

## A REDUCED CONSTRAINT $hp$ FINITE ELEMENT METHOD FOR SHELL PROBLEMS

MANIL SURI

ABSTRACT. We propose and analyze an  $hp$  finite element method for the Nagdhi shell model, based on rectangular elements. We show that for the bending-dominated case, assuming sufficient smoothness on the solution, the method is locking free in terms of both  $h$  and  $p$ , as the thickness of the shell tends to zero. Our results are established under the assumption that the geometrical coefficients appearing in the model are piecewise polynomial functions.

### 1. INTRODUCTION

In recent years, several new commercial finite element codes with  $hp$  capabilities have been developed (e.g. MSC/PROBE, STRESSCHECK, POLYFEM (IBM), Applied Structure (Rasna Co.), PHLEX, etc.). Such codes allow the use of both mesh refinement ( $h$  version) and increase of polynomial degree ( $p$  version) to attain accuracy. Due to the increasing engineering interest in  $p$  and  $h$ - $p$  methods, an important question that arises is the search for robust, locking-free  $p$  and  $hp$  elements for *plate* and *shell* problems.

In this paper, we propose and analyze an  $hp$  method based on the Nagdhi shell model [9, 10]. In this model, the unknown displacement and rotation vectors are obtained as the minimizers of an internal energy function which combines the energy due to bending, transverse shear and membrane stresses. As the thickness  $d$  of the shell approaches zero, various constraints on the exact solution appear in the “bending-dominated” case. These ensure that the energy remains finite as  $d \rightarrow 0$ . Often, finite element spaces being used to discretize the model cannot handle such constraints while still retaining their approximation properties. This leads to the phenomenon of *locking*, which manifests itself as an unacceptably large discretization error, even for highly refined meshes, when  $d$  is small. See [4, 7] for general discussions of the locking problem. In the Nagdhi shell model, both transverse shear and membrane constraints appear as  $d \rightarrow 0$ , leading to *shear* locking and *membrane* locking in the *bending-dominated* case.

The two main strategies commonly used for overcoming locking are the use of the *standard variational form* with *high-order* elements and the use of *modified variational forms* (including mixed and reduced integration methods). In the former,

---

Received by the editor August 5, 1994 and, in revised form, February 9, 1996.

1991 *Mathematics Subject Classification*. Primary 65N30, 73K15, 73V05.

*Key words and phrases*. Locking, shell,  $hp$ , finite element.

Research partially supported by the Air Force Office of Scientific Research, Bolling AFB, DC, under Grant AFOSR F49620-92-J-0100.

which we call *exact constraint* methods, the idea is to choose finite element spaces that are enriched enough to satisfy the constraints *exactly*. This is usually done by using high-order  $h$  methods or using  $p/hp$  methods. In the latter, the effect of the constraint is weakened, by enforcing it only approximately. Often, low-order ( $h$  version) elements are used. We call such methods *reduced constraint* methods. The above approaches have been rigorously investigated for various problems (other than shells) involving locking (see, e.g. [14, 16] for exact constraint approaches and [8, 15] for reduced constraint approaches).

In the case of locking for shells, however, even though there are several schemes that have been proposed in the engineering literature, the convergence properties of most schemes have not been mathematically established. A notable exception is the work of Pitkäranta (e.g. [12]) in which locking for cylindrical shells (including a reduced constraint scheme proposed by Bathe and Dvorkin) has been analyzed. One of the conclusions that emerges from Pitkäranta's work is that high-order  $p$ -type schemes do particularly well in combatting locking for shells.

In this paper, we present and analyze an  $hp$  finite element method for shells which *combines* the two approaches of using high-order elements and weakening the constraints. Our element is based on a reduced constraint method recently proposed by Arnold and Bezzi [1] which uses a mixed formulation for the Nagdhi shell model with a stabilizing term that stabilizes it in the bending-dominated case. (This is the case for which locking occurs, the membrane-dominated case, leading to “pure membrane” deformation in the limit, is free of locking for smooth solutions.) Using this stabilizing mixed formulation, Arnold and Brezzi developed a family of mixed triangular elements for which they proved that the convergence of the displacement and rotation vectors does not deteriorate as  $d \rightarrow 0$  (i.e. the method is *locking-free* in the terminology of [4]). Their results were only in terms of the  $h$  version and did not address the question of  $p$ -convergence, when successively higher-order elements from the family are used. Here, we formulate *rectangular* elements based on the modified variational form in [1]. As in the case of triangular elements in [1], we propose a *family* of such elements, one for each polynomial degree  $p$ . We consider not only the stability and convergence of this family in terms of the mesh parameter  $h$ , but also the polynomial degree  $p$ , showing that the family is *robust* in terms of *both*  $h$  and  $p$  (unlike the corresponding triangular elements from [1], no additional bubble functions are needed for stability here). Our analysis therefore validates not only the  $h$  version, but also the  $p$  and  $hp$  implementation of these reduced constraint methods. It also validates a *spectral element* approach (see Remark 3.5).

In [1], Arnold and Brezzi used an assumption of the geometrical coefficients in the variational formulation being piecewise constant. As we show here, the case when the coefficients are piecewise polynomials of degree  $m$  can also be treated, provided we use polynomials of sufficiently higher degree for the primary variables (this is easily accomplished in the  $p/hp$  versions). We show that the asymptotic  $p$ -convergence rate remains optimal in this case, and indicate, in Remark 3.7, what can be expected for the case of arbitrary geometric coefficients. Our analysis here is only carried out for the case of parallelograms, though we formulate our elements for general quadrilaterals. We believe that our results provide valuable insight into  $p$  and  $hp$  versions for more general shell and element geometries as well. (See Remark 3.4 in this context, which discusses the treatment of boundaries).

The plan of the paper is as follows. Section 2 describes the Naghdi model and introduces the modified variational form from [1]. In Section 3, we describe the approximate method and present our main theorem. The proof is presented separately in Section 4. In order to properly analyze the  $p$  convergence, we do not analyze the problem as a mixed method (as in [1]), but rather use the approach in [7] (see Remark 4.2). Essentially, we prove uniform estimates for the approximations as the solution tends to an inextensional (pure bending) one. (See Remark 3.4.)

2. THE VARIATIONAL FORMULATION FOR THE NAGHDI SHELL MODEL

In this section, we describe the Naghdi shell model in its variational form. In this we closely follow Section 2 of [1].

As in [1], we let Greek indices range over 1, 2 and Latin indices over 1, 2, 3, with products containing repeated indices being summed. Also, 3-vectors  $\vec{v} = (v_i)$  are indicated by overarrows, 2-vectors by undertildes ( $\tilde{v} = (v_\alpha)$ ) and  $2 \times 2$  symmetric tensors by double undertildes ( $\tilde{\lambda} = (\lambda_{\alpha\beta})$ ) with  $\lambda_{\alpha\beta} = \lambda_{\beta\alpha}$ . If  $V$  is a space of scalar functions, then  $\vec{v} \in \vec{V}$ ,  $\tilde{v} \in \tilde{V}$  or  $\tilde{\lambda} \in \tilde{\tilde{V}}$  will imply that each component is in  $V$ .

The Naghdi shell model involves various coefficients related to the geometry, which we now describe. Let us assume, for simplicity, that the midsurface of the shell can be represented as  $\vec{r}(\Omega)$  where  $\vec{r}$  is a smooth one-to-one mapping of  $\bar{\Omega}$  into  $\mathbf{R}^3$ , where  $\Omega$  is a bounded open set in  $\mathbf{R}^2$ . Moreover, let  $\vec{r}$  be such that  $\vec{a}_1 = \vec{r}_{x^1}$  and  $\vec{a}_2 = \vec{r}_{x^2}$  are linearly independent at each point of  $\bar{\Omega}$ . Then, defining  $\vec{a}_3$  to be the unit vector along  $\vec{a}_1 \times \vec{a}_2$ , the shell (of thickness  $d \in (0, 1]$ ) will occupy the region

$$\left\{ \vec{r}(\tilde{x}) + x^3 \vec{a}_3(\tilde{x}) \mid \tilde{x} \in \Omega, \frac{-d}{2} < x^3 < \frac{d}{2} \right\} \subset \mathbf{R}^3.$$

Next, let the first and second fundamental forms of  $\vec{r}(\Omega)$  be defined by the matrices

$$a_{\alpha\beta} = \vec{a}_\alpha \cdot \vec{a}_\beta, \quad b_{\alpha\beta} = \vec{a}_3 \cdot \frac{\partial^2 \vec{r}}{\partial x^\alpha \partial x^\beta}$$

and let  $a$  be the determinant of the matrix  $(a_{\alpha\beta})$ ,  $a \neq 0$  on  $\bar{\Omega}$ . We denote the inverse of  $(a_{\alpha\beta})$  by  $(a^{\alpha\beta})$  and set

$$\vec{a}^\alpha = a^{\alpha\beta} \vec{a}_\beta, \quad b_\gamma^\alpha = a^{\alpha\beta} b_{\beta\gamma}, \quad a^{\alpha\beta\gamma\delta} = \frac{E}{1-\nu^2} (a^{\alpha\gamma} a^{\beta\delta} + \nu a^{\alpha\beta} a^{\gamma\delta}),$$

where  $E > 0$  and  $\nu \in [0, \frac{1}{2})$  are the Young's modulus and Poisson ratio respectively. Finally, we define the Christoffel symbols  $\Gamma_{\alpha\beta}^\delta := \vec{a}^\delta \cdot \frac{\partial \vec{a}_\beta}{\partial x^\alpha}$ .

Let  $H^k(\Omega)$  denote the usual Sobolev spaces, with  $L_2(\Omega) = H^0(\Omega)$ . Let  $\Gamma_D \neq \emptyset$  be a portion of  $\partial\Omega$  and define

$$H_D^r(\Omega) = \{v \in H^r(\Omega) \mid v|_{\Gamma_D} = 0\} \quad \text{for } r \geq 1,$$

$$\mathcal{V} = \{(\vec{v}, \tilde{\psi}) \mid \vec{v} \in H_D^1(\Omega), \tilde{\psi} \in \tilde{H}_D^1(\Omega)\}$$

with the norm in  $\mathcal{V}$  given by

$$\|(\vec{v}, \tilde{\psi})\|_{\mathcal{V}}^2 = \|\vec{v}\|_1^2 + \|\tilde{\psi}\|_1^2.$$

We will assume that our solution  $(\vec{u}, \tilde{\theta}) \in \mathcal{V}$ , i.e. we will assume clamped conditions along  $\Gamma_D$ . Suppose  $d^3 f^i \vec{a}_i$  is the resultant of the forces applied on the

midsurface, then the Naghdi model determines the displacement  $\vec{u} = (u_i)$  and rotation  $\theta = (\theta_\alpha)$  as the solution of the variational problem: Find  $U = (\vec{u}, \theta) \in \mathcal{V}$  satisfying

$$(2.1) \quad \begin{aligned} a(U, V) + d^{-2}(\Phi(U), \Phi(V))_1 + d^{-2}(\Lambda(U), \Lambda(V))_2 &= F(V) \\ &\text{for all } V = (\vec{v}, \psi) \in \mathcal{V}, \end{aligned}$$

where

$$\begin{aligned} a(U, V) &= \int_{\Omega} \frac{a^{\alpha\beta\gamma\delta}}{12} \Psi_{\alpha\beta}(\vec{u}, \theta) \Psi_{\gamma\delta}(\vec{v}, \psi) \sqrt{a} \, dx, \\ (\Phi(U), \Phi(V))_1 &= \int_{\Omega} a^{\alpha\beta} \frac{E}{2(1+\nu)} \Phi_{\alpha}(\vec{u}, \theta) \Phi_{\beta}(\vec{v}, \psi) \sqrt{a} \, dx, \\ (\Lambda(U), \Lambda(V))_2 &= \int_{\Omega} a^{\alpha\beta\gamma\delta} \Lambda_{\alpha\beta}(\vec{u}) \Lambda_{\gamma\delta}(\vec{v}) \sqrt{a} \, dx, \\ F(V) &= \int_{\Omega} f^i u_i \sqrt{a} \, dx. \end{aligned}$$

In the above,  $\Psi$ ,  $\Phi$ , and  $\Lambda$  are the change of curvature tensor, the transverse shear strain tensor, and the membrane strain tensor, respectively. For any  $\vec{u} = (u_i)$ ,  $\theta = (\theta_\alpha)$  with  $u_i, \theta_\alpha \in H^1(\Omega)$ , these are defined by

$$(2.2) \quad \begin{aligned} \Psi_{\alpha\beta}(\vec{u}, \theta) &= \frac{1}{2}[\theta_{\alpha,\beta} + \theta_{\beta,\alpha} - b_{\alpha}^{\gamma}(u_{\gamma,\beta} - \Gamma_{\gamma\beta}^{\delta} u_{\delta}) - b_{\beta}^{\gamma}(u_{\gamma,\alpha} - \Gamma_{\gamma\alpha}^{\delta} u_{\delta}) + b_{\alpha}^{\gamma} b_{\gamma\beta} u_3] \\ &\quad - \Gamma_{\alpha\beta}^{\delta} \theta_{\delta}, \end{aligned}$$

$$(2.3) \quad \Phi_{\alpha}(\vec{u}, \theta) = u_{3,\alpha} + b_{\alpha}^{\gamma} u_{\gamma} + \theta_{\alpha},$$

$$(2.4) \quad \Lambda_{\alpha\beta}(\vec{u}) = \frac{1}{2}(u_{\alpha,\beta} + u_{\beta,\alpha}) - \Gamma_{\alpha\beta}^{\delta} u_{\delta} - b_{\alpha\beta} u_3.$$

In the bending-dominated case, as  $d \rightarrow 0$ , the loads are scaled as  $d^3 f^i \vec{a}_i$  and the bending energy  $a(U, V)$  dominates. In the membrane-dominated case, the load is scaled as  $d f^i \vec{a}_i$  and as  $d \rightarrow 0$ , the bending energy term  $d^2 a(U, V)$  is the one that tends to zero, with the solution tending to a pure membrane state. The latter case is actually the more common in applications (since the shell can be loaded by a larger amount this way). Smooth pure membrane deformations are free of locking and standard finite element methods work well here.

Here, we are interested in developing a method that is uniformly robust as  $d \rightarrow 0$  in the bending-dominated case. It may be seen from (2.1) that for the energy to remain finite as  $d \rightarrow 0$ , we must now have the following constraints in the limit:

$$(2.5) \quad \Phi(U) = \Phi(\vec{u}, \theta) = 0, \quad \Lambda(U) = \Lambda(\vec{u}) = 0.$$

When (2.1) is discretized, the finite element space used may not have enough functions satisfying (2.5) (for instance, the only such function in the space may be  $(\vec{u}, \theta) \equiv 0$ ). This is what causes locking, and to remove it, as we see below, the finite element solution will be required to satisfy (2.5) only in a weaker sense.

Let us note another problem with the formulation (2.1): the form  $a(U, V)$  is not coercive over the space  $\mathcal{V}$ . This makes the development of locking-free approximation methods more difficult. In [1], a simple rewriting of (2.1) was used to overcome this problem. Let  $c_0$  be an arbitrary positive constant, independent of  $d$  (say  $c_0 < 1$ ). Define

$$A(U, V) = a(U, V) + c_0(\tilde{\Phi}(U), \tilde{\Phi}(V))_1 + c_0(\tilde{\Lambda}(U), \tilde{\Lambda}(V))_2, \quad t^{-2} = d^{-2} - c_0.$$

Then (2.1) is equivalent to

$$(2.6) \quad A(U, V) + t^{-2}(\tilde{\Phi}(U), \tilde{\Phi}(V))_1 + t^{-2}(\tilde{\Lambda}(U), \tilde{\Lambda}(V))_2 = F(V)$$

where  $t \rightarrow 0$  as  $d \rightarrow 0$ , and by the results in [5],

$$(2.7) \quad A(U, V) \geq C\|U\|_{\mathcal{V}}^2.$$

(The above rewriting idea has been used in other contexts before—see e.g. [12] and [4, equation (3.1)].) By Theorem 2 of [1], problem (2.6) will have a unique solution, with  $\|U\|_{\mathcal{V}}$  bounded by  $\|F\|_{\mathcal{V}^*}$ , the norm of  $F$  in the dual space of  $\mathcal{V}$ .

### 3. THE APPROXIMATE METHOD

Suppose now that  $\mathcal{V}_N \subset \mathcal{V}$  is a sequence of finite element spaces, parametrized by  $N$ . While finding the finite element solution, in order to reduce the effect of the constraints (2.5), we will only require that

$$(3.1) \quad P_N^1 \tilde{\Phi}(U) = 0, \quad P_N^2 \tilde{\Lambda}(U) = 0,$$

where for any  $X \in L_2(\Omega)$ ,  $Y \in L_2(\Omega)$ ,

$$P_N^1 X \in S_N^1 \subset L_2(\Omega) \quad \text{and} \quad P_N^2 Y \in S_N^2 \subset L_2(\Omega)$$

are defined as orthogonal projections,

$$(3.2) \quad (\tilde{X} - P_N^1 \tilde{X}, \tilde{W})_1 = 0 \quad \forall \tilde{W} \in S_N^1, \quad (\tilde{Y} - P_N^2 \tilde{Y}, \tilde{V})_2 = 0 \quad \forall \tilde{V} \in S_N^2.$$

Assuming that a sequence of such spaces  $S_N^1, S_N^2$  has been defined, the finite element approximation to  $U$  is given by the solution  $U_N \in \mathcal{V}$  of

$$(3.3) \quad \begin{aligned} A(U_N, V) + t^{-2}(P_N^1 \tilde{\Phi}(U_N), P_N^1 \tilde{\Phi}(V))_1 \\ + t^{-2}(P_N^2 \tilde{\Lambda}(U_N), P_N^2 \tilde{\Lambda}(V))_2 = F(V) \quad \forall V \in \mathcal{V}_N. \end{aligned}$$

Using (2.7), it is easily seen that problem (3.3) has a unique solution for each  $t$  and  $N$ . Method (3.3) has been termed as “partial selective reduced integration” in [1]. Note that in terms of implementation, the projections  $P_N^1, P_N^2$  in (3.2) can be calculated elementwise (if  $S_N^1, S_N^2$  are free of interelement continuity constraints). Also, introducing two auxiliary variables, (2.6), (3.3) can be formulated as equivalent mixed formulations—see [1].

In the above, the “reduction” operators  $P_N^1, P_N^2$  are clearly characterized completely by  $S_N^1, S_N^2$ . To see what properties  $S_N^1, S_N^2$  should possess, we note the following result. The proof is essentially the same as that for equation (27) of [7].

**Theorem 3.1.** *Let  $U \in \mathcal{V}$  be the solution of (2.6) and  $U_N \in \mathcal{V}_N$  the solution of (3.3). Then, for any  $W_N \in \mathcal{V}_N$ ,*

$$\begin{aligned} \|U - U_N\|_{\mathcal{V}} &\leq C(\|U - W_N\|_{\mathcal{V}} + \|(I - P_N^1)(t^{-2}\Phi(U))\|_0 \\ &\quad + \|(I - P_N^2)(t^{-2}\Lambda(U))\|_0 + t^{-2}\|P_N^1\Phi(U - W_N)\|_0 \\ &\quad + t^{-2}\|P_N^2\Lambda(U - W_N)\|_0) \end{aligned}$$

where  $C$  is a constant independent of  $U, N$ .

Let us define

$$(3.4) \quad \xi(U) = t^{-2}\Phi(U) \quad \text{and} \quad \eta(U) = t^{-2}\Lambda(U)$$

( $\xi, \eta$ ) are related to the transverse shear stresses and membrane stresses, respectively). Then, Theorem 3.1 gives the following corollary.

**Corollary 3.1.** *Let  $Z_N(U) = \{W_N \in \mathcal{V}_N | P_N^1\Phi(U - W_N) = 0, P_N^2\Lambda(U - W_N) = 0\} \subset \mathcal{V}_N$ . Then,*

$$(3.5) \quad \begin{aligned} \|U - U_N\|_{\mathcal{V}} \\ \leq C \left( \inf_{W_N \in Z_N(U)} \|U - W_N\|_{\mathcal{V}} + \|(I - P_N^1)(\xi(U))\|_0 + \|(I - P_N^2)(\eta(U))\|_0 \right). \end{aligned}$$

*Remark 3.1.* If we take  $P_N^i \equiv I$  in (3.5), Corollary 3.1 reduces essentially to the corresponding theorem for standard (exact constraint) methods, see Theorem 2.2 of [4]. The last two terms are additional *consistency* terms when  $P_N^i \neq I$ . A more general form of Corollary 3.1, when  $P_N^i$  are more general reduction operators (as opposed to projections in an  $L_2$  type inner product), may be found in [13].

From Corollary 3.1, we see the requirements on  $S_N^1, S_N^2$ . These should be large enough so that  $\xi(U), \eta(U)$  can be approximated well. Also, they should be small enough so that  $Z_N(U)$  has enough functions to ensure the infimum in (3.5) is small. We now present a choice of spaces  $\mathcal{V}_N, S_N^1, S_N^2$  for which the above criteria hold.

Assume that  $\Omega$  is a polygon and let  $\{J_N\}$  be a sequence of meshes of parallelograms on  $\Omega$  such that  $\Gamma_D$  is a union of edges of the elements for each  $J_N$ . Let  $Q$  be the reference square  $(-1, 1)^2$  and assume that each  $K \in J_N$  is given by  $K = \vec{F}_K(Q)$ , where  $\vec{F}_K$  is an affine invertible mapping. Let  $h_K, \rho_K$  be the diameter of  $K$ , and the largest circle that can be inscribed in  $K$ , respectively, and let  $h_N = \max_{K \in J_N} h_K$ . We assume there exists a constant  $C$ , independent of  $N$ , such that for all  $K \in J_N$ ,  $(h_K/\rho_K) \leq C$ . Further, we assume that for each pair  $K_1, K_2 \in J_N$ ,  $K_1 \cap K_2$  is either an entire side, a vertex, or the empty set.

Let  $\tilde{v}$  be a function defined on  $Q$  (or  $\partial Q$ ). With  $\tilde{v}$ , we associate the function  $v$  defined on  $K$  (or  $\partial K$ ) by

$$(3.6) \quad v = \tilde{v} \circ \vec{F}_K^{-1}$$

for  $\tilde{v}$  a scalar. For  $\tilde{v}$  a vector or  $2 \times 2$  symmetric tensor, each component of  $v$  is defined as in (3.6).

For an interval  $I \subset \mathbf{R}$ , we denote  $\mathcal{P}_p(I)$  to be the set of polynomials of degree  $\leq p$  on  $I$ . We denote the set of polynomials of degree  $\leq p$  in each variable on  $Q$  by

$\mathcal{Q}_p(Q)$ . For each  $N$ , let  $p_N \geq 0$  be an associated integer. We define

$$(3.7) \quad V_k^r(J_N) = \{v \in H^r(Q) \mid v|_K \circ \vec{F}_K \in \mathcal{Q}_k(Q) \text{ for all } K \in J_N\}.$$

We make the assumption that the quantities  $a_{\alpha\beta}, b_{\alpha\beta}, \Gamma_{\alpha\beta}^\gamma$  occurring in our bilinear form belong to  $V_m^0(J_N)$  for some integer  $m \geq 0$ . Then we set

$$(3.8) \quad \mathcal{V}_N = \{(\vec{v}, \phi) \in \mathcal{V} \mid v_i, \phi_\alpha \in V_{p_N+m+2}^1(J_N)\},$$

$$(3.9) \quad S_N^1 = \{\mu \in L_2(\Omega) \mid \mu_\alpha \in V_{p_N}^0(J_N)\}, \quad S_N^2 = \{\lambda \in L_2(\Omega) \mid \lambda_{\alpha\beta} \in V_{p_N}^0(J_N)\}.$$

Hence the components in  $S_N^i$  are of  $m+2$  degrees less than those in  $\mathcal{V}_N$ . Note that for  $m=0$ , unlike the triangular elements in [1], no extra bubble functions are used in the definition of  $\mathcal{V}_N$  here. (The case  $m=0$  was the only one considered in [1].)

Equations (3.7)–(3.9) could also be used to define general quadrilateral or curvilinear elements, in which case  $\vec{F}_K$  would no longer be affine. Our analysis here will be restricted, however, to parallelograms. We have the following theorem, the proof of which is given in the next section.

**Theorem 3.2.** *Let  $a_{\alpha\beta}, b_{\alpha\beta}, \Gamma_{\alpha\beta}^\gamma \in V_m^0(J_N), m \geq 0$ . Let  $U = (\vec{u}, \theta) \in \mathcal{V}$ ,  $U_N = (\vec{u}_N, \theta_N) \in \mathcal{V}_N$  be the solutions of (2.6), (3.3) respectively and let  $\xi(U), \eta(U)$  be as in (3.4). Then for any  $0 < c_0 < 1$ ,  $d \in (0, 1]$  and  $\varepsilon > 0$ ,<sup>1</sup>*

$$(3.10) \quad \begin{aligned} \|U - U_N\|_{\mathcal{V}} &= \|\vec{u} - \vec{u}_N\|_1 + \|\theta - \theta_N\|_1 \\ &\leq Ch_N^\mu p_N^{-\gamma} (\|\vec{u}\|_r + \|\theta\|_r + \|\xi(U)\|_{r-1} + \|\eta(U)\|_{r-1}), \end{aligned}$$

where  $r$  (real) satisfies  $r > 3/2$ ,  $\mu = \min(p_N + 1, r - 1)$ ,  $\gamma = r - 1$  if  $r \geq 2$ ,  $\gamma = 2r - 3 - \varepsilon$  if  $3/2 < r \leq 2$  and where  $C$  is a constant independent of  $h_N, p_N, U$  and  $d$  but dependent on  $r$  and  $\varepsilon$ .

*Remark 3.2.* We see from the above theorem that the rate of convergence is  $O(h_N^{p_N+1} p_N^{-(r-1)})$  for the case that the norms on the right-hand side are bounded for  $r \geq p_N + 2$ . Technically speaking, this is optimal in terms of  $h_N$ , since by (3.5), the best rate of convergence that can be expected with this method when  $d$  is not small is  $O(h_N^{p_N+1})$ , due to the use of polynomials of degree  $p_N$  for  $S_N^i$ . Hence, by the definition in [4], there is no locking in terms of  $h_N$ . However, effectively, we are seeing a loss of  $O(h_N^{-1-m})$  compared to the optimal approximation rate by polynomials of degree  $p_N + m + 2$ , which is what we are using for the primary variables. In terms of  $p_N$ , the uniform estimate we get is the expected rate of  $O(p_N^{-(r-1)})$ , i.e. there is no locking in the asymptotic rate, provided  $r \geq 2$ . For  $3/2 < r \leq 2$ , there is possible locking of  $O(p_N^{1/2+\varepsilon})$ .

*Remark 3.3.* Let us mention that in practice, the right-hand side of (3.10) may not be bounded except for small  $r$ , due to the presence of boundary layers. In this case, the above rates could be recovered only if some special care (e.g. extra refinement) is given to the approximation of the boundary layers. Such a strategy is discussed in [13] for plate problems. Essentially, the grid in a strip of  $\Omega_0$  adjacent to  $\partial\Omega$  should be of mesh spacing  $h_N^0$  ( $h_N^0 \sim d^{1-\frac{1}{2p_N}} h_N$  for the methods discussed

<sup>1</sup>Here and in the sequel,  $\varepsilon$  will denote a small positive number.

in [13] for the Reissner-Mindlin plate). Then the rate of convergence that could be expected in the absence of boundary layers is preserved. The results in [13] do not depend upon explicit formulas for the boundary layer. Based on recent work by Piila and Pitkäranta (see e.g. [11]), in which asymptotic estimates for shells as  $d \rightarrow 0$  are derived, we expect that a similar treatment would be successful in the case of shell problems.

*Remark 3.4.* If the above refinement is carried out near the boundary, our parallelogram elements could be combined with the triangular elements of [1] to better approximate the boundary. The rate of convergence in terms of  $h_N$  is the same for the two types of elements, so that using a refined mesh spacing  $h_N^0$  in  $\Omega_0$  would lead once more to approximation of the boundary layer. Meanwhile, the effect of using high  $p_N$  for the parallelogram elements in  $\Omega \setminus \Omega_0$  would still be preserved (we believe that similar rates in  $p_N$  could be expected from triangles as well). Hence, although we have carried out our  $hp$  analysis only for parallelograms, this is not a serious limitation when taken in the context of the above ideas.

*Remark 3.5.* The elements we use are the usual quadrilateral ones that are popular in the field of spectral element methods [6]. This suggests implementation by spectral element techniques, which often exploit the tensor product nature of the spaces to good advantage in terms of computational time.

*Remark 3.6.* Our analysis here is limited to the bending-dominated case, and does not answer the question of how well the “partial selective reduced integration” may behave in the membrane-dominated case. In fact, as noted by Pitkäranta (private communication), the coercivity can break down in the limit for this case (depending on the choice of spaces). Hence, the best strategy might be to combine the method here (for the bending-dominated case) with the standard FE method (for the membrane-dominated case). In fact, the standard FE method is easily accomplished, simply by setting  $c_0 = d^{-2}$  in the above method. (Note, also, that shear and membrane locking effects could be treated separately as well as taking two separate parameters  $c_1, c_2$  instead of  $c_0$ .)

*Remark 3.7.* For the case of arbitrary geometric coefficients, we can first approximate them by piecewise polynomials of degree  $m_N$ , and then use the above spaces (with  $m = m_N$  increasing with  $N$ ). This introduces an additional  $O(d^{-2}f(N))$  consistency error. If the coefficients are smooth,  $f(N)$  decreases superexponentially as  $N$  increases, so that this term will not dominate in the error. Such an approach can also be used to analyze non-parallelogram elements, where the non-constant (smooth) Jacobians may be lumped with the general geometric coefficients. A full analysis using this approach will be given in a future work. The question of finding the minimal spaces  $\mathcal{V}_N$  that can be used will also be considered. (The basic idea is to use the  $p$  version to control locking by adding the minimal number of extra functions to  $\mathcal{V}_N$  while keeping the spaces  $S_N^i$  fixed.)

#### 4. PROOF OF THEOREM 3.2

To prove Theorem 3.2, we need a series of technical lemmas, culminating in Lemma 4.4, which gives an error estimate for a projection operator  $\tilde{\Pi}_N$ . Once this estimate is established, the proof of Theorem 3.2 follows in a few lines (see the end of this section). We start by defining a projection operator,  $\Pi_k : H^{1+\varepsilon}(Q) \rightarrow \mathcal{Q}_k(Q)$



as follows:

$$(4.1) \quad (\Pi_k u)(N_i) = u(N_i) \quad \text{for each node } N_i \text{ of } Q,$$

$$(4.2) \quad \int_e (\Pi_k u)(s)q(s)ds = \int_e u(s)q(s)ds, \quad \forall q \in \mathcal{P}_{k-2}(e), \text{ for each edge } e \text{ of } Q,$$

$$(4.3) \quad \iint_Q (\Pi_k u)(\vec{x})q(\vec{x})d\vec{x} = \iint_Q u(\vec{x})q(\vec{x})d\vec{x}, \quad \forall q \in \mathcal{Q}_{k-2}(Q).$$

Equations (4.1)–(4.3) give a unique  $\Pi_k u \in \mathcal{Q}_k(Q)$  for any  $u \in H^{1+\varepsilon}(Q)$ . Using Green's Theorem, we see from the above definition that for  $|\alpha| \leq 1$ ,

$$(4.4) \quad \iint_Q D^\alpha(u - \Pi_k u)q d\vec{x} = 0 \quad \text{for all } q \in \mathcal{Q}_{k-2}(Q).$$

Next, we prove some further properties of  $\Pi_k$ .

Let  $L_i(x)$  be the Legendre polynomial of degree  $i \geq 0$  and define  $L_{-1}(x) = L_{-2}(x) = 0$ . Also, let for  $i \geq 0$ ,  $\gamma_i = (2i + 1)^{-1}$  and define  $\gamma_{-1} = 1$ . Then we set

$$(4.5) \quad U_i(x) = \gamma_{i-1}(L_i(x) - L_{i-2}(x)), \quad i \geq 0.$$

We may easily verify that

$$(4.6) \quad U_i'(x) = L_{i-1}(x), \quad i \geq 0.$$

Obviously,  $\{U_i\}_{i=0}^k$  forms a basis for  $\mathcal{P}_k(I)$ ,  $I = (-1, 1)$ , and so the products  $U_i(x)U_j(y)$ ,  $0 \leq i, j \leq k$  will form a basis for  $\mathcal{Q}_k(Q)$ . Hence, if  $v \in \mathcal{Q}(Q)$ , the set of *all* polynomials on  $Q$ , then we may write

$$(4.7) \quad v(x, y) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_{ij} U_i(x) U_j(y)$$

where the sum will be finite. Also, for  $v_k(x, y) \in \mathcal{Q}_k(Q)$ , we may write

$$(4.8) \quad v_k(x, y) = \sum_{i=0}^k \sum_{j=0}^k b_{ij} U_i(x) U_j(y).$$

For  $v \in \mathcal{Q}(Q)$  given by (4.7), we now calculate  $v_k = \Pi_k v$  (assumed to be given by formula (4.8)). First, since  $U_i(\pm 1) = 0$  for  $i > 1$ , we may obtain using (4.1), (4.7), (4.8),

$$(4.9) \quad b_{ij} = a_{ij}, \quad 0 \leq i, j \leq 1.$$

Next, consider the sides  $x = \pm 1$ . Then, using (4.9), we may integrate (4.2) by parts to obtain

$$(4.10) \quad \int_{-1}^1 v_{ky}(\pm 1, y)r(y)dy = \int_{-1}^1 v_y(\pm 1, y)r(y)dy \quad \forall r \in \mathcal{P}_{k-1}(I).$$

Let  $r(y) = L_{n-1}(y)$ ,  $n = 2, \dots, k$ , in (4.10). Using (4.7), (4.8), with  $x = \pm 1$ ,

$$\int_{-1}^1 \sum_{j=1}^k (b_{0j} \pm b_{1j}) L_{j-1}(y) L_{n-1}(y) dy = \int_{-1}^1 \sum_{j=1}^{\infty} (a_{0j} \pm a_{1j}) L_{j-1}(y) L_{n-1}(y) dy$$

from which

$$b_{0n} = a_{0n}, \quad b_{1n} = a_{1n}, \quad n = 2, \dots, k.$$

Similarly, from the sides  $y = \pm 1$ ,

$$b_{n0} = a_{n0}, \quad b_{n1} = a_{n1}, \quad n = 2, \dots, k.$$

Finally, it may be easily verified (see Lemma 4.1 of [15]) that conditions (4.3) give

$$b_{ij} = a_{ij}, \quad 2 \leq i, j \leq k.$$

Hence, for  $v$  given by (4.7), we have

$$(4.11) \quad (\Pi_k v)(x, y) = \sum_{i=0}^k \sum_{j=0}^k a_{ij} U_i(x) U_j(y).$$

We will also use the one-dimensional projection  $v \rightarrow \pi_k v(\pm 1) \in \mathcal{P}_k(I)$  defined on  $H^{1/2+\varepsilon}(I)$  by

$$\pi_k v(\pm 1) = v(\pm 1), \quad \int_{-1}^1 (\pi_k v - v) q = 0, \quad q \in \mathcal{P}_{k-2}(I).$$

For  $v \in H_0^1(I)$ ,  $\pi_k$  is seen to just be the  $H^1$  projection of  $v$  into  $\mathcal{P}_k(I)$ . Using this fact and a duality argument, it may be shown that

$$(4.12) \quad \|(\pi_k v)'\|_{0,I} \leq \|v'\|_{0,I},$$

$$(4.13) \quad \|\pi_k v - v\|_{0,I} \leq \frac{C}{(k-1)} \|v'\|_{0,I}.$$

Denoting  $\pi_k^x, \pi_k^y$  to be the projection in variables  $x, y$  respectively, we may verify that

$$(4.14) \quad \Pi_k = \pi_k^x \pi_k^y = (\pi_k^x \otimes I)(I \otimes \pi_k^y).$$

Let us now estimate  $\|v - \Pi_k v\|_{1,Q}$ .

**Lemma 4.1.** *For  $v$  given by (4.7),*

$$(4.15) \quad \|v\|_{0,Q}^2 = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} 4\gamma_i \gamma_j (\gamma_{i-1} \gamma_{j-1} a_{ij} + \gamma_{i+1} \gamma_{j+1} a_{i+2,j+2} - \gamma_{i+1} \gamma_{j-1} a_{i+2,j} - \gamma_{i-1} \gamma_{j+1} a_{i,j+2})^2,$$

$$(4.16) \quad \|v_x\|_{0,Q}^2 = \sum_{i=1}^{\infty} \sum_{j=0}^{\infty} 4\gamma_{i-1} \gamma_j (\gamma_{j-1} a_{ij} - \gamma_{j+1} a_{i,j+2})^2,$$

$$(4.17) \quad \|v_y\|_{0,Q}^2 = \sum_{i=0}^{\infty} \sum_{j=1}^{\infty} 4\gamma_i \gamma_{j-i} (\gamma_{i-1} a_{ij} - \gamma_{i+1} a_{i+2,j})^2.$$

*Proof.* The above results follow easily by the orthogonality properties of Legendre polynomials and their derivatives. In particular, they are established as in Lemma 4.2 of [15].  $\square$

Obviously, a similar lemma holds for (4.8) or (4.11). Using this lemma and the characterization (4.11), we obtain the following result, which generalizes Lemma 4.3 of [15].

**Lemma 4.2.** *For any  $v \in H^r(Q)$ ,  $r > 3/2$ , let  $\Pi_k v \in \mathcal{Q}_k(Q)$  be defined by (4.1)–(4.3). Then*

$$(4.18) \quad \|\Pi_k v\|_{1,Q} \leq C \left( k^{1/2} \|v\|_{1,Q} + k^{-1/2} \sum_{e \in \partial Q} \|v\|_{1,e} \right),$$

where  $e$  are the edges of  $Q$ .

*Proof.* It is sufficient to prove the result for the operator  $\Pi_k$  defined only on  $\mathcal{Q}(Q)$ . Then by [2, pages 13–14], there is a unique norm preserving extension of  $\Pi_k$  to  $H^r(Q)$ ,  $r > 3/2$ , which coincides with the definition via (4.1)–(4.3). Thus, let  $v \in \mathcal{Q}(Q)$  be given by (4.7), so that  $\Pi_k v$  is given by (4.11). Then, using (4.16),

$$(4.19) \quad \begin{aligned} \|(\Pi_k v)_x\|_{0,Q}^2 &= \sum_{i=1}^k \sum_{j=0}^{k-2} 4\gamma_{i-1}\gamma_j (\gamma_{j-1}a_{ij} - \gamma_{j+1}a_{i,j+2})^2 \\ &\quad + \sum_{i=1}^k (4\gamma_{i-1}\gamma_{k-1}(\gamma_{k-2}a_{i,k-1})^2 + 4\gamma_{i-1}\gamma_k(\gamma_{k-1}a_{ik})^2) \\ &\leq \|v_x\|_{0,Q}^2 + B. \end{aligned}$$

To bound  $B$ , let  $k$  be odd (say). Then for each  $i = 1, \dots, k$ , we can use a telescoping sum to write

$$\gamma_{k-2}a_{i,k-1} = - \sum_{m=0}^{(k-3)/2} (\gamma_{2m-1}a_{i,2m} - \gamma_{2m+1}a_{i,2m+2}) + a_{i0}.$$

Squaring and using the fact that  $\gamma_{k-1} \leq \gamma_j$  for  $j \leq k-1$ ,

$$\begin{aligned} &4\gamma_{i-1}\gamma_{k-1}(\gamma_{k-2}a_{i,k-1})^2 \\ &\leq C \left( k \sum_{m=0}^{(k-3)/2} 4\gamma_{i-1}\gamma_{2m}(\gamma_{2m-1}a_{i,2m} - \gamma_{2m+1}a_{i,2m+2})^2 + \gamma_{i-1}\gamma_{k-1}a_{i0}^2 \right). \end{aligned}$$

Using a similar argument for  $4\gamma_{i-1}\gamma_k(\gamma_{k-1}a_{ik})^2$ , and substituting in (4.19), we have

$$\|(\Pi_k v)_x\|_{0,Q}^2 \leq C \left( k \|v_x\|_{0,Q}^2 + \sum_{i=1}^k (\gamma_{i-1}\gamma_{k-1}a_{i0}^2 + \gamma_{i-1}\gamma_k a_{i1}^2) \right).$$

Treating the terms  $\|(\Pi_k v)_y\|_{0,Q}^2$  and  $\|(\Pi_k v)\|_{0,Q}^2$  by a similar method, we obtain

$$(4.20) \quad \begin{aligned} \|(\Pi_k v)\|_{1,Q}^2 &\leq C \left( k \|v\|_{1,Q}^2 + \sum_{i=1}^k (\gamma_{i-1}\gamma_{k-1}a_{i0}^2 + \gamma_{i-1}\gamma_k a_{i1}^2) \right. \\ &\quad \left. + \sum_{j=1}^k (\gamma_{k-1}\gamma_{j-1}a_{0j}^2 + \gamma_k \gamma_{j-1}a_{1j}^2) \right). \end{aligned}$$

We now bound the last two terms in (4.20). We have, using (4.7),

$$\begin{aligned} \int_{-1}^{+1} v_y^2(\pm 1, y) dy &= \int_{-1}^{+1} \left( \sum_{j=0}^{\infty} (a_{0j} \pm a_{1j}) L_{j-1}(y) \right)^2 dy \\ &= \sum_{j=0}^{\infty} 2\gamma_{j-1} (a_{0j} \pm a_{1j})^2. \end{aligned}$$

Hence,

$$(4.21) \quad \sum_{j=1}^k (\gamma_{k-1} \gamma_{j-1} a_{0j}^2 + \gamma_k \gamma_{j-1} a_{1j}^2) \leq \frac{\gamma_{k-1}}{2} \left( \int_{-1}^{+1} v_y^2(+1, y) dy + \int_{-1}^{+1} v_y^2(-1, y) dy \right).$$

Using a similar argument for the other term in (4.20), and combining with (4.21), we get the lemma.  $\square$

*Remark 4.1.* For functions  $v$  that vanish on  $\partial Q$ , the projection  $\Pi_k$  coincides with the projection  $\mathcal{S}_k$  defined in [15]. By the sharpness of the bound for  $\mathcal{S}_k$  (Remark 4.1 of [15]), we conclude that the loss of the factor  $k^{1/2}$  in (4.18) is sharp as well.

**Lemma 4.3.** *For any  $v \in H^r(Q)$ ,  $r > 3/2$ , and any  $\varepsilon > 0$ , there exists a constant  $C$  depending on  $r, \varepsilon$  such that*

$$\|v - \Pi_k v\|_{1,Q} \leq C k^{-\gamma} \|v\|_{r,Q},$$

where  $\gamma = r - 1$  if  $r \geq 2$ , and  $\gamma = 2r - 3 - \varepsilon$  if  $3/2 < r \leq 2$ .

*Proof.* Given  $v \in H^r(Q)$ , let  $v_k \in \mathcal{Q}_k(Q)$  satisfy conditions (4.1), (4.2) (with  $v$  instead of  $u$ ), and also

$$\iint_Q \nabla v_k \cdot \nabla q \, d\vec{x} = \iint_Q \nabla v \cdot \nabla q \, d\vec{x} \quad \forall q \in \mathcal{Q}_k(Q).$$

Then,  $v_k$  is the projection of  $v$  on  $\mathcal{Q}_k(Q)$  discussed in [3], and by the results in that reference,

$$(4.22) \quad \|v - v_k\|_{1,Q} \leq C k^{-(r-1)} \|v\|_{r,Q},$$

$$(4.23) \quad \|v - v_k\|_{1,e} \leq C k^{-(r-3/2)} \|v\|_{r,Q} \quad \text{for each edge of } Q.$$

Hence, using Lemma 4.2, (4.22) and (4.23),

$$(4.24) \quad \begin{aligned} \|v - \Pi_k v\|_{q,Q} &\leq \|v - v_k\|_{1,Q} + \|\Pi_k(v - v_k)\|_{1,Q} \\ &\leq \|v - v_k\|_{1,Q} + C \left( k^{1/2} \|v - v_k\|_{1,Q} + k^{-1/2} \sum_{e \in \partial Q} \|v - v_k\|_{1,e} \right) \\ &\leq C k^{-(r-3/2)} \|v\|_{r,Q}. \end{aligned}$$

Now, suppose  $r \geq 2$ . Then, using (4.12)–(4.14),

$$\begin{aligned} \|v - \Pi_k v\|_{0,Q} &= \|(v - q) - \Pi_k(v - q)\|_{0,Q} \\ &= \|(v - q) - (\pi_k^x \otimes I)(I \otimes \pi_k^y)(v - q)\|_{0,Q} \\ &\leq \|v - q\|_{0,Q} + \frac{C}{(k-1)} \|D_x(I \otimes \pi_k^y)(v - q)\|_{0,Q} \\ &= \|v - q\|_{0,Q} + \frac{C}{(k-1)} \|(I \otimes \pi_k^y)D_x(v - q)\|_{0,Q} \\ &\leq \|v - q\|_{0,Q} + \frac{C^2}{(k-1)^2} \|D_{xy}^2(v - q)\|_{0,Q}. \end{aligned}$$

Similarly, using (4.12)–(4.14) again,

$$\|\nabla v - \nabla \Pi_k v\|_{0,Q} \leq \|\nabla(v - q)\|_{0,Q} + \frac{C}{(k-1)} \|D_{xy}^2(v - q)\|_{0,Q}.$$

By equation (4.1a) of [3], we then get

$$(4.25) \quad \|v - \Pi_k v\|_{1,Q} \leq Ck^{-(r-1)} \|v\|_{r,Q}, \quad r \geq 2.$$

Interpolating between (4.24) and (4.25) gives the result.  $\square$

Using the projection  $\Pi_k$  on the reference element, we now define the operator  $\tilde{\Pi}_N : H_D^{1+\varepsilon} \rightarrow V_{p_N+m+2}^1(J_N)$  by

$$(4.26) \quad (\tilde{\Pi}_N u)|_K = (\Pi_{p_N+m+2}(u|_K \circ \vec{F}_K)) \circ \vec{F}_K^{-1} \quad \text{for all } K \in J_N.$$

(Note that (4.1)–(4.2) ensure continuity across element boundaries.) Since the mappings  $\vec{F}_K$  are affine, (4.4) implies that for  $|\alpha| \leq 1$ ,

$$(4.27a) \quad \iint_K D^\alpha(u - \tilde{\Pi}_N u)q \, d\vec{x} = 0 \quad \text{for all } K \in J_N, q \in V_{p_N+m}^0(J_N).$$

Also, taking  $\alpha = 0$  in (4.27a), we see that

$$(4.27b) \quad \iint_K c(u - \tilde{\Pi}_N u)q \, d\vec{x} = 0 \quad \text{for all } K \in J_N, c \in V_m^0(J_N), q \in V_{p_N}^0(J_N).$$

Moreover, the following lemma holds.

**Lemma 4.4.** *Let  $\tilde{\Pi}_N : H_D^{1+\varepsilon} \rightarrow V_{p_N+m+2}^1(J_N)$  be defined by (4.26). Then for any  $v \in H_D^r$ ,  $r > \frac{3}{2}$ ,  $\varepsilon > 0$ ,*

$$\|v - \tilde{\Pi}_N v\|_1 \leq Ch_N^{\min(p_N+m+2, r-1)} p_N^{-\gamma} \|v\|_r,$$

where  $\gamma$  is defined as in Lemma 4.3 and  $C$  depends on  $r, \varepsilon$ .

*Proof.* For any  $K \in J_N$ , and  $v, \tilde{v}$  related by (3.6), we have by Lemma 4.4 of [3]

$$(4.28) \quad \inf_{\tilde{w} \in Q_k(Q)} \|\tilde{v} - \tilde{w}\|_{r,Q} \leq Ch_K^{\min(k+1, r)} \|v\|_{r,K},$$

where  $\tilde{v} \in H^r(Q)$ ,  $r \geq 0$ . Hence, by Lemma 4.3 and (4.28),

$$(4.29) \quad \|\tilde{v} - \Pi_k \tilde{v}\|_{1,Q} \leq \|(\tilde{v} - \tilde{w}) - \Pi_k(\tilde{v} - \tilde{w})\|_{1,Q} \leq Ck^{-\gamma} h_K^{\min(k+1, r)} \|v\|_{r,K}.$$

Let  $k = p_N + m + 2$ . Then by (4.26) and a standard scaling result, (4.29) gives

$$\|v - \tilde{\Pi}_N v\|_{1,K} \leq Cp_N^{-\gamma} h_K^{\min(p_N+m+2, r-1)} \|v\|_{r,K}.$$

The lemma follows by squaring and summing over all  $K \in J_N$  and noting that  $h_K \leq h_N$ .  $\square$

*Proof of Theorem 3.2.* We must bound the terms in (3.5). First, we note that by the usual approximation theory for the  $h$ - $p$  version (Lemma 4.5 of [3]), with  $\mu = \min(p_N + 1, r - 1)$ ,

$$(4.30) \quad \|(I - P_N^1)\xi(U)\|_0 \leq Ch_N^\mu p_N^{-(r-1)} \|\xi(U)\|_{r-1},$$

$$(4.31) \quad \|(I - P_N^2)\eta(U)\|_0 \leq Ch_N^\mu p_N^{-(r-1)} \|\eta(U)\|_{r-1}.$$

Next, for  $U = (\vec{u}, \theta)$  we define  $W_N = (\vec{z}, \psi)$  by

$$z_i = \Pi_N u_i, \quad \psi_\alpha = \Pi_N \theta_\alpha.$$

Then, since we assume  $a_{\alpha\beta}, b_{\alpha\beta}, \Gamma_{\alpha\beta}^\gamma$  are in  $V_m^0(J_N)$ , equation (4.27a), the definitions (2.3), (2.4) of  $\Phi, \Lambda$ , and (3.2) imply that  $W_N \in Z_N(U)$ . Applying Lemma 4.4, we see that

$$(4.32) \quad \|U - W_N\|_{\mathcal{V}} \leq Ch_N^{\min(p_N+m+2, r-1)} p_N^{-\gamma} (\|\vec{u}\|_r + \|\theta\|_r).$$

Theorem 3.2 follows from (4.30)–(4.32).  $\square$

*Remark 4.2.* Instead of studying convergence in terms of Corollary 3.1 (as we did here), we could, instead, have framed our analysis in the context of mixed methods, as in [1]. For the pure  $h$  version, this would give us precisely Theorem 5 and Corollary 6 of [1], which includes estimates for the auxiliary variables in the norm  $\|\cdot\|$  from [1].

#### ACKNOWLEDGEMENT

The author would like to thank an anonymous referee for the improvement in Lemma 4.3 when  $r \geq 2$ .

#### REFERENCES

1. D. N. Arnold and F. Brezzi, *Locking-free finite element methods for shells*, Math. Comp. **66** (1997), (to appear).
2. J.-P. Aubin, *Applied Functional Analysis*, Wiley, New York, 1979. MR **81a**:46083
3. I. Babuška and M. Suri, *The  $h$ - $p$  version of the finite element method with quasiuniform meshes*, RAIRO Modél. Math. Anal. Numér. **21** (1987), 199–238. MR **88d**:65154
4. ———, *On locking and robustness in the finite element method*, SIAM J. Numer. Anal. **29** (1992), 1261–1293. MR **94c**:65128
5. M. Bernadou, P. G. Ciarlet and B. Miara, *Existence theorems for two-dimensional linear shell theories*, J. Elasticity **34** (1994), 111–138. MR **95f**:73050
6. C. Bernardi and Y. Maday, *Approximations spectrales de problèmes aux limites elliptiques*, Springer-Verlag, Berlin and New York, 1992. MR **94f**:65112
7. F. Brezzi and K. J. Bathe, *Studies of finite element procedures—the inf-sup condition, equivalent forms and applications*, Reliability of Methods for Engineering Analysis (I. J. Bathe and D. R. J. Owen, eds.), Pineridge Press, Swansea, 1986, pp. 1787–1801.
8. F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer-Verlag, New York-Heidelberg-Berlin, 1991. MR **92d**:65187
9. P. M. Naghdi, *Foundations of elastic shell theory*, Progress in Solid Mechanics, Vol. 4 (I. N. Sneddon and R. Hill, eds.), North-Holland, Amsterdam, 1963, pp. 1–90. MR **29**:790
10. ———, *The theory of shells and plates*, Handbuch der Physik, Vol. VI a-2, Springer-Verlag, New York-Heidelberg-Berlin, 1972, pp. 425–640.
11. J. Piila and J. Pitkäranta, *Energy estimates relating different linear elastic models of a thin cylindrical shell*. I, II, I: SIAM J. Math. Anal. **24** (1993), 1–22, II: SIAM J. Math. Anal. **26** (1995), 820–849. MR **94a**:73032; MR **96e**:73068

12. J. Pitkäranta, *The problem of membrane locking in finite element analysis of cylindrical shells*, Numer. Math. **61** (1992), 523–542. MR **93b**:65178
13. J. Pitkäranta and M. Suri, *Design principles and error analysis for reduced-shear plate bending finite elements*, Numer. Math. (to appear)
14. L. R. Scott and M. Vogelius, *Norm estimates for a maximal right inverse of the divergence operator in spaces of piecewise polynomials*, RAIRO Modél. Math. Anal. Numér. **19** (1985), 111–143. MR **87i**:65190
15. R. Stenberg and M. Suri, *Mixed  $hp$  finite element methods for problems in elasticity and Stokes flow*, Numer. Math. **72** (1996), 367–389. CMP 96:06
16. M. Suri, I. Babuška and C. Schwab, *Locking effects in the finite element approximation of plate models*, Math. Comp. **64** (1995), 461–482. MR **95f**:65207

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF MARYLAND BALTIMORE  
COUNTY, 1000 HILLTOP CIRCLE, BALTIMORE, MARYLAND 21250  
*E-mail address:* `suri@umbc2.umbc.edu`