

Effective postprocessing for equilibration a posteriori error estimators

C. Carstensen · C. Merdon

Received: 22 June 2011 / Revised: 2 May 2012 / Published online: 13 September 2012
© Springer-Verlag 2012

Abstract Guaranteed error control via fully discrete a posteriori error estimators is possible with typical overestimation between 1.25 and 2 in simple computer benchmarks. The equilibration techniques due to Braess and that due to Luce–Wohlmuth are efficient tools with an accuracy limited by the hyper-circle threshold. This motivates postprocessing strategies and the analysis of successive improvements of guaranteed upper error bounds with a few pcg iterations result in reduced overestimation factors between 1 and 1.25. Numerical simulations for three classes of applications illustrate the efficiency for the Poisson model problem with and without jumping coefficients or a simple obstacle problem.

Mathematics Subject Classification 65N30 · 65R20 · 65N15

1 Introduction

The a posteriori error control of the energy norms of errors in computational PDEs has attracted high attention over the last decades [1, 6, 7, 9, 17, 25, 28]. The particular aspect

This work was supported by DFG Research Center MATHEON and by the World Class University (WCU) program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology R31-2008-000-10049-0. C. Merdon was also supported by the German Academic Exchange Service (DAAD, D/10/44641) during his stay at the Yonsei University in 2010.

C. Carstensen (✉) · C. Merdon
Humboldt-Universität zu Berlin, Unter den Linden 6, 10099 Berlin, Germany
e-mail: cc@math.hu-berlin.de

C. Merdon
e-mail: merdon@math.hu-berlin.de

C. Carstensen · C. Merdon
Department of Computational Science and Engineering,
Yonsei University, Seoul 120-749, Korea

of *guaranteed* upper bounds with all explicit constants has risen particular interest and is also relevant for effective simulations even in terms of goal functionals [1,7,21].

This paper addresses the most accurate energy norm error estimators which can be written as equilibrium error estimators [4,12]. At least since the unified approach [10,11] it is clear and visible more and more [29] that the essential task in a posteriori error control is the computation of upper bounds for some residual $\text{Res} \in H^{-1}(\Omega)$, the dual of the standard first-order Sobolev space $H_0^1(\Omega)$ with homogeneous boundary values, of the form

$$\text{Res}(\varphi) = \int_{\Omega} (f\varphi - g \cdot \nabla\varphi) \, dx \quad \text{for all } \varphi \in H_0^1(\Omega). \tag{1.1}$$

The given 2D data are the Lebesgue integrable functions $f \in L^2(\Omega)$ and the numerical flux $g \in L^2(\Omega; \mathbb{R}^2)$ (piecewise constant in practical examples). The form (1.1) results from a weak formulation of some equilibration equation

$$f + \text{div } \sigma = 0 \tag{1.2}$$

with exact flux $\sigma \in H(\text{div}, \Omega)$ and its residual

$$\text{Res} := f + \text{div } g = \text{div}(g - \sigma) \in H^{-1}(\Omega).$$

Here and throughout this paper, we use standard notation for Lebesgue and Sobolev spaces and their norms; $V := H_0^1(\Omega)$ is endowed with the energy norm $\|\cdot\| := \|\nabla \cdot\|_{L^2(\Omega)} = |\cdot|_{H^1(\Omega)}$ and the dual norm $\|\cdot\|_{\star}$ in $H^{-1}(\Omega)$. The relevant dual norm

$$\|\text{Res}\|_{\star} = \|\text{div}(\sigma - g)\|_{\star} \leq \eta, \tag{1.3}$$

is the targeted quantity and the aim is to find some computable upper bound η . This is the essential step in the guaranteed error control for many problems like the benchmarks for the two Poisson model problem, two interface problems and one obstacle problem of this paper. The standard modification of the energy norm $\|\cdot\| := \|\chi^{1/2} \nabla \cdot\|_{L^2(\Omega)}$ is understood for the interface problems with jumping coefficients. Although all experiments employ conforming first-order finite element methods, the theory also applies to nonconforming methods [14].

The class of equilibration techniques takes the input data f and g and computes some $q \in H(\text{div}, \Omega)$ such that the triangle inequality implies

$$\|\text{div}(\sigma - g)\|_{\star} = \|f + \text{div } g\|_{\star} \leq \|f + \text{div } q\|_{\star} + \|\text{div}(q - g)\|_{\star}.$$

This leads to the explicit bound η in (1.3) equal to

$$\eta := \|f + \text{div } q\|_{\star} + \|\text{div}(q - g)\|_{\star}.$$

In case that certain piecewise integrals of $f + \text{div } q$ vanish, e.g. for the first three examples of Table 1, one may further deduce some local oscillation term $\text{osc}(f)$ and

Table 1 Equilibration a posteriori error estimators

No.	Error estimator	Equilibration	Mesh $\widehat{\mathcal{T}}$	References
1	η_B	$\text{div } q_B = -f_{\mathcal{T}}$	$\mathcal{T}, \text{red}(\mathcal{T})$	[6,8]
2	η_{MFEM}	$\text{div } q_{\text{MFEM}} = -f_{\mathcal{T}}$	$\text{red}(\mathcal{T})$	[6,8]
3	η_{LW}	$\text{div } q_{\text{LW}} = -f^*$	\mathcal{T}^*	[19]
4	η_{LS}	None	$\text{red}(\mathcal{T})$	[12]
5	η_{Repin}	None	$\text{red}(\mathcal{T})$	[25,26]

an explicit constant C with

$$\|f + \text{div } q\|_{\star} \leq C \text{osc}(f).$$

For instance, the Braess a posteriori error estimator from Table 1 designs some $q_B \in \text{RT}_0(\mathcal{T})$ and leads to elementwise oscillations $\text{osc}(f) = \text{osc}(f, \mathcal{T}) := \|h_{\mathcal{T}}(f - f_{\mathcal{T}})\|_{L^2(\Omega)}$ and $C = 1/\pi$. Here, $f_{\mathcal{T}}$ denotes the piecewise integral mean of f and $h_{\mathcal{T}}$ denotes the local mesh-size. This results in the computable guaranteed upper bound

$$\eta_B := \text{osc}(f, \mathcal{T})/\pi + \|q_B - g\|_{L^2(\Omega)} \tag{1.4}$$

for the error estimate (1.3). In many benchmark examples, the efficiency of this estimator

$$\text{eff}(\eta_B) := \eta_B / \|\text{Res}\|_{\star}$$

lies in the range of 1.3–2. Section 2 discusses the hyper-circle identity and the related threshold of the efficiency indices for all the error estimators of Table 1. To overcome this hyper-circle threshold further improvements of q say on refined meshes or of higher polynomial degrees are necessary to minimise the upper bound $\|q - g\|_{L^2(\Omega)}$ under the side restriction that $f + \text{div } q$ maintains the above integral mean properties. This paper makes the explicit alternative ansatz that, given any q from Table 1 as well as for any possible future suggestion of this kind, q is substituted by $q - \text{Curl } v$ for some $v \in H^1(\Omega)/\mathbb{R}$. Since $\text{Curl } v = (\partial v/\partial x_2, -\partial v/\partial x_1)$ is divergence-free in $n = 2$ dimensions, the volume term $f + \text{div } q$ remains unchanged, while the total new upper bound

$$\mu_{\text{new}} := \text{osc}(f, \mathcal{T})/\pi + \min_{v \in H^1(\Omega)/\mathbb{R}} \|q - g - \text{Curl } v\|_{L^2(\Omega)} \tag{1.5}$$

may be much smaller than η from (1.4). The theoretical main results of this paper state that this improvement is significant and even asymptotic exactness of μ_{new} is possible in the Poisson model problem.

The estimate (1.5) coincides with a particular form of the estimate [25, (3.5.20)] with y replaced by $q - \text{Curl } v$. The numerical realisation of (1.5), however, requires a further discretisation of $H^1(\Omega)$ based on the same or even on a refined mesh $\widehat{\mathcal{T}}$

from Table 1 and some global minimisation. The striking empirical result of this paper is that a significant improvement of the overall efficiency can be obtained with a few pcg iterations and even one iteration (i.e. one line search along the gradient) in the discretised minimisation leads to amazing results. For the three model classes, the efficiency is cheaply improved without any change of the subtle design of the equilibration function q . Moreover, the improvement by successive red-refinement of \widehat{T} in Theorem 4.1 follows from arguments from the convergence of adaptive mixed finite element methods.

The remaining parts of the paper are organised as follows. Section 2 explains the hyper-circle threshold and how it limits the efficiency of the equilibration error estimators of Table 1. This motivates the derivation of an improved error control in Sect. 3. Section 4 proves a saturation property for the postprocessing based on red-refined meshes. Section 5 outlines the a posteriori error estimators from Table 1 and the design of the flux q . Section 6 describes the numerical realisation while Sects. 7, 8, and 9 give numerical evidence for the improved efficiency of the postprocessed error estimators in three different model problems. Section 10 draws some conclusions to round up the paper.

Although the examples are all in $n = 2$ dimensions for simplicity, the ansatz is feasible in any dimension as long as the problem is in divergence form and $\text{Curl } v$ is replaced by any divergence-free field. Finally $a \lesssim b$ abbreviates $a \leq Cb$ for some generic constant C that depends only on the shape regularity of the triangulation, while $a \approx b$ means $a \lesssim b \lesssim a$.

2 Hyper-circle threshold

Consider a regular triangulation \mathcal{T} of the simply connected and bounded Lipschitz domain $\Omega \subset \mathbb{R}^2$ into triangles with edges \mathcal{E} , nodes \mathcal{N} and free nodes \mathcal{M} . The midpoints of all edges are denoted by $\text{mid}(\mathcal{E}) := \{\text{mid}(E) \mid E \in \mathcal{E}\}$ and the boundary edges along $\partial\Omega$ are denoted by $\mathcal{E}(\partial\Omega) := \{E \in \mathcal{E} \mid E \subseteq \partial\Omega\}$ while $\mathcal{E}(\Omega) := \mathcal{E} \setminus \mathcal{E}(\partial\Omega)$ denotes the set of interior edges. The diameter $\text{diam}(T)$ of a triangle T is denoted by h_T , $\mathcal{E}(T)$ consists of all three edges of the triangle $T \in \mathcal{T}$ and $\mathcal{N}(T)$ consists of all of its vertices. The open set $\omega_z := \{\varphi_z > 0\}$ for some node function φ_z is the interior of its support on the subtriangulation $\mathcal{T}(z) := \{T \in \mathcal{T} \mid z \in \mathcal{N}(T)\}$. Similarly, all edges that share $z \in \mathcal{N}$ give rise to $\mathcal{E}(z) := \{E \in \mathcal{E} \mid z \in E\}$. The red-refinement $\text{red}(\mathcal{T})$ of \mathcal{T} is a regular triangulation that refines each triangle $T \in \mathcal{T}$ into four congruent sub-triangles by straight lines through the midpoints of the three edges. With the set $P_k(\mathcal{T})$ of elementwise polynomials of total degree $\leq k$, the lowest-order Raviart–Thomas finite element space is given by

$$\text{RT}_0(\mathcal{T}) := \{q \in H(\text{div}, \Omega) \mid \forall T \in \mathcal{T} \exists a_T, b_T, c_T \in P_0(T) \forall x \in T, \\ q(x) = a_T x + (b_T, c_T)\}.$$

The standard reference [6] for the FEM advertises the hyper-circle principle or Prager–Synge estimate from [22] for a posteriori error control and gives details for an easy postprocessing to compute q_B in the lowest-order Raviart–Thomas mixed finite

element space $RT_0(\mathcal{T})$ (cf. Sect. 5 below for details on the design of q_B). Lemma 9.1 of [6] compares this with the lowest-order Raviart–Thomas mixed FEM and its solution q_{MFEM} in $RT_0(\mathcal{T}) \subseteq H(\text{div}, \Omega)$ in a Poisson model problem scenario $f + \Delta u = 0$ with exact flux $\sigma := \nabla u$ and residual (1.1) with data $f \in P_0(\mathcal{T})$ and $g := \nabla u_h \in P_0(\mathcal{T}; \mathbb{R}^2)$ for the conforming first-order approximation $u_h \in P_1(\mathcal{T}) \cap C(\Omega)$. It follows

$$\|q_{MFEM} - g\|_{L^2(\Omega)} \leq \|q_B - g\|_{L^2(\Omega)} = \eta_B.$$

The hyper-circle principle leads for piecewise constant f (and otherwise up to oscillation terms for f which are neglected in Sect. 2 for the ease of this discussion) to the identity

$$\|q_{MFEM} - g\|_{L^2(\Omega)}^2 = \|\sigma - g\|_{L^2(\Omega)}^2 + \|\sigma - q_{MFEM}\|_{L^2(\Omega)}^2. \tag{2.1}$$

This identity is obtained by an integration by parts to show $\int_{\Omega} \nabla(u - u_h)(\sigma - q) \, dx = 0$ for any $q \in H(\text{div}, \Omega)$ with $\text{div } q + f = 0$. One immediate consequence of the aforementioned identities reads

$$\|\sigma - g\|_{L^2(\Omega)}^2 + \|\sigma - q_{MFEM}\|_{L^2(\Omega)}^2 \leq \|q_B - g\|_{L^2(\Omega)}^2 \equiv \eta_B^2.$$

Hence the efficiency index

$$\text{eff}(\eta_B) := \eta_B / \|\sigma - g\|_{L^2(\Omega)}$$

of the Braess a posteriori error estimator η_B is bounded from below,

$$\begin{aligned} \sqrt{1 + \kappa^2} &:= \left(1 + \|\sigma - q_{MFEM}\|_{L^2(\Omega)}^2 / \|\sigma - g\|_{L^2(\Omega)}^2\right)^{1/2} \\ &\leq \|q_B - g\|_{L^2(\Omega)} / \|\sigma - g\|_{L^2(\Omega)} = \text{eff}(\eta_B). \end{aligned}$$

A compactness argument in [15] proves for piecewise constant right-hand sides that

$$\|\sigma - g\| \leq C(\mathcal{T}) \|q_{MFEM} - g\|_{L^2(\Omega)}.$$

The constant $C(\mathcal{T})$ depends on the triangulation but not on the data nor on the exact or discrete solution. Hence,

$$\kappa := \|\sigma - q_{MFEM}\|_{L^2(\Omega)} / \|\sigma - g\|_{L^2(\Omega)} \geq 1/C(\mathcal{T}).$$

In other words, there is no reason to believe that the efficiency index is close to 1. Comparing numerical experiments in [12] and below in this paper reveal for the equilibration error estimates of Table 1, that $\text{eff}(\eta_B)$ lies in the range of 1.3–1.7. This fundamental lower bound $\sqrt{1 + \kappa^2}$ is therefore called the *hyper-circle threshold* and limits the efficiency.

It is the purpose of this paper to improve the efficiency dramatically below that hyper-circle threshold of $\sqrt{1 + \kappa^2}$ and even allow asymptotic exactness of guaranteed upper bounds. To overcome this threshold for even higher accuracy, extra calculations are required such as higher accuracy of q from higher polynomial degrees or refined meshes. The alternative key observation is that all known equilibration a posteriori error estimators use the simple but coarse estimate

$$\| \operatorname{div}(q - g) \|_{\star} = \sup_{\substack{\varphi \in H_0^1(\Omega) \\ \|\varphi\|=1}} \int_{\Omega} (q - g) \cdot \nabla \varphi \, dx \leq \|q - g\|_{L^2(\Omega)}$$

based on the Cauchy inequality. Instead, a Helmholtz decomposition leads to the identity

$$\| \operatorname{div}(q - g) \|_{\star} = \min_{v \in H^1(\Omega)/\mathbb{R}} \|q - g - \operatorname{Curl} v\|_{L^2(\Omega)}.$$

Here, v is some test function in $H^1(\Omega)$ and, more importantly, any choice of v leads to some guaranteed upper bound,

$$\| \operatorname{div}(q - g) \|_{\star} \leq \|q - g - \operatorname{Curl} v\|_{L^2(\Omega)}.$$

The optimal v in the improved previous estimate allows a posteriori error control beyond the hyper-circle threshold.

Theorem 2.1 *Under the assumptions of this section for the Poisson model problem with piecewise constant right-hand side $f \equiv f_{\mathcal{T}} \in P_0(\mathcal{T})$, it holds, for all $q \in Q(f_{\mathcal{T}}) := \{q \in H(\operatorname{div}, \Omega) \mid f_{\mathcal{T}} + \operatorname{div} q = 0\}$,*

$$\| \operatorname{Res} \|_{\star} = \| \sigma - g \|_{L^2(\Omega)} = \mu := \min_{v \in H^1(\Omega)/\mathbb{R}} \|q - g - \operatorname{Curl} v\|_{L^2(\Omega)}.$$

Proof This follows with $\| f + \operatorname{div} q \|_{\star} = 0$ in Theorem 3.1; cf. Remark 3.3. □

The global minimisation for the computation of μ leads to an elliptic PDE and appears as costly as the computation of μ . In order to approximate μ , any choice of $v \in H^1(\Omega)$ yields an upper bound and so an improved guaranteed a posteriori error control.

The following Theorem 2.2 underlines the significance of the MFEM equilibration: Without mesh-refinement there is no improvement beyond

$$\eta_{\text{MFEM}} := \|q_{\text{MFEM}} - g\|_{L^2(\Omega)}.$$

Theorem 2.2 *Under the assumptions of this section for the Poisson model problem with piecewise constant right-hand side $f \in P_0(\mathcal{T})$ and $V_H := P_1(\mathcal{T}) \cap C(\Omega)$, it holds*

$$\min_{v_H \in V_H/\mathbb{R}} \|q_B - g - \operatorname{Curl} v_H\|_{L^2(\Omega)} = \|q_{\text{MFEM}} - g\|_{L^2(\Omega)}.$$

Proof This follows from the fact that q_{MFEM} minimises the distance to g amongst any $q_{\text{RT}} \in Q(f, \mathcal{T}) := \{q \in \text{RT}_0(\mathcal{T}) \mid f_{\mathcal{T}} + \text{div } q = 0\}$ [6]. Since $q_{\text{B}} \in Q(f, \mathcal{T})$ and $\text{Curl } v_H \in P_0(\mathcal{T}; \mathbb{R}^2)$, the difference $q_{\text{B}} - \text{Curl } v_H$ belongs to $Q(f, \mathcal{T})$. \square

Numerical experiments in Sect. 7 below confirm that the improvement of the efficiency of η_{B} compared to η_{MFEM} is not significant. This is an indication for the amazing accuracy of the Braess postprocessing which defines q_{B} .

Theorem 2.3 below implies that more significant improvements follow from further mesh-refinements. In case of a red-refined triangulation $\mathcal{T}_h := \text{red}(\mathcal{T})$, the post-processed a posteriori error estimator for $V_h = P_1(\mathcal{T}_h) \cap C(\Omega)$ reads

$$\eta_{\text{Br}} := \min_{v_h \in V_h/\mathbb{R}} \|q_{\text{B}} - g - \text{Curl } v_h\|_{L^2(\Omega)}.$$

The data oscillations on edge patches $\omega_E := \bigcup\{T \in \mathcal{T} \mid E \in \mathcal{E}(T)\}$ for an edge $E \in \mathcal{E}$ read

$$\text{osc}(f, \omega_E) := \text{diam}(\omega_E) \|f - f_{\omega_E}\|_{L^2(\omega_E)}, \tag{2.2}$$

$$\text{OSC}^2(\mathcal{T}, f) := \sum_{E \in \mathcal{E}(\Omega)} \text{osc}^2(f, \omega_E) + \sum_{E \in \mathcal{E}(\partial\Omega)} |E|^2 \|f\|_{L^2(\omega_E)}^2. \tag{2.3}$$

Small data oscillations lead to a significant improvement of the postprocessed a posteriori error estimator and so may overcome the hyper-circle threshold.

Theorem 2.3 *Under the assumptions of this section for the Poisson model problem with piecewise constant right-hand side $f \in P_0(\mathcal{T})$, there exist constants $0 < \varrho < 1$ and $0 < \Lambda < \infty$ which depend on the interior angles of \mathcal{T} and neither on the mesh-sizes nor on the number of triangles with*

$$\eta_{\text{Br}}^2 - \|\sigma - g\|_{L^2(\Omega)}^2 \leq \varrho \left(\eta_{\text{MFEM}}^2 - \|\sigma - g\|_{L^2(\Omega)}^2 \right) + \Lambda \text{OSC}^2(\mathcal{T}, f).$$

Proof This follows from Theorem 4.1 below for $r_{\mathcal{T}} = g, \kappa \equiv 1, f_{\text{eq}} = f$ and $q_{\text{eq}} = q_{\text{B}}$. \square

So far, this section discussed the Braess equilibration technique. The remaining a posteriori error estimators of Table 1 also suffer from the hyper-circle threshold. The Luce–Wohlmuth a posteriori error estimator, for instance, leads to an estimate which is bigger than the mixed error estimator η_{MFEM^*} with respect to the dual mesh \mathcal{T}^* . (Recall that f is piecewise constant in Sect. 2 so that the right-hand side of the MFEM and f^* in the Luce–Wohlmuth error estimation of Sect. 5.3 coincide.) The significant improvements of Theorem 2.3 immediately apply to \mathcal{T}^* and $\text{red}(\mathcal{T}^*)$ as well.

Furthermore, the reduction properties can be iterated. For instance, the minimisation in $v_h \in P_1(\text{red}^k(\mathcal{T})) \cap C(\Omega)$ leads for $k = 1, 2, \dots$ to

$$\begin{aligned} \eta_{\underbrace{\text{Br} \dots \Gamma}_k}^2 - \|\sigma - g\|_{L^2(\Omega)}^2 &\leq \varrho \left(\eta_{\underbrace{\text{Br} \dots \Gamma}_{k-1}}^2 - \|\sigma - g\|_{L^2(\Omega)}^2 \right) + \Lambda \text{OSC}(\text{red}^{k-1}(\mathcal{T}))^2 \\ &\leq \varrho^k \left(\eta_{\text{MFEM}}^2 - \|\sigma - g\|_{L^2(\Omega)}^2 \right) \end{aligned}$$

$$+ \sum_{j=0}^{k-1} \varrho^j \Lambda \text{OSC}(\text{red}^{k-j-1}(\mathcal{T}))^2.$$

This section concludes with a discussion of the associated efficiency indices

$$\text{eff}(\eta_{\text{Br}}) := \eta_{\text{Br}} / \|\sigma - g\|_{L^2(\Omega)} \quad \text{resp.} \quad \text{eff}(\eta_{\text{MFEM}}) := \eta_{\text{MFEM}} / \|\sigma - g\|_{L^2(\Omega)}.$$

Their distances to 1 are reduced by

$$\frac{\text{eff}(\eta_{\text{Br}})^2 - 1}{\text{eff}(\eta_{\text{MFEM}})^2 - 1} \leq \frac{\varrho(\eta_{\text{MFEM}}^2 - \|\sigma - g\|_{L^2(\Omega)}^2) + \Lambda \text{OSC}^2(\mathcal{T}, f)}{\eta_{\text{MFEM}}^2 - \|\sigma - g\|_{L^2(\Omega)}^2}.$$

Under the assumption of small oscillations in the sense of

$$\varrho \left(\eta_{\text{MFEM}}^2 - \|\sigma - g\|_{L^2(\Omega)}^2 \right) + \Lambda \text{OSC}^2(\mathcal{T}, f) \leq \vartheta \left(\eta_{\text{MFEM}}^2 - \|\sigma - g\|_{L^2(\Omega)}^2 \right)$$

for some $0 < \vartheta < 1$, this implies

$$\frac{\text{eff}(\eta_{\text{Br}})^2 - 1}{\text{eff}(\eta_{\text{B}})^2 - 1} \leq \frac{\text{eff}(\eta_{\text{Br}})^2 - 1}{\text{eff}(\eta_{\text{MFEM}})^2 - 1} \leq \vartheta < 1.$$

In other words, the efficiency indices are significantly reduced.

The striking numerical evidence of Sects. 7–9 in this paper suggests that only a few iterations of some iterative solver improve the efficiency substantially. Tables 2 and 4 in Sect. 7 display typical values for ϑ that are clearly far below 1.

3 Refined error control

This section is devoted to a rigorous analysis and the asymptotic exactness of the suggested error estimator (1.5) based on the data $f \in L^2(\Omega)$, $g \in L^2(\Omega; \mathbb{R}^2)$, and $q \in H(\text{div}, \Omega)$. For all $\varphi \in V := H_0^1(\Omega)$ and $v \in H^1(\Omega)$, an integration by parts leads to

$$\text{Res}(\varphi) = \int_{\Omega} (f + \text{div } q)\varphi \, dx + \int_{\Omega} (q - g - \text{Curl } v) \cdot \nabla \varphi \, dx.$$

Consider some $q \in H(\text{div}, \Omega)$ with the assumption that $\|f + \text{div } q\|_{\star}$ is small in the sense that the dual norm of the residual $\|\text{Res}\|_{\star}$ is much larger,

$$\delta := \|f + \text{div } q\|_{\star} / \|\text{Res}\|_{\star} \ll 1. \tag{3.1}$$

The affirmative examples η_{LW} , η_{B} and η_{MFEM} from Table 1 allow for

$$\|f + \text{div } q\|_{\star} \leq \|h_{\mathcal{T}}(f + \text{div } q)\|_{L^2(\Omega)} / \pi \lesssim \text{osc}(f, \mathcal{T}) := \|h_{\mathcal{T}}(f - f_{\mathcal{T}})\|_{L^2(\Omega)}$$

(from a piecewise Poincaré inequality with Payne-Weinberger constant [23]). Here, $h_{\mathcal{T}} \in P_0(\mathcal{T})$ denotes the local mesh size and $f_{\mathcal{T}} \in P_0(\mathcal{T})$ denotes the piecewise integral mean of f (i.e. $h_{\mathcal{T}}|_T = h_T$ and $f_{\mathcal{T}}|_T := \int_T f \, dx / |T|$ for every $T \in \mathcal{T}$). Then $\|f + \operatorname{div} q\|_{\star}$ is of quadratic order in terms of the mesh-size $h_{\mathcal{T}}$ for piecewise smooth right-hand sides f . Compared to linear convergence in the mesh size $h_{\mathcal{T}}$ for first-order approximations and residuals, δ is small and tends to zero as $\|h_{\mathcal{T}}\|_{L^\infty(\Omega)} \rightarrow 0$.

Given the Helmholtz decomposition [16] (for simply connected domains Ω)

$$q - g = \nabla a + \operatorname{Curl} b$$

with some unique $a \in H_0^1(\Omega)$ and remainder $b \in H^1(\Omega)/\mathbb{R}$, the optimal postprocessing of

$$\eta := \|f + \operatorname{div} q\|_{\star} + \|q - g\|_{L^2(\Omega)} = \|f + \operatorname{div} q\|_{\star} + (\|a\|^2 + \|b\|^2)^{1/2}$$

with $v = b$ results in

$$\mu := \|f + \operatorname{div} q\|_{\star} + \|a\| \leq \eta.$$

Assumption (3.1) and the following theorem imply asymptotic exactness in the sense that

$$\mu / (1 + 2\delta) \leq \|\operatorname{Res}\|_{\star} \leq \mu.$$

Hence, the new error estimator μ overcomes the hyper-circle threshold.

Theorem 3.1 *Under the aforementioned notation it holds*

$$\begin{aligned} \|\operatorname{Res}\|_{\star} &\leq \eta \leq \left((\|\operatorname{Res}\|_{\star} + \|f + \operatorname{div} q\|_{\star})^2 + \|b\|^2 \right)^{1/2} + \|f + \operatorname{div} q\|_{\star}, \\ \|\operatorname{Res}\|_{\star} &\leq \mu \leq \|\operatorname{Res}\|_{\star} + 2\|f + \operatorname{div} q\|_{\star}. \end{aligned}$$

For $\|a\| > 0$ and $\kappa := \|b\| / \|a\|$, it holds

$$0 \leq \eta - \mu = \|a\| \left(\sqrt{1 + \kappa^2} - 1 \right) \leq (\|f + \operatorname{div} q\|_{\star} + \|\operatorname{Res}\|_{\star}) \left(\sqrt{1 + \kappa^2} - 1 \right).$$

Proof The Helmholtz decomposition shows, for all $\varphi \in H_0^1(\Omega)$,

$$\operatorname{Res}(\varphi) = \int_{\Omega} (f + \operatorname{div} q)\varphi \, dx + \int_{\Omega} (q - g - \operatorname{Curl} b) \cdot \nabla \varphi \, dx \leq (\|f + \operatorname{div} q\|_{\star} + \|a\|) \|\varphi\|.$$

Hence,

$$\|\operatorname{Res}\|_{\star} \leq \mu = \|f + \operatorname{div} q\|_{\star} + \|a\|.$$

Moreover,

$$\mu \leq \eta = \|f + \operatorname{div} q\|_{\star} + (\|a\|^2 + \|b\|^2)^{1/2}.$$

The improvement factor of the second term is

$$\|a\|/\sqrt{\|a\|^2 + \|b\|^2} = 1/\sqrt{1 + \kappa^2}.$$

Moreover,

$$\begin{aligned} \|a\|^2 &= \int_{\Omega} \nabla a \cdot \nabla a \, dx = \int_{\Omega} (q - g) \cdot \nabla a \, dx = \operatorname{Res}(a) - \int_{\Omega} (f + \operatorname{div} q) a \, dx \\ &\leq (\|\operatorname{Res}\|_{\star} + \|f + \operatorname{div} q\|_{\star}) \|a\|. \end{aligned}$$

This concludes the proof. □

Remark 3.2 For the optimal $q = \sigma - \operatorname{Curl} \beta$ for σ from (1.2) and some remainder with $\beta \in H^1(\Omega)/\mathbb{R}$ from the Helmholtz decomposition

$$\sigma - g = \nabla \alpha + \operatorname{Curl} \beta,$$

it follows that $\|b\| = 0 = \|f + \operatorname{div} q\|_{\star}$, and hence $\|\operatorname{Res}\|_{\star} = \eta = \mu$. In other words, an optimal q cannot be improved by the proposed postprocessing. Theorem 2.2 is an example for this observation.

Remark 3.3 Theorem 3.1 implies Theorem 2.1. Indeed $q = q_{\text{MFEM}}$ and $\operatorname{div}(\sigma - q_{\text{MFEM}}) \equiv 0$ shows $-\sigma + q_{\text{MFEM}} = \operatorname{Curl} b$. Hence (2.1) is the Helmholtz decomposition of this section with $a = u - u_h$.

Remark 3.4 The modification of the residual has been discussed on a rather abstract level in [24] and for general $H(\operatorname{div}, \Omega)$ functions in [25]. Algorithmic details as the improvements via successive refinements are not reported therein.

Remark 3.5 (on multiply connected domains). The Helmholtz decomposition for some multiply connected domain involves singular functions $\theta_1, \dots, \theta_J \in H^1(\Omega)$ with $\Delta \theta_j = 0$, $\theta_j|_{\Gamma_j} = 1$ and $\theta_j|_{\Gamma_k} = 0$ for $j = 1, \dots, J$ and $k = 0, \dots, J$ for $\partial\Omega = \Gamma_0 \cup \dots \cup \Gamma_J$. For any $q - g \in L^2(\Omega; \mathbb{R}^2)$ there exist $a \in H_0^1(\Omega)$ and $b \in H^1(\Omega)/\mathbb{R}$ as well as $\alpha_1, \dots, \alpha_J \in \mathbb{R}$ such that

$$q - g = \nabla a + \sum_{j=1}^J \alpha_j \nabla \theta_j + \operatorname{Curl} b.$$

Since this decomposition is $L^2(\Omega)$ -orthogonal, the suggested postprocessing may involve $\theta_1, \dots, \theta_J$ as well and reads

$$\mu := \|f + \operatorname{div} q\|_{\star} + \|a\|.$$

The practical realisation via discrete harmonic approximations of $\theta_1, \dots, \theta_J$, however, involves further discretisation errors to be evaluated in order to compute a valid upper error bound.

The remaining part of this section is devoted to inhomogeneous Dirichlet data u_D as they arise in the numerical examples of Sects. 7–9. Suppose for the remainder of this section that along the boundary edges $\mathcal{E}(\partial\Omega) := \{E \in \mathcal{E} \mid E \subset \partial\Omega\}$ the discrete solution u_h satisfies $u_h = \mathcal{I}u_D := \sum_{z \in \mathcal{N}} u_D(z)\varphi_z$.

Theorem 3.6 *Assume that $u_D \in H^1(\Omega) \cap C(\Omega)$ satisfies $u_D \in H^2(E)$ for all $E \in \mathcal{E}(\partial\Omega)$ and let $\partial_{\xi}^2 u_D / \partial s^2$ denote the edgewise second partial derivative of u_D along $\partial\Omega$. Then there exists $w_D \in H^1(\Omega)$ with*

$$\begin{aligned} w_D|_{\partial\Omega} &= u_D|_{\partial\Omega} - \mathcal{I}u_D|_{\partial\Omega}, \\ \text{supp}(w_D) &\subset \bigcup \{T \in \mathcal{T} \mid T \cap \partial\Omega \neq \emptyset\}, \\ \|w_D\|_{L^\infty(\Omega)} &= \|u_D - \mathcal{I}u_D\|_{L^\infty(\partial\Omega)}, \\ \|w_D\| &\leq C_\gamma \|h_{\mathcal{E}}^{3/2} \partial_{\xi}^2 u_D / \partial s^2\|_{L^2(\partial\Omega)}. \end{aligned}$$

Furthermore it holds

$$\|e\|^2 \leq \|\text{Res}\|_{\star}^2 + \|w_D\|^2.$$

Proof For the proof of the existence see [3, 14]. For the proof of the last equation, assume the optimal $w \in H^1(\Omega)$ with $w_D|_{\partial\Omega} = u_D|_{\partial\Omega} - \mathcal{I}u_D|_{\partial\Omega}$ and $\text{div } \nabla w_D \equiv 0$. Then, it holds the orthogonality from [3],

$$\|e\|^2 = \|e - w\|^2 + \|w\|^2 \leq \|\text{Res}\|_{\star}^2 + \|w\|^2 \leq \|\text{Res}\|_{\star}^2 + \|w_D\|^2.$$

□

Remark 3.7 More involved calculations show in [14] that $C_\gamma \leq 0.7043$ for triangulations with right isosceles triangles. However, for the numerical examples in this paper, we use $C_\gamma = 1$.

4 Improvement via red-refinements

This section analyses the reduction property of the postprocessed a posteriori error estimator under successive red-refinements and thereby contributes to the convergence analysis of adaptive mixed finite element methods.

Recall definition (2.2) of the edge-related oscillations of some $f_{\text{eq}} \in P_0(\mathcal{T})$

$$\text{OSC}^2(\mathcal{T}, f_{\text{eq}}) := \sum_{E \in \mathcal{E}(\Omega)} \text{osc}^2(f_{\text{eq}}, \omega_E) + \sum_{E \in \mathcal{E}(\partial\Omega)} |E|^2 \|f_{\text{eq}}\|_{L^2(\omega_E)}^2.$$

Theorem 4.1 (Main result). *Given a triangulation $\mathcal{T} \equiv \mathcal{T}_H$, its red-refinement $\mathcal{T}_h := \text{red}(\mathcal{T})$ and the data $\kappa \in P_0(\mathcal{T})$, $f_{\text{eq}} \in P_0(\mathcal{T})$, $q_{\text{eq}} \in \text{RT}_0(\mathcal{T})$ with $\text{div } q_{\text{eq}} + f_{\text{eq}} = 0$ and $r_{\mathcal{T}} \equiv \kappa \nabla u_C \in P_0(\mathcal{T}; \mathbb{R}^2)$, let $a \in H_0^1(\Omega)$ and $b \in H^1(\Omega)/\mathbb{R}$ be defined in the Helmholtz decomposition $q_{\text{eq}} - r_{\mathcal{T}} = \kappa \nabla a + \text{Curl } b$. Then there exist some constants $0 < \varrho < 1$ and $0 < \Lambda < \infty$, such that*

$$\begin{aligned} & \min_{v_h \in V(\text{red}(\mathcal{T}))} \|\kappa^{-1/2}(r_{\mathcal{T}} - q_{\text{eq}} - \text{Curl } v_h)\|_{L^2(\Omega)}^2 - \|\kappa^{1/2} \nabla a\|_{L^2(\Omega)}^2 \\ & \leq \varrho \left(\min_{v_H \in V(\mathcal{T})} \|\kappa^{-1/2}(r_{\mathcal{T}} - q_{\text{eq}} - \text{Curl } v_H)\|_{L^2(\Omega)}^2 - \|\kappa^{1/2} \nabla a\|_{L^2(\Omega)}^2 \right) \\ & \quad + \Lambda \text{OSC}(\mathcal{T}, f_{\text{eq}})^2. \end{aligned}$$

Remark 4.2 Note that the proof of Theorem 3.1 shows that $\|\kappa^{1/2} \nabla a\|_{L^2(\Omega)}$ equals $\|\text{Res}\|_{\star}$ up to $\|f_{\text{eq}} + \text{div } q_{\text{eq}}\|_{\star} = 0$.

Two interpolation operators J and I will be required for the proof at the end of this section. Let \mathcal{T} be a regular triangulation of the polygonal Lipschitz domain Ω into triangles with its set of edges \mathcal{E} and its red-refinement $\text{red}(\mathcal{T})$. Given any $v \in H^1(\Omega)$ let $Jv \in P_1(\mathcal{T}) \cap C(\overline{\Omega})$ be some quasi-interpolant [6] with

$$\|h_{\mathcal{T}}^{-1}(v - Jv)\|_{L^2(\Omega)} + \|h_{\mathcal{E}}^{-1/2}(v - Jv)\|_{L^2(\cup \mathcal{E})} + \|Jv\| \lesssim \|v\|. \tag{4.1}$$

Moreover, for any $w \in H^1(\Omega)$, set

$$Iw := \sum_{E \in \mathcal{E}} \left(\int_E w \, ds \right) 2\varphi_E \tag{4.2}$$

with the integral mean $\int_E w \, ds := \int_E w \, ds / |E|$ of w along the edge E and φ_E the nodal basis function of the Courant FEM with respect to the red-refined triangulation $\text{red}(\mathcal{T})$ and the midpoint $\text{mid}(E)$ of an edge $E \in \mathcal{E}$, $\varphi_E(\text{mid}(E)) = 1$ and $\varphi_E(y) = 0$ for all other nodes y of $\text{red}(\mathcal{T})$.

Lemma 4.3 *Given any $v \in H^1(\Omega)$ with Jv from (4.1) and $w := v - Jv$ let Iw be defined in (4.2). Then, it holds*

$$\|h_{\mathcal{T}}^{-1}Iw\|_{L^2(\Omega)} + \|h_{\mathcal{E}}^{-1/2}Iw\|_{L^2(\cup \mathcal{E})} + \|Iw\| \lesssim \|v\|.$$

Proof From (4.2), $\|\varphi_E\|_{L^2(T)} = |T|^{1/2} / \sqrt{8}$, and a Cauchy inequality along the edges $\mathcal{E}(T)$ of $T \in \mathcal{T}$, it follows that

$$\|Iw\|_{L^2(T)} \leq \sum_{E \in \mathcal{E}(T)} 2\|\varphi_E\|_{L^2(T)} \left| \int_E w \, ds \right| \leq \sum_{E \in \mathcal{E}(T)} |T|^{1/2} |E|^{-1/2} \|w\|_{L^2(E)} / \sqrt{2}.$$

Hence,

$$\|h_{\mathcal{T}}^{-1}Iw\|_{L^2(\Omega)}^2 \lesssim \sum_{E \in \mathcal{E}} |E|^{-1} \|v - Jv\|_{L^2(E)}^2.$$

Therefore, (4.1) implies

$$\|h_{\mathcal{T}}^{-1}Iw\|_{L^2(\Omega)} \lesssim \|v\|.$$

An inverse inequality for the piecewise affine Iw shows

$$\|Iw\| \lesssim \|h_{\mathcal{T}}^{-1}Iw\|_{L^2(\Omega)} \lesssim \|v\|.$$

A trace inequality [9] concludes the proof,

$$\|h_{\mathcal{E}}^{-1/2}Iw\|_{L^2(\cup \mathcal{E})} \lesssim \|h_{\mathcal{T}}^{-1}Iw\|_{L^2(\Omega)} + \|Iw\| \lesssim \|v\|.$$

□

Lemma 4.4 *Given $q_{\text{eq}} \in \text{RT}_0(\mathcal{T})$ with $\text{div } q_{\text{eq}} + f_{\text{eq}} = 0$, $\kappa \in P_0(\mathcal{T})$, $r_{\mathcal{T}} \in P_0(\mathcal{T}; \mathbb{R}^2)$, and $v \in H^1(\Omega)$, set $v_h := I(v - Jv)$. Then, it holds*

$$\left| \int_{\Omega} \kappa^{-1}(q_{\text{eq}} - r_{\mathcal{T}}) \cdot \text{Curl}(v - Jv - v_h) \, dx \right| \lesssim \text{OSC}(\mathcal{T}, f_{\text{eq}}) \|\kappa^{-1/2} \nabla v\|_{L^2(\Omega)}.$$

Proof It is important to notice that $2\int_E \varphi_E \, ds = 1$ and so

$$\int_E (w - Iw) \, ds = 0 \quad \text{for all } E \in \mathcal{E}. \tag{4.3}$$

Since $w := v - Jv - v_h$ satisfies $\int_E w \, ds = 0$ for all $E \in \mathcal{E}$, a piecewise integration by parts leads to

$$\begin{aligned} \int_{\Omega} \kappa^{-1}(q_{\text{eq}} - r_{\mathcal{T}}) \cdot \text{Curl } w \, dx &= \sum_{T \in \mathcal{T}} \int_{\partial T} w \kappa^{-1}|_T (q_{\text{eq}} - r_{\mathcal{T}}) \cdot \tau_T \, ds \\ &= \sum_{E \in \mathcal{E}} \int_E w [\kappa^{-1}q_{\text{eq}}]_E \cdot \tau_E \, ds. \end{aligned}$$

Here, τ_E denotes the tangent along $E \in \mathcal{E}$. Any 2D Raviart–Thomas finite element function allows for some representation [20]

$$q_{\text{eq}}(x)|_T = 1/2 (x - \text{mid}(T)) \text{div } q_{\text{eq}}|_T + \kappa \nabla_{\text{NC}} \tilde{u}_{\text{CR}} \quad \text{for } x \in T \in \mathcal{T}$$

with the piecewise gradient ∇_{NC} of some Crouzeix–Raviart function $u_{\text{CR}} \in \text{CR}(\mathcal{T})$. Hence $[\kappa^{-1}q_{\text{eq}}]_E \cdot \tau_E$ equals $1/2 [\text{div } q_{\text{eq}}]_E (x - \text{mid}(T)) \cdot \tau_E$ plus some constant for $x \in E \in \mathcal{E}$ from $\partial \tilde{u}_{\text{CR}} / \partial s$. The integral of the latter multiplied by w vanishes

$$\int_E w [\kappa^{-1}q_{\text{eq}}]_E \cdot \tau_E \, ds = 1/2 \int_E ([\kappa^{-1} \text{div } q_{\text{eq}}]_E \cdot \tau_E) w(x) ((x - \text{mid}(E)) \cdot \tau_E) \, ds_x.$$

The modulus of the previous term is bounded from above by

$$\begin{aligned} & 1/2 \left| [\varkappa^{-1} \operatorname{div} q_{\text{eq}}]_E \cdot \tau_E \right| \|(\bullet - \operatorname{mid}(E)) \cdot \tau_E\|_{L^2(E)} \|w\|_{L^2(E)} \\ &= |E|^{3/2} / (4\sqrt{3}) \left| [\varkappa^{-1} \operatorname{div} q_{\text{eq}}]_E \cdot \tau_E \right| \|w\|_{L^2(E)} \\ &\lesssim \begin{cases} \operatorname{osc}(\varkappa^{-1} \operatorname{div} q_{\text{eq}}, \omega_E) \| |E|^{-1/2} w \|_{L^2(E)} & \text{if } E \in \mathcal{E}(\Omega), \\ |E| \| \varkappa^{-1} \operatorname{div} q_{\text{eq}} \|_{L^2(\omega_E)} \| |E|^{-1/2} w \|_{L^2(E)} & \text{if } E \in \mathcal{E}(\partial\Omega). \end{cases} \end{aligned}$$

This plus a Cauchy inequality in $\mathbb{R}^{|\mathcal{E}|}$ followed by Lemma 4.3 to bound the sum of all $\| |E|^{-1/2} w \|_{L^2(E)}^2$ leads to

$$\left| \int_{\Omega} \varkappa^{-1} (q_{\text{eq}} - r_{\mathcal{T}}) \cdot \operatorname{Curl} w \, dx \right| \lesssim \operatorname{OSC}(\mathcal{T}, f_{\text{eq}}) \|v - Jv\|.$$

In the sequence all constants hidden in the notation \lesssim may depend on \varkappa . The proof concludes by (4.1) and $\|v - Jv\| \lesssim \|\varkappa^{-1/2} \nabla v\|_{L^2(\Omega)}$. □

Proof of Theorem 4.1 The application to the postprocessing concerns the term

$$\textcircled{i} := \|\varkappa^{-1/2} (q_{\text{eq}} - r_{\mathcal{T}} - \operatorname{Curl} v_H)\|_{L^2(\Omega)}$$

on the triangulation \mathcal{T}_H with an optimal $v_H \in (P_1(\mathcal{T}) \cap C(\overline{\Omega}))/\mathbb{R}$. On the red-refined triangulation $\mathcal{T}_h \equiv \operatorname{red}(\mathcal{T})$, the optimal postprocessing leads to

$$\textcircled{ii} := \|\varkappa^{-1/2} (q_{\text{eq}} - r_{\mathcal{T}} - \operatorname{Curl} (v_H - v_h))\|_{L^2(\Omega)}$$

for some minimiser $v_h \in (P_1(\mathcal{T}_h) \cap C(\overline{\Omega}))/\mathbb{R}$. The Helmholtz decomposition

$$q_{\text{eq}} - r_{\mathcal{T}} - \operatorname{Curl} v_H = \varkappa \nabla a + \operatorname{Curl} b$$

for unique $a \in H_0^1(\Omega)$ and $b \in H^1(\Omega)/\mathbb{R}$ allows for the representations

$$\begin{aligned} \textcircled{i} &= \|\varkappa^{1/2} \nabla a\|_{L^2(\Omega)}^2 + \|\varkappa^{-1/2} \operatorname{Curl} b\|_{L^2(\Omega)}^2, \\ \textcircled{ii} &= \|\varkappa^{1/2} \nabla a\|_{L^2(\Omega)}^2 + \|\varkappa^{-1/2} \operatorname{Curl} (b - v_h)\|_{L^2(\Omega)}^2. \end{aligned}$$

The optimality of v_H implies, for all $w_H \in (P_1(\mathcal{T}_H) \cap C(\overline{\Omega}))/\mathbb{R}$, that

$$\int_{\Omega} \varkappa^{-1} \operatorname{Curl} b \cdot \operatorname{Curl} w_H \, dx = \int_{\Omega} \varkappa^{-1} (q_{\text{eq}} - r_{\mathcal{T}} - \operatorname{Curl} v_H) \cdot \operatorname{Curl} w_H \, dx = 0 \tag{4.4}$$

and that of v_h implies, for all $w_h \in (P_1(\mathcal{T}_h) \cap C(\overline{\Omega}))/\mathbb{R}$, that

$$\int_{\Omega} \chi^{-1} \text{Curl}(b - v_h) \cdot \text{Curl } w_h \, dx = 0. \tag{4.5}$$

Since $Jb \in V_H$, (4.4) leads to

$$\|\chi^{-1/2} \text{Curl } b\|_{L^2(\Omega)}^2 = \int_{\Omega} \chi^{-1} \text{Curl } b \cdot \text{Curl } b \, dx = \int_{\Omega} \chi^{-1} \text{Curl } b \cdot \text{Curl } (b - Jb) \, dx.$$

With $b_h := I(b - Jb)$ this term equals

$$\begin{aligned} & \int_{\Omega} \chi^{-1} (q_{\text{eq}} - r_{\mathcal{T}} - \text{Curl } v_H) \cdot \text{Curl } b_h \, dx \\ & + \int_{\Omega} \chi^{-1} (q_{\text{eq}} - r_{\mathcal{T}} - \text{Curl } v_H) \cdot \text{Curl } (b - Jb - b_h) \, dx. \end{aligned}$$

Lemma 4.4 shows

$$\begin{aligned} \|\chi^{-1/2} \text{Curl } b\|_{L^2(\Omega)}^2 & \leq \int_{\Omega} \chi^{-1} (q_{\text{eq}} - r_{\mathcal{T}} - \text{Curl } v_H) \cdot \text{Curl } b_h \, dx \\ & + c_1 \|\chi^{-1/2} \text{Curl } b\|_{L^2(\Omega)} \text{OSC}(\mathcal{T}, f_{\text{eq}}). \end{aligned}$$

Equation (4.5) for $w_h = b_h$ implies

$$\begin{aligned} & \int_{\Omega} \chi^{-1} (q_{\text{eq}} - r_{\mathcal{T}} - \text{Curl } v_H) \cdot \text{Curl } b_h \, dx \\ & = \int_{\Omega} \chi^{-1} \text{Curl } b \cdot \text{Curl } b_h \, dx \\ & = \int_{\Omega} \chi^{-1} \text{Curl } v_h \cdot \text{Curl } b_h \, dx \\ & \leq \|\chi^{-1/2} \text{Curl } v_h\|_{L^2(\Omega)} \|\chi^{-1/2} \text{Curl } b_h\|_{L^2(\Omega)}. \end{aligned}$$

Lemma 4.3 shows

$$\|\chi^{-1/2}\text{Curl } b_h\|_{L^2(\Omega)} \leq c_2 \|\chi^{-1/2}\text{Curl } b\|_{L^2(\Omega)}.$$

Alltogether, it follows

$$\|\chi^{-1/2}\text{Curl } b\|_{L^2(\Omega)} \leq c_1 \text{OSC}(\mathcal{T}, f_{\text{eq}}) + c_2 \|\chi^{-1/2}v_h\|_{L^2(\Omega)}. \tag{4.6}$$

The orthogonality $\chi^{-1/2}\text{Curl}(b - v_h) \perp \chi^{-1/2}\text{Curl}(P_1(\mathcal{T}_h) \cap C(\overline{\Omega}))/\mathbb{R}$ from (4.5) leads to

$$\|\chi^{-1/2}\text{Curl } b\|_{L^2(\Omega)}^2 = \|\chi^{-1/2}\text{Curl } (b - v_h)\|_{L^2(\Omega)}^2 + \|\chi^{-1/2}\text{Curl } v_h\|_{L^2(\Omega)}^2.$$

The last term is bounded by (4.6) and hence

$$\begin{aligned} &\|\chi^{-1/2}\text{Curl } (b - v_h)\|_{L^2(\Omega)}^2 + 1/(2c_2^2) \|\chi^{-1/2}\text{Curl } b\|_{L^2(\Omega)}^2 - c_1^2/c_2^2 \text{OSC}(\mathcal{T}, f_{\text{eq}})^2 \\ &\leq \|\chi^{-1/2}\text{Curl } b\|_{L^2(\Omega)}^2. \end{aligned}$$

With $\varrho := 1 - 1/(2c_2^2)$ and $\Lambda := c_1^2/c_2^2$ this reads

$$\|\chi^{-1/2}\text{Curl } (b - v_h)\|_{L^2(\Omega)}^2 \leq \varrho \|\chi^{-1/2}\text{Curl } b\|_{L^2(\Omega)}^2 + \Lambda \text{OSC}(\mathcal{T}, f_{\text{eq}})^2.$$

□

Remark 4.5 With $q = q_B, \chi \equiv 1, r_{\mathcal{T}} = g$ and $f_{\text{eq}} = f \in P_0(\mathcal{T})$ it holds $\|\nabla a\|_{L^2(\Omega)} = \|\sigma - g\|_{L^2(\Omega)}$ and therefore

$$\eta_{\text{Br}}^2 = \|q_B - g - \text{Curl } v_h\|_{L^2(\Omega)}^2 = \|\sigma - g\|_{L^2(\Omega)}^2 + \|\text{Curl } (b - v_h)\|_{L^2(\Omega)}^2$$

and

$$\eta_{\text{MFEM}}^2 = \|\sigma - g\|_{L^2(\Omega)}^2 + \|\text{Curl } (b - v_H)\|_{L^2(\Omega)}^2.$$

Theorem 4.1 then implies Theorem 2.3, i.e.,

$$\eta_{\text{Br}}^2 - \|\sigma - g\|_{L^2(\Omega)}^2 \leq \varrho \left(\eta_{\text{MFEM}}^2 - \|\sigma - g\|_{L^2(\Omega)}^2 \right) + \Lambda \text{OSC}(\mathcal{T}, f_{\text{eq}})^2.$$

5 A posteriori error estimators

This section recalls some details of known a posteriori error estimators for the design of admissible averagings q and so presents the setting behind Table 1.

5.1 Braess equilibration error estimator

For some piecewise or broken Raviart–Thomas element $g \in \text{RT}_{-1}(\mathcal{T}) := \{q \in L^2(\Omega; \mathbb{R}^2) \mid \forall T \in \mathcal{T}, q|_T \in \text{RT}_0(T)\}$ with

$$\int_{\omega_z} f \varphi_z \, dx = \int_{\omega_z} g \cdot \nabla \varphi_z \, dx \quad \text{for all } z \in \mathcal{M},$$

Braess [6] designs patchwise broken Raviart–Thomas functions $r_z \in \text{RT}_{-1}(\mathcal{T}(z))$ with

$$\begin{aligned} \text{div } r_z|_T &= - \int_T f \varphi_z \, dx / |T| \quad \text{for } T \in \mathcal{T}(z), \\ [r_z \cdot \nu_E]_E &= -[g \cdot \nu_E]_E / 2 \quad \text{on } E \in \mathcal{E}(z) \cap \mathcal{E}(\partial\Omega), \\ r_z \cdot \nu &= 0 \quad \text{along } \partial\omega_z \setminus \mathcal{E}(\partial\Omega). \end{aligned}$$

The solution r_z of these problems is unique up to multiplicatives of $\text{Curl } \varphi_z$ and may be chosen such that $\|r_z\|_{L^2(\omega_z)}$ is minimal. Eventually, the quantity $q_B := g + \sum_{z \in \mathcal{N}} r_z \in \text{RT}_0(\mathcal{T})$ satisfies

$$\text{div } q_B = -f_{\mathcal{T}}$$

and so allows the dual norm estimate $\|f + \text{div } q_B\|_{\star} \leq \text{osc}(f, \mathcal{T})/\pi$. The error estimator reads

$$\eta_B := \|q_B - g\|_{L^2(\Omega)} + \text{osc}(f, \mathcal{T})/\pi.$$

Assume $g = \nabla u_h$ equals the gradient of the solution u_h of the P_1 finite element method for the Poisson problem. Then, the best Raviart–Thomas function $q \in \text{RT}_0(\mathcal{T})$ with $\text{div } q + f_{\mathcal{T}} = 0$ equals the gradient of the mixed finite element solution q_{MFEM} . This motivates the error estimator

$$\eta_{\text{MFEM}} := \|q_{\text{MFEM}} - g\|_{L^2(\Omega)} + \text{osc}(f, \mathcal{T})/\pi := \min_{\substack{q \in \text{RT}_0(\mathcal{T}) \\ \text{div } q = -f_{\mathcal{T}}}} \|q - g\|_{L^2(\Omega)} + \text{osc}(f, \mathcal{T})/\pi.$$

5.2 Least square error estimators

An integration by parts yields, for any $q \in H(\text{div}, \Omega)$ and elementwise integral mean $f_{\mathcal{T}} \in P_0(\mathcal{T})$ of f , that

$$\int_{\Omega} (\sigma - g) \cdot \nabla v \, dx = \int_{\Omega} (f - f_{\mathcal{T}})v \, dx + \int_{\Omega} (f_{\mathcal{T}} + \text{div } q)v \, dx + \int_{\Omega} (g - q) \cdot \nabla v \, dx.$$

After [12,25,26], this results in the error estimator

$$\eta_{LS} := \min_{q \in RT_0(T)} (C_F \|f_T + \operatorname{div} q\|_{L^2(\Omega)} + \|q - g\|_{L^2(\Omega)}) + \operatorname{osc}(f, T)/\pi$$

with Friedrichs’ constant $C_F := \sup_{v \in V \setminus \{0\}} \|v\|_{L^2(\Omega)} / \|v\| \leq \operatorname{width}(\Omega)/\pi$. The Repin variant (without the oscillation split) reads

$$\eta_{\operatorname{Repin}} := \min_{q \in RT_0(T)} (C_F \|f + \operatorname{div} q\|_{L^2(\Omega)} + \|q - g\|_{L^2(\Omega)}).$$

In practise $\eta_{\operatorname{Repin}}$ and η_{LS} are approximated by a series of least-square problems [27]. For the numerical experiments documented in this paper, Algorithm 5.1 approximates η_{LS} with three iterations.

Algorithm 5.1 INPUT $g \in P_0(T; \mathbb{R}^2)$, $C_F > 0$ and $f \in L^2(\Omega)$. Set $\lambda = 1$.

For $j = 1, 2, 3$ do

$$q_{LS} = \operatorname{argmin}_{q \in RT_0(T)} \left((1 + \lambda) C_F^2 \|f_T + \operatorname{div} q\|_{L^2(\Omega)}^2 + (1 + 1/\lambda) \|q - g\|_{L^2(\Omega)}^2 \right),$$

$$\lambda = \|q_{LS} - g\|_{L^2(\Omega)} / (C_F \|f_T + \operatorname{div} q_{LS}\|_{L^2(\Omega)}). \quad \text{od}$$

OUTPUT $\eta_{LS} := C_F \|f_T + \operatorname{div} q\|_{L^2(\Omega)} + \|q_{LS} - g\|_{L^2(\Omega)} + \operatorname{osc}(f, T)/\pi$

5.3 Luce–Wohlmuth equilibration error estimator

The design of q_{LW} from [19] assumes some piecewise or broken Raviart–Thomas element $g \in RT_{-1}(T)$ with

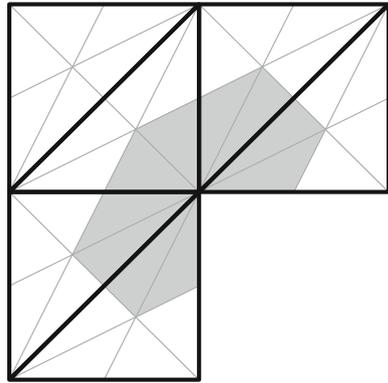
$$\int_{\omega_z} f \varphi_z \, dx = \int_{\omega_z} g \cdot \nabla \varphi_z \, dx \quad \text{for all } z \in \mathcal{M}$$

and employs the dual triangulation \mathcal{T}^* which connects each $\operatorname{mid}(T)$ with adjacent nodes and edge midpoints and so divides every $T \in \mathcal{T}$ into six triangles of area $|T|/6$ (Fig. 1).

Consider some node $z \in \mathcal{N}(T)$ and its nodal basis function φ_z^* with the fine patch $\omega_z^* := \{\varphi_z^* > 0\}$ of the dual triangulation \mathcal{T}^* and its neighbouring triangles $\mathcal{T}^*(z) := \{T^* \in \mathcal{T}^* \mid z \in \mathcal{N}^*(T^*)\}$. Since $g \in P_0(T; \mathbb{R}^2)$ is continuous along $\partial\omega_z^* \cap T$ for any $T \in \mathcal{T}$, $q \cdot v = g \cdot v \in P_0(\mathcal{E}^*(\partial\omega_z^*))$ is well-defined on the boundary edges $\mathcal{E}^*(\partial\omega_z^*)$ of ω_z^* . The further design employs an interpolation $f^* \in P_0(\mathcal{T}^*)$ of f defined by

$$f^*|_{T_\pm^*} := 3 \int_T f \varphi_z \, dx / |T| \quad \text{for the two } T_\pm^* \in \mathcal{T}^* \text{ with } \mathcal{N}^*(T_\pm^*) \cap \mathcal{N}(T) = \{z\}$$

Fig. 1 Exemplary triangulation \mathcal{T} (thick lines) and dual mesh \mathcal{T}^* (thin lines) and one node patch ω_z^* for the L-shaped domain



With the set

$$Q(\mathcal{T}^*(z)) := \{\tau_h \in \text{RT}_0(\mathcal{T}^*(z)) \mid f^* + \text{div } \tau_h = 0 \text{ in } \omega_z^* \text{ and } \tau_h \cdot \nu = g \cdot \nu \text{ along } \partial\omega_z^* \setminus \partial\Omega\},$$

one computes the minimiser

$$q_{\text{LW}}|_{\omega_z^*} := \underset{\tau_h \in Q(\mathcal{T}^*(z))}{\operatorname{argmin}} \|g - \tau_h\|_{L^2(\omega_z^*)}.$$

The choice of f^* differs from the original one of [19] for an improved bound for $\|f + \text{div } q_{\text{LW}}\|_\star$ with explicitly known constants, namely

$$\|f + \text{div } q_{\text{LW}}\|_\star \leq \|h_{\mathcal{T}}(f + \text{div } q_{\text{LW}})\|_{L^2(\Omega)}/\pi.$$

In our preferred modification, the Luce–Wohlmuth error estimator reads

$$\eta_{\text{LW}} := \|q_{\text{LW}} - g\|_{L^2(\Omega)} + \|h_{\mathcal{T}}(f + \text{div } q_{\text{LW}})\|_{L^2(\Omega)}/\pi.$$

6 Numerical realisation

This section concerns the calculation of some postprocessing γ and the adaptive mesh design in the numerical experiments.

6.1 Realisation of the postprocessing

The postprocessing is based on a minimisation within continuous and piecewise affine functions $v \in P_1(\tilde{\mathcal{T}}) \cap C(\Omega)$, namely

$$\|\nabla u - q - \text{Curl } v\|_{L^2(\Omega)} := \min_{v \in P_1(\tilde{\mathcal{T}}) \cap C(\Omega)} \|\nabla u - q - \text{Curl } v\|_{L^2(\Omega)}.$$

Given an equilibrated quantity $q \in \text{RT}_0(\widehat{T})$ on some triangulation \widehat{T} of Ω (e.g. $\widehat{T} \in \{\mathcal{T}, \mathcal{T}^*, \text{red}(\mathcal{T}), \text{red}(\mathcal{T}^*)\}$), the minimisation of the right-hand side over $v \in P_1(\widehat{T}) \cap C(\Omega)$ results in some linear system of equations $Ax = b$. Given some basis $\{\varphi_{z_1}, \dots, \varphi_{z_N}\}$ of $P_1(\widehat{T}) \cap C(\Omega)$, the stiffness matrix and right-hand side vector read

$$A_{jk} := \int_{\Omega} \text{Curl } \varphi_{z_j} \cdot \text{Curl } \varphi_{z_k} \, dx \quad \text{and} \quad b_j := \int_{\Omega} (g - q) \cdot \text{Curl } \varphi_{z_j} \, dx.$$

The Matlab routine `pcg` with Jacobi (or diagonal) preconditioner $D = \text{diag}(A_{11}, \dots, A_{NN})$ and initial value $x = 0$ solves this system iteratively in k iterations. The first iterate reads

$$x_1 := \frac{b^T D^{-1/2} b}{b^T D^{-1/2} A D^{-1/2} b} D^{-1/2} b.$$

Since, $\text{Curl}(P_1(\mathcal{T}) \cap C(\Omega)) \subset \text{RT}_0(\mathcal{T})$ and $\|g - q_{\text{MFEM}}\|_{L^2(\Omega)}$ is already the best-approximation in $\text{RT}_0(\mathcal{T})$, there is no improvement by the postprocessing with $\widehat{T} = \mathcal{T}$ in case of the MFEM error estimator, unless one refines the mesh or increases the polynomial degrees. But the postprocessing with $\widehat{T} = \mathcal{T}$ may reduce the gap between η_{MFEM} and the Braess equilibration error estimator η_B . We suggest to use $\widehat{T} = \mathcal{T}^*$ for the Luce–Wohlmuth error estimator and a red-refinement $\widehat{T} = \text{red}(\mathcal{T})$ for all other error estimators, see also Table 1. Postprocessed error estimators based on η_{xyz} with $\widehat{T} = \text{red}(\mathcal{T})$ and k iterations are labelled as $\eta_{xyz(k)}$. The Luce–Wohlmuth error estimator with postprocessing $\widehat{T} = \mathcal{T}^*$ and k iterations is labelled as $\eta_{\text{LW}(k)}$. The Braess error estimator with postprocessing on $\widehat{T} = \mathcal{T}$ and k iterations is labelled as $\eta_{\text{B}(k)}$.

Remark 6.1 In the 3D case, the minimisation problem in Sect. 6.1 involves the $\text{Curl} := \nabla \times \psi$ of functions ψ in $H^1(\Omega; \mathbb{R}^3)$. This causes modifications in the realisation of the postprocessing, either by the choice of a proper basis of $P^1(\Omega; \mathbb{R}^3) \cap C(\Omega)$ or by $H(\text{curl}, \Omega)$ -conforming finite elements.

6.2 AFEM algorithm

Automatic mesh refinement generates a sequence of meshes $\mathcal{T}_0, \mathcal{T}_1, \mathcal{T}_2 \dots$ by successive mesh refining according to a bulk criterion with parameter $0 < \Theta \leq 1$.

Algorithm 6.2 INPUT coarse mesh \mathcal{T}_0 For any level $\ell = 0, 1, 2, \dots$ do
 COMPUTE discrete solution u_ℓ on \mathcal{T}_ℓ with $\text{ndof} := |\mathcal{N}_\ell(\Omega)|$ degrees of freedom, global upper bounds $\eta_{\text{LW}}, \eta_B, \eta_{\text{MFEM}}, \eta_{\text{LS}}$ and the postprocessed quantities $\eta_{xyz(k)}$ from Sect. 6 for $k \in \{1, 3, 5, \infty\}$, and refinement indicators

$$\eta_\ell(T)^2 = |T| \|f\|_{L^2(T)}^2 + |T|^{1/2} \sum_{E \in \mathcal{E}(T)} \|[\nabla u_\ell]_E \cdot \nu_E\|_{L^2(E)}^2 \quad \text{for all } T \in \mathcal{T}_\ell.$$

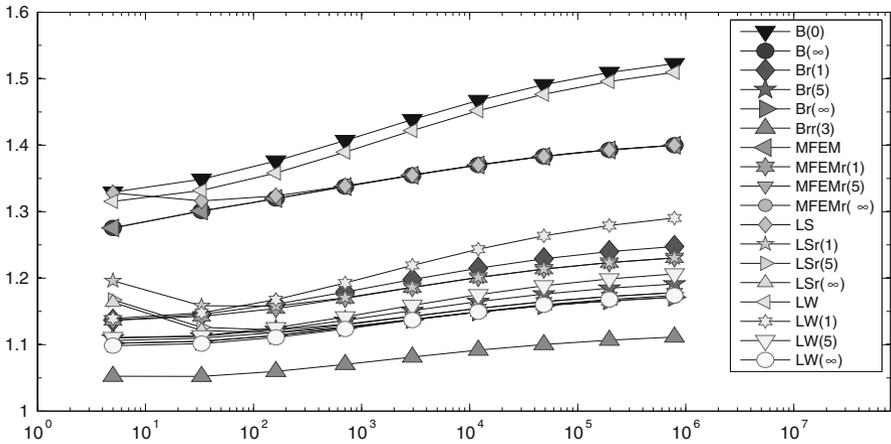


Fig. 2 History of efficiency indices $\eta_{xyz}/\|e\|$ of various a posteriori error estimators η_{xyz} labelled xyz as functions of ndof on uniform meshes in Example 7.1

MARK minimal set (for adaptive mesh-refinement) $\mathcal{M}_\ell \subseteq \mathcal{T}_\ell$ of elements such that

$$1/2 \sum_{T \in \mathcal{T}_\ell} \eta_\ell(T)^2 \leq \sum_{T \in \mathcal{M}_\ell} \eta_\ell(T)^2.$$

(For uniform mesh-refinement set $\mathcal{M}_\ell = \mathcal{T}_\ell$.)

REFINE \mathcal{T}_ℓ by *red*-refinement of elements in \mathcal{M}_ℓ and *red-green-blue*-refinement of further elements to avoid hanging nodes and compute $\mathcal{T}_{\ell+1}$. od

OUTPUT efficiency indices $\text{eff} := \eta_{xyz}/\|e\|$ for error $e = u - u_\ell$ of exact solution u and

$$\rho_{xyz,r(k)} := \left(\eta_{xyzr(k)}^2 - \|e\|^2 \right) / \left(\eta_{xyz}^2 - \|e\|^2 \right). \tag{6.1}$$

The quantity $\rho_{xyz,k}$ measures the improvement of the relative error of η_{xyz} by $\eta_{xyzr(k)}$ after k iterations.

7 Numerical examples for Poisson problems

The first two numerical example concern the Poisson problem and the residual for the solution u_h of the P_1 finite element method

$$\text{Res}(v) := \int_{\Omega} f v \, dx - \int_{\Omega} \nabla u_h \cdot \nabla v \, dx.$$

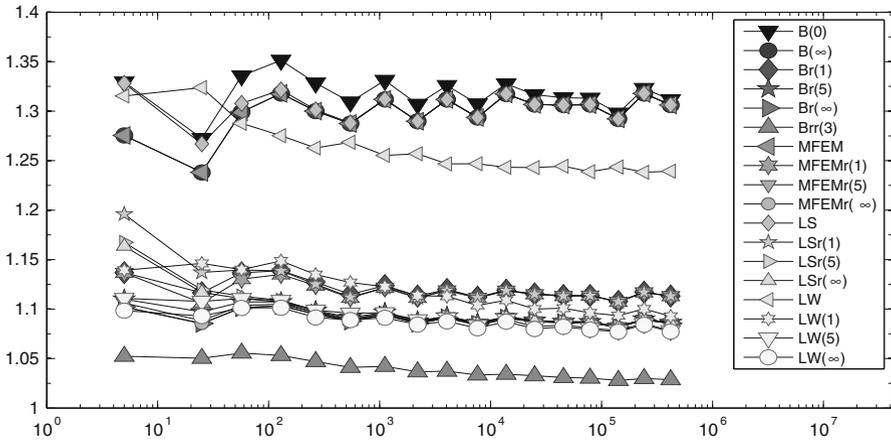


Fig. 3 History of efficiency indices $\eta_{xyz}/\|e\|$ of various a posteriori error estimators η_{xyz} labelled xyz as functions of ndof on adaptive meshes in Example 7.1

7.1 L-shaped domain example

The first benchmark problem employs $f \equiv \kappa \equiv 1$ and homogeneous Dirichlet data $u_D \equiv 0$. The exact energy error can be calculated by

$$\|e\|^2 = \|u\|^2 - \|u_h\|^2 \quad \text{and} \quad \|u\|^2 = 0.214075802680976$$

(computed with higher-order FEMs). The problem involves a typical corner singularity and shows an experimental convergence rate of $1/3$ for uniform mesh refinement.

Figures 2 and 3 display the efficiency indices of the a posteriori error estimators η_B , η_{LW} and η_{MFEM} from Sect. 5 and their postprocessed quantities from Sect. 6.

One main observation is that, even after a single iteration $k = 1$, the postprocessing significantly improves efficiency indices. The efficiency indices of all error estimators are reduced by 10 to 15%. The gap between $\eta_{B(0)}$ and $\eta_{MFEM(0)}$ allows a significant improvement by the postprocessing $\eta_{B(\infty)}$ based on \mathcal{T} . The potential of efficiency improvement appears even larger in case of uniform mesh refinement. The postprocessing of the Luce–Wohlmuth equilibration error estimator and the MFEM or LS error estimator leads to very good efficiency indices around 1.15. Table 2 lists the improvement numbers from (6.1) typically by 0.5 or below. In agreement with the results of Sect. 4, the further red-refinement in the postprocessing results in a significant improvement of the accuracy. Table 2 also displays the quotient $\varrho_{B,r(1)}/\varrho_{B,r(\infty)}$. It increases on coarse meshes, but stabilises with an increasing number of degrees of freedom. Hence, we assume a pre-asymptotic phenomenon and the postprocessing by one iteration should maintain its effectivity also for more complex problems. Table 3 for adaptive mesh-refinement hardens this assumption.

Table 2 Improvement numbers $\varrho_{X,y,z,r}(k)$ from (6.1) for three postprocessed error estimators and uniform mesh-refinement in Example 7.1

ndof	$\varrho_{B,r}(1)$	$\varrho_{B,r}(\infty)$	$\varrho_{B,r}(1)/\varrho_{B,r}(\infty)$	$\varrho_{B,rr}(3)$	$\varrho_{MFEM,r}(1)$	$\varrho_{MFEM,r}(\infty)$	$\varrho_{LW,(1)}$	$\varrho_{LW,(\infty)}$
5	0.38092	0.27966	1.3621	0.14012	0.46529	0.34225	0.40774	0.28313
33	0.38049	0.26985	1.41	0.13074	0.44053	0.31897	0.41017	0.27559
161	0.3888	0.26927	1.4439	0.13763	0.4474	0.32428	0.43019	0.27713
705	0.39866	0.27212	1.465	0.14839	0.46601	0.33784	0.45411	0.28168
2,945	0.407	0.27534	1.4782	0.15838	0.48643	0.35241	0.47581	0.28624
12,033	0.41322	0.27798	1.4865	0.16629	0.50431	0.36514	0.49304	0.28991
48,641	0.41753	0.2799	1.4917	0.17203	0.51826	0.37509	0.50565	0.29259
195,585	0.42039	0.28123	1.4948	0.17599	0.52834	0.38234	0.51441	0.29443
784,385	0.42225	0.28212	1.4967	0.17863	0.53529	0.38736	0.52029	0.29566

Table 3 Improvement numbers $\varrho_{x,y,z,r}(k)$ from (6.1) for three postprocessed error estimators and adaptive mesh-refinement in Example 7.1

ndof	$\varrho_{B,r}(1)$	$\varrho_{B,r}(\infty)$	$\varrho_{B,r}(1)/\varrho_{B,r}(\infty)$	$\varrho_{B,r}(3)$	$\varrho_{MFEM,r}(1)$	$\varrho_{MFEM,r}(\infty)$	$\varrho_{LW,r}(1)$	$\varrho_{LW,r}(\infty)$
5	0.38092	0.27966	1.3621	0.14012	0.46529	0.34225	0.40774	0.28313
25	0.39806	0.28803	1.382	0.16641	0.41961	0.33449	0.41696	0.25861
57	0.37253	0.27164	1.3714	0.14544	0.40368	0.30979	0.45472	0.32275
129	0.35914	0.26308	1.3651	0.13195	0.39098	0.29501	0.51038	0.34006
265	0.35384	0.25941	1.364	0.12557	0.38107	0.28751	0.4842	0.32202
544	0.34031	0.25691	1.3246	0.11747	0.3576	0.27893	0.44258	0.30469
1,114	0.34046	0.25505	1.3349	0.11142	0.35794	0.27302	0.45204	0.33111
2,195	0.33795	0.25216	1.3402	0.10547	0.35449	0.26902	0.41358	0.3034
4,043	0.33528	0.25112	1.3351	0.10003	0.34904	0.26399	0.43015	0.32974
7,616	0.33376	0.24899	1.3405	0.095774	0.3489	0.26206	0.39348	0.30257
13,839	0.33108	0.24931	1.328	0.091424	0.34142	0.25832	0.4212	0.33415
24,939	0.33262	0.24814	1.3405	0.090495	0.34348	0.25715	0.3855	0.30526
45,374	0.33103	0.24825	1.3334	0.08625	0.33856	0.25509	0.38632	0.31184
78,791	0.33192	0.24827	1.3369	0.084756	0.33929	0.2541	0.37714	0.308
141,567	0.33117	0.24786	1.3361	0.082686	0.33744	0.25345	0.35795	0.29315
241,800	0.33047	0.24817	1.3316	0.080876	0.33589	0.25259	0.39436	0.32889
419,238	0.33218	0.24775	1.3408	0.080926	0.33828	0.2526	0.36197	0.29886

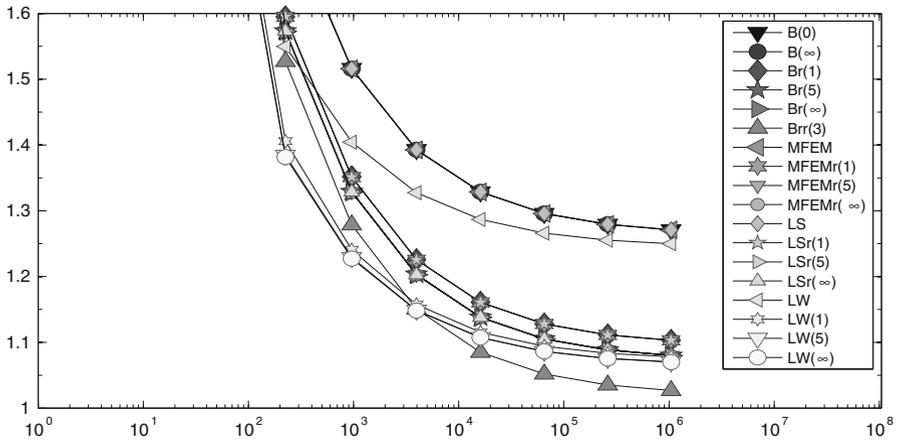


Fig. 4 History of efficiency indices $\eta_{xyz}/\|e\|$ of various a posteriori error estimators η_{xyz} labelled xyz as functions of ndof on uniform meshes in Example 7.2

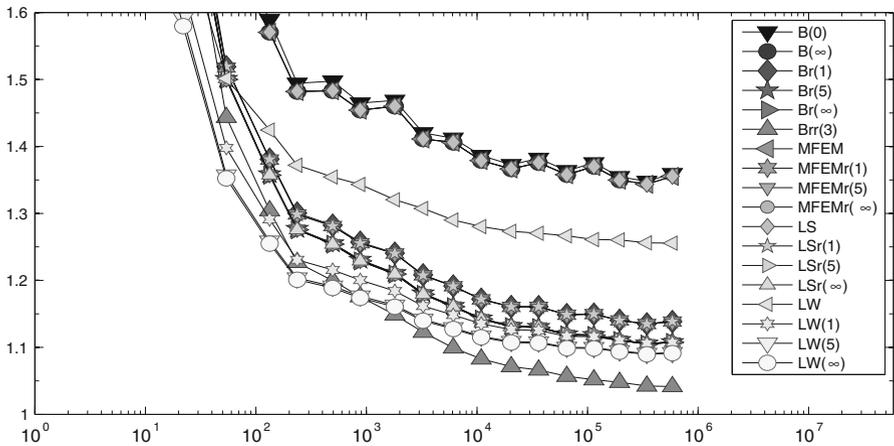


Fig. 5 History of efficiency indices $\eta_{xyz}/\|e\|$ of various a posteriori error estimators η_{xyz} labelled xyz in the figure as functions of ndof n adaptive meshes in Example 7.2

7.2 Square domain with big oscillations

Our second benchmark problem employs homogeneous boundary data $u_D \equiv 0$, $\kappa \equiv 1$, and an oscillating source term f that matches the exact solution

$$u(x, y) = x(x - 1)y(y - 1) \exp(-100(x - 1/2)^2 - 100(y - 117/1,000)^2)$$

on the square domain $\Omega = (0, 1)^2$.

Figures 4 and 5 show efficiency indices for uniform and adaptive mesh refinement. The results are similar to the results from the first example. Since the oscillations dominate the global upper bound on coarse meshes, the efficiency improvement by the

Table 4 Improvement numbers $\varrho_{X,Y,z,r}(k)$ from (6.1) for three postprocessed error estimators and niform mesh-refinement in Example 7.2

ndof	$\varrho_{B,r}(1)$	$\varrho_{B,r}(\infty)$	$\varrho_{B,r}(1)/\varrho_{B,r}(\infty)$	$\varrho_{B,rr}(3)$	$\varrho_{MFEM,r}(1)$	$\varrho_{MFEM,r}(\infty)$	$\varrho_{LW,(1)}$	$\varrho_{LW,(\infty)}$
1	0.9748	0.95736	1.0182	0.96292	0.98371	0.98239	0.98255	0.96298
9	0.92468	0.88492	1.0449	0.89092	0.93634	0.9268	0.94122	0.90648
49	0.90304	0.88813	1.0168	0.86253	0.91751	0.90635	0.86475	0.81815
225	0.74503	0.71101	1.0478	0.64259	0.7479	0.71394	0.69657	0.64799
961	0.63695	0.58873	1.0819	0.48875	0.63614	0.58915	0.55487	0.52132
3,969	0.53462	0.4749	1.1258	0.34425	0.53405	0.47502	0.44614	0.41818
16,129	0.45414	0.38526	1.1788	0.2313	0.45394	0.38528	0.37194	0.34435
65,025	0.40058	0.32538	1.2311	0.15637	0.40052	0.32538	0.32666	0.29821
261,121	0.36899	0.28999	1.2724	0.11226	0.36897	0.29	0.30126	0.27204
1,046,529	0.35172	0.27063	1.2996	0.088152	0.35171	0.27063	0.28775	0.25805

Table 5 Improvement numbers $\varrho_{xyz,r}(k)$ from (6.1) for three postprocessed error estimators and adaptive mesh-refinement in Example 7.2

ndof	$\varrho_{B,r}(1)$	$\varrho_{B,r}(\infty)$	$\varrho_{B,r}(1)/\varrho_{B,r}(\infty)$	$\varrho_{B,r}(3)$	$\varrho_{MFEM,r}(1)$	$\varrho_{MFEM,r}(\infty)$	$\varrho_{LW,r}(1)$	$\varrho_{LW,r}(\infty)$
1	0.9748	0.95736	1.0182	0.96292	0.98371	0.98239	0.98255	0.96298
4	0.92414	0.89274	1.0352	0.88761	0.93137	0.92434	0.94284	0.9168
7	0.83495	0.8115	1.0289	0.76424	0.85102	0.84055	0.92309	0.88212
22	0.76494	0.71928	1.0635	0.68212	0.7716	0.76122	0.85176	0.75915
54	0.66343	0.63443	1.0457	0.54996	0.69846	0.66954	0.76051	0.66032
133	0.59427	0.55157	1.0774	0.45911	0.61778	0.57415	0.64785	0.55811
236	0.56064	0.50878	1.1019	0.40963	0.57324	0.52426	0.58303	0.50115
494	0.51874	0.46085	1.1256	0.35348	0.53449	0.47723	0.57244	0.49521
877	0.50622	0.44556	1.1361	0.33358	0.51675	0.45795	0.54835	0.46968
1,803	0.46589	0.39963	1.1658	0.27543	0.47549	0.40882	0.54082	0.46667
3,249	0.45205	0.38398	1.1773	0.2557	0.46196	0.39399	0.49307	0.42156
6,100	0.42219	0.34937	1.2084	0.2093	0.42889	0.35569	0.47947	0.40742
10,923	0.405	0.33009	1.227	0.18701	0.41345	0.33766	0.45142	0.37982
20,394	0.39309	0.31543	1.2462	0.16656	0.39934	0.32172	0.43282	0.36294
36,113	0.38164	0.30309	1.2591	0.15045	0.38859	0.30907	0.43388	0.36536
64,799	0.37201	0.29213	1.2734	0.13578	0.37875	0.29783	0.408	0.34298
114,059	0.36063	0.27951	1.2902	0.11864	0.36681	0.28436	0.41565	0.35004
195,028	0.35951	0.27856	1.2906	0.11733	0.36512	0.2831	0.39408	0.33252
343,772	0.353	0.27079	1.3036	0.10592	0.3573	0.27459	0.38514	0.32436
585,026	0.34962	0.26707	1.3091	0.099906	0.35388	0.27051	0.39078	0.3303

postprocessing is not that significant in the beginning. On finer meshes, the improvement is as significant as in the first example, also for $k = 1$. In this example there is almost no visible gap between η_{MFEM} , η_{LS} , η_{B} and the postprocessing of η_{B} based on \mathcal{T} . The postprocessing based on $\text{red}(\mathcal{T})$ is almost as good as the postprocessing of η_{LW} based on \mathcal{T}^* . The efficiency indices are reduced by about 20%. The postprocessing $\eta_{\text{Brr}(3)}$ of η_{B} based on two red-refinements $\text{red}^2(\mathcal{T})$ and $k = 3$ iterations leads to higher efficiency. The improvements numbers from Tables 4 and 5 support this observation.

8 Numerical examples for discontinuous coefficients

This section concerns the Poisson model interface problem.

8.1 Setting

Given a right-hand side $f \in L^2(\Omega)$, Dirichlet data $u_D \in H^1(\Omega)$ and piecewise constant diffusion coefficients

$$0 < \underline{\kappa} \leq \kappa(x) \leq \bar{\kappa} < \infty \quad \text{for a.e. } x \in \Omega$$

in the domain Ω , seek $u \in H^1(\Omega)$ such that

$$-\text{div}(\kappa \nabla u) = f \quad \text{in } \Omega \quad \text{and} \quad u = u_D \quad \text{on } \partial\Omega.$$

With $g := \kappa \nabla u_h$ this leads to the usual residual (1.1) with weighted energy norm $\|v\| := \|\kappa^{1/2} \nabla v\|_{L^2(\Omega)}$ and dual norm

$$\|\text{Res}\|_{\star} := \sup_{\varphi \in H_0^1(\Omega)} \text{Res}(\varphi) / \|\varphi\|.$$

The weight κ in the energy norm results in the weight $1/\kappa$ in the a posteriori error estimators $\|\kappa^{-1/2}(q_{\text{xyz}} - g - \text{Curl } v)\|_{L^2(\Omega)}$.

8.2 Square domain

The first benchmark involves $f \equiv 0$ and u_D which match the exact quadratic function $u(x, y) = (x^2 - y^2)/\kappa$ on the square domain $\Omega = (-1, 1)^2$. The diffusion parameter κ assumes the values 1, 100, 10,000 on subdomains as depicted in Fig. 8.

Qualitatively there are no new results compared to the standard Poisson model problem examples. Figures 6 and 7 show a strong improvement by the postprocessings. While the difference of the efficiency indices of η_{LW} and its postprocessed quantities is about 0.25 for uniform and adaptive mesh refinement, there is a surprisingly huge reduction for η_{B} , η_{MFEM} and η_{LS} from 1.5 to 1.15 in case of adaptive mesh refinement!

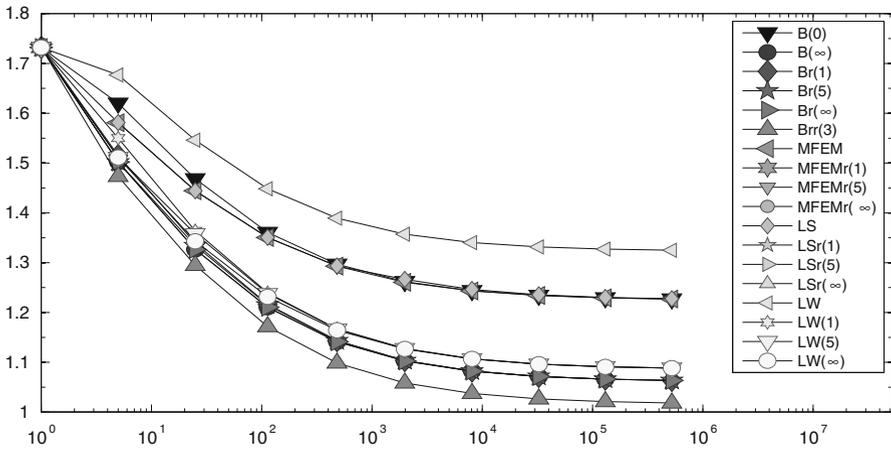


Fig. 6 History of efficiency indices $\eta_{xyz}/\|e\|$ of various a posteriori error estimators η_{xyz} labelled xyz as functions of ndof on uniform meshes in Sect. 8.2

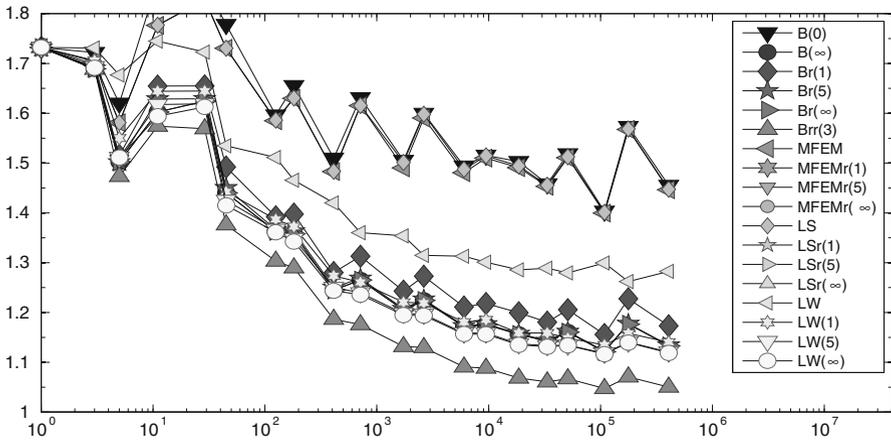


Fig. 7 History of efficiency indices $\eta_{xyz}/\|e\|$ of various a posteriori error estimators η_{xyz} labelled xyz as functions of ndof on adaptive meshes in Sect. 8.2

8.3 Octagon domain

The second benchmark problem from [18] employs $f \equiv 0$ and u_D matching the exact solution $u(x, y) = ((ax^2 - y^2)(ay^2 - x^2))/\kappa$ with $a = \tan((3\pi)/8)^2$ on the octagon domain

$$\Omega = \text{conv}\{(\cos((2j + 1)\pi/8), \sin((2j + 1)\pi/8)), j = 0, 1, \dots, 7\}.$$

The diffusion coefficients κ take alternately the values 1 and 1,000 as depicted in Fig. 8.

Fig. 8 Distribution of $\kappa = 1$ (white) and $\kappa = 1,000$ (lightgray) in octagon domain of Sect. 8.3 (left) and distribution of $\kappa = 1$ (white), $\kappa = 100$ (lightgray) and $\kappa = 10,000$ (darkgray) in square domain of Sect. 8.2

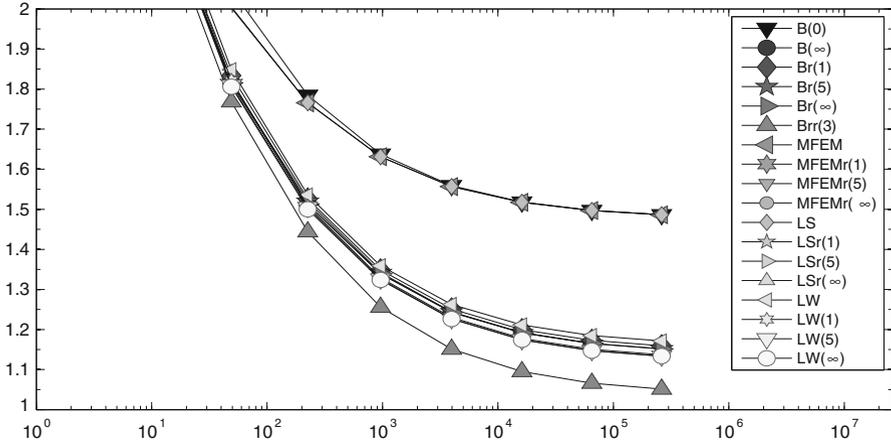
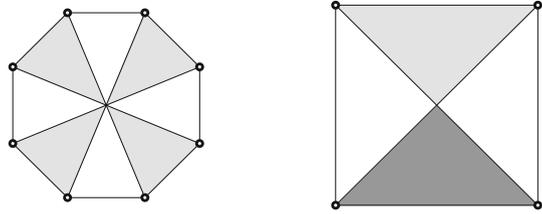


Fig. 9 History of efficiency indices $\eta_{xyz}/\|e\|$ of various a posteriori error estimators η_{xyz} labelled xyz as functions of ndof on uniform meshes in Sect. 8.3

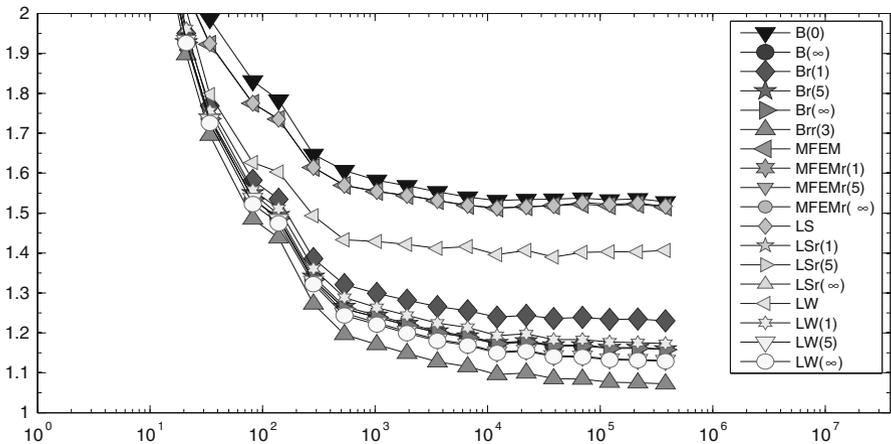


Fig. 10 History of efficiency indices $\eta_{xyz}/\|e\|$ of various a posteriori error estimators η_{xyz} labelled xyz as functions of ndof on adaptive meshes in Sect. 8.3

In contrast to the last experiment, the efficiency of the Luce–Wohlmuth equilibration error estimate does not improve much by the postprocessing in case of uniform refinement as seen in Fig. 9. The reason is that $\eta_{LW(0)}$ yields already very good efficiency of about 1.2, the postprocessing $\eta_{LW(1)}$ arrives at 1.15 together with the postprocessed quantities of the other error estimators based on $\text{red}(\mathcal{T})$. In case of adaptive mesh refinement, Fig. 10 shows a larger scattering similar to former examples. There is even some little gap between η_B and η_{MFEM} .

9 Numerical example for obstacle problems

This section is devoted to conforming obstacle problems with an affine obstacle.

9.1 Setting and global upper bound

The unique exact weak solution $u \in K$ of the obstacle problem inside the closed and convex set of admissible functions,

$$K := \{v \in H^1(\Omega) \mid v = 0 \text{ on } \Gamma_D \text{ and } \chi \leq v \text{ a.e. in } \Omega\} \neq \emptyset$$

satisfies the variational inequality

$$\int_{\Omega} \nabla u \cdot \nabla(u - v) \, dx \leq \int_{\Omega} f(u - v) \, dx \quad \text{for all } v \in K.$$

After [5] and for the particular choice of Λ_h [13], the discrete solution of the obstacle problem u_h in

$$K(\mathcal{T}) := \{v_h \in P_1(\mathcal{T}) \cap C(\Omega) \mid v_h = 0 \text{ on } \Gamma_D \text{ and } \mathcal{I}\chi \leq v_h \text{ in } \Omega\}$$

solves the discrete version of the Poisson problem for $w \in V$ with

$$\int_{\Omega} \nabla w \cdot \nabla v \, dx = \int_{\Omega} (f - \Lambda_h)v \, dx \quad \text{for all } v \in V. \tag{9.1}$$

The energy norm difference $\|w - u_h\| = \|\text{Res}\|_{\star}$ between u_h and the exact solution w of the Poisson problem (9.1) can be estimated by any a posteriori error estimator. In the conforming case $\chi \leq \mathcal{I}\chi$, [13] leads, for any a posteriori estimator η for $\|w - u_h\|$, to the reliable global upper bound (GUB) in the strict sense of

$$\begin{aligned} \|e\| \leq \text{GUB}(\eta) &:= (\eta + \|\Lambda_h - J\Lambda_h\|_{\star})/2 \\ &+ \sqrt{\int_{\Omega} (\chi - u_h)J\Lambda_h \, dx + (\eta + \|\Lambda_h - J\Lambda_h\|_{\star})^2}. \end{aligned}$$

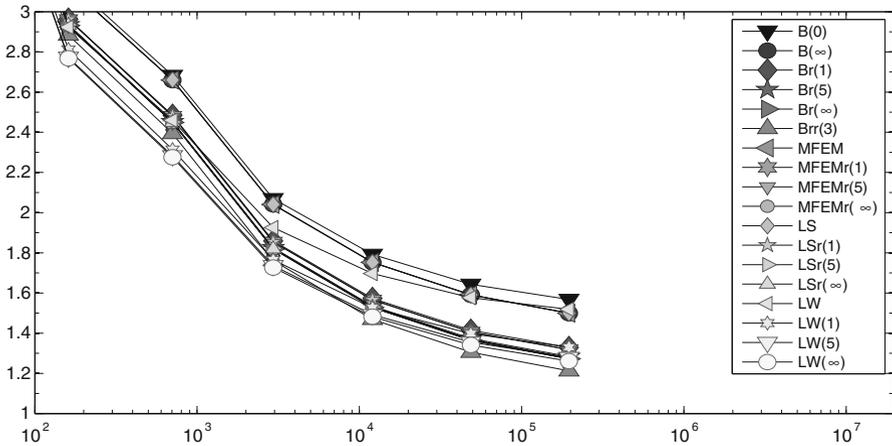


Fig. 11 History of efficiency indices $GUB(\eta_{xyz})/\|e\|$ of various a posteriori error estimators $GUB(\eta_{xyz})$ labelled xyz as functions of $ndof$ on uniform meshes in Sect. 9.2

The patchwise oscillations

$$osc(\Lambda_h, \mathcal{N}) := \left(\sum_{z \in \mathcal{N}} \text{diam}(\omega_z)^2 \min_{f_z \in \mathbb{R}} \|\Lambda_h - f_z\|_{L^2(\omega_z)}^2 \right)^{1/2}$$

are a computable bound for

$$\|\Lambda_h - J\Lambda_h\|_* := \sup_{v \in V \setminus \{0\}} \int_{\Omega} (\Lambda_h - J\Lambda_h)v \, dx / \|v\| \lesssim osc(\Lambda_h, \mathcal{N}).$$

The competition in [13] compares five classes of error estimators.

9.2 Numerical example on L-shaped domain with oscillations and constant obstacle

The benchmark example from [2] mimics a typical corner singularity on the L-shaped domain $\Omega = (-2, 2)^2 \setminus ([0, 2] \times [-2, 0])$ with constant obstacle $\chi = \mathcal{I}\chi \equiv 0$ and homogeneous Dirichlet data $u_D \equiv 0$ along $\partial\Omega$. With the right-hand side

$$f(r, \varphi) := -r^{2/3} \sin(2\varphi/3) \left(\frac{7g'(r)}{3r} + g''(r) \right) - H(r - 5/4),$$

$$g(r) := \max \left\{ 0, \min \left\{ 1, -6s^5 + 15s^4 - 10s^3 + 1 \right\} \right\} \quad \text{for } s := 2(r - 1/4)$$

and the Heaviside function H , the exact solution reads $u(r, \varphi) := r^{2/3}g(r) \sin(2\varphi/3)$.

Figures 11 and 12 compare the efficiency indices for the original equilibration error estimators and for their postprocessed modifications. Also in this nonlinear application,

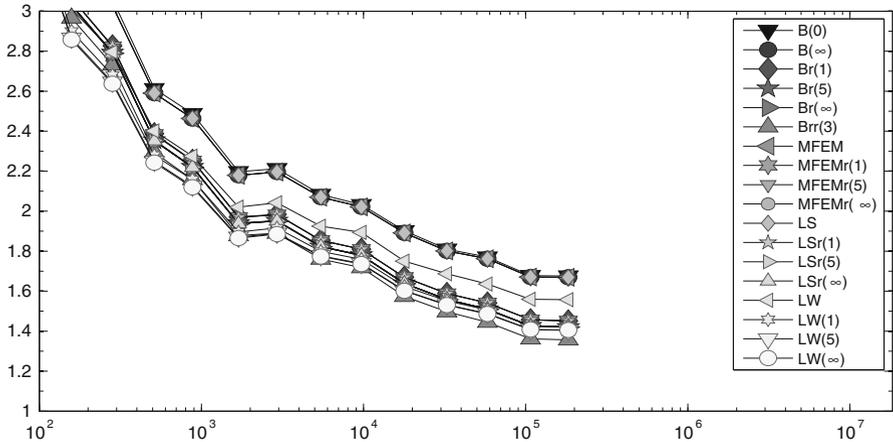


Fig. 12 History of efficiency indices $GUB(\eta_{xyz})/\|e\|$ of various a posteriori error estimators $GUB(\eta_{xyz})$ labelled xyz as functions of $ndof$ on adaptive meshes in Sect. 9.2

there is a significant improvement from efficiency indices about 1.5 to efficiency indices about 1.3 due to the postprocessing.

10 Conclusions

10.1 Braess versus Luce–Wohlmuth

One interpretation of the numerical examples below and our overall experience is that the Braess error control is already very accurate and it is the quality of that simple Braess equilibration which causes the minimal improvements in the numerical experiments for $\widehat{T} = T$. The choice $\widehat{T} = \text{red}(T)$ leads to remarkable improvements throughout all benchmarks. The slight superiority of the Luce–Wohlmuth technique appears the consequence of the dual mesh $\widehat{T} = T^*$ which is refined and hence leads to more space for improvements.

10.2 Computational costs

The striking empirical observation is that already a few iterations dramatically improve the efficiency indices. The more detailed examples in Sect. 7 convey that even a single iteration leads to substantial improvement also for very large numbers of degrees of freedom. The improvement numbers $\varrho_{xyz,r(k)}$ from (6.1) are mostly below 0.5, so the relative error is halved.

If the number of red-refinements is increased, we also suggest to increase the number of cg iterations. In our experiments, $\eta_{Brr(3)}$ employs two red-refinements and three cg iterations. This combination results in efficiency indices very close to 1 in the examples for the linear problems.

10.3 Postprocessing in presence of overhead terms

In the presence of overhead terms in form of oscillations on coarse meshes in Sect. 7.2 or in form of the quantities related to the Lagrange multiplier in the obstacle problem of Sect. 9, the improvement by the postprocessing is less significant. While the oscillations possibly may be reduced by some more elaborate choice of q , the aforementioned other overhead terms are somehow intrinsic and cannot be improved by the postprocessing. In the latter case the effectivity of the postprocessing is limited by the contribution of those overhead terms relative to the global upper bound.

References

1. Ainsworth, M., Oden, J.T.: *A Posteriori Error Estimation in Finite Element Analysis*. Wiley, New York (2000)
2. Bartels, S., Carstensen, C.: Averaging techniques yield reliable a posteriori finite element error control for obstacle problems. *Numer. Math.* **99**(2), 225–249 (2004)
3. Bartels, S., Carstensen, C., Dolzmann, G.: Inhomogeneous Dirichlet conditions in a priori and a posteriori finite element error analysis. *Numer. Math.* **99**(1), 1–24 (2004)
4. Bartels, S., Carstensen, C., Klose, R.: An experimental survey of a posteriori Courant finite element error control for the Poisson equation. *Adv. Comput. Math.* **15**(1–4), 79–106 (2001)
5. Braess, D.: A posteriori error estimators for obstacle problems—another look. *Numer. Math.* **101**, 415–421 (2005)
6. Braess, D.: *Finite Elements—Theory, Fast Solvers, and Applications in Solid Mechanics*. Cambridge University Press, New York (2007)
7. Babuška, I., Strouboulis, T.: *The Finite Element Method and Its Reliability*. Numerical Mathematics and Scientific Computation. The Clarendon Press Oxford University Press, New York (2001)
8. Braess, D., Schöberl, J.: Equilibrated residual error estimator for edge elements. *Math. Comp.* **77**(262), 651–672 (2008)
9. Brenner, S.C., Scott, L.R.: *The Mathematical Theory of Finite Element Methods*, 3rd edn. Springer, New York (2008)
10. Carstensen, C.: A unifying theory of a posteriori finite element error control. *Numer. Math.* **100**(4), 617–637 (2005)
11. Carstensen, C., Eigel, M., Hoppe, R.H.W., Löbhard, C.: A Review of Unified A Posteriori Finite Element Error Control. *Numer. Math. Theor. Meth. Appl.* **5**(4), 509–558 (2010). doi:[10.4208/nmtma.2011.m1032](https://doi.org/10.4208/nmtma.2011.m1032)
12. Carstensen, C., Merdon, C.: Estimator competition for Poisson problems. *J. Comput. Math.* **28**(3), 309–330 (2010)
13. Carstensen, C., Merdon, C.: A posteriori error estimator competition for conforming obstacle problems. *Numer. Methods Partial Differ. Equ.* (2012). doi:[10.1002/num.21728](https://doi.org/10.1002/num.21728)
14. Carstensen, C., Merdon, C.: Computational survey on a posteriori error estimators for nonconforming finite element methods for the Poisson problem (2012, submitted)
15. Carstensen, C., Peterseim, D., Schedensack, M.: Comparison results of three first-order finite element methods for the poisson model problem. *SIAM J. Numer. Anal.* <http://opus4.kobv.de/opus4-matheon/frontdoor/index/index/docId/940> (2012, preprint)
16. Girault, V., Raviart, P.A.: *Finite element methods for Navier Stokes equations*. Springer-Verlag, Berlin (1986)
17. Han, W.: *A Posteriori Error Analysis Via Duality Theory*. Advances in Mechanics and Mathematics. Springer, New York (2005)
18. Hoppe, R.H.W., Wohlmuth, B.: Element-oriented and edge-oriented local error estimators for non-conforming finite element methods. *RAIRO Modél. Math. Anal. Numér.* **30**(2), 237–263 (1996). MR 1382112 (97e:65124)
19. Luce, R., Wohlmuth, B.I.: A local a posteriori error estimator based on equilibrated fluxes. *SIAM J. Numer. Anal.* **42**(4), 1394–1414 (2004)

20. Marini, L.D.: An inexpensive method for the evaluation of the solution of the lowest order Raviart–Thomas mixed method. *SIAM J. Numer. Anal.* **22**, 493–496 (1985)
21. Mommer, M.S., Stevenson, R.: A goal-oriented adaptive finite element method with convergence rates. *SIAM J. Numer. Anal.* **47**(2), 861–886 (2009)
22. Prager, W., Synge, J. L.: Approximations in elasticity based on the concept of function space. *Quart. Appl. Math.* **5**, 241–269 (1947). MR 0025902 (10,81b)
23. Payne, L.E., Weinberger, H.F.: An optimal Poincaré inequality for convex domains. *Arch. Rat. Mech. Anal.* **5**, 286–292 (1960)
24. Repin, S.I.: Two-sided estimates of deviation from exact solutions of uniformly elliptic equations. In: *Proceedings of the St. Petersburg Mathematical Society, Vol. IX (Providence, RI), American Mathematical Society. Translations: Series 2, vol. 209. American Mathematical Society, pp. 143–171 (2003). MR 2018375 (2004k:35081)*
25. Repin, S.: *A Posteriori Estimates for Partial Differential Equations. Radon Series on Computational and Applied Mathematics. Walter de Gruyter, Berlin (2008)*
26. Repin, S., Sauter, S., Smolianski, A.: A posteriori error estimation for the Dirichlet problem with account of the error in the approximation of boundary conditions. *Computing* **70**(3), 205–233 (2003)
27. Valdman, J.: Minimization of functional majorant in a posteriori error analysis based on H(div) multigrid-preconditioned CG method. *Adv. Numer. Anal.* (2009). Art ID 164519, MR 2739760 (2011j:65290)
28. Verfürth, R.: *A Review of A-Posteriori Error Estimation and Adaptive Mesh Refinement Techniques. Wiley-Teubner, Amsterdam (1996)*
29. Veese, A., Verfürth, R.: Explicit upper bounds for dual norms of residuals. *SIAM J. Numer. Anal.* **47**(3), 2387–2405 (2009)