European Society of Computational Methods in Sciences and Engineering (ESCMSE)

Journal of Numerical Analysis, Industrial and Applied Mathematics (JNAIAM)
vol. 2, no. 3-4, 2007, pp. 157-176
ISSN 1790–8140

Computational Technologies for Reliable Control of Global and Local Errors for Linear Elliptic Type Boundary Value Problems[1]

A. Hannukainen, S. Korotov[2]

Institute of Mathematics,
Helsinki University of Technology,
P.O. Box 1100, FI-02015 TKK, Finland

Received January 12, 2006; accepted in revised form April 17, 2007

Abstract: The paper is devoted to the problem of reliable control of accuracy of approximate solutions obtained in computer simulations. This task is strongly related to the so-called a posteriori error estimates, giving computable bounds for computational errors and detecting zones in the solution domain, where such errors are too large and certain mesh refinements should be performed. A mathematical model described by a linear elliptic equation with mixed boundary conditions is considered. We derive in a simple way two-sided (upper and lower) easily computable estimates for global (in terms of the energy norm) and local (in terms of linear functionals) control of the computational error, understood as the deviation between the exact solution of the model and the approximation. Such two-sided estimates are completely independent of the numerical technique used to obtain approximations and can be made as close to the true errors as resources of a concrete computer used for computations allow. Main issues of practical realization of the estimation procedures proposed are discussed, several tests and numerical comparison with other popular error estimates based on the gradient averaging are presented.

Keywords: a posteriori error estimation, error control in energy norm, error control in terms of linear functionals, differential equation of elliptic type, mixed boundary conditions, gradient averaging

Mathematics Subject Classification: 65N15, 65N30

1 Introduction

Many physical and mechanical phenomena can be described by means of mathematical models presenting boundary value problems of elliptic type [7, 10, 17]. Various numerical techniques (the finite element method (FEM), the finite difference method, the finite volume method, etc) are well developed for finding approximate solutions of such problems. However, in order to be

[1]Published electronically October 27, 2007
[2]Corresponding author. E-mail: sergey.korotov@hut.fi

practically meaningful, computer simulations always require an accuracy verification of computed approximations. Such a verification is the main purpose of a posteriori error estimation methods.

In the present paper, we first recall two different ways of measuring the computational error, which is understood as the deviation $u - \bar{u}$ between the exact solution u and its approximation $\bar{u}$, in the global (energy) norm and in terms of linear bounded functionals. These two ways of measurement (and also control – via a posteriori error estimation procedures) of the error are very natural and commonly used nowadays in both mathematical and engineering communities. The global error estimation normally gives a general presentation on the quality of approximation and a stopping criterion to terminate the calculations [2, 3, 4, 5, 15, 18, 19, 24, 31, 33]. However, practitioners are often interested not only in the value of the overall error, but also in errors over certain critical (and usually local) parts of the solution domain (for example, in fracture mechanics – see [27, 28, 29] and references therein). This reason initiated another trend in a posteriori error estimation which is based on the concept of control of the computational error locally. One common way to perform such a control is to introduce a suitable linear functional ℓ related to subdomain of interest and to construct a posteriori computable estimate for $\ell(u - \bar{u})$, see [5, 6, 11, 13, 14, 16, 23].

Further, on the base of a model elliptic (reaction-diffusion) problem with mixed boundary conditions, we present relatively simple technologies for obtaining *computable guaranteed two-sided (upper and lower) estimates* needed for reliable error control in both global (in the energy norm) and local (in terms of linear functionals) ways. The estimates derived are valid for *any conforming approximations* independently of numerical methods used to obtain them, and can be made *arbitrarily close* to the true errors. In real-life calculations this closeness only depends on resources of a concrete computer used. The estimates proposed naturally generalize results of [23, 24, 25, 26] obtained by another techniques and for simpler problems. We shall discuss main issues of a practical realization of the error control procedures, present several tests and make a numerical comparison with another popular error estimation techniques based on gradient averaging.

It is worth to mention here that most of the other estimates available (see, e.g., [1, 30, 32] dealing with the same type of models – diffusion-reaction) strongly rely on the fact that computed solutions are true finite element (FE) approximations which, in fact, hardly happens in real computations, e.g., due to quadrature rules, forcibly stopped iterative processes, various round-off errors, or even possible bugs in FE codes.

2 Formulation of Problem

2.1 Model problem

We introduce the model elliptic problem to be considered, which consists of the governing (reaction-diffusion) equation (1) and mixed (Dirichlet/Neumann) boundary conditions (2)–(3): Find a function u such that

$$-\mathrm{div}(A\nabla u) + cu = f \quad \text{in} \quad \Omega, \tag{1}$$

$$u = u_0 \quad \text{on} \quad \Gamma_D, \tag{2}$$

$$\nu^T \cdot A\nabla u = g \quad \text{on} \quad \Gamma_N, \tag{3}$$

where Ω is a bounded domain in $\mathbf{R}^d$ with Lipschitz continuous boundary $\partial\Omega$, such that $\overline{\partial\Omega} = \overline{\Gamma}_D \cup \overline{\Gamma}_N$, $\Gamma_D \cap \Gamma_N = \emptyset$, $\mathrm{meas}_{d-1}\Gamma_D > 0$, ν is the outward normal to the boundary, $f \in L_2(\Omega)$, $u_0 \in H^1(\Omega)$, $g \in L_2(\Gamma_N)$, $c \in L_\infty(\Omega)$, the matrix of coefficients A is symmetric with entries $a_{ij} \in L_\infty(\Omega)$, $i, j = 1, ..., d$, and is such that

$$C_2|\xi|^2 \geq A(x)\xi \cdot \xi \geq C_1|\xi|^2 \qquad \forall \xi \in \mathbf{R}^d \quad \text{a. e. in } \Omega, \tag{4}$$

where $C_1, C_2 > 0$. In addition, we assume that almost everywhere in Ω

$$c \geq 0, \tag{5}$$

introduce the set

$$\Omega^c := \{x \in \Omega \mid c(x) \geq C_0 > 0\}, \tag{6}$$

where C_0 is a fixed (positive) constant, and suppose that $c = 0$ almost everywhere in $\Omega \setminus \overline{\Omega^c}$. The set Ω^c can be, in particular, empty or equal to Ω.

It is a common practice to pose problem (1)–(3) in the so-called weak form: Find $u \in u_0 + H^1_{\Gamma_D}(\Omega)$ such that

$$\int_\Omega A\nabla u \cdot \nabla w \, dx + \int_\Omega cuw \, dx = \int_\Omega fw \, dx + \int_{\Gamma_N} gw \, ds \qquad \forall w \in H^1_{\Gamma_D}(\Omega), \tag{7}$$

where

$$H^1_{\Gamma_D}(\Omega) := \{v \in H^1(\Omega) \mid v = 0 \text{ on } \Gamma_D\}. \tag{8}$$

If we define the bilinear form $a(\cdot,\cdot)$ and linear form $F(\cdot)$ as follows:

$$a(v,w) := \int_\Omega A\nabla v \cdot \nabla w \, dx + \int_\Omega cvw \, dx, \quad v, w \in H^1(\Omega), \tag{9}$$

$$F(w) := \int_\Omega fw \, dx + \int_{\Gamma_N} gw \, ds, \quad w \in H^1(\Omega), \tag{10}$$

then the weak formulation (7) can be written in a short form:

Find $u \in u_0 + H^1_{\Gamma_D}(\Omega)$ such that $a(u,w) = F(w) \quad \forall w \in H^1_{\Gamma_D}(\Omega)$.

The weak solution defined by (7) exists and is unique in view of the well-known Lax-Milgram lemma. Further, the so-called *energy functional* J of problem (7) is defined as follows

$$J(w) := \frac{1}{2} a(w,w) - \bar{F}(w), \qquad w \in H^1_{\Gamma_D}(\Omega), \tag{11}$$

where $\bar{F}(w) := F(w) - a(u_0, w)$, and the corresponding *energy norm* is defined as $\sqrt{a(\cdot,\cdot)}$.

2.2 Types of error control

Let $\bar{u}$ be *any function* from the set $u_0 + H^1_{\Gamma_D}(\Omega)$ (e.g., computed by some numerical method or be some postprocessed solution) considered as an approximation of u. It is natural to measure the overall accuracy of the approximation $\bar{u}$ in terms of the above-defined energy norm. Thus, our first goal is to construct reliable and easily computable two-sided estimates for controlling the following value:

$$a(u-\bar{u}, u-\bar{u}) = \int_\Omega A\nabla(u-\bar{u}) \cdot \nabla(u-\bar{u}) \, dx + \int_\Omega c(u-\bar{u})^2 \, dx. \tag{12}$$

The second type of error control considered in this paper is the two-sided estimation of the value of the deviation $u - \bar{u}$ in terms of some bounded linear functional ℓ:

$$\ell(u-\bar{u}). \tag{13}$$

Remark 2.1 If the functional ℓ in (13) is defined as some integral over a small subdomain (or line/surface) in $\overline{\Omega}$, then the reliable estimation of $\ell(u-\bar{u})$ helps to control the behaviour of the error $u-\bar{u}$ locally in that subdomain (or over the line/surface). For example, one can be interested in the estimation of

$$\ell(u-\bar{u}) = \int_S \varphi(u-\bar{u})\,dx$$

with S being a subdomain in Ω or a line/surface in Γ_N (where the solution is also unknown) and φ being some weight-function chosen to fit certain computational (or modelling) goal.

Remark 2.2 It is clear that the existence of an estimate for (13) also allows to estimate of the value $\ell(u)$ (often called *"quantity of interest"* or *"goal-oriented quantity"* [2]). Really, $\ell(u) = \ell(u-\bar{u})+\ell(\bar{u})$ where $\ell(\bar{u})$ is computable and $\ell(u-\bar{u})$ is estimated. The value $\ell(u)$ can be sometimes more important to know than the solution u itself (cf. [20, Chapt. VII] and also [27, 28, 29]).

2.3 Inequalities and global constants

In what follows we shall need the Friedrichs-Poincaré inequality

$$\|w\|_{0,\Omega} \leq C_{\Omega,\Gamma_D}\|\nabla w\|_{0,\Omega} \quad \forall w \in H^1_{\Gamma_D}(\Omega), \tag{14}$$

and the inequality in the trace theorem

$$\|w\|_{0,\partial\Omega} \leq C_{\partial\Omega}\|w\|_{1,\Omega} \quad \forall w \in H^1(\Omega), \tag{15}$$

where C_{Ω,Γ_D} and $C_{\partial\Omega}$ are positive constants, depending solely on Ω, Γ_D, and $\partial\Omega$. Here the notation, $\|\cdot\|_{0,\Omega}$ and $\|\cdot\|_{1,\Omega}$, stands for the standard norms in $L_2(\Omega)$ (or $L_2(\Omega,\mathbf{R}^d)$) and $H^1(\Omega)$, respectively. The symbol $\|\cdot\|_{0,\partial\Omega}$ means the norm in $L_2(\partial\Omega)$. Proofs of inequalities (14) and (15) can be found, e.g., in [22].

3 Two-Sided Estimates of Error in Energy Norm

In this section we shall employ the notation χ_S for the characteristic function of a set S, i.e., $\chi_S(x) = 1$ if $x \in S$, and $\chi_S(x) = 0$ if $x \notin S$, and also use the notation

$$|\!|\!| y |\!|\!|_\Omega := \Big(\int_\Omega Ay\cdot y\,dx\Big)^{1/2}$$

for $y \in L_2(\Omega,\mathbf{R}^d)$.

3.1 Upper estimate

Theorem 3.1 For the error in the energy norm (12) we have the following upper estimate:

$$a(u-\bar{u},u-\bar{u}) \leq \|\frac{1}{\sqrt{c}}(f+\operatorname{div}y^*-c\bar{u})\|^2_{0,\Omega^c}+$$

$$+(1+\alpha)|\!|\!|A^{-1}y^*-\nabla\bar{u}|\!|\!|^2_\Omega+(1+\frac{1}{\alpha})(1+\beta)\frac{C^2_{\Omega,\Gamma_D}}{C_1}\|f+\operatorname{div}y^*\|^2_{0,\Omega\setminus\overline{\Omega}^c} \tag{16}$$

$$+(1+\frac{1}{\alpha})(1+\frac{1}{\beta})C^2_{\Omega;\partial\Omega}\|g-\nu^T\cdot y^*\|^2_{0,\Gamma_N},$$

where α and β are arbitrary positive real numbers, y^* is any function from

$$H_N(\Omega, \mathrm{div}) := \{y \in L_2(\Omega, \mathbf{R}^d) \,|\, \mathrm{div} y \in L_2(\Omega),\ \nu^T \cdot y \in L_2(\Gamma_N)\},$$

and the constant

$$C_{\Omega;\partial\Omega} := \frac{C_{\partial\Omega}\sqrt{1 + C^2_{\Omega,\Gamma_D}}}{\sqrt{C_1}}.$$

P r o o f : First of all, we notice that it actually holds (cf. (6))

$$a(u - \bar{u}, u - \bar{u}) = \|\!|\nabla(u - \bar{u})|\!\|^2_{\Omega} + \|\sqrt{c}(u - \bar{u})\|^2_{0,\Omega^c}. \tag{17}$$

Further, using the fact that $u - \bar{u} \in H^1_{\Gamma_D}(\Omega)$, the integral identity (7) with $w = u - \bar{u}$, the Green formula, and simple regrouping of terms, we observe that

$$\begin{aligned}
a(u - \bar{u}, u - \bar{u}) &= \int\limits_{\Omega} f(u - \bar{u})dx + \int\limits_{\Gamma_N} g(u - \bar{u})\, ds - \int\limits_{\Omega} A\nabla\bar{u} \cdot \nabla\,(u - \bar{u})\, dx \\
&\quad - \int\limits_{\Omega} c\bar{u}(u - \bar{u})\, dx = \int\limits_{\Omega} (f - c\bar{u})(u - \bar{u})\ dx + \int\limits_{\Gamma_N} g(u - \bar{u})\, ds \\
&\quad - \int\limits_{\Omega} (A\nabla\bar{u} - y^*) \cdot \nabla(u - \bar{u})\ dx - \int\limits_{\Omega} y^* \cdot \nabla(u - \bar{u})\ dx = \qquad (18)\\
&= \int\limits_{\Omega} (f + \mathrm{div}\, y^* - c\bar{u})(u - \bar{u})\ dx - \int\limits_{\Omega} A(\nabla\bar{u} - A^{-1}y^*) \cdot \nabla(u - \bar{u})\ dx \\
&\quad + \int\limits_{\Gamma_N} g(u - \bar{u})\, ds - \int\limits_{\Gamma_N} \nu^T \cdot y^*(u - \bar{u})\, ds = \\
&= \int\limits_{\Omega} A(A^{-1}y^* - \nabla\bar{u}) \cdot \nabla(u - \bar{u})\ dx + \int\limits_{\Omega} (f + \mathrm{div}\, y^* - c\bar{u})(u - \bar{u})\ dx \\
&\quad + \int\limits_{\Gamma_N} (g - \nu^T \cdot y^*)(u - \bar{u})\, ds,
\end{aligned}$$

where y^* is any function from the space $H_N(\Omega, \mathrm{div})$ defined in the conditions of the theorem.

Now, the right-hand side (RHS) of equality (18) can be estimated, using the Cauchy-Schwarz inequality, notation (6), and the trace inequality (15), from above as follows

$$\begin{aligned}
\text{RHS of (18)} &\leq \|\!|A^{-1}y^* - \nabla\bar{u}|\!\|_{\Omega}\, \|\!|\nabla(u - \bar{u})|\!\|_{\Omega} + \|g - \nu^T \cdot y^*\|_{0,\Gamma_N} \|u - \bar{u}\|_{0,\Gamma_N} \\
&\quad + \int\limits_{\Omega} (f + \mathrm{div}\, y^* - c\bar{u})(u - \bar{u})\ dx \leq \\
&\leq \|\!|A^{-1}y^* - \nabla\bar{u}|\!\|_{\Omega}\, \|\!|\nabla(u - \bar{u})|\!\|_{\Omega} + \|g - \nu^T \cdot y^*\|_{0,\Gamma_N} C_{\partial\Omega} \|u - \bar{u}\|_{1,\Omega} \qquad (19)\\
&\quad + \int\limits_{\Omega^c} \frac{1}{\sqrt{c}}(f + \mathrm{div}\, y^* - c\bar{u})\, \sqrt{c}(u - \bar{u})\ dx + \int\limits_{\Omega \backslash \overline{\Omega}^c} (f + \mathrm{div}\, y^* - c\bar{u})\,(u - \bar{u})\ dx.
\end{aligned}$$

Further, using the ellipticity condition (4), the Friedrichs inequality (14), and the inequality

$$|a\,b| \leq \frac{1}{2}a^2 + \frac{1}{2}b^2, \tag{20}$$

we observe that

$$\text{RHS of (19)} \leq \Big(\| A^{-1}y^* - \nabla \bar{u}) \|_\Omega + \frac{C_{\partial\Omega}\sqrt{1 + C^2_{\Omega,\Gamma_D}}}{\sqrt{C_1}} \|g - \nu^T \cdot y^*\|_{0,\Gamma_N} \Big) \| \nabla(u - \bar{u}) \|_\Omega$$

$$+\frac{1}{2}\|\sqrt{c}(u-\bar{u})\|^2_{0,\Omega^c} + \frac{1}{2}\|\frac{1}{\sqrt{c}}(f + \operatorname{div} y^* - c\bar{u})\|^2_{0,\Omega^c} + \int\limits_\Omega \chi_{\Omega\setminus\overline{\Omega}^c}(f + \operatorname{div} y^* - c\bar{u})\,(u - \bar{u})\;dx \leq$$

$$\leq \Big(\| A^{-1}y^* - \nabla \bar{u}) \|_\Omega + C_{\Omega;\partial\Omega}\|g - \nu^T \cdot y^*\|_{0,\Gamma_N} \Big) \| \nabla(u - \bar{u}) \|_\Omega \tag{21}$$

$$+\frac{1}{2}\|\sqrt{c}(u-\bar{u})\|^2_{0,\Omega^c} + \frac{1}{2}\|\frac{1}{\sqrt{c}}(f + \operatorname{div} y^* - c\bar{u})\|^2_{0,\Omega^c} + \|\chi_{\Omega\setminus\overline{\Omega}^c}(f + \operatorname{div} y^* - c\bar{u})\|_{0,\Omega}\,\|u - \bar{u}\|_{0,\Omega}.$$

Regrouping terms in RHS of (21) and using again the inequality (20), we get an estimate

$$\text{RHS of (21)} \leq \Big(\| A^{-1}y^* - \nabla \bar{u}) \|_\Omega + C_{\Omega;\partial\Omega}\|g - \nu^T \cdot y^*\|_{0,\Gamma_N} + \frac{C_{\Omega,\Gamma_D}}{\sqrt{C_1}}\|f + \operatorname{div} y^* - c\bar{u}\|_{0,\Omega\setminus\overline{\Omega}^c} \Big) \times$$

$$\times \| \nabla(u - \bar{u}) \|_\Omega + \frac{1}{2}\|\sqrt{c}(u-\bar{u})\|^2_{0,\Omega^c} + \frac{1}{2}\|\frac{1}{\sqrt{c}}(f + \operatorname{div} y^* - c\bar{u})\|^2_{0,\Omega^c} \leq$$

$$\leq \frac{1}{2}\Big(\| A^{-1}y^* - \nabla \bar{u}) \|_\Omega + C_{\Omega;\partial\Omega}\|g - \nu^T \cdot y^*\|_{0,\Gamma_N} + \frac{C_{\Omega,\Gamma_D}}{\sqrt{C_1}}\|f + \operatorname{div} y^* - c\bar{u}\|_{0,\Omega\setminus\overline{\Omega}^c} \Big)^2 \tag{22}$$

$$+\frac{1}{2}\| \nabla(u - \bar{u}) \|^2_\Omega + \frac{1}{2}\|\sqrt{c}(u-\bar{u})\|^2_{0,\Omega^c} + \frac{1}{2}\|\frac{1}{\sqrt{c}}(f + \operatorname{div} y^* - c\bar{u})\|^2_{0,\Omega^c}.$$

Using now (17) and the final inequality resulting from (18)–(19) and (21)–(22), multiplying it by two and regrouping, we immediately get for the error in the energy norm that

$$a(u - \bar{u}, u - \bar{u}) = \| \nabla(u - \bar{u}) \|^2_\Omega + \|\sqrt{c}(u-\bar{u})\|^2_{0,\Omega^c} \leq \|\frac{1}{\sqrt{c}}(f + \operatorname{div} y^* - c\bar{u})\|^2_{0,\Omega^c}$$

$$+\Big(\| A^{-1}y^* - \nabla \bar{u} \|_\Omega + \frac{C_{\Omega,\Gamma_D}}{\sqrt{C_1}}\|f + \operatorname{div} y^*\|_{0,\Omega\setminus\overline{\Omega}^c} + C_{\Omega;\partial\Omega}\|g - \nu^T \cdot y^*\|_{0,\Gamma_N} \Big)^2. \tag{23}$$

Finally, using twice the inequality $(a + b)^2 \leq (1 + \lambda)a^2 + (1 + \frac{1}{\lambda})b^2$ $(\lambda > 0)$ for the terms in the round brackets in the left-hand side of (23), we get estimate (16). □

Remark 3.1 We note here that in spite of introducing two extra free parameters α and β in order to get (16) from (23), the upper estimate in the form (16) seems to be more suitable for practical purposes. Indeed, for fixed α and β it presents a quadratic functional with respect to the parameter y^*, and the problem of minimization of quadratic functionals is well studied and is a quite standard task in numerical analysis.

3.2 Lower estimate

We introduce the following notations:

$$u_\star := u - u_0, \quad \bar{u}_\star := \bar{u} - u_0,$$

from which it is clear that both $u_\star$ and $\bar{u}_\star$ belong to $H^1_{\Gamma_D}(\Omega)$.

Theorem 3.2 For the error in the energy norm (12) we have the following lower bound

$$a(u-\bar{u}, u-\bar{u}) \geq 2(J(\bar{u}_\star) - J(w_\star)), \tag{24}$$

where $w_\star$ is any function from $H^1_{\Gamma_D}(\Omega)$ and the functional J is defined in (11).

P r o o f : First, we prove that (cf. [20, p. 333])

$$a(u-\bar{u}, u-\bar{u}) = 2(J(\bar{u}_\star) - J(u_\star)). \tag{25}$$

Indeed, we have

$$2(J(\bar{u}_\star) - J(u_\star)) = a(\bar{u}_\star, \bar{u}_\star) - 2\bar{F}(\bar{u}_\star) - a(u_\star, u_\star) + 2\bar{F}(u_\star)$$

$$= a(\bar{u}_\star, \bar{u}_\star) - a(u_\star, u_\star) + 2\bar{F}(u_\star - \bar{u}_\star) = a(\bar{u}_\star, \bar{u}_\star) - a(u_\star, u_\star) + 2a(u_\star, u_\star - \bar{u}_\star)$$

$$= a(\bar{u}_\star, \bar{u}_\star) + a(u_\star, u_\star) - 2a(u_\star, \bar{u}_\star) = a(u_\star - \bar{u}_\star, u_\star - \bar{u}_\star) = a(u-\bar{u}, u-\bar{u}).$$

Since $u_\star$, in fact, minimizes the energy functional, we have $J(u_\star) \leq J(w_\star)$ for any $w_\star$ from $H^1_{\Gamma_D}(\Omega)$, which immediately proves (24). □

3.3 Comments on two-sided estimates (16) and (24)

- In order to derive the upper (16) and the lower (24) estimates, we did not specify the function $\bar{u}$ to be a finite element approximation (or computed by some another numerical method). In fact, it is simply any function from the set $u_0 + H^1_{\Gamma_D}(\Omega)$.

- The upper estimate (16) is sharp. Really, if one takes $y^* = A\nabla u$, which obviously belongs to $H_N(\Omega, \mathrm{div})$, then the last two terms in the right-hand side of (16) vanish. Further, taking $\alpha = 0$, we finally observe that the inequality (16) holds as equality. To prove that the lower estimate (24) is sharp too, we should, obviously, take $w = u_\star \in H^1_{\Gamma_D}(\Omega)$ and use (25).

- The upper estimate (16) contains only two global constants, C_{Ω,Γ_D} and $C_{\partial\Omega}$, which do not depend on the computational process. They have to be computed (or accurately estimated from above) only once when the problem is posed.

- In many works, devoted to a posteriori error estimation, one often considers diffusion equation only, i.e., $c \equiv 0$. In this case $a(u-\bar{u}, u-\bar{u}) = \|\nabla(u-\bar{u})\|^2_\Omega$, the set $\Omega^c = \emptyset$, and the estimate (16) takes a simpler form

$$a(u-\bar{u}, u-\bar{u}) \leq (1+\alpha)\|A^{-1}y^* - \nabla\bar{u}\|^2_\Omega + (1+\frac{1}{\alpha})(1+\beta)\frac{C^2_{\Omega,\Gamma_D}}{C_1}\|f + \mathrm{div}\, y^*\|^2_{0,\Omega}$$

$$+(1+\frac{1}{\alpha})(1+\frac{1}{\beta})C^2_{\Omega;\partial\Omega}\|g - \nu^T \cdot y^*\|^2_{0,\Gamma_N}. \tag{26}$$

- For the pure Dirichlet boundary condition, the third term in RHS of (26) does not exist, and, since the estimate is valid for any positive β, we can take it be zero. Then, we get the estimate

$$a(u-\bar{u},u-\bar{u}) \leq (1+\alpha)|\!|\!|A^{-1}y^* - \nabla\bar{u}|\!|\!|^2_{\Omega} + (1+\frac{1}{\alpha})\frac{C^2_{\Omega,\Gamma_D}}{C_1}\|f+\operatorname{div} y^*\|^2_{0,\Omega}. \tag{27}$$

- The upper estimate (27) was first obtained in [24] using quite complicated tools of the duality theory, and later it was obtained in [25] for the Poisson equation, using an idea of the Helmholtz decomposition of $L_2(\Omega,\mathbf{R}^d)$. The estimate (26) was derived in [26] using the duality theory again. Our approach of derivation of the estimates is different from those used in the above mentioned works, is applied to more general situations, and seems simpler.

- In the case of pure Dirichlet condition we have to compute, or estimate from above, only one constant C_{Ω,Γ_D}.

- In the case of pure Dirichlet condition and if $c(x) \geq C_0 > 0$ almost everywhere in Ω, no computation/estimation of any constants is needed at all.

- We note that the lower estimate (24) is different from that one presented in [23] and seems to be more natural for the considered type of problems.

In what follows we shall use the following notations for the upper and lower bounds of the error in the energy norm (12):

$$\begin{gathered} M^{\oplus}(\bar{u},y^*,\alpha,\beta) := \|\frac{1}{\sqrt{c}}(f+\operatorname{div} y^* - c\bar{u})\|^2_{0,\Omega^c} + \\ +(1+\alpha)|\!|\!|A^{-1}y^* - \nabla\bar{u}|\!|\!|^2_{\Omega} + (1+\frac{1}{\alpha})(1+\beta)\frac{C^2_{\Omega,\Gamma_D}}{C_1}\|f+\operatorname{div} y^*\|^2_{0,\Omega\setminus\overline{\Omega}^c} \\ +(1+\frac{1}{\alpha})(1+\frac{1}{\beta})C^2_{\Omega;\partial\Omega}\|g-\nu^T\cdot y^*\|^2_{0,\Gamma_N}, \end{gathered} \tag{28}$$

and

$$M^{\ominus}(\bar{u},w_\star) := 2(J(\bar{u}_\star) - J(w_\star)). \tag{29}$$

Sometimes we shall use only a short notation $M^{\oplus}$ or $M^{\ominus}$ (possibly with some subindices) for the error bounds if it does not lead to misunderstanding.

4 Two-Sided Estimates for Local Errors

Two-sided estimates for controlling the error $u-\bar{u}$ in terms of the bounded linear functional ℓ are essentially based on the usage of an auxiliary (often called *adjoint*) problem formulated below.

Adjoint Problem: Find $v \in H^1_{\Gamma_D}(\Omega)$ such that

$$\int_{\Omega} A\nabla v\cdot\nabla w\,dx + \int_{\Omega} c\,vw\,dx = \ell(w) \quad \forall w \in H^1_{\Gamma_D}(\Omega). \tag{30}$$

The adjoint problem can be rewritten in a shorter form similarly to the main problem (7):

Find $v \in H^1_{\Gamma_D}(\Omega)$ such that $a(v,w) = \ell(w) \quad \forall w \in H^1_{\Gamma_D}(\Omega)$.

The adjoint problem is uniquely solvable due to the assumption that ℓ is a linear bounded functional and the Lax-Milgram lemma. However, the exact solution v of it is usually very hard (or even impossible) to find in analytical form and, thus, we can only have some approximation for v, which we denote by the symbol $\bar{v}$ in what follows, assuming only that $\bar{v} \in H^1_{\Gamma_D}(\Omega)$. The following results hold (cf. [11, 16, 23]).

Theorem 4.1 We have

$$\ell(u-\bar{u}) = E_0(\bar{u},\bar{v}) + E_1(u-\bar{u}, v-\bar{v}), \tag{31}$$

where

$$E_0(\bar{u},\bar{v}) = F(\bar{v}) - a(\bar{u},\bar{v}), \tag{32}$$

$$E_1(u-\bar{u}, v-\bar{v}) = a(u-\bar{u}, v-\bar{v}). \tag{33}$$

P r o o f : In view of integral identities (30) and (7), and using the fact that $u-\bar{u} \in H^1_{\Gamma_D}(\Omega)$, we observe that

$$\begin{aligned}\ell(u-\bar{u}) = a(v, u-\bar{u}) = a(v-\bar{v}, u-\bar{u}) + a(\bar{v}, u-\bar{u}) = \\ = E_1(u-\bar{u}, v-\bar{v}) + a(\bar{v}, u) - a(\bar{v},\bar{u}) = E_1(u-\bar{u}, v-\bar{v}) + F(\bar{v}) - a(\bar{v},\bar{u}) = \\ = E_0(\bar{u},\bar{v}) + E_1(u-\bar{u}, v-\bar{v}). \qquad \square\end{aligned}$$

The first term E_0 is, obviously, directly computable once we have $\bar{u}$ and $\bar{v}$ computed, but the term E_1 contains unknown functions u and v, and their unknown gradients ∇u and ∇v. In order to estimate this term , we notice first that $E_1(u-\bar{u}, v-\bar{v}) \equiv a(u-\bar{u}, v-\bar{v})$. Further, the following relation obviously holds for any positive real number γ:

$$\begin{aligned}2E_1(u-\bar{u}, v-\bar{v}) = a(\gamma(u-\bar{u}) + \frac{1}{\gamma}(v-\bar{v}), \gamma(u-\bar{u}) + \frac{1}{\gamma}(v-\bar{v})) \\ -\gamma^2 a(u-\bar{u}, u-\bar{u}) - \frac{1}{\gamma^2} a(v-\bar{v}, v-\bar{v}).\end{aligned} \tag{34}$$

The last two terms in the above identity present the errors in the energy norm for the main and adjoint problems. Thus, we can immediately use the two-sided estimates from Section 3, written in somewhat simplified form:

$$M^{\ominus} \leq a(u-\bar{u}, u-\bar{u}) \leq M^{\oplus}, \quad M^{\ominus}_{ad} \leq a(v-\bar{v}, v-\bar{v}) \leq M^{\oplus}_{ad}, \tag{35}$$

where subindex "*ad*" means that the corresponding estimate is obtained for the adjoint problem.

Concerning the first term in the right-hand side of (34), we observe that

$$\begin{aligned}a(\gamma(u-\bar{u}) + \frac{1}{\gamma}(v-\bar{v}), \gamma(u-\bar{u}) + \frac{1}{\gamma}(v-\bar{v})) = \\ = a((\gamma u + \frac{1}{\gamma} v) - (\gamma \bar{u} + \frac{1}{\gamma}\bar{v}), (\gamma u + \frac{1}{\gamma} v) - (\gamma \bar{u} + \frac{1}{\gamma}\bar{v})).\end{aligned} \tag{36}$$

The function $\gamma u + \frac{1}{\gamma} v$ is the solution of the following problem (called as the *mixed problem* in what follows): Find $u_\gamma \in u_0 + H^1_{\Gamma_D}(\Omega)$ such that

$$\int\limits_\Omega A\nabla u_\gamma \cdot \nabla w \, dx + \int\limits_\Omega c u_\gamma w \, dx = \gamma F(w) + \frac{1}{\gamma}\ell(w) \quad \forall w \in H^1_{\Gamma_D}(\Omega), \tag{37}$$

which is uniquely solvable due to the fact that $\gamma F(w) + \frac{1}{\gamma}\ell(w)$ is, obviously, a linear bounded functional.

The function $\gamma\bar{u} + \frac{1}{\gamma}\bar{v} \in H^1_{\Gamma_D}(\Omega)$ can be considered as an approximation for u_γ, and we can again apply the techniques of Section 3 in order to obtain the following two-sided estimates (written again in a simplified form)

$$M^{\ominus}_{mix} \leq a(\gamma(u-\bar{u}) + \frac{1}{\gamma}(v-\bar{v}), \gamma(u-\bar{u}) + \frac{1}{\gamma}(v-\bar{v})) \leq M^{\oplus}_{mix}, \tag{38}$$

where subindex "*mix*" means that the estimates are obtained for the mixed problem.

Then we can conclude that

$$\frac{1}{2}(M^{\ominus}_{mix} - \gamma^2 M^{\oplus} - \frac{1}{\gamma^2} M^{\oplus}_{ad}) \leq E_1(u-\bar{u}, v-\bar{v}), \tag{39}$$

and

$$E_1(u-\bar{u}, v-\bar{v}) \leq \frac{1}{2}(M^{\oplus}_{mix} - \gamma^2 M^{\ominus} - \frac{1}{\gamma^2} M^{\ominus}_{ad}). \tag{40}$$

The above considerations can be summarised as the following result.

Theorem 4.2 For the error in terms of linear functional $\ell(u-\bar{u})$ we have the following upper estimate:

$$\ell(u-\bar{u}) \leq E_0(\bar{u}, \bar{v}) + \frac{1}{2}(M^{\oplus}_{mix} - \gamma^2 M^{\ominus} - \frac{1}{\gamma^2} M^{\ominus}_{ad}), \tag{41}$$

and the following lower estimate:

$$\ell(u-\bar{u}) \geq E_0(\bar{u}, \bar{v}) + \frac{1}{2}(M^{\ominus}_{mix} - \gamma^2 M^{\oplus} - \frac{1}{\gamma^2} M^{\oplus}_{ad}), \tag{42}$$

where the directly computable term $E_0(\bar{u}, \bar{v})$ is defined in (32), and γ is any positive real number.

5 Practical Realization of Error Estimation

In this section, we briefly discuss main issues of the practical realization of the above described error estimation technologies.

5.1 On the computation of the global constants

The constant C_{Ω,Γ_D} is determined via the smallest eigenvalue $\lambda_{\Omega,\Gamma_D}$ of the Laplacian in Ω with homogeneous boundary conditions, $C_{\Omega,\Gamma_D} = \frac{1}{\sqrt{\lambda_{\Omega,\Gamma_D}}}$. Thus, only estimation of $\lambda_{\Omega,\Gamma_D}$ from below is needed. In the case of homogeneous Dirichlet boundary condition (i.e., $\Gamma_N = \emptyset$) this task is easily solved as proposed by S. Mikhlin in [21, p. 8] by enclosing the solution domain into a rectangular parallelepiped, for which we can easily obtain the exact value of the smallest eigenvalue which is smaller than $\lambda_{\Omega,\Gamma_D}$. Also, suitable upper estimates of C_{Ω,Γ_D} for some conical domains are presented in [7]. On the contrary, the estimation of the constant $C_{\partial\Omega}$ seems to be still an open problem for a general case. However, one trick on estimation of this constant for a quite special case is proposed in [26, Remark 3.3]. More sophisticated techniques for the estimation of C_{Ω,Γ_D} and $C_{\partial\Omega}$ from above, suitable for the purposes of a posteriori error analysis, will be presented in our subsequent paper.

Remark 5.1 We note that another popular error estimation technique of the residual type involves *many unknown constants*, mostly related to patches of computational meshes used. Those constants are very hard to estimate (from above) and computation of them, in general, leads to a very big overestimation of the error even in simple cases (cf. [9]). Moreover, such constants have to be always recomputed if we perform adaptive computations and change the computational mesh. On the contrary, the constants C_{Ω,Γ_D} and $C_{\partial\Omega}$ remain the same under any change of the mesh as depending solely on the given data of the problem.

Remark 5.2 Due to the form of the upper estimate (16), we observe that computation of the global constants (if not provided analytically a priori) can be done in parallel to computing the main "ingredients" of the estimate as both constants are just factors multiplying the integrals.

5.2 Construction and tuning of the error estimates

As all three problems – original, adjoint and mixed ones, required for two-sided error estimation in terms of linear functionals – are of the same nature, we shall discuss in detail how to construct and optimize two-sided bounds of the error in the energy norm for the original problem (1)–(3) only.

It is clear that it is hardly possible to develop a single universal procedure for tuning all the parameters involved in order to get the most optimal (i.e., sufficiently close to the true error) bounds. In fact, such a tuning can be very much dependent of the nature of a particular numerical technique used to compute the concrete approximation $\bar{u}$.

However, for the sake of completeness, we present here one general scenario of optimizing the bounds in the case when $\bar{u}$ is obtained by FEM. Let it be further denoted by $u_h = u_0 + u_h^\star$, and let us assume that it is computed on the mesh $\mathcal{T}_h$. In addition, we suppose that for the problem considered there also exists a series of other (often finer) meshes $\mathcal{T}_{h_1}, \mathcal{T}_{h_2}, \ldots, \mathcal{T}_{h_k}$, with corresponding FE approximations $u_{h_i} = u_0 + u_{h_i}^\star$, $i = 1, \ldots, k$, such that

$$J(u_h^\star) \geq J(u_{h_1}^\star) \geq J(u_{h_2}^\star) \geq \cdots \geq J(u_{h_k}^\star). \tag{43}$$

In fact, such a situation is quite typical in most of practical calculations which use modern software packages. Depending on each particular case, the above mentioned meshes can be constructed either completely independent of each other or successively formed, e.g., from the main mesh $\mathcal{T}_h$ by using some local or global refinements.

Computation of the lower bound: The estimate (24) has a practical sense only if it provides with a nonnegative lower bound for the (nonnegative) error. This immediately suggests to construct reasonably good lower bounds as follows:

$$a(u - u_h, u - u_h) \geq 2(J(u_h^\star) - J(u_\mu^\star)) \geq 0, \tag{44}$$

where $\mu = h_1, h_2, \ldots, h_k$.

Minimization of the upper bound: Certain upper bounds for the error in the energy norm can be fastly computed using the following values for the "free" parameter y^*: $y_\mu^* = G_\mu(A\,\nabla u_\mu)$, where $\mu = h, h_1, h_2, \ldots, h_k$, and G_μ is some *gradient averaging operator* [8, 12] related to the mesh $\mathcal{T}_\mu$. Then, having y^* defined, we can easily compute (even analytically) the corresponding optimal values of the remaining parameters α and β. In many cases, the upper bounds obtained in this way can be of a quite good quality. However, more sharp estimates require a further minimization of the upper bound with respect to y^*, α, β. One possible strategy for this purpose shall be described in Section 6.

5.3 Error indicators based on superconvergence

First of all, we note that in the case when the coefficient $c \equiv 0$ (diffusion models), the error representations (12) and (31) (or, actually, the term E_1 in (33)) contain only unknown values of the gradients ∇u and ∇v and no unknown values of the functions u and v themselves.

This observation suggested a procedure of replacing the (unknown) gradients of the exact solutions of the main and adjoint problems by the corresponding averaged gradients if the approximations $\bar{u}$ and $\bar{v}$ are computed by FEM (in this case we denote them as u_h and v_τ, respectively). This idea is essentially based on the well-known phenomenon of superconvergence of averaged gradients (see [8, 12, 31] for details and more references on this topic) and it has been used in many works by now, which include, for example, [11, 13, 14, 16, 27, 28, 29, 33], where quite effective error indicators have been proposed for both types of the error control and for various problems of the elliptic type.

For the problem (1)–(3) (with $c \equiv 0$), it is natural to define such indicators as follows

$$\mathrm{I}_{gl}(u_h, G_h) := \int\limits_\Omega \mathrm{A}(G_h(\nabla u_h) - \nabla u_h) \cdot (G_h(\nabla u_h) - \nabla u_h)\, dx, \tag{45}$$

and

$$\mathrm{I}_{loc}(u_h, G_h; v_\tau, G_\tau) := E_0(u_h, v_\tau) + \int\limits_\Omega \mathrm{A}(G_h(\nabla u_h) - \nabla u_h) \cdot (G_\tau(\nabla v_\tau) - \nabla v_\tau)\, dx, \tag{46}$$

for the global and local error control purposes, respectively. In the above, G_h and G_τ are some gradient averaging operators, see, e.g., works [8, 12] and references therein, for various definitions of such operators.

Remark 5.3 However, it is also well-known that the phenomena of superconvergence are presented most strongly only in the case when the exact solution is sufficiently smooth, which, in fact, considerably limits the quality and usage of the error indicators. In particular, in Section 5.3 we show that the indicators of type (45) and (46) completely fail, for example, if there exist considerable jumps in the coefficients of the problem and the smoothnes of u deteriorates (which is quite typical for most of real-life cases) even for a simple Poisson problem posed in a square domain.

6 Numerical tests

In this section we present numerical tests demonstrating high effectivity of the two-sided estimation procedures proposed in our work, and also discuss the performance and failure of error indicators (45) and (46). All four tests are performed in planar domains, i.e., $d = 2$. For simplicity and purposes of comparison of the two-sided estimation and the error indicators, we take the coefficient $c \equiv 0$, define the right-hand side function $f \equiv 10$, and consider only the case of homogeneous Dirichlet boundary condition, i.e., $\Gamma_D = \partial\Omega$ and $u_0 \equiv 0$ in all the tests. In this situation, we need an accurate upper bound only for one global constant C_{Ω,Γ_D}, which can be easily obtained using Mikhlin's trick [21, p. 8]. The approximations $\bar{u}$ and $\bar{v}$ are assumed to be computed by the linear FEM using the PDE Toolbox of Matlab, i.e., they are continuous piecewise linear functions defined by nodal values at vertices of computational meshes $\mathcal{T}_h$ (for the main problem) and $\mathcal{T}_\tau$ (for the adjoint problem) and denoted by u_h and v_τ, respectively, later on. The corresponding averaging operators are defined as follows (cf. [12]): the operator G_h maps the gradient $\nabla u_h = \left[\frac{\partial u_h}{\partial x_1}, \ldots, \frac{\partial u_h}{\partial x_d}\right]^T$, which is constant over each triangular element of $\mathcal{T}_h$, into a vector-valued continuous piecewise affine function

$$G_h(\nabla u_h) = [G_h^1(\nabla u_h), \ldots, G_h^d(\nabla u_h)]^T,$$

whose each nodal value is the weighted mean (with respect to areas of elements) of the values of ∇u_h on all elements of the patch associated with the given node in the mesh $\mathcal{T}_h$. The operator G_τ is defined similarly.

Remark 6.1 In all the tests below, we compute the values of the exact errors using instead of the unknown exact solutions the so-called *reference solutions* which are obtained by solving the problem (1)–(3) on a very fine mesh with respect to both meshes $\mathcal{T}_h$ and $\mathcal{T}_\tau$.

6.1 Error control in global energy norm

Test 1: We consider problem (1)–(3) posed in a planar domain Ω with a reentrant corner (see Fig. 1 (left)). Let A be equal to the unit matrix (denoted by the symbol I later on). The approximation $\bar{u} = u_h$ is computed on the mesh $\mathcal{T}_h$ having 92 nodes.

To obtain sufficiently sharp upper and lower bounds for the error in the corresponding energy norm, we employ several successive meshes (with computation of the corresponding FE solutions) obtained from $\mathcal{T}_h$ by global red-refinements (when each triangular element is splitted into four similar smaller ones). As each new mesh in this refinement process is, in general, much finer than the previous one, one can expect that the FE solution on the new mesh is more accurate than the FE approximation calculated on the previous (more coarse) mesh. In practice, this means nothing else but that the sequence of values of the corresponding energy functional taken on such successive FE solutions is monotonically decaying (cf. (43)), which is used to obtain more and more accurate lower error bounds in a view of relation (44).

In the present test, we used five extra meshes (i.e., $k = 5$) constructed by the global red-refinements of the main mesh $\mathcal{T}_h$ (see Fig. 1 (right), where circles over numbers 1, 2, ..., 5 show the values of the resulting lower bounds). However, most of the other numerical experiments showed that the usage of only one or two extra meshes is quite sufficient to get reasonably sharp estimates of the error from below (see, for example, similar tuning of the lower bound for Test 2 sketched in Fig. 3 (right)), and much more efforts are often required for tuning the upper bounds. We mention here that a more economical way of constructing the required extra meshes $\mathcal{T}_{h_i}$, $i = 1, \dots, k$, based on suitable local (e.g., red-green) refinements can also be recommended.

In order to minimize the upper bound, we can easily use the idea based on the gradient averaging as proposed in Section 5.2. It is really quite simple as averaging procedures are usually "computationally cheap", and this approach works very effectively if the approximation u_h (and also all the other u_{h_i}) is close to the sufficiently smooth exact solution u. Moreover, most of required ingredients (extra meshes together with corresponding FE solutions) are already in hands due to optimization of the lower bounds. However, in a more general case, the upper estimates obtained in such a straightforward way can be too pessimistic. Therefore, more refined strategy for tuning the upper bound is needed. One possible way for doing this is as follows. First, we can prescribe $y^* := G_h(A\nabla u_h)$ and find corresponding parameters α and β. Then, using the computed values of α and β we can minimize the upper bound, which is the quadratic functional with respect to y^*, in the finite-dimensional subspace of $H_N(\Omega, \mathrm{div})$ generated, e.g., by the linear finite elements over the mesh $\mathcal{T}_h$. For a new value of y^* we can find new values of α and β, etc. This process can be repeated several times (we used it up to five times in the computations). If the gap between the upper and lower bounds is still too big, we map the last obtained value of y^* into a larger finite-dimensional subspace of $H_N(\Omega, \mathrm{div})$, now generated by the linear finite elements over the finer mesh $\mathcal{T}_{h_1}$ and perform the same iterative procedure as before. In this way, in the present test, besides $\mathcal{T}_h$, we have used four extra meshes, $\mathcal{T}_{h_1}, \dots, \mathcal{T}_{h_4}$ (see the results marked by "stars" in Fig. 1 (right)).

The behaviour of the two-sided estimates is schematically presented in Fig. 1 (right) with the corresponding values given in Table 1: the upper bound is decreasing from 4.64 to 2.44, the lower bound grows from 1.22 to 1.74. The error $|\!|\!|\nabla(u - u_h)|\!|\!|^2_\Omega = 1.7514$ in our case. We can clearly observe that the estimates are approaching each other in the process of error estimation, so the analyst can easily decide when to terminate the computational process. The gradient averaging indicator $\mathrm{I}_{gl} = 1.7547$, which is quite close to the exact error. Such a high effectivity of the indicator could be really expected in this test due to high smoothness of the problem data implying a sufficient smoothness of the solution u.

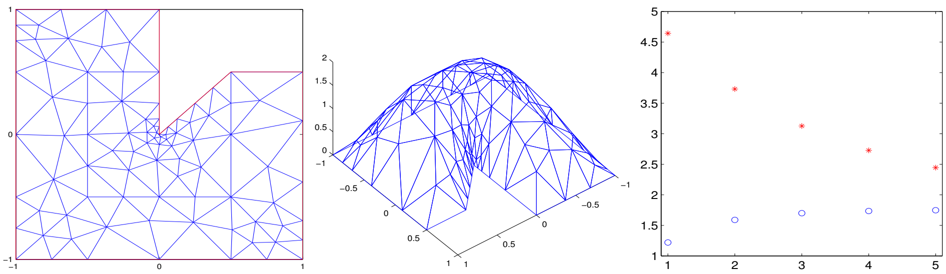

Figure 1: Solution domain Ω with computational mesh $\mathcal{T}_h$ having 92 nodes (left), the finite element approximation u_h (center), and optimization of the upper ("stars") and lower ("circles") error estimates (right) in Test 1.

$M^{\oplus}$	$\|\nabla(u-u_h)\|_{\Omega}^2$	$M^{\ominus}$
4.6426	1.7514	1.2191
3.7313	1.7514	1.5893
3.1266	1.7514	1.6998
2.7266	1.7514	1.7347
2.4443	1.7514	1.7469

Table 1: Two-sided estimates versus the error $\|\nabla(u-u_h)\|_{\Omega}^2$ in Test 1.

Test 2: In this test we demonstrate that the global error indicator I_{gl} defined in (45), can completely fail (i.e., considerably overestimate or underestimate the true error) if the problem data is not sufficiently smooth. For this purpose, we consider problem (1)–(3) posed in a simple square domain $\Omega := (-1,1)\times(-1,1)$ (see Fig. 3). Further, let the coefficient matrix A be defined as sketched in Fig. 2, i.e., A has high jumps in the entries (coefficients) along the diagonals and in the center of Ω. The approximation $\bar{u} = u_h$ is computed on the mesh $\mathcal{T}_h$ having 78 nodes (see Fig. 3 (left and center)).

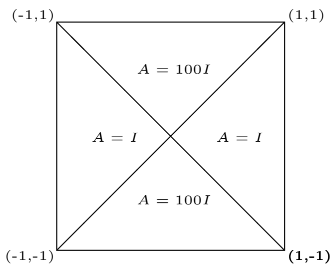

Figure 2: Definition of the coefficient matrix A in Test 2 and Test 4.

The indicator gives the value $\mathrm{I}_{gl} = 17.9628$, which considerably overestimates the exact error $\|\nabla(u-u_h)\|_{\Omega}^2 = 0.8584$. However, our two-sided estimates optimized in the same fashion as in the previous test still provide with reliable and guaranteed estimation of the global error from above and below (cf. Fig. 3 (right) and Table 2).

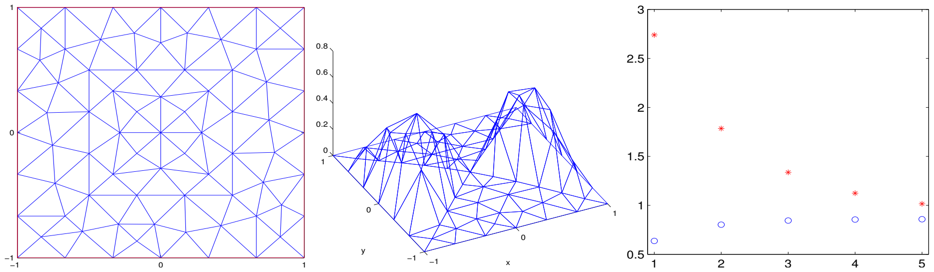

Figure 3: Solution domain Ω with computational mesh $\mathcal{T}_h$ having 78 nodes (left), the finite element approximation u_h (center), and optimization of the upper ("stars") and lower ("circles") error estimates (right) in Test 2.

$M^{\oplus}$	$\|\!\|\nabla(u-u_h)\|\!\|^2_{\Omega}$	$M^{\ominus}$
2.7394	0.8584	0.6364
1.7842	0.8584	0.8025
1.3367	0.8584	0.8445
1.1241	0.8584	0.8551
1.0157	0.8584	0.8577

Table 2: Two-sided estimates versus the error $\|\!\|\nabla(u-u_h)\|\!\|^2_{\Omega}$ in Test 2.

6.2 Error control in terms of linear functionals

Test 3: We consider the problem from Test 1 with the same approximation u_h. However, now we want to control the deviation $u - u_h$ in terms of a bounded linear functional ℓ defined by

$$\ell(w) = \int_{\Omega} \varphi w \, dx, \tag{47}$$

where $\varphi \in L_2(\Omega)$ and $\operatorname{supp}\varphi := \omega \subset \Omega$. The subdomain ω is taken in the neigbourhood of the reentrant vertex (which always is a zone of special interest for elliptic type boundary-value problems), and it is marked by the bold line in Fig. 4 (left). The weight-function in (47) satisfies $\varphi = 1$ in ω and vanishes outside of ω. Obviously, the function φ belongs to the space $L_2(\Omega)$, i.e., the corresponding adjoint problem is uniquely solvable.

The optimization of two-sided estimates is presented in Fig. 4 (right) and Table 3, where different choices of computational meshes (six in total) for the adjoint problem with 20, 47, 66, 103, 151, and 208 nodes (versus 92 nodes in $\mathcal{T}_h$) are used. The subindices 1, 2, 3 for $M^{\oplus}$ and $M^{\ominus}$ in Table 3 mean successive steps in the bounds' optimization like in Test 1 and Test 2. The error $\ell(u - u_h) = 0.0527$ is computed as before using the reference solution. Due to smoothness of the problem data the indicator I_{loc} gives values close to the true error (see Fig. 4 (right)).

In the computation of two-sided bounds for the local error we clearly observe the following phenomenon (which also appears in Test 4): the upper and lower estimates converge to each other faster if a finer computational mesh is used to approximately solve the adjoint problem. It can be

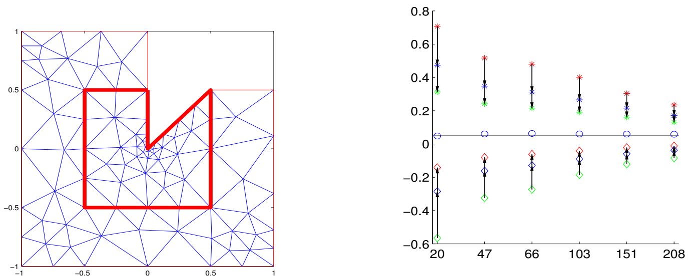

Figure 4: Solution domain Ω with ω marked by bold line, and mesh $\mathcal{T}_h$ (92 nodes) (left). The behaviour of the upper ("stars") and lower ("diamonds") error estimates versus the error $\ell(u-u_h)$ ("line") and the error indicator I_{loc} ("circles") values (right) in Test 3.

$\mathcal{T}_\tau$	$M_1^\oplus$	$M_2^\oplus$	$M_3^\oplus$	$\ell(u-u_h)$	$M_3^\ominus$	$M_2^\ominus$	$M_1^\ominus$	I_{loc}
20	0.7059	0.4730	0.3129	0.0527	-0.1419	-0.2850	-0.5648	0.0480
47	0.5165	0.3474	0.2416	0.0527	-0.0808	-0.1617	-0.3247	0.0604
66	0.4786	0.3121	0.2161	0.0527	-0.0619	-0.1286	-0.2734	0.0635
103	0.4003	0.2653	0.1907	0.0527	-0.0418	-0.0904	-0.1852	0.0603
151	0.3031	0.2153	0.1612	0.0527	-0.0213	-0.0585	-0.1209	0.0599
208	0.2352	0.1704	0.1313	0.0527	-0.0119	-0.0378	-0.0843	0.0588

Table 3: Two-sided error estimates and the indicator I_{loc} versus the error $\ell(u-u_h)$ in Test 3 for various "adjoint meshes".

explained by the fact that in the error decomposition (31) the second term E_1 is getting smaller with respect to the first term E_0 if the approximation $\bar{v}$ is getting more accurate (i.e., closer to the exact solution v). In this case the influence of the inaccuracy in two-sided estimation of the term E_1 is less critical as, in fact, the exactly computable term E_0 dominates. However, computations of approximations (and the estimates' values) on very dense meshes for adjoint problems can be quite expensive, so certain balance between computational costs coming from the computations of the approximate solution in the adjoint problem and costs appearing in the process of optimization of the two-sided bounds proposed should be found, it may depend much on the software used for concrete calculations.

For completeness, we also present in Fig. 5 examples of meshes used for computations of v_τ in the adjoint problem.

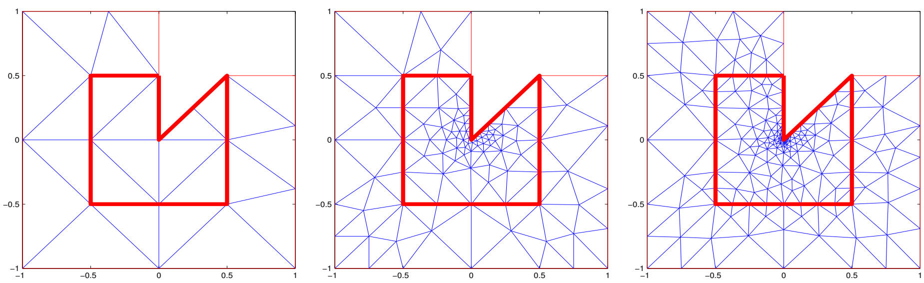

Figure 5: Meshes for the adjoint problem with 20, 103, and 208 nodes used in Test 3.

Test 4: To demonstrate that the error indicator (46), designed for the local error control, can fail, we take the same problem as in Test 2. The zone of interest ω is the square $(-0.5, 0.5) \times (-0.5, 0.5)$ in the center of Ω (see Fig. 6 (left)).

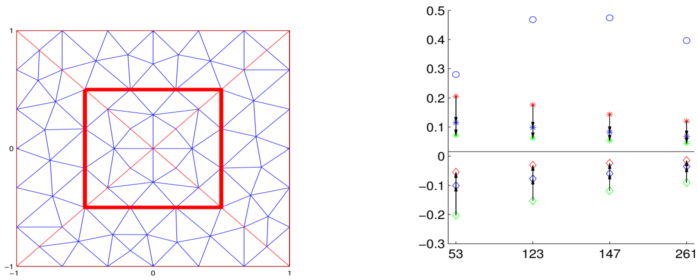

Figure 6: Solution domain Ω with ω marked by bold line and mesh $\mathcal{T}_h$ (78 nodes) (left). The behaviour of the upper ("stars") and lower ("diamonds") error estimates versus the error $\ell(u-u_h)$ ("line") and the error indicator I_{loc} ("circles") values (right) in Test 4.

Similarly to Test 3, we performed error estimation using several (four) different meshes for the adjoint problem. The results are reported in Table 4, see also Fig. 6 (right). We clearly see that in all the cases the values of I_{loc} are considerably larger than $\ell(u-u_h) = 0.0157$. However, two-sided estimation procedures provide with guaranteed estimation of the local error $\ell(u-u_h)$ from below and from above in all situations. For completeness, in Fig. 7 we present several meshes used in computation of approximations for the corresponding adjoint problem.

$\mathcal{T}_\tau$	$M_1^\oplus$	$M_2^\oplus$	$M_3^\oplus$	$\ell(u-u_h)$	$M_3^\ominus$	$M_2^\ominus$	$M_1^\ominus$	I_{loc}
53	0.2050	0.1148	0.0715	0.0157	-0.0547	-0.1008	-0.2026	0.2800
123	0.1757	0.0979	0.0620	0.0157	-0.0304	-0.0770	-0.1528	0.4681
147	0.1432	0.0823	0.0537	0.0157	-0.0240	-0.0594	-0.1187	0.4742
261	0.1197	0.0677	0.0441	0.0157	-0.0145	-0.0350	-0.0905	0.3961

Table 4: Two-sided estimates and the indicator I_{loc} versus the error $\ell(u-u_h)$ in Test 4 for various "adjoint meshes".

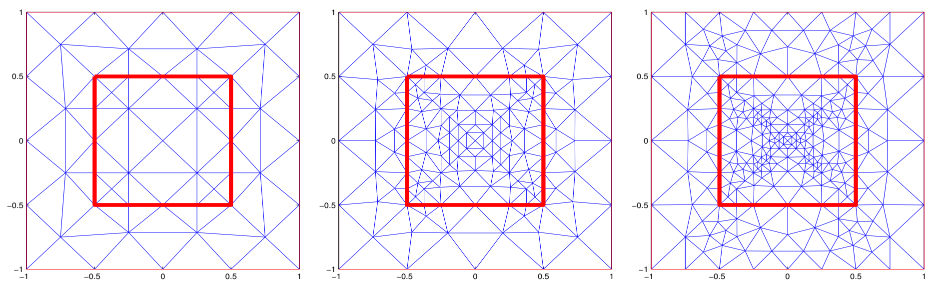

Figure 7: Meshes for the adjoint problems with 53, 147, and 261 nodes used in Test 4.

7 Conclusions

From the construction of the two-sided estimates and also from the tests presented, we observe that both presented estimates are really guaranteed and can, in principle, be made arbitrarily close to the true errors. Such closeness only depends on the available resources (memory, computational velocity) of the concrete computer used for calculations. On the contrary, popular error indicators, based on the superconvergence effect, are not very reliable if the problem data is not smooth enough. Nevertheless, due to simplicity of computation the indicators can be used, for example, for the first "rough" estimation of the error, or for the mesh adaptation purposes at initial steps, when the exact value of the error is not yet crucial.

Both guaranteed error estimates (16) and (24) have integral form, i.e., they can be represented as integrals over the solution domain Ω. This suggests a straightforward way for a mesh adaptation. Roughly speaking, we refine only those elements of the mesh whose contributions to the integrals in our two-sided estimates are too high.

The presented technologies can be adapted to treating other boundary conditions and other linear problems (e.g., in the linear elasticity (cf. [14]) and for convection-diffusion problems (see [15, 18])). The two-sided estimation proposed in this work can be easily coded and added as some block-checker to most of existing educational and industrial software products like MATLAB, FEMLAB, ANSYS, etc.

Acknowledgment

The authors wish to thank the anonymous referees for their careful reading of the manuscript and their fruitful comments and suggestions. The first author was supported by the project no. 211512 from the Academy of Finland. The second author was supported by the Academy Research Fellowship no. 208628 from the Academy of Finland.

References

[1] M. Ainsworth and I. Babuška, Reliable and robust a posteriori error estimating for singularly perturbed reaction-diffusion problems, *SIAM Journal on Numerical Analysis* **36** 331-353(1999).

[2] M. Ainsworth and J.T. Oden, *A Posteriori Error Estimation in Finite Element Analysis.* John Wiley & Sons, Inc., 2000.

[3] I. Babuška and W.C. Rheinbold, Error estimates for adaptive finite element computations, *SIAM Journal on Numerical Analysis* **15** 736-754(1978).

[4] I. Babuška and T. Strouboulis, *The Finite Element Method and Its Reliability.* Oxford University Press Inc., New York, 2001.

[5] W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations.* Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2003.

[6] R. Becker and R. Rannacher, A feed-back approach to error control in finite element methods: Basic approach and examples, *East-West Journal of Numerical Mathematics* **4** 237-264(1996).

[7] M. Borsuk and V. Kondratiev, *Elliptic Boundary Value Problems of Second Order in Piecewise Smooth Domains.* North-Holland Mathematical Library, vol. **69**. Elsevier, 2006.

[8] J. Brandts and M. Křížek, Gradient superconvergence on uniform simplicial partitions of polytopes, *IMA Journal of Numerical Analysis* **23** 489-505(2003).

[9] C. Carstensen and S. A. Funken, Constants in Clément-interpolation error and residual based a posteriori error estimates in finite element methods, *East-West Journal of Numerical Mathematics* **8** 153-175(2000).

[10] I. Faragó and J. Karátson, *Numerical Solution of Nonlinear Elliptic Problems via Preconditioning Operators: Theory and Applications*. Advances in Computation: Theory and Practice, vol. **11**. Nova Science Publishers, Inc., Hauppauge, NY, 2002.

[11] A. Hannukainen and S. Korotov, Techniques for a posteriori error estimation in terms of linear functionals for elliptic type boundary value problems, *Far East Journal of Applied Mathematics* **21** 289-304(2005).

[12] I. Hlaváček and M. Křížek, On a superconvergent finite element scheme for elliptic systems. I, II, III, *Aplikace Matematiky* **32** 131-154, 200-213, 276-289(1987).

[13] S. Korotov, A posteriori error estimation of goal-oriented quantities for elliptic type BVPs, *Journal of Computational and Applied Mathematics* **191** 216-227(2006).

[14] S. Korotov, Error control in terms of linear functionals based on gradient averaging techniques, *Computing Letters* **3(1)** 35-44(2007).

[15] S. Korotov, Global a posteriori error estimates for convection-reaction-diffusion problems, *Applied Mathematical Modelling* (in press).

[16] S. Korotov, P. Neittaanmäki and S. Repin, A posteriori error estimation of goal-oriented quantities by the superconvergence patch recovery, *Journal of Numerical Mathematics* **11** 33-59(2003).

[17] M. Křížek and P. Neittaanmäki, *Finite Element Approximation of Variational Problems and Applications*. Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. **50**. Longman Scientific & Technical, Harlow; copublished in USA with John Wiley & Sons, Inc., New York, 1990.

[18] D. Kuzmin, A. Hannukainen and S. Korotov, A new a posteriori error estimate for convection-reaction-diffusion problems, *Journal of Computational and Applied Mathematics* (in press).

[19] C. Lovadina and R. Stenberg, Energy norm a posteriori error estimates for mixed finite element methods, *Mathematics of Computation* **75** 1659-1674(2006).

[20] S.G. Mikhlin, *Variational Methods in Mathematical Physics*. Pergamon Press, 1964.

[21] S.G. Mikhlin, *Constants in Some Inequalities of Analysis*. A Wiley-Interscience Publication. John Wiley & Sons, Ltd., Chichester, 1986.

[22] J. Nečas, *Les Méthodes Directes en Théorie des Équations Elliptiques*. Academia, Prague, 1967.

[23] P. Neittaanmäki and S. Repin, *Reliable Methods for Computer Simulation. Error Control and A Posteriori Estimates*. Studies in Mathematics and its Applications, 33, Elsevier Science B.V., Amsterdam, 2004.

[24] S. Repin, A posteriori error estimation for nonlinear variational problems by duality theory, *Zapiski Nauchnykh Seminarov (POMI)* **243** 201-214(1997).

[25] S. Repin, S. Sauter and A. Smolianski, A posteriori error estimation for the Dirichlet problem with account of the error in the approximation of boundary conditions, *Computing* **70** 205-233(2003).

[26] S. Repin, S. Sauter and A. Smolianski, A posteriori error estimation for the Poisson equation with mixed Dirichlet/Neumann boundary conditions, *Journal of Computational and Applied Mathematics* **164/165** 601-612(2004).

[27] M. Rüter, S. Korotov and Ch. Steenbock, Goal-oriented error estimates based on different FE-solution spaces for the primal and the dual problem with application to linear elastic fracture mechanics, *Computational Mechanics* **39** 787-797(2007).

[28] M. Rüter and E. Stein, Goal-oriented a posteriori error estimates in linear fracture mechanics, *Computer Methods in Applied Mechanics and Engineering* **195** 251-278(2006).

[29] E. Stein, M. Rüter and S. Ohnimus, Adaptive finite element analysis and modelling of solids and structures. Findings, problems and trends, *International Journal for Numerical Methods in Engineering* **60** 103-138(2004).

[30] T. Vejchodský, Guaranteed and locally computable a posteriori error estimate, *IMA Journal of Numerical Analysis* **26** 525-540(2006).

[31] R. Verfürth, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques.* Wiley-Teubner, 1996.

[32] R. Verfürth, Robust a posteriori error estimators for a singularly perturbed reaction-diffusion equation, *Numerische Mathematik* **78** 479-493(1998).

[33] O.C. Zienkiewicz and J.Z. Zhu, The superconvergence patch recovery and a posteriori error estimates, Part 1 and Part 2, *International Journal for Numerical Methods in Engineering* **33** 1331-1364, 1365-1382(1992).