A MIXED METHOD FOR THE BIHARMONIC PROBLEM BASED ON A SYSTEM OF FIRST-ORDER EQUATIONS

EDWIN M. BEHRENS AND JOHNNY GUZMÁN

Abstract. We introduce a new mixed method for the biharmonic problem. The method is based on a formulation where the biharmonic problem is re-written as a system of four first-order equations. A hybrid form of the method is introduced which allows to reduce the globally coupled degrees of freedom to only those associated with Lagrange multipliers which approximate the solution and its derivative at the faces of the triangulation. For $k \geq 1$ a projection of the primal variable error superconverges with order $k+3$ while the error itself converges with order $k+1$ only. This fact is exploited by using local postprocessing techniques that produce new approximations to the primal variable converging with order $k+3$. We provide numerical experiments that validate our theoretical results.

1. Introduction

We consider the biharmonic problem

$$\triangle^2 u = f \quad \text{in } \Omega, \tag{1.1a}$$

$$u = 0 \quad \text{on } \partial\Omega, \tag{1.1b}$$

$$\nabla u \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega, \tag{1.1c}$$

where $\Omega \subset R^d$ is a polyhedral domain and $f \in L^2(\Omega)$. Our method is based on the following formulation of the above problem

$$\begin{aligned} \boldsymbol{q} &= \nabla u, & \underline{\boldsymbol{z}} &= \nabla \boldsymbol{q} & &\text{in } \Omega, \\ \boldsymbol{\sigma} &= \nabla \cdot \underline{\boldsymbol{z}}, & \nabla \cdot \boldsymbol{\sigma} &= f & &\text{in } \Omega, \\ u &= 0, & \boldsymbol{q} \cdot \boldsymbol{n} &= 0 & &\text{on } \partial\Omega. \end{aligned}$$

Following our convention $(\nabla \boldsymbol{q})_{ij} = \partial_{x_j}(q_i)$ for $1 \leq i, j \leq d$ where q_i is the i-th component of $\boldsymbol{q}$. Moreover, $(\nabla \cdot \underline{\boldsymbol{z}})_i = \sum_{j=1}^{d} \partial_{x_j} z_{ij}$ where the z_{ij} is the ij-entry of $\underline{\boldsymbol{z}}$. The method we propose will approximate $u, \boldsymbol{q}, \underline{\boldsymbol{z}}, \boldsymbol{\sigma}$ simultaneously. However, we introduce a hybrid form of the mixed method that will allow us to eliminate all the interior variables locally to obtain a system for the Lagrange multipliers which have domain the interfaces of the triangulation.

There are several mixed methods for the biharmonic problem; see [8, 18, 16, 3]. They are based on introducing the variable $z = \triangle u$ or the variable $\underline{\boldsymbol{z}} = \nabla(\nabla u)$. For example, the Ciarlet and Raviart (C-R) method [8] chooses as unknowns u and $\triangle u$ and obtains a

2000 *Mathematics Subject Classification.* 65M60,65N30,35L65.

Key words and phrases. finite element methods, Kirchoff-Love plate problem, biharmonic problem, mixed methods, hybrid methods.

The first author was partially supported by CONICYT-Chile through the FONDECYT Grant 11070085, and by the Dirección de Investigación of the Universidad Católica de la Santísima Concepción. The second author was partially supported by the National Science Foundation (grant DMS-0914596).

coupled system of Poisson problems. Therefore, piece-wise continuous approximations are used for both variables. The error analysis of the C-R method can be found in [17, 2, 23, 22]. Optimal convergence rates for the approximation to u are obtained however sub-optimal convergence rates are proved for the approximation to $\triangle u$. More precisely, assuming certain L^∞ smoothness assumptions on the derivatives of the solutions, the approximation to $\triangle u$ convergences with rate $h^{k-1/2}$ if piecewise polynomials of degree k are used; see [23]. It is well known that this result is in fact sharp. However, Scholz [22] proved that in fixed sub-domains optimal convergence rates can be recovered.

On the other hand, the Hellan-Hermmann-Johnson (HHJ) method analyzed by Johnson [18] takes as the new unknown $\underline{\boldsymbol{z}} = \nabla(\nabla u)$. The method uses continuous piecewise polynomial approximations to u of degree k and normal-normal continuous symmetric approximations to $\underline{\boldsymbol{z}}$ of degree $k-1$. For this method optimal error estimates are proved for both variables. Moreover, by using the hybrid form of the method one can eliminate the approximation to $\underline{\boldsymbol{z}}$ locally to get a final system involving the approximation to u (which is continuous) and a Lagrange multiplier that approximates the normal components of u on the interfaces of the triangulation. Moreover, one can postprocess the approximate solution locally to get a new approximation to u which converges with order $k+2$; see [15, 24]. We would like to mention that as far as we know optimal estimates have only been proving in two dimensions for the HHJ method.

Recently a hybridizable discontinuous Galerkin (HDG) method was developed by Cockburn et al. [9] for the biharmonic problem based on an HDG method for second-order problems [10]. Similar to the C-R method it has as unknowns u, and $\triangle u$, but in addition it also approximates ∇u and $\nabla \triangle u$. In fact, equal order approximations where used for all the variables. Optimal error estimates were proved for u and ∇u; however, similar to the C-R method, only sub-optimal error estimates where obtained for $\triangle u$. Nonetheless, the approximation to $\triangle u$ converges with order $k + 1/2$ when polynomials of degree k are used which is an improvement of the C-R method which converges with order $k - 1/2$. However, again, L^∞ regularity of higher order derivatives of the exact solution were assumed. A postprocessing technique was used to compute locally a new approximation to u that converges with order $k + 2$ for $k \geq 1$. Also, hybridization of the method was discussed that shows that the only globally coupled degrees of freedom are those of the Lagrange multipliers that approximate u and $\triangle u$ on the interfaces of the triangulation. We would like to mention that HDG methods are similar to mixed methods in their hybrid form; see [13]. In fact, mixed method techniques were used to analyze HDG methods in [14].

In search of a method that would see an improvement on convergence rates as compared to the method in [9], while retaining some of the positive properties of that method we devised the method we present in this paper. Indeed, our method will approximate the second derivatives of u, namely $\underline{\boldsymbol{z}}$, with optimal order $k+1$ while assuming the correct regularity for $\underline{\boldsymbol{z}}$. Moreover, the hybrid form of our method will allow us to eliminate all the variables local to obtain a final system for Lagrange multipliers that approximate u and $\boldsymbol{q}$ on the interfaces of the triangulation. Finally, we also develop a postprocessing technique that produces a new approximation to u that converges with order $k+3$ for $k \geq 1$ compared to order $k+2$ obtained for the method in [9] and the HHJ method.

Our method uses the formulation used by HHJ method, but instead writes the problem as four first-order equations instead of two second-order equations. In fact, our method and the HHJ method achieve optimal convergence rates for the approximation of $\underline{\boldsymbol{z}}$ whereas the C-R method and the method in [9] converge in a sub-optimal way to $\triangle u$.

In two dimensions the biharmonic problem (1.1) is a model of a clamped plate under a vertical load. In a forthcoming paper we extend our methodology to the Reissner-Mindlin plate model, a more complicated plate model.

The paper is organized as follows. In the next section we present our method. In Section 3 we provide error estimates. In the fourth section postprocessing is discussed. In Section 5 the hybrid form of the mixed method is presented. Section 6 contains numerical experiments. Finally, in the last section we conclude with a few final remarks.

2. The Method

We assume that $\mathcal{T}_h$ is a shape-regular simplicial decompositions of Ω. Moreover, we define the following function spaces.

$$\begin{aligned}
W_h :=&\{w \in L^2(\Omega) : w|_K \in \mathcal{P}^k(K), \text{ for all } K \in \mathcal{T}_h\},\\
\boldsymbol{Q}_h :=&\{\boldsymbol{m} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{m}|_K \in \boldsymbol{\mathcal{P}}^k(K), \text{ for all } K \in \mathcal{T}_h\},\\
\boldsymbol{\Sigma}_h :=&\{\boldsymbol{v} \in \boldsymbol{H}(\mathrm{div}\,,\Omega) : \boldsymbol{v}|_K \in \boldsymbol{RT}^k(K) \text{ for all } K \in \mathcal{T}_h\},\\
\underline{\boldsymbol{Z}}_h :=&\{\underline{\boldsymbol{s}} \in \underline{\boldsymbol{H}}(\mathrm{div}\,,\Omega) : \text{ each row of } \underline{\boldsymbol{s}} \text{ belongs to } \boldsymbol{\Sigma}_h\}.
\end{aligned}$$

We will also need the space

$$W_h^\ell := \{w \in L^2(\Omega) : w|_K \in \mathcal{P}^\ell(K), \text{ for all } K \in \mathcal{T}_h\},$$

$\ell \geq -1$.

Here $\boldsymbol{L}^2(\Omega) = [L^2(\Omega)]^d$. The space of polynomials of degree less than or equal to $k \geq 0$ is denoted by $\mathcal{P}^k(K)$ and $\boldsymbol{\mathcal{P}}^k(K) = [\mathcal{P}^k(K)]^d$. Furthermore, we let $\mathcal{P}^{-1}(K) := \{0\}$. The space $\boldsymbol{RT}^k(K) = \boldsymbol{\mathcal{P}}^k(K) + \mathcal{P}^k(K)\boldsymbol{x}$ is the Raviart-Thomas space of index k. Finally, $\underline{\boldsymbol{H}}(\mathrm{div}\,,\Omega)$ $d \times d$ denotes all matrix-valued functions such that each row belongs to the space $\boldsymbol{H}(\mathrm{div}\,,\Omega)$.

The finite element method finds $(u_h, \boldsymbol{q}_h, \underline{\boldsymbol{z}}_h, \boldsymbol{\sigma}_h) \in W_h \times \boldsymbol{Q}_h \times \underline{\boldsymbol{Z}}_h \times \boldsymbol{\Sigma}_h$ that satisfy

$$(\boldsymbol{q}_h, \boldsymbol{v}) + (u_h, \nabla\cdot\boldsymbol{v}) = 0 \tag{2.2a}$$

$$(\underline{\boldsymbol{z}}_h, \underline{\boldsymbol{s}}) + (\boldsymbol{q}_h, \nabla\cdot\underline{\boldsymbol{s}}) = 0 \tag{2.2b}$$

$$-(\boldsymbol{\sigma}_h, \boldsymbol{m}) + (\boldsymbol{m}, \nabla\cdot\underline{\boldsymbol{z}}_h) = 0 \tag{2.2c}$$

$$(w, \nabla\cdot\boldsymbol{\sigma}_h) = (f, w) \tag{2.2d}$$

for all $(w, \boldsymbol{m}, \underline{\boldsymbol{s}}, \boldsymbol{v}) \in W_h \times \boldsymbol{Q}_h \times \underline{\boldsymbol{Z}}_h \times \boldsymbol{\Sigma}_h$.

For matrix-valued functions we used the notation

$$(\underline{\boldsymbol{z}}, \underline{\boldsymbol{s}}) := \sum_{K\in\mathcal{T}_h} (\underline{\boldsymbol{z}}, \underline{\boldsymbol{s}})_K, \text{ where } (\underline{\boldsymbol{z}}, \underline{\boldsymbol{s}})_K := \int_K \underline{\boldsymbol{z}}(\boldsymbol{x}) : \underline{\boldsymbol{s}}(\boldsymbol{x}) d\boldsymbol{x},$$

where : is the Froebenius inner product. For vector-valued and scalar-valued functions we take a similar definition.

Since $u = 0$ on $\partial\Omega$ it will also be true that the tangential component of $\boldsymbol{q}$ is equal to zero at the $\partial\Omega$. Hence, $\boldsymbol{q}$ vanishes on $\partial\Omega$ and this will make the above equations consistent.

We prove that the method is well defined, but we first state a standard result. For the proof in two dimensions see for example [12].

Proposition 2.1. *If* $\boldsymbol{v} \in \boldsymbol{\Sigma}_h$ *and* $\nabla \cdot \boldsymbol{v} \in W_h^{k-1}$ *then* $\boldsymbol{v} \in \boldsymbol{\Sigma}_h \cap \boldsymbol{Q}_h$.

Theorem 2.2. *The mixed method* (2.2) *is well defined.*

Proof. Since (2.2) is a square linear system it is enough to prove uniqueness. To this end, we assume that $f \equiv 0$ and we have

$$\begin{aligned}\|\underline{\boldsymbol{z}}_h\|^2_{L^2(\Omega)} &= -(\boldsymbol{q}_h, \nabla \cdot \underline{\boldsymbol{z}}_h) && \text{by (2.2b)}\\ &= -(\boldsymbol{\sigma}_h, \boldsymbol{q}_h) && \text{by (2.2c)}\\ &= (u_h, \nabla \cdot \boldsymbol{\sigma}_h) && \text{by (2.2a)}\\ &= 0 && \text{by (2.2d).}\end{aligned}$$

This shows that $\underline{\boldsymbol{z}}_h = 0$. From (2.2b) we get

$$(\boldsymbol{q}_h, \nabla \cdot \underline{\boldsymbol{s}}) = 0, \qquad \text{for all } \underline{\boldsymbol{s}} \in \underline{\boldsymbol{Z}}_h.$$

By the property of the Raviart-Thomas space we have that the divergence operator is onto from $\underline{\boldsymbol{Z}}_h$ to $\boldsymbol{Q}_h$; see [7]. Hence, $\boldsymbol{q}_h = 0$. Similarly, we conclude that $u_h = 0$ since by (2.2a) we have

$$(u_h, \nabla \cdot \boldsymbol{v}) = 0, \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{\Sigma}_h.$$

Finally, from (2.2d) we easily get that $\nabla \cdot \boldsymbol{\sigma}_h = 0$ and by Proposition 2.1 we have $\boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_h \cap \boldsymbol{Q}_h$. Hence, using (2.2c) we obtain $\boldsymbol{\sigma}_h = 0$. □

3. Error Estimates

In this section we prove error estimates for all the variables. We start by writing the error equations

$$(\boldsymbol{q} - \boldsymbol{q}_h, \boldsymbol{v}) + (u - u_h, \nabla \cdot \boldsymbol{v}) = 0, \tag{3.3a}$$

$$(\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h, \underline{\boldsymbol{s}}) + (\boldsymbol{q} - \boldsymbol{q}_h, \nabla \cdot \underline{\boldsymbol{s}}) = 0, \tag{3.3b}$$

$$-(\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{m}) + (\boldsymbol{m}, \nabla \cdot (\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h)) = 0, \tag{3.3c}$$

$$(w, \nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)) = 0, \tag{3.3d}$$

for all $(w, \boldsymbol{m}, \boldsymbol{v}, \underline{\boldsymbol{s}}) \in W_h \times \boldsymbol{Q}_h \times \boldsymbol{\Sigma}_h \times \underline{\boldsymbol{Z}}_h$.

We also need to define the some projections. We let $\boldsymbol{\Pi} : \boldsymbol{H}(\mathrm{div}, \Omega) \cap L^p(\Omega) \to \boldsymbol{\Sigma}_h$ (for $p > 2$) be the Raviart-Thomas projection [21, 19] of index k defined by on each $K \in \mathcal{T}_h$ by

$$(\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{v})_K = 0 \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{\mathcal{P}}^{k-1}(K), \tag{3.4a}$$

$$\langle (\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}) \cdot \boldsymbol{n}, \mu \rangle_F = 0 \qquad \text{for all } \mu \in \mathcal{P}^k(F), \text{ for all faces } F \text{ of } K. \tag{3.4b}$$

Here we used the notation $\langle \mu, m \rangle_F = \int_F \mu(s) m(s) ds$. Moreover, we let $\underline{\boldsymbol{\Pi}}$ denote the matrix version of $\underline{\boldsymbol{\Pi}}$ as it acts on matrix-valued functions where $\boldsymbol{\Pi}$ acts on each row. We let $\boldsymbol{P}$ be the L^2-projection onto $\boldsymbol{Q}_h$. We let $\boldsymbol{P}^0$ be the L^2-projection onto piecewise constant vector-valued functions. Finally, P is the L^2-projection onto W_h and P^ℓ is the L^2-projection onto W_h^ℓ. Throughout this paper we will assume that $\boldsymbol{\sigma}$ belongs to the domain of $\boldsymbol{\Pi}$ and $\underline{\boldsymbol{z}}$ belongs to the domain of $\underline{\boldsymbol{\Pi}}$.

We will need a few properties of $\boldsymbol{\Pi}$. First, the commutative property says

$$\nabla \cdot (\boldsymbol{\Pi}\boldsymbol{\sigma}) = P \, \nabla \cdot \boldsymbol{\sigma}. \tag{3.5}$$

The following approximation properties hold

$$\|\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}\|_{L^2(K)} \le h_K^{r+1} \|\boldsymbol{\sigma}\|_{H^{r+1}(K)}, \tag{3.6}$$

for $0 \le r \le k$ and $K \in \mathcal{T}_h$.

Before proving the error estimates we prove an important lemma that gives error estimates for $\boldsymbol{q}$ in terms of $\underline{\boldsymbol{z}}$.

Lemma 3.1. *There exists a constant C such that*

$$\|(\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h) - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h)\|_{L^2(K)} \le C\, h_K \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(K)}, \tag{3.7}$$

for all $K \in \mathcal{T}_h$. Here h_K is the diameter of K. Moreover, we have the global estimate

$$\|\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(\Omega)} \le C\, \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)}. \tag{3.8}$$

Proof. It is well known that there exists a $\underline{\boldsymbol{\psi}} \in \underline{\boldsymbol{H}}^1(K)$ such that

$$\begin{aligned} \nabla \cdot \underline{\boldsymbol{\psi}} &= (\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h) - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h) && \text{in } K \\ \underline{\boldsymbol{\psi}} &= 0 && \text{on } \partial K \end{aligned}$$

with

$$\|\underline{\boldsymbol{\psi}}\|_{H^1(K)} \le C\, \|(\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h) - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h)\|_{H^1(K)}.$$

Since $(\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h) - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h) \in \mathcal{P}^k(K)$, (3.5) gives us that

$$\begin{aligned} \nabla \cdot (\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}}) &= (\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h) - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h) && \text{in } K, \\ (\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}})\boldsymbol{n} &= 0 && \text{on } \partial K, \end{aligned}$$

where we also used (3.4b).

By (3.3b) we have

$$\begin{aligned} \|(\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h) - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h)\|_{L^2(K)}^2 &= ((\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h) - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h), \nabla \cdot (\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}}))_K \\ &= (\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h, \nabla \cdot (\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}}))_K \\ &= (\boldsymbol{q} - \boldsymbol{q}_h, \nabla \cdot (\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}}))_K \\ &= -(\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h, \underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}})_K \\ &\le \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(K)}\, \|\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}}\|_{L^2(K)}. \end{aligned}$$

In the second equation we used integration by parts, the fact that $\nabla \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h) = 0$ on K, and that $(\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}})\boldsymbol{n} = 0$ on ∂K. Next we use approximation properties of $\underline{\boldsymbol{\Pi}}$ (3.6) and Poincare's inequality to get

$$\begin{aligned} \|\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}}\|_{L^2(K)} &\le \|\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\psi}} - \underline{\boldsymbol{\psi}}\|_{L^2(K)} + \|\underline{\boldsymbol{\psi}}\|_{L^2(K)} \\ &\le C\, h_K \|\underline{\boldsymbol{\psi}}\|_{H^1(K)} \\ &\le C\, h_K \|(\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h) - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h)\|_{L^2(K)}. \end{aligned}$$

This proves (3.7). The proof of (3.8) is similar, but instead there exists a globally defined $\underline{\boldsymbol{\psi}}$ (not necessarily with zero boundary conditions) with right-hand side $\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h$. ☐

3.1. **Error estimates for $\underline{\boldsymbol{z}}$.** We start this section by stating the main theorem of this section plus a simple corollary.

Theorem 3.2. *We have*

$$\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)} \leq C\, \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}}\|_{L^2(\Omega)} + C \left(\sum_{K\in\mathcal{T}_h} h_K^{2j_k} \|\boldsymbol{\sigma} - \underline{\boldsymbol{\mathit{\Pi}}}\boldsymbol{\sigma}\|^2_{L^2(K)} \right)^{1/2},$$

where $j_k = 0$ if $k = 0$ and $j_k = 1$ if $k \geq 1$.

The following corollary easily follows from this theorem.

Corollary 3.3. *For any $0 \leq r \leq k$ we have*

$$\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)} \leq C\, h^{r+1} \|\underline{\boldsymbol{z}}\|_{H^{r+1+\ell_k}(\Omega)},$$

where $\ell_k = 1$ if $k = 0$ and $\ell_k = 0$ if $k \geq 1$.

Before proving Theorem 3.2 we first prove a simple but important lemma.

Lemma 3.4. *We have,*

$$\nabla \cdot (\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}_h) = 0, \tag{3.9}$$

and

$$\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}_h \in \boldsymbol{\Sigma}_h \cap \boldsymbol{Q}_h. \tag{3.10}$$

Proof. Using (3.3a) and (3.5) we have

$$(\nabla \cdot (\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}_h), w) = 0 \quad \text{for all } w \in W_h.$$

This proves (3.9), and (3.10) follows form Proposition 2.1. □

In the remainder of this section we prove Theorem 3.2.

Proof. (Theorem 3.2)
We have

$$\begin{aligned}
\|\underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|^2_{L^2(\Omega)} =&(\underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}, \underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h) - (\boldsymbol{Pq} - \boldsymbol{q}_h, \nabla \cdot (\underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h)) && \text{by (3.3b)}\\
=&(\underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}, \underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h) - (\boldsymbol{Pq} - \boldsymbol{q}_h, \nabla \cdot (\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h)) && \text{by (3.5)}\\
=&(\underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}, \underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h) - (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{Pq} - \boldsymbol{q}_h) && \text{by (3.3c)}\\
=&(\underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}, \underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h) - (\boldsymbol{\sigma} - \boldsymbol{\Pi\sigma}, \boldsymbol{Pq} - \boldsymbol{q}_h)\\
&- (\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{q} - \boldsymbol{q}_h) && \text{by (3.10)}\\
=&(\underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}, \underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h) - (\boldsymbol{\sigma} - \boldsymbol{\Pi\sigma}, \boldsymbol{Pq} - \boldsymbol{q}_h)\\
&+ (Pu - u_h, \nabla \cdot (\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}_h)) && \text{by (3.3a)}\\
=&(\underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}, \underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h) - (\boldsymbol{\sigma} - \boldsymbol{\Pi\sigma}, \boldsymbol{Pq} - \boldsymbol{q}_h) && \text{by (3.3d).}
\end{aligned}$$

Hence, we obtain

$$\|\underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|^2_{L^2(\Omega)} = (\underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}, \underline{\boldsymbol{\mathit{\Pi}}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h) - (\boldsymbol{\sigma} - \boldsymbol{\Pi\sigma}, \boldsymbol{Pq} - \boldsymbol{q}_h). \tag{3.11}$$

To bound the last term we first consider the case $k \geq 1$

$$
\begin{aligned}
-(\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}, \boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h) &= \sum_{K \in \mathcal{T}_h} -(\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}, \boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h))_K && \text{by (3.4a)} \\
&\leq \sum_{K \in \mathcal{T}_h} \|\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}\|_{L^2(K)} \|\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h)\|_{L^2(K)} \\
&\leq C \sum_{K \in \mathcal{T}_h} h_K \|\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}\|_{L^2(K)} \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(K)} && \text{by Lemma 3.7} \\
&\leq C \, \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)} \left(\sum_{K \in \mathcal{T}_h} h_K^2 \|\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}\|_{L^2(K)} \right)^{1/2}.
\end{aligned}
$$

Combining this inequality with (3.11) proves the theorem for $k \geq 1$. For $k = 0$, we instead use (3.8) to get

$$
(\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}, \boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h) \leq C \, \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)} \|\boldsymbol{\sigma} - \boldsymbol{\Pi}\boldsymbol{\sigma}\|_{L^2(\Omega)},
$$

which combined with (3.11) will prove the theorem for $k = 0$. □

3.2. **Error estimate for $\boldsymbol{q}$ and $\boldsymbol{\sigma}$.** The next theorem is consequence of Theorem 3.2.

Theorem 3.5. *Let $\ell_k = 1$ if $k = 0$ and $\ell_k = 0$ if $k \geq 1$. Then, for $0 \leq r \leq k$, we have*

$$
\|\boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(\Omega)} \leq C \, h^{r+1} (\|\underline{\boldsymbol{z}}\|_{H^{r+1+\ell_k}(\Omega)} + \|\boldsymbol{q}\|_{H^{r+1}(\Omega)}). \tag{3.12}
$$

Also,

$$
\|\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(K)} \leq \|\nabla \cdot (\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h)\|_{L^2(K)} + \|\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}\|_{L^2(K)} \tag{3.13}
$$

for all $K \in \mathcal{T}_h$.

If we assume the mesh is quasi-uniform

$$
\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)} \leq C \, h^r \, \|\underline{\boldsymbol{z}}\|_{H^{r+1+\ell_k}(\Omega)}. \tag{3.14}
$$

Before proving this theorem we make a few remarks. From this theorem we see that we get optimal estimates for $\boldsymbol{q}$ for any $k \geq 0$. However, we get sub-optimal estimates for $\boldsymbol{\sigma}$. Later we show that $\|\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(\Omega)}$ converges with order $k + 2$ for $k \geq 1$ on quasi-uniform meshes.

Proof. (Theorem 3.5)
By (3.8) we have

$$
\|\boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(\Omega)} \leq C \, (\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)} + \|\boldsymbol{q} - \boldsymbol{P}\boldsymbol{q}\|_{L^2(\Omega)}),
$$

which proves (3.12) after we use Theorem 3.2 and approximation properties of $\boldsymbol{P}$.

In order to prove (3.13) we use (3.10) and in (3.3c) we choose $\boldsymbol{m}|_K = (\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)|_K$ and define $\boldsymbol{m} = 0$ outside of K to get

$$
\begin{aligned}
\|\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(K)}^2 &= (\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)_K + (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)_K \\
&= (\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)_K + (\nabla \cdot (\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h), \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)_K \\
&= (\boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}, \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)_K + (\nabla \cdot (\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h), \boldsymbol{\Pi}\boldsymbol{\sigma} - \boldsymbol{\sigma}_h)_K
\end{aligned}
$$

where we used (3.5). This proves (3.13).

We can then use inverse estimates and Theorem 3.2 to get

$$
\|\nabla \cdot (\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h)\|_{L^2(\Omega)} \leq C \, h^r \, \|\underline{\boldsymbol{z}}\|_{H^{r+1+\ell_k}(\Omega)}.
$$

This proves (3.14). □

3.3. **Error estimate for u.** We prove estimates for u via a duality argument.

Consider the dual problem

$$\triangle^2\theta = \gamma \quad \text{in } \Omega, \tag{3.15a}$$
$$\theta = 0 \quad \text{on } \partial\Omega, \tag{3.15b}$$
$$\nabla\theta \cdot \boldsymbol{n} = 0 \quad \text{on } \partial\Omega. \tag{3.15c}$$

In order to get the best possible estimates we assume the following elliptic regularity result

$$\|\theta\|_{H^4(\Omega)} \leq C\,\|\gamma\|_{L^2(\Omega)}. \tag{3.16}$$

Such estimates are known to hold for polygonal domains with inner-angle conditions; see [4]. Moreover, we assume that $\underline{\boldsymbol{\xi}}$ belongs to the domain of $\underline{\boldsymbol{\Pi}}$ and $\boldsymbol{\phi}$ belongs to the domain of $\mathbf{\Pi}$ where $\boldsymbol{\psi} = \nabla\theta, \underline{\boldsymbol{\xi}} = \nabla\boldsymbol{\psi}$ and $\boldsymbol{\phi} = \nabla\cdot\underline{\boldsymbol{\xi}}$.

We first prove an estimate for $P^{k-1}(u-u_h)$ for $k \geq 1$ then we prove a weaker estimate for $Pu - u_h$.

Theorem 3.6. *Assuming the regularity result* (3.16) *and that* $\underline{\boldsymbol{\xi}}$ *and* $\boldsymbol{\phi}$ *belong to the domains of* $\underline{\boldsymbol{\Pi}}$ *and* $\mathbf{\Pi}$, *respectively, we have for* $k \geq 1$

$$\|P^{k-1}(u-u_h)\|_{L^2(\Omega)} \leq C\,h^{r+3}(\|f\|_{H^{r+1}(\Omega)} + \|\underline{\boldsymbol{z}}\|_{H^{r+1}(\Omega)}),$$

for any $0 \leq r \leq k$.

Proof. We let θ solve (3.15) with $\gamma = P^{k-1}(u-u_h)$. Note that by (3.5) $\nabla\cdot(\mathbf{\Pi}\boldsymbol{\phi}) = \gamma$. Since $\gamma \in W_h^{k-1}$, Proposition (2.1) give us that

$$\mathbf{\Pi}\boldsymbol{\phi} \in \mathbf{\Sigma}_h \cap \boldsymbol{Q}_h. \tag{3.17}$$

If we use (3.3a) we have that

$$\begin{aligned}
\|P^{k-1}(u-u_h)\|^2_{L^2(\Omega)} =& (P^{k-1}(u-u_h), \nabla\cdot\boldsymbol{\phi}) \\
=& (P^{k-1}(u-u_h), \nabla\cdot(\mathbf{\Pi}\boldsymbol{\phi})) \\
=& (u-u_h, \nabla\cdot(\mathbf{\Pi}\boldsymbol{\phi}))
\end{aligned}$$

where we used (3.5). Then,

$$\begin{aligned}
\|P^{k-1}(u-u_h)\|^2_{L^2(\Omega)} =& (\boldsymbol{q}-\boldsymbol{q}_h, \mathbf{\Pi}\boldsymbol{\phi}) && \text{by (3.3a)} \\
=& (\boldsymbol{P}\boldsymbol{q}-\boldsymbol{q}_h, \mathbf{\Pi}\boldsymbol{\phi}) && \text{by (3.17)} \\
=& (\boldsymbol{P}\boldsymbol{q}-\boldsymbol{q}_h, \boldsymbol{\phi}) + (\boldsymbol{P}\boldsymbol{q}-\boldsymbol{q}_h, \mathbf{\Pi}\boldsymbol{\phi}-\boldsymbol{\phi}), && \\
=& (\boldsymbol{P}\boldsymbol{q}-\boldsymbol{q}_h, \nabla\cdot\underline{\boldsymbol{\xi}}) + (\boldsymbol{P}\boldsymbol{q}-\boldsymbol{q}_h, \mathbf{\Pi}\boldsymbol{\phi}-\boldsymbol{\phi}) && \\
=& (\boldsymbol{P}\boldsymbol{q}-\boldsymbol{q}_h, \nabla\cdot(\underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\xi}})) + (\boldsymbol{P}\boldsymbol{q}-\boldsymbol{q}_h, \mathbf{\Pi}\boldsymbol{\phi}-\boldsymbol{\phi}) && \text{by (3.5)} \\
=& -(\underline{\boldsymbol{z}}-\underline{\boldsymbol{z}}_h, \underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\xi}}) + (\boldsymbol{P}\boldsymbol{q}-\boldsymbol{q}_h, \mathbf{\Pi}\boldsymbol{\phi}-\boldsymbol{\phi}) && \text{by (3.3b)} \\
=& -(\underline{\boldsymbol{z}}-\underline{\boldsymbol{z}}_h, \nabla\boldsymbol{\psi}) - (\underline{\boldsymbol{z}}-\underline{\boldsymbol{z}}_h, \underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\xi}}-\underline{\boldsymbol{\xi}}), && \\
&+ (\boldsymbol{P}\boldsymbol{q}-\boldsymbol{q}_h, \mathbf{\Pi}\boldsymbol{\phi}-\boldsymbol{\phi}) && \text{since } \nabla\boldsymbol{\psi} = \underline{\boldsymbol{\xi}}.
\end{aligned}$$

Moreover,

$$
\begin{aligned}
(\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h, \nabla \boldsymbol{\psi}) =& (\nabla \cdot (\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h), \boldsymbol{\psi}) && \text{integration by parts, } \boldsymbol{\psi} = 0 \text{ on } \partial\Omega \\
=& (\nabla \cdot (\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h), \boldsymbol{P}\boldsymbol{\psi}), \\
&+ (\nabla \cdot (\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h), \boldsymbol{\psi} - \boldsymbol{P}\boldsymbol{\psi}) \\
=& (\boldsymbol{\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{P}\boldsymbol{\psi}) \\
&+ (\nabla \cdot (\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h)\boldsymbol{\psi} - \boldsymbol{P}\boldsymbol{\psi}) && \text{by (3.3c)} \\
=& (\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{P}\boldsymbol{\psi}) \\
&+ (\boldsymbol{\sigma} - \boldsymbol{\Pi\sigma}, \boldsymbol{P}\boldsymbol{\psi}) \\
&+ (\boldsymbol{\sigma}, \boldsymbol{\psi} - \boldsymbol{P}\boldsymbol{\psi}) && \text{definition of } \boldsymbol{P},\ \nabla \cdot \underline{\boldsymbol{z}} = \boldsymbol{\sigma}.
\end{aligned}
$$

We show the first term in the right is zero. Indeed,

$$
\begin{aligned}
(\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{P}\boldsymbol{\psi}) =& (\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}_h, \boldsymbol{\psi}) && \text{property of } \boldsymbol{P} \text{ and (3.10)} \\
=& (\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}_h, \nabla\theta) && \text{since } \boldsymbol{\psi} = \nabla\theta \\
=& -(\nabla \cdot (\boldsymbol{\Pi\sigma} - \boldsymbol{\sigma}_h), \theta) && \text{since } \theta = 0 \text{ on } \partial\Omega \\
=& 0 && \text{by (3.9).}
\end{aligned}
$$

Moreover, using (3.5), the fact that $\nabla\theta = \boldsymbol{\psi}$ and $\nabla \cdot \boldsymbol{\sigma} = f$, and the properties of the L^2-projections $\boldsymbol{P}$ and P we get

$$
\begin{aligned}
(\boldsymbol{\sigma} - \boldsymbol{\Pi\sigma}, \boldsymbol{P}\boldsymbol{\psi}) + (\boldsymbol{\sigma}, \boldsymbol{\psi} - \boldsymbol{P}\boldsymbol{\psi}) =& (\boldsymbol{\Pi\sigma}, \boldsymbol{\psi} - \boldsymbol{P}\boldsymbol{\psi}) + (\boldsymbol{\sigma} - \boldsymbol{\Pi\sigma}, \nabla\theta) \\
=& (\boldsymbol{\Pi\sigma}, \boldsymbol{\psi} - \boldsymbol{P}\boldsymbol{\psi}) - (\nabla \cdot (\boldsymbol{\sigma} - \boldsymbol{\Pi\sigma}), \theta) \\
=& (\boldsymbol{\Pi\sigma}, \boldsymbol{\psi} - \boldsymbol{P}\boldsymbol{\psi}) - (\nabla \cdot \boldsymbol{\sigma} - P\nabla \cdot \boldsymbol{\sigma}, \theta) \\
=& (\boldsymbol{\Pi\sigma}, \boldsymbol{\psi} - \boldsymbol{P}\boldsymbol{\psi}) - (f - Pf, \theta - P\theta) \\
=& (\boldsymbol{\Pi\sigma} - \boldsymbol{v}, \boldsymbol{\psi} - \boldsymbol{P}\boldsymbol{\psi}) - (f - Pf, \theta - P\theta),
\end{aligned}
$$

for any $\boldsymbol{v} \in \boldsymbol{Q}_h$.

Therefore,

$$
(\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h, \nabla \boldsymbol{\psi}) = (\boldsymbol{\Pi\sigma} - \boldsymbol{v}, \boldsymbol{P}\boldsymbol{\psi} - \boldsymbol{\psi}) - (f - Pf, \theta - P\theta)
$$

Also, using (3.4a) we have

$$
(\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h, \boldsymbol{\Pi\phi} - \boldsymbol{\phi}) = (\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi\phi} - \boldsymbol{\phi}),
$$

where we used that $k \geq 1$. Hence,

$$
\begin{aligned}
\|P^{k-1}(u - u_h)\|^2_{L^2(\Omega)} =& -(\boldsymbol{\Pi\sigma} - \boldsymbol{v}, \boldsymbol{P}\boldsymbol{\psi} - \boldsymbol{\psi}) + (f - Pf, \theta - P\theta) \\
&- (\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h, \underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\xi}} - \underline{\boldsymbol{\xi}}) \\
&+ (\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h - \boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h), \boldsymbol{\Pi\phi} - \boldsymbol{\phi}).
\end{aligned}
$$

for any $\boldsymbol{v} \in \boldsymbol{Q}_h$.

It is not difficult to show that

$$
\inf_{\boldsymbol{v} \in \boldsymbol{Q}_h} \|\boldsymbol{\Pi\sigma} - \boldsymbol{v}\|_{L^2(\Omega)} \leq C\, h^{r+1} \|\nabla \cdot \boldsymbol{\sigma}\|_{H^r(\Omega)} = C\, h^{r+1} \|f\|_{H^r(\Omega)}, \tag{3.18}
$$

for any $0 \leq r \leq k$; see for instance (vi) of Proposition 2.1 in [11] .

Therefore,

$$\begin{aligned}\|P^{k-1}(u-u_h)\|^2_{L^2(\Omega)} \leq & C\, h^{r+3}\|f\|_{H^r(\Omega)}\|\boldsymbol{\psi}\|_{H^2(\Omega)} \\ & + C\, h^{r+3}\|f\|_{H^{r+1}(\Omega)}\|\theta\|_{H^2(\Omega)} \\ & + C\, h^2\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)}\, \|\underline{\boldsymbol{\xi}}\|_{H^2(\Omega)} \\ & + C\, h^2\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)}\|\boldsymbol{\phi}\|_{H^1(\Omega)},\end{aligned}$$

for $0 \leq r \leq k$. Here we used (3.7), (3.8) and approximation properties of $\underline{\boldsymbol{\Pi}}$. If we use the regularity assumption (3.16) and Corollary 3.3 we arrive at our result. □

We would like to note that if one inspects the proof we can replace $\|f\|_{H^{r+1}(\Omega)}$ with $\|f\|_{H^{r-1}(\Omega)}$ in the above estimate for $k \geq 3$.

Next we prove an estimate for $Pu - u_h$. The result will show that $\|Pu - u_h\|_{L^2(\Omega)}$ converges with order $k + 2$. Numerical experiments show that this in fact is sharp.

Theorem 3.7. *Assuming the regularity result* (3.16) *and that* $\underline{\boldsymbol{\xi}}$ *and* $\boldsymbol{\phi}$ *belong to the domains of* $\underline{\boldsymbol{\Pi}}$ *and* $\mathbf{\Pi}$, *respectively, we have*

$$\|Pu - u_h\|_{L^2(\Omega)} \leq C\, h^{r+2}(\|f\|_{H^{r+1}(\Omega)} + \|\underline{\boldsymbol{z}}\|_{H^{r+1+\ell_k}(\Omega)}),$$

for any $0 \leq r \leq k$. *Here* $\ell_k = 1$ *if* $k = 0$ *and* $\ell_k = 0$ *for* $k \geq 1$.

Before proving this result we note that this theorem give us that $\|u - u_h\|_{L^2(\Omega)}$ convergence with optimal order $k + 1$. This follows easily from the above result and the triangle inequality.

Proof. (Theorem 3.7) We let $\gamma = Pu - u_h$ in (3.15). Similar to the proof of Theorem 3.6 we can easily show that

$$\begin{aligned}\|Pu - u_h\|^2_{L^2(\Omega)} = & -(\mathbf{\Pi}\boldsymbol{\sigma} - \boldsymbol{v}, \boldsymbol{P}\boldsymbol{\psi} - \boldsymbol{\psi}) + (f - Pf, \theta - P\theta) \\ & - (\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h, \underline{\boldsymbol{\Pi}}\underline{\boldsymbol{\xi}} - \underline{\boldsymbol{\xi}}) \\ & + (\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h, \mathbf{\Pi}\boldsymbol{\phi} - \boldsymbol{\phi}) \\ & + (\boldsymbol{q} - \boldsymbol{P}\boldsymbol{q}, \mathbf{\Pi}\boldsymbol{\phi} - \boldsymbol{P}\boldsymbol{\phi}).\end{aligned}$$

for any $\boldsymbol{v} \in \boldsymbol{Q}_h$. Note that the last term appears here and not in the estimate of $P^{k-1}(u-u_h)$ in Theorem (3.6). This is indeed the term that reduces the order of convergence of $Pu - u_h$ to $k + 2$.

It easily follows that

$$\begin{aligned}\|Pu - u_h\|^2_{L^2(\Omega)} \leq & C\, h^{r+2}\|f\|_{H^r(\Omega)}\|\boldsymbol{\psi}\|_{H^1(\Omega)} \\ & + C\, h^{r+2}\|f\|_{H^{r+1}(\Omega)}\|\theta\|_{H^1(\Omega)} \\ & + C\, h\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)}\, \|\underline{\boldsymbol{\xi}}\|_{H^1(\Omega)} \\ & + C\, h\|\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(\Omega)}\|\boldsymbol{\phi}\|_{H^1(\Omega)} \\ & + C\, h\|\boldsymbol{q} - \boldsymbol{P}\boldsymbol{q}\|_{L^2(\Omega)}\|\boldsymbol{\phi}\|_{H^1(\Omega)},\end{aligned}$$

where we used (3.18). Finally, our result now follows if we use (3.12) and Corollary 3.3. □

3.4. **Error Estimates for $\boldsymbol{Pq}-\boldsymbol{q}_h$.** In this section we prove superconvergence results for $\boldsymbol{Pq}-\boldsymbol{q}_h$ with $k \geq 1$ and quasi-uniform meshes in two and three dimensions. We start with a lemma.

Lemma 3.8. *Let F common face of two elements $K, K' \in \mathcal{T}_h$. Then*

$$\|(\boldsymbol{Pq}-\boldsymbol{q}_h)|_K - (\boldsymbol{Pq}-\boldsymbol{q}_h)|_{K'}\|_{L^2(F)} \leq C\, h_F^{1/2} \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(K\cup K')}, \tag{3.19}$$

where h_F is the diameter of F. Moreover, if F is a face of $K \in \mathcal{T}_h$ and F belongs to the boundary $\partial\Omega$ then

$$\|\boldsymbol{Pq}-\boldsymbol{q}_h\|_{L^2(F)} \leq C\, h_F^{1/2} \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(K)}. \tag{3.20}$$

Proof. We only prove (3.19). To this end, let $\boldsymbol{r} = (\boldsymbol{Pq}-\boldsymbol{q}_h)|_K - (\boldsymbol{Pq}-\boldsymbol{q}_h)|_{K'}$ and define $\underline{\boldsymbol{s}} \in \underline{\boldsymbol{Z}}_h$ in the following way: First, $\underline{\boldsymbol{s}}|_K \in \underline{\boldsymbol{RT}}^k(K)$ solves

$$
\begin{aligned}
(\underline{\boldsymbol{s}}, \underline{\boldsymbol{v}})_K &= 0 && \text{for all } \underline{\boldsymbol{v}} \in \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K), && (3.21a)\\
\langle \underline{\boldsymbol{s}}\boldsymbol{n}_K, \boldsymbol{\mu}\rangle_F &= \langle \boldsymbol{r}, \boldsymbol{\mu}\rangle_F && \text{for all } \boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}^k(F), && (3.21b)\\
\langle \underline{\boldsymbol{s}}\boldsymbol{n}_K, \boldsymbol{\mu}\rangle_G &= 0, && \text{for all } \boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}^k(G), \text{ for all faces } G \text{ of } K \text{ and } G \neq F, && (3.21c)
\end{aligned}
$$

where here $\boldsymbol{n}_K$ is the outward unit normal to K. Here $\underline{\boldsymbol{RT}}^k(K)$ is the set of matrix-valued functions such that each row belongs to $\boldsymbol{RT}^k(K)$.

Define $\underline{\boldsymbol{s}}|_{K'} \in \underline{\boldsymbol{RT}}^k(K')$

$$
\begin{aligned}
(\underline{\boldsymbol{s}}, \underline{\boldsymbol{v}})_{K'} &= 0 && \text{for all } \underline{\boldsymbol{v}} \in \underline{\boldsymbol{\mathcal{P}}}^{k-1}(K'), && (3.22a)\\
\langle \underline{\boldsymbol{s}}\boldsymbol{n}_{K'}, \boldsymbol{\mu}\rangle_F &= -\langle \boldsymbol{r}, \boldsymbol{\mu}\rangle_F && \text{for all } \boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}^k(F), && (3.22b)\\
\langle \underline{\boldsymbol{s}}\boldsymbol{n}_{K'}, \boldsymbol{\mu}\rangle_G &= 0 && \text{for all } \boldsymbol{\mu} \in \boldsymbol{\mathcal{P}}^k(G), \text{ for all faces } G \text{ of } K' \text{ and } G \neq F, && (3.22c)
\end{aligned}
$$

where here $\boldsymbol{n}_{K'}$ is the outward unit normal to K'.

Finally, set

$$\underline{\boldsymbol{s}}|_{\Omega\setminus K\cup K'} \equiv 0. \tag{3.23}$$

A standard scaling argument gives

$$\|\underline{\boldsymbol{s}}\|_{L^2(K\cup K')} \leq C\, h_F^{1/2} \|\boldsymbol{r}\|_{L^2(F)}. \tag{3.24}$$

Hence, we have

$$
\begin{aligned}
\|\boldsymbol{r}\|^2_{L^2(F)} =& \langle \boldsymbol{r}, \boldsymbol{r}\rangle_F \\
=& \langle \boldsymbol{r}, \underline{\boldsymbol{s}}\boldsymbol{n}_K\rangle_F && \text{by (3.21b)}\\
=& \langle (\boldsymbol{Pq}-\boldsymbol{q}_h)|_K, \underline{\boldsymbol{s}}\boldsymbol{n}_K\rangle_F + \langle (\boldsymbol{Pq}-\boldsymbol{q}_h)|_{K'}, \underline{\boldsymbol{s}}\boldsymbol{n}_{K'}\rangle_F \\
=& \int_{\partial K} (\boldsymbol{Pq}-\boldsymbol{q}_h)\cdot \underline{\boldsymbol{s}}\boldsymbol{n}_K + \int_{\partial K'} (\boldsymbol{Pq}-\boldsymbol{q}_h)\cdot \underline{\boldsymbol{s}}\boldsymbol{n}_{K'} && \text{by (3.21c),(3.22c)}\\
=& (\boldsymbol{Pq}-\boldsymbol{q}_h, \nabla\cdot\underline{\boldsymbol{s}})_K + (\boldsymbol{Pq}-\boldsymbol{q}_h, \nabla\cdot\underline{\boldsymbol{s}})_{K'} && \text{by (3.21a), (3.22a)}\\
=& (\boldsymbol{Pq}-\boldsymbol{q}_h, \nabla\cdot\underline{\boldsymbol{s}}) && \text{by (3.23)}\\
=& -(\underline{\boldsymbol{z}}-\underline{\boldsymbol{z}}_h, \underline{\boldsymbol{s}}) && \text{by (3.3b)}\\
=& -(\underline{\boldsymbol{z}}-\underline{\boldsymbol{z}}_h, \underline{\boldsymbol{s}})_K - (\underline{\boldsymbol{z}}-\underline{\boldsymbol{z}}_h, \underline{\boldsymbol{s}})_{K'} && \text{by (3.23).}
\end{aligned}
$$

Therefore,

$$\|\boldsymbol{r}\|_{L^2(F)} \leq \|\underline{\boldsymbol{z}}-\underline{\boldsymbol{z}}_h\|_{L^2(K\cup K')} \|\underline{\boldsymbol{s}}\|_{L^2(K\cup K')}.$$

The result now follows if we apply (3.24). □

We can now state the main results of this section.

Theorem 3.9. *For $k \geq 1$ we have*

$$\|\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(\Omega)} \leq h\,\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)} + \Big(\sum_{K \in \mathcal{T}_h} \frac{1}{h_K^2} \|P^{k-1}(u - u_h)\|_{L^2(K)}^2\Big)^{1/2}.$$

Before proving this we can state a simple consequence that follows from Theorems 3.3 and 3.6.

Corollary 3.10. *Suppose the hypotheses of Theorem 3.6 hold. Furthermore, assume the mesh $\mathcal{T}_h$ is quasi-uniform. Then, for $k \geq 1$*

$$\|\boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(\Omega)} \leq C\,h^{r+2}(\|\underline{\boldsymbol{z}}\|_{H^{r+1}(\Omega)} + \|f\|_{H^{r+1}(\Omega)}),$$

for $0 \leq r \leq k$.

Proof. (Theorem 3.9)

We need some notation. Let F be a common face of $K \in \mathcal{T}_h$ and $K' \in \mathcal{T}_h$. Let $\boldsymbol{v}_K$ be the restriction of $\boldsymbol{v}|_K$ to F and $\boldsymbol{v}_{K'}$ be the restriction of $\boldsymbol{v}|_{K'}$ to F. Define $\{\!\{\boldsymbol{v}\}\!\}$ and $[\![\boldsymbol{v}]\!]$ on F as

$$\{\!\{\boldsymbol{v}\}\!\} = \frac{1}{2}(\boldsymbol{v}_K + \boldsymbol{v}_{K'}), \quad [\![\boldsymbol{v}]\!] = \boldsymbol{v}_K \cdot \boldsymbol{n}_K + \boldsymbol{v}_{K'} \cdot \boldsymbol{n}_{K'}. \tag{3.25}$$

where $\boldsymbol{n}_K$ is the outward unit-normal to K and $\boldsymbol{n}_{K'}$ is the outward unit-normal to K'.

We let $\boldsymbol{r} = \boldsymbol{P}\boldsymbol{q} - \boldsymbol{q}_h$. Using this notation define a function $\boldsymbol{\omega} \in \boldsymbol{\Sigma}_h \cap \boldsymbol{Q}_h$ in the following way: For every $K \in \mathcal{T}_h$ define $\boldsymbol{\omega}|_K \in \boldsymbol{\mathcal{P}}^k(K)$ as the unique solution to

$$(\boldsymbol{\omega}, \boldsymbol{v})_K = (\boldsymbol{r}, \boldsymbol{v}) \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{N}^{k-1}(K), \tag{3.26a}$$

$$\langle \boldsymbol{\omega} \cdot \boldsymbol{n}_K, \mu \rangle_F = \langle \{\!\{\boldsymbol{r}\}\!\} \cdot \boldsymbol{n}_K, \mu \rangle_F \qquad \text{for all } \mu \in \mathcal{P}^k(F) \text{ and all faces } F \text{ of } K, \tag{3.26b}$$

where $\boldsymbol{n}_K$ is the outward unit normal to K. Here $\boldsymbol{N}^{k-1}(K)$ is the Nédélec space of index $k-1$; see [20]. On each element $K \in \mathcal{T}_h$ we have

$$(\boldsymbol{\omega} - \boldsymbol{r}, \boldsymbol{v})_K = 0 \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{N}^{k-1}(K), \tag{3.27a}$$

$$\langle (\boldsymbol{\omega} - \boldsymbol{r}) \cdot \boldsymbol{n}_K, \mu \rangle_F = -\langle \frac{1}{2}[\![\boldsymbol{r}]\!], \mu \rangle_F \qquad \text{for all } \mu \in \mathcal{P}^k(F), \text{ and all faces } F \text{ of } K \tag{3.27b}$$

A scaling argument then gives

$$\|\boldsymbol{\omega} - \boldsymbol{r}\|_{L^2(K)} \leq C\,h_K^{1/2}\|[\![\boldsymbol{r}]\!]\|_{L^2(\partial K)}.$$

However, by Corollary 3.10 we have

$$\|\boldsymbol{\omega} - \boldsymbol{r}\|_{L^2(K)} \leq C\,h_K \sum_{K' \in D(K)} \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(K')}. \tag{3.28}$$

where $D(K)$ denotes the set of all $K' \in \mathcal{T}_h$ that share a face with K.

We can then write

$$\begin{aligned}
\|\boldsymbol{r}\|_{L^2(\Omega)}^2 =& (\boldsymbol{r}, \boldsymbol{\omega}) + (\boldsymbol{r}, \boldsymbol{r} - \boldsymbol{\omega}) \\
=& (\boldsymbol{q} - \boldsymbol{q}_h, \boldsymbol{\omega}) + (\boldsymbol{r}, \boldsymbol{r} - \boldsymbol{\omega}) && \text{since } \boldsymbol{\omega} \in \boldsymbol{Q}_h \\
=& (u - u_h, \nabla \cdot \boldsymbol{\omega}) + (\boldsymbol{r}, \boldsymbol{r} - \boldsymbol{\omega}) && \text{since } \boldsymbol{\omega} \in \boldsymbol{\Sigma}_h, \text{ and (3.3a)} \\
=& (P^{k-1}(u - u_h), \nabla \cdot \boldsymbol{\omega}) + (\boldsymbol{r}, \boldsymbol{r} - \boldsymbol{\omega}) && \text{since } \boldsymbol{\omega} \in \boldsymbol{Q}_h
\end{aligned}$$

We first bound the first term on the right hand side using an inverse estimate

$$\begin{aligned}(P^{k-1}(u-u_h), \nabla\cdot\boldsymbol{\omega}) &= \sum_{K\in\mathcal{T}_h} (P^{k-1}(u-u_h), \nabla\cdot\boldsymbol{r})_K \\ &\quad + \sum_{K\in\mathcal{T}_h} (P^{k-1}(u-u_h), \nabla\cdot(\boldsymbol{\omega}-\boldsymbol{r}))_K \\ &\le \sum_{K\in\mathcal{T}_h} \frac{C}{h_K}\|P^{k-1}(u-u_h)\|_{L^2(K)}\,\|\boldsymbol{r}\|_{L^2(K)} \\ &\quad + \sum_{K\in\mathcal{T}_h} \frac{C}{h_K}\|P^{k-1}(u-u_h)\|_{L^2(K)}\,\|\boldsymbol{r}-\boldsymbol{\omega}\|_{L^2(K)}.\end{aligned}$$

Also, we have

$$(\boldsymbol{r}, \boldsymbol{r}-\boldsymbol{\omega}) \le \|\boldsymbol{r}\|_{L^2(\Omega)}\,\|\boldsymbol{r}-\boldsymbol{\omega}\|_{L^2(\Omega)}.$$

Hence,

$$\|\boldsymbol{r}\|_{L^2(\Omega)} \le C\,\|\boldsymbol{r}-\boldsymbol{\omega}\|_{L^2(\Omega)} + C\,(\sum_{K\in\mathcal{T}_h}\frac{1}{h_K^2}\|P^{k-1}(u-u_h)\|^2_{L^2(K)})^{1/2}.$$

The result now follows if we apply (3.28).

□

4. Postprocessing

Postprocessing solutions of mixed methods have been widely used; see for example [1, 24]. Here we give a postprocessed approximation to u which is calculated locally on each element $K\in\mathcal{T}_h$. First, we provide the postprocessed approximation for $k\ge 2$ which converges with order $k+3$. Then, we give a postprocessing technique for $\boldsymbol{q}$ that converges with order $k+2$. Finally, we give postprocessed approximation of u for $k=0,1$. We will need the following space of functions.

$$\mathcal{P}^{\ell,m}_{\perp}(K) := \{v\in\mathcal{P}^{\ell}(K) : (v,\omega)_K = 0 \text{ for all } \omega\in\mathcal{P}^m(K)\}.$$

and also $\boldsymbol{\mathcal{P}}^{\ell,m}_{\perp} = [\mathcal{P}^{\ell,m}_{\perp}(K)]^d$.

4.1. Postprocessing for u: The case $k\ge 2$.

Let $\underline{\boldsymbol{D}}^2(u)$ denote the matrix containing the second order derivatives of u. In other words, $(\underline{\boldsymbol{D}}^2(u))_{ij} := \partial_{x_j}\partial_{x_i}u$ for $i,j=1,\dots,d$.

We define the postprocessed approximation $u_h^\star\in\mathcal{P}^{k+2}(K)$ for $k\ge 1$ locally by

$$(\underline{\boldsymbol{D}}^2(u_h^\star), \underline{\boldsymbol{D}}^2(v))_K = (\underline{\boldsymbol{z}}_h, \underline{\boldsymbol{D}}^2(v))_K \quad \forall v\in\mathcal{P}^{k+2,1}_{\perp}(K) \tag{4.29a}$$

$$(u_h^\star, w)_K = (u_h, w)_K \quad \forall w\in\mathcal{P}^1(K). \tag{4.29b}$$

We prove that $u_h^\star$ is well defined.

Theorem 4.1. *The problem* (4.29) *is well defined.*

Proof. Since (4.29) is a square system we only need to prove uniqueness. To this end, assume $\underline{\boldsymbol{z}}_h = 0$ and $u_h = 0$. Then, (4.29b) gives us that $u_h^\star\in\mathcal{P}^{k+2,1}_{\perp}(K)$ and $P^1u_h^\star = 0$. Approximation properties of P^1 give us that

$$\|u_h^\star\|_{L^2(K)} = \|u_h^\star - P^1u_h^\star\|_{L^2(K)} \le Ch_K^2\|\underline{\boldsymbol{D}}^2(u_h^\star)\|_{L^2(K)}.$$

The first equation (4.29a) gives us that

$$\|\underline{\boldsymbol{D}}^2(u_h^\star)\|_{L^2(K)} = 0,$$

which proves the theorem. □

Although we defined the postprocessed approximation for $k \geq 1$ we can only prove $k+3$ order of convergence for $k \geq 2$. The reason for this is that $P^1(u - u_h)$ is of order $k+3$ as long as $k \geq 2$; see Theorem 3.6.

Theorem 4.2. *For $k \geq 2$ and $0 \leq r \leq k$ we have*

$$\|u - u_h^\star\|_{L^2(\Omega)} \leq C\, h^{r+3}(\|u\|_{H^{r+3}(\Omega)} + \|f\|_{H^{r+1}(\Omega)}).$$

Proof. We start by applying the triangle inequality to get

$$\|u - u_h^\star\|_{L^2(K)} \leq \|P^1(u - u_h)\|_{L^2(K)} + \|(u - u_h^\star) - P^1(u - u_h^\star)\|_{L^2(K)},$$

where we used that $P^1 u_h^\star = P^1 u_h$ which follows from (4.29b). However,

$$\|(u - u_h^\star) - P^1(u - u_h^\star)\|_{L^2(K)} \leq C h_K^2 \|\underline{\boldsymbol{D}}^2(u - u_h^\star)\|_{L^2(K)}.$$

Hence,

$$\|u - u_h^\star\|_{L^2(K)} \leq \|P^1(u - u_h)\|_{L^2(K)} + C\, h_K^2 \|\underline{\boldsymbol{D}}^2(u - u_h^\star)\|_{L^2(K)}.$$

In order to approximate $\|\underline{\boldsymbol{D}}^2(u - u_h^\star)\|_{L^2(K)}$ we use

$$\begin{aligned}
\|\underline{\boldsymbol{D}}^2(u - u_h^\star)\|_{L^2(K)}^2 =&(\underline{\boldsymbol{D}}^2(u - u_h^\star), \underline{\boldsymbol{D}}^2(u - u_h^\star))_K \\
=&(\underline{\boldsymbol{D}}^2(u - u_h^\star), \underline{\boldsymbol{D}}^2(P^{k+2}u - u_h^\star - P^1(u - u_h^\star)))_K \\
&+ (\underline{\boldsymbol{D}}^2(u - u_h^\star), \underline{\boldsymbol{D}}^2(u - P^{k+2}u))_K,
\end{aligned}$$

Noting that $P^{k+2}u - u_h^\star - P^1(u - u_h^\star) \in \mathcal{P}_\perp^{k+2,1}(K)$ and using (4.29a) we have

$$\begin{aligned}
(\underline{\boldsymbol{D}}^2(u - u_h^\star), \underline{\boldsymbol{D}}^2(P^{k+2}u - u_h^\star - P^1(u - u_h^\star)))_K =&(\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h, \underline{\boldsymbol{D}}^2(P^{k+2}u - u_h^\star - P^1(u - u_h^\star)))_K \\
=&(\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h, \underline{\boldsymbol{D}}^2(P^{k+2}u - u))_K \\
&+ (\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h, \underline{\boldsymbol{D}}^2(u - u_h^\star))_K.
\end{aligned}$$

Therefore,

$$\|\underline{\boldsymbol{D}}^2(u - u_h^\star)\|_{L^2(K)} \leq C\,(\|\underline{\boldsymbol{D}}^2(u - P^{k+2}u)\|_{L^2(K)} + \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(K)}).$$

Hence,

$$\begin{aligned}
\|u - u_h^\star\|_{L^2(K)} \leq& C\,(\|P^1(u - u_h)\|_{L^2(K)} + h_K^2\|\underline{\boldsymbol{D}}^2(u - P^{k+2}u)\|_{L^2(K)} + h_K^2\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(K)}) \\
\leq& C\,(\|P^1(u - u_h)\|_{L^2(K)} + h_K^2\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(K)} + h_K^{r+3}\|u\|_{H^{r+3}(K)}),
\end{aligned}$$

for $0 \leq r \leq k$.

Adding the contribution of all $K \in \mathcal{T}_h$ we get

$$\|u - u_h^\star\|_{L^2(\Omega)} \leq C\,(\|P^1(u - u_h)\|_{L^2(\Omega)} + h^2\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)} + h^{r+3}\|u\|_{H^{r+3}(\Omega)}).$$

The proof is complete if we apply Theorems 3.6 and 3.2. □

4.2. **Postprocessing for $\boldsymbol{q}$.** In this section we will take advantage of the Theorem 3.10 and define a postprocessed approximtion to $\boldsymbol{q}$ that converges with order $k+2$ for $k \geq 1$.

We define $\boldsymbol{q}_h^\star|_K \in \boldsymbol{\mathcal{P}}^{k+1}(K)$ as the solution to

$$(\nabla \boldsymbol{q}_h^\star, \nabla \boldsymbol{v})_K = (\underline{\boldsymbol{z}}_h, \nabla \boldsymbol{v})_K \qquad \text{for all } \boldsymbol{v} \in \boldsymbol{\mathcal{P}}_\perp^{k+1,0}(K) \tag{4.30a}$$

$$(\boldsymbol{q}_h^\star, \boldsymbol{w})_K = (\boldsymbol{q}_h, \boldsymbol{w})_K \qquad \text{for all } \boldsymbol{w} \in \boldsymbol{\mathcal{P}}^0(K), \tag{4.30b}$$

for all $K \in \mathcal{T}_h$. The proof of the next theorem is similar to the proof of Theorem 4.2 so we omit the details.

Theorem 4.3. *The post processing approximation $\boldsymbol{q}_h^\star$ defined by* (4.30) *is well defined. Moreover, the following estimate holds for $0 \leq r \leq k$*

$$\|\boldsymbol{q} - \boldsymbol{q}_h^\star\|_{L^2(K)} \leq C\, h_K^{r+2} \|\boldsymbol{q}\|_{H^{r+2}(K)} + C\, h_K \|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(K)} + \|\boldsymbol{P}^0(\boldsymbol{q} - \boldsymbol{q}_h)\|_{L^2(K)},$$

for all $K \in \mathcal{T}_h$.

The following corollary easily follows if we use Corollary 3.3 and Corollary 3.10.

Corollary 4.4. *Assume the hypotheses of Theorem 3.6 hold. Moreover, assume the mesh $\mathcal{T}_h$ is quasi-uniform. Then, for $k \geq 1$*

$$\|\boldsymbol{q} - \boldsymbol{q}_h^\star\|_{L^2(\Omega)} \leq C\, h^{r+2} (\|\underline{\boldsymbol{z}}\|_{H^{r+1}(\Omega)} + \|f\|_{H^{r+1}(\Omega)}).$$

where $0 \leq r \leq k$.

4.3. **Postprocessing for u: The case $k = 1$.** We present the following postprocessing technique for $k \geq 1$. However, since we already have an optimal postprocessing technique for $k \geq 2$ we would numerically only use this one for $k = 1$. We define the post-processing approximation $u_h^\star|_K \in \mathcal{P}^{k+2}(K)$

$$(\nabla u_h^\star, \nabla v)_K = (\boldsymbol{q}_h^\star, \nabla v)_K \qquad \text{for all } v \in \mathcal{P}_\perp^{k+2,0}(K), \tag{4.31a}$$

$$(u_h^\star, w)_K = (u_h, w)_K \qquad \text{for all } w \in \mathcal{P}^0(K), \tag{4.31b}$$

for all $K \in \mathcal{T}_h$. Here $\boldsymbol{q}_h^\star$ is defined in (4.30).

The proof of the next theorem is similar to the proof of Theorem 4.2 so we omit the details.

Theorem 4.5. *The solution $u_h^\star$ of* (4.31) *is well defined. Moreover,*

$$\|u - u_h^\star\|_{L^2(K)} \leq C\, h_K^{r+3} \|u\|_{H^{r+3}(K)} + C\, h_K \|\boldsymbol{q} - \boldsymbol{q}_h^\star\|_{L^2(K)} + C\, \|P^0(u - u_h)\|_{L^2(K)}.$$

for $0 \leq r \leq k$.

The following corollary is a simple consequence of the above theorem, Theorem 3.6 and Corollary 4.4.

Corollary 4.6. *Assume the hypotheses of Theorem 3.6 hold. Moreover, assume the mesh $\mathcal{T}_h$ is quasi-uniform. Then, for $k \geq 1$*

$$\|u - u_h^\star\|_{L^2(\Omega)} \leq C\, h^{r+3} (\|f\|_{H^{r+1}(\Omega)} + \|\underline{\boldsymbol{z}}\|_{H^{r+1}(\Omega)}),$$

for $0 \leq r \leq k$.

4.4. **Postprocessing for u: The case** $k = 0$**.** Since $P^0(u - u_h)$ does not converge with order $k+3$ for $k = 0$ we cannot take advantage of the post-processing technique defined in the previous section. Instead we define the following simple postprocessed approximation. For $k = 0$ find $u_h^\star|_K \in \mathcal{P}^1(K)$ that solves

$$(\nabla u_h^\star, \nabla v)_K = (\boldsymbol{q}_h, \nabla v)_K \quad \text{for all } v \in \mathcal{P}_\perp^{1,0}(K), \tag{4.32a}$$

$$(u_h^\star, w)_K = (u_h, w)_K \quad \text{for all } w \in \mathcal{P}^0(K), \tag{4.32b}$$

We can easily prove the following result.

Theorem 4.7. *The approximation* $u_h^\star$ *defined by* (4.32) *is well defined. Moreover, if we assume the hypothesis of Theorem 3.7 we have that*

$$\|u - u_h^\star\|_{L^2(\Omega)} \leq Ch^2(\|\underline{\boldsymbol{z}}\|_{H^1(\Omega)} + \|f\|_{H^1(\Omega)}).$$

5. Implementation issues: Hybrid Form

In this section we introduce a more efficient way of implementing (2.2a). We do this by introducing the hybrid form of (2.2) which is done by relaxing the continuity requirements of the spaces $\boldsymbol{\Sigma}_h$ and $\underline{\boldsymbol{Z}}_h$. As a consequence Lagrange multipliers are introduced that approximate u and $\boldsymbol{q}$ on the interfaces of $\mathcal{T}_h$. We then show that the only globally coupled degrees of freedom are those associated with the Lagrange multipliers. Moreover, the other variables can be recovered element by element. We note that using the hybrid form of mixed methods has been well studied; see for example [25, 15, 7, 5, 1, 12]. We follow more closely the notation used in [12].

Here we write the hybrid form of (2.2). In order to do so, we define the following non-conforming versions of $\boldsymbol{\Sigma}_h$ and $\underline{\boldsymbol{Z}}_h$.

$$\tilde{\boldsymbol{\Sigma}}_h := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{v}|_K \in \boldsymbol{RT}^k(K) \text{ for all } K \in \mathcal{T}_h\},$$

$$\tilde{\underline{\boldsymbol{Z}}}_h := \{\underline{\boldsymbol{s}} \in \underline{\boldsymbol{L}}^2(\Omega) : \text{ each row of } \underline{\boldsymbol{s}} \text{ belongs to } \tilde{\boldsymbol{\Sigma}}_h\}.$$

We also need to define the Lagrange multiplier spaces

$$M_h := \{\mu : \mu|_F \in \mathcal{P}^k(F) \text{ for all faces of } F \text{ of } \mathcal{T}_h, \mu = 0 \text{ on } \partial\Omega\},$$

$$\boldsymbol{M}_h := \{\boldsymbol{\mu} : \boldsymbol{\mu}|_F \in \boldsymbol{\mathcal{P}}^k(F) \text{ for all faces of } F \text{ of } \mathcal{T}_h, \boldsymbol{\mu} = 0 \text{ on } \partial\Omega\}.$$

The hybrid method finds $(\tilde{u}_h, \tilde{\boldsymbol{q}}_h, \tilde{\underline{\boldsymbol{z}}}_h, \tilde{\boldsymbol{\sigma}}_h, \lambda_h, \boldsymbol{\alpha}_h) \in W_h \times \boldsymbol{Q}_h \times \tilde{\underline{\boldsymbol{Z}}}_h \times \tilde{\boldsymbol{\Sigma}}_h \times M_h \times \boldsymbol{M}_h$ that satisfy

$$(\tilde{\boldsymbol{q}}_h, \boldsymbol{v}) + (\tilde{u}_h, \nabla \cdot \boldsymbol{v}) - \langle \lambda_h, \boldsymbol{v} \cdot \boldsymbol{n} \rangle = 0, \tag{5.33a}$$

$$(\tilde{\underline{\boldsymbol{z}}}_h, \underline{\boldsymbol{s}}) + (\tilde{\boldsymbol{q}}_h, \nabla \cdot \underline{\boldsymbol{s}}) - \langle \boldsymbol{\alpha}_h, \underline{\boldsymbol{s}}\, \boldsymbol{n} \rangle = 0, \tag{5.33b}$$

$$-(\tilde{\boldsymbol{\sigma}}_h, \boldsymbol{m}) + (\boldsymbol{m}, \nabla \cdot \tilde{\underline{\boldsymbol{z}}}_h) = 0, \tag{5.33c}$$

$$(w, \nabla \cdot \tilde{\boldsymbol{\sigma}}_h) = (f, w), \tag{5.33d}$$

$$\langle \tilde{\boldsymbol{\sigma}}_h \cdot \boldsymbol{n}, \mu \rangle = 0, \tag{5.33e}$$

$$\langle \tilde{\underline{\boldsymbol{z}}}_h \boldsymbol{n}, \boldsymbol{\mu} \rangle = 0, \tag{5.33f}$$

for all $(w, \boldsymbol{m}, \underline{\boldsymbol{s}}, \boldsymbol{v}, \mu, \boldsymbol{\mu}) \in W_h \times \boldsymbol{Q}_h \times \tilde{\underline{\boldsymbol{Z}}}_h \times \tilde{\boldsymbol{\Sigma}}_h \times M_h \times \boldsymbol{M}_h$.

Here we used the notation

$$\langle \mu, \lambda \rangle := \sum_{K \in \mathcal{T}_h} \int_{\partial K} \mu(s)\, \lambda(s)\, ds.$$

We state a trivial but important result. We leave the details to the reader.

Theorem 5.1. *The problem* (5.33) *is well defined. Moreover, let* $(\tilde{u}_h, \tilde{\boldsymbol{q}}_h, \tilde{\underline{\boldsymbol{z}}}_h, \tilde{\boldsymbol{\sigma}}_h, \lambda_h, \boldsymbol{\alpha}_h) \in W_h \times \boldsymbol{Q}_h \times \tilde{\underline{\boldsymbol{Z}}}_h \times \tilde{\Sigma}_h \times M_h \times \boldsymbol{M}_h$ *be the solution to* (5.33), *then* $(\tilde{u}_h, \tilde{\boldsymbol{q}}_h, \tilde{\underline{\boldsymbol{z}}}_h, \tilde{\boldsymbol{\sigma}}_h) = (u_h, \boldsymbol{q}_h, \underline{\boldsymbol{z}}_h, \boldsymbol{\sigma}_h)$ *where* $(u_h, \boldsymbol{q}_h, \underline{\boldsymbol{z}}_h, \boldsymbol{\sigma}_h)$ *is the solution to* (2.2).

In order to see the advantage of using the hybrid formulation (5.33) we need to introduce local solvers. First for $m \in M_h$ let $(\mathsf{u}_1(m), \mathbf{Q}_1(m), \underline{\mathbf{Z}}_1(m), \mathbf{S}_1(m)) \in W_h \times \boldsymbol{Q}_h \times \tilde{\underline{\boldsymbol{Z}}}_h \times \tilde{\Sigma}_h$ solve

$$\begin{aligned}
(\mathbf{Q}_1(m), \boldsymbol{v}) + (\mathsf{u}_1(m), \nabla\cdot\boldsymbol{v}) &= \langle m, \boldsymbol{v}\cdot\boldsymbol{n}\rangle, && (5.34a)\\
(\underline{\mathbf{Z}}_1(m), \underline{\boldsymbol{s}}) + (\mathbf{Q}_1(m), \nabla\cdot\underline{\boldsymbol{s}}) &= 0, && (5.34b)\\
-(\mathbf{S}_1(m), \boldsymbol{m}) + (\boldsymbol{m}, \nabla\cdot\underline{\mathbf{Z}}_1(m)) &= 0, && (5.34c)\\
(w, \nabla\cdot\mathbf{S}_1(m)) &= 0, && (5.34d)
\end{aligned}$$

for all $(w, \boldsymbol{m}, \underline{\boldsymbol{s}}, \boldsymbol{v}) \in W_h \times \boldsymbol{Q}_h \times \tilde{\underline{\boldsymbol{Z}}}_h \times \tilde{\Sigma}_h$.

Similarly, for $\boldsymbol{\mu} \in \boldsymbol{M}_h$ let $(\mathsf{u}_2(\boldsymbol{\mu}), \mathbf{Q}_2(\boldsymbol{\mu}), \underline{\mathbf{Z}}_2(\boldsymbol{\mu}), \mathbf{S}_2(\boldsymbol{\mu})) \in W_h \times \boldsymbol{Q}_h \times \tilde{\underline{\boldsymbol{Z}}}_h \times \tilde{\Sigma}_h$ solve

$$\begin{aligned}
(\mathbf{Q}_2(\boldsymbol{\mu}), \boldsymbol{v}) + (\mathsf{u}_2(\boldsymbol{\mu}), \nabla\cdot\boldsymbol{v}) &= 0, && (5.35a)\\
(\underline{\mathbf{Z}}_2(\boldsymbol{\mu}), \underline{\boldsymbol{s}}) + (\mathbf{Q}_2(\boldsymbol{\mu}), \nabla\cdot\underline{\boldsymbol{s}}) &= \langle \boldsymbol{\mu}, \underline{\boldsymbol{s}}\,\boldsymbol{n}\rangle, && (5.35b)\\
-(\mathbf{S}_2(\boldsymbol{\mu}), \boldsymbol{m}) + (\boldsymbol{m}, \nabla\cdot\underline{\mathbf{Z}}_2(\boldsymbol{\mu})) &= 0, && (5.35c)\\
(w, \nabla\cdot\mathbf{S}_2(\boldsymbol{\mu})) &= 0, && (5.35d)
\end{aligned}$$

for all $(w, \boldsymbol{m}, \underline{\boldsymbol{s}}, \boldsymbol{v}) \in W_h \times \boldsymbol{Q}_h \times \tilde{\underline{\boldsymbol{Z}}}_h \times \tilde{\Sigma}_h$.

Finally, let $(\mathsf{u}_3(f), \mathbf{Q}_3(f), \underline{\mathbf{Z}}_3(f), \mathbf{S}_3(f)) \in W_h \times \boldsymbol{Q}_h \times \tilde{\underline{\boldsymbol{Z}}}_h \times \tilde{\Sigma}_h$ solve

$$\begin{aligned}
(\mathbf{Q}_3(f), \boldsymbol{v}) + (\mathsf{u}_3(f), \nabla\cdot\boldsymbol{v}) &= 0, && (5.36a)\\
(\underline{\mathbf{Z}}_3(f), \underline{\boldsymbol{s}}) + (\mathbf{Q}_3(f), \nabla\cdot\underline{\boldsymbol{s}}) &= 0, && (5.36b)\\
-(\mathbf{S}_3(f), \boldsymbol{m}) + (\boldsymbol{m}, \nabla\cdot\underline{\mathbf{Z}}_3(f)) &= 0, && (5.36c)\\
(w, \nabla\cdot\mathbf{S}_3(f)) &= (f, w), && (5.36d)
\end{aligned}$$

for all $(w, \boldsymbol{m}, \underline{\boldsymbol{s}}, \boldsymbol{v}) \in W_h \times \boldsymbol{Q}_h \times \tilde{\underline{\boldsymbol{Z}}}_h \times \tilde{\Sigma}_h$.

The important fact is that the local solvers are defined locally on each element since the spaces $W_h \times \boldsymbol{Q}_h \times \tilde{\underline{\boldsymbol{Z}}}_h \times \tilde{\Sigma}_h$ are completely discontinuous. It is simple to show that they are well defined.

Now that we have the local solvers we define three bilinear forms. For $m, \mu \in M_h$ and $\boldsymbol{\mu}, \boldsymbol{r} \in \boldsymbol{M}_h$ define

$$\begin{aligned}
a(m, \mu) &:= (\underline{\mathbf{Z}}_1(m), \underline{\mathbf{Z}}_1(\mu)),\\
c(\boldsymbol{\mu}, \boldsymbol{r}) &:= (\underline{\mathbf{Z}}_2(\boldsymbol{\mu}), \underline{\mathbf{Z}}_2(\boldsymbol{r})),\\
b(m, \boldsymbol{\mu}) &:= (\underline{\mathbf{Z}}_1(m), \underline{\mathbf{Z}}_2(\boldsymbol{\mu})).
\end{aligned}$$

We next define another problem which will allow us to find the Lagrange multipliers λ_h and $\boldsymbol{\alpha}_h$.

Let $(\tilde{\lambda}_h, \tilde{\boldsymbol{\alpha}}_h) \in M_h \times \boldsymbol{M}_h$ solve

$$a(\tilde{\lambda}_h, m) + b(m, \tilde{\boldsymbol{\alpha}}_h) =(f, \mathsf{u}_1(m)) \tag{5.37a}$$

$$b(\tilde{\lambda}_h, \boldsymbol{\mu}) + c(\tilde{\boldsymbol{\alpha}}_h, \boldsymbol{\mu}) =(f, \mathsf{u}_2(\boldsymbol{\mu})), \tag{5.37b}$$

for all $(m, \boldsymbol{\mu}) \in M_h \times \boldsymbol{M}_h$.

We next prove that this problem is well defined.

Theorem 5.2. *The problem* (5.37) *is well defined.*

Proof. Since (5.37) is a square system we need to show uniqueness, so we let $f = 0$. If we let $m = \tilde{\lambda}_h$ and $\boldsymbol{\mu} = \tilde{\boldsymbol{\alpha}}_h$ and add the two equations (5.37a) and (5.37b) we get

$$a(\tilde{\lambda}_h, \tilde{\lambda}_h) + 2\, b(\tilde{\lambda}_h, \tilde{\boldsymbol{\alpha}}_h) + c(\tilde{\boldsymbol{\alpha}}_h, \tilde{\boldsymbol{\alpha}}_h) = 0.$$

Using the definition of the bilinear forms a, b, c, this gives exactly

$$\|\underline{\mathbf{Z}}_1(\tilde{\lambda}_h) + \underline{\mathbf{Z}}_2(\tilde{\boldsymbol{\alpha}}_h)\|^2_{L^2(\Omega)} = 0,$$

which in turn gives us that

$$\underline{\mathbf{Z}}_1(\tilde{\lambda}_h) + \underline{\mathbf{Z}}_2(\tilde{\boldsymbol{\alpha}}_h) = 0. \tag{5.38}$$

Next we show that this implies that $\tilde{\lambda}_h = 0$ and $\tilde{\boldsymbol{\alpha}}_h = 0$. Indeed, by (5.34b), (5.35b) (5.38) we have that

$$(\mathbf{Q}_1(\tilde{\lambda}_h) + \mathbf{Q}_2(\tilde{\boldsymbol{\alpha}}_h), \nabla \cdot \underline{\boldsymbol{s}}) = \langle \tilde{\boldsymbol{\alpha}}_h, \underline{\boldsymbol{s}}\, \boldsymbol{n}\rangle,$$

In an integration by parts gives us that

$$-(\nabla(\mathbf{Q}_1(\tilde{\lambda}_h) + \mathbf{Q}_2(\tilde{\boldsymbol{\alpha}}_h)), \underline{\boldsymbol{s}}) = \langle \tilde{\boldsymbol{\alpha}}_h - (\mathbf{Q}_1(\tilde{\lambda}_h) + \mathbf{Q}_2(\tilde{\boldsymbol{\alpha}}_h)), \underline{\boldsymbol{s}}\, \boldsymbol{n}\rangle,$$

for all $\underline{\boldsymbol{s}} \in \underline{\tilde{\boldsymbol{Z}}}_h$. Using the degrees of freedom of the Raviart-Thomas space $\underline{\tilde{\boldsymbol{Z}}}_h$ we can easily show that $\nabla(\mathbf{Q}_1(\tilde{\lambda}_h) + \mathbf{Q}_2(\tilde{\boldsymbol{\alpha}}_h)) = 0$ on Ω and that $\tilde{\boldsymbol{\alpha}}_h = (\mathbf{Q}_1(\tilde{\lambda}_h) + \mathbf{Q}_2(\tilde{\boldsymbol{\alpha}}_h))$ on the faces of $\mathcal{T}_h$. This implies that $\mathbf{Q}_1(\tilde{\lambda}_h) + \mathbf{Q}_2(\tilde{\boldsymbol{\alpha}}_h)$ is constant on Ω, and since it vanishes on $\partial\Omega$ because it is equal to $\tilde{\boldsymbol{\alpha}}_h$ there, it must be identically zero. This in turn implies that $\tilde{\boldsymbol{\alpha}}_h$ is identically zero. Now that we have that $\mathbf{Q}_1(\tilde{\lambda}_h) + \mathbf{Q}_2(\tilde{\boldsymbol{\alpha}}_h) = 0$ we can combine (5.34b) and (5.35a) to show that $\tilde{\lambda}_h = 0$ in a similar way. □

We can now prove the main result of this section.

Theorem 5.3. *Let* $(\tilde{\lambda}_h, \tilde{\boldsymbol{\alpha}}_h) \in M_h \times \boldsymbol{M}_h$ *solve* (5.37), *then* $(\lambda_h, \boldsymbol{\alpha}_h) = (\tilde{\lambda}_h, \tilde{\boldsymbol{\alpha}}_h)$ *where* $(\tilde{u}_h, \tilde{\boldsymbol{q}}_h, \tilde{\underline{\boldsymbol{z}}}_h, \tilde{\underline{\boldsymbol{\sigma}}}_h, \lambda_h, \boldsymbol{\alpha}_h) \in W_h \times \boldsymbol{Q}_h \times \underline{\tilde{\boldsymbol{Z}}}_h \times \tilde{\boldsymbol{\Sigma}}_h \times M_h \times \boldsymbol{M}_h$ *solves* (5.33). *Moreover,*

$$\begin{aligned}
\tilde{u}_h &=\mathsf{u}_1(\tilde{\lambda}_h) + \mathsf{u}_2(\tilde{\boldsymbol{\alpha}}_h) + \mathsf{u}_3(f),\\
\tilde{\boldsymbol{q}}_h &=\mathbf{Q}_1(\tilde{\lambda}_h) + \mathbf{Q}_2(\tilde{\boldsymbol{\alpha}}_h) + \mathbf{Q}_3(f),\\
\tilde{\underline{\boldsymbol{z}}}_h &=\underline{\mathbf{Z}}_1(\tilde{\lambda}_h) + \underline{\mathbf{Z}}_2(\tilde{\boldsymbol{\alpha}}_h) + \underline{\mathbf{Z}}_3(f),\\
\tilde{\underline{\boldsymbol{\sigma}}}_h &=\mathbf{S}_1(\tilde{\lambda}_h) + \mathbf{S}_2(\tilde{\boldsymbol{\alpha}}_h) + \mathbf{S}_3(f).
\end{aligned}$$

Proof. Let us use the notation

$$\begin{aligned}
\mathsf{u} &=\mathsf{u}_1(\tilde{\lambda}_h) + \mathsf{u}_2(\tilde{\boldsymbol{\alpha}}_h) + \mathsf{u}_3(f),\\
\mathbf{Q} &=\mathbf{Q}_1(\tilde{\lambda}_h) + \mathbf{Q}_2(\tilde{\boldsymbol{\alpha}}_h) + \mathbf{Q}_3(f),\\
\underline{\mathbf{Z}} &=\underline{\mathbf{Z}}_1(\tilde{\lambda}_h) + \underline{\mathbf{Z}}_2(\tilde{\boldsymbol{\alpha}}_h) + \underline{\mathbf{Z}}_3(f),\\
\mathbf{S} &=\mathbf{S}_1(\tilde{\lambda}_h) + \mathbf{S}_2(\tilde{\boldsymbol{\alpha}}_h) + \mathbf{S}_3(f).
\end{aligned}$$

Then using the definition of the local solvers, we easily can show that $(\mathsf{u}, \mathbf{Q}, \underline{\mathbf{Z}}, \mathbf{S}, \tilde{\lambda}_h, \tilde{\boldsymbol{\alpha}}_h)$ satisfies (5.33a)-(5.33d). Hence, by the uniqueness of (5.33) it is enough to show that

$$\begin{aligned}\langle \mathbf{S} \cdot \boldsymbol{n}, \mu \rangle =& 0,\\ \langle \underline{\mathbf{Z}}\, \boldsymbol{n}, \boldsymbol{\mu} \rangle =& 0.\end{aligned}$$

In other words, if $(\tilde{\lambda}_h, \tilde{\boldsymbol{\alpha}}_h)$ solves (5.37) we need to show that

$$\begin{aligned}\langle \Big(\mathbf{S}_1(\tilde{\lambda}_h) + \mathbf{S}_2(\tilde{\boldsymbol{\alpha}}_h) + \mathbf{S}_3(f)\Big) \cdot \boldsymbol{n}, \mu \rangle =& 0,\\ \langle \Big(\underline{\mathbf{Z}}_1(\tilde{\lambda}_h) + \underline{\mathbf{Z}}_2(\tilde{\boldsymbol{\alpha}}_h) + \underline{\mathbf{Z}}_3(f)\Big)\, \boldsymbol{n}, \boldsymbol{\mu} \rangle =& 0.\end{aligned}$$

for all $(\mu, \boldsymbol{\mu}) \in M_h \times \boldsymbol{M}_h$.

This in turn follows from the following identities

$$\langle \mathbf{S}_3(f) \cdot \boldsymbol{n}, \mu \rangle = (f, \mathsf{u}_1(\mu)), \tag{5.39a}$$
$$\langle \underline{\mathbf{Z}}_3(f)\, \boldsymbol{n}, \boldsymbol{\mu} \rangle = -(f, \mathsf{u}_2(\boldsymbol{\mu})), \tag{5.39b}$$
$$\langle \mathbf{S}_1(m) \cdot \boldsymbol{n}, \mu \rangle = -(\underline{\mathbf{Z}}_1(m), \underline{\mathbf{Z}}_1(\mu)), \tag{5.39c}$$
$$\langle \underline{\mathbf{Z}}_2(\boldsymbol{\mu})\, \boldsymbol{n}, \boldsymbol{r} \rangle = (\underline{\mathbf{Z}}_2(\boldsymbol{\mu}), \underline{\mathbf{Z}}_2(\boldsymbol{r})), \tag{5.39d}$$
$$\langle \mathbf{S}_2(\boldsymbol{\mu}) \cdot \boldsymbol{n}, \mu \rangle = -(\underline{\mathbf{Z}}_2(\boldsymbol{\mu}), \underline{\mathbf{Z}}_1(\mu)), \tag{5.39e}$$
$$\langle \underline{\mathbf{Z}}_1(\mu)\, \boldsymbol{n}, \boldsymbol{\mu} \rangle = (\underline{\mathbf{Z}}_2(\boldsymbol{\mu}), \underline{\mathbf{Z}}_1(\mu)), \tag{5.39f}$$

which hold for all $m, \mu \in M_h$ and $\boldsymbol{\mu}, \boldsymbol{r} \in \boldsymbol{M}_h$.

Since the proof of the above identities are similar we only prove (5.39a), (5.39c) and (5.39e). To this end,

$$\begin{aligned}(f, \mathsf{u}_1(\mu)) =& (\mathsf{u}_1(\mu), \nabla \cdot \mathbf{S}_3(f)) && \text{by (5.36d)}\\ =& -(\mathbf{Q}_1(\mu), \mathbf{S}_3(f)) + \langle \mu, \mathbf{S}_3(f) \cdot \boldsymbol{n} \rangle && \text{by (5.34a)}\\ =& -(\mathbf{Q}_1(\mu), \nabla \cdot \underline{\mathbf{Z}}_3(f)) + \langle \mu, \mathbf{S}_3(f) \cdot \boldsymbol{n} \rangle && \text{by (5.36c)}\\ =& (\underline{\mathbf{Z}}_1(\mu), \underline{\mathbf{Z}}_3(f)) + \langle \mu, \mathbf{S}_3(f) \cdot \boldsymbol{n} \rangle && \text{by (5.34b)}\\ =& -(\mathbf{Q}_3(f), \nabla \cdot \underline{\mathbf{Z}}_1(\mu)) + \langle \mu, \mathbf{S}_3(f) \cdot \boldsymbol{n} \rangle && \text{by (5.36b)}\\ =& -(\mathbf{Q}_3(f), \mathbf{S}_1(\mu)) + \langle \mu, \mathbf{S}_3(f) \cdot \boldsymbol{n} \rangle && \text{by (5.34c)}\\ =& (\mathsf{u}_3(f), \nabla \cdot \mathbf{S}_1(\mu)) + \langle \mu, \mathbf{S}_3(f) \cdot \boldsymbol{n} \rangle && \text{by (5.34c)}\\ =& \langle \mu, \mathbf{S}_3(f) \cdot \boldsymbol{n} \rangle && \text{by (5.34d).}\end{aligned}$$

This proves (5.39a).

Also,

$$\begin{aligned}(\underline{\mathbf{Z}}_1(m), \underline{\mathbf{Z}}_1(\mu)) =& -(\mathbf{Q}_1(\mu), \nabla \cdot \underline{\mathbf{Z}}_1(m)) && \text{by (5.34b)}\\ =& -(\mathbf{Q}_1(\mu), \mathbf{S}_1(m)) && \text{by (5.34c)}\\ =& (\mathsf{u}_1(\mu), \nabla \cdot (\mathbf{S}_1(m))) - \langle \mu, \mathbf{S}_1(m) \cdot \boldsymbol{n} \rangle && \text{by (5.34a)}\\ =& -\langle \mu, \mathbf{S}_1(m) \cdot \boldsymbol{n} \rangle && \text{by (5.34d).}\end{aligned}$$

This proves (5.39c).

Next we prove (5.39e).

$$
\begin{aligned}
(\underline{\mathbf{Z}}_2(\boldsymbol{\mu}), \underline{\mathbf{Z}}_1(\mu)) &= -(\mathbf{Q}_1(\mu), \nabla \cdot \underline{\mathbf{Z}}_2(\boldsymbol{\mu})) && \text{by (5.34b)}\\
&= -(\mathbf{Q}_1(\mu), \mathbf{S}_2(\boldsymbol{\mu})) && \text{by (5.35c)}\\
&= (\mathsf{u}_1(\mu), \nabla \cdot \mathbf{S}_2(\boldsymbol{\mu})) - \langle \mu, \mathbf{S}_2(\boldsymbol{\mu}) \cdot \boldsymbol{n} \rangle && \text{by (5.34a)}\\
&= -\langle \mu, \mathbf{S}_2(\boldsymbol{\mu}) \cdot \boldsymbol{n} \rangle && \text{by (5.35d).}
\end{aligned}
$$

This proves (5.39e). Since the proof of the other identities are similar we omit the details. This completes the proof of theorem.

□

6. Numerical Experiments

In this section we provide numerical experiments that validate our theoretical results. We chose $\Omega = [0,1] \times [0,1]$ and f so that the solution $u(x,y) = u = 10\,(y-1)^3\,y^3\,(x-1)^2\,x^2$. We provide the results for uniform meshes and we denote the mesh with mesh size $\frac{1}{2^i}$ simply by mesh i; see Figure 1 for an example. Table 1 shows that the order of convergence of the errors for u, $\boldsymbol{q}$ and $\underline{\boldsymbol{z}}$ are optimal and that the order of convergence for the error in $\boldsymbol{\sigma}$ is sub-optimal. In Table 2 we see that $\|P^{k-1}(u-u_h)\|_{L^2(\Omega)}$ converges order $k+3$ while $\|P(u-u_h)\|_{L^2(\Omega)}$ converges with order $k+2$. Moreover, we see that $\|\boldsymbol{P}(\boldsymbol{q}-\boldsymbol{q}_h)\|_{L^2(\Omega)}$ converges with order $k+2$ for $k \geq 0$. Note, however, that we were only able to prove this for $k \geq 1$; see Corollary 3.10. Also, in Table 2 we display the order of convergence for $\|\boldsymbol{q}-\boldsymbol{q}_h^\star\|_{L^2(\Omega)}$ and $\|u-u_h^\star\|_{L^2(\Omega)}$. As one can see from Table 2 the order of convergence for $\|\boldsymbol{q}-\boldsymbol{q}_h^\star\|_{L^2(\Omega)}$ is $k+2$. The approximation $u_h^\star$ we calculated is given by (4.29) for $k=2$, by (4.31) for $k=1$, and (4.32) for $k=0$. Table 2 shows that $\|u-u_h^\star\|_{L^2(\Omega)}$ converges with order $k+3$ for $k=1,2$ and $k+2$ for $k=0$ just as the theory predicted.

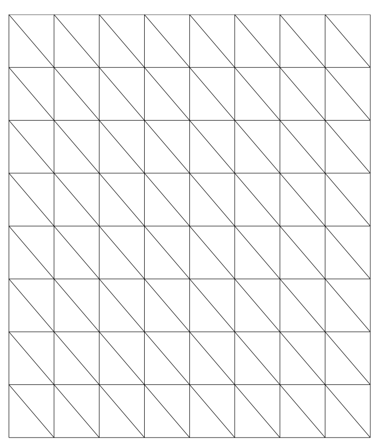

Figure 1. Mesh 3, $h = \frac{1}{2^3}$

Table 1. History of convergence of the errors

	mesh	$\|u - u_h\|_{L^2(\Omega)}$		$\|\boldsymbol{q} - \boldsymbol{q}_h\|_{L^2(\Omega)}$		$\|\underline{\boldsymbol{z}} - \underline{\boldsymbol{z}}_h\|_{L^2(\Omega)}$		$\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{L^2(\Omega)}$	
k	i	error	order	error	order	error	order	error	order
0	1	.41e-2	0.00	.21e-1	0.00	.16e+0	0.00	.14e+1	0.00
	2	.15e-2	1.44	.93e-2	1.20	.80e-1	0.97	.89e+0	0.63
	3	.62e-3	1.27	.44e-2	1.09	.41e-1	0.95	.49e+0	0.88
	4	.29e-3	1.10	.21e-2	1.03	.21e-1	0.96	.25e+0	0.96
	5	.14e-3	1.03	.11e-2	1.01	.11e-1	0.99	.13e+0	0.96
	6	.70e-4	1.01	.53e-3	1.00	.54e-2	1.00	.67e-1	0.93
	7	.35e-4	1.00	.26e-3	1.00	.27e-2	1.00	.37e-1	0.88
	8	.18e-4	1.00	.13e-3	1.00	.13e-2	1.00	.21e-1	0.81
	9	.88e-5	1.00	.66e-4	1.00	.67e-3	1.00	.13e-1	0.72
1	1	.13e-2	0.00	.86e-2	0.00	.63e-1	0.00	.69e+0	0.00
	2	.28e-3	2.20	.22e-2	1.97	.20e-1	1.65	.26e+0	1.37
	3	.68e-4	2.04	.62e-3	1.84	.55e-2	1.86	.12e+0	1.15
	4	.17e-4	1.99	.16e-3	1.95	.14e-2	1.96	.60e-1	1.00
	5	.43e-5	1.99	.40e-4	1.99	.36e-3	1.99	.30e-1	0.99
	6	.11e-5	2.00	.10e-4	2.00	.90e-4	1.99	.15e-1	0.99
	7	.27e-6	2.00	.25e-5	2.00	.23e-4	2.00	.76e-2	0.99
2	1	.36e-3	0.00	.34e-2	0.00	.20e-1	0.00	.28e+0	0.00
	2	.51e-4	2.81	.55e-3	2.63	.37e-2	2.41	.67e-1	2.06
	3	.71e-5	2.85	.78e-4	2.81	.55e-3	2.76	.17e-1	2.02
	4	.92e-6	2.95	.10e-4	2.95	.73e-4	2.92	.41e-2	2.03
	5	.12e-6	2.99	.13e-5	2.99	.92e-5	2.97	.99e-3	2.04
	6	.15e-7	3.00	.16e-6	3.00	.12e-5	2.99	.24e-3	2.04

7. Conclusions

We have developed a method that approximates u, $\boldsymbol{q}$ and $\underline{\boldsymbol{z}}$ with optimal order. Moreover, we used postprocessed approximations which can be calculated locally on each element which converge with order $k+3$ for $k \geq 1$ and $k+2$ for $k=0$.

A natural question is if we can use other spaces such as the Brezzi-Marini-Douglas spaces (or Brezzi-Douglas-Durán-Fortin in 3d) spaces [5, 6]. For instance, we can define the spaces for $k \geq 1$ as follows

$$\begin{aligned}
W_h :=&\{w \in L^2(\Omega) : w|_K \in \mathcal{P}^{k-1}(K), \text{ for all } K \in \mathcal{T}_h\},\\
\boldsymbol{Q}_h :=&\{\boldsymbol{m} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{m}|_K \in \boldsymbol{\mathcal{P}}^{k-1}(K), \text{ for all } K \in \mathcal{T}_h\},\\
\boldsymbol{\Sigma}_h :=&\{\boldsymbol{v} \in \boldsymbol{H}(\text{div}\,,\Omega) : \boldsymbol{v}|_K \in \boldsymbol{\mathcal{P}}^{k}(K) \text{ for all } K \in \mathcal{T}_h\},\\
\underline{\boldsymbol{Z}}_h :=&\{\underline{\boldsymbol{s}} \in \underline{\boldsymbol{H}}(\text{div}\,,\Omega) : \text{ each row of } \underline{\boldsymbol{s}} \text{ belongs to } \boldsymbol{\Sigma}_h\}.
\end{aligned}$$

The method (2.2) with these spaces are well-defined. However, we cannot prove optimal error estimates for $\underline{\boldsymbol{z}}$ with these spaces.

In a forthcoming paper we extend the methodology of this paper to the Reissner-Mindlin plate problem.

TABLE 2. History of convergence of projections of errors and postprocessing

	mesh	$\|P^{k-1}(u-u_h)\|_{L^2(\Omega)}$		$\|P(u-u_h)\|_{L^2(\Omega)}$		$\|\boldsymbol{P}(\boldsymbol{q}-\boldsymbol{q}_h)\|_{L^2(\Omega)}$		$\|u-u_h^\star\|_{L^2(\Omega)}$		$\|q-q_h^\star\|_{L^2(\Omega)}$	
k	i	error	order	error	order	error	order	error	order	error	order
0	1	–	–	.33e-2	–	.18e-1	–	.42e-2	–	.22e-1	–
	2	–	–	.10e-2	1.66	.48e-2	1.91	.11e-2	1.90	.66e-2	1.75
	3	–	–	.28e-3	1.92	.14e-2	1.82	.29e-3	1.97	.19e-2	1.77
	4	–	–	.70e-4	1.97	.37e-3	1.90	.73e-4	1.98	.52e-3	1.89
	5	–	–	.18e-4	1.99	.93e-4	1.97	.18e-4	2.00	.13e-3	1.97
	6	–	–	.44e-5	2.00	.23e-4	1.99	.46e-5	2.00	.33e-4	1.99
	7	–	–	.11e-5	2.00	.59e-5	2.00	.11e-5	2.00	.84e-5	2.00
	8	–	–	.28e-6	2.00	.15e-5	2.00	.29e-6	2.00	.21e-5	2.00
	9	–	–	.69e-7	2.00	.37e-6	2.00	.72e-7	2.00	.52e-6	2.00
1	1	.80e-3	–	.13e-2	–	.59e-2	–	.11e-2	–	.67e-2	–
	2	.82e-4	3.29	.13e-3	3.33	.68e-3	3.12	.93e-4	3.52	.95e-3	2.80
	3	.63e-5	3.70	.12e-4	3.51	.75e-4	3.19	.68e-5	3.76	.13e-3	2.93
	4	.42e-6	3.91	.11e-5	3.41	.85e-5	3.14	.45e-6	3.93	.16e-4	2.99
	5	.27e-7	3.98	.12e-6	3.18	.10e-5	3.05	.28e-7	3.98	.20e-5	3.00
	6	.17e-8	4.00	.14e-7	3.06	.13e-6	3.01	.18e-8	4.00	.25e-6	3.00
	7	.10e-9	4.02	.18e-8	3.01	.16e-7	3.00	.11e-9	4.02	.31e-7	3.00
2	1	.17e-3	–	.25e-3	–	.11e-2	–	.16e-3	–	.16e-2	–
	2	.56e-5	4.93	.12e-4	4.42	.59e-4	4.25	.68e-5	4.58	.14e-3	3.54
	3	.12e-6	5.55	.69e-6	4.09	.37e-5	4.00	.20e-6	5.08	.91e-5	3.90
	4	.24e-8	5.65	.44e-7	3.98	.24e-6	3.94	.61e-8	5.05	.58e-6	3.96
	5	.56e-10	5.42	.27e-8	3.99	.16e-7	3.95	.19e-9	5.01	.37e-7	3.99
	6	.17e-11	5.04	.17e-9	4.00	.10e-8	3.97	.59e-11	5.00	.23e-8	4.00

REFERENCES

[1] D.N. Arnold and F. Brezzi, *Mixed and nonconforming finite element methods: implementation, post-processing and error estimates*, RAIRO Modél. Math. Anal. Numér. 19 (1985), no. 1, 7–32.

[2] I. Babuska, J. Osborn, J. Pitkäranta, *Analysis of mixed methods using mesh dependent norms*, Math. Comp. 35 (1980), no. 152, 1039–1062.

[3] M. Amara, C. Capatina-Papaghiuc, A. Chatti, *Bending moment mixed method for the Kirchhoff-Love plate model* SIAM J. Numer. Anal. 40 (2002), no. 5, 1632–1649.

[4] H. Blum and R. Rannacher, *On the boundary value problem of the biharmonic operator on domains with angular corners*, Math. Methods Appl. Sci. 2 (1980), no. 4, 556–581.

[5] F. Brezzi, J. Douglas and D. Marini, *Two families of mixed finite elements for second order elliptic problems*, Numer. Math., 47 (1985), 217-235.

[6] F. Brezzi, J. Douglas, R. Durán and M. Fortin, *Mixed finite elements for second order elliptic problems in three variables*, Numer. Math. 51 (1987), no. 2, 237–250.

[7] F. Brezzi and M. Fortin, *Mixed and hybrid finite element methods*, Springer Series in Computational Mathematics, 15. Springer-Verlag, New York, 1991.

[8] P.G. Ciarlet and P.-A. Raviart, *A mixed finite element method for the biharmonic equation*, In: Mathematical Aspects of Finite Elements in Partial Differential Equations. Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, 1974, pp. 125145 (1974).

[9] B. Cockburn, B. Dong, J. Guzmán, *A hybridizable and superconvergent discontinuous Galerkin method for biharmonic problems*, J. Sci. Comput. 40 (2009), no. 1-3, 141–187.

[10] B. Cockburn, B. Dong and J. Guzmán, *A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems*, Math. Comp. 77 (2008), no. 264, 1887–1916.

[11] B. Cockburn, B. Dong, J. Guzmán, *A superconvergent LDG-hybridizable Galerkin method for second-order elliptic problems*, Math. Comp. 77 (2008), no. 264, 1887–1916.

[12] B. Cockburn and J. Gopalakrishnan, *A characterization of hybridized mixed methods for second order elliptic problems* SIAM J. Numer. Anal. 42 (2004), no. 1, 283–301.

[13] B. Cockburn, J. Gopalakrishnan and R. Lazarov, *Unified hybridization of discontinuous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*, SIAM J. Numer. Anal. 47 (2009), no. 2, 1319–1365.

[14] B. Cockburn, J. Guzmán and H. Wang, *Superconvergent discontinuous Galerkin methods for secondorder elliptic problems*, Math. Comp., 78 (2009), 1-24.

[15] M.I. Comodi, *The Hellan-Herrmann-Johnson method: some new error estimates and postprocessing*, Math. Comp. 52 (1989), no. 185, 17–29.

[16] R.S. Falk, *Approximation of the biharmonic equation by a mixed finite element method*, SIAM J. Numer. Anal. 15 (1978), no. 3, 556-567.

[17] R. Glowinski and O. Pironneau, *Numerical methods for the first biharmonic equation and the two-dimensional Stokes problem*, SIAM Rev. 21(2), 167-212 (1979).

[18] C. Johnson, *On the convergence of a mixed finite-element method for plate bending problems* Numer. Math. 21 (1973), 43–62.

[19] J.-C. Nédélec, *Mixed finite elements in* R^3, Numer. Math. 35 (1980), no. 3, 315–341

[20] J.-C. Nédélec, *A new family of mixed finite elements in* R^3, Numer. Math. 50 (1986), no. 1, 57–81.

[21] P.-A. Raviart and J.M. Thomas *A mixed finite element method for 2nd order elliptic problems*, Mathematical aspects of finite element methods (Proc. Conf., Consiglio Naz. delle Ricerche (C.N.R.), Rome, 1975), pp. 292–315. Lecture Notes in Math., Vol. 606, Springer, Berlin, 1977.

[22] R. Scholz, *Interior error estimates for a mixed finite element method*, Numer. Funct. Anal. Optim. 1 (1979), no. 4, 415–429.

[23] R. Scholz, *A mixed method for 4th order problems using linear finite elements*, RAIRO Anal. Numér. 12 (1978), no. 1, 85–90.

[24] R. Stenberg, *Postprocessing schemes for some mixed finite elements*, RAIRO Modél. Math. Anal. Numér. 25 (1991), no. 1, 151–167.

[25] B. Fraejis de Veubeke, *Displacement and equilibrium models in the finite element method, in Stress Analysis*, O. C. Zienkiewicz and G. Holister, eds., Wiley, New York, 1965.

Departamento de Ingeniería Civil, Universidad Católica de la Santísima Concepción, Casilla 297, Concepción, Chile
E-mail address: ebehrens@ucsc.cl

Division of Applied Mathematics, Brown University, Providence, RI 02912
E-mail address: johnny_guzman@brown.edu