DISCRETE MIRANDA–TALENTI ESTIMATES AND APPLICATIONS TO LINEAR AND NONLINEAR PDES

MICHAEL NEILAN AND MOHAN WU

Abstract. In this article, we construct simple and convergent finite element methods for linear and nonlinear elliptic differential equations in non-divergence form with discontinuous coeffients. The methods are motivated by discrete Miranda-Talenti estimates, which relate the H^2 semi-norm of piecewise polynomials with the L^2 norm of its Laplacian on convex domains. We develop a stability and convergence theory of the proposed methods, and back up the theory with numerical experiments.

1. Introduction

In this article, we consider finite element methods for linear and nonlinear second order elliptic problems in non-divergence form. A prototypical (linear) example is given by

$$A(x) : D^2u(x) = g(x) \qquad \text{in } \Omega \subset \mathbb{R}^d, \tag{1.1}$$

accompanied with Dirichlet boundary conditions. Here, $g \in L^2(\Omega)$ is a given source term, and $A : \Omega \to \mathbb{R}^{d\times d}$ is symmetric and uniformly positive definite on a bounded domain Ω. Such problems naturally appear in stochastic optimal control in the form of the Hamilton–Jacobi–Bellman equation, and they also arise as linearizations of fully nonlinear second-order PDEs.

This paper focuses on the case in which the coefficient matrix is not differentiable, in particular, integration-by-parts cannot be performed on (**??**), and weak solutions based on variational principles are not applicable. In this setting, there are several distinct theories and solution concepts concerning the well-posedness of the problem, each depending on the regularity of the matrix A. For example:

- If A is Hölder continuous and if the boundary of the domain is sufficiently smooth, then there exists a classical solution satisfying the PDE pointwise in Ω [**?**, Chapter 6].
- If A is uniformly continuous on Ω or if A has vanishing mean oscillation, then there exists a unique strong solution $u \in W^{2,p}(\Omega)$ to the problem, i.e., u satisfies the PDE a.e. in Ω [**?**, **?**].
- If A is essentially bounded and it satisfies the so-called Cordes condition (cf. Definition **??**), and if the domain Ω is convex, then there exists a strong solution $u \in H^2(\Omega)$ [**?**, **?**].

In this paper we consider the third case which assumes the least regularity conditions on the coefficient matrix, i.e., we assume that the matrix is possibly discontinuous but satisfies the Cordes condition. The Cordes condition, a type of anisotropy condition on A, is a crucial assumption to establish the well-posedness of the elliptic problem (**??**); counterexamples in $d \geq 3$ show that solutions to (**??**) are not unique for general discontinuous A (cf. Example **??**). Another key ingredient to show the well-posedness of (**??**) is the Miranda-Talenti estimate, which relates the H^2 semi-norm of a function with the L^2 norm of its Laplacian on convex domains.

Lemma 1 (Miranda-Talenti inequality [**?**, **?**]). *Suppose that $\Omega \subset \mathbb{R}^d$ is a bounded convex domain. Then there holds*

$$\|D^2v\|_{L^2(\Omega)} \leq \|\Delta v\|_{L^2(\Omega)} \qquad \forall v \in H^2(\Omega) \cap H^1_0(\Omega). \tag{1.2}$$

The important feature of the estimate (**??**), and crucial to the analysis of problem (**??**), is that the equivalence constant is exactly one on convex domains.

This work was supported in part by the National Science Foundation through grants DMS-1417980 and DMS-1719829.

The goal of this paper is to develop finite element methods for PDEs in non-divergence form and a convergence theory by extending Lemma **??** to piecewise polynomial functions. In particular, we shall prove the following discrete Miranda-Talenti inequality. A more detailed explanation of the notation is given in subsequent sections.

Theorem 1 (Discrete Miranda-Talenti inequality)**.** *Let $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ be a convex polytope. Let $V_h \subset H_0^1(\Omega)$ denote the kth degree Lagrange finite element space with respect to a simplicial mesh $\mathcal{T}_h$ with $2 \le k \le 3$ if $d = 3$, and $k \ge 2$ for $d = 2$. Then for any $v_h \in V_h$, we have*

$$\|D^2 v_h\|_{L^2(\mathcal{T}_h)} \le \|\Delta v_h\|_{L^2(\mathcal{T}_h)} + C_\dagger \left(\sum_{f \in \mathcal{F}_h^I} h_f^{-1} \|[\![\partial v_h/\partial n_f]\!]\|_{L^2(f)}^2 \right)^{1/2}, \tag{1.3}$$

where the constant $C_\dagger > 0$ is independent of h and v_h.

We shall show that the estimate (**??**) naturally leads to simple and efficient finite element methods for linear and fully nonlinear problems in non-divergence form as well as a stability and convergence theory.

Despite its non-variational structure, a flurry of finite element methods have recently been developed for problems in non-divergence form (**??**). In the case that the coefficient matrix is continuous, finite element methods have been developed in [**?**, **?**, **?**]; these methods and their analysis are based on discrete Calderon-Zygmund estimates. The first Galerkin method in the case of discontinuous coefficients was done in [**?**], where an intricate *hp*-discontinuous Galerkin (DG) method was proposed for elliptic PDEs satisfying the Cordes condition. There, the authors bypass a discrete Miranda–Talenti estimate by adding auxiliary terms in their formulation. This method was extended to the fully nonlinear Hamilton–Jacobi–Bellman equation with continuous coefficients satisfying the Cordes condition in [**?**, **?**]. Much of the present work is influenced by these results and techniques. A related but simpler DG method for elliptic problems in non-divergence form based on a least-squares formulation is proposed in [**?**]. However, it is unclear whether this method extends to fully nonlinear problems. A weak Galerkin method was presented in [**?**], and a mixed discretization based on stable finite element Stokes spaces is proposed in [**?**]. A finite element method based on the convolution of finite differences for (**??**) was proposed in [**?**], and the stability of the method was shown via discrete Alexandrov-Bakelman-Pucci estimates. Extensions of these results to fully nonlinear problems was done in [**?**].

An advantage of the proposed methods is their relative simplicity; the methods can be readily implemented on standard finite element method software packages. Furthermore, in contrast to [**?**, **?**, **?**], the methods are provably convergent for linear problems with discontinuous coefficients satisfying the Cordes condition. Finally, as far as we are aware, the methods have the fewest number of global degrees of freedom on simplicial meshes for problems with discontinuous coefficients. On the other hand, the cost of the simplicity of the methods includes restrictions on the finite element spaces and the mesh. For example, in contrast to [**?**, **?**, **?**], we require that the mesh is simplicial and does not contain hanging nodes, and that the polynomial degree does not vary between elements. Furthermore, we do not track the dependence of the polynomial degree in the stability and error estimates.

The rest of the paper is organized as follows. In Section **??** we establish the notation and state some preliminary results. In Section **??** we build an enriching operator that connects the Lagrange finite element space with a cubic spline space in two and three dimensions. With this enriching operator, we prove Theorem **??** in Section **??**. In Section **??** we propose a finite element method for linear PDEs in non-divergence form, prove the well-posedness of the method, and derive optimal order estimates. These results are extended to the fully nonlinear Hamilton–Jacobi–Bellman equation in Section **??**. Finally, numerical experiments are presented in Section **??**.

3) be a bounded, convex polytope, and let $\mathcal{T}_h$ be a
gulation of Ω without hanging nodes. For each ele
$_{\in\mathcal{T}_h} h_T$. We denote by $\mathcal{V}_h$ and $\mathcal{F}_h$ the set of vertice
$_h$, respectively. We write $\mathcal{F}_h = \mathcal{F}_h^I \cup \mathcal{F}_h^B$, wher
denotes the set of boundary faces. Likewise, we de
as $\mathcal{V}_h^I$ and $\mathcal{V}_h^B$, respectively. Let $\mathcal{V}_T$ and $\mathcal{M}_T$ be th
points, respectively, of a simplex $T \in \mathcal{T}_h$. Let $\mathcal{T}_p$ be
vertex $p \in \mathcal{V}_h$. We also denote by $\mathcal{F}_p^I$ (resp., $\mathcal{F}_p^B$)
ensional faces in $\mathcal{F}_h^I$ (resp., $\mathcal{F}_h^B$) that share the co
y edge midpoints in $\mathcal{T}_h$ is denoted by $\mathcal{M}_h^I$ and $\mathcal{M}_h^B$

, let $n_f \in \mathbb{R}^d$ denote a fixed choice of a (constant)
that n_f coincides with the outward unit normal o
perator $[\![\cdot]\!]$ on f by

$$[\![v]\!] := v|_{T_{out}} - v|_{T_{in}} \quad \text{if } f = \partial T_{out} \cap \partial T_{in} \in \mathcal{F}_h^I,$$
$$[\![v]\!] := v|_{T_{out}} \quad \text{if } f = \partial T_{out} \cap \partial\Omega \in \mathcal{F}_h^B,$$

regular scalar valued or vector-valued function. He

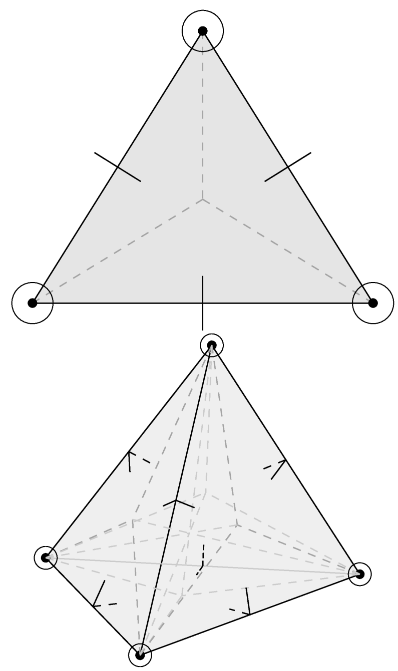

FIGURE 1. The two–dimensional (left) and three–dimensional (right) Clough-Tocher element. Solid circles indicate function evaluation, large circles indicate derivative evaluation, and straight lines indicate directional derivative evaluation.

- For $d = 2$, the simplex is split by connecting the vertices to the barycenter of the simplex (cf. Figure **??** left).
- For $d = 3$, each of the 2–dimensional faces of the simplex is split by connecting the vertices to the barycenter of the face. Then the vertices and the barycenters of the faces are connected to the barycenter of the tetrahedron (cf. Figure **??** right).

Thus we see that each simplex is split into $N_d := (d+1)!/2$ subsimplices. We denote the set of subsimplices of this split by $T_r := \{T_i\}_{i=1}^{N_d}$ and define

$$\mathbb{P}_k(T_r) := \prod_{i=1}^{N_d} \mathbb{P}_k(T_i),$$

to be the space of (local) piecewise polynomials of degree k with respect to the split. A unisolvent set of degrees of freedom which induces a globally C^1 piecewise polynomial space is given in the next proposition. We refer to [**?**] for a proof.

Proposition 1. *For a 1–dimensional edge e, denote by $\{s_i^e\}_{i=1}^{d-1} \subset \mathbb{R}^d$ a set of unit vectors such that, together with the direction determined by the edge, they provide a basis for $\mathbb{R}^d$. Then*

(1) The dimension of $C^1(T) \cap \mathbb{P}_3(T_r)$ is $\frac{1}{2}(d+1)(d^2+d+2)$.

(2) A function $v_h \in C^1(T) \cap \mathbb{P}_3(T_r)$ is uniquely determined by

   (a) the values $v_h(p)$ and $\nabla v_h(p)$ for all $p \in \mathcal{V}(T)$;

   (b) the values $\nabla v_h(p) \cdot s_i^e$ for all $p \in \mathcal{M}_T$ and $i \in \{1, \ldots, d-1\}$ such that p is the edge midpoint of e.

Remark 2. *For the convenience of the proofs below, we set the direction vectors $\{s_i^e\}_{i=1}^{d-1}$ associated with a (boundary) flat edge midpoint as follows: In two and three dimensions, we take s_1^e to be the common normal vector given in Definition* **??**. *In three dimensions, we take s_2^e to be orthogonal to* *both the common normal and the edge. Note that s_2^e is tangent to the boundary faces associated with the edge. Furthermore, we omit the superscript e in the notation when the context is clear.*

The degrees of freedom given in Proposition **??** induce a global piecewise cubic space

$$\tilde{V}_h := \{v_h \in H^2(\Omega) : \ v_h|_T \in C^1(T) \cap \mathbb{P}_3(T_r), \ \forall T \in \mathcal{T}_h\}.$$

A characterization of the associated space with zero Dirichlet boundary conditions with respect to the degrees of freedom is summarized in the next lemma.

Lemma 2. *A function $v_h \in \tilde{V}_h$ satisfies $v_h \in \tilde{V}_{h,0} := \tilde{V}_h \cap H_0^1(\Omega)$ if and only if (i) $v_h(p) = 0$ for $p \in \mathcal{V}_h^B$; (ii) $\frac{\partial v_h}{\partial t}(p) = 0$ for all $p \in \mathcal{V}_h^\flat$ and tangent vectors t with respect to faces/edges in $\mathcal{F}_p^B$; (iii) $\nabla v_h(p) = 0$ for all $p \in \mathcal{V}_h^\# \cup \mathcal{M}_h^\#$; and in three dimensions (iv) $\frac{\partial v_h}{\partial s_2}(p) = 0$ for all $p \in \mathcal{M}_h^\flat$.*

Proof. It is clear that if $v_h \in \tilde{V}_{h,0}$ then (i) is satisfied.

Let $p \in \mathcal{V}_h^\flat$ be a flat vertex, and denote the common normal vector at p by n. If $v_h \in \tilde{V}_{h,0}$, then $v_h = 0$ on $\mathcal{F}_p^B$, and so the tangential derivatives of v_h are zero along the boundary faces in $\mathcal{F}_p^B$; thus (ii) and (iv) hold. Note that the derivative of v_h in the direction of n is not restricted on $\mathcal{V}_h^\flat \cup \mathcal{M}_h^\flat$.

We now show that if $v_h \in \tilde{V}_{h,0}$ then $\nabla v_h = 0$ on sharp nodes. Let $p \in \mathcal{V}_h^\# \cup \mathcal{M}_h^\#$ be a sharp node, and denote by $f_1, \ f_2 \in \mathcal{F}_p^B$ two faces with nonparallel unit normal vectors $n_1, \ n_2$. Then there exist two orthogonal bases of $\mathbb{R}^d$, $\{n_1, t_{1,1}, \dots, t_{1,d-1}\}$ and $\{n_2, t_{2,1}, \dots, t_{2,d-1}\}$, where $\{t_{1i}\}_{i=1}^{d-1}$ are the tangential vectors of f_1, $\{t_{2,j}\}_{j=1}^{d-1}$ are the tangential vectors of f_2. Therefore, for each $i = 1, \dots, d-1$, there exist a unique decomposition $t_{1,i} = \sum_{j=1}^{d-1} c_{i,j} t_{2,j} + c_{i,d} n_2$ ($c_{i,j} \in \mathbb{R}$). We claim that there exists $1 \le i \le d-1$ such that $c_{i,d} \neq 0$.

Suppose the claim is not true, i.e., $c_{i,d} = 0$ for all i, which implies that $\operatorname{span}\{t_{1,i}\}_{i=1}^{d-1} \subset \operatorname{span}\{t_{2,j}\}_{j=1}^{d-1}$ for all j. Since the dimensions of the (linearly independent) sets are the same, we conclude that $\operatorname{span}\{t_{1,i}\}_{i=1}^{d-1} = \operatorname{span}\{t_{2,j}\}_{j=1}^{d-1}$, and therefore $n_1 \cdot t_{2,j} = 0$ for all j. Hence, since $\{n_2, t_{2,1}, \dots, t_{2,d-1}\}$ is a orthogonal basis, we have

$$n_1 = \sum_{j=1}^{d-1} (n_1 \cdot t_{2,j}) t_{2,j} + (n_1 \cdot n_2) n_2 = (n_1 \cdot n_2) n_2,$$

implying that n_1 and n_2 are parallel, a contradiction. Thus there exists $1 \le i \le d-1$ such that $c_{i,d} \neq 0$.

Now since $v_h = 0$ on $\mathcal{F}_p^B$, the tangential derivatives of v_h are zero along the boundary faces f_1 and f_2, i.e., $\frac{\partial v_h}{\partial t_{1,i}}(p) = 0$ and $\frac{\partial v_h}{\partial t_{2,j}}(p) = 0$ for $i, j = 1, \dots, d-1$. We then have

$$0 = \frac{\partial v_h}{\partial t_{1,i}}(p) = \sum_{j=1}^{d-1} c_{i,j} \frac{\partial v_h}{\partial t_{2,j}}(p) + c_{i,d} \frac{\partial v_h}{\partial n_2}(p) = c_{i,d} \frac{\partial v_h}{\partial n_2}(p) \Rightarrow \frac{\partial v_h}{\partial n_2}(p) = 0.$$

Therefore, the directional derivatives of v_h at p are zero along $\{n_2, t_{2,1}, t_{2,2} ..., t_{2,d-1}\}$, the basis of $\mathbb{R}^d$, and it thus follows that (iii) is satisfied.

Finally, suppose that $v_h \in \tilde{V}_h$ vanishes at the values (i)–(iv). Since v_h, respected to an edge, is a one-dimensional cubic polynomial, we conclude from (i)–(iii) that v_h vanishes on the boundary edges. Therefore in the case $d = 2$, $v_h = 0$ on $\partial\Omega$. In three dimensions, we use condition (iv) and the two-dimensional unisolvency result in Proposition **??** to conclude that $v_h = 0$ on $\partial\Omega$ when $d = 3$ as well. ☐

Remark 3. *Let $p \in \mathcal{V}_h^\sharp \cup \mathcal{M}_h^\sharp$, and let $\{t_{1,i}\}_{i=1}^{d-1}$ and $\{t_{2,j}\}_{j=1}^{d-1}$ span the tangent space of some $f_1, f_2 \in \mathcal{F}_p^B$ with nonparallel unit normal vectors. Then the proceeding proof shows that there exists an i such that $\{t_{1,i}, t_{2,1}, \dots, t_{2,d-1}\}$ forms a basis of $\mathbb{R}^d$.*

3.1. **Construction of map $E_h : V_h \to \tilde{V}_{h,0}$.** In this section, we construct a linear operator connecting the Lagrange finite element space to the Clough–Tocher finite element space by averaging. This is done by assigning the values specified in Proposition **??** and Lemma **??**.

Let N be any (global) degree of freedom of $\tilde{V}_{h,0}$. If N is an interior degree of freedom, then we set

$$N(E_h v_h) = \frac{1}{|\mathcal{T}_N|} \sum_{T \in \mathcal{T}_N} N(v_T), \tag{3.1}$$

where $v_T := v_h|_T$ is the function v_h restricted to the simplex T, $\mathcal{T}_N$ is the set of simplexes in $\mathcal{T}_h$ that share the degree of freedom N, and $|\mathcal{T}_N|$ is the number of elements in $\mathcal{T}_N$.

If N corresponds to a function evaluation at a boundary vertex $p \in \mathcal{V}_h^B$, we set $N(E_h v_h) = 0$. If N is a boundary degree of freedom corresponding to the function gradient at a flat vertex $p \in \mathcal{V}_h^\flat$, let the common unit normal vector of faces in $\mathcal{F}_h^b$ be n, and set

$$N(E_h v_h) = \frac{1}{|\mathcal{T}_N|} \sum_{T \in \mathcal{T}_N} (N(v_T) \cdot n) n. \tag{3.2}$$

Thus, $N(E_h v_h)$ is a vector with direction n and magnitude $\frac{1}{|\mathcal{T}_N|} \sum_{T \in \mathcal{T}_N} \frac{\partial v_T}{\partial n}(p)$.

If N is a boundary degree of freedom corresponding to a function directional derivative at $p \in \mathcal{M}_h^\flat$ with direction unit vector s_i, let the common unit normal vector be n, and set

$$N(E_h v_h) = \frac{1}{|\mathcal{T}_N|} \sum_{T \in \mathcal{T}_N} (s_i \cdot n) \frac{\partial v_T}{\partial n}(p). \tag{3.3}$$

Finally, if N is a boundary degree of freedom corresponding to the function derivative or directional derivative at some $p \in \mathcal{V}_h^\# \cup \mathcal{M}_h^\#$, we set $N(E_h v_h) = 0$. Note that this construction and Lemma **??** show that $E_h v_h \in \tilde{V}_{h,0}$.

Lemma 3. *For $k = 2$ or 3, the map E_h satisfies the estimate*

$$|v_h - E_h v_h|^2_{H^2(\mathcal{T}_h)} \lesssim \sum_{f \in \mathcal{F}_h^I} h_f^{-1} \left\| [\![\partial v_h / \partial n_f]\!] \right\|^2_{L^2(f)} \quad \forall v_h \in V_h. \tag{3.4}$$

Proof. The proof of (**??**) in the two dimensional setting is given in [**?**, **?**], thus it suffices to prove the result when $d = 3$.

Let $v_h \in V_h$ be arbitrary and set $w_h = v_h - E_h v_h$. Fix $T \in \mathcal{T}_h$ and set $w_T = w_h|_T$. From Proposition **??**, the inclusion $\mathbb{P}_3(T) \subset \mathbb{P}_3(T_r) \cap C^1(T)$, scaling and shape regularity, and since $w_h(p) = 0$ for all $p \in \mathcal{V}_h$, we have

$$\begin{aligned} \|v_h - E_h v_h\|^2_{L^2(T)} &= \|w\|^2_{L^2(T)} \\ &\lesssim \sum_{p \in \mathcal{V}_T} \left(h_T^3 |w(p)|^2 + h_T^5 |\nabla w_T(p)|^2\right) + \sum_{m \in \mathcal{M}_T} \sum_{i=1}^2 h_T^5 \left| \frac{\partial w_T}{\partial s_i}(m) \right|^2 \\ &= \sum_{p \in \mathcal{V}_T} h_T^5 |\nabla w_T(p)|^2 + \sum_{m \in \mathcal{M}_T} \sum_{i=1}^2 h_T^5 \left| \frac{\partial w_T}{\partial s_i}(m) \right|^2. \end{aligned} \tag{3.5}$$

By (**??**), for an interior point p (i.e., a point that is not on $\partial\Omega$), we have

$$\begin{aligned} |\nabla w_T(p)|^2 &= \Big(\frac{1}{|\mathcal{T}_p|} \sum_{T' \in \mathcal{T}_p} |\nabla v_T(p) - \nabla v_{T'}(p)| \Big)^2 \\ &\lesssim \sum_{T' \in \mathcal{T}_p} |\nabla v_T(p) - \nabla v_{T'}(p)|^2. \end{aligned} \tag{3.6}$$

For any $T' \in \mathcal{T}_p$, there exist a finite sequence of simplices $\{T_j\}_{j=0}^M \subset \mathcal{T}_p$ labeled such that $T_0 = T$, $T_M = T'$, and $\partial T_j \cap \partial T_{j+1} \in \mathcal{F}_h^I$. We emphasize that M is is bounded uniformly in h by the shape regularity of $\mathcal{T}_h$ (cf. Remark **??**). Hence, by an inverse estimate, and since v_h is continuous across the faces,

$$\begin{aligned} |\nabla v_T(p) - \nabla v_{T'}(p)| &\le \sum_{j=0}^{M-1} |\nabla v_{T_j}(p) - \nabla v_{T_{j+1}}(p)| \\ &\le \sum_{f\in\mathcal{F}_p^I} \big\|[\![\nabla v_h]\!]\big\|_{L^\infty(f)} \\ &\lesssim \sum_{f\in\mathcal{F}_p^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}. \end{aligned} \tag{3.7}$$

Applying (**??**) to (**??**), we find that

$$|\nabla w(p)|^2 \lesssim \sum_{f\in\mathcal{F}_p^I} h_f^{-2}\big\|[\![\partial v_h/\partial n_f]\!]\big\|^2_{L^2(f)}. \tag{3.8}$$

Using similar arguments, we have for any interior midpoint m and $i \in \{1,2\}$,

$$\Big|\frac{\partial v_T}{\partial s_i}(m) - \frac{\partial v_{T'}}{\partial s_i}(m)\Big| \le \sum_{f\in\mathcal{F}_m^I} \big\|[\![\nabla v_h]\!]\big\|_{L^\infty(f)} \lesssim \sum_{f\in\mathcal{F}_m^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)},$$

and therefore,

$$\Big|\frac{\partial w_T}{\partial s_i}(m)\Big|^2 = \Big|\frac{1}{|\mathcal{T}_m|}\sum_{T'\in\mathcal{T}_m} |\frac{\partial v_T}{\partial s_i}(m) - \frac{\partial v_{T'}}{\partial s_i}(m)|\Big|^2 \lesssim \sum_{f\in\mathcal{F}_m^I} h_f^{-2}\,\|[\![\partial v/\partial n_f]\!]\|^2_{L^2(f)}\,. \tag{3.9}$$

At a sharp vertex p, $\nabla E_h v_h$ vanishes, and thus

$$|\nabla w_T(p)|^2 = |\nabla v_T(p)|^2.$$

Since p is a sharp vertex, there exist two simplexes T', $T'' \in \mathcal{T}_p$, boundary faces $f_1 \subset \partial T' \cap \partial\Omega$, $f_2 \subset \partial T'' \cap \partial\Omega$, and f_1, f_2 do not have a common normal vector. Hence, by Remark **??**, there exist a tangential vector $t_{1,i}$ of f_1 and two tangential vectors $\{t_{2,1}, t_{2,2}\}$ of f_2 such that together, the three vectors form a basis of $\mathbb{R}^3$.

By connecting T through a sequence of simplex in $\mathcal{T}_p$ to T', we have

$$\begin{aligned} \Big|\frac{\partial v_T}{\partial t_{1,i}}(p)\Big|^2 &\lesssim \sum_{f\in\mathcal{F}_p^I} h_f^{-2}\big\|[\![\partial v_h/\partial t_{1,i}]\!]\big\|^2_{L^2(f)} + h_{f_1}^{-2}\,\|\partial v_{T'}/\partial t_{1,i}\|^2_{L^2(f_1)} \\ &\lesssim \sum_{f\in\mathcal{F}_p^I} h_f^{-2}\big\|[\![\partial v_h/\partial n_f]\!]\big\|^2_{L^2(f)} \end{aligned} \tag{3.10}$$

because the tangential derivatives of v vanish on $\partial\Omega$.

Similarly, by connecting T through a sequence of simplex in $\mathcal{T}_p$ to T'', we have

$$\Big|\frac{\partial v_T}{\partial t_{2,j}}(p)\Big|^2 \lesssim \sum_{f\in\mathcal{F}_p^I} h_f^{-2}\big\|[\![\partial v_h/\partial n_f]\!]\big\|^2_{L^2(f)} \quad \text{for } j = 1,2,$$

and therefore

$$|\nabla w_T(p)|^2 \lesssim \Big|\frac{\partial v_T}{\partial t_{1,i}}(p)\Big|^2 + \sum_{j=1}^{2}\Big|\frac{\partial v_T}{\partial t_{2,j}}(p)\Big|^2 \lesssim \sum_{f\in\mathcal{F}_p^I} h_f^{-2}\,\|\partial v_h/\partial n_f\|^2_{L^2(f)} \qquad \forall p \in \mathcal{V}_T \cap \mathcal{V}_h^\sharp. \tag{3.11}$$

Next, for a boundary flat vertex $p \in \mathcal{V}_h^\flat$ with common unit normal vector n, we first write

$$|\nabla w_T(p)|^2 \lesssim \Big|\frac{\partial v_T}{\partial n}(p)\,n - \nabla E_h v_h(p)\Big|^2 + \Big|\nabla v_T(p) - \frac{\partial v_T}{\partial n}(p)\,n\Big|^2. \tag{3.12}$$

By (**??**) and by applying similar steps in (**??**), (**??**), and (**??**), we have

$$\begin{aligned}|\frac{\partial v_T}{\partial n}(p)\, n - \nabla E_h v_h|^2 &\lesssim \sum_{T'\in\mathcal{T}_p} \Big|\big(\frac{\partial v_T}{\partial n}(p) - \nabla v_{T'}(p)\cdot n\big) n\Big|^2 \\ &= \sum_{T'\in\mathcal{T}_p} \Big|\frac{\partial v_T}{\partial n}(p) - \frac{\partial v_{T'}}{\partial n}(p)\Big|^2 \\ &\lesssim \sum_{f\in\mathcal{F}_p^I} h_f^{-2}\, \|[\![\partial v_h/\partial n_f]\!]\|_{L^2(f)}^2 .\end{aligned} \tag{3.13}$$

Since p is a flat vertex, there exist a simplex T' with a boundary face $f_3 \in \mathcal{F}_p^B$. Let $\{t_{3,1}, t_{3,2}\}$ denote an orthonormal basis of f_3. Then by marching to the boundary (as in (**??**)), we have

$$\begin{aligned}|\nabla v_T(p) - \frac{\partial v_T}{\partial n}(p)\, n|^2 &= |\sum_{i=1}^2 \frac{\partial v_T}{\partial t_{3,i}}(p)\, t_{3i}|^2 \\ &\lesssim \sum_{i=1}^2 |\frac{\partial v_T}{\partial t_{3,i}}(p)|^2 \lesssim \sum_{f\in\mathcal{F}_p^I} h_f^{-2}\, \|[\![\partial v/\partial n_f]\!]\|_{L^2(f)}^2 .\end{aligned} \tag{3.14}$$

Hence, by (**??**)–(**??**), we have

$$|\nabla w(p)|^2 \lesssim \sum_{f\in\mathcal{F}_p^I} h_f^{-2}\, \|[\![\partial v/\partial n_f]\!]\|_{L^2(f)}^2 \quad \forall p \in \mathcal{V}_T \cap \mathcal{V}_h^\flat. \tag{3.15}$$

For a boundary flat midpoint m, by (**??**) and (**??**), we have

$$\begin{aligned}\Big|\frac{\partial w}{\partial s_1}(m)\Big|^2 &\lesssim \Big|(s_1\cdot n)\frac{\partial v_T}{\partial n}(m) - \frac{\partial E_h v_h}{\partial s_1}(m)\Big|^2 + \Big|\frac{\partial v_T}{\partial s_1}(m) - (s_1\cdot n)\frac{\partial v_T}{\partial n}(m)\Big|^2 \\ &\lesssim \sum_{T'\in\mathcal{T}_m} \Big|\frac{\partial v_T}{\partial n}(m) - \frac{\partial v_{T'}}{\partial n}(m)\Big|^2 \\ &\lesssim \sum_{f\in\mathcal{F}_m^I} h_f^{1-d}\, \|[\![\partial v/\partial n_f]\!]\|_{L^2(f)}^2 .\end{aligned} \tag{3.16}$$

Likewise, we have that

$$\begin{aligned}\Big|\frac{\partial w}{\partial s_2}(m)\Big|^2 &\lesssim \Big|(s_2\cdot n)\frac{\partial v_T}{\partial n}(m) - \frac{\partial E_h v_h}{\partial s_2}(m)\Big|^2 + \Big|\frac{\partial v_T}{\partial s_2}(m) - (s_2\cdot n)\frac{\partial v_T}{\partial n}(m)\Big|^2 \\ &= \Big|\frac{\partial v_T}{\partial s_2}(m)\Big|^2 \lesssim \sum_{f\in\mathcal{F}_m^I} h_f^{-2}\, \|[\![\partial v/\partial n_f]\!]\|_{L^2(f)}^2 .\end{aligned} \tag{3.17}$$

Combining (**??**), (**??**), (**??**), (**??**), (**??**), (**??**), and (**??**) yields

$$\|v_h - E_h v_h\|_{L^2(T)}^2 \lesssim h_T^5 \Big(\sum_{p\in\mathcal{V}_T} \sum_{f\in\mathcal{F}_p^I} h_f^{-2}\, \|[\![\partial v/\partial n_f]\!]\|_{L^2(f)}^2 + \sum_{m\in\mathcal{M}_T} \sum_{f\in\mathcal{F}_m^I} h_f^{-2}\, \|[\![\partial v/\partial n_f]\!]\|_{L^2(f)}^2 \Big).$$

Finally, by an inverse estimate and the shape regularity of $\mathcal{T}_h$, we obtain

$$
\begin{aligned}
|v_h - E_h v_h|^2_{H^2(\mathcal{T}_h)} &\lesssim \sum_{T\in\mathcal{T}_h} h_T^{-4} \|v_h - E_h v_h\|^2_{L^2(T)} \\
&\lesssim \sum_{T\in\mathcal{T}_h} h_T \Big(\sum_{p\in\mathcal{V}_T} \sum_{f\in\mathcal{F}_p^I} h_f^{-2} \, \|[\![\partial v/\partial n_f]\!]\|^2_{L^2(f)} \\
&\qquad + \sum_{m\in\mathcal{M}_T} \sum_{f\in\mathcal{F}_m^I} h_f^{-2} \, \|[\![\partial v/\partial n_f]\!]\|^2_{L^2(f)} \Big) \\
&\lesssim \sum_{f\in\mathcal{F}_h^I} h_f^{-1} \, \|[\![\partial v/\partial n_f]\!]\|^2_{L^2(f)} .
\end{aligned}
$$

□

Remark 4. *In two dimensions, there exists a family of C^1 Clough-Tocher spaces of degree greater than or equal to three [?]. As a result, the estimate (**??**) can be generalized to arbitrary $k \geq 2$ [?]. However, as far as we are aware, degrees of freedom for higher-order Clough-Tocher spaces in three dimensions are not found in the literature; see [?] for partial results. As a result, the estimate (**??**) is restricted to $2 \leq k \leq 3$ if $d = 3$.*

Remark 5. *An operator that maps piecewise polynomials to $H^2(\Omega) \cap H^1_0(\Omega)$-conforming functions has been recently been constructed in [?]. There, the mesh is allowed to have hanging nodes, and the dependence of the polynomial degree is explicitly stated in the estimate. On the other hand, the operator is constructed in a global fashion, and as such, it seems that the mesh must be quasi-uniform in order to get an estimate analogous to (**??**) by directly using [?, Theorem 4].*

4. Proof of Theorem 1

With the result of Lemma **??**, we are able to prove Theorem **??**.

Proof. For $v_h \in V_h$, we have $E_h v_h \in H^2(\Omega) \cap H^1_0(\Omega)$, and therefore, by the Miranda-Talenti and triangle inequalities,

$$
\begin{aligned}
|v_h|_{H^2(\mathcal{T}_h)} &\leq |E_h v_h|_{H^2(\mathcal{T}_h)} + |v_h - E_h v_h|_{H^2(\mathcal{T}_h)} \\
&\leq \|\Delta E_h v_h\|_{L^2(\mathcal{T}_h)} + |v_h - E_h v_h|_{H^2(\mathcal{T}_h)} \\
&\leq \|\Delta(v_h - E_h v_h)\|_{L^2(\mathcal{T}_h)} + \|\Delta v_h\|_{L^2(\mathcal{T}_h)} + |v_h - E_h v_h|_{H^2(\mathcal{T}_h)}.
\end{aligned} \tag{4.1}
$$

We then use the identity $\|\Delta v_h\|_{L^2(\mathcal{T}_h)} \leq \sqrt{d}|v_h|_{H^2(\mathcal{T}_h)}$ and Lemma **??** to get

$$
\begin{aligned}
\|D^2 v_h\|_{L^2(\mathcal{T}_h)} &\leq \|\Delta v_h\|_{L^2(\mathcal{T}_h)} + (1+\sqrt{d})|v_h - E_h v_h|_{H^2(\mathcal{T}_h)} \\
&\leq \|\Delta v_h\|_{L^2(\mathcal{T}_h)} + C(1+\sqrt{d}) \Big(\sum_{f\in\mathcal{F}_h^I} h_f^{-1} \|[\![\partial v_h/\partial n_f]\!]\|^2_{L^2(f)} \Big)^{1/2}.
\end{aligned}
$$

The proof is complete with $C_\dagger = C(1+\sqrt{d})$. □

Corollary 1. *There holds, for all $\tau \in (0,1)$,*

$$
\|\Delta v_h\|^2_{L^2(\mathcal{T}_h)} \geq (1-\tau)\|D^2 v_h\|^2_{L^2(\mathcal{T}_h)} - \frac{C_\dagger^2}{\tau} \sum_{f\in\mathcal{F}_h^I} h_f^{-1} \big\|[\![\partial v_h/\partial n_f]\!]\big\|^2_{L^2(f)}.
$$

Proof. Applying the Cauchy–Schwarz inequality to Theorem **??** yields

$$
\|D^2 v_h\|^2_{L^2(\mathcal{T}_h)} \leq \big(1+\rho\big)\|\Delta v_h\|^2_{L^2(\mathcal{T}_h)} + C_\dagger^2\big(1+\frac{1}{\rho}\big) \sum_{f\in\mathcal{F}_h^I} h_f^{-1} \big\|[\![\partial v_h/\partial n_f]\!]\big\|^2_{L^2(f)}
$$

for any $\rho > 0$. Letting $\tau = \rho/(1+\rho) \in (0,1)$ and rearranging terms, we have

$$\begin{aligned}\|\Delta v_h\|_{L^2(\mathcal{T}_h)}^2 &\ge (1-\tau)\|D^2 v_h\|_{L^2(\mathcal{T}_h)}^2 - C_\dagger^2 \frac{(1-\tau)}{\tau} \sum_{f\in\mathcal{F}_h^I} h_f^{-1} \|[\![\partial v_h/\partial n_f]\!]\|_{L^2(f)}^2 \\ &\ge (1-\tau)\|D^2 v_h\|_{L^2(\mathcal{T}_h)}^2 - \frac{C_\dagger^2}{\tau} \sum_{f\in\mathcal{F}_h^I} h_f^{-1} \|[\![\partial v_h/\partial n_f]\!]\|_{L^2(f)}^2.\end{aligned}$$

□

5. Applications to linear problems in nondivergence form

In this section, motivated by the discrete Miranda-Talenti estimate, we construct simple convergent finite element methods to approximate strong solutions for elliptic problems in non-divergence form:

$$\mathcal{L}u := A : D^2u + \boldsymbol{b}\cdot\nabla u - cu = g \qquad \text{in } \Omega, \tag{5.1a}$$

$$u = 0 \qquad \text{on } \partial\Omega. \tag{5.1b}$$

Here, $A : B := \sum_{i,j=1}^d A_{i,j}B_{i,j}$ denotes the Frobenius inner product of two matrices. We recall that u is a strong solution to (**??**) if it has regularity

$$u \in V := H^2(\Omega)\cap H_0^1(\Omega),$$

and satisfies (**??**) almost everywhere in Ω.

To ensure the well-posedness of problem (**??**) we assume that $g \in L^2(\Omega)$, that the coefficients satisfy $A \in [L^\infty(\Omega)]^{d\times d}$, $\boldsymbol{b} \in [L^\infty(\Omega)]^d$, $c \in L^\infty(\Omega)$ with $c \ge 0$, and that A is uniformly positive definite in Ω, i.e., there exists $\underline{\nu}, \bar{\nu} > 0$ such that

$$\underline{\nu}|\boldsymbol{\xi}|^2 \le \boldsymbol{\xi}^t A(x)\boldsymbol{\xi} \le \bar{\nu}|\boldsymbol{\xi}|^2 \quad \text{a.e. } x\in\Omega,$$

for all $\boldsymbol{\xi}\in\mathbb{R}^d$. Here, $|\boldsymbol{\xi}|$ denotes the Euclidean distance of $\boldsymbol{\xi}$ from the origin. More importantly, we assume that the coefficients satisfy the *Cordes Condition*.

Definition 2. *The coefficients satisfy the* Cordes Condition *if*

(i) whenever $c \not\equiv 0$ or $\boldsymbol{b}\not\equiv 0$, there exists $\lambda > 0$ and $\epsilon \in (0,1)$ such that

$$\frac{|A|^2 + |\boldsymbol{b}|^2/2\lambda + (c/\lambda)^2}{(\operatorname{tr} A + c/\lambda)^2} \le \frac{1}{d+\epsilon} \quad \text{a.e. } x\in\Omega. \tag{5.2a}$$

Here, $|A| = \sqrt{A:A}$, and $\operatorname{tr} A = \sum_{i=1}^d A_{ii}$ is the trace of A.

(ii) whenever $c\equiv 0$ and $\boldsymbol{b}\equiv 0$, there exists $\epsilon\in(0,1)$ such that

$$\frac{|A|^2}{(\operatorname{tr} A)^2} \le \frac{1}{d-1+\epsilon} \quad \text{a.e. } x\in\Omega. \tag{5.2b}$$

Remark 6. *In two dimensions, and in the case $c\equiv 0$ and $\boldsymbol{b}\equiv 0$, uniform ellipticity of A implies the Cordes condition with $\epsilon = 2\underline{\nu}/(\underline{\nu}+\bar{\nu})$ [**?**, Example 2].*

Remark 7. *Uniformly ellipticity of $A\in[L^\infty(\Omega)]^{d\times d}$ is not sufficient to ensure that there exists a unique strong solution to problem (**??**), at least in three dimensions. The following classical example illustrates this feature.*

Example 1. *Let $d = 3$, $\Omega = B_1(0)$ be the unit ball, $c\equiv 0$, $\boldsymbol{b}\equiv 0$, and*

$$A(x) = I_3 + \Big(\frac{1+\alpha}{1-\alpha}\Big)\frac{xx^t}{|x|^2},$$

where $1/2 < \alpha < 1$ and I_3 denotes the 3×3 identity matrix. Clearly A is essentially bounded and uniformly positive definite with $\underline{\nu} = 1$ and $\bar{\nu} = 2/(1-\alpha)$.

The function $u(x) = |x|^\alpha - 1$ satisfies $u \in V$,

$$D^2u(x) = \alpha(\alpha-2)|x|^{\alpha-4}xx^t + \alpha|x|^{\alpha-2}I_3,$$

and since $I_3 : (xx^t) = (xx^t) : (xx^t)/|x|^2 = |x|^2$,

$$A(x) : D^2u(x) = \alpha(\alpha-2)|x|^{\alpha-2}\Big(1 + \frac{1+\alpha}{1-\alpha}\Big) + \alpha|x|^{\alpha-2}\Big(3 + \frac{1+\alpha}{1-\alpha}\Big) = 0.$$

*Therefore both $u = |x|^\alpha - 1$ and the zero function are strong solutions to (**??**) with $g \equiv 0$.*

Note that

$$|A|^2 = 3 + 2\Big(\frac{1+\alpha}{1-\alpha}\Big) + \Big(\frac{1+\alpha}{1-\alpha}\Big)^2 = \frac{2(\alpha^2-2\alpha+3)}{(1-\alpha)^2},$$
$$(\operatorname{tr} A)^2 = \Big(3 + \frac{1+\alpha}{1-\alpha}\Big)^2 = \frac{4(\alpha^2-4\alpha+4)}{(1-\alpha)^2},$$

and therefore

$$|A|^2 - \frac{1}{2}(\operatorname{tr} A)^2 = \frac{2(2\alpha-1)}{(1-\alpha)^2} > 0.$$

Thus, the coefficients do not satisfy the Cordes condition.

Define the function $\gamma \in L^\infty(\Omega)$ by

$$\gamma := \frac{\operatorname{tr} A + c/\lambda}{|A|^2 + |\boldsymbol{b}|^2/2\lambda + (c/\lambda)^2}. \tag{5.3}$$

Since A is positive definite and c is non-negative, we clearly see that $\gamma > 0$. In particular, if the Cordes condition (**??**) is satisfied, then

$$\gamma \geq \frac{d+\epsilon}{\operatorname{tr} A + c/\lambda} \geq \frac{d+\epsilon}{d\bar{\nu} + \|c\|_{L^\infty(\Omega)}/\lambda} =: \gamma_0.$$

To state the well-posedness of problem (**??**), we define the operators $\mathcal{L}_\gamma, \mathcal{L}_\lambda : V \to L^2(\Omega)$ by

$$\mathcal{L}_\gamma v := \gamma \mathcal{L} v, \qquad \mathcal{L}_\lambda v := \Delta v - \lambda v, \tag{5.4}$$

where in the case that $\boldsymbol{b} \equiv 0$ and $c \equiv 0$, we set $\lambda = 0$ in (**??**). Note that, since λ is nonnegative and Ω is convex, the mapping $\mathcal{L}_\lambda : V \to L^2(\Omega)$ is surjective. Moreover, since $\gamma \geq \gamma_0 > 0$ a.e. in Ω, simple arguments show that $u \in V$ satisfies (**??**) if and only if $\mathcal{L}_\gamma u = \gamma g$ a.e. in Ω. Thus, these two observations show that $u \in V$ is a strong solution to (**??**) if and only if

$$B(u,v) := \int_\Omega (\mathcal{L}_\gamma u)(\mathcal{L}_\lambda v)\,dx = \int_\Omega \gamma g(\mathcal{L}_\lambda v)\,dx \qquad \forall v \in V. \tag{5.5}$$

Lemma 4. *Under the given assumptions, there holds the following inequality a.e. in Ω:*

$$|\mathcal{L}_\gamma w - \mathcal{L}_\lambda w| \leq \sqrt{1-\epsilon}\sqrt{|D^2w|^2 + 2\lambda|\nabla w|^2 + \lambda^2|w|^2}. \tag{5.6}$$

Proof. Suppose that $\boldsymbol{b} \not\equiv 0$ or $c \not\equiv 0$.

Applying the definitions of the operators and the Cauchy–Schwarz inequality, we have

$$\begin{aligned} |\mathcal{L}_\gamma w - \mathcal{L}_\lambda w| &\leq |\gamma A - I_d||D^2w| + |\gamma||\boldsymbol{b}||\nabla w| + |\lambda - c\gamma||w| \\ &\leq \sqrt{M}\sqrt{|D^2w|^2 + 2\lambda|\nabla w|^2 + \lambda^2|w|^2}, \end{aligned} \tag{5.7}$$

with

$$M := |\gamma A - I_d|^2 + |\gamma|^2\frac{|\boldsymbol{b}|^2}{2\lambda} + \frac{|\lambda - c\gamma|^2}{\lambda^2}.$$

Expanding this expression out and using the definition of γ and the Cordes condition (**??**), we have

$$\begin{aligned} M &= d+1-2\gamma(\operatorname{tr} A + \frac{c}{\lambda}) + |\gamma|^2(|A|^2 + \frac{|\boldsymbol{b}|^2}{2\lambda} + \frac{|c|^2}{\lambda^2}) \\ &= d+1-\gamma\big(\operatorname{tr} A + c/\lambda\big) \\ &= d+1-\frac{\big(\operatorname{tr} A + c/\lambda\big)^2}{|A|^2+|\boldsymbol{b}|^2/(2\lambda)+(c/\lambda)^2} \\ &\le 1-\epsilon. \end{aligned}$$

Combining this inequality with (**??**) yields (**??**).

Likewise, for the special case $\boldsymbol{b}\equiv 0$, $c\equiv 0$ and $\lambda = 0$, we have by (**??**),

$$\begin{aligned} |\mathcal{L}_\gamma w - \mathcal{L}_\lambda w| &\le |\gamma A - I_d||D^2 w| \\ &= \sqrt{d-2\gamma \operatorname{tr} A + |\gamma|^2|A|^2}|D^2 w| \\ &= \sqrt{d-\gamma \operatorname{tr} A}|D^2 w| \\ &\le \sqrt{1-\epsilon}|D^2 w|. \end{aligned}$$

□

Lemma 5. *If Ω is convex, then there holds*

$$\|\mathcal{L}_\lambda v\|^2_{L^2(\Omega)} \ge \int_\Omega \big(|D^2 v|^2 + 2\lambda|\nabla v|^2 + \lambda^2|v|^2\big)\,dx \quad \forall v\in V.$$

Proof. Integration by parts gives

$$\|\mathcal{L}_\lambda v\|^2_{L^2(\Omega)} = \int_\Omega \big(|\Delta v|^2 + 2\lambda|\nabla v|^2 + \lambda^2|v|^2\big)\,dx.$$

An application of the Miranda-Talenti estimate now yields the result. □

Theorem 2. *There exists a unique strong solution to* (**??**) *provided the Cordes condition is satisfied.*

Proof. A proof of this result is given in [**?**] (also see [**?**, **?**]). However, we give it here for completeness and to motivate the numerical analysis of the method given in the next section.

Let $v\in V$, and write

$$B(v,v) = \int_\Omega |\mathcal{L}_\lambda v|^2 + \int_\Omega \big(\mathcal{L}_\gamma v - \mathcal{L}_\lambda v\big)(\mathcal{L}_\lambda v)\,dx.$$

Applying Lemmas **??**–**??** yields

$$B(v,v) \ge \big(1-\sqrt{1-\epsilon}\big)\big\|\mathcal{L}_\lambda v\big\|^2_{L^2(\Omega)},$$

and therefore $B(\cdot,\cdot)$ is coercive on V. Since $v\to\int_\Omega \gamma g\mathcal{L}_\lambda v\,dx$ is clearly a bounded linear form on V, with $\big|\int_\Omega \gamma g \mathcal{L}_\lambda v\,dx\big| \le \|\gamma\|_{L^\infty(\Omega)}\|g\|_{L^2(\Omega)}\|\mathcal{L}_\lambda v\|_{L^2(\Omega)}$, the Lax–Milgram theorem shows that there exists a unique $u\in V$ satisfying $B(u,v) = \int_\Omega \gamma g\mathcal{L}_\lambda v\,dx$ for all $v\in V$. Equivalently, there exists a unique solution $u\in V$ satisfying (**??**). □

5.1. **Finite element method.** Based on the discrete Miranda–Talenti estimate and the arguments given in Theorem **??**, we propose the following finite element scheme to approximate the solution to (**??**): Find $u_h\in V_h$ such that

$$B_h(u_h,v_h) = \sum_{T\in\mathcal{T}_h}\int_T \gamma g\mathcal{L}_\lambda v_h\,dx \qquad \forall v_h\in V_h, \tag{5.8}$$

where the bilinear form $B(\cdot,\cdot)$ is given by

$$B_h(w,v) = \sum_{T\in\mathcal{T}_h}\int_T (\mathcal{L}_\gamma w)(\mathcal{L}_\lambda v_h)\,dx + \sigma\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\int_f [\![\partial w/\partial n_f]\!][\![\partial v/\partial n_f]\!]\,ds,$$

$\sigma > 0$ is a positive penalization parameter, and we recall that V_h is the Lagrange finite element space of degree k.

We immediately notice that the scheme (**??**) is consistent. Indeed, if $u \in V$ is a strong solution to (**??**) then $\mathcal{L}_\gamma u = \gamma g$ a.e. in Ω and $[\![\partial u/\partial n_f]\!] = 0$ on $\mathcal{F}_h^I$; thus,

$$B_h(u, v_h) = \sum_{T\in\mathcal{T}_h} \int_T \gamma g \mathcal{L}_\lambda v_h \, dx \qquad \forall v_h \in V_h.$$

To analyze method (**??**) and to show that there exists a unique solution, we introduce the following norm on $V + V_h$:

$$\|v\|_h^2 := \|D^2 v\|_{L^2(\mathcal{T}_h)}^2 + 2\lambda\|\nabla v\|_{L^2(\Omega)}^2 + \lambda^2\|v\|_{L^2(\Omega)}^2 + \sum_{f\in\mathcal{F}_h^I} h_f^{-1} \|[\![\nabla v]\!]\|_{L^2(f)}^2 . \tag{5.9}$$

Note that if $\|v\|_h = 0$ with $v \in V + V_h$, then the Hessian of v vanishes on each element $T \in \mathcal{T}_h$, and $[\![\partial v/\partial n_f]\!] = 0$ on all $f \in \mathcal{F}_h^I$. This implies that v is a linear polynomial on Ω. Since v vanishes on $\partial\Omega$, then we conclude that $v \equiv 0$. Thus, $\|\cdot\|_h$ is indeed a norm on $V + V_h$ for $\lambda \geq 0$.

The next lemma, a discrete analogue of Lemma **??**, relates the discrete norm $\|\cdot\|_h$ with the operator $\mathcal{L}_\lambda$ on V_h.

Lemma 6. *There exists a constant $C_1 > 0$, depending on k and the shape-regularity of the mesh such that, for all $\tau \in (0, 1)$,*

$$\|\mathcal{L}_\lambda v_h\|_{L^2(\mathcal{T}_h)}^2 \geq (1-\tau)\|v_h\|_h^2 - C_1\tau^{-1} \sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2 \quad \forall v_h \in V_h.$$

Proof. Using the definition of $\mathcal{L}_\lambda$ and integrating by parts, we have

$$\begin{aligned}
\|\mathcal{L}_\lambda v_h\|_{L^2(\mathcal{T}_h)}^2 &= \sum_{T\in\mathcal{T}_h} \int_T \Big(|\Delta v_h|^2 + \lambda^2|v_h|^2 - 2\lambda v_h \Delta v_h\Big)\, dx \\
&= \|\Delta v_h\|_{L^2(\mathcal{T}_h)}^2 + 2\lambda\|\nabla v_h\|_{L^2(\Omega)}^2 + \lambda^2\|v_h\|_{L^2(\Omega)}^2 \\
&\quad - 2\lambda \sum_{f\in\mathcal{F}_h^I} \int_f v_h [\![\partial v_h/\partial n_f]\!]\, ds.
\end{aligned}$$

Therefore by Corollary **??**,

$$\begin{aligned}
\|\mathcal{L}_\lambda v_h\|_{L^2(\mathcal{T}_h)}^2 &\geq (1-\tau)\|D^2 v_h\|_{L^2(\mathcal{T}_h)}^2 + 2\lambda\|\nabla v_h\|_{L^2(\Omega)}^2 + \lambda^2\|v_h\|_{L^2(\Omega)}^2 \\
&\quad - \frac{C_\dagger^2}{\tau} \sum_{f\in\mathcal{F}_h^I} \big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2 - 2\lambda \sum_{f\in\mathcal{F}_h^I} \int_f v_h [\![\partial v_h/\partial n_f]\!]\, ds
\end{aligned} \tag{5.10}$$

for all $\tau \in (0, 1)$.

By the Cauchy-Schwarz inequality and scaling, we find that

$$\begin{aligned}
2\lambda \sum_{f\in\mathcal{F}_h^I} \int_f v_h [\![\partial v_h/\partial n_f]\!]\, ds &\leq C\lambda\|v_h\|_{L^2(\Omega)} \Big(\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2\Big)^{1/2} \\
&\leq \lambda^2\tau\|v_h\|_{L^2(\Omega)}^2 + \frac{C^2}{4\tau} \sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2.
\end{aligned}$$

Applying this estimate to (**??**) and applying the definition of $\|\cdot\|_h$ yields

$$\begin{aligned}
\|\mathcal{L}_\lambda v_h\|_{L^2(\mathcal{T}_h)}^2 &\geq (1-\tau)\|D^2 v_h\|_{L^2(\mathcal{T}_h)}^2 + 2\lambda\|\nabla v_h\|_{L^2(\Omega)}^2 + \lambda^2(1-\tau)\|v_h\|_{L^2(\Omega)}^2 \\
&\quad - \Big(\frac{C_\dagger^2}{\tau} + \frac{C^2}{4\tau}\Big) \sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2
\end{aligned}$$

$$\geq (1-\tau)\|v_h\|_h^2 - \Big(1-\tau+\frac{C_\dagger^2}{\tau}+\frac{C^2}{4\tau}\Big)\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2$$

$$\geq (1-\tau)\|v_h\|_h^2 - \Big(\tau^{-1}+\frac{C_\dagger^2}{\tau}+\frac{C^2}{4\tau}\Big)\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2.$$

Setting $C_1 = 1 + C_\dagger^2 + C^2/4$ yields the result. □

Lemma 7. *For any $\alpha \in (0,1)$, there exists $\sigma_\alpha > 0$, independent of h, such that if $\sigma \geq \sigma_\alpha$, there holds*

$$B_h(v_h, v_h) \geq \alpha\big(1-\sqrt{1-\epsilon}\big)\|v_h\|_h^2.$$

*Consequently, there exists a unique solution $u_h \in V_h$ to (**??**) provided σ is sufficiently large.*

Proof. We add and subtract $\mathcal{L}_\lambda v_h$ and apply Lemma **??** and the Cauchy–Schwarz inequality to obtain

$$\begin{aligned}B_h(v_h,v_h) &= \sum_{T\in\mathcal{T}_h}\int_T \big(\mathcal{L}_\gamma v_h - \mathcal{L}_\lambda v_h\big)\big(\mathcal{L}_\lambda v_h\big)\,dx + \|\mathcal{L}_\lambda v_h\|_{L^2(\mathcal{T}_h)}^2 + \sigma\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2\\
&\geq \|\mathcal{L}_\lambda v_h\|_{L^2(\mathcal{T}_h)}^2 - \sqrt{1-\epsilon}\|v_h\|_h\|\mathcal{L}_\lambda v_h\|_{L^2(\mathcal{T}_h)} + \sigma\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2\\
&\geq \big(1-\frac{1}{2}\sqrt{1-\epsilon}\big)\|\mathcal{L}_\lambda v_h\|_{L^2(\mathcal{T}_h)}^2 - \frac{1}{2}\sqrt{1-\epsilon}\|v_h\|_h^2 + \sigma\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2.\end{aligned}$$

Using Lemma **??** we find that

$$\begin{aligned}B_h(v_h,v_h) &\geq \Big((1-\tau)\big(1-\frac{1}{2}\sqrt{1-\epsilon}\big) - \frac{1}{2}\sqrt{1-\epsilon}\Big)\|v_h\|_h^2\\
&\quad + \Big(\sigma - C_1\tau^{-1}\big(1-\frac{1}{2}\sqrt{1-\epsilon}\big)\Big)\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2\end{aligned}$$

for any $\tau \in (0,1)$. For given $\alpha \in (0,1)$, we set $\tau = (1-\alpha)(1-\sqrt{1-\epsilon})/(1-\frac{1}{2}\sqrt{1-\epsilon})$. This yields

$$\begin{aligned}B_h(v_h,v_h) &\geq \alpha\big(1-\sqrt{1-\epsilon}\big)\|v_h\|_h^2\\
&\quad + \Big(\sigma - C_1\frac{(1-\frac{1}{2}\sqrt{1-\epsilon})^2}{(1-\alpha)(1-\sqrt{1-\epsilon})}\Big)\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v_h/\partial n_f]\!]\big\|_{L^2(f)}^2.\end{aligned}$$

This inequality provides the desired result provided that

$$\sigma \geq \sigma_\alpha := 1 + \frac{C_1}{(1-\alpha)(1-\sqrt{1-\epsilon})}. \tag{5.11}$$

□

Lemma 8. *There holds*

$$|B_h(v,w_h)| \leq C\|v\|_h\|w_h\|_h$$

for all $v \in V + V_h$ and $w_h \in V_h$.

Proof. We assume that $\lambda > 0$; the other case is proved in a similar fashion.

Applying the definition of $B_h(\cdot,\cdot)|$ together with the Cauchy–Schwarz inequality yields

$$\begin{aligned}|B_h(v,w_h)| \leq&\|\mathcal{L}_\gamma v\|_{L^2(\mathcal{T}_h)}\|\mathcal{L}_\lambda w_h\|_{L^2(\mathcal{T}_h)}\\
&+\sigma\Big(\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v/\partial n_f]\!]\big\|_{L^2(f)}^2\Big)^{1/2}\Big(\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial w_h/\partial n_f]\!]\big\|_{L^2(f)}^2\Big)^{1/2}.\end{aligned}$$

We easily find that

$$\begin{aligned}\|\mathcal{L}_\gamma v\|_{L^2(\mathcal{T}_h)}^2 &\le 2\|\gamma\|_{L^\infty(\Omega)}^2\big(\|A\|_{L^\infty(\Omega)}^2\|D^2 v\|_{L^2(\mathcal{T}_h)}^2 + \|\boldsymbol{b}\|_{L^\infty(\Omega)}^2\|\nabla v\|_{L^2(\Omega)}^2 + \|c\|_{L^\infty(\Omega)}^2\|v\|_{L^2(\Omega)}^2\big)\\ &\le 2\|\gamma\|_{L^\infty(\Omega)}^2 \max\{\|A\|_{L^\infty(\Omega)}^2, \|\boldsymbol{b}\|_{L^\infty(\Omega)}^2/(2\lambda), \|c\|_{L^\infty(\Omega)}^2/\lambda^2\}\|v\|_h^2,\end{aligned}$$

and

$$\|\mathcal{L}_\lambda w_h\|_{L^2(\mathcal{T}_h)}^2 \le 2\big(\|\Delta w_h\|_{L^2(\mathcal{T}_h)}^2 + \lambda^2\|w_h\|_{L^2(\Omega)}^2\big) \le 2d\|w_h\|_h^2.$$

Thus, we find that

$$\begin{aligned}|B_h(v,w_h)| &\le C_*\|v\|_h\|w_h\|_h + \sigma\Big(\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial v/\partial n_f]\!]\big\|_{L^2(f)}^2\Big)^{1/2}\Big(\sum_{f\in\mathcal{F}_h^I} h_f^{-1}\big\|[\![\partial w_h/\partial n_f]\!]\big\|_{L^2(f)}^2\Big)^{1/2}\\ &\le \big(\sigma + C_*\big)\|v\|_h\|w_h\|_h,\end{aligned}$$

with

$$C_* = 2\sqrt{d}\|\gamma\|_{L^\infty(\Omega)}\max\{\|A\|_{L^\infty(\Omega)}, \|\boldsymbol{b}\|_{L^\infty(\Omega)}/\sqrt{2\lambda}, \|c\|_{L^\infty(\Omega)}/\lambda\}.$$

□

Theorem 3. *Suppose that the solution to* (**??**) *has regularity* $u \in H^s(\Omega)$ *for some* $2 \le s \le k+1$, *and let* $u_h \in V_h$ *satisfy* (**??**). *Then there holds*

$$\|u - u_h\|_h^2 \le C \inf_{v_h\in V_h}\|u - v_h\|_h^2 \le \sum_{T\in\mathcal{T}_h} h_T^{2s-4}\|u\|_{H^s(T)}^2 \tag{5.12}$$

Proof. The first inequality is a result of Lemmas **??**–**??** and Cea's Lemma. The second inequality follows from standard approximation theory and scaling [**?**]. □

6. Applications to the Hamilton–Jacobi–Bellman Equation

In this section, we extend the method and analysis of Section **??**, and consider numerical approximations of the Hamilton–Jacobi–Bellman equation:

$$\mathcal{F}[u] := \sup_{\alpha\in\mathcal{A}}[\mathcal{L}^\alpha u - g^\alpha] = 0 \qquad \text{in } \Omega, \tag{6.1a}$$

$$u = 0 \qquad \text{on } \partial\Omega, \tag{6.1b}$$

where $\mathcal{A}$ is a compact metric space, and $\{\mathcal{L}^\alpha\}_{\alpha\in\mathcal{A}}$ is a family of second–order operators in non–divergence form, namely,

$$\mathcal{L}^\alpha v = A^\alpha : D^2 v + \boldsymbol{b}^\alpha\cdot\nabla v - c^\alpha v. \tag{6.2}$$

As in the previous section, $\Omega\subset\mathbb{R}^d$ is a convex domain, but we assume the coefficients satisfy the stronger conditions $\boldsymbol{b}^\alpha \in [C(\bar\Omega)]^d$, $c^\alpha\in C(\bar\Omega)$, and $A^\alpha \in [C(\bar\Omega)]^{d\times d}$ for all $\alpha\in\mathcal{A}$, and that the data is continuous with respect to α, e.g., the function $\alpha\to A^\alpha(x)$ is continuous on $\mathcal{A}$ for fixed $x\in\bar\Omega$. In addition, we assume that c^α is nonnegative, and that the family of operators $\{A^\alpha\}_{\alpha\in\mathcal{A}}$ is uniformly positive definite and uniformly satisfies the Cordes condition with respect to α, i.e.,

$$\nu|\boldsymbol{\xi}|^2 \le \sum_{i,j=1}^d A_{ij}^\alpha(x)\xi_i\xi_j \le \bar\nu|\boldsymbol{\xi}|^2 \quad \forall\boldsymbol{\xi}\in\mathbb{R}^d, \forall x\in\Omega, \forall\alpha\in\Lambda, \tag{6.3}$$

and if $c^\alpha\not\equiv 0$ or $\boldsymbol{b}^\alpha \not\equiv 0$ for some $\alpha\in\mathcal{A}$, there exists $\lambda > 0$ and $\epsilon\in(0,1)$ such that

$$\frac{|A^\alpha|^2 + |\boldsymbol{b}^\alpha|^2/2\lambda + (c^\alpha/\lambda)^2}{(\operatorname{tr}A^\alpha + c^\alpha/\lambda)^2} \le \frac{1}{d+\epsilon} \quad \forall x\in\Omega. \tag{6.4a}$$

Otherwise, if $c^\alpha\equiv 0$ and $\boldsymbol{b}^\alpha\equiv 0$ for all $\alpha\in\mathcal{A}$, there exists $\epsilon\in(0,1)$ such that

$$\frac{|A^\alpha|^2}{(\operatorname{tr}A^\alpha)^2} \le \frac{1}{d-1+\epsilon} \quad \forall x\in\Omega. \tag{6.4b}$$

Under these conditions, there holds the following result [**?**, Theorem 3].

Theorem 4. *Under the given conditions, there exists a unique strong solution $u \in V$ to (**??**).*

We refer to [**?**, Theorem 3] for a complete proof of this result. Here, we just state the main ideas of the proof.

Analogous to (**??**), for each $\alpha \in \mathcal{A}$, we define the (positive) function

$$\gamma^\alpha := \frac{\operatorname{tr} A^\alpha + c^\alpha/\lambda}{|A^\alpha|^2 + |\boldsymbol{b}^\alpha|^2/2\lambda + (c^\alpha/\lambda)^2} \quad \text{in } \Omega, \tag{6.5}$$

and in the special case where $\boldsymbol{b}^\alpha \equiv 0$ and $c^\alpha \equiv 0$ for all $\alpha \in \mathcal{A}$, we set

$$\gamma^\alpha := \frac{\operatorname{tr} A^\alpha}{|A^\alpha|^2} \quad \text{in } \Omega.$$

The Cordes condition and the uniform ellipticity of A shows that there exists γ_0 such that $\gamma \geq \gamma_0$ for all $\alpha \in \mathcal{A}$.

Define the operator $\mathcal{F}_\gamma : H^2(\Omega) \to L^2(\Omega)$ by

$$\mathcal{F}_\gamma[u] := \sup_{\alpha\in\mathcal{A}} \gamma^\alpha \big(\mathcal{L}^\alpha u - g^\alpha\big).$$

With $\mathcal{L}_\lambda$ defined by (**??**), one concludes that $u \in V$ is a strong solution to (**??**) if and only if

$$\langle \mathcal{M}[u], v\rangle := \int_\Omega \mathcal{F}_\gamma[u]\mathcal{L}_\lambda v \, dx = 0 \qquad \forall v \in V, \tag{6.6}$$

where $\langle\cdot,\cdot\rangle$ is the dual pairing between V^* and V. Continuity of the data and the compactness of $\mathcal{A}$ implies that $\mathcal{M}$ is Lipschitz continuous, and, by using the Cordes condition, one can show that $\mathcal{M}$ is strongly monotone. The Browder–Minty Theorem then gives the existence and uniqueness of $u \in V$ satisfying (**??**), and thus (**??**). We adopt this framework in the finite element analysis below.

6.1. **Finite element method.** Define the operator $\mathcal{M}_h : V_h + V \to V_h^*$ such that

$$\langle \mathcal{M}_h[w], v\rangle := \sum_{T\in\mathcal{T}_h} \int_T \mathcal{F}_\gamma[w]\mathcal{L}_\lambda v_h \, dx + \sigma \sum_{f\in\mathcal{F}_h^I} h_f^{-1} \int_f [\![\partial w/\partial n_f]\!][\![\partial v_h/\partial n_f]\!]\, ds.$$

We consider the following finite element method for problem (**??**): Find $u_h \in V_h$ such that

$$\langle \mathcal{M}_h[u_h], v_h\rangle = 0 \qquad \forall v_h \in V_h. \tag{6.7}$$

Note that since the exact (strong) solution to (**??**) has regularity $u \in H^2(\Omega)$, and therefore, since u satisfies (**??**) almost everywhere in Ω, we conclude that $\langle \mathcal{M}_h[u], v_h\rangle = 0$ for all $v_h \in V_h$, i.e., the method is consistent.

Lemma 9. *Let Ω be a bounded convex polygonal domain of $\mathbb{R}^d$. Suppose that (**??**) holds, and that the coefficients and continuous and satisfy the Cordes condition (**??**). Then there holds the following inequality:*

$$|\mathcal{F}_\gamma[v] - \mathcal{F}_\gamma[z] - \mathcal{L}_\lambda(v-z)| \leq \sqrt{1-\epsilon}\sqrt{|D^2(v-z)|^2 + 2\lambda|\nabla(v-z)|^2 + \lambda^2|(v-z)|^2}. \tag{6.8}$$

Proof. Using Lemma **??**, we have

$$\big|\gamma^\alpha\mathcal{L}^\alpha w - \mathcal{L}_\lambda w\big| \leq \sqrt{1-\epsilon}\sqrt{|D^2 w|^2 + 2\lambda|\nabla w|^2 + \lambda^2|w|^2} \qquad \forall \alpha \in \mathcal{A},$$

and therefore

$$\sup_{\alpha\in\mathcal{A}} \big|\gamma^\alpha\mathcal{L}^\alpha w - \mathcal{L}_\lambda w\big| \leq \sqrt{1-\epsilon}\sqrt{|D^2 w|^2 + 2\lambda|\nabla w|^2 + \lambda^2|w|^2}.$$

It then follows, with $w = v - z$, that

$$\begin{aligned} |\mathcal{F}_\gamma[v] - \mathcal{F}_\gamma[z] - \mathcal{L}_\lambda w| &= \big|\sup_{\alpha\in\mathcal{A}}(\gamma^\alpha(L^\alpha v - g^\alpha)) - \sup_{\alpha\in\mathcal{A}}(\gamma^\alpha(\mathcal{L}^\alpha z - g^\alpha) - \mathcal{L}_\lambda w\big| \\ &\leq \sup_{\alpha\in\mathcal{A}} |\gamma^\alpha\mathcal{L}^\alpha w - \mathcal{L}_\lambda w| \end{aligned}$$

$$\leq \sqrt{1-\epsilon}\sqrt{|D^2w|^2+2\lambda|\nabla w|^2+\lambda^2|w|^2}.$$

□

Theorem 5. *Let Ω be a bounded convex polygonal domain of $\mathbb{R}^d$, let $\mathcal{T}_h$ be a simplicial, conforming, and shape-regular mesh of Ω without hanging nodes. Suppose that the coefficients are continuous in $\overline{\Omega}$ and satisfy the Cordes condition (??). Then there exist a unique solution $u_h \in V_h$ satisfying (??) provided σ is sufficiently large. Moreover, there holds*

$$\|u-u_h\|_h \leq C \inf_{v_h\in V_h}\|u-v_h\|_h \leq C\sum_{T\in\mathcal{T}_h} h_T^{2s-4}|u|_{H^s(T)} \tag{6.9}$$

provided that $u \in H^s(\Omega)$ for some $2 \leq s \leq k+1$.

Proof. Let $v_h, z_h \in V_h$, and set $w_h = v_h - z_h$, then by the Cauchy-Schwarz inequality, we have

$$\begin{aligned}\langle\mathcal{M}[v_h],w_h\rangle-\langle\mathcal{M}[z_h],w_h\rangle &= \sum_{T\in\mathcal{T}_h}\int_T\left(\mathcal{F}_\gamma[v_h]-\mathcal{F}_\gamma[z_h]-\mathcal{L}_\lambda w_h\right)\mathcal{L}_\lambda w_h\,dx \\ &\quad + \|\mathcal{L}_\lambda w\|^2_{L^2(\mathcal{T}_h)}+\sigma\sum_{f\in\mathcal{F}_h^I}h_f^{-1}\int_f [\![\partial w_h/\partial n_f]\!][\![\partial w_h/\partial n_f]\!]\,ds\end{aligned} \tag{6.10}$$

Continuing as in the proof of Lemma ??, we conclude that, for any $\alpha \in (0,1)$, we have

$$\langle\mathcal{M}[v_h],w_h\rangle-\langle\mathcal{M}[z_h],w_h\rangle \geq \alpha\big(1-\sqrt{1-\epsilon}\big)\|w_h\|_h^2,$$

provided that (??) is satisfied; thus, $\mathcal{M}_h$ is strongly monotone. Continuing as in Lemma ??, we also conclude that $\mathcal{M}_h$ is Lipschitz continuous (with respect to $\|\cdot\|_h$). By the Browder-Minty theorem there exist a unique solution $u_h \in V_h$ to (??). Finally, the error estimate (??) follows from the consistency of the scheme, the monotonicity and Lipschitz continuity of $\mathcal{M}_h$, and standard interpolation estimates. □

Remark 8. *To implement the method, we use Howard's algorithm to solve the nonlinear system; see Algorithm ??. By [?] (also see [?, Section 5.3]), Howard's algorithm converges superlinearly to u_h with a good initial guess α_0.*

Algorithm 1 *Howard's algorithm*

Initialize $\alpha_0 \in \mathcal{A}$,
while $i \geq 0$ ***do***
  Find u_h^i *such that* $\forall v_h \in V_h$,

$$\sum_{T\in\mathcal{T}_h}\int_T \gamma^{\alpha_i}(\mathcal{L}^{\alpha_i}u_h^i-g^{\alpha_i})\mathcal{L}_\lambda v_h\,dx+\sigma\sum_{f\in\mathcal{F}_h^I}h_f^{-1}\int_f[\![\partial u_h^i/\partial n_f]\!][\![\partial v_h/\partial n_f]\!]\,ds=0.$$

  if $i \geq 1$ *and* $\|u_h^i-u_h^{i-1}\|_h \leq$ *tolerance* ***then*** *Stop.*
  end if
  $\alpha_{i+1} = argmax_{\alpha\in\mathcal{A}}(\mathcal{L}^\alpha u_h^i - g^\alpha)$,
  $i = i+1$.
end while

7. Numerical experiments

In this section we perform some numerical experiments and test the accuracy of the finite element methods for linear and nonlinear problems in non-divergence form. The penalty parameter is taken to be $\sigma = 10$ in all experiments.

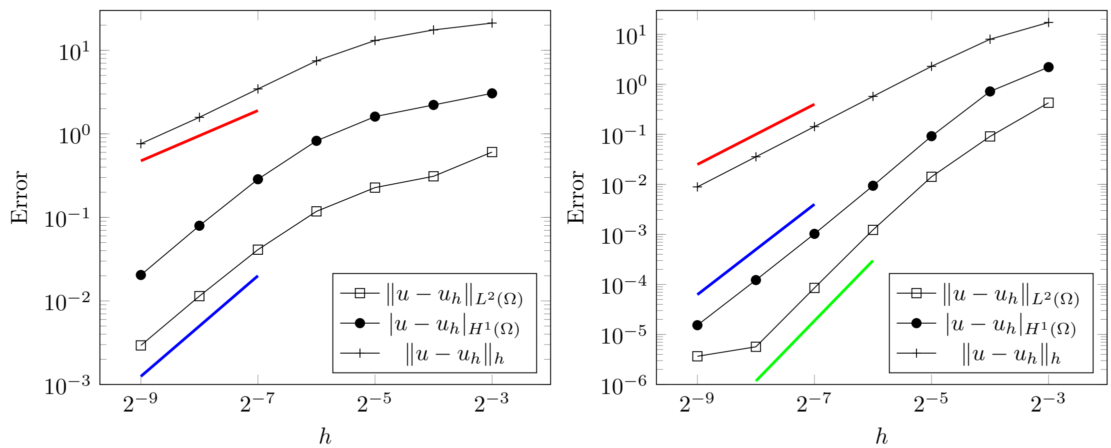

FIGURE 2. Test 1: Convergence plot of the two-dimensional linear problem with $k = 2$ (left) and $k = 3$ (right). The red reference line has slope $(k-1)$, the blue reference line has slope k, and the green reference line has slope $(k+1)$. The behavior of the L^2 error in the case $k = 3$ (right) is due to round–off error.

Test 1. In the first experiment we solve the linear problem (**??**) in two dimensions on the domain $\Omega = (-\pi, \pi)^2$. The coefficients are taken to be

$$A = 10I_2 + \frac{xx^t}{|x|^2}, \quad \boldsymbol{b} = \boldsymbol{0}, \quad c = 0. \tag{7.1}$$

The right-hand side function g is chosen such that the exact solution to (**??**)

$$u(x_1, x_2) = \sin(5x_1)\sin(5x_2)/(3x_1^2 + x_2^4 + 2). \tag{7.2}$$

It is easy to see that $9|\boldsymbol{\xi}|^2 \le \boldsymbol{\xi}^t A(x)\boldsymbol{\xi} \le 11|\boldsymbol{\xi}|^2$ for all $\boldsymbol{\xi} \in \mathbb{R}^2$, and therefore the Cordes condition is satisfied with $\epsilon = 0.9$ (cf. Remark **??**).

We compute the numerical scheme (**??**) for polynomial degrees $k = 2$ and $k = 3$ and report the resulting errors in Figure **??**. The figure clearly shows asymptotic $(k-1)$th order convergence in the H^2-type norm; this agrees with the theoretical results given in Theorem **??**. In addition, the experiments indicate that the method converges with optimal kth order convergence in the H^1 norm. The L^2 error converges with (sub-optimal) second order convergence when $k = 2$ and (optimal) fourth order convergence when $k = 3$.

Test 2. We again solve the linear problem (**??**) but in three dimensions with $\Omega = (-\pi, \pi)^3$, and with lower order terms:

$$A = 10I_3 + \frac{xx^t}{|x|^2}, \quad \boldsymbol{b} = (1\ 0\ 0)^t, \quad c = 10. \tag{7.3}$$

Note that $\operatorname{tr} A = 31$, $|A|^2 = 321$, and therefore

$$\frac{|A|^2 + |\boldsymbol{b}|^2/(2\lambda) + (c/\lambda)^2}{(\operatorname{tr} A + c/\lambda)^2} = \frac{321 + 1/(2\lambda) + (10/\lambda)^2}{(31 + 10/\lambda)^2}.$$

Taking (for example) $\lambda = 1/2$ yields

$$\frac{|A|^2 + |\boldsymbol{b}|^2/2\lambda + (c/\lambda)^2}{(\operatorname{tr} A + c/\lambda)^2} = \frac{722}{2601},$$

and therefore the Cordes condition is satisfied with $\epsilon = 435/722$ (cf. Definition **??**). In the numerical experiments, the right-hand side function g is chosen such that the exact solution to (**??**) is given

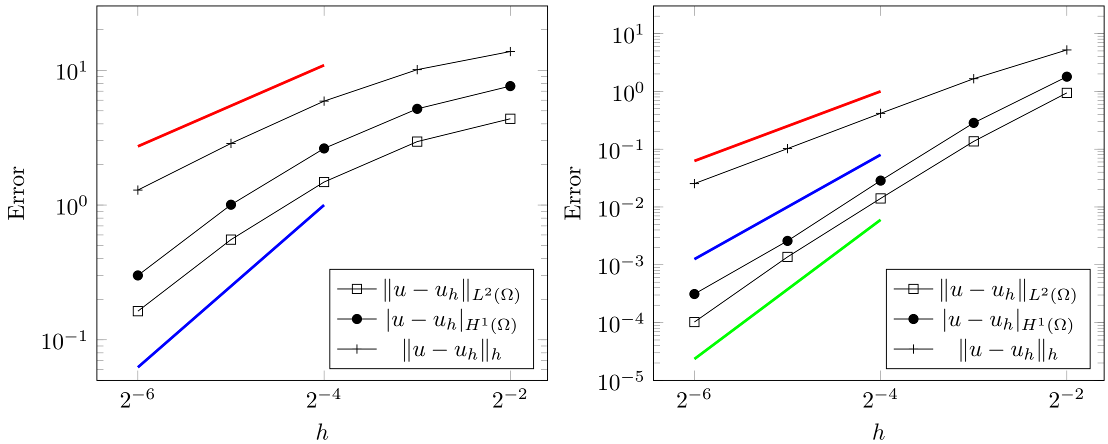

FIGURE 3. Test 2: Convergence plot of the three-dimensional linear problem with $k = 2$ (left) and $k = 3$ (right). The red reference line has slope $(k-1)$, the blue reference line has slope k, and the green reference line has slope $(k+1)$.

TABLE 1. The errors and rates of convergence for Test 2.

	h	$\|u-u_h\|_{L^2(\Omega)}$	rate	$\|u-u_h\|_{H^1(\Omega)}$	rate	$\|u-u_h\|_h$	rate
$k=2$	2^{-2}	4.3663		7.6303		13.769	
	2^{-3}	2.9558	0.56	5.1724	0.56	10.104	0.45
	2^{-4}	1.4819	1.00	2.6302	0.98	5.9030	0.78
	2^{-5}	0.5539	1.42	1.0053	1.39	2.8660	1.04
	2^{-6}	0.1631	1.76	0.3004	1.74	1.2913	1.15
$k=3$	2^{-2}	9.41E-01		1.79E+00		5.20E+00	
	2^{-3}	1.36E-01	2.79	2.85E-01	2.65	1.65E+00	1.65
	2^{-4}	1.40E-02	3.28	2.86E-02	3.32	4.17E-01	1.99
	2^{-5}	1.37E-03	3.36	2.59E-03	3.46	1.02E-01	2.03
	2^{-6}	1.02E-04	3.74	3.12E-04	3.05	2.53E-02	2.01

by

$$u(x_1,x_2,x_3) = \sin(5x_1)\sin(5x_2)\sin(5x_3)/(3x_1^2 + x_2^4 + 2).$$

The computed errors, listed in Figure **??** and Table **??**, show similar behavior as the previous two-dimensional experiments. Namely, we observe asymptotic $(k-1)$th order convergence in the H^2-type norm, and Table **??** indicate that

$$\|u-u_h\|_{H^1(\Omega)} = \mathcal{O}(h^k), \qquad \|u-u_h\|_{L^2(\Omega)} = \begin{cases} \mathcal{O}(h^2) & k=2, \\ \mathcal{O}(h^{k+1}) & k \ge 3. \end{cases}$$

Test 3. In these series of experiments, we solve the nonlinear Hamilton-Jacobi-Bellman problem (**??**) with $d=2$, $\Omega = (-\pi,\pi)^2$, $\mathcal{A} = \{1,2\}$, and

$$A^1 = \begin{pmatrix} 2 & 1/2 \\ 1/2 & 3/2 \end{pmatrix} + \frac{x_1}{|x_1|}\frac{x_2}{|x_2|}\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 3/2 & 1/2 \\ 1/2 & 2 \end{pmatrix} + \frac{x_1}{|x_1|}\frac{x_2}{|x_2|}\begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1 \end{pmatrix},$$
$$\boldsymbol{b}^1 = \boldsymbol{b}^2 = (1\ 0)^t, \qquad c^1 = c^2 = 1.$$

The source functions $\{g^\alpha\}_{\alpha\in\mathcal{A}}$ are chosen so that the solution of (**??**) is

$$u(x_1,x_2) = \sin(x_1)\sin(x_2). \tag{7.4}$$

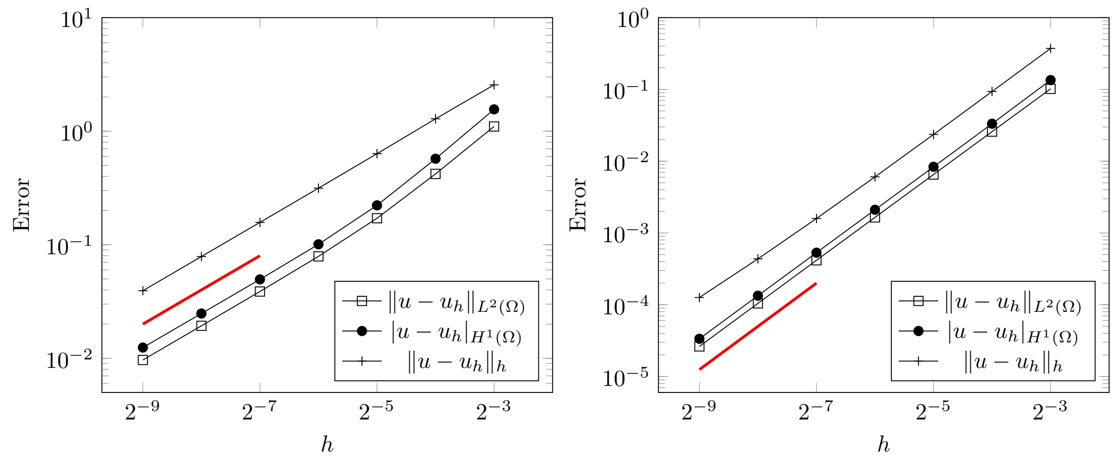

FIGURE 4. Test 3: Convergence plot of the two-dimensional nonlinear problem with $k=2$ (left) and $k=3$ (right). The red reference line has slope $(k-1)$.

With this data, we can verify that the Cordes condition (**??**) holds with $\lambda = 1$ and $\epsilon = 1/6$. Note that the matrices are discontinuous at the lines $x_1 = 0$ and $x_2 = 0$. Therefore the problem does not satisfy the conditions assumed in Theorems **??** and **??**. Nonetheless, the plots of the errors given in Figure **??** show that the method converges as the mesh is refined, albeit at sub-optimal rates. Namely, the numerical experiments indicate that the method converges with the following convergence rates:

$$\|u-u_h\|_{L^2(\Omega)} = \mathcal{O}(h^{k-1}), \quad \|u-u_h\|_{H^1(\Omega)} = \mathcal{O}(h^{k-1}), \quad \|u-u_h\|_h = \mathcal{O}(h^{k-1}).$$

REFERENCES

[1] Olivier Bokanowski, Stefania Maroso, and Hasnaa Zidani. Some convergence results for Howard's algorithm. *SIAM J. Numer. Anal.*, 47(4):3001–3026, 2009.

[2] Susanne C. Brenner, Thirupathi Gudi, and Li-yeng Sung. An a posteriori error estimator for a quadratic C^0-interior penalty method for the biharmonic problem. *IMA J. Numer. Anal.*, 30(3):777–798, 2010.

[3] Filippo Chiarenza, Michele Frasca, and Placido Longo. Interior $W^{2,p}$ estimates for nondivergence elliptic equations with discontinuous coefficients. *Ricerche Mat.*, 40(1):149–168, 1991.

[4] Philippe G. Ciarlet. *The finite element method for elliptic problems*, volume 40 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002. Reprint of the 1978 original [North-Holland, Amsterdam; MR0520174 (58 #25001)].

[5] Jim Douglas, Jr., Todd Dupont, Peter Percell, and Ridgway Scott. A family of C^1 finite elements with optimal approximation properties for various Galerkin methods for 2nd and 4th order problems. *RAIRO Anal. Numér.*, 13(3):227–255, 1979.

[6] Xiaobing Feng, Lauren Hennings, and Michael Neilan. Finite element methods for second order linear elliptic partial differential equations in non-divergence form. *Math. Comp.*, 86(307):2025–2051, 2017.

[7] Xiaobing Feng, Michael Neilan, and Schnake Stephan. Interior penalty discontinuous Galerkin methods for second order linear non-divergence form elliptic PDEs. submitted.

[8] Dietmar Gallistl. Variational formulation and numerical analysis of linear elliptic equations in nondivergence form with Cordes coefficients. *SIAM J. Numer. Anal.*, 55(2):737–757, 2017.

[9] Emmanuil H. Georgoulis, Paul Houston, and Juha Virtanen. An *a posteriori* error indicator for discontinuous Galerkin approximations of fourth-order elliptic problems. *IMA J. Numer. Anal.*, 31(1):281–298, 2011.

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

[11] Pierre Grisvard. *Elliptic problems in nonsmooth domains*, volume 69 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011. Reprint of the 1985 original [MR0775683], With a foreword by Susanne C. Brenner.

[12] Antonino Maugeri, Dian K. Palagachev, and Lubomira G. Softova. *Elliptic and parabolic equations with discontinuous coefficients*, volume 109 of *Mathematical Research*. Wiley-VCH Verlag Berlin GmbH, Berlin, 2000.

[13] Antonino Maugeri, Dian K. Palagachev, and Lubomira G. Softova. *Elliptic and parabolic equations with discontinuous coefficients*, volume 109 of *Mathematical Research*. Wiley-VCH Verlag Berlin GmbH, Berlin, 2000.
[14] Lin Mu and Xiu Ye. A simple finite element method for non-divergence form elliptic equations. *Int. J. Numer. Anal. Model.*, 14(2):306–311, 2017.
[15] Michael Neilan. Convergence analysis of a finite element method for second order non-variational elliptic problems. *J. Numer. Math.*, 2017. to appear.
[16] Michael Neilan, Abner J. Salgado, and Wujun Zhang. Numerical analysis of strongly nonlinear pdes. *Acta Numerica*, 26:137–303, 2017.
[17] Ricardo H. Nochetto and Wujun Zhang. Discrete ABP estimate and convergence rates for linear elliptic equations in non-divergence form. *Found. Comput. Math.*, 2017. to appear.
[18] Abner J. Salgado and Wujun Zhang. Finite element approximation of the Isaacs equation. arXiv:1512.09091, 2017.
[19] Iain Smears. Nonoverlapping domain decomposition preconditioners for discontinuous Galerkin approximations of Hamilton-Jacobi-Bellman equations. *J. Sci. Comput.*, 74(1):145–174, 2018.
[20] Iain Smears and Endre Süli. Discontinuous Galerkin finite element approximation of nondivergence form elliptic equations with Cordès coefficients. *SIAM J. Numer. Anal.*, 51(4):2088–2106, 2013.
[21] Iain Smears and Endre Süli. Discontinuous Galerkin finite element approximation of Hamilton-Jacobi-Bellman equations with Cordes coefficients. *SIAM J. Numer. Anal.*, 52(2):993–1016, 2014.
[22] Iain Smears and Endre Süli. Discontinuous Galerkin finite element methods for time-dependent Hamilton-Jacobi-Bellman equations with Cordes coefficients. *Numer. Math.*, 133(1):141–176, 2016.
[23] Tatyana Sorokina and Hal Schenck. Subdivision and spline spaces. *Constr. Approx.*, 2017. to appear.
[24] Chunmei Wang and Junping Wang. A primal-dual weak Galerkin finite element method for second order elliptic equations in non-divergence form. *Math. Comp.*, 2017. to appear.
[25] A. J. Worsey and G. Farin. An n-dimensional Clough-Tocher interpolant. *Constr. Approx.*, 3(2):99–110, 1987.

Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260
E-mail address: neilan@pitt.edu
E-mail address: mow11@pitt.edu