



# A posteriori error estimates for wave maps into spheres

Jan Giesselmann<sup>1</sup> · Elena Mäder-Baumdicker<sup>2</sup> · David Jakob Stonner<sup>2</sup>

Received: 11 April 2022 / Accepted: 23 May 2023 / Published online: 17 July 2023  
© The Author(s) 2023

## Abstract

We provide a posteriori error estimates in the energy norm for temporal semi-discretisations of wave maps into spheres that are based on the angular momentum formulation. Our analysis is based on novel weak–strong stability estimates which we combine with suitable reconstructions of the numerical solution. We present time-adaptive numerical simulations based on the a posteriori error estimators for solutions involving blow-up.

**Keywords** Wave maps · Weak–strong stability · A posteriori error estimates · Blow-up

**Mathematics Subject Classification (2010)** 65M16 · 35L71 · 35B44

## 1 Introduction

This paper is concerned with the numerical approximation of wave maps, i.e., semi-linear wave equations with the point-wise constraint that the solution takes values in some given target manifold. They arise as critical points of a Lagrange functional for manifold valued functions and serve as model problems in general relativity [1] and in particle physics [2]. We refer to [3–5] and the introduction in [6] for an overview on

---

Communicated by: Ilaria Perugia

---

David Jakob Stonner and Elena Mäder-Baumdicker contributed equally to this work.

---

✉ Jan Giesselmann  
jan.giesselmann@tu-darmstadt.de  
Elena Mäder-Baumdicker  
maeder-baumdicker@mathematik.tu-darmstadt.de

<sup>1</sup> Mathematics, Technical University of Darmstadt, Dolivostr. 15, 64293 Darmstadt, Hesse, Germany

<sup>2</sup> Mathematics, Technical University of Darmstadt, Schlossgartenstr. 7, 64289 Darmstadt, Hesse, Germany

the general theory of wave maps. The monograph [7] contains a detailed introduction to the recent development in the analysis of wave maps.

There are two key challenges in numerically approximating wave maps: One is the point-wise constraint that make the function spaces in which solutions are sought non-linear, and the second is gradient blow-up that leads to highly localized phenomena in space and time that need to be suitably resolved by numerical methods. A variety of different numerical methods that deal with the point-wise constraint using different approaches such as projections, penalties, and Lagrange multipliers have been proposed [8–14], and for the methods of Bartels and coworkers, a priori convergence analysis is available, in the sense that stability estimates are proven that imply convergence of subsequences to weak solutions.

It seems desirable to obtain more quantitative information on the accuracy of numerical approximations, and given the highly localized dynamics of gradient blow-up, we aim to provide a posteriori error estimates. We will focus on a scheme whose a priori analysis was studied in [15] and [16]. For the schemes at hand, convergence results are available even beyond gradient blow-up, i.e., limits of subsequences of numerical solutions are weak solutions, but quantitative estimates beyond singularity formation seem to be out of reach due to discontinuity of the solution operator [17], see Section 2 for details. Thus, we focus on estimates for the solution up to the blow-up time.

Similarly to what was done in [18, 19], we study errors entering via temporal discretization, and indeed, we restrict our study to semi-discretization in time. While the development of estimators for spatial discretization errors is certainly an important task in its own right, it is beyond the scope of this work. Indeed, since our stability analysis requires a strong solution that is in  $H^2(\Omega, S^2)$ , deriving a posteriori error estimators for numerical schemes using finite differences or  $C^0$  finite elements for spatial discretization will require suitable sufficiently regular spatial reconstructions of the numerical solution. For wave equations in flat domains, this has been achieved by elliptic reconstruction [20], i.e., the reconstruction is defined as the solution of an elliptic equation, and a posteriori error estimates for numerical solutions of elliptic problems are crucial in order to proceed. However, the corresponding elliptic problem for the case at hand would be a harmonic map problem, and for these, there do not seem to be any a posteriori error estimators available. Indeed, little has been proven concerning convergence of numerical schemes for harmonic maps. In certain situations, uniqueness and regularity of harmonic maps can be ensured, which allows to show high order convergence of so-called geodesic finite elements [21]. In the general case, only existence of weak harmonic maps is guaranteed, and for general triangulations and minimal regularity solutions, an a posteriori criterion is needed in order to guarantee weak convergence of numerical solutions [22]. To the best of our knowledge, no quantitative a posteriori error bounds are available for numerical approximation schemes for harmonic maps and harmonic map heat flows. Another option for spatial reconstruction might be to extend smoothing by convolution, as investigated in [23], to  $S^2$ -valued functions.

For a long time, a posteriori error control for (linear) wave equations was limited to first order schemes [20, 24]. Earlier works on adaptivity for wave equations can be found in [25–27]. Quite recently, a posteriori error estimates for second order multi-step time discretisations of the linear wave equation were derived [18].

Due to the appearance of singularities, our goals are similar to those pursued in [19, 28] that study blow-up for semi-linear parabolic equations. We also refer to these papers for earlier works on numerical approximation of blow-up solutions of nonlinear PDEs such as nonlinear Schrödinger equations or semi-linear parabolic problems. Let us mention that the blow-up mechanism in wave maps is rather different from the blow-up mechanisms in nonlinear Schrödinger equations or semi-linear parabolic problems where the  $L^\infty$ -norm blows up in finite time.

In order to derive the desired a posteriori error estimates, we use two main ingredients: firstly, a suitable reconstruction of the numerical solution that can be understood as the exact solution of a perturbed version of the *angular momentum formulation* for wave maps into spheres and, secondly, a novel weak–strong stability principle. One key feature of our reconstruction is to ensure that the estimator is formally of optimal order, i.e., that it converges to zero with the same rate as the true error on equidistant meshes.

The remainder of this work is organized as follows: We review some facts from the analysis of wave maps in Sect. 2 and introduce the problem and basic notation in Sect. 3. Section 4 provides two stability estimates: one is based on a first order reformulation of the problem that is available when the target manifold is  $S^2$ , and the other theorem covers the general case. In Sect. 5, we provide a posteriori error estimates for a numerical scheme that is based on the first order reformulation of the problem. The main contribution of this section is the construction of suitable reconstructions of the numerical solution, whereas computable bounds for the residuals, that appear when the reconstruction is inserted into the wave map problem, are postponed to the Appendix. Finally, in Sect. 6, we report on numerical experiments using adaptive time stepping based on the a posteriori error estimators derived before.

## 2 Background on wave maps

A specific feature of wave maps is that depending on the size of initial data (in suitable Sobolev norms) and the dimension of the target manifold either strong solutions may exist on arbitrarily long time intervals or solutions may exhibit gradient blow-up in finite time; see [6]. But there are also larger classes of solutions, namely *distributional* or *weak solutions* and in particular finite energy weak solutions. Note that weak solutions, for example, exist as accumulation points of subsequences of numerical schemes in [15]. The existence of global weak solutions for finite energy initial data in  $2 + 1$  dimensions was established in [29]. We will introduce our precise notion of finite energy weak solution in the next section. If a (global) finite energy weak solution and a non-global strong solution exist, then uniqueness results as in [30, 31] guarantee that both solutions agree until the appearance of the singularity. These, under certain circumstances, strong solutions can be extended as weak solutions through the singularity. Conditions such as an energy inequality are needed to get this uniqueness result because, in general, weak solutions are not unique [31].

We provide weak–strong stability results in Theorems 1 and 2 that can be seen as more quantitative versions of the (finite energy) weak–strong-uniqueness results in [30, 31]. Indeed, Theorems 1 and 2 assert that as long as there exists a sufficiently

regular solution for certain initial data, there are explicit bounds for the difference between this solution and solutions to perturbed problems even if those are only finite energy weak solutions. Note that this quantitative control requires the same regularity as the uniqueness result [30]; see Remark 4 for more details.

We employ Theorem 1 (that assumes that the target manifold is  $S^2$ ) in our a posteriori error analysis, but since weak–strong stability results are interesting in their own right, we present a weak–strong stability result for general target manifolds in Theorem 2. While it can be thought of as an extension of Theorem 1 (where the target manifold is  $S^2$ ), there are some significant differences on the technical level that will be discussed in Remark 5 and Remark 6. Those are the reason why we base our a posteriori error analysis on Theorem 1. This is discussed in more detail in Remark 7.

Uniqueness and stability properties of wave maps do not only depend on the target manifold but also on the domain. The main interest in analysis is currently in wave maps with the whole of  $\mathbb{R}^m$  as the spatial domain, and this is the setup the works mentioned below address. Uniqueness for wave maps in  $1 + 1$  dimensions was shown in [32], and non-uniqueness in the supercritical dimension  $3 + 1$  was shown in [31, 33]. To the best of our knowledge, in the critical dimension  $2 + 1$ , uniqueness of weak solutions to finite energy data is unknown, and it is unclear whether imposing an energy inequality restores uniqueness in 2 or more space dimensions. An interesting observation in  $2 + 1$  dimensions is that solutions (even if they are unique) do not depend continuously on the initial data in the energy norm [17]. It should be noted that our weak–strong stability results use the energy norm and are valid in arbitrarily many space dimensions. It follows from [17] that any such stability result, i.e., any bound for the difference of two solutions, measured in the energy norm, needs to involve a stronger norm (than the energy norm) of at least one of the solutions. This is a fundamental obstacle to deriving a posteriori error estimates (in the energy norm) that are convergent (i.e., go to 0 for  $\tau, h \searrow 0$ , where  $\tau, h$  denote temporal and spatial mesh width respectively) in case the exact solution does not have any additional regularity (beyond being a finite energy weak solution).

### 3 Problem statement and notation

For some bounded Lipschitz domain  $\Omega \subset \mathbb{R}^m$ , some final time  $\mathbb{T} > 0$  and some  $n$ -dimensional submanifold without boundary  $N \subset \mathbb{R}^\ell$  a wave-map is a map

$$u : (0, \mathbb{T}) \times \Omega \rightarrow N, \quad \text{satisfying} \quad \partial_t^2 u - \Delta u \perp T_u N \quad \text{in } (0, \mathbb{T}) \times \Omega \quad (1)$$

where  $T_u N$  denotes the tangent space of  $N$  at  $u$ . Equation (1) needs to be complemented with initial and boundary data. To this end, maps  $u_0 : \Omega \rightarrow N$  and  $u_1 : \Omega \rightarrow \mathbb{R}^\ell$  such that  $u_1(x) \in T_{u_0(x)} N$  for all  $x \in \Omega$  are fixed, and one requires

$$u(0, \cdot) = u_0, \quad \partial_t u(0, \cdot) = u_1, \quad (2)$$

and homogeneous Neumann boundary conditions

$$\partial_n u = 0 \quad \text{on } (0, \mathbb{T}) \times \partial\Omega. \tag{3}$$

Strong solutions of the wave map equation satisfy an energy conservation principle

$$E[u(t), \partial_t u(t)] := \frac{1}{2} \int_{\Omega} |\partial_t u(t)|^2 + |\nabla u(t)|^2 dx = E[u_0, u_1] \tag{4}$$

and are critical points of the Lagrangian

$$L[u(t), \partial_t u(t)] := \frac{1}{2} \int_0^{\mathbb{T}} \int_{\Omega} |\partial_t u(t)|^2 - |\nabla u(t)|^2 dx dt. \tag{5}$$

We reformulate the wave map (1) in order to see that this is a semi-linear wave equation. Let  $A_p(\cdot, \cdot) : T_p N \times T_p N \rightarrow (T_p N)^\perp$  be the second fundamental form of the compact submanifold  $N$  at a point  $p \in N$ . We denote the variables on  $[0, \mathbb{T}) \times \Omega$  by  $(t, x) = (x^\alpha)$ ,  $0 \leq \alpha \leq m$ . We raise and lower indices with the Minkowski metric  $(\eta_{\alpha\beta}) = \text{diag}(-1, 1, \dots, 1)$ , and we sum over repeated indices. Then, a (strong) wave map is a map  $u = (u^1, \dots, u^\ell) : [0, \mathbb{T}) \times \Omega \rightarrow N \hookrightarrow \mathbb{R}^\ell$  that satisfies

$$\partial_t^2 u - \Delta u = A[u](Du, Du), \tag{6}$$

where  $A[u](Du, Du)$  stands for  $(A^i_{jk}|_u \partial_\alpha u^j \partial^\alpha u^k)|_{1 \leq i \leq \ell}$ ; see [3].

A significant part of our analysis will consider the case in which the target manifold  $N$  is the 2-sphere  $S^2 \subset \mathbb{R}^3$  and in this case (6) reduces to

$$\partial_t^2 u - \Delta u = (|\nabla u|^2 - |\partial_t u|^2)u \tag{7}$$

$$\text{with point-wise constraint } |u(t, x)| = 1. \tag{8}$$

Let us also mention that, using angular momentum  $\omega := \partial_t u \times u$ , the wave map equation can be phrased as [16]

$$\partial_t u = u \times \omega \quad \text{and} \quad \partial_t \omega = \Delta u \times u. \tag{9}$$

This variant is the one underlying the numerical scheme that we will study.

### 4 A quantitative stability estimate

In this section, we establish a weak–strong stability result that complements weak–strong uniqueness results in [30, 31] by providing bounds for differences between solutions. In particular, our estimates quantify the impact of residuals that is crucial for the use of stability results in proving a posteriori error estimates. We prove weak–strong stability in the general case as well as in the special case of  $N = S^2$  since

we believe that the former nicely highlights the general geometric structure while the latter proof uses very elementary techniques and does not require any background in differential geometry.

We denote by  $(\cdot, \cdot)$  the  $L^2$  inner product on  $(0, \mathbb{T}) \times \Omega$  and by  $(\cdot, \cdot)_\Omega$  the  $L^2$  inner product on  $\Omega$ .

**Definition 1** Given  $u_0 \in H^1(\Omega, N)$  and  $u_1 \in L^2(\Omega, \mathbb{R}^\ell)$  so that  $u_1(x) \in T_{u_0(x)}N$  for almost all  $x$ , we call a function  $u \in L^2([0, \mathbb{T}) \times \Omega)$  with values in  $N$  and  $\nabla u, \partial_t u \in L^\infty(0, \mathbb{T}; L^2(\Omega))$  a *finite energy weak solution* of (1)-(3) provided the following conditions are satisfied

1.  $u(0, \cdot) = u_0$
2.  $(\partial_t u, \partial_t \psi) - (\nabla u, \nabla \psi) = -(A[u](Du, Du), \psi) - (u_1, \psi(0, \cdot))_\Omega$  for all  $\psi \in C_c^\infty([0, \mathbb{T}) \times \bar{\Omega})$
3.  $E[u(t), \partial_t u(t)] \leq E[u_0, u_1]$  for almost all  $0 < t < \mathbb{T}$ .

We say that  $u$  satisfies the *local energy condition* provided

$$E[u(t), \partial_t u(t)] \leq E[u(s), \partial_t u(s)] \text{ for almost all } 0 < s < t < \mathbb{T}$$

**Remark 1** Note that the weak formulation in Definition 1 actually holds for all  $\psi \in W^{1,1}(0, \mathbb{T}; L^2(\Omega)) \cap L^1(0, \mathbb{T}; L^\infty \cap H^1(\Omega))$  with  $\psi(\mathbb{T}, \cdot) = 0$ , due to a density argument where passing to the limit in  $(A[u](Du, Du), \psi)$  is done using dominated convergence; see the Appendix of [34] for details.

We are going to compare a finite energy weak solution  $u$  to a more regular solution  $(\tilde{u}, \tilde{w})$  of the perturbed problem

$$\partial_t \tilde{u} = \tilde{u} \times \tilde{w} + r_u, \quad \partial_t \tilde{w} = \Delta \tilde{u} \times \tilde{u} + r_w. \tag{10}$$

**Lemma 1** For any two sufficiently regular functions  $\tilde{u} : \Omega \rightarrow S^2, \tilde{w} : \Omega \rightarrow \mathbb{R}^3$ , the following identities hold:

- (a)  $(\tilde{u} \times \tilde{w}) \times \tilde{w} = (\tilde{u} \cdot \tilde{w})\tilde{w} - |\tilde{w}|^2\tilde{u}$
- (b)  $\tilde{u} \times (\Delta \tilde{u} \times \tilde{u}) = \Delta \tilde{u} + |\nabla \tilde{u}|^2\tilde{u}$

**Proof** Part (a) immediately follows from  $(a \times b) \times c = (a \cdot c)b - (b \cdot c)a$  for any  $a, b, c \in \mathbb{R}^3$ . In order to prove (b), we note that  $(\partial_{x_j} \tilde{u}) \cdot \tilde{u} = \frac{1}{2} \partial_{x_j} |\tilde{u}|^2 = 0$  so that

$$\begin{aligned} \tilde{u} \times (\Delta \tilde{u} \times \tilde{u}) &= \sum_{j=1}^3 [\partial_{x_j} (\tilde{u} \times (\partial_{x_j} \tilde{u} \times \tilde{u})) - (\partial_{x_j} \tilde{u}) \times (\partial_{x_j} \tilde{u} \times \tilde{u})] \\ &= \sum_{j=1}^3 [\partial_{x_j} (|\tilde{u}|^2 \partial_{x_j} \tilde{u} - (\partial_{x_j} \tilde{u} \cdot \tilde{u})\tilde{u}) - (\partial_{x_j} \tilde{u} \cdot \tilde{u}) \partial_{x_j} \tilde{u} + |\partial_{x_j} \tilde{u}|^2 \tilde{u}] \\ &= \sum_{j=1}^3 [\partial_{x_j} (\partial_{x_j} \tilde{u}) + |\partial_{x_j} \tilde{u}|^2 \tilde{u}]. \end{aligned}$$

□

**Theorem 1** *Let  $u$  be a finite energy weak solution of (1)-(3) satisfying the local energy condition. Let  $s = m$  for  $m \geq 3$  and  $s \in (2, \infty]$  for  $m = 2$ . Let  $(\tilde{u}, \tilde{w})$  be a solution of the perturbed problem (10). We ask for the following regularity:  $\tilde{u} \in L^\infty((0, \mathbb{T}) \times \Omega, S^2)$  and*

$$\begin{aligned} \nabla \tilde{u}, \tilde{w} \in & \left( L^1(0, \mathbb{T}; H^1(\Omega)) \cap L^\infty(0, \mathbb{T}; L^2(\Omega)) \right) \\ & \cap L^k(0, \mathbb{T}; L^\infty(\Omega)) \cap L^{\frac{k}{k-1}}(0, \mathbb{T}; L^s(\Omega)) \end{aligned}$$

for some  $k \in [1, 2]$  and with given functions  $r_w \in L^1(0, \mathbb{T}; L^2(\Omega))$  and  $r_u \in L^1(0, \mathbb{T}; H^1(\Omega)) \cap L^{\frac{k}{k-1}}(0, \mathbb{T}; L^2(\Omega))$ . Then, the difference at time  $t$  can be controlled via the difference in initial data and perturbation terms. Indeed,  $\mathcal{H}$  defined by

$$\mathcal{H}(t) := \frac{1}{2} \int_{\Omega} |\partial_t u(t) - \tilde{u}(t) \times \tilde{w}(t)|^2 + |\nabla u(t) - \nabla \tilde{u}(t)|^2 + |u(t) - \tilde{u}(t)|^2 dx,$$

satisfies

$$\sqrt{\mathcal{H}(t_2)} \leq \left( \sqrt{\mathcal{H}(t_1)} + \left( \int_{t_1}^{t_2} \alpha(t) dt \right) \right) \times \exp \left( \frac{1}{2} \int_{t_1}^{t_2} \delta(t) dt \right) \tag{11}$$

for almost all  $0 < t_1 < t_2 < \mathbb{T}$  with

$$\begin{aligned} \alpha &:= \|r_g + r_u \times \tilde{w} + \tilde{u} \times r_w\|_{L^2(\Omega)} + \|r_u\|_{L^2(\Omega)} + \|\nabla r_u\|_{L^2(\Omega)} \\ r_g &:= (\tilde{u} \cdot \tilde{w})\tilde{w} - |\tilde{u} \cdot \tilde{w}|^2\tilde{u}; \quad \mathcal{A}[\tilde{u}] := |\nabla \tilde{u}|^2 - |\tilde{u} \times \tilde{w}|^2 \\ \delta &:= \left( 1 + c_q \|\mathcal{A}[\tilde{u}]\|_{L^s(\Omega)} + 2c_q |\tilde{u} \times \tilde{w}|_{L^{2s}(\Omega)}^2 \right. \\ &\quad \left. + 2c_q \|\nabla \tilde{u}\|_{L^{2s}(\Omega)} |\tilde{u} \times \tilde{w}|_{L^{2s}(\Omega)} + 4|\tilde{u} \times \tilde{w}|_{L^\infty(\Omega)} \right) \end{aligned}$$

where  $c_q$  is the squared constant of the Sobolev embedding  $H^1(\Omega) \rightarrow L^q(\Omega)$  with  $q = \frac{2s}{s-2}$ . This implies that for almost any sequence  $0 = t_0 < t_1 < \dots < t_N \leq \mathbb{T}$

$$\begin{aligned} \sqrt{\mathcal{H}(t_N)} \leq & \sqrt{\mathcal{H}(0)} \exp \left( \frac{1}{2} \int_0^{t_N} \delta(t) dt \right) \\ & + \sum_{j=1}^N \left( \int_{t_{j-1}}^{t_j} \alpha(t) dt \right) \times \exp \left( \frac{1}{2} \int_{t_{j-1}}^{t_N} \delta(t) dt \right) \end{aligned} \tag{12}$$

**Remark 2** (Constants in the estimate) Note that  $s, k$  are given by assumption and fixed. Explicit upper bounds for  $c_q$  (with  $q = 2s/(s - 2)$ ), for a variety of domains, can be found in [35].

**Remark 3** (Regularity of weak solution) Note that if  $u$  is a finite energy weak solution, not necessarily satisfying the local energy condition, then (11) still holds in the special case  $t_1 = 0$ .

**Remark 4** (Regularity of strong solution) Note that our regularity assumptions with  $k = 1$  correspond to the conditions in Struwe’s result [30, Thm 2.2]. The case studied in [30, Thm 2.2] corresponds to  $m = 2$  and  $r_u = r_\omega = r_g = 0$  in our notation so that the space-time gradient  $D\tilde{u}$  corresponds to the vector  $(\tilde{u} \times \tilde{w}, \nabla\tilde{u})$ . Thus, choosing  $k = 1, s = 2 + \epsilon$ , we assume  $\nabla\tilde{u}, \tilde{w} \in (L^1(0, \mathbb{T}; H^1(\Omega) \cap L^\infty(\Omega)) \cap L^\infty(0, \mathbb{T}; L^{2+\epsilon}(\Omega))) =: X$ .

Struwe assumes  $D\tilde{u} \in X$  which also holds under our assumptions mainly because  $\tilde{u} \in L^\infty((0, \mathbb{T}) \times \Omega)$ . There is the difference that we consider Neumann boundary data and Struwe has a solution on full  $\mathbb{R}^2$ .

Note furthermore that the conditions in our theorem allow us to use  $\tilde{u} \times \tilde{w}$  as test function in (2) of Definition 1, i.e.  $\tilde{u} \times \tilde{w} \in W^{1,1}(0, \mathbb{T}; L^2(\Omega)) \cap L^1(0, \mathbb{T}; L^\infty \cap H^1(\Omega))$ . This can be seen as follows: First, we observe that  $\tilde{u} \times \tilde{w} \in L^1(0, \mathbb{T}; L^\infty(\Omega))$  and  $\nabla(\tilde{u} \times \tilde{w}) \in L^1(0, \mathbb{T}; L^2(\Omega))$  since, using Young’s inequality,  $\int \|\nabla\tilde{u}\|_{L^s(\Omega)} \|\tilde{w}\|_{L^q(\Omega)} dt \leq c \int \|\nabla\tilde{u}\|_{L^s(\Omega)}^{\frac{k}{k-1}} + \|\tilde{w}\|_{L^q(\Omega)}^k dt < \infty$  ( $\frac{1}{s} + \frac{1}{q} = \frac{1}{2}$ ). It remains to check  $\partial_t(\tilde{u} \times \tilde{w}) \in L^1(0, \mathbb{T}; L^2(\Omega))$ : Using Lemma 1, we know that

$$\begin{aligned} \partial_t(\tilde{u} \times \tilde{w}) &= \partial_t\tilde{u} \times \tilde{w} + \tilde{u} \times \partial_t\tilde{w} \\ &= (\tilde{u} \times \tilde{w}) \times \tilde{w} + r_u \times \tilde{w} + \tilde{u} \times (\Delta\tilde{u} \times \tilde{u}) + \tilde{u} \times r_w \\ &= r_g + (|\nabla\tilde{u}|^2 - |\tilde{u} \times \tilde{w}|^2)\tilde{u} + r_u \times \tilde{w} + \Delta\tilde{u} + \tilde{u} \times r_w, \end{aligned} \tag{13}$$

where we have used  $|\tilde{u} \times \tilde{w}|^2 = |\tilde{w}|^2 - |\tilde{w} \cdot \tilde{u}|^2$ . Note that  $r_g \sim |\tilde{w}|^2$ . We always have that  $\tilde{w}, \nabla\tilde{u} \in L^2(0, \mathbb{T}; L^4(\Omega))$  because of

$$\begin{aligned} &\int \|\tilde{w}\|_{L^4(\Omega)}^2 dt \leq \int \|\tilde{w}\|_{L^\infty(\Omega)} \|\tilde{w}\|_{L^2(\Omega)} dt \\ &\leq \begin{cases} \text{esssup}_t \|\tilde{w}\|_{L^2(\Omega)} \int \|\tilde{w}\|_{L^\infty(\Omega)} dt < \infty & \text{if } k = 1 \\ c \int \|\tilde{w}\|_{L^\infty(\Omega)}^k + \|\tilde{w}\|_{L^2(\Omega)}^{\frac{k}{k-1}} dt < \infty & \text{if } k > 1. \end{cases} \end{aligned}$$

The conditions on  $r_u$  imply that also  $r_u \times \tilde{w} \in L^1(0, \mathbb{T}; L^2(\Omega))$ . Thus, all terms in (13) are controlled in  $L^1(0, \mathbb{T}; L^2(\Omega))$ .

**Proof of Theorem 1** Let us fix some (arbitrary)  $0 < t_1 < t_2 < \mathbb{T}$  and let us define for any  $0 < \epsilon < \min\{\mathbb{T} - t_2, t_2 - t_1\}$  the map

$$\phi_\epsilon(t) := \begin{cases} 0 & : t \leq t_1 \\ \frac{t-t_1}{\epsilon} & : t_1 \leq t < t_1 + \epsilon \\ 1 & : t_1 + \epsilon \leq t \leq t_2 \\ 1 - \frac{t-t_2}{\epsilon} & : t_2 \leq t \leq t_2 + \epsilon \\ 0 & : t_2 + \epsilon < t \end{cases}$$



We will study  $\lim_{\varepsilon \searrow 0} \int_0^\mathbb{T} \mathcal{H}(t) \partial_t \phi_\varepsilon(t) dt$ . On the one hand, for every pair of Lebesgue points of  $t \mapsto E[u(t), \partial_t u(t)]$ , we have

$$\lim_{\varepsilon \searrow 0} \int_0^\mathbb{T} \mathcal{H}(t) \partial_t \phi_\varepsilon(t) dt = - \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon} \left( \int_{t_2}^{t_2+\varepsilon} \mathcal{H}(t) dt - \int_{t_1}^{t_1+\varepsilon} \mathcal{H}(t) dt \right) = -\mathcal{H}(t_2) + \mathcal{H}(t_1). \tag{14}$$

On the other hand, we may decompose the integral at hand as

$$\begin{aligned} \int_0^\mathbb{T} \mathcal{H}(t) \partial_t \phi_\varepsilon(t) dt &= \int_0^\mathbb{T} E[u(t), \partial_t u(t)] \partial_t \phi_\varepsilon(t) dt \\ &\quad - \int_0^\mathbb{T} (\partial_t u \cdot \tilde{u} \times \tilde{w} + \nabla u \cdot \nabla \tilde{u}) \partial_t \phi_\varepsilon(t) dt \\ &\quad + \int_0^\mathbb{T} E[\tilde{u}(t), \tilde{u}(t) \times \tilde{w}(t)] \partial_t \phi_\varepsilon(t) dt + \int_0^\mathbb{T} \frac{1}{2} |u(t) - \tilde{u}(t)|^2 \partial_t \phi_\varepsilon(t) dt \\ &=: E_\varepsilon^1 - E_\varepsilon^2 + E_\varepsilon^3 + E_\varepsilon^4. \end{aligned} \tag{15}$$

Concerning  $E_\varepsilon^1$ , we observe that for every pair of Lebesgue points of  $t \mapsto E[u(t), \partial_t u(t)]$

$$\begin{aligned} \lim_{\varepsilon \searrow 0} E_\varepsilon^1 &= \lim_{\varepsilon \searrow 0} \int_0^\mathbb{T} E[u(t), \partial_t u(t)] \partial_t \phi_\varepsilon(t) dt \\ &= - \lim_{\varepsilon \searrow 0} \frac{1}{\varepsilon} \left( \int_{t_2}^{t_2+\varepsilon} E[u(t), \partial_t u(t)] dt - \int_{t_1}^{t_1+\varepsilon} E[u(t), \partial_t u(t)] dt \right) \\ &= -E[u(t_2), \partial_t u(t_2)] + E[u(t_1), \partial_t u(t_1)] \geq 0. \end{aligned} \tag{16}$$

Next, we consider  $E_\varepsilon^2$  for fixed  $\varepsilon$  and observe

$$\begin{aligned} E_\varepsilon^2 &= (\partial_t u, \tilde{u} \times \tilde{w} \partial_t \phi_\varepsilon) + (\nabla u, \nabla \tilde{u} \partial_t \phi_\varepsilon) \\ &= (\partial_t u, \partial_t (\tilde{u} \times \tilde{w} \phi_\varepsilon)) - (\partial_t u, \partial_t (\tilde{u} \times \tilde{w})) \phi_\varepsilon + (\nabla u, \nabla \tilde{u} \partial_t \phi_\varepsilon) \\ &= (\nabla u, \nabla (\tilde{u} \times \tilde{w} \phi_\varepsilon)) - (|\nabla u|^2 - |\partial_t u|^2) u, \tilde{u} \times \tilde{w} \phi_\varepsilon \\ &\quad - (\partial_t u, \partial_t (\tilde{u} \times \tilde{w}) \phi_\varepsilon) + (\nabla u, \nabla \tilde{u} \partial_t \phi_\varepsilon), \end{aligned} \tag{17}$$

where we have used the weak formulation in the last step. We also note that

$$\begin{aligned} (\nabla u, \nabla (\tilde{u} \times \tilde{w} \phi_\varepsilon)) + (\nabla u, \nabla \tilde{u} \partial_t \phi_\varepsilon) &= (\nabla u, \nabla (\partial_t \tilde{u} \phi_\varepsilon)) - (\nabla u, \nabla r_u \phi_\varepsilon) \\ &\quad + (\nabla u, \nabla \tilde{u} \partial_t \phi_\varepsilon) \\ &= (\nabla u, \nabla \partial_t (\tilde{u} \phi_\varepsilon)) - (\nabla u, \nabla r_u \phi_\varepsilon) \end{aligned} \tag{18}$$

due to (10)<sub>1</sub> and the fact that  $\phi_\varepsilon$  is independent of  $x$ . We insert (13) and (18) into (17) and obtain

$$\begin{aligned}
 E_\varepsilon^2 &= (\nabla u, \nabla \partial_t(\tilde{u}\phi_\varepsilon)) - (\nabla u, \nabla r_u\phi_\varepsilon) - ( (|\nabla u|^2 - |\partial_t u|^2)u, \tilde{u} \times \tilde{w}\phi_\varepsilon) \\
 &\quad - (\partial_t u, (r_g + (|\nabla \tilde{u}|^2 - |\tilde{u} \times \tilde{w}|^2)\tilde{u} + r_u \times \tilde{w} + \Delta \tilde{u} + \tilde{u} \times r_w)\phi_\varepsilon) \\
 &= -(\nabla u, \nabla r_u\phi_\varepsilon) - ( (|\nabla u|^2 - |\partial_t u|^2)u, \tilde{u} \times \tilde{w}\phi_\varepsilon) \\
 &\quad - (\partial_t u, (r_g + (|\nabla \tilde{u}|^2 - |\tilde{u} \times \tilde{w}|^2)\tilde{u} + r_u \times \tilde{w} + \tilde{u} \times r_w)\phi_\varepsilon)
 \end{aligned} \tag{19}$$

where we have used integration by parts and  $\int_0^{\mathbb{T}} \partial_t(\nabla u, \nabla \tilde{u}\phi_\varepsilon)_\Omega = 0$  in the last equality. Equation (19) allows us to conclude

$$\begin{aligned}
 \lim_{\varepsilon \searrow 0} E_\varepsilon^2 &= - \int_{t_1}^{t_2} \int_\Omega \{ \partial_t u \cdot [r_g + (|\nabla \tilde{u}|^2 - |\tilde{u} \times \tilde{w}|^2)\tilde{u} + r_u \times \tilde{w} + \tilde{u} \times r_w] \\
 &\quad + \nabla u \cdot \nabla r_u + (|\nabla u|^2 - |\partial_t u|^2)u \cdot (\tilde{u} \times \tilde{w}) \} dx dt.
 \end{aligned} \tag{20}$$

Concerning  $E_\varepsilon^3$ , we use (13) and integration by parts to obtain

$$\begin{aligned}
 E_\varepsilon^3 &= \frac{1}{2}(\tilde{u} \times \tilde{w}, \tilde{u} \times \tilde{w}\partial_t\phi_\varepsilon) + \frac{1}{2}(\nabla \tilde{u}, \nabla \tilde{u}\partial_t\phi_\varepsilon) \\
 &= -(\tilde{u} \times \tilde{w}\phi_\varepsilon, r_g + (|\nabla \tilde{u}|^2 - |\tilde{u} \times \tilde{w}|^2)\tilde{u} + r_u \times \tilde{w} + \Delta \tilde{u} + \tilde{u} \times r_w) + (\partial_t \tilde{u}, \Delta \tilde{u}\phi_\varepsilon) \\
 &= -(\tilde{u} \times \tilde{w}\phi_\varepsilon, r_g + r_u \times \tilde{w} + \tilde{u} \times r_w) + (r_u, \Delta \tilde{u}\phi_\varepsilon),
 \end{aligned} \tag{21}$$

where we have used point-wise orthogonality of  $\tilde{u}$  to  $\tilde{u} \times \tilde{w}$  in the last equality. Equation (21) allows us to conclude

$$\lim_{\varepsilon \searrow 0} E_\varepsilon^3 = - \int_{t_1}^{t_2} \int_\Omega (\tilde{u} \times \tilde{w}) \cdot [r_g + r_u \times \tilde{w} + \tilde{u} \times r_w] + \nabla r_u \cdot \nabla \tilde{u} dx dt. \tag{22}$$

Finally, we find

$$\lim_{\varepsilon \searrow 0} E_\varepsilon^4 = - \int_{t_1}^{t_2} \int_\Omega (u - \tilde{u}) \cdot (\partial_t u - \tilde{u} \times \tilde{w} - r_u) dx dt. \tag{23}$$

We combine (14), (16), (20), (22) and (23) to obtain

$$\begin{aligned}
 &\mathcal{H}(t_2) - \mathcal{H}(t_1) \\
 &\leq \int_{t_1}^{t_2} \int_\Omega (\tilde{u} \times \tilde{w} - \partial_t u) \cdot [r_g + r_u \times \tilde{w} + \tilde{u} \times r_w] - \nabla r_u \cdot (\nabla u - \nabla \tilde{u}) dx dt \\
 &\quad + \int_{t_1}^{t_2} \int_\Omega (u - \tilde{u}) \cdot (\partial_t u - \tilde{u} \times \tilde{w}) - (u - \tilde{u}) \cdot r_u dx dt \\
 &\quad - \int_{t_1}^{t_2} \int_\Omega (|\nabla \tilde{u}|^2 - |\tilde{u} \times \tilde{w}|^2)\tilde{u} \cdot \partial_t u + (|\nabla u|^2 - |\partial_t u|^2)u \cdot \tilde{u} \times \tilde{w} dx dt.
 \end{aligned} \tag{24}$$

Let us define  $\mathcal{A}[u] := (|\nabla u|^2 - |\partial_t u|^2)$  and (with a slight abuse of notation)  $\mathcal{A}[\tilde{u}] := (|\nabla \tilde{u}|^2 - |\tilde{u} \times \tilde{w}|^2)$ . Then,  $u \perp \partial_t u$  and  $\tilde{u} \perp (\tilde{u} \times \tilde{w})$  allows us to infer

$$\partial_t u \cdot \mathcal{A}[\tilde{u}]\tilde{u} + (\tilde{u} \times \tilde{w}) \cdot \mathcal{A}[u]u = -(\partial_t u - \tilde{u} \times \tilde{w}) \cdot \mathcal{A}[\tilde{u}](u - \tilde{u}) + \tilde{u} \times \tilde{w}(\mathcal{A}[u] - \mathcal{A}[\tilde{u}])(u - \tilde{u}).$$

We use  $\frac{1}{s} + \frac{1}{q} = \frac{1}{2}$  to get that

$$\begin{aligned} & \int_{t_1}^{t_2} \|\partial_t u \cdot \mathcal{A}[\tilde{u}]\tilde{u} + (\tilde{u} \times \tilde{w}) \cdot \mathcal{A}[u]u\|_{L^1(\Omega)} dt \\ & \leq \int_{t_1}^{t_2} \|\partial_t u - \tilde{u} \times \tilde{w}\|_{L^2(\Omega)} \|\mathcal{A}[\tilde{u}]\|_{L^s(\Omega)} \|u - \tilde{u}\|_{L^q(\Omega)} \\ & \quad + \|(\tilde{u} \times \tilde{w})(\mathcal{A}[u] - \mathcal{A}[\tilde{u}])(u - \tilde{u})\|_{L^1(\Omega)} dt \\ & \leq \int_{t_1}^{t_2} \|\partial_t u - \tilde{u} \times \tilde{w}\|_{L^2(\Omega)} \left( \|\mathcal{A}[\tilde{u}]\|_{L^s(\Omega)} + 2\|\tilde{u} \times \tilde{w}\|_{L^{2s}(\Omega)}^2 \right) \|u - \tilde{u}\|_{L^q(\Omega)} \\ & \quad + 2\|\nabla \tilde{u}\|_{L^{2s}(\Omega)} \|\tilde{u} \times \tilde{w}\|_{L^{2s}(\Omega)} \|\nabla u - \nabla \tilde{u}\|_{L^2(\Omega)} \|u - \tilde{u}\|_{L^q(\Omega)} \\ & \quad + 2\|\nabla u - \nabla \tilde{u}\|_{L^2(\Omega)}^2 \|\tilde{u} \times \tilde{w}\|_{L^\infty(\Omega)} + 2\|\partial_t u - \tilde{u} \times \tilde{w}\|_{L^2(\Omega)}^2 \|\tilde{u} \times \tilde{w}\|_{L^\infty(\Omega)} dt \end{aligned} \tag{25}$$

where we have used that

$$\begin{aligned} \mathcal{A}[u] - \mathcal{A}[\tilde{u}] &= 2\nabla \tilde{u} \cdot (\nabla u - \nabla \tilde{u}) + |\nabla u - \nabla \tilde{u}|^2 \\ & \quad - 2(\tilde{u} \times \tilde{w}) \cdot (\partial_t u - \tilde{u} \times \tilde{w}) - |\partial_t u - \tilde{u} \times \tilde{w}|^2 \end{aligned} \tag{26}$$

and  $u, \tilde{u} \in S^2$ . We insert (25) into (24) and obtain, using that  $H^1(\Omega)$  embeds into  $L^q(\Omega)$ ,

$$\mathcal{H}(t_2) \leq \mathcal{H}(t_1) + \int_{t_1}^{t_2} 2\sqrt{\mathcal{H}(t)}\alpha(t)dt + \int_{t_1}^{t_2} \delta(t)\mathcal{H}(t)dt. \tag{27}$$

According to [36, Thm 21], (27) implies (11). Equation (12) follows by taking the square-root of (11) and induction in  $j$ . □

We now come to the computations for a general closed target manifold  $N$ . Since for general  $N$ , we use one second order equation instead of a system of first order equations we have only one residual, but we split this into two parts, assuming that one part can be expressed as a time derivative.

**Theorem 2** *Let  $u$  be a weak solution of the problem (1)-(3) and  $\tilde{u} = (\tilde{u}^1, \dots, \tilde{u}^\ell) : [0, \mathbb{T}) \times \Omega \rightarrow N$  a strong solution of a perturbed problem*

$$\partial_t^2 \tilde{u} - \Delta \tilde{u} = A[\tilde{u}](D\tilde{u}, D\tilde{u}) + R_1 + \partial_t R_2, \quad \partial_n \tilde{u} = 0.$$

Let  $s = m$  for  $m \geq 3$  and  $s \in (2, \infty]$  for  $m = 2$ . We assume that there is a  $k \in [1, \frac{4}{3}]$  such that

$$\begin{aligned} \tilde{u} &\in W^{2,1}((0, \mathbb{T}) \times \Omega), \\ D\tilde{u} &\in L^1(0, \mathbb{T}; H^1(\Omega)) \cap L^\infty(0, \mathbb{T}; L^2(\Omega)) \cap L^k(0, \mathbb{T}; L^\infty(\Omega)) \cap L^{\frac{k}{k-1}}(0, \mathbb{T}; L^s(\Omega)), \\ R_1 &\in L^2((0, \mathbb{T}) \times \Omega), \\ R_2 &\in H^1((0, \mathbb{T}) \times \Omega) \cap L^1(0, \mathbb{T}; L^\infty(\Omega)) \cap L^{\frac{k}{k-1}}(0, \mathbb{T}; L^2(\Omega)). \end{aligned}$$

We define the energy  $\mathcal{I}[u(t, \cdot)] := \frac{1}{2} \int_\Omega |u(t, \cdot)|^2 + |\partial_t u(t, \cdot)|^2 + |\nabla u(t, \cdot)|^2 dx$ . Then, we have the inequality

$$\sup_{0 \leq t_1 \leq t_2} \mathcal{I}[u(t_1, \cdot) - \tilde{u}(t_1, \cdot)] \leq \gamma(t_2) \cdot \exp\left(\int_0^{t_2} \delta(t) dt\right), \quad \text{where} \quad (28)$$

$$\begin{aligned} \gamma(t_2) &:= 8\mathcal{I}[u - \tilde{u}]|_{t=0} + 4\|R_2(0, \cdot)\|_{L^2(\Omega)}^2 + 2 \sup_{0 \leq t_1 \leq t_2} \|R_2(t_1, \cdot)\|_{L^2(\Omega)}^2 \\ &\quad + 4 \int_0^{t_2} \|R_1 R_2\|_{L^1(\Omega)} dt \\ &\quad + 8 \left( \int_0^{t_2} \|\nabla R_2\|_{L^2(\Omega)} + 2c_1 \|D\tilde{u}\|_{L^\infty(\Omega)} \|R_2\|_{L^2(\Omega)} \right. \\ &\quad \left. + c_1 \sqrt{c_2} \|R_2\|_{L^q(\Omega)} \|D\tilde{u}\|_{L^s(\Omega)}^2 + \|R_1\|_{L^2(\Omega)} dt \right)^2, \\ \delta(t) &:= 4 + 4c_1 \left( 3c_2 \|D\tilde{u}\|_{L^{2s}(\Omega)}^2 + 2c_3 \|\partial_t \tilde{u}\|_{L^\infty(\Omega)} + 2\|R_2\|_{L^\infty(\Omega)} \right). \end{aligned}$$

The constant  $c_3$  is the diameter of  $N$ ,  $c_3 = \text{diam}(N)$ ,  $c_2$  is the squared constant from the Sobolev embedding  $H^1 \hookrightarrow L^q$ ,  $q := \frac{2s}{s-2}$ . The constant  $c_1 = c_1(N)$  is the maximum of the Lipschitz constant of  $p \mapsto A[p](\cdot, \cdot)$  and  $\sup_N |A|$ .

**Proof** We pretend that  $u$  is a strong solution in order to keep the computations simpler. For a weak solution, we use the cutoff-trick with  $\phi_\epsilon$ , that was introduced in the proof of Theorem 1. Let us denote

$$\frac{d}{dt} \mathcal{I}[u - \tilde{u}] = \frac{d}{dt} \left( \frac{1}{2} \int_\Omega |u - \tilde{u}|^2 + |\partial_t u - \partial_t \tilde{u}|^2 + |\nabla u - \nabla \tilde{u}|^2 dx \right) =: \frac{d}{dt} (I^1 + I^2 + I^3).$$

Then, we compute

$$\begin{aligned} \frac{d}{dt} (I^2 + I^3) &= (\partial_t u - \partial_t \tilde{u}, \partial_{tt} u - \partial_{tt} \tilde{u})_\Omega + (\nabla u - \nabla \tilde{u}, \partial_t \nabla u - \partial_t \nabla \tilde{u})_\Omega \\ &= (\partial_t u - \partial_t \tilde{u}, \partial_{tt} u - \partial_{tt} \tilde{u})_\Omega - (\Delta u - \Delta \tilde{u}, \partial_t u - \partial_t \tilde{u})_\Omega \\ &= (\partial_t u - \partial_t \tilde{u}, A[u](Du, Du) - A[\tilde{u}](D\tilde{u}, D\tilde{u}) - R_1 - \partial_t R_2)_\Omega. \end{aligned}$$

As  $A$  has normal values, we get that  $\partial_t u \cdot A[u](\cdot, \cdot) = 0 = \partial_t \tilde{u} \cdot A[\tilde{u}](\cdot, \cdot)$ . Using this orthogonality, we get that

$$\begin{aligned} \frac{d}{dt}(I^2 + I^3) &= (\partial_t u, A[u](D\tilde{u}, D\tilde{u}) - A[\tilde{u}](D\tilde{u}, D\tilde{u}))_\Omega - (\partial_t u - \partial_t \tilde{u}, R_1 + \partial_t R_2)_\Omega \\ &\quad - (\partial_t \tilde{u}, A[u](Du, Du) - A[\tilde{u}](Du, Du))_\Omega \\ &= (\partial_t u - \partial_t \tilde{u}, (A[u] - A[\tilde{u}])(D\tilde{u}, D\tilde{u}))_\Omega - (\partial_t u - \partial_t \tilde{u}, R_1 + \partial_t R_2)_\Omega \\ &\quad - (\partial_t \tilde{u}, (A[u] - A[\tilde{u}])(Du, Du) - (A[u] - A[\tilde{u}])(D\tilde{u}, D\tilde{u}))_\Omega. \end{aligned} \tag{29}$$

We consider the term  $(\partial_t u - \partial_t \tilde{u}, \partial_t R_2)_\Omega$  that is contained in (29) and compute

$$\begin{aligned} (\partial_t u - \partial_t \tilde{u}, \partial_t R_2)_\Omega &= \frac{d}{dt}(\partial_t u - \partial_t \tilde{u}, R_2)_\Omega - (\Delta u - \Delta \tilde{u}, R_2)_\Omega \\ &\quad - (A[u](Du, Du) - A[\tilde{u}](D\tilde{u}, D\tilde{u}) - R_1 - \partial_t R_2, R_2)_\Omega \\ &= \frac{d}{dt}(\partial_t u - \partial_t \tilde{u}, R_2)_\Omega + (\nabla u - \nabla \tilde{u}, \nabla R_2)_\Omega \\ &\quad - (A[u](Du, Du) - A[\tilde{u}](D\tilde{u}, D\tilde{u}), R_2)_\Omega + (R_1, R_2)_\Omega \\ &\quad + \frac{1}{2} \frac{d}{dt}(R_2, R_2)_\Omega. \end{aligned} \tag{30}$$

In order to estimate the term  $(A[u](Du, Du) - A[\tilde{u}](D\tilde{u}, D\tilde{u}), R_2)_\Omega$ , we do the following:

$$\begin{aligned} &(A[u](Du, Du) - A[\tilde{u}](D\tilde{u}, D\tilde{u})) \cdot R_2 \\ &= (A[u](Du, Du) - A[u](D\tilde{u}, D\tilde{u}) + (A[u] - A[\tilde{u}])(D\tilde{u}, D\tilde{u})) \cdot R_2 \\ &= (A[u](Du - D\tilde{u}, Du - D\tilde{u}) + 2A[u](Du - D\tilde{u}, D\tilde{u}) + (A[u] - A[\tilde{u}])(D\tilde{u}, D\tilde{u})) \cdot R_2 \end{aligned}$$

where we used the equation  $B(X, X) - B(Y, Y) = B(X - Y, (X - Y) + 2Y)$  for a symmetric bilinear form  $B : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ . The map  $p \mapsto A[p](\cdot, \cdot)$  is Lipschitz and  $N$  is a compact smooth manifold. Therefore, there is a constant  $c_1 = c_1(N)$  such that

$$\begin{aligned} &|(A[u](Du, Du) - A[\tilde{u}](D\tilde{u}, D\tilde{u})) \cdot R_2| \\ &\leq c_1 |R_2| \left( |Du - D\tilde{u}|^2 + 2|Du - D\tilde{u}||D\tilde{u}| + |u - \tilde{u}||D\tilde{u}|^2 \right) \end{aligned}$$

Using Hölder’s inequality and generalized Hölder’s inequality, we get that

$$\begin{aligned} &(A[u](Du, Du) - A[\tilde{u}](D\tilde{u}, D\tilde{u}), R_2)_\Omega \\ &\leq 2c_1 \|R_2\|_{L^\infty(\Omega)} \mathcal{I}[u - \tilde{u}] + 2c_1 \|D\tilde{u}\|_{L^\infty(\Omega)} \|R_2\|_{L^2(\Omega)} \sqrt{2\mathcal{I}[u - \tilde{u}]} \\ &\quad + c_1 \|R_2\|_{L^q(\Omega)} \|u - \tilde{u}\|_{L^q(\Omega)} \|D\tilde{u}\|_{L^{2q}(\Omega)}^2 \end{aligned} \tag{31}$$

for any pair  $(q, \tilde{q})$  with  $\frac{2}{q} + \frac{1}{\tilde{q}} = 1$ . For  $m \geq 3$ , we use this inequality for  $q = \frac{2m}{m-2}$  and get by Sobolev embedding  $H^1(\Omega) \hookrightarrow L^{\frac{2m}{m-2}}(\Omega)$

$$\begin{aligned} & (A[u](Du, Du) - A[\tilde{u}](D\tilde{u}, D\tilde{u}), R_2)_\Omega \\ & \leq 2c_1 \|R_2\|_{L^\infty(\Omega)} \mathcal{I}[u - \tilde{u}] + 2c_1 \|D\tilde{u}\|_{L^\infty(\Omega)} \|R_2\|_{L^2(\Omega)} \sqrt{2\mathcal{I}[u - \tilde{u}]} \quad (32) \\ & \quad + c_1 \sqrt{c_2} \|R_2\|_{L^q(\Omega)} \|D\tilde{u}\|_{L^{2\tilde{q}}(\Omega)}^2 \sqrt{2\mathcal{I}[u - \tilde{u}]}, \end{aligned}$$

where  $\tilde{q} = \frac{m}{2}$  and  $c_2$  is the squared constant from the Sobolev embedding. Thus, in the case  $m \geq 3$ , we get for  $s = m$  the inequality

$$\begin{aligned} & (A[u](Du, Du) - A[\tilde{u}](D\tilde{u}, D\tilde{u}), R_2)_\Omega \\ & \leq 2c_1 \|R_2\|_{L^\infty(\Omega)} \mathcal{I}[u - \tilde{u}] \\ & \quad + c_1 \left( 2\|D\tilde{u}\|_{L^\infty(\Omega)} \|R_2\|_{L^2(\Omega)} + \sqrt{c_2} \|R_2\|_{L^q(\Omega)} \|D\tilde{u}\|_{L^s(\Omega)}^2 \right) \sqrt{2\mathcal{I}[u - \tilde{u}]}. \quad (33) \end{aligned}$$

For  $m = 2$ , we can choose  $1 \leq q < \infty$  arbitrarily large in (31), which implies (33) holds for  $2 < s$  arbitrary.

We put (30) and (33) together and get that

$$\begin{aligned} & -(\partial_t u - \partial_t \tilde{u}, \partial_t R_2)_\Omega \leq -\frac{d}{dt} (\partial_t u - \partial_t \tilde{u}, R_2)_\Omega + 2c_1 \|R_2\|_{L^\infty(\Omega)} \mathcal{I}[u - \tilde{u}] \\ & \quad + \left( \|\nabla R_2\|_{L^2(\Omega)} + 2c_1 \|D\tilde{u}\|_{L^\infty(\Omega)} \|R_2\|_{L^2(\Omega)} + c_1 \sqrt{c_2} \|R_2\|_{L^q(\Omega)} \|D\tilde{u}\|_{L^s(\Omega)}^2 \right) \\ & \quad \quad \quad \times \sqrt{2\mathcal{I}[u - \tilde{u}]} \\ & \quad + \|R_1 R_2\|_{L^1(\Omega)} - \frac{1}{2} \frac{d}{dt} (R_2, R_2)_\Omega \quad (34) \end{aligned}$$

Coming back to (29), we again use the formula

$B(X, X) - B(Y, Y) = B(X - Y, (X - Y) + 2Y)$  for the terms with the second fundamental form there. Since  $\frac{1}{s} + \frac{1}{q} = \frac{1}{2}$ , we get that

$$\begin{aligned} & (\partial_t u - \partial_t \tilde{u}, (A[u] - A[\tilde{u}])(D\tilde{u}, D\tilde{u}))_\Omega \\ & \quad - (\partial_t \tilde{u}, (A[u] - A[\tilde{u}])(Du, Du) - (A[u] - A[\tilde{u}])(D\tilde{u}, D\tilde{u}))_\Omega \\ & \leq c_1 \|\partial_t u - \partial_t \tilde{u}\|_{L^2(\Omega)} \|u - \tilde{u}\|_{L^q(\Omega)} \|D\tilde{u}\|_{L^{2s}(\Omega)}^2 \\ & \quad + c_1 \|\partial_t \tilde{u}\|_{L^\infty(\Omega)} \|u - \tilde{u}\|_{L^\infty(\Omega)} \|Du - D\tilde{u}\|_{L^2(\Omega)}^2 \\ & \quad + 2c_1 \|(\partial_t \tilde{u} | D\tilde{u})\|_{L^s(\Omega)} \|u - \tilde{u}\|_{L^q(\Omega)} \|Du - D\tilde{u}\|_{L^2(\Omega)}. \end{aligned}$$

We define  $c_3 = c_3(N) = \text{diam}(N)$  and  $c_2$  to be the squared Sobolev embedding constant from  $H^1 \hookrightarrow L^q$ . We also use  $|\partial_t \tilde{u}| |D\tilde{u}| \leq |D\tilde{u}|^2$  and get that

$$\begin{aligned} \frac{d}{dt}(I^2 + I^3) &\leq c_1 \left( c_2 \|D\tilde{u}\|_{L^{2s}(\Omega)}^2 + 2c_3 \|\partial_t \tilde{u}\|_{L^\infty(\Omega)} \right. \\ &\quad \left. + 2c_2 \|D\tilde{u}\|_{L^{2s}(\Omega)}^2 + 2\|R_2\|_{L^\infty(\Omega)} \right) \mathcal{I}[u - \tilde{u}] \\ &\quad - \frac{d}{dt}(\partial_t u - \partial_t \tilde{u}, R_2)_\Omega + \|R_1 R_2\|_{L^1(\Omega)} - \frac{1}{2} \frac{d}{dt}(R_2, R_2)_\Omega \\ &\quad + (\|\nabla R_2\|_{L^2(\Omega)} + 2c_1 \|D\tilde{u}\|_{L^\infty(\Omega)} \|R_2\|_{L^2(\Omega)} \\ &\quad + c_1 \sqrt{c_2} \|R_2\|_{L^q(\Omega)} \|D\tilde{u}\|_{L^s(\Omega)}^2 + \|R_1\|_{L^2(\Omega)}) \sqrt{2\mathcal{I}[u - \tilde{u}]}. \end{aligned} \tag{35}$$

Together with  $\frac{d}{dt} I^1 = (u - \tilde{u}, \partial_t u - \partial_t \tilde{u})_\Omega \leq \mathcal{I}[u - \tilde{u}]$ , (35) implies

$$\begin{aligned} \mathcal{I}[u(t_1, \cdot) - \tilde{u}(t_1, \cdot)] &\leq \mathcal{I}[u - \tilde{u}]|_{t=0} - (\partial_t u - \partial_t \tilde{u}, R_2)_\Omega|_{t_1} \\ &\quad + (\partial_t u(0) - \partial_t \tilde{u}(0), R_2(0, \cdot))_\Omega - \frac{1}{2} \|R_2(t_1, \cdot)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|R_2(0, \cdot)\|_{L^2(\Omega)}^2 \\ &\quad + \int_0^{t_1} \alpha(t) \mathcal{I}[u(t, \cdot) - \tilde{u}(t, \cdot)] dt + \int_0^{t_1} \beta(t) dt + \int_0^{t_1} \gamma(t) \sqrt{2\mathcal{I}[u(t, \cdot) - \tilde{u}(t, \cdot)]} dt, \end{aligned} \tag{36}$$

where

$$\begin{aligned} \alpha(t) &:= 1 + c_1 \left( 3c_2 \|D\tilde{u}\|_{L^{2s}(\Omega)}^2 + 2c_3 \|\partial_t \tilde{u}\|_{L^\infty(\Omega)} + 2\|R_2\|_{L^\infty(\Omega)} \right) \\ \beta(t) &:= \|R_1 R_2\|_{L^1(\Omega)} \\ \gamma(t) &:= \|\nabla R_2\|_{L^2(\Omega)} + 2c_1 \|D\tilde{u}\|_{L^\infty(\Omega)} \|R_2\|_{L^2(\Omega)} + c_1 \sqrt{c_2} \|R_2\|_{L^q(\Omega)} \|D\tilde{u}\|_{L^s(\Omega)}^2 + \|R_1\|_{L^2(\Omega)}. \end{aligned}$$

From inequality (36), we get that

$$\begin{aligned} \frac{1}{2} \mathcal{I}[u(t_1, \cdot) - \tilde{u}(t_1, \cdot)] &\leq 2\mathcal{I}[u - \tilde{u}]|_{t=0} + \|R_2(0, \cdot)\|_{L^2(\Omega)}^2 + \frac{1}{2} \|R_2(t_1, \cdot)\|_{L^2(\Omega)}^2 \\ &\quad + \int_0^{t_1} \alpha(t) \mathcal{I}[u(t, \cdot) - \tilde{u}(t, \cdot)] dt + \int_0^{t_1} \beta(t) dt + \int_0^{t_1} \gamma(t) \sqrt{2\mathcal{I}[u(t, \cdot) - \tilde{u}(t, \cdot)]} dt. \end{aligned} \tag{37}$$

We set

$$\bar{\mathcal{I}}[u - \tilde{u}](t_2) := \sup_{0 \leq t_1 \leq t_2} \mathcal{I}[u(t_1, \cdot) - \tilde{u}(t_1, \cdot)]$$

and obtain, as the right hand side of (37) is monotone increasing in time

$$\begin{aligned} \frac{1}{2} \bar{\mathcal{I}}[u(t_2, \cdot) - \tilde{u}(t_2, \cdot)] &\leq 2\mathcal{I}[u - \tilde{u}]|_{t=0} + \|R_2(0, \cdot)\|_{L^2(\Omega)}^2 + \frac{1}{2} \sup_{0 \leq t_1 \leq t_2} \|R_2(t_1, \cdot)\|_{L^2(\Omega)}^2 \\ &\quad + \int_0^{t_2} \alpha(t) \bar{\mathcal{I}}[u(t, \cdot) - \tilde{u}(t, \cdot)] dt + \int_0^{t_2} \beta(t) dt + \int_0^{t_2} \gamma(t) dt \sqrt{2\bar{\mathcal{I}}[u(t_2, \cdot) - \tilde{u}(t_2, \cdot)]}, \end{aligned}$$

so that

$$\begin{aligned} \frac{1}{4} \tilde{\mathcal{I}}[u(t_2, \cdot) - \tilde{u}(t_2, \cdot)] &\leq 2\mathcal{I}[u - \tilde{u}]|_{t=0} + \|R_2(0, \cdot)\|_{L^2(\Omega)}^2 + \frac{1}{2} \sup_{0 \leq t \leq t_2} \|R_2(t, \cdot)\|_{L^2(\Omega)}^2 \\ &\quad + \int_0^{t_2} \alpha(t) \tilde{\mathcal{I}}[u(t, \cdot) - \tilde{u}(t, \cdot)] dt + \int_0^{t_2} \beta(t) dt + 2 \left( \int_0^{t_2} \gamma(t) dt \right)^2. \end{aligned}$$

Gronwall’s inequality now implies (28).

It remains to explain that the conditions on  $\tilde{u}$  are strong enough for the above computations. We use  $\partial_t \tilde{u}$  as a test function in the weak formulation for  $u$ . This is allowed because the assumptions on  $\tilde{u}$  imply  $\partial_t \tilde{u} \in W^{1,1}(0, T; L^2(\Omega)) \cap L^1(0, T; L^\infty \cap H^1(\Omega))$ .

The conditions on  $\tilde{u}$  imply that all appearing terms are finite. Note particularly that

$$\begin{aligned} \int \|R_2\|_{L^2(\Omega)} \|D\tilde{u}\|_{L^\infty(\Omega)} dt &\leq c \int \|R_2\|_{L^2(\Omega)}^{\frac{k}{k-1}} + \|D\tilde{u}\|_{L^\infty(\Omega)}^k dt < \infty \quad \text{and} \\ \int \|D\tilde{u}\|_{L^{2s}(\Omega)}^2 dt &\leq \int \|D\tilde{u}\|_{L^\infty(\Omega)} \|D\tilde{u}\|_{L^s(\Omega)} dt \\ &\leq \int c \|D\tilde{u}\|_{L^\infty(\Omega)}^k + \|D\tilde{u}\|_{L^s(\Omega)}^{\frac{k}{k-1}} dt < \infty, \\ \int \|R_2\|_{L^q(\Omega)} \|D\tilde{u}\|_{L^s}^2 dt &\leq c \int \|R_2\|_{H^1}^2 + \|D\tilde{u}\|_{L^s}^4 dt, \end{aligned}$$

and the latter is finite because  $k \leq \frac{4}{3}$  implies  $\frac{k}{k-1} \geq 4$ . □

**Remark 5** (Comparison of Theorems 1 and 2) As one can see, the estimate for general  $N$  is slightly different than the one for spheres. The reason is that there is no such formulation as the *angular momentum formulation* (9) for general  $N$ . The formulation (9) has the advantage that it transforms the wave map equation into a system of equations that are only first order in time. Our numerical scheme is based on that formulation. Our reconstruction will have  $\tilde{u}, \tilde{w}$  that are (once) weakly differentiable in time, but due to the presence of  $r_u$ , this does not make  $\tilde{u}$  twice weakly differentiable in time. Indeed, in the setting of Theorem 1, the energy  $\mathcal{I}[\tilde{u}]$  would not be weakly differentiable in time.

Our motivation for splitting the residual in Theorem 2 is that if  $\tilde{u}$  in Theorem 1 happens to be twice weakly differentiable in time, then it satisfies

$$\partial_t^2 \tilde{u} = \Delta \tilde{u} + (|\nabla \tilde{u}|^2 - |\tilde{u} \times \tilde{w}|^2) \tilde{u} + \partial_t r_u + r_g + r_u \times \tilde{w} + \tilde{u} \times r_w,$$

where we have used (13). This equation provides a connection between  $R_1, R_2$  and  $r_u, r_g$  and  $r_\omega$ . We formulated the estimates in Theorem 1 and Theorem 2 as similar as possible—in particular, there is no norm of a time derivative of the residuals involved on the right hand sides of the inequalities—but in detail, the estimates look slightly different. Indeed all terms in (11) have corresponding terms in (28) whereas all terms in (28) except  $\|R_2\|_{L^\infty(0,t;L^2(\Omega))}$  and  $\int_0^t \|D\tilde{u}\|_{L^s(\Omega)}^2 \|R_2\|_{L^q(\Omega)} ds$  have counterparts in (11).



**Remark 6** (Regularity assumptions) We note that our regularity assumptions for general  $N$  are only slightly stronger than the one for spheres in Theorem 1. On the one hand, we need a second time derivative of  $\tilde{u}$ ; see the remark above. On the other hand, we need the additional assumption  $D\tilde{u} \in L^{\frac{k}{k-1}}(0, \mathbb{T}; L^s(\Omega))$  where  $\frac{k}{k-1} \geq 4$  (and not  $k \leq 2$  as in the sphere case). The reason is the appearance of the term  $\int_0^t \|D\tilde{u}\|_{L^s(\Omega)}^2 \|R_2\|_{L^q(\Omega)} ds$  in inequality (28). Due to the different structure of the proof, this term did not appear in Theorem 1.

Note furthermore that by choosing  $k = 1$  and considering the case  $R_1 = R_2 = 0$ , we have the same regularity assumptions as Struwe has in his result [30, Thm 2.2]; see Remark 4.

**Remark 7** (Connection to a posteriori error estimates) In the next chapter, we derive an a posteriori error estimate based on the stability framework of Theorem 1 and not on Theorem 2. Since an *angular momentum formulation* is not available for general target manifolds, basing our analysis on Theorem 1 requires us to restrict our analysis to numerical methods for wave maps with values in  $S^2$ . The reason why we use this, less general, stability analysis is able to handle reconstructions  $\tilde{u}$  that are not twice weakly differentiable in time. In contrast, a posteriori results based on Theorem 2 would require some reconstruction of the numerical solution that is in  $W^{2,1}$  in time. Such a reconstruction was derived for the linear wave equation in [20], but it is not clear how to extend this construction to the wave map case.

### 5 Numerical scheme and a posteriori error estimates

We are concerned with a semi-discretization in time devised for wave maps into spheres in [16]. The scheme is based on the reformulation (9) of the wave-map equation and reads

$$d_t u^{k+1} = u^{k+1/2} \times \omega^{k+1/2}, \quad d_t \omega^{k+1} = \Delta u^{k+1/2} \times u^{k+1/2}, \tag{38}$$

where  $d_t$  is a backward difference quotient in time for a step size  $\tau_k$  and the fractional superscript denotes an average in time, i.e.,

$$d_t u^{k+1} := \frac{1}{\tau_k} (u^{k+1} - u^k), \quad u^{k+1/2} := \frac{u^{k+1} + u^k}{2}.$$

Note that (38) preserves the point-wise constraints  $|u^k(x)| = 1$  and  $u^k \perp \omega^k$  for all  $k$  provided they are satisfied for the initial data. Indeed, this follows by multiplying (38) by  $u^{k+1/2}$ .

A fully discrete version of this scheme, using a finite element discretization in space, was investigated in [15]. There, it was shown that the scheme conserves energy, in the sense that

$$\frac{1}{2} \int_{\Omega} |\omega^k|^2 + |\nabla u^k|^2 dx = \frac{1}{2} \int_{\Omega} |\omega^0|^2 + |\nabla u^0|^2 dx \quad \forall k = 0, \dots, N$$

and that in the limit of vanishing time step size subsequences of the numerical solution converge to weak solutions of the wave map problem. In contrast to the results of Bartels, we are interested in a posteriori error estimates, i.e., computable error bounds that can be evaluated once the numerical solution has been computed.

While weak solutions can be defined beyond times at which a gradient blow-up has occurred, there does not seem to be uniqueness beyond these times, and the results of [17] show that numerical schemes cannot be expected to converge with respect to the energy norm, once the exact solution has no additional regularity. As a consequence, we are only able to provide useful estimates up to the blow-up time. Beyond apparent gradient blow-up in the solution, the error bounds keep converging in  $\tau$  but blow up for  $h \rightarrow 0$ .

Our error analysis follows the approach outlined by Makridakis [37] in that it combines the “energy type” stability result derived in Theorem 1 with a suitable reconstruction of the numerical solution.

The scheme (38) is very close to a Crank-Nicolson scheme, and consequently, we apply a reconstruction that is close to the reconstruction proposed in [38]. A specific feature is that in order to employ the stability result from Theorem 1, we need to project the reconstruction into the target manifold.

## 5.1 Reconstruction

In the sequel, we will define suitable reconstructions of the numerical solutions assuming that a sequence of numerical approximations at different points in time  $0 = t_0 < t_1 < \dots < t_N$  is given:

$$\{u^n\}_{n=0}^N : \Omega \rightarrow S^2, \{\omega^n\}_{n=0}^N : \Omega \rightarrow \mathbb{R}^3.$$

Firstly, we define preliminary, globally continuous, and piecewise linear interpolants using local Lagrange polynomials

$$\ell_n^0(t) := \frac{t_{n+1} - t}{t_{n+1} - t_n}, \quad \ell_n^1(t) := \frac{t - t_n}{t_{n+1} - t_n}$$

by

$$\widehat{u}|_{[t_n, t_{n+1}]}(t) := \ell_n^0(t)u^n + \ell_n^1(t)u^{n+1}, \quad \widehat{\omega}|_{[t_n, t_{n+1}]}(t) := \ell_n^0(t)\omega^n + \ell_n^1(t)\omega^{n+1}.$$

In addition, for any  $g \in C^0([0, t_N], L^2(\Omega, \mathbb{R}^3))$ , we define piecewise constant and piecewise linear interpolants by

$$I_1[g]|_{[t_n, t_{n+1}]}(t) := g(t_n) + \frac{t - t_n}{t_{n+1} - t_n}(g(t_{n+1}) - g(t_n))$$

$$I_0[g]|_{(t_n, t_{n+1})}(t) := \frac{1}{2}(g(t_n) + g(t_{n+1})).$$

This allows us to rewrite the numerical scheme as

$$\begin{aligned} \partial_t \widehat{u}|_{(t_n, t_{n+1})} &= I_0[\widehat{u} \times \widehat{\omega}] - \frac{u^{n+1} - u^n}{2} \times \frac{\omega^{n+1} - \omega^n}{2} =: I_0[\widehat{u} \times \widehat{\omega}] - a_u^n, \\ \partial_t \widehat{\omega}|_{(t_n, t_{n+1})} &= I_0[\Delta \widehat{u} \times \widehat{u}] - \frac{\Delta u^{n+1} - \Delta u^n}{2} \times \frac{u^{n+1} - u^n}{2} =: I_0[\Delta \widehat{u} \times \widehat{u}] - a_\omega^n. \end{aligned} \tag{39}$$

Next, we define piecewise quadratic reconstructions via

$$\begin{aligned} u^*|_{(t_n, t_{n+1})}(t) &:= u^n + \int_{t_n}^t I_1[\widehat{u} \times \widehat{\omega}] - a_u^n ds, \\ \tilde{w}|_{(t_n, t_{n+1})}(t) &:= \omega^n + \int_{t_n}^t I_1[\Delta \widehat{u} \times \widehat{u}] - a_\omega^n ds. \end{aligned} \tag{40}$$

Note that  $u^*, \tilde{w}$  are globally continuous in time since the trapezoidal formula is exact for linear functions and, in particular,  $u^*(t_n) = u^n$  and, thus,  $|u^*(t_n, x)| = |u^n(x)| = 1$  for all  $n$  and all  $x \in \Omega$ . Finally, we define

$$\tilde{u} := \frac{u^*}{|u^*|}, \tag{41}$$

such that  $\tilde{u}$  is a map into  $S^2$  and  $\tilde{u}(t_n) = u^n$  for  $n = 0, \dots, N$ . A straightforward computation gives

$$\begin{aligned} \partial_t \tilde{u} &= \tilde{u} \times \tilde{w} - r_u \\ \text{with } r_u &:= \tilde{u} \times \tilde{w} - I_1(\tilde{u} \times \tilde{w}) + a_u + \partial_t u^* - \frac{\partial_t u^*}{|u^*|} + \frac{\partial_t u^* \cdot u^*}{|u^*|^3} u^*, \\ \partial_t \tilde{w} &= \Delta \tilde{u} \times \tilde{u} - r_\omega \quad \text{with } r_\omega := \Delta \tilde{u} \times \tilde{u} - I_1[\Delta \tilde{u} \times \tilde{u}] + a_\omega, \end{aligned} \tag{42}$$

where  $a_u|_{(t_n, t_{n+1})} := a_u^n$  and  $a_\omega|_{(t_n, t_{n+1})} := a_\omega^n$ .

### 5.2 Computable bounds for residuals

It should be noted that, due to the projection onto the sphere,  $\tilde{u}$  is no longer piecewise quadratic. Thus, it is not straightforward how to compute (norms of) the residuals from (42). In addition, the method at hand is formally second order, so that we should strive for a reconstruction making the a posteriori error estimator second order as well. It might seem obvious that  $\tilde{u} \times \tilde{w} - I_1(\tilde{u} \times \tilde{w}) + a_u$  and  $\Delta \tilde{u} \times \tilde{u} - I_1[\Delta \tilde{u} \times \tilde{u}] + a_\omega$  are second order in time, but this property does not seem obvious for  $\partial_t u^* - \frac{\partial_t u^*}{|u^*|} + \frac{\partial_t u^* \cdot u^*}{|u^*|^3} u^*$ .

Let us begin by decomposing the residuals into several parts:

$$\begin{aligned} r_{u,1} &= \tilde{u} \times \tilde{w} - I_1(\tilde{u} \times \tilde{w}), \\ r_{u,2} &= a_u, \end{aligned}$$

$$\begin{aligned}
 r_{u,3} &= \partial_t u^* - \frac{\partial_t u^*}{|u^*|} + \frac{\partial_t u^* \cdot u^*}{|u^*|^3} u^*, \\
 r_{\omega,1} &= \Delta \tilde{u} \times \tilde{u} - I_1[\Delta \tilde{u} \times \tilde{u}], \\
 r_{\omega,2} &= a_\omega.
 \end{aligned}
 \tag{43}$$

Let us note that  $r_{u,2}, r_{\omega,2}$  are both computable from the numerical solution, without computing any reconstruction, and are both going to converge to zero as  $\tau^2$  as long as the scheme is at least second order convergent.

We will also give easily computable (though lengthy) bounds for the other parts of the residuals, but we postpone their proof to the Appendix. We will state the bounds for a representative time interval  $(t_n, t_{n+1})$  and use the following abbreviations:

$$\begin{aligned}
 A^u &:= |u^{n+1} - u^n|, & A_x^u &:= |\nabla(u^{n+1} - u^n)|, & A_{xx}^u &:= |\Delta(u^{n+1} - u^n)|, \\
 A^\omega &:= |w^{n+1} - w^n|, & A_x^\omega &:= |\nabla(w^{n+1} - w^n)| \\
 B^u &:= |u^{n+1} \times w^{n+1} - u^n \times w^n|, & B_x^u &:= |\nabla(u^{n+1} \times w^{n+1} - u^n \times w^n)|, \\
 B_{xx}^u &:= |\Delta(u^{n+1} \times w^{n+1} - u^n \times w^n)| \\
 B^\omega &:= |\Delta u^{n+1} \times u^{n+1} - \Delta u^n \times u^n|, & B_x^\omega &:= |\nabla(\Delta u^{n+1} \times u^{n+1} - \Delta u^n \times u^n)|, \\
 C^\omega &:= \max\{|\omega^n|, |\omega^{n+1}|\}, & C_x^u &:= \max\{|\nabla u^n|, |\nabla u^{n+1}|\}, \\
 C_x^\omega &:= \max\{|\nabla \omega^n|, |\nabla \omega^{n+1}|\}, & C_{xx}^u &:= \max\{|\Delta u^n|, |\Delta u^{n+1}|\}.
 \end{aligned}$$

Let us mention that we expect  $A^u, A^\omega, A_x^u, A_x^\omega, A_{xx}^u, B^u, B^\omega, B_x^u, B_x^\omega, B_{xx}^u$  to scale like  $\tau_n := t_{n+1} - t_n$  as long as the exact solution is regular enough for the true error to be proportional to  $\tau_n^2$ . In addition, we expect  $C^\omega, C_x^u, C_x^\omega, C_{xx}^u$  to be bounded (uniformly in  $\tau_n$ ) as long as the exact solution is regular. Both expectations are confirmed by our numerical experiments.

**Lemma 2** *Let us denote  $\tau_n := t_{n+1} - t_n$  and let the time-step be chosen sufficiently small such that  $(A^u)^2 + \tau_n B^u < \frac{1}{4}$  holds point-wise. Then, the residuals defined in (42) satisfy the following point-wise estimates.*

$$|r_{u,1}|_{(t_n, t_{n+1})} \leq \tau_n B^\omega + C^\omega (A^u)^2 + C^\omega \tau_n B^u + \frac{1}{4} A^u A^\omega \tag{44}$$

$$\begin{aligned}
 |\nabla r_{u,1}|_{(t_n, t_{n+1})} &\leq (C_x^u + \tau_n B_x^u)(\tau_n B^\omega + C^\omega (A^u)^2) + \tau_n B_x^\omega \\
 &\quad + C^\omega (A_x^u A^u + \tau_n B_x^u + C_x^u \tau_n B^u + \tau_n^2 B^u B_x^u) \\
 &\quad + (A^u)^2 C_x^\omega + \tau_n B_x^u C^\omega + \tau_n B^u C_x^\omega + A_x^u A^\omega + A^u A_x^\omega
 \end{aligned}
 \tag{45}$$

$$|r_{u,2}|_{(t_n, t_{n+1})} \leq \frac{1}{4} A^u A^\omega \tag{46}$$

$$|\nabla r_{u,2}|_{(t_n, t_{n+1})} \leq \frac{1}{4} (A_x^u A^\omega + A^u A_x^\omega) \tag{47}$$

$$\begin{aligned}
 |r_{u,3}|_{(t_n, t_{n+1})} &\leq (C^\omega + \frac{1}{4} A^u A^\omega) (\frac{4}{3} (A^u)^2 + \frac{8}{3} \tau_n B^u) \\
 &\quad + 4 A^u A^\omega (2 + \tau_n B^u) + 4 C^\omega \tau_n B^u
 \end{aligned}
 \tag{48}$$

$$|\nabla r_{u,3}|_{(t_n, t_{n+1})} \leq 8C^\omega [A^u A_x^u + \tau_n B_x^u + C_x^u \tau_n B^u + \tau_n^2 C^u C_x^u] \tag{49}$$

$$+ (C_x^\omega + C_x^u C^\omega) [(A^u)^2 + 3\tau_n B^u] \\ + C_x^u A^u A^\omega + A^u A_x^\omega + \frac{1}{4} (A_x^u A^\omega + A^u A_x^\omega) \\ + \frac{1}{4} A^u A^\omega (C_x^u + \tau_n B_x^u) + (C_x^\omega + C_x^u C^\omega) \tau_n B^u + C^\omega \tau_n B_x^u$$

$$|r_\omega|_{(t_n, t_{n+1})} \leq (C_{xx}^u + \tau_n B_{xx}^u) \left( \frac{7}{3} (A^u)^2 + \frac{11}{3} \tau_n B^u \right) + \frac{9}{4} A_{xx}^u A^u + (A_x^u)^2 \tag{50}$$

$$+ (C_x^u + \tau_n B_x^u) 2[A_x^u A^u + \tau_n B_x^u + (1 + C_x^u) \tau_n B^u + \tau_n^2 B_x^u B^u]$$

$$|r_g| \leq (C^\omega + \tau_n B^\omega) [\tau_n B^\omega + C^\omega ((A^u)^2 + \tau_n B^u) + A^u A^\omega] \tag{51}$$

$$+ [\tau_n B^\omega + C^\omega ((A^u)^2 + \tau_n B^u) + A^u A^\omega]^2.$$

An a posteriori estimate for the difference between the exact solution and the reconstructions  $(\tilde{u}, \tilde{w})$  is easily obtained by combining Lemma 2 and Theorem 1. It should be noted that the definitions of the residuals and the way they enter into the error estimate are rather similar to [18, Theorem 3.1]. However, we cannot avoid exponential-in-time growth of the error due to the non-linearity of the problem, even if the solution does not exhibit gradient blow-up.

### 6 Numerical experiments

In this section, we perform numerical experiments in order to demonstrate the scaling behavior of the error estimator and to investigate time stepping strategies. We focus on the scaling (in  $\tau$ ) of the error and of the error estimator. In order to obtain a fully practical scheme, we used a finite difference discretisation in space (together with the temporal discretisation discussed above), i.e., the scheme from [16]. We have performed numerical simulations on a Cartesian grid for simplicity. This seems reasonable since none of the papers investigating numerical schemes have reported any effects of the choice of spatial grid. On this grid, the scheme from [16] is equivalent to the finite element scheme from [15] since the latter uses mass lumping.

Whenever derivatives occur in the error estimator, they are approximated using finite differences. We consider the problem that was also studied in [8]. These are smooth initial data for which a singularity forms in the numerical experiments. We believe that this reflects formation of a singularity in the exact solution, but, strictly speaking, we cannot be sure that this is what happens since the error estimator stops converging for  $h \rightarrow 0$  after a singularity seems to have formed. The fact that the error estimator does not converge once a singularity (seems to) have formed is related to the fact that our error estimator is an upper bound for the difference between the numerical solution and any finite energy weak solution in the sense of Definition 3.1, and these

solutions are expected not to be unique once a singularity has formed. Note that this is not a contradiction to convergence of numerical solutions to some weak solution. In order to obtain an error estimator that converges after singularity formation, one would need some technique to compare the numerical solutions to only one weak solution. It is currently unclear how to achieve this.

The problem data are given by  $N = S^2$ ,  $\Omega = (-\frac{1}{2}, \frac{1}{2})^2$  and

$$u(0, x) = \begin{cases} \frac{(2a(x)x_1, 2a(x)x_2, a(x)-|x|)^T}{a(x)^2+|x|^2} & \text{for } |x| \leq 1/2 \\ (0, 0, -1)^T & \text{for } |x| \geq 1/2 \end{cases} \quad \partial_t u(0, \cdot) \equiv 0 \quad (52)$$

with  $a(x) := (1 - 2|x|)^4$  and homogeneous Neumann boundary conditions (Fig. 1). The numerical solution develops a singularity at some time  $t > .2$  to be precise. Thus, we expect the scheme to converge with order  $\tau^2$  in the energy norm

$$\max_{t \in [0, \mathbb{T}]} \left( \int_{\Omega} |\partial_t u - \tilde{u} \times \tilde{\omega}|^2 + |\nabla u - \nabla \tilde{u}|^2 dx \right)^{\frac{1}{2}}$$

at least until  $\mathbb{T} = .2$ .

Since the exact solution is unknown in this case, we use a numerical solution on a very fine mesh ( $h = 2^{-9}$ ,  $\tau = 2^{-14}$ ) as reference solution in order to approximate the error. We observe that the error indeed scales as  $\tau^2$ ; see Table 1. There we present errors and experimental order of convergence (eoc)

One of the goals we had in constructing the error estimator, i.e., the upper bound provided by combining Theorem 1 and Lemma 2, was that it is formally of optimal order, i.e., that it converges to zero with the same rate as the true error on equidistant meshes. This is guaranteed as soon as the quantities  $A^u, A_x^u, A_{xx}^u, A^\omega, A_x^\omega, B^u, B_x^u, B_{xx}^u, B^\omega, B_x^\omega$  are  $\mathcal{O}(\tau)$  and  $C_x^u, C_{xx}^u, C^\omega, C_x^\omega$  are uniformly bounded in  $\tau$ . This is indeed what we observe in the experiments we carried out. We do not report these

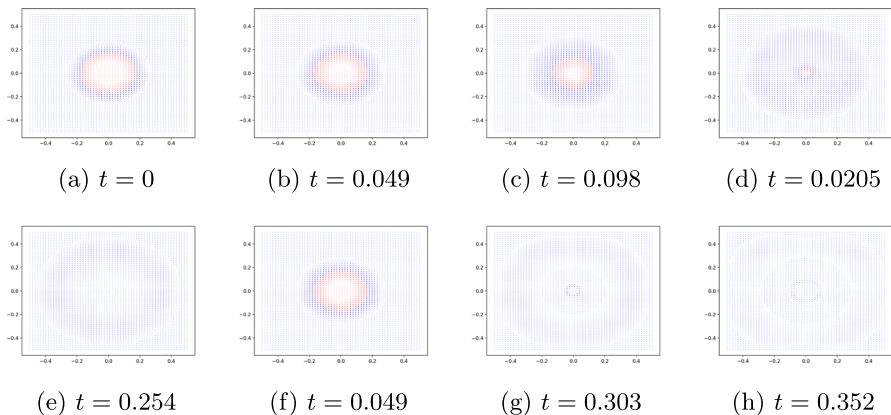


Fig. 1 Snapshots of numerical solution using  $h = 1/60$  and  $\tau = 2^{-10}$  at different times. Arrows depict first two components of  $u_h$ , and color indicates direction of third component

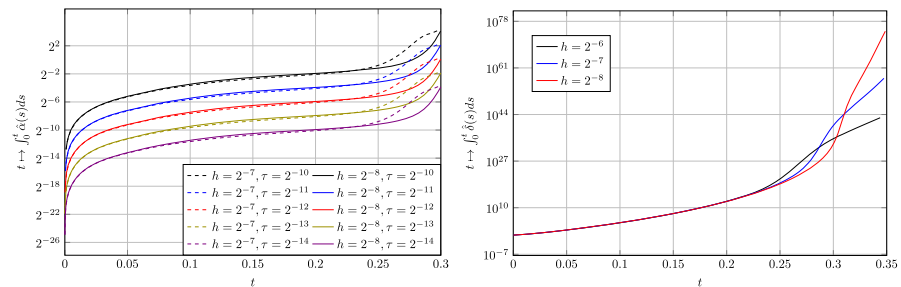
**Table 1** Error in energy norm and eoc up to time  $\mathbb{T} = 0.2$

$\tau$	$\ \omega_h - \omega_{\text{ref}}\ _{L^\infty(0, \mathbb{T}; L^2(\Omega))}$	eoc	$\ \nabla u_h - \nabla u_{\text{ref}}\ _{L^\infty(0, \mathbb{T}; L^2(\Omega))}$	eoc
$2^{-10}$	6.51e-04	—	8.10e-04	—
$2^{-11}$	2.13e-04	1.61	2.36e-04	1.78
$2^{-12}$	6.19e-05	1.78	6.51e-05	1.86
$2^{-13}$	1.28e-05	2.28	1.34e-05	2.28

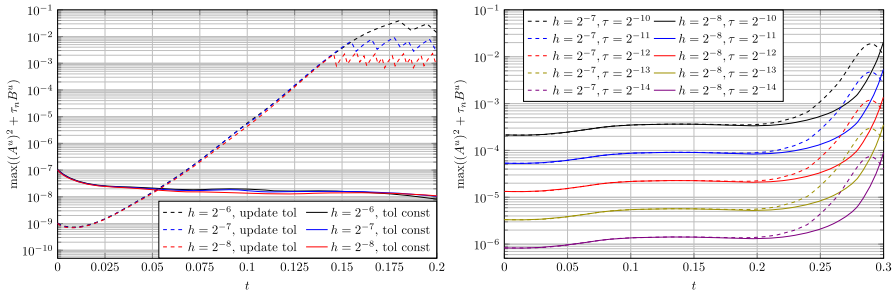
We use a mesh constant  $h = 2^{-9}$  and a reference time step  $\tau = 2^{-14}$

numbers in detail but plot  $\int_0^t \hat{\alpha}(s) ds$  and  $\int_0^t \hat{\delta}(s) ds$  for several equidistant step-sizes  $\tau$  and for different spatial mesh widths  $h$ , in Fig. 2. Here,  $\hat{\alpha}$  and  $\hat{\delta}$  are computable upper bounds for  $\alpha$  and  $\delta$  based on Lemma 2. These results show that  $\alpha$  and  $\delta$  are independent of the spatial mesh width  $h$ , as they are supposed to be since we are quantifying time discretisation errors. We also observe that  $\alpha$  is proportional to  $\tau^2$  and  $\delta$  is (essentially) independent of  $\tau$ . This confirms our theoretical predictions, i.e., optimal convergence order of the error estimator in the smooth regime.

Due to the expected non-uniqueness of finite energy weak solutions, we focus on time-stepping before and up-to singularity formation. Standard algorithms for mesh adaptation aim at equi-distributing  $\frac{1}{\tau_j} \int_{t_j}^{t_j+\tau_j} \hat{\alpha}(s) ds$  among all time steps, i.e., the local step size is reduced if  $\frac{1}{\tau_j} \int_{t_j}^{t_j+\tau_j} \hat{\alpha}(s) ds$  exceeds a given tolerance. As observed in [19], this strategy leads to excessive over-refinement close to blow-ups. Moreover, (12) shows that errors incurred in different time-steps are amplified by different factors. In order to obtain a scheme that is efficient up to blow-up, this needs to be taken into account, and we follow a strategy that is similar to the one proposed in [19], i.e., from the  $j$ -th to the  $(j + 1)$ -th time step, we increase the tolerance by a factor of  $\exp(\frac{1}{2} \int_{t_{j-1}}^{t_j} \hat{\delta}(s) ds)$ .



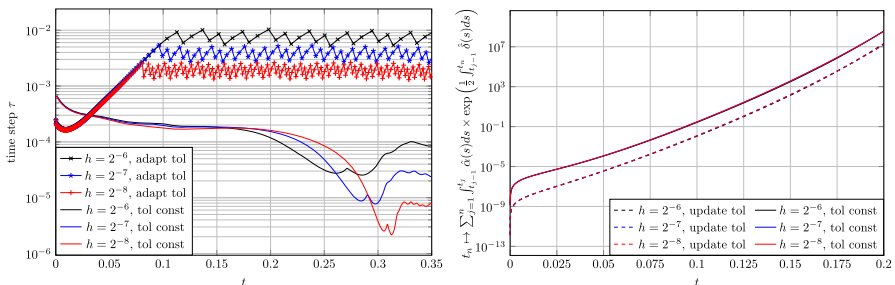
**Fig. 2** Left: temporal evolution of the integral of computable upper bounds for  $\alpha$  based on Lemma 2. Note that, as long as the numerical solutions are “smooth,”  $\alpha$  is independent of the spatial mesh width  $h$  and proportional to  $\tau^2$ . Once numerical solutions show a singularity,  $\alpha$  does depend on  $h$ . Right: temporal evolution of the integral of computable upper bounds for  $\delta$  based on Lemma 2. Note that the bounds for  $\delta$  are independent of  $h$  and  $\tau$  as long as the numerical solutions are “smooth,” whereas  $\delta$  is  $h$  dependent once the numerical solutions display singularities. It should be noted that  $\delta$  contains finite difference approximations of  $\|\nabla u\|_{L^{2p}}$ , and it is anticipated that if the gradient of the exact solution blows up the finite differences of the numerical solution approximating  $\|\nabla u\|_\infty$  scale like  $2/h$



**Fig. 3** Temporal evolution, for different spatial mesh widths, of the quantity that needs to be smaller than  $\frac{1}{4}$  for Lemma 2 to hold. This condition is satisfied for both adaptive time stepping (left) and constant time-step sizes (right)

In Fig. 4, we compare values of the error estimator and time step sizes for the two time-stepping strategies described above. Note that the error estimator hinges on Lemma 2 which is conditional on  $(A^u)^2 + \tau_n B^u$  being small enough. Figure 3 shows that this condition is indeed satisfied in the case under consideration. Moreover, Fig. 3 (right) shows that this quantity is quadratic in the time step size, i.e., we can ensure that the condition holds by making time steps small enough. The condition might be violated close to gradient blow-up of the numerical solution. It turns out that the scheme using the updated error tolerances uses much smaller step sizes in the beginning and larger step sizes at later times than the equidistribution strategy. We observe that at time  $t = 0.17$ , the strategy using updated error tolerances has used fewer time steps but leads to a smaller error bound.

We observe in Fig. 4 that for time stepping with updated tolerances, time step sizes stop increasing beyond a certain point in time. This is due to the fact that for time step sizes larger than some threshold value  $\tau_{\text{thresh}}(h) > 0$ , the fixed point iteration that we use for solving the non-linear system (38) in each time step (see [15, Algo. 2] for



**Fig. 4** Time step sizes (left) and values of the combined error estimator from Theorem 1 and Lemma 2 (right) as functions of time for different (equidistant) spatial meshes. We compare simulations with time step adaptation using a fixed tolerance  $10^{-4}$  to ones using updated tolerances starting at  $10^{-6}$ , so that at time  $t = 0.17$ , the number of computed time steps is comparable (approx. 21000)



details) does not converge. This results in the time step being automatically reduced independently of the error tolerance.

**Funding** Open Access funding enabled and organized by Projekt DEAL. J.G. thanks the German Research Foundation (DFG) for the support of the project via DFG grant GI 1131/1-1. E. M.-B. is funded by the DFG via the grant MA 7559/1-1 and appreciates the support.

## Declarations

**Conflict of interest** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## Appendix

This appendix is devoted to the proof of Lemma 2. Here, each quantity that is defined continuously in time is to be understood as its restriction to  $(t_n, t_{n+1})$ . Let us begin by giving explicit formulae for certain parts of  $r_u$  and  $r_\omega$ . The first such expressions make explicit the difference between some piecewise quadratic, globally continuous function and its piecewise linear interpolations:

$$\begin{aligned}
 u^*(t) - \widehat{u}(t) &= \frac{1}{2} \frac{(t - t_n)(t_{n+1} - t)}{t_{n+1} - t_n} (u^{n+1} \times \omega^{n+1} - u^n \times \omega^n), \\
 \tilde{w}(t) - \widehat{w}(t) &= \frac{1}{2} \frac{(t - t_n)(t_{n+1} - t)}{t_{n+1} - t_n} (\Delta u^{n+1} \times u^{n+1} - \Delta u^n \times u^n), \\
 \widehat{u}(t) \times \widehat{w}(t) - I_1(\widehat{u} \times \widehat{w})(t) &= -\frac{(t - t_n)(t_{n+1} - t)}{(t_{n+1} - t_n)^2} (u^{n+1} - u^n) \times (\omega^{n+1} - \omega^n).
 \end{aligned}
 \tag{53}$$

### Controlling $|\widehat{u}|, |u^*|$

Let us now study how far away  $\widehat{u}, u^*$  are from maps into the sphere: If  $u^{n+1} - u^n$  is sufficiently small, a geometric argument implies

$$\left| 1 - |\widehat{u}| \right| \leq |u^{n+1} - u^n|^2 = (A^u)^2, \quad \left| 1 - |u^*| \right| \leq \left| 1 - |\widehat{u}| \right| + \left| \widehat{u} - u^* \right| \leq (A^u)^2 + \tau_n B^u$$

and, therefore,

$$|\tilde{u} - u^*| = \left| \tilde{u} \left( 1 - |u^*| \right) \right| \leq |u^{n+1} - u^n|^2 = (A^u)^2 + \tau_n B^u. \tag{54}$$

Thus, the conditions in Lemma 2 imply

$$\frac{3}{4} \leq |\widehat{u}| \leq \frac{5}{4}, \quad \frac{1}{2} \leq |u^*| \leq \frac{3}{2}. \tag{55}$$

**Estimating  $r_{u,1}$**

Since  $\tilde{u}(t_n) = \widehat{u}(t_n)$  and  $\tilde{w}(t_n) = \widehat{w}(t_n)$  we may rewrite  $r_{u,1}$  as

$$r_{u,1} = \tilde{u} \times (\tilde{w} - \widehat{w}) + (\tilde{u} - u^*) \times \widehat{w} + (u^* - \widehat{u}) \times \widehat{w} + \widehat{u} \times \widehat{w} - I_1(\widehat{u} \times \widehat{w}) \tag{56}$$

such that

$$\begin{aligned} |r_{u,1}| &\leq \tau_n (\Delta \tilde{u}^{n+1} \times \tilde{u}^{n+1} - \Delta \tilde{u}^n \times \tilde{u}^n) 2|u^{n+1} - u^n| |\omega^{n+1} - \omega^n| \\ &\quad + \max\{|\omega^n|, |\omega^{n+1}|\} \left[ |u^{n+1} - u^n|^2 + |t_{n+1} - t_n| |u^{n+1} \times \omega^{n+1} - u^n \times \omega^n| \right]. \end{aligned} \tag{57}$$

This proves (44).

We obtain (45) by applying the product rule to (56) since  $|\widehat{u}(t, x)| \leq 1$ . We also use the fact that  $\nabla \tilde{u}$  is the projection of  $\nabla u^*$  onto the tangent space of the sphere, so that, provided  $|u^{n+1} - u^n| < 1/2$  holds, we have the point-wise estimate

$$|\nabla \tilde{u}| \leq 2|\nabla u^*|. \tag{58}$$

**Estimating  $r_{u,3}$**

The key to estimating  $r_{u,3}$  is to control  $\partial_t u^* \cdot u^*$ . We notice that

$$\partial_t u^* \cdot u^* = (I_1[\widehat{u} \times \widehat{w}] + a_u) \cdot (\widehat{u} + (u^* - \widehat{u})) \tag{59}$$

and, thus,

$$|\partial_t u^* \cdot u^*| \leq |I_1[\widehat{u} \times \widehat{w}] \cdot \widehat{u}| + A^u A^\omega (1 + \tau_n B^u) + C^\omega \tau_n B^u \tag{60}$$

so that it remains to understand  $I_1[\widehat{u} \times \widehat{w}] \cdot \widehat{u}$ . Using orthogonality, we obtain

$$\begin{aligned} I_1[\widehat{u} \times \widehat{w}] \cdot \widehat{u} &= [\ell_n^0(t)u^n \times \omega^n + \ell_n^1(t)u^{n+1} \times \omega^{n+1}] \cdot [\ell_n^0(t)u^n + \ell_n^1(t)u^{n+1}] \\ &= \ell_n^0(t)\ell_n^1(t)[u^{n+1} \cdot (u^n \times \omega^n) + u^n \cdot (u^{n+1} \times \omega^{n+1})] \\ &= \ell_n^0(t)\ell_n^1(t) \det(u^{n+1}, u^n, \omega^{n+1} - \omega^n) \\ &= \ell_n^0(t)\ell_n^1(t) \det(u^{n+1} - u^n, u^n, \omega^{n+1} - \omega^n). \end{aligned} \tag{61}$$

We insert (61) into (59) and obtain

$$|\partial_t u^* \cdot u^*| \leq A^u A^\omega (2 + \tau_n B^u) + C^\omega \tau_n B^u. \tag{62}$$

Moreover, due to (55) we arrive at

$$\left| 1 - \frac{1}{|u^*|} \right| = \left| 1 - \frac{1}{|\widehat{u}|} - \frac{|u^*| - |\widehat{u}|}{|u^*||\widehat{u}|} \right| \leq \frac{4}{3} (A^u)^2 + \frac{8}{3} \tau_n B^u \tag{63}$$

Using (63) and (62) we obtain:

$$|r_{u,3}| \leq (C^\omega + \frac{1}{4} A^u A^\omega) \left( \frac{4}{3} (A^u)^2 + \frac{8}{3} \tau_n B^u \right) + 4A^u A^\omega (2 + \tau_n B^u) + 4C^\omega \tau_n B^u. \tag{64}$$

Let us note that for

$$\begin{aligned} \nabla r_{u,3} &= \nabla \partial_t u^* \left( 1 - \frac{1}{|u^*|} \right) + \nabla (I_1[\widehat{u} \times \widehat{\omega}] \cdot \widehat{u}) + \nabla a_u \cdot u^* \\ &\quad + a_u \cdot \nabla u^* + \nabla (I_1[\widehat{u} \times \widehat{\omega}] \cdot (u^* - \widehat{u})) + \partial_t u^* \nabla \left( 1 - \frac{1}{|u^*|} \right) \end{aligned} \tag{65}$$

we have suitable bounds for all terms on the right hand side of (65) except for  $\nabla \left( 1 - \frac{1}{|u^*|} \right)$ , e.g.,  $\nabla (I_1[\widehat{u} \times \widehat{\omega}] \cdot \widehat{u})$  can be estimated by applying the product rule to (61). As a first step towards estimating for  $\nabla \left( 1 - \frac{1}{|u^*|} \right) = -\frac{u^* \cdot \nabla u^*}{|u^*|^3}$ , we compute

$$\begin{aligned} |u^*|^3 \partial_{x_j} \left( 1 - \frac{1}{|u^*|} \right) &= \partial_{x_j} \widehat{u} \cdot \widehat{u} + \partial_{x_j} (u^* - \widehat{u}) \cdot \widehat{u} \\ &\quad + \partial_{x_j} \widehat{u} \cdot (u^* - \widehat{u}) + \partial_{x_j} (u^* - \widehat{u}) \cdot (u^* - \widehat{u}). \end{aligned} \tag{66}$$

We recall  $|u^n| = 1$ , which implies  $\partial_{x_j} u^n \cdot u^n = 0$ , for all  $n$ , so that

$$\begin{aligned} \partial_{x_j} \widehat{u} \cdot \widehat{u} &= \partial_{x_j} (\ell_n^0(t) u^n + \ell_n^1(t) u^{n+1}) \cdot (\ell_n^0(t) u^n + \ell_n^1(t) u^{n+1}) \\ &= \ell_n^0(t) \ell_n^1(t) (\partial_{x_j} u^n \cdot u^{n+1} + \partial_{x_j} u^{n+1} \cdot u^n) \\ &= -\ell_n^0(t) \ell_n^1(t) \partial_{x_j} (u^{n+1} - u^n) \cdot (u^{n+1} - u^n). \end{aligned} \tag{67}$$

We insert (67) into (66) and obtain

$$\left| \nabla \left( 1 - \frac{1}{|u^*|} \right) \right| \leq 8[A_x^u A^u + \tau_n B_x^u + C_x^u \tau_n B^u + \tau_n^2 B_x^u B^u] \tag{68}$$

Thus, we obtain

$$\begin{aligned}
 |\nabla r_{u,3}| \leq & \left( C_x^u C^\omega + C_x^\omega + \frac{1}{4} A_x^u A^\omega + \frac{1}{4} A^u A_x^\omega \right) \left( \frac{4}{3} (A^u)^2 + \frac{8}{3} \tau_n B^u \right) \\
 & + \frac{3}{2} A_x^u A^\omega + 2 A^u C_x^u A^\omega + \frac{3}{2} A^u A_x^\omega + (C_x^u C^\omega + C_x^\omega) \tau_n B^u + C^\omega \tau_n B_x^u \\
 & + 8 \left( C^\omega + \frac{1}{4} A^u A^\omega \right) \left( A_x^u A^u + \tau_n B_x^u + C_x^u \tau_n B^u + \tau_n^2 B_x^u B^u \right).
 \end{aligned} \tag{69}$$

This completes providing bounds for the different components of  $r_u$  and  $\nabla r_u$ .

**Estimating  $r_\omega$**

Obviously,  $|r_{\omega,2}| = |a^\omega| \leq \frac{1}{4} A^u A_{xx}^u$  and

$$\begin{aligned}
 r_{\omega,1} = & (\Delta \tilde{u} - \Delta u^*) \times \tilde{u} + (\Delta u^* - \Delta \widehat{u}) \times \tilde{u} + \Delta \widehat{u} \times (\tilde{u} - u^*) \\
 & + \Delta \widehat{u} \times (u^* - \widehat{u}) + \Delta \widehat{u} \times \widehat{u} - I_1[\Delta \widehat{u} \times \widehat{u}].
 \end{aligned} \tag{70}$$

We insert (54) and (53) into (70) and obtain

$$|r_{\omega,1}| \leq |(\Delta \tilde{u} - \Delta u^*)| + \tau_n B^u + C_{xx}^u (A^u)^2 + C_{xx}^u \tau_n B^u + A_{xx}^u A^u, \tag{71}$$

where we have used that

$$\Delta \widehat{u} \times \widehat{u} - I_1[\Delta \widehat{u} \times \widehat{u}] = -\ell_n^0(t) \ell_n^1(t) (\Delta u^{n+1} - \Delta u^n) \times (u^{n+1} - u^n).$$

It remains to provide an estimate for  $|\Delta \tilde{u} - \Delta u^*|$ . We note that  $u^* = \tilde{u}|u^*|$  and, thus,

$$\Delta(\tilde{u} - u^*) = \Delta u^* \left( 1 - \frac{1}{|u^*|} \right) - \frac{\sum_j \partial_{x_j} \tilde{u} \cdot \partial_{x_j} |u^*|}{|u^*|} - \frac{\tilde{u} \Delta |u^*|}{|u^*|}, \tag{72}$$

so that

$$\begin{aligned}
 |\Delta(\tilde{u} - u^*)| \leq & (C_{xx}^u + \tau_n B_{xx}^u) \left( \frac{4}{3} (A^u)^2 + \frac{8}{3} \tau_n B^u \right) + |\nabla \tilde{u}| \frac{|u^* \cdot \nabla u^*|}{|u^*|^2} + \frac{|\Delta |u^*||}{|u^*|} \\
 \leq & (C_{xx}^u + \tau_n B_{xx}^u) \left( \frac{4}{3} (A^u)^2 + \frac{8}{3} \tau_n B^u \right) \\
 & + (C_x^u + \tau_n B_x^u) 2[A_x^u A^u + \tau_n B_x^u + C_x^u \tau_n B^u + \tau_n^2 B_x^u B^u] + \frac{|\Delta |u^*||}{|u^*|}
 \end{aligned} \tag{73}$$

where we have used (68). Orthogonality  $\partial_{x_j} u^n \perp u^n$  for all  $j$  and all  $n$  implies

$$\Delta |u^*| = -\ell_n^0(t) \ell_n^1(t) [(u^{n+1} - u^n) \cdot \partial_{x_j}^2 (u^{n+1} - u^n) + |\partial_{x_j} (u^{n+1} - u^n)|^2]. \tag{74}$$

Inserting (74) into (73) implies

$$\begin{aligned}
 |\Delta(\tilde{u} - u^*)| \leq & (C_{xx}^u + \tau_n B_{xx}^u) \left( \frac{4}{3} (A^u)^2 + \frac{8}{3} \tau_n B^u \right) + A_{xx}^u A^u + (A_x^u)^2 \\
 & + (C_x^u + \tau_n B_x^u) 2[A_x^u A^u + \tau_n B_x^u + C_x^u \tau_n B^u + \tau_n^2 B_x^u B^u]
 \end{aligned} \tag{75}$$

Inserting (75) into (71) completes the bound for  $r_\omega$ .

### Estimating $r_g$

We note that orthogonality  $u^n \perp \omega^n$  implies

$$\begin{aligned}
 \tilde{u} \cdot \tilde{w} &= \tilde{u} \cdot (\tilde{w} - \widehat{w}) - (\tilde{u} - \widehat{u}) \cdot \widehat{w} + \widehat{u} \cdot \widehat{w} \\
 &= \tilde{u} \cdot (\tilde{w} - \widehat{w}) - (\tilde{u} - \widehat{u}) \cdot \widehat{w} - \ell_n^0(t) \ell_n^1(t) (u^{n+1} - u^n) \cdot (\omega^{n+1} - \omega^n)
 \end{aligned} \tag{76}$$

so that

$$|\tilde{u} \cdot \tilde{w}| \leq \tau_n B^\omega + C^\omega ((A^u)^2 + \tau_n B^u) + A^u A^\omega. \tag{77}$$

Thus, using the definition of  $r_g$

$$\begin{aligned}
 |r_g| \leq & (C^\omega + \tau_n B^\omega) [\tau_n B^\omega + C^\omega ((A^u)^2 + \tau_n B^u) + A^u A^\omega] \\
 & + [\tau_n B^\omega + C^\omega ((A^u)^2 + \tau_n B^u) + A^u A^\omega]^2.
 \end{aligned}$$

### References

1. Aitchison, I.J.R., Hey, A.J.G.: Gauge theories in particle physics: Volume I: From Relativistic Quantum Mechanics to QED, 3rd edn. Graduate Student Series in Physics
2. Carroll, S.M.: Spacetime and geometry: an introduction to general relativity. Cambridge University Press, Cambridge (2019) <https://doi.org/10.1017/9781108770385>
3. Shatah, J., Struwe, M.: Geometric wave equations, vol. 2, p. 153. New York University, Courant Institute of Mathematical Sciences, New York, NY (1998)
4. Tataru, D.: The wave maps equation. Bull. Amer. Math. Soc. (N.S.) **41**(2), 185–204 (2004)
5. Krieger, J.: Global regularity and singularity development for wave maps. In: Surveys in differential geometry. Vol. XII. Geometric flows. Survey Differential Geometry, vol. 12, pp. 167–201. International Press
6. Rodnianski, I., Sterbenz, J.: On the formation of singularities in the critical  $O(3)$   $\sigma$ -model. Ann. Math. (2) **172**(1), 187–242 (2010)
7. Geba, D.-A., Grillakis, M.G.: An introduction to the theory of wave maps and related geometric problems, p. 477. World Scientific, Hackensack, NJ (2017)
8. Bartels, S.: Semi-implicit approximation of wave maps into smooth or convex surfaces. SIAM J. Numer. Anal. **47**(5), 3486–3506 (2009)
9. Bartels, S.: Projection-free approximation of geometrically constrained partial differential equations. Math. Comp. **85**(299), 1033–1049 (2016)
10. Bartels, S., Feng, X., Prohl, A.: Finite element approximations of wave maps into spheres. SIAM J. Numer. Anal. **46**(1), 61–87 (2007)

11. Bartels, S., Lubich, C., Prohl, A.: Convergent discretization of heat and wave map flows to spheres using approximate discrete Lagrange multipliers. *Math. Comp.* **78**(267), 1269–1292 (2009)
12. Bizoń, P., Chmaj, T., Tabor, Z.A.: Formation of singularities for equivariant (2+1)-dimensional wave maps into the 2-sphere. *Nonlinearity* **14**(5), 1041–1053 (2001)
13. Cohen, D., Verdier, O.: Multisymplectic discretization of wave map equations. *SIAM J. Sci. Comput.* **38**(2), 953–972 (2016)
14. Kycia, R.A.: On similarity in the evolution of semilinear wave and Klein-Gordon equations: numerical surveys. *J. Math. Phys.* **53**(2), 023703–18 (2012)
15. Bartels, S.: Fast and accurate finite element approximation of wave maps into spheres. *ESAIM Math. Model. Numer. Anal.* **49**(2), 551–558 (2015)
16. Karper, T.K., Weber, F.: A new angular momentum method for computing wave maps into spheres. *SIAM J. Numer. Anal.* **52**(4), 2073–2091 (2014)
17. D’Ancona, P., Georgiev, V.: On the continuity of the solution operator to the wave map system. *Comm. Pure Appl. Math.* **57**(3), 357–383 (2004)
18. Georgoulis, E.H., Lakkis, O., Makridakis, C.G., Virtanen, J.M.: A posteriori error estimates for leapfrog and cosine methods for second order evolution problems. *SIAM J. Numer. Anal.* **54**(1), 120–136 (2016)
19. Kyza, I., Makridakis, C.: Analysis for time discrete approximations of blow-up solutions of semilinear parabolic equations. *SIAM J. Numer. Anal.* **49**(1), 405–426 (2011)
20. Georgoulis, E.H., Lakkis, O., Makridakis, C.: A posteriori  $L^\infty(L^2)$ -error bounds for finite element approximations to the wave equation. *IMA J. Numer. Anal.* **33**(4), 1245–1264 (2013)
21. Grohs, P., Hardering, H., Sander, O.: Optimal a priori discretization error bounds for geodesic finite elements. *Found. Comput. Math.* **15**(6), 1357–1411 (2015)
22. Bartels, S.: Stability and convergence of finite-element approximation schemes for harmonic maps. *SIAM J. Numer. Anal.* **43**(1), 220–238 (2005)
23. Dedner, A., Giesselmann, J., Pryer, T., Ryan, J.K.: Residual estimates for post-processors in elliptic problems. *J. Sci. Comput.* **88**(2), 28 (2021)
24. Bernardi, C., Süli, E.: Time and space adaptivity for the second-order wave equation. *Math. Models Methods Appl. Sci.* **15**(2), 199–225 (2005)
25. Adjerid, S.: A posteriori finite element error estimation for second-order hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.* **191**(41–42), 4699–4719 (2002)
26. Bangerth, W., Rannacher, R.: Adaptive finite element techniques for the acoustic wave equation. *J. Comput. Acoust.* **9**(2), 575–591 (2001)
27. Süli, E.: A posteriori error analysis and global error control for adaptive finite volume approximations of hyperbolic problems. In: *Numerical Analysis 1995* (Dundee, 1995). Pitman Res. Notes Math. Ser., **344**, 169–190. Longman, Harlow (1996)
28. Cangiani, A., Georgoulis, E.H., Kyza, I., Metcalfe, S.: Adaptivity and blow-up detection for nonlinear evolution problems. *SIAM J. Sci. Comput.* **38**(6), 3833–3856 (2016)
29. Müller, S., Struwe, M.: Global existence of wave maps in 1+2 dimensions with finite energy data. *Topol. Methods Nonlinear Anal.* **7**(2), 245–259 (1996)
30. Struwe, M.: Uniqueness for critical nonlinear wave equations and wave maps via the energy inequality. *Comm. Pure Appl. Math.* **52**(9), 1179–1188 (1999)
31. Widmayer, K.: Non-uniqueness of weak solutions to the wave map problem. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **32**(3), 519–532 (2015)
32. Zhou, Y.: Uniqueness of weak solutions of 1 + 1 dimensional wave maps. *Math. Z.* **232**(4), 707–719 (1999)
33. Shatah, J., Tahvildar-Zadeh, A.S.: On the Cauchy problem for equivariant wave maps. *Comm. Pure Appl. Math.* **47**(5), 719–754 (1994)
34. Freire, A., Müller, S., Struwe, M.: Weak compactness of wave maps and harmonic maps. *Ann. Inst. Henri Poincaré, Anal. Non Linéaire* **15**(6), 725–754 (1998)
35. Mizuguchi, M., Tanaka, K., Sekine, K., Oishi, S.: Estimation of Sobolev embedding constant on a domain dividable into bounded convex domains. *J. Inequal. Appl.* **2017**, 18 (2017). <https://doi.org/10.1186/s13660-017-1571-0>, Id/No 299
36. Dragomir, S.S.: Some Gronwall type inequalities and applications, p. 193. Nova Science Publishers, Hauppauge, NY (2003)

37. Makridakis, C.: Space and time reconstructions in a posteriori analysis of evolution problems. In: ESAIM Proceedings. [Journées D'Analyse Fonctionnelle et Numérique en L'honneur De Michel Crouzeix]. ESAIM Proceeding vol. 21, pp. 31–44. EDP Science Les Ulis (2007)
38. Akrivis, G., Makridakis, C., Nochetto, R.H.: A posteriori error estimates for the Crank-Nicolson method for parabolic equations. *Math. Comp.* **75**(254), 511–531 (2006)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.