MATHEMATICS OF COMPUTATION
Volume 86, Number 307, September 2017, Pages 2025–2051
http://dx.doi.org/10.1090/mcom/3168
Article electronically published on February 13, 2017

FINITE ELEMENT METHODS FOR SECOND ORDER LINEAR ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS IN NON-DIVERGENCE FORM

XIAOBING FENG, LAUREN HENNINGS, AND MICHAEL NEILAN

Abstract. This paper is concerned with finite element approximations of $W^{2,p}$ strong solutions of second-order linear elliptic partial differential equations (PDEs) in non-divergence form with continuous coefficients. A non-standard (primal) finite element method, which uses finite-dimensional subspaces consisting of globally continuous piecewise polynomial functions, is proposed and analyzed. The main novelty of the finite element method is to introduce an interior penalty term, which penalizes the jump of the flux across the interior element edges/faces, to augment a non-symmetric piecewise defined and PDE-induced bilinear form. Existence, uniqueness and error estimate in a discrete $W^{2,p}$ energy norm are proved for the proposed finite element method. This is achieved by establishing a discrete Calderon–Zygmund-type estimate and mimicking strong solution PDE techniques at the discrete level. Numerical experiments are provided to test the performance of proposed finite element methods and to validate the convergence theory.

1. Introduction

In this paper we consider finite element approximations of the following linear elliptic PDE in non-divergence form:

$$\mathcal{L}u := -A : D^2 u = f \qquad \text{in } \Omega, \tag{1.1a}$$

$$u = 0 \qquad \text{on } \partial\Omega. \tag{1.1b}$$

Here, $\Omega \subset \mathbb{R}^n$ is an open bounded domain with boundary $\partial\Omega$, $f \in L^p(\Omega)$ ($1 < p < \infty$) is given, and $A = A(x) \in \left[C^0(\overline{\Omega})\right]^{n\times n}$ is a positive definite matrix on $\overline{\Omega}$, but not necessarily differentiable. Problems such as (1.1) arise in fully non-linear elliptic Hamilton–Jacobi–Bellman equations, a fundamental problem in the field of stochastic optimal control [12, 17, 25, 26], although, in general, Hamilton–Jacobi–Bellman equations lead to discontinuous coefficient matrices. In addition, elliptic PDEs in non-divergence form appear in the linearization and numerical methods of fully non-linear second order PDEs [6, 11, 20].

Received by the editor May 11, 2015 and, in revised form, February 1, 2016 and April 1, 2016.
2010 *Mathematics Subject Classification.* Primary 65N30, 65N12, 35J25.
Key words and phrases. Finite element methods, non-divergence form, convergence analysis.
The first author was partially supported in part by NSF grants DMS-1016173 and DMS-1318486.
This work was partially supported by the NSF through grants DMS-1016173 and DMS-1318486 (Feng), and DMS-1417980 (Neilan) and the Alfred Sloan Foundation (Neilan).
The third author was partially supported in part by NSF grant DMS-1417980 and the Alfred Sloan Foundation.

Since A is not smooth, the PDE (1.1a) cannot be written in divergence form, and therefore notions of weak solutions defined by variational principles are not applicable. Instead, the existence and uniqueness of solutions are generally sought in the classical or strong sense. In the former case, Schauder theory states the existence of a unique solution $u \in C^{2,\alpha}(\Omega)$ to (1.1) provided the coefficient matrix and source function are Hölder continuous, and if the boundary satisfies $\partial\Omega \in C^{2,\alpha}$. In the latter case, the Calderon–Zygmund theory states the existence and uniqueness of $u \in W^{2,p}(\Omega)$ satisfying (1.1) almost everywhere provided $f \in L^p(\Omega)$, $A \in [C^0(\overline{\Omega})]^{n\times n}$ and $\partial\Omega \in C^{1,1}$. In addition, the existence of a strong solution to (1.1) in two dimensions and on convex domains is proved in [2, 3, 18].

Due to their non-divergence structure, designing convergent numerical methods, in particular, Galerkin-type methods, for problem (1.1) has been proven to be difficult. Very few such results are known in the literature. Nevertheless, while problem (1.1) does not naturally fit within the standard Galerkin framework, several finite element methods have been recently proposed. In [19] the authors considered mixed finite element methods using Lagrange finite element spaces for problem (1.1). An analogous discontinuous Galerkin (DG) method was proposed in [9]. The convergence analysis of these methods for non-smooth A remains open. A discontinuous Galerkin method for problem (1.1) with discontinuous coefficients satisfying the Cordes condition was proposed and analyzed in [24]. Here, the authors established optimal order estimates in h with respect to a H^2-type norm. This method has been extended to elliptic and parabolic Hamilton–Jacobi–Bellman equations in [25, 26] in which the C^0-continuity condition on the coefficient matrix A is added to the Cordes condition. Recently a two-scale, low-order finite element method based on an regularized integral formulation of (1.1) was proposed in [21]. Here, discrete Alexandroff–Bakelman–Pucci estimates are derived and, assuming the mesh is weakly acute, the authors prove suboptimal convergence rates in the L^∞-norm.

The primary goal of this paper is to develop and analyze a structurally simple and computationally easy finite element method for problem (1.1). The method, given in Definition 3.1 below, is a primal method using Lagrange finite element spaces. The method is well defined for all polynomials with degree greater than one and can be easily implemented on current finite element software. We note that our finite element method resembles interior penalty discontinuous Galerkin (DG) methods in its formulation because its bilinear form contains the jumps of the fluxes across the element edges/faces. However, no interior penalty term is used in the formulation, hence, our method is not a DG method per se. Moreover, we prove that the proposed method is stable and converges with optimal order in a discrete $W^{2,p}$-type norm on quasi-uniform meshes provided that the polynomial degree of the finite element space is greater than or equal to two and if the mesh is sufficiently fine. However, numerical experiments given in Section 4 suggest that the polynomial restriction and mesh size restriction can be relaxed.

While the formulation and implementation of the finite element method is relatively simple, the convergence analysis is quite involved, and it requires several non-standard arguments and techniques. The overall strategy in the convergence analysis is to mimic, at the discrete level, the stability analysis of strong solutions of PDEs in non-divergence form (see [14, Section 9.5]). Namely, we exploit the fact that, locally, the finite element discretization is a perturbation of a discrete elliptic operator in divergence form with constant coefficients; see Lemma 3.1. The first

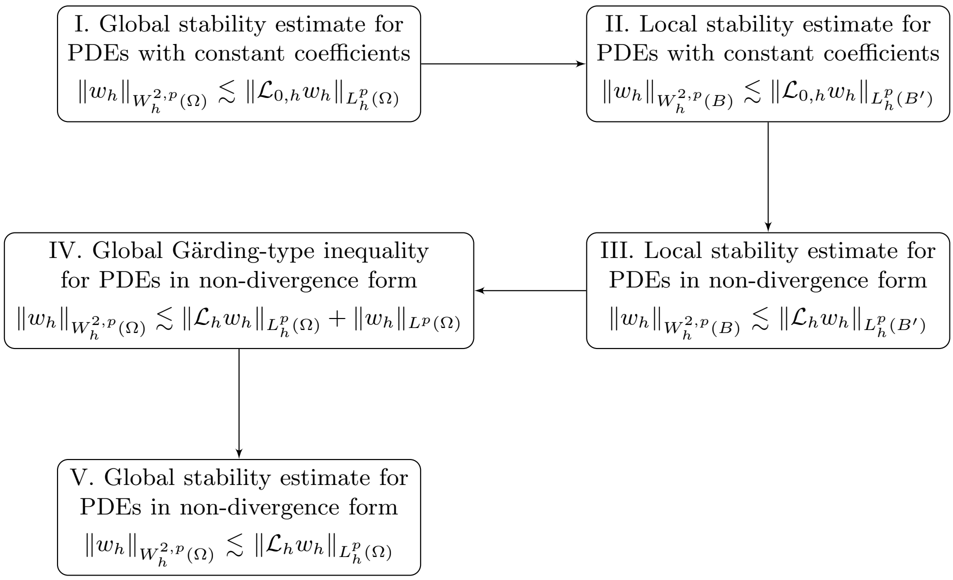

FIGURE 1. Outline of the convergence proof.

step of the stability argument is to establish a discrete Calderon–Zygmund-type estimate for the Lagrange finite element discretization of the elliptic operator in (1.1) with constant coefficients, which is equivalent to a global inf-sup condition for the discrete operator. The second step is to prove a local version of the global estimate and inf-sup condition. With these results in hand, local stability estimates for the proposed C^0 discretization of (1.1) can be easily obtained. We then glue these local stability estimates to obtain a global Gärding-type inequality. Finally, to circumvent the lack of a (discrete) maximum principle, which is often used in the PDE analysis, we use a non-standard duality argument to obtain a global inf-sup condition for the proposed C^0 discretization for problem (1.1). See Figure 1 for an outline of the convergence proof. Since the method is linear and consistent, the stability estimate naturally leads to the well-posedness of the method and the energy norm error estimate.

The organization of the paper is as follows. In Section 2 the assumptions of the PDE problem is stated, the notation is set, and some preliminary results are given. Discrete $W^{2,p}$ stability properties, including a discrete Calderon–Zygmund-type estimate, of finite element discretizations of PDEs with constant coefficients are established. In Section 3, we present the motivation and the formulation of our C^0 discontinuous finite element method for problem (1.1). Mimicking the PDE analysis from [14] at the discrete level, we prove a discrete $W^{2,p}$ stability estimate for the discretization operator. In addition, we derive an optimal order error estimate in a discrete $W^{2,p}$-norm. Finally, in Section 4, we give several numerical experiments which test the performance of the proposed C^0 finite element method and validate the convergence theory.

2. Assumptions, notation, and preliminary results

2.1. **The PDE problem.** To make the presentation clear, we state the precise assumptions on the non-divergence form PDE problem (1.1). Let $A \in [C^0(\overline{\Omega})]^{n\times n}$ be a positive definite matrix-valued function with

$$\lambda|\xi|^2 A(x)\xi\cdot\xi \le \Lambda|\xi|^2 \qquad \forall \xi\in\mathbb{R}^n,\ x\in\overline{\Omega} \tag{2.1}$$

and constants $0<\lambda\le\Lambda<\infty$. Under the above assumption, $\mathcal{L}$ is known to be uniformly elliptic, hence, strong solutions (i.e., $W^{2,p}$ solutions) of problem (1.1) must satisfy the *Aleksandrov maximum principle* for $p\ge n$ [10, 14, 16].

By the $W^{2,p}$ theory for the second order non-divergence form uniformly elliptic PDEs [14, Chapter 9], we know that if $\partial\Omega\in C^{1,1}$, for any $f\in L^p(\Omega)$ with $1<p<\infty$, there exists a unique strong solution $u\in W^{2,p}(\Omega)\cap W^{1,p}_0(\Omega)$ to (1.1) satisfying

$$\|u\|_{W^{2,p}(\Omega)} \lesssim \|f\|_{L^p(\Omega)}. \tag{2.2}$$

Moreover, when $n=2$ and $p=2$, it is also known that [2, 3, 13, 15, 18] the above conclusion holds if Ω is a convex domain.

For the remainder of the paper, we shall always assume that $A\in[C^0(\overline{\Omega})]^{n\times n}$ is positive definite satisfying (2.1), and problem (1.1) has a unique strong solution u which satisfies the Calderon–Zygmund estimate (2.2).

2.2. **Mesh and space notation.** Let $\Omega\subset\mathbb{R}^n$ be a bounded open domain. We shall use D to denote a generic subdomain of Ω and ∂D denotes its boundary. $W^{s,p}(D)$ denotes the standard Sobolev spaces for $s\ge 0$ and $1\le p\le\infty$, $W^{0,p}(D)=L^p(D)$ and $W^{s,p}_0(\Omega)$ to denote the subspace of $W^{s,p}(\Omega)$ consisting functions whose traces vanish up to order $s-1$ on $\partial\Omega$. $(\cdot,\cdot)_D$ denotes the standard inner product on $L^2(D)$ and $(\cdot,\cdot):=(\cdot,\cdot)_\Omega$. To avoid the proliferation of constants, we shall use the notation $a\lesssim b$ to represent the relation $a\le Cb$ for some constant $C>0$ independent of mesh size h.

Let $\mathcal{T}_h:=\mathcal{T}_h(\Omega)$ be a quasi-uniform, simplical, and conforming triangulation of the domain Ω. Denote by $\mathcal{E}_h^I$ the set of interior edges in $\mathcal{T}_h$, $\mathcal{E}_h^B$ the set of boundary edges in $\mathcal{T}_h$, and $\mathcal{E}_h=\mathcal{E}_h^I\cup\mathcal{E}_h^B$ the set of all edges in $\mathcal{T}_h$. We define the jump and average of a vector function $\boldsymbol{w}$ on an interior edge $e=\partial T^+\cap\partial T^-$ as follows:

$$\begin{aligned}
[\![\boldsymbol{w}]\!]\big|_e &= \boldsymbol{w}^+\cdot\boldsymbol{n}_+\big|_e + \boldsymbol{w}^-\cdot\boldsymbol{n}_-\big|_e,\\
\{\!\{\boldsymbol{w}\}\!\}\big|_e &= \frac{1}{2}\Big(\boldsymbol{w}^+\cdot\boldsymbol{n}_+\big|_e - \boldsymbol{w}^-\cdot\boldsymbol{n}_-\big|_e\Big),
\end{aligned}$$

where $\boldsymbol{w}^\pm=\boldsymbol{w}\big|_{T^\pm}$ and $\boldsymbol{n}_\pm$ is the outward unit normal of $T^\pm$.

For a normed linear space X, we denote by X^* its dual and $\langle\cdot,\cdot\rangle$ the pairing between X^* and X. The Lagrange finite element space with respect to the triangulation is given by

$$V_h := \big\{v_h\in H^1_0(\Omega):\ v_h|_T\in\mathbb{P}_k(T)\ \forall T\in\mathcal{T}_h\big\}, \tag{2.3}$$

where $\mathbb{P}_k(T)$ denotes the set of polynomials with total degree not exceeding k (≥ 1) on T. We also define the piecewise Sobolev space with respect to the mesh $\mathcal{T}_h$:

$$\begin{aligned}
&W^{s,p}(\mathcal{T}_h) := \prod_{T\in\mathcal{T}_h} W^{s,p}(T), \qquad W_h^{(p)} := W^{2,p}(\mathcal{T}_h)\cap W^{1,p}_0(\Omega),\\
&L^p_h(\mathcal{T}_h) := \prod_{T\in\mathcal{T}_h} L^p(T), \qquad W^{s,p}_h(D) := W^{s,p}(\mathcal{T}_h)\big|_D, \qquad L^p_h(D) := L^p(\mathcal{T}_h)\big|_D.
\end{aligned}$$

For a given subdomain $D \subseteq \Omega$, we also define $V_h(D) \subseteq V_h$ and $W_h^{(p)}(D) \subseteq W_h^{(p)}$ as the subspaces that vanish outside of D by

$$V_h(D) : = \{v \in V_h;\, v|_{\Omega\setminus D} = 0\}, \qquad W_h^{(p)}(D) := \{v \in W_h^{(p)};\ v|_{\Omega\setminus D} = 0\}.$$

We note that $V_h(D)$ is non-trivial for $r_D > 2h$, where r_D is the radius of the largest ball inscribed in D.

Associated with $D \subseteq \Omega$, we define a semi-norm on $W_h^{2,p}(D)$ for $1 < p < \infty$:

$$\|v\|_{W_h^{2,p}(D)} = \|D_h^2 v\|_{L^p(D)} + \Big(\sum_{e\in\mathcal{E}_h^I} h_e^{1-p} \| [\![\nabla v]\!] \|^p_{L^p(e\cap\overline{D})}\Big)^{\frac{1}{p}}. \tag{2.4}$$

Here, $D_h^2 v \in L^2(\Omega)$ denotes the piecewise Hessian matrix of v, i.e., $D_h^2 v|_T = D^2 v|_T$ for all $T \in \mathcal{T}_h$.

Let $\mathcal{Q}_h : L^p(\Omega) \to V_h$ be the L^2 projection defined by

$$\big(\mathcal{Q}_h w, v_h\big) = \big(w, v_h\big) \qquad \forall w \in L^2(\Omega),\, v_h \in V_h. \tag{2.5}$$

It is well known that [8] $\mathcal{Q}_h$ satisfies for any $w \in W^{m,p}(\Omega)$

$$\|\mathcal{Q}_h w\|_{W^{m,p}(\Omega)} \lesssim \|w\|_{W^{m,p}(\Omega)} \qquad m = 0, 1;\ 1 < p < \infty. \tag{2.6}$$

For any domain $D \subseteq \Omega$ with $r_D > 2h$ and any $w \in L_h^p(D)$, we also introduce the following mesh-dependent semi-norm

$$\|w\|_{L_h^p(D)} := \sup_{0\neq v_h\in V_h(D)} \frac{\big(w, v_h\big)_D}{\|v_h\|_{L^{p'}(D)}}, \qquad \text{where} \quad \frac{1}{p} + \frac{1}{p'} = 1. \tag{2.7}$$

By (2.5), it is easy to see that $\|\cdot\|_{L_h^p(D)}$ is a norm on $V_h(D)$. Moreover, by (2.6),

$$\begin{aligned}\|w_h\|_{L^p(\Omega)} &= \sup_{v\in L^{p'}(\Omega)} \frac{(w_h, v)}{\|v\|_{L^{p'}(\Omega)}} = \sup_{v\in L^{p'}(\Omega)} \frac{(w_h, \mathcal{Q}_h v)}{\|v\|_{L^{p'}(\Omega)}} \\ &\lesssim \sup_{v\in L^{p'}(\Omega)} \frac{(w_h, \mathcal{Q}_h v)}{\|\mathcal{Q}_h v\|_{L^{p'}(\Omega)}} \le \|w_h\|_{L_h^p(\Omega)} \qquad \forall w_h \in V_h.\end{aligned} \tag{2.8}$$

2.3. **Some basic properties of $W_h^{(p)}$ functions.** In this subsection we cite or prove some basic properties of the broken Sobolev functions in $W_h^{(p)}$, and in particular, for piecewise polynomial functions. These results, which have independent interest in themselves, will be used repeatedly in the later sections. We begin with citing a familiar trace inequality followed by proving an inverse inequality.

Lemma 2.1 ([4])**.** *For any $T \in \mathcal{T}_h$, there holds*

$$\|v\|^p_{L^p(\partial T)} \lesssim \big(h_T^{p-1} \|\nabla v\|^p_{L^p(T)} + h_T^{-1}\|v\|^p_{L^p(T)}\big) \qquad \forall v \in W^{1,p}(T) \tag{2.9}$$

for any $p \in (1, \infty)$. Therefore by scaling, there holds

$$\sum_{e\in\mathcal{E}_h^I} h_e \|v\|^p_{L^p(e\cap\bar{D})} \lesssim \begin{cases} \|v\|^p_{L^p(D)} & \forall v \in V_h(D)\ (r_D > 2h), \\ \|v\|^p_{L^p(D)} + h^p\|\nabla v\|^p_{L^p(D)} & \forall v \in W_h^{(p)}(D). \end{cases} \tag{2.10}$$

Lemma 2.2. *For any $v_h \in V_h,\ 1 < p < \infty$, and $D \subseteq \Omega$ with $r_D > 2h$, there holds*

$$\|v_h\|_{W_h^{2,p}(D)} \lesssim h^{-1}\|v_h\|_{W^{1,p}(D_h)}, \tag{2.11}$$

where

$$D_h = \{x \in \Omega : \ \mathrm{dist}(x, D) \le h\}. \tag{2.12}$$

Proof. By (2.4), (2.9) and inverse estimates [4, 7], we have

$$\begin{aligned}
\|v_h\|_{W_h^{2,p}(D)} &= \|D_h^2 v_h\|_{L^p(D)} + \Big(\sum_{e\in\mathcal{E}_h^I} h_e^{1-p} \| [\![\nabla v_h]\!] \|_{L^p(e\cap\bar{D})}^p \Big)^{\frac{1}{p}} \\
&\lesssim \|D_h^2 v_h\|_{L^p(D)} + \sum_{\substack{T\in\mathcal{T}_h \\ T\subset D_h}} \Big(h_T^{1-p} \big(h_T^{p-1} \|D^2 v_h\|_{L^p(T)}^p + h_T^{-1} \|\nabla v_h\|_{L^p(T)}^p \big) \Big)^{\frac{1}{p}} \\
&\lesssim h^{-1} \|v_h\|_{W^{1,p}(D_h)}. \qquad \square
\end{aligned}$$

The next lemma states a very simple fact about the discrete $W^{2,p}$-norm on $W_h^{2,p}(\Omega)$.

Lemma 2.3. *For any $1 < p < \infty$, there holds*

$$\|\varphi\|_{W_h^{2,p}(\Omega)} \le \|\varphi\|_{W^{2,p}(\Omega)} \qquad \forall \varphi \in W^{2,p}(\Omega). \tag{2.13}$$

Next, we state some super-approximation results of the nodal interpolant with respect to the discrete $W^{2,p}$-semi-norm. The derivation of the following results is standard [22], but for completeness we give the proof in Appendix A

Lemma 2.4. *Denote by $I_h : C^0(\overline{\Omega}) \to V_h$ the nodal interpolant onto V_h. Let $\eta \in C^\infty(\Omega)$ with $|\eta|_{W^{j,\infty}(\Omega)} \lesssim d^{-j}$ for $0 \le j \le k$. Then for each $T \in \mathcal{T}_h$ with $h \le d \le 1$, there holds*

$$h^m \|\eta v_h - I_h(\eta v_h)\|_{W^{m,p}(D)} \lesssim \frac{h}{d} \|v_h\|_{L^p(D_h)} \qquad \text{for } m = 0, 1, \tag{2.14}$$

$$\|\eta v_h - I_h(\eta v_h)\|_{W_h^{2,p}(D)} \lesssim \frac{1}{d^2} \|v_h\|_{W^{1,p}(D_h)}. \tag{2.15}$$

Moreover, if $k \ge 2$, there holds

$$\|\eta v_h - I_h(\eta v_h)\|_{W_h^{2,p}(D)} \lesssim \frac{h}{d^3} \|v_h\|_{W^{2,p}(D_h)}. \tag{2.16}$$

Here, $D \subset D_h \subset \Omega$ satisfy the conditions in Lemma 2.2.

To conclude this subsection, we state and prove a discrete Sobolev interpolation estimate.

Lemma 2.5. *There holds for all $2 \le p < \infty$,*

$$\|\nabla w\|_{L^p(\Omega)}^2 \lesssim \|w\|_{L^p(\Omega)} \|w\|_{W_h^{2,p}(\Omega)} \qquad \forall w \in W_h^{(p)}.$$

Proof. Writing $\|\nabla w\|_{L^p(\Omega)}^p = \int_\Omega |\nabla w|^{p-2} \nabla w \cdot \nabla w \, dx$ and integrating by parts, we find

$$\begin{aligned}
\|\nabla w\|_{L^p(\Omega)}^p &= -\int_\Omega \Big(|\nabla w|^{p-2} \Delta w + (p-2) |\nabla w|^{p-4} (D_h^2 w \nabla w) \cdot \nabla w \Big) w \, dx \\
&\qquad + \sum_{e\in\mathcal{E}_h^I} \int_e [\![|\nabla w|^{p-2} \nabla w]\!] w \, ds \\
&\lesssim \sum_{T\in\mathcal{T}_h} \int_T |\nabla w|^{p-2} |D_h^2 w| |w| \, dx + \sum_{e\in\mathcal{E}_h^I} \int_e [\![|\nabla w|^{p-2} \nabla w]\!] w \, ds. \qquad (2.17)
\end{aligned}$$

To bound the first term in (2.17) we apply Hölder's inequality to obtain

$$\begin{aligned}\int_\Omega |\nabla w|^{p-2}|D_h^2 w||w|\,dx &\le \big\||\nabla w|^{p-2}\big\|_{L^{\frac{p}{p-2}}(\Omega)}\|D^2 w\|_{L^p(\Omega)}\|w\|_{L^p(\Omega)}\\ &= \|\nabla w\|_{L^p(\Omega)}^{p-2}\|D_h^2 w\|_{L^p(\Omega)}\|w\|_{L^p(\Omega)}.\end{aligned}\tag{2.18}$$

Likewise, by Lemma 2.1, we have

$$\begin{aligned}&\sum_{e\in\mathcal{E}_h^I}\int_e \big[\!\big[|\nabla w|^{p-2}\nabla w\big]\!\big] w\,ds\\ &\le \sum_{e\in\mathcal{E}_h^I}\Big(h_e^{\frac{1}{p}}\big\|\big[\!\big[\nabla w\big]\!\big]\big\|_{L^p(e)}\Big)^{p-2}\Big(h_e^{\frac{1-p}{p}}\|\big[\!\big[\nabla w\big]\!\big]\|_{L^p(e)}\Big)\Big(h_e^{\frac{1}{p}}\|w\|_{L^p(e)}\Big)\\ &\lesssim \|\nabla w\|_{L^p(\Omega)}^{p-2}\|w\|_{L^p(\Omega)}\Big(\sum_{e\in\mathcal{E}_h^I} h_e^{1-p}\big\|\big[\!\big[\nabla w\big]\!\big]\big\|_{L^p(e)}^p\Big)^{\frac{1}{p}}\\ &\lesssim \|\nabla w\|_{L^p(\Omega)}^{p-2}\|w\|_{L^p(\Omega)}\|w\|_{W_h^{2,p}(\Omega)}.\end{aligned}\tag{2.19}$$

Combining (2.17)–(2.19) we obtain the desired result. The proof is complete. □

2.4. **Stability estimates for auxiliary PDEs with constant coefficients.** In this subsection, we consider a special case of (1.1a) when the coefficient matrix is a constant matrix, $A(x) \equiv A_0 \in \mathbb{R}^{n\times n}$. We introduce the finite element approximation (or projection) $\mathcal{L}_{0,h}$ of $\mathcal{L}_0$ on V_h and extend $\mathcal{L}_{0,h}$ to the broken Sobolev space $W_h^{(p)}$. We then establish some stability results for the operator $\mathcal{L}_{0,h}$. These stability results will play an important role in our convergence analysis of the proposed C^0 finite element method in Section 3.

Let $A_0 \in \mathbb{R}^{n\times n}$ be a positive definite matrix and set

$$\mathcal{L}_0 w := -A_0 : D^2 w = -\nabla\cdot\big(A_0\nabla w\big).\tag{2.20}$$

The operator $\mathcal{L}_0$ induces the following bilinear form:

$$a_0(w,v) := \big\langle \mathcal{L}_0 w, v\big\rangle = \int_\Omega A_0\nabla w\cdot\nabla v\,dx \qquad \forall w,v\in H_0^1(\Omega),\tag{2.21}$$

and the Lax–Milgram Theorem (cf. [10]) implies that $\mathcal{L}_0^{-1} : H^{-1}(\Omega) \to H_0^1(\Omega)$ exists and is bounded. Moreover, if $\partial\Omega \in C^{1,1}$, the Calderon–Zygmund theory (cf. [14, Chapter 9]) infers that $\mathcal{L}_0^{-1} : L^p(\Omega) \to W^{2,p}(\Omega)\cap W_0^{1,p}(\Omega)$ exists and there holds

$$\|\mathcal{L}_0^{-1}\varphi\|_{W^{2,p}(\Omega)} \lesssim \|\varphi\|_{L^p(\Omega)} \qquad \forall\varphi\in L^p(\Omega).\tag{2.22}$$

Equivalently,

$$\|w\|_{W^{2,p}(\Omega)} \lesssim \big\|\mathcal{L}_0 w\big\|_{L^p(\Omega)} \qquad \forall w\in W^{2,p}\cap W_0^{1,p}(\Omega).\tag{2.23}$$

The bilinear form naturally leads to a finite element approximation (or projection) of $\mathcal{L}_0$ on V_h, that is, we define the operator $\mathcal{L}_{0,h} : V_h \to V_h$ by

$$\big(\mathcal{L}_{0,h} w_h, v_h\big) := a_0(w_h, v_h) \qquad \forall v_h, w_h \in V_h.\tag{2.24}$$

Remark 2.1. When $A = I$, the identity matrix, $\mathcal{L}_{0,h}$ is exactly the finite element of the discrete Laplacian, that is, $\mathcal{L}_{0,h} = -\Delta_h$. By finite element theory [4], we know that $\mathcal{L}_{0,h} : V_h \to V_h$ is one-to-one and onto, and therefore $\mathcal{L}_{0,h}^{-1} : V_h \to V_h$ exists.

Recall the following DG integration by parts formula:

$$\int_\Omega \tau \cdot \nabla_h v\,dx = -\int_\Omega (\nabla_h \cdot \tau) v\,dx + \sum_{e\in\mathcal{E}_h^I}\Big(\int_e [\![\tau]\!]\,\{\!\{v\}\!\}\,ds + \int_e \{\!\{\tau\}\!\}\cdot [\![v]\!]\,ds\Big) + \sum_{e\in\mathcal{E}_h^B}\int_e (\tau\cdot n_e) v\,ds, \tag{2.25}$$

which holds for any piecewise scalar-valued function v and vector-valued function τ. Here, ∇_h is defined piecewise, i.e., $\nabla_h|_T = \nabla|_T$ for all $T \in \mathcal{T}_h$. For any $w_h, v_h \in V_h$, using (2.25) with $\tau = A_0\nabla w_h$, we obtain

$$a_0(w_h, v_h) = -\int_\Omega \big(A_0 : D_h^2 w_h\big) v_h\,dx + \sum_{e\in\mathcal{E}_h^I}\int_e [\![A_0\nabla w_h]\!]\, v_h\,ds. \tag{2.26}$$

We note that the above new form of $a_0(\cdot,\cdot)$ is not well defined on $H_0^1(\Omega)\times H_0^1(\Omega)$. However, it is well defined on $W_h^{(p)} \times W_h^{(p')}$ with $\frac{1}{p}+\frac{1}{p'} = 1$. Hence, we can easily extend the domain of the operator $\mathcal{L}_{0,h}$ to broken Sobolev space $W_h^{(p)}$. Precisely, (abusing the notation) we define $\mathcal{L}_{0,h} : W_h^{(p)} \to (W_h^{(p')})^*$ to be the operator induced by the bilinear form $a_0(\cdot,\cdot)$ on $W_h^{(p)}\times W_h^{(p')}$, namely,

$$\big\langle \mathcal{L}_{0,h} w, v\big\rangle := a_0(w,v) \qquad \forall w \in W_h^{(p)},\ v \in W_h^{(p')}. \tag{2.27}$$

A key ingredient in the convergence analysis of our finite element methods for PDEs in non-divergence form is to establish global and local discrete Calderon–Zygmund-type estimates similar to (2.23) for $\mathcal{L}_{0,h}$. These results are presented in the following two lemmas.

Lemma 2.6. *There exists $h_0 > 0$ such that for all $h \in (0, h_0)$, there holds*

$$\|w_h\|_{W_h^{2,p}(\Omega)} \lesssim \|\mathcal{L}_{0,h} w_h\|_{L^p(\Omega)} \qquad \forall w_h \in V_h. \tag{2.28}$$

Proof. First note that (2.28) is equivalent to

$$\big\|\mathcal{L}_{0,h}^{-1}\varphi_h\big\|_{W_h^{2,p}(\Omega)} \lesssim \|\varphi_h\|_{L^p(\Omega)} \qquad \forall \varphi_h \in V_h. \tag{2.29}$$

For any fixed $\varphi_h \in V_h$, let $w := \mathcal{L}_0^{-1}\varphi_h \in W^{2,p}(\Omega)\cap W_0^{1,p}(\Omega)$ and $w_h := \mathcal{L}_{0,h}^{-1}\varphi_h \in V_h$. Therefore, w and w_h, respectively, are the solutions of the following two problems:

$$a_0(w,v) = (\varphi_h, v) \quad \forall v \in H_0^1(\Omega), \qquad a_0(w_h, v_h) = (\varphi_h, v_h) \quad \forall v_h \in V_h, \tag{2.30}$$

and thus, w_h is the elliptic projection of w.

By (2.23) we have

$$\|w\|_{W^{2,p}(\Omega)} \lesssim \|\varphi_h\|_{L^p(\Omega)}. \tag{2.31}$$

Using well-known L^p finite element estimate results [4, Theorem 8.5.3], finite element interpolation theory, and (2.31) we obtain that there exists $h_0 > 0$ such that for all $h \in (0, h_0)$,

$$\|w - w_h\|_{W^{1,p}(\Omega)} \lesssim \|w - I_h w\|_{W^{1,p}(\Omega)} \lesssim h\|w\|_{W^{2,p}(\Omega)} \lesssim h\|\varphi_h\|_{L^p(\Omega)}. \tag{2.32}$$

It follows from the triangle inequality, an inverse inequality (see Lemma 2.2), the stability of I_h, (2.31) and (2.32) that

$$\begin{aligned}\|w-w_h\|_{W_h^{2,p}(\Omega)} &\lesssim \|w-I_hw\|_{W_h^{2,p}(\Omega)}+\|I_hw-w_h\|_{W_h^{2,p}(\Omega)}\\ &\lesssim \|w\|_{W^{2,p}(\Omega)}+h^{-1}\|I_hw-w_h\|_{W^{1,p}(\Omega)}\\ &\lesssim \|\varphi_h\|_{L^p(\Omega)}+h^{-1}\|I_hw-w\|_{W^{1,p}(\Omega)}+h^{-1}\|w-w_h\|_{W^{1,p}(\Omega)}\\ &\lesssim \|\varphi_h\|_{L^p(\Omega)}.\end{aligned}$$

Thus,

$$\|w_h\|_{W_h^{2,p}(\Omega)}\lesssim \|w-w_h\|_{W_h^{2,p}(\Omega)}+\|w\|_{W^{2,p}(\Omega)}\lesssim \|\varphi_h\|_{L^p(\Omega)},$$

which yields (2.29), and hence, (2.28). □

Lemma 2.7. *For $x_0\in\Omega$ and $R>0$, define*

$$B_R(x_0):=\{x\in\Omega:\ |x-x_0|<R\}\subset\Omega. \tag{2.33}$$

Let $R'=R+d$ with $d\geq 2h$. Then there holds

$$\|w_h\|_{W_h^{2,p}(B_R(x_0))}\lesssim \left\|\mathcal{L}_{0,h}w_h\right\|_{L_h^p(B_{R'}(x_0))}\qquad \forall w_h\in V_h(B_R(x_0)). \tag{2.34}$$

Proof. To ease notation, set $B_R:=B_R(x_0)$ and $B_{R'}:=B_{R'}(x_0)$. Recalling (2.8), we have by Lemma 2.6,

$$\begin{aligned}\|w_h\|_{W_h^{2,p}(B_R)}=\|w_h\|_{W_h^{2,p}(\Omega)}&\lesssim \|\mathcal{L}_{0,h}w_h\|_{L^p(\Omega)}\lesssim \|\mathcal{L}_{0,h}w_h\|_{L_h^p(\Omega)}\\ &=\sup_{v_h\in V_h}\frac{a_0(w_h,v_h)}{\|v_h\|_{L^{p'}(\Omega)}}.\end{aligned}$$

Set $R''=(R+R')/2$, so that $R<R''<R'$. Denote by $\chi_{B_{R''}}$ the indicator function of $B_{R''}:=B_{R''}(x_0)$. Since $w_h=0$ on $\Omega\backslash B_R$, we have

$$a_0(w_h,v_h)=a_0(w_h,\chi_{B_{R''}}v_h)=a_0(w_h,I_h(\chi_{B_{R''}}v_h))\qquad \forall v_h\in V_h.$$

Moreover, we have $I_h(\chi_{B_{R''}}v_h)\in V_h(B_{R'})$ and

$$\|I_h(\chi_{B_{R''}}v_h)\|_{L^{p'}(B_{R'})}=\|I_h(\chi_{B_{R''}}v_h)\|_{L^{p'}(\Omega)}\lesssim \|\chi_{B_{R''}}v_h\|_{L^{p'}(\Omega)}\lesssim \|v_h\|_{L^{p'}(\Omega)}.$$

Consequently,

$$\begin{aligned}\|w_h\|_{W_h^{2,p}(B_R)}&\lesssim \sup_{v_h\in V_h}\frac{a_0(w_h,I_h(\chi_{B_{R''}}v_h))}{\|I_h(\chi_{B_{R''}}v_h)\|_{L^{p'}(B_{R'})}}\leq \sup_{v_h\in V_h(B_{R'})}\frac{a_0(w_h,v_h)}{\|v_h\|_{L^{p'}(B_{R'})}}\\ &=\sup_{v_h\in V_h(B_{R'})}\frac{(\mathcal{L}_{0,h}w_h,v_h)}{\|v_h\|_{L^{p'}(B_{R'})}}=\|\mathcal{L}_{0,h}w_h\|_{L_h^p(B_{R'})}.\end{aligned}$$ □

3. Finite element methods and convergence analysis

3.1. **Formulation of finite element methods.** The formulation of our C^0 finite element method for non-divergence form PDEs is relatively simple, which is inspired by the finite element method for diffusion-convection PDEs and relies only on an unorthodox integration by parts.

To motivate its derivation, we first look at how one would construct standard finite element methods for problem (1.1) when the coefficient matrix A belongs to $[C^1(\overline{\Omega})]^{n\times n}$. In this case, since the divergence of A (taken row-wise) is well defined, we can rewrite the PDE (1.1a) in divergence form as follows:

$$-\nabla\cdot(A\nabla u)+(\nabla\cdot A)\cdot\nabla u=f. \tag{3.1}$$

Hence, the original non-divergence form PDE is converted into a "diffusion -convection equation" with the "diffusion coefficient" A and the "convection coefficient" $\nabla \cdot A$.

A standard finite element method for problem (3.1) is readily defined as seeking $u_h \in V_h$ such that

$$\int_\Omega (A\nabla u_h) \cdot \nabla v_h \, dx + \int_\Omega (\nabla \cdot A) \cdot \nabla u_h v_h \, dx = \int_\Omega f v_h \, dx \quad \forall v_h \in V_h. \tag{3.2}$$

Now come back to the case where A only belongs to $[C^0(\overline{\Omega})]^{n\times n}$. In our setting, the formulation (3.2) is not viable any more because $\nabla \cdot A$ does not exist as a function. To circumvent this issue, we apply the DG integration by parts formula (2.25) to the first term on the left-hand side of (3.2) with $\tau = A\nabla u_h$ and ∇ in (3.2) is understood piecewise, we get

$$-\int_\Omega (A : D_h^2 u_h) v_h \, dx + \sum_{e\in\mathcal{E}_h^I} \int_e [\![A\nabla u_h]\!] v_h \, ds = \int_\Omega f v_h \, dx \qquad \forall v_h \in V_h. \tag{3.3}$$

Here, we have used the fact that $[\![v_h]\!] = 0$ and $v_h|_{\partial\Omega} = 0$.

No derivative is taken on A in (3.3), so each of the terms is well defined on V_h. This indeed yields the C^0 formulation of this paper.

Definition 3.1. The C^0 finite element method is defined by seeking $u_h \in V_h$ such that

$$a_h(u_h, v_h) = (f, v_h) \qquad \forall v_h \in V_h, \tag{3.4}$$

where

$$a_h(w_h, v_h) := -\int_\Omega (A : D_h^2 w_h) v_h \, dx + \sum_{e\in\mathcal{E}_h^I} \int_e [\![A\nabla w_h]\!] v_h \, ds, \tag{3.5}$$

$$(f, v_h) := \int_\Omega f v_h \, dx \qquad \forall v_h \in V_h. \tag{3.6}$$

A few remarks are given below about the proposed C^0 finite element method.

Remark 3.1. (a) The C^0 finite element method (3.4) is a primal method with the single unknown u_h. It can be implemented on current finite element software supporting element boundary integration.

(b) From its derivation we see that (3.4) is equivalent to the standard finite element method (3.2) provided A is smooth. In addition, if A is constant, then (3.4) reduces to

$$a_0(u_h, v_h) = (f, v_h) \qquad \forall v_h \in V_h.$$

This feature will be crucially used in the convergence analysis later.

(c) In the one-dimensional and piecewise linear case (i.e., $n = 1$ and $k = 1$), the method (3.4) on a uniform mesh $\{x_i\}_{i=1}^N$ is equivalent to

$$A(x_i)\big(-c_{i-1} + 2c_i - c_{i+1}\big) = h^2 f(x_i),$$

where $u_h = \sum_{i=1}^N c_i \varphi_h^{(i)}$, and $\{\varphi_h^{(i)}\}_{i=1}^N$ represents the nodal basis for V_h.

3.2. **Stability analysis and well-posedness theorem.** As in Section 2.4, using the bilinear form $a_h(\cdot,\cdot)$ we can define the finite element approximation (or projection) $\mathcal{L}_h$ of $\mathcal{L}$ on V_h, that is, we define $\mathcal{L}_h : V_h \to V_h$ by

$$\big(\mathcal{L}_h w_h, v_h\big) := a_h(w_h, v_h) \qquad \forall v_h, w_h \in V_h. \tag{3.7}$$

Trivially, (3.4) can be rewritten as: Find $u_h \in V_h$ such that

$$\big(\mathcal{L}_h u_h, v_h\big) = (f, v_h) \qquad \forall v_h \in V_h.$$

Similar to the argument for $\mathcal{L}_{0,h}$, we can extend the domain of $\mathcal{L}_h$ to the broken Sobolev space $W_h^{(p)}$, that is, (abusing the notation) we define $\mathcal{L}_h : W_h^{(p)} \to (W_h^{(p')})^*$ by

$$\big\langle \mathcal{L}_h w, v\big\rangle := a_h(w, v) \qquad \forall w \in W_h^{(p)},\ v \in W_h^{(p')}. \tag{3.8}$$

The main objective of this subsection is to establish a $W_h^{2,p}$ stability estimate for the operator $\mathcal{L}_h$ on the finite element space V_h. From this result, the existence, uniqueness and error estimate for (3.4) will naturally follow. The stability proof relies on several technical estimates which we derive below. Essentially, the underlying strategy, known as a perturbation argument in the PDE literature, is to treat the operator $\mathcal{L}_h$ locally as a perturbation of a stable operator with constant coefficients. The following lemma quantifies this statement.

Lemma 3.1. *For any $\delta > 0$, there exists $R_\delta > 0$ and $h_\delta > 0$ such that for any $x_0 \in \Omega$ with $A_0 = A(x_0)$,*

$$\|(\mathcal{L}_h - \mathcal{L}_{0,h})w\|_{L_h^p(B_{R_\delta}(x_0))} \lesssim \delta \|w\|_{W_h^{2,p}(B_{R_\delta}(x_0))} \quad \forall w \in W_h^{(p)},\ \forall h \le h_\delta. \tag{3.9}$$

Proof. Since A is continuous on $\overline{\Omega}$, it is uniformly continuous. Therefore, for every $\delta > 0$ there exists $R_\delta > 0$ such that if $x, y \in \Omega$ satisfy $|x-y| < R_\delta$, there holds $|A(x) - A(y)| < \delta$. Consequently for any $x_0 \in \Omega$

$$\|A - A_0\|_{L^\infty(B_{R_\delta})} \le \delta, \tag{3.10}$$

with $B_{R_\delta} := B_{R_\delta}(x_0)$.

Set $h_\delta = \min\{h_0, \frac{R_\delta}{4}\}$ and consider $h \le h_\delta$, $w \in W_h^{(p)}$ and $v_h \in V_h(B_{R_\delta})$. Since $(\mathcal{L}_{0,h} - \mathcal{L}_h)w \in W_h^{(p)}$, it follows from (2.8), (2.26), (3.5), (3.10), and (2.6) that

$$\begin{aligned}
&\big((\mathcal{L}_{0,h} - \mathcal{L}_h)w, v_h\big)\\
&\quad = -\int_{B_{R_\delta}} \big((A_0 - A) : D_h^2 w\big) v_h\,dx + \sum_{e\in\mathcal{E}_h^I} \int_{e\cap \bar{B}_{R_\delta}} [\![(A_0 - A)\nabla w]\!] v_h\,ds\\
&\quad \le \|A - A_0\|_{L^\infty(B_{R_\delta})} \Big(\|D_h^2 w\|_{L^p(B_{R_\delta})} \|v_h\|_{L^{p'}(B_{R_\delta})}\\
&\qquad + \Big(\sum_{e\in\mathcal{E}_h^I} h_e^{1-p} \|[\![\nabla w]\!]\|^p_{L^p(e\cap\bar{B}_{R_\delta})}\Big)^{\frac{1}{p}} \Big(\sum_{e\in\mathcal{E}_h^I} h_e \|v_h\|^{p'}_{L^{p'}(e\cap \bar{B}_{R_\delta})}\Big)^{\frac{1}{p'}}\Big)\\
&\quad \lesssim \|A - A_0\|_{L^\infty(B_{R_\delta})} \|w\|_{W_h^{2,p}(B_{R_\delta})} \|v_h\|_{L^{p'}(B_{R_\delta})} \lesssim \delta \|w\|_{W_h^{2,p}(B_{R_\delta})} \|v_h\|_{L^{p'}(B_{R_\delta})}.
\end{aligned}$$

The desired inequality now follows from the definition of $\|\cdot\|_{L_h^p(B_{R_\delta})}$. $\square$

Lemma 3.2. *There exists $R_1 > 0$ and $h_1 > 0$ such that for any $x_0 \in \Omega$*

$$\|w_h\|_{W_h^{2,p}(B_{R_1}(x_0))} \lesssim \|\mathcal{L}_h w_h\|_{L_h^p(B_{R_2}(x_0))} \quad \forall w_h \in V_h(B_{R_1}(x_0)),\ \forall h \le h_1, \tag{3.11}$$

with $R_2 = 2R_1$.

Proof. For $\delta_0 > 0$ to be determined below, let $R_1 = \frac{1}{2}R_{\delta_0}$ as in Lemma 3.1. Let $h_1 = \frac{R_1}{2}$ and set $B_i = B_{R_i}(x_0)$. Then by Lemmas 2.7 and 3.1 with $d = R_1$ and $A_0 = A(x_0)$, we have for any $w_h \in V_h(B_1)$

$$\begin{aligned}\|w_h\|_{W_h^{2,p}(B_1)} &\lesssim \|\mathcal{L}_{0,h} w_h\|_{L_h^p(B_2)} \le \|(\mathcal{L}_{0,h} - \mathcal{L}_h) w_h\|_{L_h^p(B_2)} + \|\mathcal{L}_h w_h\|_{L_h^p(B_2)} \\ &\lesssim \delta_0 \|w_h\|_{W_h^{2,p}(B_2)} + \|\mathcal{L}_h w_h\|_{L_h^p(B_2)} \\ &= \delta_0 \|w_h\|_{W_h^{2,p}(B_1)} + \|\mathcal{L}_h w_h\|_{L_h^p(B_2)}.\end{aligned}$$

For δ_0 sufficiently small (depending only on A), we can kick back the first term on the right-hand side. This completes the proof. $\square$

Lemma 3.3. *Let R_1 and h_1 be as in Lemma* 3.2*. For any $x_0 \in \Omega$, there holds*

$$\|\mathcal{L}_h w\|_{L_h^p(B_{R_1}(x_0))} \lesssim \|w\|_{W_h^{2,p}(B_{R_1}(x_0))} \qquad \forall w \in W_h^{(p)},\ \forall h \le h_1. \tag{3.12}$$

Proof. Set $B_1 = B_{R_1}(x_0)$. By the definition of $\mathcal{L}_h$, (2.8), (2.10) and (2.6), we have for any $v_h \in V_h(B_1)$

$$\begin{aligned}(\mathcal{L}_h w, v) &= -\int_{B_1} (A : D_h^2 w) v_h\, dx + \sum_{e \in \mathcal{E}_h^I} \int_{e \cap \bar{B}_1} [\![A \nabla w]\!] v_h\, ds \\ &\lesssim \|D_h^2 w\|_{L^p(B_1)} \|v_h\|_{L^{p'}(B_1)} \\ &\qquad + \Big(\sum_{e \in \mathcal{E}_h^I} h_e^{1-p} \big\| [\![\nabla w]\!] \big\|_{L^p(e \cap \bar{B}_1)}^p \Big)^{\frac{1}{p}} \Big(\sum_{e \in \mathcal{E}_h^I} h_e \|v_h\|_{L^{p'}(e \cap \bar{B}_1)}^{p'} \Big)^{\frac{1}{p'}} \\ &\lesssim \Big(\|D_h^2 w\|_{L^p(B_1)} + \Big(\sum_{e \in \mathcal{E}_h^I} h_e^{1-p} \big\| [\![\nabla w]\!] \big\|_{L^p(e)}^p \Big)^{\frac{1}{p}} \Big) \|v_h\|_{L^{p'}(B_1)} \\ &\lesssim \|w\|_{W_h^{2,p}(B_1)} \|v_h\|_{L^{p'}(B_1)}.\end{aligned}$$

The desired inequality now follows from the definition of $\|\cdot\|_{L_h^p(B_1)}$. $\square$

Lemma 3.4. *Let h_1 be as in Lemma* 3.2*. Then there holds for $h \le h_1$*

$$\|w_h\|_{W_h^{2,p}(\Omega)} \lesssim \|\mathcal{L}_h w_h\|_{L^p(\Omega)} + \|w_h\|_{L^p(\Omega)} \qquad \forall w_h \in V_h. \tag{3.13}$$

Proof. We divide the proof into two steps.

Step 1*.* For any $x_0 \in \Omega$, let R_1 and h_1 be as in Lemma 3.2, let $R_2 = 2R_1$, $R_3 = 3R_1$, and set $B_i = B_{R_i}(x_0)$ for $i = 0, 1, 2$. Let $\eta \in C^3(\Omega)$ be a cut-off function satisfying

(3.14)
$$0 \le \eta \le 1, \quad \eta\big|_{B_1} = 1, \quad \eta\big|_{\Omega \setminus B_2} = 0, \quad \|\eta\|_{W^{m,\infty}(\Omega)} = O(d^{-m}), \quad m = 0, 1, 2.$$

We first note that $\eta w_h \in W_h^{(p)}(B_2)$ and $I_h(\eta w_h) \in V_h(B_3)$ for any $w_h \in V_h$. Therefore, by Lemmas 2.4 (with $d = R_1$) and 3.2, we have

$$\begin{aligned}\|w_h\|_{W_h^{2,p}(B_1)} &= \|\eta w_h\|_{W_h^{2,p}(B_1)} \le \|\eta w_h - I_h(\eta w_h)\|_{W_h^{2,p}(B_1)} + \|I_h(\eta w_h)\|_{W_h^{2,p}(B_1)} \\ &\lesssim \frac{1}{R_1^2}\|w_h\|_{W^{1,p}(B_2)} + \|I_h(\eta w_h)\|_{W_h^{2,p}(B_1)} \\ &\lesssim \frac{1}{R_1^2}\|w_h\|_{W^{1,p}(B_2)} + \|\mathcal{L}_h(I_h(\eta w_h))\|_{L_h^p(B_2)} \\ &\lesssim \frac{1}{R_1^2}\|w_h\|_{W^{1,p}(B_2)} + \|\mathcal{L}_h(\eta w_h)\|_{L_h^p(B_2)} \\ &\quad + \|\mathcal{L}_h(\eta w_h - I_h(\eta w_h))\|_{L_h^p(B_2)}.\end{aligned}$$

Applying Lemmas 3.3 and 2.4, we obtain

$$\begin{aligned}\|w_h\|_{W_h^{2,p}(B_1)} &\lesssim \frac{1}{R_1^2}\|w_h\|_{W^{1,p}(B_2)} + \|\mathcal{L}_h(\eta w_h)\|_{L_h^p(B_2)} \\ &\qquad + \|\eta w_h - I_h(\eta w_h)\|_{W_h^{2,p}(B_2)} \\ &\lesssim \frac{1}{R_1^2}\|w_h\|_{W^{1,p}(B_3)} + \|\mathcal{L}_h(\eta w_h)\|_{L_h^p(B_3)}.\end{aligned} \tag{3.15}$$

To derive an upper bound of the last term in (3.15), we write for $v_h \in V_h(B_3)$,

$$\begin{aligned}\big(\mathcal{L}_h(\eta w_h), v_h\big) &= -\int_{B_3} A : D_h^2(\eta w_h) v_h\,dx + \sum_{e\in\mathcal{E}_h^I}\int_{e\cap\bar{B}_3} [\![A\nabla(\eta w_h)]\!] v_h\,ds \\ &= -\int_{B_3}\big(\eta A : D_h^2 w_h + 2A\nabla\eta\cdot\nabla w_h + w_h A : D_h^2\eta\big) v_h\,dx \\ &\quad + \sum_{e\in\mathcal{E}_h^I}\int_{e\cap\bar{B}_3} [\![A\nabla w_h]\!]\eta v_h\,ds \\ &= \big(\mathcal{L}_h w_h, I_h(\eta v_h)\big) - \int_{B_3}\big(2A\nabla\eta\cdot\nabla w_h + w_h A : D_h^2\eta\big) v_h\,dx \\ &\quad - \int_{B_3}(A : D_h^2 w_h)(\eta v_h - I_h(\eta v_h))\,dx + \sum_{e\in\mathcal{E}_h^I}\int_{e\cap\bar{B}_3}[\![A\nabla w_h]\!](\eta v_h - I_h(\eta v_h))\,ds.\end{aligned}$$

By Hölder's inequality, Lemmas 2.1–2.2, 2.4, and (3.14) we obtain

$$\begin{aligned}\big(\mathcal{L}_h(\eta w_h), v_h\big) &\lesssim \|\mathcal{L}_h w_h\|_{L_h^p(B_3)}\|I_h(\eta v_h)\|_{L^{p'}(B_3)} + R_1^{-2}\|w_h\|_{W^{1,p}(B_3)}\|v_h\|_{L^{p'}(B_3)} \\ &\quad + \|w_h\|_{W_h^{2,p}(B_3)}\Big(\|\eta v_h - I_h(\eta v_h)\|_{L^{p'}(B_3)} + h\|\nabla(\eta v_h - I_h(\eta v_h))\|_{L^{p'}(B_3)}\Big) \\ &\lesssim \Big(\|\mathcal{L}_h w_h\|_{L_h^p(B_3)} + \frac{1}{R_1^2}\|w_h\|_{W^{1,p}(B_3)}\Big)\|v_h\|_{L^{p'}(B_3)},\end{aligned}$$

which implies that

$$\|\mathcal{L}_h(\eta w_h)\|_{L_h^p(B_3)} \lesssim \|\mathcal{L}_h w_h\|_{L_h^p(B_3)} + \frac{1}{R_1^2}\|w_h\|_{W^{1,p}(B_3)}.$$

Applying this upper bound to (3.15) yields

$$\|w_h\|_{W_h^{2,p}(B_1)} \lesssim \|\mathcal{L}_h w_h\|_{L_h^p(B_3)} + \frac{1}{R_1^2}\|w_h\|_{W^{1,p}(B_3)} \qquad \forall w_h \in V_h. \tag{3.16}$$

Step 2. We now use a covering argument to obtain the global estimate (3.13). To this end, let $\{x_j\}_{j=1}^N \subset \Omega$ with $N = O(R_1^{-n})$ sufficiently large (but independent of h) such that $\overline{\Omega} = \bigcup_{j=1}^N \overline{B}_{R_1}(x_j)$. Setting $S_j = B_{R_1}(x_j)$ and $\tilde{S}_j = B_{R_2}(x_j) = B_{2R_1}(x_j)$, we have by (3.16)

$$\begin{aligned}\|w_h\|^p_{W_h^{2,p}(\Omega)} &\le \sum_{j=1}^N \|w_h\|^p_{W_h^{2,p}(S_j)} \lesssim \sum_{j=1}^N \Big(\|\mathcal{L}_h w_h\|^p_{L_h^p(\tilde{S}_j)} + \frac{1}{R_1^{2p}} \|w_h\|^p_{W^{1,p}(\tilde{S}_j)} \Big) \\ &\lesssim \frac{1}{R_1^{2p}} \|w_h\|^p_{W^{1,p}(\Omega)} + \sum_{j=1}^N \|\mathcal{L}_h w_h\|^p_{L_h^p(\tilde{S}_j)}.\end{aligned}$$

Since $V_h(\tilde{S}_j) \subseteq V_h$, we have

$$\begin{aligned}\sum_{j=1}^N \|\mathcal{L}_h w_h\|^p_{L_h^p(\tilde{S}_j)} &= \sum_{j=1}^N \bigg| \sup_{0\ne v_h\in V_h(\tilde{S}_j)} \frac{\big(\mathcal{L}_h w_h, v_h\big)}{\|v_h\|_{L^{p'}(\tilde{S}_j)}} \bigg|^p \\ &= \sum_{j=1}^N \bigg| \sup_{0\ne v_h\in V_h(\tilde{S}_j)} \frac{\big(\mathcal{L}_h w_h, v_h\big)}{\|v_h\|_{L^{p'}(\Omega)}} \bigg|^p \\ &\le N \bigg| \sup_{0\ne v_h\in V_h} \frac{\big(\mathcal{L}_h w_h, v_h\big)}{\|v_h\|_{L^{p'}(\Omega)}} \bigg|^p \lesssim \frac{1}{R_1^n} \|\mathcal{L}_h w_h\|^p_{L_h^p(\Omega)}.\end{aligned}$$

Consequently, since R_1 is independent of h, we have

$$\|w_h\|_{W_h^{2,p}(\Omega)} \lesssim \frac{1}{R_1^{\frac{n}{p}}} \|\mathcal{L}_h w_h\|_{L_h^p(\Omega)} + \frac{1}{R_1^2} \|w_h\|_{W^{1,p}(\Omega)} \lesssim \|\mathcal{L}_h w_h\|_{L_h^p(\Omega)} + \|w_h\|_{W^{1,p}(\Omega)}.$$

Finally, an application of Lemma 2.5 yields

$$\|w_h\|_{W_h^{2,p}(\Omega)} \lesssim \|\mathcal{L}_h w_h\|_{L_h^p(\Omega)} + \|w_h\|^{\frac{1}{2}}_{L^p(\Omega)} \|w_h\|^{\frac{1}{2}}_{W_h^{2,p}(\Omega)}.$$

Applying the Cauchy–Schwarz inequality to the last term completes the proof. $\square$

Using arguments analogous to those in Lemma 3.4, we also have the following stability estimate for the formal adjoint operator. Due to its length and technical nature, we give the proof in the appendix.

Lemma 3.5. *There exists an $h_2 > 0$ such that*

$$\|v_h\|_{L^{p'}(\Omega)} \lesssim \sup_{0\ne w_h\in V_h} \frac{(\mathcal{L}_h w_h, v_h)}{\|w_h\|_{W_h^{2,p}(\Omega)}} \qquad \forall v_h \in V_h, \tag{3.17}$$

provided $h \le h_* := \min\{h_1, h_2\}$ and $k \ge 2$.

Remark 3.2. Denote by $\mathcal{L}_h^*$ the formal adjoint operator of $\mathcal{L}_h$. Then inequality (3.17) is equivalent to the stability estimate

$$\|v_h\|_{L^{p'}(\Omega)} \lesssim \sup_{0\ne w_h\in V_h} \frac{(\mathcal{L}_h^* v_h, w_h)}{\|w_h\|_{W_h^{2,p}(\Omega)}} \quad \forall v_h \in V_h. \tag{3.18}$$

Thus, the adjoint operator $\mathcal{L}_h^*$ is injective on V_h. Since V_h is finite dimensional, $\mathcal{L}_h^*$ on V_h is an isomorphism. This implies that $\mathcal{L}_h$ is also an isomorphism on V_h; the stability of the operator is addressed in the next theorem, the main result of this section.

Theorem 3.1. *Suppose that* $h \le \min\{h_1, h_2\}$*, and* $k \ge 2$*. Then there holds the following stability estimate:*

$$\|w_h\|_{W_h^{2,p}(\Omega)} \lesssim \|\mathcal{L}_h w_h\|_{L_h^p(\Omega)} \qquad \forall w_h \in V_h. \tag{3.19}$$

Consequently, there exists a unique solution to (3.4) *satisfying*

$$\|u_h\|_{W_h^{2,p}(\Omega)} \lesssim \|f\|_{L^p(\Omega)}. \tag{3.20}$$

Proof. For a given $w_h \in V_h$, Lemma 3.5 guarantees the existence of a unique $\psi_h \in V_h$ satisfying

$$(\mathcal{L}_h v_h, \psi_h) = \int_\Omega w_h |w_h|^{p-2} v_h \, dx \qquad \forall v_h \in V_h. \tag{3.21}$$

By (3.17) we have

$$\|\psi_h\|_{L^{p'}(\Omega)} \lesssim \sup_{0 \ne v_h \in V_h} \frac{(\mathcal{L}_h v_h, \psi_h)}{\|v_h\|_{W_h^{2,p}(\Omega)}} = \sup_{0 \ne v_h \in V_h} \frac{\int_\Omega w_h |w_h|^{p-2} v_h \, dx}{\|v_h\|_{W_h^{2,p}(\Omega)}} \lesssim \|w_h\|_{L^p(\Omega)}^{p-1}.$$

The last inequality is an easy consequence of Hölder's inequality, Lemma 2.5 and the Poincaré–Friedrichs inequality. Taking $v_h = w_h$ in (3.21), we have

$$\|w_h\|_{L^p(\Omega)}^p = (\mathcal{L}_h w_h, \psi_h) \le \|\mathcal{L}_h w_h\|_{L_h^p(\Omega)} \|\psi_h\|_{L^{p'}(\Omega)} \le \|\mathcal{L}_h w_h\|_{L_h^p(\Omega)} \|w_h\|_{L^p(\Omega)}^{p-1},$$

and therefore

$$\|w_h\|_{L^p(\Omega)} \lesssim \|\mathcal{L}_h w_h\|_{L_h^p(\Omega)}.$$

Applying this estimate in (3.4) proves (3.19).

Finally, to show existence and uniqueness of the finite element method (3.4) it suffices to show the estimate (3.20). This immediately follows from (3.19) and Hölder's inequality:

$$\begin{aligned} \|u_h\|_{W_h^{2,p}(\Omega)} \lesssim \|\mathcal{L}_h u_h\|_{L_h^p(\Omega)} &= \sup_{0 \ne v_h \in V_h} \frac{(\mathcal{L}_h u_h, v_h)}{\|v_h\|_{L^{p'}(\Omega)}} \\ &= \sup_{0 \ne v_h \in V_h} \frac{\int_\Omega f v_h \, dx}{\|v_h\|_{L^{p'}(\Omega)}} \le \|f\|_{L^p(\Omega)}. \end{aligned}$$

□

Remark 3.3. The stability result given in Theorem 3.1 requires the mesh condition $h \le \min\{h_1, h_2\}$, where the constants h_1, h_2 depend on the modulus of continuity of A. However, numerical experiments (cf. Section 4) suggest that this mesh condition as well as the polynomial condition $k \ge 2$ can be relaxed.

3.3. Convergence analysis.

The stability estimate in Theorem 3.1 immediately gives us the following error estimate in the $W_h^{2,p}$ semi-norm.

Theorem 3.2. *Assume that the hypotheses of Theorem* 3.1 *are satisfied. Let* $u \in W^{2,p}(\Omega)$ *and* $u_h \in V_h$ *denote the solution to* (1.1) *and* (3.4)*, respectively. Then there holds*

$$\|u - u_h\|_{W_h^{2,p}(\Omega)} \lesssim \inf_{w_h \in V_h} \|u - w_h\|_{W_h^{2,p}(\Omega)}. \tag{3.22}$$

Consequently, if $u \in W^{s,p}(\Omega)$*, for some* $s \ge 2$*, there holds*

$$\|u - u_h\|_{W_h^{2,p}(\Omega)} \lesssim h^{\ell-2} \|u\|_{W^{\ell,p}(\Omega)},$$

where $\ell = \min\{s, k+1\}$*.*

Proof. By Theorem 3.1 and the consistency of the method, we have $\forall v_h \in V_h$

$$\begin{aligned}\|u_h - w_h\|_{W_h^{2,p}(\Omega)} &\lesssim \|\mathcal{L}_h(u_h - w_h)\|_{L_h^p(\Omega)} = \sup_{0\neq v_h\in V_h} \frac{(\mathcal{L}_h(u_h - w_h), v_h)}{\|v_h\|_{L^{p'}(\Omega)}} \\ &= \sup_{0\neq v_h\in V_h} \frac{a_h(u_h - w_h, v_h)}{\|v_h\|_{L^{p'}(\Omega)}} = \sup_{0\neq v_h\in V_h} \frac{a_h(u - w_h, v_h)}{\|v_h\|_{L^{p'}(\Omega)}} \\ &= \sup_{0\neq v_h\in V_h} \frac{(\mathcal{L}_h(u - w_h), v_h)}{\|v_h\|_{L^{p'}(\Omega)}} \lesssim \|u - w_h\|_{W_h^{2,p}(\Omega)}.\end{aligned}$$

Applying the triangle inequality yields (3.22). □

4. Numerical experiments

In this section we present several numerical experiments to show the efficacy of the finite element method, as well as to validate the convergence theory. In addition, we perform numerical experiments where the coefficient matrix is not continuous and/or degenerate. While these situations violate some of the assumptions given in Section 2.1, the tests show that the finite element method is effective for these cases as well.

Test 1: Hölder continuous coefficients and smooth solution. In this test we take $\Omega = (-0.5, 0.5)^2$, the coefficient matrix to be

$$A = \begin{pmatrix} |x|^{1/2} + 1 & -|x|^{1/2} \\ -|x|^{1/2} & 5|x|^{1/2} + 1 \end{pmatrix}$$

and choose f such that $u = \sin(2\pi x_1)\sin(\pi x_2)\exp(x_1\cos(x_2))$ as the exact solution.

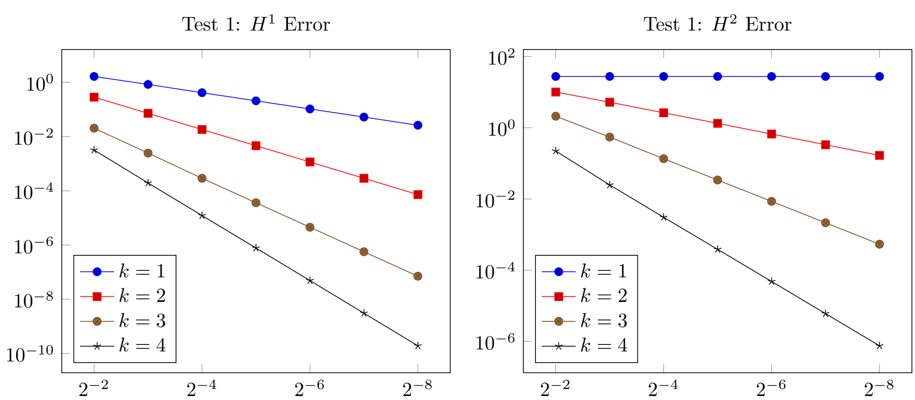

Figure 2. The H^1 (left) and piecewise H^2 (right) errors for Test Problem 1 with polynomial degree $k = 1, 2, 3, 4$. The figures show that the H^1 error converges with order $\mathcal{O}(h^k)$, and the piecewise H^2 error converges with order $\mathcal{O}(h^{k-1})$.

The resulting H^1 and piecewise H^2 errors for various values of polynomial degree k and discretization parameter h are depicted in Figure 2. The figure clearly indicates that the errors have the following behavior:

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

The second estimate is in agreement with Theorem 3.2. In addition, the numerical experiments suggest that (i) the method converges with optimal order in the H^1-norm and (ii) the method is convergent in the piecewise linear case ($k = 1$).

Test 2: Uniformly continuous coefficients and $W^{2,p}$ solution. For the second set of numerical experiments, we take the domain to be the square $\Omega = (0, 1/2)^2$, and take the coefficient matrix to be

$$A = \begin{pmatrix} -\dfrac{5}{\log(|x|)} + 15 & 1 \\ 1 & -\dfrac{1}{\log(|x|)} + 3 \end{pmatrix}.$$

We choose the data such that the exact solution is given by $u = |x|^{7/4}$. We note that $u \in W^{m,p}(\Omega)$ for $(7 - 4m)p > -8$. In particular, $u \in W^{2,p}(\Omega)$ for $p < 8$ and $u \in W^{3,p}(\Omega)$ for $p < 8/5$.

In order to apply Theorem 3.2 to this test problem, we recall that the kth degree nodal interpolant of u with $k \geq 2$ satisfies

$$\|D_h^2(u - I_h u)\|_{L^2(\Omega)} \leq C h^{2-2/p} \|u\|_{W^{3,p}(\Omega)}$$

for $p < 2$. Since $u \in W^{3,p}(\Omega)$ for $p < 8/5$, Theorem 3.2 then predicts the convergence rate

$$\|D_h^2(u - u_h)\|_{L^2(\Omega)} \leq C\|D_h^2(u - I_h u)\|_{L^2(\Omega)} = \mathcal{O}(h^{3/4-\varepsilon})$$

for any $\varepsilon > 0$. Note that a slight modification of these arguments also shows that $|u - I_h u|_{H^1(\Omega)} = \mathcal{O}(h^{7/4-\varepsilon})$.

The errors of the finite element method for Test 2 using piecewise linear, quadratic and cubic polynomials are depicted in Figure 3. As predicted by the theory, the H^2 error converges with order $\approx \mathcal{O}(h^{3/4})$ if the polynomial degree is greater than or equal to two. Similar to the first test problem, the numerical experiments also show that the H^1 error converges with optimal order, i.e., $|u - u_h|_{H^1(\Omega)} = \mathcal{O}(h^{7/4-\varepsilon})$.

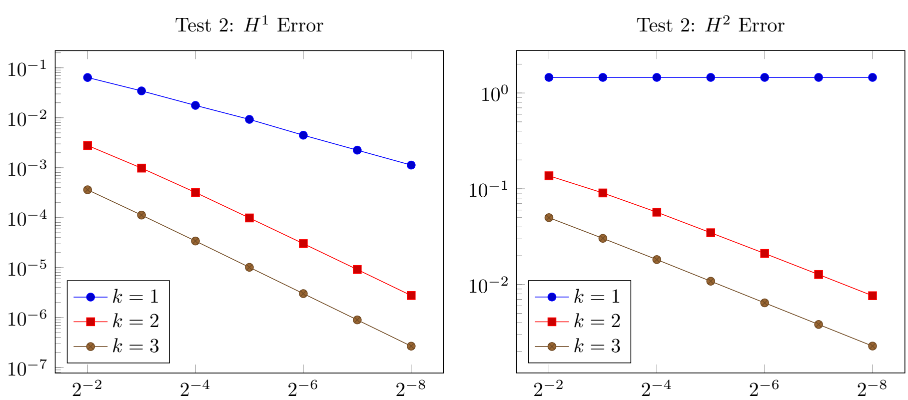

FIGURE 3. The H^1 (left) and piecewise H^2 (right) errors for Test Problem 2 with polynomial degree $k = 1, 2, 3$. The figures show that the H^1 error converges with order $\mathcal{O}(h^{\min\{k,7/4-\varepsilon\}})$, where as the piecewise H^2 error converges with order $\mathcal{O}(h^{\min\{k,7/4-\varepsilon\}-1})$.

Test 3: Degenerate coefficients and $W^{2,p}$ solution. For the third set of test problems, we take $\Omega = (0,1)^2$,

$$A = \frac{16}{9}\begin{pmatrix} x_1^{2/3} & -x_1^{1/3}x_2^{1/3} \\ -x_1^{1/3}x_2^{1/3} & x_2^{2/3} \end{pmatrix},$$

and exact solution $u = x_1^{4/3} - x_2^{4/3}$. We remark that the choice of the matrix and solution is motivated by Aronson's example for the infinity-Laplace equation. In particular, the function u satisfies the quasi-linear PDE $\Delta_\infty u = 0$, where $\Delta_\infty u := (D^2 u \nabla u)\cdot \nabla u = (D^2 u) : (\nabla u (\nabla u)^T)$. Noting that $A = \nabla u(\nabla u)^T$, we see that $-A : D^2 u = 0 =: f$.

Unlike the first two test problems, the matrix is not uniformly elliptic, as $\det(A(x)) = 0$ for all $x \in \Omega$. Therefore, the theory given in the previous sections does not apply. We also note that the exact solution satisfies the regularity $u \in W^{m,p}(\Omega)$ for $(4-3m)p > -1$, and therefore $u \in W^{2,p}(\Omega)\cap W^{1,\infty}(\Omega)$ for $p < 3/2$.

The resulting errors of the finite element method using piecewise linear and quadratic polynomials are plotted in Figure 4. While this problem is outside the scope of the theory, the experiments show that the method converges, and the following rates are observed:

$$\|u - u_h\|_{L^2(\Omega)} = \mathcal{O}(h^{4/3}), \qquad |u - u_h|_{H^1(\Omega)} = \mathcal{O}(h^{5/6}).$$

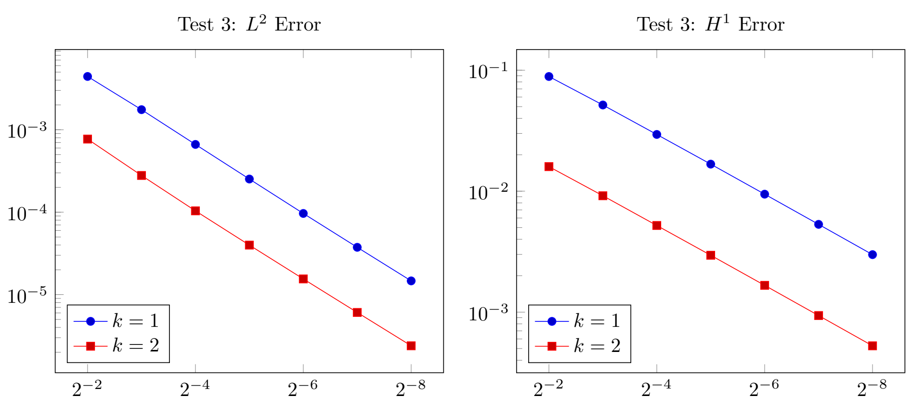

FIGURE 4. The L^2 (left) and H^1 (right) errors for the degenerate Test Problem 3 with polynomial degree $k = 1$ and $k = 2$. The figures show that the L^2 error converges with order $\approx \mathcal{O}(h^{4/3})$ and the H^1 error converges with order $\approx \mathcal{O}(h^{5/6})$.

Test 4: Discontinuous coefficients and smooth solution. In the final set of test problems, we set the domain to be $\Omega = (-1,1)^2$, the coefficient matrix as (cf. Figure 6)

$$A = \begin{pmatrix} 2 & \sin\left(\pi(20x_1x_2 + 1/2)\right)\frac{x_1x_2}{|x_1||x_2|} \\ \sin\left(\pi(20x_1x_2 + 1/2)\right)\frac{x_1x_2}{|x_1||x_2|} & 2 \end{pmatrix},$$

and choose the source function such that the exact solution is

$$u = \frac{\sin(2\pi(x_1^2 + x_2))\sin(5\pi x_2)}{2 + x_1^2\cos(2\pi x_2)}.$$

The H^1 errors and piecewise H^2 errors are plotted in Figure 5 for polynomial degrees $k = 1, 2, 3$. Again, while the theory given in the previous section does not include the case of discontinuous coefficients, the numerical experiments suggest that the method is stable and convergence rates of order $\mathcal{O}(h^k)$ and $\mathcal{O}(h^{k-1})$ are observed in the H^1- and H^2-norms, respectively.

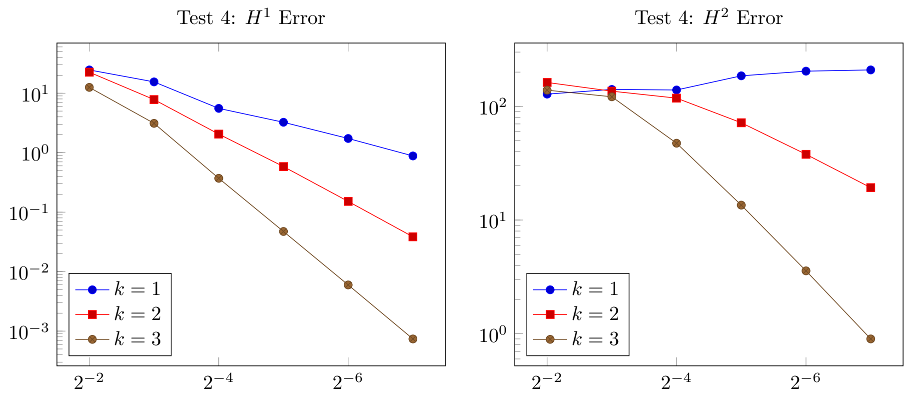

FIGURE 5. The H^1 (left) and piecewise H^2 (right) errors for Test Problem 4 with discontinuous coefficient matrix A. As $h \to 0^+$, the H^1 error converges with order $\mathcal{O}(h^k)$, and the H^2 error converges with order $\mathcal{O}(h^{k-1})$.

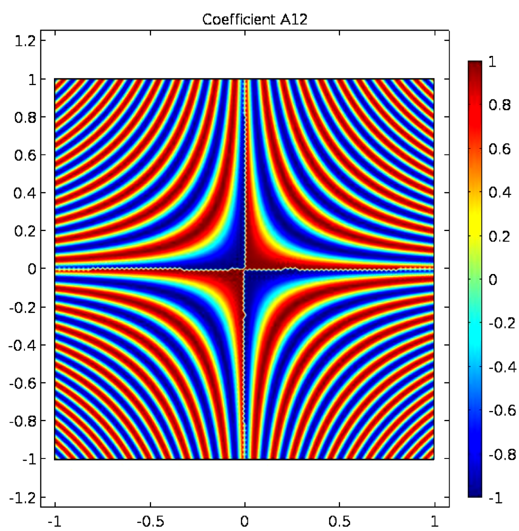

FIGURE 6. The graph of the off-diagonal entries of the coefficient matrix A in Test 4.

Appendix A. Super approximation result

Here, we provide the proof of Lemma 2.4. As a first step, we use standard interpolation estimates [4, 7] to obtain for $0 \le m \le k+1$,

$$h^{mp}\|\eta v_h - I_h(\eta v_h)\|^p_{W^{m,p}(T)} \lesssim h^{p(k+1)}|\eta v_h|^p_{W^{k+1,p}(T)}. \tag{A.1}$$

Since $|\eta|_{W^{j,\infty}(T)} \lesssim d^{-j}$ and $|v_h|_{H^{k+1}(T)} = 0$, we find

$$\begin{aligned}|\eta v_h|_{W^{k+1,p}(T)} &\lesssim \sum_{|\alpha|+|\beta|=k+1} \int_T |D^\alpha \eta|^p |D^\beta v_h|^p\,dx \\ &\lesssim \sum_{j=0}^{k} \frac{1}{d^{p(k+1-j)}}|v_h|^p_{W^{j,p}(T)} \lesssim \sum_{j=0}^{k} \frac{h^{-jp}}{d^{p(k+1-j)}}\|v_h\|^p_{L^p(T)},\end{aligned} \tag{A.2}$$

where an inverse estimate was applied to derive the last inequality. Combining (A.2) with (A.1) and using the hypothesis $h \le d$ then gives us

$$h^{mp}\|\eta v_h - I_h(\eta v_h)\|^p_{W^{m,p}(T)} \lesssim \sum_{j=0}^{k} \frac{h^{p(k+1-j)}}{d^{p(k+1-j)}}\|v_h\|^p_{L^p(T)} \lesssim \frac{h^p}{d^p}\|v_h\|^p_{L^p(T)}.$$

Therefore for $m \in \{0,1\}$ we have

$$\begin{aligned}h^{mp}\|\eta v_h - I_h(\eta v_h)\|^p_{W^{m,p}(D)} &\le \sum_{\substack{T\in\mathcal{T}_h\\ T\cap D\neq\emptyset}} h^{mp}\|\eta v_h - I_h(\eta v_h)\|^p_{W^{m,p}(T)} \\ &\lesssim \sum_{\substack{T\in\mathcal{T}_h\\ T\cap D\neq\emptyset}} \frac{h^p}{d^p}\|v_h\|^p_{L^p(T)} \le \frac{h^p}{d^p}\|v_h\|^p_{L^p(D_h)}.\end{aligned}$$

Thus, (2.14) is satisfied.

To obtain the second estimate (2.15), we first use (A.1), (A.2) as an inverse estimate to get

$$\begin{aligned}\|D^2(\eta v_h - I_h(\eta v_h))\|^p_{L^p(T)} &\lesssim h^{p(k-1)}|\eta v_h|^p_{W^{k+1,p}(T)} \lesssim \sum_{j=0}^{k} \frac{h^{p(k-1)}}{d^{p(k+1-j)}}|v_h|^p_{W^{j,p}(T)} \\ &\lesssim \frac{1}{d^{2p}}\|v_h\|^p_{L^p(T)} + \sum_{j=1}^{k} \frac{h^{k-j}}{d^{p(k+1-j)}}\|v_h\|^p_{W^{1,p}(T)} \\ &\lesssim \frac{1}{d^{2p}}\|v_h\|^p_{L^p(T)} + \frac{1}{d^p}\|v_h\|^p_{W^{1,p}(T)} \lesssim \frac{1}{d^{2p}}\|v_h\|^p_{W^{1,p}(T)}.\end{aligned} \tag{A.3}$$

By similar arguments we find

$$h^{-p}\|\nabla(\eta v_h - I_h(\eta v_h))\|^p_{L^p(T)} \lesssim h^{p(k-1)}|\eta v_h|^p_{W^{k+1,p}(T)} \lesssim \frac{1}{d^{2p}}\|v_h\|^p_{W^{1,p}(T)}. \tag{A.4}$$

Therefore by Lemma 2.9 and (A.3)–(A.4), we obtain

$$\begin{aligned}\|\eta v_h - I_h(\eta v_h)\|^p_{W^{2,p}_h(D)} &\le \sum_{\substack{T\in\mathcal{T}_h\\ T\cap D\neq\emptyset}} \|D^2(\eta v_h - I_h(\eta v_h))\|^p_{L^p(T)} \\ &\quad + \sum_{e\in\mathcal{E}_h^I} h_e^{1-p}\big\| [\![\nabla(\eta v_h - I_h(\eta v_h))]\!] \big\|^p_{L^p(e\cap\bar{D})} \\ &\lesssim \sum_{\substack{T\in\mathcal{T}_h\\ T\cap D\neq\emptyset}} \Big[\|D^2(\eta v_h - I_h(\eta v_h))\|^p_{L^p(T)} + h^{-p}\|\nabla(\eta v_h - I_h(\eta v_h))\|^p_{L^p(T)}\Big] \\ &\lesssim \sum_{\substack{T\in\mathcal{T}_h\\ T\cap D\neq\emptyset}} \frac{1}{d^{2p}}\|v_h\|^p_{W^{1,p}(T)} \le \frac{1}{d^{2p}}\|v_h\|^p_{W^{1,p}(D_h)}.\end{aligned}$$

Taking the pth root of this last expression yields the estimate (2.15). The proof of (2.16) uses the exact same arguments and is therefore omitted.

Appendix B. Proof of Lemma 3.5

To prove Lemma 3.5 we introduce the discrete $W^{-2,p}$-type norm

$$\|r\|_{W_h^{-2,p}(D)} := \sup_{0\neq v_h\in V_h(D)} \frac{(r, v_h)_D}{\|v_h\|_{W^{2,p'}(D)}}, \tag{B.1}$$

and the $W^{-1,p}$-norm (defined for L^p functions)

$$\|r\|_{W^{-1,p}(D)} = \sup_{0\neq v\in W^{1,p'}(D)} \frac{(r,v)_D}{\|v\|_{W^{1,p'}(D)}} = \sup_{\substack{v\in W^{1,p'}\\ \|v\|_{W^{1,p'}(D)}=1}} (r,v)_D\,dx. \tag{B.2}$$

The desired estimate (3.17) is then equivalent to

$$\|v_h\|_{L^{p'}(\Omega)} \lesssim \|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(\Omega)} \qquad \forall v_h\in V_h, \tag{B.3}$$

where we recall that $\mathcal{L}_h^*$ is the adjoint operator of $\mathcal{L}_h$. Due to its length, we break up the proof of (B.3) into three steps.

Step 1 (A local estimate). The first step in the derivation of (3.17) (equivalently, (B.3)) is to prove a local version of this estimate, analogous to Lemma 3.2. To this end, for fixed $x_0\in\Omega$, let δ_0, R_{δ_0}, $R_1 := \frac{1}{3}R_{\delta_0}$ and $B_1 := B_{R_1}(x_0)$ be as in Lemmas 3.1–3.2, with $\delta_0>0$ to be determined. For a fixed $v_h\in V_h(B_1)$, let $\varphi\in W^{2,p}(\Omega)\cap W_0^{1,p}(\Omega)$ satisfy $\mathcal{L}\varphi = v_h|v_h|^{p'-2}$ in Ω with

$$\|\varphi\|_{W^{2,p}(\Omega)} \lesssim \||v_h|^{p'-1}\|_{L^p(\Omega)} \lesssim \|v_h\|^{p'-1}_{L^{p'}(B_1)}. \tag{B.4}$$

Multiplying the PDE by v_h, integrating over Ω, and using the consistency of $\mathcal{L}_h$ yields

$$\|v_h\|^{p'}_{L^{p'}(B_1)} = \|v_h\|^{p'}_{L^{p'}(\Omega)} = (\mathcal{L}\varphi, v_h) = (\mathcal{L}_h\varphi, v_h).$$

Therefore, for any $\varphi_h\in V_h$, there holds

$$\begin{aligned}\|v_h\|^{p'}_{L^{p'}(B_1)} &= (\mathcal{L}_h\varphi_h, v_h) + (\mathcal{L}_h(\varphi-\varphi_h), v_h) \\ &= (\mathcal{L}_h^* v_h, \varphi_h) + (\mathcal{L}_{0,h}(\varphi-\varphi_h), v_h) + ((\mathcal{L}_h - \mathcal{L}_{0,h})(\varphi-\varphi_h), v_h),\end{aligned} \tag{B.5}$$

where $\mathcal{L}_{0,h}$ is given by (2.24) with $A_0 \equiv A(x_0)$. Now take $\varphi_h \in V_h$ to be the elliptic projection of φ with respect to $\mathcal{L}_{0,h}$, i.e.,

$$(\mathcal{L}_{0,h}(\varphi - \varphi_h), w_h) = 0 \qquad \forall w_h \in V_h.$$

Lemma 2.6 ensures that φ_h is well defined and satisfies the estimate

$$\|\varphi_h\|_{W_h^{2,p}(\Omega)} \lesssim \|\mathcal{L}_{0,h}\varphi_h\|_{L_h^p(\Omega)} = \|\mathcal{L}_{0,h}\varphi\|_{L_h^p(\Omega)} \lesssim \|\varphi\|_{W^{2,p}(\Omega)} \lesssim \|v_h\|_{L^{p'}(B_1)}^{p'-1}. \tag{B.6}$$

Combining Lemma 3.1, (B.4)–(B.6) and (B.1), we have

$$\begin{aligned}
\|v_h\|_{L^{p'}(B_1)}^{p'} &= (\mathcal{L}_h^* v_h, \varphi_h) + \big((\mathcal{L}_h - \mathcal{L}_{0,h})(\varphi - \varphi_h), v_h\big) \\
&\le \|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(\Omega)} \|\varphi_h\|_{W_h^{2,p}(\Omega)} + \|(\mathcal{L}_h - \mathcal{L}_{0,h})(\varphi - \varphi_h)\|_{L_h^p(B_1)} \|v_h\|_{L^{p'}(B_1)} \\
&\le \|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(B_1)} \|v_h\|_{L^{p'}(B_1)}^{p'-1} + \delta_0 \|\varphi - \varphi_h\|_{W_h^{2,p}(B_1)} \|v_h\|_{L^{p'}(B_1)} \\
&\le \|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(B_1)} \|v_h\|_{L^{p'}(B_1)}^{p'-1} + \delta_0 \|v_h\|_{L^{p'}(B_1)}^{p'}.
\end{aligned}$$

Taking δ_0 sufficiently small and rearranging terms gives the local stability estimate for finite element functions with compact support:

$$\|v_h\|_{L^{p'}(B_1)} \lesssim \|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(B_1)} \qquad \forall v_h \in V_h(B_1). \tag{B.7}$$

Step 2 (A global Gärding-type inequality). We now follow the proof of Lemma 3.4 to derive a global Gärding-type inequality for the adjoint problem. Let R_1 be given in the first step of the proof, $R_2 = 2R_1$, and $R_3 = 3R_1$. Let $\eta \in C^3(\Omega)$ satisfy the conditions in Lemma 3.4 (cf. (3.14)). By the triangle inequality and (B.7) we have for any $v_h \in V_h$

$$\begin{aligned}
\|v_h\|_{L^{p'}(B_1)} &= \|\eta v_h\|_{L^{p'}(B_1)} \le \|\eta v_h - I_h(\eta v_h)\|_{L^{p'}(B_1)} + \|I_h(\eta v_h)\|_{L^{p'}(B_1)} \\
&\lesssim \|\eta v_h - I_h(\eta v_h)\|_{L^{p'}(B_1)} + \|\mathcal{L}_h^*(I_h(\eta v_h))\|_{W_h^{-2,p'}(B_1)} \\
&\lesssim \|\eta v_h - I_h(\eta v_h)\|_{L^{p'}(B_1)} + \|\mathcal{L}_h^*(I_h(\eta v_h) - \eta v_h)\|_{W_h^{-2,p'}(B_1)} \\
&\quad + \|\mathcal{L}_h^*(\eta v_h)\|_{W_h^{-2,p'}(B_1)}.
\end{aligned}$$

Applying Lemmas 3.3, Lemma 2.4 (with $d = R_1$) and an inverse estimate yields

$$\begin{aligned}
\|v_h\|_{L^{p'}(B_1)} &\lesssim \|\eta v_h - I_h(\eta v_h)\|_{L^{p'}(B_2)} + \|\mathcal{L}_h^*(\eta v_h)\|_{W_h^{-2,p'}(B_1)} \\
&\lesssim \frac{h}{R_1} \|v_h\|_{L^{p'}(B_3)} + \|\mathcal{L}_h^*(\eta v_h)\|_{W_h^{-2,p'}(B_3)} \\
&\lesssim \frac{1}{R_1} \|v_h\|_{W^{-1,p'}(B_3)} + \|\mathcal{L}_h^*(\eta v_h)\|_{W_h^{-2,p'}(B_3)}.
\end{aligned} \tag{B.8}$$

The goal now is to replace $\mathcal{L}_h^*(\eta v_h)$ appearing in the right-hand side of (B.8) by $\mathcal{L}_h^* v_h$ plus low order terms. To this end, we write for $w_h \in V_h(B_3)$ (cf. 2.24),

$$\begin{aligned}
(\mathcal{L}_h^*(\eta v_h), w_h) &= a_h(w_h, \eta v_h) = a_h(w_h \eta, v_h) + \big[a_h(w_h, \eta v_h) - a_h(w_h \eta, v_h)\big] \\
&= a_h(I_h(w_h \eta), v_h) + a_h(w_h \eta - I_h(w_h \eta), v_h) + \big[a_h(w_h, \eta v_h) - a_h(w_h \eta, v_h)\big] \\
&=: I_1 + I_2 + I_3.
\end{aligned} \tag{B.9}$$

To derive an upper bound of I_1, we use (B.1) and properties of the interpolant and cut-off function η:

(B.10)

$$\begin{aligned} I_1 &= (\mathcal{L}_h^* v_h, I_h(\eta w_h)) \le \|\mathcal{L}_h^* v_h\|_{W^{-2,p'}(B_3)} \|I_h(\eta w_h)\|_{W^{2,p}(B_3)} \\ &\lesssim \|\mathcal{L}_h^* v_h\|_{W^{-2,p'}(B_3)} \|\eta w_h\|_{W_h^{2,p}(B_3)} \lesssim \frac{1}{R_1^2} \|\mathcal{L}_h^* v_h\|_{W^{-2,p'}(B_2)} \|w_h\|_{W_h^{2,p}(B_3)}. \end{aligned}$$

Next, we apply Lemmas 3.3, 2.4 and an inverse estimate to bound I_2:

$$\begin{aligned} I_2 &= (\mathcal{L}_h(\eta w_h - I_h(\eta w_h)), v_h) \lesssim \|\eta w_h - I_h(\eta w_h)\|_{W_h^{2,p}(B_3)} \|v_h\|_{L^{p'}(B_3)} \\ &\lesssim \frac{h}{R_1^3} \|w_h\|_{W_h^{2,p}(B_3)} \|v_h\|_{L^{p'}(B_3)} \lesssim \frac{1}{R_1^3} \|w_h\|_{W_h^{2,p}(B_3)} \|v_h\|_{W^{-1,p'}(B_3)}. \end{aligned} \tag{B.11}$$

To estimate I_3, we add and subtract $a_0(w_h, \eta v_h) - a_0(w_h\eta, v_h)$ and expand terms to obtain

(B.12)

$$\begin{aligned} I_3 &= a_0(w_h, \eta v_h) - a_0(w_h \eta, v_h) \\ &\qquad + \big[a_h(w_h, \eta v_h) - a_h(w_h\eta, v_h) - \big(a_0(w_h, \eta v_h) - a_0(w_h\eta, v_h)\big)\big] \\ &= -\int_{B_3} \Big(w_h A_0 : D^2\eta + 2A_0 \nabla\eta \cdot \nabla w_h\Big) v_h\, dx \\ &\qquad - \int_{B_3} \Big(w_h(A - A_0) : D^2\eta + 2(A - A_0)\nabla\eta \cdot \nabla w_h\Big) v_h\, dx =: K_1 + K_2. \end{aligned}$$

Applying Hölder's inequality and Lemmas C.1–D.1 yields

$$\begin{aligned} K_1 &\le \|w_h A_0 : D^2\eta\|_{W^{1,p}(B_3)} \|v_h\|_{W^{-1,p'}(B_3)} + 2\Big|\int_{B_3} (A_0 \nabla\eta \cdot \nabla w_h) v_h\, dx\Big| \\ &\lesssim \big(\frac{1}{R_1^3} \|w_h\|_{W^{1,p}(B_3)} + \frac{1}{R_1^2} \|w_h\|_{W_h^{2,p}(B_3)}\big) \|v_h\|_{W^{-1,p'}(B_3)} \\ &\lesssim \frac{1}{R_1^2} \|w_h\|_{W_h^{2,p}(B_3)} \|v_h\|_{W^{-1,p'}(B_3)}. \end{aligned} \tag{B.13}$$

Similarly, by Lemma C.1 and (3.10), we obtain

$$\begin{aligned} K_2 &\le \|A - A_0\|_{L^\infty(B_3)} \big(\|w_h\|_{L^p(B_3)} \|D^2\eta\|_{L^\infty(\Omega)} \\ &\qquad\qquad + \|\nabla w_h\|_{L^p(B_3)} \|\nabla\eta\|_{L^\infty(\Omega)}\big) \|v_h\|_{L^{p'}(B_3)} \\ &\lesssim \delta_0 \big(R_3^2 \|D^2\eta\|_{L^\infty(\Omega)} + R_3 \|\nabla\eta\|_{L^\infty(\Omega)}\big) \|w_h\|_{W_h^{2,p}(B_3)} \|v_h\|_{L^{p'}(B_3)} \\ &\lesssim \delta_0 \|w_h\|_{W_h^{2,p}(B_3)} \|v_h\|_{L^{p'}(B_3)}. \end{aligned} \tag{B.14}$$

Combining (B.12)–(B.14) results in the following upper bound of I_3:

$$I_3 \lesssim \frac{1}{R_1^2} \|w_h\|_{W_h^{2,p}(B_3)} \|v_h\|_{W^{-1,p'}(B_3)} + \delta_0 \|w_h\|_{W_h^{2,p}(B_3)} \|v_h\|_{L^{p'}(B_3)}. \tag{B.15}$$

Applying the estimates to (B.10)–(B.11), (B.15) to (B.9) results in

$$\begin{aligned} (\mathcal{L}_h^*(\eta v_h), w_h) &\lesssim \frac{1}{R_1^3} \Big(\|\mathcal{L}_h^* v_h\|_{W^{-2,p'}(B_3)} + \|v_h\|_{W^{-1,p'}(B_3)}\Big) \|w_h\|_{W_h^{2,p}(B_3)} \\ &\qquad + \delta_0 \|v_h\|_{L^{p'}(B_3)} \|w_h\|_{W_h^{2,p}(B_3)} \end{aligned}$$

and therefore by (B.1),

(B.16)
$$\|\mathcal{L}_h^*(\eta v_h)\|_{W_h^{-2,p'}(B_3)} \lesssim \frac{1}{R_1^3}\Big(\|\mathcal{L}_h^* v_h\|_{W^{-2,p'}(B_3)} + \|v_h\|_{W^{-1,p'}(B_3)}\Big) + \delta_0\|v_h\|_{L^{p'}(B_3)}.$$

Combining (B.16) and (B.8) yields

$$\|v_h\|_{L^{p'}(B_1)} \lesssim \frac{1}{R_1^3}\Big(\|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(B_3)} + \|v_h\|_{W^{-1,p'}(B_3)}\Big) + \delta_0\|v_h\|_{L^{p'}(B_3)}.$$

Finally, we use the exact same covering argument in the proof of Lemma 3.4 to obtain

$$\|v_h\|_{L^{p'}(\Omega)} \lesssim \|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(\Omega)} + \|v_h\|_{W^{-1,p'}(\Omega)} + \delta_0\|v_h\|_{L^{p'}(\Omega)}.$$

Taking δ_0 sufficiently small and kicking back the last term then yields the Gärding-type estimate

$$\|v_h\|_{L^{p'}(\Omega)} \lesssim \|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(\Omega)} + \|v_h\|_{W^{-1,p'}(\Omega)}. \tag{B.17}$$

Step 3 (A duality argument). In the last step of the proof, we shall combine a duality argument and (B.17) to obtain the desired result (B.3).

Define the set

$$X = \{g \in W_0^{1,p}(\Omega) : \ \|g\|_{W^{1,p}(\Omega)} = 1\}.$$

Since X is precompact in $L^p(\Omega)$, and due to the elliptic regularity estimate $\|\varphi\|_{W^{2,p}(\Omega)} \lesssim \|\mathcal{L}\varphi\|_{L^p(\Omega)}$, the set

$$W = \{\varphi \in W^{2,p} \cap W_0^{1,p}(\Omega) : \ \mathcal{L}\varphi = g, \ \exists g \in X\}$$

is precompact in $W^{2,p}(\Omega)$. Therefore by [23, Lemma 5], for every $\varepsilon > 0$, there exists a $h_2(\varepsilon, W) > 0$ such that for each $\varphi \in W$ and $h \le h_2$ there exists $\varphi_h \in V_h$ satisfying

$$\|\varphi - \varphi_h\|_{W_h^{2,p}(\Omega)} \le \varepsilon \quad \text{for } k \ge 2. \tag{B.18}$$

Note that (B.18) implies $\|\varphi_h\|_{W_h^{2,p}(\Omega)} \le \|\varphi\|_{W^{2,p}(\Omega)} + \varepsilon \lesssim 1$.

For $g \in X$ we shall use $\varphi_g \in W$ to denote the solution to $\mathcal{L}\varphi_g = g$. We then have by Lemma 3.3, for any $v_h \in V_h$ and $\varphi_h \in V_h$,

$$\begin{aligned}\int_\Omega v_h g\, dx &= (\mathcal{L}_h\varphi_g, v_h) = (\mathcal{L}_h^* v_h, \varphi_h) + (\mathcal{L}_h(\varphi_g - \varphi_h), v_h)\\ &\lesssim \|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(\Omega)}\|\varphi_h\|_{W_h^{2,p}(\Omega)} + \|\varphi_g - \varphi_h\|_{W_h^{2,p}(\Omega)}\|v_h\|_{L^{p'}(\Omega)}.\end{aligned}$$

Choosing φ_h so that (B.18) is satisfied (with $\varphi = \varphi_g$) and using the definition of the $W^{-1,p'}$-norm (B.2) results in

$$\|v_h\|_{W^{-1,p'}(\Omega)} \lesssim \|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(\Omega)} + \varepsilon\|v_h\|_{L^{p'}(\Omega)}.$$

Finally we apply this last estimate in (B.17) to obtain

$$\|v_h\|_{L^{p'}(\Omega)} \lesssim \|\mathcal{L}_h^* v_h\|_{W_h^{-2,p'}(\Omega)} + \varepsilon\|v_h\|_{L^{p'}(\Omega)}.$$

Taking ε sufficiently small and kicking back a term to the left-hand side yields (B.3). This completes the proof.

Appendix C. A discrete Poincaré estimate

Lemma C.1. *There holds for any $w_h \in V_h(D)$ with* $\operatorname{diam}(D) \geq h$,

$$\|w_h\|_{W^{m,p}(D)} \lesssim \operatorname{diam}(D)^{2-m} \|w_h\|_{W_h^{2,p}(D)} \quad m = 1, 2.$$

Proof. Denote by $V_{c,h} \subset H^2(\Omega) \cap H_0^1(\Omega)$ the Argyris finite element space [4], and let $E_h : V_h \to V_{c,h}$ be the enriching operator constructed in [5] by averaging. The arguments in [5] and scaling show that, for $w_h \in V_h(D)$,

(C.1)
$$E_h w_h \in H_0^2(D_h), \quad \|w_h - E_h w_h\|_{W^{m,p}(D)} \lesssim h^{2-m} \|w_h\|_{W_h^{2,p}(D)} \; (m = 0, 1, 2),$$

where D_h is given by (2.12). Since $E_h w_h \in H_0^2(D_h)$ and $\operatorname{diam}(D) \geq h$, the usual Poincaré inequality gives

$$\|E_h w_h\|_{W^{m,p}(D_h)} \lesssim \operatorname{diam}(D_h)^{2-m} \|E_h w_h\|_{W^{2,p}(D_h)} \lesssim \operatorname{diam}(D)^{2-m} \|w_h\|_{W_h^{2,p}(D)}.$$

Therefore by adding and subtracting terms, we obtain for $m = 0, 1$,

$$\begin{aligned}
\|w_h\|_{W^{m,p}(D)} &\leq \|E_h w_h\|_{W^{m,p}(D)} + \|w_h - E_h w_h\|_{W^{m,p}(D)} \\
&\lesssim \operatorname{diam}(D)^{2-m} \|w_h\|_{W_h^{2,p}(D)} + h^{2-m} \|w_h\|_{W_h^{2,p}(D)} \\
&\lesssim \operatorname{diam}(D)^{2-m} \|w_h\|_{W_h^{2,p}(D)},
\end{aligned}$$

where again, we have used the assumption $h \leq \operatorname{diam}(D)$. The proof is complete. □

Appendix D. A discrete Hölder inequality

Lemma D.1. *For any smooth function η, and $w_h \in V_h(D), v_h \in V_h$, there holds*

$$\int_D (\nabla \eta \cdot \nabla w_h) v_h \, dx \lesssim \|\eta\|_{W^{2,\infty}(D)} \|w_h\|_{W_h^{2,p}(D)} \|v_h\|_{W^{-1,p'}(D)}.$$

Proof. Let $E_h : V_h \to V_c$ be the enriching operator in Lemma C.1 satisfying (C.1). Since $E_h w_h \in H^2(D)$ we have

$$\begin{aligned}
\int_D (\nabla \eta \cdot \nabla (E_h w_h)) v_h \, dx &\lesssim \|\nabla \eta \cdot \nabla (E_h w_h)\|_{W^{1,p}(D)} \|v_h\|_{W^{-1,p'}(D)} \\
&\lesssim \|\eta\|_{W^{2,\infty}(D)} \|E_h w_h\|_{W^{2,p}(D)} \|v_h\|_{W^{-1,p'}(D)} \\
&\lesssim \|\eta\|_{W^{2,\infty}(D)} \|w_h\|_{W_h^{2,p}(D)} \|v_h\|_{W^{-1,p'}(D)}
\end{aligned}$$

Combining this estimate with the triangle inequality, (C.1), and an inverse estimate gives

$$\begin{aligned}
\int_D (\nabla \eta \cdot \nabla w_h) v_h \, dx &= \int_D (\nabla \eta \cdot \nabla (E_h w_h)) v_h \, dx + \int_D (\nabla \eta \cdot \nabla (w_h - E_h w_h)) v_h \, dx \\
&\lesssim \|\eta\|_{W^{2,\infty}(D)} \big(\|w_h\|_{W_h^{2,p}(D)} \|v_h\|_{W^{-1,p'}(D)} \\
&\qquad + \|w_h - E_h w_h\|_{W^{1,p}(D)} \|v_h\|_{L^{p'}(D)} \big) \\
&\lesssim \|\eta\|_{W^{2,\infty}(D)} \|w_h\|_{W_h^{2,p}(D)} \|v_h\|_{W^{-1,p'}(D)}.
\end{aligned}$$

□

References

[1] J. H. Argyris, I. Fried, and D. W. Scharpf, *The TUBA family of plate elements for the matrix displacement*, Aero. J. Roy. Soc. **72** (1968), 701–709.

[2] I. Babuška, G. Caloz, and J. E. Osborn, *Special finite element methods for a class of second order elliptic problems with rough coefficients*, SIAM J. Numer. Anal. **31** (1994), no. 4, 945–981, DOI 10.1137/0731051. MR1286212

[3] S. Bernstein, *Sur la généralisation du problème de Dirichlet* (French), Math. Ann. **69** (1910), no. 1, 82–136, DOI 10.1007/BF01455154. MR1511579

[4] 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. MR2373954

[5] 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

[6] L. A. Caffarelli and C. E. Gutiérrez, *Properties of the solutions of the linearized Monge-Ampère equation*, Amer. J. Math. **119** (1997), no. 2, 423–465. MR1439555

[7] P. G. Ciarlet, *The Finite Element Method for Elliptic Problems*, North-Holland Publishing Co., Amsterdam-New York-Oxford, 1978. Studies in Mathematics and its Applications, Vol. 4. MR0520174

[8] M. Crouzeix and V. Thomée, *The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces*, Math. Comp. **48** (1987), no. 178, 521–532, DOI 10.2307/2007825. MR878688

[9] A. Dedner and T. Pryer, *Discontinuous Galerkin methods for non variational problems*, arXiv:1304.2265v1 [math.NA].

[10] L. C. Evans, *Partial Differential Equations*, Graduate Studies in Mathematics, Volume 19, AMS, Providence, 2002.

[11] 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

[12] W. H. Fleming and H. M. Soner, *Controlled Markov processes and viscosity solutions*, 2nd ed., Stochastic Modelling and Applied Probability, vol. 25, Springer, New York, 2006. MR2179357

[13] S. J. Fromm, *Potential space estimates for Green potentials in convex domains*, Proc. Amer. Math. Soc. **119** (1993), no. 1, 225–233, DOI 10.2307/2159846. MR1156467

[14] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition. MR1814364

[15] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Monographs and Studies in Mathematics, vol. 24, Pitman (Advanced Publishing Program), Boston, MA, 1985. MR775683

[16] Q. Han and F. Lin, *Elliptic Partial Differential Equations*, Courant Lecture Notes in Mathematics, vol. 1, New York University, Courant Institute of Mathematical Sciences, New York; American Mathematical Society, Providence, RI, 1997. MR1669352

[17] M. Jensen and I. Smears, *On the convergence of finite element methods for Hamilton-Jacobi-Bellman equations*, SIAM J. Numer. Anal. **51** (2013), no. 1, 137–162, DOI 10.1137/110856198. MR3033005

[18] O. A. Ladyzhenskaya and N. N. Ural′tseva, *Linear and quasilinear elliptic equations*, Translated from the Russian by Scripta Technica, Inc. Translation editor: Leon Ehrenpreis, Academic Press, New York-London, 1968. MR0244627

[19] O. Lakkis and T. Pryer, *A finite element method for second order nonvariational elliptic problems*, SIAM J. Sci. Comput. **33** (2011), no. 2, 786–801, DOI 10.1137/100787672. MR2801189

[20] 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

[21] R. H. Nochetto and W. Zhang, *Discrete ABP estimate and convergence rates for linear elliptic equations in non-divergence form*, arXiv:1411.6036 [math.NA].

[22] A. H. Schatz, *Pointwise error estimates and asymptotic error expansion inequalities for the finite element method on irregular grids. I. Global estimates*, Math. Comp. **67** (1998), no. 223, 877–899, DOI 10.1090/S0025-5718-98-00959-4. MR1464148

[23] A. H. Schatz and J. P. Wang, *Some new error estimates for Ritz-Galerkin methods with minimal regularity assumptions*, Math. Comp. **65** (1996), no. 213, 19–27, DOI 10.1090/S0025-5718-96-00649-7. MR1308460

[24] 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

[25] I. Smears and E. Süli, *Discontinuous Galerkin finite element approximation of Hamilton-Jacobi-Bellman equations with Cordes coefficients*, SIAM J. Numer. Anal. **52** (2014), no. 2, 993–1016, DOI 10.1137/130909536. MR3196952

[26] I. Smears and E. Süli, *Discontinuous Galerkin finite element methods for time-dependent Hamilton–Jacobi–Bellman equations with Cordes coefficients*, Numer. Math. **133** (2016), no. 1, 141–176, DOI 10.1007/s00211-015-0741-6. MR3475658

[27] A. Ženíšek, *Polynomial approximation on tetrahedrons in the finite element method*, J. Approximation Theory **7** (1973), 334–351. MR0350260

[28] S. Zhang, *A family of 3D continuously differentiable finite elements on tetrahedral grids*, Appl. Numer. Math. **59** (2009), no. 1, 219–233, DOI 10.1016/j.apnum.2008.02.002. MR2474112

Department of Mathematics, The University of Tennessee, Knoxville, Tennessee 37996

E-mail address: xfeng@math.utk.edu

Department of Mathematics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

E-mail address: LNH31@pitt.edu

Department of Mathematics, University of Pittsburgh, Pittsburgh, Pennsylvania 15260

E-mail address: neilan@pitt.edu