SUPPLEMENTARY MATERIALS: THE HELLAN–HERRMANN–JOHNSON METHOD WITH CURVED ELEMENTS

DOUGLAS N. ARNOLD† AND SHAWN W. WALKER‡

For the convenience of the reader, we collect several basic results that are useful to the main paper.

SM1. Proof of Proposition 2.1. Using (2.8), we may estimate the skeleton term in (2.12) by

\[
\|v\|_{L^2(\partial T)}^2 \leq \sum_{T \in \mathcal{T}_h} \|v\|_{L^2(T)}^2 \leq C_3 \sum_{T \in \mathcal{T}_h} \|v\|_{L^2(T)}^2 + h^2 \|\nabla v\|_{L^2(T)}^2
\]

(SM1.1)

\[
\leq C_3 \left(\|v\|_{L^2(\Omega)}^2 + h^2 \|\nabla v\|_{L^2(\Omega)}^2\right),
\]

which gives the result.

SM2. Background of Curved Finite Elements. We review the theory in [SM5].

SM2.1. Parametric Elements. Let \( P_l(\mathcal{D}) \equiv P_l(\mathcal{D}; \mathbb{R}) \) be the space of polynomials of degree \( \leq l \) on the generic domain \( \mathcal{D} \) (for \( l \geq 0 \)), and let \( \hat{T} \) be the reference unit triangle. We introduce the set of Lagrange nodal variables (points) \( \mathcal{N}_l(\hat{T}) \equiv \hat{\mathcal{N}}_l \) on \( \hat{T} \) that correspond to \( P_l(\hat{T}) \equiv \hat{P}_l \), with associated (point evaluation) Lagrange interpolation operator \( \hat{I}_l \). Thus, the reference finite element is the triple \((\hat{T}, \hat{\mathcal{N}}_l, \hat{P}_l)\).

Let \( A_T : \hat{T} \to T^1 \) be an affine map that generates a straight triangle \( T^1 \in \mathcal{T}_h^1 \), with polynomial space \( P^1_l \equiv P^1_l(T^1) \), corresponding finite element triple \((T^1, \mathcal{N}^1_l, P^1_l)\), and Lagrange interpolation operator \( \tilde{I}_l \). Let \( T^m = F^m_T(T^1) \) be a curved triangle, where \( [P^m_m(\hat{T})]^2 \equiv F^m_T : T^1 \to T^m \) is a regular invertible mapping (and \( m \geq 1 \)). This induces a mapped polynomial space

\[
P^m_l \equiv P^m_l(T^m) = \{ \hat{p} \circ (F^m_T)^{-1} \mid \hat{p} \in P^1_l \},
\]

with mapped nodal set given by \( \mathcal{N}^m_l \equiv \mathcal{N}^m_l(T^m) = \{ F^m_T(a) \mid a \in \mathcal{N}^1_l \} \). Hence, the “parametric” finite element is the triple \((T^m, \mathcal{N}^m_l, P^m_l)\), with Lagrange interpolation operator given by \( \tilde{I}_l^m (f \circ F^m_T) = \tilde{I}_l^1 (f \circ F^m_T) \).

In general, there is no relation between \( l \) and \( m \). Typically, \( l = m \) refers to the iso-parametric case. We use the notation \( F^\infty_T \equiv F_T \) to indicate a general non-linear map (not necessarily a polynomial) that maps \( T^1 \) to a triangle \( T \in \mathcal{T}_h \), and the same considerations above apply to this case as well.

Note that \( m = 1 \) indicates linear (straight) triangles and there are well-known procedures for generating a conforming, shape regular triangulation, consisting of linear triangles, that approximates a smooth domain. Generating higher order triangles, 

*Submitted to the editors DATE.

Funding: This article is based upon work supported by the National Science Foundation under grants DMS-1719694 (Arnold) and DMS-155222 (Walker) and by the Simons Foundations under grant 601937 (Arnold).

†Department of Mathematics, University of Minnesota, Minneapolis (arnold@umn.edu, http://umn.edu/~arnold/).
‡Department of Mathematics, Louisiana State University, Baton Rouge (walker@lsu.edu, http://www.math.lsu.edu/~walker/).
of these maps.

**SM2.2. Approximation of \( \Omega \) by \( \Omega^m \).** The following results \([SM5, \text{Lemma 5, Propositions 2, 3}]\) give estimates on how well \( \Omega^m \) approximates \( \Omega \). Let \( \Psi_T^m : T^m \to T \), for each \( T \in T_h \), and note that \( \Psi_T^m \equiv F_T \).

**Proposition SM2.1 (Forward Map).** Let \( 1 \leq m \leq k \). The map \( \Psi_T^m \) satisfies the following properties:
1. There exist constants \( b_s > 0 \), independent of \( h \), such that
   \[
   (\text{SM2.2}) \quad \| \nabla^s((\Psi_T^m - \id_{T^m}) \circ F_T^m \circ A_T) \|_{L^\infty(T)} \leq b_s h^{m+1}, \; \forall \; s \leq m + 1.
   \]
2. There exist constants \( \gamma_s > 0 \), independent of \( h \), such that
   \[
   (\text{SM2.3}) \quad \| \nabla^s((\Psi_T^m - \id_{T^m}) \|_{L^\infty(T^m)} \leq \gamma_s h^{m+1-s}, \; \forall \; s \leq m + 1.
   \]
3. \( \Psi_T^m \) is a \( C^{m+1} \) diffeomorphism: \( T \to T^m \) satisfies the following properties:
   1. There exist constants \( \rho_s > 0 \), independent of \( h \), such that
      \[
      (\text{SM2.4}) \quad \| \det(\nabla \Psi_T^m) - 1 \|_{L^\infty(T^m)} \leq \gamma h^m.
      \]

**Proposition SM2.2 (Inverse Map).** Let \( 1 \leq m \leq k \). The inverse map \( (\Psi_T^m)^{-1} : T \to T^m \) satisfies the following properties:
1. There exist constants \( \rho_s > 0 \), independent of \( h \), such that
   \[
   (\text{SM2.5}) \quad \| \nabla^s((\Psi_T^m)^{-1} - \id_T) \|_{L^\infty(T)} \leq \rho_s h^{m+1-s}, \; \forall \; s \leq m + 1.
   \]
2. There exists \( \rho > 0 \), independent of \( h \), such that
   \[
   (\text{SM2.6}) \quad \| \det(\nabla ((\Psi_T^m)^{-1}) - 1 \|_{L^\infty(T)} \leq \rho h^m.
   \]

**SM2.3. Proof of Theorem 3.2.** We break the result up into Theorem SM2.3, Corollary SM2.4, and Corollary SM2.5.

**Theorem SM2.3.** The map \( \Phi_T^m = F_T^m \circ (F_T^l)^{-1} \), given in (3.1), satisfies a variant of (SM2.3) and (SM2.5), i.e.
\[
(\text{SM2.7}) \quad \| \nabla^s(\Phi_T^m - \id_{T^m}) \|_{L^\infty(T)} \leq Ch^{l+1-s},
\]
\[
\| \nabla^s((\Phi_T^m)^{-1} - \id_{T^m}) \|_{L^\infty(T^m)} \leq Ch^{l+1-s},
\]
for \( 0 \leq s \leq l + 1 \), where \( C \) only depends on \( \Gamma \).

**Proof.** By the triangle inequality and (SM2.3),
\[
(\text{SM2.8}) \quad \| \nabla^s(\Phi_T^m - \id_{T^m}) \|_{L^\infty(T)} \leq \| \nabla^s(\Phi_T^m - \Phi_T^l) \|_{L^\infty(T)} + \| \nabla^s(\Phi_T^l - \id_{T^l}) \|_{L^\infty(T)} \leq \| \nabla^s((\Phi_T^l)^{-1} - \id_{T^l}) \circ \Phi_T^m \|_{L^\infty(T^l)} + Ch^{l+1-s},
\]
where we used that \( \Phi_T^l = (\Phi_T^m)^{-1} \circ (F_T^l)^{-1} \circ (F_T^m)^{-1} = F_T^l \circ (F_T^m)^{-1} = \Phi_T^m \).

Next, let \( \tilde{F}_T^l := F_T^l \circ A_T : T \to T^l \) (for all \( l \leq m \)) and note that
\[
(\text{SM2.9}) \quad \| \nabla F_T^l \|_{L^\infty(T^l)} \leq \| \nabla F_T^l \|_{L^\infty(T)} \| (\nabla A_T)^{-1} \|_{L^\infty(T)} \leq C_1 h C_2 h^{-1} \leq C',
\]
where we used [SM5, Thm. 1] on $\nabla \tilde{F}_T^m$ (because $F_T^m$ is an optimal map), $A_T$ is the standard affine map, and $C' > 0$ is a uniform constant only depending on $\Gamma$. From this, we get
\[(\text{SM2.10})\]
\[
\nabla \Phi_{lm}^m = \nabla [F_{T}^m \circ (F_{T}^l)^{-1}] = [(\nabla F_{T}^m)(\nabla F_{T}^l)^{-1}] \circ (F_{T}^l)^{-1},
\]
\[
\Rightarrow \|\nabla \Phi_{lm}^m\|_{L^\infty(T^i)} \leq C\|\nabla F_{T}^m\|_{L^\infty(T^i)}\|\nabla F_{T}^l\|_{L^\infty(T^i)} \leq C'.
\]
Therefore,
\[(\text{SM2.11})\]
\[
\|\nabla^s \left( [\Psi_{T}^m - \text{id}_{T^m}] \circ \Phi_{lm}^m \right)\|_{L^\infty(T^i)} \leq C \sum_{j=1}^{s} \|\nabla^j (\Psi_{T}^m - \text{id}_{T^m})\|_{L^\infty(T^m)} \|\nabla^{s-j+1}\Phi_{lm}^m\|_{L^\infty(T^i)} \leq Ch^{l+1-s},
\]
for $0 \leq s \leq l + 1$, where we used the product rule, (SM2.3), (SM2.10), and many inverse estimates. Combining with (SM2.8) gives the first estimate in (SM2.7); the other estimate follows similarly.

**Corollary SM2.4.** The maps $F_{T}^m$, $F_{T}^l$, satisfy
\[(\text{SM2.12})\]
\[
\|\nabla^s (F_{T}^l - \text{id}_{T^l})\|_{L^\infty(T^1)} \leq Ch^{2-s}, \text{ for } s = 0, 1, 2,
\]
\[
\|\nabla^s (F_{T}^m - F_{T}^l)\|_{L^\infty(T^1)} \leq Ch^{l+1-s}, \text{ for } 0 \leq s \leq l + 1,
\]
and
\[(\text{SM2.13})\]
\[
1 - Ch \leq \|\nabla (F_{T}^l - F_{T}^m)^{-1}\|_{L^\infty(T^1)} \leq 1 + Ch, \quad \|\nabla (F_{T}^l)^{-1} - I\|_{L^\infty(T^1)} \leq Ch,
\]
where $C$ only depends on $\Gamma$.

**Proof.** The first estimate in (SM2.12) follows from (SM2.7) with $l = 1$, and also implies that
\[(\text{SM2.14})\]
\[
\|\nabla^s F_{l}\|_{L^\infty(T^1)} \leq C_s, \quad \text{for } s = 0, 1, 2,
\]
where $C_0 = C_1 = O(1)$, and $C_2 > 0$ depends on the curvature of $\Gamma$. In particular, this implies that $\|\nabla (F_{T}^l)^{-1}\|_{L^\infty(T^1)} = O(1)$, which gives
\[(\text{SM2.15})\]
\[
\|\nabla^s (F_{T}^l)^{-1} - I\|_{L^\infty(T^1)} \leq \|\nabla^s (F_{T}^l)^{-1}\|_{L^\infty(T^1)} \|I - \nabla (F_{T}^l)^{-1}\|_{L^\infty(T^1)} \leq C \|\nabla (F_{T}^l - \text{id}_{T^l})\|_{L^\infty(T^1)} \leq Ch,
\]
using the first estimate in (SM2.12). This proves (SM2.13). Then,
\[(\text{SM2.16})\]
\[
\|\nabla^s (F_{T}^m - F_{T}^l)\|_{L^\infty(T^1)} = \|\nabla^s [(\Phi_{lm}^m - \text{id}_{T^m}) \circ F_{T}^l]\|_{L^\infty(T^1)} \leq C \sum_{j=1}^{s} \|\nabla^j (\Phi_{lm}^m - \text{id}_{T^m})\|_{L^\infty(T^m)} \|\nabla^{s-j+1}F_{T}^l\|_{L^\infty(T^1)} \leq Ch^{l+1-s},
\]
for $0 \leq s \leq l + 1$, where we used the product rule, (SM2.7), (SM2.14), and many inverse estimates. 

**Corollary SM2.5.** The map $\Phi_{lm}^m$ satisfies the following identities:
\[(\text{SM2.17})\]
\[
[\Phi_{lm}^m - \text{id}_{T^m}] \circ F_{l} = F_{m} - F_{l},
\]
\[
[\nabla (\Phi_{lm}^m - \text{id}_{T^m})] \circ F_{l} = \nabla (F_{m} - F_{l}) + O(h^{l+1}),
\]
\[
[\nabla^2 (\Phi_{lm}^m - \text{id}_{T^m}) \cdot e_{\gamma}] \circ F_{l} = \nabla^2 (F_{m} - F_{l}) \cdot e_{\gamma} + O(h^{l}),
\]
for $\gamma = 1, 2$, where the constants depend on $\Gamma \in C^{k+1}$, and $l \leq m \leq k$. 

Proof. We start by noting that $\Phi^{jm} - \text{id}_{T^i} = (F^m - F^l) \circ (F^l)^{-1}$, and so
\[
\nabla (\Phi^{jm} - \text{id}_{T^i}) = (|\nabla (F^m - F^l)|) |\nabla F^l|^{-1} \circ (F^l)^{-1},
\]
(SM2.18) \[
\nabla^2 (\Phi^{jm} - \text{id}_{T^i}) = (|\nabla^2 (F^m - F^l)|) : [|\nabla F^l|^{-1} \otimes |\nabla F^l|^{-1}] \circ (F^l)^{-1}
\]
\[-(\nabla (F^m - F^l) \left[ |\nabla F^l|^{-T} |\nabla^2 F^l| |\nabla F^l|^{-1} \right] |\nabla F^l|^{-1} \circ (F^l)^{-1},
\]
and so
\[
[\nabla (\Phi^{jm} - \text{id}_{T^i})] \circ F^l = [\nabla (F^m - F^l)] [|\nabla F^l|^{-1}]
\]
(SM2.19) \[
= [\nabla (F^m - F^l)] \left[ |\nabla F^l|^{-1} - 1 \right] + \nabla (F^m - F^l)
\]
\[\leq \nabla (F^m - F^l) + Ch^{l+1},
\]
where we used (SM2.12) and (SM2.13). Next, we have
\[
[\nabla^2 (\Phi^{jm} - \text{id}_{T^i}) \cdot e_\gamma] \circ F^l = [\nabla F^l|^{-T} [\nabla^2 (F^m - F^l) \cdot e_\gamma] [\nabla F^l|^{-1}
\]
(SM2.20) \[
- (\nabla (F^m - F^l) \cdot e_\gamma) \left[ |\nabla F^l|^{-T} [\nabla^2 F^l| |\nabla F^l|^{-1} \right] [\nabla F^l|^{-1}
\]
\[\leq \nabla^2 (F^m - F^l) \cdot e_\gamma + Ch^l,
\]
where we add/subtract the identity matrix and use (SM2.12)–(SM2.14). \hfill \Box

SM2.4. Proof of Proposition 3.3.

Proof. W.L.O.G., assume $m > l$. From [SM5, Prop. 4], $\|v\|_{H^s(T^m)} \approx \|\hat{v}\|_{H^s(T)}$ for $s \geq 0$. More specifically, $\|\nabla v\|_{L^2(\Omega^m)} \approx \|\nabla \hat{v}\|_{L^2(\Omega)}$ and
\[
\|\nabla^2 v\|_{L^2(\Omega^m)} \leq C \left( \|\nabla^2 \hat{v}\|_{L^2(\Omega)} + h^{l+1} \|\nabla \hat{v}\|_{L^2(\Omega)} \right).
\]
Applying a change of variables to the jump term in (2.10) gives
(SM2.21) \[
\| \n \cdot \nabla v \|_{L^2(\Omega^m)} \leq C h^l \|\nabla \hat{v}\|_{L^2(\Omega)} + \| \n \cdot \nabla \hat{v} \|_{L^2(\Omega)},
\]
where we emphasize that we cannot put a jump in the first term on the right-hand-side because different Jacobians appear on either side of the edge. Next, we have the following scaling estimate (see (2.8))
(SM2.22) \[
\|\nabla \hat{v}\|_{L^2(\partial T^i)} \leq C_0 \left( h^{-1} \|\nabla \hat{v}\|_{L^2(T^i)} + h \|\nabla^2 \hat{v}\|_{L^2(T^i)} \right),
\]
which leads to
(SM2.23) \[
h^{-1/2} \| \n \cdot \nabla v \|_{L^2(\Omega^m)} \leq C_1 h^{-1} \|\nabla \hat{v}\|_{L^2(\Omega)} + C_1 h^l \|\nabla^2 \hat{v}\|_{L^2(\Omega)} + h^{-1/2} \| \n \cdot \nabla \hat{v} \|_{L^2(\Omega)},
\]
and implies $\|v\|_{2, h, m} \leq C_2 (\|\hat{v}\|_{2, h, l} + h^{l-1} \|\nabla \hat{v}\|_{L^2(\partial T^i)}$, giving the upper bound in (3.5) and (3.6). Combining with the Poincaré inequality in (2.14), shows the upper bound in (3.6); the lower bound follows similarly.

For $\|v\|_{0, h, m}$, the argument is simpler because there are no jump terms. \hfill \Box

SM3.1. Scaling Results.

Lemma SM3.1. Assume $1 \leq m \leq k$ (or $m = \infty$). There is a constant $C > 0$, independent of $h$ and $m$, such that

\begin{equation}
\|v\|_{L^q(\partial T^m)} \leq Ch^{1/q-2/p}\|v\|_{L^p(T^m)}, \quad \text{for any } 1 \leq q, p \leq \infty,
\end{equation}

for all $v \in P^m_l(T^m)$, where $r \geq 0$.

Corollary SM3.2. Assume $1 \leq m \leq k$ (or $m = \infty$). There is a constant $C > 0$, independent of $h$ and $m$, such that

\begin{equation}
\|\nabla v\|_{L^s(\partial T^m)} \leq Ch^{1/q-2/p}\|v\|_{L^p(T^m)}, \quad \text{for any } 1 \leq q, p \leq \infty,
\end{equation}

for all $v \in P^m_l(T^m)$, where $r \geq 0$.

The following mesh-dependent results essentially come from [SM1] (when $m = 1$).

Lemma SM3.3. Assume $1 \leq m \leq k$ (or $m = \infty$). There is a constant $C > 0$, independent of $h$ and $m$, such that

\begin{equation}
\|v\|_{0,h} \leq C\|v\|_{L^2(\Omega^m)}, \quad \text{for all } v \in W^m_h,
\end{equation}

where $\| \cdot \|_{0,h}$ is given in (2.12).

Proof. Combining Proposition 2.1 with an inverse inequality gives the assertion.\]
where $2 \leq l \leq \min(r + 2, t)$, for all $h$, $C > 0$ is an independent constant, and $1 \leq q \leq \infty$ is such that $W^{1,p}(\hat{T}) \hookrightarrow W^{2,q}(\hat{T})$.

Proof. The first inequality is standard. The second follows by recalling the definition (2.10) and combining the first inequality and (SM3.5) with $q = 2$. \qed

SM4. Approximation Results for Curved HHJ Finite Elements. We make note of some standard transformation rules for covariant and contravariant quantities, since they are critical for analyzing the geometric error when comparing similar quantities on different domains, e.g. $\Omega^m$ and $\Omega^l$ with $m \neq l$.

We use a “hat” notation to indicate a function defined on $T^1$. Recall the following transformation rules for normal and tangent vectors:

\[
\begin{align*}
\text{(covariant)} \quad v \circ F_T^m(\hat{u}) &= (B_T^m)^{-T} \hat{v}, \\
\text{(contravariant)} \quad v \circ F_T^m(\hat{u}) &= B_T^m \hat{v}, \\
\text{(covariant)} \quad w \circ F_T^m(\hat{u}) &= (B_T^m)^{-T} \hat{w}(B_T^m)^{-1}, \\
\text{(contravariant)} \quad w \circ F_T^m(\hat{u}) &= B_T^m \hat{w}(B_T^m)^T,
\end{align*}
\]

where $B_T^m := \nabla F_T^m$.

We also note the following transformation rules for normal and tangent vectors on $\partial T$:

\[
\begin{align*}
\circ F_T^m &= \frac{(B_T^m)^{-T} \hat{n}}{|(B_T^m)^{-T} \hat{n}|}, \quad \circ F_T^m = \frac{B_T^m \hat{t}}{|B_T^m \hat{t}|}.
\end{align*}
\]

SM4.1. The Reference HHJ Element. Recall (4.1) and let

\[
V_h^1 \equiv V_h^1(T^1) := \mathcal{P}_r(T^1; \mathbb{S}) \subset \mathcal{M}_{nn}^1(T^1),
\]

be a conforming finite element space on the element $T^1 \in \mathcal{T}_h$ with nodal Degrees-of-Freedom (DoFs) given by

\[
\begin{align*}
\bullet |E^1| \int_{E^1} \hat{n}^T \varphi \hat{q} \, ds, \quad \forall \hat{q} \in \mathcal{P}_r(E^1), \quad \forall E^1 \in \partial T^1, \\
\bullet \int_{T^1} \varphi \cdot \hat{\eta} \, d\hat{s}, \quad \forall \hat{\eta} \in \mathcal{P}_{r-1}(T^1; \mathbb{S}),
\end{align*}
\]

i.e. $\hat{\varphi} \in V_h^1$ is uniquely defined by (SM4.3) [SM6].

SM4.2. Matrix Piola Transform. We now recall Definition 4.2 of the matrix Piola transform, and verify a key elementary property, which shows that it preserves normal-normal continuity. Given an orientation-preserving diffeomorphism $F : \mathcal{D} \rightarrow \mathcal{D}$, and a tensor field $\varphi : \mathcal{D} \rightarrow \mathbb{R}^2$, we define $\hat{\varphi} : \mathcal{D} \rightarrow \mathbb{R}^2$ by

\[
\hat{\varphi}(\hat{x}) = (\det B)^2 B^{-1} \varphi(x) B^{-T},
\]

where $x = F(x)$, and $B = B(\hat{x}) = \nabla F(\hat{x})$. Also, denote by $\hat{t}, \hat{n} : \partial\mathcal{D} \rightarrow \mathbb{R}^2$ the positively-oriented unit tangent vector and the outward unit normal vector, and similarly for $t, n$ on $\partial\mathcal{D}$. Then the normal-normal component of $\varphi$ transforms as follows:

\[
(\hat{n}^T \hat{\varphi} \hat{n}) |B\hat{t}|^{-2} = (n^T \varphi n) \circ F.
\]
To see this, note that $t = |B\dot{t}|^{-1}B\dot{t}$ and $n = |B^{-T}\dot{n}|^{-1}B^{-T}\dot{n} = (\det B)|B\dot{t}|^{-1}B^{-T}\dot{n}$, where we have used the elementary identity $(\det B)B^{-T}\dot{n} = |B\dot{t}|$, whenever $B$ is a $2 \times 2$ matrix with positive determinant and $\dot{t}, \dot{n}$ are orthonormal. Thus $\dot{n} = (\det B)^{-1}B\dot{t}B^T\dot{n}$. Substituting this expression and the definition of $\phi$ into the left-hand size of (SM4.4) gives the result.

Note that the term $B\dot{t}$ in (SM4.4) is continuous across inter-element boundaries. Indeed, for $\phi^m$, we have

$$\phi^m \circ F^m_T \equiv (n^T \phi n) \circ F^m_T = |B^T\dot{t}|^{-2}\phi^{nn}. \quad \text{(SM4.5)}$$

Hence, one can map basis functions on the reference element to basis functions on the physical element using (4.2) and maintain normal-normal continuity. Furthermore, if $F$ is a general affine map ($m = 1$), with $E = F(E^1)$, then $|B\dot{t}| = |E|/|E^1|$, so (SM4.4) implies that

$$|E| \int_E (n^T \phi n)q \, ds = |E^1| \int_{E^1} (\dot{n}^T \phi \dot{n})\hat{q} \, ds, \quad \forall \hat{q} \in \mathcal{P}_r(E^1), \text{ and all edges } E^1 \text{ in } \partial T^1,$$

i.e. the edge DoFs are scaled when mapped.

We close with a norm equivalence implied by Lemma SM3.3:

$$\|\phi\|_{0,h}^2 \leq C_0 \sum_{\alpha,\beta=1}^2 \|\phi^{\alpha\beta}\|_{0,h}^2 \leq CC_0 \|\phi\|_{L^2(\Omega^m)}^2, \quad \forall \phi \in V_h^m, \quad \text{(SM4.7)}$$

so $\|\phi\|_{0,h} \approx \|\phi\|_{L^2(\Omega)}$ for all $\phi \in V_h^m$. By the same arguments in the proof of Proposition 3.3, we have that

$$\|\phi\|_{0,h,m} \approx \|\phi\|_{0,h,T}, \quad \forall \phi \in H^0_h(\Omega^m; \mathbb{S}), \quad \text{(SM4.8)}$$

using the Piola transform involving $\Phi$ taken from Proposition 3.3.

**SM4.3. Approximation Results of the HHJ Interpolation Operator.** The operator $\Pi_h^m$ enjoys the following approximation properties [SM4].

**LEMMA SM4.1.** Suppose $\phi \in W^{1,p}(\Omega^m; \mathbb{S})$, for $p > 1$ and $t \geq 1$ is an integer. Then,

$$\|\phi - \Pi_h^m \phi\|_{L^q(\Omega^m)} \leq C h^{t+1/q-2/p+2\min(0,1/p-1/q)} \|\phi\|_{W^{1,p}(\Omega^m)}, \quad \text{(SM4.9)}$$

where $1 \leq l \leq \min(r+1,t)$, for all $h$, and $C > 0$ is an independent constant, and $1 \leq q \leq \infty$ is such that $W^{1,p}(T) \hookrightarrow L^q(\partial T)$.

The following is a modification of [SM4, Lem. 4].

**LEMMA SM4.2.** For $2 \geq p > 1$ and $t \geq 1$ is an integer, there holds

$$\|\Pi_h^m \phi\|_{M_{\text{nn}}(\Omega^m)} \leq C \|\phi\|_{M_{\text{nn}}(\Omega^m)}, \quad \text{for all } \phi \in M_{\text{nn}}^m(\Omega^m), \quad \text{(SM4.10)}$$

$$\|\phi - \Pi_h^m \phi\|_{L^q(\Omega^m)} \leq C h^{l+2(1/q-1/p)+2\min(0,1/p-1/q)} \|\phi\|_{W^{1,p}(\Omega^m)}, \quad \text{for all } \phi \in M_{\text{nn}}^m(\Omega^m),$$

$$\|\phi - \Pi_h^m \phi\|_{L^2(\Omega^m)} \leq \|\phi - \Pi_h^m \phi\|_{0,h} \leq C h^{l+1-2/p} \|\phi\|_{W^{1,p}(\Omega^m)}, \quad 1 \leq l \leq \min(r+1,t), \quad \forall \phi \in M_{\text{nn}}^m(\Omega^m),$$

where $\|\phi\|_{M_{\text{nn}}(\Omega^m)}^p := \|\phi\|_{L^p(\Omega^m)}^p + \sum_{T \in \mathcal{M}} \|\nabla \phi\|_{L^p(T^m)}^p$, $1 \leq l \leq \min(r+1,t)$, for all $h$, $C > 0$ is an independent constant, and $1 \leq q \leq \infty$ is such that $W^{1,p}(T) \hookrightarrow L^q(\partial T)$. 


We have the following scaling result.

**Lemma SM4.3.** There is a constant $C > 0$, independent of $h$ and $m$, such that

$$
\|n^T \varphi n\|_{L^q(\partial T^m)} \leq \|\varphi\|_{L^q(\partial T^m)} \leq C h^{1/q - 2/p} \|\varphi\|_{L^p(T^m)}, \quad \text{for any } 1 \leq q, p \leq \infty,
$$

for all $\varphi := (\det B_T^m)^2 (B_T^m)^{-1} (\varphi \circ F_t^m) (B_T^m)^{-1} w \in P_r(T^1; \mathbb{R})$, where $r \geq 0$.

**SM5. Discrete Inf-Sup Condition.** We recall [SM2, Lem. 5.1].

**Lemma SM5.1.** Assume the domain $\Omega$ is piecewise linear, i.e. $m = 1$. Then,

$$\sup_{\varphi \in V_h^1} \frac{|b_h(\varphi, v)|}{\|\varphi\|_{0,h,1}} \geq C_0 \|v\|_{2,h,1}, \quad \forall v \in W_h^1, \forall h > 0,
$$

holds for any degree $r \geq 0$, where $C_0 > 0$ is independent of $h$.

**SM6. Proof of Theorem 5.2.** In lieu of Remark 4.7, we let $\|\cdot\|_h$ denote any norm on $W_h$ for which the inf-sup condition holds.

**Step 1.** First, form the usual “error equations”:

$$
a(\sigma - \sigma_h, \varphi_h) + b_h(\varphi_h, w - w_h) + b_h(\sigma - \sigma_h, v_h) = 0,
$$

for all $(v_h, \varphi_h) \in W_h \times V_h$. By the standard theory of mixed methods, one obtains the following abstract convergence result as a special case of [SM3, Thm. 5.2.1]:

$$
\|\sigma_h - \varphi\|_{0,h} \leq \frac{1}{\alpha_0} \sup_{\omega_h \in V_h} \left| a(\sigma - \varphi_h, \omega_h) + b_h(\omega_h, w - v_h) \right| \|
\omega_h\|_{0,h}
+ \frac{1}{\beta_0} \left( \frac{A_0}{\alpha_0} \right)^{1/2} \sup_{z_h \in W_h} \frac{|b_h(\sigma - \varphi_h, z_h)|}{\|z_h\|_h},
$$

$$
\|w_h - v_0\|_h \leq C \left( 1 + \frac{A_0^{1/2}}{\alpha_0^{1/2}} \right) \sup_{\omega_h \in V_h} \left| a(\sigma - \varphi_h, \omega_h) + b_h(\omega_h, w - v_h) \right| \|
\omega_h\|_{0,h}
+ \frac{C_0^2 A_0}{\beta_0} \sup_{z_h \in W_h} \frac{|b_h(\sigma - \varphi_h, z_h)|}{\|z_h\|_h},
$$

for all $v_h \in W_h$, $\varphi_h \in V_h$. Next, set $v_h = I_h w$, $\varphi_h = \Pi_h \sigma$.

**Step 2.** Let $l_1 := \min(r + 1, t - 2)$ and $l_2 := \min(r + 2, t)$. If $\|\cdot\|_h = \|\cdot\|_{2,h}$, (4.21), (3.31) and the interpolation estimates (SM3.7), (SM4.10) imply

$$
\|\sigma_h - \Pi_h \sigma\|_{0,h} + \|w_h - I_h w\|_{2,h} \leq C \|\sigma - \Pi_h \sigma\|_{0,h}
+ C \|\nabla (w - I_h w)\|_{L^2(\Omega)} + Ch \|\nabla^2 (w - I_h w)\|_{L^2(\Omega)},
$$

$$
\leq C \left( h^{l_1 + 1 - 2/p} \|\sigma\|_{W^{l_2,p}(\Omega)} + h^{l_2 - 2/p} \|w\|_{W^{l_2,p}(\Omega)} \right),
$$

where $C > 0$ is an independent constant.

Therefore, by the triangle inequality,

$$
\|\sigma - \sigma_h\|_{0,h} \leq C \left( h^{l_1 + 1 - 2/p} \|\sigma\|_{W^{l_2,p}(\Omega)} + h^{l_2 - 2/p} \|w\|_{W^{l_2,p}(\Omega)} \right),
$$

$$
\|w - w_h\|_{2,h} \leq C \left( h^{l_1 + 1 - 2/p} \|\sigma\|_{W^{l_2,p}(\Omega)} \right)
+ C \min \left\{ h^{l_2 - 1 - 2/p} \|w\|_{W^{l_2,p}(\Omega)}, \|w\|_{H^2(\Omega)} \right\},
$$

for any degree $r \geq 0$, where $C_0 > 0$ is independent of $h$. The proof is complete.
for some independent constant $C > 0$.

**Step 3.** If $\|\cdot\|_h = |\cdot|_{H^1(\Omega)}$, (4.21), (4.31) and the interpolation estimates (SM3.7), (SM4.10) imply

$$
\|\sigma - \Pi_h \sigma\|_{0,h} + \|\nabla (w_h - I_h w)\|_{L^2(\Omega)} \leq C \|\sigma - \Pi_h \sigma\|_{0,h}
$$

(SM6.7)

$$
+ C \left( \|\nabla (w - I_h w)\|_{L^2(\Omega)} + h\|\nabla^2 (w - I_h w)\|_{L^2(\sigma_{0,h})} \right)
$$

$$
\leq Ch^{1+1-2/p} \|\sigma\|_{W^{1,p}(\Omega)} + Ch^{2-2/p} \|w\|_{W^{2,p}(\Omega)},
$$

Therefore, by the triangle inequality,

(SM6.8)

$$
\|\sigma - \sigma_h\|_{0,h} + \|\nabla (w - w_h)\|_{L^2(\Omega)} \leq C \left( h^{1+1-2/p} \|\sigma\|_{W^{1,p}(\Omega)} + h^{2-2/p} \|w\|_{W^{2,p}(\Omega)} \right),
$$

for some independent constant $C > 0$.

**Step 4.** We use a duality argument to get a better estimate for $\|\nabla (w - w_h)\|_{L^2(\Omega)}$, in the low regularity case and when $r = 0$. Let $d \in H^{-1}(\Omega)$, and let $\tau \in V$, $\rho \in W$ solve (2.17) with $f$ replaced by $d$, i.e.

(SM6.9)

$$
a(\varphi, \tau) + b_h(\tau, v) + b_h(\varphi, \rho) = -\langle d, v \rangle, \quad \forall (\varphi, v) \in V \times W.
$$

Then, $\|\tau\|_{W^{1,p}(\Omega)} + \|\rho\|_{W^{2,p}(\Omega)} \leq C \|d\|_{H^{-1}(\Omega)}$. Next, set $\varphi = \sigma - \sigma_h$, $v = w - w_h$:

(SM6.10)

$$
-\langle d, w - w_h \rangle = a(\varphi, \tau) + b_h(\tau, w - w_h) + b_h(\sigma - \sigma_h, \rho).
$$

Next, combine with (SM6.1):

(SM6.11)

$$
-\langle d, w - w_h \rangle = a(\sigma - \sigma_h, \tau - \varphi_h) + b_h(\tau - \varphi_h, w - w_h)
$$

$$
+ b_h(\sigma - \sigma_h, \rho - v_h), \quad \forall (\varphi, v_h) \in V_h \times W_h.
$$

Now set $v_h = I_h \rho$, $\varphi_h = \Pi_h \tau$:

(SM6.12)

$$
|\langle d, w - w_h \rangle| \leq |a(\sigma - \sigma_h, \tau - \Pi_h \tau)| + |b_h(\tau - \Pi_h \tau, w - w_h)| + |b_h(\sigma - \sigma_h, \rho - I_h \rho)|.
$$

Let us estimate the terms involving $b_h(\cdot, \cdot)$. First, adding and subtracting $I_h w$:

(SM6.13)

$$
|b_h(\tau - \Pi_h \tau, w - w_h)| \leq |b_h(\tau - \Pi_h \tau, w - I_h w)| + |b_h(\tau - \Pi_h \tau, I_h w - w_h)|
$$

$$
\leq |b_h(\tau, w - I_h w)| + |b_h(\Pi_h \tau, w - I_h w)| + |b_h(\tau - \Pi_h \tau, I_h w - w_h)|
$$

$$
\leq |\langle d, w - I_h w \rangle| + C \|\Pi_h \tau\|_{L^2(\Omega)} \left( \|\nabla (w - I_h w)\|_{L^2(\Omega)} + h\|\nabla^2 (w - I_h w)\|_{L^2(\sigma_{0,h})} \right),
$$

$$
+ C \|\tau - \Pi_h \tau\|_{H^1_0(\Omega)} \|\nabla (I_h w - w_h)\|_{L^2(\Omega)},
$$

where we used (SM6.9) for $\tau$ and (4.21). Interpolation estimates, for $r = 0$ and minimal regularity, and (SM7.6) give

(SM6.14)

$$
|b_h(\tau - \Pi_h \tau, w - w_h)| \leq C \|d\|_{H^{-1}(\Omega)} \left( \|\nabla (w - I_h w)\|_{L^2(\Omega)} + h\|\nabla^2 (w - I_h w)\|_{L^2(\sigma_{0,h})} \right)
$$

$$
+ C \|\tau - \Pi_h \tau\|_{H^1_0(\Omega)} \|\nabla (I_h w - w_h)\|_{L^2(\Omega)}
$$

$$
\leq C \left[ h + \left( h^{2-2/p} \right)^2 \right] \|d\|_{H^{-1}(\Omega)} \|f\|_{H^{-1}(\Omega)}.
For the next term, we use (2.17) (for $\sigma$) and (4.21):

\begin{equation}
|b_h(\sigma - \sigma_h, \rho - I_h\rho)| \leq |b_h(\sigma, \rho - I_h\rho)| + |b_h(\sigma_h, \rho - I_h\rho)|
+ C\|\sigma_h\|_{L^2(\Omega)} \left( \|\nabla (\rho - I_h\rho)\|_{L^2(\Omega)} + h\|\nabla^2 (\rho - I_h\rho)\|_{L^2(T_{\partial\Omega})} \right),
\end{equation}

Again, interpolation estimates, for $r = 0$ and minimal regularity give

\begin{equation}
|b_h(\sigma - \sigma_h, \rho - I_h\rho)| \leq C\|f\|_{H^{-1}(\Omega)} \left[ \|\nabla (\rho - I_h\rho)\|_{L^2(\Omega)} + h\|\nabla^2 (\rho - I_h\rho)\|_{L^2(T_{\partial\Omega})} \right]
\leq Ch\|f\|_{H^{-1}(\Omega)} \|d\|_{H^{-1}(\Omega)}.
\end{equation}

**Step 5.** Lastly, for $r = 0$ and minimal regularity, we have

\begin{equation}
|a(\sigma - \sigma_h, \tau - \Pi_h\tau)| \leq C\|\sigma - \sigma_h\|_{L^2(\Omega)} \|\tau - \Pi_h\tau\|_{L^2(\Omega)}
\leq Ch^{2-2/p} \left( h^{2-2/p} \right) \|d\|_{H^{-1}(\Omega)} \|f\|_{H^{-1}(\Omega)},
\end{equation}

where we used (SM6.5). Therefore,

\begin{equation}
\|\nabla (w - w_h)\|_{L^2(\Omega)} \leq C \sup_{d \in H^{-1}(\Omega)} \frac{\langle d, w - w_h \rangle}{\|d\|_{H^{-1}(\Omega)}}
\leq C \max \left( h, h^{4-4/p} \right) \|f\|_{H^{-1}(\Omega)}.
\end{equation}

**REFERENCES**


