# Functional A Posteriori Error Estimate for a Nonsymmetric Stationary Diffusion Problem

**Olli Mali**

**Abstract** In this paper, a posteriori error estimates of functional type for a stationary diffusion problem with nonsymmetric coefficients are derived. The estimate is guaranteed and does not depend on any particular numerical method. An algorithm for the global minimization of the error estimate with respect to an auxiliary function over some finite dimensional subspace is presented. In numerical tests, global minimization is done over the subspace generated by Raviart-Thomas elements. The improvement of the error bound due to the $p$-refinement of these spaces is investigated.

## 1 Introduction

In this paper, we derive a posteriori error estimates of the functional type for a class of elliptic problems with nonsymmetric coefficients. Since mid 90's (see [8]), estimates of this type has been derived for a wide range of problems (see, e.g., monographs [5, 6, 9] and references therein). However, the case of a stationary diffusion problem, where coefficients are not symmetric has not been studied before. Problems of this type are not very typical among other elliptic equations but they arise in certain models (see, e.g., [1, 2]). It is shown that the derived estimate has the standard properties of a deviation estimate for a linear problem, i.e., it is guaranteed and computable. The derivation of the estimate is based on the method of integral identities and a special case of Cauchy-Schwartz-Bunyakovsky inequality.

O. Mali (✉)
Department of Mathematical Information Technology, University of Jyvaskyla,
P.O. Box 35, FI-40014 Jyvaskyla, Finland
e-mail: olli.mali@jyu.fi



P. Neittaanmäki et al. (eds.), *Mathematical Modeling and Optimization of Complex Structures*, Computational Methods in Applied Sciences 40,
DOI 10.1007/978-3-319-23564-6_2

Consider the Poisson problem

$$-\operatorname{div} \mathbf{A}\nabla u = f \quad \text{in } \Omega \subset \mathbb{R}^d \tag{1}$$

$$u = 0 \quad \text{on } \Gamma := \partial\Omega, \tag{2}$$

where $\Omega$ a is simply connected domain with a Lipschitz-continuous boundary, $f \in L^2(\Omega)$, and $\mathbf{A} \in L_\infty(\Omega, \mathbb{R}^{d\times d})$ is strictly positive definite, bounded, and has a bounded inverse $\mathbf{A}^{-1} \in \mathbb{R}^{d\times d}$ in $\Omega$. Moreover, $\mathbf{A}$ is positive definite, i.e., there exists constant $\underline{c} > 0$ such that

$$(\mathbf{A}\boldsymbol{\xi}, \boldsymbol{\xi})_{\mathbb{R}^d} \geq \underline{c}\|\boldsymbol{\xi}\|^2_{\mathbb{R}^d}, \quad \forall \boldsymbol{\xi} \in \mathbb{R}^d, \text{ a.e. in } \Omega. \tag{3}$$

The generalized solution $u \in H^1_0(\Omega)$ satisfies the integral identity,

$$(\mathbf{A}\nabla u, \nabla w)_{L^2(\Omega,\mathbb{R}^d)} = (f, w)_{L^2(\Omega)}, \quad \forall w \in H^1_0(\Omega). \tag{4}$$

## 2 Error Majorant

For symmetric problems with $\mathbf{A} \in L_\infty(\Omega, \mathbb{R}^{d\times d}_{\text{sym}})$ the respective guaranteed upper bounds (error majorants) have been presented in [5, 6, 9] and other publications cited therein. It has the form

$$\overline{\mathfrak{M}}(v, \mathbf{y}) := (\mathbf{A}\nabla v - \mathbf{y}, \nabla v - \mathbf{A}^{-1}\mathbf{y})^{1/2}_{L^2(\Omega,\mathbb{R}^d)} + \frac{C_F}{\sqrt{\underline{c}}}\|\operatorname{div}\mathbf{y} + f\|_{L^2(\Omega)},$$

where $v \in H^1_0(\Omega)$, $\mathbf{y} \in H(\operatorname{div}, \Omega)$, and $C_F$ is the constant in the Friedrichs inequality

$$\|w\|_{L^2(\Omega)} \leq C_F\|\nabla w\|_{L^2(\Omega,\mathbb{R}^d)}, \quad \forall w \in H^1_0(\Omega). \tag{5}$$

A special case of the Cauchy-Schwartz-Bunyakovsky inequality presented below is required to obtain an analogous error estimate in the nonsymmetric case.

**Lemma 1** *Let $\mathcal{U}$ be a Hilbert space whose field is real numbers, $A : \mathcal{U} \to \mathcal{U}$ is continuous, bounded, strictly positive definite, and has a continuous inverse $A^{-1}$. Moreover,*

$$B := (\mathrm{Id} + A^T A^{-1})^{-1}$$

*is continuous and bounded. Then,*

$$(y, q)_{\mathcal{U}} \leq 2(Ay, y)^{1/2}_{\mathcal{U}}(A^{-1}Bq, Bq)^{1/2}_{\mathcal{U}}, \quad \forall y, q \in \mathcal{U}. \tag{6}$$

*Proof* Since $A$ is strictly positive definite,

$$\begin{aligned}0 &\le (A(y-\gamma A^{-1}q), y-\gamma A^{-1}q)_{\mathcal{U}} \\ &= (Ay, y)_{\mathcal{U}} - \gamma(y, (\mathrm{Id} + A^T A^{-1})q)_{\mathcal{U}} + \gamma^2 (A^{-1}q, q)_{\mathcal{U}}.\end{aligned}$$

Selecting (assume $y \neq 0$ and $q \neq 0$, otherwise (6) holds trivially)

$$\gamma = \frac{2(Ay, y)_{\mathcal{U}}}{(y, (\mathrm{Id} + A^T A^{-1})q)_{\mathcal{U}}}$$

yields

$$(y, (\mathrm{Id} + A^T A^{-1})q)^2_{\mathcal{U}} \le 4(Ay, y)_{\mathcal{U}} (A^{-1}q, q)_{\mathcal{U}},$$

where setting $q = Bq = (\mathrm{Id} + A^T A^{-1})^{-1} q$ leads to (6).

**Theorem 1** *Let $v \in H^1_0(\Omega)$ and $u$ be the solution of (4). Then*

$$(\mathbf{A}\nabla(u-v), \nabla(u-v))^{1/2}_{L^2(\Omega,\mathbb{R}^d)} \le \overline{\mathfrak{M}}(v, \mathbf{y}), \quad \forall \mathbf{y} \in H(\mathrm{div}, \Omega),$$

*where*

$$\overline{\mathfrak{M}}(v, \mathbf{y}) := 2(\mathbf{A}^{-1}\mathbf{B}(\mathbf{y} - \mathbf{A}\nabla v), \mathbf{B}(\mathbf{y} - \mathbf{A}\nabla v))^{1/2}_{L^2(\Omega,\mathbb{R}^d)} + \frac{C_F}{\sqrt{\underline{c}}} \| \mathrm{div}\, \mathbf{y} + f \|_{L^2(\Omega)} \tag{7}$$

*and*

$$\mathbf{B} := (\mathbf{I} + \mathbf{A}^T \mathbf{A}^{-1})^{-1}.$$

*The constants $C_F$ and $\underline{c}$ are defined in (5) and (3), respectively.*

*Proof* Subtracting $(\mathbf{A}\nabla v, \nabla w)_{L^2(\Omega,\mathbb{R}^d)}$ from both sides of (4) and applying the integration by parts formula

$$(\mathbf{y}, \nabla w)_{L^2(\Omega,\mathbb{R}^d)} = (-\mathrm{div}\, \mathbf{y}, w)_{L^2(\Omega)}, \quad \forall \mathbf{y} \in H(\mathrm{div}, \Omega),\ w \in H^1_0(\Omega)$$

yields

$$(\mathbf{A}\nabla(u-v), \nabla w)_{L^2(\Omega,\mathbb{R}^d)} = (\mathbf{y} - \mathbf{A}\nabla v, \nabla w)_{L^2(\Omega,\mathbb{R}^d)} + (\mathrm{div}\, \mathbf{y} + f, w)_{L^2(\Omega)}.$$

The first term can be estimated from above by (6), where $\mathcal{U} := L^2(\Omega, \mathbb{R}^d)$ and $A := \mathbf{A}$. The second term is estimated from above by the Hölder inequality, (5), and (3), which leads to

$$
\begin{aligned}
&(\mathbf{A}\nabla(u-v), \nabla w)_{L^2(\Omega,\mathbb{R}^d)} \\
&\qquad \leq 2(\mathbf{A}^{-1}\mathbf{B}(\mathbf{y}-\mathbf{A}\nabla v), \mathbf{B}(\mathbf{y}-\mathbf{A}\nabla v))^{1/2}_{L^2(\Omega,\mathbb{R}^d)} (\mathbf{A}\nabla w, \nabla w)^{1/2}_{L^2(\Omega,\mathbb{R}^d)} \\
&\qquad + \frac{C_F}{\sqrt{\underline{c}}} \| \operatorname{div} \mathbf{y} + f \|_{L^2(\Omega)} (\mathbf{A}\nabla w, \nabla w)^{1/2}_{L^2(\Omega,\mathbb{R}^d)}.
\end{aligned}
$$

Setting $w = u - v$ leads to (7).

*Remark 1* Two parts of the majorant are related to the violations of the duality relation and the equilibrium condition, respectively. They are denoted by

$$
\begin{aligned}
\overline{\mathfrak{M}}_{\text{Dual}} &:= (\mathbf{A}^{-1}\mathbf{B}(\mathbf{y}-\mathbf{A}\nabla v), \mathbf{B}(\mathbf{y}-\mathbf{A}\nabla v))^{1/2}_{L^2(\Omega,\mathbb{R}^d)}, \\
\overline{\mathfrak{M}}_{\text{Equi}} &:= \| \operatorname{div} \mathbf{y} + f \|_{L^2(\Omega)}.
\end{aligned}
$$

## 3 Global Minimization of the Error Majorant

Squaring and applying the Young's inequality yields a quadratic form of the majorant, which is more suitable for the minimization over $\mathbf{y}$.

**Corollary 1** *Let $v \in H^1_0(\Omega)$ and $u$ be the solution of (4), then,*

$$
(\mathbf{A}\nabla(u-v), \nabla(u-v))_{L^2(\Omega,\mathbb{R}^d)} \leq \overline{\mathfrak{M}}^2(v, \mathbf{y}, \beta), \quad \forall \mathbf{y} \in H(\operatorname{div}, \Omega),\ \beta > 0,
$$

*where*

$$
\begin{aligned}
\overline{\mathfrak{M}}^2(v, \mathbf{y}, \beta) := 4(1+\beta)(\mathbf{A}^{-1}\mathbf{B}(\mathbf{y}-\mathbf{A}\nabla v), \mathbf{B}(\mathbf{y}-\mathbf{A}\nabla v))_{L^2(\Omega,\mathbb{R}^d)} \\
+ \frac{1+\beta}{\beta} \frac{C_F^2}{\underline{c}} \| \operatorname{div} \mathbf{y} + f \|^2_{L^2(\Omega)}. \qquad (8)
\end{aligned}
$$

**Corollary 2** *The minimizers*

$$
\begin{aligned}
\overline{\mathfrak{M}}^2(v, \hat{\mathbf{y}}, \beta) &= \min_{\mathbf{y} \in H(\operatorname{div},\Omega)} \overline{\mathfrak{M}}^2(v, \mathbf{y}, \beta), \\
\overline{\mathfrak{M}}^2(v, \mathbf{y}, \hat{\beta}) &= \min_{\beta > 0} \overline{\mathfrak{M}}^2(v, \mathbf{y}, \beta)
\end{aligned}
$$

*satisfy*

$$
\begin{aligned}
&\frac{C_F^2}{\underline{c}} (\operatorname{div} \hat{\mathbf{y}}, \operatorname{div} \mathbf{q})_{L^2(\Omega)} + 2\beta \left( (\mathbf{A}^{-1}\mathbf{B}\mathbf{q}, \mathbf{B}\hat{\mathbf{y}})_{L^2(\Omega,\mathbb{R}^d)} + (\mathbf{A}^{-1}\mathbf{B}\hat{\mathbf{y}}, \mathbf{B}\mathbf{q})_{L^2(\Omega,\mathbb{R}^d)} \right), \\
&= -\frac{C_F^2}{\underline{c}} (f, \operatorname{div} \mathbf{q})_{L^2(\Omega)} + 2\beta \left( (\mathbf{A}^{-1}\mathbf{B}\mathbf{q}, \mathbf{B}\mathbf{A}\nabla v)_{L^2(\Omega,\mathbb{R}^d)} + (\mathbf{A}^{-1}\mathbf{B}\mathbf{A}\nabla v, \mathbf{B}\mathbf{q})_{L^2(\Omega,\mathbb{R}^d)} \right), \\
&\qquad \forall \mathbf{q} \in H(\operatorname{div}, \Omega) \qquad (9)
\end{aligned}
$$

*and*

$$\hat{\beta} = \frac{\frac{C_F}{\sqrt{\underline{c}}} \| \operatorname{div} \mathbf{y} + f \|_{L^2(\Omega)}}{2(\mathbf{A}^{-1}\mathbf{B}(\mathbf{y} - \mathbf{A}\nabla v), \mathbf{B}(\mathbf{y} - \mathbf{A}\nabla v))^{1/2}_{L^2(\Omega, \mathbb{R}^d)}}, \tag{10}$$

*respectively.*

*Proof* The functional $\overline{\mathfrak{M}}^2(v, \mathbf{y}, \beta)$ is quadratic and convex w.r.t. $\mathbf{y}$. Thus the necessary and sufficient condition for the minimizer $\hat{\mathbf{y}}$ is

$$\frac{\mathrm{d}}{\mathrm{d}t} \overline{\mathfrak{M}}^2(v, \hat{\mathbf{y}} + t\mathbf{q}, \beta)\Big|_{t=0} = 0, \quad \forall \mathbf{q} \in H(\operatorname{div}, \Omega),$$

which leads to (9). Similarly,

$$\frac{\mathrm{d}}{\mathrm{d}\beta} \overline{\mathfrak{M}}^2(v, \mathbf{y}, \hat{\beta}) = 0$$

yields (10).

*Remark 2* If $\mathbf{A}$ is symmetric, then (9) reduces to

$$\frac{C_F^2}{\underline{c}} \int_\Omega \operatorname{div} \hat{\mathbf{y}} \operatorname{div} \mathbf{q} d\mathbf{x} + \beta \int_\Omega \mathbf{A}^{-1} \hat{\mathbf{y}} \cdot \mathbf{q} d\mathbf{x}$$
$$= -\frac{C_F^2}{\underline{c}} \int_\Omega f \operatorname{div} \mathbf{q} d\mathbf{x} + \beta \int_\Omega \nabla v \cdot \mathbf{q} d\mathbf{x} \quad \forall \mathbf{q} \in H(\operatorname{div}, \Omega).$$

There are many alternatives how to compute the value of the majorant (see, e.g., [5, Chap. 3]). Here, the global minimization of the majorant over finite dimensional subspace is presented. The minimization is done iteratively by solving (9) and (10) subsequently.

Let $\mathbf{y} = \sum_{j=1}^N c_j \boldsymbol{\phi}_j$ and $Q_h := \operatorname{span}(\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_N) \subset H(\operatorname{div}, \Omega)$, i.e., $\boldsymbol{\phi}_j$ $(j \in \{1, \ldots, N\})$ are the global basis functions. Then (9) leads to a system of linear equations

$$\left( \frac{C_F^2}{\underline{c}} \mathbf{S} + 2\beta \mathbf{M} \right) \mathbf{c} = -\frac{C_F^2}{\underline{c}} \mathbf{b} + 2\beta \mathbf{z}, \tag{11}$$

where

$$S_{ij} := (\operatorname{div} \boldsymbol{\phi}_j, \operatorname{div} \boldsymbol{\phi}_i)_{L^2(\Omega)}, \tag{12}$$

$$M_{ij} := (\mathbf{A}^{-1}\mathbf{B}\boldsymbol{\phi}_j, \mathbf{B}\boldsymbol{\phi}_i)_{L^2(\Omega, \mathbb{R}^d)} + (\mathbf{A}^{-1}\mathbf{B}\boldsymbol{\phi}_i, \mathbf{B}\boldsymbol{\phi}_j)_{L^2(\Omega, \mathbb{R}^d)}, \tag{13}$$

$$b_i := (f, \operatorname{div} \boldsymbol{\phi}_i)_{L^2(\Omega)}, \tag{14}$$

$$z_i := (\mathbf{A}^{-1}\mathbf{B}\boldsymbol{\phi}_i, \mathbf{B}\mathbf{A}\nabla v)_{L^2(\Omega,\mathbb{R}^d)} + (\mathbf{A}^{-1}\mathbf{B}\mathbf{A}\nabla v, \mathbf{B}\boldsymbol{\phi}_i)_{L^2(\Omega,\mathbb{R}^d)}, \tag{15}$$

and $\mathbf{c} \in \mathbb{R}^N$ is the (column) vector of unknown coefficients. The natural choice is to generate $Q_h$ using Raviart-Thomas elements (see [7]). The global minimization procedure for $\overline{\mathfrak{M}}^2$ is described in Algorithm 1.

**Algorithm 1**: Computation of the majorant for the problem (1)–(2)

---

**Input**: $v$ {approximate solution}, $\mathbf{A}$, {diffusion coefficient matrix} $f$, {RHS of the problem}, $C_F$, {Constant in (5)}, $\underline{c}$, {Constant in (3)}, $I_{\max}$ {maximum number of iterations}, $\epsilon$ {stopping criteria for $\overline{\mathfrak{M}}$}
Generate $\mathbf{S}$, $\mathbf{M}$, $\mathbf{b}$, and $\mathbf{z}$ in (12)–(15).
Compute norms $\|f\|$ and $\|\nabla v\|$.
Set $\beta_1 := 1$, $\overline{\mathfrak{M}}_k = \infty$ and $k = 0$. {initialize parameters}
**while** $k < I_{\max}$ **and** $\frac{\overline{\mathfrak{M}}_{k+1} - \overline{\mathfrak{M}}_k}{\overline{\mathfrak{M}}_k} > \epsilon$ **do**
  $k = k + 1$
  Solve $\mathbf{c}_{k+1}$ from $\left(\frac{C_F^2}{\underline{c}}\mathbf{S} + 2\beta_k\mathbf{M}\right)\mathbf{c}_{k+1} = -\frac{C_F^2}{\underline{c}}\mathbf{b} + 2\beta_k\mathbf{z}$.
  $\overline{\mathfrak{M}}_{k+1}^{\text{Equi}} = \sqrt{\mathbf{c}_{k+1}^T\mathbf{S}\mathbf{c}_{k+1} + 2\mathbf{c}_{k+1}^T\mathbf{b} + \|f\|^2}$
  $\overline{\mathfrak{M}}_{k+1}^{\text{Dual}} = \sqrt{\mathbf{c}_{k+1}^T\mathbf{M}\mathbf{c}_{k+1} - 2\mathbf{c}_{k+1}^T\mathbf{z} + \|\nabla v\|^2}$
  $\beta_{k+1} = \frac{C_F\overline{\mathfrak{M}}_{k+1}^{\text{Equi}}}{2\sqrt{\underline{c}}\,\overline{\mathfrak{M}}_{k+1}^{\text{Dual}}}$
  $\overline{\mathfrak{M}}_{k+1} = 2\overline{\mathfrak{M}}_{k+1}^{\text{Dual}} + \frac{C_F}{\sqrt{\underline{c}}}\overline{\mathfrak{M}}_{k+1}^{\text{Equi}}$
**end while**
$\mathbf{y} = \sum_{j=1}^{N} c_{kj}\boldsymbol{\phi}_j$
**Output**: $\overline{\mathfrak{M}}_{k+1}$ {Upper bound for the approximation error}, $\mathbf{y}$ {Approximation of the minimizer}

---

*Remark 3* Note that in Algorithm 1, the global matrices $\mathbf{S}$ and $\mathbf{M}$ have to be assembled only once. The coefficient matrix in (11) is symmetric regardless of the fact that $\mathbf{A}$ is not.

## 4 Numerical Tests

Algorithm 1 is very convenient to implement using any finite element software, e.g., FEniCS [4] and FREEFEM++ [3]), which allows user to define problems using weak forms. This is true for all estimates of the functional type presented in [5, 6, 9]. The following tests are computed using FEniCS finite element package. Here, we apply Algorithm 1 to estimate the error of a finite element approximation for a test example, where the exact solution is known.

*Example 1* Let $\Omega = (0,1) \times (0,1)$, $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $u(x_1,x_2) = \sin(k_1\pi x_1)$ $\sin(k_2\pi x_2)$, and

$$f(x_1,x_2) = \pi^2 \Big( (a+d)k_1^2 \sin(k_1\pi x_1)\sin(k_2\pi x_2) - (b+c)k_1k_2\cos(k_1\pi x_1)\cos(k_2\pi x_2) \Big).$$

Select $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$, then $\underline{c} = 2$, $\mathbf{A}^{-1} = \frac{1}{6}\begin{pmatrix} 3 & -1 \\ 0 & 2 \end{pmatrix}$ and $\mathbf{B} = \frac{1}{23}\begin{pmatrix} 11 & 2 \\ -3 & 12 \end{pmatrix}$.

The approximate solution $v \in V_h$ of Example 1 is computed on a mesh $\mathcal{T}_h$, using triangular Courant elements of the order $p_1$. The space $Q_h$ is generated using the Raviart-Thomas elements of order $p_2$ on the same mesh. The amount of global degrees of freedom are denoted by $N_1 = \dim(V_h)$ and $N_2 = \dim(Q_h)$. The efficiency index of the majorant is

$$I_{\text{eff}} := \sqrt{\frac{\overline{\mathfrak{M}}^2(v, \mathbf{y}, \beta)}{(\mathbf{A}\nabla(u-v), \nabla(u-v))_{L^2(\Omega,\mathbb{R}^d)}}}. \tag{16}$$

The majorant is computed for different meshes with $k_1 = 1$, $k_2 = 1$, and $p_1 = 1$ in Table 1.

The efficiency of the majorant and the number of iterations (in Algorithm 1, $\varepsilon = 10^{-6}$) do not depend on the mesh size. For $p_2 = 2$ and $p_2 = 3$, $Q_h$ can

**Table 1** Example 1: $k_1 = 1$, $k_2 = 1$, and $p_1 = 1$

| $N_1$ | $p_2$ | $N_2$ | $k$ | $\overline{\mathfrak{M}}^2(v, \mathbf{y}_k, \beta_k)$ | $\overline{\mathfrak{M}}_k^{\text{Dual}}$ | $\overline{\mathfrak{M}}_k^{\text{Equi}}$ | $I_{\text{eff}}$ |
|---|---|---|---|---|---|---|---|
| 441 | 1 | 1240 | 3 | 1.76E+00 | 2.46E−02 | 2.06E+00 | 6.6480 |
| 441 | 2 | 4080 | 3 | 3.15E−01 | 1.78E−02 | 2.09E−03 | 1.1858 |
| 441 | 3 | 8520 | 4 | 2.68E−01 | 1.78E−02 | 1.07E−06 | 1.0090 |
| 1681 | 1 | 4880 | 3 | 8.85E−01 | 6.23E−03 | 5.17E−01 | 6.6452 |
| 1681 | 2 | 16160 | 3 | 1.45E−01 | 4.44E−03 | 1.31E−04 | 1.0920 |
| 1681 | 3 | 33840 | 4 | 1.33E−01 | 4.44E−03 | 1.68E−08 | 1.0023 |
| 6561 | 1 | 19360 | 2 | 4.43E−01 | 1.56E−03 | 1.29E−01 | 6.6445 |
| 6561 | 2 | 64320 | 3 | 6.97E−02 | 1.11E−03 | 8.20E−06 | 1.0458 |
| 6561 | 3 | 134880 | 3 | 6.66E−02 | 1.11E−03 | 2.62E−10 | 1.0006 |
| 14641 | 1 | 43440 | 2 | 2.95E−01 | 6.95E−04 | 5.75E−02 | 6.6443 |
| 14641 | 2 | 144480 | 3 | 4.58E−02 | 4.93E−04 | 1.62E−06 | 1.0305 |
| 14641 | 3 | 303120 | 3 | 4.44E−02 | 4.93E−04 | 2.30E−11 | 1.0003 |
| 40401 | 1 | 120400 | 2 | 1.77E−01 | 2.50E−04 | 2.07E−02 | 6.6443 |
| 40401 | 2 | 400800 | 3 | 2.71E−02 | 1.78E−04 | 2.10E−07 | 1.0183 |
| 40401 | 3 | 841200 | 3 | 2.67E−02 | 1.78E−04 | 1.07E−12 | 1.0002 |

**Table 2** Example 1: $k_1 = 2$, $k_2 = 3$, and $p_1 = 2$

| $N_1$ | $p_2$ | $N_2$ | $k$ | $\overline{\mathfrak{M}}^2(v, \mathbf{y}_k, \beta_k)$ | $\overline{\mathfrak{M}}_k^{\text{Dual}}$ | $\overline{\mathfrak{M}}_k^{\text{Equi}}$ | $I_{\text{eff}}$ |
|---|---|---|---|---|---|---|---|
| 1681 | 1 | 1240 | 3 | 2.60E+01 | 3.94E−01 | 6.10E+02 | 189.9638 |
| 1681 | 2 | 4080 | 3 | 2.15E+00 | 6.05E−03 | 3.92E+00 | 15.6634 |
| 1681 | 3 | 8520 | 2 | 2.53E−01 | 4.71E−03 | 1.26E−02 | 1.8496 |
| 6561 | 1 | 4880 | 3 | 1.32E+01 | 9.51E−02 | 1.56E+02 | 380.2599 |
| 6561 | 2 | 16160 | 3 | 5.41E−01 | 3.89E−04 | 2.49E−01 | 15.6199 |
| 6561 | 3 | 33840 | 3 | 4.93E−02 | 3.00E−04 | 1.99E−04 | 1.4258 |
| 25921 | 1 | 19360 | 3 | 6.60E+00 | 2.36E−02 | 3.94E+01 | 760.6287 |
| 25921 | 2 | 64320 | 2 | 1.35E−01 | 2.45E−05 | 1.56E−02 | 15.6082 |
| 25921 | 3 | 134880 | 3 | 1.05E−02 | 1.88E−05 | 3.12E−06 | 1.2139 |
| 58081 | 1 | 43440 | 3 | 4.40E+00 | 1.05E−02 | 1.75E+01 | 1140.9677 |
| 58081 | 2 | 144480 | 2 | 6.02E−02 | 4.84E−06 | 3.09E−03 | 15.6060 |
| 58081 | 3 | 303120 | 2 | 4.41E−03 | 3.72E−06 | 2.74E−07 | 1.1430 |

practically present the exact minimizer of the majorant, since the efficiency index is almost one. Note that in this case $\overline{\mathfrak{M}}^{\text{Dual}}$ is almost the exact error and $\overline{\mathfrak{M}}^{\text{Equi}}$ vanishes. Results of a similar experiment in the case $k_1 = 2$, $k_2 = 3$, and $p_1 = 2$ are depicted in Table 2. It is easy to see that lowest order Raviart-Thomas elements are not able to represent the minimizer properly and in the case $p_2 = 1$, the efficiency index of the majorant is poor. Again, in the $p$-refined spaces the estimate improves significantly.

*Example 2* Let $\Omega := (0, 1) \times (0, 1) \times (0, 1)$, $f(x_1, x_2, x_3) = x_1 x_2 x_3$, and

$$\mathbf{A} = \begin{pmatrix} 1000 & 20 & -500 \\ -3 & 30 & 16 \\ 2 & 0 & 3 \end{pmatrix}.$$

Then,

$$\mathbf{A}^{-1} \approx \begin{pmatrix} 7.4490978E-04 & -4.9660652E-04 & 1.2680020E-01 \\ 3.3934779E-04 & 3.3107104E-02 & -1.2001324E-01 \\ -4.9660652E-04 & 3.3107101E-04 & 2.4879987E-01 \end{pmatrix}$$

and

$$\mathbf{B} \approx \begin{pmatrix} 1.0126139 & -0.4980245 & 2.0416897 \\ -0.0160603 & 0.5154516 & -0.0408795 \\ -0.0060666 & 0.009230 & -0.0280656 \end{pmatrix}.$$

In Example 2, the exact solution is not known. Instead a reference solution was computed using third order Courant type elements with 29791 global degrees of freedom. The approximations were computed using linear tetrahedral Courant type elements and the space $Q_h$ is generated using tetrahedral Raviart-Thomas elements

**Table 3** Example 2, $p_1 = 1$

| $N_1$ | $p_2$ | $N_2$ | $k$ | $\overline{\mathfrak{M}}^2(v, \mathbf{y}_k, \beta_k)$ | $\overline{\mathfrak{M}}_k^{\text{Dual}}$ | $\overline{\mathfrak{M}}_k^{\text{Equi}}$ | $I_{\text{eff}}$ |
|---|---|---|---|---|---|---|---|
| 125 | 1 | 864 | 4 | 4.67E−02 | 1.15E−05 | 1.57E−03 | 10.0122 |
| 125 | 2 | 3744 | 3 | 8.47E−03 | 6.99E−06 | 9.43E−06 | 1.8164 |
| 125 | 3 | 9792 | 3 | 5.26E−03 | 6.68E−06 | 8.94E−09 | 1.1284 |
| 343 | 1 | 2808 | 3 | 3.12E−02 | 5.81E−06 | 6.85E−04 | 9.2258 |
| 343 | 2 | 12312 | 3 | 5.24E−03 | 3.65E−06 | 1.86E−06 | 1.5489 |
| 343 | 3 | 32400 | 3 | 3.81E−03 | 3.65E−06 | 7.85E−10 | 1.1241 |
| 729 | 1 | 6528 | 3 | 2.35E−02 | 3.48E−06 | 3.83E−04 | 9.1082 |
| 729 | 2 | 28800 | 3 | 3.78E−03 | 2.21E−06 | 5.88E−07 | 1.4642 |
| 729 | 3 | 76032 | 3 | 2.97E−03 | 2.64E−06 | 1.40E−10 | 1.1527 |
| 1331 | 1 | 12600 | 3 | 1.88E−02 | 2.31E−06 | 2.44E−04 | 7.8120 |
| 1331 | 2 | 55800 | 3 | 2.94E−03 | 1.47E−06 | 2.41E−07 | 1.2208 |

of order $p_2$. The results were depicted on Table 3 and they show similar characteristics as in the two dimensional example.

## 5 Summary

An upper functional deviation estimate (majorant) for nonsymmetric stationary diffusion problem is derived. An algorithm for the global minimization of the majorant over a finite dimensional subspace is presented and tested. The efficiency of the majorant depends on the particular problem (i.e., the exact solution) and the relation of spaces $V_h$ and $Q_h$. The question is that how accurately $V_h$ can represent $u$ (in the energy norm) in comparison with the ability of $Q_h$ to represent the minimizer of the majorant. If $Q_h$ is "better", then the estimate is very accurate and the other way round. The crude overestimation in Table 2 shows that using a "worse" space for the computation of the minimizer.

## References

1. V.V. Denisenko, Variational methods for elliptic boundary value problems that describe transport processes with nonsymmetric tensor coefficients. Zh. Prikl. Mekh. i Tekhn. Fiz. **3**, 69–75 (1989). Translation in J. Appl. Mech. Tech. Phys. **30**(3), 404–410 (1989)
2. V.V. Denisenko, The energy method for three-dimensional elliptic equations with nonsymmetric tensor coefficients. Sibirsk. Mat. Zh., **38**(6), 1267–1281, ii (1997). Translation in Siberian Math. J. **38**(6), 1099–1111 (1997)
3. F. Hecht, New development in freefem++. J. Numer. Math. **20**(3–4), 251–265 (2012)

4. A. Logg, K.-A. Mardal, G. Wells (eds.), *Automated Solution of Differential Equations by the Finite Element Method*. Lecture Notes in Computational Science and Engineering, vol. 84 (Springer, Berlin, 2012)
5. O. Mali, P. Neittaanmäki, S. Repin, *Accuracy Verification Methods: Theory and Algorithms*. Computational Methods in Applied Sciences, vol. 32 (Springer, Dordrecht, 2014)
6. P. Neittaanmäki, S. Repin, *Reliable Methods for Computer Simulation: Error Control and a Posteriori Estimates* (Elsevier, Amsterdam, 2004)
7. P.-A. Raviart, J.M. Thomas, Primal hybrid finite element methods for 2nd order elliptic equations. Math. Comp. **31**(138), 391–413 (1977)
8. S. Repin, A posteriori error estimates for approximate solutions of variational problems with strongly convex functionals. Probl. Mat. Anal. **17**, 227–237 (1997). Translation in J. Math. Sci. **97**(4), 4311–4328 (1999)
9. S. Repin, *A Posteriori Estimates for Partial Differential Equations*. Radon Series on Computational and Applied Mathematics, vol. 4 (Walter de Gruyter, Berlin, 2008)

http://www.springer.com/978-3-319-23563-9

Mathematical Modeling and Optimization of Complex Structures
Neittaanmäki, P.; Repin, S.; Tuovinen, T. (Eds.)
2016, XXI, 328 p., Hardcover
ISBN: 978-3-319-23563-9