ELSEVIER

Comput. Methods Appl. Mech. Engrg. 190 (2001) 2483–2498

Computer methods in applied mechanics and engineering

www.elsevier.com/locate/cma

Averaging technique for FE – a posteriori error control in elasticity. Part I: Conforming FEM

Carsten Carstensen *, Stefan A. Funken

University of Kiel, Mathematisches Seminar II, Ludewig-Meyn-Str.4, 24098 Kiel, Germany

Received 20 October 1999; received in revised form 20 October 1999

Abstract

Averaging techniques are popular tools in adaptive finite element methods for the numerical treatment of second-order partial differential equations since they provide efficient a posteriori error estimates by a simple postprocessing. In this paper, the reliability of *any* averaging estimator is shown for low order finite element methods in elasticity. Theoretical and numerical evidence supports that the reliability is up to the smoothness of given right-hand sides and independent of the structure of a shape-regular mesh.

MSC: 65N30; 65R20; 73C50

Keywords: Elasticity; A posteriori error estimates; Adaptive algorithm; Reliability; Finite element method

1. Introduction

Error control and efficient mesh-design in finite element simulations of computational engineering and scientific computing is frequently based on a posteriori error estimates [2,3,14,18]. One of the more popular techniques is local or global averaging, e.g., in form of the ZZ-error estimator [20]. Efficiency and reliability of this estimator were known only for very structured grids and for solutions of higher regularity and then we have even asymptotic exactness [18]. Numerical experiments in [3] showed that averaging techniques were more reliable on irregular meshes than expected. For homogeneous Dirichlet conditions for a Laplace equation, the reliability and efficiency of the ZZ-estimator is proven on unstructured, merely shape-regular grids in [17].

In this work, we give theoretical and numerical support for the reliability of *all* averaging techniques in elasticity, robust with respect to violated (local) symmetry of meshes or super-convergence and robust with respect to *other boundary conditions* and discuss the robustness with respect to *incompressibility locking*.

For a more precise description of averaging techniques in elasticity, let us discuss a discretisation of the equilibrium condition.

$$f + \operatorname{div} \sigma = 0 \tag{1.1}$$

for a stress field $\sigma \in L^2(\Omega)^{d\times d}_{\mathrm{sym}}$, with a given right-hand side $f \in L^2(\Omega)^d$ on an elastic body $\Omega \subset \mathbb{R}^d$. Here, $\mathbb{R}^{d\times d}_{\mathrm{sym}}$ denotes the real symmetric $d \times d$-matrices and $L^2(\Omega)^{d\times d}_{\mathrm{sym}} := L^2(\Omega; \mathbb{R}^{d\times d}_{\mathrm{sym}})$ denotes the Hilbert space of all Lebesgue integrable functions on the bounded Lipschitz domain Ω with values in $\mathbb{R}^{d\times d}_{\mathrm{sym}}$.

* Corresponding author. Fax: 49-431-880-4091.
E-mail address: ttcc@numerik.uni-kiel.de (C. Carstensen).

0045-7825/01/$ - see front matter
PII: S0045-7825(00)00248-6

Suppose a known approximation $\sigma_h \in L^2(\Omega)^{d\times d}_{\rm sym}$ satisfies a discrete principle of virtual work with respect to a test function finite element space that includes continuous piecewise linears $\mathscr{S}^1_{\rm D}(\mathscr{T})^d$ (with homogeneous Dirichlet boundary conditions) based on a regular triangulation $\mathscr{T}$ of Ω, i.e., suppose

$$\int_\Omega \sigma_h : \varepsilon(v_h)\,{\rm dx} = \int_\Omega f\cdot v_h\,{\rm dx} \quad \text{for all } v_h \in \mathscr{S}^1_{\rm D}(\mathscr{T})^d. \tag{1.2}$$

Here, the (linear) Green strain $\varepsilon(u) = (Du + (Du)^{\rm T})/2$ is the symmetric part of the gradient Du of the displacement field u and colon denotes the scalar product of $d\times d$-matrices, i.e., $A : B = \sum_{j,k=1} A_{jk}B_{jk}$ with norm $|A| = \sqrt{A : A}$ for $A, B \in \mathbb{R}^{d\times d}$.

In averaging techniques, the error $\|\sigma - \sigma_h\|_{L^2(\Omega)}$ is estimated by the approximation error of a smoother approximation σ^*_h to σ_h. For instance, in linear elasticity with constant Lamé constants and with piecewise linear finite elements, σ_h is piecewise constant and approximated by continuous piecewise linear functions, written $\mathscr{S}^1(\mathscr{T})$. Then

$$\eta z := \min_{\sigma^*_h\in \mathscr{S}^1(\mathscr{T})^{d\times d}} \|\sigma_h - \sigma^*_h\|_{L^2(\Omega)} \tag{1.3}$$

may serve as a computable error estimator and the elementwise contributions as local error indicators in an adaptive mesh-refining algorithm.

This paper aims to investigate an upper and lower estimate of the form

$$c_1\eta z + \text{h.o.t.} \leqslant \|\sigma - \sigma_h\|L^2(\Omega) \leqslant c_2\eta z + \text{h.o.t.} \tag{1.4}$$

and proves that $c_1, c_2 > 0$ do *not* depend on local symmetry relations of shape-regular meshes.

We study a nonlinear model problem in elasticity and give precise conditions on the Dirichlet and Neumann boundary data, establish the necessary modifications for Neumann data, and analyse their influence on the higher order terms h.o.t. in (1.4). Furthermore, we mention equivalence of local and global averaging from [7]. Changing to energy norms in linear elasticity we show that the constants $c_1, c_2 > 0$, do *not* depend on the Poisson ratio $\nu \to 1/2$ and in this sense, the estimate (1.4) is locking-free.

Note that in practice, we may apply an averaging operator $\mathscr{A} : L^2(\Omega)^{d\times d}_{\rm sym} \to \mathscr{S}^1(\mathscr{T})^{d\times d}_{\rm sym}$ to σ_h and compute the upper bound $\|\sigma_h - \mathscr{A}\sigma_h\|_{L^2(\Omega)}$ of η_Z. Then, efficiency depends strongly on the approximation properties of $\mathscr{A}$ and deserved further investigations, but their *reliability* follows from our results and is independent of (lacking) higher-regularity of the solution.

The outline of the paper is as follows. The model problem and precise descriptions of material properties and the regularity of right-hand sides are given in Section 2 together with modified versions of (1.3) and (1.4) that include the proper treatment of boundary conditions. Three different adaptive algorithms are suggested in Section 3. Numerical evidence is provided in Section 4 for almost asymptotic exactness of our realisation of the ZZ-estimator for adapted meshes (when we start with a structured grid). For more unstructured grids the reliability and efficiency is still observed with very good constants. The proofs are given in Section 5 following arguments in [7]. The proof is essentially based on an approximation operator from [7,6,11] and (almost) avoids the use of any trace-inequality. Section 6 is devoted to the locking phenomena in linear elasticity when Young's modulus is nearly 1/2. Experimental evidence for a robust error estimation is supported by a theoretical heuristic argument. A further analysis on the locking phenomena will follow in Part II of this paper [10].

2. Model example and results

The stress field σ satisfies the equilibrium equations:

$$f + \operatorname{div}\sigma = 0 \quad \text{in } \Omega, \tag{2.1}$$

$$\sigma \cdot n = g \quad \text{on } \Gamma_N, \tag{2.2}$$

for given volume force $f \in L^2(\Omega)^d$ and applied surface load $g \in L^2(\Gamma_N)^d$. The Lipschitz boundary $\Gamma = \partial\Omega$ of the body, occupied by a bounded domain Ω in $\mathbb{R}^d$, consists of a closed Dirichlet part Γ_D with positive surface measure and a remaining, relatively open and possibly empty, Neumann part $\Gamma_N := \Gamma \setminus \Gamma_D$. The Dirichlet data $u_D \in C(\Gamma_D)$ is supposed to be differentiable at any flat piece of Γ_D such that the surface gradient is square-integrable (written $u_D \in H^1(\Gamma_D)$). Then, we suppose that the exact displacement field u belongs to $H^1(\Omega)^d$, i.e., $u \in L^2(\Omega)^d$ and the gradient Du satisfies $Du \in L^2(\Omega)^{d\times d}$, and satisfies

$$u = u_D \quad \text{on } \Gamma_D. \tag{2.3}$$

The elastic constitutive relation between the stress field σ and the (linear) Green strain $\varepsilon(u) = (\frac{1}{2}(\partial u_j/\partial x_k + \partial u_k/\partial x_j))_{j,k=1}^d$ is described by a (possibly nonlinear) function $\hat{\sigma}(x,\varepsilon)$ that depends on the material point $x \in \Omega$ and on the Green strain ε

$$\sigma(x) = \hat{\sigma}(x, \varepsilon(u)(x)) \quad \text{for a.a. } x \in \Omega. \tag{2.4}$$

The function $\hat{\sigma}(x,\varepsilon)$ is supposed to be piecewise smooth in x and Lipschitz continuous and uniformly monotone in ε, i.e., we suppose that there exist constants $c_3, c_4 > 0$ such that for almost all $x \in \Omega$ and for all $A, B \in \mathbb{R}^{d\times d}_{\text{sym}}$ there holds

$$c_3|B-A|^2 \leqslant (\sigma(x,B) - \sigma(x,A)) : (B-A), \tag{2.5}$$

$$|\sigma(x,B) - \sigma(x,A)| \leqslant c_4|B-A|. \tag{2.6}$$

A typical example for such a nonlinear material behaviour is named after Hencky [19], where

$$\hat{\sigma}(x,\varepsilon) = (\lambda+\mu)\,\operatorname{tr}(\varepsilon) I_{d\times d} + \mu D^2\varphi(|\operatorname{dev}\varepsilon|^2) \tag{2.7}$$

and $D^2\varphi$ is the Hessian matrix of all partial derivatives of second-order of a smooth monotone increasing function $\varphi : [0,\infty) \to \mathbb{R}$. Under further conditions and φ, it is shown in [19, Section 62] that $\hat{\sigma}(x,\cdot)$ is uniformly monotone and Lipschitz continuous in the sense of (2.5) and (2.6). From the general theory of monotone operators we have existence of a unique (weak) solution u to (2.1)–(2.5) in the Sobolev space $H^1(\Omega)$.

To approximate the (unknown) exact solution u with a finite element method, we consider a triangulation $\mathcal{T}$ of the bounded Lipschitz domain $\Omega \subseteq \mathbb{R}^d, d = 2, 3$, [5,12] (the domain is matched exactly) with piecewise affine Lipschitz boundary $\Gamma = \partial\Omega = \Gamma_D \cup \Gamma_N$, i.e., $\mathcal{T}$ consists of a finite number of closed subsets of $\bar{\Omega}$, that cover $\bar{\Omega} = \cup\mathcal{T}$. Each element $T \in \mathcal{T}$ is either a triangle $T = \operatorname{conv}\{a,b,c\}$ or a parallelogram $T = \operatorname{conv}\{a,b,c,d\}$ if $d = 2$ or a tetrahedron $T = \operatorname{conv}\{a,b,c,d\}$ or a parallelepiped $T = \operatorname{conv}\{a,b,c,d,e,f,g,h\}$ if $d = 3$. The extremal points $a,\ldots,h$ are called vertices, the faces $E \subseteq \partial T$ such as $E = \operatorname{conv}\{a,b\}$ if $d = 2$, or $E = \operatorname{conv}\{a,b,c\}$, respectively a parallelogram $\operatorname{conv}\{a,b,c,d\}$ if $d = 3$ are called edges. The set of all vertices and all edges appearing for some T in $\mathcal{T}$ are denoted as $\mathcal{N}$ and $\mathcal{E}$.

For $T \in \mathcal{T}$, let $P_T := \mathcal{P}_1(T)$ if $d = 1$, or if T is a triangle and $d = 2$, or if T is a tetrahedron and $d = 3$; let $P_T : \mathcal{Q}_1(T)$ if T is a parallelogram and $d = 2$, or a parallelepiped and $d = 3$. Here, $\mathcal{P}_k(K)$, respectively, $\mathcal{Q}_k(K)$ denotes the set of algebraic polynomials in d variables on K of total respectively, partial degree $\leqslant k$. The (nonconforming) discrete space $\mathcal{L}^1(\mathcal{T})$ is the set of all $U \in L^\infty(\Omega)$ with restrictions in P_T, i.e., $U|_T \in P_T$ for all T in $\mathcal{T}$ while $\mathcal{L}^0(\mathcal{T})$ denotes the set of $\mathcal{T}$-piecewise constant functions. Then, let

$$\mathcal{S}^1(\mathcal{T}) := \mathcal{L}^1(\mathcal{T}) \cap C(\Omega) \quad \text{and} \quad \mathcal{S}^1_D(\mathcal{T}) := \{v_h \in \mathcal{S}^1(\mathcal{T}) : v_h|\Gamma_D = 0\}. \tag{2.8}$$

We allow hanging nodes and suppose that there is set $\mathcal{N}$ of regular nodes (i.e., the vertices defined above which are not hanging nodes), such there exist a nodal basis $(\varphi_z : z \in \mathcal{N})$ of the space $\mathcal{S}^1(\mathcal{T})$, respectively $(\varphi_z : z \in \mathcal{K})$ of the space $\mathcal{S}^1_D(\mathcal{T})$ defined by the orthogonality-property that $\varphi_z(z) = 1$ and $\varphi_z(x) = 0$ for all $x, z \in \mathcal{N}$ with $x \neq z$. Continuity of the basis functions for all φ_z is understood such that for $d = 3$ and any face $E \in \mathcal{E}$, $\varphi_z|_E$ is either on $\mathcal{Q}^1(E)$ or $\mathcal{P}^1(E)$ for all φ_z with $\varphi_z|_E \neq 0$. That means, a parallelpiped cannot share a face of positive surface measure with a tetrahedron. Finally, $\mathcal{K} := \mathcal{N} \setminus \Gamma_D$ is the set of free nodes while $\mathcal{N} \cap \Gamma_D$ are the regular nodes on the Dirichlet boundary.

The number of hanging nodes is limited by the condition that

$$\max_{z\in\mathscr{N}} \mathrm{Lip}(\varphi_z) \leqslant c_5/h_z, \quad \text{where } h_z := \mathrm{diam}(\omega_z) \tag{2.9}$$

and $\mathrm{Lip}(\phi)$ denotes the Lipschitz constant of ϕ, $\mathrm{diam}(\omega_z)$ is the diameter of the patch

$$\omega_z := \{x \in \Omega : 0 < \varphi_z(x)\}. \tag{2.10}$$

Let $u_h \in \mathscr{S}^1(\mathscr{T})^d$ denote the Galerkin approximate to u which is defined by the Dirichlet boundary conditions

$$u_h(z) = u_{\mathrm{D}}(z) \quad \text{for } z \in \mathscr{N} \cap \Gamma_{\mathrm{D}} \tag{2.11}$$

and the Galerkin-property, for all $v_h \in \mathscr{S}^1_{\mathrm{D}}\mathscr{T}^d$,

$$\int_\Omega \hat{\sigma}(x, \varepsilon(u_h)(x)) : \varepsilon(v_h)\,\mathrm{d}x = \int_\Omega f.v_h\,\mathrm{d}x + \int_{\Gamma_{\mathrm{N}}} g.v_h\,\mathrm{d}s. \tag{2.12}$$

The existence of u_h is assumed and guaranteed, e.g., for Hencky's material (2.7). Then, we define

$$\sigma_h := \hat{\sigma}(x, \varepsilon(u_h)(x)) \tag{2.13}$$

and its piecewise constant approximation ($|T|$ denotes the area or volume of T)

$$\bar{\sigma}_h \in \mathscr{L}^0(\mathscr{T})^{d\times d}_{\mathrm{sym}} \quad \text{with} \quad \bar{\sigma}_h|_T := \int_T \sigma_h\,\mathrm{d}x/|T| \ \text{ for } T \in \mathscr{T}. \tag{2.14}$$

The proper averaging functions of $\mathscr{S}^1(\mathscr{T})^{d\times d}$ have to approximate the static boundary conditions (2.2). To make this possible, it is required that the averaging function $\sigma^* \in \mathscr{S}^1(\mathscr{T})^{d\times d}$ may be nonsymmetric and that g satisfies some compatibility conditions. Let $\mathscr{E}_{\mathrm{N}} := \{E \in \varepsilon : E \subset \overline{\Gamma_{\mathrm{N}}}\}$ denote the edges or faces at the Neumann boundary. For each $E \in \mathscr{E}_{\mathrm{N}}$, let n_E denote the (constant) outer unit normal along the flat surface piece E. To enable a nodal interpolation

$$\sigma^*(z)\cdot n_E = g(z) \quad \text{for all } z \in \mathscr{N} \cap E \ \text{ with } E \in \mathscr{E}_{\mathrm{N}} \tag{2.15}$$

we require some continuity on g. At those nodes z on Γ_{N}, where Γ_{N} is flat and so the normal vectors coincide, $n_{E_1} = n_{E_2}$, for two distinct neighbouring $E_1, E_2 \in \mathscr{E}_{\mathrm{N}}$, the continuity of σ^* at $z \in E_1 \cap E_2 \cap \mathscr{N}$ implies that the restrictions $g|_{E_1}$ and $g|_{E_2}$ coincide at z. Then, the set

$$\mathscr{Q}(\mathscr{T}, g) := \{\sigma^* \in \mathscr{S}^1(\mathscr{T})^{d\times d} : \sigma^* \text{ satisfies (2.15)}\} \tag{2.16}$$

is nonvoid. Note that $\mathscr{Q}(\mathscr{T}, g) = \mathscr{S}^1(\mathscr{T})^{(d\times d)}$ in the case of pure Dirichlet conditions.

The global averaging estimator is defined by

$$\eta_Z := \min_{\sigma^*\in\mathscr{Q}(\mathscr{T},g)} \|\sigma_h - \sigma^*\|_{L^2(\Omega)} \tag{2.17}$$

The subsequent result is a precise statement of (1.4) with $c_1 = 1$ under mixed boundary conditions. Let $h_{\mathscr{T}} \in \mathscr{L}^0(\mathscr{T})$ denote the local mesh-size, i.e., $h_{\mathscr{T}}|_T := h_T := \mathrm{diam}\,(T)$ for $T \in \mathscr{T}$, and let $h_{\mathscr{E}} \in \mathscr{L}^0(\bigcup \mathscr{E})$ define the local edge-size on the skeleton $\bigcup \mathscr{E}$ of all edges by $h_{\mathscr{E}}|_E := h_E := \mathrm{diam}(E)$ for $E \in \mathscr{E}$.

Theorem 2.1. *Let $u \in H^1(\Omega)^d$ solve (2.1)–(2.4) and let $u_h \in \mathscr{S}^1(\mathscr{T})^d$ solve (2.11) and (2.12). Suppose $f \in H^1(\Omega)^d$, $g \in H^1(\Gamma_{\mathrm{N}})$ and $u_{\mathrm{D}} \in H^2(\Gamma_{\mathrm{N}})$. Then,*

$$\eta_Z - \min_{\sigma^*\in\mathscr{Q}(\mathscr{T},g)} \|\sigma - \sigma^*_h\|_{L^2(\Omega)} \leqslant \sigma - \sigma_h\|_{L^2(\Omega)} \leqslant c_2\Big(\eta_Z + \|\sigma - \bar{\sigma}_h\|_{L^2(\Omega)} + \|h_{\mathscr{E}}^{3/2}\,\partial^2 u_{\mathrm{D}}/\partial s^2\|_{L^2(\Gamma_{\mathrm{D}})} + \|h_{\mathscr{E}}^{3/2}\,\partial g/\partial s\|_{L^2(\Gamma_{\mathrm{N}})} + \|h_{\mathscr{T}}^2\nabla f\|_{L^2(\Omega)}\Big). \tag{2.18}$$

The meshsize-independent constant $c_2 > 0$ depends on c_3, c_5 and the shape of the elements and patches only.

Remarks 2.1. (i) Since $\|u-u_h\|_{H^1(\Omega)} = \mathrm{O}(h)$ for the maximal mesh-size $h \to 0$, the remaining terms such as $||h_{\mathscr{T}}^2 \nabla f||h_{\mathscr{E}}^{3/2}\partial^2 u_{\mathrm{D}}/\partial s^2\|_{L^2(\Gamma_{\mathrm{D}})}$ or $\|h_{\mathscr{E}}^{3/2}\partial g/\partial s\|_{L^2(\Gamma_{\mathrm{N}})}$ are of higher order and so are $\min_{\sigma^* \in \mathscr{Q}(\mathscr{T},g)} \|\sigma-\sigma^*\|_{L^2(\Omega)}$ and $\|\sigma-\bar{\sigma}_h\|_{L^2(\Omega)}$. Therefore, (2.18) proves (1.4) with $c_1 = 1$. It seems interesting that h.o.t. on the right-hand side of the estimate (2.18) depends on the smoothness of the given loads only and hence it can be checked if this is true. Conversely, the h.o.t. on the left-hand side of the estimate (2.18) (unknown) smoothness of the exact solution σ.

(ii) Let us emphasise from the proof in Section 4 that the derivatives along Γ are required only $\mathscr{E}$-piecewisely while f needs to be patch-wise (not only elementwise) in H^1 and so $f \in H^1(\Omega)$. For nonsmooth right-hand side f, $\|h_{\mathscr{T}}^2 \nabla f\|_{L^2(\Omega)}$ may be replaced by a patch-wise L^2-best-approximation error in the approximation through constants of f.

(iii) It is shown in [6,11] that the edge-contributions (jump differences in the normal fluxes components across edges) dominate in standard residual a posteriori error estimates [2,4,5,9,14,18]. Arguing as in [13,16,17] one can hence derive alternative proofs of (2.18).

(iv) In an L^∞-estimate of [15] it is suggested to average over a domain of size $\mathrm{O}(h\log(1/h))$ instead of merely patches or the entire domain to obtain asymptotic exact results.

3. Adaptive algorithms

The numerical examples provide experimental evidence of the efficiency, reliability and robustness of the a posteriori error estimate as the superiority of the following adaptive refinement Algorithms (A_Z^ϑ) and $(A_{\mathscr{A}}^\vartheta)$ over a uniform mesh-refining.

Instead of η_Z we calculate $\eta_{\mathscr{A}} := \|\sigma_h - \mathscr{A}\sigma_h\|_{L^2(\Omega)}$ with the averaging operator $\mathscr{A}$ based on a function $\sigma_h^* \in \mathscr{S}^1(\mathscr{T})^{2\times 2}$ which satisfies $g(z) = \sigma_h^*(z) n_E(z)$ for each endpoint z of an edge E on $\overline{\Gamma}_{\mathrm{N}}$. We define

$$\mathscr{A}\sigma_h := \sigma_h^* := \sum_{z\in\mathscr{N}} \mathscr{I}_z(\sigma_h)\varphi_z, \tag{3.1}$$

where $\mathscr{I} : L^2(\Omega)^{2\times 2} \to \mathscr{S}^1(\mathscr{T})^{2\times 2}$ for $z \in \mathscr{N}/\overline{\Gamma}_{\mathrm{N}}$ is, with $\fint_{\omega_z}\sigma_h\,\mathrm{d}x$ denoting the integral mean of σ_h over ω_z,

$$\mathscr{I}_z(\sigma_h) := \fint_{\omega_z} \sigma_h \,\mathrm{d}x.$$

For $z \in \overline{\Gamma}_{\mathrm{N}}$ we distinguish between the following cases (i) and (ii) to fulfil the discrete Neumann condition $g(z) = \sigma_h^*(z) n_E$ at z.

(3.2.i) In case $z \in E_1 \cap E_2$ for two distinct edges $E_1, E_2 \subset \Gamma_{\mathrm{N}}$ with linearly independent outer unit normals n_{E_1} and n_{E_2} on E_1, E_2, respectively, we choose $\mathscr{I}_z(\sigma_h)$ to be the unique solution $\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22}\end{pmatrix}$ of the linear system

$$\begin{pmatrix} n_{1,E_1} & n_{2,E_1} & 0 & 0 \\ 0 & 0 & n_{1,E_1} & n_{2,E_1} \\ n_{1,E_2} & n_{2,E_2} & 0 & 0 \\ 0 & 0 & n_{1,E_2} & n_{2,E_2} \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} g_{1|E_1}(z) \\ g_{2|E_1}(z) \\ g_{1|E_2}(z) \\ g_{2|E_2}(z) \end{pmatrix}.$$

(3.2.ii) In the remaining cases $z \in E_1 \cap \Gamma_{\mathrm{D}}$ or $z \in E_1 \cap E_2$ with two parallel outer unit normals n_{E_1}, n_{E_2} we choose t_{E_1} to be the unit tangent to Ω at z that is perpendicular to n_{E_1} and let $\mathscr{I}_z(\sigma_h)$ be the solution $\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22}\end{pmatrix}$ of the uniquely solvable system

$$\begin{pmatrix} n_{1,E_1} & n_{2,E_1} & 0 & 0 \\ 0 & 0 & n_{1,E_1} & n_{2,E_1} \\ t_{1,E_2} & t_{2,E_2} & 0 & 0 \\ 0 & 0 & t_{1,E_2} & t_{2,E_2} \end{pmatrix} \begin{pmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{pmatrix} = \begin{pmatrix} g_{1|E_1}(z) \\ g_{2|E_1}(z) \\ \left(\fint_{\omega_z}(\sigma_{h,11},\sigma_{h,12})\,\mathrm{d}x\right) t_{E_1} \\ \left(\fint_{\omega_z}(\sigma_{h,21},\sigma_{h,22})\,\mathrm{d}x\right) t_{E_1} \end{pmatrix}.$$

This amounts in the error indicator $\eta_{\mathscr{A},\mathscr{T}}$, for each $T \in \mathscr{T}$,

$$\eta_{\mathscr{A},\mathscr{T}} := \|\sigma_h - \mathscr{A}\sigma_h\|_{L^2(T)}. \tag{3.3}$$

For each $T \in \mathscr{T}$ we define the error indicator $\eta_{Z,T}$ by

$$\eta_{Z,T} := \|\sigma_h - \Pi\sigma_h\|_{L^2(T)}, \tag{3.4}$$

where Π denotes the L^2-projection onto $\mathscr{Q}(\mathscr{T}, g)$.

We will illustrate the efficiency of $\eta_{\mathscr{A}}$ in comparison with η_Z and a residual based error estimate η_R, i.e.,

$$\eta_R^2 := \sum_{T\in\mathscr{T}} \eta_{R,T}^2, \quad \text{where } \eta_{R,T}^2 := \frac{1}{\mu}\|h_{\mathscr{T}} f\|_{L^2(T)}^2 + \mu\|h_{\mathscr{E}}^{1/2}[\sigma_h \cdot n]\|_{L^2(\partial T)}^2. \tag{3.5}$$

Recall that $[\sigma_h \cdot n]$ denotes the stress jumps on interior edges, define it as $\sigma_h \cdot n - g$ on Γ_{N}, and set formally $[\sigma_h \cdot n](x) := 0$ if $x \in \Gamma_{\mathrm{D}}$.

For the adaptive approaches we use the following Algorithm $(A_{\mathscr{A}}^{\vartheta}), (A_Z^{\vartheta})$, and (A_R^{ϑ}). They differ only in the used error indicator and generate meshes that are either unperturbed (relative to $\mathscr{T}_0$) for $\vartheta = 0$ and randomly perturbed for $\vartheta = 1$ in step (f).

Algorithm. $(A_{\mathscr{A}}^{\vartheta}), (A_Z^{\vartheta})$, respectively (A_R^{ϑ}):
(a) Start with a coarse mesh $\mathscr{T}_0, k = 0$.
(b) Solve the discrete problem with respect to the actual mesh $\mathscr{T}_k$ with N degrees of freedom and errror $e_{\mathrm{N}} := |||\sigma - \sigma_h||| := \|\mathbb{C}^{-1/2}\varepsilon(u - u_h)\|_{L^2(\Omega)}$.
(c) For Algorithm $(A_{\mathscr{A}}^{\vartheta})$ compute, for all $T \in \mathscr{T}_k$,

$$\eta_T = \eta_{\mathscr{A},T} := \|\sigma_h - \mathscr{A}\sigma_h\|_{L^2(T)}.$$

For Algorithm (A_Z^{ϑ}) compute, for all $T \in \mathscr{T}_k$,

$$\eta_T = \eta_{Z,T} := \|\sigma_h - \Pi\sigma_h\|_{L^2(T)}.$$

For Algorithm (A_R^{ϑ}) compute, for all $T \in \mathscr{T}_k$ $\eta_T = \eta_{R,T}$ as in (3.5)
(d) Compute a given stopping criterion and decide to terminate or to go to (e).
(e) Mark the element T (*red* refinement) provided,

$$\frac{1}{2}\max_{T'\in\mathscr{T}_k} \eta_{T'} \leqslant \eta_T.$$

(f) Mark further elements (*red–green–blue*-refinement) to avoid hanging nodes. Generate a new triangulation $\tilde{\mathscr{T}}_{k+1}$ using edge-midpoints if $\vartheta = 0$ and points on the edges at a random distance at most $0.3h_E$ from the edge-midpoints if $\vartheta = 1$. Perturb the nodes $z \in \mathscr{N}_{k+1}$ of the mesh $\tilde{\mathscr{T}}_{k+1}$ at random with values taken uniformly from a ball around z of radius $\vartheta 2^{-k}/15$. Correct boundary nodes by orthogonal projection onto that boundary piece they are expected such that $\Omega, \Gamma_{\mathrm{D}}, \Gamma_{\mathrm{N}}$ are matched by the resulting mesh $\mathscr{T}_{k+1}$ exactly. Update k and go to (b).

The implementation is performed on triangles in Matlab as suggested in [1] using analytic formulae in the calculation of the stiffness matrix. Since $\mathbb{C}$ in (4.1) is a linear operator in our examples, the linear system of equations can be solved directly. In order to approximate the right-hand side for a given function $g \in L^2(\Gamma_{\mathrm{N}})^2$ we compute $\int_{\Gamma_{\mathrm{N}}} g v_h \, \mathrm{d}s$ via a three-point Gaussian quadrature rule on any edge E. The Dirichet boundary conditions are satisfied at the boundary nodes.

4. Numerical experiments

Three examples provide numerical evidence that all three a posteriori error indicators improve the convergence rate of the discretisation, that the error estimator η_Z is reliable and efficient on unstructured (rather degenerated) meshes, that the reliability is robust with respect to $\lambda \to \infty$, and that our adaptive

scheme $(A^{\vartheta}_{\mathscr{A}})$ is efficient in practise. Experimentally, the quotient of error over estimator increases with larger perturbations as a possible consequence of lacking local symmetry in the mesh or of degenerating triangles (see Fig. 4).

4.1. L-shaped domain with analytic solution

Let us consider the linear problem (2.1) and (2.2) with

$$\sigma(\epsilon) = \lambda \operatorname{tr}(\varepsilon)\mathbb{1}_{2\times 2} + 2\mu\varepsilon \tag{4.1}$$

and $\lambda = Ev/((1+v)(1-2v)), \mu = E/(2(1+v))$ for Young's modulus $E = 100\,000$ and the Poisson coefficient $0.3 \leqslant v < 0.5$.

The model example on the L-shaped domain shown in Fig. 1 models singularities at reentrant corners. Using polar coordinates $(r,\theta), -\pi < \theta \leqslant \pi$, centered at the origin, the exact solution u with radial component u_r reads

$$u_r(r,\theta) = \frac{r^{\alpha}}{2\mu}(-(\alpha+1)\cos((\alpha+1)\theta) + (C_2 - (\alpha+1))C_1\cos((\alpha-1)\theta)),$$
$$u_\theta(r,\theta) = \frac{r^{\alpha}}{2\mu}((\alpha+1)\sin((\alpha+1)\theta) + (C_2 + \alpha - 1))C_1\sin((\alpha-1)\theta)).$$

The parameters are $C_1 = -\cos((\alpha+1)\omega)/\cos((\alpha-1)\omega), C_2 = 2(\lambda+2\mu)/(\lambda+\mu)$, where $\alpha = 0.54448373\ldots$ is the positive solution of $\alpha\sin 2\omega + \sin 2\omega\alpha = 0$ for $\omega = 3\pi/4$. The exact solution is traction free, $g = 0$, on the Neumann boundary Γ_N. Starting from the initial mesh $\mathscr{T}_0$ from Fig. 1, we run Algorithm $(A^0_{\mathscr{A}})$. The resulting mesh after 21 adaptive refinements and a zoom at the re-entrant corner is shown in Fig. 2 and displays a rather high mesh-refinement near the singularity.

Error and error estimators $\eta_{\mathscr{A}}$ are displayed versus the number of degrees of freedom N for $v = 0.3, 0.49, 0.499$ for uniform meshes and adaptively refined meshes with Algorithm $(A^0_{\mathscr{A}})$ in Fig. 3. The slope -1/2 corresponds to an experimental convergence rate 1 owing to $N \propto h^{-2}$ in two dimensions. For the sequences of uniform meshes we obtain experimentally convergence ≈ 0.54 which coincides with the theoretically expected rate α. The adaptive mesh-refining Algorithm $(A^0_{\mathscr{A}})$ improves this experimental convergence order to 1 which is expected to be optimal for the used family of finite elements spaces. This numerical experiment confirms numerically that the a posteriori error estimate is h-independent.

The used FEM for the displacement formulation shows incompressibility locking phenomena in Fig. 3, i.e., the error in energy norm is not bounded (for a given number of unknowns) as $v \to 1/2$. (Further details of this model problem and nonlocking finite elements are given in [8] and in Part II of this paper [10].)

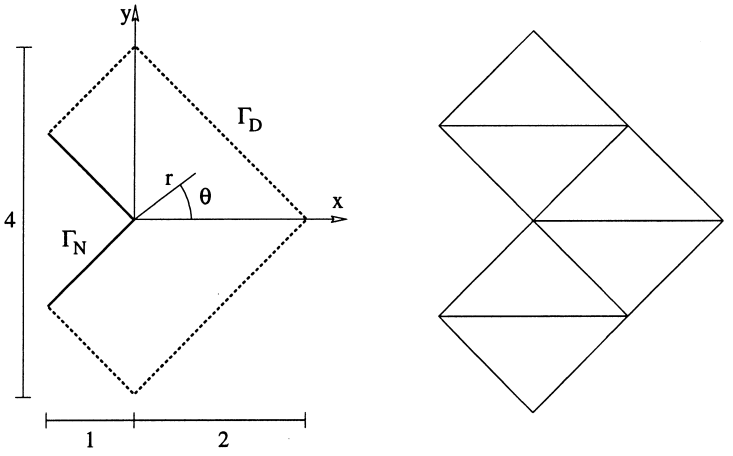

Fig. 1. System and initial mesh in Section 4.1.

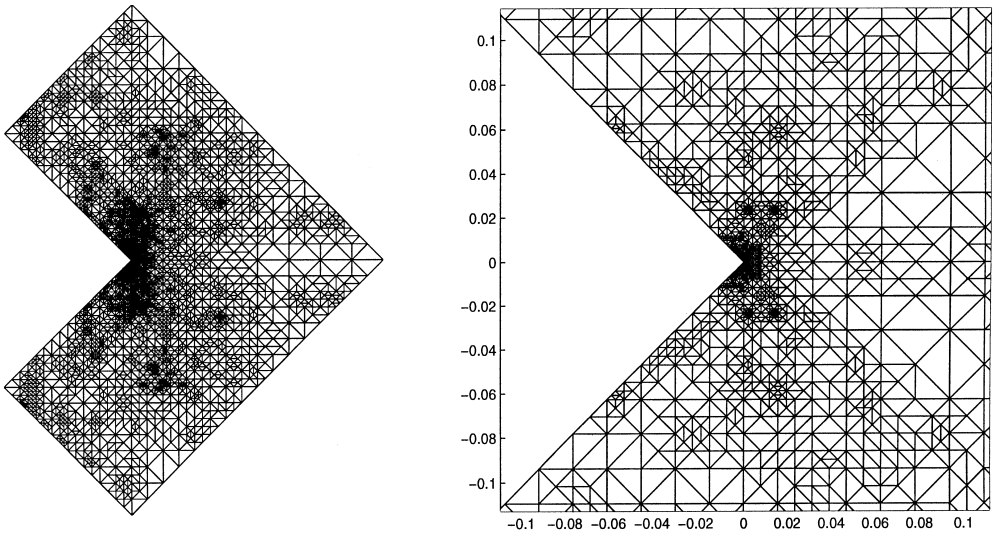

Fig. 2. Mesh $\mathscr{T}_{21}$ and magnified detail at the re-entrant corner for Section 4.1.

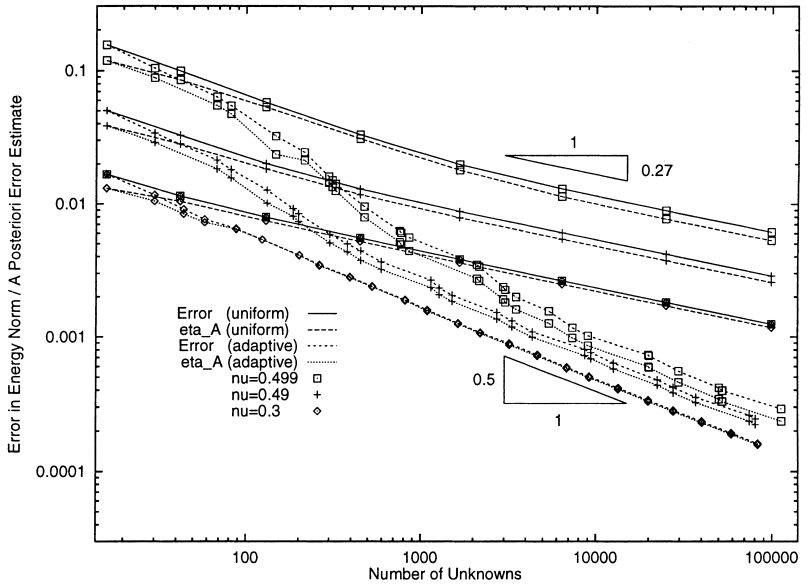

Fig. 3. Errors e_N and error indicator η_A vs N for uniform and adaptive meshes from Algorithm $(A^0_{\mathscr{A}})$ of Section 4.1.

Besides Algorithm $(A^0_{\mathscr{A}})$, we run Algorithm (A^0_Z) and obtained results (not displayed) similar to those in Fig. 3; the entries for the error indicators η_Z have the same slopes and the meshes resulting from Algorithm $(A^0_{\mathscr{A}})$ are of the same quality.

Super-convergence properties are frequently believed to be responsible for the good performance of averaging techniques for a posteriori error control in practice. To study the influence of local symmetries in the mesh, Algorithms $(A^1_{\mathscr{A}}), (A^1_Z)$, and (A^1_Z) perturb the nodes in step (f). Fig. 4 shows a sequence of perturbed refined meshes from Algorithm $(A^1_{\mathscr{A}})(v = 0.3)$ with optimal experimental convergence order 1.

For perturbed and nonperturbed meshes from Algorithms $(A^\vartheta_{\mathscr{A}})$ and (A^ϑ_Z) we display the extreme quotients of the error estimator $\eta_{\mathscr{A}}$ (respectively η_Z) over the error in energy norm e_N versus $1/2 - v$, i.e., the displayed constants are $\min\{\eta_{\mathscr{A}}/e_N\}$, $\max\{\eta_{\mathscr{A}}/e_N\}$, $\min\{\eta_Z/e_N\}$, and $\max\{\eta_Z/e_N\}$ for different values of N corresponding to $\mathscr{T}_1, \ldots, \mathscr{T}_k$ for k as implicitly shown in Fig. 3. The values for η_Z are always smaller than

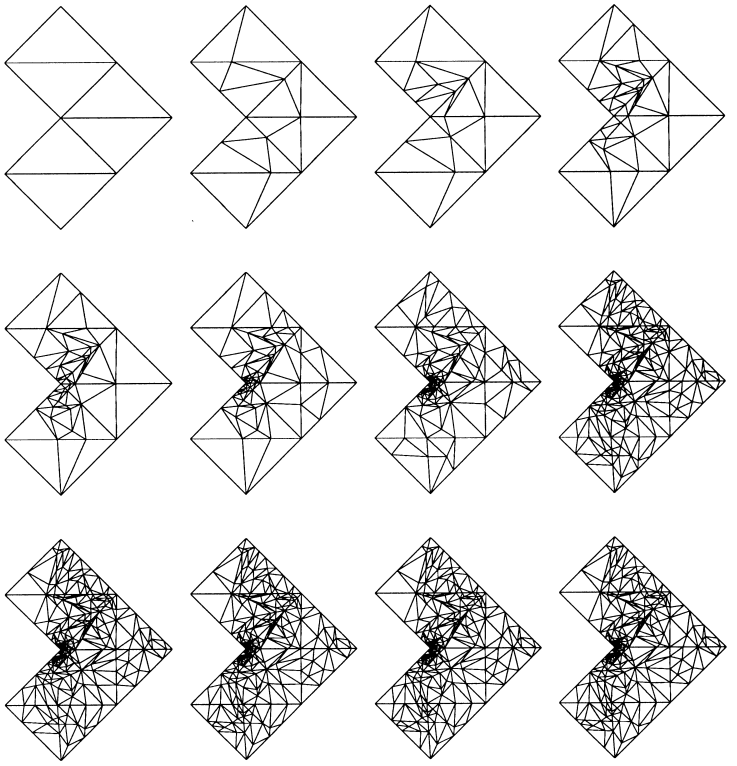

Fig. 4. $\mathscr{T}_0,\ldots,\mathscr{T}_{11}$ generated by Algorithm $(A^1_{\mathscr{A}})$ in Section 4.1.

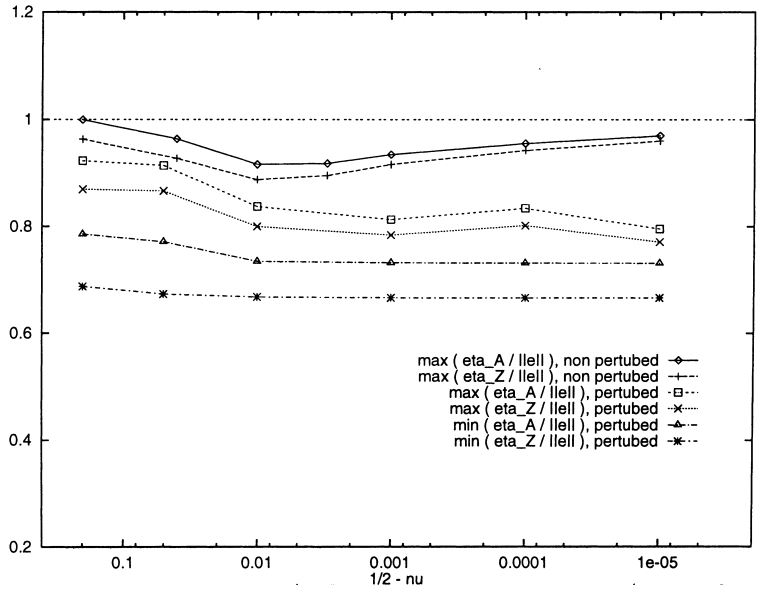

Fig. 5. Computed reliability/efficiency constants vs $1/2-\nu$ for perturbed and nonperturbed meshes from Algorithms $(A^\vartheta_{\mathscr{A}})$ and (A^ϑ_Z) of Section 4.1.

those that correspond to $\eta_{\mathscr{A}}$ as predicted by the theory. Fig. 5 shows also that the reliability constant is bounded from above and the efficiency constant from below independently from the Poisson ratio ν. This numerical example supports that also for perturbed meshes the estimate (1.4) is reliable and efficient.

4.2. Cook's membrane problem

As a further test example we consider a tapered panel clamped on the left end as depicted in Fig. 6 subject to a shearing load on the right end, i.e., $g = (0, 1000)$ on the right edge of Ω, $g = 0$ on the remaining part of Γ_N, $u = 0$, on Γ_D, and $f = 0$. The material constants are $E = 100\,000$ and $\nu = 1/3$ or 0.499 and the initial mesh $\mathscr{T}_0$ is displayed in Fig. 6.

A plot of the approximated von-Mises stress after 11 refinements generated by Algorithm $(A^0_{\mathscr{A}})$ as some magnified detail near the re-entrant corner (zoom of $(0, 10) \times (36, 46)$) is given in Fig. 7. (For visualisation the grey scalebar shows the stress only up to a fixed limit despite the maximal stress is unbounded.) The calculated von-Mises stress is in agreement with corresponding pictures in the literature.

The errors e_N for $\nu = 1/3$ (left) and $\nu = 0.499$ (right) computed with uniform and adaptive refinements are given in Fig. 8. the error e_N is computed using Galerkin orthogonality, via $e_N^2 = |||\sigma|||^2 - |||\sigma_h|||^2$. For given ν, extrapolated energies $G_\nu := |||\sigma|||$ are calculated from a sequences of approximations such as $G_{0.3} = 0.235093$ and $G_{0.499} = 0.218128$.

The adaptive mesh-refining Algorithms (A^0_Z) and (A^0_R) yield convergence order 1 which is asymptotically better than uniform refinement as observed in Fig. 8. The posteriori error estimates $\eta_{\mathscr{A}}$, respectively η_R overestimate the error in a range of (0.69, 1.02), respectively (3.3, 25, 6) for all tested

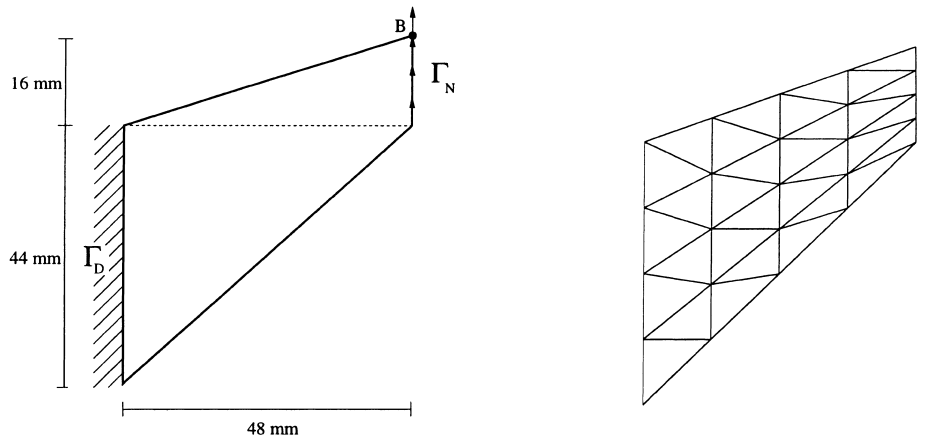

Fig. 6. Cook's membrane problem. System and initial mesh in Section 4.2.

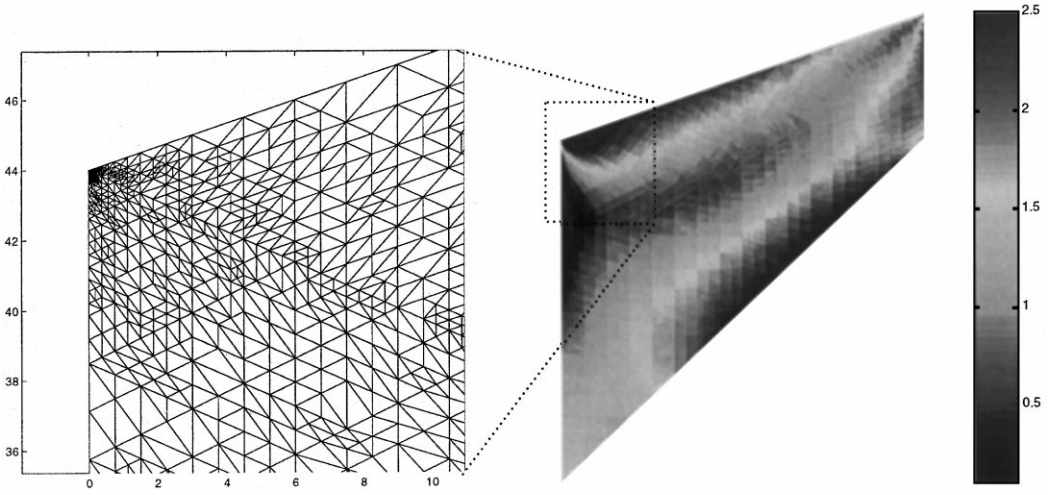

Fig. 7. Mesh $\mathscr{T}_{11}$ generated by Algorithm $(A^0_{\mathscr{A}})$ and magnified detail at (0, 44) of Section 4.2.

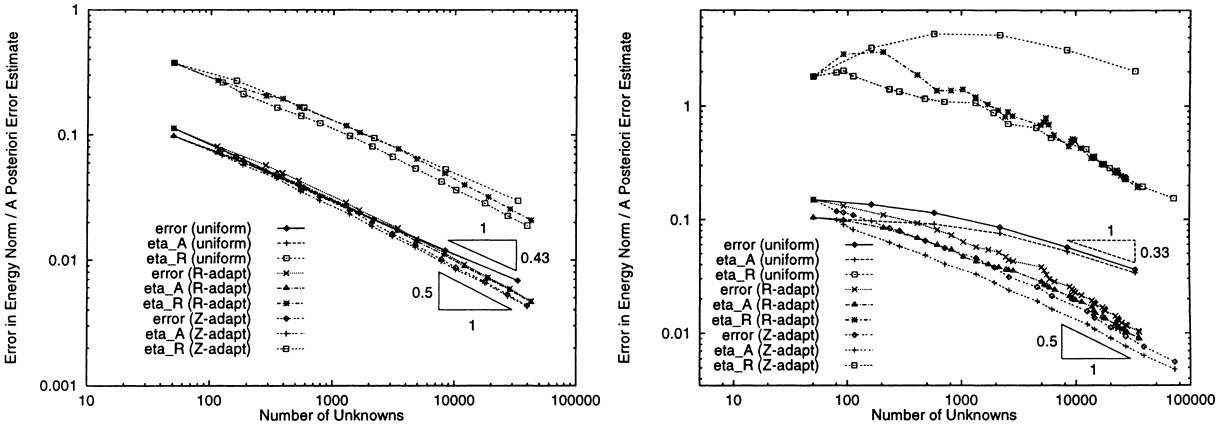

Fig. 8. Errors e_N and error indicators $\eta_{\mathscr{A}}$ and η_R vs N for uniform and adaptive meshes from Algorithms $(A^0_{\mathscr{A}})$ and (A^0_R) of Section 4.2. ($\nu = 1/3$ left , $\nu = 0.499$ right).

$\nu \in \{1/3, 0.45, 0.49, 0.499\}$. For coarse uniform meshes, η_R seems to fail for $\nu = 0.499$. By Algorithm $(A^0_{\mathscr{A}})$ we obtain meshes with slightly smaller errors e_N and quantities η_R and η_Z than generated by Algorithm (A^0_Z) and to reach a given error tolerance Algorithm (A^0_R) needs more adaptive iterations than $(A^0_{\mathscr{A}})$.

4.3. Components tension specimen

A compact tension specimen in Fig. 9 is loaded with a surface load $g = (0, 100)$ on $\Gamma_N = \{(x, y) \in \Gamma : |y| = 60\}$ and $f = 0; E = 100000$ and $\nu = 0.3$ and 0.499. The specimen is subjected to a vertical elongation. As the problem is symmetric, one half of the domain was discretized. We fixed the horizontal displacement with the constraint that the integral mean of all horizontal displacements is zero. For coarse meshes, the problem behaves like a problem with re-entrant corner at $A = (50, 0)$ and hence we expect a higher mesh-refinement. The numerical solution for this problem with $\nu = 0.3$ and $N = 21503$ and a magnification of the adaptively refined mesh around A is provided in Fig. 10. The approximated energies for this example are $G_{0.3} = 0.520590$ and $G_{0.499} = 0.490950$. The resulting improvement of the convergence is outlined in Fig. 11, where the error e_N and the a posteriori error estimates $\eta_{\mathscr{A}}$ and η_R are plotted versus the number of degrees of freedom N. The convergence rate of $n_{\mathscr{A}}$ and n_R is approximately 1 for the adaptive meshes and (computed from the last two meshes) 0.45, respectively 0.14 for uniform meshes. Similar as in Subsections 4.1 and 4.2 the overestimation factors corresponding to $n_{\mathscr{A}}$ and n_R lie here in a range of

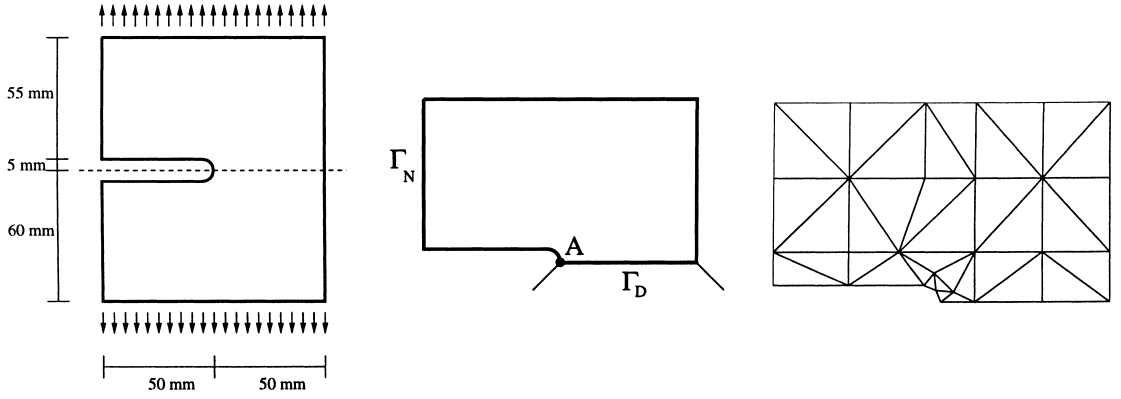

Fig. 9. System and initial mesh in Section 4.3.

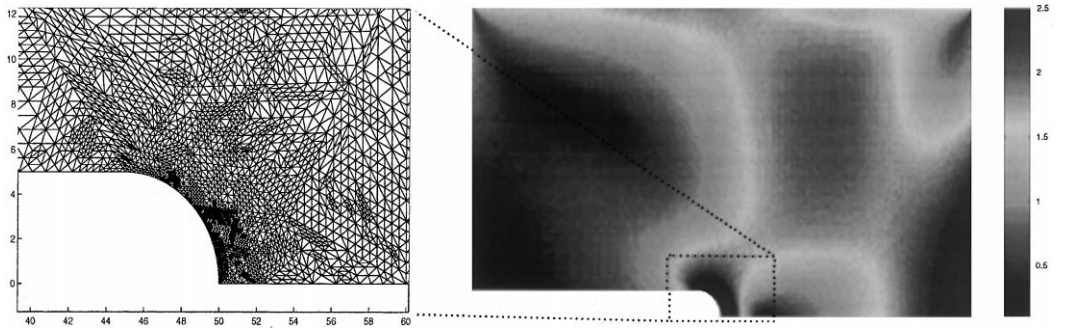

Fig. 10. Mesh $\mathscr{T}_{10}$ generated by Algorithm $(A^0_{\mathscr{A}})$ and magnified detail at A of Section 4.3.

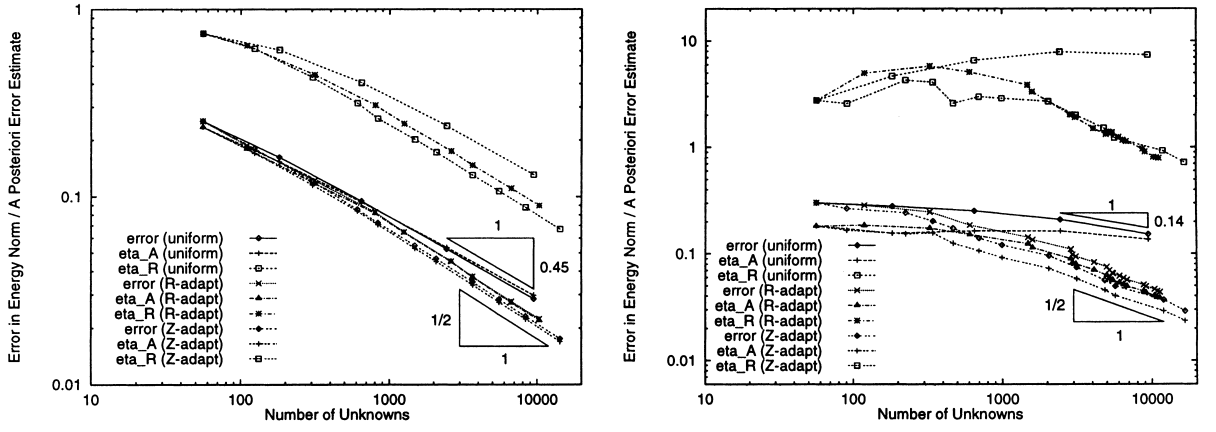

Fig. 11. Errors e_{N} and error indicators $\eta_{\mathscr{A}}$ and η_R vs N for uniform and adaptive meshes from Algorithms $(A^0_{\mathscr{A}})$ and (A^0_R) of Section 4.3. ($\nu = 0.3$ left , $\nu = 0.499$ right)

(0.6, 1.04), respectively (2, 94, 45.4) for all tested $\nu \in \{0.3, 0.45, 0.49, 0.499\}$. Both adaptive mesh-refining Algorithms $(A^0_{\mathscr{A}})$ and (A^0_R) improve this experimental convergence order to the optimal order one. The error estimator η_R fails for $\nu = 0.499$ (in the range of considered meshes) and so proves to be nonrobust for incompressibility.

5. Proof of Theorem 2.1

The proof of efficiency is relatively easy and follows from the triangle inequality

$$\|\sigma_h - \sigma^*\|_{L^2(\Omega)} \leqslant \|\sigma - \sigma_h\|_{L^2(\Omega)} + \|\sigma - \sigma^*\|_{L^2(\Omega)} \tag{5.1}$$

for all $\sigma^* \in \mathscr{Q}(\mathscr{T}, g)$. Taking the infimum over σ^* in (5.1) we deduce with (2.17)

$$\eta_Z \leqslant \|\sigma - \sigma_h\|_{L^2(\Omega)} + \inf_{\sigma^* \in \mathscr{Q}(\mathscr{T}, g)} \|\sigma - \sigma^*\|_{L^2(\Omega)} \tag{5.2}$$

and conclude the proof of the lower inequality in (2.18).

The proof of reliability is more involved and we base our proof on a certain global interpolation operator J from [6,7,11].

Definition 5.1. For each regular node $z \in \mathcal{N} \setminus \mathcal{K}$ on the Dirichlet boundary we choose a neighbouring free node $\zeta(z) \in \mathcal{K}$ and $\zeta(z) = \{z\}$ if $z \in \mathcal{K}$. For $z \in \mathcal{K}$ set $I(z) = \{\tilde{z} \in \mathcal{N} : \zeta(\tilde{z} = z\}$ and

$$\psi(z) := \sum_{\zeta \in I(z)} \varphi_\zeta \quad \text{and} \quad \Omega_z := \{x \in \Omega : 0 < \psi_z(x)\}. \tag{5.3}$$

Remarks 5.1.

(i) Note that $(\varphi_z : z \in \mathcal{N})$ and $(\psi_z : z \in \mathcal{N})$ are Lipschitz continuous partitions of unity with supports of a finite overlap. The functions ψ_z are designed by adding neighbouring hat functions in (5.3) for this purpose.

(ii) The choice of a neighbouring free node $\zeta(z)$ means Ω_z connected. If all nodes of an element belong to the Dirichlet boundary we are forced to choose a second neighbour, etc. The constants below depend on this choice and it is only required that the degree of neighbourhood stays bounded and Ω_z is connected.

(iii) If $z \in \mathcal{K}$ lies far apart from Γ_D we have $\psi_z = \varphi_z$ according to $I(z) = \{z\}$. Note that $\varphi_z \neq \psi_z$ implies that $\Gamma_D \cap (\partial\Omega_z)$ has positive surface measure.

Definition 5.2. For $w \in H^1_D(\Omega)^d$ let $Jw = \sum_{z \in \mathcal{K}} w_z \varphi_z$ be defined, for $z \in \mathcal{K}$,

$$(Jw)(z) = w_z := \int_{\Omega_z} w\psi_z \, dx \Big/ \int_{w_z} \varphi_z \, dx.$$

Lemma 5.1. ([6,7,11]) The operator $J : H^1_D(\Omega)^d \to \mathcal{S}^1_D(\mathcal{T})^d$ has the following properties for all $w \in H^1_D(\Omega)^d$

$$\|Jw\|_{H^1(\Omega)} + \|h_{\mathcal{T}}^{-1}(w - Jw)\|_{L^2(\Omega)} + \|h_{\mathcal{E}}^{-1/2}(w - Jw)\|_{L^2(\Omega)} \leqslant c_6 \|\nabla w\|_{L^2(\Omega)} \tag{5.4}$$

and locally for all $z \in \mathcal{K}$ and $h_z := \operatorname{diam}(\Omega_z)$,

$$\|w_z \varphi_z - w\psi_z\|_{L^2(\Omega)} \leqslant c_7 h_z \|\nabla w\|_{L^2(\Omega)}. \tag{5.5}$$

The constants $c_6, c_7 > 0$ depends on the shape of the elements and patches only neither on z, w or any meshsize.

We are now prepared to sketch the proof of the upper estimate of Theorem 2.1 following arguments in [7]. Let $\lesssim$ denote an inequality up to a constant factor C that is $(h_{\mathcal{T}}, \lambda)$-independent but may depend on $c_3, \ldots, c_7$, Ω, Γ_D. Let $e := u - u_h$ denote the displacement error. From (2.5) and Korn's inequality, we obtain

$$\|e\|^2_{H^1(\Omega)} \lesssim \int_\Omega (\sigma - \sigma_h) : \varepsilon(e) \, dx = \int_\Omega (\sigma - \sigma_h) : \varepsilon(w - Jw + v - u_h) \, dx, \tag{5.6}$$

where employed the Galerkin property (which follows from (2.1)–(2.4) to (2.12) and (2.13) for $w := u - v$ and $v \in H^1(\Omega)^d$ with $v = u_D$ on Γ_D and minimal $\|u_h - v\|_{H^1(\Omega)} \leqslant \|e\|_{H^1\Omega}$. Note that $\|w\|_{H^1(\Omega)} \leqslant \|e\|_{H^1(\Omega)} + \|u_h - v\|_{H^1(\Omega)} \leqslant 2\|e\|_{H^1(\Omega)}$ it is shown in [Lemma 4.1, 7] that we can find some v for u_h with (2.10) by extending the boundary values $(u_h - u_D)|\Gamma_D$ such that

$$\|u_h - v\|_{H^1(\Omega)} \leqslant \|h_{\mathcal{E}}^{3/2} \partial^2_{\mathcal{E}} u_D / \partial s^2\|_{L^2(\Gamma_D)}. \tag{5.7}$$

Similarly, for any $\sigma^* \in \mathcal{Q}(\mathcal{T}, g)$, one can prove that

$$\|h_{\mathcal{E}}^{1/2}(g - \sigma^* n)\|_{L^2(\Gamma_N)} \lesssim \|h_{\mathcal{E}}^{3/2} \partial_{\mathcal{E}} g / \partial s\|_{L^2(\Gamma_N)}. \tag{5.8}$$

Cauchy's inequality, Lemma 5.1, (2.1) and (2.2) and an integration by parts show (notice $w - Jw = 0$ on Γ_D)

$$
\begin{aligned}
&\int_\Omega (\sigma-\sigma_h) : \varepsilon(w-Jw)\,\mathrm{d}x = \int_\Omega (\sigma-\sigma_h) : \nabla(w-Jw)\,\mathrm{d}x \\
&= \int_\Omega (\sigma-\sigma^*) : \nabla(w-Jw)\,\mathrm{d}x + \int_\Omega (\sigma^*-\sigma_h) : \nabla(w-Jw)\,\mathrm{d}x \\
&\leqslant \int_{\Gamma_N} (g-\sigma_n^*)\cdot(w-Jw)\,\mathrm{d}x + \int_\Omega (f+\operatorname{div}\sigma^*)\cdot(w-Jw)\,\mathrm{d}x \\
&\quad + \|\sigma_h-\sigma^*\|_{L^2(\Omega)}\|w-Jw\|_{H^1(\Omega)} \\
&\lesssim \|h_{\mathscr{E}}^{1/2}(g-\sigma^* n)\|_{L^2(\Gamma_N)}\|h_{\mathscr{E}}^{-1/2}(w-Jw)\|_{L^2(\Gamma_N)} \\
&\quad + \int_\Omega f\cdot(w-Jw)\,\mathrm{d}x + \|h_{\mathscr{T}}\operatorname{div}\sigma^*\|_{L^2(\Omega)}\|\nabla w\|_{L^2(\Omega)} \\
&\quad + \|\sigma_h-\sigma^*\|_{L^2(\Omega)}\|\nabla w\|_{L^2(\omega)} \\
&\lesssim \|e\|_{H^1(\Omega)}\Big(\|\sigma_h-\sigma^*\|_{L^2(\Omega)} + \|h_e^{3/2}\partial_{\mathscr{E}} g/\partial s\|_{L^2(\Gamma_N)} + \|h_{\mathscr{T}}\operatorname{div}\sigma^*\|_{L^2(\Omega)}\Big) \\
&\quad + \int_\omega f\cdot(w-Jw)\,\mathrm{d}x.
\end{aligned}
\tag{5.9}
$$

The last term is analysed utilising $\sum_{z\in\mathscr{K}} w\psi_z = w$ and $Jw = \sum_{z\in\mathscr{K}} w_z\varphi_z$, where $w_z\in\mathbb{R}$ satisfies $\int_{\Omega_z}(w\psi_z - w_z\varphi_z)\,\mathrm{d}x = 0$ for all $z\in\mathscr{K}$.

Therefore, for all $f_z\in\mathbb{R}^d$

$$
\begin{aligned}
\int_\Omega f\cdot(w-Jw)\,\mathrm{d}x &= \sum_{z\in\mathscr{K}}\int_\Omega f\cdot(w\psi_z - w_z\varphi_z)\,\mathrm{d}x = \sum_{z\in\mathscr{K}}\int_{\Omega_z}(f-f_z)\cdot(w\psi_z - w_z\varphi_z)\,\mathrm{d}x \\
&\lesssim \sum_{z\in\mathscr{K}} h_z\|f-f_z\|_{L^2(\Omega_z)}\|\nabla w\|_{L^2(\Omega_z)} \lesssim \left(\sum_{z\in\mathscr{K}} h_z^2\|f-f_z\|_{L^2(\Omega_z)}^2\right)^{1/2}\|e\|_{H^1(\Omega)}.
\end{aligned}
\tag{5.10}
$$

In the last step we employed a Cauchy inequality and made essential use of the fact that the patches $(\Omega_z : z\in\mathscr{K})$ have only a finite overlap. A Poincaré inequality on Ω_z shows

$$
\min_{f_z\in\mathbb{R}}\|f-f_z\|_{L^2(\Omega_z)} \lesssim h_z\|\nabla f\|_{L^2(\Omega_z)}.
$$

Using this in (5.10), observing $h_z\lesssim h_\Gamma$ for $T\subset\bar{\Omega}_z$, and again using the finite overlap of the patches we eventually infer

$$
\int_\Omega f\cdot(w-Jw)\,\mathrm{d}x \lesssim \|\nabla w\|_{L^2(\Omega)}\|h_{\mathscr{T}}^2\nabla f\|_{L^2(\Omega)}. \tag{5.11}
$$

The proof focusses now on $\|h_{\mathscr{T}}\operatorname{div}\sigma^*\|_{L^2(\Omega)}$. On each element $T\in\mathscr{T}$, div σ^* is a constant vector, $\sigma^*|_T$ is affine. Hence an inverse inequality [4,5,12] shows for the constant $\bar{\sigma}_h|_T$

$$
\|\operatorname{div}\sigma^*\|_{L^2(T)} = \|\operatorname{div}(\sigma^*-\bar{\sigma}_h)\|_{L^2(T)} \lesssim \frac{1}{h_T}\|\sigma^*-\bar{\sigma}_h\|_{L^2(T)}
$$

and so, with a triangle inequality.

$$
\|h_{\mathscr{T}}\operatorname{div}\sigma^*\|_{L^2(\Omega)} \lesssim \|\sigma^*-\bar{\sigma}_h\|_{L^2(\Omega)}\|\sigma^*-\sigma_h\|_{L^2(\Omega)} + \|\sigma_h-\bar{\sigma}_h\|_{L^2(\Omega)}. \tag{5.12}
$$

Gathering (5.6) and (5.7) and (5.10)–(5.12) together we have shown that there holds

$$
\begin{aligned}
\|e\|_{H^1(\Omega)}^2 \lesssim \|e\|_{H^1(\Omega)}\Big(&\|h_{\mathscr{E}}^{3/2}\partial_{\mathscr{E}}^2 u_D/\partial^2 s\|_{L^2(\Gamma_D)} + \|\sigma_h-\sigma^*\|_{L^2(\Omega)} + \|\sigma_h-\bar{\sigma}_h\|_{L^2(\Omega)} \\
&+ \|h_{\mathscr{E}}^{3/2}\partial_{\mathscr{E}} g/\partial s\|_{L^2(\Gamma_N)} + \|h_{\mathscr{T}}^2\nabla f\|_{L^2(\Omega)}\Big).
\end{aligned}
$$

from which we conclude the proof.

6. Locking

Nearly incompressible linear elastic materials yield Lamé parameter $\lambda \to \infty$ and bounded positive μ in (4.1). Incompressibility locking is observed for standard conforming finite element methods in the sense that (for uniform meshes with meshsize h)

$$|||\sigma - \sigma_h||| := \left(\int_\Omega (\sigma - \sigma_h) : \mathbb{C}^{-1}(\sigma - \sigma_h)\, \mathrm{d}x \right)^{1/2} \propto \sqrt{\lambda} h^\alpha \tag{6.1}$$

($\propto$ denotes propotionality and $\alpha > 0$ depends on corner signularities as in Section 4.1.) The empirical observation of (6.1) is seen in Fig. 2 and can be explained as follows. The Galerkin scheme yields a quasi optimal a priori error estimate

$$|||\sigma - \sigma_h|||^2 = \min_{v_h \in \mathscr{S}^1_{\mathrm{D}}(\mathscr{T})^d} \left(\lambda \|\mathrm{div}(e - v_h)\|^2_{L^2(\Omega)} + 2\mu \|\mathrm{dev}\varepsilon(e - v_h)\|^2_{L^2(\Omega)} \right) \tag{6.2}$$

for the error $e := u - u_h$, respectively $\sigma - \sigma_h = \mathbb{C}\varepsilon(e)$. If μ is bounded but λ relatively large, we would expect that, $\lambda \|\mathrm{div}(e - v_h)\|^2_{L^2(\Omega)}$ on the right-hand side of (6.2) is of the same order as $2\mu \|\mathrm{dev}\,\varepsilon(e - v_h)\|^2_{L^2(\Omega)}$. But for standard conforming lower order finite element schemes this is not observed. Instead, we merely observe that $\|\mathrm{div}(e - v_h)\|_{L^2(\Omega)}$ is of the same magnitude as $\|\mathrm{dev}\,\varepsilon(e - v_h)\|_{L^2(\Omega)}$ and so

$$\|\mathrm{div}\ e\|_{L^2(\Omega)} \geqslant \kappa \|\mathrm{dev}\ \varepsilon(e)\|_{L^2(\Omega)} \tag{6.3}$$

for some $\kappa > 0$. Recall that the discretisation locks if (6.3) holds since the large penalty parameter $\lambda \to \infty$ in (6.2) enforce $\|\mathrm{div}\ \mathrm{e}\|_{L^2(\Omega)} \to 0$ which contradicts (6.3).

Depending on the smoothness of u we have generically (for sufficiently small mesh sizes)

$$c_8 h^\alpha \leqslant \|\nabla e\|_{L^2(\Omega)} \leqslant c_9 h^\alpha \tag{6.4}$$

and conclude with (6.2) that

$$c_{10}\sqrt{\lambda} h^\alpha \leqslant \sqrt{\lambda(1+\kappa^2)/2}\|\nabla e\|_{L^2(\Omega)} \leqslant c_{10}\sqrt{\lambda} h^\alpha \leqslant |||\sigma - \sigma_h||| \leqslant \sqrt{\lambda + 2\mu/\kappa^2}\|\nabla e\|_{L^2(\Omega)} \leqslant c_{11}\sqrt{\lambda} h^\alpha. \tag{6.5}$$

This is in agreement with (6.1). Notice that the failure of the scheme has nothing to do with the hopefully good estimation of the possible poor error e.

The a posteriori error estimate from Theorem 2.1 leads in the case of (4.1) to the estimate

$$|||\sigma - \sigma_h|||^2 \leqslant c_{12} \|\varepsilon(e)\|_{L^2(\Omega)} \left(\min_{\sigma_h^* \in \mathscr{S}^1_g(\mathscr{T})} \|\sigma_h - \sigma_h^*\|_{L^2(\Omega)} + \mathrm{h.o.t.} \right), \tag{6.6}$$

where the higher order terms h.o.t and the constant c_{12} are λ-independent. Suppose that we have locking in the sense that (6.3) is true for some $\kappa > 0$. Then

$$\|\varepsilon(e)\|_{L^2(\Omega)} \leqslant \left(\frac{1}{d} + \frac{1}{\kappa^2} \right)^{1/2} \|\mathrm{div}\ e\|_{L^2(\Omega)} \leqslant \lambda^{-1/2} \left(\frac{1}{d} + \frac{1}{\kappa^2} \right)^{1/2} |||\sigma - \sigma_h||| \tag{6.7}$$

which yields in (6.6) (with a κ-depending constant c_{13}) that

$$|||\sigma - \sigma_h||| \leqslant c_{13} \min_{\sigma_h^* \in \mathscr{S}^1_g(\mathscr{T})} \lambda^{-1/2} \|\sigma_h - \sigma_h^*\|_{L^2(\Omega)} + \lambda^{-1/2}\ \mathrm{h.o.t.}. \tag{6.8}$$

It is directly seen from the definition of the inverse of the elasticity tensor that

$$\lambda^{-1/2} |\sigma_h - \sigma_h^*| \leqslant \frac{c_{14}}{c_{13}} |\sigma_h - \sigma_h^*|_{\mathbb{C}^{-1}}$$

and so (6.8) yields

$$|||\sigma - \sigma_h||| \leqslant c_{14} \min_{\sigma_h^* \in \mathscr{P}_g^1(\mathscr{T})} \lambda^{-1/2} |||\sigma_h - \sigma_h^*||| + \lambda^{-1/2} \text{ h.o.t.} \tag{6.9}$$

which indicates that the Z-estimator is λ-robust reliable if locking is present in the sense of (6.3). Note that the influence of h.o.t. is decreasing if λ increases. This heuristic argument indicates the surprising robustness of the Z-estimator in practice. A detailed analysis will be reported on in Part II of this paper [10].

References

[1] J. Alberty, C. Carstensen, S.A. Funken, Remarks around 50 lines of Matlab: short finite element implementation, Numer. Alg. 20 (1999) 117–137.

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

[3] I. Babuška, T. Strouboolis, C.S. Upadhyay, S.K. Gangaraj, K. Copps, Validation of a posteriori error estimators by numerical approach, Int. J. Numer. Methods Engrg. 37 (1994) 1073–1123.

[4] D. Braess, Finite Elements, Cambridge University Press, Cambridge, 1997.

[5] S.C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, Texts Appl. Math. 15, Springer, New York, 1994.

[6] C. Carstensen, Quasi interpolation and a posteriori error analysis in finite element method. M^2AN 33 (1999) 1187–1202.

[7] C. Carstensen, S. Bartels, Each averaging technique yields reliable a posteriori error control in FEM on unstructural grids Part I: Low order conforming, nonconfirming, and mixed FEM. Berichtsreihe des Mathematischen Seminars Kiel, Technical report 99-11, Christian-Albrechts-Universität zu Kiel, Kiel (http://www.numerik.uni-kiel.de/reports/1999/99-11.html), 1999.

[8] C. Carstensen, G. Dolzmann, S.A. Funken, D. Helm, Locking-free adaptive mixed finite element methods in elasticity, Berichtstreihe des Mathematischen Seminars Kiel, Technical report 98-19 Universität Kiel (http://www.numerick.uni-kiel.de/reports/1998/98-19.html). Comp. Methods Appl. Mech. Engrg. (accepted 06/99), 1998.

[9] C. Carstensen, S.A. Funken: Constants in Clèment-interpolation error and residual based a posteriori error estimates in Finite Element Methods. Berichtsreihe des Mathematischen Seminars Kiel, Technical report 97-11, Christian-Albrechts-Universität zu Kiel, Kiel. East-West Numer. Anal. (to appear). (http://www.numerik.uni-kiel.de/reports/1997/97-11.html), 1997.

[10] C. Carstensen, S.A. Funken, Averaging technique for FE-a posteriori error control in elasticity. Part II: λ-independent estimates. Comput. Mech. Appl. Mech. Engrg., in press. Berichtsreihe des Mathematischen Seminars Kiel, Technical report 00-6, Christian-Albrechts-Universität zu Kiel, Kiel (2000) (http://www.numerik.uni-kiel.de/reports/2000/00-6.html).

[11] C. Carstensen, R. Verfürth, Edge residuals dominate a posteriori error estimates for low order finite element methods. Berichtsreihe des Mathematischen Seminars Kiel, Technical report 97-6, Christian-Albrechts-Universität zu Kiel, SIAM J. Numer. Anal. (1998) accepted for publication.

[12] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.

[13] R. Duran, M.A. Muschietti, R. Rodriguez, On the asymptotic exactness of error estimators for linear triangular elements, Numer. Math. 59 (1991) 107–127.

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

[15] W. Hoffmann, A.H. Schatz, L.B. Whalbin, G. Wittum, Asymptotically exact a posteriori error estimators for the pointwise gradient error on each element in irregular meshes. Part 1: A smooth problem and globally quasi-uniform meshes. Preprint 1999.

[16] R. Rodriguez, Some remarks on Zienkiewicz-Zhu estimator, Int. J. Numer. Methods PDE 10 (1994) 625–635.

[17] R. Rodriguez, A posteriori error analysis in the finite element method, in: M. Krizek et al. (Ed.), Finite Element Methods. 50 yrs of the Courant Element, Conference held at the University of Jyvaeskylae, Finland, 1993, Marcel Dekker, New York, Lecture Notes on Pure Appl. Math., vol. 164, 1994, pp. 389–397.

[18] R. Verfürth, A Review of a Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques, Wiley, Amsterdam, 1996.

[19] E. Zeidler, Nonlinear Functional Analysis and its Application IV, Springer, New York, 1988.

[20] O.C. Zienkiewicz, J.Z. Zhu, A simple error estimator and adaptive procedure for practical engineering analysis, Int. J. Numer. Methods Engrg. 24 (1987) 337–357.