COMMUNICATIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 4, Pages 607–640 (October 9, 2024)
https://doi.org/10.1090/cams/39

A NONLINEAR LEAST-SQUARES CONVEXITY ENFORCING C^0 INTERIOR PENALTY METHOD FOR THE MONGE–AMPÈRE EQUATION ON STRICTLY CONVEX SMOOTH PLANAR DOMAINS

SUSANNE C. BRENNER, LI-YENG SUNG, ZHIYU TAN, AND HONGCHAO ZHANG

ABSTRACT. We construct a nonlinear least-squares finite element method for computing the smooth convex solutions of the Dirichlet boundary value problem of the Monge-Ampère equation on strictly convex smooth domains in $\mathbb{R}^2$. It is based on an isoparametric C^0 finite element space with exotic degrees of freedom that can enforce the convexity of the approximate solutions. *A priori* and *a posteriori* error estimates together with corroborating numerical results are presented.

1. INTRODUCTION

The Monge-Ampère equation is a fundamental partial differential equation in geometric analysis pertaining to affine geometry (cf. [3, 4, 12, 36, 40, 56, 64, 93]). In this paper we consider the Dirichlet boundary problem of the simplest Monge-Ampère equation in two dimensions where the right-hand side is a function of the spatial variables. It is a stepping stone towards Monge-Ampère equations with more general right-hand sides and boundary conditions that appear in applications such as the prescribed Gaussian curvature problem and optimal transport.

Let $\Omega \subset \mathbb{R}^2$ be a bounded strictly convex smooth domain. The Dirichlet boundary value problem for the Monge-Ampère equation is given by

$$\det D^2 u = \psi \quad \text{in } \Omega, \tag{1.1a}$$

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

where D^2u is the Hessian of u. We assume that

$$\phi \in H^4(\Omega), \tag{1.2}$$

$$\psi \in H^2(\Omega) \text{ is strictly positive on } \bar{\Omega}, \tag{1.3}$$

Received by the editors July 12, 2023, and, in revised form, April 27, 2024, and August 5, 2024.

2020 *Mathematics Subject Classification.* Primary 65N12, 65N15, 65N30, 65K10, 90C30, 35J60.

Key words and phrases. Monge-Ampère equation, strictly convex smooth domains, nonlinear least-squares, convexity enforcing, C^0 interior penalty method, enhanced cubic Lagrange finite elements, isoparametric, *a priori* and *a posteriori* error estimates, active set algorithm.

The work of the first two authors was supported in part by the National Science Foundation under Grant No. DMS-19-13035 and Grant No. DMS-22-08404. The work of the third author was supported in part by the Fundamental Research Funds for the Central Universities under Grant No. 20720240133. The work of the fourth author was supported in part by the National Science Foundation under Grant No. DMS-2110722 and Grant No. DMS-2309549.

The third author is the corresponding author.

and

(1.4)

the boundary value problem (1.1) has a unique strictly convex solution $u \in H^4(\Omega)$.

Our goal is to design a finite element method that can capture such solutions.

Remark 1.1. The assumptions (1.2)–(1.4) are satisfied if $\psi \in C^3(\bar{\Omega})$ is strictly positive on $\bar{\Omega}$ and $\phi \in C^{4,\delta}(\bar{\Omega})$ for some $\delta \in (0,1)$ (cf. [37, p.371, Remark 2])). The smoothness and strict convexity of $\partial\Omega$ are crucial for obtaining the *a priori* estimates needed for establishing the existence of the solution. This is the motivation for considering (1.1) on domains that are not polygons.

Remark 1.2. Throughout this paper we will follow the standard notation for differential operators, function spaces and norms that can be found for example in [1, 30, 41].

Remark 1.3. The strict convexity of u means that there exists a positive constant $\alpha_\sharp$ such that

$$\alpha_\sharp|\xi|^2 \le \xi^t D^2u(x)\xi \qquad \forall\, x \in \Omega, \xi \in \mathbb{R}^2. \tag{1.5}$$

The regularity of u also ensures that there exists a positive constant $\beta_\sharp$ such that

$$\xi^t D^2u(x)\xi \le \beta_\sharp|\xi|^2 \qquad \forall\, x \in \Omega, \xi \in \mathbb{R}^2. \tag{1.6}$$

The numerical solution of (1.1) is challenging due to the fully nonlinear nature of (1.1a) and the convexity condition on the solution. There are different approaches for different solution classes that can be found for example in [2, 5–11, 13–18, 26–28, 34, 35, 39, 44–48, 52–55, 57–62, 68–75, 78–81, 84–90]. Additional references can be found in the two review articles [51, 82].

Our approach was first proposed in [32] for (1.1) on a convex polygonal domain (so that basic finite element technology can be employed), where the boundary value problem is reformulated as a nonlinear least-squares problem with equality constraints (from the boundary condition) and inequality constraints (from the convexity of the solution). The discrete problem is posed on a convexity enforcing finite element space where the equality and inequality constraints become simple box constraints that allow the least-squares problem to be solved efficiently. The construction of the convexity enforcing finite element space is based on the observations that the solution of (1.1) under the condition (1.3) is strictly convex if and only if $\Delta u \ge 0$ in Ω, and that one can use pointwise values of the Laplacian of a finite element function as degrees of freedom by going beyond the classical definition of a finite element in [41]. The elementwise convexity of the discrete solution leads to an elliptic problem in nondivergence form in the error analysis. The *a priori* and *a posteriori* analyses in [32] take advantage of the results in [83, 91] where discontinuous Galerkin methods for elliptic problems in nondivergence form on convex domains were investigated.

However the assumption that Ω is a polygonal domain is inconsistent with the conditions for the well-posedness of the boundary value problem (1.1) mentioned in Remark 1.1. We address this problem in the current paper by extending the methodology in [32] to strictly convex smooth domains. The challenges are twofold. In the case of polygonal domains, it is straightforward to use an interpolant of the exact solution of (1.1) as the boundary condition of the discrete problem, which is crucial for obtaining the *a priori* bounds used in the error analysis. Here we need to construct an isoparametric mesh carefully so that we can use (1.1b) to impose an interpolation of the solution of (1.1) as the discrete boundary condition and to obtain the correct estimate needed for establishing the elementwise convexity for a solution of the discrete problem. Secondly we need to extend many results in [83, 91] for discontinuous Galerkin methods for polynomial finite element spaces to isoparametric finite element spaces. Since the estimates for isoparametric finite element methods in the literature mostly only pertain to problems in $H^1(\Omega)$, we have to develop several new estimates for our isoparametric finite element method that involve the Sobolev space $H^2(\Omega)$.

The rest of the paper is organized as follows. The isoparametric finite element space is constructed in Section 2 and the discrete problem is presented in Section 3. The convergence analysis is carried out in Section 4, followed by numerical results in Section 5 and some concluding remarks in Section 6. Appendices A–D contain the derivations of several technical results.

Throughout the paper we will use C (with or without subscripts) to denote a generic positive constant independent of the mesh size. We also use the notation $A \lesssim B$ to represent the statement $A \leq$ (constant) B, where the positive constant is independent of the mesh size. The notation $A \approx B$ represents the statements that $A \lesssim B$ and $B \lesssim A$.

2. An isoparametric finite element space

For a given mesh parameter h, we will construct a convex domain Ω_h that approximates Ω, a special cubic isoparametric triangulation $\mathcal{T}_h$ of Ω_h and a finite element space V_h associated with $\mathcal{T}_h$. They will be used in Section 3 to define the discrete problem. We will use $\hat{T}$ to denote the reference (closed) simplex with vertices $(0,0)$, $(1,0)$ and $(0,1)$. Given two points $p_1 = (a_1, b_1)$ and $p_2 = (a_2, b_2)$, the 2×1 vector with first component $a_2 - a_1$ and second component $b_2 - b_1$ will be denoted by $\boldsymbol{p}_2 - \boldsymbol{p}_1$. The Euclidean norm is denoted by $|\cdot|$.

2.1. The domain Ω_h and the triangulation $\mathcal{T}_h$.

We begin with a convex polygon $\tilde{\Omega}_h$ equipped with a quasi-uniform triangulation $\tilde{\mathcal{T}}_h$ (cf. Figure 2.1) such that

- the vertices of $\tilde{\Omega}_h$ belong to $\partial\Omega$,
- each edge of $\tilde{\Omega}_h$ is also the edge of a triangle in $\tilde{\mathcal{T}}_h$,
- each triangle in $\tilde{\mathcal{T}}_h$ has at most two vertices on $\partial\Omega$.

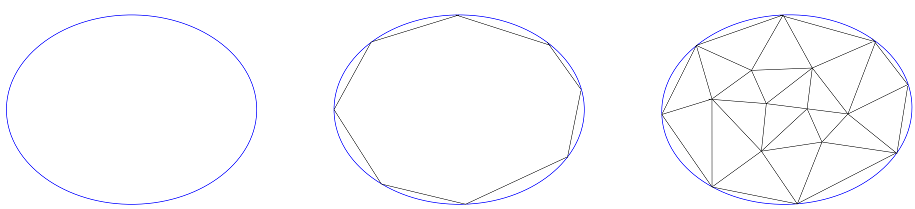

FIGURE 2.1. Ω, $\tilde{\Omega}_h$ and $\tilde{\mathcal{T}}_h$

The domain Ω_h and the triangulation $\mathcal{T}_h$ are obtained by modifying the triangles in $\tilde{\mathcal{T}}_h$ that have two vertices on $\partial\Omega$.

Recall (cf. [30,41]) the degrees of freedom (dofs) of the cubic Lagrange finite element on $\hat{T}$ are given by the values of a function $\hat{v} \in P_3(\hat{T})$ at the points $\hat{p}_1 = (0,0)$, $\hat{p}_2 = (1,0)$, $\hat{p}_3 = (0,1)$, $\hat{p}_4 = (\frac{2}{3},\frac{1}{3})$, $\hat{p}_5 = (0,\frac{2}{3})$, $\hat{p}_6 = (\frac{1}{3},0)$, $\hat{p}_7 = (\frac{1}{3},\frac{2}{3})$, $\hat{p}_8 = (0,\frac{1}{3})$, $\hat{p}_9 = (\frac{2}{3},0)$ and $\hat{p}_{10} = (\frac{1}{3},\frac{1}{3})$ (cf. Figure 2.2, where the dofs are represented by the solid dots).

We can define a modified cubic Lagrange finite element on $\hat{T}$ (cf. Figure 2.2) by replacing $\hat{v}(\hat{p}_4)$ (resp., $\hat{v}(\hat{p}_7)$) with the directional derivative of $\hat{v}$ at $\hat{p}_2$ (resp., $\hat{p}_3$) in the direction of $\overrightarrow{\hat{p}_2\hat{p}_3}$ (resp., $\overrightarrow{\hat{p}_3\hat{p}_2}$). The new dofs are presented by the arrows in Figure 2.2.

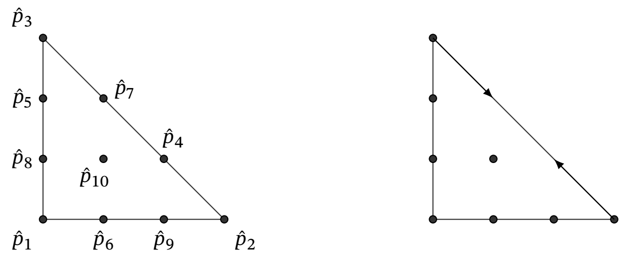

FIGURE 2.2. Cubic Lagrange finite element (left) and modified cubic Lagrange finite element (right)

We can now define the domain Ω_h and its triangulation $\mathcal{T}_h$ as follows. If the (closed) triangle $\tilde{T} \in \tilde{\mathcal{T}}_h$ has at most one vertex on $\partial\Omega$, then we include $T = \tilde{T}$ in $\mathcal{T}_h$ and take $\Phi_T : \hat{T} \longrightarrow T$ to be an affine isomorphism. If $\tilde{T} \in \tilde{\mathcal{T}}_h$ has two vertices (say p_2 and p_3) on $\partial\Omega$, then we replace $\tilde{T}$ with T, the image of $\hat{T}$ under the cubic polynomial map Φ_T (cf. Figure 2.3) which is defined below in terms of the dofs for the modified cubic Lagrange finite element.

(2.1a) $$\Phi_T(\hat{p}_i) = p_i \qquad \text{for } i = 1,2,3,5,6,8,9,$$

(2.1b) $$\Phi_T(\hat{p}_{10}) = p_{10} + \frac{1}{18}(|\mathbf{p_3} - \mathbf{p_2}|\boldsymbol{e}_{23} + |\mathbf{p_2} - \mathbf{p_3}|\boldsymbol{e}_{32}),$$

(2.1c) $$D\Phi_T(\hat{p}_2)(\hat{\mathbf{p}}_3 - \hat{\mathbf{p}}_2) = |\mathbf{p_3} - \mathbf{p_2}|\boldsymbol{e}_{23},$$

(2.1d) $$D\Phi_T(\hat{p}_3)(\hat{\mathbf{p}}_2 - \hat{\mathbf{p}}_3) = |\mathbf{p_2} - \mathbf{p_3}|\boldsymbol{e}_{32}.$$

Here the p_i's are the nodes associated with the standard cubic Lagrange element on $\tilde{T}$, $\boldsymbol{e}_{23}$ (resp., $\boldsymbol{e}_{32}$) is the unit tangent of $\partial\Omega$ at p_2 (resp., p_3) that points towards p_3 (resp., p_2), and $D\Phi_T$ is the Jacobian matrix of Φ_T.

Remark 2.1. The map defined by (2.1) is associated with a cubic Hermite isoparametric finite element space (cf. [42]) that is appropriate for error analysis involving the Sobolev space $H^2(\Omega)$.

The domain Ω_h is defined to be the interior of the union of $T \in \mathcal{T}_h$, which is a convex $C^{1,1}$ domain for h sufficiently small (which we assumed to be the case from here on). By construction $\mathcal{T}_h$ is automatically a triangulation of Ω_h. The element $T \in \mathcal{T}_h$ is a triangle if T has at most one vertex on $\partial\Omega$, otherwise T has one curved edge tangential to $\partial\Omega$ at its two vertices on $\partial\Omega$.

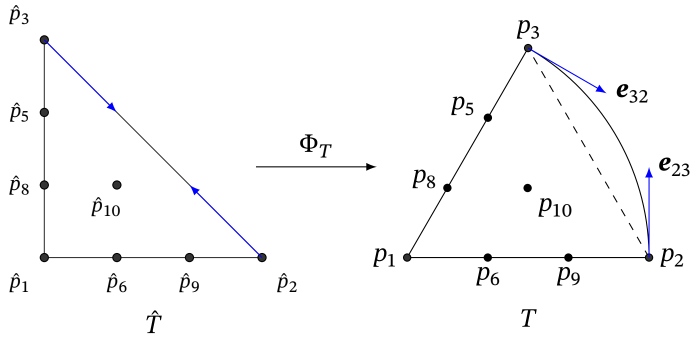

FIGURE 2.3. $\tilde{T}$ (bounded by the dotted line and the solid lines), T (bounded by the curve and the solid lines) and Φ_T

2.2. **The isoparametric map** Φ_T**.** Let $\tilde{T} \in \tilde{\mathcal{T}}_h$ have two vertices on $\partial\Omega$. The map Φ_T defined by (2.1a)–(2.1d) is identical to the one that appears in [42, Example 6, p.242–p.244]. (Figure 2.3 is identical to Fig. 5 on page 244 of [42] after relabelling.) The key to understand the behavior of Φ_T is by comparing it to an affine isomorphism $\Phi_{\tilde{T}}$: $\hat{T} \longrightarrow \tilde{T}$ defined by the following conditions:

$$\Phi_{\tilde{T}}(\hat{p}_i) = p_i \qquad \text{for } i = 1,2,3,5,6,8,9, \tag{2.2a}$$

$$\Phi_{\tilde{T}}(\hat{p}_{10}) = p_{10}, \tag{2.2b}$$

$$D\Phi_{\tilde{T}}(\hat{p}_2)(\hat{\mathbf{p}}_3 - \hat{\mathbf{p}}_2) = \mathbf{p_3} - \mathbf{p_2}, \tag{2.2c}$$

$$D\Phi_{\tilde{T}}(\hat{p}_3)(\hat{\mathbf{p}}_2 - \hat{\mathbf{p}}_3) = \mathbf{p_2} - \mathbf{p_3}, \tag{2.2d}$$

Comparing (2.1a)–(2.1d) and (2.2a)–(2.2d), we see that

$$\begin{aligned} \Phi_T(\hat{x}) = \Phi_{\tilde{T}}(\hat{x}) &+ \hat{\varphi}_4(\hat{x})\big[|\mathbf{p_3} - \mathbf{p_2}|\boldsymbol{e}_{23} - (\mathbf{p_3} - \mathbf{p_2})\big] \\ &+ \hat{\varphi}_7(\hat{x})\big[|\mathbf{p_2} - \mathbf{p_3}|\boldsymbol{e}_{32} - (\mathbf{p_2} - \mathbf{p_3})\big] \\ &+ (\hat{\varphi}_{10}(\hat{x})/18)(|\mathbf{p_3} - \mathbf{p_2}|\boldsymbol{e}_{23} + |\mathbf{p_2} - \mathbf{p_3}|\boldsymbol{e}_{32}) \qquad \forall\, \hat{x} \in \hat{T}, \end{aligned} \tag{2.3}$$

where $\hat{\varphi}_4, \hat{\varphi}_7, \hat{\varphi}_{10} \in P_3(\hat{T})$ are defined by the following conditions:

- $\hat{\varphi}_4(\hat{p}_i) = \hat{\varphi}_7(\hat{p}_i) = \hat{\varphi}_{10}(\hat{p}_i) = 0$ for $i = 1,2,3,5,6,8,9$,
- $\hat{\varphi}_4(\hat{p}_{10}) = \hat{\varphi}_7(\hat{p}_{10}) = 0$ and $\hat{\varphi}_{10}(\hat{p}_{10}) = 1$,
- $\nabla\varphi_4(\hat{p}_2) \cdot (\hat{\mathbf{p}}_3 - \hat{\mathbf{p}}_2) = 1$ and $\nabla\hat{\varphi}_7(\hat{p}_2) \cdot (\hat{\mathbf{p}}_3 - \hat{\mathbf{p}}_2) = \nabla\hat{\varphi}_{10}(\hat{p}_2) \cdot (\hat{\mathbf{p}}_3 - \hat{\mathbf{p}}_2) = 0$,
- $\nabla\hat{\varphi}_7(\hat{p}_3) \cdot (\hat{\mathbf{p}}_2 - \hat{\mathbf{p}}_3) = 1$ and $\nabla\hat{\varphi}_4(\hat{p}_3) \cdot (\hat{\mathbf{p}}_2 - \hat{\mathbf{p}}_3) = \nabla\hat{\varphi}_{10}(\hat{p}_3) \cdot (\hat{\mathbf{p}}_2 - \hat{\mathbf{p}}_3) = 0$,

i.e., $\hat{\varphi}_4$, $\hat{\varphi}_7$ and $\hat{\varphi}_{10}$ are the nodal basis functions associated with the dofs of the modified cubic Lagrange element (cf. Figure 2.2) represented by the arrow at $\hat{p}_2$, the arrow at $\hat{p}_3$ and the solid dot at $\hat{p}_{10}$ respectively.

Remark 2.2. The relation in (2.3) is also valid for $T \in \mathcal{T}_h$ that has at most one vertex on $\partial\Omega$, provided that we take $\boldsymbol{e}_{23}$ (resp., $\boldsymbol{e}_{32}$) to be the unit vector in the direction of $\mathbf{p_3} - \mathbf{p_2}$ (resp., $\mathbf{p_2} - \mathbf{p_3}$). In this case all three vectors $|\mathbf{p_3} - \mathbf{p_2}|\boldsymbol{e}_{23} - (\mathbf{p_3} - \mathbf{p_2})$, $|\mathbf{p_2} - \mathbf{p_3}|\boldsymbol{e}_{32} - (\mathbf{p_2} - \mathbf{p_3})$ and $|\mathbf{p_3} - \mathbf{p_2}|\boldsymbol{e}_{23} + |\mathbf{p_2} - \mathbf{p_3}|\boldsymbol{e}_{32}$ vanish and the discussion below applies to this case trivially.

Remark 2.3. The relation (2.3) holds for all $x \in \mathbb{R}^2$ since all the functions involved are polynomials.

It follows from Taylor's theorem that, with $\ell = |\mathbf{p_3} - \mathbf{p_2}| = |\mathbf{p_2} - \mathbf{p_3}|$,

$$|\boldsymbol{e}_{23} - \ell^{-1}(\boldsymbol{p}_3 - \boldsymbol{p}_2)| = O(\ell), \tag{2.4}$$

$$|\boldsymbol{e}_{32} - \ell^{-1}(\boldsymbol{p}_2 - \boldsymbol{p}_3)| = O(\ell), \tag{2.5}$$

$$\big|[\boldsymbol{e}_{23} - \ell^{-1}(\mathbf{p_3} - \mathbf{p_2})] - [\boldsymbol{e}_{32} - \ell^{-1}(\mathbf{p_2} - \mathbf{p_3})]\big| = O(\ell^2), \tag{2.6}$$

where the hidden constants only depend on $\partial\Omega$. Note that (2.4), (2.5) and the triangle inequality imply

$$|\boldsymbol{e}_{23} + \boldsymbol{e}_{32}| = O(\ell). \tag{2.7}$$

Note also that the maps Φ_T and $\Phi_{\tilde{T}}$ are defined on $\mathbb{R}^2$, and in particular, on the (closed) triangle $\hat{T}_\dagger$ with vertices $(1,1)$, $(-1,1)$ and $(1,-1)$ that is used in the construction of the finite element space in Section 2.5.

Combining (2.3)–(2.5) and (2.7), we find

$$\|\Phi_T - \Phi_{\tilde{T}}\|_{L^\infty(\hat{T}_\dagger)} = O(h_{\tilde{T}}^2), \tag{2.8}$$

$$\|D\Phi_T - D\Phi_{\tilde{T}}\|_{L^\infty(\hat{T}_\dagger)} = O(h_{\tilde{T}}^2), \tag{2.9}$$

where $h_{\tilde{T}}$ is the diameter of $\tilde{T}$ and the hidden constants only depend on $\partial\Omega$.

Since $\tilde{T} \subset T$, the estimate (2.8) immediately implies that

$$h_{\tilde{T}} \le h_T \le C h_{\tilde{T}}, \tag{2.10}$$

where h_T is the diameter of T and the positive constant C depends only on $\partial\Omega$.

Note that

$$\|D\Phi_{\tilde{T}}\|_{L^\infty(\mathbb{R}^2)} \approx h_{\tilde{T}} \quad \text{and} \quad \|(D\Phi_{\tilde{T}})^{-1}\|_{L^\infty(\mathbb{R}^2)} \approx h_{\tilde{T}}^{-1}, \tag{2.11}$$

where the hidden constants depend only on the shape regularity of $\tilde{T}$. Therefore we can conclude by the quasi-uniformity of $\tilde{\mathcal{T}}_h$, (2.9) and (2.11) that (for h sufficiently small) the map Φ_T is a C^∞ isomorphism between $\hat{T}_\dagger$ and $T_\dagger = \Phi_T(\hat{T}_\dagger)$, and that

$$\|D\Phi_T\|_{L^\infty(\hat{T}_\dagger)} \le Ch \qquad \forall\, T \in \mathcal{T}_h, \tag{2.12}$$

$$\|(D\Phi_T)^{-1}\|_{L^\infty(T_\dagger)} \le Ch^{-1} \qquad \forall\, T \in \mathcal{T}_h, \tag{2.13}$$

where the positive constant C is independent of h.

Finally we note that (2.3)–(2.5) and (2.7) also imply

$$|D\Phi_T|_{W^{1,\infty}(\hat{T}_\dagger)} \le Ch^2, \tag{2.14}$$

and it follows from (2.3), (2.6) and a direct calculation (cf. [42, p.242–p.244] and Appendix A) that

$$|D\Phi_T|_{W^{2,\infty}(\hat{T}_\dagger)} \le Ch^3, \tag{2.15}$$

$$\|\Delta\Phi_T\|_{L^\infty(\hat{T}_\dagger)} \le Ch^3. \tag{2.16}$$

The estimates (2.12)–(2.16) are crucial for the analyses in Sections 3 and 4.

2.3. **The sign of** $\Delta(u \circ \Phi_T)$**.** We first note that there exists a bounded linear extension map $\mathbb{E} : H^4(\Omega) \longrightarrow H^4(\mathbb{R}^2)$ (cf. [1]) such that $\mathbb{E}v = v$ on Ω, and we will denote $\mathbb{E}u$ again by u. We can also assume that u is strictly convex in a neighborhood $\Omega_\dagger$ of $\bar{\Omega}$.

The sign of $\Delta(u \circ \Phi_T)$ is addressed by Lemma 2.4.

Lemma 2.4. *The function $\Delta(u \circ \Phi_T)$ is strictly positive on $\hat{T}_\dagger$ for all $T \in \mathcal{T}_h$ provided that h is sufficiently small.*

Proof. In view of (2.8) we have, for sufficiently small h, $T_\dagger = \Phi_T(\hat{T}_\dagger) \subset \Omega_\dagger$ where u is strictly convex. Let $\phi_{T,1}$ and $\phi_{T,2}$ be the first and second components of Φ_T respectively. It follows from the chain rule that

$$D^2(u \circ \Phi_T)(\hat{x}) = D\Phi_T(\hat{x})^t (D^2 u)(\Phi_T(\hat{x})) D\Phi_T(\hat{x}) + \frac{\partial u}{\partial x_1}(\Phi_T(\hat{x})) D^2 \phi_{T,1}(\hat{x}) + \frac{\partial u}{\partial x_2}(\Phi_T(\hat{x})) D^2 \phi_{T,2}(\hat{x}) \qquad \forall\, \hat{x} \in \hat{T}_\dagger. \tag{2.17}$$

The proof is then completed by the observation that

$$\mathrm{tr}\big[D\Phi_T(\hat{x})^t (D^2 u)(\Phi_T(\hat{x})) D\Phi_T(\hat{x})\big] \gtrsim h^2 \qquad \forall\, x \in \hat{T}$$

by (1.5) and (2.13), and

$$\|\mathrm{tr} D^2 \phi_{T,i}\|_{L^\infty(\hat{T}_\dagger)} = \|\Delta(\phi_{T,i})\|_{L^\infty(\hat{T}_\dagger)} = O(h^3) \qquad \text{for } i = 1, 2$$

by (2.16). □

2.4. **The map** F_T**.** Let $\tilde{T} \in \tilde{\mathcal{T}}_h$ and T be the corresponding element in $\mathcal{T}_h$. The map

$$F_T = \Phi_T \circ \Phi_{\tilde{T}}^{-1} \tag{2.18}$$

(cf. (2.1) and (2.2)), which is a diffeomorphism between $\tilde{T}$ and T, is a useful tool for handling functions associated with the isoparametric mesh.

It follows from (2.3)–(2.5), (2.7), (2.10), (2.11) and the chain rule that

$$DF_T(\tilde{x}) = I + R(\tilde{x}) \qquad \forall\, \tilde{x} \in \tilde{T}, \tag{2.19}$$

where

$$\text{the components of } R(\tilde{x}) \text{ are quadratic polynomials in } \tilde{x} \tag{2.20}$$

and

$$\|R\|_{L^\infty(\tilde{T})} \le Ch. \tag{2.21}$$

($R = 0$ if $\tilde{T}$ has at most one vertex on $\partial\Omega$.)

In particular we have

$$\|DF_T\|_{L^\infty(\tilde{T})} \approx 1, \quad \|\det DF_T\|_{L^\infty(\tilde{T})} \approx 1 \quad \text{and} \quad \|DF_T^{-1}\|_{L^\infty(T)} \approx 1. \tag{2.22}$$

Using (2.12)–(2.15), (2.22) and the chain rule, we find

$$\|\zeta\|_{L^2(T)} \approx \|\zeta \circ F_T\|_{L^2(\tilde{T})} \qquad \forall\, \zeta \in L^2(T), \tag{2.23}$$

and

$$\sum_{j=1}^{k} |\zeta|_{H^j(T)} \approx \sum_{j=1}^{k} |\zeta \circ F_T|_{H^j(\tilde{T})} \qquad \forall\, \zeta \in H^k(T) \quad \text{and} \quad 1 \le k \le 4. \tag{2.24}$$

2.5. **The finite element space V_h.** The finite element space V_h associated with $\mathcal{T}_h$ is constructed in terms of enhanced cubic and modified cubic Lagrange finite elements. The space of shape functions for both elements is given by $P_3(\hat{T}) \oplus \varphi_{\hat{T}}^2 P_1(\hat{T})$, where $\varphi_{\hat{T}}(\hat{x}) = \hat{x}_1\hat{x}_2(1-\hat{x}_1-\hat{x}_2)$ is the cubic bubble function that vanishes on $\partial\hat{T}$.

For the enhanced cubic Lagrange element (cf. left of Figure 2.4), the dofs of $\hat{v} \in P_3(\hat{T}) \oplus \varphi_{\hat{T}}^2 P_1(\hat{T})$ are given by the 10 dofs of the cubic Lagrange element (cf. left of Figure 2.2) plus the values of $\Delta\hat{v}$ at the three vertices of the triangle $\hat{T}_\dagger$ with vertices $(1,1)$, $(-1,1)$ and $(1,-1)$, which are represented by the solid triangles.

Similarly, for the enhanced modified cubic Lagrange element (cf. right of Figure 2.4), the dofs of $\hat{v} \in P_3(\hat{T}) \oplus \varphi_{\hat{T}}^2 P_1(\hat{T})$ are given by the 10 dofs of the modified cubic Lagrange element (cf. right of Figure 2.2) plus the values of $\Delta\hat{v}$ at the three vertices of the triangle $\hat{T}_\dagger$ with vertices $(1,1)$, $(-1,1)$ and $(1,-1)$, which are represented by the solid triangles.

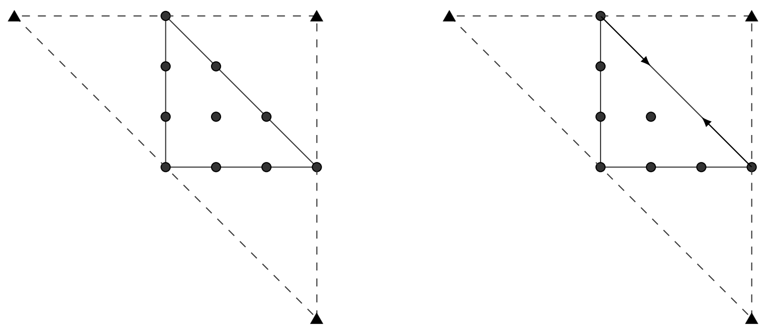

FIGURE 2.4. Enhanced cubic Lagrange element (left) and enhanced modified cubic Lagrange element (right)

It was proved in [32, Lemma 2.1] that a function $\hat{v} \in P_3(\hat{T}) \oplus \varphi_{\hat{T}}^2 P_1(\hat{T})$ is uniquely determined by the 13 dofs of the enhanced cubic Lagrange element, and the same arguments show that it is also uniquely determined by the 13 dofs of the enhanced modified cubic Lagrange element.

Remark 2.5. Since some of the dofs of the enhanced cubic and modified cubic Lagrange finite elements are associated with nodes outside the element domain, their constructions go beyond the classical constructions of finite elements.

Definition 2.6. A function v belongs to $V_h \subset H^1(\Omega_h)$ if and only if $v \circ \Phi_T$ belongs to $P_3(\hat{T}) \oplus \varphi_{\hat{T}}^2 P_1(\hat{T})$ for all $T \in \mathcal{T}_h$. The dofs of v on a triangle $T \in \mathcal{T}_h$ correspond to the dofs of the enhanced cubic Lagrange element (under the pullback by Φ_T) if T has at most one vertex on $\partial\Omega$, and to the dofs of the enhanced modified cubic Lagrange element if T has two vertices on $\partial\Omega$.

The number of global dofs for V_h is the sum of (i) the number of vertices of $\mathcal{T}_h$,(ii) $2\times$ (the number of edges in $\mathcal{T}_h$), and (iii) $4\times$ (the number of elements in $\mathcal{T}_h$).

3. THE DISCRETE PROBLEM

We assume that h is sufficiently small so that $\Phi_T : \hat{T}_\dagger \longrightarrow T_\dagger = \Phi_T(T_\dagger)$ is a C^∞ diffeomorphism for all $T \in \mathcal{T}_h$ and the estimates (2.12)–(2.16) are valid.

3.1. **The interpolation operator Π_h.** The operator $\Pi_h : H^4(\mathbb{R}^2) \longrightarrow V_h$ is defined by the condition that

$$(\Pi_h\zeta)\circ\Phi_T \text{ and } \zeta\circ\Phi_T \text{ have identical dofs for all } T\in\mathcal{T}_h, \tag{3.1}$$

where the dofs are the ones for the enhanced cubic Lagrange finite element if T has at most one vertex on $\partial\Omega$ and the ones for the enhanced modified cubic Lagrange finite element if T has two vertices on $\partial\Omega$.

Remark 3.1. The polynomial map Φ_T is actually defined on $\mathbb{R}^2$. Hence $\zeta\circ\Phi_T$ is defined on $\mathbb{R}^2$ and the 3 exotic dofs at the vertices of $\hat{T}_\dagger$ are well-defined.

The properties of Π_h are collected in Lemma 3.2, where D_h^2 denotes the piecewise Hessian operator with respect to $\mathcal{T}_h$, $\mathcal{E}_h^i$ is the set of the interior edges of $\mathcal{T}_h$, $|e|$ denotes the length of an edge e, and $[\![\partial v/\partial n]\!]$ denotes the jump of the normal derivative of v across an interior edge. The proof which involves standard arguments based on the Bramble-Hilbert lemma (cf. [19, 49]) and (2.23)–(2.24) is omitted. Details for similar estimates can be found in [42, Theorem 1].

Lemma 3.2. *The following estimates are valid for Π_h:*

(3.2)
$$\begin{aligned}\|\zeta-\Pi_h\zeta\|_{L^2(\Omega_h)} + h|\zeta-\Pi_h\zeta|_{H^1(\Omega_h)} + h\|\zeta-\Pi_h\zeta\|_{L^\infty(\Omega_h)}&\\ + h^2\|D_h^2(\zeta-\Pi_h\zeta)\|_{L^2(\Omega_h)} &\le Ch^4\|\zeta\|_{H^4(\mathbb{R}^2)},\end{aligned}$$

(3.3)
$$|\zeta-\Pi_h\zeta|_{W^{2,\infty}(T)} \le Ch\|\zeta\|_{H^4(\mathbb{R}^2)} \quad \forall\, T\in\mathcal{T}_h,$$

(3.4)
$$\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial(\Pi_h\zeta)/\partial n]\!]\|_{L^2(e)}^2 = \sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial(\zeta-\Pi_h\zeta)/\partial n]\!]\|_{L^2(e)}^2 \le Ch^4\|\zeta\|_{H^4(\mathbb{R}^2)}^2,$$

(3.5)
$$\|D_h^2(\Pi_h\zeta)\|_{L^2(\Omega_h)}^2 + \max_{T\in\mathcal{T}_h}|\Pi_h\zeta|_{W^{1,\infty}(T)}^2 + \max_{T\in\mathcal{T}_h}|\Pi_h\zeta|_{W^{2,\infty}(T)}^2 \le C\|\zeta\|_{H^4(\mathbb{R}^2)}^2,$$

(3.6)
$$\sum_{T\in\mathcal{T}_h}|D^2\big((\Pi_h\zeta)\circ\Phi_T\big)|_{H^2(\hat{T})}^2 \le Ch^6\|\zeta\|_{H^4(\mathbb{R}^2)}^2,$$

where the positive constant C is independent of h.

Remark 3.3. As noted at the beginning of Section 2.3, there exists a bounded linear extension map $\mathbb{E} : H^4(\Omega) \longrightarrow H^4(\mathbb{R}^2)$ such that $\mathbb{E}v = v$ on Ω, and we will denote $\mathbb{E}u$ (resp., $\mathbb{E}\phi$) again by u (resp., ϕ). Therefore $\Pi_h u$ and $\Pi_h\phi$ are well-defined. We assume that u is strictly convex in a neighborhood $\Omega_\dagger$ of $\bar{\Omega}$. The function ψ can also be extended to $\mathbb{R}^2$ by the relation $\det D^2u = \psi$.

Remark 3.4. It follows from (1.5), (1.6) and (3.3) that (for $h \ll 1$)

$$(\alpha_\sharp/2)|\xi|^2 \le \xi^t D_h^2(\Pi_h u)(x)\xi \le (2\beta_\sharp)|\xi|^2 \qquad \forall\, x\in\Omega_h, \xi\in\mathbb{R}^2.$$

Remark 3.5. A direct calculation using (1.1a), (3.2), (3.3) and (3.5) yields the estimate

$$\|\det D_h^2(\Pi_h u) - \psi\|_{L^2(\Omega_h)} = \|\det D_h^2(\Pi_h u) - \det D_h^2 u\|_{L^2(\Omega_h)} \le Ch^2.$$

3.2. **A nonlinear least-squares problem with box constraints.** The discrete problem is to find

$$u_h \in \operatorname{argmin}_{v\in L_h} J_h(v), \tag{3.7}$$

where

$$L_h = \{v \in V_h : v = \Pi_h\phi \text{ on } \partial\Omega_h \text{ and } \Delta(v\circ\Phi_T) \geq 0 \text{ at the vertices of } \hat{T}_\dagger \text{ for every } T \in \mathcal{T}_h\}, \tag{3.8}$$

and

$$J_h(v) = \frac{h^4}{2}\|D_h^2 v\|^2_{L^2(\Omega_h)} + \frac{h^{-2}}{2}\sum_{T\in\mathcal{T}_h}|\Delta(v\circ\Phi_T)|^2_{H^2(\hat{T})} + \frac{1}{2}\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial v/\partial n]\!]\|^2_{L^2(e)} + \frac{1}{2}\|\det D^2 v - \psi\|^2_{L^2(\Omega_h)}. \tag{3.9}$$

Remark 3.6. According to the definition of V_h and Π_h, the function $(\Pi_h\phi)\circ\Phi_T$ is a cubic polynomial on the edge $\hat{p}_2\hat{p}_3$ of $\hat{T}$ that connects $\hat{p}_2$ and $\hat{p}_3$ (cf. Figure 2.3) if T has two vertices on $\partial\Omega$. Moreover, we have

$$\begin{aligned}
(\Pi_h\phi)(\Phi_T(\hat{p}_i)) &= \phi(\Phi_T(\hat{p}_i)) \qquad \text{for} \quad i = 2,3,\\
D\big((\Pi_h\phi)\circ\Phi_T\big)(\hat{p}_2)(\hat{\mathbf{p}}_3-\hat{\mathbf{p}}_2) &= D(\phi\circ\Phi_T)(\hat{p}_2)(\hat{\mathbf{p}}_3-\hat{\mathbf{p}}_2) = |\mathbf{p_3}-\mathbf{p_2}|D\phi(p_2)\boldsymbol{e}_{23},\\
D\big((\Pi_h\phi)\circ\Phi_T\big)(\hat{p}_3)(\hat{\mathbf{p}}_2-\hat{\mathbf{p}}_3) &= D(\phi\circ\Phi_T)(\hat{p}_3)(\hat{\mathbf{p}}_2-\hat{\mathbf{p}}_3) = |\mathbf{p_2}-\mathbf{p_3}|D\phi(p_3)\boldsymbol{e}_{32},
\end{aligned}$$

where $\boldsymbol{e}_{23}$ (resp., $\boldsymbol{e}_{32}$) is the unit tangent of $\partial\Omega$ at p_2 (resp., p_3) that points towards p_3 (resp., p_2) (cf. (2.1c)–(2.1d) and Figure 2.3). Therefore $(\Pi_h\phi)\circ\Phi_T$ is the one-dimensional cubic Hermite interpolant of $\phi\circ\Phi_T$ restricted to the edge $\hat{p}_2\hat{p}_3$ and it is defined solely by the available information of ϕ $(= u)$ on $\partial\Omega$. In particular, we have

$$v = \Pi_h\phi = \Pi_h u \quad \text{on} \quad \partial\Omega_h \qquad \forall\, v \in L_h. \tag{3.10}$$

Note also that $v = \Pi_h\phi$ is a box equality constraint in the dofs of V_h.

Remark 3.7. According to the definition of V_h, the inequality constraints in the definition of L_h are also box constraints in the dofs of V_h.

Remark 3.8. The closed convex subset L_h of V_h is nonempty because

$$\Pi_h u \in L_h \tag{3.11}$$

by Lemma 2.4 and Remark 3.6.

Remark 3.9. The first two terms in (3.9) are regularization terms that are crucial for the solvability of the discrete problem and for enforcing the elementwise convexity of the discrete solutions. The third term is a penalty term (cf. [31, 50]) that compensates for the fact that $V_h \not\subset H^2(\Omega_h)$. The last term is the least-squares term for (1.1a).

We will analyze the least-squares problem defined by (3.7)–(3.9) in terms of the mesh-dependent semi-norm $\|\cdot\|_h$ defined by

$$\|v\|_h^2 = \|D_h^2 v\|_{L^2(\Omega_h)} + \sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial v/\partial n]\!]\|^2_{L^2(e)}. \tag{3.12}$$

Remark 3.10. Note that $\|\cdot\|_h$ defines a norm on $V_h \cap H_0^1(\Omega_h)$. Since $L_h \subset \Pi_h\phi + [V_h \cap H_0^1(\Omega_h)]$, the cost function $J_h(v) \to \infty$ if $v \in L_h$ and $\|v\|_h$ goes to ∞. Hence $J_h : L_h \to [0,\infty)$ has a global minimizer.

Remark 3.11. It follows from (3.2), (3.4) and (3.12) that

$$\|u - \Pi_h u\|_h \le Ch^2,$$

where the positive constant C is independent of h.

3.3. **Some *a priori* bounds.** The bounds derived in this section are crucial for the error analysis in Section 4.

Let u_h satisfy (3.7). It follows from (3.11) that

$$\begin{aligned}
& h^4\|D_h^2 u_h\|_{L^2(\Omega_h)}^2 + h^{-2}\sum_{T\in\mathcal{T}_h}|\Delta(u_h\circ\Phi_T)|_{H^2(\hat{T})}^2 + \|\det D^2 u_h - \psi\|_{L^2(\Omega_h)}^2 \\
&\qquad + \sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial u_h/\partial n]\!]\|_{L^2(e)}^2 \\
&= 2J_h(u_h) \qquad (3.13)\\
&\le 2J_h(\Pi_h u) \\
&= h^4\|D_h^2(\Pi_h u)\|_{L^2(\Omega_h)}^2 + h^{-2}\sum_{T\in\mathcal{T}_h}|\Delta\big((\Pi_h u)\circ\Phi_T\big)|_{H^2(\hat{T})}^2 \\
&\qquad + \|\det D^2(\Pi_h u) - \psi\|_{L^2(\Omega_h)}^2 + \sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial(\Pi_h u)/\partial n]\!]\|_{L^2(e)}^2 \\
&\le Ch^4,
\end{aligned}$$

where we have applied the estimates (3.4)–(3.6) and Remark 3.5.

Consequently, we have

$$\|D_h^2 u_h\|_{L^2(\Omega_h)} \le C, \tag{3.14}$$

$$\Big(\sum_{T\in\mathcal{T}_h}|\Delta(u_h\circ\Phi_T)|_{H^2(\hat{T})}^2\Big)^{\frac{1}{2}} \le Ch^3, \tag{3.15}$$

$$\Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\Big\|[\![\partial u_h/\partial n]\!]\Big\|_{L^2(e)}^2\Big)^{\frac{1}{2}} \le Ch^2, \tag{3.16}$$

$$\|\det D_h^2 u_h - \psi\|_{L^2(\Omega_h)} \le Ch^2. \tag{3.17}$$

It follows from (3.14), (3.16), the discrete Sobolev inequality in [24] and a Poincaré-Friedrichs inequality for piecewise H^1 functions in [23] that

$$\|\nabla u_h\|_{L^\infty(\Omega_h)} \le C\big(1 + |\ln h|\big). \tag{3.18}$$

Detailed arguments are provided in Appendix B.

The following lemma, which follows easily from (3.14), (3.17) and standard inverse estimates in the case of simplicial triangulations for polygonal domains, requires a careful treatment in the case of isoparametric meshes for smooth domains. Its proof is given in Appendix C.

Lemma 3.12. *There exists a positive constant C independent of h such that*

$$\|D_h^2 u_h\|_{L^\infty(\Omega_h)} \le Ch^{-1}, \tag{3.19}$$

$$\|\det D_h^2 u_h - \psi\|_{L^\infty(\Omega_h)} \le Ch. \tag{3.20}$$

4. Convergence analysis

We will develop *a priori* and *a posteriori* error estimates by exploiting the elementwise convexity of the solutions of (3.7) and the stability of C^0 interior penalty methods for elliptic problems in nondivergence form.

4.1. **Elementwise convexity of the discrete solutions.**

Let $u_h \in L_h$ be a solution of (3.7). For any $T \in \mathcal{T}_h$, we have the following analog of (2.17):

$$\begin{aligned}(D^2\hat{u}_h)(\hat{x}) = D\Phi_T(\hat{x})^t(D^2u_h)(\Phi_T(\hat{x}))D\Phi_T(\hat{x}) &+ \frac{\partial u_h}{\partial x_1}(\Phi_T(\hat{x}))D^2\phi_{T,1}(\hat{x}) \\ &+ \frac{\partial u_h}{\partial x_2}(\Phi_T(\hat{x}))D^2\phi_{T,2}(\hat{x}) \qquad \forall\, \hat{x} \in \hat{T}_\dagger,\end{aligned} \tag{4.1}$$

where $\hat{u}_h = u_h \circ \Phi_T$, and

$$\Delta\hat{u}_h \ge 0 \quad \text{at the vertices of } \hat{T}_\dagger \tag{4.2}$$

by (3.8).

We will treat a polynomial q defined on $\hat{T}$ as the restriction of a polynomial defined on $\mathbb{R}^2$ (also denoted by q), and denote by $\tilde{I}q$ the restriction of $I_{\hat{T}_\dagger} q$ to $\hat{T}$, where $I_{\hat{T}_\dagger}$ is the P_1 nodal interpolation operator associated with the vertices of the larger triangle $\hat{T}_\dagger$ (cf. Figure 2.4). It follows from (4.2) that

$$\tilde{I}(\Delta\hat{u}_h) \ge 0 \quad \text{on } \hat{T}, \tag{4.3}$$

and we also have

$$\|\Delta\hat{u}_h - \tilde{I}(\Delta\hat{u}_h)\|_{L^\infty(\hat{T})} \lesssim |\Delta\hat{u}_h|_{H^2(\hat{T})} \lesssim h^3 \tag{4.4}$$

by the Bramble-Hilbert lemma (since $P_1(\hat{T})$ is invariant under $\tilde{I}$) and (3.15).

Combining (2.16), (3.18), (4.3) and (4.4), we find

$$\begin{aligned}\Delta\hat{u}_h(\hat{x}) &- \frac{\partial u_h}{\partial x_1}(\Phi_T(\hat{x}))\Delta\phi_{T,1}(\hat{x}) - \frac{\partial u_h}{\partial x_2}(\Phi_T(\hat{x}))\Delta\phi_{T,2}(\hat{x}) \\ &\ge \Delta\hat{u}_h(\hat{x}) - \tilde{I}(\Delta\hat{u}_h)(\hat{x}) - \frac{\partial u_h}{\partial x_1}(\Phi_T(\hat{x}))\Delta\phi_{T,1}(\hat{x}) - \frac{\partial u_h}{\partial x_2}(\Phi_T(\hat{x}))\Delta\phi_{T,2}(\hat{x}) \\ &\ge -\|\Delta\hat{u}_h - \tilde{I}(\Delta\hat{u}_h)\|_{L^\infty(\hat{T})} - \frac{\partial u_h}{\partial x_1}(\Phi_T(\hat{x}))\Delta\phi_{T,1}(\hat{x}) - \frac{\partial u_h}{\partial x_2}(\Phi_T(\hat{x}))\Delta\phi_{T,2}(\hat{x}) \\ &\gtrsim -(1+|\ln h|)h^3 \qquad \forall\, \hat{x} \in \hat{T}.\end{aligned} \tag{4.5}$$

It then follows from (2.13), (4.1) and (4.5) that

$$\operatorname{tr}[D^2u_h(x)] \gtrsim -(1+|\ln h|)h \qquad \forall\, x \in T. \tag{4.6}$$

On the other hand the estimate (3.20) implies

$$\det D_h^2 u_h \ge \frac{1}{2}\min_{x\in\bar{\Omega}_h} \psi(x) > 0 \qquad \text{on } T \tag{4.7}$$

for $h \ll 1$.

We conclude from (4.6) and (4.7) that

$$D_h^2 u_h \text{ is positive definite on all } T \in \mathcal{T}_h. \tag{4.8}$$

4.2. ***A priori* error estimates.** It follows from the fundamental theorem of calculus (cf. [56, Lemma A.1]) that the relation

$$\det D_h^2(\Pi_h u) - \det D_h^2 u_h = \Big[\int_0^1 \operatorname{Cof} D_h^2\big(t(\Pi_h u) + (1-t)u_h\big)dt\Big] : D_h^2(\Pi_h u - u_h) \tag{4.9}$$

holds in the interior of all the triangles in $\mathcal{T}_h$, where the colon denotes the Frobenius inner product between matrices.

Since a symmetric 2×2 matrix and its cofactor matrix have identical eigenvalues, Remark 3.4 and (4.8) imply that the matrix-valued function

$$A_h = \int_0^1 \operatorname{Cof} D_h^2\big(t(\Pi_h u) + (1-t)u_h\big)dt = \frac{1}{2}\big(\operatorname{Cof} D_h^2(\Pi_h u) + \operatorname{Cof} D_h^2 u_h\big) \tag{4.10}$$

satisfies

$$\alpha|\xi|^2 \le \xi^t A_h(x)\xi \le \beta|\xi|^2 \qquad \forall\, \xi \in \mathbb{R}^2 \text{ a.e. on } \Omega_h, \tag{4.11}$$

where $\alpha = \alpha_\sharp/4$ and

$$\beta = \frac{1}{2}\big(2\beta_\sharp + \|D_h^2 u_h\|_{L^\infty(\Omega_h)}\big) \tag{4.12}$$

are positive constants. In particular A_h belongs to $[L^\infty(\Omega)]^{2\times 2}$.

The relation (4.9) indicates that we are in the realm of C^0 interior penalty methods for elliptic boundary value problems in nondivergence form. The following stability result, which is the consequence of (4.11) and a discrete Miranda-Talenti estimate, is related to the ones in [83, 91]. Its derivation, which requires some subtle analysis for isoparametric finite elements, is provided in Appendix D.

Lemma 4.1. *There exists a positive constant $C_\dagger$ independent of h such that*

(4.13)

$$\begin{aligned}\|D_h^2 v\|_{L^2(\Omega_h)} \le \Big(\frac{\alpha^{-1}}{1-\delta}\Big)\|A_h : D_h^2 v\|_{L_2(\Omega_h)} \\ + \Big(\frac{C_\dagger}{1-\delta}\Big)\Big[\Big(\sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}} + h^3\|\nabla v\|_{L^\infty(\Omega_h)}\Big]\end{aligned}$$

for all $v \in V_h \cap H_0^1(\Omega_h)$, where

$$\delta = \frac{\beta-\alpha}{(\alpha^2+\beta^2)^{\frac{1}{2}}} (< 1). \tag{4.14}$$

We can now combine the interpolation error estimates in Section 3.1, the *a priori* bounds in Section 3.3, the relation (4.9) and Lemma 4.1 to obtain *a priori* error estimates for any solution of (3.7).

We begin with a preliminary error estimate.

Lemma 4.2. *Let u_h be a solution of* (3.7)*. There exists a positive constant C independent of h such that*

$$\|D_h^2(\Pi_h u - u_h)\|_{L^2(\Omega_h)} \le Ch. \tag{4.15}$$

Proof. First we note that $v = \Pi_h u - u_h \in V_h \cap H_0^1(\Omega_h)$ by (3.10). Moreover, the definition (4.12) and the estimate (3.19) implies $\beta = O(1/h)$ and hence, in view of the definition of δ in (4.14),

$$(1-\delta)^{-1} \leq C_\diamond h^{-1} \tag{4.16}$$

for some positive constant $C_\diamond$ independent of h.

According to (3.4), (3.5), Remark 3.5, (3.16)–(3.18), (4.9), Lemma 4.1 and (4.16), we have

$$\begin{aligned}
\|D_h^2(\Pi_h u - u_h)\|_{L^2(\Omega_h)} &\leq \alpha^{-1} C_\diamond h^{-1} \|A_h : D_h^2(\Pi_h u - u_h)\|_{L^2(\Omega_h)} \\
&\quad + C_\dagger C_\diamond h^{-1} \Big(\sum_{e \in \mathcal{E}_h^i} |e|^{-1} \|[\![\partial(\Pi_h u - u_h)/\partial n]\!]\|_{L^2(e)}^2 \Big)^{\frac{1}{2}} \\
&\quad + C_\dagger C_\diamond h^2 \|\nabla(\Pi_h u - u_h)\|_{L^\infty(\Omega_h)} \\
&\leq \alpha^{-1} C_\diamond h^{-1} \|\det D_h^2(\Pi_h u) - \det D_h^2 u_h\|_{L^2(\Omega_h)} \\
&\quad + 2C_\dagger C_\diamond h^{-1} \Big(\sum_{e \in \mathcal{E}_h^i} |e|^{-1} \|[\![\partial(\Pi_h u)/\partial n]\!]\|_{L^2(e)}^2 \\
&\qquad + \sum_{e \in \mathcal{E}_h^i} |e|^{-1} \|[\![\partial u_h/\partial n]\!]\|_{L^2(e)}^2 \Big)^{\frac{1}{2}} \\
&\quad + C_\dagger C_\diamond h^2 \big(\|\nabla(\Pi_h u)\|_{L^\infty(\Omega_h)} + \|\nabla u_h\|_{L^\infty(\Omega_h)} \big) \\
&\leq Ch.
\end{aligned} \tag{4.17}$$

□

It follows from (3.5), (3.18), (4.15) and an inverse estimate (cf. Lemma C.1 in Appendix C) that

$$\begin{aligned}
\|D_h^2 u_h\|_{L^\infty(\Omega_h)} &\leq \|D_h^2(\Pi_h u - u_h)\|_{L^\infty(\Omega_h)} + \|D_h^2(\Pi_h u)\|_{L^\infty(\Omega_h)} \\
&\lesssim h^{-1} \|D_h^2(\Pi_h u - u_h)\|_{L^2(\Omega_h)} + h^4 \|\nabla(\Pi_h u - u_h)\|_{L^\infty(\Omega_h)} \\
&\qquad + \|D_h^2(\Pi_h u)\|_{L^\infty(\Omega_h)} \\
&\lesssim 1.
\end{aligned} \tag{4.18}$$

In view of the estimate (4.18) that improves (3.19), the estimate for β defined by (4.12) becomes $\beta \lesssim 1$ and hence δ defined by (4.14) satisfies

$$(1-\delta)^{-1} \leq C_\spadesuit \tag{4.19}$$

for some positive constant $C_\spadesuit$ independent of h.

Theorem 4.3. *Let u_h be a solution of* (3.7). *There exists a positive constant C independent of h such that*

$$\|u - u_h\|_h \leq Ch^2. \tag{4.20}$$

Proof. We can repeat the arguments in the proof of Lemma 4.2 but with (4.16) replaced by (4.19) to obtain the estimate

$$\|D_h^2(\Pi_h u - u_h)\|_{L^2(\Omega_h)} \lesssim h^2, \tag{4.21}$$

which together with (3.4), (3.12), Remark 3.11 and (3.16) implies (4.20). □

We can also derive error estimates for lower order norms from (4.21).

Corollary 4.4. *Let u_h be a solution of* (3.7)*. There exists a positive constant C independent of h such that*

$$\|u-u_h\|_{L^2(\Omega_h)} + |u-u_h|_{H^1(\Omega_h)} + \|u-u_h\|_{L^\infty(\Omega_h)} \le Ch^2. \tag{4.22}$$

Proof. First we measure the errors over the convex polygonal domain $\tilde{\Omega}_h \subset \Omega_h$ (cf. Figure 2.1).

It follows from the Poincaré-Friedrichs and Sobolev inequalities for piecewise H^2 functions in [29, 33] that

$$\begin{aligned}\|\Pi_h u - u_h\|_{L^2(\tilde{\Omega}_h)} &+ |\Pi_h u - u_h|_{H^1(\tilde{\Omega}_h)} + \|\Pi_h u - u_h\|_{L^\infty(\tilde{\Omega}_h)} \\ &\lesssim \Big(\sum_{\tilde{T}\in\tilde{\mathcal{T}}_h} |D_h^2(\Pi_h u - u_h)|^2_{L^2(\tilde{T})} + \sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial(\Pi_h u - u_h)/\partial n]\!]\|^2_{L^2(e)}\Big)^{\frac{1}{2}} \\ &\quad + \|\Pi_h u - u_h\|_{L^2(\partial\tilde{\Omega}_h)}.\end{aligned} \tag{4.23}$$

Observe that (3.4), (3.10), (3.16), (4.21) and Lemma B.1 imply

$$\|\nabla(\Pi_h u - u_h)\|_{L^\infty(\Omega_h)} \lesssim (1+|\ln h|)h^2 \tag{4.24}$$

and hence

$$\|\Pi_h u - u_h\|_{L^\infty(\partial\tilde{\Omega}_h)} \lesssim (1+|\ln h|)h^4 \tag{4.25}$$

because $\Pi_h u - u_h = 0$ on $\partial\Omega_h$ and the gap between $\partial\tilde{\Omega}_h$ and $\partial\Omega_h$ is $O(h^2)$.

Putting (3.4), (3.16), (4.21), (4.23) and (4.25) together, we have

$$\|\Pi_h u - u_h\|_{L^2(\tilde{\Omega}_h)} + |\Pi_h u - u_h|_{H^1(\tilde{\Omega}_h)} + \|\Pi_h u - u_h\|_{L^\infty(\tilde{\Omega}_h)} \lesssim h^2,$$

which, in view of (4.24), implies

$$\|\Pi_h u - u_h\|_{L^2(\Omega_h)} + |\Pi_h u - u_h|_{H^1(\Omega_h)} + \|\Pi_h u - u_h\|_{L^\infty(\Omega_h)} \lesssim h^2. \tag{4.26}$$

The estimate (4.22) follows from (3.2) and (4.26). □

Remark 4.5. Numerical results in Section 5 indicate that the error in $\|\cdot\|_{L^2(\Omega_h)}$, $|\cdot|_{H^1(\Omega_h)}$ and $\|\cdot\|_{L^\infty(\Omega_h)}$ are better than $O(h^2)$.

4.3. **An *a posteriori* error estimate.** In practice the numerical solution $\bar{u}_h$ of (3.7) obtained by an optimization algorithm is only an approximate stationary point of the cost function J_h. It is therefore important to be able to monitor the convergence of $\bar{u}_h$ by an *a posteriori* error indicator.

Under the condition that

$$\tilde{\alpha}_\flat|\xi|^2 \le \xi^t D_h^2\bar{u}_h\xi \le \tilde{\beta}_\flat|\xi|^2 \qquad \text{on all } T\in\mathcal{T}_h \text{ and for all } \xi\in\mathbb{R}^2, \tag{4.27}$$

where the positive constants $\tilde{\alpha}$ and $\tilde{\beta}$ are independent of h, we have an analog of (4.9) that follows from the fundamental theorem of calculus.

$$\det D_h^2(\Pi_h u) - \det D_h^2\bar{u}_h = \tilde{A}_h : D_h^2(\Pi_h u - \bar{u}_h) \quad \text{a.e. in } \Omega_h, \tag{4.28}$$

where

$$\tilde{A}_h = \int_0^1 \operatorname{Cof} D_h^2\big(t(\Pi_h u) + (1-t)\bar{u}_h\big)dt$$

satisfies

$$\tilde{\alpha}|\xi|^2 \leq \xi^t \tilde{A}_h(x)\xi \leq \tilde{\beta}|\xi|^2 \qquad \forall\, \xi \in \mathbb{R}^2 \text{ a.e. on } \Omega_h, \tag{4.29}$$

and the positive constants $\tilde{\alpha}$ and $\tilde{\beta}$ are independent of h.

Remark 4.6. The condition (4.27) can be verified computationally.

We can use (4.28) and (4.29) to derive an *a posteriori* error estimate for $\|u - \bar{u}_h\|_h$.

Theorem 4.7. *Under condition* (4.27) *we have*

$$\|u - \bar{u}_h\|_h \leq C(\eta_h(\bar{u}_h) + h^2), \tag{4.30}$$

where

$$\eta_h(\bar{u}_h) = \|\det D_h^2 \bar{u}_h - \psi\|_{L^2(\Omega_h)} + \Big(\sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial\bar{u}_h/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}} \tag{4.31}$$

and the constant C is independent of h.

Proof. It follows from (4.29) that we have an analog of (4.13):

$$\|D_h^2 v\|_{L^2(\Omega)} \leq C\Big[\|\tilde{A}_h : D_h^2 v\|_{L_2(\Omega_h)} + \Big(\sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}} + h^3\|\nabla v\|_{L^\infty(\Omega_h)}\Big] \tag{4.32}$$

for all $v \in V_h \cap H_0^1(\Omega_h)$, where the positive constant C is independent of h.

We can then use (1.1a), (3.4), Remark 3.5, (4.28), (4.32) and Lemma B.1 to obtain

$$\begin{aligned}
&\|D_h^2(\Pi_h u - \bar{u}_h)\|_{L^2(\Omega_h)}\\
&\quad\lesssim \|\tilde{A}_h : D_h^2(\Pi_h u - \bar{u}_h)\|_{L^2(\Omega_h)} + h^3\|\nabla(\Pi_h u - \bar{u}_h)\|_{L^\infty(\Omega_h)}\\
&\qquad + \Big(\sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial(\Pi_h u - \bar{u}_h)/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}}\\
&\quad\lesssim \|\det D_h^2(\Pi_h u) - \det D_h^2\bar{u}_h\|_{L^2(\Omega_h)} + h^3(1 + |\ln h|)\|D_h^2(\Pi_h u - \bar{u}_h)\|_{L^2(\Omega_h)}\\
&\qquad + \Big(\sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial\bar{u}_h/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}} + h^2\\
&\quad\lesssim \|\det D_h^2\bar{u}_h - \psi\|_{L^2(\Omega_h)} + h^3(1 + |\ln h|)\|D_h^2(\Pi_h u - \bar{u}_h)\|_{L^2(\Omega_h)}\\
&\qquad + \Big(\sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial\bar{u}_h/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}} + h^2,
\end{aligned}$$

and hence

$$\|D_h^2(\Pi_h u - \bar{u}_h)\|_{L^2(\Omega_h)} \lesssim \|\det D_h^2\bar{u}_h - \psi\|_{L^2(\Omega_h)} + \Big(\sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial\bar{u}_h/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}} + h^2. \tag{4.33}$$

Then we have, by (3.4), (3.12), Remark 3.11, (4.31) and (4.33),

$$\begin{aligned}\|u-\bar{u}_h\|_h &\leq \|\Pi_h u-\bar{u}_h\|_h+\|u-\Pi_h u\|_h \\ &\lesssim \|D_h^2(\Pi_h u-\bar{u}_h)\|_{L^2(\Omega_h)}+\Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial(\Pi_h u-\bar{u}_h)/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}}+h^2 \qquad (4.34)\\ &\leq C(\eta_h(\bar{u}_h)+h^2).\end{aligned}$$

□

Therefore we can use $\eta_h(\bar{u}_h)$ to monitor the convergence of $\bar{u}_h$.

5. Numerical experiments

We have tested our method on two examples. For the first example, the exact solution is known, while the exact solution is unknown for the second example. For each example, we solve the discrete problem (3.7)–(3.9) by an active set algorithm (cf. [32, Appendix B] and [65–67]) that produces an approximate stationary point of the problem. The elementwise convexity of the approximate solutions is checked numerically by the algorithm in [32, Appendix C].

We take Ω to be a disc or an elliptical domain in our tests. The boundary of the disc is given by

$$(x_1-1/2)^2+(x_2-1/2)^2=1$$

and the boundary of the elliptical domain is given by

$$x_1^2+4x_2^2=1.$$

In Example 5.1 where the exact solution is known, the relative errors of the approximate solution $\tilde{u}_h$ in various norms are defined by

$$e_{2,h}^r=\frac{|u-\tilde{u}_h|_{H^2(\Omega_h)}}{|u|_{H^2(\Omega)}},\quad e_{1,h}^r=\frac{|u-\tilde{u}_h|_{H^1(\Omega_h)}}{|u|_{H^1(\Omega)}},\quad e_{0,h}^r=\frac{\|u-\tilde{u}_h\|_{L^2(\Omega)}}{\|u\|_{L^2(\Omega)}}$$

and

$$e_{\infty,h}^r=\frac{\max_{p\in\mathcal{V}_h}|u(p)-\tilde{u}_h(p)|}{\|u\|_{L^\infty(\Omega)}},$$

where $\mathcal{V}_h$ be the set of all the vertices of the decomposition $\mathcal{T}_h$.

All the numerical experiments were carried out on a MacBook Pro laptop computer with a 2.8GHz Quad-Core Intel Core i7 processor and with 16GB 2133 MHz LPDDR3 memory. We use MATLAB (R2021a v.9.10.0) in our computations.

Example 5.1. Let $\psi=(1+|x|^2)e^{|x|^2}$ and $\phi=e^{\frac{1}{2}|x|^2}$. The exact solution of this example is $u=e^{\frac{1}{2}|x|^2}$. This example first appeared in [45] where it was posed on the unit square $(0,1)^2$. Numerical results for discs and elliptical domains can also be found in [69, 79].

The errors of the approximate solution $\tilde{u}_h$ on uniform meshes for the disc and the elliptical domain are presented in Table 5.1 and Table 5.2 respectively. The order of convergence for $e_{2,h}^r$ is about 2 for both the disc and the elliptical domain, which agrees with the estimate in Theorem 4.3. The orders of convergence for $e_{1,h}^r$, $e_{0,h}^r$ and $e_{\infty,h}^r$ are higher than the estimates in Corollary 4.4.

TABLE 5.1. Relative errors versus mesh size h and orders of convergence for Example 5.1 on the disc

h	$e^r_{2,h}$	order	$e^r_{1,h}$	order	$e^r_{0,h}$	order	$e^r_{\infty,h}$	order
1.4142e0	3.1164e-1	–	2.0470e-1	–	6.5945e-2	–	9.9749e-2	–
7.6537e-1	1.5326e-1	1.16	7.7198e-2	1.59	2.9650e-2	1.30	2.2881e-2	2.40
4.2033e-1	5.6636e-2	1.66	1.8173e-2	2.41	6.7873e-3	2.46	6.3419e-3	2.14
2.2193e-1	1.6455e-2	1.94	3.2604e-3	2.69	8.0467e-4	3.34	7.7747e-4	3.29
1.1373e-1	3.7796e-3	2.20	4.4339e-4	2.98	8.2385e-5	3.41	1.2383e-4	2.75
5.7536e-2	7.8618e-4	2.30	5.0280e-5	3.19	7.1523e-6	3.59	1.2404e-5	3.38

TABLE 5.2. Relative errors versus mesh size h and orders of convergence for Example 5.1 on the elliptical domain

h	$e^r_{2,h}$	order	$e^r_{1,h}$	order	$e^r_{0,h}$	order	$e^r_{\infty,h}$	order
1.1180e0	4.8288e-1	–	2.4353e-1	–	5.4249e-2	–	1.5862e-2	–
7.2211e-1	1.5912e-1	2.54	5.1224e-2	3.57	5.7344e-3	5.14	8.9989e-3	1.30
3.9243e-1	3.9422e-2	2.29	9.3326e-3	2.79	1.1219e-3	2.68	1.7901e-3	2.65
2.0213e-1	8.6457e-3	2.29	1.1319e-3	3.18	9.6037e-5	3.70	1.4295e-4	3.81
1.0217e-1	1.8046e-3	2.30	1.2064e-4	3.28	7.2136e-6	3.79	1.2103e-5	3.62
5.1316e-2	3.5186e-4	2.37	1.1698e-5	3.39	5.2312e-7	3.81	1.0038e-6	3.62

The residual $\eta_h(\tilde{u}_h)$ and the cost $J_h(\tilde{u}_h)$ are presented in Table 5.3 for the disc and Table 5.4 for the elliptical domain. The behavior of J_h agrees with the estimate in (3.13). The reliability estimate (4.34) can be observed by comparing $e^r_{2,h}$ in Table 5.1 (resp., Table 5.2) and $\eta_h(\tilde{u}_h)$ in Table 5.3 (resp., Table 5.4).

TABLE 5.3. Residual, Cost and CPU time for Example 5.1 on the disc

h	1.4142e0	7.6537e-1	4.2033e-1	2.2193e-1	1.1373e-1	5.7536e-2
$\eta_h(\tilde{u}_h)$	1.2261e1	4.0633e0	1.0669e0	2.7206e-1	6.9185e-2	1.7192e-2
Order	–	1.80	2.23	2.14	2.05	2.04
$J_h(\tilde{u}_h)$	1.7711e2	2.2890e1	1.9775e0	1.4741e-1	1.0113e-2	6.6290e-4
Order	–	3.33	4.09	4.07	4.01	4.00
CPU time (s)	1.4269e1	2.4154e1	1.1641e1	3.7932e1	1.3047e2	7.2279e2

TABLE 5.4. Residual, Cost and CPU time for Example 5.1 on the elliptical domain

h	1.1180e0	7.2211e-1	3.9243e-1	2.0213e-1	1.0217e-1	5.1316e-2
$\eta_h(\tilde{u}_h)$	2.2948e0	6.0810e-1	1.3225e-1	2.8456e-2	6.1127e-3	9.3132e-4
Order	–	3.04	2.50	2.32	2.25	2.73
$J_h(\tilde{u}_h)$	4.0546e0	9.0543e-1	8.1358e-2	5.6841e-3	3.6705e-4	2.3870e-5
Order	–	3.43	3.95	4.01	4.02	3.97
CPU time (s)	5.3126e1	3.5041e0	6.5679e0	3.1908e1	6.3131e1	5.3018e2

It is observed from the CPU times in Table 5.3 (resp., Table 5.4) that a good approximate solution, i.e., $e^r_{\infty,h} \leq 10^{-2}$, can be obtained in under 12 (resp., 7) seconds. The profiles of the computed solutions on the final meshes are displayed in Figure 5.1

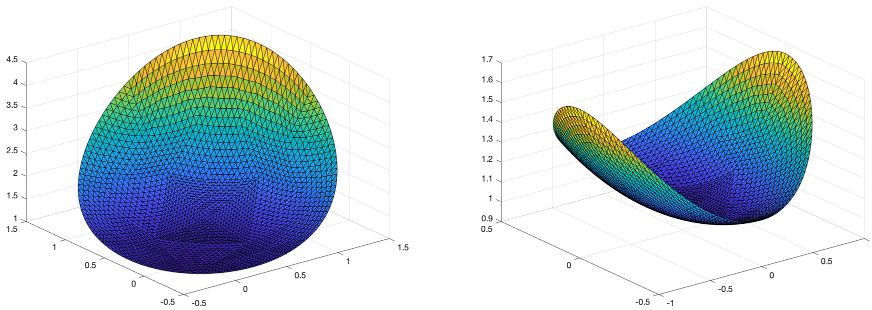

FIGURE 5.1. The profile of the computed solution on the final mesh for Example 5.1 on the disc (left) and on the elliptical domain (right)

Example 5.2. Let $\psi = 1$ and $\phi = e^{\frac{1}{2}|x|^2}$. The exact solution of this example is unknown. This example is obtained by setting $\psi = 1$ in Example 5.1. Although the exact solution is unknown, the convergence of the approximate solutions can be monitored by η_h (cf. Section 4.3).

We solve this problem on uniform meshes for the elliptical domain. The residual $\eta_h(\tilde{u}_h)$, the cost $J_h(\tilde{u}_h)$ and the CPU times are presented in Table 5.5.

We have verified that all the approximate solutions are elementwise strictly convex. According to our theory, the convergence of $\eta_h(\tilde{u}_h)$ to zero observed in Table 5.5 indicates the convergence of $\tilde{u}_h$ to the exact solution u of (1.1). The profile of the approximate solution of this example on the final mesh is shown in Figure 5.2.

TABLE 5.5. Residual, Cost and CPU time for Example 5.2 on the elliptical domain

h	1.1180e0	7.2211e-1	3.9243e-1	2.0213e-1	1.0217e-1	5.1316e-2
$\eta_h(\tilde{u}_h)$	1.4865e0	4.3033e-1	6.3078e-2	1.0410e-2	2.0413e-3	2.9115e-4
Order	–	2.84	3.15	2.72	2.39	2.83
$J_h(\tilde{u}_h)$	1.6260e0	4.5908e-1	4.3730e-2	3.0832e-3	2.0078e-4	1.3499e-5
Order	–	2.89	3.86	4.00	4.00	3.92
CPU time (s)	2.0514e1	2.2355e0	5.9253e0	1.5054e1	7.6981e1	6.2503e2

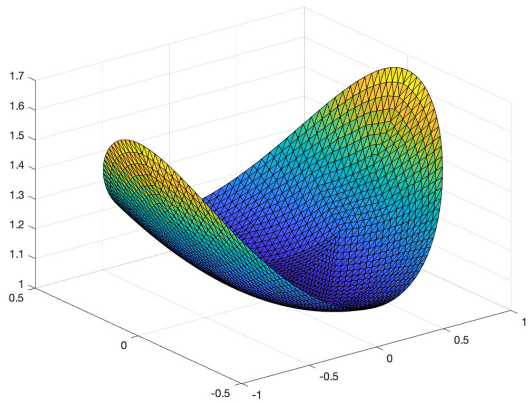

FIGURE 5.2. The profile of the computed solution on the final mesh for Example 5.2 on the elliptical domain

6. CONCLUDING REMARKS

We have constructed a nonlinear least-squares finite element method that can capture smooth convex solutions of the Dirichlet boundary value problem of the Monge-Ampère equation on two dimensional strictly convex smooth domains. It uses a bubble enriched isoparametric cubic finite element space with exotic convexity enforcing degrees of freedom and is based on the methodology of C^0 interior penalty methods.

We have obtained an optimal $O(h^2)$ *a priori* error estimate in an H^2-like energy norm. We have also shown that a simple residual-based *a posteriori* error indicator can be used to monitor the convergence of solutions computed by an optimization algorithm.

Classical isoparametric finite element methods are designed for problems in H^1. However in this paper they are applied to problems in H^2 in the spirit of discontinuous Galerkin methods. Therefore the material in Appendices B–D, which extend several results for discontinuous Galerkin methods to the isoparametric setting, is also of independent interest.

APPENDIX A. DERIVATIONS OF (2.15) AND (2.16)

For the cubic polynomials $\hat{\varphi}_4$, $\hat{\varphi}_7$ and $\hat{\varphi}_{10}$ that appear in (2.3), we have the explicit formulas

$$D^2\hat{\varphi}_4 = \begin{pmatrix} 4\hat{x}_2 & 4\hat{x}_1 + 2\hat{x}_2 - 1 \\ 4\hat{x}_1 + 2\hat{x}_2 - 1 & 2\hat{x}_1 \end{pmatrix}, \quad D^2\hat{\varphi}_7 = \begin{pmatrix} 2\hat{x}_2 & 2\hat{x}_1 + 4\hat{x}_2 - 1 \\ 2\hat{x}_1 + 4\hat{x}_2 - 1 & 4\hat{x}_1 \end{pmatrix}$$

and

$$D^2\hat{\varphi}_{10} = \begin{pmatrix} -54\hat{x}_2 & 27 - 54\hat{x}_1 - 54\hat{x}_2 \\ 27 - 54\hat{x}_1 - 54\hat{x}_2 & -54\hat{x}_1 \end{pmatrix}.$$

Let $\ell = |\boldsymbol{p}_3 - \boldsymbol{p}_2| = |\boldsymbol{p}_2 - \boldsymbol{p}_3|$. A direct calculation yields

$$\begin{aligned}
\Delta\hat{\varphi}_4(\hat{x})[\ell\boldsymbol{e}_{23} - (\boldsymbol{p}_3 - \boldsymbol{p}_2)] + \Delta\hat{\varphi}_7(\hat{x})[\ell\boldsymbol{e}_{32} - (\boldsymbol{p}_2 - \boldsymbol{p}_3)]) + \frac{\Delta\hat{\varphi}_{10}(\hat{x})}{18}\ell(\boldsymbol{e}_{23} + \boldsymbol{e}_{32})& \\
= \hat{x}_2\big[4\big(\ell\boldsymbol{e}_{23} - (\boldsymbol{p}_3 - \boldsymbol{p}_2)\big) + 2\big(\ell\boldsymbol{e}_{32} - (\boldsymbol{p}_2 - \boldsymbol{p}_3)\big) - 3\ell(\boldsymbol{e}_{23} + \boldsymbol{e}_{32})\big]& \\
+ \hat{x}_1\big[2\big(\ell\boldsymbol{e}_{23} - (\boldsymbol{p}_3 - \boldsymbol{p}_2)\big) + 4\big(\ell\boldsymbol{e}_{32} - (\boldsymbol{p}_2 - \boldsymbol{p}_3)\big) - 3\ell(\boldsymbol{e}_{23} + \boldsymbol{e}_{32})\big]& \\
= \hat{x}_2\big[\big(\ell\boldsymbol{e}_{23} - (\boldsymbol{p}_3 - \boldsymbol{p}_2)\big) - \big(\ell\boldsymbol{e}_{32} - (\boldsymbol{p}_2 - \boldsymbol{p}_3)\big)\big]& \\
+ \hat{x}_1\big[\big(\ell\boldsymbol{e}_{32} - (\boldsymbol{p}_2 - \boldsymbol{p}_3)\big) - \big(\ell\boldsymbol{e}_{23} - (\boldsymbol{p}_3 - \boldsymbol{p}_2)\big)\big].&
\end{aligned} \tag{A.1}$$

The estimate (2.16) follows from (2.3), (2.6) and (A.1).

From the explicit formulas for $D^2\hat{\varphi}_4$, $D^2\hat{\varphi}_7$ and $D^2\hat{\varphi}_{10}$, we also have

$$0 = \frac{\partial^3\hat{\varphi}_4}{\partial\hat{x}_1^3} = \frac{\partial^3\hat{\varphi}_4}{\partial\hat{x}_2^3} = \frac{\partial^3\hat{\varphi}_7}{\partial\hat{x}_1^3} = \frac{\partial^3\hat{\varphi}_7}{\partial\hat{x}_2^3} = \frac{\partial^3\hat{\varphi}_{10}}{\partial\hat{x}_1^3} = \frac{\partial^3\hat{\varphi}_{10}}{\partial\hat{x}_2^3}, \tag{A.2}$$

which implies

$$\frac{\partial^3\hat{\varphi}_i}{\partial\hat{x}_1\partial^2\hat{x}_2} = \frac{\partial}{\partial\hat{x}_1}\Delta\hat{\varphi}_i(\hat{x}) \quad\text{and}\quad \frac{\partial^3\hat{\varphi}_i}{\partial\hat{x}_2\partial^2\hat{x}_1} = \frac{\partial}{\partial\hat{x}_2}\Delta\hat{\varphi}_i(\hat{x}) \qquad\text{for}\quad i = 4, 7, 10. \tag{A.3}$$

The estimate (2.15) follows from (2.3), (2.6) and (A.1)–(A.3).

Appendix B. A discrete Sobolev inequality

Lemma B.1. *There exists a positive constant C independent of h such that*

$$\begin{aligned}
\|\nabla v\|^2_{L^\infty(\Omega_h)} \le C\Big\{(1 + |\ln h|)^2\Big(\|D_h^2 v\|^2_{L^2(\Omega_h)} + \sum_{e\in\mathcal{E}_h^i} |e|^{-1}\|[\![\partial v/\partial n]\!]\|^2_{L^2(e)}\Big)& \\
+ (1 + |\ln h|)\max_{p\in\partial\Omega_h}\Big[\frac{\partial v}{\partial s}(p)\Big]^2\Big\} \qquad \forall\, v \in V_h.&
\end{aligned} \tag{B.1}$$

Proof. Let $v \in V_h$ be arbitrary and $F_T : \tilde{T} \longrightarrow T$ be the diffeomorphism defined by (2.18). We have, by the chain rule and (2.19),

$$D(v \circ F_T)(\tilde{x}) = Dv(F_T(\tilde{x}))DF_T(\tilde{x}) = Dv(F_T(\tilde{x}))[I + R(\tilde{x})] \qquad \forall\, \tilde{x} \in \tilde{T}. \tag{B.2}$$

Let $v_1 = \partial v/\partial x_1$, $v_2 = \partial v/\partial x_2$, $\tilde{v}_1 = \partial(v \circ F_T)/\partial\tilde{x}_1$ and $\tilde{v}_2 = \partial(v \circ F_T)/\partial\tilde{x}_2$. Note that the functions $\tilde{v}_1$ and $\tilde{v}_2$ on $\tilde{\Omega}_h$ are piecewise polynomial functions with respect to $\tilde{\mathcal{T}}_h$.

It follows from (2.22)–(2.24), (B.2), and the discrete Sobolev inequality in [24] for piecewise polynomial functions that

$$\begin{aligned}
\sum_{i=1}^2 \|v_i\|^2_{L^\infty(\Omega_h)} &\approx \sum_{i=1}^2 \|\tilde{v}_i\|^2_{L^\infty(\tilde{\Omega}_h)} \\
&\lesssim (1 + |\ln h|)\Big(\sum_{\tilde{T}\in\tilde{\mathcal{T}}_h}\sum_{i=1}^2 \|\tilde{v}_i\|^2_{H^1(\tilde{T})} + \sum_{e\in\tilde{\mathcal{E}}_h^i}\sum_{i=1}^2 |e|^{-1}\|[\![\tilde{v}_i]\!]\|^2_{L^2(e)}\Big) \\
&\lesssim (1+|\ln h|)\Big(\sum_{T\in\mathcal{T}_h}\sum_{i=1}^2 \big(\|v_i\|^2_{L^2(T)} + |v_i|^2_{H^1(T)}\big) + \sum_{e\in\mathcal{E}_h^i}\sum_{i=1}^2 |e|^{-1}\|[\![\tilde{v}_i]\!]\|^2_{L^2(e)}\Big).
\end{aligned} \tag{B.3}$$

Observe that (2.21) and (B.2) imply

$$\sum_{i=1}^{2}|e|^{-1}\|[\![\tilde{v}_i]\!]\|^2_{L^2(e)} \lesssim \sum_{i=1}^{2}\Big[|e|^{-1}\|[\![v_i]\!]\|^2_{L^2(e)} + h\big(\|v_i^+\|^2_{L^2(e)} + \|v_i^-\|^2_{L^2(e)}\big)\Big], \tag{B.4}$$

where $v_i^{\pm}$ is the restriction of v_i to $T_{\pm}$, the two elements in $\mathcal{T}_h$ that share the common edge $e \in \mathcal{E}_h^i$, and it follows from the trace theorem with scaling that

$$\begin{aligned} h\big(\|v_i^+\|^2_{L^2(e)} + \|v_i^-\|^2_{L^2(e)}\big) \lesssim \|v_i^+\|^2_{L^2(T_+)} + \|v_i^-\|^2_{L^2(T_-)} \\ + h^2\big(|v_i|^2_{H^1(T_+)} + |v_i|^2_{H^1(T_-)}\big). \end{aligned} \tag{B.5}$$

Putting (B.3)–(B.5) together, we have

(B.6)

$$\sum_{i=1}^{2}\|v_i\|^2_{L^\infty(\Omega_h)} \lesssim (1+|\ln h|)\sum_{i=1}^{2}\Big(\|v_i\|^2_{L^2(\Omega_h)} + \sum_{T\in\mathcal{T}_h}|v_i|^2_{H^1(T)} + \sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![v_i]\!]\|^2_{L^2(e)}\Big).$$

Let $\bar{v}_i$ be the average of v_i over Ω_h. It follows from a Poincaré-Friedrichs inequality for piecewise H^1 functions (cf. [23]) that

$$\|v_i - \bar{v}_i\|^2_{L^2(\Omega_h)} \lesssim \sum_{T\in\mathcal{T}_h}|v_i|^2_{H^1(T)} + \sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![v_i]\!]\|^2_{L^2(e)}. \tag{B.7}$$

Let $p_i \in \partial\Omega_h$ such that $(\partial v/\partial s)(p_i) = (\partial v/\partial x_i)(p_i) = v_i(p_i)$. We have

$$\begin{aligned} \sum_{i=1}^{2}\|\bar{v}_i\|^2_{L^2(\Omega_h)} &\lesssim \sum_{i=1}^{2}\|\bar{v}_i\|^2_{L^\infty(\Omega_h)} \\ &\lesssim \sum_{i=1}^{2}\Big(\|v_i(p_i)-\bar{v}_i\|^2_{L^\infty(\Omega_h)} + |v_i(p_i)|^2\Big) \\ &\le \sum_{i=1}^{2}\|v_i-\bar{v}_i\|^2_{L^\infty(\Omega_h)} + \sum_{i=1}^{2}|v_i(p_i)|^2 \\ &\lesssim (1+|\ln h|)\sum_{i=1}^{2}\Big(\sum_{T\in\mathcal{T}_h}|v_i|^2_{H^1(T)} + \sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![v_i]\!]\|^2_{L^2(e)}\Big) + \sum_{i=1}^{2}|v_i(p_i)|^2 \end{aligned} \tag{B.8}$$

by (B.6) (applied to $v_i - \bar{v}_i = \partial w/\partial x_i$, where the function $w(x) = v(x) - \bar{v}_1x_1 - \bar{v}_2x_2$ belongs to V_h) and (B.7).

Combining (B.6)–(B.8), we arrive at the estimate

$$\begin{aligned} \sum_{i=1}^{2}\|v_i\|^2_{L^\infty(\Omega_h)} \lesssim (1+|\ln h|)^2\sum_{i=1}^{2}\Big(\sum_{T\in\mathcal{T}_h}|v_i|^2_{H^1(T)} + \sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![v_i]\!]\|^2_{L^2(e)}\Big) \\ + (1+|\ln h|)\sum_{i=1}^{2}|v_i(p_i)|^2 \end{aligned}$$

that implies (B.1). □

The estimate (3.18) follows from (3.14), (3.16), Remark 3.6 and (B.1).

APPENDIX C. DERIVATIONS OF LEMMA 3.12

The inverse estimate $\|D^2 v\|_{L^\infty(T)} \leq Ch^{-1}\|D^2 v\|_{L^2(T)}$ is not valid for $v \in V_h$ if $T \in \mathcal{T}_h$ is a triangle that has a curved edge because $D^2 v$ is in general not a polynomial on T. But we can remedy this by the observation that it behaves like a polynomial up to a perturbation involving ∇v.

Lemma C.1. *There exists a positive constant C independent of h such that*

$$\|D^2 v\|_{L^\infty(T)} \leq C\big(h^{-1}\|D^2 v\|_{L^2(T)} + h^4\|\nabla v\|_{L^\infty(T)}\big) \qquad \forall\, v \in V_h, T \in \mathcal{T}_h. \tag{C.1}$$

Proof. Let $v \in V_h$ and $T \in \mathcal{T}_h$ be arbitrary. We have the following analog of (2.17):

$$\begin{aligned} D^2(v \circ F_T)(\tilde{x}) = DF_T(\tilde{x})^t (D^2 v)(F_T(\tilde{x}))DF_T(\tilde{x}) + \frac{\partial v}{\partial x_1}(F_T(\tilde{x}))D^2\tilde{\phi}_{T,1}(\tilde{x}) \\ + \frac{\partial v}{\partial x_2}(F_T(\tilde{x}))D^2\tilde{\phi}_{T,2}(\tilde{x}) \qquad \forall\, \tilde{x} \in \tilde{T}, \end{aligned} \tag{C.2}$$

where $\tilde{\phi}_{T,1}$ and $\tilde{\phi}_{T,2}$ are the first and second components of F_T respectively.

Note that

(C.3) the components of $D^2(v \circ F_T)(\tilde{x})$, $D(v \circ F_T)(\tilde{x})$, $DF_T(\tilde{x})$, $D^2\tilde{\phi}_{T,1}(\tilde{x})$ and $D^2\tilde{\phi}_{T,2}(\tilde{x})$ are polynomials in $\tilde{x}$,

and

$$\|D^2\tilde{\phi}_{T,1}\|_{L^\infty(\tilde{T})} + \|D^2\tilde{\phi}_{T,2}\|_{L^\infty(\tilde{T})} \lesssim 1 \tag{C.4}$$

by (2.10), (2.11) and (2.14).

We also have, by (B.2),

$$(Dv)(F_T(\tilde{x}))[I - R^4(\tilde{x})] = D(v \circ F_T)(\tilde{x})[I - R(\tilde{x}) + R^2(\tilde{x}) - R^3(\tilde{x})] \qquad \forall\, \tilde{x} \in \tilde{T}. \tag{C.5}$$

It follows from (2.21) and (C.5) that

$$(Dv)(F_T(\tilde{x})) = D(v \circ F_T)(\tilde{x})[I - R(\tilde{x}) + R^2(\tilde{x}) - R^3(\tilde{x})] + S(\tilde{x}), \tag{C.6}$$

where

$$\|S\|_{L^\infty(\tilde{T})} \lesssim h^4\|\nabla v\|_{L^\infty(T)}. \tag{C.7}$$

Combining (2.19)–(2.21), Definition 2.6, (C.2)–(C.4), (C.6) and (C.7), we arrive at the following relation

$$DF_T(\tilde{x})^t (D^2 v)(F_T(\tilde{x}))DF_T(\tilde{x}) = H(\tilde{x}) + Z(\tilde{x}) \qquad \forall\, \tilde{x} \in \tilde{T}, \tag{C.8}$$

where

(C.9) the components of $H(\tilde{x})$ are polynomials in $\tilde{x}$ of total degree ≤ 13

and

$$\|Z\|_{L^\infty(\tilde{T})} \lesssim h^4\|\nabla v\|_{L^\infty(T)}. \tag{C.10}$$

It follows from (2.22), (2.23) and (C.8) that

$$\|H\|_{L^2(\tilde{T})} \lesssim \|(D^2 v) \circ F_T\|_{L^2(\tilde{T})} + \|Z\|_{L^2(\tilde{T})} \lesssim \|D^2 v\|_{L^2(T)} + \|Z\|_{L^2(\tilde{T})},$$

which together with (C.10) and a standard inverse estimate (cf. [30, 41]) implies

$$\begin{aligned}\|D^2v\|_{L^\infty(T)} &= \|(D^2v)\circ F_T\|_{L^\infty(\tilde{T})}\\ &\lesssim \|H+Z\|_{L^\infty(\tilde{T})}\\ &\lesssim \|H\|_{L^\infty(\tilde{T})}+\|Z\|_{L^\infty(\tilde{T})}\\ &\lesssim h^{-1}\|H\|_{L^2(\tilde{T})}+\|Z\|_{L^\infty(\tilde{T})}\\ &\lesssim h^{-1}\|D^2v\|_{L^2(T)}+\|Z\|_{L^\infty(\tilde{T})}\lesssim h^{-1}\|D^2v\|_{L^2(T)}+h^4\|\nabla v\|_{L^\infty(T)}\end{aligned}$$

which is the estimate (C.1). □

The estimate (3.19) follows immediately from (3.14), (3.18) and Lemma C.1.

Now we take $v=u_h$ in (C.8) and conclude that

$$DF_T(\tilde{x})^t\big[(D^2u_h)(F_T(\tilde{x}))+Q(\tilde{x})\big]DF_T(\tilde{x})=H(\tilde{x}), \tag{C.11}$$

where

$$Q(\tilde{x})=-DF_T(\tilde{x})^{-t}Z(\tilde{x})DF_T(\tilde{x})^{-1}$$

satisfies

$$\|Q\|_{L^\infty(\tilde{T})}\lesssim h^3 \tag{C.12}$$

by (3.18) and (C.10).

Note that

$$\|\det\big[(D^2u_h)\circ F_T+Q\big]-\det\big[(D^2u_h)\circ F_T\big]\|_{L^\infty(\tilde{T})}\lesssim h^2 \tag{C.13}$$

by (3.19) and (C.12).

Since $\psi\in H^2(\mathbb{R}^2)$ (cf. Remark 3.3), we can use (2.24) to show that the P_1 interpolant $\psi_{\tilde{T}}$ of $\psi\circ F_T$ on $\tilde{T}$ satisfies

$$\|\psi\circ F_T-\psi_{\tilde{T}}\|_{L^\infty(\tilde{T})}\lesssim h. \tag{C.14}$$

It then follows from (2.22), (2.23), (3.17), (C.9), (C.11), (C.13), (C.14) and a standard inverse estimate that

$$\begin{aligned}\|\det H-(\det DF_T)^2\psi_{\tilde{T}}\|_{L^\infty(\tilde{T})} &\lesssim h^{-1}\|\det H-(\det DF_T)^2\psi_{\tilde{T}}\|_{L^2(\tilde{T})}\\ &\lesssim h^{-1}\|\det\big[(D^2u_h)\circ F_T+Q\big]-\psi_{\tilde{T}}\|_{L^2(\tilde{T})}\\ &\lesssim h^{-1}\|\det\big[(D^2u_h)\circ F_T\big]-\psi\circ F_T\|_{L^2(\tilde{T})}+h\\ &\lesssim h^{-1}\|\det D^2u_h-\psi\|_{L^2(T)}+h\\ &\lesssim h.\end{aligned} \tag{C.15}$$

Finally we arrive at the estimate (3.20)

$$\begin{aligned}\|\det D^2u_h-\psi\|_{L^\infty(T)} &= \|\det\big[(D^2u_h)\circ F_T\big]-\psi\circ F_T\|_{L^\infty(\tilde{T})}\\ &\lesssim \|\det\big[(D^2u_h)\circ F_T+Q\big]-\psi_{\tilde{T}}\|_{L^\infty(\tilde{T})}+h\\ &\lesssim \|\det H-(\det DF_T)^2\psi_{\tilde{T}}\|_{L^\infty(\tilde{T})}+h\\ &\lesssim h\end{aligned}$$

by (2.22), (C.11) and (C.13)–(C.15).

APPENDIX D. DERIVATION OF LEMMA 4.1

There are two ingredients in the derivation of Lemma 4.1. The first one is a finite element space $W_h \subset H_0^1(\Omega_h)$ associated with the isoparametric mesh $\mathcal{T}_h$. The second one is a linear map E_h that connects $V_h \cap H_0^1(\Omega_h)$ and W_h.

The finite element space W_h. A function $w \in H_0^1(\Omega_h)$ belongs to W_h if and only if (i) $w \circ \Phi_T$ belongs to $P_3(\hat{T}) \oplus \varphi_{\hat{T}}^2 P_1(\hat{T})$ for all $T \in \mathcal{T}_h$, where $\Phi_T : \hat{T} \longrightarrow T$ is the cubic isoparametric map; and (ii) w is continuous up to the first order derivatives at the vertices of $\mathcal{T}_h$. The 13 dofs of $w \circ \Phi_T$ on the reference simples $\hat{T}$ are given by the values of its derivatives up to order 1 at the vertices of $\hat{T}$, its value at the center $c_{\hat{T}}$ of $\hat{T}$ and the values of its Laplacian at the vertices of $\hat{T}_\dagger$ (cf. Section 2.5).

Since the elements in $\mathcal{T}_h$ are convex and piecewise smooth, we can apply [63, Theorem 3.1.1.2] to obtain the following estimate for any function $w \in H_0^1(\Omega_h)$ that is piecewise smooth with respect to $\mathcal{T}_h$.

$$\sum_{T\in\mathcal{T}_h}\Big[\int_T |\Delta w|^2 dx - |w|_{H^2(T)}^2\Big] \geq \sum_{T\in\mathcal{T}_h}\Big[\sum_{e\in\mathcal{E}^i(T)}\int_e \Big\{\frac{\partial}{\partial s}\Big(\frac{\partial w}{\partial n}\frac{\partial w}{\partial s}\Big) - 2\frac{\partial w}{\partial s}\frac{\partial^2 w}{\partial s\partial n}\Big\}ds\Big], \tag{D.1}$$

where $\mathcal{E}^i(T)$ are the edges of T interior to Ω_h and $\partial/\partial s$ (resp., $\partial/\partial n$) denotes the counterclockwise tangential (resp., outward normal) differentiation along ∂T.

The continuity up to first order derivatives at the vertices of $\mathcal{T}_h$ for $w \in W_h$ implies that

$$\sum_{T\in\mathcal{T}_h}\Big[\sum_{e\in\mathcal{E}^i(T)}\int_e \frac{\partial}{\partial s}\Big(\frac{\partial w}{\partial n}\frac{\partial w}{\partial s}\Big)ds\Big] = 0 \qquad \forall\, w \in W_h, \tag{D.2}$$

and hence also

$$\sum_{T\in\mathcal{T}_h}\Big[\sum_{e\in\mathcal{E}^i(T)}\int_e \Big(-2\frac{\partial w}{\partial s}\frac{\partial^2 w}{\partial s\partial n}\Big)ds\Big] = \sum_{T\in\mathcal{T}_h}\Big[\sum_{e\in\mathcal{E}^i(T)}\int_e \Big(2\frac{\partial^2 w}{\partial s^2}\frac{\partial w}{\partial n}\Big)ds\Big] \qquad \forall\, w \in W_h. \tag{D.3}$$

It follows from the Cauchy-Schwarz inequality that

$$\Big|\sum_{T\in\mathcal{T}_h}\Big[\sum_{e\in\mathcal{E}^i(T)}\int_e \Big(2\frac{\partial^2 w}{\partial s^2}\frac{\partial w}{\partial n}\Big)ds\Big]\Big| \leq C_\sharp\Big(\sum_{e\in\mathcal{E}_h^i}|e|\|\partial^2 w/\partial s^2\|_{L^2(e)}^2\Big)^{\frac12}\Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial w/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac12} \qquad \forall\, w \in W_h, \tag{D.4}$$

where the positive constant $C_\sharp$ only depends on the shape regularity of $\mathcal{T}_h$.

Putting (D.1)–(D.4) together we find

$$\sum_{T\in\mathcal{T}_h}|w|_{H^2(T)}^2 \leq C_\sharp\Big(\sum_{e\in\mathcal{E}_h^i}|e|\|\partial^2 w/\partial s^2\|_{L^2(e)}^2\Big)^{\frac12}\Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial w/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac12} + \sum_{T\in\mathcal{T}_h}\|\Delta w\|_{L^2(T)}^2 \qquad \forall\, w \in W_h. \tag{D.5}$$

Let $e \in \mathcal{E}_h^i$ be an edge of $T \in \mathcal{T}_h$ and $F_T : \tilde{T} \longrightarrow T$ be the cubic polynomial map defined by (2.18), where $\tilde{T} \in \tilde{\mathcal{T}}_h$ shares the same vertices with T. The map

$$v \mapsto |v \circ F_T^{-1}|_{H^2(T)}$$

defines a semi-norm on $P_3(\tilde{T}) \oplus \varphi_{\tilde{T}}^2 P_1(\tilde{T})$ whose kernel is $K = \{v \in P_3(\tilde{T}) \oplus \varphi_{\tilde{T}}^2 P_1(\tilde{T}) : v \circ F_T^{-1} \text{ is linear}\}$ = the three dimensional subspace of $P_3(\tilde{T}) \oplus \varphi_{\tilde{T}}^2 P_1(\tilde{T})$ spanned by the constant function 1 and the two components of F_T. Since the restrictions of these functions to e are polynomials of degree ≤ 1, we can deduce from the equivalence of norms on $[P_3(\tilde{T}) \oplus \varphi_{\tilde{T}}^2 P_1(\tilde{T})]/K$ and scaling that

$$|e|^{\frac{1}{2}} \|\partial^2 v/\partial s^2\|_{L^2(e)} \lesssim |v \circ F_T^{-1}|_{H^2(T)} \qquad \forall\, v \in P_3(\tilde{T}) \oplus \varphi_{\tilde{T}}^2 P_1(\tilde{T}). \tag{D.6}$$

Letting $v = w \circ F_T$ in (D.6) for $w \in W_h$, we arrive at the estimate

$$|e|^{\frac{1}{2}} \|\partial^2 w/\partial s^2\|_{L^2(e)} \lesssim |w|_{H^2(T)}$$

and hence

$$\sum_{e \in \mathcal{E}_h^i} |e| \|\partial^2 w/\partial s^2\|_{L^2(e)}^2 \lesssim \sum_{T \in \mathcal{T}_h} |w|_{H^2(T)}^2 \qquad \forall\, w \in W_h. \tag{D.7}$$

Combining (D.5) and (D.7), we have

$$\sum_{T \in \mathcal{T}_h} |w|_{H^2(T)}^2 \leq C_\flat \Big(\sum_{T \in \mathcal{T}_h} |w|_{H^2(T)}^2 \Big)^{\frac{1}{2}} \Big(\sum_{e \in \mathcal{E}_h^i} |e|^{-1} \|[\![\partial w/\partial n]\!]\|_{L^2(e)}^2 \Big)^{\frac{1}{2}} + \Big(\sum_{T \in \mathcal{T}_h} \|\Delta w\|_{L^2(T)}^2 \Big)^{\frac{1}{2}} \Big(\sum_{T \in \mathcal{T}_h} |w|_{H^2(T)}^2 \Big)^{\frac{1}{2}}$$

and hence

$$\|D_h^2 w\|_{L^2(\Omega_h)} \leq \|\Delta_h w\|_{L^2(\Omega_h)} + C_\flat \Big(\sum_{e \in \mathcal{E}_h^i} |e|^{-1} \|[\![\partial w/\partial n]\!]\|_{L^2(e)}^2 \Big)^{\frac{1}{2}} \qquad \forall\, w \in W_h, \tag{D.8}$$

where Δ_h denotes the piecewise Laplacian operator with respect to $\mathcal{T}_h$.

The estimate (D.8) is a discrete version of the Miranda-Talenti estimate [77, 92] that plays an important role in the theory of second order elliptic problems in nondivergence form. Below we will derive an analog of (D.8) for functions in $V_h \cap H_0^1(\Omega_h)$ through a map E_h that connects $V_h \cap H_0^1(\Omega_h)$ and W_h.

The map E_h. The map $E_h : V_h \cap H_0^1(\Omega_h) \longrightarrow W_h$ is given by

$$E_h v = w,$$

where $w \in W_h$ is defined by the conditions that (i) $v \circ \Phi_T - w \circ \Phi_T \in P_3(\hat{T})$, (ii) the dofs of w at a vertex p of $\mathcal{T}_h$ interior to Ω are the averages of the corresponding dofs of v at p on the elements in $\mathcal{T}_h$ that share p as a common vertex; (iii) the dofs of w at a vertex p of $\mathcal{T}_h$ on $\partial\Omega_h$ are the averages of the corresponding dofs of v at p on the two curved elements that share p as a common vertex; and (iv) $w = v$ at $\Phi_T(c_{\hat{T}})$ for all $T \in \mathcal{T}_h$, where $c_{\hat{T}}$ is the center of the reference simplex $\hat{T}$. Note that $v = 0$ on $\partial\Omega_h$ and condition (ii) imply $w = 0$ on $\partial\Omega_h$.

Let $v \in V_h$ and $T \in \mathcal{T}_h$ be arbitrary. It follows from (by now) standard arguments (cf. [20–22, 25]), (2.10) and (2.22) that

$$\begin{aligned} \text{(D.9)} \qquad \|(v - E_h v) \circ F_T\|_{L^2(\tilde{T})}^2 &\lesssim h_{\tilde{T}}^4 \sum_{p \in \mathcal{V}_{\tilde{T}}} \big|\nabla(v \circ F_T)(p) - \nabla((E_h v) \circ F_T)(p)\big|^2 \\ &\lesssim h_T^4 \sum_{p \in \mathcal{V}_{\tilde{T}}} \big|(\nabla v)(p) - \nabla(E_h v)(p)\big|^2, \end{aligned}$$

where $\mathcal{V}_{\tilde{T}}$ is the set of the three vertices of $\tilde{T}$.

Let $p \in \mathcal{V}_{\tilde{T}}$. We separate the estimate for the right-hand side of (D.9) into two cases. In the first case, all of the elements in $\mathcal{T}_h$ that share p as a common vertex are triangles that have at most one vertex on $\partial\Omega$. In this case, we have

$$\text{(D.10)} \qquad \big|(\nabla v)(p) - \nabla(E_h v)(p)\big|^2 \lesssim \sum_{e \in \mathcal{E}_p} |e|^{-1} \|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2$$

by a standard inverse estimate, where $\mathcal{E}_p$ is the set of the edges in $\mathcal{E}_h^i$ that share p as a common vertex. Note that we can apply the inverse estimate because v is a polynomial on every triangle that shares p as a common vertex.

In the second case, at least one of the triangles that share p as a common vertex has a curved edge. In this case we have

$$\text{(D.11)} \qquad \big|(\nabla v)(p) - \nabla(E_h v)(p)\big|^2 \lesssim \sum_{e \in \mathcal{E}_p} |e|^{-1} \|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2 + h_T^8 \sum_{T' \in \mathcal{T}_p} \|\nabla v\|_{L^\infty(T')}^2$$

by (C.6) and (C.7), where $\mathcal{T}_p$ is the set of the elements in $\mathcal{T}_h$ that share p as a common vertex.

Combining (D.9)–(D.11), we find

(D.12)

$$\Big(\sum_{\tilde{T} \in \tilde{\mathcal{T}}_h} \|(v - E_h v) \circ F_T\|_{L^2(\tilde{T})}^2\Big)^{\frac{1}{2}} \lesssim h^2 \Big[\Big(\sum_{e \in \mathcal{E}_h^i} |e|^{-1} \|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}} + h^3 \|\nabla v\|_{L^\infty(\Omega_h)}\Big],$$

where we have also used the observation that the number of elements in $\mathcal{T}_h$ is $O(h^{-2})$.

Finally it follows from and standard inverse estimates that

$$\begin{aligned} \|D_h^2(v - E_h v)\|_{L^2(\Omega_h)} &\lesssim \Big(\sum_{\tilde{T} \in \tilde{\mathcal{T}}_h} \big[\|(v - E_h v) \circ F_T\|_{L^2(\tilde{T})}^2 + |(v - E_h v) \circ F_T|_{H^1(\tilde{T})}^2 \\ \text{(D.13)} \qquad &\qquad + |(v - E_h v) \circ F_T|_{H^2(\tilde{T})}^2\big]\Big)^{\frac{1}{2}} \\ &\lesssim \Big(\sum_{e \in \mathcal{E}_h^i} |e|^{-1} \|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac{1}{2}} + h^3 \|\nabla v\|_{L^\infty(\Omega_h)}, \end{aligned}$$

where the hidden constant only depends on the shape regularity of $\mathcal{T}_h$.

Similarly it follows from the trace theorem with scaling, (2.24), (D.12) and standard inverse estimates that

$$
\begin{aligned}
&\Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial(v-E_hv)/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac12}\\
&\lesssim\Big(h^{-2}\|\nabla(v-E_hv)\|_{L^2(\Omega_h)}^2+\|D_h^2(v-E_hv)\|_{L^2(\Omega_h)}^2\Big)^{\frac12}\\
&\lesssim\Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac12}+h^3\|\nabla v\|_{L^\infty(\Omega_h)}.
\end{aligned}
\tag{D.14}
$$

We are now ready to establish a discrete version of the Miranda-Talenti estimate for functions in V_h.

Lemma D.1. *There exists a positive constant $C_\dagger$ independent of h such that*

$$
\|D_h^2v\|_{L^2(\Omega_h)}\le\|\Delta_hv\|_{L^2(\Omega_h)}+C_\dagger\Big[\Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac12}+h^3\|\nabla v\|_{L^\infty(\Omega_h)}\Big]
$$

for all $v\in V_h\cap H_0^1(\Omega_h)$.

Proof. This is a simple consequence of (D.8), (D.13) and (D.14):

$$
\begin{aligned}
\|D_h^2v\|_{L^2(\Omega_h)}&\le\|D_h^2(E_hv)\|_{L^2(\Omega_h)}+\|D_h^2(v-E_hv)\|_{L^2(\Omega_h)}\\
&\le\|\Delta_h(E_hv)\|_{L^2(\Omega_h)}+C_{\maltese}\Big[\Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac12}+h^3\|\nabla v\|_{L^\infty(\Omega_h)}\Big]\\
&\le\|\Delta_hv\|_{L^2(\Omega_h)}+C_\dagger\Big[\Big(\sum_{e\in\mathcal{E}_h^i}|e|^{-1}\|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2\Big)^{\frac12}+h^3\|\nabla v\|_{L^\infty(\Omega_h)}\Big].
\end{aligned}
$$

□

Derivation of Lemma 4.1. We follow the treatment of second order elliptic equations in nondivergence form (cf. [38, 43, 76, 91]) by introducing the function

$$
\gamma_h(x)=\frac{A_h(x):I}{A_h(x):A_h(x)},
$$

where I is the 2×2 identity matrix.

We have (cf. for example [32, Appendix A])

$$
0\le\gamma_h(x)\le\frac{1}{\alpha}\quad\text{a.e. in }\Omega_h,
\tag{D.15}
$$

and

$$
|\gamma_h(x)A_h(x)-I|\le\delta=\frac{\beta-\alpha}{(\alpha^2+\beta^2)^{\frac12}}<1\quad\text{a.e. in }\Omega_h,
\tag{D.16}
$$

where $\alpha>0$ and $\beta\ge\alpha$ are the constants in (4.11).

Let $v \in V_h \cap H_0^1(\Omega_h)$ be arbitrary. It follows from Lemma D.1, (D.16), and the Cauchy-Schwarz inequality that

$$\int_{\Omega_h} (\gamma_h A_h : D_h^2 v)(\Delta_h v)dx = \|\Delta_h v\|_{L^2(\Omega_h)}^2 + \int_{\Omega_h} [(\gamma_h A_h - I) : D_h^2 v](\Delta_h v)dx$$

$$\geq \|\Delta_h v\|_{L^2(\Omega_h)}^2 - \delta \|D_h^2 v\|_{L^2(\Omega_h)} \|\Delta_h v\|_{L^2(\Omega_h)}$$

$$\geq \|\Delta_h v\|_{L^2(\Omega_h)}^2 - \delta \Big\{ \Big[\|\Delta_h v\|_{L^2(\Omega_h)}$$

$$+ C_\dagger \Big[\Big(\sum_{e \in \mathcal{E}_h^i} |e|^{-1} \|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2 \Big)^{\frac{1}{2}} + h^3 \|\nabla v\|_{L^\infty(\Omega_h)} \Big] \Big\} \|\Delta_h v\|_{L^2(\Omega_h)}$$

$$= (1-\delta)\|\Delta_h v\|_{L^2(\Omega_h)}^2 - \delta C_\dagger \Big[\Big(\sum_{e \in \mathcal{E}_h^i} |e|^{-1} \|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2 \Big)^{\frac{1}{2}} + h^3 \|\nabla v\|_{L^\infty(\Omega_h)} \Big] \|\Delta_h v\|_{L^2(\Omega_h)}.$$

Consequently, we have

(D.17)

$$\|\Delta_h v\|_{L^2(\Omega_h)} \leq \Big(\frac{\alpha^{-1}}{1-\delta} \Big) \|A_h : D_h^2 v\|_{L^2(\Omega_h)} + \Big(\frac{\delta C_\dagger}{1-\delta} \Big) \Big[\Big(\sum_{e \in \mathcal{E}_h^i} |e|^{-1} \|[\![\partial v/\partial n]\!]\|_{L^2(e)}^2 \Big)^{\frac{1}{2}} + h^3 \|\nabla v\|_{L^\infty(\Omega_h)} \Big]$$

by using (D.15) and the Cauchy-Schwarz inequality.

The estimate (4.13) follows from Lemma D.1 and (D.17).

References

[1] R. A. Adams and J. J. F. Fournier, *Sobolev spaces*, 2nd ed., Pure and Applied Mathematics (Amsterdam), vol. 140, Elsevier/Academic Press, Amsterdam, 2003. MR2424078

[2] J. Adetola, B. Ahounou, G. Awanou, and H. Guo, *Low order mixed finite element approximations of the Monge-Ampère equation*, Int. J. Numer. Anal. Model. **19** (2022), no. 5, 669–684. MR4483072

[3] A. D. Aleksandrov, *Dirichlet's problem for the equation* $\mathrm{Det}\,\|z_{ij}\| = \varphi(z_1, \cdots, z_n, z, x_1, \cdots, x_n)$. *I* (Russian, with English summary), Vestnik Leningrad. Univ. Ser. Mat. Meh. Astronom. **13** (1958), no. 1, 5–24. MR96903

[4] A. D. Alexandrov, *Convex polyhedra*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2005. Translated from the 1950 Russian edition by N. S. Dairbekov, S. S. Kutateladze and A. B. Sossinsky; With comments and bibliography by V. A. Zalgaller and appendices by L. A. Shor and Yu. A. Volkov. MR2127379

[5] G. Awanou, *Pseudo transient continuation and time marching methods for Monge-Ampère type equations*, Adv. Comput. Math. **41** (2015), no. 4, 907–935, DOI 10.1007/s10444-014-9391-y. MR3424088

[6] G. Awanou, *Spline element method for Monge-Ampère equations*, BIT **55** (2015), no. 3, 625–646, DOI 10.1007/s10543-014-0524-y. MR3401807

[7] G. Awanou, *Standard finite elements for the numerical resolution of the elliptic Monge-Ampère equations: classical solutions*, IMA J. Numer. Anal. **35** (2015), no. 3, 1150–1166, DOI 10.1093/imanum/dru028. MR3407257

[8] G. Awanou, *On standard finite difference discretizations of the elliptic Monge-Ampère equation*, J. Sci. Comput. **69** (2016), no. 2, 892–904, DOI 10.1007/s10915-016-0220-y. MR3551348

[9] G. Awanou, *Standard finite elements for the numerical resolution of the elliptic Monge-Ampère equation: Aleksandrov solutions*, ESAIM Math. Model. Numer. Anal. **51** (2017), no. 2, 707–725, DOI 10.1051/m2an/2016037. MR3626416

[10] G. Awanou and H. Li, *Error analysis of a mixed finite element method for the Monge-Ampère equation*, Int. J. Numer. Anal. Model. **11** (2014), no. 4, 745–761. MR3218347

[11] G. Awanou, H. Li, and E. Malitz, *A two-grid method for the C^0 interior penalty discretization of the Monge-Ampère equation*, J. Comput. Math. **38** (2020), no. 4, 547–546, DOI 10.4208/jcm.1901-m2018-0039. MR4087801

[12] I. J. Bakelman, *Convex analysis and nonlinear geometric elliptic equations*, Springer-Verlag, Berlin, 1994. With an obituary for the author by William Rundell; Edited by Steven D. Taliaferro, DOI 10.1007/978-3-642-69881-1. MR1305147

[13] G. Barles and P. E. Souganidis, *Convergence of approximation schemes for fully nonlinear second order equations*, Asymptotic Anal. **4** (1991), no. 3, 271–283. MR1115933

[14] J.-D. Benamou, F. Collino, and J.-M. Mirebeau, *Monotone and consistent discretization of the Monge-Ampère operator*, Math. Comp. **85** (2016), no. 302, 2743–2775, DOI 10.1090/mcom/3080. MR3522969

[15] J.-D. Benamou, B. D. Froese, and A. M. Oberman, *Two numerical methods for the elliptic Monge-Ampère equation*, M2AN Math. Model. Numer. Anal. **44** (2010), no. 4, 737–758, DOI 10.1051/m2an/2010017. MR2683581

[16] K. Böhmer, *On finite element methods for fully nonlinear elliptic equations of second order*, SIAM J. Numer. Anal. **46** (2008), no. 3, 1212–1249, DOI 10.1137/040621740. MR2390991

[17] K. Böhmer and R. Schaback, *A meshfree method for solving the Monge-Ampère equation*, Numer. Algorithms **82** (2019), no. 2, 539–551, DOI 10.1007/s11075-018-0612-1. MR4003757

[18] G. Bonnet and J.-M. Mirebeau, *Monotone discretization of the Monge-Ampère equation of optimal transport*, ESAIM Math. Model. Numer. Anal. **56** (2022), no. 3, 815–865, DOI 10.1051/m2an/2022029. MR4411481

[19] J. H. Bramble and S. R. Hilbert, *Estimation of linear functionals on Sobolev spaces with application to Fourier transforms and spline interpolation*, SIAM J. Numer. Anal. **7** (1970), 112–124, DOI 10.1137/0707006. MR263214

[20] S. C. Brenner, *Two-level additive Schwarz preconditioners for nonconforming finite elements*, Domain decomposition methods in scientific and engineering computing (University Park, PA, 1993), Contemp. Math., vol. 180, Amer. Math. Soc., Providence, RI, 1994, pp. 9–14, DOI 10.1090/conm/180/01951. MR1312372

[21] S. C. Brenner, *Two-level additive Schwarz preconditioners for nonconforming finite element methods*, Math. Comp. **65** (1996), no. 215, 897–921, DOI 10.1090/S0025-5718-96-00746-6. MR1348039

[22] S. C. Brenner, *Convergence of nonconforming multigrid methods without full elliptic regularity*, Math. Comp. **68** (1999), no. 225, 25–53, DOI 10.1090/S0025-5718-99-01035-2. MR1620215

[23] S. C. Brenner, *Poincaré-Friedrichs inequalities for piecewise H^1 functions*, SIAM J. Numer. Anal. **41** (2003), no. 1, 306–324, DOI 10.1137/S0036142902401311. MR1974504

[24] S. C. Brenner, *Discrete Sobolev and Poincaré inequalities for piecewise polynomial functions*, Electron. Trans. Numer. Anal. **18** (2004), 42–48. MR2083293

[25] S. C. Brenner, *C^0 interior penalty methods*, Frontiers in numerical analysis—Durham 2010, Lect. Notes Comput. Sci. Eng., vol. 85, Springer, Heidelberg, 2012, pp. 79–147, DOI 10.1007/978-3-642-23914-4_2. MR3051409

[26] S. C. Brenner, T. Gudi, M. Neilan, and L.-Y. Sung, *$\mathcal{C}^0$ penalty methods for the fully nonlinear Monge-Ampère equation*, Math. Comp. **80** (2011), no. 276, 1979–1995, DOI 10.1090/S0025-5718-2011-02487-7. MR2813346

[27] S. C. Brenner and E. L. Kawecki, *Adaptive C^0 interior penalty methods for Hamilton-Jacobi-Bellman equations with Cordes coefficients*, J. Comput. Appl. Math. **388** (2021), Paper No. 113241, 17, DOI 10.1016/j.cam.2020.113241. MR4199789

[28] S. C. Brenner and M. Neilan, *Finite element approximations of the three dimensional Monge-Ampère equation*, ESAIM Math. Model. Numer. Anal. **46** (2012), no. 5, 979–1001, DOI 10.1051/m2an/2011067. MR2916369

[29] S. C. Brenner, M. Neilan, A. Reiser, and L.-Y. Sung, *A C^0 interior penalty method for a von Kármán plate*, Numer. Math. **135** (2017), no. 3, 803–832, DOI 10.1007/s00211-016-0817-y. MR3606463

[30] S. C. Brenner and L. R. Scott, *The mathematical theory of finite element methods*, 3rd ed., Texts in Applied Mathematics, vol. 15, Springer, New York, 2008, DOI 10.1007/978-0-387-75934-0. MR2373954

[31] 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** (2005), 83–118, DOI 10.1007/s10915-004-4135-7. MR2142191

[32] S. C. Brenner, L.-Y. Sung, Z. Tan, and H. Zhang, *A convexity enforcing C^0 interior penalty method for the Monge-Ampère equation on convex polygonal domains*, Numer. Math. **148** (2021), no. 3, 497–524, DOI 10.1007/s00211-021-01210-x. MR4293961

[33] S. C. Brenner, K. Wang, and J. Zhao, *Poincaré-Friedrichs inequalities for piecewise H^2 functions*, Numer. Funct. Anal. Optim. **25** (2004), no. 5-6, 463–478, DOI 10.1081/NFA-200042165. MR2106270

[34] J. Brusca and B. F. Hamfeldt, *A convergent quadrature-based method for the Monge-Ampère equation*, SIAM J. Sci. Comput. **45** (2023), no. 3, A1097–A1124, DOI 10.1137/22M1494658. MR4589123

[35] A. Caboussat, R. Glowinski, and D. C. Sorensen, *A least-squares method for the numerical solution of the Dirichlet problem for the elliptic Monge-Ampère equation in dimension two* (English, with English and French summaries), ESAIM Control Optim. Calc. Var. **19** (2013), no. 3, 780–810, DOI 10.1051/cocv/2012033. MR3092362

[36] L. A. Caffarelli and X. Cabré, *Fully nonlinear elliptic equations*, American Mathematical Society Colloquium Publications, vol. 43, American Mathematical Society, Providence, RI, 1995, DOI 10.1090/coll/043. MR1351007

[37] L. Caffarelli, L. Nirenberg, and J. Spruck, *The Dirichlet problem for nonlinear second-order elliptic equations. I. Monge-Ampère equation*, Comm. Pure Appl. Math. **37** (1984), no. 3, 369–402, DOI 10.1002/cpa.3160370306. MR739925

[38] S. Campanato, *A Cordes type condition for nonlinear nonvariational systems* (English, with Italian summary), Rend. Accad. Naz. Sci. XL Mem. Mat. (5) **13** (1989), no. 1, 307–321. MR1041758

[39] H. Chen, X. Feng, and Z. Zhang, *A recovery-based linear C^0 finite element method for a fourth-order singularly perturbed Monge-Ampère equation*, Adv. Comput. Math. **47** (2021), no. 2, Paper No. 21, 37, DOI 10.1007/s10444-021-09847-w. MR4220020

[40] S. Y. Cheng and S. T. Yau, *On the regularity of the Monge-Ampère equation* $\det(\partial^2 u/\partial x_i \partial s x_j) = F(x,u)$, Comm. Pure Appl. Math. **30** (1977), no. 1, 41–68, DOI 10.1002/cpa.3160300104. MR437805

[41] P. G. Ciarlet, *The finite element method for elliptic problems*, Studies in Mathematics and its Applications, Vol. 4, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. MR520174

[42] P. G. Ciarlet and P.-A. Raviart, *Interpolation theory over curved elements, with applications to finite element methods*, Comput. Methods Appl. Mech. Engrg. **1** (1972), 217–249, DOI 10.1016/0045-7825(72)90006-0. MR375801

[43] H. O. Cordes, *Über die erste Randwertaufgabe bei quasilinearen Differentialgleichungen zweiter Ordnung in mehr als zwei Variablen* (German), Math. Ann. **131** (1956), 278–312, DOI 10.1007/BF01342965. MR91400

[44] O. Davydov and A. Saeed, *Numerical solution of fully nonlinear elliptic equations by Böhmer's method*, J. Comput. Appl. Math. **254** (2013), 43–54, DOI 10.1016/j.cam.2013.03.009. MR3061065

[45] E. J. Dean and R. Glowinski, *Numerical solution of the two-dimensional elliptic Monge-Ampère equation with Dirichlet boundary conditions: an augmented Lagrangian approach* (English, with English and French summaries), C. R. Math. Acad. Sci. Paris **336** (2003), no. 9, 779–784, DOI 10.1016/S1631-073X(03)00149-3. MR1989280

[46] E. J. Dean and R. Glowinski, *Numerical solution of the two-dimensional elliptic Monge-Ampère equation with Dirichlet boundary conditions: a least-squares approach* (English, with English and French summaries), C. R. Math. Acad. Sci. Paris **339** (2004), no. 12, 887–892, DOI 10.1016/j.crma.2004.09.018. MR2111728

[47] E. J. Dean and R. Glowinski, *An augmented Lagrangian approach to the numerical solution of the Dirichlet problem for the elliptic Monge-Ampère equation in two dimensions*, Electron. Trans. Numer. Anal. **22** (2006), 71–96. MR2208483

[48] E. J. Dean and R. Glowinski, *Numerical methods for fully nonlinear elliptic equations of the Monge-Ampère type*, Comput. Methods Appl. Mech. Engrg. **195** (2006), no. 13-16, 1344–1386, DOI 10.1016/j.cma.2005.05.023. MR2203972

[49] T. Dupont and R. Scott, *Polynomial approximation of functions in Sobolev spaces*, Math. Comp. **34** (1980), no. 150, 441–463, DOI 10.2307/2006095. MR559195

[50] G. Engel, K. Garikipati, T. J. R. Hughes, M. G. Larson, L. Mazzei, and R. L. Taylor, *Continuous/discontinuous finite element approximations of fourth-order elliptic problems in structural and continuum mechanics with applications to thin beams and plates, and strain gradient elasticity*, Comput. Methods Appl. Mech. Engrg. **191** (2002), no. 34, 3669–3750, DOI 10.1016/S0045-7825(02)00286-4. MR1915664

[51] X. Feng, R. Glowinski, and M. Neilan, *Recent developments in numerical methods for fully nonlinear second order partial differential equations*, SIAM Rev. **55** (2013), no. 2, 205–267, DOI 10.1137/110825960. MR3049920

[52] X. Feng and M. Jensen, *Convergent semi-Lagrangian methods for the Monge-Ampère equation on unstructured grids*, SIAM J. Numer. Anal. **55** (2017), no. 2, 691–712, DOI 10.1137/16M1061709. MR3623696

[53] X. Feng and M. Neilan, *Mixed finite element methods for the fully nonlinear Monge-Ampère equation based on the vanishing moment method*, SIAM J. Numer. Anal. **47** (2009), no. 2, 1226–1250, DOI 10.1137/070710378. MR2485451

[54] X. Feng and M. Neilan, *Vanishing moment method and moment solutions for fully nonlinear second order partial differential equations*, J. Sci. Comput. **38** (2009), no. 1, 74–98, DOI 10.1007/s10915-008-9221-9. MR2472219

[55] X. Feng and M. Neilan, *Analysis of Galerkin methods for the fully nonlinear Monge-Ampère equation*, J. Sci. Comput. **47** (2011), no. 3, 303–327, DOI 10.1007/s10915-010-9439-1. MR2793586

[56] A. Figalli, *The Monge-Ampère equation and its applications*, Zurich Lectures in Advanced Mathematics, European Mathematical Society (EMS), Zürich, 2017, DOI 10.4171/170. MR3617963

[57] B. D. Froese and A. M. Oberman, *Convergent finite difference solvers for viscosity solutions of the elliptic Monge-Ampère equation in dimensions two and higher*, SIAM J. Numer. Anal. **49** (2011), no. 4, 1692–1714, DOI 10.1137/100803092. MR2831067

[58] B. D. Froese and A. M. Oberman, *Fast finite difference solvers for singular solutions of the elliptic Monge-Ampère equation*, J. Comput. Phys. **230** (2011), no. 3, 818–834, DOI 10.1016/j.jcp.2010.10.020. MR2745457

[59] B. D. Froese and A. M. Oberman, *Convergent filtered schemes for the Monge-Ampère partial differential equation*, SIAM J. Numer. Anal. **51** (2013), no. 1, 423–444, DOI 10.1137/120875065. MR3033017

[60] D. Gallistl and N. T. Tran, *Convergence of a regularized finite element discretization of the two-dimensional Monge-Ampère equation*, Math. Comp. **92** (2023), no. 342, 1467–1490, DOI 10.1090/mcom/3794. MR4570330

[61] D. Gallistl and N. T. Tran, *Stability and guaranteed error control of approximations to the Monge-Ampère equation*, Numer. Math. **156** (2024), no. 1, 107–131, DOI 10.1007/s00211-023-01385-5. MR4700408

[62] R. Glowinski, H. Liu, S. Leung, and J. Qian, *A finite element/operator-splitting method for the numerical solution of the two dimensional elliptic Monge-Ampère equation*, J. Sci. Comput. **79** (2019), no. 1, 1–47, DOI 10.1007/s10915-018-0839-y. MR3936237

[63] P. Grisvard, *Elliptic problems in nonsmooth domains*, Monographs and Studies in Mathematics, vol. 24, Pitman (Advanced Publishing Program), Boston, MA, 1985. MR775683

[64] C. E. Gutiérrez, *The Monge-Ampère equation*, Progress in Nonlinear Differential Equations and their Applications, vol. 44, Birkhäuser Boston, Inc., Boston, MA, 2001, DOI 10.1007/978-1-4612-0195-3. MR1829162

[65] W. W. Hager and H. Zhang, *A new active set algorithm for box constrained optimization*, SIAM J. Optim. **17** (2006), no. 2, 526–557, DOI 10.1137/050635225. MR2247750

[66] W. W. Hager and H. Zhang, *The limited memory conjugate gradient method*, SIAM J. Optim. **23** (2013), no. 4, 2150–2168, DOI 10.1137/120898097. MR3123830

[67] W. W. Hager and H. Zhang, *An active set algorithm for nonlinear optimization with polyhedral constraints*, Sci. China Math. **59** (2016), no. 8, 1525–1542, DOI 10.1007/s11425-016-0300-6. MR3528501

[68] B. F. Hamfeldt and J. Lesniewski, *Convergent finite difference methods for fully nonlinear elliptic equations in three dimensions*, J. Sci. Comput. **90** (2022), no. 1, Paper No. 35, 36, DOI 10.1007/s10915-021-01714-6. MR4348930

[69] B. F. Hamfeldt and T. Salvador, *Higher-order adaptive finite difference methods for fully nonlinear elliptic equations*, J. Sci. Comput. **75** (2018), no. 3, 1282–1306, DOI 10.1007/s10915-017-0586-5. MR3798102

[70] M. Jensen, *Numerical solution of the simple Monge-Ampère equation with nonconvex Dirichlet data on nonconvex domains*, Hamilton-Jacobi-Bellman equations, Radon Ser. Comput. Appl. Math., vol. 21, De Gruyter, Berlin, 2018, pp. 129–142. MR3823875

[71] H. J. Kuo and N. S. Trudinger, *Discrete methods for fully nonlinear elliptic equations*, SIAM J. Numer. Anal. **29** (1992), no. 1, 123–135, DOI 10.1137/0729008. MR1149088

[72] M.-J. Lai and J. Lee, *Trivariate spline collocation methods for numerical solution to 3D Monge-Ampère equation*, J. Sci. Comput. **95** (2023), no. 2, Paper No. 56, 29, DOI 10.1007/s10915-023-02183-9. MR4572133

[73] O. Lakkis and T. Pryer, *A finite element method for nonlinear elliptic problems*, SIAM J. Sci. Comput. **35** (2013), no. 4, A2025–A2045, DOI 10.1137/120887655. MR3085125

[74] R. Li and F. Yang, *A reconstructed discontinuous approximation to Monge-Ampère equation in least square formulation*, Adv. Appl. Math. Mech. **15** (2023), no. 5, 1109–1141, DOI 10.4208/aamm.oa-2022-0047. MR4613677

[75] W. Li and R. H. Nochetto, *Optimal pointwise error estimates for two-scale methods for the Monge-Ampère equation*, SIAM J. Numer. Anal. **56** (2018), no. 3, 1915–1941, DOI 10.1137/18M1165670. MR3819162

[76] A. Maugeri, D. K. Palagachev, and L. G. Softova, *Elliptic and parabolic equations with discontinuous coefficients*, Mathematical Research, vol. 109, Wiley-VCH Verlag Berlin GmbH, Berlin, 2000, DOI 10.1002/3527600868. MR2260015

[77] C. Miranda, *Su di una particolare equazione ellittica del secondo ordine a coefficienti discontinui* (Italian, with Romanian and Russian summaries), An. şti. Univ. "Al. I. Cuza" Iaşi Secţ. I a Mat. (N.S.) **11B** (1965), 209–215. MR212369

[78] M. Neilan, *A nonconforming Morley finite element method for the fully nonlinear Monge-Ampère equation*, Numer. Math. **115** (2010), no. 3, 371–394, DOI 10.1007/s00211-009-0283-x. MR2640051

[79] M. Neilan, *Quadratic finite element approximations of the Monge-Ampère equation*, J. Sci. Comput. **54** (2013), no. 1, 200–226, DOI 10.1007/s10915-012-9617-4. MR3002620

[80] 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** (2014), 351–369, DOI 10.1016/j.cam.2013.12.027. MR3162358

[81] M. Neilan, *A unified analysis of three finite element methods for the Monge-Ampère equation*, Electron. Trans. Numer. Anal. **41** (2014), 262–288. MR3270089

[82] M. Neilan, A. J. Salgado, and W. Zhang, *Numerical analysis of strongly nonlinear PDEs*, Acta Numer. **26** (2017), 137–303, DOI 10.1017/S0962492917000071. MR3653852

[83] M. Neilan and M. Wu, *Discrete Miranda-Talenti estimates and applications to linear and nonlinear PDEs*, J. Comput. Appl. Math. **356** (2019), 358–376, DOI 10.1016/j.cam.2019.01.032. MR3921147

[84] M. Neilan and W. Zhang, *Rates of convergence in W^2_p-norm for the Monge-Ampère equation*, SIAM J. Numer. Anal. **56** (2018), no. 5, 3099–3120, DOI 10.1137/17M1160409. MR3864689

[85] R. H. Nochetto, D. Ntogkas, and W. Zhang, *Two-scale method for the Monge-Ampère equation: convergence to the viscosity solution*, Math. Comp. **88** (2019), no. 316, 637–664, DOI 10.1090/mcom/3353. MR3882279

[86] R. H. Nochetto, D. Ntogkas, and W. Zhang, *Two-scale method for the Monge-Ampère equation: pointwise error estimates*, IMA J. Numer. Anal. **39** (2019), no. 3, 1085–1109, DOI 10.1093/imanum/dry026. MR3984051

[87] R. H. Nochetto and W. Zhang, *Pointwise rates of convergence for the Oliker-Prussner method for the Monge-Ampère equation*, Numer. Math. **141** (2019), no. 1, 253–288, DOI 10.1007/s00211-018-0988-9. MR3903208

[88] A. M. Oberman, *Wide stencil finite difference schemes for the elliptic Monge-Ampère equation and functions of the eigenvalues of the Hessian*, Discrete Contin. Dyn. Syst. Ser. B **10** (2008), no. 1, 221–238, DOI 10.3934/dcdsb.2008.10.221. MR2399429

[89] V. I. Oliker and L. D. Prussner, *On the numerical solution of the equation $(\partial^2 z/\partial x^2)(\partial^2 z/\partial y^2) - ((\partial^2 z/\partial x\partial y))^2 = f$ and its discretizations. I*, Numer. Math. **54** (1988), no. 3, 271–293, DOI 10.1007/BF01396762. MR971703

[90] W. Qiu and L. Tang, *A note on the Monge-Ampère type equations with general source terms*, Math. Comp. **89** (2020), no. 326, 2675–2706, DOI 10.1090/mcom/3554. MR4136543

[91] I. Smears and E. Süli, *Discontinuous Galerkin finite element approximation of nondivergence form elliptic equations with Cordès coefficients*, SIAM J. Numer. Anal. **51** (2013), no. 4, 2088–2106, DOI 10.1137/120899613. MR3077903

[92] G. Talenti, *Sopra una classe di equazioni ellittiche a coefficienti misurabili* (Italian), Ann. Mat. Pura Appl. (4) **69** (1965), 285–304, DOI 10.1007/BF02414375. MR201816

[93] N. S. Trudinger and X.-J. Wang, *The Monge-Ampère equation and its geometric applications*, Handbook of geometric analysis. No. 1, Adv. Lect. Math. (ALM), vol. 7, Int. Press, Somerville, MA, 2008, pp. 467–524. MR2483373

Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803
Email address: brenner@math.lsu.edu

Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803
Email address: sung@math.lsu.edu

Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana
Current address: School of Mathematical Sciences and Fujian Provincial Key Laboratory on Mathematical Modeling and High Performance Scientific Computing, Xiamen University, Fujian, 361005, China
Email address: ztan@cct.lsu.edu, zhiyutan@xmu.edu.cn

Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803
Email address: hozhang@math.lsu.edu