SIAM J. NUMER. ANAL.
Vol. 46, No. 1, pp. 61–87

FINITE ELEMENT APPROXIMATIONS OF WAVE MAPS INTO SPHERES*

SÖREN BARTELS†, XIAOBING FENG‡, AND ANDREAS PROHL§

Abstract. Three fully discrete finite element methods are developed for approximating wave maps into the sphere based on two different approaches. The first method is an explicit scheme and the numerical solution satisfies the sphere-constraint exactly at every node. The second and third methods are implicit schemes which are based on a penalization approach, their numerical solutions satisfy the sphere-constraint approximately, and the quality of approximations is controlled by a small penalization parameter. Discrete energy conservation laws which mimic the underlying differential conservation law are established, and convergence of all proposed methods is proved. Computational experiments are also provided to validate the proposed methods and to present numerical evidence for possible finite-time blow-ups of the wave maps.

Key words. wave maps, penalization, finite element method, fully discrete scheme, convergence analysis

AMS subject classifications. 65M12, 65M60, 83C35

DOI. 10.1137/060659971

1. Introduction.

Let $\mathbf{R} \times \mathbf{R}^m$ for $m \geq 1$ denote the $(m+1)$-dimensional Minkowski space endowed with the Minkowski metric $\eta = (\eta_{\alpha\beta}) = \eta^{-1} = (\eta^{\alpha\beta}) = \mathrm{diag}(-1, 1, 1, \ldots, 1)$, and let $(\mathbf{N}, g)$ be an n-dimensional Riemannian manifold. A map $\mathbf{u} : (t, \mathbf{x}) \in \mathbf{R} \times \mathbf{R}^m \to \mathbf{u}(t, \mathbf{x}) \in \mathbf{N}$ is called a *wave map* if it satisfies the *wave map equation* (cf. [20, 25])

$$\Box \mathbf{u} = A(\mathbf{u})(\partial_\alpha \mathbf{u}, \partial^\alpha \mathbf{u}) \perp T_{\mathbf{u}}\mathbf{N}, \tag{1.1}$$

where $\partial^\alpha \mathbf{u} = \eta^{\alpha\beta}\partial_\beta \mathbf{u}$, and $\Box$ is the d'Alembert operator, $\Box = \partial_\alpha \partial^\alpha$. Moreover, $A(\mathbf{u})(\cdot, \cdot)$ denotes the second fundamental form of the manifold $\mathbf{N}$, and $T_{\mathbf{u}}\mathbf{N}$ denotes the tangent space of $\mathbf{N}$ at $\mathbf{u}$.

It is well known (cf. [20]) that (1.1) can also be interpreted as the Euler–Lagrange equation for the Lagrangian

$$\mathcal{L}(\mathbf{u}) := \frac{1}{2} \int_{\mathbf{R} \times \mathbf{R}^m} \langle \partial_\alpha \mathbf{u}, \partial^\alpha \mathbf{u} \rangle_g \, \mathrm{d}x\mathrm{d}t, \tag{1.2}$$

where all the first order derivatives of $\mathbf{u}$ take values in the tangent space $T\mathbf{N}$ of $\mathbf{N}$, that is,

$$\partial_\alpha \mathbf{u} : \mathbf{R} \times \mathbf{R}^m \to T_{\mathbf{u}}\mathbf{N}, \qquad \alpha = t, x_1, x_2, \ldots, x_m.$$

*Received by the editors May 15, 2006; accepted for publication (in revised form) May 11, 2007; published electronically December 19, 2007.
http://www.siam.org/journals/sinum/46-1/65997.html

†Department of Mathematics, Humboldt-Universität zu Berlin, Unter den Linden 6, D-10099 Berlin, Germany (sba@math.hu-berlin.de). This author was supported by the Deutsche Forschungsgemeinschaft through the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin. Part of this work was done when this author visited Forschungsinstitut für Mathematik (ETH Zürich) in February, 2006.

‡Department of Mathematics, The University of Tennessee, Knoxville, TN 37996 (xfeng@math.utk.edu). The work of this author was partially supported by NSF grant DMS-0410266.

§Mathematisches Institut, Universität Tübingen, Auf der Morgenstelle 10, D-72076 Tübingen, Germany (prohl@na.uni-tuebingen.de).

The wave map equation is the simplest geometric nonlinear wave equation, thus providing a good setup for studying nonlinear wave interactions. Wave maps have many applications in physics; they arise as harmonic gauges in general relativity (cf. [10]) and as the nonlinear σ-models in particle physics (cf. [1]); under suitable hypotheses, the Einstein vacuum equations with cylindrical symmetry reduce to a wave map system on $(2+1)$-dimensional Minkowski space-time (cf. [7]). Moreover, wave maps also arise in the analysis of the more difficult hyperbolic Yang–Mills equations either as special cases or as equations for certain families of gauge transformations (cf. [11]).

In this paper we consider wave maps in the case when the target manifold $\mathbf{N}$ is an $(n-1)$-sphere, i.e., $\mathbf{N} = \mathbf{S}^{n-1} \subset \mathbf{R}^n$ $(n \geq 1)$. In this case, we have (cf. [20])

$$\begin{aligned} \Box &= \partial_t^2 - \Delta, \\ \langle \partial_\alpha \mathbf{u}, \partial^\alpha \mathbf{u} \rangle_g &= |\nabla \mathbf{u}|^2 - |\partial_t \mathbf{u}|^2, \\ A(\mathbf{u})(\partial_\alpha \mathbf{u}, \partial^\alpha \mathbf{u}) &= \big(|\,\nabla \mathbf{u}\,|^2 - |\,\partial_t \mathbf{u}\,|^2\big)\mathbf{u}, \end{aligned}$$

where $\nabla \mathbf{u} = (\partial_{x_1}, \ldots, \partial_{x_m})$, and Δ is the (spatial) Laplacian. Specifically, we shall study the following initial-boundary value problem.

Given $T > 0$ and $(\mathbf{u}_0(\mathbf{x}), \mathbf{u}_1(\mathbf{x})) \in \mathbf{S}^{n-1} \times T_{\mathbf{u}_0(\mathbf{x})}\mathbf{S}^{n-1}$ for all $\mathbf{x} \in \Omega$, find a wave map $\mathbf{u} : \mathbf{R} \times \mathbf{R}^m \to \mathbf{S}^{n-1}$ such that

$$\Box \mathbf{u} = \big(|\,\nabla \mathbf{u}\,|^2 - |\,\partial_t \mathbf{u}\,|^2\big)\mathbf{u} \qquad \text{in } \Omega_T := (0,T) \times \Omega\,, \tag{1.3}$$

$$\frac{\partial \mathbf{u}}{\partial \mathbf{n}} = 0 \qquad \text{on } \partial\Omega_T := (0,T) \times \partial\Omega\,, \tag{1.4}$$

$$\mathbf{u}(0,\cdot) = \mathbf{u}_0, \qquad \partial_t \mathbf{u}(0,\cdot) = \mathbf{u}_1 \qquad \text{in } \Omega\,, \tag{1.5}$$

where $\Omega \subset \mathbf{R}^m$ is a bounded domain. Here, $\mathbf{n}$ and $\frac{\partial \mathbf{u}}{\partial \mathbf{n}}$ denote the outward unit normal to, and the normal derivative of $\mathbf{u}$ on the boundary $\partial\Omega$, of Ω.

It is well known that the wave map equation (1.1) is *integrable* (so are many other wave equations), which implies that the solutions of (1.1) satisfy some conservation law, which is, for wave maps with the sphere target manifold (cf. [20]),

$$E\big(\partial_t \mathbf{u}(t,\cdot), \mathbf{u}(t,\cdot)\big) = E\big(\mathbf{u}_1(\cdot), \mathbf{u}_0(\cdot)\big) \qquad \forall\, t \geq 0, \tag{1.6}$$

where

$$E\big(\partial_t \mathbf{u}(t,\cdot), \mathbf{u}(t,\cdot)\big) := \frac{1}{2}\Big[\|\,\partial_t \mathbf{u}(t,\cdot)\,\|_{L^2}^2 + \|\,\nabla \mathbf{u}(t,\cdot)\,\|_{L^2}^2\Big]. \tag{1.7}$$

Wave map equation (1.1) has been extensively studied in the past twenty years from a theoretical point of view. It was shown that on a $(1+1)$-dimensional Minkowski space-time base the equation has global in time smooth solutions if the initial data are smooth. It has also been known that for $m \geq 3$ the wave maps on an $(m+1)$-dimensional Minkowski space-time base can blow up in finite time in the sense that $\lim_{t \to t^*} \|\nabla \mathbf{u}(\cdot, t)\|_{L^\infty} = \infty$ for some $t^* < \infty$. On the other hand, whether the wave map equation has global in time smooth solutions on a $(2+1)$-dimensional Minkowski space-time base remains an open problem; see Open Problems 5 and 6 in [25]. The dimension analysis identifies $(2+1)$ as the critical dimension for wave maps, and so there has been a considerable amount of interest in determining whether $(2+1)$-dimensional wave maps develop singularities in finite time for smooth initial data. Numerical experiments of [8, 17] suggest that finite-time collapses, which are identified as local concentrations of energy, may occur for the special class of spherically equivariant wave maps in the case of large, smooth initial data with the spheres as target manifolds. For more discussions and up-to-date analytical results for wave maps, we refer the reader to the recent monograph of Shatah and Struwe [20] and the survey paper of Tataru [25].

Numerically, besides the works of [8, 17], no other results are known in the literature. However, despite the strong evidence of finite-time singularities in [8, 17], the numerical schemes proposed there are straightforward discretizations of reduced models and are not supported by rigorous numerical analysis, so unphysical numerical artifacts cannot be completely ruled out. The primary goal of this paper is to develop various finite element methods for approximating the initial-boundary value problem (1.3)–(1.5) based on two different approaches (see below and section 3). Rigorous convergence analysis will be provided, so our numerical solutions are guaranteed to converge to weak solutions of (1.3)–(1.5).

Like other well-known maps such as harmonic maps (cf. [4, 5, 22] and references therein), the main difficulty in analyzing and approximating wave maps is caused by the nonflatness of the target manifold $\mathbf{N}$, which is equivalent to imposing some (usually nonconvex) constraints on the solutions of the wave map equation, and is the reason for the nonlinearities. For example, when $\mathbf{N} = \mathbf{S}^{n-1}$, which is the focus of this paper, the constraint is $|\mathbf{u}| = 1$. Recall that the wave map equation (1.1) can be regarded as the Euler–Lagrange equation for the Lagrangian $\mathcal{L}$ defined in (1.2), interpreting this difficulty in the calculus of variation context, which means that the wave map problem is a variation problem with a holonomic constraint (cf. [16]).

The strong nonlinearity and nonconvex constraint make it very difficult to develop convergent numerical methods for (1.3)–(1.5). To the best of our knowledge, no such numerical method is known in the literature up to now. To overcome this difficulty, there are two approaches one can try. On the one hand, one may want to approximate (1.3)–(1.5) directly. To this end, one has to maintain the constraint $|\mathbf{u}| = 1$ at the discrete level, which is difficult to accomplish. Usually, a projection technique is employed for this purpose. Unfortunately, straightforward strategies do not give (discrete) energy conservation laws; consequently, the methods may not be convergent. On the other hand, one can approximate (1.3)–(1.5) indirectly. This approach often involves two steps: first, the wave map equation is penalized and the nonconvex constraint is relaxed by a singular perturbation technique; second, the penalized and nonconstrained problem is approximated by various numerical methods such as finite element, finite difference, and spectral methods. The trade-off is that one now needs to approximate singularly perturbed equations and to resolve the solutions on a very small scale introduced by the penalization parameter and to control numerical parameters according to the penalization parameter for both stability and convergence concerns.

In this paper, we shall develop numerical methods using both direct and indirect approaches. Our direct methods are inspired by the works of [2, 4, 6], where constraint-preserving finite element methods were developed respectively for the Landau–Lifshitz equation and the p-harmonic map heat flow. The direct methods are constructed based on the following equivalent reformulation of (1.3): find $\mathbf{u} : \Omega_T \to \mathbf{S}^{n-1}$ such that for all $\mathbf{w} \in C_0^\infty\big([0,T) \times \Omega, \mathbf{R}^n\big)$ with $\langle \mathbf{u}, \mathbf{w} \rangle = 0$ a.e. in Ω_T there holds

$$\int_0^T \big(\partial_t^2 \mathbf{u}, \mathbf{w}\big)\, \mathrm{d}t + \int_0^T \big(\nabla \mathbf{u}, \nabla \mathbf{w}\big)\, \mathrm{d}t = 0. \tag{1.8}$$

Our indirect methods are based on the following Ginzburg–Landau-type penalization ansatz (cf. [20]):

$$\Box \mathbf{u}^\varepsilon + \frac{1}{\varepsilon}\big(|\,\mathbf{u}^\varepsilon\,|^2 - 1\big)\mathbf{u}^\varepsilon = 0 \qquad \text{in } \Omega_T\,, \tag{1.9}$$

$$\frac{\partial \mathbf{u}^\varepsilon}{\partial \mathbf{n}} = 0 \qquad \text{on } \partial\Omega_T\,, \tag{1.10}$$

$$\mathbf{u}^\varepsilon(0,\cdot) = \mathbf{u}_0\,, \qquad \partial_t \mathbf{u}^\varepsilon(0,\cdot) = \mathbf{u}_1 \qquad \text{in } \Omega \tag{1.11}$$

for $0 < \varepsilon << 1$. It was shown that for each $\varepsilon > 0$ the regularized problem (1.9)–(1.11) has a weak solution $\mathbf{u}^\varepsilon : \Omega_T \to \mathbf{R}^n$, which satisfies the following conservation law (hence, (1.9)–(1.11) is also integrable):

$$E\big(\partial_t \mathbf{u}^\varepsilon(t,\cdot), \mathbf{u}^\varepsilon(t,\cdot)\big) + \frac{1}{\varepsilon} F(\mathbf{u}^\varepsilon(t,\cdot)) = E\big(\mathbf{u}_1(\cdot), \mathbf{u}_0(\cdot)\big) \qquad \forall\, t \in [0,T], \tag{1.12}$$

where

$$F(\mathbf{w}) = \frac{1}{4}\int_\Omega \big(|\,\mathbf{w}\,|^2 - 1\big)^2 \,\mathrm{d}\mathbf{x}. \tag{1.13}$$

Moreover, there exists a subsequence of $\{\mathbf{u}^\varepsilon\}$ which converges to a map $\mathbf{u}$, which turns out to be a weak solution of (1.3)–(1.5) as $\varepsilon \to 0^+$.

The remainder of this paper is organized as follows. In section 2 we introduce some notation and a weak formulation of (1.3)–(1.5). In section 3 we present our fully discrete finite element methods based on both direct and indirect approaches. The convergence analysis of the proposed numerical methods is given in sections 4 and 5. In section 6 we present several numerical experiment results and provide numerical evidence for finite-time (finite-energy) blow-ups of wave maps in the critical dimension $(2+1)$. Finally, we finish the paper with a few concluding remarks.

2. Preliminaries. Standard notations are adopted throughout this paper. $\langle\cdot,\cdot\rangle$ denotes the standard inner product of the Euclidean space $\mathbf{R}^m$, and $(\cdot,\cdot) := (\cdot,\cdot)_\Omega$ is the standard L^2-inner product over the domain Ω. $W^{m,p}(\Omega,\mathbf{R}^n)$ denotes the (m,p)-Sobolev space of vector-valued functions, and $\|\cdot\|_{W^{m,p}}$ denotes its norm. Note that $W^{0,p}(\Omega,\mathbf{R}^n) = L^p(\Omega,\mathbf{R}^n)$. Throughout this paper, $C > 0$ is used to denote a generic positive $(\,h,\tau,\varepsilon\,)$-independent constant which may take different values at different locations. We also introduce $\mathbf{u}_t := \partial_t \mathbf{u}, \nabla\mathbf{u} := (\partial_{x_1},\ldots,\partial_{x_m})$, and $D := (\partial_t, \nabla)$ and define the nonlinear Sobolev space

$$W^{1,2}(\Omega,\mathbf{S}^{n-1}) = \big\{\mathbf{v} \in W^{1,2}(\Omega,\mathbf{R}^n);\ \mathbf{v} \in \mathbf{S}^{n-1} \text{ a.e. in } \Omega\big\}.$$

We now give a definition of weak solutions to (1.3)–(1.5).

Definition 2.1. *Given $T > 0$ and $(\,\mathbf{u}_0,\mathbf{u}_1\,) \in W^{1,2}(\Omega,\mathbf{S}^{n-1}) \times L^2(\Omega, T\mathbf{S}^{n-1})$, we call $\mathbf{u} : \Omega_T \to \mathbf{S}^{n-1}$ a weak solution of* (1.3)–(1.5) *if the following hold:*

(1) $D\mathbf{u} \in L^2(\Omega_T,\mathbf{R}^n)$.
(2) $|\,\mathbf{u}\,| = 1$ *a.e. in* Ω_T.
(3) *Equation* (1.3) *holds (in wedged form) in the distributional sense; i.e., for all* $\boldsymbol{\phi} \in C_0^\infty\big([0,T);\, W^{1,2}(\Omega,\mathbf{R}^n)\big)$ *we have*

$$-\int_0^T \big(\mathbf{u}_t \wedge \mathbf{u}, \boldsymbol{\phi}_t\big)\,\mathrm{d}t + \int_0^T \big(\nabla\mathbf{u} \wedge \mathbf{u}, \nabla\boldsymbol{\phi}\big)\,\mathrm{d}t = \big(\mathbf{u}_1 \wedge \mathbf{u}_0, \boldsymbol{\phi}(0,\cdot)\big). \tag{2.1}$$

(4) *The initial conditions* $\mathbf{u}_0$ *and* $\mathbf{u}_1$ *are continuously attained by* $\mathbf{u}$ *and* $\mathbf{u}_t$ *in* $W^{1,2}(\Omega;\mathbf{R}^n)$ *and* $L^2(\Omega;\mathbf{R}^n)$*; i.e., there hold as* $t\to 0$

$$\mathbf{u}(t,\cdot)\to\mathbf{u}_0 \quad \text{in } W^{1,2}(\Omega;\mathbf{R}^n), \qquad \mathbf{u}_t(t,\cdot)\to\mathbf{u}_1 \quad \text{in } L^2(\Omega;\mathbf{R}^n).$$

(5) *The following energy inequality holds:*

$$E\big(\mathbf{u}_t(t,\cdot),\mathbf{u}(t,\cdot)\big)\le E\big(\mathbf{u}_1,\mathbf{u}_0\big) \qquad \forall\, t\in[0,T]\,. \tag{2.2}$$

Existence of weak solutions to (1.3)–(1.5) for $m\ge 1$ was proved in [19] by the Ginzburg–Landau penalization technique, and weak solutions were constructed as proper limits of $\{\mathbf{u}_\varepsilon\}$ which solve (1.9)–(1.11). Another method, which was proposed in [18], constructs weak solutions for general targets $\mathbf{N}$ as limits of convergent sequences of spatially discrete wave maps. The semidiscrete maps are direct approximations of the original wave maps and are constructed by (abstract) finite difference methods, so no penalization is involved. We remark that weak solutions to (1.3)–(1.5) are not unique in general (cf. [20]).

3. Fully discrete finite element methods.

3.1. Notation. For simplicity, let Ω be a bounded polygonal (when $m=2$) or polyhedral (when $m=3$) domain. Let $\mathcal{T}_h$ denote a quasiuniform triangulation of Ω into triangles or tetrahedrons with mesh-size $h>0$ for $m=2$ or $m=3$, respectively. For a domain $A\subset\mathbf{R}^m$, let $\mathcal{P}(A,\mathbf{R}^n)$ stand for the set of all polynomials on A of degree ≤ 1. We define the Lagrange finite element space

$$\boldsymbol{\mathcal{V}}_h := \big\{\mathbf{w}\in C(\overline{\Omega},\mathbf{R}^n);\ \mathbf{w}|_K\in\mathcal{P}(K,\mathbf{R}^n)\ \forall\, K\in\mathcal{T}_h\big\}\,.$$

Let $\mathcal{N}_h$ denote the set of all nodes associated with the finite element space $\boldsymbol{\mathcal{V}}_h$, and let $\big\{\varphi_{\mathbf{q}_i};\ \mathbf{q}_i\in\mathcal{N}_h\big\}$ denote the nodal basis for $\boldsymbol{\mathcal{V}}_h$; we define the following nodal interpolation operator $I_h: C(\overline{\Omega},\mathbf{R}^n)\to\boldsymbol{\mathcal{V}}_h$ by

$$I_h\mathbf{w} := \sum_{\mathbf{q}_i\in\mathcal{N}_h}\mathbf{w}(\mathbf{q}_i)\varphi_{\mathbf{q}_i} \qquad \forall\,\mathbf{w}\in C(\overline{\Omega},\mathbf{R}^n).$$

For any two functions $\mathbf{v},\mathbf{w}\in C(\overline{\Omega},\mathbf{R}^n)$ we define a discrete L^2-inner product by

$$\big(\mathbf{v},\mathbf{w}\big)_h := \int_\Omega I_h\big(\langle\mathbf{v},\mathbf{w}\rangle\big)\,\mathrm{d}\mathbf{x} = \sum_{\mathbf{q}_i\in\mathcal{N}_h}\beta_{\mathbf{q}_i}\langle\mathbf{v}(\mathbf{q}_i),\mathbf{w}(\mathbf{q}_i)\rangle\,,$$

where $\beta_{\mathbf{q}_i}=\int_\Omega\varphi_{\mathbf{q}_i}\,\mathrm{d}\mathbf{x}$ for all $\mathbf{q}_i\in\mathcal{N}_h$. We also define $||\mathbf{w}||_h:=(\mathbf{w},\mathbf{w})_h^{\frac12}$. It is easy to check that there hold for all $\mathbf{v}_h,\mathbf{w}_h\in\boldsymbol{\mathcal{V}}_h$

$$||\mathbf{w}_h||_{L^2}\le||\mathbf{w}_h||_h\le(m+2)^{\frac12}||\mathbf{w}_h||_{L^2}, \tag{3.1}$$
$$\big|(\mathbf{v}_h,\mathbf{w}_h)_h-(\mathbf{v}_h,\mathbf{w}_h)\big|\le Ch\,||\mathbf{v}_h||_{L^2}\,||\nabla\mathbf{w}_h||_{L^2}. \tag{3.2}$$

Let τ_1 and τ be two (small) positive numbers, and

$$t_0:=0,\ t_1:=\tau_1,\ t_j:=t_1+(j-1)\tau, \quad j=1,2,\dots,J:=\left[\frac{T-t_1}{\tau}\right].$$

Then $\{t_j\}$ forms a partition of the interval $[0,T]$. We introduce the following difference operators (cf. [13, 14]):

$$
\begin{aligned}
&d_t\mathbf{v}^1 := \frac{\mathbf{v}^1-\mathbf{v}^0}{\tau_1}, \qquad \delta_t\mathbf{v}^1 := \frac{\mathbf{v}^2-\mathbf{v}^0}{\tau_1+\tau},\\
&d_t\mathbf{v}^j := \frac{\mathbf{v}^j-\mathbf{v}^{j-1}}{\tau}, \qquad \delta_t\mathbf{v}^j := \frac{\mathbf{v}^{j+1}-\mathbf{v}^{j-1}}{2\tau}, \quad j\ge 1,\\
&\mathbf{v}^{j\pm\frac12} := \frac{\mathbf{v}^{j\pm1}+\mathbf{v}^j}{2}, \quad \mathbf{v}^{j,\theta} := \theta\mathbf{v}^{j+1}+(1-2\theta)\mathbf{v}^j+\theta\mathbf{v}^{j-1} \qquad \forall\,\theta\in\left[0,\frac12\right].
\end{aligned}
$$

Given a sequence $\{\mathbf{w}^j\}_{1\le j\le J}\in\ell^\infty\big(0,J;\boldsymbol{\mathcal{V}}_h\big)$, we define the following constant and linear interpolations of the sequence in time: for every $t\in(t_{j-1},t_j]$ set

$$
\begin{aligned}
&\mathbf{w}^-(t,\cdot) := \mathbf{w}^{j-1}(\cdot)\,, &&\mathbf{w}^+(t,\cdot) := \mathbf{w}^j(\cdot)\,,\\
&\overline{\mathbf{w}}(t,\cdot) := \frac12\big[\mathbf{w}^{j-1}(\cdot)+\mathbf{w}^j(\cdot)\big]\,, &&\mathbf{w}(t,\cdot) := \frac{t-t_{j-1}}{\tau}\mathbf{w}^j(\cdot)+\frac{t_j-t}{\tau}\mathbf{w}^{j-1}(\cdot)\,.
\end{aligned}
$$

We note that τ needs to be replaced by τ_1 in the definition of $\mathbf{w}(t,\cdot)$ when $j=1$. It is easy to check that

$$
\|\,\mathbf{w}^\pm-\mathbf{w}\,\|+\|\,\overline{\mathbf{w}}-\mathbf{w}\,\|\le 3\tau\,\|\,d_t\mathbf{w}\,\|. \tag{3.3}
$$

Hence, $\mathbf{w}^\pm,\overline{\mathbf{w}}$, and $\mathbf{w}$ converge simultaneously to the same limit provided that $\|\,d_t\mathbf{w}\,\|$ is bounded and one of the sequences converges. Finally, we introduce the following subsets of $\boldsymbol{\mathcal{V}}_h$:

$$
\begin{aligned}
\boldsymbol{\mathcal{M}}_h &:= \big\{\mathbf{w}\in\boldsymbol{\mathcal{V}}_h;\,|\,\mathbf{w}(\mathbf{q}_i)\,|=1\ \forall\,\mathbf{q}_i\in\mathcal{N}_h\big\}\,,\\
\boldsymbol{\mathcal{F}}_h(\boldsymbol{\chi}) &:= \big\{\mathbf{w}\in\boldsymbol{\mathcal{V}}_h;\,\langle\mathbf{w}(\mathbf{q}_i),\boldsymbol{\chi}(\mathbf{q}_i)\rangle=0\quad\forall\,\mathbf{q}_i\in\mathcal{N}_h\big\}\,,\quad\text{where }\boldsymbol{\chi}\in\boldsymbol{\mathcal{M}}_h\,.
\end{aligned}
$$

3.2. Fully discrete methods for (1.3)–(1.5). We introduce a family of fully discrete *explicit* finite element methods for (1.3)–(1.5). The starting point is the weak formulation (1.8). To discretize (1.8), the following fully discrete semi-implicit approximation of (1.8) is immediate: For $j=1\to(J-1)$, given $\hat{\mathbf{U}}^{j-1},\hat{\mathbf{U}}^j\in\boldsymbol{\mathcal{M}}_h$, find $\hat{\mathbf{U}}^{j+1}\in\boldsymbol{\mathcal{M}}_h$ such that

$$
(d_t^2\hat{\mathbf{U}}^{j+1},\mathbf{w})+(\nabla\hat{\mathbf{U}}^{j+1},\nabla\mathbf{w})=0 \qquad \forall\,\mathbf{w}\in\boldsymbol{\mathcal{F}}_h(\hat{\mathbf{U}}^j)\,, \tag{3.4}
$$

and $\hat{\mathbf{U}}^0,\hat{\mathbf{U}}^1\in\boldsymbol{\mathcal{M}}_h$, where $d_t^2\varphi^j=d_t(d_t\varphi^j)$. However, (3.4) is difficult to solve because of the nonconvex constraint on $\boldsymbol{\mathcal{M}}_h$. To overcome this difficulty, on noting $\langle\mathbf{u}_t,\mathbf{u}\rangle=0$ we may assume that $d_t\hat{\mathbf{U}}^{j+1}$ is almost an element of $\boldsymbol{\mathcal{F}}_h(\hat{\mathbf{U}}^j)$. This motivates the following explicit scheme for (1.8).

ALGORITHM 3.1.

(i) *Choose* $(\mathbf{U}^0,\mathbf{V}^0)\in\boldsymbol{\mathcal{M}}_h\times\boldsymbol{\mathcal{V}}_h$ *as a suitable approximation to* $(\mathbf{u}_0,\mathbf{u}_1)$.

(ii) *For* $j=0\to(J-1)$, *given* $(\mathbf{U}^j,\mathbf{V}^j)\in\boldsymbol{\mathcal{M}}_h\times\boldsymbol{\mathcal{V}}_h$, *find* $\mathbf{V}^{j+1}\in\boldsymbol{\mathcal{F}}_h(\mathbf{U}^j)$ *such that*

$$
(d_t\mathbf{V}^{j+1},\mathbf{w})_h=(\nabla\mathbf{U}^j,\nabla\mathbf{w}) \qquad \forall\,\mathbf{w}\in\boldsymbol{\mathcal{F}}_h(\mathbf{U}^j)\,,
$$

and define $\mathbf{U}^{j+1}\in\boldsymbol{\mathcal{M}}_h$ *by*

$$
\mathbf{U}^{j+1}(\mathbf{q}_i)=\frac{\mathbf{U}^j(\mathbf{q}_i)+\tau\mathbf{V}^{j+1}(\mathbf{q}_i)}{|\,\mathbf{U}^j(\mathbf{q}_i)+\tau\mathbf{V}^{j+1}(\mathbf{q}_i)\,|} \qquad \forall\,\mathbf{q}_i\in\mathcal{N}_h\,.
$$

Convergence analysis of Algorithm 3.1 will be given in section 4.

Remark 3.1. (a) The auxiliary function $\mathbf{V}^j$ is an approximation to $\mathbf{u}_t(t_j,\cdot)$, so the algorithm may be viewed as a mixed method which approximates $\mathbf{u}$ and $\mathbf{u}_t$ simultaneously.

(b) The last equation in the algorithm combines a discretization of $\mathbf{u}_t = \mathbf{v}$ with a projection to the sphere at the nodes.

(c) Note that owing to $\mathbf{U}^j(\mathbf{q}_i)\cdot\mathbf{V}^{j+1}(\mathbf{q}_i) = 0$ and $|\mathbf{U}^j(\mathbf{q}_i)| = 1$ for all $\mathbf{q}_i \in \mathcal{N}_h$ the projection step is well defined.

3.3. Fully discrete methods for (1.9)–(1.11). We construct two families of fully discrete *implicit* finite element methods for (1.9)–(1.11). Since the problem (1.9)–(1.11) converges to the problem (1.3)–(1.5) as $\varepsilon \to 0^+$ (cf. [20]), the methods constructed in this subsection are indirect methods for approximating the problem (1.3)–(1.5).

Because for each fixed $\varepsilon > 0$ (1.9) is a semilinear system of wave equations, a natural strategy is to adapt some of best known numerical schemes developed for the linear wave equation to the problem (1.9)–(1.11). Indeed, two families of such schemes will be considered in this subsection. The first family of schemes motivated by the work of Baker [3] are one-step implicit mixed methods in the sense of Remark 3.1 (a); the second family motivated by the work of Dupont [13] are two-step implicit methods.

Apart from the similarities of numerical methods to be given below and those developed in [3, 13], we emphasize that, in addition to the extra difficulty caused by the nonlinearity in (1.9), we also have to cope with a new scale which is due to the introduction of the penalization parameter ε. In particular, we need to make sure our numerical solutions satisfy a discrete conservation law which mimics the differential conservation law (1.12), and they are robust with respect to ε.

First, adapting the schemes of [3], we introduce the following fully discrete implicit discretization of (1.9)–(1.11).

ALGORITHM 3.2.

(i) *Choose* $(\mathbf{U}^0, \mathbf{V}^0) \in \big[\boldsymbol{\mathcal{V}}_h\big]^2$ *as a suitable approximation to* $(\,\mathbf{u}_0, \mathbf{u}_1\,)$.

(ii) *For* $j = 0 \to (J-1)$, *given* $(\mathbf{U}^j, \mathbf{V}^j) \in \big[\boldsymbol{\mathcal{V}}_h\big]^2$, *find* $(\mathbf{U}^{j+1}, \mathbf{V}^{j+1}) \in \big[\boldsymbol{\mathcal{V}}_h\big]^2$ *such that for all* $(\mathbf{w}, \boldsymbol{\chi}) \in [\boldsymbol{\mathcal{V}}_h]^2$

$$
\begin{aligned}
&(d_t\mathbf{V}^{j+1}, \mathbf{w})_h + (\nabla\mathbf{U}^{j+\frac{1}{2}}, \nabla\mathbf{w}) + \frac{1}{2\varepsilon}\Big(\big(|\,\mathbf{U}^{j+1}\,|^2 + |\,\mathbf{U}^j\,|^2 - 2\big)\mathbf{U}^{j+\frac{1}{2}}, \mathbf{w}\Big)_h = 0\,,\\
&(d_t\mathbf{U}^{j+1}, \boldsymbol{\chi})_h = (\mathbf{V}^{j+\frac{1}{2}}, \boldsymbol{\chi})_h\,.
\end{aligned}
$$

Next, adapting the schemes of [13], we introduce our second family of fully discrete implicit methods for (1.9)–(1.11).

ALGORITHM 3.3.

(i) *Choose* $\mathbf{U}^0, \mathbf{U}^1 \in \boldsymbol{\mathcal{V}}_h$ *as suitable approximations to* $\mathbf{u}_0$ *and* $\mathbf{u}(t_1,\cdot)$.

(ii) *For* $j = 1 \to (J-1)$, *given* $\{\mathbf{U}^i\}_{i=0}^j \subset \boldsymbol{\mathcal{V}}_h$, *find* $\mathbf{U}^{j+1} \in \boldsymbol{\mathcal{V}}_h$ *such that for all* $\mathbf{w} \in \boldsymbol{\mathcal{V}}_h$

$$
(d_t^2\mathbf{U}^{j+1}, \mathbf{w})_h + (\nabla\mathbf{U}^{j,\frac{1}{4}}, \nabla\mathbf{w}) + \frac{1}{2\varepsilon}\Big(\big(|\,\mathbf{U}^{j+\frac{1}{2}}\,|^2 + |\,\mathbf{U}^{j-\frac{1}{2}}\,|^2 - 2\big)\mathbf{U}^{j,\frac{1}{4}}, \mathbf{w}\Big)_h = 0\,.
$$

Convergence analysis of Algorithms 3.2 and 3.3 will be presented in section 5. We note that in the above equation $\mathbf{U}^{j,\frac{1}{4}}$ could be replaced by $\mathbf{U}^{j,\theta}$ for any $\theta \in [0, \frac{1}{2}]$. However, as pointed out in [13], although for $\theta \geq \frac{1}{4}$ the analogues of the above scheme are all absolutely stable, the time truncation error is minimized over this class at $\theta = \frac{1}{4}$, which is the main reason for the choice of $\theta = \frac{1}{4}$.

4. Convergence analysis for Algorithm 3.1. The main goal of this section is to prove that the solution sequence $\{\mathbf{U}^j\}$ of Algorithm 3.1 has a convergent subsequence whose limit is a weak solution of (1.3)–(1.5). To this end, we first need to establish some energy estimates for the solution sequence $\{\mathbf{U}^j\}$. Throughout this section we assume $\tau_1 = \tau$.

4.1. Stability. Recall that by definition every weak solution of (1.3)–(1.5) must satisfy the energy inequality (2.2). In what follows, we show that the solution sequence $\{(\mathbf{U}^j, \mathbf{V}^j)\}$ generated by Algorithm 3.1 satisfies a discrete energy inequality, which mimics (2.2), provided that the temporal mesh-size τ and the spatial mesh-size h satisfy the relation $\tau = o(h^{\frac{m+4}{3}})$. Specifically, we have the following lemma.

LEMMA 4.1. *Let $m \geq 1$ and $(\mathbf{U}^0, \mathbf{V}^0) \in \boldsymbol{\mathcal{M}}_h \times \boldsymbol{\mathcal{F}}_h(\mathbf{U}^0)$. Then, under the mesh condition $\tau = o(h^{\frac{4+m}{3}})$, the solution sequence $\{(\mathbf{U}^j, \mathbf{V}^j)\}_{j=0}^J$ of Algorithm 3.1 satisfies*

$$\begin{aligned} E_h\big(\mathbf{V}^{j+1}, \mathbf{U}^{j+1}\big) + \frac{\tau^2}{2} \sum_{\ell=0}^{j} \big[\| d_t \mathbf{V}^{\ell+1} \|_h^2 + \| \nabla d_t \mathbf{U}^{\ell+1} \|_{L^2}^2\big] \\ = c_0 \, E_h\big(\mathbf{V}^0, \mathbf{U}^0\big) \end{aligned} \tag{4.1}$$

with a positive constant $c_0 = 1 + o(1)$ and where

$$E_h(\mathbf{V}^j, \mathbf{U}^j) := \frac{1}{2}\big[\, \| \mathbf{V}^j \|_h^2 + \| \nabla \mathbf{U}^j \|_{L^2}^2 \,\big]. \tag{4.2}$$

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

Step 1 (auxiliary solution estimates). Setting $\mathbf{w} = \mathbf{V}^{j+1}$ in Step (ii) of Algorithm 3.1 and using the inverse inequality $\| \nabla \mathbf{v}_h \|_{L^2} \leq C h^{-1} \| \mathbf{v}_h \|_{L^2}$ yield

$$\frac{1}{2} d_t \| \mathbf{V}^{j+1} \|_h^2 + \frac{\tau}{2} \| d_t \mathbf{V}^{j+1} \|_h^2 = -(\nabla \mathbf{U}^j, \nabla \mathbf{V}^{j+1}) \leq C h^{-1} \, \| \nabla \mathbf{U}^j \|_{L^2} \| \mathbf{V}^{j+1} \|_{L^2} \,. \tag{4.3}$$

Next, we need to get an estimate for $d_t \| \nabla \mathbf{U}^{j+1} \|_{L^2}^2$, which in turn requires an estimate for $\| d_t \mathbf{U}^{j+1} \|_{L^2}^2$. We now derive these estimates by borrowing an argument from [2]. Let

$$\mathbf{R}^{j+1} := \mathbf{U}^{j+1} - \mathbf{U}^j - \tau \mathbf{V}^{j+1}.$$

Note that $\mathbf{R}^{j+1} \in \boldsymbol{\mathcal{V}}_h$ and $\tau^{-1}\mathbf{R}^{j+1}$ measures the discrepancy between $d_t \mathbf{U}^{j+1}$ and $\mathbf{V}^{j+1}$. By the definition of $\mathbf{U}^{j+1}$ in (ii) of Algorithm 3.1 we have

$$\begin{aligned} | \mathbf{R}^{j+1}(\mathbf{q}_i) | &= \left| \frac{\mathbf{U}^j(\mathbf{q}_i) + \tau \mathbf{V}^{j+1}(\mathbf{q}_i)}{| \mathbf{U}^j(\mathbf{q}_i) + \tau \mathbf{V}^{j+1}(\mathbf{q}_i) |} - \mathbf{U}^j(\mathbf{q}_i) - \tau \mathbf{V}^{j+1}(\mathbf{q}_i) \right| \\ &= \big| 1 - \big| \mathbf{U}^j(\mathbf{q}_i) + \tau \mathbf{V}^{j+1}(\mathbf{q}_i) \big| \big| \qquad \forall \, \mathbf{q}_i \in \mathcal{N}_h. \end{aligned}$$

Since $| \mathbf{U}^j(\mathbf{q}_i) + \tau \mathbf{V}^{j+1}(\mathbf{q}_i) | = \sqrt{1 + \tau^2 | \mathbf{V}^{j+1}(\mathbf{q}_i) |^2} \geq 1$, we conclude that

$$| \mathbf{R}^{j+1}(\mathbf{q}_i) | \leq \frac{\tau^2}{2} | \mathbf{V}^{j+1}(\mathbf{q}_i) |^2 \,. \tag{4.4}$$

Hence,

$$\int_\Omega |\,\mathbf{R}^{j+1}\,|\,\mathrm{d}\mathbf{x} \le \int_\Omega I_h\big[|\,\mathbf{R}^{j+1}\,|\big]\,\mathrm{d}\mathbf{x} \le \frac{\tau^2}{2}\int_\Omega I_h\big[|\,\mathbf{V}^{j+1}\,|^2\big]\,\mathrm{d}\mathbf{x} \le \frac{\tau^2(m+2)}{2}\int_\Omega |\,\mathbf{V}^{j+1}\,|^2\,\mathrm{d}\mathbf{x}\,. \tag{4.5}$$

Similarly,

$$\|\,\mathbf{R}^{j+1}\,\|_{L^2}^2 \le \|\,\mathbf{R}^{j+1}\,\|_h^2 \le \frac{\tau^4}{4}\|\,\mathbf{V}^{j+1}\,\|_{L^\infty}^2\|\,\mathbf{V}^{j+1}\,\|_h^2 \le C\tau^4 h^{-m}\,\|\,\mathbf{V}^{j+1}\,\|_{L^2}^4\,;$$

thus

$$\begin{aligned}\|\,d_t\mathbf{U}^{j+1}\,\|_{L^2}^2 &\le \big[\|\,\mathbf{V}^{j+1}\,\|_{L^2} + \tau^{-1}\,\|\,\mathbf{R}^{j+1}\,\|_{L^2}\big]^2 \\ &\le \big[1 + C\tau h^{-\frac{m}{2}}\,\|\,\mathbf{V}^{j+1}\,\|_{L^2}\big]^2\|\,\mathbf{V}^{j+1}\,\|_{L^2}^2\,.\end{aligned} \tag{4.6}$$

Now, setting $\mathbf{w} = \mathbf{V}^{j+1} = d_t\mathbf{U}^{j+1} - \tau^{-1}\mathbf{R}^{j+1}$ in Step (ii) of Algorithm 3.1, we get

$$(d_t\mathbf{V}^{j+1},\mathbf{V}^{j+1})_h + \big(\nabla\mathbf{U}^{j+1},\nabla d_t\mathbf{U}^{j+1}\big) = \tau^{-1}(\nabla\mathbf{U}^j,\nabla\mathbf{R}^{j+1}) + \tau\|\,\nabla d_t\mathbf{U}^{j+1}\,\|_{L^2}^2;$$

this, an application of the inverse inequality, and (4.6) yield

$$\begin{aligned}&\frac{1}{2}d_t\|\,\mathbf{V}^{j+1}\,\|_h^2 + \frac{1}{2}d_t\|\,\nabla\mathbf{U}^{j+1}\,\|_{L^2}^2 + \frac{\tau}{2}\|\,d_t\mathbf{V}^{j+1}\,\|_h^2 + \frac{\tau}{2}\|\,\nabla d_t\mathbf{U}^{j+1}\,\|_{L^2}^2 \\ &\quad\le C\tau^{-1}h^{-2}\,\|\,\mathbf{U}^j\,\|_{L^\infty}\|\,\mathbf{R}^{j+1}\,\|_{L^1} + C\tau h^{-2}\,\|\,d_t\mathbf{U}^{j+1}\,\|_{L^2}^2 \\ &\quad\le C\tau^2h^{-2}\,\|\,\mathbf{V}^{j+1}\,\|_h^2 + C\tau h^{-2}\,\|\,d_t\mathbf{U}^{j+1}\,\|_{L^2}^2 \\ &\quad\le C\tau h^{-2}\,\|\,\mathbf{V}^{j+1}\,\|_{L^2}^2 + C\tau h^{-2}\big[1+\tau h^{-\frac{m}{2}}\,\|\,\mathbf{V}^{j+1}\,\|_{L^2}\big]^2\|\,\mathbf{V}^{j+1}\,\|_{L^2}^2\,.\end{aligned} \tag{4.7}$$

Step 2 (finishing up). We now conclude the proof by an induction argument. The assumptions on $\mathbf{u}_0$ and $\mathbf{U}^0$ imply that there exists a positive constant C_1 independent of τ and h such that $\|\,\nabla\mathbf{U}^0\,\|_{L^2} \le C_1$. Assume that $\|\,\nabla\mathbf{U}^i\,\|_{L^2} \le C_1$ for all $0 \le i \le j$; it follows from (4.3) and a discrete version of Gronwall's lemma that

$$\|\,\mathbf{V}^{j+1}\,\|_{L^2} \le \|\,\mathbf{V}^{j+1}\,\|_h \le Ch^{-1}\,. \tag{4.8}$$

If we use this result in (4.7), together with the mesh condition, we have existence of another $C_2 = C_2(T) > 0$, such that $\tau h^{-2}\tau^2h^{-m}\,\|\,\mathbf{V}^{j+1}\,\|_{L^2}^2 \le C_2$. The discrete version of Gronwall's lemma then yields ($h \to 0$)

$$E_h\big(\mathbf{V}^{j+1},\mathbf{U}^{j+1}\big) + \frac{\tau^2}{2}\sum_{\ell=0}^{j}\big[\|\,d_t\mathbf{V}^{\ell+1}\,\|_h^2 + \|\,\nabla d_t\mathbf{U}^{\ell+1}\,\|_{L^2}^2\big] \le \big(1+o(1)\big)\,E_h\big(\mathbf{V}^0,\mathbf{U}^0\big).$$

Hence, by induction (4.1) holds for all $j = 0,1,\ldots,J$. The proof is complete. ∎

The following corollary follows immediately from Lemma 4.1 and (4.6).

COROLLARY 4.2. *The solution sequence* $\{\mathbf{U}^j\}$ *of Algorithm* 3.1 *also satisfies*

$$\|\,d_t\mathbf{U}^{j+1}\,\|_{L^2} \le C\,, \qquad j = 0,1,\ldots,J\,. \tag{4.9}$$

4.2. Convergence. To establish the convergence of Algorithm 3.1, we first need some preparations. Let $(\mathbf{U}^+_{\tau,h}, \mathbf{U}^-_{\tau,h}, \overline{\mathbf{U}}_{\tau,h}, \mathbf{U}_{\tau,h})$ and $(\mathbf{V}^+_{\tau,h}, \mathbf{V}^-_{\tau,h}, \overline{\mathbf{V}}_{\tau,h}, \mathbf{V}_{\tau,h})$ denote their respective constant and linear interpolations in time of the solution sequences $\{\mathbf{U}^j\}$ and $\{\mathbf{V}^j\}$ of Algorithm 3.1 as defined in section 3.1. However, for notational brevity, most of the time we omit subindices τ, h in this subsection and use $(\mathbf{U}^+, \mathbf{U}^-, \overline{\mathbf{U}}, \mathbf{U})$ and $(\mathbf{V}^+, \mathbf{V}^-, \overline{\mathbf{V}}, \mathbf{V})$ to stand for $(\mathbf{U}^+_{\tau,h}, \mathbf{U}^-_{\tau,h}, \overline{\mathbf{U}}_{\tau,h}, \mathbf{U}_{\tau,h})$ and $(\mathbf{V}^+_{\tau,h}, \mathbf{V}^-_{\tau,h}, \overline{\mathbf{V}}_{\tau,h}, \mathbf{V}_{\tau,h})$, respectively.

We first derive the following reformulation of Algorithm 3.1.

LEMMA 4.3. *Suppose that the assumptions of Lemma* 4.1 *are valid. There holds for all* $\mathbf{w} \in C_0^\infty\big([0,T); \boldsymbol{\mathcal{F}}_h(\mathbf{U}^-)\big)$

$$\begin{aligned}(4.10)\quad &\left| \int_0^T \big[-(\mathbf{U}_t, \mathbf{w}_t) + (\nabla \mathbf{U}, \nabla \mathbf{w})\big]\,\mathrm{d}t + \big(\mathbf{V}^0, \mathbf{w}(0,\cdot)\big) \right| \\ &\quad \le (C\tau h^{-\frac{m}{2}} + h) \left[\int_0^T \| \mathbf{w}_t \|_{L^2}^2 + \| \nabla \mathbf{w}_t \|_{L^2}^2 \,\mathrm{d}t + \| \nabla \mathbf{w}(0,\cdot) \|_{L^2}^2 \right]^{\frac{1}{2}} \\ &\quad\quad + \left| \int_0^T (\mathbf{V} - \mathbf{V}^+, \mathbf{w}_t)\,\mathrm{d}t \right| + \left| \int_0^T (\nabla[\mathbf{U} - \mathbf{U}^-], \nabla \mathbf{w})\,\mathrm{d}t \right| .\end{aligned}$$

Proof. Using the new notation we rewrite Step (ii) in Algorithm 3.1 as

$$\int_0^T \big[(\mathbf{V}_t, \mathbf{w})_h + (\nabla \mathbf{U}^-, \nabla \mathbf{w})\big]\,\mathrm{d}t = 0\,.$$

Integrating by parts in t in the first term yields

$$\int_0^T \big[-(\mathbf{V}, \mathbf{w}_t)_h + \big(\nabla \mathbf{U}^-, \nabla \mathbf{w}\big)\big]\,\mathrm{d}t + \big(\mathbf{V}^0, \mathbf{w}(0,\cdot)\big)_h = 0;$$

this and the relation $\mathbf{V}^+ = \mathbf{U}_t - \tau^{-1}\mathbf{R}^+$ imply that

$$\begin{aligned}(4.11)\quad &-\int_0^T \big[-(\mathbf{U}_t, \mathbf{w}_t) + (\nabla \mathbf{U}, \nabla \mathbf{w})\big]\,\mathrm{d}t - \big(\mathbf{V}^0, \mathbf{w}(0,\cdot)\big) \\ &\quad = \int_0^T \big[(\mathbf{V}, \mathbf{w}_t) - (\mathbf{V}, \mathbf{w}_t)_h\big]\,\mathrm{d}t + \big(\mathbf{V}^0, \mathbf{w}(0,\cdot)\big)_h - \big(\mathbf{V}^0, \mathbf{w}(0,\cdot)\big) \\ &\quad\quad + \int_0^T \big[\tau^{-1}(\mathbf{R}^+, \mathbf{w}_t) + (\mathbf{V}^+ - \mathbf{V}, \mathbf{w}_t) + \big(\nabla[\mathbf{U}^- - \mathbf{U}], \nabla \mathbf{w}\big)\big]\,\mathrm{d}t\,.\end{aligned}$$

By (4.1) and the inequality after (4.5) we have

$$\int_0^T \| \mathbf{R}^+ \|_{L^2}^2 \,\mathrm{d}t \le C\tau^4 h^{-m}\,.$$

Finally, the assertion (4.10) follows from (4.11), (4.9), and (3.2). □

We are ready to state and prove the first main convergence theorem of this paper.

THEOREM 4.4. *Let* $T > 0$, *and* $\big(\mathbf{u}_0, \mathbf{u}_1\big) \in W^{1,2}(\Omega, \mathbf{S}^{n-1}) \times L^2(\Omega, T\mathbf{S}^{n-1})$. *Suppose that* $\big(\mathbf{U}^0, \mathbf{V}^0\big) \in \boldsymbol{\mathcal{M}}_h \times \boldsymbol{\mathcal{F}}_h(\mathbf{U}^0)$ *satisfies* $\mathbf{U}^0 \to \mathbf{u}_0$ *strongly in* $W^{1,2}(\Omega, \mathbf{R}^n)$, *and* $\mathbf{V}^0 \to \mathbf{u}_1$ *strongly in* $L^2(\Omega, \mathbf{R}^n)$ *as* $h \to 0$. *Let* $\tau = o(h^{\frac{4+m}{3}})$. *Then there exist a map*

$\mathbf{u} \in L^\infty\big(0,T;W^{1,2}(\Omega,\mathbf{R}^n)\big) \cap W^{1,\infty}\big(0,T;L^2(\Omega,\mathbf{R}^n)\big)$ *and a subsequence of solutions* $\{\mathbf{U}_{\tau,h}\}$ *(denoted by the same notation) of Algorithm* 3.1 *such that as* $h \to 0$

$$\begin{aligned} \mathbf{U}_{\tau,h} &\overset{*}{\rightharpoonup} \mathbf{u} \quad && \text{in } L^\infty\big(0,T;W^{1,2}(\Omega,\mathbf{R}^n)\big)\,, \\ (\mathbf{U}_{\tau,h})_t &\rightharpoonup \mathbf{u}_t \quad && \text{in } L^2(\Omega_T,\mathbf{R}^n)\,. \end{aligned}$$

Moreover, $\mathbf{u}$ *is a weak solution of* (1.3)–(1.5).

Proof. It follows from Lemma 4.1 that under the above mesh relation, there exist a map $\mathbf{u} \in W^{1,\infty}\big(0,T;L^2(\Omega,\mathbf{R}^n)\big) \cap L^\infty\big(0,T;W^{1,2}(\Omega,\mathbf{R}^n)\big)$ and a subsequence of $\{\mathbf{U} = \mathbf{U}_{\tau,h}\}$ (denoted by the same notation) such that as $h \to 0$

$$\begin{aligned} \mathbf{U}, \mathbf{U}^-, \mathbf{U}^+ &\overset{*}{\rightharpoonup} \mathbf{u} \quad && \text{in } L^\infty\big(0,T;W^{1,2}(\Omega,\mathbf{R}^n)\big)\,, \quad (4.12) \\ &\to \mathbf{u} && \text{in } L^2(\Omega_T,\mathbf{R}^n), \\ \mathbf{U}_t &\overset{*}{\rightharpoonup} \mathbf{u}_t && \text{in } L^\infty\big(0,T;L^2(\Omega,\mathbf{R}^n)\big)\,. \quad (4.13) \end{aligned}$$

Moreover, taking $h \to 0$ in (4.1) immediately shows that $\mathbf{u}$ satisfies the energy inequality in item (5) of Definition 2.1. Furthermore, from the inequality

$$\|\,|\mathbf{U}^+(t,\cdot)|^2 - 1\,\|_{L^2} \le Ch\,\|\nabla \mathbf{U}^+\|_{L^2} \qquad \forall\, t \in [0,T]$$

we conclude that $|\,\mathbf{u}\,| = 1$ a.e. in Ω_T.

Given $\boldsymbol{\phi} \in C_0^\infty\big([0,T);C^\infty(\Omega,\mathbf{R}^n)\big)$ let $\mathbf{w}_h = I_h\big(\mathbf{U} \wedge \boldsymbol{\phi}\big)$. By the properties of the interpolation operator I_h we get (cf. [9, 12])

$$\begin{aligned} &\|\big(I_h(\mathbf{U}\wedge\boldsymbol{\phi}) - \mathbf{U}\wedge\boldsymbol{\phi}\big)_t\|_{L^2}^2 \\ &\quad\le Ch^4 \sum_{K\in\mathcal{T}_h} \Big(\|\nabla^2(\mathbf{U}_t\wedge\boldsymbol{\phi})\|_{L^2(K)}^2 + \|\nabla^2(\mathbf{U}\wedge\boldsymbol{\phi}_t)\|_{L^2(K)}^2\Big) \\ &\quad\le Ch^4\left[\|\,|\mathbf{U}_t|\,|\nabla^2\boldsymbol{\phi}|\,\|_{L^2}^2 + \|\,|\nabla\mathbf{U}_t|\,|\nabla\boldsymbol{\phi}|\,\|_{L^2}^2 + \|\nabla^2\boldsymbol{\phi}_t\|_{L^2}^2 + \|\,|\nabla\mathbf{U}|\,|\nabla\boldsymbol{\phi}_t|\,\|_{L^2}^2\right] \\ &\quad\le Ch^4\left[\|\boldsymbol{\phi}\|_{W^{2,\infty}}^2 + h^{-2}\|\mathbf{U}_t\|_{L^2}^2\|\nabla\boldsymbol{\phi}\|_{L^\infty}^2 + \|\boldsymbol{\phi}_t\|_{W^{2,\infty}}^2\right]. \end{aligned}$$

This implies that $(\mathbf{w}_h - \mathbf{U}\wedge\boldsymbol{\phi})_t \to 0$ in $L^\infty\big(0,T;L^2(\Omega,\mathbf{R}^n)\big)$. Hence, as $h \to 0$, using $\mathbf{U}_t \cdot \big(\mathbf{U}_t \wedge \boldsymbol{\phi}\big) = 0$ and (4.12) and (4.13),

$$\begin{aligned} &\int_{\Omega_T} \langle \mathbf{U}_t, (\mathbf{w}_h)_t\rangle\,\mathrm{d}x\mathrm{d}t \\ &\quad= \int_{\Omega_T} \langle \mathbf{U}_t, \big(I_h(\mathbf{U}\wedge\boldsymbol{\phi}) - \mathbf{U}\wedge\boldsymbol{\phi}\big)_t\rangle\,\mathrm{d}x\mathrm{d}t + \int_{\Omega_T} \langle \mathbf{U}_t, \mathbf{U}\wedge\boldsymbol{\phi}_t\rangle\,\mathrm{d}x\mathrm{d}t \qquad (4.14) \\ &\quad\to \int_{\Omega_T} \langle \mathbf{u}_t, \mathbf{u}\wedge\boldsymbol{\phi}_t\rangle\,\mathrm{d}x\mathrm{d}t\,. \end{aligned}$$

Similarly, we can show

$$\|\nabla\big[I_h(\mathbf{U}\wedge\boldsymbol{\phi}) - \mathbf{U}\wedge\boldsymbol{\phi}\big]\|_{L^2}^2 \le Ch^2\left[\|\boldsymbol{\phi}\|_{W^{2,2}}^2 + \|\nabla\mathbf{U}\|_{L^2}^2\|\nabla\boldsymbol{\phi}\|_{L^\infty}^2\right].$$

On noting the vector identity $\langle\nabla\mathbf{z}, \nabla\mathbf{z}\wedge\boldsymbol{\phi}\rangle = \langle\nabla\mathbf{z}, \mathbf{z}\wedge\nabla\boldsymbol{\phi}\rangle$, from (4.12) and (4.13) we conclude that as $h \to 0$

$$\begin{aligned} &\int_{\Omega_T} \langle\nabla\mathbf{U}, \nabla\mathbf{w}_h\rangle\,\mathrm{d}\mathbf{x}\mathrm{d}t \\ &\quad= \int_{\Omega_T} \langle\nabla\mathbf{U}, \nabla\big[I_h(\mathbf{U}\wedge\boldsymbol{\phi}) - \mathbf{U}\wedge\boldsymbol{\phi}\big]\rangle\,\mathrm{d}\mathbf{x}\mathrm{d}t + \int_{\Omega_T} \langle\nabla\mathbf{U}, (\mathbf{U}\wedge\nabla\boldsymbol{\phi})\rangle\,\mathrm{d}\mathbf{x}\mathrm{d}t \qquad (4.15) \\ &\quad\to \int_{\Omega_T} \langle\nabla\mathbf{u}, \mathbf{u}\wedge\nabla\boldsymbol{\phi}\rangle\,\mathrm{d}\mathbf{x}\mathrm{d}t\,. \end{aligned}$$

Straightforward calculations show that

$$\left|\int_0^T (\mathbf{V}-\mathbf{V}^+,\mathbf{w}_t)\,\mathrm{d}t\right| \le C\tau^{\frac12}\left(\tau\int_0^T \|\mathbf{V}_t\|_{L^2}^2\,\mathrm{d}t\right)^{\frac12}\left(\int_0^T \|\mathbf{w}_t\|_{L^2}^2\,\mathrm{d}t\right)^{\frac12}.$$

Similarly, we can bound the last term in (4.10) and benefit from the estimates for $\tau\int_0^T \|\mathbf{V}_t\|_{L^2}^2\,\mathrm{d}t$ and $\tau\int_0^T \|\nabla\mathbf{U}_t\|_{L^2}^2\,\mathrm{d}t$ of Lemma 4.1. Letting $h\to 0$ in (4.10) and using (4.14) and (4.15) then give (2.1).

In order to verify item (4) of Definition 2.1, we adopt an argument from [20, p. 92]. $\mathbf{u}_0=\lim_{t\to 0}\lim_{\tau,h\to 0}\mathbf{U}(t,\cdot)$ in $L^2(\Omega,\mathbf{R}^n)$ follows immediately from (5.7). It remains to verify $\mathbf{U}_t(t,\cdot)\to\mathbf{u}_1$ in $L^2(\Omega,\mathbf{R}^n)$ as $t\to 0$. To this end, multiply $\Box\mathbf{u}\wedge\mathbf{u}=0$ with $\mathbf{w}\in C_0^\infty\big([0,T);C^\infty(\Omega,\mathbf{R}^n)\big)$, integrate by parts on Ω_T, and then subtract the resulting equation from (4.10); together with Lemma 4.3 and (5.7), (5.8), as well as assumptions on $(\mathbf{U}^0,\mathbf{V}^0)$, we find for the limit $\tau,h\to 0$,

$$\big((\mathbf{u}_t(0,\cdot)-\mathbf{u}_1)\wedge\mathbf{u}_0,\mathbf{w}\big)=0 \qquad \forall\,\mathbf{w}\in C_0^\infty\big([0,T);C^\infty(\Omega,\mathbf{R}^n)\big).$$

On noting that $\langle\mathbf{u}_1(\mathbf{x}),\mathbf{u}_0(\mathbf{x})\rangle=\langle\mathbf{u}_t(0,\mathbf{x}),\mathbf{u}_0(\mathbf{x})\rangle=0$ for a.e. $\mathbf{x}\in\Omega$, it follows from the vector identity $\mathbf{v}=\langle\mathbf{u},\mathbf{v}\rangle\mathbf{u}-\mathbf{u}\wedge(\mathbf{u}\wedge\mathbf{v})$ with $\mathbf{v}=\mathbf{u}_t(0,\cdot)-\mathbf{u}_1$ that $\mathbf{u}_t(t,\cdot)\rightharpoonup\mathbf{u}_1$ in $L^2(\Omega,\mathbf{R}^n)$ as $t\to 0$.

We also need to show that $\mathbf{u}_t(t,\cdot)\to\mathbf{u}_1$ in $L^2(\Omega,\mathbf{R}^n)$ as $t\to 0$. By weak lower semicontinuity of the L^2-norm and Fatou's lemma

$$\|D\mathbf{u}(t,\cdot)\|_{L^2}\le\liminf_{\tau,h\to 0}\|D\mathbf{U}(t,\cdot)\|_{L^2},\qquad t\ge 0.$$

Hence, for all $t\ge 0$, by (3.2), Lemma 4.1, and the assumptions on initial data,

$$\begin{aligned}E\big(\mathbf{u}_t(t,\cdot),\mathbf{u}(t,\cdot)\big)&\le\liminf_{\tau,h\to 0}E\big(\mathbf{U}_t(t,\cdot),\mathbf{U}(t,\cdot)\big)\\&=\liminf_{\tau,h\to 0}E_h\big(\mathbf{U}_t(t,\cdot),\mathbf{U}(t,\cdot)\big)\le E\big(\mathbf{u}_1,\mathbf{u}_0\big).\end{aligned}$$

Therefore,

$$\limsup_{t\to 0}\|\mathbf{u}_t(t,\cdot)\|_{L^2}\le\|\mathbf{u}_1\|_{L^2},\quad \limsup_{t\to 0}\|\nabla\mathbf{u}(t,\cdot)\|_{L^2}\le\|\nabla\mathbf{u}_0\|_{L^2},$$

and the weak convergence $\mathbf{u}_t(t,\cdot)\rightharpoonup\mathbf{u}_1(\cdot)$ in $L^2(\Omega,\mathbf{R}^n)$ and $\nabla\mathbf{u}(t,\cdot)\rightharpoonup\nabla\mathbf{u}_0(\cdot)$ in $L^2(\Omega)$ imply strong convergence $\mathbf{u}_t(t,\cdot)\to\mathbf{u}_1(\cdot)$ and $\nabla\mathbf{u}(t,\cdot)\to\nabla\mathbf{u}_0(\cdot)$ in $L^2(\Omega,\mathbf{R}^{mn})$ as $t\to 0$. Consequently, $\mathbf{u}:\Omega_T\to\mathbf{R}^n$ attains prescribed initial data continuously in $W^{1,2}(\Omega,\mathbf{R}^n)\times L^2(\Omega,\mathbf{R}^n)$.

Since all requirements of Definition 2.1 are verified, the map $\mathbf{u}:\Omega_T\to\mathbf{R}^n$ is a weak solution to (1.3)–(1.5). The proof is complete. ∎

As a by-product, the above convergence result provides an alternative (to the one given in [20]) proof of existence of weak solutions to (1.3)–(1.5).

5. Convergence analysis for Algorithms 3.2 and 3.3. The goal of this section is to prove that the solution sequences $\{\mathbf{U}^j\}$ of Algorithms 3.2 and 3.3 converge uniformly in ε, and the limits of both sequences are weak solutions of (1.9)–(1.11).

5.1. Stability of Algorithms 3.2 and 3.3. Let $E_h(\cdot,\cdot)$ be the same as in (4.2), and define

$$F_h(\mathbf{U}^j) = \frac{1}{4}\|\,|\,\mathbf{U}^j\,|^2 - 1\,\|_h^2.$$

The following lemma establishes a discrete conservation law for Algorithm 3.2, which mimics the differential conservation law (1.12).

LEMMA 5.1. *For $\varepsilon > 0$, let $\tau_1 = \tau$ and $\big(\mathbf{U}^0, \mathbf{V}^0\big) \in [\boldsymbol{\mathcal{V}}_h]^2$. Then the solution sequence $\{(\mathbf{U}^j, \mathbf{V}^j)\}$ generated by Algorithm 3.2 satisfies for $j = 0 \to (J-1)$*

$$E_h\big(\mathbf{V}^{j+1}, \mathbf{U}^{j+1}\big) + \frac{1}{\varepsilon}F_h(\mathbf{U}^{j+1}) = E_h\big(\mathbf{V}^0, \mathbf{U}^0\big) + \frac{1}{\varepsilon}F_h(\mathbf{U}^0)\,. \tag{5.1}$$

Moreover, if $\mathbf{U}^0 \in \boldsymbol{\mathcal{M}}_h$, then $F_h(\mathbf{U}^0) = 0$, and

$$\|\, d_t\mathbf{U}^{j+1}\,\|_h^2 \le 2E_h(\mathbf{V}^0, \mathbf{U}^0). \tag{5.2}$$

Proof. Taking test functions $(\mathbf{w}, \boldsymbol{\chi}) = (d_t\mathbf{U}^{j+1}, d_t\mathbf{V}^{j+1})$ in (ii) of Algorithm 3.2 and using binomial formulas, we obtain

$$\begin{aligned}
&(d_t\mathbf{V}^{j+1}, d_t\mathbf{U}^{j+1})_h + \frac{1}{2}d_t\|\,\nabla\mathbf{U}^{j+1}\,\|_{L^2}^2 + \frac{1}{\varepsilon}d_tF_h(\mathbf{U}^{j+1}) = 0\,,\\
&(d_t\mathbf{U}^{j+1}, d_t\mathbf{V}^{j+1})_h = \frac{1}{2}d_t\|\,\mathbf{V}^{j+1}\,\|_h^2\,.
\end{aligned}$$

Subtracting the second equation from the first one and summing up the resulting equation over the index from 1 to j gives (5.1).

Alternatively, taking $\boldsymbol{\chi} = d_t\mathbf{U}^{j+1}$ as a test function in the second equation of (ii) of Algorithm 3.2, it follows from (5.1) and the fact $F_h(\mathbf{U}^0) = 0$ if $\mathbf{U}^0 \in \boldsymbol{\mathcal{M}}_h$ that

$$\|\, d_t\mathbf{U}^{j+1}\,\|_h^2 \le \|\,\mathbf{V}^{j+\frac{1}{2}}\,\|_h^2 \le \frac{1}{2}\|\,\mathbf{V}^{j+1}\,\|_h^2 + \frac{1}{2}\|\,\mathbf{V}^j\,\|_h^2 \le 2E_h(\mathbf{V}^0, \mathbf{U}^0)\,.$$

The proof is complete. $\square$

Likewise, we now prove that Algorithm 3.3 also satisfies a discrete conservation law which mimics the differential conservation law (1.12).

LEMMA 5.2. *Let $\varepsilon > 0, m \ge 1$, and $\big(\mathbf{U}^0, \mathbf{U}^1\big) \in [\boldsymbol{\mathcal{V}}_h]^2$. Then the solution sequence $\{\mathbf{U}^j\}$ generated by Algorithm 3.3 satisfies for $j = 1 \to (J-1)$*

$$E_h\big(d_t\mathbf{U}^{j+1}, \mathbf{U}^{j+\frac{1}{2}}\big) + \frac{1}{\varepsilon}F_h(\mathbf{U}^{j+\frac{1}{2}}) = E_h\big(d_t\mathbf{U}^1, \mathbf{U}^{\frac{1}{2}}\big) + \frac{1}{\varepsilon}F_h(\mathbf{U}^{\frac{1}{2}})\,. \tag{5.3}$$

Proof. Choose $\mathbf{w} = \delta_t\mathbf{U}^j$ in (ii) of Algorithm 3.3, and use the identities (cf. [14])

$$\delta_t\mathbf{U}^j = d_t\mathbf{U}^{j+\frac{1}{2}} = \frac{1}{2}\big(\, d_t\mathbf{U}^{j+1} + d_t\mathbf{U}^j\,\big), \qquad \mathbf{U}^{j,\frac{1}{4}} = \frac{1}{2}\big(\,\mathbf{U}^{j+\frac{1}{2}} + \mathbf{U}^{j-\frac{1}{2}}\,\big)$$

and binomial formulas to simplify the resulting equation. Assertion (5.3) follows easily from summing up the equation over the index from 1 to j. $\square$

We remark that the existence of solutions to Algorithms 3.2 and 3.3 now follows from the discrete conservation laws (5.1) and (5.3), respectively, and an application of the Brouwer fixed point theorem.

5.2. Convergence of Algorithm 3.2. The same notation as that in section 4.2 will be used in this subsection. The following lemma prepares us to prove the convergence of Algorithm 3.2.

LEMMA 5.3. *In addition to the assumptions of Lemma* 5.1*, we assume* $\mathbf{U}^0 \in \boldsymbol{\mathcal{M}}_h$*; then for all* $\mathbf{w} \in C_0^\infty([0,T);\boldsymbol{\mathcal{V}}_h)$ *there holds*

$$
\begin{aligned}
(5.4)\quad & \left| \int_0^T \left[-(\mathbf{U}_t, \mathbf{w}_t) + (\nabla \mathbf{U}, \nabla \mathbf{w}) + \frac{1}{\varepsilon}\big([|\,\mathbf{U}\,|^2 - 1]\mathbf{U}, \mathbf{w}\big) \right] \mathrm{d}t + \big(\mathbf{V}^0, \mathbf{w}(0,\cdot)\big) \right| \\
& \qquad \le C \left[h + \frac{\tau + h}{\varepsilon} \right] \left\{ \int_0^T \big(\|\, \nabla \mathbf{w} \,\|_{L^2}^2 + \|\, \nabla \mathbf{w}_t \,\|_{L^2}^2 \big) \mathrm{d}t + \|\, \nabla \mathbf{w}(0,\cdot) \,\|_{L^2}^2 \right\}^{\frac{1}{2}} \\
& \qquad\qquad + \left| \int_0^T (\mathbf{V} - \overline{\mathbf{V}}, \mathbf{w}_t)_h \,\mathrm{d}t \right| + \left| \int_0^T \big(\nabla[\mathbf{U} - \overline{\mathbf{U}}], \nabla \mathbf{w} \big) \,\mathrm{d}t \right| .
\end{aligned}
$$

Proof. Using the new notation and integration by parts, the equations in (ii) of Algorithm 3.2 can be rewritten as follows:

$$
\begin{aligned}
(5.5)\quad & \int_0^T \left\{ -(\mathbf{V}, \mathbf{w}_t)_h + (\nabla \overline{\mathbf{U}}, \nabla \mathbf{w}) + \frac{1}{2\varepsilon}\big([\,|\,\mathbf{U}^-\,|^2 + |\,\mathbf{U}^+\,|^2 - 2]\overline{\mathbf{U}}, \mathbf{w}\big)_h \right\} \mathrm{d}t \\
& \qquad\qquad = -\big(\mathbf{V}^0, \mathbf{w}(0,\cdot)\big)_h , \\
(5.6)\quad & \int_0^T (\mathbf{U}_t - \overline{\mathbf{V}}, \mathbf{w})_h \,\mathrm{d}t = 0 .
\end{aligned}
$$

Using the second equation, we restate terms in the first equation as follows:

$$
\begin{aligned}
\int_0^T (\mathbf{V}, \mathbf{w}_t)_h \,\mathrm{d}t &= \int_0^T (\mathbf{U}_t, \mathbf{w}_t) \,\mathrm{d}t - \int_0^T \big[(\mathbf{U}_t, \mathbf{w}_t) - (\mathbf{U}_t, \mathbf{w}_t)_h \big] \,\mathrm{d}t \\
&\quad + \int_0^T \big(\mathbf{V} - \overline{\mathbf{V}}, \mathbf{w}_t \big)_h \,\mathrm{d}t , \\
\int_0^T (\nabla \overline{\mathbf{U}}, \nabla \mathbf{w}) \,\mathrm{d}t &= \int_0^T (\nabla \mathbf{U}, \nabla \mathbf{w}) \,\mathrm{d}t - \int_0^T \big(\nabla[\mathbf{U} - \overline{\mathbf{U}}], \nabla \mathbf{w} \big) \,\mathrm{d}t , \\
\frac{1}{2\varepsilon} \int_0^T \big([\,|\,\mathbf{U}^-\,|^2 + |\,\mathbf{U}^+\,|^2 - 2]\overline{\mathbf{U}}, \mathbf{w} \big)_h \,\mathrm{d}t &- \frac{1}{\varepsilon} \int_0^T \big([|\,\mathbf{U}\,|^2 - 1]\mathbf{U}, \mathbf{w} \big) \,\mathrm{d}t \\
= -\frac{1}{\varepsilon} \int_0^T & \left\{ \big([\,|\,\mathbf{U}\,|^2 - 1]\mathbf{U}, \mathbf{w} \big) - \big([\,|\,\mathbf{U}\,|^2 - 1]\mathbf{U}, \mathbf{w} \big)_h \right\} \mathrm{d}t \\
- \frac{1}{2\varepsilon} \int_0^T & \big([2|\,\mathbf{U}\,|^2 - |\,\mathbf{U}^-\,|^2 - |\,\mathbf{U}^+\,|^2]\mathbf{U}, \mathbf{w} \big)_h \,\mathrm{d}t \\
- \frac{1}{2\varepsilon} \int_0^T & \big([\,|\,\mathbf{U}^-\,|^2 + |\,\mathbf{U}^+\,|^2 - 2](\mathbf{U} - \overline{\mathbf{U}}), \mathbf{w} \big)_h \,\mathrm{d}t .
\end{aligned}
$$

It follows from (3.2), the L^p-stability of I_h, and an inverse inequality that

$$\Big|\int_0^T \big[(\mathbf{V},\mathbf{w}_t)_h - (\mathbf{U}_t,\mathbf{w}_t)\big]\,\mathrm{d}t\Big| \le Ch\int_0^T \|\,\mathbf{U}_t\,\|_{L^2}\,\|\,\nabla\mathbf{w}_t\,\|_{L^2}\,\mathrm{d}t + \Big|\int_0^T (\mathbf{V}-\overline{\mathbf{V}},\mathbf{w}_t)_h\Big|,$$

$$\Big|\frac{1}{2\varepsilon}\int_0^T \Big\{\big([\,|\,\mathbf{U}^-\,|^2+|\,\mathbf{U}^+\,|^2-2]\overline{\mathbf{U}},\mathbf{w}\big)_h - 2\big([\,|\,\mathbf{U}\,|^2-1]\mathbf{U},\mathbf{w}\big)\Big\}\mathrm{d}t\Big|$$

$$\le \frac{C}{\varepsilon}\int_0^T \Big\{h\|\,I_h([\,|\,\mathbf{U}\,|^2-1]\mathbf{U})\,\|_{L^2}\|\,\nabla\mathbf{w}\,\|_{L^2} + h\|\,\nabla\big([|\,\mathbf{U}\,|^2-1]\mathbf{U}\big)\,\|_{L^1}\|\,\mathbf{w}\,\|_{L^\infty}$$

$$+\tau\,\|\,\mathbf{U}_t\,\|_{L^2}\|\,\mathbf{U}\,\|^2_{L^4}\|\,\mathbf{w}\,\|_{L^\infty} + \frac{\tau}{2}\,\|\,\mathbf{U}_t\,\|_{L^2}\|\,I_h(|\,\mathbf{U}^-\,|^2+|\,\mathbf{U}^+\,|^2-2)\,\|_{L^2}\|\,\mathbf{w}\,\|_{L^\infty}\Big\}\,\mathrm{d}t\,.$$

The assertion of the lemma now follows from combining the above estimates and appealing to Lemma 5.1. ∎

We are ready to state and prove the second main convergence theorem of this paper.

THEOREM 5.4. *Let $\varepsilon>0, m\ge 1, \tau_1=\tau$, and $\big(\mathbf{u}_0,\mathbf{u}_1\big)\in W^{1,2}(\Omega,\mathbf{S}^{n-1})\times L^2(\Omega,T\mathbf{S}^{n-1})$. Suppose that $\big(\mathbf{U}^0,\mathbf{V}^0\big)\in\boldsymbol{\mathcal{M}}_h\times\boldsymbol{\mathcal{V}}_h$ satisfies $\mathbf{U}^0\to\mathbf{u}_0$ strongly in $W^{1,2}(\Omega,\,\mathbf{R}^n)$ and $\mathbf{V}^0\to\mathbf{u}_1$ strongly in $L^2(\Omega,\mathbf{R}^n)$ as $h\to 0$. Then there exist a map $\mathbf{u}^\varepsilon\in L^\infty\big(0,T;W^{1,2}(\Omega,\,\mathbf{R}^n)\big)\cap W^{1,\infty}\big(0,T;L^2(\Omega,\mathbf{R}^n)\big)$ and a subsequence of the solution sequence $\{\mathbf{U}_{\varepsilon,\tau,h}\}$ (denoted by the same notation) of Algorithm 3.2 such that as $\tau,h\to 0$*

$$\mathbf{U}_{\varepsilon,\tau,h}\overset{*}{\rightharpoonup}\mathbf{u}^\varepsilon \quad \text{in } L^\infty\big(0,T;W^{1,2}(\Omega,\mathbf{R}^n)\big)\,,$$
$$(\mathbf{U}_{\varepsilon,\tau,h})_t\overset{*}{\rightharpoonup}\mathbf{u}^\varepsilon_t \quad \text{in } L^\infty\big(0,T;L^2(\Omega,\mathbf{R}^n)\big)$$

uniformly in ε. Moreover, $\mathbf{u}^\varepsilon$ is a weak solution of (1.9)–(1.11). Hence, $\mathbf{u}:=\lim_{\varepsilon\to0}\lim_{\tau,h\to0}\mathbf{U}_{\varepsilon,\tau,h}$ exists and is a weak solution of (1.3)–(1.5), by [20].

Proof. It follows from Lemma 5.1 that there exist $\mathbf{u}^\varepsilon\in W^{1,\infty}\big(0,T;L^2(\Omega,\mathbf{R}^n)\big)\cap L^\infty\big(0,T;W^{1,2}(\Omega,\mathbf{R}^n)\big)$ and a subsequence of $\{\mathbf{U}=\mathbf{U}_{\varepsilon,\tau,h}\}$ (denoted by the same notation) such that as $\tau,h\to 0$

$$\mathbf{U},\mathbf{U}^-,\mathbf{U}^+,\overline{\mathbf{U}}\overset{*}{\rightharpoonup}\mathbf{u}^\varepsilon \quad \text{in } L^\infty\big(0,T;W^{1,2}(\Omega,\mathbf{R}^n)\big)\,, \tag{5.7}$$
$$\to\mathbf{u}^\varepsilon \quad \text{in } L^2(\Omega_T,\mathbf{R}^n)\,,$$
$$\mathbf{U}_t\overset{*}{\rightharpoonup}\mathbf{u}^\varepsilon_t \quad \text{in } L^\infty\big(0,T;L^2(\Omega,\mathbf{R}^n)\big)\,, \tag{5.8}$$

all uniformly in ε. For a.a. $\mathbf{x}\in\Omega$ and all $j=0,1,\ldots,J-1$ there holds $\int_{t_j}^{t_{j+1}}(\mathbf{V}-\overline{\mathbf{V}})(\mathbf{x},t)\,\mathrm{d}t=0$; hence we can subtract the temporal average

$$\overline{\mathbf{w}}_t(x)=(1/\tau)\int_{t_j}^{t_{j+1}}\mathbf{w}_t(\mathbf{x},t)\,\mathrm{d}t$$

of w. Then by a Poincaré estimate we have

$$\Big|\int_0^T(\mathbf{V}-\overline{\mathbf{V}},\mathbf{w}_t)\,\mathrm{d}t\Big| = \Big|\int_0^T(\mathbf{V}-\overline{\mathbf{V}},\mathbf{w}_t-\overline{\mathbf{w}}_t)\,\mathrm{d}t\Big|$$
$$\le C\tau\left(\int_0^T||\mathbf{V}-\overline{\mathbf{V}}||^2_{L^2}\,\mathrm{d}t\right)^{1/2}\left(\int_0^T||\mathbf{w}_{tt}||^2_{L^2}\,\mathrm{d}t\right)^{1/2},$$

and this tends to zero for $h, \tau \to 0$. Similarly, we can bound the last term in (5.4). Therefore, for any $\boldsymbol{\phi} \in C_0^\infty\big(\Omega_T, \mathbf{R}^n\big)$, let $\mathbf{w} = I_h \boldsymbol{\phi}$; using (5.1), (5.2), and the above inequality we can easily pass to the limit in (5.4) by setting $\tau, h \to 0$, and hence conclude that $\mathbf{u}^\varepsilon$ satisfies (1.9)–(1.11) in the distributional sense.

We now show that initial data are continuously taken for $t \to 0$ by the argument used at the end of the proof of Theorem 4.4. Since $\{D\mathbf{U}\}_{k,h}$ is bounded in $L^2(\Omega_T, \mathbf{R}^n)$, we have $\mathbf{U} \to \mathbf{u}^\varepsilon$ in $C^0\big(0,T; L^2(\Omega, \mathbf{R}^n)\big)$ as $\tau, h \to 0$; hence, $\mathbf{u}^\varepsilon : \Omega_T \to \mathbf{R}^n$ assumes data $\mathbf{u}_0 : \Omega \to \mathbf{R}^n$ continuously in $L^2(\Omega; \mathbf{R}^n)$. Multiply (1.9) with $\mathbf{w} \in C_0^\infty\big([0,T); C^\infty(\Omega, \mathbf{R}^n)\big)$, integrate by parts on Ω_T, and then subtract the resulting equation from (5.4); together with Lemma 5.3 and (5.7), (5.8), as well as the assumptions on the initial data $\mathbf{U}^0, \mathbf{V}^0$, we find that $\mathbf{u}_t^\varepsilon(t,\cdot) \rightharpoonup \mathbf{u}_1(\cdot)$ weakly in $L^2(\Omega, \mathbf{R}^n)$ as $t \to 0$. Moreover, by weak lower semicontinuity of the L^2-norm and Fatou's lemma we have

$$\| D\mathbf{u}^\varepsilon(t,\cdot) \|_{L^2} \le \liminf_{\tau,h\to 0} \| D\mathbf{U}(t,\cdot) \|_{L^2}\,, \quad F\big(\mathbf{u}^\varepsilon(t,\cdot)\big) \le \liminf_{\tau,h\to 0} F\big(\mathbf{U}(t,\cdot)\big).$$

Hence, for all $t \ge 0$, by (3.1), Lemma 5.1, and the assumptions on initial data,

$$\begin{aligned} E\big(\mathbf{u}_t^\varepsilon(t,\cdot), \mathbf{u}^\varepsilon(t,\cdot)\big) + \frac{1}{\varepsilon} F\big(\mathbf{u}^\varepsilon(t,\cdot)\big) &\le \liminf_{\tau,h\to 0} \Big[E\big(\mathbf{U}_t(t,\cdot), \mathbf{U}(t,\cdot)\big) + \frac{1}{\varepsilon} F\big(\mathbf{U}(t,\cdot)\big)\Big] \\ &\le \liminf_{\tau,h\to 0} \Big[E_h\big(\mathbf{U}_t(t,\cdot), \mathbf{U}(t,\cdot)\big) + \frac{1}{\varepsilon} F_h\big(\mathbf{U}(t,\cdot)\big)\Big] \\ &\le E\big(\mathbf{u}_1, \mathbf{u}_0\big)\,. \end{aligned}$$

Therefore,

$$\limsup_{t\to 0} \| \mathbf{u}_t^\varepsilon(t,\cdot) \|_{L^2} \le \| \mathbf{u}_1 \|_{L^2}\,, \quad \limsup_{t\to 0} \| \nabla \mathbf{u}^\varepsilon(t,\cdot) \|_{L^2} \le \| \nabla \mathbf{u}_0 \|_{L^2},$$

and weak convergence $\mathbf{u}_t^\varepsilon(t,\cdot) \rightharpoonup \mathbf{u}_1(\cdot)$ in $L^2(\Omega, \mathbf{R}^n)$ and $\nabla \mathbf{u}^\varepsilon(t,\cdot) \rightharpoonup \nabla \mathbf{u}_0(\cdot)$ in $L^2(\Omega)$ imply strong convergence $\mathbf{u}_t^\varepsilon(t,\cdot) \to \mathbf{u}_1(\cdot)$ and $\nabla \mathbf{u}^\varepsilon(t,\cdot) \to \nabla \mathbf{u}_0(\cdot)$ in $L^2(\Omega, \mathbf{R}^{mn})$ as $t \to 0$. Hence, $\mathbf{u}^\varepsilon : \Omega \to \mathbf{R}^n$ attains prescribed initial data continuously in $W^{1,2}(\Omega, \mathbf{R}^n) \times L^2(\Omega, \mathbf{R}^n)$ and is a weak solution of (1.9)–(1.11) in the sense of [20, Definition 5.1]. ☐

5.3. Convergence of Algorithm 3.3. Let $\{\mathbf{U}^j\}$ be the solution sequence generated by Algorithm 3.3, and define for $j = 1 \to J$

$$\mathbf{V}^j := d_t \mathbf{U}^j = \frac{1}{\tau}\big(\, \mathbf{U}^j - \mathbf{U}^{j-1}\,\big), \qquad \mathbf{Z}^j := \mathbf{U}^{j-\frac{1}{2}} = \frac{1}{2}\big(\, \mathbf{U}^{j-1} + \mathbf{U}^j\,\big).$$

We note that τ needs to be replaced by τ_1 in the definition of $\mathbf{V}^j$ when $j = 1$. Let $\mathbf{U}^+, \mathbf{U}^-, \overline{\mathbf{U}}, \mathbf{U}$ (respectively, $\mathbf{V}^+, \mathbf{V}^-, \overline{\mathbf{V}}, \mathbf{V}$ and $\mathbf{Z}^+, \mathbf{Z}^-, \overline{\mathbf{Z}}, \mathbf{Z}$) be the constant and linear interpolations of $\{\mathbf{U}^j\}$ (respectively, $\{\mathbf{V}^j\}$ and $\{\mathbf{Z}^j\}$) in time as defined in section 3.1.

Using the new notation the equation in (ii) of Algorithm 3.3 can be rewritten as for any $\mathbf{w} \in \boldsymbol{\mathcal{V}}_h$

$$\begin{aligned} &(d_t \mathbf{V}^{j+1}, \mathbf{w})_h + \big(\nabla \mathbf{Z}^{j+\frac{1}{2}}, \nabla \mathbf{w}\big) + \frac{1}{2\varepsilon}\big([\,|\mathbf{Z}^j|^2 + |\mathbf{Z}^{j+1}|^2 - 2]\mathbf{Z}^{j+\frac{1}{2}}, \mathbf{w}\big)_h = 0, \\ &(d_t \mathbf{U}^{j+1} - \mathbf{V}^{j+1}, \mathbf{w})_h = 0, \end{aligned}$$

or, equivalently, for any $\mathbf{w} \in C_0^\infty([0,T);\boldsymbol{\mathcal{V}}_h)$

$$\int_0^T \left\{ -(\mathbf{V},\mathbf{w}_t)_h + \big(\nabla\overline{\mathbf{Z}},\nabla\mathbf{w}\big) + \frac{1}{2\varepsilon}\big([\,|\mathbf{Z}^+|^2 + |\mathbf{Z}^-|^2 - 2]\overline{\mathbf{Z}},\mathbf{w}\big)_h \right\} \mathrm{d}t = \big(\mathbf{V}^0,\mathbf{w}(0,\cdot)\big)_h, \tag{5.9}$$

$$\int_0^T (\mathbf{U}_t - \mathbf{V}^+,\mathbf{w})_h \,\mathrm{d}t = 0. \tag{5.10}$$

Comparing (5.9)–(5.10) with (5.5)–(5.6), we see that (5.9) can be obtained from (5.5) by replacing $\mathbf{Z}$ by $\mathbf{U}$, and (5.10) from (5.6) by replacing $\mathbf{V}^+$ by $\overline{\mathbf{V}}$. Moreover, the discrete conservation law (5.3) provides the same energy estimates for $(\mathbf{Z},\mathbf{V})$ here as those enjoyed by $(\mathbf{U},\mathbf{V})$ in subsection 5.2 as a result of the discrete conservation law (5.1). Hence, on noting $\mathbf{U}^j = \mathbf{Z}^j - \tau\mathbf{V}^j$ for Algorithm 3.3 and the inequality (3.3), by repeating the proof of Lemma 5.3 we easily get the following lemma.

LEMMA 5.5. *Let $\{\mathbf{U}^j\}$ be the solution sequence of Algorithm* 3.3 *and $\mathbf{V}^+,\mathbf{V},\mathbf{Z}^+$, $\mathbf{Z}^-$, and $\overline{\mathbf{Z}}$ be defined as above. In addition to the assumptions of Lemma* 5.2*, we assume $(\mathbf{U}^0,\mathbf{U}^1) \in [\boldsymbol{\mathcal{M}}_h]^2$ such that $E_h(d_t\mathbf{U}^1,\mathbf{U}^{\frac12}) + \frac{1}{\varepsilon}F_h(\mathbf{U}^{\frac12}) \le C$; then for all $\mathbf{w} \in C_0^\infty([0,T);\boldsymbol{\mathcal{V}}_h)$, there holds*

$$\begin{aligned} &\left| \int_0^T \left[-(\mathbf{U}_t,\mathbf{w}_t) + (\nabla\mathbf{Z},\nabla\mathbf{w}) + \frac{1}{\varepsilon}\big([|\,\mathbf{Z}\,|^2 - 1]\mathbf{Z},\mathbf{w}\big)\right] \mathrm{d}t + \big(\mathbf{V}^0,\mathbf{w}(0,\cdot)\big) \right| \\ &\qquad \le C\left[h + \frac{\tau+h}{\varepsilon}\right]\left\{ \int_0^T \big(\|\,\nabla\mathbf{w}\,\|_{L^2}^2 + \|\,\nabla\mathbf{w}_t\,\|_{L^2}^2\big)\mathrm{d}t + \|\,\nabla\mathbf{w}(0,\cdot)\,\|_{L^2}^2 \right\}^{\frac12} \\ &\qquad\quad + \left| \int_0^T (\mathbf{V}-\mathbf{V}^+,\mathbf{w}_t)_h \,\mathrm{d}t \right| + \left| \int_0^T \big(\nabla[\mathbf{Z}-\overline{\mathbf{Z}}],\nabla\mathbf{w}\big)\,\mathrm{d}t \right| . \end{aligned} \tag{5.11}$$

We now state and show the third main convergence theorem of this paper.

THEOREM 5.6. *Let $\varepsilon > 0, m \ge 1$, and $\big(\mathbf{u}_0,\mathbf{u}_1\big) \in W^{1,2}(\Omega,\mathbf{S}^{n-1}) \times L^2(\Omega,T\mathbf{S}^{n-1})$. In addition to the assumptions of Lemma* 5.5*, suppose $\big(\mathbf{U}^0,\mathbf{U}^1\big) \in [\boldsymbol{\mathcal{M}}_h]^2$ satisfies $\mathbf{U}^0 \to \mathbf{u}_0$ strongly in $W^{1,2}(\Omega,\mathbf{R}^n)$ and $d_t\mathbf{U}^1 \to \mathbf{u}_1$ strongly in $L^2(\Omega,\mathbf{R}^n)$ as $\tau_1, h \to 0$. Then there exist a map $\mathbf{u}^\varepsilon \in L^\infty\big(0,T;W^{1,2}(\Omega,\ \mathbf{R}^n)\big) \cap W^{1,\infty}\big(0,T;L^2(\Omega,\mathbf{R}^n)\big)$ and a subsequence of the solution sequence $\{\mathbf{U}_{\varepsilon,\tau,h}\}$ (denoted by the same notation) of Algorithm* 3.3 *such that as $\tau, h \to 0$*

$$\begin{aligned} \mathbf{U}_{\varepsilon,\tau,h} &\overset{*}{\rightharpoonup} \mathbf{u}^\varepsilon \quad \text{in } L^\infty\big(0,T;W^{1,2}(\Omega,\mathbf{R}^n)\big)\,, \\ (\mathbf{U}_{\varepsilon,\tau,h})_t &\overset{*}{\rightharpoonup} \mathbf{u}_t^\varepsilon \quad \text{in } L^\infty\big(0,T;L^2(\Omega,\mathbf{R}^n)\big) \end{aligned}$$

uniformly in ε. Moreover, $\mathbf{u}^\varepsilon$ is a weak solution of (1.9)–(1.11)*. Hence, $\mathbf{u} := \lim_{\varepsilon\to0}\lim_{\tau,h\to0}\mathbf{U}_{\varepsilon,\tau,h}$ exists and is a weak solution of* (1.3)–(1.5) *by* [20].

Proof. The proof is essentially the same as that of Theorem 5.4. A slight complication is that after extracting convergent subsequences of $\{\mathbf{U}\}$, $\{\mathbf{Z}\}$, and $\{\mathbf{V}\}$ (which is possible thanks to the discrete energy law (5.3)) and passing to the limit in (5.11), we now need to relate the three limiting functions $\mathbf{u}^\varepsilon, \mathbf{z}^\varepsilon$, and $\mathbf{v}^\varepsilon$ to each other.

To this end, first, on noting identity $\mathbf{U}^j = \mathbf{Z}^j - \tau\mathbf{V}^j$ (or equivalently, $\mathbf{U}^+ = \mathbf{Z}^+ - \tau\mathbf{V}^+$) and inequality (3.3), we conclude that $\mathbf{U}^+$, $\mathbf{Z}^+$, and $\mathbf{Z}$ must converge to the same limit. Hence, $\mathbf{u}^\varepsilon = \mathbf{z}^\varepsilon$. Next, it follows from identity $\mathbf{V}^j = d_t\mathbf{U}^j$ (or equivalently, $\mathbf{V}^+ = \mathbf{U}_t$) and inequality (3.3) that $\mathbf{v}^\varepsilon = \mathbf{u}_t^\varepsilon$. The proof is complete. ∎

Remark 5.1. (a) The use of Algorithm 3.3 requires choosing suitable starting values $\mathbf{U}^0$ and $\mathbf{U}^1$. It is easy to see that $\mathbf{U}^0$ can be constructed straightforwardly as the interpolation or the L^2-projection of $\mathbf{u}_0$. On the other hand, the choice of $\mathbf{U}^1$ is less obvious since U^1 must satisfy (i) $d_t\mathbf{U}^1 \to \mathbf{u}_1$ strongly in $L^2(\Omega, \mathbf{R}^n)$ as $\tau_1, h \to 0$, and (ii) $\frac{1}{\varepsilon}F_h(\mathbf{U}^{\frac{1}{2}}) \le C$. To this end, let $\mathbf{V}^0$ be the interpolation or the L^2-projection of $\mathbf{u}_1$, and define $\mathbf{U}^1 := \mathbf{U}^0 + \tau_1\mathbf{V}^0$. Clearly, requirement (i) is fulfilled. Since $\mathbf{U}^{\frac{1}{2}} = \mathbf{U}^0 + \frac{\tau_1}{2}\mathbf{V}^0$, then $|\mathbf{U}^{\frac{1}{2}}| = 1 + O(\tau_1)$ and $\frac{1}{\varepsilon}F_h(\mathbf{U}^{\frac{1}{2}}) = O(\frac{\tau_1^2}{\varepsilon})$. Hence, requirement (ii) is fulfilled provided that $\tau_1 = O(\varepsilon^{\frac{1}{2}})$. We also note that if $\mathbf{u}_0 \cdot \mathbf{u}_1 = 0$, $\mathbf{V}^0$ above can be chosen such that $\mathbf{U}^0 \cdot \mathbf{V}^0 = 0$; in this case we have $\frac{1}{\varepsilon}F_h(\mathbf{U}^{\frac{1}{2}}) \le C$ provided that $\tau_1 = O(\varepsilon^{\frac{1}{4}})$.

(b) It is interesting to note that Algorithm 3.3 is stable (and converges) uniformly in ε provided that $\tau_1 = O(\varepsilon^{\frac{1}{2}})$, which means that the *first time step* must be chosen very small. On the other hand, no restriction is imposed on h and the later time step τ for stability. Hence, as expected, the algorithm is absolutely stable.

6. Numerical experiments. Besides local existence results, the only global result is a unique smooth solution for all times, provided that the initial energy is sufficiently small [20]. In order to obtain global existence without the assumption of small energy, it would be sufficient to show that the energy cannot concentrate at the hypothetical singularity. Similar to [8], our results for the same initial data show a concentration effect in the case of large initial data at a finite time $T > 0$. All of our numerical results were obtained with MATLAB using the direct linear solver provided by MATLAB's backslash operator. Constraints were included through Lagrange multipliers, and the assembly of matrices was realized in C.

6.1. Consistent discretization versus penalized approximation. In our first numerical experiment we report the performance of Algorithms 3.1 and 3.2 on an example that leads to large gradients. In particular, we follow [6] and employ, for $\mathbf{x} = (x_1, x_2) \in \mathbf{R}^2$, $r := |\mathbf{x}|$, and $a(r) := (1-2r)^4$, the pair of initial data

$$\mathbf{u}_0(\mathbf{x}) := \begin{cases} (0,0,-1) & \text{for } r \ge 1/2, \\ (2x_1 a, 2x_2 a, a^2 - r^2)/(a^2 + r^2) & \text{for } r \le 1/2, \end{cases} \qquad \mathbf{u}_1(\mathbf{x}) := \mathbf{0}.$$

The vector field $\mathbf{u}_0$ wraps about the unit sphere once and is therefore expected to lead to large gradients. We use a uniform triangulation of $\Omega = (-1/2, 1/2)^2$ into 2048 triangles which are squares halved along the direction $(1,1)$ and with maximal diameter $h = \sqrt{2}\,2^{-5}$. The time-step size that we use is $\tau = h^{2.1}$, where the additional power 0.1 is motivated by the condition $\tau = o(h^2)$. The upper plot in Figure 1 displays the $W^{1,\infty}$ seminorm

$$|\mathbf{U}(t)|_{1,\infty} = ||\nabla \mathbf{U}(t)||_{L^\infty(\Omega)}$$

as a function of $t \in [0,2]$ and with $\mathbf{U}$ obtained from Algorithm 3.1 (indicated as "consistent") and Algorithm 3.2 for various choices of ε. In addition, we display in the lower plot of Figure 1 the discrete energies

$$E_h(\mathbf{U}(t)) = E_h(\mathbf{V}^+(t), \mathbf{U}^+(t))$$

and

$$E_h^\varepsilon(\mathbf{U}(t)) = E_h(\mathbf{V}^+(t), \mathbf{U}^+(t)) + \frac{1}{\varepsilon}F_h(\mathbf{U}^+(t))$$

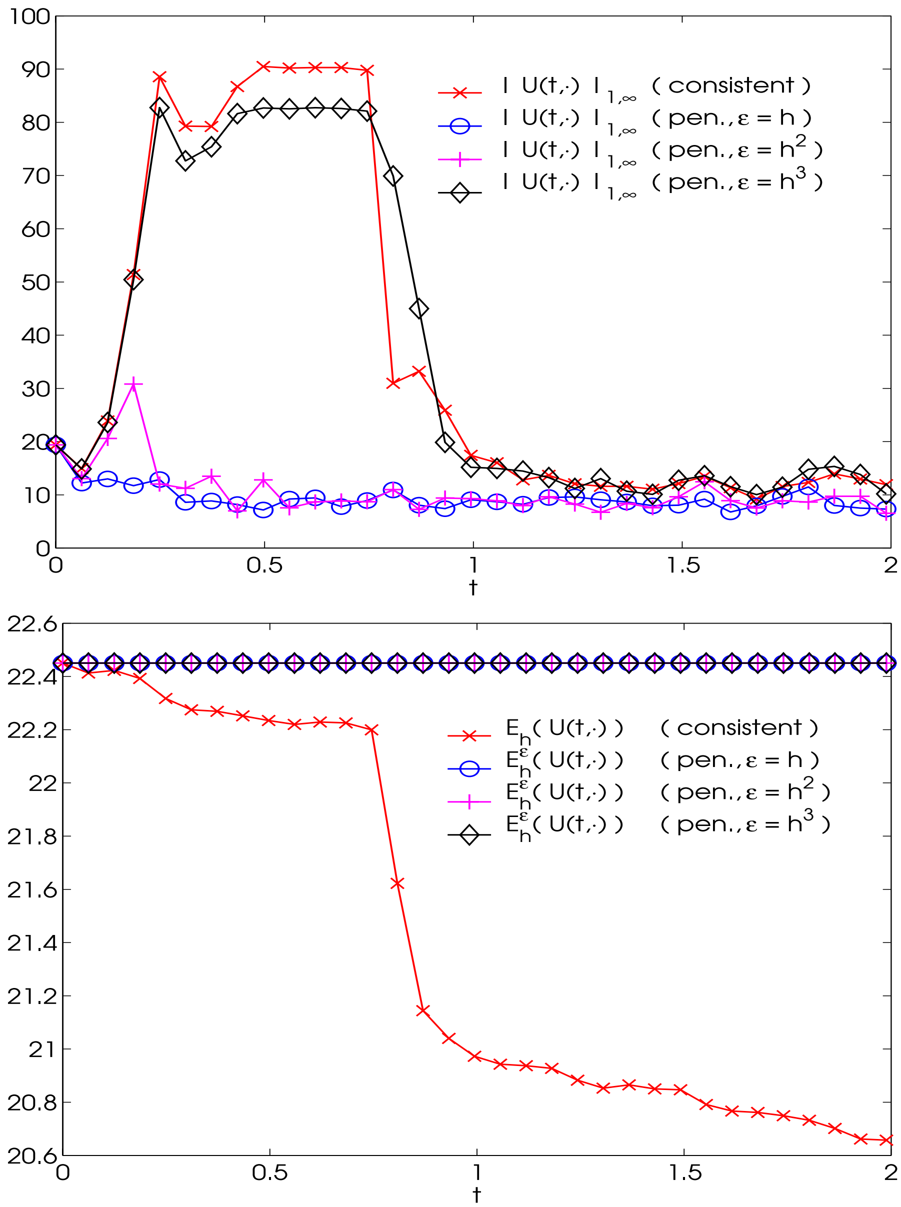

FIG. 1. *$W^{1,\infty}$ seminorm (upper plot) and discrete energy (lower plot) as functions of $t \in [0,2]$ for approximations obtained with Algorithms* 3.1 *and* 3.2 *for various choices of ε. The plots for the results obtained with the penalization method almost coincide in the lower plot.*

for the approximate solutions obtained with Algorithm 3.1 and Algorithm 3.2, respectively. We observe that the $W^{1,\infty}$ seminorm attains its maximum value

$$\max_{\mathbf{W}\in\boldsymbol{\mathcal{V}}_h;\,\forall \mathbf{q}_i\in\mathcal{N}_h,\,|\mathbf{W}(\mathbf{q}_i)|=1} ||\nabla \mathbf{W}||_{L^\infty(\Omega)} = 2^{6.5} \approx 90.51$$

among functions $\mathbf{W} \in \boldsymbol{\mathcal{V}}_h$ that satisfy $|\mathbf{W}(\mathbf{q}_i)| = 1$ for all nodes $\mathbf{q}_i \in \mathcal{N}_h$ for the solution obtained with Algorithm 3.1 at $t \approx 0.25$. Its energy $E_h\big(\mathbf{U}(t)\big)$ slightly decreases at the time of blow-up; however, the fact that it is not increasing indicates the robustness of our direct method. The numerical results differ significantly when the approximation is obtained with Algorithm 3.2 and when different values for ε are used. While the energies E_h^ε obtained with the unconstrained method appear to be constant, we observe that even for the extremely small choice $\varepsilon = h^3$ the $W^{1,\infty}$ seminorm of the numerical approximation remains bounded away from the maximum value. We note that such a choice of ε is critical since Lemma 5.3 suggests that convergence of the numerical approximations obtained by Algorithm 3.2 can be expected only if $\tau, h = o(\varepsilon)$. Nevertheless, it seems that the results obtained with the penalization approach those obtained with Algorithm 3.1 when ε becomes small.

6.2. Numerical evidence for finite-time blow-up and occurrence of singular solutions. Even though Algorithms 3.1 and 3.2 are mathematically equivalent in the sense that they both provide subsequences that converge weakly to a solution of the wave map equation, we believe that Algorithm 3.1 is better suited to studying the occurrence of singularities since the corresponding topological effects are not diffused as in the regularized (penalized) approximation scheme of Algorithm 3.2. The two plots in Figure 2 display the behavior of numerical solutions obtained with Algorithm 3.1 for decreasing mesh-sizes and with the initial data described above. As in the previous subsection we employed uniform triangulations of $\Omega = (-1/2, 1/2)^2$ into squares halved along the direction $(1, 1)$. The time-step size was defined through $\tau = h^{2.1}$ for $h = \sqrt{2}\,2^{-5}$, $\sqrt{2}\,2^{-6}$, $\sqrt{2}\,2^{-7}$ corresponding to a sequence of triangulations with 2048, 8192, and 32768 triangles, respectively. Figure 2 shows that for each triangulation the maximal $W^{1,\infty}$ seminorm is attained. This indicates that an existing exact solution for the choice of our initial data has unbounded $W^{1,\infty}$ norm, i.e., is a singular solution, but we stress that this behavior might change for sufficiently small mesh-sizes. This so-called finite-time blow-up behavior occurs in our example when the solution at the origin points in a direction different from all the surrounding vectors. We illustrate this behavior in Figures 3 and 4, where the first two components of the vector field $\mathbf{U}\big(t, \mathbf{x}\big)$ for the nodes $\mathbf{q}_i \in (-1/4, 1/4)^2 \cap \mathcal{N}_h$ and the vectors in a neighborhood of the origin are shown, respectively, for the solution obtained with the triangulation that consists of 2048 triangles. After a small period of large (maximal) gradients, the vector at the origin changes its direction. This change of direction may be unstable and mesh-dependent. It is not clear to the authors whether such an effect also occurs in the continuous situation. We note that in this experiment it was not possible to recover the initial data by reversing time. The results in the following subsection indicate that the quantitative behavior does not change significantly when other triangulations with more symmetry are used.

6.3. Effect of symmetric/asymmetric triangulations. In order to obtain a better understanding of the mechanism that forces the vector at the origin to change its direction after the occurrence of a maximal $W^{1,\infty}$ norm, we applied Algorithm 3.1 to the initial data given above for different spatial triangulations of comparable mesh-size. The four different triangulations we used are displayed in Figure 5. The first one (upper left plot) is obtained from four uniform refinements of a coarse triangulation $\mathcal{T}_0$ of $\Omega = (-1/2, 1/2)$ into two triangles. The next triangulation (upper right plot) resulted from four uniform refinements of that coarse triangulation of Ω that is obtained from dividing Ω into four similar triangles by dividing it along its diagonals. The third triangulation shown (lower left plot) is obtained from the previous one by a random perturbation of those nodes that do not belong to the boundary $\partial\Omega$ by a value smaller than $h/10$ in each direction. Finally, we also employed an approximation of

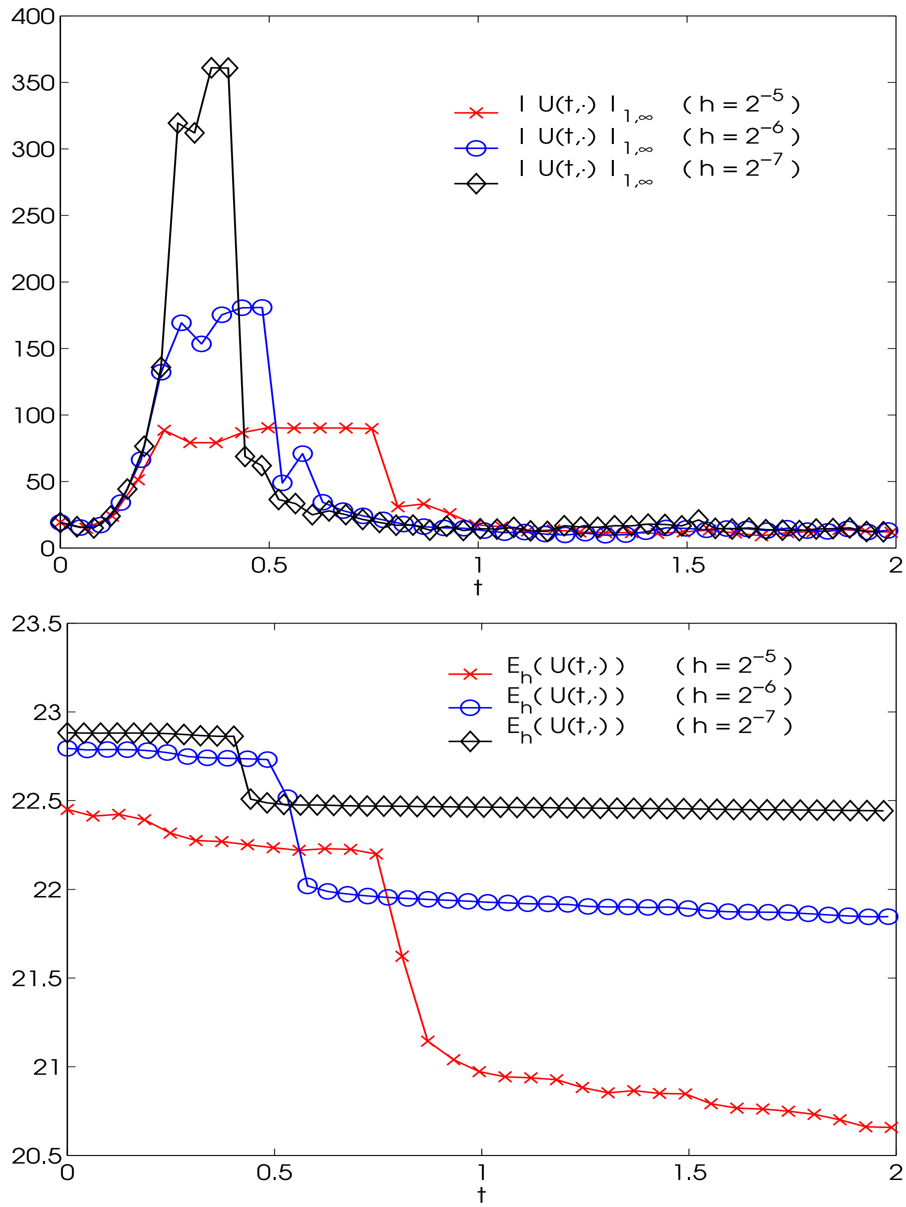

FIG. 2. $W^{1,\infty}$ *seminorm (upper plot) and discrete energy (lower plot) as functions of* $t \in [0, 2]$ *for approximations obtained with Algorithm* 3.1 *on refining triangulations.*

the disk $\Omega = \{\mathbf{x} \in \mathbf{R}^2 : |\mathbf{x}| \le 1/2\}$; cf. the last triangulation in Figure 5 (lower right plot). Figure 6 displays the $W^{1,\infty}$ seminorms of the numerical approximations using the four different triangulations. The results differ mostly after the occurrence of discrete blow-up, and we expect that this mesh-dependence reflects a loss of stability and a possibility of nonuniqueness of exact solutions. However, we observe that before the occurrence of maximal gradients, quantitatively the behavior of the numerical approximations does not change significantly when more symmetry is introduced in the grid or when symmetry is entirely destroyed by perturbing the positions of the vertices. We thus believe that our Algorithm 3.1 provides an accurate approximation of the exact solution and is capable of detecting singularities in a reliable way.

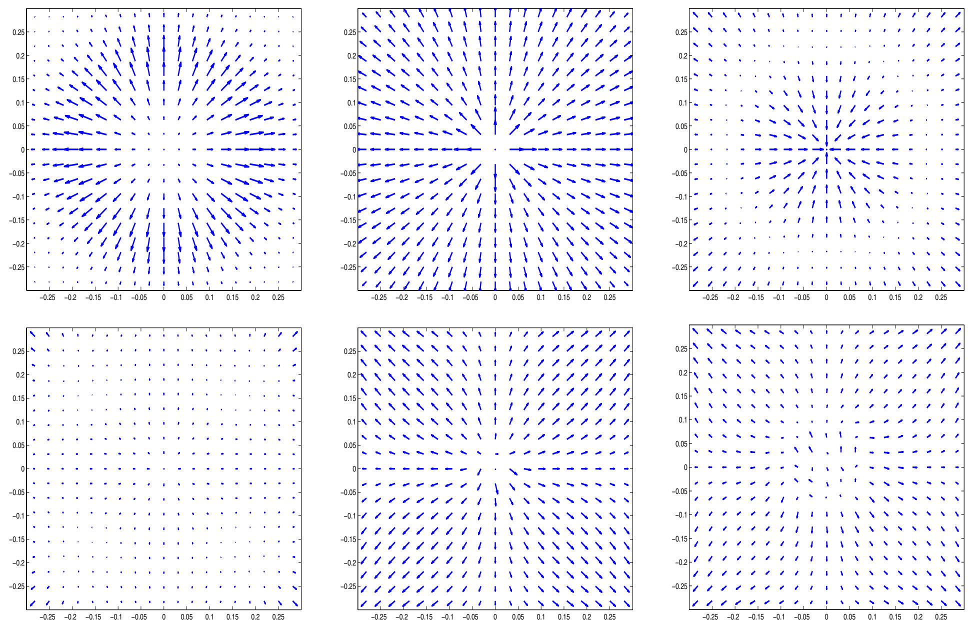

FIG. 3. *Scaled first two components of* $\mathbf{U}(t,\cdot)$ *for the solution provided by Algorithm* 3.1 *for* $t = 0$, 0.20025, 0.40051, 0.60076, 0.80102, 1.00130 *(from left to right, from top to bottom).*

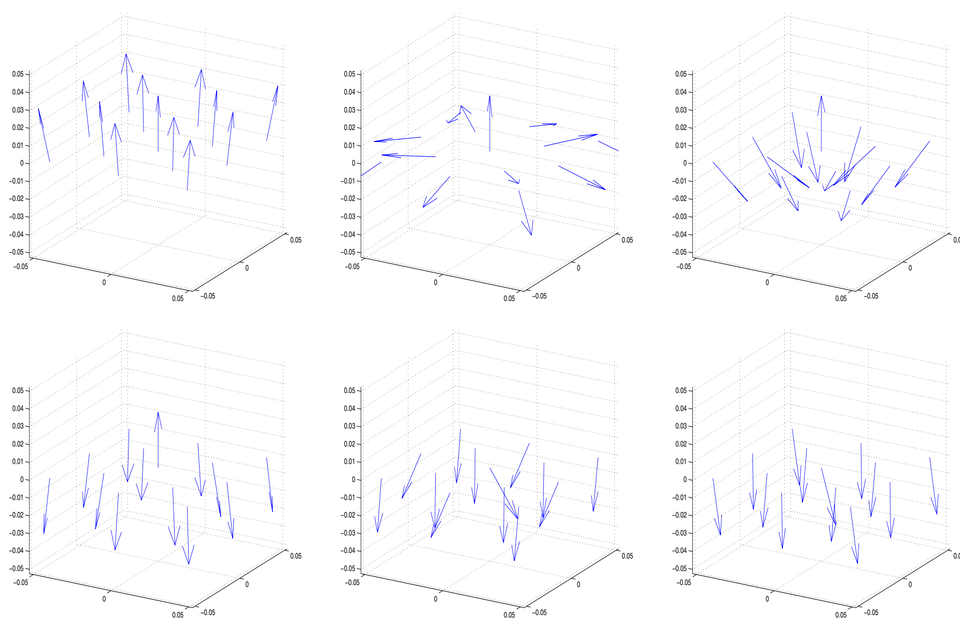

FIG. 4. *Scaled solution* $\mathbf{U}(t,\cdot)$ *obtained with Algorithm* 3.1 *in a neighborhood of the origin for* $t = 0$, 0.20025, 0.40051, 0.60076, 0.80102, 1.00130 *(from left to right, from top to bottom) on a criss-cross triangulation.*

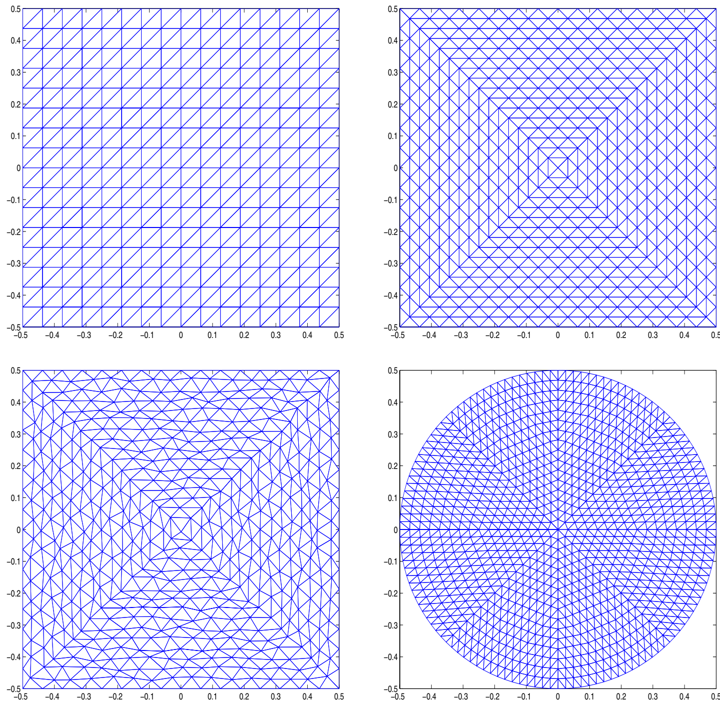

FIG. 5. *Triangulations with different symmetry properties.*

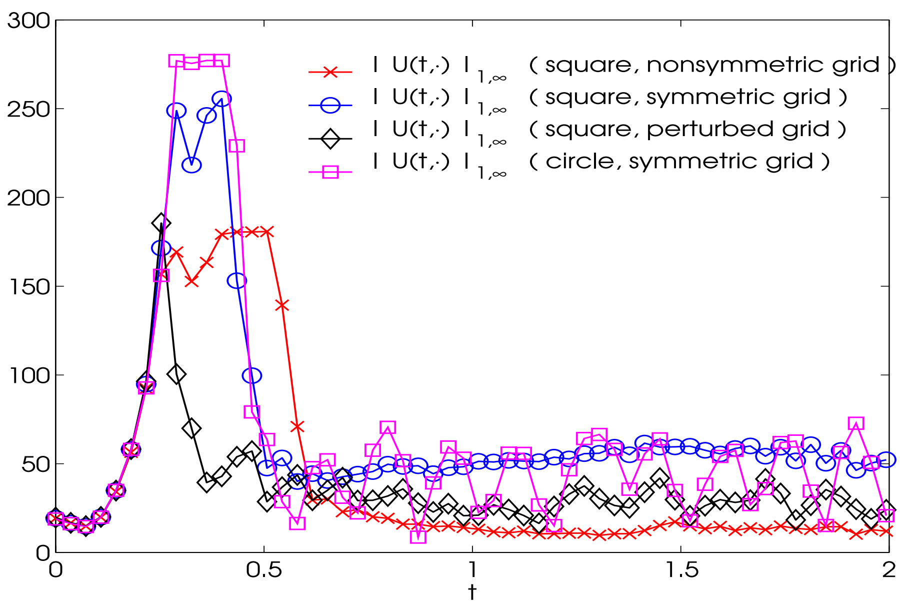

FIG. 6. $W^{1,\infty}$ *seminorm and discrete energy as functions of* $t \in [0, 2]$ *for approximations obtained with Algorithm* 3.1 *for (refinements of) the triangulations shown in Figure* 5.

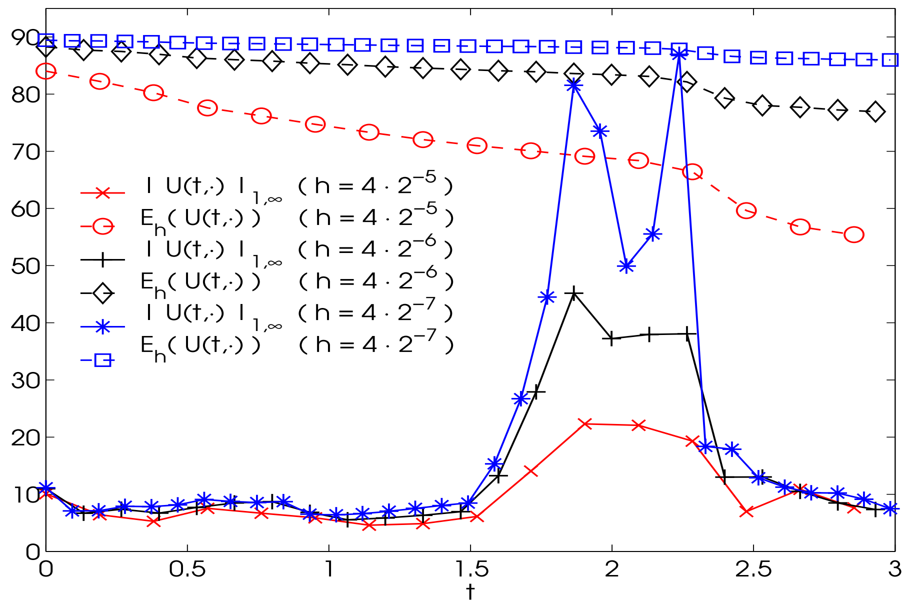

FIG. 7. *$W^{1,\infty}$ seminorm and discrete energy as functions of $t \in [0,3]$ for approximations obtained with Algorithms* 3.1 *for equivariant initial data on refining triangulations.*

6.4. Equivariant initial data. In our last series of experiments we follow the work of [8] and employ equivariant initial data defined by

$$\mathbf{u}_0(r,\theta) = \begin{pmatrix} \sin\chi(r,\theta)\,\sin\theta \\ \sin\chi(r,\theta)\,\cos\theta \\ \cos\chi(r,\theta) \end{pmatrix}, \qquad \mathbf{u}_1(r,\theta) \equiv 0,$$

where (r,θ) denotes polar coordinates in $\mathbf{R}^2$ and $\chi(r,\theta) = (r^3/4)\exp(-(4(r-2)/10)^4)$. We ran Algorithm 3.1 with these initial data on uniform triangulations of $\Omega = (-2,2)^2$ with 2048, 8192, and 32768 triangles. Figure 7 displays the $W^{1,\infty}$ seminorms and the energy as functions of $t \in [0,3]$ obtained with the three different triangulations. As in [8] we observe finite-time blow-up. We remark that in [8] the results were obtained by reducing the problem to a one-dimensional model that guarantees that the solution remains an equivariant map for all times. Here, this is not included as a constraint in the approximation scheme. Nevertheless, the sequence of snapshots of the numerical solution shown in Figure 8 for $t \approx 0, 2, 4, \ldots, 10$ indicates that this is true for our numerical solution. In particular, the vector at the origin does not change its direction on triangulations that obey certain symmetry properties. In Figure 9 we display the third component of the solution at the origin as a function of $t \in [0,10]$, i.e., the real number $\mathbf{U}^{(3)}(t,\mathbf{0})$, for the numerical solution obtained on the different triangulations described above (scaled by a factor of 4 for this experiment) and indicated in Figure 9 (for simplicity, we did not perturb the vertex located at the origin). For the unperturbed triangulations the solution vector at the origin does not change for $t \in (0,10)$, while for the perturbed triangulation the vector changes its direction for a small period of time when the blow-up occurs. These results are in good agreement with theoretical predictions stating that for equivariant data the vector at the origin remains constant [24]. Finally, we remark that in this experiment (on the nonsymmetric, unperturbed triangulation) it was possible to reverse time and to reproduce the initial data.

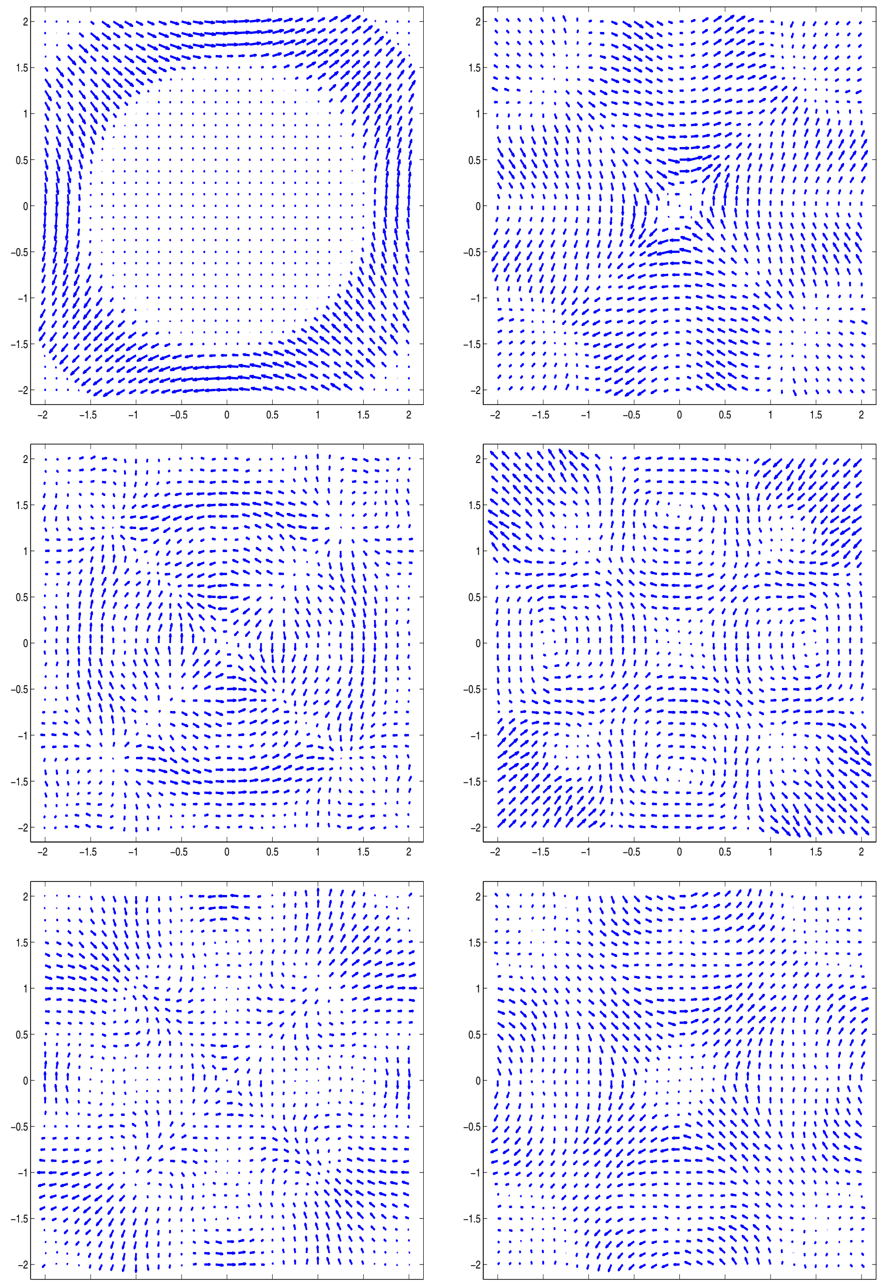

Fig. 8. *Scaled first two components of* $\mathbf{U}(t,\cdot)$ *for the solution provided by Algorithm* 3.1 *for equivariant initial data for* $t \approx 0$, 2.0, 4.0, 6.0, 8.0, 10.0 *(from left to right, from top to bottom).*

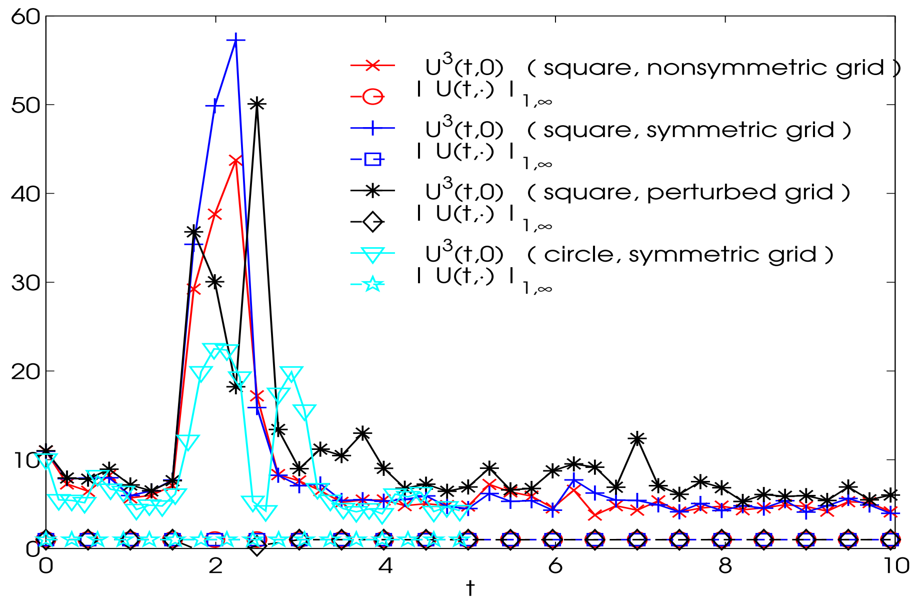

FIG. 9. $W^{1,\infty}$ *seminorm and third component of numerical solutions at origin of* $t \in [0, 10]$ *for approximations obtained with Algorithm* 3.1 *for (refinements of) the triangulations shown in Figure* 5 *and equivariant initial data.*

REFERENCES

[1] I. J. R. AITCHISON AND A. J. G. HEY, *Gauge Theories in Particle Physics, Volume* I: *A Practical Introduction: From Relativistic Quantum Mechanics to QED*, Graduate Student Series in Physics, Taylor & Francis, London, 2002.

[2] F. ALOUGES AND P. JAISSON, *Convergence of a finite element discretization for the Landau Lifshitz equations*, Math. Models Methods Appl. Sci., 16 (2006), pp. 299–316.

[3] G. A. BAKER, *Error estimates for finite element methods for second order hyperbolic equations,* SIAM J. Numer. Anal., 13 (1976), pp. 564–576.

[4] J. W. BARRETT, S. BARTELS, X. FENG, AND A. PROHL, *A convergent and constraint-preserving finite element method for the p-harmonic flow into spheres*, SIAM J. Numer. Anal., 45 (2007), pp. 905–927.

[5] J. W. BARRETT, X. FENG, AND A. PROHL, *p-harmonic map heat flows for* $1 \leq p < \infty$ *and their finite element approximations with applications to image processing*, SIAM J. Math. Anal., submitted.

[6] S. BARTELS, J. KO, AND A. PROHL, *Numerical approximation of the Landau–Lifshitz–Gilbert equation and finite time blow-up of weak solutions*, SIAM J. Numer. Anal., submitted.

[7] B. K. BERGER, P. T. CHRUSCIEL, AND V. MONCRIEF, *On "asymptotically flat" space-times with* G_2*-invariant Cauchy surfaces*, Ann. Physics, 237 (1995), pp. 322–354.

[8] P. BIZON, T. CHMAJ, AND Z. TABOR, *Formation of singularities for equivariant* $2 + 1$ *dimensional wave maps into two-sphere*, Nonlinearity, 14 (2001), pp. 1041–1053.

[9] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer-Verlag, New York, 1994.

[10] S. M. CARROLL, *Spacetime and Geometry: An Introduction to General Relativity*, Addison–Wesley, Reading, MA, 2003.

[11] T.-P. CHENG AND L.-F. LI, *Gauge Theory of Elementary Particle Physics*, Oxford University Press, Oxford, UK, 1983.

[12] P. G. CIARLET, *Basic error estimates for elliptic problems*, in Handbook of Numerical Analysis, Vol. II, North–Holland, Amsterdam, 1991, pp. 17–351.

[13] T. DUPONT, L^2*-estimates for Galerkin methods for second order hyperbolic equations*, SIAM J. Numer. Anal., 10 (1973), pp. 880–889.

[14] X. Feng, *Analysis of finite element methods and domain decomposition algorithms for a fluid-solid interaction problem*, SIAM J. Numer. Anal., 38 (2000), pp. 1312–1336.
[15] A. Freire, *Global weak solutions of the wave map system to compact homogeneous spaces*, Manuscripta Math., 91 (1996), pp. 525–533.
[16] M. Giaquinta and S. Hildebrandt, *Calculus of Variations* I, Springer-Verlag, New York, 1996.
[17] J. Isenberg and S. L. Liebling, *Singularity formation in* $2+1$ *wave maps*, J. Math. Phys., 43 (2002), pp. 678–683.
[18] S. Müller and M. Struwe, *Spatially discrete wave maps on* $(1+2)$*-dimensional space-time*, Topol. Methods Nonlinear Anal., 11 (1998), pp. 295–320.
[19] J. Shatah, *Weak solutions and development of singularities in the* $SU(2)$ σ*-model*, Comm. Pure Appl. Math., 41 (1998), pp. 459–469.
[20] J. Shatah and M. Struwe, *Geometric Wave Equations*, New York University, Courant Institute of Mathematical Sciences, New York; AMS, Providence, RI, 1998.
[21] J. Shatah and M. Struwe, *The Cauchy problem for wave maps*, Int. Math. Res. Not., 11 (2002), pp. 555–571.
[22] M. Struwe, *Geometric evolution problems*, in Nonlinear Partial Differential Equations in Differential Geometry, IAS/Park City Math. Ser. 2, AMS, Providence, RI, 1996, pp. 257–339.
[23] M. Struwe, *Recent existence and regularity results for wave maps*, in Journées "Équations aux Dérivées Partielles" (Saint-Jean-de-Monts, 1997), Exp. No. XVII, 7 pp., École Polytech., Palaiseau, France, 1997.
[24] M. Struwe, *Equivariant wave maps in two space dimensions*, Comm. Pure Appl. Math., 56 (2003), pp. 815–823.
[25] D. Tataru, *The wave maps equation*, Bull. Amer. Math. Soc. (N.S.), 41 (2004), pp. 185–204.