MOX–Report No. 02/2007

A safeguarded dual weighted residual method

Ricardo H. Nochetto, Andreas Veeser,
Marco Verani

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 29 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

A safeguarded dual weighted residual method *

Ricardo H. Nochetto[a], Andreas Veeser[b] and Marco Verani[c]

9th, january 2007

[a] Department of Mathematics and Institute for Physical Science and Technology, University of Maryland, College Park, 20742 Maryland, U.S.A.

[b] Dipartimento di Matematica "F. Enriques", Università degli Studi di Milano, Milano, Italy.

[c] MOX - Modelling and Scientific Computing - Dipartimento di Matematica "F. Brioschi", Politecnico di Milano, Milano, Italy

Abstract

The dual weighted residual (DWR) method yields reliable a posteriori error bounds for linear output functionals provided that the error incurred by the numerical approximation of the dual solution is negligible. In that case its performance is generally superior than that of global 'energy norm' error estimators which are 'unconditionally' reliable. We present a simple numerical example for which neglecting the approximation error leads to severe underestimation of the functional error, thus showing that the DWR method may be unreliable. We propose a remedy that preserves the original performance, namely a DWR method safeguarded by additional asymptotically higher order a posteriori terms. In particular, the enhanced estimator is unconditionally reliable and asymptotically coincides with the original DWR method. These properties are illustrated via the aforementioned example.

Keywords: a posteriori estimates, output functionals, finite elements, duality.

AMS subject classification: 65N15, 65N30, 65N50

1 Goal-oriented error estimation by duality and outline

Suppose that we are given the prototype elliptic boundary value problem

$$-\Delta u = f \quad \text{in} \quad \Omega, \qquad u = 0 \quad \text{on} \quad \Gamma = \partial\Omega \tag{1}$$

and a linear output functional G such that $G(u)$ is a quantity of physical, engineering, or scientific interest. In order to approximate $G(u)$, one may compute $G(u^1_{\mathcal{T}})$, where $u^1_{\mathcal{T}} \in \mathbb{V}^1_{\mathcal{T}}(\Omega)$ is the linear finite element approximation to u over a conforming triangulation

*Research partially supported by NSF grants DMS-0204670, DMS-0505454, and by Italian MIUR PRIN 2006 "Metodi numerici avanzati per il calcolo scientifico".

$\mathcal{T}$ of Ω. We are interested in estimating the approximation error $|G(u) - G(u_{\mathcal{T}}^1)| = |G(u - u_{\mathcal{T}}^1)|$.

Several approaches have been proposed for this task; see [1, 2, 12, 13, 14, 15]. A key device is duality [6, 7, 8, 9]. In fact, combining the solution of the dual problem

$$-\Delta z = G \quad \text{in} \quad \Omega, \qquad z = 0 \quad \text{on} \quad \Gamma = \partial\Omega \tag{2}$$

with Galerkin orthogonality yields the representation formula

$$G(u - u_{\mathcal{T}}^1) = \langle -\Delta z, u - u_{\mathcal{T}}^1 \rangle = \big\langle \nabla z, \nabla(u - u_{\mathcal{T}}^1) \big\rangle = \big\langle \nabla(u - u_{\mathcal{T}}^1), \nabla(z - z_{\mathcal{T}}^1) \big\rangle, \tag{3}$$

where $\langle \cdot, \cdot \rangle$ stands for the duality pairing $H^{-1}(\Omega) \times H_0^1(\Omega)$ or the scalar product in $L^2(\Omega)$ and $z_{\mathcal{T}}^1$ denotes the linear finite element approximation of z. Moreover, invoking also the residual $\mathcal{R}(u_{\mathcal{T}}^1) = f + \Delta u_{\mathcal{T}}^1 \in H^{-1}(\Omega)$ of the primal solution and splitting it into element contributions leads to the alternative representation formula:

$$G(u - u_{\mathcal{T}}^1) = \langle \mathcal{R}(u_{\mathcal{T}}^1), z - z_{\mathcal{T}}^1 \rangle = \sum_{T \in \mathcal{T}} \langle R, z - z_{\mathcal{T}}^1 \rangle_T - \langle J, z - z_{\mathcal{T}}^1 \rangle_{\partial T}, \tag{4}$$

where R and J denote the interior and jump residual of the approximate primal solution $u_{\mathcal{T}}^1$ (for their precise definitions, see (10) below). These two error representations give rise to two basic ideas:

- Estimate the right hand side of (3) by means of global 'energy norm' techniques.
- Eliminate the explicit appearance of the not computable dual solution z from the right hand side of (4).

The realization of the first idea requires a passage to global energy norms, which can be accomplished by the Cauchy-Schwarz inequality

$$\big\langle \nabla(u - u_{\mathcal{T}}^1), \nabla(z - z_{\mathcal{T}}^1) \big\rangle \leq \|\nabla(u - u_{\mathcal{T}}^1)\|_{0,\Omega} \|\nabla(z - z_{\mathcal{T}}^1)\|_{0,\Omega}; \tag{5}$$

hereafter $\|\cdot\|_{s,\omega}$ indicates the norm of $H^s(\omega)$. Since the energy norms on the right hand side can be estimated in an unconditionally reliable way, the ensuing estimator is also unconditionally reliable. However, a global step like (5) does not take into account possible 'cancellations' in the product $\langle \nabla(u - u_{\mathcal{T}}^1), \nabla(z - z_{\mathcal{T}}^1) \rangle$ and may thus entail a severe overestimation: in the case of (5), this is measured by the angle between $u - u_{\mathcal{T}}^1$ and $z - z_{\mathcal{T}}^1$ in $H_0^1(\Omega)$, which depends on the interplay of f and G.

The second idea offers a way to exploit cancellations. In fact, the appearance of $z - z_{\mathcal{T}}^1$ in the right hand side of (4) may be interpreted as a locally varying weight to the residual $\mathcal{R}(u_{\mathcal{T}}^1)$ of the primal solution. This interpretation seems to be particularly useful in instances when the dual solution exhibits complex behavior and its size varies substantially in the computational domain Ω. In this spirit Becker and Rannacher [2] have proposed the *Dual Weighted Residual* (DWR) estimate

$$|G(u - u_{\mathcal{T}}^1)| \leq \sum_{T \in \mathcal{T}} \rho_T \omega_T, \tag{6}$$

where

$$\rho_T := \|R\|_{0,T} + h_T^{-1/2}\|J\|_{0,\partial T} \quad \text{and} \quad \omega_T := \|z - z_{\mathcal{T}}^1\|_{0,T} + h_T^{1/2}\|z - z_{\mathcal{T}}^1\|_{0,\partial T}$$

and h_T is the diameter of T. Strictly speaking, (6), as it stands, is not really a classical a posteriori error bound, because it involves the unknown solution z to the dual problem. To approximate the dual solution z numerically, there are two different approaches [3]:

- Post-process the linear finite element solution $z_{\mathcal{T}}^1$ of (2) over $\mathcal{T}$, the mesh corresponding to the approximate primal solution $u_{\mathcal{T}}^1$.
- Approximately solve the dual problem (2) by means of a higher order method or on a mesh that is finer than $\mathcal{T}$.

Of course, both approaches yield a reliable variant of (6) only if the corresponding approximation error is negligible. If the approximation of z is of higher order, then this is guaranteed for a sufficiently small (global) meshsize. In this respect, E. Süli in his review [18] of the book [1] writes:

> *"The additional errors incurred through the numerical approximation of the dual solution are difficult to quantify unless one embarks on reliable a posteriori error estimation for the dual problem; for reasons of economy, this is rarely attempted in practice. Indeed, there is very little in the current literature in the way of rigorous analytical quantification of the impact of replacing the exact dual solution z in the DWR error bound by its numerical approximation; see, however, the recent analytical work of Carstensen [4] on the estimation of higher Sobolev norms from lower order approximation, and the application of this in the context of the DWR method. A second issue is that the necessity to obtain a "reasonably" accurate approximation to the dual solution adds computational work."*

In this vein, it is worth stressing that [4] gives a justification for the recovery of z from $z_{\mathcal{T}}^1$ provided that the global meshsize is sufficiently small. Thus, at least from a theoretical viewpoint, the issue of reliability remains open for relatively coarse or strongly graded meshes – two cases that are quite important in an adaptive context. Notice also that the notion 'sufficiently small meshsize', and so the guaranteed reliability of the computational DWR method, is strongly problem-dependent, quite vague, and difficult to verify in practice.

These serious theoretical drawbacks are not just 'academic'. In §2 we show with the help of a simple, far from pathological, example that the DWR method with a higher order approximation of the dual solution is unreliable on relatively coarse meshes, i.e. in early stages of adaptivity when critical decisions have to be made. Consequently, even if one restricts to the model problem class described by (1) and (2), the DWR method should not be implemented as a 'black box' algorithm.

We propose a remedy that preserves the advantages of the DWR estimate. We derive in §3 the a posteriori upper bound

$$|G(u - u_{\mathcal{T}}^1)| \leq \mathcal{E}_1(\mathcal{T}) + \mathcal{E}_2(\mathcal{T}), \tag{7}$$

where $\mathcal{E}_1(\mathcal{T})$ is a computable version of the right hand side of (6), based upon the quadratic finite element approximation of the dual solution z, and $\mathcal{E}_2(\mathcal{T})$ is an a posteriori estimate of the error incurred by the approximation of z, which in the original DWR method is neglected. Thanks to the 'safeguarding' part $\mathcal{E}_2(\mathcal{T})$, the bound (7) holds also on relatively coarse and strongly graded meshes. Consequently, $\mathcal{E}_1(\mathcal{T}) + \mathcal{E}_2(\mathcal{T})$ is unconditionally reliable. Moreover, since $\mathcal{E}_2(\mathcal{T})$ estimates an error that is formally of higher order, we can expect that $\mathcal{E}_1(\mathcal{T}) + \mathcal{E}_2(\mathcal{T})$ asymptotically coincides with the usual DWR method, thus essentially maintaining its performance. Both properties are illustrated by our numerical experiments in §4, where we apply (7) to the example of §2.

The validity of the usual DWR method hinges upon a closeness assumption on the numerical approximation to the dual solution, which is strongly problem-dependent and thus problematic: it depends on the differential operator, the domain and the output functional. It thus appears very difficult to design an abstract and general strategy for estimating the error incurred on substituting the dual solution by a numerical approximation. Therefore, our contribution is inevitably rather modest in scope since we restrict the discussion to the simple model (1). Despite its simplicity, the behavior of the DWR method can already be troublesome for (1). In order to avoid any closeness assumption on the numerical dual solution, we need to incorporate additional geometric information about the interaction of the elliptic operator and the domain, namely the structure of corner singularities. Such information is often available for most important elliptic operators in applications, both in 2D and 3D, but the ensuing safeguarded estimator (7) is far from a problem-independent tool. On the other hand, typical functionals G are unrelated to corner singularities and so their explicit use in constructing an estimator that gets around the closeness assumption on the numerical dual solution seems to be justified. We hope that this first attempt to make the DWR method reliable will promote further research on this important but unexplored territory.

2 Example: the DWR method is unreliable

We first discuss the computational DWR method to be used in what follows and then present an example for which it fails.

Define u, $u^1_\mathcal{T}$, and z as in §1. If $z^2_\mathcal{T} \in \mathbb{V}^2_\mathcal{T}(\Omega)$ is the continuous piecewise quadratic finite element approximation to z, then we may write

$$G(u - u^1_\mathcal{T}) = \langle \mathcal{R}(u^1_\mathcal{T}), z^2_\mathcal{T} \rangle + \langle \mathcal{R}(u^1_\mathcal{T}), z - z^2_\mathcal{T} \rangle. \tag{8}$$

Observe that the first term on the right hand side is computable and can be rewritten upon using the definition of $u^1_\mathcal{T}$ and integrating by parts elementwise:

$$\langle \mathcal{R}(u^1_\mathcal{T}), z^2_\mathcal{T} \rangle = \sum_{T \in \mathcal{T}} \langle R, z^2_\mathcal{T} - I^1_\mathcal{T} z^2_\mathcal{T} \rangle_T - \langle J, z^2_\mathcal{T} - I^1_\mathcal{T} z^2_\mathcal{T} \rangle_{\partial T}, \tag{9}$$

where $I^1_\mathcal{T}$ is the Lagrange interpolation operator onto $\mathbb{V}^1_\mathcal{T}(\Omega)$ and the element and interelement residuals R and J are defined as follows: for each triangle T or each side S

of $\mathcal{T}$, respectively, there holds

$$R = f, \qquad J = \begin{cases} \frac{1}{2}[\nu_S \cdot \nabla u_{\mathcal{T}}^1], & \text{if } S \subset \partial T \setminus \partial\Omega, \\ 0, & \text{if } S \subset \partial\Omega, \end{cases} \tag{10}$$

where $[\nu_S \cdot \nabla u_{\mathcal{T}}^1]$ denotes the jump of the normal flux across the interelement side S (which does not depend on the orientation of ν_S). The second term in (8), however, depends on the unknown dual solution z and thus, in general, is not computable.

According to [3, (3.60)], $\langle \mathcal{R}(u_{\mathcal{T}}^1), z - z_{\mathcal{T}}^2 \rangle$ can be neglected from (8) for a *sufficiently fine initial mesh*. In this case, (9) suggests the a posteriori error estimator

$$\mathcal{E}_1(\mathcal{T}) := \Big| \sum_{T\in\mathcal{T}} \langle R, z_{\mathcal{T}}^2 - I_{\mathcal{T}}^1 z_{\mathcal{T}}^2 \rangle_T - \langle J, z_{\mathcal{T}}^2 - I_{\mathcal{T}}^1 z_{\mathcal{T}}^2 \rangle_{\partial T} \Big| \le \sum_{T\in\mathcal{T}} \Big| \langle R, z_{\mathcal{T}}^2 - I_{\mathcal{T}}^1 z_{\mathcal{T}}^2 \rangle_T - \langle J, z_{\mathcal{T}}^2 - I_{\mathcal{T}}^1 z_{\mathcal{T}}^2 \rangle_{\partial T} \Big|. \tag{11}$$

Notice that $\mathcal{E}_1(\mathcal{T})$ allows for interelement cancellations, while the right hand side of (11) does not; however, the latter consists of positive element contributions. This suggests to mark elements to be refined with the help of the indicators

$$\Big| \langle R, z_{\mathcal{T}}^2 - I_{\mathcal{T}}^1 z_{\mathcal{T}}^2 \rangle_T - \langle J, z_{\mathcal{T}}^2 - I_{\mathcal{T}}^1 z_{\mathcal{T}}^2 \rangle_{\partial T} \Big|, \quad T \in \mathcal{T}, \tag{12}$$

but to stop the algorithm employing $\mathcal{E}_1(\mathcal{T})$. Regarding the sharpness of $\mathcal{E}_1(\mathcal{T})$ and its relationship with (6), it is instructive to observe that

$$\begin{aligned} \mathcal{E}_1(\mathcal{T}) &= |\langle \mathcal{R}(u_{\mathcal{T}}^1), z_{\mathcal{T}}^2 \rangle| = |\langle \mathcal{R}(u_{\mathcal{T}}^1), z_{\mathcal{T}}^2 - z_{\mathcal{T}}^1 \rangle| \\ &\le \sum_{T\in\mathcal{T}} \Big| \langle R, z_{\mathcal{T}}^2 - z_{\mathcal{T}}^1 \rangle_T - \langle J, z_{\mathcal{T}}^2 - z_{\mathcal{T}}^1 \rangle_{\partial T} \Big| \le \sum_{T\in\mathcal{T}} \rho_T \tilde{\omega}_T, \end{aligned} \tag{13}$$

with

$$\tilde{\omega}_T := \|z_{\mathcal{T}}^2 - z_{\mathcal{T}}^1\|_{0,T} + h_T^{1/2} \|z_{\mathcal{T}}^2 - z_{\mathcal{T}}^1\|_{0,\partial T}, \quad \forall T \in \mathcal{T}.$$

This means that $\mathcal{E}_1(\mathcal{T})$ is bounded by a computable version of the DWR estimator in (6). The steps in (13) do not incorporate additional information about the dual solution z and, thus, $\sum_{T\in\mathcal{T}} \rho_T \tilde{\omega}_T$ is just bigger than $\mathcal{E}_1(\mathcal{T})$, without actually increasing reliability. Therefore, in what follows, we consider $\mathcal{E}_1(\mathcal{T})$ as the original DWR estimator.

We implemented an adaptive algorithm aiming at $|G(u - u_{\mathcal{T}}^1)| \le \mathtt{tol}$ for a prescribed tolerance $\mathtt{tol} > 0$ within the framework of the finite element toolbox ALBERTA [16]. The algorithm stops when $\mathcal{E}_1(\mathcal{T}) \le \mathtt{tol}$ is satisfied and marks elements for refinement with the help of the equidistribution strategy, the indicators of which are given by (12).

Let us consider (1) with the L-shaped domain $\Omega := (-2,2)^2 \setminus (-2,0) \subset \mathbb{R}^2$ and the constant load $f = 1$ in Ω. We are interested in approximating $\partial_y u(x_0, y_0)$ with $(x_0, y_0) = (-0.5, 1)$. Since $u \mapsto \partial_y u(x_0, y_0)$ is not a continuous functional in $H_0^1(\Omega)$, we resort to the following approximation G of the y-derivative of the Dirac mass in $(x_0, y_0) = (-0.5, 1)$:

$$G(x,y) = -\frac{a(y - y_0)}{b + ((x - x_0)^2 + (y - y_0)^2)^{2.5}} \quad \text{with} \quad a = 3,\ b = 10^{-4};$$

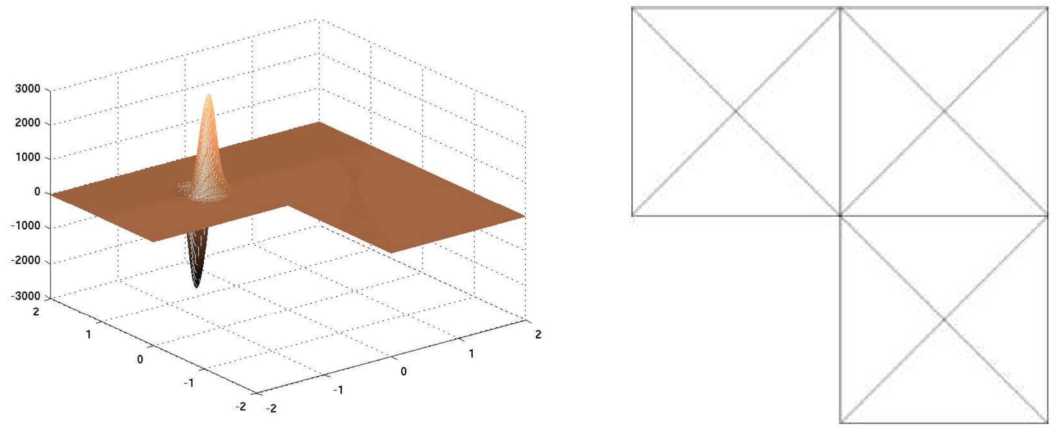

Figure 1: Plot of the regularized $\partial_y \delta_{(x_0,y_0)} \notin H^{-1}(\Omega)$ (left). Initial (and final, using original DWR) mesh (right).

see also Figure 1 (left). We require `tol` $= 0.1$ and start from the mesh $\mathcal{T}_0$ in Figure 1 with 12 triangles, for which the point (x_0, y_0) is not a (linear or quadratic) node. It turns out that

$$\mathcal{E}_1(\mathcal{T}_0) = 0.070921 \quad \text{but} \quad |G(u - u^1_{\mathcal{T}_0})| = 1.66290594.$$

The value of $\mathcal{E}_1(\mathcal{T}_0)$ causes the instantaneous stop of the DWR method, while the approximation error $|G(u - u^1_{\mathcal{T}})|$ in the quantity of interest is still large. We conclude that discarding $\langle \mathcal{R}(u^1_{\mathcal{T}}), z - z^2_{\mathcal{T}} \rangle$ in (8) leads to severe underestimation in that the actual error is about 25 times larger than the estimator. The use of $\mathcal{E}_1(\mathcal{T})$ without any further reliability checks may thus lead to unreliable results.

This behavior is not surprising since on the mesh $\mathcal{T}_0$ the functional G is quasi-orthogonal to the space $\mathbb{V}^2_{\mathcal{T}}(\Omega)$ of continuous piecewise quadratic finite elements. Therefore, neglecting the second term in (8) has a devastating effect of almost complete loss of information: the neglected term contains most of the information on G! To convince ourselves that this elementary example is not that special a simple counting argument suffices. The 'zoo' of pairs of functionals G and coarse meshes $\mathcal{T}_0$ exhibiting the same critical behavior can be (potentially) dangerous in practical situations in which cancellation, the very reason for using the DWR method, may be difficult or impossible to detect beforehand.

3 A safeguarded DWR estimator

As illustrated in §2, the DWR is in general unreliable. However, exploitation of cancellations, as described in §1, remains a valid conceptual advantage. Therefore, the design of a reliable variant of the DWR method is a highly desirable objective. To this end, we observe that the term neglected in the DWR method of §2, that is the second term in the right hand side of (8), satisfies

$$\langle \mathcal{R}(u_{\mathcal{T}}^1), z - z_{\mathcal{T}}^2 \rangle = \langle \nabla(u - u_{\mathcal{T}}^1), \nabla(z - z_{\mathcal{T}}^2) \rangle = \langle \nabla(u - u_{\mathcal{T}}^2), \nabla(z - z_{\mathcal{T}}^2) \rangle, \tag{14}$$

where $u_{\mathcal{T}}^2$ denotes the quadratic finite element approximation to u and the Galerkin orthogonality for $z_{\mathcal{T}}^2$ has been used. Identity (14) reveals that, depending on the regularity of u and z, the neglected term is of higher order. Notice also that the 'higher order contribution' associated with z hinges on G and, thus, is affected by possibile regularizations as in §2.

The right hand side of (14) can be estimated by means of global energy norm techniques. Inserting these estimates in (8) leads to a DWR estimate that is safeguarded by higher order terms. However, such an approach additionally requires the computation of $u_{\mathcal{T}}^2$, which, when at hand, suggests the use of $G(u_{\mathcal{T}}^2)$ as approximate value for $G(u)$. We thus end up with the classical dilemma of a posteriori error estimation by higher order terms.

Here we shall pursue a different approach that relies on the assumption

$$f \in H^{1+\epsilon}(\Omega) \tag{15}$$

with $\epsilon \geq 0$ to be specified below (see Remark 3.1) and the identity

$$\langle \mathcal{R}(u_{\mathcal{T}}^1), z - z_{\mathcal{T}}^2 \rangle = \langle \nabla(z - z_{\mathcal{T}}^2), \nabla(u - \Pi_{\mathcal{T}}^2 u) \rangle = \langle \mathcal{R}^*(z_{\mathcal{T}}^2), u - \Pi_{\mathcal{T}}^2 u \rangle, \tag{16}$$

where $\mathcal{R}^*(z_{\mathcal{T}}^2) := G + \Delta z_{\mathcal{T}}^2 \in H^{-1}(\Omega)$ denotes the dual residual, and $\Pi_{\mathcal{T}}^2 u$ is an appropriate second order interpolant of u. The term $\langle \mathcal{R}^*(z_{\mathcal{T}}^2), u - \Pi_{\mathcal{T}}^2 u \rangle$ can be estimated with the help of the techniques in [11]. The computation of $u_{\mathcal{T}}^2$ is not required and, thanks to the assumption (15), there is a higher order contribution that is independent of G; see Remark 3.2.

Apart from (15), we suppose that the data of (1) and (2) satisfies

$$\Omega \subset \mathbb{R}^2 \text{ is a bounded, polygonal domain and } G \in L^2(\Omega). \tag{17}$$

3.1 Corner singularities

Assumption (15) implies extra regularity for the primal solution u, which has to be exploited for a sharp estimate of the last term $\langle \mathcal{R}^*(z_{\mathcal{T}}^2), u - \Pi_{\mathcal{T}}^2 u \rangle$ in (16). However, the full shift theorem, which would imply $u \in H^{3+\epsilon}(\Omega)$, is spoiled by singularities at the corners of Ω. In this subsection we recall a decomposition of the solution u of (1) into a smooth part with regularity compatible with the shift theorem and singular functions associated to the corners of Ω; this decomposition is a main tool for our estimate of $\langle \mathcal{R}^*(z_{\mathcal{T}}^2), u - \Pi_{\mathcal{T}}^2 u \rangle$ in (16).

Let $\Gamma_1, \dots, \Gamma_M$ be the segments of $\Gamma = \partial\Omega$ and let $P_1, \dots, P_M$ be the vertices arranged in a counterclockwise order so that P_l and P_{l+1} are the two vertices of Γ_l and Γ_l, Γ_{l-1} are the two sides emanating from P_l. Let $\omega_l \in (0, 2\pi]$ denote the interior angle of Γ

at P_l, measured in a counterclockwise direction from Γ_l to Γ_{l-1}. Given the smoothness parameter $s \geq 2$, define

$$J_l(s) = \max\{j \in \mathbb{N} \mid j\alpha_l < s - 1\}$$

with $\alpha_l = \pi/\omega_l \in [1/2, \infty)$ and the convention $\max \emptyset = -\infty$. We finally introduce the singular functions $\psi_j^{(l)}$ associated with each vertex. Let the polar coordinates (r_l, θ_l) be chosen at vertex P_l so that the interior angle ω_l is spanned by the two half lines $\theta_l = 0$ and $\theta_l = \omega_l$. The singular functions can be written as

$$\psi_j^{(l)}(r_l, \theta_l) = \phi_l(r_l) v_j^{(l)}(r_l, \theta_l),$$

where $\phi_l(r_l)$ are smooth cut-off functions which equal 1 in a neighborhood of $r_l = 0$ and possess non-overlapping support, whereas $v_j^{(l)}$ are given by

$$v_j^{(l)}(r_l, \theta_l) := \begin{cases} r_l^{j\alpha_l} \sin(j\alpha_l\theta_l), & \text{if} \quad j\alpha_l \notin \mathbb{N}, \\ r_l^{j\alpha_l}\big(\theta_l \cos(j\alpha_l\theta_l) + (\log r_l) \sin(j\alpha_l\theta_l)\big), & \text{if} \quad j\alpha_l \in \mathbb{N}. \end{cases}$$

Notice that, if $j\alpha_l \in \mathbb{N}$, then $v_j^{(l)}$ and so $\psi_j^{(l)}$ does not have homogeneous Dirichlet boundary values.

Theorem 3.1 (Shift theorem with corner singularities). *Let $s \geq 2$ be such that $(s-1)/\alpha_l \notin \mathbb{N}$ for every $l = 1, \ldots, M$. Then $f \in H^{s-2}(\Omega)$ entails that the solution u of (1) satisfies*

$$u = \sum_{l=1}^{M} \sum_{j=1}^{J_l(s)} \Lambda_j^{(l)} \psi_j^{(l)} + w \tag{18}$$

with

$$\|w\|_{s,\Omega} + \sum_{l=1}^{M} \sum_{j=1}^{J_l(s)} |\Lambda_j^{(l)}| \leq C\|f\|_{s-2,\Omega}. \tag{19}$$

Proof. See [10, Thm 3.1], [5]. □

3.2 Weighted Residual Estimator

The effect of the corner singularities in §3.1 is encoded into a suitable weight for the ensuing estimator of $\langle \mathcal{R}^*(z_{\mathcal{T}}^2), u - \Pi_{\mathcal{T}}^2 u\rangle$. In this subsection we first introduce this weight and then define the estimator.

The aforementioned weight is built up with the following modified distance functions $\rho_l : \bar{\Omega} \to \mathbb{R}^+$, $l = 1, \ldots, M$:

$$\rho_l(x) := \sqrt{r_l(x)^2 + h(x)^2}, \qquad \forall x \in \bar{\Omega},$$

where $r_l(x) := |x - P_l|$ is the Euclidean distance between x and the vertex P_l of Ω and h denotes the piecewise constant meshsize of $\mathcal{T}$, that is the restriction of h to any $T \in \mathcal{T}$

equals the diameter of T, $h_{|T} = h_T$. Notice that ρ_l is strictly positive in $\bar{\Omega}$ and so in particular in P_l. Moreover, as we will see in a moment, there hold

$$P_l \notin T \quad \Longrightarrow \quad r_l(x) < \rho_l(x) \leq C r_l(x), \quad \forall x \in T \tag{20}$$

as well as the *non-oscillation* property

$$\max_{x \in N(T)} \rho_l \leq \Lambda \min_{x \in N(T)} \rho_l(x), \quad \forall T \in \mathcal{T}, \tag{21}$$

where C and Λ solely depend on the minimum angle of $\mathcal{T}$; hereafter, the neighborhood

$$N(A) := \bigcup \{T' \in \mathcal{T} : A \cap T' \neq \emptyset\} \tag{22}$$

of a given closed set A is the union of all elements of the partition $\mathcal{T}$ adjacent A. To prove the nontrivial inequality of (20), we recall that

$$\operatorname{dist}(T, \Omega \setminus N(T)) \geq C h_T, \quad \forall T \in \mathcal{T},$$

with $C > 0$ solely depending on the minimum angle of $\mathcal{T}$. Therefore, for elements T not containing corner P_l,

$$r_l(x) \geq \operatorname{dist}(T, P_l) \geq \operatorname{dist}(T, \Omega \setminus N(T)) \geq C h_T = C h(x), \quad \forall x \in T, \tag{23}$$

whence $\rho_l \leq C r_l$ in T. To prove (21), we first observe that there hold $\max_{N(T)} h \leq C \min_{N(T)} h$ with C as above and $\max_{N(T)} r_l \leq \min_{N(T)} r_l + 3 \max_{N(T)} h$ since r_l is Lipschitz with constant 1. Thus, if $N(T) \ni P_l$, then (21) follows from $\min_{N(T)} r_l = 0$ and $h \leq \rho_l$ and, if $N(T) \not\ni P_l$, from (20).

Given a M-tuple $\gamma = (\gamma_l)_{l=1}^M$ with nonnegative coefficients, we define the mesh dependent weight functions

$$\sigma_\gamma(x) := \min_{1 \leq l \leq M} \rho_l^{\gamma_l}(x) \quad \text{and} \quad \sigma_{-\gamma}(x) := \frac{1}{\sigma_\gamma(x)}, \qquad \forall x \in \Omega, \tag{24}$$

which also satisfy (21). The error indicator for a triangle $T \in \mathcal{T}$ is given by

$$\eta^*_{-1,\gamma}(T)^2 := h_T^6 \|R^* \sigma_\gamma\|^2_{0,T} + h_T^5 \|J^* \sigma_\gamma\|^2_{0,\partial T}, \tag{25}$$

where the dual element and interelements residuals R^* and J^* are defined as follows: for each triangle T and side S of $\mathcal{T}$,

$$R^* = G + \Delta z_{\mathcal{T}}^2, \qquad J^* = \begin{cases} \frac{1}{2}[\nu_S \cdot \nabla z_{\mathcal{T}}^2], & \text{if } S \subset \partial T \setminus \partial\Omega, \\ 0, & \text{if } S \subset \partial\Omega. \end{cases} \tag{26}$$

The global estimator is then defined by summing the squared indicators:

$$\eta^*_{-1,\gamma}(\mathcal{T})^2 := \sum_{T \in \mathcal{T}} \eta^*_{-1,\gamma}(T)^2. \tag{27}$$

Notice that

$$\eta^*_{-1,\gamma}(T) \leq h_T^2 (\max_T \sigma_\gamma) \eta_1^*(T), \tag{28}$$

where $\eta_1^*(T)^2 := h_T^2 \|R^*\|^2_{0,T} + h_T \|J^*\|^2_{0,\partial T}$ denotes the usual energy norm error indicator for $z_{\mathcal{T}}^2$.

3.3 Upper bound

We now prove the main result of this article, an a posteriori upper bound for the error $|G(u) - G(u^1_{\mathcal{T}})|$ in the quantity of interest in terms of the DWR estimator (11) and additional higher order safeguarding terms.

Theorem 3.2 (Upper bound). *Suppose that* (15) *and* (17) *are valid with*

$$\epsilon \geq 0 \text{ such that } (2+\epsilon)/\alpha_l \notin \mathbb{N} \text{ for every } l = 1, \dots, M, \tag{29}$$

where $\alpha_l := \pi/\omega_l$. *Then the error of the approximation* $G(u^1_{\mathcal{T}})$ *associated to the continuous linear finite element solution* $u^1_{\mathcal{T}} \in \mathbb{V}^1_{\mathcal{T}}(\Omega)$ *of* (1) *is bounded in the following way:*

$$|G(u) - G(u^1_{\mathcal{T}})| \leq \mathcal{E}_1(\mathcal{T}) + \mathcal{E}_2(\mathcal{T}),$$

where

$$\mathcal{E}_1(\mathcal{T}) = \Big| \sum_{T\in\mathcal{T}} \langle R, z^2_{\mathcal{T}} - I^1_{\mathcal{T}} z^2_{\mathcal{T}}\rangle_T - \langle J, z^2_{\mathcal{T}} - I^1_{\mathcal{T}} z^2_{\mathcal{T}}\rangle_{\partial T}\Big|, \quad \mathcal{E}_2(\mathcal{T}) = C \max_{l=1}^{M} L(h_l)^{3/2} \eta^*_{-1,-\beta}(\mathcal{T}) \|f\|_{1+\epsilon,\Omega},$$

$z^2_{\mathcal{T}}$ *denotes the continuous piecewise quadratic finite element solutions of* (2), $L(h_l) := 1 + |\log h_l|$ *with* h_l *indicating the minimum length of all sides emanating from* P_l, *and* $\eta^*_{-1,-\beta}(\mathcal{T})$ *is given by* (27) *and* $\beta := (\beta_l)^M_{l=1}$ *with*

$$\beta_l := \max\big(0, 2 - \alpha_l\big). \tag{30}$$

Remark 3.1 *(Presence of ϵ).* Condition (29) is due to the application of Theorem 3.1 with $s = 3+\epsilon$. Taking $\epsilon = 0$ would exclude the important cases $\omega_l \in \{\pi/2, 3\pi/2, 2\pi\}$ for some $l \in \{1, \dots, M\}$, but these cases are included with $\epsilon > 0$ and so whenever $f \in H^{1+\epsilon}(\Omega)$.

Remark 3.2 *(Higher order nature of $\mathcal{E}_2$).* In general, $\mathcal{E}_2$ is expected to be of higher order 'away from the corners'. In what follows, we illustrate the higher order nature of $\mathcal{E}_2$ in the case of a convex domain Ω, that is $\omega_l < \pi$ for all $l = 1, \dots, M$, and global refinement. Although the estimate

$$\mathcal{E}_1(\mathcal{T}) \leq \|\nabla(u - u^1_{\mathcal{T}})\|_{0,\Omega} \|\nabla(z^2_{\mathcal{T}} - z^1_{\mathcal{T}})\|_{0,\Omega}$$

may be quite rough for concrete meshes, it is typically sharp in terms of asymptotic convergence speed. Consequently, the convergence order (on globally refined meshes) of $\mathcal{E}_1$ is the sum of the convergence orders of $u^1_{\mathcal{T}}$ and of $z^1_{\mathcal{T}}$ with respect to the energy norm error. In view of (28), the convergence order of $\mathcal{E}_2$ is the sum of $\min^M_{l=1}(2-\beta_l)$ and the convergence order of $z^2_{\mathcal{T}}$ with respect to the energy norm. Since $\min^M_{l=1}(2-\beta_l) > 1$ and the order of $z^2_{\mathcal{T}}$ is greater or equal to the one of $z^1_{\mathcal{T}}$, $\mathcal{E}_2$ is of higher order. Notice that the possible higher order of $z^2_{\mathcal{T}}$ hinges on the nature of G and may take place only on quite fine meshes. Therefore, the higher order contribution associated with $h^2_{\mathcal{T}} \max_T \sigma_{-\beta}$ in (28) is particularly advantageous in the case of regularized output functionals as in §2.

Proof of Theorem 3.2. In view of (8), (9), and (16), we have the error representation formula

$$G(u - u_{\mathcal{T}}^1) = \langle \mathcal{R}(u_{\mathcal{T}}^1), z_{\mathcal{T}}^2 - I_{\mathcal{T}}^1 z_{\mathcal{T}}^2 \rangle + \langle \mathcal{R}^*(z_{\mathcal{T}}^2), u - \Pi_{\mathcal{T}}^2 u \rangle.$$

That the first term is bounded by $\mathcal{E}_1(\mathcal{T})$ follows from (11). Therefore, it just remains to show that, for an appropriate choice of the interpolation operator $\Pi_{\mathcal{T}}^2$,

$$\langle \mathcal{R}^*(z_{\mathcal{T}}^2), u - \Pi_{\mathcal{T}}^2 u \rangle \leq \mathcal{E}_2(\mathcal{T}) = C \max_{l=1}^{M} L(h_l)^{3/2} \eta_{-1,-\beta}^*(\mathcal{T}) \|f\|_{1+\epsilon,\Omega}. \tag{31}$$

Let $\Pi_{\mathcal{T}}^2$ be the Scott-Zhang interpolation operator onto the space $\mathbb{V}_{\mathcal{T}}^2(\Omega)$ of continuous piecewise quadratic finite elements over $\mathcal{T}$; see [17]. Then the following local stability and error estimates are valid: for all $k \in \mathbb{N}_0, m \in \mathbb{N}$ with $0 \leq k \leq m \leq 3$, $v \in H^m(\Omega)$, and $T \in \mathcal{T}$,

$$\|D^k(v - \Pi_{\mathcal{T}}^2 v)\|_{k,T} \leq C h_T^{m-k} \|D^m v\|_{0,N(T)}. \tag{32}$$

In view of the non-oscillation property (21), this readily implies in particular the weighted stability ($k = m$) and error estimate ($k < m$)

$$\|\sigma_{-\beta} D^k(v - \Pi_{\mathcal{T}}^2 v)\|_{0,T} \leq C h_T^{m-k} \|\sigma_{-\beta} D^m v\|_{0,N(T)}. \tag{33}$$

Exploiting (29), we may invoke the decomposition in Theorem 3.1 with $s = 3 + \epsilon$ to write

$$\langle \mathcal{R}^*(z_{\mathcal{T}}^2), u - \Pi_{\mathcal{T}}^2 u \rangle = \langle \mathcal{R}^*(z_{\mathcal{T}}^2), w - \Pi_{\mathcal{T}}^2 w \rangle + \sum_{l=1}^{M} \sum_{j=1}^{J_l(s)} \Lambda_j^{(l)} \langle \mathcal{R}^*(z_{\mathcal{T}}^2), \psi_j^{(l)} - \Pi_{\mathcal{T}}^2 \psi_j^{(l)} \rangle. \tag{34}$$

We thus have to estimate the various evaluations of $\mathcal{R}^*(z_{\mathcal{T}}^2)$ on the right hand side.

Let us start with $\langle \mathcal{R}^*(z_{\mathcal{T}}^2), w - \Pi_{\mathcal{T}}^2 w \rangle$, which is associated to the regular part $w \in H^3(\Omega)$ of u. Standard techniques for the residual error estimator and (32) with $k = 0, 1$ and $m = 3$ yield the following 'weight-free' estimate:

$$|\langle \mathcal{R}^*(z_{\mathcal{T}}^2), w - \Pi_{\mathcal{T}}^2 w \rangle| \leq C \eta_{-1,0}^*(\mathcal{T}) \|D^3 w\|_{0,\Omega},$$

where C depends on the minimum angle of $\mathcal{T}$. In view of $\beta_l \geq 0$ for all $l = 1, \ldots, M$, this implies

$$|\langle \mathcal{R}^*(z_{\mathcal{T}}^2), w - \Pi_{\mathcal{T}}^2 w \rangle| \leq C \eta_{-1,-\beta}^*(\mathcal{T}) \|D^3 w\|_{0,\Omega} \tag{35}$$

with C depending also on the domain Ω.

For the remaining terms in (34) associated to the singular part of u, we modify the preceding argument for $|\langle \mathcal{R}^*(z_{\mathcal{T}}^2), w - \Pi_{\mathcal{T}}^2 w \rangle|$ as follows. Fix $l \in \{1, \ldots, M\}$ and $j \in \{1, \ldots, J_l(s)\}$, and write

$$\begin{aligned} \langle \mathcal{R}^*(z_{\mathcal{T}}^2), \psi_j^{(l)} - \Pi_{\mathcal{T}}^2 \psi_j^{(l)} \rangle = &\sum_{T \in \mathcal{T} \setminus \mathcal{T}_l} \Big[\langle R^*, \psi_j^{(l)} - \Pi_{\mathcal{T}}^2 \psi_j^{(l)} \rangle_T - \langle J^*, \psi_j^{(l)} - \Pi_{\mathcal{T}}^2 \psi_j^{(l)} \rangle_{\partial T} \Big] \\ &+ \sum_{T \in \mathcal{T}_l} \Big[\langle R^*, \psi_j^{(l)} - \Pi_{\mathcal{T}}^2 \psi_j^{(l)} \rangle_T - \langle J^*, \psi_j^{(l)} - \Pi_{\mathcal{T}}^2 \psi_j^{(l)} \rangle_{\partial T} \Big]. \end{aligned} \tag{36}$$

with $\mathcal{T}_l := \{T \in \mathcal{T} \mid N(T) \ni P_l\}$ denoting those triangles of $\mathcal{T}$ whose neighborhoods $N(T)$ touch the corner P_l.

We first estimate that part of the first sum in (36) associated to the dual element residual R^*. To this end, we observe that, independently of j, the following weighted estimate holds:

$$\|\sigma_\beta D^3 \psi_j^{(l)}\|_{0,\Omega_l} \leq CL(h_l)^{3/2}, \tag{37}$$

where $\Omega_l := \cup_{T\in\mathcal{T}\setminus\mathcal{T}_l} N(T)$ and $h_l \leq C\text{dist}(P_l, \Omega \setminus \Omega_l)$. In fact, (37) follows from direct integration of $|\sigma_\beta D^3 \psi_j^{(l)}|^2$ in $\{r_l \geq h_l\}$; note that $\psi_j^{(l)}$ contains a $(\log r_l)$-term whenever $j\alpha_l \leq 2$ is an integer. Utilizing (33) with $k = 0$ and $m = 3$ and (37), we derive

$$\begin{aligned}\sum_{T\in\mathcal{T}\setminus\mathcal{T}_l} |\langle R^*, \psi_j^{(l)} - \Pi_T^2 \psi_j^{(l)}\rangle_T| &\leq \sum_{T\in\mathcal{T}\setminus\mathcal{T}_l} \|\sigma_{-\beta} R^*\|_{0,T} \|\sigma_\beta(\psi_j^{(l)} - \Pi_T^2 \psi_j^{(l)})\|_{0,T} \\ &\leq C \sum_{T\in\mathcal{T}\setminus\mathcal{T}_l} \|\sigma_{-\beta} R^*\|_{0,T} h_T^3 \|\sigma_\beta D^3 \psi_j^{(l)}\|_{0,N(T)} \\ &\leq CL(h_l)^{3/2} \Big[\sum_{T\in\mathcal{T}\setminus\mathcal{T}_l} h_T^6 \|\sigma_{-\beta} R^*\|_{0,T}^2\Big]^{1/2}.\end{aligned} \tag{38}$$

Next, we estimate that part of the second sum in (36) associated to the dual element residual R^*. Here we shall use the following property of the singular functions $\psi_j^{(l)}$, $j = 1, \ldots, J_l(s)$:

$$\int_{N_l} \sigma_\beta^2 |\nabla \psi_j^{(l)}|^2 \leq C\bar{h}_l^{\max\{0,4-2\alpha_l\}} \int_0^{\bar{h}_l} r^{2\alpha_l - 1} L(r)^2\, dr \leq C\bar{h}_l^4 L(\bar{h}_l)^2, \tag{39}$$

where $N_l := \cup_{T\in\mathcal{T}_l} N(T)$ and $\bar{h}_l$ is the smallest radius such that $N_l \subset B(P_l; \bar{h}_l)$. Exploiting (33) with $k = 0$ and $m = 1$, (39), as well as $\bar{h}_l \leq Ch_T$ for all $T \in \mathcal{T}_l$, and $h_l \leq \bar{h}_l$, we obtain

$$\begin{aligned}\sum_{T\in\mathcal{T}_l} |\langle R^*, \psi_j^{(l)} - \Pi_T^2 \psi_j^{(l)}\rangle_T| &\leq C \sum_{T\in\mathcal{T}_l} h_T \|\sigma_{-\beta} R^*\|_{0,T} \|\sigma_\beta \nabla \psi_j^{(l)}\|_{0,N(T)} \\ &\leq C \|\sigma_\beta \nabla \psi_j^{(l)}\|_{N_l} \Big[\sum_{T\in\mathcal{T}_l} h_T^2 \|\sigma_{-\beta} R^*\|_{0,T}^2\Big]^{1/2} \\ &\leq CL(h_l) \Big[\sum_{T\in\mathcal{T}_l} h_T^6 \|\sigma_{-\beta} R^*\|_{0,T}^2\Big]^{1/2}.\end{aligned} \tag{40}$$

We now turn to those parts of the two sums in (36) associated with the jump residual J^*. In view of the scaled and weighted trace inequality

$$\|\sigma_\beta v\|_{0,\partial T}^2 \leq Ch_T^{-1} \|\sigma_\beta v\|_{0,T}^2 + Ch_T \|\sigma_\beta \nabla v\|_{0,T}^2 \qquad \forall v \in H^1(T), \tag{41}$$

one can proceed similarly to the parts for R^*. In fact, after application of (41) with $v = \psi_j^{(l)} - \Pi_T^2 \psi_j^{(l)}$ we can treat the sums with the first term on the right hand side of

(41) as before, while for the other sums we employ (33) with $k = 1$ (instead of $k = 0$) and $m = 1, 3$. We thus obtain

$$\sum_{T \in \mathcal{T}} |\langle J^*, \psi_j^{(l)} - \Pi_{\mathcal{T}}^2 \psi_j^{(l)} \rangle_{\partial T}| \leq C L(h_l)^{3/2} \Big[\sum_{T \in \mathcal{T}} h_T^5 \|\sigma_{-\beta} J^*\|_{0,\partial T}^2 \Big]^{1/2}. \tag{42}$$

Upon inserting (38), (40), and (42) into (36), we readily obtain

$$|\langle \mathcal{R}^*(z_{\mathcal{T}}^2), \psi_j^{(l)} - \Pi_{\mathcal{T}}^2 \psi_j^{(l)} \rangle| \leq C L(h_l)^{3/2} \eta_{-1,-\beta}^*(\mathcal{T}).$$

Finally, replacing this and (35) into (34), and using the stability estimate of Theorem 3.1, we arrive at (31) and thus complete the proof. □

4 Back to the Example: Reliable Estimator

We now reexamine the example of §2 and illustrate that the safeguarded DWR estimate is not only reliable but also asymptotically coincides with the original DWR estimate.

The adaptive algorithm based upon the safeguarded DWR estimate of Theorem 3.2 stops when $\mathcal{E}_1(\mathcal{T}) + \mathcal{E}_2(\mathcal{T}) \leq \texttt{tol}$ is satisfied. Notice that the evaluation of $\mathcal{E}_2(\mathcal{T})$ requires a choice for the constant C. Here we use the ad hoc choice $C\|f\|_{1+\epsilon} \max_l L(h_l)^{3/2} = 1$, which means that we ignore the presence of the logarithmic term and suppose that C is of moderate size. If the stopping test is not satisfied, elements are marked for refinement with the help of the equidistribution strategy, the indicators of which are given by

$$\big|\langle R, z_{\mathcal{T}}^2 - I_{\mathcal{T}}^1 z_{\mathcal{T}}^2 \rangle_T - \langle J, z_{\mathcal{T}}^2 - I_{\mathcal{T}}^1 z_{\mathcal{T}}^2 \rangle_{\partial T}\big| + \eta_{-1,-\beta}^*(T), \qquad T \in \mathcal{T}. \tag{43}$$

We apply this algorithm to the example of §2. As there, we require $\texttt{tol} = 0.1$ and start from the same mesh $\mathcal{T}_0$ in Figure 1 with 12 triangles. Here it turns out that

$$\mathcal{E}_1(\mathcal{T}_0) = 0.070921 \quad \text{and} \quad \mathcal{E}_2(\mathcal{T}_0) = 93.490629.$$

Consequently, the algorithm does not stop and marks elements for refinement. After 8 adaptive loops, it finishes with

$$\mathcal{E}_1(\mathcal{T}_8) = 0.073332, \quad \mathcal{E}_2(\mathcal{T}_8) = 0.000923, \quad \text{and} \quad |G(u) - G(u_{\mathcal{T}}^1)| = 8.75382298e - 03.$$

Figure 2 (left) shows the corresponding convergence history of the true error $|G(u) - G(u_{\mathcal{T}}^1)|$, the safeguarded DWR estimate of Theorem 3.2, and the original DWR estimate (11) and Figure 2 (right) depicts the final mesh $\mathcal{T}_8$ with 7561 degrees of freedom (DOFs).

These numerical results

- show that the original DWR estimate (11) is unreliable not only on $\mathcal{T}_0$ but also on $\mathcal{T}_1$ and $\mathcal{T}_2$;
- corroborate the upper bound with the safeguarded DWR estimate in Theorem 3.2;

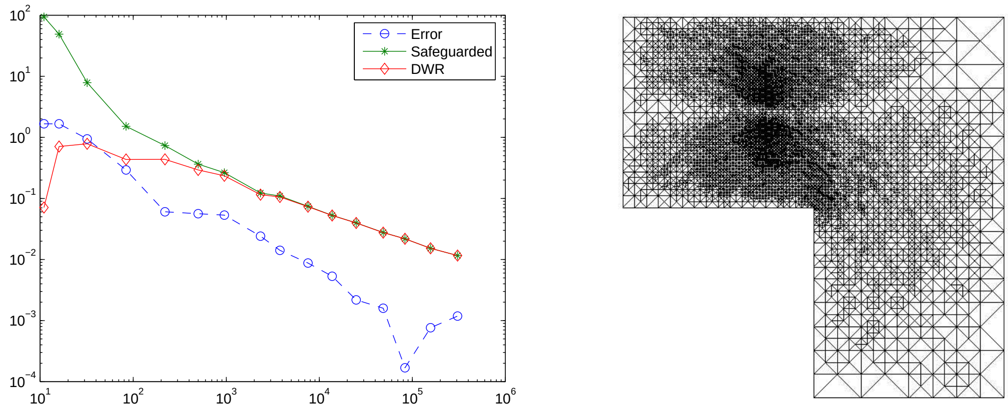

Figure 2: Using the safeguarded DWR: True error, safeguarded and original DWR estimate versus number of degrees of freedom in log-log scale (left). The original DWR may be unreliable but the safeguarded one is always reliable; the two estimates asymptotically coincide. Final mesh with 7561 degrees of freedom (DOFs) (right).

- indicate that the constant C is of moderate size for 'reasonable' domains and meshes;
- complement Remark 3.2 by demonstrating that the safeguarded DWR estimate asymptotically coincides with the original one also on graded meshes.

We may summarize our findings by saying that the safeguarded DWR estimate of Theorem 3.2 gains unconditional reliablity without loosing the advantages of the original DWR estimate.

References

[1] W. Bangerth and R. Rannacher. *Adaptive finite element methods for differential equations*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2003.

[2] 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(4):237–264, 1996.

[3] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numer.*, 10:1–102, 2001.

[4] C. Carstensen. Estimation of higher Sobolev norm from lower order approximation. *SIAM J. Numer. Anal.*, 42(5):2136–2147 (electronic), 2005.

[5] M. Dauge. *Elliptic boundary value problems on corner domains*, volume 1341 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1988. Smoothness and asymptotics of solutions.

[6] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson. Introduction to adaptive methods for differential equations, *Acta Numer.*, 4: 105–158, 1995 (A. Iserles ed., Cambridge University Press).

[7] K. Eriksson and C. Johnson. An adaptive finite element method for linear elliptic problems. *Math. Comp.*, 50(182):361–383, 1988.

[8] K. Eriksson and C. Johnson. Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.*, 28(1):43–77, 1991.

[9] C. Johnson. Adaptive finite element methods for diffusion and convection problems. *Comput. Methods Appl. Mech. Engrg.*, 82(1-3):301–322, 1990.

[10] B. Kellogg. Notes on piecewise smooth elliptic boundary value problems. Technical Note BN-1137, Institute for Physical Science and Technology, University of Maryland, College Park, 1992.

[11] X. Liao and R. H. Nochetto. Local a posteriori error estimates and adaptive control of pollution effects. *Numer. Methods Partial Differential Equations*, 19(4):421–442, 2003.

[12] Y. Maday and A. T. Patera. Numerical analysis of a posteriori finite element bounds for linear functional outputs. *Math. Models Methods Appl. Sci.*, 10(5):785–799, 2000.

[13] M. Paraschivoiu and A. T. Patera. A hierarchical duality approach to bounds for the outputs of partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 158(3-4):389–407, 1998.

[14] M. Paraschivoiu, J. Peraire, and A.T. Patera. A posteriori finite element bounds for linear-functional outputs of elliptic partial differential equations. *Comput. Methods Appl. Mech. Engrg.*, 150(1-4):289–312, 1997. Symposium on Advances in Computational Mechanics, Vol. 2 (Austin, TX, 1997).

[15] S. Prudhomme and J. T. Oden. On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. *Comput. Methods Appl. Mech. Engrg.*, 176(1-4):313–331, 1999. New advances in computational methods (Cachan, 1997).

[16] A. Schmidt and K. G. Siebert. *Design of adaptive finite element software*, volume 42 of *Lecture Notes in Computational Science and Engineering*. Springer-Verlag, Berlin, 2005. The finite element toolbox ALBERTA, With 1 CD-ROM (Unix/Linux).

[17] L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54(190):483–493, 1990.

[18] E. Süli. Review of the book *Adaptive finite element methods for differential equations* by W. Bangerth and R. Rannacher. *Math.Comp.*, 74(250):1033–1039, 2004.