFG) through SFB 359 'Reaktive Strömungen, Diffusion und Transport'.

on a polyhedral domain $\Omega \subset \mathbb{R}^d$, not necessarily convex, with boundary $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$. All results of this paper can be extended to more general equations with variable coefficients and other boundary conditions as well as to systems of such equations.

Let the goal be to compute a certain output quantity $J(u)$ from the solution u of (1.1) with a prescribed accuracy TOL. Here, we will concentrate on derivative point values and global weighted averages:

$$J(u) = \partial_i u(a), \qquad J(u) = \int_\Omega u(x)\psi(x)\,dx;$$

for more examples see [6] and [4]. To the prescribed (linear) functional $J(\cdot)$, we associate the corresponding 'dual solution' $z \in V$ determined by the auxiliary problem

$$(\nabla\varphi, \nabla z) = J(\varphi) \quad \forall \varphi \in V, \tag{1.2}$$

where $V := \{v \in H^1(\Omega) : v = 0 \text{ on } \partial\Omega_D\}$, and $(\cdot,\cdot) = (\cdot,\cdot)_\Omega$ denoting the usual L^2-inner product. Some regularization may be required to ensure that $J \in V^*$. Let u_h be the computed solution in some finite element space V_h^p defined on a mesh $\mathbb{T}_h$ with cells K of varying size h_K and varying polynomial degree p_K. With this notation, the key step of our analysis is the error representation

$$J(u) - J(u_h) = \sum_{K\in\mathbb{T}_h} \big\{ (R_h, z - I_h^p z)_K + (r_h, z - I_h^p z)_{\partial K} \big\}, \tag{1.3}$$

where $R_h := R(u_h)$ and $r_h := r(u_h)$ denote the usual equation and normal-jump residuals of u_h, and $I_h^p z \in V_h^p$ is an arbitrary interpolation. The evaluation of this identity requires guesses of the dual solution z which may be obtained by post-processing the computed solution $z_h \in V_h^p$. This yields an error estimator of the form

$$|J(u) - J(u_h)| \approx \eta(u_h) := \sum_{K\in\mathbb{T}_h} \eta_K^p, \tag{1.4}$$

with local 'error indicators' $\eta_K^p \geqslant 0$. Then, equilibrating these indicators over the mesh $\mathbb{T}_h$,

$$\eta(u_h) \approx \frac{TOL}{N_h}, \quad N_h := \#\{K \in \mathbb{T}_h\},$$

the desired goal $\eta(u_h) \approx \mathrm{TOL}$ is reached. In this process the decission of how to simultaneously adapt h_K and p_K is largely based on heuristic arguments.

The error estimate (1.4) can also be used to predict quasi-optimal distributions of the mesh size h_K and polynomial degree p_K in certain interesting special cases. Using the local approximation properties of p-th-degree polynomials together with local orthogonality arguments, the error estimate (1.4) can be rewritten in the form

$$\eta(u_h) = \sum_{K\in\mathbb{T}_h} h_K^d h_K^{2p} \Phi_K(h,p),$$

where $\Phi_K(h,p)$ are certain weight factors depending on higher-order derivatives of the primal and dual solutions u and z, respectively. Now, assuming that in the limit $TOL \to 0$, the distributed quantities $\Phi_K(h,p)$ approach a certain limit function $\Phi(h,p)$, depending on the regularity of u and z, the error estimator $\eta(u_h)$ may be replaced by an integral term

$$\eta(u_h) \approx \eta(h,p) := \int_\Omega h^{2p}\Phi(h,p,x)\,dx. \tag{1.5}$$

Here, $h = h(x)$ and $p = p(x)$ stand for continuousy distributed mesh-size and polynomial degree functions. Clearly, the above assumption remains largely heuristic since it requires the Galerkin projection to behave locally like the interpolation and the $(p+1)$-th-order derivatives of the Galerkin approximation u_h like those of its limit u. Analogously, the complexity of the approximation, i.e. $N_{h,p} := \dim(V_h^p)$, can be expressed in the form

$$N(h,p) := \int_\Omega p^d h^{-d}\,dx.$$

An 'optimal' distribution of h_K and p_K can now be determined using the calculus of variation. For pure h refinement, with p kept constant, we obtain the formula

$$h(x) = \left(\frac{TOL}{W}\right)^{1/d}\Phi(x)^{-1/(d+2p)}, \qquad W := \int_\Omega \Phi(p,x)^{d/(d+2p)}\,dx. \tag{1.6}$$

This, in turn yields estimates for the achievable effectivity of h-adaptation based on local indicator equilibration $TOL \approx N^{-2p/d}$. The task of predicting optimal distributions for simultaneous hp adaptation turns out to be much more difficult since the quantity $\eta(h,p)$ depends exponentially on p. However, for certain interesting special cases, including pointwise error estimation and corner singularities, in which the density function $\Phi(h,p)$ can be expressend in terms of negative powers of certain distance functions,

$$\Phi(h,p) \approx A_0 r_0^{-m_0} + A_1 r_1^{-m_1},$$

we are able to derive asymptotic formulas for h as well as p. These predictions for 'optimal' distributions of h and p are well confirmed by some numerical tests. The conclusion of our analysis is that the concept of the DWR method for goal-oriented adaptivity is effective also in the context of the hp finite element method.

2. THE *HP* FINITE ELEMENT METHOD

Let $\mathbb{T}_h$ be decompositions of Ω into (closed) cells K, triangles or quadrilaterals in 2*D*, and tetrahedra or hexahedra in 3*D*, with boundaries ∂K. The cells are assumed to satisfy the uniform shape condition (see Brenner and Scott [7]), but are allowed to vary in size for local mesh adaptation. In particular, the meshes are not required to be 'conforming', i.e., cells may possess 'hanging' nodes in order to facilitate local mesh refinement. We set $h_K := \mathrm{diam}(K)$. Further, to each cell $K \in \mathbb{T}_h$, we associate a parameter p_K indicating the local polynomial degree of the trial and test functions on this cell. On the mesh $\mathbb{T}_h$, we define the finite element space

$$V_h^p \;:=\; \{v_h \in V : v_{h|K} \in P(K),\, K \in \mathbb{T}_h\},$$

where $P(K)$ is a space of functions which are generated by linear or d-linear transformation from a reference cell $\hat{K}$ of a space of polynomials $\hat{P}(\hat{K})$ containing the full polynom spaces

$P_{p_K}(\hat{K})$ in the case of triangles or tetrahedrals, or the full tensorproduct polynom space $Q_{p_K}(\hat{K})$ in the case of quadrilaterals or hexahedrals; for details of this construction we refer to the literature, e.g., Babuska and Dorr [2], and Schwab [11]. For practical purposes, we impose the restriction that for neighboring cells K, K' the cell-width h and polynomial degree p can only vary within certain limits. Then, the discrete solution $u_h \in V_h^p$ is determined by

$$(\nabla u_h, \nabla \varphi_h) = (f, \varphi_h) + (g, \varphi_h)_{\partial\Omega_N} \quad \forall \varphi_h \in V_h^p. \tag{2.1}$$

In the case of constant distribution of p_K, we have the following energy-norm a priori error estimate:

$$\|\nabla(u - u_h)\| \leqslant c_p h^p \|\nabla^{p+1} u\|. \tag{2.2}$$

where $c_p \approx 1/p!$. Throughout, we use the symbol c for a generic positive constant which may vary with the context, but is always independent of h and p.

3. A POSTERIORI ERROR ANALYSIS

We will now derive *a posteriori* estimates for the hp finite element scheme (2.1). Let $J \in V^*$ be an arbitrary (for simplicity *linear*) functional on V with respect to which the error $e = u - u_h$ is to be estimated. Let $z \in V$ be the associated dual solution defined by (1.2). Taking $\varphi = e$ as a test function in (1.2), and using Galerkin orthogonality, we obtain the error representation

$$J(e) = (\nabla e, \nabla z) = (\nabla e, \nabla(z - \psi_h)),$$

for arbitrary $\psi_h \in V_h^p$. Integrating cell-wise by parts, we rewrite this in the form

$$J(e) = \sum_K \{(R_h, z - \psi_h)_K + (r_h, z - \psi_h)_{\partial K}\}, \tag{3.1}$$

with the following notation of cell and edge residuals:

$$R_{h|K} := f + \Delta u_h, \qquad r_{h|\Gamma} := \begin{cases} \frac{1}{2}[\partial_n u_h], & \text{if } \Gamma \subset \partial K \backslash \partial\Omega, \\ g, & \text{if } \Gamma \subset \partial\Omega_N, \\ 0, & \text{if } \Gamma \subset \partial\Omega_D. \end{cases}$$

where $[\partial_n u_h]$ denotes the jump of the normal derivative across a cell boundary. In evaluating the error representation (3.1), we choose the special interpolation $I_h^p z \in V_h^p$ for ψ_h. This is locally defined on each cell $K \in \mathbb{T}_h$ by requiring that $(I_h^p z - z)(a) = 0$ at vertices $a \in \partial K$ (for $p \geqslant 1$), and additionally

$$\int_\Gamma (I_h^p z - z)\chi \, ds = 0 \ \forall \chi \in Q_{p-2}(\Gamma), \quad \int_K (I_h^p z - z)\chi \, ds = 0 \quad \chi \in Q_{p-2}(K),$$

on faces $\Gamma \subset \partial K$ (for $p \geqslant 2$). From the error representation (3.1), we immediately deduce the following result.

Proposition 3.1. *For the error $e = u - u_h$ in the hp finite element scheme (2.1), there holds the a posteriori error estimate*

$$J(e) = \sum_{K\in\mathbb{T}} \left\{ \rho_K^{(1)} \omega_K^{(1)} + \rho_K^{(2)} \omega_K^{(2)} \right\}, \tag{3.2}$$

with the cell residuals $\rho_K^{(i)}$ and weight factors $\omega_K^{(i)}$ being defined by

$$\rho_K^{(1)} = \|R_h - I_h^{p-2} R_h\|_K, \qquad \omega_K^{(1)} = \|z - I_h^p z\|_K,$$
$$\rho_K^{(2)} = h_K^{-1/2} \|r_h - I_h^{p-2} r_h\|_{\partial K}, \qquad \omega_K^{(2)} = h_K^{1/2} \|z - I_h^p z\|_{\partial K}.$$

Theorem 3.1 provides a posteriori estimates for arbitrary functionals of the error. However, splitting the contribution from the cell-residuals R_h from that of the face-residuals r_h usually causes severe overestimation of the error. Therefore, we prefer to directly use the error representation (3.1) as the basis for error control. Its evaluation requires us to generate approximations to the dual solution z with accuracy better than that of the simple Ritz projection $z_h \in V_h^p$ obtained from

$$(\nabla\varphi_h, \nabla z_h) = J(\varphi_h) \quad \forall \varphi_h \in V_h^p. \tag{3.3}$$

This is usually obtained by post-processing z_h using local higher-order interpolation or defect correction. Numerical experience shows that, at least for standard elliptic problems, this approximation is not critical for the method. Let the post-processed approximation to z be denoted by $\tilde{z}_h$. Then the error is controlled by the error estimator

$$\eta_\omega(u_h) := \left| \sum_{K\in\mathbb{T}} \left\{ (R_h, \tilde{z}_h - I_h^p \tilde{z}_h)_K + (r_h, \tilde{z}_h - I_h^p \tilde{z}_h)_{\partial K} \right\} \right|, \tag{3.4}$$

while the adaptation process is driven by the local error indicators

$$\eta_K := \left| (R_h, \tilde{z}_h - I_h^p \tilde{z}_h)_K + (r_h, \tilde{z}_h - I_h^p \tilde{z}_h)_{\partial K} \right| \tag{3.5}$$

Our numerical tests indicate that the estimator $\eta_\omega(u_h)$ is indeed asymptotically sharp as $TOL \to 0$.

3.1. The adaptation process

On the basis of the a posteriori error estimator $\eta(u_h)$ in (3.4) and the resulting local error indicators η_K, 'optimal' distributions of h_K and p_K are constructed by a series of adaptation cycles such that at the final stage the equilibration property

$$\eta_K \approx \frac{TOL}{N_h}, \qquad N_h = \#\{K \in \mathbb{T}_{\text{opt}}\}, \tag{3.6}$$

is achieved. Here, we follow standard adaptation strategies described in the literature, e.g., Ainsworth and Senior [1], Oden, Patra and Feng [10], and Melenk and Wohlmuth [9].

Let a tolerance TOL be given. In solving a stati onary problem the adaptation process usually starts from a coarse mesh $\mathbb{T}_h^{(0)}$, $k = 1,2,\dots$, with mesh-size distribution $h^{(0)}$ and minimum polynomial degree $p^{(0)} \equiv 1$. Then, a sequence of meshes $T_h^{(k)}$, $k = 1,2,\dots$, with corresponding distributions $h_K^{(k)}$ and $p_K^{(k)}$ is constructed by the following process:

1. On the current mesh $\mathbb{T}_h^{(k)}$ compute $u_h^{(k)}$ and $z_h^{(k)}$, and evaluate $\eta_\omega(u_h^{(k)})$ and the cell-error indicators η_K. If $\eta(u_h^{(k)}) < TOL$, then STOP.

2. Order cells according to the size of η_K. On each cell $K \in \mathbb{T}_h^{(k)}$ test whether

$$\eta_K < \frac{1}{2}\frac{TOL}{N_h}\,? \qquad N_h = \#\{K \in \mathbb{T}_h^{(k)}\}.$$

   If YES, procede to next cell. If NO, consider the following three cases:

   - Cell K and the polynomial degree p_K had been left unchanged in the preceding cycle. Then, leave K again unchanged but increase p_K to $p_K + 1$.
   - Cell K had been left unchanged in the preceding cycle, but p_K had been increased. Check whether

$$\eta_K < h_K \eta_K^{\text{old}}.$$

     If YES, then again increase p_K to $p_K + 1$. If NO, then refine K into 2^d cells.
   - Cell K had been obtained by refinement of a mother cell $K_m \in \mathbb{T}_h^{(k-1)}$ (without changing the polynomial degree). Check whether

$$\eta_K \leqslant 2^{-p_K} \eta_{K_m}^{\text{old}}\,?$$

     If YES, increase p_K to $p_K + 1$ and keep h_K. If NO, keep p_K and refine K into 2^d cells.

3. Usually the local changes of h_K and p_K cause additional hanging nodes and discontinuity of functions. In order to restore conformity (in the sense defined above), some neighboring cells need to be refined and their polynomial degree raised.

We note that in this adaptation process, we try to raise p_K whenever possible. The philosophy underlying this rule is that usually for enhancing accuracy on a cell K it is more economical to raise p_K rather than to reduce h_K.

3.2. Mesh optimization

The information contained in the error estimate (3.4) may be used directly to construct 'optimal' h and p distributions for which the required error tolerance TOL is achieved at minimal cost. However, this ambitious goal turns out to be almost impracticable as it requires super-accurate information on the dual solution z which is rarely available on the current meshes. Therefore, we will use the error representation (3.1) only as basis of a heuristic argument for predicting the complexity of hp adaptation in certain special cases.

A mesh $\mathbb{T}_h = \{K\}$ is characterized by a distributed *mesh-size function* h, such that $h_{|K} \approx h_K$. Correspondingly, the degree p_K of the local finite element Ansatz on cell K is described in terms of a distributed *degree function* p such that $p_{|K} \approx p_K$. Motivated by the local approximation behaviour of hp-finite element *interpolation* (see Schwab [11]), we assume that the terms in the error representation behave like

$$|R(u_h) - I_h^{p-2}R(u_h)| \approx \min_{2\leqslant m\leqslant p+1} \left\{ \frac{h_K^{m-2}}{\max\{p^{m-2},(m-2)!\}} |\nabla^m u| \right\}, \tag{3.7}$$

$$|r(u_h) - I_h^{p-2}r(u_h)| \approx \min_{2\leqslant m\leqslant p+1} \left\{ \frac{h_K^{m-2}}{\max\{p^{m-2},(m-2)!\}} |\nabla^m u| \right\}, \tag{3.8}$$

where $I_h^{-1}R(u_h) := 0$, and

$$z - I_h^p z \approx \min_{0\leqslant m\leqslant p+1} \left\{ \frac{h_K^m}{\max\{p^m, m!\}} |\nabla^m z| \right\}. \tag{3.9}$$

The rigorous justification of these estimates, particularly of replacing the Ritz projection u_h by the continuous solution u, is far beyond the scope of this paper. These relations suggest to replace the discrete error representation (3.1) by a continuous analogue of the form

$$J(e) \approx \eta(h,p) := \int_\Omega h(x)^{2p(x)} \Phi(h,p)(x)\, dx, \tag{3.10}$$

with a certain function $\Phi(h,p)$. We assume that $\Phi > 0$, i.e., we exclude the exceptional case that primal or dual solutions locally coincide with a polynomial of degree p. The function $\Phi(h,p)$ may have strong singularities which will force the mesh-size $h(x)$ to develop a very heterogeneous behavior as $TOL \to 0$.

For varying $p(x)$, we denote by $N_{h,p}$ the number of degrees of freedoms in the finite element space V_h^p:

$$N_{h,p} \approx \sum_{K\in\mathbb{T}_h} p_K^d \approx N(h,p) := \int_\Omega h(x)^{-d} p(x)^d\, dx. \tag{3.11}$$

We take this quantity as a measure for the complexity of the discretization.

3.3. Optimization of h for p fixed

At first, we consider the simple case of constant polynomial degree $p \geqslant 1$, i.e., only the mesh size h is allowed to be varied for satisfying the error tolerance TOL. Further, the density function $\Phi(h,p)$ is supposed to be independent of h, of the form

$$\Phi(x) = |\nabla^{p+1}u(x)|\,|\nabla^{p+1}z(x)|. \tag{3.12}$$

Proposition 3.2. *For constant* p*, the optimization problem*

$$N(h,p) \to \min, \qquad \eta(h,p) = TOL, \quad h \geqslant 0, \tag{3.13}$$

has a unique solution which is given by

$$h(x) = W^{-1/d}\Phi(x)^{-1/(d+2p)}TOL^{1/d}, \tag{3.14}$$

provided that $W := \int_\Omega \Phi(x)^{d/(d+2p)}\,dx < \infty$*. The achievable mesh complexity is given by*

$$N_h := \dim(V_h^p) \approx p^d W^{(d+2p)/2p}TOL^{-d/2p}. \tag{3.15}$$

Proof. (i) At first, we determine possible optimal solutions by employing the classical Euler-Lagrange approach. We introduce the Lagrangian

$$L(h,\lambda) = N(h) + \lambda\{\eta(h) - TOL\}$$

with an unknown parameter $\lambda \in \mathbb{R}$. Notice that p is kept fixed. Then, any optimal mesh-size function h corresponds to stationary point $\{h,\lambda\}$ of L. Differentiation L with respect to h and λ, yields the necessary optimality conditions

$$-dh^{-d-1}p^d + \lambda 2ph^{2p-1}\Phi \equiv 0, \qquad \int_\Omega h^{2p}\Phi\,dx - TOL = 0.$$

Clearly $\lambda > 0$, since $h > 0$ and $\Phi > 0$. We conclude that

$$h = \big(2d^{-1}\lambda p^{1-d}\big)^{-1/(2p+d)}\Phi^{-1/(2p+d)}, \qquad \lambda = \tfrac{1}{2}dp^{d-1}h^{-(2p+d)}\Phi^{-1}.$$

Inserting this value for h into the second optimality condition results in

$$\big(2d^{-1}\lambda p^{1-d}\big)^{-2p/(2p+d)}\int_\Omega \Phi^{d/(2p+d)}\,dx = TOL,$$

and using the abbreviation W,

$$\big(2d^{-1}\lambda p^{1-d}\big)^{-1/(2p+d)} = (W^{-1}TOL)^{1/2p}.$$

Combining these equations, we finally obtain

$$h = (W^{-1}TOL)^{1/2p}\Phi^{-1/(2p+d)}, \qquad \lambda = \tfrac{1}{2}dp^{d-1}(W^{-1}TOL)^{-(2p+d)/2p}.$$

Clearly, the so determined pair $\{h,\lambda\}$ is the only stationary point of $L(h,\lambda)$.

(ii) Next, we show that h is actually a solution of the mesh optimization problem. For any admissible mesh-size function $h > 0$, there holds (by use of Hölder's inequality with $r = (d+2p)/d$ and $s = r/(r-1) = (d+2p)/2p$)

$$\begin{aligned}
W &= \int_\Omega \Phi^{d/(d+2p)}\,dx = \int_\Omega h^{2pd/(d+2p)}\Phi^{d/(d+2p)}h^{-2pd/(d+2p)}\,dx \\
&\leqslant \Big(\int_\Omega h^{2p}\Phi\,dx\Big)^{d/(d+2p)}\Big(\int_\Omega h^{-d}\,dx\Big)^{2p/(d+2p)} \\
&= \eta(h)^{d/(d+2p)}N(h)^{2p/(d+2p)}p^{-2pd/(d+2p)},
\end{aligned}$$

and consequently,

$$N(h) \geqslant W^{(d+2p)/2p} TOL^{-d/2p} p^d.$$

This gives us a lower bound on the optimal value. The proof is completed by showing that this lower bound is actually attained by h:

$$N(h) = \int_\Omega h^{-d} p^d \, dx = \int_\Omega (W^{-1} TOL)^{-d/2p} \Phi^{d/(2p+d)} p^d = W^{d/2p} TOL^{-d/2p} W p^d.$$

□

Remark 3.1. Alternatively to the optimization problem (3.13), we can also consider $\eta(h,p) \to \min$, under the constraints $N(h,p) = N_{\max}$, $h > 0$, which has the solution $h(x) = W^{1/d} N_{\max}^{-1/d} \Phi(x)^{-1/(2p+d)}$.

Remark 3.2. In two dimensions and for bilinear elements, i.e. for $d = 2$ and $p = 1$, the formula for the 'optimal' mesh complexity reduces to $N \approx W^2 TOL^{-1}$, provided that $\sup_{TOL \to 0} W < \infty$. This is the case even for rather 'irregular' functionals $J(\cdot)$ including the evaluation of $J(u) = \partial_i^2 u(a)$ as limit case.

Remark 3.3. The result of Proposition 3.2, namely the necessary condition

$$\tfrac{1}{2} d p^{d-1} h^{-(2p+d)} \Phi^{-1} \equiv \lambda \equiv \tfrac{1}{2} d p^{d-1} (W^{-1} TOL)^{-(2p+d)/2p}$$

and the relation

$$N_h = p^{-d} N(h,p) = W^{(2p+d)/2p} TOL^{-d/2p} = (TOL^{-1} W)^{(2p+d)/2p} TOL$$

imply that the optimal mesh-size distribution is characterized by the *equilibration property*

$$h^{2p+d} \Phi \equiv p^{d-1} \lambda^{-1} = (W^{-1} TOL)^{(2p+d)/2p} = N^{-1} TOL. \tag{3.16}$$

Hence, the strategy of equilibrating the local error indicators η_K in the error estimator (3.4), as used in the *error-balancing strategy*, seems justified.

3.4. Simultaneous optimization of h and p

Next, we consider the more complicated case of spatially varying polynomial degree $p(x)$. To simplify the analysis, we have to assume certain properties for the derivatives of the primal and dual solutions such as they are to be expected in prototypical situations. We restrict us to the two-dimensional case. Since it can be expected that the local hp adaptation will be determined by the essential singularities in the primal solution u and the dual solution z, we concentrate on the extreme cases of point value evaluation and corner singularities. Some aspects of this analysis have already been developed in Scott and Rannacher [12].

For computing the solution of the optimization problem

$$N(h,p) \;\to\; \min, \qquad \eta(h,p) = TOL, \quad h > 0, \;\; p > 0, \tag{3.17}$$

we again employ the Euler-Lagrange formalism, while the existence of a solution is assumed to be guaranteed. The Lagrangian is defined by

$$L(h,p;\lambda) := N(h,p) + \lambda\{\eta(h,p) - TOL\},$$

with a scalar $\lambda \in \mathbb{R}$. Taking derivatives of $L(h,p;\lambda)$ with respect to h, p and λ, we obtain for arbitrary variations $\{\varphi,\psi,\mu\}$ of $\{h,p,\lambda\}$ the optimality conditions

$$L'_h(h,p;\lambda) = \int_\Omega \{-dh^{-d-1}p^d + \lambda\big(2ph^{2p-1}\Phi + h^{2p}\Phi'_h\big)\}\varphi\,dx = 0,$$

$$L'_p(h,p;\lambda) = \int_\Omega \{dh^{-d}p^{d-1} + \lambda h^{2p}\big(\ln(h^2)\Phi + \Phi'_p\big)\}\psi\,dx = 0,$$

$$L'_\lambda(h,p;\lambda) = \Big\{\int_\Omega h^{2p}\Phi\,dx - TOL\Big\}\mu = 0,$$

This implies the pointwise relations

$$-dh^{-d-1}p^d + \lambda\big(2ph^{2p-1}\Phi + h^{2p}\Phi'_h\big) = 0, \tag{3.18}$$

$$dh^{-d}p^{d-1} + \lambda h^{2p}\big(\ln(h^2)\Phi + \Phi'_p\big) = 0, \tag{3.19}$$

$$\int_\Omega h^{2p}\Phi\,dx - TOL = 0. \tag{3.20}$$

to be satisfied by any optimal triple $\{h,p,\lambda\}$. The first two of these equations can be simplified to

$$-dh^{-2p-d}p^{d-1} + \lambda\big(2\Phi + hp^{-1}\Phi'_h\big) = 0, \tag{3.21}$$

$$dh^{-2p-d}p^{d-1} + \lambda\big(\ln(h^2)\Phi + \Phi'_p\big) = 0. \tag{3.22}$$

From these equations, we infer that $\lambda \neq 0$ (in fact $\lambda > 0$), and therefore adding them yields

$$2\Phi + hp^{-1}\Phi'_h + 2\ln(h)\Phi + \Phi'_p = 0. \tag{3.23}$$

Assuming that $2\Phi + hp^{-1}\Phi'_h \neq 0$, this is rewritten in the form

$$1 + \ln(h) + \frac{\Phi'_p - \ln(h)hp^{-1}\Phi'_h}{2\Phi + hp^{-1}\Phi'_h} = 0,$$

and taking exponentials,

$$h = \frac{1}{e}\exp\Big(\frac{\Phi'_p - \ln(h)hp^{-1}\Phi'_h}{2\Phi + hp^{-1}\Phi'_h}\Big). \tag{3.24}$$

Further, from (3.21), we obtain

$$\lambda = \frac{dh^{-2p-d}p^{d-1}}{2\Phi + hp^{-1}\Phi'_h}. \tag{3.25}$$

and, consequently,

$$p^{d-1} = \frac{\lambda}{d}h^{2p+d}(2\Phi + hp^{-1}\Phi'_h). \tag{3.26}$$

From these equations for particular cases of $\Phi(h,p)$, we will compute h and p, while λ is determined using additionally the constraint (3.20). This stationary point $\{h,p,\lambda\}$ of L is then assumed to actually be a solution of the hp optimization problem. In contrast to the special situation of Proposition 3.2, the proof of this must be left open in the general case.

3.4.1. Meanvalue of solution with corner singularity. At first, we consider the case of mean-error evaluation on a domain with a reentrant corner. The error functional is

$$J(e) = \int_\Omega e\psi\, dx,$$

with some weight function $\psi \in C_0^\infty(\Omega)$. In the presence of a reentrant corner $x_1 \in \partial\Omega$ with inner angle $\omega > \pi$, in general, the solution contains a leading 'corner singularity' of the form $s_1(r,\vartheta) = r^\alpha \sin(\vartheta\alpha)$, with $r = |x - x_1|$ and $\alpha = \pi/\omega$, and $\{r, \vartheta\}$ denoting polar coordinates at x_1. Hence, we assume that u behave just like this singular function:

$$|\nabla^q u(x)| \approx q!|x - x_1|^{\alpha - q}.$$

In this case the dual solution is smooth but also contains the component r^α. Hence, the singular behaviour of Φ is described by

$$\Phi(h,p)(x) = \sigma(x)^{2(\alpha-p-1)}, \qquad \sigma(x) := |x - x_1|. \tag{3.27}$$

Proposition 3.3. *Let the weight function $\Phi(h,p)$ be given by (3.27), with some $\alpha \geqslant \frac{1}{2}$. Then, the solution of the hp optimization problem has the following properties:*

$$h(x) \;=\; e^{-1}|x - x_1|, \tag{3.28}$$

$$p(x) \;\approx\; 1 + \ln(TOL^{-1}|x - x_1|^{2\alpha}). \tag{3.29}$$

The complexity of the resulting discretization is $N \;\approx\; \ln(TOL^{-1})^2$.

Proof. For $d = 2$ and $\Phi_h' = 0$ the equations (3.24) and (3.26) reduce to

$$h = e^{-1}\exp\Big(\frac{-\Phi_p'}{2\Phi}\Big) = e^{-1}\exp\Big(\frac{2\ln(\sigma)\sigma^{2(\alpha-p-1)}}{2\sigma^{2(\alpha-p-1)}}\Big) = e^{-1}\sigma, \tag{3.30}$$

and

$$p = \lambda h^{2p+2}\Phi = \lambda h^{2p+2}\sigma^{2(\alpha-p-1)}. \tag{3.31}$$

Using (3.30) in (3.31) yields $2pe^{2p} = 2\lambda e^{-2}\sigma^{2\alpha}$, and therefore

$$p = \tfrac{1}{2}W(2\lambda e^{-2}\sigma^{2\alpha}).$$

Here, $W(\cdot)$ is the so-called 'Lambert W-function' defined by $W(x)e^{W(x)} = x$ (see Corless et al. [8]), which has the asymptotic behavior

$$W(r) \approx r,\;\; r \leqslant 2, \qquad W(r) \approx \ln(r),\;\; r > 2.$$

This implies (3.29). In order to determine λ, we insert the above formulas for h and p in the constraint (3.25), obtaining

$$\begin{aligned} TOL &= \int_\Omega h^{2p}\Phi(h,p)\, dx = \int_\Omega e^{-2p}\sigma^{2(\alpha-1)}\, dx = \lambda^{-1}e^2\int_\Omega p\sigma^{-2}\, dx \\ &= \tfrac{1}{2}\lambda^{-1}e^2\int_\Omega W(2\lambda e^{-2}\sigma^{2\alpha})\sigma^{-2}\, dx \end{aligned}$$

or

$$\lambda = \tfrac{1}{2} TOL^{-1} e^2 \int_\Omega W(2\lambda e^{-2}\sigma^{2\alpha})\sigma^{-2}\, dx.$$

Since $W(r)$ is monotone with $W(r) \approx \ln(r)$ for large r, and $TOL^{-1} \gg 1$, this equation determines $\lambda > 0$ and hence also p. With $B_1 := \{x \in \Omega,\, 2\lambda e^{-2}\sigma(x)^{2\alpha} < 1\}$, we obtain

$$TOL \approx \int_{B_1} \sigma^{2(\alpha-1)}\, dx + \tfrac{1}{2}\lambda^{-1} e^2 \int_{\Omega\setminus B_1} \sigma^{-2}\, dx.$$

From this we infer that $\lambda \approx TOL^{-1}$. To estimate the complexity N, we recall that

$$N \approx \int_\Omega h^{-2} p^2\, dx \approx \int_{B_1} \lambda^2 \sigma^{4\alpha-2}\, dx + \int_{\Omega\setminus B_1} \sigma^{-2} \ln(2e^{-2}\lambda\sigma^{2\alpha})^2\, dx,$$

which implies $N \approx \ln(TOL^{-1})^2$. □

Remark 3.4. The remarkable fact, that the optimal mesh-size distribution $h \approx |x - x_1|$ does not depend on the exponent α is in agreement with a priori interpolation results (see Schwab [11]). Further, the strategy of equilibrating the local error indicators η_K used in our adaptation process is justified. Indeed, we have

$$\eta_\omega(u_h) \approx \int_\Omega h^{2p}\sigma^{2(\alpha-p-1}\, dx \approx \sum_{K\in\mathbb{T}_h} h_K^{2p_K+2}\sigma_K^{2(\alpha-p_K-1} \approx \sum_{K\in\mathbb{T}_h} e^{-2p}\sigma_K^{2\alpha}\, dx$$

and from (3.29) the relation $e^{-2p}\sigma^{2\alpha} \approx \text{const}$.

3.4.2. Gradient point values of smooth solution. Next, we consider the case of evaluation of gradient point-values $\nabla u(x_0)$, corresponding to smooth data and regular domain Ω. In this case we expect, based on results of a priori analysis, that the most economical way to achieve high pointwise accuracy is to keep h fixed and to only increase p. This should also be the result of the automatic adaptation process based on the error estimate $|J(e)| \approx \eta(h,p)$. In order to achieve this, we would like to use the error representation

$$\eta(h,p) := \int_\Omega h^{p+1}|\nabla^{p+1}u| \min_{0\leqslant m\leqslant p+1}\left\{\frac{h^m}{\max\{p^m, m!\}}|\nabla^m z|\right\} dx.$$

This seems too difficult as the corresponding function $\Phi(h,p)$ is not differentiable. Therefore, we try to mimic this behavior by using the weight function

$$\Phi(h,p)(x) = \sigma(x)^{-2-p}, \qquad \sigma(x) := |x - x_0| + h. \tag{3.32}$$

Proposition 3.4. *Let the weight function be given by (*3.32*). Then, the solution of the hp optimization problem has the following properties:*

$$h(x) \;\approx\; \begin{cases} e^{-1}, & |x-x_0| \leqslant 1, \\ e^{-1}|x-x_0|^{1/2}, & |x-x_0| > 1. \end{cases} \tag{3.33}$$

$$p(x) \;\approx\; \begin{cases} \ln\left(TOL^{-1}\right), & |x-x_0| \ll 1, \\ \ln(TOL^{-1}|x-x_0|^{-1})\ln(|x-x_0|h^{-2})^{-1}, & |x-x_0| \approx 1, \\ \ln(TOL^{-1}|x-x_0|^{-1}), & |x-x_0| \gg 1. \end{cases} \tag{3.34}$$

The complexity of the resulting discretization is $N \approx \ln(TOL^{-1})^2$.

Proof. (i) In the present case, we have

$$\Phi_h'(x) = -(p+2)\sigma^{-p-3}, \qquad \Phi_p'(x) = -\ln(\sigma)\sigma^{-p-2}.$$

Using this in equation (3.23), we obtain

$$2\sigma(x)^{-2-p} - h(p+2)p^{-1}\sigma^{-p-3} + 2\ln(h)\sigma(x)^{-2-p} - \ln(\sigma)\sigma^{-p-2} = 0,$$

and after some simplification,

$$2 - h(p+2)p^{-1}\sigma^{-1} + \ln(h^2\sigma^{-1}) = 0.$$

Taking exponentials yields

$$h^2 = e^{-2+h(p+2)p^{-1}\sigma^{-1}}(|x-x_0| + h).$$

From this, we conclude the desired asymptotic behavior $h \approx e^{-1}$ for $|x-x_0| \leqslant 1$, and $h \approx e^{-1}|x-x_0|^{1/2}$ for $|x-x_0| > 1$. Further, from (3.25), setting $x = x_0$, we see that

$$\lambda = \frac{2h^{-2p-2}p}{2\sigma^{-2-p} - hp^{-1}(p+2)\sigma^{-3-p}} = \frac{2h^{-p}p}{2-(p+2)p^{-1}} \approx e^p p,$$

for $p \geqslant 3$, which implies that λ will be large for small TOL.

(ii) To determine p, we recall equation (3.26) which in the present case reduces to

$$p = \tfrac{1}{2}\lambda h^{2p+2}(2\sigma^{-p-2} - hp^{-1}(p+2)\sigma^{-p-3}).$$

Using the result for h implies that $pe^p \approx \lambda$ for $|x-x_0| \ll 1$, $\ln(\sigma h^{-2})pe^{p\ln(\sigma h^{-2})} \approx \lambda\sigma^{-1}\ln(\sigma h^{-2})$ for $|x-x_0| \approx 1$, and $pe^{2p} \approx \lambda|x-x_0|^{-1}$ for $|x-x_0| \gg 1$. Hence,

$$p \approx \begin{cases} W(\lambda), & |x-x_0| \ll 1, \\ W(\lambda\sigma^{-1}\ln(\sigma h^{-2}))\ln(\sigma h^{-2})^{-1}, & |x-x_0| \approx 1, \\ \tfrac{1}{2}W(2\lambda|x-x_0|^{-1}), & |x-x_0| \gg 1, \end{cases}$$

where $W(\cdot)$ is again the Lambert W-function. Since $\lambda \gg 1$, we obtain that

$$p \approx \begin{cases} \ln(\lambda), & |x-x_0| \ll 1, \\ \ln(\lambda\sigma^{-1})\ln(\sigma h^{-2})^{-1}, & |x-x_0| \approx 1 \\ \tfrac{1}{2}\ln(2\lambda|x-x_0|^{-1}), & |x-x_0| \gg 1. \end{cases}$$

To determine λ, we recall the constraint (3.20). Setting $\Omega_0 := \{x \in \Omega, |x-x_0| \leqslant 1\}$ and $\Omega_1 := \{x \in \Omega, |x-x_0| > 1\}$, and using the results for h and p, we obtain

$$\begin{aligned} TOL = \int_\Omega h^{2p}\sigma^{-2-p}\,dx &\approx \int_{\Omega_0} e^{2-p}\,dx + \int_{\Omega_1} e^{-2p}\sigma^{-2}\,dx \\ &\approx \lambda^{-1}\int_{\Omega_0} e^2\,dx + \tfrac{1}{2}\lambda^{-1}\int_{\Omega_1} |x-x_0|\sigma^{-2}\,dx. \end{aligned}$$

Thus, $TOL \approx \lambda^{-1}$. Finally, the complexity is given by

$$N = \int_\Omega h^{-2}p^2\,dx \approx \ln(TOL^{-1})^2.$$

□

4. NUMERICAL TEST

In the following, we describe the results of some computational experiments in order to verify the predictions made in Propositions 3.4 and 3.3 for the 'optimal' distributions of spatial mesh-size h and polynomial degree p. The configuration for this test is chosen as shown in Fig. 1. The computational domain is $\Omega = (-1,1) \times (-1,3)$ possibly with a vertical slit with tip at $(0,0)$.

The error $J(e)$ is estimated by the error estimator $\eta_\omega := \eta_\omega(u_h)$ which is evaluated by the procedure described above. The local adaptation of h and p is done on the basis of the corresponding local error indicators η_K defined in (3.5) using the heuristic strategy as described above. The initial mesh consists of $N_0 = 45$ cells. For the purpose of this particular test, we did not go for maximum efficiency, but rather iterated on each adaptation level as long as necessary in order to achieve satisfactory equilibration of the indicators η_K over the mesh. The quality of the error estimate is expressed as usual in terms of the 'effectivity index'

$$I_{\text{eff}} := |\eta_\omega(u_h)/J(e)|$$

Further, we monitor the number of degrees of freedom $N := \dim V_h^p$, the actual error $J_\omega = J_\omega(e)$ obtained by our method, and the error $J_{\text{E}} = J_{\text{E}}(e)$ obtained by using a standard energy-type error estimator; see Melenk and Wohlmuth [9]:

$$\eta_{\text{E}}(u_h) := \Big(\sum_{K\in\mathbb{T}_h} h_K^2 p_K^{-2} \| f + \Delta u_h\|_K^2 + \tfrac{1}{2} h_K p_K^{-1} \|[\partial_n u_h]\|_{\partial K}^2 \Big)^{1/2}.$$

4.1. The 'semi-singular' case

On the domain $\Omega_0 = (-1,1) \times (-1,3)$, we compute the derivative point value $J(u) := \partial_1 u(x_0)$ at $x_0 = (0.5, 2.5)$. The solution is $u(x) = \sin(\pi(x_1+1)/2)\sin(3\pi(x_2+1)/4)$. As expected the automatic adaptation process keeps the initial mesh unrefined and only raises p. Table 1 shows that for computing point values, hp adaptivity based on the weighted error estimator η_ω is more efficient than that using the energy error estimator η_{E}. Further, the estimator η_ω seems to be asymptotically sharp, i.e., $I_{\text{eff}} \to 1$ as $TOL \to 0$. Also the complexity $N \approx \ln(TOL^{-1})^2$ is well confirmed. Figure 1 shows the predicted decay of p away from the point x_0 enforced by the weighted error indicators.

Table 1. *Computation of $\partial_1 u(x_0)$ with varying h and p for smooth solution.*

N	$J_\omega(e)$	$N/\ln(J_\omega)^2$	$\eta_w(u_h)$	I_{eff}	$J_{\text{E}}(e)$	J_{E}/J_ω
45	$4.45e-01$	68	$1.02e+01$	23.2	$4.45e-01$	1
125	$3.50e-01$	113	$7.40e+00$	21.1	$4.33e-01$	1
233	$2.59e-02$	17	$3.42e-01$	13.2	$7.93e-02$	3
297	$4.19e-03$	10	$3.10e-02$	7.4	$3.59e-02$	8
412	$3.21e-04$	6	$6.75e-04$	2.1	$9.98e-03$	3
581	$3.09e-05$	5	$4.02e-05$	1.3	$5.34e-03$	173
812	$4.32e-06$	5	$5.19e-06$	1.2	$1.34e-03$	310
1113	$4.21e-07$	5	$5.06e-07$	1.1	$3.29e-04$	781

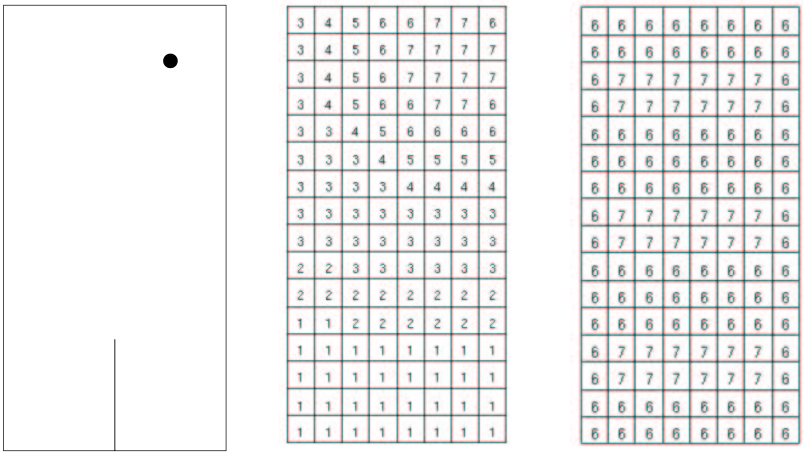

Figure 1. *Configuration of the test problem (left), optimized distribution of p by η_ω (middle), and by η_{E} (right).*

4.2. The fully singular case

On the slit domain $\Omega_1 = (-1,1) \times (-1,3) \setminus \{x \in \mathbb{R}, x_1 = 0, -1 < x_2 < 0\}$ defined above, we compute the derivative point value $J(u) := \partial_1 u(x_0)$. The exact solution is $u(x) = r^{1/2} \sin(\vartheta/2)(x_1 - 1)(x_1 + 1)(x_2 - 3)(x_2 + 1)$. Table 2 indicates that the error estimator η_ω is asymptotically sharp and more efficient for computing the point value $\partial_1 u(P)$ than η_{E}. Further the complexity seems to be $N \approx \ln(TOL^{-1})^2$. The meshes shown in Figures 2 and 3 confirm that the mesh-size behavior is $h \approx |x - x_1|$ as predicted, independet on TOL and p. Finally, Figure 4 shows a comparison of the efficiency of hp adaptation driven by the weighted error estimator η_ω and the plain energy-error estimator η_{E} for the two test cases.

Table 2. *Computation of $\partial_1 u(x_0)$ with varying h and p on the slit domain.*

N	$J_\omega(e)$	$N/\ln(J_\omega)^2$	$\eta_w(u_h)$	I_{eff}	$J_{\text{E}}(e)$	J_{E}/J_ω
740	$2.02e-01$	289	$9.61e+00$	47.6	$8.56e-02$	1
1138	$8.45e-03$	50	$1.96e-01$	23.3	$4.65e-02$	6
1467	$3.45e-03$	45	$4.31e-02$	12.5	$6.48e-03$	2
1736	$8.43e-04$	34	$6.65e-03$	7.9	$8.32e-03$	10
2284	$7.73e-05$	25	$2.47e-04$	3.2	$2.23e-03$	29
2943	$8.72e-06$	22	$2.00e-05$	2.3	$3.32e-04$	38
3752	$8.34e-07$	19	$1.50e-06$	1.8	$8.71e-05$	104
5372	$3.34e-08$	18	$3.60e-08$	1.1	$9.32e-06$	282
6156	$4.43e-10$	15	$4.87e-10$	1.1	$1.37e-07$	309

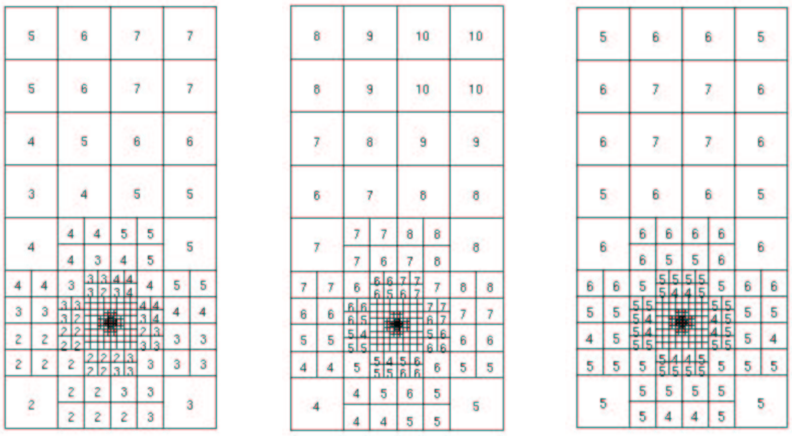

Figure 2. *Optimized h and p by η_ω for $TOL \approx 10^{-6}$ (left) and $TOL \approx 10^{-9}$ (middle), and by η_{E} (right).*

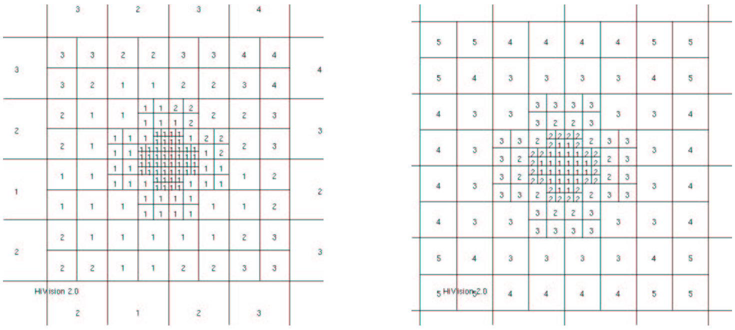

Figure 3. *Zooms into optimized meshes by eta_ω (left) and $\eta_{\rm E}$ (right).*

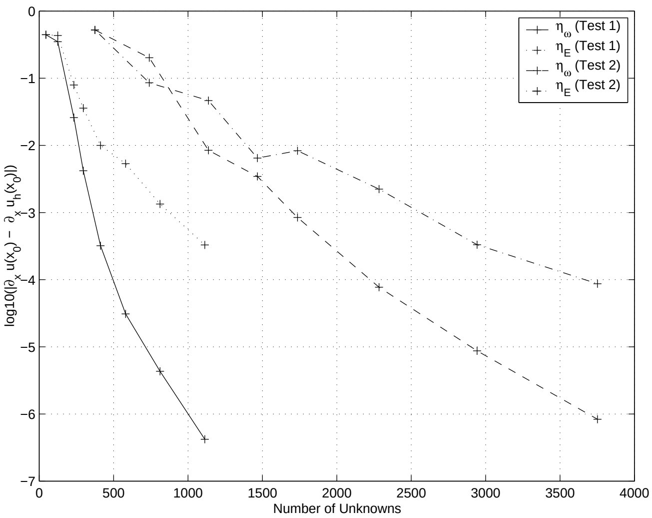

Figure 4. *Efficiency of hp adaptation using the weighted and the energy-error estimators η_ω and $\eta_{\rm E}$.*

REFERENCES

1. M. Ainsworth and B. Senior. An adaptive refinement strategy for *hp*-finite element computations. *Appl. Numer. Math*, 26:165–178, 1998.
2. I. Babuška and M. R. Dorr. Error estimaters for the combined *h*- and *p*-versions of the finite element method. *Numer. Math.*, 37:257–277, 1981.

3. I. Babuška and M. Suri. The p and $h-p$ versions of the finite element method, basic principles and properties. *SIAM Review*, 36:578–632, 1994.

4. W. Bangerth and R. Rannacher. *Adaptive Finite Element Methods for Differential Equations.* Lectures in Mathematics, ETH Zürich, Birkhäuser, 2003.

5. R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: Basic analysis and examples. *East-West J. Numer. Math.*, 4:237–264, 1996.

6. R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. In *Acta Numerica 2001* (A. Iserles, ed.), Cambridge University Press, Cambridge, 2001.

7. S. C. Brenner and R. L. Scott. *The Mathematical Theory of Finite Element Methods*, Springer, Berlin-Heidelberg-New York, 1994.

8. R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the Lambert W-function. *Adv. Comput. Math.*, 5:329–359, 1996.

9. J. M. Melenk and B. I. Wohlmuth. On residual-based a-posteriori error estimation in hp-FEM. *Adv. Comput. Math.*, 15:311–331, 2001.

10. T. J. Oden, A. Patra, and Y. S. Feng. An hp adaptive strategy. In *Adaptive, Multilevel and Hierarchical Computational Strategies* (A. K. Noor, ed.), pp. 23–46, ASME Publications, 1993.

11. C. Schwab. *p- and hp-Finite Elements Methods. Theory and Applications to Solid and Fluid Mechanics*. Oxford University Press, 1998.

12. R. Rannacher and L. R. Scott. *hp optimal adaptivity*. Manuscript, 1997.