

# ANALYSIS OF THE VANISHING MOMENT METHOD AND ITS FINITE ELEMENT APPROXIMATIONS FOR SECOND-ORDER LINEAR ELLIPTIC PDES IN NON-DIVERGENCE FORM\*

XIAOBING FENG<sup>†</sup>, THOMAS LEWIS<sup>‡</sup>, AND STEFAN SCHNAKE<sup>§</sup>

*Dedicated to Professor Roland Glowinski on the occasion of his eightieth birthday*

**Abstract.** This paper is concerned with continuous and discrete approximations of  $W^{2,p}$  strong solutions of second-order linear elliptic partial differential equations (PDEs) in non-divergence form. The continuous approximation of these equations is achieved through the Vanishing Moment Method (VMM) which adds a small biharmonic term to the PDE. The structure of the new fourth-order PDE is a natural fit for Galerkin-type methods unlike the original second order equation since the highest order term is in divergence form. The well-posedness of the weak form of the perturbed fourth order equation is shown as well as error estimates for approximating the strong solution of the original second-order PDE. A  $C^1$  finite element method is then proposed for the fourth order equation, and its existence and uniqueness of solutions as well as optimal error estimates in the  $H^2$  norm are shown. Lastly, numerical tests are given to show the validity of the method.

**Key words.** Elliptic PDEs in non-divergence form, strong solution, vanishing moment method,  $C^1$  finite element method, discrete Calderon-Zygmund estimate.

**Mathematics Subject Classification.** 65N12, 65N15, 65N30.

**1. Introduction.** In this paper, we propose  $C^1$  finite element approximations of the following second-order linear elliptic PDE in non-divergence form:

$$\begin{aligned} \mathcal{L}u &:= -A : D^2u + b \cdot \nabla u + cu = f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega, \end{aligned} \tag{P}$$

where  $\Omega$  is an open and bounded domain in  $\mathbb{R}^n$  and  $\partial\Omega$  denotes its boundary. These non-divergence form PDEs have several applications including game theory, stochastic optimal control, and mathematical finance [12]. Moreover, non-divergence PDEs explicitly appear in several second-order fully nonlinear PDEs such as Hamilton-Jacobi-Bellman and Issac's equations as well as in the linearization of the Monge-Ampère equation [5, 4].

When the coefficient matrix  $A$  is not smooth, (P) cannot be written in divergence form. Thus, any standard notion of weak solutions to (P) must be abandoned, and, indeed, the PDE theory respects this observation and seeks well-posedness of these equations in a stronger sense. There have been three main theories for the existence and uniqueness of these equations. First, Schauder (or classical) theory seeks solutions in  $C^{2,\alpha}(\Omega)$  where  $a_{ij}, b_i, c, f \in C^\alpha(\Omega)$  and  $\partial\Omega \in C^{2,\alpha}$ . Second, strong solution theory seeks solutions in  $W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$  that satisfy the PDE almost everywhere. There have been two frameworks that guarantee unique strong solutions. The first requires  $a_{ij} \in C(\bar{\Omega})$ ,  $b_i \in L^\infty(\Omega)$ ,  $c \in L^\infty(\Omega)$ , and  $f \in L^p(\Omega)$ , with  $1 < p < \infty$  and  $\partial\Omega \in C^{1,1}$

---

\*Received December 28, 2017; accepted for publication April 12, 2019. This work was partially supported by the NSF grant DMS-1620168.

<sup>†</sup>Department of Mathematics, The University of Tennessee, Knoxville, TN 37996, U.S.A. (xfeng@math.utk.edu).

<sup>‡</sup>Department of Mathematics and Statistics, The University of North Carolina at Greensboro, Greensboro, NC 27412, U.S.A. (tllewis3@uncg.edu).

<sup>§</sup>Department of Mathematics, The University of Oklahoma, Norman, OK 73019, U.S.A. (sschnake@ou.edu).

while the second requires  $p = 2$ ,  $f \in L^2(\Omega)$ ,  $\partial\Omega$  convex, and  $a_{ij} \in L^\infty(\Omega)$  where the matrix  $A$  satisfies the Córdes condition [13, 16]. The last theory, called viscosity solution theory, seeks solutions in  $C^0(\bar{\Omega})$  given  $a_{ij}, b_i, c \in L^\infty(\Omega)$  and  $f \in C(\Omega)$ , where the underlying viscosity solutions satisfy the PDE in a much weaker sense [5].

Due to the lack of a divergence structure, constructing convergent numerical methods for  $(P)$ , especially finite element methods, is not obvious. Only a handful of Galerkin-type methods have been developed, and these methods did not appear in the literature until quite recently [11, 17, 8, 16, 14]. All of these methods, however diverse they are in their construction, share a common thread: a (nonstandard) direct discretization of  $(P)$ . The method we propose, however, is based upon the Vanishing Moment Method (VMM) - a method developed by Feng and Neilan in [10] for second-order fully nonlinear PDEs such as the Hamilton-Jacobi-Bellman and Monge-Ampère equations. The main solution concept for these equations is that of the viscosity solution which requires passing the derivatives of the solutions to functions that locally lie above or below the graph of the solution [5]. This notion of a solution is not natural in the Galerkin framework since it is not based on integration by parts. The VMM seeks to approximate  $(P)$  by a fourth order, quasi-linear PDE where the fourth order term is a “nice” operator, such as the biharmonic operator. Since this new PDE is in divergence form if the biharmonic operator is chosen, it can be readily adapted to a weak solution concept and, more importantly, allow the natural formulation of Galerkin-type methods. In our case of non-divergence form PDEs, the VMM is given by

$$\begin{aligned} \varepsilon \Delta^2 u^\varepsilon - A : D^2 u^\varepsilon + b \cdot \nabla u^\varepsilon + c u^\varepsilon &= f && \text{in } \Omega, \\ u^\varepsilon &= 0 && \text{on } \partial\Omega, \\ \Delta u^\varepsilon &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{P_\varepsilon}$$

While strong solution theory for  $(P)$  is not as weak as viscosity solution theory for fully nonlinear PDEs, by first applying the VMM to  $(P)$ , the resulting approximate equation  $(P_\varepsilon)$ , whose highest order derivative is in divergence form, can be discretized using a variety of conforming and nonconforming finite element methods. These numerical solutions will converge to the solution of  $(P)$  as  $\varepsilon \rightarrow 0$ . Thus, from a numerical standpoint, the application of the VMM is just as applicable to non-divergence PDEs as they are to fully nonlinear PDEs.

Several papers have been written on ways to formally construct solutions to fully nonlinear PDEs using the Vanishing Moment Method [10, 9]. This paper is the first to offer a detailed analysis of the VMM for a particular class of PDEs. Moreover, since non-divergence PDEs have close ties to many popular second-order fully nonlinear PDEs, they serve as a natural starting point for the complete analysis of the Vanishing Moment Method.

The goals of this paper are to provide a detailed PDE analysis of the Vanishing Moment Method when it is applied to second order elliptic linear PDEs in non-divergence form and to study its finite element approximations. This analysis requires showing that the solution  $u^\varepsilon$  of  $(P_\varepsilon)$  exists and is unique; proving  $u^\varepsilon \rightarrow u$  as  $\varepsilon \rightarrow 0$ , where  $u \in H^2(\Omega) \cap H_0^1(\Omega)$  is the strong solution to  $(P)$ ; and formulating error estimates for  $\|u^\varepsilon - u\|$  in powers of  $\varepsilon$ . We also formulate a  $C^1$  conforming finite element method for approximating the weak solution to  $(P_\varepsilon)$  and derive its error estimate in the energy-norm as well as provide some numerical experiments that test the method and theory. We note that the PDE results given for the VMM will be crucial for the analysis of any Galerkin-type scheme developed for  $(P_\varepsilon)$ , including our

$C^1$  scheme.

The rest of our paper is organized as follows. In Section 2, we set the notation and provide the preliminary information about the well-posedness and stability of  $(P)$ . In Section 3, we formally introduce and analyze the VMM applied to  $(P)$ . In Section 4, we propose a simple  $C^1$  conforming finite element method for our fourth order equation as well as give some tests showing the convergence of the method.

**2. Notation and Preliminaries.** Let  $\Omega \subset \mathbb{R}^n$  be an open and bounded domain. Consider a subdomain  $D \subset \Omega$ . Let  $L^2(D)$  and  $H^k(D)$  be the standard Lebesgue and Sobolev spaces with their respective norms. We define  $(\cdot, \cdot)_D$  to be the  $L^2$ -inner product for all scalar and vector valued functions with  $(\cdot, \cdot) := (\cdot, \cdot)_\Omega$ . In addition, we let  $H_0^k(\Omega)$  be the completion of all compactly supported smooth functions in  $H^k(D)$  and  $H^{-1}(D)$  be the dual space of  $H_0^1(D)$ . Lastly, define the space  $H$  by

$$H := \{v \in H^2(\Omega) \cap H_0^1(\Omega) \text{ with } \Delta v \in H_0^1(\Omega)\}.$$

As mentioned in the introduction, there are three main well-posedness theories for  $(P)$ . We are choosing to focus on approximating  $W^{2,p}$  strong solutions for linear elliptic PDEs. To this end, we assume that  $A \in [C(\bar{\Omega})]^{n \times n}$  is uniformly elliptic; that is, there exist constants  $0 < \lambda \leq \bar{\lambda}$  such that

$$\lambda|\xi|^2 \leq A(x)\xi \cdot \xi \leq \bar{\lambda}|\xi|^2 \quad \forall \xi \in \mathbb{R}^n, x \in \bar{\Omega} \tag{2.1}$$

and  $c \geq 0$  a.e. in  $\Omega$ . Let  $f \in L^p(\Omega)$ . From [13] we have that  $(P)$  exhibits a unique strong solution  $u \in W^{2,p}(\Omega) \cap W_0^{1,p}(\Omega)$  for  $1 < p < \infty$  that satisfies the PDE almost everywhere. Moreover, we have the stability result

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

Here and in the rest of the paper we use  $a \lesssim b$  to denote  $a \leq Cb$  for some constant  $C > 0$  independent of relevant parameters. Estimate (2.2) is called the Calderón-Zygmund estimate for  $\mathcal{L}$ . For simplicity of presentation, we will assume for the remainder of the paper that  $b, c \equiv 0$ .

For the finite element method developed in Section 4, given an  $h > 0$ , we let  $\mathcal{T}_h$  be a quasi-uniform and shape-regular mesh of  $\Omega$ . We set  $V_h \subset H^2(\Omega) \cap H_0^1(\Omega)$  to be a  $C^1$  conforming finite element space over  $\mathcal{T}_h$  satisfying

$$V_h = \{v_h \in C^1(\bar{\Omega}) \cap H_0^1(\Omega) : v_h|_T \in \mathbb{P}_k(T) \ \forall T \in \mathcal{T}_h\},$$

where  $\mathbb{P}_k(T)$  is the set of all polynomials of total degree less than or equal to  $k$ . Examples of such spaces include those defined using the cubic Hermite element in 1-D where,  $k \geq 3$ , or the Argyris element in 2-D, where  $k \geq 5$  [3]. Given a sub-domain  $D \subset \Omega$ , we define

$$V_h = \{v_h \in V_h : v_h = 0 \text{ on } \Omega \setminus D\}$$

and remark that  $V_h(D)$  is non-trivial provided  $h \leq \frac{1}{3} \text{diam}(D)$ .

Define the discrete  $L^2$  norm by

$$\|v\|_{L_h^2(D)} := \sup_{w_h \in V_h(D) \setminus \{0\}} \frac{(v, w_h)_D}{\|w_h\|_{L^2(D)}}. \tag{2.3}$$

Clearly we have

$$\|v_h\|_{L^2(D)} \leq \|v_h\|_{L_h^2(D)} \quad \forall v_h \in V_h.$$

There also holds

$$\|v\|_{L_h^2(D)} \lesssim \|v\|_{L^2(D)} \quad \forall v \in L^2(D).$$

Lastly, define the discrete  $H^{-2}$  norm by

$$\|v\|_{H_h^{-2}(D)} = \sup_{w_h \in V_h(D) \setminus \{0\}} \frac{(v, w_h)}{\|w_h\|_{H^2(D)}}. \quad (2.4)$$

For notational convenience, we will often write  $w_h \in V_h(D)$  instead of  $w_h \in V_h(D) \setminus \{0\}$  in (2.3) and (2.4).

### 3. The Vanishing Moment Method and its Analysis.

**3.1. Construction of the Vanishing Moment Method.** The Vanishing Moment Method is an approximation technique originally developed for second-order fully nonlinear PDEs [10]. The approximation corresponds to converting the original second-order equation into a quasi-linear fourth order equation. Let  $\varepsilon > 0$ . In the context of our problem, the solution  $u$  to (P) will be approximated by the solution  $u^\varepsilon$ , where  $u^\varepsilon$  satisfies the fourth order problem:

$$\begin{aligned} \mathcal{L}^\varepsilon u^\varepsilon &:= \varepsilon \Delta^2 u^\varepsilon - A : D^2 u^\varepsilon = f && \text{in } \Omega, \\ u^\varepsilon &= 0 && \text{on } \partial\Omega, \\ \Delta u^\varepsilon &= 0 && \text{on } \partial\Omega. \end{aligned} \quad (P_\varepsilon)$$

Since  $(P_\varepsilon)$  is a fourth order equation, an additional boundary condition must be added in order to guarantee a unique solution. We add the simply supported boundary condition  $\Delta u^\varepsilon = 0$  for this particular method, due to being a natural boundary condition for the biharmonic equation; moreover, it allows us to achieve  $H^2$  estimates near  $\partial\Omega$ . We refer to [10] for other boundary conditions that may be used.

Using the fact that the highest order derivative of  $(P_\varepsilon)$  is in divergence form, we can easily define a weak solution concept for  $(P_\varepsilon)$  as follows.

**DEFINITION 3.1.** *A function  $u^\varepsilon \in H^2(\Omega) \cap H_0^1(\Omega)$  is called a weak solution to  $(P_\varepsilon)$  if it satisfies*

$$\varepsilon(\Delta u^\varepsilon, \Delta v) - (A : D^2 u^\varepsilon, v) = (f, v) \quad \forall v \in H^2(\Omega) \cap H_0^1(\Omega). \quad (3.2)$$

Note that the simply supported boundary condition is naturally absorbed into the weak formulation.

**3.2. Stability Estimates for  $\mathcal{L}^\varepsilon$ .** We now give the analysis required to show an estimate similar to the Calderon-Zygmund estimate (2.2) but for the operator  $\mathcal{L}^\varepsilon$  by using a freezing coefficient technique. This technique, used in [13, Chapter 9] by Gilbarg-Trudinger to show (2.2), relies on that fact that since  $A$  is continuous, it is approximately constant over small balls. For  $A$  constant, the non-divergence operator  $-A : D^2 u$  is merely a change of basis and dilation away from the Laplacian  $-\Delta u$ . Since (2.2) is true for  $\mathcal{L} = -\Delta$ , one can argue the estimate holds for all

A constant and holds locally for all  $A$  continuous over a small ball. Then, using a partition of unity argument and cut-off functions, the estimate can be shown to hold over all of  $\Omega$ . Section 3.2.1 shows the global Calderon-Zygmund estimates for  $\mathcal{L}^\varepsilon$  for constant coefficient  $A$  while Section 3.2.2 shows the analogous estimates for continuous coefficient  $A$ .

**3.2.1. Estimates for Constant Coefficient Operators.** First we consider the case  $A = A_0$  for  $A_0$  a constant matrix satisfying (2.1). This gives us the following two problems:

$$\begin{aligned} \mathcal{L}_0 u &:= -A_0 : D^2 u = f && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega \end{aligned} \tag{P'}$$

and

$$\begin{aligned} \mathcal{L}_0^\varepsilon u^\varepsilon &:= \varepsilon \Delta^2 u^\varepsilon - A_0 : D^2 u^\varepsilon = f && \text{in } \Omega, \\ u^\varepsilon &= 0 && \text{on } \partial\Omega, \\ \Delta u^\varepsilon &= 0 && \text{on } \partial\Omega. \end{aligned} \tag{P'_\varepsilon}$$

It should be noted that since  $A_0 : D^2 u = \operatorname{div}(A_0 \nabla u)$ , we immediately recover weak solutions  $u \in H_0^1(\Omega)$  and  $u^\varepsilon \in H^2(\Omega) \cap H_0^1(\Omega)$  to (P') and (P'\_\varepsilon), respectively. We seek to derive  $H^1$  estimates for  $u^\varepsilon$ .

**LEMMA 3.1.** *For  $\varepsilon > 0$ , let  $u^\varepsilon \in H^2(\Omega) \cap H_0^1(\Omega)$  be the weak solution of (P'\_\varepsilon). Then there holds the following estimate:*

$$\sqrt{\varepsilon} \|\Delta u^\varepsilon\|_{L^2(\Omega)} + \sqrt{\lambda} \|\nabla u^\varepsilon\|_{L^2(\Omega)} \lesssim \|f\|_{H^{-1}(\Omega)}. \tag{3.5}$$

Moreover,  $u^\varepsilon \rightharpoonup u$  weakly in  $H^1(\Omega)$ , where  $u$  is the weak solution to (P').

*Proof.* Testing (P'\_\varepsilon) by  $u^\varepsilon$ , using the ellipticity condition (2.1), and applying integration by parts, we have

$$\begin{aligned} \varepsilon \|\Delta u^\varepsilon\|_{L^2(\Omega)}^2 + \lambda \|\nabla u^\varepsilon\|_{L^2(\Omega)}^2 &\leq \varepsilon (\Delta u^\varepsilon, \Delta u^\varepsilon) + (A_0 \nabla u^\varepsilon, \nabla u^\varepsilon) \\ &= (f, u^\varepsilon) \\ &\leq \|f\|_{H^{-1}(\Omega)} \|u^\varepsilon\|_{H^1(\Omega)} \\ &\lesssim \|f\|_{H^{-1}(\Omega)} \|\nabla u^\varepsilon\|_{L^2(\Omega)} \\ &\leq \frac{1}{2\delta} \|f\|_{H^{-1}(\Omega)}^2 + \frac{\delta}{2} \|\nabla u^\varepsilon\|_{L^2(\Omega)}^2 \end{aligned}$$

for any  $\delta > 0$ . Choosing  $\delta$ , independent of  $\varepsilon$ , sufficiently small allows us to move  $\|\nabla u^\varepsilon\|_{L^2(\Omega)}$  on the right side to the left and obtain (3.5). This estimate immediately gives us the boundedness of  $\{u^\varepsilon\}$  in  $H_0^1(\Omega)$ . Thus, by the weak compactness of  $H_0^1(\Omega)$ , there exists a subsequence  $\{u^\varepsilon\}$  (not relabeled) and  $u^* \in H_0^1(\Omega)$  such that  $u^\varepsilon \rightharpoonup u^*$  weakly in  $H^1(\Omega)$  as  $\varepsilon \rightarrow 0$ . Since  $u^\varepsilon$  is a weak solution of (P'\_\varepsilon), we have for every  $\varphi \in C_0^\infty(\Omega)$

$$\varepsilon (\Delta u^\varepsilon, \Delta \varphi) + (A_0 \nabla u^\varepsilon, \nabla \varphi) = (f, \varphi). \tag{3.6}$$

Using the weak convergence of  $u^\varepsilon \rightharpoonup u^*$ , the uniform boundedness of  $\varepsilon \|\Delta u^\varepsilon\|_{L^2(\Omega)}^2$ , and passing  $\varepsilon \rightarrow 0$  in (3.6), we obtain

$$(A_0 \nabla u^*, \nabla \varphi) = (f, \varphi).$$

Thus,  $u^*$  is a weak solution of  $(P')$ . By uniqueness we have  $u^* = u$ , and the whole sequence  $u^\varepsilon \rightharpoonup u$  weakly in  $H^1(\Omega)$ . The proof is complete.  $\square$

We now wish to derive local  $H^2$  stability estimates for  $\mathcal{L}^\varepsilon$ .

LEMMA 3.2. *Let  $B \subset\subset \Omega$  and  $v \in H$  with  $\text{supp}(v) \subset B$ . The following estimate holds:*

$$\sqrt{\varepsilon}\|\nabla\Delta v\|_{L^2(B)} + \sqrt{\lambda}\|D^2v\|_{L^2(B)} \lesssim \|\mathcal{L}_0^\varepsilon v\|_{L^2(B)}. \quad (3.7)$$

*Proof.* By testing  $\mathcal{L}_0^\varepsilon v$  by  $-\Delta v$  and integrating by parts, we get

$$(\mathcal{L}_0^\varepsilon v, -\Delta v) = (\varepsilon\Delta^2v - A_0 : D^2v, -\Delta v) = \varepsilon\|\nabla\Delta v\|_{L^2(B)}^2 + (A_0 : D^2v, \Delta v). \quad (3.8)$$

Since  $A_0$  is symmetric and positive definite, there exists an orthogonal matrix  $Q \in \mathbb{R}^{n \times n}$  such that  $Q^T A_0 Q = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n) =: \Lambda$ , where  $\lambda \leq \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ . Let  $y = Q^T x$  and  $\hat{v}(y) = v(Qy) = v(x)$ . Since the Laplacian is preserved under an orthogonal change of basis, we have the following:

$$\Delta_x v(x) = \Delta_y \hat{v}(y), \quad \Delta_x^2 v(x) = \Delta_y^2 \hat{v}(y),$$

$$A_0 : D_x^2 v(x) = \Lambda : D_y^2 \hat{v}(y) = \sum_{j=1}^n \lambda_j \hat{v}_{y_j y_j}(y).$$

Without a loss of generality, we may assume that  $A_0 = \Lambda$  in (3.8). Hence,

$$\begin{aligned} (A_0 : D^2v, \Delta v) &= (\text{div}(A_0 \nabla v), \Delta v) \\ &= -(A_0 \nabla v, \nabla \Delta v) \\ &= -(A_0 \nabla v, \text{div}(D^2v)) \\ &= (\nabla(A_0 \nabla v), D^2v) \\ &= \sum_{j=1}^n \lambda_j \|\nabla v_{x_j}\|_{L^2(B)}^2 \geq \lambda \|D^2v\|_{L^2(B)}^2. \end{aligned} \quad (3.9)$$

Combining (3.8) and (3.9) gives us

$$\begin{aligned} \varepsilon\|\nabla\Delta v\|_{L^2(B)}^2 + \lambda\|D^2v\|_{L^2(B)}^2 &\leq \|\mathcal{L}_0^\varepsilon v\|_{L^2(B)}\|\Delta v\|_{L^2(B)} \\ &\leq \frac{\delta}{2}\|\Delta v\|_{L^2(B)}^2 + \frac{1}{2\delta}\|\mathcal{L}_0^\varepsilon v\|_{L^2(B)}^2 \\ &\leq \frac{\delta}{2}\|D^2v\|_{L^2(B)}^2 + \frac{1}{2\delta}\|\mathcal{L}_0^\varepsilon v\|_{L^2(B)}^2. \end{aligned}$$

Choosing  $\delta$  sufficiently small to move  $\|D^2v\|_{L^2(B)}^2$  to the left hand side gives the desired result. The proof is complete.  $\square$

Next, we derive a similar boundary estimate. Let  $\mathbb{R}_+^n := \{x = (x', x_n) \in \mathbb{R}^n : x_n > 0\}$ ,  $B^+ = B \cap \mathbb{R}_+^n$ , and  $(\partial B^+)^+ = \partial(B^+) \cap \mathbb{R}_+^d$ , where  $B$  is a small ball with its center on the  $x_n$ -axis.

LEMMA 3.3. *Let  $v \in H^2(B) \cap H_0^1(B)$  with  $\Delta v \in H_0^1(B)$  and  $v, \nabla v, \Delta v = 0$  near  $\partial B^+$ . Then we have the following estimate:*

$$\sqrt{\varepsilon}\|\nabla\Delta v\|_{L^2(B^+)} + \lambda\|D^2v\|_{L^2(B^+)} \lesssim \|\mathcal{L}_0^\varepsilon v\|_{L^2(B^+)}. \quad (3.10)$$

*Proof.* We will extend  $v$  from  $B^+$  to  $B$  by an odd reflection, that is  $v(x', x_n) = -v(x', -x_n)$  for all  $x \in B \setminus B^+$ . Since  $\text{supp}(v) \subset B$  after the reflection, we may test the PDE by  $\Delta v$  and use a similar argument as to the one in Lemma 3.2 to obtain

$$\sqrt{\varepsilon} \|\nabla \Delta v\|_{L^2(B)} + \sqrt{\lambda} \|D^2 v\|_{L^2(B)} \lesssim \|\mathcal{L}_0^\varepsilon v\|_{L^2(B)}. \tag{3.11}$$

Since the odd reflection is a bounded linear operator, we have

$$\begin{aligned} \sqrt{\varepsilon} \|\nabla \Delta v\|_{L^2(B^+)} + \sqrt{\lambda} \|D^2 v\|_{L^2(B^+)} &\leq \sqrt{\varepsilon} \|\nabla \Delta v\|_{L^2(B)} + \sqrt{\lambda} \|D^2 v\|_{L^2(B)} \\ &\lesssim \|\mathcal{L}_0^\varepsilon v\|_{L^2(B)} \\ &\lesssim \|\mathcal{L}_0^\varepsilon v\|_{L^2(B^+)}. \end{aligned}$$

The proof is complete.  $\square$

**3.2.2. Estimates for Continuous Coefficient Operators.** Suppose that  $A \in [C(\bar{\Omega})]^{n \times n}$  is uniformly positive definite. In this section, we seek uniform  $H^1$  and  $H^2$  stability estimates for  $\mathcal{L}^\varepsilon$ . Following the freezing coefficients technique, we first need to derive local  $H^1$  and  $H^2$  stability estimates, which in turn require the following lemma controlling the bound of the  $H^{-1}$  norm of the Hessian.

LEMMA 3.4. *Let  $B$  be an open ball and  $v \in H^2(B)$  and  $|\cdot|_1$  denote the vector 1-norm. There holds*

$$\|D^2 v\|_{H^{-1}(B)} \leq n^{\frac{1}{2}} \|\nabla v\|_{L^2(B)}, \tag{3.12}$$

$$\| |D^2 v|_1 \|_{H^{-1}(B)} \leq n \|\nabla v\|_{L^2(B)}, \tag{3.13}$$

where  $n$  is the dimension of the domain  $\Omega$  and

$$\begin{aligned} \|D^2 v\|_{H^{-1}(B)} &= \left( \sum_{i,j=1}^n \|v_{x_i x_j}\|_{H^{-1}(B)}^2 \right)^{\frac{1}{2}}, \\ \| |D^2 v|_1 \|_{H^{-1}(B)} &= \sum_{i,j=1}^n \| |v_{x_i x_j}| \|_{H^{-1}(B)}. \end{aligned}$$

*Proof.* Let  $i, j = 1, \dots, n$  and  $w \in H_0^1(B)$  with  $w \neq 0$ . Integration by parts yields

$$(v_{x_i x_j}, w) = -(v_{x_i}, w_{x_j}) \leq \|v_{x_i}\|_{L^2(B)} \|\nabla w\|_{L^2(B)}.$$

Thus, by the definition of  $\|v\|_{H^{-1}(B)}$ , we have

$$\|v_{x_i x_j}\|_{H^{-1}(B)} = \sup_{\substack{w \in H_0^1(B) \\ w \neq 0}} \frac{(v_{x_i x_j}, w)_B}{\|\nabla w\|_{L^2(B)}} \leq \|v_{x_i}\|_{L^2(B)}.$$

Summing over  $i$  and  $j$  gives us (3.12).

The proof of (3.13) follows the same lines with an additional help of some facts from [13, Chapter 7]. Let  $|v_{x_i, x_j}| = v_{x_i x_j}^+ + v_{x_i x_j}^-$ , where  $v^\pm$  denotes respectively the positive and negative parts of  $v$ . Then

$$\begin{aligned} (v_{x_i x_j}^+, w) &= \int_{\{v_{x_i, x_j} > 0\}} v_{x_i x_j}(x) w(x) \, dx = - \int_{\{v_{x_i, x_j} > 0\}} v_{x_i}(x) w_{x_j}(x) \, dx \\ &\leq \|v_{x_i}\|_{L^2(B)} \|\nabla w\|_{L^2(B)}. \end{aligned}$$

A similar result holds for  $v_{x_i x_j}^-$ . The proof is complete.  $\square$

We are now ready to prove the local  $H^1$  and  $H^2$  stability of  $\mathcal{L}^\varepsilon$ .

LEMMA 3.5. *Let  $x_0 \in \Omega$  and  $B_R(x_0) \subset \Omega$  be the ball of radius  $R$  centered at  $x_0$ . There exists  $R_\delta > 0$ , independent of  $\varepsilon$ , such that, for all  $v \in H$  with  $\text{supp}(v) \subset B := B_{R_\delta}(x_0)$ , the following estimates hold:*

$$\sqrt{\varepsilon} \|\nabla \Delta v\|_{L^2(B)} + \sqrt{\lambda} \|D^2 v\|_{L^2(B)} \lesssim \|\mathcal{L}^\varepsilon v\|_{L^2(B)}, \quad (3.14)$$

$$\sqrt{\varepsilon} \|\Delta v\|_{L^2(B)} + \sqrt{\lambda} \|\nabla v\|_{L^2(B)} \lesssim \|\mathcal{L}^\varepsilon v\|_{H^{-1}(B)}. \quad (3.15)$$

*Proof.* Let  $\delta > 0$ , and define  $A_0 := A(x_0)$ . Since  $A$  is continuous, there exists  $R_\delta > 0$  such that

$$\|A - A_0\|_{L^\infty(B_{R_\delta}(x_0))} \leq \delta.$$

By Lemma 3.2, we have

$$\begin{aligned} \sqrt{\varepsilon} \|\nabla \Delta v\|_{L^2(B)} + \sqrt{\lambda} \|D^2 v\|_{L^2(B)} &\lesssim \|\mathcal{L}_0^\varepsilon v\|_{L^2(B)} \\ &\lesssim \|\mathcal{L}^\varepsilon v\|_{L^2(B)} + \|(\mathcal{L}_0^\varepsilon - \mathcal{L}^\varepsilon)v\|_{L^2(B)} \\ &= \|\mathcal{L}^\varepsilon v\|_{L^2(B)} + \|(A - A_0) : D^2 v\|_{L^2(B)} \\ &\lesssim \|\mathcal{L}^\varepsilon v\|_{L^2(B)} + \|A - A_0\|_{L^\infty(B)} \|D^2 v\|_{L^2(B)} \\ &\leq \|\mathcal{L}^\varepsilon v\|_{L^2(B)} + \delta \|D^2 v\|_{L^2(B)}, \end{aligned}$$

and (3.14) follows for  $\delta$  sufficiently small (independent of  $\varepsilon$  and  $B$ ).

To show (3.15), we follow a similar technique. Using Lemma 3.1 and Lemma 3.4, we have

$$\begin{aligned} \sqrt{\varepsilon} \|\Delta v\|_{L^2(B)} + \sqrt{\lambda} \|\nabla v\|_{L^2(B)} &\lesssim \|\mathcal{L}_0^\varepsilon v\|_{H^{-1}(B)} \\ &\lesssim \|\mathcal{L}^\varepsilon v\|_{H^{-1}(B)} + \|(\mathcal{L}_0^\varepsilon - \mathcal{L}^\varepsilon)v\|_{H^{-1}(B)} \\ &= \|\mathcal{L}^\varepsilon v\|_{H^{-1}(B)} + \|(A - A_0) : D^2 v\|_{H^{-1}(B)} \\ &\lesssim \|\mathcal{L}^\varepsilon v\|_{H^{-1}(B)} + \|A - A_0\|_{L^\infty(B)} \|D^2 v\|_{H^{-1}(B)} \\ &\lesssim \|\mathcal{L}^\varepsilon v\|_{H^{-1}(B)} + \delta n^2 \|\nabla v\|_{L^2(B)}, \end{aligned}$$

and (3.15) follows for  $\delta$  sufficiently small (independent of  $\varepsilon$  and  $B$ ). The proof is complete.  $\square$

Finally, using a partition of unity argument, we can get an interior Gårding inequality.

LEMMA 3.6. *Let  $v \in H$ . For any  $\Omega' \subset\subset \Omega$  the following estimates hold:*

$$\begin{aligned} &\sqrt{\varepsilon} \|\nabla \Delta v\|_{L^2(\Omega')} + \sqrt{\lambda} \|D^2 v\|_{L^2(\Omega')} \\ &\lesssim \|\mathcal{L}^\varepsilon v\|_{L^2(\Omega)} + \|v\|_{L^2(\Omega)} + \varepsilon (\|\nabla v\|_{L^2(\Omega)} + \|\Delta v\|_{L^2(\Omega)} + \|\nabla \Delta v\|_{L^2(\Omega)}), \end{aligned} \quad (3.16)$$

$$\begin{aligned} &\sqrt{\varepsilon} \|\Delta v\|_{L^2(\Omega')} + \sqrt{\lambda} \|\nabla v\|_{L^2(\Omega')} \\ &\lesssim \|\mathcal{L}^\varepsilon v\|_{H^{-1}(\Omega)} + \|v\|_{L^2(\Omega)} + \varepsilon (\|\nabla v\|_{L^2(\Omega)} + \|\Delta v\|_{L^2(\Omega)}). \end{aligned} \quad (3.17)$$

*Proof.* For a ball  $B_R$  with radius  $R$ , let  $\sigma = 1/2$  and consider the cutoff function  $\eta \in C_0^\infty(B_R)$  with  $0 \leq \eta \leq 1$ ,  $\eta \equiv 1$  in  $B_{\sigma R}$ , and  $\eta \equiv 0$  on  $B_R \setminus B_{\sigma' R}$ , where  $\sigma' = 3/4$ .



Moreover,  $\|D^k \eta\|_{L^\infty(B_R)} \lesssim (1 - \sigma)^{-k} R^{-k}$  for  $k = 0, 1, 2, 3, 4$ . Applying (3.14) to the function  $\eta v$  on the ball  $B_{\sigma R}$  gives us

$$\begin{aligned} & \sqrt{\varepsilon} \|\nabla \Delta v\|_{L^2(B_{\sigma R})} + \sqrt{\lambda} \|D^2 v\|_{L^2(B_{\sigma R})} \\ &= \sqrt{\varepsilon} \|\nabla \Delta(\eta v)\|_{L^2(B_{\sigma R})} + \sqrt{\lambda} \|D^2(\eta v)\|_{L^2(B_{\sigma R})} \\ &\leq \sqrt{\varepsilon} \|\nabla \Delta(\eta v)\|_{L^2(B_{\sigma' R})} + \sqrt{\lambda} \|D^2(\eta v)\|_{L^2(B_{\sigma' R})} \\ &\lesssim \|\mathcal{L}^\varepsilon(\eta v)\|_{L^2(B_{\sigma' R})} = \|\varepsilon \Delta^2(\eta v) - A : D^2(\eta v)\|_{L^2(B_{\sigma' R})}. \end{aligned} \quad (3.18)$$

Expanding  $\Delta^2(\eta v)$  and  $A : D^2(\eta v)$  gives us

$$\Delta^2(\eta v) = \eta \Delta^2 v + 4 \nabla \Delta v \cdot \nabla \eta + 6 \Delta v \Delta \eta + 4 \nabla v \cdot \nabla \Delta \eta + v \Delta^2 \eta, \quad (3.19)$$

$$A : D^2(\eta v) = \eta A : D^2 v + 2A \nabla v \cdot \nabla \eta + v A : D^2 \eta. \quad (3.20)$$

Using (3.19) and (3.20) gives us, with the  $L^2$  norm taken over  $B_{\sigma' R}$ ,

$$\begin{aligned} \|\varepsilon \Delta^2(\eta v) - A : D^2(\eta v)\|_{L^2} &\lesssim \|\mathcal{L}^\varepsilon v\|_{L^2} + \frac{1}{(1 - \sigma)R} \|\nabla v\|_{L^2} + \frac{1}{(1 - \sigma)^2 R^2} \|v\|_{L^2} \\ &\quad + \frac{\varepsilon}{(1 - \sigma)^4 R^4} \|v\|_{L^2} + \frac{\varepsilon}{(1 - \sigma)^3 R^3} \|\nabla v\|_{L^2} \\ &\quad + \frac{\varepsilon}{(1 - \sigma)^2 R^2} \|\Delta v\|_{L^2} + \frac{\varepsilon}{(1 - \sigma)R} \|\nabla \Delta v\|_{L^2}. \end{aligned} \quad (3.21)$$

The treatment of the first three terms on the right follows from the interpolation technique in [13, p.236]. Keeping the rest of the terms on the right and using a covering argument we arrive at (3.16). Since  $\overline{\Omega'}$  is compact it will only take a finite number of balls to cover  $\Omega'$ . Thus, the estimate does not depend on  $R$ .

To show (3.17), using the same  $\eta$  as prescribed above we first recover a similar estimate to (3.18) for (3.17):

$$\begin{aligned} & \sqrt{\varepsilon} \|\Delta v\|_{L^2(B_{\sigma R})} + \sqrt{\lambda} \|\nabla v\|_{L^2(B_{\sigma R})} \\ &\lesssim \|\mathcal{L}^\varepsilon(\eta v)\|_{H^{-1}(B_{\sigma' R})} \\ &= \|\varepsilon \Delta^2(\eta v) - A : D^2(\eta v)\|_{H^{-1}(B_{\sigma' R})} \\ &= \sup_{\substack{w \in H_0^1(B_{\sigma' R}) \\ w \neq 0}} \frac{(\varepsilon \Delta^2(\eta v) - A : D^2(\eta v), w)}{\|\nabla w\|_{L^2(B_{\sigma' R})}}. \end{aligned} \quad (3.22)$$

Let  $w \in H_0^1(B_{\sigma' R})$ . By integration by parts we have

$$\begin{aligned} (\varepsilon \Delta^2(\eta v) - A : D^2(\eta v), w) &= -\varepsilon (\nabla \Delta(\eta v), \nabla w) - (A : D^2(\eta v), w) \\ &:= \varepsilon I_1 + I_2. \end{aligned} \quad (3.23)$$

We first focus on  $I_2$ . Expanding  $\nabla \Delta(\eta v)$  similar to (3.19) and integrating by parts

gives us

$$\begin{aligned}
I_1 &= -(\nabla\Delta(\eta v), \nabla w) & (3.24) \\
&= -(\eta\nabla\Delta v, \nabla w) - (3\Delta v\nabla\eta + 3\Delta\eta\nabla v + u^\varepsilon\nabla\Delta\eta, \nabla w) \\
&= (\operatorname{div}(\eta\nabla\Delta v), w) - (3\Delta v\nabla\eta + 3\Delta\eta\nabla v + u^\varepsilon\nabla\Delta\eta, \nabla w) \\
&= (\Delta^2 v, \eta w) + (\nabla\Delta v, \nabla\eta w) - (3\Delta v\nabla\eta + 3\Delta\eta\nabla v + u^\varepsilon\nabla\Delta\eta, \nabla w) \\
&= (\Delta^2 v, \eta w) - (\Delta v, \operatorname{div}(\nabla\eta w)) - (3\Delta v\nabla\eta + 3\Delta\eta\nabla v + u^\varepsilon\nabla\Delta\eta, \nabla w) \\
&= (\Delta^2 v, \eta w) - (\Delta v, w\Delta\eta + \nabla\eta \cdot \nabla w) - (3\Delta v\nabla\eta + 3\Delta\eta\nabla v + v\nabla\Delta\eta, \nabla w) \\
&\lesssim (\Delta^2 v, \eta w) + \left( \frac{1}{(1-\sigma)^2 R^2} \|\Delta v\|_{L^2} + \frac{1}{(1-\sigma)^2 R^2} \|\nabla v\|_{L^2} \right. \\
&\quad \left. + \frac{1}{(1-\sigma)^3 R^3} \|v\|_{L^2} \right) \|\nabla w\|_{L^2}.
\end{aligned}$$

Using (3.20) on  $I_2$ , we get

$$\begin{aligned}
I_2 &= -(A : D^2(\eta v), w) & (3.25) \\
&= -(\eta A : D^2 v + 2A\nabla v \cdot \nabla\eta + vA : D^2\eta, w) \\
&\lesssim -(A : D^2 v, \eta w) + \left( \frac{1}{(1-\sigma)R} \|\nabla v\|_{L^2} + \frac{1}{(1-\sigma)^2 R^2} \|v\|_{L^2} \right) \|\nabla w\|_{L^2}.
\end{aligned}$$

Combining (3.24) and (3.25) with (3.23) gives us

$$\begin{aligned}
&\sqrt{\varepsilon} \|\Delta v\|_{L^2(B_{\sigma R})} + \sqrt{\lambda} \|\nabla v\|_{L^2(B_{\sigma R})} \\
&\lesssim \|\mathcal{L}^\varepsilon v\|_{H^{-1}(\Omega)} + \frac{1}{(1-\sigma)R} \|\nabla v\|_{L^2} + \frac{1}{(1-\sigma)^2 R^2} \|v\|_{L^2} \\
&\quad + \frac{\varepsilon}{(1-\sigma)^3 R^3} \|v\|_{L^2} + \frac{\varepsilon}{(1-\sigma)^2 R^2} \|\nabla v\|_{L^2} \\
&\quad + \frac{\varepsilon}{(1-\sigma)^2 R^2} \|\Delta v\|_{L^2}.
\end{aligned} \tag{3.26}$$

Following a similar treatment as (3.21), we arrive at (3.17). The proof is complete.  $\square$

We now desire a global estimate instead of an interior estimate.

LEMMA 3.7. *Let  $\partial\Omega \in C^{2,1}$  and  $v \in H$ . The following estimates hold:*

$$\begin{aligned}
\sqrt{\varepsilon} \|\nabla\Delta v\|_{L^2(\Omega)} + \sqrt{\lambda} \|D^2 v\|_{L^2(\Omega)} &\lesssim \|\mathcal{L}^\varepsilon v\|_{L^2(\Omega)} + \|v\|_{L^2(\Omega)} + \varepsilon (\|\nabla v\|_{L^2(\Omega)} \\
&\quad + \|\Delta v\|_{L^2(\Omega)} + \|\nabla\Delta v\|_{L^2(\Omega)}),
\end{aligned} \tag{3.27}$$

$$\begin{aligned}
\sqrt{\varepsilon} \|\Delta v\|_{L^2(\Omega)} + \sqrt{\lambda} \|\nabla v\|_{L^2(\Omega)} &\lesssim \|\mathcal{L}^\varepsilon v\|_{H^{-1}(\Omega)} + \|v\|_{L^2(\Omega)} \\
&\quad + \varepsilon (\|\nabla v\|_{L^2(\Omega)} + \|\Delta v\|_{L^2(\Omega)}).
\end{aligned} \tag{3.28}$$

*Proof.* Since  $\partial\Omega \in C^{2,1}$ , for any  $x_0 \in \partial\Omega$ , we may flatten  $\partial\Omega$  near  $x_0$  and use Lemma 3.3 and the proof in Lemma 3.5 to create a local boundary estimate mimicking (3.14) and (3.15). Following the same argument as in Lemma 3.6 we can obtain estimates (3.16) and (3.17) near the boundary. These estimates combined with (3.16) and (3.17) give us (3.27) and (3.28). The proof is complete.  $\square$

Now we must deal with the terms involving  $\varepsilon$  on the right hand side of (3.27) and (3.29). Noting that  $\varepsilon/\sqrt{\varepsilon} \rightarrow 0$  as  $\varepsilon \rightarrow 0$ , we may hide these terms for  $\varepsilon$  sufficiently small.

LEMMA 3.8. *Let  $\partial\Omega \in C^{2,1}$  and  $v \in H$ . There exists  $\varepsilon_0 > 0$  such that for any  $\varepsilon < \varepsilon_0$  the following estimates hold:*

$$\sqrt{\varepsilon}\|\nabla\Delta v\|_{L^2(\Omega)} + \sqrt{\lambda}\|D^2v\|_{L^2(\Omega)} + \sqrt{\lambda}\|\nabla v\|_{L^2(\Omega)} \lesssim \|\mathcal{L}^\varepsilon v\|_{L^2(\Omega)} + \|v\|_{L^2(\Omega)}, \quad (3.29)$$

$$\sqrt{\varepsilon}\|\Delta v\|_{L^2(\Omega)} + \sqrt{\lambda}\|\nabla v\|_{L^2(\Omega)} \lesssim \|\mathcal{L}^\varepsilon v\|_{H^{-1}(\Omega)} + \|v\|_{L^2(\Omega)}. \quad (3.30)$$

*Proof.* Adding (3.27) and (3.28) and noting that  $\|\mathcal{L}^\varepsilon v\|_{H^{-1}(\Omega)} \leq \|\mathcal{L}^\varepsilon v\|_{L^2(\Omega)}$  yields

$$\begin{aligned} & \sqrt{\varepsilon}\|\nabla\Delta v\|_{L^2(\Omega)} + \sqrt{\varepsilon}\|\Delta v\|_{L^2(\Omega)} + \sqrt{\lambda}\|D^2v\|_{L^2(\Omega)} + \sqrt{\lambda}\|\nabla v\|_{L^2(\Omega)} \\ & \leq C \left( \|\mathcal{L}^\varepsilon v\|_{L^2(\Omega)} + \|v\|_{L^2(\Omega)} + \varepsilon(\|\nabla v\|_{L^2(\Omega)} + \|\Delta v\|_{L^2(\Omega)} + \|\nabla\Delta v\|_{L^2(\Omega)}) \right), \end{aligned} \quad (3.31)$$

where  $C$  is independent of  $\varepsilon$ . Choosing  $\varepsilon_0 = \min\{4/C^2, \sqrt{\lambda}/(2C)\}$  gives us  $C\varepsilon < \sqrt{\varepsilon}/2$  and  $C\varepsilon < \sqrt{\lambda}/2$  for all  $\varepsilon < \varepsilon_0$ . Letting  $\varepsilon < \varepsilon_0$ , we can move the terms with coefficient  $\varepsilon$  on the right hand side of (3.31) and obtain

$$\begin{aligned} & \sqrt{\varepsilon}\|\nabla\Delta v\|_{L^2(\Omega)} + \sqrt{\varepsilon}\|\Delta v\|_{L^2(\Omega)} + \sqrt{\lambda}\|D^2v\|_{L^2(\Omega)} + \sqrt{\lambda}\|\nabla v\|_{L^2(\Omega)} \\ & \leq C \left( \|\mathcal{L}^\varepsilon v\|_{L^2(\Omega)} + \|v\|_{L^2(\Omega)} \right). \end{aligned} \quad (3.32)$$

Dropping  $\sqrt{\varepsilon}\|\Delta v\|_{L^2(\Omega)}$  gives us (3.29). (3.30) is an immediate consequence of (3.28). The proof is complete.  $\square$

We can derive a full stability estimate from Lemma 3.8 through contradiction and showing that  $\mathcal{L}^\varepsilon$  is 1-1.

LEMMA 3.9. *Let  $\partial\Omega \in C^{2,1}$  and  $v \in H$ . For all  $\varepsilon < \varepsilon_0$ , we have the following stability estimate:*

$$\sqrt{\varepsilon}\|\nabla\Delta v\|_{L^2(\Omega)} + \sqrt{\lambda}\|D^2v\|_{L^2(\Omega)} + \sqrt{\lambda}\|\nabla v\|_{L^2(\Omega)} \leq C\|\mathcal{L}^\varepsilon v\|_{L^2(\Omega)}, \quad (3.33)$$

$$\sqrt{\varepsilon}\|\Delta v\|_{L^2(\Omega)} + \sqrt{\lambda}\|\nabla v\|_{L^2(\Omega)} \leq C\|\mathcal{L}^\varepsilon v\|_{H^{-1}(\Omega)}, \quad (3.34)$$

where  $C$  is independent of  $\varepsilon$  and  $u^\varepsilon$ .

*Proof.* Fix  $\varepsilon < \varepsilon_0$ . We first show that the operator  $\mathcal{L}^\varepsilon$  is 1-1 using eigenvalue theory. Define  $H^{-2}(\Omega) = (H^2(\Omega) \cap H_0^1(\Omega))^*$  and  $\mathcal{K} : H^2(\Omega) \cap H_0^1(\Omega) \rightarrow H^{-2}(\Omega)$  by

$$(\mathcal{K}w, v) = (\Delta w, \Delta v) \quad \forall v \in H^2(\Omega) \cap H_0^1(\Omega).$$

By the elliptic existence theory for fourth order problems,  $\mathcal{K}$  is invertible with  $\mathcal{K}^{-1}$  bounded. We see that the operator  $\mathcal{L}^\varepsilon = \varepsilon\mathcal{K} + \mathcal{L}$  satisfies

$$(\mathcal{L}^\varepsilon w, v) = \varepsilon(\Delta w, \Delta v) - (A : D^2w, v) \quad \forall v \in H^2(\Omega) \cap H_0^1(\Omega). \quad (3.35)$$

Thus, a weak solution  $u^\varepsilon$  to  $(P_\varepsilon)$  satisfies  $\mathcal{L}^\varepsilon u^\varepsilon = f$  in  $H^{-2}(\Omega)$ .

Suppose  $\mathcal{L}^\varepsilon$  is not 1-1. Then there exists  $w \neq 0$  such that  $\mathcal{L}^\varepsilon w = \mathcal{L}w + \varepsilon\mathcal{K}w = 0$  in  $H^{-2}(\Omega)$ . Hence,  $\mathcal{L}w = -\mathcal{K}(\varepsilon w)$  in  $H^{-2}(\Omega)$ . Since  $\mathcal{K}$  is invertible we have,  $\mathcal{K}^{-1}\mathcal{L}w = -\varepsilon w$  in  $H^2(\Omega)$ . Thus,  $-\varepsilon$  is an eigenvalue for  $\mathcal{K}^{-1}\mathcal{L}$ . Since  $\mathcal{K}^{-1}$  is a symmetric and compact operator, its eigenvalues are positive and tend to 0. Moreover, the eigenvalues of  $\mathcal{L}$ , while complex, have real parts greater than some positive constant  $r$  [7, p.361]. Since  $\mathcal{K}^{-1}$  is positive definite, repeating the proof given in [7, p.361], we can show that the eigenvalues of  $\mathcal{K}^{-1}\mathcal{L}$  must have positive real parts (one may verify easily the conclusion in the finite-dimensional case). Thus  $-\varepsilon$  cannot be an eigenvalue, which is a contradiction. Hence,  $\mathcal{L}^\varepsilon$  must be 1-1.

To show (3.33) we argue by contradiction. Suppose that, for all  $k \in \mathbb{N}$ , there exists  $v_k \in H$  with  $\|v_k\|_{L^2(\Omega)} = 1$  and  $\|\mathcal{L}^\varepsilon v_k\|_{L^2(\Omega)} \rightarrow 0$  as  $k \rightarrow \infty$ . By Lemma 3.8 we have  $\sqrt{\varepsilon}\|\nabla\Delta v_k\|_{L^2(\Omega)} + \|v_k\|_{H^2(\Omega)}$  is uniformly bounded in  $k$ . Thus, we may extract a convergent subsequence (not relabeled) and a  $v_* \in H$  such that  $v_k \rightharpoonup v_*$  weakly in  $H$  and  $\sqrt{\varepsilon}\|\nabla\Delta v_*\|_{L^2(\Omega)} + \|v_*\|_{H^2(\Omega)} > 0$ . Moreover, since  $v_k \rightharpoonup v_*$  and  $\mathcal{L}^\varepsilon$  is linear, we have  $\mathcal{L}^\varepsilon v_k \rightharpoonup \mathcal{L}^\varepsilon v_*$  in  $H^{-2}(\Omega)$  and

$$0 \leq \|\mathcal{L}^\varepsilon v_*\|_{H^{-2}(\Omega)} \leq \liminf_{k \rightarrow \infty} \|\mathcal{L}^\varepsilon v_k\|_{H^{-2}(\Omega)} \leq \liminf_{k \rightarrow \infty} \|\mathcal{L}^\varepsilon v_k\|_{L^2(\Omega)} = 0.$$

Thus,  $\|\mathcal{L}^\varepsilon v_*\|_{H^{-2}(\Omega)} = 0$  and consequently  $\mathcal{L}^\varepsilon v_* = 0$  in  $H^{-2}(\Omega)$ . Clearly  $u^\varepsilon \equiv 0$  is a solution to  $(P_\varepsilon)$  when  $f \equiv 0$ . Since  $\mathcal{L}^\varepsilon$  is 1-1, then  $v_* = 0$  in  $H$  which is a contradiction to the fact that  $\sqrt{\varepsilon}\|\nabla\Delta v_*\|_{L^2(\Omega)} + \|v_*\|_{H^2(\Omega)} > 0$ . The proof is complete.  $\square$

We are now ready to show the well-posedness of  $\mathcal{L}^\varepsilon$ .

**THEOREM 3.10.** *Let  $\varepsilon < \varepsilon_0$  and  $\partial\Omega \in C^{2,1}$ . Then, for every  $f \in L^2(\Omega)$ , there exists a unique weak solution  $u^\varepsilon \in H$  to  $(P_\varepsilon)$ . Moreover, the weak solution satisfies*

$$\sqrt{\varepsilon}\|\nabla\Delta u^\varepsilon\|_{L^2(\Omega)} + \sqrt{\lambda}\|D^2 u^\varepsilon\|_{L^2(\Omega)} + \sqrt{\lambda}\|\nabla u^\varepsilon\|_{L^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}, \quad (3.36)$$

$$\sqrt{\varepsilon}\|\Delta u^\varepsilon\|_{L^2(\Omega)} + \sqrt{\lambda}\|\nabla u^\varepsilon\|_{L^2(\Omega)} \leq C\|f\|_{H^{-1}(\Omega)}. \quad (3.37)$$

*Proof.* We approximate  $A$  by  $A_k \in [C^1(\bar{\Omega})]^{n \times n}$ , where  $A_k \rightarrow A$  uniformly in  $\Omega$ . Since  $A_k$  is differentiable,

$$-A_k : D^2 v = -\operatorname{div}(A_k \nabla v) + \operatorname{div}(A_k) \cdot \nabla v,$$

which converts the non-divergence operator into a sum of a diffusion operator and an advection operator. Consider weak solutions to the problem

$$\begin{aligned} \varepsilon\Delta^2 v - \operatorname{div}(A_k \nabla v) + \operatorname{div}(A_k) \cdot \nabla v &= f \text{ in } \Omega, \\ v &= 0 \text{ on } \partial\Omega, \\ \Delta v &= 0 \text{ on } \partial\Omega. \end{aligned} \quad (3.38)$$

By the fourth-order elliptic PDE theory [1, Chapter 8], there is a weak solution  $u_k^\varepsilon \in H^2(\Omega) \cap H_0^1(\Omega)$  to (3.38). Moreover,  $u_k^\varepsilon \in H$  since  $\partial\Omega \in C^{2,1}$  and, by Lemma 3.9,  $u_k^\varepsilon$  satisfies

$$\sqrt{\varepsilon}\|\nabla\Delta u_k^\varepsilon\|_{L^2(\Omega)} + \sqrt{\lambda}\|D^2 u_k^\varepsilon\|_{L^2(\Omega)} + \sqrt{\lambda}\|\nabla u_k^\varepsilon\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}, \quad (3.40)$$

$$\sqrt{\varepsilon}\|\Delta u_k^\varepsilon\|_{L^2(\Omega)} + \sqrt{\lambda}\|\nabla u_k^\varepsilon\|_{L^2(\Omega)} \lesssim \|f\|_{H^{-1}(\Omega)}. \quad (3.41)$$

Thus,  $\{u_k^\varepsilon\}_k$  is bounded in  $H$ , and it follows that  $u_k^\varepsilon$  weakly converges to some  $u^\varepsilon \in H$ . By the linearity of the PDE,  $u^\varepsilon$  is a weak solution of  $(P_\varepsilon)$ . Since  $\mathcal{L}^\varepsilon$  is 1-1, the solution  $u^\varepsilon$  is unique. Taking the  $\liminf$  as  $k \rightarrow \infty$  of (3.40) and (3.41) we arrive at (3.36) and (3.37). The proof is complete.  $\square$

**REMARK 3.1.** *In the above three lemmas we require that  $\partial\Omega \in C^{2,1}$ . This is imposed in order to achieve the global estimates using a boundary flattening technique. More precisely, we use the assumption to preserve  $\|\nabla\Delta v\|_{L^2}$  when flattening the boundary,  $\partial\Omega$ . Thus, to guarantee  $u^\varepsilon \in H$ , we need  $\partial\Omega \in C^{2,1}$ . We note that if  $\partial\Omega \in C^{1,1}$ , the argument still works, but all  $\|\nabla\Delta v\|_{L^2}$  terms must be dropped implying  $u^\varepsilon$  is only in  $H^2$ . This does not affect convergence but does affect the  $H^1$  error estimate in Theorem 3.13 below.*

We now prove the convergence of the solutions  $u^\varepsilon \rightarrow u$ , where  $u$  is a strong solution to  $(P)$ . We also give an  $H^1$  stability result which will be useful for the numerical discretization of  $(P)$  and has not yet been obtained in the literature.

**THEOREM 3.11.** *Let  $\varepsilon < \varepsilon_0$  and  $u^\varepsilon$  be the solution to  $(P_\varepsilon)$ . Then  $u^\varepsilon$  converges to  $u \in H^2(\Omega) \cap H_0^1(\Omega)$  weakly in  $H^2(\Omega)$ , where  $u$  is the strong solution to  $(P)$  with  $b, c \equiv 0$ . Moreover, we have the following  $H^1$  stability result for  $\mathcal{L}$ :*

$$\|\nabla u\|_{L^2(\Omega)} \lesssim \|\mathcal{L}u\|_{H^{-1}(\Omega)}. \quad (3.42)$$

*Proof.* Since  $\mathcal{L}^\varepsilon u^\varepsilon = f$  in  $\Omega$  and  $u^\varepsilon = 0$  on  $\partial\Omega$ , we have the boundedness of  $\|u^\varepsilon\|_{H^2(\Omega)}$  from the Poincaré inequality and Lemma 3.9. By compactness we can extract a subsequence  $\{u^\varepsilon\}$  (not relabeled) and  $u^* \in H^2(\Omega) \cap H_0^1(\Omega)$  such that  $u^\varepsilon \rightharpoonup u^*$  weakly in  $H^2(\Omega)$ . Moreover, since  $\mathcal{L}^\varepsilon u^\varepsilon = f$ , we have for any  $\varphi \in C_0^\infty(\Omega)$

$$\varepsilon(\Delta u^\varepsilon, \Delta \varphi) - (A : D^2 u^\varepsilon, \varphi) = (f, \varphi). \quad (3.43)$$

Since  $\varepsilon(\Delta u^\varepsilon, \Delta \varphi) \leq \varepsilon \|\Delta u^\varepsilon\|_{L^2(\Omega)} \|\Delta \varphi\|_{L^2(\Omega)} \rightarrow 0$  as  $\varepsilon \rightarrow 0$ , letting  $\varepsilon \rightarrow 0$  in (3.43) we obtain

$$-(A : D^2 u^*, \varphi) = (f, \varphi) \quad \forall \varphi \in C_0^\infty(\Omega). \quad (3.44)$$

Thus,  $u^*$  is a strong solution to  $(P)$ . Since  $\mathcal{L}$  is 1-1, we have  $u^* = u$ , and the whole sequence  $u^\varepsilon \rightarrow u$  as  $\varepsilon \rightarrow 0$ .

We now derive (3.42). Since  $\|\mathcal{L}^\varepsilon u^\varepsilon\|_{H^{-1}(\Omega)} = \|f\|_{H^{-1}(\Omega)}$  is constant with respect to  $\varepsilon$ , we take the  $\liminf$  of (3.34) and use  $u^\varepsilon \rightarrow u$  to obtain

$$\begin{aligned} \|\mathcal{L}u\|_{H^{-1}(\Omega)} &= \liminf_{\varepsilon \rightarrow 0} \|\mathcal{L}^\varepsilon u^\varepsilon\|_{H^{-1}(\Omega)} \\ &\gtrsim \liminf_{\varepsilon \rightarrow 0} \left( \sqrt{\varepsilon} \|\Delta u^\varepsilon\|_{L^2(\Omega)} + \sqrt{\lambda} \|\nabla u^\varepsilon\|_{L^2(\Omega)} \right) \\ &\gtrsim \liminf_{\varepsilon \rightarrow 0} \sqrt{\varepsilon} \|\Delta u^\varepsilon\|_{L^2(\Omega)} + \liminf_{\varepsilon \rightarrow 0} \sqrt{\lambda} \|\nabla u^\varepsilon\|_{L^2(\Omega)} \\ &\gtrsim \|\nabla u\|_{L^2(\Omega)}. \end{aligned}$$

The proof is complete.  $\square$

Now that we have the existence and convergence of  $u^\varepsilon$  to the strong solution  $u$  of  $(P)$ , in  $H^2(\Omega)$ , we note that indeed the regularity of  $u^\varepsilon$  can be higher than  $u$  – up to  $H^4$  dependent on the smoothness of  $\partial\Omega$  – and is shown in the theorem below.

**THEOREM 3.12.** *Let  $\varepsilon < \varepsilon_0$ ,  $\partial\Omega \in C^3$ , and  $u^\varepsilon \in H$  be the weak solution to  $(P_\varepsilon)$ . Then  $u^\varepsilon \in H^3(\Omega)$  with*

$$\varepsilon \|u^\varepsilon\|_{H^3(\Omega)} \lesssim \|f\|_{L^2(\Omega)}. \quad (3.45)$$

Moreover, if  $\partial\Omega \in C^4$ , then  $u^\varepsilon \in H^4(\Omega)$  with the estimate

$$\varepsilon \|u^\varepsilon\|_{H^4(\Omega)} \lesssim \|f\|_{L^2(\Omega)}. \quad (3.46)$$

*Proof.* Since  $u^\varepsilon \in H^2(\Omega) \cap H_0^1(\Omega)$  and  $A \in [L^\infty(\Omega)]^{d \times d}$ , then  $A : D^2 u^\varepsilon + f \in L^2(\Omega)$  with  $\|A : D^2 u^\varepsilon + f\|_{L^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}$ . Moreover, since  $u^\varepsilon \in H$  is weak solution to  $(P_\varepsilon)$ , we have

$$\varepsilon(\Delta u^\varepsilon, \Delta v) = (A : D^2 u^\varepsilon + f, v)$$

Thus  $u^\varepsilon$  is the weak solution to the biharmonic equation

$$\begin{aligned} \varepsilon \Delta^2 u^\varepsilon &= A_0 : D^2 u^\varepsilon + f && \text{in } \Omega, \\ u^\varepsilon &= \Delta u^\varepsilon = 0 && \text{on } \partial\Omega, \end{aligned}$$

and the smoothness of  $u^\varepsilon$  is entirely dependent on the smoothness of  $\partial\Omega$ , and the desired estimates (3.45)-(3.46) follow from the elliptic theory for the biharmonic equation. The proof is complete.  $\square$

We conclude this section with some error estimates of  $u^\varepsilon - u$  in various norms.

**THEOREM 3.13.** *Let  $\varepsilon < \varepsilon_0$ ,  $u \in H^2(\Omega) \cap H_0^1(\Omega)$ , and  $u^\varepsilon \in H$  be the solutions to (P) and  $(P_\varepsilon)$ , respectively. Then we have the following error estimates:*

$$\|\nabla(u^\varepsilon - u)\|_{L^2(\Omega)} \lesssim \sqrt{\varepsilon} \|f\|_{L^2(\Omega)}, \quad (3.47)$$

$$\|u^\varepsilon - u\|_{L^2(\Omega)} \lesssim \sqrt{\varepsilon} \|f\|_{L^2(\Omega)}. \quad (3.48)$$

*Proof.* Let  $e^\varepsilon = u^\varepsilon - u$ . By the linearity of  $\mathcal{L}$  we get  $\mathcal{L}e^\varepsilon = \varepsilon \Delta^2 u^\varepsilon \in H^{-1}(\Omega)$  since  $u^\varepsilon \in H$ . Using (3.42) and (3.36), we have

$$\begin{aligned} \|\nabla e^\varepsilon\|_{L^2(\Omega)} &\lesssim \varepsilon \|\Delta^2 u^\varepsilon\|_{H^{-1}(\Omega)} \\ &= \sup_{v \in H_0^1(\Omega)} \frac{-\varepsilon(\nabla \Delta u^\varepsilon, \nabla v)}{\|\nabla v\|_{L^2(\Omega)}} \\ &\leq \varepsilon \|\nabla \Delta u^\varepsilon\|_{L^2(\Omega)} \\ &\lesssim \sqrt{\varepsilon} \|f\|_{L^2(\Omega)}, \end{aligned}$$

which infers (3.47). Poincaré's inequality yields (3.48). The proof is complete.  $\square$

We note that the rates in Theorem 3.13 may be suboptimal based on the numerical experiments provided below.

**4.  $C^1$  Finite Element Methods for Second Order Elliptic Linear PDEs in Non-Divergence Form.** Since the highest order derivative of  $(P_\varepsilon)$  is in divergence form, we can easily and naturally discretize it using  $C^1$  conforming finite element methods.

**DEFINITION 4.1.** *We define our  $C^1$  finite element method for  $(P_\varepsilon)$  as seeking  $u_h^\varepsilon \in V_h$  such that*

$$\mathcal{A}^\varepsilon(u_h^\varepsilon, v_h) = (f, v_h) \quad \forall v_h \in V_h, \quad (4.1)$$

where

$$\mathcal{A}^\varepsilon(w_h, v_h) := \varepsilon(\Delta w_h, \Delta v_h) - (A : D^2 w_h, v_h).$$

**4.1. Well-posedness of  $\mathcal{L}_h^\varepsilon$ .** In this subsection, we show the well-posedness of Problem (4.1). We define the operator  $\mathcal{L}_h^\varepsilon : V_h \rightarrow V_h$  by

$$(\mathcal{L}_h^\varepsilon v_h, w_h) = \mathcal{A}^\varepsilon(v_h, w_h) \quad \forall w_h \in V_h.$$

We can also naturally extend the domain of  $\mathcal{L}_h^\varepsilon$  to  $H^2(\Omega) \cap H_0^1(\Omega)$ .

We assume that there is a unique weak solution  $u^\varepsilon \in H^2(\Omega) \cap H_0^1(\Omega)$  of  $(P_\varepsilon)$  and that the operator  $\mathcal{L}^\varepsilon$  satisfies the Calderon-Zygmund estimate

$$\|v\|_{H^2(\Omega)} \lesssim \|\mathcal{L}^\varepsilon v\|_{L^2(\Omega)}. \tag{4.2}$$

We note that Theorem 3.10 implies this result provided  $\partial\Omega \in C^{2,1}$  and  $\varepsilon$  is sufficiently small.

Our goal is to derive a similar estimate to (4.2) for  $\mathcal{L}_h^\varepsilon$ , that is,

$$\|v_h\|_{H^2(\Omega)} \lesssim \|\mathcal{L}_h^\varepsilon v_h\|_{L_h^2(\Omega)} \quad \forall v_h \in V_h. \tag{4.3}$$

To this end, we will adapt the freezing coefficient technique as in Section 3.2 but at the discrete level. Section 4.1.1 shows (4.3) for constant coefficient  $A$  which will be used in Section 4.1.2 to give us (4.3).

**4.1.1. Stability Analysis of  $\mathcal{L}_h^\varepsilon$  for Constant Coefficient  $A$ .** Suppose that  $A = A_0 \equiv A(x_0)$  for an  $x_0 \in \Omega$ . Since  $A$  is constant, we have  $A : D^2 v_h = \operatorname{div}(A \nabla v_h)$  for all  $v_h \in V_h$ . Thus, we can define  $\mathcal{L}_{0,h}^\varepsilon : V_h \rightarrow V_h$  by

$$(\mathcal{L}_{0,h}^\varepsilon v_h, w_h) = \mathcal{A}_0^\varepsilon(v_h, w_h) \quad \forall w_h \in V_h,$$

where

$$\mathcal{A}_0^\varepsilon(v_h, w_h) := \varepsilon(\Delta v_h, \Delta w_h) + (A \nabla v_h, \nabla w_h).$$

Since  $\mathcal{A}_0^\varepsilon(v_h, w_h)$  is clearly continuous and coercive on  $V_h$  with the norm

$$\|w_h\|_E^2 := \varepsilon \|\Delta w_h\|_{L^2(\Omega)}^2 + \|\nabla w_h\|_{L^2(\Omega)}^2,$$

we have the existence and uniqueness of a solution  $u_h^\varepsilon$  satisfying

$$(\mathcal{L}_{0,h}^\varepsilon u_h^\varepsilon, w_h) = (f, w_h) \quad \forall w_h \in V_h \tag{4.4}$$

and the estimate

$$\|u_h^\varepsilon - u^\varepsilon\|_E \lesssim h^{r-2}(\sqrt{\varepsilon} + h) \|u^\varepsilon\|_{H^s(\Omega)}, \tag{4.5}$$

where  $r = \min\{s, k + 1\}$  and  $u^\varepsilon \in H^s(\Omega)$  is the solution to

$$\mathcal{A}_0^\varepsilon(u^\varepsilon, v) = (f, v) \quad \forall v \in H^2(\Omega) \cap H_0^1(\Omega). \tag{4.6}$$

With all of the above in place, we can now show (4.3) for  $\mathcal{L}_{0,h}^\varepsilon$ .

**THEOREM 4.1.** *Suppose  $\varepsilon = \mathcal{O}(h^\beta)$  with  $\beta \geq 2$ . Then there holds*

$$\|w_h\|_{H^2(\Omega)} \lesssim \|\mathcal{L}_{0,h}^\varepsilon w_h\|_{L^2(\Omega)} \quad \forall w_h \in V_h. \tag{4.7}$$

*Proof.* We first note that (4.7) is equivalent to

$$\|(\mathcal{L}_{0,h}^\varepsilon)^{-1} \varphi_h\|_{H^2(\Omega)} \lesssim \|\varphi_h\|_{L^2(\Omega)} \quad \forall \varphi_h \in V_h. \tag{4.8}$$

Fix  $\varphi_h \in V_h$ , and let  $w^\varepsilon$  be the weak solution to (4.6) with  $f = \varphi_h$  and  $w_h = (\mathcal{L}_{0,h}^\varepsilon)^{-1} \varphi_h \in V_h$ . Since  $w^\varepsilon$  is a weak solution of (4.6), we have  $w^\varepsilon \in H^2(\Omega) \cap H_0^1(\Omega)$  with

$$\|w^\varepsilon\|_{H^2(\Omega)} \lesssim \|\varphi_h\|_{L^2(\Omega)} \tag{4.9}$$

by (4.2). Since  $w_h^\varepsilon$  satisfies

$$\mathcal{A}_0^\varepsilon(w_h^\varepsilon, v_h) = (\varphi_h, v_h) \quad \forall v_h \in V_h,$$

we can apply (4.5) to get

$$\|\nabla(w^\varepsilon - w_h^\varepsilon)\|_{L^2(\Omega)} \leq \|w^\varepsilon - w_h^\varepsilon\|_E \lesssim (\sqrt{\varepsilon} + h)\|w^\varepsilon\|_{H^2(\Omega)} \lesssim h\|\varphi_h\|_{L^2(\Omega)}. \quad (4.10)$$

Let  $I_h^1$  denote the  $P_1$  continuous finite element interpolation operator on  $\mathcal{T}_h$ . By the finite element interpolation theory, we have

$$\|\nabla(w^\varepsilon - I_h^1 w^\varepsilon)\|_{L^2(\Omega)} \lesssim h\|w^\varepsilon\|_{H^2(\Omega)} \lesssim h\|\varphi_h\|_{L^2(\Omega)}, \quad (4.11)$$

Let  $\|\cdot\|_{H^2(\mathcal{T}_h)}$  denote the piecewise  $H^2$ -norm on  $\mathcal{T}_h$ . From (4.10), (4.11), the triangle inequality, and an inverse inequality, we have

$$\begin{aligned} \|w^\varepsilon - w_h^\varepsilon\|_{H^2(\Omega)} &\lesssim \|w^\varepsilon - I_h^1 w^\varepsilon\|_{H^2(\mathcal{T}_h)} + \|I_h^1 w^\varepsilon - w_h^\varepsilon\|_{H^2(\mathcal{T}_h)} \\ &\lesssim \|w^\varepsilon\|_{H^2(\Omega)} + \frac{1}{h}\|\nabla(I_h^1 w^\varepsilon - w_h^\varepsilon)\|_{L^2(\Omega)} \\ &\lesssim \|\varphi_h\|_{L^2(\Omega)} + \frac{1}{h}\|\nabla(I_h^1 w^\varepsilon - w^\varepsilon)\|_{L^2(\Omega)} + \frac{1}{h}\|\nabla(w^\varepsilon - w_h^\varepsilon)\|_{L^2(\Omega)} \\ &\lesssim \|\varphi_h\|_{L^2(\Omega)}. \end{aligned}$$

Hence,

$$\|w_h^\varepsilon\|_{H^2(\Omega)} \leq \|w^\varepsilon - w_h^\varepsilon\|_{H^2(\Omega)} + \|w^\varepsilon\|_{H^2(\Omega)} \lesssim \|\varphi_h\|_{L^2(\Omega)}.$$

Since  $w_h^\varepsilon = (\mathcal{L}_{0,h}^\varepsilon)^{-1}\varphi_h$ , we have (4.5). The proof is complete.  $\square$

**4.1.2. Stability Analysis of  $\mathcal{L}_h^\varepsilon$  for Continuous Coefficient  $A$ .** We need some estimates before continuing with the freezing coefficient argument. The first is to show  $\mathcal{L}_h^\varepsilon$  is bounded independent of  $h$  which is a result of the following lemma.

LEMMA 4.2. *Let  $\varepsilon = \mathcal{O}(h^\beta)$  for  $\beta \geq 2$  and  $D \subseteq \Omega$  be a subdomain. Then, for any  $v \in H^2(D) \cap H_0^1(D)$ , there holds*

$$\|\mathcal{L}_h^\varepsilon v\|_{L_h^2(D)} \lesssim \|v\|_{H^2(D)}. \quad (4.12)$$

*Proof.* Let  $w_h \in V_h(D)$ . By Hölder's inequality and an inverse inequality we have

$$\begin{aligned} (\mathcal{L}_h^\varepsilon v, w_h) &= \mathcal{A}^\varepsilon(v, w_h) = \varepsilon(\Delta v, \Delta w_h) - (A : D^2 v, w_h) \\ &\lesssim h^\beta \|v\|_{H^2(D)} \|w_h\|_{H^2(D)} + \|A\|_{L^\infty(D)} \|v\|_{H^2(D)} \|w_h\|_{L^2(D)} \\ &\lesssim h^{\beta-2} \|v\|_{H^2(D)} \|w_h\|_{L^2(D)} + \|v\|_{H^2(D)} \|w_h\|_{L^2(D)} \\ &\lesssim \|v\|_{H^2(D)} \|w_h\|_{L^2(D)}. \end{aligned} \quad (4.13)$$

Diving both sides of (4.13) by  $\|w_h\|_{L^2(D)}$  yields (4.12). The proof is complete.  $\square$

Next we need to show that  $\mathcal{L}_{0,h}^\varepsilon$  and  $\mathcal{L}_h^\varepsilon$  are locally close to each other.

LEMMA 4.3. *Let  $x_0 \in \Omega$  and let  $\delta > 0$ . Then there exists  $R_\delta > 0$  and  $h_\delta > 0$  such that*

$$\|(\mathcal{L}_{0,h}^\varepsilon - \mathcal{L}_h^\varepsilon)v_h\|_{L_h^2(B_{R_\delta})} \lesssim \delta \|v_h\|_{H^2(B_{R_\delta})} \quad \forall v_h \in V_h(B_{R_\delta}), \quad (4.14)$$



where  $B_{R_\delta}$  is the ball of radius  $R_\delta$  centered at  $x_0$  and  $h < h_\delta$ .

*Proof.* Since  $A$  is continuous and  $\bar{\Omega}$  is compact, it is uniformly continuous on  $\bar{\Omega}$ . Thus, given  $\delta > 0$ , there exists  $R_\delta > 0$  such that

$$\|A_0 - A\|_{L^\infty(B_{R_\delta})} \leq \delta. \tag{4.15}$$

Choose  $h_\delta$  such that  $V_h(B_{R_\delta})$  is nontrivial. Let  $w_h \in V_h(B_{R_\delta})$  with  $w_h \neq 0$ . Then

$$\begin{aligned} ((\mathcal{L}_{0,h}^\varepsilon - \mathcal{L}_h^\varepsilon)v_h, w_h) &= ((A_0 - A) : D^2v_h, w_h) \\ &\lesssim \|A_0 - A\|_{L^\infty(B_{R_\delta})} \|v_h\|_{H^2(B_{R_\delta})} \|w_h\|_{L^2(B_{R_\delta})}. \end{aligned}$$

Dividing by  $\|w_h\|_{L^2(B)}$  gives us (4.14). The proof is complete.  $\square$

Lastly, we state some super approximation results similar to those given in [11]. Since the proof is almost the same by verifying the estimates on each element, we omit it.

LEMMA 4.4. *Let  $I_h$  be a standard  $C^1$  (e.g. Argyris) nodal finite element interpolation operator (cf. [3]) and  $\eta \in C^\infty(\Omega)$  with  $\|\eta\|_{W^{j,\infty}(\Omega)} \lesssim d^{-j}$  for  $0 \leq j \leq k$  and some  $d > 0$ . Then, for any subdomain  $D \subseteq \Omega$ , we have*

$$\begin{aligned} \|\eta v_h - I_h(\eta v_h)\|_{L^2(D)} &\lesssim \frac{h}{d} \|v_h\|_{L^2(D)} \quad \forall v_h \in V_h, \\ \|\eta v_h - I_h(\eta v_h)\|_{H^2(D)} &\lesssim \frac{h}{d^3} \|v_h\|_{H^2(D)} \quad \forall v_h \in V_h. \end{aligned}$$

We also cite the following inverse inequality from [11].

LEMMA 4.5. *Let  $v_h \in V_h(D)$  where  $D \subseteq \Omega$  is a subdomain. Then we have*

$$\|v_h\|_{L^2(D)} \lesssim h^{-1} \|v_h\|_{H^{-1}(D)}.$$

To follow the freezing coefficient technique from here, we could employ arguments similar to the likes of Section 3.2.2 to achieve a Gårding-type estimate similar to (3.33); namely,

$$\|v_h\|_{H^2(\Omega)} \lesssim \|\mathcal{L}_h^\varepsilon v_h\|_{L_h^2(\Omega)} + \|v_h\|_{L^2(\Omega)}, \tag{4.16}$$

which is almost (4.3) aside from the  $L^2$  norm of  $v_h$  on the right hand side. To strip this term off in the PDE theory, we used the fact that  $\mathcal{L}^\varepsilon$  was injective. However, we do not have the tools available at the discrete level to say the same for  $\mathcal{L}_h^\varepsilon$ . To overcome this difficulty, our idea is to utilize a duality argument in order to achieve (4.3). Unfortunately, a standard duality argument would require an a-priori stability estimate of  $(\mathcal{L}^\varepsilon)^*$ , the adjoint of  $\mathcal{L}^\varepsilon$ , which is unknown at this time. Instead, we apply the freezing coefficient technique to  $(\mathcal{L}_h^\varepsilon)^*$ , the adjoint of  $\mathcal{L}_h^\varepsilon$ , for two reasons. First, since  $\mathcal{L}_h^\varepsilon$  lives on a finite dimensional vector space, the invertibility of  $\mathcal{L}_h^\varepsilon$  is equivalent to the invertibility of  $(\mathcal{L}_h^\varepsilon)^*$ . Second, a duality argument applied to  $(\mathcal{L}_h^\varepsilon)^*$  uses the stability of  $\mathcal{L}^\varepsilon$ , which we have already proved in Theorem 3.10.

Define  $(\mathcal{L}_h^\varepsilon)^* : V_h \rightarrow V_h$  by

$$((\mathcal{L}_h^\varepsilon)^* w_h, v_h) = (\mathcal{L}_h^\varepsilon v_h, w_h) \quad \forall v_h \in V_h.$$

The entirety of the non-standard duality argument is given in the lemma below.

LEMMA 4.6. *There exists  $h_0 > 0$  such that for all  $h < h_0$  and  $\varepsilon = \mathcal{O}(h^\beta)$  for some  $\beta \geq 2$ ,*

$$\|v_h\|_{L^2(\Omega)} \lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(\Omega)} \quad \forall v_h \in V_h. \quad (4.17)$$

*Proof.* Since the proof is long, we divide it into three steps.

*Step 1: Local Estimates for  $(\mathcal{L}_h^\varepsilon)^*$ .* Let  $x_0 \in \Omega$ . Given  $\delta > 0$ , define  $h_\delta$  and  $R_\delta$  as in Lemma 4.3. Let  $h < h_\delta$ ,  $R_1 = (1/3)R_\delta$ , and  $B_1$  be the ball centered at  $x_0$  with radius  $R_1$ . Let  $v_h \in V_h(B_{R_1})$ . From the assumption on  $\mathcal{L}^\varepsilon$ , there exists  $\varphi^\varepsilon \in H^2(\Omega) \cap H_0^1(\Omega)$  with

$$\mathcal{A}_0^\varepsilon(\varphi^\varepsilon, w) = (v_h^\varepsilon, w) \quad \forall w \in H^2(\Omega) \cap H_0^1(\Omega)$$

and

$$\|\varphi^\varepsilon\|_{H^2(\Omega)} \lesssim \|v_h\|_{L^2(\Omega)} = \|v_h\|_{L^2(B_1)}. \quad (4.18)$$

We also note that

$$(\mathcal{L}_h^\varepsilon \psi, w_h) = \mathcal{A}^\varepsilon(\psi, w_h) = (\mathcal{L}^\varepsilon \psi, w_h) \quad \forall \psi \in H^2(\Omega) \cap H_0^1(\Omega), w_h \in V_h. \quad (4.19)$$

Let  $\varphi_h^\varepsilon = (\mathcal{L}_{0,h}^\varepsilon)^{-1} v_h \in V_h$ . By (4.19), we have

$$\begin{aligned} \|v_h\|_{L^2(B_1)}^2 &= \|v_h\|_{L^2(\Omega)}^2 = (v_h, v_h) = (\mathcal{L}^\varepsilon \varphi^\varepsilon, v_h) = (\mathcal{L}_h^\varepsilon \varphi_h^\varepsilon, v_h) \\ &= (\mathcal{L}_h^\varepsilon \varphi_h^\varepsilon, v_h) + (\mathcal{L}_h^\varepsilon (\varphi^\varepsilon - \varphi_h^\varepsilon), v_h) \\ &= ((\mathcal{L}_h^\varepsilon)^* v_h, \varphi_h^\varepsilon) + (\mathcal{L}_{0,h}^\varepsilon (\varphi^\varepsilon - \varphi_h^\varepsilon), v_h) \\ &\quad + ((\mathcal{L}_h^\varepsilon - \mathcal{L}_{0,h}^\varepsilon) (\varphi^\varepsilon - \varphi_h^\varepsilon), v_h). \end{aligned} \quad (4.20)$$

Since  $\varphi_h^\varepsilon = (\mathcal{L}_{0,h}^\varepsilon)^{-1} v_h$ , we obtain the standard Galerkin orthogonality, namely

$$\begin{aligned} (\mathcal{L}_{0,h}^\varepsilon (\varphi^\varepsilon - \varphi_h^\varepsilon), v_h) &= \mathcal{A}_0^\varepsilon(\varphi^\varepsilon - \varphi_h^\varepsilon, v_h) = \mathcal{A}_0^\varepsilon(\varphi^\varepsilon, v_h) - \mathcal{A}_0^\varepsilon(\varphi_h^\varepsilon, v_h) \\ &= (v_h, v_h) - (v_h, v_h) = 0. \end{aligned} \quad (4.21)$$

Moreover, by Theorem 4.1, (4.21), and (4.18), we have

$$\begin{aligned} \|\varphi_h^\varepsilon\|_{H^2(\Omega)} &\lesssim \|\mathcal{L}_{0,h}^\varepsilon \varphi_h^\varepsilon\|_{L_h^2(\Omega)} \leq \|\mathcal{L}_{0,h}^\varepsilon \varphi^\varepsilon\|_{L_h^2(\Omega)} = \|\mathcal{L}_{0,h}^\varepsilon \varphi^\varepsilon\|_{L^2(\Omega)} \\ &\lesssim \|\varphi^\varepsilon\|_{H^2(\Omega)} \lesssim \|v_h\|_{L^2(B_1)}. \end{aligned} \quad (4.22)$$

Employing Lemma 4.3, (4.18), (4.20), and (4.22), we obtain

$$\begin{aligned} \|v_h\|_{L^2(B_1)}^2 &= ((\mathcal{L}_h^\varepsilon)^* v_h, \varphi_h^\varepsilon) + ((\mathcal{L}_h^\varepsilon - \mathcal{L}_{0,h}^\varepsilon) (\varphi^\varepsilon - \varphi_h^\varepsilon), v_h) \\ &\leq \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(\Omega)} \|\varphi_h^\varepsilon\|_{H^2(\Omega)} + \|(\mathcal{L}_h^\varepsilon - \mathcal{L}_{0,h}^\varepsilon) (\varphi^\varepsilon - \varphi_h^\varepsilon)\|_{L_h^2(\Omega)} \|v_h\|_{L^2(\Omega)} \\ &\lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(B_1)} \|v_h\|_{L^2(B_1)} + \delta \|(\varphi^\varepsilon - \varphi_h^\varepsilon)\|_{H^2(\Omega)} \|v_h\|_{L^2(B_1)} \\ &\lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(B_1)} \|v_h\|_{L^2(B_1)} + \delta (\|\varphi^\varepsilon\|_{H^2(\Omega)} + \|\varphi_h^\varepsilon\|_{H^2(\Omega)}) \|v_h\|_{L^2(B_1)} \\ &\lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(B_1)} \|v_h\|_{L^2(B_1)} + \delta \|v_h\|_{L^2(B_1)}^2. \end{aligned}$$

Taking  $\delta$  sufficiently small and then dividing both sides by  $\|v_h\|_{L^2(B_1)}$  gives us a local version of (4.17), that is,

$$\|v_h\|_{L^2(B_1)} \lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(B_1)}. \quad (4.23)$$

*Step 2: A Gårding-type inequality.* Let  $R_3 = 3R_1$ , and  $\eta \in C^\infty(\Omega)$  be a cutoff function with the following properties:

$$0 \leq \eta \leq 1, \quad \eta|_{B_1} \equiv 1, \quad \eta|_{\Omega \setminus B_2} \equiv 0, \quad |\eta|_{W^{m,\infty}(\Omega)} = \mathcal{O}(R_1^{-m}). \quad (4.24)$$

Let  $v_h \in V_h$ . By (4.23) we have

$$\begin{aligned} \|v_h\|_{L^2(B_1)} &= \|\eta v_h\|_{L^2(B_1)} \leq \|\eta v_h - I_h(\eta v_h)\|_{L^2(B_1)} + \|I_h(\eta v_h)\|_{L^2(B_1)} \\ &\lesssim \|\eta v_h - I_h(\eta v_h)\|_{L^2(B_1)} + \|(\mathcal{L}_h^\varepsilon)^* I_h(\eta v_h)\|_{H_h^{-2}(B_1)} \\ &\lesssim \|\eta v_h - I_h(\eta v_h)\|_{L^2(B_1)} + \|(\mathcal{L}_h^\varepsilon)^*(I_h(\eta v_h) - \eta v_h)\|_{H_h^{-2}(B_1)} \\ &\quad + \|(\mathcal{L}_h^\varepsilon)^* \eta v_h\|_{H_h^{-2}(B_1)}. \end{aligned} \quad (4.25)$$

Using Lemma 4.2, we have

$$\begin{aligned} \|(\mathcal{L}_h^\varepsilon)^*(I_h(\eta v_h) - \eta v_h)\|_{H_h^{-2}(B_1)} &= \sup_{w_h \in V_h(B_1)} \frac{((\mathcal{L}_h^\varepsilon)^*(I_h(\eta v_h) - \eta v_h), w_h)}{\|w_h\|_{H^2(B_1)}} \\ &\lesssim \sup_{w_h \in V_h(B_1)} \frac{(\mathcal{L}_h^\varepsilon w_h, (I_h(\eta v_h) - \eta v_h))}{\|w_h\|_{H^2(B_1)}} \\ &\lesssim \sup_{w_h \in V_h(B_1)} \frac{\|\mathcal{L}_h^\varepsilon w_h\|_{L^2(B_1)} \|I_h(\eta v_h) - \eta v_h\|_{L^2(B_1)}}{\|w_h\|_{H^2(B_1)}} \\ &\lesssim \sup_{w_h \in V_h(B_1)} \frac{\|w_h\|_{H^2(B_1)} \|I_h(\eta v_h) - \eta v_h\|_{L^2(B_1)}}{\|w_h\|_{H^2(B_1)}} \\ &= \|I_h(\eta v_h) - \eta v_h\|_{L^2(B_1)}. \end{aligned}$$

Thus, (4.25) becomes

$$\|v_h\|_{L^2(B_1)} \lesssim \|\eta v_h - I_h(\eta v_h)\|_{L^2(B_1)} + \|(\mathcal{L}_h^\varepsilon)^* \eta v_h\|_{H_h^{-2}(B_1)}. \quad (4.26)$$

Using Lemmas 4.4 and 4.5, we have

$$\begin{aligned} \|v_h\|_{L^2(B_1)} &\lesssim \frac{h}{R_1} \|v_h\|_{L^2(B_3)} + \|(\mathcal{L}_h^\varepsilon)^* \eta v_h\|_{H_h^{-2}(B_3)} \\ &\lesssim \frac{1}{R_1} \|v_h\|_{H^{-1}(B_3)} + \|(\mathcal{L}_h^\varepsilon)^* \eta v_h\|_{H_h^{-2}(B_3)}. \end{aligned} \quad (4.27)$$

We now wish to remove  $\eta$  from  $\|(\mathcal{L}_h^\varepsilon)^* \eta v_h\|_{H_h^{-2}(B_3)}$  in (4.27). Let  $w_h \in V_h(B_3)$  and decompose  $(\mathcal{L}_h^\varepsilon)^* \eta v_h$  as follows:

$$\begin{aligned} ((\mathcal{L}_h^\varepsilon)^*(\eta v_h), w_h) &= (\mathcal{L}_h^\varepsilon w_h, \eta v_h) \\ &= (\mathcal{L}_h^\varepsilon w_h \eta, v_h) + [(\mathcal{L}_h^\varepsilon w_h, \eta v_h) - (\mathcal{L}_h^\varepsilon w_h \eta, v)] \\ &= (\mathcal{L}_h^\varepsilon (I_h(w_h \eta)), v_h) + (\mathcal{L}_h^\varepsilon (w_h \eta - I_h(w_h \eta)), v_h) \\ &\quad + [(\mathcal{L}_h^\varepsilon w_h, \eta v_h) - (\mathcal{L}_h^\varepsilon w_h \eta, v_h)] \\ &=: I_1 + I_2 + I_3. \end{aligned} \quad (4.28)$$

We now bound each  $I_i$  in order. For  $I_1$ , the stability of  $I_h$  and (4.24) give us

$$\begin{aligned} I_1 &= ((\mathcal{L}_h^\varepsilon)^* v_h, I_h(\eta w_h)) \lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(B_3)} \|I_h(\eta w_h)\|_{H^2(B_3)} \\ &\lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(B_3)} \|\eta w_h\|_{H^2(B_3)} \\ &\lesssim \frac{1}{R_1^2} \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(B_3)} \|w_h\|_{H^2(B_3)}. \end{aligned} \quad (4.29)$$

To bound  $I_2$ , we use Lemmas 4.4, 4.5, and 4.2 to obtain

$$\begin{aligned} I_2 &\lesssim \|\eta w_h - I_h(\eta w_h)\|_{H^2(B_3)} \|v_h\|_{L^2(B_3)} \lesssim \frac{h}{R_1^3} \|w_h\|_{H^2(B_3)} \|v_h\|_{L^2(B_3)} \\ &\lesssim \frac{1}{R_1^3} \|w_h\|_{H^2(B_3)} \|v_h\|_{H^{-1}(B_3)}. \end{aligned} \quad (4.30)$$

In order to bound  $I_3$ , we reintroduce the operator  $\mathcal{L}_{0,h}^\varepsilon$  and define  $\tilde{A} = A - A_0$ . We then write

$$\begin{aligned} I_3 &= (\mathcal{L}_h^\varepsilon w_h, \eta v_h) - (\mathcal{L}_h^\varepsilon w_h \eta, v_h) \\ &= (\mathcal{L}_{0,h}^\varepsilon w_h, \eta v_h) - (\mathcal{L}_{0,h}^\varepsilon \eta w_h, v_h) \\ &\quad + \left[ (\mathcal{L}_h^\varepsilon w_h, \eta v_h) - (\mathcal{L}_h^\varepsilon \eta w_h, v_h) - (\mathcal{L}_{0,h}^\varepsilon w_h, \eta v_h) + (\mathcal{L}_{0,h}^\varepsilon \eta w_h, v_h) \right] \\ &= \varepsilon \int_{B_3} v_h \Delta w_h \Delta \eta + 2 \Delta w_h (\nabla \eta \cdot \nabla v_h) - w_h \Delta \eta \Delta v_h - 2 \Delta w_h (\nabla \eta \cdot \nabla w_h) \, dx \\ &\quad - \int_{B_3} (w_h A_0 : D^2 \eta + (A_0 + A_0^T) \nabla \eta \cdot \nabla w_h) v_h \, dx \\ &\quad - \int_{B_3} (w_h \tilde{A} : D^2 \eta + (\tilde{A} + \tilde{A}^T) \nabla \eta \cdot \nabla w_h) v_h \, dx =: K_1 + K_2 + K_3. \end{aligned} \quad (4.31)$$

We now must bound each  $K_i$ . Since  $\varepsilon < Ch^2$  for some pure constant  $C > 0$ , we bound  $K_1$  using Lemma 4.5, standard inverse inequalities, and (4.24) as follows:

$$\begin{aligned} K_1 &\lesssim \frac{h^2}{R_1^2} \|v_h\|_{L^2(B_3)} \|w_h\|_{H^2(B_3)} + \frac{h_2}{R_1} \|v_h\|_{H^1(B_3)} \|w_h\|_{H^2(B_3)} \\ &\quad + h^2 \|\Delta v_h\|_{H^{-1}(B_3)} \|w_h \Delta \eta\|_{H^1(B_3)} + h^2 \|\Delta v_h\|_{H^{-1}(B_3)} \|\nabla w_h \cdot \nabla \eta\|_{H^1(B_3)} \\ &\lesssim \frac{h}{R_1^2} \|v_h\|_{H^{-1}(B_3)} \|w_h\|_{H^2(B_3)} + \frac{1}{R_1} \|v_h\|_{H^{-1}(B_3)} \|w_h\|_{H^2(B_3)} \\ &\quad + \frac{h^2}{R_1^3} \|\nabla v_h\|_{L^2(B_3)} \|w_h\|_{H^2(B_3)} + \frac{h^2}{R_1^2} \|\nabla v_h\|_{L^2(B_3)} \|w_h\|_{H^2(B_3)} \\ &\lesssim \frac{h}{R_1^2} \|v_h\|_{H^{-1}(B_3)} \|w_h\|_{H^2(B_3)} + \frac{1}{R_1} \|v_h\|_{H^{-1}(B_3)} \|w_h\|_{H^2(B_3)} \\ &\quad + \frac{1}{R_1^3} \|v_h\|_{H^{-1}(B_3)} \|w_h\|_{H^2(B_3)} + \frac{1}{R_1^2} \|v_h\|_{H^{-1}(B_3)} \|w_h\|_{H^2(B_3)} \\ &\lesssim \frac{1}{R_1^3} \|w_h\|_{H^2(B_3)} \|v_h\|_{H^{-1}(B_3)}. \end{aligned} \quad (4.32)$$

To bound  $K_2$ , we have

$$\begin{aligned} K_2 &\lesssim \left( \|w_h A_0 : D^2 \eta\|_{H^1(B_3)} + \|(A_0 + A_0^T) \nabla \eta \cdot \nabla w_h\|_{H^1(B_3)} \right) \|v_h\|_{H^{-1}(B_3)} \\ &\lesssim \frac{1}{R_1^3} \|w_h\|_{H^2(B_3)} \|v_h\|_{H^{-1}(B_3)}. \end{aligned} \quad (4.33)$$

Bounding  $K_3$  requires Hölder's inequality and  $\|\tilde{A}\|_{L^\infty(B_3)} \leq \delta$  which gives us

$$\begin{aligned} K_3 &\lesssim \left( \|w_h \tilde{A} : D^2 \eta\|_{L^2(B_3)} + \|(\tilde{A} + \tilde{A}^T) \nabla \eta \cdot \nabla w_h\|_{L^2(B_3)} \right) \|v_h\|_{L^2(B_3)} \\ &\lesssim \delta \left( \frac{1}{R_1^2} \|w_h\|_{L^2(B_3)} + \frac{1}{R_1} \|w_h\|_{H^1(B_3)} \right) \|v_h\|_{L^2(B_3)}. \end{aligned} \quad (4.34)$$

From Lemma C.1 of [8] we have

$$\|w_h\|_{H^l(B_3)} \lesssim R_1^{2-l} \|w_h\|_{H^2(B_3)}, \quad l = 0, 1. \quad (4.35)$$

Combining (4.34) and (4.35) yields

$$K_3 \lesssim \delta \|w_h\|_{H^2(B_3)} \|v_h\|_{L^2(B_3)}. \quad (4.36)$$

From (4.32), (4.33), and (4.36) we get

$$I_3 \lesssim \frac{1}{R_1^3} \|w_h\|_{H^2(B_3)} \|v_h\|_{H^{-1}(B_3)} + \delta \|w_h\|_{H^2(B_3)} \|v_h\|_{L^2(B_3)}, \quad (4.37)$$

and from (4.29), (4.30), and (4.37), we have

$$\begin{aligned} ((\mathcal{L}_h^\varepsilon)^*(\eta v_h), w_h) &\lesssim \frac{1}{R_1^3} \left( \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(B_3)} + \|v_h\|_{H^{-1}(B_3)} \right) \|w_h\|_{H^2(B_3)} \\ &\quad + \delta \|w_h\|_{H^2(B_3)} \|v_h\|_{L^2(B_3)}. \end{aligned}$$

Thus,

$$\begin{aligned} \|(\mathcal{L}_h^\varepsilon)^*(\eta v_h)\|_{H_h^{-2}(B_3)} &\lesssim \frac{1}{R_1^3} \left( \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(B_3)} + \|v_h\|_{H^{-1}(B_3)} \right) \\ &\quad + \delta \|v_h\|_{H^{-1}(B_3)}. \end{aligned} \quad (4.38)$$

From (4.27) and (4.38) we obtain

$$\|v_h\|_{L^2(B_1)} \lesssim \frac{1}{R_1^3} \left( \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(B_3)} + \|v_h\|_{H^{-1}(B_3)} \right) + \delta \|v_h\|_{H^{-1}(B_3)}.$$

Since  $\bar{\Omega}$  is compact and  $\delta$  is independent of  $R_1$ , by employing a covering argument we can show that

$$\|v_h\|_{L^2(\Omega)} \lesssim \frac{1}{R_1^3} \left( \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(\Omega)} + \|v_h\|_{H^{-1}(\Omega)} \right) + \delta \|v_h\|_{H^{-1}(\Omega)}. \quad (4.39)$$

Choosing  $\delta$  sufficiently small in (4.39) produces a Gårding type estimate, namely

$$\|v_h\|_{L^2(\Omega)} \lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(\Omega)} + \|v_h\|_{H^{-1}(\Omega)}. \quad (4.40)$$

*Step 3: A nonstandard duality argument.* We now wish to remove  $\|v_h\|_{H^{-1}(\Omega)}$  in (4.40). To control this term we use a nonstandard duality argument on the operator  $(\mathcal{L}_h^\varepsilon)^*$ . Let

$$X = \{g \in H_0^1(\Omega) : \|\nabla g\|_{L^2(\Omega)} = 1\}.$$

By the Poincaré inequality, there exists a constant  $C = C(\Omega)$  such that

$$\|g\|_{L^2(\Omega)} \leq C \|\nabla g\|_{L^2(\Omega)} < \infty.$$

Thus,  $X$  is precompact in  $H^1(\Omega)$  by Sobolev embedding. Now we define

$$W = \{(\mathcal{L}^\varepsilon)^{-1}g : g \in X\}.$$

Let  $\varphi \in W$  with  $\mathcal{L}^\varepsilon \varphi = g$ . Then by (4.2) we have

$$\|\varphi\|_{H^2(\Omega)} \lesssim \|g\|_{L^2(\Omega)} < \infty. \quad (4.41)$$

Thus,  $W$  is a bounded set in  $H^2(\Omega) \cap H_0^1(\Omega)$ . Let  $\tau > 0$ . Then from [15, Lemma 5] there exists  $h_* > 0$  only dependent upon  $\tau$  and the closure of  $W$  such that for every  $\varphi \in W$  and  $0 < h \leq h_*$ , there exists a  $\varphi_h \in V_h$  such that

$$\|\varphi - \varphi_h\|_{H^2(\Omega)} \leq \tau. \quad (4.42)$$

Note by the triangle inequality, (4.41), and (4.42), we have

$$\|\varphi_h\|_{H^2(\Omega)} \leq \|\varphi\|_{H^2(\Omega)} + \|\varphi_h - \varphi\|_{H^2(\Omega)} \lesssim C + \tau.$$

Thus, the set

$$\{\varphi_h \in V_h : \|\varphi_h - \varphi\|_{H^2(\Omega)} \leq \tau\}$$

is uniformly bounded in  $\varphi$  and  $h$ . Let  $g \in X$  and choose  $\varphi_g \in W$  with  $\mathcal{L}^\varepsilon \varphi_g = g$ . For  $v_h \in V_h$  and  $\varphi_h \in V_h$  satisfying (4.42), we have

$$\begin{aligned} (v_h, g) &= (g, v_h) \\ &= (\mathcal{L}^\varepsilon \varphi_g, v_h) = (\mathcal{L}_h^\varepsilon \varphi_g, v_h) = (\mathcal{L}_h^\varepsilon \varphi_h, v_h) + (\mathcal{L}_h^\varepsilon (\varphi_g - \varphi_h), v_h) \\ &= ((\mathcal{L}_h^\varepsilon)^* v_h, \varphi_h) + (\mathcal{L}_h^\varepsilon (\varphi_g - \varphi_h), v_h) \\ &\lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(\Omega)} \|\varphi_h\|_{H^2(\Omega)} + \|\varphi_g - \varphi_h\|_{H^2(\Omega)} \|v_h\|_{L^2(\Omega)} \\ &\lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(\Omega)} + \tau \|v_h\|_{L^2(\Omega)}. \end{aligned} \quad (4.43)$$

Taking the supremum of both sides of (4.43) over  $g \in X$  gives us

$$\|v_h\|_{H^{-1}(\Omega)} \lesssim \|(\mathcal{L}_h^\varepsilon)^* v_h\|_{H_h^{-2}(\Omega)} + \tau \|v_h\|_{L^2(\Omega)}. \quad (4.44)$$

Combining (4.40) and (4.44) and choosing  $\tau$  sufficiently small gives us (4.17) upon taking  $h_0 = \min\{h_\delta, h_*\}$ .  $\square$

Lemma 4.6 allows us to prove the discrete Calderon-Zygmund estimate (4.3). This is shown in the Lemma below.

LEMMA 4.7. *Let  $h_0$  be from Lemma 4.6 and  $\varepsilon = \mathcal{O}(h^\beta)$  for some  $\beta \geq 2$ . Then*

$$\|v_h\|_{H^2(\Omega)} \lesssim \|\mathcal{L}_h^\varepsilon v_h\|_{L_h^2(\Omega)} \quad \forall v_h \in V_h. \quad (4.45)$$

*Proof.* We first show  $(\mathcal{L}_h^\varepsilon)^*$  is 1-1. Suppose there is a  $w_h \in V_h$  with  $(\mathcal{L}_h^\varepsilon)^* w_h = 0$ . By Lemma 4.6, we have

$$\|w_h\|_{L^2(\Omega)} \lesssim \|(\mathcal{L}_h^\varepsilon)^* w_h\|_{H_h^{-2}(\Omega)} = 0. \quad (4.46)$$

Thus  $w_h = 0$ , and it follows that  $(\mathcal{L}_h^\varepsilon)^*$  is 1-1. Since  $(\mathcal{L}_h^\varepsilon)^* : V_h \rightarrow V_h$ , it is invertible on  $V_h$ . Let  $v_h \in V_h$ . Consider the problem of finding  $q_h \in V_h$  such that

$$(w_h, (\mathcal{L}_h^\varepsilon)^* q_h) = (D^2 v_h, D^2 w_h) + (\nabla v_h, \nabla w_h) + (v_h, w_h). \tag{4.47}$$

Since  $(\mathcal{L}_h^\varepsilon)^*$  is invertible, such a  $q_h$  exists. By (4.17) we have

$$\begin{aligned} \|q_h\|_{L^2(\Omega)} &\lesssim \|(\mathcal{L}_h^\varepsilon)^* q_h\|_{H_h^{-2}(\Omega)} = \sup_{w_h \in V_h} \frac{(w_h, (\mathcal{L}_h^\varepsilon)^* q_h)}{\|w_h\|_{H^2(\Omega)}} \\ &\lesssim \sup_{w_h \in V_h} \left( \frac{\|D^2 v_h\|_{L^2(\Omega)} \|D^2 w_h\|_{L^2(\Omega)} + \|\nabla v_h\|_{L^2(\Omega)} \|\nabla w_h\|_{L^2(\Omega)}}{\|w_h\|_{H^2(\Omega)}} \right. \\ &\quad \left. + \frac{\|v_h\|_{L^2(\Omega)} \|w_h\|_{L^2(\Omega)}}{\|w_h\|_{H^2(\Omega)}} \right) \\ &\lesssim \|v_h\|_{H^2(\Omega)}. \end{aligned} \tag{4.48}$$

From (4.48) we obtain

$$\begin{aligned} \|v_h\|_{H^2(\Omega)}^2 &= (D^2 v_h, D^2 v_h) + (\nabla v_h, \nabla v_h) + (v_h, v_h) = (v_h, (\mathcal{L}_h^\varepsilon)^* q_h) = (\mathcal{L}_h^\varepsilon v_h, q_h) \\ &\leq \|\mathcal{L}_h^\varepsilon v_h\|_{L_h^2(\Omega)} \|q_h\|_{L^2(\Omega)} \lesssim \|\mathcal{L}_h^\varepsilon v_h\|_{L_h^2(\Omega)} \|v_h\|_{H^2(\Omega)}. \end{aligned} \tag{4.49}$$

Dividing both sides of (4.49) by  $\|v_h\|_{H^2(\Omega)}$  gives us (4.45). The proof is complete.  $\square$

We now present the well-posedness of our  $C^1$  finite element scheme.

**THEOREM 4.8.** *Let  $h_0$  be from Lemma 4.6 and  $\varepsilon = \mathcal{O}(h^\beta)$  for some  $\beta \geq 2$ . Then there is a unique  $u_h^\varepsilon$  satisfying (4.1) and*

$$\|u_h^\varepsilon\|_{H^2(\Omega)} \lesssim \|f\|_{L^2(\Omega)}. \tag{4.50}$$

*Proof.* From a similar argument to (4.46), we assert that  $\mathcal{L}_h^\varepsilon$  is 1-1. Since  $\mathcal{L}_h^\varepsilon : V_h \rightarrow V_h$ , it is invertible. Thus, given  $f \in L^2(\Omega)$ , there is a unique  $u_h^\varepsilon \in V_h$  satisfying (4.1). By (4.45) we have

$$\|u_h^\varepsilon\|_{H^2(\Omega)} \lesssim \|\mathcal{L}_h^\varepsilon u_h\|_{L_h^2(\Omega)} = \sup_{v_h \in V_h} \frac{(\mathcal{L}_h^\varepsilon u_h, v_h)}{\|v_h\|_{L^2(\Omega)}} = \sup_{v_h \in V_h} \frac{(f, v_h)}{\|v_h\|_{L^2(\Omega)}} \leq \|f\|_{L^2(\Omega)},$$

yielding (4.50). The proof is complete.  $\square$

Since  $\mathcal{L}_h^\varepsilon$  is consistent, we have a C ea type estimate which asserts an optimal convergence in the  $H^2(\Omega)$ -norm:

**THEOREM 4.9.** *Suppose  $\varepsilon = \mathcal{O}(h^\beta)$  for some  $\beta \geq 2$ . Let  $u^\varepsilon \in H^2(\Omega) \cap H_0^1(\Omega)$  and  $u_h^\varepsilon \in V_h$  satisfy (4.1) and (3.2), respectively. Then*

$$\|u^\varepsilon - u_h^\varepsilon\|_{H^2(\Omega)} \lesssim \inf_{v_h \in V_h} \|u^\varepsilon - v_h\|_{H^2(\Omega)}. \tag{4.51}$$

Moreover, if  $u^\varepsilon \in H^s(\Omega)$  for  $s \geq 2$ , we have

$$\|u^\varepsilon - u_h^\varepsilon\|_{H^2(\Omega)} \lesssim h^{r-2} \|u^\varepsilon\|_{H^r(\Omega)}, \tag{4.52}$$

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

*Proof.* By the consistency of  $\mathcal{L}_h^\varepsilon$ , we have the usual Galerkin orthogonality:

$$\begin{aligned} (\mathcal{L}_h^\varepsilon(u^\varepsilon - u_h^\varepsilon), v_h) &= \mathcal{A}^\varepsilon(u^\varepsilon - u_h^\varepsilon, v_h) \\ &= (f, v_h) - (f, v_h) = 0 \quad \forall v_h \in V_h. \end{aligned} \quad (4.53)$$

Let  $v_h \in V_h$ . By Theorem 4.8, Lemma 4.2, and (4.53), we have

$$\begin{aligned} \|u_h^\varepsilon - v_h\|_{H^2(\Omega)} &\lesssim \|\mathcal{L}_h^\varepsilon(u_h^\varepsilon - v_h)\|_{L_h^2(\Omega)} = \sup_{w_h \in V_h} \frac{(\mathcal{L}_h^\varepsilon(u_h^\varepsilon - v_h), w_h)}{\|w_h\|_{L^2(\Omega)}} \\ &\lesssim \sup_{w_h \in V_h} \frac{(\mathcal{L}_h^\varepsilon(u^\varepsilon - v_h), w_h)}{\|w_h\|_{L^2(\Omega)}} \lesssim \|\mathcal{L}_h^\varepsilon(u^\varepsilon - v_h)\|_{L_h^2(\Omega)} \\ &\lesssim \|u^\varepsilon - v_h\|_{H^2(\Omega)}. \end{aligned} \quad (4.54)$$

Hence, by the triangle inequality and (4.54), we obtain

$$\|u^\varepsilon - u_h^\varepsilon\|_{H^2(\Omega)} \lesssim \|u^\varepsilon - v_h\|_{H^2(\Omega)} + \|u_h^\varepsilon - v_h\|_{H^2(\Omega)} \lesssim \|u^\varepsilon - v_h\|_{H^2(\Omega)}. \quad (4.55)$$

Taking the infimum of (4.55) over all  $v_h \in V_h$  yields (4.51). Estimate (4.52) follows from taking  $v_h = I_h u$  and using the standard interpolation estimates. The proof is complete.  $\square$

**4.2. Numerical Experiments.** In this section, we present a series of 2-D numerical tests using the  $C^1$ -conforming (fifth order) Argyris finite element space. All of the tests are performed using the COMSOL software package. The first three tests correspond to choosing the continuous, positive definite matrix

$$A(x, y) = \begin{bmatrix} (2x - y)^{1/3} + 4e^{2-x} & \frac{1}{2} \sin(10xy) - \frac{1}{2}(x+2)^{1/2} \\ \frac{1}{2} \sin(10xy) - \frac{1}{2}(x+2)^{1/2} & |y-2x|^{1/4} + 3 \end{bmatrix} \quad (4.56)$$

and  $b, c = 0$  in (P). The fourth test will correspond to choosing a degenerate elliptic matrix  $A$  in (4.1) to gauge the impact of the method for approximating “harder” problems with viscosity solutions instead of strong solutions. We use a very fine mesh in all of the tests so that the error is dominated by the vanishing moment approximation to the PDE problem. We will observe that the error is maximized along sets where the solution is not as regular and along the boundary due to the auxiliary boundary condition  $\Delta u^\varepsilon = 0$ . The measured error does not appear to correspond to sets where the coefficient matrix  $A$  is not as well-behaved. We also observe better rates of convergence with respect to  $\varepsilon$  than the rates guaranteed by Theorem 3.13. Numerically we observe linear convergence in the  $L^2$ -norm.

**Test 1.** Consider (P) with  $A$  defined by (4.56),  $\Omega = (-2, 2) \times (-2, 2)$ , and solution  $u(x, y) = \frac{1}{6}|x|^3 \cos(y)$ , where  $f$  and  $g$  are chosen accordingly. The given test problem has a solution in  $H^3(\Omega)$ , where the third-order partial derivative with respect to  $x$  is discontinuous along the line  $x = 0$ . The results for varying  $\varepsilon$  can be found in Table 4.1. We can also see from Figure 4.1 that the error is largest along the line  $x = 0$ , as expected.

**Test 2.** Consider (P) with  $A$  defined by (4.56),  $\Omega = (-2, 2) \times (-2, 2)$ , and solution  $u(x, y) = \frac{1}{2}x|x| \cos(y)$ , where  $f$  and  $g$  are chosen accordingly. The given test problem has a solution in  $H^2(\Omega)$ , where the second-order derivative is discontinuous along the line  $x = 0$ . The results for varying  $\varepsilon$  can be found in Table 4.2. Figure 4.2



TABLE 4.1: Rates of convergence for Test 1 using the vanishing moment method.

$\varepsilon$	$\ u_h^\varepsilon - u\ _{L^2(\Omega)}$	Order	$\ \nabla(u_h^\varepsilon - u)\ _{L^2(\Omega)}$	Order	$\ \Delta(u_h^\varepsilon - u)\ _{L^2(\Omega)}$	Order
4e-2	9.44e-3		2.25e-2		2.55e-1	
2e-2	4.89e-3	0.95	1.32e-2	0.76	2.15e-1	0.24
1e-2	2.50e-3	0.96	7.84e-3	0.76	1.82e-1	0.24
5e-3	1.27e-3	0.98	4.64e-3	0.76	1.54e-1	0.25
2.5e-7	1.65e-7		2.09e-5		3.80e-3	

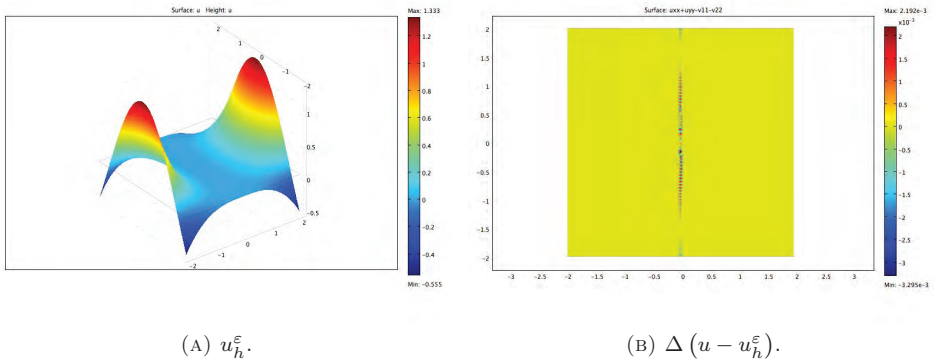


FIG. 4.1: Computed solution and corresponding error for Test 1 using the vanishing moment method.

shows that the error is once again largest along the line  $x = 0$ . The figure also shows the boundary layer due to the high-order auxiliary boundary condition.

**Test 3.** Consider  $(P)$  with  $A$  defined by (4.56),  $\Omega = B_2(0)$  (the ball of radius 2 centered at the origin), and solution  $u(x, y) = (x - y)^{8/3}$ , where  $f$  and  $g$  are chosen accordingly. Observe that the solution is once again in  $H^2(\Omega)$ , where the second order derivatives have a cusp along the line  $x = y$ . The results for varying  $\varepsilon$  can be found in Table 4.3. We can also see the finite element method does not converge to  $u$  when using  $\varepsilon = 0$ , which verifies the fact that trivial  $H^2$  conforming finite element methods do not work in general for second order linear problems of non-divergence form even when approximating an  $H^2$  solution. The plot of an approximation can be found in Figure 4.3.

It should be noted that for many biharmonic problems with the simply supported boundary condition, approximating the circular domain  $\Omega$  by a sequence of inscribed polygonal domains  $\{\Omega_h\}$ , as we have done in this numerical test, can lead to the so-called Babuška's paradox [2, 6] in the sense that the limiting solution may not satisfy the simply supported boundary condition on the circular boundary but, instead, satisfies the clamped boundary condition. As a result, problems with this paradox fail

TABLE 4.2: Rates of convergence for Test 2 using the vanishing moment method.

$\varepsilon$	$\ u_h^\varepsilon - u\ _{L^2(\Omega)}$	Order	$\ \nabla(u_h^\varepsilon - u)\ _{L^2(\Omega)}$	Order	$\ \Delta(u_h^\varepsilon - u)\ _{L^2(\Omega)}$	Order
4e-2	4.30e-3		2.26e-2		3.99e-1	
2e-2	2.30e-3	0.90	1.42e-2	0.67	3.38e-1	0.24
1e-2	1.22e-3	0.91	8.79e-3	0.69	2.86e-1	0.24
5e-7	2.25e-4		9.09e-4		1.07e-1	

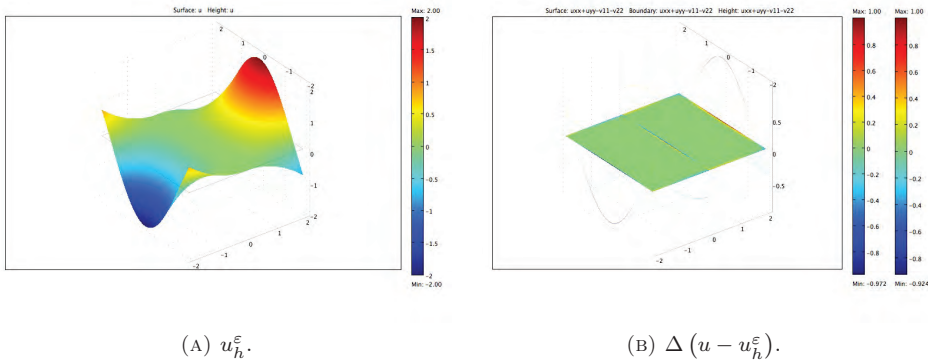


FIG. 4.2: Computed solution and corresponding error for Test 2 using the vanishing moment method.

to converge (due to the failed convergence of the simply supported boundary condition) as the approximate polygonal domains  $\{\Omega_h\}$  converge in measure to the curved boundary domain  $\Omega$ . However, we observe from this numerical test that the application of the VMM with an artificial simply supported boundary condition does not suffer the Babuška's paradox, as seen from the error  $u - u_h^\varepsilon$ ; this is largely because we are driving  $\varepsilon \rightarrow 0$  in tandem with  $h$ . Also, the flexibility of the VMM allows alternate choices for the auxiliary boundary condition such as a clamped BC to be used because its convergence does not depend on the choice of the artificial (higher order) boundary condition. As such, the VMM provides a built-in mechanism to completely avoid the difficulties associated with the Babuška's paradox.

TABLE 4.3: Rates of convergence for Test 3 using the vanishing moment method.

$\varepsilon$	$\ u_h^\varepsilon - u\ _{L^2(\Omega)}$	Order	$\ \nabla(u_h^\varepsilon - u)\ _{L^2(\Omega)}$	Order	$\ \Delta(u_h^\varepsilon - u)\ _{L^2(\Omega)}$	Order
5e-3	1.98e-2		1.22e-1		4.27	
2.5e-3	1.01e-2	0.98	7.32e-2	0.73	3.59	0.25
1.25e-3	5.08e-3	0.98	4.40e-2	0.73	3.03	0.24
0	5.40e3		1.83e6		4.99e8	

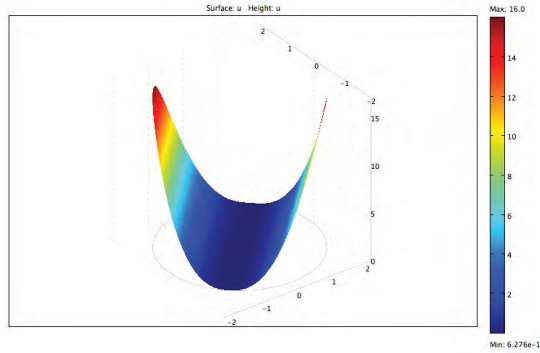


FIG. 4.3: *Computed solution for Test 3 using the vanishing moment method.*

**Test 4.** Consider  $(P)$  with  $\Omega = (-2, 2) \times (-2, 2)$ ,

$$A(x, y) = \frac{16}{9} \begin{bmatrix} x^{2/3} & -x^{1/3}y^{1/3} \\ -x^{1/3}y^{1/3} & y^{2/3} \end{bmatrix},$$

and solution  $u(x, y) = x^{4/3} - y^{4/3} \in H^1(\Omega)$ , where  $f$  and  $g$  chosen accordingly. For the above example, we have  $u \notin H^2(\Omega)$  and  $A$  is not uniformly elliptic. Thus,  $u$  is only a viscosity solution, not a strong solution. We see in Table 4.4 that the vanishing moment method appears to be working, although with unknown deteriorated rates of convergence. From Figure 4.4, we see that the finite element method with Argyris finite element space does not work for the given example. Thus, we can see that the vanishing moment method has strong potential for approximating more general second order problems that are understood in the viscosity solution framework.

TABLE 4.4: *Rates of convergence for Test 4 using the vanishing moment method.*

$\varepsilon$	$\ u_h^\varepsilon - u\ _{L^2(\Omega)}$	Order	$\ \nabla(u_h^\varepsilon - u)\ _{L^2(\Omega)}$	Order
2e-4	6.89e-2		3.94e-1	
1e-4	6.10e-2	0.18	3.53e-1	0.16
5e-5	5.46e-2	0.16	3.15e-1	0.17

## REFERENCES

- [1] S. AGMON, *Lectures on Elliptic Boundary Value Problems*, American Mathematical Society, 2010.
- [2] I. BABUŠKA, *The theory of small changes in the domain of existence in the theory of partial differential equations and its applications*, Differential equations and their applications, Proceedings of the Conference held in Prague in September 1962. Publishing House of the Czechoslovak Academy of Sciences, Prague, pp. 13–26, 1963.

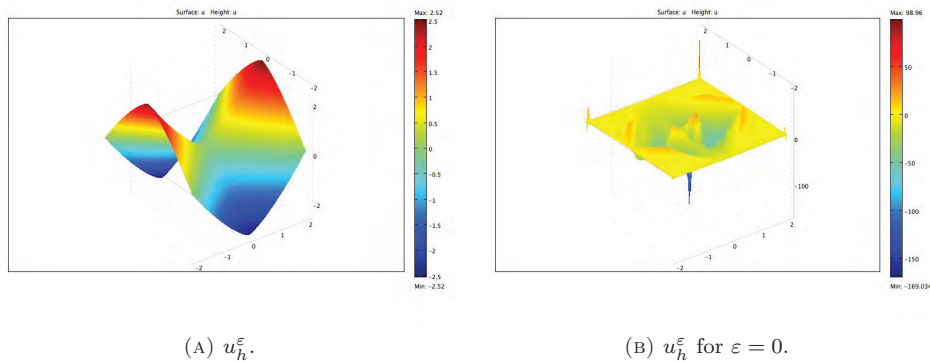


FIG. 4.4: Computed solutions for Test 4 using the vanishing moment method.

- [3] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*, Springer, New York, 2008.
- [4] L. A. CAFFARELLI AND C. E. GUTIÉRREZ, *Properties of the solutions of the linearized Monge-Ampère equation*, *American Journal of Mathematics*, pp. 423–465, 1997.
- [5] M. G. CRANDALL, H. ISHII, AND P.-L. LIONS, *User's guide to viscosity solutions of second order partial differential equations*, *Bulletin of the American Mathematical Society*, 27:1 (1992), pp. 1–67.
- [6] C. DAVINI, *Gaussian Curvature and Babuška's Paradox in the Theory of Plates*, pp. 67–87. Springer Milan, Milano, 2003.
- [7] L. C. EVANS, *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*, American Mathematical Society, Providence, RI, 2010.
- [8] X. FENG, L. HENNINGS, AND M. NEILAN, *Finite element methods for second order linear elliptic partial differential equations in non-divergence form*, *Mathematics of Computation*, 86 (2017), pp. 2025–2051.
- [9] X. FENG AND M. NEILAN, *Mixed finite element methods for the fully nonlinear monge-ampère equation based on the vanishing moment method*, *SIAM Journal on Numerical Analysis*, 47:2 (2009), pp. 1226–1250.
- [10] X. FENG AND M. NEILAN, *Vanishing moment method and moment solutions for fully nonlinear second order partial differential equations*, *Journal of Scientific Computing*, 38:1 (2009), 74.
- [11] X. FENG, M. NEILAN, AND S. SCHNAKE, *Interior penalty discontinuous Galerkin methods for second order linear non-divergence form elliptic PDEs*, *Journal of Scientific Computing*, 74:3 (2018), pp. 1651–1676.
- [12] W. H. FLEMING AND H. M. SONER, *Controlled Markov Processes and Viscosity Solutions*, Springer, Berlin, 2010.
- [13] D. GILBARG AND N. S. TRUDINGER, *Elliptic Partial Differential Equations of Second Order*, *Classics in Mathematics*. Springer, Berlin, 2001.
- [14] R. H. NOCHETTO AND W. ZHANG, *Discrete ABP estimate and convergence rates for linear elliptic equations in non-divergence form*, *Foundations of Computational Mathematics*, 18:3 (2018), pp. 537–593.
- [15] A. H. SCHATZ AND J. WANG, *Some new error estimates for Ritz-Galerkin methods with minimal regularity assumptions*, *Math. Comp.*, 65:213 (1996), pp. 19–27.
- [16] I. SMEARS AND E. SÜLI, *Discontinuous Galerkin finite element approximation of nondivergence form elliptic equations with Cordes coefficients*, *SIAM Journal on Numerical Analysis*, 51:4 (2013), pp. 2088–2106.
- [17] C. WANG AND J. WANG, *A primal-dual weak Galerkin finite element method for second order elliptic equations in non-divergence form*, *Math. Comp.*, 87:310 (2018), pp. 515–545.