A Feed-Back Approach to Error Control in Finite Element Methods: Basic Analysis and Examples

Roland Becker and Rolf Rannacher [1]

Summary. A refined approach to residual-based error control in finite element (FE) discretizations is presented. The conventional strategies for adaptive mesh refinement in FE methods are mostly based on *a posteriori* error estimates in the global energy or L^2-norm involving local residuals of the computed solution. The mesh refinement process then aims at equilibrating these local error indicators. Such estimates reflect the approximation properties of the finite element space by local *interpolation constants* while the stability properties of the continuous model enter through a global *stability constant*, which may be known explicitly in simple cases. Meshes generated on the basis of such global error estimates may not be appropriate in cases of strongly varying coefficients and for the computation of local quantities as, for example, point values or contour integrals. More detailed information about the mechanism of error propagation can be obtained by employing duality arguments specially adapted to the quantity of interest. This results in *a posteriori* error estimates in which the local information derived from the dual solution is used in the form of weights multiplied by local residuals. On the basis of such estimates, a feed-back process in which the weights are numerically computed with increasing accuracy leads to almost optimal meshes for various kinds of error functionals. This approach is developed here for a simple model problem, namely the Poisson equation in two dimensions, in order to emphasize its basic features. However, the underlying concept is rather universal and has, on a heuristic basis, already been successfully applied to much more complex problems in structural and fluid mechanics as well as in astrophysics.

1 Introduction

The strategies for mesh refinement conventionally used in FE methods are mostly based on *a posteriori* error estimates in global norms, e.g., the energy or the L^2-norm, involving local residuals of the computed solution. The mesh refinement process then aims at equilibrating these local error indicators. However, meshes generated on the basis of such *global* error estimates may not be appropriate in cases of strongly varying coefficients and for controlling the accuracy in approximating local quantities as, e.g., point values or contour integrals. For this one needs more detailed information on the mechanism of error propagation which can be obtained by employing suitable duality arguments known from the *a priori* error analysis as the so-called Aubin-Nitsche trick (see, e.g., [11]). The corresponding *dual solutions* then yield the appropriate weight factors to be used in the *a posteriori* error estimates.

Consider, for example, the Poisson equation

$$-\Delta u = f \quad \text{in } \Omega\,, \qquad u = 0 \quad \text{on } \partial\Omega\,, \tag{1.2}$$

on a bounded domain $\Omega \subset \mathbb{R}^2$ which, for the moment, is assumed to be convex polygonal. The restriction to two dimensions is only for simplicity as the extension to arbitrary dimensions is

[1] This work has been supported by the German Research Foundation (DFG), SFB 359 Reactive Flow, Diffusion and Transport, Institute of Applied Mathematics, University of Heidelberg, INF 293, D-69120 Heidelberg, Germany

straightforward. By $(\cdot,\cdot)$ we denote the L^2 inner product and by $\|\cdot\|$ the corresponding norm on Ω. Let (1.2) be discretized by a standard finite element Galerkin method using piecewise polynomial, say, linear or bilinear shape functions on meshes $\mathbb{T}_h = \{K\}$ satisfying the usual regularity conditions (see, e.g., [11]). To facilitate local mesh refinement, hanging nodes may be allowed. For each $\mathbb{T}_h$, let $h_{max} = \max_{K\in\mathbb{T}_h} h_K$, where $h_K = \mathrm{diam}(K)$. The corresponding finite element subspaces are $V_h \subset V := H_0^1(\Omega)$. Their approximation properties are characterized in terms of local approximation estimates (see ([11]),

$$\max\left\{\|v - I_h v\|_K, h_K^{1/2}\|v - I_h v\|_{\partial K}\right\} \le C_{i,K} h_K^{1+r} \|\nabla^{1+r} v\|_{\tilde{K}}\,, \quad r \in \{0,1\}\,, \tag{1.3}$$

for $v \in V \cap H^{1+r}(\Omega)$, where $I_h v \in V_h$ is some locally defined approximation to v, and $\|\cdot\|_B$ denotes the L^2-norm over a set B. For $r = 0$, $\tilde{K}$ is the union of all neighboring elements of K, while for $r = 1$, one simply has $\tilde{K} = K$.

Starting from the variational formulation of (1.2), the finite element method seeks to determine approximations $u_h \in V_h$, such that

$$(\nabla u_h, \nabla \varphi_h) = (f, \varphi_h) \quad \forall\, \varphi_h \in V_h\,. \tag{1.4}$$

Subtracting (1.4) from the variational formulation of (1.2) results in the following orthogonality relation for the error $e = u - u_h$,

$$(\nabla e, \nabla \varphi_h) = 0 \quad \forall\, \varphi_h \in V_h. \tag{1.5}$$

This discretization allows for optimal-order *a priori* estimates in the energy and L^2-norm,

$$\|e\| + h_{max}\|\nabla e\| \le C\, h_{max}^2 \|\nabla^2 u\|, \tag{1.6}$$

provided that the solution u is sufficiently regular. Corresponding estimates also hold with respect to the maximum norm (for references see, e.g., [11]).

For error control in the energy or L^2-norm, one may proceed as follows. Using Galerkin orthogonality (1.5) and integration by parts on each element K yields

$$(\nabla e, \nabla z) = \sum_{K\in\mathbb{T}_h} \left\{(f + \Delta u_h, z - z_h)_K - \tfrac{1}{2}(n\cdot[\nabla u_h], z - z_h)_{\partial K}\right\}, \tag{1.7}$$

for any $z \in V$, where $[\nabla u_h]$ denotes the jump of ∇u_h across the element boundary, and $z_h \in V_h$ is a suitable approximation of z. On edges along the boundary, we set $[\nabla u_h] = \nabla u_h$. Then, using Hölder's inequality on each element, we obtain

$$|(\nabla e, \nabla z)| \le \sum_{K\in\mathbb{T}_h} \rho_K \omega_K\,, \tag{1.8}$$

with local *residuals* ρ_K and *weights* ω_K defined by

$$\rho_K := h_K\|f + \Delta u_h\|_K + \tfrac{1}{2} h_K^{1/2}\|n\cdot[\nabla u_h]\|_{\partial K}\,, \tag{1.9}$$

$$\omega_K := \max\left\{h_K^{-1}\|z - z_h\|_K, h_K^{-1/2}\|z - z_h\|_{\partial K}\right\}. \tag{1.10}$$

By virtue of (1.3), the approximation z_h may be chosen such that for $r \in \{0,1\}$, there holds

$$\omega_K \le C_{i,K} h_K^r \|\nabla^{1+r} z\|_{\tilde{K}}\,. \tag{1.11}$$

Consequently, it follows that

$$|(\nabla e, \nabla z)| \le C_i \Big(\sum_{K \in \mathbb{T}_h} h_K^{2r} \rho_K^2 \Big)^{1/2} \|\nabla^{1+r} z\| \,, \tag{1.12}$$

for $z \in H^{1+r}(\Omega)$. The *interpolation constant* C_i is usually of size $0.1 \le C_i \le 1$, depending on the shape of the elements K. Taking now the supremum over $z \in V \cap H^{1+r}(\Omega)$, one obtains the following *a posteriori* error bound in the energy norm, for $r = 0$, or in the L^2-norm, for $r = 1$,

$$\boxed{\|\nabla^{1-r} e\| \le C_s C_i \Big(\sum_{K \in \mathbb{T}_h} h_K^{2r} \rho_K^2 \Big)^{1/2} =: \eta(u_h) \,.} \tag{1.13}$$

The *stability constant* C_s measures the stability properties of the *dual problem*

$$(\nabla \varphi, \nabla z) = (\nabla^{1-r} e, \nabla^{1-r} \varphi) \quad \forall\, \varphi \in V, \tag{1.14}$$

in terms of the global *a priori* estimate

$$\|\nabla^{1+r} z\| \le C_s \|\nabla^{1-r} e\| \,, \tag{1.15}$$

which, of course, is trivial for $r = 0$. Clearly, by analogous arguments *a posteriori* error estimates may also be derived with respect to L^q-norms including the limit case $q = \infty$.

For earlier work on *a posteriori* error estimates for FE methods, we refer to the pioneering papers of Babuška and Rheinboldt [1], Bank and Weiser [3], and Babuška and Gui [2]. The underlying approaches have been surveyed in Verfürth [28] and [30]. The concept of *a posteriori* error estimation by duality arguments has been developed by Eriksson, Johnson and their co-workers (see [13], [21], [15], and the literature cited therein). Heuristically based error indicators for elliptic problems have been devised, e.g., by Zienkiewicz and Zhu [34].

Based on an *a posteriori* error estimate of the type (1.13) the mesh refinement process may be organized according to the so-called *error per cell strategy*. For some prescribed error tolerance TOL, the goal is to reach a most economical mesh $\mathbb{T}_h$ on which $\eta(u_h) \approx TOL$. Accordingly, the mesh refinement process aims at equilibrating the *local error indicators* $\eta_K := h_K^r \rho_K$ by refining (or coarsening) the elements $K \in \mathbb{T}_h$ according to the criterion

$$\eta_K \approx \frac{TOL}{\sqrt{N} C_i C_s} \,, \qquad N = \#\{K \in \mathbb{T}_h\} \,. \tag{1.16}$$

The refinement strategies based on the conventional *a posteriori* error estimate (1.13) rely on the assumption that the local error indicators $\eta_K := h_K^r \rho_K$ properly describe the dependence of the global error on the local mesh size h_K. However, this may not be true in certain situations since the *a posteriori* error estimate (1.13) contains information about the mechanism of error propagation only through the *global* stability constant C_s. To overcome this deficiency, it has been proposed in [7] to use the quantities ω_K in the estimate (1.8) as weight-factors multiplied by the local residuals ρ_K and to compute them numerically. These weights contain all information about the local approximation properties of the spaces V_h, as well as the local stability properties of the underlying continuous problem. This correspondence can be used for the mesh refinement algorithm through a feed-back process. In the course of the refinement the dual solution is calculated on the current mesh yielding approximate weights. On the basis of the resulting

a posteriori error estimate the mesh is refined according to one of the criteria described below. This process is repeated yielding more and more accurate weights, i.e., *a posteriori* error bounds, until the prescribed stopping criterion is fulfilled. This approach allows one to construct almost optimal meshes for various kinds of error quantities, where *"optimal"* can mean *"most economical for achieving a prescribed accuracy TOL"* or *"most accurate for a given maximal number N_{max} of mesh points"*.

2 Residual based weighted a posteriori error estimates

First, we illustrate our concept at a simple model situation. Let $\Omega = (-1,1)^2$ and the point-value $u(0)$ be the quantity to be computed with best accuracy. It is not likely that the meshes obtained through energy or global L^∞-error control are optimal for this purpose. To do better, we have first to derive an *a posteriori* error estimate for the point-error. This can be done by a duality argument. Let z be the Green function corresponding to the origin, i.e., the solution of the "dual" problem

$$(\nabla\varphi, \nabla z) = \varphi(0) \quad \forall\, \varphi \in V \cap C(\bar{\Omega})\,, \tag{2.2}$$

where, in this special situation, $z(x) \approx \log(|x|)$. Then, we have

$$e(0) = (\nabla e, \nabla(z - z_h)), \tag{2.3}$$

with the nodal interpolant $z_h \in V_h$, appropriately regularized at $x = 0$. Consequently, in view of the error identity (1.7), there holds

$$|e(0)| \le \sum_{K\in\mathbb{T}_h} \omega_K \rho_K\,, \tag{2.4}$$

with residuals ρ_K and weights ω_K as defined above. Using the particular form of the dual solution, and the usual local interpolation properties of finite elements, this error estimator takes the form

$$|e(0)| \le C_i \sum_{K\in\mathbb{T}_h} \frac{h_K^2}{r_K^2}\rho_K =: \eta_{point}(u_h)\,, \tag{2.5}$$

where $r_K := \max\{\mathrm{dist}(K,0), h_K\}$ and $C_i \approx 1$.

Suppose that for some prescribed tolerance TOL the refinement process has led to a final mesh $\mathbb{T}_h^{opt}$ on which the local error indicators $\eta_K = \omega_K\rho_K$ are equilibrated,

$$\eta_K \approx \frac{TOL}{N_{opt}}\,, \qquad N_{opt} = \#\{K \subset \mathbb{T}_h^{opt}\}\,. \tag{2.6}$$

This implies the relation

$$h_K \approx \Big(\frac{TOL\, r_K^2}{N_{opt}}\Big)^{1/4} \tag{2.7}$$

if we assume that $\rho_K \approx h_K^2$. The latter assumption means that on the current mesh, there holds

$$\rho_K \approx h_K^{1/2}\|n\cdot[\nabla u_h]\|_{\partial K} \approx h_K^2 \max_{\partial K}|n\cdot h_K^{-1}[\nabla u_h]| \le c(u)h_K^2\,, \tag{2.8}$$

a hypothesis which is supported by computational experiments (see the the examples, below). As consequence, the number N_{opt} of elements of the final mesh $\mathbb{T}_h^{opt}$ is given by

$$N_{opt} \approx \sum_{K\in\mathbb{T}_h^{opt}} h_K^2 h_K^{-2} = \sum_{K\in\mathbb{T}_h^{opt}} h_K^2 \frac{N_{opt}^{1/2}}{TOL^{1/2} r_K}. \tag{2.9}$$

Since r^{-1} is integrable over Ω, it follows that $N_{opt} \approx \mathrm{TOL}^{-1}$ and the "optimal" mesh-size distribution is $h_K \approx (\mathrm{TOL} r_K)^{1/2}$. In contrast, the global energy-error estimator (1.13), for $r = 0$, would generate meshes of complexity $N_{opt} \approx TOL^{-2}$, in order to guarantee an asymptotic behavior $\|\nabla e\| \approx \mathrm{TOL}$. The corresponding global L^2-error estimator

$$\|e\| \le C_s C_i \Big(\sum_{K\in\mathbb{T}_h} h_K^2 \rho_K^2 \Big)^{1/2} =: \eta_2(u_h) \tag{2.10}$$

seems to be more appropriate as, at least, it has the right asymptotic behavior as $h \to 0$. Alternatively, one may use the L^∞-error estimator which is obtained by an L^∞/L^1-duality argument (see the proof of Lemma 4.2 in Section 4). It takes the form

$$\max_\Omega |e| \le C_i C_s |\log h_0|^{1/2} \max_{K\in\mathbb{T}_h} \rho_{\infty,K} =: \eta_\infty(u_h), \tag{2.11}$$

where h_0 is the diameter of the element containing the point $x = 0$, and

$$\rho_{\infty,K} := h_K^2 \max_K |f + \Delta u_h| + h_K \max_{\partial K} |n \cdot [\nabla u_h]| .$$

This estimator also refers to a global error norm and does not particularly emphasize the point $x = 0$ at which the solution u is to be computed. Therefore, it will not be able to produce a most economical mesh for this particular purpose. However, asymptotically, it yields meshes of the same complexity in terms of TOL, $N_{opt} \approx \mathrm{TOL}^{-1}$, as the point-value estimator $\eta_{point}(\cdot)$. This situation changes when point-values of the gradient $\nabla u(0)$ ("stress values") are to be computed. Then, the global error bounds require $N_{opt} \approx \mathrm{TOL}^{-2}$ cells to achieve a corresponding error of size TOL, while the weighted error estimator achieves this again with only $N_{opt} \approx \mathrm{TOL}^{-1}$ cells (see Example 1, below).

Of course, in more general situations, explicit bounds for the dual solution z are not known and have to be computed numerically. This essential question will be discussed in more detail in Section 4. The simplest approach is based on the local interpolation estimate (1.11) for $r = 1$. Estimates for the weights ω_K may be obtained from the approximate dual solution $\tilde{z}_h \in V_h$, computed on the current mesh, by taking appropriate 2nd-order difference quotients,

$$\omega_K \approx \tilde{\omega}_K := h_K^2 |\nabla_h^2 \tilde{z}_h(x_K)| , \tag{2.12}$$

where x_K denotes the mid-point of element K.

Heuristically, the meaning of our *a posteriori* error estimate is as follows. We expect that the weights ω_K converge to certain limits, as $h \to 0$, in the sense $h_K^{-1/2}\omega_K \to |D^2 z(x_K)|$, where x_K denotes a point contained in an infinite sequence of nested elements K. Further, the quantity $h_K^{-1} n \cdot [\nabla u_h]$ can be viewed as a certain second-order difference quotient of u_h. Hence, it may be expected that, for a sequence of properly refined meshes, the residual terms in (2.4) also converge to certain limits as $h \to 0$,

$$\begin{aligned} |K|^{-1/2}\|f + \Delta u_h\|_K &\to |(f + \Delta u)(x_K)| = 0, \\ |\partial K|^{-1/2}\|h_K^{-1} n \cdot [\nabla u_h]\|_{\partial K} &\to |D^2 u(x_K)|. \end{aligned}$$

In this sense, the *a posteriori* error estimate (2.4) asymptotically, for $h \to 0$, takes the form

$$|e(0)| \approx \sum_{K \in \mathbb{T}_h} h_K^4 |D^2 z(x_K)||D^2 u(x_K)|. \tag{2.13}$$

The local residuals ρ_K consist of two parts, the domain residual $h_K \|f + \Delta u_h\|_K$ and the normal-jump terms $h_K^{1/2} \|n \cdot [\nabla u_h]\|_{\partial K}$. For linear elements, it can be argued that, in the case of smooth f, the contribution of the normal jumps asymptotically dominates that of the domain residual, and the latter may therefore be neglected. This aspect will be discussed in more detail in Section 4. Hence, with the above notation, we are led to the practical error estimator

$$|e(0)| \approx C_i \sum_{K \in \mathbb{T}_h} \tilde{\omega}_K \tilde{\rho}_K =: \eta_{point}^{approx}(u_h)\,, \tag{2.14}$$

where

$$\tilde{\rho}_K := h_K^{1/2} \|n \cdot [\nabla u_h]\|_{\partial K}\,, \qquad \tilde{\omega}_K := h_K^{3/2} \|n \cdot [\nabla \tilde{z}_h]\|_{\partial K}\,, \tag{2.15}$$

and the interpolation constant may be set to $C_i \approx 1$.

Numerical test. The weighted error estimator (2.5) and its approximate version (2.14) have been tested, with $C_i = 0.1$, for the solution $u = 10 \sin(2x_1 + x_2 + 2)$ using bilinear elements. For successively reduced tolerances TOL, the resulting balanced meshes $\mathbb{T}_h^{opt}$ and those obtained by the global L^2-error estimator (2.10) and the L^∞-estimator (2.11) are shown in Figure 1. The corresponding errors are listed in Table 1. Here, L is the number of refinement levels of the final mesh $\mathbb{T}_h^{opt}$ and $I_{eff} := |e(0)|/\eta(u_h)$. For the problem considered, the predicted dependence $N_{opt} \approx \mathrm{TOL}^{-1}$ for the weighted *a posteriori* estimate (2.4) is confirmed. The slight tendency of our error estimator to underestimate the true error could be compensated by using a larger interpolation constant C_i.

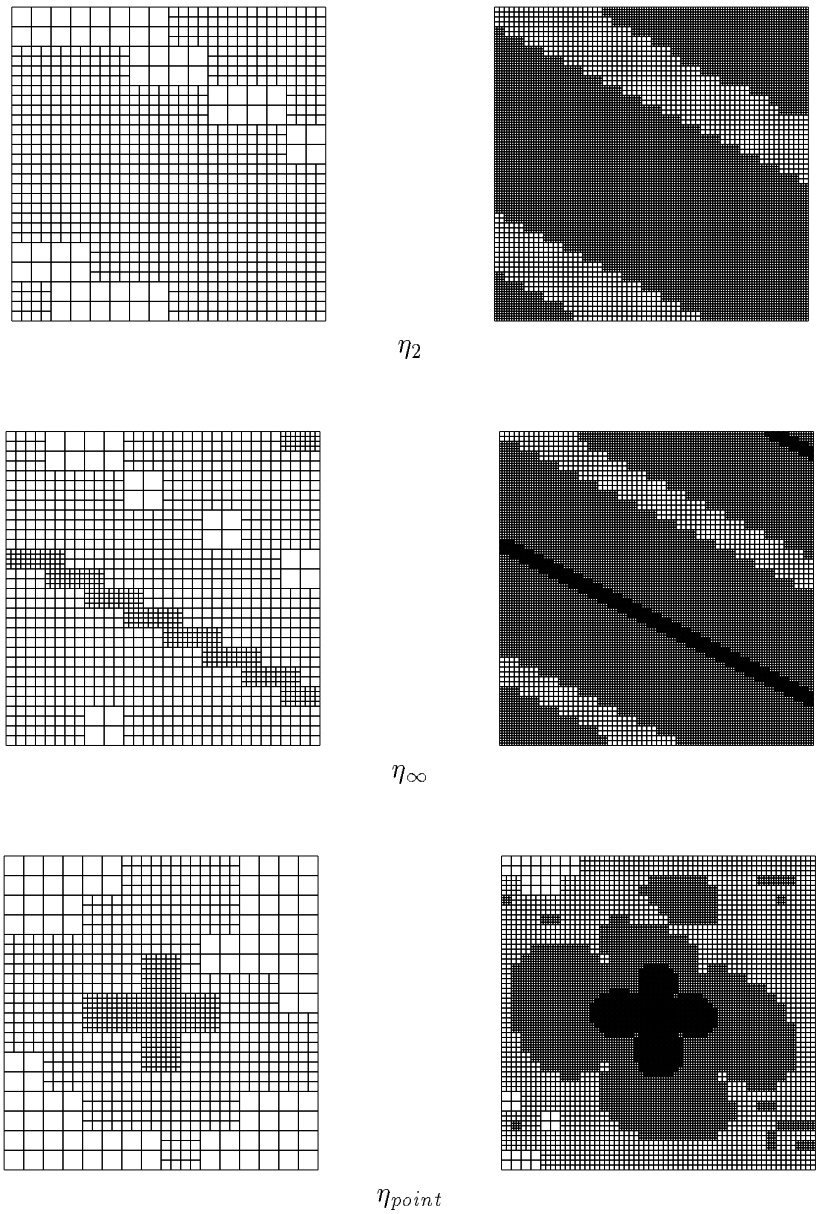

Figure 1: Refined meshes (left to right) for computing $u(0)$ for $u = 10\sin(2x_1 + x_2 + 2)$, using the L^2- and L^∞-error estimators η_2 and η_∞, respectively, compared to the weighted error estimator η_{point}

The above argument directly generalizes to the case of an arbitrary linear error functional $J(\cdot)$ defined on the space V, or on a suitable subspace containing the finite element space V_h and the exact solution u. Other relevant cases are, for instance, the moments of u over Ω (torsion moment), the value of its derivative $\beta \cdot \nabla u(x_0)$ at some point $x_0 \in \bar{\Omega}$ (stress values), or the weighted integral of the normal derivative $\partial_n u$ over the boundary $\partial\Omega$ (total surface tension),

$$J(\varphi) = \int_\Omega \varphi\psi\, dx, \qquad J(\varphi) = \beta \cdot \nabla\varphi(x_0), \qquad J(\varphi) = \int_{\partial\Omega} \partial_n\varphi\, \psi\, ds\,. \tag{2.16}$$

η_2						η_∞					
TOL	N	L	$\lvert e(0)\rvert$	η	I_{eff}	TOL	N	L	$\lvert e(0)\rvert$	η	I_{eff}
1	40	4	2.65e-1	8.62e-1	0.30	1	52	4	3.50e-1	7.53e-1	0.45
4^{-1}	136	5	8.67e-2	2.36e-1	0.37	4^{-1}	196	5	3.19e-2	2.21e-1	0.14
4^{-2}	532	6	2.06e-2	5.98e-2	0.34	4^{-2}	820	6	6.24e-3	4.18e-2	0.15
4^{-3}	2248	7	1.40e-3	1.40e-2	0.10	4^{-3}	2992	7	1.42e-3	1.21e-2	0.12
4^{-4}	8752	8	3.57e-4	3.62e-3	0.10	4^{-4}	11392	8	3.60e-4	3.40e-3	0.11
4^{-5}	33664	9	8.97e-5	9.40e-4	0.10	4^{-5}	43948	9	8.52e-5	8.59e-4	0.10
4^{-6}	133648	10	2.13e-5	2.36e-4	0.09	4^{-6}	180196	10	2.05e-5	1.99e-4	0.10

η_{point}						η_{point}^{approx}					
TOL	N	L	$\lvert e(0)\rvert$	η	I_{eff}	TOL	N	L	$\lvert e(0)\rvert$	η	I_{eff}
1	16	3	3.17e-1	2.72e-1	1.11	1	64	4	7.69e-2	5.34e-0	0.14
4^{-1}	40	4	2.65e-1	1.85e-1	1.43	4^{-1}	88	5	4.33e-2	2.29e-1	0.19
4^{-2}	112	5	1.18e-2	5.28e-2	0.22	4^{-2}	148	6	1.86e-2	4.43e-2	0.42
4^{-3}	400	7	5.17e-3	1.40e-2	0.37	4^{-3}	436	8	9.26e-3	6.67e-3	1.43
4^{-4}	1444	9	1.23e-3	3.82e-3	0.32	4^{-4}	1240	10	1.45e-3	2.13e-3	0.67
4^{-5}	6040	12	4.38e-4	8.96e-4	0.50	4^{-5}	4192	12	8.06e-4	5.77e-4	1.42
4^{-6}	23176	14	1.25e-4	2.31e-4	0.56	4^{-6}	14248	14	4.13e-4	1.62e-4	2.50

Table 1: Results obtained for computing $u(0)$ for $u = 10\sin(2x_1 + x_2 + 2)$, using the L^2- and L^∞-error estimators η_2 and η_∞, respectively, compared to the weighted estimator η_{point} and its approximate analogue η_{point}^{approx}, all with $C_i = 0.1$

Then, following the above approach, we have the *a posteriori* error estimate

$$|J(e)| \leq \sum_{K \in \mathbb{T}_h} \omega_K \rho_K \,, \tag{2.17}$$

with weights ω_K and residuals ρ_K as defined above.

We like to emphasize that, in this particular problem, it is crucial to utilize the approximation property of the finite element space to its full extent also in the evaluation of the weights ω_K, in order to obtain economical meshes. This applies even though there may be no *a priori* bound for the resulting norms of the continuous dual solution z (see [24], [10] and [19]). We will come back to this point in Example 4 at the end of this paper.

In many interesting cases, the functional $J(\cdot)$ is "singular", i.e., not properly defined on $H_0^1(\Omega)$ but only on some subspace containing the solution u and on the finite element space V_h. Typical examples are the evaluation of point-values or of contour integrals of derivatives. In such a case, it is advisable to regularize the functional according to the prescribed tolerance, in order to avoid over-refinement near the singularities. The type of regularization depends on the particular situation and should be such that

$$|J(u) - J_{TOL}(u)| \approx \text{TOL} \,. \tag{2.18}$$

For example, the point-value estimator (2.14) may be used in the regularized form

$$|e(0)| \approx \sum_{K \in \mathbb{T}_h} \frac{h_K^2}{(r_K + \text{TOL})^2} \rho_K \,. \tag{2.19}$$

Further examples of this type will be discussed in Section 5.

Since some of the examples presented below deal with domains with a curved boundary and also higher order finite elements, we have to extend our *a posteriori* error analysis to this situation. Now, the domain Ω is assumed either to have a smooth boundary or to be convex polygonal, in order to ensure that the usual L^2/H^2-shift theorem holds true. Accordingly, along the boundary $\partial\Omega$, the elements K may be curved (isoparametric elements). For the mesh size, we use the same notation h_K, etc., as above. In addition, we now introduce the mesh domain $\Omega_h := \cup\{K \in \mathbb{T}_h\}$ and assume that all nodal points of $\partial\Omega_h$ lie on $\partial\Omega$. This means that, in general, $\Omega_h \not\subset \Omega$ for a non-convex domain. In such a case, one has to cope with a boundary strip $\{\Omega \setminus \Omega_h\} \cup \{\Omega \setminus \Omega_h\}$ of local width $\delta_{h;K} \approx h_K^{m+1}$. The finite element spaces are defined by

$$V_h := \{v \in H_0^1(\Omega_h), v_{|K} \in P(K)\},$$

where $P(K)$ is a suitable vector space of polynomials over K containing $P_m(K)$. They approximate the solution space $V = H_0^1(\Omega)$, such that local approximation properties (1.3) hold for $v \in V \cap H^{r+1}(\Omega \cup \Omega_h)$ with $1 \le r \le m$. We introduce the subset of elements adjacent to the boundary, $\mathbb{T}_h^\partial := \{K \in \mathbb{T}_h, meas(\partial K \cap \partial\Omega_h) \neq 0\}$ and accordingly set $\mathbb{T}_h^0 := \mathbb{T}_h \setminus \mathbb{T}_h^\partial$. To each element $K \in \mathbb{T}_h^\partial$, there naturally belongs a section of the strip $\{\Omega \setminus \Omega_h\} \cup \{\Omega \setminus \Omega_h\}$ denoted by S_K. By a variant of the Poincaré inequality, there holds

$$\|v\|_{S_K} \le C_{i,K} h_K^{m+1} \|\nabla v\|_K . \tag{2.20}$$

In this case the finite element scheme reads

$$(\nabla u_h, \nabla \varphi_h)_{\Omega_h} = (\bar{f}, \varphi_h)_{\Omega_h} \quad \forall\, \varphi_h \in V_h , \tag{2.21}$$

where $\bar{f}$ (if necessary) is a suitable extension of f to $\Omega \cup \Omega_h$. For the following, we extend also the finite element solution u_h as well as the dual solution z by zero to $\Omega_h \cup \Omega$ (in accordance with homogeneous boundary conditions). Using this convention, there holds

$$\begin{aligned}(\nabla e, \nabla z)_\Omega &= (f, z)_\Omega - (\nabla u_h, \nabla z)_{\Omega\cap\Omega_h} - (\nabla u_h, \nabla z)_{\Omega\setminus\Omega_h} \\ &= (f, z)_{\Omega\setminus\Omega_h} + (\bar{f}, z)_{\Omega_h} - (\nabla u_h, \nabla z)_{\Omega_h} \\ &= (f, z)_{\Omega\setminus\Omega_h} + (\bar{f}, z - z_h)_{\Omega_h} - (\nabla u_h, \nabla(z - z_h))_{\Omega_h} ,\end{aligned}$$

from which by integration by parts, we conclude the following analogue of the identity (1.7),

$$\begin{aligned}(\nabla e, \nabla z)_\Omega = &\sum_{K\in\mathbb{T}_h^0} \left\{(f + \Delta u_h, z - z_h)_K - \tfrac{1}{2}(n \cdot [\nabla u_h], z - z_h)_{\partial K}\right\} \\ &+ \sum_{K\in\mathbb{T}_h^\partial} \left\{(\bar{f} + \Delta u_h, \bar{z} - z_h)_K - \tfrac{1}{2}(n \cdot [\nabla u_h], \bar{z} - z_h)_{\partial K} + (f, z)_{S_K}\right\} .\end{aligned} \tag{2.22}$$

Then, the weighted *a posteriori* estimate (2.17) holds with the local *residuals* and *weights* as defined above, with the only modification that for $K \in \mathbb{T}_h^\partial$

$$\rho_K := h_K \|\bar{f} + \Delta u_h\|_K + \tfrac{1}{2} h_K^{1/2} \|n \cdot [\nabla u_h]\|_{\partial K} + \|f\|_{S_K} , \tag{2.23}$$

$$\omega_K := \max\left\{h_K^{-1}\|z - z_h\|_K,\, h_K^{-1/2}\|z - z_h\|_{\partial K},\, \|z\|_{S_K}\right\} . \tag{2.24}$$

In view of (2.20), we have $\omega_K \le C_{i,K} h_K^{m+1} \|z\|_{H^{m+1}(K)}$ for $K \in \mathbb{T}_h^\partial$. For a detailed analysis of the effect of boundary approximation using similar arguments, we refer to [12].

3 A feed-back algorithm for adaptive mesh refinement

We briefly discuss how a mesh refinement process may be organized on the basis of an *a posteriori* error estimate of the type (2.17) or (2.14). Suppose that some error tolerance TOL and maximum number N_{max} of mesh points are given. The goal is to find a most economical mesh $\mathbb{T}_h$ on which

$$|J(e)| \approx \eta(u_h) = \sum_{K \in \mathbb{T}_h} \eta_K \approx TOL\,, \tag{3.2}$$

with the *local error indicators* $\eta_K := \omega_K \rho_K$. Usually, one starts from an initial coarse mesh which is then successively refined. There are essentially three alternative strategies:

1. *Error per cell strategy.* The mesh generation aims at equilibrating the local error indicators η_K, by refining (or coarsening) the elements $K \in \mathbb{T}_h$ according to the criterion

$$\eta_K \approx \frac{TOL}{N}\,, \qquad N = \#\{K \in \mathbb{T}_h\}.$$

Since N depends on the result of the refinement decision, this strategy is implicit and usually needs iteration. It is common practice to work with a varying value N on each refinement level which is permanently updated according to the refinement process. The result is a mesh on which $\eta(u_h) \approx TOL$, provided that N_{max} is not exceeded.

2. *Fixed fraction strategy.* In each refinement cycle, the elements are ordered according to the size of η_K and either a fixed portion (say 30%) of the elements with largest η_K or the portion of elements which make up for a certain part of the estimator, $\kappa\eta(u_h)$, is refined. The appropriate choice of the parameter κ is crucial and depends very much on the particular situation. For "regular" functionals, one may choose $\kappa = 0.6 - 0.8$, while for "singular" functions a smaller choice $\kappa = 0.1 - 0.2$ is advisable, in order to enhance local refinement. This process is repeated until the stopping criterion $\eta(u_h) \approx TOL$ is fulfilled, or N_{max} is exceeded.

3. *Tolerance reduction strategy.* One works with a varying tolerance TOL_{var}. If on a mesh $\mathbb{T}_h$ a discrete solution u_h^{old} has been obtained with corresponding error estimator $\eta(u_h^{old})$, the tolerance is set to

$$TOL_{var} := \sigma\eta(u_h^{old})\,,$$

with some reduction factor $\sigma \in (0,1)$ (usually $\sigma = 0.5$). In the next step, one (or more) cycles of the *error per cell strategy* are applied with tolerance TOL_{var} yielding a refined mesh $\mathbb{T}_h^{new}$ and the new solution u_h^{new} with corresponding error bound $\eta(u_h^{new})$. Then, the tolerance is reduced again and a new refinement cycle begins. This process is repeated until $TOL_{var} \le TOL$, or N_{max} is exceeded.

The *error per cell strategy* is very delicate as its performance strongly depends on the parameters in the refinement decision. In certain cases it may happen that an inappropriate choice leads to very slow mesh refinement and in turn to an inefficient overall solution. On the other hand, the *fixed fraction strategy* guarantees that in each refinement cycle a sufficiently large number of elements is refined. However, in its pure form, it does not allow for mesh coarsening and in certain cases may tend to over-refine the mesh. Therefore, for practical computations, we recommend the *tolerance reduction strategy* which seems to be the most robust and accurate strategy among the three. Its use is particularly advisable if the weight-function $\omega(x)$ has

nonintegrable singularities which would otherwise dominate the refinement process (see the examples at the end).

However, for simplicity, in all the test calculations presented below, the "optimal" meshes $\mathbb{T}_h^{opt}$ have been generated by the *fixed fraction strategy* based on the simplified estimator (2.14)

$$\tilde{\eta}(u_h) := C_i \sum_{K \in \mathbb{T}_h} h_K \|n \cdot [\nabla u_h]\|_{\partial K} \|n \cdot [\nabla z_h^{(1)}]\|_{\partial K} ,$$

where $z_h^{(1)}$ is the bilinear approximation to z obtained on the current mesh. The interpolation constant is usually set $C_i = 0.1$, in order to compensate for the over-estimation in this bound. Accordingly, the local error indicators are defined as

$$\eta_K := C_i h_K \|n \cdot [\nabla u_h]\|_{\partial K} \|n \cdot [\nabla z_h^{(1)}]\|_{\partial K} . \tag{3.3}$$

4 Model analysis of the weighted a posteriori error estimation

We will give some theoretical support for the proposed approach to adaptive error control by looking in more detail at the model problem introduced in the previous section. The mesh refinement process is organized as described above using (isoparametric) bilinear shape functions on quadrilateral elements. Each element is allowed to have just one *hanging node* on each of its edges.

The crucial question is whether the *a posteriori* error bounds obtained by the proposed approach are not only asymptotically optimal (*order-optimal*) but also asymptotically sharp in the sense

$$I_{\text{eff}} := \lim_{TOL \to 0} \frac{|J(e)|}{\eta(u_h)} = 1 ,$$

possibly depending on the work the user is willing to invest in the evaluation of the weights ω_K, i.e., in the numerical solution of the dual problem. First, we run some computational experiments in order to see what can be expected. Unfortunately, the result is rather negative, i.e., with acceptable effort asymptotic sharpness does not seem to be achievable. With all the methods considered the *efficiency index* I_{eff} never really tends to one, but in most relevant cases stays well below 2, what may actually be considered as good enough.

There are two separate aspects to be considered: The sharpness of the global error bound $\eta(u_h)$, and the optimality of the local error indicators $\eta_K = \rho_K \omega_K$ which are used in the mesh refinement process. Starting from the error representation (1.7),

$$J(e) = \sum_{K \in \mathbb{T}_h} \left\{ (f + \Delta u_h, z - z_h)_K - \tfrac{1}{2}(n \cdot [\nabla u_h], z - z_h)_{\partial K} \right\} , \tag{4.2}$$

there are various steps in evaluating the right hand side. First, one estimates

$$|J(e)| \le \sum_{K \in \mathbb{T}_h} \eta_K , \tag{4.3}$$

with the (positive) local error indicators

$$\eta_K := \left| (f + \Delta u_h, z - z_h)_K - \tfrac{1}{2}(n \cdot [\nabla u_h], z - z_h)_{\partial K} \right| , \tag{4.4}$$

which are used in the mesh refinement process. By this step, in general, the asymptotic sharpness of the global error estimate is lost. This is seen, for instance, in the case where the exact as well as the approximate solution are anti-symmetric with respect to the x-axis meaning that $e(0) = 0$, but $\eta(u_h) \neq 0$. Therefore, we will analyze various ways of directly evaluating the error identity (4.2). Here, $z_h \in V_h$ may be taken as the bilinear interpolation of the approximation for z on the current mesh, or simply be set zero, as its global contributions (if exactly evaluated) cancel out to zero under summation. The approximate values obtained for $J(e)$ are denoted by $\eta^{(i)}(u_h)$.

1. Approximation by higher order methods: The dual problem is solved by using biquadratic finite elements on the current mesh yielding an approximation $z_h^{(2)}$ to z. The error estimator obtained by using this approximation in (4.3) and (4.4) is denoted by $\eta^{(1)}(\cdot)$. Computing approximations to z by using a higher order finite element scheme does not appear very economical in estimating the error in the low-order scheme. Hence, one may try to work with the given method, i.e., piecewise bilinear approximation, and increase its accuracy through defect correction with biquadratic interpolation on a coarser mesh. Notice that simple Richardson extrapolation using two consecutive meshes yields improved approximations only at the nodes of the coarser mesh. However, since this procedure requires to solve at least twice on the current mesh, it appears too costly for practical purposes and is not further considered.

2. Approximation by higher order interpolation: A further simplification is achieved by simple local biquadratic interpolation of the bilinear approximation $z_h^{(1)}$ on the current mesh yielding an approximation $I_h^{(2)} z_h^{(1)}$ to z. The resulting global error estimator is denoted by $\eta^{(2)}(\cdot)$. This requires some special care on elements with hanging nodes, in order to preserve the higher order accuracy of the interpolation process.

3. Approximation by higher order difference quotients: Without loss of generality, consider a rectangular element K oriented in the cartesian directions with side lengths $h_1 = h_{K,1}$ and $h_2 = h_{K,2}$. Its vertices are numbered in the counter-clockwise sense as $a^i = (x_1^i, x_2^i)\,(i = 0, ..., 3)$. On K the error in the bilinear interpolation can be expanded like

$$(z - I_h^{(1)} z)(x) = E_K(x) h_K^2 + R_K(h_K, x)\,, \tag{4.5}$$

with an expansion coefficient $E_K(x)$ and a remainder term $R_K(h_K, x) = O(h_K^3)$. There holds

$$\begin{aligned} E_K(x) = {} & \tfrac{1}{2} h_1^2 h_K^{-2} \Big\{ \varphi(\varphi - 1)(\psi\chi z_{11}(a^0) + \psi\xi z_{11}(a^3)) + \psi(\psi - 1)(\varphi\chi z_{11}(a^1) + \varphi\psi z_{11}(a^2)) \Big\} \\ & + \tfrac{1}{2} h_2^2 h_K^{-2} \Big\{ \xi(\xi - 1)(\psi\chi z_{22}(a^0) + \varphi\chi z_{22}(a^1)) + \chi(\chi - 1)(\psi\chi z_{22}(a^1) + \psi\xi z_{22}(a^2)) \Big\}\,, \end{aligned}$$

where $\varphi = (x_1 - x_1^0)/(x_1^1 - x_1^0)$, $\psi = (x_1^1 - x_1)/(x_1^1 - x_1^0)$, $\xi = (x_2 - x_2^0)/(x_2^1 - x_2^0)$, $\chi = (x_2^1 - x_2)/(x_2^1 - x_2^0)$, and $|R_K(h_K, x)| \leq h_K^3 \max_K |\nabla^3 z|$. The elementary but tedious computation is omitted. According to (4.5), the local error indicators are estimated in the form

$$\eta_K \approx h_K^2 \Big| (f + \Delta u_h, E_K)_K - \tfrac{1}{2} (n \cdot [\nabla u_h], E_K)_{\partial K} \Big|\,. \tag{4.6}$$

The second derivatives of z in the expansion coefficient E are then approximated by suitable central difference quotients of its bilinear approximation $z_h^{(1)}$ obtained on the current mesh. Again at hanging nodes special care is required. The resulting global error estimator is denoted by $\eta^{(3)}(\cdot)$.

4. Approximation by local problems: Following the idea already used in the energy-error estimator of Bank and Weiser [3], on each element K the local problems

$$(\nabla v_K, \nabla \varphi_h)_K = (f + \Delta u_h, \varphi_h)_K - \tfrac{1}{2}(n \cdot [\nabla u_h], \varphi_h)_{\partial K} \quad \forall\, \varphi_h \in V_K\,, \tag{4.7}$$

are solved, where $V_K = Q_2(K)$ the space of biquadratic polynomials on K. Using the resulting solutions, the local residual terms are replaced by

$$\tilde{\eta}_K := |(\nabla v_K, \nabla(z - z_h))_K|\,. \tag{4.8}$$

The resulting global error estimator is denoted by $\eta^{(4)}(\cdot)$. The related error indicator based on solving patch-wise Dirichlet problems is not considered here, since it turns out to be too expensive in practice.

From the various results obtained in our tests, we quote those for the point-value computation discussed above, as well as those for the error measured in the global L^2-norm. The observed efficiency indices are listed in Tables 2 and 3. These results indicate that even for the simplest model situations, an efficiency index $I_{\text{eff}} = 1$ is achievable only at the expense of unacceptably high cost, e.g., by approximating the dual solution using higher-order elements.

L	N	$J(e) = e(0)$	$J(e)/\eta^{(1)}$	$J(e)/\eta^{(2)}$	$J(e)/\eta^{(3)}$	$J(e)/\eta^{(4)}$
1	4^3	1.195e-1	0.150	0.078	0.484	0.103
2	4^4	5.429e-2	0.798	0.074	0.022	0.073
3	4^5	1.995e-2	0.950	0.322	0.107	0.321
4	4^6	6.517e-3	0.993	0.487	0.142	0.487
5	4^7	1.997e-3	0.997	0.530	0.150	0.530

Table 2: Efficiency of various *weighted* error indicators obtained for the point-error $J(e) = |e(0)|$

L	N	$J(e) = \|e\|$	$J(e)/\eta^{(1)}$	$J(e)/\eta^{(2)}$	$J(e)/\eta^{(3)}$	$J(e)/\eta^{(4)}$
1	4^3	9.793e-0	0.446	0.089	0.089	0.089
2	4^4	2.525e-0	0.875	0.562	0.467	0.527
3	4^5	6.328e-1	0.928	0.804	0.685	0.792
4	4^6	1.586e-2	0.939	0.860	0.850	0.859
5	4^7	3.961e-3	0.943	0.863	0.861	0.863

Table 3: Efficiency of various *weighted* error indicators obtained for the L^2-error $J(e) = \|e\|$

To prove that the approximation strategies described above actually work in constructing "optimal" meshes seems very difficult, as for this one would need precise information about the local (on the element level) approximation properties of the finite element Ritz projection on general unstructured meshes. Unfortunately, results of this type are not available in literature and may not even be true at all. Below, we will prove a result for the point-error estimator which is not quite what one would like to have, but provides some partial justification. It states that, at least, on an optimally refined mesh on the basis of the "exact" estimator $\eta(u_h)$, the perturbed estimator $\tilde{\eta}(u_h)$ using a higher-order approximation to the dual solution z shows the

right behavior in terms of the tolerance TOL. To this end, we have to make some reasonable technical assumptions. First, suppose that on the "optimal" meshes $\mathbb{T}_h$, there holds

$$\rho_K \le ch_K^2 \Big\{ \max_K |f + \Delta u_h| + \max_{\partial K} |n \cdot \nabla [u_h] h_K^{-1}| \Big\} \le C(u) h_K^2 \quad K \in \mathbb{T}_h \,, \tag{4.9}$$

which indicates the (global) stability of the finite element Ritz projection with respect to the discrete W_∞^2-norm on general locally refined meshes. On quasi-uniform meshes (satisfying the uniform shape and size condition) this immediately follows by the known error estimates in the W_∞^1-norm and the inverse properties of finite elements. For more general meshes, this has to be left as a conjecture. However, (4.9) may be checked in the course of the mesh refinement process.

Lemma 4.1. *For the model problem (1.2) on the square $\Omega = (-1,1)^2$, let the dual solution z, i.e., the Green function to the point $x = 0$, be approximated on the current mesh by its Ritz projection $z_h^{(2)}$ into the subspace of piecewise quadratic functions. Further, suppose that for a sequence of tolerances, $\mathrm{TOL} \to 0$, optimally balanced meshes $\mathbb{T}_h^{opt}$ have been generated on the basis of the exact estimator $\eta_{point}(u_h)$, such that $\eta_{point}(u_h) \approx \mathrm{TOL}$ and*

$$\eta_K \approx \frac{\mathrm{TOL}}{N_{opt}} \,, \quad K \in \mathbb{T}_h^{opt} \,. \tag{4.10}$$

Then, the approximate error estimator

$$\tilde{\eta}_{point}(u_h) := \Big| \sum_{K \in \mathbb{T}_h} \Big\{ (f + \Delta u_h, z_h^{(2)})_K - \tfrac{1}{2} (n \cdot [\nabla u_h], z_h^{(2)})_{\partial K} \Big\} \Big| \tag{4.11}$$

is asymptotically exact in the sense

$$|\eta_{point}(u_h) - \tilde{\eta}_{point}(u_h)| = o(TOL) \,. \tag{4.12}$$

Proof. We only give a sketch of the proof which is based on results from local finite element error analysis (see [11]). Using the abbreviation $w_h := z - z_h^{(2)}$, we have by Galerkin orthogonality

$$\eta_{point}(u_h) - \tilde{\eta}_{point}(u_h) = \sum_{K \in \mathbb{T}_h} \Big\{ (f + \Delta u_h, w_h - I_h w_h)_K - \tfrac{1}{2} (n \cdot [\nabla u_h], w_h - I_h w_h)_{\partial K} \Big\} \,,$$

and, consequently, by the local approximation properties of the interpolation operator I_h,

$$\Big| \eta_{point}(u_h) - \tilde{\eta}_{point}(u_h) \Big| \le C_i \sum_{K \in \mathbb{T}_h} \rho_K h_K \|\nabla^2 w_h\|_K \,.$$

This is further estimated by

$$\Big| \eta_{point}(u_h) - \tilde{\eta}_{point}(u_h) \Big| \le C_i C(u) \Big(\sum_{K \in \mathbb{T}_h} \rho_K^2 h_K^2 r_K^{-4} \Big)^{1/2} \Big(\sum_{K \in \mathbb{T}_h} r_K^4 \|\nabla^2 w_h\|_K^2 \Big)^{1/2} \,. \tag{4.13}$$

The second factor on the right is bounded by

$$\Big(\sum_{K \in \mathbb{T}_h} r_K^4 \|\nabla^2 w_h\|_K^2 \Big)^{1/2} \le \Big(\int_\Omega r^4 |\nabla^2 w_h|^2 \, dx \Big)^{1/2} \,,$$

where $r(x) := \max\{|x|, h_0\}$, $h_0 = h_{min}$ being the diameter of the element containing the point $x = 0$. By standard arguments from local finite element analysis, one can prove that

$$\Big(\int_\Omega r^4 |\nabla^2 w_h|^2 \, dx \Big)^{1/2} \le c h_{max} \Big(\int_\Omega r^4 |\nabla^3 z|^2 \, dx \Big)^{1/2}$$
$$\le c h_{max} \Big(\int_\Omega r^{-2} \, dx \Big)^{1/2} \le c h_{max} \, |log(h_0)|^{1/2} .$$

We omit the rather technical details and refer to [18] and [11] for similar arguments. Under the assumptions (4.10) and (4.9), for the optimal mesh $\mathbb{T}_h$, there holds $N \approx TOL^{-1}$ and $h_K \approx (TOL r_K)^{1/2}$. Therefore, the first factor on the right of (4.13) can be bounded by

$$\Big(\sum_{K \in \mathbb{T}_h} \rho_K^2 h_K^2 r_K^{-4} \Big)^{1/2} \le C(u) TOL \Big(\sum_{K \in \mathbb{T}_h} h_K r_K^{-2} \Big)^{1/2} \le C(u) \, |log(h_0)|^{1/2} \, TOL .$$

Combining the foregoing estimates in (4.13) yields

$$\Big| \eta_{point}(u_h) - \tilde{\eta}_{point}(u_h) \Big| \le C_i C(u) \, |log(h_0)| \, h_{max} \, TOL .$$

In view of $h_{max} \approx TOL^{1/2}$, this implies the assertion. □

In view of the computational results shown in Table 3, the statement of Lemma 4.1 does not seem to be true for the approximation of the global L^2-error. Since asymptotically sharp *a posteriori* control does not seem to be feasible in general, we have to live with error estimates involving at least certain *interpolation constants* C_i. These constants may usually be assumed to be of moderate size $C_i = 0.1 - 1$, but certainly represent some factor of uncertainty. In particular situations, one may try to calibrate these constants in the course of the refinement process which, however, is a very delicate matter. The full theoretical justification of the error estimator $\tilde{\eta}(\cdot)$ would require an analysis of the effect of replacing the exact second derivatives of the dual solution z by suitable difference quotients of some approximation z_h obtained numerically on the current mesh. Yet, this is strongly supported by numerical evidence, even in the present simple case the proof must be left as an open problem.

For the residual-based energy error estimator (1.13) one has the two-sided bound

$$\|\nabla e\| \le \eta_{energy}(u_h) \le \gamma \{ \|\nabla e\| + \|f - f_h\| \} , \tag{4.14}$$

where f_h is a suitable approximation of f. This shows asymptotic optimality. An analogous bound cannot hold in general for local error quantities like the point-error. What we can hope to achieve is only an optimality estimate of the type presented in the following lemma.

Lemma 4.2. *For the finite element approximation of the model problem (1.2) on the square $\Omega = (-1, 1)^2$, the error estimator $\eta_{point}(\cdot)$ defined in (2.4) is asymptotically optimal in the sense*

$$|e(0)| \le \eta_{point}(u_h) \le c \, |log(h_0)| \max_{K \in \mathbb{T}_h} \Big(\max_K |e| + h_K^2 \max_K |\nabla^2 u| \Big) , \tag{4.15}$$

where h_0 is the diameter of the element containing the point $x = 0$.

Proof. Starting from (2.4), we use the Hölder inequality to obtain

$$\eta_{point}(u_h) \le \Big(\sum_{K\in\mathbb{T}_h} \frac{h_K^2}{r_K^2}\rho_K^2\Big)^{1/2}\Big(\sum_{K\in\mathbb{T}_h} h_K^{-2} r_K^2 \omega_K^2\Big)^{1/2}, \tag{4.16}$$

where $r_K := max\{dist(K,0), h_K\}$. The second factor can be estimated by (see [18] or [11])

$$\sum_{K\in\mathbb{T}_h} h_K^{-2} r_K^2 \omega_K^2 \le c \int_{\{h_0<|x|<1\}} |x|^2 |\nabla^2 z|^2 \, dx + c|log(h_0)| \le c\,|log(h_0)|\,. \tag{4.17}$$

Further, using the usual inverse inequality for finite elements together with the local interpolation estimate (1.3), we conclude in a standard way that

$$\rho_K \le h_K \|\Delta e\|_K + \tfrac{1}{2} h_K^{1/2} \|n\cdot[\nabla e]\|_{\partial K} \le c h_K^{-1} \|e\|_{\tilde K} + c h_K \|\nabla^2 u\|_{\tilde K}\,,$$

where $\tilde K$ denotes the union of all cells in $\mathbb{T}_h$ intersecting K. This implies

$$\sum_{K\in\mathbb{T}_h} \frac{h_K^2}{r_K^2}\rho_K^2 \le c\,|log(h_0)| \max_{K\in\mathbb{T}_h} \Big(max_{\tilde K}|e| + h_K^2 max_{\tilde K}|\nabla^2 u|\Big)^2\,. \tag{4.18}$$

Combining (4.17) and (4.18), the assertion follows. □

For a special situation (uniform tensor-product mesh) it has been shown in [32] and [33] that in the *energy-error* estimate (1.13) either the domain residual terms $\|f+\Delta u_h\|_K$ or the jump terms $\|n\cdot[\nabla u_h]\|_{\partial K}$ may be asymptotically neglected depending on whether the degree of the elements used is odd or even, respectively. This result has recently been extended in [31] to general adaptively refined meshes for lowest-order elements (odd degree). We will derive a corresponding result for the present situation which to some extent justifies the use of the simplified weighted error estimator $\tilde\eta(\cdot)$ defined in (2.14).

Lemma 4.3. *Suppose that $f \in H^1(\Omega)$ in the model problem (1.2). Then, for piecewise bilinear elements, there holds*

$$|J(e)| \le C_i \sum_{K\in\mathbb{T}_h} \Big(h_K^3 \|\nabla f\|_{\tilde K} + h_K^{3/2} \|n\cdot[\nabla u_h]\|_{\partial K}\Big) \|\nabla^2 z\|_{\tilde K}\,, \tag{4.19}$$

where $\tilde K$ is the union of all cells in $\mathbb{T}_h$ intersecting K.

Proof. For simplicity, we give the proof only for rectangular meshes. The generalization to more general meshes requires some more technical work. In this case $\Delta u_h|_K \equiv 0$ on each element K. Since the mesh $\mathbb{T}_h$ is generated from a coarse mesh by local cell-wise bisection (i.e., subdivision into four sub-cells), it can be decomposed into macro-cells consisting of four sub-cells each. These macro-cells form a coarser mesh named $\tilde{\mathbb{T}}_h$. The case of cells with hanging nodes requires obvious modifications and will not be discussed in detail. On $\tilde{\mathbb{T}}_h$ we define an interpolation operator $\tilde I_h : C(\bar\Omega) \to V_h$, by requiring $\tilde I_h v(a) = v(a)$, for vertices a of $\tilde{\mathbb{T}}_h$, and $\int_{\tilde K} \tilde I_h v\, dx = \int_{\tilde K} v\, dx$, for each macro-cell $\tilde K \in \tilde{\mathbb{T}}_h$. By standard arguments (Bramble-Hilbert lemma) we conclude the estimate

$$\|v - \tilde I_h v\|_K + h_K^{1/2} \|v - \tilde I_h v\|_{\partial K} \le C_{i,K} h_K^2 \|\nabla^2 v\|_{\tilde K}\,, \tag{4.20}$$

with certain interpolation constants $C_{i,K}$ independent of h. Now, we choose $z_h = \tilde{I}_h z$ in the error representation (4.2) and obtain

$$J(e) = \sum_{K\in\mathbb{T}_h} \Big\{ (f - \tilde{f}_h, z - \tilde{I}_h z_h)_K - \tfrac{1}{2}(n\cdot[\nabla u_h], z - \tilde{I}_h z_h)_{\partial K} \Big\}, \tag{4.21}$$

where $\tilde{f}_h$ denotes the piecewise constant interpolation f on the macro-mesh $\tilde{\mathbb{T}}_h$. Hence, in view of (4.20), we conclude that

$$|J(u_h)| \le C_i \sum_{K\in\mathbb{T}_h} \Big(h_K^3 \|\nabla f\|_{\tilde{K}} + h_K^{3/2} \|n\cdot[\nabla u_h]\|_{\partial K} \Big) \|\nabla^2 z\|_{\tilde{K}}\,, \tag{4.22}$$

which proves the assertion. □

5 Applications

The following examples are intended to support some of the claims made above on the basis of theoretical analysis. In the tables shown below, we list the results obtained for a sequence of successively reduced values of TOL, the number of points, N_{opt}, of the final mesh $\mathbb{T}_h^{opt}$, on which the error indicators are balanced, the maximum number of refinement levels, L, in this mesh, and the efficiency index $I_{\text{eff}} := |J(e)|/\eta(u_h)$.

Example 1. As first example, we consider the computation of isolated stress values, i.e., point-values of the gradient ∇u, in the model Poisson problem (1.2) on the square $\Omega = (-1,1)^2$. Let, for example, the error functional be chosen as

$$J(u) := \partial_1 u(0)\,.$$

Since the corresponding dual solution does not exist in the sense of $H_0^1(\Omega)$, we have to regularize the functional according to

$$J_\epsilon(u) := |B_\epsilon(0)|^{-1} \int_{B_\epsilon(0)} \partial_1 u\, dx\,,$$

where $B_\epsilon(0) := \{x \in \Omega, |x| < \epsilon\}$, and $\epsilon = \text{TOL}$. Then, $|J(u) - J_\epsilon(u)| \approx \text{TOL}$, and the corresponding dual solution z_ϵ is well defined in $H_0^1(\Omega)$. Observing that $z_\epsilon(x) \approx x_1/(|x|+\epsilon)^2$, we are led to the weighted *a posteriori* error estimate

$$|\partial_1 e(0)| \le C_i \sum_{K\in\mathbb{T}_h} \frac{h_K^2}{(r_K + \text{TOL})^3} \rho_K + c\text{TOL}\,, \tag{5.2}$$

where $r_K = \max\{dist(K,0), h_K\}$. We again assume that the local residuals satisfy $\rho_K \approx h_K^2$. Then, on a mesh $\mathbb{T}_h^{opt}$ on which the local error indicators $\eta_K = \omega_K \rho_K$ are equilibrated,

$$\eta_K \approx \frac{TOL}{N_{opt}}\,, \qquad N_{opt} = \#\{K \subset \mathbb{T}_h^{opt}\}\,, \tag{5.3}$$

there holds

$$h_K \approx \Big(\frac{TOL\,(r_K + \text{TOL})^3}{N_{opt}} \Big)^{1/4}\,, \tag{5.4}$$

Consequently, the number N_{opt} of elements of $\mathbb{T}_h^{opt}$ is given by

$$N_{opt} \approx \sum_{K\in\mathbb{T}_h^{opt}} h_K^2 h_K^{-2} = \sum_{K\in\mathbb{T}_h^{opt}} h_K^2 \frac{N_{opt}^{1/2}}{\mathrm{TOL}^{1/2}\,(r_K+\mathrm{TOL})^{3/2}}\,. \tag{5.5}$$

Since $r^{-3/2}$ is integrable over Ω, it follows that $N_{opt} \approx \mathrm{TOL}^{-1}$, as for the evaluation of the point value $u(0)$. Further, the mesh size is distributed like

$$h_K \approx \mathrm{TOL}^{1/2}(r_K+\mathrm{TOL})^{3/4}\,,$$

and, consequently, $h_{min} \approx \mathrm{TOL}^{5/4}$ and $h_{max} \approx \mathrm{TOL}^{1/2}$. We note that if the functional $J(\cdot)$ is not regularized, the same reasoning yields the mesh-size distribution $h_K \approx \mathrm{TOL}^{1/2} r_K^{3/4}$ and, in this case, $h_{min} \approx \mathrm{TOL}^2$, meaning a strong over-refinement at $x=0$. This effect is confirmed in our test computations.

In contrast to the weighted estimator (5.2), the global energy-error estimator (1.13), for $r=0$,

$$\|\nabla e\| \le C_s C_i \Big(\sum_{K\in\mathbb{T}_h} \rho_K^2\Big)^{1/2} =: \eta_{energy}(u_h)\,, \tag{5.6}$$

or the analogue of the L^∞-error estimator (2.11),

$$\max_\Omega |\partial_1 e| \le C_i C_s |\log h_0|^{1/2} \max_{K\in\mathbb{T}_h} \rho_{\infty,K} =: \eta_\infty(u_h)\,, \tag{5.7}$$

with

$$\rho_{\infty,K} := h_K \max_K |f+\Delta u_h| + \max_{\partial K} |n\cdot[\nabla u_h]|\,,$$

would generate meshes of complexity $N_{opt} \approx TOL^{-2}$, in order to guarantee $\eta(u_h) \approx \mathrm{TOL}$.

Figure 2 shows the balanced mesh for $\mathrm{TOL} = 4^{-4}$ and the approximation to the dual solution z_ϵ, $\epsilon = \mathrm{TOL}$, computed on this mesh. The corresponding errors for a sequence of tolerances are listed in Table 4. For the problem considered, the weighted *a posteriori* estimate (2.4) for the point error is obviously asymptotically optimal and the predicted dependence $N_{opt} \approx \mathrm{TOL}^{-1}$ is confirmed.

Example 2. As a very particular example, we consider the computation of the contour integral

$$J(u) = \int_{\partial\Omega} \partial_n u n_1\, ds\,, \tag{5.8}$$

for the solution of the model problem (1.2) on the unit disc, $\Omega = B_1 = \{x\in\mathbb{R}^2, |x|<1\}$, where n_1 is the x_1-component of the outer unit vector to $\partial\Omega$. Error functionals of this type occur, for instance, in the computation of drag and lift coefficients of blunt bodies in viscous flows modeled by the Navier-Stokes equations (see [7]). In view of the pointwise analogue of the *a priori* error estimate (1.6), we have $J(e) = O(h_{max})$, so that convergence is no problem. The question is how the mesh refinement process should be organized in order to obtain $J(u)$ with best accuracy. We will see that the answer is somewhat surprising and seems, at the first glance, to conflict with what one is used to from a finite element method.

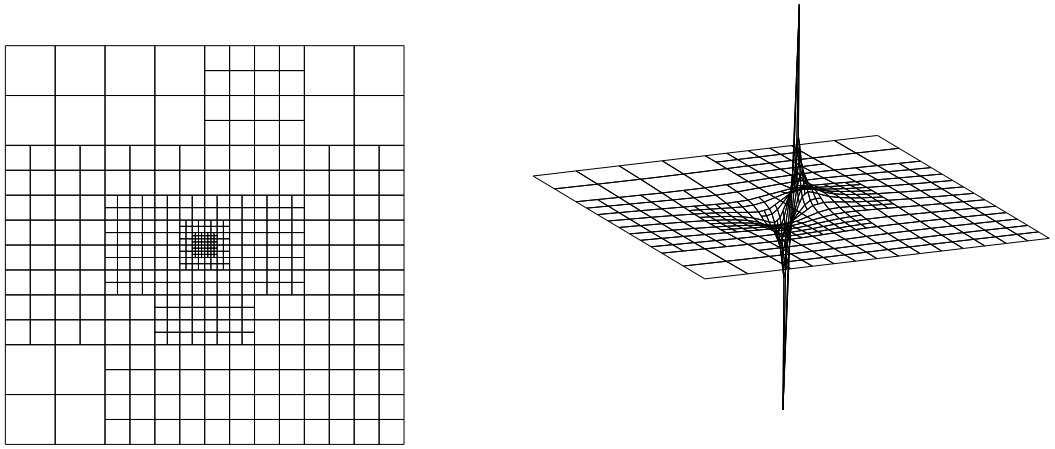

Figure 2: Refined mesh and approximate dual solution for computing $\partial_1 u(0)$ in Example 1, using the weighted error estimator η_{point}, with TOL $= 4^{-4}$

η_{energy}					
TOL	N	L	$\lvert\partial_1 e(0)\rvert$	η	I_{eff}
1	736	6	3.91e-1	8.80e-1	0.43
4^{-1}	11020	8	9.91e-2	2.28e-1	0.43
4^{-2}	166360	10	2.48e-2	5.88e-2	0.42

$\eta_{1,\infty}$					
TOL	N	L	$\lvert\partial_1 e(0)\rvert$	η	I_{eff}
1	292	6	7.65e-1	7.25e-1	1.11
4^{-1}	2848	7	1.98e-1	2.27e-1	0.90
4^{-2}	44980	9	4.96e-2	4.94e-2	1.00

$\eta_{1,point}$					
TOL	N	L	$\lvert\partial_1 e(0)\rvert$	η	I_{eff}
1	40	4	2.57e-0	5.57e-1	5.00
4^{-1}	124	6	7.38e-1	2.34e-1	3.33
4^{-2}	448	11	1.35e-3	6.09e-2	0.02
4^{-3}	1780	15	1.19e-3	1.34e-2	0.09
4^{-4}	6328	19	2.20e-3	3.61e-3	0.62
4^{-5}	25984	24	4.28e-4	8.61e-4	0.50
4^{-6}	95260	28	1.39e-4	2.33e-4	0.59

$\eta_{1,point}^{regular}$					
TOL	N	L	$\lvert\partial_1 e(0)\rvert$	η	I_{eff}
1	40	4	2.57e-0	1.09e-2	-
4^{-1}	64	4	1.47e-0	5.16e-2	-
4^{-2}	148	6	7.51e-1	5.92e-2	-
4^{-3}	940	9	4.10e-1	1.42e-2	3.33
4^{-4}	4912	12	4.14e-3	3.50e-3	1.25
4^{-5}	20980	15	2.27e-4	9.25e-4	0.24
4^{-6}	86740	17	5.82e-5	2.38e-4	0.24

Table 4: Results for computing $\partial_1 u(0)$ in Example 1, using the energy- and W^1_∞-error estimators η_{energy} and $\eta_{1,\infty}$, respectively, compared to the weighted estimator $\eta_{1,point}$ and its regularized variant $\eta^{regular}_{1,point}$

First, in oder to avoid having to deal with measures, we regularize the functional $J(\cdot)$. With the cut-off function

$$\psi_\epsilon(x) = \begin{cases} 1\,, & 0 \le |x| < 1-\epsilon\,, \\ \epsilon^{-1}(1-|x|)\,, & 1-\epsilon \le |x| < 1\,, \end{cases}$$

we make the following ansatz for the dual solution

$$z_\epsilon(x) = -x_1\psi_\epsilon(x)\,,$$

where $\epsilon = \mathrm{TOL}$. Let $\varphi \in V \cup C^2(\bar{\Omega})$ be arbitrary. We split the integration over B_1 according to $B_1 = B_{1-\epsilon} \cup S_\epsilon$ with $S_\epsilon = B_1 \setminus B_{1-\epsilon}$ and obtain, using polar coordinates (r, θ),

$$\int_{B_1} \nabla z_\epsilon \cdot \nabla \varphi \, dx = \int_{\partial B_{1-\epsilon}} \partial_n z_\epsilon \varphi \, ds + \int_{S_\epsilon} \left\{ \partial_r z_\epsilon \partial_r \varphi + r^{-2} \partial_\theta z_\epsilon \partial_\theta \varphi \right\} dx \, .$$

By construction, there holds $\partial_n z_\epsilon = -\partial_r x_1 \psi_\epsilon = -cos\theta$ on $\partial B_1 \setminus B_{1-\epsilon}$, and consequently, observing $x_1 = r\cos(\theta)$,

$$\int_{\partial B_{1-\epsilon}} \partial_n z_\epsilon \varphi \, ds = -\int_0^{2\pi} cos(\theta) \varphi(1-\epsilon, \theta) \, ds = O(\epsilon) \, .$$

Further, we have

$$\int_{S_\epsilon} \partial_r z_\epsilon \partial_r \varphi \, dx = -\int_{S_\epsilon} x_1 \partial_r \psi_\epsilon \partial_r \varphi \, dx - \int_{S_\epsilon} cos(\theta) \psi_\epsilon \partial_r \varphi \, dx = -\epsilon^{-1} \int_{S_\epsilon} x_1 \partial_r \varphi \, dx + O(\epsilon) \, ,$$

and

$$\int_{S_\epsilon} r^{-2} \partial_\theta z_\epsilon \partial_\theta \varphi \, dx = \epsilon^{-1} \int_{S_\epsilon} r^{-2} (1-r) sin(\theta) \partial_\theta \varphi \, dx = O(\epsilon) \, .$$

Finally, observing that

$$\epsilon^{-1} \int_{S_\epsilon} x_1 \partial_r \varphi \, dx = \int_{\partial B_1} n_1 \partial_n \varphi \, ds + O(\epsilon) \, ,$$

we find that

$$J(\varphi) = (\nabla z_\epsilon, \nabla \varphi) + O(\epsilon) \, . \tag{5.9}$$

This implies that the approximate weights $\tilde{\omega}_K \approx \|\nabla^2 z_\epsilon\|_K$ in the *a posteriori* error estimate (2.17) corresponding to the functional $J_\epsilon(u_h) := (\nabla z_\epsilon, \nabla \varphi)$ are all zero on elements K outside a boundary strip of width ϵ, i.e., there is no contribution to the error $J(e)$ from those interior elements. This remains true if we let $\epsilon \to 0$.

Therefore, starting from some coarse mesh, the "optimal" refinement strategy would be in each step only to refine those elements which are adjacent to the boundary and to leave the others unchanged. The resulting meshes are maximally refined along the boundary down to a minimal mesh size h_{min}, while in the interior of Ω the elements are kept of size $h_{max} \approx 1$. If only local mesh bisection is allowed in each refinement step, the mesh obtained after L steps has minimal size $h_{min} \approx 2^{-L}$ and consists of $N_L \approx 2^L$ elements. To reach an error of size $|J(e)| \leq TOL$, about $k \approx \log_2(1/TOL)$ refinement cycles are needed leading to an "optimal" mesh with $N \approx 1/TOL$ elements. However, to realize the full accuracy on the constructed mesh, the right hand side consisting of integrals over the force f multiplied by the nodal basis functions has to be calculated almost exactly also on the coarse elements.

Figure 3 shows the the finest mesh obtained after 7 refinement steps and the approximation to the dual solution $z_\epsilon, \epsilon = h_{min}$, computed on this mesh. In Table 5 the results obtained by the usual energy error estimator $\eta_{energy}(u_h)$ are compared to those of the "optimal" refinement strategy.

Example 3. By the third example, we demonstrate that carrying weights in the *a posteriori* error estimates can also be advantageous in estimating the error in global norms. This applies

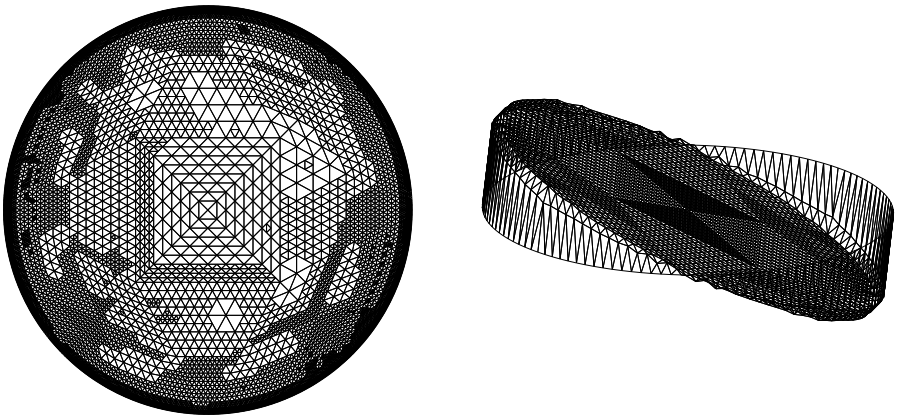

Figure 3: Refined mesh and approximate dual solution for Example 2 obtained by the approximate weighted error estimator η_{weight}^{approx}, after 7 refinement steps

η_{energy}			"optimal" strategy		
L	N	$J(e)$	L	N	$J(e)$
0	256	$5.09+0$	0	256	$5.09+0$
1	1024	$2.91+0$	1	559	$2.90+0$
2	4096	$1.50+0$	2	1294	$1.50+0$
3	16384	$7.49-1$	3	3079	$7.48-1$
4	64768	$3.77-1$	4	7174	$3.62-1$
5	253858	$1.73-1$	5	16738	$1.67-1$
6	*memory*	*exhausted*	6	38146	$6.93-2$
7	*memory*	*exhausted*	7	86128	$2.03-2$

Table 5: Results obtained with the energy norm estimator η_{energy} and by the "optimal" refinement strategy for Example 2

particularly when global *a priori* bounds for the dual solution do not appropriately reflect the local stability properties of the underlying problem, e.g., for problems with strongly varying coefficients. Consider, for example the boundary value problem

$$-\operatorname{div}\{a\nabla u\} = f \quad \text{in } \Omega\,, \qquad u = 0 \quad \text{on } \partial\Omega\,, \tag{5.10}$$

on the square $\Omega = (-1,1)^2$ with a coefficient function,

$$a(x) = 0.1 + e^{3(x_1+x_2)}\,. \tag{5.11}$$

We discretize this problem by bilinear finite elements and recall the weighted L^2-*a posteriori* error estimates following from (1.8),

$$\|e\| \le \sum_{K\in\mathbb{T}_h} \omega_K \rho_K\,, \tag{5.12}$$

where the residuals ρ_K and weights ω_K are defined by (1.9) and (1.10), where z is the solution

of the dual problem,

$$-\operatorname{div}\{a\nabla z\} = e/\|e\| \quad \text{in } \Omega\,, \qquad z = 0 \quad \text{on } \partial\Omega\,, \tag{5.13}$$

First, the weights $\tilde{\omega}_K$ are computed by taking second difference quotients of the discrete dual solution $z_h^{(1)} \in V_h$ on the current mesh $\mathbb{T}_h$, while on the right hand side of (5.13) the exact error e is used. The reference solution is as before $u = 10\sin(2x_1 + x_2 + 2)$. This gives the approximate stability constant

$$\tilde{C}_s := \Big(\sum_{K\in\mathbb{T}_h} \tilde{\omega}_K^2\Big)^{1/2}.$$

This way, we obtain a first approximate weighted L^2-error estimator

$$\tilde{\eta}_{weight}(u_h) := \sum_{K\in\mathbb{T}_h} \rho_K \tilde{\omega}_K$$

as well as the corresponding global L^2-error estimator

$$\tilde{\eta}_2(u_h) := C_i \tilde{C}_s \Big(\sum_{K\in\mathbb{T}_h} h_K^2 \rho_K^2\Big)\,,$$

the interpolation constant being chosen as $C_i = 0.2$. Next, the unknown error e in (5.13) is approximated by $\tilde{e} := I_h^{(2)} u_h - u_h$, where $I_h^{(2)} u_h$ is the patch-wise biquadratic interpolation of u_h. The resulting error estimator is denoted by $\tilde{\eta}_{weight}^{approx}$. The results obtained for the above test problem by using these three L^2-error estimators are listed in Table 6. Figure 4 shows the errors on the corresponding "optimal" meshes for the global estimator $\tilde{\eta}_2(u_h)$ and the weighted estimator $\tilde{\eta}_{weight}(u_h)$. These tests clearly demonstrate the superiority of $\tilde{\eta}_{weight}$ over $\tilde{\eta}_2$.

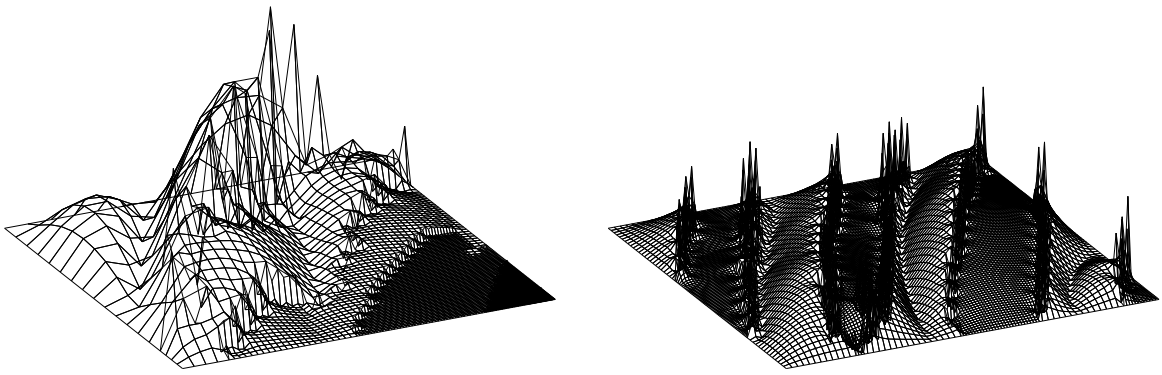

Figure 4: Errors on balanced meshes with $N \sim 10000$ for Example 3 by the L^2-error estimator $\tilde{\eta}_2(u_h)$ (left, scaled by 1:10) and the weighted error estimator $\tilde{\eta}_{weight}^{approx}(u_h)$ (right, scaled by 1:30)

Example 4. The final example is intended to support our claim that, in using higher order finite elements, one should indeed use the approximation property of the finite element space to its full extent in generating the weights ω_K in the *a posteriori* estimate (2.17). First, we

$\tilde{\eta}_2$						
TOL	N	L	$\|e\|$	η	I_{eff}	$\tilde{C}_s$
1	556	8	9.63e-2	9.15e-1	0.11	2.10
4^{-1}	1132	9	9.03e-2	4.89e-1	0.19	2.52
4^{-2}	2836	9	6.40e-2	2.32e-1	0.28	3.02
4^{-3}	5884	10	2.13e-2	1.21e-1	0.18	3.26
4^{-4}	15736	11	7.36e-3	4.76e-2	0.15	3.55
4^{-5}	23380	11	5.59e-3	3.12e-2	0.18	3.39

$\tilde{\eta}_{weight}$					
TOL	N	L	$\|e\|$	η	I_{eff}
1	40	4	4.67e-1	4.20e-1	1.11
4^{-1}	52	4	3.18e-1	2.45e-1	1.25
4^{-2}	88	5	1.34e-1	2.16e-1	0.63
4^{-3}	172	5	9.76e-2	1.08e-1	0.91
4^{-4}	400	6	2.99e-2	5.17e-2	0.59
4^{-5}	664	6	1.91e-2	3.03e-2	0.63
4^{-6}	1372	7	7.85e-3	1.54e-2	0.50
4^{-7}	2656	7	4.27e-3	7.31e-3	0.59
4^{-8}	5548	8	1.90e-3	3.89e-3	0.50
4^{-9}	10432	8	1.02e-3	1.94e-3	0.53

$\tilde{\eta}^{approx}_{weight}$					
TOL	N	L	$\|e\|$	η	I_{eff}
1	40	4	3.94e-1	2.76e-1	1.42
4^{-1}	52	4	3.18e-1	2.25e-1	1.42
4^{-2}	64	4	1.47e-1	1.52e-1	1.01
4^{-3}	148	5	1.08e-1	9.80e-2	1.11
4^{-4}	220	5	6.77e-2	5.24e-2	1.25
4^{-5}	592	6	2.21e-2	2.59e-2	1.83
4^{-6}	892	6	1.19e-2	1.54e-2	0.77
4^{-7}	2368	7	5.11e-3	7.17e-3	0.71
4^{-8}	3640	7	2.53e-3	3.72e-3	0.67
4^{-9}	9772	8	1.14e-3	1.70e-3	0.67

Table 6: Results obtained for Example 3 by the global L^2-error estimator $\tilde{\eta}_2$, compared to the weighted estimators $\tilde{\eta}_{weight}$ and $\tilde{\eta}^{approx}_{weight}$

illustrate this point by an analytical argument. Consider again the model Poisson problem (1.2) on the unit circle $\Omega = \{x \in \mathbb{R}^2, |x| < 1\}$. Suppose that the mean value

$$J(u) := \int_\Omega u(x)\, dx$$

has to be computed. In this case the corresponding dual solution is just the quadratic polynomial $z(x) = 1 - \frac{1}{2}|x|^2$, which satisfies $-\Delta z = 1$ and $z_{|\partial\Omega} = 0$. We think this problem to be solved by biquadratic finite elements (case $p = 2$). Then, the "optimal" *a posteriori* error bound

$$|J(e)| \leq C_i \sum_{K \in \mathbb{T}_h} \rho_K h_K^2 \|\nabla^3 z\|_K$$

would indicate that, similar as in Example 2, the error is concentrated along the boundary, while there is no contribution from interior mesh cells. In contrast, the sub-optimal *a posteriori* bound

$$|J(e)| \leq C_i \sum_{K \in \mathbb{T}_h} \rho_K h_K \|\nabla^2 z\|_K$$

would yield contributions to the error from all mesh cells.

For further illustration, we consider again the computation of the point value $u(0)$ in the above model situation. Accordingly, we compare the following two *a posteriori* error estimates,

$$|e(0)| \leq C_i \sum_{K \in \mathbb{T}_h} \frac{h_K^3}{r_K^3} \rho_K =: \eta^{(3)}_{point}(u_h)\,, \tag{5.14}$$

and

$$|e(0)| \leq C_i \sum_{K \in \mathbb{T}_h} \frac{h_K^2}{r_K^2} \rho_K =: \eta_{point}^{(2)}(u_h)\,. \tag{5.15}$$

The results shown in Table 7 clearly support our claim.

$\eta_{point}^{(3)}$				$\eta_{point}^{(2)}$			
TOL	N	L	η	TOL	N	L	η
4^{-1}	112	5	1.58e-1	4^{-1}	112	5	1.93e-1
4^{-2}	208	5	5.28e-2	4^{-2}	256	5	3.84e-2
4^{-3}	352	6	9.91e-3	4^{-3}	448	6	1.55e-2
4^{-4}	544	7	3.22e-3	4^{-4}	1264	8	2.68e-3
4^{-5}	1072	8	8.94e-4	4^{-5}	3280	9	7.69e-4
4^{-6}	1696	9	2.35e-4	4^{-6}	6688	10	2.35e-4
4^{-7}	3760	10	4.93e-5	4^{-7}	17152	11	5.10e-5
4^{-8}	7024	11	1.15e-5	4^{-8}	44032	12	1.48e-5
4^{-9}	14704	12	2.84e-6	4^{-9}	107776	13	3.50e-6
4^{-10}	25264	13	8.15e-7	4^{-10}	253888	13	8.98e-7

Table 7: Results for computing $e(0)$ by quadratic elements in Example 4, obtained with the local estimators $\eta_{point}^{(3)}(u_h)$ and $\eta_{point}^{(2)}(u_h)$

6 Generalization to nonlinear problems

The approach to residual-based error estimation described above can be extended to general nonlinear problems; for ideas in the same direction see [14], [22], [23], [16], as well as [29] and [26]. Related energy-error estimators have been derived in [29]. We outline our general concept in an abstract setting. Let V be a Hilbert space with inner product $(\cdot,\cdot)$ and corresponding norm $\|\cdot\|$, and $a(\cdot;\cdot)$ a semi-linear form continuously defined on $V \times V$. We seek a solution to the abstract variational problem

$$u \in V: \quad a(u;\varphi) = 0 \quad \forall\, \varphi \in V. \tag{6.2}$$

This problem is approximated by a Galerkin method using a sequence of finite dimensional subspaces $V_h \subset V$ parameterized by a discretization parameter h. The discrete problems read

$$u_h \in V_h: \quad a(u_h;\varphi_h) = 0 \quad \forall\, \varphi_h \in V_h. \tag{6.3}$$

With the derivative $a'(\cdot;\cdot,\cdot)$ of $a(\cdot;\cdot)$, we have the following orthogonality relation for the error $e = u - u_h$:

$$\int_0^1 a'(tu + (1-t)u_h; e, \varphi_h)\, dt = a(u;\varphi_h) - a(u_h;\varphi_h) = 0 \quad \forall \varphi_h \in V_h\,. \tag{6.4}$$

This suggests the use of the bilinear form

$$L(u, u_h; \varphi, z) = \int_0^1 a'(tu + (1-t)u_h; \varphi, z)\, dt\,, \tag{6.5}$$

depending on u as well as u_h in the duality arguments. Suppose that the quantity $J(u)$ has to be computed, where $J(\cdot)$ is a linear functional defined on V, or on a suitable subspace containing the spaces V_h and the solution u. For representing the error $J(e)$, we use the dual problem

$$L(u, u_h; \varphi, z) = J(\varphi) \quad \forall\, \varphi \in V. \tag{6.6}$$

Assuming that this problem has a (unique) solution $z \in V$, and using Galerkin orthogonality (6.4), we have the error representation

$$J(e) = L(u, u_h; e, z - z_h), \tag{6.7}$$

with any approximation $z_h \in V_h$. The goal is to evaluate the right hand side numerically, in order to get a sharp *a posteriori* estimate for the quantity $J(e)$ and thus a criterion for the optimal local adjustment of the discretization. Since the bilinear form $L(u, u_h; \cdot, \cdot)$ contains the unknown solution u as coefficient, it has to be approximated. The simplest way is to replace u by u_h yielding a perturbed dual solution $\tilde{z} \in V$ defined by

$$L(u_h, u_h; \varphi, \tilde{z}) = J(\varphi) \quad \forall\, \varphi \in V. \tag{6.8}$$

Controlling the effect of this perturbation on the accuracy of the resulting error estimator

$$J(e) \approx \tilde{\eta}(u_h) := L(u_h, u_h; e, \tilde{z} - \tilde{z}_h) \tag{6.9}$$

may be a delicate task and depends strongly on the particular problem considered. A rigorous analysis can be given in the case of ordinary differential equations (see [16]). Our own experience with the stationary Navier-Stokes equations (see [7]) indicates that this problem seems to be less critical in stable situations. The more crucial problem is the numerical computation of the perturbed dual solution $\tilde{z}$ as described above. The described approach can be further extended to the case of more general *nonlinear* control functionals. This may become particularly relevant in the context of adaptive mesh-size control in solving shape-optimization problems. Suppose that the control functional $J(\cdot)$ is *nonlinear* but differentiable with differential $J'(\cdot;\cdot)$. In this case, one would use the error representation

$$J(u) - J(u_h) = \int_0^1 J'(u_h + te; e)\, dt \approx L(u_h, u; e, z - z_h)\,, \tag{6.10}$$

with the solution z of the linearized *dual* problem

$$L(u, u_h; \varphi, z) = J'(u_h; \varphi) \quad \forall\, \varphi \in V\,. \tag{6.11}$$

A special example, for instance, is control of the L^2-error $\|e\|$. In this case one may set $J(\varphi) := \|u - \varphi\|$, to obtain $\|e\| = J(u_h) - J(u)$ and, then use $(\varphi, \bar{e}/\|\bar{e}\|) \approx (\varphi, e/\|e\|) = J'(u_h; \varphi)$ as right hand side of the dual problem. Here, $\bar{e}$ is a suitable approximation to e generated from the preceding computation on coarser meshes.

The proposed approach is rather universal and has, on a heuristic basis, already been successfully applied to rather complex problems like the Navier-Stokes equations in fluid mechanics (see [7] and [4]), the Navier-Lamé equations in linear elasticity (see [25]) and the Prandtl-Reuss model in elasto-plasticity (see [27]). It has proven particularly efficient in solving large dimensional radiative transfer problems in astrophysics (see [24]). The same concept can also be used in the context of ordinary differential equations where residual-based error estimation was introduced in [20], [17], [16], and then further developed in the spirit of this paper in [10]. Finally, weighted a posteriori error analysis may also be useful in deriving stopping criteria for iteration processes in finite element methods as studied in [8] and [4].

References

[1] I. Babuška and W.C. Rheinboldt: *Error estimates for adaptive finite element computations*, SIAM J. Num. Anal. 15, 736-754 (1978).

[2] I. Babuška and W. Gui: *Basic principles of feedback and adaptive approaches in the finite element method*, Comp. Meth. Appl. Mech. Engrg. 55, 27-42 (1986).

[3] R.E. Bank and A. Weiser: *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comput. 44, 283-301 (1985).

[4] R. Becker: *An Adaptive Finite Element Method for the Incompressible Navier-Stokes Equations on Time-Dependent Domains*, Dissertation, Institut für Angewandte Mathematik, Universität Heidelberg, 1995.

[5] R. Becker: *An adaptive finite element method for the Stokes equations including control of the iteration error*, Proc. ENUMATH 1995, Paris, 18-22 Sept., 1995, to appear.

[6] R. Becker and R. Rannacher: *Finite element solution of the incompressible Navier-Stokes equations on anisotropically refined meshes*, Proc. Tenth GAMM-Seminar, Kiel, Jan. 1994, (W.Hackbusch, G.Wittum, eds.), Notes on Num. Fluid Mech., Vieweg, Braunschweig, 1995.

[7] R. Becker and R. Rannacher: *Weighted a posteriori error control in FE methods*, Preprint 96-1 (SFB 359), Universität Heidelberg, Januar 1996, Proc. ENUMATH-95, Paris, 18-22 Sept., 1995, to appear.

[8] R. Becker, C. Johnson and R. Rannacher: *Adaptive error control for multigrid finite element methods*, Computing 55, 271-288 (1995).

[9] Bernardi, C., Bonnon, O., Langouët, C., and Métivet, B. : *Residual error indicators for linear problems: Extension to the Navier-Stokes equations*, Proc. 9th Int. Conf. Finite Elements in Fluids, Venezia, 1995.

[10] K. Böttcher and R. Rannacher: *Adaptive error control in solving ordinary differential equations by the discontinuous Galerkin method*, Preprint, Institut für Angewandte Mathematik, Universität Heidelberg, Oct. 1996.

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

[12] W. Dörfler and M. Rumpf: *An adaptive strategy for elliptic problems including a posteriori controlled boundary approximation*, Preprint Nr. 1/1996-16.01.1996, Mathematische Fakultät, Universität Freiburg (1996).

[13] K. Eriksson and C. Johnson: *An adaptive finite element method for linear elliptic problems*, Math. Comp. 50, 361-383 (1988).

[14] K. Eriksson and C. Johnson: *Adaptive finite element methods for parabolic problems: Nonlinear problems*, Preprint, Department of Mathematics, Chalmers University of Technology, Gothenburg, 1992.

[15] K. Eriksson, D. Estep, P. Hansbo and C. Johnson: *Adaptive Finite Elements*, Springer, Berlin-Heidelberg-New York, 1996.

[16] D. Estep: *A posteriori error bounds and global error control for approximation of ordinary differential equations*, SIAM J. Numner. Anal. 32, 1-48 (1995).

[17] D. Estep and C. Johnson: *The computation of the Lorenz system*, Preprint, Department of Mathematics, Chalmers Technical University, Gothenburg, 1994.

[18] J. Frehse and R. Rannacher: *Eine L^1-Fehlerabschätzung für diskrete Grundlösungen in der Methode der finiten Elemente*, Proc. Workshop Finite Elemente, Bonn. Math. Schr. 89, 92-114, 1976.

[19] C. Führer: *Error Control in Finite Element Methods for Hyperbolic Problems* Dissertation, Institut für Angewandte Mathematik, Universität Heidelberg, in preparation.

[20] C. Johnson: *Error estimates and adaptive time-step control for a class of one-step methods for stiff ordinary differential equations*, SIAM J. Numer. Anal. 908-926 (1988).

[21] C. Johnson and P. Hansbo: *Adaptive finite elements in computational mechanics* in Computer Methods in Applied Mechanics and Engineering 101, North-Holland, Amsterdam, 1992.

[22] C. Johnson and R. Rannacher: *On error control in CFD*, Proc. Int. Workshop Num. Meth. Navier-Stokes Equations, Heidelberg, 25-28 Oct., 1993, Notes Num. Fluid Mech., Vol. 47, Vieweg, Braunschweig, 1994.

[23] C. Johnson, R. Rannacher and M. Boman: *Numerics and hydrodynamic stability: Towards error control in CFD*, SIAM J. Numer. Anal. 32, 1058-1079 (1995).

[24] G. Kanschat: *Parallel and Adaptive Galerkin Methods for Radiative Transfer Problems*, Dissertation, Institut für Angewandte Mathematik, Universität Heidelberg, 1996.

[25] R. Rannacher and F.-T. Suttmeier: *A feed-back approach to error control in finite element methods: Application to linear elasticity*, Preprint, Institut für Angewandte Mathematik, Universität Heidelberg, Sept. 1996, to appear in Comp. Mech.

[26] K.-G. Siebert: *A posteriori error estimation for nonlinear problems by duality techniques* Preprint, Institut für Angewandte Mathematik, Universität Freiburg, 1995.

[27] F.-T. Suttmeier: *Adaptive Finite Element Approximation of Problems in Elasto-Plasticity Theory*, Dissertation, Institut für Angewandte Mathematik, Universität Heidelberg, in preparation.

[28] R. Verfürth: *A posteriori error estimation and adaptive mesh-refinement techniques*, J. Comp. Appl. Math. 50, 67-83 (1994).

[29] R. Verfürth: *A posteriori error estimates for nonlinear problems: Finite element discretizations of elliptic problems*, Math. Comp. 62, 445-475 (1994).

[30] R. Verfürth: *A Review of A Posteriori Estimation and Adaptive Mesh-Refinement Techniques*, Advances in Numerical Mathematics, Wiley/Teubner, New York-Stuttgart, 1996, to appear.

[31] R. Verfürth: *A posteriori error estimates for low order finite elements*, Preprint, Fakultät für Mathematik, Universität Bochum, March 1996.

[32] D. Yu: *Asymptotically exact a posteriori error estimators for elements of bi-odd degree*, Chinese J. Numer. Math. Appl. 13, 64-78 (1991).

[33] D. Yu: *Asymptotically exact a posteriori error estimators for elements of bi-even degree*, Chinese J. Numer. Math. Appl. 13, 82-790 (1991).

[34] O.C. Zienkiewicz and J.Z. Zhu: *A simple error estimator and adaptive procedure for practical engineering analysis*, Int. J. Numer. Methods Engrg. 24 (1987).