A FINITE ELEMENT METHOD FOR NEMATIC LIQUID CRYSTALS WITH VARIABLE DEGREE OF ORIENTATION

RICARDO H. NOCHETTO1, SHAWN W. WALKER2, WUJUN ZHANG3

Abstract. We consider the simplest one-constant model, put forward by J. Ericksen, for nematic liquid crystals with variable degree of orientation. The equilibrium state is described by a director field \( n \) and its degree of orientation \( s \), where the pair \((s, n)\) minimizes a sum of Frank-like energies and a double well potential. In particular, the Euler-Lagrange equations for the minimizer contain a degenerate elliptic equation for \( n \), which allows for line and plane defects to have finite energy.

We present a structure preserving discretization of the liquid crystal energy with piecewise linear finite elements that can handle the degenerate elliptic part without regularization, and show that it is consistent and stable. We prove \( \Gamma \)-convergence of discrete global minimizers to continuous ones as the mesh size goes to zero. We develop a quasi-gradient flow scheme for computing discrete equilibrium solutions and prove it has a strictly monotone energy decreasing property. We present simulations in two and three dimensions to illustrate the method’s ability to handle non-trivial defects.

Key words. liquid crystals, finite element method, gamma-convergence, gradient flow, line defect, plane defect

AMS subject classifications. 65N30, 49M25, 35J70

1. Introduction. Complex fluids are ubiquitous in nature and industrial processes and are critical for modern engineering systems [30, 39, 14]. An important difficulty in modeling and simulating complex fluids is their inherent microstructure. Manipulating the microstructure via external forces can enable control of the mechanical, chemical, optical, or thermal properties of the material. Liquid crystals [45, 24, 20, 3, 2, 12, 6, 31, 32, 4, 44] are a relatively simple example of a material with microstructure that may be immersed in a fluid with a free interface [51, 50].

Several numerical methods for liquid crystals have been proposed in [9, 28, 22, 33, 1] for harmonic mappings and liquid crystals with fixed degree of orientation, i.e. a unit vector field \( n(x) \) (called the director field) is used to represent the orientation of liquid crystal molecules. See [27, 34, 47] for methods that couple liquid crystals to Stokes flow. We also refer to the survey paper [5] for more numerical methods.

In this paper, we consider the one-constant model for liquid crystals with variable degree of orientation [25, 24, 45]. The state of the liquid crystal is described by a director field \( n(x) \) and a scalar function \( s(x) \), \(-1/2 < s < 1\), that represents the degree of alignment that molecules have with respect to \( n \). The equilibrium state is given by \((s, n)\) which minimizes the so-called one-constant Ericksen’s energy (2.1).

Despite the simple form of the one-constant Ericksen’s model, its minimizer may have non-trivial defects. If \( s \) is a non-vanishing constant, then the energy reduces to the Oseen-Frank energy whose minimizers are harmonic maps that may exhibit point defects (depending on boundary conditions) [13, 15, 19, 32, 31, 40]. If \( s \) is part of the minimization of (2.1), then \( s \) may vanish to allow for line (and plane) defects in dimension \( d = 3 \) [4, 44], and the resulting Euler-Lagrange equation for \( n \) is degenerate. However, in [32], it was shown that both \( s \) and \( u = sn \) have strong limits, which enabled the study of regularity properties of minimizers and the size of defects. This inspired the study of dynamics [20] and corresponding numerics [7], which are most relevant to our paper. However, in both cases they regularize the model to avoid the degeneracy introduced by the \( s \) parameter.

1rhn@math.umd.edu, 2walker@math.lsu.edu, 3wujun@umd.edu.
We design a finite element method (FEM) without any regularization. We prove stability and convergence properties and explore equilibrium configurations of liquid crystals via quasi-gradient flows. Our method builds on [11, 8, 10] and consists of a structure preserving discretization of (2.1). Given a weakly acute mesh $\mathcal{T}_h$ with meshsize $h$ (see Section 2.2), we use the subscript $h$ to denote continuous piecewise linear functions defined over $\mathcal{T}_h$, e.g. $(s_h, n_h)$ is a discrete approximation of $(s, n)$.

Our discretization of the energy is defined in (2.13) and requires that $\mathcal{T}_h$ be weakly acute. This discretization preserves the underlying structure and converges to the continuous energy in the sense of $\Gamma$-convergence [16] as $h$ goes to zero. Next, we develop a quasi-gradient flow scheme for computing discrete equilibrium solutions. We prove that this scheme has a strictly monotone energy decreasing property. Finally, we carry out numerical experiments and show that our finite element method, and gradient flow, allows for computing minimizers that exhibit line and plane defects.

The paper is organized as follows. In Section 2, we describe the Ericksen model for liquid crystals with variable degree of orientation, as well as the details of our discretization. Section 3 shows the $\Gamma$-convergence of our numerical method. A quasi-gradient flow scheme is given in Section 4, where we also prove a strictly monotone energy decreasing property. Section 5 presents simulations in two and three dimensions that exhibit non-trivial defects in order to illustrate the method’s capabilities.

2. Discretization of Ericksen’s model. We review the model [25] and relevant analysis results from the literature. We then develop our discretization strategy and show it is stable. In principle, the space dimension $d$ can be arbitrary $d \geq 2$, but for some of the proofs we require $d = 2, 3$.

2.1. Ericksen’s one constant model. Let the director field $n : \Omega \subset \mathbb{R}^d \rightarrow S^{d-1}$ be a vector-valued function with unit length, and the degree of orientation $s : \Omega \subset \mathbb{R}^d \rightarrow [-1/2, 1]$ be a real valued function. The case $s = 1$ represents the state of perfect alignment in which all molecules are parallel to $n$. Likewise, $s = -1/2$ represents the state of microscopic order in which all molecules are orthogonal to the orientation $n$. When $s = 0$, the molecules do not lie along any preferred direction which represents the state of an isotropic distribution of molecules.

The equilibrium state of the liquid crystals is described by the pair $(s, n)$ minimizing a bulk-energy functional which in the simplest one-constant model reduces to

\begin{equation}
E[s, n] := \int_\Omega \left( \kappa |\nabla s|^2 + s^2 |\nabla n|^2 \right) dx + \int_\Omega \psi(s) dx,
\end{equation}

with $\kappa > 0$ and double well potential $\psi$, which is a $C^2$ function defined on $-1/2 < s < 1$ that satisfies

1. $\lim_{s \rightarrow 1} \psi(s) = \lim_{s \rightarrow -1/2} \psi(s) = \infty$,
2. $\psi(0) > \psi(s^*) = \min_{s \in [-1/2, 1]} \psi(s)$ for some $s^* \in (0, 1)$,
3. $\psi'(0) = 0$;

see [25]. By introducing an auxiliary variable $u = sn$, we rewrite the energy as

\begin{equation}
E_1[s, n] = E_1[s, u] := \int_\Omega \left( (\kappa - 1) |\nabla s|^2 + |\nabla u|^2 \right) dx,
\end{equation}

which follows from the orthogonal splitting $\nabla u = n \otimes \nabla s + s \nabla n$ due to the constraint $|n| = 1$. Accordingly, we define the admissible class

\begin{equation}
\mathcal{A} := \{(s, n) : \Omega \rightarrow [-1/2, 1] \times S^{d-1}, \text{ where } s \in H^1(\Omega), \ u = sn \in H^1(\Omega)^d \}.
\end{equation}
Moreover, we may also enforce boundary conditions on \( s \) and \( \mathbf{n} \), possibly on different parts of the boundary. Let \( \Gamma_s (\Gamma_n) \) be an open subset of \( \partial \Omega \) where we will set Dirichlet boundary conditions for \( s \) (\( n \)). Then we have the following restricted admissible class

\[
A(g, r) := \{ (s, n) \in A : s|_{\Gamma_s} = g, \quad n|_{\Gamma_n} = r \},
\]

for some given smooth functions \( g \) and \( r \) such that \(|r| = 1\) and both \( g \) and \( r \) are the traces of some \( H^1(\Omega) \) functions.

Note that when the degree of orientation \( s \) equals a non-zero constant, the energy (2.1) effectively reduces to the Oseen-Frank energy \( \int_\Omega |\nabla s|^2 \, dx \). The introduction of the degree of orientation relaxes the energy of defects. In fact, with finite energy \( E[s, n] \), defects (i.e. discontinuities in \( n \)) may still occur in the singular set

\[
S := \{ x \in \Omega, \ s(x) = 0 \}.
\]

The existence of such a minimizer in the admissible class subject to Dirichlet boundary conditions is shown in [32, 2]. It is worth mentioning that the constant \( \kappa \) in \( E[s, n] \) (2.1) plays a significant role in the occurrence of the defects. Roughly speaking, if \( \kappa \) is large, then \( \int_\Omega \kappa |\nabla s|^2 \, dx \) dominates the energy and \( s \) is close to a constant. In this case, defects with finite energy are less likely to occur. But if \( \kappa \) is small, then \( \int_\Omega s^2 |\nabla n|^2 \, dx \) dominates the energy, and \( s \) may become zero. In this case, defects are more likely to occur. (This heuristic argument is later confirmed in the numerical experiments.) Since the investigation of defects is of primary interest in this paper, we consider the most significant case to be \( 0 < \kappa < 1 \).

We now describe our discretization \( E_h[s_h, n_h] \) of the energy (2.1) and its finite element minimizer \( (s_h, n_h) \).

### 2.2. Discretization of the energy.

Let \( T_h = \{ T \} \) be a conforming simplicial triangulation of the domain \( \Omega \). We denote by \( N_h \) the set of nodes (vertices) of \( T_h \) and the cardinality of \( N_h \) by \( N \) (with some abuse of notation). We demand that \( T_h \) be **weakly acute**, namely

\[
k_{ij} := -\int_\Omega \nabla \phi_i \cdot \nabla \phi_j \, dx \geq 0 \quad \text{for all } i \neq j,
\]

where \( \phi_i \) is the standard “hat” function associated with node \( x_i \in N_h \). We indicate with \( \omega_i = \text{supp} \phi_i \) the patch of a node \( x_i \) (i.e. the “star” of elements in \( T_h \) that contain the vertex \( x_i \)). Condition (2.6) imposes a severe geometric restriction on \( T_h \) [21, 43]. We recall the following characterization of (2.6) for \( d = 2 \).

**Proposition 2.1** (weak acuteness in two dimensions). For any pair of triangles \( T_1, T_2 \) in \( T_h \) that share a common edge \( e \), let \( \alpha_i \) be the angle in \( T_i \) opposite to \( e \) (for \( i = 1, 2 \)). If \( \alpha_1 + \alpha_2 \leq \pi \) for every edge \( e \), then (2.6) holds.

Generalizations of Proposition 2.1 to three dimensions, involving interior dihedral angles of tetrahedra, can be found in [29, 18].

We construct continuous piecewise affine spaces associated with the mesh, i.e.

\[
S_h := \{ s_h \in H^1(\Omega) : s_h|_T \text{ is affine for all } T \in T_h \},
\]

\[
U_h := \{ u_h \in H^1(\Omega)^d : u_h|_T \text{ is affine in each component for all } T \in T_h \},
\]

\[
N_h := \{ n_h \in U_h : |n_h(x_i)| = 1 \text{ for all nodes } x_i \in N_h \}.
\]

We also have the discrete spaces that include (Dirichlet) boundary conditions:

\[
S_h(\Gamma_s, g_h) := \{ s_h \in S_h : s_h|_{\Gamma_s} = g_h \},
\]

\[
N_h(\Gamma_n, r_h) := \{ n_h \in N_h : n_h|_{\Gamma_n} = r_h \},
\]

\[
A(g, r) := \{ (s, n) \in A : s|_{\Gamma_s} = g, \quad n|_{\Gamma_n} = r \},
\]

\[
S := \{ x \in \Omega, \ s(x) = 0 \}.
\]
where \((g_h, r_h)\) are the Lagrange interpolations of \((g, r)\) and \(g\) and \(r\) are the traces of some \(W_0^1(\Omega)\) functions.

In order to motivate our discrete version of \(E_1[s, n]\), note that \(\sum_{j=1}^N k_{ij} = 0\) for all \(x_i \in \mathcal{N}_h\). Therefore, for piecewise linear \(s_h = \sum_{i=1}^N s_h(x_i)\phi_i\), we have

\[
\int_{\Omega} |\nabla s_h|^2 dx = -\sum_{i=1}^N k_{ii} s_h(x_i)^2 - \sum_{i,j=1, i \neq j}^N k_{ij} s_h(x_i) s_h(x_j),
\]

whence, exploiting \(k_{ii} = -\sum_{j \neq i} k_{ij}\) and the symmetry \(k_{ij} = k_{ji}\), we get

\[
\int_{\Omega} |\nabla s_h|^2 dx = \sum_{i,j=1}^N k_{ij} s_h(x_i)(s_h(x_i) - s_h(x_j))
\]

\[
= \frac{1}{2} \sum_{i,j=1}^N k_{ij} (s_h(x_i) - s_h(x_j))^2 = \frac{1}{2} \sum_{i,j=1}^N k_{ij} (\delta_{ij} s_h)^2,
\]

where we define

\[
\delta_{ij} s_h := s_h(x_i) - s_h(x_j), \quad \delta_{ij} n_h := n_h(x_i) - n_h(x_j).
\]

With this in mind, we define the discrete energy using

\[
E^h_1[s_h, n_h] := \frac{\kappa}{2} \sum_{i,j=1}^N k_{ij} (\delta_{ij} s_h)^2 + \frac{1}{2} \sum_{i,j=1}^N k_{ij} \left( \frac{s_h(x_i)^2 + s_h(x_j)^2}{2} \right) |\delta_{ij} n_h|^2,
\]

and

\[
E^h_2[s_h] := \int_{\Omega} \psi(s_h(x)) dx.
\]

The second summation in (2.11) does not come from applying the standard discretization of \(\int_{\Omega} s^2 |\nabla n|^2 dx\) by piecewise linear elements. It turns out that this special form of the discrete energy allows us to handle the degenerate coefficient \(s^2\) without regularization. Eventually, we seek an approximation \((s_h, n_h) \in \mathcal{S}_h(\Gamma_s, g_h) \times \mathcal{N}_h(\Gamma_n, r_h)\) of the pair \((s, n)\) such that the discrete pair \((s_h, n_h)\) minimizes the discrete version of the bulk energy (2.1) given by

\[
E_h[s_h, n_h] := E^h_1[s_h, n_h] + E^h_2[s_h].
\]

The following result shows that definition (2.11) preserves the key structure (2.2) of [2, 32] at the discrete level, and turns out to be crucial for our analysis as well. We first introduce two discrete versions of the auxiliary vector field \(u\)

\[
u_h := I_h[s_h, n_h] \in \mathbb{U}_h, \quad \tilde{u}_h := I_h ||n_h| n_h| \in \mathbb{U}_h,
\]

where \(I_h\) denotes the piecewise linear Lagrange interpolation operator on mesh \(\mathcal{T}_h\).

**Lemma 2.2** (Energy Inequality). Let the mesh \(\mathcal{T}_h\) satisfy (2.6). If \((s_h, n_h) \in \mathcal{S}_h \times \mathcal{N}_h\), then, for any \(\kappa > 0\), the discrete energy (2.11) satisfies

\[
E^h_1[s_h, n_h] \geq (\kappa - 1) \int_{\Omega} |\nabla s_h|^2 dx + \int_{\Omega} |\nabla u_h|^2 dx =: \tilde{E}^h_1[s_h, u_h],
\]
as well as

\( E^h_t[\mathbf{s}_h, \mathbf{n}_h] \geq (\kappa - 1) \int_{\Omega} |\nabla I_h| s_h|^2 dx + \int_{\Omega} |\nabla \mathbf{u}_h|^2 dx =: E^h_t[|I_h| s_h, \mathbf{u}_h] \).

Proof. Since

\[
s_h(x_i) \mathbf{n}_h(x_i) - s_h(x_j) \mathbf{n}_h(x_j) = \frac{s_h(x_i) + s_h(x_j)}{2} \left( \mathbf{n}_h(x_i) - \mathbf{n}_h(x_j) \right) + \left( s_h(x_i) - s_h(x_j) \right) \frac{\mathbf{n}_h(x_i) + \mathbf{n}_h(x_j)}{2},
\]

using the orthogonality relation \( (\mathbf{n}_h(x_i) - \mathbf{n}_h(x_j)) \cdot (\mathbf{n}_h(x_i) + \mathbf{n}_h(x_j)) = |\mathbf{n}_h(x_i)|^2 - |\mathbf{n}_h(x_j)|^2 = 0 \) and (2.9) yields

\[
\int_{\Omega} |\nabla \mathbf{u}_h|^2 dx = \frac{1}{2} \sum_{i,j=1}^{N} k_{ij} |s_h(x_i) \mathbf{n}_h(x_i) - s_h(x_j) \mathbf{n}_h(x_j)|^2
\]

\[
= \frac{1}{2} \sum_{i,j=1}^{N} k_{ij} \left( \frac{s_h(x_i) + s_h(x_j)}{2} \right)^2 |\delta_{ij} \mathbf{n}_h|^2 + \frac{1}{2} \sum_{i,j=1}^{N} k_{ij} (\delta_{ij} s_h)^2 \left| \frac{\mathbf{n}_h(x_i) + \mathbf{n}_h(x_j)}{2} \right|^2.
\]

Employing again this orthogonality, this time in the form \( |\mathbf{n}_h(x_i) - \mathbf{n}_h(x_j)|^2 + |\mathbf{n}_h(x_i) + \mathbf{n}_h(x_j)|^2 = 4 \), we obtain

\[
\int_{\Omega} |\nabla \mathbf{u}_h|^2 dx = \frac{1}{2} \sum_{i,j=1}^{N} k_{ij} \left( \frac{s_h(x_i) + s_h(x_j)}{2} \right)^2 |\delta_{ij} \mathbf{n}_h|^2 + \frac{1}{2} \sum_{i,j=1}^{N} k_{ij} (\delta_{ij} s_h)^2 \left| \frac{\mathbf{n}_h(x_i) - \mathbf{n}_h(x_j)}{2} \right|^2.
\]

Since \( (s_h(x_i) + s_h(x_j))^2 = (s_h(x_i)^2 + s_h(x_j)^2) - (s_h(x_i) - s_h(x_j))^2 \), we infer that

\[
E^h_t[\mathbf{s}_h, \mathbf{n}_h] = \int_{\Omega} (\kappa - 1) |\nabla s_h|^2 + |\nabla \mathbf{u}_h|^2 dx + \sum_{i,j=1}^{N} k_{ij} (\delta_{ij} s_h)^2 \left| \frac{\mathbf{n}_h(x_i) - \mathbf{n}_h(x_j)}{2} \right|^2.
\]

The inequality (2.15) follows directly from \( k_{ij} \geq 0 \) for \( i \neq j \).

To prove (2.16), we note that the argument above still holds if we replace \( \mathbf{u}_h \) with \( \bar{\mathbf{u}}_h \) and \( s_h \) with \( |s_h| \) to get

\[
\int_{\Omega} |\nabla \bar{\mathbf{u}}_h|^2 dx \leq \frac{1}{2} \sum_{i,j=1}^{N} k_{ij} \frac{s_h(x_i)^2 + s_h(x_j)^2}{2} |\delta_{ij} \mathbf{n}_h|^2 + \frac{1}{2} \sum_{i,j=1}^{N} k_{ij} (\delta_{ij} |s_h|)^2.
\]

We finally find that

\[
\int_{\Omega} |\nabla \bar{\mathbf{u}}_h|^2 dx + (\kappa - 1) \int_{\Omega} |\nabla I_h| s_h|^2 dx = \int_{\Omega} |\nabla \mathbf{u}_h|^2 dx + \frac{\kappa - 1}{2} \sum_{i,j=1}^{N} k_{ij} (\delta_{ij} |s_h|)^2
\]

\[
\leq \frac{1}{2} \sum_{i,j=1}^{N} k_{ij} \frac{s_h(x_i)^2 + s_h(x_j)^2}{2} |\delta_{ij} \mathbf{n}_h|^2 + \frac{\kappa}{2} \sum_{i,j=1}^{N} k_{ij} (\delta_{ij} |s_h|)^2 \leq E^h_t[\mathbf{s}_h, \mathbf{n}_h],
\]
where we have used the triangle inequality \((|s_h(x_i)| - |s_h(x_j)|)^2 \leq (s_h(x_i) - s_h(x_j))^2\).

This completes the proof. □

**Remark 2.3** (purpose of (2.16)). The presence of the Lagrange interpolation operator \(I_h\) in (2.16) might seem strange, but accounts for the variational crime committed when enforcing \(\tilde{u}_h|_n = |s, u_h|\) only at the vertices. This is necessary to prove the \(\Gamma\)-convergence of our discrete energy (2.11) to the original continuous energy in (2.1).

In fact, we later exploit the weak lower semicontinuity in \(\Gamma\)-convergence of our discrete energy (2.11) to the original continuous energy in (2.1) (Lemma 3.4 below), which is a consequence of its convexity with respect to \(\nabla u_h\).

This is not obvious when \(\kappa < 1\), the most significant case for formation of defects.

### 3. \(\Gamma\)-convergence of the discrete energy

In this section, we show that our discrete energy (2.11) converges to the continuous energy (2.1) in the sense of \(\Gamma\)-convergence. Let the product space \(X := L^2(\Omega) \times L^2(\Omega)^d\) be equipped with the \(L^2\)-norm and let \(X_h := S_h \times N_h\). We define \(E[s, n]\) as in (2.1) for \((s, n) \in A\) and \(E[s, n] = \infty\) for \((s, n) \in X \setminus A\). Likewise, we define \(E^b_1[s_h, n_h]\) as in (2.11) for \((s_h, n_h) \in X_h\) and \(E^b_1[s, n] = \infty\) for all \((s, n) \in X \setminus X_h\).

**Theorem 3.1** (\(\Gamma\)-convergence). Let \(\{T_h\}\) be a sequence of weakly acute meshes. Then, for every \((s, n) \in X\) the following two properties hold:

- **Lim-inf inequality:** for every sequence \(\{(s_h, n_h)\}\) converging strongly to \((s, n)\) in \(X\), we have

\[
E_1[s, n] \leq \liminf_{h \to 0} E^b_1[s_h, n_h];
\]

- **Lim-sup inequality:** there exists a sequence \(\{(s_h, n_h)\}\) such that \((s_h, n_h)\) converges strongly to \((s, n)\) in \(X\) and

\[
E_1[s, n] \geq \limsup_{h \to 0} E^b_1[s_h, n_h].
\]

The proof of this theorem is split into several lemmas. We start with the lim-sup inequality (or consistency). We first observe that if \(E_1(s, n) = \infty\), then the assertion (3.2) is valid for any sequence \((s_h, n_h)\). Consequently, we consider the nontrivial case \(E_1(s, n) = \tilde{E}_1[s, u] < \infty\) or equivalently \((s, n) \in A\). Since \(H^2\)-functions are dense in \(A\), given \(\epsilon > 0\) there exists \((s_\epsilon, u_\epsilon) \in H^2(\Omega) \times H^2(\Omega)^d\) such that

\[
\| (s, u) - (s_\epsilon, u_\epsilon) \|_{H^1(\Omega)} \leq \epsilon \quad \Rightarrow \quad |\tilde{E}[s, u] - \tilde{E}[s_\epsilon, u_\epsilon]| \leq C\epsilon.
\]

Therefore, we can assume that \((s, u) \in H^2(\Omega)^{1+d}\) for \(d = 2, 3\), and let \((s_h, u_h)\) be the Lagrange interpolants of \((s, u)\), which are well defined because \(H^2(\Omega) \subset C^0(\Omega)\). Since \((s_h, u_h) \to (s, u)\) in \(H^1(\Omega)\), in view of the energy identity (2.18), we must show

\[
\sum_{i,j=1}^N k_{ij} (\delta_{ij} s_h)^2 |\delta_{ij} n_h|^2 \to 0, \quad \text{as } h \to 0.
\]

Heuristically, if \(n(x)\) is smooth, then the sum (3.3) is of order \(h^2 \int_{\Omega} |
abla s_h|^2 dx\) which obviously converges to zero. However, such an argument fails if the director field \(n(x)\) lacks high regularity, which is the case with defects. These are discontinuities of \(n(x)\) which occur in the singular set \(S\) defined in (2.5). Since \(n(x)\) is not regular in general, the proof of consistency requires a separate treatment of the region where \(n(x)\) is regular and the region where \(n(x)\) is singular. The heuristic argument can
be carried out in the regular region, while in the singular region we appeal to basic measure theory. With this motivation in mind, we now prove the following lemma.

**Lemma 3.2 (lim-sup inequality).** Let the pair \((s, n)\) belong to the admissible class \(A\) and \((s, u) \in H^2(\Omega)^{1+d}\). If \((s_h, u_h)\) are the Lagrange interpolants of \((s, u)\), then \(E_h^0[s_h, u_h] \to E_{\Omega}[s, u]\) as \(h \to 0\).

**Proof.** Given \(\epsilon > 0\), we divide the domain \(\Omega\) into two regions, \(S_\epsilon = \{x \in \Omega, |s(x)| < \epsilon\}\) and \(K_\epsilon = \overline{\Omega} \setminus S_\epsilon\), and the sum in (3.3) into two parts

\[
\mathcal{I}_h(K_\epsilon) := \sum_{i,j \in K_\epsilon} k_{ij} \left( \delta_{ij}s_h \right)^2 |\delta_{ij}u_h|^2, \quad \mathcal{I}_h(S_\epsilon) := \sum_{x_i, x_j \in S_\epsilon} k_{ij} \left( \delta_{ij}s_h \right)^2 |\delta_{ij}u_h|^2,
\]

where \(u_h(x_i) = n(x_i)\) is well defined provided \(s(x_i) \neq 0\) and otherwise \(u_h(x_i)\) is an arbitrary vector of unit length; thus \(n_h \in \mathbb{N}_h\).

**Step 1:** Estimate on \(K_\epsilon\). Note that \(K_\epsilon\) is a compact set. Since \(n = s^{-1}u\) is continuous everywhere except on the singular set \(S\), the field \(n\) is uniformly continuous on \(K_\epsilon\). Thus, we have \(|u_h(x_i) - u_h(x_j)| \to 0\) uniformly as the meshsize \(h \to 0\) because \(x_i\) and \(x_j\) are connected by a single edge of the mesh. Therefore, as \(h \to 0\)

\[
\mathcal{I}_h(K_\epsilon) \leq \left( \max_{i,j \in K_\epsilon} |u_h(x_i) - u_h(x_j)| \right)^2 \sum_{i,j = 1}^N k_{ij} \left( \delta_{ij}s_h \right)^2 = o(1) \int_{\Omega} |\nabla s_h|^2 dx \to 0.
\]

**Step 2:** Estimate on \(S_\epsilon\). If either \(x_i\) or \(x_j\) is in \(S_\epsilon\), without loss of generality, we assume that \(x_i \in S_\epsilon\). Since \(s(x)\) is uniformly continuous, and \(s_h(x)\) is the Lagrange interpolant of \(s(x)\), there is a meshsize \(h\) such that for any \(x\) in the star \(\omega_i\) of \(x_i\), \(|s_h(x) - s_h(x_i)| \leq \epsilon\), which implies that \(\omega_i \subset S_{2\epsilon}\). Thus, by the triangle inequality,

\[
\mathcal{I}_h(S_\epsilon) \leq 4 \sum_{x_i \text{ or } x_j \in S_\epsilon} k_{ij} \left( \delta_{ij}s_h \right)^2 \leq 8 \int_{\bigcup \omega_i} |\nabla s_h|^2 dx \leq 8 \int_{S_{2\epsilon}} |\nabla s_h|^2 dx,
\]

where the union \(\bigcup \omega_i\) is taken over all nodes \(x_i\) in \(S_\epsilon\). Since \(s \in H^2(\Omega)\), we have

\[
\int_{S_{2\epsilon}} |\nabla s_h|^2 dx \to \int_{S_{2\epsilon}} |\nabla s|^2 dx \quad \text{as } h \to 0.
\]

**Step 3:** The limit \(\epsilon \to 0\). Combining Steps 1 and 2 gives

\[
limit_{h \to 0} \sum_{i,j = 1}^N k_{ij} \left( \delta_{ij}s_h \right)^2 |\delta_{ij}u_h|^2 \leq 8 \int_{S_{2\epsilon}} |\nabla s|^2 dx,
\]

for all \(\epsilon > 0\). We finally show that

\[
\int_{S_{2\epsilon}} |\nabla s|^2 dx = \int_{\Omega} |\nabla s|^2 \chi_{\{|s| \leq 2\epsilon\}} dx \to 0 \quad \text{as } \epsilon \to 0,
\]

where \(\chi_A\) is the characteristic function of the set \(A\). By virtue of the Lebesgue’s dominated convergence theorem, we obtain

\[
\lim_{\epsilon \to 0} \int_{\Omega} |\nabla s|^2 \chi_{\{|s| \leq 2\epsilon\}} dx = \int_{\Omega} |\nabla s|^2 \chi_{\{|s| = 0\}} dx = 0,
\]

where the last equality follows by basic measure theory, i.e. \(\nabla s(x) = 0\) for a.e. \(x\) in \(\{s(x) = 0\}\) [26, Ch. 5, exercise 17, p. 292.]. This proves the lemma. ☐
To prove the lim-inf inequality (3.1), we need first to show coercivity (Lemma 3.3) and weak lower semi-continuity (Lemma 3.4). We do this next.

**Lemma 3.3 (coercivity).** For any \((s_h, n_h) \in S_h \times N_h\), we have

\[
E_h^1[s_h, n_h] \geq \min\{\kappa, 1\} \int_{\Omega} |\nabla u_h|^2 dx \geq \min\{\kappa, 1\} \int_{\Omega} |\nabla I_h s_h|^2 dx.
\]

**Proof.** Inequality (2.16) of Lemma 2.2 shows that

\[
E_h^1[s_h, n_h] \geq (\kappa - 1) \int_{\Omega} |\nabla I_h s_h|^2 dx + \int_{\Omega} |\nabla u_h|^2 dx.
\]

If \(\kappa \geq 1\), then \(E_h^1[s_h, n_h]\) obviously controls the \(H^1\)-norm of \(u_h\) with constant 1.

If \(0 < \kappa < 1\), then combining (2.11) with (2.19) yields

\[
E_h^1[s_h, n_h] \geq \frac{\kappa}{2} \sum_{i,j=1}^N k_{ij} (\delta_{ij}|s_h|)^2 + \frac{\kappa}{2} \sum_{i,j=1}^N k_{ij} \left( \frac{1}{2} \left( s_h(x_i)^2 + s_h(x_j)^2 \right) \right) |\delta_{ij} n_h|^2
\]

\[
\geq \kappa \int_{\Omega} |\nabla u_h|^2 dx,
\]

whence \(E_h^1[s_h, n_h] \geq \min\{\kappa, 1\} \int_{\Omega} |\nabla u_h|^2 dx\) as asserted. Since

\[
\sum_{i,j=1}^N k_{ij} |\bar{u}_h(x_i) - u_h(x_j)|^2 \geq \sum_{i,j=1}^N k_{ij} (|\bar{u}_h(x_i)| - |\bar{u}_h(x_j)|)^2,
\]

and \(|s_h(x_i)| = |\bar{u}_h(x_i)|\) for all nodes, we obtain

\[
\int_{\Omega} |\nabla u_h|^2 dx \geq \int_{\Omega} |\nabla I_h u_h|^2 dx = \int_{\Omega} |\nabla I_h s_h|^2 dx,
\]

which is the desired second estimate. □

Weak lower semi-continuity usually follows from convexity. While it is obvious that the discrete energy \(E_h^1[s_h, u_h]\) in (2.15) is convex with respect to \(\nabla u_h\) and \(\nabla s_h\) if \(\kappa \geq 1\), the convexity is not clear if \(0 < \kappa < 1\). It is worth mentioning that if \(\kappa < 1\), the convexity of the continuous energy (2.2) is based on the fact that \(|u| = |s|\) a.e. in \(\Omega\) and hence the convex part \(\int_{\Omega} |\nabla u|^2 dx\) controls the concave part \((\kappa - 1) \int_{\Omega} |\nabla s|^2 dx\) [32]. However, for the discrete energy (2.15), the equality \(|u_h| = |s_h|\) holds only at the vertices. Therefore, it is nontrivial to establish the weak lower semi-continuity of \(E_h^1[s_h, u_h]\). This is why we exploit the nodal relation \(|s_h| = |u_h|\) to derive the alternative formula (2.16) for \(E_h^1[I_h |s_h|, \tilde{u}_h]\). Our next lemma hinges on (2.16) and makes the convexity of \(E_h^1[I_h |s_h|, \tilde{u}_h]\) with respect to \(\nabla \tilde{u}_h\) completely explicit.

**Lemma 3.4 (weak lower semi-continuity).** The energy \(\int_{\Omega} L_h(w_h, \nabla w_h) dx\), with

\[
L_h(w_h, \nabla w_h) := (\kappa - 1) |\nabla I_h w_h|^2 + |\nabla w_h|^2,
\]

is well defined for any \(w_h \in U_h\) and is weakly lower semi-continuous in \(H^1(\Omega)\), i.e. for any weakly convergent sequence \(w_h \rightharpoonup w\) in the \(H^1\) norm, we have

\[
\liminf_{h \to 0} \int_{\Omega} L_h(w_h, \nabla w_h) dx \geq \int_{\Omega} (\kappa - 1) |\nabla w|^2 + |\nabla w|^2 dx.
\]
Proof. If $\kappa \geq 1$, then the assertion follows from standard arguments. Here, we only dwell upon $0 < \kappa < 1$ and dimension $d = 2$, because the case $d = 3$ is similar. After extracting a subsequence (not relabeled) we can assume that $w_h$ converges to $w$ strongly in $L^2(\Omega)$ and pointwise a.e. in $\Omega$.

Step 1: Equivalent form of $L_h$. We let $T$ be any triangle in the mesh $T_h$, label its three vertices as $x_0$, $x_1$, $x_2$, and define $e_1 := x_1 - x_0$ and $e_2 := x_2 - x_0$. After denoting $w_h^i = w_h(x_i)$ for $i = 0, 1, 2$, a simple calculation yields

$$
\nabla w_h = (w_h^1 - w_h^0) \otimes e_1^* + (w_h^2 - w_h^0) \otimes e_2^*,
$$
$$
\nabla I_h|_{w_h} = (|w_h^1| - |w_h^0|)e_1^* + (|w_h^2| - |w_h^0|)e_2^*,
$$

where $\{e_i^*\}$ is the dual basis of $\{e_i\}$, that is, $e_i^* \cdot e_j = I_{ij}$, and $I = (I_{ij})_{i,j=1}^2$ is the identity matrix. Assuming $|w_h^1| + |w_h^0| \neq 0$, we realize that

$$
|w_h^i| - |w_h^0| = \frac{w_h^i + w_h^0}{|w_h^1| + |w_h^0|} \cdot (w_h^i - w_h^0).
$$

We then obtain $\nabla I_h|_{w_h} = G_h(w_h) : \nabla w_h$ where $G_h(w_h)$ is the 3-tensor:

$$
G_h(w_h) := \frac{w_h^1 + w_h^0}{|w_h^1| + |w_h^0|} \otimes e_1 \otimes e_1^* + \frac{w_h^2 + w_h^0}{|w_h^2| + |w_h^0|} \otimes e_2 \otimes e_2^*,
$$

and the contraction between a 3-tensor and a 2-tensor in dyadic form is given by

$$
(g_1 \otimes g_2 \otimes g_3) : (m_1 \otimes m_2) := (g_1 \cdot m_1)(g_2 \cdot m_2)g_3.
$$

Therefore, we have

$$
L_h(w_h, \nabla w_h) = |\nabla w_h|^2 + (\kappa - 1)|G_h(w_h) : \nabla w_h|^2,
$$

which expresses $L_h(w_h, \nabla w_h)$ directly in terms of $\nabla w_h$ and nodal values of $w_h$.

Step 2: Convergence of $G_h(w_h)$. Given $\epsilon > 0$, Egoroff’s Theorem [48] asserts that

$$
w_h \rightarrow w \text{ uniformly on } E_\epsilon,
$$

for some subset $E_\epsilon$ and $|\Omega \setminus E_\epsilon| \leq \epsilon$. We now consider the set $A_\epsilon := \{|w(x)| \geq 2\epsilon\} \cap E_\epsilon$, and observe that there exists a sufficiently small $h_\epsilon$ such that for any $x \in A_\epsilon$

$$
|w_h(x)| \geq \epsilon \quad \text{for all } h \leq h_\epsilon.
$$

If $G(w) := \frac{w}{|w|} \otimes I$, then we claim that

$$
\int_{A_\epsilon} |G_h(w_h) - G(w)|^2 dx \rightarrow 0, \quad \text{as } h \rightarrow 0.
$$

For any $x \in A_\epsilon$, let $\{T_h\}$ be a sequence of triangles such that $x \in \overline{T_h}$. Since $|w_h(x)| \geq \epsilon$ and $w_h$ is piecewise linear, there exists a vertex of $T_h$, which we label as $x_h^0$, such that $|w_h^0| \geq \epsilon$. To compare $G_h(w_h)$ with $\frac{w_h(x)}{|w_h(x)|} \otimes I$, we use that $I = e_1 \otimes e_1^* + e_2 \otimes e_2^*$:

$$
G_h(w_h) - \frac{w_h(x)}{|w_h(x)|} \otimes I = \sum_{i=1,2} \left( \frac{w_h^i + w_h^0}{|w_h^1| + |w_h^0|} - \frac{w_h(x)}{|w_h(x)|} \right) \otimes e_i \otimes e_i^*.
$$
We define $H(x, y) := \frac{x + y}{|x| + |y|}$ and observe that for all $x \in A_{\varepsilon}$, we have

$$G_h(w_h) - \frac{w_h(x)}{|w_h(x)|} \otimes I = \sum_{i=1,2} \left( H(w^0_h, w^1_h) - H(w_h(x), w_h(x)) \right) \otimes e_i \otimes e_i^T.$$ 

Next, we estimate

$$|H(w^0_h, w^1_h) - H(w_h(x), w_h(x))| = \left| \frac{w_h(x)(w^0_h + w^1_h) - (|w^0_h| + |w^1_h|)w_h(x)}{(|w^0_h| + |w^1_h|)w_h(x)} \right| \leq \frac{|w^0_h + w^1_h - 2w_h(x)|}{|w^0_h| + |w^1_h|} + \left| \frac{|w^0_h(x)| + |w^1_h(x)|}{(|w^0_h| + |w^1_h|)w_h(x)} \right|.$$ 

Since $|w^0_h|, |w_h(x)| \geq \varepsilon$, and $w_h(x) - w^j_h(x_h) = \nabla w_h \cdot (x - x_h)$ for all $x \in T_h$, we have

$$|H(w^0_h, w^1_h) - H(w_h(x), w_h(x))| \leq C \frac{\varepsilon}{\varepsilon} |\nabla w_h| \quad \forall x \in A_{\varepsilon} \cap \overline{T_h}.$$ 

Integrating on $A_{\varepsilon}$, we obtain

$$\int_{A_{\varepsilon}} G_h(w_h) - \frac{w_h(x)}{|w_h(x)|} \otimes I \ dx \leq C \frac{\varepsilon^2}{\varepsilon^2} \int_{A_{\varepsilon}} |\nabla w_h(x)|^2 dx \to 0, \quad \text{as} \ h \to 0.$$ 

Since $w_h \to w$ a.e. in $\Omega$, and $\frac{w_h}{|w_h|} - \frac{w}{|w|}$ is bounded, applying the dominated convergence theorem, we infer that

$$\int_{A_{\varepsilon}} \left( \frac{w_h}{|w_h|} - \frac{w}{|w|} \right)^2 \ dx \to 0, \quad \text{as} \ h \to 0.$$ 

Combining these two limits, we deduce (3.5).

**Step 3: Convexity.** We now prove that the energy density

$$L(w, M) := |M|^2 + (\kappa - 1)|G(w) : M|^2$$

is convex with respect to any matrix $M$ for any vector $w$. Note that $L(w, M)$ is a quadratic function of $M$, so we only need to show that $L(w, M) \geq 0$ for any $M$ and $w$. Thus, it suffices to show that $|G(w) : M| \leq |M|$.

Assume that $M = \sum_{i,j} m_{ij} v_i \otimes v_j$ where $\{v_i\}_{i=1}^2$ is the canonical basis on $\mathbb{R}^2$. Then we have $|M|^2 = \sum_{i,j=1}^2 m^2_{ij}$ and a simple calculation yields

$$G(w) : M = \sum_{i} \frac{w_i}{|w|} v_i \otimes (v_1 \otimes v_1 + v_2 \otimes v_2) : \left( \sum_{k,l} m_{kl} v_k \otimes v_l \right) = \frac{1}{|w|^2} \sum_{i,k,l} w_i m_{kl} \delta_{ik} v_l = \frac{1}{|w|^2} \sum_{i,l} w_i m_{il} v_l,$$

where $w = \sum_{i=1}^2 w_i v_i$. Therefore, we obtain

$$|G(w) : M|^2 = \frac{1}{|w|^2} \sum_{j=1}^2 \left( \sum_{i=1}^2 w_i m_{ij} \right)^2.$$
The Cauchy-Schwarz inequality yields
\[
\left( \sum_{i=1}^{2} w_i m_{ij} \right)^2 \leq \left( \sum_{i=1}^{2} w_i^2 \right) \left( \sum_{i=1}^{2} m_{ij}^2 \right) = |w|^2 \left( \sum_{i=1}^{2} m_{ij}^2 \right),
\]
which implies \(|G(w) : M|^2 \leq |M|^2\) and \(L(w, M) \geq 0\) for any matrix \(M\) and vector \(w\).

A similar argument shows that \(L_h(w_h, M) \geq 0\) for any matrix \(M\) and vector \(w_h\).

**Step 4: Weak lower semi-continuity.** Since \(G_h(w_h) \to G(w)\) in \(L^2(A_e)\) according to (3.5), Egoroff’s theorem yields
\[
G_h(w_h) \to G(w) \quad \text{uniformly on } B_{\epsilon},
\]
where \(B_{\epsilon} \subset A_e\) and \(|A_e \setminus B_{\epsilon}| \leq \epsilon\). We claim that
\[
(3.6) \quad \liminf_{h \to 0} \int_{\Omega} L_h(w_h, \nabla w_h)dx \geq \int_{B_{\epsilon}} L(w, \nabla w)dx.
\]

Step 3 implies \(L_h(w_h, \nabla w_h) \geq 0\) for all \(x \in \Omega\). Hence,
\[
\int_{\Omega} L_h(w_h, \nabla w_h)dx \geq \int_{B_{\epsilon}} \left( |\nabla w_h|^2 + (\kappa - 1)|G_h(w_h) : \nabla w_h|^2 \right)dx.
\]

A simple calculation yields
\[
\int_{\Omega} L_h(w_h, \nabla w_h)dx \geq \int_{B_{\epsilon}} L(w, \nabla w_h)dx + (\kappa - 1)Q_h(w, w_h)
\]
where
\[
Q_h(w, w_h) := \int_{B_{\epsilon}} \left( |(G_h(w_h) - G(w)) : \nabla w_h|^2 |G_h(w_h) : \nabla w_h| + (G(w) : \nabla w_h)^2 |(G_h(w_h) - G(w)) : \nabla w_h| \right)dx.
\]

Since \(L(w, \nabla w_h)\) is convex with respect to \(\nabla w_h\) (Step 3), we have [26, pg. 446, Sec. 8.2.2]
\[
\liminf_{h \to 0} \int_{B_{\epsilon}} L(w, \nabla w_h)dx \geq \int_{B_{\epsilon}} L(w, \nabla w)dx.
\]

To prove (3.6), it remains to show that \(Q_h(w, w_h) \to 0\) as \(h \to 0\). Since \(G(w)\) and \(G_h(w_h)\) are bounded and \(\int_{\Omega} |\nabla w_h(x)|^2 dx\) is uniformly bounded, we have
\[
Q_h(w, w_h) \leq C \int_{B_{\epsilon}} |G_h(w_h) - G(w)||\nabla w_h|^2 dx
\]
\[
\leq C \max_{B_{\epsilon}} |G_h(w_h) - G(w)| \int_{B_{\epsilon}} |\nabla w_h|^2 dx \to 0 \quad \text{as } h \to 0,
\]
due to the uniform convergence of \(G_h(w_h)\) to \(G(w)\) in \(B_{\epsilon}\). Therefore, we infer that
\[
\liminf_{h \to 0} \int_{\Omega} L_h(w_h, \nabla w_h)dx \geq \int_{B_{\epsilon}} L(w, \nabla w)dx.
\]

Since the inequality above holds for arbitrarily small \(\epsilon\), taking \(\epsilon \to 0\) yields
\[
\liminf_{h \to 0} \int_{\Omega} L_h(w_h, \nabla w_h)dx \geq \int_{\Omega \setminus \{w(x) = 0\}} L(w, \nabla w)dx = \int_{\Omega} L(w, \nabla w)dx.
\]
where the last equality follows from $\nabla w = 0$ a.e. in the set $\{w(x) = 0\}$ [26, Ch. 5, exercise 17, p. 292.]. Finally, noting that $G(w) : \nabla w = \nabla |w|$, we get the assertion. \[\square\]

**Lemma 3.5** (characterizing limits). Let $\{T_h\}$ satisfy (2.6) and let $(s_h, n_h) \in \mathbb{X}_h$ converge strongly to $(s, n) \in \mathbb{X}$. Suppose that there exists a constant $C > 0$ such that

\begin{equation}
E_h^1(s_h, n_h) \leq C \quad \text{for all } h > 0,
\end{equation}

and let $u_h, \tilde{u}_h \in U_h$ be defined in (2.14). Then $(s, n) \in A$ and there is a subsequence of $\{(h|s_h|, \tilde{u}_h)\}$ that converges weakly in $H^1(\Omega)$, strongly in $L^2(\Omega)$, and pointwise a.e. to $(|s|, \tilde{u})$, where $\tilde{u} = |s|n$. In addition, there is also a subsequence of $\{(s_h, u_h)\}$ that converges weakly in $H^1(\Omega)$, strongly in $L^2(\Omega)$, and pointwise a.e. to $(s, u)$, where $u = sn$.

**Proof.** We define $\tilde{s}_h := I_h|s_h|$ and use Lemma 3.3 (coercivity) in conjunction with (3.7) to realize that the $H^1$-norm of $(\tilde{s}_h, \tilde{u}_h)$ is uniformly bounded in $h$. We can thus extract a subsequence (not relabeled) converging weakly in $H^1(\Omega)$, strongly in $L^2(\Omega)$, and pointwise a.e. to $(\tilde{s}, \tilde{u})$, where $\tilde{u} = |s|n$. We must relate this limit to $(s, n)$.

Since $s_h \to s$ in $L^2(\Omega)$, $s_h$ and $|s_h|$ converge pointwise a.e. to $s$ and $|s|$. Moreover,

$$
\|\tilde{s}_h - s_h\|_{L^2(\Omega)} = \|I_h|s_h| - |s_h|\|_{L^2(\Omega)} \leq C\|\nabla |s_h|\|_{L^2(\Omega)} \leq Ch\|\nabla s_h\|_{L^2(\Omega)},
$$

where $\|\nabla s_h\|_{L^2(\Omega)}$ is uniformly bounded because of (3.7) as well as (2.9) and (2.11). Consequently, $|s| = \tilde{s} = \tilde{s}_h \in H^1(\Omega)$ and so $s \in H^1(\Omega)$ as well.

Next, we show that $\tilde{u} = |s|n$. Since $\tilde{s}_h$ and $n_h$ are piecewise linear functions over $T_h$, we observe that

$$
\|\tilde{s}_h n_h - I_h[\tilde{s}_h n_h]\|_{L^1(\Omega)} \leq C h^2 \|\nabla \tilde{s}_h \otimes \nabla n_h\|_{L^1(\Omega)} \leq C h^2 \|\nabla \tilde{s}_h\|_{L^2(\Omega)} \|\nabla n_h\|_{L^2(\Omega)}.
$$

An inverse estimate shows that $\|\nabla n_h\|_{L^2(\Omega)} \leq C^{-1}|\Omega|^{1/2}$ because $|n_h| \leq 1$. Hence, $\|\tilde{s}_h n_h - I_h[\tilde{s}_h n_h]\|_{L^1(\Omega)} = O(h)$. Now write

$$
\tilde{u}_h - |s|n = (I_h[\tilde{s}_h n_h] - \tilde{s}_h n_h) + (\tilde{s}_h n_h - |s|n),
$$

and note that the first term goes to zero in $L^1(\Omega)$ with rate $h$ and the second one goes to zero in $L^1(\Omega)$ because $(\tilde{s}_h, n_h)$ converges to $(|s|, n)$ in $L^2(\Omega)$; ergo, $\tilde{u} = |s|n$.

Since $\|\nabla s_h\|_{L^2(\Omega)} \leq C$, (3.7) together with (2.15) gives $\|\nabla u_h\|_{L^2(\Omega)} \leq C$. The preceding argument thus shows that a subsequence of $u_h$ converges weakly in $H^1(\Omega)$, strongly in $L^2(\Omega)$ and pointwise a.e. to $u = sn \in H^1(\Omega)^d$.

We finally prove that $|n| = 1$ a.e. in $\Omega$, which implies $(s, n) \in A$. We first observe that $I_h[\tilde{u}_h] = I_h|s_h|$ along with

$$
\|\tilde{u}_h - I_h|s_h|\|_{L^2(\Omega)} \leq C h |\nabla \tilde{u}_h|_{L^2(\Omega)} \leq C h |\nabla \tilde{u}_h|_{L^2(\Omega)} \leq Ch.
$$

Since $I_h|s_h| \to |s|$ and $I_h|\tilde{u}_h| \to |\tilde{u}|$ as $h \to 0$, we deduce that $|s| = |\tilde{u}|$ a.e. in $\Omega$, or equivalently $|n| = 1$ a.e. in $\Omega$ as asserted. \[\square\]

We now prove the main theorem.

**Proof of Theorem 3.1.** The lim-sup inequality (3.2) follows directly from Lemma 3.2 provided $(s, n) \in A$; otherwise $E[s, n] = \infty$ and (3.2) is obvious.

As for the lim-inf inequality (3.1), let $(s, n) \in \mathbb{X}$ and take $\{(s_h, n_h) \in \mathbb{X}_h\}$ to be any sequence that converges to $(s, n)$ in the $L^2$-norm. If $\liminf_{h \to 0} E_h^1(s_h, n_h) = \infty$, then there is nothing to prove. So assume that $E_h^1(s_h, n_h)$ is uniformly bounded for a subsequence (not relabeled). By Lemma 3.5, $(s, n) \in A$ is an admissible pair and there
exists a subsequence \((I_h|s_h|, \mathbf{\tilde{u}}_h)\) (not relabeled) converging weakly in \(H^1\), strongly in \(L^2\) and pointwise a.e. to \((|s|, \mathbf{\bar{u}})\) with \(\mathbf{\bar{u}} = |s|\mathbf{n}\). Since \(I_h|s_h| = I_h|\mathbf{\tilde{u}}_h|\), invoking (2.16) in conjunction with Lemma 3.4 (weak lower semi-continuity), we obtain

\[
\liminf_{h \to 0} E_h^n[s_h, \mathbf{n}_h] \geq \int_\Omega (\kappa - 1)|\nabla \mathbf{\tilde{u}}|^2 + |\nabla \mathbf{\bar{u}}|^2 dx.
\]

Exploiting the properties \(\mathbf{\bar{u}} = |s|\mathbf{n}\) and \(|\mathbf{n}| = 1\) a.e. in \(\Omega\), we deduce the orthogonal decomposition \(\nabla \mathbf{\bar{u}} = \nabla |s| \otimes \mathbf{n} + |s|\nabla \mathbf{n}\) a.e. in \(\Omega\). Hence,

\[
\int_\Omega (\kappa - 1)|\nabla \mathbf{\tilde{u}}|^2 + |\nabla \mathbf{\bar{u}}|^2 dx = \int_\Omega \kappa|\nabla |s||^2 + |s|^2|\nabla \mathbf{n}|^2 dx
\]

\[
= \int_\Omega \kappa|\nabla |s||^2 + s^2|\nabla \mathbf{n}|^2 dx \equiv E_1[s, \mathbf{n}].
\]

This completes the proof. \(\square\)

The \(\Gamma\)-convergence result immediately yields the following corollary [17, 23].

**Corollary 3.6** (convergence of global discrete minimizers). Let \(\{T_h\}\) satisfy (2.6). If \(\{(s_n, \mathbf{n}_h)\} \subset \mathcal{X}_h\) is a sequence of global minimizers of the discrete energy \(E_h^n[s_h, \mathbf{n}_h]\) in (2.13), then there is a subsequence that converges weakly in \(H^1(\Omega)\), strongly in \(L^2(\Omega)\), and pointwise a.e. in \(\Omega\) to an admissible pair \((s, \mathbf{n})\) \(\in \mathcal{A}\), which is a global minimizer of the continuous energy \(E[s, \mathbf{n}]\) in (2.1). In addition,

\[
E_h^n[s_h, \mathbf{n}_h] \to E[s, \mathbf{n}] \quad \text{as } h \to 0.
\]

This corollary is about global minimizers, both discrete and continuous. In the next section, we design a quasi-gradient flow to compute discrete local minimizers, and show its convergence (see Theorem 4.2). In general, convergence to a global minimizer is not available, nor are rates of convergence due to the lack of continuous dependence results. However, if local minimizers of \(E[s, \mathbf{n}]\) are isolated, then there exists local minimizers of \(E_h^n[s_h, \mathbf{n}_h]\) that \(\Gamma\)-converge to \((s, \mathbf{n})\) [17, 23].

4. **Quasi-Gradient Flow.** We consider a gradient flow methodology consisting of a gradient flow in \(s\) and a minimization in \(\mathbf{n}\) as a way to compute minimizers of (2.1) and (2.13). We begin with its description for the continuous system and verify that it has a monotone energy decreasing property. We then do the same for the discrete system.

4.1. **Continuous case.** We introduce the following subspace to enforce Dirichlet boundary conditions on open subsets \(\Gamma\) of \(\partial \Omega\):

\[
H^1_\Gamma(\Omega) = \{v \in H^1(\Omega) : v = 0 \text{ on } \Gamma\}.
\]

4.1.1. **First order variation.** Consider the bulk energy \(E[s, \mathbf{n}]\) where the pair \((s, \mathbf{n})\) is in the admissible class \(\mathcal{A}\) defined in (2.3). We take a variation \(z \in H^1_\Gamma(\Omega)\) of \(s\) and obtain \(\delta_s E[s, \mathbf{n}; z] = \delta_s E_1[s, \mathbf{n}; z] + \delta_s E_2[s; z]\), the first variation of \(E\) in the direction \(z\), where

\[
\delta_s E_1[s, \mathbf{n}; z] = 2 \int_\Omega (\nabla s \cdot \nabla z + |\nabla \mathbf{n}|^2 s z) \quad \text{and} \quad \delta_s E_2[s; z] = \int_\Omega \psi'(s) z.
\]
Next, we introduce the space of tangential variations of $n$:

$$
V^\perp(n) = \{ v \in H^1(\Omega)^d : v \cdot n = 0 \text{ a.e. in } \Omega \}.
$$

In order to satisfy the constraint $|n| = 1$, we take a variation $v \in V^\perp(n)$ of $n$ and get

$$
\delta_n E[s, n; v] = \delta_n E_1[s, n; v] = 2 \int_\Omega s^2(\nabla n \cdot \nabla v).
$$

Note that variations in $V^\perp(n)$ preserve the unit length constraint up to second order accuracy [45]: $|n + tv|^2 = 1 + t^2|v|^2$ and $|n + tv| \geq 1$ for all $t \in \mathbb{R}$.

4.1.2. Quasi-gradient flow. We consider an $L^2$-gradient flow for $E$ with respect to the scalar variable $s(t)$:

$$
\int_\Omega \partial_t s z := -\delta_s E_1[s, n; z] - \delta_s E_2[s; z] \quad \text{for all } z \in H^1(\Omega);
$$

here, we enforce stationary Dirichlet boundary conditions for $s$ on the set $\Gamma_s \subset \partial \Omega$, whence $z = 0$ on $\Gamma_s$. A simple but formal integration by parts yields

$$
\int_\Omega \partial_t s z = -\int_\Omega (-2\Delta s + 2|\nabla n|^2 s + \psi'(s)) z \quad \text{for all } z \in H^1(\Omega),
$$

where we use the implicit Neumann condition $\nu \cdot \nabla s = 0$ on $\partial \Omega \setminus \Gamma_s$, $\nu$ being the outer unit normal on $\partial \Omega$. Therefore, $s$ satisfies the (nonlinear) parabolic PDE:

$$
\partial_t s = 2\Delta s - 2|\nabla n|^2 s - \psi'(s).
$$

Given $s$, let $n$ satisfy $|n| = 1$ a.e. in $\Omega$, a stationary Dirichlet boundary condition on the open set $\Gamma_n \subset \partial \Omega$, and the following degenerate minimization problem:

$$
E[s, n] \leq E[s, m] \quad \text{for all } |m| = 1 \text{ a.e. } \Omega,
$$

with the same boundary condition as $n$. This implies

$$
\delta_n E[s, n; v] = 0 \quad \text{for all } v \in V^\perp(n) \cap H^1(\Omega)^d.
$$

4.1.3. Formal energy decreasing property. Differentiating the energy $E[s, n]$ with respect to time, we obtain

$$
\partial_t E[s, n] = \delta_s E[s, n; \partial_t s] + \delta_n E[s, n; \partial_t n].
$$

By virtue of (4.3) and (4.4), we deduce that

$$
\partial_t E[s, n] = -\delta_s E[s, n; \partial_t s] = -\int_\Omega |\partial_t s|^2.
$$

Hence, the bulk energy $E$ is monotonically decreasing for our quasi-gradient flow.

4.2. Discrete case. Let $s_h^k \in S_h(\Gamma_s, g_h)$ and $n_h^k \in N_h(\Gamma_n, r_h)$ denote finite element functions, where $k$ indicates a “time-step” index (see Section 4.2.2 for the discrete gradient flow algorithm). To simplify notation, we use the following:

$$
s_i^k := s_h^k(x_i), \quad n_i^k := n_h^k(x_i), \quad z_i := z_h(x_i), \quad v_i := v_h(x_i).
$$
4.2.1. First order variation. First, we introduce the discrete version of (4.2):

$$(4.6) \quad V_{h}^{-1}(n_h) = \{v_h \in U_h : v_h(x_i) \cdot n_h(x_i) = 0 \text{ for all nodes } x_i \in \mathcal{N}_h \}.$$ 

Next, the first order variation of $E^h_1$ in the direction $v_h \in V^{-1}_h(n_h^k) \cap H^1_{\Gamma_h}(\Omega)$ at the director variable $n_h^k$ reads

$$(4.7) \quad \delta_{n_h} E^h_1(s^k_h, n_h^k; v_h) = \sum_{i,j=1}^{N} k_{ij} \left( \frac{(s^k_i)^2 + (s^k_j)^2}{2} \right) (\delta_{ij} n_h^k) \cdot (\delta_{ij} v_h),$$

whereas the first order variation of $E^h_1$ in the direction $z_h \in S_h \cap H^1_{L_2}(\Omega)$ at the degree of orientation variable $s^k_h$ consists of two terms

$$(4.8) \quad \delta_{s_h} E^h_1(s^k_h, n_h^k; z_h) = \kappa \sum_{i,j=1}^{N} k_{ij} \left( \delta_{ij} s^k_i \right) \left( \delta_{ij} z_h \right) + \sum_{i,j=1}^{N} k_{ij} |\delta_{ij} n_h^k|^2 \left( \frac{s^k_i z_i + s^k_j z_j}{2} \right).$$

To design an unconditionally stable scheme for the discrete gradient flow, we employ the convex splitting technique in [49, 41, 42]. We split the double well potential into a convex and concave part: let $\psi_c$ and $\psi_e$ be both convex for all $s \in (-1/2,1)$ so that $\psi(s) = \psi_c(s) - \psi_e(s)$, and set

$$(4.9) \quad \delta_{s_h} E^h_2(s^{k+1}_h; z_h) := \int_{\Omega} \left[ \psi'_c(s^{k+1}_h) - \psi'_e(s^{k}_h) \right] z_h dx.$$

**Lemma 4.1** (convex-concave splitting). For any $s^k_h$ and $s^{k+1}_h$ in $S_h$, we have

$$\int_{\Omega} \psi(s^{k+1}_h) dx - \int_{\Omega} \psi(s^k_h) dx \leq \delta_{s_h} E^h_2(s^{k+1}_h; s^{k+1}_h - s^k_h).$$

**Proof.** A simple calculation, based on the mean-value theorem and the convex splitting $\psi = \psi_c - \psi_e$, yields

$$\int_{\Omega} (\psi(s^{k+1}_h) - \psi(s^k_h)) dx = \delta_{s_h} E^h_2(s^{k+1}_h; s^{k+1}_h - s^k_h) + T,$$

where

$$T = \int_{\Omega} \int_{0}^{1} \left[ \psi'_c(s^k_h + \theta(s^{k+1}_h - s^k_h)) - \psi'_e(s^{k+1}_h) \right] (s^{k+1}_h - s^k_h) d\theta dx.$$

The convexity of both $\psi_c$ and $\psi_e$ implies $T \leq 0$, as desired. \( \square \)

4.2.2. Discrete quasi-gradient flow algorithm. Our scheme for minimizing the discrete energy $E_h|_{S_h, n_h}$ is an alternating direction method, which minimizes with respect to $n_h$ and evolves $s_h$ separately in the steepest descent direction during each iteration. Therefore, this algorithm is not a standard gradient flow but rather a quasi-gradient flow.

**Algorithm:** Given $(s^0_h, n^0_h)$ in $S_h(\Gamma_s, g) \times N_h(\Gamma_n, r)$, iterate Steps (a)-(c) for $k \geq 0.$
Step (a): Minimization. Find \( t^k_h \in V^+_h(n^k_h) \cap \Gamma^1_h) \) such that \( n^k_h + t^k_h \) minimizes the energy \( E^1_h[n^k_h, n^k_h + v_h] \) for all \( v_h \) in \( V^+_h(n^k_h) \cap \Gamma^1_h(\Omega) \), i.e. \( t^k_h \) satisfies
\[
\delta_n_j E^1_h[n^k_h, n^k_h + t^k_h, v_h] = 0, \quad \forall v_h \in V^+_h(n^k_h) \cap \Gamma^1_h(\Omega).
\]

Step (b): Projection. Normalize \( n^{k+1}_h := \frac{n^k_h + t^k_h}{|n^k_h + t^k_h|} \) at all nodes \( x_i \in N_h \).

Step (c): Gradient flow. Using \( (s^k_h, n^{k+1}_h) \), find \( s^{k+1}_h \) in \( S_h(\Gamma_s, g) \) such that
\[
\frac{1}{\Omega} \int \frac{s^{k+1}_h - s^k_h}{\delta t} z_h = -E^1_h[s^k_h, n^{k+1}_h, z_h; v_h] - \delta_n_j E^1_h[s^{k+1}_h, z_h], \quad \forall z_h \in S_h(\Gamma_s, g).
\]

We impose Dirichlet boundary conditions for both \( s^k_h \) and \( n^k_h \). Note that the scheme has no restriction on the time step thanks to the implicit Euler method in Step (c).

4.3. Energy decreasing property. The quasi-gradient flow scheme in Section 4.2.2 has a monotone energy decreasing property, a discrete version of (4.5), provided the mesh \( T_h \) is weakly acute, namely it satisfies (2.6) [21, 43].

**Theorem 4.2 (energy decrease).** Let \( T_h \) satisfy (2.6). The iterate \( (s^{k+1}_h, n^{k+1}_h) \) of the Algorithm (discrete quasi-gradient flow) of Section 4.2.2 exists and satisfies
\[
E^h[s^{k+1}_h, n^{k+1}_h] \leq E^h[s^k_h, n^k_h] - \frac{1}{\delta t} \int_{\Omega} (s^{k+1}_h - s^k_h)^2 dx.
\]
Equality holds if and only if \( (s^{k+1}_h, n^{k+1}_h) = (s^k_h, n^k_h) \) (equilibrium state).

**Proof.** The Steps (a) and (b) are monotone whereas Step (c) decreases the energy.

**Step (a): Minimization.** Since \( E^1_h \) is convex in \( n^k_h \) for fixed \( s^k_h \), there exists a tangential variation \( t^k_h \) which minimizes \( E^1_h[s^k_h, n^k_h + v^k_h] \) among all tangential variations \( v^k_h \). The fact that \( E^2_h \) is independent of the director field \( n^k_h \) implies
\[
E^h[s^k_h, n^k_h + t^k_h] \leq E^h[s^k_h, n^k_h].
\]

**Step (b): Projection.** Since the mesh \( T_h \) is weakly acute, we claim that
\[
n^{k+1}_h = \frac{n^k_h + t^k_h}{|n^k_h + t^k_h|} \quad \Rightarrow \quad E^1_h[s^k_h, n^{k+1}_h] \leq E^h[s^k_h, n^k_h + t^k_h].
\]
We follow [1, 8]. Let \( v_h = n^k_h + t^k_h \), \( w_h = \frac{v_h}{|v_h|} \), and observe that \( |v_h| \geq 1 \) and \( w_h \) is well-defined. By (2.11) (definition of discrete energy), we only need to show that
\[
k_{ij}\frac{(s^k_i)^2 + (s^k_j)^2}{2} |w_h(x_i) - w_h(x_j)|^2 \leq k_{ij}\frac{(s^{k+1}_i)^2 + (s^{k+1}_j)^2}{2} |v_h(x_i) - v_h(x_j)|^2.
\]
for all \( x_i, x_j \in N_h \). Because \( k_{ij} \geq 0 \) for \( i \neq j \), this is equivalent to showing that \( |w_h(x_i) - w_h(x_j)| \leq |v_h(x_i) - v_h(x_j)| \). This follows from the fact that the mapping \( a \rightarrow a/|a| \) defined on \( \{ a \in \mathbb{R}^d : |a| \geq 1 \} \) is Lipschitz continuous with constant 1. Note that equality above holds if and only if \( n^{k+1}_h = n^k_h \) or equivalently \( t^k_h = 0 \).

**Step (c): Gradient flow.** Since \( E^1_h \) is quadratic in terms of \( s^k_h \), and
\[
2s^{k+1}_h(s^{k+1}_h - s^k_h) = (s^{k+1}_h - s^k_h)^2 + |s^{k+1}_h|^2 - |s^k_h|^2,
\]
reordering terms gives
\[
E^1_h[s^{k+1}_h, n^{k+1}_h] - E^1_h[s^k_h, n^{k+1}_h] = R_1 - E^1_h[s^{k+1}_h - s^k_h, n^{k+1}_h] \leq R_1,
\]
where
\[ R_1 := \delta_{s_h} E_1^{h}[s_h^{k+1}, n_h^{k+1}, s_h^k - s_h^k]. \]

On the other hand, Lemma 4.1 implies
\[ E_2^{h}[s_h^{k+1}] - E_2^{h}[s_h^k] = \int_{\Omega} \psi(s_h^{k+1}) dx - \int_{\Omega} \psi(s_h^k) dx \leq R_2 := \delta_{s_h} E_2^{h}[s_h^{k+1}, s_h^{k+1} - s_h^k]. \]

Combining both estimates and invoking Step (c) of the Algorithm yields
\[ E_h^{h}[s_h^{k+1}, n_h^{k+1}] - E_h^{h}[s_h^k, n_h^{k+1}] \leq R_1 + R_2 = -\frac{1}{\delta t} \int_{\Omega} (s_h^{k+1} - s_h^k)^2 \leq 0, \]
which is the assertion. Note finally that equality occurs if and only if \( s_h^{k+1} = s_h^k \) and \( n_h^{k+1} = n_h^k \), which corresponds to an equilibrium state. This completes the proof. \( \Box \)

5. Numerical experiments. We present computational experiments to illustrate our method, which was implemented with the MATLAB/C++ finite element toolbox FELICITY [46]. For all 3-D simulations, we used the algebraic multi-grid solver (AGMG) [37, 35, 36, 38] to solve the linear systems in parts (a) and (c) of the quasi-gradient flow algorithm. In 2-D, we simply used the “backslash” command in MATLAB.

5.1. Tangential variations. Solving step (a) of the Algorithm requires a tangential basis for the test function and the solution. However, forming the matrix system is easily done by first ignoring the tangential variation constraint (i.e. arbitrary variations), followed by a simple modification of the matrix system.

Let \( \mathbf{A} \mathbf{t}_h^k = \mathbf{B} \) represent the linear system in Step (a) and suppose \( d = 3 \). Multiplying by a discrete test function \( \mathbf{v}_h \), we have
\[ \mathbf{v}_h^T \mathbf{A} \mathbf{t}_h^k = \mathbf{v}_h^T \mathbf{B}, \quad \text{for all } \mathbf{v}_h \in \mathbb{R}^{dN}. \]

Next, using \( \mathbf{n}_h^k \), find \( \mathbf{r}_1, \mathbf{r}_2 \) such that \( \{\mathbf{n}_h^k, \mathbf{r}_1, \mathbf{r}_2\} \) forms an orthonormal basis of \( \mathbb{R}^3 \) at each node \( x_i \), i.e. find an orthonormal basis of \( V_h^1(\mathbf{n}_h^k) \). Next, expand \( \mathbf{t}_h^k = \Phi_1 \mathbf{r}_1 + \Phi_2 \mathbf{r}_2 \) and make a similar expansion for \( \mathbf{v}_h \). After a simple rearrangement and partitioning of the linear system, one finds it decouples into two smaller systems: one for \( \Phi_1 \) and one for \( \Phi_2 \). After solving for \( \Phi_1, \Phi_2 \), define the nodal values of \( \mathbf{t}_h^k \) by the formula \( \mathbf{t}_h^k = \Phi_1 \mathbf{r}_1 + \Phi_2 \mathbf{r}_2 \).

5.2. Point defect in 2-D. For the classic Frank energy \( \int_{\Omega} |\nabla \mathbf{n}|^2 \), a point defect in two dimensions has infinite energy [45]. This is not the case for the energy (2.1), because \( s \) can go to zero at the location of the point defect, so the term \( \int_{\Omega} s^2 |\nabla \mathbf{n}|^2 \) will be finite.

We simulate the gradient flow evolution of a point defect moving to the center of the domain (\( \Omega \) is the unit square). We set \( \kappa = 2 \) and take the double well potential to have the following splitting:
\[ \psi(s) = \psi_c(s) - \psi_c(s) = 63.0s^2 - (-16.0s^4 + 21.33333333333s^2 + 57.0s^2), \]
with a local minimum at \( s = 0 \) and global minimum at \( s = s^* := 0.750025 \) (see Section 2.1 for more information). The following Dirichlet boundary conditions on \( \partial \Omega \) are imposed for \( s \) and \( \mathbf{n} \):
\[ s = s^*, \quad \mathbf{n} = \frac{(x, y) - (0.5, 0.5)}{|(x, y) - (0.5, 0.5)|}. \]
Fig. 1: Evolution of a point defect toward its equilibrium state (Section 5.2). Time step is $\delta t = 0.02$. The minimum value of $s$, at time index 230, is $2.0226 \cdot 10^{-2}$.

Initial conditions on $\Omega$ for the gradient flow are: $s = s^*$ and a regularized point defect away from the center.

Figure 1 shows the evolution of the director field $\mathbf{n}$ and the scalar degree of orientation parameter $s$. One can see the regularizing effect that $s$ has. We note that an $L^2$ gradient flow scheme, instead of the quasi (weighted) gradient flow we use, yields a much slower evolution to equilibrium.
5.3. Plane defect in 3-D. Next, we simulate the gradient flow evolution of the liquid crystal director field toward a plane defect on a cube domain ($\Omega = (0,1)^3$ is the unit cube). This is motivated by an exact solution found in [45, Sec. 6.4]. We set $\kappa = 0.2$ and remove the double well potential. The following Dirichlet boundary conditions on $\partial \Omega \cap \{(z = 0) \cup \{z = 1\}\}$ are imposed for $(s,n)$:

\begin{align*}
  z = 0 & : s = s^*, \quad n = (1,0,0), \\
  z = 1 & : s = s^*, \quad n = (0,1,0),
\end{align*}

(5.2)

and Neumann conditions are imposed on the remaining part of $\partial \Omega$, i.e. $\nu \cdot \nabla s = 0$ and $\nu \cdot \nabla n = 0$. The exact solution $(s,n)$ (at equilibrium) only depends on $z$ and is given by

\begin{align*}
  n(z) = (1,0,0), & \text{ for } z < 0.5, \quad n(z) = (0,1,0), & \text{ for } z > 0.5, \\
  s(z) = 0, & \text{ at } z = 0.5, \text{ and } s(z) \text{ is linear for } z \in (0,0.5) \cup (0.5,1.0).
\end{align*}

(5.3)

Initial conditions on $\Omega$ for the gradient flow are: $s = s^*$ and a regularized point defect away from the center of the cube.

Figure 2 shows the evolution of the director field $n$ toward the plane defect. Only a few slices are shown in Figure 2 because of the simple form of the equilibrium solution.

Figure 3 (left) shows the components of $n$ evaluated along a one dimensional vertical slice. Clearly, the numerical solution approximates the exact solution well,
Fig. 3: Evolution toward an (equilibrium) plane defect (Section 5.3); time step is $\delta t = 0.02$. Left: plots of the three components of $\mathbf{n}$, evaluated along the vertical line $x = 0.5, y = 0.5$, are shown at three time indices (solid blue curve: $\mathbf{n} \cdot \mathbf{e}_1$, dashed black curve: $\mathbf{n} \cdot \mathbf{e}_2$, dotted red curve: $\mathbf{n} \cdot \mathbf{e}_3$). At equilibrium, $\mathbf{n}$ is nearly piecewise constant with a narrow transition region around $z = 0.5$. Right: plots of the degree-of-orientation $s$, corresponding to $\mathbf{n}$, are shown. The equilibrium solution is piecewise linear, with a kink at $z = 0.5$ where $s \approx 0.008$.

except at the narrow transition region near $z = 0.5$. Furthermore, Figure 3 (right) shows the corresponding evolution of the degree of orientation parameter $s$ (evaluated along the same one dimensional vertical slice). One can see the regularizing effect that $s$ has, i.e. at equilibrium, $s \approx 0.008$ at the $z = 0.5$ plane (the defect plane of $\mathbf{n}$). Our numerical experiments suggest that $s_{|z=0.5} \rightarrow 0$ as the mesh size goes to zero.

5.4. Fluting effect and propeller defect. This example further investigates the effect of $\kappa$ on the presence of defects. An exact solution of a line defect in a right circular cylinder is given in [45, Sec. 6.5]. They show that for $\kappa$ sufficiently large (say $\kappa > 1$) the director field is smooth, but if $\kappa$ is sufficiently small, then a line defect in $\mathbf{n}$ appears along the axis of the cylinder. Our numerical experiments confirm this.

To further illustrate this effect, we conducted a similar experiment for a unit cube domain $\Omega = (0,1)^3$. Again, for simplicity, we remove the double well potential. The following Dirichlet boundary conditions on the vertical sides of the cube $\partial \Omega \cap \{x = 0\} \cup \{x = 1\} \cup \{y = 0\} \cup \{y = 1\}$ are imposed for $(s, \mathbf{n})$:

\begin{equation}
  s = s^*, \quad \mathbf{n}(x,y,z) = \frac{(x,y) - (0.5,0.5)}{|(x,y) - (0.5,0.5)|},
\end{equation}

and Neumann conditions are imposed on the top and bottom parts of $\partial \Omega$, i.e. $\mathbf{\nu} \cdot \nabla s = 0$ and $\mathbf{\nu} \cdot \nabla \mathbf{n} = 0$. Figure 4 shows the equilibrium solution when $\kappa = 2$. The $z$-component of $\mathbf{n}$ is not zero, i.e. it points out of the plane of the horizontal slice that we plot. This is referred to as the “fluting effect” (or escape to the third dimension.
Fig. 4: Equilibrium state (Section 5.4) of \( n \) and \( s \). One horizontal slice \((z = 0.5)\) is plotted: \( n \) on the left, \( s \) on the right \((n \text{ and } s \text{ are approximately independent of } z)\). The director field points out of the plane \((i.e. n \cdot e_3 \neq 0)\) and \( s > 0.278 \), so there is no defect.

\[ s \approx 2 \times 10^{-5} \] near where \( n \) has a discontinuity.

The 3-D shape of the defect resembles two planes intersecting near the \( x = 0.5, y = 0.5 \) vertical line, i.e. the defect looks like an “X” extruded in the \( z \) direction.

5.5. Floating plane defect. This example investigates the effect of the domain shape on the defect. The setup here is essentially the same as in Section 5.4, with \( \kappa = 0.1 \), except the domain is a rectangular box: \( \Omega = (0,1) \times (0,0.7143) \times (0,1) \). Figure 7 shows \( n \) and \( s \) in their final equilibrium state at the \( z = 0.5 \) plane. Both \( n \) and \( s \) are approximately uniform with respect to the \( z \) variable. Instead of the propeller defect, we get a “floating” plane defect aligned with the major axis of the box. Again, the regularizing effect of \( s \) is apparent, i.e. \( s \approx 7 \times 10^{-5} \) near where \( n \) has a discontinuity.

6. Conclusion. We introduced and analyzed a robust finite element method for a degenerate energy functional that models nematic liquid crystals with variable degree of orientation. We also developed a gradient flow scheme for computing energy minimizers, with a strict monotone energy decreasing property. The numerical experiments show a variety of defect structures that Ericksen’s model exhibits. Some of the defect structures are high dimensional with surprising shapes (see Figure 6). An interesting extension of this work is to couple the effect of external fields (e.g. magnetic and electric fields) to the liquid crystal as way to drive and manipulate the defect structures.

Acknowledgements: Nochetto and Zhang acknowledge financial support by
Fig. 5: Evolution toward an (equilibrium) “propeller” defect (Section 5.4). Director field $n$ is shown at five different horizontal slices through the cube. The time step used was $\delta t = 0.02$.

the National Science Foundation through the following grant DMS-1411808. Walker acknowledges financial support by the National Science Foundation through DMS-1418994. Zhang also acknowledges support through the Brin post-doctoral fellowship.

REFERENCES


Fig. 6: Equilibrium state of a “propeller” defect (Section 5.4). One horizontal slice \((z = 0.5)\) is plotted: \(n\) on the left, \(s\) on the right (\(n\) and \(s\) are nearly independent of \(z\)). The \(z\)-component of \(n\) is zero and \(s \approx 2 \times 10^{-5}\) near the discontinuity in \(n\).

Fig. 7: Equilibrium state of a floating plane defect (Section 5.5). One horizontal slice \((z = 0.5)\) is plotted: \(n\) on the left, \(s\) on the right (\(n\) and \(s\) are approximately independent of \(z\)). The \(z\)-component of \(n\) is zero and \(s > 0\) with \(s \approx 7 \times 10^{-5}\) near the discontinuity in \(n\).

[38] ———, Aggregation-based algebraic multigrid for convection-diffusion equations, SIAM Journal


