Research Article

Giovanni Di Fratta, Carl-Martin Pfeiler*, Dirk Praetorius and Michele Ruggeri

# The Mass-Lumped Midpoint Scheme for Computational Micromagnetics: Newton Linearization and Application to Magnetic Skyrmion Dynamics

**Abstract:** We discuss a mass-lumped midpoint scheme for the numerical approximation of the Landau–Lifshitz–Gilbert equation, which models the dynamics of the magnetization in ferromagnetic materials. In addition to the classical micromagnetic field contributions, our setting covers the non-standard Dzyaloshinskii–Moriya interaction, which is the essential ingredient for the enucleation and stabilization of magnetic skyrmions. Our analysis also includes the inexact solution of the arising nonlinear systems, for which we discuss both a constraint-preserving fixed-point solver from the literature and a novel approach based on the Newton method. We numerically compare the two linearization techniques and show that the Newton solver leads to a considerably lower number of nonlinear iterations. Moreover, in a numerical study on magnetic skyrmions, we demonstrate that, for magnetization dynamics that are very sensitive to energy perturbations, the midpoint scheme, due to its conservation properties, is superior to the dissipative tangent plane schemes from the literature.

**Keywords:** Landau–Lifshitz–Gilbert Equation, Dzyaloshinskii–Moriya Interaction, Magnetic Skyrmions, Newton Linearization, Computational Micromagnetics, Finite Elements

**MSC 2010:** 35K55, 65M12, 65M22, 65M60, 65Z05

## 1 Introduction

### 1.1 Energetics of a Ferromagnet

In the continuum theory of micromagnetism, whose origin dates back to the seminal work of Landau–Lifshitz [34] on small ferromagnetic particles, the amount of magnetic moment (per unit volume) of a rigid ferromagnetic body occupying a bounded region $\Omega \subset \mathbb{R}^3$ is represented by a classical vector field, the *magnetization* $\boldsymbol{M} : \Omega \to \mathbb{R}^3$. Its module $M_s := |\boldsymbol{M}|$ describes the so-called *saturation magnetization*. In single-crystal ferromagnets [2, 6], $M_s$ depends only on the temperature and is assumed to be constant when the specimen is

**\*Corresponding author: Carl-Martin Pfeiler,** Institute of Analysis and Scientific Computing, TU Wien, Wiedner Hauptstrasse 8–10, 1040, Vienna, Austria, e-mail: carl-martin.pfeiler@asc.tuwien.ac.at
**Giovanni Di Fratta,** Dipartimento di Matematica e Applicazioni "R. Caccioppoli", Università degli Studi di Napoli "Federico II", Via Cintia, Complesso Monte S. Angelo, 80126 Napoli, Italy, e-mail: giovanni.difratta@unina.it.
https://orcid.org/0000-0003-0254-2957
**Dirk Praetorius,** Institute of Analysis and Scientific Computing, TU Wien, Wiedner Hauptstrasse 8–10, 1040, Vienna, Austria, e-mail: dirk.praetorius@asc.tuwien.ac.at. https://orcid.org/0000-0002-1977-9830
**Michele Ruggeri,** Department of Mathematics and Statistics, University of Strathclyde, 26 Richmond Street, Glasgow G1 1XH, United Kingdom, e-mail: michele.ruggeri@strath.ac.uk. https://orcid.org/0000-0001-6213-1602

well below the so-called Curie temperature of the material. In this case, the magnetization can be represented in the form $\boldsymbol{M} := M_s \boldsymbol{m}$, where $\boldsymbol{m} \colon \Omega \to \mathbb{S}^2$ is a vector field with values in the unit sphere of $\mathbb{R}^3$, and the observable magnetization states minimize the micromagnetic energy functional [16, 31]

$$\mathcal{E}(\boldsymbol{m}) := \mathcal{E}_\Omega(\boldsymbol{m}) + \mathcal{K}_\Omega(\boldsymbol{m}) + \mathcal{W}_\Omega(\boldsymbol{m}) + \mathcal{A}_\Omega(\boldsymbol{m}) + \mathcal{Z}_\Omega(\boldsymbol{m})$$

$$:= \int_\Omega A|\nabla \boldsymbol{m}|^2 + D(\nabla \times \boldsymbol{m}) \cdot \boldsymbol{m} - \frac{\mu_0}{2} M_s \boldsymbol{H}_s(\boldsymbol{m}) \cdot \boldsymbol{m} + \varphi_{an}(\boldsymbol{m}) - \mu_0 M_s \boldsymbol{H}_{ext} \cdot \boldsymbol{m} \, d\boldsymbol{x}, \tag{1.1}$$

defined for every $\boldsymbol{m} \in H^1(\Omega; \mathbb{S}^2)$.

The *exchange energy* $\mathcal{E}_\Omega(\boldsymbol{m})$ penalizes spatial variations of the direction of the magnetization, with $A > 0$ representing a material-dependent constant that summarizes the stiffness of short-range (symmetric) exchange interactions. The second term $\mathcal{K}_\Omega(\boldsymbol{m})$ represents the *bulk Dzyaloshinskii–Moriya interaction (DMI)* [23, 35] and accounts for antisymmetric exchange interactions caused by possible lack of inversion symmetry in the crystal structure of the ferromagnet. The sign of the constant $D \in \mathbb{R}$ affects the chirality of the ferromagnetic system [42, 43]. The third term $\mathcal{W}_\Omega(\boldsymbol{m})$ is the *magnetostatic self-energy*, i.e., the energy due to the stray field $\boldsymbol{H}_s(\boldsymbol{m})$ induced by $M_s \boldsymbol{m}$. From the mathematical point of view, $\boldsymbol{H}_s(\boldsymbol{m})$ can be characterized as the projection of $(-M_s \chi_\Omega \boldsymbol{m}) \in L^2(\mathbb{R}^3; \mathbb{R}^3)$ on the closed subspace of gradient vector fields $\nabla H^1(\mathbb{R}^3; \mathbb{R}) := \{\nabla u : u \in H^1(\mathbb{R}^3; \mathbb{R})\}$ (see, e.g., [21, 38]). Here, $\mu_0$ denotes the vacuum permeability. Additionally, the micromagnetic energy includes two additional energy contributions: the *magnetocrystalline anisotropy energy* $\mathcal{A}_\Omega(\boldsymbol{m})$ and the *Zeeman energy* $\mathcal{Z}_\Omega(\boldsymbol{m})$. The energy density $\varphi_{an} \colon \mathbb{S}^2 \to \mathbb{R}_{\geq 0}$ models the existence of easy directions of the magnetization due to the crystallographic structure of the ferromagnet, while $\mathcal{Z}_\Omega(\boldsymbol{m})$ models the tendency of a specimen to have the magnetization aligned with the external applied field $\boldsymbol{H}_{ext} \in L^2(\Omega; \mathbb{R}^3)$, assumed to be unaffected by variations of $\boldsymbol{m}$. The competition among the energy contributions in (1.1) explains *most* of the striking spin textures observable in ferromagnetic materials [31], in particular, the emergence of magnetic skyrmions [25, 26].

## 1.2 A More General Energy Functional

When a ferromagnetic system consists of several magnetic materials, the material-dependent quantities $A$, $D$, and $M_s$ are no longer constant in the region $\Omega$ occupied by the ferromagnet, and one has to model spin interactions among different magnetic materials at their touching interface [2]. The easiest way is to assume a *strong coupling condition* [5, 6, 20]: although $M_s$ can be discontinuous across an interface, the direction of the magnetization never jumps through it. Under this constitutive assumption, the analysis of the composite can be carried out under the classical conditions $M_s \in L^\infty(\Omega; \mathbb{R}_{>0})$ and $\boldsymbol{m} \in H^1(\Omega; \mathbb{S}^2)$. In this setting, the observable states of a rigid ferromagnetic body can be characterized as the local minimizers of the micromagnetic energy functional still defined by (1.1), but with the quantities $A = A(\boldsymbol{x})$, $D = D(\boldsymbol{x})$, and $M_s = M_s(\boldsymbol{x})$ to be understood as functions defined on $\Omega$.

In this paper, we are interested in a more general energy functional which, other than incorporating the previous one as a special case, also accounts for the presence of anisotropies in the lattice structures of the constituents. To introduce the model, we first observe that the bulk DMI energy density can be equivalently rewritten as

$$D(\nabla \times \boldsymbol{m}) \cdot \boldsymbol{m} = D \sum_{d=1}^{3} (\boldsymbol{e}_d \times \partial_d \boldsymbol{m}) \cdot \boldsymbol{m},$$

where $\{\boldsymbol{e}_d\}_{d=1,2,3}$ denotes the standard basis of $\mathbb{R}^3$. It is therefore a special case of the energy density

$$g_{asym}(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{\xi}) = \sum_{d=1}^{3} \boldsymbol{K}_d(\boldsymbol{x}) \boldsymbol{\xi}_d \cdot \boldsymbol{s} \quad \text{for all } \boldsymbol{x} \in \Omega, \ \boldsymbol{s} \in \mathbb{R}^3, \text{ and } \boldsymbol{\xi} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) \in \mathbb{R}^{3 \times 3},$$

with $\{\boldsymbol{K}_d\}_{d=1,2,3}$ being 3-by-3 antisymmetric matrices, i.e., $\boldsymbol{K}_d = -\boldsymbol{K}_d^T$. Similarly, the symmetric exchange energy density can be generalized to the density

$$g_{sym}(\boldsymbol{x}, \boldsymbol{\xi}) = \frac{1}{2} \sum_{d=1}^{3} \boldsymbol{A}_d(\boldsymbol{x}) \boldsymbol{\xi}_d \cdot \boldsymbol{\xi}_d \quad \text{for all } \boldsymbol{x} \in \Omega \text{ and } \boldsymbol{\xi} = (\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3) \in \mathbb{R}^{3 \times 3},$$

with $\{A_d\}_{d=1,2,3}$ being 3-by-3 invertible symmetric matrices, i.e., $A_d = A_d^T$. Hence, for $g := g_{\text{sym}} + g_{\text{asym}}$, it holds that

$$
\begin{aligned}
g(\boldsymbol{x}, \boldsymbol{s}, \boldsymbol{\xi}) &= \frac{1}{2} \sum_{d=1}^{3} \left( A_d(\boldsymbol{x})\boldsymbol{\xi}_d \cdot \boldsymbol{\xi}_d - 2K_d(\boldsymbol{x})\boldsymbol{s} \cdot \boldsymbol{\xi}_d \right) \\
&= \frac{1}{2} \sum_{d=1}^{3} A_d(\boldsymbol{x})(\boldsymbol{\xi}_d - A_d^{-1}(\boldsymbol{x})K_d(\boldsymbol{x})\boldsymbol{s}) \cdot (\boldsymbol{\xi}_d - A_d^{-1}(\boldsymbol{x})K_d(\boldsymbol{x})\boldsymbol{s}) + \frac{1}{2} \sum_{d=1}^{3} K_d(\boldsymbol{x})A_d^{-1}(\boldsymbol{x})K_d(\boldsymbol{x})\boldsymbol{s} \cdot \boldsymbol{s}. \quad (1.2)
\end{aligned}
$$

Note that $K_d(\boldsymbol{x})A_d^{-1}(\boldsymbol{x})K_d(\boldsymbol{x})$ is a symmetric matrix. This discussion suggests the opportunity to investigate an energy functional covering the above generalized form; see (2.1) below. It is worth to notice that the structure of this energy functional does not only allow for the description of a mixture of ferromagnetic materials, but also covers typical homogeneous models arising as Γ-limit of composite ferromagnetic materials with highly oscillating heterogeneities [5, 20].

## 1.3 Landau–Lifshitz–Gilbert Equation and Its Numerical Integration

When the magnetization $\boldsymbol{m}$ does not minimize the micromagnetic energy functional, the ferromagnetic system is in a non-equilibrium state. A well-accepted model for its time evolution is the Landau–Lifshitz–Gilbert equation (LLG) [28, 34], which in the so-called Gilbert form reads

$$
\partial_t \boldsymbol{m} = -\gamma_0 \boldsymbol{m} \times \boldsymbol{H}_{\text{eff}}(\boldsymbol{m}) + \alpha \boldsymbol{m} \times \partial_t \boldsymbol{m}. \tag{1.3}
$$

This phenomenological equation describes the magnetization dynamics as a dissipative precession driven by the effective field $\boldsymbol{H}_{\text{eff}}(\boldsymbol{m}) := -\mu_0^{-1} M_s^{-1} \frac{\delta \mathcal{E}(\boldsymbol{m})}{\delta \boldsymbol{m}}$, and modulated by the gyromagnetic ratio of the electron $\gamma_0 > 0$ and the Gilbert damping parameter $\alpha > 0$. The numerical approximation of LLG is not a trivial task. Nonlinearities, the numerical realization of the unit-length constraint, the possible coupling with other (nonlinear) partial differential equations, and the need of unconditionally stable numerical schemes make the problem very challenging. For this reason, in the last twenty years, the problem has been the subject of several mathematical studies; see, e.g., [1, 3, 4, 7–9, 14, 15, 17–19, 22, 24, 27, 30, 32, 33, 40].

In this work, we consider the mass-lumped midpoint scheme proposed in [15]. The method is based on a mass-lumped first-order finite element method for the spatial discretization and the second-order midpoint rule for the time discretization, and involves the solution of one nonlinear system per time step. Besides introducing the method, the work [15] proves unconditional convergence of the finite element approximation towards a weak solution of LLG in the sense of [10] and proposes a fixed-point iteration to linearize the nonlinear problem arising from the scheme. The scheme has also been the subject of further research. On the one hand, the works [12, 19] incorporate the inexact solution of the nonlinear system into the convergence result. On the other hand, the work [39] focuses on the design and the analysis of effective approaches to treat the nonlocal field contributions.

## 1.4 Contributions

In this work, as a novel contribution, we extend the midpoint scheme and its analysis to more general energy contributions; see the discussion in Section 1.2. In particular, the present analysis covers DMI, which is not covered by the analysis in [12, 15, 19, 39]. We note that DMI is the essential ingredient for the enucleation and the stabilization of magnetic skyrmions. At this point, it is worth pointing out that DMI contributions represent a challenging testing ground for numerical schemes for LLG. Indeed, besides requiring accurate adaptations in the numerical analysis, they determine magnetization configurations – magnetic skyrmions – that turn out to be very sensitive to small perturbations of the micromagnetic energy. In addition, we also discuss the linearization of the nonlinear scheme: we extend the fixed-point iteration proposed in [12] to the present setting and propose an approach based on the Newton method, for which we provide a first full

analysis (well-posedness, stability, convergence). Finally, in a collection of numerical experiments, first, we accurately test the energy conservation properties of the mass-lumped midpoint scheme and extensively compare it with the tangent plane schemes from [4, 8, 22, 30]. Second, we emphasize the faster convergence of the Newton linearization over the fixed-point solver in a numerical study, which additionally indicates that the restrictive conditions required by our analysis of the Newton linearization are likely to be suboptimal and might be weakened to match those for the fixed-point solver by an improved analysis.

## 1.5 Outline

The remainder of the work is organized as follows. We conclude this section by collecting the notation used throughout the paper. In Section 2, we describe the mathematical problem under consideration. In Section 3, we present the proposed algorithms and state their stability and convergence results. Section 4 is devoted to numerical experiments. Finally, in Sections 5–6, we collect the proofs of the results stated in Section 3.

## 1.6 Notation

Throughout the paper, we use the standard notation for Lebesgue, Sobolev, and Bochner spaces and norms. To highlight (spaces of) vector-valued or matrix-valued functions, we use bold letters, e.g., we denote both $L^2(\Omega; \mathbb{R}^3)$ and $L^2(\Omega; \mathbb{R}^{3\times3})$ by $\boldsymbol{L}^2(\Omega)$. We denote by $\langle\,\cdot\,,\cdot\,\rangle_\Omega$ the scalar product in $\boldsymbol{L}^2(\Omega)$ and by $\langle\,\cdot\,,\cdot\,\rangle$ the duality pairing between $\boldsymbol{H}^1(\Omega)$ and its dual. By $C > 0$, we always denote a generic constant, which is independent of the discretization parameters, but not necessarily the same at each occurrence.

# 2 Problem Formulation

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

$$\mathcal{E}(\boldsymbol{m}) = \frac{1}{2}a(\boldsymbol{m}, \boldsymbol{m}) - \langle\boldsymbol{f}, \boldsymbol{m}\rangle_\Omega, \tag{2.1}$$

where $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$, while the bilinear form $a\colon \boldsymbol{H}^1(\Omega) \times \boldsymbol{H}^1(\Omega) \to \mathbb{R}$ is defined, for all $\boldsymbol{\psi}, \boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega)$, by

$$a(\boldsymbol{\psi}, \boldsymbol{\phi}) = \sum_{d=1}^{3}\langle\boldsymbol{A}_d(\partial_d\boldsymbol{\psi} - \boldsymbol{J}_d\boldsymbol{\psi}), \partial_d\boldsymbol{\phi} - \boldsymbol{J}_d\boldsymbol{\phi}\rangle_\Omega - \langle\boldsymbol{\pi}(\boldsymbol{\psi}), \boldsymbol{\phi}\rangle_\Omega. \tag{2.2}$$

Here, $\boldsymbol{\pi}\colon \boldsymbol{L}^2(\Omega) \to \boldsymbol{L}^2(\Omega)$ is a linear, bounded, and self-adjoint operator, while, for $d = 1, 2, 3$, the 3-by-3 matrices $\boldsymbol{A}_d$ and $\boldsymbol{J}_d$ have coefficients in $L^\infty(\Omega)$, with $\boldsymbol{A}_d$ being also symmetric and uniformly positive definite, i.e., it holds that $\boldsymbol{A}_d^T = \boldsymbol{A}_d$ and

$$\boldsymbol{A}_d(\boldsymbol{x})\boldsymbol{u} \cdot \boldsymbol{u} \geq A_0|\boldsymbol{u}|^2 \quad \text{for almost all } \boldsymbol{x} \in \Omega \text{ and all } \boldsymbol{u} \in \mathbb{R}^3,$$

where $A_0 > 0$ is a fixed constant. The energy (2.1) covers the extensions of the classical micromagnetic functional discussed in Section 1.2; cf. the expression in (1.2).

The existence of minimizers of (2.1) in $\boldsymbol{H}^1(\Omega; \mathbb{S}^2)$ follows from the direct method of calculus of variations. Moreover, any minimizer $\boldsymbol{m} \in \boldsymbol{H}^1(\Omega; \mathbb{S}^2)$ satisfies the Euler–Lagrange equations

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

Here, $\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) := -\frac{\delta\mathcal{E}(\boldsymbol{m})}{\delta\boldsymbol{m}}$ is the (negative) Gâteaux derivative of the energy, i.e.,

$$-\langle\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}), \boldsymbol{\phi}\rangle = \left\langle\frac{\delta\mathcal{E}(\boldsymbol{m})}{\delta\boldsymbol{m}}, \boldsymbol{\phi}\right\rangle = \lim_{\delta\to0}\frac{\mathcal{E}(\boldsymbol{m} + \delta\boldsymbol{\phi}) - \mathcal{E}(\boldsymbol{m})}{\delta} \overset{(2.1)}{=} a(\boldsymbol{m}, \boldsymbol{\phi}) - \langle\boldsymbol{f}, \boldsymbol{\phi}\rangle_\Omega. \tag{2.3}$$

Turning to the dynamical case, a non-equilibrium configuration $\boldsymbol{m}(t) \in \boldsymbol{H}^1(\Omega; \mathbb{S}^2)$ evolves according to (1.3), which, after a suitable rescaling, reads

$$\partial_t \boldsymbol{m} = -\boldsymbol{m} \times \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}) + \alpha \boldsymbol{m} \times \partial_t \boldsymbol{m}, \tag{2.4}$$

with $\alpha > 0$ being the Gilbert damping parameter. The dynamics is dissipative in the sense that any sufficiently smooth solution of (2.4) satisfies the energy law

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

We conclude this section by recalling the notion of a weak solution of (2.4); see [10].

**Definition 2.1.** Let $\boldsymbol{m}^0 \in \boldsymbol{H}^1(\Omega; \mathbb{S}^2)$. A vector field $\boldsymbol{m} \colon \Omega \times \mathbb{R}_{>0} \to \mathbb{S}^2$ is called a global weak solution of (2.4) if $\boldsymbol{m} \in L^\infty(\mathbb{R}_{>0}; \boldsymbol{H}^1(\Omega; \mathbb{S}^2))$ and, for all $T > 0$, with $\Omega_T := \Omega \times (0, T)$, the following properties are satisfied:
(i) $\boldsymbol{m} \in \boldsymbol{H}^1(\Omega_T)$;
(ii) $\boldsymbol{m}(0) = \boldsymbol{m}^0$ in the sense of traces;
(iii) for all $\boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega_T)$, it holds that

$$\int_0^T \langle \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle_\Omega \, \mathrm{d}t = -\int_0^T \langle \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}(t)), \boldsymbol{\varphi}(t) \times \boldsymbol{m}(t) \rangle \, \mathrm{d}t + \alpha \int_0^T \langle \boldsymbol{m}(t) \times \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle_\Omega \, \mathrm{d}t; \tag{2.6}$$

(iv) it holds that

$$\mathcal{E}(\boldsymbol{m}(T)) + \alpha \int_0^T \|\partial_t \boldsymbol{m}(t)\|^2_{\boldsymbol{L}^2(\Omega)} \, \mathrm{d}t \leq \mathcal{E}(\boldsymbol{m}^0). \tag{2.7}$$

We note that (2.3) implicitly includes natural boundary conditions on $\boldsymbol{m}$, which are homogeneous Neumann boundary conditions $\partial_{\boldsymbol{n}} = \boldsymbol{0}$ if $\boldsymbol{A}_d = \ell_{\text{ex}}^2 \, \mathbf{Id}$ and $\boldsymbol{J}_d = \boldsymbol{0}$ for $d = 1, 2, 3$. For a more explicit presentation, we refer to [30]. The variational formulation (2.6) comes from a weak formulation of (2.4) in the space-time cylinder. The energy inequality (2.7) is a weak counterpart of the dissipative energy law (2.5).

**Remark 2.2.** (i) For ease of presentation, we restrict ourselves to the case of a time-independent field $\boldsymbol{f} \in \boldsymbol{L}^2(\Omega)$. For time-dependent fields, the strong form (2.5) and the weak form (2.7) of the energy law of LLG read

$$\frac{\mathrm{d}}{\mathrm{d}t} \tilde{\mathcal{E}}(\boldsymbol{m}(t)) = -\alpha \|\partial_t \boldsymbol{m}(t)\|^2_{\boldsymbol{L}^2(\Omega)} + \langle \boldsymbol{f}(t), \partial_t \boldsymbol{m}(t) \rangle_\Omega$$

and

$$\tilde{\mathcal{E}}(\boldsymbol{m}(T)) + \alpha \int_0^T \|\partial_t \boldsymbol{m}(t)\|^2_{\boldsymbol{L}^2(\Omega)} \, \mathrm{d}t - \int_0^T \langle \boldsymbol{f}(t), \partial_t \boldsymbol{m}(t) \rangle_\Omega \, \mathrm{d}t \leq \tilde{\mathcal{E}}(\boldsymbol{m}^0),$$

respectively, where $\tilde{\mathcal{E}}(\boldsymbol{m}) = \mathcal{E}(\boldsymbol{m}) + \langle \boldsymbol{f}, \boldsymbol{m} \rangle_\Omega = a(\boldsymbol{m}, \boldsymbol{m})/2$.

(ii) The present setting covers and generalizes the model problems considered in previous mathematical works on the numerical integration of LLG.

- With the choices $\boldsymbol{A}_d = \ell_{\text{ex}}^2 \, \mathbf{Id}$ and $\boldsymbol{J}_d = \boldsymbol{0}$ for $d = 1, 2, 3$, where $\ell_{\text{ex}} > 0$ is the so-called exchange length and $\mathbf{Id}$ is the 3-by-3 identity matrix, $\boldsymbol{\pi} \equiv \boldsymbol{0}$, and $\boldsymbol{f} \equiv \boldsymbol{0}$, we obtain that

$$\langle \boldsymbol{h}_{\text{eff}}(\boldsymbol{\psi}), \boldsymbol{\varphi} \rangle = -\ell_{\text{ex}}^2 \langle \nabla \boldsymbol{\psi}, \nabla \boldsymbol{\varphi} \rangle_\Omega.$$

This is the so-called small particle limit of LLG, for which finite element schemes have been proposed, e.g., in the seminal papers [4, 15].
- With the choices $\boldsymbol{A}_d = \ell_{\text{ex}}^2 \, \mathbf{Id}$ and $\boldsymbol{J}_d = \boldsymbol{0}$ for $d = 1, 2, 3$, we obtain that

$$\langle \boldsymbol{h}_{\text{eff}}(\boldsymbol{\psi}), \boldsymbol{\phi} \rangle = -\ell_{\text{ex}}^2 \langle \nabla \boldsymbol{\psi}, \nabla \boldsymbol{\phi} \rangle_\Omega + \langle \boldsymbol{\pi}(\boldsymbol{\psi}), \boldsymbol{\phi} \rangle_\Omega + \langle \boldsymbol{f}, \boldsymbol{\phi} \rangle_\Omega.$$

This setting covers the classical energy contributions considered in micromagnetics (exchange, uniaxial anisotropy, magnetostatic, Zeeman) and numerical integrators for this case have been analyzed, e.g., in [8, 9, 17, 22, 39].

- With the choices $\boldsymbol{A}_d = \ell_{\mathrm{ex}}^2 \, \mathbf{Id}$ and $\boldsymbol{J}_d = \ell_{\mathrm{dm}}[\boldsymbol{e}_d]_\times/(2\ell_{\mathrm{ex}}^2)$ for $d = 1, 2, 3$, where $\ell_{\mathrm{dm}} \in \mathbb{R}$ is a characteristic length associated with DMI, $\boldsymbol{\pi}(\boldsymbol{m}) = \ell_{\mathrm{dm}}\boldsymbol{m}/(2\ell_{\mathrm{ex}}^2)$, and $\boldsymbol{f} \equiv \boldsymbol{0}$, we obtain that

$$\langle \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{\psi}), \boldsymbol{\phi} \rangle = -\ell_{\mathrm{ex}}^2 \langle \nabla \boldsymbol{\psi}, \nabla \boldsymbol{\phi} \rangle_\Omega - \frac{\ell_{\mathrm{dm}}}{2} \langle \nabla \times \boldsymbol{\psi}, \boldsymbol{\phi} \rangle_\Omega - \frac{\ell_{\mathrm{dm}}}{2} \langle \boldsymbol{\psi}, \nabla \times \boldsymbol{\phi} \rangle_\Omega,$$

which is the setting analyzed in [30] for the simulation of chiral magnetic skyrmions by the means of a family of tangent plane integrators. Here, $[\boldsymbol{e}_d]_\times \in \mathbb{R}^{3\times3}$ denotes the matrix such that $[\boldsymbol{e}_d]_\times \boldsymbol{u} = \boldsymbol{e}_d \times \boldsymbol{u}$ for all $\boldsymbol{u} \in \mathbb{R}^3$.

# 3 Numerical Algorithms and Main Results

## 3.1 Preliminaries

Let $\kappa \geq 1$. For the spatial discretization, assuming $\Omega$ to be a polyhedral domain, we consider a $\kappa$-quasi-uniform family $\{\mathcal{T}_h\}_{h>0}$ of regular tetrahedral triangulations of $\Omega$ parametrized by the mesh size $h = \max_{K \in \mathcal{T}_h} \mathrm{diam}(K)$, i.e., $\kappa^{-1}h \leq \mathrm{vol}(K)^{1/3} \leq h$ for all $K \in \mathcal{T}_h$. We denote by $\mathcal{N}_h$ the set of vertices of $\mathcal{T}_h$. For any $K \in \mathcal{T}_h$, we denote by $\mathcal{P}^1(K)$ the space of first-order polynomials on $K$. We consider the space of $\mathcal{T}_h$-piecewise affine and globally continuous finite elements

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

The classical basis for this finite-dimensional subspace of $H^1(\Omega)$ is the set of nodal hat functions $\{\varphi_{\boldsymbol{z}}\}_{\boldsymbol{z} \in \mathcal{N}_h}$, which satisfy $\varphi_{\boldsymbol{z}}(\boldsymbol{z}') = \delta_{\boldsymbol{z},\boldsymbol{z}'}$ for all $\boldsymbol{z}, \boldsymbol{z}' \in \mathcal{N}_h$. The nodal interpolant $\mathcal{I}_h \colon C^0(\overline{\Omega}) \to \mathcal{S}^1(\mathcal{T}_h)$ is defined by $\mathcal{I}_h[u] = \sum_{\boldsymbol{z} \in \mathcal{N}_h} u(\boldsymbol{z})\varphi_{\boldsymbol{z}}$ for all $u \in C^0(\overline{\Omega})$.

Let $\boldsymbol{V}_h := \mathcal{S}^1(\mathcal{T}_h)^3$. For each time step, approximate solutions of (2.4) are sought in the *set of admissible approximate magnetizations*

$$\mathcal{M}_h := \{\boldsymbol{\phi}_h \in \boldsymbol{V}_h : |\boldsymbol{\phi}_h(\boldsymbol{z})| = 1 \text{ for all } \boldsymbol{z} \in \mathcal{N}_h\},$$

which consists of all elements of $\boldsymbol{V}_h$ satisfying the unit-length constraint at the nodes of the triangulation.

Besides the standard scalar product $\langle \cdot, \cdot \rangle_\Omega$, given a mesh $\mathcal{T}_h$ and the associated nodal interpolant $\mathcal{I}_h[\cdot]$, we consider the mass-lumped product $\langle \cdot, \cdot \rangle_h$ defined by

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

Using the definition of the nodal interpolant, we see that

$$\langle \boldsymbol{\psi}, \boldsymbol{\phi} \rangle_h = \sum_{\boldsymbol{z} \in \mathcal{N}_h} \beta_{\boldsymbol{z}} \boldsymbol{\psi}(\boldsymbol{z}) \cdot \boldsymbol{\phi}(\boldsymbol{z}), \quad \text{where } \beta_{\boldsymbol{z}} := \int_\Omega \varphi_{\boldsymbol{z}} \, \mathrm{d}\boldsymbol{x} > 0.$$

On $\boldsymbol{V}_h$, $\langle \cdot, \cdot \rangle_h$ is a scalar product and the induced norm $\|\cdot\|_h$ is equivalent to the standard norm of $\boldsymbol{L}^2(\Omega)$. In particular, it holds that

$$\|\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)} \leq \|\boldsymbol{\phi}_h\|_h \leq \sqrt{5}\|\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \boldsymbol{\phi}_h \in \boldsymbol{V}_h; \tag{3.1}$$

see [13, Lemma 3.9]. Finally, we define the mapping $\mathbb{P}_h \colon \boldsymbol{H}^1(\Omega)^\star \to \boldsymbol{V}_h$ by

$$\langle \mathbb{P}_h \boldsymbol{u}, \boldsymbol{\phi}_h \rangle_h = \langle \boldsymbol{u}, \boldsymbol{\phi}_h \rangle \quad \text{for all } \boldsymbol{u} \in \boldsymbol{H}^1(\Omega)^\star \text{ and } \boldsymbol{\phi}_h \in \boldsymbol{V}_h, \tag{3.2}$$

i.e., $\mathbb{P}_h \boldsymbol{u} \in \boldsymbol{V}_h$ is the Riesz representative of $\langle \boldsymbol{u}, \cdot \rangle \in \boldsymbol{V}_h^\star$ in the Hilbert space $(\boldsymbol{V}_h, \langle \cdot, \cdot \rangle_h)$.

For the time discretization, we consider a partition of the positive real axis $\mathbb{R}_{>0}$ with constant time-step size $k > 0$, i.e., $t_i := ik$ for all $i \in \mathbb{N}_0$. Given a sequence $\{\boldsymbol{\phi}_h^i\}_{i \in \mathbb{N}_0} \subset \boldsymbol{V}_h$, we define

$$\boldsymbol{\phi}_h^{i+1/2} := \frac{\boldsymbol{\phi}_h^{i+1} + \boldsymbol{\phi}_h^i}{2} \quad \text{and} \quad d_t\boldsymbol{\phi}_h^{i+1} := \frac{\boldsymbol{\phi}_h^{i+1} - \boldsymbol{\phi}_h^i}{k} \quad \text{for all } i \in \mathbb{N}_0, \tag{3.3}$$

as well as the piecewise linear time reconstruction

$$\boldsymbol{\phi}_{hk}(t) := \frac{t - t_i}{k} \boldsymbol{\phi}_h^{i+1} + \frac{t_{i+1} - t}{k} \boldsymbol{\phi}_h^i \quad \text{for all } i \in \mathbb{N}_0 \text{ and } t \in [t_i, t_{i+1}], \tag{3.4}$$

which satisfies $\boldsymbol{\phi}_{hk} \in \boldsymbol{H}^1(\Omega \times (0, T))$ for any $T > 0$.

## 3.2 Ideal Midpoint Scheme

In the following algorithm, we adapt the scheme initially proposed in [15] to the present setting. The fundamental ingredients are the midpoint rule for the time discretization, the finite element space $\boldsymbol{V}_h$ endowed with the mass-lumped scalar product $\langle \cdot, \cdot \rangle_h$ for the spatial integration, and the mapping (3.2) for the discrete realization of the effective field. We refer to the method as *ideal midpoint scheme* in the sense that, as we will see in the next section, practical implementations require suitable modifications.

**Algorithm 3.1** (Ideal Midpoint Scheme). Input: $\boldsymbol{m}_h^0 \in \mathcal{M}_h$.
Loop: for all $i \in \mathbb{N}_0$, compute $\boldsymbol{m}_h^{i+1} \in \boldsymbol{V}_h$ such that

$$\langle d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{\phi}_h \rangle_h = -\langle \boldsymbol{m}_h^{i+1/2} \times \mathbb{P}_h \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_h^{i+1/2}), \boldsymbol{\phi}_h \rangle_h + \alpha \langle \boldsymbol{m}_h^{i+1/2} \times d_t \boldsymbol{m}_h^{i+1}, \boldsymbol{\phi}_h \rangle_h \tag{3.5}$$

for all $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$.
Output: sequence of approximations $\{\boldsymbol{m}_h^i\}_{i \in \mathbb{N}_0}$. ☐

With the sequence of approximations $\{\boldsymbol{m}_h^i\}_{i \in \mathbb{N}_0}$ delivered by Algorithm 3.1, we define the piecewise linear time reconstruction $\boldsymbol{m}_{hk}$ via (3.4). In the following theorem, we establish the stability and convergence of the approximations obtained with Algorithm 3.1. The proof is postponed to Section 5.

**Theorem 3.2.** *The following statements hold.*
(i) *Suppose that $\boldsymbol{m}_h^0 \in \mathcal{M}_h$. Then, for all $i \in \mathbb{N}_0$, (3.5) admits a solution $\boldsymbol{m}_h^{i+1} \in \mathcal{M}_h$. In particular, the scheme preserves the unit-length constraint at any time step at the nodes of the triangulation.*
(ii) *The scheme is unconditionally stable in the sense that, for all $J \in \mathbb{N}$, it holds that*

$$\mathcal{E}(\boldsymbol{m}_h^J) + \alpha k \sum_{i=0}^{J-1} \|d_t \boldsymbol{m}_h^i\|_h^2 = \mathcal{E}(\boldsymbol{m}_h^0). \tag{3.6}$$

(iii) *Suppose that $\boldsymbol{m}_h^0 \to \boldsymbol{m}^0$ in $\boldsymbol{H}^1(\Omega)$ as $h \to 0$. Then there exist a global weak solution $\boldsymbol{m} \colon \Omega \times \mathbb{R}_{>0} \to \mathbb{S}^2$ of (2.4) in the sense of Definition 2.1 and a subsequence of $\{\boldsymbol{m}_{hk}\}$ (not relabeled) which unconditionally converges towards $\boldsymbol{m}$. Specifically, as $h, k \to 0$, $\boldsymbol{m}_{hk} \overset{\star}{\rightharpoonup} \boldsymbol{m}$ in $L^\infty(\mathbb{R}_{>0}; \boldsymbol{H}^1(\Omega; \mathbb{S}^2))$ and $\boldsymbol{m}_{hk}|_{\Omega_T} \rightharpoonup \boldsymbol{m}|_{\Omega_T}$ in $\boldsymbol{H}^1(\Omega_T)$ for all $T > 0$.*

**Remark 3.3.** Note that, differently from the corresponding estimates for tangent plane schemes [4, 8, 22, 30], the stability result for Algorithm 3.1 (Theorem 3.2 (ii)) does not require any geometric assumption on the mesh. Moreover, (3.6) holds with equality and without any artificial dissipative term on the left-hand side.

Theorem 3.2 (i) establishes unconditional existence of a solution of (3.5), but does not provide information about its uniqueness. If $k = o(h^2)$, one can show that a suitable fixed-point iteration is a contraction provided that the discretization parameters are sufficiently small. With the Banach fixed-point theorem, this implies that each time step of Algorithm 3.1 is well-posed.

**Proposition 3.4.** *Suppose that $k = o(h^2)$ as $h, k \to 0$. Then there exist thresholds $h_0 > 0$ and $k_0 > 0$ such that, for all $h < h_0$ and $k < k_0$, the variational problem (3.5) admits a unique solution $\boldsymbol{m}_h^{i+1} \in \mathcal{M}_h$ for all $i \in \mathbb{N}_0$.*

The proof of Proposition 3.4 can be obtained simplifying the argument of the proofs of Proposition 3.5 and Theorem 3.7 below; therefore, we omit it.

## 3.3 Practical Midpoint Schemes

Each time step of Algorithm 3.1 requires the solution of a *nonlinear* system and the computation of *nonlocal* field contributions.

Nonlinearity is a consequence of the first term on the right-hand side of (3.5). The second term on the right-hand side, at first glance also nonlinear in $m_h^{i+1}$, turns out to be actually linear. Indeed, it holds that

$$m_h^{i+1/2} \times d_t m_h^{i+1} \stackrel{(3.3)}{=} \frac{m_h^{i+1} + m_h^i}{2} \times \frac{m_h^{i+1} - m_h^i}{k} = -\frac{1}{k} m_h^{i+1} \times m_h^i.$$

However, using an arbitrary off-the-shelf nonlinear solver for (3.5), the conservation and stability properties of Algorithm 3.1 established in Theorem 3.2 (i)–(ii) are in general lost. Moreover, $\pi$ can be nonlocal and non-exactly computable (e.g., for the stray field) so that the field contribution $\pi(m_h^{i+1/2})$ must be numerically approximated. Hence, a direct implementation of Algorithm 3.1 should be based on an inner iteration performing the solution of the nonlinear system (3.5) and the approximate computation of $\pi(m_h^{i+1/2})$.

In the remainder of this section, we discuss and analyze an effective treatment of the nonlocal contribution, which we combine with two approaches for the linearization of (3.5), from which we will obtain two *practical midpoint schemes*.

To start with, we define the bilinear form $a^{\mathrm{loc}} \colon H^1(\Omega) \times H^1(\Omega) \to \mathbb{R}$ by

$$a^{\mathrm{loc}}(\boldsymbol{\psi}, \boldsymbol{\phi}) = \sum_{d=1}^3 \langle A_d(\partial_d \boldsymbol{\psi} - J_d \boldsymbol{\psi}), \partial_d \boldsymbol{\phi} - J_d \boldsymbol{\phi} \rangle_\Omega \quad \text{for all } \boldsymbol{\psi}, \boldsymbol{\phi} \in H^1(\Omega). \tag{3.7}$$

We consider the local parts of the energy and the effective field given by

$$\mathcal{E}^{\mathrm{loc}}(m) := \frac{1}{2} a^{\mathrm{loc}}(m, m) - \langle f, m \rangle_\Omega \stackrel{(2.1)}{=} \mathcal{E}(m) + \frac{1}{2} \langle \pi(m), m \rangle_\Omega$$

and

$$
\begin{aligned}
-\langle h_{\mathrm{eff}}^{\mathrm{loc}}(m), \boldsymbol{\phi} \rangle := \left\langle \frac{\delta \mathcal{E}^{\mathrm{loc}}(m)}{\delta m}, \boldsymbol{\phi} \right\rangle &\stackrel{(3.7)}{=} a^{\mathrm{loc}}(m, \boldsymbol{\phi}) - \langle f, \boldsymbol{\phi} \rangle_\Omega \\
&\stackrel{(2.2)}{=} a(m, \boldsymbol{\phi}) + \langle \pi(m), \boldsymbol{\phi} \rangle_\Omega - \langle f, \boldsymbol{\phi} \rangle_\Omega \\
&\stackrel{(2.3)}{=} -\langle h_{\mathrm{eff}}(m), \boldsymbol{\phi} \rangle + \langle \pi(m), \boldsymbol{\phi} \rangle_\Omega,
\end{aligned}
$$

respectively. Then, for $i \in \mathbb{N}_0$, we rewrite (3.5) in terms of the new unknown $\eta_h^i := m_h^{i+1/2} \in V_h$. Since $d_t m_h^{i+1} = 2(\eta_h^i - m_h^i)/k$, it is easy to see that (3.5) is equivalent to the following problem: first, compute $\eta_h^i \in V_h$ such that, for all $\boldsymbol{\phi}_h \in V_h$, it holds that

$$\langle \eta_h^i, \boldsymbol{\phi}_h \rangle_h + \frac{k}{2} \langle \eta_h^i \times \mathbb{P}_h h_{\mathrm{eff}}^{\mathrm{loc}}(\eta_h^i), \boldsymbol{\phi}_h \rangle_h + \frac{k}{2} \langle \eta_h^i \times \mathbb{P}_h \pi(\eta_h^i), \boldsymbol{\phi}_h \rangle_h + \alpha \langle \eta_h^i \times m_h^i, \boldsymbol{\phi}_h \rangle_h = \langle m_h^i, \boldsymbol{\phi}_h \rangle_h. \tag{3.8}$$

Then define

$$m_h^{i+1} := 2\eta_h^i - m_h^i. \tag{3.9}$$

To treat the nonlocal contribution $\pi(\eta_h^i)$, we adopt the implicit-explicit (IMEX) approach introduced in [39]. Let $\pi_h \colon V_h \to V_h$ be an operator approximating $\pi$, assumed to be linear and uniformly bounded in $L^2(\Omega)$ in the sense that $\|\pi_h\|_{L(L^2(\Omega);L^2(\Omega))} \le C_\pi$ for some $C_\pi > 0$ independent of $h$. Moreover, we say that $\pi_h$ is *consistent* with $\pi$ if, for all $\boldsymbol{\phi} \in L^2(\Omega)$ and all $(\boldsymbol{\phi}_h)_{h>0} \subset V_h$ with $\boldsymbol{\phi}_h \to \boldsymbol{\phi}$ in $L^2(\Omega)$ as $h \to 0$, it holds that

$$\pi_h(\boldsymbol{\phi}_h) \to \pi(\boldsymbol{\phi}) \quad \text{in } L^2(\Omega) \quad \text{as } h \to 0. \tag{3.10}$$

We define $m_h^{-1} := m_h^0$ and

$$\Pi_h(m_h^i, m_h^{i-1}) := \frac{3}{2} \pi_h(m_h^i) - \frac{1}{2} \pi_h(m_h^{i-1}).$$

Then, in (3.8), we replace $\pi(\eta_h^i)$ with its approximation $\Pi_h(m_h^i, m_h^{i-1})$ to obtain

$$
\begin{aligned}
&\langle \eta_h^i, \boldsymbol{\phi}_h \rangle_h + \frac{k}{2} \langle \eta_h^i \times \mathbb{P}_h h_{\mathrm{eff}}^{\mathrm{loc}}(\eta_h^i), \boldsymbol{\phi}_h \rangle_h + \frac{k}{2} \langle \eta_h^i \times \mathbb{P}_h \Pi_h(m_h^i, m_h^{i-1}), \boldsymbol{\phi}_h \rangle_h \\
&\quad + \alpha \langle \eta_h^i \times m_h^i, \boldsymbol{\phi}_h \rangle_h = \langle m_h^i, \boldsymbol{\phi}_h \rangle_h.
\end{aligned}
\tag{3.11}
$$

In particular, the nonlocal contribution, treated explicitly, becomes independent of the unknown $\boldsymbol{\eta}_h^i$. We now discuss two strategies to linearize (3.11) in order to arrive at two practical midpoint schemes. To emphasize the inexact solution of (3.11) up to some accuracy $\varepsilon > 0$, we write $\boldsymbol{m}_{h\varepsilon}^i$ rather than $\boldsymbol{m}_h^i$ for the iterates of the practical (linearized) midpoint schemes.

### 3.3.1 Constraint-Preserving Fixed-Point Iteration

We solve (3.11) with the following fixed-point iteration. Let $\varepsilon > 0$ denote some prescribed tolerance. Set $\boldsymbol{\eta}_h^{i,0} := \boldsymbol{m}_{h\varepsilon}^i$. For $\ell \in \mathbb{N}_0$, given $\boldsymbol{\eta}_h^{i,\ell} \in \boldsymbol{V}_h$, compute $\boldsymbol{\eta}_h^{i,\ell+1} \in \boldsymbol{V}_h$ such that, for all $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$, it holds that

$$\langle \boldsymbol{\eta}_h^{i,\ell+1}, \boldsymbol{\phi}_h \rangle_h + \frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell}), \boldsymbol{\phi}_h \rangle_h + \frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h \Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}), \boldsymbol{\phi}_h \rangle_h$$
$$+ \alpha\langle \boldsymbol{\eta}_h^{i,\ell+1} \times \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h = \langle \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h, \tag{3.12}$$

until

$$\big\| \mathfrak{I}_h[\boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell}))] \big\|_h \leq \varepsilon, \tag{3.13}$$

where $\mathfrak{I}_h[\cdot]$ denotes the vector-valued nodal interpolant. If $\ell^* \in \mathbb{N}_0$ is the smallest integer for which the stopping criterion (3.13) is satisfied, in view of (3.9), the approximate magnetization at the new time step is defined as $\boldsymbol{m}_{h\varepsilon}^{i+1} := 2\boldsymbol{\eta}_h^{i,\ell^*+1} - \boldsymbol{m}_{h\varepsilon}^i$.

In the following proposition, we collect the properties of the proposed fixed-point iteration. The proof is postponed to Section 6.

**Proposition 3.5.** *Let $i \in \mathbb{N}_0$.*
(i) *For all $\ell \in \mathbb{N}_0$, the variational problem* (3.12) *admits a unique solution $\boldsymbol{\eta}_h^{i,\ell+1} \in \boldsymbol{V}_h$. Moreover, it holds that*
$\|\boldsymbol{\eta}_h^{i,\ell+1}\|_{\boldsymbol{L}^\infty(\Omega)} \leq 1$.
(ii) *If $k = o(h^2)$ as $h, k \to 0$, there exist a contraction constant $0 < q < 1$ and thresholds $h_0, k_0 > 0$ such that, for all $h < h_0$ and $k < k_0$, it holds that*

$$\|\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}\|_h \leq q\|\boldsymbol{\eta}_h^{i,\ell+1} - \boldsymbol{\eta}_h^{i,\ell}\|_h \quad \text{for all } \ell \in \mathbb{N}_0. \tag{3.14}$$

*The constants $q, h_0, k_0$ depend only on the mesh parameter $\kappa$ 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 $\ell^* \in \mathbb{N}_0$ denotes the smallest integer for which* (3.13) *is satisfied, the new approximation*

$$\boldsymbol{m}_{h\varepsilon}^{i+1} := 2\boldsymbol{\eta}_h^{i,\ell^*+1} - \boldsymbol{m}_{h\varepsilon}^i \in \boldsymbol{V}_h$$

*belongs to $\mathcal{M}_h$.*

For all $i \in \mathbb{N}_0$, let $\boldsymbol{r}_{h\varepsilon}^i := \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell^*+1}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell^*})) \in \boldsymbol{V}_h$. Because of the stopping criterion (3.13), it holds that $\|\mathfrak{I}_h[\boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i]\|_h \leq \varepsilon$. With this definition, the proposed linearization of Algorithm 3.1 is covered by the following algorithm.

**Algorithm 3.6** (Practical Midpoint Scheme, Constraint-Preserving Fixed-Point Iteration).
Input: $\boldsymbol{m}_{h\varepsilon}^0 := \boldsymbol{m}_h^0 \in \mathcal{M}_h$.
Loop: for all $i \in \mathbb{N}_0$, use the constraint-preserving fixed-point iteration (3.12)–(3.13) to compute $\boldsymbol{m}_{h\varepsilon}^{i+1} \in \mathcal{M}_h$ and $\boldsymbol{r}_{h\varepsilon}^i \in \boldsymbol{V}_h$ with $\|\mathfrak{I}_h[\boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i]\|_h \leq \varepsilon$ such that

$$\langle d_t\boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\phi}_h \rangle_h = -\langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \mathbb{P}_h \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}), \boldsymbol{\phi}_h \rangle_h + \alpha\langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times d_t\boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\phi}_h \rangle_h$$
$$+ \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times [\boldsymbol{r}_{h\varepsilon}^i + \mathbb{P}_h(\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}))], \boldsymbol{\phi}_h \rangle_h \tag{3.15}$$

for all $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$.
Output: sequence of approximations $\{\boldsymbol{m}_{h\varepsilon}^i\}_{i\in\mathbb{N}_0}$. □

In the following theorem, we establish the stability and convergence of the approximations obtained with Algorithm 3.6. The proof is postponed to Section 6.

**Theorem 3.7.** *The following statements hold.*

(i) *Suppose that $\boldsymbol{m}_h^0 \in \mathcal{M}_h$. If $k = o(h^2)$ as $h, k \to 0$, there exist thresholds $h_0 > 0$ and $k_0 > 0$ such that, for all $h < h_0$ and $k < k_0$, (3.15) admits solutions $\boldsymbol{m}_{h\varepsilon}^{i+1} \in \mathcal{M}_h$ and $\boldsymbol{r}_{h\varepsilon}^i \in V_h$ with $\|\mathfrak{I}_h[\boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i]\|_h \le \varepsilon$ for all $i \in \mathbb{N}_0$. In particular, the scheme preserves the unit-length constraint at the nodes of the triangulation for all time steps. The thresholds $h_0, k_0$ depend only on the mesh parameter $\kappa$ and the problem data.*

(ii) *Under the assumptions of part (i), for all $h < h_0$, $k < k_0$, and $J \in \mathbb{N}$, the scheme satisfies the discrete energy identity*

$$\mathcal{E}(\boldsymbol{m}_{h\varepsilon}^J) + \alpha k \sum_{i=0}^{J-1} \|d_t \boldsymbol{m}_{h\varepsilon}^{i+1}\|_h^2 = \mathcal{E}(\boldsymbol{m}_h^0) - k \sum_{i=0}^{J-1} \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times [\boldsymbol{r}_{h\varepsilon}^i + \mathbb{P}_h(\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}))],$$
$$\mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \alpha d_t \boldsymbol{m}_{h\varepsilon}^{i+1} \rangle_h. \tag{3.16}$$

(iii) *Let $T > 0$. Let $J \in \mathbb{N}$ be the smallest integer such that $T \le kJ$. Under the assumptions of part (i), if $\varepsilon = \mathcal{O}(h)$ and $\{\boldsymbol{m}_h^0\}_{h>0}$ is bounded in $\boldsymbol{H}^1(\Omega)$ as $h, \varepsilon \to 0$, there exist thresholds $0 < h_0^* \le h_0$, $0 < k_0^* \le k_0$, and $\varepsilon_0^* > 0$ such that, for all $h < h_0^*$, $k < k_0^*$, $\varepsilon < \varepsilon_0^*$, and $1 \le j \le J$, we have the stability estimate*

$$\|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{H}^1(\Omega)}^2 + k \sum_{i=0}^{j-1} \|d_t \boldsymbol{m}_{h\varepsilon}^{i+1}\|_{\boldsymbol{L}^2(\Omega)}^2 \le C. \tag{3.17}$$

*The constant $C > 0$ and the thresholds $h_0^*, k_0^*, \varepsilon_0^*$ depend only on the mesh parameter $\kappa$, the final time $T$, and the problem data.*

(iv) *In addition to the assumptions of part (iii), assume $\boldsymbol{m}_h^0 \to \boldsymbol{m}^0$ in $\boldsymbol{H}^1(\Omega)$ as $h \to 0$, and suppose that $\boldsymbol{\pi}_h$ is consistent (3.10) with $\boldsymbol{\pi}$. Then there exist $\boldsymbol{m} \in \boldsymbol{H}^1(\Omega_T) \cap L^\infty(0, T; \boldsymbol{H}^1(\Omega; \mathbb{S}^2))$ and a subsequence of $\{\boldsymbol{m}_{h\varepsilon k}\}$ (not relabeled) which converges towards $\boldsymbol{m}$ as $h, k, \varepsilon \to 0$. Specifically,*

$$\boldsymbol{m}_{h\varepsilon k}|_{\Omega_T} \stackrel{\star}{\rightharpoonup} \boldsymbol{m} \quad \text{in } L^\infty(0, T; \boldsymbol{H}^1(\Omega; \mathbb{S}^2)) \quad \text{and} \quad \boldsymbol{m}_{h\varepsilon k}|_{\Omega_T} \rightharpoonup \boldsymbol{m} \quad \text{in } \boldsymbol{H}^1(\Omega_T) \quad \text{as } h, k, \varepsilon \to 0.$$

*The limit function $\boldsymbol{m}$ satisfies conditions* (i)–(iv) *of Definition 2.1.*

### 3.3.2 Newton Iteration

Based on the Newton scheme, in [11, Section 1.4.1], the authors employ a linearization of the nonlinear system (3.5) in the ideal midpoint scheme with simplified effective field, i.e., without nonlocal contributions and without DMI. Their 2D numerical experiments give hope for a less restrictive CFL condition than for the fixed-point iteration from Section 3.3.1.

For three-dimensional micromagnetics and considering the full effective field (2.3), in Section 7.3, we apply Newton's method to the nonlinear system of equations (3.11) resulting in the following iteration. Let $\varepsilon > 0$ denote some tolerance. Set $\boldsymbol{\eta}_h^{i,0} := \boldsymbol{m}_{h\varepsilon}^i$. For $\ell \in \mathbb{N}_0$, given $\boldsymbol{\eta}_h^{i,\ell} \in V_h$, compute $\boldsymbol{u}_h^{i,\ell} \in V_h$ such that, for all $\boldsymbol{\phi}_h \in V_h$, it holds that

$$\langle \boldsymbol{u}_h^{i,\ell}, \boldsymbol{\phi}_h \rangle_h + \frac{k}{2} \langle \boldsymbol{u}_h^{i,\ell} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell}) + \Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})), \boldsymbol{\phi}_h \rangle_h$$
$$+ \alpha \langle \boldsymbol{u}_h^{i,\ell} \times \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h + \frac{k}{2} \langle \boldsymbol{\eta}_h^{i,\ell} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{u}_h^{i,\ell}) - \boldsymbol{f}), \boldsymbol{\phi}_h \rangle_h$$
$$= \langle \boldsymbol{m}_{h\varepsilon}^i - \boldsymbol{\eta}_h^{i,\ell}, \boldsymbol{\phi}_h \rangle_h - \frac{k}{2} \langle \boldsymbol{\eta}_h^{i,\ell} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\eta}_h^{i,\ell}) + \Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})), \boldsymbol{\phi}_h \rangle_h - \alpha \langle \boldsymbol{\eta}_h^{i,\ell} \times \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h, \tag{3.18}$$

and define $\boldsymbol{\eta}_h^{i,\ell+1} := \boldsymbol{\eta}_h^{i,\ell} + \boldsymbol{u}_h^{i,\ell}$ until

$$\|\mathfrak{I}_h[\boldsymbol{u}_h^{i,\ell} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{u}_h^{i,\ell}) - \boldsymbol{f})]\|_h \le \varepsilon. \tag{3.19}$$

If $\ell^* \in \mathbb{N}_0$ is the smallest integer for which the stopping criterion (3.19) is satisfied, the approximate magnetization at the new time step is defined as $\boldsymbol{m}_{h\varepsilon}^{i+1} := 2\boldsymbol{\eta}_h^{i,\ell^*+1} - \boldsymbol{m}_{h\varepsilon}^i$.

For all $i \in \mathbb{N}_0$, let $\boldsymbol{r}_{h\varepsilon}^i = \mathfrak{I}_h[\boldsymbol{u}_h^{i,\ell^*} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{u}_h^{i,\ell^*}) - \boldsymbol{f})] \in V_h$. In view of the stopping criterion (3.19), it holds that $\|\boldsymbol{r}_{h\varepsilon}^i\|_h \le \varepsilon$. With this definition, the proposed linearization of Algorithm 3.1 based on the Newton method is covered by the following algorithm.

**Algorithm 3.8** (Practical Midpoint Scheme, Newton Iteration).  Input: $\boldsymbol{m}_{h\varepsilon}^0 := \boldsymbol{m}_h^0 \in \boldsymbol{V}_h$.
Loop:  for all $i \in \mathbb{N}_0$, use Newton's method (3.18)–(3.19) with initial guess $\boldsymbol{\eta}_h^{i,0} := \boldsymbol{m}_{h\varepsilon}^i$ to compute $\boldsymbol{m}_{h\varepsilon}^{i+1} \in \boldsymbol{V}_h$
  and $\boldsymbol{r}_{h\varepsilon}^i \in \boldsymbol{V}_h$ with $\|\boldsymbol{r}_{h\varepsilon}^i\|_h \leq \varepsilon$ such that

$$\langle d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\phi}_h \rangle_h = -\langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \mathbb{P}_h \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}), \boldsymbol{\phi}_h \rangle_h + \alpha \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\phi}_h \rangle_h$$
$$+ \langle \boldsymbol{r}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h + \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \mathbb{P}_h(\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})), \boldsymbol{\phi}_h \rangle_h \qquad (3.20)$$

  for all $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$.
Output:  sequence of approximations $\{\boldsymbol{m}_{h\varepsilon}^i\}_{i \in \mathbb{N}_0}$.                                                          □

The results on Algorithm 3.8 are stated in Lemma 3.9 ($\boldsymbol{L}^\infty$-bound), Theorem 3.10 (stability), and Theorem 3.11 (well-posedness) below. Compared to Algorithm 3.6, our analysis is more involved: precisely, for $i < J$, the proof of Theorem 3.11 requires $i$-independent bounds on $\|\boldsymbol{m}_{h\varepsilon}^i\|_{\boldsymbol{L}^\infty(\Omega)}$ and $\mathcal{E}(\boldsymbol{m}_{h\varepsilon}^i)$ in order to guarantee well-posedness of Algorithm 3.8, while Lemma 3.9 and Theorem 3.10 require termination of (3.18)–(3.19) so that $\boldsymbol{m}_{h\varepsilon}^{i+1}$ is well-defined.

   In contrast to the fixed-point iteration (3.12)–(3.13), the Newton iteration (3.18)–(3.19) does not inherently preserve discrete unit length, i.e., in general, $|\boldsymbol{m}_{h\varepsilon}^{i+1}(\boldsymbol{z})| \neq |\boldsymbol{m}_{h\varepsilon}^i(\boldsymbol{z})|$ for $\boldsymbol{z} \in \mathcal{N}_h$. However, assuming well-posedness of Algorithm 3.8, in the following lemma, we establish uniform $\boldsymbol{L}^\infty(\Omega)$-boundedness of the approximations obtained with Algorithm 3.8.

**Lemma 3.9.**  *Suppose $h, k, \varepsilon > 0$, $\boldsymbol{m}_h^0 \in \mathcal{M}_h$, and let $J \in \mathbb{N}$ be the smallest integer such that $T \leq kJ$. Let $0 \leq i < J$, and suppose that the Newton iteration (3.18)–(3.19) in Algorithm 3.8 terminates for all $0 \leq n \leq i$, i.e., the sequences $\{\boldsymbol{m}_{h\varepsilon}^n\}_{n=0}^{i+1}, \{\boldsymbol{r}_{h\varepsilon}^n\}_{n=0}^i \subset \boldsymbol{V}_h$ are the output of Algorithm 3.8 and satisfy (3.20) with $\|\boldsymbol{r}_{h\varepsilon}^n\|_h \leq \varepsilon$ for all $0 \leq n \leq i$.*

(i) *If $\varepsilon = \mathcal{O}(h^{3/2})$ as $h, k, \varepsilon \to 0$, then there exist a constant $C_\infty > 0$ and thresholds $h_0 > 0$, $k_0 > 0$, and $\varepsilon_0 > 0$ such that, for all $h < h_0$, $k < k_0$, and $\varepsilon < \varepsilon_0$, it holds that $\|\boldsymbol{m}_{h\varepsilon}^{n+1}\|_{\boldsymbol{L}^\infty(\Omega)} \leq C_\infty$ uniformly for all $0 \leq n \leq i$. The thresholds $h_0, k_0, \varepsilon_0$ depend only on the mesh parameter $\kappa$ and the problem data, while the bound $C_\infty > 0$ depends only on $\kappa$, $\varepsilon h^{-3/2} \lesssim 1$, and the final time $T > 0$, but not on the integer $i < J$.*

(ii) *If $\varepsilon = o(h^{3/2})$, then there holds $\|\mathcal{I}_h(|\boldsymbol{m}_{h\varepsilon k}|^2) - 1\|_{L^\infty([0, t_{i+1}] \times \Omega)} \to 0$ as $h, k, \varepsilon \to 0$.*

Assuming well-posedness of Algorithm 3.8, in the following theorem, we establish the stability and convergence of the approximations obtained with Algorithm 3.8. The proof is postponed to Section 7.1.

**Theorem 3.10.**  *Let $T > 0$, and suppose that $\{\mathcal{I}_h\}_{h>0}$ is a $\kappa$-quasi-uniform family of triangulations. Suppose $h, k, \varepsilon > 0$, $\boldsymbol{m}_h^0 \in \mathcal{M}_h$, and let $J \in \mathbb{N}$ be the smallest integer such that $T \leq kJ$. Let $0 \leq i < J$, and suppose that the Newton iteration (3.18)–(3.19) in Algorithm 3.8 terminates for all $0 \leq n \leq i$, i.e., the sequences $\{\boldsymbol{m}_{h\varepsilon}^n\}_{n=0}^{i+1}, \{\boldsymbol{r}_{h\varepsilon}^n\}_{n=0}^i \subset \boldsymbol{V}_h$ are the output of Algorithm 3.8 and satisfy (3.20) with $\|\boldsymbol{r}_{h\varepsilon}^n\|_h \leq \varepsilon$ for all $0 \leq n \leq i$.*
(i) *Under these assumptions, the scheme satisfies the discrete energy identity*

$$\mathcal{E}(\boldsymbol{m}_{h\varepsilon}^{i+1}) + \alpha k \sum_{n=0}^i \|d_t \boldsymbol{m}_{h\varepsilon}^{n+1}\|_h^2 = \mathcal{E}(\boldsymbol{m}_h^0) - k \sum_{n=0}^i \langle \boldsymbol{r}_{h\varepsilon}^n + \boldsymbol{m}_{h\varepsilon}^{n+1/2} \times \mathbb{P}_h(\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{n+1/2}) - \Pi_h(\boldsymbol{m}_{h\varepsilon}^n, \boldsymbol{m}_{h\varepsilon}^{n-1})),$$
$$\mathbb{P}_h \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{n+1/2}) - \alpha d_t \boldsymbol{m}_{h\varepsilon}^{n+1} \rangle_h.$$

(ii) *If $k = o(h^2)$, $\varepsilon = \mathcal{O}(h^{3/2})$, and $\{\boldsymbol{m}_h^0\}_{h>0}$ is bounded in $\boldsymbol{H}^1(\Omega)$ as $k, h, \varepsilon \to 0$, there exist thresholds $h_0 > 0$, $k_0 > 0$, and $0 < \varepsilon_0 \leq \alpha$ such that, for all $h < h_0$, $k < k_0$, $\varepsilon < \varepsilon_0$, we have the stability estimate*

$$\|\boldsymbol{m}_{h\varepsilon}^{i+1}\|_{\boldsymbol{H}^1(\Omega)}^2 + k \sum_{n=0}^i \|d_t \boldsymbol{m}_{h\varepsilon}^{n+1}\|_{\boldsymbol{L}^2(\Omega)}^2 \leq C.$$

*The constant $C > 0$ and the thresholds $h_0, k_0, \varepsilon_0$ depend only on the mesh parameter $\kappa$, the final time $T$, and the problem data.*

(iii) *In addition to the assumptions of part (ii), suppose that $\boldsymbol{m}_h^0 \to \boldsymbol{m}^0$ in $\boldsymbol{H}^1(\Omega)$ as $h \to 0$, and that $\boldsymbol{\pi}_h$ is consistent (3.10) with $\boldsymbol{\pi}$. Then there exist $\boldsymbol{m} \in \boldsymbol{H}^1(\Omega_T) \cap L^\infty(0, T; \boldsymbol{H}^1(\Omega; \mathbb{S}^2))$ and a subsequence of $\{\boldsymbol{m}_{h\varepsilon k}\}$ (not relabeled) which converges towards $\boldsymbol{m}$ as $h, k, \varepsilon \to 0$. Specifically, $\boldsymbol{m}_{h\varepsilon k} \overset{*}{\rightharpoonup} \boldsymbol{m}$ in $L^\infty(0, T; \boldsymbol{H}^1(\Omega; \mathbb{S}^2))$ and $\boldsymbol{m}_{h\varepsilon k} \rightharpoonup \boldsymbol{m}$ in $\boldsymbol{H}^1(\Omega_T)$ as $h, k, \varepsilon \to 0$. The limit function $\boldsymbol{m}$ satisfies conditions* (i)–(iv) *of Definition 2.1.*

The following theorem guarantees that, under appropriate CFL conditions, Algorithm 3.8 is well-posed, which is required by Lemma 3.9 and Theorem 3.10.

**Theorem 3.11.** *Let $T > 0$. Suppose $h, k, \varepsilon > 0$, $\boldsymbol{m}_h^0 \in \mathcal{M}_h$, and let $J \in \mathbb{N}$ be the smallest integer such that $T \leq kJ$.*

(i) *If $k = o(h^{7/3})$ and $\varepsilon = \mathcal{O}(h^{3/2})$ as $h, k, \varepsilon \to 0$, then there exist thresholds $h_0 > 0$, $k_0 > 0$, and $\varepsilon_0 > 0$ such that, for all $h < h_0$, $k < k_0$, and $\varepsilon < \varepsilon_0$, Algorithm 3.8 is well-defined, i.e., for all $i = 0, \ldots, J - 1$, it provides after finitely many iterations of Newton's method (3.18)–(3.19) solutions $\boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{r}_{h\varepsilon}^i \in \boldsymbol{V}_h$ to (3.20) with $\|\boldsymbol{r}_{h\varepsilon}^i\|_h \leq \varepsilon$.*

(ii) *In particular, there exists a constant $C_\star > 0$ such that the number of Newton iterations (3.18)–(3.19) required to solve (3.20) is bounded by $\log_2 \log_2 (C_\star k h^{-7/2} \varepsilon^{-1})$. The thresholds $h_0, k_0, \varepsilon_0$ and the constant $C_\star$ depend only on the mesh parameter $\kappa$ and the problem data.*

### 3.3.3 Coupling Conditions on Practical Midpoint Schemes

While the ideal midpoint scheme (Algorithm 3.1) is unconditionally convergent towards a weak solution of LLG, the analysis of the practical midpoint schemes (Algorithm 3.6 and Algorithm 3.8) crucially relies on CFL conditions imposed on the discretization parameters $h, k, \varepsilon > 0$. We conclude this section by Table 1, giving an overview on the imposed coupling conditions sufficient to establish a rigorous analysis of the practical midpoint schemes.

|  | **Fixed-point linearization**<br>**Algorithm 3.6** | **Newton linearization**<br>**Algorithm 3.8** |
| --- | --- | --- |
| Well-posedness | $k = o(h^2)$ | $k = o(h^{7/3})$, $\varepsilon = \mathcal{O}(h^{3/2})$ |
| $\boldsymbol{L}^\infty(\Omega)$-bound | none | $\varepsilon = \mathcal{O}(h^{3/2})$ |
| Stability | $\varepsilon = \mathcal{O}(h)$ | $\varepsilon = \mathcal{O}(h^{3/2})$ |
| Convergence | $\varepsilon = \mathcal{O}(h)$ | $\varepsilon = \mathcal{O}(h^{3/2})$ |
| Total | $k = o(h^2)$, $\varepsilon = \mathcal{O}(h)$ | $k = o(h^{7/3})$, $\varepsilon = \mathcal{O}(h^{3/2})$ |

**Table 1:** Sufficient CFL conditions for the analysis of the practical midpoint schemes of Section 3.3.

# 4 Numerical Experiments

The goal of this section is threefold. First, in Section 4.1, we verify the extension of the midpoint scheme to the DMI contribution and its correct implementation by simulating an experiment on skyrmion dynamics from [42]. The simulation results with the midpoint scheme are compared to the results from [42] and to simulations with the tangent plane scheme from [30]. In Section 4.2, we introduce a variation of the experiment from [42] in order to compare reliability of the midpoint scheme to the generally cheaper tangent plane schemes in simulating sensitive skyrmion dynamics susceptible to slight (artificial) disturbances. By doing this, we emphasize the advantages of discrete energy conservation realized by the midpoint scheme. Finally, in an academic setting, the CFL conditions arising from our analysis sufficient to prove well-posedness of the practical midpoint schemes are experimentally verified in Section 4.3. In particular, the numerical CFL study hints that the CFL condition $k = o(h^{7/3})$ derived for the practical midpoint scheme based on the Newton iteration is likely pessimistic and might be weakened to $k = o(h^2)$ with a sharper analysis. Moreover, we compare the number of iterations in the nonlinear solvers of the two practical midpoint schemes, as well as the impact of the solver accuracy $\varepsilon > 0$ on the deviation from discrete unit length. All experiments in this section were performed with Commics [36, 37].

## 4.1 Stability of Isolated Skyrmions in Nanodisks

To validate the extension of the midpoint scheme incorporating the DMI contribution, we reproduce a numerical experiment from [42] for both the practical midpoint scheme based on the constraint-preserving fixed-point iteration (Algorithm 3.6) and the practical midpoint scheme based on Newton's method (Algorithm 3.8). There, the relaxed states of a thin nanodisk of diameter 80 nm (aligned with $x_1 x_2$-plane) and thickness 0.4 nm ($x_3$-direction) centered at $(0, 0, 0)$ for different values of the DMI constant are investigated. The effective field consists of exchange interaction, perpendicular uniaxial anisotropy, interfacial DMI, and stray field, i.e.,

$$\boldsymbol{H}_{\text{eff}}(\boldsymbol{m}) = \frac{2A}{\mu_0 M_{\text{s}}} \Delta \boldsymbol{m} + \frac{2K}{\mu_0 M_{\text{s}}} (\boldsymbol{a} \cdot \boldsymbol{m}) \boldsymbol{a} - \frac{2D}{\mu_0 M_{\text{s}}} \begin{pmatrix} -\partial_1 m_3 \\ -\partial_2 m_3 \\ \partial_1 m_1 + \partial_2 m_2 \end{pmatrix} + \boldsymbol{H}_{\text{s}}(\boldsymbol{m}).$$

The involved material parameters mimic those of cobalt: $M_{\text{s}} = 5.8 \cdot 10^5$ A/m, $\alpha = 0.3$, $A = 1.5 \cdot 10^{-11}$ J/m, $K = 8 \cdot 10^5$ J/m$^3$, and $\boldsymbol{a} = (0, 0, 1)$. For the DMI constant, the range $D = 0, 1, \ldots, 8$ mJ/m$^2$ is considered. The initial condition is a skyrmion-like state, i.e., given $r = \sqrt{x_1^2 + x_2^2}$, we define $\boldsymbol{m}^0(\boldsymbol{x}) = (0, 0, -1)$ if $r \in [0, 15]$ nm and $\boldsymbol{m}^0(\boldsymbol{x}) = (0, 0, 1)$ if $r \in (15, 40]$ nm. For all simulations, we choose $T = 1$ ns, which experimentally turns out to be a sufficiently large time to relax the system. The computational domain is discretized by a regular partition generated by Netgen [45] consisting of 34 596 tetrahedra and 11 797 vertices, which corresponds to a prescribed mesh size of 1 nm. For the time discretization, we consider a uniform partition of the time interval $(0, T)$ with a time-step size of 2.5 fs. We note that the time-step size has to be chosen considerably smaller than, e.g., for (different variants of) the tangent plane scheme; see our previous work [30, Section 4.3]. This is due to the more restrictive CFL conditions required for convergence of the nonlinear solvers in the practical midpoint schemes; see Theorem 3.7 and Theorem 3.11. The accuracy for the nonlinear solver is chosen as $\varepsilon = 10^{-8}$.

The stable state is a quasi-uniform ferromagnetic state for the values $D = 0, 1, 2$ mJ/m$^2$, a skyrmion for the values $D = 3, \ldots, 6$ mJ/m$^2$, and a multidomain state (target skyrmion) for the values $D = 7, 8$ mJ/m$^2$; see
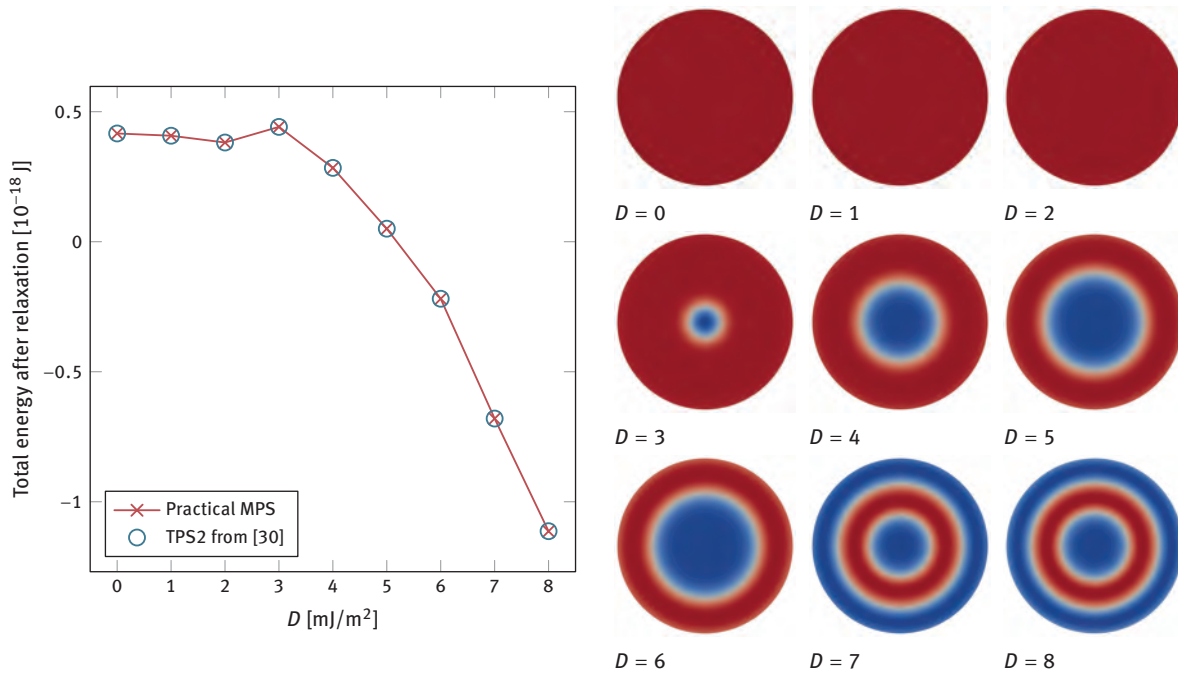


**Figure 1:** Experiment of Section 4.1. Left: final energy for different values of the DMI constant. Right: magnetization component $m_3$ ranging from $-1$ (blue) to $+1$ (red) of the relaxed state for different values of the DMI constant (in mJ/m$^2$). The results computed with either of the practical midpoint schemes, Algorithm 3.6 or Algorithm 3.8, coincide.

Figure 1. The skyrmion size, i.e., the diameter of the circle $\{m_3 = 0\}$ in the $x_1x_2$-plane, increases from the minimum value of circa 14 nm for $D = 3\,\mathrm{mJ/m^2}$ to the maximum value of circa 48 nm for $D = 6\,\mathrm{mJ/m^2}$.

In Figure 1, the relaxed states computed with the practical midpoint scheme for different values of the DMI constant are given. The energy values and the magnetization profiles are in perfect quantitative and qualitative agreement with those reported in [42, Figure 1] and [30, Section 4.2]. This validates both the extension of the midpoint scheme to DMI energy contributions and its implementation in Commics [36, 37].

## 4.2 Reliable Schemes for Energy Sensitive Dynamics

We recall the discrete energy equality (3.6) achieved by the ideal midpoint scheme

$$\mathcal{E}(\boldsymbol{m}_h^J) + \alpha k \sum_{i=0}^{J-1} \|d_t \boldsymbol{m}_h^i\|_h^2 = \mathcal{E}(\boldsymbol{m}_h^0).$$

Differently, for the first-order tangent plane scheme [4], we recite from [30, Proposition 2] the discrete energy inequality

$$\mathcal{E}(\boldsymbol{m}_h^J) + (\alpha - Ck/h)k \sum_{i=0}^{J-1} \|\boldsymbol{v}_h^i\|_h^2 + \ell_{\mathrm{ex}}^2(\theta - 1/2)k^2 \sum_{i=0}^{J-1} \|\nabla \boldsymbol{v}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \leq \mathcal{E}(\boldsymbol{m}_h^0), \tag{4.1}$$

where $\boldsymbol{v}_h^i$ denotes the discrete time derivative computed in the $i$-th time step of the tangent plane scheme to define the update $\boldsymbol{m}_h^{i+1}(\boldsymbol{z}) = (\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z}))/|\boldsymbol{m}_h^i(\boldsymbol{z}) + k\boldsymbol{v}_h^i(\boldsymbol{z})| \in \mathbb{S}^2$ for all $\boldsymbol{z} \in \mathcal{N}_h$. We note that the generic constant $C > 0$ in (4.1) stems from an inverse estimate used in the analysis of [30] to control the discrete energy in presence of a DMI energy contribution. The third term on the left-hand side in (4.1) corresponds to artificial damping introduced by implicit treatment in time of the Laplacian for $1/2 < \theta \leq 1$, while the inequality (instead of equality) is a result of the nodal projection in each time step. As a third integrator, we consider the (almost) second-order tangent plane scheme from [8], which provides a discrete energy inequality, which, although not identical to (4.1), introduces similar artificial energy dissipation due to implicit treatment of the Laplacian and the nodal projection update. For the second-order schemes, i.e., for the midpoint scheme and the second-order tangent plane scheme, an IMEX treatment of the lower-order terms is employed, which results in a perturbation of order $\mathcal{O}(k^2)$ of the respective discrete energy identity [22]. While the discrete energy identity for the midpoint scheme mimics the continuous law

$$\mathcal{E}(\boldsymbol{m}(\tau)) + \alpha \int_0^\tau \|\partial_t \boldsymbol{m}(t)\|_{\boldsymbol{L}^2(\Omega)}^2 \, \mathrm{d}t = \mathcal{E}(\boldsymbol{m}^0),$$

due to the severe CFL condition $k = o(h^2)$, the practical midpoint schemes are very restrictive on the time-step size. In contrast, the tangent plane integrators allow for considerably larger time-step sizes, but introduce artificial damping to the system. Hence, we expect decreased reliability of the tangent plane integrators for accurately simulating processes, which are particularly sensitive to slight inaccuracies in the discrete energy evolution.

To quantify the effects of this artificial damping introduced by the tangent plane integrators, we extend the experiment of Section 4.1: considering the different relaxed states in Figure 1 (right), one infers that, between $D = 2\,\mathrm{mJ/m^2}$ and $D = 3\,\mathrm{mJ/m^2}$, there is a (qualitative) discontinuity, corresponding to a jump in Figure 1 (left) if the resolution on the $D$-axis was increased. Analogously, this applies to the interval from $D = 6\,\mathrm{mJ/m^2}$ to $D = 7\,\mathrm{mJ/m^2}$. The goal of this experiment is the determination of the points of transition $D_{\mathrm{crit}}^{2-3}$ and $D_{\mathrm{crit}}^{6-7}$ from the quasi-uniform relaxed state to the skyrmion state between $D = 2\,\mathrm{mJ/m^2}$ and $D = 3\,\mathrm{mJ/m^2}$, as well as from the skyrmion state to the target skyrmion state between $D = 6\,\mathrm{mJ/m^2}$ and $D = 7\,\mathrm{mJ/m^2}$, respectively. We will evaluate and compare the reliability of the midpoint scheme (MPS), the first-, and the second-order tangent plane scheme (TPS1 and TPS2) in determining $D_{\mathrm{crit}}^{2-3}$ and $D_{\mathrm{crit}}^{6-7}$. For all three schemes, dynamics are simulated with identical parameters.
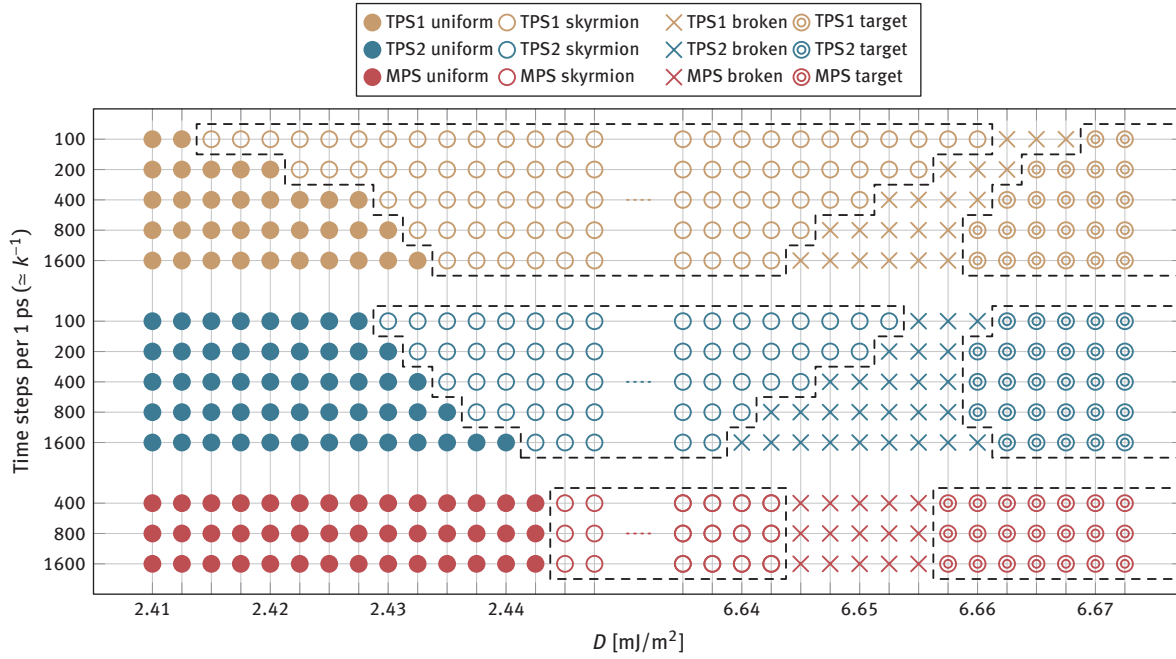
**Figure 2:** Experiment of Section 4.2. Each marker corresponds to one simulation carried out with one of the three integrators, for a DMI parameter $D$ with one particular time-step size $k$. The shape of a marker characterizes the qualitative result after relaxing the skyrmion-like initial state until the equilibrium state is reached, i.e., whether a quasi-uniform, a skyrmion-like, a broken unsymmetrical, or a target skyrmion state is obtained. The results computed with either of the practical midpoint schemes, Algorithm 3.6 or Algorithm 3.8, coincide.

We consider the fixed mesh from Section 4.1. Although this mesh does not satisfy the so-called *angle condition* ensuring validity of (4.1), stability of the tangent plane integrators is still recovered for the smaller time-step sizes meeting $k = o(h^2)$ in this experiment; see [30, (15) and Remark 3 (iv)]. To narrow down the critical values $D_{\text{crit}}^{2-3}$ and $D_{\text{crit}}^{6-7}$, we simulate the relaxation dynamics for different values of the DMI constant $D$ corresponding to a resolution of $0.0025 \, \text{mJ/m}^2$ as seen in Figure 2. For each of the integrators and all considered DMI constants $D$, we relax the initial state using time-step sizes $k = 1/100 \, \text{ps}$, $1/200 \, \text{ps}$, $1/400 \, \text{ps}$, $1/800 \, \text{ps}$, $1/1600 \, \text{ps}$, where the two largest time-step sizes are omitted for the midpoint scheme because experimentally they do not fulfill the CFL constraint $k = o(h^2)$, i.e., neither of the nonlinear solvers converges for $k = 1/100 \, \text{ps}$, $1/200 \, \text{ps}$. We expect the simulations to be more and more accurate as the time-step size $k > 0$ decreases. The accuracy for the nonlinear solver is chosen as $\varepsilon = 10^{-8}$.

The results of this experiment displayed in Figure 2 show a sharp transition $D_{\text{crit}}^{2-3}$ between the uniform state and the skyrmion state. There is no sharp transition from the skyrmion state to the target skyrmion state, as the experiment reveals a small interval of DMI parameters $D$ for which relaxation leads to states we call *broken (symmetry) states* – neither a skyrmion nor a target skyrmion; see Figure 3 for a compilation of simulation details on this interval of broken states. While for the tangent plane integrators the determined transition value $D_{\text{crit}}^{2-3}$ and the transition interval of broken states clearly show a dependence on the used time discretization $k > 0$, the results for the midpoint scheme are robust and, in particular, are identical for all investigated time-step sizes. We draw the conclusion that the varying transition thresholds obtained for decreasing time-step size $k > 0$ by simulations with either of the tangent plane integrators are a consequence of the artificial energy dissipation quantified in (4.1).

We conclude that the tangent plane schemes are preferable for uncritical simulations as in Section 4.1 or [30, Section 4.2], where small deteriorations of the energy are acceptable, as they lead to already accurate results for much coarser time discretizations. However, when it comes to the simulation of dynamics, which are very sensitive to small inaccuracies and crucially depend on an accurate energy evolution, the midpoint scheme yields the most reliable results.
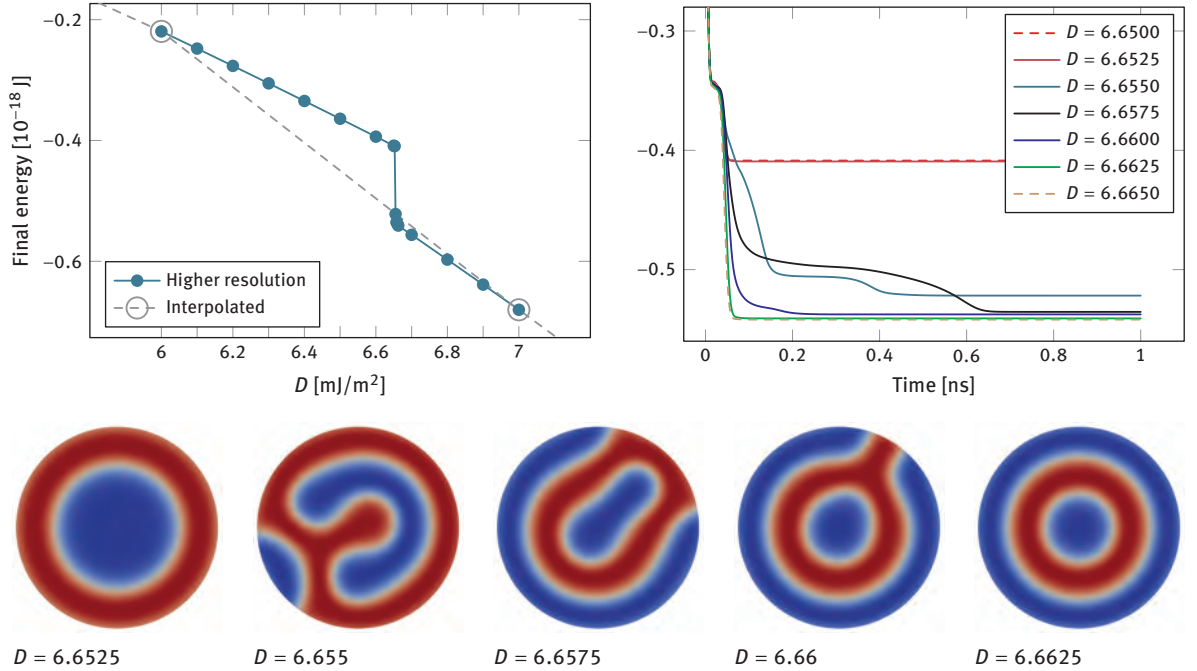
**Figure 3:** Experiment of Section 4.2. Details on the simulations for critical transition values between $D = 6\,\mathrm{mJ/m^2}$ and $D = 7\,\mathrm{mJ/m^2}$ with the second-order tangent plane scheme and time-step size $k = 1/100$ ps. Left: critical area of Figure 1 (left) recomputed with higher resolution reveals the predicted jump. Right: evolution of the total energy for different DMI constants $D$ emphasizing the transition dynamics. Bottom: relaxed states colored by $\boldsymbol{m}_3$ ranging from $-1$ (blue) to $+1$ (red). We observed that any relaxed state with broken symmetry obtained in the experiment of this section (marked with a cross $\times$ in Figure 2), qualitatively coincides with one of the three broken-symmetry states obtained by TPS2 and $k = 1/100$ ps displayed here.

## 4.3 Numerical Study on the CFL Conditions

Our results from Section 3.3.1 and Section 3.3.2, respectively, provide sufficient CFL conditions guaranteeing well-posedness and stability of the practical midpoint schemes in Theorem 3.7 (i) and Theorem 3.11. In this section, we investigate whether the CFL conditions arising from theory are also necessary in practice, or if they are technical artifacts possibly caused by unsharp estimates.

We consider the unit cube $\Omega \subset \mathbb{R}^3$ centered at the origin. Steered by the exchange-only effective field $\boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}) = \ell_{\mathrm{ex}}^2 \Delta \boldsymbol{m}$, the so-called initial *hedgehog* state $\boldsymbol{m}^0 \in \boldsymbol{H}^1(\Omega; \mathbb{S}^2)$ with $\boldsymbol{m}^0(\boldsymbol{x}) := \boldsymbol{x}/|\boldsymbol{x}| \in \mathbb{S}^2$ is relaxed towards equilibrium. The exchange length $\ell_{\mathrm{ex}} > 0$ and the Gilbert damping parameter $\alpha > 0$ are fixed at 1. The other discretization parameters – namely the mesh size $h > 0$, the time-step size $k > 0$, and the nonlinear solver accuracy $\varepsilon > 0$ – are subject to the numerical studies and are specified separately for each experiment. Linear systems are solved with GMRES and accuracy $10^{-14}$. For given $N \in \mathbb{N}$, the geometry is discretized by a structured mesh consisting of $(N + 1)^3$ vertices and $6N^3$ elements as described in [37, Section 5.2], leading to a uniform mesh $\mathcal{T}_h$ of congruent tetrahedra, each of diameter $h_{\mathrm{max}} = \sqrt{3}/N$ and with shortest edge length $h_{\mathrm{min}} = 1/N$. To break symmetry, the discontinuity of the hedgehog state at the origin is discretized via $\boldsymbol{m}_h^0(\boldsymbol{0}) := \boldsymbol{e}_3 \in \mathbb{S}^2$, while $\boldsymbol{m}_h^0(\boldsymbol{z}) := \boldsymbol{z}/|\boldsymbol{z}| \in \mathbb{S}^2$ for all other $\boldsymbol{0} \neq \boldsymbol{z} \in \mathcal{N}_h$.

### 4.3.1 Feasible Discretization Parameters for Nonlinear Solvers

In the next section, we carry out a numerical study on the CFL coupling of the time-step size to the mesh size arising from our analysis. Since the constants hidden in CFL conditions are usually not readily available, we need to propose an appropriate criterion for the classification of given discretization parameters as feasible or non-feasible. Hence, the goal is to derive such a criterion from the numerical experiment in this section.
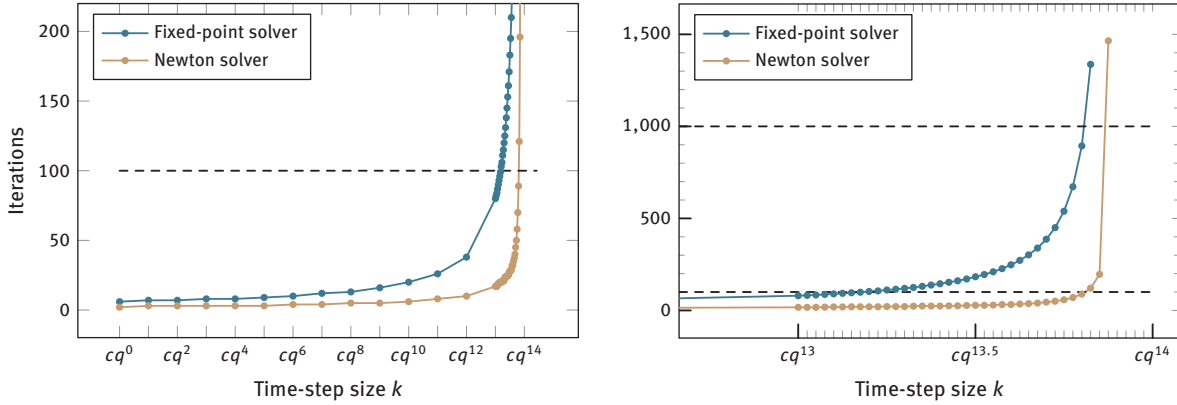
**Figure 4:** Experiment of Section 4.3.1. Left: the number of nonlinear iterations rapidly grows as the time-step size $k = cq^j > 0$ with $c = 0.00016$ and $q = 5/4$ approaches the threshold value $k \to k_{\mathrm{thresh}}(h)$. Right: zoom into the critical area between $cq^{13}$ and $cq^{14}$ where the blow-up occurs.

For fixed mesh size $h_{\min} = 1/8$, nonlinear solver tolerance $\varepsilon = 10^{-8}$, and starting from a rather fine time discretization $k = 0.00016$, we iteratively increase the time-step size by 25 % multiple times and track the number of nonlinear iterations required to meet the stopping criterion (3.13) or (3.19), respectively, in the first time step of Algorithm 3.6 or Algorithm 3.8.

The results depicted in Figure 4 show that, for both practical midpoint schemes, the number of nonlinear iterations stays well-bounded until a certain threshold value $k_{\mathrm{thresh}}(h) > 0$ is approached. Close to the threshold value, however, an increase of the time-step size by 25 % impacts the number of nonlinear iterations by numbers of magnitude, if the solver converges at all. Hence, it is reasonable to classify time-step sizes $k > 0$ with $k < k_{\mathrm{thresh}}(h)$ as feasible, and those with $k > k_{\mathrm{thresh}}(h)$ as non-feasible. Surprisingly, despite the different theoretical CFL conditions $k = o(h^2)$ and $k = o(h^{7/3})$ imposed in Proposition 3.5 and Theorem 3.11, respectively, the threshold value $k_{\mathrm{thresh}}(h)$ seems to coincide for Algorithm 3.6 and Algorithm 3.8. This observation is investigated further in Section 4.3.2. Finally, we note that, in view of the quadratic convergence of Newton's method, it is not surprising that the Newton solver clearly outperforms the fixed-point iteration in terms of nonlinear iteration numbers.

Motivated by the results of this experiment, in Section 4.3.2, we will use the following criterion to classify feasibility of discretization parameters: if, for any given $(h, k, \varepsilon)$, the respective stopping criterion (3.13) or (3.19) is not met after at most 100 iterations of the nonlinear solver in Algorithm 3.6 or Algorithm 3.8, we consider the practical midpoint scheme as non-feasible for this combination of discretization parameters $h, k, \varepsilon > 0$. Given $h > 0$, this classification of feasibility is an estimate for the threshold value $k_{\mathrm{thresh}}(h) > 0$ such that the nonlinear solver converges for $0 < k < k_{\mathrm{thresh}}(h)$ and diverges for $k > k_{\mathrm{thresh}}(h)$. Although only an approximation, Figure 4 shows that, in view of practical applicability, this estimation of $k_{\mathrm{thresh}}(h)$ seems quite appropriate as nonlinear iteration numbers increase drastically as $k$ approaches $k_{\mathrm{thresh}}(h)$.

### 4.3.2 Coupling of Time-Step Size to Mesh Size

We consider the CFL conditions $k = o(h^2)$ and $k = o(h^{7/3})$ from Theorem 3.7 (i) and Theorem 3.11, respectively, sufficient to guarantee convergence of the fixed-point iteration and the Newton solver. For different mesh sizes $h_{\min} \in \{2^{-j} : j = 1, \dots, 5\}$, time-step sizes $k \in \{0.00016 \cdot (\frac{5}{4})^j : j = 0, \dots, 27\}$, and nonlinear solver accuracy $\varepsilon > 0$ fixed at $10^{-8}$, we investigate convergence of the nonlinear solver for one time step of relaxing the initial hedgehog state. As argued in Section 4.3.1, the threshold value of 100 nonlinear iterations is used to classify feasibility of the discretization parameters.

The results of this experiment shown in Figure 5 give insight to the applicability of the practical midpoint schemes. First, for Algorithm 3.6, the theoretically sufficient CFL condition $k = o(h^2)$ is shown to be
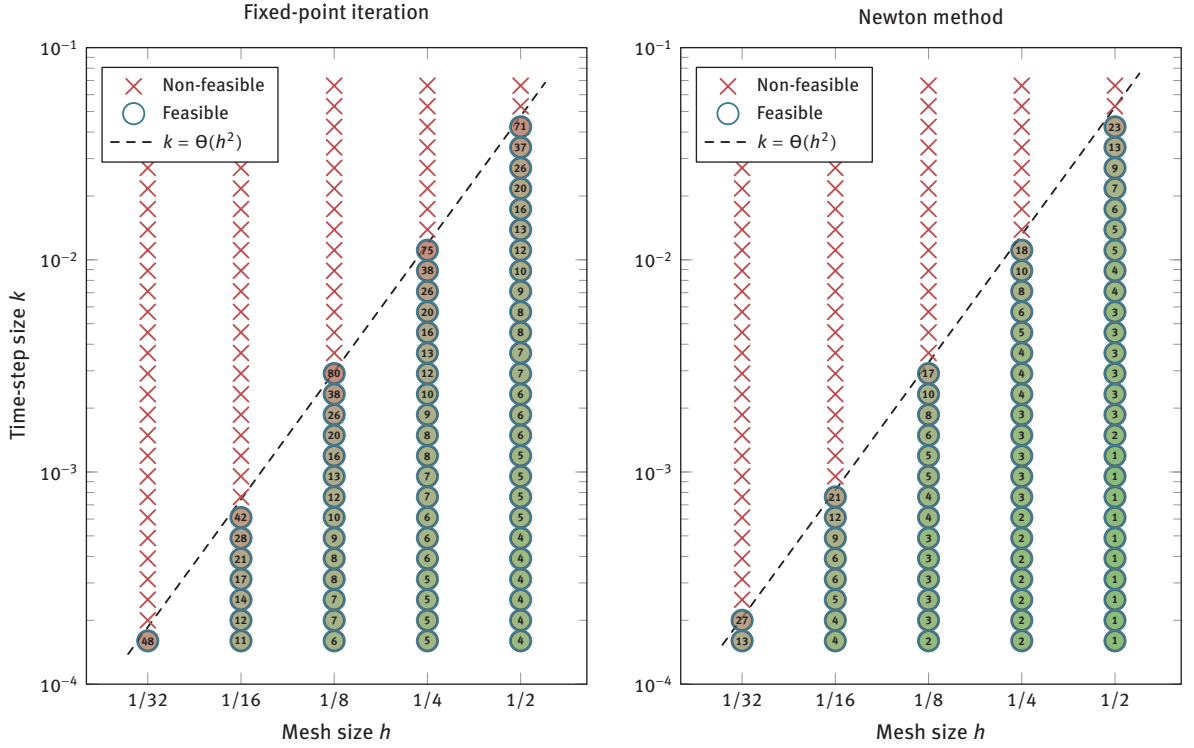
**Figure 5:** Experiment of Section 4.3.2. Convergence of the nonlinear solvers in the practical midpoint schemes is investigated. For feasible parameters, the number of nonlinear iterations is given inside the circle. Left: practical midpoint scheme based on the fixed-point iteration (Algorithm 3.6). Right: practical midpoint scheme based on the Newton iteration (Algorithm 3.8). The data points show feasibility if $k = \Theta(h^\beta)$ with possible slopes $1.93 \leq \beta_{\text{fixed-point}} \leq 2.09$ and $1.85 \leq \beta_{\text{newton}} \leq 2.01$.

sharp in practice. Further, since the experiment reveals the same CFL condition $k = o(h^2)$ to be sufficient for convergence of the Newton solver, we expect that the well-posedness analysis of the Newton iteration can be improved weakening the CFL condition in Theorem 3.11 from $k = o(h^{7/3})$ to $k = o(h^2)$. We note that, also in the simulation of skyrmion dynamics in Section 4.2, both practical midpoint schemes were equivalently restrictive on the time discretization. Lastly, this experiment shows that, in terms of iteration numbers, the Newton solver outperforms the fixed-point solver as expected from theory (quadratic vs. linear convergence).

### 4.3.3 Constraint Violation Induced by Nonlinear Solver Accuracy

In contrast to the fixed-point iteration from Section 3.3.1, the Newton iteration from Section 3.3.2 does not inherently preserve discrete unit length, i.e., $\boldsymbol{m}_{h\varepsilon}^i \notin \mathcal{M}_h$ for the Newton linearization. To quantify the impact of the Newton solver on the discrete magnetization length, the initial hedgehog state is relaxed to equilibrium ($T = 5$) using different nonlinear solver accuracies $\varepsilon \in \{10^{-j/2} : j = 0, \ldots, 24\}$. We simulate the dynamics for $h_{\min} = 1/4$ and $h_{\min} = 1/8$ with time-step sizes chosen roughly half the value of $k_{\text{thresh}}(h)$ from Section 4.3.1. In Figure 6, we plot the deviations

$$\max_{\boldsymbol{z} \in \mathcal{N}_h} |\boldsymbol{m}_{h\varepsilon k}(T, \boldsymbol{z})| - 1 \quad \text{and} \quad 1 - \min_{\boldsymbol{z} \in \mathcal{N}_h} |\boldsymbol{m}_{h\varepsilon k}(T, \boldsymbol{z})| \tag{4.2}$$

over the nonlinear solver accuracy $\varepsilon > 0$. In this experiment for the practical midpoint scheme based on the Newton iteration, deviation from unit length decreases with rate between $\Theta(\varepsilon^{9/10})$ and $\Theta(\varepsilon^{8/10})$ as $\varepsilon \to 0$. In contrast to that, for the practical midpoint scheme based on the fixed-point iteration, the deviation from unit length is unaffected by the choice of $\varepsilon > 0$ as expected from theory.
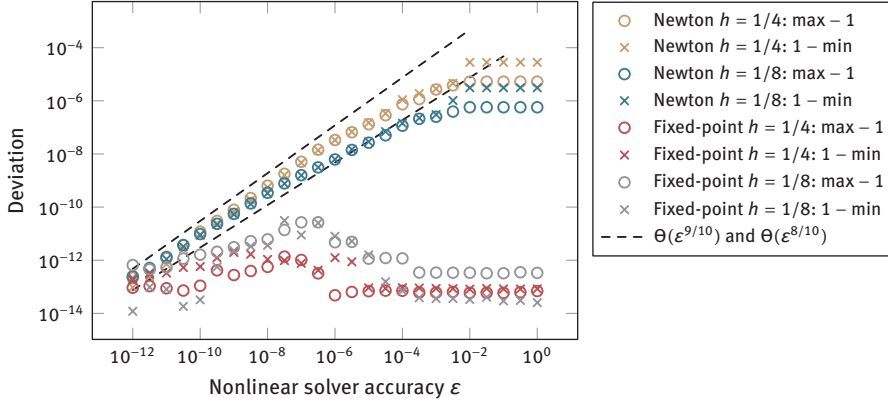
**Figure 6:** Experiment of Section 4.3.3. Dependence of the constraint violation (4.2) on the nonlinear solver accuracy $\varepsilon > 0$ is investigated. For the Newton solver, the deviation from unit length decreases as $\varepsilon \to 0$. No obvious correlation is observed for the fixed-point iteration, which is expected since it is designed to be constraint preserving; see Proposition 3.5 (iii).

As in this experiment both practical midpoint schemes were stable (i.e., non energy-increasing) even for nonlinear solver accuracies as large as $\varepsilon = 1$, an experimental setup for the investigation of the coupling to the mesh size $\varepsilon = \mathcal{O}(h)$ and $\varepsilon = \mathcal{O}(h^{3/2})$ from Theorem 3.7 (i) and Theorem 3.11, respectively, is yet to be proposed in a future numerical study.

# 5 Proof of Theorem 3.2 for the Ideal Midpoint Scheme

## 5.1 Existence of Solutions, Unit-Length Constraint, and Stability

*Proof of Theorem 3.2* (i). Let $i \in \mathbb{N}_0$ be arbitrary. Define $\boldsymbol{F} \colon \boldsymbol{V}_h \to \boldsymbol{V}_h$ by

$$\boldsymbol{F}(\boldsymbol{\phi}_h) := \boldsymbol{\phi}_h - \boldsymbol{m}_h^i + \frac{k}{2}\mathcal{I}_h[\boldsymbol{\phi}_h \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{\phi}_h) + \alpha \boldsymbol{\phi}_h \times \boldsymbol{m}_h^i] \quad \text{for all } \boldsymbol{\phi}_h \in \boldsymbol{V}_h.$$

If $\boldsymbol{\eta}_h \in \boldsymbol{V}_h$ satisfies $\boldsymbol{F}(\boldsymbol{\eta}_h) = \boldsymbol{0}$, then $\boldsymbol{m}_h^{i+1} := 2\boldsymbol{\eta}_h - \boldsymbol{m}_h^i$ satisfies (3.5). Since

$$\langle \boldsymbol{F}(\boldsymbol{\phi}_h), \boldsymbol{\phi}_h \rangle_h = \langle \boldsymbol{\phi}_h - \boldsymbol{m}_h^i, \boldsymbol{\phi}_h \rangle_h \geq 0 \quad \text{for all } \boldsymbol{\phi}_h \in \boldsymbol{V}_h \text{ with } \|\boldsymbol{\phi}_h\|_h = \|\boldsymbol{m}_h^i\|_h > 0,$$

an application of the Brouwer fixed-point theorem (see, e.g., [29, Chapter IV, Corollary 1.1]) ensures the existence of $\boldsymbol{\eta}_h \in \boldsymbol{V}_h$ such that $\|\boldsymbol{\eta}_h\|_h \leq \|\boldsymbol{m}_h^i\|_h$ and $\boldsymbol{F}(\boldsymbol{\eta}_h) = \boldsymbol{0}$. This proves that (3.5) admits a solution $\boldsymbol{m}_h^{i+1} \in \boldsymbol{V}_h$.

Let $z \in \mathcal{N}_h$ be arbitrary. We test (3.5) with $\boldsymbol{\phi}_h = \boldsymbol{m}_h^{i+1/2}(z)\varphi_z \in \boldsymbol{V}_h$ to obtain that

$$\langle d_t \boldsymbol{m}_h^{i+1}, \varphi_z \boldsymbol{m}_h^{i+1/2}(z) \rangle_h = \frac{\beta_z}{2k}\left(|\boldsymbol{m}_h^{i+1}(z)|^2 - |\boldsymbol{m}_h^i(z)|^2\right) = 0.$$

We conclude that $|\boldsymbol{m}_h^{i+1}(z)| = |\boldsymbol{m}_h^i(z)|$. Since $\boldsymbol{m}_h^0 \in \mathcal{M}_h$ by assumption, we conclude that $\boldsymbol{m}_h^{i+1} \in \mathcal{M}_h$. □

*Proof of Theorem 3.2* (ii). Let $J \in \mathbb{N}$. To show (3.6), we choose the test function

$$\boldsymbol{\phi}_h = \alpha d_t \boldsymbol{m}_h^{i+1} - \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_h^{i+1/2}) \in \boldsymbol{V}_h$$

in (3.5). We obtain the equality

$$\langle \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_h^{i+1/2}), d_t \boldsymbol{m}_h^{i+1} \rangle_h = \alpha \|d_t \boldsymbol{m}_h^{i+1}\|_h^2.$$

For the left-hand side, it holds that

$$\langle \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_h^{i+1/2}), d_t \boldsymbol{m}_h^{i+1} \rangle_h \overset{(3.2)}{=} \langle \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_h^{i+1/2}), d_t \boldsymbol{m}_h^{i+1} \rangle$$

$$\overset{(2.3)}{=} -a(\boldsymbol{m}_h^{i+1/2}, d_t \boldsymbol{m}_h^{i+1}) + \langle \boldsymbol{f}, d_t \boldsymbol{m}_h^{i+1} \rangle_\Omega$$

$$\overset{(2.1)}{=} -\frac{1}{k}\left(\mathcal{E}(\boldsymbol{m}_h^{i+1}) - \mathcal{E}(\boldsymbol{m}_h^i)\right). \tag{5.1}$$

We conclude that

$$\mathcal{E}(\boldsymbol{m}_h^{i+1}) - \mathcal{E}(\boldsymbol{m}_h^i) = -\alpha k \|d_t \boldsymbol{m}_h^{i+1}\|_h^2.$$

Summation over $i = 0, \ldots, J - 1$ yields (3.6). □

## 5.2 Weak Convergence Result

To start with, we note that the bilinear forms $a(\cdot, \cdot)$ and $a^{\mathrm{loc}}(\cdot, \cdot)$ are continuous, i.e., there exists $C_1 > 0$ such that

$$a(\boldsymbol{\psi}, \boldsymbol{\varphi}) \le (C_1 + \|\boldsymbol{\pi}\|_{L(\boldsymbol{L}^2(\Omega); \boldsymbol{L}^2(\Omega))}) \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega)} \|\boldsymbol{\varphi}\|_{\boldsymbol{H}^1(\Omega)} \quad \text{for all } \boldsymbol{\psi}, \boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega), \tag{5.2a}$$

$$a^{\mathrm{loc}}(\boldsymbol{\psi}, \boldsymbol{\varphi}) \le C_1 \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega)} \|\boldsymbol{\varphi}\|_{\boldsymbol{H}^1(\Omega)} \quad \text{for all } \boldsymbol{\psi}, \boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega), \tag{5.2b}$$

and satisfy the Gårding inequality, i.e., there exist $C_2 > 0$ and $C_3 \in \mathbb{R}$ such that

$$a(\boldsymbol{\psi}, \boldsymbol{\psi}) \ge a^{\mathrm{loc}}(\boldsymbol{\psi}, \boldsymbol{\psi}) \ge C_2 \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega)}^2 - C_3 \|\boldsymbol{\psi}\|_{\boldsymbol{L}^2(\Omega)}^2 \quad \text{for all } \boldsymbol{\psi} \in \boldsymbol{H}^1(\Omega). \tag{5.2c}$$

The constants $C_1, C_2, C_3$ in (5.2) depend on $\|\boldsymbol{A}_n\|_{\boldsymbol{L}^\infty(\Omega)}$ and $\|\boldsymbol{J}_n\|_{\boldsymbol{L}^\infty(\Omega)}$ ($n = 1, 2, 3$), and $A_0$. Finally, we consider, besides (3.4), the piecewise constant time reconstruction $\overline{\boldsymbol{m}}_{hk}$ defined by $\overline{\boldsymbol{m}}_{hk}(t) := \boldsymbol{m}_h^{i+1/2}$ for all $i \in \mathbb{N}_0$ and $t \in [t_i, t_{i+1})$.

With these ingredients, we prove the convergence result for Algorithm 3.1.

*Proof of Theorem 3.2* (iii). The proof follows the lines of [15, 39]; therefore, we only sketch it. Let $J \in \mathbb{N}$. Since $\boldsymbol{m}_h^J \in \mathcal{M}_h$, $\|\boldsymbol{m}_h^J\|_{\boldsymbol{L}^2(\Omega)} \le |\Omega|^{1/2}$. Hence, combining the inequalities (5.2) and the norm equivalence (3.1) with (3.6), we obtain the estimate

$$\|\boldsymbol{m}_h^J\|_{\boldsymbol{H}^1(\Omega)}^2 + k \sum_{i=0}^{J-1} \|d_t \boldsymbol{m}_h^i\|_{\boldsymbol{L}^2(\Omega)}^2 \le C, \tag{5.3}$$

where $C > 0$ depends only on the problem data. We infer the uniform boundedness of the sequences of time reconstructions $\{\boldsymbol{m}_{hk}\}$ and $\{\overline{\boldsymbol{m}}_{hk}\}$ in $L^\infty(\mathbb{R}_{>0}; \boldsymbol{H}^1(\Omega))$. Let $T > 0$ be arbitrary. From (5.3), it also follows the uniform boundedness of $\{\boldsymbol{m}_{hk}|_{\Omega_T}\}$ (resp., $\{\overline{\boldsymbol{m}}_{hk}|_{\Omega_T}\}$) in $\boldsymbol{H}^1(\Omega_T)$ and in $L^\infty(0, T; \boldsymbol{H}^1(\Omega))$ (resp., only in $L^\infty(0, T; \boldsymbol{H}^1(\Omega))$). With successive extractions of convergent subsequences (not relabeled), one can show that there exists a common limit $\boldsymbol{m} \in L^\infty(\mathbb{R}_{>0}; \boldsymbol{H}^1(\Omega))$ with $\boldsymbol{m}|_{\Omega_T} \in \boldsymbol{H}^1(\Omega_T)$ for which we have the convergences $\boldsymbol{m}_{hk}, \overline{\boldsymbol{m}}_{hk} \stackrel{*}{\rightharpoonup} \boldsymbol{m}$ in $L^\infty(\mathbb{R}_{>0}; \boldsymbol{H}^1(\Omega))$, $\boldsymbol{m}_{hk}|_{\Omega_T}, \overline{\boldsymbol{m}}_{hk}|_{\Omega_T} \stackrel{*}{\rightharpoonup} \boldsymbol{m}|_{\Omega_T}$ in $L^\infty(0, T; \boldsymbol{H}^1(\Omega))$, and $\boldsymbol{m}_{hk}|_{\Omega_T} \rightharpoonup \boldsymbol{m}|_{\Omega_T}$ in $\boldsymbol{H}^1(\Omega_T)$, With the argument of [39, Sections 3.2–3.3], one also gets that the limit function $\boldsymbol{m}$ is $\mathbb{S}^2$-valued and satisfies the initial condition $\boldsymbol{m}(0) = \boldsymbol{m}^0$ in the sense of traces.

To verify the variational formulation (2.6), let $\boldsymbol{\varphi} \in \boldsymbol{C}^\infty(\overline{\Omega_T})$. Let $J \in \mathbb{N}$ be the smallest integer such that $T \le kJ$. We define the semi-discrete function $\boldsymbol{\varphi}_h \in \boldsymbol{C}^\infty([0, kJ]; \boldsymbol{V}_h)$ by $\boldsymbol{\varphi}_h(t) = \mathcal{I}_h[\boldsymbol{\varphi}(t)]$ for all $t \in [0, kJ]$. For $i = 0, \ldots, J - 1$ and $t \in (t_i, t_{i+1})$, we test (3.5) with $\boldsymbol{\phi}_h = \boldsymbol{\varphi}_h(t) \in \boldsymbol{V}_h$. Then, integrating in time over $(t_i, t_{i+1})$ and summing over $i = 0, \ldots, J - 1$, we obtain that

$$\int_0^{kJ} \langle \partial_t \boldsymbol{m}_{hk}(t), \boldsymbol{\varphi}_h(t) \rangle_h \, dt = - \int_0^{kJ} \langle \overline{\boldsymbol{m}}_{hk}(t) \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\overline{\boldsymbol{m}}_{hk}(t)), \boldsymbol{\varphi}_h(t) \rangle_h \, dt$$
$$+ \alpha \int_0^{kJ} \langle \overline{\boldsymbol{m}}_{hk}(t) \times \partial_t \boldsymbol{m}_{hk}(t), \boldsymbol{\varphi}_h(t) \rangle_h \, dt. \tag{5.4}$$

The argument in [15, Section 3] shows that

$$\int_0^{kJ} \langle \partial_t \boldsymbol{m}_{hk}(t), \boldsymbol{\varphi}_h(t) \rangle_h \, dt \to \int_0^T \langle \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle_\Omega \, dt \qquad \text{as } h, k \to 0,$$

$$\int_0^{kJ} \langle \overline{\boldsymbol{m}}_{hk}(t) \times \partial_t \boldsymbol{m}_{hk}(t), \boldsymbol{\varphi}_h(t) \rangle_h \, dt \to \int_0^T \langle \boldsymbol{m}(t) \times \partial_t \boldsymbol{m}(t), \boldsymbol{\varphi}(t) \rangle_\Omega \, dt \quad \text{as } h, k \to 0.$$

For the first term on the right-hand side of (5.4), simple algebraic manipulations together with (3.2) show that

$$
\int_0^{kJ} \langle \overline{\boldsymbol{m}}_{hk}(t) \times \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\overline{\boldsymbol{m}}_{hk}(t)), \boldsymbol{\varphi}_h(t) \rangle_h \, \mathrm{d}t = \int_0^{kJ} \langle \boldsymbol{h}_{\mathrm{eff}}(\overline{\boldsymbol{m}}_{hk}(t)), (\mathfrak{I}_h - 1)[\boldsymbol{\varphi}_h(t) \times \overline{\boldsymbol{m}}_{hk}(t)] \rangle \, \mathrm{d}t
$$
$$
+ \int_0^{kJ} \langle \boldsymbol{h}_{\mathrm{eff}}(\overline{\boldsymbol{m}}_{hk}(t)), \boldsymbol{\varphi}_h(t) \times \overline{\boldsymbol{m}}_{hk}(t) \rangle \, \mathrm{d}t. \tag{5.5}
$$

Since

$$
\left| \int_0^{kJ} \langle \boldsymbol{h}_{\mathrm{eff}}(\overline{\boldsymbol{m}}_{hk}(t)), (\mathfrak{I}_h - 1)[\boldsymbol{\varphi}_h(t) \times \overline{\boldsymbol{m}}_{hk}(t)] \rangle \, \mathrm{d}t \right|
$$
$$
\leq C \int_0^{kJ} (\|\overline{\boldsymbol{m}}_{hk}(t)\|_{\boldsymbol{H}^1(\Omega)} + \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}) \|(\mathfrak{I}_h - 1)[\boldsymbol{\varphi}_h(t) \times \overline{\boldsymbol{m}}_{hk}(t)]\|_{\boldsymbol{H}^1(\Omega)} \, \mathrm{d}t
$$

and $\|(\mathfrak{I}_h - 1)[\boldsymbol{\varphi}_h \times \overline{\boldsymbol{m}}_{hk}]\|_{\boldsymbol{L}^\infty(0,T;\boldsymbol{H}^1(\Omega))} \leq Ch$ (see [39, equations (39)–(40)]), the first term on the right-hand side of (5.5) tends to 0 as $h, k \to 0$. Moreover, owing to the available convergence results and the convergence properties of $a(\cdot, \cdot)$ and $\boldsymbol{\pi}(\cdot)$, it holds that

$$
\int_0^{kJ} \langle \boldsymbol{h}_{\mathrm{eff}}(\overline{\boldsymbol{m}}_{hk}(t)), \boldsymbol{\varphi}_h(t) \times \overline{\boldsymbol{m}}_{hk}(t) \rangle \, \mathrm{d}t \to \int_0^T \langle \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}(t)), \boldsymbol{\varphi}(t) \times \boldsymbol{m}(t) \rangle \, \mathrm{d}t \quad \text{as } h, k \to 0.
$$

Hence, passing (5.4) to the limit as $h, k \to 0$, we obtain (2.6) for any smooth test function $\boldsymbol{\varphi}$. By density, we obtain the desired result.

Finally, the energy inequality (2.7) is obtained by passing to the limit as $h, k \to 0$ the discrete energy identity (3.6) and using standard lower semicontinuity arguments in combination with the available convergence results. $\qquad \square$

# 6 Analysis of the Practical Midpoint Scheme: Constraint-Preserving Fixed-Point Iteration

To start with, we recall that, for quasi-uniform families of triangulations, we have the inverse estimate

$$
\|\nabla \boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)} \leq C_{\mathrm{inv}} h^{-1} \|\boldsymbol{\phi}_h\|_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \boldsymbol{\phi}_h \in \boldsymbol{V}_h, \tag{6.1a}
$$

from which it follows that

$$
\|\mathbb{P}_h \boldsymbol{\phi}\|_h \leq (1 + C_{\mathrm{inv}}^2 h^{-2})^{1/2} \|\boldsymbol{\phi}\|_{\boldsymbol{H}^1(\Omega)^\star} \quad \text{for all } \boldsymbol{\phi} \in \boldsymbol{H}^1(\Omega)^\star. \tag{6.1b}
$$

Here, $C_{\mathrm{inv}} > 0$ depends only on $\kappa$. Moreover, the following inequalities are direct consequences of (5.2a)–(5.2b):

$$
\|\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\psi}) - \boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\varphi})\|_{\boldsymbol{H}^1(\Omega)^\star} \leq C_1 \|\boldsymbol{\psi} - \boldsymbol{\varphi}\|_{\boldsymbol{H}^1(\Omega)} \quad \text{for all } \boldsymbol{\psi}, \boldsymbol{\varphi} \in \boldsymbol{H}^1(\Omega), \tag{6.1c}
$$
$$
\|\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\psi})\|_{\boldsymbol{H}^1(\Omega)^\star} \leq C_1 \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega)} + \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)} \quad \text{for all } \boldsymbol{\psi} \in \boldsymbol{H}^1(\Omega), \tag{6.1d}
$$
$$
\|\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{\psi}) - \boldsymbol{f}\|_{\boldsymbol{H}^1(\Omega)^\star} \leq C_1 \|\boldsymbol{\psi}\|_{\boldsymbol{H}^1(\Omega)} \quad \text{for all } \boldsymbol{\psi} \in \boldsymbol{H}^1(\Omega). \tag{6.1e}
$$

## 6.1 Well-Posedness

We now prove Proposition 3.5, which establishes the properties of the constraint-preserving fixed-point iteration proposed in Section 3.3.1.

*Proof of Proposition 3.5* (i). Since the bilinear form on the left-hand side of (3.12) is elliptic with respect to the norm $\|\cdot\|_h$, the variational problem admits a unique solution $\boldsymbol{\eta}_h^{i,\ell+1} \in \boldsymbol{V}_h$ for each $\ell \geq 0$.

Let $\ell \in \mathbb{N}_0$, and let $\boldsymbol{z} \in \mathcal{N}_h$ be an arbitrary node. Testing (3.12) with $\boldsymbol{\phi}_h = \boldsymbol{\eta}_h^{i,\ell+1}(\boldsymbol{z})\varphi_{\boldsymbol{z}} \in \boldsymbol{V}_h$, we obtain that

$$\beta_{\boldsymbol{z}}|\boldsymbol{\eta}_h^{i,\ell+1}(\boldsymbol{z})|^2 = \beta_{\boldsymbol{z}}\boldsymbol{\eta}_h^{i,\ell+1}(\boldsymbol{z}) \cdot \boldsymbol{m}_{h\varepsilon}^i(\boldsymbol{z}).$$

Hence, $|\boldsymbol{\eta}_h^{i,\ell+1}(\boldsymbol{z})| \leq |\boldsymbol{m}_{h\varepsilon}^i(\boldsymbol{z})| = 1$. We conclude that $\|\boldsymbol{\eta}_h^{i,\ell+1}\|_{\boldsymbol{L}^\infty(\Omega)} \leq 1$. □

*Proof of Proposition 3.5* (ii). Let $\ell \in \mathbb{N}_0$. Subtracting the equations satisfied by two consecutive iterates $\boldsymbol{\eta}_h^{i,\ell+1}, \boldsymbol{\eta}_h^{i,\ell+2} \in \boldsymbol{V}_h$ in (3.12), we obtain that

$$\langle \boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}, \boldsymbol{\phi}_h \rangle_h \overset{(3.12)}{=} -\frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+2} \times \mathbb{P}_h\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}), \boldsymbol{\phi}_h \rangle_h + \frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell}), \boldsymbol{\phi}_h \rangle_h$$
$$- \frac{k}{2}\langle (\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}) \times \mathbb{P}_h\Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}), \boldsymbol{\phi}_h \rangle_h$$
$$- \alpha\langle (\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}) \times \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h \rangle_h.$$

Choosing $\boldsymbol{\phi}_h = \boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1} \in \boldsymbol{V}_h$, we obtain that

$$\|\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}\|_h^2 = -\frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+2} \times \mathbb{P}_h\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}), \boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1} \rangle_h$$
$$+ \frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell}), \boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1} \rangle_h$$
$$= -\frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell})), \boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1} \rangle_h,$$

where the second equality can be seen by adding and subtracting the quantity

$$\frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}), \boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1} \rangle_h.$$

It follows that

$$\|\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}\|_h^2 = -\frac{k}{2}\langle \boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell})), \boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1} \rangle_h$$
$$\leq \frac{k}{2}\|\boldsymbol{\eta}_h^{i,\ell+1}\|_{\boldsymbol{L}^\infty(\Omega)}\|\mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell}))\|_h\|\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}\|_h$$
$$\leq \frac{k(1 + C_{\text{inv}}^2 h^{-2})^{1/2}}{2}\|\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell})\|_{\boldsymbol{H}^1(\Omega)^\star}\|\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}\|_h,$$

where the last inequality follows from $\|\boldsymbol{\eta}_h^{i,\ell+1}\|_{\boldsymbol{L}^\infty(\Omega)} \leq 1$ and (6.1b). Moreover, it holds that

$$\|\boldsymbol{\eta}_h^{i,\ell+2} - \boldsymbol{\eta}_h^{i,\ell+1}\|_h \leq \frac{k(1 + C_{\text{inv}}^2 h^{-2})^{1/2}}{2}\|\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell})\|_{\boldsymbol{H}^1(\Omega)^\star}$$
$$\overset{(6.1c)}{\leq} \frac{C_1 k(1 + C_{\text{inv}}^2 h^{-2})^{1/2}}{2}\|\boldsymbol{\eta}_h^{i,\ell+1} - \boldsymbol{\eta}_h^{i,\ell}\|_{\boldsymbol{H}^1(\Omega)}$$
$$\overset{(6.1a)}{\leq} \frac{C_1 k(1 + C_{\text{inv}}^2 h^{-2})}{2}\|\boldsymbol{\eta}_h^{i,\ell+1} - \boldsymbol{\eta}_h^{i,\ell}\|_h.$$

Since $k = o(h^2)$ as $h, k \to 0$, there exist $h_0, k_0 > 0$ and a constant $0 < q < 1$ for which (3.14) holds for all $h < h_0$ and $k < k_0$. □

*Proof of Proposition 3.5* (iii). Let $\ell \in \mathbb{N}_0$. Using (6.1b), (6.1c), (3.1) as well as the fact that $\|\boldsymbol{\eta}_h^{i,\ell+1}\|_{\boldsymbol{L}^\infty(\Omega)} \leq 1$, we obtain that

$$\|\mathcal{I}_h[\boldsymbol{\eta}_h^{i,\ell+1} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell}))]\|_h \leq \|\mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell}))\|_h$$
$$\leq C_1(1 + C_{\text{inv}}^2 h^{-2})\|\boldsymbol{\eta}_h^{i,\ell+1} - \boldsymbol{\eta}_h^{i,\ell}\|_h$$
$$\overset{(3.14)}{\leq} C_1(1 + C_{\text{inv}}^2 h^{-2})q^\ell\|\boldsymbol{\eta}_h^{i,1} - \boldsymbol{\eta}_h^{i,0}\|_h$$
$$\leq 2C_1|\Omega|^{1/2}(1 + C_{\text{inv}}^2 h^{-2})q^\ell.$$

Hence, $\|\mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell+1}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell}))\|_h \le \varepsilon$ for all $\ell \in \mathbb{N}_0$ satisfying

$$\ell \ge \frac{\log(2C_1|\Omega|^{1/2}(1 + C_{\text{inv}}^2 h^{-2})/\varepsilon)}{\log(1/q)}.$$

Since $\boldsymbol{m}_{h\varepsilon}^{i+1} := 2\boldsymbol{\eta}_h^{i,\ell^*+1} - \boldsymbol{m}_{h\varepsilon}^i$, there holds $\boldsymbol{m}_{h\varepsilon}^{i+1/2} = \boldsymbol{\eta}_h^{i,\ell^*+1}$. From (3.12), it follows that $\boldsymbol{m}_{h\varepsilon}^{i+1}$ solves

$$\langle d_t\boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\phi}_h\rangle_h = -\langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\boldsymbol{\eta}_h^{i,\ell^*}) + \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})), \boldsymbol{\phi}_h\rangle_h + \alpha\langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times d_t\boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\phi}_h\rangle_h$$

for all $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$. Testing with $\boldsymbol{\phi}_h = \boldsymbol{m}_{h\varepsilon}^{i+1/2}(\boldsymbol{z})\varphi_{\boldsymbol{z}} \in \boldsymbol{V}_h$ then reveals that $\boldsymbol{m}_{h\varepsilon}^{i+1} \in \mathcal{M}_h$ (see the proof of Theorem 3.2 (i)). □

## 6.2 Stability and Weak Convergence

Next, we provide the proof of Theorem 3.7, which establishes the stability and convergence of Algorithm 3.6.

*Proof of Theorem 3.7.* Part (i) is a direct consequence of Proposition 3.5. The proof of part (ii) follows the lines of the one of Theorem 3.2 (ii).

Let us now consider the proof of part (iii). Testing (3.15) with

$$\boldsymbol{\phi}_h = \alpha d_t\boldsymbol{m}_{h\varepsilon}^{i+1} - \mathbb{P}_h\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) + \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]$$

yields

$$\alpha\|d_t\boldsymbol{m}_{h\varepsilon}^{i+1}\|_h^2 - \langle d_t\boldsymbol{m}_{h\varepsilon}^{i+1}, \mathbb{P}_h\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2})\rangle_h + \langle d_t\boldsymbol{m}_{h\varepsilon}^{i+1}, \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h$$
$$= \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i, \alpha d_t\boldsymbol{m}_{h\varepsilon}^{i+1} - \mathbb{P}_h\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) + \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h.$$

Using (5.1) and rearranging the terms, we obtain that

$$\mathcal{E}(\boldsymbol{m}_{h\varepsilon}^{i+1}) + \alpha k\|d_t\boldsymbol{m}_{h\varepsilon}^{i+1}\|_h^2$$
$$= \mathcal{E}(\boldsymbol{m}_{h\varepsilon}^i) - k\langle d_t\boldsymbol{m}_{h\varepsilon}^{i+1}, \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h$$
$$+ k\langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i, \alpha d_t\boldsymbol{m}_{h\varepsilon}^{i+1} - \mathbb{P}_h\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) + \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h.$$

Let $1 \le j \le J$. Summation over $i = 0, \ldots, j - 1$ leads to

$$\mathcal{E}(\boldsymbol{m}_{h\varepsilon}^j) + \alpha k \sum_{i=0}^{j-1} \|d_t\boldsymbol{m}_{h\varepsilon}^{i+1}\|_h^2$$
$$= \mathcal{E}(\boldsymbol{m}_h^0) - k \sum_{i=0}^{j-1} \langle d_t\boldsymbol{m}_{h\varepsilon}^{i+1}, \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h$$
$$+ k \sum_{i=0}^{j-1} \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i, \alpha d_t\boldsymbol{m}_{h\varepsilon}^{i+1} - \mathbb{P}_h\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) + \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h.$$

Applying the Gårding inequality (5.2c) and continuity (5.2a) and using the fact that

$$\|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{L}^\infty(\Omega)} = \|\boldsymbol{m}_h^0\|_{\boldsymbol{L}^\infty(\Omega)} = 1,$$

we obtain that

$$C_2\|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{H}^1(\Omega)}^2 + 2\alpha k \sum_{i=0}^{j-1} \|d_t\boldsymbol{m}_{h\varepsilon}^{i+1}\|_h^2$$
$$\le (C_1 + \|\boldsymbol{\pi}\|_{L(\boldsymbol{L}^2(\Omega), \boldsymbol{L}^2(\Omega))})\|\boldsymbol{m}_h^0\|_{\boldsymbol{H}^1(\Omega)}^2 + 2\|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}(\|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{L}^2(\Omega)} + \|\boldsymbol{m}_h^0\|_{\boldsymbol{L}^2(\Omega)})$$
$$+ C_3\|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{L}^2(\Omega)}^2 - 2k \sum_{i=0}^{j-1} \langle d_t\boldsymbol{m}_{h\varepsilon}^{i+1}, \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h$$
$$+ 2k \sum_{i=0}^{j-1} \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i, \alpha d_t\boldsymbol{m}_{h\varepsilon}^{i+1} - \mathbb{P}_h\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) + \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h.$$

Using Young's inequality, the first sum on the right-hand side can be estimated as

$$-2k \sum_{i=0}^{j-1} \langle d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h$$

$$= -2k \sum_{i=0}^{j-1} \langle d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})\rangle$$

$$\leq \alpha k \sum_{i=0}^{j-1} \|d_t \boldsymbol{m}_{h\varepsilon}^{i+1}\|_{\boldsymbol{L}^2(\Omega)}^2 + \frac{k}{\alpha} \sum_{i=0}^{j-1} \|\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})\|_{\boldsymbol{L}^2(\Omega)}^2.$$

Since $\boldsymbol{m}_{h\varepsilon}^i \in \mathcal{M}_h$ for all $i = 0, \ldots, j-1$, it holds that

$$\|\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})\|_{\boldsymbol{L}^2(\Omega)} \leq (\|\boldsymbol{\pi}\|_{L(\boldsymbol{L}^2(\Omega); \boldsymbol{L}^2(\Omega))} + 2C_{\boldsymbol{\pi}})|\Omega|^{1/2},$$

and hence

$$\frac{k}{\alpha} \sum_{i=0}^{j-1} \|\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})\|_{\boldsymbol{L}^2(\Omega)}^2 \leq \frac{C|\Omega|(T + k_0)}{\alpha},$$

where $C > 0$ depends only on $\boldsymbol{\pi}$ and $C_{\boldsymbol{\pi}}$. Hence, using the norm equivalence (3.1), we obtain the estimate

$$-2k \sum_{i=0}^{j-1} \langle d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h \leq \alpha k \sum_{i=0}^{j-1} \|d_t \boldsymbol{m}_{h\varepsilon}^{i+1}\|_h^2 + \frac{C|\Omega|(T + k_0)}{\alpha}.$$

Using the estimates

$$\|\mathbb{P}_h \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2})\|_h \overset{(6.1b)}{\leq} (1 + C_{\text{inv}}^2 h^{-2})^{1/2} \|\boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2})\|_{\boldsymbol{H}^1(\Omega)^\star}$$

$$\overset{(6.1d)}{\leq} (1 + C_{\text{inv}}^2 h^{-2})^{1/2} (C_1 \|\boldsymbol{m}_{h\varepsilon}^{i+1/2}\|_{\boldsymbol{H}^1(\Omega)} + \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}),$$

$$\sum_{i=0}^{j-1} \|\boldsymbol{m}_{h\varepsilon}^{i+1/2}\|_{\boldsymbol{H}^1(\Omega)} \leq \frac{1}{2} \sum_{i=0}^{j-1} (\|\boldsymbol{m}_{h\varepsilon}^{i+1}\|_{\boldsymbol{H}^1(\Omega)} + \|\boldsymbol{m}_{h\varepsilon}^i\|_{\boldsymbol{H}^1(\Omega)})$$

$$= \frac{1}{2} \|\boldsymbol{m}_h^0\|_{\boldsymbol{H}^1(\Omega)} + \sum_{i=1}^{j-1} \|\boldsymbol{m}_{h\varepsilon}^i\|_{\boldsymbol{H}^1(\Omega)} + \frac{1}{2} \|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{H}^1(\Omega)}$$

$$\leq \frac{j}{2} + \frac{1}{2} \|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{H}^1(\Omega)}^2 + \frac{1}{2} \sum_{i=0}^{j-1} \|\boldsymbol{m}_{h\varepsilon}^i\|_{\boldsymbol{H}^1(\Omega)}^2,$$

and

$$\|d_t \boldsymbol{m}_{h\varepsilon}^{i+1}\|_{\boldsymbol{L}^2(\Omega)} \leq \frac{1}{4} + \|d_t \boldsymbol{m}_{h\varepsilon}^{i+1}\|_{\boldsymbol{L}^2(\Omega)}^2,$$

together with the stopping criterion $\|\mathcal{I}_h[\boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i]\|_h \leq \varepsilon$ of Algorithm 3.6, if $h$ is sufficiently small, we obtain that

$$2k \sum_{i=0}^{j-1} \langle \boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i, \alpha d_t \boldsymbol{m}_{h\varepsilon}^{i+1} - \mathbb{P}_h \boldsymbol{h}_{\text{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) + \mathbb{P}_h[\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})]\rangle_h$$

$$\leq C'\varepsilon(1 + h^{-1}) + 2k\varepsilon\alpha \sum_{i=0}^{j-1} \|d_t \boldsymbol{m}_{h\varepsilon}^{i+1}\|_h^2$$

$$+ C_1 k\varepsilon(1 + C_{\text{inv}}^2 h^{-2})^{1/2} \|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{H}^1(\Omega)}^2 + C_1 k\varepsilon(1 + C_{\text{inv}}^2 h^{-2})^{1/2} \sum_{i=0}^{j-1} \|\boldsymbol{m}_{h\varepsilon}^i\|_{\boldsymbol{H}^1(\Omega)}^2,$$

where the constant $C' > 0$ depends only on $T$, $|\Omega|$, $\boldsymbol{f}$, $\kappa$, $\boldsymbol{\pi}$, and $C_{\boldsymbol{\pi}}$. Altogether, exploiting the assumption $\varepsilon = \mathcal{O}(h)$ as $h, \varepsilon \to 0$, there exist thresholds $0 < h_0^* \leq h_0$, $0 < k_0^* \leq k_0$, and $\varepsilon_0^* > 0$ as well as constants $A, B > 0$ (depending only on $\kappa$, $T$, and the problem data) such that

$$\|\boldsymbol{m}_{h\varepsilon}^j\|_{\boldsymbol{H}^1(\Omega)}^2 + k \sum_{i=0}^{j-1} \|d_t \boldsymbol{m}_{h\varepsilon}^i\|_h^2 \leq A + Bk \sum_{i=0}^{j-1} \|\boldsymbol{m}_{h\varepsilon}^i\|_{\boldsymbol{H}^1(\Omega)}^2$$

for all $h < h_0^*$, $k < k_0^*$, and $\varepsilon < \varepsilon_0^*$. Then the discrete Gronwall lemma (see, e.g., [44, Lemma 10.5]) and the norm equivalence (3.1) yield (3.17). This concludes the proof of part (iii).

The proof of part (iv) follows the lines of [12, 15, 19, 39]; see also the proof of Theorem 3.2 (iii). In particular, (2.6) and (2.7) are obtained by passing to the limit as $h, k, \varepsilon \to 0$ the discrete identities (3.15) and (3.16), respectively, where the additional contributions arising from the linearization of the nonlinear system (resp., from the explicit treatment of $\boldsymbol{\pi}$), which do not appear in the proof of Theorem 3.2 (iii), vanish in the limit because they are bounded by $\varepsilon$ (resp., because $\boldsymbol{\pi}_h$ is assumed to be consistent with $\boldsymbol{\pi}$). □

# 7 Analysis of the Practical Midpoint Scheme: Newton Iteration

## 7.1 Stability of Algorithm 3.8

Lemma 3.9 and Theorem 3.10 assume well-posedness of Algorithm 3.8 up to time step $i < J$, i.e., for all $n = 0, \dots, i$, the Newton solver (3.18) returns after finitely many iterations the solutions $\boldsymbol{m}_{h\varepsilon}^{n+1}, \boldsymbol{r}_{h\varepsilon}^{n} \in \boldsymbol{V}_h$ such that (3.20) holds with $\|\boldsymbol{r}_{h\varepsilon}^{n}\|_h \leq \varepsilon$. Later, in Sections 7.4.1–7.4.4, Theorem 3.11 is proved, guaranteeing that, given appropriate CFL-conditions, this well-posedness assumption is always satisfied.

### 7.1.1 Boundedness of Magnetization Length, Lemma 3.9 (i)–(ii)

For $0 \leq n \leq i$ and $\boldsymbol{z} \in \mathcal{N}_h$, testing (3.20) with $\boldsymbol{\phi}_h = \boldsymbol{m}_{h\varepsilon}^{n+1/2}(\boldsymbol{z})\varphi_{\boldsymbol{z}} \in \boldsymbol{V}_h$ yields

$$\frac{1}{2k}\beta_{\boldsymbol{z}}(|\boldsymbol{m}_{h\varepsilon}^{n+1}(\boldsymbol{z})|^2 - |\boldsymbol{m}_{h\varepsilon}^{n}(\boldsymbol{z})|^2) = \langle \boldsymbol{r}_{h\varepsilon}^{n}, \boldsymbol{m}_{h\varepsilon}^{n+1/2}(\boldsymbol{z})\varphi_{\boldsymbol{z}} \rangle_h \leq \|\boldsymbol{r}_{h\varepsilon}^{n}\|_h \|\boldsymbol{m}_{h\varepsilon}^{n+1/2}(\boldsymbol{z})\varphi_{\boldsymbol{z}}\|_h$$
$$\leq \varepsilon\beta_{\boldsymbol{z}}^{1/2}|\boldsymbol{m}_{h\varepsilon}^{n+1/2}(\boldsymbol{z})|$$
$$\leq \varepsilon\beta_{\boldsymbol{z}}^{1/2}\left(\frac{1}{2}|\boldsymbol{m}_{h\varepsilon}^{n}(\boldsymbol{z})|^2 + \frac{1}{2}|\boldsymbol{m}_{h\varepsilon}^{n+1}(\boldsymbol{z})|^2 + \frac{1}{4}\right).$$

Rearranging the terms and using $\varepsilon\beta_{\boldsymbol{z}}^{-1/2} \leq C_1\varepsilon h^{-3/2} =: C_{h\varepsilon}$ uniformly for all $\boldsymbol{z} \in \mathcal{N}_h$ shows that, for $k < 1/(2C_{h\varepsilon})$, it holds that

$$|\boldsymbol{m}_{h\varepsilon}^{n+1}(\boldsymbol{z})|^2 \leq \frac{1 + C_{h\varepsilon}k}{1 - C_{h\varepsilon}k}|\boldsymbol{m}_{h\varepsilon}^{n}(\boldsymbol{z})|^2 + C_{h\varepsilon}k = \left(1 + \frac{2C_{h\varepsilon}k}{1 - C_{h\varepsilon}k}\right)|\boldsymbol{m}_{h\varepsilon}^{n}(\boldsymbol{z})|^2 + C_{h\varepsilon}k$$
$$\leq (1 + 4C_{h\varepsilon}k)|\boldsymbol{m}_{h\varepsilon}^{n}(\boldsymbol{z})|^2 + \frac{C_{h\varepsilon}k}{2}.$$

Using $n \leq i < J = T/k$ implies

$$|\boldsymbol{m}_{h\varepsilon}^{n+1}(\boldsymbol{z})|^2 \leq (1 + 4C_{h\varepsilon}k)^{n+1}|\boldsymbol{m}_h^0(\boldsymbol{z})|^2 + C_{h\varepsilon}k\sum_{p=0}^{n}(1 + 4C_{h\varepsilon}k)^p \leq \exp(4C_{h\varepsilon}T)(|\boldsymbol{m}_h^0(\boldsymbol{z})|^2 + C_{h\varepsilon}T).$$

Using $\boldsymbol{m}_h^0 \in \mathcal{M}_h$ and uniform boundedness of $C_{h\varepsilon}$ due to $\varepsilon = \mathcal{O}(h^{3/2})$ concludes the proof of (i). Analogously to the estimate above on $|\boldsymbol{m}_{h\varepsilon}^{n+1}(\boldsymbol{z})|^2$, starting from

$$\frac{1}{2k}\beta_{\boldsymbol{z}}(|\boldsymbol{m}_{h\varepsilon}^{n+1}(\boldsymbol{z})|^2 - |\boldsymbol{m}_{h\varepsilon}^{n}(\boldsymbol{z})|^2) = \langle \boldsymbol{r}_{h\varepsilon}^{n}, \boldsymbol{m}_{h\varepsilon}^{n+1/2}(\boldsymbol{z})\varphi_{\boldsymbol{z}} \rangle_h \geq -\|\boldsymbol{r}_{h\varepsilon}^{n}\|_h \|\boldsymbol{m}_{h\varepsilon}^{n+1/2}(\boldsymbol{z})\varphi_{\boldsymbol{z}}\|_h,$$

by a similar computation, one derives an estimate below via

$$|\boldsymbol{m}_{h\varepsilon}^{n+1}(\boldsymbol{z})|^2 \geq (1 - 4C_{h\varepsilon}k)^{n+1}|\boldsymbol{m}_h^0(\boldsymbol{z})|^2 - C_{h\varepsilon}k\sum_{p=0}^{n}(1 - 4C_{h\varepsilon}k)^p$$
$$\geq \exp(-8C_{h\varepsilon}T)|\boldsymbol{m}_h^0(\boldsymbol{z})|^2 - \exp(-4C_{h\varepsilon}T)C_{h\varepsilon}T \quad \text{for all } 0 < k < k_0,$$

where $k_0$ can be uniformly chosen since $\varepsilon = \mathcal{O}(h^{3/2})$. If $\varepsilon = o(h^{3/2})$, then in both estimates $C_{h\varepsilon}$ tends to zero as $h, \varepsilon \to 0$. Hence, also statement (ii) holds true. □

### 7.1.2 Stability and Weak Convergence, Theorem 3.10 (i)–(iii)

For $0 \le i < J$ testing (3.20) with $\boldsymbol{\phi}_h = \alpha d_t \boldsymbol{m}_{h\varepsilon}^{i+1} - \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) + \mathbb{P}_h(\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}))$ yields

$$\alpha \| d_t \boldsymbol{m}_{h\varepsilon}^{i+1} \|_h^2 - \langle d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) \rangle_h + \langle d_t \boldsymbol{m}_{h\varepsilon}^{i+1}, \mathbb{P}_h(\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})) \rangle_h$$
$$= \langle \boldsymbol{r}_{h\varepsilon}^i, \alpha d_t \boldsymbol{m}_{h\varepsilon}^{i+1} - \mathbb{P}_h \boldsymbol{h}_{\mathrm{eff}}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) + \mathbb{P}_h(\boldsymbol{\pi}(\boldsymbol{m}_{h\varepsilon}^{i+1/2}) - \boldsymbol{\Pi}_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})) \rangle_h.$$

Up to replacing $\boldsymbol{r}_{h\varepsilon}^i$ by $\boldsymbol{m}_{h\varepsilon}^{i+1/2} \times \boldsymbol{r}_{h\varepsilon}^i$, this identity resembles the first identity in Section 6.2, where Theorem 3.7 (ii)–(iv) is proved. Hence, using $\boldsymbol{L}^\infty(\Omega)$-boundedness of the iterates from Lemma 3.9 (i) and that the stopping criterion (3.19) guarantees $\|\boldsymbol{r}_{h\varepsilon}^i\|_h \le \varepsilon$, the proof of Theorem 3.10 (i)–(iii) directly follows the lines of Section 6.2. □

## 7.2 Main theorem on Newton's method

Newton's method is an iterative scheme to generate a converging sequence of approximate solutions to the following problem: given $\boldsymbol{F} \colon \mathbb{R}^n \to \mathbb{R}^n$,

$$\text{find } \boldsymbol{x}^* \in \mathbb{R}^n \text{ such that } \boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{0}. \tag{7.1}$$

Here, $\boldsymbol{F}$ is considered to be $C^1$-continuous on a convex open set $D \subseteq \mathbb{R}^n$ containing $\boldsymbol{x}^*$ and the Jacobian of $\boldsymbol{F}$ evaluated at $\boldsymbol{x} \in \mathbb{R}^n$ is denoted by $\nabla \boldsymbol{F}(\boldsymbol{x}) \in \mathbb{R}^{n \times n}$. Given a starting value $\boldsymbol{x}^0 \in \mathbb{R}^n$, Newton's method applied to (7.1) iterates, for all $\ell \in \mathbb{N}_0$,

$$\begin{aligned} \text{solve} \quad & \nabla \boldsymbol{F}(\boldsymbol{x}^\ell) \boldsymbol{\delta x}^\ell = -\boldsymbol{F}(\boldsymbol{x}^\ell), \\ \text{set} \quad & \boldsymbol{x}^{\ell+1} = \boldsymbol{x}^\ell + \boldsymbol{\delta x}^\ell. \end{aligned} \tag{7.2}$$

Given a vector norm $\| \cdot \|$ on $\mathbb{R}^n$, by $B(\| \cdot \|; \boldsymbol{x}, R)$, the open unit ball of radius $R > 0$ around $\boldsymbol{x} \in \mathbb{R}^n$ with respect to the norm $\| \cdot \|$ is denoted. In accordance with [41, Definition 1.20], a matrix norm $\| \cdot \|_{\mathbb{R}^{n \times n}}$ and a vector norm $\| \cdot \|_{\mathbb{R}^n}$ are called *consistent* if it holds that $\| \boldsymbol{A x} \|_{\mathbb{R}^n} \le \| \boldsymbol{A} \|_{\mathbb{R}^{n \times n}} \| \boldsymbol{x} \|_{\mathbb{R}^n}$ for all $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ and all $\boldsymbol{x} \in \mathbb{R}^n$. Clearly, any vector norm is consistent with the natural matrix norm induced by the vector norm defined as

$$\| \boldsymbol{A} \| = \sup_{\boldsymbol{x} \in \mathbb{R}^n \setminus \{\boldsymbol{0}\}} \frac{\| \boldsymbol{A x} \|}{\| \boldsymbol{x} \|} \quad \text{for all } \boldsymbol{A} \in \mathbb{R}^{n \times n}. \tag{7.3}$$

Using the above notation, we recall the classical local convergence result for Newton's method.

**Theorem 7.1** ([41, Theorem 7.1]). *For a convex open set $D \subseteq \mathbb{R}^n$ with $\boldsymbol{x}^* \in D$, let $\boldsymbol{F} \in C^1(D; \mathbb{R}^n)$ with $\boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{0}$. Suppose that $(\nabla \boldsymbol{F}(\boldsymbol{x}^*))^{-1} \in \mathbb{R}^{n \times n}$ exists and that there exist constants $C, R, L > 0$ such that*

$$\| (\nabla \boldsymbol{F}(\boldsymbol{x}^*))^{-1} \| \le C \tag{7.4a}$$

*and*

$$\| \nabla \boldsymbol{F}(\boldsymbol{x}) - \nabla \boldsymbol{F}(\boldsymbol{y}) \| \le L \| \boldsymbol{x} - \boldsymbol{y} \| \quad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in B(\| \cdot \|; \boldsymbol{x}^*, R), \tag{7.4b}$$

*where the symbol $\| \cdot \|$ denotes two consistent vector and matrix norms. Then there holds that, for any*

$$\boldsymbol{x}^0 \in B(\| \cdot \|; \boldsymbol{x}^*, \min\{R, 1/(2CL)\}),$$

*the sequence $(\boldsymbol{x}^\ell)_{\ell \in \mathbb{N}}$ generated by Newton's method (7.2) is uniquely defined and converges to $\boldsymbol{x}^*$ with*

$$\| \boldsymbol{x}^{\ell+1} - \boldsymbol{x}^* \| \le CL \| \boldsymbol{x}^\ell - \boldsymbol{x}^* \|^2. \tag{7.5}$$

**Remark 7.2.** In particular, (7.5) and $\boldsymbol{x}^0 \in B(\| \cdot \|; \boldsymbol{x}^*, \min\{R, 1/(2CL)\})$ imply

$$\| \boldsymbol{x}^\ell - \boldsymbol{x}^* \| \le \left( \prod_{j=0}^{\ell-1} (CL)^{2^j} \right) \| \boldsymbol{x}^0 - \boldsymbol{x}^* \|^{2^\ell} = (CL)^{2^\ell - 1} \| \boldsymbol{x}^0 - \boldsymbol{x}^* \|^{2^\ell} \le \frac{2 \| \boldsymbol{x}^0 - \boldsymbol{x}^* \|}{2^{2^\ell}}$$

for all $\ell \in \mathbb{N}_0$. Hence, there holds $\| \boldsymbol{x}^\ell - \boldsymbol{x}^* \|, \| \boldsymbol{\delta x}^\ell \| \to 0$ for $\ell \to \infty$.

## 7.3 Newton's Method Applied to the Nonlinear Midpoint Scheme

We aim to apply Newton's method (7.2) to the nonlinear system of equations (3.11), i.e., to the IMEX version of the ideal midpoint scheme where the lower order terms are integrated explicitly in time

$$\pi(m_h^{i+1/2}) \approx \Pi_h(m_h^i, m_h^{i-1}).$$

Consider a numbering of the nodes $\{z_j : j = 1, \ldots, N\} = \mathcal{N}_h$ of the mesh $\mathcal{T}_h$, and associate with a given vector $x \in (\mathbb{R}^3)^N$ the finite element function defined by $\hat{x} := \sum_{j=1}^N x_j \varphi_{z_j} \in V_h$. Further, for a finite element function $u = \sum_{j=1}^N u(z_j)\varphi_{z_j} \in V_h$, we write $[u] \in (\mathbb{R}^3)^N \simeq \mathbb{R}^{3N}$ for the vector of nodal values, i.e., $[u]_j := u(z_j) \in \mathbb{R}^3$.

The mass lumped scalar product $\langle \cdot, \cdot \rangle_h$ gives rise to the matrix $\mathbf{M}_h \in (\mathbb{R}^{3\times3})^{N\times N} \simeq \mathbb{R}^{3N\times3N}$, defined via $(\mathbf{M}_h)_{jk} := \delta_{j,k}\beta_{z_j}\mathbf{I}_{3\times3} \in \mathbb{R}^{3\times3}$. Given $m_h^i \in V_h$, the solution $m_h^{i+1/2}$ of (3.11) satisfies $F([m_h^{i+1/2}]) = \mathbf{0}$, with

$$F(x) := \mathbf{M}_h\Big(x - [m_h^i] + \Big[\mathcal{I}_h\Big(\frac{k}{2}\hat{x} \times \mathbb{P}_h(h_{\text{eff}}^{\text{loc}}(\hat{x}) + \Pi_h(m_h^i, m_h^{i-1})) + \alpha\hat{x} \times m_h^i\Big)\Big]\Big). \tag{7.6}$$

The Jacobian $\nabla F: \mathbb{R}^{3N} \to \mathbb{R}^{3N\times3N}$ satisfies for all $x, u, v \in \mathbb{R}^{3N}$ that

$$u^\top \nabla F(x)v = \langle \hat{u}, \hat{v} \rangle_h + \frac{k}{2}\langle \hat{u} \times \mathbb{P}_h h_{\text{eff}}^{\text{loc}}(\hat{x}), \hat{v} \rangle_h + \frac{k}{2}\langle \hat{x} \times \mathbb{P}_h(h_{\text{eff}}^{\text{loc}}(\hat{u}) - f), \hat{v} \rangle_h$$

$$+ \frac{k}{2}\langle \hat{u} \times \mathbb{P}_h\Pi_h(m_h^i, m_h^{i-1}), \hat{v} \rangle_h + \alpha\langle \hat{u} \times m_h^i, \hat{v} \rangle_h. \tag{7.7}$$

Newton's method (7.2) applied to system (3.11) in the $i$-th time step now can be written as: given $m_{h\varepsilon}^i \in V_h$ and initial value $\eta_h^{i,0} \in V_h$, for all $\ell \in \mathbb{N}_0$, compute $u_h^{i,\ell} \in V_h$ such that, for all $\phi_h \in V_h$, it holds that

$$\langle u_h^{i,\ell}, \phi_h \rangle_h + \frac{k}{2}\langle u_h^{i,\ell} \times \mathbb{P}_h h_{\text{eff}}^{\text{loc}}(\eta_h^{i,\ell}), \phi_h \rangle_h + \frac{k}{2}\langle \eta_h^{i,\ell} \times \mathbb{P}_h(h_{\text{eff}}^{\text{loc}}(u_h^{i,\ell}) - f), \phi_h \rangle_h$$

$$+ \frac{k}{2}\langle u_h^{i,\ell} \times \mathbb{P}_h\Pi_h(m_{h\varepsilon}^i, m_{h\varepsilon}^{i-1}), \phi_h \rangle_h + \alpha\langle u_h^{i,\ell} \times m_{h\varepsilon}^i, \phi_h \rangle_h$$

$$= \langle m_{h\varepsilon}^i - \eta_h^{i,\ell}, \phi_h \rangle_h - \frac{k}{2}\langle \eta_h^{i,\ell} \times \mathbb{P}_h h_{\text{eff}}^{\text{loc}}(\eta_h^{i,\ell}), \phi_h \rangle_h$$

$$- \frac{k}{2}\langle \eta_h^{i,\ell} \times \mathbb{P}_h\Pi_h(m_{h\varepsilon}^i, m_{h\varepsilon}^{i-1}), \phi_h \rangle_h - \alpha\langle \eta_h^{i,\ell} \times m_{h\varepsilon}^i, \phi_h \rangle_h, \tag{7.8a}$$

and define

$$\eta_h^{i,\ell+1} := \eta_h^{i,\ell} + u_h^{i,\ell}. \tag{7.8b}$$

In the remainder of this section, to improve readability, we omit the $h$-subscript of the iteration variables $\eta^{i,\ell}$ and $u^{i,\ell}$. Note that, by (7.6)–(7.7), we see that (7.8a)–(7.8b) resembles Newton's method (7.2) with $x^\ell = [\eta^{i,\ell}]$ and $\delta x^\ell = [u^{i,\ell}]$. Given some tolerance $\varepsilon > 0$, the iteration will be stopped once

$$\big\|\mathcal{I}_h(u^{i,\ell} \times \mathbb{P}_h(h_{\text{eff}}^{\text{loc}}(u^{i,\ell}) - f))\big\|_h \leq \varepsilon. \tag{7.9}$$

If $\ell^* \in \mathbb{N}_0$ is the first index for which the stopping criterion (7.9) is satisfied, the approximate magnetization at the new time step is defined as $m_{h\varepsilon}^{i+1} := 2\eta^{i,\ell^*+1} - m_{h\varepsilon}^i$.

For all $i \in \mathbb{N}_0$, let $r_{h\varepsilon}^i := \mathcal{I}_h(u^{i,\ell^*} \times \mathbb{P}_h(h_{\text{eff}}^{\text{loc}}(u^{i,\ell^*}) - f)) \in V_h$ so that $\langle r_{h\varepsilon}^i, \phi_h \rangle_h$ equals the difference of (3.11) and (7.8a). In view of the stopping criterion (7.9), it holds that $\|r_{h\varepsilon}^i\|_h \leq \varepsilon$. With this definition, the proposed linearization of one iteration of Algorithm 3.1 based on the Newton method is covered by Algorithm 3.8.

## 7.4 Well-Posedness of Algorithm 3.8

We show Theorem 3.11 (i) by induction: for $0 \leq i < J$, assume that Algorithm 3.8 is well-defined for all $n = 0, \ldots, i - 1$. In particular, by Lemma 3.9 and Theorem 3.10, we have the bounds

$$\|m_{h\varepsilon}^i\|_{L^\infty(\Omega)} \leq C_\infty \quad \text{and} \quad \mathcal{E}(m_{h\varepsilon}^i) \leq C. \tag{7.10}$$

Now the inductive step is to prove convergence of the Newton iteration (3.18)–(3.19) for time step $n = i$. We do this by verifying the assumptions (7.4a)–(7.4b) of Theorem 7.1 for the Newton solver (7.8a) with the initial value chosen as $\boldsymbol{\eta}^{i,0} := \boldsymbol{m}_{h\varepsilon}^i$. In Section 7.4.1, we verify the Lipschitz continuity (7.4b). Invertibility (7.4a) is shown in Section 7.4.2. In Section 7.4.3, we prove that, under the assumed CFL-conditions, the initial guess $\boldsymbol{\eta}^{i,0} := \boldsymbol{m}_{h\varepsilon}^i$ is an appropriate choice, which guarantees convergence of Newton's method. Finally, in Section 7.4.4, we conclude by estimating the maximum number of Newton iterations required to achieve the required tolerance (7.9), in particular showing that the number is finite. Hence, Sections 7.4.1–7.4.4 prove Theorem 3.11.

Throughout the proof, we use the notation of Sections 7.2–7.3 and consider the $\ell^2$-norm on $(\mathbb{R}^3)^N \simeq \mathbb{R}^{3N}$ defined by $\|\boldsymbol{x}\|_2 = \sum_{j=1}^N |\boldsymbol{x}_j|^2$, as well as the induced matrix norm on $(\mathbb{R}^{3\times3})^{N\times N} \simeq \mathbb{R}^{3N\times3N}$ also denoted by $\|\cdot\|_2$; cf. (7.3).

### 7.4.1 Lipschitz Continuity of $\nabla F$

By (7.3), it holds for arbitrary $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{3N}$ that

$$\|\nabla \boldsymbol{F}(\boldsymbol{x}) - \nabla \boldsymbol{F}(\boldsymbol{y})\|_2 = \sup_{\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^{3N}\setminus\{\boldsymbol{0}\}} \frac{\boldsymbol{u}^\top(\nabla \boldsymbol{F}(\boldsymbol{x}) - \nabla \boldsymbol{F}(\boldsymbol{y}))\boldsymbol{v}}{\|\boldsymbol{u}\|_2\|\boldsymbol{v}\|_2}.$$

With representation (7.7) and estimates (6.1), we see

$$\begin{aligned}
\boldsymbol{u}^\top(\nabla \boldsymbol{F}(\boldsymbol{x}) - \nabla \boldsymbol{F}(\boldsymbol{y}))\boldsymbol{v} &= \frac{k}{2}\langle\hat{\boldsymbol{u}} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\hat{\boldsymbol{x}}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\hat{\boldsymbol{y}})), \hat{\boldsymbol{v}}\rangle_h + \frac{k}{2}\langle(\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}) \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\hat{\boldsymbol{u}}) - \boldsymbol{f}), \hat{\boldsymbol{v}}\rangle_h \\
&\lesssim k\|\hat{\boldsymbol{u}}\|_{\boldsymbol{L}^\infty(\Omega)}\|\mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\hat{\boldsymbol{x}}) - \boldsymbol{h}_{\text{eff}}^{\text{loc}}(\hat{\boldsymbol{y}}))\|_h\|\hat{\boldsymbol{v}}\|_h + k\|\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}\|_{\boldsymbol{L}^\infty(\Omega)}\|\mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\hat{\boldsymbol{u}}) - \boldsymbol{f})\|_h\|\hat{\boldsymbol{v}}\|_h \\
&\lesssim kh^{-2}(\|\boldsymbol{u}\|_2\|\hat{\boldsymbol{x}} - \hat{\boldsymbol{y}}\|_{\boldsymbol{L}^2(\Omega)} + \|\boldsymbol{x} - \boldsymbol{y}\|_2\|\hat{\boldsymbol{u}}\|_{\boldsymbol{L}^2(\Omega)})\|\hat{\boldsymbol{v}}\|_h.
\end{aligned}$$

With the norm equivalence $h^{3/2}\|\cdot\|_2 \simeq \|\hat{\cdot}\|_h \simeq \|\hat{\cdot}\|_{\boldsymbol{L}^2(\Omega)}$ on $\mathbb{R}^{3N}$, we get uniformly for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^{3N}$ that

$$\|\nabla \boldsymbol{F}(\boldsymbol{x}) - \nabla \boldsymbol{F}(\boldsymbol{y})\|_2 \lesssim kh\|\boldsymbol{x} - \boldsymbol{y}\|_2.$$

In particular, (7.4b) holds for $\|\cdot\|_2$ with $R = +\infty$ and $L \simeq kh$. $\qquad\square$

### 7.4.2 Invertibility of $\nabla F(x^*)$

The unknown $\boldsymbol{x}^* \in \mathbb{R}^{3N}$ is defined by $\boldsymbol{F}(\boldsymbol{x}^*) = \boldsymbol{0}$. Hence,

$$0 = [\widehat{\boldsymbol{x}^*}(\boldsymbol{z}_j)\varphi_{\boldsymbol{z}_j}]^\top \boldsymbol{F}(\boldsymbol{x}^*) \overset{(7.6)}{=} \beta_{\boldsymbol{z}_j}(|\boldsymbol{x}_j^*|^2 - \boldsymbol{x}_j^* \cdot \boldsymbol{m}_{h\varepsilon}^i(\boldsymbol{z}_j)) \quad \text{for all } j = 1, \dots, N$$

together with (7.10) guarantees boundedness

$$\|\widehat{\boldsymbol{x}^*}\|_{\boldsymbol{L}^\infty(\Omega)} \leq \|\boldsymbol{m}_{h\varepsilon}^i\|_{\boldsymbol{L}^\infty(\Omega)} \leq C_\infty. \tag{7.11}$$

Now the assumption $k = o(h^2)$ guarantees invertibility of $\nabla \boldsymbol{F}(\boldsymbol{x}^*)$ by ellipticity

$$\boldsymbol{y}^\top\nabla \boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{y} \overset{(7.7)}{=} \|\hat{\boldsymbol{y}}\|_h^2 + \frac{k}{2}\langle\widehat{\boldsymbol{x}^*} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\hat{\boldsymbol{y}}) - \boldsymbol{f}), \hat{\boldsymbol{y}}\rangle_h \geq (1 - kh^{-2})\|\hat{\boldsymbol{y}}\|_h^2 \gtrsim \|\hat{\boldsymbol{y}}\|_h^2,$$

where we used (7.11) and the Cauchy–Schwarz inequality together with estimates (6.1). To show boundedness of $(\nabla \boldsymbol{F}(\boldsymbol{x}^*))^{-1}$, we write

$$\|(\nabla \boldsymbol{F}(\boldsymbol{x}^*))^{-1}\|_2 = \sup_{\boldsymbol{x}\in\mathbb{R}^{3N}\setminus\{\boldsymbol{0}\}} \frac{\|(\nabla \boldsymbol{F}(\boldsymbol{x}^*))^{-1}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = \sup_{\boldsymbol{x}\in\mathbb{R}^{3N}\setminus\{\boldsymbol{0}\}} \frac{\|\boldsymbol{y}(\boldsymbol{x})\|_2}{\|\boldsymbol{x}\|_2},$$

with $\boldsymbol{y} := \boldsymbol{y}(\boldsymbol{x}) := (\nabla \boldsymbol{F}(\boldsymbol{x}^*))^{-1}\boldsymbol{x}$. Using (7.7), it holds that

$$\boldsymbol{y}^\top\boldsymbol{x} = \boldsymbol{y}^\top\nabla \boldsymbol{F}(\boldsymbol{x}^*)\boldsymbol{y} = \|\hat{\boldsymbol{y}}\|_h^2 + \frac{k}{2}\langle\widehat{\boldsymbol{x}^*} \times \mathbb{P}_h(\boldsymbol{h}_{\text{eff}}^{\text{loc}}(\hat{\boldsymbol{y}}) - \boldsymbol{f}), \hat{\boldsymbol{y}}\rangle_h.$$

Using norm equivalences $h^{3/2}\|\cdot\|_2 \simeq \|\hat{\cdot}\|_h \simeq \|\hat{\cdot}\|_{L^2(\Omega)}$ on $\mathbb{R}^{3N}$ and an inverse estimate, it follows that

$$h^3\|\boldsymbol{y}\|_2^2 \lesssim \|\hat{\boldsymbol{y}}\|_h^2 = \boldsymbol{y}^\top \boldsymbol{x} - \frac{k}{2}\langle \widehat{\boldsymbol{x}^*} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\hat{\boldsymbol{y}}) - \boldsymbol{f}), \hat{\boldsymbol{y}}\rangle_h$$

$$\lesssim \|\boldsymbol{y}\|_2\|\boldsymbol{x}\|_2 + k\|\widehat{\boldsymbol{x}^*}\|_{\boldsymbol{L}^\infty(\Omega)}\|\mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\hat{\boldsymbol{y}}) - \boldsymbol{f})\|_h\|\hat{\boldsymbol{y}}\|_h$$

$$\overset{(6.1)}{\lesssim} \|\boldsymbol{y}\|_2\|\boldsymbol{x}\|_2 + kh^{-2}\|\hat{\boldsymbol{y}}\|_h^2 \lesssim \|\boldsymbol{y}\|_2\|\boldsymbol{x}\|_2 + kh\|\boldsymbol{y}\|_2^2.$$

With the CFL condition $k = o(h^2)$, we estimate $h^3(1 - kh^{-2})\|\boldsymbol{y}\|_2 \lesssim h^3\|\boldsymbol{y}\|_2 \lesssim \|\boldsymbol{x}\|_2$ and conclude that

$$\|(\nabla \boldsymbol{F}(\boldsymbol{x}^*))^{-1}\|_2 \lesssim h^{-3}.$$

In particular, (7.4a) holds for $\|\cdot\|_2$ with $C \simeq h^{-3}$. ☐

### 7.4.3 Initial Guess Leads to Convergence

We recall the results from Section 7.4.1 and Section 7.4.2: the Newton iteration (7.8a) satisfies the assumptions of Theorem 7.1 for $\|\cdot\|_2$ with $C \simeq h^{-3}$, $R = +\infty$ and $L \simeq kh$. The theorem now guarantees convergence $\boldsymbol{x}^\ell \to \boldsymbol{x}^*$ in $\|\cdot\|_2$ as $\ell \to \infty$ of the Newton iteration (7.8a) for any initial guess

$$\boldsymbol{x}^0 \in \mathbb{R}^{3N} \quad \text{with } \|\boldsymbol{x}^* - \boldsymbol{x}^0\|_2 \le 1/(2CL) \simeq h^2/k.$$

Given $\boldsymbol{m}_{h\varepsilon}^i \in \boldsymbol{V}_h$, Algorithm 3.8 defines the initial guess as $\boldsymbol{x}^0 := [\boldsymbol{\eta}^{i,0}] := [\boldsymbol{m}_{h\varepsilon}^i]$. Let $\boldsymbol{x}^* \in \mathbb{R}^{3N}$ be the solution of (7.1), i.e., by (7.6), it holds for all $\boldsymbol{\phi}_h \in \boldsymbol{V}_h$ that

$$\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{\phi}_h\rangle_h = -\frac{k}{2}\langle \widehat{\boldsymbol{x}^*} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\widehat{\boldsymbol{x}^*}) + \Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})), \boldsymbol{\phi}_h\rangle_h + \alpha\langle \widehat{\boldsymbol{x}^*} \times (\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i), \boldsymbol{\phi}_h\rangle_h.$$

Using $\boldsymbol{\phi}_h = \alpha(\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i) - (k/2)\mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\widehat{\boldsymbol{x}^*}) + \Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1})) \in \boldsymbol{V}_h$ shows

$$\alpha\|\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i\|_h^2 = \frac{k}{2}\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\widehat{\boldsymbol{x}^*}) + \Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}))\rangle_h$$

$$= \frac{k}{2}\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \mathbb{P}_h\boldsymbol{h}_{\mathrm{eff}}(\widehat{\boldsymbol{x}^*})\rangle_h + \frac{k}{2}\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \mathbb{P}_h(\Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}) - \boldsymbol{\pi}(\widehat{\boldsymbol{x}^*}))\rangle_h.$$

We rewrite the first term on the right-hand side as

$$4\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \mathbb{P}_h\boldsymbol{h}_{\mathrm{eff}}(\widehat{\boldsymbol{x}^*})\rangle_h$$

$$\overset{(3.2)}{=} 4\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{h}_{\mathrm{eff}}(\widehat{\boldsymbol{x}^*})\rangle$$

$$\overset{(2.3)}{=} -4a(\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \widehat{\boldsymbol{x}^*}) + 4\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{f}\rangle_\Omega$$

$$= -a((2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i) - \boldsymbol{m}_{h\varepsilon}^i, (2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i) + \boldsymbol{m}_{h\varepsilon}^i) + 2\langle 2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{f}\rangle_\Omega - 2\langle \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{f}\rangle_\Omega$$

$$\overset{(2.1)}{=} 2\mathcal{E}(\boldsymbol{m}_{h\varepsilon}^i) - a(2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, 2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i) + 2\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{f}\rangle_\Omega.$$

With (7.11) and the Gårding inequality (5.2c), we estimate

$$-a(2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, 2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i) \le C_3\|2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i\|_{\boldsymbol{L}^2(\Omega)}^2 - C_2\|2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i\|_{\boldsymbol{H}^1(\Omega)}^2 \le 3C_3C_\infty^2|\Omega|.$$

Now combination with the generous estimates

$$\langle \widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \mathbb{P}_h(\Pi_h(\boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{m}_{h\varepsilon}^{i-1}) - \boldsymbol{\pi}(\widehat{\boldsymbol{x}^*}))\rangle_h \le 2C_\infty^2|\Omega|(2C_{\boldsymbol{\pi}} + \|\boldsymbol{\pi}\|_{L(\boldsymbol{L}^2(\Omega);\boldsymbol{L}^2(\Omega))}),$$

$$\langle 2\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i, \boldsymbol{f}\rangle_\Omega \le 3C_\infty\|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)}$$

yields

$$\|\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i\|_h^2 \lesssim \mathcal{E}(\boldsymbol{m}_{h\varepsilon}^i)k + C(C_3, C_\infty, |\Omega|, C_{\boldsymbol{\pi}}, \|\boldsymbol{\pi}\|_{L(\boldsymbol{L}^2(\Omega);\boldsymbol{L}^2(\Omega))}, \|\boldsymbol{f}\|_{\boldsymbol{L}^2(\Omega)})k \overset{(7.10)}{\lesssim} k.$$

Due to the norm equivalence $h^{3/2}\|\cdot\|_2 \simeq \|\hat{\cdot}\|_h$ on $\mathbb{R}^{3N}$, the claim $\boldsymbol{x}^0 \in B(\|\cdot\|_2; \boldsymbol{x}^*, 1/(2CL))$ follows for $h, k \to 0$ from $k = o(h^{7/3})$ via

$$\|\boldsymbol{x}^* - \boldsymbol{x}^0\|_2 \simeq h^{-3/2}\|\widehat{\boldsymbol{x}^*} - \boldsymbol{m}_{h\varepsilon}^i\|_h \lesssim k^{1/2}h^{-3/2} = k^{-1}h^2(k^{3/2}h^{-7/2}) < k^{-1}h^2 \simeq 1/(2CL).$$

Hence, the choice $\boldsymbol{\eta}^{i,0} := \boldsymbol{m}_{h\varepsilon}^i$ implies convergence $\boldsymbol{x}^\ell \to \boldsymbol{x}^*$ in $\|\cdot\|_2$ as $\ell \to \infty$. ☐

### 7.4.4 Finite Number of Newton Iterations

In the previous section, we showed $\boldsymbol{x}^\ell \to \boldsymbol{x}^*$ in $\|\cdot\|_2$ and therefore also $\|\boldsymbol{\delta x}^\ell\|_2 \to 0$ as $\ell \to \infty$. Now let $\ell^* \in \mathbb{N}$ be the smallest integer such that (7.9) is satisfied. The index $\ell^*$ is well-defined due to $\widehat{\boldsymbol{\delta x}^\ell} = \boldsymbol{u}^{i,\ell}$ and

$$\left\|\mathcal{I}_h\big(\widehat{\boldsymbol{\delta x}^\ell} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\widehat{\boldsymbol{\delta x}^\ell}) - \boldsymbol{f})\big)\right\|_h \lesssim \|\boldsymbol{\delta x}^\ell\|_\infty h^{-2} \|\widehat{\boldsymbol{\delta x}^\ell}\|_h \lesssim h^{-1/2} \|\boldsymbol{\delta x}^\ell\|_2^2 \to 0 \quad \text{as } \ell \to \infty.$$

Recalling that, by Remark 7.2, it holds that

$$\|\boldsymbol{x}^* - \boldsymbol{x}^\ell\|_2 \le \frac{2\|\boldsymbol{x}^0 - \boldsymbol{x}^*\|_2}{2^{2^\ell}},$$

we estimate the index $\ell^* \in \mathbb{N}$: with estimates (6.1) and the norm equivalence $h^{3/2}\|\cdot\|_2 \simeq \|\hat{\cdot}\|_h$ on $\mathbb{R}^{3N}$, it holds for the error $\boldsymbol{r}_{h\varepsilon}^i$ that

$$\begin{aligned}
\|\boldsymbol{r}_{h\varepsilon}^i\|_h &= \left\|\mathcal{I}_h\big(\boldsymbol{u}^{i,\ell^*} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\boldsymbol{u}^{i,\ell^*}) - \boldsymbol{f})\big)\right\|_h = \left\|\mathcal{I}_h\big(\widehat{\boldsymbol{\delta x}^{\ell^*}} \times \mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\widehat{\boldsymbol{\delta x}^{\ell^*}}) - \boldsymbol{f})\big)\right\|_h \\
&\le \|\widehat{\boldsymbol{\delta x}^{\ell^*}}\|_{\boldsymbol{L}^\infty(\Omega)} \|\mathbb{P}_h(\boldsymbol{h}_{\mathrm{eff}}^{\mathrm{loc}}(\widehat{\boldsymbol{\delta x}^{\ell^*}}) - \boldsymbol{f})\|_h \lesssim \|\boldsymbol{\delta x}^{\ell^*}\|_\infty h^{-2} \|\widehat{\boldsymbol{\delta x}^{\ell^*}}\|_h \\
&\lesssim h^{-1/2} \|\boldsymbol{\delta x}^{\ell^*}\|_2^2 = h^{-1/2} \|\boldsymbol{x}^{\ell^*+1} - \boldsymbol{x}^{\ell^*}\|_2^2 \lesssim h^{-1/2}(\|\boldsymbol{x}^* - \boldsymbol{x}^{\ell^*+1}\|_2^2 + \|\boldsymbol{x}^* - \boldsymbol{x}^{\ell^*}\|_2^2) \\
&\lesssim 2^{-2^{\ell^*}} h^{-1/2} \|\boldsymbol{x}^0 - \boldsymbol{x}^*\|_2^2 \lesssim 2^{-2^{\ell^*}} k h^{-7/2}.
\end{aligned}$$

Since $\ell^* \in \mathbb{N}$ is defined as the smallest integer such that (7.9) is satisfied, $\ell^*$ is estimated from above by $\log_2 \log_2(C_\star k h^{-7/2} \varepsilon^{-1})$ with a generic constant $C_\star > 0$. □

# 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), no. 6, 639–654.

[2] E. Acerbi, I. Fonseca and G. Mingione, Existence and regularity for mixtures of micromagnetic materials, *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **462** (2006), no. 2072, 2225–2243.

[3] G. Akrivis, M. Feischl, B. Kovács and C. Lubich, Higher-order linearly implicit full discretization of the Landau–Lifshitz–Gilbert equation, *Math. Comp.* **90** (2021), no. 329, 995–1038.

[4] F. Alouges, A new finite element scheme for Landau–Lifchitz equations, *Discrete Contin. Dyn. Syst. Ser. S* **1** (2008), no. 2, 187–196.

[5] F. Alouges, A. de Bouard, B. Merlet and L. Nicolas, Stochastic homogenization of the Landau–Lifshitz–Gilbert equation, *Stoch. Partial Differ. Equ. Anal. Comput.* **9** (2021), no. 4, 789–818.

[6] F. Alouges and G. Di Fratta, Homogenization of composite ferromagnetic materials, *Proc. A.* **471** (2015), no. 2182, Article ID 20150365.

[7] F. Alouges and P. Jaisson, Convergence of a finite element discretization for the Landau–Lifshitz equations in micromagnetism, *Math. Models Methods Appl. Sci.* **16** (2006), no. 2, 299–316.

[8] 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), no. 3, 407–430.

[9] F. Alouges, E. Kritsikis and J.-C. Toussaint, A convergent finite element approximation for Landau–Lifshitz–Gilbert equation, *Phys. B* **407** (2012), no. 9, 1345–1349.

[10] F. Alouges and A. Soyeur, On global weak solutions for Landau–Lifshitz equations: Existence and nonuniqueness, *Nonlinear Anal.* **18** (1992), no. 11, 1071–1084.

[11] Ľ. Baňas, Z. A. Brzeźniak, M. Neklyudov and A. Prohl, *Stochastic Ferromagnetism: Analysis and Numerics*, De Gruyter Stud. Math. 58, De Gruyter, Berlin, 2014.

[12] S. Bartels, Constraint preserving, inexact solution of implicit discretizations of Landau–Lifshitz–Gilbert equations and consequences for convergence, *PAMM. Proc. Appl. Math. Mech.* **6** (2006), no. 1, 19–22.

[13] S. Bartels, *Numerical Methods for Nonlinear Partial Differential Equations*, Springer Ser. Comput. Math. 47, Springer, Cham, 2015.

[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), no. 262, 773–788.

[15] S. Bartels and A. Prohl, Convergence of an implicit finite element method for the Landau–Lifshitz–Gilbert equation, *SIAM J. Numer. Anal.* **44** (2006), no. 4, 1405–1419.

[16] W. F. Brown, *Micromagnetics*, Interscience, New York, 1963.

[17] F. Bruckner, D. Suess, M. Feischl, T. Führer, P. Goldenits, M. Page, D. Praetorius and M. Ruggeri, Multiscale modeling in micromagnetics: Existence of solutions and numerical integration, *Math. Models Methods Appl. Sci.* **24** (2014), no. 13, 2627–2662.

[18] I. Cimrák, A survey on the numerics and computations for the Landau–Lifshitz equation of micromagnetism, *Arch. Comput. Methods Eng.* **15** (2008), no. 3, 277–309.

[19] I. Cimrák, Convergence result for the constraint preserving mid-point scheme for micromagnetism, *J. Comput. Appl. Math.* **228** (2009), no. 1, 238–246.

[20] E. Davoli and G. Di Fratta, Homogenization of chiral magnetic materials: A mathematical evidence of Dzyaloshinskii's predictions on helical structures, *J. Nonlinear Sci.* **30** (2020), no. 3, 1229–1262.

[21] G. Di Fratta, C. B. Muratov, F. N. Rybakov and V. V. Slastikov, Variational principles of micromagnetics revisited, *SIAM J. Math. Anal.* **52** (2020), no. 4, 3580–3599.

[22] 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), no. 4, 2802–2838.

[23] I. Dzyaloshinskii, A thermodynamic theory of "weak" ferromagnetism of antiferromagnetics, *J. Phys. Chem. Solids* **4** (1958), no. 4, 241–255.

[24] M. Feischl and T. Tran, The Eddy current–LLG equations: FEM-BEM coupling and a priori error estimates, *SIAM J. Numer. Anal.* **55** (2017), no. 4, 1786–1819.

[25] A. Fert, V. Cros and J. Sampaio, Skyrmions on the track, *Nat. Nanotechnol.* **8** (2013), no. 3, 152–156.

[26] A. Fert, N. Reyren and V. Cros, Magnetic skyrmions: Advances in physics and potential applications, *Nat. Rev. Mater.* **2** (2017), Article ID 17031.

[27] C. J. García-Cervera, Numerical micromagnetics: A review, *Bol. Soc. Esp. Mat. Apl. SeMA* (2007), no. 39, 103–135.

[28] T. L. Gilbert, A Lagrangian formulation of the gyromagnetic equation of the magnetization fields, *Phys. Rev.* **100** (1955), Article ID 1243.

[29] V. Girault and P.-A. Raviart, *Finite Element Methods for Navier–Stokes Equations: Theory and Algorithms*, Springer Ser. Comput. Math. 5, Springer, Berlin, 1986.

[30] 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), no. 3, 1329–1368.

[31] A. Hubert and R. Schäfer, *Magnetic Domains: The Analysis of Magnetic Microstructures*, Springer, Berlin, 1998.

[32] E. Kim and J. Wilkening, Convergence of a mass-lumped finite element method for the Landau–Lifshitz equation, *Quart. Appl. Math.* **76** (2018), no. 2, 383–405.

[33] M. Kružík and A. Prohl, Recent developments in the modeling, analysis, and numerics of ferromagnetism, *SIAM Rev.* **48** (2006), no. 3, 439–483.

[34] L. Landau and E. Lifshitz, On the theory of the dispersion of magnetic permeability in ferromagnetic bodies, *Phys. Z. Sowjetunion* **8** (1935), 153–168.

[35] T. Moriya, Anisotropic superexchange interaction and weak ferromagnetism, *Phys. Rev.* **120** (1960), Article ID 91.

[36] C.-M. Pfeiler, Commics – A Python module for computational micromagnetics, https://gitlab.asc.tuwien.ac.at/cpfeiler/commics, 2022.

[37] C.-M. Pfeiler, M. Ruggeri, B. Stiftner, L. Exl, M. Hochsteger, G. Hrkac, J. Schöberl, N. J. Mauser and D. Praetorius, Computational micromagnetics with Commics, *Comput. Phys. Commun.* **248** (2020), Article ID 106965.

[38] D. Praetorius, Analysis of the operator $\Delta^{-1}\mathrm{div}$ arising in magnetic models, *Z. Anal. Anwend.* **23** (2004), no. 3, 589–605.

[39] D. Praetorius, M. Ruggeri and B. Stiftner, Convergence of an implicit-explicit midpoint scheme for computational micromagnetics, *Comput. Math. Appl.* **75** (2018), no. 5, 1719–1738.

[40] A. Prohl, *Computational Micromagnetism*, Adv. Numer. Math., B. G. Teubner, Stuttgart, 2001.

[41] A. Quarteroni, R. Sacco and F. Saleri, *Numerical Mathematics*, 2nd ed., Texts Appl. Math. 37, Springer, Berlin, 2007.

[42] J. Sampaio, V. Cros, S. Rohart, A. Thiaville and A. Fert, Nucleation, stability and current-induced motion of isolated magnetic skyrmions in nanostructures, *Nat. Nanotechnol.* **8** (2013), no. 11, 839–844.

[43] A. Thiaville, S. Rohart, V. Jué, É. Cros and A. Fert, Dynamics of Dzyaloshinskii domain walls in ultrathin magnetic films, *Europhys. Lett.* **100** (2012), no. 5, Article ID 57002.

[44] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, 2nd ed., Springer Ser. Comput. Math. 25, Springer, Berlin, 2006.

[45] Netgen/NGSolve Finite Element Library, https://ngsolve.org/, 2022.