SIAM J. NUMER. ANAL.
Vol. 58, No. 5, pp. 2829–2855

THE HELLAN–HERRMANN–JOHNSON METHOD WITH CURVED ELEMENTS*

DOUGLAS N. ARNOLD† AND SHAWN W. WALKER‡

Abstract. We study the finite element approximation of the Kirchhoff plate equation on domains with curved boundaries using the Hellan–Herrmann–Johnson (HHJ) method. We prove optimal convergence on domains with piecewise C^{k+1} boundary for $k \geq 1$ when using a parametric (curved) HHJ space. Computational results are given that demonstrate optimal convergence and how convergence degrades when curved triangles of insufficient polynomial degree are used. Moreover, we show that the lowest order HHJ method on a polygonal approximation of the disk does not succumb to the classic Babuška paradox, highlighting the geometrically nonconforming aspect of the HHJ method.

Key words. Kirchhoff plate, simply supported, parametric finite elements, mesh-dependent norms, geometric consistency error, Babuška paradox

AMS subject classifications. 65N30, 35J40, 35Q72

DOI. 10.1137/19M1288723

1. Introduction.

The fourth order Kirchhoff plate bending problem presents notorious difficulties for finite element discretization. Among the many approaches that have been proposed, the Hellan–Herrmann–Johnson (HHJ) mixed method is one of the most successful. In simple situations (polygonal domains and smooth solutions), it provides stable discretization of arbitrary order and has been analyzed by many authors [9, 11, 10, 3, 2, 15, 6, 24, 16, 21]. However, in realistic applications the plate domain may well have a curved boundary, and additional errors arise from geometric approximation of the domain. In this paper, we analyze the effect of this geometric approximation and show that if handled correctly, the full discretization converges at the same optimal rate as is achieved for polygonal plates.

The well-known Babuška paradox demonstrates that there may be difficulties with low degree approximation of the geometry. Specifically, the paradox considers the effect of approximating the geometry only, without further numerical error. It considers a uniformly loaded simply supported circular plate and approximates the solution by the *exact solution* of the same problem on an inscribed regular polygon. As the number of sides of the polygon increases, the solution does *not* converge to the solution on the disk. The errors arising from linear approximation of the geometry lead to nonconvergence. However, as we shall show below, if the problem on the polygon is solved using the lowest order HHJ method, then convergence is restored. We further show that if higher order approximation of the geometry is combined with HHJ discretization of higher degree, the resulting method achieves any desired order.

While, to the best of our knowledge, the effect of domain approximation has not been studied before for the HHJ discretization of plates, its effect on the solution

*Received by the editors September 20, 2019; accepted for publication (in revised form) July 21, 2020; published electronically October 13, 2020.

https://doi.org/10.1137/19M1288723

Funding: The work of the first author was supported by National Science Foundation grant DMS-1719694 and by Simons Foundation grant 601397. The work of the second author was supported by National Science Foundation grant DMS-155222.

†Department of Mathematics, University of Minnesota, Minneapolis, MN 55455 USA (arnold@umn.edu, http://umn.edu/~arnold/).

‡Department of Mathematics, Louisiana State University, Baton Rouge, LA 70803-4918 USA (walker@math.lsu.edu, http://www.math.lsu.edu/~walker/).

of second order problems by standard finite elements is classical. See, for instance, [25, 22, 13, 18]. For simplicity, consider the Poisson equation $-\Delta u = f$ on Ω with the Dirichlet boundary condition $u = 0$ on $\partial\Omega$. Suppose we approximate the domain using parametric curved elements of some degree m to obtain an approximate domain Ω^m consisting of elements with maximum diameter h. To compare approximate solutions obtained on the approximation domain to the true solution on the true domain, we require a diffeomorphic mapping $\boldsymbol{\Psi} : \Omega^m \to \Omega$. The error analysis then depends on the identity [18, sect. 6]

$$\int_{\Omega} \nabla u \cdot \nabla v = \int_{\Omega^m} \nabla \tilde{u} \cdot \nabla \tilde{v} + \int_{\Omega^m} \nabla \tilde{u} \cdot \left[\det(\boldsymbol{J})\boldsymbol{J}^{-1}\boldsymbol{J}^{-T} - \mathbf{I}\right] \nabla \tilde{v},$$

where $\tilde{u} = u \circ \boldsymbol{\Psi}$ (and similarly for v) and $\boldsymbol{J} = \nabla \boldsymbol{\Psi}$ is the Jacobian matrix. The mapping $\boldsymbol{\Psi}$ is defined so that $\|\boldsymbol{J} - \mathbf{I}\|_{L^\infty} = O(h^m)$, where m is the degree of the polynomials used for the domain approximation. This leads to an $O(h^m)$ bound on the geometric consistency error term, the second integral on the right-hand side of the above identity. Choosing m to equal or exceed the degree of the finite elements used to approximate the solution (isoparametric or superparametric approximation) then ensures that full approximation order is maintained with curved elements.

Convergence for fourth order problems is less well established. In [20], the biharmonic problem is split into two second order equations with curved isoparametric elements and slightly modified boundary conditions. For plate problems, analyses of C^1 domain approximations have been considered (see [23, 12, 19, 26]).

The purpose of this paper is to give a rigorous estimate of the error between the continuous solution on the true domain and the discrete solution on the approximate domain. *The main difficulty* in this is dealing with higher derivatives of the nonlinear map that appear in the analysis (for instance, see [7, p. 78] and [14, Thm. 4.4.3]). For example, when mapping the Hessian, we have $\nabla^2 v = \boldsymbol{J}^{-1}\left[\nabla^2 \tilde{v} - \partial_\gamma \tilde{v} \boldsymbol{\Gamma}^\gamma\right] \boldsymbol{J}^{-T}$, where $\boldsymbol{\Gamma}^\gamma$ is a 2×2 matrix whose entries are the Christoffel symbols $\Gamma^\gamma_{\alpha\beta}$ of the second kind for the induced metric. These depend on *second* derivatives of $\boldsymbol{\Psi}$, and, consequently, $\|\boldsymbol{\Gamma}^\gamma\|_{L^\infty} = O(h^{m-1})$, so a naive handling of this term would yield suboptimal results or no convergence at all for $m = 1$. Another related issue is the handling of jump terms (appearing in some mesh-dependent norms) when affected by the nonlinear map.

The crucial tools needed to overcome these difficulties is the use of a Fortin-like operator (4.20) together with a particular optimal map (5.1) that is different from the curved element map given in [22, 18]. The results we present here should be of relevance to simulating plate problems on smooth and piecewise smooth domains.

We close the introduction with a brief outline of the remainder of the paper. Section 2 reviews the Kirchhoff plate problem and the mesh-dependent weak formulation behind the HHJ method. Section 3 provides a quick review of curved finite elements, and section 4 shows how to extend the classic HHJ method to curved elements. Section 5 provides the error analysis, which follows the framework of [3, 6], where we use a formulation of the Kirchhoff plate problem based on mesh-dependent spaces and analyze it with mesh-dependent norms. Section 6 gives numerical results, and we conclude in section 7 with some remarks. We also collect several basic or technical results in the supplementary materials, linked from the main article webpage.

2. Preliminaries and statements of results. We begin by recalling the Kirchhoff plate problem and the HHJ discretization and establishing our notations. The domain of the plate, i.e., its undeformed midsurface, is denoted by $\Omega \subset \mathbb{R}^2$ and its

boundary by $\Gamma := \partial\Omega$. Denoting the vertical displacement by w and the bending moment tensor by $\boldsymbol{\sigma}$, the plate equations [17, pp. 44–51] are

$$\boldsymbol{\sigma} = \mathbf{C}\,\nabla^2 w, \quad \operatorname{div}\operatorname{div}\boldsymbol{\sigma} = f \text{ in } \Omega, \quad \mathbf{C}\boldsymbol{\tau} := D\left[(1-\nu)\boldsymbol{\tau} + \nu\,\operatorname{tr}(\boldsymbol{\tau})\mathbf{I}\right], \tag{2.1}$$

where $\nabla^2 w$ denotes the Hessian of w, the iterated divergence div div takes a matrix field to a scalar function, f denotes the load function, and $\mathbf{C}$ is the constitutive tensor, with the bending modulus D given by $Et^3/12(1-\nu^2)$ in terms of Young's modulus E, the Poisson ratio ν, and the plate thickness t. We assume that $\nu \in (-1,1)$, so $\mathbf{C}$ is a symmetric positive-definite operator on the space $\mathbb{S}$ of symmetric 2×2 tensors. For a standard material, $0 \le \nu < 1/2$.

The differential equations (2.1) are supplemented by boundary conditions on $\partial\Omega$, such as

$$w = \partial w/\partial n = 0 \text{ for a clamped plate} \quad \text{or} \quad w = \sigma^{\mathrm{nn}} = 0 \text{ for a simply supported one.}$$

Here $\sigma^{\mathrm{nn}} = \boldsymbol{n}^T\boldsymbol{\sigma}\boldsymbol{n}$ denotes the normal-normal component of $\boldsymbol{\sigma}$.

The Kirchhoff plate problem can be formulated weakly. Taking $\mathcal{W} = \mathring{H}^2(\Omega)$ or $H^2(\Omega)\cap\mathring{H}^1(\Omega)$ for clamped and simply supported boundary conditions, respectively, $w \in \mathcal{W}$ is uniquely determined by the weak equations

$$\left(\mathbf{C}\nabla^2 w, \nabla^2 v\right) = \langle f, v\rangle \quad \forall v \in \mathcal{W} \tag{2.2}$$

for any f in $L^2(\Omega)$ or, more generally, in $\mathcal{W}^*$. Note that we use standard notations $H^m(\Omega)$ and $\mathring{H}^m(\Omega)$ for Sobolev spaces, with the latter subject to vanishing traces.

Next, we recall the HHJ method, first in the case of a polygonal domain (or a polygonal approximation to the true domain), and then for higher order polynomial approximations to the domain, which is the main subject here. Let $\mathcal{T}_h$ be a triangulation of the polygonal domain Ω, and let the degree $r \ge 0$ be fixed. The transverse displacement w will be approximated in the usual Lagrange finite element space

$$W_h = \{v \in \mathring{H}^1(\Omega) \mid v|_T \in \mathcal{P}_{r+1}(T)\ \forall T \in \mathcal{T}_h\},$$

while the bending moment tensor $\boldsymbol{\sigma}$ will be sought in the HHJ space

$$V_h = \{\boldsymbol{\varphi} \in L^2(\Omega;\mathbb{S}) \mid \boldsymbol{\varphi}|_T \in \mathcal{P}_r(T;\mathbb{S})\ \forall T \in \mathcal{T}_h,\ \boldsymbol{\varphi} \text{ normal-normal continuous}\}. \tag{2.3}$$

The normal-normal continuity condition means that if two triangles T_1 and T_2 share a common edge E, then $\boldsymbol{n}^T(\boldsymbol{\varphi}|_{T_1})\boldsymbol{n} = \boldsymbol{n}^T(\boldsymbol{\varphi}|_{T_2})\boldsymbol{n}$ on E. For simply supported boundary conditions, the space V_h also incorporates the vanishing of φ^{nn} on boundary edges.

Assume that $\boldsymbol{\sigma}$ belongs to $H^1(\Omega;\mathbb{S})$ and f belongs to $L^2(\Omega)$ (this is for simplicity; it can be weakened). Multiplying the second equation in (2.1) by a test function $v \in W_h$ and integrating over a triangle T, we obtain

$$\begin{aligned}(f,v)_T &= (\operatorname{div}\operatorname{div}\boldsymbol{\sigma}, v)_T = -(\operatorname{div}\boldsymbol{\sigma}, \nabla v)_T = (\boldsymbol{\sigma}, \nabla^2 v)_T - \langle \boldsymbol{\sigma}\boldsymbol{n}, \nabla v\rangle_{\partial T}\\ &= (\boldsymbol{\sigma}, \nabla^2 v)_T - \langle \boldsymbol{n}^T\boldsymbol{\sigma}\boldsymbol{n}, \partial v/\partial n\rangle_{\partial T} - \langle \boldsymbol{t}^T\boldsymbol{\sigma}\boldsymbol{n}, \partial v/\partial t\rangle_{\partial T}.\end{aligned}$$

Next, we sum this equation over all the triangles T. The penultimate term gives

$$\sum_T \langle \boldsymbol{n}^T\boldsymbol{\sigma}\boldsymbol{n}, \partial v/\partial n\rangle_{\partial T} = \sum_{E\in\mathcal{E}_h} \langle \sigma^{\mathrm{nn}}, [\![\partial v/\partial n]\!]\rangle_E\,.$$

Here $[\![\eta]\!]$ denotes the jump in a quantity η across a mesh edge E, so if the edge E is shared by two triangles T_1 and T_2 with outward normals $\boldsymbol{n}_1$ and $\boldsymbol{n}_2$, then $[\![\partial v/\partial n]\!] = \boldsymbol{n}_1 \cdot \nabla v|_{T_1} + \boldsymbol{n}_2 \cdot \nabla v|_{T_2}$ on E. For E a boundary edge, we set $[\![\eta]\!] = \eta|_E$. For the final term above, we obtain $\sum_T \langle \boldsymbol{t}^T \boldsymbol{\sigma} \boldsymbol{n}, \partial v/\partial t \rangle_{\partial T} = 0$, since $\partial v/\partial t$ is continuous across interior edges and the normal vector switches sign; moreover, $\partial v/\partial t$ vanishes on boundary edges. Thus, if we define the bilinear form

$$b_h(\boldsymbol{\varphi}, v) = -\sum_{T \in \mathcal{T}_h} \left(\boldsymbol{\varphi}, \nabla^2 v\right)_T + \sum_{E \in \mathcal{E}_h} \langle \varphi^{\mathrm{nn}}, [\![\partial v/\partial n]\!] \rangle_E, \quad \boldsymbol{\varphi} \in V_h,\ v \in W_h,$$

we have $b_h(\boldsymbol{\sigma}, v) = -\langle f, v \rangle$ for all $v \in W_h$. We define as well a second bilinear form

$$a(\boldsymbol{\tau}, \boldsymbol{\varphi}) = (\mathbf{K}\boldsymbol{\tau}, \boldsymbol{\varphi}), \quad \boldsymbol{\tau}, \boldsymbol{\varphi} \in V_h, \quad \text{where } \mathbf{K}\boldsymbol{\tau} := \frac{1}{D}\left[\frac{1}{1-\nu}\boldsymbol{\tau} - \frac{\nu}{1-\nu^2}\operatorname{tr}(\boldsymbol{\tau})\mathbf{I}\right],$$

with $\mathbf{K}$ the inverse of $\mathbf{C}$. Using the first equation in (2.1) and the continuity of $\partial w/\partial n$, we have $a(\boldsymbol{\sigma}, \boldsymbol{\tau}) + b_h(\boldsymbol{\tau}, w) = 0$ for any $\boldsymbol{\tau} \in V_h$. This leads us to the HHJ mixed method, which defines $\boldsymbol{\sigma}_h \in V_h$, $w_h \in W_h$ by

$$\begin{aligned} a(\boldsymbol{\sigma}_h, \boldsymbol{\tau}) + b_h(\boldsymbol{\tau}, w_h) &= 0 \quad \forall \boldsymbol{\tau} \in V_h, \\ b_h(\boldsymbol{\sigma}_h, v) &= -\langle f, v \rangle \quad \forall v \in W_h. \end{aligned} \tag{2.4}$$

This method has been analyzed by numerous authors with different techniques. The present analysis owes the most to [3, 6]. In particular, optimal $O(h^{r+1})$ convergence for $\boldsymbol{\sigma}$ in L^2 and w in H^1 has been established for smooth solutions.

If the plate domain Ω is not polygonal, a simple possibility is to construct a polygonal approximate domain. For this, we let $\mathcal{T}_h^1$ denote a triangulation consisting of straight-edged triangles with interior vertices belonging to Ω and boundary vertices belonging to $\partial\Omega$. The approximate domain Ω^1 is the region triangulated by $\mathcal{T}_h^1$. We assume further that no element of $\mathcal{T}_h^1$ has more than one edge on the boundary of Ω^1, and we call those that have such an edge boundary triangles.

As we shall see, in the case of the lowest order HHJ elements, $r = 0$, such a polygonal approximation of the geometry does not degrade the rate of convergence of the numerical scheme (the Babuška paradox notwithstanding). For higher order elements, however, we need to make a better approximation of the geometry in order to obtain the approximate rate, just as is true when solving the Poisson problem with standard Lagrange finite elements [18]. We now briefly describe the procedure (see Figure 1), the full specification and analysis of which will occupy the remainder of the paper. Let $m \geq 1$ denote the integer degree of approximation of the geometry (so $m = 1$ corresponds to the polygonal approximation). To each triangle $T^1 \in \mathcal{T}_h^1$ we associate a curvilinear triangle T^m and a diffeomorphism $\boldsymbol{F}_T^m : T^1 \to T^m$ which is a polynomial map of degree m. In the case where T^1 is a boundary triangle, we require that $\boldsymbol{F}_T^m$ restricts itself to the identity on the two nonboundary edges of T^1, and in case T^1 is not a boundary triangle, we simply take $T^m = T^1$ and $\boldsymbol{F}_T^m$ to be the identity. We require that the set of all such curvilinear triangles forms a triangulation $\mathcal{T}_h^m$ of the domain $\Omega^m = \bigcup_{T^1 \in \mathcal{T}_h} \boldsymbol{F}_T^m(T^1)$, which is a polynomial approximation of the true domain of degree m. Note that the map $\boldsymbol{F}^m : \Omega^1 \to \Omega^m$ given by $\boldsymbol{F}^m|_{T^1} = \boldsymbol{F}_T^m$, for all $T^1 \in \mathcal{T}_h^1$, is a diffeomorphism of the polygonal approximate domain onto the approximate domain of degree m.

Using the mapping $\boldsymbol{F}^m$, we may transform the finite element spaces W_h and V_h from the polygonal approximate domain to the degree m approximation Ω^m. For W_h,

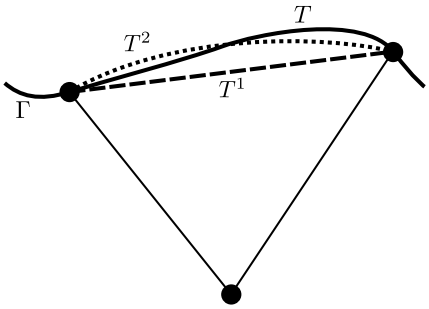

FIG. 1. *A straight-edged triangle T^1, degree 2 curvilinear triangle T^2, and curvilinear triangle T exactly conforming to the boundary Γ. All three share the two straight edges, with the corresponding boundary edges shown as dashed, dotted, and solid, respectively.*

the transformation is a simple composition, but for the tensor space V_h we must use the *matrix Piola transform*, which preserves normal-normal continuity. In this way, we obtained a mixed discretization of the plate problem based on elements of degrees $r+1$ and r for w and $\boldsymbol{\sigma}$, respectively, and geometric approximation of degree m. The integers $r \geq 0$ and $m \geq 1$ can be taken arbitrarily, but we show that to obtain the same optimal rates of convergence on a curved domain, as occurs for smooth solutions on a polygonal domain, it is sufficient to take $m \geq r+1$; e.g., polygonal approximation ($m = 1$) is sufficient for the lowest order HHJ elements ($r = 0$), but approximation must be at least quadratic to obtain optimality when $r = 1$, and so forth. Numerical experiments are included to show the necessity of this restriction.

2.1. Boundary assumptions. We shall allow for mixed boundary conditions, clamped on part of the domain and simply supported on the rest. To this end, we assume that Γ is piecewise smooth with a *finite* number of corners, where the interior angle α_i of the ith corner satisfies $\alpha_i \in (0, 2\pi]$ (see Figure 2). In particular, Γ is globally continuous and can be parameterized by a piecewise C^{k+1} curve for some $k \geq 1$, i.e., $\Gamma = \bigcup_{p \in \mathcal{V}_\Gamma} p \cup \bigcup_{\zeta \in \mathcal{C}_\Gamma} \zeta$, where $\mathcal{V}_\Gamma$ is the set of corner vertices and $\mathcal{C}_\Gamma$ is the set of (open) C^{k+1} curves that make up Γ. Moreover, we assume $\Gamma = \overline{\Gamma_{\mathrm{c}}} \cup \overline{\Gamma_{\mathrm{s}}}$ partitions into two mutually disjoint one-dimensional components Γ_{c} (clamped) and Γ_{s} (simply supported). Each open curve $\zeta \in \mathcal{C}_\Gamma$ belongs to only one of the sets Γ_{c} or Γ_{s}, and each curve is maximal such that two distinct curves contained in the same component do not meet at an angle of π. At the expense of small additional technical and notational complications, we could allow a partition of the boundary into three sets rather than two, imposing free boundary conditions on the third portion.

With the above partition of Γ, we have the following set of boundary conditions:

$$w = \partial w/\partial n = 0 \text{ on } \Gamma_{\mathrm{c}}, \quad w = \boldsymbol{n}^T \boldsymbol{\sigma} \boldsymbol{n} = 0 \text{ on } \Gamma_{\mathrm{s}}. \tag{2.5}$$

Extending the definition of the energy space $\mathcal{W}$ to account for these mixed boundary conditions,

$$\mathcal{W}(\Omega) := \{v \in H^2(\Omega) \mid v = 0 \text{ on } \Gamma_{\mathrm{c}} \cup \Gamma_{\mathrm{s}},\ \partial_{\boldsymbol{n}} v = 0 \text{ on } \Gamma_{\mathrm{c}}\}, \tag{2.6}$$

we have that $\left(\mathbf{C}\nabla^2 v, \nabla^2 v\right) \geq a_0 \|v\|_{H^2(\Omega)}$ for all $v \in \mathcal{W}$ and some constant $a_0 > 0$. Consequently, there exists a unique $w \in \mathcal{W}$ satisfying the plate equations in the weak formulation (2.2).

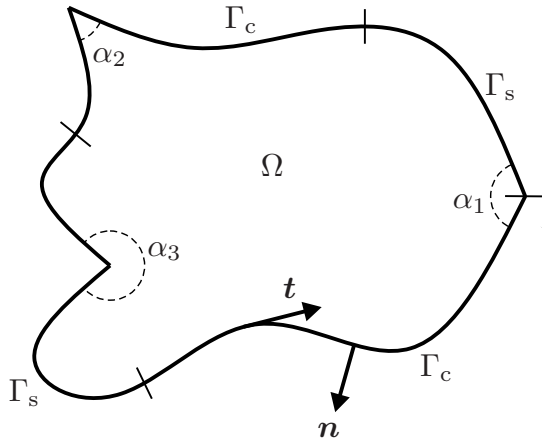

FIG. 2. *Illustration of plate domain* Ω. *The boundary* Γ *decomposes as* $\Gamma = \overline{\Gamma_c} \cup \overline{\Gamma_s}$ *and has a finite number of corners with interior angles* α_i*; the corners may (or may not) lie at the intersection of two boundary components. The outer unit normal vector is* $\boldsymbol{n}$*, and the oriented unit tangent vector is* $\boldsymbol{t}$.

2.2. Continuous mesh-dependent formulation. The main difficulty in solving (2.2) numerically is that $\mathcal{W} \subset H^2(\Omega)$, and so C^1 elements are required for a conforming discretization. We adopt the approach in [11, 3, 2, 6] and use a mesh-dependent version of $H^2(\Omega)$. We start by partitioning the domain Ω with a mesh $\mathcal{T}_h = \{T\}$ of triangles such that $\Omega = \bigcup_{T\in\mathcal{T}_h} T$, where $h_T := \operatorname{diam}(T)$ and $h := \max_T h_T$, and assume throughout that the mesh is quasi-uniform and shape regular. We further assume the corners of the domain are captured by vertices of the mesh.

Next, we have the *skeleton* of the mesh, i.e., the set of mesh edges $\mathcal{E}_h := \partial\mathcal{T}_h$. Let $\mathcal{E}_{\partial,h} \subset \mathcal{E}_h$ denote the subset of edges that are contained in the boundary Γ and respect the boundary condition partition of Γ. The internal edges are given by $\mathcal{E}_{0,h} := \mathcal{E}_h \setminus \mathcal{E}_{\partial,h}$. Note that elements in $\mathcal{T}_h$, $\mathcal{E}_h$ may be curved. For now, we assume Γ is piecewise smooth (at least C^2) with a *finite* number of corners to which the mesh conforms (see subsection 2.1 for more detailed assumptions).

The spaces in the following sections are infinite dimensional but defined in a "broken" way with respect to the partition. Thus, we adopt standard dG notation for writing inner products and norms over the partition, e.g.,

$$
\begin{aligned}
(f,g)_{\mathcal{T}_h} &:= \sum_{T\in\mathcal{T}_h} (f,g)_T\,, \quad (f,g)_{\mathcal{E}_h} := \sum_{E\in\mathcal{E}_h} (f,g)_E\,,\\
\|f\|^p_{L^p(\mathcal{T}_h)} &:= \sum_{T\in\mathcal{T}_h} \|f\|^p_{L^p(T)}, \quad \|f\|^p_{L^p(\mathcal{E}_h)} := \sum_{E\in\mathcal{E}_h} \|f\|^p_{L^p(E)}.
\end{aligned}
\tag{2.7}
$$

We shall make repeated use of the following scaling/trace estimate [1, Thm. 3.10]:

$$
\|v\|^2_{L^2(\partial T)} \le C\left(h^{-1}\|v\|^2_{L^2(T)} + h\|\nabla v\|^2_{L^2(T)}\right) \quad \forall v \in H^1(T). \tag{2.8}
$$

2.2.1. Skeleton spaces. We follow [3] in defining infinite-dimensional but mesh dependent spaces and norms. A mesh-dependent version of $H^2(\Omega)$ is given by

$$H_h^2(\Omega) := \{v \in H^1(\Omega) \mid v|_T \in H^2(T) \text{ for } T \in \mathcal{T}_h\}, \tag{2.9}$$

with the seminorm

$$\|v\|_{2,h}^2 := \|\nabla^2 v\|_{L^2(\mathcal{T}_h)}^2 + h^{-1} \|[\![\boldsymbol{n} \cdot \nabla v]\!]\|_{L^2(\mathcal{E}_{0,h})}^2 + h^{-1} \|[\![\boldsymbol{n} \cdot \nabla v]\!]\|_{L^2(\Gamma_c)}^2, \tag{2.10}$$

where $[\![\eta]\!]$ is the jump in quantity η across mesh edge E, and $\boldsymbol{n}$ is the unit normal on $E \in \mathcal{E}_h$; on a boundary edge, $[\![\eta]\!] \equiv \eta$. Next, for any $\boldsymbol{\varphi} \in H^1(\Omega;\mathbb{S})$, define

$$\|\boldsymbol{\varphi}\|_{0,h}^2 := \|\boldsymbol{\varphi}\|_{L^2(\Omega)}^2 + h \left\|\boldsymbol{n}^T \boldsymbol{\varphi} \boldsymbol{n}\right\|_{L^2(\mathcal{E}_h)}^2, \tag{2.11}$$

and define H_h^0 to be the completion: $H_h^0(\Omega;\mathbb{S}) := \overline{H^1(\Omega;\mathbb{S})}^{\|\cdot\|_{0,h}}$. Note that $H_h^0(\Omega;\mathbb{S}) \equiv L^2(\Omega;\mathbb{S}) \oplus L^2(\mathcal{E}_h;\mathbb{R})$, i.e., $\boldsymbol{\varphi} \in H_h^0(\Omega;\mathbb{S})$ is actually $\boldsymbol{\varphi} \equiv (\boldsymbol{\varphi}', \varphi^{\mathrm{nn}})$, where $\boldsymbol{\varphi}' \in L^2(\Omega;\mathbb{S})$ and $\varphi^{\mathrm{nn}} \in L^2(\mathcal{E}_h)$, with no connection between $\boldsymbol{\varphi}'$ and φ^{nn}. We also have that $\boldsymbol{\varphi} \in H^1(\Omega;\mathbb{S}) \subset H_h^0(\Omega;\mathbb{S})$ implies $\boldsymbol{n}^T \boldsymbol{\varphi}' \boldsymbol{n}|_{\mathcal{E}_h} = \varphi^{\mathrm{nn}}$ [3]. Furthermore, we have a scalar-valued function version of $\|\cdot\|_{0,h}$:

$$\|v\|_{0,h}^2 := \|v\|_{L^2(\Omega)}^2 + h\|v\|_{L^2(\mathcal{E}_h)}^2 \quad \forall v \in H^1(\Omega), \tag{2.12}$$

which satisfies the following estimate (proved in section SM1 of the supplementary materials).

PROPOSITION 2.1. *For all $v \in H^1(\Omega)$, $\|v\|_{0,h}^2 \le C\big(\|v\|_{L^2(\Omega)}^2 + h^2\|\nabla v\|_{L^2(\Omega)}^2\big)$ for some independent constant C.*

Next, we introduce the skeleton subspaces

$$\mathcal{W}_h := H_h^2(\Omega) \cap \mathring{H}^1(\Omega), \quad \mathcal{V}_h := \{\boldsymbol{\varphi} \in H_h^0(\Omega;\mathbb{S}) \mid \varphi^{\mathrm{nn}} = 0 \text{ on } \Gamma_s\}, \tag{2.13}$$

where $\mathcal{W}_h$ is a mesh-dependent version of (2.6) and $\mathcal{V}_h$ is used for the stress $\boldsymbol{\sigma}$. Note how essential and natural boundary conditions are imposed differently in (2.13) than in (2.6). In addition, we have the following Poincaré inequality, which follows by standard integration by parts arguments [6].

PROPOSITION 2.2. *Define $|||v|||_{2,h}^2 := \|\nabla^2 v\|_{L^2(\mathcal{T}_h)}^2 + h^{-1} \|[\![\boldsymbol{n} \cdot \nabla v]\!]\|_{L^2(\mathcal{E}_{0,h})}^2$. Then $|||\cdot|||_{2,h}$ is a norm on $\mathcal{W}_h$. Moreover, there is a constant $C_{\mathrm{P}} > 0$, depending only on Ω, such that*

$$\|\nabla v\|_{L^2(\Omega)} \le C_{\mathrm{P}} |||v|||_{2,h} \quad \forall v \in \mathcal{W}_h. \tag{2.14}$$

2.2.2. Mixed skeleton formulation. Following [3, 6], we define a broken version of the Hessian operator. Recalling the earlier discussion, we extend $b_h(\boldsymbol{\varphi}, v)$ to all $\boldsymbol{\varphi} \in H_h^0(\Omega;\mathbb{S})$ and $v \in H_h^2(\Omega)$, i.e.,

$$b_h(\boldsymbol{\varphi}, v) = -\left(\boldsymbol{\varphi}', \nabla^2 v\right)_{\mathcal{T}_h} + \langle \varphi^{\mathrm{nn}}, [\![\boldsymbol{n} \cdot \nabla v]\!] \rangle_{\mathcal{E}_h}, \tag{2.15}$$

and extend $a(\boldsymbol{\varphi}, \boldsymbol{\sigma})$ to all $\boldsymbol{\tau}, \boldsymbol{\varphi} \in H_h^0(\Omega;\mathbb{S})$:

$$a(\boldsymbol{\tau}, \boldsymbol{\varphi}) = \sum_{T \in \mathcal{T}_h} (\boldsymbol{\tau}, \mathbf{K}\boldsymbol{\varphi})_T \equiv (\boldsymbol{\tau}, \mathbf{K}\boldsymbol{\varphi})_{\mathcal{T}_h}. \tag{2.16}$$

Thus, we pose the following mixed weak formulation of the Kirchhoff plate problem. Given $f \in H^{-1}(\Omega)$, find $\boldsymbol{\sigma} \in \mathcal{V}_h$, $w \in \mathcal{W}_h$ such that

$$\begin{aligned} a(\boldsymbol{\sigma}, \boldsymbol{\varphi}) + b_h(\boldsymbol{\varphi}, w) &= 0 \quad \forall \boldsymbol{\varphi} \in \mathcal{V}_h, \\ b_h(\boldsymbol{\sigma}, v) &= -\langle f, v \rangle \quad \forall v \in \mathcal{W}_h, \end{aligned} \tag{2.17}$$

where $\langle \cdot, \cdot \rangle$ is the duality pairing between H^{-1} and $\mathring{H}^1$.

Remark 2.3. Assume for simplicity that the domain is smooth and $f \in H^{-1}$. Then the solution $(\boldsymbol{\sigma}, w)$ of (2.1) satisfies $w \in H^3$ and $\boldsymbol{\sigma} \in H^1$ and the pair $(\boldsymbol{\sigma}, w)$ solves (2.17). In addition, any solution of (2.17) is also a solution of (2.1), and so the formulations are equivalent. See [6, sect. 3] for details.

3. Curved finite elements. The basic theory of curved elements was initiated in [13] in two dimensions, with specific procedures for some low degree isoparametric Lagrange elements. In [30, 29, 28], a theory for arbitrarily curved (two-dimensional) elements was given, while the author of [22] gave a general procedure for arbitrary order isoparametric elements. Later, the author of [18] generalized the theory to any dimension and gave a method of constructing the curved elements. The following sections give the essential parts of [18] that we need for this paper; section SM2 gives a more complete review.

3.1. Curved triangulations. We recall the parametric approach to approximating a domain with a curved boundary by a curvilinear triangulation $\mathcal{T}_h^m$ of order $m \geq 1$, following [18]. The process begins with a conforming, shape-regular, straight-edged triangulation $\mathcal{T}_h^1$ which triangulates a polygon Ω^1 interpolating Ω (in the sense that the boundary vertices of Ω^1 lie on the boundary of Ω). We define $\mathcal{T}_{\partial,h}^1$ to be the set of triangles with at least one vertex on the boundary. We make the following assumption.

Hypothesis 3.1. Each triangle in $\mathcal{T}_h^1$ has at most two vertices on the boundary, and so it has at most one edge contained in Γ^1.

Next, for each $T^1 \in \mathcal{T}_h^1$, we define a map $\boldsymbol{F}_T^m : T^1 \to \mathbb{R}^2$ of polynomial degree m which maps T^1 diffeomorphically onto a curvilinear triangle T^m. The map is determined by specifying the images of the Lagrange nodes of degree m on T^1. Nodes on an interior edge of T^1 are specified to remain fixed, while those on a boundary edge have their image determined by interpolation of a chart defining the boundary. Nodes interior to T^1 are mapped in an intermediate fashion through their barycentric coordinates. See equation (14) of [18] for an explicit formula for $\boldsymbol{F}_T^m \circ \widehat{\boldsymbol{F}_T^1}$, where $\widehat{\boldsymbol{F}_T^1}$ is the affine map from the standard reference triangle to T^1. The maps $\boldsymbol{F}_T^m$ so determined satisfy optimal bounds on their derivatives, as specified in [18, Thms. 1 and 2]. Moreover, the triangulation $\mathcal{T}_h^m$ consisting of all the curvilinear triangles $T^m = \boldsymbol{F}_T^m(T^1)$, $T^1 \in \mathcal{T}_h^1$, is itself a conforming, shape-regular triangulation that approximates Ω by $\Omega^m := \bigcup_{T^m \in \mathcal{T}_h^m} T^m$. We also denote by $\mathcal{E}_h^m$ the set of edges of the triangulation $\mathcal{T}_h^m$, which is partitioned into interior edges $\mathcal{E}_{0,h}^m$ (all straight) and boundary edges $\mathcal{E}_{\partial,h}^m$ (possibly curved). Thus, $\Gamma^m := \bigcup_{E^m \in \mathcal{E}_{\partial,h}^m} E^m$ is an mth order approximation of Γ. Note that, by construction, (i) $\boldsymbol{F}_T^1 \equiv \mathrm{id}_{T^1}$; (ii) if T^1 has no side on Γ, then $\boldsymbol{F}_T^m \equiv \mathrm{id}_{T^1}$; and (iii) $\boldsymbol{F}_T^m|_E = \mathrm{id}_{T^1}|_E$ for all interior edges $E \in \mathcal{E}_{0,h}^m$.

Of course, the polynomial maps $\boldsymbol{F}_T^m$ may be combined to define a piecewise polynomial diffeomorphism $\boldsymbol{F}^m : \Omega^1 \to \Omega^m$. Moreover, two of these maps, for degrees l and m, may be combined to give a map between the corresponding approximate

domains. Referring to Figure 3(a), it is defined piecewise by

$$\boldsymbol{\Phi}^{lm}|_T = \boldsymbol{\Phi}_T^{lm} : T^l \to T^m, \quad \text{where} \quad \boldsymbol{\Phi}_T^{lm} := \boldsymbol{F}_T^m \circ (\boldsymbol{F}_T^l)^{-1}, \ \text{so} \ \boldsymbol{\Phi}_T^{1m} \equiv \boldsymbol{F}_T^m. \tag{3.1}$$

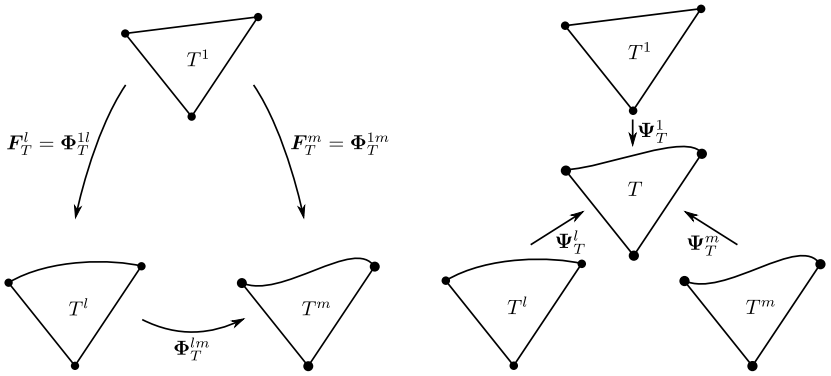

FIG. 3. (a) *Mappings between linear approximate triangles and approximate triangles of higher degree.* (b) *Mappings between the approximate triangles and the exact curvilinear triangle.*

In order to compare the exact solution, defined on the exact domain Ω, with an approximation defined on the approximate domain Ω^m, we require a map from the approximate domain to the true one. These can be defined elementwise in close analogy to $\boldsymbol{F}_T^m$. Specifically, given a triangle $T^m \in \mathcal{T}_h^m$, we define a map $\boldsymbol{\Psi}_T^m : T^m \to \mathbb{R}^2$ which maps T^m diffeomorphically onto a curvilinear triangle T exactly fitting Ω. If T^m has no boundary edges, the map is taken to be the identity. Otherwise, T^m has one edge $E^m \subset \Gamma^m$, and the map is defined by [18, eq. (32)]. It restricts itself to the identity on the interior edges of T^m and satisfies Propositions SM2.1 and SM2.2. The curvilinear triangulation $\mathcal{T}_h := \{\boldsymbol{\Psi}_T^m(T^m)\}_{T^m \in \mathcal{T}_h^m}$ then exactly triangulates Ω. The $\boldsymbol{\Psi}_T^m$ may be pieced together to give a global map $\boldsymbol{\Psi}^m : \Omega^m \to \Omega$.

We may view the exact domain and the corresponding triangulation as the limiting case of the approximate domain, and its triangulation, as $m \to \infty$. This leads to alternative notations $\Omega^\infty \equiv \Omega$, $\mathcal{T}_h^\infty \equiv \mathcal{T}_h$, $\boldsymbol{\Phi}^{l\infty} \equiv \boldsymbol{\Psi}^l$, $\boldsymbol{F}_T^\infty \equiv \boldsymbol{\Psi}^1$, etc., which will sometimes be convenient. Note that the use of the superscript infinity in the notation for these quantities is suggestive: the exact domain can be thought of as an infinite order approximation of itself. However, this is merely a choice of notation. We are not asserting here some sort of convergence of the polynomial approximate domains to the true domain.

Section SM2 of the supplementary materials gives further details on these maps, with the main results summarized in the next theorem (proved in subsection SM2.3).

THEOREM 3.2. *Assume Hypothesis* 3.1. *Then for all* $1 \le l \le m \le k$ *and* $m = \infty$, *the maps* $\boldsymbol{F}_T^m$, $\boldsymbol{F}_T^l$ *described above satisfy*

$$\begin{aligned} \|\nabla^s(\boldsymbol{F}_T^l - \mathrm{id}_{T^1})\|_{L^\infty(T^1)} &\le Ch^{2-s} \quad \textit{for} \quad s = 0, 1, 2, \\ \|\nabla^s(\boldsymbol{F}_T^m - \boldsymbol{F}_T^l)\|_{L^\infty(T^1)} &\le Ch^{l+1-s} \quad \textit{for} \quad 0 \le s \le l+1, \\ 1 - Ch \le \|[\nabla \boldsymbol{F}_T^l]^{-1}\|_{L^\infty(T^1)} &\le 1 + Ch, \quad \|[\nabla \boldsymbol{F}_T^l]^{-1} - \mathbf{I}\|_{L^\infty(T^1)} \le Ch, \end{aligned} \tag{3.2}$$

and the map $\boldsymbol{\Phi}^{lm}$ *satisfies the estimates*

$$\begin{aligned} \|\nabla^s(\boldsymbol{\Phi}_T^{lm} - \mathrm{id}_{T^l})\|_{L^\infty(T^l)} &\le Ch^{l+1-s} \quad \textit{for} \quad 0 \le s \le l+1, \\ \|\nabla^s((\boldsymbol{\Phi}_T^{lm})^{-1} - \mathrm{id}_{T^m})\|_{L^\infty(T^m)} &\le Ch^{l+1-s} \quad \textit{for} \quad 0 \le s \le l+1, \end{aligned} \tag{3.3}$$

$$\begin{aligned}\big[\nabla(\boldsymbol{\Phi}^{lm}-\mathrm{id}_{T^l})\big]\circ\boldsymbol{F}^l &= \nabla(\boldsymbol{F}^m-\boldsymbol{F}^l)+O(h^{l+1}),\\ \big[\nabla^2(\boldsymbol{\Phi}^{lm}-\mathrm{id}_{T^l})\cdot\mathbf{e}_\gamma\big]\circ\boldsymbol{F}^l &= \nabla^2(\boldsymbol{F}^m-\boldsymbol{F}^l)\cdot\mathbf{e}_\gamma+O(h^l)\quad \textit{for}\ \ \gamma=1,2,\end{aligned}\tag{3.4}$$

where all constants depend on the piecewise C^{k+1} norm of Γ.

Analyzing the geometric error of the HHJ mixed formulation is delicate (recall section 1). Indeed, the identity (3.4) will play an important role.

We close with a basic result relating norms on different order approximations of the same domain. The following result extends [14, Thm. 4.3.4] to the mesh-dependent norms in subsection 2.2.1 and is proved in subsection SM2.4.

PROPOSITION 3.3. *Assume the hypothesis of Proposition* 2.2. *Let $v\in H_h^2(\Omega^m)$, and define $\hat v = v\circ\boldsymbol{\Phi}\in H_h^2(\Omega^l)$, $\boldsymbol{\Phi}|_T := \boldsymbol{\Phi}_T^{lm}$ for any choice of $l,m\in\{1,2,\dots,k,\infty\}$. Let $\|v\|_{2,h,m}$, $\|v\|_{0,h,m}$ denote the norms in* (2.10), (2.12) *defined on Ω^m. Then $\|\nabla^2 v\|_{L^2(\mathcal{T}_h^m)}\le C\big(\|\nabla^2\hat v\|_{L^2(\mathcal{T}_h^l)}+h^{l-1}\|\nabla\hat v\|_{L^2(\mathcal{T}_h^l)}\big)$, and*

$$\|v\|_{2,h,m}\le C\left(\|\hat v\|_{2,h,l}+h^{l-1}\|\nabla\hat v\|_{L^2(\Omega^l)}\right),\quad \|v\|_{0,h,m}\approx C\|\hat v\|_{0,h,l},\tag{3.5}$$

$$\|v\|_{2,h,m}\approx C\|\hat v\|_{2,h,l}\quad \textit{if } v\in H_h^2(\Omega^m)\cap\mathring{H}^1(\Omega^m)\tag{3.6}$$

for some constant $C>0$ depending on the domain, where we modify the norm subscript to indicate the order of the domain.

3.2. Curved Lagrange spaces. Let r be a positive integer and m be a positive integer or ∞. The (continuous) Lagrange finite element space of degree r is defined on Ω^m via the mapping $\boldsymbol{F}_T^m$:

$$W_h^m\equiv W_h^m(\Omega^m) := \{v\in\mathring{H}^1(\Omega^m)\mid v|_T\circ\boldsymbol{F}_T^m\in\mathcal{P}_{r+1}(T^1)\ \forall T\in\mathcal{T}_h^m\}.\tag{3.7}$$

For the case $m=\infty$ (the exact domain), we simply write W_h.

If $v\in H_h^2(\Omega^1)$, then, on each triangle T^1, v is in H^2 and hence continuous up to the boundary of T^1. Globally, $v\in H^1(\Omega^1)$, and so it has a well-defined trace on each edge. Consequently, v is continuous on $\overline{\Omega^1}$ and we can define the Lagrange interpolation operator $\mathcal{I}_h^1: H_h^2(\Omega^1)\to W_h^1$ [3] defined on each element $T^1\in\mathcal{T}_h^1$ by

$$\begin{aligned}&(\mathcal{I}_h^1 v)(\boldsymbol{u})-v(\boldsymbol{u})=0\quad\forall\text{ vertices }\boldsymbol{u}\text{ of }T^1,\\ &\int_{E^1}(\mathcal{I}_h^1 v-v)q\,d\mathrm{s}=0\quad\forall q\in\mathcal{P}_{r-1}(E^1),\ \forall E^1\in\partial T^1,\\ &\int_{T^1}(\mathcal{I}_h^1 v-v)q\,d\mathrm{S}=0\quad\forall q\in\mathcal{P}_{r-2}(T^1).\end{aligned}\tag{3.8}$$

Then, given $v\in H_h^2(\Omega^m)$, we define the global interpolation operator, $\mathcal{I}_h^m: H_h^2(\Omega^m)\to W_h^m$, elementwise through $\mathcal{I}_h^m v\big|_{T^m}\circ\boldsymbol{F}_T^m := \mathcal{I}_h^1(v\circ\boldsymbol{F}_T^m)$. Note that $v\circ\boldsymbol{F}^m\in C^0(\Omega^1)$ because $v\in C^0(\Omega^m)$ and $\boldsymbol{F}^m$ is continuous over Ω^1. Approximation results for $\mathcal{I}_h^m$ are given in subsection SM3.2. We also denote $\mathcal{I}_h^{m,s}$ to be the above Lagrange interpolant on Ω^m onto continuous piecewise polynomials of degree s. Thus, $\mathcal{I}_h^{m,r+1}\equiv\mathcal{I}_h^m$.

4. The HHJ method. We start with a space of tensor-valued functions, defined on curved domains, with special continuity properties, followed by a transformation rule for the forms in (2.15) and (2.16). Next, we state the finite element approximation spaces for (2.17), which conform to $H_h^0(\Omega^m;\mathbb{S})$ and $H_h^2(\Omega^m)$, and define interpolation operators for these spaces while accounting for the effect of curved elements (recall that $1\le m\le k$ or $m=\infty$).

4.1. A tensor-valued space on curved domains. For $p > 3/2$, let

$$\mathcal{M}^m_{\mathrm{nn}}(\Omega^m) := \{\boldsymbol{\varphi} \in L^2(\Omega^m;\mathbb{S}) \mid \boldsymbol{\varphi}|_{T^m} \in W^{1,p}(T^m;\mathbb{S})\ \forall T^m \in \mathcal{T}^m_h, \ \boldsymbol{\varphi} \text{ normal-normal continuous}\}. \tag{4.1}$$

Note that $\mathcal{M}^m_{\mathrm{nn}}(\Omega^m) \subset H^0_h(\Omega^m;\mathbb{S})$, with $\varphi^{\mathrm{nn}} \equiv \boldsymbol{n}^T\boldsymbol{\varphi}'\boldsymbol{n}$ on each mesh edge.

Remark 4.1. The assumption that $p > 3/2$ is a technical simplification to ensure that the trace of a function in $\mathcal{M}^m_{\mathrm{nn}}(\Omega^m)$ onto the mesh skeleton $\mathcal{E}^m_h$ is in $L^2(\mathcal{E}^m_h)$.

In order to map between $\mathcal{M}^m_{\mathrm{nn}}(\Omega^m)$ and $\mathcal{M}^l_{\mathrm{nn}}(\Omega^l)$ (with $m \neq l$) such that normal-normal continuity is preserved, we need the following transformation rule.

DEFINITION 4.2 (matrix Piola transform). *Let $\boldsymbol{F} : \widehat{\mathcal{D}} \to \mathcal{D}$ be an orientation-preserving diffeomorphism between domains in $\mathbb{R}^2$. Given $\boldsymbol{\varphi} : \mathcal{D} \to \mathbb{S}$, we define its* matrix Piola transform $\hat{\boldsymbol{\varphi}} : \widehat{\mathcal{D}} \to \mathbb{S}$ *by*

$$\hat{\boldsymbol{\varphi}}(\hat{\mathbf{x}}) = (\det \boldsymbol{B})^2 \boldsymbol{B}^{-1}\boldsymbol{\varphi}(\mathbf{x})\boldsymbol{B}^{-T}, \tag{4.2}$$

where $\mathbf{x} = \boldsymbol{F}(\hat{\mathbf{x}})$, and $\boldsymbol{B} = \boldsymbol{B}(\hat{\mathbf{x}}) = \nabla \boldsymbol{F}(\hat{\mathbf{x}})$.

Note that (4.2) is analogous to the Piola transform for $H(\mathrm{div},\Omega)$ functions.

By elementary arguments (see (SM4.5)), we find that

$$\varphi^{\mathrm{nn}} \circ \boldsymbol{F} = \hat{\varphi}^{\mathrm{nn}}\, |(\nabla \boldsymbol{F})\hat{\boldsymbol{t}}|^{-2}. \tag{4.3}$$

We shall apply the transform when the diffeomorphism is $\boldsymbol{F}^m$, which is piecewise smooth and continuous with respect to the mesh. It follows that $(\nabla \boldsymbol{F})\hat{\boldsymbol{t}}$ is single-valued at interelement edges, so $\boldsymbol{\varphi}$ is normal-normal continuous if and only if $\hat{\boldsymbol{\varphi}}$ is.

We close with the following norm equivalences (see (SM4.7) and (SM4.8)):

$$\|\boldsymbol{\varphi}\|_{0,h,m} \approx \|\boldsymbol{\varphi}\|_{L^2(\Omega^m)}\ \forall\, \boldsymbol{\varphi} \in V^m_h, \quad \|\boldsymbol{\varphi}\|_{0,h,m} \approx \|\hat{\boldsymbol{\varphi}}\|_{0,h,l}\ \forall\, \boldsymbol{\varphi} \in H^0_h(\Omega^m;\mathbb{S}) \tag{4.4}$$

for all $1 \le l, m \le k, \infty$.

4.2. Mapping forms. The following result is crucial for analyzing the geometric error when approximating the solution on an approximate domain and also for deriving the discrete inf-sup condition on curved elements. We define $\Omega_S = \bigcup_{T\in\mathcal{T}_{\partial,h}} T$ for the "strip" domain contained in Ω. In addition, we generalize the definitions (2.16) and (2.15) of the bilinear forms $a(\cdot,\cdot)$ and $b_h(\cdot,\cdot)$ to include a superscript m to indicate that they are defined on the approximate domain Ω^m.

THEOREM 4.3. *Let $1 \le l \le k$ such that $m > l$, recall $\boldsymbol{\Phi} \equiv \boldsymbol{\Phi}^{lm} : \Omega^l \to \Omega^m$ from (3.1) for $1 < m \le k$ and $m = \infty$, and set $\boldsymbol{J} := \nabla\boldsymbol{\Phi}$. For all $\boldsymbol{\sigma}, \boldsymbol{\varphi} \in \mathcal{M}^m_{\mathrm{nn}}(\Omega^m)$ and $v \in H^2_h(\Omega^m)$, there holds that*

$$a^m(\boldsymbol{\sigma},\boldsymbol{\varphi}) = a^l(\hat{\boldsymbol{\sigma}},\hat{\boldsymbol{\varphi}}) + O(h^l)\|\hat{\boldsymbol{\sigma}}\|_{L^2(\Omega^l_S)}\|\hat{\boldsymbol{\varphi}}\|_{L^2(\Omega^l_S)}, \tag{4.5}$$

$$\begin{aligned} b^m_h(\boldsymbol{\varphi},v) = {} & b^l_h(\hat{\boldsymbol{\varphi}},\hat{v}) + \sum_{T^l\in\mathcal{T}^l_{\partial,h}} \langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot\nabla\left[(\mathrm{id}_{T^l} - \boldsymbol{\Phi}_T)\cdot \mathrm{P}_0\nabla\hat{v}\right]\rangle_{\partial T^l} \\ & + O(h^l)\|\hat{\boldsymbol{\varphi}}\|_{0,h,l}\|\nabla\hat{v}\|_{H^1(\mathcal{T}^l_{\partial,h})} - \sum_{T^l\in\mathcal{T}^l_{\partial,h}} \left(\hat{\boldsymbol{\varphi}}, \nabla^2[(\mathrm{id}_{T^l} - \boldsymbol{\Phi}_T)\cdot\mathrm{P}_0\nabla\hat{v}]\right)_{T^l} \\ & + O(h^l)\sum_{E^l\in\mathcal{E}^l_{\partial,h}} \|\hat{\varphi}^{\mathrm{nn}}\|_{L^2(E^l)}\|\nabla\mathcal{I}^{l,1}_h\hat{v}\|_{L^2(E^l)}, \end{aligned} \tag{4.6}$$

where $\boldsymbol{\sigma}, \boldsymbol{\varphi}$ and $\hat{\boldsymbol{\sigma}}, \hat{\boldsymbol{\varphi}}$ are related by the matrix Piola transform (4.2) *involving $\boldsymbol{\Phi}_T$, $v|_T \circ \boldsymbol{\Phi}_T = \hat{v}$, $\mathcal{I}_h^{l,1}$ is the Lagrange interpolation operator onto piecewise linears on Ω^l, and $\mathrm{P}_0 : L^2(\Omega^l) \to L^2(\Omega^l)$ is the projection onto piecewise constants.*

Proof. To derive (4.5), we use (4.2) to obtain

$$a^m(\boldsymbol{\sigma}, \boldsymbol{\varphi}) = \left(\varphi^{\gamma\omega}, K_{\gamma\omega\alpha\beta}\sigma^{\alpha\beta}\right)_{\Omega^m} = \sum_{T^l \in \mathcal{T}_h^l} \left((\det \boldsymbol{J}_T)^{-1}\hat{\varphi}^{\gamma\omega}, \widehat{K}_{\gamma\omega\alpha\beta}\hat{\sigma}^{\alpha\beta}\right)_{T^l}, \tag{4.7}$$

where $[\mathbf{K}]_{\gamma\omega\alpha\beta} \equiv K_{\gamma\omega\alpha\beta} = 1/(2\bar{\mu})\delta_{\gamma\alpha}\delta_{\omega\beta} - (\nu/E)\delta_{\gamma\omega}\delta_{\alpha\beta}$, with $\delta_{\alpha\beta}$ being the Kronecker delta, $\widehat{K}_{\gamma\omega\alpha\beta} = 1/(2\bar{\mu})g_{\gamma\alpha}g_{\omega\beta} - (\nu/E)g_{\gamma\omega}g_{\alpha\beta}$, $\boldsymbol{g} := \boldsymbol{J}^T\boldsymbol{J}$ is the induced metric, and $[\boldsymbol{g}]_{\alpha\beta} \equiv g_{\alpha\beta}$. The result follows by adding and subtracting terms, noting that $\|K_{\gamma\omega\alpha\beta} - \widehat{K}_{\gamma\omega\alpha\beta}\|_{L^\infty(\Omega^l)} \le Ch^l$, and using that $\boldsymbol{\Phi}_T = \mathrm{id}_{T^l}$ for all $T^l \notin \mathcal{T}_{\partial,h}^l$.

As for (4.6), we start with (2.15) and write it as

$$b_h^m(\boldsymbol{\varphi}, v) = -\sum_{T^m \in \mathcal{T}_h^m} \left[\left(\boldsymbol{\varphi}', \nabla^2 v\right)_{T^m} - \langle \varphi^{\mathrm{nn}}, \boldsymbol{n}\cdot\nabla v\rangle_{\partial T^m}\right], \tag{4.8}$$

noting that $\boldsymbol{n}^T\boldsymbol{\varphi}'\boldsymbol{n} \equiv \varphi^{\mathrm{nn}}$. It is only necessary to consider elements adjacent to the boundary, i.e., let $T^m \in \mathcal{T}_{\partial,h}^m$. Then, mapping the first term in (4.8) from Ω^m to Ω^l, we see that

$$\left(\boldsymbol{\varphi}, \nabla^2 v\right)_{T^m} = \left((\det \boldsymbol{J})^{-1}\hat{\varphi}^{\alpha\beta}, \left[\partial_\alpha\partial_\beta\hat{v} - \partial_\gamma\hat{v}\Gamma_{\alpha\beta}^\gamma\right]\right)_{T^l}, \tag{4.9}$$

where $\Gamma_{\alpha\beta}^\gamma$ are the Christoffel symbols of the second kind (depending on the induced metric $\boldsymbol{g}$). Note that $\Gamma_{\alpha\beta}^\gamma = g^{\mu\gamma}\partial_\alpha\partial_\beta(\boldsymbol{\Phi}\cdot\mathbf{e}_\mu)$, where $\mathbf{e}_\mu$ is a canonical basis vector, and $g^{\mu\gamma} \equiv [\boldsymbol{g}^{-1}]_{\mu\gamma}$ is the inverse metric. Using the estimates in (3.3) for $\boldsymbol{\Phi}$, we can express (4.9) as

$$\begin{aligned}\left(\boldsymbol{\varphi}, \nabla^2 v\right)_{T^m} &= \left(\hat{\boldsymbol{\varphi}}, \nabla^2\hat{v}\right)_{T^l} + \left([(\det \boldsymbol{J})^{-1} - 1]\hat{\varphi}^{\alpha\beta}, \partial_\alpha\partial_\beta\hat{v}\right)_{T^l} \\ &\quad - \left(\hat{\varphi}^{\alpha\beta}, \partial_\gamma\hat{v}\,\partial_\alpha\partial_\beta(\boldsymbol{\Phi}\cdot\mathbf{e}_\gamma)\right)_{T^l} - \left(\hat{\varphi}^{\alpha\beta}, \partial_\gamma\hat{v}(q^{\mu\gamma} - \delta^{\mu\gamma})\partial_\alpha\partial_\beta(\boldsymbol{\Phi}\cdot\mathbf{e}_\mu)\right)_{T^l} \\ &= \left(\hat{\boldsymbol{\varphi}}, \nabla^2\hat{v}\right)_{T^l} - \left(\hat{\varphi}^{\alpha\beta}\nabla\hat{v}, \partial_\alpha\partial_\beta\boldsymbol{\Phi}\right)_{T^l} + O(h^l)\|\hat{\boldsymbol{\varphi}}\|_{L^2(T^l)}\|\nabla\hat{v}\|_{H^1(T^l)},\end{aligned}$$

where we introduced $q^{\mu\gamma} = (\det \boldsymbol{J})^{-1}g^{\mu\gamma}$, and note that $\|q^{\mu\gamma} - \delta^{\mu\gamma}\|_{L^\infty(T^l)} \le Ch^l$ for all $T^l \in \mathcal{T}_h^l$. Furthermore, using the piecewise projection $\mathrm{P}_0|_{T^l} : L^2(T^l) \to \mathbb{R}$ onto constants, we have that

$$\begin{aligned}\left(\hat{\varphi}^{\alpha\beta}\nabla\hat{v}, \partial_\alpha\partial_\beta\boldsymbol{\Phi}\right)_{T^l} &= \left(\hat{\varphi}^{\alpha\beta}\mathrm{P}_0\nabla\hat{v}, \partial_\alpha\partial_\beta\boldsymbol{\Phi}\right)_{T^l} \\ &\quad + \left(\hat{\varphi}^{\alpha\beta}[\nabla\hat{v} - \mathrm{P}_0\nabla\hat{v}], \partial_\alpha\partial_\beta(\boldsymbol{\Phi} - \mathrm{id}_{T^l})\right)_{T^l} \\ &\le \left(\hat{\varphi}^{\alpha\beta}\mathrm{P}_0\nabla\hat{v}, \partial_\alpha\partial_\beta\boldsymbol{\Phi}\right)_{T^l} + Ch^l\|\hat{\boldsymbol{\varphi}}\|_{L^2(T^l)}\|\nabla^2\hat{v}\|_{L^2(T^l)},\end{aligned} \tag{4.10}$$

further noting that $\left(\hat{\varphi}^{\alpha\beta}\mathrm{P}_0\nabla\hat{v}, \partial_\alpha\partial_\beta\boldsymbol{\Phi}\right)_{T^l} = \left(\hat{\boldsymbol{\varphi}}, \nabla^2[(\boldsymbol{\Phi} - \mathrm{id}_{T^l})\cdot\mathrm{P}_0\nabla\hat{v}]\right)_{T^l}$.

Next, consider the second term in (4.8). Express $\partial T^m =: E_1^m \cup E_2^m \cup \widetilde{E}^m$, where $\widetilde{E}^m$ is the curved side, and map from ∂T^m to ∂T^l:

$$\langle \varphi^{\mathrm{nn}}, \boldsymbol{n}\cdot\nabla v\rangle_{\partial T^m} = \left\langle \frac{\det \boldsymbol{J}}{|\boldsymbol{J}\hat{\boldsymbol{t}}|^2}\hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot\boldsymbol{g}^{-1}\nabla\hat{v}\right\rangle_{\widetilde{E}^l} + \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot\boldsymbol{J}^{-T}\nabla\hat{v}\right\rangle_{E_1^l \cup E_2^l}, \tag{4.11}$$

where $\hat{\boldsymbol{n}}$ is the unit normal on ∂T^l and we used (4.3). Mapping the noncurved edges is simpler because $\boldsymbol{\Phi}_T = \mathrm{id}_{T^l}$ on $E_1^l \cup E_2^l$, so $E^m \equiv E^l$ and $\boldsymbol{n} \equiv \hat{\boldsymbol{n}}$. For convenience,

define $\boldsymbol{R} = \boldsymbol{J}^{-1}(\det \boldsymbol{J})|\boldsymbol{J}\hat{\boldsymbol{t}}|^{-2}$, which implies, by (3.3), that $\|\boldsymbol{R}-\mathbf{I}_{2\times 2}\|_{L^\infty(\tilde{E}^l)} \le C_1 h^l$. Since $\boldsymbol{g}^{-1} = \boldsymbol{J}^{-1}\boldsymbol{J}^{-T}$, we get

$$\begin{aligned}\langle \varphi^{\mathrm{nn}}, \boldsymbol{n}\cdot\nabla v\rangle_{\partial T^m} &= \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot\boldsymbol{J}^{-T}\nabla\hat{v}\right\rangle_{\partial T^l} + \|\hat{\boldsymbol{n}}\cdot[\boldsymbol{R}-\mathbf{I}]\boldsymbol{J}^{-T}\|_{L^\infty(\tilde{E}^l)} \\ &\quad\times \|\hat{\varphi}^{\mathrm{nn}}\|_{L^2(\tilde{E}^l)}\left[\|\nabla\hat{v}-\nabla\mathcal{I}_h^{l,1}\hat{v}\|_{L^2(\tilde{E}^l)} + \|\nabla\mathcal{I}_h^{l,1}\hat{v}\|_{L^2(\tilde{E}^l)}\right] \\ &= \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot\boldsymbol{J}^{-T}\nabla\hat{v}\right\rangle_{\partial T^l} + O(h^l)h^{1/2}\|\hat{\varphi}^{\mathrm{nn}}\|_{L^2(\tilde{E}^l)}\|\nabla^2\hat{v}\|_{L^2(T^l)} \\ &\quad + O(h^l)\|\hat{\varphi}^{\mathrm{nn}}\|_{L^2(\tilde{E}^l)}\|\nabla\mathcal{I}_h^{l,1}\hat{v}\|_{L^2(\tilde{E}^l)}\end{aligned} \tag{4.12}$$

for all $T^l \in \mathcal{T}^l_{\partial,h}$, where $\mathcal{I}_h^{l,1}$ satisfies, by (2.8), (SM3.6), and (SM3.7),

$$\begin{aligned}&h^{1/2}\|\nabla\hat{v}-\nabla\mathcal{I}_h^{l,1}\hat{v}\|_{L^2(\partial T^l)} \\ &\quad\le C\left(\|\nabla(\hat{v}-\mathcal{I}_h^{l,1}\hat{v})\|_{L^2(T^l)} + h\|\nabla^2(\hat{v}-\mathcal{I}_h^{l,1}\hat{v})\|_{L^2(T^l)}\right) \le Ch\|\nabla^2\hat{v}\|_{L^2(T^l)}.\end{aligned} \tag{4.13}$$

Furthermore,

$$\left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot\boldsymbol{J}^{-T}\nabla\hat{v}\right\rangle_{\partial T^l} = \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot[\boldsymbol{J}^{-T}-\mathbf{I}]\nabla\hat{v}\right\rangle_{\partial T^l} + \langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot\nabla\hat{v}\rangle_{\partial T^l},$$

and expanding further, and using (3.3), gives

$$\begin{aligned}\left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot[\boldsymbol{J}^{-T}-\mathbf{I}]\nabla\hat{v}\right\rangle_{\partial T^l} &= \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot\boldsymbol{J}^{-T}[\mathbf{I}-\boldsymbol{J}^T]\nabla\hat{v}\right\rangle_{\partial T^l} \\ &= \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot[\mathbf{I}-\boldsymbol{J}^T]\nabla\hat{v}\right\rangle_{\partial T^l} + \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot[\boldsymbol{J}^{-T}-\mathbf{I}][\mathbf{I}-\boldsymbol{J}^T]\nabla\hat{v}\right\rangle_{\partial T^l} \\ &= \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot[\mathbf{I}-\boldsymbol{J}^T](\mathrm{P}_0\nabla\hat{v})\right\rangle_{\partial T^l} + \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot[\boldsymbol{J}^{-T}-\mathbf{I}][\mathbf{I}-\boldsymbol{J}^T]\nabla\hat{v}\right\rangle_{\partial T^l} \\ &\quad + \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot[\mathbf{I}-\boldsymbol{J}^T](\nabla\hat{v}-\mathrm{P}_0\nabla\hat{v})\right\rangle_{\partial T^l} \\ &\le \left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot[\mathbf{I}-\boldsymbol{J}^T](\mathrm{P}_0\nabla\hat{v})\right\rangle_{\partial T^l} + Ch^l h^{1/2}\|\hat{\varphi}^{\mathrm{nn}}\|_{L^2(\partial T^l)}h^{1/2}\|\nabla\hat{v}\|_{L^2(\partial T^l)} \\ &\quad + Ch^{l-1}h^{1/2}\|\hat{\varphi}^{\mathrm{nn}}\|_{L^2(\partial T^l)}h^{1/2}\|\nabla\hat{v}-\mathrm{P}_0\nabla\hat{v}\|_{L^2(\partial T^l)}.\end{aligned} \tag{4.14}$$

Combining with (4.14), noting that P_0 satisfies an estimate similar to (4.13), and using (2.8) again, we obtain

$$\begin{aligned}\left\langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot[\boldsymbol{J}^{-T}-\mathbf{I}]\nabla\hat{v}\right\rangle_{\partial T^l} &= \langle \hat{\varphi}^{\mathrm{nn}}, \hat{\boldsymbol{n}}\cdot\nabla\left[(\mathrm{id}_{T^l}-\boldsymbol{\Phi})\cdot(\mathrm{P}_0\nabla\hat{v})\right]\rangle_{\partial T^l} \\ &\quad + O(h^l)\left(h^{1/2}\|\hat{\varphi}^{\mathrm{nn}}\|_{L^2(\partial T^l)}\right)\|\nabla\hat{v}\|_{H^1(T^l)}.\end{aligned} \tag{4.15}$$

Combining the above results and summing over all $T^m \in \mathcal{T}^m_{\partial,h}$ completes the proof. ∎

A simple consequence of Theorem 4.3 is

$$b_h^m(\boldsymbol{\varphi}, v) = b_h^l(\hat{\boldsymbol{\varphi}}, \hat{v}) + O(h^{l-1})\|\hat{\boldsymbol{\varphi}}\|_{0,h,l}\|\nabla\hat{v}\|_{2,h,l}. \tag{4.16}$$

4.3. The HHJ curved finite element space. We can use (4.2) to build the global, conforming, HHJ finite element space (on curved elements) by mapping from a reference element (see subsection SM4.2 of the supplementary materials for details); i.e., $V_h^m \equiv V_h^m(\Omega^m) \subset \mathcal{M}^m_{\mathrm{nn}}(\Omega^m)$ is defined by

$$\begin{aligned}V_h^m(\Omega^m) := \{\boldsymbol{\varphi} \in \mathcal{M}^m_{\mathrm{nn}}(\Omega^m) \mid \boldsymbol{\varphi}|_T \circ \boldsymbol{F}_T^m := (\det\nabla\boldsymbol{F}_T^m)^{-2}(\nabla\boldsymbol{F}_T^m)\hat{\boldsymbol{\varphi}}(\nabla\boldsymbol{F}_T^m)^T, \\ \hat{\boldsymbol{\varphi}} \in \mathcal{P}_r(T^1;\mathbb{S})\ \forall T^m \in \mathcal{T}_h^m\}.\end{aligned} \tag{4.17}$$

Note that V_h^m is isomorphic to V_h^1 for $1 \le m \le k$ and $m = \infty$.

We also have the following tensor-valued interpolation operator $\Pi_h^1 : \mathcal{M}_{\rm nn}^1(\Omega^1) \to V_h^1$ [11, 3], defined on each element $T^1 \in \mathcal{T}_h^1$ by

$$\begin{aligned} \int_{E^1} \boldsymbol{n}^T \left[\Pi_h^1 \boldsymbol{\varphi} - \boldsymbol{\varphi}\right] \boldsymbol{n}\, q \, \mathrm{d}s = 0 \quad \forall q \in \mathcal{P}_r(E^1),\ \forall E^1 \in \partial T^1, \\ \int_{T^1} \left[\Pi_h^1 \boldsymbol{\varphi} - \boldsymbol{\varphi}\right] : \boldsymbol{\eta}\, d\mathrm{S} = 0 \quad \forall \boldsymbol{\eta} \in \mathcal{P}_{r-1}(T^1; \mathbb{S}). \end{aligned} \tag{4.18}$$

Recall Theorem 4.3, and set $\boldsymbol{\Phi}_T \equiv \boldsymbol{\Phi}_T^{1m} := \boldsymbol{F}_T^m : T^1 \to T^m$, with $\boldsymbol{J}_T := \nabla \boldsymbol{\Phi}_T$. Now, given $\boldsymbol{\varphi} \in \mathcal{M}_{\rm nn}^m(\Omega^m)$, we define the global interpolation operator, $\Pi_h^m : \mathcal{M}_{\rm nn}^m(\Omega^m) \to V_h^m$, elementwise through

$$\Pi_h^m \boldsymbol{\varphi}\big|_{T^m} \circ \boldsymbol{\Phi}_T = (\det \boldsymbol{J}_T)^{-2} \boldsymbol{J}_T \left(\Pi_h^1 \hat{\boldsymbol{\varphi}}\right) \boldsymbol{J}_T^T, \tag{4.19}$$

where $\hat{\boldsymbol{\varphi}} := (\det \boldsymbol{J}_T)^2 \boldsymbol{J}_T^{-1} (\boldsymbol{\varphi} \circ \boldsymbol{\Phi}_T) \boldsymbol{J}_T^{-T}$ (i.e., see (4.2)). The operator Π_h^m satisfies many basic approximation results which can be found in subsection SM4.3.

On affine elements, we have a Fortin-like property involving $b_h^1(\cdot,\cdot)$ [11, 3, 6]:

$$\begin{aligned} b_h^1\left(\boldsymbol{\varphi} - \Pi_h^1 \boldsymbol{\varphi}, \theta_h v_h\right) = 0 \quad \forall \boldsymbol{\varphi} \in \mathcal{M}_{\rm nn}^1(\Omega^1), \quad v_h \in W_h^1, \\ b_h^1\left(\theta_h \boldsymbol{\varphi}_h, v - \mathcal{I}_h^1 v\right) = 0 \quad \forall \boldsymbol{\varphi}_h \in V_h^1, \quad v \in H_h^2(\Omega^1), \end{aligned} \tag{4.20}$$

which holds for *any* piecewise constant function θ_h defined on $\mathcal{T}_h^1$; in [11, 3, 6], it is assumed that $\theta_h \equiv 1$. However, (4.20) does not hold on curved elements, but instead we have the following result.

LEMMA 4.4. *Let $1 \le m \le k$, or $m = \infty$, and set $r \ge 0$ to be the degree of HHJ space V_h^m and $r+1$ to be the degree of the Lagrange space W_h^m. Moreover, assume V_h^m and W_h^m impose* no *boundary conditions. Then the following estimates hold:*

$$\begin{aligned} |b_h^m(\boldsymbol{\varphi}_h, v - \mathcal{I}_h^m v)| &\le C \|\boldsymbol{\varphi}_h\|_{L^2(\Omega_S^m)} \\ &\quad \times \left(\|\nabla (v - \mathcal{I}_h^m v)\|_{L^2(\Omega_S^m)} + h \|\nabla^2 (v - \mathcal{I}_h^m v)\|_{L^2(\mathcal{T}_{\partial,h}^m)} \right), \\ |b_h^m(\boldsymbol{\varphi} - \Pi_h^m \boldsymbol{\varphi}, v_h)| &\le C \|\boldsymbol{\varphi} - \Pi_h^m \boldsymbol{\varphi}\|_{H_h^0(\Omega_S^m)} \|\nabla v_h\|_{L^2(\Omega_S^m)} \end{aligned} \tag{4.21}$$

for all $\boldsymbol{\varphi} \in \mathcal{M}_{\rm nn}(\Omega^m)$, $v_h \in W_h^m$ and all $\boldsymbol{\varphi}_h \in V_h^m$, $v \in H_h^2(\Omega^m)$, where C is an independent constant. Note that $C = 0$ if $m = 1$.

Proof. Consider the map $\boldsymbol{\Phi}_T : T^l \to T^m$ defined in Theorem 4.3 with approximation properties given by (3.3). Let $v \in H_h^2(\Omega^m)$; so, by definition, there exists $\hat{v} \in H_h^2(\Omega^l)$ such that $v \circ \boldsymbol{\Phi}_T = \hat{v}$ on each $T^l \in \mathcal{T}_h^l$. By Proposition 3.3, $\|\nabla^2 v\|_{L^2(\mathcal{T}_h^m)} \approx C\big(\|\nabla^2 \hat{v}\|_{L^2(\mathcal{T}_h^l)} + h^{l-1} \|\nabla \hat{v}\|_{L^2(\mathcal{T}_h^l)}\big)$. Moreover, given $\boldsymbol{\varphi} \in \mathcal{M}_{\rm nn}(\Omega^m)$, there exists $\hat{\boldsymbol{\varphi}} \in \mathcal{M}_{\rm nn}(\Omega^l)$ given by $\boldsymbol{\varphi} \circ \boldsymbol{\Phi}_T := (\det \nabla \boldsymbol{\Phi}_T)^{-2} (\nabla \boldsymbol{\Phi}_T) \hat{\boldsymbol{\varphi}} (\nabla \boldsymbol{\Phi}_T)^T$ (cf. (4.2)). By (4.4), $\|\boldsymbol{\varphi}\|_{0,h,m} \approx \|\hat{\boldsymbol{\varphi}}\|_{0,h,l}$.

Next, we estimate the "problematic" terms in (4.6). First, we have

$$\left(\hat{\boldsymbol{\varphi}}, \nabla^2 [(\mathrm{id}_{T^l} - \boldsymbol{\Phi}_T) \cdot \mathrm{P}_0 \nabla \hat{v}]\right)_{T^l} \le C h^{l-1} \|\hat{\boldsymbol{\varphi}}\|_{L^2(T^l)} \|\nabla \hat{v}\|_{L^2(T^l)}, \tag{4.22}$$

where we used the optimal mapping properties in (3.3). In addition, we have

$$\begin{aligned} \langle \hat{\varphi}^{\rm nn}, \hat{\boldsymbol{n}} \cdot \nabla \left[(\mathrm{id}_{T^l} - \boldsymbol{\Phi}_T) \cdot \mathrm{P}_0 \nabla \hat{v}\right] \rangle_{\partial T^l} &\le C h^l \|\hat{\varphi}^{\rm nn}\|_{L^2(\partial T^l)} \|\mathrm{P}_0 \nabla \hat{v}\|_{L^2(\partial T^l)} \\ &\le C h^{l-1} h^{1/2} \|\hat{\varphi}^{\rm nn}\|_{L^2(\partial T^l)} \|\nabla \hat{v}\|_{L^2(T^l)} \end{aligned} \tag{4.23}$$

by the inverse estimate $\|\mathrm{P}_0\nabla\hat{v}\|_{L^2(\partial T^l)} \le Ch^{-1/2}\|\mathrm{P}_0\nabla\hat{v}\|_{L^2(T^l)}$. Last,

$$O(h^l)\|\hat{\varphi}^{\mathrm{nn}}\|_{L^2(E^l)}\|\nabla\mathcal{I}_h^{l,1}\hat{v}\|_{L^2(E^l)} \le Ch^{l-1}h^{1/2}\|\hat{\varphi}^{\mathrm{nn}}\|_{L^2(E^l)}\|\nabla\hat{v}\|_{L^2(T^l)} \tag{4.24}$$

for all $E^l \in \mathcal{E}^l_{\partial,h}$ (again using an inverse estimate and stability of the interpolant). Plugging (4.22)–(4.24) into (4.6) yields

$$|b_h^m(\boldsymbol{\varphi}, v)| \le \left|b_h^l(\hat{\boldsymbol{\varphi}}, \hat{v})\right| + Ch^{l-1}\|\hat{\boldsymbol{\varphi}}\|_{0,h,l}\left(\|\nabla\hat{v}\|_{L^2(\Omega^l_S)} + h\|\nabla^2\hat{v}\|_{L^2(\mathcal{T}^l_{\partial,h})}\right). \tag{4.25}$$

For the first estimate in (4.21), set $l = 1$, replace v with $v - \mathcal{I}_h^m v$, set $\boldsymbol{\varphi} = \boldsymbol{\varphi}_h \in V_h^m$, and use (4.20) and (4.4) to get

$$|b_h^m(\boldsymbol{\varphi}_h, v - \mathcal{I}_h^m v)| \le C\|\hat{\boldsymbol{\varphi}}_h\|_{L^2(\Omega^1_S)}\left(\|\nabla(\hat{v} - \mathcal{I}_h^1\hat{v})\|_{L^2(\Omega^1_S)} + h\|\nabla^2(\hat{v} - \mathcal{I}_h^1\hat{v})\|_{L^2(\mathcal{T}^1_{\partial,h})}\right). \tag{4.26}$$

Then use equivalence of norms (see Proposition 3.3 and (4.4)).

For the second estimate, replace $\boldsymbol{\varphi}$ with $\boldsymbol{\varphi} - \Pi_h^m\boldsymbol{\varphi}$, set $v = v_h \in W_h^m$, use (4.20), and use an inverse inequality to get

$$|b_h^m(\boldsymbol{\varphi} - \Pi_h^m\boldsymbol{\varphi}, v_h)| \le C\|\hat{\boldsymbol{\varphi}} - \Pi_h^1\hat{\boldsymbol{\varphi}}\|_{0,h,1}\|\nabla\hat{v}_h\|_{L^2(\Omega^1_S)}, \tag{4.27}$$

followed by an equivalence of norms argument. ∎

4.4. The HHJ mixed formulation. We pose (2.17) on Ω^m with continuous skeleton spaces denoted by $\mathcal{V}_h^m \equiv \mathcal{V}_h^m(\Omega^m)$ and $\mathcal{W}_h^m \equiv \mathcal{W}_h^m(\Omega^m)$. Fixing the polynomial degree $r \ge 0$, the conforming finite element spaces are

$$V_h^m \subset \mathcal{V}_h^m, \qquad W_h^m \subset \mathcal{W}_h^m. \tag{4.28}$$

The conforming finite element approximation to (2.17) is as follows. Given $f \in H^{-1}(\Omega^m)$, find $\boldsymbol{\sigma}_h \in V_h^m$, $w_h \in W_h^m$ such that

$$\begin{aligned} a^m(\boldsymbol{\sigma}_h, \boldsymbol{\varphi}) + b_h^m(\boldsymbol{\varphi}, w_h) &= 0 \quad \forall \boldsymbol{\varphi} \in V_h^m, \\ b_h^m(\boldsymbol{\sigma}_h, v) &= -\langle f, v\rangle_{\Omega^m} \quad \forall v \in W_h^m. \end{aligned} \tag{4.29}$$

The well-posedness of (4.29) is established in the next section; i.e., we prove the classic LBB conditions [7]. With this, we have the following a priori estimate:

$$\|w_h\|_{2,h,m} + \|\boldsymbol{\sigma}_h\|_{0,h,m} \le C\|f\|_{H^{-1}(\Omega^m)}. \tag{4.30}$$

Note that LBB conditions for (4.29) for the case $m = 1$ were originally shown in [6].

4.4.1. Well-posedness. Obviously, we have

$$a^m(\boldsymbol{\sigma}, \boldsymbol{\varphi}) \le A_0\|\boldsymbol{\sigma}\|_{L^2(\Omega^m)}\|\boldsymbol{\varphi}\|_{L^2(\Omega^m)} \quad \forall \boldsymbol{\sigma}, \boldsymbol{\varphi} \in H_h^0(\Omega^m; \mathbb{S}) \supset V_h^m, \tag{4.31}$$

$$|b_h^m(\boldsymbol{\varphi}, v)| \le B_0\|\boldsymbol{\varphi}\|_{0,h,m}\|v\|_{2,h,m} \quad \forall \boldsymbol{\varphi} \in V_h^m,\ v \in W_h^m, \tag{4.32}$$

and we have coercivity of $a^m(\cdot,\cdot)$, which is a curved element version of [3, Thm. 2].

LEMMA 4.5. *Assume the domain $\Omega^m \subset \mathbb{R}^2$ is piecewise smooth, consisting of curved elements as described in section 3. Then there is a constant $\alpha_0 > 0$, independent of h and m, such that*

$$a^m(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \ge \min(|\mathbf{K}|)\|\boldsymbol{\sigma}\|^2_{L^2(\Omega^m)} \ge \alpha_0\|\boldsymbol{\sigma}\|^2_{0,h} \quad \forall \boldsymbol{\sigma} \in V_h^m,\ \forall h > 0, \tag{4.33}$$

where α_0 depends on $\mathbf{K}$.

Proof. Clearly, $a^m(\boldsymbol{\sigma}, \boldsymbol{\sigma}) \ge C_0\|\boldsymbol{\sigma}\|^2_{L^2(\Omega^m)}$, where C_0 depends on $K_{\gamma\omega\alpha\beta}$. Furthermore, by (4.4), $\|\boldsymbol{\sigma}\|_{L^2(\Omega^m)} \ge C^{-1}\|\boldsymbol{\sigma}\|_{0,h}$, so then $\alpha_0 := C_0/C^2$. ∎

4.4.2. Inf-sup. Next, we have a curved element version of the inf-sup condition in [6, Lem. 5.1].

LEMMA 4.6. *Assume the domain $\Omega^m \subset \mathbb{R}^2$ is piecewise smooth, consisting of curved elements as described in section 3. Then there is a constant $\beta_0 > 0$, independent of h and m, such that for all h sufficiently small*

$$\sup_{\boldsymbol{\varphi}\in V_h^m} \frac{|b_h^m(\boldsymbol{\varphi}, v)|}{\|\boldsymbol{\varphi}\|_{0,h,m}} \geq \beta_0 \|v\|_{2,h,m} \quad \forall v \in W_h^m,\ \forall h > 0. \tag{4.34}$$

Proof. We start with the case $m = 1$ in [6, Lem. 5.1]:

$$\sup_{\hat{\boldsymbol{\varphi}}\in V_h^1} \frac{|b_h^1(\hat{\boldsymbol{\varphi}}, \hat{v})|}{\|\hat{\boldsymbol{\varphi}}\|_{0,h,1}} \geq C_0 \|\hat{v}\|_{2,h,1} \quad \forall \hat{v} \in W_h^1,\ \forall h > 0 \tag{4.35}$$

on the piecewise linear domain Ω^1 with triangulation $\mathcal{T}_h^1$ and holds for any degree $r \geq 0$ of the HHJ space.

Consider the map $\boldsymbol{\Phi}_T : T^1 \to T^m$ from Theorem 4.3 with approximation properties given by (3.3). Let $v \in W_h^m$; so, by definition, there exists $\hat{v} \in W_h^1$ such that $v \circ \boldsymbol{\Phi}_T = \hat{v}$ on each $T^1 \in \mathcal{T}_h^1$. By (3.6), $\|v\|_{2,h,m} \approx \|\hat{v}\|_{2,h,1}$. By (4.2) (using $\boldsymbol{\Phi}_T$), for any $\boldsymbol{\varphi} \in V_h^m$, there exists $\hat{\boldsymbol{\varphi}} \in V_h^1$, such that $\|\boldsymbol{\varphi}\|_{0,h,m} \approx \|\hat{\boldsymbol{\varphi}}\|_{0,h,1}$, by (4.4).

We will use (4.6) to estimate $|b_h^m(\boldsymbol{\varphi}, v)|$ when $l = 1$. Upon recalling the norm (2.10), because of boundary conditions, we have that the last term in (4.6) bounds as

$$O(h) \sum_{E^1 \in \mathcal{E}_{\partial,h}^1} \|\hat{\varphi}^{\mathrm{nn}}\|_{L^2(E^1)} \|\hat{\boldsymbol{n}} \cdot \nabla \mathcal{I}_h^{1,1} \hat{v}\|_{L^2(E^1)} \leq Ch \|\hat{\boldsymbol{\varphi}}\|_{0,h,1} \|\hat{v}\|_{2,h,1}. \tag{4.36}$$

Moreover, applying basic estimates to (4.6) yields

$$b_h^m(\boldsymbol{\varphi}, v) \geq b_h^1(\hat{\boldsymbol{\varphi}}, \hat{v}) - Ch \|\hat{\boldsymbol{\varphi}}\|_{0,h,1} \|\hat{v}\|_{2,h,1} - C \|\hat{\boldsymbol{\varphi}}\|_{L^2(\Omega_S^1)} \|\nabla \hat{v}\|_{L^2(\Omega_S^1)}. \tag{4.37}$$

Since $\|\nabla \hat{v}\|_{L^2(\Omega_S^1)} \leq \|\nabla \hat{v}\|_{L^\infty(\Omega^1)} \|1\|_{L^2(\Omega_S^1)} \leq \|\nabla \hat{v}\|_{L^\infty(\Omega^1)} |\Gamma|^{1/2} h^{1/2}$, by the quasi-uniform mesh assumption, and $\|\nabla \hat{v}\|_{L^\infty(\Omega^1)}^2 \leq C(1 + \ln h)(\|\hat{v}\|_{2,h,1}^2 + \|\nabla \hat{v}\|_{L^2(\Omega^1)}^2)$ (see [8, eq. (4)]), we get

$$b_h^m(\boldsymbol{\varphi}, v) \geq b_h^1(\hat{\boldsymbol{\varphi}}, \hat{v}) - Ch^{1/2-\epsilon} \|\hat{\boldsymbol{\varphi}}\|_{0,h,1} \|\hat{v}\|_{2,h,1} \tag{4.38}$$

for some small $\epsilon > 0$. Dividing by $\|\boldsymbol{\varphi}\|_{0,h,m}$ and using equivalence of norms, we get

$$\frac{b_h^m(\boldsymbol{\varphi}, v)}{\|\boldsymbol{\varphi}\|_{0,h,m}} \geq \frac{b_h^1(\hat{\boldsymbol{\varphi}}, v)}{\|\hat{\boldsymbol{\varphi}}\|_{0,h,1}} - Ch^{1/2-\epsilon} \|\hat{v}\|_{2,h,1}. \tag{4.39}$$

Taking the supremum, using (4.35) and equivalence of norms, proves (4.34) when h is sufficiently small. ∎

Remark 4.7. By (2.14), (4.34) holds with $\|v\|_{2,h,m}$ replaced by $|v|_{H^1(\Omega^m)}$ with a different inf-sup constant.

Therefore, (4.31), (4.32), (4.33), and (4.34) imply by the standard theory of mixed methods that (4.29) is well-posed in the mesh-dependent norms.

5. Error analysis. We now prove convergence of the HHJ method while accounting for the approximation of the domain using the theory of curved elements described in section 3. The main difficulties are dealing with higher derivatives of the nonlinear map and handling the jump terms in the mesh-dependent norms when affected by a nonlinear map. The key ingredients here are Theorem 3.2, (4.20), and the following crucial choice of optimal map: let $\widetilde{\boldsymbol{F}_T^m} : T^1 \to T^m$, for all $T^1 \in \mathcal{T}_h^1$ and $1 \le m \le k$, be given by

$$\boldsymbol{F}_T^m \equiv \widetilde{\boldsymbol{F}_T^m} := \mathcal{I}_h^{1,m} \boldsymbol{F}_T \equiv \mathcal{I}_h^{1,m} \boldsymbol{\Psi}_T^1, \tag{5.1}$$

where $\mathcal{I}_h^{1,m}$ is the Lagrange interpolation operator in (3.8) onto degree m polynomials, and we abuse notation by writing $\boldsymbol{F}_T^m \equiv \widetilde{\boldsymbol{F}_T^m}$.

Remark 5.1. Note that $\widetilde{\boldsymbol{F}_T^m}$ is an optimal map because of the approximation properties of $\mathcal{I}_h^{1,m}$; hence, the results of Theorem 3.2 apply to $\widetilde{\boldsymbol{F}_T^m}$. This choice is necessary to guarantee optimal convergence of the HHJ method when $m = r + 1$. If $m > r + 1$, the standard Lenoir map suffices.

In deriving the error estimates, we use the following regularity result for the Kirchhoff plate problem (see [5, Thm. 2], [6, Table 1]).

THEOREM 5.2. *Assume Ω satisfies the assumptions in subsection* 2.1*, and let $f \in H^{-1}(\Omega)$. Then the weak solution $w \in \mathcal{W}$ of* (2.2) *always satisfies $w \in W^{3,p}(\Omega)$ for some value of $p \in (p_0, 2]$, where $1 \le p_0 < 2$ depends on the angles at the corners of Ω.*

For technical reasons, we also assume $p > 3/2$ in Theorem 5.2 (recall Remark 4.1). Higher regularity (e.g., $w \in H^3(\Omega)$) is achieved if the corner angles are restricted. In addition, if $f \in L^2(\Omega)$, then $w \in H^4(\Omega)$. See [5, 6] for more details.

5.1. Estimate the PDE error. We start with an error estimate that ignores the geometric error; i.e., the continuous and discrete problems are posed on the exact domain.

THEOREM 5.3. *Adopt the boundary assumptions in subsection* 2.1*. Let $\boldsymbol{\sigma} \in V$ and $w \in W$ solve* (2.17) *on the true domain Ω, and assume $w \in W^{t,p}(\Omega)$, so then $\boldsymbol{\sigma} \in W^{t-2,p}(\Omega; \mathbb{S})$, $t \ge 3$, $3/2 < p \le 2$ (recall Theorem* 5.2*). Furthermore, let $r \ge 0$ be the degree of V_h, and let $\boldsymbol{\sigma}_h \in V_h$, $w_h \in W_h$ be the discrete solutions of* (4.29) *on Ω. Then we obtain*

$$\begin{aligned} &\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,h} + \|\nabla(w - w_h)\|_{L^2(\Omega)} \le C h^{\min(r+2,t-1)-2/p}, \\ &\quad \textit{when } r \ge 1\colon \quad \|w - w_h\|_{2,h} \le C h^{\min(r+1,t-1)-2/p}, \\ &\textit{when } r = 0\colon \quad \|\nabla(w - w_h)\|_{L^2(\Omega)} \le C h, \end{aligned} \tag{5.2}$$

where $C > 0$ depends on f, the domain Ω, and the shape regularity of the mesh.

Proof. With coercivity and the inf-sup condition in hand, the proof is a standard application of error estimates for mixed methods and is given in section SM6 of the supplementary materials. ∎

Note that the last line of (5.2) generalizes [6, Thm. 5.1] to curved domains.

5.2. Estimate the geometric error. We now approximate the domain using curved, Lagrange mapped triangle elements.

LEMMA 5.4. *Recall the map* $\boldsymbol{\Psi}^m : \Omega^m \to \Omega$, *with* $\boldsymbol{\Psi}^m_T := \boldsymbol{\Psi}^m|_T$, *from subsection* 3.1, *and adopt* (5.1). *For convenience, set* $\boldsymbol{J} = \nabla\boldsymbol{\Psi}^m$. *Adopt the boundary assumptions in subsection* 2.1. *Let* $\hat{\boldsymbol{\sigma}}_h \in V_h^m$, $\hat{w}_h \in W_h^m$ *be the discrete solutions of* (4.29), *with* f *replaced by* $\tilde{f} := f \circ \boldsymbol{\Psi}^m(\det \boldsymbol{J})$. *Take* $(\boldsymbol{\sigma}_h, w_h)$ *from Theorem* 5.3, *and let* $\widetilde{\boldsymbol{\sigma}}_h \in V_h^m$, $\tilde{w}_h \in W_h^m$ *be the mapped discrete solutions onto* Ω^m *using* (4.2). *In other words,* $\boldsymbol{\sigma}_h \circ \boldsymbol{\Psi}^m = (\det \boldsymbol{J})^{-2}\boldsymbol{J}\widetilde{\boldsymbol{\sigma}}_h\boldsymbol{J}^T$ *and* $\tilde{w}_h = w_h \circ \boldsymbol{\Psi}^m$, *defined elementwise. Similarly, we map the test functions* $\boldsymbol{\varphi}_h \in V_h$, $v_h \in W_h$ *to* $\hat{\boldsymbol{\varphi}}_h \in V_h^m$, $\hat{v}_h \in W_h^m$. *Then we obtain the error equations for the geometric error:*

$$a^m\left(\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h, \hat{\boldsymbol{\varphi}}_h\right) + b_h^m\left(\hat{\boldsymbol{\varphi}}_h, \tilde{w}_h - \hat{w}_h\right) + b_h^m\left(\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h, \hat{v}_h\right) = \mathrm{E}_0(\hat{\boldsymbol{\varphi}}_h, \hat{v}_h) \tag{5.3}$$

for all $(\hat{v}_h, \hat{\boldsymbol{\varphi}}_h) \in W_h^m \times V_h^m$, *where*

$$|\mathrm{E}_0(\hat{\boldsymbol{\varphi}}_h, \hat{v}_h)| \le Ch^q\left(\|\hat{\boldsymbol{\varphi}}_h\|_{0,h,m} + \|\hat{v}_h\|_{2,h,m}\right)\|f\|_{H^{-1}(\Omega)}, \tag{5.4}$$

where $q = m$ *when* $m = r+1$; *otherwise,* $q = m-1$.

Proof. We will use (4.6) with m, l replaced by ∞, m, respectively. First, note that $(\mathrm{id}_{\Omega^m} - \boldsymbol{\Psi}^m) \cdot \mathrm{P}_0\nabla\hat{v}_h \in H_h^2(\Omega^m)$ because $(\mathrm{id}_{T^m} - \boldsymbol{\Psi}_T^m)$ is zero at all internal edges and is identically zero on all elements not in $\mathcal{T}_{\partial,h}^m$. Upon noting $v_h \in \mathring{H}^1(\Omega)$ and (4.36), straightforward manipulation gives

$$\begin{aligned} b_h\left(\boldsymbol{\varphi}_h, v_h\right) = b_h^m\left(\hat{\boldsymbol{\varphi}}_h, \hat{v}_h\right) + b_h^m\left(\hat{\boldsymbol{\varphi}}_h, (\mathrm{id}_{\Omega^m} - \boldsymbol{\Psi}^m) \cdot \mathrm{P}_0\nabla\hat{v}_h\right) \\ + O(h^m)\|\hat{\boldsymbol{\varphi}}_h\|_{0,h,m}\|\hat{v}_h\|_{2,h,m}. \end{aligned} \tag{5.5}$$

Taking advantage of (3.4), we get

$$\begin{aligned} b_h\left(\boldsymbol{\varphi}_h, v_h\right) = b_h^m\left(\hat{\boldsymbol{\varphi}}_h, \hat{v}_h\right) + b_h^1\left(\check{\boldsymbol{\varphi}}_h, (\boldsymbol{F}^m - \boldsymbol{F}) \cdot \mathrm{P}_0\nabla\hat{v}_h\right) \\ + O(h^m)\|\hat{\boldsymbol{\varphi}}_h\|_{0,h,m}\|\hat{v}_h\|_{2,h,m}, \end{aligned} \tag{5.6}$$

where $\check{\boldsymbol{\varphi}}_h \in V_h^1$ and $\boldsymbol{F}^m := \mathcal{I}_h^{1,m}\boldsymbol{F}$, by (5.1). If $m = r+1$, the Fortin property (4.20) yields $b_h^1\left(\check{\boldsymbol{\varphi}}_h, (\boldsymbol{F}^m - \boldsymbol{F}) \cdot \mathrm{P}_0\nabla\hat{v}_h\right) = 0$. If $m \ne r+1$, then a straightforward estimate shows that $b_h^1\left(\check{\boldsymbol{\varphi}}_h, (\boldsymbol{F}^m - \boldsymbol{F}) \cdot \mathrm{P}_0\nabla\hat{v}_h\right) \le Ch^{m-1}\|\hat{\boldsymbol{\varphi}}_h\|_{0,h,m}\|\hat{v}_h\|_{2,h,m}$, where we used equivalence of norms (3.6) and (4.4).

Therefore, using (4.5) and (4.16), the first line in (4.29) (with $m = \infty$) maps to

$$a^m\left(\widetilde{\boldsymbol{\sigma}}_h, \hat{\boldsymbol{\varphi}}_h\right) + b_h^m\left(\hat{\boldsymbol{\varphi}}_h, \tilde{w}_h\right) = \mathrm{I}_1 \quad \forall\hat{\boldsymbol{\varphi}}_h \in V_h^m, \tag{5.7}$$

where $1 \le m \le k$ and $C > 0$ is a constant depending only on Ω such that

$$\mathrm{I}_1 \le Ch^q\|\hat{\boldsymbol{\varphi}}_h\|_{L^2(\Omega_S^m)}\left(\|\widetilde{\boldsymbol{\sigma}}_h\|_{L^2(\Omega_S^m)} + \|\tilde{w}_h\|_{2,h,m}\right), \tag{5.8}$$

where q was defined earlier. The second equation in (4.29) (with $m = \infty$) maps to

$$b_h^m\left(\widetilde{\boldsymbol{\sigma}}_h, \hat{v}_h\right) = -\left\langle f \circ \boldsymbol{\Psi}^m(\det \boldsymbol{J}), \hat{v}_h\right\rangle_{\Omega^m} + \mathrm{I}_2 \quad \forall\hat{v}_h \in W_h^m, \tag{5.9}$$

where, for some constant $C > 0$ depending only on Ω,

$$\mathrm{I}_2 \le Ch^q\|\widetilde{\boldsymbol{\sigma}}_h\|_{L^2(\Omega_S^m)}\|\hat{v}_h\|_{2,h,m}. \tag{5.10}$$

Then subtracting (4.29) (with $1 \le m \le k$) for the solution $(\hat{\boldsymbol{\sigma}}_h, \hat{w}_h)$ from the above equations, combining everything, and noting the a priori estimate (4.30) gives (5.3) and (5.4). ∎

THEOREM 5.5. *Adopt the hypothesis of Lemma* 5.4. *Then the following error estimate holds:*

$$\|\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h\|_{0,h,m} + \|\tilde{w}_h - \hat{w}_h\|_{2,h,m} \le Ch^q \|f\|_{H^{-1}(\Omega)} \tag{5.11}$$

for some uniform constant $C > 0$.

Proof. From (5.3), choose $\hat{v}_h = 0$ and use Lemma 4.6 to get

$$\begin{aligned} \beta_0 \|\tilde{w}_h - \hat{w}_h\|_{2,h,m} &\le \sup_{\hat{\boldsymbol{\varphi}}_h \in V_h^m} \frac{|b_h^m(\hat{\boldsymbol{\varphi}}_h, \tilde{w}_h - \hat{w}_h)|}{\|\hat{\boldsymbol{\varphi}}_h\|_{0,h,m}} \\ &\le \sup_{\hat{\boldsymbol{\varphi}}_h \in V_h^m} \frac{|a^m(\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h, \hat{\boldsymbol{\varphi}}_h)| + |\mathrm{E}_0(\hat{\boldsymbol{\varphi}}_h, 0)|}{\|\hat{\boldsymbol{\varphi}}_h\|_{0,h,m}} \\ &\le CA_0 \|\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h\|_{0,h,m} + Ch^q \|f\|_{H^{-1}(\Omega)}, \end{aligned} \tag{5.12}$$

where we used the norm equivalence (4.4). Next, choose $\hat{\boldsymbol{\varphi}}_h = \widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h$ and $\hat{v}_h = -(\tilde{w}_h - \hat{w}_h)$ in (5.3) to get

$$\begin{aligned} \alpha_0 \|\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h\|^2_{L^2(\Omega^m)} &\le a^m(\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h, \widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h) \\ &\le Ch^q \left(\|\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h\|_{0,h,m} + \|\tilde{w}_h - \hat{w}_h\|_{2,h,m}\right) \|f\|_{H^{-1}(\Omega)} \\ &\le Ch^q \left(\|\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h\|_{0,h,m} + Ch^q \|f\|_{H^{-1}(\Omega)}\right) \|f\|_{H^{-1}(\Omega)} \\ &\le C(h^q)^2 \|f\|^2_{H^{-1}(\Omega)} + Ch^q \|\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h\|_{L^2(\Omega^m)} \|f\|_{H^{-1}(\Omega)} \\ &\le C(h^q)^2 \|f\|^2_{H^{-1}(\Omega)} + \frac{\alpha_0}{2} \|\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h\|^2_{L^2(\Omega^m)}, \end{aligned} \tag{5.13}$$

where we used (5.12), norm equivalence (4.4), and a weighted Cauchy inequality. Then, by combining the above results, we get the assertion. ∎

5.3. Estimate the total error. We will combine Theorems 5.3 and 5.5 to get the total error.

THEOREM 5.6 (general error estimate). *Adopt the hypotheses of Theorem* 5.3 *and Lemma* 5.4. *If* $m \ge r + 1$, *then*

$$\begin{aligned} \|\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}}_h \circ (\boldsymbol{\Psi}^m)^{-1}\|_{0,h} + \|\nabla(w - \hat{w}_h \circ (\boldsymbol{\Psi}^m)^{-1})\|_{L^2(\Omega)} &\le Ch^{\min(r+2, t-1)-2/p}, \\ r \ge 1\colon \quad \|w - \hat{w}_h \circ (\boldsymbol{\Psi}^m)^{-1}\|_{2,h} &\le Ch^{\min(r+1, t-1)-2/p}, \\ r = 0\colon \quad \|\nabla(w - \hat{w}_h \circ (\boldsymbol{\Psi}^m)^{-1})\|_{L^2(\Omega)} &\le Ch, \end{aligned} \tag{5.14}$$

where $C > 0$ *depends on* f, *the domain* Ω, *and the shape regularity of the mesh.*

Proof. By the triangle inequality and using the properties of the map $\boldsymbol{\Psi}^m$, we have

$$\begin{aligned} \|\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}}_h \circ (\boldsymbol{\Psi}^m)^{-1}\|_{0,h} &\le \|\boldsymbol{\sigma} - \widetilde{\boldsymbol{\sigma}}_h \circ (\boldsymbol{\Psi}^m)^{-1}\|_{0,h} \\ &\quad + \|\widetilde{\boldsymbol{\sigma}}_h \circ (\boldsymbol{\Psi}^m)^{-1} - \hat{\boldsymbol{\sigma}}_h \circ (\boldsymbol{\Psi}^m)^{-1}\|_{0,h} \\ &\le \|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{0,h} + \|\boldsymbol{\sigma}_h - \widetilde{\boldsymbol{\sigma}}_h \circ (\boldsymbol{\Psi}^m)^{-1}\|_{0,h} + C\|\widetilde{\boldsymbol{\sigma}}_h - \hat{\boldsymbol{\sigma}}_h\|_{0,h,m}. \end{aligned} \tag{5.15}$$

Focusing on the middle term, we have

$$\begin{aligned} \|\boldsymbol{\sigma}_h - \widetilde{\boldsymbol{\sigma}}_h \circ (\boldsymbol{\Psi}^m)^{-1}\|_{0,h} &\le \|\boldsymbol{\sigma}_h \circ \boldsymbol{\Psi}^m - \widetilde{\boldsymbol{\sigma}}_h\|_{0,h,m} \\ &\le \|(\det \boldsymbol{J})^{-2} \boldsymbol{J} \widetilde{\boldsymbol{\sigma}}_h \boldsymbol{J}^T - \widetilde{\boldsymbol{\sigma}}_h\|_{0,h,m} \le Ch^{r+1} \|\widetilde{\boldsymbol{\sigma}}_h\|_{0,h,m} \le Ch^{r+1} \|f\|_{H^{-1}(\Omega)}. \end{aligned} \tag{5.16}$$

Combining everything, we get

$$\|\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}}_h \circ (\boldsymbol{\Psi}^m)^{-1}\|_{0,h} \leq C \max\left(h^{r+1}, h^{\min(r+2,t-1)-2/p}\right), \tag{5.17}$$

where $C > 0$ depends on f. Taking a similar approach for the other terms involving $w - \hat{w}_h \circ (\boldsymbol{\Psi}^m)^{-1}$ delivers the estimates. ∎

COROLLARY 5.7. *Adopt the hypothesis of Theorem* 5.6*, but assume* Γ *is globally smooth and* f*,* w*, and* $\boldsymbol{\sigma}$ *are smooth. If* $r \geq 0$ *is the degree of* V_h*, then*

$$\begin{aligned}\|\boldsymbol{\sigma} - \hat{\boldsymbol{\sigma}}_h \circ (\boldsymbol{\Psi}^m)^{-1}\|_{0,h} + \|\nabla(w - \hat{w}_h \circ (\boldsymbol{\Psi}^m)^{-1})\|_{L^2(\Omega)} \\ + h\|w - \hat{w}_h \circ (\boldsymbol{\Psi}^m)^{-1}\|_{2,h} \leq Ch^{r+1},\end{aligned} \tag{5.18}$$

where $C > 0$ *depends on* w*, the domain* Ω*, and the shape regularity of the mesh.*

Remark 5.8. From Theorem 5.5, if $m < r + 1$, the error is suboptimal, i.e., is $O(h^{m-1})$ for a smooth solution. However, this only occurs in $\mathcal{T}_{\partial,h}$; in the rest of the mesh, it is $O(h^{r+1})$. Since the mesh is quasi-uniform, a straightforward estimate gives that the error measured over the entire domain is $O(h^{m-1/2})$. This is verified in the simply supported numerical example in subsection 6.2.2, as well as in both examples in subsections 6.3.1 and 6.3.2.

5.4. Inhomogeneous boundary conditions. We now explain how to extend the above theory to handle nonvanishing boundary conditions. First, construct a function $g \in W^{t,p}(\Omega)$, such that the displacement satisfies $w = g$ on Γ and $\partial_{\boldsymbol{n}} w = \partial_{\boldsymbol{n}} g$ on Γ_{c}, and construct a function $\boldsymbol{\rho} \in W^{t-2,p}(\Omega; \mathbb{S})$, such that the normal-normal moment satisfies $\boldsymbol{n}^T \boldsymbol{\sigma} \boldsymbol{n} = \boldsymbol{n}^T \boldsymbol{\rho} \boldsymbol{n}$ on Γ_{s}, where $t \geq 3$, $3/2 < p \leq 2$ (recall Theorem 5.2).

Then (2.17) is replaced by the problem of determining $(\boldsymbol{\sigma}, w) = (\mathring{\boldsymbol{\sigma}} + \boldsymbol{\rho}, \mathring{w} + g)$, with $\mathring{\boldsymbol{\sigma}} \in \mathcal{V}_h$, $\mathring{w} \in \mathcal{W}_h$ (i.e., with homogeneous boundary conditions) such that

$$\begin{aligned} a(\mathring{\boldsymbol{\sigma}}, \boldsymbol{\varphi}) + b_h(\boldsymbol{\varphi}, \mathring{w}) &= -a(\boldsymbol{\rho}, \boldsymbol{\varphi}) - b_h(\boldsymbol{\varphi}, g) + (\varphi^{\mathrm{nn}}, \boldsymbol{n} \cdot \nabla g)_{\Gamma_{\mathrm{c}}} \quad \forall \boldsymbol{\varphi} \in \mathcal{V}_h, \\ b_h(\mathring{\boldsymbol{\sigma}}, v) &= -\langle f, v \rangle_\Omega - b_h(\boldsymbol{\rho}, v) \quad \forall v \in \mathcal{W}_h. \end{aligned} \tag{5.19}$$

Note that the right-hand side in the first equation of (5.19) simplifies to $-a(\boldsymbol{\rho}, \boldsymbol{\varphi}) - \mathring{b}_h(\boldsymbol{\varphi}, g)$, where $\mathring{b}_h(\boldsymbol{\varphi}, v) := b_h(\boldsymbol{\varphi}, v) - (\varphi^{\mathrm{nn}}, \boldsymbol{n} \cdot \nabla v)_{\Gamma_{\mathrm{c}}}$ (i.e., it has no boundary term).

Similarly, the corresponding (intermediate) discrete problem (4.29), on the exact domain, is replaced by finding $(\boldsymbol{\sigma}_h, w_h) = (\mathring{\boldsymbol{\sigma}}_h + \boldsymbol{\rho}_h, \mathring{w}_h + g_h)$, with $\mathring{\boldsymbol{\sigma}}_h \in V_h$, $\mathring{w}_h \in W_h$ such that

$$\begin{aligned} a(\mathring{\boldsymbol{\sigma}}_h, \boldsymbol{\varphi}_h) + b_h(\boldsymbol{\varphi}_h, \mathring{w}_h) &= -a(\boldsymbol{\rho}_h, \boldsymbol{\varphi}_h) - \mathring{b}_h(\boldsymbol{\varphi}_h, g_h) \\ &\quad - (\varphi_h^{\mathrm{nn}}, \boldsymbol{n} \cdot \nabla g_h)_{\Gamma_{\mathrm{c}}} + (\varphi_h^{\mathrm{nn}}, \boldsymbol{n} \cdot \nabla g)_{\Gamma_{\mathrm{c}}} \quad \forall \boldsymbol{\varphi}_h \in V_h, \\ b_h(\mathring{\boldsymbol{\sigma}}_h, v_h) &= -\langle f, v_h \rangle_\Omega - b_h(\boldsymbol{\rho}_h, v_h) \quad \forall v_h \in W_h, \end{aligned} \tag{5.20}$$

where $\boldsymbol{\rho}_h = \mathrm{P}_h \boldsymbol{\rho}$, and $\mathrm{P}_h : H_h^0(\Omega) \to V_h$ is the $L^2(\Omega)$ projection; i.e., $\boldsymbol{\rho}_h$ satisfies

$$(\boldsymbol{\rho}_h - \boldsymbol{\rho}, \boldsymbol{\varphi}_h)_{\mathcal{T}_h} + \langle \boldsymbol{n}^T [\boldsymbol{\rho}_h - \boldsymbol{\rho}] \boldsymbol{n}, \varphi_h^{\mathrm{nn}} \rangle_{\mathcal{E}_h} = 0 \quad \forall \boldsymbol{\varphi}_h \in V_h, \tag{5.21}$$

and $g_h = \mathcal{I}_h g$. An error estimate between the solutions of (5.19) and (5.20), analogous to Theorem 5.3, follows similarly with the following additional steps. First, estimate $b_h(\boldsymbol{\rho} - \boldsymbol{\rho}_h, v_h) \leq \|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,h} \|v_h\|_{2,h}$, note that $\|\boldsymbol{\rho} - \boldsymbol{\rho}_h\|_{0,h} \leq \|\boldsymbol{\rho} - \Pi_h \boldsymbol{\rho}\|_{0,h}$, and use the approximation properties of Π_h in subsection SM4.3 of the supplementary materials. Next, estimate $\mathring{b}_h(\boldsymbol{\varphi}_h, g - g_h)$ and $(\varphi_h^{\mathrm{nn}}, \boldsymbol{n} \cdot \nabla(g - g_h))_{\Gamma_{\mathrm{c}}}$ with (4.21).

Finally, the discrete problem on the discrete domain is to find $(\hat{\boldsymbol{\sigma}}_h, \hat{w}_h) = (\mathring{\hat{\boldsymbol{\sigma}}}_h + \hat{\boldsymbol{\rho}}_h, \mathring{\hat{w}}_h + \hat{g}_h)$, with $\mathring{\hat{\boldsymbol{\sigma}}}_h \in V_h^m$, $\mathring{\hat{w}}_h \in W_h^m$ such that

$$\begin{aligned} a^m\left(\mathring{\hat{\boldsymbol{\sigma}}}_h, \hat{\boldsymbol{\varphi}}_h\right) + b_h^m\left(\hat{\boldsymbol{\varphi}}_h, \mathring{\hat{w}}_h\right) &= -a^m\left(\hat{\boldsymbol{\rho}}_h, \hat{\boldsymbol{\varphi}}_h\right) - \mathring{b}_h^m\left(\hat{\boldsymbol{\varphi}}_h, \hat{g}_h\right) \\ &\quad - \left(\hat{\varphi}_h^{\mathrm{nn}}, \hat{\boldsymbol{n}} \cdot [\nabla \hat{g}_h - \tilde{\boldsymbol{\xi}}]\right)_{\Gamma_{\mathrm{c}}^m} \quad \forall \hat{\boldsymbol{\varphi}}_h \in V_h^m, \\ b_h^m\left(\mathring{\hat{\boldsymbol{\sigma}}}_h, \hat{v}_h\right) &= -\left\langle \tilde{f}, \hat{v}_h \right\rangle_{\Omega^m} - b_h^m\left(\hat{\boldsymbol{\rho}}_h, \hat{v}_h\right) \quad \forall \hat{v}_h \in W_h^m, \end{aligned} \tag{5.22}$$

where $\hat{\boldsymbol{\rho}}_h := \mathrm{P}_h^m \widetilde{\boldsymbol{\rho}}$, with $\widetilde{\boldsymbol{\rho}}$ given by $\boldsymbol{\rho} \circ \boldsymbol{\Psi}^m|_T = (\det \boldsymbol{J}_T)^{-2} \boldsymbol{J}_T \widetilde{\boldsymbol{\rho}} \boldsymbol{J}_T^T$, $\boldsymbol{J}_T = \nabla \boldsymbol{\Psi}^m|_T$, and $\mathrm{P}_h^m : H_h^0(\Omega^m) \to V_h^m$ is the $L^2(\Omega^m)$ projection on Ω^m, $\hat{g}_h := \mathcal{I}_h^m \tilde{g}$, with $\tilde{g} := g \circ \boldsymbol{\Psi}^m$, and $\tilde{\boldsymbol{\xi}} = (\nabla g) \circ \boldsymbol{\Psi}^m$. To obtain a result analogous to Theorem 5.6, we need to generalize Lemma 5.4. The argument is mostly the same as the proof of Lemma 5.4, except there is an additional step to show that

$$a^m\left(\widetilde{\boldsymbol{\rho}} - \hat{\boldsymbol{\rho}}_h, \hat{\boldsymbol{\varphi}}_h\right) + \mathring{b}_h^m\left(\hat{\boldsymbol{\varphi}}_h, \tilde{g} - \hat{g}_h\right) + b_h^m\left(\widetilde{\boldsymbol{\rho}} - \hat{\boldsymbol{\rho}}_h, \hat{v}_h\right) = \mathrm{E}_1(\hat{\boldsymbol{\varphi}}_h, \hat{v}_h) \tag{5.23}$$

for all $(\hat{v}_h, \hat{\boldsymbol{\varphi}}_h) \in W_h^m \times V_h^m$, where

$$|\mathrm{E}_1(\hat{\boldsymbol{\varphi}}_h, \hat{v}_h)| \leq C h^q \left(\|\hat{\boldsymbol{\varphi}}_h\|_{0,h,m} + \|\hat{v}_h\|_{2,h,m}\right) \left(\|\boldsymbol{\rho}\|_{W^{1,p}(\Omega;\mathbb{S})} + \|g\|_{W^{3,p}(\Omega)}\right), \tag{5.24}$$

where $q = m$ when $m = r+1$; otherwise, $q = m-1$. This also follows the same outline, but we note the following. (1) Estimating $\mathring{b}_h^m\left(\hat{\boldsymbol{\varphi}}_h, \tilde{g} - \hat{g}_h\right)$ with (4.6) is simpler because the last boundary term in (4.6) does not appear; then use Lemma 4.4. (2) Noting that $\hat{g}_h = g_h \circ \boldsymbol{\Psi}^m$, $(\varphi_h^{\mathrm{nn}}, \boldsymbol{n} \cdot \nabla(g_h - g))_{\Gamma_{\mathrm{c}}}$ is mapped to $(\hat{\varphi}_h^{\mathrm{nn}}, \hat{\boldsymbol{n}} \cdot \nabla(\hat{g}_h - \tilde{g}))_{\Gamma_{\mathrm{c}}^m}$ (plus residual terms) and is compared against $\left(\hat{\varphi}_h^{\mathrm{nn}}, \hat{\boldsymbol{n}} \cdot [\nabla \hat{g}_h - \tilde{\boldsymbol{\xi}}]\right)_{\Gamma_{\mathrm{c}}^m}$. (3) Finally, estimate $\left(\hat{\varphi}_h^{\mathrm{nn}}, \hat{\boldsymbol{n}} \cdot [\nabla \tilde{g} - \tilde{\boldsymbol{\xi}}]\right)_{\Gamma_{\mathrm{c}}^m}$ using arguments similar to those in the proof of Theorem 4.3. With this and the obvious generalization of Theorem 5.5, we obtain the following.

THEOREM 5.9 (inhomogeneous boundary conditions). *Adopt the hypotheses of Theorem* 5.6*, except assume that* $(\boldsymbol{\sigma}, w)$ *solves* (5.19) *and* $(\hat{\boldsymbol{\sigma}}_h, \hat{w}_h)$ *solves* (5.22)*. If* $m \geq r+1$*, then* $(\boldsymbol{\sigma}, w)$ *and* $(\hat{\boldsymbol{\sigma}}_h, \hat{w}_h)$ *satisfy the same estimates as in* (5.14)*.*

We also obtain a corollary directly analogous to Corollary 5.7.

6. Numerical results. We present numerical examples computed on a disk, as well as on a nonsymmetric domain. The discrete domains were generated by a successive uniform refinement scheme, with curved elements generated using a variant of the procedure in [18, sect. 3.2]. As above, the finite element spaces V_h and W_h are of degrees r and $r+1$, respectively, where $r \geq 0$, and the geometric approximation degree is denoted by m. All computations were done with the MATLAB/C++ finite element toolbox FELICITY [27], where we used the "backslash" command in MATLAB to solve the linear systems.

From (5.1), recall that $\boldsymbol{F}^m := \mathcal{I}_h^{1,m} \boldsymbol{\Psi}^1$, which is plausible to implement but inconvenient. Instead, we first compute $\boldsymbol{F}^{m+1}$ using the procedure in [18, sect. 3.2]; then we define $\boldsymbol{F}^m := \mathcal{I}_h^{1,m} \boldsymbol{F}^{m+1}$, which is easy to implement because they are standard Lagrange spaces. Moreover, the accuracy is not affected. As for the boundary data, $\hat{g}_h$, $\tilde{\boldsymbol{\xi}}$, and $\hat{\boldsymbol{\rho}}_h$ only need to be computed on the boundary Γ^m; in fact, only the boundary part of the L^2 projection P_h^m needs to be computed.

6.1. Practical error estimates. For convenience, the errors we compute are $\|\nabla(w-\hat{w}_h)\|_{L^2(\Omega^m)}$, $|w-\hat{w}_h|_{H^2(\mathcal{T}_h^m)}$, $\|\boldsymbol{\sigma}-\hat{\boldsymbol{\sigma}}_h\|_{L^2(\Omega^m)}$, and $h^{1/2}\|\boldsymbol{n}_h^T(\boldsymbol{\sigma}-\hat{\boldsymbol{\sigma}}_h)\boldsymbol{n}_h\|_{L^2(\mathcal{E}_h^m)}$, where the exact solution has been extended by analytic continuation. These errors can be related to the ones in (5.18) by basic arguments and a triangle inequality. Essentially, we need to bound the error between $(\boldsymbol{\sigma}_h, w_h)$ and $(\tilde{\boldsymbol{\sigma}}_h, \tilde{w}_h)$ in a sense clarified by the following result.

PROPOSITION 6.1. *Let f_h be a piecewise (possibly mapped) polynomial defined over $\mathcal{T}_h$, and let $\boldsymbol{\Psi}^m : \Omega^m \to \Omega$ be the piecewise element mapping from subsection* 3.1. *Assume f_h has a bounded extension to Ω^m. Then*

$$\begin{aligned} \|\nabla^s(f_h - f_h\circ\boldsymbol{\Psi}^m)\|_{L^2(\Omega^m)} &\le Ch^m\|\nabla^s f_h\|_{L^2(\widetilde{\Omega}^m)} \text{ for } s=0,1,\\ \|f_h - f_h\circ\boldsymbol{\Psi}^m\|_{H_h^2(\Omega^m)} &\le Ch^{m-1}\|f_h\|_{H_h^2(\widetilde{\Omega}^m)}, \end{aligned} \tag{6.1}$$

where $\widetilde{\Omega}^m = \bigcup_{T^m\in\mathcal{T}_h^m}\operatorname{conv}(T^m)$ and $\operatorname{conv}(T^m)$ is the convex hull of T^m.

6.2. Unit disk domain. The disk has an unexpected symmetry with respect to curved elements. When approximating Γ by polynomials of degree $m = 2q$, the approximation order is actually $m = 2q+1$. This is because each circular arc of Γ, when viewed as a graph over a flat edge in $\mathcal{E}^1_{\partial,h}$, is symmetric about the midpoint of the edge. Thus, since the Lagrange interpolation nodes are placed symmetrically on the edge, the resulting interpolant must be of even degree. In other words, Ω^2, Ω^4 have the same approximation order as Ω^3, Ω^5, respectively. Our numerical results reflect this.

6.2.1. The homogeneous clamped disk. In this example, the exact solution with clamped boundary conditions on Γ, written in polar coordinates, is taken to be

$$w(r,\theta) = \sin^2(\pi r). \tag{6.2}$$

Table 1 shows the estimated orders of convergence (EoC), which were computed by evaluating the ratio of the errors between the last two meshes in a sequence of successively, uniformly refined meshes. The optimal orders of convergence, based on the degree of the elements, is $r+1$ for the three quantities $|\hat{w}_h|_{H^1}$, $|\hat{\boldsymbol{\sigma}}_h|_{L^2}$, and $h^{1/2}|\hat{\sigma}_h^{\mathrm{nn}}|_{L^2(\mathcal{E}_h^m)}$ and is r for $|\hat{w}_h|_{H_h^2}$; note that we use the abbreviation $|\hat{w}_h|_{H^1} \equiv |w-\hat{w}_h|_{H^1(\Omega^m)}$, etc. The convergence is better than expected, in that we do not see reduced order convergence when $m = r$, possibly due to the clamped boundary conditions and/or the choice of exact solution (also recall the symmetry discussion earlier). We do see reduced convergence when $m = r-1$.

6.2.2. The homogeneous simply supported disk. The exact solution with simply supported boundary conditions on Γ, written in polar coordinates, is

$$w(r,\theta) = \cos((3/2)\pi r). \tag{6.3}$$

Table 2 shows the EoC. The convergence order is consistent with the error estimate in (5.18) (accounting for the symmetry of the disk). For example, when $m = r = 1$, we see $O(h^{1/2})$ for $\hat{\boldsymbol{\sigma}}_h$ (see Remark 5.8). The convergence rate for $\hat{w}_h$ is not reduced, but it is not optimal. When $m = r = 3$, $\boldsymbol{\sigma}$ converges with $O(h^{5/2})$ (consistent with Remark 5.8), yet $\hat{w}_h$ performs better. The "improved" error convergence for $\hat{w}_h$ could be due to the particular choice of exact solution.

TABLE 1

EoC for the clamped disk. N_T is the number of triangles in the final mesh after multiple uniform refinements. Asterisks indicate suboptimal orders, and italics indicate the case $m = r + 1$, for which optimality is proven in this paper.

N_T	m	r	$\lvert\hat{w}_h\rvert_{H^1}$	$\lvert\hat{w}_h\rvert_{H_h^2}$	$\lvert\hat{\boldsymbol{\sigma}}_h\rvert_{L^2}$	$h^{1/2}\lvert\hat{\sigma}_h^{\mathrm{nn}}\rvert_{L^2(\mathcal{E}_h^m)}$
2^{17}	1	0	*1.0002*	*0.0000*	*0.9997*	*1.0007*
2^{15}	1	1	2.0006	0.9998	1.9978	2.0312
2^{15}	1	2	2.0052*	1.9980*	1.5121*	1.5022*
2^{15}	2	1	*2.0002*	*0.9990*	*1.9976*	*2.0317*
2^{15}	2	2	2.9984	1.9985	2.9994	2.9934
2^{13}	2	3	4.0039	3.0007	3.9907	4.0853
2^{15}	3	2	*2.9984*	*1.9985*	*2.9994*	*2.9934*
2^{13}	3	3	4.0039	3.0007	3.9906	4.0746
2^{13}	3	4	4.9862	3.9881	3.8387*	3.5173*
2^{13}	4	3	*4.0038*	*3.0006*	*3.9908*	*4.0975*
2^{13}	4	4	4.9868	3.9883	5.0022	4.9824
2^{13}	5	4	*4.9868*	*3.9883*	*5.0022*	*4.9823*

TABLE 2

EoC for the simply supported disk. N_T is the number of triangles in the final mesh after multiple uniform refinements. Asterisks indicate suboptimal orders, and italics indicate $m = r + 1$.

N_T	m	r	$\lvert\hat{w}_h\rvert_{H^1}$	$\lvert\hat{w}_h\rvert_{H_h^2}$	$\lvert\hat{\boldsymbol{\sigma}}_h\rvert_{L^2}$	$h^{1/2}\lvert\hat{\sigma}_h^{\mathrm{nn}}\rvert_{L^2(\mathcal{E}_h^m)}$
2^{17}	1	0	*1.0002*	*0.0000*	*0.9997*	*1.0016*
2^{15}	1	1	1.0827*	0.6840*	0.4976*	0.4835*
2^{15}	1	2	1.0297*	0.4920*	0.4926*	0.4775*
2^{15}	2	1	*1.9997*	*0.9996*	*1.9988*	*2.0202*
2^{15}	2	2	3.0001	1.9987	2.9974	2.9918
2^{13}	2	3	3.9793	2.9819	2.5704*	2.4795*
2^{15}	3	2	*3.0001*	*1.9987*	*2.9976*	*2.9930*
2^{13}	3	3	3.9789	2.9820	2.5780*	2.4783*
2^{13}	3	4	3.5159*	2.5344*	2.5008*	2.4952*
2^{13}	4	3	*3.9896*	*2.9916*	*4.0010*	*4.0354*
2^{13}	4	4	5.0107	4.0067	4.9846	4.9747
2^{13}	5	4	*5.0107*	*4.0067*	*4.9849*	*4.9772*

6.3. Three-leaf domain. The boundary of Ω is parameterized by

$$x(t) = [1 + 0.4\cos(3t)]\cos(t), \quad y(t) = [1 + (0.4 + 0.22\sin(t))\cos(3t)]\sin(t) \tag{6.4}$$

for $0 \le t \le 2\pi$ (see Figure 4). This domain does not have the additional symmetry of the disk.

6.3.1. Inhomogeneous clamped boundary conditions. The exact solution is taken to be

$$w(x, y) = \sin(2\pi x)\cos(2\pi y), \tag{6.5}$$

with the corresponding boundary conditions. Table 3 shows the EoC. The order of convergence is reduced, as expected, for $\hat{\boldsymbol{\sigma}}_h$; e.g., when $m = 4$, the convergence rate *goes down* (from $O(h^4)$ to $O(h^{3.5})$) when r *increases* from $r = 3$ to $r = 4$. One would expect the convergence rate to at least stay the same. The reason for this is connected to estimating the term (5.6) in the proof of Lemma 5.4, where if $m \neq r + 1$, then the geometric error is $O(h^{m-1})$. So, in the example above, the error should go down to $O(h^3)$. However, this geometric error is concentrated in the elements adjacent to the

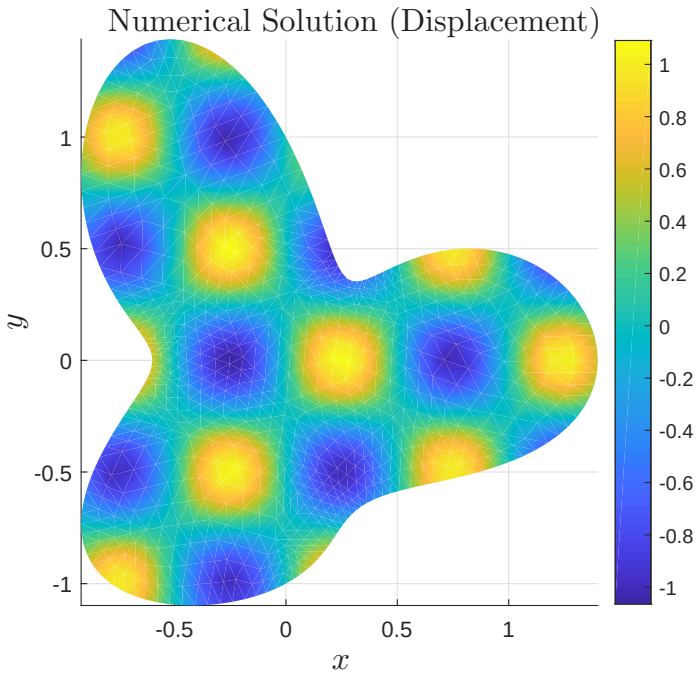

FIG. 4. *Illustration of the three-leaf domain. The numerical solution $\hat{w}_h$ approximating* (6.5) *is shown.*

boundary only, so it is only $O(h^{3.5})$; see Remark 5.8. Recall that when $m = r+1$, the Fortin property (4.20) applies, and the geometric error is $O(h^m)$. When $m > r+1$, then $O(h^{m-1}) = O(h^{r+1})$, which is sufficient (see the hypothesis of Theorem 5.9). However, the convergence order of $\hat{w}_h$ is better than expected, although it is reduced when $m = r-1$.

TABLE 3

EoC for the clamped three-leaf domain. N_T is the number of triangles in the final mesh after multiple uniform refinements. Asterisks indicate suboptimal orders, and italics indicate $m = r+1$.

N_T	m	r	$\lvert\hat{w}_h\rvert_{H^1}$	$\lvert\hat{w}_h\rvert_{H^2_h}$	$\lvert\hat{\boldsymbol{\sigma}}_h\rvert_{L^2}$	$h^{1/2}\lvert\hat{\sigma}_h^{\mathrm{nn}}\rvert_{L^2(\mathcal{E}_h^m)}$
368640	1	0	*1.0009*	*0.0000*	*0.9993*	*1.0072*
92160	1	1	1.8821*	0.8326*	0.5016*	0.4787*
92160	1	2	1.5191*	0.5045*	0.4966*	0.4809*
92160	2	1	*2.0009*	*1.0010*	*1.9942*	*2.0633*
92160	2	2	2.9808	1.9271*	1.5048*	1.5069*
23040	2	3	2.5621*	1.5496*	1.5001*	1.5195*
92160	3	2	*2.9996*	*2.0000*	*2.9983*	*2.9908*
23040	3	3	3.9900	2.9619	2.5301*	2.4545*
23040	3	4	3.5672*	2.5239*	2.4872*	2.4317*
23040	4	3	*3.9972*	*2.9985*	*3.9933*	*4.1447*
23040	4	4	4.9893	3.9460	3.5022*	3.5105*
23040	5	4	*4.9989*	*4.0003*	*4.9959*	*4.9710*

6.3.2. Inhomogeneous simply supported boundary conditions. The exact solution is taken to be

$$w(x,y) = \sin(2\pi x)\cos(2\pi y) \tag{6.6}$$

with the corresponding boundary conditions. Table 4 shows the EoC. The convergence order is consistent with the error estimate in (5.18). When $m = r = 1$, we see $O(h^{1/2})$ for $\hat{\boldsymbol{\sigma}}_h$ (see Remark 5.8). The convergence rate for $\hat{w}_h$ is not reduced, but it is not optimal. When $m = r = 3$, $\boldsymbol{\sigma}$ converges with $O(h^{5/2})$ (consistent with Remark 5.8), yet $\hat{w}_h$ performs better. The "improved" error convergence for $\hat{w}_h$ could be due to the particular choice of exact solution.

TABLE 4

EoC for the simply supported three-leaf domain. N_T is the number of triangles in the final mesh after multiple uniform refinements. Asterisks indicate suboptimal orders, and italics indicate $m = r + 1$.

N_T	m	r	$\lvert\hat{w}_h\rvert_{H^1}$	$\lvert\hat{w}_h\rvert_{H_h^2}$	$\lvert\hat{\boldsymbol{\sigma}}_h\rvert_{L^2}$	$h^{1/2}\lvert\hat{\sigma}_h^{\mathrm{nn}}\rvert_{L^2(\mathcal{E}_h^m)}$
368640	1	0	*1.0007*	*0.0000*	*0.9995*	*1.0048*
92160	1	1	1.1860*	0.8325*	0.4992*	0.4680*
92160	1	2	0.9950*	0.4968*	0.4912*	0.4669*
92160	2	1	*2.0009*	*1.0010*	*1.9944*	*2.0575*
92160	2	2	2.9810	1.9271*	1.5055*	1.6780*
23040	2	3	2.5891*	1.5470*	1.5055*	1.9092*
92160	3	2	*2.9996*	*2.0000*	*2.9985*	*2.9837*
23040	3	3	3.9901	2.9617	2.5296*	2.3845*
23040	3	4	3.5682*	2.5217*	2.4828*	2.3819*
23040	4	3	*3.9972*	*2.9985*	*3.9939*	*4.1086*
23040	4	4	4.9893	3.9457	3.5029*	3.8619*
23040	5	4	*4.9989*	*4.0003*	*4.9961*	*4.9544*

7. Final remarks. We have shown that the classic HHJ method can be extended to curved domains using parametric approximation of the geometry to solve the Kirchhoff plate problem on a curved domain. Moreover, optimal convergence rates are achieved so long as the degree of geometry approximation m exceeds the degree of polynomial approximation r by at least 1 (recall that the degree of the Lagrange space is $r + 1$). Smaller values of m generally lead to some deterioration of the convergence rates, although our estimates are not always sharp in this situation.

In particular, we have shown that solving the simply supported plate problem on a curved domain using *polygonal* approximation of the domain and lowest order HHJ elements gives optimal first order convergence, the well-known Babuška paradox notwithstanding [4]. Perhaps surprisingly, if the triangulation is sufficiently fine, the HHJ method computed on the fixed polygon will yield a good approximation of the exact solution of the plate problem on the smooth domain, *not* a good approximation of the exact solution on the polygonal domain. One explanation of the Babuška paradox is that the polygonal approximating domains do not converge to the true domain in the sense of curvature, with curvature being crucial to the simply supported boundary conditions. However, our results show that the HHJ method does not require convergence of the curvatures. In this sense, we might refer to the HHJ method as *geometrically nonconforming*.

REFERENCES

[1] S. Agmon, *Lectures on Elliptic Boundary Value Problems*, Van Nostrand Math. Stud. 2, Van Nostrand, Princeton, Toronto, London, 1965.

[2] D. N. Arnold and Brezzi, F., *Mixed and nonconforming finite element methods: Implementation, postprocessing and error estimates*, RAIRO Modél. Math. Anal. Numér., 19 (1985), pp. 7–32, https://doi.org/10.1051/m2an/1985190100071.

[3] I. Babuška, J. Osborn, and J. Pitkäranta, *Analysis of mixed methods using mesh dependent norms*, Math. Comp., 35 (1980), pp. 1039–1062, http://www.jstor.org/stable/2006374.

[4] I. Babuška and J. Pitkäranta, *The plate paradox for hard and soft simple support*, SIAM J. Math. Anal., 21 (1990), pp. 551–576, https://doi.org/10.1137/0521030.

[5] 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), pp. 556–581, https://doi.org/10.1002/mma.1670020416.

[6] H. Blum and R. Rannacher, *On mixed finite element methods in plate bending analysis. Part 1: The first Herrmann scheme*, Comput. Mech., 6 (1990), pp. 221–236, https://doi.org/10.1007/BF00350239.

[7] D. Boffi, F. Brezzi, and M. Fortin, *Mixed Finite Element Methods and Applications*, Springer Ser. Comput. Math. 44, Springer-Verlag, New York, 2013.

[8] S. C. Brenner, *Discrete Sobolev and Poincaré inequalities for piecewise polynomial functions*, Electron. Trans. Numer. Anal., 18 (2004), pp. 42–48.

[9] F. Brezzi and L. D. Marini, *On the numerical solution of plate bending problems by hybrid methods*, R.A.I.R.O. Analyse Numérique, 9 (1975), pp. 5–50, https://doi.org/10.1051/m2an/197509R300051.

[10] F. Brezzi, L. D. Marini, A. Quarteroni, and P. A. Raviart, *On an equilibrium finite element method for plate bending problems*, Calcolo, 17 (1980), pp. 271–291.

[11] F. Brezzi and P. A. Raviart, *Mixed finite element methods for 4th order elliptic equations*, in Topics in Numerical Analysis III: Proceedings of the Royal Irish Academy Conference on Numerical Analysis, J. J. H. Miller, ed., Academic Press, London, 1976, pp. 33–56.

[12] M. W. Chernuka, G. R. Cowper, G. M. Lindberg, and M. D. Olson, *Finite element analysis of plates with curved edges*, Internat. J. Numer. Methods Engrg., 4 (1972), pp. 49–65, https://doi.org/10.1002/nme.1620040108.

[13] P. Ciarlet and P.-A. Raviart, *Interpolation theory over curved elements, with applications to finite element methods*, Comput. Methods Appl. Mech. Engrg., 1 (1972), pp. 217 – 249, https://doi.org/10.1016/0045-7825(72)90006-0.

[14] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, Classics Appl. Math. 40, SIAM, Philadelphia, 2002, https://doi.org/10.1137/1.9780898719208.

[15] M. I. Comodi, *The Hellan-Herrmann-Johnson method: Some new error estimates and postprocessing*, Math. Comp., 52 (1989), pp. 17–29, http://www.jstor.org/stable/2008650.

[16] W. Krendl, K. Rafetseder, and W. Zulehner, *A decomposition result for biharmonic problems and the Hellan-Herrmann-Johnson method*, Electron. Trans. Numer. Anal., 45 (2016), pp. 257–282.

[17] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, 2nd ed., Course Theoret. Phys. 7, Addison-Wesley, Reading, MA, 1970.

[18] M. Lenoir, *Optimal isoparametric finite elements and error estimates for domains involving curved boundaries*, SIAM J. Numer. Anal., 23 (1986), pp. 562–580, https://doi.org/10.1137/0723036.

[19] L. Mansfield, *Approximation of the boundary in the finite element solution of fourth order problems*, SIAM J. Numer. Anal., 15 (1978), pp. 568–579, https://doi.org/10.1137/0715037.

[20] P. Monk, *A mixed finite element method for the biharmonic equation*, SIAM J. Numer. Anal., 24 (1987), pp. 737–749, https://doi.org/10.1137/0724048.

[21] K. Rafetseder and W. Zulehner, *A decomposition result for Kirchhoff plate bending problems and a new discretization approach*, SIAM J. Numer. Anal., 56 (2018), pp. 1961–1986, https://doi.org/10.1137/17M1118427.

[22] L. R. Scott, *Finite Element Techniques for Curved Boundaries*, Ph.D. thesis, MIT, Cambridge, MA, 1973.

[23] L. R. Scott, *Survey of displacement methods for the plate bending problem*, in US-Germany Symposium on Finite Element Analysis (Cambridge, MA), Brookhaven National Lab., Upton, NY, 1976; available online from https://www.osti.gov/servlets/purl/7268336.

[24] R. Stenberg, *Postprocessing schemes for some mixed finite elements*, RAIRO Modél. Math. Anal. Numér., 25 (1991), pp. 151–167, https://doi.org/10.1051/m2an/1991250101511.

[25] G. Strang and A. E. Berger, *The change in solution due to change in domain*, in Partial

Differential Equations (Proc. Sympos. Pure Math., Vol. XXIII), AMS, Providence, RI, 1971, pp. 199–205.

[26] T. Tiihonen, *Shape calculus and finite element method in smooth domains*, Math. Comp., 70 (2001), pp. 1–15, http://www.jstor.org/stable/2698922.

[27] S. W. Walker, *FELICITY: A MATLAB/C++ toolbox for developing finite element methods and simulation modeling*, SIAM J. Sci. Comput., 40 (2018), pp. C234–C257, https://doi.org/10.1137/17M1128745.

[28] M. Zlámal, *Curved elements in the finite element method.* I, SIAM J. Numer. Anal., 10 (1973), pp. 229–240, https://doi.org/10.1137/0710022.

[29] M. Zlámal, *The finite element method in domains with curved boundaries*, Internat. J. Numer. Methods Engrg., 5 (1973), pp. 367–373, https://doi.org/10.1002/nme.1620050307.

[30] M. Zlámal, *Curved elements in the finite element method.* II, SIAM J. Numer. Anal., 11 (1974), pp. 347–362, https://doi.org/10.1137/0711031.