# NUMERICAL ANALYSIS OF THE LANDAU–LIFSHITZ–GILBERT EQUATION WITH INERTIAL EFFECTS

MICHELE RUGGERI*

**Abstract.** We consider the numerical approximation of the inertial Landau–Lifshitz–Gilbert equation (iLLG), which describes the dynamics of the magnetisation in ferromagnetic materials at subpicosecond time scales. We propose and analyse two fully discrete numerical schemes: The first method is based on a reformulation of the problem as a linear constrained variational formulation for the linear velocity. The second method exploits a reformulation of the problem as a first order system in time for the magnetisation and the angular momentum. Both schemes are implicit, based on first-order finite elements, and generate approximations satisfying the unit-length constraint of iLLG at the vertices of the underlying mesh. For both methods, we prove convergence of the approximations towards a weak solution of the problem. Numerical experiments validate the theoretical results and show the applicability of the methods for the simulation of ultrafast magnetic processes.

## 1. INTRODUCTION

### 1.1. Magnetisation dynamics with inertial effects

The understanding of the magnetisation dynamics and the capability to perform reliable numerical simulations of magnetic systems play a fundamental role in the design of many technological applications, *e.g.* hard disk drives. A well-accepted model to describe the magnetisation dynamics in ferromagnetic materials is the Landau–Lifshitz–Gilbert equation (LLG), which, in the so-called Gilbert form, is given by

$$\partial_t \boldsymbol{m} = -\gamma_0 \, \boldsymbol{m} \times \boldsymbol{H}_{\mathrm{eff}}[\boldsymbol{m}] + \alpha \, \boldsymbol{m} \times \partial_t \boldsymbol{m}. \tag{1.1}$$

Here, $\boldsymbol{m}$ denotes the normalised magnetisation (dimensionless and satisfying $|\boldsymbol{m}| = 1$), the effective field $\boldsymbol{H}_{\mathrm{eff}}[\boldsymbol{m}]$ (in $\mathrm{A\,m^{-1}}$), up to a negative multiplicative constant, is the functional derivative of the micromagnetic energy $\mathcal{E}[\boldsymbol{m}]$ (in J) with respect to the magnetisation, while $\gamma_0 > 0$ and $\alpha > 0$ denote the gyromagnetic ratio (in $\mathrm{m\,A^{-1}\,s^{-1}}$) and the Gilbert damping parameter (dimensionless), respectively. The first term on the right-hand side of (1.1) describes the precession of the magnetisation around the effective field. The second term is dissipative and pushes the magnetisation towards the effective field. The resulting dynamics is a damped precession, where the magnetisation rotates around the effective field while being damped towards it; see Figure 1a.
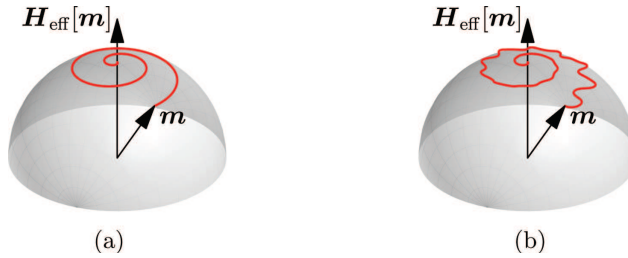
FIGURE 1. Schematic of the magnetisation dynamics: (a) LLG (precession and damping); (b) iLLG (precession, damping, and nutation).

In 1996, a pioneering experiment showed that, using femtosecond laser excitations, it is possible to manipulate the magnetisation of a nickel sample at subpicosecond time scales [16]. This discovery gave impulse to several theoretical and experimental studies, which gave rise to the field that nowadays is referred to as *ultrafast magnetism* [36].

The standard LLG (1.1) is not capable to describe the dynamics of the magnetisation at such short time scales. Based on the concept of angular momentum in magnetic spin systems, a novel evolution equation has been recently proposed [19]. This equation, called *inertial* LLG (iLLG), reads as

$$\partial_t \boldsymbol{m} = -\gamma_0\, \boldsymbol{m} \times \boldsymbol{H}_{\text{eff}}[\boldsymbol{m}] + \alpha\, \boldsymbol{m} \times \partial_t \boldsymbol{m} + \tau\, \boldsymbol{m} \times \partial_{tt}\boldsymbol{m}. \tag{1.2}$$

Here, $\tau > 0$ denotes the angular momentum relaxation time (in s). In addition to the classical precession and damping contributions, the right-hand side of (1.2) comprises a third term involving the second time derivative of the magnetisation. It has been predicted that the effect of this additional contribution on the magnetisation dynamics consists in the appearance of nutation dynamics – superimposed magnetisation oscillations occurring at a frequency much higher than the one of the damped precession dynamics; see Figure 1b. Such inertial dynamics has been experimentally observed for the first time only very recently [32].

## 1.2. Numerical approximation of LLG and wave map equation

This work is concerned with the numerical analysis of (1.2). As LLG has some similarities with the harmonic map heat flow into the sphere [30]

$$\partial_t \boldsymbol{u} - \boldsymbol{\Delta} \boldsymbol{u} = |\boldsymbol{\nabla} \boldsymbol{u}|^2 \boldsymbol{u},$$

it turns out that iLLG is related to the wave map equation into the sphere [35]

$$\partial_{tt} \boldsymbol{u} - \boldsymbol{\Delta} \boldsymbol{u} = \left( |\boldsymbol{\nabla} \boldsymbol{u}|^2 - |\partial_t \boldsymbol{u}|^2 \right) \boldsymbol{u}. \tag{1.3}$$

The numerical approximation of this class of partial differential equations (PDEs) poses several challenges: nonuniqueness, possible blow-up in time, and low regularity of weak solutions, geometric nonlinearities, a nonconvex pointwise unit-length constraint, intrinsic energy laws as well as (for LLG) the possible coupling with other PDEs, *e.g.* the Maxwell equations.

In the last twenty years, several numerical integrators have been proposed. Without claiming to be exhaustive and, in particular, restricting ourselves to the methods that are akin to the ones proposed in the present work, we refer to the works [2, 3, 12, 14, 28] for LLG and to [8, 9, 11, 13, 15, 27] for the wave map equation.

## 1.3. Contributions and outline of the present work

In this work, combining techniques developed for LLG and the wave map equation, we introduce, analyse, and numerically compare two fully discrete numerical schemes for iLLG. For both methods, the spatial discretisation is based on first-order finite elements. The first scheme (Algorithm 3.1) is an extension of the tangent

plane scheme proposed for (1.1) in [2]. The scheme is based on an equivalent reformulation of (1.2) in the tangent space. The unit-length constraint is enforced at the vertices of the mesh by projecting to the sphere the nodal values of the computed approximation at each time-step. The second method (Algorithm 3.2) extends to iLLG the constraint-preserving angular momentum method proposed for (1.3) in [27]. Following [9], the spatial discretisation based on finite differences considered in [27] is replaced by a mass-lumped finite element approximation. Since the resulting method leads to the solution of a nonlinear system of equations per time-step, a linearisation based on a convergent constraint-preserving fixed-point iteration – similar to those considered in [9,12] for the methods proposed therein – is discussed and analysed (Algorithm 3.4).

We study well-posedness and stability of the proposed schemes, and determine sufficient conditions which guarantee that the algorithms satisfy discrete energy laws resembling the one of the continuous problem (see (2.7) below). Moreover, we prove that they generate sequences of finite element solutions that, upon extraction of a subsequence, converge towards a weak solution of the problem. The proof is constructive and provides an alternative proof of existence of weak solutions to (1.2) (first established in [25]). The numerical analysis of iLLG has been considered so far only in [31], where a semi-implicit method has been proposed and its conservation properties have been analysed. The present work thus proposes the first numerical schemes that are proven to be convergent towards a weak solution of iLLG. We note that our analysis, in the spirit of the aforementioned works [2,3,8,9,11–15,27,28] (concerned with the wave map equation and/or LLG), establishes plain convergence (without rates) of the finite element approximations under minimal regularity assumptions on the weak solution. We refer, *e.g.* to the papers [6,23] (both concerned with LLG) for a different type of analysis showing convergence (with rates) of the finite element approximations towards a sufficiently regular strong solution to the problem.

The remainder of the work is organised as follows: We conclude this section by collecting some useful notation used throughout the paper. In Section 2, we present the mathematical model under consideration in detail. In Section 3, we introduce the proposed numerical schemes and state the main results of the work. Section 4 is devoted to numerical experiments. Finally, in Section 5, we collect the proofs of the results presented in the paper.

## 1.4. Notation

We denote by $\mathbb{N} = \{1, 2, \ldots\}$ the set of natural numbers and set $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. We denote the unit sphere by $\mathbb{S}^2 = \{x \in \mathbb{R}^3 : |x| = 1\}$. We denote by $\{e_1, e_2, e_3\}$ the standard basis of $\mathbb{R}^3$. For (spaces of) vector-valued or matrix-valued functions, we use bold letters, *e.g.* for a generic domain $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$, we denote both $L^2(\Omega; \mathbb{R}^3)$ and $L^2(\Omega; \mathbb{R}^{3 \times 3})$ by $\boldsymbol{L}^2(\Omega)$. We denote by $\langle \cdot, \cdot \rangle$ both the scalar product of $\boldsymbol{L}^2(\Omega)$ and the duality pairing between $\boldsymbol{H}^1(\Omega)$ and its dual, with the ambiguity being resolved by the arguments. The set of sphere-valued functions in $\boldsymbol{H}^1(\Omega)$ is denoted by $H^1(\Omega; \mathbb{S}^2)$. We also use the notation $\lesssim$ to denote *smaller than or equal to up to a multiplicative constant*, *i.e.* we write $A \lesssim B$ if there exists a constant $c > 0$, which is clear from the context and always independent of the discretisation parameters, such that $A \leq cB$. Finally, we write $A \simeq B$ if $A \lesssim B$ and $B \lesssim A$ hold simultaneously.

## 2. Mathematical model

Let $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ be a bounded Lipschitz domain. The energy of $\boldsymbol{m} \in H^1(\Omega; \mathbb{S}^2)$ is described by the Dirichlet energy functional

$$\mathcal{E}[\boldsymbol{m}] = \frac{1}{2} \|\boldsymbol{\nabla} \boldsymbol{m}\|_{\boldsymbol{L}^2(\Omega)}^2. \tag{2.1}$$

Minimisers $\boldsymbol{m} \in H^1(\Omega; \mathbb{S}^2)$ of (2.1) satisfy the Euler–Lagrange equations

$$\langle \boldsymbol{h}_{\text{eff}}[\boldsymbol{m}], \boldsymbol{\phi} \rangle = 0 \quad \text{for all } \boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega) \text{ such that } \boldsymbol{m} \cdot \boldsymbol{\phi} = 0 \text{ a.e. in } \Omega,$$

which, in strong form, take the form

$$\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}[\boldsymbol{m}] = \boldsymbol{0} \quad \text{in } \Omega, \tag{2.2a}$$

$$\partial_{\boldsymbol{\nu}} \boldsymbol{m} = \boldsymbol{0} \quad \text{on } \partial\Omega, \tag{2.2b}$$

where $\boldsymbol{\nu} : \partial\Omega \to \mathbb{S}^2$ denote the outward-pointing unit normal vector to $\partial\Omega$. Here, $\boldsymbol{h}_{\text{eff}}[\boldsymbol{m}]$ is defined as the opposite of the Gâteaux derivative of the energy, *i.e.*

$$-\langle \boldsymbol{h}_{\text{eff}}[\boldsymbol{m}], \boldsymbol{\phi} \rangle = \left\langle \frac{\delta \mathcal{E}[\boldsymbol{m}]}{\delta \boldsymbol{m}}, \boldsymbol{\phi} \right\rangle \overset{(2.1)}{=} \langle \boldsymbol{\nabla} \boldsymbol{m}, \boldsymbol{\nabla} \boldsymbol{\phi} \rangle. \tag{2.3}$$

Nonequilibrium magnetisation configurations $\boldsymbol{m}(t) \in H^1(\Omega; \mathbb{S}^2)$ evolve according to LLG (see (1.1)), which in rescaled form reads as

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} \times (\boldsymbol{h}_{\text{eff}}[\boldsymbol{m}] - \alpha\, \partial_t \boldsymbol{m}) \qquad \text{in } \Omega, \tag{2.4}$$

where $\alpha > 0$. Note that stationary solutions to (2.4) satisfy (2.2a). A simple formal computation reveals the orthogonality $\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0$, from which it follows that the dynamics inherently preserves the unit-length constraint. Moreover, taking the scalar product of (2.4) with $\boldsymbol{h}_{\text{eff}}[\boldsymbol{m}] - \alpha\, \partial_t \boldsymbol{m}$, one can show that any sufficiently smooth solution of (2.4) satisfies the energy law

$$\frac{\mathrm{d}}{\mathrm{d}t} \mathcal{E}[\boldsymbol{m}(t)] = -\alpha \|\partial_t \boldsymbol{m}(t)\|^2_{\boldsymbol{L}^2(\Omega)} \leq 0 \quad \text{for all } t > 0. \tag{2.5}$$

Hence, the dynamics is dissipative with the dissipation being modulated by the parameter $\alpha$.

Inertial effects can be included in the model by adding a term on the right-hand side of (2.4) (see (1.2)). The resulting equation, iLLG, is given by

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} \times (\boldsymbol{h}_{\text{eff}}[\boldsymbol{m}] - \alpha\, \partial_t \boldsymbol{m} - \tau\, \partial_{tt} \boldsymbol{m}) \qquad \text{in } \Omega, \tag{2.6}$$

where $\tau > 0$. Let $\boldsymbol{v} := \partial_t \boldsymbol{m}$ and $\boldsymbol{w} := \boldsymbol{m} \times \partial_t \boldsymbol{m} = \boldsymbol{m} \times \boldsymbol{v}$. Using the jargon of kinematics, we refer to these two quantities as *linear velocity* and *angular momentum*, respectively. By construction, $\boldsymbol{m}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ are mutually orthogonal. Moreover, since $|\boldsymbol{m}| = 1$ and $\boldsymbol{m} \cdot \boldsymbol{v} = 0$, it holds that $|\boldsymbol{v}| = |\boldsymbol{w}|$. The same computation leading to (2.5) yields the energy law of iLLG:

$$\frac{\mathrm{d}}{\mathrm{d}t} \left( \mathcal{E}[\boldsymbol{m}(t)] + \frac{\tau}{2} \|\partial_t \boldsymbol{m}(t)\|^2_{\boldsymbol{L}^2(\Omega)} \right) = -\alpha \|\partial_t \boldsymbol{m}(t)\|^2_{\boldsymbol{L}^2(\Omega)} \leq 0 \quad \text{for all } t > 0. \tag{2.7}$$

This motivates the definition of the extended energy functional

$$\mathcal{J}(\boldsymbol{m}, \boldsymbol{u}) = \mathcal{E}[\boldsymbol{m}] + \frac{\tau}{2} \|\boldsymbol{u}\|^2_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \boldsymbol{m} \in H^1(\Omega; \mathbb{S}^2) \text{ and } \boldsymbol{u} \in \boldsymbol{L}^2(\Omega). \tag{2.8}$$

With this definition, the quantity decaying over time in the dynamics governed by (2.6) becomes $\mathcal{J}(\boldsymbol{m}, \partial_t \boldsymbol{m})$. Staying within the framework of kinematics, we can interpret $\mathcal{J}(\boldsymbol{m}, \partial_t \boldsymbol{m})$ as the *total energy* of the magnetisation, which comprises the *potential energy* $\mathcal{E}[\boldsymbol{m}]$ and the *kinetic energy* $\tau \|\partial_t \boldsymbol{m}\|^2_{\boldsymbol{L}^2(\Omega)}/2$.

The initial boundary value problem considered in this work consists of (2.6) supplemented with homogeneous Neumann boundary conditions (2.2b), which make the dynamic problem compatible with the stationary case (2.2), and suitable initial conditions $\boldsymbol{m}(0) = \boldsymbol{m}^0$ and $\partial_t \boldsymbol{m}(0) = \boldsymbol{v}^0$.

We conclude this section by presenting the definition of a weak solution of (2.6), which is obtained by extending Definition 1.2 of [4] to the present setting; see also [25].

**Definition 2.1.** Let $\boldsymbol{m}^0 \in H^1(\Omega; \mathbb{S}^2)$ and $\boldsymbol{v}^0 \in \boldsymbol{L}^2(\Omega)$ such that $\boldsymbol{m}^0 \cdot \boldsymbol{v}^0 = 0$ a.e. in $\Omega$. A vector field $\boldsymbol{m} : \Omega \times (0, \infty) \to \mathbb{S}^2$ is called a *global weak solution* of iLLG (2.6) if $\boldsymbol{m} \in L^\infty(0, \infty; \boldsymbol{H}^1(\Omega)) \cap W^{1,\infty}(0, \infty; \boldsymbol{L}^2(\Omega))$ and, for all $T > 0$, the following properties are satisfied:

(i) $\boldsymbol{m} \in \boldsymbol{H}^1(\Omega_T)$, where $\Omega_T := \Omega \times (0, T)$;
(ii) $\boldsymbol{m}(t) \to \boldsymbol{m}^0$ in $\boldsymbol{H}^1(\Omega)$ and $\partial_t \boldsymbol{m}(t) \to \boldsymbol{v}^0$ in $\boldsymbol{L}^2(\Omega)$ as $t \to 0$;

(iii) For all $\boldsymbol{\varphi} \in C_c^\infty([0,T]; \boldsymbol{H}^1(\Omega))$, it holds that

$$
\begin{aligned}
\int_0^T \langle \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t = &-\int_0^T \langle \boldsymbol{h}_{\mathrm{eff}}[\boldsymbol{m}(t)], \boldsymbol{\varphi}(t) \times \boldsymbol{m}(t) \rangle \, \mathrm{d}t + \alpha \int_0^T \langle \boldsymbol{m}(t) \times \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t \\
&- \tau \int_0^T \langle \boldsymbol{m}(t) \times \partial_t \boldsymbol{m}(t), \partial_t \boldsymbol{\varphi}(t) \rangle \, \mathrm{d}t - \tau \langle \boldsymbol{m}^0 \times \boldsymbol{v}^0, \boldsymbol{\varphi}(0) \rangle;
\end{aligned}
\tag{2.9}
$$

(iv) It holds that

$$
\mathcal{J}(\boldsymbol{m}(T), \partial_t \boldsymbol{m}(T)) + \alpha \int_0^T \|\partial_t \boldsymbol{m}(t)\|_{\boldsymbol{L}^2(\Omega)}^2 \mathrm{d}t \le \mathcal{J}(\boldsymbol{m}^0, \boldsymbol{v}^0).
\tag{2.10}
$$

In Definition 2.1, equation (2.9) comes from a variational formulation of (2.6) in the space-time cylinder $\Omega_T$, where we integrate by parts in time the inertial contribution in order to lower the requested regularity in time of $\boldsymbol{m}$. The energy inequality (2.10) is the weak counterpart of (2.7).

**Remark 2.2.** (i) The setting discussed in this section can be obtained from the original equations expressed in physical units after a suitable rescaling. Let $t$ and $x$ denote the time and spatial variables (measured in s and m, respectively). First, we perform the change of variables $t' = \gamma_0 M_{\mathrm{s}} t$ and $x' = x/\ell_{\mathrm{ex}}$, where $M_{\mathrm{s}} > 0$ and $\ell_{\mathrm{ex}} > 0$ denote the saturation magnetisation (in $\mathrm{A\,m}^{-1}$) and the exchange length (in m) of the material, respectively. Second, we rescale the energy $\mathcal{E}[\boldsymbol{m}]$ (in J), the effective field $\boldsymbol{H}_{\mathrm{eff}}[\boldsymbol{m}]$ (in $\mathrm{A\,m}^{-1}$), and the angular momentum relaxation time $\tau$ (in s) in the following way: $\mathcal{E}'[\boldsymbol{m}] = \mathcal{E}[\boldsymbol{m}]/(\mu_0 M_{\mathrm{s}}^2 \ell_{\mathrm{ex}}^3)$, $\boldsymbol{h}_{\mathrm{eff}}[\boldsymbol{m}] = \boldsymbol{H}_{\mathrm{eff}}[\boldsymbol{m}]/M_{\mathrm{s}}$, $\tau' = \gamma_0 M_{\mathrm{s}} \tau$. Then, using the chain rule, (1.2) can be rewritten as (2.6), where all "primes" are omitted from the rescaled quantities in order to simplify the notation.

(ii) For ease of presentation, in the micromagnetic energy functional (2.1), we consider only the leading-order exchange contribution. The numerical treatment of standard lower-order energy contributions (*e.g.* magnetocrystalline anisotropy, Zeeman energy, magnetostatic energy, Dzyaloshinskii–Moriya interaction) is well understood; see, *e.g.* [17, 21, 26, 34].

## 3. NUMERICAL ALGORITHMS AND MAIN RESULTS

In this section, we introduce two fully discrete algorithms for the numerical approximation of iLLG and we state the corresponding stability and convergence results.

### 3.1. Preliminaries

For the time discretisation, we consider a uniform partition of the positive real axis $(0, \infty)$ with time-step size $k > 0$, *i.e.* $t_i := ik$ for all $i \in \mathbb{N}_0$. Given a sequence $\{\phi^i\}_{i \in \mathbb{N}_0}$, for all $i \in \mathbb{N}_0$, we define $d_t \phi^{i+1} := (\phi^{i+1} - \phi^i)/k$ and $\phi^{i+1/2} := (\phi^{i+1} + \phi^i)/2$. Interpreting the sequence $\{\phi^i\}_{i \in \mathbb{N}_0}$ as a collection of snapshots of a time-dependent function, we consider the time reconstructions $\phi_k$, $\phi_k^-$, $\overline{\phi}_k$, $\phi_k^+$ defined, for all $i \in \mathbb{N}_0$ and $t \in [t_i, t_{i+1})$, as

$$
\phi_k(t) := \frac{t - t_i}{k} \phi^{i+1} + \frac{t_{i+1} - t}{k} \phi^i, \quad \phi_k^-(t) := \phi^i, \quad \overline{\phi}_k(t) := \phi^{i+1/2}, \quad \text{and} \quad \phi_k^+(t) := \phi^{i+1}.
\tag{3.1}
$$

Note that $\partial_t \phi_k(t) = d_t \phi^{i+1}$ for all $i \in \mathbb{N}_0$ and $t \in [t_i, t_{i+1})$.

For the spatial discretisation, we assume $\Omega$ to be a polytopal domain with Lipschitz boundary and consider a shape-regular family $\{\mathcal{T}_h\}_{h>0}$ of tetrahedral meshes of $\Omega$ parametrised by the mesh size $h = \max_{K \in \mathcal{T}_h} \mathrm{diam}(K)$. Moreover, let $h_{\min} = \min_{K \in \mathcal{T}_h} \mathrm{diam}(K)$. We denote by $\mathcal{N}_h$ the set of vertices of $\mathcal{T}_h$. For any $K \in \mathcal{T}_h$, let $\mathcal{P}^1(K)$ be the space of first-order polynomials on $K$. We denote by $\mathcal{S}^1(\mathcal{T}_h)$ the space of piecewise affine and globally continuous functions from $\Omega$ to $\mathbb{R}$, *i.e.*

$$
\mathcal{S}^1(\mathcal{T}_h) = \{v_h \in C^0(\overline{\Omega}) : v_h|_K \in \mathcal{P}^1(K) \text{ for all } K \in \mathcal{T}_h\}.
$$

Its classical basis is given by the set of the nodal hat functions $\{\varphi_z\}_{z \in \mathcal{N}_h}$, which satisfy $\varphi_z(z') = \delta_{z,z'}$ for all $z, z' \in \mathcal{N}_h$. Let $\mathcal{I}_h : C^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)$ denote the nodal interpolant defined by $\mathcal{I}_h[v] = \sum_{z \in \mathcal{N}_h} v(z)\varphi_z$ for all $v \in C^0(\overline{\Omega})$. We denote by $\boldsymbol{\mathcal{I}}_h : \boldsymbol{C}^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)^3$ its vector-valued counterpart. We consider the mass-lumped product $\langle \cdot, \cdot \rangle_h$ defined by

$$\langle \boldsymbol{\psi}, \boldsymbol{\phi} \rangle_h = \int_\Omega \mathcal{I}_h[\boldsymbol{\psi} \cdot \boldsymbol{\phi}] \quad \text{for all } \boldsymbol{\psi}, \boldsymbol{\phi} \in \boldsymbol{C}^0(\overline{\Omega}). \tag{3.2}$$

Moreover, we define the mapping $\mathbb{P}_h : \boldsymbol{H}^1(\Omega)^\star \to \mathcal{S}^1(\mathcal{T}_h)^3$ by

$$\langle \mathbb{P}_h \boldsymbol{u}, \boldsymbol{\phi}_h \rangle_h = \langle \boldsymbol{u}, \boldsymbol{\phi}_h \rangle \quad \text{for all } \boldsymbol{u} \in \boldsymbol{H}^1(\Omega)^\star \text{ and } \boldsymbol{\phi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3. \tag{3.3}$$

Finally, we say that a mesh satisfies the *angle condition* if all off-diagonal entries of the so-called stiffness matrix are nonpositive, *i.e.*

$$\langle \nabla\varphi_z, \nabla\varphi_{z'} \rangle \leq 0 \quad \text{for all } z, z' \in \mathcal{N}_h \text{ with } z \neq z'. \tag{3.4}$$

Meshes satisfying (3.4) are sometimes also referred to as *weakly acute* or *nonobtuse* meshes and are relevant for the construction of finite element approximations of elliptic problems satisfying discrete maximum principles. The validity of (3.4) usually requires restrictive geometric properties of the mesh: For example, a sufficient condition to hold in 3D is that the measure of all dihedral angles of all tetrahedra of the mesh is smaller than or equal to $\pi/2$. For more details, we refer the interested reader, *e.g.* to the papers [7,29] as well as to Section 3.2.3 of [10].

## 3.2. Numerical algorithms

In the following algorithms, the main identities satisfied by any solution $\boldsymbol{m}$ of LLG/iLLG, *i.e.* $|\boldsymbol{m}| = 1$ and $\boldsymbol{m} \cdot \partial_t \boldsymbol{m} = 0$, are imposed only at the vertices of the mesh $\mathcal{T}_h$. To this end, we define the *set of admissible discrete magnetisations*

$$\boldsymbol{\mathcal{M}}_h := \left\{ \boldsymbol{\phi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3 : |\boldsymbol{\phi}_h(z)| = 1 \text{ for all } z \in \mathcal{N}_h \right\}$$

and, for $\boldsymbol{\psi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, the *discrete tangent space* of $\boldsymbol{\psi}_h$

$$\boldsymbol{\mathcal{K}}_h[\boldsymbol{\psi}_h] := \left\{ \boldsymbol{\phi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3 : \boldsymbol{\psi}_h(z) \cdot \boldsymbol{\phi}_h(z) = 0 \text{ for all } z \in \mathcal{N}_h \right\}. \tag{3.5}$$

The discrete counterpart of the functional (2.8) is defined by

$$\mathcal{J}_h(\boldsymbol{m}_h, \boldsymbol{u}_h) = \mathcal{E}[\boldsymbol{m}_h] + \frac{\tau}{2} \|\boldsymbol{u}_h\|_h^2 \quad \text{for all } \boldsymbol{m}_h, \boldsymbol{u}_h \in \mathcal{S}^1(\mathcal{T}_h)^3.$$

### 3.2.1. Tangent plane scheme

The first method uses the linear velocity $\boldsymbol{v} = \partial_t \boldsymbol{m}$ as an auxiliary variable. Using the vector identity

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c} \quad \text{for all } \mathbf{a}, \mathbf{b}, \mathbf{c} \in \mathbb{R}^3, \tag{3.6}$$

together with the properties $|\boldsymbol{m}| = 1$ and $\boldsymbol{m} \cdot \boldsymbol{v} = 0$, iLLG can be formally rewritten as

$$\tau \, \partial_t \boldsymbol{v} + \alpha \, \boldsymbol{v} + \boldsymbol{m} \times \boldsymbol{v} = \boldsymbol{h}_{\text{eff}}[\boldsymbol{m}] - (\boldsymbol{h}_{\text{eff}}[\boldsymbol{m}] \cdot \boldsymbol{m})\boldsymbol{m} - \tau|\boldsymbol{v}|^2 \boldsymbol{m}. \tag{3.7}$$

Following the tangent plane paradigm [2,3,5,14], to obtain a numerical scheme for iLLG, we consider a finite element approximation of a mass-lumped variational formulation of (3.7) based on test functions fulfilling the same orthogonality property satisfied by $\boldsymbol{v}$. This yields a natural linearisation of (3.7), as the contributions associated with last two (nonlinear) terms on the right-hand side vanish by orthogonality. More precisely, for all time-steps $i \in \mathbb{N}_0$, given the current approximations $\boldsymbol{m}_h^i \approx \boldsymbol{m}(t_i)$ and $\boldsymbol{v}_h^i \approx \boldsymbol{v}(t_i)$, we compute $\boldsymbol{v}_h^{i+1} \approx \boldsymbol{v}(t_{i+1})$ using a discretised version of (3.7), which is based on the discrete tangent space $\boldsymbol{\mathcal{K}}_h[\boldsymbol{m}_h^i]$ introduced in (3.5) for the spatial discretisation and on the backward Euler method for the temporal discretisation. Then, using

the available approximations $\boldsymbol{m}_h^i$ and $\boldsymbol{v}_h^{i+1}$, we obtain $\boldsymbol{m}_h^{i+1} \approx \boldsymbol{m}(t_{i+1})$ *via* a first-order time-stepping, *i.e.* $\mathcal{I}_h\big[\big(\boldsymbol{m}_h^i + k\boldsymbol{v}_h^{i+1}\big)/\big|\boldsymbol{m}_h^i + k\boldsymbol{v}_h^{i+1}\big|\big]$, where the nodal projection is employed to ensure that the new approximation belongs to $\boldsymbol{\mathcal{M}}_h$. Unlike [2, 3, 5] and as in [14], we use the mass-lumped product $\langle\cdot,\cdot\rangle_h$, which enhances the efficiency of the scheme without affecting its formal convergence order. The resulting scheme is summarised in the following algorithm.

**Algorithm 3.1** (Tangent plane scheme). *Input:* $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ and $\boldsymbol{v}_h^0 \in \boldsymbol{\mathcal{K}}_h\big[\boldsymbol{m}_h^0\big]$.
*Loop:* For all $i \in \mathbb{N}_0$, iterate (i)–(ii):

(i) Compute $\boldsymbol{v}_h^{i+1} \in \boldsymbol{\mathcal{K}}_h\big[\boldsymbol{m}_h^i\big]$ such that, for all $\boldsymbol{\phi}_h \in \boldsymbol{\mathcal{K}}_h\big[\boldsymbol{m}_h^i\big]$, it holds that

$$\tau\big\langle d_t\boldsymbol{v}_h^{i+1}, \boldsymbol{\phi}_h\big\rangle_h + \alpha\big\langle \boldsymbol{v}_h^{i+1}, \boldsymbol{\phi}_h\big\rangle_h + \big\langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^{i+1}, \boldsymbol{\phi}_h\big\rangle_h - k\big\langle \mathbb{P}_h\boldsymbol{h}_{\text{eff}}\big[\boldsymbol{v}_h^{i+1}\big], \boldsymbol{\phi}_h\big\rangle_h = \big\langle \mathbb{P}_h\boldsymbol{h}_{\text{eff}}\big[\boldsymbol{m}_h^i\big], \boldsymbol{\phi}_h\big\rangle_h. \quad (3.8)$$

(ii) Define $\boldsymbol{m}_h^{i+1} = \mathcal{I}_h\big[\big(\boldsymbol{m}_h^i + k\boldsymbol{v}_h^{i+1}\big)/\big|\boldsymbol{m}_h^i + k\boldsymbol{v}_h^{i+1}\big|\big] \in \boldsymbol{\mathcal{M}}_h$.

*Output:* Sequence of approximations $\big\{\big(\boldsymbol{m}_h^{i+1}, \boldsymbol{v}_h^{i+1}\big)\big\}_{i\in\mathbb{N}_0}$.

Algorithm 3.1 is well-defined: The (nonsymmetric) bilinear form on the left-hand side of (3.8) in step (i) is elliptic, so that existence and uniqueness of a solution $\boldsymbol{v}_h^{i+1} \in \boldsymbol{\mathcal{K}}_h\big[\boldsymbol{m}_h^i\big]$ are guaranteed by the Lax–Milgram theorem; In the nodal projection appearing in step (ii), the denominator is bounded from below by 1, so that division by zero never occurs.

*3.2.2. Angular momentum method*

The second method, following [9, 27], uses the angular momentum $\boldsymbol{w} = \boldsymbol{m} \times \partial_t\boldsymbol{m}$ as an auxiliary variable. First, note that

$$\boldsymbol{m} \times \boldsymbol{w} = \boldsymbol{m} \times (\boldsymbol{m} \times \partial_t\boldsymbol{m}) = -\partial_t\boldsymbol{m},$$

where the second identity follows from (3.6), together with $|\boldsymbol{m}| = 1$ and $\boldsymbol{m} \cdot \partial_t\boldsymbol{m} = 0$. Then, we have that

$$\tau\,\partial_t\boldsymbol{w} = \tau\,\boldsymbol{m} \times \partial_{tt}\boldsymbol{m} \overset{(2.6)}{=} \boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}[\boldsymbol{m}] - \alpha\,\boldsymbol{m} \times \partial_t\boldsymbol{m} + \partial_t\boldsymbol{m}.$$

We infer the first-order (in time) system

$$\partial_t\boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{w},$$
$$\tau\,\partial_t\boldsymbol{w} = \boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}[\boldsymbol{m}] - \alpha\,\boldsymbol{m} \times \partial_t\boldsymbol{m} - \boldsymbol{m} \times \boldsymbol{w}.$$

For all $i \in \mathbb{N}_0$, given $\boldsymbol{m}_h^i \approx \boldsymbol{m}(t_i)$ and $\boldsymbol{w}_h^i \approx \boldsymbol{w}(t_i)$, approximations $\boldsymbol{m}_h^{i+1} \approx \boldsymbol{m}(t_{i+1})$ and $\boldsymbol{w}_h^{i+1} \approx \boldsymbol{w}(t_{i+1})$ are computed by solving a mass-lumped variational formulation of this first-order system, where the time discretisation is based on the midpoint rule. The resulting scheme is stated in the following algorithm.

**Algorithm 3.2** (Nonlinear angular momentum method). *Input:* $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ and $\boldsymbol{v}_h^0 \in \boldsymbol{\mathcal{K}}_h\big[\boldsymbol{m}_h^0\big]$.
*Initialisation:* Define $\boldsymbol{w}_h^0 = \mathcal{I}_h\big[\boldsymbol{m}_h^0 \times \boldsymbol{v}_h^0\big] \in \boldsymbol{\mathcal{K}}_h\big[\boldsymbol{m}_h^0\big]$.
*Loop:* For all $i \in \mathbb{N}_0$, compute $\big(\boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h^{i+1}\big) \in \boldsymbol{\mathcal{M}}_h \times \boldsymbol{\mathcal{K}}_h\big[\boldsymbol{m}_h^{i+1}\big]$ such that, for all $(\boldsymbol{\phi}_h, \boldsymbol{\psi}_h) \in \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that

$$\big\langle d_t\boldsymbol{m}_h^{i+1}, \boldsymbol{\phi}_h\big\rangle_h = -\big\langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{w}_h^{i+1/2}, \boldsymbol{\phi}_h\big\rangle_h, \quad (3.9a)$$

$$\tau\big\langle d_t\boldsymbol{w}_h^{i+1}, \boldsymbol{\psi}_h\big\rangle_h = \big\langle \boldsymbol{m}_h^{i+1/2} \times \mathbb{P}_h\boldsymbol{h}_{\text{eff}}\big[\boldsymbol{m}_h^{i+1/2}\big], \boldsymbol{\psi}_h\big\rangle_h - \alpha\big\langle \boldsymbol{m}_h^{i+1/2} \times d_t\boldsymbol{m}_h^{i+1}, \boldsymbol{\psi}_h\big\rangle_h$$
$$- \big\langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{w}_h^{i+1/2}, \boldsymbol{\psi}_h\big\rangle_h. \quad (3.9b)$$

*Output:* Sequence of approximations $\big\{\big(\boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h^{i+1}\big)\big\}_{i\in\mathbb{N}_0}$.

In the following proposition, we show that the pointwise constraints $|\boldsymbol{m}| = 1$ and $\boldsymbol{m} \cdot \boldsymbol{w} = 0$ are inherently preserved by Algorithm 3.2 (at the vertices of the mesh). Its proof is deferred to Section 5.1.

**Proposition 3.3.** *Let $i \in \mathbb{N}_0$. The approximations generated by Algorithm 3.2 satisfy $\boldsymbol{m}_h^{i+1} \in \boldsymbol{\mathcal{M}}_h$ and $\boldsymbol{w}_h^{i+1} \in \boldsymbol{\mathcal{K}}_h[\boldsymbol{m}_h^{i+1}]$.*

The computation of $(\boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h^{i+1})$ satisfying (3.9) involves the solution of a nonlinear system of equations. An effective implementation requires a linearisation.

Let $\boldsymbol{u}_h^i := \boldsymbol{m}_h^{i+1/2}$ and $\boldsymbol{z}_h^i := \boldsymbol{w}_h^{i+1/2}$. Performing simple algebraic manipulations, we rewrite (3.9) with respect to the unknowns $\boldsymbol{u}_h^i$ and $\boldsymbol{z}_h^i$:

$$2\langle \boldsymbol{u}_h^i, \boldsymbol{\phi}_h \rangle_h + k\langle \boldsymbol{u}_h^i \times \boldsymbol{z}_h^i, \boldsymbol{\phi}_h \rangle_h = 2\langle \boldsymbol{m}_h^i, \boldsymbol{\phi}_h \rangle_h, \tag{3.10a}$$

$$2\tau\langle \boldsymbol{z}_h^i, \boldsymbol{\psi}_h \rangle_h - k\langle \boldsymbol{u}_h^i \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}[\boldsymbol{u}_h^i], \boldsymbol{\psi}_h \rangle_h$$
$$-2\alpha\langle \boldsymbol{u}_h^i \times \boldsymbol{m}_h^i, \boldsymbol{\psi}_h \rangle_h + k\langle \boldsymbol{u}_h^i \times \boldsymbol{z}_h^i, \boldsymbol{\psi}_h \rangle_h = 2\tau\langle \boldsymbol{w}_h^i, \boldsymbol{\psi}_h \rangle_h. \tag{3.10b}$$

Starting from this formulation, in the following algorithm we introduce a linear fixed-point iteration (similar in spirit to those considered in [9,12]), which provides an effective implementation of Algorithm 3.2.

**Algorithm 3.4** (Linearised angular momentum method). *Input:* $\boldsymbol{m}_h^0 \in \boldsymbol{\mathcal{M}}_h$ and $\boldsymbol{v}_h^0 \in \boldsymbol{\mathcal{K}}_h[\boldsymbol{m}_h^0]$.
*Initialisation:* Define $\boldsymbol{w}_h^0 = \boldsymbol{\mathcal{I}}_h[\boldsymbol{m}_h^0 \times \boldsymbol{v}_h^0] \in \boldsymbol{\mathcal{K}}_h[\boldsymbol{m}_h^0]$.
*Loop:* For all $i \in \mathbb{N}_0$, iterate (i) and (ii):

(i) Let $\boldsymbol{u}_h^{i,0} = \boldsymbol{m}_h^i$ and $\boldsymbol{z}_h^{i,0} = \boldsymbol{w}_h^i$. For all $\ell \in \mathbb{N}_0$, iterate (i-a) and (i-b):
(i-a) Compute $\boldsymbol{u}_h^{i,\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that, for all $\boldsymbol{\phi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that

$$2\left\langle \boldsymbol{u}_h^{i,\ell+1}, \boldsymbol{\phi}_h \right\rangle_h + k\left\langle \boldsymbol{u}_h^{i,\ell+1} \times \boldsymbol{z}_h^{i,\ell}, \boldsymbol{\phi}_h \right\rangle_h = 2\langle \boldsymbol{m}_h^i, \boldsymbol{\phi}_h \rangle_h; \tag{3.11}$$

(i-b) Compute $\boldsymbol{z}_h^{i,\ell+1} \in \mathcal{S}^1(\mathcal{T}_h)^3$ such that, for all $\boldsymbol{\psi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3$, it holds that

$$2\tau\left\langle \boldsymbol{z}_h^{i,\ell+1}, \boldsymbol{\psi}_h \right\rangle_h + k\left\langle \boldsymbol{u}_h^{i,\ell+1} \times \boldsymbol{z}_h^{i,\ell+1}, \boldsymbol{\psi}_h \right\rangle_h$$
$$= k\left\langle \boldsymbol{u}_h^{i,\ell+1} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{u}_h^{i,\ell+1}\right], \boldsymbol{\psi}_h \right\rangle_h + 2\alpha\left\langle \boldsymbol{u}_h^{i,\ell+1} \times \boldsymbol{m}_h^i, \boldsymbol{\psi}_h \right\rangle_h + 2\tau\langle \boldsymbol{w}_h^i, \boldsymbol{\psi}_h \rangle_h; \tag{3.12}$$

until

$$\left\|\boldsymbol{u}_h^{i,\ell+1} - \boldsymbol{u}_h^{i,\ell}\right\|_h + \left\|\boldsymbol{z}_h^{i,\ell+1} - \boldsymbol{z}_h^{i,\ell}\right\|_h \le \varepsilon. \tag{3.13}$$

(ii) Let $\ell_i \in \mathbb{N}_0$ be the smallest integer for which the stopping criterion (3.13) is met. Define $\boldsymbol{m}_h^{i+1} := 2\boldsymbol{u}_h^{i,\ell_i+1} - \boldsymbol{m}_h^i$ and $\boldsymbol{w}_h^{i+1} := 2\boldsymbol{z}_h^{i,\ell_i+1} - \boldsymbol{w}_h^i$.

*Output:* Sequence of approximations $\left\{\left(\boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h^{i+1}\right)\right\}_{i \in \mathbb{N}_0}$.

The well-posedness and the conservation properties of Algorithm 3.4 are the subject of the following proposition. Its proof is postponed to Section 5.1.

**Proposition 3.5.** *Let $i \in \mathbb{N}_0$. Suppose that $\boldsymbol{m}_h^i \in \boldsymbol{\mathcal{M}}_h$.*

(i) *For all $\ell \in \mathbb{N}_0$, (3.11) and (3.12) admit unique solutions $\boldsymbol{u}_h^{i,\ell+1}$ and $\boldsymbol{z}_h^{i,\ell+1}$ in $\mathcal{S}^1(\mathcal{T}_h)^3$. Moreover, it holds that $\left\|\boldsymbol{u}_h^{i,\ell+1}\right\|_{\boldsymbol{L}^\infty(\Omega)} \le 1$.*

(ii) *There exist $k_0 > 0$ and $C > 0$ such that, if $k < k_0$ and $k < Ch_{\min}$, then, for all $\ell \in \mathbb{N}_0$, it holds that*

$$\left\| \boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1} \right\|_h + \left\| \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1} \right\|_h \leq q\left( \left\| \boldsymbol{u}_h^{i,\ell+1} - \boldsymbol{u}_h^{i,\ell} \right\|_h + \left\| \boldsymbol{z}_h^{i,\ell+1} - \boldsymbol{z}_h^{i,\ell} \right\|_h \right) \qquad (3.14)$$

*for some $0 < q < 1$. The constants $k_0$, $C$, and $q$ depend only on the shape-regularity of $\mathcal{T}_h$ and the problem data.*

(iii) *Under the assumptions of part* (ii), *the stopping criterion* (3.13) *is met in a finite number of iterations. If $\ell_i \in \mathbb{N}_0$ denotes the smallest integer for which* (3.13) *holds, the new approximations $\boldsymbol{m}_h^{i+1} = 2\boldsymbol{u}_h^{i,\ell_i+1} - \boldsymbol{m}_h^i$ and $\boldsymbol{w}_h^{i+1} = 2\boldsymbol{z}_h^{i,\ell_i+1} - \boldsymbol{w}_h^i$ belong to $\mathcal{M}_h$ and $\mathcal{K}_h\left[\boldsymbol{m}_h^{i+1}\right]$, respectively.*

Proposition 3.5(ii) shows that, under suitable assumptions, the mapping defining the fixed-point iteration is a contraction. Therefore, under the same assumptions, Banach fixed-point theorem ensures that (3.9) admits a unique solution so that Algorithm 3.2 is well-posed.

In view of the stability and convergence analysis, we observe that, for all $i \in \mathbb{N}_0$, the iterates $\left( \boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h^{i+1} \right) \in \mathcal{M}_h \times \mathcal{K}_h\left[\boldsymbol{m}_h^{i+1}\right]$ of Algorithm 3.4 satisfy

$$\left\langle d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{\phi}_h \right\rangle_h = -\left\langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{w}_h^{i+1/2}, \boldsymbol{\phi}_h \right\rangle_h + \left\langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{r}_h^i, \boldsymbol{\phi}_h \right\rangle_h,$$

$$\tau\left\langle d_t \boldsymbol{w}_h^{i+1}, \boldsymbol{\psi}_h \right\rangle_h = \left\langle \boldsymbol{m}_h^{i+1/2} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right], \boldsymbol{\psi}_h \right\rangle_h - \alpha\left\langle \boldsymbol{m}_h^{i+1/2} \times d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{\psi}_h \right\rangle_h$$
$$- \left\langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{w}_h^{i+1/2}, \boldsymbol{\psi}_h \right\rangle_h$$

for all $(\boldsymbol{\phi}_h, \boldsymbol{\psi}_h) \in \mathcal{S}^1(\mathcal{T}_h)^3 \times \mathcal{S}^1(\mathcal{T}_h)^3$, where $\boldsymbol{r}_h^i := \boldsymbol{z}_h^{i,\ell_i+1} - \boldsymbol{z}_h^{i,\ell_i} \in \mathcal{S}^1(\mathcal{T}_h)^3$ satisfies $\left\| \boldsymbol{r}_h^i \right\|_h \leq \varepsilon$.

### 3.3. Stability and convergence results

In the following proposition, we establish the discrete energy laws satisfied by the algorithms. Its proof is postponed to Section 5.2.

**Proposition 3.6** (Discrete energy law and stability). *Let $i \in \mathbb{N}_0$.*

(i) *Suppose that the mesh $\mathcal{T}_h$ satisfies angle condition* (3.4). *The approximations generated by Algorithm 3.1 satisfy the discrete energy law*

$$\mathcal{J}_h\left(\boldsymbol{m}_h^{i+1}, \boldsymbol{v}_h^{i+1}\right) + \alpha k\left\|\boldsymbol{v}_h^{i+1}\right\|_h^2 + \frac{\tau k^2}{2}\left\|d_t \boldsymbol{v}_h^{i+1}\right\|_h^2 + \frac{k^2}{2}\left\|\boldsymbol{\nabla}\boldsymbol{v}_h^{i+1}\right\|_{\boldsymbol{L}^2(\Omega)}^2 \leq \mathcal{J}_h\left(\boldsymbol{m}_h^i, \boldsymbol{v}_h^i\right). \qquad (3.15)$$

(ii) *The approximations generated by Algorithm 3.2 satisfy the discrete energy law*

$$\mathcal{J}_h\left(\boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h^{i+1}\right) + \alpha k\left\|d_t \boldsymbol{m}_h^{i+1}\right\|_h^2 = \mathcal{J}_h\left(\boldsymbol{m}_h^i, \boldsymbol{w}_h^i\right). \qquad (3.16)$$

(iii) *The approximations generated by Algorithm 3.4 satisfy the discrete energy law*

$$\mathcal{J}_h\left(\boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h^{i+1}\right) + \alpha k\left\|d_t \boldsymbol{m}_h^{i+1}\right\|_h^2 + k\left\langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{r}_h^i, \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right] - \alpha\, d_t \boldsymbol{m}_h^{i+1} \right\rangle_h = \mathcal{J}_h\left(\boldsymbol{m}_h^i, \boldsymbol{w}_h^i\right). \quad (3.17)$$

In the energy law satisfied by Algorithm 3.1, besides the LLG-intrinsic dissipation, we observe the presence of numerical dissipation due to the use of the backward Euler method; *cf.* the last two terms on the left-hand side of (3.15). Algorithm 3.2 fulfils a discrete energy identity, which reflects the fact that the midpoint rule is symplectic. The same identity, apart from an additional term coming from the inexact solution of the nonlinear system, is satisfied by Algorithm 3.4.

From each algorithm, we obtain a sequence of approximations $\{\boldsymbol{m}_h^i\}_{i\in\mathbb{N}_0}$, which we can use to define the piecewise affine time reconstruction $\boldsymbol{m}_{hk} : (0,\infty) \to \mathcal{S}^1(\mathcal{T}_h)^3$ (denoted by $\boldsymbol{m}_{hk}^\varepsilon$ in the case of Algorithm 3.4) as

$$\boldsymbol{m}_{hk}(t) := \frac{t - t_i}{k}\boldsymbol{m}_h^{i+1} + \frac{t_{i+1} - t}{k}\boldsymbol{m}_h^i \quad \text{for all } i \in \mathbb{N}_0 \text{ and } t \in [t_i, t_{i+1});$$

see (3.1). In the following theorem, we show that, under appropriate assumptions, the sequence $\{\boldsymbol{m}_{hk}\}$ (resp., $\{\boldsymbol{m}_{hk}^\varepsilon\}$ for Algorithm 3.4) converges in a suitable sense towards solutions of iLLG as $h$, $k$ (and $\varepsilon$) go to 0. Its proof is postponed to Section 5.3.

**Theorem 3.7.** *Let the approximate initial conditions satisfy*

$$\boldsymbol{m}_h^0 \to \boldsymbol{m}^0 \quad \text{in } \boldsymbol{H}^1(\Omega) \quad \text{and} \quad \boldsymbol{v}_h^0 \to \boldsymbol{v}^0 \quad \text{in } \boldsymbol{L}^2(\Omega) \quad \text{as } h \to 0. \tag{3.18}$$

(i) *For Algorithm 3.1, assume that each mesh $\mathcal{T}_h$ satisfies the angle condition (3.4) and that $k = o\left(h_{\min}^{d/2}\right)$ as $h, k \to 0$. For Algorithm 3.2, assume that the scheme is well-posed. Then, there exist a global weak solution $\boldsymbol{m} : \Omega \times (0,\infty) \to \mathbb{S}^2$ of iLLG in the sense of Definition 2.1 and a (nonrelabeled) subsequence of $\{\boldsymbol{m}_{hk}\}$ which converges towards $\boldsymbol{m}$ as $h, k \to 0$. In particular, as $h, k \to 0$, it holds that $\boldsymbol{m}_{hk} \stackrel{*}{\rightharpoonup} \boldsymbol{m}$ in $L^\infty(0, \infty; \boldsymbol{H}^1(\Omega; \mathbb{S}^2))$ and $\boldsymbol{m}_{hk}|_{\Omega_T} \rightharpoonup \boldsymbol{m}|_{\Omega_T}$ in $\boldsymbol{H}^1(\Omega_T)$ for all $T > 0$.*

(ii) *For Algorithm 3.4, assume that the scheme is well-posed and that $\varepsilon = \mathcal{O}(h_{\min})$ as $h, \varepsilon \to 0$. For all $T > 0$, there exist $\boldsymbol{m} : \Omega \times (0, T) \to \mathbb{S}^2$, which satisfies the requirements* (i)–(iv) *of Definition 2.1, and a (nonrelabeled) subsequence of $\{\boldsymbol{m}_{hk}^\varepsilon\}$ such that $\boldsymbol{m}_{hk}^\varepsilon|_{\Omega_T} \stackrel{*}{\rightharpoonup} \boldsymbol{m}$ in $L^\infty(0, T; \boldsymbol{H}^1(\Omega; \mathbb{S}^2))$ and $\boldsymbol{m}_{hk}^\varepsilon|_{\Omega_T} \rightharpoonup \boldsymbol{m}$ in $\boldsymbol{H}^1(\Omega_T)$ as $h, k, \varepsilon \to 0$.*

In order to refer to the limit of the approximations generated by Algorithm 3.4, we have not used the expression "global weak solution". This is a consequence of the fact that, for this algorithm, the boundedness result we are able to show (see Prop. 5.1 below) is not uniform with respect to the (arbitrary but fixed) final time $T > 0$.

To conclude, we summarise the results of our analysis for the proposed algorithms:

– Algorithm 3.1 is unconditionally well-posed, unconditionally stable (under the angle condition (3.4)), and its convergence towards a global weak solution of iLLG requires the CFL condition $k = o\left(h_{\min}^{d/2}\right)$ as $h, k \to 0$.

– Algorithm 3.2 is well-posed if $k$ is sufficiently small and the CFL condition $k < Ch_{\min}$ holds. Assuming its well-posedness, it is unconditionally stable and unconditionally convergent towards a global weak solution of iLLG.

– Algorithm 3.4 is well-posed if $k$ is sufficiently small and the CFL condition $k < Ch_{\min}$ holds. Assuming its well-posedness and choosing a stopping tolerance $\varepsilon$ having the same order of $h_{\min}$, for all $T > 0$, it is unconditionally stable and unconditionally convergent towards a function $\boldsymbol{m} : \Omega \times (0, T) \to \mathbb{S}^2$ which fulfils the properties (i)–(iv) of Definition 2.1.

**Remark 3.8.** For the sake of brevity, we restrict ourselves to the case of algorithms generating approximations which satisfy the unit-length constraint at the vertices of the mesh, *i.e.* $\boldsymbol{m}_h^i \in \mathcal{M}_h$ for all $i \in \mathbb{N}_0$. For the tangent plane scheme, this property is guaranteed by the use of the nodal projection. A theoretical consequence is that the stability analysis requires the assumption of the angle condition (3.4), which turns out to be quite restrictive in 3D. For LLG, a tangent plane scheme which avoids the nodal projection (and the related angle condition) was proposed in [1]; see also [11,22]. We believe that a similar approach can be pursued to construct a projection-free tangent plane scheme for iLLG.

## 4. Numerical results

Before presenting the proof of the results stated in Section 3, we aim to show the effectivity of the proposed algorithms by means of two numerical experiments. For the sake of brevity, in this section, we refer to the

TABLE 1. Experiment of Section 4.1: Average number of fixed-point iterations needed to reach the prescribed tolerance $\varepsilon = 1 \times 10^{-12}$ for $\ell = 5, 6, 7, 8$ and $\delta = 0.1, 0.2, 0.3, 0.4$.

|  | $\delta = 0.1$ | $\delta = 0.2$ | $\delta = 0.3$ | $\delta = 0.4$ |
|---|---|---|---|---|
| $\ell = 5$ | 4.99 | 7.99 | 13.87 | 35.66 |
| $\ell = 6$ | 4.93 | 7.89 | 13.82 | 35.39 |
| $\ell = 7$ | 4.84 | 7.75 | 13.66 | 34.55 |
| $\ell = 8$ | 4.78 | 7.72 | 13.44 | 33.74 |

tangent plane scheme (Algorithm 3.1) as TPS and to the angular momentum method (Algorithm 3.2 or, more appropriately, its effective realisation given in Algorithm 3.4) as AMM. The computation presented in this section were obtained with a MATLAB implementation of the proposed algorithms. All linear systems were solved using the direct solver provided by MATLAB's backslash operator.

### 4.1. Finite-time blow-up of weak solutions

We investigate the performance of the algorithms for different choices of the discretisation parameters $h$ and $k$ (and $\varepsilon$ for AMM). At the same time, we numerically study for a weak solution $\boldsymbol{m}$ of iLLG the occurrence of a so-called *finite-time blow-up*, *i.e.* whether there exists $T^* > 0$ such that

$$\lim_{t \to T^*} \|\boldsymbol{\nabla}\boldsymbol{m}(t)\|_{\boldsymbol{L}^\infty(\Omega)} = \infty.$$

To this end, we adapt to iLLG the model problem studied in [9, 13, 27] for the wave map equation and in [14] for LLG.

We consider the nondimensional setting presented in Section 2 for the unit square domain $\Omega = (-1/2, 1/2)^2$ in the time interval $(0, 2)$. We set $\alpha = \tau = 1$ in (2.6) and consider homogeneous Neumann boundary conditions (2.2b) as well as the initial conditions $\boldsymbol{m}^0(x) = \left(2a(x)x_1, 2a(x)x_2, a(x)^2 - |x|^2\right) / \left(a(x)^2 + |x|^2\right)$ with $a(x) = \max\{0, (1 - 2|x|)^4\}$ for all $x = (x_1, x_2) \in \Omega$ and $\boldsymbol{v}^0 \equiv \boldsymbol{0}$. For the wave map equation (1.3) and LLG (2.4), this setting leads to numerical approximations with large gradients, which suggests the occurrence of a finite-time blow-up. For snapshots of numerical approximations which illustrate this phenomenon, we refer to, *e.g.* Figures 3 and 4 of [13] or Figures 1 and 2 of [14].

First, we investigate the convergence of the fixed-point iteration in AMM analysed in Proposition 3.5(ii). For $\ell = 5, 6, 7, 8$, we consider a uniform mesh $\mathcal{T}_{h_\ell}$ of the unit square consisting of $2^{2\ell+1}$ rectangular triangles. The resulting mesh size is $h_\ell = \sqrt{2}\, 2^{-\ell}$. In the stopping criterion (3.13), in order to better evaluate the convergence of the fixed-point iteration, we use the small tolerance $\varepsilon = 1 \times 10^{-12}$. The iteration is terminated either when the stopping criterion (3.13) is met or when the number of iterations exceeds 1000. We use the time-step size $k_\ell = \delta h_\ell$ for different values of $0 < \delta < 1$.

In Table 1, we show the average number of fixed-point iterations needed to reach the prescribed tolerance $\varepsilon$ for $\delta = 0.1, 0.2, 0.3, 0.4$. The number of iterations increases as $\delta$ increases and decreases (very slightly) as the mesh size decreases. For all $\ell = 5, 6, 7, 8$, the fixed-point iteration does not converge (within the prescribed maximum number of iterations) if $\delta$ is larger than a threshold value located between 0.46 and 0.47. This behaviour is in agreement with the dependence of the contraction constant on the discretisation parameters which can be inferred from the proof of Proposition 3.5(ii), *i.e.* $q \simeq \delta(1 + h)$ (recall that the tolerance $\varepsilon$ is fixed).

Next, we compare the performance of TPS and AMM. We consider the uniform mesh $\mathcal{T}_{h_6}$ (8192 elements and mesh size $h_6 = 0.0221$) and $k = h_6/10$. For AMM, we set $\varepsilon = h_6/10$ in (3.13).

In Figure 2, we plot the evolutions of the spatial average of the third magnetisation component, *i.e.* $\langle m_3(t) \rangle := |\Omega|^{-1} \int_\Omega \boldsymbol{m}_{hk}(t) \cdot \boldsymbol{e}_3$, of the $W^{1,\infty}$-seminorm $\|\boldsymbol{\nabla}\boldsymbol{m}_{hk}(t)\|_{\boldsymbol{L}^\infty(\Omega)}$, and of the total discrete energy $\mathcal{J}_h(\boldsymbol{m}_{hk}(t), \partial_t \boldsymbol{m}_{hk}(t))$ for $t \in [0, 2]$. We observe that the algorithms capture the same average magnetisation
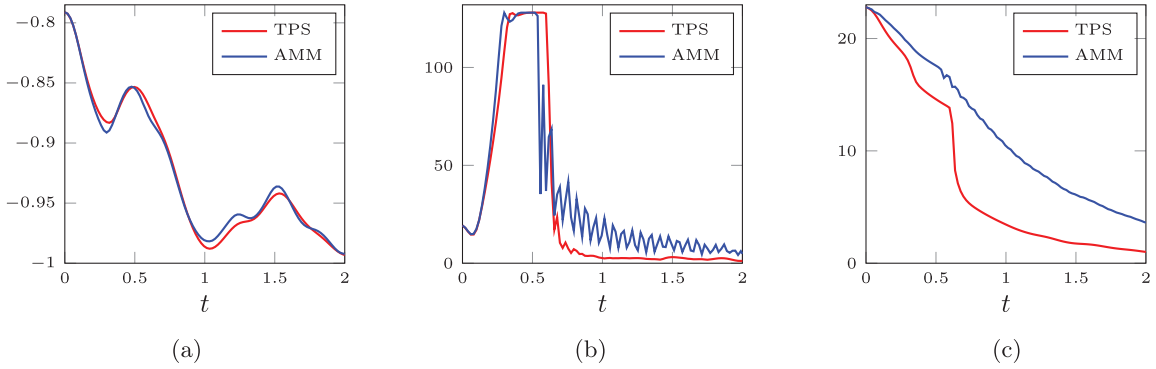
FIGURE 2. Experiment of Section 4.1: comparison of the results obtained with TPS and AMM for $\ell = 6$, $k = h_6/10$, and $\varepsilon = h_6/10$. (a) Evolution of the average magnetisation $\langle m_3 \rangle$. (b) Evolution of the $W^{1,\infty}$-seminorm. (c) Evolution of the total energy.
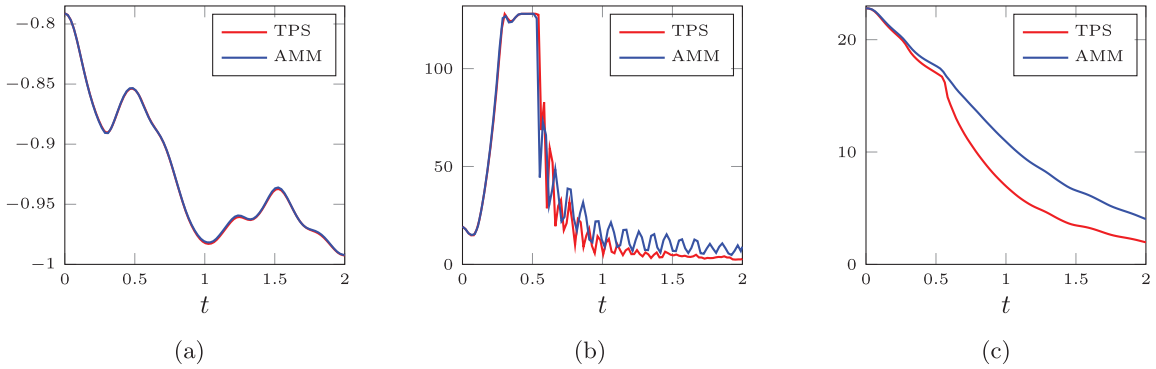


FIGURE 3. Experiment of Section 4.1: comparison of the results obtained with TPS and AMM for $\ell = 6$, $k = h_6/100$, and $\varepsilon = h_6/100$. (a) Evolution of the average magnetisation $\langle m_3 \rangle$. (b) Evolution of the $W^{1,\infty}$-seminorm. (c) Evolution of the total energy.

dynamics. In particular, at $t \approx 0.3$, the approximations attain the largest possible value of the $W^{1,\infty}$-seminorm for functions in $\mathcal{M}_h$ residing in $\mathcal{T}_{h_6}$, which in general, for all $\ell = 5, 6, 7, 8$, is given by

$$\max_{\phi_h \in \mathcal{M}_h} \|\nabla \phi_h\|_{\mathbf{L}^\infty(\Omega)} = \max_{\phi_h \in \mathcal{M}_h} \max_{T \in \mathcal{T}_{h,\ell}} |\nabla \phi_h|_T| = 2/2^{-\ell} = 2^{\ell+1}. \tag{4.1}$$

This value is obtained when the magnetisations of two neighboring vertices point to opposite directions. Indeed, in our case, the magnetisation at $(0,0)$ points to the out-of-plane direction $(1,0,0)$, while all surrounding vectors point to the opposite direction. This configuration lasts for some time (see the "plateau" in Fig. 2b). Then, at $t \approx 0.8$, the magnetisation at $(0,0)$ is reversed. This gives rise to oscillations of decaying amplitude. Looking at Figure 2c, we observe that, in agreement with (3.15), the total energy decays monotonically in the case of TPS. In the case of AMM, the decay is nonmonotone. Note that possible lack of monotonicity is predicted by the energy law of AMM; *cf.* the (unsigned) third term on the left-hand side (3.17). Moreover, we see that the curve for TPS is well below the one of AMM. This fact can be justified by the dissipation of the backward Euler method; *cf.* the second and third terms on the left-hand side of (3.15).

In Figure 3, we show the results obtained repeating the experiment using the same mesh $\mathcal{T}_{h_6}$, but smaller time-step size $k = h_6/100$ and tolerance $\varepsilon = h_6/100$. The overall behaviour remains the same. However, we see
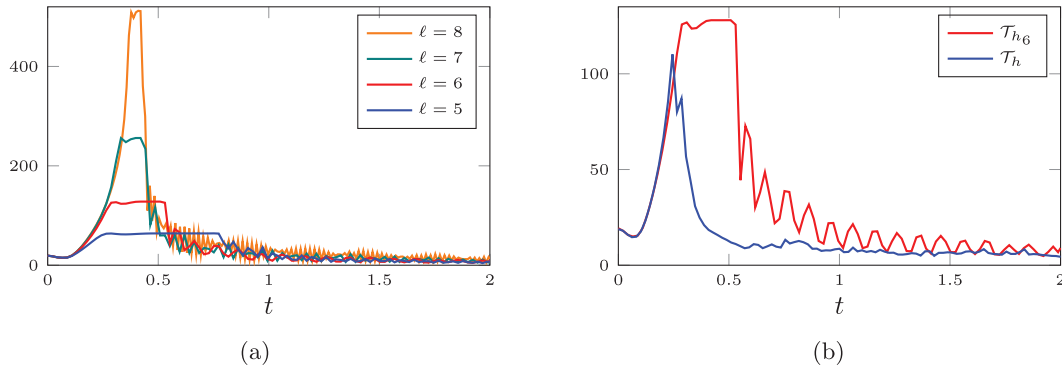
(a)                                    (b)

FIGURE 4. Experiment of Section 4.1: evolution of the $W^{1,\infty}$-seminorm obtained with AMM for different mesh resolutions (a) and different symmetry properties (b). (a) $W^{1,\infty}$-seminorm for $\ell = 5, 6, 7, 8$. (b) $W^{1,\infty}$-seminorm for $\mathcal{T}_{h_6}$ and $\mathcal{T}_h$.

that the numerical dissipation of TPS and the nonmonotonicity of the energy decay of AMM are reduced. This observation confirms the validity of the energy laws established in Proposition 3.6, as the terms responsible for the two above effects can indeed be controlled by time-step size $k$ and the tolerance $\varepsilon$, respectively.

Finally, we investigate whether the resolution or the symmetry of the mesh have an influence on the detection of the blow-up. In Figure 4a, we compare the evolution of the $W^{1,\infty}$-seminorm obtained using the uniform meshes $\mathcal{T}_{h_\ell}$ for $\ell = 5, 6, 7, 8$ ($k = h_\ell/100$ and $\varepsilon = h_\ell/100$). The appearance of blow-up of the three approximations occurs at the same time ($t \approx 0.3$), but the length of the plateaux, *i.e.* the duration of the configuration in which the magnetisation of the origin has the opposite direction of the surrounding vertices, decreases with the mesh size. Moreover, for all $\ell = 5, 6, 7, 8$, the maximum value attained by the $W^{1,\infty}$-seminorm is always the maximum value (4.1) allowed by the discrete space.

In Figure 4b, we compare the evolution of the $W^{1,\infty}$-seminorm computed using $\mathcal{T}_{h_6}$ with the one obtained using an unstructured mesh $\mathcal{T}_h$ of comparable number of elements and mesh size (7234 elements and $h = 0.0262$). We observe that a finite-time blow-up at $t \approx 0.3$ occurs also for the approximation residing in the unstructured mesh $\mathcal{T}_h$. However, the magnetisation configuration with maximum gradient is quickly left (no plateau in the evolution of the $W^{1,\infty}$-seminorm). We believe that the stagnation of the configuration with maximum gradient observed for uniform meshes is a numerical artefact related to their symmetry.

Since the computations performed with TPS lead to the same conclusions, in order not to overload the plots, in Figure 4 we have shown only the results computed using AMM.

It is not clear to the author whether the observed finite-time blow-up also occurs for the weak solution of iLLG towards which the computed approximations converge as $h, k, \varepsilon \to 0$. However, the fact that the phenomenon has been observed for approximations computed using two different schemes and various choices of the discretisation parameters seems to provide a clear evidence in this direction.

## 4.2. Nutation dynamics in ferromagnetic thin films

With this experiment, we aim to illustrate the differences in the magnetisation dynamics induced by the standard LLG (1.1) and iLLG (1.2). Moreover, we compare the performance of TPS and AMM in a simulation with physically relevant geometry and material parameters.

The domain $\Omega$ is a planar thin film of the form $\omega \times (0, c)$ with cross section $\omega \subset \mathbb{R}^2$ (parallel to the $x_1 x_2$-plane) and thickness $c = 3\,\mathrm{nm}$ (aligned with $\boldsymbol{e}_3$). The cross section $\omega$ is an elliptic domain with semiaxis lengths $a = 100\,\mathrm{nm}$ and $b = 50\,\mathrm{nm}$. The axes of the ellipse are parallel to $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$, with the major axis being parallel to $\boldsymbol{e}_1$. It is well known [18, 20, 24] that the magnetisation of thin films is usually homogeneous in the out-of-plane component, so that an adequate description of the energetics of the magnet can be obtained using a 2D energy

functional $\mathcal{E}[\boldsymbol{m}]$ posed only on the cross section $\omega$. Accordingly, we consider the energy functional

$$\mathcal{E}[\boldsymbol{m}]/c = A \int_\omega |\boldsymbol{\nabla}\boldsymbol{m}|^2 + K \int_\omega \left[1 - (\boldsymbol{m} \cdot \boldsymbol{e}_1)^2\right] - \mu_0 M_{\mathrm{s}} \int_\omega \boldsymbol{H}_{\mathrm{ext}} \cdot \boldsymbol{m} + \frac{\mu_0 M_{\mathrm{s}}^2}{2} \int_\omega (\boldsymbol{m} \cdot \boldsymbol{e}_3)^2.$$

Here, $\mu_0 = 4\pi \cdot 10^{-7}\,\mathrm{N\,A}^{-2}$ is the vacuum permeability, $M_{\mathrm{s}}$, $A$, $K$ are positive material parameters (see below), while $\boldsymbol{H}_{\mathrm{ext}}$ denotes an applied magnetic field (in $\mathrm{A\,m}^{-1}$). The four energy contributions in $\mathcal{E}[\boldsymbol{m}]$ are exchange interaction, uniaxial anisotropy (with easy axis $\boldsymbol{e}_1$ parallel to the major axis of the ellipse), Zeeman energy, and magnetostatic interaction, respectively. Note that the nonlocal magnetostatic energy is replaced by a local planar anisotropy contribution penalising out-of-plane magnetisation configurations, which is admissible for magnetic thin films [18, 20, 24]. In (1.1) and (1.2), for the gyromagnetic ratio, we consider the value $\gamma_0 = 2.211 \times 10^5\,\mathrm{m\,A}^{-1}\,\mathrm{s}^{-1}$, while the effective field $\boldsymbol{H}_{\mathrm{eff}}[\boldsymbol{m}]$ is related to the energy $\mathcal{E}[\boldsymbol{m}]$ *via* the relation $\mu_0 M_{\mathrm{s}}\,\boldsymbol{H}_{\mathrm{eff}}[\boldsymbol{m}] = -\frac{\delta \mathcal{E}[\boldsymbol{m}]}{\delta \boldsymbol{m}}$. For the material parameters, we use the values of permalloy (see, *e.g.* [33]): $M_{\mathrm{s}} = 8 \times 10^5\,\mathrm{A\,m}^{-1}$, $A = 1.3 \times 10^{-11}\,\mathrm{J\,m}^{-1}$, $K = 5 \times 10^2\,\mathrm{J\,m}^{-2}$, and $\alpha = 0.023$. For the angular momentum relaxation time $\tau$ in (1.2), we consider the value $\tau = \alpha\,\xi$ with $\xi = 12.3\,\mathrm{ps}$; see [32]. As initial conditions, we consider the constant fields $\boldsymbol{m}^0 \equiv \boldsymbol{e}_1$ and $\boldsymbol{v}^0 \equiv \boldsymbol{0}$. Note that $\boldsymbol{m}^0$ is a global minimum for the energy $\mathcal{E}[\boldsymbol{m}]$ if $\boldsymbol{H}_{\mathrm{ext}} \equiv \boldsymbol{0}$.

The overall simulation time is 30 ps. The experiment consists in perturbing the equilibrium state $\boldsymbol{m}^0 \equiv \boldsymbol{e}_1$ with a perpendicular and spatially uniform high-frequency pulse field $\boldsymbol{H}_{\mathrm{ext}}(t) = F(t) M_{\mathrm{s}}\,\boldsymbol{e}_2$, where $F(t) = 0.01 \sin(2\pi f t)\,\chi_{\{0 \le t \le 2 \times 10^{-12}\}}(t)$ with $f = 500\,\mathrm{GHz}$; see Figure 5a. In order to assess the resulting magnetisation dynamics, we analyse the time evolution of the spatial average of the third magnetisation component $\langle m_3 \rangle$.

For the spatial discretisation we consider a triangular mesh of $\omega$ made of 5998 elements. Its mesh size (3.760 nm) is well below the exchange length of the material $\ell_{\mathrm{ex}} = \sqrt{2A/(\mu_0 M_{\mathrm{s}}^2)} = 5.686\,\mathrm{nm}$. For the time discretisation, we consider three different time-step sizes ($\Delta t = 1, 10, 100\,\mathrm{fs}$). In the stopping criterion (3.13), we use the tolerance $\varepsilon = 1 \times 10^{-6}$.

In Figure 5b, we compare the evolution of $\langle m_3 \rangle$ for LLG and iLLG. We show the results computed using TPS with $\Delta t = 1\,\mathrm{fs}$. Note that TPS for LLG can be obtained from Algorithm 3.1 by omitting the first term on the left-hand side of (3.8); see [2, 17]. The dynamics induced by the two models are completely different. For LLG, the magnetisation reacts to the pulse field and returns straight to the equilibrium state. For iLLG, the deflection from the equilibrium state gives rise to oscillations with approximately the same frequency of the inducing pulse field (500 GHz). Due to damping, the amplitude of the oscillations decays with time and the magnetisation regains the initial equilibrium state. This experiment provides a numerical evidence of the inertial nutation dynamics predicted by the model, which has been experimentally observed only very recently [32].

In Figure 5c, we plot the evolution of $\langle m_3 \rangle$ computed using TPS with $\Delta t = 1, 10, 100\,\mathrm{fs}$. For larger time-step sizes, we observe a faster decay of the oscillations. This phenomenon is a consequence of the artificial damping of the backward Euler method used for the time discretisation. The observed dependence on $\Delta t$ reflects the fact that the artificial damping can be controlled by the time-step size; see (3.15). Finally, in Figure 5d, we show the same plot for AMM. Unlike TPS, AMM is robust with respect to variations of the time-step size. This reflects the energy conservation properties of the symplectic midpoint rule; see (3.16)–(3.17). All the considered time-step sizes are sufficiently small to guarantee the convergence of the fixed-point iteration, which requires 1–2 iterations for $\Delta t = 1, 10\,\mathrm{fs}$ and 2–3 iterations for $\Delta t = 100\,\mathrm{fs}$. Note that using larger time-step sizes is not advisable, as they cannot resolve the pulse field and the resulting magnetisation dynamics.

This experiment shows the importance of designing a numerical scheme which respects the energy law of the underlying model. This general statement, which holds true for any PDE with a physical background (and, in particular, for LLG), is a key aspect for iLLG due to the small extent and the ultrafast time scale of the nutation dynamics. As a consequence, this experiment suggests that for iLLG, between TPS and AMM, AMM should be the method of choice. Indeed, besides ensuring an accurate evolution of the discrete energy, its main drawback, *i.e.* the requirement of smaller time-step sizes to have a well-posed fixed-point iteration, does not represent a restriction, since small time-step sizes are necessary anyway in order to resolve the ultrafast dynamics of the magnetisation.
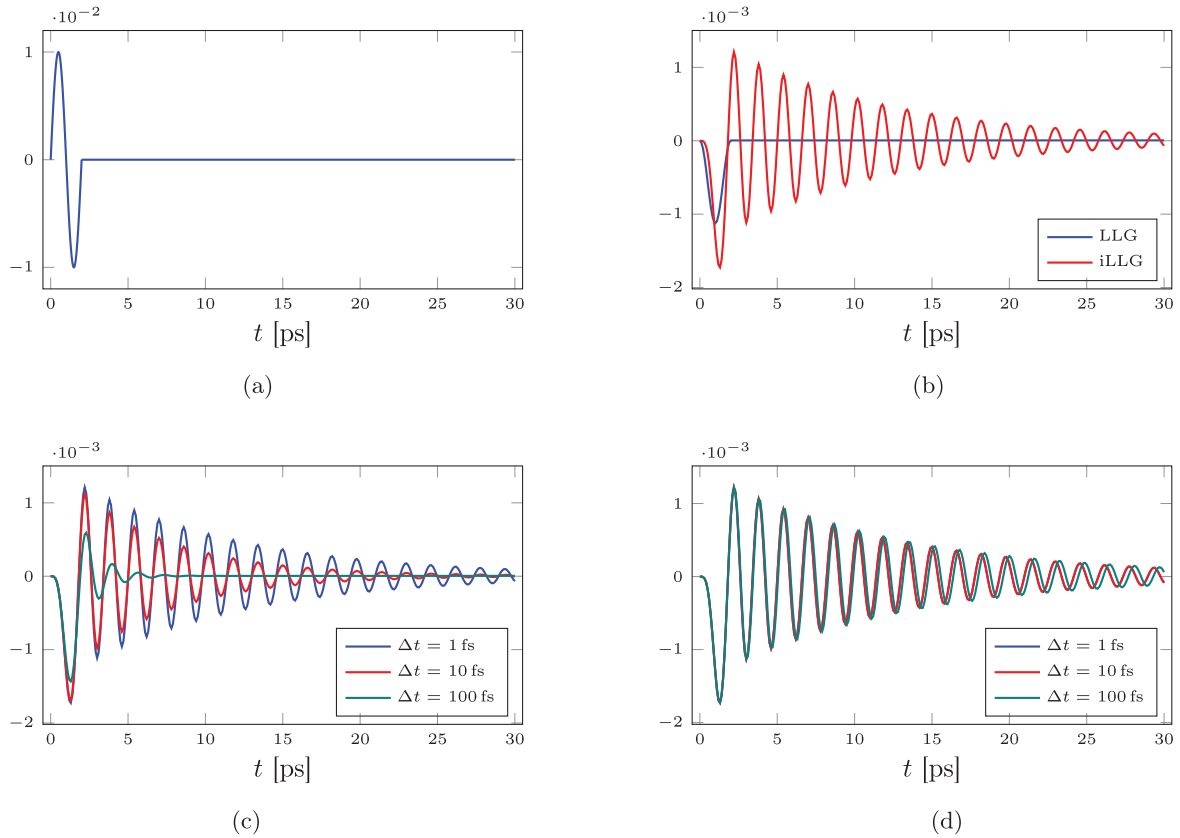
FIGURE 5. Experiment of Section 4.2: (a) Plot of the function $t \mapsto F(t)$ which modulates the amplitude of the pulse field. (b) Evolution of $\langle m_3 \rangle$ for LLG and iLLG computed with TPS for $\Delta t = 1\,\text{fs}$. (c) Evolution of $\langle m_3 \rangle$ computed with TPS for different time-step sizes. (d) Evolution of $\langle m_3 \rangle$ computed with AMM for different time-step sizes.

## 5. PROOFS

In this section, we present the proofs of the results stated in Section 3. In view of their later use, we recall some facts: the mass-lumped product $\langle \cdot, \cdot \rangle_h$ defined in (3.2) is a scalar product on $\mathcal{S}^1(\mathcal{T}_h)^3$ and the induced norm $\|\cdot\|_h$ satisfies the norm equivalence

$$\|\phi_h\|_{\boldsymbol{L}^2(\Omega)} \leq \|\phi_h\|_h \leq \sqrt{d+2}\,\|\phi_h\|_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \phi_h \in \mathcal{S}^1(\mathcal{T}_h)^3.$$

Moreover, there holds the error estimate

$$\left| \langle \phi_h, \psi_h \rangle - \langle \phi_h, \psi_h \rangle_h \right| \leq Ch^2 \|\boldsymbol{\nabla}\phi_h\|_{\boldsymbol{L}^2(\Omega)} \|\boldsymbol{\nabla}\psi_h\|_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \phi_h, \psi_h \in \mathcal{S}^1(\mathcal{T}_h)^3, \tag{5.1}$$

where $C > 0$ depends only on the shape-regularity of $\mathcal{T}_h$; see Lemma 3.9 of [10]. Finally, we recall the following relations between the $L^p$-norm of a discrete function and the $\ell^p$-norm of the vector collecting its nodal values (see [10], Lem. 3.4):

$$\|\phi_h\|_{\boldsymbol{L}^p(\Omega)}^p \simeq \sum_{z \in \mathcal{N}_h} h_z^d |\phi_h(z)|^p \quad \text{and} \quad \|\phi_h\|_{\boldsymbol{L}^\infty(\Omega)} = \max_{z \in \mathcal{N}_h} |\phi_h(z)| \quad \text{for all } \phi_h \in \mathcal{S}^1(\mathcal{T}_h)^3. \tag{5.2}$$

Here, $h_z > 0$ denotes the diameter of the nodal patch associated with $z \in \mathcal{N}_h$.

## 5.1. Properties of Algorithm 3.2 and well-posedness of Algorithm 3.4

We start with proving the conservation properties of Algorithm 3.2.

*Proof of Proposition 3.3.* Let $z \in \mathcal{N}_h$. Choosing $\boldsymbol{\phi}_h = \varphi_z \boldsymbol{m}_h^{i+1/2}(z)$ in (3.9a), we infer that $\left|\boldsymbol{m}_h^{i+1}(z)\right| = \left|\boldsymbol{m}_h^i(z)\right|$. Since $\boldsymbol{m}_h^0 \in \mathcal{M}_h$ by assumption, we conclude that $\boldsymbol{m}_h^{i+1} \in \mathcal{M}_h$.

Choosing $\boldsymbol{\phi}_h = \varphi_z \boldsymbol{w}_h^{i+1/2}(z)$ in (3.9a) and $\boldsymbol{\psi}_h = \varphi_z \boldsymbol{m}_h^{i+1/2}(z)$ in (3.9b), we obtain the identities $d_t \boldsymbol{m}_h^{i+1}(z) \cdot \boldsymbol{w}_h^{i+1/2}(z) = 0$ and $d_t \boldsymbol{w}_h^{i+1}(z) \cdot \boldsymbol{m}_h^{i+1/2}(z)$, respectively. It follows that

$$\boldsymbol{m}_h^{i+1}(z) \cdot \boldsymbol{w}_h^{i+1}(z) - \boldsymbol{m}_h^i(z) \cdot \boldsymbol{w}_h^i(z) = d_t \boldsymbol{m}_h^{i+1}(z) \cdot \boldsymbol{w}_h^{i+1/2}(z) + d_t \boldsymbol{w}_h^{i+1}(z) \cdot \boldsymbol{m}_h^{i+1/2}(z) = 0.$$

Since $\boldsymbol{m}_h^0(z) \cdot \boldsymbol{w}_h^0(z) = \boldsymbol{m}_h^0(z) \cdot (\boldsymbol{m}_h^0(z) \times \boldsymbol{v}_h^0(z)) = 0$, we conclude that $\boldsymbol{w}_h^{i+1} \in \mathcal{K}_h\left[\boldsymbol{m}_h^{i+1}\right]$. $\qquad\square$

Next, we show that the fixed-point iteration designed for the solution of (3.10) is well-posed and converges.

*Proof of Proposition 3.5.* Let $\ell \in \mathbb{N}_0$. The bilinear forms on the left-hand side of both (3.11) and (3.12) are elliptic. Therefore, existence and uniqueness of solutions $\boldsymbol{u}_h^{i,\ell+1}$ and $\boldsymbol{z}_h^{i,\ell+1}$ in $\mathcal{S}^1(\mathcal{T}_h)^3$ follow from the Lax–Milgram theorem.

Let $z \in \mathcal{N}_h$ be an arbitrary vertex. Testing (3.11) with $\boldsymbol{\phi}_h = \varphi_z \boldsymbol{u}_h^{i,\ell+1}(z) \in \mathcal{S}^1(\mathcal{T}_h)^3$, we obtain that $\left|\boldsymbol{u}_h^{i,\ell+1}(z)\right|^2 = \boldsymbol{u}_h^{i,\ell+1}(z) \cdot \boldsymbol{m}_h^i(z)$. Hence, $\left|\boldsymbol{u}_h^{i,\ell+1}(z)\right| \leq \left|\boldsymbol{m}_h^i(z)\right| = 1$. This shows that $\left\|\boldsymbol{u}_h^{i,\ell+1}\right\|_{\boldsymbol{L}^\infty(\Omega)} \leq 1$ and concludes the proof of part (i).

Let $\boldsymbol{u}_h^{i,\ell+1}$ and $\boldsymbol{u}_h^{i,\ell+2}$ (resp., $\boldsymbol{z}_h^{i,\ell+1}$ and $\boldsymbol{z}_h^{i,\ell+2}$) be two consecutive iterates satisfying (3.11) (resp., (3.12)). Taking the difference of the equations satisfied by $\boldsymbol{z}_h^{i,\ell+2}$ and $\boldsymbol{z}_h^{i,\ell+1}$ and choosing $\boldsymbol{\psi}_h = \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}$, we obtain the identity

$$\begin{aligned}
2\tau\left\|\boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\|_h^2 &= k\left\langle \boldsymbol{u}_h^{i,\ell+2} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{u}_h^{i,\ell+2}\right] - \boldsymbol{u}_h^{i,\ell+1} \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{u}_h^{i,\ell+1}\right], \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\rangle_h \\
&\quad + 2\alpha\left\langle \left(\boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1}\right) \times \boldsymbol{m}_h^i, \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\rangle_h \\
&\quad - k\left\langle \boldsymbol{u}_h^{i,\ell+2} \times \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1} \times \boldsymbol{z}_h^{i,\ell+1}, \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\rangle_h.
\end{aligned} \tag{5.3}$$

Taking the difference of the equations satisfied by $\boldsymbol{u}_h^{i,\ell+2}$ and $\boldsymbol{u}_h^{i,\ell+1}$ and choosing the test functions $\boldsymbol{\phi}_h = \boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1}$ and $\boldsymbol{\phi}_h = \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}$, we obtain the identities

$$2\left\|\boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1}\right\|_h^2 = -k\left\langle \boldsymbol{u}_h^{i,\ell+1} \times \left(\boldsymbol{z}_h^{i,\ell+1} - \boldsymbol{z}_h^{i,\ell}\right), \boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1}\right\rangle_h, \tag{5.4}$$

$$2\left\langle \boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1}, \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\rangle_h + k\left\langle \boldsymbol{u}_h^{i,\ell+2} \times \boldsymbol{z}_h^{i,\ell+1} - \boldsymbol{u}_h^{i,\ell+1} \times \boldsymbol{z}_h^{i,\ell}, \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\rangle_h = 0. \tag{5.5}$$

From (5.4), since $\left\|\boldsymbol{u}_h^{i,\ell+1}\right\|_{\boldsymbol{L}^\infty(\Omega)} \leq 1$ from part (i), we deduce that

$$\left\|\boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1}\right\|_h \leq \frac{k}{2}\left\|\boldsymbol{z}_h^{i,\ell+1} - \boldsymbol{z}_h^{i,\ell}\right\|_h. \tag{5.6}$$

Combining (5.3) and (5.5), we obtain that

$$\begin{aligned}
2\tau\left\|\boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\|_h^2 &= k\left\langle \left(\boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1}\right) \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{u}_h^{i,\ell+2}\right], \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\rangle_h \\
&\quad + k\left\langle \boldsymbol{u}_h^{i,\ell+1} \times \left(\mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{u}_h^{i,\ell+2}\right] - \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{u}_h^{i,\ell+1}\right]\right), \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\rangle_h \\
&\quad + 2\alpha\left\langle \left(\boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1}\right) \times \boldsymbol{m}_h^i, \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\rangle_h \\
&\quad + 2\left\langle \boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1}, \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\rangle_h + k\left\langle \boldsymbol{u}_h^{i,\ell+1} \times \left(\boldsymbol{z}_h^{i,\ell+1} - \boldsymbol{z}_h^{i,\ell}\right), \boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\rangle_h.
\end{aligned}$$

It follows that

$$2\tau\left\|\boldsymbol{z}_h^{i,\ell+2} - \boldsymbol{z}_h^{i,\ell+1}\right\|_h \leq \left(2 + 2\alpha\left\|\boldsymbol{m}_h^i\right\|_{\boldsymbol{L}^\infty(\Omega)} + k\left\|\mathbb{P}_h\boldsymbol{h}_{\text{eff}}\left[\boldsymbol{u}_h^{i,\ell+2}\right]\right\|_{\boldsymbol{L}^\infty(\Omega)}\right)\left\|\boldsymbol{u}_h^{i,\ell+2} - \boldsymbol{u}_h^{i,\ell+1}\right\|_h + k\left\|\boldsymbol{u}_h^{i,\ell+1}\right\|_{\boldsymbol{L}^\infty(\Omega)}$$
$$\times\left\|\mathbb{P}_h\boldsymbol{h}_{\text{eff}}\left[\boldsymbol{u}_h^{i,\ell+2}\right] - \mathbb{P}_h\boldsymbol{h}_{\text{eff}}\left[\boldsymbol{u}_h^{i,\ell+1}\right]\right\|_h + k\left\|\boldsymbol{u}_h^{i,\ell+1}\right\|_{\boldsymbol{L}^\infty(\Omega)}\left\|\boldsymbol{z}_h^{i,\ell+1} - \boldsymbol{z}_h^{i,\ell}\right\|_h.$$

Using that $\left\|\boldsymbol{m}_h^i\right\|_{\boldsymbol{L}^\infty} = 1$ and $\left\|\boldsymbol{u}_h^{i,\ell+1}\right\|_{\boldsymbol{L}^\infty} \leq 1$, together with the estimates [9,12]

$$\left\|\mathbb{P}_h\boldsymbol{h}_{\text{eff}}\left[\boldsymbol{u}_h^{i,\ell+2}\right]\right\|_{\boldsymbol{L}^\infty} \leq C_{\text{inv}}^2 h_{\min}^{-2}\left\|\boldsymbol{u}_h^{i,\ell+2}\right\|_{\boldsymbol{L}^\infty},$$
$$\left\|\mathbb{P}_h\boldsymbol{h}_{\text{eff}}\left[\boldsymbol{u}_h^{i,\ell+2}\right]\right\|_h \leq C_{\text{inv}}^2 h_{\min}^{-2}\left\|\boldsymbol{u}_h^{i,\ell+2}\right\|_h$$

(where $C_{\text{inv}} > 0$ depends only on the shape-regularity of $\mathcal{T}_h$) and (5.6), we obtain (3.14) with $q = \left[(2 + \alpha + \tau)k + C_{\text{inv}}^2 k^2 h_{\min}^{-2}\right]/(2\tau)$. Hence, if $k < k_0 := \tau/(2 + \alpha + \tau)$ and $k < Ch_{\min}$ (with $C = \sqrt{\tau}/C_{\text{inv}}$), then $0 < q < 1$. This proves part (ii).

By construction, $\boldsymbol{m}_h^{i+1} = 2\boldsymbol{u}_h^{i,\ell_i+1} - \boldsymbol{m}_h^i$ satisfies

$$\left\langle d_t\boldsymbol{m}_h^{i+1}, \boldsymbol{\phi}_h\right\rangle_h = \left\langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{z}_h^{i,\ell}, \boldsymbol{\phi}_h\right\rangle_h \quad \text{for all } \boldsymbol{\phi}_h \in \mathcal{S}^1(\mathcal{T}_h)^3,$$

while $\boldsymbol{w}_h^{i+1} = 2\boldsymbol{z}_h^{i,\ell_i+1} - \boldsymbol{w}_h^i$ satisfies (3.9b). The argument used to prove Proposition 3.3 shows that $\boldsymbol{m}_h^{i+1} \in \mathcal{M}_h$ and $\boldsymbol{w}_h^{i+1} \in \mathcal{K}_h\left[\boldsymbol{m}_h^{i+1}\right]$. This shows part (iii) and concludes the proof. $\qquad\square$

## 5.2. Discrete energy laws and stability

We prove the discrete energy laws satisfied by the approximations generated by the algorithms.

*Proof of Proposition 3.6.* Let $i \in \mathbb{N}_0$. We test (3.8) with $\boldsymbol{\phi}_h = \boldsymbol{v}_h^{i+1} \in \mathcal{K}_h\left[\boldsymbol{m}_h^i\right]$ and multiply the resulting equation by $k$. We obtain the identity

$$\tau\left\langle \boldsymbol{v}_h^{i+1} - \boldsymbol{v}_h^i, \boldsymbol{v}_h^{i+1}\right\rangle_h + \alpha k\left\|\boldsymbol{v}_h^{i+1}\right\|_h^2 + k^2\left\|\boldsymbol{\nabla}\boldsymbol{v}_h^{i+1}\right\|_{\boldsymbol{L}^2(\Omega)}^2 = -k\left\langle\boldsymbol{\nabla}\boldsymbol{m}_h^i, \boldsymbol{\nabla}\boldsymbol{v}_h^{i+1}\right\rangle$$

Since the angle condition (3.4) is satisfied, Lemma 3.2 of [7] yields that

$$\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^{i+1}\right\|_{\boldsymbol{L}^2(\Omega)} \leq \left\|\boldsymbol{\nabla}\boldsymbol{m}_h^i + k\boldsymbol{\nabla}\boldsymbol{v}_h^{i+1}\right\|_{\boldsymbol{L}^2(\Omega)}. \tag{5.7}$$

Hence, we obtain that

$$\mathcal{E}\left[\boldsymbol{m}_h^{i+1}\right] \overset{(5.7)}{\leq} \frac{1}{2}\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^i + k\boldsymbol{\nabla}\boldsymbol{v}_h^{i+1}\right\|_{\boldsymbol{L}^2(\Omega)}^2 = \mathcal{E}\left[\boldsymbol{m}_h^i\right] + k\left\langle\boldsymbol{\nabla}\boldsymbol{m}_h^i, \boldsymbol{\nabla}\boldsymbol{v}_h^{i+1}\right\rangle + \frac{k^2}{2}\left\|\boldsymbol{\nabla}\boldsymbol{v}_h^{i+1}\right\|_{\boldsymbol{L}^2(\Omega)}^2.$$

Altogether, we obtain that

$$\mathcal{E}\left[\boldsymbol{m}_h^{i+1}\right] + \tau\left\langle\boldsymbol{v}_h^{i+1} - \boldsymbol{v}_h^i, \boldsymbol{v}_h^{i+1}\right\rangle_h + \alpha k\left\|\boldsymbol{v}_h^{i+1}\right\|_h^2 + \frac{k^2}{2}\left\|\boldsymbol{\nabla}\boldsymbol{v}_h^{i+1}\right\|_{\boldsymbol{L}^2(\Omega)}^2 \leq \mathcal{E}\left[\boldsymbol{m}_h^i\right].$$

Applying the vector identity

$$(\mathbf{a} - \mathbf{b}) \cdot \mathbf{a} = \frac{1}{2}|\mathbf{a}|^2 - \frac{1}{2}|\mathbf{b}|^2 + \frac{1}{2}|\mathbf{a} - \mathbf{b}|^2 \quad \text{for all } \mathbf{a}, \mathbf{b} \in \mathbb{R}^3$$

to the second term on the left-hand side yields (3.15). This proves part (i).

To show the stability of Algorithm 3.2, we choose $\phi_h = d_t \boldsymbol{m}_h^{i+1}$ in (3.9a), $\phi_h = \mathbb{P}_h \boldsymbol{h}_{\text{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right]$ in (3.9a), and $\boldsymbol{\psi}_h = \boldsymbol{w}_h^{i+1/2}$ in (3.9b) to obtain the identities

$$\left\|d_t \boldsymbol{m}_h^{i+1}\right\|_h^2 = -\left\langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{w}_h^{i+1/2}, d_t \boldsymbol{m}_h^{i+1}\right\rangle_h,$$

$$\left\langle d_t \boldsymbol{m}_h^{i+1}, \mathbb{P}_h \boldsymbol{h}_{\text{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right]\right\rangle_h = -\left\langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{w}_h^{i+1/2}, \mathbb{P}_h \boldsymbol{h}_{\text{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right]\right\rangle_h,$$

$$\tau\left\langle d_t \boldsymbol{w}_h^{i+1}, \boldsymbol{w}_h^{i+1/2}\right\rangle_h = \left\langle \boldsymbol{m}_h^{i+1/2} \times \mathbb{P}_h \boldsymbol{h}_{\text{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right], \boldsymbol{w}_h^{i+1/2}\right\rangle_h - \alpha\left\langle \boldsymbol{m}_h^{i+1/2} \times d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{w}_h^{i+1/2}\right\rangle_h,$$

respectively. Combining these three equations, we obtain that

$$\left\langle d_t \boldsymbol{m}_h^{i+1}, \mathbb{P}_h \boldsymbol{h}_{\text{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right]\right\rangle_h = \tau\left\langle d_t \boldsymbol{w}_h^{i+1}, \boldsymbol{w}_h^{i+1/2}\right\rangle_h + \alpha\left\|d_t \boldsymbol{m}_h^{i+1}\right\|_h^2.$$

Since

$$\left\langle \mathbb{P}_h \boldsymbol{h}_{\text{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right], d_t \boldsymbol{m}_h^{i+1}\right\rangle_h \overset{(3.3),(2.3),(2.1)}{=} -\frac{1}{k}\left(\mathcal{E}\left[\boldsymbol{m}_h^{i+1}\right] - \mathcal{E}\left[\boldsymbol{m}_h^i\right]\right),$$

$$\left\langle d_t \boldsymbol{w}_h^{i+1}, \boldsymbol{w}_h^{i+1/2}\right\rangle_h = \frac{1}{2k}\left(\left\|\boldsymbol{w}_h^{i+1}\right\|_h^2 - \left\|\boldsymbol{w}_h^i\right\|_h^2\right),$$

we obtain (3.16). This proves part (ii). The proof of (3.17) from part (iii) can be obtained with the very same argument. $\square$

### 5.3. Convergence results

First, we prove the convergence of Algorithm 3.1.

*Proof of Theorem 3.7 for Algorithm 3.1.* The proof is largely based on the argument of [2,17,26]. We start with recalling the estimates

$$\left|\boldsymbol{m}_h^{i+1}(z) - \boldsymbol{m}_h^i(z)\right| \le k\left|\boldsymbol{v}_h^{i+1}(z)\right| \quad \text{and} \quad \left|\boldsymbol{m}_h^{i+1}(z) - \boldsymbol{m}_h^i(z) - k\boldsymbol{v}_h^{i+1}(z)\right| \le \frac{k^2}{2}\left|\boldsymbol{v}_h^{i+1}(z)\right|^2, \qquad (5.8)$$

which hold for all $i \in \mathbb{N}_0$ and $z \in \mathcal{N}_h$; see [3,14]. The first inequality in (5.8), together with (5.2), yields that $\left\|d_t \boldsymbol{m}_h^{i+1}\right\|_h \le \left\|\boldsymbol{v}_h^{i+1}\right\|_h$ for all $i \in \mathbb{N}_0$.

Let $T > 0$ be arbitrary. With the uniform boundedness guaranteed by (3.15) and (3.18), we can construct $\boldsymbol{m} \in L^\infty(0, \infty; \boldsymbol{H}^1(\Omega)) \cap W^{1,\infty}(0, \infty; \boldsymbol{L}^2(\Omega))$, satisfying $|\boldsymbol{m}| = 1$ a.e. in $\Omega \times (0, \infty)$, such that, upon extraction of (nonrelabeled) subsequences, there hold the convergences $\boldsymbol{m}_{hk}, \boldsymbol{m}_{hk}^\pm \overset{*}{\rightharpoonup} \boldsymbol{m}$ in $L^\infty(0, \infty; \boldsymbol{H}^1(\Omega))$, $\boldsymbol{m}_{hk}|_{\Omega_T} \rightharpoonup \boldsymbol{m}|_{\Omega_T}$ in $\boldsymbol{H}^1(\Omega_T)$, $\partial_t \boldsymbol{m}_{hk}, \boldsymbol{v}_{hk}^+ \overset{*}{\rightharpoonup} \partial_t \boldsymbol{m}$ in $L^\infty(0, \infty; \boldsymbol{L}^2(\Omega))$, as well as $\boldsymbol{m}_{hk}, \boldsymbol{m}_{hk}^\pm \to \boldsymbol{m}$ in $\boldsymbol{L}^2(\Omega \times (0, \infty))$ and pointwise almost everywhere in $\Omega \times (0, \infty)$. Moreover, we have that $k\boldsymbol{\nabla}\boldsymbol{v}_{hk}^+ \to 0$ in $L^2(0, \infty; \boldsymbol{L}^2(\Omega))$ as $h, k \to 0$.

Let $\boldsymbol{\varphi} \in C_c^\infty([0, T); \boldsymbol{C}(\overline{\Omega}))$ be an arbitrary smooth test function. Let $N \in \mathbb{N}$ be the smallest integer such that $T \le kN = t_N$. Let $i \in \{0, \dots, N-1\}$. We choose the test function $\phi_h = \mathcal{I}_h[\boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i)] \in \mathcal{K}_h\left[\boldsymbol{m}_h^i\right]$ in (3.8) to obtain

$$\tau\left\langle d_t \boldsymbol{v}_h^{i+1}, \mathcal{I}_h[\boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i)]\right\rangle_h + \alpha\left\langle \boldsymbol{v}_h^{i+1}, \mathcal{I}_h[\boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i)]\right\rangle_h + \left\langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^{i+1}, \mathcal{I}_h[\boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i)]\right\rangle_h$$
$$- k\left\langle \mathbb{P}_h \boldsymbol{h}_{\text{eff}}\left[\boldsymbol{v}_h^{i+1}\right], \mathcal{I}_h[\boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i)]\right\rangle_h = \left\langle \mathbb{P}_h \boldsymbol{h}_{\text{eff}}\left[\boldsymbol{m}_h^i\right], \mathcal{I}_h[\boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i)]\right\rangle_h,$$

where we extend $\boldsymbol{\varphi}$ by zero in $(T, t_N)$. Due to the presence of the mass-lumped scalar product, we can remove the nodal interpolant from the first three terms on the left-hand side without affecting the value of the integrals.

Then, multiplying the latter by $k$, summing over $i = 0, \ldots, N-1$, and using (2.3) and (3.3), we obtain the identity

$$\tau k \sum_{i=0}^{N-1} \left\langle d_t \boldsymbol{v}_h^{i+1}, \boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i) \right\rangle_h + \alpha k \sum_{i=0}^{N-1} \left\langle \boldsymbol{v}_h^{i+1}, \boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i) \right\rangle_h + k \sum_{i=0}^{N-1} \left\langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^{i+1}, \boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i) \right\rangle_h$$

$$= -k \sum_{i=0}^{N-1} \left\langle \boldsymbol{\nabla} \left( \boldsymbol{m}_h^i + k \boldsymbol{v}_h^{i+1} \right), \boldsymbol{\nabla} \boldsymbol{\mathcal{I}}_h [\boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i)] \right\rangle.$$

Next, we rewrite the first term on the left-hand side using the summation by parts formula

$$\sum_{i=0}^{N-1} (a_{i+1} - a_i) b_i = -\sum_{i=0}^{N-1} a_{i+1}(b_{i+1} - b_i) + a_N b_N - a_0 b_0 \quad \text{for all sequences } \{a_i\}_{i=0}^N, \{b_i\}_{i=0}^N \tag{5.9}$$

and performing some algebraic manipulations:

$$\tau k \sum_{i=0}^{N-1} \left\langle d_t \boldsymbol{v}_h^{i+1}, \boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i) \right\rangle_h = -\tau \sum_{i=0}^{N-1} \left\langle \boldsymbol{v}_h^{i+1}, \boldsymbol{m}_h^{i+1} \times \boldsymbol{\varphi}(t_{i+1}) - \boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i) \right\rangle_h$$

$$+ \tau \left\langle \boldsymbol{v}_h^N, \boldsymbol{m}_h^N \times \boldsymbol{\varphi}(t_N) \right\rangle_h - \tau \left\langle \boldsymbol{v}_h^0, \boldsymbol{m}_h^0 \times \boldsymbol{\varphi}(0) \right\rangle_h$$

$$= -\tau k \sum_{i=0}^{N-1} \left\langle \boldsymbol{v}_h^{i+1}, d_t \boldsymbol{m}_h^{i+1} \times \boldsymbol{\varphi}(t_{i+1}) \right\rangle_h - \tau k \sum_{i=0}^{N-1} \left\langle \boldsymbol{v}_h^{i+1}, \boldsymbol{m}_h^i \times d_t \boldsymbol{\varphi}(t_{i+1}) \right\rangle_h$$

$$+ \tau \left\langle \boldsymbol{v}_h^N, \boldsymbol{m}_h^N \times \boldsymbol{\varphi}(t_N) \right\rangle_h - \tau \left\langle \boldsymbol{v}_h^0, \boldsymbol{m}_h^0 \times \boldsymbol{\varphi}(0) \right\rangle_h$$

$$= -\tau k \sum_{i=0}^{N-1} \left\langle \boldsymbol{v}_h^{i+1}, (d_t \boldsymbol{m}_h^{i+1} - \boldsymbol{v}_h^{i+1}) \times \boldsymbol{\varphi}(t_{i+1}) \right\rangle_h + \tau k \sum_{i=0}^{N-1} \left\langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^{i+1}, d_t \boldsymbol{\varphi}(t_{i+1}) \right\rangle_h$$

$$- \tau \left\langle \boldsymbol{m}_h^N \times \boldsymbol{v}_h^N, \boldsymbol{\varphi}(t_N) \right\rangle_h + \tau \left\langle \boldsymbol{m}_h^0 \times \boldsymbol{v}_h^0, \boldsymbol{\varphi}(0) \right\rangle_h.$$

Using Hölder inequality, a combination of the second inequality in (5.8) and the norm equivalence (5.2), and inverse estimates (see, *e.g.* [10], Lem. 3.5), the first term on the right-hand side can be estimated as

$$\tau k \sum_{i=0}^{N-1} \left\langle \boldsymbol{v}_h^{i+1}, \left( d_t \boldsymbol{m}_h^{i+1} - \boldsymbol{v}_h^{i+1} \right) \times \boldsymbol{\varphi}(t_{i+1}) \right\rangle_h \lesssim k \sum_{i=0}^{N-1} \left\| \boldsymbol{v}_h^{i+1} \right\|_{\boldsymbol{L}^3(\Omega)} \left\| d_t \boldsymbol{m}_h^{i+1} - \boldsymbol{v}_h^{i+1} \right\|_{\boldsymbol{L}^{3/2}(\Omega)} \left\| \boldsymbol{\varphi}(t_{i+1}) \right\|_{\boldsymbol{L}^\infty(\Omega)}$$

$$\lesssim k^2 \sum_{i=0}^{N-1} \left\| \boldsymbol{v}_h^{i+1} \right\|_{\boldsymbol{L}^3(\Omega)}^3 \lesssim k^2 h_{\min}^{-d/2} \sum_{i=0}^{N-1} \left\| \boldsymbol{v}_h^{i+1} \right\|_{\boldsymbol{L}^2(\Omega)}^3 \lesssim k h_{\min}^{-d/2}.$$

Altogether, using also the identity

$$\left[ \boldsymbol{m}_h^i(z) \times \boldsymbol{v}_h^{i+1}(z) \right] \cdot \left[ \boldsymbol{m}_h^i(z) \times \boldsymbol{\varphi}(z, t_i) \right] = \boldsymbol{v}_h^{i+1}(z) \cdot \boldsymbol{\varphi}(z, t_i) \quad \text{for all } z \in \mathcal{N}_h$$

(which follows from (3.6), $\boldsymbol{m}_h^i \in \boldsymbol{\mathcal{M}}_h$ and $\boldsymbol{v}_h^{i+1} \in \boldsymbol{\mathcal{K}}_h[\boldsymbol{m}_h^i]$) and observing that $\boldsymbol{\varphi}(t_N) = \boldsymbol{0}$, we thus obtain that

$$k \sum_{i=0}^{N-1} \left\langle \boldsymbol{v}_h^{i+1}, \boldsymbol{\varphi}(t_i) \right\rangle_h = -k \sum_{i=0}^{N-1} \left\langle \boldsymbol{\nabla} \left( \boldsymbol{m}_h^i + k \boldsymbol{v}_h^{i+1} \right), \boldsymbol{\nabla} \boldsymbol{\mathcal{I}}_h [\boldsymbol{m}_h^i \times \boldsymbol{\varphi}(t_i)] \right\rangle + \alpha k \sum_{i=0}^{N-1} \left\langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^{i+1}, \boldsymbol{\varphi}(t_i) \right\rangle_h$$

$$- \tau k \sum_{i=0}^{N-1} \left\langle \boldsymbol{m}_h^i \times \boldsymbol{v}_h^{i+1}, d_t \boldsymbol{\varphi}(t_{i+1}) \right\rangle_h - \tau \left\langle \boldsymbol{m}_h^0 \times \boldsymbol{v}_h^0, \boldsymbol{\varphi}(0) \right\rangle_h + o(1).$$

Using the approximation properties of the nodal interpolant and estimate (5.1) (see the argument of [2, 12]), we can replace all mass-lumped scalar products by $L^2$-products and remove the nodal interpolant from the first term on the left-hand side at the price of an error which goes to zero in the limit. Rewriting the space-time integrals of the resulting equation in terms of the time reconstructions (3.1), we obtain that

$$\int_0^T \langle \boldsymbol{v}_{hk}^+(t), \boldsymbol{\varphi}_k^-(t) \rangle \mathrm{d}t$$

$$= -\int_0^T \langle \boldsymbol{\nabla}(\boldsymbol{m}_{hk}^-(t) + k\boldsymbol{v}_{hk}^+(t)), \boldsymbol{\nabla}(\boldsymbol{m}_{hk}^-(t) \times \boldsymbol{\varphi}_k^-(t)) \rangle \mathrm{d}t + \alpha \int_0^T \langle \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{v}_{hk}^+(t), \boldsymbol{\varphi}_k^-(t) \rangle \mathrm{d}t$$

$$- \tau \int_0^T \langle \boldsymbol{m}_{hk}^-(t) \times \boldsymbol{v}_{hk}^+(t), \partial_t \boldsymbol{\varphi}_k(t) \rangle \mathrm{d}t - \tau \langle \boldsymbol{m}_h^0 \times \boldsymbol{v}_h^0, \boldsymbol{\varphi}(0) \rangle + o(1).$$

Using the available convergence results, we can pass to the limit as $h, k \to 0$ the latter and obtain that each term converges towards the corresponding one in (2.9). By density, it follows that $\boldsymbol{m}$ satisfies (2.9) for all $\boldsymbol{\varphi} \in C_c^\infty([0, T); \boldsymbol{H}^1(\Omega))$. This shows that $\boldsymbol{m}$ satisfies part (iii) of Definition 2.1.

The proof that $\boldsymbol{m}$ attains the prescribed initial data $(\boldsymbol{m}^0, \boldsymbol{v}^0)$ continuously in $\boldsymbol{H}^1(\Omega) \times \boldsymbol{L}^2(\Omega)$ (part (ii) of Def. 2.1) follows the argument of [13], page 72. The energy inequality (2.10) (part (iv) of Def. 2.1) can be obtained from the discrete energy law (3.15) using the available convergence results and standard lower semicontinuity arguments. □

Next, we prove the convergence result for the nonlinear angular momentum method.

*Proof of Theorem 3.7 for Algorithm 3.2.* The proof follows the ideas of [9, 12, 27]. Testing (3.9a) with $\boldsymbol{\phi}_h = d_t \boldsymbol{m}_h^{i+1}$, we infer that $\left\| d_t \boldsymbol{m}_h^{i+1} \right\|_h \leq \left\| \boldsymbol{w}_h^{i+1/2} \right\|_h$ for all $i \in \mathbb{N}_0$.

Let $T > 0$ be arbitrary. With the uniform boundedness guaranteed by (3.16) and (3.18), we can construct $\boldsymbol{m} \in L^\infty(0, \infty; \boldsymbol{H}^1(\Omega)) \cap W^{1,\infty}(0, \infty; \boldsymbol{L}^2(\Omega))$ and $\boldsymbol{w} \in L^\infty(0, \infty; \boldsymbol{L}^2(\Omega))$ such that, upon extraction of a (nonrelabeled) subsequence, we have the convergences $\boldsymbol{m}_{hk}, \boldsymbol{m}_{hk}^+, \overline{\boldsymbol{m}}_{hk} \overset{*}{\rightharpoonup} \boldsymbol{m}$ in $L^\infty(0, \infty; \boldsymbol{H}^1(\Omega))$, $\partial_t \boldsymbol{m}_{hk} \overset{*}{\rightharpoonup} \partial_t \boldsymbol{m}$ in $L^\infty(0, \infty; \boldsymbol{L}^2(\Omega))$, $\boldsymbol{m}_{hk}|_{\Omega_T} \rightharpoonup \boldsymbol{m}|_{\Omega_T}$ in $\boldsymbol{H}^1(\Omega_T)$, $\boldsymbol{m}_{hk}, \boldsymbol{m}_{hk}^+, \overline{\boldsymbol{m}}_{hk} \to \boldsymbol{m}$ in $\boldsymbol{L}^2(\Omega \times (0, \infty))$ and pointwise almost everywhere in $\Omega \times (0, \infty)$, as well as $\overline{\boldsymbol{w}}_{hk}, \boldsymbol{w}_{hk}^+ \overset{*}{\rightharpoonup} \boldsymbol{w}$ in $L^\infty(0, \infty; \boldsymbol{L}^2(\Omega))$. Moreover, it holds that $|\boldsymbol{m}| = 1$ and $\boldsymbol{m} \cdot \boldsymbol{w} = 0$ a.e. in $\Omega \times (0, \infty)$.

Let $\boldsymbol{\zeta}, \boldsymbol{\varphi} \in C_c^\infty([0, T); \boldsymbol{C}(\overline{\Omega}))$ be arbitrary smooth test functions. Let $N \in \mathbb{N}$ be the smallest integer such that $T \leq kN = t_N$. Let $i \in \{0, \ldots, N-1\}$. We choose the test function $\boldsymbol{\phi}_h = \mathcal{I}_h[\boldsymbol{\zeta}(t_i)] \in \mathcal{S}^1(\mathcal{T}_h)^3$ in (3.9a) and $\boldsymbol{\psi}_h = \mathcal{I}_h[\boldsymbol{\varphi}(t_i)] \in \mathcal{S}^1(\mathcal{T}_h)^3$ in (3.9b). Multiplying the resulting equation by $k$, summing over $i = 0, \ldots, N-1$, and using (2.3) and (3.3) on the term which involves the effective field, we obtain the identities

$$k \sum_{i=0}^{N-1} \langle d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{\zeta}(t_i) \rangle_h = -k \sum_{i=0}^{N-1} \langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{w}_h^{i+1/2}, \boldsymbol{\zeta}(t_i) \rangle_h,$$

$$\tau k \sum_{i=0}^{N-1} \langle d_t \boldsymbol{w}_h^{i+1}, \boldsymbol{\varphi}(t_i) \rangle_h = -k \sum_{i=0}^{N-1} \langle \boldsymbol{\nabla} \boldsymbol{m}_h^{i+1/2}, \boldsymbol{\nabla}\left( \boldsymbol{\varphi}(t_i) \times \boldsymbol{m}_h^{i+1/2} \right) \rangle$$

$$- \alpha k \sum_{i=0}^{N-1} \langle \boldsymbol{m}_h^{i+1/2} \times d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{\varphi}(t_i) \rangle_h - k \sum_{i=0}^{N-1} \langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{w}_h^{i+1/2}, \boldsymbol{\varphi}(t_i) \rangle_h,$$

where we extend $\boldsymbol{\zeta}, \boldsymbol{\varphi}$ by zero in $(T, t_N)$. Next, we rewrite the first term on the left-hand side of both equations using the summation by parts formula (5.9):

$$-k \sum_{i=0}^{N-1} \langle \boldsymbol{m}_h^{i+1}, d_t \boldsymbol{\zeta}(t_{i+1}) \rangle_h - \langle \boldsymbol{m}_h^0, \boldsymbol{\zeta}(0) \rangle_h = -k \sum_{i=0}^{N-1} \langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{w}_h^{i+1/2}, \boldsymbol{\zeta}(t_i) \rangle_h,$$

$$-\tau k \sum_{i=0}^{N-1} \left\langle \boldsymbol{w}_h^{i+1}, d_t\boldsymbol{\varphi}(t_{i+1})\right\rangle_h - \tau\left\langle \boldsymbol{w}_h^0, \boldsymbol{\varphi}(0)\right\rangle_h = -k\sum_{i=0}^{N-1}\left\langle \boldsymbol{\nabla m}_h^{i+1/2}, \boldsymbol{\nabla}\left(\boldsymbol{\varphi}(t_i) \times \boldsymbol{m}_h^{i+1/2}\right)\right\rangle$$

$$-\alpha k \sum_{i=0}^{N-1}\left\langle \boldsymbol{m}_h^{i+1/2} \times d_t\boldsymbol{m}_h^{i+1}, \boldsymbol{\varphi}(t_i)\right\rangle_h - k\sum_{i=0}^{N-1}\left\langle \boldsymbol{m}_h^{i+1/2} \times \boldsymbol{w}_h^{i+1/2}, \boldsymbol{\varphi}(t_i)\right\rangle_h.$$

Rewriting the space-time integrals of both equations in terms of the time reconstructions defined in (3.1), we obtain that

$$-\int_0^T \left\langle \boldsymbol{m}_{hk}^+(t), \partial_t\boldsymbol{\zeta}_k(t)\right\rangle_h \mathrm{d}t - \left\langle \boldsymbol{m}_h^0, \boldsymbol{\zeta}(0)\right\rangle_h = -\int_0^T \left\langle \overline{\boldsymbol{m}}_{hk}(t) \times \overline{\boldsymbol{w}}_{hk}(t), \boldsymbol{\zeta}_k^-(t)\right\rangle_h \mathrm{d}t,$$

$$-\tau\int_0^T \left\langle \boldsymbol{w}_{hk}^+(t), \partial_t\boldsymbol{\varphi}_k(t)\right\rangle_h \mathrm{d}t - \tau\left\langle \boldsymbol{w}_h^0, \boldsymbol{\varphi}(0)\right\rangle_h = -\int_0^T \left\langle \boldsymbol{\nabla}\overline{\boldsymbol{m}}_{hk}(t), \boldsymbol{\nabla}\left(\boldsymbol{\varphi}_k^-(t) \times \overline{\boldsymbol{m}}_{hk}(t)\right)\right\rangle \mathrm{d}t$$

$$-\alpha\int_0^T \left\langle \overline{\boldsymbol{m}}_{hk}(t) \times \partial_t\boldsymbol{m}_{hk}(t), \boldsymbol{\varphi}_k^-(t)\right\rangle_h \mathrm{d}t - \int_0^T \left\langle \overline{\boldsymbol{m}}_{hk}(t) \times \overline{\boldsymbol{w}}_{hk}(t), \boldsymbol{\varphi}_k^-(t)\right\rangle_h \mathrm{d}t.$$

Using the available convergence results and (5.1), we can proceed as in Section 3 of [12] and pass the latter equations to the limit as $h, k \to 0$. Rearranging the terms, we obtain that

$$-\int_0^T \left\langle \boldsymbol{m}(t), \partial_t\boldsymbol{\zeta}(t)\right\rangle \mathrm{d}t = -\int_0^T \left\langle \boldsymbol{m}(t) \times \boldsymbol{w}(t), \boldsymbol{\zeta}(t)\right\rangle \mathrm{d}t + \left\langle \boldsymbol{m}^0, \boldsymbol{\zeta}(0)\right\rangle, \tag{5.10}$$

$$-\int_0^T \left\langle \boldsymbol{m}(t) \times \boldsymbol{w}(t), \boldsymbol{\varphi}(t)\right\rangle \mathrm{d}t = -\int_0^T \left\langle \boldsymbol{\nabla m}(t), \boldsymbol{\nabla}(\boldsymbol{\varphi}(t) \times \boldsymbol{m}(t))\right\rangle \mathrm{d}t$$

$$+\alpha\int_0^T \left\langle \boldsymbol{m}(t) \times \partial_t\boldsymbol{m}(t), \boldsymbol{\varphi}(t)\right\rangle \mathrm{d}t - \tau\int_0^T \left\langle \boldsymbol{w}(t), \partial_t\boldsymbol{\varphi}(t)\right\rangle \mathrm{d}t - \tau\left\langle \boldsymbol{m}^0 \times \boldsymbol{v}^0, \boldsymbol{\varphi}(0)\right\rangle. \tag{5.11}$$

To conclude, following Section 3.3 of [27], we observe that (5.10) reveals that

$$\partial_t\boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{w} \quad \text{a.e. in } \Omega \times (0, T). \tag{5.12}$$

Since $|\boldsymbol{m}| = 1$ and $\boldsymbol{m} \cdot \boldsymbol{w} = 0$, using (3.6), it follows that

$$\boldsymbol{w} = \boldsymbol{m} \times \partial_t\boldsymbol{m} \quad \text{a.e. in } \Omega \times (0, T). \tag{5.13}$$

Using (5.12) and (5.13) in the term on the left-hand side and in the third term on the right-hand side of (5.11), respectively, we obtain the variational formulation (2.9). By density, it follows that $\boldsymbol{m}$ satisfies (2.9) for all $\boldsymbol{\varphi} \in C_c^\infty([0, T); \boldsymbol{H}^1(\Omega))$. This shows that $\boldsymbol{m}$ satisfies part (iii) of Definition 2.1. The verification of part (ii) and (iv) can be performed using the argument employed for Algorithm 3.2. This proves the desired convergence and concludes the proof. □

In the following proposition, we establish the boundedness of the approximations generated by Algorithm 3.4.

**Proposition 5.1.** *Let $j \in \mathbb{N}$. Assume that $\varepsilon = \mathcal{O}(h_{\min})$ as $h, \varepsilon \to 0$. Then, there exist thresholds $h_0, k_0, \varepsilon_0 > 0$ such that, if $h < h_0$, $k < k_0$, and $\varepsilon < \varepsilon_0$, the approximations generated by Algorithm 3.4 satisfy the inequality*

$$\left\|\boldsymbol{\nabla m}_h^j\right\|_{\boldsymbol{L}^2(\Omega)}^2 + \left\|\boldsymbol{w}_h^j\right\|_h^2 + k\sum_{i=0}^{j-1}\left\|d_t\boldsymbol{m}_h^{i+1}\right\|_h^2 \leq C(1 + jk)\exp(jk). \tag{5.14}$$

*The thresholds $h_0, k_0, \varepsilon_0 > 0$ and the constant $C > 0$ depend only on the shape-regularity of $\mathcal{T}_h$ and the problem data.*

*Proof.* Let $i = 0, \ldots, j-1$. Proposition (3.6)(iii) yields (3.17). Taking the sum over $i = 0, \ldots, j-1$, we obtain the identity

$$\frac{1}{2}\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^j\right\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\tau}{2}\left\|\boldsymbol{w}_h^j\right\|_h^2 + \alpha k \sum_{i=0}^{j-1}\left\|d_t\boldsymbol{m}_h^{i+1}\right\|_h^2$$
$$= \frac{1}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\tau}{2}\|\boldsymbol{w}_h^0\|_h^2 - k\sum_{i=0}^{j-1}\left\langle\boldsymbol{m}_h^{i+1/2}\times\boldsymbol{r}_h^i, \mathbb{P}_h\boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right] - \alpha\,d_t\boldsymbol{m}_h^{i+1}\right\rangle_h.$$

A straightforward application of the discrete Young inequality yields the inequalities

$$\sum_{i=0}^{j-1}\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^{i+1/2}\right\|_{\boldsymbol{L}^2(\Omega)} \leq \frac{j}{2} + \frac{1}{2}\sum_{i=0}^{j-1}\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^i\right\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{1}{4}\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^j\right\|_{\boldsymbol{L}^2(\Omega)}^2,$$
$$\left\|d_t\boldsymbol{m}_h^{i+1}\right\|_h \leq \frac{1}{2} + \frac{1}{2}\left\|d_t\boldsymbol{m}_h^{i+1}\right\|_h^2.$$

Since $\left\|\boldsymbol{m}_h^{i+1/2}\right\|_{\boldsymbol{L}^\infty(\Omega)} \leq 1$, $\left\|\boldsymbol{r}_h^i\right\|_h \leq \varepsilon$, and $\left\|\mathbb{P}_h\boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right]\right\|_h \leq Ch_{\min}^{-1}\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^{i+1/2}\right\|_{\boldsymbol{L}^2(\Omega)}$ (where $C > 0$ depends only on the shape-regularity of $\mathcal{T}_h$), we obtain that

$$k\sum_{i=0}^{j-1}\left\langle\boldsymbol{m}_h^{i+1/2}\times\boldsymbol{r}_h^i, \mathbb{P}_h\boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right] - \alpha\,d_t\boldsymbol{m}_h^{i+1}\right\rangle_h$$
$$\leq k\sum_{i=0}^{j-1}\left\|\boldsymbol{m}_h^{i+1/2}\right\|_{\boldsymbol{L}^\infty(\Omega)}\left\|\boldsymbol{r}_h^i\right\|_h\left(\left\|\mathbb{P}_h\boldsymbol{h}_{\mathrm{eff}}\left[\boldsymbol{m}_h^{i+1/2}\right]\right\|_h + \alpha\left\|d_t\boldsymbol{m}_h^{i+1}\right\|_h\right)$$
$$\leq \varepsilon k\sum_{i=0}^{j-1}\left(Ch_{\min}^{-1}\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^{i+1/2}\right\|_{\boldsymbol{L}^2(\Omega)} + \alpha\left\|d_t\boldsymbol{m}_h^{i+1}\right\|_h\right)$$
$$\leq \frac{(Ch_{\min}^{-1}+\alpha)\varepsilon jk}{2} + \frac{C\varepsilon h_{\min}^{-1}k}{4}\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^j\right\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{C\varepsilon h_{\min}^{-1}k}{2}\sum_{i=0}^{j-1}\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^i\right\|_{\boldsymbol{L}^2(\Omega)}^2 + \alpha\varepsilon k\sum_{i=0}^{j-1}\left\|d_t\boldsymbol{m}_h^{i+1}\right\|_h^2.$$

Overall, we thus obtain that

$$\frac{1}{2}\left(1 - \frac{C\varepsilon h_{\min}^{-1}k}{2}\right)\left\|\boldsymbol{\nabla}\boldsymbol{m}_h^j\right\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\tau}{2}\left\|\boldsymbol{w}_h^j\right\|_h^2 + \alpha(1-\varepsilon)k\sum_{i=0}^{j-1}\left\|d_t\boldsymbol{m}_h^{i+1}\right\|_h^2$$
$$\leq \frac{1}{2}\|\boldsymbol{\nabla}\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{\tau}{2}\|\boldsymbol{w}_h^0\|_h^2 + \frac{(Ch_{\min}^{-1}+\alpha)\varepsilon jk}{2} + \frac{C\varepsilon h_{\min}^{-1}k}{2}\sum_{i=0}^{j-1}\|\boldsymbol{\nabla}\boldsymbol{m}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2.$$

Using this estimate, the convergence (3.18), and the assumption $\varepsilon = \mathcal{O}(h_{\min})$ as $h, \varepsilon \to 0$, (5.14) can be shown by applying the discrete Gronwall lemma (assuming the discretisation parameters to be sufficiently small). $\quad\square$

The boundedness result established in Proposition 5.1 is the starting point to prove Theorem 3.7(ii). We omit the presentation of the proof, since this follows line-by-line the argument used to show the convergence of Algorithm 3.4. We only stress that, in the proof of the variational formulation (2.9) and the energy inequality (2.10), the additional contributions arising from the inexact solution of the nonlinear system are always uniformly bounded by $\varepsilon$ and therefore vanish in the limit.

# References

[1] C. Abert, G. Hrkac, M. Page, D. Praetorius, M. Ruggeri and D. Suess, Spin-polarized transport in ferromagnetic multilayers: an unconditionally convergent FEM integrator. *Comput. Math. Appl.* **68** (2014) 639–654.

[2] F. Alouges, A new finite element scheme for Landau–Lifchitz equations. *Discrete Contin. Dyn. Syst. Ser. S* **1** (2008) 187–196.

[3] F. Alouges and P. Jaisson, Convergence of a finite element discretization for the Landau–Lifshitz equation in micromagnetism. *Math. Models Methods Appl. Sci.* **16** (2006) 299–316.

[4] F. Alouges and A. Soyeur, On global weak solutions for Landau–Lifshitz equations: existence and nonuniqueness. *Nonlinear Anal.* **18** (1992) 1071–1084.

[5] F. Alouges, E. Kritsikis, J. Steiner and J.-C. Toussaint, A convergent and precise finite element scheme for Landau–Lifschitz–Gilbert equation. *Numer. Math.* **128** (2014) 407–430.

[6] R. An, Optimal error estimates of linearized Crank–Nicolson Galerkin method for Landau–Lifshitz equation. *J. Sci. Comput.* **69** (2016) 1–27.

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

[8] S. Bartels, Semi-implicit approximation of wave maps into smooth or convex surfaces. *SIAM J. Numer. Anal.* **47** (2009) 3486–3506.

[9] S. Bartels, Fast and accurate finite element approximation of wave maps into spheres. *ESAIM: Math. Model. Numer. Anal.* **49** (2015) 551–558.

[10] S. Bartels, Numerical Methods for Nonlinear Partial Differential Equations. Vol. 47 of *Springer Series in Computational Mathematics*. Springer (2015).

[11] S. Bartels, Projection-free approximation of geometrically constrained partial differential equations. *Math. Comp.* **85** (2016) 1033–1049.

[12] S. Bartels and A. Prohl, Convergence of an implicit finite element method for the Landau–Lifshitz–Gilbert equation. *SIAM J. Numer. Anal.* **44** (2006) 1405–1419.

[13] S. Bartels, X. Feng and A. Prohl, Finite element approximations of wave maps into spheres. *SIAM J. Numer. Anal.* **46** (2007) 61–87.

[14] S. Bartels, J. Ko and A. Prohl, Numerical analysis of an explicit approximation scheme for the Landau–Lifshitz–Gilbert equation. *Math. Comp.* **77** (2008) 773–788.

[15] S. Bartels, C. Lubich and A. Prohl, Convergent discretization of heat and wave map flows to spheres using approximate discrete Lagrange multipliers. *Math. Comp.* **78** (2009) 1269–1292.

[16] E. Beaurepaire, J.-C. Merle, A. Daunois and J.-Y. Bigot, Ultrafast spin dynamics in ferromagnetic nickel. *Phys. Rev. Lett.* **76** (1996) 4250–4253.

[17] F. Bruckner, M. Feischl, T. Führer, P. Goldenits, M. Page, D. Praetorius, M. Ruggeri and D. Suess, Multiscale modeling in micromagnetics: Existence of solutions and numerical integration. *Math. Models Methods Appl. Sci.* **24** (2014) 2627–2662.

[18] G. Carbou, Thin layers in micromagnetism. *Math. Models Methods Appl. Sci.* **11** (2001) 1529–1546.

[19] M.-C. Ciornei, J.M. Rubí and J.-E. Wegrowe, Magnetization dynamics in the inertial regime: nutation predicted at short time scales. *Phys. Rev. B* **83** (2011) 020410.

[20] G. Di Fratta, Micromagnetics of curved thin films. *Z. Angew. Math. Phys.* **71** (2020) 111.

[21] G. Di Fratta, C.-M. Pfeiler, D. Praetorius, M. Ruggeri and B. Stiftner, Linear second order IMEX-type integrator for the (eddy current) Landau–Lifshitz–Gilbert equation. *IMA J. Numer. Anal.* **40** (2020) 2802–2838.

[22] M. Feischl and T. Tran, The Eddy Current–LLG equations: FEM-BEM coupling and a priori error estimates. *SIAM J. Numer. Anal.* **55** (2017) 1786–1819.

[23] H. Gao, Optimal error estimates of a linearized backward Euler FEM for the Landau–Lifshitz equation. *SIAM J. Numer. Anal.* **52** (2014) 2574–2593.

[24] G. Gioia and R.D. James, Micromagnetics of very thin films. *Proc. Roy. Soc. Lond. A* **453** (1997) 213–223.

[25] M. Hadda and M. Tilioua, On magnetization dynamics with inertial effects. *J. Eng. Math.* **88** (2014) 197–206.

[26] G. Hrkac, C.-M. Pfeiler, D. Praetorius, M. Ruggeri, A. Segatti and B. Stiftner, Convergent tangent plane integrators for the simulation of chiral magnetic skyrmion dynamics. *Adv. Comput. Math.* **45** (2019) 1329–1368.

[27] T.K. Karper and F. Weber, A new angular momentum method for computing wave maps into spheres. *SIAM J. Numer. Anal.* **52** (2014) 2073–2091.

[28] E. Kim and J. Wilkening, Convergence of a mass-lumped finite element method for the Landau–Lifshitz equation. *Quart. Appl. Math.* **76** (2018) 383–405.

[29] S. Korotov, M. Křížek and P. Neittaanmaki, Weakened acute type condition for tetrahedral triangulations and the discrete maximum principle. *Math. Comput.* **70** (2001) 107–119.

[30] F. Lin and C. Wang, The Analysis of Harmonic Maps and Their Heat Flows. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ (2008).

[31] M. Moumni and M. Tilioua, A finite-difference scheme for a model of magnetization dynamics with inertial effects. *J. Eng. Math.* **100** (2016) 95–106.

[32] K. Neeraj, N. Awari, S. Kovalev, D. Polley, N.Z. Hagström, S.S. Arekapudi, A. Semisalova, K. Lenz, B. Green, J.-C. Deinert, I. Ilyakov, M. Chen, M. Bawatna, V. Scalera, M. d'Aquino, C. Serpico, O. Hellwig, J.-E. Wegrowe, M. Gensch and S. Bonetti, Inertial spin dynamics in ferromagnets. *Nat. Phys.* **17** (2021) 245–250.

[33] NIST, Micromagnetic modeling activity group website. http://www.ctcms.nist.gov/~rdm/mumag.html. Accessed on November 25, 2020.

[34] D. Praetorius, M. Ruggeri and B. Stiftner, Convergence of an implicit–explicit midpoint scheme for computational micromagnetics. *Comput. Math. Appl.* **75** (2018) 1719–1738.

[35] D. Tataru, The wave maps equation. *Bull. Amer. Math. Soc. (N.S.)* **41** (2004) 185–204.

[36] J. Walowski and M. Münzenberg, Perspective: ultrafast magnetism and THz spintronics. *J. Appl. Phys.* **120** (2016) 140901.