Journal of Computational Mathematics
Vol.38, No.4, 2020, 547–564.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1901-m2018-0039

A TWO-GRID METHOD FOR THE C^0 INTERIOR PENALTY DISCRETIZATION OF THE MONGE-AMPÈRE EQUATION*

Gerard Awanou
Department of Mathematics, Statistics, and Computer Science, M/C 249.
University of Illinois at Chicago, Chicago, IL 60607, USA
Email: awanou@uic.edu
Hengguang Li
Department of Mathematics, Wayne State University, 656 W. Kirby, Detroit, MI 48202, USA
Email: hli@math.wayne.edu
Eric Malitz
Department of Mathematics, DePaul University, Chicago, IL 60614, USA
Email: EMALITZ@depaul.edu

Abstract

The purpose of this paper is to analyze an efficient method for the solution of the nonlinear system resulting from the discretization of the elliptic Monge-Ampère equation by a C^0 interior penalty method with Lagrange finite elements. We consider the two-grid method for nonlinear equations which consists in solving the discrete nonlinear system on a coarse mesh and using that solution as initial guess for one iteration of Newton's method on a finer mesh. Thus both steps are inexpensive. We give quasi-optimal $W^{1,\infty}$ error estimates for the discretization and estimate the difference between the interior penalty solution and the two-grid numerical solution. Numerical experiments confirm the computational efficiency of the approach compared to Newton's method on the fine mesh.

Mathematics subject classification: 65N30, 65N55, 35J96.
Key words: Two-grid discretization, Interior penalty method, Finite element, Monge-Ampère.

1. Introduction

In this paper, we prove the convergence of a two grid method for solving the nonlinear system resulting from the discretization of the elliptic Monge-Ampère equation

$$\det(D^2 u) = f \quad \text{in } \Omega, \quad u = g \quad \text{on } \partial\Omega, \tag{1.1}$$

with a version of the C^0 interior penalty discretization proposed in [5]. The domain Ω is assumed to be a convex polygonal domain of $\mathbb{R}^2$ and (1.1) is assumed to have a strictly convex smooth solution $u \in C^{k+1}(\overline{\Omega})$ for an integer $k \geq 3$. The function $f \in C^{k-1}(\overline{\Omega})$ is given and satisfies $f \geq c_0$ for a constant $c_0 > 0$ and the function $g \in C(\partial\Omega)$ is also given and assumed to extend to a $C^{k+1}(\overline{\Omega})$ function G. In (1.1), $D^2 u = \big(\partial^2 u/(\partial x_i \partial x_j)\big)_{i,j=1,2}$ is the Hessian matrix of u and det denotes the determinant operator. Let $V_h \subset H^1(\Omega)$ denote the Lagrange finite element space of degree $k \geq 3$. Let Dv denote the gradient of the function v. Recall that

* Received March 26, 2018 / Revised version received October 23, 2018 / Accepted January 23, 2019 / Published online January 17, 2020 /

cof D^2v denotes the matrix of cofactors of D^2v. The C^0 interior penalty discretization can be written in abstract form as: find $u_h \in V_h$ such that $u_h = g_h$ on $\partial\Omega$ and

$$A(u_h, \phi) = 0, \quad \forall \phi \in V_h \cap H_0^1(\Omega).$$

Here g_h denotes the canonical interpolant in V_h of a continuous extension of g and A is defined in (3.1) below. The discretization has the property that if we denote by $A'(u; v, \phi)$ the Fréchet derivative evaluated at u of the mapping $v \to A(v, \phi)$, then

$$A'(u; v, \phi) = \int_\Omega \big((\operatorname{cof} D^2u)Dv\big) \cdot D\phi \, dx, \tag{1.2}$$

which gives the weak form of a standard linear elliptic operator. We exploit this property to give quasi-optimal $W^{1,\infty}$ error estimates, and the convergence of a two-grid numerical scheme for solving the discrete nonlinear system. Numerical experiments confirm the computational efficiency of the two-grid method compared to Newton's method on the fine mesh. Two-grid methods were initially analyzed in [13] for quasi-linear problems, and (1.1) is a fully nonlinear equation. The numerical results in [11] used a two-grid method.

Monge-Ampère type equations with smooth solutions on polygonal domains appear in many problems of practical interest. For example they appear in the study of von Kármán model for plate buckling [6]. In addition, for meteorological applications for which other differential operators are discretized with a finite element method, it would be advantageous to use a finite element discretization for the Monge-Ampère operator as well. It is known that when Ω is strictly convex with a smooth boundary, and with our smoothness assumptions on f and g, (1.1) has a smooth solution. There are several discretizations for smooth solutions of (1.1). Provably convergent schemes for non smooth solutions can be used for smooth solutions as well. However the latter have a low order of approximation for smooth solutions. We refer to [8] for example for a review. Because the interior penalty term involves the cofactor matrix of the Hessian, it is very likely that the method proposed in [5] is suitable only for smooth solutions. It does not seem possible to put it in the framework of approximation by smooth solutions proposed in [2], where the right hand side of (1.1) is viewed as a measure.

There has been no previous study of multilevel methods for finite element discretizations of (1.1). A key tool in the proof of convergence of the two-grid method is a $W^{1,\infty}$ norm error estimate for $k \geq 3$. Such estimates were obtained in [10] for quadratic and higher order elements on a smooth domain. But the proof therein relies on an elliptic regularity property of the linearized problem [10, (2.21)]. Unless the domain is a rectangle, we do not expect such an elliptic regularity property to hold for general polygonal domains considered in this paper.

With the quasi-optimal $W^{1,\infty}$ error estimates we obtain a new proof of the optimal H^1 estimates obtained in [5]. Although these estimates are not new, we include nethertheless the proof since its ideas are also used in the proof of the convergence of the two-grid method. The version of the C^0 interior penalty discretization proposed in [5] we consider, consists in imposing the boundary condition through interpolation, instead of weakly with a penalty term. In this context, as expected, the proof of the H^1 estimates is simpler than the one given in [5]. In particular, no mesh-dependent norms are used.

The two-grid method consists in solving (1.1) on a coarse mesh of size H and using that solution as initial guess for one iteration of Newton's method on a finer mesh of size h with $H = h^\lambda, 0 < \lambda < 1$. Thus both steps are inexpensive. We prove that the convergence rate does not deteriorate provided that $k \geq 3$ and $\lambda > 1/2 + (2+\epsilon)/(2k)$ for $0 < \epsilon < 1$ and h sufficiently small. Thus for $k = 6$ and $h = 1/2^n$ for an integer $n \geq 6$, we may take $H = 1/2^{n-1}$. In the case $k = 3$ the theoretical result dictates for these choices $n \geq 12$. However our numerical experiments indicate that the mesh size h does not need to be that small for computational efficiency. In addition much coarser mesh sizes H can be taken, for example $H = 1/2^{n-2}$. Thus our lower bound for λ is not sharp. In addition numerical results indicate that the two-grid method is effective for $k = 2$. The quadratic case is not covered by our theoretical result.

The paper is organized as follows. In the next section, we introduce additional notation and recall some preliminary results. In Section 3 we give the $W^{1,\infty}$ error estimates for the discretization. In Section 4 we present the two-grid algorithm and its error analysis. In Section 5, we present numerical results which confirm the computational efficiency of the two-grid algorithm.

2. Additional Notation and Preliminaries

Let $\mathcal{T}_h$ denote a conforming, shape regular and quasi-uniform triangulation of Ω into simplices K. We denote by $\mathcal{E}_h$ the set of edges of $\mathcal{T}_h$, by $\mathcal{E}_h^b$ the set of boundary edges and by $\mathcal{E}_h^i$ the set of interior edges.

Let h_K denote the diameter of the element K and put $h = \max_{K \in \mathcal{T}_h} h_K$. We assume that $0 < h \leq 1$. We recall that, for a shape regular and quasi-uniform triangulation, there exists a constant $\sigma > 0$, independent of h such that $h_K/\rho_K \leq \sigma$, for all $K \in \mathcal{T}_h$ where ρ_K denotes the radius of the largest ball inside K and there is another constant C, independent of h, such that $h \leq Ch_K$ for all $K \in \mathcal{T}_h$. Throughout the paper, we will use the letter C for a generic constant, independent of h, which may change from occurrences. Many of the constants depend on the finite element space degree k. Since we have assumed that k is constant, we do not always specifically indicate the dependence.

We use the usual notation $W^{s,p}(\Omega)$, $1 \leq s, p \leq \infty$ for the Sobolev spaces of functions in $L^p(\Omega)$, $1 \leq p \leq \infty$, with weak derivatives up to order s in $L^p(\Omega)$. The standard notation $H^s(\Omega)$ is used for $W^{s,2}(\Omega)$ and $H_0^1(\Omega)$ denotes the subspace of elements in $H^1(\Omega)$ with vanishing trace on the boundary $\partial\Omega$. Similarly, we define $W_0^{1,\infty}(\Omega)$. The norm of $v \in W^{k,p}(\Omega)$ is denoted $||v||_{W^{k,p}(\Omega)}$ and its seminorm $|v|_{W^{k,p}(\Omega)}$. We will omit the argument Ω when it is understood from context.

Denote by $P_k(K)$ the space of polynomials of degree $k \geq 3$ on the element K and let V_h denote the Lagrange finite element space of degree k

$$V_h = \Big\{ v \in C^0(\Omega) : v_{|K} \in P_k(K) \text{ for all } K \in \mathcal{T}_h \Big\}.$$

We will need the broken Sobolev norm defined for $1 \leq p < \infty$ by

$$||v||_{W^{k,p}(\mathcal{T}_h)} = \left(\sum_{K \in \mathcal{T}_h} ||v||^p_{W^{k,p}(K)} \right)^{1/p},$$

and for $p = \infty$ by

$$||v||_{W^{k,\infty}(\mathcal{T}_h)} = \max_{K \in \mathcal{T}_h} ||v||_{W^{k,\infty}(K)}.$$

We recall the inverse estimates [7, Lemma 4.5.3]

$$||v||_{W^{s,p}(\mathcal{T}_h)} \leq Ch^{t-s+\min(0,\frac{2}{p}-\frac{2}{q})}||v||_{W^{t,q}(\mathcal{T}_h)}, \quad \forall v \in V_h, \tag{2.1}$$

valid for $0 \leq t \leq s$ and $1 \leq p, q \leq \infty$. We also recall the trace inequality [7, Theorem 1.6.6], $||v||_{L^p(\partial\Omega)} \leq C||v||_{W^{1,p}(\Omega)}, 1 \leq p \leq \infty$ which gives by a scaling argument

$$||v||_{L^p(\partial K)} \leq Ch^{-\frac{1}{p}}\Big(||v||_{L^p(K)} + h||Dv||_{L^p(K)}\Big). \tag{2.2}$$

For $\phi \in W^{1,1}(\Omega)$, by the trace estimate (2.2), we have

$$\sum_{e\in\mathcal{E}_h^i} ||\phi||_{L^1(e)} \leq Ch^{-1} \sum_{K\in\mathcal{T}_h} ||\phi||_{W^{1,1}(K)} = Ch^{-1}||\phi||_{W^{1,1}}. \tag{2.3}$$

By an inverse estimate one has from (2.2)

$$||v||_{L^2(\partial K)} \leq Ch^{-\frac{1}{2}}||v||_{L^2(K)} \, \forall v \in V_h. \tag{2.4}$$

We will also need the following properties of the Lagrange interpolant operator I_h [7, Corollary 4.4.24]

$$||v - I_h v||_{W^{s,p}(\mathcal{T}_h)} \leq Ch^{k+1-s}||v||_{W^{k+1,p}}, \quad s = 0, 1, 2 \quad \text{and} \quad 1 \leq p \leq \infty, v \in W^{k+1,p}. \tag{2.5}$$

This follows from our assumptions on the triangulation and [7, (4.4.5)], i.e. for $v \in W^{k+1,p}(K)$

$$|v - I_h v|_{W^{s,p}(K)} \leq Ch_K^{k+1-s}|v|_{W^{k+1,p}(K)}, \quad s = 0, 1, 2 \quad \text{and} \quad 1 \leq p \leq \infty. \tag{2.6}$$

It follows from (2.2) and (2.6) that

$$\begin{aligned} &||D(I_h u - u)||_{L^2(\partial K)} \\ \leq &Ch^{-\frac{1}{2}}||D(I_h u - u)||_{L^2(K)} + Ch^{\frac{1}{2}}||D(I_h u - u)||_{H^1(K)} \\ \leq &Ch^{k-\frac{1}{2}}||u||_{H^{k+1}(K)}. \end{aligned} \tag{2.7}$$

For two matrices A and B, $A : B = \sum_{i,j=1}^2 A_{ij}B_{ij}$ denotes their Frobenius inner product. The divergence of a matrix field is understood as the vector obtained by taking the divergence of each row.

The following results can be checked by simple algebraic computations and can also be found in [3]. For v sufficiently smooth we have

$$\det D^2 v = \frac{1}{2}(\operatorname{cof} D^2 v) : D^2 v, \tag{2.8}$$

and if $F(v) = \det D^2 v - f$, the Fréchet derivative of F at v is given by

$$F'(v)w = (\operatorname{cof} D^2 v) : D^2 w, \tag{2.9}$$

for v, w sufficiently smooth. Under the same assumptions

$$\operatorname{div}\big((\operatorname{cof} D^2 w)Dv\big) = (\operatorname{cof} D^2 w) : D^2 v, \tag{2.10}$$

which is a consequence of the product rule and the row divergence-free property of the Hessian, i.e. $\operatorname{div}(\operatorname{cof} D^2 w) = 0$. It then follows that

$$F'(u)v = \operatorname{div}\big((\operatorname{cof} D^2 u) Dv\big). \tag{2.11}$$

Note that

$$\operatorname{cof}(D^2 v + D^2 w) = \operatorname{cof}(D^2 v) + \operatorname{cof}(D^2 w), \tag{2.12}$$

since we restrict our discussion to the two dimensional case. We have [5]

$$\det D^2 v - \det D^2 w = \frac{1}{2}\Big(\operatorname{cof}(D^2 v) + \operatorname{cof}(D^2 w)\Big) : (D^2 v - D^2 w). \tag{2.13}$$

Using (2.10), (2.12) and (2.13), we obtain

$$\det D^2 v - \det D^2 w = \frac{1}{2}\operatorname{div}\Big(\big(\operatorname{cof} D^2(v+w)\big) D(v-w)\Big). \tag{2.14}$$

We recall that u is strictly convex and thus $\operatorname{cof} D^2 u$ is uniformly positive definite; that is, there exists positive constants α_0 and α_1 such that $\forall x \in \mathbb{R}^2$

$$\alpha_0 |r|^2 \leq r^T\Big(\operatorname{cof} D^2 u(x)\Big) r \leq \alpha_1 |r|^2, \quad \forall r \in \mathbb{R}^2. \tag{2.15}$$

For $v \in V_h$, we will make the abuse of notation of denoting by $D^2 v$ the discrete Hessian of v computed element by element.

Next, we recall some algebraic manipulations of discontinuous functions. For $e \subset \partial K$, let n_K denote the outward normal to K and let $v|_K$ denote the restriction of the field v to K. For $e = K^+ \cap K^-$, we define the jump of the vector field v across e as

$$[\![v]\!]_e = v|_{K^+} \cdot n_{K^+} + v|_{K^-} \cdot n_{K^-}, \tag{2.16}$$

and its average on e as

$$\{\{v\}\}_e = \frac{1}{2}(v|_{K^+} + v|_{K^-}). \tag{2.17}$$

The jump and average of a matrix field E on e are defined respectively as

$$[\![E]\!]_e = n_{K^+}^T E|_{K^+} + n_{K^-}^T E|_{K^-}, \tag{2.18}$$

$$\{\{E\}\}_e = \frac{1}{2}(E|_{K^+} + E|_{K^-}). \tag{2.19}$$

For a matrix field E and a vector field v it is not difficult to check that for $e \in \mathcal{E}_h^i$

$$[\![Ev]\!]_e = [\![\{\{E\}\}_e v]\!]_e + [\![E]\!]_e \{\{v\}\}_e. \tag{2.20}$$

We will omit below the subscript e as it will be clear from the context.

Let $\hat{P}_h : H_0^1(\Omega) \to V_h$ denote the projection with respect to the bilinear form $A'(u;\cdot,\cdot)$ given by (1.2) and recall that G denotes a $C^{k+1}(\overline{\Omega})$ extension of g. We define

$$P_h u = \hat{P}_h(u - G) + I_h G,$$

where I_h denotes the canonical Lagrange interpolant operator into V_h. Then $P_h u = I_h u$ on $\partial\Omega$ and

$$A'(u; P_h u - u, \phi) = A'(u; I_h G - G, \phi), \quad \forall \phi \in V_h \cap H_0^1(\Omega). \tag{2.21}$$

Put $w = u - G$. Since $w = 0$ on $\partial\Omega$ and w is smooth, we have [7, Corollary 8.1.12]

$$||w - \hat{P}_h(w)||_{W^{1,\infty}} \leq Ch^k ||w||_{W^{k+1,\infty}}, \quad \text{for } w \in W^{k+1,\infty}(\Omega), w = 0 \text{ on } \partial\Omega.$$

Therefore

$$\begin{aligned} ||u - P_h u||_{W^{1,\infty}} &= ||w + G - \hat{P}_h w - I_h G||_{W^{1,\infty}} \\ &\leq ||w - \hat{P}_h w||_{W^{1,\infty}} + ||G - I_h G||_{W^{1,\infty}}. \end{aligned}$$

It thus follows from the approximation properties of I_h that

$$||u - P_h u||_{W^{1,\infty}} \leq Ch^k ||u||_{W^{k+1,\infty}} \equiv C_1(k) h^k, \quad \text{for } u \in W^{k+1,\infty}(\Omega). \tag{2.22}$$

By an inverse estimate, (2.22), and the approximation properties of I_h, we have

$$\begin{aligned} ||u - P_h u||_{W^{2,\infty}(\mathcal{T}_h)} &\leq ||u - I_h u||_{W^{2,\infty}(\mathcal{T}_h)} + ||I_h u - P_h u||_{W^{2,\infty}(\mathcal{T}_h)} \\ &\leq Ch^{k-1} ||u||_{W^{k+1,\infty}} + Ch^{-1}\Big(||I_h u - u||_{W^{1,\infty}} + ||u - P_h u||_{W^{1,\infty}} \Big), \end{aligned}$$

that is

$$||u - P_h u||_{W^{2,\infty}(\mathcal{T}_h)} \leq Ch^{k-1} ||u||_{W^{k+1,\infty}}. \tag{2.23}$$

Following [13], we will obtain pointwise estimates via the use of discrete Green's functions. For $z \in \Omega \setminus \cup_{K \in \mathcal{T}_h} \partial K$, let $g_{h,i}^z \in V_h \cap H_0^1(\Omega)$, $i = 1, 2$ be defined by:

$$A'(u; g_{h,i}^z, \phi) = \frac{\partial \phi}{\partial x_i}(z), \quad \forall \phi \in V_h \cap H_0^1(\Omega), \tag{2.24}$$

and let G_h^z be defined by

$$A'(u; G_h^z, \phi) = \phi(z), \quad \forall \phi \in V_h \cap H_0^1(\Omega). \tag{2.25}$$

We have for h sufficiently small

$$||g_{h,i}^z||_{W^{1,1}} \leq C |\ln h|, \quad ||G_h^z||_{L^2} \leq C, \tag{2.26}$$

where the constant C is independent of z. In the case cof $D^2 u$ is the identity matrix, the proof is given in [12, Lemmas 2.1 and 3.3]. The proof of the general case is similar [9]. Moreover we have [9], see also [7, Exercise 8.x.19],

$$||G_h^z||_{W^{1,1}} \leq C |\ln h|. \tag{2.27}$$

It is enough to prove that $|G_h^z|_{W^{1,1}} \leq C |\ln h|$ which follows from the bound $|G_h^z|_{W^{1,2}} \leq C |\ln h|^{\frac{1}{2}}$ and Hölder's inequality. By the discrete Sobolev inequality [7, (4.9.2)] and the coercivity of the form $A'(u;.,.)$, we have for h sufficiently small

$$|A'(u;G_h^z,G_h^z)| = |G_h^z(z)| \leq ||G_h^z||_{L^\infty}$$
$$\leq C|\ln h|^{\frac{1}{2}}||G_h^z||_{H^1} \leq C|\ln h|^{\frac{1}{2}}|A'(u;G_h^z,G_h^z)|^{\frac{1}{2}}.$$

It follows that

$$|G_h^z|_{W^{1,2}} \leq C|A'(u;G_h^z,G_h^z)|^{\frac{1}{2}} \leq C|\ln h|^{\frac{1}{2}},$$

giving the claimed bound.

3. $W^{1,\infty}$ Error Estimates for the C^0 Interior Penalty Discretization

We first describe the interior penalty discretization proposed in [5] for polygonal domains and with the boundary condition enforced strongly. For $\phi \in H_0^1(\Omega)$ and $v \in H^3(K)$ for all $K \in \mathcal{T}_h$, we define

$$A(v,\phi) := \sum_{K\in\mathcal{T}_h} \int_K (f - \det D^2 v)\phi\, dx + \sum_{e\in\mathcal{E}_h^i} \int_e [\![\{\{\operatorname{cof} D^2 v\}\} Dv]\!]\phi\, ds. \tag{3.1}$$

Recall that the discrete problem is given by

$$A(u_h,\phi) = 0, \qquad \forall \phi \in V_h \cap H_0^1(\Omega). \tag{3.2}$$

The addition of the penalty terms, the second term on the right of (3.1), to the natural discretization of (1.1) is the reason of the name C^0 Interior penalty discretization for (3.2). The addition of these terms is motivated by the need in the analysis that the Fréchet derivative evaluated at u of the mapping $v \to A(v,\phi)$ is given by (1.2). This is proven in [5, p. 5]. For the convenience of the reader, we give the proof in the next lemma.

Let $R(w;v,\phi)$ denote the remainder of the Taylor expansion at w of $w \mapsto A(w,\phi)$, i.e.

$$A(w+v,\phi) = A(w,\phi) + A'(w;v,\phi) + R(w;v,\phi). \tag{3.3}$$

Lemma 3.1. *For $v,w \in H^3(K)$ for all $K \in \mathcal{T}_h$ and $\phi \in H_0^1(\Omega)$, we have*

$$\begin{aligned} A'(w;v,\phi) = &\sum_{K\in\mathcal{T}_h} \int_K \big((\operatorname{cof} D^2 w) Dv\big)\cdot D\phi\, dx \\ &- \sum_{e\in\mathcal{E}_h^i} \int_e [\![(\operatorname{cof} D^2 w)]\!]\{\{Dv\}\}\phi\, ds + \sum_{e\in\mathcal{E}_h^i} \int_e [\![\{\{\operatorname{cof} D^2 v\}\} Dw]\!]\phi\, ds, \end{aligned} \tag{3.4}$$

$$R(w;v,\phi) = -\sum_{K\in\mathcal{T}_h} \int_K (\det D^2 v)\phi\, dx + \sum_{e\in\mathcal{E}_h^i} \int_e [\![\{\{\operatorname{cof} D^2 v\}\} Dv]\!]\phi\, ds. \tag{3.5}$$

In particular, for $u \in C^3(\Omega)$, (1.2) holds.

Proof. For $w, v \in H^3(K)$ for all $K \in \mathcal{T}_h$ we have

$$A(w+v,\phi) - A(w,\phi) = -\sum_{K\in\mathcal{T}_h}\int_K \big(\det D^2(w+v) - \det D^2 w\big)\phi\,dx + \sum_{e\in\mathcal{E}_h^i}\int_e [\![\{\{\operatorname{cof} D^2(w+v)\}\}D(w+v)]\!]\phi\,ds - \sum_{e\in\mathcal{E}_h^i}\int_e [\![\{\{\operatorname{cof} D^2 w\}\}Dw]\!]\phi\,ds,$$

for all $\phi \in H_0^1(\Omega)$. Since D^2w is a 2×2 matrix, $\det D^2(w+v) = \det D^2 w + \det D^2 v + \operatorname{cof} D^2 w : D^2 v$. Thus

$$A(w+v,\phi) - A(w,\phi) = -\sum_{K\in\mathcal{T}_h}\int_K (\det D^2 v)\phi\,dx - \sum_{K\in\mathcal{T}_h}\int_K (\operatorname{cof} D^2 w : D^2 v)\phi\,dx + \sum_{e\in\mathcal{E}_h^i}\int_e [\![\{\{\operatorname{cof} D^2 w\}\}Dv]\!]\phi\,ds + \sum_{e\in\mathcal{E}_h^i}\int_e [\![\{\{\operatorname{cof} D^2 v\}\}Dw]\!]\phi\,ds + \sum_{e\in\mathcal{E}_h^i}\int_e [\![\{\{\operatorname{cof} D^2 v\}\}Dv]\!]\phi\,ds.$$

By (2.10) $\operatorname{cof} D^2 w : D^2 v = \operatorname{div}\big((\operatorname{cof} D^2 w)Dv\big)$. This implies that

$$A'(w;v,\phi) = -\sum_{K\in\mathcal{T}_h}\int_K \operatorname{div}\big((\operatorname{cof} D^2 w)Dv\big)\phi\,dx + \sum_{e\in\mathcal{E}_h^i}\int_e [\![\{\{\operatorname{cof} D^2 w\}\}Dv]\!]\phi\,ds + \sum_{e\in\mathcal{E}_h^i}\int_e [\![\{\{\operatorname{cof} D^2 v\}\}Dw]\!]\phi\,ds,$$

and $R(w;v,\phi)$ is given by (3.5). By integration by parts and using the fact $\phi = 0$ on $\partial\Omega$,

$$A'(w;v,\phi) = \sum_{K\in\mathcal{T}_h}\int_K \big((\operatorname{cof} D^2 w)Dv\big)\cdot D\phi\,dx - \sum_{e\in\mathcal{E}_h^i}\int_e [\![(\operatorname{cof} D^2 w)Dv]\!]\phi\,dx + \sum_{e\in\mathcal{E}_h^i}\int_e [\![\{\{\operatorname{cof} D^2 w\}\}Dv]\!]\phi\,ds + \sum_{e\in\mathcal{E}_h^i}\int_e [\![\{\{\operatorname{cof} D^2 v\}\}Dw]\!]\phi\,ds.$$

By (2.20), $[\![(\operatorname{cof} D^2 w)Dv]\!] = [\![\{\{\operatorname{cof} D^2 w\}\}Dv]\!] + [\![\operatorname{cof} D^2 w]\!]\{\{Dv\}\}$. It follows that

$$A'(w;v,\phi) = \sum_{K\in\mathcal{T}_h}\int_K \big((\operatorname{cof} D^2 w)Dv\big)\cdot D\phi\,dx - \sum_{e\in\mathcal{E}_h^i}\int_e [\![(\operatorname{cof} D^2 w)]\!]\{\{Dv\}\}\phi\,ds + \sum_{e\in\mathcal{E}_h^i}\int_e [\![\{\{\operatorname{cof} D^2 v\}\}Dw]\!]\phi\,ds.$$

Finally, since by assumption $u \in C^3(\Omega)$, $[\![(\operatorname{cof} D^2 u)]\!] = 0$. In addition, by definition $\{\{\operatorname{cof} D^2 v\}\}$ is continuous and Du is continuous by the assumption on u. Thus $[\![\{\{\operatorname{cof} D^2 v\}\}Du]\!] = 0$ on each interior edge. The statement about $A'(u;v,\phi)$ easily follows. $\square$

Lemma 3.2. *We have for $\phi \in H_0^1(\Omega)$ and $v, w \in H^3(K)$ for all $K \in \mathcal{T}_h$,*

$$\begin{aligned} R(w; v, \phi) =& \frac{1}{2} \sum_{K \in \mathcal{T}_h} \int_K [(\operatorname{cof} D^2 v) Dv] \cdot D\phi \, dx \\ & - \frac{1}{2} \sum_{e \in \mathcal{E}_h^i} \int_e [\![(\operatorname{cof} D^2 v)]\!] \{\{Dv\}\} \phi \, ds + \frac{1}{2} \sum_{e \in \mathcal{E}_h^i} \int_e [\![\{\{\operatorname{cof} D^2 v\}\} Dv]\!] \phi \, ds. \end{aligned} \tag{3.6}$$

Proof. Using (2.8) and integration by parts

$$\begin{aligned} \int_K (\det D^2 v) \phi \, dx &= \frac{1}{2} \int_K (\operatorname{cof} D^2 v) : (D^2 v) \, \phi \, dx = \frac{1}{2} \int_K \operatorname{div} \big((\operatorname{cof} D^2 v) Dv \big) \phi \, dx \\ &= -\frac{1}{2} \int_K [(\operatorname{cof} D^2 v) Dv] \cdot D\phi \, dx + \frac{1}{2} \int_{\partial K} \big[(\operatorname{cof} D^2 v) Dv \big] \cdot n_K \phi \, ds. \end{aligned}$$

By definition of jump and since $\phi = 0$ on $\partial\Omega$, we have

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} \big((\operatorname{cof} D^2 v) Dv \big) \cdot n_K \phi \, ds = \sum_{e \in \mathcal{E}_h^i} \int_e [\![(\operatorname{cof} D^2 v) Dv]\!] \phi \, ds. \tag{3.7}$$

We conclude that

$$\begin{aligned} R(w; v, \phi) = \frac{1}{2} \sum_{K \in \mathcal{T}_h} \int_K [(\operatorname{cof} D^2 v) Dv] \cdot D\phi \, dx - \frac{1}{2} \sum_{e \in \mathcal{E}_h^i} \int_e [\![(\operatorname{cof} D^2 v) Dv]\!] \phi \, ds \\ + \sum_{e \in \mathcal{E}_h^i} \int_e [\![\{\{\operatorname{cof} D^2 v\}\} Dv]\!] \phi \, ds. \end{aligned}$$

Therefore, using (2.20) to expand the term $[\![(\operatorname{cof} D^2 v) Dv]\!]$ we obtain (3.6). □

We define a nonlinear operator $\Phi : V_h \to V_h$ by $v_h = \Phi(v_h)$ on $\partial\Omega$ and

$$A'(u; v_h - \Phi(v_h), \phi) = A(v_h, \phi), \qquad \forall \phi \in V_h \cap H_0^1(\Omega). \tag{3.8}$$

A fixed point of Φ is a solution of the nonlinear finite element problem (3.2). We note that by (1.2), (2.5), (2.26) and (2.27)

$$\begin{aligned} |A'(u; I_h G - G, G_h^z)| &\le C_2(k) h^k |\ln h|, \\ |A'(u; I_h G - G, g_{h,i}^z)| &\le C_3(k) h^k |\ln h|. \end{aligned}$$

We then define $C_4 = \max\{C_1, C_2, C_3\}$ where the constant C_1 is defined in (2.22). Consider the closed set

$$B_h = \Big\{ v \in V_h : v = g_h \text{ on } \partial\Omega, ||v - u||_{W^{1,\infty}} \le 3 C_4 h^k |\ln h| \Big\}. \tag{3.9}$$

By (2.22) $P_h u \in B_h$ and hence B_h is non-empty.

Lemma 3.3. *We have $\Phi(B_h) \subset B_h$ for h sufficiently small and $k \ge 3$.*

Proof. For $v_h \in B_h$, we have using (3.8) and (2.21)

$$\begin{aligned} A'(u;\Phi(v_h)-P_hu,\phi) &= A'(u;\Phi(v_h)-v_h,\phi)+A'(u;v_h-P_hu,\phi) \\ &= -A(v_h,\phi)+A'(u;v_h-u,\phi)+A'(u;u-P_hu,\phi) \\ &= -A(v_h,\phi)+A'(u;v_h-u,\phi)-A'(u;I_hG-G,\phi). \end{aligned}$$

By definition of the residual (3.3), and since $A(u,\phi)=0$, we have

$$\begin{aligned} &-A(v_h,\phi)+A'(u;v_h-u,\phi) \\ =&A(u,\phi)-A(v_h,\phi)+A'(u;v_h-u,\phi)=-R(u;v_h-u,\phi). \end{aligned}$$

We conclude that

$$A'(u;\Phi(v_h)-P_hu,\phi)=-R(u;v_h-u,\phi)-A'(u;I_hG-G,\phi). \tag{3.10}$$

Therefore, using (3.6), (2.3) and (1.2), we have

$$\begin{aligned} &|A'(u;\Phi(v_h)-P_hu,\phi)| \\ \le& C||v_h-u||_{W^{2,\infty}(\mathcal{T}_h)}||v_h-u||_{W^{1,\infty}}||\phi||_{W^{1,1}} \\ &+C||v_h-u||_{W^{2,\infty}(\mathcal{T}_h)}||v_h-u||_{W^{1,\infty}}\sum_{e\in\mathcal{E}_h^i}||\phi||_{L^1(e)}+|A'(u;I_hG-G,\phi)| \\ \le& Ch^{-1}||v_h-u||_{W^{2,\infty}(\mathcal{T}_h)}||v_h-u||_{W^{1,\infty}}||\phi||_{W^{1,1}}+|A'(u;I_hG-G,\phi)|. \end{aligned}$$

By definition of B_h, $||v_h-u||_{W^{1,\infty}}\le Ch^k|\ln h|$. Moreover, by triangle inequality, (2.23) and an inverse estimate $||v_h-u||_{W^{2,\infty}(\mathcal{T}_h)}\le||v_h-P_hu||_{W^{2,\infty}(\mathcal{T}_h)}+||P_hu-u||_{W^{2,\infty}(\mathcal{T}_h)}\le Ch^{k-1}|\ln h|+Ch^{k-1}\le Ch^{k-1}|\ln h|$. Thus

$$\begin{aligned} |A'(u;\Phi(v_h)-P_hu,\phi)| \le& Ch^{k-2}|\ln h|\,||\phi||_{W^{1,1}}h^k|\ln h|+|A'(u;I_hG-G,\phi)| \\ \le&(Ch^{k-2}|\ln h|^2)h^k||\phi||_{W^{1,1}}+|A'(u;I_hG-G,\phi)|. \end{aligned}$$

Taking $\phi=g^z_{h,i}$ with the estimate (2.26), and taking $\phi=G^z_h$ with the estimate (2.27), we obtain using the definition of C_4

$$||\Phi(v_h)-P_hu||_{W^{1,\infty}}\le(Ch^{k-2}|\ln h|^2+C_4)h^k|\ln h|.$$

Since $Ch^{k-2}|\ln h|^2\le C_4$ for h sufficiently small and $k\ge3$, we get $||\Phi(v_h)-P_hu||_{W^{1,\infty}}\le 2C_4h^k|\ln h|$. By triangular inequality and (2.22), the result follows. □

We will use below a certain algebraic manipulation which is encoded in the following lemma.

Lemma 3.4. *Let L_1 and L_2 be linear functionals and let L denote their product, i.e. $L(v)=L_1(v)L_2(v)$. We have*

$$L(w-u)-L(v-u)=L_1(w-v)L_2(w-u)+L_1(v-u)L_2(w-v).$$

Proof. We have using the linearity of L_1 and L_2

$$\begin{aligned} L(w-u)-L(v-u) &= L_1(w-u)L_2(w-u)-L_1(v-u)L_2(v-u) \\ &= [L_1(w-v)+L_1(v-u)]L_2(w-u)-L_1(v-u)L_2(v-u) \\ &= L_1(w-v)L_2(w-u)+L_1(v-u)[L_2(w-u)-L_2(v-u)], \end{aligned}$$

from which the desired result follows. □

Lemma 3.5. *The mapping Φ is a strict contraction in B_h for h sufficiently small and $k \geq 3$.*

Proof. For v_h and w_h in B_h, we have

$$\begin{aligned} &A'(u;\Phi(v_h)-\Phi(w_h),\phi) \\ =&A'(u;\Phi(v_h)-v_h,\phi)+A'(u;v_h-w_h,\phi)+A'(u;w_h-\Phi(w_h),\phi) \\ =&A(w_h,\phi)-A(v_h,\phi)+A'(u;v_h-w_h,\phi) \\ =&A(w_h,\phi)-A(v_h,\phi)+A'(u;v_h-u,\phi)+A'(u;u-w_h,\phi). \end{aligned}$$

Since $A(u,\phi)=0$, by definition of the residual (3.3), we have

$$A'(u;\Phi(v_h)-\Phi(w_h),\phi)=R(u;w_h-u,\phi)-R(u;v_h-u,\phi).$$

Using algebraic manipulations of the type identified in Lemma 3.4 and (3.6), we obtain

$$\begin{aligned} &A'(u;\Phi(v_h)-\Phi(w_h),\phi) \qquad (3.11)\\ =&\frac{1}{2}\sum_{K\in\mathcal{T}_h}\int_K [(\operatorname{cof} D^2(w_h-v_h))D(w_h-u)]\cdot D\phi\,dx \\ &+\frac{1}{2}\sum_{K\in\mathcal{T}_h}\int_K [(\operatorname{cof} D^2(v_h-u))D(w_h-v_h)]\cdot D\phi\,dx \\ &-\frac{1}{2}\sum_{e\in\mathcal{E}_h^i}\int_e \Big([\![(\operatorname{cof} D^2(w_h-v_h))]\!]\{\{D(w_h-u)\}\} + [\![(\operatorname{cof} D^2(v_h-u))]\!]\{\{D(w_h-v_h)\}\}\Big)\phi\,ds \\ &+\frac{1}{2}\sum_{e\in\mathcal{E}_h^i}\int_e \Big([\![\{\{\operatorname{cof} D^2(w_h-v_h)\}\}D(w_h-u)]\!] + [\![\{\{\operatorname{cof} D^2(v_h-u)\}\}D(w_h-v_h)]\!]\Big)\phi\,ds. \end{aligned}$$

Arguing as in the proof of Lemma 3.3 and using (2.3), we obtain

$$\begin{aligned} |A'(u;\Phi(v_h)-\Phi(w_h),\phi)| \leq& C\Big(||w_h-v_h||_{W^{2,\infty}(\mathcal{T}_h)}||w_h-u||_{W^{1,\infty}} \\ &+||v_h-u||_{W^{2,\infty}(\mathcal{T}_h)}||v_h-w_h||_{W^{1,\infty}}\Big)||\phi||_{W^{1,1}}. \end{aligned}$$

As in the proof of Lemma 3.3, we have $||v_h-u||_{W^{2,\infty}(\mathcal{T}_h)} \leq Ch^{k-1}|\ln h|$ and recall that $||w_h-u||_{W^{1,\infty}} \leq Ch^k|\ln h|$ by definition of B_h. Moreover, by an inverse estimate $||w_h-v_h||_{W^{2,\infty}(\mathcal{T}_h)} \leq Ch^{-1}||w_h-v_h||_{W^{1,\infty}}$. We conclude that

$$|A'(u;\Phi(v_h)-\Phi(w_h),\phi)| \leq C\Big(h^{k-1}|\ln h|+h^{k-1}|\ln h|\Big)||v_h-w_h||_{W^{1,\infty}}||\phi||_{W^{1,1}}.$$

Taking $\phi = g_{h,i}^z$ with the estimate (2.26), and taking $\phi = G_h^z$ with the estimate (2.27), we obtain

$$||\Phi(v_h) - \Phi(w_h)||_{W^{1,\infty}} \leq C(h^{k-1} + h^{k-1})|\ln h|^2 ||v_h - w_h||_{W^{1,\infty}},$$

that is, for $k \geq 3$ and h sufficiently small, we have $||\Phi(v_h) - \Phi(w_h)||_{W^{1,\infty}} \leq \frac{1}{2}||v_h - w_h||_{W^{1,\infty}}$. This completes the proof of the lemma. □

The following theorem follows from Lemmas 3.5 and 3.3 and the Banach fixed point theorem.

Theorem 3.1. *Problem* (3.2) *has a unique solution* u_h *in* B_h *for* $k \geq 3$, h *sufficiently small and*

$$||u - u_h||_{W^{1,\infty}} \leq Ch^k |\ln h|.$$

We note that in the case of a homogeneous boundary condition, $G = 0$ and the right hand side of (2.21) vanishes. In that case, the right hand side of (3.10) simplifies and the rate of convergence in the $W^{1,\infty}$ norm can be shown to be optimal. In other words, Theorem 3.1 can be improved with suitable estimates of the Ritz projection with a non homogeneous boundary condition. The following optimal error estimate in the H^1 norm is derived from Theorem 3.1. A different proof was given in [5].

Theorem 3.2. *Problem* (3.2) *has a unique solution* u_h *in* B_h *for* $k \geq 3$, h *sufficiently small and*

$$||u - u_h||_{H^1} \leq Ch^k.$$

Proof. The proof is based on the expression (3.11) of $A'(u; \Phi(v_h) - \Phi(w_h), \phi)$ derived in the proof of Lemma 3.5 and the expression (3.10) of $A'(u; \Phi(v_h) - P_h u, \phi)$ derived in the proof of Lemma 3.3. Since $\Phi(u_h) = u_h$, we have

$$\begin{aligned} A'(u; u_h - P_h u, \phi) &= A'(u; \Phi(u_h) - P_h u, \phi) \\ &= A'(u; \Phi(u_h) - \Phi(P_h u), \phi) + A'(u; \Phi(P_h u) - P_h u, \phi). \qquad (3.12) \end{aligned}$$

In view of (3.11), we obtain

$$\begin{aligned} &|A'(u; \Phi(u_h) - \Phi(P_h u), \phi)| \\ \leq& C||u_h - P_h u||_{W^{2,\infty}(\mathcal{T}_h)} ||P_h u - u||_{H^1} ||\phi||_{H^1} + C||u_h - u||_{W^{2,\infty}(\mathcal{T}_h)} ||u_h - P_h u||_{H^1} ||\phi||_{H^1} \\ &+ C||u_h - P_h u||_{W^{2,\infty}(\mathcal{T}_h)} \sum_{K \in \mathcal{T}_h} |P_h u - u|_{H^1(\partial K)} ||\phi||_{L^2(\partial K)} \\ &+ C||u_h - u||_{W^{2,\infty}(\mathcal{T}_h)} \sum_{K \in \mathcal{T}_h} ||P_h u - u_h||_{H^1(\partial K)} ||\phi||_{L^2(\partial K)}. \qquad (3.13) \end{aligned}$$

By an inverse estimate, Theorem 3.1, triangle inequality and (2.22), we have $||u_h - P_h u||_{W^{2,\infty}(\mathcal{T}_h)} \leq Ch^{k-1}|\ln h|$. Similarly, using (2.23), we have $||u_h - u||_{W^{2,\infty}(\mathcal{T}_h)} \leq Ch^{k-1}|\ln h|$. Next, by the scaled trace inverse inequality (2.4) and Cauchy-Schwarz inequality

$$\begin{aligned} &\sum_{K \in \mathcal{T}_h} ||P_h u - u_h||_{H^1(\partial K)} ||\phi||_{L^2(\partial K)} \\ \leq& Ch^{-1} \sum_{K \in \mathcal{T}_h} ||P_h u - u_h||_{H^1(K)} ||\phi||_{H^1(K)} \leq Ch^{-1} ||P_h u - u_h||_{H^1} ||\phi||_{H^1}. \qquad (3.14) \end{aligned}$$

By (2.7), an inverse estimate and (2.6)

$$\begin{aligned}|P_h u - u|_{H^1(\partial K)} &\leq |P_h u - I_h u|_{H^1(\partial K)} + |I_h u - u|_{H^1(\partial K)}\\ &\leq Ch^{-\frac{1}{2}}|P_h u - I_h u|_{H^1(K)} + Ch^{k-\frac{1}{2}}||u||_{H^{k+1}(K)}\\ &\leq Ch^{-\frac{1}{2}}|P_h u - u|_{H^1(K)} + Ch^{-\frac{1}{2}}|u - I_h u|_{H^1(K)} + Ch^{k-\frac{1}{2}}||u||_{H^{k+1}(K)}\\ &\leq Ch^{-\frac{1}{2}}|P_h u - u|_{H^1(K)} + Ch^{k-\frac{1}{2}}||u||_{H^{k+1}(K)}.\end{aligned}$$

We have $||P_h u - u||_{H^1} \leq Ch^k||u||_{H^{k+1}}$. This follows from [7, Theorem 5.4.4] in the case of homogeneous boundary conditions. The proof of the general case is similar to (2.22). Thus

$$\begin{aligned}&\Big(\sum_{K\in\mathcal{T}_h} |P_h u - u|^2_{H^1(\partial K)}\Big)^{\frac{1}{2}}\\ \leq &Ch^{-\frac{1}{2}}|P_h u - u|_{H^1} + Ch^{k-\frac{1}{2}}||u||_{H^{k+1}} \leq Ch^{k-\frac{1}{2}}||u||_{H^{k+1}}.\end{aligned}$$

As with (3.14), by (2.4) and Cauchy-Schwarz inequality, it follows that

$$\begin{aligned}&\sum_{K\in\mathcal{T}_h} |P_h u - u|_{H^1(\partial K)}||\phi||_{L^2(\partial K)}\\ \leq &Ch^{-\frac{1}{2}}\sum_{K\in\mathcal{T}_h} |P_h u - u|_{H^1(\partial K)}||\phi||_{H^1(K)} \leq Ch^{k-1}||u||_{H^{k+1}}||\phi||_{H^1}. \end{aligned}\tag{3.15}$$

We conclude from (3.13)–(3.15) that

$$\begin{aligned}&|A'(u;\Phi(u_h) - \Phi(P_h u), \phi)|\\ \leq &Ch^{k-1}|\ln h|\,||P_h u - u||_{H^1}||\phi||_{H^1} + Ch^{k-1}|\ln h|\,||u_h - P_h u||_{H^1}||\phi||_{H^1}\\ &\quad + Ch^{k-1}||u_h - P_h u||_{W^{2,\infty}}||\phi||_{H^1} + Ch^{k-2}|\ln h|\,||u_h - P_h u||_{H^1}||\phi||_{H^1}.\end{aligned}$$

Therefore since by Theorem 3.1 we have the suboptimal estimate $||u_h - P_h u||_{H^1} \leq C||u_h - P_h u||_{W^{1,\infty}} \leq Ch^k|\ln h|$

$$\begin{aligned}&|A'(u;\Phi(u_h) - \Phi(P_h u), \phi)| \leq Ch^{k-1}|\ln h|h^k||\phi||_{H^1} + Ch^{k-1}|\ln h|^2h^k||\phi||_{H^1}\\ &\quad +Ch^{k-2}|\ln h|h^k||\phi||_{H^1} + Ch^{k-2}|\ln h|^2h^k||\phi||_{H^1}\\ &\leq Ch^{k-2}|\ln h|^2h^k||\phi||_{H^1} \leq Ch^k||\phi||_{H^1},\end{aligned}\tag{3.16}$$

for $k \geq 3$. By (3.10), we have

$$A'(u;\Phi(P_h u) - P_h u, \phi) = -R(u; P_h u - u, \phi) + A'(u; I_h G - G, \phi).$$

Using (3.6) and inverse estimates as for (3.15), we obtain

$$\begin{aligned}&|A'(u;\Phi(P_h u) - P_h u, \phi)|\\ &\leq C||P_h u - u||_{W^{2,\infty}(\mathcal{T}_h)}||P_h u - u||_{H^1}||\phi||_{H^1}\\ &\quad +Ch^{-1}||P_h u - u||_{W^{2,\infty}(\mathcal{T}_h)}||P_h u - u||_{H^1}||\phi||_{H^1} + C||I_h G - G||_{H^1}||\phi||_{H^1},\end{aligned}$$

i.e.

$$\begin{aligned} &|A'(u;\Phi(P_hu)-P_hu,\phi)| \\ \leq& C(h^{k-1}|\ln h|h^k+h^{k-2}|\ln h|h^k+h^k)||\phi||_{H^1}\leq Ch^k||\phi||_{H^1}, \end{aligned} \tag{3.17}$$

for $k\geq 3$. Taking $\phi=u_h-P_hu$ in (3.12) and using (3.16) and (3.17), we get from Poincaré's inequality and $k\geq 3$, for h sufficiently small

$$||u_h-P_hu||_{H^1}\leq Ch^k.$$

This completes the proof by a triangle inequality. □

4. Analysis of the Two-Grid Algorithm

The two-grid discretization for solving nonlinear problems is a well established technique. The nonlinear problem (3.2) is first solved on a coarse mesh of size H. The solution u_H is used as an initial guess for one step of Newton's method on the finer mesh of size h. Both steps are inexpensive and the method is more efficient than solving the problem through multiple iterations of Newton's method directly on the fine mesh.

Since u is smooth and strictly convex, the smallest eigenvalue of D^2u is uniformly bounded from below. Thus by the continuity of the eigenvalues of the Hessian as a function of its entries and by approximation, D^2u_H is uniformly positive definite on each element for H sufficiently small. A detailed argument was given in [1, Lemma 4] in the context of C^1 approximations. We consider the version of [13, Algorithm 5.5].

Two-grid algorithm

1. Find $u_H\in V_H$, $u_H=g_H$ on $\partial\Omega$, and $A(u_H,\chi)=0,\ \forall\ \chi\in V_H\cap H_0^1(\Omega)$.
2. Find $u^h\in V_h$, $u^h=g_h$ on $\partial\Omega$, and $A'(u_H;u^h-u_H,\phi)=-A(u_H,\phi),\quad \forall\ \phi\in V_h\cap H_0^1(\Omega)$.

Our goal is to show that the two-grid method is optimal in the sense that $||u-u^h||_{H^1}\leq Ch^k$.

Theorem 4.1. *We have the estimate*

$$||u^h-u_h||_{H^1}\leq Ch^k, \tag{4.1}$$

for $k\geq 3, H=h^\lambda, 1>\lambda>1/2+(2+\epsilon)/(2k), 0<\epsilon<1$ *and* h *sufficiently small.*

Proof. By definition of the two-grid algorithm, the definition of the residual (3.3), and $A(u_h,\phi)=0$ for $\phi\in V_h\cap W_0^{1,\infty}(\Omega)$, we have

$$\begin{aligned} A'(u_H;u_h-u^h,\phi)&=A'(u_H;u_h-u_H,\phi)+A'(u_H;u_H-u^h,\phi)\\ &=A'(u_H;u_h-u_H,\phi)+A(u_H,\phi)\\ &=A(u_h,\phi)-R(u_H;u_h-u_H,\phi)=-R(u_H;u_h-u_H,\phi). \end{aligned}$$

It follows that

$$A'(u;u_h-u^h,\phi)=A'(u-u_H;u_h-u^h,\phi)+A'(u_H;u_h-u^h,\phi)$$
$$=A'(u-u_H;u_h-u^h,\phi)-R(u_H;u_h-u_H,\phi).$$

With arguments similar to the ones used in the proof of Theorem 3.2, we have

$$\begin{aligned}&|A'(u;u_h-u^h,\phi)| \\ \leq &C||u-u_H||_{W^{2,\infty}(\mathcal{T}_h)}||u_h-u^h||_{H^1}||\phi||_{H^1}+Ch^{-1}||u-u_H||_{W^{2,\infty}(\mathcal{T}_h)}||u_h-u^h||_{H^1}||\phi||_{H^1} \\ &+C||\phi||_{L^\infty}\sum_{K\in\mathcal{T}_h}|u-u_H|_{H^1(\partial K)}||u_h-u^h||_{H^2(\partial K)}+|R(u_H;u_h-u_H,\phi)|.\end{aligned} \tag{4.2}$$

But, using (2.7), (2.6) and an inverse estimate

$$\begin{aligned}&|u-u_H|_{H^1(\partial K)}\leq|u-I_Hu|_{H^1(\partial K)}+|I_Hu-u_H|_{H^1(\partial K)} \\ \leq&CH^{k-\frac{1}{2}}||u||_{H^{k+1}(K)}+CH^{-\frac{1}{2}}|I_Hu-u_H|_{H^1(K)} \\ \leq&CH^{k-\frac{1}{2}}||u||_{H^{k+1}(K)}+CH^{-\frac{1}{2}}|I_Hu-u|_{H^1(K)}+CH^{-\frac{1}{2}}|u-u_H|_{H^1(K)}.\end{aligned}$$

It follows that

$$\begin{aligned}&\Big(\sum_{K\in\mathcal{T}_h}|u-u_H|^2_{H^1(\partial K)}\Big)^{\frac{1}{2}} \\ \leq&CH^{k-\frac{1}{2}}||u||_{H^{k+1}}+CH^{-\frac{1}{2}}|I_Hu-u|_{H^1}+CH^{-\frac{1}{2}}|u-u_H|_{H^1}.\end{aligned}$$

We therefore obtain from (2.5) and Theorem 3.2

$$\Big(\sum_{K\in\mathcal{T}_h}|u-u_H|^2_{H^1(\partial K)}\Big)^{\frac{1}{2}}\leq CH^{k-\frac{1}{2}}||u||_{H^{k+1}}.$$

By Cauchy-Schwarz's inequality, (2.4) and an inverse estimate, it follows that

$$\sum_{K\in\mathcal{T}_h}|u-u_H|_{H^1(\partial K)}||u_h-u^h||_{H^2(\partial K)}\leq Ch^{-\frac{3}{2}}H^{k-\frac{1}{2}}||u||_{H^{k+1}}||u_h-u^h||_{H^1}.$$

Therefore, since $||u-u_H||_{W^{2,\infty}(\mathcal{T}_h)}\leq CH^{k-1}|\ln H|$ and by the discrete Sobolev inequality, c.f. [4], $||\phi||_{L^\infty}\leq C(1+|\ln h|^{1/2})||\phi||_{H^1}$, we obtain from (4.2)

$$\begin{aligned}&A'(u;u_h-u^h,\phi)|\leq Ch^{-1}||u-u_H||_{W^{2,\infty}(\mathcal{T}_h)}||u_h-u^h||_{H^1}||\phi||_{H^1} \\ &\quad+CH^{k-\frac{1}{2}}h^{-\frac{3}{2}}|\ln h|\,||u_h-u^h||_{H^1}||u||_{H^{k+1}}||\phi||_{H^1}+|R(u_H;u_h-u_H,\phi)| \\ \leq&Ch^{-1}H^{k-1}|\ln H|\,||u_h-u^h||_{H^1}||\phi||_{H^1}+CH^{k-\frac{1}{2}}h^{-\frac{3}{2}}|\ln h|\,||u_h-u^h||_{H^1}||\phi||_{H^1} \\ &\quad+|R(u_H;u_h-u_H,\phi)|.\end{aligned} \tag{4.3}$$

Using (3.6) and the trace estimates (2.4) and (2.2), we have

$$\begin{aligned}&|R(u_H; u_h - u_H, \phi)|\\ \le& C||u_h - u_H||_{W^{2,\infty}(\mathcal{T}_h)}||u_h - u_H||_{H^1}||\phi||_{H^1} + Ch^{-1}||u_h - u_H||_{W^{2,\infty}(\mathcal{T}_h)}||u_h - u_H||_{H^1}||\phi||_{H^1}\\ \le& Ch^{-2}||u_h - u_H||_{W^{1,\infty}}||u_h - u_H||_{H^1}||\phi||_{H^1}. \end{aligned} \tag{4.4}$$

Taking $\phi = u_h - u^h$ in (4.3) and using (4.4), we get from Poincaré's inequality, Theorems 3.1 and 3.2 for $k \ge 3$ and h sufficiently small

$$\begin{aligned}||u_h - u^h||_{H^1} \le& C(h^{-1}H^{k-1}|\ln H| + H^{k-\frac{1}{2}}h^{-\frac{3}{2}}|\ln h|)||u_h - u^h||_{H^1}\\ &+ Ch^{-2}(h^k|\ln h| + H^k|\ln H|)(h^k + H^k).\end{aligned}$$

We conclude that for $H = h^\lambda$, $\lambda > \max\{1/(k-1), 3/(2k-1)\} = 3/(2k-1)$ and $k \ge 3$,

$$||u_h - u^h||_{H^1} \le Ch^{-2}|\ln H|H^{2k}.$$

We therefore get $||u^h - u_h||_{H^1} \le Ch^k$ provided $\lambda > 3/(2k-1)$ and $2\lambda k - 2 - \epsilon > k$ for some $\epsilon \in (0,1)$, that is $\lambda > \max\{(k+2+\epsilon)/(2k), 3/(2k-1)\} = 1/2 + (2+\epsilon)/(2k)$ for $k \ge 3$. □

5. Numerical Experiments

The computational domain is taken to be the unit square $[0,1]^2$. A uniform grid is obtained by dividing the domain into smaller equal size squares, then dividing each square into two triangles by taking the diagonal with positive slope. We consider a smooth convex test function $u(x,y) = e^{(x^2+y^2)/2}$ so that $f(x,y) = (1+x^2+y^2)e^{(x^2+y^2)}$ and $g(x,y) = e^{(x^2+y^2)/2}$ on $\partial\Omega$. While our convergence analysis is only for cubic and higher order elements and for fine meshes h, we believe the results should be true for quadratic elements, reasonable values of h and allow for much coarse meshes H. Thus, we provide numerical results for both P_2 and P_3 finite elements. On the coarse grid of size H, we first seek an initial guess u_H^0 of u_H as the standard finite element approximation of the solution u_0 of

$$\Delta u_0 = 2\sqrt{f}, \quad u_0 = g \text{ on } \partial\Omega.$$

For solving the coarse grid problem, we perform Newton's method on the coarse grid, setting the maximum iterations to 10 and we impose that the algorithm terminates when $||u_H||_{L^\infty}/||u_H^0||_{L^\infty} \le 10^{-6}$. We report computation times (in seconds) for the two-grid method and Newton's method on the fine grid, as well as H^1 errors and associated rate of convergence. See Table 5.1 for $\lambda = 1 + 2\ln 2/(\ln h) = 1 - 2/n, h = 1/2^n, n = 2, 3\ldots$, and Table 5.2 for $\lambda = 1 + \ln 2/(\ln h) = 1 - 1/n, h = 1/2^n, n = 2, 3\ldots$ for P_2 elements. See Table 5.3 and Table 5.4 for the analogous results using P_3 elements. Note that λ in these numerical results is allowed to be much lower than what our theory predicts.

We also attempted several multigrid experiments, where we interpolate between a series of meshes before ending on the fine grid. However these gave results comparable to the two-grid algorithm. At the cost of extra computation time, there is a slight increase in accuracy if a second iteration is performed on the fine grid. We do not report these results since a second

Table 5.1: P_2; $\lambda = 1 + \frac{2\ln 2}{\ln h}$.

H	h	$\|\|u-u_h\|\|_{H^1}$	rate	$\|\|u-u^h\|\|_{H^1}$	rate	two-grid time	Newton time
$1/2^0$	$1/2^2$	$2.19\ 10^{-2}$	-	$3.12\ 10^{-1}$	-	$2.80\ 10^{-2}$	$4.90\ 10^{-2}$
$1/2^1$	$1/2^3$	$5.55\ 10^{-3}$	1.98	$3.22\ 10^{-2}$	3.28	$7.10\ 10^{-2}$	$1.67\ 10^{-1}$
$1/2^2$	$1/2^4$	$1.39\ 10^{-3}$	2.00	$3.69\ 10^{-3}$	3.12	$2.37\ 10^{-1}$	$8.58\ 10^{-1}$
$1/2^3$	$1/2^5$	$3.48\ 10^{-4}$	2.00	$6.34\ 10^{-4}$	2.54	$9.92\ 10^{-1}$	$3.25\ 10^{0}$
$1/2^4$	$1/2^6$	$8.70\ 10^{-5}$	2.00	$1.48\ 10^{-4}$	2.10	$4.15\ 10^{0}$	$1.31\ 10^{1}$
$1/2^5$	$1/2^7$	$2.18\ 10^{-5}$	2.00	$3.65\ 10^{-5}$	2.02	$1.87\ 10^{1}$	$5.59\ 10^{1}$
$1/2^6$	$1/2^8$	$5.44\ 10^{-6}$	2.00	$9.11\ 10^{-6}$	2.00	$8.57\ 10^{1}$	$2.73\ 10^{2}$

Table 5.2: P_2; $\lambda = 1 + \frac{\ln 2}{\ln h}$.

H	h	$\|\|u-u_h\|\|_{H^1}$	rate	$\|\|u-u^h\|\|_{H^1}$	rate	two-grid time	Newton time
$1/2^1$	$1/2^2$	$2.19\ 10^{-2}$	-	$2.72\ 10^{-2}$	-	$3.10\ 10^{-2}$	$4.90\ 10^{-2}$
$1/2^2$	$1/2^3$	$5.55\ 10^{-3}$	1.98	$5.97\ 10^{-3}$	2.19	$7.80\ 10^{-2}$	$1.67\ 10^{-1}$
$1/2^3$	$1/2^4$	$1.39\ 10^{-3}$	2.00	$1.43\ 10^{-3}$	2.06	$3.44\ 10^{-1}$	$8.58\ 10^{-1}$
$1/2^4$	$1/2^5$	$3.48\ 10^{-4}$	2.00	$3.54\ 10^{-4}$	2.01	$1.69\ 10^{0}$	$3.25\ 10^{0}$
$1/2^5$	$1/2^6$	$8.70\ 10^{-5}$	2.00	$8.83\ 10^{-5}$	2.00	$6.44\ 10^{0}$	$1.31\ 10^{1}$
$1/2^6$	$1/2^7$	$2.18\ 10^{-5}$	2.00	$2.21\ 10^{-5}$	2.00	$3.18\ 10^{1}$	$5.59\ 10^{1}$
$1/2^7$	$1/2^8$	$5.44\ 10^{-6}$	2.00	$5.51\ 10^{-6}$	2.00	$1.20\ 10^{2}$	$2.73\ 10^{2}$

Table 5.3: P_3; $\lambda = 1 + \frac{2\ln 2}{\ln h}$.

H	h	$\|\|u-u_h\|\|_{H^1}$	rate	$\|\|u-u^h\|\|_{H^1}$	rate	two-grid time	Newton time
$1/2^0$	$1/2^2$	$9.82\ 10^{-4}$	-	$3.43\ 10^{-1}$	-	$9.30\ 10^{-2}$	$2.50\ 10^{-1}$
$1/2^1$	$1/2^3$	$1.13\ 10^{-4}$	3.11	$3.84\ 10^{-3}$	6.48	$1.88\ 10^{-1}$	$7.96\ 10^{-1}$
$1/2^2$	$1/2^4$	$1.36\ 10^{-5}$	3.06	$3.70\ 10^{-5}$	6.70	$6.09\ 10^{-1}$	$2.30\ 10^{0}$
$1/2^3$	$1/2^5$	$1.67\ 10^{-6}$	3.03	$2.24\ 10^{-6}$	4.05	$2.32\ 10^{0}$	$8.31\ 10^{0}$
$1/2^4$	$1/2^6$	$2.07\ 10^{-7}$	3.01	$2.23\ 10^{-7}$	3.33	$9.80\ 10^{0}$	$3.94\ 10^{1}$
$1/2^5$	$1/2^7$	$2.57\ 10^{-8}$	3.01	$2.62\ 10^{-8}$	3.09	$4.95\ 10^{1}$	$2.04\ 10^{2}$

Table 5.4: P_3; $\lambda = 1 + \frac{\ln 2}{\ln h}$.

H	h	$\|\|u-u_h\|\|_{H^1}$	rate	$\|\|u-u^h\|\|_{H^1}$	rate	two-grid time	Newton time
$1/2^1$	$1/2^2$	$9.82\ 10^{-4}$	-	$1.85\ 10^{-3}$	-	$1.29\ 10^{-1}$	$2.50\ 10^{-1}$
$1/2^2$	$1/2^3$	$1.13\ 10^{-4}$	3.11	$1.16\ 10^{-4}$	3.99	$3.16\ 10^{-1}$	$7.96\ 10^{-1}$
$1/2^3$	$1/2^4$	$1.36\ 10^{-5}$	3.06	$1.37\ 10^{-5}$	3.09	$9.08\ 10^{-1}$	$2.30\ 10^{0}$
$1/2^4$	$1/2^5$	$1.67\ 10^{-6}$	3.03	$1.67\ 10^{-6}$	3.03	$3.87\ 10^{0}$	$8.31\ 10^{0}$
$1/2^5$	$1/2^6$	$2.07\ 10^{-7}$	3.01	$2.07\ 10^{-7}$	3.01	$1.44\ 10^{1}$	$3.94\ 10^{1}$
$1/2^6$	$1/2^7$	$2.57\ 10^{-8}$	3.01	$2.57\ 10^{-8}$	3.01	$7.10\ 10^{1}$	$2.04\ 10^{2}$

iteration on the fine grid does not affect the rate of convergence of the method.

The two-grid computations are accurate and fast compared with Newton's method. The computations were done in FreeFEM++ on an HP computer with Pentium dual-core 2.60 GHz processor running Windows 10.

Acknowledgments. Gerard Awanou was partially supported by NSF DMS grant # 1720276 and Hengguang Li by NSF DMS grant # 1819041 and by NSF China Grant 11628104.

References

[1] G. Awanou, Pseudo transient continuation and time marching methods for Monge-Ampère type equations, *Adv. Comput. Math.*, **41**:4 (2015), 907–935.

[2] G. Awanou, Standard finite elements for the numerical resolution of the elliptic Monge-Ampère equation: Aleksandrov solutions, *ESAIM Math. Model. Numer. Anal.*, **51**:2 (2017), 707–725.

[3] G. Awanou and H. Li, Error analysis of a mixed finite element method for the Monge-Ampère equation, *Int. J. Num. Analysis and Modeling*, **11** (2014), 745–761.

[4] J.H. Bramble, J.E. Pasciak and A. H. Schatz, The construction of preconditioners for elliptic problems by substructuring. I, *Math. Comp.*, **47**:175 (1986), 103–134.

[5] S.C. Brenner, T. Gudi, M. Neilan and L.Y. Sung, C^0 penalty methods for the fully nonlinear Monge-Ampère equation, *Math. Comp.*, **80**:276 (2011), 1979–1995.

[6] 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**:3 (2017), 803–832.

[7] S.C. Brenner and L.R. Scott, *The mathematical theory of finite element methods, volume 15 of Texts in Applied Mathematics*, Springer, New York, third edition, 2008.

[8] X. Feng, R. Glowinski and M. Neilan, Recent Developments in Numerical Methods for Fully Nonlinear Second Order Partial Differential Equations, *SIAM Rev.*, **55**:2 (2013), 205–267.

[9] E. Malitz, Two-grid discretization for interior penalty and mixed finite element approximations of the elliptic Monge-Ampère equation, Ph.D. Dissertation, University of Illinois at Chicago. USA, 2019.

[10] M. Neilan, Quadratic finite element approximations of the Monge-Ampère equation, *J. Sci. Comput.*, **54**:1 (2013), 200–226.

[11] 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.

[12] V. Thomée, J. Xu and N.Y. Zhang, Superconvergence of the gradient in piecewise linear finite-element approximation to a parabolic problem, *SIAM J. Numer. Anal.*, **26**:3 (1989), 553–573.

[13] J. Xu, Two-grid discretization techniques for linear and nonlinear PDEs, *SIAM J. Numer. Anal.*, **33**:5 (1996), 1759–1777.