A Locally Conservative Stress Recovery Technique for Continuous Galerkin FEM in Linear Elasticity

Lawrence Bush^{1,†}, Quanling Deng^{1,*}, and Victor Ginting^{1,†}

¹ *Department of Mathematics, University of Wyoming, Laramie, WY 82071.*

Abstract

The standard continuous Galerkin finite element method (FEM) is a versatile and well understood method for solving partial differential equations. However, one shortcoming of the method is lack of continuity of derivatives of the approximate solution at element boundaries. This leads to undesirable consequences for a variety of problems such as a lack of local conservation. A two-step postprocessing technique is developed in order to obtain a local conservation from the standard continuous Galerkin FEM on a vertex centered dual mesh relative to the finite element mesh when applied to displacement based linear elasticity. The postprocessing requires an auxiliary fully Neumann problem to be solved on each finite element where local problems are independent of each other and involve solving two small linear algebra systems whose sizes are 3-by-3 when using linear finite elements on a triangular mesh for displacement based linear elasticity. The postprocessed stresses then satisfy local conservation on the dual mesh. An a priori error analysis and numerical simulations are provided.

Keywords: stress recovery, local conservation, two-step postprocessing, linear finite element, displacement, linear elasticity

---

1. Introduction

Linear elasticity is important for many engineering applications. Given an elastic body with associated constraints and forces applied to it, we would like

---

^{0*} Correspondence to: Quanling Deng, University of Wyoming, Email: qdeng@uwyo.edu

^{0†} This work is partially supported by the grants from DOE (DE-FE0004832 and DE-SC0004982), the Center for Fundamentals of Subsurface Flow of the School of Energy Resources of the University of Wyoming (WYDEQ49811GNTG), and from NSF (DMS-1016283).

to know the resulting deformation of the body. In many applications, secondary information is also needed such as the stress in the body. The equations governing linear elasticity are elliptic in category. Three of the main methods for solving the displacement based elasticity problem are finite differences, finite volume, and continuous Galerkin FEMs [16, 11, 15, 23]. More recently, discontinuous Galerkin FEMs and mixed FEMs have been used [19, 3]. In the case of finite differences, finite volume, and FEMs for displacement based linear elasticity, a postprocessing is needed to recover the stress. The next few paragraphs discuss some of the techniques associated with FEMs.

Many methods have been proposed to recover stress from displacement based finite element. The two most common methods are simple averaging [5], where the stress is computed using the basis functions on each element in a finite element discretization and the second uses the enhanced accuracy properties of the recovery at Gauss points in each element [5]. The former method may be easy to implement, but since the finite element formulation of the problem is to obtain a C^0 representation of the displacement, there are discontinuities in the stress at the finite element boundaries. Typically this is corrected through averaging element stresses to compute stress at element boundaries. The second method benefits from additional accuracy in the stress calculations at Gauss points, which is then used to extend the stresses to other points in the element.

Another category of methods which have been proposed are smoothing techniques as presented in [22, 21] and the references therein. Generally the techniques use global or local least squares fitting to obtain a stress field with the desired smoothness. One advantage of the methods presented there is that it allows a direct way to estimate the error which can be used for mesh refinement procedures in order to reduce the error in critical regions.

Another more recent procedure uses nodal point forces to improve the accuracy of the recovered stress [18, 17]. The advantage of this method is that the stresses obtained are more accurate than those obtained directly from the finite element solutions. The procedure can be viewed as a postprocessing on a patch of elements the authors call the stress calculation domain, hence it is not a global formulation.

Some postprocessing techniques have been applied for conducting a posteriori error analysis. Zienkiewicz and Zhu [25, 26, 27] proposed a technique to postprocess the discontinuous gradient of the numerical solution using interpolation functions to obtain a recovery-based a posteriori error estimator. The error representation obtained there is close to the procedure proposed by Sussman and Bathe [20], who suggest to find an estimate for the true error by comparing the

postprocessed gradient and the non-postprocessed gradient. Some a posteriori error estimates for the finite element solution of elliptic problems are reviewed in [14]. For more details on the procedures and postprocessing techniques applied to obtain the a posteriori error estimation for finite element solutions, see for example Bank and Weiser [4], Ainsworth and Oden [1], Carstensen et al. [8], Verfürth [24], and the references therein.

All of the methods mentioned above do not necessarily satisfy the local conservation property of the governing equation due to the nature of the solution representation except for the finite volume method in which conservation is guaranteed by the formulation. The two-step postprocessing method presented in this paper provides a locally conservative stress on a vertex centered dual mesh by posing a fully Neumann problem on each element whose boundary conditions come from the finite element displacement solution. This is a similar postprocessing technique that has been developed in [6, 7] for scalar elliptic problems. Here we extend it to a two-step postprocessing technique for the displacement based formulation of the linear elasticity problem since it is elliptic in category. The local problem and its solution technique are posed in such a way so that the local conservation is obtained on the dual mesh. The main motivation for seeking a locally conservative solution is to represent the physics of the governing equation as accurately as possible.

The rest of the paper is organized as follows. First, the model problem is presented in Section 2. Second the Galerkin finite element formulation of the problem is presented in Section 3 followed by the two-step postprocessing to recover the stresses in Section 4. Section 5 presents an analysis of the technique to establish the convergence of the postprocessed quantity to the true one. Finally numerical examples are presented in Section 6 to verify the analysis presented in Section 5 as well as a general two-step postprocessing discussion.

2. Model problem

We will focus on linear elasticity on two-dimensional space, but the technique is directly applicable to the three dimensional case also. The governing equation for (σ, u) is given by

$$\begin{cases} -\nabla \cdot \sigma = f & \text{in } \Omega \subset \mathbb{R}^2 \\ \sigma = \lambda(\nabla \cdot u)I + \mu\epsilon(u) & \text{in } \Omega \subset \mathbb{R}^2 \\ u_i = u_{D,i} & \text{on } \Gamma_{D,i} \text{ for } i = 1, 2 \\ (\sigma \cdot n)_i = g_{N,i} & \text{on } \Gamma_{N,i} \text{ for } i = 1, 2, \end{cases} \quad (2.1)$$

where $\boldsymbol{\sigma}$ is the stress tensor with components σ_{ij} , $i, j = 1, 2$ that represent a force per unit area in the x_i direction with unit normal in the x_j direction, $\mathbf{u} = (u_1, u_2)$ is the displacement in the x_1 and x_2 directions respectively, $\mathbf{f} = (f_1, f_2)$ is a body force, $\boldsymbol{\epsilon}(\mathbf{u})$ is the strain tensor given below, I is the 2-by-2 identity matrix, λ and μ are material parameters, and boundary conditions are given by $\mathbf{u}_D = (u_{D,1}, u_{D,2})$ and $\mathbf{g}_N = (g_{N,1}, g_{N,2})$. It should also be noted here that the boundary conditions are such that $\Gamma_{D,i} \cap \Gamma_{N,i} = \emptyset$ and $\Gamma_{D,i} \cup \Gamma_{N,i} = \partial\Omega$ for $i = 1, 2$. The goal is to find the stress $\boldsymbol{\sigma}$ and displacement $\mathbf{u}$.

One typical approach to solving (2.1) is to formulate the problem in terms of the displacement $\mathbf{u}$, which is given below. We start by first defining the stress and strain components

$$\begin{cases} \sigma_{ij} = \lambda(\nabla \cdot \mathbf{u})\delta_{ij} + \mu\epsilon_{ij}(\mathbf{u}), \text{ and} \\ \epsilon_{ij}(\mathbf{u}) = \frac{1}{2} \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \text{ for } i, j = 1, 2. \end{cases} \quad (2.2)$$

Here δ_{ij} is the usual Kronecker delta and it should be noted that

$$(\boldsymbol{\sigma} \cdot \mathbf{n})_i = \sum_{j=1}^2 \sigma_{ij} n_j.$$

This allows (2.1) to be written in component form

$$\begin{cases} -\nabla \cdot (\mathbf{a}_{i1} \nabla u_1) - \nabla \cdot (\mathbf{a}_{i2} \nabla u_2) = f_i & \text{in } \Omega \subset \mathbb{R}^2 \text{ for } i = 1, 2 \\ u_i = u_{D,i} & \text{on } \Gamma_{D,i} \\ (\mathbf{a}_{i1} \nabla u_1 + \mathbf{a}_{i2} \nabla u_2) \cdot \mathbf{n} = g_{N,i} & \text{on } \Gamma_{N,i}, \end{cases} \quad (2.3)$$

where the matrices $\mathbf{a}_{ij}$ are given by

$$\mathbf{a}_{11} = \begin{bmatrix} \lambda + \mu & 0 \\ 0 & \mu \end{bmatrix}, \mathbf{a}_{12} = \begin{bmatrix} 0 & \lambda \\ \mu & 0 \end{bmatrix} = \mathbf{a}_{21}^\top, \mathbf{a}_{22} = \begin{bmatrix} \mu & 0 \\ 0 & \lambda + \mu \end{bmatrix}. \quad (2.4)$$

This is the displacement based formulation of the problem. The next section will derive the variational form of the problem, which will result in the Galerkin finite element formulation of the problem.

3. Galerkin finite element formulation

We start by introducing the variational formulation associated with (2.3). We define

$$\mathcal{V} = \{(v_1, v_2) : v_i \in H^1(\Omega), v_i|_{\Gamma_{D,i}} = 0, \ i = 1, 2\}.$$

Multiplying (2.3) (i.e taking the Euclidean dot product) by $\mathbf{v} \in \mathcal{V}$ and integration by parts yields the variational equation for (2.3), where we find $\mathbf{u}$ with $(\mathbf{u} - \mathbf{u}_D) \in \mathcal{V}$ satisfying

$$a(\mathbf{u}, \mathbf{v}) = (\mathbf{f}, \mathbf{v}) + \langle \mathbf{g}_N, \mathbf{v} \rangle \quad \forall \mathbf{v} \in \mathcal{V},$$

where

$$\begin{aligned} a(\mathbf{u}, \mathbf{v}) &= \sum_{i=1}^2 \sum_{j=1}^2 a_{ij}(u_j, v_i), \\ (\mathbf{f}, \mathbf{v}) &= \sum_{i=1}^2 (f_i, v_i), \quad \text{and} \\ \langle \mathbf{g}_N, \mathbf{v} \rangle &= \sum_{i=1}^2 \int_{\Gamma_{N,i}} g_i v_i \, dl, \end{aligned} \tag{3.1}$$

i.e., respectively the bilinear form and usual $(L^2(\cdot))^2$ inner product on the domain Ω and Neumann boundary Γ_N . Here, $a_{ij}(u_j, v_i) = \int_{\Omega} \mathbf{a}_{ij} \nabla u_j \cdot \nabla v_i \, dA$ and $(f_i, v_i) = \int_{\Omega} f_i v_i \, dA$. Existence and uniqueness for this problem are established in [11].

We let Ω be a polygonal domain with triangulation $\mathcal{T}_h$ consisting of closed triangle (or quadrilateral) elements τ such that $\overline{\Omega} = \cup_{\tau \in \mathcal{T}_h} \tau$, where $h = \max_{\tau \in \mathcal{T}_h} \{h_{\tau}\}$ and h_{τ} is the diameter of τ . We then define the conforming linear finite element space $\mathcal{V}_h$ over $\mathcal{T}_h$ as

$$\mathcal{V}_h = \{\mathbf{v}_h \in C(\overline{\Omega})^2 : \mathbf{v}_h|_{\tau} \in (\mathcal{P}_1(\tau))^2 \text{ and } \mathbf{v}_h|_{\Gamma_{D,i}} = 0, \ i = 1, 2\},$$

where $\mathcal{P}_1(\tau)$ is a linear polynomial when restricted to τ . We denote the set of vertices in Ω by Z_{in} , $Z_{d,j}$, and $Z_{n,j}$, namely the set of vertices on the interior, $\Gamma_{D,j}$, and $\Gamma_{N,j}$ respectively. We denote $\mathcal{V}_h = \text{span}\{\{\Phi_z^{(1)}\}_{z \in Z_{\text{in}} \cup Z_{n,1}}, \{\Phi_z^{(2)}\}_{z \in Z_{\text{in}} \cup Z_{n,2}}\}$, where

$$\Phi_z^{(1)} = \begin{bmatrix} \phi_z \\ 0 \end{bmatrix} \quad \text{and} \quad \Phi_z^{(2)} = \begin{bmatrix} 0 \\ \phi_z \end{bmatrix} \tag{3.2}$$

with ϕ_z the usual scalar linear basis function so that the finite element solution $\mathbf{u}_h$ is expressed as the linear combination

$$\mathbf{u}_h = \sum_{j=1}^2 \sum_{\zeta \in Z \cup Z_{n,j}} \alpha_{\zeta}^{(j)} \Phi_{\zeta}^{(j)}, \quad (3.3)$$

which satisfies

$$\begin{aligned} a(\mathbf{u}_h, \Phi_z^{(i)}) &= (\mathbf{f}, \Phi_z^{(i)}), \quad \forall z \in Z_{\text{in}}, \quad i = 1, 2 \\ a(\mathbf{u}_h, \Phi_z^{(i)}) &= (\mathbf{f}, \Phi_z^{(i)}) + \langle \mathbf{g}_N, \Phi_z^{(i)} \rangle, \quad \forall z \in Z_{n,i}, \quad i = 1, 2 \end{aligned} \quad (3.4)$$

This yields a linear system of the form $\mathbf{A}\alpha = \mathbf{F}$, where α is a vector whose entries are from the set $\{\{\alpha_{\zeta}^{(1)}\}_{\zeta \in Z_{\text{in}} \cup Z_{n,1}}, \{\alpha_{\zeta}^{(2)}\}_{\zeta \in Z_{\text{in}} \cup Z_{n,2}}\}$. $\mathbf{A}$ is a square matrix whose entries are $a(\Phi_{\zeta}^{(j)}, \Phi_z^{(i)})$, and $\mathbf{F}$ is a vector whose entries are $(\mathbf{f}, \Phi_{\zeta}^{(i)})$ with the additional Neumann boundary term where applicable.

4. A two-step postprocessing for the numerical displacement

Once the displacement solution to (2.1) has been found, we would like to compute the stress σ that satisfies local conservation property to be described next. We do this by constructing control volumes around each node in the finite element mesh and imposing local conservation there. To the best of the author's knowledge, compared with the other stress recovery methods mentioned in the introduction, this is the only one that satisfies the local conservation inherent in the governing equation. The conservation on each control volume can then be turned into a local element-by-element calculation that results in locally conservative stress. The postprocessed quantity is shown to inherit the convergence property of the original finite element solution. The following paragraphs describe the two-step postprocessing procedure, which is similar to the technique presented in [6] for a scalar elliptic problem. For three-dimensional linear elasticity problem, this procedure can be extended to a three-step postprocessing technique and the corresponding analysis in Section 5 follows naturally.

For an interior vertex $z \in Z_{\text{in}}$, let Ω^z be the support of the basis function $\Phi_z^{(j)}$. We denote the support of this basis function by $\Omega^z = \cup_{i=1}^N \tau_i^z$, where τ_i^z is a finite element and z is a vertex of the finite element. N is the number of elements that have z as a vertex. The number of elements containing z will vary from one vertex to the next in the case of triangular elements.

Figure 1: Ω^z : support of $\Phi_z^{(j)}$ (left), control volume C^z (middle), a finite element τ (right) for triangular meshes.

We also define

$$a_{\tau}(\mathbf{u}, \mathbf{v}) = \sum_{i=1}^2 \sum_{j=1}^2 a_{ij}^{\tau}(u_j, v_i), \quad (4.1)$$

where

$$a_{ij}^{\tau}(v, w) = \int_{\tau} \mathbf{a}_{ij} \nabla v \cdot \nabla w \, dA. \quad (4.2)$$

With this in mind, we write

$$a(\mathbf{u}_h, \Phi_z^{(j)}) = \sum_{i=1}^N Q_i^{z,j}, \quad j = 1, 2, \quad z \in Z_{\text{in}} \quad \text{and} \quad (4.3)$$

$$(\mathbf{f}, \Phi_z^{(j)}) = \sum_{i=1}^N F_i^{z,j} \quad j = 1, 2, \quad z \in Z_{\text{in}},$$

where $Q_i^{z,j} = a_{\tau_i}^{z,j}(\mathbf{u}_h, \Phi_z^{(j)})$ and $F_i^{z,j} = (\mathbf{f}, \Phi_z^{(j)})_{\tau_i}$, i.e. the bilinear form and inner products from (3.1) restricted to τ_i . Equation (3.4) becomes

$$\sum_{i=1}^N Q_i^{z,j} = \sum_{i=1}^N F_i^{z,j}, \quad \text{for } j = 1, 2, \quad z \in Z_{\text{in}}. \quad (4.4)$$

Equation (4.4) allows conservation to be satisfied on the vertex centered control volumes shown in Figure 1. For a vertex $z \in Z_{n,j}$ the term associated with the Neumann condition, $\langle \mathbf{g}_N, \Phi_z^{(j)} \rangle$, must be added to the right hand side of (4.4).

We would like to obtain a locally conservative stress from the Galerkin FEM solution that satisfies

$$-\int_{\partial C^z} \tilde{\sigma}_h \cdot \mathbf{n} \, dl = \int_{C^z} f \, dA, \quad \forall z \in Z, \quad (4.5)$$

where Z is the set of all vertices. The control volumes on which we would like the local conservation to hold are the previously mentioned vertex centered control volumes C^z , each of which is associated with vertex z , and constructed as follows. For a finite element τ , a quadrilateral associated with a vertex z is formed by connecting the barycenter of τ to the midpoint of each of the edges sharing z (see t_z in the third column of Figure 1). C^z is constructed by taking the union of all quadrilaterals from each of the elements sharing z . The boundary of control volume ∂C^z is piecewise differentiable can be broken up into the individual pieces making up the boundary of the volume, i.e., $\partial C^z = \cup_{i=1}^N e_i^z$ (see Figure 1). This allows (4.5) to be written as

$$-\sum_{i=1}^N \int_{e_i^z} \tilde{\sigma}_h \cdot \mathbf{n} \, dl = \int_{C^z} f \, dA, \quad \forall z \in Z_{\text{in}}, \quad (4.6)$$

or in component form

$$-\sum_{i=1}^N \int_{e_i^z} \tilde{\sigma}_{h,j1} n_1 + \tilde{\sigma}_{h,j2} n_2 \, dl = \int_{C^z} f_j \, dA, \quad j = 1, 2, \quad \forall z \in Z_{\text{in}}. \quad (4.7)$$

We would like to have element based calculations that are independent of each other and avoid the need to do a calculation on each control volume or a global calculation. This is the subject of the next subsection.

4.1. Elemental calculation

We propose to construct $\tilde{\sigma}_h \cdot \mathbf{n}$ satisfying (4.7) using element-by-element calculations. An auxiliary variational boundary value problem posed in an element τ having vertices $v(\tau) = \{x, y, z\}$ in the triangular element (see the right plots of Figure 1). We designate $\partial\tau = \cup_{\zeta \in v(\tau)} E_\zeta^\tau$, where $E_\zeta^\tau = \partial\tau \cap \partial t_\zeta$ (i.e. half of each element edge containing the vertex ζ).

To state the elemental variational equation, we set $V(\tau) = \text{span}\{\phi_\zeta\}_{\zeta \in v(\tau)}$ and a map I_τ with

$$I_\tau : V(\tau) \rightarrow V^0(\tau)$$

defined as

$$I_\tau w = \sum_{\zeta \in v(\tau)} w(\zeta) \psi_\zeta \quad \text{and} \quad \psi_\zeta(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in t_\zeta \\ 0 & \text{if } \mathbf{x} \notin t_\zeta \end{cases}.$$

This map has a property [9, 10] that

$$\|w - I_\tau w\|_{L^2(\tau)} \leq Ch_\tau \|\nabla w\|_{L^2(\tau)}, \quad \text{for } w \in V(\tau). \quad (4.8)$$

This way, we have $V^0(\tau) = \text{span}\{\psi_\zeta\}_{\zeta \in \mathcal{V}(\tau)}$, i.e., $I_\tau w$ is a piecewise constant function on τ . Moreover, we consider again the right plots of Figure 1, which is now seen as a discretization of τ into quadrilaterals t_ζ , i.e., $\tau = \cup_{\zeta \in \mathcal{V}(\tau)} t_\zeta$, each of which can be written as $t_\zeta = C^\zeta \cap \tau$.

The elemental calculation consists of two steps. The first step is the postprocessing for finding a solution $\widetilde{u}_{h,1}$ satisfying (4.7) when $j = 1$ with $u_{h,2}$ served as a data in the equation. The second step is the postprocessing for finding a solution $\widetilde{u}_{h,2}$ satisfying (4.7) when $j = 2$ with $u_{h,1}$ served as a data in the equation. For this second step, we could have used $\widetilde{u}_{h,1}$ obtained from the first step instead and it still yields the conservative stresses with a same error convergence rate. The analysis for this case is very similar to the analysis in Section 5. Here, however, we use Galerkin FEM solution $u_{h,1}$ for the second step. The details of the two-step postprocessing are given as follows.

4.1.1. First step

The elemental calculation for the first step is to find $\widetilde{u}_{\tau,h,1} \in V(\tau)$ satisfying

$$b_\tau^1(\widetilde{u}_{\tau,h,1}, I_\tau w) = \ell_\tau^1(I_\tau w) \quad \forall I_\tau w \in V^0(\tau), \quad (4.9)$$

where

$$\begin{aligned} b_\tau^1(\widetilde{u}_{\tau,h,1}, I_\tau w) &= \sum_{\zeta \in \mathcal{V}(\tau)} \int_{\partial t_\zeta} (-\mathbf{a}_{11} \nabla \widetilde{u}_{\tau,h,1} \cdot \mathbf{n}) I_\tau w \, dl \\ \ell_\tau^1(I_\tau w) &= \int_\tau f_1 I_\tau w \, dA + \sum_{\zeta \in \mathcal{V}(\tau)} \int_{\partial t_\zeta} (\mathbf{a}_{12} \nabla u_{h,2} \cdot \mathbf{n}) I_\tau w \, dl, \end{aligned}$$

with the boundary condition imposed on $\partial\tau$ to satisfy

$$\int_{\partial\tau} (-\mathbf{a}_{11} \nabla \widetilde{u}_{\tau,h,1} \cdot \mathbf{n} - \mathbf{a}_{12} \nabla u_{h,2} \cdot \mathbf{n}) I_\tau w \, dl = \int_{\partial\tau} \widetilde{g}_{\tau,1} I_\tau w \, dl, \quad (4.10)$$

where

$$\widetilde{g}_{\tau,1} = \sum_{\zeta \in \mathcal{V}(\tau)} (F_{\tau,1}^\zeta - Q_{\tau,1}^\zeta) P_\zeta \psi_\zeta, \quad (4.11)$$

where $F_{\tau,1}^\zeta = (f_1, \phi_\zeta)_\tau$ and $Q_{\tau,1}^\zeta = \sum_{j=1}^2 a_{1j}^\tau(u_{h,j}, \phi_\zeta)$ (i.e., the inner product and bilinear form given in (3.1) restricted to the element τ for the first component), and P_ζ are polynomials of degree two in τ satisfying

$$\int_{E_\zeta^\tau} P_\zeta dl = 1 \quad \text{and} \quad \int_{E_\zeta^\tau} P_\zeta \phi_\eta dl = \delta_{\zeta\eta} \quad \text{for } \eta, \zeta \in v(\tau), \quad (4.12)$$

where $\delta_{\zeta\eta}$ is the usual Dirac delta.

4.1.2. Second step

The elemental calculation for the second step is to find $\tilde{u}_{\tau,h,2} \in V(\tau)$ satisfying

$$b_\tau^2(\tilde{u}_{\tau,h,2}, I_\tau w) = \ell_\tau^2(I_\tau w) \quad \forall I_\tau w \in V^0(\tau), \quad (4.13)$$

where

$$\begin{aligned} b_\tau^2(\tilde{u}_{\tau,h,2}, I_\tau w) &= \sum_{\zeta \in v(\tau)} \int_{\partial t_\zeta} (-a_{22} \nabla \tilde{u}_{\tau,h,2} \cdot \mathbf{n}) I_\tau w dl \\ \ell_\tau^2(I_\tau w) &= \int_\tau f_2 I_\tau w dA + \sum_{\zeta \in v(\tau)} \int_{\partial t_\zeta} (a_{21} \nabla u_{h,1} \cdot \mathbf{n}) I_\tau w dl, \end{aligned}$$

with the boundary condition imposed on $\partial\tau$ to satisfy

$$\int_{\partial\tau} (-a_{21} \nabla u_{h,1} \cdot \mathbf{n} - a_{22} \nabla \tilde{u}_{\tau,h,2} \cdot \mathbf{n}) I_\tau w dl = \int_{\partial\tau} \tilde{g}_{\tau,2} I_\tau w dl, \quad (4.14)$$

where

$$\tilde{g}_{\tau,2} = \sum_{\zeta \in v(\tau)} (F_{\tau,2}^\zeta - Q_{\tau,2}^\zeta) P_\zeta \psi_\zeta, \quad (4.15)$$

where $F_{\tau,2}^\zeta = (f_2, \phi_\zeta)_\tau$ and $Q_{\tau,2}^\zeta = \sum_{j=1}^2 a_{2j}^\tau(u_{h,j}, \phi_\zeta)$ (i.e., the inner product and bilinear form given in (3.1) restricted to the element τ for the second component) and P_ζ are polynomials of degree two in τ satisfying (4.12).

4.2. Existence and uniqueness of the local problems

The following lemma establishes the existence and uniqueness of the solutions to those two local problems.

Lemma 4.1. *The Neumann boundary value problems (4.9)-(4.10) and (4.13)-(4.14) have unique solutions up to a constant and*

$$\int_{\partial\tau} (-\mathbf{a}_{ii}\nabla\tilde{u}_{\tau,h,i} - \mathbf{a}_{ij}\nabla u_{h,j}) \cdot \mathbf{n} \, dl = \int_{\partial\tau} -\sum_{k=1}^2 \mathbf{a}_{ik}\nabla u_k \cdot \mathbf{n} \, dl, \quad (4.16)$$

where $j = 3 - i$ with $i = 1, 2$.

Proof. The existence and uniqueness of the solution are established by verifying the compatibility condition [13]. Noticing that $\sum_{\zeta \in v(\tau)} \phi_\zeta = 1$, and $\nabla(\sum_{\zeta \in v(\tau)} \phi_\zeta) = \mathbf{0}$, simple calculation confirms the compatibility condition

$$\int_{\partial\tau} \tilde{g}_{\tau,i} \, dl = \sum_{\zeta \in v(\tau)} F_{\tau,i}^\zeta - Q_{\tau,i}^\zeta = \int_\tau f_i \, dA. \quad (4.17)$$

Using (2.3) and Divergence theorem, we obtain the identity

$$\int_{\partial\tau} (-\mathbf{a}_{ii}\nabla\tilde{u}_{\tau,h,i} - \mathbf{a}_{ij}\nabla u_{h,j}) \cdot \mathbf{n} \, dl = \int_{\partial\tau} \tilde{g}_{\tau,i} \, dl = \int_\tau f_i \, dA = \int_{\partial\tau} -\sum_{k=1}^2 \mathbf{a}_{ik}\nabla u_k \cdot \mathbf{n} \, dl.$$

This completes the proof. $\square$

4.3. Elemental systems

The elemental systems consist of the systems derived from each step described in Section 4.1. Generally the local problem (4.9) and (4.13) can be written as linear algebra system of the form $\tilde{\mathbf{A}}\tilde{\boldsymbol{\alpha}} = \tilde{\mathbf{f}}$ where $\tilde{\mathbf{A}}$ is 3-by-3 for triangular element. If the element is adjacent to the boundary of Ω , then the systems should be modified accordingly (shown in the next section) to take into account the global boundary conditions.

To derive the elemental systems, replacing $I_\tau w$ in (4.9) and (4.13) by the basis of $V^0(\tau)$, rearranging the terms yields

$$-\int_{\partial t_\zeta} \mathbf{a}_{ii}\nabla\tilde{u}_{\tau,h,i} \cdot \mathbf{n} \, dl = \int_{t_\zeta} f_i \, dA + \int_{\partial t_\zeta} \mathbf{a}_{ij}\nabla u_{h,j} \cdot \mathbf{n} \, dl, \quad \forall \zeta \in v(\tau), \quad (4.18)$$

where $j = 3 - i$ with $i = 1, 2$.

With this in mind, we set $\tilde{u}_{\tau,h,i} = \sum_{\zeta \in v(\tau)} \tilde{\alpha}_{\zeta,i} \phi_\zeta$. Using the boundary condition (4.10) for $i = 1$ and (4.14) for $i = 2$, (4.18) yields the system $\tilde{\mathbf{A}}^i \tilde{\boldsymbol{\alpha}}^i = \tilde{\mathbf{f}}^i$ with

entries defined by

$$\begin{aligned}\widetilde{A}_{\zeta\eta}^i &= - \int_{\partial C^\zeta \cap \tau} \mathbf{a}_{ii} \nabla \phi_\eta \cdot \mathbf{n} \, dl, \\ \widetilde{f}_\zeta^i &= \int_{\tau_\zeta} f_i \, dA - F_\tau^{\zeta,i} + Q_\tau^{\zeta,i} + \int_{\partial C^\zeta \cap \tau} \mathbf{a}_{ij} \nabla u_{h,j} \cdot \mathbf{n} \, dl,\end{aligned}\tag{4.19}$$

where ϕ_η is the usual linear basis function (see (3.2)) for $\zeta, \eta \in v(\tau)$. Notice that (4.19) gives a 3-by-3 linear systems for the triangular finite element.

4.4. Recovering the stresses

Once the solution to (4.18) is found, we compute the stress components through (2.2) on each of the control volume edges by direct calculations. Generally the postprocessed stresses are

$$\begin{aligned}\widetilde{\sigma}_{\tau,h,11} &= (\lambda + \mu) \frac{\partial \widetilde{u}_{\tau,h,1}}{\partial x_1} + \lambda \frac{\partial u_{\tau,h,2}}{\partial x_2}, \quad \widetilde{\sigma}_{\tau,h,12} = \frac{\mu}{2} \left(\frac{\partial \widetilde{u}_{\tau,h,1}}{\partial x_2} + \frac{\partial u_{\tau,h,2}}{\partial x_1} \right), \\ \widetilde{\sigma}_{\tau,h,21} &= \frac{\mu}{2} \left(\frac{\partial u_{\tau,h,1}}{\partial x_2} + \frac{\partial \widetilde{u}_{\tau,h,2}}{\partial x_1} \right), \quad \widetilde{\sigma}_{\tau,h,22} = (\lambda + \mu) \frac{\partial \widetilde{u}_{\tau,h,2}}{\partial x_2} + \lambda \frac{\partial u_{\tau,h,1}}{\partial x_1},\end{aligned}\tag{4.20}$$

where $u_{\tau,h,i}$ is $u_{h,i}$ restricted to τ for $i = 1, 2$.

4.5. Conservation on the global boundary

In this section we cover the case of the control volumes that are adjacent to the global boundary of the domain Ω . In the case of a Neumann boundary condition, the value is just assigned to the vertex. Conservation must be computed on the control volume associated with a boundary vertex if a Dirichlet boundary condition is specified. Figure 2 shows a typical example of a control volume associated with a global boundary vertex. It should be noted that some of the stresses have already been computed from the control volumes associated with interior vertices. This leaves only the stress to be computed on control volume edges that overlap the global boundary or control volume edges coincident with the global boundary. For example, we must find σ_{21}^{zx} and σ_{21}^{yz} . In order to do this, we compute the stresses i.e.,

$$\sigma_{21}^{zx} = (\mathbf{a}_{21} \nabla u_{h,1} + \mathbf{a}_{22} \nabla u_{h,2}) \cdot \mathbf{n}_x \quad \text{and} \quad \sigma_{21}^{yz} = (\mathbf{a}_{21} \nabla u_{h,1} + \mathbf{a}_{22} \nabla u_{h,2}) \cdot \mathbf{n}_y,$$

where $\mathbf{n}_x$ and $\mathbf{n}_y$ are normals for E_{zx} and E_{zy} respectively (see Figure 2). Here $u_{h,1}$ and $u_{h,2}$ are from the Galerkin finite element solution $\mathbf{u}_h$ coming from (3.4). Once

the stresses are found, conservation is computed on the control volume edges coinciding with the global boundary. The conservation is given by

$$\int_{E_z} \sigma_{h,22}^z dl = \int_{E_{zx}} \sigma_{21}^{zx} dl - \int_{E_{yz}} \sigma_{21}^{yz} dl - \int_{E_c} \sigma_{22,h}^i dl + (f, 1)_{C^z},$$

where $\sigma_{h,22}^z$ is the flux on E_z (E_z is the intersection of the global boundary and the boundary of the control volume) and $\sigma_{h,22}^i$ is the known stress on $E_c = C^z \setminus \{E_z \cup E_{zx} \cup E_{yz}\}$ from previous calculation on the interior vertices.

Figure 2: Example for triangular boundary element. Here the thicker line represents a global boundary and the dashed lines represent the continuing mesh.

5. An error analysis for the two-step postprocessing

In this section, we focus on establishing a convergence property for the post-processed solution $\tilde{\mathbf{u}}_{\tau,h} = (\tilde{\mathbf{u}}_{\tau,h,1}, \tilde{\mathbf{u}}_{\tau,h,2})$ in H^1 semi-norm by following the analysis framework in [6]. Here the corresponding induced H^1 semi-norm for a function $\mathbf{v} = (v_1, v_2)$ is denoted as

$$|\mathbf{v}|_{H^1} = \|\nabla \mathbf{v}\|_{L^2} = \sqrt{\sum_{i=1}^2 \|\nabla v_i\|_{L^2}^2} = \sqrt{\sum_{i=1}^2 |v_i|_{H^1}^2}. \quad (5.1)$$

We estimate the desired error for $\tilde{\mathbf{u}}_{\tau,h}$ as follows and we start with some properties and then present the theorem for the error.

Lemma 5.1. Let $\mathbf{u} = (u_1, u_2)$ be the true solution of (2.3) and $\mathbf{u}_h = (u_{h,1}, u_{h,2})$ be the solution to (3.4). For $i = 1, 2$, assume $w_i \in V(\tau)$ and $\widetilde{g}_{\tau,i}$ is defined in (4.11) and in (4.15). Then respectively for $i = 1, 2$,

$$\int_{\partial\tau} \left(\sum_{j=1}^2 \mathbf{a}_{ij} \nabla u_j \cdot \mathbf{n} + \widetilde{g}_{\tau,i} \right) w_i \, dl = \sum_{j=1}^2 a_{ij}^\tau (u_j - u_{h,j}, w_i) \quad (5.2)$$

and

$$\int_{\partial\tau} \widetilde{g}_{\tau,i} (w_i - I_\tau w_i) \, dl = 0. \quad (5.3)$$

Proof. Since $w_i \in V(\tau)$, we write $w_i = \sum_{\eta \in \mathcal{V}(\tau)} w_i(\eta) \phi_\eta$. Using (4.12), we calculate

$$\begin{aligned} \int_{\partial\tau} \widetilde{g}_{\tau,i} w_i \, dl &= \int_{\partial\tau} \widetilde{g}_{\tau,i} \left(\sum_{\eta \in \mathcal{V}(\tau)} w_i(\eta) \phi_\eta \right) \, dl = \sum_{\eta \in \mathcal{V}(\tau)} w_i(\eta) \int_{\partial\tau} \widetilde{g}_{\tau,i} \phi_\eta \, dl \\ &= \sum_{\xi \in \mathcal{V}(\tau)} (F_{\tau,i}^\xi - Q_{\tau,i}^\xi) \sum_{\eta \in \mathcal{V}(\tau)} w_i(\eta) \int_{\partial\tau} P_\xi \phi_\eta \, dl \\ &= \sum_{\eta \in \mathcal{V}(\tau)} w_i(\eta) F_{\tau,i}^\eta - \sum_{\eta \in \mathcal{V}(\tau)} w_i(\eta) Q_{\tau,i}^\eta \\ &= \sum_{\eta \in \mathcal{V}(\tau)} w_i(\eta) (f_i, \phi_\eta)_\tau - \sum_{\eta \in \mathcal{V}(\tau)} w_i(\eta) \sum_{j=1}^2 a_{ij}^\tau (u_{h,j}, \phi_\eta) \\ &= (f_i, w_i)_\tau - \sum_{j=1}^2 a_{ij}^\tau (u_{h,j}, w_i). \end{aligned}$$

Now by multiplying the PDE in (2.3) by the test function $w_i \in V(\tau)$ and using integration by parts, we obtain

$$\sum_{j=1}^2 a_{ij}^\tau (u_j, w_i) = (f_i, w_i)_\tau + \int_{\partial\tau} \left(\sum_{j=1}^2 \mathbf{a}_{ij} \nabla u_j \right) \cdot \mathbf{n} w_i \, dl. \quad (5.4)$$

By using (5.4) and (2.3), we then have

$$\begin{aligned} \int_{\partial\tau} \left(\sum_{j=1}^2 \mathbf{a}_{ij} \nabla u_j \cdot \mathbf{n} + \widetilde{g}_{\tau,i} \right) w_i \, dl &= \sum_{j=1}^2 a_{ij}^\tau (u_j, w_i) - (f_i, w_i)_\tau + (f_i, w_i)_\tau - \sum_{j=1}^2 a_{ij}^\tau (u_{h,j}, w_i) \\ &= \sum_{j=1}^2 a_{ij}^\tau (u_j - u_{h,j}, w_i), \end{aligned}$$

which establishes (5.2). Similarly since $I_\tau w_i = \sum_{\eta \in v(\tau)} w_i(\eta) \psi_\eta$, we have

$$\begin{aligned}
\int_{\partial\tau} \widetilde{g}_{\tau,i} I_\tau w_i \, dl &= \int_{\partial\tau} \widetilde{g}_{\tau,i} \sum_{\eta \in v(\tau)} w_i(\eta) \psi_\eta \, dl = \sum_{\eta \in v(\tau)} w_i(\eta) \int_{\partial\tau} \widetilde{g}_{\tau,i} \psi_\eta \, dl \\
&= \sum_{\xi \in v(\tau)} (F_{\tau,i}^\xi - Q_{\tau,i}^\xi) \sum_{\eta \in v(\tau)} w_i(\eta) \int_{\partial\tau} P_\xi \psi_\eta \, dl \\
&= \sum_{\eta \in v(\tau)} w_i(\eta) F_{\tau,i}^\eta - \sum_{\eta \in v(\tau)} w_i(\eta) Q_{\tau,i}^\eta \\
&= \sum_{\eta \in v(\tau)} w_i(\eta) (f_i, \phi_\eta)_\tau - \sum_{\eta \in v(\tau)} w_i(\eta) \sum_{j=1}^2 a_{ij}^\tau(u_{h,j}, \phi_\eta) \\
&= (f_i, w_i)_\tau - \sum_{j=1}^2 a_{ij}^\tau(u_{h,j}, w_i).
\end{aligned}$$

Therefore

$$\int_{\partial\tau} \widetilde{g}_{\tau,i} (w_i - I_\tau w_i) \, dl = 0,$$

which completes the proof. $\square$

Now we are ready to present and prove the lemma and theorem for the error of the postprocessed solution $\widetilde{\mathbf{u}}_{\tau,h}$ in H^1 semi-norm. The tools for the proof mainly consist of triangle inequality, integration by parts, divergence theorem, and subtracting-adding a term.

Lemma 5.2. *Let $\mathbf{u}$ be the true solution of (2.3) and $\widetilde{\mathbf{u}}_{\tau,h}$ be the postprocessed solution satisfying (4.9) and (4.13) on the element τ . For sufficiently smooth and bounded coefficient matrix $\mathbf{a}_{ij}$ with $i, j = 1, 2$ in τ and sufficiently small h_τ , we have the error estimate*

$$\|\nabla(\mathbf{u} - \widetilde{\mathbf{u}}_{\tau,h})\|_{L^2(\tau)} \leq C_1 R_{1,\tau} + C_2 R_{2,\tau} + C_3 R_{3,\tau},$$

where

$$\begin{aligned}
R_{1,\tau} &= \|\nabla(\mathbf{u} - \mathbf{w})\|_{L^2(\tau)}, \\
R_{2,\tau} &= h_\tau (\|\mathbf{f}\|_{L^2(\tau)} + \|\nabla \mathbf{w}\|_{L^2(\tau)} + \|\mathbf{L} \mathbf{u}_h\|_{L^2(\tau)}), \\
R_{3,\tau} &= \|\nabla(\mathbf{u} - \mathbf{u}_h)\|_{L^2(\tau)},
\end{aligned}$$

and $\mathbf{w} = (w_1, w_2)$ with $w_1, w_2 \in V(\tau)$, $\mathbf{L} \mathbf{u}_h = (\nabla \cdot (\mathbf{a}_{12} \nabla u_{h,2}), \nabla \cdot (\mathbf{a}_{21} \nabla u_{h,1}))$, and C_1, C_2, C_3 are constants independent of h_τ .

Proof. Setting $\mathbf{e}_{\tau,h} = \mathbf{w} - \widetilde{\mathbf{u}}_{\tau,h}$, triangle inequality gives

$$\|\nabla(\mathbf{u} - \widetilde{\mathbf{u}}_{\tau,h})\|_{L^2(\tau)} \leq \|\nabla(\mathbf{u} - \mathbf{w})\|_{L^2(\tau)} + \|\nabla(\mathbf{w} - \widetilde{\mathbf{u}}_{\tau,h})\|_{L^2(\tau)} = R_{1,\tau} + \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)}. \quad (5.5)$$

Using (5.4), $\mathbf{e}_{\tau,h} = (\mathbf{w} - \mathbf{u}) + (\mathbf{u} - \widetilde{\mathbf{u}}_{\tau,h})$ and $\mathbf{e}_{\tau,h} = \mathbf{e}_{\tau,h} - I_\tau \mathbf{e}_{\tau,h} + I_\tau \mathbf{e}_{\tau,h}$, we then get

$$a_{\min,\tau} \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)}^2 \leq \sum_{i=1}^2 \int_\tau \mathbf{a}_{ii} \nabla e_{\tau,h,i} \cdot \nabla e_{\tau,h,i} \, dA = J_1 + J_2 + J_3, \quad (5.6)$$

where

$$\begin{aligned} J_1 &= \sum_{i=1}^2 \int_\tau \mathbf{a}_{ii} \nabla(w_i - u_i) \cdot \nabla e_{\tau,h,i} \, dA, \\ J_2 &= \sum_{i=1}^2 \int_\tau f_i(e_{\tau,h,i} - I_\tau e_{\tau,h,i}) \, dA, \\ J_3 &= \sum_{i=1}^2 \left(\int_\tau f_i I_\tau e_{\tau,h,i} \, dA + \int_{\partial\tau} \left(\sum_{j=1}^2 \mathbf{a}_{ij} \nabla u_j \right) \cdot \mathbf{n} e_{\tau,h,i} \, dl \right. \\ &\quad \left. - \int_\tau \mathbf{a}_{ij} \nabla u_j \cdot \nabla e_{\tau,h,i} \, dA - \int_\tau \mathbf{a}_{ii} \nabla \widetilde{u}_{\tau,h,i} \cdot \nabla e_{\tau,h,i} \, dA \right), \end{aligned}$$

with $j = 3 - i$ (same as follows), and $a_{\min,\tau} = \min\{\lambda + \mu, \frac{\mu}{2}\}$, where the minimum is taken over τ . Here, the Lamé coefficients λ and μ satisfy $\mu > 0$ and $\lambda + \frac{2}{3}\mu \geq 0$ due to thermodynamic stability; see [12]. The boundedness of $\mathbf{a}_{ii}$ for $i = 1, 2$ and the Cauchy-Schwarz inequality give

$$J_1 \leq a_{\max,\tau} \sum_{i=1}^2 \|\nabla(w_i - u_i)\|_{L^2(\tau)} \|\nabla e_{\tau,h,i}\|_{L^2(\tau)} = C_1 R_{1,\tau} \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)}, \quad (5.7)$$

where $a_{\max,\tau} = \max\{\lambda + \mu, \frac{\mu}{2}\}$, where the maximum is taken over τ . Also by Cauchy-Schwarz inequality and (4.8), we get

$$J_2 \leq \sum_{i=1}^2 \|f_i\|_{L^2(\tau)} \|e_{\tau,h,i} - I_\tau e_{\tau,h,i}\|_{L^2(\tau)} \leq Ch_\tau \|f\|_{L^2(\tau)} \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)}. \quad (5.8)$$

To estimate J_3 , we first use (4.9), (4.13), and Green's formula to write the first

summation term in J_3 as

$$\begin{aligned}
\sum_{i=1}^2 \int_{\tau} f_i I_{\tau} e_{\tau,h,i} \, dA &= \sum_{i=1}^2 \sum_{\xi \in \mathcal{V}(\tau)} \int_{\partial t_{\xi}} (-\mathbf{a}_{ii} \nabla \tilde{u}_{\tau,h,i} - \mathbf{a}_{ij} \nabla u_{h,j}) \cdot \mathbf{n} \, I_{\tau} e_{\tau,h,i} \, dl \\
&= \sum_{i=1}^2 \sum_{\xi \in \mathcal{V}(\tau)} \int_{I_{\xi}} \nabla \cdot (-\mathbf{a}_{ii} \nabla \tilde{u}_{\tau,h,i} - \mathbf{a}_{ij} \nabla u_{h,j}) \, I_{\tau} e_{\tau,h,i} \, dA \quad (5.9) \\
&= \sum_{i=1}^2 \int_{\tau} \nabla \cdot (-\mathbf{a}_{ii} \nabla \tilde{u}_{\tau,h,i} - \mathbf{a}_{ij} \nabla u_{h,j}) \, I_{\tau} e_{\tau,h,i} \, dA.
\end{aligned}$$

Next, from (5.2), (5.3), (4.10), and (4.14), the second summation term in J_3 is expressed as

$$\begin{aligned}
\sum_{i=1}^2 \int_{\partial \tau} \left(\sum_{j=1}^2 \mathbf{a}_{ij} \nabla u_j \cdot \mathbf{n} \right) e_{\tau,h,i} \, dl &= \sum_{i=1}^2 \left(\sum_{j=1}^2 \mathbf{a}_{ij}^{\tau} (u_j - u_{h,j}, w_i) - \int_{\partial \tau} \tilde{g}_{\tau,i} I_{\tau} e_{\tau,h,i} \, dl \right) \\
&= a_{\tau}(\mathbf{u} - \mathbf{u}_h, \mathbf{e}_{\tau,h}) + \sum_{i=1}^2 \int_{\partial \tau} (\mathbf{a}_{ii} \nabla \tilde{u}_{\tau,h,i} + \mathbf{a}_{ij} \nabla u_{h,j}) \cdot \mathbf{n} \, I_{\tau} e_{\tau,h,i} \, dl \quad (5.10)
\end{aligned}$$

Furthermore, we apply integration by parts to the last summation term in J_3 to get

$$\begin{aligned}
-\sum_{i=1}^2 \int_{\tau} \mathbf{a}_{ii} \nabla \tilde{u}_{\tau,h,i} \cdot \nabla e_{\tau,h,i} \, dA &= \sum_{i=1}^2 \int_{\tau} \nabla \cdot (\mathbf{a}_{ii} \nabla \tilde{u}_{\tau,h,i}) e_{\tau,h,i} \, dA \\
&\quad - \sum_{i=1}^2 \int_{\partial \tau} \mathbf{a}_{ii} \nabla \tilde{u}_{\tau,h,i} \cdot \mathbf{n} e_{\tau,h,i} \, dl. \quad (5.11)
\end{aligned}$$

Putting (5.9), (5.10), and (5.11) back into J_3 gives

$$J_3 = J_{31} + J_{32},$$

where

$$\begin{aligned}
J_{31} &= \sum_{i=1}^2 \int_{\tau} \nabla \cdot (\mathbf{a}_{ii} \nabla \tilde{u}_{\tau,h,i}) (e_{\tau,h,i} - I_{\tau} e_{\tau,h,i}) \, dA, \\
J_{32} &= a_{\tau}(\mathbf{u} - \mathbf{u}_h, \mathbf{e}_{\tau,h}) - \sum_{i=1}^2 \int_{\tau} \mathbf{a}_{ij} \nabla u_j \cdot \nabla e_{\tau,h,i} \, dA - \sum_{i=1}^2 \int_{\partial \tau} \mathbf{a}_{ii} \nabla \tilde{u}_{\tau,h,i} \cdot \mathbf{n} e_{\tau,h,i} \, dl \\
&\quad - \sum_{i=1}^2 \int_{\tau} \nabla \cdot (\mathbf{a}_{ij} \nabla u_{h,j}) \, I_{\tau} e_{\tau,h,i} \, dA + \sum_{i=1}^2 \int_{\partial \tau} (\mathbf{a}_{ii} \nabla \tilde{u}_{\tau,h,i} + \mathbf{a}_{ij} \nabla u_{h,j}) \cdot \mathbf{n} \, I_{\tau} e_{\tau,h,i} \, dl.
\end{aligned}$$

Now, for J_{31} , using (4.8), Cauchy-Schwarz inequality and triangle inequality give

$$\begin{aligned}
J_{31} &\leq \sum_{i=1}^2 a_{\infty} \|\nabla \widetilde{u}_{\tau,h,i}\|_{L^2(\tau)} \|e_{\tau,h,i} - I_{\tau} e_{\tau,h,i}\|_{L^2(\tau)} \\
&\leq Ch_{\tau} \sum_{i=1}^2 \|\nabla \widetilde{u}_{\tau,h,i}\|_{L^2(\tau)} \|\nabla e_{\tau,h,i}\|_{L^2(\tau)} \\
&\leq Ch_{\tau} \sum_{i=1}^2 \|\nabla w_i\|_{L^2(\tau)} \|\nabla e_{\tau,h,i}\|_{L^2(\tau)} + Ch_{\tau} \sum_{i=1}^2 \|\nabla e_{\tau,h,i}\|_{L^2(\tau)}^2 \\
&\leq Ch_{\tau} \|\nabla w\|_{L^2(\tau)} \|\nabla e_{\tau,h,i}\|_{L^2(\tau)} + Ch_{\tau} \|\nabla e_{\tau,h,i}\|_{L^2(\tau)}^2,
\end{aligned} \tag{5.12}$$

where $a_{\infty} = \max \{ \|\nabla(\lambda + \mu)\|_{L^{\infty}(\tau)}, \|\nabla(\frac{\mu}{2})\|_{L^{\infty}(\tau)} \}$. Obviously, if λ and μ are constants, then $a_{\infty} = 0$ and $J_{31} = 0$. Expanding the first term in J_{32} , applying integration by parts, and using (2.3) give

$$\begin{aligned}
a_{\tau}(\mathbf{u} - \mathbf{u}_h, \mathbf{e}_{\tau,h}) - \sum_{i=1}^2 \int_{\tau} \mathbf{a}_{ij} \nabla u_j \cdot \nabla e_{\tau,h,i} \, dA &= \sum_{i=1}^2 \int_{\tau} \mathbf{a}_{ii} \nabla(u_i - u_{h,i}) \cdot \nabla e_{\tau,h,i} \, dA \\
&\quad + \sum_{i=1}^2 \int_{\tau} \nabla \cdot (\mathbf{a}_{ij} \nabla u_{h,j}) e_{\tau,h,i} \, dA - \sum_{i=1}^2 \int_{\partial\tau} \mathbf{a}_{ij} \nabla u_{h,j} \cdot \mathbf{n} e_{\tau,h,i} \, dl.
\end{aligned} \tag{5.13}$$

Putting (5.13) back to J_{32} and appropriately rearranging the terms give the following decomposition

$$J_{32} = J_{41} + J_{42} + J_{43},$$

where

$$\begin{aligned}
J_{41} &= \sum_{i=1}^2 \int_{\tau} \mathbf{a}_{ii} \nabla(u_i - u_{h,i}) \cdot \nabla e_{\tau,h,i} \, dA, \\
J_{42} &= \sum_{i=1}^2 \int_{\tau} \nabla \cdot (\mathbf{a}_{ij} \nabla u_{h,j}) (e_{\tau,h,i} - I_{\tau} e_{\tau,h,i}) \, dA, \\
J_{43} &= \sum_{i=1}^2 \int_{\partial\tau} (\mathbf{a}_{ii} \nabla \widetilde{u}_{\tau,h,i} + \mathbf{a}_{ij} \nabla u_{h,j}) \cdot \mathbf{n} (I_{\tau} e_{\tau,h,i} - e_{\tau,h,i}) \, dl,
\end{aligned}$$

Cauchy-Schwarz inequality gives

$$J_{41} \leq a_{\max,\tau} \sum_{i=1}^2 \|\nabla(u_i - u_{h,i})\|_{L^2(\tau)} \|\nabla e_{\tau,h,i}\|_{L^2(\tau)} \leq CR_{3,\tau} \|\nabla e_{\tau,h}\|_{L^2(\tau)}. \tag{5.14}$$

Using Cauchy-Schwarz inequality and (4.8), we get

$$J_{42} \leq \sum_{i=1}^2 Ch_{\tau} \|\nabla \cdot (\mathbf{a}_{ij} \nabla u_{h,j})\|_{L^2(\tau)} \|\nabla e_{\tau,h,i}\|_{L^2(\tau)} \leq Ch_{\tau} \|\mathbf{L}u_h\|_{L^2(\tau)} \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)}. \quad (5.15)$$

Using Lemma 6.1 in [9] and triangle inequality yield

$$\begin{aligned} J_{43} &\leq \sum_{i=1}^2 Ch_{\tau} (a_{\infty} \|\nabla \tilde{u}_{\tau,h,i}\|_{L^2(\tau)} + \|\nabla \cdot (\mathbf{a}_{ij} \nabla u_{h,j})\|_{L^2(\tau)}) \|\nabla e_{\tau,h,i}\|_{L^2(\tau)} \\ &\leq Ch_{\tau} (\|\nabla \mathbf{w}\|_{L^2(\tau)} + \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)} + \|\mathbf{L}u_h\|_{L^2(\tau)}) \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)}. \end{aligned} \quad (5.16)$$

Based on the above derivation, (5.6) can be rewritten as

$$a_{\min,\tau} \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)}^2 \leq J_1 + J_2 + J_{31} + J_{41} + J_{42} + J_{43}.$$

We collect the like-terms of all the estimates (5.7) for J_1 , (5.8) for J_2 , (5.12) for J_{31} , (5.14) for J_{41} , (5.15) for J_{42} , and (5.16) for J_{43} to give

$$a_{\min,\tau} \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)}^2 \leq (C_1 R_{1,\tau} + C_2 R_{2,\tau} + C_3 R_{3,\tau}) \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)} + Ch_{\tau} \|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)}^2.$$

Choosing sufficiently small h_{τ} , we can combine the last term on the right hand side of the last inequality to the left hand side to get

$$\|\nabla \mathbf{e}_{\tau,h}\|_{L^2(\tau)} \leq (C_1 R_{1,\tau} + C_2 R_{2,\tau} + C_3 R_{3,\tau}),$$

which is then combined with (5.5) to give the desired result. $\square$

Theorem 5.1. *Let $\mathbf{u}$ be the true solution of (2.3) and $\tilde{\mathbf{u}}_{\tau,h}$ be the postprocessed solution satisfying (4.9) and (4.13), then we have*

$$\|\nabla(\mathbf{u} - \tilde{\mathbf{u}}_{\tau,h})\|_{L^2(\Omega)} \leq Ch, \quad (5.17)$$

where C is a constant independent of h .

Proof. Let $\mathbf{w}_h$ be the nodal interpolation of $\mathbf{u}$ such that $\mathbf{w}_h|_{\tau} = \mathbf{w}$ with $w_1, w_2 \in V(\tau)$. Using Lemma 5.2, we have

$$\begin{aligned} \|\nabla(\mathbf{u} - \tilde{\mathbf{u}}_{\tau,h})\|_{L^2(\Omega)} &= \sum_{\tau} \|\nabla(\mathbf{u} - \tilde{\mathbf{u}}_{\tau,h})\|_{L^2(\tau)}^2 \\ &\leq C \left(\sum_{\tau} R_{1,\tau}^2 + \sum_{\tau} R_{2,\tau}^2 + \sum_{\tau} R_{3,\tau}^2 \right) \\ &= C \left(\|\nabla(\mathbf{u} - \mathbf{w}_h)\|_{L^2(\Omega)}^2 + h^2 (\|\mathbf{f}\|_{L^2(\Omega)}^2 + \|\nabla \mathbf{w}\|_{L^2(\tau)}^2 \right. \\ &\quad \left. + \|\mathbf{L}u_h\|_{L^2(\Omega)}^2) + \|\nabla(\mathbf{u} - \mathbf{u}_h)\|_{L^2(\Omega)}^2 \right). \end{aligned}$$

Moreover, the continuous Galerkin FEM solution $\mathbf{u}$ satisfies $\|\mathbf{u} - \mathbf{u}_h\|_{H^1(\Omega)} \leq Ch\|\mathbf{u}\|_{H^2(\Omega)}$; see [12]. Also, interpolation $\mathbf{w}_h$ has a standard property that it converges linearly to $\mathbf{u}$ with respect to h . All these combined give the desired result. $\square$

Through the definition in (5.1) and (4.20), we note that Theorem 5.1 immediately implies that $\|\nabla(u_i - \tilde{u}_{h,i})\|_{L^2(\Omega)} \leq Ch$ for $i = 1, 2$.

6. Numerical examples

Numerical examples are presented in this section to verify that the finite element formulation and the two-step postprocessing perform as expected. Comparisons are also made between the postprocessed stresses, computed stresses by the finite element solution and basis functions, and the true stresses when they are known. In each example we solve (2.1) by the continuous Galerkin FEM using triangular elements followed with the two-step postprocessing technique as described in previous sections.

6.1. Example 1

The first example verifies that the stress is captured correctly when the displacement is a cubic polynomial in one spatial direction and the stress quadratically depends on the spatial position. In this case the example is derived from the desired solution $\mathbf{u} = (0.1x_1^3, 0)$ on $\Omega = [0, 1] \times [0, 1]$. Here $\mathbf{f} = (-0.6(\lambda + \mu)x_1, 0)$ and the problem is treated as fully Dirichlet with the values set from the exact solution. The stress components are given by

$$\begin{cases} \sigma_{11} = 0.3(\lambda + \mu)x_1^2 \\ \sigma_{21} = \sigma_{12} = 0 \\ \sigma_{22} = 0.3\lambda x_1^2. \end{cases}$$

For simplicity $\lambda = \mu = 1$ is used in this example on a uniform triangular mesh with 72 elements and 98 degrees of freedom. As expected, the finite element solution is not exact. The stress computed from the postprocessing has an associated error as well. The original finite element mesh and the deformation (with a scale factor 1) are plotted in Figure 3. The example was also computed on an arbitrary triangular mesh with 168 elements and 202 degrees of freedom. The mesh and resulting deformation (with a scale factor 1) are shown in Figure 4. The H^1 semi-norm errors of the postprocessed solution $\tilde{u}_{h,1}$ and $\tilde{u}_{h,2}$ are shown in Figure 5, which confirms the linear convergence order analyzed in Section 5. The second

Figure 3: The original finite element mesh is shown on the left and the deformed mesh is shown on the right for Example 1 with a uniform triangular mesh.

Figure 4: The original finite element mesh is shown on the left and the deformed mesh is shown on the right for Example 1 with an arbitrary triangular mesh.

plot in Figure 5 tells that numerically the gradient of the postprocessed solution $\tilde{u}_{h,2}$ is the same the gradient of the true solution u_2 .

A representative stress and the $L^2(\Omega)$ error of the postprocessed stress $\tilde{\sigma}_{11,h}$ are shown in Figure 6. In this case, the stress is dependent on x_1 only so it is the same for any fixed x_2 . The postprocessed solution is faithful to this property, so the stress σ_{11} is plotted as a single cross section of the triangular finite element mesh. As expected there is a numerical error associated with the postprocessed stress, but there is a good agreement with the true solution.

6.2. Example 2

Our second example is more general for the numerical investigation of the error analysis presented in Section 5. We consider (2.1) with an exact solution $\mathbf{u} = (u_1, u_2) = (x_2^2 e^{-x_1}, x_1^2 x_2^3)$ and coefficients $\lambda = 1, \mu = 2$. From (2.1), we can calculate $\mathbf{f} = (f_1, f_2)$ where $f_1 = -3x_2^2 e^{-x_1} - 12x_1 x_2^2 - 2e^{-x_1}$ and $f_2 = 4x_2 e^{-x_1} -$

Figure 5: The left plot shows $H^1(\Omega)$ semi-norm error of the postprocessed solution $\tilde{u}_{h,1}$ while the right plot shows $H^1(\Omega)$ semi-norm error of the postprocessed solution $\tilde{u}_{h,2}$.

Figure 6: The left plot shows postprocessed stress for Example 1 with a uniform triangular mesh where a cross section of the postprocessed stress is plotted along with the exact stress when the mesh size is 0.1. The right plot shows $L^2(\Omega)$ error of the postprocessed stress $\tilde{\sigma}_{11,h}$ plotted against the mesh size for Example 1 with a uniform triangular mesh. As expected, the convergence is linear.

$2x_2^3 - 18x_1^2x_2$. We impose full Dirichlet boundary condition in this example. The H^1 semi-norm errors of the postprocessed solution $\tilde{u}_{h,1}$ and $\tilde{u}_{h,2}$ are shown in Figure 7. The convergence orders of these errors for both $\tilde{u}_{h,1}$ and $\tilde{u}_{h,2}$ are approximately 1, which confirm the convergence analysis presented in Section 5.

Now we use (4.20) in Section 4.4 to recover the postprocessed stresses $\tilde{\sigma}_{11,h}, \tilde{\sigma}_{12,h}, \tilde{\sigma}_{21,h}$, and $\tilde{\sigma}_{22,h}$. These postprocessed stresses are locally conservative. With a simple

Figure 7: The left plot shows $H^1(\Omega)$ semi-norm error of the postprocessed solution $\tilde{u}_{h,1}$ while the right plot shows $H^1(\Omega)$ semi-norm error of the postprocessed solution $\tilde{u}_{h,2}$.

calculation from the true solutions, the true stresses are

$$\begin{cases} \sigma_{11} = 3x_2^2(x_1^2 - e^{-x_1}) \\ \sigma_{21} = \sigma_{12} = 2x_2e^{-x_1} + 2x_1x_2^3 \\ \sigma_{22} = x_2^2(9x_1^2 - e^{-x_1}). \end{cases}$$

The $L^2(\Omega)$ norm errors of these postprocessed stresses are shown in Figure 8. As expected, these convergence orders are linear.

6.3. Example 3

A beam fixed at one end and simply supported at the other is considered in the third example. Again (2.1) is solved by continuous Galerkin FEM followed with the two-step postprocessing technique in $\Omega = [0, 5] \times [0, 1]$. Let $\mathbf{f} = (0, -1)$ and the boundary condition is

$$\begin{cases} u_1(0, x_2) = u_2(0, x_2) = u_2(5, x_2) = 0 \\ g_{N,2}(x_1, 1) = \frac{x_1(5 - x_1)}{5} \\ g_{N,i} = 0 \text{ elsewhere, } i = 1, 2. \end{cases}$$

Here 20-by-8 triangular finite elements are used with elasticity modulus $E = 1000$ and Poisson's ratio $\nu = 0.3$. This leads to $\lambda \approx 576.9$ and $\mu \approx 769.2$ through the relations

$$\lambda = \frac{Ev}{(1 + \nu)(1 - 2\nu)}$$

$$\mu = \frac{E}{1 + \nu}.$$

Figure 8: These plots show the $L^2(\Omega)$ errors of the postprocessed stresses for Example 2.

The deformation with a scale factor 1 of the beam is plotted in Figure 9.

To demonstrate one of the advantages of the postprocessing, we next plot the local conservation errors defined as follows where the stresses are directly calculated through the basis functions and the Galerkin FEM solutions. We define the local conservation errors as

$$\text{error}_i = \int_{\partial C^z} (\sigma_h \cdot \mathbf{n})_i \, dl + \int_{C^z} f_i \, dA \quad \text{for } i = 1, 2, \quad (6.1)$$

for a control volume C^z . Figure 10 shows these errors, which are not negligible for any problem for which local conservation needs to be satisfied. Since the postprocessing is constructed in such a way that local conservation is satisfied for control volumes surrounding vertices in the finite element mesh, $\text{error}_1 = 0$ and $\text{error}_2 = 0$ if we replace the directly calculated stresses with the postprocessed stresses in (6.1).

Figure 9: The original finite element mesh is shown at the top and the deformed mesh is shown at the bottom for Example 3.

6.4. Example 4

A final example with a singularity is provided, where an L-shaped domain Ω is described by the corner points at $(-1, -1)$, $(0, -2)$, $(2, 0)$, $(0, 2)$, $(-1, 1)$, and $(0, 0)$. The exact solution to the problem with $\mathbf{f} = (0, 0)$ is given in polar coordinates by

$$\begin{aligned} u_r(r, \theta) &= \frac{1}{2\mu} r^\alpha (-(\alpha + 1) \cos((\alpha + 1)\theta) + (C_2 - (\alpha + 1))C_1 \cos((\alpha - 1)\theta)) \\ u_\theta(r, \theta) &= \frac{1}{2\mu} r^\alpha ((\alpha + 1) \sin((\alpha + 1)\theta) + (C_2 + \alpha - 1)C_1 \sin((\alpha - 1)\theta)), \end{aligned} \quad (6.2)$$

with $\alpha \approx 0.544483737$, $C_1 = -\cos((\alpha + 1)0.75\pi) / \cos((\alpha - 1)0.75\pi)$, and $C_2 = 2(\lambda + 2\mu) / (\lambda + \mu)$ [2]. The boundary condition is imposed as fully Dirichlet with values prescribed by the exact solution given above.

We ran this example on a triangular mesh with 467 elements and 514 degrees of freedom. In this example $E = 100000$ and $\nu = 0.3$ is used with the relations

$$\begin{aligned} \lambda &= \frac{E\nu}{(1 + \nu)(1 - 2\nu)} \\ \mu &= \frac{E}{2(1 + \nu)}. \end{aligned}$$

The original finite element mesh and resulting deformation with a scale factor of 3000 are shown in Figure 11.

Figure 10: Local conservation errors associated with computed stresses by the finite element solution and basis functions for Example 3.

Figure 12 shows the conservation errors defined in (6.1) when the stresses are computed directly from the finite element solution and basis functions. Clearly without the two-step postprocessing, the local conservation is not satisfied and the corresponding Galerkin FEM solutions need a postprocessing to obtain local conservation.

7. Conclusion

The two-step postprocessing is presented in the context of a stress recovery technique for the problem of displacement based linear elasticity. Numerical examples are also presented to verify that the method works as expected. It is shown that the convergence order of the postprocessed stress to the true stress is linear with respect to the mesh size. The main advantage of the method is that it has the desirable property of local conservation in the recovered stresses, which is inherent in the governing partial differential equation. This is in contrast to the

Figure 11: The original finite element mesh is shown on the left and the deformation (scaled by 3000) is shown on the right for Example 4.

commonly used stress recovery methods.

There are several directions for future work. The method can be applied in the multiscale context for heterogeneous materials. This is a direct analog of the material presented in [7]. The technique can also be used to include geomechanical coupling in multiphase flow models presented in [6, 7]. Moreover, the technique can be developed into an a posteriori error estimator and we will concentrate on developing such an estimator in our future work.

References

- [1] M. A $\bar{\text{Zuccherato}}$, J. T. O $\bar{\text{Neil}}$, *A posteriori error estimation in finite element analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2000.
- [2] J. A $\bar{\text{Bergs}}$, C. C $\bar{\text{Carstensen}}$, S. A. F $\bar{\text{ried}}$, R. K $\bar{\text{Wentz}}$, *Matlab-implementation of the finite element method in elasticity*, Computing, 69 (2000), p. 2002.
- [3] D. N. A $\bar{\text{Brezis}}$, R. S. F $\bar{\text{ort}}$, R. W $\bar{\text{heeler}}$, *Mixed finite element methods for linear elasticity with weakly imposed symmetry*, Math. Comp., 76 (2007), pp. 1699–1723 (electronic).

Figure 12: Local conservation error associated with computed stresses by the finite element solution and basis functions for Example 4.

- [4] R. E. B $\quad$ A. W $\quad$, *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comp., 44 (1985), pp. 283–301.
- [5] K. B $\quad$, *Finite Element Procedures*, Prentice Hall, 2006.
- [6] L. B $\quad$ V. G $\quad$, *On the application of the continuous galerkin finite element method for conservation problems*, SIAM Journal on Scientific Computing, 35 (2013), pp. A2953–A2975.
- [7] L. B $\quad$, V. G $\quad$, $\quad$ M. P $\quad$, *Application of a conservative, generalized multiscale finite element method to flow models*, Journal of Computational and Applied Mathematics, 260 (2014), pp. 395 – 409.
- [8] C. C $\quad$ S. A. F $\quad$, *Averaging technique for FE—a posteriori error control in elasticity. I. Conforming FEM*, Comput. Methods Appl. Mech. Engrg., 190 (2001), pp. 2483–2498.

- [9] P. C   , *Finite volume methods for elliptic PDE's: a new approach*, M2AN Math. Model. Numer. Anal., 36 (2002), pp. 307–324.
- [10] S.-H. C   Q. L , *Error estimates in L^2 , H^1 and L^∞ in covolume methods for elliptic and parabolic problems: a unified approach*, Math. Comp., 69 (2000), pp. 103–120.
- [11] P. G. C   , *The finite element method for elliptic problems*, vol. 40 of Classics in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.
- [12] A. E   J.-L. G   , *Theory and practice of finite elements*, vol. 159 of Applied Mathematical Sciences, Springer-Verlag, New York, 2004.
- [13] L. C. E   , *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2010.
- [14] T. G   K.-J. B   , *A posteriori error estimation techniques in practical finite element analysis*, Comput. & Structures, 83 (2005), pp. 235–265.
- [15] H. J   H. G. W   , *Application of the finite volume method and unstructured meshes to linear elasticity*, International Journal for Numerical Methods in Engineering, 48 (2000), pp. 267–287.
- [16] C. J   , *Numerical solution of partial differential equations by the finite element method*, Dover Publications Inc., Mineola, NY, 2009. Reprint of the 1987 edition.
- [17] D. J. P   K.-J. B   , *Improved stresses for the 4-node tetrahedral element*, Computers & Structures, 89 (2011), pp. 1265 – 1273.
- [18] D. J. P   K.-J. B   , *The use of nodal point forces to improve element stresses*, Computers & Structures, 89 (2011), pp. 485 – 495.
- [19] B. R   , S. S   , M. F. W   ,   J. W   , *Discontinuous galerkin finite element methods for linear elasticity and quasistatic linear viscoelasticity*, Numerische Mathematik, 95 (2003), pp. 347–376.

- [20] T. S. K.-J. B., *Studies of finite element procedures stress band plots and the evaluation of finite element meshes*, Engineering computations, 3 (1986), pp. 178–191.
- [21] A. T., H. R., C. E. F., G. M. C., *An improved variational method for finite element stress recovery and a posteriori error estimation*, Computer Methods in Applied Mechanics and Engineering, 155 (1998), pp. 15 – 30.
- [22] A. T., H. R., S. M., *A variational method for finite element stress recovery and error estimation*, Computer Methods in Applied Mechanics and Engineering, 111 (1994), pp. 369 – 382.
- [23] J. W. T., *Numerical partial differential equations*, vol. 33 of Texts in Applied Mathematics, Springer-Verlag, New York, 1999. Conservation laws and elliptic equations.
- [24] R. V., *A review of a posteriori error estimation techniques for elasticity problems*, Comput. Methods Appl. Mech. Engrg., 176 (1999), pp. 419–440. New advances in computational methods (Cachan, 1997).
- [25] O. C. Z. J. Z. Z., *A simple error estimator and adaptive procedure for practical engineering analysis*, Internat. J. Numer. Methods Engrg., 24 (1987), pp. 337–357.
- [26] O. C. Z. J. Z. Z., *The superconvergent patch recovery and a posteriori error estimates. I. The recovery technique*, Internat. J. Numer. Methods Engrg., 33 (1992), pp. 1331–1364.
- [27] O. C. Z. J. Z. Z., *The superconvergent patch recovery and a posteriori error estimates. II. Error estimates and adaptivity*, Internat. J. Numer. Methods Engrg., 33 (1992), pp. 1365–1382.