SIAM J. NUMER. ANAL.
Vol. 40, No. 1, pp. 180–197
© 2002 Society for Industrial and Applied Mathematics

A STABLE AFFINE-APPROXIMATE FINITE ELEMENT METHOD*

K. ARUNAKIRINATHAR† AND B. D. REDDY‡

Abstract. The notion of the affine figure closest to a given quadrilateral can be given a precise mathematical definition. The resulting figure is referred to as the equivalent parallelogram associated with a quadrilateral. Equipped with such a concept, it is then feasible to consider finite element approximations in which the original quadrilateral elements are replaced by their equivalent parallelograms. The idea is appealing, not least because of the resulting economy arising from computations performed on an element generated by an affine map. Furthermore, numerical experiments reported recently indicate that highly efficient and accurate schemes result when such a concept is combined with the enhanced strain method or the method of incompatible modes. The purpose of this work is to analyze finite element schemes resulting from approximation of quadrilaterals by their equivalent parallelograms. The focus is on low-order (bilinear) elements, and the analysis is carried out in the context of linear elasticity for standard approximations as well as for those which use enhanced strains. The affine approximation applies only to the element map, and the primary unknown (the displacement vector in the context of elasticity) is approximated by conventional piecewise bilinear functions. The analysis confirms convergence at the optimal rate, provided that the deviations of the quadrilaterals from their equivalent parallelograms are at most $O(h)$.

Key words. equivalent parallelogram, affine approximations, four-noded quadrilateral, error estimates

AMS subject classifications. 65N30, 65N15

PII. S0036142900382442

1. Introduction. The use of low-order elements in finite element analyses of complex problems carries with it significant advantages. Most particularly, such schemes are highly economical and for this reason are attractive.

Finite element analyses based on four-noded quadrilaterals in two dimensions, and on eight-noded hexahedral elements in three, are widely used. Unfortunately, they are not without their drawbacks. In problems of solid mechanics in which bending deformations dominate, analyses based on these elements exhibit poor accuracy, at least when coarse meshes are used. In addition, in the incompressible limit, or when the compressibility is small, locking behavior is experienced.

There is a vast literature that is devoted to the construction of methods which are intended to overcome the problems referred to, while retaining the advantages of using low-order elements. One commonly used set of remedies is that based on a combination of underintegration plus stabilization (see, for example, the work of [4] and [8]). The great advantage of this approach is its efficiency, in that only a single integration point is used. However, the eigenvalues of the stiffness matrix are required in the process, and it is not possible to evaluate these without a relatively high degree of effort, for nonaffine elements.

Another popular approach is that associated with the enrichment or enhancement of the strain by the addition of suitably chosen basis functions. The key work dealing with enhanced strain formulations is [14], which in turn contains as a special case

*Received by the editors December 13, 2000; accepted for publication (in revised form) October 9, 2001; published electronically April 12, 2002.
http://www.siam.org/journals/sinum/40-1/38244.html

†Department of Mathematics, University of Swaziland, PO Luyengo M205, Kwaluseni, Swaziland.
‡Faculty of Science, University of Cape Town, 7701 Rondebosch, South Africa (bdr@science.uct.ac.za). The work of this author was supported by the National Research Foundation of South Africa.

the nonconforming method of incompatible modes due to Wilson, Taylor, Doherty, and Ghaboussi [17] for rectangular elements and extended by Taylor, Beresford, and Wilson [16] to incorporate arbitrary quadrilaterals. The method has been successfully extended to nonlinear problems (see, for example, [15]). Reddy and Simo [11] have shown, for linear problems and for affine elements, that the enhanced strain method is stable and convergent, while Arunakirinathar and Reddy [3] have extended that work to include the case of arbitrary quadrilaterals.

The enhanced strain method is still not without its drawbacks. For example, the quality of approximations for arbitrary elements declines with an increase in distortion of the elements. Consideration of this shortcoming, together with a desire to improve the efficiency of computations associated with arbitrary quadrilaterals or hexahedra, leads naturally to the notion of replacing the arbitrary quadrilateral by the affine element that is closest to it, in a manner that can be made precise. Such an affine element is known as the equivalent parallelogram in two dimensions, and the equivalent parallelepiped in three. It is important to bear in mind that the affine approximation applies only to the element map and that the primary unknown (the displacement vector in the context of elasticity) is still approximated by piecewise bilinear functions in two dimensions and by piecewise trilinear functions in three.

It has been shown in [2] that the interpolation error obtained by using the equivalent parallelogram instead of the original quadrilateral is of the same order as that corresponding to the usual interpolation error. The element stiffness matrices associated with the equivalent elements are therefore admissible alternatives to the "exact" stiffness matrices of the original elements, while at the same time they are far easier to construct. This set of ideas has been proposed, and then tested numerically, first in the context of problems of linear elasticity in [9] and subsequently for problems involving nonlinearly elastic materials in [12, 13]. In all cases the numerical results are encouraging and suggest a significant improvement in efficiency and accuracy when this approach is used, particularly in circumstances in which element distortions are significant.

The purpose of this work is to carry out a detailed analysis of these affine-approximate finite element methods, for linear problems. The analysis is confined to plane problems but can be extended to three dimensions with little difficulty, though the details are messy. The analysis includes treatment of finite elements without and with the inclusion of enhanced strains. The key results, with respect to both classes of approximations, is that the method converges at the optimal rate provided that the element distortion is sufficiently small—more precisely, provided that deviation of the quadrilateral from the equivalent parallelogram is of the order of mesh size. This notion will be made precise in what follows.

The plan of the remainder of this work is as follows. In section 2 the problem is formulated. Finite element approximations are introduced in section 3, as is the notion of the equivalent parallelogram. The analysis of the affine-approximate method is carried in section 4 for the problem without enhancement, while section 5 is devoted to an analysis of the problem with enhancement.

2. The boundary-value problem of elasticity. The model problem of relevance is the displacement boundary-value problem of linear elasticity. Suppose that a linear elastic body occupies a region $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$. The body has boundary Γ. Then the governing equations which are required to hold on Ω are the equation of equilibrium

$$-\operatorname{div} \boldsymbol{\sigma} = \boldsymbol{b}, \tag{1}$$

the strain-displacement relation

$$\boldsymbol{\epsilon}(\boldsymbol{u}) = \tfrac{1}{2}(\nabla \boldsymbol{u} + [\nabla \boldsymbol{u}]^T), \tag{2}$$

and the elastic constitutive equation

$$\boldsymbol{\sigma} = \mathbb{C}[\boldsymbol{\epsilon}]. \tag{3}$$

Here $\boldsymbol{\sigma}$ is the symmetric Cauchy stress tensor, $\boldsymbol{\epsilon}$ is the infinitesimal strain tensor, $\boldsymbol{u}$ is the displacement vector, $\mathbb{C}$ is the fourth-order elasticity tensor, and $\boldsymbol{b}$ is a prescribed body force vector.

For convenience we assume that the displacement satisfies the homogeneous Dirichlet boundary condition, that is,

$$\boldsymbol{u} = \boldsymbol{0} \quad \text{on } \Gamma. \tag{4}$$

The tensor $\mathbb{C}$ has the symmetries

$$\mathbb{C}_{ijkl} = \mathbb{C}_{jikl} = \mathbb{C}_{ijlk} = \mathbb{C}_{klij} \tag{5}$$

and is assumed to be pointwise stable, in the sense that there exists a constant $c_0 > 0$ such that

$$\mathbb{C}_{ijkl} M_{ij} M_{kl} \geq c_0 M_{ij} M_{ij} \tag{6}$$

for all symmetric matrices $\boldsymbol{M}$, the summation convention for repeated indices being invoked here and henceforth. In addition, the components of $\mathbb{C}$ are assumed to be bounded measurable functions; that is,

$$\mathbb{C}_{ijkl} \in L^\infty(\Omega) \tag{7}$$

for all indices i, j, k, l ranging over 1 to d, with

$$c_\infty := \max_{i,j,k,l} \operatorname{ess\,sup}\{\mathbb{C}_{ijkl}(\boldsymbol{x}) : \ \boldsymbol{x} \in \Omega \ \}. \tag{8}$$

For isotropic elastic materials the elasticity tensor takes the simple form

$$\mathbb{C}[\boldsymbol{\epsilon}] = \lambda \operatorname{tr} \boldsymbol{\epsilon} + 2\mu\, \boldsymbol{\epsilon},$$

in which λ and μ are the Lamé constants. Pointwise stability of $\mathbb{C}$ is equivalent to the condition that the Lamé constants satisfy the inequalities [10]

$$\mu > 0 \quad \text{and} \quad \lambda > -\tfrac{2}{3}\mu.$$

We will make use of the space $L^2(\Omega)$ of square-integrable functions defined on Ω. The inner product and norm on this space are denoted, respectively, by $(\cdot,\cdot)_0$ and $\|\cdot\|_0$. We recall also the definition of the Sobolev spaces $H^m(\Omega)$, where m is an integer; for nonnegative m, these are equivalence classes of functions which, together with their generalized derivatives up to and including those of order m, are members of $L^2(\Omega)$. The Sobolev spaces are Hilbert spaces with inner product and associated norm given by

$$(u, v)_m = \int_\Omega \sum_{|\alpha| \leq m} D^\alpha u(\boldsymbol{x}) D^\alpha v(\boldsymbol{x}) \; d\boldsymbol{x}, \quad \|u\|_m = (u, u)_m^{1/2}$$

for all $u, v \in H^m(\Omega)$. Here $\alpha = (\alpha_1, \ldots, \alpha_d)$ is a multi-index whose components α_i are nonnegative integers, $|\alpha| = \alpha_1 + \cdots + \alpha_d$, and $D^\alpha = \partial^{|\alpha|}/\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}$. The seminorm $|\cdot|_m$ on $H^m(\Omega)$ is defined by

$$|u|_m = \int_\Omega \sum_{|\alpha|=m} D^\alpha u(\boldsymbol{x}) D^\alpha v(\boldsymbol{x})\ d\boldsymbol{x}.$$

We define the space $H_0^m(\Omega)$ of functions in $H^m(\Omega)$ which, together with their derivatives of order up to and including those of order $m-1$, vanish on the boundary in the sense of traces. The seminorm $|\cdot|_m$ is a norm on $H_0^m(\Omega)$, equivalent to the standard norm $\|\cdot\|_m$. Finally, the space $H^{-m}(\Omega)$, for m a nonnegative integer, may be identified with the topological dual space of $H_0^m(\Omega)$.

We are now in a position to define the standard variational problem in linear elasticity. For this purpose we denote by $V := [H_0^1(\Omega)]^d$ the space of admissible displacements and define the bilinear form $a(\cdot,\cdot)$ and linear functional $\ell(\cdot)$ by

$$a : V \times V \to \mathbb{R}, \qquad a(\boldsymbol{u}, \boldsymbol{v}) = \int_\Omega \mathbb{C}\boldsymbol{\epsilon}(\boldsymbol{u}) : \boldsymbol{\epsilon}(\boldsymbol{v})\ d\boldsymbol{x}, \tag{9}$$

$$\ell : V \to \mathbb{R}, \qquad \ell(\boldsymbol{v}) = \int_\Omega \boldsymbol{b} \cdot \boldsymbol{v}\ d\boldsymbol{x}. \tag{10}$$

The properties of $\mathbb{C}$ guarantee that $a(\cdot,\cdot)$ is symmetric, continuous, and V-elliptic; that is, $a(\boldsymbol{v}, \boldsymbol{u}) = a(\boldsymbol{u}, \boldsymbol{v})$, and there exist positive constants M and α such that

$$|a(\boldsymbol{u}, \boldsymbol{v})| \leq M \|\boldsymbol{u}\|_V \|\boldsymbol{v}\|_V \quad \text{and} \quad a(\boldsymbol{v}, \boldsymbol{v}) \geq \alpha \|\boldsymbol{v}\|_V^2$$

for all $\boldsymbol{u}$, $\boldsymbol{v} \in V$.

The topological dual of V is denoted by V'.

The standard variational problem is as follows.

Problem S. Given $\boldsymbol{b} \in V'$, find $\boldsymbol{u} \in V$ which satisfies

$$a(\boldsymbol{u}, \boldsymbol{v}) = \ell(\boldsymbol{v}) \tag{11}$$

for all $\boldsymbol{v} \in V$.

It is well known (see, for example, [7]) that Problem S has a unique solution, which satisfies the bound

$$\|\boldsymbol{u}\|_V \leq (1/\alpha)\, \|\ell\|_{V'}. \tag{12}$$

3. Finite element approximations. We confine attention to plane situations, so that $d = 2$. The domain Ω is assumed to be polygonal, and a finite element mesh $\mathcal{T}$ of quadrilateral elements is constructed on Ω in the usual manner. A typical element K in $\mathcal{T}$ is generated by an isoparametric map F from a reference element $\hat{K} \equiv (-1, 1) \times (-1, 1)$. The mesh parameter h is defined by

$$h = \max_{K \in \mathcal{T}} \ \sup\{|\boldsymbol{x} - \boldsymbol{y}| : \ \ \boldsymbol{x}, \boldsymbol{y} \in K\}. \tag{13}$$

We define basis functions $\hat{N}_A \ \ (A = 1, \ldots, 4)$ on $\hat{K}$ by

$$\hat{N}_A(\boldsymbol{\xi}) = \tfrac{1}{4}(1 + \xi\xi_A)(1 + \eta\eta_A),$$

where $\boldsymbol{\xi}_A \equiv (\xi_A, \eta_A)$ are the nodal coordinates on $\hat{K}$ with $(\xi_1 \cdots \xi_4) = (1\ -1\ -1\ 1)$ and $(\eta_1 \cdots \eta_4) = (1\ 1\ -1\ -1)$. Denote by Q_1 the space of bilinear functions spanned by $\hat{N}_A$. Then it is convenient to express the map F in the form

$$F : \hat{K} \to K, \qquad \boldsymbol{x} = F(\boldsymbol{\xi}) = \sum_{A=1}^{4} \boldsymbol{x}_A \hat{N}_A(\boldsymbol{\xi}), \tag{14}$$

in which $\boldsymbol{x}_A$ are the nodal points of K.

The Jacobian matrix $\boldsymbol{J}$ is defined to be the gradient of the map F and is the matrix with components

$$J_{ij} = \frac{\partial F_i(\boldsymbol{\xi})}{\partial \xi_j}, \quad i, j = 1, 2.$$

We also set

$$j = \det \boldsymbol{J}.$$

Next we define the space V^h by

$$V^h = \{\boldsymbol{v}_h \in V : \ (v_h)_i|_K \circ F \in Q_1\}. \tag{15}$$

In other words, if we define the function $\hat{\boldsymbol{v}}$ on $\hat{K}$ by

$$\hat{\boldsymbol{v}}(\boldsymbol{\xi}) = \boldsymbol{v}|_K(\boldsymbol{x}),$$

in which $\boldsymbol{\xi}$ and $\boldsymbol{x}$ are related through (14), then $\hat{\boldsymbol{v}} \in Q_1$.

The standard discrete variational problem takes the following form.

Problem $\mathrm{S^h}$. Given $\boldsymbol{b} \in V'$, find $\boldsymbol{u}_h \in V^h$ which satisfies

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = \ell(\boldsymbol{v}_h) \tag{16}$$

for all $\boldsymbol{v}_h \in V^h$.

It is well known (see [7]) that Problem $\mathrm{S^h}$ has a unique solution, and furthermore, provided that $\boldsymbol{u} \in [H^2(\Omega)]^2$, there exists a constant $C > 0$, depending on Ω and on u, but independent of h, such that

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_V \le Ch.$$

It is instructive, and relevant to the developments that follow, to note that the bilinear form and linear functional appearing in (16) are usually evaluated on the reference element. In order to do this it is necessary to carry out transformations of the functions appearing in $a(\cdot,\cdot)$ and $\ell(\cdot)$. Thus, if we define $\hat{\boldsymbol{v}}$ as above, the chain rule gives

$$\frac{\partial \hat{v}_i}{\partial \xi_j} = \sum_{k=1}^{2} \frac{\partial v_i}{\partial x_k} \frac{\partial F_k}{\partial \xi_j},$$

or if we define the tensors or matrices $\boldsymbol{L}$ and $\hat{\boldsymbol{L}}$ by

$$L_{ij} = \frac{\partial v_i}{\partial x_j}, \quad \hat{L}_{ij} = \frac{\partial \hat{v}_i}{\partial \xi_j},$$

then

$$\hat{\boldsymbol{L}} = \boldsymbol{L}\boldsymbol{J}.$$

The transformation of the strain tensor may now be easily carried out, and we have, using an obvious notation,

$$\begin{aligned}\boldsymbol{\epsilon}(\boldsymbol{v}) &= \tfrac{1}{2}(\boldsymbol{L} + \boldsymbol{L}^T)\\ &= \tfrac{1}{2}(\hat{\boldsymbol{L}}\boldsymbol{J}^{-1} + \boldsymbol{J}^{-T}\hat{\boldsymbol{L}}^T) := \hat{\boldsymbol{\epsilon}}(\hat{\boldsymbol{v}}).\end{aligned} \tag{17}$$

Hence the bilinear form may be evaluated according to

$$a(\boldsymbol{u}_h, \boldsymbol{v}_h) = \sum_{K \in \mathcal{T}} a_K(\boldsymbol{u}_h, \boldsymbol{v}_h) \tag{18}$$

in which

$$\begin{aligned}a_K(\boldsymbol{u}_h, \boldsymbol{v}_h) &= \int_K \mathbb{C}[\boldsymbol{\epsilon}(\boldsymbol{u}_h)] : \boldsymbol{\epsilon}(\boldsymbol{v}_h)\ dxdy\\ &= \int_{\hat{K}} \hat{\mathbb{C}}[\hat{\boldsymbol{\epsilon}}(\hat{\boldsymbol{u}}_h)] : \hat{\boldsymbol{\epsilon}}(\hat{\boldsymbol{v}}_h)\ j\, d\xi d\eta\end{aligned} \tag{19}$$

and where $\hat{\mathbb{C}} = \mathbb{C} \circ F$.

3.1. The equivalent parallelogram. The notion of the equivalent parallelogram associated with a quadrilateral arises naturally when one seeks to define the parallelogram that is closest to the quadrilateral in a sense that can be made mathematically precise. Assuming that such a parallelogram can be constructed, the quadrilateral can then be viewed as a perturbation of the parallelogram.

This problem was considered by Arunakirinathar and Reddy [2], who showed that the equivalent parallelogram can be constructed as follows. Suppose that the map from the reference element $\hat{K}$ to an arbitrary quadrilateral K is given by (14); then the equivalent parallelogram $\tilde{K}$ associated with K is defined by the *affine* map $\tilde{F}$ obtained simply by discarding the bilinear terms in (14). That is, if we define the vector $\boldsymbol{k}$ by

$$\boldsymbol{k} = \tfrac{1}{4}(\boldsymbol{x}_1 - \boldsymbol{x}_2 + \boldsymbol{x}_3 - \boldsymbol{x}_4),$$

then the map $\tilde{F}$ may be expressed in the form

$$\begin{aligned}\tilde{F}(\boldsymbol{\xi}) &= F(\boldsymbol{\xi}) - \boldsymbol{k}\xi\eta\\ &= \sum_{A=1}^{4} N_A(\boldsymbol{\xi})\tilde{\boldsymbol{x}}_A,\end{aligned}$$

in which the nodal points $\tilde{\boldsymbol{x}}_A$ of the equivalent parallelogram are defined by

$$\tilde{\boldsymbol{x}}_A = \tfrac{3}{4}\boldsymbol{x}_A + \tfrac{1}{4}(\boldsymbol{x}_{A+1} - \boldsymbol{x}_{A+2} + \boldsymbol{x}_{A+3}), \quad A = 1, \ldots, 4 \text{ (modulo 4)}.$$

These notions are illustrated in Figure 1.

The equivalent parallelogram has some interesting properties [2]: for example, K and $\tilde{K}$ have the same areas, their sides intersect at midpoints, and the lengths of the corresponding diagonals are equal. These last two properties are evident in Figure 1.

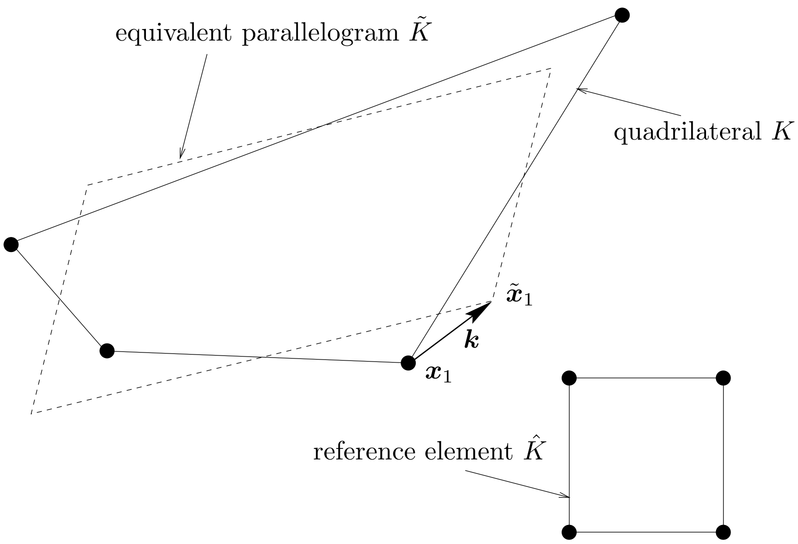

FIG. 1. *The equivalent parallelogram associated with a quadrilateral.*

It is necessary to characterize mathematically the relationship between the quadrilateral and its equivalent parallelogram; more particularly, we need to characterize the notion of closeness between K and $\tilde{K}$. Suppose that the affine map from $\hat{K}$ to $\tilde{K}$ takes the form

$$\tilde{\boldsymbol{x}} = \boldsymbol{C}\boldsymbol{\xi} + \boldsymbol{c}, \tag{20}$$

in which $\boldsymbol{C}$ and $\boldsymbol{c}$ are, respectively, a constant matrix and vector; then the *distortion parameter* τ_K for element K is defined by

$$\tau_K = |\boldsymbol{C}^{-1}\boldsymbol{k}|. \tag{21}$$

The distortion is thus measured by mapping the vector $\boldsymbol{k}$ back to the reference element $\hat{K}$, using $\boldsymbol{C}^{-1}$. From Figure 1 and the definition of $\boldsymbol{k}$ it is clear that $\boldsymbol{k} = \boldsymbol{0}$, and $\tau_K = 0$, if and only if K is a parallelogram.

The distortion parameter associated with a finite element mesh may now be defined by

$$\tau = \max_{K \in \mathcal{T}} |\tau_K|. \tag{22}$$

We will also require the notion of an *h-regular mesh*, which is defined to be a finite element mesh for which $\tau = O(h)$.

The role played by τ_K in characterizing the difference between K and $\tilde{K}$ may be seen more clearly by examining the properties of the map $G : \tilde{K} \to K$ from the parallelogram to the quadrilateral. If we set $\boldsymbol{J}' = DG$ and $j' = \det \boldsymbol{J}'$, then since

$$\boldsymbol{x} = G(\tilde{\boldsymbol{x}}) = \tilde{\boldsymbol{x}} + \boldsymbol{k}\xi\eta,$$

it is straightforward to show, also using (20), that

$$\boldsymbol{J}' = \boldsymbol{I} + \boldsymbol{C}^{-1}\boldsymbol{k} \otimes \boldsymbol{l}$$

and

$$j' = 1 + \boldsymbol{C}^{-1}\boldsymbol{k} \cdot \boldsymbol{l},$$

in which $\boldsymbol{l} = (\eta, \xi)$. It follows from (21) that

$$\sup_{\boldsymbol{x} \in K} |j'(\boldsymbol{x})| := \|j'\|_\infty \le 1 + \sqrt{2}\,\tau_K. \tag{23}$$

If $|\boldsymbol{A}| = (\sum_{i,j=1}^2 A_{ij}A_{ij})^{1/2}$ for any matrix $\boldsymbol{A}$, then

$$\begin{aligned} |\boldsymbol{J}'| &\le |\boldsymbol{I}| + |\boldsymbol{C}^{-1}\boldsymbol{k} \otimes \boldsymbol{l}| \\ &\le |\boldsymbol{I}| + |\boldsymbol{C}^{-1}\boldsymbol{k}|\,|\boldsymbol{l}| \\ &\le \sqrt{2}\,(1 + \tau_K). \end{aligned} \tag{24}$$

The inverse of the map $\boldsymbol{J}'$ is given by

$$(\boldsymbol{J}')^{-1} = \boldsymbol{I} - \boldsymbol{R}, \tag{25}$$

in which

$$\boldsymbol{R} = \frac{\boldsymbol{C}^{-1}\boldsymbol{k} \otimes \boldsymbol{l}}{1 + \boldsymbol{C}^{-1}\boldsymbol{k} \cdot \boldsymbol{l}}. \tag{26}$$

It follows from (21) that

$$|\boldsymbol{R}| \le \frac{|\boldsymbol{C}^{-1}\boldsymbol{k}|\,|\boldsymbol{l}|}{|1 + \boldsymbol{C}^{-1}\boldsymbol{k} \cdot \boldsymbol{l}|} \le \frac{\sqrt{2}\tau_K}{1 - \sqrt{2}\tau_K}. \tag{27}$$

Likewise, using the identity

$$\det(\boldsymbol{I} + \boldsymbol{B}) = 1 + \det \boldsymbol{B}(1 + \operatorname{tr} \boldsymbol{B}^{-1}) + \operatorname{tr} \boldsymbol{B}$$

and (25) and (26), we find that

$$\begin{aligned} (j')^{-1} &= \det((\boldsymbol{J}')^{-1}) \\ &= 1 - \frac{\boldsymbol{C}^{-1}\boldsymbol{k} \cdot \boldsymbol{l}}{1 + \boldsymbol{C}^{-1}\boldsymbol{k} \cdot \boldsymbol{l}} \\ &\le \frac{1}{1 - \sqrt{2}\tau_K} \end{aligned}$$

so that

$$\|(j')^{-1}\|_\infty \le \frac{1}{1 - \sqrt{2}\tau_K}. \tag{28}$$

Thus, roughly, it is seen that the Jacobians of $\tilde{K}$ and K differ by a term that is $O(\tau_K)$.

Finally, a transformation similar to that in (17) may be obtained by defining on $\tilde{K}$ a coordinate system $\tilde{\boldsymbol{x}} = (\tilde{x}, \tilde{y})$ and by defining the gradient $\tilde{\nabla}$ and strain $\tilde{\boldsymbol{\epsilon}}$ relative to this coordinate system by

$$(\tilde{\nabla}\tilde{\boldsymbol{v}})_{ij} = \frac{\partial \tilde{v}_i}{\partial \tilde{x}_j}, \quad \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}) = \frac{1}{2}(\tilde{\nabla}\tilde{\boldsymbol{v}} + [\tilde{\nabla}\tilde{\boldsymbol{v}}]^T). \tag{29}$$

For convenience we set

$$\tilde{\boldsymbol{L}}(\tilde{\boldsymbol{v}}) = \tilde{\nabla}\tilde{\boldsymbol{v}};$$

then we have

$$\tilde{\boldsymbol{L}}(\tilde{\boldsymbol{v}}) = \boldsymbol{L}(\boldsymbol{v})\tilde{\boldsymbol{J}}. \tag{30}$$

4. Affine finite element approximations. Since the equivalent parallelogram associated with a quadrilateral is "close" to that quadrilateral, it is natural to enquire whether approximate solutions to Problem S of the desired degree of accuracy might be obtained if Problem $\mathrm{S^h}$ were modified by replacing the integral in (18) with an integral over $\tilde{K}$ for each element K in the mesh. Such a procedure would have the advantage that, since the map from K to $\tilde{K}$ is affine, the associated Jacobian matrix and determinant are constant, and the integrals can be evaluated exactly, for homogeneous materials at least, for which case $\mathbb{C}$ is constant.

We now show that such a procedure does in fact lead to finite element approximations that converge at the usual rate. Numerical results presented by Küssner and Reddy [9] in the case of linear elasticity and by Reese and coworkers [12, 13] for problems involving nonlinear elasticity and finite deformations show in addition that, when this concept is applied to enhanced assumed strain formulations, the results represent in many cases an improvement over those obtained by the conventional approach.

From (25), (29), and (30), we have

$$\begin{aligned}\boldsymbol{\epsilon}(\boldsymbol{v}) &= \tfrac{1}{2}(\boldsymbol{L}(\boldsymbol{v}) + \boldsymbol{L}^T(\boldsymbol{v}))\\ &= \tfrac{1}{2}(\tilde{\boldsymbol{L}}(\boldsymbol{J}')^{-1} + (\boldsymbol{J}')^{-T}\tilde{\boldsymbol{L}}^T)\\ &= \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}) - \Delta(\tilde{\boldsymbol{v}}),\end{aligned} \tag{31}$$

in which

$$\Delta(\tilde{\boldsymbol{v}}) := \tfrac{1}{2}[\tilde{\boldsymbol{L}}\boldsymbol{R} + \boldsymbol{R}^T\tilde{\boldsymbol{L}}^T], \tag{32}$$

and $\boldsymbol{R}$ is defined by (26).

Next, for continuous functions $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ on $\tilde{K}$ we define the bilinear form $a_{\tilde{K}}(\cdot,\cdot)$ and linear functional $\ell_{\tilde{K}}(\cdot)$ by

$$a_{\tilde{K}}(\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}}) = \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{u}})] : \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}})\, d\tilde{x}d\tilde{y}, \quad \ell_{\tilde{K}}(\tilde{\boldsymbol{v}}) = \int_{\tilde{K}} \boldsymbol{b}\cdot\tilde{\boldsymbol{v}}\, d\tilde{x}d\tilde{y}. \tag{33}$$

Here, and henceforth, $\tilde{\mathbb{C}}(\tilde{\boldsymbol{x}}) := \mathbb{C}(G(\tilde{\boldsymbol{x}}))$, and we write $d\tilde{x}d\tilde{y}$ for dx_1dx_2 for convenience. Likewise, for functions $\boldsymbol{u}$, $\boldsymbol{v} \in V$ we set

$$\tilde{a}(\boldsymbol{u},\boldsymbol{v}) = \sum_{K\in\mathcal{T}} a_{\tilde{K}}(\tilde{\boldsymbol{u}},\tilde{\boldsymbol{v}}), \quad \tilde{\ell}(\boldsymbol{v}) = \sum_{K\in\mathcal{T}} \ell_{\tilde{K}}(\tilde{\boldsymbol{v}}), \tag{34}$$

where it is to be understood that $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ are the maps to $\tilde{K}$ of the restrictions $\boldsymbol{u}|_K$ and $\boldsymbol{v}|_K$. In particular, in what follows, for any function $\boldsymbol{v}_h \in V^h$ we set

$$\tilde{\boldsymbol{v}}_h := \boldsymbol{v}_h|_K \circ G^{-1}.$$

We are now in a position to define the affine-approximate problem.

Problem $\tilde{\mathrm{S}}^{\mathrm{h}}$. Given $\boldsymbol{b} \in V'$, find $\boldsymbol{w}_h \in V^h$ which satisfies

$$\tilde{a}(\boldsymbol{w}_h,\boldsymbol{v}_h) = \tilde{\ell}(\boldsymbol{v}_h) \tag{35}$$

for all $\boldsymbol{v}_h \in V^h$.

Remark. It is important to note that Problem $\tilde{\mathrm{S}}^{\mathrm{h}}$ is defined not by *transforming* the integrals in (18) onto $\tilde{K}$ but rather by *replacing* the integrals over K with those over $\tilde{K}$. Thus the relevant functions are mapped from K to $\tilde{K}$, after which the appropriate gradients are evaluated on $\tilde{K}$. So, for example, the integrals in the definitions of $a_{\tilde{K}}$ and $\ell_{\tilde{K}}$ contain no Jacobian determinants.

LEMMA 1. *The bilinear form $\tilde{a}$ is V^h-elliptic; that is, there exists a constant $\tilde{\alpha} > 0$, independent of τ, such that*

$$\tilde{a}(\boldsymbol{v}_h, \boldsymbol{v}_h) \geq \tilde{\alpha}(1 + \tau + \tau^2)\|\boldsymbol{v}_h\|_V^2 \quad \textit{for all } \boldsymbol{v}_h \in V^h, \tag{36}$$

where τ is defined by (22).

Proof. From the V^h-ellipticity of $a(\cdot,\cdot)$ we have

$$\alpha \,\|\boldsymbol{v}_h\|_V^2 \leq a(\boldsymbol{v}_h, \boldsymbol{v}_h) = \sum_{K \in \mathcal{T}} a_K(\boldsymbol{v}_h, \boldsymbol{v}_h), \tag{37}$$

where $a_K(\cdot,\cdot)$ is given by

$$a_K(\boldsymbol{u}, \boldsymbol{v}) := \int_K \mathbb{C}[\boldsymbol{\epsilon}(\boldsymbol{u})] : \boldsymbol{\epsilon}(\boldsymbol{v}) \; dxdy. \tag{38}$$

Using (31) we have

$$a_K(\boldsymbol{v}_h, \boldsymbol{v}_h) = \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h) - \Delta(\tilde{\boldsymbol{v}}_h)] : (\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h) - \Delta(\tilde{\boldsymbol{v}}_h))\, j' \; d\tilde{x}d\tilde{y}. \tag{39}$$

We now examine each of the terms on the right-hand side of (39) in turn.

We set

$$(\boldsymbol{E}, \boldsymbol{F})_{\mathbb{C},\tilde{K}} := \int_{\tilde{K}} \tilde{\mathbb{C}}[\boldsymbol{E}] : \boldsymbol{F} \; d\tilde{x}d\tilde{y}$$

for any $\boldsymbol{E}$, $\boldsymbol{F} \in \mathbb{L}^2(\tilde{K})$, and we note that this is an inner product on $\mathbb{L}^2(\tilde{K})$, equivalent to the standard inner product, as a result of the properties (6) and (7) of $\mathbb{C}$. We will also make use of the symmetry property $\mathbb{C}[\boldsymbol{E}] : \boldsymbol{F} = \mathbb{C}[\boldsymbol{F}] : \boldsymbol{E}$, which follows from (5).

Now, setting $\tilde{\boldsymbol{L}}\boldsymbol{R} := \boldsymbol{A}$ in the definition (32) of Δ, consider the expression

$$\begin{aligned}
(\mathrm{I}) &:= \int_{\tilde{K}} \tilde{\mathbb{C}}[\Delta] : \Delta \, j' \, d\tilde{x}d\tilde{y} \\
&\leq \|j'\|_\infty (\Delta, \Delta)_{\mathbb{C},\tilde{K}} \\
&= \tfrac{1}{4}\|j'\|_\infty (\boldsymbol{A} + \boldsymbol{A}^T, \boldsymbol{A} + \boldsymbol{A}^T)_{\mathbb{C},\tilde{K}} \\
&= \tfrac{1}{2}\|j'\|_\infty \left(\|\boldsymbol{A}\|_{\mathbb{C},\tilde{K}}^2 + (\boldsymbol{A}, \boldsymbol{A}^T)_{\mathbb{C},\tilde{K}}\right) \\
&\leq \|j'\|_\infty \|\boldsymbol{A}\|_{\mathbb{C},\tilde{K}}^2.
\end{aligned} \tag{40}$$

Next, from the definition of $\boldsymbol{A}$ we have

$$\|\boldsymbol{A}\|_{\mathbb{C},\tilde{K}} \leq \|\tilde{\boldsymbol{L}}\|_{\mathbb{C},\tilde{K}} \|\boldsymbol{R}\|_{\mathbb{C},\tilde{K}}. \tag{41}$$

Furthermore, we have, using the definition of the norm $\|\cdot\|_{\mathbb{C},\tilde{K}}$, (8), and (27),

$$\begin{aligned}\|\boldsymbol{R}\|^2_{\mathbb{C},\tilde{K}} &\le c_\infty \mu_K \max_{\tilde{\boldsymbol{x}}\in\tilde{K}} |\boldsymbol{R}|^2(\tilde{\boldsymbol{x}})\\ &\le \frac{2c_\infty \mu_K \tau_K^2}{(1-\sqrt{2}\,\tau_K)^2},\end{aligned} \tag{42}$$

in which μ_K is the area of K (and of $\tilde{K}$).

We also have

$$\begin{aligned}\|\tilde{\boldsymbol{L}}\|^2_{\mathbb{C},\tilde{K}} &= \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{L}}] : \tilde{\boldsymbol{L}}\; d\tilde{x}d\tilde{y}\\ &= \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}] : \tilde{\boldsymbol{\epsilon}}\; d\tilde{x}d\tilde{y}\\ &= a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{v}}_h).\end{aligned} \tag{43}$$

From (41)–(43), it follows that

$$(\mathrm{I}) \le c_K^2\, a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{v}}_h), \tag{44}$$

where the constant c_K is defined by

$$c_K = \tau_K (1-\sqrt{2}\,\tau_K)^{-1} \sqrt{2c_\infty \mu_K (1+\sqrt{2}\,\tau_K)}. \tag{45}$$

Next, consider the expression

$$\begin{aligned}(\mathrm{II}) &:= \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h) j'\; d\tilde{x}d\tilde{y}\\ &\le (1+\sqrt{2}\,\tau_K) \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}] : \tilde{\boldsymbol{\epsilon}}\; d\tilde{x}d\tilde{y}\\ &= (1+\sqrt{2}\,\tau_K) a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{v}}_h),\end{aligned} \tag{46}$$

where we have used the positivity of the integrand and (23).

Finally, consider the term

$$\begin{aligned}(\mathrm{III}) &:= -\int_{\tilde{K}} \Big[\tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \Delta(\tilde{\boldsymbol{v}}_h) + \tilde{\mathbb{C}}[\Delta(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)\Big]\, j'\; d\tilde{x}d\tilde{y}\\ &= -2\int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \Delta(\tilde{\boldsymbol{v}}_h) j'\; d\tilde{x}d\tilde{y}\\ &= -2(j'\tilde{\boldsymbol{\epsilon}}, \Delta)_{\mathbb{C},\tilde{K}}\\ &\le 2\|j'\|_\infty \left|(\tilde{\boldsymbol{\epsilon}}, \Delta)_{\mathbb{C},\tilde{K}}\right|\\ &= 2\|j'\|_\infty \left|(\tilde{\boldsymbol{\epsilon}}, \boldsymbol{A})_{\mathbb{C},\tilde{K}}\right|\\ &\le 2\|j'\|_\infty \|\tilde{\boldsymbol{\epsilon}}\|_{\mathbb{C},\tilde{K}} \|\boldsymbol{A}\|_{\mathbb{C},\tilde{K}},\end{aligned}$$

where we have used the representation $\Delta = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T)$ and the Cauchy–Schwarz inequality. Now we know that $\|\tilde{\boldsymbol{\epsilon}}\|_{\mathbb{C},\tilde{K}} = \sqrt{a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{v}}_h)}$ and, from (41)–(43),

$$\|\boldsymbol{A}\|^2_{\mathbb{C},\tilde{K}} \le c_K^2 (1+\sqrt{2}\,\tau_K)^{-1} a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{v}}_h).$$

So we have

$$\text{(III)} \leq 2c_K \sqrt{1+\sqrt{2}\,\tau_K}\, a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{v}}_h). \tag{47}$$

Expanding (39) and substituting the estimates for (I), (II), and (III), we have

$$\begin{aligned}
\alpha \|\boldsymbol{v}_h\|_V^2 &\leq a(\boldsymbol{v}_h, \boldsymbol{v}_h) = \sum_{K\in\mathcal{T}} a_K(\boldsymbol{v}_h, \boldsymbol{v}_h) \\
&\leq \sum_{K\in\mathcal{T}} \text{(I)} + \text{(II)} + \text{(III)} \\
&\leq \sum_{K\in\mathcal{T}} \left[c_K + \sqrt{1+\sqrt{2}\,\tau_K}\right]^2 a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{v}}_h) \\
&\leq \gamma\, \tilde{a}(\boldsymbol{v}_h, \boldsymbol{v}_h),
\end{aligned}$$

where the constant γ is given by

$$\gamma = [1 + \sqrt{2c_\infty}\,(1-\sqrt{2}\,\tau_K)^{-1} h\tau]^2 (1+\sqrt{2}\,\tau).$$

It follows that $\tilde{a}(\cdot,\cdot)$ is V^h-elliptic. $\square$

Since $\tilde{a}$ and $\tilde{\ell}$ are clearly continuous, it follows therefore that Problem $\tilde{\mathrm{S}}^h$ has a unique solution $\boldsymbol{w}_h$.

Remark. The consistency and convergence of the solution to Problem S^{h} depends in a fundamental way on the following lemma, which is analogous to the first Strang lemma (see [7, Theorem 4.4.1]) associated with errors induced by numerical quadrature in finite element approximations.

LEMMA 2. *There exists a positive constant C, independent of h, such that*

$$\begin{aligned}
\|\boldsymbol{u} - \boldsymbol{w}_h\|_V \leq C \Bigg(&\inf_{\boldsymbol{v}_h \in V^h} \left\{ \|\boldsymbol{u} - \boldsymbol{v}_h\|_V + \sup_{\boldsymbol{z}_h \in V^h} \frac{|a(\boldsymbol{v}_h, \boldsymbol{z}_h) - \tilde{a}(\boldsymbol{v}_h, \boldsymbol{z}_h)|}{\|\boldsymbol{z}_h\|_V} \right\} \\
&+ \sup_{\boldsymbol{z}_h \in V^h} \frac{|\ell(\boldsymbol{z}_h) - \tilde{\ell}(\boldsymbol{z}_h)|}{\|\boldsymbol{z}_h\|_V} \Bigg).
\end{aligned} \tag{48}$$

Proof. The proof follows that of the Strang lemma very closely, and we therefore merely sketch the details. Using the V^h-ellipticity of $\tilde{a}$, we have

$$\begin{aligned}
\tilde{\alpha} \|\boldsymbol{w}_h - \boldsymbol{v}_h\|_V^2 &\leq \tilde{a}(\boldsymbol{w}_h - \boldsymbol{v}_h, \boldsymbol{w}_h - \boldsymbol{v}_h) \\
&= a(\boldsymbol{u} - \boldsymbol{v}_h, \boldsymbol{w}_h - \boldsymbol{v}_h) + [a(\boldsymbol{v}_h, \boldsymbol{w}_h - \boldsymbol{v}_h) - \tilde{a}(\boldsymbol{v}_h, \boldsymbol{w}_h - \boldsymbol{v}_h)] \\
&\quad + [\tilde{\ell}(\boldsymbol{w}_h - \boldsymbol{v}_h) - \ell(\boldsymbol{w}_h - \boldsymbol{v}_h)].
\end{aligned} \tag{49}$$

Here we have used (11) and (35). Now

$$\frac{a(\boldsymbol{v}_h, \boldsymbol{w}_h - \boldsymbol{v}_h) - \tilde{a}(\boldsymbol{v}_h, \boldsymbol{w}_h - \boldsymbol{v}_h)}{\|\boldsymbol{w}_h - \boldsymbol{v}_h\|_V} \leq \sup_{\boldsymbol{z}_h \in V^h} \frac{|a(\boldsymbol{v}_h, \boldsymbol{z}_h) - \tilde{a}(\boldsymbol{v}_h, \boldsymbol{z}_h)|}{\|\boldsymbol{z}_h\|_V}, \tag{50}$$

and a similar inequality exists for $\ell - \tilde{\ell}$. The triangle inequality gives

$$\|\boldsymbol{u} - \boldsymbol{w}_h\|_V \leq \|\boldsymbol{u} - \boldsymbol{v}_h\|_V + \|\boldsymbol{w}_h - \boldsymbol{v}_h\|_V. \tag{51}$$

We divide throughout in (49) by $\|\boldsymbol{w}_h - \boldsymbol{v}_h\|_V$ and make use of the continuity of a in the first term on the right-hand side of (49); next, we use (50) and the corresponding expression for $|\ell - \tilde{\ell}|$ and take the supremum of the terms in square brackets in

(49). Finally, (48) is obtained by using (51) and by taking the infimum over all $\boldsymbol{v}_h \in V^h$. ∎

Next, we address the task of estimating the expression

$$\sup_{\boldsymbol{z}_h \in V^h} \frac{|a(\boldsymbol{v}_h, \boldsymbol{z}_h) - \tilde{a}(\boldsymbol{v}_h, \boldsymbol{z}_h)|}{\|\boldsymbol{z}_h\|_V}.$$

From (33) and (39) we have

$$\begin{aligned}
&a_K(\boldsymbol{v}_h, \boldsymbol{z}_h) - a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{z}}_h)\\
&\quad = \int_K \mathbb{C}[\boldsymbol{\epsilon}(\boldsymbol{v}_h)] : \boldsymbol{\epsilon}(\boldsymbol{z}_h)\ dxdy - \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{z}}_h)\ d\tilde{x}d\tilde{y}\\
&\quad = \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{z}}_h) j'\ d\tilde{x}d\tilde{y} - \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \Delta(\tilde{\boldsymbol{z}}_h)\, j'\ d\tilde{x}d\tilde{y}\\
&\qquad - \int_{\tilde{K}} \tilde{\mathbb{C}}[\Delta(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{z}}_h)\, j'\ d\tilde{x}d\tilde{y} + \int_{\tilde{K}} \tilde{\mathbb{C}}[\Delta(\tilde{\boldsymbol{v}}_h)] : \Delta(\tilde{\boldsymbol{z}}_h)\, j'\ d\tilde{x}d\tilde{y}\\
&\qquad - \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{z}}_h)\ d\tilde{x}d\tilde{y}\\
&\quad = \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{z}}_h)(j'-1)\ d\tilde{x}d\tilde{y} - \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \Delta(\tilde{\boldsymbol{z}}_h) j'\ d\tilde{x}d\tilde{y}\\
&\qquad - \int_{\tilde{K}} \tilde{\mathbb{C}}[\Delta(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{z}}_h) j'\ d\tilde{x}d\tilde{y} + \int_{\tilde{K}} \tilde{\mathbb{C}}[\Delta(\tilde{\boldsymbol{v}}_h)] : \Delta(\tilde{\boldsymbol{z}}_h) j'\ d\tilde{x}d\tilde{y}
\end{aligned}$$

$$\leq \sqrt{2}\,\tau_K\, a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{z}}_h) + \sqrt{a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{v}}_h) a_{\tilde{K}}(\tilde{\boldsymbol{z}}_h, \tilde{\boldsymbol{z}}_h)}\left[c_K^2 + 2c_K\sqrt{1+\sqrt{2}\,\tau_K}\right]. \tag{52}$$

Here we have used the estimates leading to (44) and (47).

Now

$$a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{z}}_h) \leq M_{\tilde{K}}\, \|\tilde{\boldsymbol{v}}_h\|_{H^1(\tilde{K})} \|\tilde{\boldsymbol{z}}_h\|_{H^1(\tilde{K})}$$

for some positive constant $M_{\tilde{K}}$, independent of h_K and τ_K, and, using (24), (28), and (30), we have

$$\begin{aligned}
\|\tilde{\boldsymbol{v}}_h\|^2_{H^1(\tilde{K})} &= \int_{\tilde{K}} \left(|\tilde{\boldsymbol{v}}_h|^2 + |\tilde{\nabla}\tilde{\boldsymbol{v}}_h|^2\right) d\tilde{x}d\tilde{y}\\
&= \int_K [|\boldsymbol{v}_h|^2 + |(\nabla \boldsymbol{v}_h)\boldsymbol{J}'|^2](j')^{-1}\, dxdy\\
&\leq C_K \|\boldsymbol{v}_h\|^2_{H^1(K)},
\end{aligned}$$

where $C_K = 2(1+\tau_K)^2(1-\sqrt{2}\,\tau_K)^{-1}$.

Hence

$$\begin{aligned}
&\frac{|a(\boldsymbol{v}_h, \boldsymbol{z}_h) - \tilde{a}(\boldsymbol{v}_h, \boldsymbol{z}_h)|}{\|\boldsymbol{z}_h\|_V}\\
&\quad \leq \frac{\sum_{K\in\mathcal{T}} M_{\tilde{K}} C_K \left[\left(c_K + \sqrt{1+\sqrt{2}\,\tau_K}\right)^2 - 1\right] \|\boldsymbol{v}_h\|_{H^1(K)} \|\boldsymbol{z}_h\|_{H^1(K)}}{\|\boldsymbol{z}_h\|_V}\\
&\quad \leq C\tau \frac{\sum_{K\in\mathcal{T}} \|\boldsymbol{v}_h\|_{H^1(K)} \|\boldsymbol{z}_h\|_{H^1(K)}}{\|\boldsymbol{z}_h\|_V}
\end{aligned}$$

$$\leq C\tau \|\boldsymbol{v}_h\|_V, \tag{53}$$

where the constant C is independent of h and τ.

Choosing $\boldsymbol{v}_h = \Pi_h \boldsymbol{u}$, where Π_h is the interpolation operator onto V^h, and noting that $\|\Pi_h \boldsymbol{u}\|_V \le \|\boldsymbol{u}\|_V + \|\boldsymbol{u} - \Pi_h \boldsymbol{u}\|_V \le ch|\boldsymbol{u}|_{H^2} + \|\boldsymbol{u}\|_V$, we finally obtain

$$\inf_{\boldsymbol{v}_h \in V^h} \sup_{\boldsymbol{z}_h \in V^h} \frac{|a(\boldsymbol{v}_h, \boldsymbol{z}_h) - \tilde{a}(\boldsymbol{v}_h, \boldsymbol{z}_h)|}{\|\boldsymbol{z}_h\|} \le C\tau,$$

where the constant C depends on the geometry and on the solution $\boldsymbol{u}$. In exactly the same way we can derive the estimate

$$\inf_{\boldsymbol{v}_h \in V^h} \sup_{\boldsymbol{z}_h \in V^h} \frac{|\ell(\boldsymbol{z}_h) - \tilde{\ell}(\boldsymbol{z}_h)|}{\|\boldsymbol{z}_h\|} \le C\tau,$$

where again C depends on the geometry and on $\boldsymbol{u}$. We therefore have the following result.

THEOREM 3. *Let $\mathcal{T}$ be an h-regular finite element mesh of quadrilaterals, with the maximum distortion of quadrilaterals being bounded in the sense that $\tau \le ch$ for some constant c, independent of h, as $h \to 0$. Then problem $\tilde{\mathrm{S}}^h$ has a unique solution $\boldsymbol{w}_h$ which satisfies*

$$\|\boldsymbol{u} - \boldsymbol{w}_h\|_V \le Ch,$$

the constant C depending on the geometry and the solution $\boldsymbol{u}$ to the continuous problem, but not on h.

5. The enhanced strain problem. In the context of the finite element method, the enhanced strain method refers to an approach proposed by Simo and Rifai [14], in which the discrete strain $\boldsymbol{\epsilon}_h$ takes the form

$$\boldsymbol{\epsilon}_h = \boldsymbol{\epsilon}(\boldsymbol{u}_h) + \boldsymbol{\eta}_h, \tag{54}$$

the first term on the right-hand side being evaluated as in (2), while the second term on the right-hand side is the enhanced strain, which is required to have the property $\boldsymbol{\eta}_h \to \mathbf{0}$ as $h \to 0$.

In order to formulate the problem in weak form it is necessary to add to the spaces already defined the space Γ^h of enhanced strains, which is defined by

$$\Gamma^h := \left\{ \boldsymbol{\gamma} = (\gamma_{ij}) : \ \gamma_{ij} \in L^2(\Omega), \ \gamma_{ji} = \gamma_{ij}, \ \int_K \mathbb{C}\boldsymbol{\gamma}|_K \ dxdy = \mathbf{0} \text{ for all } K \in \mathcal{T} \right\}. \tag{55}$$

In practice Γ^h will comprise functions of the form $\boldsymbol{\gamma} = j^{-1}\hat{\boldsymbol{\gamma}}$ on each element, in which the components of $\hat{\boldsymbol{\gamma}}$ are simple polynomials defined on the reference element $\hat{K}$. A consequence of this definition is that $\int_{\hat{K}} \mathbb{C}\hat{\boldsymbol{\gamma}} \ d\xi d\eta = \mathbf{0}$ [3]. Concrete examples of bases for Γ^h, together with applications, may be found in [14, 1].

We set $\boldsymbol{\phi}_h = (\boldsymbol{u}_h, \boldsymbol{\eta}_h)$ and $\boldsymbol{\psi}_h = (\boldsymbol{v}_h, \boldsymbol{\gamma}_h)$ for $\boldsymbol{u}_h$, $\boldsymbol{v}_h \in V^h$ and $\boldsymbol{\eta}_h$, $\boldsymbol{\gamma}_h \in \Gamma^h$. Also, we define the product space

$$\Psi^h := V^h \times \Gamma^h,$$

which is a Hilbert space with the natural norm

$$\|\boldsymbol{\psi}_h\|_\Psi := \left(\|\boldsymbol{v}_h\|_V^2 + \|\boldsymbol{\gamma}_h\|_{L^2}^2 \right)^{1/2}.$$

The bilinear form $A : \Psi^h \times \Psi^h \longrightarrow \mathbb{R}$ is defined by

$$A(\boldsymbol{\phi}_h, \boldsymbol{\psi}_h) = \int_\Omega \mathbb{C}(\boldsymbol{\epsilon}(\boldsymbol{u}_h) + \boldsymbol{\eta}_h) : (\boldsymbol{\epsilon}(\boldsymbol{v}_h) + \boldsymbol{\gamma}_h)\, dxdy, \tag{56}$$

and we recall the definition (10) of the linear functional ℓ.

The weak formulation of the problem then takes the following form [11, 14].

Problem $\mathrm{E^h}$. Find $(\boldsymbol{u}_h, \boldsymbol{\eta}_h) \in V^h \times \Gamma^h$ such that

$$A(\boldsymbol{\phi}_h, \boldsymbol{\psi}_h) = \ell(\boldsymbol{v}_h) \quad \text{for all } \boldsymbol{\psi}_h \in \Psi^h. \tag{57}$$

We have the following result, proved in [11] for affine-equivalent meshes (see also [5]) and in [3] for isoparametric meshes and stated here for the special case in which V^h is chosen as in (15).

THEOREM 4. *Let $\mathcal{T}$ be a regular mesh of quadrilaterals on a bounded polygonal domain $\Omega \in \mathbb{R}^2$. Let the space V^h be defined by (15) and the space Γ^h by (55). Assume, in addition, that*

(a) $\boldsymbol{\epsilon}(V^h) \cap \Gamma^h = \{\mathbf{0}\}$,

(b) *there exists a constant c_1 with $0 < c_1 < 1$ such that, for any $\boldsymbol{\gamma}_h \in \Gamma^h$, $\|P\boldsymbol{\gamma}_h\|_\Gamma \le c_1 \|\boldsymbol{\gamma}_h\|$, where P is the L^2-orthogonal projection onto $\boldsymbol{\epsilon}(V^h)$.*

Then there exists a unique solution to Problem $\mathrm{E^h}$. *Furthermore, if $\boldsymbol{u} \in [H^2(\Omega)]^2$, then there exists a constant $C > 0$, independent of h, such that*

$$\|\boldsymbol{u} - \boldsymbol{u}_h\|_V + \|\boldsymbol{\eta}_h\|_\Gamma \le Ch|\boldsymbol{u}|_{H^2}.$$

We now consider the affine-approximate problem analogous to Problem $\tilde{\mathrm{S}}^h$.

Define

$$A_{\tilde{K}}(\tilde{\boldsymbol{\chi}}_h, \tilde{\boldsymbol{\psi}}_h) = \int_{\tilde{K}} \tilde{\mathbb{C}}(\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{w}}_h) + \tilde{\boldsymbol{\beta}}_h) : (\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h) + \tilde{\boldsymbol{\gamma}}_h)\, d\tilde{x}d\tilde{y} \tag{58}$$

and

$$\ell_{\tilde{K}}(\tilde{\boldsymbol{\psi}}_h) = \int_{\tilde{K}} \boldsymbol{b} \cdot \tilde{\boldsymbol{v}}_h \, d\tilde{x}d\tilde{y}, \tag{59}$$

where $\boldsymbol{\chi}_h = (\boldsymbol{w}_h, \boldsymbol{\beta}_h)$ and superposed tildes have the same interpretation as previously. Set

$$\tilde{A}(\boldsymbol{\chi}_h, \boldsymbol{\psi}_h) = \sum_{K \in \mathcal{T}} A_{\tilde{K}}(\tilde{\boldsymbol{\chi}}_h, \tilde{\boldsymbol{\psi}}_h) \tag{60}$$

and

$$\tilde{\ell}(\boldsymbol{\psi}_h) = \sum_{K \in \mathcal{T}} \ell_{\tilde{K}}(\tilde{\boldsymbol{\psi}}_h). \tag{61}$$

Problem $\tilde{\mathrm{E}}^{\mathrm{h}}$. Given $\boldsymbol{b} \in V'$, find $\boldsymbol{\chi}_h := (\boldsymbol{w}_h, \boldsymbol{\beta}_h) \in V^h \times \Psi^h$ which satisfy

$$\tilde{A}(\boldsymbol{\chi}_h, \boldsymbol{\psi}_h) = \tilde{\ell}(\boldsymbol{\psi}_h) \tag{62}$$

for all $\boldsymbol{\psi}_h \in \Psi^h$.

We need to show that $\tilde{A}$ is Ψ^h-elliptic. We proceed as in the case of Lemma 1, beginning with the observation that $A(\cdot,\cdot)$ is Ψ^h-elliptic (see [11]), from which it follows that

$$\begin{aligned} \alpha_A \|\boldsymbol{\psi}_h\|_\Psi^2 &\le A(\boldsymbol{\psi}_h, \boldsymbol{\psi}_h) \\ &= \sum_{K\in\mathcal{T}} A_K(\boldsymbol{\psi}_h, \boldsymbol{\psi}_h). \end{aligned} \tag{63}$$

Now we have, from the positivity of the integrand of A_K and the symmetry of $\mathbb{C}$,

$$\begin{aligned} A_K(\boldsymbol{\psi}_h,\boldsymbol{\psi}_h) &= \int_K \mathbb{C}(\boldsymbol{\epsilon}(\boldsymbol{v}_h)+\boldsymbol{\gamma}_h) : (\boldsymbol{\epsilon}(\boldsymbol{v}_h)+\boldsymbol{\gamma}_h)\ dxdy \\ &= \int_{\tilde{K}} \mathbb{C}(\boldsymbol{\epsilon}(\boldsymbol{v}_h)+\boldsymbol{\gamma}_h) : (\boldsymbol{\epsilon}(\boldsymbol{v}_h)+\boldsymbol{\gamma}_h)\ j'\, d\tilde{x}d\tilde{y} \\ &\le \|j'\|_\infty \int_{\tilde{K}} \mathbb{C}(\boldsymbol{\epsilon}(\boldsymbol{v}_h)+\boldsymbol{\gamma}_h) : (\boldsymbol{\epsilon}(\boldsymbol{v}_h)+\boldsymbol{\gamma}_h)\ d\tilde{x}d\tilde{y} \\ &= \|j'\|_\infty \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h) - \Delta(\tilde{\boldsymbol{v}}_h) + \tilde{\boldsymbol{\gamma}}_h) : (\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h) - \Delta(\tilde{\boldsymbol{v}}_h) + \tilde{\boldsymbol{\gamma}}_h)\ d\tilde{x}d\tilde{y} \\ &= \|j'\|_\infty \Bigg[A_{\tilde{K}}(\tilde{\boldsymbol{\psi}}_h, \tilde{\boldsymbol{\psi}}_h) \underbrace{-2\int_{\tilde{K}} \tilde{\mathbb{C}}[\Delta(\tilde{\boldsymbol{v}}_h)] : [\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h) - \tfrac{1}{2}\Delta(\tilde{\boldsymbol{v}}_h) + \tilde{\boldsymbol{\gamma}}_h]\ d\tilde{x}d\tilde{y}}_{\mathcal{F}} \\ &\qquad + 2\int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\gamma}}_h] : (\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h) + \tfrac{1}{2}\tilde{\boldsymbol{\gamma}}_h)\ d\tilde{x}d\tilde{y} \Bigg] \\ &\le \|j'\|_\infty \Big[A_{\tilde{K}}(\tilde{\boldsymbol{\psi}}_h, \tilde{\boldsymbol{\psi}}_h) + 2(\tilde{\boldsymbol{\epsilon}}, \tilde{\boldsymbol{\gamma}})_{\mathbb{C},\tilde{K}} + (\tilde{\boldsymbol{\gamma}}, \tilde{\boldsymbol{\gamma}})_{\mathbb{C},\tilde{K}} + \mathcal{F}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{\gamma}}_h, \tau_K) \Big]. \end{aligned}$$

Here the term $\mathcal{F}$ is of the form $\mathcal{F} = c_K \cdot$[terms depending on $(\tilde{\boldsymbol{v}}_h,\ \tilde{\boldsymbol{\gamma}}_h,\ \tau_K)$], as can be deduced by a series of manipulations similar to those carried out in (44)–(46). From the definition of $A_{\tilde{K}}$ it follows therefore that

$$A_K(\boldsymbol{\psi}_h,\boldsymbol{\psi}_h) \le (1+\sqrt{2}\,\tau_K)[2A_{\tilde{K}}(\tilde{\boldsymbol{\psi}}_h, \tilde{\boldsymbol{\psi}}_h) + \mathcal{F}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{\gamma}}_h, \tau_K)].$$

We therefore have the following result.

LEMMA 5. *The bilinear form $\tilde{A}$ is Ψ^h-elliptic for sufficiently small τ. Furthermore, Problem $\tilde{\mathrm{E}}^h$ has a unique solution $\boldsymbol{\chi}_h = (\boldsymbol{w}_h, \boldsymbol{\beta}_h)$ in Ψ^h.*

Next, we have the following counterpart to Lemma 2.

LEMMA 6. *Set $\boldsymbol{\phi} = (\boldsymbol{u}, \mathbf{0})$, where $\boldsymbol{u}$ is the solution to Problem* S, *and denote the solution to Problem $\tilde{\mathrm{E}}^{\mathrm{h}}$ by $\boldsymbol{\chi}_h = (\boldsymbol{w}_h, \boldsymbol{\beta}_h)$. Then there exists a constant C, independent of h, such that*

$$\|\boldsymbol{\phi} - \boldsymbol{\chi}_h\|_\Psi \le C\left[\inf_{\boldsymbol{\psi}_h\in\Psi^h}\left\{\|\boldsymbol{\phi}-\boldsymbol{\psi}_h\|_\Psi + \sup_{\boldsymbol{\omega}_h=(\boldsymbol{z}_h,\boldsymbol{\rho}_h)\in\Psi^h} \frac{\diamond}{\|\boldsymbol{\omega}_h\|_\Psi}\right\}\right], \tag{64}$$

where $\diamond$ is given by

$$\diamond = |A(\boldsymbol{\psi}_h,\boldsymbol{\omega}_h) - \tilde{A}(\boldsymbol{\psi}_h,\boldsymbol{\omega}_h)| + |\ell(\boldsymbol{\psi}_h) - \tilde{\ell}(\boldsymbol{\psi}_h)| + \left|\int_\Omega \mathbb{C}\boldsymbol{\epsilon}(\boldsymbol{u}) : \boldsymbol{\gamma}_h\ dxdy\right|. \tag{65}$$

Proof. From the Ψ^h-ellipticity of $\tilde{A}$, we have

$$\begin{aligned}
\tilde{\alpha}\|\boldsymbol{\chi}_h - \boldsymbol{\psi}_h\|_\Psi^2 &\leq \tilde{A}(\boldsymbol{\chi}_h - \boldsymbol{\psi}_h, \boldsymbol{\chi}_h - \boldsymbol{\psi}_h) \\
&= A(\boldsymbol{\phi} - \boldsymbol{\psi}_h, \boldsymbol{\chi}_h - \boldsymbol{\psi}_h) + [A(\boldsymbol{\psi}_h, \boldsymbol{\chi}_h - \boldsymbol{\psi}_h) - \tilde{A}(\boldsymbol{\psi}_h, \boldsymbol{\chi}_h - \boldsymbol{\psi}_h)] \\
&\quad - A(\boldsymbol{\phi}, \boldsymbol{\chi}_h - \boldsymbol{\psi}_h) + \tilde{A}(\boldsymbol{\chi}_h, \boldsymbol{\chi}_h - \boldsymbol{\psi}_h) \\
&= A(\boldsymbol{\phi} - \boldsymbol{\psi}_h, \boldsymbol{\chi}_h - \boldsymbol{\psi}_h) + [A(\boldsymbol{\psi}_h, \boldsymbol{\chi}_h - \boldsymbol{\psi}_h) - \tilde{A}(\boldsymbol{\psi}_h, \boldsymbol{\chi}_h - \boldsymbol{\psi}_h)] \\
&\quad + [\tilde{\ell}(\boldsymbol{\chi}_h - \boldsymbol{\psi}_h) - \ell(\boldsymbol{\chi}_h - \boldsymbol{\psi}_h)] + [\ell(\boldsymbol{\chi}_h - \boldsymbol{\psi}_h) - A(\boldsymbol{\phi}, \boldsymbol{\chi}_h - \boldsymbol{\psi}_h)].
\end{aligned}$$

The rest of the proof proceeds in much the same way as the proof of Lemma 2. ◻

Finally, the expression $\diamond$ in (64) must be estimated. We have, from (58) and (64),

$$\begin{aligned}
&A_K(\boldsymbol{\psi}_h, \boldsymbol{\omega}_h) - A_{\tilde{K}}(\tilde{\boldsymbol{\psi}}_h, \tilde{\boldsymbol{\omega}}_h) \\
&= a_K(\boldsymbol{v}_h, \boldsymbol{z}_h) - a_{\tilde{K}}(\tilde{\boldsymbol{v}}_h, \tilde{\boldsymbol{z}}_h) \quad \text{(i)} \\
&\quad + \int_K (\mathbb{C}[\boldsymbol{\epsilon}(\boldsymbol{v}_h)] : \boldsymbol{\rho}_h + \mathbb{C}[\boldsymbol{\epsilon}(\boldsymbol{z}_h)] : \boldsymbol{\gamma}_h)\; dxdy \\
&\quad - \int_{\tilde{K}} (\tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\rho}}_h - \tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{z}}_h)] : \tilde{\boldsymbol{\gamma}}_h)\; d\tilde{x}d\tilde{y} \quad \text{(ii)} \\
&\quad + \int_K \mathbb{C}[\boldsymbol{\gamma}_h] : \boldsymbol{\rho}_h\; dxdy - \int_{\tilde{K}} \tilde{\mathbb{C}}[\tilde{\boldsymbol{\gamma}}_h] : \tilde{\boldsymbol{\rho}}_h\; d\tilde{x}d\tilde{y}. \quad \text{(iii)}
\end{aligned} \tag{66}$$

We now examine the expressions (i)–(iii) in (66) in turn. First, we see that (i) is estimated in (52). Next, we have

$$\begin{aligned}
\text{(ii)} &\leq \int_{\tilde{K}} (j' - 1)\tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\rho}}_h\; d\tilde{x}d\tilde{y} + \int_{\tilde{K}} (j' - 1)\tilde{\mathbb{C}}[\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{z}}_h)] : \tilde{\boldsymbol{\gamma}}_h\; d\tilde{x}d\tilde{y} \\
&\quad - \int_{\tilde{K}} \tilde{\mathbb{C}}[\Delta(\tilde{\boldsymbol{v}}_h)] : \tilde{\boldsymbol{\rho}}_h\; \tilde{j}\; d\tilde{x}d\tilde{y} - \int_{\tilde{K}} \tilde{\mathbb{C}}[\Delta(\tilde{\boldsymbol{z}}_h)] : \tilde{\boldsymbol{\gamma}}_h\; \tilde{j}\; d\tilde{x}d\tilde{y} \\
&\leq \left[\sqrt{2}\tau_K + \sqrt{1 + \sqrt{2}\tau_K}\, c_K\right] \left\{\|\tilde{\boldsymbol{\rho}}_h\|_{\mathbb{C},\tilde{K}} \|\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{v}}_h)\|_{\mathbb{C},\tilde{K}} + \|\tilde{\boldsymbol{\gamma}}_h\|_{\mathbb{C},\tilde{K}} \|\tilde{\boldsymbol{\epsilon}}(\tilde{\boldsymbol{z}}_h)\|_{\mathbb{C},\tilde{K}}\right\}.
\end{aligned} \tag{67}$$

Here we have used (23) and the manipulations leading to (44). Finally,

$$\text{(iii)} \leq \sqrt{2}\tau_K \|\tilde{\boldsymbol{\rho}}_h\|_{\mathbb{C},\tilde{K}} \|\tilde{\boldsymbol{\gamma}}_h\|_{\mathbb{C},\tilde{K}}. \tag{68}$$

It is not difficult to see, from the arguments leading to (53) and from (67) and (68), that

$$\sup_{\boldsymbol{\omega}_h \in \Psi^h} \frac{A(\boldsymbol{\psi}_h, \boldsymbol{\omega}_h) - \tilde{A}(\tilde{\boldsymbol{\psi}}_h, \tilde{\boldsymbol{\omega}}_h)}{\|\boldsymbol{\omega}_h\|_\Psi} \leq C\tau, \tag{69}$$

in which the constant C depends on the geometry and on the solution $\boldsymbol{u}$ but not on h nor on τ. In the same way, one may derive an estimate of the form (69) for the second term in the definition (65) of $\diamond$.

Finally, the last term on the right-hand side of (65) is shown in [11] to be bounded, up to a constant, by $h|\boldsymbol{u}|_1$.

THEOREM 7. *Let $\mathcal{T}$ be an h-regular finite element mesh of quadrilaterals, with the maximum distortion of quadrilaterals being bounded in the sense that $\tau \leq ch$ for* *some constant c, independent of h, as* $h \to 0$. *Let* $\boldsymbol{\phi} = (\boldsymbol{u}, \mathbf{0}) \in \Psi$, *where* $\boldsymbol{u}$ *is the solution to Problem* S. *Then Problem* $\tilde{\mathrm{E}}^{\mathrm{h}}$ *has a unique solution* $\boldsymbol{\chi}_h$ *which satisfies*

$$\|\boldsymbol{\phi} - \boldsymbol{\chi}_h\|_\Psi \leq Ch,$$

the constant C depending on the geometry, on the material tensor $\mathbb{C}$, *and on* $\boldsymbol{u}$, *but not on h.*

Remark. The analysis presented here has been carried out for the case of compressible materials, for which the components of $\mathbb{C}$ are bounded. A modified approach, such as that presented in [11, section 5] or in [6], is required for the limiting cases of incompressibility or near incompressibility. Such an analysis would combine the approaches presented here and in those works.

REFERENCES

[1] U. Andelfinger and E. Ramm, *EAS-elements for two-dimensional, three-dimensional, plate and shell structures and their equivalence to HR-elements*, Internat. J. Numer. Methods Engrg., 36 (1993), pp. 1311–1337.

[2] K. Arunakirinathar and B. D. Reddy, *Some geometrical results and estimates for quadrilateral finite elements*, Comput. Methods Appl. Mech. Engrg., 122 (1995), pp. 307–314.

[3] K. Arunakirinathar and B. D. Reddy, *Further results for enhanced strain methods with isoparametric elements*, Comput. Methods Appl. Mech. Engrg., 127 (1995), pp. 127–143.

[4] T. Belytschko and W. Bachrach, *Efficient implementation of quadrilaterals with high coarse-mesh accuracy*, Comput. Methods Appl. Mech. Engrg., 54 (1986), pp. 279–301.

[5] D. Braess, *Enhanced assumed strain elements and locking in membrane problems*, Comput. Methods Appl. Mech. Engrg., 165 (1998), pp. 155–174.

[6] D. Braess, C. Carstensen, and B. D. Reddy, *Uniform convergence and a posteriori estimators for the enhanced strain finite element method*, submitted.

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

[8] D. P. Flanagan and T. Belytschko, *A uniform strain hexahedron and quadrilateral with orthogonal hourglass control*, Internat. J. Numer. Methods Engrg., 17 (1981), pp. 679–706.

[9] M. Küssner and B. D. Reddy, *The equivalent parallelogram and parallelipiped, and their application to stabilized finite elements in two and three dimensions*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 1967–1983.

[10] J. E. Marsden and T. J. R. Hughes, *Mathematical Foundations of Elasticity*, Prentice-Hall, Englewood Cliffs, NJ, 1983.

[11] B. D. Reddy and J. C. Simo, *Stability and convergence of a class of enhanced strain methods*, SIAM J. Numer. Anal., 32 (1995), pp. 1705–1728.

[12] S. Reese, M. Küssner, and B. D. Reddy, *A new stabilization technique for finite elements in non-linear elasticity*, Internat. J. Numer. Methods Engrg., 44 (1999), pp. 1617–1652.

[13] S. Reese, P. Wriggers, and B. D. Reddy, *A new locking-free brick element technique for large deformation problems in elasticity*, Comput. & Structures, 75 (2000), pp. 291–304.

[14] J. C. Simo and S. Rifai, *A class of mixed assumed strain methods and the method of incompatible modes*, Internat. J. Numer. Methods Engrg., 29 (1990), pp. 1595–1638.

[15] J. C. Simo, F. Armero, and R. L.Taylor, *Improved version of assumed enhanced strain tri-linear elements for 3D finite deformation problems*, Comput. Methods Appl. Mech. Engrg., 110 (1993), pp. 359–366.

[16] R. L. Taylor, P. J. Beresford, and E. L. Wilson, *A non-conforming element for stress analysis*, Internat. J. Numer. Methods Engrg., 10 (1976), pp. 1211–1220.

[17] E. L. Wilson, R. L. Taylor, W. P. Doherty, and J. Ghaboussi, *Incompatible displacement models*, in Numerical and Computer Models in Structural Mechanics, Academic Press, New York, 1973, pp. 43–57.