A DGFEM for Uniformly Elliptic Two Dimensional Oblique boundary-value problems

Ellya Kawecki∗

October 1, 2018

Abstract

In this paper we present and analyse a discontinuous Galerkin finite element method (DGFEM) for the approximation of solutions to elliptic partial differential equations in nondivergence form, with oblique boundary conditions, on curved domains. In “E. Kawecki, A DGFEM for Non-divergence Form Elliptic Equations with Cordes Coefficients on Curved Domains”, the author introduced a DGFEM for the approximation of solutions to elliptic partial differential equations in nondivergence form, with Dirichlet boundary conditions. In this paper, we extend the framework further, allowing for the oblique boundary condition. The method also provides an approximation for the constant occurring in the compatibility condition for the elliptic problems under consideration.

1 Introduction

The model problem that we consider in this paper is the following oblique boundary-value problem, find $u : \Omega \rightarrow \mathbb{R}$ such that

$$\sum_{i,j=1}^{2} A_{ij} D_{ij}^2 u = f, \quad \text{in } \Omega,$$

$$\beta \cdot \nabla u \text{ is constant, on } \partial \Omega,$$

where $\Omega \subset \mathbb{R}^2$ is a given, convex $C^2$ domain, $f \in L^2(\Omega)$, and $A \in L^\infty(\Omega; \mathbb{R}^{2 \times 2}_{\text{Sym}})$ satisfies, for some constant $\lambda > 0$,

$$x^T A(\xi) x \geq \lambda |x|^2 \quad \forall x \in \mathbb{R}^2, \quad \text{for a.e. } \xi \in \Omega.$$

(1.1)

The constant present in the boundary condition of (1.1) is there to absorb potentially arising compatibility conditions (consider solving the Poisson problem with a homogeneous Neumann boundary condition imposed). Finally, we assume that the vector-valued function $\beta \in C^1(\partial \Omega; S^1)$ (hereby called the “oblique vector”).

The oblique boundary-value problem appears in several interesting applications, often dependent upon which dimension, $d$, is considered, and whether or not the oblique boundary-value problem is strict. For $d \geq 2$, and $\beta \in C^1(\partial \Omega; S^d)$, the boundary-value problem is referred to as strictly oblique, if there exists a constant $\delta > 0$ such that

$$\beta \cdot n_{\partial \Omega} \geq \delta \quad \text{on } \partial \Omega,$$

where $n_{\partial \Omega} : \partial \Omega \rightarrow S^1$ is the unit outward normal to $\partial \Omega$. If (1.3) may only hold with $\delta = 0$, then the boundary-value problem is called degenerate oblique.

For $d \geq 3$, (1.3) is necessary for the well-posedness of the boundary-value problem, with (35) (pg 13–14) providing counter examples to uniqueness for the Poisson problem, in the case that the oblique vector becomes tangential to the boundary, even on a set of zero boundary measure.

∗EK acknowledges support of the Engineering and Physical Sciences Research Council [EP/L015811/1].
That said, the degenerate (or tangential) oblique problem (falling into the class of degenerate elliptic problems), arises naturally in the (geodetic) problem of determining the gravitational fields of celestial bodies [32]. This problem was discovered by Poincaré [34] during his work on the theory of tides. In the case that \( d = 2 \), the oblique boundary-value problem arises in systems of conservation laws in [44] [11], where the latter focuses on a mixed elliptic-hyperbolic problem that requires the boundary condition to be strictly oblique. For an overview of the case \( d = 2 \), one should refer to [27], and for the case \( d \geq 3 \), one should seek [28]. A particular, and broad subclass of the oblique boundary-value problem is the case when \( \beta \equiv n_{\partial \Omega} \), which is in fact the Neumann boundary-value problem.

The author’s interest in this type of boundary-value problem, stems from applications to fully nonlinear second order elliptic partial differential equations (PDEs). In particular equations of Monge–Ampère (MA) and Hamilton–Jacobi–Bellman (HJB) type. Upon linearising such equations (for instance by the application of Newton’s method), one arrives at an infinite sequence of problems of the form [14], and as such, the linear theory contained in this paper will be applicable when considering these nonlinear problems. The MA problem arises in areas such as optimal transport and differential geometry, and has been an area of interest, both from an analytical and a numerical computation point of view for many years, see [42] [10] [33] and [8] [31]; while the HJB problem arises in applications to mean field games, engineering, physics, economics, optimal control, and finance [18] [26], where [39] [23] mark recent developments in the numerical analysis of such problems.

It is clear that the linearisation of HJB and MA type equations results in a sequence of nondivergence form elliptic equations. What is not immediately clear, is how the oblique boundary condition may also arise. In the applications outlined above (geodetic problems, and conservation laws), the oblique boundary condition arises, but these problems are not typically cast in nondivergence form.

The nondivergence form oblique boundary-value problem (1.4) arises in the linearisation of the MA optimal transport problem, which can be posed as follows: given two uniformly convex domains, \( \Omega, \Upsilon \subset \mathbb{R}^2 \), and uniformly positive functions \( f_1: \Omega \to \mathbb{R}, f_2: \Upsilon \to \mathbb{R} \), find \( u: \Omega \to \mathbb{R} \) such that

\[
\begin{aligned}
\det D^2 u(x) - \frac{f_1(x)}{f_2(\nabla u(x))} &= 0 \quad x \in \Omega, \\
B_{\Upsilon}(\nabla u(x)) &= 0 \quad x \in \partial \Omega.
\end{aligned}
\]

(1.4)

where \( B_{\Upsilon}: \mathbb{R}^2 \to \mathbb{R} \) is a convex defining function for \( \Upsilon \) (take, for example the signed distance function to \( \Upsilon \)). For simplicity, we assume that \( f_2 \equiv c \) is constant. Upon applying Newton’s method to (1.4), we arrive at the following sequence of inhomogeneous oblique boundary-value problems: given \( u_n \in C^2(\Omega) \), uniformly convex, find \( u_{n+1} \) such that

\[
\begin{aligned}
\text{Cof } D^2 u_n(x) : D^2 u_{n+1} &= \text{Cof } D^2 u_n(x) : D^2 u_n + \frac{f_1(x)}{c} - \det(D^2 u_n) \quad x \in \Omega, \\
DB_{\Upsilon}(\nabla u_n(x)) \cdot \nabla u_{n+1} &= DB_{\Upsilon}(\nabla u_n(x)) \cdot \nabla u_n - B_{\Upsilon}(\nabla u_n(x)) \quad x \in \partial \Omega,
\end{aligned}
\]

(1.5)

where \( \text{Cof} \) denotes the cofactor matrix. Notice that there is no free constant in the boundary condition for (1.5). This is due to the fact that the functions \( f_1, f_2 \) are assumed to satisfy the compatibility condition \( \int_{\Omega} f_1 = \int_{\Upsilon} f_2 \), which is necessary for the existence of a unique (up to a constant) convex solution to the MA optimal transport problem. One can see that the boundary condition in (1.5) is of the (inhomogeneous) oblique type, with \( \beta := DB_{\Upsilon}(\nabla u_n(x))/|DB_{\Upsilon}(\nabla u_n(x))| \). However, when considering the discretisation of (1.5), the gradient of an arbitrary finite element function may take arbitrary values in \( \mathbb{R}^2 \), but, even when \( \partial \Upsilon \) is smooth, the derivative of the defining function, \( B_{\Upsilon} \), may not be well defined everywhere. This can occur when one takes \( B_{\Upsilon} \) to be the signed distance function to \( \partial \Upsilon \) (for example, the signed distance function to the unit disk is not smooth at the origin); one may appeal to a different representation of this function in terms of the supporting hyperplanes of \( \Upsilon \). That is,

\[
B(q) := \sup_{\|n\| = 1} \{q \cdot n - H^*(n)\}, \quad \text{where } H^*(n) := \sup_{q \in \partial \Upsilon} \{q \cdot n\}.
\]

(1.6)

See [5] for further justification of this representation. In [5], the authors use this representation of the boundary condition (with further modifications, so that the supremum in (1.6) is taken over
vectors that make an acute angle with the unit normal to \( \partial \Omega \), and iteratively solve (via a finite difference method) the MA problem with Neumann boundary conditions. Utilising (1.6), we arrive at the following MA problem:

\[
\begin{aligned}
\det D^2 u(x) - \frac{f_1(x)}{c} &= 0 \quad x \in \Omega, \\
\sup_{\|n\|=1} \{ \nabla u \cdot n - H^*(n) \} &= 0 \quad x \in \partial \Omega.
\end{aligned}
\] (1.7)

This brings us to the intersection of MA and HJB problems. Taking into account the representation (1.6), one can see that the application of Newton’s method in (1.5) is not justified, in general, since we may not be able to make sense of the derivative of \( B_\Upsilon \). However, a semismooth Newton’s method may be justified, which relies only on the subdifferential of the operator under consideration. This type of linearisation scheme is implemented in [39] in order to design a convergent scheme for the DG approximation of HJB equations with Dirichlet boundary conditions. To move into the context of HJB equations, one may utilise a longstanding result proven by Krylov [25], that allows one to characterise the MA problem (1.7), as the following HJB problem

\[
\begin{aligned}
\sup_{W \in X} \left\{ -W : D^2 u + 2(\det W)^{1/2} \frac{f_1^{1/2}(x)}{c^{1/2}} \right\} &= 0 \quad \text{in } \Omega, \\
\sup_{\|n\|=1} \{ \nabla u \cdot n - H^*(n) \} &= 0 \quad \text{on } \partial \Omega,
\end{aligned}
\] (1.8)

where \( X := \{ W \in \mathbb{R}^{2 \times 2} : W = W^T, W \geq 0, \text{Trace}(W) = 1 \} \). Applying a semismooth Newton’s method to the operator in the domain and on the boundary, we arrive at a sequence of nondivergence form oblique boundary-value problems.

Exporting the ideas of this paper to the context of (1.8) will be the focus of future work. This will involve recasting the HJB problem in terms of one with uniformly elliptic coefficients (the set \( X \) contains degenerate matrices), and, furthermore, a proof of convergence of the semismooth Newton’s method with a Hamilton–Jacobi type boundary condition. A key motivation for Krylov’s HJB formulation of the MA equation, is that the numerical solution may in fact be unique, which is not always the case when considering numerical methods for the MA equation (see [17], for example).

Another motivation for considering HJB type equation with oblique boundary conditions arises in the area of mean field games, in the modelling of pedestrian crowds (the elliptic counterpart of this problem is to determine the long term behaviour of the crowd). In this case the HJB equation is coupled with a Fokker–Planck equation (see [26]), and the boundary condition models the interaction the crowd will have with the boundary of the space that they inhabit. In the case of oblique boundary conditions, one can think of this interaction as the requirement that the direction in which one is incentivised to exit the domain is determined by the oblique vector.

Now that we have discussed why one is motivated to approximate solutions problems of the form (1.1), it is pertinent to discuss why standard approaches may not apply in this case. The problem (1.1) poses several difficulties, both analytically and numerically. The problem is in nondivergence form (and due to our assumptions, it cannot be written in divergence form, in general, and so the standard weak formulation cannot be used here), and as such, well-posedness is not guaranteed (see [21, 30, 37]), and the use of standard conforming finite element methods is also not applicable (as they also rely on the weak formulation of the PDE).

As we will see, on \( C^2 \) domains, one can show well-posedness by assuming that the coefficients satisfy the “Cordes condition” (defined in Section 3), and a condition between the oblique vector, and the curvature of the boundary. In [30], well-posedness is proven using the method of continuity; in this paper we use a different method, analogous to that seen in [35] Theorem 3. This technique relies upon a variant of the Miranda–Talenti estimate (see Lemma 3.11). The motivation for the use of a different technique is the extension of this technique to problems of HJB type, for which the method present in [30] does not apply.

The Miranda–Talenti estimate is known to hold for \( d \geq 2 \) when considering functions in a suitable Sobolev space, whose trace or normal derivative vanishes on \( \partial \Omega \). However, for \( d \geq 3 \), this estimate
remains an open problem if one assumes that the oblique derivative $(\beta \cdot \nabla u)$ is constant on the boundary \([14]\), restricting us to a two-dimensional framework.

In terms of PDE analysis, our goal is to prove existence of a unique strong solution \((H^2\text{-regular, so that the PDE is satisfied a.e.})\) \(u \in H\) satisfying

\[
A_\gamma(u, v) := \int_\Omega \gamma A : D^2u \Delta v = \int_\Omega \gamma f \Delta v =: \ell(v) \quad \forall v \in H,
\]

where \(H\) is a particular subset of \(H^2(\Omega)\) (see Section 2), and \(\gamma\) is a uniformly positive renormalisation factor defined in Section 3. This motivates the construction of the finite element method given in Section 4, with the goal of finding a suitable numerical analogue of the operator \(A_\gamma\).

This is another example where the techniques present in this paper divert from the standard concepts of conforming finite element methods. A direct discretisation of the operator \(A_\gamma\) in a finite element setting (by replacing the derivatives present in \(A_\gamma\) with piecewise derivatives) would fail, since general finite element functions do not satisfy the Miranda–Talenti estimate, and thus the corresponding operator would, in general, fail to be coercive. The alternative would be to construct a finite element space whose members are both \(H^2\)-regular, and satisfy the oblique boundary condition. In particular, it is not immediately clear how one would do the latter. We instead design a discontinuous Galerkin scheme that numerically enforces the Miranda–Talenti estimates (via internal and boundary jump penalisation terms), and is also consistent (i.e., that the true solution to the PDE is also a solution to the finite element method, assuming the solution has sufficient broken-Sobolev regularity, allowing one to substitute it into the formulation), resulting in optimal a priori error estimates.

It is important to consider the expected regularity of solutions to \((1.1)\). In general, since the coefficient matrix, \(A\) is only \(L^\infty(\Omega)\), and the right-hand side function \(f \in L^2(\Omega)\), standard elliptic regularity theory implies that a solution would, in general, be at most \(H^2\)-regular. In Section 5, we provide an a priori error estimate for solutions of such regularity. In the case that solutions are smoother, we are able to prove a priori error estimates in a \(H^2\)-style norm, that are optimal with respect to the polynomial degree of the finite element space. Even though the linearised problems may admit solutions of lower regularity, the limiting problems (by way of Newton’s method) may indeed be of higher regularity. For example, in the case that \(\partial \Omega, \partial \Upsilon\) are \(C^{2,1}\) regular, the MA optimal transport problem \((1.4)\) admits a \(C^{3,\alpha}(\Omega) \cap C^{2,\alpha}(\Upsilon)\) convex solution (see [41]).

Interestingly, our method also gives an approximation of the constant that arises in the compatibility condition for problems of the form \((1.1)\). In the case of conormal boundary-value problems, i.e.,

\[
\begin{aligned}
-\nabla \cdot (A \nabla u) &= f, \quad \text{in } \Omega, \\
(A \nabla u) \cdot n_{\partial \Omega} &= \text{constant on } \partial \Omega,
\end{aligned}
\]

it is quite straightforward (by an application of the Divergence Theorem) to determine the constant in terms of the function \(f\). In the problems we consider, however, there does not appear to be any explicit relationship between the constant and the right hand side, in general.

Our approach extends the framework of [55] and [24] (the first of which applies to the Dirichlet boundary condition on polytopal domains, and the second applies to the Dirichlet boundary condition on curved, uniformly convex domains) to the oblique case. To the author’s knowledge, there is a sparse amount of work on finite element methods for oblique boundary-value problems present in the existing literature, and, as such, this paper provides the first DGFEM for the approximation of solutions to \((1.1)\). As such, a motivation of this paper is to widen the scope of the current numerical framework for oblique boundary-value problems.

The papers [13] [16] [4] provide examples of finite element approximations of oblique boundary-value problems, where [16] and [4] apply to a particular geodetic and free boundary problem respectively. Close to the timing of this paper, the author of [19] introduced a mixed finite element method for the approximation of solutions to \((1.1)\), and proved a priori and a posteriori error estimates. The hypotheses of [19] with respect to the data for problem \((1.1)\) are identical to that of this paper (i.e., the assumptions upon \(A, \partial \Omega, f, \text{ and } \beta\)), however, the consideration of curved finite elements provides
a difference in our computationally sufficient assumptions. This stems from the fact that the finite element functions that we consider are polynomials on the flat reference simplex, whereas the finite element functions in [19] are polynomials on the curved physical element. In particular, computationally, [19] assumes that $\partial \Omega \in C^2$, whereas we require that $\partial \Omega$ is also piecewise $C^3$. However, in [19] the function approximating $\nabla u$ must be piecewise affine, whereas we are able to employ piecewise polynomials of degree $p \geq 2$ (as we are approximating $u$, one may consider piecewise quadratics in our case to be the analogue of piecewise affine in the case of [19]).

This paper is organised as follows: In Section 2 we begin by introducing the notation needed, as well as the function spaces used in the analysis of the PDE. In Section 3 we prove several important estimates, such as the Miranda–Talenti estimate, and then proceed to prove an existence and uniqueness result for problems of the form (1.1). In Section 4 we prove an important consistency result, and proceed to prove a stability result for our numerical scheme; this stability result is then used as a main tool in the proof of existence and uniqueness of a numerical solution. In Section 5 we prove an error estimate that is optimal in terms of the mesh size, for piecewise sufficiently smooth solutions, as well as an estimate for solutions of conformal regularity. In Section 6 we perform several numerical experiments, validating the estimates of Section 5. Section 7 is the final section, where we give concluding remarks on what has been accomplished in this paper.

2 Set-up

Before we prove existence and uniqueness of solutions to (1.1), we must first determine a suitable function space in which to do so. In general, we will consider strong solutions of (1.1). That is, where all weak derivatives of the solution, up to second order, are square integrable, so that (1.1) may hold a.e. in $\Omega$.

2.1 Function spaces

We consider the standard Sobolev space
$$H^2(\Omega) := \{ v \in L^2(\Omega) : D^\alpha v \in L^2(\Omega), \forall 1 \leq |\alpha| \leq 2 \}.$$ We define the following subset (determined by the oblique vector $\beta$) of $H^2(\Omega)$:
$$H^2_\beta(\Omega) := \bigcup_{c \in \mathbb{R}} \{ v \in H^2(\Omega) : \beta \cdot \nabla v \text{ is constant on } \partial \Omega \},$$
as well as a further subspace
$$H^2_{\beta,0}(\Omega) := \left\{ v \in H^2_\beta(\Omega) : \int_{\Omega} v = 0 \right\}.$$ Furthermore, we will endow all three spaces with the following norm:
$$\|u\|_{H^2(\Omega)}^2 := \sum_{|\alpha| \leq 2} \|D^\alpha u\|_{L^2(\Omega)}^2.$$ We will seek to prove existence and uniqueness of solutions to (1.1) in $H^2_{\beta,0}(\Omega)$.

2.2 Tangential differential operators

As mentioned previously, we denote by $n_{\partial \Omega} = (n_1, n_2)^T : \partial \Omega \rightarrow \mathbb{S}^1$, the unit outward normal to $\partial \Omega$. We define the unit tangent vector $T : \partial \Omega \rightarrow \mathbb{S}^1$ as follows,
$$T := (-|n_{\partial \Omega}|_2, |n_{\partial \Omega}|_1)^T.$$
For a given vector-valued function $v$ we denote:

$$v^\perp := (-v_2, v_1)^T.$$

**Definition 2.1 (Tangential gradient)** We define the tangential gradient, $\nabla_T : H^s(\partial \Omega) \to H_T^{s-1}(\partial \Omega)$, as follows:

$$\nabla_T v = \nabla v - \frac{\partial v}{\partial n_{\partial \Omega}} n_{\partial \Omega},$$

where $\frac{\partial v}{\partial n_{\partial \Omega}} = \nabla v \cdot n_{\partial \Omega}$, and $H_T^s(\partial \Omega) := \{ v \in H^s(\partial \Omega) \times H^s(\partial \Omega) : v \cdot n_{\partial \Omega} = 0 \}$.

**Definition 2.2 (Directional derivative)** For a sufficiently smooth function $v$, and a vector $s$, we define the directional derivative of $u$ with respect to $s$ as follows:

$$\partial_s u := \nabla u \cdot s.$$

**Definition 2.3 (Tangential Laplacian)** We define the tangential Laplacian, $\Delta_T : H^s(\partial \Omega) \to H^{s-2}(\partial \Omega)$, as follows:

$$\Delta_T := \text{div}_T \nabla_T = \nabla_T \cdot \nabla_T.$$

**Definition 2.4 (Mean curvature)** We define the mean curvature, $H_{\partial \Omega} : \partial \Omega \to \mathbb{R}$, of $\partial \Omega$ as follows:

$$H_{\partial \Omega} := \nabla_T \cdot n_{\partial \Omega}.$$

### 3 PDE analysis

We begin this section by defining the “Cordes condition”, which is crucial to the technique we will use to analyse problems in the form of (1.1); remarkably, the condition is a consequence of uniform ellipticity in two dimensions.

**Remark 3.1 (Minimal domain and oblique vector regularity)** We will always assume that the domain, $\Omega$, has a $C^2$ boundary, and that the oblique vector, $\beta \in C^1(\partial \Omega; \mathbb{S}^1)$.

**Definition 3.2 (Cordes condition)** Let $L := \sum_{i,j=1}^2 A_{ij} D_{ij}^2$. The operator $L$ satisfies the Cordes condition if there exists $\varepsilon \in (0, 1)$ such that

$$\frac{|A|^2}{(\text{Tr}(A))^2} \leq \frac{1}{1 + \varepsilon} \quad \text{a.e. in } \Omega. \quad (3.1)$$

**Lemma 3.3** Assume that $A$ is uniformly elliptic. Then, $A$ satisfies the Cordes condition (3.1).

**Proof:** Denote the lower and upper ellipticity constants of $A$ by $\lambda$ and $\Lambda$, respectively. Then,

$$\frac{|A|^2}{(\text{Tr}(A))^2} = \frac{\lambda^2 + \Lambda^2}{(\lambda + \Lambda)^2} = \frac{1}{1 + 2\lambda\Lambda/(\lambda + \Lambda)^2}.$$

Choosing $\varepsilon \in (0, 2\lambda\Lambda/(\lambda + \Lambda)^2)$ yields the claim. $\Box$

**Lemma 3.4** Assume that $L$ satisfies the Cordes condition (3.1). Then, defining

$$\gamma := \frac{\text{Tr} A}{|A|^2},$$

we have that

$$|\gamma L v - \Delta v| \leq \sqrt{1 - \varepsilon} |D^2 v|,$$

for any $v \in H^2(\Omega)$. 

Proof: See [38], Lemma 1.

**Definition 3.5 (Oblique angle)** We define the “oblique angle”, $\Theta : \partial \Omega \to \mathbb{R}$, to be the (anticlockwise) oriented angle between the oblique vector, $\beta$, and the unit outward normal, $n$.

The following lemma will allow us to prove the Miranda–Talenti estimate, which can be found in [30]. It also generalises the techniques present in [30] to the following bilinear form

$$\mathcal{B}(u, v) := \int_\Omega D_{11}^2 u D_{22}^2 v + D_{22}^2 u D_{11}^2 v - 2D_{12}^2 u D_{12}^2 v,$$

which will be used to prove an important consistency result in Section 4.

**Lemma 3.6** Assume that $E \subset \mathbb{R}^2$ is a bounded, Lipschitz, piecewise $C^2$ domain, and that $\beta \in C^1(\Gamma_n; \mathbb{S}^1)$ for each $C^2$ portion $\Gamma_n$ of $\partial E$, $n = 1, \ldots, N$, $N \in \mathbb{N}$. Then, for any $u, v \in H^s(E)$, $s > 5/2$, we have that

$$\int_E D_{11}^2 u D_{22}^2 v + D_{22}^2 u D_{11}^2 v - 2D_{12}^2 u D_{12}^2 v = \int_{\partial E} \left( \beta_1 \partial_T \beta_2 - \beta_2 \partial_T \beta_1 \right) (\beta^\perp \cdot \nabla u \beta^\perp \cdot \nabla v + \beta \cdot \nabla u \beta \cdot \nabla v) + \int_{\partial E} \left( \partial_T (\beta^\perp \cdot \nabla u) \beta \cdot \nabla v - \partial_T (\beta \cdot \nabla u) \beta^\perp \cdot \nabla v \right). \quad (3.2)$$

**Proof:** Let us momentarily assume that $u, v \in C^3(\overline{E})$, and note that an application of integration by parts gives us

$$\int_E D_{11}^2 u D_{22}^2 v + D_{22}^2 u D_{11}^2 v - 2D_{12}^2 u D_{12}^2 v = \int_E D_1 (D_1 v D_{22}^2 u - D_2 v D_{11}^2 u) - D_2 (D_1 v D_{22}^2 u - D_2 v D_{11}^2 u) \quad (3.3)$$

$$= \int_{\partial E} (D_1 v D_{22}^2 u - D_2 v D_{11}^2 u) n_1 - (D_1 v D_{22}^2 u - D_2 v D_{11}^2 u) n_2.$$

Now, denoting $C_1^1 := \beta^\perp \cdot \nabla u$, $C_1^2 := \beta^\perp \cdot \nabla v$, $C_1^3 := \beta \cdot \nabla u$, and $C_1^4 := \beta \cdot \nabla v$, we obtain the following linear systems on $\Gamma_n$, $n = 1, \ldots, N$:

$$\begin{cases} \beta_1 D_1 u + \beta_2 D_2 u = C_1^2, \\ -\beta_2 D_1 u + \beta_1 D_2 u = C_1^3, \end{cases} \quad (3.4)$$

$$\begin{cases} \beta_1 D_1 v + \beta_2 D_2 v = C_1^4, \\ -\beta_2 D_1 v + \beta_1 D_2 v = C_1^1, \end{cases} \quad (3.5)$$

with corresponding unique solutions

$$\begin{cases} D_1 u = -C_1^4 \beta_2 + C_1^2 \beta_1, \\ D_2 u = C_1^3 \beta_1 + C_1^2 \beta_2, \\ D_1 v = -C_1^4 \beta_2 + C_1^2 \beta_1, \\ D_2 v = C_1^1 \beta_1 + C_1^4 \beta_2. \end{cases} \quad (3.6)$$
Substituting in the values of $D_1v$ and $D_2v$ present in (3.6) into (3.3), (denoting $n_i = [u_{0i}]^t$, $i = 1, 2$) we obtain the following:

\[
\int_E D_1^2 u D_2^2 v + D_2^2 u D_1^2 v - 2 D_2^2 u D_2^2 v
\]
\[= \int_{\partial E} -C_v^1 \beta_2 D_2^2 u n_1 + C_v^2 \beta_1 D_2^2 u n_1 - C_v^1 \beta_1 D_2^2 u n_1 - C_v^2 \beta_2 D_2^2 u n_1
\]
\[+ \int_{\partial E} C_v^1 \beta_2 D_2^2 u n_2 - C_v^2 \beta_1 D_2^2 u n_2 + C_v^1 \beta_1 D_2^2 u n_2 + C_v^2 \beta_2 D_2^2 u n_2
\]
\[= \int_{\partial E} C_v^1 (-\beta_2 D_2^2 u n_1 - \beta_1 D_2^2 u n_1 + \beta_2 D_2^2 u n_2 + \beta_1 D_2^2 u n_1)
\]
\[+ \int_{\partial E} C_v^2 (\beta_1 D_2^2 u n_1 - \beta_2 D_2^2 u n_1 - \beta_1 D_2^2 u n_2 + \beta_2 D_2^2 u n_2).
\]

Taking the directional derivative of the equations in (3.3) with respect to $\mathbf{T}$ yields (denoting $\cdot := \partial_T$)

\[
\begin{align*}
\beta_1 D_1 u + \beta_1 D_1^2 u (-n_2) + \beta_1 D_1^2 u n_1 + \beta_2 D_2 u + \beta_2 D_2^2 u (-n_2) + \beta_2 D_2^2 u n_1 &= C_u^2, \\
-\beta_2 D_1 u - \beta_2 D_1^2 u (-n_2) - \beta_2 D_1^2 u n_1 + \beta_1 D_2 u + \beta_1 D_2^2 u (-n_2) + \beta_1 D_2^2 u n_1 &= C_u^1,
\end{align*}
\]

thus,

\[
\begin{align*}
-(\beta_1 D_1^2 u n_2 - \beta_1 D_1^2 u n_1 + \beta_2 D_2^2 u n_2 - \beta_2 D_2^2 u n_1) &= C_u^2 - (\beta_1 D_1 u + \beta_2 D_2 u), \\
(\beta_1 D_2^2 u n_1 - \beta_2 D_2^2 u n_1 - \beta_1 D_2^2 u n_2 + \beta_2 D_2^2 u n_2) &= C_u^1 + \beta_2 D_1 u - \beta_1 D_2 u.
\end{align*}
\]

Substituting the values for $D_1u$ and $D_2u$ present in (3.6) gives us

\[
\begin{align*}
\beta_1 \hat{\beta}_1 + \beta_2 \hat{\beta}_2 &= 1/2 \partial_T(\beta^2) = 1/2 \partial_T(1) = 0.
\end{align*}
\]

Similarly, we obtain

\[
\hat{\beta}_2 D_1 u - \hat{\beta}_1 D_2 u = C_u^2 (\beta_1 \hat{\beta}_2 - \beta_2 \hat{\beta}_1).
\]

Substituting (3.9) and (3.10) into (3.8), and then substituting the resulting equations into (3.7), we obtain

\[
\int_E D_1^2 u D_2^2 v + D_2^2 u D_1^2 v - 2 D_2^2 u D_2^2 v
\]
\[+ \int_{\partial E} C_v^1 (C_v^1 (\beta_2 \hat{\beta}_2 - \beta_2 \hat{\beta}_1) - C_v^1) + C_v^2 (C_v^2 (\beta_1 \hat{\beta}_2 - \beta_2 \hat{\beta}_1) + C_v^2)
\]
\[= \int_{\partial E} (\hat{\beta}_1 \beta_1 \beta_2 - \hat{\beta}_2 \beta_1) (\beta^2) + \beta \cdot \nabla u \beta \cdot \nabla v + \beta \cdot \nabla u \beta \cdot \nabla v
\]
\[+ \int_{\partial E} (\hat{\beta}_1 \beta_1 \beta_2 - \hat{\beta}_2 \beta_1) (\beta^2) + \beta \cdot \nabla u \beta \cdot \nabla v - \partial_T (\beta \cdot \nabla u \beta^2) + \beta \cdot \nabla v),
\]

which is exactly (3.2). This identity extends to $u, v \in H^s(E)$, $s > 5/2$, by density, which concludes the proof. □
Corollary 3.7 Assume that $\Omega \subset \mathbb{R}^2$ is a bounded $C^2$ domain, and that $\beta \in C^1(\partial \Omega; S^1)$. Then, for all $u, v \in H^s(\Omega)$, $s > 5/2$, we have that

$$
\int_{\Omega} \frac{D^2 u}{D^2 v} \frac{D^2 v}{D^2 u} - 2D^2 u D^2 v = \int_{\partial \Omega} (\partial \Theta + H_{\partial \Omega}) (\beta^\perp \cdot \nabla u \beta^\perp \cdot \nabla v + \beta \cdot \nabla u \beta \cdot \nabla v)
+ \int_{\partial \Omega} (\partial \Theta (\beta^\perp \cdot \nabla u \beta \cdot \nabla v - \partial \Theta (\beta \cdot \nabla u) \beta^\perp \cdot \nabla v).
$$

(3.11)

Proof: This follows from applying Lemma 3.6 and applying the following equality

$$
\beta_1 \partial \Theta \beta_2 - \beta_2 \partial \Theta \beta_1 = \partial \Theta + H_{\partial \Omega} \text{ on } \partial \Omega.
$$

Note that the above equality follows from identities (1.83) and (1.84) on page 48 in [30] (note that to pass from the present notation, to the notation found in [30], one must denote $\ell := \beta$, $\theta := \Theta$, and $\chi := -H_{\partial \Omega}$.) □

Corollary 3.8 Assume that $E \subset \mathbb{R}^2$ is a bounded, Lipschitz, piecewise $C^2$ domain, and that $\beta \in C^1(\Gamma_n; S^1)$ for each $C^2$ portion $\Gamma_n$ of $\partial E$, $n = 1, \ldots, N$, $N \in \mathbb{N}$. Then, for any $u, v \in H^s(E)$, $s > 5/2$, we have that

$$
\int_E D^2 u \cdot D^2 v + \int_{\partial E} (\beta_1 \partial \Theta \beta_2 - \beta_2 \partial \Theta \beta_1) (\beta^\perp \cdot \nabla u \beta^\perp \cdot \nabla v + \beta \cdot \nabla u \beta \cdot \nabla v)
+ \int_{\partial E} (\partial \Theta (\beta^\perp \cdot \nabla u \beta \cdot \nabla v - \partial \Theta (\beta \cdot \nabla u) \beta^\perp \cdot \nabla v)
= \int_E \Delta u \Delta v.
$$

Proof: First note that for $u, v \in H^2(E)$,

$$
\int_E D^2 u \cdot D^2 v + \int_{\partial E} (\beta_1 \partial \Theta \beta_2 - \beta_2 \partial \Theta \beta_1) (\beta^\perp \cdot \nabla u \beta^\perp \cdot \nabla v + \beta \cdot \nabla u \beta \cdot \nabla v)
+ \int_{\partial E} (\partial \Theta (\beta^\perp \cdot \nabla u \beta \cdot \nabla v - \partial \Theta (\beta \cdot \nabla u) \beta^\perp \cdot \nabla v)
= \int_E \Delta u \Delta v,
$$

and apply Lemma 3.6 □

Lemma 3.9 (Miranda–Talenti estimate) Let $\Omega \subset \mathbb{R}^2$ be a $C^2$ domain, and assume that $\beta \in C^1(\partial \Omega; S^1)$. Furthermore, assume that

$$
\partial \Theta + H_{\partial \Omega} \geq 0 \text{ on } \partial \Omega.
$$

Then, we have that

$$
|u|_{H^2(\Omega)} := \int_{\Omega} \sum_{i,j=1}^2 (D^2 u)^2 \leq \int_{\Omega} (\Delta u)^2 = \|\Delta u\|_{L^2(\Omega)},
$$

(3.13)

for all $u \in H^2(\Omega)$.

Proof: First, we note that since $\Omega \subset \mathbb{R}^2$ is a $C^2$ domain, it is both Lipschitz continuous and piecewise $C^2$. Furthermore, as $\beta \in C^1(\partial \Omega; S^1)$, it follows that $\beta \in C^1(\Gamma_n; S^1)$ for each $C^2$ portion $\Gamma_n$ of $\partial \Omega$; indeed one may take $\Gamma_n = \partial \Omega$. Now, let us assume that $u \in C^2(\Omega) \cap H^2_{\beta,0}(\Omega)$, so that $u \in H^s(\Omega)$, with $s > 5/2$. Setting $v = u$, it then follows that the hypotheses of Corollaries 3.7 and 3.8 are satisfied. It then follows that

$$
\int_{\Omega} D^2 u : D^2 u + \int_{\partial \Omega} (\partial \Theta + H_{\partial \Omega}) (\beta^\perp \cdot \nabla u \beta^\perp \cdot \nabla v + \beta \cdot \nabla u \beta \cdot \nabla v)
+ \int_{\partial \Omega} (\partial \Theta (\beta^\perp \cdot \nabla u \beta \cdot \nabla v - \partial \Theta (\beta \cdot \nabla u) \beta^\perp \cdot \nabla v)
= \int_{\Omega} \Delta u \Delta u.
$$
Since $\beta \cdot \nabla u|_{\partial \Omega} = C$ for some constant $C$, it follows that
$$\partial_T(\beta \cdot \nabla u)|_{\partial \Omega} = 0.$$ Furthermore, we see that
$$\beta^\perp \cdot \nabla u \beta^\perp \cdot \nabla u + \beta \cdot \nabla u \beta \cdot \nabla u = |\nabla u|^2.$$ Thus, we obtain
$$\int_{\Omega} |D^2 u|^2 + \int_{\partial \Omega} (\partial_T \Theta + \mathcal{H}_{\partial \Omega}) |\nabla u|^2 + C \partial_T(\beta^\perp \cdot \nabla u) = \int_\Omega (\Delta u)^2.$$ Since $\partial \Omega$ is a compact hyper-surface, an application of integration by parts yields
$$\int_{\partial \Omega} C \partial_T(\beta^\perp \cdot \nabla u) = 0.$$ Finally, since $\partial_T \Theta + \mathcal{H}_{\partial \Omega} \geq 0$ on $\partial \Omega$, we obtain
$$|u|^2_{H^2(\Omega)} := \int_{\Omega} \sum_{i,j=1}^2 (D^2_{ij} u)^2 = \int_\Omega |D^2 u|^2 \leq \int_\Omega (\Delta u)^2 = \|\Delta u\|^2_{L^2(\Omega)},$$ as desired. □

**Lemma 3.10 (Gradient estimate)** Under the assumptions of Lemma 3.9 with the additional assumption that
$$\partial_T \Theta + \mathcal{H}_{\partial \Omega} > 0 \text{ on } \partial \Omega,$$ we have that
$$\int_{\Omega} |\nabla u|^2 \leq C \int_{\Omega} (\Delta u)^2,$$ for all $u \in H^2(\Omega)$, where $C$ is a constant independent of $u$.

**Proof:** See [30], Lemma 1.5.8. □

**Corollary 3.11** Under the assumptions of Lemma 3.10 we have that
$$\|u\|^2_{H^2(\Omega)} \leq C \|\Delta u\|^2_{L^2(\Omega)}, \quad (3.14)$$ for all $u \in H^2_{\beta,0}(\Omega)$, where the constant $C$ is independent of $u$.

**Proof:** Applying Lemmas 3.9 and 3.10 we obtain
$$\|u\|^2_{H^2(\Omega)} = \|u\|^2_{L^2(\Omega)} + \|\nabla u\|^2_{L^2(\Omega)} + \|D^2 u\|^2_{L^2(\Omega)} \leq \|u\|^2_{L^2(\Omega)} + (C + 1)\|\Delta u\|^2_{L^2(\Omega)}. \quad (3.15)$$
Now, by Poincaré’s inequality, we obtain
$$\|u\|^2_{L^2(\Omega)} = \left\| u - \frac{1}{|\Omega|} \int_\Omega u \right\|^2_{L^2(\Omega)} \leq C \int_{\Omega} |\nabla u|^2 \leq C\|\Delta u\|^2_{L^2(\Omega)},$$ combining this with (3.15), we obtain the desired result. □
Theorem 3.12 Under the assumptions of Lemma 3.10, there exists a unique \( u \in H^2_{\beta,0}(\Omega) \) that is a strong solution of (1.1).

Proof: Define \( A_\gamma : H^2_{\beta,0}(\Omega) \times H^2_{\beta,0}(\Omega) \to \mathbb{R} \), and \( l_\gamma : H^2_{\beta,0}(\Omega) \to \mathbb{R} \) as follows:

\[
A_\gamma(u, v) := \int_\Omega \gamma A : D^2u \Delta v \quad \forall u, v \in H^2_{\beta,0}(\Omega),
\]

and

\[
l_\gamma(v) := \int_\Omega \gamma f \Delta v \quad \forall v \in H^2_{\beta,0}(\Omega).
\]

We will first prove that there exists a unique \( u \in H^2_{\beta,0}(\Omega) \) such that

\[
A_\gamma(u, v) = l_\gamma(v) \quad \forall v \in H^2_{\beta,0}(\Omega).
\]

We see that

\[
A_\gamma(u, v) \leq \| \gamma \|_{L^\infty(\Omega)} \| A \|_{L^\infty(\Omega)} \| u \|_{H^2(\Omega)} \| \Delta v \|_{L^2(\Omega)} \leq C \| u \|_{H^2(\Omega)} \| v \|_{H^2(\Omega)},
\]

similarly

\[
l_\gamma(v) \leq C \| v \|_{H^2(\Omega)}.
\]

Furthermore,

\[
A_\gamma(u, u) = \int_\Omega (\gamma Lu - \Delta u) \Delta u + \int_\Omega (\Delta u)^2 \geq -\sqrt{1 - \epsilon} \int_\Omega |D^2u| \Delta u + \int_\Omega (\Delta u)^2 \geq -\sqrt{1 - \epsilon} \| D^2u \|_{L^2(\Omega)} \| \Delta u \|_{L^2(\Omega)} + \| \Delta u \|_{L^2(\Omega)}^2 \geq (1 - \sqrt{1 - \epsilon}) \| \Delta u \|_{L^2(\Omega)}^2 \geq C \| u \|_{H^2(\Omega)}^2.
\]

Note that the second inequality follows from Lemma 3.4 and the final inequalities follow from Lemma 3.9 and Corollary 3.11.

We have shown that \( A_\gamma \) is both coercive and bounded, and \( l_\gamma \) is bounded. In view of the fact that \( H^2_{\beta,0}(\Omega) \) is a Hilbert space when endowed with the inner product

\[
\langle u, v \rangle_\Delta := \int_\Omega \Delta u \Delta v \quad \forall u, v \in H^2_{\beta,0}(\Omega),
\]

the Lax Milgram Theorem ([15] Section 6.2.1) yields existence and uniqueness of \( u \in H^2_{\beta,0}(\Omega) \) such that

\[
A_\gamma(u, v) = l_\gamma(v) \quad \forall v \in H^2_{\beta,0}(\Omega),
\]

i.e.,

\[
\int_\Omega \gamma A : D^2u \Delta v = \int_\Omega \gamma f \Delta v \quad \forall v \in H^2_{\beta,0}(\Omega).
\]

Now, it follows from [30] Page 56, that \( \Delta : H^2_{\beta,0}(\Omega) \to L^2(\Omega) \) is a surjection, and thus we obtain

\[
\int_\Omega \gamma (A : D^2u - f) v = 0 \quad \forall v \in L^2(\Omega).
\]

An application of the Fundamental Lemma of the calculus of variations yields

\[
\gamma (A : D^2u - f) = 0 \quad \text{a.e. in } \Omega,
\]

and since \( \gamma \) is uniformly positive, we obtain

\[
A : D^2u = f \quad \text{a.e. in } \Omega. \]
4 Numerical method

4.1 Notation

Definition 4.1 (Edge and vertex sets) Given a triangulation $\mathcal{T}_h$, we denote by $E_h^b$, the set of boundary edges of $\mathcal{T}_h$, by $E_h^i$ the set of interior edges of $\mathcal{T}_h$, by $E_h^{i,b} := E_h^i \cup E_h^b$, and by $V_h^b$ the set of boundary vertices of $\mathcal{T}_h$.

Definition 4.2 (Piecewise $C^k$ domain) A domain $\Omega \subset \mathbb{R}^d$ is piecewise $C^k$ for $k \in \mathbb{N}$, if we may express the boundary of $\Omega$, $\partial \Omega$, as a finite union

$$\partial \Omega = \bigcup_{n=1}^{N} \Gamma_n,$$  \hspace{1cm} (4.1)

where each $\Gamma_n \subset \mathbb{R}^d$ is of zero $d$-dimensional Lebesgue measure, and admits a local representation as the graph of a uniformly $C^k$ function. That is, for each $n$, and at each $x \in \Gamma_n$ there exists an open neighbourhood $V_n$ of $x$ in $\mathbb{R}^d$ and an orthogonal coordinate system $(y_1, \ldots, y_d)$, such that

$$V_n = \{(y_1, \ldots, y_d) : -a_j < y_j < a_j, 1 \leq j \leq d \};$$

as well as a uniformly $C^k$ function $\phi_n$ defined on $V_n = \{(y_1, \ldots, y_{d-1}) : -a_j < y_j < a_j, 1 \leq j \leq d-1 \}$ and such that

$$|\phi_n(y)| \leq a_n^d / 2 \text{ for every } y \in V_n,$$

$$\Omega \cap V = \{y^n = (y^n, y_d) \in V : y_d < \phi_n(y^n)\},$$

$$\Gamma_n \cap V = \{y^n = (y^n, y_d) \in V : y_d = \phi_n(y^n)\}.$$  

Jump and average operators. For each face $F = \overline{K} \cap \overline{K'}$ for some $K, K' \in \mathcal{T}_h$ (in the case that $F \in E_h^b$ take $F = \partial K \cap \partial \Omega$), with corresponding unit normal vector $n_F$ (which, for convention is chosen so that it is the outward normal to $K$), we define the jump operator, $[\cdot]$ over $F$, by

$$[\phi] = \begin{cases} 
(\phi|_K)|_F - (\phi|_{K'})|_F & \text{if } F \in E_h^i \\
(\phi|_K)|_F & \text{if } F \in E_h^b,
\end{cases}$$

and the average operator, $\langle \cdot \rangle$, by

$$\langle \phi \rangle = \begin{cases} 
\frac{1}{2}(\phi|_K)|_F + (\phi|_{K'})|_F & \text{if } F \in E_h^i \\
(\phi|_K)|_F & \text{if } F \in E_h^b.
\end{cases}$$

Each vertex $e = \overline{F} \cap \overline{F'}$ for some $F, F' \in E_h^b$. We thus define the jump and average over a vertex $e \in E_h^b$ analogously.

For two matrices $A, B \in \mathbb{R}^{m \times n}$, we set $A : B := \sum_{i,j=1}^{m,n} A_{ij} B_{ij}$. For an element $K$, we define the bilinear form $\langle \cdot, \cdot \rangle_K$ by

$$\langle u, v \rangle_K := \int_K u \cdot v \text{ if } u, v \in L^2(K; \mathbb{R}^{m \times n}).$$

Any ambiguity in this notation will be resolved by arguments of the bilinear form. The bilinear forms $\langle \cdot, \cdot \rangle_{\partial K}$ and $\langle \cdot, \cdot \rangle_F$ for $F \in E_h^{i,b}$, are defined similarly.

4.2 Exact domain approximation

We will continue this section by providing the details of [6], which provides us with a notion of exact domain approximation.
Definition 4.3 (Curved d-simplex) An open set \( K \subset \mathbb{R}^d \) is called a curved d-simplex if there exists a \( C^1 \) mapping \( F_K \) that maps a straight reference d-simplex \( \tilde{K} \) onto \( K \), and that is of the form

\[
F_K = \tilde{F}_K + \Phi_K,
\]

where

\[
\tilde{F}_K : \hat{x} \mapsto \hat{B}_K \hat{x} + \hat{b}_K
\]

is an invertible map and \( \Phi_K \in C^1(\tilde{K} ; \mathbb{R}^d) \) satisfies

\[
C_K := \sup_{\hat{x} \in K} \| D\Phi_K(\hat{x})\hat{B}_K^{-1} \| < 1,
\]

where \( \| \cdot \| \) denotes the induced Euclidean norm on \( \mathbb{R}^d \times \mathbb{R}^d \).

Definition 4.4 (Class \( m \) curved d-simplex) A curved d-simplex \( K \) is of class \( C^m \), \( m \geq 1 \), if the mapping \( F_K \) is of class \( C^m \) on \( \tilde{K} \).

Definition 4.5 (Mesh size) For each element \( K \in \mathcal{T}_h \), let \( h_K := \text{diam} \tilde{K} \geq C_T \| \hat{B}_K \| \) (where \( \tilde{K} = \hat{B}_K(\tilde{K}) \)). It is assumed that \( h = \max_{K \in \mathcal{T}_h} h_K \) for each mesh \( \mathcal{T}_h \). Furthermore, for each face \( F \in \mathcal{E}_h^{i,b} \), we define

\[
\tilde{h}_F := \begin{cases} 
\min(h_K, h_{K'}) & \text{if } F \in \mathcal{E}_h^i, \\
h_K & \text{if } F \in \mathcal{E}_h^b.
\end{cases}
\]

where \( K \) and \( K' \) are such that \( F = \partial K \cap \partial K' \) if \( F \in \mathcal{E}_h^i \), or \( F \subset \partial K \cap \partial \Omega \) if \( F \in \mathcal{E}_h^b \). Finally, for each \( e \in \mathcal{T}_h \), we define

\[
h_e := \min_{F \in \mathcal{E}_h^i, F \cap e \neq \emptyset} \tilde{h}_F.
\]

Definition 4.6 The family \( (\mathcal{T}_h)_h \) of meshes is said to be regular if there exist two constants, \( \sigma \) and \( c \), independent of \( h \), such that, for each \( h \), any \( K \in \mathcal{T}_h \) satisfies

\[
h_K / \rho_K \leq \sigma,
\]

where \( \rho_K \) is the diameter of the sphere inscribed in \( \tilde{K} \). Furthermore, we have

\[
\sup_{h} \sup_{K \in \mathcal{T}_h} C_K \leq c < 1.
\]

Definition 4.7 The family \( (\mathcal{T}_h)_h \) of meshes is said to be regular of order \( m \) if it is regular and if, for each \( h \), any \( K \in \mathcal{T}_h \) is of class \( C^{m+1} \), with

\[
\sup_{h} \sup_{K \in \mathcal{T}_h} \sup_{\hat{x} \in K} \| D^l F_K(\hat{x})\| \| \hat{B}_K \|^{-l} < \infty, \quad 2 \leq l \leq m + 1.
\]

Assumption 4.8 The meshes are allowed to be irregular, i.e., there may be hanging nodes. We assume that there is a uniform upper bound on the number of faces composing the boundary of any given element; in other words, there is a \( C_T > 0 \), independent of \( h \), such that

\[
\max_{K \in \mathcal{T}_h} \{ F \in \mathcal{E}_h^{i,b} : F \subset \partial K \} \leq C_T \quad \forall K \in \mathcal{T}_h, \forall h > 0.
\]

We assume that any two elements sharing a face have commensurate diameters, i.e., there is a \( C_T \geq 1 \), independent of \( h \), such that

\[
\max(h_K, h_{K'}) \leq C_T \min(h_K, h_{K'}),
\]

for any \( K \) and \( K' \) in \( \mathcal{T}_h \) that share a face.
Finally, we assume that each \( F \in \mathcal{E}_h^b \) satisfies
\[
F = F \cap \Gamma_n, \quad (4.12)
\]
for some \( n \in \{1, \ldots, N\}, \) with \( \Gamma_n \) given as in (4.1). This implies that each boundary face is completely contained in a boundary portion \( \Gamma_n, \) as well as ensuring that our approximation of the domain \( \Omega \) is exact.

**Remark 4.9** The assumptions on the mesh given by Assumption 4.8, in particular (4.11), show that if \( F \) is a face of \( K, \) then
\[
h_K \leq C_T h_F. \quad (4.13)
\]

### 4.3 Finite element spaces

For each \( K \in \mathcal{T}_h, \) we denote by \( \mathbb{P}^p(K) \) the space of all polynomials on \( K \) with total degree less than or equal to \( p. \) The discontinuous Galerkin finite element space \( V_{h,p} \) is defined by
\[
V_{h,p} := \{ v \in L^2(\Omega) : v|_K = \rho \circ F^{-1}_K, \rho \in \mathbb{P}^p(K), \forall K \in \mathcal{T}_h \}, \quad (4.14)
\]
we also define the subspace, \( V_{h,p,0}, \) of \( V_{h,p} \) as follows
\[
V_{h,p,0} := \{ v \in V_{h,p} : \int_{\Omega} v = 0 \}. \quad (4.15)
\]

Let \( s = (s_K : K \in \mathcal{T}_h) \) denote a vector of nonnegative real numbers and let \( r \in [1, \infty]. \) The broken Sobolev space \( W^{s,r}(\Omega; \mathcal{T}_h) \) is defined by
\[
W^{s,r}(\Omega; \mathcal{T}_h) := \{ v \in L^2(\Omega) : v|_K \in W^{s_K,r}(K) \forall K \in \mathcal{T}_h \}. \quad (4.16)
\]
We denote \( H^s(\Omega; \mathcal{T}_h) := W^{s,2}(\Omega; \mathcal{T}_h), \) and set \( W^{s,r}(\Omega; \mathcal{T}_h) := W^{s,r}(\Omega; \mathcal{T}_h), \) in the case that \( s_K = s, \) \( s \geq 0, \) for all \( K \in \mathcal{T}_h. \) For \( v \in W^{1,r}(\Omega; \mathcal{T}_h), \) let \( \nabla_h v \in L^r(\Omega; \mathbb{R}^d) \) denote the discrete (also known as broken) gradient of \( v, \) i.e., \( (\nabla_h v)|_K = \nabla v|_K \) for all \( K \in \mathcal{T}_h. \) Higher order discrete derivatives are defined in a similar way. We define a norm on \( W^{s,r}(\Omega; \mathcal{T}_h) \) by
\[
\|v\|_{W^{s,r}(\Omega; \mathcal{T}_h)} := \sum_{K \in \mathcal{T}_h} \|v\|_{W^{s,r}(K)} \quad (4.17)
\]
with the usual modification when \( r = \infty. \)

The following Lemma is a direct application of Lemma 4 from [38].

**Lemma 4.10** Let \( \Omega \) be a bounded domain, and let \( \mathcal{T}_h \) be a mesh on \( \Omega \) consisting of possibly curved simplices. Then, for each \( K \in \mathcal{T}_h \) and each face \( F \subset \partial K \) that belongs to \( \mathcal{E}_h^b, \) the following identities hold:
\[
\tau_F(\nabla v) = \nabla_T(\tau_F v) + \left( \tau_F \frac{\partial v}{\partial n_F} \right) n_F \quad \forall v \in H^s(K), s > 3/2, \quad (4.18)
\]
\[
\tau_F(\Delta v) = \text{div}_T \nabla_T(\tau_F v) + \tau_F \frac{\partial}{\partial n_F}(\nabla v \cdot n_F), \quad \forall v \in H^s(K), s > 5/2. \quad (4.19)
\]

### 4.4 The design of the numerical method

We shall now discuss how the terms in the definition bilinear form that defines the finite element method of this paper, arise. We are motivated by the desire to numerically enforce the Miranda–Talenti (MT) estimates (3.13) and (3.14), whilst producing a scheme that is both consistent and symmetric (the latter occurs when the operator \( A: D^2 \) is isotropic).

We solve for both \( u_h \in V_{h,p,0} := V_{h,p} \cap L^2_0(\Omega), \) which approximates the strong solution \( u \in H^2_{\beta,0}(\Omega) \) of (1.1), and \( c_h \in V_{h,0}, \) which approximates the compatibility constant of (1.1), that is, \( c_h \)
approximates the value of $C = \beta \cdot \nabla u |_{\partial \Omega}$ (the value of $C$ is a priori unknown). As such, our finite element space will be

$$M_h := V_{h,p,0} \times V_{h,0}.$$

We first note that the bilinear form, $A_h : M_h \times M_h \rightarrow \mathbb{R}$, that we use to define the finite element method, will take the following structure:

$$A_h((u_h, c_h); (v_h, \mu_h)) := \sum_{K \in \mathcal{T}_h} \langle \gamma A : D^2 u_h, \Delta v_h \rangle_K + B_{h,1/2}((u_h, c_h); (v_h, \mu_h))$$

$$- \sum_{K \in \mathcal{T}_h} \langle \Delta u_h, \Delta v_h \rangle_K \quad \forall (u_h, c_h), (v_h, \mu_h) \in M_h. \quad (4.18)$$

We claim that the bilinear form $B_{h,1/2}$ is coercive on $M_h \times M_h$, and that

$$B_{h,1/2}((w, c); (v_h, \mu)) = \sum_{K \in \mathcal{T}_h} \langle \Delta w, \Delta v_h \rangle_K, \quad \forall (v_h, \mu) \in V_{h,p,0} \times \mathbb{R}, \quad (4.19)$$

when $w \in H^2_\beta(\Omega) \cap H^s(\Omega; \mathcal{B}_h)$, $s > 5/2$, and $c = \beta \cdot \nabla w |_{\partial \Omega}$.

It is then clear that $(4.19)$ implies that

$$A_h((w, c); (v_h, \mu)) = \sum_{K \in \mathcal{T}_h} \langle \gamma A : D^2 u_h, \Delta v_h \rangle_K =: A_{\gamma,h}(w, v_h), \quad \forall (v_h, \mu) \in V_{h,p,0} \times \mathbb{R}, \quad (4.20)$$

for the aforementioned choice of $w$ and $c$. Note that $A_{\gamma,h}$ is a numerical discretisation of $A_{\gamma}$, defined by $(1.3)$.

The bilinear form $B_{h,1/2} : M_h \times M_h \rightarrow \mathbb{R}$ takes the following form

$$B_{h,1/2}((u_h, c_h); (v_h, \mu_h)) := \frac{1}{2} B_{h,*}((u_h, c_h); (v_h, \mu_h)) + \frac{1}{2} \sum_{K \in \mathcal{T}_h} \langle \Delta u_h, \Delta v_h \rangle_K$$

$$+ J_h((u_h, c_h); (v_h, \mu_h)), \quad (4.21)$$

where the bilinear forms $B_{h,*}, J_h : M_h \times M_h \rightarrow \mathbb{R}$ satisfy

$$B_{h,*}((w, c), (v_h, \mu)) = \sum_{K \in \mathcal{T}_h} \langle \Delta w, \Delta v_h \rangle_K, \quad \text{and} \quad J_h((w, c), (v_h, \mu)) = 0, \quad (4.22)$$

for all $(v_h, \mu) \in V_{h,p,0} \times \mathbb{R}$, when $w \in H^2_\beta(\Omega) \cap H^s(\Omega; \mathcal{B}_h)$, $s > 5/2$, and $c = \beta \cdot \nabla w |_{\partial \Omega}$.

Moreover, one can see that $(4.22)$ implies $(4.19)$, which in turn implies $(4.20)$. We also remark that the bilinear form $J_h$ is a jump penalty term that enforces regularity that is consistent with that of the true solution. In particular, if $w \in H^2_\beta(\Omega) \cap H^s(\Omega)$ (which is the space that the strong solution of $(1.1)$ belongs to), and $c = \beta \cdot \nabla w |_{\partial \Omega}$, then we see that

$$[c] = [w] = [\nabla w \cdot n_F] = [\nabla T w] = 0 \quad \forall F \in \mathcal{E}^b_h, \quad (4.23)$$

and furthermore, since $\tau_F(\beta \cdot \nabla w) = c$ for all $F \in \mathcal{E}_h$, it follows that

$$[\beta \cdot \nabla w - c] = [\partial T (\beta \cdot \nabla w)] = 0 \quad \forall F \in \mathcal{E}^b_h, \quad (4.24)$$

$J_h$ also enforces the oblique boundary condition, and leads to the bilinear form $B_{h,1/2}$ being provably coercive (in a particular $H^2$-type norm on $M_h$). In particular we define $J_h$ as follows:

$$J_h((u_h, \lambda), (v_h, \mu)) := \sum_{F \in \mathcal{E}^b_h} \mu_F \langle [\nabla_T u_h], [\nabla_T v_h] \rangle_F$$

$$+ \sum_{F \in \mathcal{E}^b_h} [\mu_F \langle [\nabla u_h \cdot n_F], [\nabla v_h \cdot n_F] \rangle_F + \eta_F \langle [u_h], [v_h] \rangle_F + \ell_F \langle [\lambda], [\mu] \rangle_F]$$

$$+ \sum_{F \in \mathcal{E}^b_h} \sigma_F (\beta \cdot \nabla u_h - \lambda, \beta \cdot \nabla v_h - \mu)_F, \quad (4.25)$$
where the positive edge-dependent quantities \( \mu_F, \eta_F, \ell_F, \) and \( \sigma_F \) will be specified later, and their particular choice will be made clear when we prove that \( B_{h,1/2} \) is coercive (see Lemma 4.14). Furthermore, (4.23) and (4.24) imply that
\[
J_h((w, e), (v_h, \mu)) = 0, \quad \forall (v_h, \mu) \in V_{h,p,0} \times \mathbb{R},
\]
when \( w \in H^2_{\tilde{\partial}, \tilde{\partial}}(\Omega) \cap H^s(\Omega; r_h), \ s > 5/2, \) and \( e = \beta \cdot \nabla w \big|_{\partial \Omega}. \)

The bilinear form \( B_{h,*} \) plays a key role in identity (4.19), and its structure is motivated by Corollary 3.8 and identity (4.12), the statements of which we recall.

**Statement of Corollary 3.8:** Assume that \( E \subset \mathbb{R}^2 \) is a bounded, Lipschitz, piecewise \( C^2 \) domain, and that \( \beta \in C^1(\Gamma_n; S^1) \) for each \( C^2 \) portion \( \Gamma_n \) of \( \partial E, \ n = 1, \ldots, N, \ N \in \mathbb{N} \). Then, for any \( u, v \in H^s(E), s > 5/2, \) we have that
\[
\int_{\partial K} \beta \cdot \nabla u \nabla v + \int_{\partial K} \beta \cdot \nabla u \nabla v - \beta \cdot \nabla u \beta \cdot \nabla v
\]
\[
= \int_{\partial K} \Delta u \Delta v.
\]

**Identity (4.12):** Let \( \Omega \subset \mathbb{R}^2 \) be a \( C^2 \) domain, and assume that \( \beta \in C^1(\partial \Omega; S^1) \). Then, on \( \partial \Omega \), we have that
\[
\beta_1 (\beta_2 - \beta_2 \delta_1) \beta_2 = \beta_1 \Theta + \mathcal{H} \delta_1.
\]

**Designing the bilinear form:** Let us assume that \( v_h \in V_{h,p,0}, w \in H^2_{\tilde{\partial}, \tilde{\partial}}(\Omega) \cap H^s(\Omega; r_h), \ s > 5/2, \) and \( e = \beta \cdot \nabla w \big|_{\partial \Omega}. \) Furthermore, we will assume that \( \Omega \) is a \( C^2 \) domain that is also piecewise \( C^3 \), and that \( (\mathcal{F}_h)_{h>0} \) is a regular of order 2 family of triangulations on \( \Omega \), satisfying assumption 4.8.

Let us consider \( K \in \mathcal{F}_h \) that satisfies \( |\partial K \cap \partial \Omega| \neq 0 \) (this allows for elements with one curved edge that lies on \( \partial \Omega \), but excludes elements that only intersect \( \partial \Omega \) at a vertex). Note that \( K \subset \mathbb{R}^2 \) is bounded with a Lipschitz continuous, piecewise \( C^1 \) boundary. \( K \) also has three edges \( F^1_K, F^2_K, F^3_K \), (each of which are \( C^3 \) and hence \( C^2 \) portions of \( \partial K \)) and three vertices \( e^1_K, e^2_K, e^3_K \). Let \( e^1_K \) and \( e^2_K \) be the two vertices that lie on \( \partial \Omega \), and let \( F^2_K \) be the curved side that lies on \( \partial \Omega \) and also connects \( e^1_K \) and \( e^2_K \). Finally, let \( F^3_K \) be the straight edge of \( K \) that connects \( e^2_K \), and \( e^3_K \). It then follows that \( F^3_K \) is the remaining straight edge that connects \( e^1_K \), and \( e^3_K \).

Now define \( \tilde{\beta} = \partial K \rightarrow S^1 \) by
\[
\begin{cases}
\tilde{\beta}|_{F^1_K} = \beta, \\
\tilde{\beta}|_{F^2_K} = \beta(e^2_K), \\
\tilde{\beta}|_{F^3_K} = \beta(e^1_K),
\end{cases}
\]
and so \( \tilde{\beta} \in C^1(F^1_K; S^1), j = 1, 2, 3, \) where \( \partial K = \cup_{j=1}^3 F^j_K \).

Then, noting that \( w, v_h \in H^s(K), s > 5/2, \) applying (4.27) with \( E := K, \) and \( \beta := \tilde{\beta}, \) we obtain
\[
\int_K D^2_{1w}wD^2_{2v}v + D^2_{2w}wD^2_{1v}v - 2D^2_{1w}wD^2_{1v}v =
\int_{\partial K} \beta_1 (\beta_2 - \beta_2 \delta_1) (\beta_2 \cdot \nabla w \beta \cdot \nabla v + \beta \cdot \nabla u \beta \cdot \nabla v)
\]
\[
+ \int_{\partial K} (\beta \cdot \nabla u \beta \cdot \nabla v - \beta (\beta \cdot \nabla w) \beta \cdot \nabla v)
\]
\[
= \int_{F^2_K} (\beta_1 (\beta_2 - \beta_2 \delta_1) (\beta_2 \cdot \nabla w \beta \cdot \nabla v + \beta \cdot \nabla u \beta \cdot \nabla v)
\]
\[
+ \int_{F^2_K \cup F^3_K} (\beta \cdot \nabla u \beta \cdot \nabla v)
\]
\[
= \int_{F^2_K \cup F^3_K} (\beta \cdot \nabla u \beta \cdot \nabla v - \beta \cdot \nabla u \beta \cdot \nabla v)
\]
Furthermore, upon noting that \(\beta, \beta^\perp\), and the unit normal to \(\partial K\), are all constant on \(F_K^2\) and \(F_K^3\), one can calculate the following:

\[
\int_{F_K^2 \cup F_K^3} \left( \partial_T (\beta^\perp \cdot \nabla w) \beta \cdot \nabla v_h - \partial_T (\beta \cdot \nabla w) \beta^\perp \cdot \nabla v_h \right) = \int_{F_K^2 \cup F_K^3} \Delta w (\nabla v_h \cdot n_{\partial K}) - \nabla (\nabla w \cdot n_{\partial K}) \cdot \nabla v_h. \tag{4.30}
\]

Since \(F_K^1 \subset \partial \Omega\), we may apply identity \(4.28\), obtaining

\[
\int_{F_K^1} (\beta_1 \partial_T \beta_2 - \beta_2 \partial_T \beta_1) (\beta^\perp \cdot \nabla w \beta^\perp \cdot \nabla v_h + \beta \cdot \nabla w \beta \cdot \nabla v_h) = \int_{F_K^1} (\partial_T \Theta + H_{F_K^1}) (\beta^\perp \cdot \nabla w \beta^\perp \cdot \nabla v_h + \beta \cdot \nabla w \beta \cdot \nabla v_h), \tag{4.31}
\]

where \(H_{F_K} = H_{\partial \Omega | F_K}\).

We now consider an element \(K \in \mathcal{T}_h\) that satisfies \(|\partial K \cap \partial \Omega| = 0\). An application of integration by parts (noting that the unit outward normal to \(\partial K\) is constant on each edge of \(K\)) yields

\[
\int_K (D^2 w : D^2 v_h) + \int_{\partial K} (\Delta w (\nabla v_h \cdot n_{\partial K}) - \nabla (\nabla w \cdot n_{\partial K}) \cdot \nabla v_h) = \int_K \Delta w \Delta v_h. \tag{4.32}
\]

One can also see that for any \(K \in \mathcal{T}_h\), and thus in particular, for those \(K \in \mathcal{T}_h\) that satisfy \(|\partial K \cap \partial \Omega| \neq 0\),

\[
\int_K (D^2 w : D^2 v_h + D_{11}^2 w D_{22}^2 v_h + D_{22}^2 w D_{11}^2 v_h - 2D_{12}^2 w D_{12}^2 v_h) = \int_K \Delta w \Delta v_h. \tag{4.33}
\]

Applying identities \(4.30\) and \(4.31\) to \(4.29\), and summing \(4.32\) over all \(K \in \mathcal{T}_h\) such that \(|\partial K \cap \partial \Omega| = 0\), and \(4.33\) over all \(K \in \mathcal{T}_h\) such that \(|\partial K \cap \partial \Omega| \neq 0\), we obtain

\[
\sum_{K \in \mathcal{T}_h} \int_K D^2 w : D^2 v_h + \sum_{F \in \mathcal{E}_h^0} \int_F [\Delta w \nabla v_h \cdot n_F - \nabla (\nabla w \cdot n_F) \cdot \nabla v_h] + \sum_{F \in \mathcal{E}_h^0} \int_F (\partial_T (\beta^\perp \cdot \nabla w) \beta \cdot \nabla v_h - \partial_T (\beta \cdot \nabla w) \beta^\perp \cdot \nabla v_h) + \sum_{F \in \mathcal{E}_h^0} \int_F ((\partial_T \Theta + H_F) \beta^\perp \cdot \nabla w \beta^\perp \cdot \nabla v_h + (\partial_T \Theta + H_F) \beta \cdot \nabla w \beta \cdot \nabla v_h)
\]

\[
\sum_{K \in \mathcal{T}_h} \int_K \Delta w \Delta v_h,
\]

where \(n_F\) is now a fixed choice of unit normal to \(F\), and \(H_F := H_{\partial \Omega | F}\). Utilising the tangential operator identities in \(4.17\), we obtain

\[
\sum_{F \in \mathcal{E}_h^0} \int_F [\Delta w \nabla v_h \cdot n_F - \nabla (\nabla w \cdot n_F) \cdot \nabla v_h] = \sum_{F \in \mathcal{E}_h^0} \int_F [\Delta_T w \nabla v_h \cdot n_F - \nabla_T (\nabla w \cdot n_F) \cdot \nabla_T v_h]. \tag{4.35}
\]

We then apply the identity (valid for any \(f, g \in H^2(\Omega; \mathcal{T}_h), s > 1/2\))

\[
\sum_{F \in \mathcal{E}_h^0} \int_F [fg] = \sum_{F \in \mathcal{E}_h^0} \int_F [f] [g] + \sum_{F \in \mathcal{E}_h^0} \int_F [f] [g],
\]

17
along with (4.23), to (4.35), which gives us

\[ \sum_{F \in \mathcal{E}_h^i} \int_{F} [\Delta w \nabla v_h \cdot n_F - \nabla (\nabla w \cdot n_F) \cdot \nabla v_h] \]

\[ = \sum_{F \in \mathcal{E}_h^i} \int_{F} (\langle \Delta_T w \rangle \langle \nabla v_h \cdot n_F \rangle + \langle \Delta_T w \rangle \langle \nabla v_h \rangle) \]

\[ - \sum_{F \in \mathcal{E}_h^i} \int_{F} (\langle \nabla_T (\nabla w \cdot n_F) \rangle \langle \nabla_T v_h \rangle + \langle \nabla_T (\nabla w \cdot n_F) \rangle \cdot \nabla_T v_h \rangle) \]

\[ = \sum_{F \in \mathcal{E}_h^i} \int_{F} (\langle \Delta_T w \rangle \langle \nabla v_h \cdot n_F \rangle - \langle \nabla_T (\nabla w \cdot n_F) \rangle \cdot \nabla_T v_h \rangle). \] (4.36)

Then, again by (4.23) we may consistently symmetrise the final right-hand side of (4.36), yielding

\[ \sum_{F \in \mathcal{E}_h^i} \int_{F} [\Delta w \nabla v_h \cdot n_F - \nabla (\nabla w \cdot n_F) \cdot \nabla v_h] \]

\[ = \sum_{F \in \mathcal{E}_h^i} \int_{F} (\langle \Delta_T w \rangle \langle \nabla v_h \cdot n_F \rangle + \langle \Delta_T v_h \rangle \langle \nabla w \cdot n_F \rangle) \]

\[ - \sum_{F \in \mathcal{E}_h^i} \int_{F} (\langle \nabla_T (\nabla w \cdot n_F) \rangle \cdot \nabla_T v_h \rangle + \langle \nabla_T (\nabla v_h \cdot n_F) \rangle \cdot \nabla_T v_h \rangle). \] (4.37)

From (4.24), we obtain

\[ \sum_{F \in \mathcal{E}_h^i} \int_{F} \partial_T (\beta \cdot \nabla w) \beta \cdot \nabla v_h = 0, \] (4.38)

for all \( F \in \mathcal{E}_h^b \). Furthermore, on \( F \in \mathcal{E}_h^b \), one has that

\[ \beta \cdot \nabla w \beta \cdot \nabla v_h + \beta \cdot \nabla w \beta \cdot \nabla v_h = \nabla w \cdot \nabla v_h. \] (4.39)

Applying (4.37), (4.38), and (4.39) to (4.34) we obtain

\[ \sum_{K \in \mathcal{K}_h} \int_{K} (D^2 w : D^2 v_h) + \sum_{F \in \mathcal{E}_h^i} \int_{F} (\langle \Delta_T w \rangle \langle \nabla v_h \cdot n_F \rangle + \langle \Delta_T v_h \rangle \langle \nabla w \cdot n_F \rangle) \]

\[ - \sum_{F \in \mathcal{E}_h^i} \int_{F} (\langle \nabla_T (\nabla w \cdot n_F) \rangle \cdot \nabla_T v_h \rangle + \langle \nabla_T (\nabla v_h \cdot n_F) \rangle \cdot \nabla_T v_h \rangle) \]

\[ + \sum_{F \in \mathcal{E}_h^i} \int_{F} (\partial_T (\beta \cdot \nabla w) \beta \cdot \nabla v_h + (\partial_T \Theta + H_F) \nabla w \cdot \nabla v_h) \]

\[ = \sum_{K \in \mathcal{K}_h} \int_{K} \Delta w \Delta v_h. \] (4.40)

So far, all of the applications of (4.23) and (4.24) have been made with consistency and symmetry in mind. We make a penultimate alteration, which is necessary for the coercivity of \( B_{h,1/2} \). In particular, notice that each term of each integrand on the left-hand side of (4.40) either has a sign if we take \( w = v_h \) (in particular, \( D^2 w : D^2 v_h \) and \( (\partial_T \Theta + H_F) \nabla w \cdot \nabla v_h \)), or consists of the product of two terms, one of which is present in the definition (4.25) of the jump stabilisation bilinear form, \( J_h \), except for the integrand \( \partial_T (\beta \cdot \nabla w) \beta \cdot \nabla v_h \).
To this end, let us denote by $e_F^+$ and $e_F^-$ the two vertices of an edge $F \in \mathcal{E}_h$, and notice that for any $\mu \in \mathbb{R}$,

$$
\sum_{F \in \mathcal{E}_h} \int_F \partial_T (\beta^{\perp} \cdot \nabla w) \mu = \sum_{F \in \mathcal{E}_h} (\beta^{\perp} \cdot \nabla w) \mu |_{e_F^+} = \mu \sum_{e \in \mathcal{V}_h} \| \beta^\perp \cdot \nabla w \| = 0,
$$

(4.41)

where the jumps in (4.41) are considered across boundary vertices $e \in \mathcal{V}_h$. Note that the final equality holds, due to the fact that $\beta^\perp \in C^1(\partial \Omega)$, and $\nabla w \in H^{1/2}(\partial \Omega)$, and thus neither function may jump across boundary vertices.

Applying (4.41) to (4.40), we obtain

$$
\sum_{K \in \mathcal{T}_h} \int_K (D^2 w : D^2 v_h) + \sum_{F \in \mathcal{E}_h} \int_F \left( \langle \langle \Delta_T w \rangle \rangle [\nabla v_h \cdot n_F] + \langle \langle \Delta_T v_h \rangle \rangle [\nabla w \cdot n_F] \right)
- \sum_{F \in \mathcal{E}_h} \int_F \left( \langle \langle \nabla_T (\nabla w \cdot n_F) \rangle \rangle \cdot [\nabla_T v_h] + \langle \langle \nabla_T (\nabla v_h \cdot n_F) \rangle \rangle \cdot [\nabla_T w] \right)
+ \sum_{F \in \mathcal{E}_h} \int_F (\partial_T (\beta^\perp \cdot \nabla w) (\beta \cdot \nabla v_h - \mu) + (\partial_T \Theta + \mathcal{H}_F) \nabla w \cdot \nabla v_h)
= \sum_{K \in \mathcal{T}_h} \int_K \Delta w \Delta v_h.
$$

(4.42)

Alas, the left-hand side of (4.42) is not symmetric. However, by (4.43), it follows that

$$
\sum_{F \in \mathcal{E}_h} \langle \partial_T (\beta^\perp \cdot \nabla v_h), \beta \cdot \nabla w - c \rangle_F = 0,
$$

(4.43)

which, when applied to (4.42), gives us

$$
\sum_{K \in \mathcal{T}_h} \int_K (D^2 w : D^2 v_h) + \sum_{F \in \mathcal{E}_h} \int_F \left( \langle \langle \Delta_T w \rangle \rangle [\nabla v_h \cdot n_F] + \langle \langle \Delta_T v_h \rangle \rangle [\nabla w \cdot n_F] \right)
- \sum_{F \in \mathcal{E}_h} \int_F \left( \langle \langle \nabla_T (\nabla w \cdot n_F) \rangle \rangle \cdot [\nabla_T v_h] + \langle \langle \nabla_T (\nabla v_h \cdot n_F) \rangle \rangle \cdot [\nabla_T w] \right)
+ \sum_{F \in \mathcal{E}_h} \int_F (\partial_T (\beta^\perp \cdot \nabla w) (\beta \cdot \nabla v_h - \mu) + (\partial_T \Theta + \mathcal{H}_F) \nabla w \cdot \nabla v_h)
= \sum_{K \in \mathcal{T}_h} \int_K \Delta w \Delta v_h,
$$

(4.44)

consistently restoring symmetry.
We then define $B_{h,*}$ by the left-hand side of (4.33). That is,

$$
B_{h,*}((u_h, \lambda), (v_h, \mu)) := \sum_{K \in \mathcal{T}_h} \langle D^2u_h, D^2v_h \rangle_K \\
+ \sum_{F \in \Sigma^*_h} [\langle \nabla \nabla T \langle \langle u_h \rangle \rangle, [\nabla v_h \cdot n_F] \rangle_F + \langle \nabla \nabla T \langle \langle v_h \rangle \rangle, [\nabla u_h \cdot n_F] \rangle_F] \\
- \sum_{F \in \Sigma_h^*} [\langle \nabla \nabla T \langle \langle u_h \rangle \rangle, [\nabla v_h] \rangle_F + \langle \nabla \nabla T \langle \langle v_h \rangle \rangle, [\nabla u_h] \rangle_F] \\
+ \sum_{F \in \Sigma_h^*} [\langle \partial_T(\beta^+ \cdot \nabla u_h) \rangle \beta \cdot \nabla v_h - \mu \rangle_F + \langle \partial_T(\beta^+ \cdot \nabla v_h) \rangle \beta \cdot \nabla u_h - \lambda \rangle_F],
$$

(4.45)

for all $(u_h, \lambda), (v_h, \mu) \in M_h$, and we recall that $n_F$ is a fixed choice of unit normal to $F$, and $H_F := \partial_T \cdot n_F = [\partial_T n] F$ for $F \in \Sigma_h^*$. It follows from (4.33) that the bilinear form $B_{h,*}$ satisfies the first identity of (4.22). We are now ready to define the numerical method of this chapter.

4.5 Finite element method

Recall that we define $M_h := V_{h,p,0} \times V_{h,0}$. The definition of the finite element method first requires the definition of the jump stabilisation bilinear form $J_h : M_h \times M_h \to \mathbb{R}$, given by

$$
J_h((u_h, \lambda), (v_h, \mu)) := \sum_{F \in \Sigma^*_h} \mu_F [\langle \nabla \nabla T u_h \rangle, [\nabla \nabla T v_h] \rangle_F \\
+ \sum_{F \in \Sigma^*_h} \mu_F [\langle \nabla u_h \cdot n_F \rangle, [\nabla v_h \cdot n_F] \rangle_F + \eta_F [\langle \nabla u_h \rangle, \langle \nabla v_h \rangle] \rangle_F + \ell_F [\langle \lambda \rangle, [\mu] \rangle_F] \\
+ \sum_{\partial F \in \mathcal{E}_h} \sigma_F \langle \beta \cdot \nabla u_h - \lambda, \beta \cdot \nabla v_h - \mu \rangle_F,
$$

where $\mu_F, \eta_F, \ell_F, \sigma_F$ are positive, face dependent parameters to be provided. Furthermore for $\theta \in [0, 1]$, we define the bilinear form $B_{h,\theta} : M_h \times M_h \to \mathbb{R}$

$$
B_{h,\theta}((u_h, \lambda), (v_h, \mu)) := \theta B_{h,*}((u_h, \lambda), (v_h, \mu)) + (1 - \theta) \sum_{K \in \mathcal{T}_h} \langle \Delta u_h, \Delta v_h \rangle_K \\
+ J_h((u_h, \lambda), (v_h, \mu)),
$$

where $B_{h,*}$ is given by (4.35). We now define

$$
A_h((u_h, \lambda), (v_h, \mu)) := \sum_{K \in \mathcal{T}_h} \langle [\gamma L u_h, \Delta v_h] K - \langle \Delta u_h \rangle, \Delta v_h \rangle K \rangle + B_{h,1/2}((u_h, \lambda), (v_h, \mu)).
$$

(4.46)

We can now state the finite element method: find $(u_h, c_h) \in M_h$ such that

$$
A_h((u_h, c_h), (v_h, \mu)) = \sum_{K \in \mathcal{T}_h} \langle \gamma f, \Delta v_h \rangle K \forall (v_h, \mu) \in M_h.
$$

(4.47)

4.6 Consistency of the method

Remark 4.11 (Extension of the bilinear forms) The bilinear forms $B_{h,*}$ and $J_h$ are both defined on $M_h \times M_h$, but one must note that they are both well defined on $(H^s(\Omega; \mathcal{T}_h) \cap H^2_{\beta,0}(\Omega) \times V_{h,0}) \times M_h$ for $s > 5/2$, of which $(H^4(\Omega; \mathcal{T}_h) \cap H^2_{\beta,0}(\Omega) \times \mathbb{R}) \times (V_{h,p,0} \times \mathbb{R})$ is a proper subset, that the functions in the following lemma belong to.
Lemma 4.12 Let $\Omega \subset \mathbb{R}^2$ be a $C^2$ and piecewise $C^3$ domain, and let $\beta \in C^1(\partial \Omega; S^1)$. Furthermore, assume that $\{\mathcal{T}_h\}_h$ is a regular of order 2 family of triangulations on $\overline{\Omega}$ satisfying Assumption 4.8. Let $(w, c) \in H^s(\Omega; \mathcal{T}_h) \cap H^{s/2}_0(\Omega) \times \mathbb{R}$, $s > 5/2$, where $\beta \cdot \nabla w |_{\partial \Omega} = c$. Then, for every $(v_h, \mu) \in V_{h,p} \times \mathbb{R}$, we have the identities

$$B_h,\ast((w, c), (v_h, \mu)) = \sum_{K \in \mathcal{T}_h} \langle \Delta w, \Delta v_h \rangle_K \quad \text{and} \quad J_h((w, c), (v_h, \mu)) = 0. \quad (4.48)$$

Proof: Assume that the pair $(w, c)$ satisfies the hypotheses of the lemma. Then, the identities of (4.48) follow from (4.11) and (4.20). \(\square\)

4.7 Stability of the method

We now aim to show that $B_{h,\theta}$ is coercive in a particular norm on $M_h$. Before we prove that $B_{h,\theta}$ is coercive, we must define the norm in which the bilinear form is coercive. To this end, let us define the following family of functionals, $\| \cdot \|_{h,\theta} : M_h \to [0, \infty)$ for $\theta \in (0, 1)$:

$$\| (u_h, \lambda) \|_{h,\theta}^2 := \sum_{K \in \mathcal{T}_h} [\theta |u_h|_{H^2(K)}^2 + (1 - \theta) \| \Delta u_h \|_{L^2(K)}^2] + c_* J_h((u_h, \lambda), (u_h, \lambda)) + \frac{\theta}{2} \sum_{F \in \mathcal{E}_h^i} \| (\partial \Theta + \mathcal{H}_F) \|_{L^2(F)}^{1/2} \| \nabla u_h \|_{L^2(F)}^2. \quad (4.49)$$

where $c_*$ is a positive constant to be determined.

Lemma 4.13 Let $\Omega \subset \mathbb{R}^2$ be a $C^2$ and piecewise $C^3$ domain, and let $\beta \in C^1(\partial \Omega; S^1)$. Assume that $\partial \Theta + \mathcal{H}_{\partial \Omega} > 0$ on $\partial \Omega$.

Furthermore, assume that $\{\mathcal{T}_h\}_h$ is a regular of order 2 family of triangulations on $\overline{\Omega}$ satisfying Assumption 4.8. Then, for each $\theta \in (0, 1)$, $\| \cdot \|_{h,\theta} : M_h \to [0, \infty)$ defines a norm on $M_h$.

Proof: First we note that homogeneity and the triangle inequality are clear. Now let us assume that the pair $(v_h, \mu) \in M_h$ satisfies

$$\| (v_h, \mu) \|_{h,\theta} = 0,$$

for some $\theta \in (0, 1]$. It then follows that $|v_h|_{H^s(\Omega; \mathcal{T}_h)} = 0$, and so $v_h$ is piecewise affine. Moreover

$$[\mu]_F = [v_h]_F = [\nabla v_h]_F = 0 \quad \text{for} \quad F \in \mathcal{E}_h^i,$$

and, as $\partial \Theta + \mathcal{H}_{\partial \Omega} > 0$ on $\partial \Omega$, it follows that

$$[\nabla v_h]_F = 0 \quad \text{for} \quad F \in \mathcal{E}_h^b.$$ 

It then follows that $v_h$ is affine, i.e., $v_h = a^T x + b$, with $a \in \mathbb{R}^d$, $b \in \mathbb{R}$, and that $\mu$ is constant. But then we see that

$$0 = \nabla v_h |_F = a,$$

for $F \in \mathcal{E}_h^b$, and thus $a = 0$, i.e., $v_h = b$. Then, since $v_h \in V_{h,p,0}$, $0 = \int_{\Omega} v_h = |\Omega| b$, and so $b = 0$, i.e., $v_h \equiv 0$.

Finally, we see that $J_h((v_h, \mu), (v_h, \mu)) = 0$, and it follows that

$$0 = \beta \cdot \nabla v_h = \mu \quad \text{on} \quad \partial \Omega,$$

and so $\mu \equiv 0$. Overall, we have obtained $(v_h, \mu) \equiv (0, 0)$. \(\square\)
Lemma 4.14 Let \( \Omega \subset \mathbb{R}^2 \) be a \( C^2 \) and piecewise \( C^3 \) domain, and let \( \beta \in C^4(\partial \Omega; S^1) \). Assume that

\[
\partial_T \Theta + \mathcal{H}_{\partial \Omega} > 0 \quad \text{on } \partial \Omega.
\]

Furthermore, assume that \( \{ \mathcal{T}_h \}_{h} \) is a regular of order 2 family of triangulations on \( \Omega \) satisfying Assumption 4.3. Then, for each constant \( \kappa > 1 \), there exist positive constants \( c_{\text{stab}} \) and \( c_\star \), independent of \( h \) and \( \theta \), such that

\[
B_{h, \theta}((u_h, \lambda), (u_h, \lambda)) \geq \kappa^{-1} \|(u_h, \lambda)\|_{h, \theta}^2 \quad \forall (u_h, \lambda) \in M_h, \forall \theta \in [0, 1],
\]

whenever

\[
\mu_F \geq \frac{c_{\text{stab}}}{h_F}, \sigma_F \geq \frac{c_{\text{stab}}}{h_F} \quad \text{and } \eta_F, \ell_F > 0.
\]

Proof: We see that for \((u_h, \lambda) \in V_{h, \theta} \times V_{h, 0},\)

\[
B_{h, \theta}((u_h, \lambda), (u_h, \lambda)) = \sum_{K \in \mathcal{T}_h} \left[ \theta (D^2 u_h, D^2 u_h)_K + (1 - \theta) (\Delta u_h, \Delta u_h)_K \right] + 2\theta \sum_{F \in \mathcal{E}_h^k} \left[ \langle \text{div}_F \nabla \langle u_h \rangle, \nabla u_h \cdot n_F \rangle_F - \langle \nabla \langle u_h \rangle \cdot n_F, \nabla u_h \rangle_F \right]
\]

\[
+ \theta \sum_{F \in \mathcal{E}_h^k} \left( \| (\partial_T \Theta + \mathcal{H}_F) \|_{1/2} \beta^\perp \cdot \nabla u_h \|_{L^2(F)}^2 + \| (\partial_T \Theta + \mathcal{H}_F) \|_{1/2} \beta \cdot \nabla u_h \|_{L^2(F)}^2 \right)
\]

\[
+ 2\theta \sum_{F \in \mathcal{E}_h^k} \langle \partial_T (\beta^\perp \cdot \nabla u_h), \beta \cdot \nabla u_h - \lambda \rangle_F
\]

\[
+ \sum_{F \in \mathcal{E}_h^k} \mu_F \| \nabla u_h \|_{L^2(F)}^2 + \mu_F \| \nabla u_h \cdot n_F \|_{L^2(F)}^2 + \eta_F \| u_h \|_{L^2(F)}^2 + \ell_F \| \lambda \|_{L^2(F)}^2
\]

\[
+ \sum_{F \in \mathcal{E}_h^k} \sigma_F \| \beta \cdot \nabla u_h - \lambda \|_{L^2(F)}^2.
\]

Now notice that for any \( \alpha > 0, \)

\[
|I_1| := \left| \sum_{F \in \mathcal{E}_h^k} \langle \text{div}_F \nabla \langle u_h \rangle, \nabla u_h \cdot n_F \rangle_F \right|
\]

\[
\leq \left( \sum_{F \in \mathcal{E}_h^k} \alpha h_F \| \text{div}_F \nabla \langle u_h \rangle \|_{L^2(F)}^2 \right)^{1/2} \left( \sum_{F \in \mathcal{E}_h^k} \frac{1}{\alpha h_F} \| \nabla u_h \cdot n_F \|_{L^2(F)}^2 \right)^{1/2},
\]

and (associating \( F = K \cap K' \) for some \( K, K' \in \mathcal{T}_h \))

\[
\| \text{div}_F \nabla \langle u_h \rangle \|_{L^2(F)}^2 \leq \frac{1}{2} \| \text{div}_K \nabla u_h |_K \|_{L^2(F)}^2 + \frac{1}{2} \| \text{div}_K \nabla u_h |_{K'} \|_{L^2(F)}^2.
\]

Therefore, the trace inequality gives us

\[
\frac{\alpha}{2} \sum_{F \in \mathcal{E}_h^k} \frac{h_F}{\alpha} \| \text{div}_F \nabla \langle u_h \rangle \|_{L^2(F)}^2 \leq \frac{\alpha}{2} \sum_{F \in \mathcal{E}_h^k} \frac{h_F}{\alpha} \sum_{K \in \mathcal{T}_h, F \subset \partial K} \| D^2 u_h \|_{L^2(\partial K)}^2
\]

\[
\leq \frac{\alpha C_{\text{Tr}}}{2} \sum_{F \in \mathcal{E}_h^k} h_F \sum_{K \in \mathcal{T}_h, F \subset \partial K} h_K^{-1} | u_h |_{H^2(K)}^2 + h_K | u_h |_{H^3(K)}^2,
\]

where \( C_{\text{Tr}} \) is the constant of the trace inequality, and is independent of \( K \) and \( h_K \). We now apply an inverse estimate above, noting that since a given function of the finite element space is a polynomial.
composed with a nonaffine function \( F_K \), by the chain rule, the inverse estimate takes the following form

\[
|u_h|^2_{H^2(K)} \leq C_I h_K^{-2} (|u_h|_{H^2(K)}^2 + |u_h|_{H^1(K)}^2),
\]

where \( C_I \) a constant independent of \( K \) and \( h_K \). This results in

\[
\frac{\alpha}{2} \sum_{F \in \mathcal{E}_h^1} h_F \| \text{div}_T \nabla_T \langle u_h \rangle \|^2_{L^2(F)} \leq \frac{\alpha C_T C_I}{2} \sum_{F \in \mathcal{E}_h^1} h_F \sum_{K \in \mathcal{K}_h : F \subset \partial K} h_K^{-1} (|u_h|_{H^2(K)}^2 + |u_h|_{H^1(K)}^2),
\]

\[
\leq \frac{\alpha C_T C_I C_F}{2} \sum_{K \in \mathcal{K}_h} |u_h|_{H^2(K)}^2 + |u_h|_{H^1(K)}^2,
\]

where the final inequality is due to the fact that the number of faces that make up a simplex \( K \in \mathcal{K}_h \) is bounded by \( C_F \). Applying the above estimate to (4.53), we obtain

\[
|I_1| \leq \frac{\alpha C_T C_I C_F}{2} \sum_{K \in \mathcal{K}_h} |u_h|_{H^2(K)}^2 + |u_h|_{H^1(K)}^2 + \sum_{F \in \mathcal{E}_h^1} h_F^{-1} \| \nabla u_h \cdot n_F \|^2_{L^2(F)}. \tag{4.54}
\]

Similarly, for any \( \alpha > 0 \)

\[
|I_2| := \sum_{F \in \mathcal{E}_h^0} \| \nabla_T (\nabla u_h \cdot n_F) - \nabla_T u_h \|^2_F \leq \frac{\alpha C_T C_I C_F}{2} \sum_{K \in \mathcal{K}_h} |u_h|_{H^2(K)}^2 + |u_h|_{H^1(K)}^2 + \sum_{F \in \mathcal{E}_h^1} \frac{1}{2\alpha h_F} \| \nabla u_h \|^2_{L^2(F)}. \tag{4.55}
\]

Since \( \beta \in C^1(\partial \Omega ; \mathbb{S}^1) \), utilising the trace and inverse inequalities, we also see that for any \( \alpha > 0 \),

\[
|I_3| := \sum_{F \in \mathcal{E}_h^0} \langle \partial \Omega \cdot \nabla u_h - \nabla u_h \rangle \leq \sum_{F \in \mathcal{E}_h^1} \langle \partial \Omega \cdot \nabla u_h \rangle \leq \sum_{F \in \mathcal{E}_h^0} \| \nabla u_h \|^2_{L^2(F)} + \| D^2 u_h \|^2_{L^2(F)} \| \beta \cdot \nabla u_h - \lambda \|^2_{L^2(F)} \leq \sum_{F \in \mathcal{E}_h^0} \frac{C(\beta)^2}{2\alpha h_F} \| \beta \cdot \nabla u_h - \lambda \|^2_{L^2(F)} + \sum_{F \in \mathcal{E}_h^1} \frac{h_F}{2} \| u_h \|^2_{H^2(F)} + |u_h|_{H^1(\partial K)}^2 \leq \sum_{F \in \mathcal{E}_h^0} \frac{C(\beta)^2}{2\alpha h_F} \| \beta \cdot \nabla u_h - \lambda \|^2_{L^2(F)} + \frac{\alpha C_T C_I C_F}{2} \sum_{K \in \mathcal{K}_h} |u_h|_{H^2(K)}^2 + |u_h|_{H^1(K)}^2, \tag{4.56}
\]

Now, by the discrete Poincaré–Friedrichs’ inequality of [2], we obtain

\[
\sum_{K \in \mathcal{K}_h} |u_h|_{H^1(K)}^2 \leq C(\sigma) \left( \sum_{K \in \mathcal{K}_h} |u_h|_{H^2(K)}^2 + \sum_{F \in \mathcal{E}_h^0} h_F^{-1} \| \nabla u_h \|^2_{L^2(F)} + \frac{1}{|\mathcal{K}|} \sum_{F \in \mathcal{E}_h^1} \| \nabla u_h \|^2_{L^2(F)} \right)
\leq C(\sigma) \left( \sum_{K \in \mathcal{K}_h} |u_h|_{H^2(K)}^2 + \sum_{F \in \mathcal{E}_h^0} h_F^{-1} \| \nabla u_h \|^2_{L^2(F)} + \frac{\Theta}{|\mathcal{K}|} \sum_{F \in \mathcal{E}_h^1} \| \Omega \|_{H^2(F)} \right) \| \partial_T \Theta + H_f \|^2_{L^2(F)} \right)^{1/2} \| \nabla u_h \|^2_{L^2(F)}, \tag{4.57}
\]

where \( C(\sigma) \) depends only upon the shape-regularity constant of the family of meshes \( (\mathcal{K}_h)_h \), and the final inequality follows from the following observation: for any \( F \in \mathcal{E}_h^0 \),

\[
\| \nabla u \|^2_{L^2(F)} \leq \int_F \frac{1}{\partial_T \Theta + H_f} \| \partial_T \Theta + H_f \|^{1/2} \| \nabla u \|^2 \leq \left( \min_{F \in \mathcal{E}_h^0} \frac{1}{\partial_T \Theta + H_f} \right)^{-1} \| \partial_T \Theta + H_f \|^{1/2} \| \nabla u \|^2_{L^2(F)},
\]

23
where \( \Theta_\ast := \left( \min_{F \in \mathcal{E}_h^k} \inf_{\Gamma} (\partial_\Gamma \Theta + \mathcal{H}_F) \right)^{-1} \) is uniformly positive, due to (4.50). Applying (4.57) to (4.54)-(4.56), and summing the resulting inequalities, we obtain

\[
\sum_{i=1}^{3} |I_i| \leq \frac{3\alpha C_T C_I C_F (1 + C(\sigma))}{2} \sum_{K \in \mathcal{T}_h} ||D^2 u_h||^2_{L^2(K)}
+ \sum_{F \in \mathcal{E}_h^k} \tilde{h}_F^{-1} \left( \frac{1}{2\alpha} + \frac{3\alpha C_T C_I C_F C(\sigma)}{2} \right) \left( ||\nabla T u_h||^2_{L^2(F)} + ||\nabla u_h \cdot n_F||^2_{L^2(F)} \right)
+ \sum_{F \in \mathcal{E}_h^k} C(\beta)^2 \frac{2}{\alpha h_F} ||\beta \cdot \nabla u_h - \lambda||^2_{L^2(F)} + \frac{3\alpha C_T C_I C_F C(\sigma) \Theta_\ast}{2|\partial \Omega|} ||\nabla u_h||^2_{L^2(F)}.
\]

The above estimate implies that \( B_{h,\theta}((u_h, \lambda), (u_h, \lambda)) \geq \sum_{i=1}^{7} A_i \), where

\[
A_1 := \theta (1 - 3\alpha C_T C_I C_F (1 + C(\sigma))) \sum_{K \in \mathcal{T}_h} ||D^2 u_h||^2_{L^2(K)}, \quad A_2 := (1 - \theta) \sum_{K \in \mathcal{T}_h} ||\Delta u_h||^2_{L^2(K)},
\]

\[
A_3 := \sum_{F \in \mathcal{E}_h^k} \left( \mu_F - \frac{\theta}{h_F} (\alpha^{-1} + 3\alpha C_T C_I C_F C(\sigma)) \right) ||\nabla u_h \cdot n_F||^2_{L^2(F)},
\]

\[
A_4 := \sum_{F \in \mathcal{E}_h^k} \eta_F |||u_h|||_{L^2(F)}^2 + \ell_F |||\lambda|||_{L^2(F)}^2,
\]

\[
A_5 := \sum_{F \in \mathcal{E}_h^k} \left( \mu_F - \frac{\theta}{h_F} (\alpha^{-1} + 3\alpha C_T C_I C_F C(\sigma)) \right) ||\nabla T u_h||^2_{L^2(F)},
\]

\[
A_6 := \sum_{F \in \mathcal{E}_h^k} \left( \sigma_F - \frac{\theta C(\beta)^2}{\alpha h_F} \right) ||\beta \cdot \nabla u_h - \lambda||^2_{L^2(F)},
\]

\[
A_7 := \theta \left( 1 - \frac{3\alpha C_T C_I C_F C(\sigma) \Theta_\ast}{|\partial \Omega|} \right) \sum_{F \in \mathcal{E}_h^k} ||(\partial_\Gamma \Theta + \mathcal{H}_F)^{1/2} \nabla u_h||^2_{L^2(F)}.
\]

Now let \( \kappa > 1 \) be given. Then, since \( \kappa^{-1} < 1 \), there exists \( \alpha > 0 \) sufficiently small such that

\[
\min \left\{ 1 - 3\alpha C_T C_I C_F (1 + C(\sigma)), 1 - \frac{3\alpha C_T C_I C_F C(\sigma) \Theta_\ast}{|\partial \Omega|} \right\} > \kappa^{-1},
\]

we then choose \( c_{stab} := 2 \max \{ \alpha^{-1} + 3\alpha C_T C_I C_F C(\sigma), \alpha C(\beta)^2 \} \), \( c_\ast = \kappa / 2 \) and note that by assumption, \( \mu_F \geq c_{stab} / h_F \) and \( \sigma_F \geq c_{stab} / h_F \). Therefore, for any \( \theta \in [0,1] \),

\[
A_3 \geq \frac{2}{\alpha} \sum_{F \in \mathcal{E}_h^k} \mu_F |||\nabla u_h \cdot n_F|||_{L^2(F)}^2 = \kappa^{-1} c_\ast \sum_{F \in \mathcal{E}_h^k} |||\nabla u_h \cdot n_F|||_{L^2(F)},
\]

\[
A_4 \geq \frac{1}{2} A_4 = \kappa^{-1} c_\ast \sum_{F \in \mathcal{E}_h^k} \eta_F (||u_h||_{L^2(F)}^2 + ||F||_{L^2(F)}^2),
\]

\[
A_5 \geq \frac{1}{2} \sum_{F \in \mathcal{E}_h^k} \mu_F |||\nabla T u_h|||_{L^2(F)}^2 = \kappa^{-1} c_\ast \sum_{F \in \mathcal{E}_h^k} \mu_F |||\nabla T u_h|||_{L^2(F)}^2,
\]

\[
A_6 \geq \frac{1}{2} \sum_{F \in \mathcal{E}_h^k} \sigma_F ||\beta \cdot \nabla u_h - \lambda||_{L^2(F)}^2 = \kappa^{-1} c_\ast \sum_{F \in \mathcal{E}_h^k} \sigma_F ||\beta \cdot \nabla u_h - \lambda||_{L^2(F)}^2.
\]

Thus, we obtain

\[
\kappa B_{h,\theta}((u_h, \lambda), (u_h, \lambda)) \geq \sum_{K \in \mathcal{T}_h} [\theta ||D^2 u_h||^2_{L^2(K)} + (1 - \theta)||\Delta u_h||^2_{L^2(K)}]
+ c_\ast J_h((u_h, \lambda), (u_h, \lambda)) + \theta \sum_{F \in \mathcal{E}_h^k} ||(\partial_\Gamma \Theta + \mathcal{H}_F)^{1/2} \nabla u_h||^2_{L^2(F)}.
\]

\[
\square
\]
We will now prove that $A_h$ is coercive in $\| \cdot \|_{h,1}$.

**Theorem 4.15** Under the assumptions of Lemma 4.14, let $c_{stab}$ and $c_*, \mu_F, \eta_F, \sigma_F,$ and $\ell_F$ be chosen so that (4.51) and (4.52) hold with $\kappa < (1 - \varepsilon)^{1/2}$. Let the operator $L$ be uniformly elliptic (and thus satisfy the Cordes condition (3.1)). Then, the operator $A_h$ is coercive in $\| \cdot \|_{h,1}$. In particular, for any $(v_h, \mu) \in M_h$, we have

$$\| (v_h, \mu) \|^2_{h,1} \leq \frac{2\kappa}{1 - \kappa^2(1 - \varepsilon)} A_h((v_h, \mu), (v_h, \mu)). \tag{4.58}$$

Therefore, there exists a unique solution pair $(u_h, c_h) \in M_h$ of the numerical scheme (4.37).

Moreover, the pair $(u_h, c_h)$ satisfies

$$\| (u_h, c_h) \|_{h,1} \leq \frac{2\sqrt{2}k_1 \gamma L_{\infty}(\Omega)}{1 - \kappa^2(1 - \varepsilon)} \| f \|_{L^2(\Omega)}. \tag{4.59}$$

**Proof:** Let $(v_h, \mu) \in M_h$, then we have that for any $K \in \mathcal{T}_h$:

$$\langle \gamma L v_h - \Delta v_h, \Delta v_h \rangle_K \leq \| (\gamma L - \Delta) v_h \|_{L^2(\Omega)} \| \Delta v_h \|_{L^2(\Omega)} \leq \sqrt{1 - \varepsilon} \| D^2 v_h \|_{L^2(\Omega)} \| \Delta v_h \|_{L^2(\Omega)}.$$

Applying the Cauchy–Schwarz inequality with a parameter, and (4.51) then gives us

$$A_h((v_h, \mu); (v_h, \mu)) = B_{h,1/2}((v_h, \mu); (v_h, \mu)) + \sum_{K \in \mathcal{T}_h} \langle \gamma L v_h - \Delta v_h, \Delta v_h \rangle_K + \| \Delta v_h \|^2_{L^2(\Omega)}$$

$$\geq \kappa^{-1} \| (v_h, \mu) \|^2_{h,1/2} + \sum_{K \in \mathcal{T}_h} \| \Delta v_h \|^2_{L^2(\Omega)} - \sqrt{1 - \varepsilon} \| D^2 v_h \|_{L^2(\Omega)} \| \Delta v_h \|_{L^2(\Omega)}$$

$$\geq \kappa^{-1} \left( \sum_{K \in \mathcal{T}_h} \frac{1}{2} \| D^2 v_h \|^2_{L^2(\Omega)} + \frac{1}{2} \| \Delta v_h \|^2_{L^2(\Omega)} \right)$$

$$+ \frac{1}{2} \sum_{F \in \mathcal{E}_h^b} \| (\partial_T \Theta + \mathcal{H}_F)^{1/2} \nabla u_h \|^2_{L^2(F)} - \frac{1}{2} \sum_{K \in \mathcal{T}_h} \left[ \frac{\kappa(1 - \varepsilon)}{2} \| D^2 v_h \|^2_{L^2(\Omega)} + \frac{\kappa^{-1}}{2} \| \Delta v_h \|^2_{L^2(\Omega)} \right]$$

$$= \kappa^{-1} \sum_{K \in \mathcal{T}_h} \frac{1 - \kappa^2(1 - \varepsilon)}{2} \| D^2 v_h \|^2_{L^2(\Omega)}$$

$$+ \kappa^{-1} \left( c_* J_h((v_h, \mu), (v_h, \mu)) + \frac{1}{2} \sum_{F \in \mathcal{E}_h^b} \| (\partial_T \Theta + \mathcal{H}_F)^{1/2} \nabla u_h \|^2_{L^2(F)} \right)$$

$$\geq \frac{1 - \kappa^2(1 - \varepsilon)}{2\kappa} \left[ \sum_{K \in \mathcal{T}_h} \| D^2 v_h \|^2_{L^2(\Omega)} + c_* J_h((v_h, \mu), (v_h, \mu)) + \frac{1}{2} \sum_{F \in \mathcal{E}_h^b} \| (\partial_T \Theta + \mathcal{H}_F)^{1/2} \nabla u_h \|^2_{L^2(F)} \right]$$

$$= \frac{1 - \kappa^2(1 - \varepsilon)}{2\kappa} \| (v_h, \mu) \|^2_{h,1},$$

where the penultimate inequality follows from the fact that $1 > (1 - \kappa^2(1 - \varepsilon))/2 > 0$. Thus, we obtain

$$\| (v_h, \mu) \|^2_{h,1} \leq \frac{2\kappa}{1 - \kappa^2(1 - \varepsilon)} A_h((v_h, \mu), (v_h, \mu)). \tag{4.60}$$

By Lemma 4.13, $\| \cdot \|_{h,1}$ is a norm on $M_h$, and so it follows that there exists a unique pair $(u_h, c_h) \in M_h$ such that

$$A_h((u_h, c_h), (v_h, \mu)) = \sum_{K \in \mathcal{T}_h} \langle \gamma f, \Delta v_h \rangle_K \quad \forall (v_h, \mu) \in M_h.$$

25
Finally, taking \((v_h, \mu) = (u_h, c_h)\) in (4.60) gives us:
\[
\| (u_h, c_h) \|_{h, 1} \leq \frac{2\kappa}{1 - \kappa^2 (1 - \varepsilon)} A_h((u_h, c_h), (u_h, c_h))
\]
\[
= \frac{2\kappa}{1 - \kappa^2 (1 - \varepsilon)} \sum_{K \in \mathcal{T}_h} \langle \gamma f, \Delta u_h \rangle_K
\]
\[
\leq \frac{2\kappa \| \gamma \|_{L^\infty(\Omega)}}{1 - \kappa^2 (1 - \varepsilon)} \sum_{K \in \mathcal{T}_h} \| f \|_{L^2(K)} \| \Delta u_h \|_{L^2(K)}
\]
\[
\leq \frac{2\sqrt{2} \kappa \| \gamma \|_{L^\infty(\Omega)}}{1 - \kappa^2 (1 - \varepsilon)} \| f \|_{L^2(K)} \| (u_h, c_h) \|_{h, 1},
\]

note that the factor of \(\sqrt{2}\) comes from the fact that \(\| \Delta u_h \|_{L^2(K)} \leq \sqrt{2} \| D^2 u_h \|_{L^2(K)}\) for \(K \in \mathcal{T}_h\). Dividing through by \(\| (u_h, c_h) \|_{h, 1}\), we obtain (4.59). \(\square\)

5 Error analysis

Herein we will denote \(a \lesssim b\) for \(a, b \in \mathbb{R}\), if there exists a constant \(C > 0\), such that
\[a \leq C b,\]

independent of \(h := \{h_K : K \in \mathcal{T}_h\}\), and \(u\), but otherwise possibly dependent on the polynomial degree, \(p\), the shape-regularity constants of \(\mathcal{T}_h\), \(C_T\), \(\sigma\), \(c\), \(s\), etc.

**Theorem 5.1** Let \(\Omega \subset \mathbb{R}^2\) be a \(C^2\) and piecewise \(C^{m+1}\) domain, \(m \in \mathbb{N}, m \geq 2\), and let \(\beta \in C^1(\partial \Omega; \mathbb{S})\). Assume that
\[\partial T \Theta + \mathcal{H}_{\partial \Theta} > 0 \quad \text{on} \quad \partial \Omega.\]

Furthermore, assume that \(\{\mathcal{T}_h\}_h\) is a regular of order \(m\) family of triangulations on \(\overline{\Omega}\) satisfying Assumption 4.8. Let \((u, c) \in H^2_{\beta, 0}(\Omega) \times \mathbb{R}\) be the unique strong solution of (1.1). Assume that \(u \in H^s(\Omega; \mathcal{T}_h)\) with \(s_K > 5/2\) for all \(K \in \mathcal{T}_h\).

Let \(c_{stab}, c_*, \mu_F, \sigma_F\) be chosen as in Theorem 4.15 and choose \(\eta_F \lesssim 1/\tilde{h}^2_F\), \(\mu_F, \sigma_F \lesssim 1/\tilde{h}_F\), \(F \in \mathcal{E}_h\). Furthermore, for \(F \in \mathcal{E}_h\), let \(\tilde{h}_F^{21\kappa} \lesssim \ell_F\), where \(t_F := \max_{K \in \mathcal{T}_h \mid \partial K \cap \partial F \neq \emptyset} t_K\), and \(t_K := \min\{p + 1, s_K, m + 1\}. \)

Then, there exists a constant \(C > 0\), independent of \(h\), and \(u\), but depending on \(\max_k s_K\), such that
\[
\| (u - u_h, c - c_h) \|_{h, 1} \leq C \left( \sum_{K \in \mathcal{T}_h} h^{21\kappa - 4}_K \| u \|_{H^{s, \kappa}(K)}^2 \right)^{1/2} \right) + \left( \sum_{e \in \mathcal{E}_h^e \mid [c_{ch}]_e \neq 0} h^{21\kappa - 4}_K \| u \|_{H^{s, \kappa}(K_e)}^2 \right)^{1/2},
\]

where, for a given \(e \in \mathcal{E}_h^e\) such that \([c_{ch}]_e \neq 0\), \(K_e \in \mathcal{T}_h\) has \(e\) as a vertex. Note that for the special case of quasi-uniform meshes, denoting \(s := \min_k s_K\), the a priori error bound (5.7) simplifies to
\[
\| (u - u_h, c - c_h) \|_{h, 1} \leq C h^{\min(p + 1, s, m) - 2} \| u \|_{H^s(\Omega)}.\]

**Proof:** Let us take \(z_h \in V_{h,p}\), and denote by \(\psi_h := z_h - u_h\), \(\xi_h := z_h - u\), and \(\mu_h := c - c_h\). Then, we see that
\[
\| (u - u_h, c - c_h) \|_{h, 1} = \| (\xi_h + \psi_h, \mu_h) \|_{h, 1} \leq \| (\xi_h, 0) \|_{h, 1} + \| (\psi_h, \mu_h) \|_{h, 1}.
\]

The proof we present relies on the existence of a \(z_h \in V_{h,p}\) and a constant \(C\), independent of \(u, h_K\), but dependent on \(\max_k s_K\), such that for each \(K \in \mathcal{T}_h\), each nonnegative integer \(q \leq \min\{s_K, m\},\)
We see that $|\alpha| < s_K - 1/2$, we have
\[
\|u - z_h\|_{H^s(K)} \lesssim h^{s_K - q}\|u\|_{H^{s_K}(K)}, \quad \|D^\alpha (u - z_h)\|_{L^2(\partial K)} \lesssim C h^{s_K - |\alpha| - 1/2}\|u\|_{H^{s_K}(K)}. \tag{5.3}
\]

The error estimates given by the first inequality in \((5.3)\) is given in \([1]\) in the context of meshes consisting of simplices that do not have curved faces. These results, however, still hold when elements of the mesh are curved. First one must note that the second inequality in \((5.3)\) follows from the trace inequality, followed by an application of the first inequality in \((5.3)\). Furthermore, in \([6]\), the first bound in \((5.3)\) is derived (see Corollary 4.1 in \([6]\)) for integer values of $s_K$. However, for non-integer values of $s_K$, the estimate can be proven via scaling.

Due to our assumptions upon the parameters $\mu_F, \eta_F, \ell_F$ and $\sigma_F$, by applying the estimates in \((5.3)\), we obtain
\[
\|\xi_h,0\|_{h,1} \lesssim \left( \sum_{K \in \mathcal{T}_h} h^{2s_K - 4}\|u\|_{H^{s_K}(K)}^2 \right)^{1/2},
\]
thus, by \((5.2)\), it is sufficient to obtain the following estimate:
\[
\|(\psi_h, \mu_h)\|_{h,1} \lesssim \left( \sum_{K \in \mathcal{T}_h} h^{2s_K - 4}\|u\|_{H^{s_K}(K)}^2 \right)^{1/2} + \left( \sum_{c \in \mathcal{T}_h, [c_h] \neq 0} h^{2s_K - 4}\|u\|_{H^{s_K}(K_c)}^2 \right)^{1/2}. \tag{5.4}
\]

Now, applying the coercivity result from Theorem 4.13, we obtain
\[
\|(\psi_h, \mu_h)\|_{h,1}^2 \lesssim A_h((\psi_h, \mu_h), (\psi_h, \mu_h)) = A_h((z_h - u_h, c - c_h), (\psi_h, \mu_h)) = A_h((z_h, c), (\psi_h, \mu_h)) - A_h((u_h, c_h), (\psi_h, \mu_h)). \tag{5.5}
\]

We then utilise the consistency identity 4.38, noting that $c$ is constant, and the fact that the pair $(u_h, c_h) \in M_h$ satisfies (1.47), yielding
\[
A_h((u_h, c_h), (\psi_h, \mu_h)) = \sum_{K \in \mathcal{T}_h} \langle D^2 \xi_h, D^2 \psi_h \rangle_K = A_h((u_h, c), (\psi_h, c)) = A_h((u_h, c), (\psi_h, c) - c_h) + A_h((u_h, c), (0, c_h)) = A_h((u_h, c), (\psi_h, \mu_h)) + A_h((u_h, c), (0, c_h)).
\]

We apply the above identity to \((5.5)\), which results in
\[
\|(\psi_h, \mu_h)\|_{h,1}^2 \lesssim A_h((z_h, c), (\psi_h, \mu_h)) - A_h((u_h, c), (\psi_h, \mu_h)) - A_h((u_h, c), (0, c_h)) = A_h((\xi_h, c), (\psi_h, \mu_h)) - A_h((u_h, c), (0, c_h)).
\]

From this, we obtain $\|(\psi_h, \mu_h)\|_{h,1} \lesssim \sum_{i=1}^6 A_i$, where
\[
A_1 := \sum_{K \in \mathcal{T}_h} \langle D^2 \xi_h, D^2 \psi_h \rangle_K, \quad A_2 := \sum_{K \in \mathcal{T}_h} \langle (\gamma L - \Delta)\xi_h, \Delta \psi_h \rangle_K, \quad A_3 := \sum_{K \in \mathcal{T}_h} \frac{1}{2}\langle \Delta \xi_h, \Delta \psi_h \rangle_K, \quad A_4 := \frac{1}{2} B_{\xi_h,0}(\xi_h, (\psi_h, \mu_h)), \quad A_5 := J_h((\xi_h, 0), (\psi_h, \mu_h)), \quad A_6 := -B_{\xi_h,1/2}(u_h, (0, c_h)).
\]

We see that
\[
|A_1|, |A_2|, |A_3| \lesssim \left( \sum_{K \in \mathcal{T}_h} \|D^2 \xi_h\|_{L^2(K)}^2 \right)^{1/2} \|(\psi_h, \mu_h)\|_{h,1}, \quad \text{and}
\]
\[
|A_5| \leq J_h((\xi_h, 0), (\xi_h, 0))^{1/2}\|(\psi_h, \mu_h)\|_{h,1}.
\]
Applying the first estimate in \((5.3)\) to the estimates in \((5.6)\), we obtain

\[
|A_1|, |A_2|, |A_3| \lesssim \left( \sum_{K \in \mathcal{K}_h} h_K^{2^{1-k}} \|u\|_{H^{2+k}(K)}^2 \right)^{1/2} \|\langle \psi_h, \mu_h \rangle\|_{h,1}.
\]

We also see that \(|A_5| \lesssim (e_1 + e_2 + e_3)^{1/2} \|\psi_h\|_{h,1}\), where, based on the assumption that \(\eta_F \lesssim 1/h_F\), and \(\sigma_F \lesssim 1/h_F\),

\[
e_1 := \sum_{F \in \mathcal{E}_h} \mu_F \left( \|\nabla \xi_h \cdot n_F\|_{L^2(F)}^2 + \|\nabla_T \xi_h\|_{L^2(F)}^2 \right) \lesssim \sum_{F \in \mathcal{E}_h} \frac{1}{h_F} \|\nabla \xi_h\|_{L^2(F)}^2,
\]

\[
e_2 := \sum_{F \in \mathcal{E}_h} \eta_F \|\nabla \xi_h\|_{L^2(F)}^2 \lesssim \sum_{F \in \mathcal{E}_h} \frac{1}{h_F} \|\nabla \xi_h\|_{L^2(F)}^2,
\]

\[
e_3 := \sum_{F \in \mathcal{E}_h} \sigma_F \|\xi_h\|_{L^2(F)}^2 \lesssim \sum_{F \in \mathcal{E}_h} \frac{1}{h_F} \|\nabla \xi_h\|_{L^2(F)}^2,
\]

which are all bounded above by \(C \sum_{K \in \mathcal{K}_h} h_K^{2^{1-k}} \|u\|_{H^{2+k}(K)}^2\), due to \((5.3)\). Now we must obtain a bound for \(A_4 = (1/2)B_{h,*}(\langle \xi_h, 0 \rangle, \langle \psi_h, \mu_h \rangle)\). One can see that \(B_{h,*}(\langle \xi_h, 0 \rangle, \langle \psi_h, \mu_h \rangle) = \sum_{j=1}^6 I_j\); for which

\[
|I_1| := \left| \sum_{K \in \mathcal{K}_h} (D^2 \xi_h, D^2 \psi_h) \right| \lesssim \left( \sum_{K \in \mathcal{K}_h} \|D^2 \xi_h\|_{L^2(K)}^2 \right)^{1/2} \left( \sum_{K \in \mathcal{K}_h} \|D^2 \psi_h\|_{L^2(K)}^2 \right)^{1/2},
\]

\[
|I_2| := \left| \sum_{F \in \mathcal{E}_h} \left( \langle \text{div}_T \nabla \xi_h, \nabla \psi_h \rangle, \nabla \psi_h \cdot n_F \rangle \right)_F + \langle \text{div}_T \nabla \xi_h, [\nabla \psi_h]_F \rangle \right| \lesssim \left( \sum_{F \in \mathcal{E}_h} \frac{1}{h_F} \|\nabla \xi_h\|_{L^2(F)}^2 \right)^{1/2} \left( \sum_{F \in \mathcal{E}_h} \frac{1}{h_F} \|\nabla \psi_h \cdot n_F\|_{L^2(F)}^2 \right)^{1/2} + \left( \sum_{F \in \mathcal{E}_h} \frac{1}{h_F} \|\nabla \xi_h\|_{L^2(F)}^2 \right)^{1/2} \left( \sum_{F \in \mathcal{E}_h} \frac{1}{h_F} \|\text{div}_T \nabla \xi_h\|_{L^2(F)}^2 \right)^{1/2} \|\psi_h, \mu_h\|_{h,1},
\]

\[
|I_3| := \left| - \sum_{F \in \mathcal{E}_h} \left( \langle \nabla_T \xi_h, \nabla_T \psi_h \rangle, \nabla_T \psi_h \cdot n_F \rangle \right)_F + \langle \nabla_T \xi_h, [\nabla_T \psi_h]_F \rangle \right| \lesssim \left( \sum_{F \in \mathcal{E}_h} \frac{1}{h_F} \|\nabla \xi_h\|_{L^2(F)}^2 \right)^{1/2} \left( \sum_{F \in \mathcal{E}_h} \frac{1}{h_F} \|\nabla \psi_h\|_{L^2(F)}^2 \right)^{1/2} \|\psi_h, \mu_h\|_{h,1},
\]

28
where, for a given $\ell$ note that obtaining the final inequality in the estimate for $I_5$ is analogous to \((5.56)\).

Applying both estimates from (5.3) to the estimates for $I_1, \ldots, I_6$, we obtain

$$|A_4| \leq \sum_{i=1}^{6} |I_i| \lesssim \left( \sum_{K \in \mathcal{F}_h} h_{K}^{2}\| u \|^2_{H^{s} (K)} \right)^{1/2} \| (\psi_h, \mu_h) \|_{h, 1}.$$  \hfill (5.7)

It then follows that

$$\| (\psi_h, \mu_h) \|^2_{h, 1} \leq \sum_{i=1}^{6} |A_i| \lesssim \left( \sum_{K \in \mathcal{F}_h} h_{K}^{2}\| u \|^2_{H^{s} (K)} \right)^{1/2} \| (\psi_h, \mu_h) \|_{h, 1} + A_6.$$  \hfill (5.8)

Based upon our assumptions upon $\ell_F$, and Sobolev embeddings, it follows that

$$A_6 \lesssim \left( \sum_{e \in \mathcal{F}_h : |c_e| = 0} (h_e \ell_F)^{-1} |\nabla u(e)|^2 \right)^{1/2} \| (\psi_h, \mu_h) \|_{h, 1}$$

and

$$\lesssim \left( \sum_{e \in \mathcal{F}_h : |c_e| = 0} h_{K_e}^{2\| u \|^2_{H^{s} (K_e)}} \right)^{1/2} \| (\psi_h, \mu_h) \|_{h, 1},$$

where, for a given $e \in \mathcal{F}_h$ such that $|c_e| = 0$, $K_e \in \mathcal{F}_h$ has $e$ as a vertex, and a face $F \in \mathcal{F}_h$ satisfies $F \subset \partial K_e$, and $h_e = h_F$ (recall the definition (4.6) of $h_e$).
The first inequality of (5.8) holds due to the following argument. We see that

\[ A_6 = -B_{h,1/2}((u,c),(0,c_h)) \]

\[ = \sum_{F \in \mathcal{E}_h^b} \left[ \frac{1}{2} \langle \partial_T(\beta^\perp \cdot \nabla u), c_h \rangle_F + \sigma_F \langle \beta \cdot \nabla u - c, c_h \rangle_F \right] + \sum_{F \in \mathcal{E}_h^c} \ell_F \langle [c], [c_h] \rangle_F \]

\[ = \frac{1}{2} \sum_{F \in \mathcal{E}_h^b} \langle \partial_T(\beta^\perp \cdot \nabla u), c_h \rangle_F , \]

where the final equality holds due to the fact that \( \beta \cdot \nabla u - c|_F = 0 \) for all \( F \in \mathcal{E}_h^b \), and as \( c \) is constant, it cannot jump across internal edges. Upon integrating by parts on each \( F \in \mathcal{E}_h^b \), we see that

\[ A_6 = \frac{1}{2} \sum_{F \in \mathcal{E}_h^b} \langle \partial_T(\beta^\perp \cdot \nabla u), c_h \rangle_F \]

\[ = \frac{1}{2} \sum_{e \in \mathcal{E}_h^b} \langle (\beta^\perp \cdot \nabla u) c_h \rangle_e \]

\[ = \frac{1}{2} \sum_{e \in \mathcal{E}_h^b} \langle \beta^\perp \cdot \nabla u \rangle_e [c_h]_e + \langle \langle \beta^\perp \cdot \nabla u \rangle]_e [c_h]_e \]

\[ = \frac{1}{2} \sum_{e \in \mathcal{E}_h^b} \langle \beta^\perp \cdot \nabla u \rangle_e [c_h]_e \]

\[ = \frac{1}{2} \sum_{e \in \mathcal{E}_h^b} \langle [\beta^\perp \cdot \nabla u]_e [c_h]_e \].

The penultimate equality holds, due to the fact that \( \beta^\perp \in C^1(\partial \Omega, S^1) \), and so cannot jump across vertices, furthermore, \( \nabla u \in H^{1/2}(\partial \Omega) \), and thus, since \( \partial \Omega \) is a one-dimensional hypersurface, neither can \( \nabla u \).

For a given \( e \in \mathcal{E}_h^b \), we have two cases to consider: either \( e \cap \mathcal{E}_h^b = \emptyset \) or \( e \cap \mathcal{E}_h^b \neq \emptyset \). In the first case, \( e = F \cap F' \), for some \( F, F' \in \mathcal{E}_h^b \), where \( F \) and \( F' \) are faces of one \( K \in \mathcal{T}_h \). Since \( c_h \) is piecewise constant, it must be constant on \( K \). Thus, \( [c_h]|_F = [c_h]|_{F'} \), and so, \( [c_h]_e = 0 \). This tells us that if \( e \in \mathcal{E}_h^b \) such that \( [c_h]_e \neq 0 \), then \( e \cap \mathcal{E}_h^b = \emptyset \).

When \( e \cap \mathcal{E}_h^b \neq \emptyset \), we see that, \( e = \bigcap_{m=1}^{N_e} F_m \), for some finite collection \( F_1, \ldots, F_{N_e} \in \mathcal{E}_h^b \) such that \( F_m \cap \partial \Omega \neq \emptyset \) for each \( m = 1, \ldots, N_e \) (that is, \( e \) may be expressed as the intersection of several faces that have a nonempty intersection with the boundary of \( \Omega \)). Furthermore, ordering the collection of faces (without relabelling them) with respect to an anticlockwise orientation, each \( F_i \) in this collection satisfies

\[ F_m = \overline{K_m} \cap \overline{K_{m+1}}, \quad m = 1, \ldots, N_e, \]

for some \( K_m, K_{m+1} \in \mathcal{T}_h \) that have a nonempty intersection with \( \partial \Omega \). We thus see that

\[ [c_h]_e = c_h|_{K_{N_e+1}} - c_h|_{K_1} \]

\[ = \sum_{m=1}^{N_e} c_h|_{K_{m+1}} - c_h|_{K_m} \]

\[ \leq \sum_{m=1}^{N_e} |c_h|_{K_{m+1}} - c_h|_{K_m} | \]

\[ = \sum_{m=1}^{N_e} |[c_h]|_{F_m} |. \]
which yields,
\[
|\langle c_h \rangle_e|^2 \leq \left( \sum_{m=1}^{N_e} |\langle c_h \rangle_{E_m}| \right)^2
\]
\[
= \sum_{m,n=1}^{N_e} |\langle c_h \rangle_{E_m}| |\langle c_h \rangle_{F_n}| \leq \frac{1}{2} \sum_{m,n=1}^{N_e} |\langle c_h \rangle_{E_m}|^2 + |\langle c_h \rangle_{F_n}|^2
\]
\[
= N_e \sum_{m=1}^{N_e} |\langle c_h \rangle_{E_m}|^2
\]
\[
= N_e \sum_{F \in \mathcal{E}_h^i : F \cap e \neq \emptyset} |\langle c_h \rangle_{F}|^2.
\]

Thus, applying the Cauchy–Schwarz inequality for vectors in $\mathbb{R}^n$, we obtain
\[
A_6 = \frac{1}{2} \sum_{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0} \langle \beta^+ \cdot \nabla u \rangle_{e \langle c_h \rangle_e}
\]
\[
\leq \frac{1}{2} \left( \sum_{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0} \alpha_e^{-1} |\langle \beta^+ \cdot \nabla u \rangle_{e \langle c_h \rangle_e}|^2 \right)^{1/2} \left( \sum_{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0} \alpha_e |\langle c_h \rangle_{e \langle c_h \rangle_e}|^2 \right)^{1/2}
\]
\[
\leq \frac{1}{2} \left( \sum_{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0} \alpha_e^{-1} |\langle \beta^+ \cdot \nabla u \rangle_{e \langle c_h \rangle_e}|^2 \right)^{1/2} \left( \sum_{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0} \alpha_e N_e \left( \sum_{F \in \mathcal{E}_h^i : F \cap e \neq \emptyset} |\langle c_h \rangle_{F}|^2 \right) \right)^{1/2}
\]
for any collection of positive real numbers $\{\alpha_e\}_{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0}$. Let us denote
\[
r_h := \frac{1}{2} \left( \sum_{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0} \alpha_e^{-1} |\langle \beta^+ \cdot \nabla u \rangle_{e \langle c_h \rangle_e}|^2 \right)^{1/2}.
\]

Note that the value $\max_{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0} N_e$ is bounded independently of the mesh size, due to the mesh regularity assumptions, and so
\[
A_6 \lesssim r_h \left( \sum_{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0} \alpha_e \left( \sum_{F \in \mathcal{E}_h^i : F \cap e \neq \emptyset} |\langle c_h \rangle_{F}|^2 \right) \right)^{1/2}.
\]

We now choose $\alpha_e$ such that
\[
\alpha_e = \min_{F \in \mathcal{E}_h^i : F \cap e \neq \emptyset} \ell_F |F|.
\]

We also see that if $F \in \mathcal{E}_h^i$, then $F$ can intersect at most one vertex belonging to $\{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0\}$; thus
\[
\bigcup_{e \in \mathcal{Y}^h_i : [c_h]_e \neq 0} \{F \in \mathcal{E}_h^i : F \cap e \neq \emptyset\} \subset \{F \in \mathcal{E}_h^i : F \cap \partial \Omega \neq \emptyset\}.
\]

It then follows that
\[
A_6 \lesssim r_h \left( \sum_{F \in \mathcal{E}_h^i : F \cap \partial \Omega \neq \emptyset} \ell_F |F| |\langle c_h \rangle_{F}|^2 \right)^{1/2}.
\]
Since \([c_h]_F\) is constant on \(F\), one also has that
\[
|F| [c_h]_F^2 = [c_h]_F^2 \int_F 1 = \int_F [c_h]_F^2 = \|c_h\|_{L^2(F)}^2,
\]
hence,
\[
\sum_{F \in \mathcal{E}_h : F \cap \Omega \neq \emptyset} \ell_F |F| \|c_h\|_F^2 = \left( \sum_{F \in \mathcal{E}_h : F \cap \Omega \neq \emptyset} \ell_F \|c_h\|_{L^2(F)}^2 \right) 
\leq \left( \sum_{F \in \mathcal{E}_h} \ell_F \|c_h\|_{L^2(F)}^2 \right) 
= \left( \sum_{F \in \mathcal{E}_h} \ell_F \|c - c_h\|_{L^2(F)}^2 \right) 
\lesssim \|\psi_h, \mu_h\|_{h,1}^2.
\]
Applying this to (5.9), we obtain
\[
A_0 \lesssim r_h \|\psi_h, \mu_h\|_{h,1}. \tag{5.10}
\]
Overall, we have obtained
\[
\|\psi_h, \mu_h\|_{h,1}^2 \lesssim \left( r_h + \left( \sum_{K \in \mathcal{T}_h} \beta_h^{-4} \|u\|_{H^s(K)}^2 \right) \right)^{1/2} \|\psi_h, \mu_h\|_{h,1}.
\]
Thanks to the mesh condition (4.11), we have that
\[
\alpha^{-1} \lesssim |F|^{-1} \ell_F^{-1}
\]
for any \(F \in \mathcal{E}_h\) such that \(\overline{F} \cap e \neq \emptyset\). Furthermore, \(|F|^{-1} \lesssim h_e^{-1}\). Thus,
\[
r_h = \frac{1}{2} \left( \sum_{e \in \mathcal{E}_h : [\beta],_e \neq 0} \alpha^{-1} \|\beta_\perp \cdot \nabla u\|_e^2 \right)^{1/2} 
\lesssim \left( \sum_{e \in \mathcal{E}_h : [\beta],_e \neq 0} \alpha^{-1} |\nabla u(e)|^2 \right)^{1/2} \tag{5.11}
\]
applying (5.11) to (5.10), we obtain
\[
A_0 \lesssim \left( \sum_{e \in \mathcal{E}_h : [\beta],_e \neq 0} (h_e \ell_F)^{-1} |\nabla u(e)|^2 \right)^{1/2} \|\psi_h, \mu_h\|_{h,1},
\]
which is the first inequality of (5.8). We obtain the final estimate of (5.8), by first noting that for \(e \in \mathcal{E}_h\) such that \([c_h],_e \neq 0\), \(e\) is a vertex of some \(K_e \in \mathcal{T}_h\), and \(u \in H^{s_{K_e}}(K_e)\), with \(s_{K_e} > 5/2\), and so, by Sobolev embeddings and scaling, we obtain
\[
|\nabla u(e)|^2 \lesssim h_e^{-1} \|u\|_{H^{s_{K_e}}(K_e)}^2.
\]

32
Furthermore, we may choose $K_e$ such that there is an $F \in \mathcal{E}_h^i$ that satisfies $F \subset \partial K_e$, and $h_e^{-1} = \tilde{h}_F \leq h_{K_e}^{-1}$, and thus by our assumptions on $\ell_F$, we have $\ell_F^{-1} \leq \tilde{h}_F^{-1} \leq h_{K_e}^{-1}$. Thus, we obtain
\[
\sum_{e \in \mathcal{T}_h^k: |c_e| \neq 0} (h_e \ell_F)^{-1} |\nabla u(e)|^2 \lesssim \sum_{e \in \mathcal{T}_h^k: |c_e| \neq 0} h_e^{-1} \ell_F^{-1} h_{K_e}^2 \|u\|_{H^2(K_e)}^2 \lesssim \sum_{e \in \mathcal{T}_h^k: |c_e| \neq 0} 2^{2K_e} \|u\|_{H^2(K_e)}^2.
\]

Combining our estimates for $|A_1|, \ldots, |A_6|$, and $A_6$, yields
\[
\| (\psi_h, \mu_h) \|_{h,1}^2 \lesssim \left( \sum_{K \in \mathcal{T}_h} h_{2K}^{-2} \|u\|_{H^2(K)}^2 \right)^{1/2} + \left( \sum_{c \in \mathcal{T}_h^k: |c| \neq 0} h_{2K}^{-2} \|u\|_{H^2(K)}^2 \right)^{1/2} \| (\psi_h, \mu_h) \|_{h,1}.
\]

Finally, upon noting that $\| (\xi_h, 0) \|_{h,1} \lesssim \left( \sum_{K \in \mathcal{T}_h} h_{2K}^{-2} \|u\|_{H^2(K)}^2 \right)^{1/2}$, we obtain
\[
\| (u - u_h, c - c_h) \|_{h,1} \lesssim \| (\xi_h, 0) \|_{h,1} + \| (\psi_h, \mu_h) \|_{h,1}
\lesssim \left( \sum_{K \in \mathcal{T}_h} h_{2K}^{-2} \|u\|_{H^2(K)}^2 \right)^{1/2} + \sum_{e \in \mathcal{T}_h^k: |c_e| \neq 0} h_{2K}^{-2} \|u\|_{H^2(K_e)}^2,
\]
as desired. \( \square \)

5.1 An error estimate in the case of conformal regularity

The hypotheses of Theorem 5.1 includes the sufficient condition that the strong solution, $u$, is piecewise-sufficiently regular, so that one may substitute $(u, c)$ into the left hand argument of the operator, $A_6$. In the following lemma, we provide an error estimate for strong solutions $u \in H^2_{\sigma_0}$, i.e., the expected conformal regularity of strong solutions implied by Theorem 5.1. As in estimate (5.1), one can see the error contribution arising from the inconsistency of $c_h$ belonging to $V_{h,0}$ as opposed to $\mathbb{R}$. Similarly, this contribution is zero if $c_h$ does not jump across boundary vertices. This shows that our method provides an approximation that is at least as accurate in the $\| \cdot \|_{h,1}$-norm, as a $H^2$-conforming finite element method.

Lemma 5.2 Let $\Omega \subset \mathbb{R}^d$ be a $C^2$ and piecewise $C^3$ domain, and let $\beta \in C^3(\partial \Omega; \mathbb{S})$. Assume that $\partial_T \Theta + \mathcal{H}_{\partial \Omega} > 0$ on $\partial \Omega$.

Furthermore, assume that $\{ \mathcal{T}_h \}$ is a regular of order 2 family of triangulations on $\mathbb{P}$ satisfying Assumption 4.3. Let $(u, c) \in H^2_{\sigma_0}(\Omega) \times \mathbb{R}$ be the unique strong solution of (4.7). Let $c_{stab}$, $c_s$, and $\mu_F$ be chosen as in Theorem 4.12 and choose $\eta_F \leq 1/\tilde{h}_F^3$, $\sigma_F \leq 1/\tilde{h}_F$, and $\tilde{h}_F^{1+r+p'} \lesssim \ell_F$, where $r > 0$, and $p^* := 2 \text{sgn}(p - 2)$. Then, we have the following error estimate
\[
\| (u - u_h, c - c_h) \|_{h,1} \lesssim \inf_{z_h \in V} \left\{ \| u - z_h \|_{H^2(\Omega)} + \left[ \sum_{F \in \mathcal{E}_h^k} \frac{1}{\tilde{h}_F} \| \partial_T (\beta \cdot \nabla (u - z_h)) \|_{L^2(F)}^2 \right]^{1/2} + \left( \sum_{F \in \mathcal{E}_h^k} \frac{1}{\tilde{h}_F^{1+r+p'}} \| u \|_{H^2(K_e)}^2 \right)^{1/2} \right\},
\]
where $V := V_{h,p,0} \cap H^2(\Omega)$.

33
Furthermore, we have the following bounds:

\[ \| (\psi, \mu_h) \|_{h,1} \lesssim A_h((z_h, c, (\psi, \mu_h)) - A_h((u_h, c_h, (\psi, \mu_h))) \]

Proof: First we assume that \( z_h \in H^2(\Omega) \cap V_{h,p,0} \). Then, we see that

\[ \|(u - u_h, c - c_h)\|_{h,1} \leq \| (\xi_h, 0) \|_{h,1} + \| (\psi_h, \mu_h) \|, \]

where \( \xi_h, \psi_h \) and \( \mu_h \) are given as in the proof of Theorem 5.1. Since we only assume that \( z_h \) is in \( H^2(\Omega) \cap V_{h,p,0} \), only the consistency properties of the bilinear form \( A_h \) that depend on the piecewise regularity and \( H^2 \)-regularity of \( z_h \) hold. In particular \( (1.38) \) does not hold. This results in:

\[
\begin{align*}
\| (\psi, \mu_h) \|_{h,1} & \lesssim A_h((z_h, c, (\psi, \mu_h)) - A_h((u_h, c_h, (\psi, \mu_h))) \\
& = \sum_{K \in \mathcal{T}_h} \langle (L z_h - f), \Delta \psi_h \rangle_K \\
& \quad + \frac{1}{2} \sum_{F \in \mathcal{E}_h} \langle \partial_T (\beta \cdot \nabla z_h), c_h \rangle_F - \langle \partial_T (\beta \cdot \nabla \psi_h), \beta \cdot \nabla z_h - c \rangle_F \\
& \quad + \frac{1}{2} \sum_{F \in \mathcal{E}_h} \langle \partial_T (\beta \cdot \nabla z_h), \beta \cdot \nabla \psi_h \rangle_F + \langle \partial_T (\beta \cdot \nabla \psi_h), \beta \cdot \nabla z_h - c \rangle_F \\
& \quad + \sum_{F \in \mathcal{E}_h} \sigma_F \langle \beta \cdot \nabla z_h - c, \beta \cdot \nabla \psi_h - \mu_h \rangle_F \\
& \lesssim \left( |z_h - u|_{H^2(\Omega)} + \left( \sum_{e \in \mathcal{E}_h} h_e^{-1} \| z_h \|_{H^2(K_e)} \right)^{1/2} \right) \| (\psi, \mu_h) \|_{h,1} \\
& \quad + \frac{1}{2} \sum_{F \in \mathcal{E}_h} \langle \partial_T (\beta \cdot \nabla z_h - c), \beta \cdot \nabla \psi_h \rangle_F + \langle \partial_T (\beta \cdot \nabla \psi_h), \beta \cdot \nabla z_h - c \rangle_F \\
& \quad + \sum_{F \in \mathcal{E}_h} \sigma_F \langle \beta \cdot \nabla z_h - c, \beta \cdot \nabla \psi_h - \mu_h \rangle_F,
\end{align*}
\]

where the bound for the term \( B_{h,1/2}((z_h, c), (0, c_h)) \) is obtained analogously to the first estimate of \( (5.8) \) in the proof of Theorem 5.1 (except we utilise the fact that \( z_h \) is piecewise \( H^1 \)-regular), and we recall that \( K_e \in \mathcal{T}_h \) has \( e \) as a vertex, and \( F \in \mathcal{E}_h \) is an edge of \( K_e \) that satisfies \( h_e = h_F \).

Furthermore, we have the following bounds:

\[
\begin{align*}
\sum_{F \in \mathcal{E}_h} \sigma_F \langle \beta \cdot \nabla z_h - c, \beta \cdot \nabla \psi_h - \mu_h \rangle_F & \lesssim \left( \sum_{F \in \mathcal{E}_h} \sigma_F \| \beta \cdot \nabla z_h - c \|_{L^2(F)}^2 \right)^{1/2} \| (\psi, \mu_h) \|_{h,1}, \\
\sum_{F \in \mathcal{E}_h} \langle \partial_T (\beta \cdot \nabla z_h - c), \beta \cdot \nabla \psi_h - \mu_h \rangle_F & \lesssim \left( \sum_{F \in \mathcal{E}_h} \sigma_F \| \beta \cdot \nabla z_h - c \|_{L^2(F)}^2 \right)^{1/2} \left( \sum_{F \in \mathcal{E}_h} \| \nabla \psi_h \|_{L^2(F)}^2 + h_F \| D^2 \psi_h \|_{L^2(F)}^2 \right)^{1/2} \| (\psi, \mu_h) \|_{h,1},
\end{align*}
\]

Note that the first bound follows directly from an application of the Cauchy–Schwarz inequality, and the definition of the \( \| (\cdot, \cdot) \|_{h,1} \)-norm, and the second inequality is obtained analogously to \( (1.56) \) in...
the proof of Lemma 4.14. We also see that
\[
\sum_{F \in \mathcal{E}_h^+} (\partial_T (\beta \cdot \nabla z_h - c), \beta \cdot \nabla \psi_h)_F \lesssim \left( \sum_{F \in \mathcal{E}_h^+} \frac{1}{h_F} \|\partial_T (\beta \cdot \nabla z_h - c)\|_{L^2(F)}^2 \right)^{1/2} \|\psi_h, \mu_h\|_{h,1}. \tag{5.13}
\]
Since \(\sigma_F \lesssim 1/h_F\), overall, we have obtained
\[
\|\psi_h, \mu_h\|_{h,1} \lesssim \left( \left| z_h - u \right|_{H^2(\Omega)} + \left( \sum_{F \in \mathcal{E}_h^+} \frac{1}{h_F} \|\beta \cdot \nabla z_h - c\|_{L^2(F)}^2 \right)^{1/2} \right) + \left( \sum_{e \in \mathcal{T}_h} \frac{\ell_e^{-1}}{h_{K_e}} \|z_h\|_{H^2(K_e)}^2 \right)^{1/2} \|\psi_h, \mu_h\|_{h,1}. \tag{5.14}
\]
Finally, by Lemma 2.3 of [R], if \(p = 2\), we have that \(\|z_h\|_{H^2(K)} \lesssim \|z_h\|_{H^2(\Omega)} \) for all \(K \in \mathcal{T}_h\), and if \(p \geq 3\), \(\|z_h\|_{H^2(K)} \lesssim h_{K}^{1/2} \|z_h\|_{H^2(\Omega)} \). i.e., for any \(p \geq 2\), \(\|z_h\|_{H^2(K)} \lesssim h_{K}^{1/p^*} \|z_h\|_{H^2(\Omega)} \), where \(p^* := 2 \text{ sgn}(p - 2)\). Furthermore, by our assumptions on \(\ell_F\), it follows that \(\ell_F^{-1}/h_{K}^{p^*+1} \lesssim 1\), thus
\[
\frac{\ell_F^{-1}}{h_{K}} \|z_h\|_{H^2(K)}^2 \lesssim \frac{\ell_F^{-1}}{h_{K}^{p^*+1}} \|z_h\|_{H^2(\Omega)}^2 \lesssim \frac{\ell_F^{-1}}{h_{K}^{p^*+1}} \|u - z_h\|_{H^2(\Omega)}^2 + \|u\|_{H^2(\Omega)}^2 \tag{5.15}
\]
Applying (5.15) to (5.14), dividing through by \(\|\psi_h, \mu_h\|_{h,1}\), and recalling that \(c = \beta \cdot \nabla u|_{\partial \Omega}\), we obtain
\[
\|\psi_h, \mu_h\|_{h,1} \lesssim \|z_h - u\|_{H^2(\Omega)} + \left( \sum_{F \in \mathcal{E}_h^+} \frac{1}{h_F} \|\beta \cdot \nabla (z_h - u)\|_{L^2(F)}^2 \right)^{1/2} \left( \sum_{e \in \mathcal{T}_h} \frac{\ell_e^{-1}}{h_{K_e}} \|u\|_{H^2(K_e)}^2 \right)^{1/2} \|\psi_h, \mu_h\|_{h,1}. \tag{5.16}
\]
We then see that
\[
\| (\xi_h, 0) \|_{h,1}^2 \lesssim |u - z_h|_{H^2(\Omega)} + \sum_{F \in \mathcal{E}_h^+} \frac{1}{h_F} \|\beta \cdot \nabla (z_h - u)\|_{L^2(F)}^2 + \| (\partial_T \Theta + H_F)^{1/2} \nabla (u - z_h) \|_{L^2(F)}^2 .
\]
Furthermore, the trace operator is continuous from \(H^1(\Omega) \rightarrow L^2(\partial \Omega)\), and so
\[
\sum_{F \in \mathcal{E}_h^+} \| (\partial_T \Theta + H_F)^{1/2} \nabla (u - z_h) \|_{L^2(F)}^2 \lesssim \| \nabla (u - z_h) \|_{L^2(\partial \Omega)}^2 \lesssim \|u - z_h\|_{H^2(\Omega)}^2.
\]
Thus, we obtain
\[
\|u - u_h, c - c_h\|_{h,1} \lesssim \|z_h - u\|_{H^2(\Omega)} + \left( \sum_{F \in \mathcal{E}_h^+} \frac{1}{h_F} \|\beta \cdot \nabla (z_h - u)\|_{L^2(F)}^2 \right)^{1/2} \left( \sum_{e \in \mathcal{T}_h} \frac{\ell_e^{-1}}{h_{K_e}} \|u\|_{H^2(K_e)}^2 \right)^{1/2}.
\]
Note that our choice of \(z_h \in H^2(\Omega) \cap V_{h,p,0}\) was arbitrary, thus we may take an infimum over \(V\) above, yielding estimate (5.12). \(\square\)
Remark 5.3 (Stabilisation parameter choice) Note that our assumption \( F^{-1} \lesssim h^p + r \) implies that \( h^{p+1} F^{1/2} \lesssim h^p \), controlling the contribution of the final term on the right-hand side of (5.12).

Remark 5.4 (Conforming finite element methods) Notice that the error bound (5.12) incorporates the error arising from the approximation of the oblique derivative (i.e., the second and third term in the infimum). Firstly, one must note that if \( \beta \cdot \nabla z_h \big|_F \) is constant, for each \( F \in E_{bh} \), then the third term in the infimum vanishes. This occurs if \( z_h \in V_{c} = V_{h,p,0} \cap H^{2} \beta,0(\Omega) \). However, it is not immediately clear that the space \( V_{c} \) is non-empty for arbitrary \( \beta \in C^1(\partial \Omega; S^1) \) satisfying the hypotheses of Lemma 3.11.

Furthermore, if \( \beta \cdot \nabla z_h \big|_e = c_z \) for some constant \( c_z \), and all \( e \in \gamma^h \) then one may subtract an arbitrary constant on each face in (5.13) (i.e., \( \beta^\perp \cdot \nabla \psi_h \) may be replaced with \( \beta^\perp \cdot \nabla \psi_h - c_F \), for arbitrary constants \( c_F \)), and then applying the Cauchy–Schwarz inequality with a parameter, and Poincaré’s inequality, improves the order of the estimate from \( O(\tilde{h}^{-1} F) \) to \( O(\tilde{h} F) \).

These considerations also imply that the scheme introduced in this paper is at least as accurate as any conforming method seeking a numerical solution \( u_h \in V \), where \( V_{h,p,0} \cap H^{2} \beta,0(\Omega) \subset V \subset V_{h,p,0} \cap H^{2}(\Omega) \).

6 Numerical results

6.1 Implementation

Software and code: The experiments in this section have been implemented in the most recent version of the Firedrake software [36, 29] (as of 3rd July 2018), which interfaces directly with PETSc [2, 3] running through a Python interface [13, 22]. A working Firedrake script, Curved-oblique-DGFEM.py, used to generate the experiments of this Chapter is available in the Github repository: https://github.com/ekawecki/FiredrakeNDV.

Linear systems and condition numbers: The bilinear form \( A_h \) defined by (4.46) can also be considered to be similar to those present in finite element methods for fourth-order elliptic boundary-value problems (see [40, 9] for example), in the sense that the evaluation of \( A_h((u_h, \lambda_h); (v_h, \mu_h)) \) for \( (u_h, \lambda_h), (v_h, \mu_h) \in M_h \) involves the integration of products of second order partial derivatives. This typically leads to the matrix \( A_h \), describing the linear system given by (4.47), to have a Euclidean norm condition number of order \( h^{-4} \). This can pose difficulties when applying iterative methods to solve the linear system, and thus to ensure that we solve the linear system with sufficiently high accuracy as the mesh size \( h \) decreases, we apply the Iterative refinement algorithm, i.e., Algorithm 1.1 of [12]. We implement the Iterative refinement algorithm by using the choices depicted in the code snippet below, in the Firedrake “solve” function.

```python
# implementing nullspace, as solution should have zero sum
V_basis = VectorSpaceBasis(constant=True)
nullspace = MixedVectorSpaceBasis(S, [V_basis, S[1]])

# begin timing of linear system solve
start = time

# solving linear system
solve(A_gamma == L,Uh, nullspace = nullspace, solver_parameters = {"mat_type": "aij", "snes_type": "newtonls", "ksp_type": "preonly", "pc_type": "lu", "snes_monitor": False, "snes_rtol": 1e-16, "snes_atol": 1e-25})

# end timing of linear system solve
end = time

# computing runtimes
runtimes.append(end - start)
```

One can also see that when executing the script in Firedrake, we record the runtimes by way of the sixth and last line above, so that we only record the time that it takes to solve the linear system.
Furthermore, in the solver choices we include “nullspace = nullspace”, where “nullspace” is defined on line 3, and “mat_type = aij”. The first choice imposes that the numerical solution \( u_h \) from the pair \((u_h, c_h) \in M_h \) that satisfies \((1.17)\) has a zero-sum, and the latter informs the solver that the solution consists of two parts, i.e., \( u_h \) and \( c_h \), and that the system may be treated in block formation.

Two-dimensional curved boundary approximation: When implementing curved finite elements, we use a piecewise quadratic polynomial mapping to obtain a higher order approximation of the domain boundary. This is implemented in Firedrake by first using Gmsh [20] (version 3.0.1) to generate an affine triangulation \( \Omega \) that approximates \( \Omega \). We then define the continuous Lagrange finite element space \( V := \{ v \in C(\Omega) : v \in P^2(K ; \mathbb{R}^2) \forall K \in \Omega_h \} \). Then, we take \( \psi_i : \omega_i \to \mathbb{R}^2, \omega_i \subset \mathbb{R}, \) for \( i = 1, \ldots, n \), to be the collection of charts that locally describe \( \partial \Omega \), and denote \( \{x_j\}_{j=1}^N \) to be the degrees of freedom of \( V \). We partition the collection of degrees of freedom by defining \( J_{\text{ext}} = \{ j \in \{1, \ldots, N \} : x_j \in \partial \Omega \}, \) and \( J_{\text{int}} = \{1, \ldots, N \} \setminus J_{\text{ext}} \), and so \( \{x_j\}_{j=1}^N = \{x_j\}_{j \in J_{\text{int}}} \cup \{x_j\}_{j \in J_{\text{ext}}} \). We then define the the function \( T \in V \) by

\[
\begin{align*}
T(x_j) &= x_j, \quad j \in J_{\text{int}}, \\
T(x_j) &= \psi_i(x_j), \quad j \in J_{\text{ext}}, \quad i \in \{1, \ldots, n\} \text{ such that } x_j \in \omega_i.
\end{align*}
\]

Finally, we define our computational finite element space \( \mathcal{V}_{h,p}^{\text{comp}} := \{ v \in L^2(\Omega) : v \circ T^{-1} \in \mathbb{P}(K) \} \). This procedure is implemented in Firedrake, in the code snippet below, utilising the Firedrake “Mesh” function. In this case \( \Omega \) is the unit disk, and so there is only one chart, \( \psi := x/|x| \).

Furthermore, when we refine the mesh in our experiments, the meshes at each refinement level are not related to one another (the one exception being Experiment \(6.2.3\)). That is, there is no hierarchical mesh structure, i.e., at each refinement level, we “remesh”. A collection of the meshes used for the computations of this thesis can be found in the folder “Meshes” in the Github repository: https://github.com/ekawecki/FiredrakeNDV.

```python
# Affine mesh of the unit disk, generated in Gmsh
mesh = Mesh("quasiumifiredisk.msh")
# Implementing quadratic domain approximation
V = FunctionSpace(mesh, "CG", 2)
# Defining a function that identifies the curved portion of the boundary
bdry_indicator = Function(V)
bc = DirichletBC(V, Constant(1.0), 1)
bcry = apply(bdry_indicator)
# Defining the space \( V_{h,p} \)
VV = VectorFunctionSpace(mesh, "CG", 2)
T = Function(VV)
T.interpolate(SpatialCoordinate(mesh))
# Defining the function \( T \) given by (6.1)
T.interpolate(conditional(abs(1-bdry_indicator) < 1e-5, T/sqrt(inner(T,T)), T))
# Defining the curved mesh
mesh = Mesh(T)
# Defining the space \( V_{h,p}^{\text{comp}} \)
FES = FunctionSpace(mesh,"DG",deg)
```

Remark 6.1 (Computational parameters) In the following experiments, we employ the following parameter choices: \( c_{\text{stab}} = 2.5 \), \( \mu_F = 2c_{\text{stab}}(p - 1)^2/\tilde{h}_F \), \( \eta_F = 15(p - 1)^4/16\tilde{h}_F^2 \), \( \sigma_F = 2c_{\text{stab}}\tilde{h}_F^2 \), and \( \ell_F = c_{\text{stab}}\tilde{h}_F^{-3} \). The order of the computational parameters with respect to \( \tilde{h}_F \) were guided by the hypotheses of Theorem 5.1. The orders with respect to \( p \) for \( \eta_F \) and \( \mu_F \) were guided by the experiments of [28]. Finally, the value of \( c_{\text{stab}} \), and the orders with respect to \( p \) of \( \sigma_F \) and \( \ell_F \) (in the case of \( \ell_F \), the parameter is in fact independent of \( p \)) were obtained experimentally. Furthermore, we have that \( H_F = 1 \) on each \( F \in \mathcal{E}_h \), due to the fact that the experiments are on the unit disc \( \Omega := \{ x \in \mathbb{R}^2 : |x| < 1 \} \).
6.2 Experiments

In this section, we test the robustness of the scheme (4.47), with the computational domain $\Omega$ taken to be the unit disk, approximated in the same manner as in present in [24], Section 3.4. We consider various elliptic operators, $L$, that satisfy the Cordes condition (3.1). In each case, we see that the convergence rates are of the expected order in the various broken Sobolev norms considered, and in particular in the $\| \cdot \|_{h,1}$-norm, for which we have proven the error bound (5.1).

6.2.1 Experiment 1

In this experiment, we consider the following problem

$$\begin{cases}
\Delta u = f, & \text{in } \Omega, \\
\beta \cdot \nabla u \text{ is constant on } \partial \Omega,
\end{cases}$$

(6.2)

where $\Omega = \{ x \in \mathbb{R}^2 : |x| < 1 \}$, and $\beta \equiv n_{\partial \Omega}$. In this case $f$ is chosen so that the solution of (6.2) is given by

$$u(x) = \frac{1}{6}|x|^6 - \frac{1}{2}|x|^2 + \frac{5}{24}.$$

Notice that in this case, the compatibility constant $c = 0$, $\gamma = 1$, and $\Theta \equiv \partial \Gamma \equiv 0$.

In this experiment, we successively increase the degree, $p$, of the finite element space $V_{h,p,0}^{\text{comp}}$ from 2 to 4, and for each fixed degree we refine the mesh quasi-uniformly, we observe that the experimental orders of convergence in the $\| \cdot \|_{h,1}$-norm are optimal, that is

$$\| (u - u_h, c - c_h) \|_{h,1} = O(h^{p-1}),$$

where $(e_h^u, e_h^c) := (u - u_h, c - c_h)$. We also observe that $\| e_h^c \|_{L^2(\partial \Omega)} = O(h^{p})$. We plot the error values in the $\| \cdot \|_{h,1}$-norm, and plot the error arising in the approximation of the compatibility constant in Figure 1 and report the exact values in Tables 1 and 2, with the corresponding experimental orders of convergence given in brackets.

![Figure 1](image-url)

Figure 1: Convergence rates for the numerical scheme applied to problem (6.2). We provide the error values $\| (u - u_h, c - c_h) \|_{h,1}$ (left), and $\| c - c_h \|_{L^2(\partial \Omega)}$ (right). We observe that the convergence rates in the $\| \cdot \|_{h,1}$ norm are optimal with respect to the choice of polynomial degree, $p$. That is, $\| (u - u_h, c - c_h) \|_{h,1} = O(h^{p-1})$. Furthermore, we observe that $\| c - c_h \|_{L^2(\partial \Omega)} = O(h^{p})$. 38
where $\Omega = \{ x \in \mathbb{R}^2 : |x| < 1 \}$. We take $\beta$ to be the anti-clockwise rotation of the normal by the angle $\varphi(x_1, x_2) := \pi/4 + \arctan(\frac{x_2}{x_1})$, for $(x_1, x_2) \in \partial \Omega$. Furthermore, the function $f$ on the right-hand side of (6.3) is chosen so that the solution $u$ is given by $u(x_1, x_2) = \frac{1}{4} \cos(\pi(x_1^2 + x_2^2)) - \frac{1}{4} \int_0^1 \frac{1}{2} \cos(\pi(x_1^2 + x_2^2))$. Notice that in this case the compatibility constant $c = 0$, $\gamma = 2/5$, $\varepsilon = 3/5$, $\Theta = \pi/4 + \varphi(x_1, x_2)$, and $\partial T \Theta \equiv 1$. This experiment serves to demonstrate the robustness of this method with respect to the choice of oblique vector, $\beta$, and choice of discontinuous coefficients, $A \in L^\infty(\Omega; \mathbb{R}^{2 \times 2})$. In particular, $\beta$ performs a full rotation around the normal vector, and the coefficients $A_{12}, A_{21}$, are discontinuous across the lines $\{ x \in \Omega : x_1 = 0 \}$ and $\{ x \in \Omega : x_2 = 0 \}$.

In this experiment, we successively increase the degree, $p$, of the finite element space $V^{\text{comp}}_{h,p,0}$ from 2 to 4, and for each fixed degree we refine the mesh quasi-uniformly. In Tables 6.2.1 and 6.2.2 we report the error values in the $\| \cdot \|_{h,1}$-norm and the $| \cdot |_{H^1(\Omega, \mathbb{R}^2)}$-seminorm, respectively, with the corresponding experimental orders of convergence given in brackets. We observe the optimal convergence rates $\| (u - u_h, c - c_h) \|_{h,1} = \mathcal{O}(h^{p-1})$, and $|u - u_h|_{H^1(\Omega, \mathbb{R}^2)} = \mathcal{O}(h^p)$.

### Table 1: Error values in the $\| \cdot \|_{h,1}$-norm and EOCs for Experiment 6.2.1

| Mesh size | $p = 2$ | $p = 3$ | $p = 4$ |
|----------|---------|---------|---------|
| 0.4981   | 2.75    | 1.16    | 3.09 × 10^{-1} |
| 0.2828   | 1.84 (0.70) | 3.61 × 10^{-1} (2.06) | 4.33 × 10^{-2} (3.47) |
| 0.1627   | 1.10 (0.94) | 1.17 × 10^{-1} (2.03) | 7.35 × 10^{-3} (3.21) |
| 0.0973   | 6.00 × 10^{-1} (1.17) | 3.45 × 10^{-2} (2.39) | 1.06 × 10^{-3} (3.76) |
| 0.0508   | 2.93 × 10^{-1} (1.10) | 8.46 × 10^{-3} (2.16) | 1.27 × 10^{-4} (3.27) |
| 0.0269   | 1.47 × 10^{-1} (1.08) | 2.11 × 10^{-3} (2.18) | 1.47 × 10^{-5} (3.38) |
| 0.0138   | 7.24 × 10^{-2} (1.06) | 5.12 × 10^{-4} (2.11) | 1.70 × 10^{-6} (3.22) |

### Table 2: $\| c - c_h \|_{L^2(\partial \Omega)}$ error values and EOCs for Experiment 6.2.1

| Mesh size | $p = 2$ | $p = 3$ | $p = 4$ |
|----------|---------|---------|---------|
| 0.4981   | 1.06 × 10^{-1} | 2.63 × 10^{-2} | 4.82 × 10^{-4} |
| 0.2828   | 7.06 × 10^{-2} (0.72) | 1.09 × 10^{-2} (1.56) | 9.19 × 10^{-5} (6.99) |
| 0.1627   | 3.42 × 10^{-2} (1.31) | 2.83 × 10^{-3} (2.44) | 5.32 × 10^{-6} (5.16) |
| 0.0973   | 1.22 × 10^{-2} (2.01) | 4.44 × 10^{-4} (3.60) | 3.94 × 10^{-7} (5.06) |
| 0.0508   | 3.53 × 10^{-3} (1.91) | 7.48 × 10^{-5} (2.74) | 3.16 × 10^{-7} (0.34) |
| 0.0269   | 9.70 × 10^{-4} (2.03) | 8.79 × 10^{-6} (3.36) | 2.34 × 10^{-8} (4.08) |
| 0.0138   | 2.50 × 10^{-4} (2.02) | 7.72 × 10^{-7} (3.63) | 1.48 × 10^{-9} (4.13) |

### 6.2.2 Experiment 2

In this experiment, we consider the following problem

$$
\sum_{i,j=1}^{2} (1 + \delta_{ij}) \frac{x_i}{|x|} \frac{x_i}{|x|} D^2 u = f, \quad \text{in} \quad \Omega,
\beta \cdot \nabla u \text{ is constant on } \partial \Omega,
$$

where $\Omega = \{ x \in \mathbb{R}^2 : |x| < 1 \}$. We take $\beta$ to be the anti-clockwise rotation of the normal by the angle $\varphi(x_1, x_2) := \pi/4 + \arctan(\frac{x_2}{x_1})$, for $(x_1, x_2) \in \partial \Omega$. Furthermore, the function $f$ on the right-hand side of (6.3) is chosen so that the solution $u$ is given by $u(x_1, x_2) = \frac{1}{4} \cos(\pi(x_1^2 + x_2^2)) - \frac{1}{4} \int_0^1 \frac{1}{2} \cos(\pi(x_1^2 + x_2^2))$. Notice that in this case the compatibility constant $c = 0$, $\gamma = 2/5$, $\varepsilon = 3/5$, $\Theta = \pi/4 + \varphi(x_1, x_2)$, and $\partial T \Theta \equiv 1$. This experiment serves to demonstrate the robustness of this method with respect to the choice of oblique vector, $\beta$, and choice of discontinuous coefficients, $A \in L^\infty(\Omega; \mathbb{R}^{2 \times 2})$. In particular, $\beta$ performs a full rotation around the normal vector, and the coefficients $A_{12}, A_{21}$, are discontinuous across the lines $\{ x \in \Omega : x_1 = 0 \}$ and $\{ x \in \Omega : x_2 = 0 \}$.
In this experiment, we consider problem (6.3), where $\Omega = \{ x \in \mathbb{R}^2 : |x| < 1 \}$, and $\beta$ is a $\pi/4$ anticlockwise rotation of the normal, $n_{\partial \Omega}$. That is $\beta = \frac{1}{\sqrt{2}}(n_{\partial \Omega}^1 - |n_{\partial \Omega}|^2, |n_{\partial \Omega}|^1 + |n_{\partial \Omega}|^2)^T$. In this case, $f$ is chosen so that the solution of (6.3) is given by $u(x) = |x|^{1.5} - 0.75|x|^2 - 0.5 \int_{\Omega} |x|^{1.5} - 0.75|x|^2$. Notice that in this case the compatibility constant $c = 0$, $\gamma = 2/5$, $\varepsilon = 3/5$, $\Theta = \pi/4$, and $\partial \Omega \Theta = 0$.

In this experiment, the true solution $u \in H^2(\Omega)$, and, in particular, $u \in H^{5/2-\delta}(\Omega)$ for arbitrary $\delta > 0$. However, the $H^s$-broken Sobolev regularity of $u$ fails for $s \geq 5/2$, and we must appeal to the conformal regularity estimate of Lemma 6.3. In this experiment we successively increase the degree, $p$, of the finite element space $V_{h,p,0}$ from 2 to 4.

Furthermore, we compute the numerical solution both on sequence of meshes refined towards the origin (where the solution lacks regularity), an example of such a mesh is given in Figure 6.2.3, and on a sequence of quasi-uniformly refined meshes (that in particular does not prioritise refinement towards the origin). We plot the error arising in both cases (adapted mesh refinement and non adapted mesh refinement) in the broken $H^2$-seminorm, against the number of DoFs in Figure 6.2.3. For $p = 2, 3, 4$, we see a reduction in error from the adapted mesh sequence. In particular, for $p = 3$ and $p = 4$ we see a reduction in the order of error from $O(ndofs^{-1/4})$ to $O(ndofs^{-1/2})$.

### 7 Conclusion

We have extended the framework introduced in [38], and [24] allowing for domains with curved boundaries, as well as oblique boundary conditions. In doing so, we have introduced a new DGFEM for elliptic equations in nondivergence form, that satisfy the Cordes condition.

The computational domain we considered was the unit disc; in order to verify the error estimates present in Section 3 we used a mesh consisting of curved triangles with edges defined by polynomial mappings. It would be an interesting avenue for future research to consider oblique boundary-value problems in dimensions three and higher; this would require one to prove the Miranda–Talenti estimates [38] in higher dimensions, which is currently an open problem.
Figure 2: Convergence rates for the numerical scheme applied to problem (6.3), with a true solution of conformal regularity. On the left, we provide the error values in the $\| \cdot \|_{H^2(\Omega)}$ seminorm, where the numerical scheme is implemented on a quasiuniformly refined mesh, and an adapted mesh, with refinement towards the origin. On the right we provide an example of this adapted mesh, at refinement level 7, consisting of 4532 elements.

The finite element approximation of solutions to elliptic problems in non-divergence form with oblique boundary conditions is a challenging problem, and as such appears to be underrepresented in the available literature. This paper provides and analyses a new method, which appears to be the first discontinuous Galerkin finite element method for oblique boundary-value problems; we were successful in proving both a stability estimate (4.58), guaranteeing existence and uniqueness of the numerical solution, and an apriori error estimate (5.1) that is optimal with respect to the polynomial degree.

References

[1] I. Babuška and M. Suri. The $hp$ version of the finite element method with quasiuniform meshes. *RAIRO-Modélisation mathématique et analyse numérique*, 21(2):199–238, 1987.

[2] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, D. A. May, L. C. McInnes, R. T. Mills, T. Munson, K. Rupp, P. Sanan, B. F. Smith, S. Zampini, and H. Zhang. PETSc users manual. Technical Report ANL-95/11 - Revision 3.9, Argonne National Laboratory, 2018.

[3] S. Balay, W. D. Gropp, L. C. McInnes, and B. F. Smith. Efficient management of parallelism in object oriented numerical software libraries. In E. Arge, A. M. Bruaset, and H. P. Langtangen, editors, *Modern Software Tools in Scientific Computing*, pages 163–202. Birkhäuser Press, 1997.

[4] J. W. Barrett and C. M. Elliott. Fixed mesh finite element approximations to a free boundary problem for an elliptic equation with an oblique derivative boundary condition. *Comput. Math. Appl.*, 11(4):335–345, 1985.

[5] J-D. Benamou, B. D. Froese, and A. M. Oberman. A viscosity solution approach to the Monge–Ampère formulation of the Optimal Transportation Problem. *arXiv preprint arXiv:1208.4873*, 2012.

[6] C. Bernardi. Optimal finite-element interpolation on curved domains. *SIAM J. Numer. Anal.*, 26(5):1212–1240, 1989.
[7] S. C. Brenner. Poincaré-Friedrichs inequalities for piecewise $H^1$ functions. *SIAM J. Numer. Anal.*, 41(1):306–324, 2003.

[8] S. C. Brenner and M. Neilan. Finite element approximations of the three dimensional Monge-Ampère equation. *ESAIM Math. Model. Numer. Anal.*, 46(5):979–1001, 2012.

[9] S. C. Brenner and L-Y. Sung. $C^0$ interior penalty methods for fourth order elliptic boundary value problems on polygonal domains. *J. Sci. Comput.*, 22/23:83–118, 2005.

[10] L. A. Caffarelli. Boundary regularity of maps with convex potentials. II. *Ann. of Math. (2)*, 144(3):453–496, 1996.

[11] S. Čanić, B. L. Keyfitz, and G. M. Lieberman. A proof of existence of perturbed steady transonic shocks via a free boundary problem. *Comm. Pure Appl. Math.*, 53(4):484–511, 2000.

[12] E. Carson and N. J. Higham. A new analysis of iterative refinement and its application to accurate solution of ill-conditioned sparse linear systems. *SIAM Journal on Scientific Computing*, 39(6):A2834–A2856, 2017.

[13] L. D. Dalcin, R. R. Paz, P. A. Kler, and A. Cosimo. Parallel distributed computing using Python. *Advances in Water Resources*, 34(9):1124–1139, 2011. New Computational Methods and Software Tools.

[14] G. Devillanova and F. Pugliese. A variant on Miranda-Talenti estimate. *Matematiche (Catania)*, 54(1):91–97 (2000), 1999.

[15] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, second edition, 2010.

[16] Z. Fašková, R. Čunderlík, and K. Mikula. Finite element method for solving geodetic boundary value problems. *Journal of geodesy*, 84(2):135–144, 2010.

[17] X. Feng, R. Glowinski, and M. Neilan. Recent developments in numerical methods for fully nonlinear second order partial differential equations. *SIAM Rev.*, 55(2):205–267, 2013.

[18] W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25 of *Stochastic Modelling and Applied Probability*. Springer, New York, second edition, 2006.

[19] D. Gallistl. Numerical approximation of planar oblique derivative problems in nondivergence form. *Math. Comp.*, 2018.

[20] C. Geuzaine and J. F. Remacle. GMSH: A three-dimensional finite element mesh generator with built-in pre- and post-processing facilities, version 2.2. 4, 2008.

[21] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Classics in Mathematics. Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[22] B. Hendrickson and R. Leland. A multilevel algorithm for partitioning graphs. In *Supercomputing ’95: Proceedings of the 1995 ACM/IEEE Conference on Supercomputing (CDROM)*, page 28, New York, 1995. ACM Press.

[23] M. Jensen and I. Smears. On the convergence of finite element methods for Hamilton-Jacobi-Bellman equations. *SIAM J. Numer. Anal.*, 51(1):137–162, 2013.

[24] E. Kawecki. A DGFEM for Nondivergence Form Elliptic Equations with Cordes Coefficients on Curved Domains. *arXiv preprint arXiv:1709.05028*, 2017.

[25] N. V. Krylov. *Nonlinear elliptic and parabolic equations of the second order*, volume 7 of *Mathematics and its Applications (Soviet Series)*. D. Reidel Publishing Co., Dordrecht, 1987. Translated from the Russian by P. L. Buzytsky [P. L. Buzytskii].
[26] Aimé Lachapelle and Marie-Therese Wolfram. On a mean field game approach modeling congestion and aversion in pedestrian crowds. *Transportation research part B: methodological*, 45(10):1572–1589, 2011.

[27] G. M. Lieberman. Two-dimensional nonlinear boundary value problems for elliptic equations. *Trans. Amer. Math. Soc.*, 300(1):287–295, 1987.

[28] G. M. Lieberman. *Oblique derivative problems for elliptic equations*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2013.

[29] F. Luporini, A. L. Varbanescu, F. Rathgeber, G-T. Bercea, J. Ramanujam, D. A. Ham, and P. H. J. Kelly. Cross-loop optimization of arithmetic intensity for finite element local assembly. *ACM Transactions on Architecture and Code Optimization*, 11(4):57:1–57:25, 2015.

[30] A. Maugeri, D. K. Palagachev, and L. G. Softova. *Elliptic and parabolic equations with discontinuous coefficients*, volume 109 of *Mathematical Research*. Wiley-VCH Verlag Berlin GmbH, Berlin, 2000.

[31] M. Neilan. Finite element methods for fully nonlinear second order PDEs based on a discrete Hessian with applications to the Monge-Ampère equation. *J. Comput. Appl. Math.*, 263:351–369, 2014.

[32] D. K. Palagachev. The Poincaré problem in $L^p$-Sobolev spaces. I. Codimension one degeneracy. *J. Funct. Anal.*, 229(1):121–142, 2005.

[33] A. V. Pogorelov. *Monge-Ampère equations of elliptic type*. Translated from the first Russian edition by Leo F. Boron with the assistance of Albert L. Rabenstein and Richard C. Bollinger. P. Noordhoff, Ltd., Groningen, 1964.

[34] H. Poincaré. Lecons de Mécanique Céleste, Tome III, Théorie de Marées. *Gauthiers–Villars, Paris*, 1910.

[35] Peter R. Popivanov and Dian K. Palagachev. *The degenerate oblique derivative problem for elliptic and parabolic equations*, volume 93 of *Mathematical Research*. Akademie Verlag, Berlin, 1997.

[36] F. Rathgeber, D. A. Ham, L. Mitchell, M. Lange, F. Luporini, A. T. T. McRae, G-T. Bercea, G. R. Markall, and P. H. J. Kelly. Firedrake: automating the finite element method by composing abstractions. *ACM Trans. Math. Softw.*, 43(3):24:1–24:27, 2016.

[37] M. V. Safonov. Nonuniqueness for second-order elliptic equations with measurable coefficients. *SIAM J. Math. Anal.*, 30(4):879–895, 1999.

[38] I. Smears and E. Süli. Discontinuous Galerkin finite element approximation of nondivergence form elliptic equations with Cordès coefficients. *SIAM J. Numer. Anal.*, 51(4):2088–2106, 2013.

[39] I. Smears and E. Süli. Discontinuous Galerkin finite element approximation of Hamilton-Jacobi-Bellman equations with Cordès coefficients. *SIAM J. Numer. Anal.*, 52(2):993–1016, 2014.

[40] E. Süli and I. Mozolevski. $hp$-version interior penalty DG FEMs for the biharmonic equation. *Computer methods in applied mechanics and engineering*, 196(13-16):1851–1863, 2007.

[41] J. Urbas. On the second boundary value problem for equations of Monge-Ampère type. *J. Reine Angew. Math.*, 487:115–124, 1997.

[42] J. Urbas. Oblique boundary value problems for equations of Monge-Ampère type. *Calc. Var. Partial Differential Equations*, 7(1):19–39, 1998.
[43] G. C. Wen and C. J. Yang. Finite element solutions of the oblique derivative problem for elliptic complex equations of second order. *Sichuan Shifan Daxue Xuebao Ziran Kexue Ban*, 17(3):20–28, 1994.

[44] Y. Zheng. A global solution to a two-dimensional Riemann problem involving shocks as free boundaries. *Acta Math. Appl. Sin. Engl. Ser.*, 19(4):559–572, 2003.