A Posteriori Error Estimates for Finite Volume Approximations

S. Cochez-Dhondt [a]**, S. Nicaise** [a] [1] **and S. Repin** [b]

[a] Université de Valenciennes et du Hainaut Cambrésis, LAMAV, FR CNRS 2956, ISTV
F–59313 - Valenciennes Cedex 9, France
[b] Steklov Institute of Mathematics in St. Petersburg, Fontanka 27, 191023, St. Petersburg, Russia

Abstract. We present new a posteriori error estimates for the finite volume approximations of elliptic problems. They are obtained by applying functional a posteriori error estimates to natural extensions of the approximate solution and its flux computed by the finite volume method. The estimates give guaranteed upper bounds for the errors in terms of the primal (energy) norm, dual norm (for fluxes), and also in terms of the combined primal–dual norms. It is shown that the estimates provide sharp upper and lower bounds of the error and their practical computation requires solving only finite–dimensional problems.

Key words: finite volume methods, elliptic problems, a posteriori error estimates of the functional type
AMS subject classification: 65N30

1.   Introduction

The finite volume method is a well adapted method for the discretization of various partial differential equations and is very popular in the engineering community [23]. The mathematical community started recently to analyse it in detail. Presently existence and uniqueness results as well as a priori error estimates are available for a quite large class of problems; we refer to [9] and the references cited there. Contrary to the finite element methods [2, 4, 29], a posteriori error estimates for finite volume methods are less developed and only few results are obtained in that

---

[1]Corresponding author. E-mail: snicaise@univ-valenciennes.fr

direction. See [12, 21, 1, 10, 11, 18, 19, 30] for cell centered finite volume methods, [15, 16, 22] for vertex-centered methods, and [5, 13, 14] for finite volume element methods.

The goal of our paper is to present the natural framework for functional a posteriori estimates of some cell centered finite volume methods for linear elliptic equations. In a first attempt we restrict ourselves to a diffusion equation in $\mathbb{R}^2$. The case of diffusion-convection-reaction equations is postponed to a forthcoming paper and will be based on some ideas developed in [20]. The key idea is to obtain error bounds in a functional framework as presented in [17] (see also [24, 25, 26, 27, 28]). This point of view yields error majorants in a general framework. As particular case we reconstruct interpolants of the finite volume solution (based on piecewise polynomials), and of its gradient (using Raviart-Thomas elements), this last one has as principal property that the mean of its flux through any edge of the mesh is equal to the numerical flux through that edge.

The idea to interpolate the finite volume solution by a smoother function having the above mentioned property on the flux was presented in [10] in a L^1 framework for time-dependent non-linear convection-diffusion equations in $\mathbb{R}^d \times \mathbb{R}^+$. In that paper the authors obtain a reliable estimator in a L^1-norm, instead of the energy norm. Furthermore their interpolant is a piecewise linear Lagrange interpolant on a dual mesh. More recently, in [18], the author proposes a residual estimator based on the use of a reconstructed approximation of the finite volume solution with the help of a Morley interpolant and this is extended to the convection-diffusion-reaction problems in [19]. Finally in [30] the author considers convection-diffusion-reaction equations, presenting a general inhomogeneous diffusion tensor, and builds an a posteriori error estimator by using a non conforming reconstruction of the approximated solution.

Here, we propose to directly compute our fluxes by using the cell-centered finite volume scheme and apply the functional approach to a posteriori error estimation to obtain our lower and upper bounds.

The outline of the paper is as follows: in section 2. we recall our boundary value problem. In section 3. we describe the so-called cell centered method for the diffusion equation on a mesh made of triangles. Section 4. is devoted to the descriptions of the different error estimates. Finally in section 5. we present various numerical experiments that confirm our theoretical considerations.

2.   Basic problem

Let $\Omega \in \mathbb{R}^2$ be a bounded domain with a Lipschitz boundary Γ. Consider the problem

$$-\mathrm{div}(\alpha(x)\nabla u) = f \qquad \text{in } \Omega, \tag{2.1}$$

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

$$\alpha\nabla u \cdot \nu = g_N \qquad \text{on } \Gamma_N. \tag{2.3}$$

Here Γ_D and Γ_N are two measurable nonintersecting parts of Γ associated with the Dirichlet and Neumann boundary conditions, respectively, and

$$u_0 \in H^1(\Omega), \quad f \in L^2(\Omega), \quad g_N \in L^2(\Gamma_N), \tag{2.4}$$

$$\alpha \in L^\infty(\Omega), \quad 0 < \alpha_\ominus \le \alpha(x) \le \alpha_\oplus < +\infty, \ \forall x \in \Omega, \tag{2.5}$$

are the given data. The respective integral relation

$$\int_\Omega \alpha \nabla u \cdot \nabla w \, dx = \int_\Omega f w \, dx + \int_{\Gamma_N} g_N w \, ds, \qquad \forall w \in V_0, \tag{2.6}$$

defines the generalized solution $u \in V_0 + u_0$, where

$$V_0 := \left\{ w \in H^1(\Omega) \mid w = 0 \text{ on } \Gamma_D \right\}$$

and $\cdot$ denotes the scalar product of vectors. In what follows $\| \, . \, \|$ stands for the L^2–norm.

3.   Finite volume approximations

Let Ω be divided into a collection of simplicial cells K_i, $i = 1, 2, ...N$ that we suppose to be triangles. We denote by Γ_i the boundary of an element K_i and by ν_i its outward normal vector. The finite volume method is based on the *conservation law principle*

$$-\int_{K_i} \text{div} p \, dx = -\int_{\Gamma_i} p \cdot \nu_i \, ds = \int_{K_i} f \, dx, \tag{3.1}$$

which holds for the true flux $p = \alpha \nabla u$ on each cell that has no common boundary with Γ_N and on a modified relation

$$-\int_{\Gamma_i} p \cdot \nu_i \, ds = \int_{K_i} f \, dx + \int_{\Gamma_{Ni}} g_N \, ds \tag{3.2}$$

for a cell that has common boundary Γ_{Ni} with Γ_N.

For any element K_i, α_i denotes the mean value of the function α on K_i, namely

$$\alpha_i = \frac{1}{|K_i|} \int_{K_i} \alpha(x) \, dx.$$

In the finite volume method, the system of linear equations is constructed on the basis of (3.1) and (3.2) and a special presentation of the normal fluxes throughout the values of the approximation u_h on the cells. In the simplest case, we assume that $u_h \in P^0(K_i)$. For the control volume K_i, we select a certain point x_i called "cell center" (see Fig. 1) where u_i denotes $u_h(x_i)$.

The flux along an edge of the triangulation is approximated using the following scheme (see [9]) of the finite volume method:
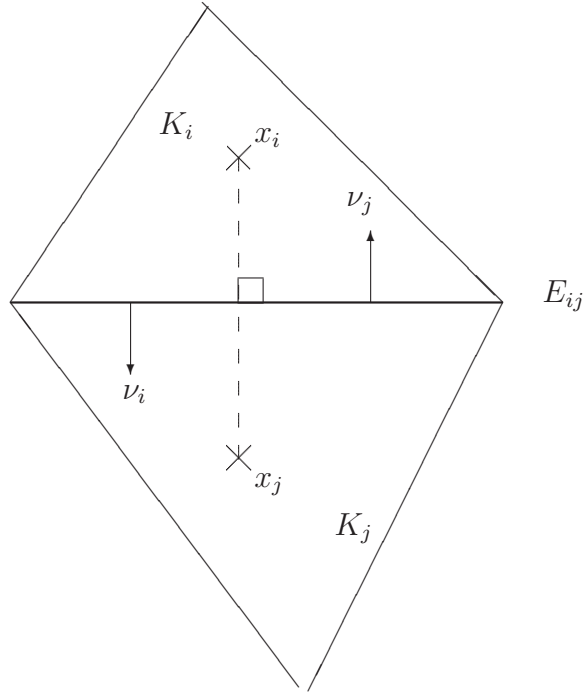
Figure 1:

Let E_{ij} be a fixed interior edge of the triangulation, shared by two elements K_i and K_j of respective centers x_i and x_j. If we suppose that this edge E_{ij} is orthogonal to the straightline joining x_i and x_j, then

$$-\int_{E_{ij}} \alpha \nabla u \cdot \nu_i \, ds \cong -\tau_{E_{ij}}(u_j - u_i) = F_{K_i, E_{ij}}, \tag{3.3}$$

where

$$\tau_{E_{ij}} := |E_{ij}| \frac{\alpha_i \alpha_j}{\alpha_i d_{K_j, E_{ij}} + \alpha_j d_{K_i, E_{ij}}}$$

and $d_{K_i, E_{ij}}$ denotes the minimal distance between the center x_i and the edge E_{ij}. If we suppose that an edge E_{ik} is not orthogonal to the straightline joining the centers x_i and x_k of the two elements K_i and K_k sharing that edge, then we consider the straightline orthogonal to the edge E_{ik} that comes from the intersection between (x_i, x_k) and E_{ik}. If we look at the element K_i, this straigthline has an intersection with one of the two segments joining x_i with the centers of the others neighbours of K_i. This intersection is a point denoted by x_{E_{ik}, K_i} (see Fig. 2) and it is constructed by linear combination from x_i and the involved center. The value of the solution is approximated in that point by u_{E_{ik}, K_i}. Then

$$-\int_{E_{ik}} \alpha \nabla u \cdot \nu_i \, ds \cong -\tau_{E_{ik}}(u_{E_{ik}, K_k} - u_{E_{ik}, K_i}) = F_{K_i, E_{ik}}, \tag{3.4}$$

where

$$\tau_{E_{ik}} = |E_{ik}| \frac{\alpha_i \alpha_k}{\alpha_i d_{K_k, E_{ik}} + \alpha_k d_{K_i, E_{ik}}},$$

and $d_{K_i, E_{ik}}$ stands for the minimal distance between the point x_{E_{ik}, K_i} and the edge E_{ik}, and

$$u_{E_{ik}, K_i} := \frac{d_{K_i, E_{ik}} u_i + d_{K_k, E_{ik}} u_k}{d_{K_i, E_{ik}} + d_{K_k, E_{ik}}}.$$

If an edge E_{Di} of an element K_i belongs to the Dirichlet boundary of the domain, then we denote by $d_{K_i, E_{Di}}$ the distance between the center x_i and $\bar{x}_i$ where $\bar{x}_i$ is defined by the relation

$$|x_i - \bar{x}_i| \leq |x_i - x| \quad \forall x \in E_{Di}.$$

Along this edge, the flux is approximated by

$$-\int_{E_{Di}} \alpha \nabla u \cdot \nu_i \, ds \cong -\tau_{E_{Di}}(u_0 - u_i) = F_{K_i, E_{Di}}, \tag{3.5}$$

where $\tau_{E_{Di}} = |E_{Di}| \frac{\alpha_i}{d_{K_i, E_{Di}}}$.

If we are along an edge E_{Ni} that belongs to the Neumann boundary, since

$$-\int_{E_{Ni}} \alpha \nabla u \cdot \nu_i \, ds = -\int_{E_{Ni}} g_N \, ds, \tag{3.6}$$

the flux is given by:

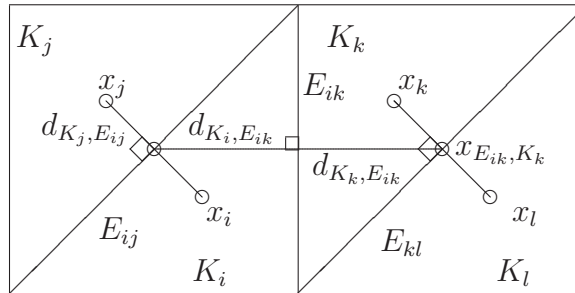$$F_{K_i, E_{Ni}} = -\int_{E_{Ni}} g_N \, ds. \tag{3.7}$$

Figure 2: Finite Volume scheme

Equations (3.1) and (3.2) on a given element K_i of the triangulation then lead to the following finite volume scheme:

$$-\sum_{E_{ij}\subset\Gamma_i} F_{K_i,E_{ij}} = \int_{K_i} f\,dx, \tag{3.8}$$

$$-\sum_{E_{ij}\subset\Gamma_i} F_{K_i,E_{ij}} = \int_{K_i} f\,dx + \int_{E_{Ni}} g_N\,ds. \tag{3.9}$$

Then, the numerical flux associated with an internal edge E_{ij} common to K_i and K_j is given by

$$q_{E_{ij}} := -F_{K_i,E_{ij}}.$$

On Γ_{Ni} the respective normal flux is given by the formula

$$q_{E_{Ni}} = \int_{E_{Ni}} g_N\,ds.$$

On Γ_D the flux is defined as

$$q_{E_{Di}} := -F_{K_i,E_{Di}}$$

Thus, the approximations produced by the finite volume method are presented by a set of piecewise constant function

$$u_h(x) = u_i \qquad \text{for } x \in K_i$$

and the set

$$\mathfrak{Q}_h := \{q_{ij} \mid q_{ij} \in P^0(E_{ij})\}$$

of normal fluxes on the edges.

By these data we can construct a pair of functions

$$(\widetilde{u}_h, \widetilde{q}_h) \in H^1(\Omega) \times H(\Omega, \mathrm{div})$$

viewed as approximations of u and p, respectively.

Let $\mathfrak{P}_k$ be the patch related to a common node k. We define the value of $\widetilde{u}_h$ at the interior node k as follows:

$$\widetilde{u}_h(x_k) := \frac{\displaystyle\sum_{s=i_1,\ldots i_{m_k}} |K_s| u_s}{\displaystyle\sum_{s=i_1,\ldots i_{m_k}} |K_s|},$$

where m_k is the number of elements in the patch $\mathfrak{P}_k$. Inside K_i the function $\widetilde{u}_h$ is defined as the affine function having the above defined values at all the nodes. For a boundary node k, we take

$$\widetilde{u}_h(x_k) := u_0(x_k),$$

assuming that $u_{0|\Gamma} \in C(\Gamma)$.

The function $\widetilde{q}_h$ is defined by the extension of edge fluxes inside K_i with the help of Raviart–Thomas elements of the lowest order, denoted by RT_0.

4.   A posteriori error estimates

To measure the quality of the obtained approximate solutions we apply the so–called functional a posteriori estimates (see [24, 25, 26, 27, 28]; a consequent exposition of the theory is presented in [17]). These estimates were derived by purely functional arguments without attracting Galerkin orthogonality or some other special properties of an approximate solution. Therefore, they are valid for any conforming approximations of a boundary–value problem under consideration. We will exploit the latter property of functional a posteriori estimates and apply them to the approximations obtained as extensions of the FV solutions.

4.1.   A posteriori error estimate for the primal variable

For the problem (2.1)–(2.3) with mixed Dirichlet–Neumann boundary conditions the respective error majorant is presented by the following theorem the proof of which can be found in the above-cited publications (see for instance section 6.4.3 of [17]):

Theorem 1. *Let v be an arbitrary element in $V_0 + u_0$. Then,*

$$\|\nabla(u-v)\|_\alpha^2 := \int_\Omega \alpha |\nabla(u-v)|^2 dx \leq \mathfrak{M}_u^2(v,y), \tag{4.1}$$

where

$$\mathfrak{M}_u(v,y) := \|\alpha\nabla v - y\|_{1/\alpha} + C_\Omega\sqrt{\|f + \mathrm{div}y\|_{1/\alpha}^2 + \|g_N - y\cdot\nu\|_{1/\alpha,\Gamma_N}^2}$$

is the majorant of the error (evaluated in terms of the primal energy norm), in which y is an arbitrary function in the space $H(\Omega, \mathrm{div})$, and C_Ω is any constant greater than $\lambda_1^{-1}(\Omega, \Gamma_N)$, where

$$\lambda_1^2(\Omega, \Gamma_N) = \inf_{w\in V_0} \frac{\|\nabla w\|_\alpha^2}{\|w\|_\alpha^2 + \|w\|_{\alpha,\Gamma_N}^2}, \tag{4.2}$$

where clearly

$$\|w\|_{\alpha,\Gamma_N}^2 := \int_{\Gamma_N} \alpha |w(x)|^2 ds.$$

If $\Gamma = \Gamma_D$ then C_Ω coincides with the constant $C_F(\Omega)$ in the Friedrichs type inequality

$$\|w\| \leq C_F(\Omega)\|\nabla w\|_\alpha \tag{4.3}$$

and (4.1) comes in a simplified form

$$\|\nabla(u-v)\|_\alpha \leq \mathfrak{M}_u(v,y) := \|\alpha\nabla v - y\|_{1/\alpha} + C_F(\Omega)\|f + \mathrm{div}y\|_{1/\alpha}. \tag{4.4}$$

Note further that if $y \in H(\Omega, \mathrm{div})$ is chosen such that $y \cdot \nu = g_N$ on Γ_N, then the constant C_Ω can be chosen greater than $\bar{\lambda}_1^{-1}(\Omega, \Gamma_N)$, where

$$\bar{\lambda}_1^2(\Omega, \Gamma_N) = \inf_{w\in V_0} \frac{\|\nabla w\|_\alpha^2}{\|w\|_\alpha^2}. \tag{4.5}$$

Let us apply this estimate to the finite volume approximations. For the sake of simplicity, we assume that the trace of u_0 on Γ_D is piecewise affine and continuous function (then $\widetilde{u}_h \in V_0 + u_0$). In this case, we use (4.1) and obtain

$$\|\nabla(u - \widetilde{u}_h)\|_\alpha \quad \leq \quad \|\alpha\nabla\widetilde{u}_h - \widetilde{q}_h\|_{1/\alpha} + C_\Omega\sqrt{\|f+\operatorname{div}\widetilde{q}_h\|^2_{1/\alpha} + \|g_N - \widetilde{q}_h \cdot \nu\|^2_{1/\alpha,\Gamma_N}}, \quad (4.6)$$

where $\widetilde{q}_h \in H(\Omega, \operatorname{div})$ and has a square summable trace on Γ_N. We note that the right–hand side of (4.6) is directly computable and provides a guaranteed upper bound of the error computed in terms of the primal variable (i.e., $\nabla(u - \widetilde{u}_h)$). If g_N is piecewise constant on the edges that belong to Γ_N then (in view of the definition of $Q_{E_{Ni}}$) we have $\widetilde{q}_h \cdot \nu = g_N$ on Γ_N and therefore the very last term in (4.6) vanishes. Thus, in this case, the upper bound is given by the estimate

$$\|\nabla(u - \widetilde{u}_h)\|_\alpha \leq \|\alpha\nabla v - y\|_{1/\alpha} + C_\Omega\|f+\operatorname{div}y\|_{1/\alpha}, \quad (4.7)$$

which is similar to (4.4) but with a different factor at the second term. In the finite volume method, the second term in the right-hand side of (4.6) has a simple and easily computable form. Indeed,

$$\|g_N - \widetilde{q}_h \cdot \nu\|^2_{1/\alpha,\Gamma_N} = \sum_{E_{Ni}\in\Gamma_N}\int_{E_{Ni}} \frac{1}{\alpha}|g_N - \{\!| g_N |\!\}_{E_{Ni}}|^2 ds,$$

where $\{\!| g |\!\}_S$ stands for the mean value of g on a set S. Analogously, on any element K_i that have no common boundaries with Γ_N the fluxes defined by (3.2) are such that the respective RT_0 extension has its divergence equal to $\{\!| f |\!\}_{K_i}$. Therefore,

$$\|f + \operatorname{div}\widetilde{q}_h\|^2_{K_i} = \int_{K_i} |f - \{\!| f |\!\}_{K_i}|^2 dx$$

and we observe that if the approximations $\widetilde{u}_h$ and $\widetilde{q}_h$ are computed sharply (i.e., do not contain computational errors) then the second part of the majorant is defined by the respective oscillation terms.

4.2. A posteriori error estimate for the dual variable

Approximation errors related to the dual variable (i.e., to the flux) can be measured in the dual energy norm $\|q\|_{1/\alpha}$ or in $H(\Omega, \operatorname{div})$-type norms

$$\|q\|_{\operatorname{div}} := \|q\|_{1/\alpha} + \sqrt{\|\operatorname{div}q\|^2_{1/\alpha} + \|q \cdot \nu\|^2_{1/\alpha,\Gamma_N}},$$

or

$$\|[q]\|_{\operatorname{div}} := \|q\|_{1/\alpha} + C_\Omega\sqrt{\|\operatorname{div}q\|^2_{1/\alpha} + \|q \cdot \nu\|^2_{1/\alpha,\Gamma_N}}.$$

It is easy to see that

$$\|p - \widetilde{q}_h\|_{1/\alpha} \leq \|p - \alpha\nabla\widetilde{u}_h\|_{1/\alpha} + \|\widetilde{q}_h - \alpha\nabla\widetilde{u}_h\|_{1/\alpha}$$
$$\leq \|\nabla(u - \widetilde{u}_h)\|_\alpha + \|\widetilde{q}_h - \alpha\nabla\widetilde{u}_h\|_{1/\alpha}$$
$$\leq \mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h) + \|\alpha\nabla\widetilde{u}_h - \widetilde{q}_h\|_{1/\alpha} := \mathfrak{M}_p(\widetilde{u}_h, \widetilde{q}_h), \quad (4.8)$$

where $\mathfrak{M}_p(\widetilde{u}_h, \widetilde{q}_h)$ is the computable error majorant for the dual variable. We note that

$$\mathfrak{M}_p(\widetilde{u}_h, \widetilde{q}_h) \leq 2\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h).$$

Since

$$\|\mathrm{div}(p - \widetilde{q}_h)\|_{1/\alpha} = \|f + \mathrm{div}\widetilde{q}_h\|_{1/\alpha}$$

we find that

$$\|p - \widetilde{q}_h\|_{\mathrm{div}} \leq \mathfrak{M}_p(\widetilde{u}_h, \widetilde{q}_h) + \|f + \mathrm{div}\widetilde{q}_h\|_{1/\alpha} := \mathfrak{M}_{p,\mathrm{div}}(\widetilde{u}_h, \widetilde{q}_h) \quad (4.9)$$

Note that

$$\mathfrak{M}_{p,\mathrm{div}}(\widetilde{u}_h, \widetilde{q}_h) \leq 2\|\alpha\nabla\widetilde{u}_h - \widetilde{q}_h\|_{1/\alpha} +$$
$$+ (C_\Omega + 1)\sqrt{\|f + \mathrm{div}\widetilde{q}_h\|_{1/\alpha}^2 + \|g_N - \widetilde{q}_h \cdot \nu\|_{1/\alpha,\Gamma_N}^2} \leq C_{1\oplus}\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h) \quad (4.10)$$

where $C_{1\oplus} = \max\left\{2, \frac{C_\Omega + 1}{C_\Omega}\right\}$. For another norm we have

$$\|[p - \widetilde{q}_h]\|_{\mathrm{div}} \leq 2\|\alpha\nabla\widetilde{u}_h - \widetilde{q}_h\|_{1/\alpha} +$$
$$+ 2C_\Omega\sqrt{\|f + \mathrm{div}\widetilde{q}_h\|_{1/\alpha}^2 + \|g_N - \widetilde{q}_h \cdot \nu\|_{1/\alpha,\Gamma_N}^2} = 2\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h). \quad (4.11)$$

In the numerical tests it was observed that $\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h)$ quite adequately represents the error regardless of whether or not the term $\mathrm{div}\widetilde{q}_h + f$ is big.

4.3.   A posteriori estimates in combined primal–dual norms

From (4.4) and (4.9) we obtain the estimate in a combined primal–dual norm

$$\| (u - \widetilde{u}_h, p - \widetilde{q}_h) \| := \|\nabla(u - \widetilde{u}_h)\|_\alpha + \|p - \widetilde{q}_h\|_{\mathrm{div}} \leq \mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h) + \mathfrak{M}_{p,\mathrm{div}}(\widetilde{u}_h, \widetilde{q}_h). \quad (4.12)$$

Denote the right-hand side of (4.12) by $\mathfrak{M}_{u,p}^{(1)}(\widetilde{u}_h, \widetilde{q}_h)$. This functional is also subject to the majorant $\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h)$. Indeed,

$$\mathfrak{M}_{u,p}^{(1)}(\widetilde{u}_h, \widetilde{q}_h) \leq 3\|\alpha\nabla\widetilde{u}_h - \widetilde{q}_h\|_{1/\alpha} +$$
$$+ (2C_\Omega + 1)\sqrt{\|f + \mathrm{div}\widetilde{q}_h\|_{1/\alpha}^2 + \|g_N - \widetilde{q}_h \cdot \nu\|_{1/\alpha,\Gamma_N}^2} \leq C_{2\oplus}\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h), \quad (4.13)$$

where $C_{2\oplus} = \max\left\{3, \frac{2C_\Omega + 1}{C_\Omega}\right\}$.

Another combined norm (generated by the norm $\|[p - \widetilde{q}_h]\|_{\mathrm{div}}$) is estimated quite similarly. We have

$$\|[(u - \widetilde{u}_h, p - \widetilde{q}_h)]\| := \|\nabla(u - \widetilde{u}_h)\|_\alpha + \|[p - \widetilde{q}_h]\|_{\mathrm{div}}$$

$$\leq 3(\|\alpha\nabla\widetilde{u}_h - \widetilde{q}_h\|_{1/\alpha} + C_\Omega\sqrt{\|f + \mathrm{div}\widetilde{q}_h\|_{1/\alpha}^2 + \|g_N - \widetilde{q}_h \cdot \nu\|_{1/\alpha,\Gamma_N}^2}) = 3\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h). \quad (4.14)$$

It is worth emphasizing that the majorant $\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h)$ also provides a lower bound for the error in the combined primal–dual norm. Indeed,

$$\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h) \leq \|\alpha\nabla(\widetilde{u}_h - u)\|_{1/\alpha} + \|p - \widetilde{q}_h\|_{1/\alpha}$$

$$+ C_\Omega\sqrt{\|f + \mathrm{div}\widetilde{q}_h\|_{1/\alpha}^2 + \|g_N - \widetilde{q}_h \cdot \nu\|_{1/\alpha,\Gamma_N}^2}$$

$$= \|\nabla(\widetilde{u}_h - u)\|_\alpha + \|p - \widetilde{q}_h\|_{1/\alpha} + C_\Omega\sqrt{\|f + \mathrm{div}\widetilde{q}_h\|_{1/\alpha}^2 + \|g_N - \widetilde{q}_h \cdot \nu\|_{1/\alpha,\Gamma_N}^2}.$$

Therefore,

$$\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h) \leq \|[(u - \widetilde{u}_h, p - \widetilde{q}_h)]\|. \quad (4.15)$$

Thus, we observe that the majorant $\mathfrak{M}_u$ is *equivalent to the error evaluated in the combined primal–dual norm*. Similar result for mixed approximations of elliptic type boundary–value problems was established in [28]. Moreover, the relations (4.14) and (4.15) show that if the error is controlled in the norm $\|[(u - \widetilde{u}_h, p - \widetilde{q}_h)]\|$ then the efficiency index is always between 1 and 3. These bounds are observed in the experiments, see below.

The majorant $\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h)$ is also equivalent to the norm

$$\|[(u - \widetilde{u}_h, p - \widetilde{q}_h)]\|.$$

Indeed,

$$C_\ominus\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h) \leq \|(u - \widetilde{u}_h, p - \widetilde{q}_h)\|, \quad (4.16)$$

where $C_\ominus = 1/\max\{1, C_\Omega\}$. Recalling (4.13), we conclude that for the norm $\|(u - \widetilde{u}_h, p - \widetilde{q}_h)\|$ the upper bound of the efficiency index is given by the quantity $C_{2\oplus}/C_\ominus$.

5. Numerical results

In this section, we present numerical tests showing the efficiency and the reliability of our majorants $\mathfrak{M}_u$, $\mathfrak{M}_p$ and combinations of them. In all the following examples, the different errors and the error indicator are drawn with respect to the degrees of freedom using log-log scales. We plot here the convergence of the approximation errors in the above-introduced norms, namely, the error in the primal variable $\|\nabla(u - \widetilde{u}_h)\|_\alpha$, in the dual variable $\|p - \widetilde{q}_h\|_{1/\alpha}$ and in the primal-dual variables

$\| (u - \widetilde{u}_h, p - \widetilde{q}_h) \|$, that we compare with the error indicator $\| \alpha \nabla \widetilde{u}_h - \tilde{q}_h \|_{1/\alpha}$, which is in fact the first term of $\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h)$.

For all our tests, we further present the following efficiency indices, namely the ratios numbered as follows and corresponding to the respective estimates (4.6), (4.8), (4.9), (4.11), (4.12) and (4.14):

$$ratio1 = \frac{\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h)}{\| \nabla(u - \widetilde{u}_h) \|_\alpha} \qquad\qquad ratio2 = \frac{\mathfrak{M}_p(\widetilde{u}_h, \widetilde{q}_h)}{\| p - \widetilde{q}_h \|_{1/\alpha}}$$

$$ratio3 = \frac{\mathfrak{M}_{p,\mathrm{div}}(\widetilde{u}_h, \widetilde{q}_h)}{\| p - \widetilde{q}_h \|_{\mathrm{div}}} \qquad\qquad ratio4 = \frac{2\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h)}{\| [p - \widetilde{q}_h] \|_{\mathrm{div}}}$$

$$ratio5 = \frac{\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h) + \mathfrak{M}_{p,\mathrm{div}}(\widetilde{u}_h, \widetilde{q}_h)}{\| (u - \widetilde{u}_h, p - \widetilde{q}_h) \|} \qquad ratio6 = \frac{3\mathfrak{M}_u(\widetilde{u}_h, \widetilde{q}_h)}{\| [(u - \widetilde{u}_h, p - \widetilde{q}_h)] \|}$$

According to the above mentioned estimates, these ratios are always larger than 1, moreover according to the estimate (4.15), $ratio6$ is also less than 3, i.e., $1 \le ratio6 \le 3$. Note further that the ratios are not related to optimal computable majorants, but to upper bounds of these majorants; see comments above.

5.1. Dirichlet problem, square domain

First, we consider the simplest case of the Laplace equation in the unit square $\Omega = (0, 1)^2$ with homogeneous Dirichlet boundary condition. This means that we consider problem (2.1) to (2.3) in the unit square with $\alpha = 1$ and $\Gamma_N = \emptyset$. As exact solution we take $u(x, y) = \sin(\pi x)\sin(\pi y)$, the right-hand side being fixed accordingly. This solution is the eigenvector of the Laplace operator associated with its smallest eigenvalue $2\pi^2$. Accordingly the constant C_Ω appearing in (4.3) is equal to $\frac{1}{\sqrt{2}\pi}$. In Fig. 3 left, we have presented the different errors and the error indicator in a log-log scale. Their convergence rates are around 1 for the dual variable error and of order $\frac{1}{2}$ for the other ones. On Fig. 3 right, we have plotted the effectivity indices, there we can check that they are always greater than 1 and the property $1 \le ratio6 \le 3$.

5.2. Dirichlet problem, L–shaped domain

For the second example, we consider the Laplace equation in the L–shaped domain $\Omega =]-1, 1[^2 \backslash]0, 1[\times]-1, 0[$ with non homogeneous Dirichlet boundary condition. In other words, we consider problem (2.1) to (2.3) in the L–shaped domain with $\alpha = 1$ and $\Gamma_N = \emptyset$. Here as solution, we take the so-called singular function of this problem given in polar coordinates by $u = r^{2/3} \sin(\frac{2\theta}{3})$, which is solution of the problem (2.1) to (2.3) (with $\alpha = 1$), $f = 0$ and non homogeneous Dirichlet boundary conditions. In Fig. 4, we plot at left the errors and the error indicator with respect to the number of degrees of freedom in a log-log scale and at right the corresponding ratios. In the left we see that the error indicator and the primal and dual errors converge with a convergence rate of $\frac{1}{2}$. On the contrary, we may remark that the primal-dual error does not converge (convergence rate of $-\frac{1}{2}$), this is due to the term $\| \mathrm{div}(p - \tilde{q}_h) \| = \| f + \mathrm{div}\tilde{q}_h \|$. In theory, this term should be
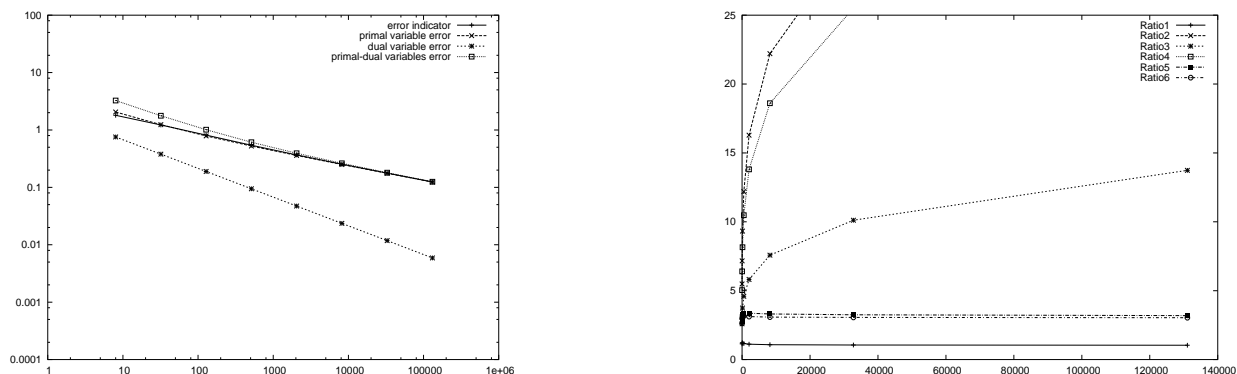
Figure 3: Example 1: on the left, the errors and the error indicator wrt the dofs; on the right, the ratios.

zero (for this example), but it turns out that the bad convergence of u_h into u leads to the fact that this term is not zero (and even blows up). This phenomenon is related to the numerical method rather than to error control. It is not surprising that such a phenomenon appears for this example, because the exact solution $u = r^{2/3}\sin\left(\frac{2\theta}{3}\right)$ is very irregular near the corner (and the divergence of $\tilde{q}_h$ is related to its second derivative). The numerical result indicates that the simplest (i.e., P^0 on cells) FV scheme is unable to provide convergence in H_{div} norm. This fact is quite predictable and is not in contradiction with the theoretical results related to error control. On the other hand, at Fig. 4 right, we may notice that all ratios are larger than 1, and again that $1 \leq ratio6 \leq 3$, as theoretically expected.
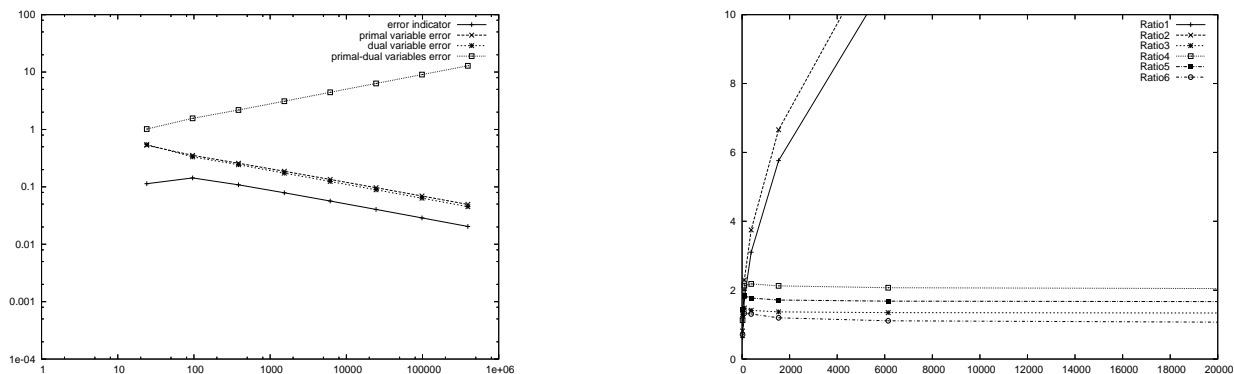

Figure 4: Example 2: on the left, the errors and the error indicator wrt the dofs; on the right, the ratios.

5.3. Mixed Dirichlet-Neumann problem, square domain

Thirdly, we demonstrate the performance of our error estimation method on the Laplace equation in the unit square $\Omega = (0,1)^2$ with mixed boundary condition. Hence we consider problem (2.1)

to (2.3) in the unit square with $\alpha = 1$, where homogeneous Dirichlet conditions are taken along the axes $x = 0$ and $x = 1$, and homogeneous Neumann boundary conditions are imposed along the axes $y = 0$ and $y = 1$. As exact solution we here take $u(x, y) = \sin(\pi x)\cos(\pi y)$. Since the smallest eigenvalue of problem (4.5) is equal to π^2, the constant C_Ω is equal to $\frac{1}{\pi}$. On Fig. 5, we have plotted all the errors and the error indicator as well as the effectivity indices. From it, we can make the same conclusions than for example 1.
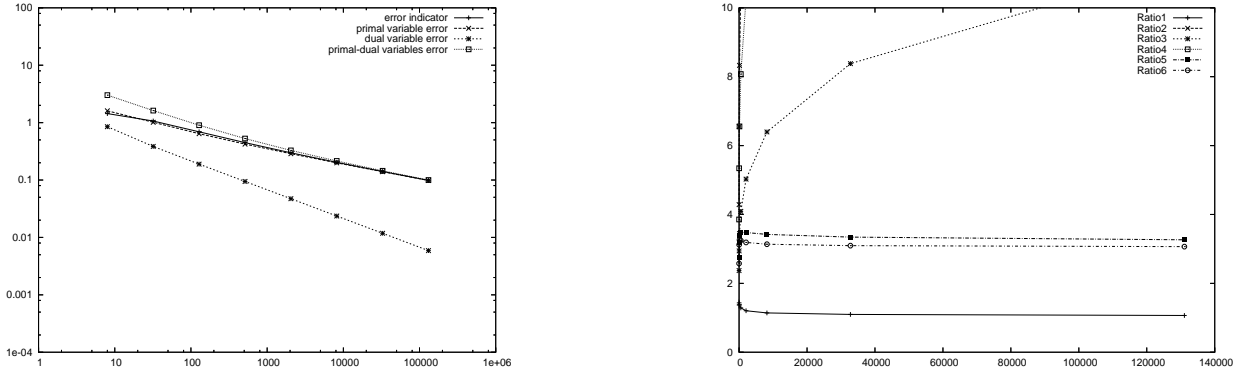

Figure 5: Example 3: on the left, the errors and the error indicator wrt the dofs; on the right, the ratios.

5.4.   The checkerboard

Here we want to consider the so-called checkerboard, namely we consider problem (2.1) to (2.3) in the unit square partitioned in four Lipschitz subdomains Ω_i, $i = 1, \ldots, 4$ as presented in Fig. 6 but with piecewise coefficient α, namely $\alpha = \alpha_i$ on Ω_i. In our test, we take $\alpha_1 = \alpha_3 = 10^{-4}$ and $\alpha_2 = \alpha_4 = 1$. The considered solution is the polynomial function $u(x, y) = x^2(1 - x)^2(1 + x)^2 y^2(1 - y)^2(1 + y)^2$ so that it is a smooth solution of (2.1) to (2.3) with homogeneous Dirichlet boundary conditions on Γ. For such a problem the constant in (4.2) (smallest eigenvalue of the operator $-\text{div}(\alpha\nabla)$ with Dirichlet boundary conditions) is not known explicitly but according to Proposition 4.3 of [6], this constant is close to the eigenvalue of the smallest eigenvalue of the Laplace equation in the subdomains $\Omega_1 \cup \Omega_3$ with Dirichlet boundary conditions on part of the boundary in common with the boundary of Ω (i.e., the part of the boundary included into $x = 0$, or $x = 1$, or $y = 0$, or $y = 1$) and Neumann boundary conditions on the remaining part of its boundary. Since this eigenvalue is equal to $\frac{\pi^2}{2}$, the constant C_Ω is then chosen equal to $\frac{\sqrt{2}}{\pi}$. All the errors and the error indicator as well as the effectivity indices are presented in Fig 7. From it, we can make the same conclusions than for example 1 with even convergence order of around 1 for the four quantities.
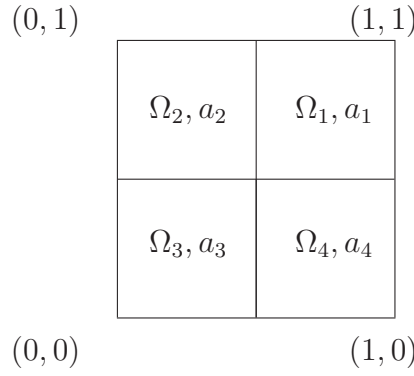
$(0, 1)$ $(1, 1)$

Ω_2, a_2	Ω_1, a_1
Ω_3, a_3	Ω_4, a_4

$(0, 0)$ $(1, 0)$

Figure 6: The checkerboard.

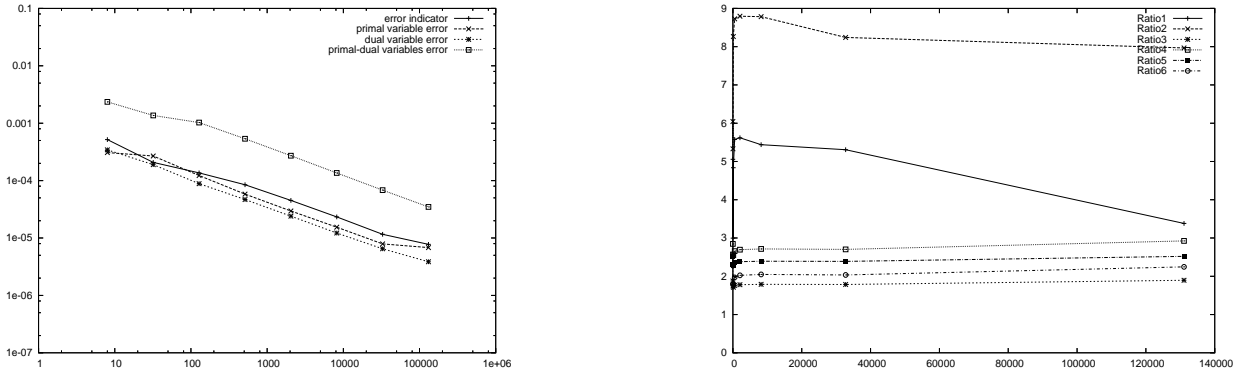
Figure 7: Example 4: on the top, the errors and the error indicator wrt the dofs; on the bottom, the ratios.

5.5. A domain with a crack

To end this series of tests, we consider a boundary value problem on the crack domain $\Omega = (-1, 1)^2 \setminus \sigma$, where the crack $\sigma = [0, 1)$. On this domain we consider problem (2.1) to (2.3) with $\alpha = 1$, non-homogeneous Dirichlet condition on the exterior boundary and homogeneous Neumann boundary conditions on σ (note that σ is unfolded in the sense that boundary conditions are imposed from above and from below of σ). As exact solution we take the so-called singular function $u(x, y) = r^{\frac{1}{2}} \cos(\frac{\theta}{2})$ so that it is a non smooth solution of (2.1) to (2.3) with $f = 0$. For such a problem the constant in (4.2) is equal to $\frac{\sqrt{2}}{\pi}$. Indeed if u is an eigenvector of eigenvalue λ^2 then $v(x, y) = u(x, y) + u(x, -y)$ is an eigenvector of the Laplace operator with the same eigenvalue in the domain $(-1, 1) \times (0, 1)$ with Neumann boundary conditions on $(-1, 1) \times \{0\}$ and Dirichlet boundary conditions on the remainder of the boundary, hence $\lambda^2 \geq \frac{\pi^2}{2}$. On the other hand we easily see that the function $u(x, y) = \sin(\frac{\pi}{2}(x + 1)) \sin(\frac{\pi}{2}(y + 1))$ is an eigenvector of the original problem of eigenvalue $\frac{\pi^2}{2}$, which implies that the smallest eigenvalue is exactly $\frac{\pi^2}{2}$.

All the errors and the error indicator as well as the effectivity indices are presented in Fig 8. From it, we can make the same conclusions than for example 2, due to the singular behavior of the solution.
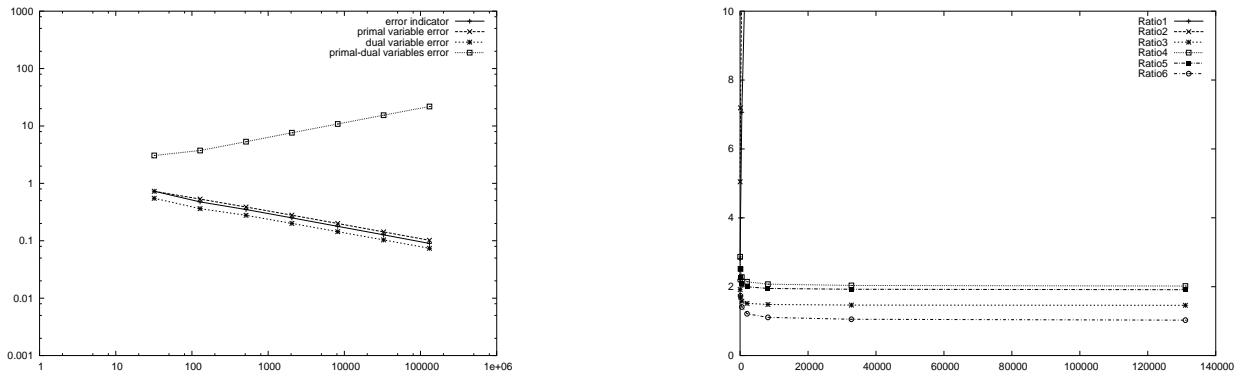
Figure 8: Example 5: on the top, the errors and the error indicator wrt the dofs; on the bottom, the ratios.

References

[1] A. Agouzal, F. Oudin. *A posteriori error estimator for finite volume methods.* Appl. Math. Comp., 110 (2000), 239–250.

[2] M. Ainsworth, J. T. Oden. A posteriori error estimation in finite element analysis, Wiley, New York, 2000.

[3] I. Babuška, W. C. Rheinboldt. *Error estimates for adaptive finite element computations.* SIAM J. Numer. Anal., 15 (1978), 736–754.

[4] I. Babuška, T. Strouboulis. The finite element method and its reliability. Clarendon Press, New York, 2001.

[5] A. Bergam, Z. Mghazli, and R. Verfürth. *A posteriori estimators for the finite volume discretization of an elliptic problem.* Numer. Math., 95 (2003), No. 4, 599–624.

[6] M. Costabel, M. Dauge, S. Nicaise. *Singularities of eddy current problems.* RAIRO Modél. Math. Anal. Numér., 37 (2003), 807–831.

[7] E. Dari, R. Duran, C. Padra. *Error estimators for nonconforming finite element approximations of the Stokes problem.* Math. Comp., 64 (1995), No. 211, 1017–1033.

[8] I. Ekeland, R. Temam. Convex analysis and variational problems, Stud. Math. Appl., Vol. 1, North-Holland, Amsterdam–Oxford, 1976.

[9] R. Eymard, T. Gallouët, R. Herbin. The finite volume method. Handbook for Numerical Analysis, Ph. Ciarlet J.L. Lions editors, North Holland, 2000, 715–1022.

[10] R. Herbin and M. Ohlberger. *A posteriori error estimate for finite volume approximation of convection-diffusion problems.* In R. Herbin and D. Kröner, editors, Finite Volume for Complex Applications, 753–760. Hermès, 2002.

[11] N. Jullian. *An error indicator for cell-centered finite volumes for linear convection-diffusion problems.* In R. Herbin and D. Kröner, editors, Finite Volume for Complex Applications, 753–760. Hermès, 2002.

[12] D. Kröner and M. Ohlberger. *A posteriori error estimates for upwind finite volume schemes for nonlinear conservation laws in multi dimensions.* Math. Comp., 69 (1999), 25–39.

[13] R. Lazarov and S. Tomov. *Adaptive finite volume element method for convection-diffusion-reaction problems in 3-d.* In Y. W. P. Minev and Y. Lin, editors, Scientific Computing and Application, 91–106. Nova Science Publishing House, 2001.

[14] R. Lazarov and S. Tomov. *A posteriori error estimates for finite volume approximations of convection-diffusion-reaction equations.* Comput. Geosci., 6 (2002), 483–503.

[15] J. Mackenzie, T. Sonar, and G. Warnecke. *A posteriori error estimates for the cell-vertex finite volume method.* In W. Hackbusch and G. Wittum, editors, Adaptive Methods: Algorithms, Theory and Applications, 221–235, Vieweg, 1994.

[16] K. W. Morton and E. Süli. *A posteriori and a priori error analysis of finite volume methods.* In J. R. Whiteman, editor, The Mathematics of Finite Elements and Applications, 267–288, J. Wiley and Sons, 1994.

[17] P. Neittaanmäki, S. Repin. Reliable methods for computer simulation, Error control and a posteriori estimates. Elsevier, New York, 2004.

[18] S. Nicaise. *A posteriori error estimations of some cell centered finite volume methods.* Siam J. Numer. Anal., 43 (2005), No. 4, 1481–1503.

[19] S. Nicaise. *A posteriori error estimations of some cell centered finite volume methods for diffusion-convection-reaction problems.* Siam J. Numer. Anal., 44 (2006), No. 3, 949–978.

[20] S. Nicaise, S. Repin. *Functional a posteriori error estimates for the reaction–convection–diffusion problem.* J. Math. Sciences, 152 (2008), 690–701.

[21] M. Ohlberger. *A posteriori error estimate for finite volume approximations to singularly perturbed nonlinear convection-diffusion equations.* Numer. Math., 87 (2001), 737–761.

[22] M. Ohlberger. *A posteriori error estimates for vertex centered finite volume approximations of convection-diffusion-reaction equations.* M2AN, 35 (2001), 355–387.

[23] S. V. Patankar. Numerical Heat transfer and fluid flow. Series in Comp. methods Mechanics and Thermal Sciences. Mc Graw Hill, 1980.

[24] S. Repin. *Two-sided estimates of deviation from exact solutions of uniformly elliptic equations.* Proceedings of the St. Petersburg Mathematical Society, Vol. IX, 143–171, Amer. Math. Soc. Transl. Ser. 2, 209, Amer. Math. Soc., Providence, RI, 2003.

[25] S. Repin. *A posteriori error estimation for nonlinear variational problems by duality theory.* Zapiski Nauchn. Semin. Pomi., 243 (1997), 201–214.

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

[27] S. Repin, S. Sauter. *Functional A posteriori estimates for the reaction-diffusion Problem.* C. R. Acad. Sci. Paris, Sér. 1, 343 (2006), 349–354.

[28] S. Repin, S. Sauter, A. Smolianski. *Two-Sided a posteriori error estimates for mixed formulations of elliptic problems.* Preprint 21-2005, Institute of Mathematics, University of Zurich (to appear in SIAM J. Numer. Anal.).

[29] R. Verfürth. A review of a posteriori error estimation and adaptive mesh–refinement techniques. Wiley, Teubner, New York, 1996.

[30] M. Vohralík. *A posteriori error estimates for finite volume and mixed finite element discretizations of convection-diffusion-reaction equations.* ESAIM: Proc., 18 (2007), 57–69.