Mathematical Modelling and Numerical Analysis Will be set by the publisher
Modélisation Mathématique et Analyse Numérique

VARIATIONAL AND NUMERICAL ANALYSIS OF A Q-TENSOR MODEL FOR SMECTIC-A LIQUID CRYSTALS *

JINGMIN XIA [1] AND PATRICK E. FARRELL [2]

Abstract. We analyse an energy minimisation problem recently proposed for modelling smectic-A liquid crystals. The optimality conditions give a coupled nonlinear system of partial differential equations, with a second-order equation for the tensor-valued nematic order parameter $\mathbf{Q}$ and a fourth-order equation for the scalar-valued smectic density variation u. Our two main results are a proof of the existence of solutions to the minimisation problem, and the derivation of a priori error estimates for its discretisation of the decoupled case (i.e., $q = 0$) using the $\mathscr{C}^0$ interior penalty method. More specifically, optimal rates in the H^1 and L^2 norms are obtained for $\mathbf{Q}$, while optimal rates in a mesh-dependent norm and L^2 norm are obtained for u. Numerical experiments confirm the rates of convergence.

1991 Mathematics Subject Classification. 76A15, 49J10, 35B45, 65N12, 65N30.

The dates will be set by the publisher.

INTRODUCTION

Smectic liquid crystals (LC) are layered mesophases that have a periodic modulation of the mass density along one spatial direction. Informally, they can be thought of as one-dimensional solids along the direction of periodicity and two-dimensional fluids along the other two remaining directions. According to different in-layer structures, several types of smectic LC are recognised, e.g., smectic-A, smectic-B and smectic-C (see [26, Figure 9] for illustrations of different smectic phases). In particular, in the smectic-A phase, the molecules tend to align parallel to the normals of the smectic layers. For a broader review of liquid crystals, see [2, 16, 28].

Several models have been proposed to describe smectic-A liquid crystals. The classical de Gennes model [15] employs a complex-valued order parameter to describe the magnitude and phase of the density variation. This complex-valued parameterisation leads to some key modelling difficulties, which motivated the development by Pevnyi, Selinger & Sluckin (PSS) of a smectic-A model with a real-valued smectic order parameter and a vector-valued nematic order parameter [24]. The use of a vector-valued nematic order parameter limits the kinds of defects the model can permit [2], and hence Ball & Bedford (BB) proposed a version of the PSS model employing a tensor-valued nematic order parameter [3] instead. The model considered in this work is similar to the BB model, but with additional modifications to render it amenable to numerical simulation (see [31] for details). The model was used to numerically simulate several key smectic defect structures, such as oily streaks and focal conic domains, for the first time.

Keywords and phrases: $\mathscr{C}^0$ interior penalty method, a priori error estimates, finite element methods, smectic liquid crystals

* *The work of JX is supported by the National Science Foundation of China (No. 12201636) and the Research Fund of National University of Defense Technology [grant number ZK22-37]. The work of PEF is supported by EPSRC [grant numbers EP/R029423/1 and EP/W026163/1].*

[1] College of Meteorology and Oceanography, National University of Defense Technology, Changsha, China; e-mail: `jingmin.xia@nudt.edu.cn`

[2] Mathematical Institute, University of Oxford, Oxford, UK; e-mail: `patrick.farrell@maths.ox.ac.uk`

While the numerical modelling of nematic liquid crystals is now mature, relatively little work has considered smectics. Garcìa-Cervera and Joo [19] perform numerical simulations of the classical de Gennes model [15] in the presence of a magnetic field, using a combined Fourier-finite difference approach. Wittmann et al. use density functional theory to investigate the topology of smectic liquid crystals in complex confinement [30]. Monderkamp et al. examine the topology of defects in two-dimensional confined smectics with the help of extensive Monte Carlo simulations [22]. Our goal in this work is to analyse a model for smectic-A liquid crystals, and its finite element discretization, that was recently proposed by Xia, MacLachlan, Atherton and Farrell in [31].

We consider an open, bounded and convex spatial domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$ with Lipschitz boundary $\partial\Omega$. The smectic-A free energy we analyse in this work is given by

$$\mathscr{J}(u,\mathbf{Q}) = \int_\Omega \left(f_s(u) + B\left|\mathscr{D}^2 u + q^2\left(\mathbf{Q}+\frac{\mathbf{I}_d}{d}\right)u\right|^2 + f_n(\mathbf{Q},\nabla\mathbf{Q})\right), \tag{1}$$

where $\mathbf{I}_d$ denotes the $d\times d$ identity matrix, $\mathscr{D}^2$ is the Hessian operator, $f_s(u)$ is the smectic bulk energy density

$$f_s(u) := \frac{a_1}{2}u^2 + \frac{a_2}{3}u^3 + \frac{a_3}{4}u^4$$

and $f_n(\mathbf{Q},\nabla\mathbf{Q})$ is the nematic bulk energy density

$$\begin{aligned} f_n(\mathbf{Q},\nabla\mathbf{Q}) &= f_n^e(\nabla\mathbf{Q}) + f_n^b(\mathbf{Q}) \\ &:= \frac{K}{2}|\nabla\mathbf{Q}|^2 + \begin{cases} \left(-l\left(\operatorname{tr}(\mathbf{Q}^2)\right) + l\left(\operatorname{tr}(\mathbf{Q}^2)\right)^2\right), & \text{if } d=2, \\ \left(-\frac{l}{2}\left(\operatorname{tr}(\mathbf{Q}^2)\right) - \frac{l}{3}\left(\operatorname{tr}(\mathbf{Q}^3)\right) + \frac{l}{2}\left(\operatorname{tr}(\mathbf{Q}^2)\right)^2\right), & \text{if } d=3. \end{cases} \end{aligned} \tag{2}$$

Here, $B>0$ is the nematic-smectic coupling parameter, a_1,a_2,a_3 represent smectic bulk constants with $a_3>0$, $K>0$ and $l>0$ are nematic elastic and bulk constants, respectively, $q\geq 0$ is the wave number of the smectic layers, and the trace of a matrix is given by $\operatorname{tr}(\cdot)$.

The two main contributions of this paper are to prove existence of minimisers to the problem of minimising $\mathscr{J}$ over a suitably-defined admissible set, and to derive a priori error estimates for its discretisation using the $\mathscr{C}^0$ interior penalty method [7, 9]. We show the existence result of minimising the free energy eq. (1) in section 1. We then derive a priori error estimates for both $\mathbf{Q}$ and u in the simplified case $q=0$ in section 2. We confirm that the theoretical predictions match numerical experiments in section 3, for both $q=0$ and $q>0$.

In order to avoid the proliferation of constants throughout this work, we use the notation $a \lesssim b$ (respectively, $b \gtrsim a$) to represent the relation $a \leq Cb$ (respectively, $b \geq Ca$) for some generic constant C (possibly not the same constant on each appearance) independent of the mesh.

1. Existence of minimisers

To formulate the minimisation problem for eq. (1), we must first define the admissible space $\mathscr{A}$ in which minimisers will be sought. We define $\mathscr{A}$ as

$$\mathscr{A} = \left\{(u,\mathbf{Q}) \in H^2(\Omega,\mathbb{R}) \times H^1(\Omega,S_0) : \mathbf{Q} = \mathbf{Q}_b, u = u_b \text{ on } \partial\Omega\right\}, \tag{3}$$

with the specified Dirichlet boundary data $\mathbf{Q}_b \in H^{1/2}(\partial\Omega,S_0)$ and $u_b \in H^{3/2}(\partial\Omega,\mathbb{R})$. Here, S_0 denotes the space of symmetric and traceless $d\times d$ real-valued matrices. For simplicity of the analysis, we only consider Dirichlet boundary conditions for $\mathbf{Q}$ and u here, but other types of boundary conditions (e.g., a mixture of the Dirichlet and natural boundary conditions for u as illustrated in the implementations in [31]) can be taken. With this admissible space, we consider the minimisation problem

$$\min_{(u,\mathbf{Q})\in\mathscr{A}} \mathscr{J}(u,\mathbf{Q}). \tag{4}$$

Notice that $f_n(\mathbf{Q},\nabla\mathbf{Q})$ is the classical Landau de Gennes (LdG) free energy density [16, 23] for nematic LC and it is proven by Davis & Gartland [14, Corollary 4.4] that there exists a minimiser of the functional $\int_\Omega f_n(\mathbf{Q},\nabla\mathbf{Q})$ for $\mathbf{Q}\in H^1(\Omega,S_0)$ in three dimensions. Moreover, Bedford [4, Theorem 5.18] gives an existence result for the BB model in three dimensions:

$$\min_{(u,\mathbf{Q})\in\mathscr{A}^{BB}} \mathscr{J}^{BB}(u,\mathbf{Q}) = \int_\Omega \left\{ \frac{K}{2}|\nabla\mathbf{Q}|^2 + B\left|\mathscr{D}^2 u + q^2\left(\frac{\mathbf{Q}}{s}+\frac{\mathbf{I}_3}{3}\right)u\right|^2 + \frac{a_1}{2}u^2 + \frac{a_2}{3}u^3 + \frac{a_3}{4}u^4 \right\},$$

with an admissible space $\mathscr{A}^{BB} := \{\mathbf{Q}\in SBV(\Omega,S_0), u\in H^2(\Omega,\mathbb{R}) : \mathbf{Q} = s\left(\mathbf{n}\otimes\mathbf{n} - \frac{\mathbf{I}_d}{d}\right), s\in[0,1], |\mathbf{n}|=1\}$, where SBV denotes special functions of bounded variation. For simplicity, we have ignored the surface integral here in the energy functional of the BB model. One can observe its resemblance to eq. (1). Motivated by the above results, we prove the existence of minimisers to eq. (1) via the direct method of calculus of variations.

Theorem 1.1. *(Existence of minimisers) Let $\mathscr{J}$ be of the form eq.* (1) *with positive parameters a_3, B, K, l and non-negative wave number q. Then there exists a solution pair $(u^\star,\mathbf{Q}^\star)$ that solves eq.* (4)*.*

Proof. Note that both the smectic density f_s and the nematic bulk density f_n^b are bounded from below as $a_3, l > 0$. Thus, $\mathscr{J}$ is also bounded from below and we can choose a minimising sequence $\{(u_j,\mathbf{Q}_j)\}$, i.e.,

$$\begin{gathered}(u_j,\mathbf{Q}_j)\in\mathscr{A},\ u_j-\tilde{u}\in H^2\cap H_0^1(\Omega,\mathbb{R}),\ \mathbf{Q}_j-\tilde{\mathbf{Q}}\in H_0^1(\Omega,S_0),\\ \mathscr{J}(u_j,\mathbf{Q}_j)\overset{j\to\infty}{\longrightarrow}\inf\{\mathscr{J}(u,\mathbf{Q}) : (u,\mathbf{Q})\in\mathscr{A},\ u-\tilde{u}\in H^2\cap H_0^1(\Omega,\mathbb{R}),\ \mathbf{Q}-\tilde{\mathbf{Q}}\in H_0^1(\Omega,S_0)\}<\infty.\end{gathered} \tag{5}$$

Here we set $\tilde{\mathbf{Q}}\in H^1(\Omega,S_0)$ (resp. $\tilde{u}\in H^2(\Omega,\mathbb{R})$) to be any function with trace $\mathbf{Q}_b$ (resp. u_b). We tackle the three terms in $\mathscr{J}(u,\mathbf{Q})$ separately in the following.

First, for the nematic energy term $\int_\Omega f_n(\mathbf{Q},\nabla\mathbf{Q})$, we can follow the proof of [14, Theorem 4.3] to obtain that $f_n(\mathbf{Q}_j,\nabla\mathbf{Q}_j)$ is coercive in $H^1(\Omega,S_0)$, i.e., f_n grows unbounded as $\|\mathbf{Q}_j\|_1\to\infty$, and thus the minimising sequence $\{\mathbf{Q}_j\}$ must be bounded in $\mathbf{H}^1(\Omega,S_0)$. Since $H^1(\Omega)$ is a reflexive Banach Space, we have a subsequence (also denoted as $\{\mathbf{Q}_j\}$) that weakly converges to $\mathbf{Q}^\star\in H^1(\Omega,S_0)$ such that $\mathbf{Q}^\star-\tilde{\mathbf{Q}}\in H_0^1(\Omega,S_0)$, and from the Rellich–Kondrachov theorem it follows that $\mathbf{Q}_j\to\mathbf{Q}^\star$ in $L^2(\Omega)$ and $\nabla\mathbf{Q}_j\rightharpoonup\nabla\mathbf{Q}^\star$ in $L^2(\Omega)$. The weak lower semi-continuity of the nematic energy density f_n in eq. (2) is guaranteed by [14, Lemma 4.2], therefore,

$$\liminf_{j\to\infty}\int_\Omega f_n(\mathbf{Q}_j,\nabla\mathbf{Q}_j) \geq \int_\Omega f_n(\mathbf{Q}^\star,\nabla\mathbf{Q}^\star). \tag{6}$$

For the smectic bulk term $\int_\Omega f_s(u)$, we can follow the proof in [4, Theorem 5.19]. By eq. (5), we have

$$\sup_j \int_\Omega \left(\left|\mathscr{D}^2 u_j\right|^2 + |u_j|^2\right) < \infty,$$

which implies an upper bound for ∇u_j using [4, Ineq. (5.42)]. Hence, $\{u_j\}$ is bounded in $H^2(\Omega)$ and thus there is a subsequence (also denoted as $\{u_j\}$) such that $u_j\rightharpoonup u^\star$ in $H^2(\Omega)$ and $u^\star-\tilde{u}\in H^2\cap H_0^1(\Omega)$. Moreover, one can readily check that $\|u^\star\|_\infty<\infty$ by the Sobolev embedding of $H^2(\Omega)$ into the Hölder spaces $\mathscr{C}^{\mathfrak{t},\varkappa_0}(\Omega)$ ($\mathfrak{t}+\varkappa_0=1$ for $d=2$ and $\mathfrak{t}+\varkappa_0=1/2$ for $d=3$) and the boundedness of domain Ω. Again, it follows from the Rellich–Kondrachov theorem that $u_j\to u^\star$ in $L^2(\Omega)$ and $\mathscr{D}^2 u_j\rightharpoonup\mathscr{D}^2 u^\star$ in $L^2(\Omega)$. Noting $f_s(u)$ is bounded from below for all $u\in H^2(\Omega)$, we then obtain

$$\liminf_{j\to\infty}\int_\Omega f_s(u_j) \geq \int_\Omega f_s(u^\star). \tag{7}$$

Now, we consider the nematic-smectic coupling term $\int_\Omega B\left|\mathscr{D}^2 u+q^2\left(\mathbf{Q}+\frac{\mathbf{I}_d}{d}\right)u\right|^2$ in $\mathscr{J}(u,\mathbf{Q})$. Note that when the wave number $q=0$, this term reduces to $\int_\Omega B|\mathscr{D}^2 u|^2$ and it is straightforward to obtain the weak lower semi-continuity property. Therefore, we discuss the case of $q>0$ in detail as follows. By the $\mathbf{H}^1$-boundedness property of the minimising sequence $\{\mathbf{Q}_j\}$ and the fact that $\|u^\star\|_\infty<\infty$, we can deduce

$$\begin{aligned}\int_\Omega |u_j\mathbf{Q}_j-u^\star\mathbf{Q}^\star|^2 &= \int_\Omega \left|(u_j-u^\star)\mathbf{Q}_j+u^\star(\mathbf{Q}_j-\mathbf{Q}^\star)\right|^2\\ &\le 2\int_\Omega \left(|u_j-u^\star|^2|\mathbf{Q}_j|^2+|u^\star|^2|\mathbf{Q}_j-\mathbf{Q}^\star|^2\right)\\ &\to 0 \quad \text{as } u_j\to u^\star, \mathbf{Q}_j\to\mathbf{Q}^\star \text{ in } L^2.\end{aligned}$$

Hence, $u_j\mathbf{Q}_j\to u^\star\mathbf{Q}^\star$ in $L^2(\Omega)$, and further,

$$\begin{aligned} u_j\left(\mathbf{Q}_j+\frac{\mathbf{I}_d}{d}\right) &\to u^\star\left(\mathbf{Q}^\star+\frac{\mathbf{I}_d}{d}\right) && \text{in } L^2(\Omega),\\ u_j\left(\mathbf{Q}_j+\frac{\mathbf{I}_d}{d}\right):\mathscr{D}^2u_j &\rightharpoonup u^\star\left(\mathbf{Q}^\star+\frac{\mathbf{I}_d}{d}\right):\mathscr{D}^2u^\star && \text{in } L^1(\Omega).\end{aligned}$$

Therefore, we have

$$\begin{aligned}\liminf_{j\to\infty}\int_\Omega\left|\mathscr{D}^2u_j+q^2\left(\mathbf{Q}_j+\frac{\mathbf{I}_d}{d}\right)u_j\right|^2 &= \liminf_{j\to\infty}\int_\Omega\left(\left|\mathscr{D}^2u_j\right|^2+2q^2u_j\left(\mathbf{Q}_j+\frac{\mathbf{I}_d}{d}\right):\mathscr{D}^2u_j+q^4\left|u_j\left(\mathbf{Q}_j+\frac{\mathbf{I}_d}{d}\right)\right|^2\right)\\ &\ge \int_\Omega\left(\left|\mathscr{D}^2u^\star\right|^2+2q^2u^\star\left(\mathbf{Q}^\star+\frac{\mathbf{I}_d}{d}\right):\mathscr{D}^2u^\star+q^4\left|u^\star\left(\mathbf{Q}^\star+\frac{\mathbf{I}_d}{d}\right)\right|^2\right)\\ &= \int_\Omega\left|\mathscr{D}^2u^\star+q^2\left(\mathbf{Q}^\star+\frac{\mathbf{I}_d}{d}\right)u^\star\right|^2. \end{aligned}\tag{8}$$

Hence, we can conclude that $\mathscr{J}(u^\star,\mathbf{Q}^\star)$ achieves its minimum in the admissible space $\mathscr{A}$ by combining eq. (6), eq. (7) and eq. (8). □

Remark 1.2. *We briefly compare the proof with that of Bedford [4, Theorem 5.18]. In that work, the admissible space $\mathscr{A}^{BB}$ included a uniaxial constraint.This assumption is necessary to guarantee the $\mathbf{H}^1$-boundedness property of the minimising sequence $\{\mathbf{Q}_j\}$, since that work seeks $\mathbf{Q}\in SBV(\Omega,S_0)$ instead of $\mathbf{H}^1$. Enforcing this constraint numerically is difficult [6]; in this work we prefer the choice $\mathbf{Q}\in\mathbf{H}^1$, which enables us to remove the uniaxiality assumption.*

2. A PRIORI ERROR ESTIMATES

We now consider the discretisation of the minimisation problem eq. (4). For simplicity, we analyse the decoupled case with $q=0$, where eq. (4) splits into two independent problems: one for the smectic density variation u:

$$\min_{u\in H^2\cap H^1_b(\Omega,\mathbb{R})}\quad \mathscr{J}_1(u)=\int_\Omega\left(B\left|\mathscr{D}^2u\right|^2+f_s(u)\right),$$

and one for the tensor field $\mathbf{Q}$,

$$\min_{\mathbf{Q}\in H^1_b(\Omega,S_0)}\quad \mathscr{J}_2(\mathbf{Q})=\int_\Omega\left(f_n(\mathbf{Q},\nabla\mathbf{Q})\right).$$

Here, $H_b^1(\Omega,\mathbb{R}) := \{u \in H^1(\Omega,\mathbb{R}) : u = u_b \text{ on } \partial\Omega\}$ and $H_b^1(\Omega,S_0) := \{\mathbf{Q} \in H^1(\Omega,S_0) : \mathbf{Q} = \mathbf{Q}_b \text{ on } \partial\Omega\}$. One can derive the following strong forms of their equilibrium equations using integration by parts. The optimality condition for u yields a fourth-order partial differential equation (PDE)

$$\begin{cases} 2B\nabla\cdot\left(\nabla\cdot\mathscr{D}^2u\right) + a_1u + a_2u^2 + a_3u^3 = 0 & \text{in } \Omega, \\ u = u_b, \quad \mathscr{D}^2u\cdot\nu = \mathscr{D}^2u_b\cdot\nu & \text{on } \partial\Omega, \end{cases}$$

where ν denotes the outward unit normal. Note that the second boundary condition of u is a natural boundary condition on the second derivative of u. For simplicity, we only consider the case of a cubic nonlinearity (i.e., $a_2 = 0$) here as the quadratic term can be tackled similarly. Therefore, we consider the following governing equations

$$(\mathscr{P}1) \quad \begin{cases} 2B\nabla\cdot\left(\nabla\cdot\mathscr{D}^2u\right) + a_1u + a_3u^3 = 0 & \text{in } \Omega, \\ u = u_b, \quad \mathscr{D}^2u\cdot\nu = \mathscr{D}^2u_b\cdot\nu & \text{on } \partial\Omega. \end{cases} \tag{9}$$

Meanwhile, the optimality condition for $\mathbf{Q}$ yields a second-order PDE

$$(\mathscr{P}2) \quad \begin{cases} d = 2 \Rightarrow -K\Delta\mathbf{Q} + 2l\left(2|\mathbf{Q}|^2 - 1\right)\mathbf{Q} = 0 & \text{in } \Omega, \\ d = 3 \Rightarrow -K\Delta\mathbf{Q} + l\left(-\mathbf{Q} - |\mathbf{Q}|^2 + 2|\mathbf{Q}|^2\mathbf{Q}\right) = 0 & \text{in } \Omega, \\ \mathbf{Q} = \mathbf{Q}_b & \text{on } \partial\Omega, \end{cases} \tag{10}$$

We now consider these two problems $(\mathscr{P}1)$ and $(\mathscr{P}2)$ in turn.

Remark 2.1. *The uniqueness of solutions is not expected. It is well-known that eq.* (10) *can support multiple solutions [25], while numerical experiments in [31] indicate the existence of multiple solutions to the optimality conditions for eq.* (4)*.*

2.1. A priori error estimates for $(\mathscr{P}1)$

Since the PDE eq. (9) for the density variation u is a fourth-order problem, a conforming discretisation requires a finite dimensional subspace of the Sobolev space $H^2(\Omega)$, which necessitates the use of $\mathscr{C}^1$-continuous elements. The construction of these elements is quite involved, particularly in three dimensions; without a special mesh structure, the lowest-degree conforming elements are the Argyris [1] and Zhang [32] elements, of degree 5 and 9 in two and three dimensions respectively. One approach to avoid such complexity is to use mixed formulations by solving two second-order systems, and we refer to [12, 27] for instance. However, this substantially increases the size of the linear systems to be solved. Alternatively, one can directly tackle the fourth-order problem with non-conforming elements, that do not satisfy the $\mathscr{C}^1$-requirement. For instance, the so-called *continuous/discontinuous Galerkin* methods and $\mathscr{C}^0$ *interior penalty* methods ($\mathscr{C}^0$-IP) are analysed in [9, 17], combining concepts from the theory of continuous and discontinuous Galerkin methods. Essentially, these methods use $\mathscr{C}^0$-conforming elements and penalise inter-element jumps in first derivatives to weakly enforce $\mathscr{C}^1$-continuity. This has the advantages of both convenience and efficiency: the weak form is simple, with only minor modifications from a conforming method, and fewer degrees of freedom are used than with a fully discontinuous Galerkin method.

We thus adopt the idea of $\mathscr{C}^0$-IP methods to solve the nonlinear fourth-order problem $(\mathscr{P}1)$ and derive a priori error estimates regarding u. We adapt the techniques of [21] to prove convergence rates with the use of familiar continuous Lagrange elements for the problem $(\mathscr{P}1)$. The weak form of eq. (9) is defined as: find $u \in H^2(\Omega)\cap H_b^1(\Omega;\mathbb{R})$ such that

$$\mathscr{N}^s(u)t := A^s(u,t) + B^s(u,u,u,t) + C^s(u,t) = L^s(t) \quad \forall t \in H^2(\Omega)\cap H_0^1(\Omega), \tag{11}$$

where for $t, w \in H^2(\Omega)$,

$$A^s(t,w) = 2B\int_\Omega \mathscr{D}^2t : \mathscr{D}^2w,\ C^s(t,w) = a_1\int_\Omega tw,\ L^s(t) := 2B\int_{\partial\Omega}\left(\mathscr{D}^2u_b\cdot\nabla t\right)\cdot\nu,$$

and for $\mu,\zeta,\eta,\xi \in H^2(\Omega)$,

$$B^s(\mu,\zeta,\eta,\xi) = a_3 \int_\Omega \mu\zeta\eta\xi.$$

Since eq. (11) is nonlinear, we derive its linearisation: find $v \in H^2(\Omega)\cap H^1_0(\Omega)$ such that

$$\langle \mathscr{D}\mathscr{N}^s(u)v,w\rangle_{H^2} := A^s(v,w)+3B^s(u,u,v,w)+C^s(v,w) = L^s(w) \quad \forall w \in H^2(\Omega)\cap H^1_0(\Omega), \tag{12}$$

where $\langle\cdot,\cdot\rangle_{H^2}$ represents the dual pairing between $\left(H^2(\Omega)\cap H^1_0(\Omega)\right)^\star$ and $H^2(\Omega)\cap H^1_0(\Omega)$.

It is straightforward to derive the coercivity and boundedness of the bilinear operator $A^s(\cdot,\cdot)$ with the semi-norm $|\cdot|_2$ (in fact, this is indeed a norm in $H^2(\Omega)\cap H^1_0(\Omega)$).

Lemma 2.2. *For $v,w \in H^2(\Omega)\cap H^1_0(\Omega)$, there holds $A^s(v,w) \lesssim |v|_2|w|_2$ and $A^s(v,v) \gtrsim |v|_2^2$.*

Let $\mathscr{T}_h$ be a mesh of Ω with T denoting an element, and let $\mathscr{E}_I$ (resp. $\mathscr{E}_B$) denote the set of all interior (resp. boundary) edges/faces e of the mesh, and $\mathscr{E} := \mathscr{E}_I \cup \mathscr{E}_B$. Define the broken Sobolev space $H^2(\mathscr{T}_h) := \{v \in H^1(\Omega) : v|_T \in H^2(T)\ \forall T \in \mathscr{T}_h\}$ equipped with the broken norm $\|v\|^2_{2,\mathscr{T}_h} = \sum_{T\in\mathscr{T}_h}\|v\|^2_{2,T}$. We take a nonconforming but continuous approximation u_h for the solution u of eq. (11), that is to say, $u_h \in W_{h,b} \subset H^2(\mathscr{T}_h)\cap H^1_b(\Omega)$ defined for deg ≥ 2 (since ($\mathscr{P}1$) is a fourth-order problem) by

$$\begin{aligned}
W_h &:= \{v \in H^2(\mathscr{T}_h)\cap H^1(\Omega) : v \in \mathbb{Q}_{\deg}(T)\ \forall T \in \mathscr{T}_h\},\\
W_{h,0} &:= \{v \in H^2(\mathscr{T}_h)\cap H^1(\Omega) : v = 0 \text{ on } \partial\Omega, v \in \mathbb{Q}_{\deg}(T)\ \forall T \in \mathscr{T}_h\},\\
W_{h,b} &:= \{v \in H^2(\mathscr{T}_h)\cap H^1(\Omega) : v = u_b \text{ on } \partial\Omega, v \in \mathbb{Q}_{\deg}(T)\ \forall T \in \mathscr{T}_h\},
\end{aligned}$$

with $\mathbb{Q}_{\deg}$ denoting piecewise continuous polynomials of degree deg on a mesh of quadrilaterals or hexahedra. Following the derivation of the $\mathscr{C}^0$-IP formulation given in [7, Section 3], we introduce the discrete nonlinear weak form: find $u_h \in W_{h,b}$ such that

$$\mathscr{N}^s_h(u_h)t_h := A^s_h(u_h,t_h)+P^s_h(u_h,t_h)+B^s(u_h,u_h,u_h,t_h)+C^s(u_h,t_h) = L^s(t_h) \quad \forall t_h \in W_{h,0}, \tag{13}$$

where for all $u_h,t_h \in W_h$,

$$A^s_h(u_h,t_h) := 2B\left(\sum_{T\in\mathscr{T}_h}\int_T \mathscr{D}^2u_h : \mathscr{D}^2t_h - \sum_{e\in\mathscr{E}_I}\int_e \left\{\!\!\left\{\frac{\partial^2 u_h}{\partial\nu^2}\right\}\!\!\right\}[\![\nabla t_h]\!] - \sum_{e\in\mathscr{E}_I}\int_e \left\{\!\!\left\{\frac{\partial^2 t_h}{\partial\nu^2}\right\}\!\!\right\}[\![\nabla u_h]\!]\right),$$

and

$$P^s_h(u_h,t_h) := \sum_{e\in\mathscr{E}_I}\frac{2B\varepsilon}{h_e^3}\int_e [\![\nabla u_h]\!][\![\nabla t_h]\!]. \tag{14}$$

Here, ε is the penalty parameter (to be specified in section 3), h_e indicates the size of the edge/face e and the average of the second derivatives of u along the normal direction across the edge/facet e is defined as $\left\{\!\!\left\{\frac{\partial^2 u}{\partial\nu^2}\right\}\!\!\right\} = \frac{1}{2}\left(\frac{\partial^2 u_+}{\partial\nu^2}\Big|_e + \frac{\partial^2 u_-}{\partial\nu^2}\Big|_e\right)$. For any interior edge $e \in \mathscr{E}_I$ shared by two cells T_- and T_+, we define the jump $[\![\mathbf{v}]\!]$ by $[\![\mathbf{v}]\!] = \mathbf{v}_-\cdot\nu_- + \mathbf{v}_+\cdot\nu_+$ with ν_-,ν_+ representing the restriction of outward normals in T_-,T_+ respectively. On the boundary edge/face $e \in \mathscr{E}_B$, we define $[\![\mathbf{v}]\!] = \mathbf{v}\cdot\nu$. The operator P^s_h penalises the first derivatives across the interior edge/facet since the function in $H^1(\Omega)$ is not necessarily continuously differentiable.

The linearised version is to seek $v_h \in W_{h,0}$ such that

$$\langle \mathscr{D}\mathscr{N}^s_h(u_h)v_h,w_h\rangle = L^s(w_h) \quad \forall w_h \in W_{h,0}, \tag{15}$$

where

$$\langle \mathscr{D}\mathscr{N}^s_h(u_h)v_h,w_h\rangle := A^s_h(v_h,w_h)+P^s_h(v_h,w_h)+3B^s(u_h,u_h,v_h,w_h)+C^s(v_h,w_h). \tag{16}$$

We also define the mesh-dependent H^2-like semi-norm for $v \in W_h$,

$$|||v|||_h^2 := \sum_{T\in\mathscr{T}_h} |v|_{H^2(T)}^2 + \sum_{e\in\mathscr{E}_I} \int_e \frac{1}{h_e^3} |[\![\nabla v]\!]|^2. \tag{17}$$

Note that $|||\cdot|||_h$ is indeed a norm on $W_{h,0}$. This norm will be used in the well-posedness and convergence analysis below.

We first give an immediate result about the consistency of the discrete form eq. (13).

Theorem 2.3. *(Consistency) Assuming that $u \in H^4(\Omega)$. The solution u of the continuous weak form eq.* (11) *solves the discrete weak problem eq.* (13)*.*

Proof. Multiplying the fourth-order term $2B\nabla\cdot(\nabla\cdot(\mathscr{D}^2u))$ in eq. (9) with a test function $t \in W_{h,0}$ and using piecewise integration by parts with the boundary condition specified in eq. (9) for u, one can obtain

$$2B \sum_{T\in\mathscr{E}_h} \int_T \nabla\cdot(\nabla\cdot(\mathscr{D}^2u))t = 2B \sum_{T\in\mathscr{E}_h} \int_T \mathscr{D}^2u \colon \mathscr{D}^2t - 2B \sum_{e\in\mathscr{E}_I} \int_e \left\{\!\!\left\{ \frac{\partial^2 u}{\partial \nu^2} \right\}\!\!\right\} [\![\nabla t]\!]. \tag{18}$$

Since $u \in H^4(\Omega)$ implies ∇u is continuous on the whole domain Ω, the jump term $[\![\nabla u]\!]$ then becomes zero and we can thus symmetrise and penalise the form eq. (18). This leads to the presence of $A_h^s(u,t) + P_h^s(u,t)$. The remaining terms involving B^s and C^s are straightforward as one takes the test function $t \in W_{h,0}$. Therefore, u satisfies eq. (13). □

By noting that $\left(\mathscr{D}^2 \colon \mathscr{D}^2\right)u = \left[\left(\partial_x^2\right)^2 + \left(\partial_y^2\right)^2 + 2\left(\partial_{xy}^2\right)^2\right]u = \Delta^2 u$, it is natural to extend the classical elliptic regularity result [5] for the biharmonic operator Δ^2 to the case of the bi-Hessian operator $\mathscr{D}^2 \colon \mathscr{D}^2$. If the domain Ω is smooth, the weak solutions of the biharmonic problem (e.g., [7, Example 2]) belong to $H^4(\Omega)$ by classical elliptic regularity results and thus we make this assumption henceforth.

Moreover, to facilitate the analysis, we further assume that u is an isolated solution, i.e., there is only one solution u satisfying eq. (9) within a sufficiently small ball $\{v \in H^2(\Omega)\cap H_0^1(\Omega) : |v-u|_2 \le r_b\}$ with radius r_b. These assumptions then imply that the linearised operator $\langle \mathscr{D}\mathscr{N}^s(u)\cdot,\cdot\rangle_{H^2}$ satisfies the following inf-sup condition:

$$0 < \beta_u = \inf_{\substack{v\in H^2(\Omega)\cap H_0^1(\Omega)\\ |v|_2=1}} \sup_{\substack{w\in H^2(\Omega)\cap H_0^1(\Omega)\\ |w|_2=1}} \langle \mathscr{D}\mathscr{N}^s(u)v,w\rangle_{H^2} = \inf_{\substack{w\in H^2(\Omega)\cap H_0^1(\Omega)\\ |w|_2=1}} \sup_{\substack{v\in H^2(\Omega)\cap H_0^1(\Omega)\\ |v|_2=1}} \langle \mathscr{D}\mathscr{N}^s(u)v,w\rangle_{H^2}. \tag{19}$$

2.1.1. *Well-posedness of the discrete form*

Recalling [7, Eq. (3.20)], we can obtain for $v,w \in W_{h,0}$,

$$\sum_{e\in\mathscr{E}_I} \left| \int_e \left\{\!\!\left\{ \frac{\partial^2 w}{\partial \nu^2} \right\}\!\!\right\} [\![\nabla v]\!] \right| \lesssim \left(\sum_{T\in\mathscr{T}_h} \int_T \mathscr{D}^2 w \colon \mathscr{D}^2 w \right)^{1/2} \left(\sum_{e\in\mathscr{E}_I} \frac{1}{h_e^3} \int_e ([\![\nabla v]\!])^2 \right)^{1/2}, \tag{20}$$

as the edge/facet size $h_e < 1$. With the estimate eq. (20) at hand, we then apply the Cauchy–Schwarz inequality and use the definition eq. (17) of $|||\cdot|||_h$ to obtain the boundedness of $A_h^s(\cdot,\cdot)$ and $P_h^s(\cdot,\cdot)$. We omit the details of the proofs here and only illustrate the boundedness result for $B^s(\cdot,\cdot,\cdot,\cdot)$ and $C^s(\cdot,\cdot)$ below.

Lemma 2.4. *(Boundedness of $B^s(\cdot,\cdot,\cdot,\cdot)$ and $C^s(\cdot,\cdot)$.) For $u,v,w,p \in W_{h,0}$, we have*

$$|B^s(u,v,w,p)| \lesssim |||u|||_h |||v|||_h |||w|||_h |||p|||_h \text{ and } |C^s(u,v)| \lesssim |||u|||_h |||v|||_h. \tag{21}$$

For $u,v \in H^2(\Omega)$, $w,p \in W_h$,

$$|B^s(u,v,w,p)| \lesssim \|u\|_2 \|v\|_2 |||w|||_h |||p|||_h. \tag{22}$$

Proof. By Hölder's inequality, the Sobolev embedding $H^1(\Omega) \hookrightarrow L^4(\Omega)$, and the fact that the H^1 semi-norm $|\cdot|_1$ is a norm in $H_0^1(\Omega)$, we deduce

$$\begin{aligned}|B^s(u,v,w,p)| &\lesssim \|u\|_{L^4}\|v\|_{L^4}\|w\|_{L^4}\|p\|_{L^4}\\ &\lesssim |u|_1|v|_1|w|_1|p|_1.\end{aligned}$$

It then follows from a Poincaré inequality [11, Eq. (5.7)] for piecewise H^2 functions that

$$\sum_{T\in\mathscr{T}_h}|v|_{1,T}^2 \lesssim \sum_{T\in\mathscr{T}_h}|v|_{2,T}^2 + \sum_{e\in\mathscr{E}_I}\frac{1}{h_e^3}\|[\![\nabla v]\!]\|_{0,e}^2 = |\!|\!|v|\!|\!|_h^2 \quad \forall v\in W_{h,0}. \tag{23}$$

Thus, we obtain $|B^s(u,v,w,p)| \lesssim |\!|\!|u|\!|\!|_h|\!|\!|v|\!|\!|_h|\!|\!|w|\!|\!|_h|\!|\!|p|\!|\!|_h$.

The boundedness of $C^s(\cdot,\cdot)$ follows similarly by the Cauchy–Schwarz inequality, the Sobolev embedding $H^1(\Omega)\hookrightarrow L^2(\Omega)$ and the use of eq. (23). The proof of eq. (22) is analogous to that of eq. (21) with a use of the embedding result $H^2(\Omega)\hookrightarrow L^\infty(\Omega)$ and the Cauchy–Schwarz inequality. □

We give the coercivity result for the bilinear form $\left(A_h^s(\cdot,\cdot)+P_h^s(\cdot,\cdot)\right)$.

Lemma 2.5. *(Coercivity of $A_h^s+P_h^s$) For a sufficiently large penalty parameter ε, there holds*

$$|\!|\!|v_h|\!|\!|_h^2 \lesssim A_h^s(v_h,v_h)+P_h^s(v_h,v_h) \quad \forall v_h\in W_{h,0}. \tag{24}$$

Proof. By eq. (20) and the inequality of geometric and arithmetic means, we deduce for $v\in W_h$,

$$\begin{aligned}A_h^s(v,v)+P_h^s(v,v) &\geq 2B\sum_{T\in\mathscr{T}_h}|v|_{H^2(T)}^2 - 2BC\left(\sum_{T\in\mathscr{T}_h}|v|_{2,T}^2\right)^{1/2}\left(\sum_{e\in\mathscr{E}_I}\frac{1}{h_e^3}\|[\![\nabla v]\!]\|_{0,e}^2\right)^{1/2} + 2B\left(\sum_{e\in\mathscr{E}_I}\int_e\frac{\varepsilon}{h_e^3}|[\![\nabla v]\!]|^2\right)\\ &\geq 2B\left[\frac{1}{2}\sum_{T\in\mathscr{T}_h}|v|_{H^2(T)}^2 + \left(\varepsilon-\frac{C^2}{2}\right)\sum_{e\in\mathscr{E}_I}\frac{1}{h_e^3}\|[\![\nabla v]\!]\|_{0,e}^2\right] \geq B|\!|\!|v|\!|\!|_h^2,\end{aligned}$$

provided the penalty parameter ε is sufficiently large with the generic constant C from eq. (20). □

An important question about the well-posedness is the coercivity of the bilinear operator $\langle\mathscr{D}\mathscr{N}_h^s(u_h)\cdot,\cdot\rangle$. Due to the presence of B^s and C^s terms in $\langle\mathscr{D}\mathscr{N}_h^s(u_h)\cdot,\cdot\rangle$, it is not trivial to derive its coercivity. We first discuss the weak coercivity of the bilinear form $\langle\mathscr{D}\mathscr{N}_h^s(u)\cdot,\cdot\rangle$ defined as

$$\langle\mathscr{D}\mathscr{N}_h^s(u)v_h,w_h\rangle := A_h^s(v_h,w_h)+P_h^s(v_h,w_h)+3B^s(u,u,v_h,w_h)+C^s(v_h,w_h) \quad \forall v_h,w_h\in W_{h,0}. \tag{25}$$

To this end, we will employ the enrichment operator $E_h: W_h\to W_C\subset H^2(\Omega)$ with W_C being the Hsieh–Clough–Tocher macro finite element space [7]. The following lemma is adapted to our notation and definition of $|\!|\!|\cdot|\!|\!|_h$ from [7, Lemma 1].

Lemma 2.6. *[7, Lemma 1] For $v_h\in W_{h,0}$, there holds*

$$\sum_{T\in\mathscr{T}_h}\left(h^{-4}\|v_h-E_hv_h\|_{L^2(T)}^2 + h^{-2}|v_h-E_hv_h|_{H^1(T)}^2 + |v_h-E_hv_h|_{H^2(T)}^2\right) \lesssim \sum_{e\in\mathscr{E}_I}\frac{1}{h_e^3}\|[\![\nabla v_h]\!]\|_{L^2(e)}^2 \lesssim |\!|\!|v_h|\!|\!|_h^2.$$

With this, we obtain the following discrete inf-sup condition for the discrete bilinear operator $\langle\mathscr{D}\mathscr{N}_h^s(u)\cdot,\cdot\rangle$.

Theorem 2.7. *(Weak coercivity of $\langle\mathscr{D}\mathscr{N}_h^s(u)\cdot,\cdot\rangle$) Let u be a regular isolated solution of the nonlinear continuous weak form eq.* (13). *For a sufficiently large ε and a sufficiently small mesh size h, the following discrete inf-sup condition holds with a positive constant $\beta_c>0$:*

$$0<\beta_c\leq \inf_{\substack{v\in W_{h,0}\\ |\!|\!|v_h|\!|\!|_h=1}}\ \sup_{\substack{w\in W_{h,0}\\ |\!|\!|w_h|\!|\!|_h=1}}\langle\mathscr{D}\mathscr{N}_h^s(u)v_h,w_h\rangle. \tag{26}$$

Proof. For $v \in H^2(\Omega)\cap H_0^1(\Omega)$, it follows from the boundedness result of B^s, C^s that $B^s(u,u,v,\cdot)$ and $C^s(v,\cdot) \in (L^2(\Omega))^\star$. Furthermore, since $A^s(\cdot,\cdot)$ is bounded and coercive as given by lemma 2.2, for a given $v_h \in W_h$ with $|\!|\!|v_h|\!|\!|_h = 1$, there exists ξ and $\eta \in H^4(\Omega)\cap H_0^1(\Omega)$ that solve the linear systems:

$$A^s(\xi,w) = 3B^s(u,u,v_h,w) + C^s(v_h,w) \quad \forall w \in H^2(\Omega)\cap H_0^1(\Omega), \tag{27a}$$

$$A^s(\eta,w) = 3B^s(u,u,E_hv_h,w) + C^s(E_hv_h,w) \quad \forall w \in H^2(\Omega)\cap H_0^1(\Omega). \tag{27b}$$

We require the H^4-regularity for the derivation of eq. (34) below. It then follows from the standard elliptic regularity result that $\|\eta\|_4 \lesssim C_{BC}$ with a constant C_{BC} depending on $\|u\|_2$.

Subtracting eq. (27a) from eq. (27b), then taking $w = \eta - \xi$ and using the coercivity of $A^s(\cdot,\cdot)$, boundedness of $B^s(\cdot,\cdot,\cdot,\cdot)$, $C^s(\cdot,\cdot)$ and enrichment estimate lemma 2.6, we get

$$|\eta - \xi|_2 \lesssim \left(3\|u\|_2^2 + 1\right)\|E_hv_h - v_h\|_0 \lesssim h^2|\!|\!|v_h|\!|\!|_h = h^2. \tag{28}$$

Since u is a regular isolated solution of eq. (11), it yields by eq. (19), eq. (27b), lemma 2.2, the fact that $[\![\nabla(E_hv_h+\eta)]\!] = 0$ and the triangle inequality, that there exists $w \in H^2(\Omega)\cap H_0^1(\Omega)$ with $|w|_2 = 1$ such that

$$\begin{aligned}|E_hv_h|_2 &\lesssim \langle \mathscr{D}\mathscr{N}^s(u)E_hv_h, w\rangle_{H^2} = A^s(E_hv_h,w) + 3B^s(u,u,E_hv_h,w) + C^s(E_hv_h,w)\\ &= A^s(E_hv_h+\eta,w) \lesssim |E_hv_h+\eta|_2|w|_2 = |\!|\!|E_hv_h+\eta|\!|\!|_h\\ &\le |\!|\!|E_hv_h - v_h|\!|\!|_h + |\!|\!|v_h + I_h\xi|\!|\!|_h + |\!|\!|I_h\xi - \xi|\!|\!|_h + \underbrace{|\!|\!|\xi-\eta|\!|\!|_h}_{=|\xi-\eta|_2}.\end{aligned} \tag{29}$$

Note that $[\![\nabla\xi]\!] = 0$ on $\mathscr{E}_I$ since $\xi \in H^4(\Omega)$. We can thus calculate using lemma 2.6 that

$$|\!|\!|E_hv_h - v_h|\!|\!|_h^2 \lesssim \sum_{e\in\mathscr{E}_I}\int_e \frac{1}{h_e^3}|[\![\nabla v_h]\!]|^2 \lesssim \sum_{e\in\mathscr{E}_I}\int_e \frac{1}{h_e^3}|[\![\nabla(v_h+\xi)]\!]|^2 \le |\!|\!|v_h+\xi|\!|\!|_h^2.$$

Further, by the triangle inequality, we get

$$|\!|\!|E_hv_h - v_h|\!|\!|_h \lesssim |\!|\!|v_h+\xi|\!|\!|_h \le |\!|\!|v_h + I_h\xi|\!|\!|_h + |\!|\!|\xi - I_h\xi|\!|\!|_h. \tag{30}$$

Since $v_h + I_h\xi \in W_h$, it follows from the coercivity result eq. (24) that there exists $w_h \in W_h$ with $|\!|\!|w|\!|\!|_h = 1$ such that

$$\begin{aligned}|\!|\!|v_h + I_h\xi|\!|\!|_h &\lesssim A_h^s(v_h+I_h\xi,w_h) + P_h^s(v_h+I_h\xi,w_h)\\ &= \langle \mathscr{D}\mathscr{N}_h^s(u)v_h,w_h\rangle - 3B^s(u,u,v_h,w_h) - C^s(v_h,w_h)\\ &\quad + A_h^s(I_h\xi-\xi,w_h) + P_h^s(I_h\xi-\xi,w_h) + A_h^s(\xi,w_h) + P_h^s(\xi,w_h)\\ &= \langle \mathscr{D}\mathscr{N}_h^s(u)v_h,w_h\rangle + 3B^s(u,u,v_h,E_hw_h-w_h) + C^s(v_h,E_hw_h-w_h)\\ &\quad + A_h^s(I_h\xi-\xi,w_h) + P_h^s(I_h\xi-\xi,w_h) + A_h^s(\xi,w_h-E_hw_h) + P_h^s(\xi,w_h-E_hw_h),\end{aligned} \tag{31}$$

where in the last equality we have used the fact that

$$3B^s(u,u,v_h,E_hw_h) + C^s(v_h,E_hw_h) = A^s(\xi,E_hw_h) = A_h^s(\xi,E_hw_h) + P^s(\xi,E_hw_h)$$

because of eq. (27a) and $[\![\nabla\xi]\!] = [\![\nabla E_hw_h]\!] = 0$.

Using the boundedness result lemma 2.4 and the enrichment estimate lemma 2.6, we obtain

$$3B^s(u,u,v_h,E_hw_h-w_h) + C^s(v_h,E_hw_h-w_h) \lesssim \underbrace{\|v_h\|_0}_{\lesssim|v_h|_1\lesssim|\!|\!|v_h|\!|\!|_h=1}\ \underbrace{\|E_hw_h-w_h\|_0}_{\lesssim h^2|\!|\!|w_h|\!|\!|_h=h^2}. \tag{32}$$

By the boundedness of the bilinear form $A_h^s + P_h^s$ and standard interpolation estimate, we have

$$A_h^s(I_h\xi - \xi, w_h) + P_h^s(I_h\xi - \xi, w_h) \lesssim |||I_h\xi - \xi|||_h \underbrace{|||w_h|||_h}_{=1} \lesssim h^{\min\{\deg -1,2\}} \|\xi\|_4. \tag{33}$$

Moreover, by the fact that $[\![\nabla\xi]\!] = [\![\nabla(E_h w_h)]\!] = 0$, the enrichment estimate lemma 2.6 and eq. (18), there holds

$$\begin{aligned}
&A_h^s(\xi, w_h - E_h w_h) + P_h^s(\xi, w_h - E_h w_h) \\
&= 2B \sum_{T\in\mathcal{T}_h} \int_T \mathscr{D}^2\xi : \mathscr{D}^2(w_h - E_h w_h) - 2B \sum_{e\in\mathcal{E}_I} \int_e \left\{\!\!\left\{ \frac{\partial^2 \xi}{\partial \nu^2} \right\}\!\!\right\} [\![\nabla(w_h - E_h w_h)]\!] \\
&= 2B \sum_{T\in\mathcal{T}_h} \nabla\cdot\left(\nabla\cdot(\mathscr{D}^2\xi)\right)(w_h - E_h w_h) \lesssim \|\xi\|_4 \|w_h - E_h w_h\|_0 \lesssim h^2 \|\xi\|_4.
\end{aligned} \tag{34}$$

Combine eqs. (32) to (34) in eq. (31) to obtain

$$|||v_h + I_h\xi|||_h \lesssim \langle \mathscr{D}\mathscr{N}_h^s(u) v_h, w_h\rangle + h^2 + h^{\min\{\deg -1,2\}}. \tag{35}$$

Substituting eq. (35) into eq. (30) and using standard interpolation estimates yield that

$$|||E_h v_h - v_h|||_h \lesssim \langle \mathscr{D}\mathscr{N}_h^s(u) v_h, w_h\rangle + h^2 + h^{\min\{\deg -1,2\}}. \tag{36}$$

A use of eqs. (35) and (36), standard interpolation estimates and eq. (28) in eq. (29) leads to

$$|E_h v_h|_2 \lesssim \langle \mathscr{D}\mathscr{N}_h^s(u) v_h, w_h\rangle + h^2 + h^{\min\{\deg -1,2\}}.$$

Then, by the triangle inequality, we have

$$1 = |||v_h|||_h \le |||v_h - E_h v_h|||_h + \underbrace{|||E_h v_h|||_h}_{=|E_h v_h|_2} \le C_t \left(\langle \mathscr{D}\mathscr{N}_h^s(u) v_h, w_h\rangle + h^2 + h^{\min\{\deg -1,2\}} \right).$$

Therefore, for the mesh size h satisfying $h^2 + h^{\min\{\deg -1,2\}} < \frac{1}{2C_t}$, the discrete inf-sup condition eq. (26) holds for $\beta_c = \frac{1}{2C_t}$. □

We can now obtain the discrete inf-sup condition for the perturbed bilinear form $\langle \mathscr{D}\mathscr{N}_h^s(I_h u)\cdot, \cdot\rangle$ given by

$$\langle \mathscr{D}\mathscr{N}_h^s(I_h u) v_h, w_h\rangle = A_h^s(v_h, w_h) + P_h^s(v_h, w_h) + 3B^s(I_h u, I_h u, v_h, w_h) + C^s(v_h, w_h) \quad \forall v_h, w_h \in W_{h,0}. \tag{37}$$

Here, we employ the interpolation operator $I_h : H^2(\Omega) \cap H_b^1(\Omega;\mathbb{R}) \to W_{h,b}$.

Theorem 2.8. *(Weak coercivity of $\langle \mathscr{D}\mathscr{N}_h^s(I_h u)\cdot, \cdot\rangle$) Let u be a regular isolated solution of the nonlinear continuous weak form eq.* (13) *and $I_h u$ the interpolation of u. For a sufficiently large ε and a sufficiently small mesh size h, the following discrete inf-sup condition holds:*

$$0 < \frac{\beta_c}{2} \le \inf_{\substack{v_h \in W_{h,0} \\ |||v_h|||_h = 1}} \sup_{\substack{w_h \in W_{h,0} \\ |||w_h|||_h = 1}} \langle \mathscr{D}\mathscr{N}_h^s(I_h u) v_h, w_h\rangle. \tag{38}$$

Proof. Denote $\tilde{u} = u - I_h u$. By the definition eq. (37) of the bilinear form $\langle \mathscr{D}\mathscr{N}_h^s(I_h u)\cdot, \cdot\rangle$, we have $\langle \mathscr{D}\mathscr{N}_h^s(I_h u) v_h, w_h\rangle = A_h^s(v_h, w_h) + P_h^s(v_h, w_h) + 3B^s(u - \tilde{u}, u - \tilde{u}, v_h, w_h) + C^s(v_h, w_h)$. It follows from the definition of B^s and its boundedness result lemma 2.4 that

$$\begin{aligned}
B^s(u - \tilde{u}, u - \tilde{u}, v_h, w_h) &= B^s(u, u, v_h, w_h) + B^s(\tilde{u}, \tilde{u}, v_h, w_h) - 2B^s(u, \tilde{u}, v_h, w_h) \\
&\ge B^s(u, u, v_h, w_h) + B^s(\tilde{u}, \tilde{u}, v_h, w_h) - 2C_1 |||u|||_h |||\tilde{u}|||_h |||v_h|||_h |||w_h|||_h,
\end{aligned}$$

where C_1 is the generic constant arising in the boundedness result lemma 2.4 for $B^s(\cdot,\cdot,\cdot,\cdot)$. Therefore, we obtain that

$$\langle \mathscr{D}\mathscr{N}_h^s(I_hu)v_h, w_h\rangle \geq \langle \mathscr{D}\mathscr{N}_h^s(u)v_h, w_h\rangle + 3B^s(\tilde{u},\tilde{u},v_h,w_h) - 6C_1 |||u|||_h |||\tilde{u}|||_h |||v_h|||_h |||w_h|||_h.$$

Now using the inf-sup condition theorem 2.7 for the bilinear form $\langle \mathscr{D}\mathscr{N}_h^s(u)\cdot,\cdot\rangle$, boundedness result lemma 2.4 and interpolation estimates, we get

$$\begin{aligned}\sup_{\substack{|||w_h|||_h=1\\ w_h\in W_{h,0}}} \langle \mathscr{D}\mathscr{N}_h^s(I_hu)v_h, w_h\rangle &\geq \sup_{\substack{|||w_h|||_h=1\\ w_h\in W_{h,0}}} \langle \mathscr{D}\mathscr{N}_h^s(u)v_h, w_h\rangle \\ &\quad - 3|B^s(\tilde{u},\tilde{u},v_h,w_h)| - 6C_1 h^{\min\{\deg-1,\mathbb{k}_u-2\}} |||u|||_h |||v_h|||_h \\ &\geq \left(\beta_c - C_2 h^{\min\{\deg-1,\mathbb{k}_u-2\}}\right) |||v_h|||_h \geq \frac{\beta_c}{2} |||v_h|||_h,\end{aligned}$$

for a sufficiently small mesh size h such that $h^{\min\{\deg-1,\mathbb{k}_u-2\}} < \frac{\beta_c}{2C_2}$. Here, C_2 depends on C_1 and $\|u\|_{\mathbb{k}_u}$ and $\mathbb{k}_u \geq 4$ gives the regularity of u, i.e., $u \in H^{\mathbb{k}_u}(\Omega)$. Therefore, the inf-sup condition eq. (38) holds. ☐

2.1.2. *Convergence analysis*

We proceed to the error analysis for the discrete nonlinear problem eq. (13). Let $\mathscr{B}_\rho(I_hu) := \{v_h \in W_h : |||I_hu - v_h|||_h \leq \rho\}$. We define the nonlinear map $\mu_h : W_h \to W_h$ by

$$\langle \mathscr{D}\mathscr{N}_h^s(I_hu)\mu_h(v_h), w_h\rangle = 3B^s(I_hu, I_hu, v_h, w_h) + L^s(w_h) - B^s(v_h,v_h,v_h,w_h) \tag{39}$$

for $v_h, w_h \in W_{h,0}$. Due to the weak coercivity property in theorem 2.8, the nonlinear map μ_h is well-defined.

The existence and local uniqueness of the solution u_h to the discrete nonlinear problem eq. (13) will be proven via an application of Brouwer's fixed point theorem, which necessitates the use of two auxiliary lemmas illustrating that (i) μ_h maps from a ball to itself; and (ii) the map μ_h is contracting.

Lemma 2.9. *(Mapping from a ball to itself) Let u be a regular isolated solution of the continuous nonlinear weak problem eq.* (11). *For a sufficiently large ε and a sufficiently small mesh size h, there exists a positive constant $R(h) > 0$ such that:*

$$|||v_h - I_hu|||_h \leq R(h) \Rightarrow |||\mu_h(v_h) - I_hu|||_h \leq R(h) \quad \forall v_h \in W_{h,0}.$$

Proof. Note that the solution $u \in H^2(\Omega)\cap H_0^1(\Omega)$ of eq. (11) satisfies the discrete weak formulation eq. (13) due to the consistency result theorem 2.3, that is to say, there holds that

$$A_h^s(u,w_h) + P_h^s(u,w_h) + B^s(u,u,u,w_h) + C^s(u,w_h) = L^s(w_h) \quad \forall w_h \in W_{h,0}. \tag{40}$$

By the linearity of $\langle \mathscr{D}\mathscr{N}_h^s(I_hu)\cdot,\cdot\rangle_{H^2}$, the definition eq. (39) of the nonlinear map μ_h and eq. (40), we calculate

$$\begin{aligned}\langle \mathscr{D}\mathscr{N}_h^s(I_hu)(I_hu - \mu_h(v_h)), w_h\rangle &= \langle \mathscr{D}\mathscr{N}_h^s(I_hu)I_hu, w_h\rangle - \langle \mathscr{D}\mathscr{N}_h^s(I_hu)\mu_h(v_h), w_h\rangle \\ &= A_h^s(I_hu,w_h) + P_h^s(I_hu,w_h) + 3B^s(I_hu,I_hu,I_hu,w_h) + C^s(I_hu,w_h) \\ &\quad - 3B^s(I_hu,I_hu,v_h,w_h) + B^s(v_h,v_h,v_h,w_h) - L^s(w_h) \\ &= \underbrace{A_h^s(I_hu-u,w_h) + P_h^s(I_hu-u,w_h)}_{=:\mathfrak{N}_1} + \underbrace{C^s(I_hu-u,w_h)}_{=:\mathfrak{N}_2} + \underbrace{\left(B^s(I_hu,I_hu,I_hu,w_h) - B^s(u,u,u,w_h)\right)}_{=:\mathfrak{N}_3} \\ &\quad + \underbrace{\left(2B^s(I_hu,I_hu,I_hu,w_h) - 3B^s(I_hu,I_hu,v_h,w_h) + B^s(v_h,v_h,v_h,w_h)\right)}_{=:\mathfrak{N}_4}\end{aligned}$$

In what follows, we give upper bounds for each $\mathfrak{N}_i, i=1,2,3,4$. Using the boundedness of $A_h^s+P_h^s, C^s$ and the interpolation estimate [9, Eq. (5.3)] in the $|||\cdot|||$-norm, we obtain

$$\mathfrak{N}_1 \lesssim |||I_h u - u|||_h |||w_h|||_h \lesssim h^{\min\{\deg-1,\Bbbk_u-2\}} |||w_h|||_h,$$
$$\mathfrak{N}_2 \lesssim |||I_h u - u|||_h |||w_h|||_h \lesssim h^{\min\{\deg-1,\Bbbk_u-2\}} |||w_h|||_h.$$

We rearrange terms in $\mathfrak{N}_3$ and use the boundedness result lemma 2.4 and the interpolation result [9, Eq. (5.3)] to obtain

$$\begin{aligned}
\mathfrak{N}_3 &= B^s(I_h u-u, I_h u-u, I_h u, w_h) + 2B^s(I_h u-u, I_h u-u, u, w_h) + 3B^s(u,u,I_h u-u, w_h) \\
&\lesssim \left(|||I_h u-u|||_h^2 |||I_h u|||_h + |||I_h u-u|||_h^2 |||u|||_h + \|u\|_2^2 \|I_h u-u\|_0 \right) |||w_h|||_h \\
&\lesssim \left(h^{2\min\{\deg-1,\Bbbk_u-2\}} + h^{\min\{\deg+1,\Bbbk_u\}} \right) |||w_h|||_h.
\end{aligned}$$

Let $e_I = v_h - I_h u$. We use the definition of $B^s(\cdot,\cdot,\cdot,\cdot)$ and its boundedness to deduce that

$$\begin{aligned}
\mathfrak{N}_4 &= a_3 \int_\Omega \left\{ 2(I_h u)^3 w_h - 3(I_h u)^2 v_h w_h + v_h^3 w_h \right\} \\
&= a_3 \int_\Omega \left\{ \left(v_h^2 - (I_h u)^2\right) v_h w_h + 2(I_h u)^2 (I_h u - v_h) w_h \right\} \\
&= a_3 \int_\Omega \left\{ e_I (e_I + 2I_h u)(e_I + I_h u) w_h - 2(I_h u)^2 e_I w_h \right\} \\
&= a_3 \int_\Omega \left\{ e_I \left(e_I^2 + 3e_I I_h u + 2(I_h u)^2\right) w_h - 2(I_h u)^2 e_I w_h \right\} \\
&= a_3 \int_\Omega \left(e_I^3 + 3e_I^3 I_h u\right) w_h = B^s(e_I, e_I, e_I, w_h) + 3B^s(e_I, e_I, I_h u, w_h) \\
&\lesssim |||e_I|||_h^2 \left(|||e_I|||_h + |||I_h u|||_h\right) |||w_h|||_h.
\end{aligned}$$

Hence, we combine the above bounds for $\mathfrak{N}_i$, $i=1,2,3,4$ to have

$$\langle D\mathcal{N}_h^s(I_h u)(I_h u - \mu_h(v_h)), w_h \rangle \lesssim \left(h^{\min\{\deg-1,\Bbbk_u-2\}} + h^{\min\{2\deg-2, 2\Bbbk_u-4, \deg+1, \Bbbk_u\}} + |||e_I|||_h^2 \left(|||e_I|||_h + 1\right) \right) |||w_h|||_h.$$

By the inf-sup condition eq. (38) for the perturbed bilinear form, we further deduce that there exists a $w_h \in W_h$ with $|||w_h|||_h = 1$ such that $|||I_h u - \mu_h(v_h)|||_h \lesssim \langle D\mathcal{N}_h^s(I_h u)(I_h u - \mu_h(v_h)), w_h \rangle$. Since $|||e_I|||_h \le R(h)$, we obtain

$$\begin{aligned}
|||I_h u - \mu_h(v_h)|||_h &\lesssim \left(h^{\min\{\deg-1,\Bbbk_u-2\}} + h^{\min\{2\deg-2, 2\Bbbk_u-4, \deg+1, \Bbbk_u\}} + R(h)^2 \left(R(h)+1\right) \right) \\
&\le \begin{cases} C_u \left(2h^{\min\{\deg-1,\Bbbk_u-2\}} + R(h)^2(1+R(h))\right) & \text{for } 2 \le \deg \le 3, \Bbbk_u \le 4, \\ C_u \left(h^{\min\{\deg-1,\Bbbk_u-2\}} + h^{\min\{\deg+1, 2\Bbbk_u-4\}} + R(h)^2(1+R(h))\right) & \text{for } \deg > 3, \Bbbk_u \le 4. \end{cases}
\end{aligned}$$

Note that there are other cases when $\Bbbk_u > 4$ and we only focus on the case of $\Bbbk_u \le 4$ here for brevity. The idea of the remainder of the proof is to choose an appropriate $R(h)$ so that $|||I_h u - \mu_h(v_h)|||_h \le R(h)$. For simplicity of the calculation, we illustrate the case when $2 \le \deg \le 3, \Bbbk_u \le 4$. To this end, we take $R(h) = 4C_u h^{\min\{\deg-1,\Bbbk_u-2\}}$ and choose h satisfying $h^{2\min\{\deg-1,\Bbbk_u-2\}} \le \frac{1}{32C_u} - \frac{1}{16}$. This yields

$$\begin{aligned}
|||I_h u - \mu_h(v_h)|||_h &\le 2C_u h^{\min\{\deg-1,\Bbbk_u-2\}} \left(1 + C_u R(h)^2 + C_u\right) \\
&= 2C_u h^{\min\{\deg-1,\Bbbk_u-2\}} \left(1 + 32C_u^3 h^{2\min\{\deg-1,\Bbbk_u-2\}} + 2C_u\right) \le R(h).
\end{aligned}$$

This completes the proof. □

Lemma 2.10. *(Contraction result) For a sufficiently large ε, a sufficiently small mesh size h and any $v_1, v_2 \in \mathscr{B}_{R(h)}(I_h u)$, there holds*

$$|||\mu_h(v_1) - \mu_h(v_2)|||_h \lesssim h^{\min\{\deg -1, \Bbbk_u - 2\}} |||v_1 - v_2|||_h. \tag{41}$$

Proof. For $w_h \in W_h$, we use the definition eq. (39) of the nonlinear map μ_h, the definition eq. (37) and linearity of $\langle \mathscr{D}\mathscr{N}_h^s(I_h u)\cdot, \cdot\rangle$ to calculate

$$\begin{aligned}
&\langle \mathscr{D}\mathscr{N}_h^s(I_h u)(\mu_h(v_1) - \mu_h(v_2)), w_h\rangle \\
&= 3B^s(I_h u, I_h u, v_1, w_h) - B^s(v_1, v_1, v_1, w_h) - 3B^s(I_h u, I_h u, v_2, w_h) + B^s(v_2, v_2, v_2, w_h) \\
&= a_3 \int_\Omega \left(3(I_h u)^2 v_1 w_h - v_1^3 w_h\right) - a_3 \int_\Omega \left(3(I_h u)^2 v_2 w_h - v_2^3 w_h\right) \\
&= a_3 \int_\Omega \left(\left((I_h u)^2 - v_1^2\right) v_1 w_h + 2(I_h u)^2 (v_1 - v_2) w_h - \left((I_h u)^2 - v_2^2\right) v_2 w_h\right) \\
&= a_3 \int_\Omega ((I_h u - v_1)(v_1 - I_h u)(v_1 - v_2) w_h + 2(I_h u - v_1) I_h u (v_1 - v_2) w_h + (I_h u - v_1)(I_h u + v_1) v_2 w_h) \\
&\quad + 2a_3 \int_\Omega (I_h u (v_1 - v_2)(I_h u - v_2) w_h + I_h u (v_1 - v_2) v_2 w_h) - a_3 \int_\Omega (I_h u - v_2)(I_h u + v_2) v_2 w_h \\
&= a_3 \int_\Omega (I_h u - v_1)(v_1 - I_h u)(v_1 - v_2) w_h + 2a_3 \int_\Omega (I_h u - v_1) I_h u (v_1 - v_2) w_h \\
&\quad + 2a_3 \int_\Omega (I_h u - v_2) I_h u (v_1 - v_2) w_h + a_3 \int_\Omega (v_1 - v_2)\left((I_h u - v_1) + (I_h u - v_2)\right)\left((v_2 - I_h u) + I_h u\right) w_h.
\end{aligned}$$

Let $e_1 = I_h u - v_1$, $e_2 = I_h u - v_2$ and $e = v_1 - v_2$. We make some elementary manipulations and use the boundedness of B^s and the inequality of geometric and arithmetic means to yield

$$\begin{aligned}
&\langle \mathscr{D}\mathscr{N}_h^s(I_h u)(\mu_h(v_1) - \mu_h(v_2)), w_h\rangle \\
&= a_3 \int_\Omega (-e_1^2) e w_h + 2a_3 \int_\Omega e_1 (I_h u) e w_h + 2a_3 \int_\Omega e_2 (I_h u) e w_h + a_3 \int_\Omega \{e w_h (e_1 I_h u + e_2 I_h u - e_1 e_2 - e_2^2)\} \\
&\lesssim \left(|||e_1|||_h^2 + |||I_h u|||_h |||e_1|||_h + |||e_2|||_h |||I_h u|||_h + |||e_1|||_h |||e_2|||_h + |||e_2|||_h^2\right) |||e|||_h |||w_h|||_h \\
&\lesssim \left(|||e_1|||_h^2 + |||e_2|||_h^2 + |||e_1|||_h + |||e_2|||_h\right) |||e|||_h |||w_h|||_h \lesssim \left(R(h)^2 + R(h)\right) |||e|||_h |||w_h|||_h.
\end{aligned}$$

By the inf-sup condition eq. (38), we know that there exists $w_h \in W_h$ with $|||w_h|||_h = 1$ such that

$$\frac{\beta_c}{2} |||\mu_h(v_1) - \mu_h(v_2)|||_h \lesssim \langle \mathscr{D}\mathscr{N}_h^s(I_h u)(\mu_h(v_1) - \mu_h(v_2)), w_h\rangle.$$

Therefore, we have $|||\mu_h(v_1) - \mu_h(v_2)|||_h \lesssim R(h)(1 + R(h))|||e|||_h$. Noting that $R(h)(1 + R(h)) < 1$ for $0 < R(h) < \frac{1}{2}$ completes the proof. □

The existence and local uniqueness of the discrete solution u_h can now be obtained via the application of Brouwer's fixed point theorem [20].

Theorem 2.11. *(Convergence in $|||\cdot|||_h$-norm) Let u be a regular isolated solution of the nonlinear problem eq.* (11). *For a sufficiently large ε and a sufficiently small h, there exists a unique solution u_h of the discrete nonlinear problem eq.* (13) *within the local ball $\mathscr{B}_{R(h)}(I_h u)$. Furthermore, we have $|||u - u_h|||_h \lesssim h^{\min\{\deg -1, \Bbbk_u - 2\}}$.*

Proof. A use of lemma 2.9 yields that the nonlinear map μ_h maps a closed convex set $\mathscr{B}_{R(h)}(I_h u) \subset W_h$ to itself. Moreover it is a contracting map. Therefore, an application of Brouwer's fixed point theorem yields that μ_h has at least one fixed point, say u_h, in this ball $\mathscr{B}_{R(h)}(I_h u)$. The uniqueness of the solution to eq. (13) in that ball $\mathscr{B}_{R(h)}(I_h u)$ follows from the contraction result in lemma 2.10. Meanwhile, we have by lemma 2.9 that

$$|||u_h - I_h u|||_h \lesssim h^{\min\{\deg -1, \Bbbk_u - 2\}}. \tag{42}$$

The error estimate is then obtained straightforwardly using the triangle inequality $|||u - u_h|||_h \le |||u - I_h u|||_h + |||I_h u - u_h|||_h$ combined with eq. (42) and the interpolation estimate [9, Eq. (5.3)]. □

It follows from theorem 2.11 that optimal convergence rates are achieved in the mesh-dependent norm $|||\cdot|||_h$. This will be numerically verified in section 3.

2.1.3. *Estimates in the L^2-norm*

We derive an L^2 error estimate using a duality argument in this subsection. To this end, we consider the following linear dual problem to the primal nonlinear problem eq. (9):

$$\begin{cases} 2B\nabla \cdot (\nabla \cdot (\mathscr{D}^2 \chi)) + a_1 \chi + 3a_3 u^2 \chi = f_{dual} & \text{in } \Omega, \\ \chi = 0, \quad \mathscr{D}^2 \chi = \mathbf{0} & \text{on } \partial\Omega, \end{cases} \tag{43}$$

for $f_{dual} \in L^2(\Omega)$. For smooth domains Ω, it can be deduced by a classical elliptic regularity result that $\chi \in H^4(\Omega)$. The corresponding weak form is: find $\chi \in H^2(\Omega) \cap H_0^1(\Omega)$ such that

$$2B \int_\Omega \mathscr{D}^2 \chi : \mathscr{D}^2 v + a_1 \int_\Omega \chi v + 3a_3 \int_\Omega u^2 \chi v = \int_\Omega f_{dual} v \quad \forall v \in H^2(\Omega) \cap H_0^1(\Omega),$$

that is to say,

$$\langle \mathscr{D}\mathscr{N}^s(u)\chi, v \rangle_{H^2} = \langle \mathscr{D}\mathscr{N}_h^s(u)\chi, v \rangle = (f_{dual}, v)_0. \tag{44}$$

Remark 2.12. *The first equality in eq.* (44) *holds since $u \in H^2(\Omega), \chi \in H^2(\Omega)$ and $v \in H^2(\Omega)$.*

We give two auxiliary results in the following.

Lemma 2.13. *For $u \in H^{\Bbbk_u}(\Omega)$, $\Bbbk_u > 2$, $\chi \in H^4(\Omega) \cap H_0^1(\Omega)$ and $I_h u \in W_{h,0} \subset H_0^1(\Omega)$, there holds that*

$$A_h^s(I_h u - u, \chi) + P_h^s(I_h u - u, \chi) \lesssim h^{\min\{\deg +1, \Bbbk_u\}} \|\chi\|_4.$$

Proof. Note that $[\![\nabla \chi]\!] = 0$ since $\chi \in H^4(\Omega)$ and $\chi = 0$ on $\partial\Omega$. We calculate

$$\begin{aligned} A^s(I_h u - u, \chi) + P_h^s(I_h u - u, \chi) &= \sum_{T \in \mathscr{E}_h} \int_T 2B\mathscr{D}^2(I_h u - u) : \mathscr{D}^2 \chi - 2B \sum_{e \in \mathscr{E}_I} \left\{\!\!\left\{ \frac{\partial^2 (I_h u - u)}{\partial \nu^2} \right\}\!\!\right\} [\![\nabla \chi]\!] \\ &\quad - 2B \sum_{e \in \mathscr{E}_I} \left\{\!\!\left\{ \frac{\partial^2 \chi}{\partial \nu^2} \right\}\!\!\right\} [\![\nabla (I_h u - u)]\!] + \sum_{e \in \mathscr{E}_I} \frac{2B\varepsilon}{h_e^3} \int_e [\![\nabla (I_h u - u)]\!] [\![\nabla \chi]\!] \\ &= \sum_{T \in \mathscr{E}_h} \int_T 2B\mathscr{D}^2(I_h u - u) : \mathscr{D}^2 \chi - 2B \sum_{e \in \mathscr{E}_I} \left\{\!\!\left\{ \frac{\partial^2 \chi}{\partial \nu^2} \right\}\!\!\right\} [\![\nabla (I_h u - u)]\!] \\ &= \sum_{T \in \mathscr{E}_h} \int_T 2B(I_h u - u) \nabla \cdot (\nabla \cdot (\mathscr{D}^2 \chi)) \\ &\lesssim \|I_h u - u\|_0 \|\nabla \cdot (\nabla \cdot (\mathscr{D}^2 \chi))\|_0 \lesssim h^{\min\{\deg +1, \Bbbk_u\}} \|\chi\|_4. \end{aligned}$$

Here, the last, second last, and third last steps follow from standard interpolation estimates, the Cauchy–Schwarz inequality, and integration by parts twice, respectively. □

Lemma 2.14. *The solution χ of the linear dual problem eq.* (43) *belongs to $H^4(\Omega)$ on a smooth domain Ω and it holds that*

$$\|\chi\|_4 \lesssim \|f_{dual}\|_0. \tag{45}$$

Proof. We can use the inf-sup condition eq. (19) for the linear operator $\langle \mathscr{D}\mathscr{N}^s(u)\cdot,\cdot\rangle$, the weak form eq. (44) and the Cauchy–Schwarz inequality to obtain

$$|\chi|_2 \lesssim \sup_{\substack{w\in H^2\cap H_0^1\\ |w|_2=1}} \langle \mathscr{D}\mathscr{N}^s(u)\chi, w\rangle_{H^2} = \sup_{\substack{w\in H^2\cap H_0^1\\ |w|_2=1}} (f_{dual}, w)_0 \lesssim \|f_{dual}\|_0 \underbrace{\|w\|_0}_{\lesssim |w|_2=1}. \tag{46}$$

By the form of eq. (44), the boundedness of $B^s(u,u,\cdot,\cdot)$ and $C^s(\cdot,\cdot)$, and eq. (46), we have

$$\|\nabla\cdot(\nabla\cdot(\mathscr{D}^2\chi))\|_0 = \|-3B^s(u,u,\chi,\cdot) - C^s(\chi,\cdot) + (f_{dual},\cdot)_0\|_0 \lesssim \|\chi\|_0 + \|f_{dual}\|_0 \lesssim \|f_{dual}\|_0. \tag{47}$$

Using a bootstrapping argument in elliptic regularity (see, e.g., [18, Section 6.3]), we can deduce that $\chi \in H^4(\Omega)$ in a smooth domain Ω. The regularity estimate eq. (45) follows from eq. (47). □

We are ready to derive the L^2 a priori error estimates.

Theorem 2.15. *(L^2 error estimate) Under the same conditions as theorem 2.11 and assuming further that* $\deg > 1$ *(since the problem is fourth-order), the discrete solution u_h approximates u such that*

$$\|u-u_h\|_0 \lesssim \begin{cases} h^{\min\{\deg+1,\Bbbk_u\}} & \textit{for } \deg \geq 3,\\ h^{2\min\{\deg-1,\Bbbk_u-2\}} & \textit{for } \deg = 2. \end{cases} \tag{48}$$

Proof. Taking $f_{dual} = I_h u - u_h \in W_h \subset H^1(\Omega)\cap H^2(\mathscr{T}_h)$ in eq. (43) and multiplying eq. (43) by a test function $v_h = I_h u - u_h$ and integrating by parts, we obtain $\langle \mathscr{D}\mathscr{N}_h^s(u)\chi, I_h u - u_h\rangle = \|I_h u - u_h\|_0^2$. It follows from the fact that $u \in H^{\Bbbk_u}(\Omega)$, $\Bbbk_u \geq 4$, and the definition eq. (11) of the nonlinear continuous weak form $\mathscr{N}^s(u)\cdot$ that

$$\begin{aligned}
\|I_h u - u_h\|_0^2 &= \langle \mathscr{D}\mathscr{N}_h^s(u)\chi, I_h u - u_h\rangle + \mathscr{N}_h^s(u_h)(I_h\chi) - \mathscr{N}_h^s(u)(I_h\chi)\\
&= A_h^s(\chi, I_h u - u_h) + P_h^s(\chi, I_h u - u_h) + C^s(\chi, I_h u - u_h) + 3B^s(u,u,\chi, I_h u - u_h)\\
&\quad + A_h^s(u_h, I_h\chi) + P_h^s(u_h, I_h\chi) + C^s(u_h, I_h\chi) + B^s(u_h,u_h,u_h,I_h\chi)\\
&\quad - A_h^s(u, I_h\chi) - P_h^s(u, I_h\chi) - C^s(u, I_h\chi) - B^s(u,u,u,I_h\chi)\\
&= \underbrace{A_h^s(I_h u - u, \chi) + A_h^s(u-u_h, \chi - I_h\chi) + P_h^s(I_h u - u, \chi) + P_h^s(u-u_h, \chi - I_h\chi)}_{=:\mathfrak{U}_1}\\
&\quad + \underbrace{C^s(I_h u - u, \chi) + C^s(u-u_h, \chi - I_h\chi)}_{=:\mathfrak{U}_2}\\
&\quad + \underbrace{3B^s(u,u,I_h u - u_h, \chi - I_h\chi) + 3B^s(u,u,I_h u - u, I_h\chi)}_{=:\mathfrak{U}_3}\\
&\quad + \underbrace{B^s(u_h,u_h,u_h,I_h\chi) - 3B^s(u,u,u_h,I_h\chi) + 2B^s(u,u,u,I_h\chi)}_{=:\mathfrak{U}_4}.
\end{aligned}$$

We bound each $\mathfrak{U}_i$ separately using the boundedness of A_h^s, P_h^s, B^s and C^s, theorem 2.11 and standard interpolation estimates. This leads to

$$\begin{aligned}
\mathfrak{U}_1 &\lesssim h^{\min\{\deg+1,\Bbbk_u\}}\|\chi\|_4 + |||u-u_h|||_h |||\chi - I_h\chi|||_h \lesssim h^{\min\{\deg+1,\Bbbk_u\}}\|\chi\|_4,\\
\mathfrak{U}_2 &\lesssim \|I_h u - u\|_0\|\chi\|_0 + |||u-u_h|||_h |||\chi - I_h\chi|||_h \lesssim h^{\min\{\deg+1,\Bbbk_u\}}\|\chi\|_4,
\end{aligned}$$

$$\mathfrak{U}_3 \lesssim \|u\|_2^2 |||I_h u - u_h|||_h |||\chi - I_h\chi|||_h + \|u\|_2^2 \|I_h u - u\|_0 \|I_h\chi\|_0 \lesssim h^{\min\{\deg+1,\Bbbk_u\}}\|\chi\|_4.$$

Setting $e_3 = u_h - u$ and estimating $\mathfrak{U}_4$ as in $\mathfrak{R}_4$ of lemma 2.9 with the use of theorem 2.11 and $|||I_h\chi|||_h \lesssim \|\chi\|_2 \le \|\chi\|_4$ yields

$$\mathfrak{U}_4 \lesssim |||e_3|||_h^2 (|||e_3|||_h + |||u|||_h)\, |||I_h\chi|||_h \lesssim h^{2\min\{\deg-1,\Bbbk_u-2\}}(h^{\min\{\deg-1,\Bbbk_u-2\}}+1)\|\chi\|_4.$$

Combining the above estimates for $\mathfrak{U}_i$ $(i = 1,2,3,4)$ and using the regularity estimate eq. (45) and $\|\chi\|_4 \lesssim \|I_h u - u_h\|_0$, we obtain

$$\|I_h u - u_h\|_0 \lesssim \begin{cases} h^{\min\{\deg+1,\Bbbk_u\}} & \text{if } \deg \ge 3, \\ h^{2\min\{\deg-1,\Bbbk_u-2\}} & \text{if } \deg = 2. \end{cases}$$

Hence, eq. (48) follows from the triangle inequality and standard interpolation estimates. □

theorem 2.15 implies that for quadratic approximations to the sufficiently regular solution of eq. (9), there is a sub-optimal convergence rate in the L^2-norm while for higher order (≥ 3) approximations, we expect optimal L^2 error rates. We shall see numerical verifications of this in the subsequent sections.

2.1.4. *The inconsistent discrete form*

The above analysis considers the consistent weak formulation eq. (13). In practice, Xia et al. [31] adopted the inconsistent discrete weak form in the implementation due to its cheaper assembly cost: find $u_h \in W_{h,b}$ such that

$$\tilde{\mathscr{N}}_h^s(u_h)t_h = \tilde{A}_h^s(u_h,t_h) + B^s(u_h,u_h,u_h,t_h) + C^s(u_h,t_h) + P_h^s(u_h,t_h) = 0 \quad \forall t_h \in W_{h,0}, \tag{49}$$

where $\tilde{A}_h^s(u,t) := 2B\sum_{T\in\mathscr{T}_h}\int_T \mathscr{D}^2u : \mathscr{D}^2t$. Comparing $\tilde{A}_h^s$ and A_h^s, the missing terms are the interior facet integrals arising from piecewise integration by parts and symmetrisation. Due to the absence of these terms in $\tilde{A}_h^s$, one can immediately notice that the discrete weak formulation eq. (49) is inconsistent in the sense that the solution u of the strong form eq. (9) does not satisfy the weak form eq. (49), as opposed to the result of theorem 2.3.

Despite this inconsistency, in practice this also leads to a convergent numerical scheme with similar convergence rates, as illustrated in section 3. This is not surprising; a similar idea has also been applied and introduced as *weakly over-penalised symmetric interior penalty* (WOPSIP) methods in [10] for second-order elliptic PDEs and in [8] for biharmonic equations.

Remark 2.16. *The excessive size of the penalty parameter in the WOPSIP method could induce ill-conditioned linear systems. No such effects are observed in our numerical results.*

2.2. A priori error estimates for $(\mathscr{P}2)$

Problem $(\mathscr{P}2)$ is a special form of the classical LdG model of nematic LC. Finite element analysis for a more general form using conforming discretisations has been studied in [13, 14]. More specifically, Davis and Gartland [14] gave an abstract nonlinear finite element convergence analysis where an optimal H^1 error bound is proved on convex domains with piecewise linear polynomial approximations, but do not derive an error bound in the L^2 norm. Recently, Maity, Majumdar and Nataraj [21] analysed the discontinuous Galerkin finite element method for a two-dimensional reduced LdG free energy, where optimal a priori error estimates in the L^2-norm with exact solutions in H^2 are achieved for a piecewise linear discretisation. Both works only focus on piecewise linear approximations. In this section, we will follow similar steps to section 2.1 to prove the H^1- and L^2-convergence rates for the problem $(\mathscr{P}2)$ with the use of common continuous Lagrange elements of arbitrary positive degree. Since the approach is similar to the previous subsections, we omit some details for brevity.

The continuous weak formulation of $(\mathscr{P}2)$ in two dimensions (the three-dimensional case can be tackled similarly) is given by: find $\mathbf{Q} \in H_b^1(\Omega, S_0)$ such that

$$\mathscr{N}^n(\mathbf{Q})\mathbf{P} := A^n(\mathbf{Q},\mathbf{P}) + B^n(\mathbf{Q},\mathbf{Q},\mathbf{Q},\mathbf{P}) + C^n(\mathbf{Q},\mathbf{P}) = 0 \quad \forall \mathbf{P} \in \mathbf{H}_0^1(\Omega), \tag{50}$$

where the bilinear forms are $A^n(\mathbf{Q},\mathbf{P}) := K\int_\Omega \nabla\mathbf{Q} \vdots \nabla\mathbf{P}$, $C^n(\mathbf{Q},\mathbf{P}) := -2l\int_\Omega \mathbf{Q} : \mathbf{P}$, and the nonlinear operator is given by

$$B^n(\Psi,\Phi,\Theta,\Xi) := \frac{4l}{3}\int_\Omega ((\Psi:\Phi)(\Theta:\Xi)+2(\Psi:\Theta)(\Phi:\Xi)). \tag{51}$$

Since eq. (50) is nonlinear, we need to approximate the solution of its linearised version, i.e., find $\Theta \in \mathbf{H}_0^1(\Omega)$ such that

$$\langle \mathscr{D}\mathscr{N}^n(\mathbf{Q})\Theta,\Phi\rangle := A^n(\Theta,\Phi)+3B^n(\mathbf{Q},\mathbf{Q},\Theta,\Phi)+C^n(\Theta,\Phi) = -\mathscr{N}^n(\mathbf{Q})\Phi \quad \forall\Phi \in \mathbf{H}_0^1(\Omega), \tag{52}$$

where $\langle\cdot,\cdot\rangle$ represents the dual pairing between $\mathbf{H}^{-1}(\Omega)$ and $\mathbf{H}_0^1(\Omega)$.

Suppose $\mathbf{Q}_h \in \mathbf{V}_h$ approximates the solution of eq. (50) with the conforming finite element method on a finite dimensional space $\mathbf{V}_h := \{\mathbf{P} \in \mathbf{H}^1(\Omega) : \mathbf{P} \in \mathbb{Q}_{\deg}(T), \deg \geq 1, \forall T \in \mathscr{T}_h\}$. Throughout this subsection we take deg ≥ 1. Furthermore, we denote $\mathbf{V}_{h,0} := \{\mathbf{P} \in \mathbf{V}_h : \mathbf{P} = \mathbf{0} \text{ on } \partial\Omega\}$ and $\mathbf{V}_{h,b} := \{\mathbf{P} \in \mathbf{V}_h : \mathbf{P} = \mathbf{Q}_b \text{ on } \partial\Omega\}$. We assume that the minimiser $\mathbf{Q}$ to be approximated is isolated, i.e., the linearised operator $\langle \mathscr{D}\mathscr{N}^n(\mathbf{Q})\cdot,\cdot\rangle$ is nonsingular. This is equivalent to the following continuous inf-sup condition [21, Eq. (2.8)]:

$$0 < \beta_Q := \inf_{\substack{\Theta\in\mathbf{H}_0^1(\Omega)\\ \|\Theta\|_1=1}} \sup_{\substack{\Phi\in\mathbf{H}_0^1(\Omega)\\ \|\Phi\|_1=1}} \langle \mathscr{D}\mathscr{N}^n(\mathbf{Q})\Theta,\Phi\rangle = \inf_{\substack{\Phi\in\mathbf{H}_0^1(\Omega)\\ \|\Phi\|_1=1}} \sup_{\substack{\Theta\in\mathbf{H}_0^1(\Omega)\\ \|\Theta\|_1=1}} \langle \mathscr{D}\mathscr{N}^n(\mathbf{Q})\Theta,\Phi\rangle. \tag{53}$$

With this inf-sup condition for $\langle \mathscr{D}\mathscr{N}^n(\mathbf{Q})\cdot,\cdot\rangle$, we can obtain a stability result for the perturbed bilinear form $\langle \mathscr{D}\mathscr{N}^n(I_h\mathbf{Q})\cdot,\cdot\rangle$ by following similar steps as in the proof of theorem 2.8.

Theorem 2.17. *(Stability of the perturbed bilinear form) Let $\mathbf{Q}$ be a regular isolated solution of the nonlinear continuous weak form eq.* (50) *and $I_h\mathbf{Q}$ the interpolant of $\mathbf{Q}$. For a sufficiently small mesh size h, the following discrete inf-sup condition holds:*

$$0 < \frac{\beta_Q}{2} \leq \inf_{\substack{\Theta\in\mathbf{H}_0^1(\Omega)\\ \|\Theta\|_1=1}} \sup_{\substack{\Phi\in\mathbf{H}_0^1(\Omega)\\ \|\Phi\|_1=1}} \langle \mathscr{D}\mathscr{N}^n(I_h\mathbf{Q})\Theta,\Phi\rangle. \tag{54}$$

We give some auxiliary results about the operators $A^n(\cdot,\cdot)$, $B^n(\cdot,\cdot,\cdot,\cdot)$ and $C^n(\cdot,\cdot)$ that can be verified via the Cauchy–Schwarz inequality, the Poincaré inequality, and Sobolev embeddings.

Lemma 2.18. *(Boundedness and coercivity of $A^n(\cdot,\cdot)$) For $\Theta,\Phi \in \mathbf{H}_0^1(\Omega)$, there holds*

$$A^n(\Theta,\Phi) \lesssim \|\Theta\|_1\|\Phi\|_1 \text{ and } \|\Theta\|_1^2 \lesssim A^n(\Theta,\Theta).$$

Lemma 2.19. *(Boundedness of $B^n(\cdot,\cdot,\cdot,\cdot)$, $C^n(\cdot,\cdot)$) For $\Psi,\Phi,\Theta,\Xi \in \mathbf{H}^1(\Omega)$, there holds*

$$B^n(\Psi,\Phi,\Theta,\Xi) \lesssim \|\Psi\|_1\|\Phi\|_1\|\Theta\|_1\|\Xi\|_1, \quad C^n(\Psi,\Phi) \lesssim \|\Psi\|_1\|\Phi\|_1, \tag{55}$$

and for $\Psi,\Phi \in \mathbf{H}^{\Bbbk}(\Omega)$, $\Bbbk \geq 2$, $\Theta,\Xi \in \mathbf{H}^1(\Omega)$,

$$B^n(\Psi,\Phi,\Theta,\Xi) \lesssim \|\Psi\|_{\Bbbk}\|\Phi\|_{\Bbbk}\|\Theta\|_1\|\Xi\|_1. \tag{56}$$

To proceed to error estimates for the nonlinear problem eq. (50), we define the nonlinear map $\psi : \mathbf{V}_h \to \mathbf{V}_h$ by

$$\langle \mathscr{D}\mathscr{N}^n(I_h\mathbf{Q})\psi(\Theta_h),\Phi_h\rangle = 3B^n(I_h\mathbf{Q},I_h\mathbf{Q},\Theta_h,\Phi_h) - B^n(\Theta_h,\Theta_h,\Theta_h,\Phi_h)$$

for $\Theta_h,\Phi_h \in \mathbf{V}_{h,0}$. Due to the stability result of theorem 2.17, the nonlinear map ψ is well-defined. We define the local ball $\mathscr{B}_\rho(I_h\mathbf{Q}) := \{\mathbf{P}_h \in \mathbf{V}_h : \|I_h\mathbf{Q} - \mathbf{P}_h\|_1 \leq \rho\}$. The following two auxiliary lemmas provide the necessary components for the application of Brouwer's fixed point theorem.

Lemma 2.20. *(Mapping from a ball to itself) Let* $\mathbf{Q}$ *be a regular isolated solution of the continuous nonlinear weak problem eq.* (50)*. For a sufficiently small mesh size h, there exists a positive constant* $r(h) > 0$ *such that:*

$$\|\mathbf{P}_h - I_h\mathbf{Q}\|_1 \leq r(h) \Rightarrow \|\psi(\mathbf{P}_h) - I_h\mathbf{Q}\|_1 \leq r(h) \quad \forall \mathbf{P}_h \in \mathbf{V}_{h,0}.$$

Remark 2.21. *In fact, the choice of* $r(h)$ *can be taken as* $r(h) = \mathscr{O}(h^{\min\{\deg,\Bbbk_Q-1\}})$ *in the proof of lemma 2.20. Here,* $\Bbbk_Q \geq 2$ *denotes the regularity index of* $\mathbf{Q}$*, i.e.,* $\mathbf{Q} \in \mathbf{H}^{\Bbbk_Q}(\Omega)$.

Lemma 2.22. *(Contraction result) For a sufficiently small mesh size h and any* $\mathbf{P}_1, \mathbf{P}_2 \in \mathscr{B}_{r(h)}(I_h\mathbf{Q})$*, there holds*

$$\|\psi(\mathbf{P}_1) - \psi(\mathbf{P}_2)\|_1 \lesssim h^{\min\{\deg,\Bbbk_Q-1\}}\|\mathbf{P}_1 - \mathbf{P}_2\|_1. \tag{57}$$

Remark 2.23. *In the proof of lemma 2.22, we have particularly used the stability property of the perturbed bilinear form as given by theorem 2.17.*

Hence, the existence and local uniqueness of the discrete solution $\mathbf{Q}_h$ can be derived by following similar steps as in the proof of theorem 2.11.

Theorem 2.24. *(Convergence in* $\|\cdot\|_1$*-norm) Let* $\mathbf{Q}$ *be a regular isolated solution of the nonlinear problem eq.* (50)*. For a sufficiently small h, there exists a unique solution* $\mathbf{Q}_h$ *of the discrete nonlinear problem eq.* (50) *within the local ball* $\mathscr{B}_{r(h)}(I_h\mathbf{Q})$*. Furthermore, we have* $\|\mathbf{Q} - \mathbf{Q}_h\|_1 \lesssim h^{\min\{\deg,\Bbbk_Q-1\}}$.

We again employ an Aubin–Nitsche duality argument to derive L^2 error estimates. To this end, we consider the following linear dual problem to the primal nonlinear problem eq. (10): find $\mathbf{N} \in \mathbf{H}_0^1(\Omega)$ such that

$$\begin{cases} -K\Delta\mathbf{N} + 4l|\mathbf{Q}|^2\mathbf{N} + 8l(\mathbf{Q}:\mathbf{N})\mathbf{Q} - 2l\mathbf{N} = \mathbf{G} & \text{in } \Omega, \\ \mathbf{N} = \mathbf{0} & \text{on } \partial\Omega, \end{cases} \tag{58}$$

for a given $\mathbf{G} \in \mathbf{L}^2(\Omega)$ (we will make a particular choice for $\mathbf{G}$ in the proof of theorem 2.27). The weak form of eq. (58) is to find $\mathbf{N} \in \mathbf{H}_0^1(\Omega)$ such that

$$\langle \mathscr{D}\mathscr{N}^n(\mathbf{Q})\mathbf{N}, \Phi\rangle = A^n(\mathbf{N}, \Phi) + 3B^n(\mathbf{Q}, \mathbf{Q}, \mathbf{N}, \Phi) + C^n(\mathbf{N}, \Phi) = (\mathbf{G}, \Phi)_0 \quad \forall \Phi \in \mathbf{H}_0^1(\Omega). \tag{59}$$

To derive the L^2 a priori error estimates, we need two more auxiliary results.

Lemma 2.25. *For* $\mathbf{Q} \in \mathbf{H}^{\Bbbk_Q}(\Omega) \cap \mathbf{H}_b^1(\Omega)$*,* $\Bbbk_Q \geq 2$*,* $\mathbf{N} \in \mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega)$ *and* $I_h\mathbf{Q} \in \mathbf{V}_h \subset \mathbf{H}_b^1(\Omega)$*, it holds that*

$$A^n(I_h\mathbf{Q} - \mathbf{Q}, \mathbf{N}) \lesssim h^{\min\{\deg+1,\Bbbk_Q\}}\|\mathbf{Q}\|_{\Bbbk_Q}\|\mathbf{N}\|_2.$$

Lemma 2.26. *(Boundedness of the dual solution in the* H^2*-norm) The solution* $\mathbf{N}$ *to the weak form eq.* (59) *of the dual linear problem belongs to* $\mathbf{H}^2(\Omega) \cap \mathbf{H}_0^1(\Omega)$ *and it holds that*

$$\|\mathbf{N}\|_2 \lesssim \|\mathbf{G}\|_0. \tag{60}$$

Finally, we are ready to deduce an optimal L^2 error estimate.

Theorem 2.27. *(*L^2 *error estimate) Let* $\mathbf{Q}$ *be a regular solution of the nonlinear weak problem eq.* (50) *and* $\mathbf{Q}_h$ *be the approximate solution to the discrete problem (having the same weak formulation as eq.* (50)*). Then*

$$\|\mathbf{Q} - \mathbf{Q}_h\|_0 \lesssim h^{\min\{\deg+1,\Bbbk_Q\}}\left(2 + \left(3 + h + h^2 + h^{\min\{\deg,\Bbbk_Q-1\}} + h^{\min\{\deg+1,\Bbbk_Q\}}\right)\|\mathbf{Q}\|_{\Bbbk_Q}^2\right)\|\mathbf{Q}\|_{\Bbbk_Q}. \tag{61}$$

We will verify these results in the next section.

3. NUMERICAL EXPERIMENTS

The proceeding section presents some a priori error estimates for both $\mathbf{Q}$ and u in the decoupled case $q=0$. We now test the convergence rate of the finite element approximations by the method of manufactured solutions (MMS) and experimentally investigate the coupled case $q\neq 0$ in two dimensions. To this end, we choose a nontrivial solution for each state variable and add an appropriate source term to the equilibrium equations, thus modifying the energy accordingly. We can then compute the numerical convergence order.

3.1. Test 1: on the unit square

In this test, the numerical runs are performed on the unit square $\Omega=(0,1)^2$ and we take the following exact expressions for each state variable,

$$\begin{aligned}
Q_{11}^e &= \left(\cos\left(\frac{\pi(2y-1)(2x-1)}{8}\right)\right)^2 - \frac{1}{2},\\
Q_{12}^e &= \cos\left(\frac{\pi(2y-1)(2x-1)}{8}\right)\sin\left(\frac{\pi(2y-1)(2x-1)}{8}\right),\\
u^e &= 10\left((x-1)x(y-1)y\right)^3.
\end{aligned} \tag{62}$$

Then, in conducting the MMS, we are to solve the following governing equations

$$\begin{cases}
4Bq^4u^2Q_{11}+2Bq^2u\left(\partial_x^2 u-\partial_y^2 u\right)-2K\Delta Q_{11}-4lQ_{11}+16lQ_{11}\left(Q_{11}^2+Q_{12}^2\right)=\mathfrak{s}_1,\\
4Bq^4u^2Q_{12}+4Bq^2u\left(\partial_x\partial_y u\right)-2K\Delta Q_{12}-4lQ_{12}+16lQ_{12}\left(Q_{11}^2+Q_{12}^2\right)=\mathfrak{s}_2,\\
a_1u+a_2u^2+a_3u^3+2B\nabla\cdot(\nabla\cdot(\mathscr{D}^2u))+Bq^4\left(4\left(Q_{11}^2+Q_{12}^2\right)+1\right)u+2Bq^2(t_1+t_2)=\mathfrak{s}_3,
\end{cases}$$

subject to Dirichlet boundary conditions for both u and $\mathbf{Q}$ and a natural boundary condition for u. Here, source terms $\mathfrak{s}_1$, $\mathfrak{s}_2$ and $\mathfrak{s}_3$ are derived by substituting eq. (62) to the left hand sides, and t_1 and t_2 are given by

$$\begin{aligned}
t_1 &:= (Q_{11}+1/2)\partial_x^2 u+(-Q_{11}+1/2)\partial_y^2 u+2Q_{12}\partial_x\partial_y u,\\
t_2 &:= \partial_x^2\left(u\left(Q_{11}+1/2\right)\right)+\partial_y^2(u(-Q_{11}+1/2))+2\partial_x\partial_y(uQ_{12}).
\end{aligned}$$

We partition the domain $\Omega=(0,1)^2$ into $N\times N$ squares with uniform mesh size $h=\frac{1}{N}$ ($N=6$, 12, 24, 48) and denote the numerical solutions by u_h, $Q_{11,h}$ and $Q_{12,h}$. The numerical errors of u and $\mathbf{Q}$ in the $\|\cdot\|_0$-, $\|\cdot\|_1$- and $|||\cdot|||_h$-norms are defined as

$$\begin{aligned}
\|\mathbf{e}_u\|_0 = \|u^e-u_h\|_0,\quad \|\mathbf{e}_u\|_1 = \|u^e-u_h\|_1,\quad |||\mathbf{e}_u|||_h = |||u^e-u_h|||_h,\\
\|\mathbf{e}_\mathbf{Q}\|_0 = \|(Q_{11}^e,Q_{12}^e)-(Q_{11,h},Q_{12,h})\|_0,\quad \|\mathbf{e}_\mathbf{Q}\|_1 = \|(Q_{11}^e,Q_{12}^e)-(Q_{11,h},Q_{12,h})\|_1.
\end{aligned}$$

The convergence order is then calculated from the formula $\log_2\left(\frac{\text{error}_{h/2}}{\text{error}_h}\right)$. Throughout this section, we use the parameter values $a_1=-10$, $a_2=0$, $a_3=10$, $B=10^{-5}$, $K=0.3$ and $l=30$, similar to the simulations of oily streaks in [31].

Remark 3.1. *Since this is purely a numerical verification exercise, the manufactured solution can be physically unrealistic. However, we must specify a reasonable initial guess for Newton's method, due to the nonlinearity of the problem. The initial guess throughout this section is taken to be* $\left(\frac{1}{2}(\text{exact solution})+10^{-9}\right)$.

3.1.1. *Convergence rate for* $q=0$

For $\mathbf{Q}$ we expect both optimal H^1 and L^2 rates, as illustrated in theorems 2.24 and 2.27. table 1 presents the numerical convergence rate for the finite elements $[\mathbb{Q}_1]^2$, $[\mathbb{Q}_2]^2$ and $[\mathbb{Q}_3]^2$. Optimal L^2 and H^1 rates are shown with all choices of finite elements, as predicted.

Regarding the density variation u, we first present the convergence behaviour of the consistent discrete formulation eq. (13) with penalty parameter $\varepsilon=1$, since we have proven the optimal error rate in the mesh-dependent norm $|||\cdot|||_h$. The

	$N=\frac{1}{h}$	$\|\mathbf{e_Q}\|_0$	rate	$\|\mathbf{e_Q}\|_1$	rate
$[\mathbb{Q}_1]^2$	6	8.12×10^{-4}	–	3.78×10^{-2}	–
	12	2.02×10^{-4}	2.01	1.88×10^{-2}	1.01
	24	5.05×10^{-5}	2.00	9.39×10^{-3}	1.00
	48	1.26×10^{-5}	2.00	4.69×10^{-3}	1.00
$[\mathbb{Q}_2]^2$	6	2.92×10^{-5}	–	1.11×10^{-3}	–
	12	3.90×10^{-6}	2.90	2.71×10^{-4}	2.04
	24	5.02×10^{-7}	2.96	6.72×10^{-5}	2.01
	48	6.36×10^{-8}	2.99	1.68×10^{-5}	2.00
$[\mathbb{Q}_3]^2$	6	3.02×10^{-7}	–	2.25×10^{-5}	–
	12	2.17×10^{-8}	3.80	2.72×10^{-6}	3.05
	24	1.45×10^{-9}	3.90	3.34×10^{-7}	3.03
	48	9.33×10^{-11}	3.96	4.13×10^{-8}	3.01

TABLE 1. Test 1: Convergence rates for $\mathbf{Q}$ with different degrees of polynomial approximation, in the decoupled case $q=0$.

errors and convergence orders are listed in table 2. Optimal rates are observed in the $|||\cdot|||_h$-norm. Furthermore, optimal orders of convergence in the L^2-norm are shown for approximating polynomials of degree greater than 2, while a sub-optimal rate in the L^2-norm is given for piecewise quadratic polynomials, exactly as expected. Sub-optimal convergence rates for quadratic polynomials were also illustrated in the numerical results of [29]. We also tested the convergence with the penalty parameter $\varepsilon = 5\times10^4$ and found that the discrete norms are very similar to table 2. We therefore avoid repeating the details here.

	$N=\frac{1}{h}$	$\|\mathbf{e}_u\|_0$	rate	$\|\mathbf{e}_u\|_1$	rate	$|||\mathbf{e}_u|||_h$	rate
$\mathbb{Q}_2$	6	1.17×10^{-5}	–	3.46×10^{-4}	–	1.36×10^{-2}	–
	12	2.60×10^{-6}	2.17	9.81×10^{-5}	1.82	7.25×10^{-3}	0.91
	24	6.37×10^{-7}	2.03	2.54×10^{-5}	1.95	3.54×10^{-3}	1.03
	48	1.82×10^{-7}	1.80	6.88×10^{-6}	1.88	1.76×10^{-3}	1.01
$\mathbb{Q}_3$	6	4.73×10^{-6}	–	1.32×10^{-4}	–	4.98×10^{-3}	–
	12	3.32×10^{-7}	3.83	1.41×10^{-5}	3.23	9.96×10^{-4}	2.32
	24	2.12×10^{-8}	3.97	1.63×10^{-6}	3.12	2.46×10^{-4}	2.02
	48	1.32×10^{-9}	4.00	1.99×10^{-7}	3.03	6.14×10^{-5}	2.00
$\mathbb{Q}_4$	6	2.01×10^{-7}	–	7.76×10^{-6}	–	3.94×10^{-4}	–
	12	5.40×10^{-9}	5.22	4.30×10^{-7}	4.17	4.88×10^{-5}	3.01
	24	1.68×10^{-10}	5.00	2.68×10^{-8}	4.00	6.11×10^{-6}	2.99
	48	5.27×10^{-12}	4.99	1.68×10^{-9}	3.99	7.64×10^{-7}	3.00

TABLE 2. Test 1: Convergence rates using the consistent discrete formulation eq. (13) with penalty parameter $\varepsilon = 1$ and different polynomial degrees, in the decoupled case $q=0$.

We next give the error rates for the inconsistent discrete formulation eq. (49). We illustrate the discrete norms and the computed convergence rates in table 3 with the penalty parameter $\varepsilon = 1$. It can be observed that only first order convergence is obtained in the H^2-like norm $|||\cdot|||_h$ even with different approximating polynomials. Moreover, we notice by comparing tables 2 and 3 that the convergence rate deteriorates slightly for polynomials of degree 3 (although not for degree 4). This, however, can be improved by choosing a larger penalty parameter, as shown in table 4 with $\varepsilon = 5\times10^4$, where optimal rates are shown for the discrete norms $|||\cdot|||_h$, $\|\cdot\|_1$ and $\|\cdot\|_0$ for all polynomial degrees (except only sub-optimal in $\|\cdot\|_0$ when a piecewise quadratic polynomial is used as the approximation). The inconsistent discrete formulation appears to be a reasonable choice when a sufficiently large penalty parameter is used.

	$N=\frac{1}{h}$	$\|\mathbf{e}_u\|_0$	rate	$\|\mathbf{e}_u\|_1$	rate	$|||\mathbf{e}_u|||_h$	rate
$\mathbb{Q}_2$	6	3.50×10^{-6}	–	1.06×10^{-4}	–	5.60×10^{-3}	–
	12	8.76×10^{-8}	5.32	5.41×10^{-6}	4.29	2.56×10^{-3}	1.13
	24	1.77×10^{-8}	2.31	7.47×10^{-7}	2.86	1.28×10^{-3}	0.99
	48	4.35×10^{-9}	2.02	1.24×10^{-7}	2.56	6.42×10^{-4}	1.00
$\mathbb{Q}_3$	6	6.47×10^{-6}	–	1.86×10^{-4}	–	7.59×10^{-3}	–
	12	3.40×10^{-7}	4.25	1.73×10^{-5}	3.43	2.74×10^{-3}	1.47
	24	1.98×10^{-8}	4.10	2.03×10^{-6}	3.09	1.31×10^{-3}	1.07
	48	3.73×10^{-9}	2.39	2.63×10^{-7}	2.95	6.45×10^{-4}	1.02
$\mathbb{Q}_4$	6	2.05×10^{-7}	–	7.85×10^{-6}	–	3.93×10^{-4}	–
	12	5.40×10^{-9}	5.24	4.31×10^{-7}	4.19	4.88×10^{-5}	3.01
	24	1.68×10^{-10}	5.00	2.68×10^{-8}	4.01	6.11×10^{-6}	3.00
	48	5.27×10^{-12}	5.00	1.67×10^{-9}	4.00	7.64×10^{-7}	3.00

TABLE 3. Test 1: Convergence rates using the inconsistent discrete formulation eq. (49) with penalty parameter $\varepsilon=1$ and different polynomial degrees, in the decoupled case $q=0$.

	$N=\frac{1}{h}$	$\|\mathbf{e}_u\|_0$	rate	$\|\mathbf{e}_u\|_1$	rate	$|||\mathbf{e}_u|||_h$	rate
$\mathbb{Q}_2$	6	1.17×10^{-5}	–	3.48×10^{-4}	–	1.36×10^{-2}	–
	12	2.62×10^{-6}	2.16	9.86×10^{-5}	1.82	7.26×10^{-3}	0.91
	24	6.38×10^{-7}	2.04	2.54×10^{-5}	1.96	3.54×10^{-3}	1.03
	48	1.82×10^{-7}	1.81	6.88×10^{-6}	1.88	1.76×10^{-3}	1.01
$\mathbb{Q}_3$	6	4.80×10^{-6}	–	1.35×10^{-4}	–	4.92×10^{-3}	–
	12	3.35×10^{-7}	3.84	1.43×10^{-5}	3.23	9.86×10^{-4}	2.32
	24	2.14×10^{-8}	3.97	1.63×10^{-6}	3.13	2.45×10^{-4}	2.01
	48	1.33×10^{-9}	4.01	1.99×10^{-7}	3.04	6.13×10^{-5}	2.00
$\mathbb{Q}_4$	6	2.05×10^{-7}	–	7.85×10^{-6}	–	3.93×10^{-4}	–
	12	5.40×10^{-9}	5.24	4.31×10^{-7}	4.19	4.88×10^{-5}	3.01
	24	1.68×10^{-10}	5.00	2.68×10^{-8}	4.01	6.11×10^{-6}	3.00
	48	5.27×10^{-12}	5.00	1.67×10^{-9}	4.00	7.64×10^{-7}	3.00

TABLE 4. Test 1: Convergence rates using the inconsistent discrete formulation eq. (49) with penalty parameter $\varepsilon=5\times10^4$ and different polynomial degrees, in the decoupled case $q=0$.

3.1.2. *Convergence rate for $q\neq0$*

We next investigate the numerical convergence behaviour in the coupled case, i.e., $q\neq0$, in this subsection. Its analysis remains future work, but since it is the coupled case that is solved in practice it is important to assure ourselves that the discretisation is sensible. For brevity, we fix the model parameter $q=30$.

We examine the inconsistent discretisation for u with penalty parameter $\varepsilon=5\times10^4$. In unreported preliminary experiments, we observed that the error in $\mathbf{Q}$ is governed by the lower of the degrees of the polynomials used for $\mathbf{Q}$ and u. We thus give the convergence rates for u and $\mathbf{Q}$ separately in tables 5 and 6, with the other degree fixed appropriately. It can be seen that **Q** retains optimal rates in both the H^1 and L^2 norms, and though there are some fluctuations of the order for u, it still possesses very similar convergence rates when compared with the decoupled case described in table 4.

	$N=\frac{1}{h}$	$\|\mathbf{e}_u\|_0$	rate	$\|\mathbf{e}_u\|_1$	rate	$\|\|\|\mathbf{e}_u\|\|\|_h$	rate
$\mathbb{Q}_2$	6	1.21×10^{-5}	–	3.59×10^{-4}	–	1.37×10^{-2}	–
	12	3.98×10^{-6}	1.61	1.42×10^{-4}	1.34	8.30×10^{-3}	0.72
	24	1.57×10^{-6}	1.35	4.99×10^{-5}	1.51	3.89×10^{-3}	1.09
	48	2.58×10^{-7}	2.60	9.06×10^{-6}	2.46	1.78×10^{-3}	1.13
$\mathbb{Q}_3$	6	7.36×10^{-6}	–	2.25×10^{-4}	–	9.10×10^{-3}	–
	12	4.13×10^{-7}	4.16	1.86×10^{-5}	3.60	1.11×10^{-3}	3.03
	24	4.23×10^{-8}	3.29	2.24×10^{-6}	3.05	2.53×10^{-4}	2.14
	48	3.01×10^{-9}	3.81	2.28×10^{-7}	3.29	6.15×10^{-5}	2.04

TABLE 5. Test 1: Convergence rates for u with $q=30$ and penalty parameter $\varepsilon=5\times10^4$ with the inconsistent discretisation eq. (49) for u, fixing the approximation for **Q** to be with the $[\mathbb{Q}_2]^2$ element.

	$N=\frac{1}{h}$	$\|\mathbf{e_Q}\|_0$	rate	$\|\mathbf{e_Q}\|_1$	rate
$[\mathbb{Q}_1]^2$	6	8.12×10^{-4}	–	3.78×10^{-2}	–
	12	2.02×10^{-4}	2.01	1.88×10^{-2}	1.01
	24	5.05×10^{-5}	2.00	9.39×10^{-3}	1.00
	48	1.26×10^{-5}	2.00	4.69×10^{-3}	1.00
$[\mathbb{Q}_2]^2$	6	2.92×10^{-5}	–	1.11×10^{-3}	–
	12	3.90×10^{-6}	2.90	2.71×10^{-4}	2.04
	24	5.02×10^{-7}	2.96	6.72×10^{-5}	2.01
	48	6.37×10^{-8}	2.98	1.68×10^{-5}	2.00
$[\mathbb{Q}_3]^2$	6	3.02×10^{-7}	–	2.25×10^{-5}	–
	12	2.17×10^{-8}	3.80	2.72×10^{-6}	3.05
	24	1.45×10^{-9}	3.90	3.34×10^{-7}	3.03
	48	9.32×10^{-11}	3.96	4.13×10^{-08}	3.01

TABLE 6. Test 1: Convergence rates for **Q** with $q=30$ and penalty parameter $\varepsilon=5\times10^4$ with the inconsistent discretisation eq. (49) for u, fixing the approximation for u to be with the $\mathbb{Q}_3$ element.

Remark 3.2. *We also tested the convergence with the consistent weak formulation for u under the same numerical settings as in tables 5 and 6. We found that in both cases they present very similar convergence behaviour and thus we omit the details here.*

3.2. Test 2: on the unit disc

For the second set of experiments, we provide numerical results for an exact solution u^e with only H^3-regularity, instead of C^∞ as in eq. (62). Our goal is to investigate whether the H^4-regularity assumption on u can be relaxed. To this end, we consider a triangular mesh of $\Omega=\{(x,y)\mid x^2+y^2<1\}$ and choose u^e to be

$$u^e=(x^2+y^2)^{3/2}, \tag{63}$$

and choose the same exact solution for Q_{11}^e and Q_{12}^e as in eq. (62). The exact solution given by eq. (63) is in $H^3(\Omega)$ but not in $H^4(\Omega)$, hence violating the regularity assumption of the analysis in section 2.1.

The resulting convergence rates are reported in tables 7 and 8. Table 7 shows that optimal H^1 and L^2 rates are achieved for $\mathbf{Q}$ with three different choices of finite elements $[\mathbb{P}_1]^2, [\mathbb{P}_2]^2, [\mathbb{P}_3]^2$. Table 8 shows the convergence behaviour with penalty parameter $\varepsilon = 1$ when using the inconsistent discrete formulation eq. (49). In contrast to table 3, only first order convergence is obtained for the discrete norm $|||\cdot|||_h$ and second-order convergence for $\|\cdot\|_0$ and $\|\cdot\|_1$, with both $\mathbb{P}_3$ and $\mathbb{P}_4$. Interestingly, table 8 indicates no convergence when using $\mathbb{P}_2$ elements. It appears that the assumption $u \in H^4(\Omega)$ is necessary for our analysis, and that a different analysis should be carried out when this assumption no longer holds.

	No. of triangles	$\|\mathbf{e_Q}\|_0$	rate	$\|\mathbf{e_Q}\|_1$	rate
$[\mathbb{P}_1]^2$	60	6.08×10^{-2}	–	1.09	–
	240	1.56×10^{-2}	1.96	5.80×10^{-1}	0.91
	960	3.92×10^{-3}	2.00	2.93×10^{-1}	0.99
	3840	9.83×10^{-4}	2.00	1.47×10^{-1}	1.00
	15360	2.47×10^{-4}	1.99	7.34×10^{-2}	1.00
$[\mathbb{P}_2]^2$	60	8.97×10^{-3}	–	2.11×10^{-1}	–
	240	1.51×10^{-3}	2.57	5.87×10^{-2}	1.84
	960	2.22×10^{-4}	2.77	1.52×10^{-2}	1.95
	3840	3.02×10^{-5}	2.88	3.85×10^{-3}	1.98
	15360	3.93×10^{-6}	2.94	9.67×10^{-4}	1.99
$[\mathbb{P}_3]^2$	60	1.08×10^{-3}	–	3.21×10^{-2}	–
	240	8.21×10^{-5}	3.72	4.58×10^{-3}	2.81
	960	5.52×10^{-6}	3.89	5.92×10^{-4}	2.95
	3840	3.54×10^{-7}	3.96	7.44×10^{-5}	2.99
	15360	2.23×10^{-8}	3.99	9.31×10^{-6}	3.00

TABLE 7. Test 2: Convergence rates for $\mathbf{Q}$ with different degrees of polynomial approximation, in the decoupled case $q = 0$.

REFERENCES

[1] J. H. ARGYRIS, I. FRIED, AND D. W. SCHARPF, *The TUBA family of plate elements for the matrix displacement method*, Aeronaut. J., 72 (1968), pp. 701–709.

[2] J. M. BALL, *Mathematics and liquid crystals*, Mol. Cryst. Liq. Cryst., 647 (2017), pp. 1–27.

[3] J. M. BALL AND S. J. BEDFORD, *Discontinuous order parameters in liquid crystal theories*, Mol. Cryst. Liq. Cryst., 612 (2015), pp. 1–23.

[4] S. J. BEDFORD, *Calculus of variations and its application to liquid crystals*, PhD thesis, University of Oxford, 2014.

[5] H. BLUM AND R. R. BONN, *On the boundary value problem of the biharmonic operator on domains with angular corners*, Math. Mech. in the Appli. Sci., 2 (1980), pp. 556–581.

[6] J. P. BORTHAGARAY, R. H. NOCHETTO, AND S. W. WALKER, *A structure-preserving FEM for the uniaxially constrained* ***Q****-tensor model of nematic liquid crystals*, Numer. Math., 145 (2020), pp. 837–881.

[7] S. C. BRENNER, *C^0 interior penalty methods*, in Frontiers in Numerical Analysis - Durham 2010. Lecture Notes in Computational Science and Engineering, J. Blowey and M. Jensen, eds., vol. 85, Springer, Berlin, Heidelberg, 2011.

[8] S. C. BRENNER, T. GUDI, AND L. SUNG, *A weakly over-penalized symmetric interior penalty method for the biharmonic problem*, Electon. Trans. Numer. Anal., 37 (2010), pp. 214–238.

[9] S. C. BRENNER AND L. SUNG, *C^0 interior penalty methods for fourth order elliptic boundary value problems on polygonal domains*, J. Sci. Comput., 22 (2005), pp. 83–118.

[10] S. C. BRENNER AND L. SUNG, *A weakly over-penalized symmetric interior penalty method*, Electon. Trans. Numer. Anal., 30 (2008), pp. 107–127.

[11] S. C. BRENNER, K. WANG, AND J. ZHAO, *Poincaré-Friedrichs inequalities for piecewise H^2 functions*, Numer. Funct. Anal. Optim., 25 (2004), pp. 463–478.

[12] X. CHENG, W. HAN, AND H. HUANG, *Some mixed finite element methods for the biharmonic equation*, J. Comp. Appl. Math., 126 (2000), pp. 91–109.

	No. of triangles	$\|\mathbf{e}_u\|_0$	rate	$\|\mathbf{e}_u\|_1$	rate	$\|\|\|\mathbf{e}_u\|\|\|_h$	rate
$\mathbb{P}_2$	60	1.36×10^{-2}	–	2.77×10^{-1}	–	7.39	–
	240	1.58×10^{-3}	3.11	3.86×10^{-2}	2.84	2.39	1.63
	960	7.76×10^{-4}	1.02	1.57×10^{-2}	1.30	1.44	0.73
	3840	1.84×10^{-3}	-1.25	3.42×10^{-2}	-1.12	1.26	0.19
	15360	2.77×10^{-3}	-0.59	5.29×10^{-2}	-0.63	1.60	-0.34
$\mathbb{P}_3$	60	9.94×10^{-3}	–	2.21×10^{-1}	–	6.15	–
	240	3.99×10^{-3}	1.32	8.49×10^{-2}	1.38	3.03	1.02
	960	1.33×10^{-4}	4.90	4.57×10^{-3}	4.22	1.27	1.26
	3840	3.21×10^{-5}	2.06	8.47×10^{-4}	2.43	0.66	0.93
	15360	9.22×10^{-6}	1.80	2.11×10^{-4}	2.00	0.34	0.97
$\mathbb{P}_4$	60	7.17×10^{-3}	–	1.65×10^{-1}	–	4.70	–
	240	1.34×10^{-3}	2.42	3.84×10^{-2}	2.10	2.39	0.98
	960	1.18×10^{-4}	3.50	4.54×10^{-3}	3.08	1.28	0.90
	3840	2.98×10^{-5}	1.99	8.14×10^{-4}	2.48	0.67	0.94
	15360	8.15×10^{-6}	1.87	1.86×10^{-4}	2.13	0.34	0.97

TABLE 8. Test 2: Convergence rates using the inconsistent discrete formulation eq. (49) with penalty parameter $\varepsilon = 1$ and different polynomial degrees, in the decoupled case $q = 0$.

[13] T. A. DAVIS, *Finite element analysis of the Landau–de Gennes minimization problem for liquid crystals in confinement*, PhD thesis, Kent State University, 1994.

[14] T. A. DAVIS AND J. E. C. GARTLAND, *Finite element analysis of the Landau-de Gennes minimization problem for liquid crystals*, SIAM J. Numer. Anal., 35 (1998), pp. 336–362.

[15] P. G. DE GENNES, *An analogy between superconductors and smectic A*, Solid State Commun., 10 (1972), pp. 753–756.

[16] P. G. DE GENNES, *The Physics of Liquid Crystals*, Oxford University Press, Oxford, 1974.

[17] G. ENGEL, K. GARIKIPATI, T. J. R. HUGHES, M. G. LARSON, L. MAZZEI, AND R. L. TAYLOR, *Continuous/discontinuous finite element approximations of fourth-order elliptic problems in structural and continuum mechanics with applications to thin beams and plates, and strain gradient elasticity*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 3669–3750.

[18] L. C. EVANS, *Partial Differential Equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2nd ed., 2010.

[19] C. J. GARCÍA-CERVERA AND S. JOO, *Layer undulations in Smectic A liquid crystals*, J. Comp. Theo. Nano., 7 (2010), pp. 795–801.

[20] S. KESAVAN, *Topics in Functional Analysis and Applications*, John Wiley & Sons, New York, 1989.

[21] R. R. MAITY, A. MAJUMDAR, AND N. NATARAJ, *Discontinuous Galerkin finite element methods for the Landau–de Gennes minimization problem of liquid crystals*, IMA J. Numer. Anal., 41 (2021), pp. 1130–1163.

[22] P. MONDERKAMP, R. WITTMANN, L. B. G. CORTES, D. G. A. L. AARTS, AND H. LÖWEN, *Topology of orientational defects in confined smectic liquid crystals*, Phys. Rev. Lett., 127 (2021), pp. 1–10.

[23] N. J. MOTTRAM AND C. J. P. NEWTON, *Introduction to* **Q***-tensor theory*, 2014, `https://arxiv.org/abs/1409.3542`.

[24] M. Y. PEVNYI, J. SELINGER, AND T. J. SLUCKIN, *Modeling smectic layers in confined geometries: order parameter and defects*, Phys. Rev. E, 90 (2014), pp. 1–8.

[25] M. ROBINSON, C. LUO, P. E. FARRELL, R. ERBAN, AND A. MAJUMDAR, *From molecular to continuum modelling of bistable liquid crystal devices*, Liq. Cryst., 44 (2017), pp. 2267–2284.

[26] H. SACKMANN, *Plenary lecture. Smectic liquid crystals. A historical review*, Liq. Cryst., 5 (1989), pp. 43–55.

[27] R. SCHOLTZ, *A mixed method for fourth-order problems using linear finite elements*, RAIRO Numer. Anal., 15 (1978), pp. 85–90.

[28] I. W. STEWART, *The Static and Dynamic Continuum Theory of Liquid Crystals: A Mathematical Introduction*, CPC Press, 2004.

[29] E. SÜLI AND I. MOZOLEVSKI, *hp-version interior penalty DGFEMs for the biharmonic equation*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 1851–1863.

[30] R. WITTMANN, L. B. G. CORTES, H. LÖWEN, AND D. G. A. L. AARTS, *Particle-resolved topological defects of smectic colloidal liquid crystals in extreme confinement*, Nat. Comm., 12 (2021), pp. 1–10.

[31] J. XIA, S. MACLACHLAN, T. J. ATHERTON, AND P. E. FARRELL, *Structural landscapes in geometrically frustrated smectics*, Phys. Rev. Lett., 126 (2021), pp. 1–6.

[32] S. ZHANG, *A family of 3D continuously differentiable finite elements on tetrahedral grids*, Appl. Numer. Math., 59 (2009), pp. 219–233.