Probabilistic Analysis of an a Posteriori Error Estimator for Finite Elements

P. Díez
J. J. Egozcue

Publication CIMNE Nº PB-175, October 1999

International Center for Numerical Methods in Engineering
Gran Capitán s/n, 08034 Barcelona, Spain

Mathematical Models and Methods in Applied Sciences
© World Scientific Publishing Company

PROBABILISTIC ANALYSIS OF AN A POSTERIORI ERROR ESTIMATOR FOR FINITE ELEMENTS

Pedro DÍEZ and Juan José EGOZCUE
Departament de Matemàtica Aplicada III
E.T.S. Ingenieros de Caminos
Universitat Politècnica de Catalunya
Modulo C-2 Campus Nord, E-08034 Barcelona, SPAIN
e-mail: pedro.diez@upc.es

A residual type a posteriori error estimator for Finite Elements is analyzed using a new technique. In this case, the error estimate is the result of two consecutive projections of the exact error on two finite-dimensional subspaces. The analysis introduced in this paper is based on a probabilistic approach, that is, the idea is to assess the average value of the effectivity index (the ratio estimated error over exact error) by assuming the randomness of the exact error. The average value characterizes the mean behavior of the estimator and it is found to be related with some geometric properties of the subspaces. These geometric properties are obtained from the standard matrices of the linear systems arising in the formulation of the Finite Element Method.

Keywords: Error estimation, probabilistic analysis, finite elements, adaptivity

1. Introduction

A posteriori error estimators are needed to perform practical Finite Element (FE) computations and to control the quality of the numerical solution. They are also required to drive adaptive procedures leading to optimal meshes[5].

The analysis of a posteriori error estimators for finite elements is usually performed in terms of finding lower and upper bounds of the effectivity index[1,7] which is the ratio of the estimated error and the exact error. This may be seen, in fact, as an a priori analysis of the a posteriori error estimator. This kind of analysis is used to ensure a good behavior of the estimator in the asymptotic range. Nevertheless, the practical application of an error estimator to a FE computation is usually far from this asymptotic range and, consequently, this analysis does not furnish any clue about the actual behavior of the estimator in a current case. Moreover, often the proofs of the bounds of the effectivity index assume that some Superconvergence properties are verified and, usually, these properties are valid only in very particular cases.

This work introduces a new approach to the analysis of error estimators using a probabilistic viewpoint. Instead of finding pessimistic bounds of the behavior of the studied error estimator under certain assumptions, the exact error is assumed to be a random function. Thus, the goal of the analysis of the error estimator is to assess the expected values of the effectivity index. This analysis is applied to a specific residual type error estimator based on the approximation of a h-refined reference solution[2,3]. In the analyzed error estimator the error is undervaluated, that is, the (measure of the) estimated error is lower than (measure of the) exact error and, consequently, the effectivity index is lower than one. The effectivity index is therefore a random variable that ranges from zero to one. If the expected value of this random variable is close to one, the estimator shows a good average behavior.

In the case of the analyzed error estimator, the average behavior may be predicted a priori estimating some geometric magnitudes related to the subspaces used in the error estimation procedure.

The remainder of the paper is structured as follows: in section 2 the analyzed error estimator is briefly described, in section 3 the main idea and the basis of the probabilistic approach to the error estimator analysis are introduced, in section 4 average behavior of the estimator is predicted using some geometric characteristics of the involved projection subspaces, finally, in section 5, a numerical example is presented that shows the agreement of the predictions with the real behavior of the estimator.

2. Description of the error estimator: multiprojection strategy

The Finite Element Method (FEM) is used to solve a partial differential equation with an unknown solution u. The exact solution u belongs to a functional space V and the FEM provides an approximate solution u_h, lying in a finite-dimensional interpolation space V_h. The interpolation space V_h is generated by a mesh of finite elements with characteristic size h. Since the exact solution, u, is unknown, the error $e := u - u_h$ is also unknown. Nevertheless, the solution can be as accurate as desired: it suffices to reduce enough the element size, h (or, alternatively, to increase enough the order of the interpolation, p). Then, if a new mesh is considered of characteristic size $\tilde{h}$ ($\tilde{h}$ much smaller than h) the associated solution $u_{\tilde{h}}$, lying in the interpolation space $V_{\tilde{h}}$, is much more accurate than u_h. This new fine mesh is denoted as the reference mesh and $u_{\tilde{h}}$ is denoted as the reference solution. Consequently, a reference

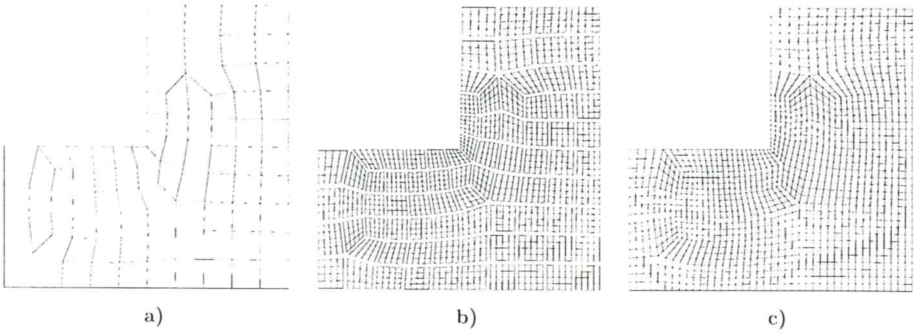
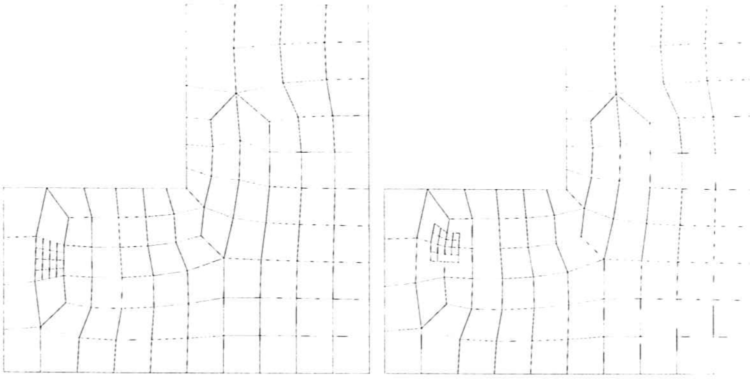


Fig. 1. Computational mesh of characteristic size h, **a)**, set of elementary submeshes discretizing the elements, **b)**, and refined reference mesh of characteristic size $\tilde{h}$, **c)**.

error, $e_{\tilde{h}} := u_{\tilde{h}} - u_h$, which fairly approximates the exact error e is introduced. For elliptic self-adjoint problems the Finite Element solution, u_h, is the projection of the exact solution, u, on V_h following the energetic scalar product that appears in the weak form of the problem. In the following, this scalar product is denoted by $< \cdot, \cdot >$ and the induced energy norm is denoted by $\| \cdot \|$.

The goal of a posteriori error estimators is to assess the magnitude and the distribution of e. However, the problem of obtaining e is as difficult as to obtain u, because both u and e belong to the infinite-dimensional functional space V. On the contrary, the reference error $e_{\tilde{h}}$ lies in a finite-dimensional space $V_{\tilde{h}}$ and may replace the exact error e. Thus, the goal of error estimators is switch to assess $e_{\tilde{h}}$. However, the computational cost of determining $e_{\tilde{h}}$ by standard finite element analysis is unaffordable for error estimation purposes. The mesh generating $V_{\tilde{h}}$ is much finer than the computational mesh generating $V_{\tilde{h}}$ and, consequently, the cost of obtaining $u_{\tilde{h}}$ (or $e_{\tilde{h}}$) is much larger than the cost of computing u_h. Then, a practical error estimator must preclude the unaffordable global computation of $e_{\tilde{h}}$ and has to approximate $e_{\tilde{h}}$ by inexpensive local computations.

The error estimator introduced by Díez et al. [3,8] is based on the previous idea. A reference mesh is build up assembling a set of elementary submeshes discretizing each one of the elements of the computational mesh, see Fig. 1. The error estimation is splitted into two phases: the interior estimation and the patch estimation. In the interior estimation, the error $e_{\tilde{h}}$ is projected on the space $V^{\mathrm{I}} \subset V_{\tilde{h}}$ generated by all the elementary submeshes, see Fig. 1 **b)**. The space V^{I} is smaller than $V_{\tilde{h}}$ because the nodal values of the nodes lying in the boundary of the elements of the computational mesh are set to zero. The projection



a) b)

Fig. 2. Elementary submesh discretizing the interior of one element, **a)** and patch submesh, centered on one node and covering the edges of the adjacent elements, **b)**.

of $e_{\tilde{h}}$ on V^{I}, ε, is the interior estimate. A new set of submeshes, the patch submeshes, is defined covering the boundary of the elements of the computational mesh, see Fig. 2. The patch estimation phase is carried out to obtain an enriched estimate: the error is projected on the space $V^{\mathrm{P}} \subset V_{\tilde{h}}$ generated by all the patch submeshes. The projection of $e_{\tilde{h}}$ on V^{P}, η, is the patch estimate which has nonzero values on the boundary of the elements. In order to be able to add the interior and the patch estimates to get a complete estimate, the patch estimate, η, is forced to be orthogonal to the interior estimate, ε. The complete estimate, $e_{_L} := \varepsilon + \eta$, is a residual type error estimate obtained by local computations.

Thus, in fact the complete estimate is computed through two projections, the first is free (interior estimate, projection on V^{I}), the latter is restricted with an orthogonality constrain (patch estimate, projection on V^{P} orthogonal to ε). This process is denoted multiprojection strategy.

For analysis purposes, the error estimation process is splitted into three phases.

1. From e to $e_{\tilde{h}}$ (from an infinite-dimensional problem to a finite-dimensional one).

2. From $e_{\tilde{h}}$ to the projection of $e_{\tilde{h}}$ on $V^{\mathrm{I}} + V^{\mathrm{P}}$, denoted by $e^*_{\tilde{h}}$. This phase accounts for the effect of the remaining hidden points.

3. From the latter to $e_{_L}$ (pure multiprojection).

None of these phases correspond to the practical computation of the estimate. However, each one of these phases is a projection and can be analyzed independently. The undervaluation introduced in each projection is analyzed separately.

The analysis of each one of these phases is carried out using different techniques.

1. The first phase can be studied using a priori error estimates and Richardson extrapolation to obtain an explicit expression:

$$\|u_{\tilde{h}} - u_h\| = \|e_{\tilde{h}}\| \simeq \left[1 - \left(\frac{\tilde{h}}{h}\right)^{2p}\right]^{1/2} \|e\|, \tag{2.1}$$

   where p stands for the degree of the interpolating polynomial. That is, if $\tilde{h}$ is one fourth of h and p is one, the reference error, $e_{\tilde{h}}$, is 97% of the exact error e.

2. The second phase is, in fact, a cancellation of a small amount of degrees of freedom. It accounts for the effect of the points where the error estimate is forced to be zero. This can be seen as a single projection (from $V_{\tilde{h}}$ on to $V^{\mathrm{I}} + V^{\mathrm{P}}$) and the undervaluation introduced in this phase is assessed in section 4.1.

3. The third phase contains the essence of the multiprojection strategy: the assessment of the efficiency of this phase is carried out in section 4.2.

3. Probabilistic analysis

In order to assess the efficiency of the error estimator introduced above, the obtained estimate should be compared with the exact error. This is usually done in terms of the effectivity index, ν, introduced by Zienkiewicz and Zhu [10] and defined as

$$\nu := \frac{\text{estimated error}}{\text{exact error}} = \frac{\|e_{L}\|}{\|e\|} \,. \tag{3.2}$$

Of course, the exact error is unknown and the previous definition is often replaced by the following

$$\nu := \frac{\text{estimated error}}{\text{reference error}} = \frac{\|e_{L}\|}{\|e_{\tilde{h}}\|} \,. \tag{3.3}$$

Nevertheless, even $e_{\tilde{h}}$ may only be computed in academic problems.

The theoretical analysis of the a posteriori error estimators is based on the a priori behavior of the finite element solution. As previously said, this analysis furnishes pessimistic bounds of the effectivity index depending on unknown constants, see for instance [1,7]. Thus, this kind of analysis is only a theoretical tool allowing to describe the asymptotic behavior of the a posteriori estimators. The constants appearing in the bounds cannot be computed. Moreover, although the values of the bounds could be obtained, the exact value of the effectivity index would be unknown. Then, for a given problem, no piece of information regarding the expected behavior of the estimators is supplied.

The error is unknown and, therefore, some additional assumption must be used to characterize the behavior of the error estimator. Instead of considering the worse situation, leading to the pessimistic bounds mentioned above, here, the error is considered as a random function. That allows to assess the expected value of the effectivity index.

In the following the randomness is characterized in the simplest way. Then, without any previous consideration about the error, its probability distribution is assumed to be uniform. In fact, the study is carried out using the definition of Eq. (3.3) because $e_{\tilde{h}}$ lays in the finite dimensional space $V_{\tilde{h}}$. Then, in the remainder of the paper $e_{\tilde{h}}$ is a random vector in a finite dimensional space.

The goal of this paper is to assess the behavior of the effectivity index ν which is understood as a scalar function defined over $V_{\tilde{h}}$. The estimator can be seen as a function from $V_{\tilde{h}}$ to $V_{\tilde{h}}$, mapping the reference error function $e_{\tilde{h}}$ into the estimate e_{L}. In the following, the value of the effectivity index defined in Eq. (3.3) is denoted by $\nu(e_{\tilde{h}})$. Since $e_{\tilde{h}}$ is assumed to be a random vector in $V_{\tilde{h}}$, $\nu(e_{\tilde{h}})$ is a random number. The average value of $\nu(e_{\tilde{h}})$ has to be assessed.

Notice that, for any real number β ($\beta \neq 0$),

$$\nu(\beta e_{\tilde{h}}) = \nu(e_{\tilde{h}}) \ , \tag{3.4}$$

that is, the value of the effectivity index does not depend on the "size" of the error but only on the "direction" of $e_{\tilde{h}}$ (if $e_{\tilde{h}}$ is seen as a vector). Then, in order to study the mean behavior of the effectivity index, the norm of vector $e_{\tilde{h}}$ is taken constant, that is, $e_{\tilde{h}}$ is assumed to yield in a hypersphere $S_{R}(V_{\tilde{h}}) := \{e_{\tilde{h}} \in V_{\tilde{h}} \text{ such that } \|e_{\tilde{h}}\| = R\}$ of radius R. Thus, the mean value of $\nu(e_{\tilde{h}})$ is defined as

$$\phi := \frac{1}{\text{meas}(S_{R}(V_{\tilde{h}}))} \int_{S_{R}(V_{\tilde{h}})} \nu(e_{\tilde{h}}) dS \ , \tag{3.5}$$

where $\text{meas}(S_{R}(V_{\tilde{h}}))$ stands for the measure of $S_{R}(V_{\tilde{h}})$ and the density of probability of $e_{\tilde{h}}$ is assumed to be uniform over $S_{R}(V_{\tilde{h}})$. The measure of $S_R(V_{\bar{h}})$ depends on the dimension of the ambient space $V_{\bar{h}}$, n, and may be analytically computed[2]:

$$\text{meas}(S_R) = \begin{cases} R^{n-1} \frac{(2\pi)^{n/2}}{2\cdot 4\cdots(n-2)} & \text{for even } n \\ R^{n-1} \frac{2(2\pi)^{(n-1)/2}}{3\cdot 5\cdots(n-2)} & \text{for odd } n \ . \end{cases} \tag{3.6}$$

Since the error estimator is a combination of error projections, the obtained estimate undervaluates the exact error and, consequently, ϕ is lower than one. In fact, if $\phi = 1$, the projection strategy is said to be optimal because for every $e_{\bar{h}}$, $\|e_L\| = \|e_{\bar{h}}\|$. Usually the squared norms of the error magnitudes are easier to handle. Consequently, it is convenient to introduce the mean value of $[\nu(e_{\bar{h}})]^2$,

$$\psi := \frac{1}{\text{meas}(S_R(V_{\bar{h}}))} \int_{S_R(V_{\bar{h}})} [\nu(e_{\bar{h}})]^2 dS \ . \tag{3.7}$$

The definition of ψ is introduced because the expressions for ψ are much simpler than the expressions for ϕ (see section 4). In fact, both ϕ and ψ can be used to evaluate the average underestimation and the optimality of the projection strategy is also equivalent to $\psi = 1$.

4. Analysis of the average behavior of the estimator

This section deals with the assessment of the main undervaluation introduced in the transformation from $e_{\bar{h}}$ to e_L. Attending to the three steps of the analysis introduced in section 2, this transformation is splitted into two partial transformations: from $e_{\bar{h}}$ to $e^*_{\bar{h}}$ and from $e^*_{\bar{h}}$ to $e_L := \varepsilon + \eta$. The first one is a single projection from the space $V_{\bar{h}}$ to the space $V^{\text{I}} + V^{\text{P}}$ (recall that $V^{\text{I}} + V^{\text{P}} \subset V_{\bar{h}}$) and, therefore, the average efficiency of the projection depends only on the dimensions of $V_{\bar{h}}$ and $V^{\text{I}} + V^{\text{P}}$. The second one is more complex because ε is a single projection of $e^*_{\bar{h}}$ on V^{I} but η is obtained via a restricted projection on V^{P}.

4.1. *Analysis of the efficiency of a single projection*

Let us denote by n the dimension on $V_{\bar{h}}$ and by $n-m$ the dimension of $V^{\text{I}} + V^{\text{P}}$. Thus, m dimensions are lost in the transformation from $e_{\bar{h}}$ to $e^*_{\bar{h}}$. The average efficiency of this transformation is obviously a function of n and m. The expression of ϕ as a function of n and m is explicitly found using a set of generalized hyperspherical coordinates[2,4]. This expression is quite complex and depends on the parity of n and m:

$$\phi = \frac{(n-m)\cdot(n-m+2)\cdots(n-2)}{(n-m+1)\cdot(n-m+3)\cdots(n-1)} \tag{4.8}$$

stands for even m, while the expression for odd m must be splitted:

$$\phi = \frac{2}{\pi} \frac{[2 \cdot 4 \cdots (n-m-1)]^2 (n-m+1) \cdot (n-m+3) \cdots (n-2)}{[3 \cdot 5 \cdots (n-m-2)]^2 (n-m) \cdot (n-m+2) \cdots (n-1)} \tag{4.9}$$

for even n, and

$$\phi = \frac{\pi}{2} \frac{[3 \cdot 5 \cdots (n-m-1)]^2 (n-m+1) \cdot (n-m+3) \cdots (n-2)}{[2 \cdot 4 \cdots (n-m-2)]^2 (n-m) \cdot (n-m+2) \cdots (n-1)} \tag{4.10}$$

for odd n.

The expression for ψ (recall Eq. (3.7)) is much simpler and does not depend on the parity on n and m:

$$\psi = 1 - \frac{m}{n} \, . \tag{4.11}$$

Recall that the dimension of $V_{\tilde{h}}$, n, is directly related with the number of nodes in the reference mesh generating $V_{\tilde{h}}$. On the other hand, m is related with the number of hidden points. Thus, the values of n and m are determined by simple node (that is, degrees of freedom) counting. Once n and m are known, ϕ and ψ are obtained using Eqs. (4.8-4.10) and Eq. (4.11), respectively. Using simple numerical experimentation it can be found that, for large values of n ($n \geq 100$), ϕ is very similar to $\sqrt{\psi}$ (error less than 1%), that is, the variance of ν, given by $\psi - \phi^2$, is small.

Thus, two main conclusions may be extracted of the study the mean effectivity of the single projection. First, it is worth to note that analytical expressions are available for the expected value of the efficiency of the single projection. Second,the resulting analytical expression for ψ is much simpler than the expression for ϕ. As it is shown in section 4.2,this stands also for the evaluation of the efficiency of the multiprojection strategy and, consequently, ψ is preferred to measure the mean undervaluation of the complete multiprojection.

4.2. *Analysis of the efficiency of the multiprojection*

The main goal of this section is to describe the expected behavior of the multiprojection strategy and to relate it with some properties of the involved subspaces V^{I} and V^{P}. These properties are measured by a set of magnitudes that may be interpreted under a geometric viewpoint and that can be evaluated for given subspaces. Then, once the values of these magnitudes are known, they are used to compute the expected value of the (squared) effectivity index ψ and assess the performance of the error estimator.

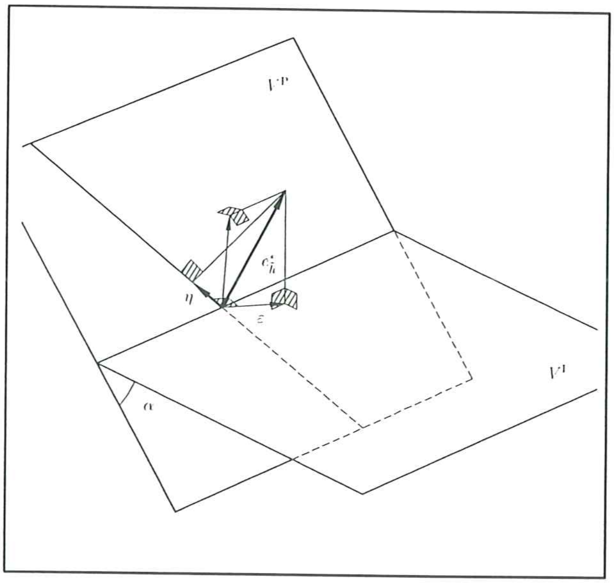


Fig. 3. Geometric illustration of the multiprojection strategy in a simple case.

Let us study first a very particular case of multiprojection. For this purpose we set $\dim V^{\mathrm{I}} = 2$, $\dim V^{\mathrm{P}} = 2$ and $\dim V^{\mathrm{I}} + V^{\mathrm{P}} = 3$, that is, V^{I} and V^{P} may be seen as two planes in a three dimensional space. Then, the multiprojection may be seen as a way to approximate a vector $e^*_{\tilde{h}}$ in $\mathbb{R}^3$ by the sum of two vectors, $e^*_{\tilde{h}} \approx e_{L} := \varepsilon + \eta$. The first, ε, is the projection of $e^*_{\tilde{h}}$ on the plane V^{I} and the second, η, is the projection of $e^*_{\tilde{h}}$ of the straight line in V^{P} orthogonal to ε, see Fig.3. The aim is to characterize the main value, ϕ, of $\|e_{L}\|/\|e^*_{\tilde{h}}\|$ or, alternatively, the main value, ψ, of $\|e_{L}\|^2/\|e^*_{\tilde{h}}\|^2$ for $e^*_{\tilde{h}}$ ranging in the unit sphere.

It is worth noting that, in this case, the relative position of the two planes V^{I} and V^{P} is characterized by the diedric angle α and, consequently, the values of ψ and ϕ are functions of α. In the limit case $\alpha = 0$, V^{I} and V^{P} are the same plane and the second projection is useless. Then, for $\alpha = 0$, we are in the case of section 4.1with $n = 3$ and $m = 1$, consequently $\phi = \pi/4 \approx 0.785$ and $\psi = 2/3 \approx 0.667$. Moreover, for the particular case of perpendicular planes ($\alpha = \pi/2$), it can be easily seen, see Fig. 4, that the multiprojection is optimal and therefore $\psi = \phi = 1$. In this simple case both ϕ and ψ may be easily computed for every intermediate value of α between 0 and π. The curves representing ϕ and ψ versus α are plotted in Fig. 5.

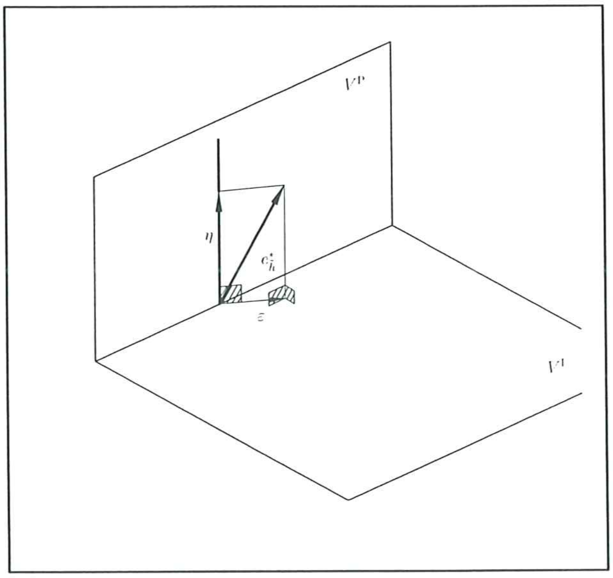


Fig. 4. Geometric illustration of the behavior of the multiprojection strategy for perpendicular subspaces.

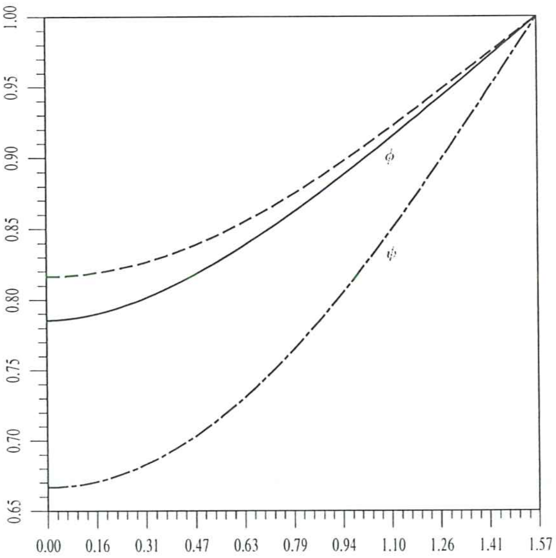


Fig. 5. Representation of ϕ and ψ versus α. The upper dashed line is the representation of $\sqrt{\psi}$.

The values of ϕ plotted in Fig. 5 have been numerically computed using a standard quadrature over the sphere because we did not find any analytical expression for ϕ as a function of α. On the contrary, it can be easily find as a particular case of the developments showed in the following, that ψ has a very simple expression:

$$\psi = 1 - \frac{1}{3} \cos \alpha \ . \tag{4.12}$$

Of course, for $\alpha = 0$ the value given by Eq. (4.12) coincides with Eq. (4.11) for $n = 3$ and $m = 1$. The variance of ν (recall Eq. (3.3)) is $\psi - \phi^2$ and must be positive, consequently $\sqrt{\psi} \geq \phi \geq \psi$. This is valid for all the cases and not only for this particular case with geometric interpretation. However, in this case the relative position of the subspaces V^{I} and $V^{\mathrm{P}}$is controlled by only one parameter, the angle α, and the relationship $\sqrt{\psi} \geq \phi \geq \psi$ may be easily illustrated in Fig. 5 .

From this example we can conclude that, in this particular case, the average behavior of the multiprojection process (that is, the values of ϕ and ψ) is directly related with the angle α that describes the relative position of V^{I} and V^{P}. Moreover, while ϕ requires to be computed using a numerical quadrature, ψ has a very simple analytical expression.

In the remainder of this section this result is generalized for any dimensions of the subspaces V^{I} and V^{P}. First, an extended definition of perpendicularity is given that ensures optimality of the multiprojection strategy. Second, the general multiprojection is analyzed in a proper basis in order to obtain an algebraic expression for ψ as simple as possible. Third, this expressions are related with geometric magnitudes, analogous to the diedric angle α, that characterize the relative position of the two subspaces.

Let us denote by n_c the dimension of $V^{\mathrm{I}} \cap V^{\mathrm{P}}$ and introduce the following notation:

$$n_c := \dim(V^{\mathrm{I}} \cap V^{\mathrm{P}}),\ n_i := \dim V^{\mathrm{I}} - n_c \ \text{ and } \ n_p := \dim V^{\mathrm{P}} - n_c \tag{4.13}$$

Thus, $\dim(V^{\mathrm{I}} + V^{\mathrm{P}}) = n_i + n_c + n_p$. Let us build up a particular basis of $V^{\mathrm{I}} + V^{\mathrm{P}}$, $\mathcal{B}$ that simplifies the algebraic expressions in the following. We first select an orthonormal basis, $\mathcal{B}_c$ of $V^{\mathrm{I}} \cap V^{\mathrm{P}}$. Then, two families of vectors $\mathcal{B}_i$ and $\mathcal{B}_p$ are given such that $\mathcal{B}_i \cup \mathcal{B}_c$ is an orthonormal basis of V^{I} ($\mathcal{B}_i$ completes an orthonormal basis of V^{I}) and $\mathcal{B}_c \cup \mathcal{B}_p$ is an orthonormal basis of V^{P} ($\mathcal{B}_p$ completes an orthonormal basis of V^{P}). Then, $\mathcal{B} := \mathcal{B}_i \cup \mathcal{B}_c \cup \mathcal{B}_p$ is a basis of $V^{\mathrm{I}} + V^{\mathrm{P}}$ such that the matrix of the scalar product, $< \cdot, \cdot >$, in $\mathcal{B}$ has a quite simple shape:

$$[< \cdot, \cdot >]_{\mathcal{B}} = \begin{pmatrix} I_{n_i} & 0_{n_i \times n_c} & A \\ 0_{n_c \times n_i} & I_{n_c} & 0_{n_c \times n_p} \\ A^T & 0_{n_p \times n_c} & I_{n_p} \end{pmatrix} , \tag{4.14}$$

where for any value of n_1 and n_2, I_{n_1} is the $n_1 \times n_1$ identity matrix, $0_{n1 \times n_2}$ is the $n_1 \times n_2$ null matrix, and A is a rectangular $n_i \times n_p$ which contains the cross scalar products of the elements of the bases $\mathcal{B}_i$ and $\mathcal{B}_p$. For the particular simple case with geometrical interpretation discussed above, we have $n_i = n_c = n_p = 1$ and A is a scalar that coincides with $\cos\alpha$, such that the expression of the matrix of Eq. (4.14) particularizes in

$$[< \cdot, \cdot >]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & \cos\alpha \\ 0 & 1 & 0 \\ \cos\alpha & 0 & 1 \end{pmatrix}. \tag{4.15}$$

Using the basis $\mathcal{B}$, the error $e^*_{\bar{h}}$ to be approximated with the multiprojection is expressed in vectorial form:

$$\left[e^*_{\bar{h}}\right]_{\mathcal{B}} = \begin{pmatrix} e_i \\ e_c \\ e_p \end{pmatrix}. \tag{4.16}$$

Then, the projections ε and η may be expressed in terms of $[< \cdot, \cdot >]_{\mathcal{B}}$ and $\left[e^*_{\bar{h}}\right]_{\mathcal{B}}$. After some algebra, explicit expressions for the components of the multiprojection are found[2]:

$$[\varepsilon]_{\mathcal{B}} = \begin{pmatrix} e_i + Ae_p \\ e_c \\ 0 \end{pmatrix} \tag{4.17}$$

and

$$[\eta]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \eta_c \\ \eta_p \end{pmatrix}, \tag{4.18}$$

where

$$\begin{pmatrix} \eta_c \\ \eta_p \end{pmatrix} = \left\{ I_{n_c+n_p} - \frac{1}{\mathbf{a}^T\mathbf{a}}\mathbf{a}\mathbf{a}^T \right\} \begin{pmatrix} e_c \\ A^T e_i + e_p \end{pmatrix} \tag{4.19}$$

and $\mathbf{a}$ is a $n_c + n_p$ vector defined by:

$$\mathbf{a} := \begin{pmatrix} e_c \\ A^T e_i + A^T A e_p \end{pmatrix}. \tag{4.20}$$

As shown in Eqs. (4.17) and (4.19 the result of the multiprojection strategy depends on (and only on) the rectangular matrix A. In fact, all the information regarding the relative position of the two subspaces V^{I} and V^{P} is contained in A. Hence, the expected values ϕ and ψ are also functions of A. The matrix A may be decomposed as

$$A = B\Sigma C^T, \tag{4.21}$$

using the Singular Value Decomposition (SVD)[9], where B and C are $n_i \times n_i$ and $n_p \times n_p$ unit matrices and Σ is a $n_i \times n_p$ diagonal matrix with diagonal entries σ_i, $i = 1, \dots, \min(n_i, n_p)$ such that

$$\Sigma^T\Sigma = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_{n_p}^2 \end{pmatrix} \text{ and } \Sigma\Sigma^T = \begin{pmatrix} \sigma_1^2 & & & \\ & \sigma_2^2 & & \\ & & \ddots & \\ & & & \sigma_{n_i}^2 \end{pmatrix}, \tag{4.22}$$

where the scalars σ_i, $i = 1, \dots, \max(n_i, n_p)$, are the singular values of A. Denoting by r the rank of A, $\sigma_i = 0$ for $i > r$. Moreover, due to the definition of the basis $\mathcal{B}$ ($\mathcal{B}_i$ and $\mathcal{B}_p$ are both orthonormal) the singular values σ_i are lower than 1. The set of singular values, σ_i, $i = 1, \dots, r$, may be interpreted as the cosinus of a set of angles describing the relative position of V^{I} and V^{P}. Indeed, for $n_i = n_c = n_p = 1$ we are in the simple 3D case previously introduced and, for $\alpha \neq 0$, we have $r = 1$ and $\sigma_1 = \cos\alpha$. If all the singular values are close to 0, the spaces V^{I} and V^{P} are nearly "perpendicular". The notion of perpendicularity between two vectorial subspaces must be here understood in the sense of the following definition:

Definition 1 *Two vectorial subspaces V^{I} and V^{P} are said to be perpendicular if the orthogonal to the first, $(V^{\mathrm{I}})^{\perp}$, is included in the latter ($(V^{\mathrm{I}})^{\perp} \subset V^{\mathrm{P}}$).*

Remark 1 *The orthogonal space is defined in the ambient space $V^{\mathrm{I}} + V^{\mathrm{P}}$, that is,*

$$(V^{\mathrm{I}})^{\perp} = \left\{ \mathbf{x} \in V^{\mathrm{I}} + V^{\mathrm{P}} \textit{ such that } \forall \mathbf{y} \in V^{\mathrm{P}}, < \mathbf{x}, \mathbf{y} >= 0 \right\} \tag{4.23}$$

Remark 2 *The definition of perpendicularity is symmetric because*

$$(V^{\mathrm{I}})^{\perp} \subset V^{\mathrm{P}} \Leftrightarrow (V^{\mathrm{P}})^{\perp} \subset V^{\mathrm{I}} \tag{4.24}$$

As it has already been stated for the simple case with geometrical interpretation, perpendicularity implies optimality also in the general case.

Theorem 1. Let V^{I} and V^{P} be two perpendicular subspaces, then the multiprojection strategy is optimal, that is, for every $e_h^* \in V^{\mathrm{I}} + V^{\mathrm{P}}$, $e_h^* \neq 0$, $\nu(e_h^*) = 1$.

Proof. ε is the projection of e_h^* in V^{I}. Then, $\varepsilon^{\perp} := e_h^* - \varepsilon$ is in $(V^{\mathrm{I}})^{\perp}$. Using the perpendicularity condition, $(V^{\mathrm{I}})^{\perp} \subset V^{\mathrm{P}}$, it is found that $\varepsilon^{\perp} \in V^{\mathrm{P}}$ and, consequently, $\eta = \varepsilon^{\perp}$. Then, $e_{_L} = \varepsilon + \eta = e^*_{\tilde{h}}$ and $\nu(e^*_{\tilde{h}}) = \|e_{_L}\|/\|e^*_{\tilde{h}}\| = 1$ □.

As previously said, the singular values of A are related with the relative position of V^{I} and V^{P} and represent the cosinus of a set of generalized angles between V^{I} and V^{P}. In particular, if all the singular are 0 (A is a null matrix), the spaces V^{I} and V^{P} are perpendicular. If all the singular values are close to one (recall they must be lower than one) the angles between the subspaces are small, the spaces V^{I} and V^{P} are almost coincident and, consequently, the contribution of the second part of the projection is small. In this case, the projection strategy behaves almost as a single projection on the first space.

Thus, the expected value of the squared effectivity index, ψ is found to be ^{2}

$$\psi = 1 - \frac{1}{n}\sum_{i=1}^{r}\sigma_i^2 - \frac{1}{\mathrm{meas}(S_{_R})}\int_{e^*_{\tilde{h}} \in S_{_R}} \frac{\left[e_i^T\Sigma(I_{n_p} - \Sigma^T\Sigma)^{1/2}e_p\right]^2}{e_i^T\Sigma\Sigma^T e_i + e_c^T e_c} \tag{4.25}$$

where $n := n_i + n_c + n_p$ is the dimension of $V^{\mathrm{I}} + V^{\mathrm{P}}$ and the variable $e^*_{\tilde{h}}$ of the integral ranges in the hypersphere $S_{_R}$, recall Eq. (4.16). The integral in the right-hand-side term of Eq. (4.25) is a function of A and, in particular, of its singular values σ_i. However, this term is difficult to handle and we did not find any analytical expression for it. In the simple 3D case, setting $\sigma_1 = \cos\alpha$ and using the proper spherical coordinates, the expression of Eq. (4.12) is obtained from Eq. (4.25). Indeed, in this case, the integral in Eq. (4.25) is such that

$$\begin{aligned} &\frac{1}{\mathrm{meas}(S_{_R})}\int_{e^*_{\tilde{h}} \in S_{_R}} \frac{\left[e_i^T\Sigma(I_{n_p} - \Sigma^T\Sigma)^{1/2}e_p\right]^2}{e_i^T\Sigma\Sigma^T e_i + e_c^T e_c} = \\ &= \frac{1}{4\pi R^2}\int_{e^*_{\tilde{h}} \in S_R} \frac{\left[e_i \cos\alpha(1-\cos^2\alpha)^{1/2}e_p\right]^2}{\cos^2\alpha e_i^2 + e_c^2} \\ &= \frac{1}{3}\cos\alpha - \frac{1}{3}\cos^2\alpha, \end{aligned} \tag{4.26}$$

recall that, in this case, $n_i = n_c = n_p = r = 1$ and $n = 3$.

Thus, in the general case, the right-hand-side term of Eq. (4.25) is unknown. Nevertheless we can use this equation to find an approximate expression for ψ. Such approximate expression of ψ must be valid for the limit cases that have already been mentioned. For instance, if $r = 0$ (all $\sigma_i = 0$), $\psi = 1$ (perpendicular subspaces). Moreover, if $r = n_p$ and $\sigma_i = 1$, $i = 1, \ldots, n_p$ ($V^{\mathrm{P}} \subset V^{\mathrm{I}}$), ψ must coincide with the expression of Eq. (4.11) with $n = n_i + n_c + n_p$ and $m = n_p$.

All this requirements are fulfilled if the following approximation is assumed:

$$\psi \approx 1 - \frac{1}{n} \sum_{i=1}^{r} \sigma_i . \tag{4.27}$$

It must be remarked that the expression of Eq. (4.27) is not exact, that is, the integral of Eq. (4.25) has not been explicitly computed as a (simple) function of the singular values of A. However, the very simple expression of Eq. (4.27) gives a quite good answer to evaluate a number that must be greater than $1 - \frac{r}{n}$ and lower than $1 - \frac{1}{n}\sum_{i=1}^{r} \sigma_i^2$. Especially if it is noticed that, for the simple 3D case with analytical expression, the exact answer is a particular case of this general approximate expression.

Moreover, the next section shows an example demonstrating that the approximate evaluation fits in a general and realistic situation the exact value.

5. Numerical example

In this section the technique introduced above is used to a priori evaluate the expected value of the effectivity index of the studied estimator in a concrete problem. This can be used to predict and, hence, to correct the estimate. The problem is a standard example used in the validation of error estimators. The Poisson equation is solved in an L-shaped domain discretized with the mesh shown in Fig. 1 a). The proper source term and Dirichlet boundary conditions are imposed to obtain a given exact solution (in this case $u(x, y) = x^2 + y^2$). Thus, the exact error is evaluated and the behavior of the error estimate may be studied and compared with the exact error. This example has been used to demonstrate the robustness of the studied error estimator[3]. The global effectivity index is found to be 91.9% and the distribution of the local effectivity index (element by element) is quite uniform (ranging from 82% to 95%).

As previously said the analysis of the average undervaluation introduced in the error estimation is splitted into three phases that are studied independently.

First, the effect of approximating the exact error by a reference error belonging to a finite dimensional (even if fine) space is accounted for using Eq. (2.1). In this case bilinear elements are used ($p = 1$) and the refinement factor is 4 ($\tilde{h}/h = 0.25$). Thus,

$$\frac{\|e_{\tilde{h}}\|}{\|e\|} \simeq \left[1 - \left(\frac{\tilde{h}}{h}\right)^{2p}\right]^{1/2} = 0.968. \tag{5.28}$$

Second, the effect of the points where the estimated is forced to vanish is accounted for. The number of free nodes in reference mesh is 1425. The number of points where the estimate is forced to vanish (center points of the interior edges of the elements of the computational mesh) is 148. Then, the expected value of the undervaluation introduced in this second phase is described either by $\phi^1 425_{148} = 0.946628$ or $\psi^1 425_{148} = 089614$. Note that, as previously said, the value of $\sqrt{\psi} = 0.946647$ is very close to ϕ. Up to the required accuracy (3 significant digits) $\sqrt{\psi}$ and ϕ may be considered equal. Thus, the undervaluation associated with this phase is taken to be 0.947. That is,

$$\frac{\|e_{\tilde{h}}^*\|}{\|e_{\tilde{h}}\|} \simeq 0.947. \tag{5.29}$$

Third, the undervaluation introduced in the multiprojection strategy is assessed. Recall that the multiprojection approximates a vector in $V^{\mathrm{I}} + V^{\mathrm{P}}$, say $e_{\tilde{h}}^*$, by the sum of the two locally computed projections, $e_{_L} := \varepsilon + \eta$.

In the considered case the dimension of V^{I} is $n_i + n_c = 765$, the dimension of V^{P} is $n_p + n_c = 802$ and the dimension of the intersection, $V^{\mathrm{I}} \cup V^{\mathrm{P}}$, is $n_c = 340$. Thus, the A matrix is a 425×462 matrix. The study of the Singular Value Decomposition of this matrix and the corresponding application of Eq. (4.27) to obtain an approximate value for the average underestimation is cumbersome and computationally expensive. However, both V^{I} and V^{P} are generated as a sum of vectorial spaces associated with (almost) identical meshes. Denoting by M the number of elements in the original computational mesh and by M' the number of patches (which coincides with the number of free nodes)

$$V^{\mathrm{I}} = \bigoplus_{k=1}^{M} V_k^I \text{ and } V^{\mathrm{P}} = \bigoplus_{l=1}^{M'} V_l^P, \tag{5.30}$$

where V_k^I is the interpolation space associated with the submesh discretizing the k-th element and V_l^P is the interpolation space associated with the submesh discretizing the l-th patch. The set of element spaces $\{V_k^I\}_{k=1,\ldots,M}$ is orthogonal as well as the set of the set patch spaces $\{V_l^P\}_{l=1,\ldots,M'}$. Moreover, if the k-th element and the l-th patch are disjoint,V_k^I and V_l^P are also orthogonal. On the other hand, the value of ψ depends only on the "relative position" of V^{I} and V^{P}. The singular values of A are interpreted as the cosinus of a set of angles describing precisely this relative position. Due to the decomposition described in Eq. (5.30), the singular values different to zero, that is, the angles different to the right angle, must correspond to local spaces V_k^I and V_l^P

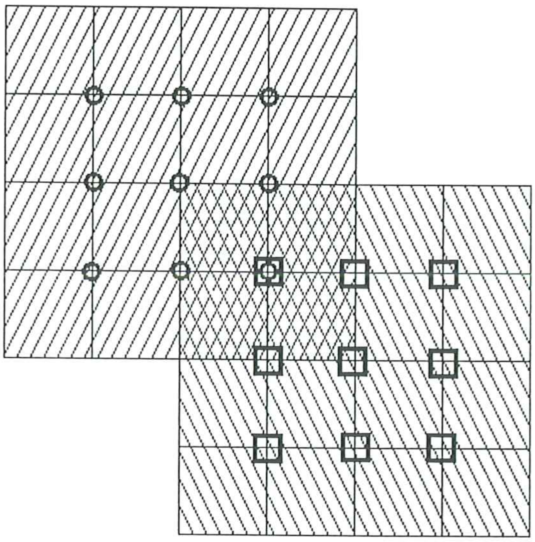


Fig. 6. Sample meshes generating V_1 (nodes marked by circles) and V_2 (nodes marked by squares). Measuring the "relative position" between V_1 and V_2 suffices to characterize the relative position between V^{I} and V^{P} and, hence, the average behavior of the multiprojection strategy.

associated with an element and a patch which are non-disjoint. These angles are essentially a function of the topology of the intersection between the element and the patch and therefore will be the same for every couple of local spaces V_k^I and V_l^P. In the following it is assumed that the singular values of A are represented by the singular values arising in a much simpler problem related with the sample spaces V_1 and V_2 generated by the mesh shown in Fig. 6. Note that if the singular values of A are simply a repetition of the singular values of the sample (and simple) problem, the result obtained applying Eq. (4.27) is the same for the original problem and the simplified one.

Thus, in the following, the average undervaluation introduced in the multiprojection strategy is analyzed replacing $V^{\mathrm{I}}$and V^{P} by V_1 and V_2, respectively. The dimension of V_1 is $n_i + n_c = 8 + 1 = 9$, the dimension of V_2 is $n_p + n_c = 8 + 1 = 9$. Consequently, in this case, A is a 8×8 matrix. The analysis of the matrix shows that the range is $r = 3$ and the corresponding singular values are

$$\sigma_1 = 0.23932, \qquad \sigma_2 = 0.11452 \quad \text{and} \quad \sigma_3 = 0.01276.$$

Recall that small singular values of A are associated with nearly perpendicular spaces and a multiprojection strategy with a good behavior.

The approximate expression of Eq. (4.27) yields

$$\psi \simeq 0.978435.$$

In order to verify that the approximation introduced in Eq. (4.27) is acceptable, at least in this simple case, 100,000 vectors are randomly generated on the unit sphere of the 17 dimensions space $V_1 + V_2$. The effectivity index of the projection strategy, ν, is computed for each of these vectors and the mean value of ν, ϕ, and ν^2, ψ, are evaluated. The obtained values are

$$\phi = 0.990121 \text{ and } \psi = 0.980736.$$

Consequently, the variance of ν is $0.019924 \simeq 0.02$. It is worth noting that the predicted value of ψ (0.978435) and the computed value (0.980736) are equal up to the second significant digit, that is, the error introduced in the approximation of Eq. (4.27) is much less than the variance of ν.

Thus, the predicted undervaluation in the transformation from $e^*_{\tilde{h}}$ to $e_{_L}$ is taken as the square root of the predicted value for ψ , that is,

$$\frac{\|e_{_L}\|}{\|e^*_{\tilde{h}}\|} \simeq 0.989. \tag{5.31}$$

Resuming the results of Eqs. (5.28), (5.29) and (5.31), an evaluation of the expected value of the effectivity index is found:

$$\frac{\|e_{_L}\|}{\|e\|} = \frac{\|e_{\tilde{h}}\|}{\|e\|}\frac{\|e^*_{\tilde{h}}\|}{\|e_{\tilde{h}}\|}\frac{\|e_{_L}\|}{\|e^*_{\tilde{h}}\|} \simeq 0.968 \times 0.947 \times 0.989 = 0.907. \tag{5.32}$$

Recall that, in this problem the obtained effectivity index is 0.919. If the prediction on the behavior of the effectivity index is used to improve the error estimate, the corrected estimate would be $e_{_L}/0.907$. The new value for the effectivity index is $0.919/0.907 = 1.013$. That results on improving the quality of the error estimate. The original estimate has an error of 8.1% in the evaluation of the error. The "error in the error" for the corrected estimate is 1.3%.

6. Concluding remarks

This paper introduces a new approach to the analysis of a posteriori error estimators for finite elements. The main goal of this analysis is to a priori characterize the average behavior of the error estimator in front of a random error. This philosophy is applied to the analysis of a residual type error estimator based on local projections of the error.

The formulation and the geometrical interpretation of the estimator make this analysis specially simple in this case. The main undervaluation introduced in the consecutive projections is, in fact, a function of some geometrical properties of the Finite Element spaces used to approximate the error.

The numerical example shows that the predicted expected value of the effectivity index approximates well the obtained effectivity index in a problem where the exact solution is known. This technique may also be used to correct the obtained estimate and to reduce the error on the approximation of the error.

References

1. I. Babuška, R. Durán and R. Rodríguez , *Analysis of the efficiency of an a posteriori error estimator for linear triangular finite elements, SIAM Journal of Numerical Analysis*, **29** (1992) 947–964.
2. P. Díez, *Un nuevo estimador de error para el método de los elementos finitos*, Doctoral Thesis, Universitat Politècnica de Catalunya, Barcelona (1996).
3. P. Díez, J.J. Egozcue and A. Huerta, *A posteriori error estimation for standard finite element analysis, Computer Methods in Applied Mechanics and Engineering*, **163** (1998) 141–157.
4. P. Díez, J.J. Egozcue and A. Huerta, *Analysis of the average efficiency of an error estimator*, in *Finite Element Methods: Superconvergence, Post-processing and a Posteriori Error Estimates,* eds. M. Křížek et al. (Marcel Dekker, 1997), pp. 113–126.
5. P. Díez and A. Huerta, *A unified approach to remeshing strategies for finite element h-adaptivity, Computer Methods in Applied Mechanics and Engineering*, **176** (1999) 215–229.
6. P. Díez, M. Arroyo and A. Huerta, *Adaptivity based on error estimation for viscoplastic softening materials, Mechanics of Cohesive-Frictional Materials* (in press).
7. R. Durán, R. & R. Rodríguez, *On the asymptotic exactness of Bank-Weiser's estimator, Numerische Mathematik*, **62** (1992) 297–303.
8. A. Huerta and P. Díez, *Error estimation including pollution assessment for nonlinear finite element analysis, Computer Methods in Applied Mechanics and Engineering* (in press).
9. D.W. Lewis, *Matrix Theory*, World Scientific (1991)
10. O.C. Zienkiewicz, O.C. and J.Z. Zhu, *A simple error estimator and adaptive procedure for practical engineering analysis, International Journal for Numerical Methods in Engineering*, **24** (1987) 337-357.