A Newton-Penalty Method for A Simplified Liquid Crystal Model

Qiya Hu* and Long Yuan†

November 14, 2011

Abstract

In this paper we are concerned with the computation of a liquid crystal model defined by a simplified Oseen-Frank energy functional and a (sphere) nonlinear constraint. A particular case of this model defines the well known *harmonic maps*. We design an new iterative method for solving such a minimization problem with the nonlinear constraint. The main ideas are to linearize the nonlinear constraint by Newton's method and to define a suitable penalty functional associated with the original minimization problem. It is shown that the solution sequence of the new minimization problems with the linear constraints converges to the desired solutions provided that the penalty parameters are chosen by a suitable rule. Numerical results confirm the efficiency of the new method.

Keywords: harmonic maps, nonlinear constraint, Newton's method, regularized functional, saddle points.

AMS subject classifications. 35A40, 65C20, 65N30

1 Introduction

Let Ω be a bounded and Lipschitz domain in $\mathbb{R}^d$ $(d = 2, 3)$, and let $\mathcal{E}$ be the simplified Oseen-Frank energy functional defined by (see, for example, [11])

$$\mathcal{E}(\mathbf{v}) = \frac{1}{2}\int_\Omega (\kappa_1|\nabla \mathbf{v}|^2 + \kappa_2|\nabla \times \mathbf{v}|^2)dx, \quad \mathbf{v} \in H^1(\Omega)^d \ \ (d = 2, 3), \tag{1.1}$$

with $\kappa_1 > 0$ and $\kappa_2 \geq 0$. The target manifold S^{d-1} is the sphere

$$S^{d-1} = \{\mathbf{v} \in \mathbb{R}^d \,|\, F(\mathbf{v}) = 0 \ \ a.e.\ in\ \Omega\},$$

with $F(\mathbf{v}) = |\mathbf{v}|^2 - 1$. We shall study the problem of finding local minima of a constrained minimization problem of the form:

$$\min_{\mathbf{v}\in \mathbf{H}^1_{\mathbf{g}}(\Omega;S^{d-1})} \mathcal{E}(\mathbf{v}). \tag{1.2}$$

*LSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. This author was supported by Natural Science Foundation of China G10771178, The Key Project of Natural Science Foundation of China G11031006 and National Basic Research Program of China G2011309702. (email: hqy@lsec.cc.ac.cn

†Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China. (yuanlong@lsec.cc.ac.cn)

Here, $\mathbf{H}^1_{\mathbf{g}}(\Omega; S^{d-1})$ denotes the set of vector fields with values in the manifold S^{d-1}, and function values and first derivatives in $L^2(\Omega)^d$, such that each element $\mathbf{v}$ in $\mathbf{H}^1_{\mathbf{g}}(\Omega; S^{d-1})$ satisfies $\mathbf{v}|_{\partial\Omega} = \mathbf{g}$ in the sense of trace for a fixed vector field $\mathbf{g}$ defined on the boundary $\partial\Omega$. We assume that the problem (1.2) has a solution $\mathbf{u}$ at least. Problems of the form (1.2) can be found in many practical applications, for example, numerical simulation of liquid crystals. In particular, when $\kappa_2 = 0$, the critical points of the functional $\mathcal{E}$ over $\mathbf{H}^1_{\mathbf{g}}(\Omega; S^{d-1})$ are frequently referred as harmonic maps from Ω into S^{d-1} (see [7] and [8]). For convenience, the critical points of the functional $\mathcal{E}$ over $\mathbf{H}^1_{\mathbf{g}}(\Omega; S^{d-1})$ for all $\kappa_2 \geq 0$ are called *generalized harmonic maps*. The computation of the generalized harmonic maps is a difficult topic, since the target manifold S^{d-1} is not a convex set, the problem (1.2) usually possesses many solutions (when $d = 3$), some of which may be non-smooth (see [13] for some details).

The goal of this paper is to develop an iterative method for computing the *generalized harmonic maps*. The simplest case of (1.2) with $\kappa_2 = 0$ has been studied by many researchers (see, for example, [1], [3], [4] and [16]). It is known that the projection method is the most natural one for solving this problem. The projection method was first introduced in [1], and was further developed by [3], [10] (for the Landau-Lifshitz equation) and [16]. A variant of the projection method was proposed in [4]. An advantage of the projection method is that it is globally convergent (design of a globally convergent iterative method for solving complicated nonlinear problems is both important and difficult). However, the discretization version of the projection method may be not convergent for the general regular and quasi-uniform triangulation even if $\kappa_2 = 0$ (see [3] for detailed discussion). To our knowledge, there is few work to study numerical methods for solving (1.2) with $\kappa_2 > 0$. The first important attempt to solve this problem was made in [11] where (1.2) was transformed into a time-evolution problem with a penalty term by using the heat flow (gradient flow) method [8] and the penalty method, and the resulting variational problem was solved by the operator-splitting method after discretization of the time variable. The heat flow method for computation of p harmonic maps was also studied recently in [5]. The main advantage of the heat flow method is that, when the time step size is small enough, the resulting nonlinear problem at one time step can be solved more easily than the original independent-time problem. But, one has to iteratively solve a nonlinear problem at each time step in the gradient flow method. A general method for solving (1.2) is the penalty method, namely, use Ginzburg-Landau free energy instead of $\mathcal{E}(\mathbf{v})$ (see, for example, [6]). However, it is difficult to design an efficient iterative method for solving the variational problem arising from Ginzburg-Landau free energy, since the penalty term is too complicated. A saddle-point method is studied in [14] for the case $d = 2$ and $\kappa_2 = 0$, but for more general function F.

In this paper we propose a new iterative method for solving a discrete problem associated with (1.2). We consider the more general case with $\kappa_2 \geq 0$. Our main ideas can be described as follows: use Newton method to linearize the constraint $F(\mathbf{v}) = 0$, and replace the energy functional $\mathcal{E}(\mathbf{v})$ by a new energy functional with a *nonstandard* penalty term associated with the constraint $F(\mathbf{v}) = 0$. Then we solve the minimization problems with the penalized energy on the *linearized* constraint spaces. We find that the new minimization problem can be solved easily, and the cost of computation is almost same with that in the projection method. It will be shown that the resulting solution sequence is always *globally* convergent without particular requirement to triangulation, provided that the penalty term is designed properly. Our idea can be extended to more general liquid crystal model by designing a different penalty term (which will be done in another paper).

The outline of this paper is as follows. In Section 2, we give some notations. In Section 3, we introduce a finite element discretization for (1.2), and analyze the convergence of the resulting approximate solution. In section 4, we describe the motivation to designing the new iterative method. In Section 5, we present the Newton-penalty method, and prove an important property of the method: the energy is decreasing. The convergence of the new method is proved in Section 6. In Section 7, we discuss some details about the solution of the minimization problem associated with the Newton-penalty method. Some numerical results are given in Section 8.

2 Notations and Preliminaries

Throughout this paper we use c and C to denote generic positive constants, not necessarily the same at different occurrences. It is assumed that the constants are independent of the mesh size h which will be introduced later. For vectors $\mathbf{v}, \mathbf{w} \in \mathbb{R}^d$ we use $\mathbf{v} \cdot \mathbf{w}$ to denote the Euclidian inner product, while the notation $\mathcal{A} : \mathcal{B}$ is used to denote the Frobenius inner product of two matrices $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{d\times d}$. The corresponding norms are given by $|\mathbf{v}|$ and $|\mathcal{A}|$, respectively. For a vector or matrix $\mathcal{A}$, $\mathcal{A}^t$ is the transpose of $\mathcal{A}$.

For $m \geq 0$ we use $H^m = H^m(K)$ to denote the real valued L^2– based Sobolev spaces on domain $K \subset \mathbb{R}^d$, with the corresponding norm by $\|\cdot\|_{m,K}$, and use $|\cdot|_{m,K}$ to denote the semi norm involving only the mth order derivatives. The subspace H_0^m is the closure of $C_0^\infty(K)$ in H^m, while H^{-m} is the dual of H_0^m with respect to an extension of the L^2 inner product $(\cdot,\cdot)$. The corresponding L^∞–based Sobolev spaces are denoted by $W^{m,\infty}(K)$, with the norm $\|\cdot\|_{m,\infty,K}$. The notation $\mathbf{H}^1(\Omega)$ and $\mathbf{W}^{1,\infty}(\Omega)$ will be used for the vector version of the corresponding spaces.

In general, we use boldface symbols for vector or matrix valued functions. For convenience, we give the exact definitions of more notations only for the case with $S^{d-1} \subset \mathbb{R}^3$ (i.e., $d = 3$) in the rest of this section. The exact definitions of notations for $d = 2$ can be given with obvious modification.

The gradient operator with respect to the spatial variable $\mathbf{x} = (x_1, x_2, x_3)$ is denoted as $\nabla = (\partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3)^t$. The gradient of a vector valued function $\mathbf{v} = (v_1, v_2, v_3)^t$, $\nabla\mathbf{v}$, is the matrix valued function obtained by taking the gradient row–wise, i.e.,

$$\nabla\mathbf{v} = (\nabla v_1, \nabla v_2, \nabla v_3)^t \quad \text{or} \quad (\nabla\mathbf{v})_{ij} = \partial v_i/\partial x_j.$$

For two vector valued functions $\mathbf{v}$ and $\mathbf{w} = (w_1\ w_2\ w_3)^t$, we define $\mathbf{v} \times \nabla\mathbf{w}$ as the matrix valued function obtained by taking the vector product row–wise, i.e.,

$$\mathbf{v} \times \nabla\mathbf{w} = (\mathbf{v} \times \nabla w_1,\ \mathbf{v} \times \nabla w_2,\ \mathbf{v} \times \nabla w_3)^t.$$

It can be verified that

$$\nabla(\mathbf{v} \times \mathbf{w}) = \nabla\mathbf{v} \times \mathbf{w} + \mathbf{v} \times \nabla\mathbf{w}. \tag{2.1}$$

As usual, we use $\mathbf{D}F$ to denote the gradient of F, i.e.,

$$\mathbf{D}F(\mathbf{v}) = (\partial F/\partial v_1, \partial F/\partial v_2, \partial F/\partial v_3)^t = (2v_1, 2v_2,\ 2v_3)^t = 2\mathbf{v} \quad (d = 3). \tag{2.2}$$

The corresponding Hessian is denoted by

$$\mathbf{D}^2F(\mathbf{v}) = (\partial^2 F/\partial v_i \partial v_j)_{i,j=1}^3 = 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.3}$$

For the boundary function $\mathbf{g}$ of (1.2) we assume that it has been extended into the interior of Ω such that $\mathbf{g} \in \mathbf{H}^1(\Omega)$. For such $\mathbf{g}$, we let

$$\mathbf{H}^1_{\mathbf{g}}(\Omega) = \{\mathbf{v} \in \mathbf{H}^1(\Omega) : \mathbf{v} = \mathbf{g} \text{ on } \partial\Omega\}.$$

For a functional $\mathbf{J} : \mathbf{H}^1(\Omega) \to \mathbb{R}$, we use $\mathbf{J}'(\mathbf{v}) : \mathbf{H}^1(\Omega) \to \mathbb{R}$ to represent the Gateaux derivative of $\mathbf{J}$ at $\mathbf{v} \in \mathbf{H}^1(\Omega)$. Similarly, we use $\mathbf{J}''(\mathbf{v}) : \mathbf{H}^1(\Omega) \to \mathbf{H}^1(\Omega)$ to denote the second-order Gateaux derivative of $\mathbf{J}$ at $\mathbf{v} \in \mathbf{H}^1(\Omega)$. Let $\mathbf{J}'(\mathbf{v})\mathbf{w}$ and $\mathbf{J}''(\mathbf{v})\mathbf{w}$ denote the images of the maps $\mathbf{J}'(\mathbf{v})$ and $\mathbf{J}''(\mathbf{v})$ at $\mathbf{w} \in \mathbf{H}^1(\Omega)$, respectively. It is easy to see that

$$\mathcal{E}'(\mathbf{v})\mathbf{w} = \int_\Omega \big[\kappa_1 \nabla\mathbf{v} : \nabla\mathbf{w} + \kappa_2(\nabla \times \mathbf{v}) \cdot (\nabla \times \mathbf{w})\big] d\mathbf{x}, \quad \mathbf{v} \in \mathbf{H}^1(\Omega)$$

and

$$\mathcal{E}''(\mathbf{v})\mathbf{w} \cdot \mathbf{w} = \int_\Omega \big[\kappa_1 \nabla\mathbf{w} : \nabla\mathbf{w} + \kappa_2(\nabla \times \mathbf{w}) \cdot (\nabla \times \mathbf{w})\big] d\mathbf{x}, \quad \mathbf{v} \in \mathbf{H}^1(\Omega).$$

Note that

$$0 \le \|\nabla \times \mathbf{w}\|^2_{0,\Omega} \le C\|\nabla\mathbf{w}\|^2_{0,\Omega},$$

we have

$$\kappa_1\|\nabla\mathbf{w}\|^2_{0,\Omega} \le \mathcal{E}''(\mathbf{v})\mathbf{w} \cdot \mathbf{w} \le C(\kappa_1 + \kappa_2)\|\nabla\mathbf{w}\|^2_{0,\Omega}, \quad \mathbf{v},\ \mathbf{w} \in \mathbf{H}^1(\Omega). \tag{2.4}$$

The critical points of the functional $\mathcal{E}$ over $\mathbf{H}^1_{\mathbf{g}}(\Omega; S^{d-1})$, i.e., the stationary points of the minimization problem (1.2), are called *generalized harmonic maps* from Ω into S^{d-1}. A vector field $\mathbf{u} \in \mathbf{H}^1_{\mathbf{g}}(\Omega; S^{d-1})$ is such a critical point if it satisfies

$$\mathcal{E}'(\mathbf{u})\mathbf{v} = 0 \tag{2.5}$$

for any $\mathbf{v}$ in the tangential space of $\mathbf{H}^1_{\mathbf{g}}(\Omega; S^{d-1})$ at $\mathbf{u}$, i.e., for any $\mathbf{v} \in \mathbf{H}^1_0(\Omega)$ such that $\mathbf{D}F(\mathbf{u}) \cdot \mathbf{v} \equiv 0$ (refer to [8]).

For a two-dimensional vector $\mathbf{v}$, let $\mathbf{v}^\perp$ denote the vector obtained by a rotation of 90 degrees for $\mathbf{v}$. The following result can be obtained directly from the above definition.

Proposition 2.1 The vector function $\mathbf{u} \in \mathbf{H}^1_{\mathbf{g}}(\Omega; S^{d-1})$ is a harmonic map if and only if

$$\mathcal{E}'(\mathbf{u})(\phi\mathbf{D}F(\mathbf{u})^\perp) = 0, \quad \forall\phi \in C^\infty_0(\Omega) \quad (d = 2), \tag{2.6}$$

or

$$\mathcal{E}'(\mathbf{u})(\mathbf{v} \times \mathbf{D}F(\mathbf{u})) = 0, \quad \forall\mathbf{v} \in \mathbf{C}^\infty_0(\Omega) \quad (d = 3). \tag{2.7}$$

□

3 The finite element discretizations

In this section we introduce a discrete version of the problem (1.2) and study convergence of the resulting approximate solution.

In the rest of the paper we assume that the domain $\Omega \subset \mathbb{R}^d$ is a polygon or a polyhedron. Given a family of shape regular and quasi–uniform triangulation $\{\mathcal{T}_h\}$ in Ω with a mesh size $h < 1$, let $\mathcal{N}_h = \{p_k\}_{k=1}^N$ denote the set of nodes associated with $\mathcal{T}_h$. As usual, we use φ_k to denote the nodal basis function associated with the node p_k.

We use V_h to denote the space of continuous piecewise linear functions with respect to $\mathcal{T}_h$ and $V_{h,0} = V_h \cap H_0^1(\Omega)$. The notation $\mathbf{V}_h$ and $\mathbf{V}_{h,0}$ will be used for the vector version of the corresponding spaces. We will use π_h to denote the usual nodal interpolation operators onto the spaces V_h and $\mathbf{V}_h$. Let π_h^∂ denote the restriction of π_h on $\partial\Omega$.

The following inverse inequality for finite element functions in V_h will be used later:

$$\|v_h\|_{1,\Omega} \le Ch^{-1}\|v_h\|_{0,\Omega} \quad \text{and} \quad \|v_h\|_{0,\infty,\Omega} \le C\beta_h(d)\|v_h\|_{1,\Omega}, \quad \forall v_h \in V_h. \tag{3.1}$$

Here, $\beta_h(d) = \log^{\frac{1}{2}}(1/h)$ for $d = 2$, and $\beta_h(d) = h^{-\frac{1}{2}}$ for $d = 3$.

As in [3], we assume that $\mathbf{g}$ is continuous on $\partial\Omega$. Set $\mathbf{g}_h = \pi_h^\partial \mathbf{g}$ (on $\partial\Omega$), and define

$$\mathbf{V}_{h,\mathbf{g}} = \{\mathbf{v} \in \mathbf{V}_h : \ \mathbf{v}|_{\partial\Omega} = \mathbf{g}_h\}.$$

We will consider the following discretized minimization problem:

$$\min_{\mathbf{v}\in\mathbf{V}_{h,\mathbf{g}}} \mathcal{E}(\mathbf{v}) \ \text{ subject to } F(\mathbf{v}) = 0 \text{ on } \mathcal{N}_h. \tag{3.2}$$

Since the minimization functional is convex and continuous, and the set

$$\mathbf{K}_h = \{\mathbf{v} \in \mathbf{V}_{h,\mathbf{g}} : \ F(\mathbf{v}) = 0 \text{ on } \mathcal{N}_h\}$$

is closed and compact, the problem (3.2) has a solution at least for a fixed h (refer to Chapter 7 of [17]).

For convenience, the critical points of the functional $\mathcal{E}$ over the discrete space $\mathbf{K}_h$, i.e., stationary points of the minimization problem (3.2), are called *discrete generalized harmonic maps* in $\mathbf{K}_h$. It is well known that $\mathbf{u}_h^* \in \mathbf{K}_h$ is a stationary point of the minimization problem (3.2) if and only if there exists a Lagrange multiplier $\chi = (b_1, b_2, \cdots, b_N)^t$ such that

$$\mathcal{E}'(\mathbf{u}_h^*)\mathbf{v} + \sum_{k=1}^{N} b_k \mathbf{D}F(\mathbf{u}_h^*(p_k)) \cdot \mathbf{v}(p_k) = 0, \quad \forall \mathbf{v} \in \mathbf{V}_{h,0}. \tag{3.3}$$

The following result can be viewed as the discrete version of **Proposition 2.1**.

Proposition 3.1 A vector function $\mathbf{u}_h^* \in \mathbf{K}_h$ is a *discrete generalized harmonic map* if and only if $\mathbf{u}_h^*$ satisfies

$$\mathcal{E}'(\mathbf{u}_h^*)\pi_h(\phi_h \mathbf{D}F(\mathbf{u}_h^*)^\perp) = 0, \quad \forall \phi_h \in V_{h,0} \quad (d = 2), \tag{3.4}$$

or

$$\mathcal{E}'(\mathbf{u}_h^*)\pi_h(\mathbf{v}_h \times \mathbf{D}F(\mathbf{u}_h^*)) = 0, \quad \forall \mathbf{v}_h \in \mathbf{V}_{h,0} \quad (d = 3). \tag{3.5}$$

Proof. It suffices to prove that the condition (3.4) or (3.5) is equivalent to the existence of numbers $\{b_k\}$ satisfying the relation (3.3).

Without loss of generality, we consider only the case of $d = 3$. Assume that $\mathbf{u}_h^* \in \mathbf{K}_h$ satisfies (3.5). We try to verify (3.3). Let $\Phi_k^r \in \mathbf{V}_{h,0}$ $(r = 1, 2, 3)$ be the three nodal basis vectors associated with an *interior* node p_k. Namely,

$$\Phi_k^1 = (\varphi_k, 0, 0)^t, \ \ \Phi_k^2 = (0, \varphi_k, 0)^t \ \text{ and } \ \Phi_k^3 = (0, 0, \varphi_k)^t.$$

In the rest of the proof, we consider only the index k associated to an interior nodes p_k. Then equality (3.5) is equivalent to

$$\mathcal{E}'(\mathbf{u}_h^*)\pi_h(\Phi_k^r \times \mathbf{D}F(\mathbf{u}_h^*)) = 0, \quad \forall k \ \ (r = 1, 2, 3). \tag{3.6}$$

By the definition of π_h, one can verify that

$$\pi_h(\Phi_k^r \times \mathbf{D}F(\mathbf{u}_h^*)) = \Phi_k^r \times \mathbf{D}F(\mathbf{u}_h^*(p_k)).$$

Thus one gets from (3.6) that

$$\mathcal{E}'(\mathbf{u}_h^*)(\Phi_k^r \times \mathbf{D}F(\mathbf{u}_h^*(p_k))) = 0, \quad \forall k \ \ (r = 1, 2, 3).$$

It can be verified, by the above equality, that the vector

$$(\mathcal{E}'(\mathbf{u}_h^*)\Phi_k^1,\ \mathcal{E}'(\mathbf{u}_h^*)\Phi_k^2,\ \mathcal{E}'(\mathbf{u}_h^*)\Phi_k^3)^t$$

is parallel to the vector $\mathbf{D}F(\mathbf{u}_h^*)(p_k)$. Thus, there exists a number b_k such that

$$\mathcal{E}'(\mathbf{u}_h^*)\Phi_k^r = -b_k\big[\mathbf{D}F(\mathbf{u}_h^*)(p_k) \cdot \Phi_k^r(p_k)\big], \quad \forall k \ \ (r = 1, 2, 3). \tag{3.7}$$

This implies (3.3) because in the sum, the terms corresponding to boundary nodes vanish.

On the other hand, the equality (3.5) obviously follows by (3.3).

Remark 3.1 *From (2.2), we have* $|\mathbf{D}F(\mathbf{u}_h^*(p_k))|^2 = 4 \neq 0$. *It follows by (3.7) that*

$$b_k = -\frac{\mathcal{E}'(\mathbf{u}_h^*)(\mathbf{D}F(\mathbf{u}_h^*(p_k))\varphi_k)}{|\mathbf{D}F(\mathbf{u}_h^*(p_k))|^2}$$

for each interior node p_k. *The stationary point* $\mathbf{u}_h^*$ *is a local strict minimizer of (3.2) if and only if* $\mathbf{u}_h^*$ *satisfies the condition (for such* b_k*)*

$$\mathcal{E}''(\mathbf{u}_h^*)\mathbf{v} \cdot \mathbf{v} + \sum_{k=1}^{N} b_k \mathbf{D}^2 F(\mathbf{u}_h^*(p_k))\mathbf{v}(p_k) \cdot \mathbf{v}(p_k) > 0, \tag{3.8}$$

for all $\mathbf{v} \in \{\mathbf{v} : \mathbf{D}F(\mathbf{u}_h^*(z)) \cdot \mathbf{v}(z) = 0,\ \forall z \in \mathcal{N}_h\}$. *Some details on this result can be found in Chapter 8 of [17].*

In the rest of this section, we derive some approximate properties of $\mathbf{u}_h$. To this end, we first give two simple lemmata.

Lemma 3.1 *For any* $\mathbf{v}_h \in \mathbf{K}_h$, *we have* $F(\mathbf{v}_h) \leq 0$ *in* Ω.

Proof. Let φ_k be the basis function on the node p_k. Then,

$$\mathbf{v}_h(p) = \sum_{k=1}^{N} \mathbf{v}_h(p_k)\varphi_k(p), \quad \forall p \in \Omega. \tag{3.9}$$

Since

$$\varphi_k(p) > 0 \quad \text{and} \quad \sum_{k=1}^{N} \varphi_k(p) = 1,$$

we get by (3.9) and the convexity of the functional F

$$F(\mathbf{v}_h(p)) \leq \sum_{k=1}^{N} \varphi_k(p) F(\mathbf{v}_h(p_k)) \leq 0.$$

□

Lemma 3.2 *Let $\mathbf{w}_h \in \mathbf{K}_h$. Assume that the sequence $\{|\mathbf{w}_h|_{1,\Omega}\}$ is uniformly bounded with h. Then the sequence $\{\|\mathbf{w}_h\|_{1,\Omega}\}$ is also uniformly bounded with h. Moreover, we have*

$$\|\mathbf{D}F(\mathbf{w}_h)\|_{0,\infty,\Omega} \le C, \quad |\mathbf{D}F(\mathbf{w}_h)|_{1,\Omega} \le C, \tag{3.10}$$

and

$$|F(\mathbf{w}_h)|_{1,\Omega} \le C. \tag{3.11}$$

Proof. Since $F(\mathbf{w}_h) = 0$, we have $|\mathbf{w}_h|^2 = 1$ in Ω. We further get by (2.2)

$$|\mathbf{D}F(\mathbf{w}_h)|^2 = 4 \text{ in } \Omega. \tag{3.12}$$

These imply the first conclusion of the lemma and the first inequality in (3.10).

The second inequality of (3.10) follows by (2.2) and the assumption.

It can be verified that

$$|F(\mathbf{w}_h)|_{1,\Omega} \le C\|\mathbf{D}F(\mathbf{w}_h)\|_{0,\infty,\Omega} \cdot \|\nabla \mathbf{w}_h\|_{0,\Omega}. \tag{3.13}$$

Then we deduce (3.11) by (3.12) and the assumption.

□

Theorem 3.1 *Let $\mathbf{u}_h \in \mathbf{K}_h$ be a discrete harmonic map associated with the mesh size h. Assume that the sequence $\{|\mathbf{u}_h|_{1,\Omega}\}$ is bounded with respect to h. Then*

(i) there exists a subsequence of the sequence $\{\mathbf{u}_h\}_{h>0}$, which is denoted by $\{\mathbf{u}_h\}_{h>0}$ itself, such that $\{\mathbf{u}_h\}_{h>0}$ converges to $\mathbf{u} \in \mathbf{H}^1_{\mathbf{g}}(\Omega; S^{d-1})$ weakly in H^1 and strongly in L^4;

(ii) $\mathbf{u}$ is a harmonic map.

Proof. (i) By Lemma 3.2, the sequence $\{\|\mathbf{u}_h\|_{1,\Omega}\}$ is also uniformly bounded with h. Thus there exists a subsequence of the sequence $\{\mathbf{u}_h\}_{h>0}$, which converges to a $\mathbf{u} \in \mathbf{H}^1(\Omega)$ weakly in H^1 and strongly in L^4. It suffices to prove $\mathbf{u} \in \mathbf{H}^1_{\mathbf{g}}(\Omega,\ S^{d-1})$.

Since $\pi_h F(\mathbf{u}_h) = 0$, we have

$$\|F(\mathbf{u}_h)\|_{0,\Omega} = \|(I - \pi_h)F(\mathbf{u}_h)\|_{0,\Omega} \le Ch|F(\mathbf{u}_h)|_{1,\Omega}. \tag{3.14}$$

This, together with (3.11), leads to

$$\|F(\mathbf{u}_h)\|_{0,\Omega} \le Ch \to 0^+ \quad \text{when } h \to 0^+. \tag{3.15}$$

On the other hand, we have

$$F(\mathbf{u}) - F(\mathbf{u}_h) = 2\mathbf{u}_h \cdot (\mathbf{u} - \mathbf{u}_h) + |\mathbf{u} - \mathbf{u}_h|^2.$$

Hence

$$\|F(\mathbf{u}) - F(\mathbf{u}_h)\|_{0,\Omega} \le C(\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega} + \|\mathbf{u} - \mathbf{u}_h\|^2_{L^4(\Omega)}) \to 0^+ \quad \text{when } h \to 0^+.$$

This, together with (3.15), yields

$$\|F(\mathbf{u})\|_{0,\Omega} = 0.$$

Thus $F(\mathbf{u}) = 0$ a.e. in Ω. Similarly, we can prove $\mathbf{u}|_{\partial\Omega} = \mathbf{g}$ (refer to [3]), and so $\mathbf{u} \in \mathbf{H}^1_{\mathbf{g}}(\Omega,\ S^{d-1})$.

(ii) The proof is similar to that of Theorem 3.8 in [3]. Let $\mathbf{v} \in \mathbf{C}_0^\infty(\Omega)$, and set $\mathbf{v}_h = \pi_h\mathbf{v} \in \mathbf{V}_{h,0}$. It follows by Proposition 3.1 that

$$\begin{aligned}
\mathcal{E}'(\mathbf{u})(\mathbf{v}\times\mathbf{D}F(\mathbf{u})) &= \mathcal{E}'(\mathbf{u})(\mathbf{v}\times\mathbf{D}F(\mathbf{u})) - \mathcal{E}'(\mathbf{u}_h)\pi_h(\mathbf{v}_h\times\mathbf{D}F(\mathbf{u}_h)) \\
&= [\mathcal{E}'(\mathbf{u})(\mathbf{v}\times\mathbf{D}F(\mathbf{u})) - \mathcal{E}'(\mathbf{u}_h)(\mathbf{v}\times\mathbf{D}F(\mathbf{u}_h))] \\
&+ [\mathcal{E}'(\mathbf{u}_h)(\mathbf{v}\times\mathbf{D}F(\mathbf{u}_h)) - \mathcal{E}'(\mathbf{u}_h)(\mathbf{v}_h\times\mathbf{D}F(\mathbf{u}_h))] \\
&+ [\mathcal{E}'(\mathbf{u}_h)(\mathbf{v}_h\times\mathbf{D}F(\mathbf{u}_h)) - \mathcal{E}'(\mathbf{u}_h)\pi_h(\mathbf{v}_h\times\mathbf{D}F(\mathbf{u}_h))] \\
&= I_1 + I_2 + I_3. \qquad (3.16)
\end{aligned}$$

We first verify $I_2, I_3 \to 0$ when $h \to 0^+$. In fact, we have by the assumption

$$\begin{aligned}
|I_2| &= |\mathcal{E}'(\mathbf{u}_h)((\mathbf{v}-\mathbf{v}_h)\times\mathbf{D}F(\mathbf{u}_h))| \\
&\le C|\mathbf{u}_h|_{1,\Omega}\cdot|(\mathbf{v}-\mathbf{v}_h)\times\mathbf{D}F(\mathbf{u}_h)|_{1,\Omega} \\
&\le C|(\mathbf{v}-\mathbf{v}_h)\times\mathbf{D}F(\mathbf{u}_h)|_{1,\Omega}. \qquad (3.17)
\end{aligned}$$

It follows by (2.1) that

$$\begin{aligned}
|(\mathbf{v}-\mathbf{v}_h)\times\mathbf{D}F(\mathbf{u}_h)|_{1,\Omega} &\le 2[\|\nabla(\mathbf{v}-\mathbf{v}_h)\times\mathbf{D}F(\mathbf{u}_h)\|_{0,\Omega} + \|(\mathbf{v}-\mathbf{v}_h)\times\nabla(\mathbf{D}F(\mathbf{u}_h))\|_{0,\Omega}] \\
&\le C\|\mathbf{D}F(\mathbf{u}_h)\|_{0,\infty,\Omega}\cdot|\mathbf{v}-\mathbf{v}_h|_{1,\Omega} + |\mathbf{D}F(\mathbf{u}_h)|_{1,\Omega}\cdot|\mathbf{v}-\mathbf{v}_h|_{0,\infty,\Omega}.
\end{aligned}$$

Thus, by (3.10) and the convergence of the interpolation operator, we get

$$|(\mathbf{v}-\mathbf{v}_h)\times\mathbf{D}F(\mathbf{u}_h)|_{1,\Omega} \to 0^+ \quad (h\to 0^+).$$

This, together with (3.17), leads to $I_2 \to 0$ $(h \to 0^+)$. Similarly, we have $I_3 \to 0$ $(h \to 0^+)$.

Now we consider the term I_1. By the definitions, we have

$$\begin{aligned}
I_1 &= [\int_\Omega \nabla\mathbf{u} : (\nabla\mathbf{v}\times\mathbf{D}F(\mathbf{u}))d\mathbf{x} - \int_\Omega \nabla\mathbf{u}_h : (\nabla\mathbf{v}\times\mathbf{D}F(\mathbf{u}_h))d\mathbf{x}] \\
&+ [\int_\Omega (\nabla\times\mathbf{u}) : ((\nabla\times\mathbf{v})\times\mathbf{D}F(\mathbf{u}))d\mathbf{x} - \int_\Omega (\nabla\times\mathbf{u}_h) : ((\nabla\times\mathbf{v})\times\mathbf{D}F(\mathbf{u}_h))d\mathbf{x}] \\
&= I_{11} + I_{12}. \qquad (3.18)
\end{aligned}$$

Here we have used the equalities $\mathbf{D}F(\mathbf{u}) = 2\mathbf{u}$ and $\mathbf{D}F(\mathbf{u}_h) = 2\mathbf{u}_h$.

It is clear that

$$I_{11} = \int_\Omega \nabla(\mathbf{u}-\mathbf{u}_h) : (\nabla\mathbf{v}\times\mathbf{D}F(\mathbf{u}))d\mathbf{x} + \int_\Omega \nabla\mathbf{u}_h : (\nabla\mathbf{v}\times(\mathbf{D}F(\mathbf{u}) - \mathbf{D}F(\mathbf{u}_h)))d\mathbf{x}. \quad (3.19)$$

Since $\mathbf{v} \in C^\infty(\Omega)$, we have $\nabla\mathbf{v}\times\mathbf{D}F(\mathbf{u}) \in \mathbf{L}^2(\Omega)$, therefore the linear functional

$$J(\cdot) = \int_\Omega (\nabla\ \cdot) : (\nabla\mathbf{v}\times\mathbf{D}F(\mathbf{u}))d\mathbf{x}$$

is bounded on $\mathbf{H}^1(\Omega)$. Thus we get by the result (i)

$$\int_\Omega \nabla(\mathbf{u}-\mathbf{u}_h) : (\nabla\mathbf{v}\times\mathbf{D}F(\mathbf{u}))d\mathbf{x} \to 0 \quad (h\to 0^+). \qquad (3.20)$$

On the other hand, we have by (2.2)

$$\|\mathbf{D}F(\mathbf{u}) - \mathbf{D}F(\mathbf{u}_h)\|_{0,\Omega} = 2\|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}.$$

Then we can obtain by using Holder inequality

$$\begin{aligned} |\int_\Omega \nabla \mathbf{u}_h : (\nabla \mathbf{v} \times (\mathbf{D}F(\mathbf{u}) - \mathbf{D}F(\mathbf{u}_h)))d\mathbf{x}| &\leq |\mathbf{u}_h|_{1,\Omega} \cdot \|\nabla \mathbf{v}\|_{0,\infty,\Omega} \cdot \|\mathbf{D}F(\mathbf{u}) - \mathbf{D}F(\mathbf{u}_h)\|_{0,\Omega} \\ &\leq C|\mathbf{u}_h|_{1,\Omega} \cdot \|\mathbf{v}\|_{1,\infty,\Omega} \cdot \|\mathbf{u} - \mathbf{u}_h\|_{0,\Omega}. \end{aligned}$$

Hence, by the assumption and the conclusion (i), we get

$$\int_\Omega \nabla \mathbf{u}_h : (\nabla \mathbf{v} \times (\mathbf{D}F(\mathbf{u}) - \mathbf{D}F(\mathbf{u}_h)))d\mathbf{x} \to 0 \quad (h \to 0^+).$$

Substituting (3.19) by (3.20) and the above relation, we obtain $I_{11} \to 0$ $(h \to 0^+)$. Similarly, we have $I_{12} \to 0$ $(h \to 0^+)$. All these, together with (3.18), yields $I_1 \to 0$ $(h \to 0^+)$.

Finally we obtain the desired result by (3.16).

□

4 Motivations

In the rest of the paper, we develop an efficient method to solve the problem (3.2). To explain our idea more clearly, we first investigate two possible methods.

In the commonly used penalty approach, c.f. [11], one is seeking a minimizer for the following regularized problem:

$$\min_{\mathbf{v}_h \in \mathbf{V}_{h,\mathbf{g}}} \mathcal{E}(\mathbf{v}) + \frac{1}{2\epsilon} \int_\Omega |\pi_h F(\mathbf{v}_h)|^2 d\mathbf{x},$$

where the penalty parameter $\epsilon > 0$ is small. Formally, the necessary equilibrium condition for this problem is: Find $\mathbf{u}_h^\epsilon \in \mathbf{V}_{h,\mathbf{g}}$, such that

$$\int_\Omega \nabla \mathbf{u}_h^\epsilon : \nabla \mathbf{v}_h d\mathbf{x} + \frac{1}{\epsilon} \int_\Omega \pi_h F(\mathbf{u}_h^\epsilon) \pi_h \mathbf{D}F(\mathbf{u}_h^\epsilon) \cdot \mathbf{v}_h d\mathbf{x} = 0, \quad \mathbf{v}_h \in \mathbf{V}_{h,0}.$$

A difficulty with this approach is that the penalty parameter ϵ needs to be chosen sufficiently small in order to resolve the constraint, and usually it also needs to be related to the discretization parameter h. However, for small penalty parameters, numerical instabilities may occur. Moreover, the functional defined by the above regularized problem is not convex, which increase the difficulty for solving the corresponding nonlinear variational problem.

It is well known that the difficulty on the problem (3.2) is to find an efficient method to deal with the constraint $\pi_h F(\mathbf{v}) = 0$. A possible way is to use Newton method to linearize the constraint and then solve a minimization problem with a linear constraint. More specifically, let $\mathbf{u}_n$ be a solution obtained at iteration n, we need to find a new iterative solution $\mathbf{u}_{n+1}$ from $\mathbf{u}_n$. The constrain we shall impose for $\mathbf{u}_{n+1}$ is

$$\mathbf{D}F(\mathbf{u}_n) \cdot (\mathbf{u}_{n+1} - \mathbf{u}_n) = -F(\mathbf{u}_n) \quad \text{on } \mathcal{N}_h. \tag{4.1}$$

i.e. $\mathbf{u}_{n+1}$ does not satisfy the full constraint $\pi_h F(\mathbf{u}_{n+1}) = 0$, instead, we require $\mathbf{u}_{n+1}$ to satisfy a linearized constraint which is coming from a Newton's linearization of $\pi_h F(\mathbf{v}) = 0$

at $\mathbf{u}_n$. We also need $\mathbf{u}_{n+1}$ to minimize the energy functional. Thus, it is necessary to let $\mathbf{u}_{n+1}$ be the solution of the following problem:

$$\min_{\mathbf{v}_h \in \mathbf{V}_{h,\mathbf{g}}} \mathcal{E}(\mathbf{v}_h) \ \text{ subject to } \ \mathbf{D}F(\mathbf{u}_n)\cdot(\mathbf{v}_h - \mathbf{u}_n) = -F(\mathbf{u}_n) \ \text{ on } \ \mathcal{N}_h. \tag{4.2}$$

This minimization problem (with linear constraint) can be solved more easily, but it seems that the solution sequence $\{\mathbf{u}_{n+1}\}$ does not converge except in some particular situations.

In the following we propose another way to solve (3.2) based on an important observation.

As we will see in Lemma 5.1, we have

$$\pi_h F(\mathbf{v}) \geq 0 \ \text{ if } \ \mathbf{D}F(\mathbf{u}_n)\cdot(\mathbf{v} - \mathbf{u}_n) = -F(\mathbf{u}_n) \ \text{ on } \ \mathcal{N}_h. \tag{4.3}$$

Due to this special property, one can add a particular regularization term into the cost functional of (4.2) and consider the following minimization problem:

$$\min_{\mathbf{v}_h \in \mathbf{V}_{h,\mathbf{g}}} [\mathcal{E}(\mathbf{v}_h) + \gamma(h)\int_\Omega \pi_h F(\mathbf{v}_h) d\mathbf{x}] \\ \text{subject to } \ \mathbf{D}F(\mathbf{u}_n)\cdot(\mathbf{v}_h - \mathbf{u}_n) = -F(\mathbf{u}_n) \ \text{ on } \ \mathcal{N}_h. \tag{4.4}$$

Here, $\gamma(h)$ is a nonnegative real number. The parameter $\gamma(h)$ should be chosen as $\gamma(h) \cong h^{-2}$ except some particular cases. Note that $\gamma(h)\int_\Omega \pi_h F(\mathbf{v}) d\mathbf{x}$ is not the standard penalty term. Such design enhances the global convexity of the minimization functional of (4.4).

The solution of (4.4) is still denoted by $\mathbf{u}_{n+1}$. Fortunately, we can prove that the new solution sequence $\{\mathbf{u}_{n+1}\}$ possesses global convergence for an initial data $\mathbf{u}_0 \in \mathbf{V}_{h,\mathbf{g}}$ which satisfies $\pi_h F(\mathbf{u}_0) \geq 0$. Since the constraint for the problem (4.4) is linear, we can use a Lagrange multiplier to deal with the constraint.

We emphasize that the regularization term

$$\gamma(h)\int_\Omega \pi_h F(\mathbf{v}) d\mathbf{x}$$

is related to the original nonlinear constraint $\pi_h F(\mathbf{v}_h) = 0$, instead of the current linear constraint

$$\mathbf{D}F(\mathbf{u}_n)\cdot(\mathbf{v}_h - \mathbf{u}_n) = -F(\mathbf{u}_n) \ \text{ on } \ \mathcal{N}_h.$$

This means that the new method is neither augmented Lagrange method nor penalty method associated with the current linear constraint. For convenience, we call the new method as *Newton-penalty method*.

5 Minimization problems with a penalized functional

For a given triangulation $\mathcal{T}_h$, we want to compute a sequence $\{\mathbf{u}_n\} \subset \mathbf{V}_{h,\mathbf{g}}$, such that $\{\mathbf{u}_n\}$ converges to a discrete harmonic map. As pointed in the last section, this sequence is determined by new minimization problems with penalized cost functional and linear constraints. The energy decreasing property of the sequence $\{\mathbf{u}_n\}$ will play a key role in its convergence. This section is devoted to designing the penalty term mentioned in the last section, which can guarantee the decreasing energy of the sequence $\{\mathbf{u}_n\}$.

5.1 Main result

Before defining the penalty term, we give some properties of the linear constraint spaces. Let $\mathbf{u}_n \in \mathbf{V}_{h,\mathbf{g}}$ such that $F(\mathbf{u}_n) = |\mathbf{u}_n|^2 - 1 \geq 0$ on $\mathcal{N}_h$, and define the linear constraint space

$$\begin{aligned} \mathbf{K}_n &= \{\mathbf{v} \in \mathbf{V}_{h,\mathbf{g}} : \ \mathbf{D}F(\mathbf{u}_n) \cdot \mathbf{v} = \mathbf{D}F(\mathbf{u}_n) \cdot \mathbf{u}_n - F(\mathbf{u}_n) \ on \ \mathcal{N}_h\} \\ &= \{\mathbf{v} \in \mathbf{V}_{h,\mathbf{g}} : \ 2\mathbf{u}_n \cdot \mathbf{v} = |\mathbf{u}_n|^2 + 1 \ on \ \mathcal{N}_h\}. \end{aligned} \tag{5.1}$$

The new minimization problem at $(n+1)$th step will be defined in the constraint space $\mathbf{K}_n$.

Lemma 5.1 *The constraint space $\mathbf{K}_n$ possesses the following properties:*

(a) For any $\mathbf{v} \in \mathbf{K}_n$, $F(\mathbf{v}) \geq 0$ on $\mathcal{N}_h$.

(b) For any $\mathbf{v} \in \mathbf{K}_n$, we have

$$\mathbf{D}F(\mathbf{v}) \cdot \mathbf{v} \geq 2 > 0 \quad and \quad |\mathbf{D}F(\mathbf{v})|^2 \geq 4 > 0 \quad on \ \mathcal{N}_h. \tag{5.2}$$

Proof. Since $\mathbf{v} \in \mathbf{K}_n$, we have

$$F(\mathbf{u}_n) + \mathbf{D}F(\mathbf{u}_n) \cdot (\mathbf{v} - \mathbf{u}_n) = 0 \ \ on \ \mathcal{N}_h.$$

Thus, we get by the generalized Taylor formula and (2.3)

$$\begin{aligned} F(\mathbf{v}) &= F(\mathbf{u}_n) + \mathbf{D}F(\mathbf{u}_n) \cdot (\mathbf{v} - \mathbf{u}_n) + \frac{1}{2}\mathbf{D}^2F(\xi)(\mathbf{v} - \mathbf{u}_n) \cdot (\mathbf{v} - \mathbf{u}_n) \\ &\geq |\mathbf{v} - \mathbf{u}_n|^2 \geq 0 \ \ \text{on } \mathcal{N}_h, \end{aligned}$$

which implies (a).

Note that $F(\mathbf{v}) = |\mathbf{v}|^2 - 1$, we deduce (b) by (2.2) and (a) directly.

□

It follows, by Lemma 5.1 (a), that $\pi_h F(\mathbf{v}) \geq 0$ for any $\mathbf{v} \in \mathbf{K}_n$. As pointed out in the last section, one can consider the regularization term $\gamma(h) \int_\Omega \pi_h F(\mathbf{v}_h) d\mathbf{x}$. But, in practice, the integration $\int_\Omega \pi_h F(\mathbf{v}_h) d\mathbf{x}$ needs to be computed by some quadrature formula. Thus, we do not consider $\gamma(h) \int_\Omega \pi_h F(\mathbf{v}_h) d\mathbf{x}$ itself, instead, we design a discrete version of $\gamma(h) \int_\Omega \pi_h F(\mathbf{v}_h) d\mathbf{x}$. To define the new penalty term, we need more notations.

Let φ_k denote the nodal basis function of the k-th node $p_k \in \mathcal{N}_h$ $(k = 1, \cdots, N)$, and set

$$a_{ij} = \int_\Omega \nabla\varphi_i \cdot \nabla\varphi_j dx \quad (i, j = 1, \cdots, N).$$

Given a node p_k, we use O_k to denote the set of other nodes that are neighbor to p_k. Define

$$\rho_{d,k} = \sum_{r \neq k} (a_{kr} + |a_{kr}|) \ \text{ and } \ \tilde{\rho}_{d,k} = -a_{kk} + \sum_{p_r \in O_k} (a_{kk}a_{rr})^{\frac{1}{2}}. \tag{5.3}$$

Proposition 5.1 For each k, we have

$$\tilde{\rho}_{d,k} \geq \rho_{d,k} \geq 0. \tag{5.4}$$

Proof. For every $r \neq k$ the sub-matrix

$$\begin{pmatrix} a_{rr} & a_{rk} \\ a_{kr} & a_{kk} \end{pmatrix}$$

is symmetric and positive semi-definite, so its determinent is nonnegative, i.e., $a_{rr}a_{kk} \geq a_{kr}^2$. Moreover, we have $\sum_{r=1}^{N} a_{kr} = 0$. Thus

$$\tilde{\rho}_{d,k} \geq \sum_{r \neq k} a_{kr} + \sum_{p_r \in O_k} |a_{kr}|,$$

which together with the fact that $a_{kr} = 0$ for $r \notin O_k$ gives (5.4).

□

Based on the notations described above, we define the local penalty parameter $\gamma_{d,k}$ by

$$\gamma_{d,k} = \frac{1}{2}(\kappa_1 \rho_{d,k} + \kappa_2 \tilde{\rho}_{d,k}).$$

Let $G_h^d(\cdot)$ be the linear functional defined by

$$G_h^d(\phi) = \sum_{k=1}^{N} \gamma_{d,k} \phi(p_k), \quad \phi \in C(\Omega).$$

Then the new penalty term is designed as $G_h^d(F(\mathbf{v}))$. Define

$$\mathcal{E}_h(\mathbf{v}) = \mathcal{E}(\mathbf{v}) + G_h^d(F(\mathbf{v})).$$

For convenience, we assume that the initial guess $\mathbf{u}_0$ satisfies $\pi_h F(\mathbf{u}_0) = 0$. Let K_n be the set defined in (5.1). The Newton-penalty method is to consider the following minimization problem: to find $\mathbf{u}_{n+1} \in \mathbf{K}_n$, such that

$$\mathcal{E}_h(\mathbf{u}_{n+1}) = \min_{\mathbf{v} \in \mathbf{K}_n} \mathcal{E}_h(\mathbf{v}). \tag{5.5}$$

Due to the strict convexity of $\mathcal{E}_h$ and the linearity of the constraint, the minimization problem (5.5) has a unique solution for every n. Moreover, the minimizer satisfies a saddle-point system, which can be solved easily (see Section 7). Note that, when the functional $G_h^d \equiv 0$, the method described by (5.5) becomes Newton linearization method.

Theorem 5.1 *(Decreasing energy) Let* $\mathbf{u}_{n+1}$ *be the solution of (5.5). Then* $F(\mathbf{u}_{n+1}) = |\mathbf{u}_{n+1}|^2 - 1 \geq 0$ *on* $\mathcal{N}_h$ *and*

$$\mathcal{E}_h(\mathbf{u}_{n+1}) \leq \mathcal{E}_h(\mathbf{u}_n) \leq \cdots \leq \mathcal{E}_h(\mathbf{u}_0) = \mathcal{E}(\mathbf{u}_0). \tag{5.6}$$

The above theorem will be proved in the next subsection. As we will see, this theorem is the key result of the paper. Based on this result, we can prove that the sequence $\{\mathbf{u}_n\}$ converges to a discrete harmonic map in some sense (see Section 6 below). A natural question is whether the simple Newton linearization method may possess the decreasing energy? The following result gives a reply to the question.

Corollary 5.1. Assume that $\kappa_2 = 0$. If $a_{kr} \leq 0$ for every $r \neq k$ $(k = 1, \cdots, N)$, then $\mathcal{E}(\mathbf{u}_{n+1}) \leq \mathcal{E}(\mathbf{u}_n)$.

Proof. Since $a_{kr} \leq 0$, which yields $|a_{kr}| = -a_{kr}$, we deduce $\rho_{d,k} = 0$. Then, we get $\gamma_{d,k} = 0$ for each k by the assumption $\kappa_2 = 0$. Thus, $G_h^d \equiv 0$, which implies that $\mathcal{E}_h(\mathbf{u}_n) = \mathcal{E}(\mathbf{u}_n)$. The desired result follows directly by Theorem 5.1.

□

Remark 5.1 *Assume that $\kappa_2 = 0$. If $\mathcal{T}_h$ is an acute triangulation (see [3] and [18] for details), then $a_{kr} \leq 0$ for every $r \neq k$ $(k = 1, \cdots, N)$. Then, the conditions in* **Corollary 5.1** *are satisfied. This means that the simple Newton linearization method possesses the decreasing energy for the case with $\kappa_2 = 0$, provided that the triangulation $\mathcal{T}_h$ has some particular structure.*

Remark 5.2 *In the design of the penalty term $G_h^d(F(\mathbf{v}))$, we need to define the parameters $\gamma_{d,k}$ carefully. For smaller $\gamma_{d,k}$, Theorem 5.1 may be not valid. For larger $\gamma_{d,k}$, Theorem 5.1 is still valid, but the sequence $\{\mathbf{u}_n\}$ converges more slowly.*

5.2 Analysis

In the analysis of Theorem 5.1, we shall associate $\mathbf{u}_n$ with a special function $\bar{\mathbf{u}}_n$ which is defined below.

By induction and by Lemma 5.1 (a), we have $F(\mathbf{u}_n) = |\mathbf{u}_n|^2 - 1 \geq 0$ on $\mathcal{N}_h$. By (5.2) we get

$$\mathbf{D}F(\mathbf{u}_n) \cdot \mathbf{u}_n \geq 2 > 0 \quad \text{on } \mathcal{N}_h \text{ for all } n.$$

Thus we can define $\alpha_n \in \mathbf{V}_{h,0}$ and $\bar{\mathbf{u}}_n \in \mathbf{V}_{h,\mathbf{g}}$ by

$$\alpha_n = \frac{F(\mathbf{u}_n)}{\mathbf{D}F(\mathbf{u}_n) \cdot \mathbf{u}_n} = \frac{|\mathbf{u}_n|^2 - 1}{2|\mathbf{u}_n|^2} \quad \text{and} \quad \bar{\mathbf{u}}_n = (1 - \alpha_n)u_n \quad \text{on } \mathcal{N}_h.$$

Namely,

$$\alpha_n = \pi_h(\frac{|\mathbf{u}_n|^2 - 1}{2|\mathbf{u}_n|^2}) \quad \text{and} \quad \bar{\mathbf{u}}_n = \pi_h((1 - \alpha_n)\mathbf{u}_n). \tag{5.7}$$

It is clear that

$$0 \leq \alpha_n < \frac{1}{2} < 1 \quad (\text{on } \mathcal{N}_h).$$

Moreover, $\alpha_n(p) = 0$ if and only if $F(\mathbf{u}_n(p)) = 0$ $(p \in \mathcal{N}_h)$.

It is easy to verify that the linearized constraint set $\mathbf{K}_n$ is the tangential plane at $\bar{\mathbf{u}}_n$ for the following level set:

$$\{\mathbf{v} \in \mathbf{V}_{h,\mathbf{g}} : \quad F(\mathbf{v}) = F(\bar{\mathbf{u}}_n) \text{ on } \mathcal{N}_h\}.$$

In order to prove Theorem 5.1, we need to estimate $\mathcal{E}(\bar{\mathbf{u}}_n) - \mathcal{E}(\mathbf{u}_n)$ carefully. To this end, we first derive some auxiliary results.

For convenience, we define $\alpha_{n,k} = \alpha_n(p_k)$ and $\mathbf{u}_{n,k} = \mathbf{u}_n(p_k)$ $(k = 1, \cdots, N)$ in the rest of this subsection.

Lemma 5.2 *For $i \neq j$, set*

$$\begin{aligned} r_{ij} &= [(2\alpha_{n,i} - \alpha_{n,i}^2) - (\alpha_{n,i} + \alpha_{n,j} - \alpha_{n,i}\alpha_{n,j})]|\mathbf{u}_{n,i}|^2 \\ &+ [(2\alpha_{n,j} - \alpha_{n,j}^2) - (\alpha_{n,i} + \alpha_{n,j} - \alpha_{n,i}\alpha_{n,j})]|\mathbf{u}_{n,j}|^2. \end{aligned}$$

Then,

$$r_{ij} \geq 0. \tag{5.8}$$

Proof. By the direct calculation, we get

$$r_{ij} = (\alpha_{n,i} - \alpha_{n,j})(1 - \alpha_{n,i})|\mathbf{u}_{n,i}|^2 + (\alpha_{n,j} - \alpha_{n,i})(1 - \alpha_{n,j})|\mathbf{u}_{n,j}|^2. \tag{5.9}$$

By the definition of $\alpha_{n,k}$, we have

$$1 - \alpha_{n,k} = \frac{|\mathbf{u}_{n,k}|^2 + 1}{2|\mathbf{u}_{n,k}|^2} \quad (k = i, j) \tag{5.10}$$

and

$$\alpha_{n,i} - \alpha_{n,j} = \frac{1}{2}\big(\frac{1}{|\mathbf{u}_{n,j}|^2} - \frac{1}{|\mathbf{u}_{n,i}|^2}\big), \quad \alpha_{n,j} - \alpha_{n,i} = \frac{1}{2}\big(\frac{1}{|\mathbf{u}_{n,i}|^2} - \frac{1}{|\mathbf{u}_{n,j}|^2}\big). \tag{5.11}$$

It follows by (5.10) that

$$(1 - \alpha_{n,k})|\mathbf{u}_{n,k}|^2 = \frac{1}{2}(|\mathbf{u}_{n,k}|^2 + 1) \quad (k = i, j).$$

Using (5.9), together with the above equality and (5.11), yields

$$r_{ij} = \frac{1}{4}\big(\frac{|\mathbf{u}_{n,i}|^2}{|\mathbf{u}_{n,j}|^2} - 1\big) + \frac{1}{4}\big(\frac{|\mathbf{u}_{n,j}|^2}{|\mathbf{u}_{n,i}|^2} - 1\big) = \frac{1}{4}\big(\frac{|\mathbf{u}_{n,i}|}{|\mathbf{u}_{n,j}|} - \frac{|\mathbf{u}_{n,j}|}{|\mathbf{u}_{n,i}|}\big)^2 \geq 0.$$

□

The following result gives a simple equality, which will be used later repeatedly.

Lemma 5.3 *Let Λ be a set of finite ordered indices k. Let η_k denote a number depending on the index $k \in \Lambda$, and let τ_{ij} denote a number depending on two different indices i, j in Λ. Then,*

$$\sum_k \sum_{r \neq k} \tau_{kr}\eta_k = \sum_{i<j} (\tau_{ij}\eta_i + \tau_{ji}\eta_j). \tag{5.12}$$

In particular, when $\tau_{ij} = \tau_{ji}$, we have

$$\sum_k \sum_{r \neq k} \tau_{kr}\eta_k = \sum_{i<j} \tau_{ij}(\eta_i + \eta_j). \tag{5.13}$$

Proof. The equality (5.12) can be derived by direct calculations.

□

Lemma 5.4 *Let α_n be defined by (5.7), and let $\rho_{d,k}$ be defined by (5.3). Then,*

$$-2\int_\Omega \nabla \mathbf{u}_n : \nabla \pi_h(\alpha_n \mathbf{u}_n) d\mathbf{x} + \int_\Omega |\nabla \pi_h(\alpha_n \mathbf{u}_n)|^2 d\mathbf{x} \leq \sum_{k=1}^N \rho_{d,k}[2\alpha_n(p_k) - \alpha_n^2(p_k)]|\mathbf{u}_n(p_k)|^2, \tag{5.14}$$

where $\rho_{d,k}$ denotes the parameter defined by (5.3).

Proof. Let a_{ij} be the real numbers defined in the last subsection. It can be verified that

$$\int_\Omega \nabla \mathbf{u}_n : \nabla \pi_h(\alpha_n \mathbf{u}_n) d\mathbf{x} \;=\; \int_\Omega (\sum_{k=1}^N \mathbf{u}_{n,k}(\nabla \varphi_k)^t) : (\sum_{k=1}^N \alpha_{n,k}\mathbf{u}_{n,k}(\nabla \varphi_k)^t) d\mathbf{x}$$

$$= \sum_{k=1}^{N} a_{kk}\alpha_{n,k}|\mathbf{u}_{n,k}|^2 + \sum_{i<j} a_{ij}(\alpha_{n,i}+\alpha_{n,j})\mathbf{u}_{n,i}\cdot\mathbf{u}_{n,j},$$

and

$$\begin{aligned}\int_\Omega |\nabla\pi_h(\alpha_n\mathbf{u}_n)|^2 d\mathbf{x} &= \int_\Omega |\sum_{k=1}^{N}\alpha_{n,k}\mathbf{u}_{n,k}(\nabla\varphi_k)^t|^2 d\mathbf{x}\\ &= \sum_{k=1}^{N} a_{kk}\alpha_{n,k}^2|\mathbf{u}_{n,k}|^2 + 2\sum_{i<j} a_{ij}\alpha_{n,i}\alpha_{n,j}\mathbf{u}_{n,i}\cdot\mathbf{u}_{n,j}.\end{aligned}$$

Then,

$$\begin{aligned}&-2\int_\Omega \nabla\mathbf{u}_n : \nabla\pi_h(\alpha_n\mathbf{u}_n)d\mathbf{x} + \int_\Omega |\nabla\pi_h(\alpha_n\mathbf{u}_n)|^2 d\mathbf{x}\\ =& -\sum_{k=1}^{N} a_{kk}(2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2 - 2\sum_{i<j} a_{ij}(\alpha_{n,i}+\alpha_{n,j}-\alpha_{n,i}\alpha_{n,j})\mathbf{u}_{n,i}\cdot\mathbf{u}_{n,j}\\ \le& -\sum_{k=1}^{N} a_{kk}(2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2 + 2\sum_{i<j} |a_{ij}|(\alpha_{n,i}+\alpha_{n,j}-\alpha_{n,i}\alpha_{n,j})|\mathbf{u}_{n,i}|\cdot|\mathbf{u}_{n,j}|.\end{aligned} \tag{5.15}$$

Here we have used the fact that (since $\alpha_{n,k}\in[0,\frac{1}{2}]$)

$$\alpha_{n,i}+\alpha_{n,j}-\alpha_{n,i}\alpha_{n,j} \ge 0. \tag{5.16}$$

Note that

$$a_{kk} = -\sum_{r\neq k} a_{kr} \quad (k=1,\cdots,N),$$

we have

$$\begin{aligned}-\sum_{k=1}^{N} a_{kk}(2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2 &= \sum_{k=1}^{N}\sum_{r\neq k} a_{kr}(2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2\\ &= \sum_{k=1}^{N}\rho_{d,k}(2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2\\ &- \sum_{k=1}^{N}\sum_{r\neq k}|a_{kr}|(2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2.\end{aligned} \tag{5.17}$$

Plugging (5.17) in (5.15), leads to

$$-2\int_\Omega \nabla\mathbf{u}_n : \nabla\pi_h(\alpha_n\mathbf{u}_n)d\mathbf{x} + \int_\Omega |\nabla\pi_h(\alpha_n\mathbf{u}_n)|^2 d\mathbf{x} = \sum_{k=1}^{N}\rho_{d,k}(2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2 + R_N, \tag{5.18}$$

where

$$R_N = -\sum_{k=1}^{N}\sum_{r\neq k}|a_{kr}|(2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2 + 2\sum_{i<j}|a_{ij}|(\alpha_{n,i}+\alpha_{n,j}-\alpha_{n,i}\alpha_{n,j})|\mathbf{u}_{n,i}|\cdot|\mathbf{u}_{n,j}|.$$

It suffices to estimate R_N. Note that $a_{kr} = a_{rk}$, we get by (5.13)

$$\sum_{k=1}^{N}\sum_{r\neq k}|a_{kr}|(2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2 = \sum_{i<j}|a_{ij}|\big\{(2\alpha_{n,i}-\alpha_{n,i}^2)|\mathbf{u}_{n,i}|^2 + (2\alpha_{n,j}-\alpha_{n,j}^2)|\mathbf{u}_{n,j}|^2\big\}.$$

Moreover, we have

$$2|\mathbf{u}_{n,i}|\cdot|\mathbf{u}_{n,j}| = |\mathbf{u}_{n,i}|^2 + |\mathbf{u}_{n,j}|^2 - (|\mathbf{u}_{n,i}| - |\mathbf{u}_{n,j}|)^2.$$

Using the above two equalities and (5.16), we deduce

$$R_N = -\sum_{i<j}|a_{ij}|r_{ij} - \sum_{i<j}|a_{ij}|(\alpha_{n,i}+\alpha_{n,j}-\alpha_{n,i}\alpha_{n,j})(|\mathbf{u}_{n,i}| - |\mathbf{u}_{n,j}|)^2 \leq -\sum_{i<j}|a_{ij}|r_{ij}, \quad (5.19)$$

where r_{ij} is defined by Lemma 5.2. Combining (5.19) and (5.8), yields $R_N \leq 0$. Then the inequality (5.14) follows by (5.18).

□

Lemma 5.5 *Let e be an element. For the basis φ_k associated with a vertex p_k of the element e, set*

$$a_{kk}^e = \int_e |\nabla\varphi_k|^2 dx = \|\nabla\varphi_k\|_{0,e}^2.$$

Let p_i and p_j be two different vertices of the element e, define

$$\begin{aligned} r_{ij}^e = &(\tfrac{a_{jj}^e}{a_{ii}^e})^{\frac{1}{2}}[(\alpha_{n,i}^2 - 2\alpha_{n,i}) + (\alpha_{n,i}+\alpha_{n,j}-\alpha_{n,i}\alpha_{n,j})]\|\mathbf{u}_{n,i}\times\nabla\varphi_i\|_{0,e}^2 \\ &+(\tfrac{a_{ii}^e}{a_{jj}^e})^{\frac{1}{2}}[(\alpha_{n,j}^2 - 2\alpha_{n,j}) + (\alpha_{n,i}+\alpha_{n,j}-\alpha_{n,i}\alpha_{n,j})]\|\mathbf{u}_{n,j}\times\nabla\varphi_j\|_{0,e}^2. \end{aligned}$$

Then,

$$r_{ij}^e \leq 0. \quad (5.20)$$

Proof. Since each φ_k is linear on the element e, each complement of the vector $\nabla\varphi_k|_e$ is a number independent of the coordinate variables. It is known that

$$\mathbf{u}_{n,k}\times\nabla\varphi_k|_e = \mathbf{0}$$

if $\mathbf{u}_{n,k}$ is parallel to $\nabla\varphi_k|_e$. Thus, without loss of generality, we assume that

$$\mathbf{u}_{n,k}\perp\nabla\varphi_k|_e. \quad (5.21)$$

Otherwise, $\mathbf{u}_{n,k}$ can be decomposed into a sum of two vectors, one of which is parallel to $\nabla\varphi_k|_e$, and another one is orthogonal to $\nabla\varphi_k|_e$.

By direct calculations, and using the assumption (5.21), we obtain

$$\|\mathbf{u}_{n,k}\times\nabla\varphi_k\|_{0,e}^2 = |\mathbf{u}_{n,k}|^2\cdot\|\nabla\varphi_k\|_{0,e}^2. \quad (5.22)$$

Then,

$$(\frac{a_{jj}^e}{a_{ii}^e})^{\frac{1}{2}}\|\mathbf{u}_{n,i}\times\nabla\varphi_i\|_{0,e}^2 = (a_{ii}^e a_{jj}^e)^{\frac{1}{2}}|\mathbf{u}_{n,i}|^2 \quad \text{and} \quad (\frac{a_{ii}^e}{a_{jj}^e})^{\frac{1}{2}}\|\mathbf{u}_{n,j}\times\nabla\varphi_j\|_{0,e}^2 = (a_{ii}^e a_{jj}^e)^{\frac{1}{2}}|\mathbf{u}_{n,j}|^2.$$

By the definitions of r_{ij} and r_{ij}^e, we have

$$r_{ij}^e = -(a_{ii}^e a_{jj}^e)^{\frac{1}{2}} r_{ij}.$$

Now, the inequality (5.20) is a direct result of (5.8).

□

Lemma 5.6 *Let α_n be defined by (5.7). Then,*

$$\begin{aligned} -2\int_\Omega(\nabla\times\mathbf{u}_n)\cdot(\nabla\times\pi_h(\alpha_n\mathbf{u}_n))d\mathbf{x}+\int_\Omega|\nabla\times\pi_h(\alpha_n\mathbf{u}_n)|^2d\mathbf{x}\\ \le\sum_{k=1}^{N}\tilde{\rho}_{d,k}[2\alpha_n(p_k)-\alpha_n^2(p_k)]|\mathbf{u}_n(p_k)|^2, \end{aligned} \tag{5.23}$$

where $\tilde{\rho}_{d,k}$ is defined by (5.3).

Proof. We use a similar idea with the proof of Lemma 5.3. But, there are some necessary changes, since the underlying integrations involve a different operator. We need to estimate the integrations in elements instead of in the global domain Ω. For an element e, let $\mathcal{V}_e$ denote the set of $d+1$ vertices of e, and set

$$\Lambda_e=\{(i,j):\ i<j,\ p_i\ \text{ and }\ p_j\ \text{ are two vertices of } e\}.$$

It is easy to see that

$$\begin{aligned} &\int_e(\nabla\times\mathbf{u}_n)\cdot(\nabla\times\pi_h(\alpha_n\mathbf{u}_n))d\mathbf{x}=\int_e(\sum_{k=1}^{N}\mathbf{u}_{n,k}\times\nabla\varphi_k)\cdot(\sum_{k=1}^{N}\alpha_{n,k}\mathbf{u}_{n,k}\times\nabla\varphi_k)d\mathbf{x}\\ &=\sum_{p_k\in\mathcal{V}_e}\alpha_{n,k}\int_e|\mathbf{u}_{n,k}\times\nabla\varphi_k|^2d\mathbf{x}+\sum_{(i,j)\in\Lambda_e}(\alpha_{n,i}+\alpha_{n,j})\int_e(\mathbf{u}_{n,i}\times\nabla\varphi_i)\cdot(\mathbf{u}_{n,j}\times\nabla\varphi_j)d\mathbf{x}, \end{aligned}$$

and

$$\begin{aligned} &\int_e|\nabla\times\pi_h(\alpha_n\mathbf{u}_n)|^2d\mathbf{x}=\int_e|\sum_{k=1}^{N}\alpha_{n,k}\mathbf{u}_{n,k}\times\nabla\varphi_k|^2d\mathbf{x}\\ &=\sum_{p_k\in\mathcal{V}_e}\alpha_{n,k}^2\int_e|\mathbf{u}_{n,k}\times\nabla\varphi_k|^2d\mathbf{x}+2\sum_{(i,j)\in\Lambda_e}\alpha_{n,i}\alpha_{n,j}\int_e(\mathbf{u}_{n,i}\times\nabla\varphi_i)\cdot(\mathbf{u}_{n,j}\times\nabla\varphi_j)d\mathbf{x}. \end{aligned}$$

Thus,

$$\begin{aligned} &-2\int_e(\nabla\times\mathbf{u}_n)\cdot(\nabla\times\pi_h(\alpha_n\mathbf{u}_n))d\mathbf{x}+\int_e|\nabla\times\pi_h(\alpha_n\mathbf{u}_n)|^2d\mathbf{x}\\ &=\sum_{p_k\in\mathcal{V}_e}(\alpha_{n,k}^2-2\alpha_{n,k})\int_e|\mathbf{u}_{n,k}\times\nabla\varphi_k|^2d\mathbf{x}\\ &+2\sum_{(i,j)\in\Lambda_e}(\alpha_{n,i}\alpha_{n,j}-\alpha_{n,i}-\alpha_{n,j})\int_e(\mathbf{u}_{n,i}\times\nabla\varphi_i)\cdot(\mathbf{u}_{n,j}\times\nabla\varphi_j)d\mathbf{x}\\ &\le\sum_{p_k\in\mathcal{V}_e}(\alpha_{n,k}^2-2\alpha_{n,k})\|\mathbf{u}_{n,k}\times\nabla\varphi_k\|_{0,e}^2\\ &+2\sum_{(i,j)\in\Lambda_e}(\alpha_{n,i}+\alpha_{n,j}-\alpha_{n,i}\alpha_{n,j})\|\mathbf{u}_{n,i}\times\nabla\varphi_i\|_{0,e}\cdot\|\mathbf{u}_{n,j}\times\nabla\varphi_j\|_{0,e}. \end{aligned} \tag{5.24}$$

Let O_k be defined as in Subsection 5.1, and set

$$\delta_{d,k}^e=\sum_{p_r\in O_k}(\frac{a_{rr}^e}{a_{kk}^e})^{\frac{1}{2}}\quad(p_k\in\mathcal{V}_e).$$

Then,

$$\begin{aligned}\sum_{p_k\in\mathcal{V}_e}(\alpha_{n,k}^2-2\alpha_{n,k})\|\mathbf{u}_{n,k}\times\nabla\varphi_k\|_{0,e}^2 &= \sum_{p_k\in\mathcal{V}_e}\delta_{d,k}^e(\alpha_{n,k}^2-2\alpha_{n,k})\|\mathbf{u}_{n,k}\times\nabla\varphi_k\|_{0,e}^2\\ &+ \sum_{p_k\in\mathcal{V}_e}(1-\delta_{d,k}^e)(\alpha_{n,k}^2-2\alpha_{n,k})\|\mathbf{u}_{n,k}\times\nabla\varphi_k\|_{0,e}^2.\end{aligned}$$

Applying (5.12) to the first sum in the right hand side of the above equality, we further have

$$\begin{aligned}\sum_{p_k\in\mathcal{V}_e}(\alpha_{n,k}^2-2\alpha_{n,k})\|\mathbf{u}_{n,k}\times\nabla\varphi_k\|_{0,e}^2 &= \sum_{(i,j)\in\Lambda_e}\big\{(\frac{a_{jj}^e}{a_{ii}^e})^{\frac{1}{2}}(\alpha_{n,i}^2-2\alpha_{n,i})\|\mathbf{u}_{n,i}\times\nabla\varphi_i\|_{0,e}^2\\ &+ (\frac{a_{ii}^e}{a_{jj}^e})^{\frac{1}{2}}(\alpha_{n,j}^2-2\alpha_{n,j})\|\mathbf{u}_{n,j}\times\nabla\varphi_j\|_{0,e}^2\big\}\\ &+ \sum_{p_k\in\mathcal{V}_e}(\delta_{d,k}^e-1)(2\alpha_{n,k}-\alpha_{n,k}^2)\|\mathbf{u}_{n,k}\times\nabla\varphi_k\|_{0,e}^2.\end{aligned}$$

Combining this with (5.24), leads to

$$\begin{aligned}&-2\textstyle\int_e(\nabla\times\mathbf{u}_n)\cdot(\nabla\times\pi_h(\alpha_n\mathbf{u}_n))d\mathbf{x}+\int_e|\nabla\times\pi_h(\alpha_n\mathbf{u}_n)|^2d\mathbf{x}\\ &\le R_N^e+\sum_{p_k\in\mathcal{V}_e}(\delta_{d,k}^e-1)(2\alpha_{n,k}-\alpha_{n,k}^2)\|\mathbf{u}_{n,k}\times\nabla\varphi_k\|_{0,e}^2,\end{aligned}\tag{5.25}$$

with

$$\begin{aligned}R_N^e=\sum_{(i,j)\in\Lambda_e}&\big\{(\tfrac{a_{jj}^e}{a_{ii}^e})^{\frac{1}{2}}(\alpha_{n,i}^2-2\alpha_{n,i})\|\mathbf{u}_{n,i}\times\nabla\varphi_i\|_{0,e}^2+(\tfrac{a_{ii}^e}{a_{jj}^e})^{\frac{1}{2}}(\alpha_{n,j}^2-2\alpha_{n,j})\|\mathbf{u}_{n,j}\times\nabla\varphi_j\|_{0,e}^2\\ &-2(\alpha_{n,i}+\alpha_{n,j}-\alpha_{n,i}\alpha_{n,j})\|\mathbf{u}_{n,i}\times\nabla\varphi_i\|_{0,e}\cdot\|\mathbf{u}_{n,j}\times\nabla\varphi_j\|_{0,e}\big\}.\end{aligned}$$

It suffices to estimate R_N^e. By direct calculations, we deduce

$$\begin{aligned}R_N^e=\sum_{(i,j)\in\Lambda_e}&\big\{r_{ij}^e-(\alpha_{n,i}+\alpha_{n,j}-\alpha_{n,i}\alpha_{n,j})[(\tfrac{a_{jj}^e}{a_{ii}^e})^{\frac{1}{4}}\|\mathbf{u}_{n,i}\times\nabla\varphi_i\|_{0,e}-(\tfrac{a_{ii}^e}{a_{jj}^e})^{\frac{1}{4}}\|\mathbf{u}_{n,j}\times\nabla\varphi_j\|_{0,e}]^2\big\}\\ &\le\sum_{(i,j)\in\Lambda_e}r_{ij}^e.\end{aligned}$$

This, together with Lemma 5.5 and (5.16), leads to $R_N^e\le 0$. We further get by (5.25) and (5.22)

$$\begin{aligned}&-2\int_e(\nabla\times\mathbf{u}_n)\cdot(\nabla\times\pi_h(\alpha_n\mathbf{u}_n))d\mathbf{x}+\int_e|\nabla\times\pi_h(\alpha_n\mathbf{u}_n)|^2d\mathbf{x}\\ &\le\sum_{p_k\in\mathcal{V}_e}(\delta_{d,k}^e-1)(2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2\cdot\|\nabla\varphi_k\|_{0,e}^2\\ &=\sum_{p_k\in\mathcal{V}_e}[\sum_{p_r\in O_k}(a_{rr}^e a_{kk}^e)^{\frac{1}{2}}-a_{kk}^e](2\alpha_{n,k}-\alpha_{n,k}^2)|\mathbf{u}_{n,k}|^2.\end{aligned}\tag{5.26}$$

Note that, when the assumption (5.21) does not hold and we have to make the orthogonal decomposition for $\mathbf{u}_{n,k}$, the final equality at the above would become an inequality, since the module of the orthogonal part of $\mathbf{u}_{n,k}$ is not larger than the module of $\mathbf{u}_{n,k}$ itself.

For a node p_k, let Ξ_k denote the set of elements that contain p_k as a vertex. Then,

$$\sum_e\sum_{p_k\in\mathcal{V}_e}=\sum_{k=1}^N\sum_{e\in\Xi_k}.$$

Besides, we have

$$\sum_{e\in\Xi_k}(a^e_{rr}a^e_{kk})^{\frac{1}{2}} \le (\sum_{e\in\Xi_k}a^e_{rr})^{\frac{1}{2}}(\sum_{e\in\Xi_k}a^e_{kk})^{\frac{1}{2}}.$$

Now the inequality (5.23) follows by summing (5.26) over all elements e.

□

Proof of Theorem 5.1. By the definitions of $\mathcal{E}$ and $\bar{\mathbf{u}}_n$, we can deduce that

$$\begin{aligned}\mathcal{E}(\bar{\mathbf{u}}_n) &= \mathcal{E}(\mathbf{u}_n)+\frac{\kappa_1}{2}[-2\int_\Omega \nabla\mathbf{u}_n:\nabla\pi_h(\alpha_n\mathbf{u}_n)d\mathbf{x}+\int_\Omega|\nabla\pi_h(\alpha_n\mathbf{u}_n)|^2d\mathbf{x}]\\ &+ \frac{\kappa_2}{2}[-2\int_\Omega(\nabla\times\mathbf{u}_n)\cdot(\nabla\times\pi_h(\alpha_n\mathbf{u}_n))d\mathbf{x}+\int_\Omega|\nabla\times\pi_h(\alpha_n\mathbf{u}_n)|^2d\mathbf{x}]\big\}. \quad (5.27)\end{aligned}$$

This, together with Lemma 5.4 and Lemma 5.6, leads to

$$\mathcal{E}(\bar{\mathbf{u}}_n)\le\mathcal{E}(\mathbf{u}_n)+\frac{1}{2}\sum_{k=1}^N(\kappa_1\rho_{d,k}+\kappa_2\tilde{\rho}_{d,k})[2\alpha_n(p_k)-\alpha_n^2(p_k)]|\mathbf{u}_n(p_k)|^2.$$

By the definition of $\gamma_{d,k}$, we further get

$$\mathcal{E}(\bar{\mathbf{u}}_n)\le\mathcal{E}(\mathbf{u}_n)+\sum_{k=1}^N\gamma_{d,k}[2\alpha_n(p_k)-\alpha_n^2(p_k)]|\mathbf{u}_n(p_k|^2. \quad (5.28)$$

On the other hand, we have by the definitions of F and $\bar{\mathbf{u}}_n$

$$F(\bar{\mathbf{u}}_n)=(1-\alpha_n)^2|\mathbf{u}_n|^2-1 \ \ on\ \mathcal{N}_h.$$

Thus, by the definition of the functional $G_h^d(\cdot)$, we deduce

$$\begin{aligned}G_h^d(F(\bar{\mathbf{u}}_n)) &= \sum_{k=1}^N\gamma_{d,k}F(\bar{\mathbf{u}}_n(p_k))\\ &= \sum_{k=1}^N\gamma_{d,k}[(1-2\alpha_n(p_k)+\alpha_n^2(p_k))|\mathbf{u}_n(p_k)|^2-1].\end{aligned}$$

Together with (5.28), this leads to

$$\mathcal{E}(\bar{\mathbf{u}}_n)+G_h^d(F(\bar{\mathbf{u}}_n))\le\mathcal{E}(\mathbf{u}_n)+\sum_{k=1}^N\gamma_{d,k}[|\mathbf{u}_n(p_k)|^2-1]=\mathcal{E}(\mathbf{u}_n)+G_h^d(F(\mathbf{u}_n)).$$

This means that

$$\mathcal{E}_h(\bar{\mathbf{u}}_n)\le\mathcal{E}_h(\mathbf{u}_n). \quad (5.29)$$

It is easy to see that $\bar{\mathbf{u}}_n\in\mathbf{K}_n$. Since $\mathbf{u}_{n+1}$ is the minimizer of (5.5), we get by (5.29)

$$\mathcal{E}_h(\mathbf{u}_{n+1})\le\mathcal{E}_h(\bar{\mathbf{u}}_n)\le\mathcal{E}_h(\mathbf{u}_n),\quad n=0,1,\cdots.$$

□

6 Convergence of the solution sequence $\{\mathbf{u}_n\}$

This section is devoted to proving convergence of the sequence $\{\mathbf{u}_n\}$ generated by the minimization problem (5.5). To this end, we introduce an auxiliary minimization problem.

For the sequence $\mathbf{u}_n$ generated by (5.5), define

$$\bar{\mathbf{K}}_n = \{\mathbf{v} \in \mathbf{V}_{h,0} : \ \mathbf{D}F(\mathbf{u}_n) \cdot \mathbf{v} = 0 \ \ on \ \mathcal{N}_h\}.$$

Let $\bar{\mathbf{u}}_n$ be defined by (5.7). Consider the minimization problem: Find $\mathbf{w}_n \in \bar{\mathbf{K}}_n$, such that

$$\mathcal{E}_h(\bar{\mathbf{u}}_n + \mathbf{w}_n) = \min_{\mathbf{v} \in \bar{\mathbf{K}}_n} \mathcal{E}_h(\bar{\mathbf{u}}_n + \mathbf{v}). \tag{6.1}$$

It is easy to see that $\mathbf{u}_{n+1} = \bar{\mathbf{u}}_n + \mathbf{w}_n$.

Lemma 6.1 *Let $\mathbf{w}_n$ be the sequence generated by (6.1). Then, we have*

$$\|\nabla \mathbf{w}_n\|_{0,\Omega}^2 + G_h^d(|\mathbf{w}_n|^2) \to 0 \quad when \ n \to +\infty. \tag{6.2}$$

In particular, we have $\mathbf{w}_n(p_k) \to 0$ for each node p_k when $n \to +\infty$ (for a fixed h).

Proof. It follows by (6.1) that

$$\mathcal{E}_h'(\bar{\mathbf{u}}_n + \mathbf{w}_n)\mathbf{w}_n = 0. \tag{6.3}$$

Using the generalized Taylor formula, we get

$$\mathcal{E}_h(\bar{\mathbf{u}}_n) = \mathcal{E}_h(\bar{\mathbf{u}}_n + \mathbf{w}_n) + \mathcal{E}_h'(\bar{\mathbf{u}}_n + \mathbf{w}_n)\mathbf{w}_n + \frac{1}{2}\mathcal{E}_h''(\xi_n)\mathbf{w}_n \cdot \mathbf{w}_n.$$

This, together with (6.3), leads to

$$\mathcal{E}_h(\bar{\mathbf{u}}_n) = \mathcal{E}_h(\bar{\mathbf{u}}_n + \mathbf{w}_n) + \frac{1}{2}\mathcal{E}_h''(\xi_n)\mathbf{w}_n \cdot \mathbf{w}_n.$$

Thus,

$$\frac{1}{2}\mathcal{E}_h''(\xi_n)\mathbf{w}_n \cdot \mathbf{w}_n = \mathcal{E}_h(\bar{\mathbf{u}}_n) - \mathcal{E}_h(\bar{\mathbf{u}}_n + \mathbf{w}_n) \le \mathcal{E}_h(\bar{\mathbf{u}}_n) - \mathcal{E}_h(\bar{\mathbf{u}}_{n+1}). \tag{6.4}$$

Here we have used the fact that $\mathcal{E}_h(\bar{\mathbf{u}}_{n+1}) \le \mathcal{E}_h(\mathbf{u}_{n+1})$ from (5.29). It is easy to see that

$$\mathcal{E}_h''(\xi_n)\mathbf{w}_n \cdot \mathbf{w}_n = \mathcal{E}''(\xi_n)\mathbf{w}_n \cdot \mathbf{w}_n + G_h^d(\mathbf{D}^2F(\xi_n)\mathbf{w}_n \cdot \mathbf{w}_n).$$

Hence we get by (2.4) and (2.3)

$$\mathcal{E}_h''(\xi_n)\mathbf{w}_n \cdot \mathbf{w}_n \ge c[\|\nabla \mathbf{w}_n\|_{0,\Omega}^2 + G_h^d(|\mathbf{w}_n|^2)].$$

Combining this with (6.4), we deduce

$$\|\nabla \mathbf{w}_n\|_{0,\Omega}^2 + G_h^d(|\mathbf{w}_n|^2) \le C(\mathcal{E}_h(\bar{\mathbf{u}}_n) - \mathcal{E}_h(\bar{\mathbf{u}}_{n+1})).$$

Then, we have for any positive integer M

$$\sum_{n=1}^{M}[\|\nabla \mathbf{w}_n\|_{0,\Omega}^2 + G_h^d(|\mathbf{w}_n|^2)] \le C\mathcal{E}_h(\bar{\mathbf{u}}_0).$$

This implies (6.2).

Since $\mathbf{w}_n|_{\partial\Omega} = \mathbf{0}$, we get by (6.2) and Poincare inequality

$$\|\mathbf{w}_n\|_{0,\Omega}^2 \to 0 \ \text{ when } n \to \infty.$$

Thus, $\mathbf{w}_n(p_k) \to 0$ for each node p_k when $n \to +\infty$ (for a fixed h).

□

The following result indicates that the solution $\mathbf{u}_n$ approximately satisfies the original nonlinear constraint.

Lemma 6.2 *Let $\mathbf{u}_{n+1}$ be the sequence generated by (5.5). Then we have*

$$\lim_{n\to+\infty} F(\mathbf{u}_n(p)) = 0,$$

which is uniform for $p \in \mathcal{N}_h$ (for a fixed h).

Proof. The proof of this lemma is a bit technical. We will prove it by three steps.

Step 1. A recursive relation.

Let z denote a node in $\mathcal{N}_h$, and set

$$\alpha^n = \frac{F(\mathbf{u}_n(z))}{\mathbf{D}F(\mathbf{u}_n(z)) \cdot \mathbf{u}_n(z)} = \frac{|\mathbf{u}_n(z)|^2 - 1}{2|\mathbf{u}_n(z)|^2}.$$

Then,

$$\mathbf{u}_{n+1}(z) = \bar{\mathbf{u}}_n(z) + \mathbf{w}_n(z) = (1-\alpha^n)\mathbf{u}_n(z) + \mathbf{w}_n(z).$$

Note that $\alpha^n \in [0, 1)$, one can verify that

$$\begin{aligned} F(\mathbf{u}_{n+1}(z)) &= (1-\alpha^n)^2|\mathbf{u}_n(z)|^2 - 1 \\ &+ 2(1-\alpha^n)\mathbf{u}_n(z)) \cdot \mathbf{w}_n(z) + |\mathbf{w}_n(z)|^2 \\ &\le (1-\alpha^n)F(\mathbf{u}_n(z)) - \alpha^n \\ &+ 2(1-\alpha^n)\mathbf{u}_n(z)) \cdot \mathbf{w}_n(z) + |\mathbf{w}_n(z)|^2. \end{aligned} \tag{6.5}$$

For convenience, set

$$x_n = F(\mathbf{u}_n(z)) \ \text{ and } \ \varepsilon_n = 2(1-\alpha^n)\mathbf{u}_n(z)) \cdot \mathbf{w}_n(z) + |\mathbf{w}_n(z)|^2.$$

Then the inequality (6.5) can be written as

$$x_{n+1} \le (1-\alpha^n)x_n - \alpha^n + \varepsilon_n. \tag{6.6}$$

Step 2. Verify that the sequence $\{\varepsilon_n\}$ converges to zero.

By Lemma 5.1 (a) and Theorem 5.1, we have for any n (note that $\pi_h F(\mathbf{u}_0) = 0$)

$$c\|\nabla \mathbf{u}_n\|_{0,\Omega}^2 \le \mathcal{E}_h(\mathbf{u}_n) \le \mathcal{E}_h(\mathbf{u}_0) \le C\|\nabla \mathbf{u}_0\|_{0,\Omega}^2.$$

This, together with (3.1) and Poincare inequality, leads to

$$\begin{aligned} |\mathbf{u}_n(z)| &\le C\beta_h(d)\|\mathbf{u}_n\|_{1,\Omega} \le C\beta_h(d)(\|\nabla \mathbf{u}_n\|_{0,\Omega} + \|\pi_h \mathbf{g}\|_{0,\partial\Omega}) \\ &\le C(h), \quad \text{for any } n \text{ and } z \in \mathcal{N}_h. \end{aligned}$$

Namely,

$$|\mathbf{u}_n(z)| \le C(h) = \sqrt{\frac{\beta_0}{2}} \ \text{(for a fixed } h), \ \text{for any } n \text{ and } z \in \mathcal{N}_h. \tag{6.7}$$

On the other hand, we have from Lemma 6.1

$$|\mathbf{w}_n(z)| \to 0 \quad \text{when} \ \ n \to +\infty \ \ (\forall z \in \mathcal{N}_h).$$

This, together with (6.7), leads to

$$\varepsilon_n \to 0, \quad \text{when} \ \ n \to +\infty \ \ (\text{for a fixed } h). \tag{6.8}$$

Step 3. Prove the desired result by the reduction to absurdity.

Let β_0 be defined by (6.7). Suppose the sequence $\{x_n\}$ does not converge to zero. Then, there exists a number $\delta_0 \in (0, \beta_0)$ and a subsequence $\{x_{n_i}\}$, such that

$$x_{n_i} \ge \delta_0 \quad \text{for all } i = 1, 2, \cdots. \tag{6.9}$$

It follows by (6.8) there exists a positive integer M, such that

$$|\varepsilon_n| < \frac{\delta_0}{2\beta_0} \min\{\delta_0, 1\} \ \ \text{when } n \ge \text{M}. \tag{6.10}$$

Without loss of generality, we assume that $\text{M} = n_1$.

Step 3.1. A relation between $x_{\text{M}+1}$ and x_{M}.

By the definition of α^n, together with (6.7) and (6.9), we have

$$1 > \alpha^{\text{M}} \ge \frac{\delta_0}{\beta_0} > 0.$$

This, together with (6.6), (6.10) and (6.9), leads to

$$\begin{aligned} x_{\text{M}+1} &\le (1-\alpha^{\text{M}})x_{\text{M}} + \varepsilon_{\text{M}} \le (1 - \frac{\delta_0}{\beta_0})x_{\text{M}} + \frac{\delta_0}{2\beta_0} \cdot \delta_0 \\ &\le (1 - \frac{\delta_0}{\beta_0})x_{\text{M}} + \frac{\delta_0}{2\beta_0} \cdot x_{\text{M}} = (1 - \frac{\delta_0}{2\beta_0})x_{\text{M}}. \end{aligned} \tag{6.11}$$

Step 3.2. A relation between $x_{\text{M}+2}$ and x_{M}.

There exist two possibilities for $x_{\text{M}+1}$:

$$\text{case}(i) \ \ x_{\text{M}+1} < \frac{\delta_0}{2}; \quad \text{case}(ii) \ \ x_{\text{M}+1} \ge \frac{\delta_0}{2}.$$

If the case (i) occurs, we get by (6.6) and (6.10) (note that $\delta_0 < \beta_0$)

$$\begin{aligned} x_{\text{M}+2} &\le (1-\alpha^{\text{M}+1})x_{\text{M}+1} - \alpha^{\text{M}+1} + \frac{\delta_0}{\beta_0} \cdot \frac{\delta_0}{2} \\ &< x_{\text{M}+1} + \frac{\delta_0}{2} < \frac{\delta_0}{2} + \frac{\delta_0}{2} = \delta_0. \end{aligned} \tag{6.12}$$

If the case (ii) occurs, we have by the definition of α^n and (6.7)

$$\alpha^{\text{M}+1} \ge \frac{\delta_0}{2\beta_0}.$$

Then, we further get by (6.6) and (6.10)-(6.11)

$$\begin{aligned} x_{\mathrm{M}+2} &\leq (1-\frac{\delta_0}{2\beta_0})x_{\mathrm{M}+1} - \frac{\delta_0}{2\beta_0} + \frac{\delta_0}{2\beta_0} \\ &\leq (1-\frac{\delta_0}{2\beta_0})x_{\mathrm{M}+1} \leq (1-\frac{\delta_0}{2\beta_0})^2 x_{\mathrm{M}}. \end{aligned} \tag{6.13}$$

Step 3.3. A relation between $x_{\mathrm{M}+3}$ and x_{M}.
Similarly, we can consider two possibilities of $x_{\mathrm{M}+2}$ and prove that

$$x_{\mathrm{M}+3} < \delta_0$$

or

$$x_{\mathrm{M}+3} \leq (1-\frac{\delta_0}{2\beta_0})x_{\mathrm{M}+2} \leq \begin{cases} (1-\frac{\delta_0}{2\beta_0})\delta_0 < \delta_0, \\ (1-\frac{\delta_0}{2\beta_0})^3 x_{\mathrm{M}}. \end{cases}$$

In the second inequality, we have used (6.12) and (6.13)). Namely,

$$x_{\mathrm{M}+3} < \delta_0 \quad \text{or} \quad x_{\mathrm{M}+3} \leq (1-\frac{\delta_0}{2\beta_0})^3 x_{\mathrm{M}}.$$

Step 3.4. The desired conclusion.
Using this process repeatedly, we obtain

$$x_{n_2} < \delta_0 \quad \text{or} \quad x_{n_2} \leq (1-\frac{\delta_0}{2\beta_0})^{n_2-\mathrm{M}} x_{\mathrm{M}}.$$

But, we have $x_{n_2} \geq \delta_0$ by (6.9). Thus,

$$x_{n_2} \leq (1-\frac{\delta_0}{2\beta_0})^{n_2-\mathrm{M}} x_{\mathrm{M}}.$$

Finally, we can prove, by a repeated process, that

$$x_{n_i} \leq (1-\frac{\delta_0}{2\beta_0})^{n_i-\mathrm{M}} x_{\mathrm{M}}, \quad i = 2, 3, \cdots$$

It is clear that x_{M} is bounded for a fixed h, and $\frac{\delta_0}{2\beta_0} \in (0,1)$ is a constant independent of the index i. Then, we have for sufficiently large n_i

$$x_{n_i} < \delta_0.$$

This contradicts the inequality (6.9). Therefore, the assumption that the sequence $\{x_n\}$ does not converge to zero is false.

□

Now we can prove the final convergence result of this paper.

Theorem 6.1 *Let $\{u_n\}$ be the solution sequence of the problem (5.5) with a fixed h. Then*

(i) there exists a subsequence of $\{\mathbf{u}_n\}$, which is denoted by $\{\mathbf{u}_n\}$ itself, such that $\{\mathbf{u}_n\}$ converges to $\mathbf{u}_h$ strongly both in $\mathbf{H}^1(\Omega)$ and in $\mathbf{L}^\infty(\Omega)$;

(ii) the limit point $\mathbf{u}_h$ *is a discrete harmonic map, i.e., a stationary point of the minimization problem (3.2);*

(iii) when the initial guess $\mathbf{u}_0$*, which may depend on the mesh size* h*, is chosen such that* $|\mathbf{u}_0|_{1,\Omega}$ *is bounded with respect to* h*, there exist a subsequence of* $\{\mathbf{u}_h\}$*, which is denoted by* $\{\mathbf{u}_h\}_{h>0}$ *itself, such that* $\{\mathbf{u}_h\}_{h>0}$ *converges to a harmonic map* $\mathbf{u} \in \mathbf{H}^1_{\mathbf{g}}(\Omega; S^{d-1})$ *weakly in* H^1 *and strongly in* L^4*.*

Proof. (i) It follows by Theorem 5.1 that there exists a subsequence of $\{\mathbf{u}_n\}$ such that $\{\mathbf{u}_n\}$ converges to $\mathbf{u}_h$ weakly in $\mathbf{H}^1(\Omega)$. Since $\{\mathbf{u}_n\}$ and $\mathbf{u}_h$ belong to the finite dimensional space $\mathbf{V}_h$, the convergence is also strong. Moreover, the convergence holds in $\mathbf{L}^\infty(\Omega)$, which will be used in the proof later. Using Lemma 6.2, we have $F(\mathbf{u}_h) = 0$ on $\mathcal{N}_h$. Besides, it is easy to see that $\mathbf{u}_h|_{\partial\Omega} = \pi_h\mathbf{g}$. Hence $\mathbf{u}_h \in \mathbf{K}_h$.

(ii) Without loss of generality, we consider only the case with $d = 3$. It suffices to prove (3.5). Let $\mathbf{w}_n$ be the sequence generated by (6.1) with the current approximation $\mathbf{u}_n$. For $\Phi_h \in \mathbf{V}_{h,0}$, define $\bar{\mathbf{v}}_n = \pi_h(\Phi_h \times \mathbf{D}F(\mathbf{u}_n))$. It is easy to see that $\bar{\mathbf{v}}_n \in \bar{\mathbf{K}}_n$. Thus, we have by the definition of $\mathbf{w}_n$

$$\mathcal{E}'(\bar{\mathbf{u}}_n + \mathbf{w}_n)\bar{\mathbf{v}}_n + G_h^d(\mathbf{D}F(\bar{\mathbf{u}}_n + \mathbf{w}_n) \cdot \bar{\mathbf{v}}_n) = 0, \quad \forall \Phi_h \in \mathbf{V}_{h,0}. \tag{6.14}$$

It can be verified directly that

$$\|\bar{\mathbf{v}}_n - \pi_h(\Phi_h \times \mathbf{D}F(\mathbf{u}_h))\|_{1,\Omega} \to 0^+, \quad n \to +\infty. \tag{6.15}$$

From (5.2), we know that

$$\mathbf{D}F(\mathbf{u}_n) \cdot \mathbf{u}_n \geq 2 > 0 \quad \text{on } \mathcal{N}_h.$$

Then, it follows by Lemma 6.2 that

$$\alpha_n \to 0^+ \quad \text{when } n \to +\infty \ \ (\text{on } \mathcal{N}_h).$$

Moreover, $\mathbf{u}_n$ is uniformly bounded for n (for a fixed h). Thus,

$$\bar{\mathbf{u}}_n - \mathbf{u}_n = -\alpha_n\mathbf{u}_n \to 0^+ \quad \text{when } n \to +\infty \ \ (\text{on } \mathcal{N}_h).$$

This means that $\|\bar{\mathbf{u}}_n - \mathbf{u}_n\|_{\infty,\Omega} \to 0$, and so $\|\bar{\mathbf{u}}_n - \mathbf{u}_n\|_{1,\Omega} \to 0$ (for a fixed h). Furthermore, we get

$$\|\bar{\mathbf{u}}_n - \mathbf{u}_h\|_{1,\Omega} \to 0, \quad n \to +\infty. \tag{6.16}$$

This, together with Lemma 6.1, leads to ($p_k \in \mathcal{N}_h$)

$$\begin{aligned}\mathbf{D}F(\bar{\mathbf{u}}_n + \mathbf{w}_n)(p_k) \cdot \bar{\mathbf{v}}_n(p_k) &= \mathbf{D}F(\bar{\mathbf{u}}_n + \mathbf{w}_n)(p_k) \cdot (\Phi_h \times \mathbf{D}F(\mathbf{u}_n))(p_k) \\ &\to \mathbf{D}F(\mathbf{u}_h)(p_k) \cdot (\Phi_h \times \mathbf{D}F(\mathbf{u}_h))(p_k) = 0. \end{aligned} \tag{6.17}$$

Let $n \to +\infty$ in (6.14), and using (6.15)-(6.17), yields

$$\mathcal{E}'(\mathbf{u}_h)\pi_h(\Phi_h \times \mathbf{D}F(\mathbf{u}_h)) = 0, \quad \forall \Phi_h \in \mathbf{V}_{h,0}. \tag{6.18}$$

Now the desired result follows by **Proposition 3.1** and (6.18).

(iii) By the assumption, Theorem 5.1 and (ii), we know that the sequence $\{|\mathbf{u}_h|_{1,\Omega}\}$ is bounded with respect to h. Then we further deduce the desired result by Theorem 3.1.

□

Remark 6.1 *For the case $d = 2$, we need to verify (3.4) by revising the proof of Theorem 6.1 in an obvious manner.*

Remark 6.2 *From the analysis in this section, we can see that the simple Newton linearization method is convergent when the conditions in* **Corollary 5.1** *are satisfied (In some sense, the simple Newton linearization method can be viewed as a variant of the projection method analyzed in [3]). For the more general situation, the penalty term $G_h^d(F(\mathbf{v}))$ in the Newton-penalty method introduced in this paper seems necessary.*

Remark 6.3 *For the case $d = 2$, it has been shown in [14] that the discrete harmonic map $\mathbf{u}_h$ satisfies the condition (3.8), and so $\mathbf{u}_h$ defined by Theorem 6.1 is just the minimizer of (3.2). Since the minimization problem (3.2) may possess many stationary points for the case $d = 3$, we can not guarantee that $\mathbf{u}_h$ defined by Theorem 6.1 is just the desired (global) minimizer for $d = 3$, unless the initial data $\mathbf{u}_0$ is chosen such that $\mathbf{u}_0$ belongs to some small neighborhood of the minimizer.*

7 The saddle-point problem associated with (5.5)

In this section, we transform the minimization problem (5.5) into a saddle point problem, and investigate the spectral properties of the saddle point problem.

Let $G_h^d(\cdot)$ be defined as in Subsection 5.1, and let $\langle \cdot, \cdot \rangle$ denote the standard *discrete* L^2 inner product. The minimization problem (5.5) with the penalty functional $\mathcal{E}_h(\mathbf{v})$ is equivalent to find $(\mathbf{u}_{n+1}, \lambda_{n+1}) \in \mathbf{V}_{h,\mathbf{g}} \times V_{h,0}$ such that

$$\begin{aligned} \mathcal{E}'(\mathbf{u}_{n+1})\mathbf{v} + G_h^d(\mathbf{D}F(\mathbf{u}_{n+1}) \cdot \mathbf{v}) + \langle \mathbf{D}F(\mathbf{u}_n) \cdot \mathbf{v}, \lambda_{n+1} \rangle = 0, \quad \forall \mathbf{v} \in \mathbf{V}_{h,0}, \\ \langle \mathbf{D}F(\mathbf{u}_n) \cdot (\mathbf{u}_{n+1} - \mathbf{u}_n) + F(\mathbf{u}_n), \mu \rangle = 0, \ \forall \mu \in V_{h,0}. \end{aligned} \tag{7.1}$$

Since $F(\mathbf{w}) = |\mathbf{w}|^2 - 1$, we have $\mathbf{D}F(\mathbf{w}) = 2\mathbf{w}$. Note that the above saddle point system is the same with that in Subsection 4.1 of [3] when $\kappa_2 = 0$ and $G_h^d(\cdot) = 0$. We would like to stress that the system (7.1) is a *linear* saddle-point system, which can be solved by any existing iterative method, for example, the Uzawa-type method introduced in [15]. To improve the convergence of the iterative method, we need to construct preconditioners for the primal system and the Schur complement (this topic did not considered in [3]).

Let $\tilde{\mathbf{g}}_h \in \mathbf{V}_{h,\mathbf{g}}$ denote the zero extension of $\mathbf{g}_h$. Then $\mathbf{u}_{n+1}$ can be written as $\mathbf{u}_{n+1} = \mathbf{u}_{n+1}^0 + \tilde{\mathbf{g}}_h$ with $\mathbf{u}_{n+1}^0 \in \mathbf{V}_{h,0}$. Define

$$\mathbf{g}_n = \pi_h(|\mathbf{u}_n|^2 + 1 - 2\mathbf{u}_n\tilde{\mathbf{g}}_h)$$

and $\mathbf{f}_h \in \mathbf{V}_{h,0}$ by

$$(\mathbf{f}_h, \mathbf{v}) = -\mathcal{E}'(\tilde{\mathbf{g}}_h)\mathbf{v} - 2G_h^d(\tilde{\mathbf{g}}_h \cdot \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_{h,0}.$$

It can be verified that (7.1) can be written as

$$\begin{aligned} \mathcal{E}'(\mathbf{u}_{n+1}^0)\mathbf{v} + 2G_h^d(\mathbf{u}_{n+1}^0 \cdot \mathbf{v}) + 2\langle \mathbf{u}_n \cdot \mathbf{v}, \lambda_{n+1} \rangle = (\mathbf{f}_h, \mathbf{v}), \quad \forall \mathbf{v} \in \mathbf{V}_{h,0}, \\ 2\langle \mathbf{u}_n \cdot \mathbf{u}_{n+1}^0, \mu \rangle = \langle \mathbf{g}_n, \mu \rangle, \quad \forall \mu \in V_{h,0}. \end{aligned} \tag{7.2}$$

For convenience, we transform (7.2) into operator form. Define the operators $\mathbf{A} : \mathbf{V}_{h,0} \to \mathbf{V}_{h,0}$ and $\mathbf{B}_n : \mathbf{V}_{h,0} \to V_{h,0}$ by

$$(\mathbf{A}\mathbf{v}, \mathbf{w}) = \mathcal{E}'(\mathbf{v})\mathbf{w} + G_h^d(\mathbf{D}F(\mathbf{v}) \cdot \mathbf{w}) = \mathcal{E}'(\mathbf{v})\mathbf{w} + 2G_h^d(\mathbf{v} \cdot \mathbf{w}), \quad \forall \mathbf{w} \in \mathbf{V}_{h,0},$$

and

$$\langle \mathbf{B}_n \cdot \mathbf{v}, \mu \rangle = 2\langle \mathbf{v}, \mu \rangle, \quad \forall \mu \in \mathbf{V}_{h,0},$$

respectively. Let $\mathbf{B}_n^* : V_{h,0} \to \mathbf{V}_{h,0}$ denote the dual operator of $\mathbf{B}_n$. Then the saddle point problem (7.2) can be written in the operator form

$$\begin{aligned} \mathbf{A}\mathbf{u}_{n+1}^0 + \mathbf{B}_n^* \lambda_{n+1} &= \mathbf{f}_h, \\ \mathbf{B}_n \mathbf{u}_{n+1}^0 &= \mathbf{g}_n. \end{aligned} \tag{7.3}$$

The rest of this section is devoted to the construction of preconditioners $\mathbf{B}$ and K_n for $\mathbf{A}$ and the Schur complement $S_n = \mathbf{B}_n \mathbf{A}^{-1} \mathbf{B}_n^*$, respectively. For simplicity of exposition, let Λ denote the set of all the indices k of $p_k \in \mathcal{N}_h$. Let $\gamma_{d,k}$ be the local penalty parameter given in Subsection 5.1. Define

$$\Lambda_1 = \{k \in \Lambda;\ \gamma_{d,k} > 0\} \ \text{ and } \ \Lambda_2 = \{k \in \Lambda;\ \gamma_{d,k} = 0\}.$$

Proposition 7.1. Assume that the triangulation $\mathcal{T}_h$ is regular and quasi-uniform. Then there exist constants C_0 and c_0 independent of $\mathcal{T}_h$, such that

$$c_0 h^{d-2} \le \gamma_{d,k} \le C_0 h^{d-2}, \quad \forall k \in \Lambda_1. \tag{7.4}$$

Proof. From the definition of the number a_{ij} (see Subsection 5.1), we know that each a_{ij} has the scale h^{d-2}, i.e., a_{ij} can be written as $a_{ij} = c_{ij} h^{d-2}$, with c_{ij} being a constant independent of h. Since the triangulation $\mathcal{T}_h$ is regular and quasi-uniform, every constants c_{ij} satisfying $c_{ij} > 0$ has both a positive upper bound and a positive lower bound which depend on only the constants in the definitions of the regularity and the quasi-uniformity of the triangulation $\mathcal{T}_h$. From (5.3) we have

$$\rho_{d,k} = h^{d-2} \sum_{r \neq k} (c_{rk} + |c_{rk}|).$$

It is clear that $c_{rk} + |c_{rk}| \ge 0$, and $c_{rk} + |c_{rk}| > 0$ if and only if $c_{rk} > 0$. For each index k, there exist finite indices r at most such that $c_{rk} > 0$. All of these show that

$$c_0 h^{d-2} \le \rho_{d,k} \le C_0 h^{d-2}, \quad \forall k \in \Lambda_1,$$

where C_0 and c_0 are constants independent of h. Similarly, we have

$$c_0 h^{d-2} \le \tilde{\rho}_{d,k} \le C_0 h^{d-2}, \quad \forall k \in \Lambda_1.$$

The above two inequalities imply (7.4).

□

We will construct preconditioners $\mathbf{B}$ and K_n for three different cases, which correspond different triangulations:

Case (i) for all of the nodes p_k, we have $\gamma_{d,k} > 0$, i.e., $\Lambda_2 = \emptyset$;

Case (ii) for all of the nodes p_k, we have $\gamma_{d,k} = 0$ (so $G_h^d(\cdot) = 0$), i.e., $\Lambda_1 = \emptyset$;

Case (iii) both $\Lambda_1 \neq \emptyset$ and $\Lambda_2 \neq \emptyset$.

It is easy to see that the operator $\mathbf{A}$ satisfies

$$c[\|\nabla \mathbf{v}_h\|_{0,\Omega}^2 + G_h^d(|\mathbf{v}_h|^2)] \le (\mathbf{A}\mathbf{v}_h, \mathbf{v}_h) \le C[\|\nabla \mathbf{v}_h\|_{0,\Omega}^2 + G_h^d(|\mathbf{v}_h|^2)], \tag{7.5}$$

for $\mathbf{v}_h \in \mathbf{V}_{h,0}$.

We first consider **Case (i)**. Let $\mathbf{T} : \mathbf{V}_{h,0} \to \mathbf{V}_{h,0}$ and $T_n : V_{h,0} \to V_{h,0}$ denote the linear operators defined by

$$\langle \mathbf{Tv},\ \mathbf{w} \rangle = G_h^d(\mathbf{v} \cdot \mathbf{w}) = \sum_{k=1}^{N} \gamma_{d,k} \mathbf{v}(p_k) \cdot \mathbf{w}(p_k), \quad \mathbf{v} \in \mathbf{V}_{h,0},\ \forall \mathbf{w} \in \mathbf{V}_{h,0}$$

and

$$\langle T_n v,\ w \rangle = 4 \sum_{k=1}^{N} \frac{|\mathbf{u}_n(p_k)|^2}{\gamma_{d,k}} v(p_k) w(p_k), \quad \mathbf{v} \in V_{h,0},\ \forall w \in V_{h,0},$$

respectively. It is easy to see that the stiffness matrices of $\mathbf{T}$ and T_n are *diagonal* matrices with positive diagonal entries (note Lemma 5.1 (b)).

Theorem 7.1 *For* **Case (i)**, *the operator* $\mathbf{A}$ *is spectrally equivalent to the operator* $\mathbf{T}$ *and the Schur complement* S_n *is spectrally equivalent to the operator* T_n. *Namely, we can define preconditioners for* $\mathbf{A}$ *and* S_n *as* $\mathbf{B} = \mathbf{T}$ *and* $K_n = T_n$.

Proof. Since all $\gamma_{d,k} > 0$, by the definition of $G_h^d(\cdot)$ and (7.4), we deduce

$$G_h^d(|\mathbf{v}_h|^2) \geq c h^{-2} \|\mathbf{v}_h\|_{0,\Omega}^2, \quad \forall \mathbf{v}_h \in \mathbf{V}_{h,0}.$$

Then, by the inverse inequality (3.1), we obtain

$$\|\nabla \mathbf{v}_h\|_{0,\Omega}^2 \leq C h^{-2} \|\mathbf{v}_h\|_{0,\Omega}^2 \leq C G_h^d(|\mathbf{v}_h|^2), \quad \forall \mathbf{v}_h \in \mathbf{V}_{h,0}.$$

Plugging this in (7.5), leads to

$$c G_h^d(|\mathbf{v}_h|^2) \leq (\mathbf{A}\mathbf{v}_h, \mathbf{v}_h) \leq C G_h^d(|\mathbf{v}_h|^2), \quad \mathbf{v}_h \in \mathbf{V}_{h,0}.$$

Thus the operator $\mathbf{A}$ is spectrally equivalent to the operator $\mathbf{T}$ by the definition of $\mathbf{T}$. Moreover, the Schur complement S_n is spectrally equivalent to the operator $\mathbf{B}_n \mathbf{T}^{-1} \mathbf{B}_n^*$.

It suffices to prove that $\mathbf{B}_n \mathbf{T}^{-1} \mathbf{B}_n^*$ is spectrally equivalent to the operator T_n. From the definition of $\mathbf{B}_n^*$, we have

$$\mathbf{B}_n^* \mu = 2\pi_h(\mathbf{u}_n \mu), \quad \mu \in V_{h,0}.$$

Then we get by the particular design of $\mathbf{T}$

$$\begin{aligned} \langle \mathbf{B}_n \mathbf{T}^{-1} \mathbf{B}_n^* \mu, \mu \rangle &= \langle \mathbf{T}^{-1} \mathbf{B}_n^* \mu, \mathbf{B}_n^* \mu \rangle = \sum_{k=1}^{N} \gamma_{d,k}^{-1} |(\mathbf{B}_n^* \mu)(p_k)|^2 \\ &= 4 \sum_{k=1}^{N} \frac{|\mathbf{u}_n(p_k)|^2}{\gamma_{d,k}} |\mu(p_k)|^2 = \langle T_n \mu, \mu \rangle, \quad \mu \in V_{h,0}. \end{aligned}$$

This yields the desired result.

□

We next consider **case (ii)**. Let $\mathbf{\Delta}_h : \mathbf{V}_{h,0} \to \mathbf{V}_{h,0}$ and $\Delta_h : V_{h,0} \to V_{h,0}$ denote the discrete Laplacian defined by

$$(\mathbf{\Delta}_h \mathbf{v}_h, \mathbf{w}_h) = (\nabla \mathbf{v}_h, \nabla \mathbf{w}_h), \quad \mathbf{v}_h,\ \mathbf{w}_h \in \mathbf{V}_{h,0}$$

and

$$(\Delta_h v_h, w_h) = (\nabla v_h, \nabla w_h), \quad v_h,\ w_h \in V_{h,0},$$

respectively.

Theorem 7.2 *For* **Case (ii)**, *the operator* $\mathbf{A}$ *is spectrally equivalent to* $\mathbf{\Delta}_h$. *Moreover, for the case* $d = 2$, *the Schur complement* S_n *is spectrally equivalent to* Δ_h^{-1} *in the sense*

$$c\log^{-1}(1/h)(\Delta_h^{-1}v_h, v_h) \le (S_n v_h, v_h) \le C\log(1/h)(\Delta_h^{-1}v_h, v_h), \quad v_h \in V_{h,0}. \tag{7.6}$$

Namely, we have $cond(\mathbf{\Delta}_h^{-1}\mathbf{A}) \le C$ *and* $cond(\Delta_h S_n) \le C\log^2(1/h)$ *for the case* $d = 2$. *In particular, any spectrally equivalent operator with the Laplacian* $\mathbf{\Delta}_h$ *can be chosen as a preconditioner for* $\mathbf{A}$.

Proof. Since $\gamma_{d,k} = 0$ for every nodes p_k, we have $G_h^d(\cdot) = 0$. Then, by (7.5), we get

$$c(\mathbf{\Delta}_h \mathbf{v}_h, \mathbf{v}_h) \le (\mathbf{A}\mathbf{v}_h, \mathbf{v}_h) \le C(\mathbf{\Delta}_h \mathbf{v}_h, \mathbf{v}_h), \quad \mathbf{v}_h \in \mathbf{V}_{h,0}, \tag{7.7}$$

which implies the first result.

From (7.7), we know that the Schur complement S_n is spectrally equivalent to $\mathbf{B}_n\mathbf{\Delta}_h^{-1}\mathbf{B}_n^*$. We need only to investigate the operator $\mathbf{B}_n\mathbf{\Delta}_h^{-1}\mathbf{B}_n^*$. It is easy to see that

$$\langle \mathbf{B}_n\mathbf{\Delta}_h^{-1}\mathbf{B}_n^*\mu_h, \mu_h\rangle = \sup_{\mathbf{v}_h\in\mathbf{V}_{h,0}} \frac{(\langle \mathbf{B}_n\mathbf{v}_h, \mu_h\rangle)^2}{\langle \mathbf{\Delta}_h\mathbf{v}_h, \mathbf{v}_h\rangle}, \quad \mu_h \in V_{h,0}.$$

It suffices to verify that

$$c\log^{-\frac{1}{2}} \|\mu_h\|_{H^{-1}(\Omega)} \le \sup_{\mathbf{v}_h\in\mathbf{V}_{h,0}} \frac{\langle \mathbf{B}_n\mathbf{v}_h, \mu_h\rangle}{\|\mathbf{v}_h\|_{1,\Omega}} \le C\log^{\frac{1}{2}} \|\mu_h\|_{H^{-1}(\Omega)}, \quad \mu_h \in V_{h,0}. \tag{7.8}$$

It follows by Theorem 5.1 that

$$\|\nabla \mathbf{u}_n\|_{0,\Omega} \le \|\nabla \mathbf{u}_0\|_{0,\Omega}.$$

Then, by Poincare inequality, we get

$$\|\mathbf{u}_n\|_{1,\Omega} \le C. \tag{7.9}$$

This, together with the inverse estimate (3.1), yields (noting $d = 2$)

$$\|\mathbf{u}_n\|_{0,\infty,\Omega} \le C\log^{\frac{1}{2}}(1/h). \tag{7.10}$$

We further deduce that

$$\|\mathbf{u}_n \cdot \mathbf{v}_h\|_{1,\Omega} \le C\log^{\frac{1}{2}}(1/h)\|\mathbf{v}_h\|_{1,\Omega}.$$

Thus,

$$\sup_{\mathbf{v}_h\in\mathbf{V}_{h,0}} \frac{\langle \mathbf{B}_n\mathbf{v}_h, \mu_h\rangle}{\|\mathbf{v}_h\|_{1,\Omega}} \le C\log^{\frac{1}{2}}(1/h)\|\mu_h\|_{H^{-1}(\Omega)}, \quad \mu_h \in V_{h,0}. \tag{7.11}$$

For any $\mu_h \in V_{h,0}$, we have

$$\|\mu_h\|_{-1} = \frac{\langle \mu_h, \mu_h\rangle}{\|\mu_h\|_{1,\Omega}}. \tag{7.12}$$

Define $\mathbf{v}_h \in \mathbf{V}_{h,0}$ as

$$\mathbf{v}_h = \pi_h(\mu_h \frac{\mathbf{u}_n}{2|\mathbf{u}_n|^2}),$$

which implies that

$$\mu_h = 2\pi_h(\mathbf{u}_n \cdot \mathbf{v}_h).$$

Then,

$$\langle \mu_h, \mu_h \rangle = 2\langle \mathbf{u}_n \cdot \mathbf{v}_h, \mu_h \rangle = \langle \mathbf{B}_n \mathbf{v}_h, \mu_h \rangle. \tag{7.13}$$

As in Lemma 4.2 of [14], one can prove by (5.2) and (7.9)-(7.10)

$$\|\mathbf{v}_h\|_{1,\Omega} \leq C \log^{\frac{1}{2}}(1/h) \|\mu_h\|_{1,\Omega}.$$

Combining this with (7.12)-(7.13), we obtain

$$\sup_{\mathbf{v}_h \in \mathbf{V}_{h,0}} \frac{\langle \mathbf{B}_n \mathbf{v}_h, \mu_h \rangle}{\|\mathbf{v}_h\|_{1,\Omega}} \geq c \log^{-\frac{1}{2}}(1/h) \|\mu_h\|_{H^{-1}(\Omega)}, \quad \mu_h \in V_{h,0}.$$

Then the relation (7.8) follows by (7.11) and the above inequality.

□

Remark 7.1 *We can image that the iteration (5.5) for* **Case (i)** *possesses slower convergence than that for* **Case (ii)**, *but Theorem 7.1-7.2 tell us that the computationca at each iteration step for* **Case (i)** *is much cheaper than that for* **Case (ii)** *since both* $\mathbf{T}^{-1}$ *and* T_n^{-1} *correspond diagonal stiffness matrices. This means that introduction of the penalty term* $G_h^d(\cdot)$ *does not significantly increase the cost of calculation.*

Remark 7.2 *Unfortunately, we have not obtained a satisfactory spectrally equivalent result for Schur complement* S_n *for* $d = 3$ *in* **Case** *(ii). The main difficulty comes from the "bad" inverse estimate*

$$\|\mu_h\|_{0,\infty,\Omega}^2 \leq Ch^{-1} \|\mu_h\|_{1,\Omega}^2, \quad \mu_h \in V_{h,0} \quad (d = 3).$$

But, if the triangulation $\mathcal{T}_h$ *has nested structure, and we use the well known BPX multilevel preconditioner* $\mathbf{M}$ *for* $\mathbf{\Delta}_h$, *then the resulting Schur complement* $\mathbf{B}_n \mathbf{M}^{-1} \mathbf{B}_n^*$ *has also multilevel structure. Since all the coarse solvers in* $\mathbf{M}$ *correspond diagonal stiffness matrices, we can derive a simple (multilevel) expression for* $\mathbf{B}_n \mathbf{M}^{-1} \mathbf{B}_n^*$ *as in Theorem 7.1. Based on this, it is possible to construct an multilevel preconditioner for* $\mathbf{B}_n \mathbf{M}^{-1} \mathbf{B}_n^*$.

Now we consider **Case (iii)**. As in the proof of **Proposition 3.1**, we use $\Phi_k^r \in \mathbf{V}_{h,0}$ $(r = 1, 2, 3)$ to denote the three nodal basis vectors associated with an interior node p_k. Define

$$\mathbf{V}_{h,0}^{(1)} = span\{\Phi_k^r;\ r = 1, 2, 3;\ k \in \Lambda_1\} \ \text{ and } \ \mathbf{V}_{h,0}^{(2)} = span\{\Phi_k^r;\ r = 1, 2, 3;\ k \in \Lambda_2\}.$$

Then we have the direct sum decomposition

$$\mathbf{V}_{h,0} = \mathbf{V}_{h,0}^{(1)} + \mathbf{V}_{h,0}^{(2)}.$$

Let $\mathbf{T}^{(1)} : \mathbf{V}_{h,0}^{(1)} \to \mathbf{V}_{h,0}^{(1)}$ and $\mathbf{\Delta}_h^{(2)} : \mathbf{V}_{h,0}^{(2)} \to \mathbf{V}_{h,0}^{(2)}$ be the restrictions of $\mathbf{T}$ on $\mathbf{V}_{h,0}^{(1)}$ and $\mathbf{\Delta}_h$ on $\mathbf{V}_{h,0}^{(2)}$, respectively. It is clear that $\mathbf{T}^{(1)}$ and $\mathbf{\Delta}_h^{(2)}$ are inverse operators (but $\mathbf{T}$ is not inverse when $\Lambda_2 \neq \emptyset$). Then the preconditioner for $\mathbf{A}$ is defined as

$$\mathbf{B}^{-1} = (\mathbf{T}^{(1)})^{-1} \mathbf{Q}_1 + (\mathbf{\Delta}_h^{(2)})^{-1} \mathbf{Q}_2,$$

where $\mathbf{Q}_i : \mathbf{V}_{h,0} \to \mathbf{V}_{h,0}^{(i)}$ $(i = 1, 2)$ denotes the L^2 projector.

Let $\varphi_k \in V_{h,0}$ denote the nodal basis function associated with an interior node p_k. Set

$$V_{h,0}^{(1)} = span\{\varphi_k; \quad k \in \Lambda_1\} \quad \text{and} \quad V_{h,0}^{(2)} = span\{\varphi_k; \quad k \in \Lambda_2\}.$$

Then we have

$$V_{h,0} = V_{h,0}^{(1)} + V_{h,0}^{(2)}.$$

Define $T_n^{(1)} : V_{h,0}^{(1)} \to V_{h,0}^{(1)}$ by

$$\langle T_n v, \ w\rangle = 4 \sum_{k\in\Lambda_1} \frac{|\mathbf{u}_n(p_k)|^2}{\gamma_{d,k}} v(p_k)w(p_k), \quad \mathbf{v} \in V_{h,0}^{(1)}, \ \forall w \in V_{h,0}^{(1)}.$$

Let $\Delta_h^{(2)} : V_{h,0}^{(2)} \to V_{h,0}^{(2)}$ be the restrictions of Δ_h on $V_{h,0}^{(2)}$. Then the preconditioner for S_n is defined as

$$K_n^{-1} = (T_n^{(1)})^{-1} Q_1 + \Delta_h^{(2)} Q_2,$$

where $Q_i : V_{h,0} \to V_{h,0}^{(i)}$ $(i = 1, 2)$ denotes the L^2 projector.

Theorem 7.3 *Let* $\mathbf{B}$ *and* K_n *be the preconditioners defined above. Then we have cond*$(\mathbf{B}^{-1}\mathbf{A}) \le C$ *and cond*$(K_n^{-1}S_n) \le C \log^2(1/h)$ *for the case* $d = 2$.

Proof. We prove the first result only. By the standard theory (see, for example, [19]) one needs only to verify that

(a) for any $\mathbf{v} \in \mathbf{V}_{h,0}$ there exists a decomposition $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$ with $\mathbf{v}_i \in \mathbf{V}_{h,0}^{(i)}$ $(i = 1, 2)$ such that

$$(\mathbf{T}^{(1)}\mathbf{v}_1, \mathbf{v}_1) + (\mathbf{\Delta}_h^{(2)}\mathbf{v}_2, \mathbf{v}_2) \le C_1(\mathbf{A}\mathbf{v}, \mathbf{v}) \tag{7.14}$$

and

(b) for any $\mathbf{w}_i \in \mathbf{V}_{h,0}^{(i)}$ $(i = 1, 2)$ we have

$$(\mathbf{A}(\mathbf{w}_1 + \mathbf{w}_2), \mathbf{w}_1 + \mathbf{w}_2) \le C_2[(\mathbf{T}^{(1)}\mathbf{w}_1, \mathbf{w}_1) + (\mathbf{\Delta}_h^{(2)}\mathbf{w}_2, \mathbf{w}_2)]. \tag{7.15}$$

We first verify the condition (a). For any $\mathbf{v} \in \mathbf{V}_{h,0}$, define $\mathbf{v}_i \in \mathbf{V}_{h,0}^{(i)}$ such that $\mathbf{v}_i(p_k) = \mathbf{v}(p_k)$ for any $k \in \Lambda_i$ and $\mathbf{v}_i(p_k) = 0$ for any $k \notin \Lambda_i$ $(i = 1, 2)$. Then we have $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2$. By the definition of $\mathbf{T}^{(1)}$, we get

$$(\mathbf{T}^{(1)}\mathbf{v}_1, \mathbf{v}_1) = (\mathbf{T}\mathbf{v}_1, \mathbf{v}_1) = (\mathbf{T}\mathbf{v}, \mathbf{v}) \le (\mathbf{A}\mathbf{v}, \mathbf{v}). \tag{7.16}$$

On the other hand, we have

$$(\mathbf{\Delta}_h^{(2)}\mathbf{v}_2, \mathbf{v}_2) = (\mathbf{\Delta}_h\mathbf{v}_2, \mathbf{v}_2) = |\nabla\mathbf{v}_2|_{0,\Omega}^2 \le 2(|\nabla\mathbf{v}|_{0,\Omega}^2 + |\nabla\mathbf{v}_1|_{0,\Omega}^2).$$

This, together with (3.1) and (7.4), leads to

$$(\mathbf{\Delta}_h^{(2)}\mathbf{v}_2, \mathbf{v}_2) \le C(|\nabla\mathbf{v}|_{0,\Omega}^2 + h^{-2}|\mathbf{v}_1|_{0,\Omega}^2) \le C(|\nabla\mathbf{v}|_{0,\Omega}^2 + G_h^d(|\mathbf{v}_1|^2)).$$

By (7.5), we further deduce

$$(\mathbf{\Delta}_h^{(2)}\mathbf{v}_2, \mathbf{v}_2) \le C(|\nabla\mathbf{v}|_{0,\Omega}^2 + G_h^d(|\mathbf{v}|^2)) \le C(\mathbf{A}\mathbf{v}, \mathbf{v}).$$

Combining this with (7.16), gives (7.14).

Now we consider the condition (b). Let $\mathbf{w}_1 \in \mathbf{V}_{h,0}^{(1)}$. By (7.5) and (3.1), we get

$$(\mathbf{A}\mathbf{w}_1, \mathbf{w}_1) \leq C(\|\nabla \mathbf{w}_1\|_{0,\Omega}^2 + G_h^d(|\mathbf{w}_1|^2)) \leq C(h^{-2}\|\mathbf{w}_1\|_{0,\Omega}^2 + G_h^d(|\mathbf{w}_1|^2)).$$

This, together with (7.4), leads to

$$(\mathbf{A}\mathbf{w}_1, \mathbf{w}_1) \leq C(\mathbf{T}^{(1)}\mathbf{w}_1, \mathbf{w}_1), \quad \forall \mathbf{w}_1 \in \mathbf{V}_{h,0}^{(1)}.$$

Besides, since $(\mathbf{T}\mathbf{w}_2, \mathbf{w}_2) = 0$ for $\mathbf{w}_2 \in \mathbf{V}_{h,0}^{(2)}$, we have

$$(\mathbf{A}\mathbf{w}_2, \mathbf{w}_2) \leq C(\mathbf{\Delta}_h^{(2)}\mathbf{w}_2, \mathbf{w}_2), \quad \forall \mathbf{w}_2 \in \mathbf{V}_{h,0}^{(2)}.$$

Combining the above two inequalities, yields (7.15).

The second result can be proved in a similar manner, by using some results obtained in the proofs of Theorem 7.1-7.2.

□

Remark 7.3 *It follows by Theorem 7.1-7.3 that both the operators* $\mathbf{A}$ *and the Schur complement* $S_n = \mathbf{B}_n(\mathbf{A}^{-1}\mathbf{B}_n^*$ *are positive definite. Thus, the saddle-point problem (7.3) has a unique solution* $(\mathbf{u}_{n+1}^0, \lambda_{n+1}) \in \mathbf{V}_{h,0} \times V_{h,0}$.

8 Numerical experiments

In this section we shall use the proposed algorithm to solve the problem (1.2). We shall report some numerical results to illustrate efficiency of the new iteration method (5.5).

Let $\mathbf{u}_0$ denote a suitable initial guess satisfying $\pi_h F(\mathbf{u}_0) = 0$, and let $\{\mathbf{u}_n\}_{n\geq 1}$ denote the solution sequence generated by the iteration (5.5). Set

$$\epsilon_n = \frac{\|\nabla(\mathbf{u}_{n+1} - \mathbf{u}_n)\|_{0,\Omega}}{\|\nabla(\mathbf{u}_1 - \mathbf{u}_0)\|_{0,\Omega}} \quad (n \geq 1).$$

The stopping criterion in the iteration (5.5) is that the tolerance $\epsilon_n < 1.e-4$ in the case of subsection 8.1.2, or $\epsilon_n < 1.e-3$ in other cases.

8.1 Examples with $\kappa_1 = 1$ and $\kappa_2 = 0$

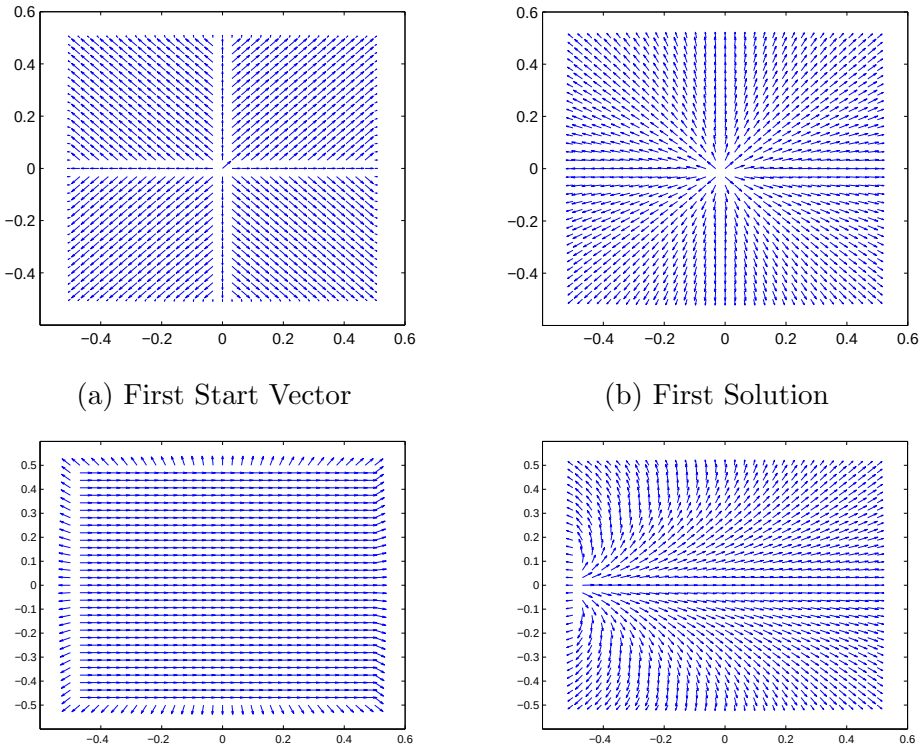
In this subsection we set $\kappa_1 = 1$ and $\kappa_2 = 0$ in the functional (1.1), which was considered in [14] (for the case of two dimensions) and [3] (for the case of three dimensions).

8.1.1 A two-dimensional harmonic map with singularity

It is well known that the solution of the harmonic map problem is generally not unique and may have singularities even with smooth data. In order to show the applicability of our algorithms for these problems, we test a problem with a singular solution (see [14]): $\mathbf{u}(\mathbf{x}) = (x_1/r, x_2/r)$, with $r(\mathbf{x}) = \sqrt{(x_1^2 + x_2^2)}$ on $\Omega = (-0.5, 0.5) \times (-0.5, 0.5)$. The Dirichlet boundary conditions are obtained from the analytical solution.

We adopt a uniform triangulation $\mathcal{T}_h$ as follows: we first divide Ω into small squares with side-length h, and then divide each small square into two equal triangles. The triangulation corresponds to **Case (ii)** that $\gamma_{d,k} = 0$ for every nodes p_k.

The initial guess $\mathbf{u}_0$ for the new iteration method (5.5) is shown in Figure 8.1.1(a,c). The computed solution is shown in Figure 8.1.1(b,d).



(a) First Start Vector

(b) First Solution

(c) Second Start Vector

(d) Second Solution

Plot of the initial solutions and the computed solutions. a) The first initial solution. b) The solution for (a). c) The second initial solution.d) The solution for (c).

The numerical errors are given in Table 1. The errors indicate that $\mathbf{u}_h$ converges quasi linearly to the solution when measured in l^2. It is interesting to observe that we get convergence for $\| \mathbf{u} - \mathbf{u}_h \|_0$ even without mesh refinement around the singularity. In Table 2 we list the iteration counts of the new iteration algorithm (5.5) with different mesh sizes.

Table 1: The L_2 error of $\mathbf{u_h}$ with respect to h

h	2^{-3}	2^{-4}	2^{-5}
$\| \mathbf{u} - \mathbf{u_h} \|_0$	2.5e-1	1.4e-1	7.7e-2

Table 2: Iteration counts with respect to h

h	2^{-3}	2^{-4}	2^{-5}
iter	14	14	13

8.1.2 A three-dimensional harmonic map with singularity

Set $\Omega = (-0.5, 0.5)^3$ and $\mathbf{g}(\mathbf{x}) = \mathbf{x}/|\mathbf{x}|, \mathbf{x} \in \partial\Omega$. Then, $\mathbf{u}(\mathbf{x}) = \mathbf{x}/|\mathbf{x}|$ $(\mathbf{x} \in \Omega)$ is the unique solution of the problem (1.2) (see [3]). We consider the following triangulations of Ω in this subsubsection.

Let Ω be divided into some small cuboid with three sides of lengths being $\frac{1}{2}h$,h and h, and then let each small cuboid be further divided into five smaller tetrahedrons (see Figure 1). Let v_k $(k = 1, \cdots, 8)$ denote the 8 vertices of a cuboid, then the five tetrahedrons in the cuboid are given as follows:

$T_1 := conv\{v_4, v_1, v_2, v_6\}$, $T_2 := conv\{v_8, v_6, v_4, v_7\}$, $T_3 := conv\{v_3, v_7, v_1, v_4\}$,

$T_4 := conv\{v_5, v_7, v_1, v_6\}$, $T_5 := conv\{v_4, v_7, v_1, v_6\}$.

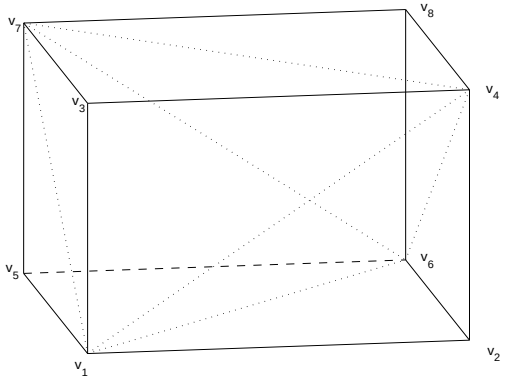


Figure 1: The division in a cuboid of the triangulation

Note that the above tetrahedrons do not satisfy the assumption described in [3]. The triangulation corresponds to **Case (iii)** that both $\gamma_{d,k} > 0$ and $\gamma_{d,k} = 0$ happen for different nodes p_k. For convenience, let $\mathcal{T}_h$ denote the resulting triangulation, and let $\mathcal{N}_h$ denote the set of the nodes. For the triangulation $\mathcal{T}_h$, the initial guess $\mathbf{u}_h^0$ in the algorithm (5.5) is defined by (see [3])

$$\mathbf{u}_h^0(\mathbf{x}) := \begin{cases} \mathbf{x}/|\mathbf{x}|, & \text{for} \quad \mathbf{x} \in \mathcal{N}_h \cap \partial\Omega \\ (0,1,0), & \text{for} \quad \mathbf{x} \in \mathcal{N}_h \cap \Omega. \end{cases}$$

The saddle-point system (7.2) is solved by Uzawa algorithm introduced in [15], with the preconditioners described in Section 7. In Table 3 we report the L^2 errors of the approximations $\mathbf{u}_h$ in terms of h. We observe that the L^2 error of the approximated solutions $\mathbf{u}_h$ decreases linearly with respect to h.

Table 3: The L_2 error of $\mathbf{u}_h$ with respect to h

.

h	2^{-3}	2^{-4}	2^{-5}
$\| \mathbf{u} - \mathbf{u_h} \|_{\mathbf{0}}$	2.0e-1	1.1e-1	5.7e-2

In Table 4 we list the iteration counts of the method (5.5). We observe that the iteration counts increases linearly with respect to h.

Table 4: Iteration counts with respect to h

h	2^{-3}	2^{-4}	2^{-5}
iter	131	244	513

Let $\mathbf{u}_h^{(j)}$ denote the approximation generated by j-th steps iteration (5.5) with the starting value $\mathbf{u}_h^{(0)}$ defined above, and let $\mathbf{u}_h^{(j)}(0,.,.)$ denote the projection of the vector fields $\mathbf{u}_h^{(j)}$ onto the plane $\{(x_1,x_2,x_3) \in \mathbb{R}^3 : x_1 = 0,\ (x_2,x_3) \in (-0.5,0.5)^2\}$. The following Figure 2 shows the projections $\mathbf{u}_h^{(j)}(0,.,.)$ with $h = 1/32$ and $j = 0, 10, 50, 513$.

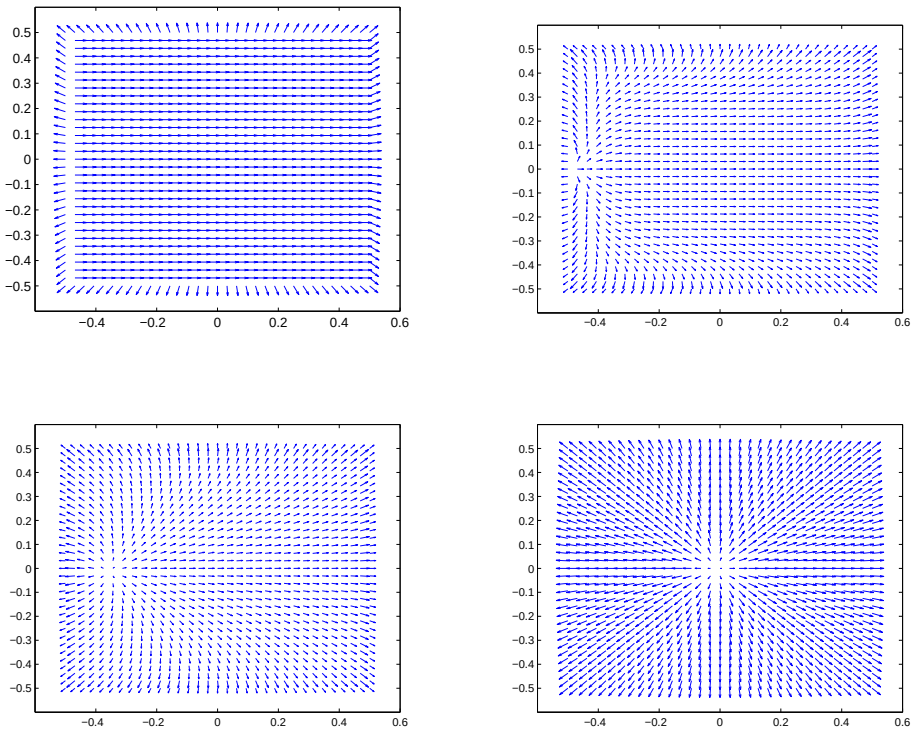


Figure 2: The projections $\mathbf{u}_h^{(j)}(0,.,.)$ for $h = 1/32$ and $j = 0, 10, 50, 513$

We observe that, for the case with $h = 1/32$, the output $\mathbf{u}_h^{513}$ generated by 513 iterations appears to be very close to the exact solution away from 0. The value of the numerical solution at 0, where the exact solution has a singularity, has no particular meaning and seems to depend on the triangulation and the initial value.

As pointed out in [3], the definition of $\mathbf{u}_h^{(0)}$ is suboptimal as it admits large gradients in a neighborhood of $\partial\Omega$. As in [3], we choose another starting value. Let $\xi(\mathbf{p})$, for all $\mathbf{p} \in \Omega$, be a random unit vector in $\mathbb{R}^3$, and let the starting value $\tilde{\mathbf{u}}_h^{(0)}$ is defined by (see [3])

$$\tilde{\mathbf{u}}_h^0(\mathbf{p}) := \begin{cases} \mathbf{p}/|\mathbf{p}|, & \text{for} \quad \mathbf{p} \in \mathcal{N}_h \cap \partial\Omega \\ \xi(\mathbf{p}), & \text{for} \quad \mathbf{p} \in \mathcal{N}_h \cap \Omega \end{cases}$$

Let $\tilde{\mathbf{u}}_h^{(j)}(0,.,.)$, associated with the starting value $\tilde{\mathbf{u}}_h^0$, be defined as $\mathbf{u}_h^{(j)}(0,.,.)$. The following Figure 3 shows the projections $\tilde{\mathbf{u}}_h^{(j)}(0,.,.)$ with $h = 1/32$ and $j = 0, 10, 100, 510$.

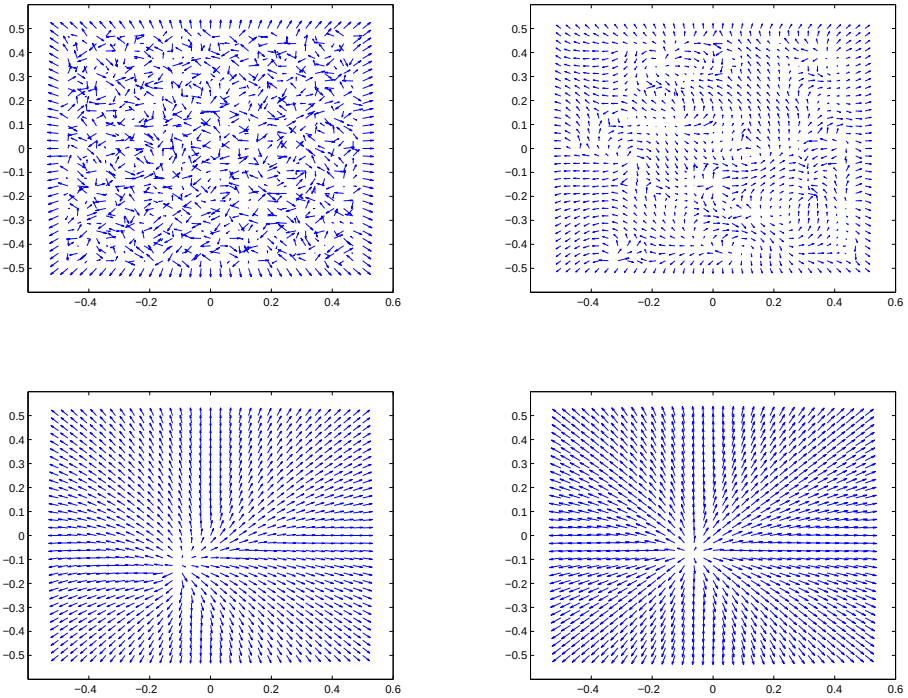


Figure 3: The projections $\tilde{\mathbf{u}}_{32}^{(j)}(0,.,.)$ for $h = 1/32$ and $j = 0, 10, 100, 510$

For such case, we observe that the algorithm immediately changes the highly unordered initial configuration into a more stable one; after one hundred iterations only one degree of one singularity with high symmetry can be seen. The subsequent iterations move the singularity to the origin.

8.2 Examples with $\kappa_2 = 1$

In this subsection we set $\kappa_2 = 1$ in the functional (1.1), but κ_1 may be different, which was considered in [11].

8.2.1 Another two-dimensional harmonic map with singularity

Set $\Omega = (0,1) \times (0,1)$. We adopt a uniform triangulation $\mathcal{T}_h$ as follows: we first divide Ω into small squares with side-length h, and then divide each small square into two equal triangles. The triangulation corresponds to **Case (i)** that $\gamma_{d,k} > 0$ for all the nodes p_k since $\kappa_2 \neq 0$.

For the case with $\kappa_2 \neq 0$, it seems difficult to construct an analytic solution of the problem (1.2). It is clear that the solutions of (1.2) are determined by the boundary conditions $\mathbf{u} = \mathbf{g}$. Thus we only need to define the boundary value function $\mathbf{g}$. We consider two cases for the boundary conditions $\mathbf{u} = \mathbf{g}$:

(i) $\mathbf{g} = \mu$, where μ is the unit outward normal vector on Γ.

(ii) $\mathbf{g}$ satisfies $|\mathbf{g}| = 1$ and $\mathbf{g} \cdot \mu = sin\gamma$. Here γ is a constant angle between vectors $\mathbf{g}$ and μ.

We define initial values $\mathbf{u}_h^0$ in the iteration (5.5) as the interpolation of the following function

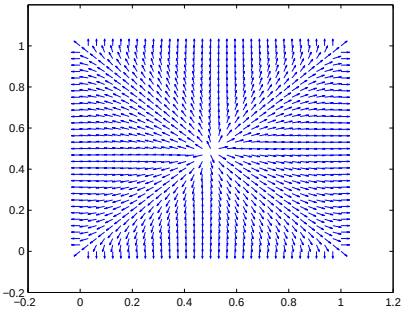
$$\mathbf{u}^0(x_1, x_2) = \begin{cases} (-sin(\phi), cos(\phi)) & \text{if } x_2 \geq x_1 \text{ and } x_2 \geq 1 - x_1; \\ (cos(\phi), sin(\phi)) & \text{if } x_2 < x_1 \text{ and } x_2 \geq 1 - x_1; \\ (sin(\phi), -cos(\phi)) & \text{if } x_2 < x_1 \text{ and } x_2 < 1 - x_1; \\ (-cos(\phi), -sin(\phi)) & \text{if } x_2 \geq x_1 \text{ and } x_2 < 1 - x_1. \end{cases}$$

Here we choose $\phi = 0$ for the case (i) and $\phi = \pi/4$ for the case (ii). This initial values for the case (ii) have the same formulas as the boundary value function $\mathbf{g}$ everywhere except in $(1/2, 1/2)$, where the initial value is taken $(1/\sqrt{2}, 1/\sqrt{2})$.

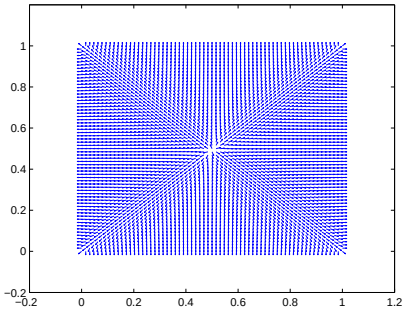
The computational results for the case (i) are depicted in Figure 4. Our computed solutions verify the expectation of the researchers, namely, singularities take place on the diagonals of the square (see,e.g.,[11]). In Table 5 we list the iteration counts of the new iteration algorithm (5.5) with different mesh sizes and κ_1. As we can see, when we fix h, the convergence is quasi independent of the κ_1; when we fix κ_1, the iteration counts almost linearly increase with respect to $1/h$.

Table 5: Iteration counts with respect to h and κ_1

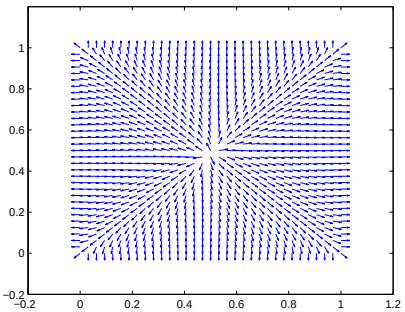
$h \setminus \kappa_1$	0.1	0.01	0.001
2^{-4}	49	59	60
2^{-5}	95	103	104
2^{-6}	167	177	179



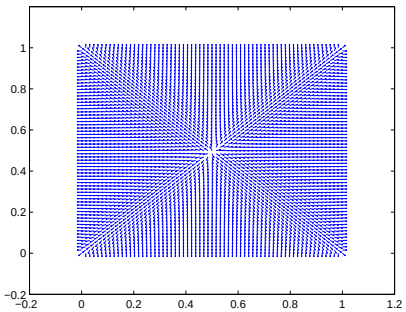
$h = 1/32, \kappa_1 = 0.1, \kappa_2 = 1$



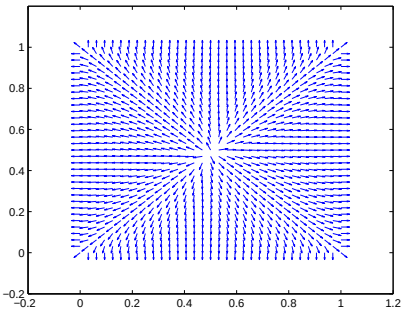
$h = 1/64, \kappa_1 = 0.1, \kappa_2 = 1$



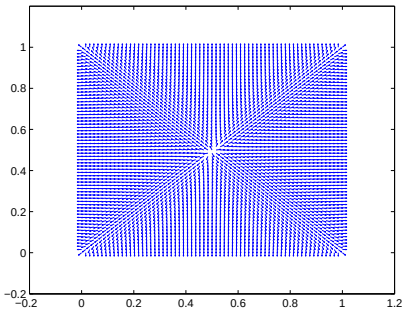
$h = 1/32, \kappa_1 = 0.01, \kappa_2 = 1$



$h = 1/64, \kappa_1 = 0.01, \kappa_2 = 1$



$h = 1/32, \kappa_1 = 0.001, \kappa_2 = 1$



$h = 1/64, \kappa_1 = 0.001, \kappa_2 = 1$

Figure 4: $\mathbf{u}$ (director field) on a square liquid crystal slab with outward normal boundary values

The result for the case (ii) is depicted in Figure 5. After 57 iterations algorithm (5.5) with initial values $\mathbf{u}_h^0$ terminates, and we observe that the numerical solution has one point singularity.

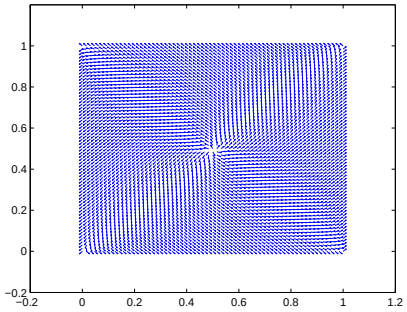


Figure 5: **u** (director field) on a square liquid crystal slab with boundary values possessing constant angle $\sin\gamma = 0.7$ to the outward normal

8.2.2 Another three-dimensional harmonic map with singularity

Set $\Omega = (0,1)^3$. We adopt the uniform triangulation $\mathcal{T}_h$ as shown in Figure 6. The triangulation corresponds to **Case (i)** that $\gamma_{d,k} > 0$ for all the nodes p_k since $\kappa_2 \neq 0$. Let Ω be divided into small regular hexahedrons with the same size h, and then let each small hexahedron be further divided into six smaller tetrahedrons. Let v_k $(k = 1, \cdots, 8)$ denote the 8 vertices of a hexahedron, then the six tetrahedrons in the hexahedron are given as follows:

$T_1 := conv\{v_1, v_2, v_3, v_6\}$, $T_2 := conv\{v_2, v_4, v_3, v_6\}$, $T_3 := conv\{v_3, v_4, v_8, v_6\}$,

$T_4 := conv\{v_3, v_8, v_7, v_6\}$, $T_5 := conv\{v_7, v_5, v_3, v_6\}$, $T_6 := conv\{v_3, v_5, v_1, v_6\}$.

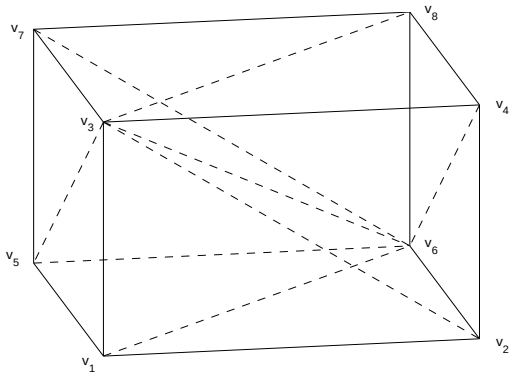


Figure 6: The division in a hexahedron of the triangulation

We shall consider two cases for the boundary conditions $\mathbf{u} = \mathbf{g}$.

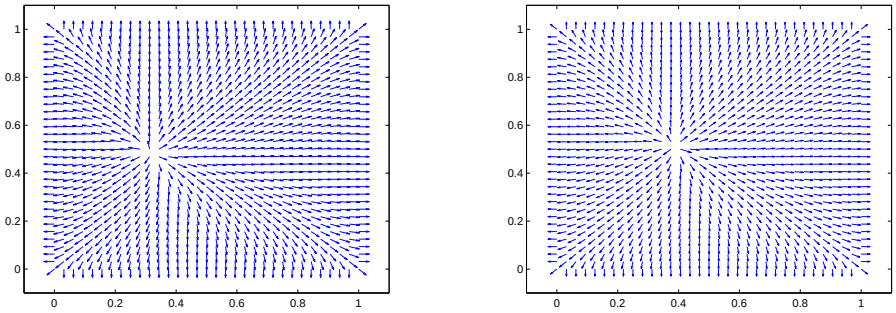
(i) $\mathbf{g} = \mu$, where μ is the unit outward normal vector on Γ.

(ii) $\mathbf{g}$ satisfies $|\mathbf{g}| = 1$ and $\mathbf{g} \cdot \mu = sin\gamma, \mathbf{g} \times \mu = (\xi, \xi, 0), (\xi, 0, \xi), (0, \xi, \xi)$. Here γ is a constant angle between vectors $\mathbf{g}$ and μ, $\xi = cos(\gamma)/\sqrt{2}$.

The initial guess $\mathbf{u}_h^0$ for the iteration (5.5) is defined by

$$\mathbf{u}_h^0(\mathbf{x}) := \begin{cases} \mathbf{g}(\mathbf{x}), & \text{for} \quad \mathbf{x} \in \mathcal{N}_h \cap \partial\Omega, \\ (0,1,0), & \text{for} \quad \mathbf{x} \in \mathcal{N}_h \cap \Omega. \end{cases}$$

The computational results for the case (i) are depicted in Figure 7, which depicts the projections $\mathbf{u}_h^{(j)}(0.5,.,.)$ onto the plane $\{(x_1,x_2,x_3) \in \mathbb{R}^3 : x_1 = 0.5,\ (x_2,x_3) \in (0,1)^2\}$. For $\kappa_1 = 0.1$ and 1, they have just one singularity. In Table 6 we list the iteration counts of the new iteration algorithm (5.5) with different κ_1 and h. As we can see, when we fix κ_1, the iteration counts almost linearly increases with respect to $1/h$.



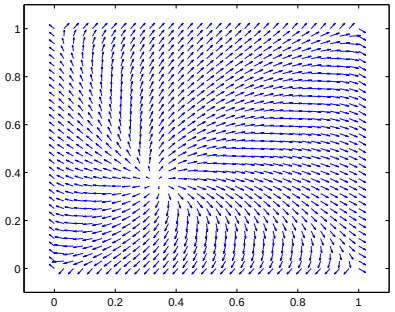
$h = 1/32, k_1 = 0.1, k_2 = 1$ $\qquad$ $h = 1/32, k_1 = 1, k_2 = 1$

Figure 7: $\mathbf{u}$ (director field) on a square liquid crystal slab with the boundary condition (i)

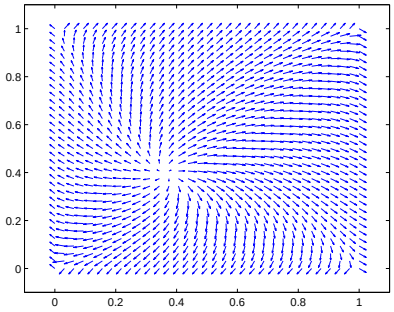
Table 6: Iteration counts with respect to κ_1

$h \setminus \kappa_1$	0.1	1
2^{-3}	260	155
2^{-4}	522	505
2^{-5}	828	888

The computational result for the case (ii) is depicted in Figure 8, which shows the projections $\mathbf{u}_h^{(j)}(0.5,.,.)$ onto $\{(x_1,x_2,x_3) \in \mathbb{R}^3 : x_1 = 0.5,\ (x_2,x_3) \in (0,1)^2\}$. For $\kappa_1 = 0.1$ and 1, they have also one singularity. In Table 7 we list the iteration counts of the new iteration algorithm (5.5) with different κ_1 and h. As we can see, when we fix κ_1, the iteration counts increase almost linearly with respect to $1/h$.



$h = 1/32, k_1 = 0.1, k_2 = 1$



$h = 1/32, k_1 = 1, k_2 = 1$

Figure 8: **u** (director field) on a square liquid crystal slab with the boundary condition (ii)

Table 7: Iteration counts with respect to κ_1

$h \setminus \kappa_1$	0.1	1
2^{-3}	396	328
2^{-4}	1011	789
2^{-5}	1199	1227

8.3 Concluding remarks

In this section we apply the Newton-penalty method introduced in Section 5 to solving some test problems defining harmonic maps in two dimensions or three dimensions, especially both $\kappa_2 = 0$ and $\kappa_2 \neq 0$ are considered. The reported numerical results confirm the *global* convergence of the proposed approach whenever $\kappa_2 = 0$ or $\kappa_2 = 1$. The results indicate that the iteration counts is almost h-independent for the two dimensional case with $\kappa_2 = 0$, and is only linearly increasing with respect to $1/h$ for the other cases. The new method not only is easy to implement but also is efficient to more general models with $\kappa_2 \neq 0$, especially without any particular requirement to triangulations.

Acknowledge The most parts of this work were finished when the first author visited in Centre of Mathematics for Applications, University of Oslo, Norway. This article is a further work of the paper [14]. During the visit, the first author had thorough discussion on the work with Professor Xuecheng Tai and Professor Ragnar Winther, who proposed many insightful suggestions to this work. In fact this article was motivated partly by their idea that the simple Newton's method perhaps can be used to computation of harmonic maps (see Section 4). Also they read the initial version of the article carefully and gave many useful comments for the improvement of the article.

References

[1] F. Alouges, *A new algorithm for computing liquid crystal stable configurations: the harmonic mapping case*, SIAM J. Numer. Anal., **34**(1997), 1708-1726

[2] F. Alouges, *A new finite element scheme for Landau-Lifchitz equations*. Discrete Contin. Dyn. Syst. Ser. S **1**(2008), No. 2, 187196

[3] S. Bartels, *Stability and convergence of finite element approximation schemes for harmonic maps*, SIAM J. Numer. Anal., **43**(2005), 220-238

[4] S. Bartels, *Numerical analysis of a finite element scheme for the approximation of harmonic maps into surfaces*, Math. Comp., **79**(2010):1263-1301.

[5] J. Barrett, X. Feng and A. Prohl, *On p-harmonic map heat flows for* $1 \leq p < \infty$ *and their finite element approximations*, SIAM J. Math. Anal., **40**(2008), No. 4, pp. 1471C1498

[6] F. Bethuel, H. Brezis and F. Hélein, Ginzburg-Landau vortices. Progress in Nonlinear Differential Equations and their Applications, 13. Birkhauser Boston, Inc., Boston, MA, 1994.

[7] H. Brezis, *The interplay between analysis and topology in some nonlinear PDE problems*, Bull. Amer. Math. Soc. **40**(2003), 179-201

[8] Y. Chen and M. Struwe, *Existence and partial regularity results for the heat flow for harmonic maps*, Math. Z., **201**(1989), 83-103

[9] X. Chen, *Global and superlinear convergence of inexact Uzawa methods for saddle-point problems with nondifferentiable mappings*, SIAM J. Numer. Anal., **35** (1998), pp. 1130–1148.

[10] W. E and X. Wang, *Numerical Methods for the Landau-Lifshitz equation*, SIAM J. Numer. Anal., **38**(2000), 1647-1665

[11] R. Glowinski, P. Lin and X. Pan, *An operator-splitting method for a liquid crystal model*, Computer Physics Communications, **152** (2003), 242-252

[12] F. Hélein, *Régularité des applications faiblement harmoniques une surface et une variété riemannienne*, C. R. Acad. Sci. Paris 312 (1991), 591–596.

[13] F. Hélein, Harmonic maps, conservation laws and moving frames, Cambridge University Press, Cambridge, 2002

[14] Q. Hu, X.-C. Tai and R. Winther, *A saddle-point approach to the computation of harmonic maps*, SIAM J. Numer. Anal., **47**(2009), 1500-1523

[15] Q. Hu and J. Zou, *Nonlinear Inexact Uzawa Algorithms for Linear and Nonlinear Saddle-point Problems*, SIAM J. Optim., **16**(2006), 798-825

[16] M. Pierre, *Newton and conjugate gradient for harmonic maps from the disc into the sphere*, ESAIM: Control, Optimisation and Calculus of Variations, **10**(2004), No.1, 142-167

[17] B. T. Polyak, Introduction to Optimization. Optimization Software, Inc., Publications Division, New York, 1987

[18] R. Vanselow, *About Delaunay triangulations and discrete maximum principles for the linear conforming FEM applied to the Poisson equation*. Appl. Math., **46**(2001), No. 1, 13-28

[19] J. Xu, *Iterative methods by space decomposition and subspace correction*, SIAM Review, **34**(1992), 581-613.