Title

The Extended-Stencil Finite Element Method

Permalink

https://escholarship.org/uc/item/01j9q57h

Author

Brinkman, Leah Ann

Publication Date

2023

Peer reviewed|Thesis/dissertation

The Extended-Stencil Finite Element Method

By

LEAH A. BRINKMAN
DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Civil & Environmental Engineering

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

_____

Mark M. Rashid, Chair

_____

Boris Jeremić

_____

John B. Bolander

Committee in Charge

2023

ABSTRACT

The Extended-Stencil Finite Element Method

Leah A. Brinkman

This dissertation introduces a novel finite element method called the extended-stencil finite element method (ESFEM). The ESFEM uses conventional finite element meshes to produce moderately high order polynomial element interpolants. The "nodal stencil" of the ESFEM is extended such that elements reference nodal values of other elements. This extension increases the nodal data available to an element. On each face of the element, a polynomial fit is performed to the nodal data. These "face polynomials" are then used as constraints in the polynomial fit of the element interpolant to the nodal data. Because the element interpolants are formulated independently from other element interpolants, the ESFEM is generally nonconforming. However, convergence is achieved by constraining the interpolant to ensure passage of the so called F-E-M-Test. The ESFEM was implemented into a finite element code base, and the results of the numerical examples show improved accuracy over the conventional finite element method in problems exhibiting shear locking, volumetric locking, and mesh distortion. In particular, the ESFEM enables an efficient use of a mesh's degrees of freedom by using a polynomial fit to form the element interpolants.

ACKNOWLEDGEMENTS

This body of text is not just a creation of my own efforts. Rather, it was only made possible through the support of others throughout my years. I have had the pleasure of interacting with so many wonderful, supporting people throughout my life. Not all can be mentioned here, but I am eternally grateful for all who have contributed to developing me as a person, and as a student.

My mother, father, and sister all contributed to my drive and set up for educational success. My family always was supportive of whatever life path I wanted to go down. And my sister paved the way to higher education, achieving her Master's. This helped me see that school beyond an undergraduate education was attainable. In addition to my family, my close friends have supported me as I traversed through the stressors of higher education.

As a GSR and intern, I acknowledge the support and mentorship provided by the Sandia National Laboratory team 1542. Thank you to Professors John Bolander and Boris Jeremić for serving on my qualifying examination and dissertation committees. Moreover, I acknowledge the support of the Civil & Environmental Engineering department at UC Davis as a whole.

Finally, I acknowledge the guidance and support of my advisor, Professor Mark Rashid, who has provided valuable mentorship throughout my PhD program. He has always been willing to take time to explain concepts thoroughly, which I believe has contributed immensely to the knowledge I've gained over the past few years.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The finite element method (FEM) has undergone much algorithmic and mathematical development, allowing extensive usage in engineering today. In the early days, Hrennikoff's framework of bar elements is an example of the segmentation of a material into individual elements to estimate solutions to elastic problems [14]. Today, the FEM has extended past structural mechanics into areas of electromagnetics, fluid flow, heat conduction, and flow through porous media. Even with these advancements, there are still some problems that remain difficult to resolve. In the realm of solid mechanics, chief among the remaining challenges are contact between two different moving bodies, highly nonlinear problems, and locking.

The problem of interest in this dissertation is locking, which arises in two main circumstances: bending of thin structural elements, and near-incompressible materials. The former is referred to as shear locking and the latter, volumetric locking. Locking presents itself as large errors in the displacement solution that can often be multiple orders of magnitude less than what would occur in reality. In Section 1.1, an example of locking in a cantilever beam is presented. Moreover, Chapter 3 discusses locking and strategies for its mitigation more in depth.

In this dissertation, the extended-stencil FEM (ESFEM) is introduced to increase the

polynomial order of the basis functions such that locking is mitigated while at the same time avoiding rank deficiency. The full scope of the ESFEM's potential has yet to be determined and this work presents mainly the method's capabilities involving shear and volumetric locking.

1.1 Illustrative Example

This section investigates a cantilever beam with different element types to illustrate locking in the Conventional Finite Element Method (CFEM). "CFEM" is utilized in this dissertation in order to describe the FEM whose origins are described in [10] and is summarized in Chapter 2. Generally, shape functions of the CFEM are mapped polynomials of the lowest possible order such that the Kronecker-delta property holds on the element's nodes. Now, the crucial distinction between the CFEM and the ESFEM is the sets of nodes associated with the element. In the ESFEM, a larger set of nodes is used than the CFEM, which requires a higher polynomial order. For an 8-node hexahedral (hex-8) element in the CFEM, shape functions are trilinear while the 27-node hexahedral (hex-27) element shape functions are tri-quadratic. It is commonly thought that a hex-27 can be used to mitigate locking, however the example below reveals a distorted hex-27 mesh does not mitigate locking. Please note that this example is not meant to be an in depth analysis of shear locking. Rather, it is intended to illustrate the phenomenon of locking.

Consider a cantilever beam with a square cross-section subject to a traction force at the end as shown in Fig. 1.1. A mesh of 1x1x10 hex-27 elements is used. With undistorted elements and $\nu = 0.3$, the hex-27 shows no locking, as the quadratic basis functions are easily able to capture the bending deformation. This results in a displacement of node A to be 0.984 (1.6% error). However, when the elements' vertex nodes along the long edges are randomly moved by a distance of up to 0.4 to introduce distortion, we see the return of shear locking, with node A displacing only 0.402 (59.8% error).
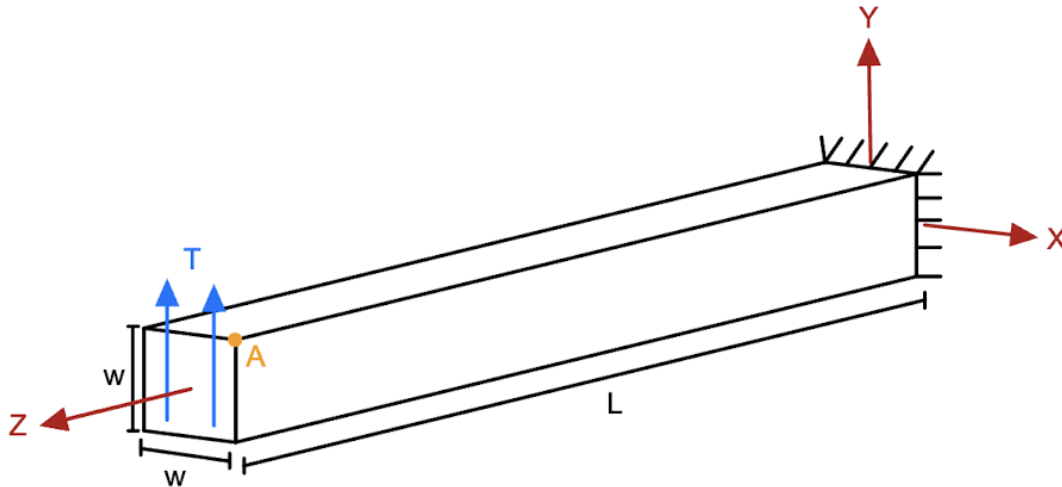
Figure 1.1: Square cross-section cantilever beam with dimensions w = 2 and L = 100. Traction T = 1000 units force/area on the end face in the y-direction and the other end fixed. All values are in consistent units.
Material properties: E = 10^9, ν = 0.3 - 0.4999
Theoretical Displacement of Node A: 1.0 in the Y-direction

Now, consider this same mesh but with ν = 0.4999, which corresponds to a nearly-incompressible material. For an undistorted mesh, node A displaces 0.982 (1.8% error) and with distortion, node A displaces 0.257 (74.3% error). This indicates the addition of not only shear locking, but volumetric locking as well for the distorted mesh.

Comparing the displacement of node A in a hex-27 mesh to a hex-8 mesh, the displacement is a significantly smaller value, with the undistorted mesh displacing 0.099 (90.1% error) and 0.115 (88.5% error) for ν = 0.3 and 0.4999, respectively. Furthermore, the distorted hex-8 mesh displacements are 0.031 (96.9% error) and 0.035 (96.5% error) for ν = 0.3 and 0.4999, respectively. See Table 1.1 for a summary.

Displacement of Node A							
Hex-8				Hex-27			
undistorted		distorted		undistorted		distorted	
ν = 0.3	0.4999	0.3	0.4999	0.3	0.4999	0.3	0.4999
0.099	0.115	0.031	0.035	0.984	0.982	0.402	0.257

Table 1.1: Displacements of node A in Fig. 1.1 for distorted and undistorted hex-8 and hex-27 meshes with different values of Poisson's ratio.

These values show that the use of hex-8 elements results in the presence of shear locking. This is expected for hex-8 elements in this mesh since there is only one element through the thickness, and the basis functions are tri-linear. While the consideration of a nearly-incompressible material actually shows a slight increase in displacement for both distorted and undistorted meshes, they are still highly inaccurate results due to the severe locking. However, it is interesting that the hex-27 elements under these circumstances appear to have compounding effects of both shear locking and near-incompressibility. This problem displays that, while an undistorted hex-27 element may not lock, there is still potential for locking to occur if the element is distorted. Meshing flexibility is a critical aspect of the FEM's overall utility. And, an element's allowance of mesh distortion is how such flexibility is obtained.

1.2   Shear Locking

Shear locking is identified in displacement results that are markedly diminished, frequently by an order of magnitude or more, in comparison to the exact solution. This phenomenon manifests in thin structures (such as beams, plates, and shells) that have been discretized using conventional continuum elements. Furthermore, it occurs when the elements themselves are thin and the through-thickness discretization is coarse (e.g. one element through the thickness).

Shear locking has plagued the FEM since the beginning, however it is fairly sparsely treated in the literature, perhaps surprisingly so. For example, in an account of the 80-year history of the FEM, shear locking was not mentioned once [22]. This may be due, in part, to the fact that it is not well understood mathematically and can be hard to see without experience. Xia et al. remarked that "[i]t is interesting to observe that if one looks only at the calculated deflection of the beam, it offers no clue that the solution is not correct." [39] Without careful and experienced examination of the FEM results, shear locking can be overlooked and lead ultimately to wrong conclusions and poor engineering judgements.

In practice, three to four elements through the thickness, along with aspect ratios below 2-3, is often sufficient to avoid shear locking in hex-8 elements. However, this can lead to excessive numbers of elements when attempting to alleviate shear locking with mesh refinement alone. So-called structural elements, e.g. shell or beam elements, are often preferred, though they present challenges of their own. Shell element nodes have rotational degrees of freedom which complicates mesh generation when shell and continuum elements must coexist in the same mesh. The extra rotational degrees of freedom at the interface of the two element types may have to be set manually to ensure proper connection. Moreover, in simulations involving sheet metal forming, it has been found that shell elements fail to correctly capture the actual stress state in circumstances where the bending radii of the metal is comparable with the sheet thickness, and thus continuum elements must be used [21, 39, 42].

Another element specifically designed to overcome shear locking is the so-called solid-shell element as presented in [7, 8]. In contrast to shell elements, solid-shell elements only contain displacement degrees of freedom at the nodes, such that the degrees of freedom match a hex-8 continuum element. Due to the similar degrees of freedom at the nodes, solid-shell elements can easily connect to continuum elements, unlike shell elements. The "shell" part of a solid-shell element relates to the restrictions applied that enforce constant strain throughout the thickness. Consequently, as noted in [8], this brings about volumetric and Poisson locking that is addressed by supplementing with methods such as the B-bar method [7] and the enhanced assumed strain (EAS) method [8].

Upon a comprehensive analysis of recent literature, the potential causation of locking phenomena has been correlated with a solution subspace incapable of reproducing the exact polynomial order of the strain solution. Analyses of this form have been explored in various studies employing EAS techniques and other methodologies. These investigations seek to expand the solution subspace through the incorporation of additional displacement-like degrees of freedom. The additional degrees of freedom exist to enhance the strain field by increasing

its polynomial order [3, 4, 6, 8]. The overall objective herein is to employ a mathematical subspace analysis to determine the optimal count of EAS parameters necessary for the mitigation of shear locking, while introducing a minimal number of supplementary degrees of freedom. These extra degrees of freedom have no physical significance beyond granting increased flexibility to the strain field. Furthermore, sensible contemplation is important to ensure the minimal inclusion of degrees of freedom, a process that involves intricacies and therefore potential cost detriments.

The intricacies of finding the "proper number" of degrees of freedom to mitigate locking arises with the potential occurrence of rank deficiency in the stiffness matrix. Rank deficiency refers to a mathematical property of a matrix where its rank, denoting the maximum count of linearly independent rows or columns, falls short of its overall count of rows or columns. In other words, a matrix is rank deficient when it fails to have the full number of independent dimensions that are theoretically possible for a matrix of its size. In the case of the FEM, the stiffness matrix is often rank deficient in cases where an element's integration rule is unable to accommodate the number of degrees of freedom.

Further, the integration rule implemented on an element is an important factor in the flexibility of the displacement field. Integration rules utilize integration points to partition an element by defining where the governing equations are to be evaluated, which inform the element's strain field. Furthermore, integration rules dictate the locations and weights of the integration points, influencing the accuracy and representation of strain behavior throughout the element's domain. Shear locking often occurs in thin structures with higher numbers of integration points, due to the "over-sampling" of the strain field. Over-sampling entails the addition of an excess number of integration points beyond the minimum required to adequately represent the solution. In contrast, rank deficiency arises with an "under-sampling", where there are not enough integration points to adequately represent the solution.

Evaluation of the literature reveals formulations that provide locking mitigation, however introduce rank deficiency due to the under-sampling of the stiffness matrix. For example,

in reduced integration (RI) and selective-reduced integration (SRI) [7, 16, 25, 42, 43], terms within the stiffness matrix are under-sampled as a way to "relax" the elements such that locking is mitigated. However, the presence rank deficiency and spurious modes under bending dominated deformation, causes need for hourglass control [33, 38, 42]. Hourglass control requires stability parameters developed through testing and assumptions to provide the solution stability. To help mitigate locking without risk of rank deficiency or the need for stabilizing factors, the basis functions themselves must be reworked in order to allow for higher order interpolation and therefore better handle locking-prone circumstances.

Hughes et al. have investigated isogeometric analysis (IGA) with NURBS [18] to increase the polynomial order of basis functions. While these methods increase the polynomial order of the basis functions, and work to better align CAD geometries to FEA geometries, it has been shown to not mitigate locking [28]. Thus, IGA must be combined with other methods to alleviate locking [4, 6]. This implies that increasing the polynomial order of the basis functions does not offer a definitive resolution, so further exploration is required.

1.3   Volumetric Locking

Volumetric locking emerges in situations where the material's behavior is either incompressible or exhibits near-incompressibility. This phenomenon occurs due to the imposed requirement that deformation at the integration points remains isochoric or nearly so, consequently leading to an excessive constraint on the element's overall deformation behavior. It is noteworthy that volumetric locking is better-known and understood, as compared to shear locking. Increasing the number of integration points increases the isochoric constraint in the presence of incompressible material behavior, and therefore increases the volumetric locking potential. Conversely, employing too few integration points leads to a state of "under-integration," a condition characterized by an inadequate number of integration points to sufficiently capture the true behavior of the simulation. This can lead to a deficiency in

7

the rank of the element's stiffness matrix that decreases the accuracy and stability of the numerical solution.

As an attempt to mitigate volumetric locking, some methods employ additional degrees of freedom, as in EAS methods [6, 7, 19, 31, 32]. However, if too many degrees of freedom are added, this brings the element into a rank-deficient state that gives way to zero-energy modes. Striking the balance between an appropriate number of integration points, the nature of the material, and the specific element formulation is an essential consideration avoiding volumetric locking and rank deficiency.

1.4   The Extended-Stencil Finite Element Method

As with the conventional FEM, the extended-stencil FEM (ESFEM) discretizes the problem domain into non-overlapping elements, each associated with a set of nodes. These nodes hold displacement degrees of freedom, and are associated with corresponding shape functions. In fact, ESFEM meshes can be, and within this work are, identical to CFEM meshes consisting of quasi-linear elements. However, the essence of the ESFEM is that it allows typically quasi-linear elements of the CFEM to receive nodal data from adjacent elements. This nodal data is then utilized to formulate shape functions via polynomial fitting.

As a result of the augmented nodal data available to an element, the polynomial fit allows for higher-order interpolation of the displacement solution. Moreover, through experimentation, the ESFEM displays favorable locking mitigation capabilities. From what has been gleaned from literature and experimentation, the ESFEM's locking mitigation tendencies stem from the implementation of full integration, in combination with a prudent increase in the degrees of freedom associated with the system, if any increase is to take place. Consider a problem subject to a pure bending load (i.e. a cantilever beam with an end moment). The exact displacement solution for pure bending, in the context of linear elasticity, is quadratic. Now, analyzing a hex-8 element in the context of the CFEM, the

shape functions are all quasi-linear. Therefore, under pure bending, the linear functions are not capable of representing a quadratic function fully. Additionally, it was shown that when distortion is introduced, hex-27 meshes may result in the presence of shear and/or volumetric locking. This is where the ESFEM is demonstrating potential value, giving higher prospects of locking mitigation by increasing an element's nodal dependencies. Moreover, it will be shown that all this is achieved without the need for artificial stability treatments, such as hourglass stiffness.

1.5 ESFEM: Novelty & Significance

The novelty in the ESFEM stems from the difference in the shape functions. The shape functions in the conventional finite element method are defined based on the element type (i.e. hex-8 versus hex-27). In the ESFEM, the shape functions are obtained by performing a polynomial fit to the nodal data referenced by an element. Additionally, the ESFEM elements reference nodal data of other elements to better obtain higher-order polynomial functions than conventional shape functions provide. The ESFEM methodology can obtain these higher-order polynomial shape functions using the same conventional FEM mesh.

The main objective of the ESFEM development has been to mitigate shear and volumetric locking that often occurs with conventional finite element meshes. As shown in the numerical examples of Chapter 5, the ESFEM hex-8 and 4-node tetrahedral (tet-4) elements are capable of converging to the theoretical displacement solution at less refined meshes than their conventional counterparts. A resulting significance of the ESFEM is the ability to obtain accurate nodal value results with plate, shell, and beam meshes containing maximum aspect ratios greater than 3, and one element through the thickness. By decreasing the aspect ratio and layers of elements needed for accurate results, time required for mesh generation decreases. This, balanced with decreased CPU time for smaller meshes, facilitates improved cost-effectiveness for the ESFEM.

Another significant improvement is the ESFEM's robust handling of mesh distortion. As discussed in Section 1.1, distortion may reintroduce locking to a hex-27 mesh where its undistorted counterpart does not exhibit locking. In Chapter 5, the ESFEM elements are shown to still produce accurate results, even with distorted meshes.

1.6   Outline

In this dissertation, the ESFEM will be presented, explained, and tested rigorously. In Chapter 2, the CFEM will be summarized to introduce the notational framework that will then be used to explain the methodology behind the ESFEM. Chapter 3 then gives an in depth account of existing locking mitigation methods. Chapter 4 provides a full explanation of the ESFEM with particular emphasis on its performance in relation to locking scenarios. Then, Chapter 5 will give numerical exploration of the ESFEM to demonstrate its locking mitigation capabilities. Finally, in Chapter 6, the future of the ESFEM will be discussed to speculate potential areas of continued exploration.

Chapter 2

The Conventional Finite Element Method

This dissertation utilizes the conventional finite element method (CFEM) in accordance with the description found in [10], implemented in a finite deformation setting. This chapter is dedicated to a comprehensive explanation of the CFEM, serving to establish a notational framework and foundation that will be useful in later chapters. Subsequently, Chapter 4 extends the conventional framework via the the extended-stencil finite element method (ES-FEM), which is the main contribution of this dissertation.

The CFEM is a useful approach for approximating solutions to boundary value problems (BVPs) that are too complex to solve analytically. The technique involves partitioning a body into smaller pieces, termed elements, thereby facilitating the development of a practical solution strategy. These elements consist of a defined arrangement of nodes, which hold essential information pertaining to material deformation. The integration points, on the other hand, quantify the material state throughout the body, and facilitate evaluation of the weak-form integrals over the problem domain.

Within this chapter, finite deformation kinematics are utilized to set up the deformation gradient tensor, which described the deformation of the body. Moreover, the boundary

value problem is described for quasi-static loading, and the strong form of the problem is developed. Subsequently, the weak form problem statement is formulated with a Galerkin approximation to allow for numerical analysis to solve the unknowns. Elements and their properties in the CFEM are defined. Further, the numerical integration technique, Gaussian quadrature, is defined. To conclude the chapter, limitations of the CFEM are discussed.

2.1   The Deformation Gradient Tensor

To establish the formulation of the CFEM within the scope of this dissertation, some elements of finite-deformation continuum mechanics are needed. This section defines the deformation gradient denoted as $\mathbf{F}$, which serves as the basis for describing the deformation of a body. Other, related kinematical quantities are defined as well.
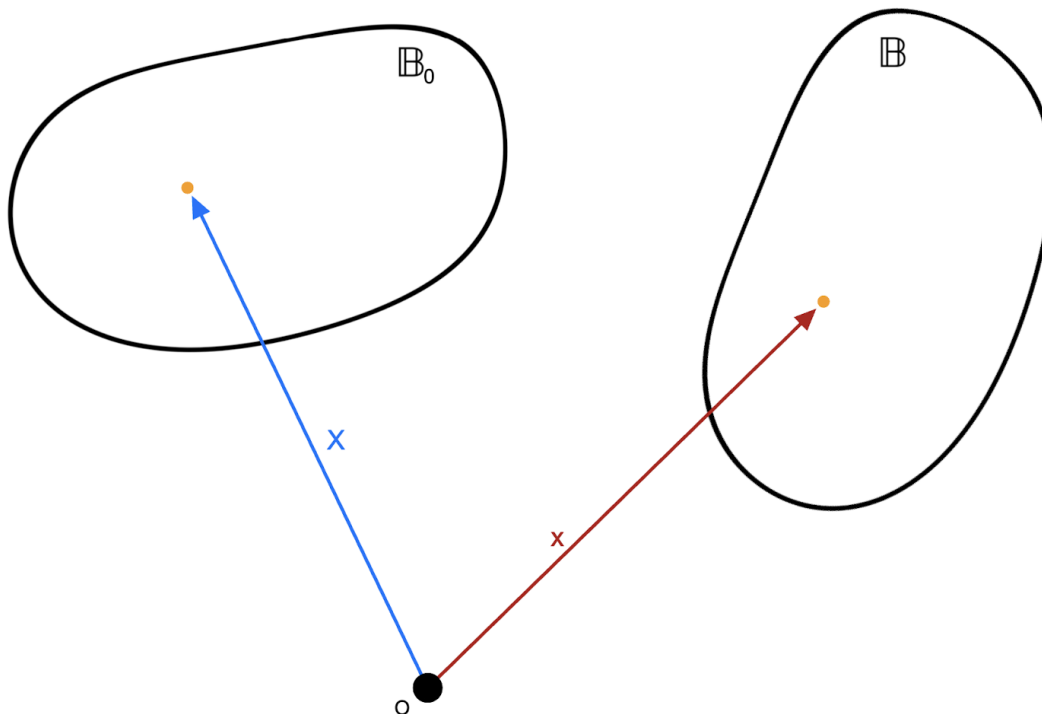
Figure 2.1: Deformation map from the reference configuration to the current configuration.

Consider the initial, or reference, configuration of the body denoted as $\mathbb{B}_0$ at time $t = 0$.

Through a time sequence of deformations, it transforms into the configuration of the body $\mathbb{B}$ at a subsequent time $t > 0$, as visually represented in Fig. 2.1. The process involves the establishment of a deformation map, characterized by the function $\mathbf{x}\left(\mathbf{X}, t\right)$, facilitating the mapping from the reference configuration $\mathbb{B}_0$ to the current configuration $\mathbb{B}$. In this context, the deformation gradient tensor, $\mathbf{F}$, is defined by Eq. 2.1. This tensor defines the spatial relationship between infinitesimal material line elements in the reference configuration $\mathbb{B}_0$ and their corresponding counterparts in the deformed configuration $\mathbb{B}$, accounting for both length and orientation changes.

$$\mathbf{F} = \frac{\partial \mathbf{x}}{\partial \mathbf{X}} \tag{2.1}$$

Furthermore, it is recognized that the ratio of volumes between two material volume elements, one originating from $\mathbb{B}_0$ and the other, from $\mathbb{B}$, is governed by the determinant of the deformation gradient tensor $\mathbf{F}$. This volumetric change, symbolized as J, defines the change in volume and is expressed through the relationship in Eq. 2.2.

$$J = \det\left(\mathbf{F}\right) = \frac{\text{volume in } \mathbb{B}}{\text{volume in } \mathbb{B}_0} \tag{2.2}$$

The utilization of the polar decomposition theorem affords a decomposition of the deformation gradient tensor $\mathbf{F}$ into pure rotation and pure stretch components, as depicted in Eq. 2.3. Herein, $\mathbf{U}$ and $\mathbf{V}$ emerge as positive definite symmetric tensors, while $\mathbf{R}$ denotes a proper orthogonal tensor corresponding to rotations. Specifically, $\mathbf{U}$ and $\mathbf{V}$ are termed the right and left stretch tensors, respectively, with $\mathbf{R}$ serving as the rotation tensor. Eq. 2.3 then signifies that the overall local deformation described by $\mathbf{F}$ can be decomposed into a sequence involving either a pure stretch followed by a rigid rotation or vice versa.

$$\text{Right Decomposition } \mathbf{F} = \mathbf{R}\mathbf{U}$$
$$\text{Left Decomposition } \mathbf{F} = \mathbf{V}\mathbf{R} \tag{2.3}$$

Furthermore, for the subsequent sections, the left Cauchy-Green tensor, $\mathbf{B}$, is needed as

defined below:

$$\mathbf{B} = \mathbf{F}\mathbf{F}^{T} = \mathbf{V}^2 \tag{2.4}$$

2.2 A Static Boundary Value Problem

2.2.1 Eulerian Frame of Reference
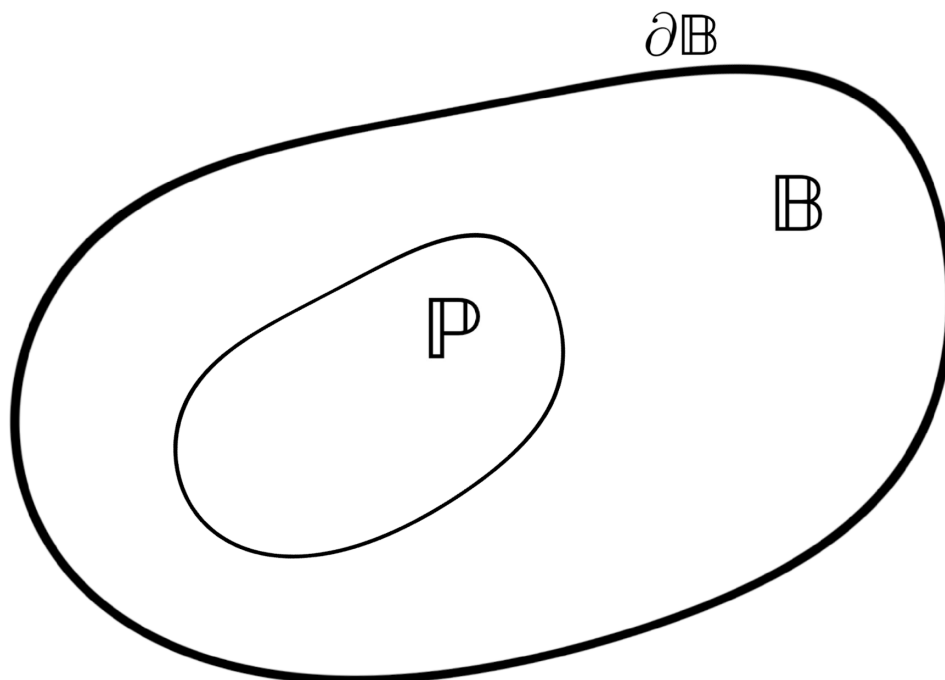
Figure 2.2: Continuum potato body $\mathbb{B}$ with the boundary $\partial\mathbb{B}$ and the subset $\mathbb{P} \subset \mathbb{B}$

To commence, the scope of this dissertation resides within the domain of solid mechanics, specifically focusing on scenarios characterized by quasi-static loading. Within the framework of an Eulerian reference frame, the central interest revolves around a boundary value problem (BVP) featuring a body in space, designated as $\mathbb{B}$, herein termed a "continuum potato." Force equilibrium can be stated as follows. Consider a smaller section, $\mathbb{P}$, encapsulated within $\mathbb{B}$,

as illustrated in Fig. 2.2. This configuration is then subject to the following:

$$\int_{\partial \mathbb{P}} \mathbf{t}\left(\mathbf{x}, \mathbf{n}\right) \, da + \int_{\mathbb{P}} \rho \, \mathbf{b}\left(\mathbf{x}\right) \, dv = \mathbf{0} \qquad \forall \quad \mathbb{P} \subset \mathbb{B} \tag{2.5}$$

Eq. 2.5 represents the equilibrium of forces on the smaller continuum potato $\mathbb{P}$, where $\mathbf{t}\left(\mathbf{x}, \mathbf{n}\right)$ is the traction (force per unit current-configuration area) on the boundary $\partial \mathbb{P}$ that has the unit normal $\mathbf{n}$. Moreover, ρ is the density (mass per unit current-configuration volume) of the material, and $\mathbf{b}\left(\mathbf{x}\right)$ is the body force per unit mass, such as gravity. Manipulating Eq. 2.5 using the divergence theorem and localization theorem, Eq. 2.5 results in the local statement

$$\nabla \cdot \mathbf{T} + \rho \, \mathbf{b} = \mathbf{0} \qquad \forall \quad \mathbf{x} \in \mathbb{B}. \tag{2.6}$$

Herein, $\mathbf{T}$ is the Cauchy stress tensor, related to the traction by Eq. 2.7. Note that $\mathbf{n}$ is the same unit normal vector as defined above.

$$\mathbf{t} = \mathbf{T}\mathbf{n} \tag{2.7}$$

Following the establishment of force equilibrium, a complete statement of the strong form BVP necessitates a description of boundary conditions (BCs). BCs involve displacement and traction boundary conditions, which must not overlap one another as shown in Fig. 2.3. Furthermore, for an isothermal material subject to quasi-static loading, Eqs. 2.8 and 2.9 define the BVP strong form for this dissertation along with that, in the most general case, $\mathbf{T}$ is a local function of the deformation-gradient ($\mathbf{F}$) history up to the current time t. Eq. 2.8 is written in index notation with i and j spanning the values 1, 2, and 3 to represent the Cartesian coordinate directions in three-dimensional space.

$$\text{Equilibrium Equation} \quad \left. T_{ij,j} + \rho \, b_i = 0 \right\} \quad \forall \quad \mathbf{x} \in \mathbb{B} \tag{2.8}$$

15
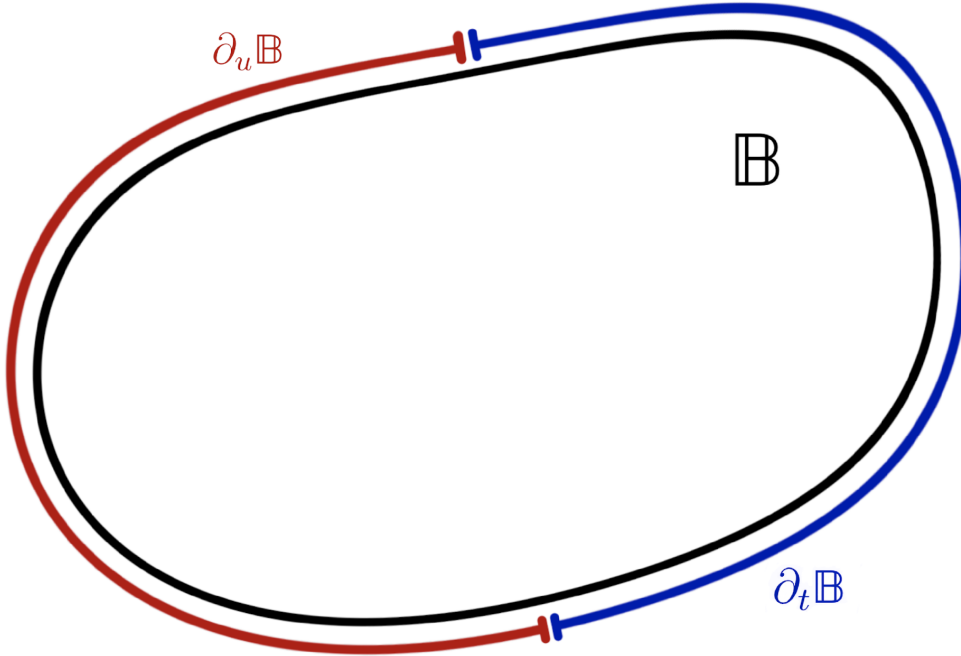
Figure 2.3: Boundaries on the continuum potato split into two sections: one containing displacement boundary conditions ($\partial_u \mathbb{B}$), and the other the traction boundary conditions ($\partial_t \mathbb{B}$).

$$\text{Boundary Conditions} \qquad \begin{aligned} \mathbf{u} &= \overline{\mathbf{u}} \quad on \quad \partial_u \mathbb{B} \\ \mathbf{t} &= \overline{\mathbf{t}} \quad on \quad \partial_t \mathbb{B} \\ \partial_u \mathbb{B} &\cap \partial_t \mathbb{B} = \emptyset \\ \partial_u \mathbb{B} &\cup \partial_t \mathbb{B} = \mathbb{B} \end{aligned} \qquad (2.9)$$

2.2.2 Lagrangian Frame of Reference

In the context of solid mechanics, selection of the Lagrangian frame of reference is motivated by the necessity to track the trajectories of individual material points as they traverse through space. Unlike the Eulerian frame of reference in which field quantities are regarded to be functions of points in space, the Lagrangian frame of reference facilitates the continuous tracking of specific material points as they experience spatial displacement. Moreover, the preference for Lagrangian reference frames in solid mechanics flows from the fact that, for

a solid, the stress response is a function of the deformation history for a particular material particle. And in FEM, we need the stress response at the integration points (IPs). If the IPs coincide with fixed material particles for all time, then the calculation of the stress state via time-integration of complex constitutive models becomes an entirely local process. In fluid mechanics, on the other hand, the stress state is a function of the local, instantaneous deformation rate; not of the entire history of the deformation experienced by the material up to the current time. This is the key distinction between the two reference frames, and why Lagrangian is favored in solid mechanics.

To commence, consider first the first Piola-Kirchhoff stress, as defined in Eq. 2.10. Herein, $\mathbf{T}$ symbolizes the Cauchy stress tensor, conventionally used within the Eulerian frame of reference. Note that J and $\mathbf{F}$ are defined in Section 2.1.

$$\mathbf{P} = J\,\mathbf{T}\,\mathbf{F}^{-1} \tag{2.10}$$

An equivalent equation of motion within the Lagrangian framework, comparable to the Eulerian equation defined in Eq. 2.8, can be derived using analogous procedures. In this context, the derivation centers on the utilization of the reference configuration $\mathbb{B}_0$, whose current configuration is $\mathbb{B}$. The outcome of this derivation is Eq. 2.11, where the indices i and j span the Cartesian coordinate axes in three-dimensional space. The fundamental components of this formulation encompass the Piola-Kirchhoff stress tensor $\mathbf{P}$, the reference configuration coordinates $\mathbf{X}$, the mass per unit volume of $\mathbb{B}_0$ ρ_0, and the body force per unit mass $\mathbf{b}$. Additionally, $\mathbf{p} = \mathbf{P}\mathbf{N}$ is the Piola traction vector, where $\mathbf{N}$ is the unit normal vector of material surfaces in the reference configuration, upon which the traction is applied. $\mathbf{p}$ is related to the Cauchy traction vector $\mathbf{t}$ by Eq. 2.13, wherein da is an infinitesimal area in the current configuration $\mathbb{B}$, and dA is an infinitesimal area in the reference configuration $\mathbb{B}_0$.

$$\text{Equilibrium Equation} \quad \left.\frac{\partial P_{ij}}{\partial X_j} + \rho_0\,b_i = 0\right\} \quad \forall \quad \mathbf{X} \in \mathbb{B}_0 \tag{2.11}$$

$$\mathbf{u} = \overline{\mathbf{u}} \quad on \quad \partial_u \mathbb{B}_0$$

Boundary Conditions $\qquad \mathbf{p} = \overline{\mathbf{p}} \quad on \quad \partial_p \mathbb{B}_0 \qquad\qquad (2.12)$

$$\partial_u \mathbb{B}_0 \cap \partial_p \mathbb{B}_0 = \emptyset \partial_u \mathbb{B} \cup \partial_p \mathbb{B} = \mathbb{B}$$

$$\mathbf{p}\,dA = \mathbf{t}\,da \qquad\qquad (2.13)$$

2.3   The Galerkin Approximation

In order to practically solve complex boundary value problems, a transformation is applied to the strong form boundary value problem (BVP) to develop the corresponding weak form. The result of this transformation is outlined in Eq. 2.14, wherein the trial solution is represented by the vector field denoted as $\mathbf{u}$, and the test function is represented by $\mathbf{v}$. It is crucial to recognize that the the first Piola-Kirchoff stress, $\mathbf{P}$, depends on the trial solution, $\mathbf{u}$, through the material's constitutive relation. The weak form is the basis for powerful and flexible approximation methods in solid mechanics, the FEM chief among them.

$$\text{Find } \mathbf{u} \in \mathcal{S} = \{\mathbf{w} \mid w_i \in C^0\left(\mathbb{B}_0\right), w_i = \overline{u}_i \text{ on } \partial_u \mathbb{B}_0\} \text{ such that:}$$
$$\int_{\mathbb{B}_0} P_{ij} \frac{\partial v_i}{\partial X_j} dV = \int_{\partial_p \mathbb{B}_0} \overline{p}_i v_i dA + \int_{\mathbb{B}_0} \rho_0 b_i v_i dV \qquad (2.14)$$
$$\forall\, \mathbf{v} \in \mathcal{V} = \{\mathbf{w} \mid w_i \in C^0\left(\mathbb{B}_0\right), w_i = 0 \text{ on } \partial_u \mathbb{B}_0\}$$

To establish the Galerkin approximation, the construction of a finite element mesh and the derivation of suitable basis functions are necessary steps in the process of approximating the solution. The FE mesh is used to generate finite-dimensional subspaces of the trial-solution and test-function spaces, denoted as $\mathcal{S}^h \subset \mathcal{S}$ and $\mathcal{V}^h \subset \mathcal{V}$, respectively. The mesh consists of non-overlapping elements, each of which is associated with a small set of nodes. The nodes carry the unknowns of the approximation problem as well as basis functions. For the CFEM, the basis functions ϕ_a are defined at each node a, and satisfy the following properties:

1. Kronecker Delta Property: $\phi_a\left(\mathbf{x}_b\right) = \delta_{ab}$

2. Compact Support: $\phi_a \neq 0$ only on elements containing node a

3. $\phi_a \in C^0\left(\mathbb{B}_0\right)$

With the foundational principles in place, the formulation of the Galerkin approximation to the weak form can now be developed. The Galerkin approximation is dependent upon the definition of approximations to both the trial solution, denoted as $\mathbf{u}^h \in \mathcal{S}^h$, and the test function, represented as $\mathbf{v}^h \in \mathcal{V}^h$. These approximations are structured as linear combinations encompassing the set of basis functions ϕ_a, as expressed in Equation 2.15, where N signifies the count of nodes present within the finite element mesh.

$$
\begin{aligned}
\mathbf{u}^h &= \sum_{a=1}^{N} \mathbf{u}_a \phi_a\left(\mathbf{x}\right) \\
\mathbf{v}^h &= \sum_{b=1}^{N} \mathbf{v}_b \phi_b\left(\mathbf{x}\right)
\end{aligned}
\tag{2.15}
$$

Subsequently, the set of nodes in the mesh can be broken up into two subsets, A_0 and A_u, such that:

- A_u represents the subset of nodes on $\partial_u \mathbb{B}$

- $A_u \cup A_0 = \{1, 2, 3, ..., N\}$

- $A_u \cap A_0 = \emptyset$

Utilizing these subsets, along with information known about $\mathbf{u}^h$ and $\mathbf{v}^h$ on $\partial_u \mathbb{B}$, Eq. 2.15 can be written as shown in Eq. 2.16.

$$
\begin{aligned}
\mathbf{u}^h &= \sum_{a \in A_0} \mathbf{u}_a \phi_a\left(\mathbf{x}\right) + \sum_{a \in A_u} \bar{\mathbf{u}}_a \phi_a(\mathbf{x}) \\
\mathbf{v}^h &= \sum_{b \in A_0} \mathbf{v}_b \phi_b\left(\mathbf{x}\right)
\end{aligned}
\tag{2.16}
$$

This manipulation facilitates a transformation of the weak form, yielding a reformulated

expression as demonstrated in Eq. 2.17.

$$\text{Find } \mathbf{u}_a, \, a \in A_0 \text{ such that:}$$

$$\int_{\mathbb{B}_0} P_{ij}\frac{\partial v_i^h}{\partial X_j}dV = \int_{\partial_p\mathbb{B}_0} \bar{p}_i v_i^h dA + \int_{\mathbb{B}_0} \rho_0 b_i v_i^h dV \qquad (2.17)$$

$$\forall \, \mathbf{v}_b, \, b \in A_0$$

The Galerkin approximation for $\mathbf{v}^h$ is now substituted into the reformulated weak form, resulting in the residual, as exhibited in Eq. 2.18.

$$0 = \sum_b v_{ib}R_{ib} = \sum_{b\in A_0} v_{ib}R_{ib}$$

$$R_{ib} = 0 = \int_{\mathbb{B}_0} P_{ij}\frac{\partial \phi_b}{\partial X_j}dV - \int_{\partial_p\mathbb{B}_0} \bar{p}_i\phi_b dA - \int_{\mathbb{B}_0} \rho_0 b_i\phi_b dV \quad \forall \quad b \in A_0 \qquad (2.18)$$

The residual is "assembled" from element contributions, as follows. Consider a domain $\Omega_M \subset \mathbb{B}_0$, representing an element M within the body. This element definition contributes a corresponding element force vector and stiffness matrix. Assuming zero body force, the force vector, (f_{ia}), contibutes to the residual R_{ib}, and the stiffness matrix, $(k_{ia,rs})$, is the derivative of this residual with respect to u_{rs}^h. This relationship is presented, assuming index notation form, in Eq. 2.19. Here, N_a^M is the shape function on element M for node a. N_a^M is related to the basis functions by $\phi_a|_{elementM} = N_a^M$

$$f_{ia} = \int_{\Omega_M} P_{ij}\frac{\partial N_a^M}{\partial X_j}dV$$

$$k_{ia,rs} = \int_{\Omega_M} \frac{\partial P_{ij}}{\partial u_{rs}^h}\frac{\partial N_a^M}{\partial X_j}dV \qquad (2.19)$$

While obtaining f_{ia} is straightforward, obtaining $\frac{\partial P_{ij}}{\partial u_{rs}^h}$ in $k_{ia,rs}$ requires a chain rule of operations, as depicted in Eq. 2.20. Here, $\mathbf{B}$ is the left Cauchy-Green tensor as is defined in Section 2.1.

$$\frac{\partial P_{ij}}{\partial u_{rs}^h} = \frac{\partial P_{ij}}{\partial B_{mn}}\frac{\partial B_{mn}}{\partial u_{rs}^h} \qquad (2.20)$$

By invoking Eq. 2.10, the first term in Eq. 2.20 can be expanded, yielding the expression

defined in Eq. 2.21.

$$\frac{\partial P_{ij}}{\partial B_{mn}} = \overline{J}\frac{\partial \hat{J}}{\partial B_{mn}}T_{iq}F_{jq}^{-1} + J\frac{\partial T_{iq}}{\partial B_{mn}}F_{iq}^{-1} - JT_{iq}\hat{F}_{nq}^{-1}F_{jm}^{-1}$$
$$\frac{\partial \hat{J}}{\partial B_{mn}} = \frac{\partial \det(\hat{\mathbf{F}})}{\partial B_{mn}} = \frac{\partial \det(\mathbf{B+I})}{\partial B_{mn}} = \frac{\partial \det(\mathbf{B})}{\partial B_{mn}} \qquad (2.21)$$

Here, $\overline{J} = \det(\overline{\mathbf{F}})$ and $\hat{J} = \det(\hat{\mathbf{F}})$, wherein $\overline{\mathbf{F}}$ is the deformation gradient of the beginning-step configuration relative to the global reference configuration, while $\hat{\mathbf{F}}$ corresponds to the deformation gradient of the end-step configuration relative to the beginning-step configuration. It is noteworthy that the total deformation gradient $\mathbf{F}$ is defined as the product of $\hat{\mathbf{F}}$ and $\overline{\mathbf{F}}$. Conceptually, it can be defined as the composite outcome of the successive deformation gradients.

In the FEM, the Newton-Raphson iteration is a mathematical method used to find the approximate solutions through a step-by-step procedure. During each iteration, the residual is evaluated to determine if convergence is achieved. The residual is a nonlinear function of the nodal displacements, so the residual can be linearized by expanding it in a Taylor series, and truncating the series after the linear term. This linearized residual equation (see Eq. 2.22) is then solved for the corrections to the nodal degrees of freedom, $u_{jb} - u_{jb}^0$, where u_{jb}^0 is the current "best estimate" to the nodal displacements.

$$0 = R_{ib}(u_{jb}^0) + \frac{\partial R_{ia}}{\partial u_{jb}}|_{u_{jb}^0}\left[u_{jb} - u_{jb}^0\right] \qquad (2.22)$$

Let $\Delta u_{jb} = u_{jb} - u_{jb}^0$, $R_{ib}(u_{jb}^0) = R_{ib}^0$, and $\frac{\partial R_{ia}}{\partial u_{jb}}|_{u_{jb}^0} = K_{ia,jb}$ such that the linearized residual equation takes the form of Eq. 2.23.

$$\mathbf{K}\Delta\mathbf{u} = -\mathbf{R}^0 \qquad (2.23)$$

Expanding upon the given context, the derivative $\frac{\partial \hat{J}}{\partial B_{mn}}$ is defined in Eq. 2.21. Additionally, the expression $\frac{\partial T_{iq}}{\partial B_{mn}}$ can be established in terms of the incremental rotation tensor $\hat{\mathbf{R}}$

and stress values evaluated at a specific integration point. The second term of the right hand side of Eq. 2.20, with the use of Eq. 2.24, becomes as shown in Eq. 2.25 where δ_{mr} is the Kronecker delta.

$$B_{mn} = \sum_{s=1}^{N} u_{ms}^h \frac{\partial N_s^M}{\partial X_k} \overline{F}_{kn}^{-1} \tag{2.24}$$

$$\frac{\partial B_{mn}}{\partial u_{rs}^h} = \delta_{mr} \frac{\partial N_s^M}{\partial X_k} \overline{F}_{kn}^{-1} \tag{2.25}$$

2.4   Elements

Recall that elements in the CFEM are non-overlapping subregions that discretize the problem domain, and are regular in the sense that they all map from a single parent element. All elements have a small number of nodes associated with them. Nodes are geometric points, usually at the vertices of the elements, with which there are associated shape functions and problem unknowns. In Chapter 5, as part of the numerical exploration, two element types will be employed: the hexahedral element with eight nodes (hex-8) and the tetrahedral element with four nodes (tet-4). Within the framework of the CFEM, both of these elements exhibit quasi-linear shape functions denoted as N_α^m, wherein α is the local node number for element m. From this definition, the basis function for a global node a, on element m, is given by:

$$\phi_a|_{element\,m} = N_\alpha^m \tag{2.26}$$

To provide a visual representation of the parent element local node ids and integration points, see Figs. 2.4 and 2.5. For hex-8 elements, the parent element is a cube, while for tet-4 elements, it is a tetrahedron. The local node coordinates within the parent element are defined in the (ξ, η, ζ) coordinate system, ranging from -1 to 1. In the context of the CFEM, an isoparametric mapping utilizes the shape functions to transform coordinates from the parent element to the physical element. This transformation is depicted by Eq. 2.27, where $x_i(\xi, \eta, \zeta)$ is the ith coordinate in the physical element, corresponding to the point (ξ, η, ζ)
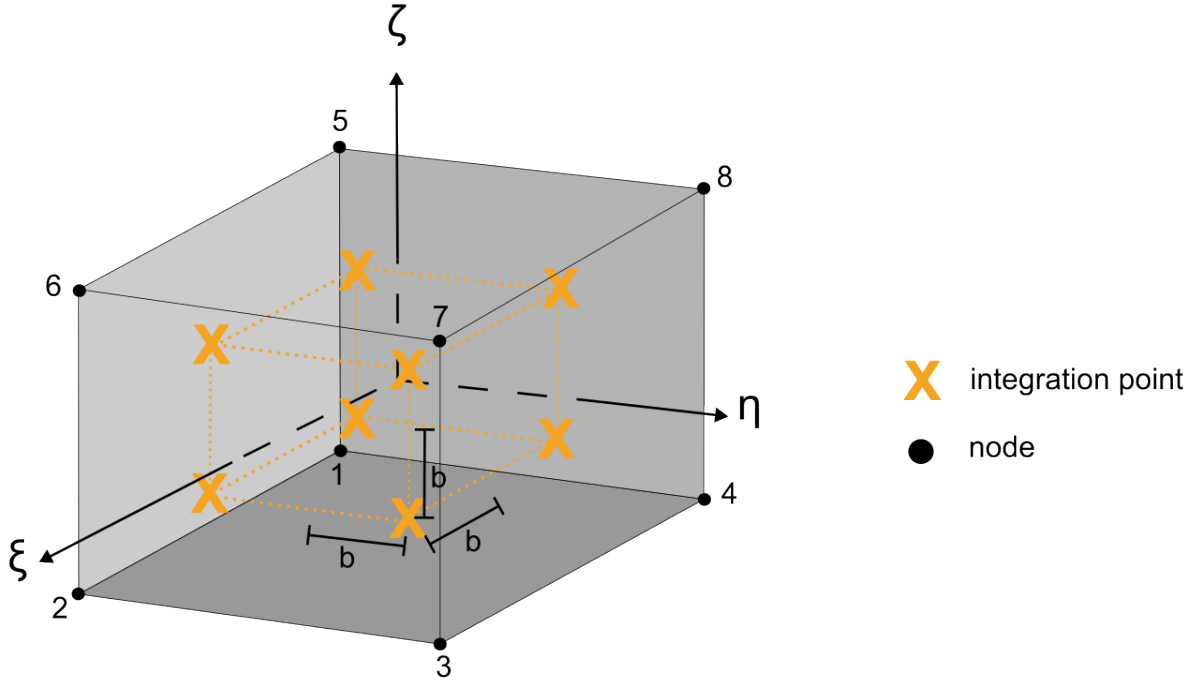
Figure 2.4: CFEM hex-8 parent element used in this dissertation for numerical exploration with nodes and integration points (IP)s labeled. Local node numbers α are labeled 1-8. b $= \frac{1}{\sqrt{3}}$

within the parent element. Additionally, $x_{i\alpha}^m$ is the ith coordinate of local node α on element m, with shape function N_α^m.

$$x_i\left(\xi, \eta, \zeta\right) = \sum_{\alpha=1}^{n} x_{i\alpha}^m N_\alpha^m\left(\xi, \eta, \zeta\right) \tag{2.27}$$

In developing an isoparametric map, the map becomes bijective provided that the element is not excessively distorted relative to the parent element, and exhibits the following set of requirements on the shape functions:

- Kronecker Delta Property: $N_\alpha^m\left(\mathbf{x}_\beta\right) = \delta_{\alpha\beta}$

- Interior Smoothness: $N_\alpha^m \in C^\infty\left(\Omega_m\right)$ where Ω_m is the domain containing physical element m

- Facet Autonomy:

23
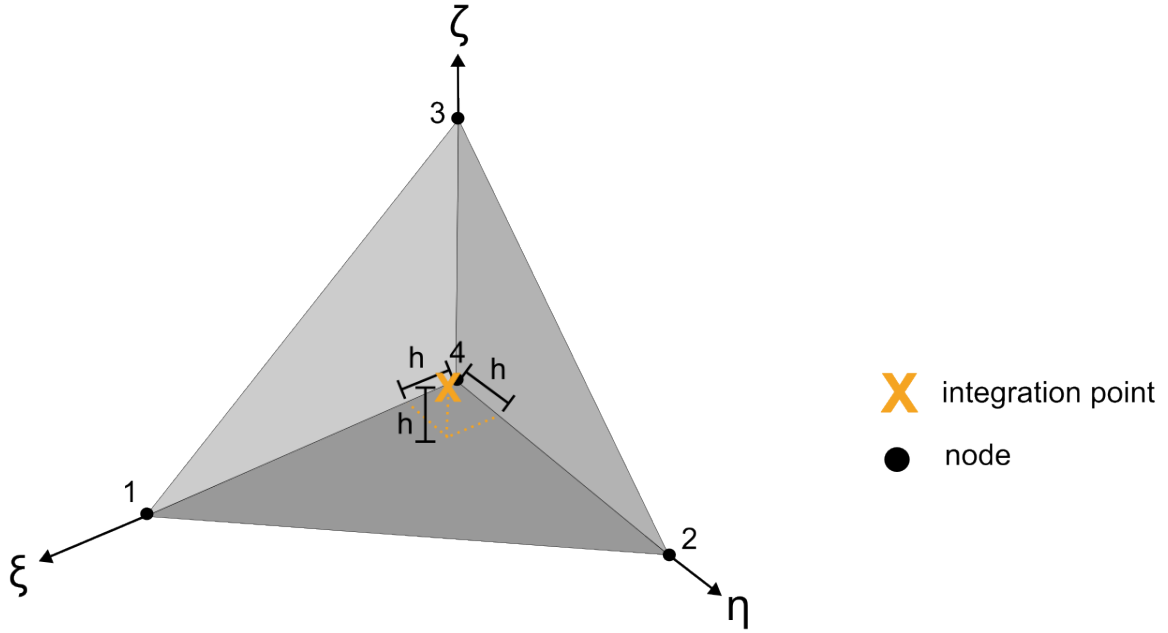
Figure 2.5: CFEM tet-4 parent element used in this dissertation for numerical exploration with nodes and integration points (IP)s labeled. Local node numbers α are labeled 1-4. h $= \frac{1}{4}$

1. $N_\alpha^m \neq 0$ only on edges that local node α belongs to

2. The variation of N_α^m on a facet depends only on the facet type. For example, if the facet is a triangle with 3 nodes, N_α^m must be linear.

- Partition of Unity:

$$\sum_{\alpha=1}^{N} N_\alpha^m (\xi, \eta, \zeta) = 1$$

- Linear Completeness: Nodal values that are consistent with a linear function produce exactly the same linear function.

The Kronecker Delta Property is important for two reasons: it facilitates convenient enforcement of displacement boundary conditions, and it also facilitates the isoparametric mapping. Interior smoothness is not strictly required, but continuity on the whole domain is, and interior smoothness certainly is consistent with that. Facet autonomy is the shape function

property that guarantees inter-element continuity. Partition of unity of the shape functions is required for the isoparametric mapping to map the parent to the physical element, and for the interpolant to be linearly complete in the physical coordinates. And finally, linear completeness is important to enable convergence as the mesh is refined.

The trial solution is written as a linear combination of the shape functions within each element, where the nodal values for the element are the coefficients. This construction shown in Eq. 2.28 leads to a globally continuous, linearly complete trial solution that supports convergent approximations under mesh refinement. It is also of importance to mention that shape functions for commonly used elements, such as the hex-8, are comprehensively documented in [17].

$$u_i\left(\xi, \eta, \zeta\right) = \sum_{\alpha=1}^{N} u_{i\alpha}^m N_\alpha^m\left(\xi, \eta, \zeta\right) \tag{2.28}$$

2.5 Gaussian Quadrature

Gaussian quadrature is a widely used numerical integration technique, evaluating the integrand at sample points, and accumulating the weighted results to approximate the integral value. The approximation is given by Eq. 2.29, where $\left(\tilde{\xi}_\ell, \tilde{\eta}_\ell, \tilde{\zeta}_\ell\right)$ is the integration point (IP) location of IP ℓ, n_{int} is the number of IPs, $g\left(\tilde{\xi}_\ell, \tilde{\eta}_\ell, \tilde{\zeta}_\ell\right)$ is a arbitrary integrand, and W_ℓ is the weight of IP ℓ. IP locations and weights of common elements can be found in [17].

$$\int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} g\left(\xi, \eta, \zeta\right) \, d\xi \, d\eta \, d\zeta = \sum_{\ell=1}^{n_{int}} g\left(\tilde{\xi}_\ell, \tilde{\eta}_\ell, \tilde{\zeta}_\ell\right) W_\ell \tag{2.29}$$

Full integration is a quadrature rule of sufficiently high order that the stiffness matrix is exactly integrated on undistorted elements. Full integration always leads to a full-rank stiffness matrix. And in 3D, a stiffness matrix of full-rank has a rank 6 less than the dimension due to the 6 rigid body modes. For full integration, hex-8 elements require 2x2x2-point Gauss quadrature while tet-4 elements require only 1-point. Full integration prevents the occurence of zero-energy modes and the need for hourglass stiffness. Zero-

energy modes, also referred to as spurious modes, represent non-physical solutions that have unresisted deformation. They can have a significant impact on the accuracy and reliability of the results. It is commonly accepted to mend this problem with the use of an artificial stiffness, termed hourglass stiffness, that stiffens the element's response in its zero-energy modes. An example of this phenomenon is a hex-8 element with one integration point in the middle, subject to a bending moment. The resulting deformation in the case of a hex-8 is a trapezoidal prism. However, when the strain is evaluated at the single integration point, there is zero strain, which results in zero nodal forces.

2.6   Limitations of the CFEM

Standard hex-8 elements are widely used and are often effective in practice, but not for all problems. Specifically, they are known to exhibit severe locking when the material is nearly-incompressible, and/or deformation is bending dominated, and the aspect ratio of the element is large. In addition, locking is not always obvious since a numerical solution is produced, with little indication that something is wrong.

Tet-4 elements are often used in combination with hex-8 elements in order to facilitate automatic meshing of complex domains. However, tet-4 elements are more prone to locking than hex-8 elements are as seen in . And, while the solutions produced by tet-4 elements provably converge under mesh refinement, the error is often unacceptably large for practical-size meshes.

In Chapter 3, the focus shifts towards exploring various strategies designed to mitigate locking phenomena. By addressing the issue of locking, these methods aim to expand the utility of the finite element method to a broader range of problems involving complex geometries, thin structures subjected to bending, and nearly-incompressible materials.

Chapter 3

Locking Mitigation

The primary focus of this dissertation is an investigation of two distinct forms of locking: shear and volumetric. Shear locking is a common occurrence in the analysis of thin structures, where the aspect ratio of an element surpasses 3, mainly affecting problems dominated by bending. Volumetric locking, on the other hand, affects scenarios featuring nearly-incompressible materials.

This chapter surveys techniques devised to alleviate shear and volumetric locking within the conventional finite element method. By illuminating the factors contributing to the emergence of locking phenomena in simulations, the aim of this chapter is to provide mitigation insights. The surveyed methods have been categorized into five main groups: assumed strain, F-bar, smoothed FEM, multifield mixed variational formulations, and "other" approaches that do not conform to any of the other groups. Throughout these diverse strategies, researchers have cultivated a theoretical understanding of locking and its potential origins.

As earlier mentioned in Chapter 1, the mathematical comprehension of volumetric locking is more highly-developed than that of shear locking, thus resulting in more substantial advancements in devising mitigation strategies. However, recent progress has been made in attempts to mathematically define shear locking. For example, subspace analyses have been employed to identify prohibited deformation modes, leading to solutions that increase the

nullity of matrices associated with incompressibility and the Kirchhoff hypothesis [6, 8, 4]. This extension of the polynomial order of the strain field serves to mitigate locking. Nonetheless, addressing shear locking involves more than merely increasing the polynomial order of the solution, a point discussed in greater detail within Section 3.6.

3.1 Assumed Strain Methods

The classification termed "assumed strain" encompasses a range of methodologies wherein the conventional strain-displacement matrix, $\mathbf{B}$ (see Eq. 3.5), is replaced with a modified version based on a specific set of assumptions. In the CFEM, the strain-displacement matrix relates the displacements at the nodes to the resultant strain at the element's integration points. Within this grouping, three distinctive subgroups emerge: enhanced assumed strain (EAS), reduced integration (RI), and B-bar. Note that the replacement of $\mathbf{B}$ directly applies only to small-deformation problems, in which the IP strains are linearly related to the nodal displacements. The more generic term "enhanced assumed strain" encompasses general finite-deformation kinematics.

Finite-deformation kinematics have been discussed in Chapter 2, but small-deformation allows one to define $\mathbf{B}$ in terms of the element basis functions. Consider first the weak form in the Eulerian reference frame shown in Eq. 3.1.

$$\text{Find } \mathbf{u}_a, a \in A_0 \text{ such that:}$$

$$\int_{\mathbb{B}} T_{ij} v_{i,j} \, dv = \int_{\mathbb{B}} \rho b_i v_i dv + \int_{\partial_t \mathbb{B}} v_i \bar{t}_i da \qquad (3.1)$$

$$\forall \, \mathbf{v}_a, a \in A_0$$

In this equation, $\mathbf{u}_a$ is the displacement at node a, A_0 is the space of nodal values without displacement boundary conditions, $\mathbf{T}$ is the cauchy stress tensor, ρ is the density, $\mathbf{b}$ is the body force, $\bar{t}_i$ is the applied traction force, and $\mathbf{v}_a$ is the test function. The strain, $\mathbf{e}$, and

the Cauchy stress, $\mathbf{T}$, can be defined as column vectors, as shown in Eq. 3.2.

$$
\mathbf{e} = \begin{bmatrix} e_{11} \\ e_{22} \\ e_{33} \\ 2e_{23} \\ 2e_{13} \\ 2e_{12} \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} T_{11} \\ T_{22} \\ T_{33} \\ T_{23} \\ T_{13} \\ T_{12} \end{bmatrix}
\tag{3.2}
$$

Now, for small deformations, the strain can be approximated to first order as

$$
e_{ij} = \frac{1}{2} \left(u_{i,j} + u_{j,i} \right).
\tag{3.3}
$$

Thus, using the Galerkin approximation for the displacement in Eq. 2.28, the strain becomes

$$
e_{ij} = \frac{1}{2} \sum_{a=1}^{N} \left(u_{ia}\phi_{a,j} + u_{ja}\phi_{a,i} \right),
\tag{3.4}
$$

where $()_{,j} = \frac{\partial ()}{\partial x_j}$. Define then the formulation of the strain-displacement matrix at node a, $\mathbf{B}_a$, for small deformations in Eq. 3.5, in terms of the element basis functions at each node a, $\phi_a(\mathbf{x})$.

$$
\mathbf{B}_a = \begin{bmatrix} \phi_{a,1} & 0 & 0 \\ 0 & \phi_{a,2} & 0 \\ 0 & 0 & \phi_{a,3} \\ 0 & \phi_{a,3} & \phi_{a,2} \\ \phi_{a,3} & 0 & \phi_{a,1} \\ \phi_{a,2} & \phi_{a,1} & 0 \end{bmatrix}
\tag{3.5}
$$

Therefore, the strain can be defined in terms of $\mathbf{B}_a$:

$$\mathbf{e} = \sum_{a=1}^{N} \mathbf{B}_a \mathbf{u}_a, \quad \mathbf{u}_a = \begin{bmatrix} u_{1a} \\ u_{2a} \\ u_{3a} \end{bmatrix}. \tag{3.6}$$

In addition, linear elasticity allows the Cauchy stress to relate to the strain using the relation

$$\mathbf{T} = \mathbf{Ce} = \sum_{a=1}^{N} \mathbf{CB}_a \mathbf{u}_a, \tag{3.7}$$

where $\mathbf{C}$ is a rank-4 modulus tensor that holds material relations. For an isotropic material,

$$\mathbf{C} = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & & & \\ \lambda & \lambda + 2\mu & \lambda & & \mathbf{0} & \\ \lambda & \lambda & \lambda + 2\mu & & & \\ & & & \mu & 0 & 0 \\ & \mathbf{0} & & 0 & \mu & 0 \\ & & & 0 & 0 & \mu \end{bmatrix}. \tag{3.8}$$

Consider now the integrand on the left-hand side of Eq. 3.1. Recalling the Galerkin approximation for the test function in Eq. 2.28,

$$T_{ij} v_{i,j} \Rightarrow \sum_{b \in A_0} \mathbf{v}_b^T \mathbf{B}_b^T \mathbf{T} \tag{3.9}$$

$$\Rightarrow \sum_{b \in A_0} \mathbf{v}_b^T \mathbf{B}_b^T \left[\sum_{a=1}^{N} \mathbf{CB}_a \mathbf{u}_a \right] \tag{3.10}$$

$$\Rightarrow \sum_{b \in A_0} \sum_{a=1}^{N} \mathbf{v}_b^T \mathbf{B}_b^T \mathbf{CB}_a \mathbf{u}_a. \tag{3.11}$$

Plugging in the results for $T_{ij} v_{i,j}$, along with the Galerkin approximation to the displacement and the test function to the weak form allows the formulation of the residual. From

30

the residual, the stiffness matrix becomes

$$\mathbf{K}_{ba} = \int_{\mathbb{B}} \mathbf{B}_b^T \mathbf{C} \mathbf{B}_a dv \quad \in \mathbb{R}^{3 \times 3} \tag{3.12}$$

This stiffness matrix has been outlined to assist in the explanation of reduced integration and B-bar methods.

3.1.1   Enhanced Assumed Strain

EAS methods are grounded in the idea that the strain field can be expressed as the sum of two distinct components. These components encompass the conventional strain field derived from the local strain-displacement relations at the integration points, alongside an augmentation that remains independent of the displacement field. The augmented strain field arises from the introduction of a displacement-like field into the problem, which scales a modified version of the strain-displacement matrix $\mathbf{B}$. This displacement-field augmentation is not continuous at the element boundary, and therefore is subject to restrictions. This is since discontinuous displacement fields do not arbitrarily lead to convergent FE solutions.

The precise count of supplementary degrees of freedom introduced to develop this enhancement may vary, demanding thoughtful consideration to attain the desired effect. Consequently, methods may intentionally introduce rank deficiency to mitigate locking issues. It is seen here that balancing rank deficiency and locking reoccurs regularly in locking mitigation methods, as discussed in Section 1.2. Thus, many of these approaches acknowledge the need to counteract rank deficiency through stabilization strategies [32, 19, 7, 6, 8].

The rank of an element stiffness matrix in three dimensions should be 6 less than the dimension of the matrix to be considered "full rank" for the CFEM. The 6-dimensional nullspace contains the rigid body modes, and any higher dimension can be influenced by the number of integration points (IPs). Full rank can always be achieved by increasing the number of IPs, but doing so can re-introduce locking. Returning to the discussion of

31

over-sampling and under-sampling in Section 1.2, under-sampling can lead to rank deficiency and over-sampling can lead to locking. The methods increasing the number of degrees of freedom with EAS often results in element stiffness matrices that are under-sampled to avoid locking. Thus, rank deficiency is introduced, and generally this means zero-energy modes will be present in the solution. To counteract these zero-energy modes, artificial stabilization, such as hourglass stiffness, must be added to the method.

The methodology described in [32] introduces hourglass force components to the internal forces of the element, effectively employing stabilization techniques. Additionally, [19] delves into the establishment of sets of orthogonality conditions that must be satisfied by the displacement gradient. The fulfillment of these conditions mitigates the necessity for external stabilization, such as hourglass stiffness. Moreover, [7] combines multiple techniques to formulate an axisymmetric solid-shell element. In this work, EAS is harnessed to counteract volumetric locking, Selective Reduced Integration (SRI) is used to alleviate shear locking, and B-bar methodology is deployed to mitigate volumetric locking within the hourglass field, thus providing stabilization.

A recent advancement in the domain of EAS methods, introduced in [31], eliminates the demand for stabilization by introducing an additional field termed the spatial-displacement gradient (h-enhancement). The integration of this supplementary field is designed to avoid zero-energy modes, achieved through coupling it with a Petrov-Galerkin approximation to produce robust handling of mesh distortion. A notable result of this approach is the inevitable unsymmetric element stiffness matrices, leading to increased computational cost as compared to symmetric element stiffness matrices.

3.1.2 Reduced Integration

RI techniques utilized for the mitigation of locking aim to eliminate from the integration points what is colloquially referred to as "parasitic shear" [9]. To achieve this elimination, RI uses a lesser number of Gaussian IPs, resulting in an under-sampled stiffness matrix with

the intention to mitigate locking. In small-deformations, this involves under-sampling the stiffness matrix defined in Eq. 3.12. The application of RI for locking mitigation often draws from earlier attempts for volumetric and shear locking mitigation [25, 16, 43, 21].

However, in contemporary contexts, RI is frequently used in the form of SRI. In SRI, the strain-displacement matrix of Eq. 3.12 is split into its devatoric and dilatational parts by first defining the dilatational strain:

$$\mathbf{e}^{dil} = \frac{1}{3} \begin{bmatrix} e_{11} + e_{22} + e33 \\ e_{11} + e_{22} + e33 \\ e_{11} + e_{22} + e33 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{3.13}$$

The dilatational part of $\mathbf{B}$ is defined through $\mathbf{e}^{dil} = \mathbf{B}^{dil}\mathbf{u}$, and the deviatoric part is then

$$\mathbf{B}^{dev} = \mathbf{B} - \mathbf{B}^{dil}. \tag{3.14}$$

With the split of the two parts, the stiffness matrix can be redefined to have two terms:

$$\mathbf{K} = \int_{\mathbb{B}} \left[\mathbf{B}^{dev}\right]^T \mathbf{C}\mathbf{B}^{dev} dv + \int_{\mathbb{B}} \left[\mathbf{B}^{dil}\right]^T \mathbf{C}\mathbf{B}^{dil} dv \tag{3.15}$$

Now, the second term on the right hand side is integrated using one integration point to selectively reduce the integration of the stiffness matrix. In the contemporary texts, SRI is utilized exclusively to counteract shear locking. In such cases, it is frequently combined with other methodologies to address volumetric locking [7, 42].

3.1.3   B-Bar

The B-bar subgroup encompasses methods that revolve around the replacement of the strain-displacement matrix, $\mathbf{B}$, with an altered matrix, $\overline{\mathbf{B}}$, featuring an "improved" dilatational component. See Eq. 3.5 for the definition of $\mathbf{B}$. This enhancement involves decomposing $\mathbf{B}$ into its deviatoric and dilatational components, subsequently selecting a dilatational part that aligns with specific requirements. Notably, a generalized form of SRI can be derived from this framework by opting to utilize a solitary integration point at the center to compute the dilatational contribution [17]. There is, therefore, a duality between selective reduced integration and the B-bar approach, at least for small deformations and elastic material behavior.

B-bar methodologies are frequently combined with other approaches to improve their capabilities in mitigating locking issues [7, 33, 34]. For instance, in [33], the Element-Free Galerkin Method (EFGM) [2] is coupled with a B-bar technique to mitigate volumetric locking. Furthermore, as alluded to in Section 3.1.1, [7] incorporates a B-bar approach alongside EAS and SRI to alleviate volumetric locking within the hourglass field. Given the potential vulnerability of B-Bar methods to rank deficiency, stabilization measures are often warranted. Moreover, a recent study by [34] adopts a B-bar method in conjunction with a mixed formulation, thereby allowing independent strain and displacement interpolations. Here, B-bar is essential in addressing volumetric locking.

Lastly, the work of [20] draws inspiration from [15], wherein $\overline{\mathbf{B}}$ is computed as an average value of multiple $\mathbf{B}$ matrices across a specified region. While [15] performed this averaging over a single element, [20] conducts a weighted average of strain-displacement matrices from multiple elements, as established within the CFEM. These elements used in the average occur in a small patch such that they are adjacent to one another. Furthermore, this approach incorporates nodal integration instead of Gaussian quadrature.

3.2 F-bar Methods

In the original F-bar method, proposed in [11], the computation of the Cauchy stress tensor, $\mathbf{T}$, involves substituting the deformation gradient, $\mathbf{F}$, with an assumed and adjusted gradient, denoted as $\overline{\mathbf{F}}$ [27]. However, within the domain of linear elements, where $\mathbf{F}$ remains constant, the original F-bar approach alone fails to effectively address volumetric locking with nearly-incompressible materials. To overcome this limitation, the F-bar-patch method was introduced in [27]. This variant modifies $\mathbf{F}$ across a group of adjacent elements, thereby extending the condition of isochoric deformation to encompass a group of elements instead of being applied to each individual element. The choice of the count of elements forming a patch introduces difficulties into the implementation of this technique.

Observations indicate that increasing the number of elements constituting the patch increases the flexibility of the mesh. This increased flexibility may potentially lead to deformations exceeding theoretical predictions. Conversely, a reduction in the element count within the patch results in diminished flexibility, thereby impeding the effective alleviation of locking. Thus, achieving a balance in determining the optimal count of elements within a patch becomes paramount. Striking this balance is necessary to sufficiently increase mesh flexibility for the purposes of locking mitigation, while simultaneously avoiding rank deficiency.

3.3 Smoothed FEMs

Smoothed FEM (S-FEM) is an approach that combines the principles of the mesh-free methods' weakened weak (W2) formulation with the traditional FEM [40]. This amalgamation proves particularly effective when applied to triangular elements, benefiting from the ease of triangular mesh generation (i.e. tet-4 elements). In S-FEM, strains are computed through integration utilizing a smoothing function that operates over a designated smoothing domain. The process involves several key steps, as outlined by [40]:

1. Generation of a Smoothing Domain: A smoothing domain is constructed based on adjacent elements sharing edges, faces, and/or nodes within the CFEM mesh.

2. Development of Shape Functions: Shape function values are determined at a point using point interpolation. This formulation avoids the need to explicitly generate shape functions.

3. Calculation of Smoothed Strain: The smoothed strain is then evaluated in terms of the shape function values at specific points and the nodal displacements for the elements within the smoothing domain.

The S-FEM methodology exists in various forms, with the primary distinctions arising from the nature of the smoothing domains. A summary of several S-FEM variations, along with their characteristics, is provided below based on [40]:

- Cell-Based S-FEM (CS-FEM)

    - smoothing done inside elements

    - good for element distortion

    - for each element type, there is a different optimal number of smoothing domains to discretize the element with

- Edge-Based S-FEM (ES-FEM)

    - smoothing domains based on edges of tetrahedral and polyhedral elements

    - to mitigate locking, it must be paired with other methods such as F-bar [30]

- Node-Based S-FEM (NS-FEM)

    - smoothing domains occur over the nodes

    - not fully locking-free for incompressible materials

    - paired with selective FS-FEM to mitigate shear and volumetric locking [29]

- Face-Based S-FEM (FS-FEM)

    - smoothing domains are associated with the faces of tetrahedral elements

    - paired with NS-FEM to mitigate shear and volumetric locking [29]

- Selective S-FEM

    - NS-FEM for volumetric part to avoid volumetric locking

    - ES-FEM/FS-FEM for deviatoric part for good accuracy

In essence, S-FEM introduces a novel perspective by integrating mesh-free concepts into the traditional FEM framework, providing a range of variations tailored to different scenarios and locking concerns.

3.4   Multifield Mixed Variational Formulations

Methods falling under the category of multifield mixed variational formulations are rooted in either the Hellinger-Reissner (HR) or Hu-Washizu (HW) variational principles. Unlike the CFEM, which is based on a single unknown displacement field, these formulations incorporate multiple unknown fields, such as displacement and stress. The choice between HR and HW approaches depends on their underlying principles, as described in [26]. The key distinction lies in the complementary energy function for HR and the strain energy for HW. HR involves both displacement and stress fields, rendering it suitable for addressing shear and volumetric locking [3, 4, 26, 41]. HW encompasses displacement, strains, and stress fields, making it better suited for nonlinear materials [5, 43].

While HR methods, as in [26], exhibit favorable performance on distorted meshes, not all HR-based approaches explicitly address the effects of mesh distortion [3, 4]. The consideration of mesh distortion is paramount for thin, bending-dominated problems, where such distortion is often prevalent. Therefore, understanding and mitigating the impact of mesh

distortion can greatly enhance the robustness and accuracy of multifield mixed variational formulations, especially in scenarios where thin structures are involved.

3.5   Other Methods

3.5.1   Generalized Modal Element Method

The Generalized Modal Element Method (GMEM) is an approach that constructs the element stiffness matrix by decomposing element deformation patterns into generalized modes, as detailed in [13]. In this method, each generalized mode is defined by a nodal displacement vector $\mathbf{U}$, and its corresponding nodal force vector $\mathbf{F}$, which are referred to as modal displacement and force vectors, respectively. Using these generalized modes, the element stiffness matrix is formulated as $\mathbf{K} = \mathbf{F}\mathbf{U}^{-1}$, where $\mathbf{U}$ is a matrix containing modal displacement vectors $\mathbf{U} = \left[\mathbf{U}^1, \mathbf{U}^2, \ldots, \mathbf{U}^N\right]$, and $\mathbf{F}$ contains modal force vectors $\mathbf{F} = \left[\mathbf{F}^1, \mathbf{F}^2, \ldots, \mathbf{F}^N\right]$ for N generalized modes.

For an 8-node solid element, GMEM identifies a total of 24 modes, classified into various categories: 6 rigid body motions, 3 tensile modes, 3 shear modes, 3 torsional modes, 6 bending modes, and 3 "unphysical" modes. While the first 21 modes are determined using solid mechanics equations, the unphysical modes are computed using the assumed displacement method to prevent volumetric locking. The assumed displacement method in [13] redefines the internal displacement functions to be expressed in terms of the Cartesian coordinate system, and this enables a more robust handling of mesh distortion. Moreover, the assumed displacement method ensures that the matrix $\mathbf{U}$ remains invertible. In more complex modes, as the "unphysical" modes commonly are, it is difficult to find displacement functions that satisfy the governing equations by solving the problem analytically.

The element stiffness matrix produced by GMEM is unsymmetric for any mesh that is not strictly cuboidal. However, it can be transformed into a symmetric form by using different

modal force vectors. The potentially unsymmetric stiffness matrix is illustrated below:

$$\mathbf{K}_{unsymmetric} = \int_{\mathbb{B}} \mathbf{B}^T \mathbf{C} \mathbf{E} \, dv \, \mathbf{U}^{-1} \tag{3.16}$$

where $\mathbf{B}$ is the strain-displacement matrix, $\mathbf{C}$ is the elasticity tensor, and $\mathbf{E}$ is termed the modal strain matrix. The modal strain matrix contains column vectors of the modal strain fields of each of the 24 modes. Now, to produce a symmetric stiffness matrix, $\mathbf{E}\mathbf{U}^{-1}$ serves the role of the strain-displacement matrix such that

$$\mathbf{K}_{symmetric} = \left[\mathbf{U}^T\right]^{-1} \int_{\mathbb{B}} \mathbf{E}^T \mathbf{C} \mathbf{E} \, dv \, \mathbf{U}^{-1}. \tag{3.17}$$

One limitation of an unsymmetric element stiffness matrix is its inability to extend to frequency analysis. Therefore, the numerical investigation in [13] examines both the unsymmetric and symmetric versions of the GMEM element stiffness matrix formulation. The results of the study demonstrate effective mitigation of shear and volumetric locking in bending problems, as well as robustness in the presence of mesh distortion for both cases.

While the unsymmetric case passes the patch test for solids proposed by MacNeal and Harder [24], the symmetric case does not. Additionally, there is no theoretical analysis of convergence, such as the F-E-M-Test [35]. The absence of a theoretical consideration of convergence raises questions about the overall convergence behavior of the symmetric GMEM element stiffness matrix formulation.

3.5.2 Generalized Finite Element Method

The Generalized Finite Element Method (GFEM) builds upon the Partition of Unity Method (PUM) [1] and extends its capabilities to incorporate more generalized types of functions. In PUM, overlapping subdomains or "patches" are combined with desirable approximation functions to approximate solutions across the entire domain. A desirable approximation function is smooth, continuous, and differentiable on each subdomain. These approximation

functions are weighted by functions that form a C^0 partition of unity on the domain Ω, ensuring continuity across the subdomains. The PUM approximation of the solution is given by Eq. 3.18, where $\psi_j^{(i)}$ represents the jth desirable approximation function of patch i, $a_j^{(i)}$ are unknown coefficients, and the ϕ_i are a collection of continuous basis functions. In fact, the CFEM basis functions form good candidates. A patch $\Omega_i \subset \Omega$ contains elements with vertex V_i such that $\phi_i = 1$ at V_i and zero on the boundary of the patch. This set of conditions for ϕ_i is what produces a C^0 partition of unity that enforces inter-element continuity.

$$u_{PUM} = \sum_i \phi_i \left(\sum_j a_j^{(i)} \psi_j^{(i)} \right) \tag{3.18}$$

GFEM further extends PUM by adding another set of functions that can include higher order shape functions, as shown in Eq. 3.19. Here, n_{vert} represents the number of basis functions at element vertices, n_i is the number of desirable functions on patch i with vertex V_i, n_{FEM} is the number of additional side and interior nodes used to increase the polynomial order of the solution, b_k are unknown coefficients, and $\tilde{\phi}_k$ are special functions. These special functions can be higher order polynomials or can even be discontinuous. This formulation generalizes the FEM by allowing any family of special functions to be included in the solution.

$$u_{GFEM} = \sum_i^{n_{vert}} \phi_i \left(\sum_j^{n_i} a_j^{(i)} \psi_j^{(i)} \right) + \sum_k^{n_{FEM}} b_k \tilde{\phi}_k \tag{3.19}$$

It is worth noting that the addition of n_{FEM} introduces extra degrees of freedom (DOFs), which can potentially lead to zero-energy modes, as discussed in Chapter 1. This issue becomes pronounced in GFEM, especially for large deformations of elastoplastic solids. In an effort to mitigate volumetric locking, [23] utilized GFEM without these extra DOFs. This approach, proposed in [37], is demonstrated in Eq. 3.20, where $L_j^{\Omega_i}$ represents the enrichment

function on patch Ω_i, and u_j is the conventional nodal unknown of the CFEM.

$$u_{GFEM} = \sum_i^{n_{vert}} \phi_i \left(\sum_j^{n_i} L_j^{\Omega_i} u_j \right) \tag{3.20}$$

Results in [23] show this method's potential for robustness against mesh distortion, stable convergence, and less-severe volumetric locking than the CFEM when handling nearly-incompressible materials.

3.5.3   Field-Consistent Formulation

The field-consistent formulation introduced in [9] represents an early effort to address locking in finite element analysis. While it may be superseded by more recent approaches, it provides valuable insights into the locking phenomenon and its mitigation. This approach focuses on tackling shear locking by removing the "parasitic" shear terms through the assumption of plane stress conditions. This adjustment helps alleviate shear locking and allows for more accurate results with one element through the thickness of thin plate, beam, and shell structures.

However, this solution introduces a new challenge: volumetric locking. With only one hexahedral element through the thickness, and the use of quasi-linear shape functions, the method encounters an issue known as "Poisson's ratio stiffening," which is referred to as volumetric locking in this work. This phenomenon is particularly pronounced in problems involving nearly-incompressible materials. In these cases, linear elements with the assumption of plane stress exhibit constant strain throughout the thickness, leading to volumetric locking.

To mitigate volumetric locking, the approach in [9] incorporates parabolic bubble functions into the displacement fields. By introducing these additional functions, the strain distribution can be adjusted to become linear through the element's thickness, helping to alleviate volumetric locking. This emphasizes the significance of increasing the polynomial

41

order of the displacement field, especially in bending-dominated problems.

Although, it is important to note that increasing the polynomial order of the displacement is not always an effective solution to mitigate locking alone. As discussed in Section 3.6, other factors such as the element formulation, choice of integration schemes, and material behavior also play crucial roles in determining whether locking can be effectively mitigated. Therefore, while the approach in [9] sheds light on the importance of higher polynomial order interpolation of the displacement solutions, it also highlights the complex interaction of various factors that need to be considered when addressing locking in finite element analysis.

3.6   Locking Source Contemplation

As discussed in sections 1.2 and 1.3, expanding the solution subspace represents a sound strategy for mitigating shear and volumetric locking. The underlying principle is to introduce greater flexibility within the displacement field associated with a given set of nodal displacements. This flexibility enables the nodal displacements to conform to the constraint associated with the locking phenomenon. In the case of volumetric locking, this constraint is the isochoricity in the case of near-incompressibility. This greater flexibility was achieved in many of the above methods by augmenting the strain field within each element to obtain additional terms controlled by independent unknowns.

Although, navigating locking mitigation involves a delicate balance between the strategic increase of the polynomial order of the solution, and the avoidance of rank deficiency. The excessive use of supplementary degrees of freedom can give way to under-sampling the stiffness matrix, which in turn may introduce rank deficiency. This domain of balance is where ESFEM is established, by increasing the polynomial order of the solution primarily through thoughtful utilization of existing degrees of freedom. Thus, ESFEM may keep the same Gauss quadrature rules for the CFEM quasi-linear hex-8 element without under-sampling. However, for the ESFEM tet-4 element, one integration point is so little that the quadrature

42

rule may need to change for improved results. ESFEM provides new strategies to address these complexities by balancing precision and stability in the aim to overcome challenges provided in locking mitigation without inducing rank deficiency.

Chapter 4

The Extended-Stencil Finite Element Method

As discussed in Chapter 3, mitigating locking and at the same time avoiding rank deficiency requires a novel approach. The ESFEM makes use of nodal data available in a conventional finite element mesh. This is done by extending the concept of an element to reference nodes of other elements in the mesh. A more comprehensive explanation of this extension is given in later sections. By simply modifying an element's nodal dependencies, the formulation of the weak form and the Galerkin approximation to the Boundary Value Problem (BVP) remains the same. However, the basis functions are now moderately high-order polynomials, and formulated with a polynomial fitting algorithm.

The central aim of this methodology is to mitigate locking by way of these higher-order polynomial basis functions. A resulting benefit discussed in Chapter 5 is improved accuracy for the same conventional finite element mesh. Another desired aim is to increase the utility of the tet-4 element. Currently, tet-4 elements are easily auto-generated; however yield inaccurate results. The prospect of dependable outcomes from an auto-generated mesh may decrease time spent developing a "high quality" mesh. Such meshes for thin plates and shells are conventionally known to be hexahedral elements with 3-5 layers through the thickness

and aspect ratios less than 3.

4.1   Increasing an Element's Nodal Dependencies

To modify an element's nodal dependencies to include those of other elements, two distinct nodal categories are defined. Within this framework, nodes belonging to an element are designated as its "primary nodes", while those of other elements become "secondary nodes." There exist circumstances where the demand for more comprehensive data surpasses that which the mesh provides. A notable example of this scenario is in the context of the example in Section 1.1. Here, the mesh has only a single element through its thickness and width. In such scenarios, the lack of nodes distributed across the width inhibits an accurate representation of anticlastic curvature. The exact solution to this curvature is a quadratic function, so at least three nodes through the width is needed to represent the exact solution.

To address this limitation, "ghost" secondary nodes are added into the mesh to supplement the nodal data. Ghost nodes are exclusively incorporated as secondary nodes, linked to elements, and do not have any vertex association with elements in the mesh. This addition increases the available nodal data, developing increased accuracy in cases where the conventional mesh representation is insufficient alone.

4.1.1   Secondary Nodes

In the context of the ESFEM, secondary nodes are explicitly defined as nodes of other elements. The input given to the ESFEM program is equivalent to the CFEM-compatible mesh framework. An element's own nodes assume the role of primary nodes. After the program obtains the nodal locations and element connectivity, the identification of secondary nodes may begin. This is achieved through the identification of nodes positioned precisely one "element edge" away from the current element of interest. In this context, an "element edge" defines the connection between an element's vertices. In the context of hex-8 and tet-

4 elements, this then consists of 2 nodes. To append secondary nodes to an element, each primary node is considered individually. All edges of the current primary node are considered. If the other node of the edge belongs to another element, then this node is appended as a secondary node. Upon completion of this procedure, the resulting configuration provides all elements with both primary and secondary nodes. This augmented nodal dataset serves as a foundation to obtain the polynomial basis functions, as explained in Section 4.3.

Pictorially, in Fig. 4.1, the element highlighted is shown with its nodes conforming to the CFEM arrangement as compared against the incorporation of secondary nodes in the ESFEM. The ESFEM framework extends the stencil of the element beyond the the highlighted element to encompass the vertices of adjacent elements as secondary nodes. Thus, the highlighted element references 24 nodes instead of 8. In instances where adjacent elements exist on all faces of the highlighted element, the count increases to 32 nodes (comprising 8 primary nodes and 24 secondary nodes).

4.1.2   Ghost Secondary Nodes

Incorporating secondary nodes using solely the vertices of adjacent elements has limitations dictated by the mesh. This constraint becomes apparent when considering the mesh of the example in Section 1.1, which features a single element in two directions and a longitudinal extension of 10 elements. In this mesh, two nodes are along the width of the beam. Given that the analytical solution for anticlastic curvature is a quadratic polynomial, a minimum of three points along an edge is necessary for an accurate representation of this curvature in the ESFEM. To address this discrepancy, the concept of "ghost" secondary nodes is introduced.

Ghost secondary nodes are strategically introduced into the mesh to augment the available nodal data. The goal is for these additional nodes to facilitate a more precise representation of deformation without the need for additional elements. However, the inclusion of ghost nodes must be executed thoughtfully to avoid rank deficiency. By limiting ghost node additions to one per node on the external boundary of the mesh, the hex-8 stiffness matrix
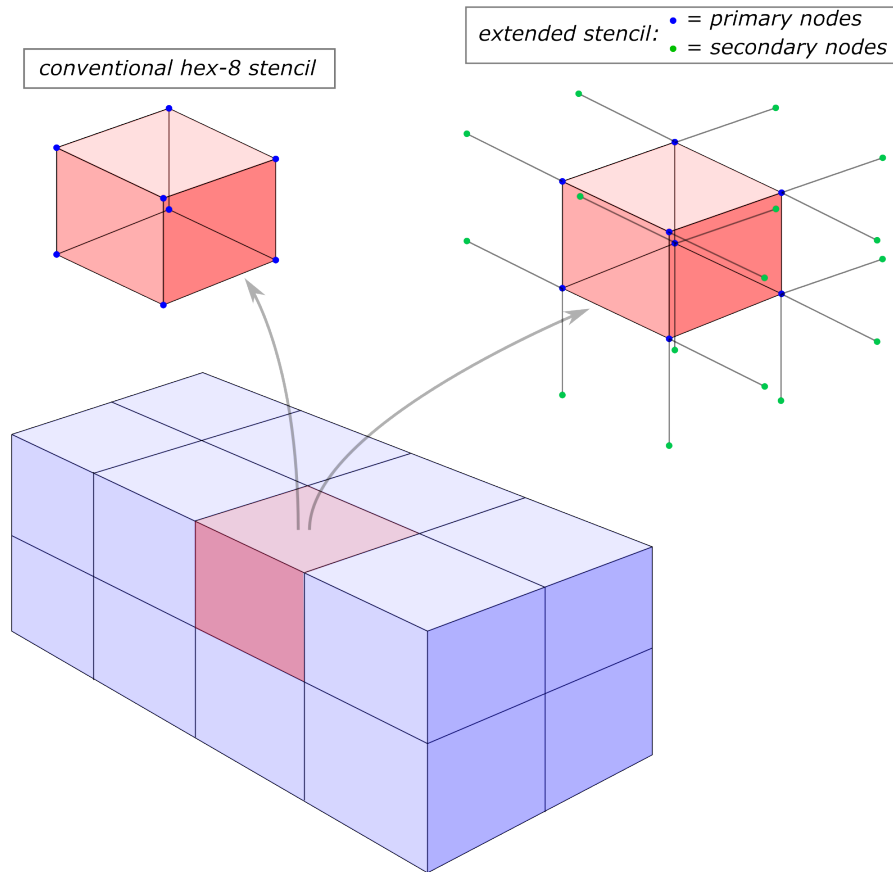
Figure 4.1: Comparison of a CFEM hex-8 compared to an ESFEM hex-8 with primary and secondary nodes.

can remain of full-rank while maintaining 2x2x2 point quadrature. Increasing quadrature to 3x3x3 points would increase computational time, and should be avoided if possible.

Initial attempts involved the addition of ghost nodes until each primary node was connected to at least six other nodes by an edge. These edges could be either an element edge or a "ghost edge" (connecting the primary node to the ghost node). Although, this approach led to an under-integrated stiffness matrix, and therefore rank deficiency.

So, a more refined strategy was adopted to frugally introduce ghost nodes while still avoiding rank deficiency. The addition of ghost nodes is done by systematically traversing each node within the mesh, guided by a specific set of rules outlined below. To facilitate comprehension, the node under consideration during the traversal will be designated as the

"planet" node. All nodes connected via edges will be designated as "satellites" to the planet node. Note that the details of these rules rely on the element type. The prescribed rules for the addition of ghost nodes are as follows:

1. The planet node must have less than six (hex-8) or eight (tet-4) existing satellites

2. The angle between the vector originating from the planet node to the ghost node (ghost vector) and the vectors from the planet node to the existing satellites (satellite vectors) must exceed a predefined threshold (i.e. the existing satellites must be heavily weighted in one direction relative to the planet node)

3. Each planet node is permitted up to one ghost node

Upon satisfying rule one, and progressing to the second, vectors are generated to define the directions extending from the planet node to each satellite. Fig. 4.2 depicts the front top corner primary node of the ESFEM element in Fig. 4.1 as the planet node with its corresponding satellites. This planet node has four satellite nodes, making it a candidate for a ghost node. A trial ghost node is deployed to maximize the minimal angle formed between the trial ghost vector and any of the satellite vectors. This angle is termed here to be the "maximal minimum angle." To further clarify, the maximal minimum angle is the minimum angle between the trial ghost vector and all satellite vector that is the largest. This strategic selection aims to increase the spatial distribution of nodes neighboring the planet node.

Subsequently, the angle derived from this procedure is evaluated and compared to a threshold value of 15°. This is the minimum angular separation required between the ghost vector and the nearest satellite vector. If this threshold is met, one ghost node is placed along the ghost vector direction at a distance equivalent to the average distance between all existing satellites and the planet node. If the threshold is not satisfied, meaning an inadequate angular separation, the introduction of a ghost node is surrendered, and the next primary node assumes the role of the planet node. The implementation of ghost nodes

through this approach increases the spatial dispersion of nodes around the planet node. The intention here is to facilitate a more accurate representation of complex deformations.
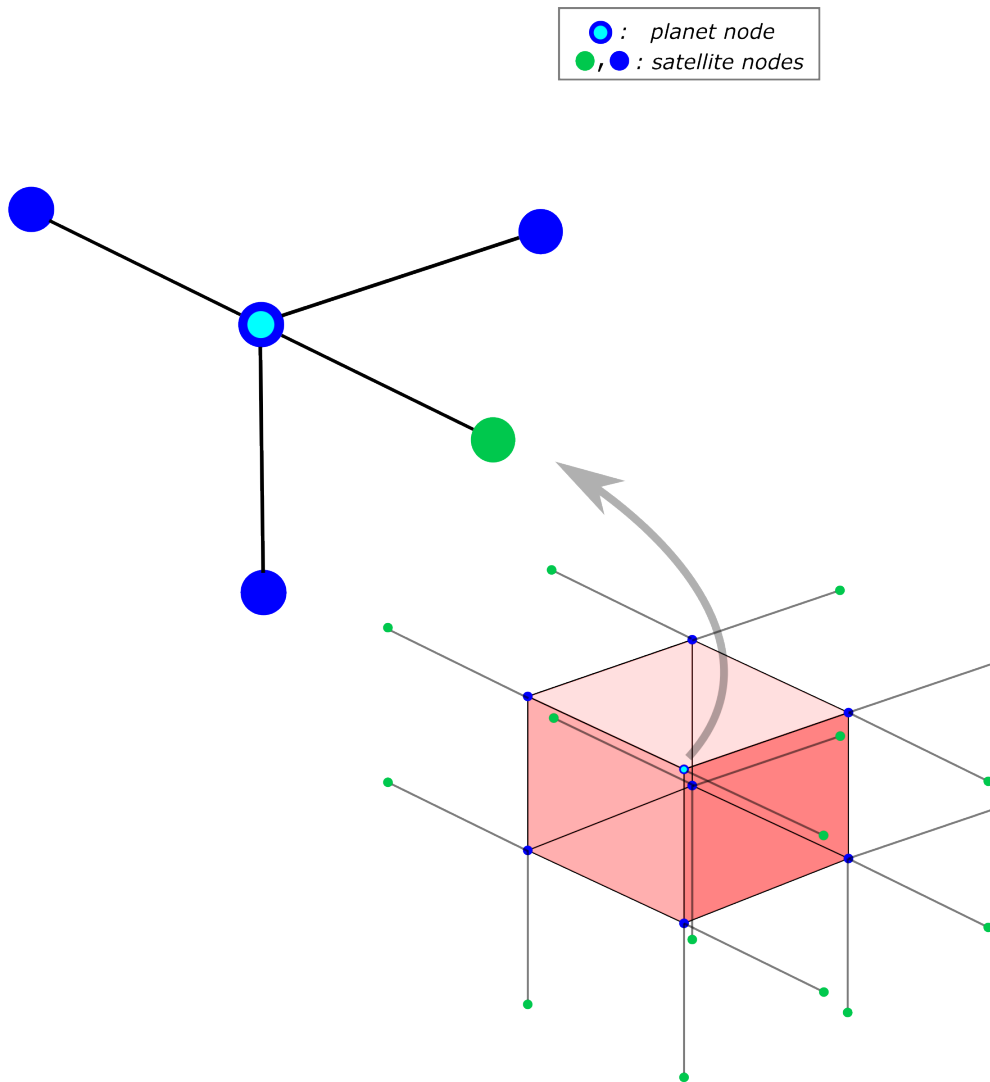
Figure 4.2: The front top corner node of the highlighted element of Fig. 4.1 as a planet node and its satellite nodes before ghost node generation.

To examine the maximal minimum angle in greater detail, consider the planet node depicted in Fig. 4.2. Then, satellite vectors are established as depicted in Fig. 4.3. The trial ghost node is strategically positioned in such a manner that the trial ghost vector is the largest minimum angle possible from any of the satellite vectors. In this mesh, and under the assumption of zero mesh distortion, the ghost vector would yield an identical minimum

angle of 135° for the primary nodes positioned both to the right and below the planet node, as depicted in Fig. 4.3.
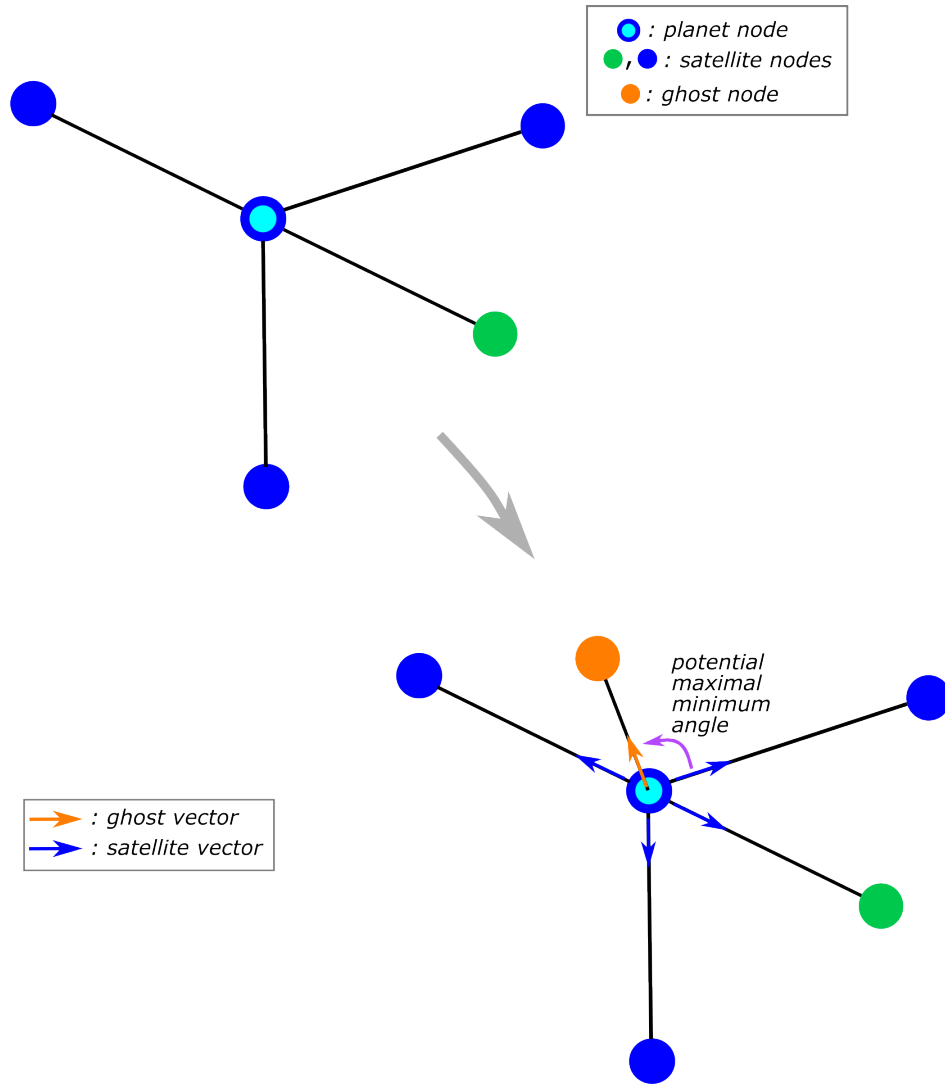
Figure 4.3: A ghost node is placed in this configuration if the maximal minimum angle is at least 15°

The above procedure is iteratively applied to each primary node within the mesh, ensuring that each node is considered as a planet node only once. In cases where ghost secondary nodes are introduced, the elements and faces containing the planet node are augmented to include the ghost node as a secondary node. In the mesh of Fig. 4.1, the repetition of this process yields the incorporation of ghost nodes as depicted in Fig. 4.4.
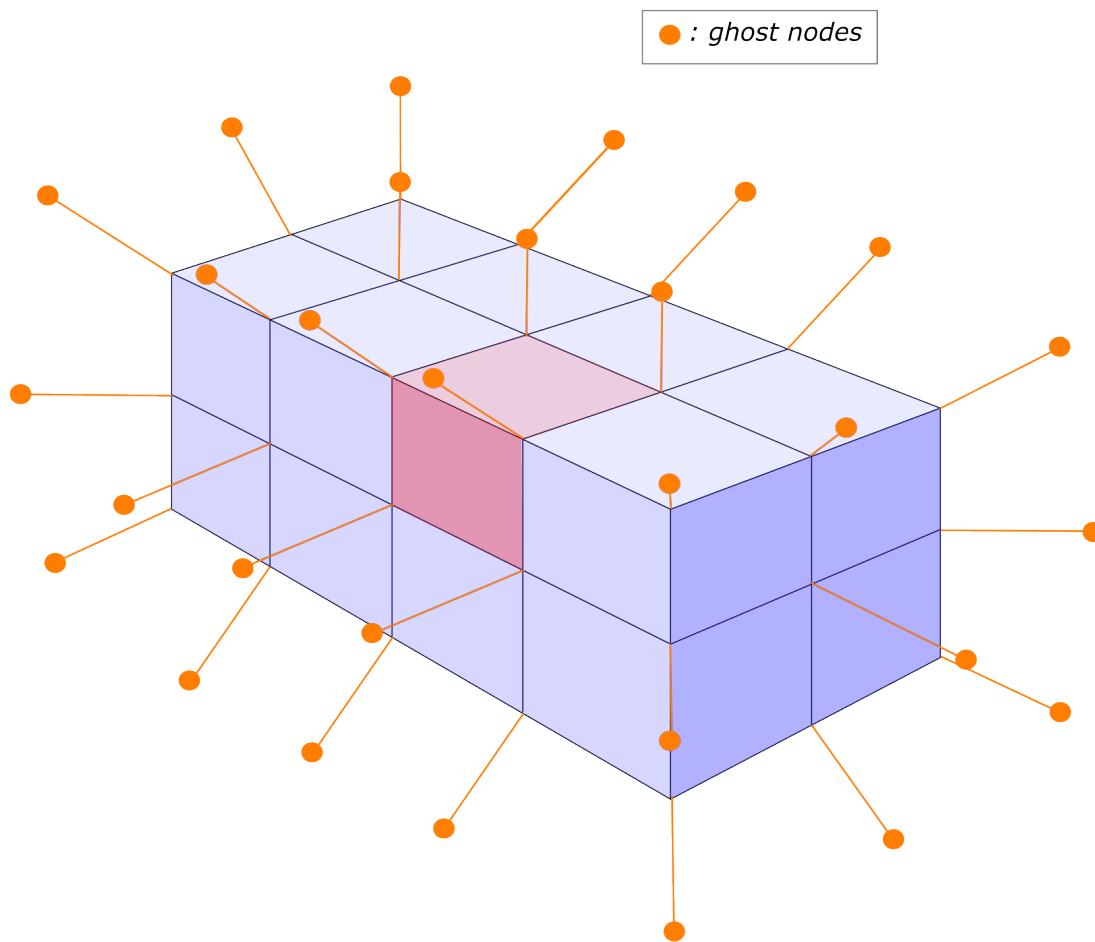
Figure 4.4: Hex-8 mesh with ghost nodes added that adhere to the rules of ghost node generation described in Section 4.1.2. While it cannot be seen, the ghost nodes exist in a similar fashion on the back side and bottom of the mesh.

4.2 Polynomial Fitting to a 3D Cloud of Points

To formulate the element interpolant in the ESFEM framework, a polynomial fit to the nodal values is needed. Consider a set of N points in three-dimensional space (nodes in this case), nodal coordinates $\mathbf{x}_a$, and associated nodal values $\mathbf{p}_a$, with subscript a ranging from 1 to N.

The objective is to fit a complete polynomial expression of the form

$$f(\mathbf{x}) = \sum_{i=1}^{n} c_i m_i(\mathbf{x}) \tag{4.1}$$

to the nodal data, where $m_i(\mathbf{x})$ are monomials in the Cartesian coordinates (x_1, x_2, x_3) and c_i are monomial coefficients. Since working with complete polynomial interpolants,

$$n = \frac{(k+1)(k+2)(k+3)}{6} \tag{4.2}$$

where k is the polynomial degree. The aim is to establish a mapping denoted as $\mathbf{K}$, which establishes a linear relationship between a vector encompassing nodal values and a vector composed of polynomial coefficients as

$$\mathbf{c} = \mathbf{Kp}. \tag{4.3}$$

The basic tool used to calculate $\mathbf{K}$ is least squares, which seeks to minimize the nodal interpolation error given by

$$\mathcal{F} = \sum_{a} (c_i m_{ia} - p_a)^2, \tag{4.4}$$

where $m_{ia} = m_i(\mathbf{x}_a)$, and summation over a repeated monomial index $()_i$ is implied.

It will prove instructive to first consider a straightforward minimization of Eq. 4.4. Minimization of this expression with respect to c_i yields

$$\mathbf{Ac} = \mathbf{Mp}, \quad \mathbf{A} = \mathbf{MM}^T. \tag{4.5}$$

The general solution to this equation for $\mathbf{c}$ is

$$\mathbf{c} = \mathbf{Kp} + \mathbf{d}, \quad \mathbf{K} = \mathbf{A}^+\mathbf{M}, \tag{4.6}$$

where $\mathbf{M}$ is a matrix of m_{ia} values,

$$\mathbf{M} = \begin{bmatrix} m_1(\mathbf{x}_1) & m_1(\mathbf{x}_2) & m_1(\mathbf{x}_3) & ... & m_1(\mathbf{x}_N) \\ m_2(\mathbf{x}_1) & m_2(\mathbf{x}_2) & m_2(\mathbf{x}_3) & ... & m_2(\mathbf{x}_N) \\ m_3(\mathbf{x}_1) & m_3(\mathbf{x}_2) & m_3(\mathbf{x}_3) & ... & m_3(\mathbf{x}_N) \\ ... & ... & ... & ... & ... \\ m_n(\mathbf{x}_1) & m_n(\mathbf{x}_2) & m_n(\mathbf{x}_3) & ... & m_n(\mathbf{x}_N) \end{bmatrix}, \tag{4.7}$$

and $\mathbf{A}^+$ is the Moore-Penrose pseudo-inverse of $\mathbf{A}$. The Moore-Penrose pseudo-inverse is a mathematical concept that extends the idea of matrix inversion to encompass cases where the matrix is not square, may be singular, or may not have a unique inverse. In this specific case, it provides a generalized solution for solving systems of linear equations, even when a true matrix inverse does not exist. The vector $\mathbf{d}$ is an arbitrary member of a particular subspace of $\mathbb{R}^n$, and will be discussed shortly. For now, it bears emphasis that the choice $\mathbf{d} = \mathbf{0}$ results in satisfaction of Eq. 4.5.

The $n \times N$ least-squares map $\mathbf{K}$ can be reduced through the use of a pseudo-inverse identity that holds for all matrices:

$$\mathbf{B}^+ = (\mathbf{B}^T \mathbf{B})^+ \mathbf{B} \quad \forall \quad \mathbf{B} \quad \Rightarrow \quad \mathbf{K} = \mathbf{M}^{T+}. \tag{4.8}$$

Note that the transpose and pseudo-inverse operations commute.

We now take up the possible nonuniqueness of the map from nodal values to polynomial coefficients in Eq. 4.6, as expressed by the presence of $\mathbf{d}$. Begin by recalling some facts about linear maps and the Moore-Penrose pseudo-inverse. $\mathbf{M}^T$ creates the subspace map $M_C \to M_R$ and $\mathbf{M}^{T+}$ maps $M_R \to M_C$ (note that the C and R are swapped due to the

transpose). The following facts about these vector subspaces can then be defined:

$$M_C \subseteq \mathbb{R}^n = M_C \bigoplus M_C^{\perp}$$
$$M_R \subseteq \mathbb{R}^N = M_R \bigoplus M_R^{\perp} \qquad (4.9)$$
$$\dim(M_C) = \dim(M_R) \leq \min(n, N).$$

$M_C^{\perp}$ and $M_R^{\perp}$ are referred to as orthogonal complements of M_C and M_R, respectively. In the case of equality on the last line of Eq. 4.9, $\mathbf{M}^T$ is said to be of full rank, otherwise it is rank deficient. If $\mathbf{d} \in M_C^{\perp}$, then it follows that $\mathbf{M}^T\mathbf{d} = \mathbf{0}$, and the first of Eq. 4.5 remains satisfied. On the other hand, if $\mathbf{d} \in M_C$ and $\mathbf{d} \neq \mathbf{0}$, then $\mathbf{Ad} \in M_C$ and is nonzero, thereby the solution to Eq. 4.5 is no longer adequate. We conclude then that $\mathbf{d}$ can be chosen from $M_C^{\perp}$ arbitrarily.

The best-fit coefficient vector $\mathbf{c}$ is nonunique under the condition that $M_C^{\perp} \neq \emptyset$, which in turn means that $M_C \neq \mathbb{R}^n$. In essence, this condition states that the set of monomials evaluated at each node does not provide a spanning set of n-vectors for all of $\mathbb{R}^n$. This is certain to happen when there are not enough nodes for the number of monomials (i.e. $N < n$). Although, it may also happen when $N \geq n$, if the nodal positions exhibit certain degeneracies. As an example, consider the case where $k \geq 2$, giving $n = 10$, and the nodes are placed exclusively on two parallel planes that are normal to the z-direction (see Fig. 4.5). Even in scenarios where an abundance of nodes is distributed across these planes, the available nodal values remain inadequate to uniquely establish a variation in the z-direction beyond a linear relationship.

In scenarios where nonuniqueness frequently arises, it remains vital to establish a procedure to compute a mapping from nodal values to polynomial coefficients. The overarching strategy is then to select $\mathbf{d}$ by involving the minimization of an auxiliary quadratic function. The chosen function here is the mean-square-magnitude of the gradient evaluated at the nodal points, with respect to $\mathbf{d}$. This approach proves advantageous when the dimensionality of the system satisfies $n > N$.
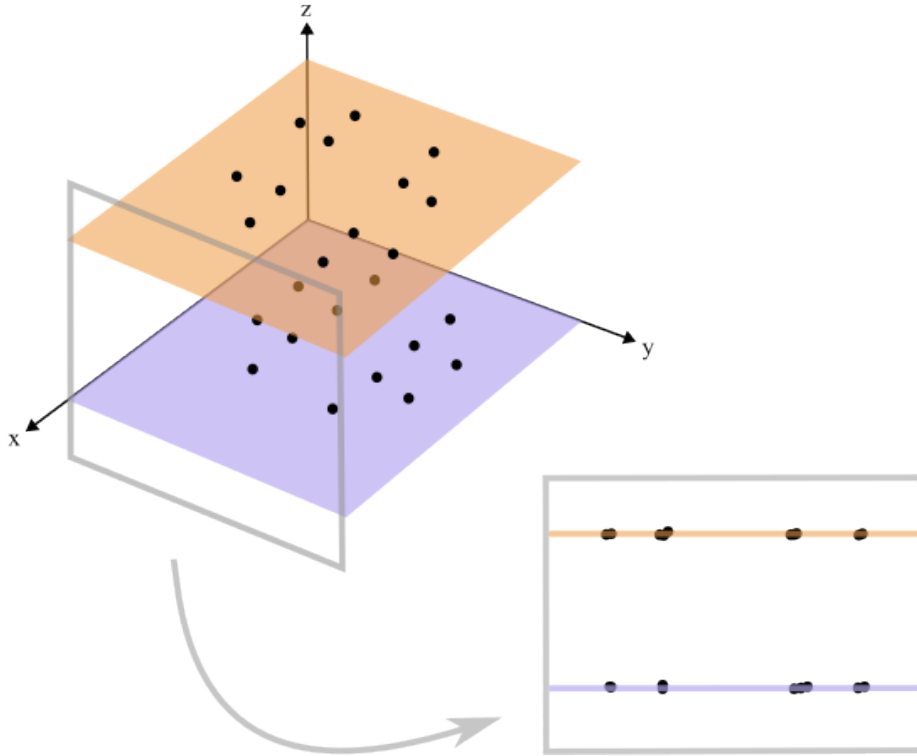
Figure 4.5: Two parallel planes normal to the z-direction with nodes placed strictly on each plane

In Section 4.3, the polynomial fit on each individual face is developed. The analysis extends to a fitting procedure that involves all nodal values referenced by the element. This fitting process is subject to a set of constraints dependent on the vector $\mathbf{d}$ that help to achieve higher levels of accuracy in the displacement solution.

4.3   Element Basis Functions

The main distinction between the ESFEM and the CFEM is the element basis functions. In the CFEM, these functions are defined by predetermined shape functions associated with the particular element type. In the ESFEM, the basis functions utilize a polynomial fit to a distribution of points, governed by a set of specific constraints.

Two distinct categories of constraints are enforced: quadrature consistency and Shi's F-Test constraint. The quadrature consistency constraint is imperative to maintain alignment with the exact weak form, ensuring a consistent mathematical formulation. Shi's F-Test constraint guarantees convergence in accordance with the F-E-M-Test as introduced by Shi [35]. The F-E-M-Test is considerd a practical implementation of Stummel's Generalized Patch Test [36].

To enforce these constraints, polynomial expressions are developed for each individual face of the element. These polynomials are determined through a polynomial fitting process that depends on the nodal values attributed to the primary and secondary nodes of the respective face. The primary and secondary nodes of a face are depicted in Fig. 4.6.

The following sections provide comprehensive explanations, including both the methodology behind the polynomial face fitting process, and the development of the element basis functions. The amalgamation of these features is a crucial aspect of the ESFEM framework, contributing to its novel approach and capabilities.

4.3.1 Polynomial Fit to Nodal Values

In this section, a generalized approach for polynomial fitting is developed, applicable to both the polynomial fitting process for the face polynomials and the formulation of element basis functions. In the context of the ESFEM, the number of available primary and secondary nodes is often substantially fewer than the count of polynomial coefficients. As a result, this approach encounters complexity in the three-dimensional polynomial fit to an underdetermined system of points. This fitting leads to a nonunique solution.

While the polynomial fitting process is straightforward when the size of the point cloud surpasses the number of coefficients in the polynomial, this is not the case within the ESFEM. Notably, the polynomial being determined is a moderately high-order polynomial. A 5th order polynomial is used throughout the numerical examples in Chapter 5, and is characterized by 56 coefficients. However, each individual face of a hex-8 element references at

most 20 primary and secondary nodes.

Consider the mesh depicted in Fig. 4.6, where the face between elements 1 and 2 features 16 associated nodes. In this configuration, a challenge arises: a unique solution cannot be obtained, and the task of determining the lowest order polynomial solution that minimizes oscillations becomes difficult. This predicament may lead to the generation of polynomial coefficient values of considerable magnitude, inevitably resulting in inaccurate results.
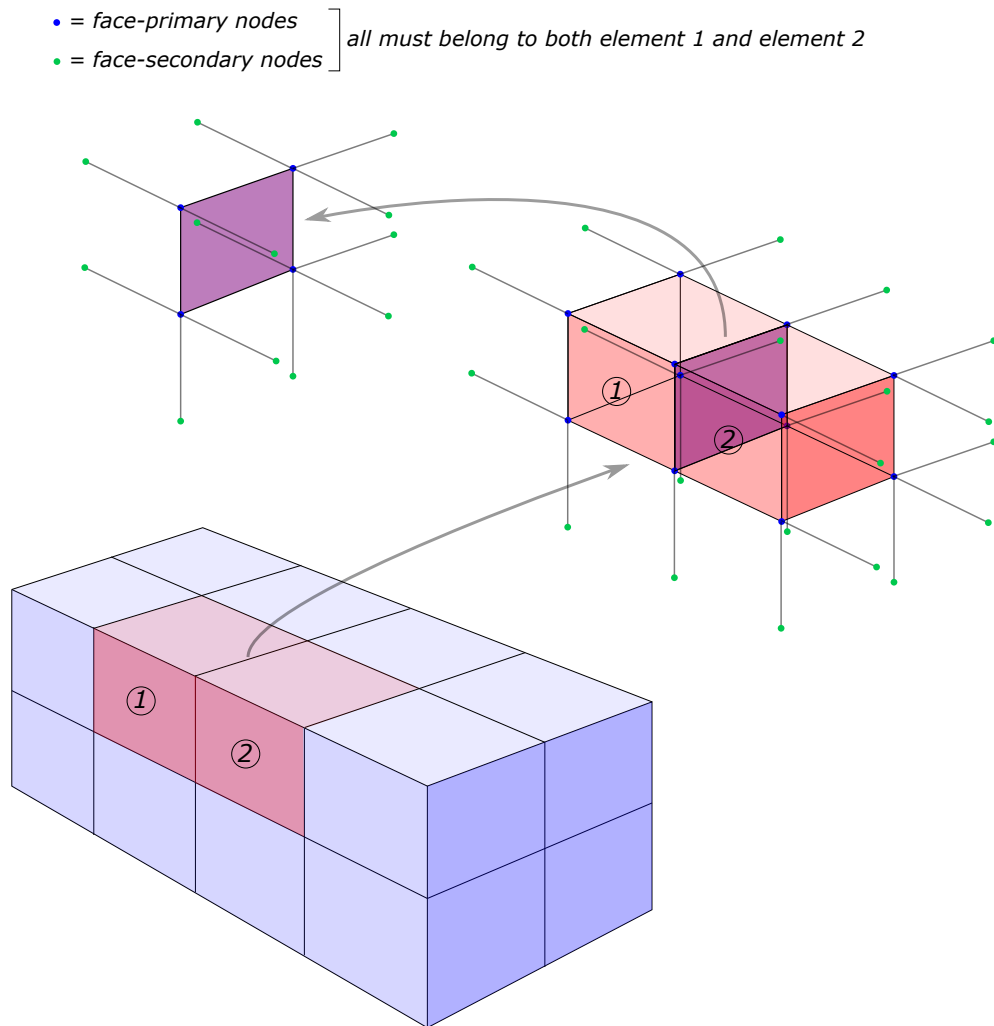

Figure 4.6: Face-secondary and primary nodes of the face that is shared between elements 1 and 2

Consider Eq. 4.10,

$$\mathbf{p} \in \mathbb{R}^N, \ \mathbf{c} \in \mathbb{R}^n, \tag{4.10}$$

where the nodal values attributed to the associated primary and secondary nodes are organized within the vector $\mathbf{p}$, such that N denotes the total count of primary and secondary nodes for the given object (face or element). The vector of polynomial coefficients is represented by $\mathbf{c}$, with n denoting the number of monomials encompassed within a complete polynomial.

Furthermore, define the matrix $\mathbf{M}$ to be that which is given in Eq. 4.7. Then, the nodal values $\mathbf{p}$ for a given $\mathbf{c}$ is represented by $\mathbf{M}^T\mathbf{c}$. Also define the inner product $< \cdot, \cdot >$ on $\mathbb{R}^n$ by

$$< \mathbf{c}, \mathbf{d} >= \mathbf{c}^T \mathbf{M}\mathbf{M}^T \mathbf{d} = \mathbf{p} \cdot \mathbf{q}, \tag{4.11}$$

where $\mathbf{M}\mathbf{M}^T$ is positive semi-definite and

$$\mathbf{p} = \mathbf{M}^T\mathbf{c}, \ \mathbf{q} = \mathbf{M}^T\mathbf{d}. \tag{4.12}$$

Then, the Gram-Schmidt orthogonalization process is utilized on the identity tensor $\mathbf{I} \in \mathbb{R}^{n \times n}$ with the inner product defined by Eq. 4.11. The outcome is $\mathbf{B} \in \mathbb{R}^{n \times m}$ in the form:

$$\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots \mathbf{b}_m], \tag{4.13}$$

where $\mathbf{b}_i$ represents the orthonormal basis vectors of the subspace $\mathbb{R}^n$ containing $\mathbf{c}$'s that produce nonzero values at the nodes. Thus, the columns of $\mathbf{B}$ span the subset $\mathcal{V} \subset \mathbb{R}^n$, where $\mathbf{c} \in \mathcal{V}$.

Now, define the subspace $\mathcal{V}^\perp$, wherein the $\mathbf{c} \in \mathcal{V}^\perp$ yield zero values at the nodes. Referring to Eq. 4.9, it is evident the relationship $\mathbb{R}^n = \mathcal{V} \oplus \mathcal{V}^\perp$ holds, as $\mathcal{V}$ and $\mathcal{V}^\perp$ are orthogonal complements.

Now, let the subspace $\mathcal{W} \subset \mathbb{R}^N$ define the sets of nonzero nodal values $\mathbf{p}, \mathbf{q} \in \mathcal{W}$. Thus,

$\mathbf{M}^T$ maps $\mathcal{V} \to \mathcal{W}$. Then suppose $\mathbf{c}, \mathbf{d} \in \mathcal{V}$ and $< \mathbf{c}, \mathbf{d} > = 0$. Moreover, let $\mathbf{p} = \mathbf{M}^T\mathbf{c}$ and $\mathbf{q} = \mathbf{M}^T\mathbf{d}$. Then,

$$
\begin{aligned}
\mathbf{p} \cdot \mathbf{q} &= \mathbf{c}^T\mathbf{M}^T\mathbf{M}\mathbf{d} \\
&= < \mathbf{c}, \mathbf{d} > \\
&= 0,
\end{aligned}
\tag{4.14}
$$

and $\mathbf{B}$ generates an orthonormal basis of $\mathcal{W}$, under the inner product in Eq. 4.11, by way of $\mathbf{M}^T$. To elaborate further, it can be specified that the columns of $\mathbf{M}^T\mathbf{B}$ constitute an orthonormal basis for $\mathcal{W}$ with respect to the inner product $< \cdot, \cdot >$. Projection operators are as outlined below:

$$
\begin{aligned}
\mathbf{M}^T\mathbf{B}(\mathbf{M}^T\mathbf{B})^T = \mathbf{M}^T\mathbf{B}\mathbf{B}^T\mathbf{M} \, &\text{projects} \, \mathbb{R}^n \, \text{onto} \, \mathcal{W} \\
\mathbf{B}\mathbf{B}^T\mathbf{M}\mathbf{M}^T \, &\text{projects} \, \mathbb{R}^n \, \text{onto} \, \mathcal{V} \\
\mathbf{I} - \mathbf{B}\mathbf{B}^T\mathbf{M}\mathbf{M}^T \, &\text{projects} \, \mathbb{R}^n \, \text{onto} \, \mathcal{V}^\perp
\end{aligned}
$$

It is imperative to note that $\mathbf{M}^T$ maps $\mathcal{V}^\perp \to \emptyset$. More specifically, $\mathcal{V}^\perp$ is the nullspace of $\mathbf{M}^T$.

Now, suppose the following:

$$
\begin{aligned}
\mathbf{c} &= \mathbf{B}\mathbf{a}, \quad \mathbf{a} \in \mathbb{R}^m \\
\mathbf{p} &= \mathbf{B}\mathbf{M}^T\mathbf{a},
\end{aligned}
\tag{4.15}
$$

such that $\mathbf{a}$ contains the coefficients in the orthonormal basis. Taking into account the orthonormality of the columns of $\mathbf{B}$ with respect to the inner product defined in Eq. 4.11, the following is presented. Begin with the relation

$$
\mathbf{B}^T\mathbf{M}\mathbf{M}^T\mathbf{B} = \mathbf{I} \in \mathbb{R}^{m \times m}.
\tag{4.16}
$$

$$\therefore \mathbf{B}^T \mathbf{M} \mathbf{p} = \mathbf{B}^T \mathbf{M} \mathbf{B} \mathbf{M}^T \mathbf{a}$$

$$= \mathbf{B}^T \mathbf{M} \mathbf{M}^T \mathbf{B} \mathbf{a}$$

$$= \mathbf{I} \mathbf{a}$$

$$= \mathbf{a}$$

$$\Rightarrow \mathbf{a} = \mathbf{B}^T \mathbf{M} \mathbf{p} \tag{4.17}$$

$$\Rightarrow \mathbf{c} = \mathbf{B} \mathbf{B}^T \mathbf{M} \mathbf{p} \tag{4.18}$$

$$\mathbf{K} = \mathbf{B} \mathbf{B}^T \mathbf{M} \tag{4.19}$$

In the above statement, the resulting $\mathbf{K}$ is a mapping from nodal values to polynomial coefficients that minimizes error at the nodes and produces linear exactness. However, it is crucial to note that this map does not reflect any constraints. Thus, the matrix $\mathbf{K}$ described above serves as the appropriate mapping for the development of the polynomial fit on the faces. Although, for the element interpolants within the ESFEM framework, further elaboration and effort are required.

Let $\mathbf{x} \in \mathbb{R}^n$, and thus the projection under the inner product defined in Eq. 4.11 of $\mathbf{x}$ onto $\mathcal{V}$ be defined as $\overline{\mathbf{x}} \in \mathcal{V}$:

$$\overline{\mathbf{x}} = \sum_{i=1}^{m} <\mathbf{b}_i, \mathbf{x}> \mathbf{b}_i, \tag{4.20}$$

where $\mathbf{b}_i$ is the ith column of $\mathbf{B}$. Hence, we can conclude that

$$\overline{\mathbf{x}} = \sum_{i=1}^{m} (\mathbf{b}_i \mathbf{M} \mathbf{M}^T \mathbf{x}) \mathbf{b}_i \tag{4.21}$$

$$\Rightarrow \overline{x}_l = [B_{ji} M_{ja} M_{ka} x_k] B_{li}$$

$$= B_{li} B_{ji} M_{ja} M_{ka} x_k$$

$$\therefore \overline{\mathbf{x}} = \mathbf{BB}^T \mathbf{MM}^T \mathbf{x} \tag{4.22}$$

Given the relationship of Eq. 4.22, it is now possible to demonstrate that $\mathbf{x} - \overline{\mathbf{x}}$ has a zero projection onto subspace $\mathcal{V}$:

$$\mathbf{BB}^T \mathbf{MM}^T (\mathbf{x} - \overline{\mathbf{x}}) = \overline{\mathbf{x}} - \mathbf{BB}^T \mathbf{MM}^T [\mathbf{BB}^T \mathbf{MM}^T \mathbf{x}]$$

$$= \overline{\mathbf{x}} - \mathbf{B} \underbrace{[\mathbf{B}^T \mathbf{MM}^T \mathbf{B}]}_{\mathbf{I},\, \text{Eq. 4.16}} \mathbf{B}^T \mathbf{MM}^T \mathbf{x}$$

$$= \overline{\mathbf{x}} - \mathbf{BB}^T \mathbf{MM}^T \mathbf{x}$$

$$= \overline{\mathbf{x}} - \overline{\mathbf{x}}$$

$$= \mathbf{0}$$

Subsequently, let $\mathbf{P} = \mathbf{I} - \mathbf{BB}^T \mathbf{MM}^T$ be the projection from $\mathbb{R}^n$ onto $\mathcal{V}^\perp$, and the matrix $\mathbf{D} \in \mathbb{R}^{n \times n}$ be the mean square gradient on the element:

$$\mathbf{D} = \int_\Omega \nabla \mathbf{m}(\mathbf{x}) \cdot \nabla \mathbf{m}(\mathbf{x}) \, dv \tag{4.23}$$

It becomes feasible to derive an orthonormal basis for $\mathcal{V}^\perp$ by use of the Gram-Schmidt orthogonalization process on the matrix $\mathbf{P}$. It is noteworthy that this instance of Gram-Schmidt employs the inner product, $< \cdot, \cdot >_D$ defined in Eq. 4.24, to yield $\overline{\mathbf{B}} \in \mathbb{R}^{n \times \overline{m}}$, whose columns are an orthonormal basis of $\mathcal{V}^\perp$ under $< \cdot, \cdot >_D$.

$$< \mathbf{c}, \mathbf{d} >_D = \mathbf{c}^T \mathbf{D} \, \mathbf{d} \tag{4.24}$$

Recall that the desired outcome of the polynomial fit is the establishment of a mapping denoted by $\mathbf{K}$, which maps the nodal values $\mathbf{p}$ to the polynomial coefficients $\mathbf{c}$. In essence, the objective is to obtain a matrix $\mathbf{K} \in \mathbb{R}^{n \times N}$ that aligns with the formulation presented in Eq. 4.25:

$$\mathbf{c} = \mathbf{Kp}. \tag{4.25}$$

In Eq. 4.19, no constraints were considered in its formulation. The following discussion outlines the approach taken to enforce the constraints of the form $\mathbf{Gc} = \mathbf{Hp}$.

Drawing inspiration from the utilization of Eq. 4.6 and the insights regarding orthogonal complements in Section 4.2, the vector $\mathbf{c}$ is partitioned into two distinct components:

$$\mathbf{c} = \hat{\mathbf{c}} + \overline{\mathbf{c}}, \tag{4.26}$$

where $\hat{\mathbf{c}} \in \mathcal{V}$ is chosen to minimize the sum of squared errors associated with nodal values, as presented in Eq. 4.4. $\overline{\mathbf{c}} \in \mathcal{V}^\perp$ is determined through the utilization of the matrix $\mathbf{D}$, wherein $\overline{\mathbf{c}}^T \mathbf{D} \overline{\mathbf{c}}$ is minimized. This minimization process is conducted while enforcing constraints of the type $\mathbf{Gc} = \mathbf{Hp}$, a process facilitated by the adjustment of any residual degrees of freedom that remain within the polynomial fit after the minimization of nodal errors. Due to the typically underdetermined nature of the system in the ESFEM framework, there is a certain degree of "residual freedom" that can be harnessed to satisfy these constraints.

Stemming from the partition of $\mathbf{c}$, the concept of the map $\mathbf{K}$ is partitioned into $\hat{\mathbf{K}}$ and $\overline{\mathbf{K}}$. These two map nodal values to polynomial coefficients $\hat{\mathbf{c}}$ and $\overline{\mathbf{c}}$, respectively. Eq. 4.25 then becomes the following:

$$\mathbf{c} = \hat{\mathbf{K}}\mathbf{p} + \overline{\mathbf{K}}\mathbf{p} \tag{4.27}$$

Given that $\overline{\mathbf{c}} \in \mathcal{V}^\perp$, $\overline{\mathbf{c}}$ is zero at the nodes, and $\hat{\mathbf{c}} \in \mathcal{V}$ exhibits strictly nonzero values at the nodes. As a result, the matrix $\hat{\mathbf{K}}$ in Eq. 4.28 takes care of the mapping process that involves minimizing nodal errors and produces linear exactness, while $\overline{\mathbf{K}}$ manages the mapping process related to the enforcement of constraints.

$$\hat{\mathbf{K}} = \mathbf{B}\mathbf{B}^T\mathbf{M} \tag{4.28}$$

Now, proceed with the process of enforcing the constraint $\mathbf{Gc} = \mathbf{Hp}$ using the combined

efforts of $\overline{\mathbf{K}}$ and minimizing $< \overline{\mathbf{c}}, \overline{\mathbf{c}} >_D$. Begin with

$$\mathbf{G}(\hat{\mathbf{c}} + \overline{\mathbf{c}}) = \mathbf{Hp}. \tag{4.29}$$

Through rearrangement, and inclusion of $\hat{\mathbf{c}} = \hat{\mathbf{K}}\mathbf{p}$, $\overline{\mathbf{H}}$ is defined:

$$\overline{\mathbf{H}} = \mathbf{H} - \mathbf{G}\hat{\mathbf{K}} \tag{4.30}$$

Now, obtain $\overline{\mathbf{B}}$ using the Gram-Schmidt orthogonalization process on the matrix $\mathbf{P}$, as defined to be the projection of $\mathbb{R}^n$ onto $\mathcal{V}^\perp$, with inner product $< \cdot, \cdot >_D$. Given $\overline{\mathbf{B}}$, along with $\overline{\mathbf{c}} = \overline{\mathbf{B}}\overline{\mathbf{a}}$, the constraints are defined as follows:

$$\overline{\mathbf{G}}\overline{\mathbf{a}} = \overline{\mathbf{H}}\mathbf{p}, \quad \overline{\mathbf{G}} \equiv \mathbf{G}\overline{\mathbf{B}} \tag{4.31}$$

Utilizing the Moore-Penrose pseudo-inverse to solve for $\overline{\mathbf{a}}$, one obtains Eq. 4.32 where $\overline{\mathbf{N}} = \mathbf{I} - \overline{\mathbf{G}}^+ \overline{\mathbf{G}}$ and $\mathbf{t}$ is an arbitrary vector. If $\overline{\mathbf{G}}$ is invertible, then $\overline{\mathbf{N}}$ is just a zero matrix.

$$\overline{\mathbf{a}} = \overline{\mathbf{G}}^+ \overline{\mathbf{H}}\mathbf{p} + \overline{\mathbf{N}}\mathbf{t} \tag{4.32}$$

Then, define the objective function for minimization to be

$$\mathcal{F} = < \overline{\mathbf{c}}, \overline{\mathbf{c}} >_D = \overline{\mathbf{a}} \cdot \overline{\mathbf{a}}. \tag{4.33}$$

By substituting in the expression from Eq. 4.32, then minimizing this modified objective function with respect to the vector $\mathbf{t}$, the following results are obtained:

$$\overline{\mathbf{a}} = \overline{\mathbf{G}}^+ \overline{\mathbf{H}}\mathbf{p} + \overline{\mathbf{N}}\overline{\mathbf{N}}^T \overline{\mathbf{G}}^+ \overline{\mathbf{H}}\mathbf{p} \tag{4.34}$$

$$= (\mathbf{I} - \overline{\mathbf{N}}\overline{\mathbf{N}}^T)\overline{\mathbf{G}}^+ \overline{\mathbf{H}}\mathbf{p} \tag{4.35}$$

Finally, using $\overline{\mathbf{c}} = \overline{\mathbf{B}}\overline{\mathbf{a}}$ along with $\overline{\mathbf{c}} = \overline{\mathbf{K}}\mathbf{p}$, the resulting map is formed:

$$\overline{\mathbf{K}} = \overline{\mathbf{B}}(\mathbf{I} - \overline{\mathbf{N}}\overline{\mathbf{N}}^T)\overline{\mathbf{G}}^+\overline{\mathbf{H}} \tag{4.36}$$

With the polynomial fit established and the constraints $\mathbf{Gc} = \mathbf{Hp}$ in mind, the mapping matrix $\mathbf{K}$ can be determined, which leads to the formulation of the element interpolant:

$$\begin{aligned}
\phi^e(\mathbf{x}) &= \sum_{i=1}^{n} c_i m_i(\mathbf{x}) \\
&= \sum_{i=1}^{n}\sum_{a=1}^{N} K_{ia} p_a m_i(\mathbf{x}) \\
&= \sum_{i=1}^{n}\sum_{a=1}^{N} (\hat{K}_{ia} + \overline{K}_{ia}) p_a m_i(\mathbf{x}),
\end{aligned} \tag{4.37}$$

where the element interpolant is now constructed based on the polynomial fit discussed above. The next section addresses how to populate the $\mathbf{G}$ and $\mathbf{H}$ matrices with the specific constraints required of the ESFEM.

4.3.2 Element Interpolant Constraints

Element interpolant constraints are expressed in the form $\mathbf{Gc} = \mathbf{Hp}$ and depend on the existence of "residual freedom." In the context of a quintic polynomial, characterized by 56 monomial coefficients, there exist 56 "knobs of freedom" that can be manipulated to achieve the desired solution.

Consider an element on the interior of a mesh, entirely surrounded by other elements. This element will have what can is termed a "complete set" of secondary nodes. Then, a total of 32 nodes are referenced for a hex-8 mesh. Even if the process of minimizing nodal errors involves the utilization of 32 out of the available 56 degrees of freedom, a surplus of up to 24 degrees still remains.

The constraints are then utilized to manage this residual freedom. More specifically, up

to 3 constraints contribute to ensuring quadrature consistency, while an additional 4 (tet-4) or 6 (hex-8) constraints belong to what is called here, Shi's F-E-M-Test constraint. These constraints work together to fine-tune the polynomial fit.

The constraints associated with quadrature consistency ensure consistency with the exact weak form. This is obtained by guaranteeing the validity of the divergence theorem for the element interpolant. This correlates to the following equality:

$$\int_{\Omega} \phi_{,i}^e \, dv = \int_{\partial \Omega} \phi^f n_i \, da \quad \forall \quad i = 1, 2, 3, \tag{4.38}$$

where ϕ^e represents the element interpolant, ϕ^f symbolizes the polynomial fit on each face, and n_i denotes the ith component of the normal vector pertaining to the ith face of ϕ^f. This equation fills the first three rows of the matrices $\mathbf{G}$ and $\mathbf{H}$.

To follow, Shi's F-E-M-Test constraint involves a distinct constraint for each face of the element, ensuring convergence through the F-E-M-Test [35], in the context of non-conforming elements. The test used for verifying the convergence of this method is known as the F_1-Test [35], which requires

$$\lim_{h_K \to 0} \frac{|\int_{\Gamma} [\phi_h^e] \, ds|}{h_K^{3/2} ||\phi_h^e||_{1, K_1 \cup K_2}} = 0 \tag{4.39}$$

across the shared face of two elements, denoted by $\Gamma = K_1 \cap K_2$. Here, h_K denotes the maximum of the diameters of the smallest spheres capable of encompassing the individual elements K_1 and K_2. The subscript "h" in Eq. 4.39 denotes the interpolation is dependent upon the mesh size. Moreover, $||\phi_h^e||_{1, K_1 \cup K_2}$ corresponds to the H^1 norm over the union of the two elements, defined as:

$$||\phi_h^e||_{1, K_1 \cup K_2} = \int_{K_1 \cup K_2} |\phi_h^e|^2 dx + \int_{K_1 \cup K_2} |\nabla \phi_h^e|^2 dx. \tag{4.40}$$

The term $[\phi_h^e]$ denotes the jump between the element interpolants on each individual element,

evaluated at the shared face Γ, given by

$$[\phi_h^e] = \phi_h^{e,K_1}|_\Gamma - \phi_h^{e,K_2}|_\Gamma. \tag{4.41}$$

The F-E-M-Test constraint implemented for the ESFEM requires that the mean difference of ϕ^e on each element and ϕ^f on each face of said element, Γ, is zero such that

$$\begin{aligned} \frac{1}{|\Gamma|} \int_\Gamma \phi_h^{e,K_1} + \phi_h^f \, ds &= 0 \\ \frac{1}{|\Gamma|} \int_\Gamma \phi_h^{e,K_2} + \phi_h^f \, ds &= 0. \end{aligned} \tag{4.42}$$

So, beginning with the numerator of Eq. 4.39,

$$\begin{aligned} |\int_\Gamma [\phi_h^e] \, ds| &= |\int_\Gamma \phi_h^{e,K_1} - \phi_h^{e,K_2} \, ds| \\ &= |\int_\Gamma \phi_h^{e,K_1} - \phi_h^{e,K_2} + \phi_h^f - \phi_h^f \, ds| \\ &= |\int_\Gamma \phi_h^{e,K_1} + \phi_h^f - \phi_h^{e,K_2} - \phi_h^f \, ds| \\ &= |\int_\Gamma \underbrace{(\phi_h^{e,K_1} + \phi_h^f)}_{0,\,\text{Eq. } 4.42} - \underbrace{(\phi_h^{e,K_2} + \phi_h^f)}_{0,\,\text{Eq. } 4.42} \, ds| \\ \Rightarrow |\int_\Gamma [\phi_h^e] \, ds| &= 0. \end{aligned}$$

Referring to Eq. 4.39, this gives

$$\frac{|\int_\Gamma [\phi_h^e] \, ds|}{h_K^{3/2} \, ||\phi_h^e||_{1,K_1 \cup K_2}} = 0, \tag{4.43}$$

and exact satisfaction of the F_1-Test. This condition contributes to the next 4 or 6 rows of the matrices $\mathbf{G}$ and $\mathbf{H}$ for tet-4 and hex-8 elements, respectively.

Chapter 5

Numerical Exploration of ESFEM

In this chapter, numerical examples are used to evaluate the ESFEM's accuracy in a variety of geometries, materials, and boundary conditions. Meshes have all been developed with hex-8 elements and some with tet-4 elements. All meshes in this chapter have been developed with Gmsh [12]. Note that all dimensions in the examples are of compatible units. The different element types and their short-hand descriptor are listed in Table 5.1.

Descriptor	Element Type	Formulation	Secondary Nodes?	Ghost Nodes?
C-8	Hex-8	CFEM	×	×
C-4	Tet-4	CFEM	×	×
C-27	Hex-27	CFEM	×	×
ES-8-N	Hex-8	ESFEM	×	×
ES-8-S	Hex-8	ESFEM	✓	×
ES-8-GS	Hex-8	ESFEM	✓	✓
ES-4-N	Tet-4	ESFEM	×	×
ES-4-S	Tet-4	ESFEM	✓	×
ES-4-GS	Tet-4	ESFEM	✓	✓

Table 5.1: Descriptors for the different element types used in this chapter, along with information about the capability for secondary and ghost nodes.

To evaluate the accuracy of a solution, displacement will be one measurement of solution convergence and accuracy. For problems with a theoretical displacement solution, the

displacement results will be presented as a normalized displacement such that

$$\text{normalized displacement} = \frac{\text{test results}}{\text{theoretical solution}}. \tag{5.1}$$

Additionally, an energy norm error is used to measure convergence of stresses for some hex-8 problems. The development of the energy norm error is explained in Section 5.1. Not all hexahedral problems will undergo this evaluation due to the large number of elements needed to produce a converged solution with the C-8 mesh.

5.1 Energy Norm Error

The energy norm error of a solution is based on the comparison of a "reference" mesh and a "test" mesh. The reference mesh is a refined C-8 mesh that has converged, while the test mesh is a coarser hex-8 mesh. Let the stress of the reference mesh be $\hat{\mathbf{T}}$, and the stress of the test mesh be $\mathbf{T}$. The energy norm error is shown in Eq. 5.3, where $\mathbb{B}_i$ is element i and $\mathbf{D}^{-1}$ is the inverse of

$$\mathbf{D} = \begin{bmatrix} \lambda + 2\mu & \lambda & \lambda & & & \\ \lambda & \lambda + 2\mu & \lambda & & \mathbf{0} & \\ \lambda & \lambda & \lambda + 2\mu & & & \\ & & & \mu & 0 & 0 \\ & \mathbf{0} & & 0 & \mu & 0 \\ & & & 0 & 0 & \mu \end{bmatrix} \tag{5.2}$$

$$e = \sum_{i}^{\#\,ref\,elems} \int_{\mathbb{B}_i} [\mathbf{T} - \hat{\mathbf{T}}]^T D^{-1} [\mathbf{T} - \hat{\mathbf{T}}] dv \tag{5.3}$$

The integration is done over the reference mesh. Integration points (IPs) of each element are given an "IP region," which partitions a hex-8 parent element into 8 equal cubes. Each one of these 8 cubes is associated with the IP that exists within it. To calculate the energy

norm error, the volume of the overlap between the test mesh and reference mesh IP regions is needed. To obtain these overlapping volumes, the test mesh elements are each divided into 1000 "microelements" in the parent configuration with a 10x10x10 grid of elements.

The microelements will reference the location of its centroid in the current configuration, volume in the current configuration, and the stress values of the closest test mesh and reference mesh IPs. The volume and centroid are calculated in the parent configuration, and the centroid is mapped to the current configuration using the element's isoparametric map. The current configuration volume is estimated with the Jacobian matrix evaluated at the centroid of the microelement: (parent volume) × det(Jacobian). The "closest" test and reference mesh IPs are those closest in distance to the microelement centroid in the current configuration. With this information, the energy norm error is evaluated as shown in Eq. 5.4, where w_I represents the volume of the microelement in the current configuration, $\hat{\mathbf{T}}_I$ is the stress at the closest reference mesh IP, and $\mathbf{T}_I$ is the stress at the closest test mesh IP.

$$e = \sum_{I=1}^{\#\,microelements} w_I [\mathbf{T}_I - \hat{\mathbf{T}}_I]^T D^{-1} [\mathbf{T}_I - \hat{\mathbf{T}}_I] \tag{5.4}$$

5.2   Distortion Parameter

To quantify the distortion in a mesh, a "distortion parameter" (DP) (see Eq. 5.5) is used. This parameter averages the departure of the determinant of the Jacobian (J_k) of the isoparametric transformation at each integration point (IP), k, from its average value over the element (J_{avg}). In Eq. 5.5, NIP is the number of integration points. For undistorted elements, J_k is constant at all element IPs, and therefore the parameter is zero. Since the tet-4 elements used in this dissertation have only one IP, the DP is always zero for these meshes. Therefore, DPs are only calculated for the hex-8 and hex-27 meshes.

$$\text{DP} = \frac{1}{\#\,elements} \times \sum_{\#\,elements} \sqrt{\frac{\sum_{k=1}^{NIP} (J_k - J_{avg})^2}{\sum_{k=1}^{NIP} (J_{avg})^2}} \tag{5.5}$$

5.3 Patch Test

The patch test for the hexahedral elements from [24] is set up with the following displacement boundary conditions (BCs):

$$u = 10^{-3}\left(\frac{2x+y+z}{2}\right)$$
$$v = 10^{-3}\left(\frac{x+2y+z}{2}\right) \tag{5.6}$$
$$w = 10^{-3}\left(\frac{x+y+2z}{2}\right)$$

where u is the displacement BC in the x-direction, v in the y-direction, and w in the z-direction. Moreover, the (x, y, z) values are the Cartesian coordinate locations of the nodes on the boundary of the system. The theoretical solution of the stresses is

$$\sigma_x = \sigma_y = \sigma_z = 2000$$
$$\tau_{xy} = \tau_{yz} = \tau_{zx} = 400. \tag{5.7}$$

The tet-4 patch test is inspired by the hexahedral patch test, consisting of a single, off-center interior node that the rest of the mesh is generated around to give a total of 12 elements. If a patch test produces correct results (in this case, appropriate stresses), then the results of any problem will converge to the correct solution as the mesh is refined. However, it is possible that even with passage of the patch test, an element's rate of convergence may be too slow for practical use [24]. Thus, a full suite of tests is required to verify an element's versatility.

5.3.1 Hex

This patch test passes for C-8, C-27, and ES-8-N with the method as-is. Note that the C-27 patch test is carried about in the same manner, just altered to account for the additional nodes. For ES-8-S and ES-8-GS, the effect of displacement BCs on the face of an element changes between the CFEM and ESFEM methods. In the CFEM treatment of displacement BCs, if all nodes on the face of an element are fixed, then the entire face is fixed. In
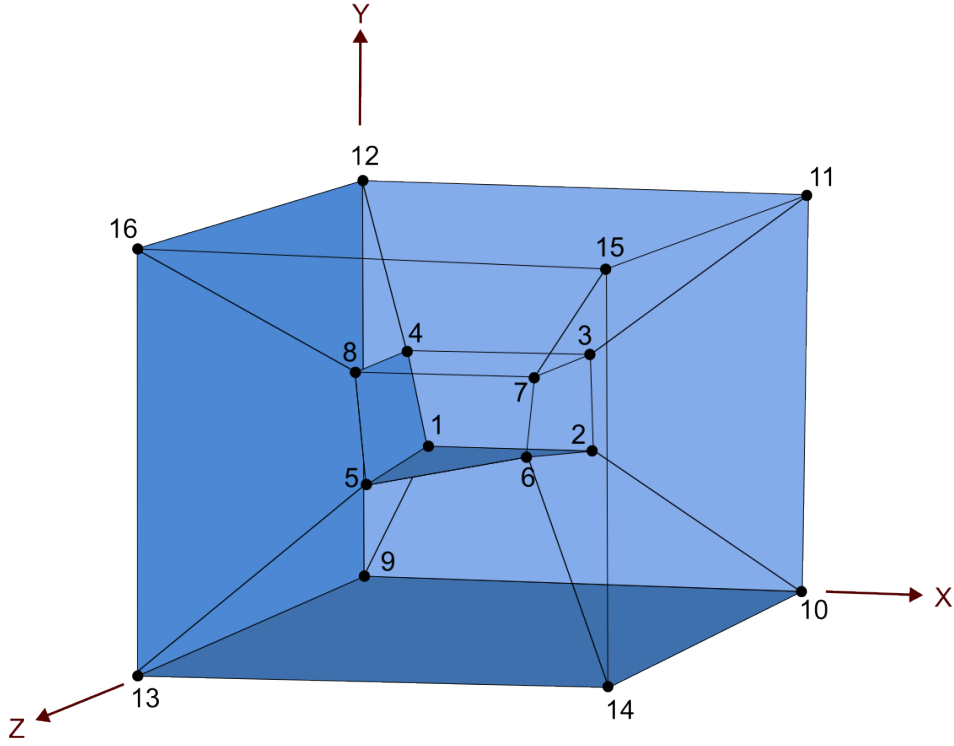
Figure 5.1: Patch test geometry for hex-8 element with displacement BCs in Eq. 5.6. See Table 5.2 for the nodal coordinates of the internal nodes numbered 1-8. Nodes 9 and 15 are at $(X,Y,Z) = (0,0,0)$ and $(1,1,1)$, respectively.

Material Properties: $E = 10^6, \nu = 0.25$

	Nodal Coordinates		
Node	X	Y	Z
1	0.249	0.342	0.192
2	0.826	0.288	0.288
3	0.850	0.649	0.263
4	0.273	0.750	0.230
5	0.320	0.186	0.643
6	0.677	0.305	0.683
7	0.788	0.693	0.644
8	0.165	0.745	0.702

Table 5.2: Nodal coordinates of internal nodes numbered 1-8 for Fig. 5.1.

contrast, the ESFEM allows for movement within the face due to dependence on secondary nodes that may not be fixed. In this particular problem, the internal nodes are not fixed, and thus movement is granted on the face.

A temporary modification was added to remove secondary nodes on faces where all primary nodes have a displacement BC in at least one direction. With the secondary nodes removed from such faces, and therefore from the element, ES-8-S and ES-8-GS then pass the patch test. The effects of this modification on the flexibility of the mesh should be evaluated before being implemented permanently. Note that without the modification, the ESFEM still satisfies the F-E-M-Test [35], guaranteeing its convergence. See Table 5.3 for a summary of the results.

Descriptor	Secondaries on BC Faces?	Pass or Fail
C-8	N/A	Pass
C-27	N/A	Pass
ES-8-N	N/A	Pass
ES-8-S	✓	Fail
ES-8-GS	✓	Fail
ES-8-S	✕	Pass
ES-8-GS	✕	Pass

Table 5.3: Results for the hexahedral patch test.

5.3.2 Tet-4

Here, the C-4 and ES-4-N elements pass as is. In the same manner as the hexahedral elements, the elements ES-4-S and ES-4-GS do not pass the patch test unless the secondary nodes are turned off for faces that contain displacement BCs at all nodes of the face. See Table 5.4 for a summary of the results.

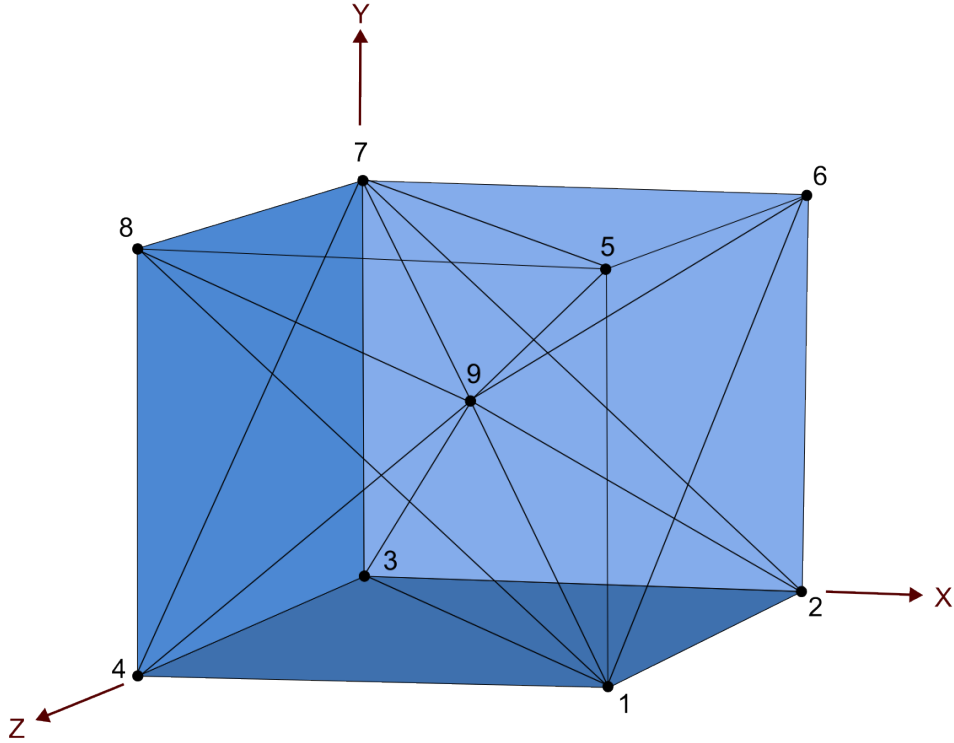| Descriptor	Secondaries on BC Faces?	Pass or Fail
C-4	N/A	Pass
ES-4-N	N/A	Pass
ES-4-S	✓	Fail
ES-4-GS	✓	Fail
ES-4-S	✕	Pass
ES-4-GS	✕	Pass

Table 5.4: Results for the tetrahedral patch test.

Figure 5.2: Patch test geometry for tet-4 element with displacement BCs in line with Eq. 5.6. Internal node 9 is at (X,Y,Z) = (0.374,0.467,0.542) while nodes 3 and 5 are at (X,Y,Z) = (0,0,0) and (1,1,1), respectively.

Material Properties: $E = 10^6, \nu = 0.25$

5.4   Twisted Beam: Effect of Warping

The twisted beam problem has been included to study the effect of warping [24]. The setup of the problem is illustrated in Fig. 5.3. Properties of the beam are as listed in the caption and the loading includes nodal forces in the -Y-direction adding up to a magnitude of 1.0 at the Y = 0 and Z = L location on the free end of the beam. The left end at Z = 0 is fixed, and the beam begins in the warped configuration where the free end is rotated 90° counter-clockwise about the Z-direction.

The different mesh discretizations are given an ID in Table 5.5, with L, M, and N corresponding to the discretizations illustrated in Figs. 5.4 and 5.5. The "element diameter" is the diameter of the smallest sphere that can enclose an element. Thus, the maximum element diameter is the maximum value among all elements in the mesh.
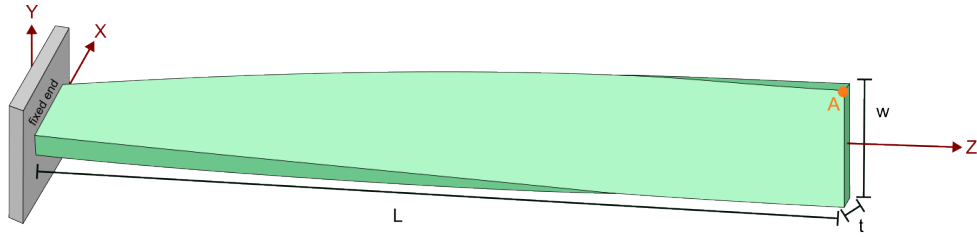
Figure 5.3: Thin cantilever beam with a beginning state twisted configuration, subject to a tip force, applied as nodal forces at the nodes along Y = 0 and Z = L.
Material Properties: $E = 29 \times 10^6, \nu = 0.22$
Geometrical Properties: $L = 12, w = 1.1, t = 0.32$
Theoretical Displacement of Node A: -0.005424 in the Y-direction

Mesh ID	L	M	N	Max Element Diameter	Hex-8 DP
1	6	1	1	2.30001	9.85E-08
2	12	1	2	1.19478	1.51E-05
3	12	2	4	1.06150	8.57E-08
4	24	4	8	5.32473E-01	6.57E-08
5	80	6	10	1.97882E-01	3.98E-08
6	200	6	20	1.00202E-01	2.19E-08

Table 5.5: Mesh discretizations for the twisted beam.
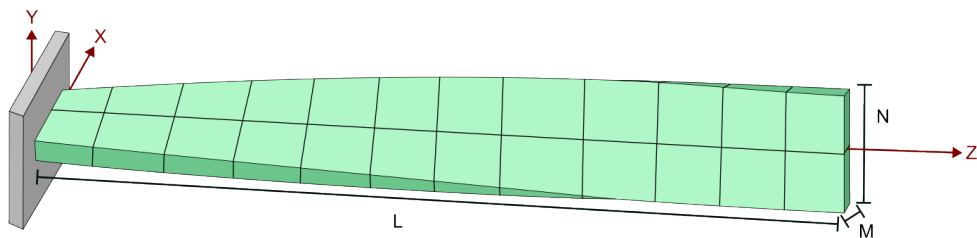
Figure 5.4: Discretization of hexehedral mesh on the twisted beam. In this example, N = 2, M = 1, and L = 12.
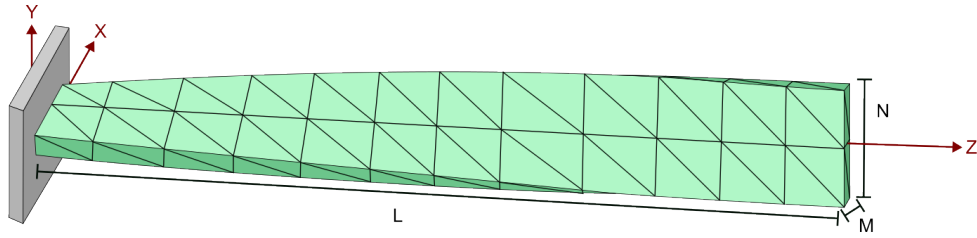
Figure 5.5: Discretization of tetrahedral mesh on the twisted beam. In this example, N = 2, M = 1, and L = 12.

5.4.1 Normalized Displacement Results

The normalized displacement results are illustrated in Fig. 5.6. The ES-8-GS element is converged for mesh ID 1, and remains at most 5% greater than the theoretical displacement. The ES-8-S element needs more refinement than ES-8-GS and converges at mesh ID 3. This need for additional refinement can be linked to the limited nodal data available with the ES-8-S element as compared to the ES-8-GS element. As the N and M mesh discretizations increase from mesh ID 2 to 3, the displacement reaches convergence.

Note that the ES-4-GS element converges faster than C-8, which indicates an improved utility for the tet-4 element when given ghost secondary nodes. The test for the ES-4-S element failed to give results for mesh ID 4, and therefore only meshes 1-3 are shown. As expected, the normalized displacement of ES-4-S approaches the theoretical displacement at a lesser slope than the ES-4-GS element due to the increased nodal data.

Note that even at mesh ID 6, the C-4 element is not converged to the theoretical solution. These results indicate that the ES-4-GS, ES-8-S, and ES-8-GS elements all converge to the theoretical solution at a coarser mesh than the C-8 element. The ES-4-S element does have a greater slope than the C-8 element, but the failure at mesh ID 4 indicates a need for further investigation of this element.

Note the dotted vertical line at mesh ID 2. From mesh ID 1 to 2 the hex-8 DP increases and from 2 to 3, it decreases. In both jumps, the hex-8 DP of mesh ID 2 is 3 orders of magnitude greater than mesh IDs 1 and 3 (see Table 5.5). There is a notable variation in the results at mesh ID 2 that likely relates to this increased DP. However, the ES-8-S and ES-8-GS elements seem to be unaffected by this outlier DP. And, if they are affected, it is minimal. This suggests that the ES-8-S and ES-8-GS elements handle the presence of distortion well in this problem.
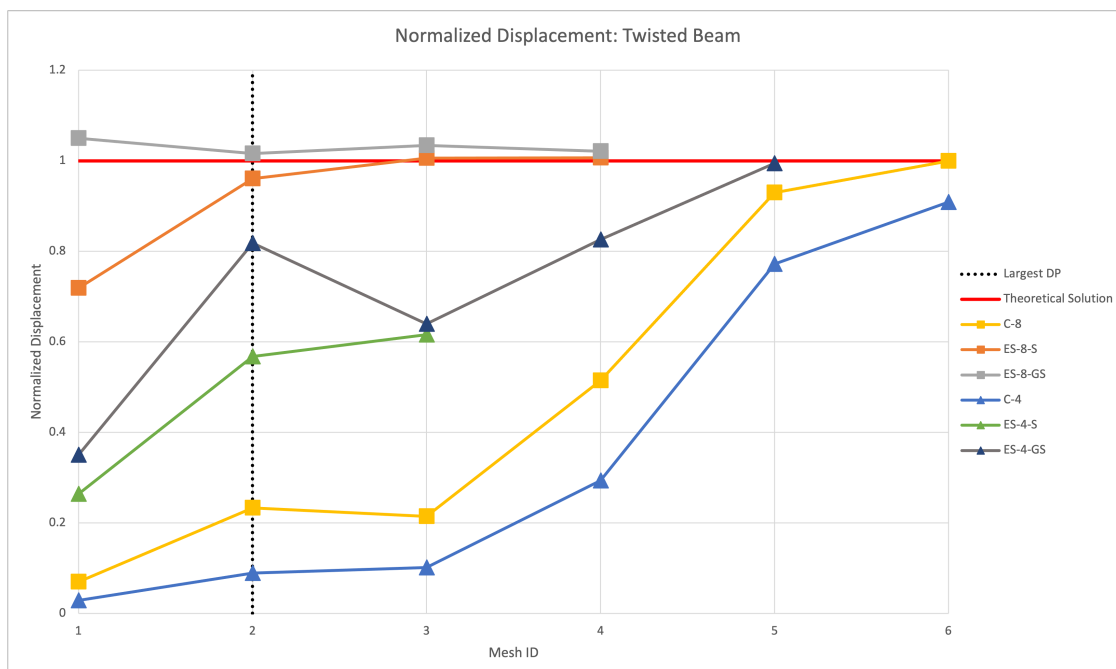
Figure 5.6: Plot of the normalized displacements for the twisted beam problem for different mesh discretizations.

5.4.2 Energy Norm Error

The energy norm error explained in Section 5.1 is used to evaluate the accuracy of the stresses evaluated at the IPs. Fig. 5.7 shows the energy norm error using all hex-8 elements for mesh IDs 1-4. The energy norm error decreases for the C-8 element. For the Es-8-S and ES-8-GS elements, the energy norm error increases from mesh ID 1 to 4. Ideally, the energy norm error would decrease and approach zero at the same time the displacement solution reaches
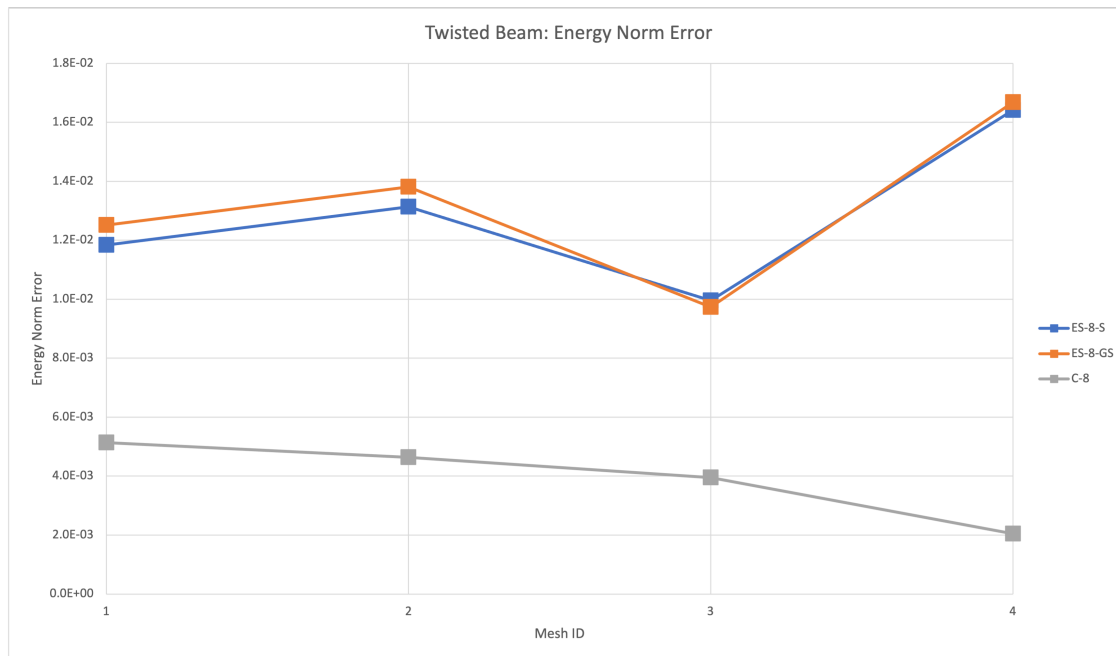
the theoretical value.

Figure 5.7: Plot of the energy norm error for the twisted beam problem for different mesh discretizations.

5.5   Square Cantilever Beam

A square cantilever beam is included for multiple analyses that compare CFEM elements to ESFEM elements. Included is an analysis of the stiffness matrix eigenvalues, a comparison of CPU time, effects of a nearly-incompressible material, and the evaluation of stress convergence. Stress convegence is quantified by using the energy norm error in Section 5.1 for the hex-8 elements. All elements listed in Table 5.1 will be used in each test, except for the eigenvalue analysis and stress convergence. For hex-27 elements, the package LPACK utilized to calculate the eigenvalues gives an error that stops the run of the problem. For the stress convergence, only hex-8 elements are considered.

A couple things to note about Table 5.6. The reference mesh to normalize the CPU time with is mesh ID 7 of the C-8 element. For mesh ID 4 for the ES-4-S element, the number of zero eigenvalues is marked N/A because the function package used to calculate

the eigenvalues gave an error. Additionally, only one mesh discretization is given for ES-4-GS since the more refined meshes would not converge. Looking at mesh ID 1 of the ES-8-S element, it results in a normalized displacement 1.57. If a problem encounters shear locking, it will produce smaller values than the theoretical result. So, while not an accurate result as this discretization, this indicates that shear locking is not present for the ES-8-S element.
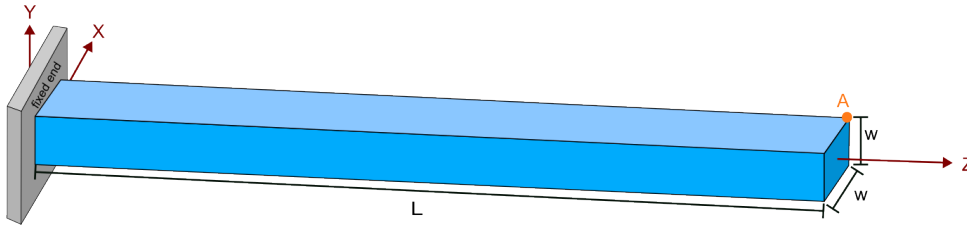

Figure 5.8: Thin cantilever beam with a square cross-section, subject to a traction of 10^3 at Z = L.
Material Properties: $E = 10^9, \nu = 0.3 - 0.4999$
Geometrical Properties: $L = 100, w = 2$
Theoretical Displacement at Node A: 1.0 in the Y-direction

5.5.1   Analysis of Eigenvalues

A stiffness matrix of full rank in the CFEM has 6 zero eigenvalues for the 6 rigid body modes. Thus, the rank is 6 less than the dimension of the stiffness matrix. In the CFEM, a rank-deficient stiffness matrix with more than 6 zero eigenvalues leads to zero-energy modes. The maximum number of zero eigenvalues, and the normalized displacements are displayed in Table 5.6. The number of zero eigenvalues for the ES-8-S, ES-8-GS, ES-4-S, and ES-4-GS element stiffness matrix are all much larger than 6. However, the displacements and convergence of the ES-8-S, ES-8-GS, and ES-4-S element meshes indicate no evidence of rank-deficiency. While the element stiffness matrices are rank deficient, it is generally the

Descriptor	Mesh ID	L	M	N	Number of Zero Eigenvalues	Normalized Displacement at Node A	Normalized CPU Time
C-8	1	10	1	1	6	0.09946	0.00038
	2	10	2	1	6	0.09980	0.00076
	3	20	4	2	6	0.30109	0.00939
	4	100	4	4	6	0.92700	0.10460
	6	200	4	4	6	0.99340	0.20598
	7	300	6	6	6	0.99682	1.00000
ES-8-S	1	10	1	1	18	1.57321	0.00643
	2	10	2	1	21	0.83895	0.00855
	3	20	4	2	39	0.97955	0.09451
	4	100	4	4	36	1.00011	1.37081
ES-8-GS	1	10	1	1	18	0.99012	0.00635
	2	10	2	1	21	1.01399	0.01385
	3	20	4	2	39	0.99982	0.13595
	4	100	4	4	36	1.00206	1.97374
C-27	1	10	1	1	6	0.98406	0.00641
	2	10	2	1	6	0.98406	0.01302
	3	20	4	2	6	0.98322	0.13263
	4	100	4	4	6	0.98281	1.85233
	5	200	4	4	6	0.98284	3.68289
C-4	1	10	1	1	6	0.02036	0.00048
	2	10	2	1	6	0.00884	0.00094
	3	20	4	2	6	0.03014	0.00767
	4	100	4	4	6	0.56824	0.19253
	5	200	4	4	6	0.75434	0.53284
	6	300	6	6	6	0.86745	4.14568
ES-4-S	1	10	1	1	30	0.15642	0.01619
	2	10	2	1	49	0.07817	0.02981
	3	20	4	2	48	0.22136	0.28563
	4	100	4	4	N/A	0.96481	5.44524
ES-4-GS	1	10	1	1	33	0.51247	0.01724

Table 5.6: Results for a thin, square cantilever beam.

case that the global stiffness matrix is not rank deficient for the ESFEM.

5.5.2 Comparison of CPU Times

CPU time has been calculated for each run of the problem by adding together the CPU times of the problem set up and the solution computations. It is notable that each problem

run will result in different CPU times for the same mesh. However, the variation is only a maximum of 10% for the same mesh. In contrast, the variation between different mesh discretizations is at least 75%. Thus, only one run is used as a representation of the CPU time. Since the code based used has not been optimized for speed, the CPU times have been normalized with the converged C-8 CPU time. See Table 5.6 for the results.

For the same mesh, the normalized CPU times of the ES-8-S and ES-8-GS elements are 10-18 times greater than the C-8 element. However, the ES-8-S and ES-8-GS CPU times at convergence are 0.46 and 0.031 of the C-8 time at convergence, respectively. This indicates that while the ESFEM uses more CPU time for the same mesh than the CFEM, the method may provide a lower CPU time in practice due to a decreased need for mesh refinement.

Comparing ES-8-GS and C-27 results, the discretization of convergence and the CPU time are close. The normalized CPU times of ES-8-GS elements are less than 7% different from the C-27 normalized CPU times of the same mesh. Moreover, both are converged for the most coarse mesh (ID 1). In fact, the ES-8-GS result for this mesh is only 0.6% different from the C-27 result.

Results indicate that use of a C-4 element would be impractical since the CPU time of the most refine mesh is 4.1 times that of the same C-8 mesh. Although, for the same tet-4 mesh, the ES-4-S element approaches convergence, with a normalized displacement of 0.96481. The CPU time of this mesh is 5.4 times higher than the most refined C-8 mesh. While this is a much larger CPU time demand, it must be considered in conjunction with other factors. One being that for complex geometries, an auto-generated tet-4 mesh could save an analyst much time as compared to manually generating a "high quality" hex-8 mesh.

5.5.3   Nearly-Incompressible Material Effects

To evaluate the ESFEM's handling of a nearly-incompressible material for this problem, the Poisson's ratio is varied from 0.3 to 0.4999. This is all done with distorted meshes. The distortion is applied as the addition of a random parameter to the Z-coordinate of nodal

coordinates. Note that the nodal coordinates at either end of the beam are unaffected by the distortion. The distortion parameters (DPs) are given for mesh IDs 1-4 in Table 5.7.

Mesh ID	Hex-8 DP	Hex-27 DP
1	6.27E-03	8.80E-03
2	3.47E-03	5.98E-03
3	2.53E-03	4.31E-03
4	4.37E-03	7.02E-03

Table 5.7: Distortion parameters (DPs) for the hex-8 and hex-27 mesh IDs 1-4.

Plots of the normalized displacement for mesh IDs 1-4 are displayed in Figs. 5.9- 5.12. The C-8 graph shows a strong presence of shear locking, but the results seem to be unaffected by the change in Poisson's ratio. This is concluded since all of the Poisson's ratio lines overlap closely.

The C-27 graph only has mesh IDs 1-3 included for Poisson's ratio values greater than 0.3 because the solution would not converge. As shown in Section 1.1, the C-27 element would not result in any forms of locking without distortion. Now, with distortion, there are clear indicators of shear and volumetric locking. Even the most refined mesh ID 4 does not converge for the distorted C-27 elements.

The ES-8-S graph depicts the presence of shear and volumetric locking that is only avoided as the mesh is refined. As mentioned previously, the coarse meshes with only 1 element through the width and thickness do not provide the ES-8-S element with enough nodal data to produce accurate results. However, the element converges for all Poisson's ratio values at mesh ID 4. This indicates that the mesh distortion does not affect the ES-8-S's capability to produce accurate results for a nearly-incompressible material.

In the ES-8-GS graph, the solution decreases in accuracy at mesh ID 3. Interestingly, this is the mesh with the lowest hex-8 DP. This suggests that the presence of distortion could benefit the accuracy of the ES-8-GS elements. A study of the resulting normalized displacements by varying the DP would help determine if this was the case.
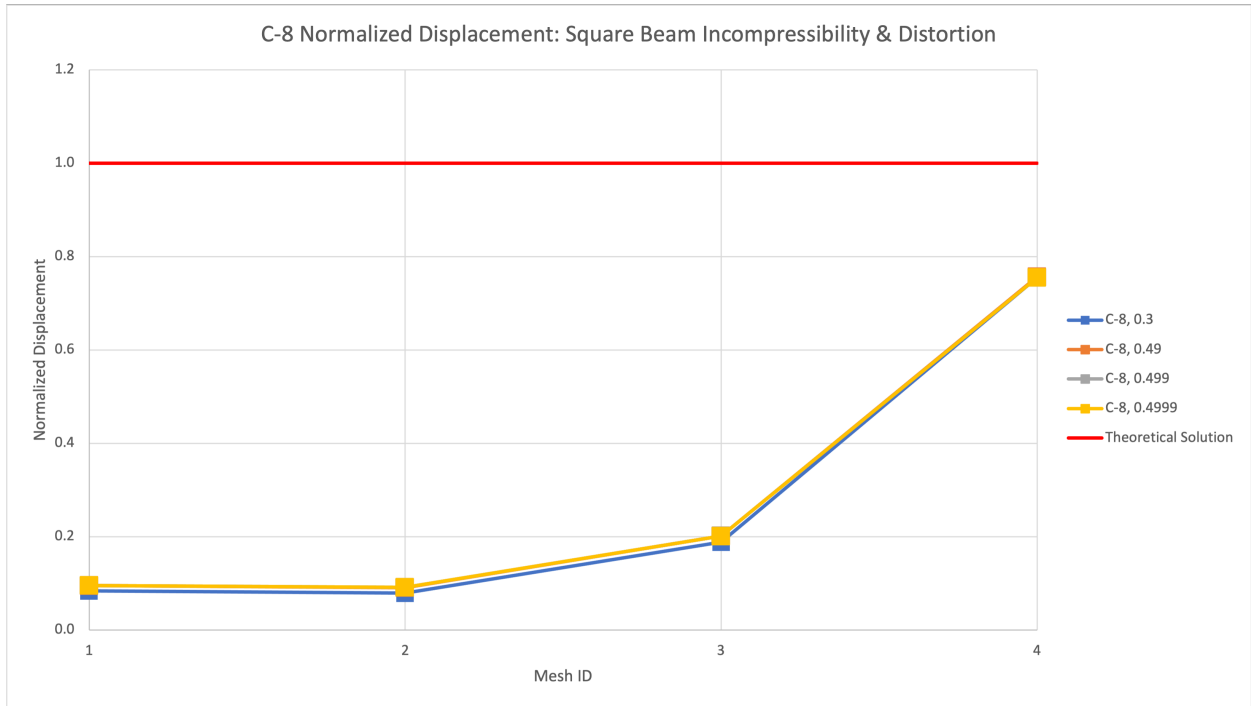
Figure 5.9: Comparison of normalized displacements of the square beam problem for varying Poisson's ratios in distorted C-8 elements.
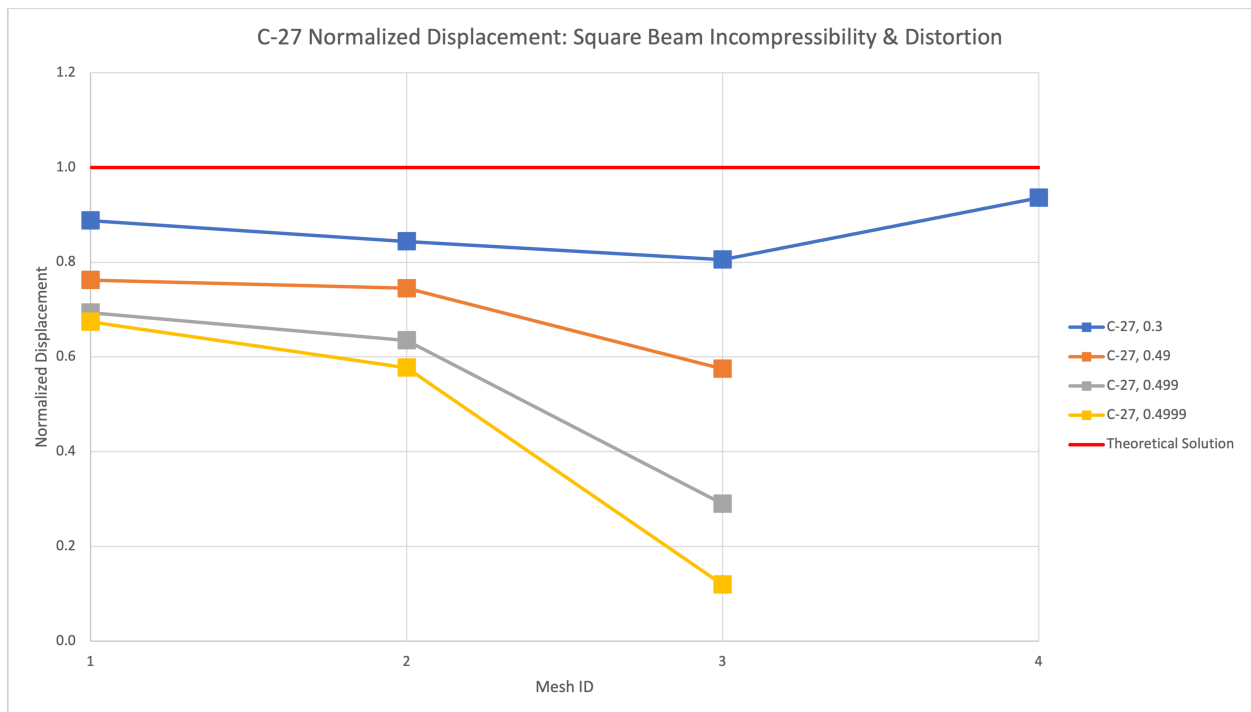

Figure 5.10: Comparison of normalized displacements of the square beam problem for varying Poisson's ratios in distorted C-27 elements.
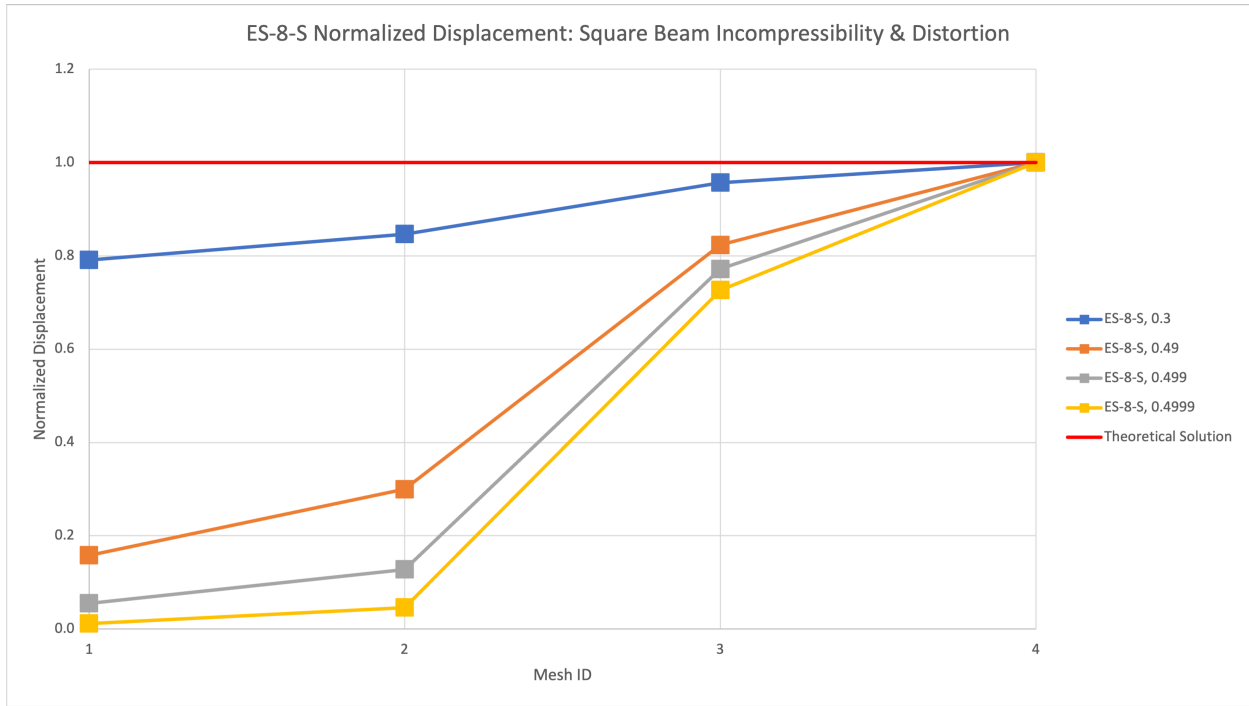
Figure 5.11: Comparison of normalized displacements of the square beam problem for varying Poisson's ratios in distorted ES-8-S elements.
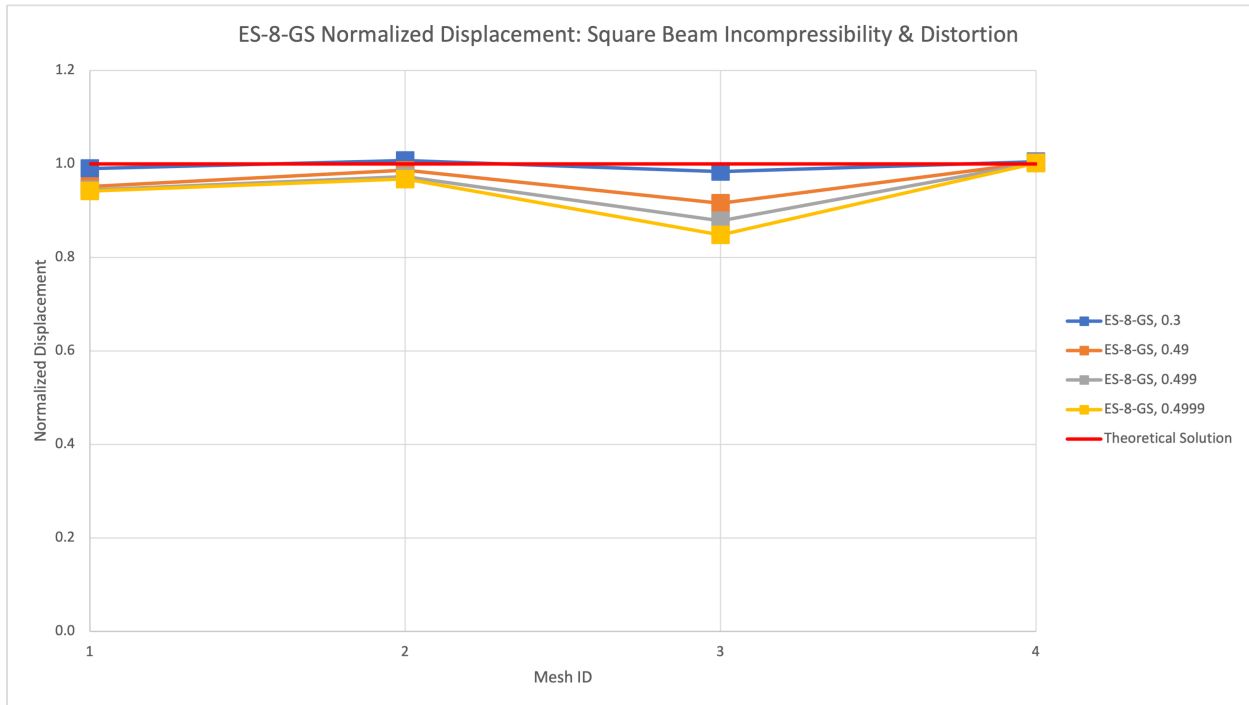

Figure 5.12: Comparison of normalized displacements of the square beam problem for varying Poisson's ratios in distorted ES-8-GS elements.

5.5.4 Energy Norm Error: Stress Convergence

The energy norm error described in Section 5.1 is used here to evaluate the convergence of the stress solution at the integration points (see Fig. 5.13). As expected, the C-8 solution approaches zero as the mesh is refined, where mesh ID 7 (see Table 5.6) is the reference mesh. For both the ES-8-S and ES-8-GS elements, the energy norm error remains about the same as the mesh is refined. Due to the increased CPU times of the ES-8-S and ES-8-GS elements, the most refined mesh tested was mesh ID 4. Since the polynomial fitting is performed with an underdetermined system, large oscillations in the solution are possible. Without nodal data inside of the element to inform the polynomial fit, it is likely that while the nodal displacements are accurate, the stress solution inside the element remains inaccurate.

A potential remedy for these poor stress convergence results is the addition of a node in the center of the element. As displayed by the use of ghost nodes, the ESFEM allows the addition of nodes anywhere in the mesh. With a node at the centroid of the element, the basis functions may be more likely to avoid oscillations that lead to inaccurate stress results.
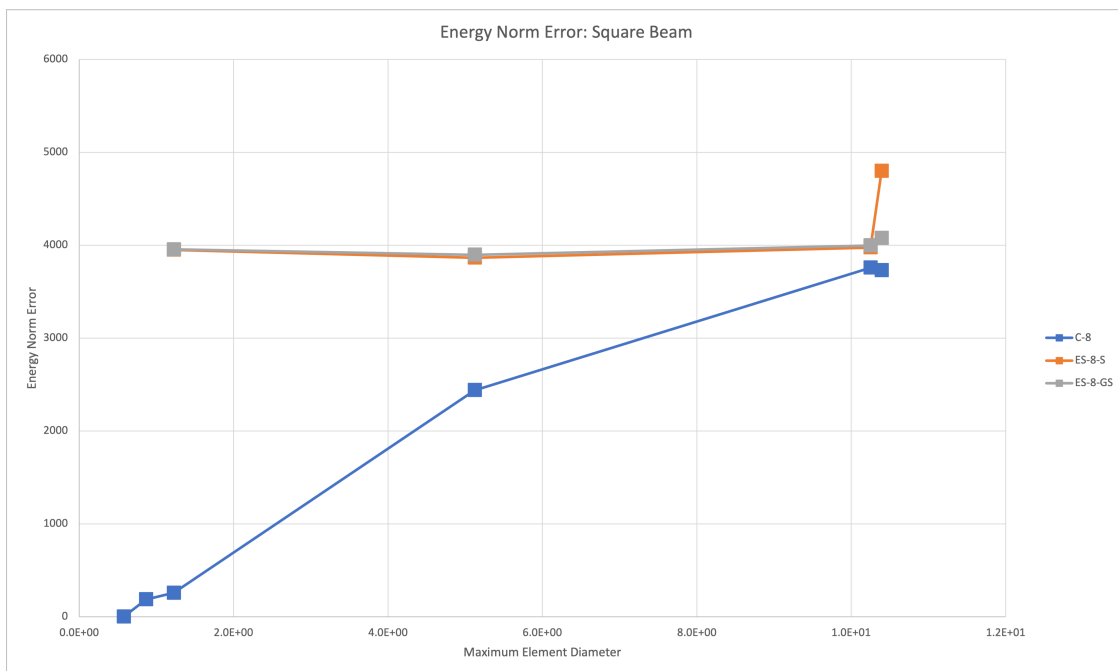

Figure 5.13: The convergence of stress for the hex-8 elements.

5.6   Curved Beam: Effect of Slight Irregularities

The curved beam problem has been chosen to test the effect of slight irregularities in a mesh
[24]. The geometry of the problem is illustrated in Fig. 5.14. The material and geometrical
properties are listed in the caption of Fig. 5.14. The loading in this problem is a traction
with magnitude 50 applied to the faces at X = 0 in the -Y-direction. The hex and tet meshes
for the curved beam configuration are illustrated in Figs. 5.15 and 5.16.

Mesh ID	L	M	N	Max Element Diameter	Hex-8 DP
1	6	1	1	1.13	5.59E-03
2	12	2	2	5.70E-01	9.88E-04
3	24	4	4	2.87E-01	1.75E-04
4	48	6	6	1.46E-01	5.49E-05
5	100	8	6	7.40E-02	2.47E-05
6	200	8	6	4.53E-02	1.75E-05

Table 5.8: Mesh discretizations for the curved beam.

5.6.1   Normalized Displacement Results

The normalized displacement results are illustrated in Fig. 5.17. Both the ES-8-S and ES-8-
GS elements converge to the theoretical displacement at a coarser mesh than C-8. At mesh
ID 2, ES-8-GS is already converged with a normalized displacement of 1.0301, while ES-8-S
converges at mesh ID 3 with a normalized displacement of 1.0263.

   The C-4 mesh for this problem is not converged even for the most refined mesh. Addi-
tionally, all other elements approach the theoretical displacement with a larger slope. The
ES-4-S and ES-4-GS problems were only refined to mesh ID 3 due to the impractical time
required to run the problems. While they do not show improvement as compared to C-8
elements, they do show promise to potentially converge at a more coarse mesh than the C-4
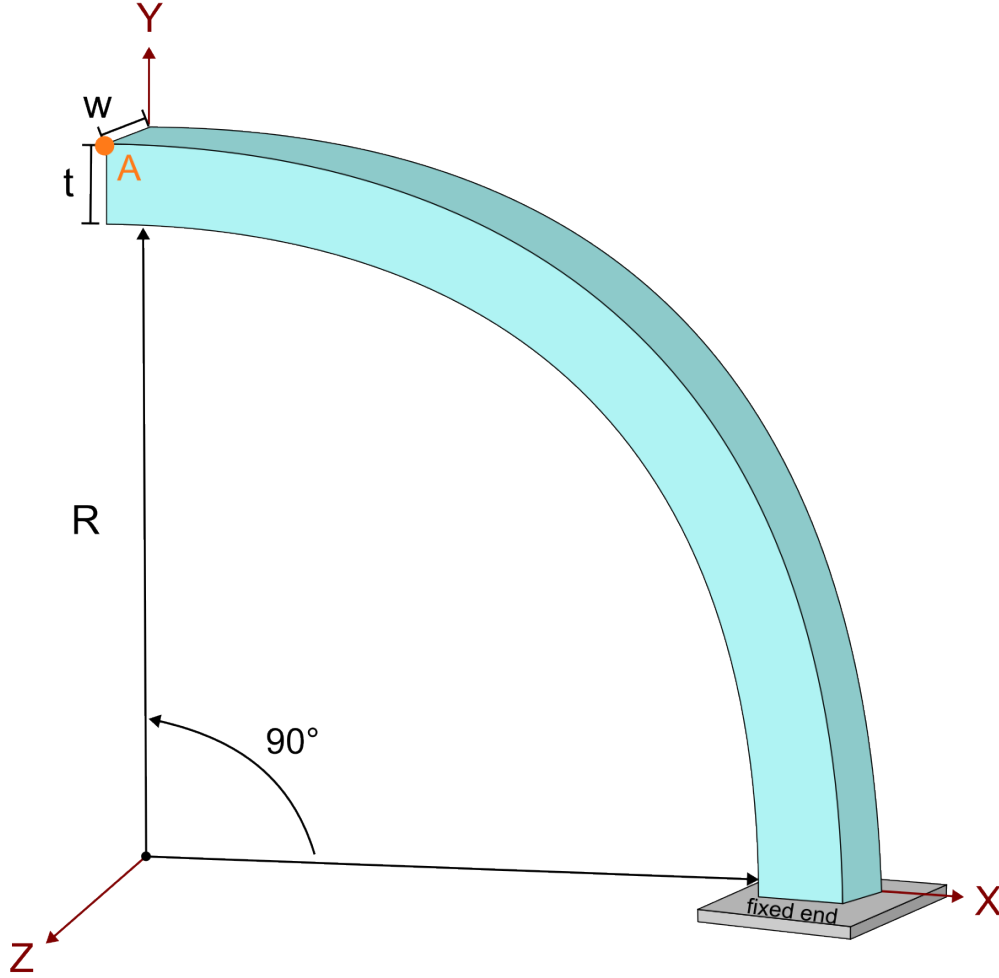elements.

Figure 5.14: Thin beam with a beginning state curved configuration subject to a tip force of magnitude 1 applied as a traction magnitude of 50 in the -Y-direction along the faces at X = 0.

Material Properties: $E = 1 \times 10^7, \nu = 0.25$

Geometrical Properties: $R = 4.12, w = 0.1, t = 0.2$

Theoretical Displacement at Node A: -0.08734 in the Y-direction

5.7 Scordelis-Lo Roof: Singly-Curved Thin Sheet

The Scordelis-Lo Roof problem is an example of a singly-curved thin sheet that is subject to uniform traction on the top surface in the -Z-direction with magnitude 90 [24]. The inner radius of the geometry is 25, while the outer radius is 25.25, as illustrated in Fig. 5.18.
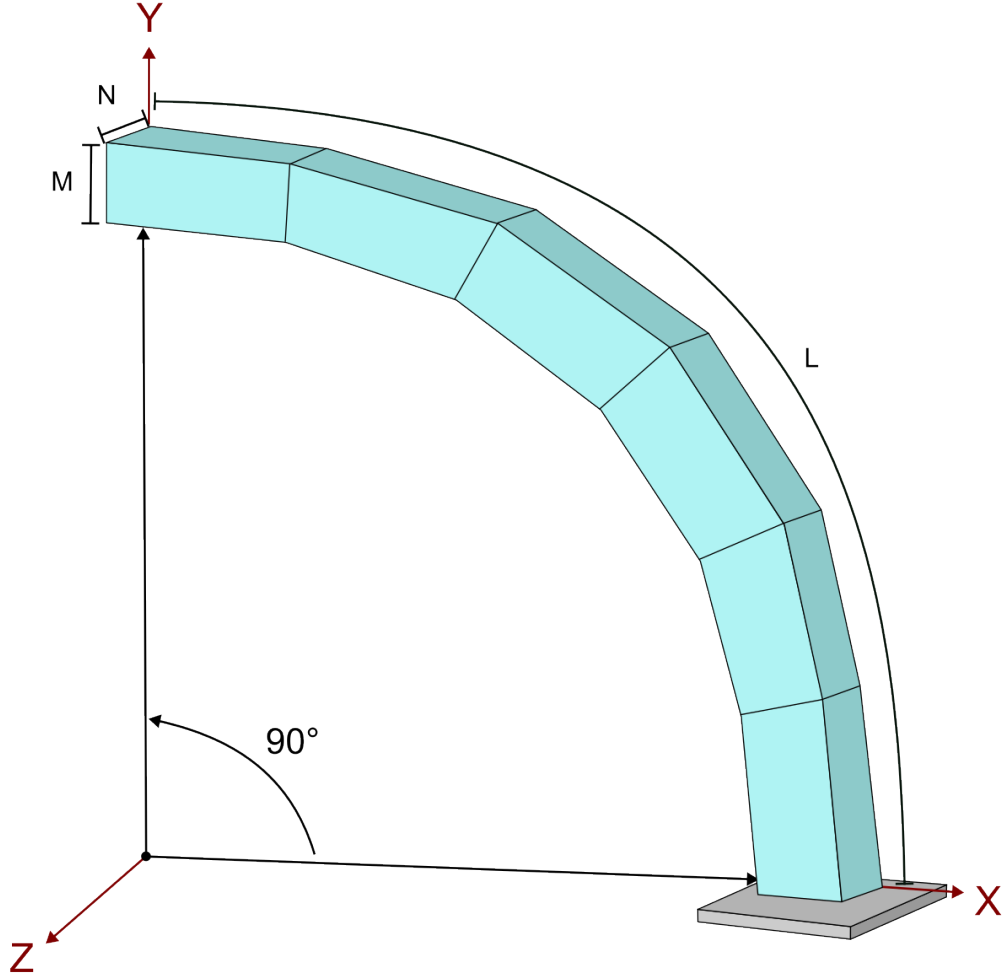
Figure 5.15: Discretization of hexehedral mesh on the twisted beam. In this example, N = 1, M = 1, and L = 6.

Symmetry allows one quarter of the geometry to be meshed, as indicated by the dotted lines.

Mesh ID	N	M	Maximum Element Diameter	Hex Aspect Ratio	Hex-8 DP
1	4	1	7.64	25.0	7.18E-04
2	8	2	3.82	25.0	1.27E-04
3	16	4	1.91	25.0	2.40E-05
4	24	1	1.30	4.17	1.20E-04
5	24	4	1.28	16.7	9.77E-05
6	36	1	1.18	4.17	2.40E-05

Table 5.9: Mesh discretizations for the Scordelis-Lo roof problem.
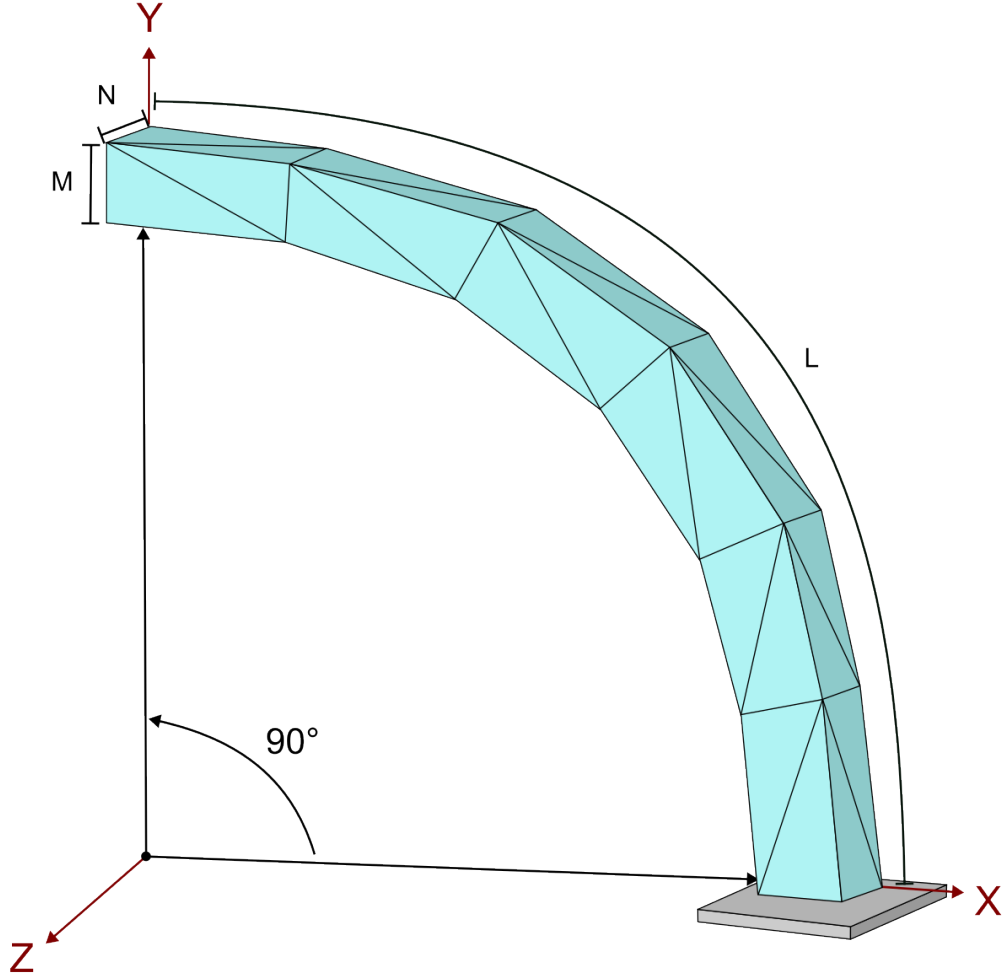
Figure 5.16: Discretization of tetrahedral mesh on the twisted beam. In this example, N = 1, M = 1, and L = 6.

The meshes used in this problem vary in number of elements through the thickness (M), and in-plane of the roof (N). See Table 5.9 for the varying meshes and their maximum element diameter. The normalized displacement results is given in Fig. 5.19. None of the elements converge to the theoretical solution in the meshes used.

Since the aspect ratios exceed the commonly used maximum of 3, it would be valuable to experiment with aspect ratios closer to 1. Another potential cause of the inaccurate convergence of the ES-8-S and ES-8-GS elements is the presence of large oscillations in the
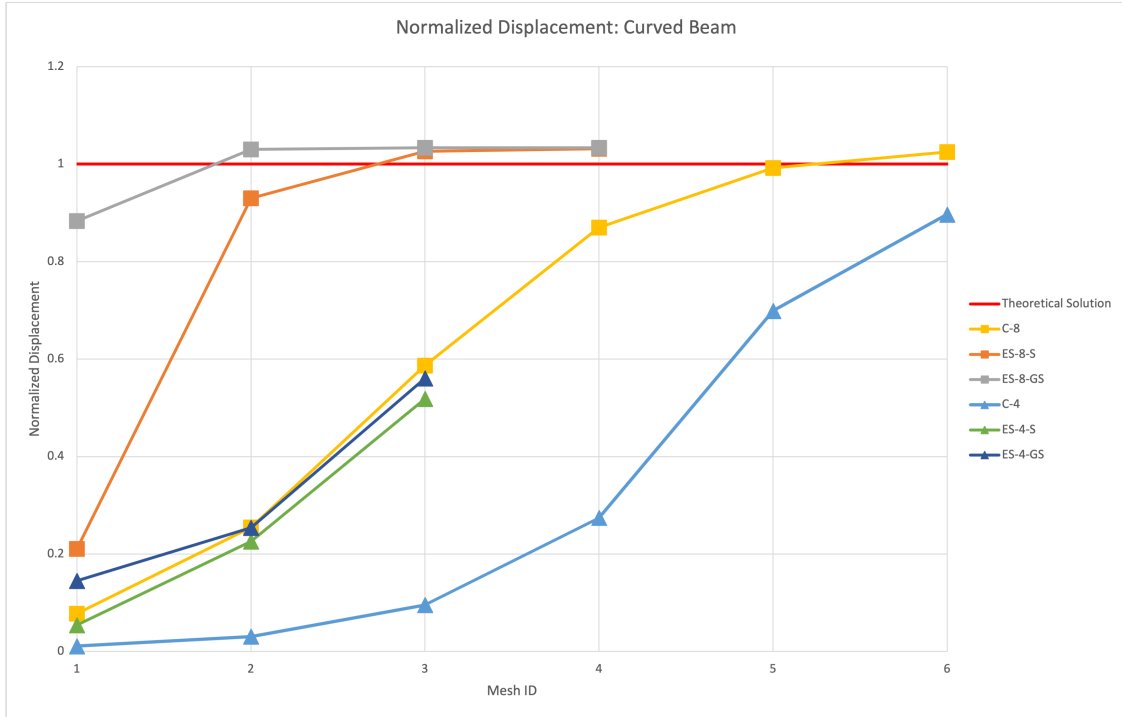
Figure 5.17: Plot of the normalized displacements for the curved beam problem for different mesh discretizations.

basis functions. As discussed in Section 5.5.4, adding a node to the centroid of the extended-stencil elements may reduce the gyrations and therefore produce more accurate results.

The dotted vertical line indicates the maximum element diameter of mesh ID 4. The DP of mesh ID 4 is an order of magnitude greater than mesh IDs 5 and 6. While the element diameters for mesh IDs 4 and 5 are close, it was verified that the large jump in solutions occurs at mesh ID 4. In other problems, a change in DP does not affect the ES-8-S problem much, but the normalized displacement for mesh ID 4 is 28% less than the result of mesh ID 3 and 39% less than the result of mesh ID 5. Another potential cause of this jump is the use of only one element through the thickness. This is considered since mesh ID 6 also causes a jump in solutions for the ES-8-S element. To evaluate the root cause, it would be valuable to test the ES-8-S element with varying DP values and number of elements through the thickness.
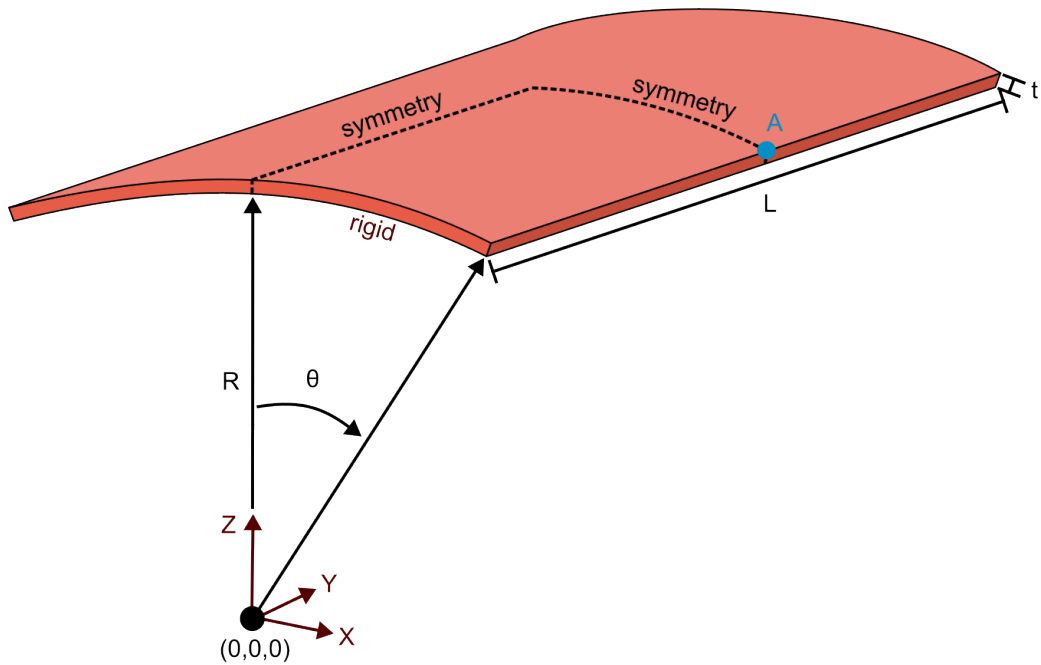
Figure 5.18: A singly-curved surface that is subject to a traction of 90 units force per area in the -Z-direction. Nodes along X = 0 and Y = L/2 are given to symmetric BCs. The curved edges are rigid such that the nodes at Y = 0 are fixed in the X and Z-directions.
Material Properties: $E = 4.32 \times 10^8, \nu = 0.0$
Geometrical Properties: $R = 25, t = 0.25, \theta = 40°, L = 50$
Theoretical Displacement of Node A: -0.3024 in the Z-direction

5.8 Lateral Rectangle: Shear Locking & Distortion Effects

The lateral rectangle problem has been included to evaluate performance in the presence of shear locking when considering varying mesh distortion and aspect ratios. The geometry of this problem involves a square plate, that allows only one quarter of the square to be meshed due to symmetry (see Fig. 5.20). A uniform traction of 10^{-4} (property set 1) or 10^3 (property set 2) is applied across the front surface in the -X-direction, and has simply supported displacement BCs.
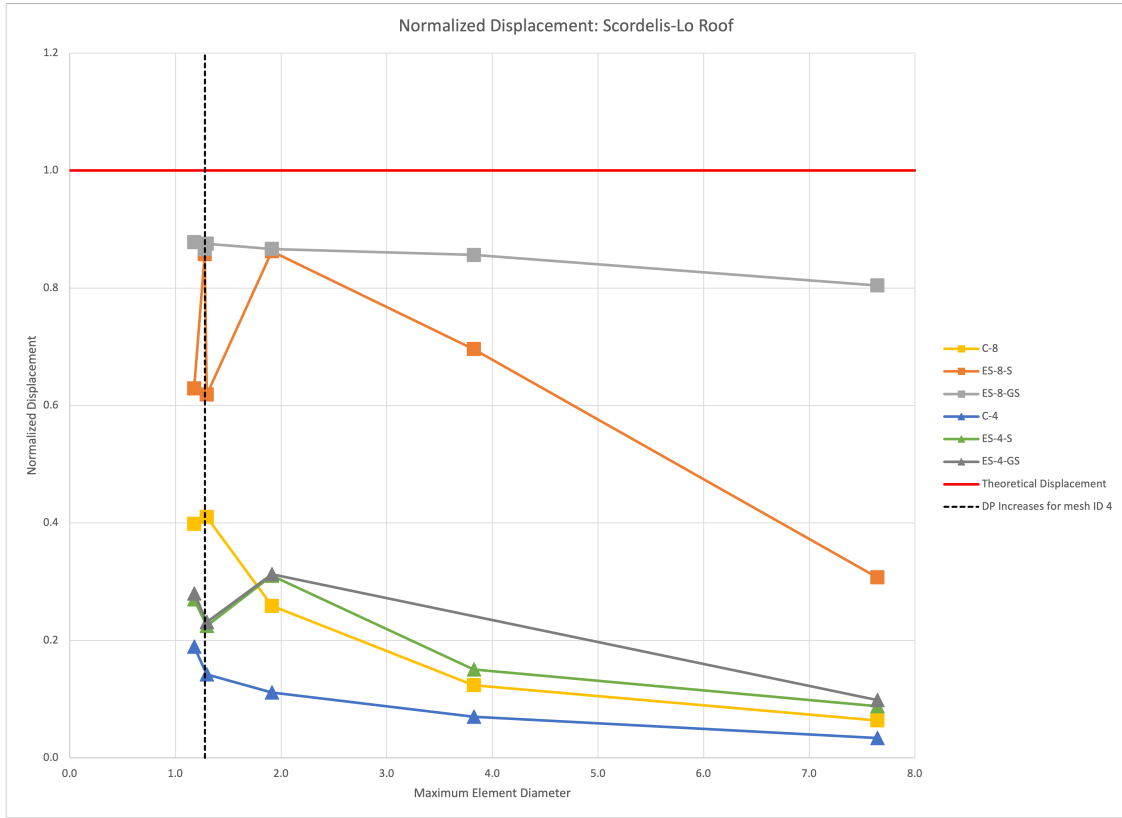
Figure 5.19: A plot of the Z-direction normalized displacement of the Scordelis-Lo roof at node A.

5.8.1 Distortion Effects

All hexahedral elements are used to compare the effects of distortion in the mesh and the results are displayed in Figs. 5.21 - 5.24. To indicate the number of elements through the thickness, the meshes are tagged with a "_T#" in the legend where # is the number of element layers. The meshes distorted in-plane allow for random nodal perturbations in the Y and Z direction and are tagged with "_YP" in the legend. The meshes distorted through the thickness allow for random nodal perturbations in the X direction and are tagged with "_YT" in the legend. Meshes with no distortion are tagged with "_N" in the legend.

As seen in Section 1.1, the C-27 element (see Fig. 5.22) no longer provides accurate results for coarse meshes if distortion is introduced. Moreover, the C-27 results indicate a dependence on the distortion being in-plane vs through the thickness. This is determined by noticing the grouping of the curves that are tagged with "_YP" vs "_YT." Distortion
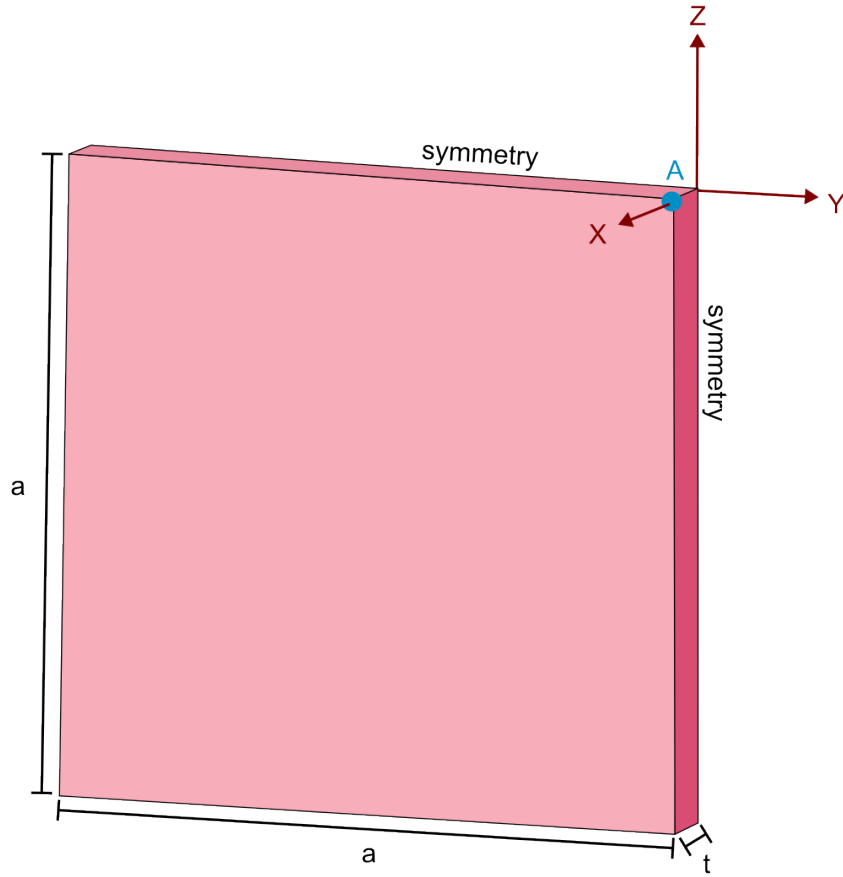
Figure 5.20: A rectangular plate with simply supported (SS) BCs, that due to symmetry allows 1/4 of the plate to be meshed. The bottom and left edges at Y = -1 and Z = -1 where X = 0 are fixed in the X-direction for the SS BCs while the right and top edges are given symmetric BCs across the surfaces of Y = 0 and Z = 0. There is a uniform traction of 10^{-4} (set 1) or 10^3 (set 2) across the entire front surface of the rectangle in the -X-direction.
Properties Set 1: $a = 1.0, t = 0.01, E = 1.7472 \times 10^7, \nu = 0.3$
Properties Set 2: $a = 1.0, t = 0.1, E = 10^9, \nu = 0.3$
Theoretical Displacement of Node A: -4.06E-06 (Set 1) or -7.10E-04 (Set 2) in the X-direction

through the thickness results in more accurate solutions than distortion in-plane.

The C-8, ES-8-S, and ES-8-GS results do not indicate any strong effects on the normalized displacement for one form of distortion versus the other. Looking specifically at the ES-8-GS_T1_N results (see Fig. 5.24, the displacements are highly over-estimated as the aspect ratio decreases. As distortion is added, the accuracy improves. This was also seen in the square beam problem of Fig. 5.12. It would be valuable to test if the ES-8-GS problem has

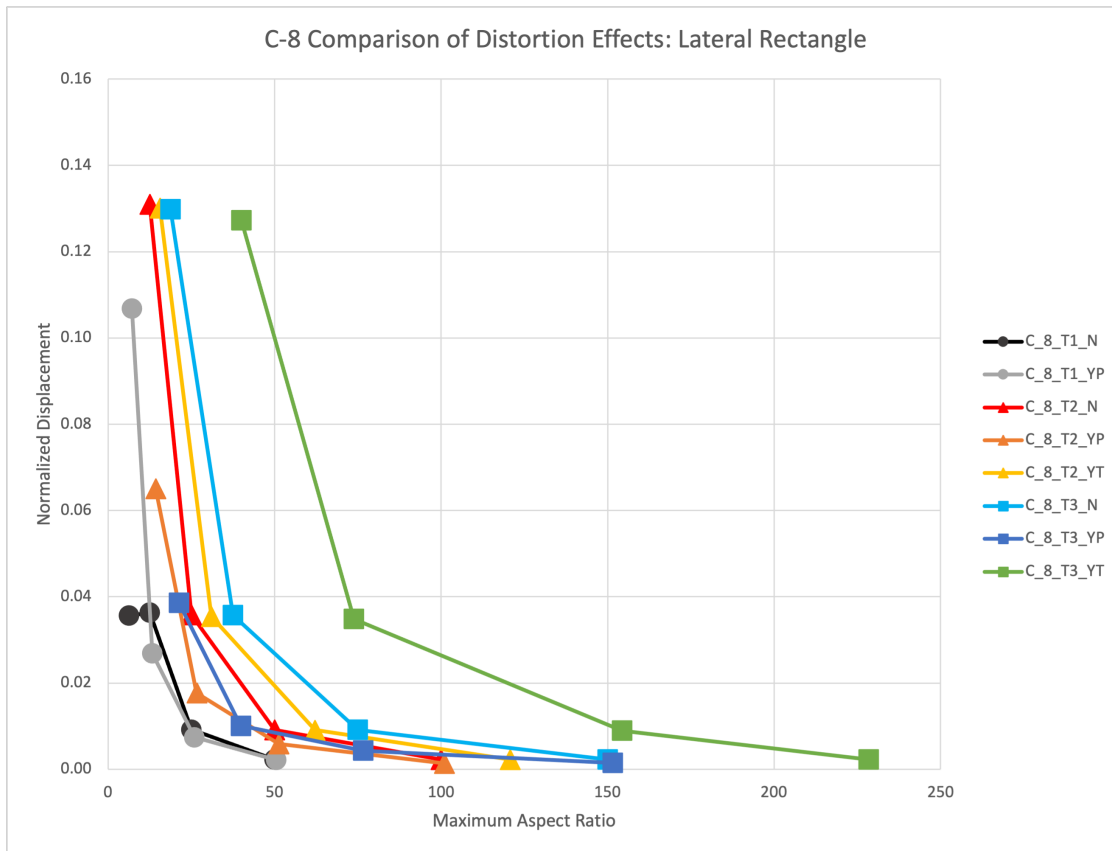an increased accuracy for distorted meshes, as compared to undistorted meshes.

Figure 5.21: Comparison of normalized displacements of the lateral rectangle problem for varying aspect ratios in C-8 elements. Note the different vertical axis scales for the element as compared to Figs. 5.22 - 5.24.

5.8.2 Aspect Ratio Effects

The lateral rectangle problem with property set 2 is used here to evaluate the effects of the aspect ratio on the results. The normalized results are plotted against the aspect ratio of undistorted hex-8 meshes ranging from 1.667 to 15. All meshes have six elements through the thickness and vary the Y-Z plane discretization to vary the aspect ratio. Fig. 5.25 displays the results for the C-8, ES-8-S, and ES-8-GS elements. Only the first five meshes were used for the ES-8-S and ES-8-GS elements since the CPU times were 10 times that of the C-8 mesh and therefore not practical.
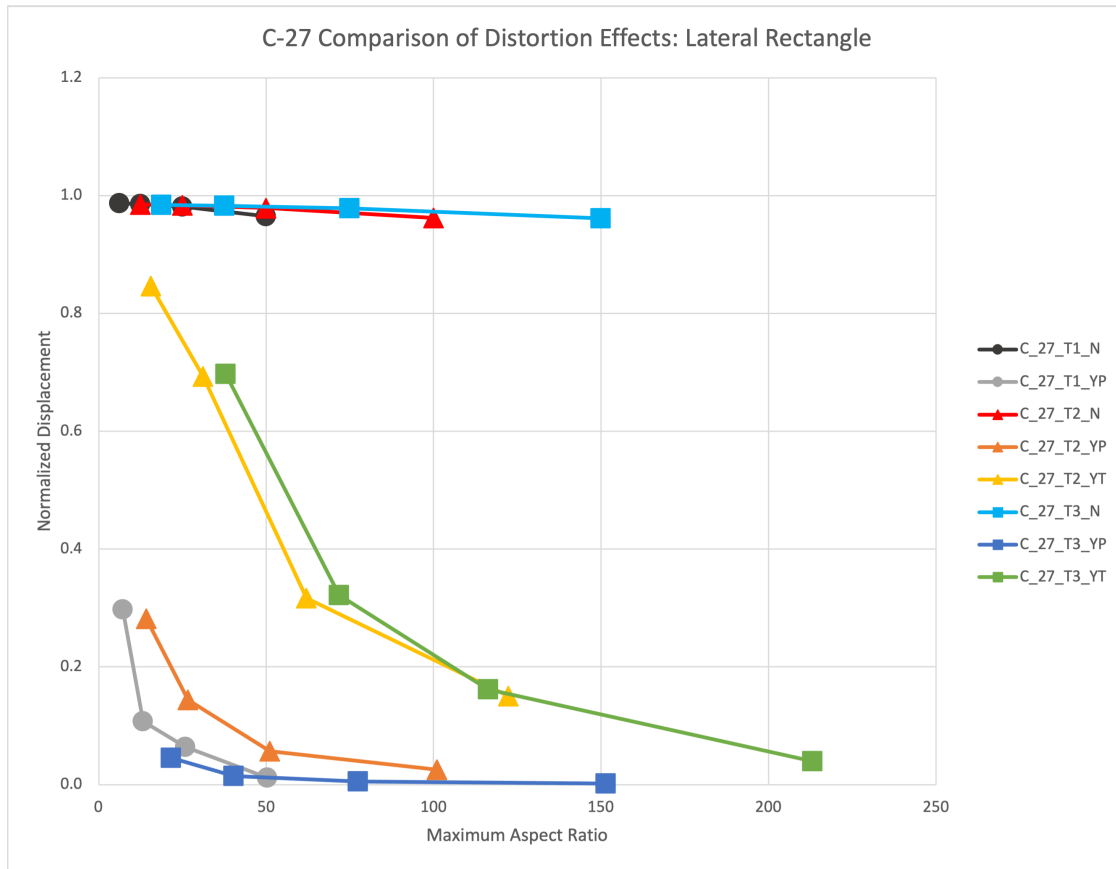
Figure 5.22: Comparison of normalized displacements of the lateral rectangle problem for varying aspect ratios in C-27 elements.

The value converged to by the C-8 element is slightly higher than the theoretical displacement. A horizontal line has been added to indicate this value. This value the C-8 element converges to is the same normalized displacement that the ES-8-S and ES-8-GS elements converge to.

The C-8 element results show agreement with convention that an aspect ratio of 3 or less is preferred. The normalized displacement is at or above 1 for all aspect ratios less than 3. For the ES-8-S and ES-8-GS elements with an aspect ratio of 15, the normalized displacements are only 3.6% and 2.6% different from the converged C-8 solution, respectively. This indicates the potential for the ESFEM elements to produce accurate results with aspect ratios much larger than 3.
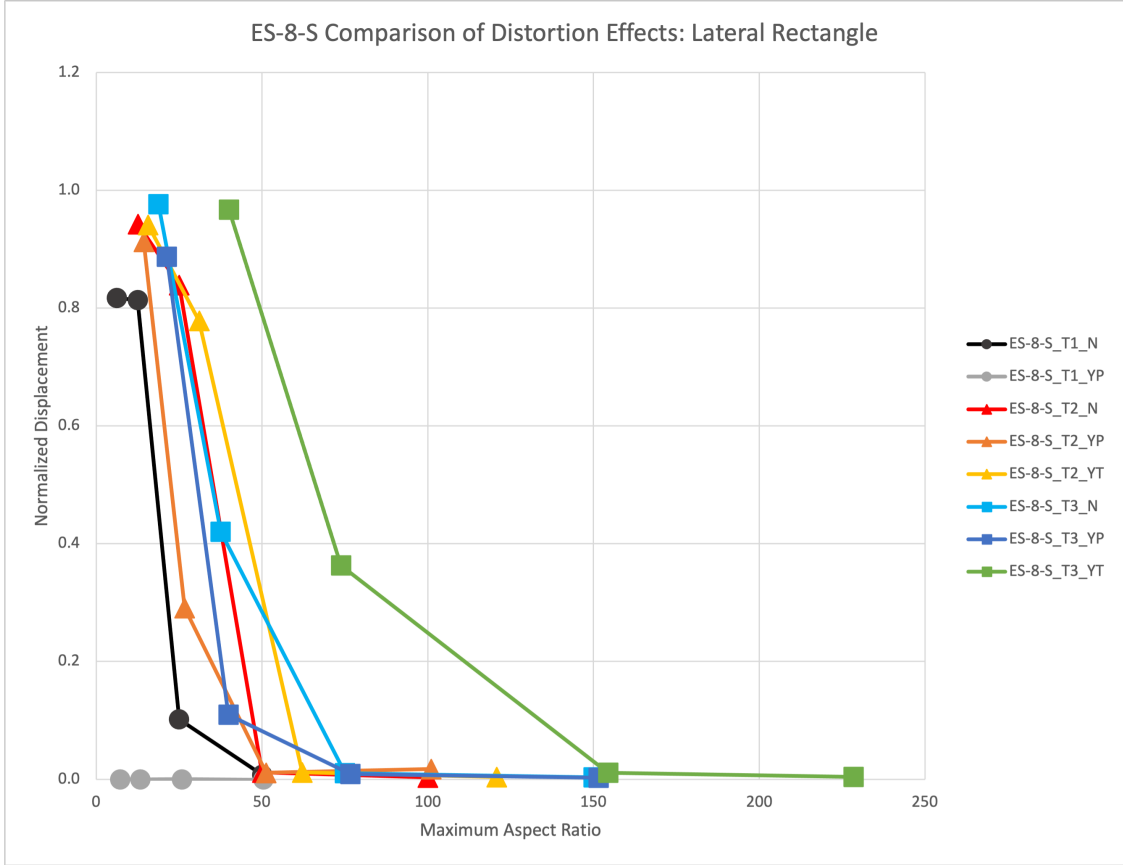
Figure 5.23: Comparison of normalized displacements of the lateral rectangle problem for varying aspect ratios in ES-8-S elements.

5.9   Thick Cylinder: Volumetric Locking Effects

The thick cylinder problem is used to evaluate the different elements as the material approaches near-incompressibility. This is done by applying an outward pressure of magnitude 1 on the inner surface of the cylinder shown in Fig. 5.26. Due to symmetry, one one quarter of the cylinder was meshed. This problem was inspired by the thick cylinder problem in [24] that meshes only $\frac{1}{36}$ of the cylinder, but the theoretical results remain the same.

The equation to calculate the theoretical displacement for each value of the Poisson's ratio is given in [24] to be

$$u_{theory} = \frac{(1+\nu)pr^2}{E\left(R^2 - r^2\right)}\left[\frac{R^2}{x} + (1-2\nu)x\right], \tag{5.8}$$
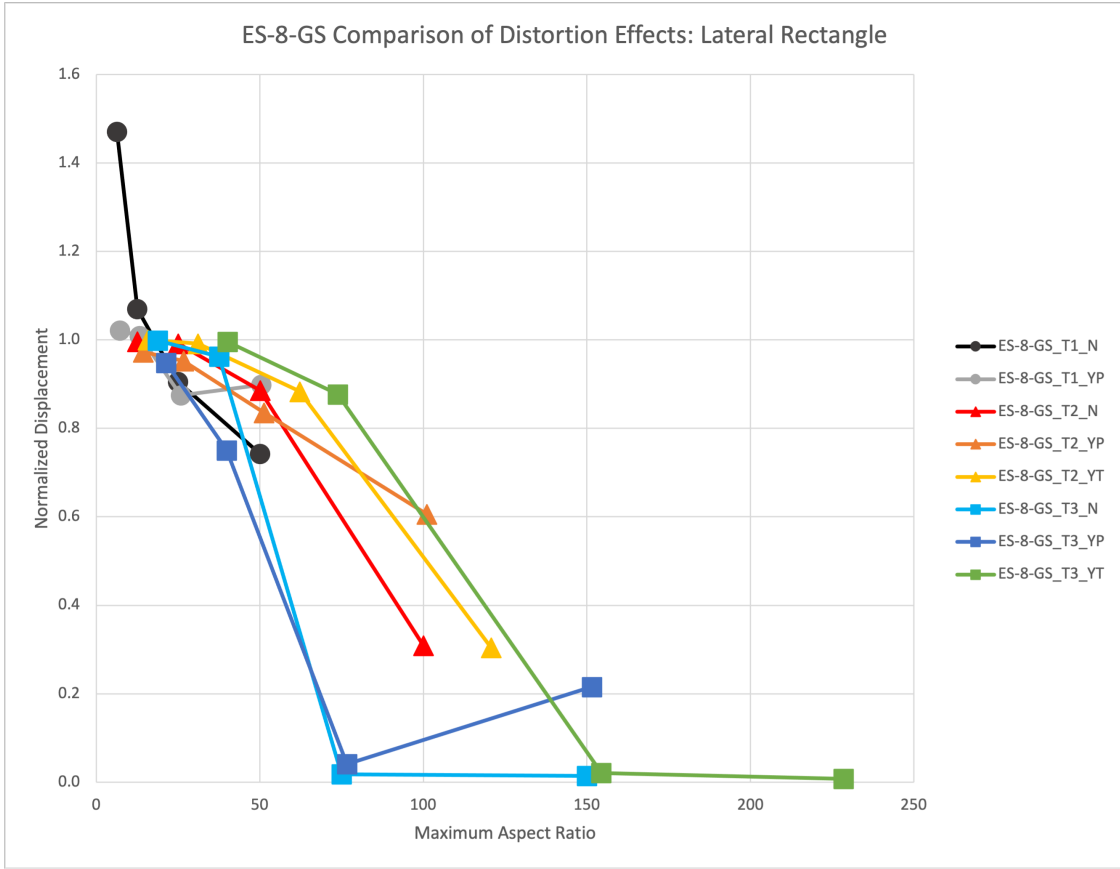
Figure 5.24: Comparison of normalized displacements of the lateral rectangle problem for varying aspect ratios in ES-8-GS elements.

where ν is the Poisson's ratio, p is the applied pressure, E is Young's modulus, r is the inner radius, R is the outer radius, and x is the radius at which the theoretical displacement is calculated. In this problem, $x = r$ since node A is at the inner radius. The calculated values are given in Table 5.10.

Poisson's Ratio	Theoretical Displacement
0.3	4.58250E-03
0.49	5.03993E-03
0.499	5.06025E-03
0.4999	5.06227E-03

Table 5.10: Theoretical displacement of node A as calculated with Eq. 5.8 with $x = r = 3, R = 9, E = 10^3$, and ν varying.
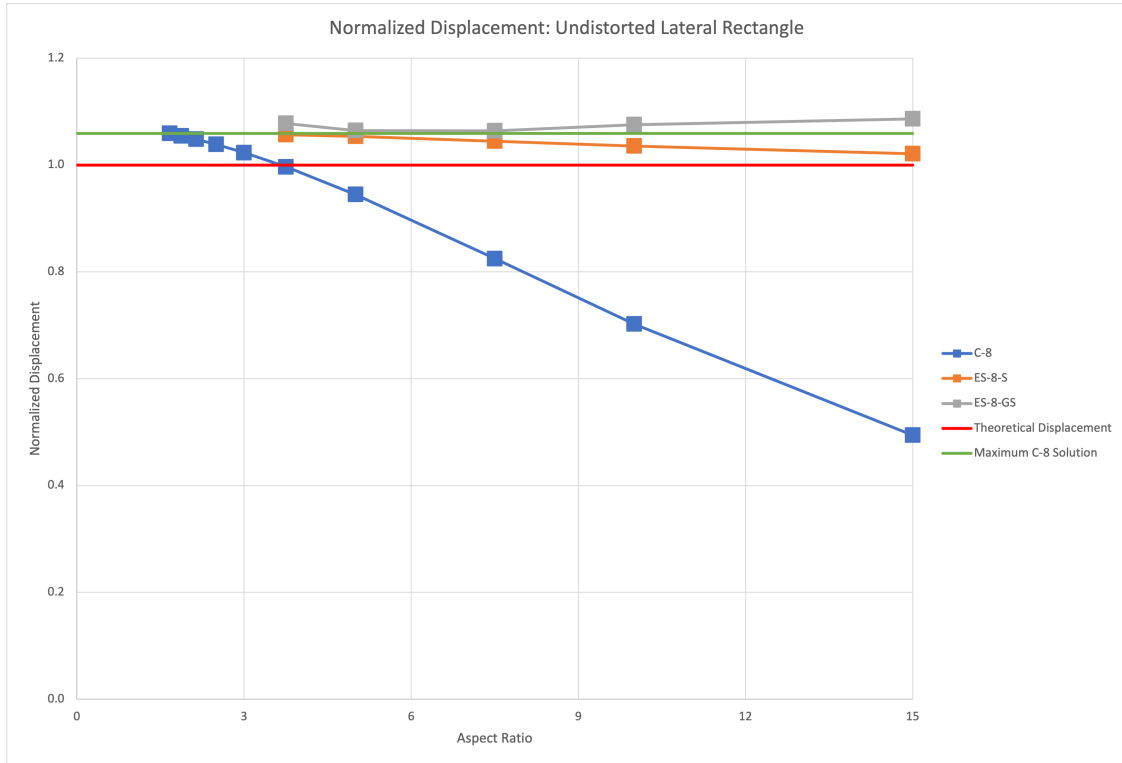
Figure 5.25: Normalized displacement results for varying aspect ratios in the lateral rectangle problem.

5.9.1 Normalized Displacement Results

The normalized displacement results are illustrated in Fig. 5.27. Both of the elements with ghost nodes, ES-8-GS and ES-4-GS, overestimate the solution by up to 32.5% and 16.4%, respectively. Interestingly, as the Poisson's ratio approaches 0.5, the accuracy of these elements slightly improves. Although, all tet-4 element displacement results have a large decrease in results as the Poisson's ratio increases beyond 0.49. This decrease in displacement results is indicative volumetric locking is present. In addition to the C-8 and C-27 elements, the ES-8-S element continues to reproduce accurate results as the Poisson's ratio approaches 0.5.

5.10 Tapered Beam: Effect of No Secondary Nodes

To verify the utility of the ES-8-N element, a tapered cantilever beam (see Fig. 5.28) with a traction of 10^3 in the Y-direction on the free end is compared to the C-8, ES-8-S, and
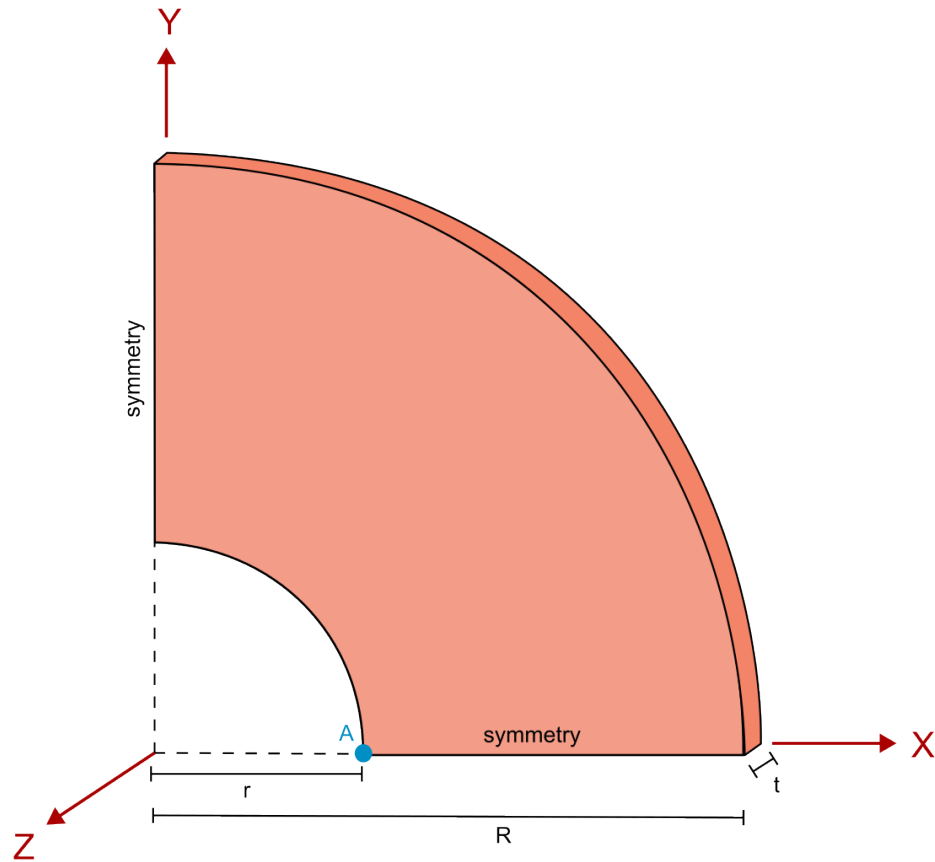
Figure 5.26: A cylinder subject to outward pressure of magnitude 1 on the faces at the inner radius. The edges at X = 0 and Y = 0 are given symmetric BCs.
Material Properties: $E = 10^3, \nu = 0.3 - 0.4999$
Geometrical Properties: $r = 3, R = 9, t = 0.1$
Theoretical Displacement of Node A: In the X-direction, see Table 5.10

ES-8-GS elements. The results of the two elements are compared across multiple different mesh discretizations, characterized by the maximum element diameter. The varying meshes are defined in Table 5.11, where N defines the number of elements on an X-Y plane edge, and M defines the number of elements along the length.

The displacement results at node A are given in Fig. 5.29. The displacement of the
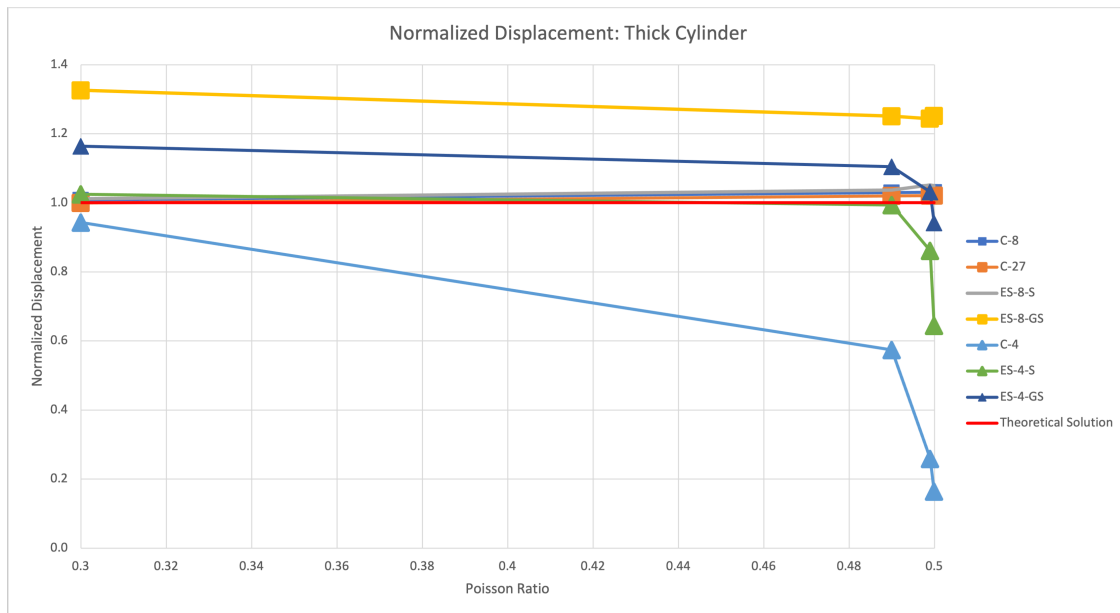
Figure 5.27: Plot of the normalized displacements for the thick cylinder problem for different Poisson's ratio values.
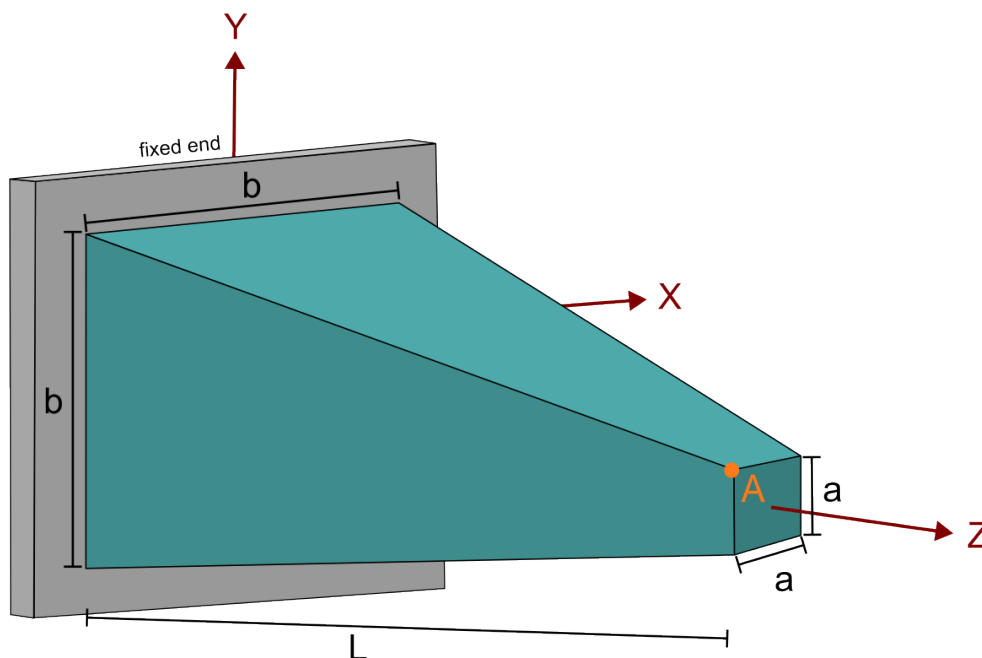

Figure 5.28: A tapered cantilever beam with a traction of 10^3 in the Y-direction at Z = 50.
Material Properties: $E = 10^9, \nu = 0.3$
Geometrical Properties: $L = 50, a = 2, b = 10$

Mesh ID	N	M	Maximum Element Diameter	Hex-8 DP
1	2	30	7.26	2.83E-03
2	4	60	3.68	5.00E-04
3	6	100	2.45	2.00E-04
4	10	60	1.71	1.55E-04
5	15	100	1.11	6.19E-05

Table 5.11: Mesh discretizations for the tapered beam problem.

ES-8-N element is 10% greater than the result for the C-8 element for mesh ID 1. Ideally, the results of these two elements would be similar, however this is not the case for coarse meshes in this problem. As the mesh is refined, the solutions do become much closer. For mesh ID 2, which is still more coarse than the other meshes, the ES-8-N element results in a displacement only 2.6% less than the C-8 element. It would be valuable to study further the coarse mesh results of the ES-8-N element as compared to the C-8 element. However, the quick convergence of the two graph lines indicates that the two elements give very similar results for mildly refined meshes. Note additionally that the ES-8-S and ES-8-GS elements have results close in comparison to the C-8 results.

5.11   Rectangle with Circular Hole: Effect of Plasticity

In order to evaluate the ESFEM's compatibility with plastic deformations, a rectangle with a circular hole cut out is given a uniform traction pulling in opposite directions along the Y-axis. Due to symmetry, only one quarter of the problem is meshed, as shown in Fig. 5.30. For the material model, J2 plasticity is used.

The accuracy of the ESFEM hex-8 elements will be quantitatively compared to a refined C-8 mesh's nodal displacement at node A. The refinement of the mesh is described by varying the number of elements along each edge (N) and through the thickness (M). The maximum displacements of node A is given in Table 5.12. The ES-8-S element mesh ID 2 results in a displacement that is only 1% less than the refined C-8 element. This result is achieved with only one element through the thickness, as opposed to six in mesh ID 3. This result
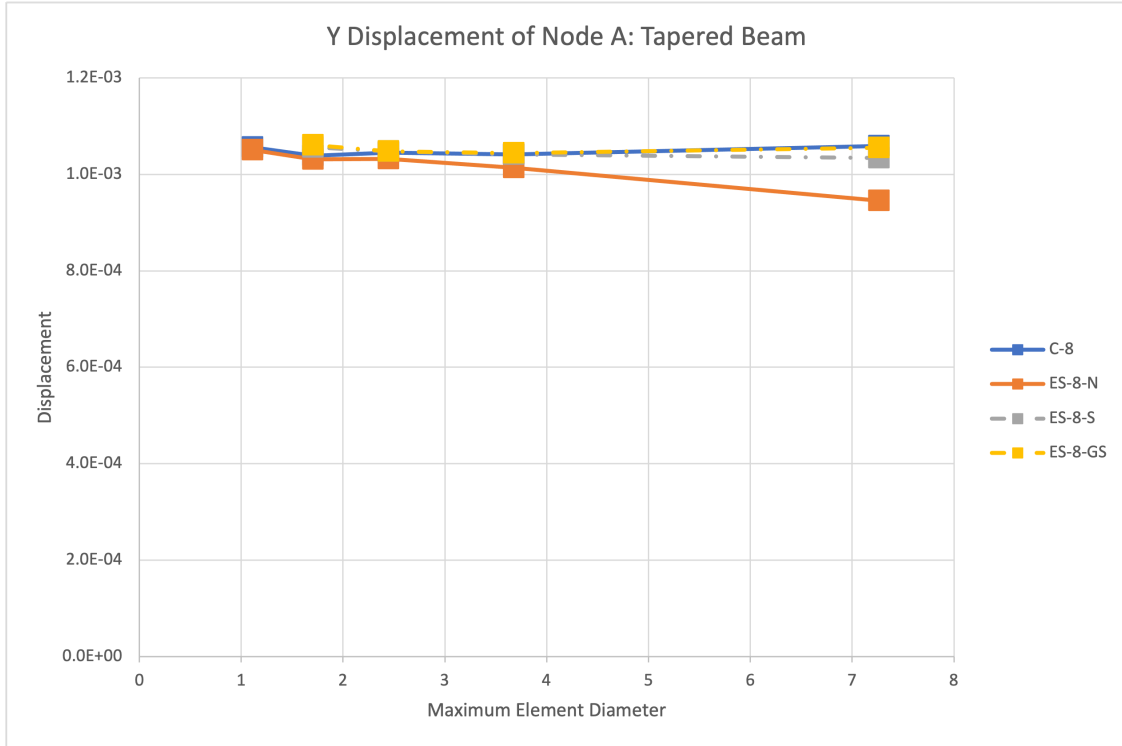
Figure 5.29: A plot of the Y-direction displacement of a tapered cantilever beam at node A.

shows good compatibility with the J2 plasticity material model for the ES-8-S element. For the ES-8-GS element, it does not result in a similar value, with a result 101% greater than the C-8 mesh ID 3 result. Further research to determine if the addition of ghost nodes is increasing the solution would be valuable.

Descriptor	Mesh ID	N	M	Displacement	Hex-8 DP
C-8	1	2	1	-5.21455E-04	7.69E-02
	2	8	1	-5.86955E-04	5.45E-03
	3	15	6	-5.98895E-04	6.39E-04
ES-8-S	1	2	1	-3.40754E-04	7.69E-02
	2	8	1	-5.92027E-04	5.45E-03
ES-8-GS	1	2	1	-3.95670E-04	7.69E-02
	2	8	1	-1.20675E-03	5.45E-03

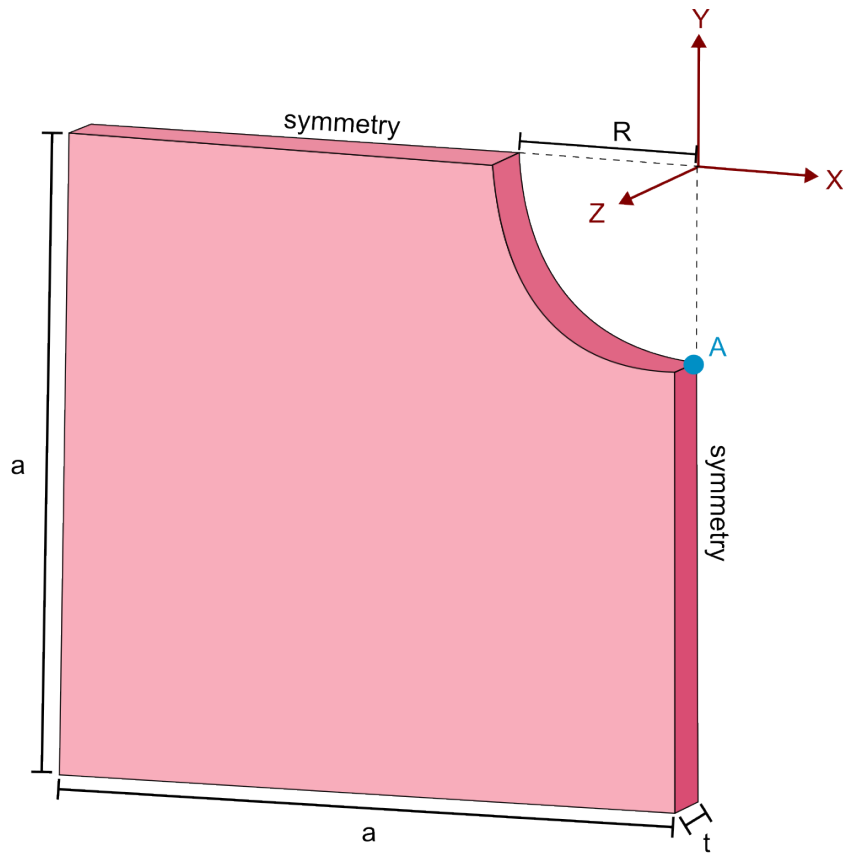Table 5.12: Maximum displacement of node A in the Y-direction for the rectangular plate with a hole.

Figure 5.30: A rectangular plate with a circular hole. Symmetric displacement BCs are given along the edges where X = 0 and Y = 0. A uniform traction is applied along the faces where Y = -1 with a magnitude of -20×10^3 in the Y-direction.
Material Properties: $E = 30 \times 10^6, \nu = 0.3,$ hardening modulus $= 120 \times 10^3$
Geometrical Properties: $R = 0.25, a = 1, t = 0.1$

Chapter 6

Summary and Future Work

The extended-stencil finite element method (ESFEM) is a novel method that takes a conventional finite element mesh and improves its utility by augmenting an element's nodal dependencies. The literature and problems involving distorted hex-27 elements demonstrate that simply increasing the polynomial order of the element interpolants may not always provide a locking-free solution. Avoiding locking and rank deficiency seems to require an "efficient" use of a problem's degrees of freedom. This "efficiency" could be measured by how many degrees of freedom are needed to produce an accurate solution. In this context, the ESFEM for hex-8 elements has shown to be more efficient than the CFEM in Chapter 5.

Future work potential for this method involves further investigation of its capabilities and applications to different element forms. This work focused on locking mitigation and robust handling of distortion in hex-8 and tet-4 elements. There are many other problem domains in the finite element method that the ESFEM could be tested, such as dynamics. And many other element formulations to test, including the 20-node hex.

The current ES-8-S and ES-8-GS elements avoid locking and rank deficiency, however the stress solution at the integration points did not converge as the mesh was refined. The polynomial fit on the faces will likely always be an underdetermined system. And this is where the ESFEM has faced the most challenges. The polynomial fitting method itself seems

to be good, but only if the right set of nodes is referenced by the elements. Ghost nodes provide the faces with more nodal data outside of the mesh to improve the polynomial fit. However, there are many more potential ways to add nodes due to the flexibility of ESFEM.

One potential node addition is a centroidal node. The centroidal node would exist as a primary node at the centroid of the element. This addition may provide more control over the displacement solution inside of the element. Currently, there are still larger oscillations in the basis functions for some problems and that can cause inaccurate stress values. The ESFEM would be of more value if the stresses were accurate and converged to the same value as a refined conventional hex-8 mesh when the displacement solution converges.

Another option is to apply the ESFEM to other element types. For example, a 20-node hex element with mid-edge nodes. The presence of these mid-edge nodes allow for an accurate representation of anticlastic curvature through the thickness of a thin beam without any node additions. The ESFEM may be applied to any CFEM element type, and may be applicable to other FEM elements as well.

While results of the ES-4-S and ES-4-GS elements showed potential improvement from the C-4 element, the ES-4-GS element meshes did not always converge to a solution before the maximum Newton-Raphson iterations was reached. There is a potential that the addition of the ghost nodes for the tet-4 element is not ideal. Additionally, only one integration point in the tet-4 elements poses a risk for rank deficiency when adding ghost nodes. It would be valuable to test 4-point integration for the ES-4-GS elements to verify only 1-point integration is not causing the lack of convergence to a solution.

The problems used in Chapter 5 assumed quasi-static loading, so there was no need for dynamics. Although, dynamics are an important part of many problems, such as airplanes and cars. Value would come from implementing the ESFEM in a dynamic FEM numerical example.

The performance of the rectangular plate with a hole problem reveals a need to further analyze the ESFEM in conjunction with a J2 plasticity material model. Recall that the ES-8-

GS element resulted in a nodal displacement 101% greater than the more refined C-8 result. An analysis with more refined meshes would be able to better convey what is happening with the ES-8-GS element.

The extended-stencil finite element method shows promise for an improved utility of conventional finite element meshes. More numerical examples would assist in the discovery of potential benefits besides locking mitigation and robust handling of mesh distortion. Additionally, expanding the method to other element types may improve accuracy past what is shown for the extended-stencil hex-8 elements.

Bibliography

[1] I. Babuška and J. M. Melenk. The partition of unity method. *International Journal for Numerical Methods in Engineering*, 40(4):727–758, 1997.

[2] T. Belytschko, Y. Y. Lu, and L. Gu. Element-free Galerkin methods. *International Journal for Numerical Methods in Engineering*, 37(2):229–256, jan 1994.

[3] Simon Bieber, Bastian Oesterle, Ekkehard Ramm, and Manfred Bischoff. A variational method to avoid locking—independent of the discretization scheme. *International Journal for Numerical Methods in Engineering*, 114(8):801–827, 2018.

[4] Dhiraj S. Bombarde, Manish Agrawal, Sachin S. Gautam, and Arup Nandy. A locking-free formulation for three-dimensional isogeometric analysis. *Materials Today: Proceedings*, 66:1710–1715, 2022.

[5] Yan Ping Cao, Ning Hu, Jian Lu, Hisao Fukunaga, and Zhen Han Yao. A 3D brick element based on Hu-Washizu variational principle for mesh distortion. *International Journal for Numerical Methods in Engineering*, 53(11):2529–2548, 2002.

[6] Rui P. R. Cardoso and J. M. A. Cesar de Sa. The enhanced assumed strain method for the isogeometric analysis of nearly incompressible deformation of solids. *International Journal for Numerical Methods in Engineering*, 92(1):56–78, jun 2012.

[7] Rui P. R. Cardoso, Jeong Whan Yoon, and Robert E. Dick. A new axi-symmetric element for thin walled structures. *Computational Mechanics*, 45(4):281–296, mar 2010.

[8] J. F. Caseiro, R. J. Alves De Sousa, and R. A.F. Valente. A systematic development of EAS three-dimensional finite elements for the alleviation of locking phenomena. *Finite Elements in Analysis and Design*, 73:30–41, 2013.

[9] Satish Chandra and Gangan Prathap. A field-consistent formulation for the eight-noded solid finite element. *Computers and Structures*, 33(2):345–355, 1989.

[10] Ray W. Clough. Original formulation of the finite element method. *Finite Elements in Analysis and Design*, 7(2):89–101, nov 1990.

[11] E. A. De Souza Neto, D. Perić, M. Dutko, and D. R.J. Owen. Design of simple low order finite elements for large strain analysis of nearly incompressible solids. *International Journal of Solids and Structures*, 33(20-22):3277–3296, 1996.

[12] Gmsh: a three-dimensional finite element mesh generator with built-in pre- and post-processing facilities.

[13] P. Q. He, Q. Sun, and K. Liang. Generalized modal element method: part-I—theory and its application to eight-node asymmetric and symmetric solid elements in linear analysis. *Computational Mechanics*, 63(4):755–781, 2019.

[14] A. Hrennikoff. Solution of Problems of Elasticity by the Framework Method. *Journal of Applied Mechanics, Transactions ASME*, 8(4):A169–A175, 1941.

[15] Thomas J. R. Hughes. Generalization of selective integration procedures to anisotropic and nonlinear media. *International Journal for Numerical Methods in Engineering*, 15(9):1413–1418, sep 1980.

[16] Thomas J. R. Hughes, Robert L. Taylor, and Worsak Kanoknukulchai. A simple and efficient finite element for plate bending. *International Journal for Numerical Methods in Engineering*, 11(10):1529–1543, 1977.

[17] Thomas J.R. Hughes. *The Finite Element Method: Linear and Static Finite Element Analysis.* Dover Publications Inc., Mineola, New York, 2000.

[18] T.J.R. Hughes, J.A. Cottrell, and Y Bazilevs. Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Computer Methods in Applied Mechanics and Engineering*, 194(39-41):4135–4195, oct 2005.

[19] Jože Korelc, Urša Šolinc, and Peter Wriggers. An improved EAS brick element for finite deformation. *Computational Mechanics*, 46(4):641–659, 2010.

[20] P Krysl and B Zhu. Locking-free continuum displacement finite elements with nodal integration. *International Journal for Numerical Methods in Engineering*, 76(7):1020–1043, apr 2008.

[21] Wing Kam Liu, Yong Guo, Sing Tang, and Ted Belytschko. A multiple-quadrature eight-node hexahedral finite element for large deformation elastoplastic analysis. *Computer Methods in Applied Mechanics and Engineering*, 154(1-2):69–132, feb 1998.

[22] Wing Kam Liu, Shaofan Li, and Harold S. Park. Eighty Years of the Finite Element Method: Birth, Evolution, and Future. *Archives of Computational Methods in Engineering*, 29(6):4431–4453, 2022.

[23] Jinwei Ma, Qinglin Duan, and Rong Tian. A generalized finite element method without extra degrees of freedom for large deformation analysis of three-dimensional elastic and elastoplastic solids. *Computer Methods in Applied Mechanics and Engineering*, 392, mar 2022.

[24] Richard H. Macneal and Robert L Harder. A proposed standard set of problems to test finite element accuracy. *Finite Elements in Analysis and Design*, 1(1):3–20, apr 1985.

[25] David S Malkus and Thomas J.R. Hughes. Mixed finite element methods — Reduced and selective integration techniques: A unification of concepts. *Computer Methods in Applied Mechanics and Engineering*, 15(1):63–81, jul 1978.

[26] D. Mijuca. On hexahedral finite element HC8/27 in elasticity. *Computational Mechanics*, 33(6):466–480, may 2004.

[27] E. A.de Souza Neto, F. M. Andrade Pires, and D. R. J. Owen. F-bar-based linear triangles and tetrahedra for finite strain analysis of nearly incompressible solids. Part I: formulation and benchmarking. *International Journal for Numerical Methods in Engineering*, 62(3):353–383, jan 2005.

[28] Vinh Phu Nguyen, Cosmin Anitescu, Stéphane P.A. Bordas, and Timon Rabczuk. Isogeometric analysis: An overview and computer implementation aspects. *Mathematics and Computers in Simulation*, 117:89–116, 2015.

[29] Y. Onishi and K. Amaya. A locking-free selective smoothed finite element method using tetrahedral and triangular elements with adaptive mesh rezoning for large deformation problems. *International Journal for Numerical Methods in Engineering*, 99(5):354–371, aug 2014.

[30] Y. Onishi, R. Iida, and K. Amaya. F-bar aided edge-based smoothed finite element method using tetrahedral elements for finite deformation analysisof nearly incompressible solids. *International Journal for Numerical Methods in Engineering*, 109(11):1582–1606, 2017.

[31] Robin Pfefferkorn and Peter Betsch. Hourglassing- and locking-free mesh distortion insensitive Petrov–Galerkin EAS element for large deformation solid mechanics. *International Journal for Numerical Methods in Engineering*, 124(6):1307–1343, mar 2023.

[32] Michael Anthony Puso. A highly efficient enhanced assumed strain physically stabilized hexahedral element. *International Journal for Numerical Methods in Engineering*, 49(8):1029–1064, 2000.

[33] David Per Recio, R. M. Natal Jorge, and L. M. S. Dinis. Locking and hourglass phenomena in an element-free Galerkin context: the B-bar method with stabilization and an enhanced strain method. *International Journal for Numerical Methods in Engineering*, 68(13):1329–1357, dec 2006.

[34] Savvas Saloustros, Miguel Cervera, Sungchul Kim, and Michele Chiumenti. Accurate and locking-free analysis of beams, plates and shells using solid elements. *Computational Mechanics*, 67(3):883–914, mar 2021.

[35] Zhong-ci Shi. The F-E-M-Test for Convergence of Nonconforming Finite Elements. *Mathematics of Computation*, 49(180):391–405, oct 1987.

[36] Friedrich Stummel. The Generalized Patch Test. *SIAM Journal on Numerical Analysis*, 16(3):449–471, 1979.

[37] Rong Tian. Extra-dof-free and linearly independent enrichments in GFEM. *Computer Methods in Applied Mechanics and Engineering*, 266:1–22, 2013.

[38] Jue Wang and R. H. Wagoner. A practical large-strain solid finite element for sheet forming. *International Journal for Numerical Methods in Engineering*, 63(4):473–501, may 2005.

[39] Z. Cedric Xia and Sing C. Tang. Locking Phenomena in the Use of Solid Elements for Sheet Metal Forming Simulation. In *SAE Technical Papers*, volume 108, pages 862–868, mar 1999.

[40] W. Zeng and G. R. Liu. Smoothed Finite Element Methods (S-FEM): An Overview and Recent Developments. *Archives of Computational Methods in Engineering*, 25(2):397–435, 2018.

[41] Shiquan Zhang and Xiaoping Xie. Accurate 8-node hybrid hexahedral elements with energy-compatible stress modes. *Advances in Applied Mathematics and Mechanics*, 2(3):333–354, 2010.

[42] Wen Zhong, Yu Qi Liu, Yun Ming Hu, Sheng Qiang Li, and Heng Jian Xu. A Selective Reduced Integration 8-Node Hexahedral Element by Assumed Strain for Coining Simulation. *Advanced Materials Research*, 472-475:533–537, feb 2012.

[43] Y. Y. Zhu and S. Cescotto. Unified and mixed formulation of the 8-node hexahedral elements by assumed strain method. *Computer Methods in Applied Mechanics and Engineering*, 129(1-2):177–209, 1996.