A Posteriori Error Estimation for the Dirichlet Problem with Account of the Error in the Approximation of Boundary Conditions

S. Repin, St. Petersburg, **S. Sauter**, Zürich,
and **A. Smolianski**, Zürich

Abstract

The present work is devoted to the a posteriori error estimation for 2nd order elliptic problems with Dirichlet boundary conditions. Using the duality technique we derive the reliable and efficient a posteriori error estimator that measures the error in the energy norm. The estimator can be used in assessing the error of any approximate solution which belongs to the Sobolev space H^1, independently of the discretization method chosen. In particular, our error estimator can be applied also to problems and discretizations where the Galerkin orthogonality is not available. We will present different strategies for the evaluation of the error estimator. Only one constant appears in its definition which is the one from Friedrichs' inequality; that constant depends solely on the domain geometry, and the estimator is quite non-sensitive to the error in the constant evaluation. Finally, we show how accurately the estimator captures the local error distribution, thus, creating a base for a justified adaptivity of an approximation.

2000 *Mathematics Subject Classification:* 35J20, 65N15, 65N30

Key words and phrases: a posteriori error estimate, duality technique, reliability, efficiency, local error distribution

1 Introduction

A posteriori error estimation has proved to be very essential for reliable scientific computing, and the literature on this subject is vast (see, e.g., the monographs [2], [5], [21] and also [6], [16]). Many different methods have been developed over the past decade, each having its own advantages and drawbacks. Some estimators possess the asymptotic exactness property, i.e. tend to the exact error with the growth of the dimension of the approximation space, but they may fail to provide a reliable upper bound for the exact error if some particular (relatively rough) approximation is considered. In addition, some of those estimators require a sufficient smoothness of the exact solution. Other estimators are free from the latter requirement and should, theoretically, yield a two-sided bound for the exact error, but they do not exhibit the asymptotic exactness. In this work we derive the a posteriori error estimator that has the asymptotic exactness property and, at the same time, provides a guaranteed upper bound for the exact error. The estimator is also completely independent of any particular discretization method and, hence, can be used in combination not only with finite element but also with finite difference or finite volume method, as well as for assessing the error in a post-processed solution.

In its original form, the estimator (like most existing a posteriori error estimators) is derived under the assumption that the approximate solution satisfies the Dirichlet boundary condition exactly.

However, in many practically interesting cases, the essential boundary condition can be satisfied merely approximately either owing to complicated, e.g., non-polynomial Dirichlet data or because of accounting the boundary condition in a weak sense, like in fictitious domain methods (see, e.g., [13]). Thus, the approximate solution does not, in general, belong to the set of admissible functions of the original problem, i.e. presents a *non-conforming* approximation to the exact solution (see [20]). In this paper, we show how the estimator can be modified to take into account the error in the approximation of the Dirichlet boundary condition. The resulting estimator measuring the error in the energy norm has a very clear form: it consists of three separate terms, the first of which controls the boundary condition error, the second corresponds to the error in the duality relation ("the error in the constitutive law"), and the third represents the error in the equation for the dual variable ("the error in the equilibrium equation"). The estimator could be very useful for adaptive improvement of approximations on the domains with complicated structures (see, e.g., [14]).

The a posteriori error estimation for the non-homogeneous Dirichlet problem has been recently investigated in [1] but with a different technique, namely, in the framework of the equilibrated residual method. In this paper, we will focus on Poisson's model problem but emphasize that our approach has the potential of being extended to more general elliptic problems and further types of the discretization error, such as the error in the geometric approximation of the domain (see [10], [12] concerning the estimation of the geometric error). These topics will be addressed in a forthcoming paper.

The paper is organized as follows. We start in Section 2 with the preliminaries on the primal and dual variational formulations and the links between them, and obtain the basic estimate for the energy norm of the error in the primal variable. Then, in Section 3, we proceed with deriving the a posteriori error majorant for approximate solutions exactly satisfying the boundary condition. We also discuss the strategy for evaluating the majorant and the estimation of the local error distribution. Next, in Section 4, we present the modified error majorant capable of estimating the error for non-conforming approximations. Finally, we illustrate the developed theory by numerical examples (Sections 5.1–5.3), discuss the computational complexity of the estimator (Section 5.4) and draw the conclusions (Section 6).

2 Preliminaries

Let Ω be a bounded domain with Lipschitz continuous boundary Γ. Consider the Dirichlet problem: Find a function u such that

$$-\Delta u = f \qquad \text{in } \Omega\,, \tag{2.1}$$

$$u = u_0 \qquad \text{on } \Gamma\,, \tag{2.2}$$

where the trace of the given function $u_0 \in H^1(\Omega)$ defines the boundary condition on Γ. Throughout this paper, we assume that

$$f \in L_2(\Omega).$$

Let $V_0 := H^1_0(\Omega)$ and $V_0 + u_0 := \{v \in H^1(\Omega) \mid v = u_0 + w,\ w \in V_0\}$. A weak formulation of the problem (2.1), (2.2) is: Find $u \in V_0 + u_0$, such that

$$\int\limits_{\Omega} \nabla u \cdot \nabla w \, dx = \int\limits_{\Omega} f w \, dx \quad \forall w \in V_0\,. \tag{2.3}$$

It is well known that the solution to this problem exists and is unique. This solution can be characterized equivalently as the minimizer of the following variational problem:

Problem $\mathcal{P}$. Find $u \in V_0 + u_0$ such that

$$J(u) = \inf_{v \in V_0 + u_0} J(v)\,,$$

where

$$J(v) = \frac{1}{2}\int\limits_{\Omega} \mid \nabla v \mid^2 \; dx - \int\limits_{\Omega} f v \, dx \, .$$

To derive the dual variational formulation we employ the relation

$$J(u) = \inf_{v \in V_0 + u_0} \sup_{y^* \in L^2(\Omega, \mathbb{R}^n)} \int\limits_{\Omega} \left(\nabla v \cdot y^* - \frac{1}{2} \mid y^* \mid^2 - f v \right) \, dx \, .$$

Then we have

Problem $\mathcal{P}^*$. Find $p^* \in Q^*_f$ such that

$$I^*(p^*) = \sup_{q^* \in Q^*_f} I^*(q^*) \, ,$$

where

$$I^*(q^*) = \int\limits_{\Omega} \left(\nabla u_0 \cdot q^* - \frac{1}{2} \mid q^* \mid^2 - f u_0 \right) \, dx$$

is the dual variational functional and

$$Q^*_f := \{ q^* \in L^2(\Omega, \mathbb{R}^n) \mid \int\limits_{\Omega} q^* \cdot \nabla w \, dx = \int\limits_{\Omega} f w \, dx \; \forall w \in V_0 \} \, . \tag{2.4}$$

Both problems $\mathcal{P}$ and $\mathcal{P}^*$ have unique solutions u and p^*, which satisfy the duality relations (see, e.g., [15])

$$J(u) = I^*(p^*) \, , \tag{2.5}$$

$$\nabla u = p^* \, . \tag{2.6}$$

In view of (2.3), we have

$$\begin{aligned} J(v) - J(u) &= \frac{1}{2}\int\limits_{\Omega} \mid \nabla (v-u) \mid^2 \; dx + \int\limits_{\Omega} \nabla u \cdot \nabla (v-u) \, dx \\ &- \int\limits_{\Omega} f(v-u) \, dx = \frac{1}{2}\int\limits_{\Omega} \mid \nabla (v-u) \mid^2 \; dx \quad \forall v \in V_0 + u_0 \, . \end{aligned}$$

Hence, using (2.5), one can derive

$$\begin{aligned} \frac{1}{2}\|\nabla(v-u)\|^2 &= J(v) - I^*(p^*) = J(v) - \sup_{q^* \in Q^*_f} I^*(q^*) \\ &= J(v) + \inf_{q^* \in Q^*_f} \{-I^*(q^*)\} = \inf_{q^* \in Q^*_f} \{J(v) - I^*(q^*)\} \, , \end{aligned}$$

and, since

$$\begin{aligned} J(v) - I^*(q^*) &= \frac{1}{2}\int\limits_{\Omega} \mid \nabla v - q^* \mid^2 \; dx + \int\limits_{\Omega} q^* \cdot \nabla v \, dx - \int\limits_{\Omega} q^* \cdot \nabla u_0 \, dx \\ &- \int\limits_{\Omega} f v \, dx + \int\limits_{\Omega} f u_0 \, dx = \frac{1}{2}\int\limits_{\Omega} \mid \nabla v - q^* \mid^2 \; dx \, , \forall v \in V_0 + u_0 \, , \, \forall q^* \in Q^*_f \, , \end{aligned}$$

we obtain

$$\|\nabla(v-u)\|^2 = \inf_{q^*\in Q_f^*} \|\nabla v - q^*\|^2 \quad \forall v \in V_0 + u_0\,. \tag{2.7}$$

From (2.7) we immediately deduce the estimate

$$\|\nabla(v-u)\|^2 \le \|\nabla v - q^*\|^2 \quad \forall v \in V_0 + u_0\,,\, \forall q^* \in Q_f^*\,, \tag{2.8}$$

showing that $\nabla(v-u)$ in the L^2-norm (which may be viewed as the approximation error) is majorized by the L^2-norm of the difference $(\nabla v - q^*)$ with any $q^* \in Q_f^*$. However, if q^* does not belong to Q_f^*, the L^2-norm of the difference $(\nabla v - q^*)$ does not, in general, provide an upper bound for the error. This means that any numerical approximation of q^* should satisfy the constraint $q^* \in Q_f^*$ with very high accuracy in order to guarantee a reliability of the error estimate; thus, the estimate (2.8) is not very useful for practical application. Our further efforts will be focused on finding a computable upper bound for the right-hand side of (2.8).

3 Error majorant for conforming approximations

Functional-type a posteriori error majorants for conforming approximations have been derived by using general minimax theorems of convex analysis in [18], [19]. For linear elliptic problems under consideration we will present in this section a much simplified way of deriving functional-type a posteriori estimates using a variant of the Helmholtz decomposition for the space $L^2(\Omega, \mathbb{R}^n)$.

Theorem 3.1 *Let $u \in V_0 + u_0$ be the solution to the problem (2.1), (2.2) and v any function from $V_0 + u_0$.*
Then

$$\|\nabla(v-u)\|^2 \le (1+\beta)\|\nabla v - y^*\|^2 + (1+\frac{1}{\beta})\, C_\Omega^2 \|\operatorname{div} y^* + f\|^2\,, \tag{3.1}$$

where β is an arbitrary positive number, y^* is any function from $H(\Omega, \operatorname{div}) := \{y^* \in L^2(\Omega, \mathbb{R}^n) \mid \operatorname{div} y^* \in L^2(\Omega)\}$ and C_Ω is the constant from Friedrichs' inequality for the domain Ω.

Proof. Let us first consider the homogeneous equation, i.e. the problem

$$-\Delta u = 0 \qquad \text{in } \Omega\,, \tag{3.2}$$

$$u = u_0 \qquad \text{on } \Gamma\,. \tag{3.3}$$

After repeating the derivations of the preceding section with $f \equiv 0$ we obtain (see (2.8))

$$\|\nabla(v-u)\|^2 \le \|\nabla v - q^*\|^2 \quad \forall v \in V_0 + u_0\,,\, \forall q^* \in Q_0^*\,, \tag{3.4}$$

where $Q_0^* := \{q^* \in L^2(\Omega, \mathbb{R}^n) \mid \int_\Omega q^* \cdot \nabla w \, dx = 0 \;\forall w \in V_0\}$.

In order to estimate the right-hand side of (3.4) for any fixed $v \in V_0 + u_0$, we take an arbitrary function y^* from the space $H(\Omega, \operatorname{div})$. As $y^* \in L^2(\Omega, \mathbb{R}^n)$, we have for y^* the Helmholtz orthogonal decomposition $y^* = q^* + \nabla w$, where $q^* \in Q_0^*$ and $w \in V_0$ (see, e.g., [17] for this particular variant of the Helmholtz decomposition). Since $y^* \in H(\Omega, \operatorname{div})$, we conclude that $\Delta w = \operatorname{div} y^* \in L^2(\Omega)$. Then, using Young's inequality, we obtain

$$\|\nabla v - q^*\|^2 \le (1+\beta)\|\nabla v - y^*\|^2 + (1+\frac{1}{\beta})\|\nabla w\|^2 \quad \forall \beta > 0\,. \tag{3.5}$$

Since $w \in V_0$ and $\Delta w \in L_2(\Omega)$, the second term on the right-hand side of the last inequality can be estimated by

$$\|\nabla w\|^2 = \left| \int_\Omega (\Delta w) w \, dx \right| \le \|\Delta w\|\|w\| \le C_\Omega \|\Delta w\|\|\nabla w\|, \tag{3.6}$$

that is

$$\|\nabla w\| \leq C_\Omega \|\Delta w\| \,, \tag{3.7}$$

where C_Ω is the constant in Friedrichs' inequality for the domain Ω. As $\Delta w = \operatorname{div} y^*$, we arrive at the inequality

$$\|\nabla v - q^*\|^2 \leq (1+\beta)\|\nabla v - y^*\|^2 + (1+\frac{1}{\beta})\, C_\Omega^2 \|\operatorname{div} y^*\|^2 \quad \forall \beta > 0\,. \tag{3.8}$$

Using (3.4) we finally obtain the a posteriori error estimate

$$\|\nabla(v-u)\|^2 \leq (1+\beta)\|\nabla v - y^*\|^2 + (1+\frac{1}{\beta})\, C_\Omega^2 \|\operatorname{div} y^*\|^2 \,, \qquad \forall \beta > 0\,,\ \forall y^* \in H(\Omega, \operatorname{div})\,, \tag{3.9}$$

which is valid for any approximate solution v from $V_0 + u_0$.

To derive an analogous estimate for the general inhomogeneous equation, i.e. for the problem (2.1), (2.2), we consider an auxiliary problem

$$-\Delta \bar{u} = f \qquad \text{in } \Omega\,, \tag{3.10}$$

$$\bar{u} = 0 \qquad \text{on } \Gamma\,. \tag{3.11}$$

Then the solution to the problem (3.2), (3.3) is given by $\tilde{u} = u - \bar{u}$. Applying (3.9) to any $\tilde{v} \in V_0 + u_0$ we obtain

$$\|\nabla(\tilde{v} - u + \bar{u})\|^2 \leq (1+\beta)\|\nabla \tilde{v} - \tilde{y}^*\|^2 + (1+\frac{1}{\beta})\, C_\Omega^2 \|\operatorname{div} \tilde{y}^*\|^2 \,, \tag{3.12}$$

valid for any $\tilde{y}^* \in H(\Omega, \operatorname{div})$ and for any $\beta > 0$.

The solution $\bar{u}$ is related to the solution $\bar{p}^*$ of the corresponding dual problem by

$$\nabla \bar{u} = \bar{p}^*\,. \tag{3.13}$$

However, in this case, $\bar{p}^*$ belongs not to Q_0^*, but to the affine manifold Q_f^* (see (2.4)). The inequality (3.12) can be rewritten in the form

$$\|\nabla(\tilde{v} - u + \bar{u})\|^2 \leq (1+\beta)\|\nabla(\tilde{v} + \bar{u}) - (\tilde{y}^* + \bar{p}^*)\|^2 + (1+\frac{1}{\beta}) C_\Omega^2 \|\operatorname{div} \tilde{y}^*\|^2\,. \tag{3.14}$$

Note that $v := \tilde{v} + \bar{u} \in V_0 + u_0$ and $y^* := \tilde{y}^* + \bar{p}^* \in H(\Omega, \operatorname{div})$. Then

$$\operatorname{div} \tilde{y}^* = \operatorname{div} y^* - \operatorname{div} \bar{p}^* = \operatorname{div} y^* + f\,, \tag{3.15}$$

and

$$\|\nabla(v-u)\|^2 \leq (1+\beta)\|\nabla v - y^*\|^2 + (1+\frac{1}{\beta})\, C_\Omega^2 \|\operatorname{div} y^* + f\|^2\,.$$

■

Since u is the exact solution of the problem (2.1), (2.2), v is any function from $V_0 + u_0$ and y^* is any function from $H(\Omega, \operatorname{div})$, the estimate (3.1) is an a posteriori error estimate valid for any conforming approximation of the problem (2.1), (2.2).

Corollary 3.1 *Let $u \in V_0 + u_0$ be the solution to the problem (2.1), (2.2) and v any function from $V_0 + u_0$.*
Then

$$\|\nabla(v-u)\| \leq M(v; y^*) \quad \forall y^* \in H(\Omega, \operatorname{div})\,, \tag{3.16}$$

where $M(v; y^*) = \|\nabla v - y^*\| + C_\Omega \|\operatorname{div} y^* + f\|$.

Proof. Since (3.1) is valid for any $\beta > 0$, we can minimize the right-hand side of (3.1) with respect to β. This immediately leads to the estimate (3.16).

■

Remark 3.1 It is evident that the error majorant $M(v; y^*)$ is not only always reliable (as far as $y^* \in H(\Omega, \mathrm{div})$) but also possesses the property of *asymptotic exactness* in the sense that, if $y^* \to p^* = \nabla u$ in $H(\Omega, \mathrm{div})$, the majorant $M(v; y^*)$ tends to the exact error $\|\nabla(v - u)\|$.

Remark 3.2 It immediately follows from (3.16) that

$$M(v; y^*) \leq \|\nabla(v - u)\| + \|y^* - p^*\| + C_\Omega \|\mathrm{div}\,(y^* - p^*)\| \,. \tag{3.17}$$

This indicates that, if y^* is close enough to the exact dual variable p^* in $H(\Omega, \mathrm{div})$-norm, the error majorant $M(v; y^*)$ provides also a reasonable *lower bound* for the exact error.

Although the estimate (3.16) is more practical than (2.8) (since y^* should belong not to the set Q^*_f but to the much wider space $H(\Omega, \mathrm{div})$), it is still useless as long as one has no good choice for y^* at hand. Next, we will present three strategies for finding a good choice of y^*:
(a) If, in addition, the dual problem is discretized, we may directly plug the approximate dual solution in the error majorant, hence, obtaining a good error estimate;
(b) The much cheaper way is to project ∇v, where $v \in V_0 + u_0$ is the found approximate solution to our problem, onto the space $H(\Omega, \mathrm{div})$ by virtue of some suitable projection procedure. This leads to the well-known gradient recovery based error estimation (see [7], [22], [23]), but, unlike the standard estimators of that class, our majorant provides a guaranteed upper bound for the exact error owing to the presence of the additional, residual term $C_\Omega \|\mathrm{div}\, y^* + f\|$. The success of a gradient recovery, however, heavily relies on the superconvergence phenomenon that, in turn, can be observed only if the exact solution has a sufficient regularity;
(c) Another way of obtaining a good choice for y^*, which is less costly than solving the dual problem and does not require an additional smoothness of the exact solution, is given by the following

Theorem 3.2 *Let $M^2(v; y^*, \beta)$ denote the right-hand side of the inequality (3.1), i.e.* $M^2(v; y^*, \beta) := (1 + \beta)\|\nabla v - y^*\|^2 + (1 + \frac{1}{\beta})\, C_\Omega^2 \|\mathrm{div}\, y^* + f\|^2$.
Let $y^*_\beta \in H(\Omega, \mathrm{div})$ be the minimizer of $M^2(v; y^*, \beta)$ for any fixed $v \in V_0 + u_0$ and $\beta > 0$ $(M^2(v; y^*_\beta, \beta) = \inf\limits_{y^* \in H(\Omega, \mathrm{div})} M^2(v; y^*, \beta))$.
Then y^*_β converges to $p^* = \nabla u$ in $H(\Omega, \mathrm{div})$ as $\beta \to 0$, and moreover

$$\|y^*_\beta - p^*\| \leq C\beta^{1/2} \,,$$
$$\|\mathrm{div}\,(y^*_\beta - p^*)\| \leq C\beta \,,$$

where the constant C depends only on f, v and C_Ω.
(*The theorem implies that* $\|y^*_\beta - p^*\|_{H(\Omega, \mathrm{div})} = \mathcal{O}(\beta^{1/2})$ *when* $\beta \to 0$)

Proof. First, we see that for any fixed $v \in V_0 + u_0$ and $\beta > 0$ the functional $M^2(y^*) := M^2(v; y^*, \beta)$ is continuous, strictly convex and coercive on $H(\Omega, \mathrm{div})$, which immediately implies the existence of the unique minimizer $y^*_\beta \in H(\Omega, \mathrm{div})$. This minimizer satisfies the Euler equation for the functional $M^2(y^*)$

$$(1 + \beta) \int\limits_\Omega y^*_\beta \cdot z^* \, dx + (1 + \frac{1}{\beta})\, C_\Omega^2 \int\limits_\Omega (\mathrm{div}\, y^*_\beta + f) \mathrm{div}\, z^* \, dx =$$
$$(1 + \beta) \int\limits_\Omega \nabla v \cdot z^* \, dx \,,\ \forall z^* \in H(\Omega, \mathrm{div}) \,,$$

that can be rewritten in the form

$$\int_\Omega y^*_\beta \cdot z^*\,dx + \frac{1}{\beta} C^2_\Omega \int_\Omega (\operatorname{div} y^*_\beta + f) \operatorname{div} z^*\,dx = \int_\Omega \nabla v \cdot z^*\,dx\,,\ \forall z^* \in H(\Omega, \operatorname{div})\,. \tag{3.18}$$

In the preceding section it has been proved (see (2.7)) that p^* is the minimizer of the functional $F(q^*) := \|\nabla v - q^*\|^2$ on the set Q^*_f (the minimizer is obviously unique). The corresponding Euler equation for the functional F reads

$$\int_\Omega p^* \cdot q^*\,dx = \int_\Omega \nabla v \cdot q^*\,dx\,,\ \forall q^* \in Q^*_0\,. \tag{3.19}$$

Now we consider the function $(y^*_\beta - p^*) \in H(\Omega, \operatorname{div})$. Exactly as it has been done above in this section, we may use the Helmholtz decomposition $(y^*_\beta - p^*) = q^* + \nabla w$, where $q^* \in Q^*_0$ and $w \in V_0$, $\Delta w = \operatorname{div}(y^*_\beta - p^*) \in L^2(\Omega)$, to obtain from (3.19)

$$\int_\Omega p^* \cdot (y^*_\beta - p^*)\,dx - \int_\Omega p^* \cdot \nabla w\,dx = \int_\Omega \nabla v \cdot (y^*_\beta - p^*)\,dx - \int_\Omega \nabla v \cdot \nabla w\,dx\,. \tag{3.20}$$

If we substitute $(y^*_\beta - p^*)$ instead of z^* into (3.18) and subtract (3.20) from the resulting equation, we obtain

$$\|y^*_\beta - p^*\|^2 + \frac{1}{\beta} C^2_\Omega \|\operatorname{div}(y^*_\beta - p^*)\|^2 = -\int_\Omega f w\,dx + \int_\Omega \nabla v \cdot \nabla w\,dx\,. \tag{3.21}$$

The fact that $\operatorname{div} p^* = -f$ in $L^2(\Omega)$ has been also used to derive this equation.

We can estimate the right-hand side of (3.21) from above by means of the Cauchy-Schwarz inequality, Friedrichs' inequality for $w \in V_0$ and the inequality (3.7) to obtain

$$\|y^*_\beta - p^*\|^2 + \frac{1}{\beta} C^2_\Omega \|\operatorname{div}(y^*_\beta - p^*)\|^2 \leq C^2_\Omega \|f\|\,\|\Delta w\| + C_\Omega \|\nabla v\|\,\|\Delta w\|\,, \tag{3.22}$$

from where, using the fact $\Delta w = \operatorname{div}(y^*_\beta - p^*)$, we deduce

$$\|y^*_\beta - p^*\|^2 + \frac{1}{\beta} C^2_\Omega \|\operatorname{div}(y^*_\beta - p^*)\|^2 \leq (C^2_\Omega \|f\| + C_\Omega \|\nabla v\|)\,\|\operatorname{div}(y^*_\beta - p^*)\|\,.$$

It immediately follows that

$$\|\operatorname{div}(y^*_\beta - p^*)\| \leq (\|f\| + \frac{1}{C_\Omega}\|\nabla v\|)\,\beta$$

and, consequently,

$$\|y^*_\beta - p^*\|^2 \leq C^2_\Omega (\|f\| + \frac{1}{C_\Omega}\|\nabla v\|)^2\,\beta\,,$$

yielding the statement of the theorem.

■

Remark 3.3 Theorem 3.2 suggests that, taking a small value β_0 of the parameter β, we can obtain a reasonably good $y^*_{\beta_0}$ to be used in the error majorant M. This $y^*_{\beta_0}$ is found simply by minimizing the quadratic functional $M^2(v; y^*, \beta_0)$ with fixed v and β_0. It is important to note that, unlike in the penalty method for solving the dual problem, we do not have to set β_0 very small as it is not necessary for our purposes to strongly enforce the constraint $\operatorname{div} y^* = -f$ in $L^2(\Omega)$. The majorant provides an upper bound for the exact error as long as $y^* \in H(\Omega, \operatorname{div})$, and it is close to the exact error if the term $C_\Omega \|\mathrm{div}\, y^* + f\|$ is dominated by the other term $\|\nabla v - y^*\|$ in the majorant. The optimal value of β_0 should, obviously, depend on the chosen discretization. By setting the value of β_0 only moderately small, the typical problems of a penalty method, like a severe deterioration of the condition number and the locking effect for the discretized problem, can be circumvented.

Remark 3.4 The constant C_Ω stemming from Friedrichs' inequality is related to the smallest eigenvalue λ_Ω of the Laplace operator in Ω (equipped with the homogeneous Dirichlet boundary condition) by the expression $\lambda_\Omega = C_\Omega^{-2}$. Calculating the smallest eigenvalue in a larger domain leads to a lower bound for λ_Ω, and, consequently, to an upper bound for the constant C_Ω. It is worth noting that C_Ω depends solely on the domain geometry, and, thus, must be evaluated only once for the particular domain.
In the majorant $M(v; y^*_{\beta_0})$ the constant C_Ω is multiplied by the term $\|\mathrm{div}\, y^*_{\beta_0} + f\| = \|\mathrm{div}\,(y^*_{\beta_0} - p^*)\|$ which is of order of β_0 according to Theorem 3.2. Thus, the majorant is generally not sensitive to possible errors in the constant evaluation.

Next we will investigate the local estimation properties of our error majorant.

Evaluation of local errors by the error majorant

The function

$$\varepsilon(x) = \mid \nabla(v(x) - u(x)) \mid^2$$

gives the actual distribution of the energy error in the points of the domain Ω. As a matter of fact, this function provides the information needed, e.g., for performing an adaptive mesh refinement. However, since u is unknown, this function is not at our disposal. The theorem we prove below shows that $\varepsilon(x)$ can be successfully approximated by an explicitly computable function stemming from our error majorant.

Define the function

$$\mu_\beta(x) := (1+\beta) \mid \nabla v(x) - y^*_\beta(x) \mid^2 + \left(1 + \frac{1}{\beta}\right) C_\Omega^2 \mid \mathrm{div}\, y^*_\beta(x) + f(x) \mid^2 ,$$

where y^*_β is the minimizer of $M^2(v; y^*, \beta)$ as in Theorem 3.2. Define, also, for any $\sigma > 0$ the set

$$\Omega_\sigma = \{x \in \Omega \mid \ \mid \mu_\beta(x) - \varepsilon(x) \mid \geq \sigma\} .$$

Then the following theorem holds true:

Theorem 3.3 *For any $\sigma > 0$*

$$\mathrm{meas}\,(\Omega_\sigma) \to 0 \quad \text{as}\, \beta \to 0 .$$

Proof. For almost all $x \in \Omega$, we have

$$\begin{aligned} \mid \mu_\beta(x) - \varepsilon(x) \mid \leq\ & \Big|(1+\beta) \mid \nabla v(x) - y^*_\beta(x) \mid^2 - \mid \nabla(v(x) - u(x)) \mid^2\Big| \\ & + \left(1 + \frac{1}{\beta}\right) C_\Omega^2 \mid \mathrm{div}\, y^*_\beta(x) + f(x) \mid^2 \leq\ \beta \mid \nabla v(x) \mid^2 + 2 \mid (1+\beta) y^*_\beta(x) - \nabla u(x) \mid \\ & \cdot \mid \nabla v(x) \mid + \Big|(1+\beta) \mid y^*_\beta(x) \mid^2 - \mid \nabla u(x) \mid^2\Big| + \left(1 + \frac{1}{\beta}\right) C_\Omega^2 \mid \mathrm{div}\, y^*_\beta(x) + f(x) \mid^2 . \end{aligned}$$

Since, according to Theorem 3.2,

$$\begin{aligned} \|y^*_\beta - \nabla u\| &= \mathcal{O}(\beta^{1/2}) , \\ \|\mathrm{div} y^*_\beta + f\| &= \mathcal{O}(\beta) , \end{aligned}$$

it is easy to see that

$$\int\limits_{\Omega} \mid \mu_\beta(x) - \varepsilon(x) \mid \, dx \to 0 \quad \text{as}\, \beta \to 0\,. \tag{3.23}$$

As

$$\int\limits_{\Omega} \mid \mu_\beta(x) - \varepsilon(x) \mid \, dx \geq \int\limits_{\Omega_\sigma} \mid \mu_\beta(x) - \varepsilon(x) \mid \, dx \geq \sigma \,\text{meas}\, (\Omega_\sigma)\,,$$

the statement of the theorem follows from (3.23).

■

The theorem shows that $\mu_\beta(x)$ may essentially differ from $\varepsilon(x)$ only in "small" subdomains, whose Lebesgue measure will tend to zero if y^*_β becomes sufficiently close to the exact dual variable p^*. From the practical viewpoint, this means that the function $\mu_\beta(x)$ standing under the integral sign in $M^2(v; y^*_\beta, \beta)$ can be used as an indicator of local errors.

Below we discuss this issue more precisely. Let

$$\Omega = \bigcup_{k=1}^{N} \Omega_k\,,$$

where $\{\Omega_k\}$ is a certain collection of subdomains. The problem of our interest is to evaluate the quantities

$$e_k := \int\limits_{\Omega_k} \varepsilon(x)\, dx\,,$$

which provide necessary information for further improvement of the finite-dimensional approximation. For this purpose, we can use the quantities

$$m^\beta_k := \int\limits_{\Omega_k} \mu_\beta(x)\, dx\,.$$

The accuracy of the representation of the quantities e_k by m^β_k can be measured by the local effectivity indices

$$(\mathrm{i_{eff}})^\beta_k = \left(\frac{m^\beta_k}{e_k} \right)^{1/2} \qquad \forall k = \overline{1, N}\,,$$

if $e_k \neq 0$.

Corollary 3.2 *Let y^*_β be the same as in Theorem 3.2.*
Then, for any given sampling $\{\Omega_k\}_{k=1}^N$ of the domain Ω and for any $k = \overline{1, N}$, we have

$$m^\beta_k \to e_k \quad (\textit{i.e.}\ (\mathrm{i_{eff}})^\beta_k \to 1 \ \textit{when}\ e_k \neq 0) \quad \text{as}\, \beta \to 0\,. \tag{3.24}$$

Proof. In Theorem 3.3 it has been proved that $\int\limits_{\Omega} \mid \mu_\beta(x) - \varepsilon(x) \mid \, dx \to 0$ if $\beta \to 0$ (see (3.23)). This means that

$$\int\limits_{\Omega_k} \mid \mu_\beta(x) - \varepsilon(x) \mid \, dx \to 0 \qquad \forall\, \Omega_k\,,$$

which implies (3.24).

■

Thus, we see that explicitly computable quantities m^β_k will provide correct information on the distribution of subdomain errors once y^*_β and the respective parameter β are appropriately defined.

Remark 3.5 It is evident that everywhere above we could replace the function $\mu_\beta(x)$ with the simpler function $\mid \nabla v(x) - y^*_\beta(x) \mid^2$, where $y^*_\beta \to p^*$ in $L^2(\Omega, \mathbb{R}^n)$ as $\beta \to 0$. Theorem 3.2 addresses the possible way of constructing such y^*_β.

4 A posteriori error estimate for functions that do not exactly satisfy the Dirichlet boundary condition

Consider the problem (2.1), (2.2) and assume that v is any function from $H^1(\Omega)$ that satisfies the boundary condition (2.2) only approximately. Then, our aim is to control this additional error by an a posteriori error estimate.

First, we introduce the following notation:

$$m_0^2(v) := \inf_{\tilde{v}\in V_0+u_0} \|\nabla(v-\tilde{v})\|^2 \,, \tag{4.1}$$

$$m_d^2(v,y^*) := \|\nabla v - y^*\|^2 \,, \tag{4.2}$$

$$m_f^2(y^*) := \|\operatorname{div} y^* + f\|^2 \,. \tag{4.3}$$

The required a posteriori error estimate is given by

Theorem 4.1 *Let $u \in V_0 + u_0$ be the solution to the problem (2.1), (2.2) and v any function from $H^1(\Omega)$.*
Then

$$\|\nabla(v-u)\| \leq M(v;y^*) \quad \forall y^* \in H(\Omega,\operatorname{div}) \,, \tag{4.4}$$

where $M(v;y^*) = 2m_0(v) + m_d(v,y^*) + C_\Omega m_f(y^*)$.

Proof. In view of Young's inequality, we have

$$\|\nabla(v-u)\|^2 \leq (1+\frac{1}{\alpha_1})\|\nabla(v-\tilde{v})\|^2 + (1+\alpha_1)\|\nabla(\tilde{v}-u)\|^2 \,, \tag{4.5}$$

where α_1 is any positive number and $\tilde{v}$ any function from $V_0 + u_0$. For the second term on the right-hand side of (4.5) we can apply the estimate (3.1)

$$\|\nabla(\tilde{v}-u)\|^2 \leq (1+\beta)\|\nabla\tilde{v} - y^*\|^2 + (1+\frac{1}{\beta})\, C_\Omega^2\|\operatorname{div} y^* + f\|^2 \tag{4.6}$$

yielding

$$\begin{aligned}\|\nabla(v-u)\|^2 \leq{}& (1+\frac{1}{\alpha_1})\|\nabla(v-\tilde{v})\|^2 + (1+\alpha_1)(1+\beta)\|\nabla\tilde{v}-y^*\|^2 \\ &+ (1+\alpha_1)(1+\frac{1}{\beta})\, C_\Omega^2\|\operatorname{div} y^* + f\|^2 \,. \end{aligned} \tag{4.7}$$

This is valid for all $\tilde{v} \in V_0 + u_0$ and $y^* \in H(\Omega,\operatorname{div})$. The right-hand side represents the error majorant, and, evidently, if $\tilde{v} \to u$ in $H^1(\Omega)$ and $y^* \to \nabla u$ in $H(\Omega,\operatorname{div})$, we can choose $\alpha_1(\tilde{v},y^*)$ so that the majorant will converge to the exact error.

However, it is difficult to minimize the majorant from (4.7) with respect to two functions $\tilde{v}$ and y^*, and we use once again Young's inequality to eliminate the function $\tilde{v}$ from the majorant. Indeed, we have

$$\|\nabla\tilde{v} - y^*\|^2 \leq (1+\frac{1}{\alpha_2})\|\nabla v - y^*\|^2 + (1+\alpha_2)\|\nabla\tilde{v} - \nabla v\|^2 \,, \tag{4.8}$$

where α_2 is any positive number. Inequalities (4.7) and (4.8) imply the estimate

$$\|\nabla(v-u)\|^2 \leq M^2(v;y^*,\alpha_1,\alpha_2,\beta) \,, \tag{4.9}$$

where

$$M^2(v; y^*, \alpha_1, \alpha_2, \beta) := (1 + \frac{1}{\alpha_1} + (1+\alpha_1)(1+\alpha_2)(1+\beta))$$
$$\cdot \inf_{\tilde{v}\in V_0+u_0} \|\nabla(v-\tilde{v})\|^2$$
$$+ (1+\alpha_1)(1+\beta)(1+\frac{1}{\alpha_2})\|\nabla v - y^*\|^2$$
$$+ (1+\alpha_1)(1+\frac{1}{\beta})C_\Omega^2\|\operatorname{div} y^* + f\|^2 \,. \tag{4.10}$$

By using the notations (4.1)–(4.3), the latter inequality can be rewritten as

$$\|\nabla(v-u)\|^2 \le (1 + \frac{1}{\alpha_1} + (1+\alpha_1)(1+\alpha_2)(1+\beta))m_0^2(v)$$
$$+ (1+\alpha_1)(1+\beta)(1+\frac{1}{\alpha_2})m_d^2(v,y^*)$$
$$+ (1+\alpha_1)(1+\frac{1}{\beta})C_\Omega^2 m_f^2(y^*) \,. \tag{4.11}$$

Minimizing the right-hand side of (4.11) over the parameters α_1, α_2 and β we obtain (4.4); the corresponding optimal values of the parameters are given by the formulae

$$\alpha_1 = \frac{m_0(v)}{m_0(v) + m_d(v,y^*) + C_\Omega m_f(y^*)} \,, \tag{4.12}$$

$$\alpha_2 = \frac{m_d(v,y^*)}{m_0(v)} \,, \tag{4.13}$$

$$\beta = \frac{C_\Omega m_f(y^*)}{m_0(v) + m_d(v,y^*)} \,. \tag{4.14}$$

■

According to the above theorem, the error majorant consists of three terms in the general case of a non-conforming approximation. The term $m_0(v)$ is non-negative and vanishes if and only if v exactly satisfies boundary condition (2.2). Thus, it is a measure of the *error in the boundary condition*. The quantity $m_d(v,y^*)$ is a measure of the error in the *duality relation* for the pair (v,y^*) and $m_f(y^*)$ measures the error in the *equation for the dual variable ("equilibrium equation")*.

It is also easy to verify that the estimate (4.4) is *consistent* in the sense that its right-hand side tends to zero when $v \to u$ in $H^1(\Omega)$ and $y^* \to p^*$ in $H(\Omega, \operatorname{div})$.

Remark 4.1 The quantity $m_0^2(v)$ cannot, in general, be computed directly, but one can estimate it from above, for example, by choosing the function $\tilde{v} \in V_0 + u_0$ as $\tilde{v}(x) = \frac{\Phi(x)}{\Phi_0}v(x) + (1 - \frac{\Phi(x)}{\Phi_0})u_0(x)$ in the domain $\{x \in \Omega \,|\, \Phi(x) \le \Phi_0\}$ and $\tilde{v} \equiv v$ on the rest of Ω, where $\Phi(x)$ is the distance from x to Γ and $\Phi_0 > 0$ is some fixed number.

5 Numerical examples

In this section we consider several numerical examples highlighting different features of the proposed error estimator.

In Section 5.1, we investigate systematically the approximation properties of the error majorant in dependence on the parameters such as the polynomial degree of the dual variable and the number of iteration steps for minimizing the majorant. In order to study these dependencies we have chosen a very simple model problem which allows us to vary diverse approximation parameters; the problem is also well-suited to validate the convergence of the error majorant to the exact error, since it is possible to reach the "asymptotic range" of the approximation.

In Section 5.2, we apply the error estimator (with parameter choices found in Section 5.1) to the discretization of the Poisson model problem by the penalty/fictitious-domain method taking into account the error in the approximation of Dirichlet boundary conditions.

We compare our error majorant with some standard residual-based error estimator in Section 5.3 and address the issue of computational complexity for our error estimator in Section 5.4.

5.1 The parameter studies. Error estimation for non-Galerkin approximate solution

5.1.1 Model problem

Consider the 2D model problem:

$$\Delta u = 1 \qquad \text{in } \Omega \tag{5.1}$$
$$u = 1 \qquad \text{on } \Gamma_1 \tag{5.2}$$
$$u = 0 \qquad \text{on } \Gamma_2 \, , \tag{5.3}$$

where Γ_1 is the unit circle with the center at the origin, Γ_2 is the circle with the same center and of the radius 0.5, Ω is the domain bounded by Γ_1 and Γ_2. Since the exact solution in polar coordinates depends only on ρ, i.e. $u = u(\rho)$, it can be found explicitly:

$$u(\rho) = \frac{1}{4}(\rho^2 - \frac{1}{4}) + C_1 \ln(2\rho) \, , \tag{5.4}$$

with $C_1 = 13/(16 \ln 2)$.

As the approximate solution we take

$$v(\rho) = \sum_{k=0}^{n_1} v_k \rho^k \, , \tag{5.5}$$

with $n_1 = 2$, and choose the coefficients v_k ($k = 0, 1, 2$) by the conditions $v(1) = 1$, $v(0.5) = \varepsilon$ and $v'(1) = 0.5$, where ε is some given number. We see that while the Dirichlet boundary condition on the boundary piece Γ_1 is satisfied exactly, the condition on Γ_2 is violated; the magnitude of the corresponding error, i.e. the parameter ε, will be varied in the numerical tests reported below.

The variable y^* is sought as a polynomial

$$y^*(\rho) = \sum_{k=0}^{n_2} y_k^* \rho^k \, , \tag{5.6}$$

and we have chosen different values for the polynomial degree n_2.

In the present case, the term m_0 can be computed directly:

$$m_0^2(v) = \frac{2\pi}{\ln 2} \, v^2(0.5) \, . \tag{5.7}$$

5.1.2 Choice of constant C_Ω

In the light of Remark 3.4, we have evaluated the smallest eigenvalue of the Laplace operator in the circular domain $\tilde{\Omega} = \{x \in \mathbb{R}^2 \mid |x| < 1\}$ ($\tilde{\Omega} \supset \Omega$) to obtain the estimate $C_\Omega \leq 1/3$. We emphasize that the overestimation of the constant is not only harmless for the reliability of the majorant (this is obvious) but, also, does not pollute the majorant's efficiency, see Remark 3.4. The last statement is clearly confirmed by numerical experiments.

5.1.3 Minimization of error majorant M^2

To find a good y^*-approximation, we minimize the "squared error majorant" $M^2(v; y^*, \alpha_1, \alpha_2, \beta)$ (see (4.10)) with respect to the last four variables. The algorithm of the minimization is given below:

$$
\begin{aligned}
&\text{Set } \alpha_1^{(0)} = \alpha_2^{(0)} = \beta^{(0)} = 0.5 \\
&\text{For } n = 1, 2, \ldots : \\
&\quad \text{Minimize } M^2(v; y^*, \alpha_1^{(n-1)}, \alpha_2^{(n-1)}, \beta^{(n-1)}) \ \text{ w.r.t. } y^* \implies y^*_{(n)} \\
&\quad \text{Using } y^*_{(n)}, \text{update } \alpha_1^{(n)}, \alpha_2^{(n)}, \beta^{(n)} \ \text{ with (4.12)–(4.14)}\,,
\end{aligned}
$$

where the minimization problem can be solved in a standard way since the majorant M^2 is a quadratic functional with respect to y^*. The minimization is performed over the set of the polynomial functions (5.6) of fixed degree n_2. Just one or two iterations were found to be sufficient for a reasonable accuracy of the error majorant. Below we will present a detailed study demonstrating the dependence of the majorant quality on the number of iterations.

It is worth noticing that the convergence of the iterational algorithm is not a primary issue here; we use one or two iterations only to improve the y^*-approximation. Indeed, one can easily see that $M^2(v; y^*_{(i)}, \alpha_1^{(i)}, \alpha_2^{(i)}, \beta^{(i)}) \leq M^2(v; y^*_{(j)}, \alpha_1^{(j)}, \alpha_2^{(j)}, \beta^{(j)})$ for any $i \geq j$, and, consequently, $M(v; y^*_{(2)}) \leq M(v; y^*_{(1)})$, if the algorithm does not stop because of the degenerate situation when either $\alpha_2^{(1)}$ or $\beta^{(1)}$ is zero. Thus, $y^*_{(2)}$ is, usually, a better approximation than $y^*_{(1)}$, although the quality of $y^*_{(1)}$ can already suffice in many cases.

We may also note that, if the Dirichlet boundary conditions are satisfied exactly, the term m_0 simply vanishes and the majorant $M^2(v; y^*, \alpha_1, \alpha_2, \beta)$ reduces to the functional $M^2(v; y^*, \beta)$ defined as the right-hand side of inequality (3.1). While one or two iterations will still yield a good result in this case, one can gain a better majorant accuracy by chosing a suitable, moderately small value of $\beta^{(0)}$ (see Remark 3.3); then only one iteration is, in fact, needed.

5.1.4 Choice of polynomial degree for y^* and asymptotic exactness of the majorant

In all tests presented here the iterations of the algorithm were done until

$$
\mid M^2(v; y^*_{(n-1)}, \alpha_1^{(n-1)}, \alpha_2^{(n-1)}, \beta^{(n-1)}) - M^2(v; y^*_{(n)}, \alpha_1^{(n)}, \alpha_2^{(n)}, \beta^{(n)}) \mid < 10^{-4}\,,
$$

so as to eliminate the influence of the iteration error. The numerical results corresponding to three different values of the error ε in the Dirichlet boundary condition on Γ_2 are presented in Figure 1 (left). They clearly indicate that an improvement of the approximation to the y^*-variable leads to a very accurate error estimation by the majorant M. For small values of ε, i.e. when the total error is dominated by the errors in the domain, the asymptotic exactness of the majorant with respect to the y^*-approximation is observed (the effectivity index $M(v; y^*)/\|\nabla e\|$, where $e := v - u$ is the exact error, tends to 1 with increasing polynomial degree n_2); for larger values of ε the influence of the boundary error becomes more pronounced. The overestimation can be possibly explained by the factor 2 in front of the term $m_0(v)$ in the majorant M: it occured as a result of the repeated application of the triangle inequality and, thus, in many cases, exceeds the optimal value of the factor. It is noteworthy, however, that the effectivity index monotonically decreases (down to the value ≈ 1.6 when $\varepsilon = 0.1$) with increasing accuracy of the approximation of y^*; since M cannot be smaller than the norm of the exact error, the reliability of our error estimator is always preserved, i.e. the estimator always provides an upper bound for the norm of the error.

Figure 1 (left) clearly shows that the improvement in the effectivity index caused by the increase of degree n_2 from 1 to 2 is apparently greater than the improvement due to the increase of n_2 from 2 to 3 (the relative decrease of the effectivity index is about 20% in the first and only 5% in the second case). This suggests that the value $n_2 = 2 = n_1$ for the polynomial degree of y^*-variable is optimal from the viewpoint of both the accuracy and computational cost, as a further increase of n_2 does not bring a significant merit.

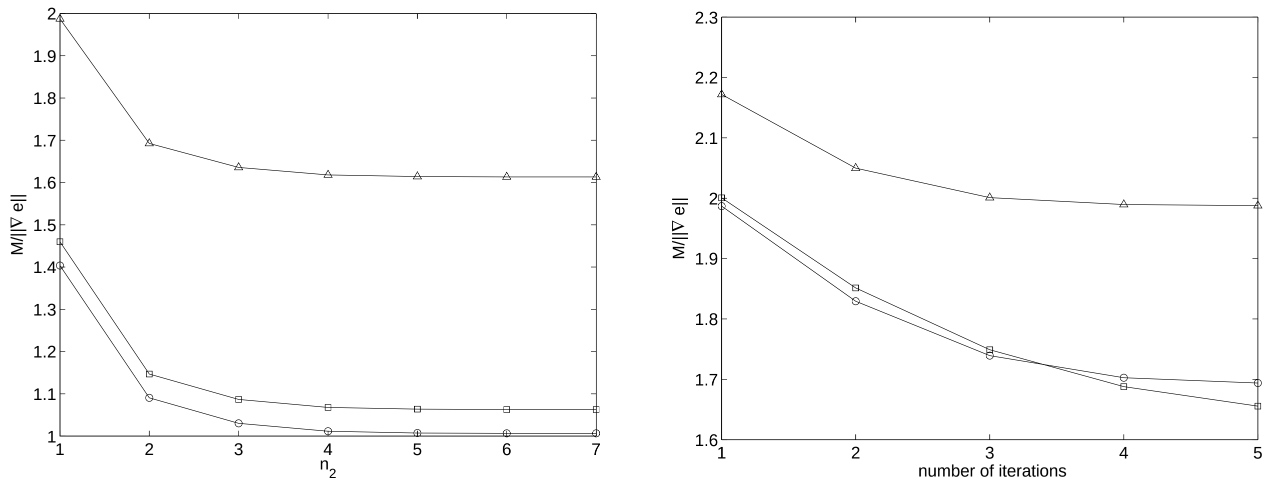


Figure 1: (left) The effectivity index vs. the polynomial degree n_2 of the y^*-approximation: the error in the Dirichlet boundary condition $\varepsilon = 0.1$ ($\triangle$), $\varepsilon = 0.01$ ($\Box$), $\varepsilon = 0.001$ ($\circ$); (right) the effectivity index vs. the number of iterations done for minimization of majorant M^2 (here $\varepsilon = 0.1$): 1st degree y^*-approximation ($\triangle$), 2nd degree ($\circ$), 3rd degree ($\Box$).

5.1.5 Choice of the number of iterations for minimization of the error majorant M^2

In Figure 1 (right) the dependence of the effectivity index on the number of iterations done for minimization of M^2 is shown. We experimented with different initial values $\alpha_1^{(0)}$, $\alpha_2^{(0)}$ and $\beta^{(0)}$ and found that, for any $\alpha_1^{(0)}$, not too small $\alpha_2^{(0)}$ (i.e., roughly, $\alpha_2^{(0)} > 10^{-2}$) and not too large $\beta^{(0)}$ (roughly, $\beta^{(0)} < 10^2$), already the first iteration yields the effectivity index exceeding by only 15% its minimal value for the chosen degree n_2 (this minimal value was normally reached within 5–6 iterations). The second iteration brought further $\approx 8\%$ decrease of the effectivity index, and the relative decrease of the effectivity index due to the third iteration was about 4%. Thus, just one iteration can be recommended in all cases, and the second iteration may be used when one needs an improved accuracy of the error estimation; the third iteration brings too small accuracy improvement to pay off the increased computational cost.

5.1.6 Asymptotic behavior of the weights $\alpha_1, \alpha_2, \beta$ in the majorant

Here we analyze the behavior of the parameters α_1, α_2 and β depending on the polynomial degree n_2 of the y^*-approximation for the case $\varepsilon = 0.1$. The parameters α_1 and α_2 remain nearly constant for all values of n_2, which reflects the fact that the term $m_0^2(v)$ is completely independent of y^* and the term $m_d^2(v, y^*)$ is practically independent of the degree of y^*-approximation (it was observed in numerical experiments). Figure 2 (left) illustrates the behavior of β and of the term $m_f^2(y^*)$ in dependence on the degree n_2. It is clearly seen that the parameter β is monotonically decreasing and becomes very small when the approximation of y^* is sufficiently good; the term $m_f^2(y^*)$ behaves like $\mathcal{O}(\beta^2)$ in a full agreement with the predictions of Theorem 3.2.

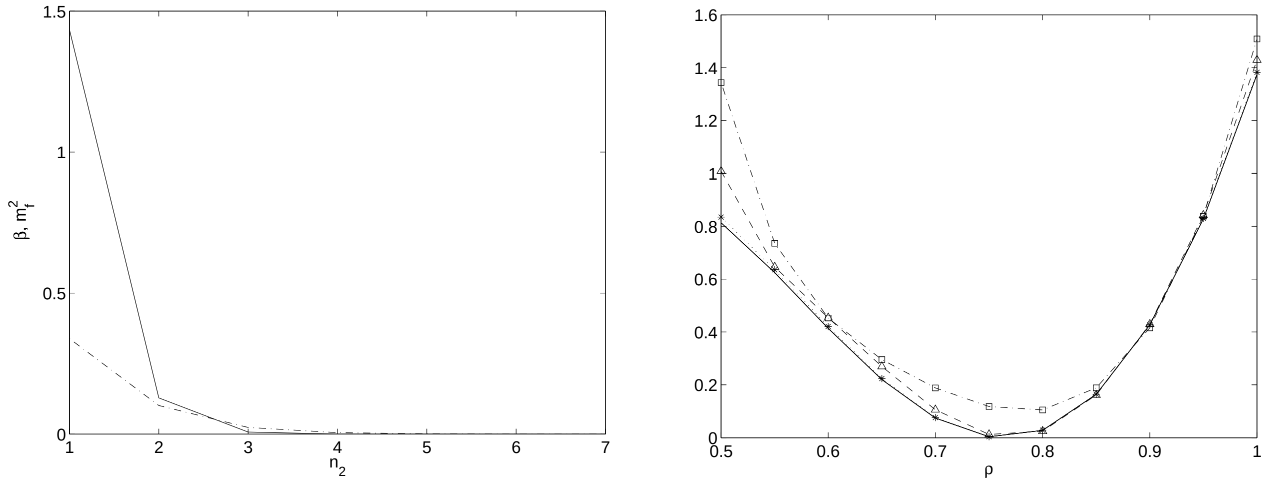


Figure 2: (left) Dependence of the parameter β (dash-dot line) and of the term $m_f^2(y^*)$ (solid line) on the polynomial degree n_2 of the y^*-approximation; (right) the exact energy error (solid line) and its estimate produced by the majorant (drawn as functions of ρ): 2nd degree approximation of y^* ($\square$), 3rd degree ($\triangle$), 5th degree ($*$).

5.1.7 The local error representation by the majorant

The majorant reproduces the local behavior of the exact error with a good quality, as can be seen in Figure 2 (right). The error in the boundary condition on Γ_2 was taken small here ($\varepsilon = 0.001$). In the figure we plot the functions $\mid \nabla v - \nabla u \mid^2$ (the exact energy error) and $(1+\alpha_1)(1+\beta)(1+\frac{1}{\alpha_2}) \cdot \mid \nabla v - y^* \mid^2 + (1+\alpha_1)(1+\frac{1}{\beta})C_\Omega^2 \mid \operatorname{div} y^* + f \mid^2$ (the integrand in the majorant $M^2(v; y^*, \alpha_1, \alpha_2, \beta)$) as functions of the variable ρ. One can see that even with the 2nd degree y^*-approximation the qualitative behavior of the exact energy error is captured correctly, while the 5th degree yields also the quantitatively correct behavior of the error.

5.1.8 Some conclusions concerning the selection of control parameters

Having in mind the obvious trade-off between the accuracy of the error estimation and the computational cost of evaluating the majorant, we can summarize the main results of the presented above numerical experiments:

- it can be recommended to choose the polynomial degree of approximation for the dual variable y^* equal to the degree for the primal variable v;
- one or two iterations are generally sufficient for minimizing the error majorant M^2, and, hence, for obtaining a good y^*-approximation.

In addition, we observed that, if the Dirichlet boundary conditions are satisfied precisely, the asymptotic exactness of the error majorant is achieved with an increase of the polynomial degree for y^*.

5.2 Error estimation for penalty/fictitious-domain finite element method

Consider the problem

$$-\Delta u = f \qquad \text{in } \Omega\,, \tag{5.8}$$
$$u = u_0 \qquad \text{on } \Gamma_1\,, \tag{5.9}$$
$$u = 0 \qquad \text{on } \Gamma_2\,, \tag{5.10}$$

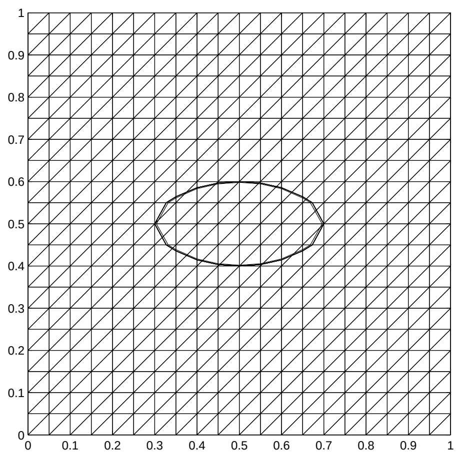
where $\Omega = (0,1)\times(0,1)\setminus\bar{\omega}$, $\omega = \{(x;y)\in\mathbb{R}^2 : \frac{(x-0.5)^2}{(0.2)^2} + \frac{(y-0.5)^2}{(0.1)^2} < 1\}$, $\Gamma_2 = \partial\omega$, Γ_1 is the boundary of the square $(0,1)\times(0,1)$ (see Figure 3), having the exact solution

$$u_{\text{exact}}(x,y) = x\left(1 - e^{-0.5\left(1-\frac{(x-0.5)^2}{(0.2)^2}-\frac{(y-0.5)^2}{(0.1)^2}\right)^2}\right).$$

The solution is shown in Figure 3; the function u_0 is specified as

$$u_0(x,y) = x \quad \text{on } \Gamma_1\,. \tag{5.11}$$

We have to note that u_{exact} satisfies the condition (5.11) with the error $\approx 10^{-8}$ that can be considered negligible in the present numerical tests.



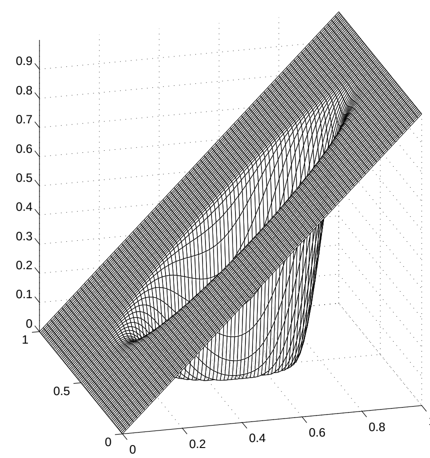


Figure 3: (left) The domain geometry and computational mesh; (right) the exact solution u_{exact}.

To approximate the problem (5.8)–(5.10) we make use of the so-called penalty/fictitious-domain method which is rather popular in computational fluid dynamics community (see, e.g. [3], [4] and the references therein). The penalized problem reads as follows:

$$-\Delta u_\varepsilon + \tfrac{1}{\varepsilon}\chi_\omega\, u_\varepsilon = \tilde{f} \qquad \text{in } (0,1)\times(0,1)\,, \tag{5.12}$$

$$u_\varepsilon = u_0 \qquad \text{on } \Gamma_1\,, \tag{5.13}$$

where χ_ω is the characteristic function of the domain ω, and $\tilde{f}$ is the original right-hand side f extended into ω (for example, by some constant value; we set $\tilde{f} = 30.0$ in ω). The main peculiarity of the method, rendering it as particularly attractive for computing the flows with moving solid objects, is that the mesh does not have to be adjusted to the curvilinear parts of the domain boundary (see Figure 3). However, this convenience is accompanied by the unavoidable error in the approximation of the Dirichlet boundary condition on the curvilinear parts, where the condition is satisfied in a weak sense only. It is also worth noting that the finite element solution of the penalized problem does not, in general, possesses the Galerkin orthogonality property with respect to the original problem (which is a requirement of most existing error estimators).

We start with $\varepsilon = 10^{-7}$, which is small enough to make the "boundary error" (m_0) almost negligible as compared to the "domain error" ($m_d + C_\Omega m_f$), and observe the convergence rate of our majorant with respect to the mesh width in Figure 4 (left). Obviously, the majorant decreases with optimal rate for both $\mathbb{P}_1$ and $\mathbb{P}_2$ finite elements (we used the same approximation for the primal and dual variables and a uniform mesh refinement).

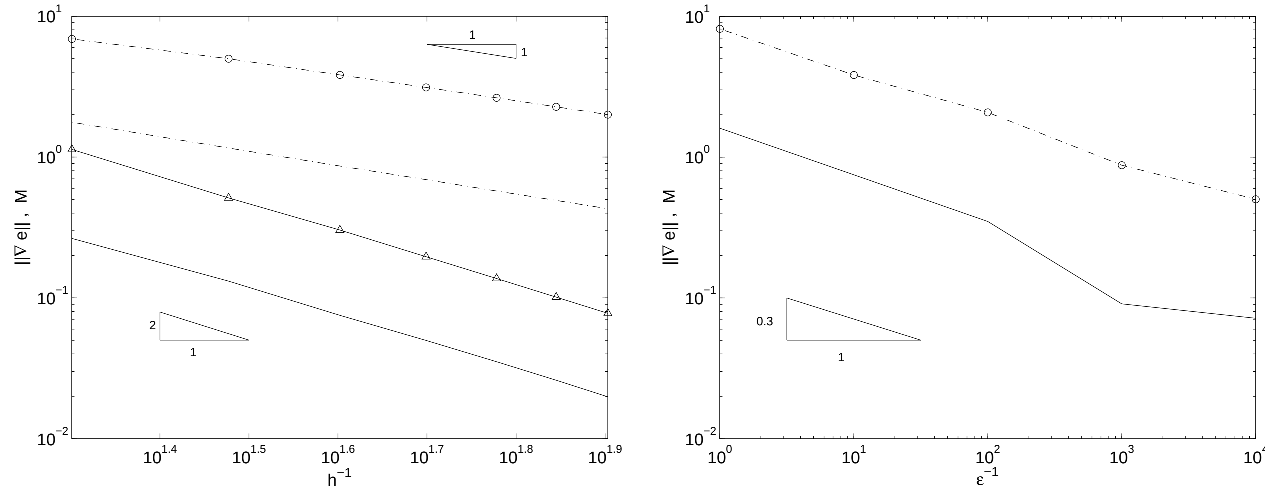


Figure 4: (left) The logarithmic plot of the decrease rates for the exact energy error $\|\nabla e\|$ and for the majorant M ($\varepsilon = 10^{-7}$): the dash-dotted lines correspond to $\mathbb{P}_1$ and the solid lines to $\mathbb{P}_2$ elements, "$\circ$" and "$\triangle$" correspond to the majorant M; (right) the logarithmic plot of the convergence rate for the exact energy error $\|\nabla e\|$ (solid line) and for the majorant M (dash-dotted line) with respect to ε ($\mathbb{P}_1$ elements, $\quad h = 1/240$).

In Figure 5 the local error distribution delivered by the elementwise computation of the m_d term of the majorant is shown. It agrees very well with the exact error distribution. The addition of the residual term $C_\Omega m_f$ slightly spoils the local picture, hence, we use a simple error indicator m_d in accordance with Remark 3.5.

Next, we investigate the influence of the boundary error by taking sufficiently rough penalty parameter, namely $\varepsilon = 10^{-1}$. We can obtain a detailed information on the local distribution of the error near the boundary Γ_2 by computing the term m_0 elementwise in the narrow band around Γ_2 (see Remark 4.1). The evaluated local distribution of the boundary error is depicted in Figure 6 together with the exact error distribution obtained via the $H^{1/2}$-norm of the approximate solution on Γ_2 (the exact solution is zero on Γ_2).

The detailed information delivered by the term m_0 can be used to adaptively improve the approximation near Γ_2 until the boundary error will become smaller than the error in the domain.

In Figure 4 (right) we demonstrate the convergence of the majorant with respect to the penalty parameter ε, when the mesh size is sufficiently small ($h = 1/240$). The convergence rate is slightly better than theoretically predicted $\mathcal{O}(\varepsilon^{1/4})$ (see [3] and [4]). The most important fact is that the majorant obviously provides a reliable upper bound for the exact error even when the error in the approximation of boundary conditions is very large.

5.3 Comparison with the residual error estimator

In order to assess the computational efficiency of our error majorant, we compare it with the standard residual-based error estimator on the model problem in the L-shaped domain. Namely, we consider the problem (2.1), (2.2) with $u_0 \equiv 0$ and $f \equiv 1$ in the domain $\Omega = (-1,1) \times (-1,1) \setminus [0,1] \times [0,1]$ (see Figure 7).

Let $\mathcal{T}$ be a regular triangulation of $\bar{\Omega}$ and v the Galerkin finite element solution to the problem, $v \in \mathcal{S}_0 := \{v \in C^0(\bar{\Omega}) \mid v\big|_T \in \mathbb{P}_1 \ \forall T \in \mathcal{T}\,,\ v = 0 \text{ on } \Gamma\}$. The residual-based a posteriori error

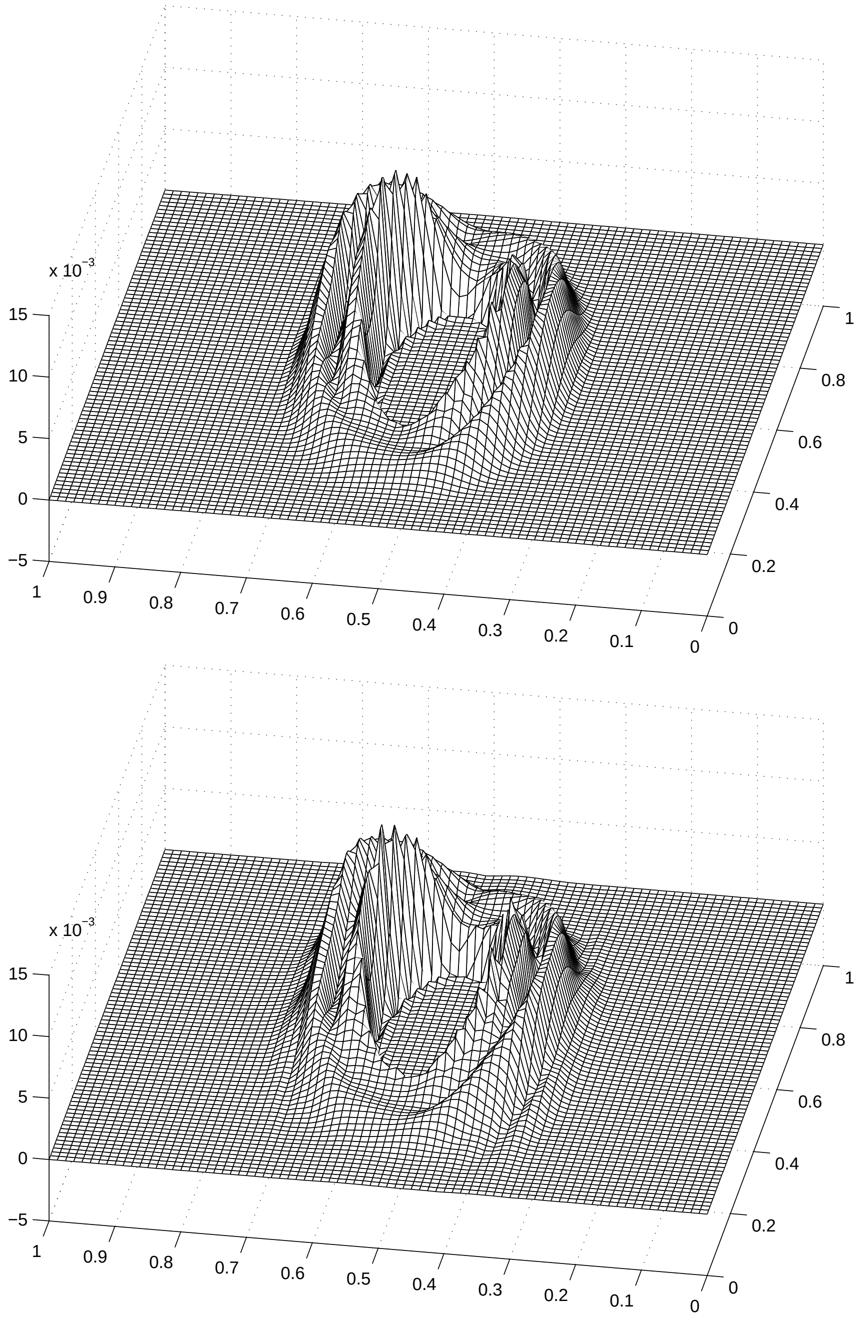
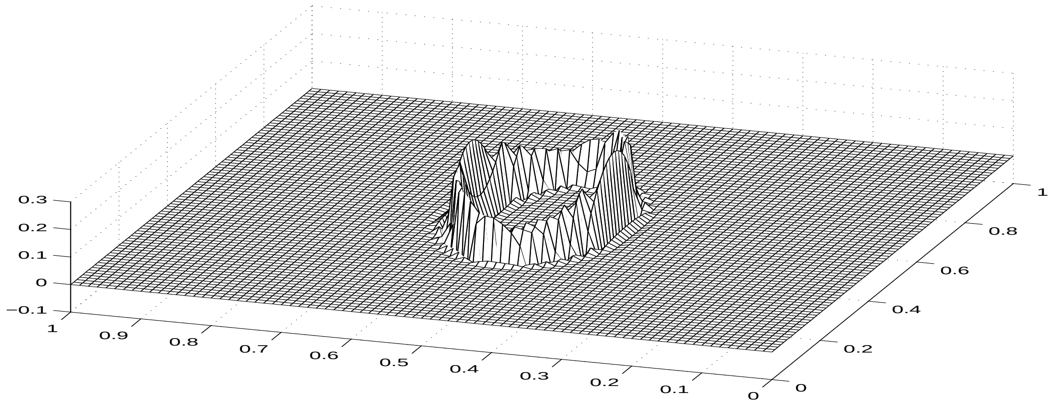


Figure 5: Local error distribution for $\varepsilon = 10^{-7}$ ($\mathbb{P}_1$ elements, $h = 1/80$): exact energy error (top) and the m_d term of the majorant (bottom).



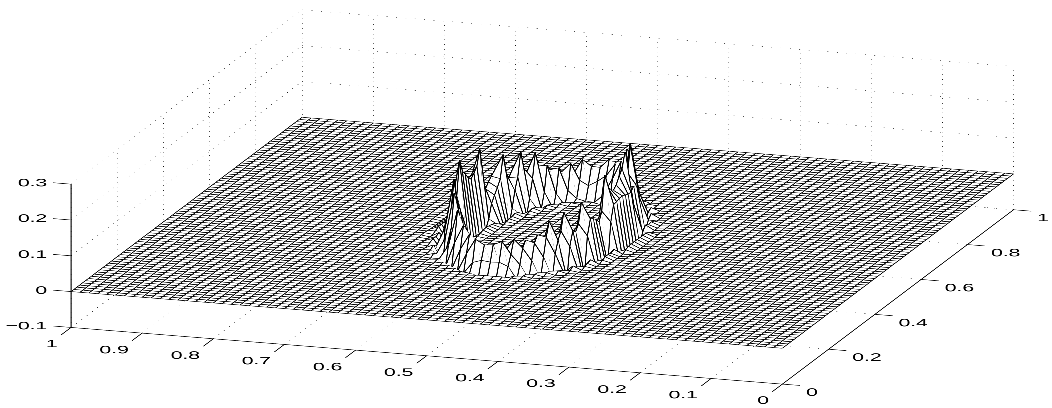


Figure 6: Local boundary-error distribution: exact error (top) and the m_0 term of the majorant (bottom); $\varepsilon = 10^{-1}$, $\mathbb{P}_1$ elements, $h = 1/80$.

estimator (see, e.g., [2], [5], [8], [21]) states

$$\|\nabla(v-u)\| \leq c_1 \left(\sum_{T\in\mathcal{T}} h_T^2 \|f + \Delta v\|^2 \right)^{1/2} + c_2 \left(\sum_{E\in\mathcal{E}} h_E \|J(\nabla v \cdot n_E)\|^2_{L^2(E)} \right)^{1/2}, \tag{5.14}$$

where h_T is the diameter of triangle T, h_E the length of edge E of T ($\mathcal{E}$ being the set of all edges), $J(\dots)$ the jump across E and n_E the normal to E. While the estimator is easy-to-use and became very popular, its real performance strongly depends on the evaluation of the constants c_1 and c_2. In [8] it is shown that c_1 and c_2 in (5.14) can be estimated from the constants in the Clément interpolation estimates (see [11]) which depend on the geometry of Γ and on the triangulation $\mathcal{T}$ (namely, on the shape of the triangles of $\mathcal{T}$). The rigorous analysis of [8] provides the explicit upper bounds for c_1 and c_2; these bounds were used in [8] for evaluating the effectivity index of the estimator (5.14).

We consider here the same model problem as in Section 6 of [8] and evaluate the effectivity index of our error majorant. In accordance with the conclusions of Section 5.1.8, we utilize the equal-order $\mathbb{P}_1$ approximation for the primal variable v and the dual variable y^* and, starting with the initial value of the parameter β equal 0.5, perform only one iteration for the minimization of the majorant M^2.

For a sequence of uniform meshes, Table 1 displays the number N of degrees of freedom, the energy norm of the exact error $e = v - u$, the value of the error majorant $M(v; y^*)$ and the effectivity index $M(v;y^*)/\|\nabla e\|$. The 3rd and the 5th columns of the table illustrate the convergence rate: for two subsequent meshes with N_1 and N_2 degrees of freedom, the 3rd (resp. the 5th) column contains the ratio of $-\log(\|\nabla e_1\|/\|\nabla e_2\|)$ (resp. $-\log(M_1/M_2)$) to $\log(N_1/N_2)$, where the index "1" corresponds to the first and the index "2" to the second mesh. The results of Table 1 indicate that the Galerkin scheme with a uniform mesh refinement converges like $\mathcal{O}(h^{2/3})$, i.e. $\mathcal{O}(N^{-1/3})$, where $h = \mathcal{O}(N^{-1/2})$ is the mesh width; this agrees with the well-known theoretical predictions for the problems with re-entering corner. The error majorant decreases with the same rate and the effectivity index remains between 1.06 and 1.22. This is in a contrast to the effectivity index of the residual-based error estimator (see Table 8 in [8], where the results for exactly the same numbers N are presented); the index of the latter estimator is in the range of 26 to 35. Thus, the residual error estimator overestimates the exact error at least in ≈ 20 times more than our error majorant does.

N	$\|\nabla e\|$	$-\frac{\log(\|\nabla e_1\|/\|\nabla e_2\|)}{\log(N_1/N_2)}$	M	$-\frac{\log(M_1/M_2)}{\log(N_1/N_2)}$	$M/\|\nabla e\|$
5	0.28400		0.3032		1.0676
33	0.15801	0.31	0.1844	0.26	1.1670
161	0.08621	0.38	0.1030	0.36	1.1948
705	0.04757	0.40	0.0579	0.39	1.2172
2945	0.02679	0.40	0.0315	0.42	1.1756
12033	0.01539	0.39	0.0182	0.38	1.1802

Table 1: Numerical results for the uniform mesh refinement.

To improve the convergence rate one can use the adaptive meshes; we have generated those meshes by virtue of the error indicator $m_d\big|_T = \|\nabla v - y^*\|_{L^2(T)}$ (see Remark 3.5). The triangle T was marked for the refinement if $m_d\big|_T \geq 0.5 \max_{T' \in \mathcal{T}} m_d\big|_{T'}$ (the same refinement strategy was chosen in [8]). Figure 7 demonstrates some steps of the mesh refinement process, while the results on the convergence of the exact error and the error majorant are shown in Table 2 (the notations are the same as in Table 1). Now the convergence rate is improved to the optimal $\mathcal{O}(h)$, i.e. $\mathcal{O}(N^{-1/2})$, for both the exact error and the majorant, which justifies the mesh refinement guided by the error indicator $m_d\big|_T$. Since we used some Delaunay-type triangulation algorithm that is different from the one utilized in [8], we could not obtain exactly the same numbers of degrees of freedom as in Table 9 of [8]; however, the range for N in our experiments was roughly the same as in Table 9 of [8]. It is clear from Table 2 that the effectivity index of our majorant stays between 1.06 and 1.2, while the index of the residual-based estimator ranges from 46 to 70 (see [8], Table 9). Hence, with the adaptive meshes, the overestimation of the residual error estimator is at least 38 times greater than the overestimation due to the use of the error majorant.

In [9], the constants c_1 and c_2 were estimated by solving some local eigenvalue problems on the typical patches of the triangles (the configuration of the patches depends, of course, on the geometry of the domain boundary and on the particular triangulation). The computations performed in [9] for the L-shaped domain used the high (typically, 4th) degree elements to solve the local eigenvalue problems and resulted in the improved values for c_1 and c_2 as compared to those in [8]; the corresponding effectivity index of the residual-based error estimator was between 2.5 and 3.0 (resp. 2.5 and 3.8) with the uniform (resp. adaptive) meshes. This shows that the effectivity index of our error estimator is at least 2 times better than the index of the residual-based error estimator, even when the latter estimator is equipped with the more accurate values of the constants c_1 and c_2.

5.4 Computational complexity

In order to address the computational complexity of the presented error estimator we denote by N the total number of degrees of freedom used for the discretization of the primal variable. We assume that both, the linear system for the primal and for the dual variable, can be solved in linear

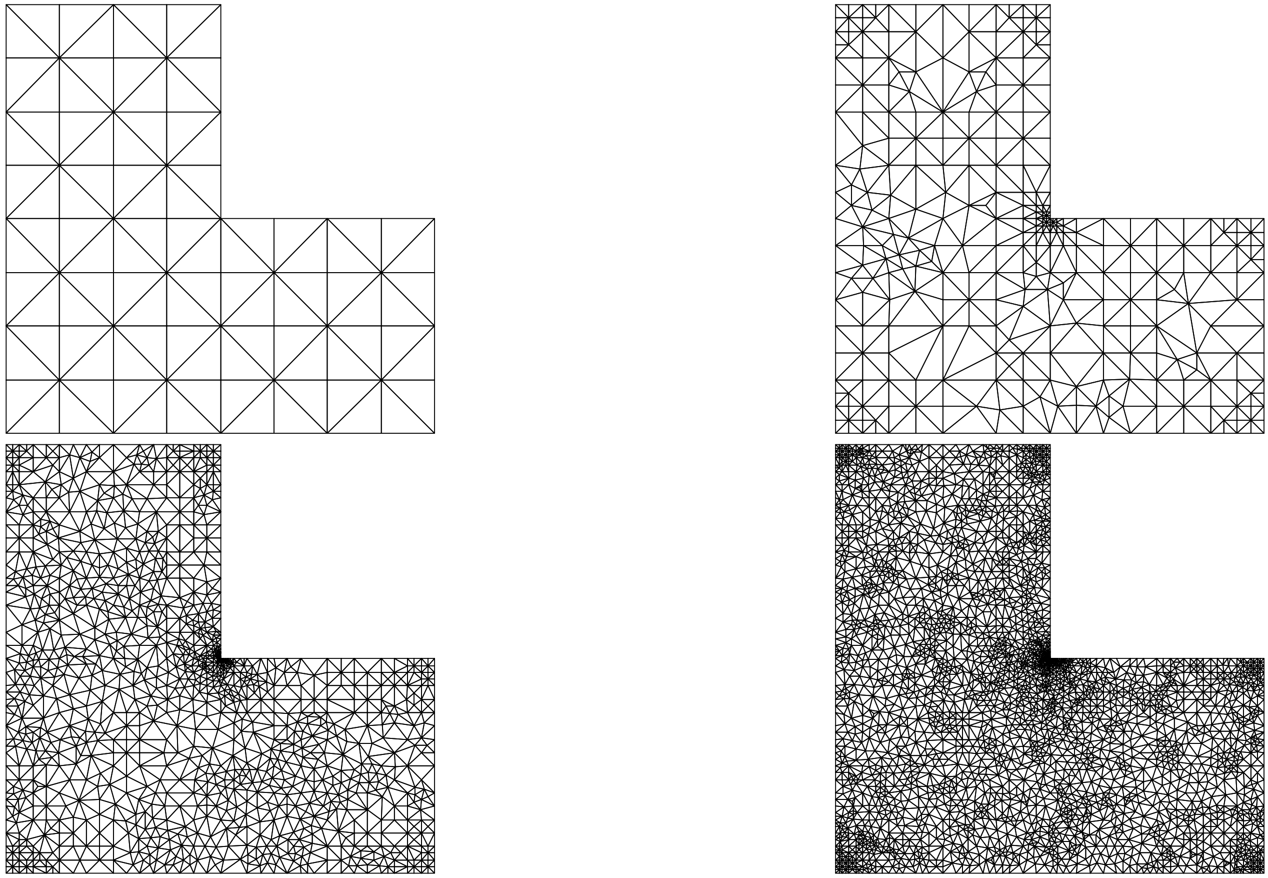


Figure 7: The process of mesh adaptation.

complexity, e.g. by the multigrid method. The polynomial degrees of approximation for the primal and dual variables are chosen to be equal. Hence, the dimension of the dual system is $2N$ and the solver requires $C \cdot 2 \cdot N$ arithmetic operations for the dual and $C \cdot N$ operations for the primal system.

The iterative algorithm suggested in Example 5.1 for minimizing M^2 has been shown to yield a reasonably accurate error majorant after 1–2 iterations. The same number of iterations has been also used in other numerical examples. Since each iteration requires a solution of the linear system for the y^*-variable, the total number of arithmetic operations needed for evaluation of our error estimator ranges between $2CN$ and $4CN$. Thus, we may conclude that the computational cost of evaluating our error majorant is from 2 to 4 times higher than the cost of computing the solution to the original problem.

Although, at the first glance, the proposed error estimator may seem to be relatively costly, the example of Section 5.3 shows that our estimator may be even cheaper than the standard residual-based error estimator, if we compare the amount of work needed to guarantee a prescribed accuracy of the solution. This is a consequence of the absence of any discretization-dependent constants in the estimator: the presented estimator contains a single constant stemming from Friedrichs' inequality and, hence, depending only on the domain geometry. The constant can be easily estimated from above and the estimator is quite non-sensitive to the constant overestimation.

6 Conclusions

The most salient feature of the presented error estimator is its independence of the discretization method; it may work with any approximate solution that belongs to the Sobolev space H^1. Moreover,

N	$\|\nabla e\|$	$-\frac{\log(\|\nabla e_1\|/\|\nabla e_2\|)}{\log(N_1/N_2)}$	M	$-\frac{\log(M_1/M_2)}{\log(N_1/N_2)}$	$M/\|\nabla e\|$
5	0.28400		0.3032		1.0676
33	0.15801	0.31	0.1844	0.26	1.1670
94	0.09561	0.48	0.1135	0.46	1.1876
167	0.07175	0.50	0.0823	0.55	1.1476
348	0.05230	0.43	0.0612	0.41	1.1709
1217	0.03173	0.40	0.0368	0.41	1.1594
1897	0.02557	0.49	0.0306	0.42	1.1972
4069	0.01728	0.51	0.0202	0.54	1.1675
5166	0.01524	0.52	0.0179	0.51	1.1776
9591	0.01112	0.51	0.0131	0.50	1.1802
11566	0.01002	0.55	0.0120	0.47	1.1927
13624	0.00917	0.54	0.0110	0.53	1.1996

Table 2: Numerical results for the adaptive mesh refinement.

we emphasize that our error estimator is by no means restricted to the discretizations obeying the Galerkin orthogonality and the Dirichlet boundary conditions as it is required for most other error estimators. It is also worth noting that the estimator does not use any regularity assumptions concerning the exact solution.

The other remarkable strength of the proposed error estimator seems to be the guaranteed upper bound for the exact error. The feature is extremely important for providing the stopping criterion of an adaptive process as well as for an assurance that the approximate solution has been computed with the accuracy not worse than the prescribed one. While the indication of the local error distribution needed for an adaptivity is a strong side of many existing error estimators, the most of them may fail to provide an upper bound for the exact error when they work with some particular (rough) approximations. In this respect, the proposed error estimator may be rendered as trully reliable, since it always provides an upper bound for the exact error. Thus, the estimator can be utilized as both the error indicator and the stopping criterion for a mesh adaptation.

The future research directions are the extension of the presented method to more general boundary value problems and its combination with more general finite element discretizations (see [14]).

References

[1] Ainsworth, M., Kelly, D.W.: A posteriori error estimators and adaptivity for finite element approximation of the non-homogeneous Dirichlet problem. Advances in Comp. Math. *15*, no. 1-4, 3–23 (2001).

[2] Ainsworth, M., Oden, J.T.: A Posteriori Error Estimation in Finite Element Analysis. John Wiley & Sons, New York, 2000.

[3] Amiez, G., Gremaud, P.-A.: On a penalty method for the Navier-Stokes problem in regions with moving boundaries. Comp. Appl. Math. *12*, no. 2, 113–122 (1993).

[4] Angot, Ph.: Analysis of singular perturbations on the Brinkman problem for fictitious domain models of viscous flows. Math. Meth. Appl. Sci. *22*, 1395–1412 (1999).

[5] Babuška, I., Strouboulis, T.: The Finite Element Method and its Reliability. Clarendon Press, Oxford, 2001.

[6] Becker, R., Rannacher, R.: A feed-back approach to error control in finite element methods: basic analysis and examples. East-West J. Numer. Math. *4*, no. 4, 237–264 (1996).

[7] Carstensen, C., Bartels, S.: Each averaging technique yields reliable a posteriori error control in FEM on unstructured grids. Part I: Low order conforming, nonconforming and mixed FEM. Math. Comp. *71*, no. 239, 945–969 (2002).

[8] Carstensen, C., Funken, S.A.: Constants in Clément-interpolation error and residual based a posteriori error estimates in finite element methods. East-West J. Numer. Math. *8*, no. 3, 153–175 (2000).

[9] Carstensen, C., Funken, S.A.: Fully reliable localized error control in the FEM. SIAM. J. Sci. Comput. *21*, no. 4, 1465–1484 (2000).

[10] Carstensen, C., Sauter, S.A.: A posteriori error analysis for elliptic PDEs on domains with complicated structures. Preprint no. 10-01, Institut für Mathematik, Universität Zürich, 2001. URL : http://www.math.unizh.ch/index.php?preprint

[11] Clément, P.: Approximation by the finite element functions using local regularization. RAIRO Sér. Rouge Anal. Numér. *9*, 77–84 (1975).

[12] Dörfler, W., Rumpf, M.: An adaptive strategy for elliptic problems including a posteriori controlled boundary approximation. Math. Comp. *67*, no. 224, 1361–1382 (1998).

[13] Glowinski, R., Pan, T.-W., Hesla, T.I., Joseph, D.D., Periaux, J.: A distributed Lagrange multiplier/fictitious domain method for the simulation of flow around moving rigid bodies: application to particulate flow. Comput. Methods Appl. Mech. Engrg. *184*, no. 2-4, 241–267 (2000).

[14] Hackbusch, W., Sauter, S.A.: Composite finite elements for the approximation of PDEs on domains with complicated micro-structures. Numer. Math. *75*, 447–472 (1997).

[15] Ekeland, I., Temam, R.: Convex Analysis and Variational Problems. North-Holland, New York, 1976.

[16] Eriksson, K., Estep, D., Hansbo, P., Johnson, C.: Introduction to adaptive methods for differential equations. Acta Numerica, 105–158 (1995).

[17] Mikhlin, S.G.: Variational Methods in Mathematical Physics. Pergamon Press, Oxford, 1964.

[18] Repin, S.I.: A posteriori error estimation for nonlinear variational problems by duality theory. Zapiski Nauchnih Seminarov, V.A. Stekov Mathematical Institute (POMI) *243*, 201–214 (1997).

[19] Repin, S.I.: A posteriori error estimation for variational problems with uniformly convex functionals. Math. Comp. *69*, no. 230, 481–500 (2000).

[20] Strang, G., Fix, G.J.: An Analysis of the Finite Element Method. Prentice-Hall, New Jersey, 1973.

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

[22] Zienkiewicz, O., Zhu, J.: A simple error estimator and adaptive procedure for practical engineering analysis. Int. J. Num. Meth. Engrg. *24*, 337–357 (1987).

[23] Zienkiewicz, O., Zhu, J.: The superconvergent patch recovery and a posteriori error estimates. Parts 1 and 2. Int. J. Num. Meth. Engrg. *33*, 1331–1382 (1992).

Sergey Repin
V.A. Steklov Mathematical Institut, Fontanka 27, 191 011 St. Petersburg, Russia.
E-mail address: repin@pdmi.ras.ru

Stefan Sauter
Institut für Mathematik, Universität Zürich, Winterthurerstr. 190, CH-8057, Zürich, Switzerland.
E-mail address: stas@amath.unizh.ch

Anton Smolianski
Institut für Mathematik, Universität Zürich, Winterthurerstr. 190, CH-8057, Zürich, Switzerland.
E-mail address: antsmol@amath.unizh.ch